ORIGINS OF ASYMMETRIC STRESS-STRAIN RESPONSE IN PHASE TRANSFORMATIONS

Huseyin Sehitoglu and Ken Gall

Department of Mechanical and Industrial Engineering
University of Illinois, Urbana, IL 61801, USA

ABSTRACT

It has been determined that the transformation stress-strain behavior of CuZnAl and NiTi shape memory alloys is dependent on the applied stress state. The uniaxial compressive stress necessary to macroscopically trigger the transformation is approximately 34% (CuZnAl) and 26% (NiTi) larger than the required uniaxial tensile stress. For three dimensional stress states, the response of either alloy system is dependent on the directions of the dominant principal stresses along with the hydrostatic stress component of the stress state. The stress state effects are dominated by the favored growth and nucleation of more martensite plates in tension versus compression. The effect of different hydrostatic pressure levels between stress states on martensite plates volume change is considered small.

INTRODUCTION

The purpose of this work is to determine the physical origins of the tension-compression asymmetry and the hydrostatic stress effect with novel experiments and measurements in two technologically important materials (CuZnAl and NiTi). Using unique equipment, considerable sensitivity to hydrostatic stress state has been obtained experimentally for the first time. It should be noted that there are currently no studies available in which both effective and hydrostatic stresses were systematically changed. Since shape memory alloys (SMA) can store large amounts of recoverable pseudo-elastic energy, they could be used in many applications where large strains are essential, but permanent deformation and energy loss due to plastic dissipation is undesirable. Additionally, SMA's have an advantage over traditional materials since the large pseudo-elastic mechanical strains can be triggered thermally, electrically, or mechanically.

SMA's owe their unique stress-strain behavior to a reversible thermoelastic martensitic transformation. It is widely accepted that the stress-induced martensitic transformation produces two unique macroscopic stress-strain responses, pseudoelasticity and the shape memory effect [1, 2] Analogous to stress-strain curves in the plastic regime, pseudoelastic and shape memory stress-strain curves demonstrate macroscopic yield points, hardening regions, and mechanical hysteresis upon unloading. The primary difference between the three curves concerns the mechanism of recoverable strain. Plastically deformed materials recover strains upon reverse loading, pseudoelastic materials recover strains immediately upon unloading, and materials exhibiting the shape memory effect recover strains after being subsequently heated. The existence of one phenomenon over another in any given alloy system is a function of test temperature, material composition, processing technique, and heat treatment.

Despite the wide ranging applicability of SMA's, there is a limited amount of experimental work on the response of SMA's to stress states other than tension [1-3]. This gap in research efforts is intriguing since the dependence of stress-induced martensitic transformations on the applied stress state was observed some 40 years ago in steels [4]. With this in mind, the purpose of the current study is to expose the issues related to the transformation behavior of CuZnAl and NiTi shape memory alloys under different stress states. More precisely, this work will focus on the dependence of the critical transformation stress level on the applied stress state.

EXPERIMENTAL TECHNIQUES

Polycrystalline $Cu_{59.1}Zn_{27.0}Al_{13.8}$ and $Ni_{50.0}Ti_{50.0}$ weight percent alloys were employed for the study. The normal to the habit plane and the twinning direction have the direction cosines (.199, .6804, .705) and (.1817,-.7457,.6411) respectively. CuZnAl demonstrates a small negative change (-0.3%) in volume upon transformation from the parent phase to the martensitic phase. The habit plane normal and transformation direction are given as (-.8889,.404,.215) and (.435,.7543,.4874) respectively for NiTi. This results in a small positive volume change (.19%). The heat treatment in both cases consisted of a solution heat treatment followed by an aging treatment. The treatment was performed to keep the martensite start temperature, M_s, at a reasonable level below room temperature. This assures that the transformation will be stress-induced. On average, the NiTi M_s was about -18 °C, while the CuZnAl M_s was about -10 °C. The tests in this study were conducted at room temperature where the sample is fully austenitic ($T > A_f$). Details of the unique experimental equipment used for the triaxial tests can be found in a recent publication [3] and are also summarized below.

In our work, a servohydraulic test machine fitted with a unique high pressure vessel is used for triaxial testing of NiTi and CuZnAl specimens. The schematic of the test system is provided in Figure 1. As Figure 1 indicates, axial stresses are applied to the specimen by the servohydraulic actuator of the MTS test machine; diametral stresses are applied to the specimen through the introduction of pressurized fluid into the pressure vessel. The axial stress is changed by applying force in the longitudinal direction, and circumferential and radial stresses are related to applied pressure ($=-p$). The ability of the present triaxial testing apparatus to simultaneously ramp the lateral and axial stresses on the specimen represents one of its main advantages over previous triaxial research efforts. In previous works, hydrostatic compression was typically applied first and the uniaxial stress was increased in a secondary operation. The present scheme circumvents any arguments regarding the role of initial hydrostatic compression on the material behavior. A personal computer was used for all test definition, command generation, and data acquisition tasks. More details of the pressure intensifier, load and strain measurements can be found in a recent publication [3].

EXPERIMENTAL RESULTS

The effective stress-strain curves for the CuZnAl and NiTi are shown in Figures 2 and 3 respectively. Stress states #1 and #3 are simple uniaxial tension and compression. Stress state

#2 has the following combination of principal stresses: $\sigma_1 = 2p$, $\sigma_2 = -p$, $\sigma_3 = -p$, while stress state #5 is governed by: $\sigma_1 = -2p$, $\sigma_2 = -p$, $\sigma_3 = -p$. Several other triaxial stress states were studied, and the results are discussed more thoroughly in two recent publications [13, 14]. The stress-strain curves are only shown up to 3% strain because strains much larger that this introduce considerable plastic deformation and non-recoverability [13]. In general, the NiTi has a much higher transformation yield point while the CuZnAl demonstrates a larger post-yield hardening modulus. Although the difference in hardening behavior is not as drastic as Figures 1 and 2 might indicate (Figure 2 has a scale twice as large as Figure 1) the difference is still notable. Both CuZnAl and NiTi show transformation yield points which are much higher in compression than in tension. In addition, the yield point of the zero hydrostatic case lies close to the yield point in pure tension for both materials. The yield point of CuZnAl under triaxial compression lies in between the tension and compression yield points. However, in NiTi, the yield point of the triaxial compression test lies considerably above both the tensile and compressive yield points.

DISCUSSION

For the most part, the stress state effects in CuZnAl and NiTi can be directly linked to micro-mechanical phenomenon. When a particular stress state is imposed on a SMA specimen, transformation strains are accumulated through the nucleation and subsequent growth of several preferred martensite plates (variants) [5]. Figures 4a and 4b show the typical arrangement and number of martensite plates caused by an applied stress in CuZnAl. The first image (a) is a magnified view of the plates in a single grain while the image (b) show the formation of different plates in several grains. The purpose of the two images is simply to demonstrate that two variants usually control the stress-induced transformation and that different grains favor the formation of selected variants. Unfortunately, it is not trivial to compare Figures 2-4 and completely understand the stress-state effects. To link the experimental behavior to the microscopic observations, a micro-mechanical model must be incorporated [1,6].

Although the model will not be extensively discussed here, the predictions of the model are a cornerstone in the understanding of stress-state effects in these alloys. As in a real material the model has the possibility of forming 24 martensite variants per grain. However, consistent with experimental observations, the model predicts that only 2 or 3 of these variants actually control the transformation under an applied stress state [1]. The advantage of the model is that it allows the "observation" of microscopic variables controlling the transformation which are not easily observed experimentally. One of the key predictions of the model is that more variants will activate under an applied tensile stress versus a compressive stress (Figures 5(a) and 5(b)). Clearly, the favored activation of martensite variants between stress states is one cause of stress state effects in these alloys. Simply stated, if a particular stress state has dominant principal stresses in tension, more variants will activate, the transformation will microscopically proceed quicker, and the macroscopic transformation yield stress will be lower.

Balancing the effect of the number of transforming variants is the relationship between the volume change during transformation and the hydrostatic component of the applied stress state. Thus, if the applied stress state has a negative hydrostatic component then the transformation will be triggered at a lower effective stress for CuZnAl. Macroscopically, the difference in the transformation yield point caused by differences in the hydrostatic stress component between stress states is not visible unless the hydrostatic stress difference is substantial. For example, the hydrostatic stress component due to pure compression is slightly more negative than the hydrostatic stress component due to pure tension. However, from Figures 2 and 3 it is clear that transformation in tension is favored over transformation in compression. One would have expected compression to have a lower yield point since its hydrostatic stress component is compressing in the direction which the transformation wants to proceed. Through Figure 5(a), the model demonstrates that the transformation is indeed "microscopically triggered" at a lower effective stress in compression in the CuZnAl case (note the very small offset in the number of transforming variants curve). However, this small offset is quickly overshadowed when more variants begin contributing to the transformation. In the case of NiTi the habit plane normal and the transformation direction lead to a positive volume change, consequently, the effect of volume change and transforming variants are additive leading to a higher sensitivity of the results to hydrostatic stress (Figure 5(b)).

Although the origin of stress state effects is clearly related to the microscopic aspects of the transformation, there still exists some experimental phenomenon that are not completely accounted for. The current theory of the authors is that texture is playing an intense role in the 3-D transformation behavior. At any rate, research is now in progress to experimentally view microscopic martensite growth *in situ* to better understand the dependence of martensite growth on the stress state.

CONCLUSIONS

(1) The uniaxial compressive stress necessary to macroscopically trigger the transformation is approximately 34% (CuZnAl) and 26% (NiTi) larger than the required uniaxial tensile stress. For three dimensional stress states, the response of either alloy system is dependent on the directions of the dominant principal stresses along with the hydrostatic stress component of the stress state.

(2) Stress state effects in CuZnAl and NiTi alloys are a balance between the number of transforming variants and the hydrostatic pressure (volume change) effect. The variant effects are more pronounced when two stress states have a small difference in hydrostatic stress components and the principal stresses are in different directions. The hydrostatic pressure effects become evident when there are extremely large differences in hydrostatic pressures between stress states.

ACKNOWLEDGMENTS

The research is supported by Department of Energy (DOE DEFG02-93ER14393), Basic Energy Sciences Division, Germantown, Maryland.

REFERENCES

1. K. Gall, H. Sehitoglu, H. J. Maier, and K. Jacobus, *Submitted to Acta. Met.*, (1997).
2. K. Jacobus, H. Sehitoglu and M. Balzer, *Met. Trans.*, **27A**, 3066, (1996).
3. M. Balzer and H. Sehitoglu, *Experimental Mechanics*, **37-1**, (1997).
4. S.A. Kulin, M. Cohen and B.L. Averbach, *J. Metals.* , **4,** 661, (1952).
5. T.A. Schroeder and C.M. Wayman, "Pseudoelastic Effects in Cu-Zn Single Crystals" *Acta Met.* **27**, 405, (1979).
6. E. Patoor, A. Eberhardt and M. Berveiller, Mechanics of Phase Transformations and Shape Memory Alloys, eds. L.C. Brinson and B. Moran, ASME, New York, NY, 23, (1994).

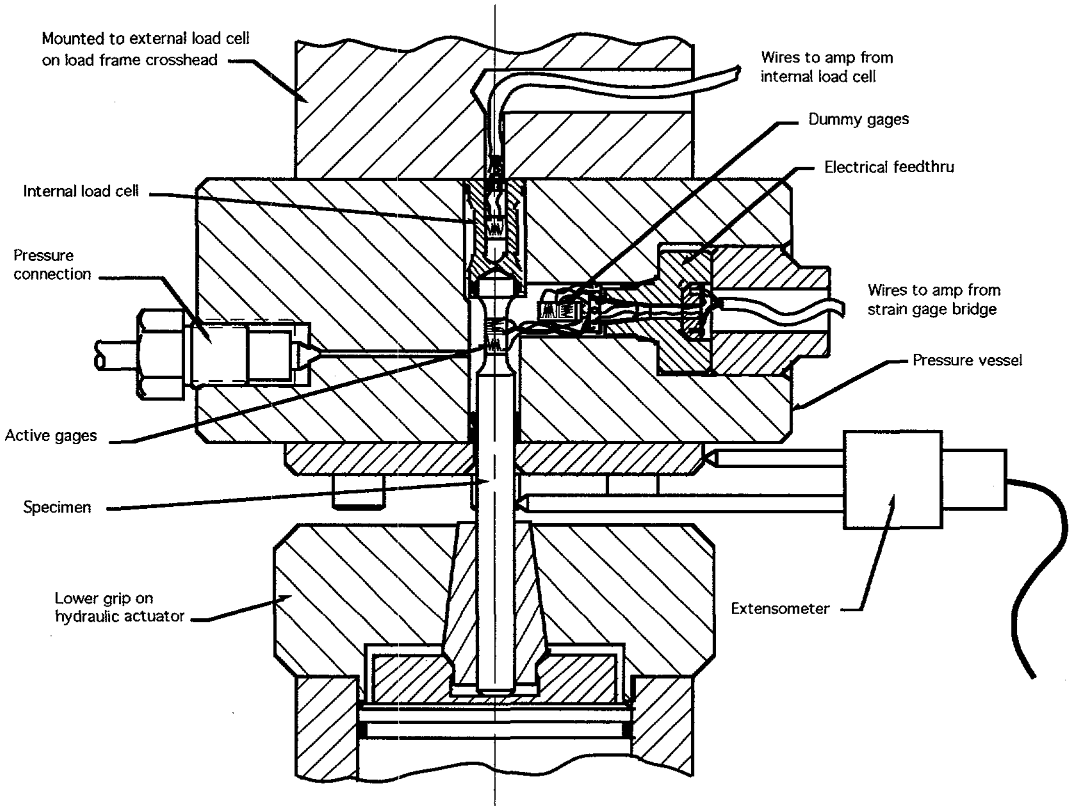

Mounted to external load cell on load frame crosshead
Wires to amp from internal load cell
Dummy gages
Electrical feedthru
Internal load cell
Pressure connection
Wires to amp from strain gage bridge
Pressure vessel
Active gages
Specimen
Lower grip on hydraulic actuator
Extensometer

Figure 1. Schematic of the pressure vessel for studying mechanical behavior at high pressures [3].

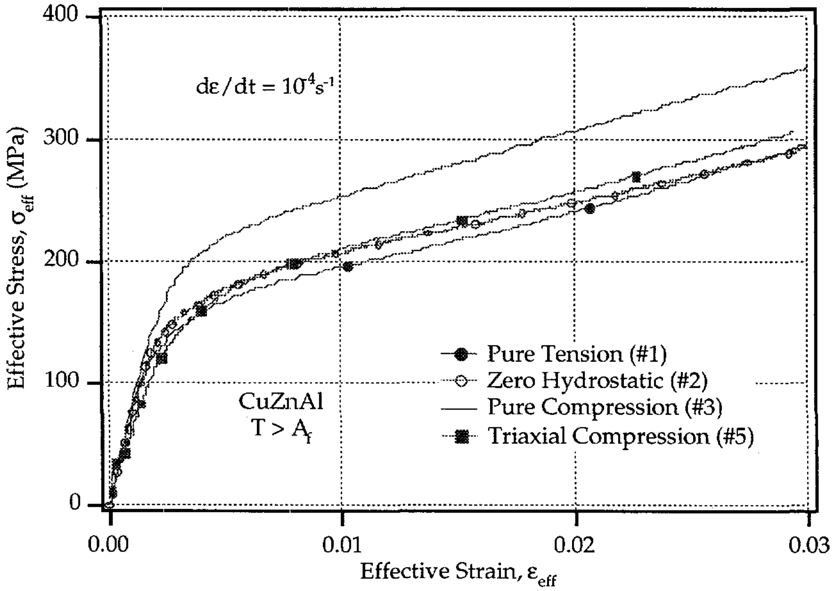


Figure 2. Effective stress-strain plots for a polycrystalline CuZnAl shape memory alloy specimen above the austenite finish temperature, A_f [1].

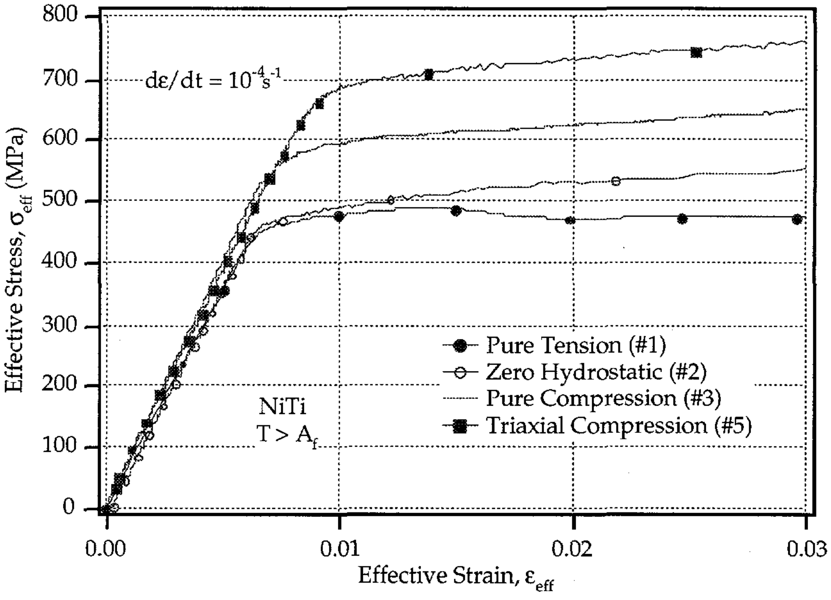


Figure 3: Effective stress-strain plots for a polycrystalline NiTi shape memory alloy specimen above the austenite finish temperature, A_f [2].

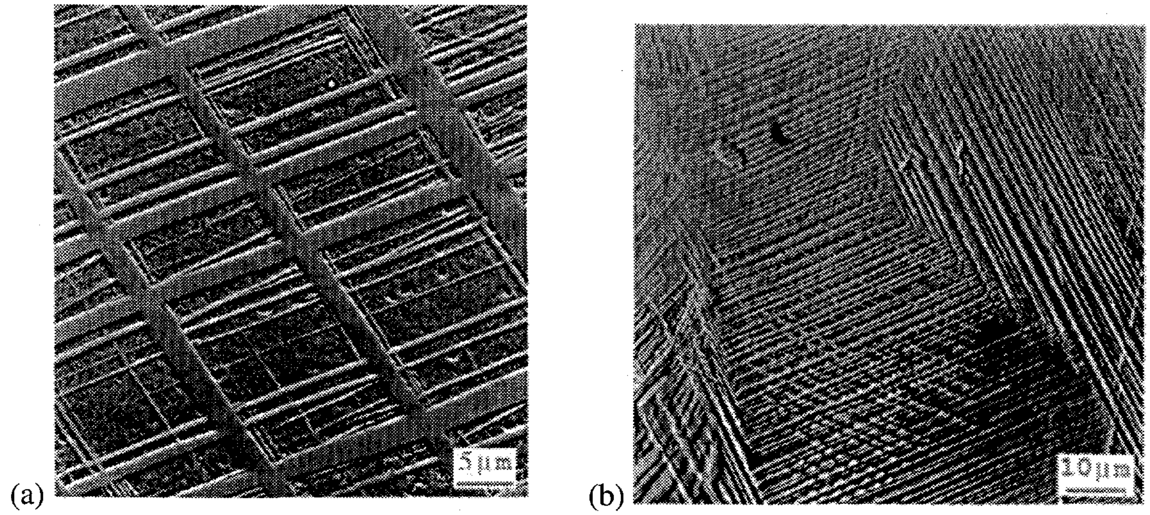
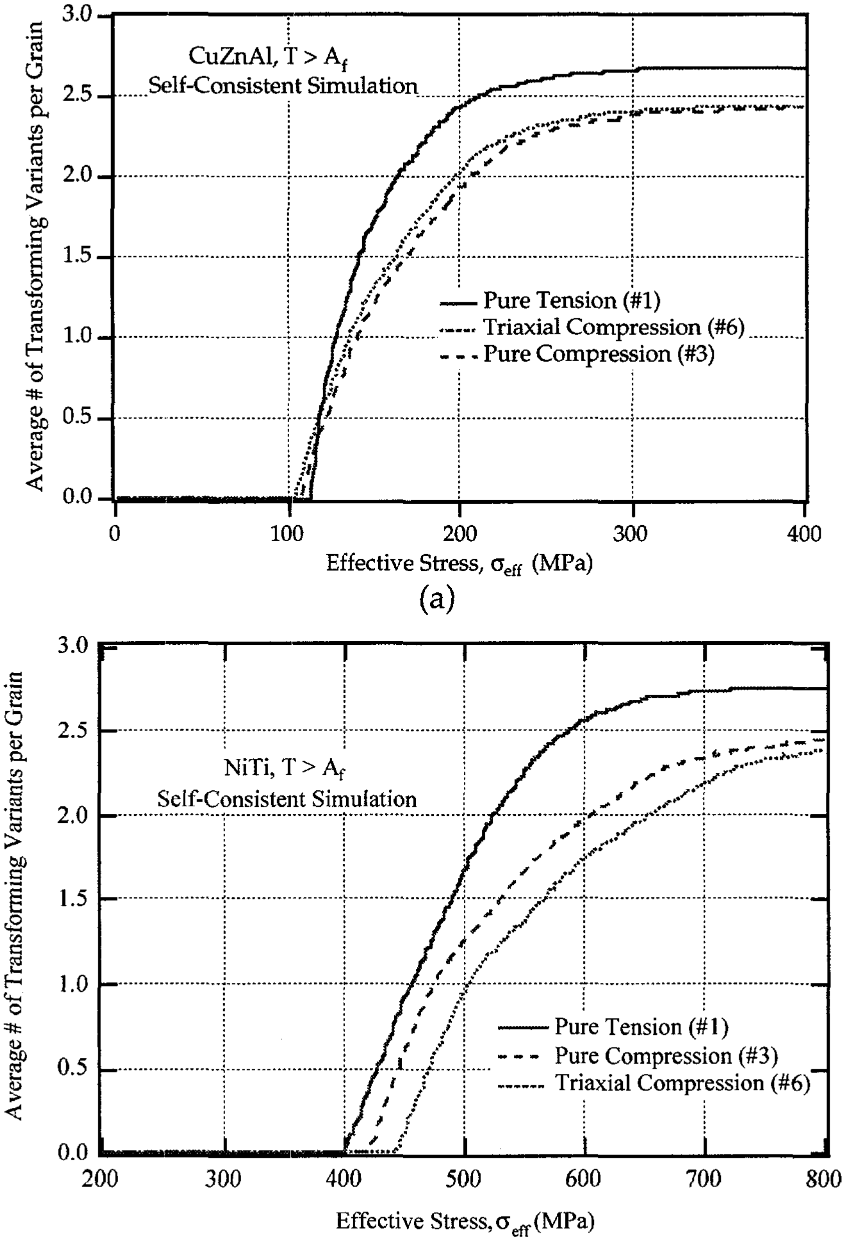


Figure 4. Scanning electron microscope image of martensite plates (a) in a single grain and (b) in several grains of CuZnAl [1].



(a)

(b)

Figure 5. Plot of the average number of active martensite variants versus the applied effective stress for a polycrystalline CuZnAl SMA [1], (b) results for NiTi .

CUTTING STATE IDENTIFICATION

B. S. Berger, I. Minis, M. Rokni, M. Papadopoulos, K. Deng, A. Chavalli

University of Maryland, Mechanical Engineering Department
College Park, MD 20742-3035 U.S.A.

Argonne National Laboratory
Argonne, Illinois 60439, U.S.A.

A Summary Reprinted from *Journal of Sound and Vibration* Vol. 200(1), B.S. Berger, et al., CUTTING STATE IDENTIFICATION, pp. 15-29, 1997, by permission of the publisher Academic Press Limited, London

ABSTRACT

Cutting states associated with the orthogonal cutting of stiff cylinders are identified through an analysis of the singular values of a Toeplitz matrix of third order cumulants of acceleration measurements. The ratio of the two pairs of largest singular values is shown to differentiate between light cutting, medium cutting, pre-chatter and chatter states. Sequences of cutting experiments were performed in which either depth of cut or turning frequency was varied. Two sequences of experiments with variable turning frequency and five with variable depth of cut, 42 cutting experiments in all, provided a database for the calculation of third order cumulants. Ratios of singular values of cumulant matrices find application in the analysis of control of orthogonal cutting

INTRODUCTION

The identification of cutting states, associated with the orthogonal cutting of stiff cylinders, is realized in the following through an analysis of the behavior of the singular values of a Toeplitz matrix of third order cumulants of acceleration measurements. A bispectral analysis of cutting tool acceleration measurements has shown, [3], that the cutting process is quadratically phase coupled. The determination of coefficients in an autoregressive approximation of the bispectrum, [20], involves the construction of an unsymmetric Toeplitz matrix, **R**, of third order cumulants. It is shown that the behavior of the dominant pairs of singular values of **R** provides a basis for the identification of cutting states. In particular, the ratio of the two pairs of largest singular values, the

R-ratio, is shown to differentiate between light cutting, medium cutting, pre-chatter and chatter states. Sequences of cutting experiments were performed in which either depth of cut or turning frequency was varied while all other cutting parameters were held constant. Two sequences of experiments with variable turning frequency and five with variable depth of cut, a total of forty-two cutting experiments, were studied. Results typical of the entire set are presented for a sequence of variable cutting depth and a sequence of variable turning frequency. The R-ratio evaluated at maxlag = 100, (4), is close to one for all cases of light cutting and two or greater for chatter. For intermediate states the ratio increases as the chatter state is approached.

EXPERIMENTAL APPARATUS

A schematic diagram of the experimental apparatus employed is shown in Figure 1 and consists of a Hardinge CNC lathe, a special force dynamometer (utilizing three Kistler 9068 force transducers) and its associated electronics, and a digital spectrum analyzer (Hewlett Packard 3566A) for data acquisition and real-time analysis.

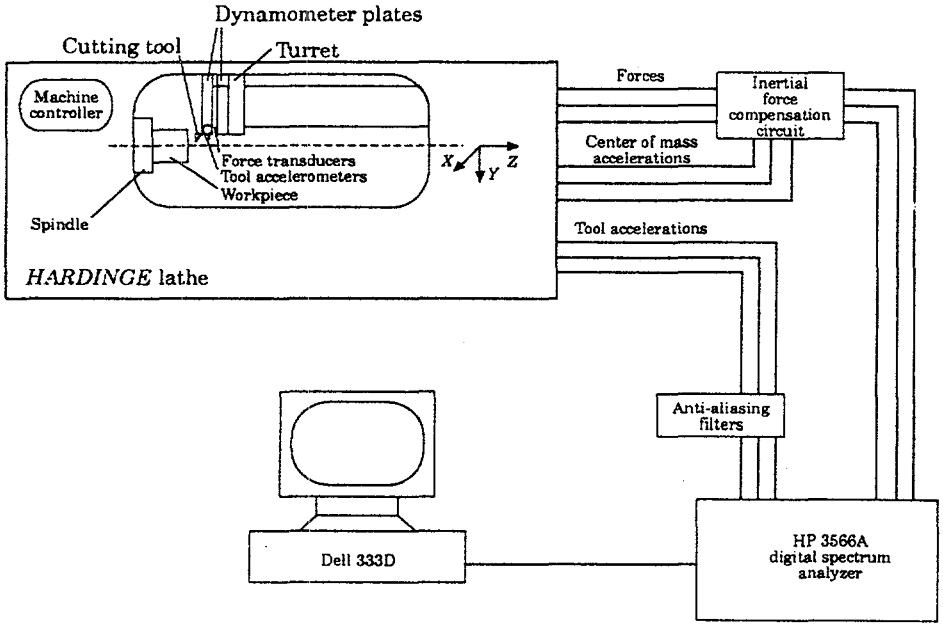


Figure 1. The experimental system.

THIRD ORDER RECURSION

Let $c_3(\tau_1, \tau_2) \equiv$ the third order cumulant of the real third order stationary random process X(k), k = 0, ±1, ±2. ... . If the mean of X(k) vanishes then $c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2)$ where $m_3(\tau_1, \tau_2)$

= E(X(k), X(k+τ_1) X(k+τ_2)), E is the expected value, which may be estimated by

$$m_3(\tau_1,\tau_2) \ = \ (1/2n) \sum_{k=-n}^{+n} \ X(k) \ X(k+\tau_1) \ X(k+\tau_2) \tag{1}$$

where n→ +∞. The bispectrum of X(k), $C_3(\omega_1, \omega_2)$ is defined by

$$C_3(\omega_1,\omega_2) \ = \ \sum_{\tau_1=-\infty}^{+\infty} \ \sum_{\tau_2=-\infty}^{+\infty} \ c_3(\tau_1,\tau_2) \ \exp \ [-j(\omega_1\tau_1+\omega_1\tau_2)] \tag{2}$$

$|C_3(\omega_1, \omega_2)|$ ≡ the bispectral index.

Consider an autoregressive, AR, estimation of the bispectrum, $C_3(\omega_1, \omega_2)$, (2) [16,17]. A p - th order AR process is described by

$$X(k) \ + \sum_{i=1}^{p} \ a(i) \ X(k-i) \ = \ W(k) \tag{3}$$

where it is assumed that W(k) is non-Gaussian, E(W(k)) = 0, E(W^3(k)) = β. Multiplying through (3), summing and noting (1) gives

$$c_3^x(-k,-l) \ + \sum_{i=1}^{p} \ a(i) \ c_3^x(i-k, \ i-l) \ = \ \beta \ \delta(k,l) \tag{4}$$

where k, l > 0. Letting k=l in the third order recursion equation, (4), with k = 0, ..., p yields p+1 equations for the p+1 unknowns a(i) and β; p+1 ≡ maxlag. In matrix notation

$$\boldsymbol{R} \ \boldsymbol{a} \ = \ \boldsymbol{b} \tag{5}$$

where

$$\boldsymbol{R} \ = \begin{vmatrix} g(o,o) & g(1,1) & ...g(p,p) \\ g(-1,-1) & g(o,o) & ...g(p-1,p-1) \\ \vdots & & \vdots \\ g(-p,-p) & g(-p+1,-p+1) & ...g(o,o) \end{vmatrix} \tag{6}$$

g(i,j) ≡ c_3^x(i,j), **a** ≡ [1, a(1), ..., a(p)]^T and **b** ≡ [β, 0, ..., 0]^T. **R** is in general a nonsymmetric Toeplitz matrix. A sufficient but not necessary condition for the representation in (5) to exist is the symmetry and positive definiteness of **R**. A discussion of this and related conditions is given in [17]. The bispectrum corresponding to (3) is given by, [4],

$$C_3^x(\omega_1,\omega_2) = \beta\ H(\omega_1)\ H(\omega_2)\ H^*(\omega_1+\omega_2) \tag{7}$$

where

$$H(\omega) = 1/(\ 1 + \sum_{n=1}^{p} a(i)\ \exp\ (-j\ \omega\ n)) \tag{8}$$

and $H^*(\omega) \equiv$ complex conjugate of $H(\omega)$.

An estimate of the **R** matrix, (6), and bispectrum, (7), for a data set X(I), I=1,...,N may be formed [16,17], as follows:

1. Segment the data set into K records of M samples each. $X^i(k)$, k=1,2,...,M are data points associated with the i-th record.
2. Compute $c_3^x{}_{,i}$ (m,n) for the i-th record as

$$c_{3,i}^{x} = (1/M) \sum_{l=a}^{b} X^{(i)}(l)\ X^{(i)}(l+m)\ X^{(i)}(l+n) \tag{9}$$

where i = 1,2,...,K, a ≡ max (1,1-m,1-n) and b≡ min(M, M-m, M-n).

3. Average $c_3^x{}_{,i}$ (m,n) over all K records,

$$\hat{c}_3(m,n) = (1/K) \sum_{i=1}^{K} c_{3,i}^{x}(m,n) \tag{10}$$

to yield the estimate $\hat{c}_3$(m,n) of c_3(m,n). Form an estimated **R** matrix by replacing c_3(m,n) by $\hat{c}_3$(m,n) in (6). Estimated values of **a** follow from (5). These results implemented in [22] are subsequently applied to orthogonal cutting data.

SINGULAR VALUE DECOMPOSITION

If **A** is a real mxn matrix then there exist orthogonal matrices **U** ∊ R^{mxm} and **V** ∊ $\mathbf{R}^{nxn}$ such that

$$U^T A\ V = diag.(\sigma_1,\cdots,\sigma_q) \in R^{mxn} \tag{11}$$

where q = min(m,n), $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_q \geq 0$ are the singular values and R^{mxn} denotes a real mxn matrix. A criterion for selecting the autoregressive order, p, in (3) is given in [17,22]. p is chosen to equal the number of singular values of the R matrix which are above the noise floor. Note that if $\sigma_1 \geq ... > \sigma_r > \sigma_{r+1} = ... = \sigma_q = 0$ then rank (**A**) = r, [5,7].

Relationships between phase coupled trigonometric functions and the singular values of the corresponding **R** matrix were established through a study of three functions $f_i(t)$ where

$$f_1(t) = \cos(2\pi \cdot 100t+\phi_1) + \cos(2\pi \cdot 100t+\phi_2) + 0.2\cos(2\pi \cdot 200t+\phi_1+\phi_2) \quad \textbf{(12)}$$

$$f_2(t) = 0.9\cos(2\pi \cdot 90t+\phi_1) + 1.0\cos(2\pi \cdot 100t+\phi_2) + 0.2\cos(2\pi \cdot 190t+\phi_1+\phi_2) \quad \textbf{(13)}$$

$$\begin{aligned} f_3(t) &= 1.0\cos(2\pi\cdot 90t+\phi_1) + 1.0\cos(2\pi\cdot 100t+\phi_2) \\ &+ 1.0\cos(2\pi\cdot 190t+\phi_1+\phi_2) + 1.0\cos(2\pi\cdot 100t+\phi_2) \\ &+ 1.0\cos(2\pi\cdot 110t+\phi_3) + 0.5\cos(2\pi\cdot 210t+\phi_2+\phi_3) \end{aligned} \quad \textbf{(14)}$$

and ϕ_i are mutually independent and uniformly distributed over $[0, 2\pi]$. The $f_i(t)$ functions were sampled at 1024 Hz over an interval of 10 sec. **R** matrices were evaluated for each $f_i(t)$ by averaging over 10 one sec. intervals, (6), (9).

$f_1(t)$, (12) is an example of the self phase coupling of a 100 Hz frequency component. In the experimental data studied frequency components in the neighborhood of 100 and 200 Hz were always observed in the power spectra of cutting states close to chatter. A peak with frequency coordinates in the neighborhood of (100 Hz, 100 Hz), appeared in the bispectrum of cutting states in the neighborhood of chatter. The ratio of the mean of the largest pair of singular values to the mean of the second largest pair defines a non-dimensional ratio of invariants of **R**, the R-ratio. This ratio is shown as a function of maxlag for $f_1(t)$ in Figure 2(c). R-ratio $\approx$ 2.0 for maxlag > 30.

$f_2(t)$, (13), involves the phase coupling of 90 and 100 Hz components. A bispectral peak at (100, 90) indicates phase coupling between the 90 and 100 Hz components. The mean of the first pair of singular values is nearly equal to the mean of the second pair of singular values for maxlag > 80. Note Figure 3(b) for maxlag > 90 for which $1.0 < \text{R-ratio} \leq 1.2$.

$f_3(t)$, (14). is the sum of a phase coupled component at 100 Hz and 110 Hz and a phase coupling of 90 and 100 Hz components. The bispectrum of $f_3(t)$ has peaks at (100, 110), (100, 90) and (110, 100), (90, 100) because of symmetry. In Figure 4(b) the R-ratio is plotted as a function of maxlag from which it is seen that R-ratio $\approx$ 1.5 for maxlag > 80.

CUTTING STATE CHARACTERIZATION

Sequences of cutting experiments were performed in which either depth of cut or the turning frequency was varied with all other cutting parameters held constant. Singular values of **R**, (6), were computed for two sequences with variable turning frequency and five sequences with variable depth of cut over a turning frequency range of 290-852 rpm. Each variable cutting depth sequence ended in chatter while each variable turning frequency sequence contained at least one chatter state. A total of 42 cutting experiments were performed. Typical sequences have been selected from the set.

For sequence 1, s-1, the turning frequency = 460 rpm, rake angle = 5°, surface speed = 90 m/min, feed rate = .007 in/rev, resampling rate = 1024 Hz, frequency cut-off = 1100 Hz and the depth of cut = 2.5, 2.6, 2.7 and 2.8 mm at which depth chatter was observed.

The R-ratio vs. maxlag is shown in Figure 5(b) for a depth of cut of 2.5 mm which corresponds to light cutting. The R-ratio, Figure 5(b), is close to 1.0. For $70 \leq$ maxlag ≤ 100, $1.12 \geq R \geq 1.08$. The behavior of the R-ratio as a function of maxlag has similarities with that of $f_2(t)$, (13), Figure 3(b). $f_2(t)$ contains two phase coupled trigonometric functions of 90 and 100 Hz which approximates phase coupling between the first natural frequency of the system at 98 Hz and a lower frequency component of the sideband structure.

Chatter was observed for a depth of cut of 2.8 mm. The R-ratio vs. maxlag is shown in Figure 6(b). One pair of singular values is dominant. For $20 \leq$ maxlag ≤ 100, $2.0 \leq$ R-ratio ≤ 2.4. The R-ratio as a function of maxlag is similar to that of $f_1(t)$, (12), Figures 2(b) and 6(b), which represents self phase coupling of the 100 Hz component. The presence of self-phase coupling in the time series is confirmed by peaks in the power spectrum at 100 and 200 Hz and a peak in the bicoherence index at (100, 100). These results are consistent across all sequences of experiments in which the depth of cut varies. For increasing depth of cut the R-ratio clearly differentiates between light cutting and chatter.

Two intermediate states with depths of cut of 2.6 and 2.7 mm complete the sequence s-1. The R-ratio vs. maxlag is shown in Figure 7(b), for the 2.6 mm case. For $50 \leq$ maxlag ≤ 110 the R-ratio $\simeq 1.6$, Figure 7(b). There is a similarity between Figure 4(b), R-ratio for $f_3(t)$ and Figure 7(b). The R-ratios are close to one another for $50 \leq$ maxlag ≤ 110.

CONCLUSIONS

The ratio of the mean of the two dominant pairs of singular values, R-ratio, evaluated for maxlag = 100, approximates one for light cutting, two or more for chatter and near chatter states and takes intermediate values for intermediate states, increasing from one to two as chatter is approached. This behavior was observed in an analysis of tool acceleration time series for five sequences of cutting experiments with increasing depth of cut and two sequences with variable turning frequency. For chatter and light and intermediate cutting the R-ratio is seen to be a constant or slowly changing function of maxlag for maxlag > 40 (4), Figures 5(b), 6(b), and 7(b).

ACKNOWLEDGEMENTS

The authors acknowledge the support of D.O.E. through DE-FG02-93ER14335 and the N.S.F. through GER-9354956. The encouragement of O. Manley, D.O.E. and P. Grootenhuis and D.J. Ewins, Imperial College of Science and Technology, London, is appreciated.

REFERENCES

3. B. S. Berger, I. Minis, K. Deng, Y. S. Chen, A. Chavali and M. Rokni 1996 *Journal of Sound and Vibration* 191(5), 986-992. Phase Coupling in orthogonal cutting.

4. D. R. Brillinger and M. Rosenblatt 1967a in Spectral Analysis of Time Series (B. Harris, editor). New York : John Wiley. Asymptotic theory of k-th order spectra. 1967b in Spectral Analysis of Time Series (B. Harris, editor). New York : John Wiley. Computation and interpretation of k-th order spectra.

5. G. H. Golub and C. F. Van Loan 1993 Matrix Computations. Baltimore : The Johns Hopkins University Press.

7. R. A. Horn and C. R. Johnson 1991 Topics in Matrix Analysis. Cambridge : Cambridge University Press.

16. C. L. Nikias and J. M. Mendel 1993 *IEEE Signal Processing Magazine* July, 10-37. Signal processing with higher-order spectra.

17. C. L. Nikias and A. P. Petropulu 1993 Higher-Order Spectra Analysis. Englewood Cliffs, New Jersey : Prentice Hall.

20. M. R. Raghuveer and C. L. Nikias 1985 *IEEE Transactions on Acoustics, Speech and Signal Processing* ASSP-33(4), 1213-1230. Bispectrum estimation : A parametric approach.

22. A. Swami, J. M. Mendel and C. L. Nikias 1993 Hi Spec Toolbox. Matick Massachusetts : Math Works, Inc.

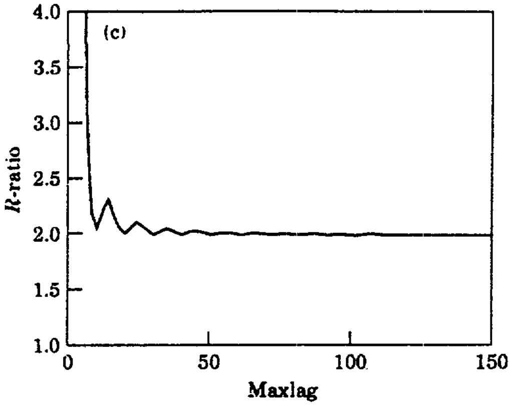


Figure 2. Test Function $f_1(t)$: (c) R-ratio vs. maxlag for $f_1(t)$.

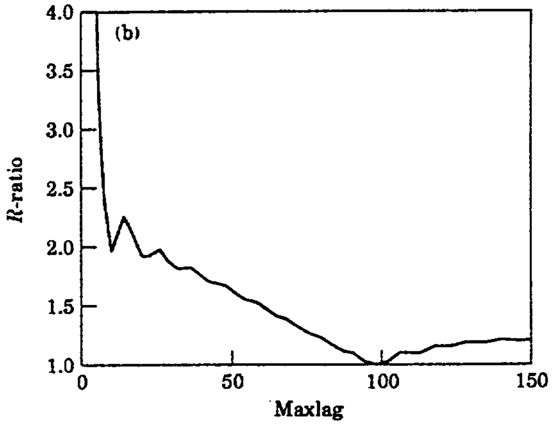


Figure 3. Test Function $f_2(t)$. (b) R-ratio vs. maxlag for $f_2(t)$.

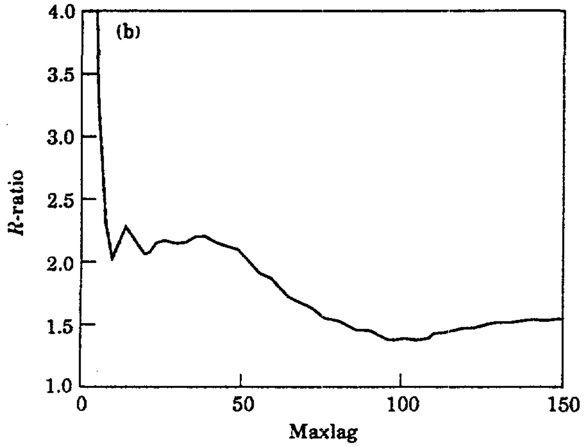


Figure 4. Test Function $f_3(t)$. (b) R-ratio vs. maxlag for $f_3(t)$.

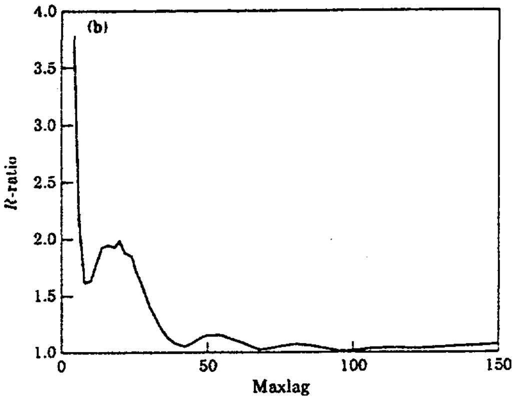


Figure 5. Data Set s-1. 2.5 mm. (b) R-ratio vs. maxlag.

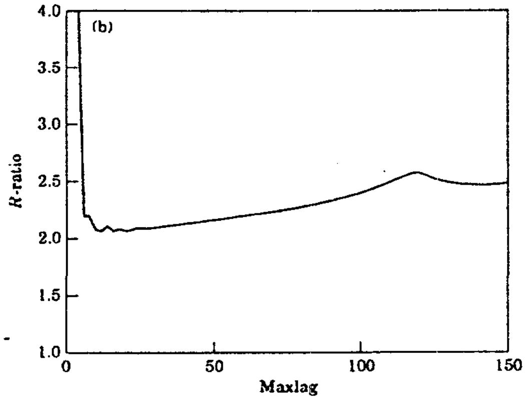


Figure 6. Data Set s-1, 2.8 mm. (b) R-ratio vs. maxlag.

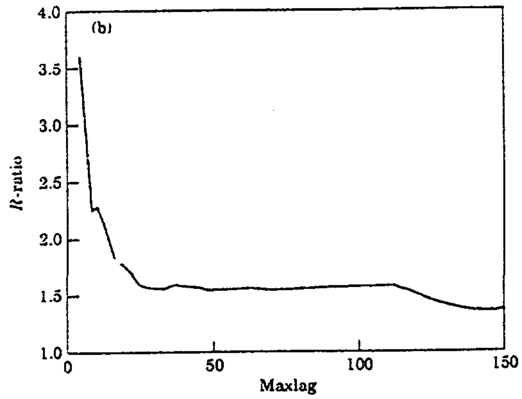


Figure 7. Data Set s-1, 2.6 mm. (b) R-ratio vs. maxlag.

MULTI-ROBOT MOTION CONTROL FOR COOPERATIVE OBSERVATION

Lynne E. Parker

Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

ABSTRACT

An important issue that arises in the automation of many security, surveillance, and reconnaissance tasks is that of monitoring (or observing) the movements of targets navigating in a bounded area of interest. A key research issue in these problems is that of sensor placement — determining where sensors should be located to maintain the targets in view. In complex applications involving limited-range sensors, the use of multiple sensors dynamically moving over time is required. In this paper, we investigate the use of a cooperative team of autonomous sensor-based robots for the observation of multiple moving targets. We focus primarily on developing the distributed control strategies that allow the robot team to attempt to minimize the total time in which targets escape observation by some robot team member in the area of interest. This paper first formalizes the problem and discusses related work. We then present a distributed approximate approach to solving this problem that combines low-level multi-robot control with higher-level reasoning control based on the ALLIANCE formalism. We analyze the effectiveness of our approach by comparing it to three other feasible algorithms for cooperative control, showing the superiority of our approach for a large class of problems.

INTRODUCTION

An important issue that arises in the automation of many security, surveillance, and reconnaissance tasks is that of monitoring (or observing) the movements of targets navigating in a bounded area of interest. A key research issue in these problems is that of sensor placement — determining where sensors should be located to maintain the targets in view. In the simplest version of this problem, the number of sensors and sensor placement can be fixed in advance to ensure adequate sensory coverage of the area of interest. However, in more complex applications, a number of factors may prevent fixed sensory placement in advance. For example, there may be little prior information on the location of the area to be monitored, the area may be sufficiently large that economics prohibit the placement of a large number of sensors, the available sensor range may be limited, or the area may not be physically accessible in advance of the mission. In the general case, the combined coverage capabilities of the available robot sensors will be insufficient to cover the entire terrain of interest. Thus, the above constraints force the use of multiple sensors dynamically moving over time.

In this paper, we investigate the use of a cooperative team of autonomous sensor-based robots for applications in this domain. We focus primarily on developing the distributed control strategies that allow the team to attempt to minimize the total time in which targets escape observation by some robot team member in the area of interest. Of course, many variations of this dynamic, distributed sensory coverage problem are possible. For example, the relative numbers and speeds of the robots and the targets to be tracked can vary, the availability of inter-robot communication can vary, the robots can differ in their sensing and movement capabilities, the terrain may be either enclosed or have entrances that allow targets to enter and exit the area of interest, the terrain may be either indoor (and thus largely planar) or outdoor (and thus 3D), and so forth. Many other subproblems must also be addressed, including the physical tracking of targets (e.g. using vision, sonar, IR, or laser range), prediction of target movements, multi-sensor fusion, and so forth. Thus, while our ultimate goal is to develop distributed algorithms that address all of these problem variations, we first focus on the aspects of distributed control in homogeneous robot teams with equivalent sensing and movement capabilities working in an uncluttered, bounded area.

The following section defines the multitarget observation problem of interest in this paper, and is followed by a discussion of related work. We then describe our approach, discussing each of the subcomponents of the system. Next, we describe and analyze the results of our approach, compared to three other feasible algorithms for cooperative motion control. Finally, we offer concluding remarks.

PROBLEM DESCRIPTION

The problem of interest in this paper — the cooperative multi-robot observation of multiple moving targets (or *CMOMMT* for short) — is defined as follows. Given:

$\mathcal{S}$: a two-dimensional, bounded, enclosed spatial region, with entrances/exits
$\mathcal{R}$: a team of m robots with 360^0 field of view observation sensors
that are noisy and of limited range
$\mathcal{O}(t)$: a set of n targets $o_j(t)$, such that $In(o_j(t), \mathcal{S})$ is true (where $In(o_j(t), \mathcal{S})$
means that target $o_j(t)$ is located within region $\mathcal{S}$ at time t)

Define an $m \times n$ matrix $A(t)$, where

$$a_{ij}(t) = \begin{cases} 1 & \text{if robot } r_i \text{ is monitoring target } o_j(t) \text{ in } \mathcal{S} \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

We further define the *logical OR* operator over a vector H as:

$$\bigvee_{i=1}^{k} h_i = \begin{cases} 1 & \text{if there exists an } i \text{ such that } h_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

We say that a robot is *monitoring* a target when the target is within that robot's observation sensory field of view. Then, the goal is to maximize:

$$\sum_{t=0}^{T} \sum_{j=1}^{n} \bigvee_{i=1}^{m} a_{ij}(t)$$

over time steps Δt under the assumptions listed below. In other words, the goal of the robots is to maximize the collective time during which targets in $\mathcal{S}$ are being monitored by at least one robot during the mission from $t = 0$ to $t = T$. Note that we do not assume that the membership of $\mathcal{O}(t)$ is known in advance.

In addressing this problem, we assume the following: Define *sensor_coverage*(r_i) as the area visible to robot r_i's observation sensors, for $r_i \in \mathcal{R}$. Then we assume that, in general,

$$\bigcup_{r_i \in \mathcal{R}} \textit{sensor\_coverage}(r_i) \ll \mathcal{S}.$$

That is, the maximum area covered by the observation sensors of the robot team is much less than the total area to be monitored. This implies that fixed robot sensing locations or sensing paths will not be adequate in general, and that, instead, the robots must move dynamically as targets appear in order to maintain observational contact with them and to maximize the coverage of the area $\mathcal{S}$.

We further assume the following:

- The robots have a broadcast communication mechanism that allows them to send (receive) messages to (from) each other within the area $\mathcal{S}$.

- For all $r_i \in \mathcal{R}$ and for all $o_j(t) \in \mathcal{O}(t)$, $max\_v(r_i) > max\_v(o_j(t))$, where $max\_v(a)$ gives the maximum possible velocity of entity a, for $a \in \mathcal{R} \cup \mathcal{O}(t)$.

- Targets in $\mathcal{O}$ can enter and exit region $\mathcal{S}$ through distinct entrances/exits.

- The robot team members share a known global coordinate system.

To somewhat simplify the problem initially, we report here the results of the case of an omnidirectional 2D sensory system (such as a ring of cameras or sonars), in which the robot sensory system is of limited range, but is available for the entire 360^o around the robot.

RELATED WORK

Research related to the multiple target observation problem can be found in a number of domains, including art gallery and related problems, multitarget tracking, and multi-robot surveillance tasks. While a complete review of these fields is not possible in a short paper, we will briefly outline the previous work that is most closely related to the topic of this paper.

The work most closely related to the *CMOMMT* problem falls into the category of the *art gallery* and related problems [1], which deal with issues related to polygon visibility. The basic art gallery problem is to determine the minimum number of guards required to ensure the visibility of an interior polygonal area. Variations on the problem include fixed point guards or mobile guards that can patrol a line segment within the polygon. Most research in this area typically utilizes centralized approaches to the placement of sensors, uses ideal sensors (noise-free and infinite range), and assumes the availability of sufficient numbers of sensors to cover the entire area of interest. Several authors have looked at the static placement of sensors for target tracking in known polygonal environments (e.g. [2]). These works differ from the *CMOMMT* problem, in that our robots must dynamically shift their positions over time to ensure that as many targets as possible remain under surveillance, and their sensors are noisy and of limited range.

Sugihara *et al.* [3] address the *searchlight scheduling problem*, which involves searching for a mobile "robber" (which we call *target*) in a simple polygon by a number of fixed searchlights, regardless of the movement of the target. They develop certain necessary and sufficient conditions for the existence of a search schedule in certain situations, under the assumption of a single target, no entrances/exits to the polygon, and fixed searcher positions

Suzuki and Yamashita [4] address the *polygon search* problem, which deals with searching for a mobile target in a simple polygon by a single mobile searcher. They examine two cases: one in which the searcher's visibility is restricted to k rays emanating from its position, and one in which the searcher can see in all directions simultaneously. Their work assumes no entrances/exits to the polygon and a single searcher.

LaValle *et al.* [5] introduces the visibility-based motion planning problem of locating an unpredictable target in a workspace with one or more robots, regardless of the movements of the target. They define a visibility region for each robot, with the goal of guaranteeing that the target will eventually lie in at least one visibility region. In LaValle *et al.* [6], they address the related question of maintaining the visibility of a moving target in a cluttered workspace by a single robot. They are also able to optimize the path along additional criteria, such as the total distance traveled. The problems they address in these papers are closely related to the problem of interest here. The primary difference is that their work does not deal with multiple robots maintaining visibility of multiple targets, nor a domain in which targets may enter and exit the area of interest.

Another large area of related research has addressed the problem of multitarget tracking (e.g. Bar-Shalom [7, 8], Blackman [9], Fox *et al.* [10]). This problem is concerned with computing the trajectories of multiple targets by associating observations of current target locations with previously detected target locations. In the general case, the sensory input can come from multiple sensory platforms. Our task in this paper differs from this work in that our goal is not to calculate the trajectories of the targets, but rather to find dynamic sensor placements that minimize the collective time that any target is not being monitored (or observed) by at least one of the mobile sensors.

APPROACH

Overview

Since the *CMOMMT* problem can be shown to be NP-complete, and thus intractable for computing optimal solutions, we propose an approximate control mechanism that is shown to work well in practice. This approximate control mechanism is based upon our previous work, described in [11, 12], which defines a fully distributed, behavior-based software architecture called ALLIANCE that enables fault tolerant, adaptive multi-robot action selection. This architecture is a hybrid approach to robotic control that incorporates a distributed, real-time reasoning system utilizing behavioral motivations above a layer of low-level, behavior-based control mechanisms. This architecture for cooperative control utilizes no centralized control; instead, it enables each individual robot to select its current actions based upon its own capabilities, the capabilities of its teammates, a previous history of interaction with particular team members, the current state of the environment, and the robot's current sensory readings. ALLIANCE does not require any use of negotiation among robots, but rather relies upon broadcast messages from robots to announce their current activities. The ALLIANCE approach to communication and action selection results in multi-robot cooperation that gracefully degrades and/or adapts to real-world problems, such as robot failures, changes in the team mission, changes in the robot team, or failures or noise in the communication system. This approach has been successfully applied to a variety of cooperative robot problems, including mock hazardous waste cleanup, bounding overwatch, janitorial service, box pushing, and cooperative manipulation, implemented on both physical and simulated robot teams.

Our proposed approach to the *CMOMMT* problem is based upon the same philosophy of control that was utilized in ALLIANCE. In this approach, we enable each robot team member to make its own action selections, without the need for any centralized control or negotiation. The low-level, behavior based control of each robot calculates local force vectors that attract the robot to nearby targets and repel the robot from nearby teammates. Added above the low-level control is a higher-level reasoning system that generates weights to be applied to the force vectors. These weights are based upon previous experiences of the robot, and can be in the form of motivations of behavior or rule-based heuristics. The high-level reasoning system of an individual robot is thus able to influence the local, low-level control of that robot, with the aim of generating an improved collective behavior across robots when utilized by all robot team members.

Target and robot detection

Ideally, robot team members would be able to passively observe nearby robots and targets to ascertain their current positions and velocities. Research fields such as machine vision have dealt extensively with this topic, and have developed algorithms for this type of passive position calculation. However, since the physical tracking and 2D positioning of visual targets is not the focus of this research, we instead assume that robots use a global positioning system (such as GPS for outdoors, or the laser-based MTI indoor positioning system [13] that is in use at our CESAR laboratory) to determine their own position and the position of targets within their sensing range, and communicate this information to the robot team members within their communication range[1].

For each robot r_i, we define the *predictive tracking range* as the range in which targets localized by other robots $r_k \neq r_i$ can affect r_i's movements. Thus, a robot can know about two types of targets: those that are directly sensed or those that are "virtually" sensed through predictive tracking. When a robot receives a communicated message regarding the location and velocity of a sighted target that is within its predictive tracking range, it begins a predictive tracking of that target's location, assuming that the target will continue linearly from its current state. We

[1] This approach to communication places an upper limit on the total allowable number of robots and targets at about 400. Since the communication is $O(nm)$, we compute this upper limit by assuming a 1.6 Mbps Proxim radio ethernet system (such as the one in our laboratory) and assuming that messages of length 10 bytes per robot per target are transmitted every 2 seconds. With these numbers, we find that nm must be less than 4×10^4 bps to avoid saturation of the communication bandwidth.

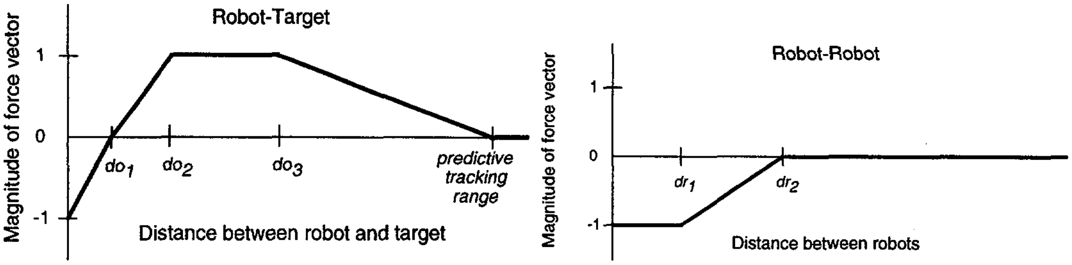


Figure 1: Functions defining the magnitude of the force vectors to nearby targets and robots.

assume that if the targets are dense enough that their position estimations do not supply enough information to disambiguate distinct targets, then existing tracking approaches (e.g. Bar-Shalom [8]) should be used to uniquely identify each target based upon likely trajectories.

Local force vector calculation

The local control of a robot team member is based upon a summation of force vectors which are attractive for nearby targets and repulsive for nearby robots. The first function in figure 1 defines the relative magnitude of the attractive forces of a target within the predictive tracking range of a given robot. Note that to minimize the likelihood of collisions, the robot is repelled from a target if it is too close to that target ($distance < do_1$). The range between do_2 and do_3 defines the preferred tracking range of a robot from an object. In practice, this range will be set according to the type of tracking sensor used and its range for optimal tracking. The attraction to the object falls off linearly as the distance to the object varies from do_2. The attraction goes to 0 beyond the predicted tracking range, indicating that this object is too far to have an effect on the robot's movements.

The second function of figure 1 defines the magnitude of the repulsive forces between robots. If the robots are too close together ($distance < dr_1$), they repel strongly. If the robots are far enough apart ($distance > dr_2$), they have no effect upon each other in terms of the force vector calculations. The magnitude scales linearly between these values.

One problem with using only force vectors, however, is that of local minima. As defined so far, the force vector computation is equivalent for all targets, and for all robots. Thus, we need to inject additional high-level reasoning control into the system to take into account more global information. This reasoning is modeled as predictive weights that are factored into the force vector calculation, and are described in the next subsection.

High-level reasoning control

To help resolve the problems of local minima, the higher-level reasoning control differentially weights the contributions of each target's force field on the total computed field. This higher-level knowledge can express any information or heuristics that are known to result in more effective global control when used by each robot team member locally. Our present approach expresses this high-level knowledge in the form of two types of probabilities: the probability that a given target actually exists, and the probability that no other robot is already monitoring a given target. Combining these two probabilities helps reduce the overlap of robot sensory areas toward the goal of minimizing the likelihood of a target escaping detection.

The probability that a target exists is modeled as a decay function based upon when the target was most recently seen, and by whom. In general, the probability decreases inversely with distance from the current robot. Beyond the predictive tracking range of the robot, the probability becomes zero.

The probability that no other robot is already monitoring a nearby target is based upon the target's position and the location of nearby robots. If the target is in range of another robot, then this probability is generally high. In the future, we plan to incorporate the ALLIANCE motivation of "impatience", if a nearby robot does not appear to be satisfactorily observing its local targets (perhaps due to faulty sensors). This impatience will effectively reduces the probability that the other robot is already monitoring nearby targets. In more complex versions of the *CMOMMT* problem, robots could also learn about the viewing capabilities of their teammates, and discount their teammates' observations if that teammate has been unreliable in the past.

The higher-level weight information is combined with the local force vectors to generate the commanded direction of robot movement. This direction of movement is given by:

$$\sum_{i=0}^{N}(FVO_i \times P(exists_i) \times P(NT_i)) + \sum_{j=0}^{M} FVR_j$$

where FVO_k is the force vector attributed to target o_k, $P(exists_k)$ is the probability that target o_k exists, $P(NT_k)$ is the probability that target o_k is not already being tracked, and FVR_l is the force vector attributed to robot r_l. This movement command is then sent to the robot actuators to cause the appropriate robot movements. We also incorporate a low-level obstacle avoidance behavior that overrides these movement commands if it would likely result in a collision.

EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate the effectiveness of the algorithm we designed for the *CMOMMT* problem (which we will refer to as *A-CMOMMT*, we conducted experiments both in simulation and on a team of mobile robots. In the simulation studies, we compared four possible cooperative observation algorithms: (1) *A-CMOMMT* (high-level plus local control), (2) *Local control only*, (3) *Random/linear robot movement*, and (4) *Fixed robot positions.*

In all of these experiments, targets moved according to a "random/linear" movement, which causes the target to move in a straight-line until an obstacle is met, followed by random turns until the target is able to again move forward without collision. The *local control only* algorithm computed the motion of the robots by calculating the unweighted local force vectors between robots and targets. This approach was studied to determine the effectiveness of the high-level reasoning that is incorporated into the *A-CMOMMT* algorithm. The last two algorithms are control cases for the purposes of comparison: the *random/linear robot movement* approach caused robots to move according the the "random/linear" motion defined above, while the *fixed robot positions* algorithms distributed the robots uniformly over the area $\mathcal{S}$, where they maintained fixed positions. In both of these control approaches, robot movements were not dependent upon target locations or movements (other than obstacle avoidance).

We compared these 4 approaches by measuring the average value of the $A(t)$ matrix (see PROBLEM DESCRIPTION section) during the execution of the algorithm. Since the algorithm performance is expected to be a function f of the number of robots n, number of targets m, the range of a given robot's sensor r, and the relative size of the area $\mathcal{S}$, we collected data for a wide range of values of these variables. To simplify the analysis of our results, we defined the area $\mathcal{S}$ as the area within a circle of radius R, fixed the range of robot sensing at 2,600 units of distance, and included no obstacles within $\mathcal{S}$ (other than the robots and targets themselves, and the boundary of $\mathcal{S}$).

We collected data by varying n from 1 to 10, m from 1 to 20, and R from 1,000 to 50,000 units. For each instantiation of variables n, m, and R, we computed the average $A(t)$ value every $\Delta t = 2$ seconds of a run of length 2 minutes; we then repeated this process for 250 runs for each instantiation to derive an average $A(t)$ value for the given values of n, m, and R. In all runs of all 4 algorithms, the targets were placed randomly at the center of $\mathcal{S}$ within a circle of radius 1,000. In all runs of all algorithms (except for *fixed robot positions*), the robots were also placed randomly within the same area as the targets.

To analyze the results of these experiments, we speculated that the function $f(n, m, r, R)$ would be proportional to ratio of the total collective area that could be covered by the robot sensors (i.e. $n\pi r^2$) over the area that would be allotted to one target (call it a *target slot*), were $\mathcal{S}$ divided equally over all targets (i.e. $\frac{\pi R^2}{m}$), we have:

$$f(n, m, r, R) = \frac{n\pi r^2}{\frac{\pi R^2}{m}} = \frac{nmr^2}{R^2}$$

Thus, this function was used to compare the similarity of experiments that varied in their instantiations of n, m, and R.

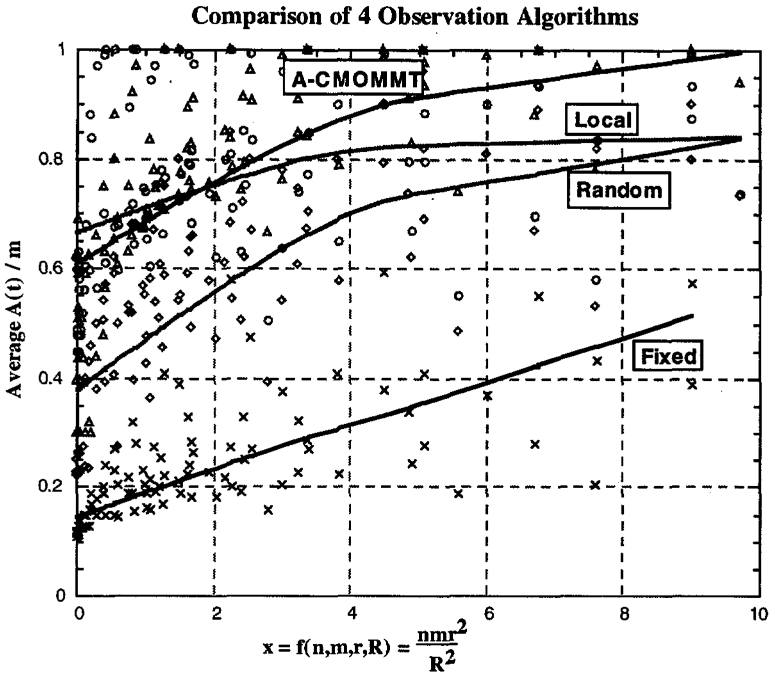


Figure 2: Comparison of 4 cooperative observation algorithms.

Since the optimum value of the average $A(t)$ for a given experiment depends upon the value of m (and, in fact, equals m), we normalized the experiments by plotting the average $A(t)/m$ which is the average percentage of targets that are within some robot's view at a given instant of time.

Figure 2 gives the results of our experiments, plotting the average $A(t)/m$ versus $f(n,m,r,R)$ for all of our experimental data. For each algorithm, we fit a curve to the data using the locally weighted Least Squared error method. Since there is considerable deviation in the data points for given values of $f(n,m,r,R)$, we computed the statistical significance of the results using the Student's t distribution, comparing the algorithms two at a time for all 6 possible pairings. In these computations, we used the null hypothesis: $H_0 : \mu_1 = \mu_2$, *and there is essentially no difference between the two algorithms.* Under hypothesis H_0:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad where \quad \sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1+n_2-2}}$$

Then, on the basis of a two-tailed test at a 0.01 level of significance, we would reject H_0 if T were outside the range $-t_{.995}$ to $t_{.995}$, which for $n_1 + n_2 - 2 = 250 + 250 - 2 = 498$ degrees of freedom, is the range -2.58 to 2.58. For the data given in figure 2, we found that we could reject H_0 at a 0.01 level of significance for all pairing of algorithms that show a visible difference in performance in this figure. Thus, we can conclude that the variation in performance of the algorithms illustrated by the fitted curves in figure 2 is significant.

We see from figure 2 that the *A-CMOMMT* and *local control only* algorithms perform better than the two naive control algorithms, which is expected since the naive algorithms use no information about target positions. Note that all approaches improve as the value of $f(n,m,r,R)$ increases, corresponding to a higher level of robot coverage available per target. The *random/linear robot movement* approach performed better than the *fixed robot positions*, most likely due to the proximity of the initial starting locations of the robots and objects in the *random/linear robot movement* approach. This seems to suggest that much benefit can be gained by learning areas of the environment $\mathcal{S}$ where targets are more likely to be found, and concentrate on locating robots in those areas.

Of more interest, we see that the *A-CMOMMT* approach is superior to the *local control only* approach for values of $f(n,m,r,R)$ greater than about 2; the *local control only* approach is slightly better for $f(n,m,r,R)$ less than 2. This means that when the fraction of robot coverage available per target is low (< 2), relative to the size of $\mathcal{S}$, then robots are better off *not* ignoring any targets, which is essentially what happens due to the high-level control of *A-CMOMMT*. Examples of experimental scenarios where the *local control only* approach is better than the *A-CMOMMT* approach are (n,m,R) = (2,1,5000-50000), (2,2,4000-50000), (3,1,5000-50000), (3,2,5000-50000), (3,3,8000-50000), and (3,4,8000-50000). However, for more complex cases, where the number of targets is much greater than the number of robots, and the environmental area is not "too large", we find that the higher-level reasoning provided by *A-CMOMMT* works better. Examples of scenarios where *A-CMOMMT* is better include (n,m,R) = (2,4,1000-5000), (2,6,1000-6000) (2,20,1000-10000),(3,3,1000-5000), (3,4,1000-6000), (3,6, 1000-7000), and (3,12,1000-11000). Note that *A-CMOMMT* approaches perfect performance as $f(n,m,r,R)$ reaches 10, whereas the results of the *random/linear robot movement* and *local control only* approaches begin to level off at around 85%. In continuing and future work, we are determining the impact of these results on multi-robot cooperative algorithm design.

We have also implemented the *A-CMOMMT* algorithm on a team of a team of four Nomadic Technologies robots to illustrate the feasibility of our approach for physical robot teams. We have demonstrated a very simple case of cooperative tracking using these robots. Refer to [14] for details.

CONCLUSIONS

Many real-world applications in security, surveillance, and reconnaissance tasks require multiple targets to be monitored using mobile sensors. We have presented an approximate, distributed approach based upon the philosophies of the ALLIANCE architecture and have illustrated its effectiveness in a wide range of cooperative observation scenarios. This approach is based upon a combination of high-level reasoning control and lower-level force vector control that is fully distributed across all robot team members and involves no centralized control. Empirical investigations of our cooperative control approach have shown it to be effective at achieving the goal of maximizing target observation for most experimental scenarios, as compared to three other feasible control algorithms.

ACKNOWLEDGEMENTS

This research is sponsored by the Office of Engineering Research, Basic Energy Sciences of the U.S. Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

REFERENCES

[1] J. O'Rourke. Art Gallery Theorems and Algorithms. Oxford University Press, 1987.

[2] Amy J. Briggs. Efficient Geometric Algorithms for Robot Sensing and Control. PhD thesis, Cornell University, 1995.

[3] Kzuo Sugihara, Ichiro Suzuki, and Masafumi Yamashita. The searchlight scheduling problem. SIAM Journal of Computing, 19(6):1024–1040, 1990.

[4] Ichiro Suzuki and Masafumi Yamashita. Searching for a mobile intruder in a polygonal region. SIAM Journal of Computing, 21(5):863–888, 1992.

[5] S. M. LaValle, D. Lin, L. J. Guibas, J-C. Latombe, and R. Motwani. Finding an unpredictable target in a workspace with obstacles. In submitted to 1997 International Conference on Robots and Automation, 1997.

[6] S. M. LaValle, H. H. Gonzalez-Banos, C. Becker, and J-C. Latombe. Motion strategies for maintaining visibility of a moving target. In submitted to 1997 International Conference on Robots and Automation, 1997.

[7] Yaakov Bar-Shalom. Tracking methods in a multitarget environment. IEEE Transactions on Automatic Control, AC-23(4):618–626, 1978.

[8] Yaakov Bar-Shalom. Multitarget Multisensor Tracking: Advanced Applications. Artech House, 1990.

[9] S. S. Blackman. Multitarget Tracking with Radar Applications. Artech House, 1986.

[10] G.C. Fox, R.D. Williams, and P.C. Messina. Parallel Computing Works. Morgan Kaufmann, 1994.

[11] Lynne E. Parker. ALLIANCE: An architecture for fault tolerant, cooperative control of heterogeneous mobile robots. In Proc. of the 1994 IEEE/RSJ/GI Int'l Conf. on Intelligent Robots and Systems (IROS '94), pages 776–783, Munich, Germany, Sept. 1994.

[12] Lynne E. Parker. Heterogeneous Multi-Robot Cooperation. PhD thesis, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Cambridge, MA, February 1994. MIT-AI-TR 1465 (1994).

[13] MTI Research Inc. Conac 3-D tracking system. Operating manual, Chelmsford, MA, 1995.

[14] Lynne E. Parker and Brad Emmons. Cooperative multi-robot observation of multiple moving targets. In Proceedings of the 1997 IEEE International Conference on Robotics and Automation, pages 2082–2089, Albuquerque, New Mexico, April 1997.

GLOBAL OPTIMIZATION FOR MULTISENSOR FUSION IN SEISMIC IMAGING

Jacob Barhen, Vladimir Protopopescu, and David Reister

Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6355

ABSTRACT

The accurate imaging of subsurface structures requires the fusion of data collected from large arrays of seismic sensors. The fusion process is formulated as an optimization problem and yields an extremely complex "energy surface". Due to the very large number of local minima to be explored and escaped from, the seismic imaging problem has typically been tackled with stochastic optimization methods based on Monte Carlo techniques. Unfortunately, these algorithms are very cumbersome and computationally intensive. Here, we present TRUST - a novel deterministic algorithm for global optimization that we apply to seismic imaging. Our excellent results demonstrate that TRUST may provide the necessary breakthrough to address major scientific and technological challenges in fields as diverse as seismic modeling, process optimization, and protein engineering.

INTRODUCTION

In many geophysical tasks, seismic energy is detected by receivers which are regularly spaced along a grid that covers the explored domain. A source is positioned at some grid node to produce a shot. Time series data is collected from the detectors for each shot; then the source is moved to another grid node for the next shot. A major degradation of seismic signals usually arises from near-surface geologic irregularities [1, 2]. These include uneven soil densities, topography, and significant lateral variations in the velocity of seismic waves. The most important consequence of such irregularities is a *distorted image of the subsurface structure*, due to misalignment of signals caused by unpredictable delays in recorded travel times of seismic waves in a vertical neighborhood of every source and receiver. To improve the quality of the seismic analysis, timing adjustments (called "statics corrections") must be performed. One typically distinguishes between "*field statics*", which correspond to corrections that can be derived directly from topographic and well measurements, and "*residual statics*", which incorporate adjustments that must be inferred statistically from the seismic data. The common occurrence of severe residual statics (where the dominant period of the recorded data is significantly exceeded), and the significant noise contamination render the automatic identification of large static shifts extraordinarily difficult. Thus, *multisensor fusion* must be invoked [3]. This problem has generally been formulated in terms of global optimization and, to date, Monte-Carlo techniques (e.g., simulated annealing, genetic algorithms) have provided the primary tools for seeking a potential solution.

The objective function associated with the task of fusing data from an array of seismic sensors depends on a very large number of parameters. Finding the extrema and, in particular, the absolute extremum of such a function turns out to be painstaking difficult. The primary difficulty stems from the fact that the global extremum, say minimum, of a real function is - despite its name - a local property. In other words, significant alteration of the location and magnitude of the global minimum can be carried out without affecting at all the locations and magnitudes of the other minima. Short of exhaustive search, it would then appear extraordinarily unlikely to design unfallible methods to locate the absolute minimum for an arbitrary function. In recent years there has been a remarkable surge of interest in global optimization [5 - 8]. Although significant progress has been achieved in breaking new theoretical ground [9 - 19], the need for efficient and reliable global optimization methods remains as urgent as ever. In particular, a major need exists for a breakthrough paradigm which would enable the accurate and efficient solution of *large-scale* problems. In response to that need, we have been focusing, at ORNL's Center for Engineering Systems Advanced Research (CESAR), on two innovative concepts, namely subenergy tunneling and non-Lipschitzian terminal repellers, to ensure escape from local minima in a fast, reliable, and computationally efficient manner. The generally applicable methodology is embodied in the TRUST algorithm [4], which is deterministic, scalable, and easy to implement. Benchmark results show that TRUST is considerably faster and more accurate than previously reported global optimization techniques. Hence, TRUST may provide the enabling means for addressing major scientific and technological challenges in fields as diverse as seismic modeling, process optimization, and protein engineering.

The classical theory of optimization started to develop almost concomitantly with classical mechanics by trying to find extremal values (minima or maxima) of certain functions that bear special physical meaning and practical significance. For instance, Newton studied projectile trajectories and obtained their maximum range by taking into account the friction with the atmosphere. He was also interested in minimizing resistance by modifying the shape of an object propelled through water. The Bernoulli brothers, who were active in Switzerland between 1670 and 1720, discovered that the shortest time of descent between two points under gravity is achieved not on the straight line joining the two points, but on a convex curve, called brachistocrone. Another famous optimization problem is to find the greatest area enclosed between a straight line and an arbitrary curve of fixed length joining two points on the line. By Virgil's account (Aeneid, Book I, line 367), Queen Dido solved the problem by determining the shape of the curve and the position of the points, thereby founding Carthage.

The completion of the main body of classical physics around the turn of the century came with the realization that many natural processes take place according to extremal principles such as: (i) the principle of stationary (minimum) action in mechanics and electrodynamics; (ii) the principle of minimal potential energy in stable mechanical equilibrium states; (iii) the principle of maximal entropy in isolated thermodynamic systems at equilibrium; and (iv) the principle of motion along geodesics (Fermat's principle in geometrical optics and Einstein's principle in relativity theory). Thenceforth extremal principles and, more generally, optimization problems have been perceived as a systematic and elegant framework for addressing and solving more complex problems with applications to economy, investment policies, and social or political negotiations. In these domains, optimization is, in turn, used to determine "the best" model for a complex situation , to make "the best" choice within a given model, and to solve the associated, purely technical, sub-problems that occur in the mathematical analysis and implementation of the model. In this context, optimality is, almost always, to be obtained under certain constraints and/or at the expense of a certain price.

The generic *global optimization* problem can be stated as follows. The overall performance of a system is described by a multivariate function, called the objective function. Optimality of the system is reached when the objective function attains its global extremum, which can be a maximum or a minimum, depending on the problem under consideration. From an algorithmic perspective, however, there is essentially no difference between the two.

THE TRUST ALGORITHM

We now define the global optimization problem considered in more rigorous terms. Let $f(\mathbf{x}) : \mathcal{D} \to \mathcal{R}$ be a function with a finite number of discontinuities, and $\mathbf{x}$ be an n-dimensional state vector. At any discontinuity point, $\mathbf{x}^\delta$, the function $f(\cdot)$ is required to satisfy the inequality $\lim_{\mathbf{x}\to\mathbf{x}^\delta} \inf f(\mathbf{x}) \geq f(\mathbf{x}^\delta)$ (lower semicontinuity condition). Hereafter, $f(\mathbf{x})$ will be referred to as the objective function, and the set $\mathcal{D}$ as the set of feasible solutions (or the solution space). The goal is to find location of the global minimum, i.e. the value $\mathbf{x}^{gm}$ of the state variables which minimizes $f(\mathbf{x})$,

$$f(\mathbf{x}^{gm}) = \min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathcal{D}\} \ . \tag{1}$$

Without loss of generality, we shall take $\mathcal{D}$ as the hyperparallelepiped

$$\mathcal{D} = \{x_i \ |\beta_i^- \leq x_i \leq \beta_i^+ \ ; \ \ i = 1, 2, \ldots, n\} \ . \tag{2}$$

where β_i^- and β_i^+ denote, respectively, the lower and upper bound of the i-th state variable.

We define the *subenergy tunneling* transformation of the function $f(\mathbf{x})$ by the following nonlinear monotonic mapping:

$$E_{sub}(\mathbf{x}, \mathbf{x}^*) = \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) \ . \tag{3}$$

In Eq. (3), $\hat{f}(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*)$, a is a constant that affects the asymptotic behavior, but not the monotonicity, of the transformation, and $\mathbf{x}^*$ is a fixed value of $\mathbf{x}$, whose selection will be discussed in the sequel. Whenever f is differentiable, the derivative of $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ with respect to $\mathbf{x}$ is given by

$$\partial E_{sub}(\mathbf{x}, \mathbf{x}^*)/\partial \mathbf{x} = (\partial f(\mathbf{x})/\partial \mathbf{x})(1/[1 + \exp(\hat{f}(\mathbf{x}) + a)]) \ , \tag{4}$$

which yields

$$\partial E_{sub}(\mathbf{x}, \mathbf{x}^*)/\partial \mathbf{x} = 0 \quad \Leftrightarrow \quad \partial f(\mathbf{x})/\partial \mathbf{x} = 0 \ . \tag{5}$$

It is clear that $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ has the same discontinuity and critical points as $f(\mathbf{x})$, and the same relative ordering of the local and global minima. In other words, $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ is a transformation of $f(\mathbf{x})$ which preserves all properties relevant for optimization. In addition, this transformation

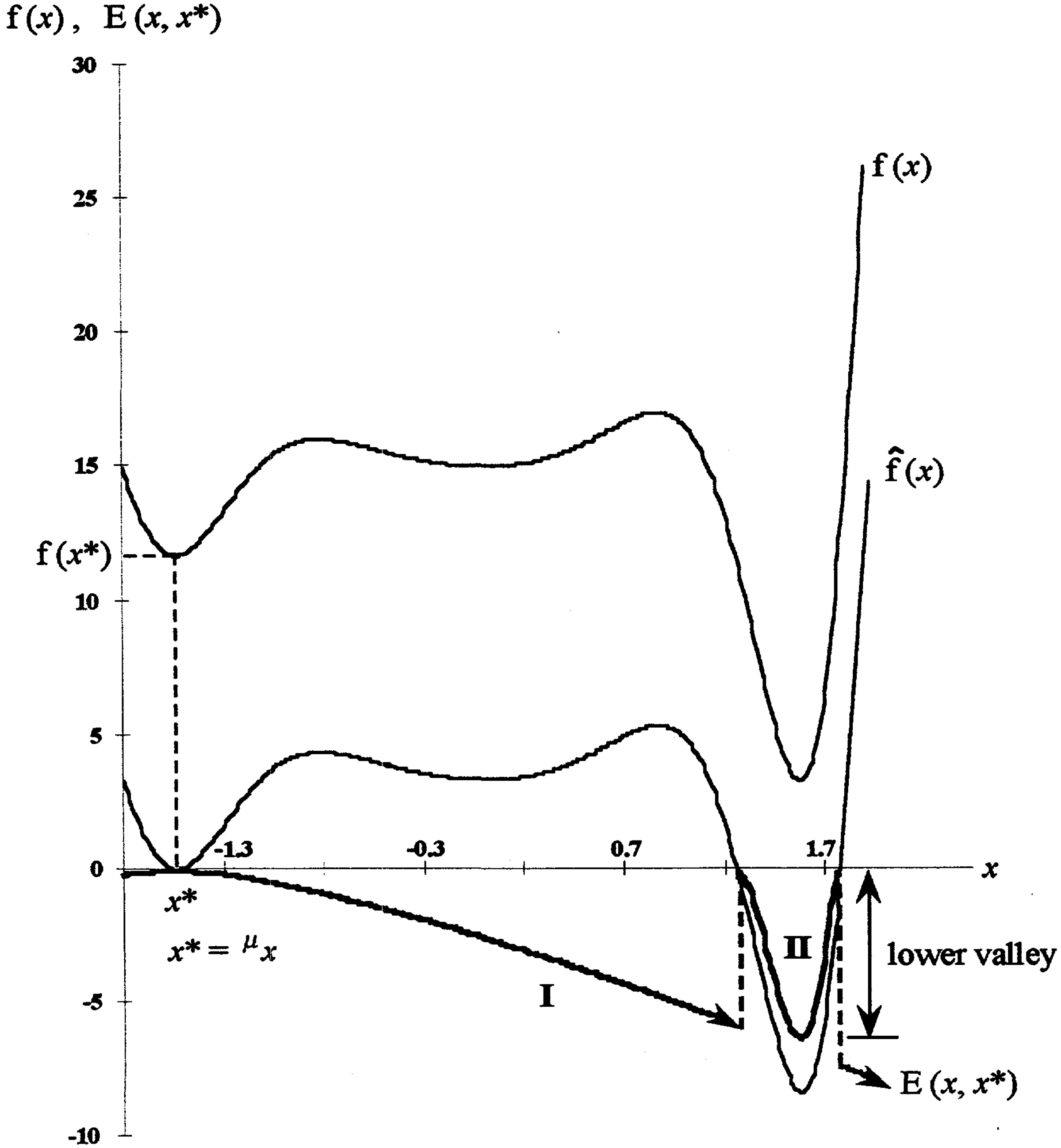


Figure1 . Operation of TRUST, illustrated on the function $f(x) = 4x^2 e^{2\alpha(x-1)} \sin[\frac{\pi}{8}(4x^2+3)]$, with $\alpha \simeq 1.22$.

is designed to ensure that: (i) $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ quickly approaches zero for large positive $\hat{f}(\mathbf{x})$; and (ii) $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ rapidly tends toward $\hat{f}(\mathbf{x})$, whenever $\hat{f}(\mathbf{x}) \ll 0$.

An equilibrium point $\mathbf{x}_{eq}$ of the dynamical system $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ is termed an attractor (repeller) if no (at least one) eigenvalue of the $n \times n$ matrix $\mathcal{M}$, $\mathcal{M} = \partial \mathbf{g}(\mathbf{x}_{eq})/\partial \mathbf{x}$ has a positive real part. Typically, a certain amount of regularity (Lipschitz condition) is required to guarantee the existence of a unique solution for each initial condition $\mathbf{x}(0)$ and, in those cases, the system's relaxation time to an attractor, or escape time from a repeller, is theoretically infinite. If the regularity condition at equilibrium points is violated, singular solutions are induced, such that each solution approaches a *terminal attractor* or escapes from a *terminal repeller* in finite time. The above concepts are at the foundation of our Terminal Repeller Unconstrained Subenegy Tunneling (TRUST) global optimization algorithm.

Let $f(\mathbf{x})$ be a function one wishes to globally minimize over $\mathcal{D}$. We define the TRUST *virtual objective function*

$$\begin{aligned} E(\mathbf{x}, \mathbf{x}^*) &= \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) - (3/4)\rho(\mathbf{x} - \mathbf{x}^*)^{4/3}\theta(\hat{f}(\mathbf{x})) \\ &= E_{sub}(\mathbf{x}, \mathbf{x}^*) + E_{rep}(\mathbf{x}, \mathbf{x}^*). \end{aligned} \tag{6}$$

In the above expression $\theta(\cdot)$ denotes the Heaviside function, that is equal to one for positive values of the argument and zero otherwise. The first term on the right-hand side of Eq. (6) corresponds to the subenergy tunneling function; the second term is referred to as the repeller energy term. The parameter $\rho > 0$ quantifies the strength of the repeller. Application of gradient descent to $E(\mathbf{x}, \mathbf{x}^*)$ results in the dynamical system($i = 1, ..n$)

$$\dot{x}_i = -(\partial f(\mathbf{x})/\partial x_i)(1/[1 + \exp(\hat{f}(\mathbf{x}) + a)]) + \rho(x_i - x_i^*)^{1/3}\theta(\hat{f}(\mathbf{x})) \ . \tag{7}$$

Figure 1 illustrates the main characteristics of TRUST for a one-dimensional problem objective function $E(x, x^*)$. A schematical representation of a sufficiently smooth $f(x)$ is shown, which has three local minima, one of which is the global minimum. We assume that the solution flows in the positive direction (i.e., away from the left boundary), and that the local minimum at $x = {}^{\mu}x$ is encountered by a local minimization method, gradient descent for example. The task under consideration is to escape this local minimum, in order to reach the valley of another minimum with a lower value. We set $x^* = {}^{\mu}x$; then the objective function in Eq. (6) performs the following transformation (see Figure 1):

- the offset function $\hat{f}(x) = f(x) - f(x^*)$ creates the curve parallel to $f(x)$, such that the local minimum at $x = x^*$ intersects with the x-axis tangentially;
- the term $E_{sub}(x, x^*)$ forms the portion of the thick line denoted by **II** (i.e., the lower valley) as a result of the properties of the subenergy transformation;
- the repeller energy term $E_{rep}(x, x^*)$ essentially constitutes the portion of the thick line denoted by **I**;
- finally, as the complete thick line (i.e., **I** and **II**) shows, the virtual objective function $E(x, x^*)$, which is a superposition of these two terms, creates a discontinuous but well-defined function with a *global maximum* located at the previously specified local minimum ${}^{\mu}x$.

To summarize, as seen in Figure 1, $E(x, x^*)$ of Eq. (7) transforms the current local minimum of $f(x)$ into a global maximum, but preserves all lower local minima. Thus, when gradient descent is applied to the function $E(x, x^*)$, the new dynamics, initialized at a small perturbation from the local minimum of $f(x)$ (i.e., at $x = x^* + d$, with $x^* = {}^{\mu}x$), will escape this critical point (which is also the global maximum of $E(x, x^*)$) to a lower valley of $f(x)$ with a lower local minimum. It is important to note that the discontinuity of $E(x, x^*)$ does not affect this desired operation, since the flow of the gradient descent dynamics follows the gradient of $E(x, x^*)$, which is well-defined at every point in the region. It is clear that if gradient descent were to be applied to the objective function $f(x)$ under the same conditions, escaping the local minimum at $x = {}^{\mu}x$ would not be accomplished.

Hence, application of gradient descent to the function $E(x, x^*)$ as defined in Eq. (6), as opposed to the original function $f(x)$, results in a system that has a *global descent property*, i.e., the new system escapes the encountered local minimum to another one with a lower functional value. This is the main idea behind constructing the TRUST virtual objective function of Eq. (6). Additional details and formal derivations can be found in [4, 15, 18].

BENCHMARKS AND COMPARISONS TO OTHER METHODS

This section presents results of benchmarks carried out to assess the TRUST algorithm using several standard test functions taken from the literature. A description of each test function is given in Table 1. In Tables 2-3, the performance of TRUST is compared to the best competing global optimization methods, where the term "best" indicates the best widely reported reproducible results the authors could find for the particular test function. The criterion for comparison is the number of function evaluations.

One of the primary limitations of conventional global optimization algorithms is their lack of stopping criteria. This limitation is circumvented in benchmark problems, where the value and coordinates of the global minima are known in advance. The achievement of a desired accurracy (e.g., $\epsilon = 10^{-6}$) is then considered as a suitable termination condition [6]. For consistent comparisons, this condition has also been used in TRUST, rather than its general stopping criterion described earlier. For each function, corners of the domain were taken as initial conditions; each reported result then represents the average number of evaluations required for convergence to the global minimum of the particular function. The TRUST calculations were performed using the value $a = 2$, for which the subenergy tunneling transformation achieves its most desirable asymptotic behavior [15]. The dynamical equations were integrated using an adapitive scheme, that, within the basin of attraction of a local minimum, considers the local minimum as a terminal attractor. Typical base values for the key parameters Δ_t and ρ were 0.05 and 10., respectively.

In Table 2, the benchmark labels, i.e. BR (Branin), CA (Camelback), GP (Goldstein-Price), RA (Rastrigin), SH (Shubert) and H3 (Hartman), refer to the test functions specified in Table 1. The following abbreviations are also used: SDE is the stochastic method of Aluffi-Pentini [9]; EA is the annealing evolution algorithms of Yong, Lishan, and Evans [17] and Schneider [19]; MLSL is the multiple level single linkage method of Kan and Timmer [10]; IA is the interval arithmetic technique of Ratschek and Rokne [19]; TUN is the tunneling method of Levy and Montalvo [11]; and TS refers to the Taboo Search scheme of Cvijovic and Klinowski [16]. The results demonstrate that TRUST is substantially faster than these state-of-the-art methods.

Table 1. Standard Test Functions used for global optimization benchmarks.

Name	Definition	Domain
Branin	$f(\mathbf{x}) = [x_2 - (5.1/4\pi^2)x_1^2 + (5/\pi)x_1 - 6]^2 + 10(1 - 1/8\pi)\cos x_1 + 10$	$-5. \leq x_1 \leq +10.$ $0. \leq x_2 \leq +15.$
Camelback	$f(\mathbf{x}) = [4 - 2.1x_1^2 + (x_1^4/3)]\,x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2$	$-3. \leq x_1 \leq +3.$ $-2. \leq x_2 \leq +2.$
Goldstein-Price	$f(\mathbf{x}) = [1 + (x_1 + x_2 + 1)^2\,(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times[30 + (2x_1 - 3x_2)^2\,(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	$-2. \leq x_i \leq +2.$
Rastrigin	$f(\mathbf{x}) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2)$	$-1. \leq x_i \leq +1.$
Shubert	$f(\mathbf{x}) = \left\{\sum_{i=1}^{5} i\cos[(i+1)x_1 + i]\right\}\left\{\sum_{i=1}^{5} i\cos[(i+1)x_2 + i]\right\}$	$-10. \leq x_i \leq +10.$
Hartman*	$f(\mathbf{x}) = \sum_{i-1}^{i=4} c_i \exp[-\sum_{j=1}^{j=N} a_{ij}(x_j - p_{ij})^2]$	$0. \leq x_i \leq 1.$
Styblinski and Tang	$f(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{i=2}(x_i^4 - 16x_1^2 + 5x_i) + \sum_{i=3}^{i=5}(x_i - 1)^2$	$-4.6 \leq x_i \leq +4.6$

(*)The values of the parameters are given in ([6], p. 185).

Table 2. Number of function evaluations required by different methods to reach a global minimum of Standard Test Functions.

Method	BR	CA	GP	RA	SH	H3
SDE	2700	10822	5439	–	241215	3416
EA	430	–	460	5917	–	–
MLSL	206	–	148	–	–	197
IA	1354	326	–	–	7424	–
TUN	–	1469	–	–	12160	–
TS	492	–	486	540	727	508
TRUST	55	31	103	59	72	58

Table 3. Number of function evaluations and precision for Styblinski and Tang function. Global minimum FSA and SAS results taken from Ref. [13].

Method	FSA	SAS	TRUST	Exact
Cost	100,000	3,710	89	n/a
x_1	-2.702844	-2.903506	-2.90353	-2.903534
x_2	-3.148829	-2.903527	- 2.90353	-2.903534
x_3	1.099552	1.000241	1.00004	1.
x_4	1.355916	0.999855	0.99997	1.
x_5	1.485936	1.000194	0.99997	1.

In Table 3, FSA is the fast simulated annealing algorithm of Szu [12], and SAS denotes the stochastic approximation paradigm of Styblinski and Tang [13]. As can be observed, TRUST is not only much faster, but produces very consistent and accurate results. Therefore, it seemed the ideal candidate for the solution of the notoriously difficult problem of multisensor fusion for seismic imaging, formulated as residual statics optimiation.

RESIDUAL STATICS CORRECTIONS FOR SEISMIC DATA

Statics optimization is typically done in a surface consistent manner to seismic traces corrected for *normal moveout* [3]; consequently, the correction time shifts depend only on the shot and receiver positions, and not on the ray path from shot to receiver. Shot corrections **S** correspond to wave propagation times from the shot locations to a reference plane, while the receiver corrections **R** are propagation times from the reference plane to receiver locations. From an operational perspective, data D_{ft} are provided by trace $(t = 1, \dots N_t)$, and sorted to midpoint offset coordinates (common midpoint *stacking*). For each trace, the data consist of the complex Fourier components $(f = 1, \dots N_f)$ of the collected time series. Each trace t corresponds to seismic energy travel from a source s_t to a receiver r_t via a midpoint k_t. Assuming the availability of N_k common midpoints, we seek statics corrections **S** and **R** that maximize the total power E in the stacked data:

$$E = \sum_k \sum_f | \sum_t \exp[2\pi i \, f \, (S_{s_t} + R_{r_t})] D_{ft} \delta_{kk_t} |^2 \quad . \tag{8}$$

The above expression highlights the multimodal nature of E which, even for relatively low dimensional **S** and **R**, exhibits a very large number of local minima. This is illustrated in Figure 2.

To assess the performance of TRUST, we considered a problem involving 77 shots and 77 receivers. A dataset consisting of 1462 synthetic seismic traces folded over 133 common midpoint gathers was obtained from CogniSeis Corporation (J. DuBose). It uses 49 Fourier components for data representation. Even though this set is somewhat smaller than typical collections obtained during seismic surveys by the oil industry, it is representative of the extreme complexity underlying residual statics problems. To derive a quantitative estimate of TRUST's impact, let E_k denote the total contribution to the stack power arising from midpoint k, and let B_k refer to the upper bound of E_k in terms of **S** and **R**. Using a polar coordinates representation for the trace data D_{ft}, i.e., writing $D_{ft} = \alpha_{ft} \exp(iw_{ft})$, we can prove that

$$B_k = \sum_f (\sum_t \alpha_{ft} \, \delta_{kk_t})^2 \quad . \tag{9}$$

The TRUST results, illustrated in Figure 3, show the dramatic improvement in the coherence factor of each common gather. This factor is defined as the ratio $\kappa_k = E_k / B_k$, and characterizes the overall quality of the seismic image.

CONCLUSIONS

TRUST is a novel methodology for unconstrained global function optimization, that combines the concepts of subenergy tunneling and non-Lipschitzian "terminal repellers." The evolution of a deterministic nonlinear dynamical system incorporating these concepts provides the computational

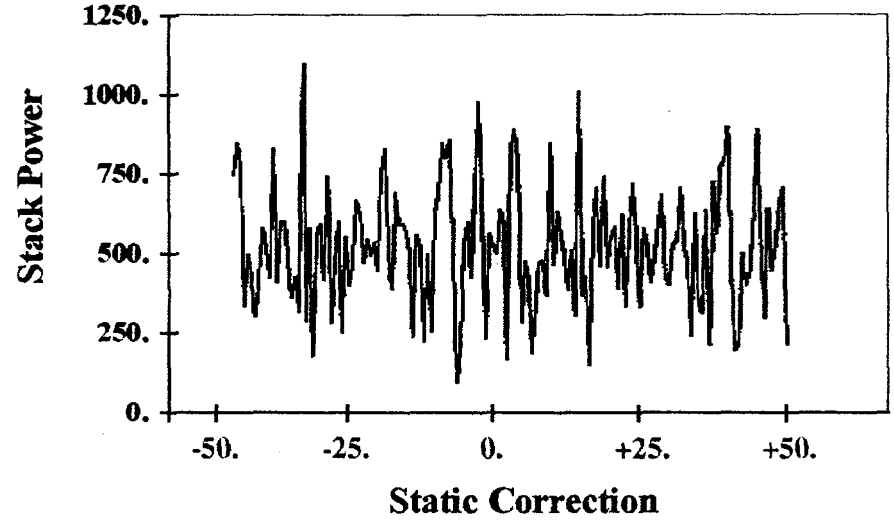


Figure 2. One-dimensional slice through a 154-dimensional objective function associated with a residual statics problem.

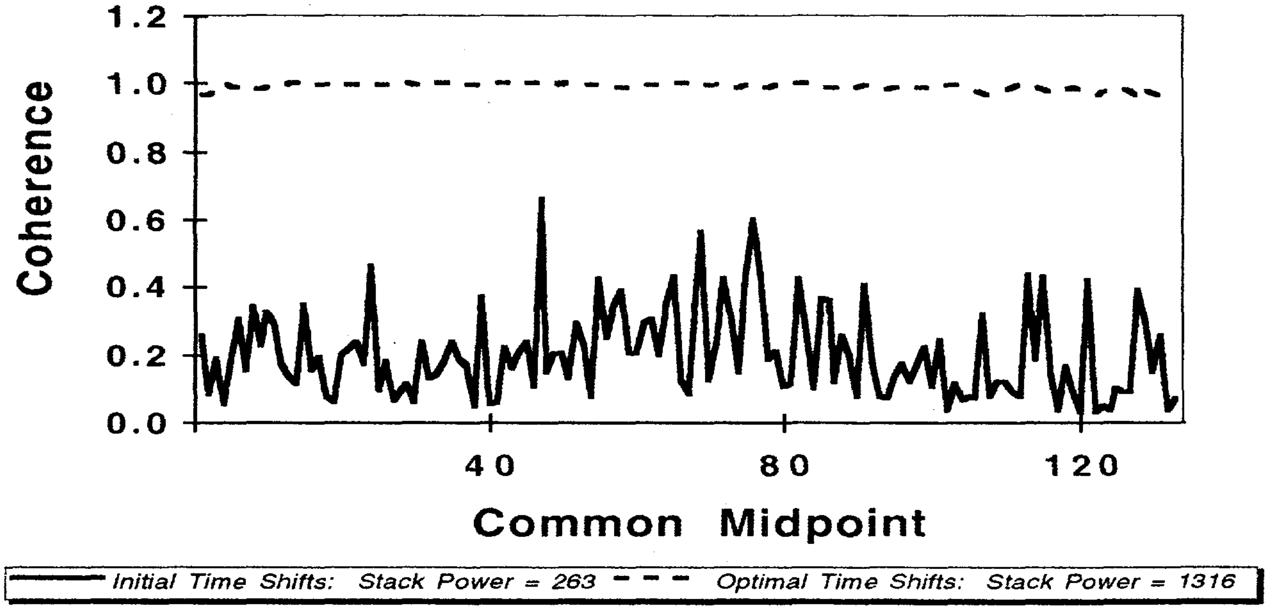


Figure 3. The coherence factors, i.e., the dimensionless ratios E_k/B_k, are plotted for each common gather using the initial and the optimal time shifts ("residual statics"). Ideally, at the global optimum, these ratios should be equal to one.

mechanism for reaching the global minima. The benchmark results demonstrate that TRUST is substantially faster, as measured by the number of function evaluations, than other global optimization techniques for which reproducible results have been published in the open literature. The application of TRUST to the problem of multisensor fusion for accurate seismic imaging (residual statics corrections) proves that the method is not a mere academic exercise for toy problems, but has the robustness and consistency required by large-scale, real-life applications.

Acknowledgments

The research described in this paper was performed at the Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory. Funding for the effort was provided by the Engineering Research Program of the Office of Basic Energy Sciences, and by the Advanced Computing Technology Initiative Program of the U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research, Corp.

References

[1] Rothman, D., *Geophysics*, **50**(12), 2784-2796 (1985).

[2] DuBose, J., *Geophysics*, **58**(3), 399-407 (1993).

[3] Yilmaz, O., *Seismic Data Processing*, Society of Exploration Geophysicists, 1988.

[4] Barhen, J., V. Protopopescu, and D. Reister, *Science*, **276**, 1094-1097 (1997).

[5] Ratschek, H. and J. Rokne,*New Computer Methods for Global Optimization*, Ellis Horwood, 1988.

[6] Törn, A. and A. Zilinskas, *Global Optimization*, Springer-Verlag, 1989.

[7] Horst, R., P. M. Pardalos, and N. V. Thoai, Kluwer Academin Publishers, 1996.

[8] Floudas, C. A. and P. M. Pardalos, *State of the Art in Global Optimization: Computational Methods and Applications*, Kluwer Academic Publishers, 1996.

[9] Aluffi-Pentini, F., V. Parisi, and F. Zirilli, *Journal of Optimization Theory and Applications*, **47**, 1–15 (1985).

[10] Kan, A. H. and G. T. Timmer, in *Numerical Optimization*, eds. P. T. Boggs *et al*, pp. 245–262 SIAM, 1985.

[11] Levy, A. V. and A. Montalvo, *SIAM Journal on Scientific and Statistical Computing*, **6**, 15–29 (1985).

[12] Szu, H. and R. Hartley, *Physics Letters*, **A 122**, 157–162 (1987).

[13] Styblinski, M. A. and T. S. Tang, *Neural Networks*, **3**, 467–483 (1990).

[14] Ammar, H. and Y. Cherruault, *Math. Comp. and Modeling*, **18**, 17–21 (1993).

[15] Cetin, B., J. Barhen, and J. Burdick, *J. Opt. Theory and Appl.*, **77**, 97–126 (1993).

[16] Cvijovic, D. and J. Klinowski, *Science*, **267**, 664-666 (1995).

[17] Yong, L., K. Lishan, and D. J. Evans, *Parallel Computing*, **21**, 389-400 (1995).

[18] Barhen, J. and V. Protopopescu, in *State of the Art in Global Optimization*, C.A. Floudas and P.M. Pardalos eds., pp. 163-180, Kluwer Academic Press, 1996.

[19] Schneider, J., and P. Schuchhardt, *Biol. Cybern*, **74**, 203-207 (1996).

ALGORITHMS FOR FUSION OF MULTIPLE SENSORS HAVING UNKNOWN ERROR DISTRIBUTIONS

Nageswara S. V. Rao

Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

ABSTRACT

The sensor S_i, $i = 1, 2 \ldots, N$, of a multiple sensor system outputs $Y^{(i)} \in \Re$ in response to input $X \in \Re$ according to an unknown probability distribution $P_{Y^{(i)}|X}$. For a fusion rule $f : \Re^N \mapsto \Re$ the expected square error is given by $I(f) = E[(X - f(Y))^2]$. When only a training sample is available, f^* that minimizes $I(.)$ over a family of functions $\mathcal{F}$ cannot be computed since the underlying distributions are unknown. We consider methods to compute an estimator $\hat{f}$ such that $I(\hat{f}) - I(f^*) < \epsilon$ with probability $1 - \delta$, for any $\epsilon > 0$ and $0 < \delta < 1$. We present a general method based on the scale-sensitive dimension of $\mathcal{F}$. We then review two computational methods based on the Nadaraya-Watson estimator, and the finite dimensional vector spaces.

INTRODUCTION

In a number of engineering applications, there has been an increased need for solving difficult sensor fusion problems. The fuser is very critical in these problems since an inappropriate fuser can render the system worse than the worst individual sensor. Additionally, the fuser must be efficiently computable in order to be of practical use. Early sensor fusion methods require either independence of sensor errors or closed-form analytical expressions for error densities. Under the first condition a general majority rule suffices, while under the second condition the Bayesian methods can be used to design the fuser. Furthermore, there have been only a limited number of studies on the computational aspects of sensor fusion problems. In practical applications, however, independence can seldom be assured and, in fact, may not be satisfied. The fusion rules are typically obtained from a specific function class which can be chosen to make the estimation problem simple, while the underlying distributions cannot be so chosen since they depend on the sensors. As a result, the problem of obtaining the probability densities required by the Bayesian methods can be more difficult than the fusion problem itself (in an information theoretic sense). When sensors are available for operation, one can collect "empirical data" by sensing objects with known parameters. Such data can then be exploited to solve the fusion rule estimation problems under very general conditions as shown in [1]. In this paper, we generalize the results of [1] in terms of function classes by using the scale-sensitive dimension [2], and also by removing the requirement of sensor error densities.

Consider a system of N sensors such that corresponding to input $X \in \Re$, the sensor S_i, $i = 1, 2, \ldots, N$, outputs $Y^{(i)} \in \Re$ according to an *unknown* distribution $P_{Y^{(i)}|X}$. An independently and identically distributed (iid) training n-sample $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ is given where $Y_i = (Y_i^{(1)}, Y_i^{(2)}, \ldots, Y_i^{(N)})$ and $Y_i^{(j)}$ is the output of S_j in response to input X_i. We consider the *expected square error*

$$I(f) = \int [X - f(Y))]^2 dP_{Y,X}, \tag{1.1}$$

where $Y = (Y^{(1)}, Y^{(2)}, \ldots, Y^{(N)})$, to be minimized over a family of fusion rules $\mathcal{F}$, based on the given n-sample. In general, the *expected best* fusion rule f^* that minimizes $I(.)$ over $\mathcal{F}$ cannot be computed since the underlying distributions are unknown. We consider conditions under which, based on a sufficiently large sample, an estimator $\hat{f}$ can be computed such that

$$P[I(\hat{f}) - I(f^*) > \epsilon] < \delta, \tag{1.2}$$

where $\epsilon > 0$ and $0 < \delta < 1$. Informally, Eq. (1.2) states that the "error" of $\hat{f}$ is within ϵ of the optimal error (of f^*) with arbitrary high probability $1 - \delta$, *irrespective* of the underlying sensor distributions.

The sensor fusion problem (1.1) is solved under the criteria (1.2) in [1] using empirical estimation methods of Vapnik [3] when $\mathcal{F}$ has a finite capacity. The computational problems associated with this approach are intractable even for simple function classes. Under additional conditions, the stochastic approximation algorithms are shown to solve this problem [4], but, these conditions are hard to verify in practical cases. Sample size estimates to ensure the criterion (1.2) using feedforward sigmoidal neural networks are derived in [5] based on three different properties. A polynomial-time solution is obtained in [6] using the classical Nadaraya-Watson estimator when: (a) the densities corresponding to $P_{Y^{(i)}|X}$ exist and are smooth, and (b) the function class itself is smooth. When the function class $\mathcal{F}$ forms a finite dimensional vector space, finite sample results as well as polynomial-time computation are obtained in [7]. A review of the last two methods will be presented in this paper.

PRELIMINARIES

Let S be a set equipped with a pseudodometric ν. The *covering number* $N(\epsilon, \nu, S)$ is defined as the smallest number of closed balls of radius ϵ, and centers in S, whose union covers S. Let d_∞^m be a specific pseudometric defined on $[0,1]^m$ such that for $a, b \in [0,1]^m$, we have $d_\infty^m(a, b) = \max\limits_{1 \leq i \leq m} |a_i - b_i|$ where $a = (a_1, a_2, \ldots, a_m)$ and $b = (b_1, b_2, \ldots, b_m)$.

Let $\mathcal{F}$ be a class of $[0,1]$-valued functions on some domain set D and let ρ be a positive real number. We say that $\mathcal{F}$ *P_ρ-shatters* a set $A \subseteq D$ if there exists a function $s : A \mapsto [0,1]$ such that for every $E \subseteq A$ there exists some $f_E \in \mathcal{F}$ satisfying: for every $x \in A - E$, $f_E(x) \leq s(x) - \rho$, and for every $x \in E$, $f_E(x) \geq s(x) + \rho$. Let the P_ρ-dimension of $\mathcal{F}$, denoted by P_ρ-dim $(\mathcal{F})$, be the maximal cardinality of a set $A \subseteq D$ that is P_ρ-shattered by $\mathcal{F}$.

Let Q denote the unit cube $[0,1]^N$ and $\mathcal{C}(Q)$ denote the set of all continuous functions defined on Q. The modulus of smoothness of $f \in \mathcal{C}(Q)$ is defined as

$$\omega_\infty(f;r) = \sup_{\|y-z\|_\infty < r,\ y,z\in Q} |f(y) - f(z)|$$

where $\| y - z \|_\infty = \max_{i=1}^{M} |y_i - z_i|$. For $m = 0, 1, \ldots$, let Q_m denote a family of diadic cubes (Haar system) such that $Q = \bigcup_{J\in Q_m} J$, $J \cap J' = \emptyset$ for $J \neq J'$, and the N-dimensional volume of J, denoted by $|J|$, is 2^{-Nm}. Let $1_J(y)$ denote the indicator function of $J \in Q_m$: $1_J(y) = 1$ if $y \in J$, and $1_J(y) = 0$ otherwise. For given m, we define the map P_m on $\mathcal{C}(Q)$ as follows: for $f \in \mathcal{C}(Q)$, we have $P_m(f) = P_m f$ defined by $P_m f(y) = \frac{1}{|J|}\int_J f(z)dz$ for $y \in J$ and $J \in Q_m$ [8]. Note that $P_m f : Q \mapsto [0,1]$ is a discontinuous (in general) function which takes constant values on each $J \in Q_m$. The *Haar kernel* is given by

$$P_m(y,z) = \frac{1}{|J|} \sum_{J\in Q_m} 1_J(y) 1_J(z)$$

for $y, z \in Q$.

GENERAL SOLUTIONS FOR FUSER DESIGN

In this section, we consider conditions for solving the general sensor fusion problem in (1.1) under the criterion (1.2). Consider the *empirical error* of $f \in \mathcal{F}$ given by

$$I_{emp}(f) = \frac{1}{l}\sum_{i=1}^{n}[X_i - f(Y_i)]^2 \tag{3.1}$$

based on the sample $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$. To approximate $f^* \in \mathcal{F}$ that minimizes the expected error in (1.1), we minimize instead the empirical error in (3.1) to obtain the *empirical best* fusion rule $\hat{f}$. The following theorem presents an estimate of the sample size to ensure the condition (1.2) when $\mathcal{F}$ has finite scale-sensitive dimension [2] and $X \in [0,1]$.

Theorem 1 *Let f^* and $\hat{f}$ denotes the expected best and empirical best fusion rules chosen from a function class $\mathcal{F}$ with range* $[0,1]$. *Given an iid sample of size*

$$\frac{5040}{\epsilon^2} \max\left\{ d \ln^2 \frac{50d}{\epsilon}, \ln\frac{48}{\delta} \right\}$$

where $d = P_{\epsilon/4}$-dim ($\mathcal{F}$), we have $P[I(\hat{f}) - I(f^*) > \epsilon] < \delta$.

Proof: From Vapnik [9] we have $P\left\{I(\hat{f}) - I(f^*) > \epsilon\right\} < P\left\{\sup_{g\in\mathcal{G}} |I_{emp}(f) - I(f)| > \epsilon/2\right\}$, where $\mathcal{G} = \{(x - f(y))^2 : f \in \mathcal{F}\}$.

Let $\mathcal{G}_{2n} = \{g(X_1, Y_1), g(X_2, Y_2), \ldots, g(X_{2n}, Y_{2n}) | g \in \mathcal{G}\} \subseteq [0,1]^{2n}$, based on an iid sample of size $2n$. From Lemma 3.3 and 3.4 of [2], we have.

$$\begin{aligned} P\left\{\sup_{g\in\mathcal{G}} |I_{emp}(f) - I(f)| > \epsilon/2\right\} &\leq 24nE_{X^{2n}}[N(\epsilon/12, d_{\infty}^{2n}, \mathcal{G}_{2n})]e^{-\epsilon^2 n/144} \\ &\leq 48n\left(\frac{4608n}{\epsilon^2}\right)^{d\log_2(96en/(d\epsilon))} e^{-\epsilon^2 n/144} \end{aligned}$$

where $d = P_{\epsilon/4}$-dim $(\mathcal{F})$. By equating the right hand side to δ, we obtain our sample size estimate. The derivation closely follows that of Theorem 1 of [2]. □

The result of Theorem 1 is more general than that in [1] which is based on the capacity of $\mathcal{F}$ [9] in that finiteness of capacity implies that of scale-sensitive dimension but not vice versa. This theorem can be generalized in a straight forward manner to handle the cases: (a) $Y^{(i)}$ is a multi-dimensional vector from $\Re^d$, and/or (b) $X \in [0, \tau]$, $\tau > 0$. The cost function can also be generalized to Lipschitz cost functions with an appropriate change in the sample size (see [10]).

The sample bound of Theorem 1 is based on uniform convergence of empirical means to their expectations for function classes, which are available from the empirical process theory [11, 12] and its applications to machine learning [3, 13]. Results of this kind are available based on a number of characterizations of $\mathcal{F}$ such as pseudo-dimension [11], fat VC-dimension [14], scale-sensitive dimension [2], graph dimension [15], and Euclidean parameters [12], which can be used to obtain sample size estimates along the lines of Theorem 1. Finiteness of these parameters is only sufficient for the "learnability" of bounded functions, while that of the scale sensitive dimension is both necessary and sufficient [2]. Moreover, the latter is only such deterministic quantity known to us, while other similar quantities are based on expected capacity or entropy [3].

A solution based on Theorem 1 simply requires that $\hat{f}$ minimize the empirical error, and does not specify methods to *compute* it. The problem of computing $\hat{f}$ in this general framework is intractable; for example in the special case that $\mathcal{F}$ is set of feedforward neural networks with threshold hidden units, this problem is NP-complete even for simple architectures. In the next sections, we consider more restrictive cases where $\mathcal{F}$ is chosen to be a special class to make the computational problems easier.

FUSERS BASED ON NADARAYA-WATSON ESTIMATOR

We now present a polynomial-time (in sample size n) computable estimator which guarantees the criterion (1.2) under additional smoothness conditions. Given an n-sample, the Nadaraya-Watson estimator based on Haar kernels is defined by

$$\hat{f}_{m,n}(y) = \frac{\sum\limits_{j=1}^{n} X_j P_m(y, Y_j)}{\sum\limits_{j=1}^{n} P_m(y, Y_j)} = \frac{\sum\limits_{Y_j \in J} X_j}{\sum\limits_{Y_j \in J} 1_J(Y_j)}$$

for $y \in J$ [16] (see also Engel [17]). The second expression indicates that $\hat{f}_{m,n}(y)$ is the mean of the function values corresponding to Y_j's in J that contains y. This property is the key to efficient computation of the estimate [18].

Theorem 2 [6] *Consider a family of functions $\mathcal{F} \subseteq \mathcal{C}(Q)$ with range $[0,1]$ such that $\omega_\infty(f;r) \leq kr$ for some $0 < k < \infty$. We assume that: (i) there exists a family of densities $\mathcal{P} \subseteq \mathcal{C}(Q)$; (ii) for each $p \in \mathcal{P}$, $\omega_\infty(p;r) \leq kr$; and (iii) there exists $\mu > 0$ such that for each $p \in \mathcal{P}$, $p(y) > \mu$ for all $y \in [0,1]^N$. Suppose that the sample size, n, is larger than*

$$\frac{2^{2m+4}}{\epsilon_1^2}\left[\left(\frac{k2^m}{\epsilon_1}\left[\left(\frac{k2^m}{\epsilon_1}-1\right)^{N-1}+1\right]+m\right)\ln\left(2^{m+1}k/\epsilon_1\right)+\ln\left(\frac{2^{2m+6}}{(\delta-\lambda)\epsilon_1^4}\right)\right]$$

where $\epsilon_1 = \epsilon(\mu-\epsilon)/4$, $0 < \beta < \frac{N}{2(N+1)}$, $m = \lceil\frac{\log n\beta}{N}\rceil$ and $\lambda = b\left(\frac{2}{\epsilon}\right)^{1/N+1-1/2\beta} + b\left(\frac{2}{\epsilon_1}\right)^{1/N+1-1/2\beta}$. Then for any $f \in \mathcal{F}$, we have $P\left[I(\hat{f}_{m,n}) - I(f^*)| > \epsilon\right] < \delta$.

The value of $\hat{f}_{m,n}(y)$ can be computed in $O((\log n)^N)$ time after a preprocessing step in $O(n(\log n)^{N-1})$ time (see [18]). The smoothness conditions required in Theorem 2 are not very easy to verify in practice. However, this estimators is found to perform well in a number of applications including those that do not have smoothness properties.

VECTOR SPACE METHODS

We now consider the case when $\mathcal{F}$ forms a finite dimensional vector space. The advantages of vector space methods over the existing methods are three-fold: (a) the sample size estimate is a simple function of the dimensionality of $\mathcal{F}$, (b) the estimate can be easily computed by well-known least square methods in polynomial time, and (c) no smoothness conditions are required on the functions or distributions.

Theorem 3 [7] *Let f^* and $\hat{f}$ denote the expected best and empirical best fusion functions chosen from a vector space $\mathcal{F}$ of dimension d_V and range $[0,1]$. Given an iid sample of size*

$$\frac{512}{\epsilon^2}\left[d_V \ln\left(\frac{64e}{\epsilon} + \ln\frac{64e}{\epsilon}\right) + \ln(8/\delta)\right],$$

we have $P[I(\hat{f}) - I(f^*) > \epsilon] < \delta$.

Let $\{f_1, f_2, \ldots, f_{d_V}\}$ be a basis of $\mathcal{F}$ such that $f \in \mathcal{F}$ can be written as $f(y) = \sum_{i=1}^{d_V} a_i f_i(y)$ for $a_i \in \Re$. Then consider $\hat{f} = \sum_{i=1}^{d_V} \hat{a}_i f_i(y)$ such that $\hat{a} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{d_V})$ minimizes the cost expressed as (with abuse of notation)

$$I_{emp}(a) = \frac{1}{n}\sum_{k=1}^{n}\left(X_k - \sum_{i=1}^{d_V} a_i f_i(Y_k)\right)^2,$$

where $a = (a_1, a_2, \ldots, a_{d_V})$. Now $I_{emp}(a)$ can be written in the quadratic form $a^T C a + a^T D$, where $C = [c_{ij}]$ is a positive definite symmetric matrix, and D is a vector. This problem can be solved in polynomial-time using quadratic programming methods [19].

The potential functions of Aizerman *et al.* [20], where $f_i(y)$ is of the form $exp((y-\alpha)^2/\beta)$ for suitably chosen constants α and β, constitute an example of the vector space methods. Note that the above sample size is valid only for the method that minimizes $I_{emp}(.)$ and is not valid for the original incremental algorithm of the potential functions.

The two-layer sigmoidal networks of Kurkova [21], where the unknown weights are only in the output layer, constitute another example for the vector space methods. The specific form of these networks enables us to express each network in the form $\sum_{k=1}^{d_V} a_i\eta_i(y)$ where $\eta_i(.)$'s are universal.

APPLICATION

We consider the problem of recognizing a door (an opening) wide enough for a mobile robot to move through. The mobile robot (TRC Labmate) is equipped with an array of four ultrasonic and four infrared Boolean sensors on each of four sides as shown in Figure 1. The sensors are periodically polled while the robot is in motion. This example deals with only the problem of detecting a wide enough door when the sensor array of any side is facing it. The ultrasonic sensors return a measurement corresponding to distance to an object within a certain cone as illustrated in Figure 1. The infrared sensors return Boolean value based on the light reflected by an object in the line-of-sight of the sensor; white smooth objects are detected due to high reflectivity, while objects with black or rough surface are generally not detected. Both ultrasonic and infrared sensors are unreliable. It is very difficult to derive accurate probabilistic models for these sensors since it requires a detailed knowledge of the physics and engineering of the device as well as a priori statistical information. Thus a Bayesian solution to this problem is very hard to implement. We employ the Nadaraya-Watson estimator to derive a non-linear relationship between the width of the door and the sensor readings. Here the training sample is generated by actually recording the measurements while the sensor system is facing the door. Positive examples are generated if the door is wide enough for the robot, and the sensory system is facing the door. Negative examples are generated when the door is not wide enough or the sensory system is not correctly facing a door (wide enough or not). The robot is manually located in various positions to generate the data. Consider the sensor array of a particular side of the mobile robot. Here Y_1, Y_2, Y_3, Y_4 correspond to the normalized distance measurements from the four ultrasonic sensors, and Y_5, Y_6, Y_7, Y_8 correspond to the Boolean measurements of the infrared sensors. X is 1 if the sensor system is correctly facing a wide enough door, and is 0 otherwise. The training data included 6 positive examples and 12 negative examples. The test data included 3 positive examples and 7 negative examples. The Nadaraya-Watson estimator predicted the correct output in all examples of test data.

CONCLUSIONS

We presented recent results on a general sensor fusion problem, where the underlying sensor error distributions are not known, but a sample is available. We presented a general method for obtaining a fusion rule based on scale-sensitive dimension of the function class.

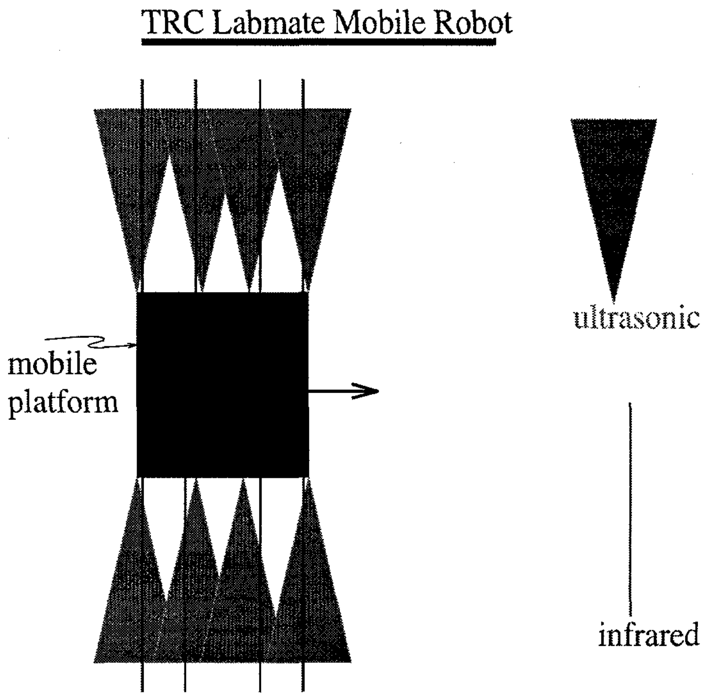


Figure 1: Schematic of sensory system (only the side sensor arrays are shown for simplicity).

Two computationally viable methods are reviewed based on the Nadaraya-Watson estimator, and the finite dimensional vector spaces.

Several computational issues of the fusion rule estimation are open problems. It would be interesting to obtain necessary and sufficient conditions under which polynomial-time algorithms can be used to solve the fusion rule estimation problem under the criterion (1.2). Also, conditions under which the composite system is "significantly" better than best sensor would be extremely useful. Finally, lower bound estimates for various sample sizes will be very important in judging the optimality of sample size estimates.

ACKNOWLEDGEMENTS

This research is sponsored by the Engineering Research Program of the Office of Basic Energy Sciences, of the U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp., and by the Office of Naval Research under order N00014-96-F-0415.

REFERENCES

[1] N. S. V. Rao. Fusion methods for multiple sensor systems with unknown error densities. *Journal of Franklin Institute*, 331B(5):509–530, 1994.

[2] N. Alon, S. Ben-David, N. Cesa-Bianchi, and D. Hausler. Scale-sensitive dimensions, uniform convergence, and learnability. In *Proc. of 1993 IEEE Symp. on Foundations of Computer Science*, 1993.

[3] V. N. Vapnik. *The Nature of Statistical Learning Theory.* Springer-Verlag, New York, 1995.

[4] N. S. V. Rao. Fusion rule estimation in multiple sensor systems using training. In H. Bunke, T. Kanade, and H. Noltemeier, editors, *Modelling and Planning for Sensor Based Intelligent Robot Systems*, pages 179–190. World Scientific Pub., 1995.

[5] N. S. V. Rao. Fusion methods in multiple sensor systems using feedforward neural networks. *Intelligent Automation and Soft Computing*, 1996. submitted.

[6] N. S. V. Rao. Nadaraya-Watson estimator for sensor fusion. *Optical Engineering*, 36(3):642–647, 1997.

[7] N. S. V. Rao. Fusion rule estimation using vector space methods. In *Proceedings of SPIE Conference on Sensor Fusion: Architecture and Applications*. 1997.

[8] Z. Ciesielski. Haar system and nonparametric density estimation in several variables. *Probability and Mathematical Statistics*, 9:1–11, 1988.

[9] V. Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer-Verlag, New York, 1982.

[10] N. S. V. Rao and V. Protopopescu. Function estimation by feedforward sigmoidal networks with bounded weights. 1997. manuscript, submitted for publication.

[11] D. Pollard. *Empirical Processes: Theory and Applications*. Institute of Mathematical Statistics, Haywood, California, 1990.

[12] M. Talagrand. Sharper bounds for Gaussian and empirical processes. *Annals of Probability*, 22(1):28–76, 1994.

[13] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. *Information and Computation*, 100:78–150, 1992.

[14] M. J. Kearns and R. E. Schapire. Efficient distribution-free learning of probabilistic concepts. *Journal Computer and Systems Sciences*, 48(3):464–, 1994.

[15] R. Dudley. Universal Donsker classes and metric entropy. *Annals of Probability*, 15:1306–1326, 1987.

[16] B. L. S. Prakasa Rao. *Nonparametric Functional Estimation*. Academic Press, New York, 1983.

[17] J. Engel. A simple wavelet approach to nonparametric regression from recursive partitioning schemes. *Journal of Multivariate Analysis*, 49:242–254, 1994.

[18] N. S. V. Rao and V. Protopopescu. On PAC learning of functions with smoothness properties using feedforward sigmoidal networks. *Proceedings of the IEEE*, 84(10):1562–1569, 1996.

[19] S. A. Vavasis. *Nonlinear Optimization*. Oxford University Press, New York, 1991.

[20] M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer. *Extrapolative problems in automatic control and method of potential functions*, volume 87 of *American Mathematical Society Translations*, pages 281–303. 1970.

[21] V. Kurkova. Kolmogorov's theorem and multilayer neural networks. *Neural Networks*, 5:501–506, 1992.

AN ALGORITHM FOR NOISY IMAGE SEGMENTATION

(Extended Abstract)

Ying Xu, Victor Olman, and Edward C. Uberbacher
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

Abstract

This paper presents a segmentation algorithm for gray-level images and addresses issues related to its performance on noisy images. It formulates an image segmentation problem as a partition of an image into (arbitrarily-shaped) connected regions to minimize the sum of gray-level variations over all partitioned regions, under the constraints that (1) each partitioned region has at least a specified number of pixels, and (2) two adjacent regions have significantly different "average" gray-levels. To overcome the computational difficulty of directly solving this problem, a minimum spanning tree representation of a gray-level image has been developed. With this tree representation, an image segmentation problem is effectively reduced to a tree partitioning problem, which can be solved efficiently. To evaluate the algorithm, we have studied how noise affects the performance of the algorithm. Two types of noise, transmission noise and Gaussian additive noise, are considered, and their effects on both phases of the algorithm, construction of a tree representation and partition of a tree, are studied. Evaluation results have shown that the algorithm is stable and robust in the presence of these types of noise.

1 Introduction

Image segmentation is one of the most fundamental problems in low-level image processing. The problem is to partition (segment) an image into connected regions of similar textures or similar colors/gray-levels, with adjacent regions having significant dissimilarity. Many algorithms have been proposed to solve this problem (see surveys [1, 2]). Most of these algorithms fit into two categories: (1) boundary detection-based approaches, which partition an image by discovering closed boundary contours, and (2) region clustering-based approaches, which group "similar" neighboring pixels into clusters. Rigorous mathematical solutions to the image segmentation problems are generally difficult to achieve due to their (intrinsic) computational complexity. Hence many researchers have exploited either probabilistic/stochastic methods, which guarantee only asymptotic results, or heuristic methods while sacrificing the mathematical rigor.

In this paper, we present an efficient region-based segmentation algorithm. We formulate an image segmentation problem as a partition of an image into a number (not predetermined) of arbitrarily-shaped connected regions to minimize the sum of gray-level variations over all partitioned regions under the constraints that (1) each partitioned region has at least a specified number of pixels, and (2) two adjacent regions have significantly different "average" gray-levels. To overcome the computational difficulty of directly solving this problem, we have developed a minimum spanning tree representation of an image. The minimum spanning tree representation, though simple, captures the essential information of an image for the purpose of segmentation, and it facilitates a fast segmentation algorithm. The technical contribution of our approach includes (1) a new spanning tree representation of an image that captures all the key information for the purpose of segmentation, and (2) a fast and mathematically rigorous tree partitioning algorithm.

To evaluate the algorithm, we have studied how two types of noise, transmission noise and Gaussian additive noise, affect the performance of the algorithm. We have shown, both analytically and experimentally, that (1) both types of noise have very little effect on the minimum spanning tree construction algorithm, i.e., the property that an originally homogeneous region corresponds to one subtree of the spanning tree will generally not be affected by noise; (2) transmission noise, in general, has less effect on the performance of our tree-partitioning algorithm than Gaussian additive noise does.

2 Image Segmentation: the problem formulation

Consider a gray-level image I. Each pixel x of I has a gray level $\mathcal{G}(x) \in [0, \mathcal{K}]$. An image *segmentation problem* can be naturally formulated as follows: find a partition $\{I_1, ..., I_k\}$ of I with each I_i being a connected region of I, such that

$$\text{minimize} \qquad \textstyle\sum_{i=1}^{k} \sum_{x_i^j \in I_i} (average(I_i) - \mathcal{G}(x_i^j))^2$$

$$\begin{array}{lll} \text{subject to:} & (1) & \|I_i\| \geq L, \text{ for each } I_i, \\ & (2) & |average(I_i) - average(I_{i'})| \geq D, \text{ for all adjacent } I_i \text{ and } I_{i'}. \end{array}$$

where $\|\cdot\|$ denotes the cardinality of a set, $average(I_i)$ denotes the average gray-level of region I_i, and L and D are two (application-dependent) parameters.

Though this formulation captures the intuition of segmenting an image it is computationally difficult to solve due to two reasons: (1) segmenting a 2-D object to optimize some non-trivial function is always difficult, and (2) explicit calculation of averages implicitly requires to consider all the possible partitions. Two strategies have been developed to overcome these difficulties: a tree representation of an image, and an approximation scheme to avoid explicit calculation of averages.

2.1 Spanning tree representation of an image

For a given image I, we define a weighted (undirected) planar graph $G(I) = (V, E)$ as follows: The vertex set $V = \{$ all pixels of I $\}$ and the edge set $E = \{(u,v) | u, v \in V$ and $distance(u,v) \leq \sqrt{2}$ $\}$, with $distance(u,v)$ representing the Euclidean distance in terms of the coordinates of the image array; Each edge $(u,v) \in E$ has a weight $w(u,v) = |\mathcal{G}(u) - \mathcal{G}(v)|$.

A *spanning tree* T of a connected graph $G(I)$ (note that $G(I)$ is connected) is a connected subgraph of $G(I)$ such that (1) T contains every vertex of $G(I)$, and (2) T does not contain cycles. A *minimum* spanning tree is a spanning tree with a minimum total weight.

A minimum spanning tree of a weighted graph can be found using greedy methods, like in the classical Kruskal's algorithm [3]: the initial solution is a singleton set containing an edge with the smallest weight, and then the current partial solution is repeatedly expanded by adding the next smallest weighted edge (from the unconsidered edges) under the constraint that no cycles are formed until no more edges can be added. For the above defined planar graph $G(I)$, a minimum spanning tree can be constructed in $O(\|V\| \log(\|V\|))$ time and in $O(\|V\|)$ space.

A key property of a minimum spanning tree representation obtained by Kruskal's algorithm is that *pixels of a homogeneous region are connected in the tree structure only through pixels of this region*, i.e., pixels of a homogeneous region form a (connected) subtree of the minimum spanning tree. The following theorem formalizes this statement.

Consider an object A in a given image I. Let $G(I)$ be the planar graph representation of I and T be its minimum spanning tree obtained by Kruskal's algorithm. A is called *T-connected* if every pair of pixels of A are connected in T only through pixels of A. We use $G(A)$ to denote the subgraph of $G(I)$ induced by the pixels of A. A set of edges C of $G(A)$ is called a *cutset* if deleting C divides $G(A)$ into two unconnected parts.

Theorem 1 *A is not T-connected if and only if there exists a cutset C of $G(A)$ and a path P in $G(I)$ that has its two end vertices on two sides of the cut of $G(A)$ and has its remaining vertices outside of $G(A)$ such that every edge of P has smaller*[1] *weight than every edge of C.* □

2.2 An approximation scheme

To formulate the image segmentation problem in a natural and intuitive way, we have explicitly used the average gray-levels of a region in the problem formulation, which makes the computation difficult. This subsection presents an approximation scheme to avoid the explicit calculation of averages.

Consider the following formulation of an image segmentation problem. Given an image I and two parameters D and L, find a partition $\{I_1, ..., I_k\}$ of I with each I_i being a connected region of I, and a $g_i \in \mathcal{R}$ (real value) for each I_i, such that

$$\begin{array}{lll} \text{minimize} & & \sum_{i=1}^{k} \sum_{x_i^j \in I_i} (g_i - \mathcal{G}(x_i^j))^2 \\ & & \\ \text{subject to:} & (1) & \|I_i\| \geq L, \text{ for each } I_i, \\ & (2) & |g_i - g_{i'}| \geq D, \text{ for all adjacent } I_i \text{ and } I_{i'}. \end{array}$$

The relationship between this formulation, which does not involve explicit calculation of averages, and the original one can be intuitively described as follows: if a solution to this formulation is stable around the given parameter D, then the two formulations are equivalent. This can be stated more rigorously as in the following theorem. Let

$$F(k, I, g) = \sum_{i=1}^{k} \sum_{x_i^j \in I_i} (g_i - \mathcal{G}(x_i^j))^2,$$

and

$$R(D, L) = \{(k, I, g)| \text{ which satisfies constraints (1) and (2)}\},$$

[1] We ignore the case of equality for the simplicity of discussion.

where $I = \bigcup_{i=1}^{k} I_i$ and $g = (g_1, ..., g_k)$. Hence the above formulation can be rewritten as

$$\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(D,L)\}.$$

Theorem 2 *For the given parameters D and L, if there is an $\epsilon > 0$ such that*

$$\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(d,L)\} = F_0 \tag{1}$$

for some constant F_0, for all $d \in [D, D+\epsilon]$, then any minimum solution $I^* = \{I_1^*, ..., I_k^*\}$ and $g^* = \{g_1^*, ..., g_k^*\}$ to $\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(D+\epsilon, L)\}$ has $g_i^* = average(I_i^*)$, for all $i \in [1,k]$. □

Note that g_i's, as defined above, are real values $\in [0, \mathcal{K}]$. To facilitate a fast algorithm, we restrict g_i's to integer values $\in [0, \mathcal{K}]$. Now we can give the tree-based image segmentation problem as follows. Given a minimum spanning tree representation T of an image and two parameters D and L, find a partition $\{T_1, ..., T_k\}$ of T with each T_i being a connected subtree of T, and an integer $g_i \in [0, \mathcal{K}]$ for each T_i, such that

$$\begin{array}{lll} \text{minimize} & \sum_{i=1}^{k} \sum_{x_i^j \in T_i} (g_i - \mathcal{G}(x_i^j))^2 & \\ & & (P) \\ \text{subject to:} & (1) \quad \|T_i\| \geq L, \text{ for each } T_i, & \\ & (2) \quad |g_i - g_{i'}| \geq D, \text{ for all adjacent } T_i \text{ and } T_{i'}. & \end{array}$$

To estimate how close the approximation problem is to the original problem, we have the following result:

$$\frac{E(\sum_{i=1}^{k} \|T_i\| (average(T_i) - g_i)^2)}{E(\sum_{i=1}^{k} \sum_{x_i^j \in T_i} (average(x_i^j) - \mathcal{G}(x_i^j))^2)} \leq k/N, \tag{2}$$

which indicates the minimum value of the approximation problem is fairly close to the minimum value of the original optimization problem, where $E()$ represents the mathematical expectation.

3 A Tree-based Image Segmentation Algorithm

A dynamic programming algorithm is developed to solve the optimization problem (P). The algorithm first converts the given tree into a *rooted* tree by selecting an arbitrary vertex as the root. Hence the *parent-children* relation is defined. We assume that the vertices of T are labeled consecutively from 1 to $\|T\|$ with the tree root labeled as 1. We use T^i to denote the subtree rooted at vertex i. For each tree vertex i, the dynamic programming algorithm solves a number of constraint version of the problem (P) on T^i by combining solutions to the "corresponding" problems on T^j's, with j's being i's children. It does this repeatedly in such a bottom-up fashion and stops when it reaches the tree root.

Let $score(i,k,g)$ denote the minimum value of (P) on T^i, under the additional constraint that the partitioned subtree of T^i containing i has at least k vertices and is mapped to a fixed value g, for $k \in [0, L]$ and $g \in [0, \mathcal{K}]$. These quantities can be efficiently calculated using the following lemma and can be used to construct an optimum partition of T.

Lemma 1 *(a) If $i_1, i_2, ..., i_n$ are the children of vertex i, $n \leq 8$ and $1 \leq k \leq L$, we have*

$$\begin{array}{l} score(i,k,g) = \min \sum_{j=1}^{n} score(i_j, k_j, g) + (g - \mathcal{G}(i))^2, \\ \quad for\ k = \sum_{j=1}^{n} k_j, k_j \geq 0, \quad when\ \|T^i\| \geq L \\ scores(i,k,g) = \begin{cases} \sum_{p \in \mathcal{D}(i)} (g - \mathcal{G}(p))^2, & \|T^i\| = k \\ +\infty, & \|T^i\| \neq k \end{cases} \\ \quad when\ \|T^i\| < L \end{array}$$

where $\mathcal{D}(i)$ is the set of all i's descendants, i is defined to be $\in \mathcal{D}(i)$ and $score(i_j, 0, g)$ is defined to be

$$\min_{|g'-g| \geq D} score(i_j, L, g').$$

(b) $\min_g score(1, L, g)$ is a minimum solution of (P), where 1 represents the tree root. □

Based on Lemma 1, we can solve the optimization problem (P) by calculating *score*() for each tree vertex in a bottom-up fashion using the recurrence from Lemma 1(a), and stopping at the tree root.

Theorem 3 *$\min_g score(1, L, g)$ can be correctly calculated by the above algorithm in $O(\max\{(\|T\| - L), 1\}\mathcal{K}(\log(\mathcal{K}) + L^2))$ time and in $O(\|T\|L\mathcal{K})$ space.* □

4 Algorithm Evaluation on Noisy Images

Potentially noise affects the algorithm's performance in both stages of the algorithm: spanning tree construction and tree partitioning. We will show that noise has greater effects on the performance in the tree partitioning stage than in the spanning tree construction stage. In this study, we consider two types of noise: transmission noise and Gaussian additive noise.

The model for generating *transmission noise* is defined as follows: each pixel of the image has a probability $\mathcal{P}$ to keep its original gray level during transmission and the probability $1 - \mathcal{P}$ to randomly change to arbitrary gray level $\in [0, \mathcal{K}]$. *Gaussian additive noise* adds to each pixel independently a random normal value (using the floor function for real-to-integer conversion) according to a normal distribution $N(0, \sigma^2)$ censored to $[-\mathcal{K}/2, \mathcal{K}/2]$.

4.1 Effect of noise on tree representation

One basic premise for our image segmentation algorithm to be effective is that each object, given as a homogeneous region in an image, is represented as one subtree of the spanning tree representation. In the following, we show how noise affects this property. Theorem 1 provides the basic framework for such a study.

To estimate how probable the if-and-only-if condition in Theorem 1 is we have conducted the following computer simulation. The experiment is done on a 256-gray-level image I having one object A in the center of the image. I is a 256×256 image and A is a 30×30 square. The background has a uniform gray level 100 and A has a uniform gray level 150. We add transmission noise and Gaussian additive noise, respectively, to I as follows. When adding transmission noise, each pixel of I has a probability 0.3 to keep its original gray level and the probability 0.7 to randomly and uniformly change to arbitrary gray level $\in [0, 255]$. When adding Gaussian additive noise, each pixel of I is added by a value $\lfloor \delta + 1/2 \rfloor$ (modulo 256), where δ is random number generated according to the normal distribution $N(0, \sigma^2)$ censored to [-128, 128] with $\sigma = 50$.

For each type of noise, we estimated the probability that there exist a path P connecting two pixels a and b, and a cutset C of A separating a and b such that every edge of P has smaller weight than every edge of C, where a and b are two randomly chosen pixels both of which are 5-pixels from the left boundary of A and are at least 5 pixels from the lower and upper boundaries of A, and P has at least 20 edges.

We have observed, for this particular experiment, that the probability that there exist such a P and a cutset C is very small ($< 10^{-3}$), for both types of noise. This experiment suggested that both types of noise have very little effect on the property that a homogeneous region corresponds to one subtree of the minimum spanning tree constructed by Kruskal's algorithm.

4.2 Effect of noise on tree partitioning

Though both types of noise have little effect on the property that a homogeneous region corresponds to one subtree of the spanning tree representation they could affect the tree partitioning result in a form we call *corrosions.* Consider an object A in a given image and its representing subtree T_A. With noise, T_A may contain a subtree (or subtrees) that has a (significantly) different average gray level than the rest of T_A, and contains more than enough vertices ($\geq L$) to be partitioned into a separate region. This subsection presents a comparative study on how the two types of noise affect the formation of corrosions.

Let $g(A)$ be the (uniform) gray level of A before noise is added. We compare the probabilities, P_1 and P_2, that a connected region A' of A will have its gray level changed to the same value $g(A) + k$, for any $k \neq 0$, when transmission noise and Gaussian additive noise are added, respectively. Let p_k denote the probability that one pixel of A' changes its gray level from $g(A)$ to $g(A) + k$ when Gaussian additive noise is added. For the simplicity of discussion, we assume that $g(A) = 0$, hence $k \in [1, \mathcal{K}]$. Recall $\mathcal{P}$ denotes the probability that a pixel keeps its original gray level in the presence of transmission noise. It can be shown by a simple calculation that

$$P_1 = \left(\frac{1-\mathcal{P}}{\mathcal{K}-1}\right)^n (\mathcal{K}-1), \qquad \text{and} \qquad P_2 = \sum_{k=1}^{\mathcal{K}} p_k^n,$$

where $n = \|A'\|$ (note that $P_2 = \sum_{k=1}^{\mathcal{K}} p_k^n$ is true for any type of independent noise). Theorem 4 shows the relationship between P_1 and P_2, which can be proved using Jensen's inequality [5] (page 433).

Theorem 4 *For any $\mathcal{N} \in [2, \mathcal{K}]$ and $n > 0$, when $\sum_{k=0}^{\mathcal{N}} p_k = 1$ and $p_0 = \mathcal{P}$,*

$$\sum_{k=1}^{\mathcal{N}} p_k^n \geq \left(\frac{1-\mathcal{P}}{\mathcal{N}-1}\right)^n (\mathcal{N}-1).$$

□

Theorem 4 implies that transmission noise is the least possible to form corrosions among all possible forms of noises (including Gaussian additive noise) when $\mathcal{P}$ or p_0 is fixed.

4.3 Tests on noisy images

This subsection presents a case-study on an aerial image of 202x503 pixels and with 256 gray levels, and on how noise of different types affects the performance of the segmentation algorithm. Throughout this study, the same set of parameters D and L are used. Segmentation on each image takes less than 1 CPU minute on a SPARC-20 workstation. Figure 1 gives the test examples on the image with noise added. For each figure, the image on the left is the original image with added noise and the one on the right represents the segmentation results.

Table 1 summarizes the performance of algorithm and the effect of the averaging operation on the two types of noise. Each entry of the first row represents the correlational coefficient between the original image and the image with noise, and each entry of the second row represents the correlational coefficient between the segmentation result of the original image and the segmentation of the noisy image.

Table 1: Performance summary of segmentations

	Transmission noise				**Gaussian additive noise**			
	$\mathcal{P} = 0.1$	$\mathcal{P} = 0.3$	$\mathcal{P} = 0.5$	$\mathcal{P} = 0.7$	$\sigma = 40$	$\sigma = 60$	$\sigma = 80$	$\sigma = 100$
noisy image	0.86	0.62	0.41	0.24	0.82	0.69	0.57	0.47
segmentation	0.95	0.89	0.80	0.70	0.87	0.84	0.81	0.76

Acknowledgements

This research was supported by the United States Department of Energy, under contract DE-AC05-84OR21400 with Lockheed Martin Energy Systems, Inc. The authors would like to thank Dr. Reinhold C. Mann for many helpful discussions related to the work presented in this paper.

References

[1] N. Pal and S. Pal, "A review on image segmentation techniques", *Pattern Recognition*, Vol. 26, No. 9, pp. 1277 - 1294, 1993.

[2] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques", *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 100 - 132, 1982.

[3] J. B. Jr. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem", *Proc. Amer. Math Soc*, Vol. 7, No. 1, pp. 48 - 50, 1956.

[4] Y. Xu and E. C. Uberbacher, "2-D Image Segmentation Using Minimum Spanning Trees", *Image and Vision Computing*, Vol. 15 pp. 47 - 57, 1997.

[5] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day Inc., 1977.

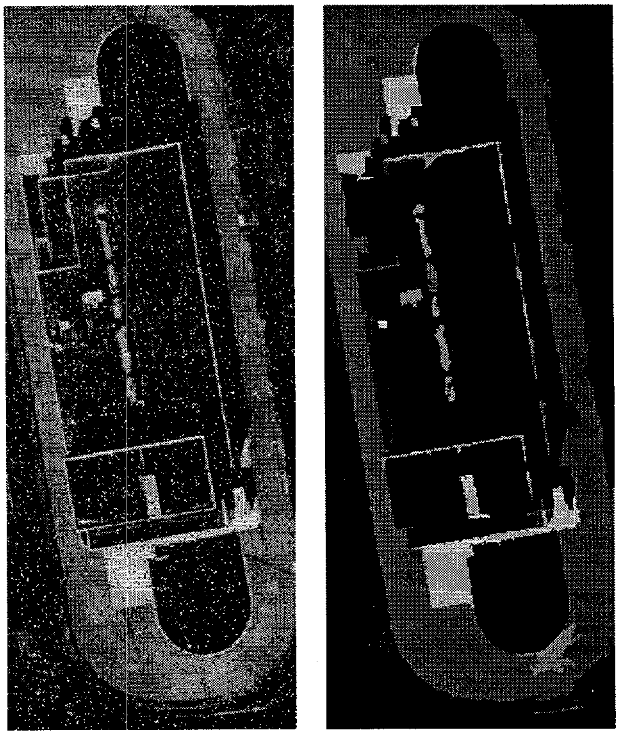


Figure 1: (a) Aerial image with added transmission noise and $\mathcal{P} = 0.1$.

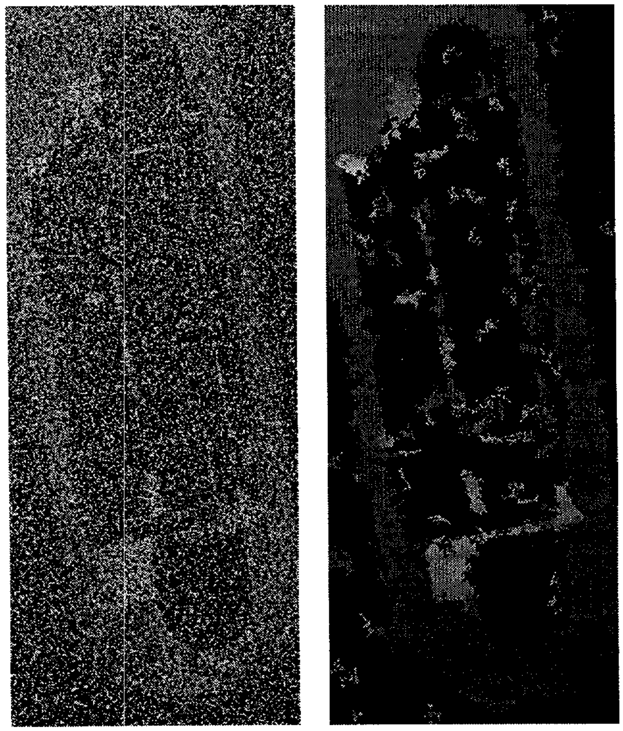


Figure 1: (b) Aerial image with added transmission noise and $\mathcal{P} = 0.7$.

ADAPTATION WITH DISTURBANCE ATTENUATION IN NONLINEAR CONTROL SYSTEMS

Tamer Başar

Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering
University of Illinois
Urbana, IL 61801-2307

ABSTRACT

We present an optimization-based adaptive controller design for nonlinear systems exhibiting parametric as well as functional uncertainty. The approach involves the formulation of an appropriate cost functional that places positive weight on deviations from the achievement of desired objectives (such as tracking of a reference trajectory while the system exhibits good transient performance) and negative weight on the energy of the uncertainty. This cost functional also translates into a disturbance attenuation inequality which quantifies the effect of the presence of uncertainty on the desired objective, which in turn yields an interpretation for the optimizing control as one that optimally attenuates the disturbance, viewed as the collection of unknown parameters and unknown signals entering the system dynamics. In addition to this *disturbance attenuation* property, the controllers obtained also feature *adaptation* in the sense that they help with identification of the unknown parameters, even though this has not been set as the primary goal of the design. In spite of this adaptation/identification role, the controllers obtained are not of certainty-equivalent type, which means that the identification and the control phases of the design are not decoupled.

INTRODUCTION AND PROBLEM DESCRIPTION

We consider in this paper the problem of control of partially unknown, uncertain nonlinear systems so that the system output tracks (at least asymptotically) a given reference trajectory while all internal states remain bounded and the system exhibits acceptable transient performance. The uncertainty is due to the presence of unknown deterministic signals entering the system dynamics, and unknown (and unmeasurable) noise in the measurements. To capture the presence of all these factors that impact the overall performance of the system, and to quantify various tradeoffs that exist, we base our control design on the minimization of a carefully selected cost functional, which leads to a systematic construction of *robust adaptive controllers* that attenuate the disturbances optimally. These adaptive controllers have distinguishable identifier and control dynamics, which however are not decoupled, and hence the controllers are not certainty equivalent — in contrast to many existing designs in the literature. This "noncertainty equivalence" structure, which comes about naturally as a result of the optimization procedure, brings with it many appealing features such as robustness to unmodeled dynamics, attenuation of disturbances, and excellent transient performance.

To introduce the approach adopted in this paper in general terms, consider the n-dimensional dynamic system described by

$$\dot{x} \quad = \quad f(x,\theta) + G(x,\theta)u + \sigma(x)w, \quad x(0) = x_0 \tag{1}$$

where θ is a p-dimensional unknown parameter vector, x is the n-dimensional state, u is an r-dimensional control, w is a q-dimensional unknown disturbance, f, G and σ are appropriate dimensional vectors and matrices, continuous in x, and with f and G linear in θ. Let us assume for the moment that the control uses state feedback with memory, that is for some measurable function μ,

$$u(t) \quad = \quad \mu(t, x_{[0,t]})\,, \tag{2}$$

and that the objective is for an m-dimensional ($m \leq n$) output of the system,

$$z \quad = \quad h(x), \tag{3}$$

to track a given m-dimensional reference trajectory, z_r, in spite of the presence of the disturbance w, and regardless of what the true value of θ is. Hence, what is being sought is a controller that achieves the desired objective (of tracking) while attenuating the disturbances at the output of the error systems, which is the tracking error, and at the same time keeping all internal states of the system bounded. A criterion that captures this objective is now given in the following.

Let us first introduce the notation

$$\|y\|_t^2 \quad := \quad \int_0^t |y(\tau)|^2 d\tau, \qquad |y(\tau)|_Q^2 \quad := \quad y'(\tau)Qy(\tau)\,,$$

for any vector-valued $\mathcal{L}_2$ function y, where $'$ stands for transpose of a vector (or a matrix), the latter is the square of a weighted Euclidean norm of $y(\tau)$, where Q is a positive definite weighting matrix, and the former is the square of the $\mathcal{L}_2$ norm of the function $y(\tau)$ restricted to interval $[0,t]$. Then, consider for each $t > 0$,

$$\mathcal{I}_t(\mu) \quad = \quad \sup_{w,\theta,x_0} \frac{\|z - z_r\|_t^2 + \tilde{\ell}_t(x_{[0,t]}; u_{[0,t]})}{\|w\|_t^2 + |\theta - \bar{\theta}|_{Q_0}^2 + \ell_0(x_0, \theta - \bar{\theta})} \tag{4}$$

as the performance index to be minimized by the controller μ for each $t > 0$. Here

$$\tilde{\ell}_t \quad := \quad \int_0^t \ell(x_{[0,\tau]}; u(\tau), \tau) d\tau$$

is a nonnegative integral cost on the state and the control, $z - z_r$ is the tracking error, $\bar{\theta}$ is an initial estimate for θ, Q_0 is a positive definite matrix, and ℓ_0 is a nonnegative cost on x_0 and $\theta - \bar{\theta}$, vanishing at $x_0 = 0$ and $\theta = \bar{\theta}$. Note that $\mathcal{I}_t$ involves a maximization operation with respect to the unknowns, w, θ and x_0, and hence characterizes a worst case scenario. By minimizing this index with respect to μ we would be minimizing the worst-case effect of w, θ and ℓ_0 on the tracking error $z - z_r$, the state x and the control u. Now let

$$\inf_{\mu} \mathcal{I}_t(\mu) \quad =: \quad \gamma_t^{*^2}, \quad t > 0\,,$$

and pick $\gamma > 0$ such that $\gamma > \gamma_t^*$ for all $t > 0$. Let μ_γ be a controller that achieves a better (lower) level of disturbance attenuation than γ for all $t > 0$, that is

$$\mathcal{I}_t(\mu_\gamma) \quad \leq \quad \gamma^2. \tag{5}$$

Then, (4) implies that for all $w \in \mathcal{L}_2[0,\infty)$, $\theta \in \mathbf{R}^p$, $x_0 \in \mathbf{R}^n$, the following dissipation inequality holds, for all $t > 0$, with $u = \mu(\cdot)$:

$$J_\gamma^t(\mu;\omega) := \|z - z_r\|_t^2 + \tilde{\ell}_t(x_{[0,t]}; u_{[0,t]}) - \gamma^2\|w\|_t^2 - \gamma^2\,|\theta - \bar{\theta}|_{Q_0}^2 - \gamma^2\,\ell_0(x_0, \theta - \bar{\theta}) \quad \leq \quad 0. \tag{6}$$

Denote the left hand side of this inequality for an arbitrary μ and $\omega := (w, \theta, x_0)$ by $\mathcal{I}^t_\gamma(\mu;\omega)$. Then, clearly, a μ_γ satisfying (5) can be obtained by solving the *minmax* problem:

$$\inf_{\mu} \sup_{\omega} J^t_\gamma(\mu;\omega), \qquad t > 0. \tag{7}$$

This can be viewed as a zero-sum differential game between two players [2], with the minimizer choosing μ and the maximizer ω, and the quantity in (7) is the upper value of such a game. Note that whenever this value is bounded, it has to be zero, since by picking $w = 0$, $\theta = \bar{\theta}$ and $x_0 = 0$ the maximizer can force it to be nonnegative, and on the other hand we know from inequality (5) that it cannot be positive.

Our approach to this problem is based on the recognition that the supremization part of (7) can be broken into two sequential supremizations,

$$\sup_{\omega} J^t_\gamma(\mu;\omega) \quad = \quad \sup_{\theta, x_{[0,t]}} \quad \sup_{(w_{[0,\infty)}|x_{[0,t]},\theta)} \quad J^t_\gamma(\mu;\omega) \tag{8}$$

where the inner supremization is over all disturbance trajectories consistent with the observed state trajectory $x_{[0,t]}$ and for a fixed value of the parameter vector θ, and the outer supremization is over all possible values of $\theta \in \mathbf{R}^p$ and all continuous state trajectories $x_{[0,t]}$. If the controller did not have access to full state measurements, but only to partial measurements, possibly corrupted with (unknown) noise, such as

$$y(t) \quad = \quad h(x,\theta) + n(x)w$$

where h and n are continuous functions of their arguments, then (8) would be replaced by the more general relationship

$$\sup_{\omega} J^t_\gamma(\mu;\omega) \quad = \quad \sup_{\theta, x_t, y_{[0,t]}} \quad \sup_{(x_0, w_{[0,\infty)}|y_{[0,t]},\theta,x(t)=x_t)} \quad J^t_\gamma(\mu;\omega) \tag{9}$$

with the inner and outer supremizations interpreted in a similar way (as in (8)). Now, this splitting of J^t_γ into two parts leads to a sequential design procedure that generates worst-case identifiers and robust adaptive controllers. The inner supremization (maximization) is the *worst-case identification step* which can be solved using the recently developed tool of *cost-to-come function* [4,5] which leads for some important classes of problems (as to be discussed shortly) to closed-form expressions for an identifier for the unknown parameters and an estimator for the unmeasured states. During this identification step the control $u_{[0,t]}$, generated by μ, can be regarded as an open-loop time function since it is merely a causal function of the given output waveform $y_{[0,t]}$ (or of the state $x_{[0,t]}$, if state measurements are available). After thus completing the inner supremization, we then proceed with the outer supremization of J^t_γ over all measurement waveforms $y_{[0,t]}$, terminal states x_t and parameter values θ, while structuring the control in such a way that J^t_γ remains nonpositive. This is the *control design step* which leads to a robust disturbance attenuating controller.

The problem just formulated above can also be viewed as a nonlinear H^∞ control problem with partial state information [1], by adjoining to the system dynamics (1) the natural parameter dynamics

$$\dot{\theta} = 0, \qquad \theta(0) = \theta_0 \tag{10}$$

where now θ_0 is the unknown parameter vector. This H^∞ control problem is one with partial state information even if full state measurements are available, because θ is now considered a part of the extended system dynamics, which is not measured directly. Nonlinear H^∞ control problems with partial state information are known to be inherently difficult to solve, and generally they do not admit finite-dimensional solutions [1,3]. It turns out, however, that (as shown in our recent research [9,10]) for some special subclasses of the robust adaptive control problems formulated above, finite-dimensional closed-form solutions do exist; this will be discussed also in the following sections.

Our approach to robust adaptive control as delineated above is inherently different from other existing approaches which are either Lyapunov-based or estimation-based. The former places restrictions on the selection of parameter update laws, whereas the latter (which generally makes use of the "certainty equivalent" principle) uses a wide variety of estimation/identification tools, among which are the standard gradient and least-squares algorithms. Any stabilizing controller can in fact be combined with any such identifier, as long as the identifier guarantees certain boundedness properties independently of the controller module. This modularity feature has made estimation-based schemes more popular in linear adaptive control than their Lyapunov-based counterparts, but efforts to extend this to nonlinear systems have failed to a large extent. The source of this failure is mainly the fact that nonlinear systems exhibit different instability characteristics (than linear systems) such as finite escape. Various measures have been taken to overcome this difficulty [6,7,8], but the designs have involved certainty-equivalent controllers, which are known to have weaknesses in the framework of nonlinear systems, in particular as regards robustness against model uncertainty and external disturbance inputs.

The approach presented above, and discussed in some detail (as length restrictions permit) in the following sections, is a direct optimization-based approach that brings in robustness as an essential component of the design procedure. To carry out the details of the two-step procedure outlined above, and to obtain explicit expressions for the optimally disturbance attenuating controllers, we focus on a special, but important, class of systems where the dynamics (1) are in triangular form, the control is of dimension *one*, and the system output is the first component of the state. Such systems are called "systems in parametric strict feedback form," and one of their appealing features is that a recursive technique called *backstepping* can be used to construct an optimizing controller. Formulation of this specific problem is provided in the next section, followed by presentation of some explicit results.

SYSTEMS IN PARAMETRIC-STRICT-FEEDBACK FORM

In view of the discussion above, consider now the class of single input-single output (SISO) nonlinear systems described by (as a special case of (1), and by a possible abuse of notation):

$$\begin{aligned}
\dot{x}_1 &= x_2 + f_1(x_1) + \phi_1'(x_1)\theta_1 + \sigma_1'(x_1)w_1 \\
\vdots\ &\ \vdots \\
\dot{x}_{n-1} &= x_n + f_{n-1}(x_1,\ldots,x_{n-1}) + \phi_{n-1}'(x_1,\ldots,x_{n-1})\theta_{n-1} \\
&\quad +\sigma_{n-1}'(x_1,\ldots,x_{n-1})w_{n-1} \\
\dot{x}_n &= f_n(x_1,\ldots,x_n) + \phi_n'(x_1,\ldots,x_n)\phi_n + b(x_1,\ldots,x_n)u + \sigma_n'(x_1,\ldots,x_n)w_n \\
z &= x_1.
\end{aligned} \tag{11}$$

where $w := (w_1',\ldots,w_n')'$ is the q-dimensional disturbance input, where w_i is of dimension q_i, $i = 1,\ldots,n$; $\theta = (\theta_1',\ldots,\theta_n')$ is the p-dimensional vector of unknown parameters, where θ_i is of dimension p_i, $i = 1,\ldots,n$; z is the scalar output; and the nonlinear functions f_i, ϕ_i, σ_i, $i = 1,\ldots,n$, are known and satisfy the triangular structure depicted above. We assume that

A1 f_i, ϕ_i, $\sigma_i \in \mathcal{C}^{n-i+1}$, $i = 1,\ldots,n$; $b \in \mathcal{C}^1$, $(1/b) \in \mathcal{C}^1$.

A2 $\sigma_i'(x)\sigma_i(x) > c$, $\forall x \in \mathbf{R}^n$, $i = 1,\ldots,n$, for some $c > 0$.

A3 The reference trajectory $z_r \in \mathcal{C}^n$, and both z_r and its first n derivatives are uniformly bounded on $[0,\infty)$.

Let us first endow the controller (2) with also the derivative of the state, $\dot{x}$, under which the inner maximization of (8) can be performed to yield [5] the identifier dynamics (for θ):

$$\begin{aligned}
\dot{\hat{\theta}}_i &= \Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1}(\dot{x}_i - \mathcal{X}_i - \phi_i'\hat{\theta}_i); \qquad \hat{\theta}_i(0) = \bar{\theta}_i & (12)\\
\dot{\Sigma}_i &= -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - Q_i)\Sigma_i; \qquad \Sigma_i(0) = Q_{0i}^{-1} & (13)\\
i &= 1,\ldots,n
\end{aligned}$$

where

$$\mathcal{X}_i := f_i + x_{i+1}, \qquad i = 1,\ldots,n-1; \qquad \mathcal{X}_n := f_n + bu \tag{14}$$

and Q_i, Q_{0i} are positive-definite matrices constituting the i-th diagonal blocks of Q and Q_{oi}, respectively. The identifier $\hat{\theta}_i$, $i = 1,\ldots,n$, above is asymptotically convergent to the true value of the parameter vector θ

provided that $\Sigma_i(t) > 0 \; \forall t \in [0, \infty)$, $i = 1, \ldots, n$, which is a *persistency of excitation* condition, that can be guaranteed by restricting the disturbance inputs to a particular set [5]. This set can be made unconstrained, and equal to $\mathcal{L}_\infty$, by choosing the design matrix Q_i as

$$Q_i \;\; = \;\; c_i \phi_i (\sigma_i' \sigma_i)^{-1} \phi_i', \qquad c_i \in [0, 1), \quad i = 1, \ldots, n. \tag{15}$$

The identification error $\tilde{\theta} := \theta - \hat{\theta}$ satisfies:

$$\dot{\tilde{\theta}}_i \;\; = \;\; -\Sigma_i \phi_i (\sigma_i' \sigma_i)^{-1} \sigma_i' v_i, \qquad i = 1, \ldots, n \tag{16}$$

where

$$v_i \;\; := \;\; w_i + \sigma_i (\sigma_i' \sigma_i)^{-1} \phi_i' \tilde{\theta}_i, \quad i = 1, \ldots, n \tag{17}$$

is a transformed disturbance input. This converts the original attenuation problem with respect to w to a new (but equivalent) attenuation problem with respect to v, with dynamics described by (in place of (11)):

$$\dot{x}_i \;\; = \;\; \mathcal{X}_i + \phi_i' \hat{\theta}_i + \sigma_i' v_i, \quad i = 1, \ldots, n \tag{18}$$

$$\dot{\hat{\theta}}_i \;\; = \;\; \Sigma_i \phi_i (\sigma_i' \sigma_i)^{-1} \sigma_i' v_i, \quad i = 1, \ldots, n \tag{19}$$

along with (13). Note that there is no parametric uncertainty here, and hence the problem has been converted to one with perfect state measurements, where x_i, $\hat{\theta}_i$ and Σ_i, $i = 1, \ldots, n$, constitute the new states. This new nonlinear H^∞ control problem (with perfect state measurements) corresponds to the outer maximization problem in (8), where the cost to be maximized is

$$\int_0^t \left((z - z_r)^2 + \ell(\tau, x_{[0,\tau]}) - \gamma^2 \sum_{i=1}^n \mid \tilde{\theta}_i \mid_{Q_i}^2 - \gamma^2 |v|^2 \right) d\tau - \gamma^2 \sum_{i=1}^n \mid \tilde{\theta}_i(t) \mid_{\Sigma_i^{-1}(t)}^2 - \ell_0(x(0), \theta - \bar{\theta}) \tag{20}$$

where we have dropped the control dependence in ℓ, and have absorbed γ^2 in ℓ_0.

This is now the control design step, which we carry out under assumptions **A1-A3**. The design procedure here is *backstepping*, which proceeds as follows: We first consider the first subsystem ($i = 1$), and treat x_2 as an input to this system. Introducing the transformed variable $y_1 := x_1 - z_r$, one can show that this decoupled scalar system with x_2 as an input can be made to achieve arbitrarily small levels of disturbance attenuation by picking x_2 appropriately. A corresponding value function for (20) for only this subsystem is $\bar{V}_1(y_1) = \frac{1}{2} y_1^2$. However, since x_2 is not a control input this is not exact, and hence we proceed to the next subsystem ($i = 2$) and choose x_3 as the new control input, where the dynamics for x_2 are now replaced by the dynamics of y_2, which stands for the difference between x_2 and its ideal value, had it been the control variable at step 1. At this step, again there exist choices for x_3 that make the attenuation level arbitrarily close to zero, with a corresponding value function being $\bar{V}_2 = \frac{1}{2}(y_1^2 + y_2^2)$. Again x_3 is not the true control and hence this result is not exact .... Proceeding in this manner, we arrive at step n at the last subsystem where the real control appears, a proper choice for which makes the overall attenuation level again arbitrarily small. Because of space limitations, expressions for the construction of this controller (which are quite lengthy) are not given here; they can be found in an internal report available from the author. These steps now lead us to the following theorem.

Theorem 1. *Consider the nonlinear system described by (11) and with the performance index (4) where* $\tilde{\ell}_t$ *does not depend on* u. *Let assumptions* **A1-A3** *hold, and disturbances belong to a set (say* $\mathcal{W}$*) that makes* $\Sigma_i(t)$ *positive definite for all* $t \geq 0$, $i = 1, \ldots, n$. *Finally, let the derivative of* x *also be available for feedback. Then:*

(i) The control law generated by the backstepping procedure outlined above achieves asymptotic tracking with an arbitrarily small level of disturbance attenuation, γ, *for all* $w \in \mathcal{W}$.

(ii) For any $w_{[0,\infty)} \in \mathcal{L}_\infty$, $x(0)$, θ *and* $t \geq 0$, *if the covariance matrices* Σ_i*'s are uniformly upper bounded on* $[0, t]$, *then the expanded state vector* ζ *(consisting of* x, $\hat{\theta}$ *and* Σ*) is uniformly bounded on* $[0, t]$, *and* Σ_i*'s are further uniformly bounded from below by some positive-definite matrices.*

(iii) For any uncertainty triple in the set $\mathcal{W}$ such that $w_{[0,\infty)} \equiv 0$, if the covariance matrices Σ_i, $i = 1, \ldots, n$, become uniformly upper bounded on $[0, \infty)$, then the parameter estimates are uniformly bounded and the transformed state variable $y := (y_1, \ldots, y_n)'$ converges to zero as $t \to \infty$; if, in addition, the reference signal z_r is persistently exciting, i.e., $\lim_{t\to\infty} \lambda_{\max}\Sigma_i = 0$, $i = 1, \ldots, n$, then ζ converges to zero as $t \to \infty$.

Remark 1. The controller above is not a certainty equivalent controller, that is it does not correspond to the controller obtained by assuming full knowledge of parameter values and then replacing the true values of the parameters by their estimates. It is, however, asymptotically certain equivalent, as $\Sigma_i \to 0$, $i = 1, \ldots, n$, under a specified choice of the design parameters. ◇

A disadvantage of the controller presented above is that it depends (through the identifier dynamics) on the derivative of the state, which may not be available. To remove this dependence, so as to obtain a controller under the original measurement scheme (2), we first consider a noise-perturbed measurement:

$$u(t) = \mu(t, y(t)), \qquad y(t) = x(t) + \epsilon v(t),$$

where ϵ is a small positive parameter and v is an unknown disturbance. The identifier dynamics corresponding to this measurement (as the counterpart of (12)-(13)), and after v is set equal to *zero* are [5]:

$$\begin{aligned}
\dot{\hat{\theta}}_i &= \Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1/2}\frac{1}{\epsilon}(x_i - \hat{x}_i); \qquad \hat{\theta}_i(0) = \bar{\theta}_i \quad . && (21)\\
\dot{\Sigma}_i &= -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - Q_i)\Sigma_i; \qquad \Sigma_i(0) = Q_{0i}^{-1} && (22)\\
\dot{\hat{x}} &= \chi_i + \phi_i'\hat{\theta}_i + \frac{1}{\epsilon}(\sigma_i'\sigma_i)^{1/2}(x_i - \hat{x}_i) : \qquad \hat{x}_i(0) = x_i(0) && (23)\\
i &= 1, \ldots, n
\end{aligned}$$

where we now have additional dynamics representing the estimate for x. An appropriate choice for the design matrix Q_i in this case turns out to be

$$Q_i = \Sigma_i^{-1}\Delta_i\Sigma_i^{-1} + \check{Q}_i; \quad \Delta_i > \kappa_Q I_{p_i} \tag{24}$$

for some symmetric matrices Δ_i and $\check{Q}_i$. Now, introducing

$$e := (x - \hat{x})/\epsilon$$

whose i-th component is $e_i = (x_i - \hat{x}_i)/\epsilon$, we can equivalently write (21)-(22) as (using also the specific choice made for Q_i):

$$\begin{aligned}
\dot{\hat{\theta}}_i &= \Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1/2}e_i && (25)\\
\dot{\Sigma}_i &= -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - \check{Q}_i)\Sigma_i + \Delta_i && (26)\\
\epsilon\dot{e}_i &= -(\sigma_i'\sigma_i)^{1/2}e_i + \phi_i'\tilde{\theta}_i + \sigma_i'\omega_i, && (27)
\end{aligned}$$

which involves singularly perturbed dynamics. It should be noted that formally setting $\epsilon = 0$ in (27) and substituting the resulting expression for e_i into (25) yields precisely the identifier dynamics (19). Using this limiting relationship (which can be made precise using singular perturbations analysis), and the same backstepping design tool as in the earlier case, we obtain a robust disturbance attenuating controller in exactly the same form as in Theorem 1 but with the identifier now generated by (25)-(27). For a precise statement of this result, which would be the counterpart of Theorem 1 here, we first introduce a class of admissible uncertainties, $\mathcal{W}_C$, as the counterpart of $\mathcal{W}$ introduced in Theorem 1. For some arbitrary positive constant C, let

$$\mathcal{W}_C := \left\{(x(0), \theta, w_{[0,\infty)}) : \lambda_{\max}\Sigma_i(t) \le C, |x(0)| \le C, |\theta| \le C, |w(t)| \le C, \forall t \in [0, \infty), \forall i = 1, \ldots, n\right\} \tag{28}$$

Theorem 2. *Consider the nonlinear system described by (11) with perfect state (but not derivative) information, and with performance index (4) where $\tilde{\ell}_t$ does not depend on u and is positive definite. Let assumptions* **A1-A3** *hold, Q_i be given by (24), and $\mathcal{W}_C$ be as defined by (28). Then:*

(i) There exists a positive scalar $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0]$, the control law of Theorem 1, with identifier (25)-(27), achieves asymptotic tracking with disturbance attenuation level γ for any uncertainty triple in the set $\mathcal{W}_C$. Furthermore, the closed-loop signals generated by the overall system are uniformly bounded on $[0, \infty)$.

(ii) For any uncertainty triple in the set $\mathcal{W}_C$ such that $w_{[0,\infty)} \equiv 0$, the expanded state vector, including the system state, and both slow and fast parameter errors, converges to 0 as $t \to \infty$ for any $\epsilon \in (0, \epsilon_0]$.

Remark 2. The passage from Theorem 1 to Theorem 2 has involved (in order to avoid the use of derivative information, and singularity in the optimization problem under pure state measurements) the introduction of small noise in the measurement equation, obtaining a controller along with a worst-case identifier under this noise-perturbed measurement, and then setting the disturbance (noise) entering the measurement equation to *zero*. The resulting identifier dynamics still depend on the small parameter ϵ multiplying the measurement disturbance even after the disturbance has been eliminated. This way, any performance achieved under derivative information can be achieved by using only state information. We should also note that in this case, due to the requirement that the error covariance matrices be bounded away from zero, the robust adaptive controller will not be certainty equivalent, even asymptotically. $\diamond$

THE CASE OF OUTPUT MEASUREMENTS

Let us now turn to the case of output measurements, that is the case when not all state variables but only a subset of them is available for control purposes. In particular, let us consider in the context of the parametric strict feedback form (11) only the output, z, to be available. As in the previous section, let us first assume that the derivative of z is also measurable and is available for control purposes. Then, the first subsystem of (11) serves as the measurement equation:

$$\dot{z} - f_1(z) \quad = \quad x_2 + \phi_1'(z)\theta_1 + \sigma_1'(z)w_1, \tag{29}$$

through which noisy information is available on x_2 and θ_1 — with the noise being due to the presence of the disturbance w_1. Denote the remaining components of x by x_F, and let ξ denote the extended state $(\theta', x_F')'$, which satisfies an equation of the form

$$\dot{\xi} \quad = \quad A\xi + f + Hw, \quad \xi(0) = (\theta', x_F(0)')' \tag{30}$$

with obvious definitions for A, f and H. It should be noted that A depends on u (linearly) , and the dependence of f on x_F is in a lower triangular form. In terms of this notation, (29) can be rewritten as

$$\dot{z} - f_1(z) \quad = \quad C'(z)\xi + \sigma_i'(z)w \tag{31}$$

where again the definition of C should be obvious.

Now, with (30) serving as a state equation and (31) as the measurement equation, the inner supremization of (9) becomes an H^∞ filtering problem which can be solved using the theory of [1, chapter 7], leading to the following optimal (worst-case) observer and error covariance equations:

$$\dot{\hat{\xi}} \quad = \quad A\xi + f + (\gamma^2\Sigma C + L)N(\dot{z} - f_1 - C'\hat{\xi}), \quad \hat{\xi}(0) = \begin{pmatrix} \hat{\theta}_0 \\ \hat{x}_{F_0} \end{pmatrix} \tag{32}$$

$$\begin{aligned} \dot{\bar{\Sigma}} \quad &= \quad (A - LNC')\bar{\Sigma} + \bar{\Sigma}(A - LNC')' - \bar{\Sigma}(\gamma^2 CNC' - Q)\bar{\Sigma} + \gamma^{-2}(HH' - LNL'), \\ &\qquad \bar{\Sigma}(0) = \gamma^{-2}\text{blockdiag}(\Sigma_0 \;\; \pi_0) \end{aligned} \tag{33}$$

where $\hat{\theta}_0$ and $\hat{x}_{F_0}$ denote the initial (a *priori*) estimates for θ and $x_F(0)$, respectively, $L := H\sigma_1$, $N := (\sigma_1'\sigma_1)^{-1}$, and Q is a nonnegative-definite matrix, serving as a Euclidean weighting on the estimation error

$\xi - \hat{\xi}$, which is a part of $\tilde{\ell}_t$ in (4) (or equivalently (6)). To ensure boundedness of parameter error, it is generally useful to add to the right hand side of (32) a smooth function that forces the parameter estimate (the first p components of $\hat{\xi}$) to stay within an a priori known set θ_0 (where all the parameters lie); see [10] for details. This then completes the design of the identifier/estimator, and brings us to the control design stage (i.e., the outer maximization in (9)). The combined state, estimator and error covariance dynamics are again in strict feedback form, which makes it possible to apply the backstepping tool of the previous sections; the details are lengthy and have not been included here due to page limitations.

The procedure outlined above leads to a controller that depends not only on z but also on $\dot{z}$. To remove the dependence on $\dot{z}$ we again follow a procedure similar to that carried out in the previous section, to go from derivative measurements to the state measurement case. We introduce a new measurement, y, which is a noise-perturbed version of z: $y = z + \epsilon v$, where v is a scalar unknown disturbance, and ϵ is a small positive parameter. The inner maximization problem of (9) can be solved as in the derivative measurement case, to which we subsequently apply singular perturbations analysis to obtain estimators that are well-defined when $v \equiv 0$ and ϵ is small. Then, the solution of the outer maximization problem again involves backstepping, leading to a robust adaptive controller which uses only the given scalar output measurement. Under some technical conditions, one can then prove a result similar to Theorem 2, assuring asymptotic tracking property of the derived controller for sufficiently small $\epsilon > 0$, with an arbitrary level of disturbance attenuation, and for all uncertainty belonging to a given set. Furthermore, all closed-loop signals remain bounded for all time.

ACKNOWLEDGEMENT

This work was performed under the auspices of the U.S. Department of Energy. Contributions of Z. Pan and I.E. Tezcan to this effort are also acknowledged.

REFERENCES

1. T. BAŞAR and P. BERNHARD, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkhäuser, Boston, MA, 2nd edition (1995).

2. T. BAŞAR and G. J. OLSDER, *Dynamic Noncooperative Game Theory*, Academic Press, London, 2nd edition (1995).

3. J. A. BALL, J. W. HELTON, and M. WALKER, "H^∞ control for nonlinear systems with output feedback," *IEEE Trans. Automatic Control*, 38(4):546-559 (1993).

4. G. DIDINSKY and T. BAŞAR, "Minimax adaptive control of uncertain plants," *Proceedings of the 33rd IEEE Conference on Decision and Control*, 2839-2844, Orlando, FL (1994).

5. G. DIDINSKY, Z. PAN, and T. BAŞAR, "Parameter identification for uncertain plants using H^∞ methods," *Automatica*, 31(9):1227-1250 (1995).

6. I. KANELLAKOPOULOS, P. V. KOKOTOVIĆ, and A. S. MORSE, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Trans. Automatic Control*, 36:1241-1253 (1991).

7. M. KRSTIĆ, I. KANELLAKOPOULOS, and P. V. KOKOTOVIĆ, *Nonlinear and Adaptive Controllers Design*, Wiley, NY (1995).

8. M. KRSTIĆ and P. V. KOKOTOVIĆ, "Adaptive nonlinear design with controller-identifier separation and swapping, *IEEE Trans. Automatic Control*, 40(3):426-440 (1995).

9. Z. PAN and T. BAŞAR, "Adaptive controller design for tracking and disturbance attenuation in parametric-strict-feedback nonlinear systems," *Proc. 13th IFAC World Congress*, San Francisco, CA, F:323-328 (1996).

10. I. E. TEZCAN and T. BAŞAR, "Disturbance attenuating adaptive controllers for parametric strict feedback nonlinear systems with output measurements," *Proc. 1997 American Control Conference*, Albuquerque, NM (1997).

USE OF LASER DIODES IN CAVITY RING-DOWN SPECTROSCOPY

R. N. Zare, B. A. Paldus, Y. Ma, and J. Xie

Department of Chemistry, Stanford University
Stanford, CA 94305-5080, USA

ABSTRACT

We have demonstrated that cavity ring-down spectroscopy (CRDS), a highly sensitive absorption technique, is versatile enough to serve as a complete diagnostic for materials process control. In particular, we have used CRDS in the ultraviolet to determine the concentration profile of methyl radicals in a hot-filament diamond reactor; we have applied CRDS in the mid-infrared to detect 50 ppb of methane in a N_2 environment; and, we have extended CRDS so that we can use continuous-wave diode laser sources. Using a laser diode at 810 nm, we were able to achieve a sensitivity of 2×10^{-8} cm^{-1}. Thus, CRDS can be used not only as an *in situ* diagnostic for investigating the chemistry of diamond film deposition, but it can also be used as a gas purity diagnostic for any chemical vapor deposition system.

INTRODUCTION

Present-day technology is dominated by the synthesis of materials, ranging from biocompatible plastics, to metal-semiconductor heterostructures for lasers used in telecommunications, to silicon oxides and nitrides that provide the backbone of the electronics industry. Materials process control is rapidly becoming more important in industry, and is triggering fundamental research of materials and their chemistries.

Diamond films, because of their mechanical hardness, high thermal conductivity, and excellent optical properties are commercially important in a wide set of applications, ranging from the more traditional tool coating to integrated circuit fabrication to even modern sound system manufacturing. Diamond deposition by plasma, oxy-acetylene flame, and hot-filament chemical vapor deposition is a rapidly growing technology. Intense interest exists in the study of the basic reaction mechanisms in both the gas-phase and surface chemistries, because presently diamond synthesis remains more an art than an empirical process. New laser diagnostics developed during this study are being directly applied to various diamond deposition environments, such as inductively coupled plasma torch and hot-filament chemical vapor deposition, under the continuing collaboration with the nonequilibrium plasma chemistry program of Prof. Charles H. Kruger at the High Temperature Gas Dynamics Laboratory, Stanford University.

Our principal diagnostic tool is based on cavity ring-down spectroscopy (CRDS). CRDS is a high-sensitivity absorption technique with potential for absolute concentration measurements of trace gases and impurities[1]. CRDS is usually practiced by coupling a pulsed laser source into a high-finesse optical resonator (Fabry-Perot cavity) that encloses the sample of interest, and detecting the decay of light in the resonator. Under many conditions, the decay is exponential, and a plot of the ring-down lifetime versus frequency gives the absorption spectrum[2]. The ring-down lifetime is controlled by the resonator finesse, and changes wherever the sample absorbs the wavelength of the incident radiation.

Most diagnostics used in research, however, tend to rely on expensive equipment that is difficult to maintain. To increase the utility of our diagnostics, we have begun to investigate practical schemes for CRDS. In particular, laser diodes, owing to their small size, low cost and relative ease of use, have begun to play a more dominant role in our research, and will open the possibility of portable diagnostics.

A MODEL SYSTEM: DIAMOND FILM GROWTH

A particularly suitable system for study of energy-related phenomena is the diamond film reactor, where the growth mechanism directly involves plasma chemistry. Two of the commonly used diamond film deposition methods are a CVD reactor using hot-filament chemical vapor deposition (HFCVD) or an inductively coupled atmospheric plasma torch. Both techniques are already under investigation at the Stanford High Temperature Gas Dynamics Laboratory. In order to understand the elementary growth mechanisms involved in diamond deposition, data bases of information about the numerous radicals present (e.g., hydrocarbon radicals as CH_3, CH_2, CH, C_2H, C_2, etc., or atomic hydrogen) are being compiled and will be used in future computer modeling, and subsequent numerical simulation of the complex plasma chemistry (e.g., gas-phase reactions of atomic hydrogen with hydrocarbon radicals or diamond interface reactions of atomic hydrogen selectively with graphite).

A CRDS setup has been designed to measure trace radical species generated in a hot-filament reactor for diamond deposition[3]. The methyl (CH_3) radical is an important free radical present during the initial stages of hydrocarbon combustion: it is believed to be a precursor for diamond growth by CVD. *In situ* measurements of methyl radical concentrations (cf. schematic diagram of reactor in figure 1a) have been carried out under various conditions[4,5]. Typically, a mixture of H_2 of CH_4 is flowed through the previously evacuated reactor. A tungsten filament is positioned vertically inside the reactor chamber and is resistively heated to a specified brightness temperature. Methyl radical absorption is observed near 216 nm, where feature lines are a few nm wide (cf. figure 1b)[3]. It is also important that the ground-state population of the absorber molecule is not significantly depleted by excitation during the time the laser pulse is circulating inside the optical cavity. In our experiment, for 216 nm light pulse of energy about 0.2 mJ and TEM_{00} mode radius w = 250 mm, for mirrors reflectivity R = 0.991, and for CH_3 absorption cross-section s < 10^{-17} cm^2 / molecule, the fraction of molecules excited by the laser pulse inside the cavity is less than $3x10^{-3}$, which is sufficient for accurate CRDS measurements.

A profile of CH_3 absolute concentration near the hot filament has been determined by CRDS using a topological method - Abel inversion of the spatial profile of CH_3 absorbance (cf. figure 1c)[3,4,5].

This approach allowed us to estimate the uncertainty in the inverted profile. The error bars represent one standard deviation. The shaded part of the figure indicates radial distances from the filament where the gas temperature is between 1250 K and 2000 K. Based on a hydrogen diffusion model, methyl concentration should peak at the filament. It was unexpectedly observed, however, to peak about 5 mm from the filament. This behavior can possibly be explained by the Soret effect or dissociation of methyl near the filament (cf. figure 1d)[4,5].

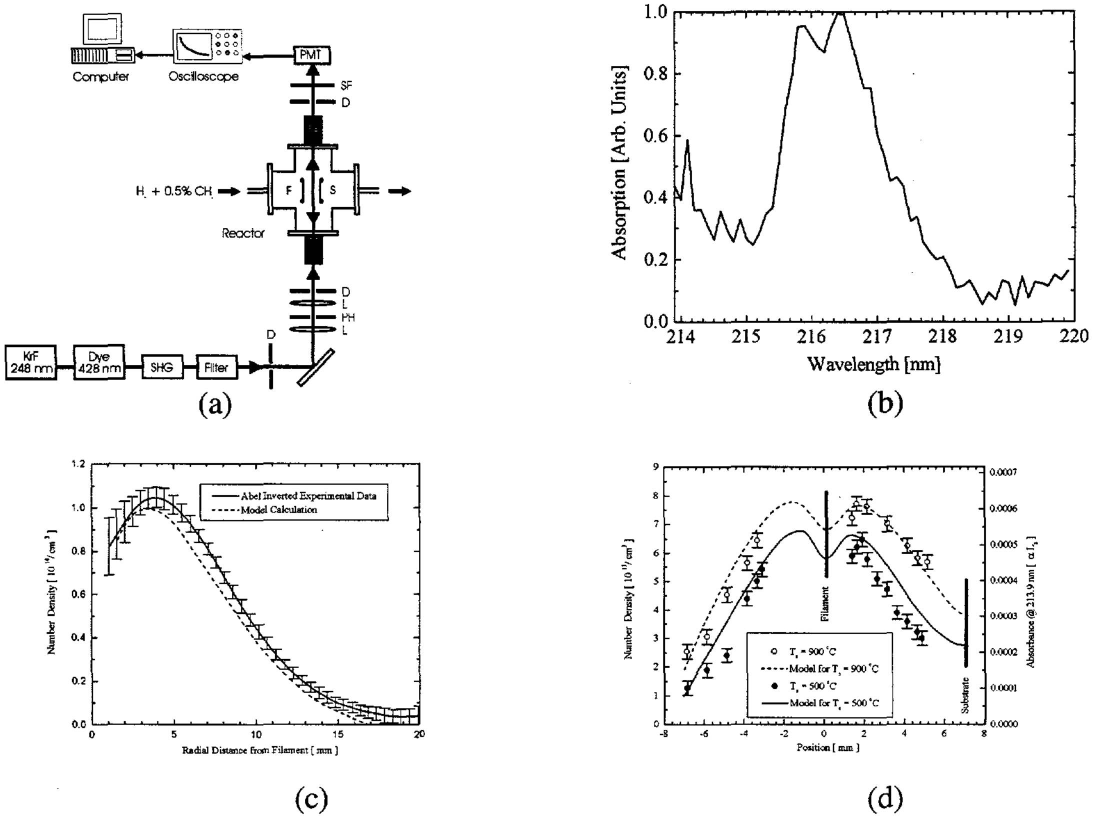


Figure 1: (a) CRDS setup for radical concentration measurements, (b) spectrum of methyl absorption at 216 nm, (c) radial distribution of CH_3, and (d) spatial profiles of the measured number density within the hot-filament reactor at two different substrate temperatures.

EXTENSIONS TO THE MID-INFRARED

The 1.5 to 10 μm region of the electromagnetic spectrum is rich in rovibrational transitions forming molecular "fingerprints" that are well known to be a means for identifying and characterizing specific species. This region is therefore rather ideal for mapping species concentration or temperature gradients in hot-filament reactors and arc jets. We have begun to exploit the high sensitivity, linearity, and simplicity in quantifying number densities provided by CRDS in the mid-infrared.

The application of CRDS to a problem presupposes the existence and availability of suitable light sources and cavity mirrors. With the advent of nonlinear optical devices, it has recently become possible to obtain tunable coherent light sources in the mid-infrared based on optical parametric oscillators (OPO)[2]. Simultaneously, highly reflecting mirrors with only minute scattering and absorption losses have become available for wavelengths in the visible and the near infrared regions.

Our light source is a Nd:YAG laser-pumped OPO system (Continuum Mirage 3000) that can generate nearly Fourier transform-limited nanosecond Gaussian pulses with a manufacturer-specified bandwidth of 500 MHz (0.017 cm^{-1}) at a repetition rate of 10 Hz[2]. The wavelength can be tuned continuously from 1.5 to 4.0 µm, with the pulse energy decreasing from 8 mJ to 1 mJ, respectively. The OPO system architecture is shown in figure 2a[2].

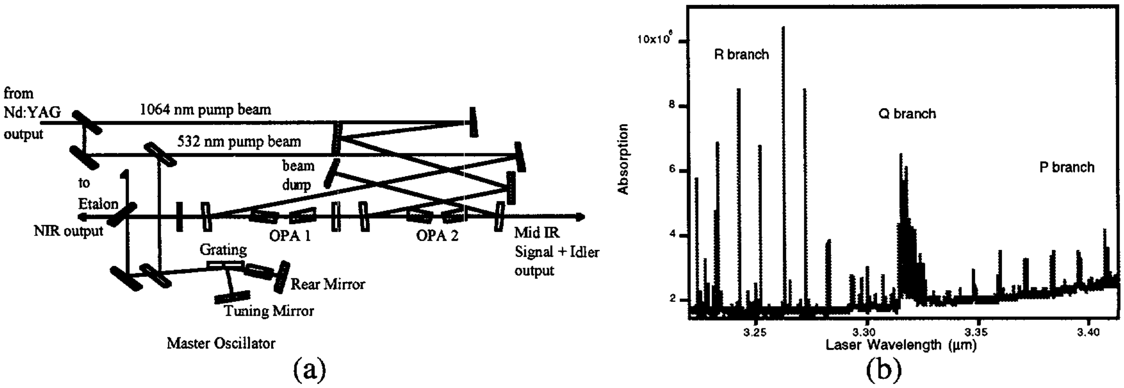


Figure 2: (a) Continuum Mirage 3000 OPO system diagram, and (b) absorption spectrum of a 100 ppm CH_4 in N_2 mixture at 50 Torr pressure.

We are currently pursuing CRDS studies of the well-known methane fundamental C-H stretching mode (n3), that occurs around 3.17 mm, and should serve as a good reference for future calibrating purposes. A typical absorption spectrum is given in figure 2b. All recorded spectra showed a very strong absorption, allowing us to record methane lines below 10^{-8} Torr partial pressure in N_2.

We have also applied our OPO system to the measurement of water vapor in various types of flames, to demonstrate the effectiveness of CRDS as a diagnostic tool for hostile environments such as flames, discharges, flashes, or plasmas[6]. A strong need exists for spectroscopic methods that can serve as remote diagnostics in these environments because they remain difficult to characterize, owing to their wide range of extreme physical conditions: high temperatures and consequently strong luminous background, sharp gradients in both temperature and density, and a reactive medium with ions, electrons and a variety of free radicals or intermediate states. CRDS, a laser-based spectroscopy, which is noninvasive, species specific, and spatially resolved, is ideally suited for probing environments like these.

We have measured the spectrum of water vapor in air from 810 to 820 nm, from atmospheric pressure to 20 mTorr, with a resolution of 0.001 nm (0.015 cm^{-1}). This demonstrates a nominal measurement sensitivity (with R=99.99% mirrors) to absorption coefficients as low as 1.7 x 10^{-7} cm^{-8}. We have also been able to extract accurate species partial pressure measurements of water vapor in a regulated cell (figure 3)[6]. We have subsequently measured a similar spectra of water vapor generated at the tip of a propane torch flame (T = 2000 K), and at various heights above a controlled plane methane-air burner[6]. By using the HITEMP database, we can extract rotational temperatures of water vapor at different heights above the plane burner surface. Figure 4 compares spectra of water vapor at room temperature to those in the propane flame, while figure 5 illustrates changes in the water spectrum, caused by the decreasing temperature gradient in a controlled flame.

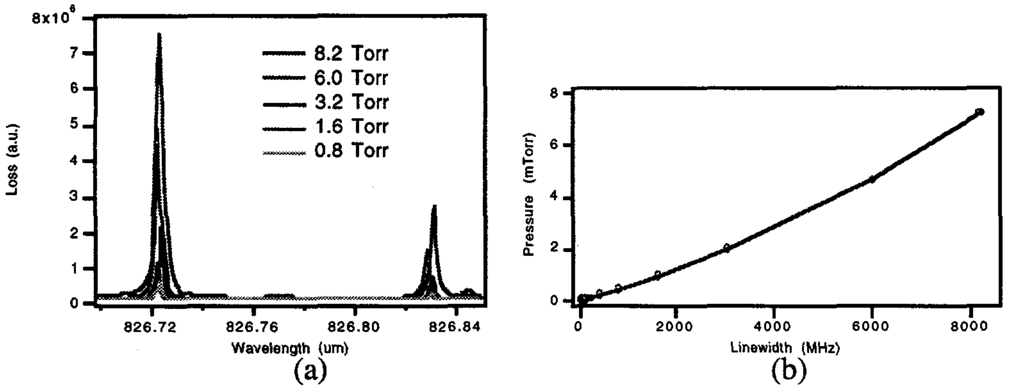


Figure 3: (a) Absorption spectra of water vapor at various cell pressures, and (b) variation of linewidth with pressure.

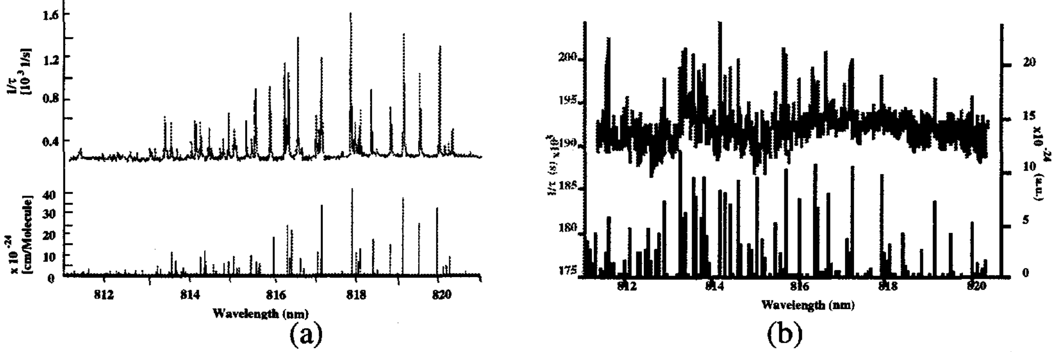


Figure 4: Absorption spectra of water vapor at (a) room temperature, and (b) at the tip of a propane torch. Measured spectra are at the top, while spectra from HITRAN96 are on the bottom.

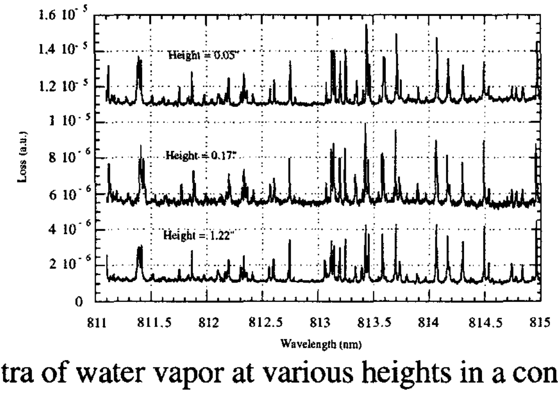


Figure 5: Absorption spectra of water vapor at various heights in a controlled methane burner.

MINIATURIZATION WITH LASER DIODE SOURCES

Much of current ring-down spectroscopy still relies on fairly costly laser sources. As solid state lasers (e.g., Ti:Sapphire lasers, Nd:Yag-pumped OPOs, and ECDLs) have gained in reliability, tuning range, and output power, they have started to replace the more traditional tunable dye lasers, although they are no less expensive. Simultaneously, semiconductor laser diodes (LDs) have also been improving in power, wavelength coverage, and reliability. The rapid growth of the communications industry in recent years has resulted in the availability of tunable UV-, near- and mid-infrared LDs at a rapidly diminishing cost (< $2000). In fact, owing to their compactness, low cost, durability, high wallplug efficiency, and compatibility with both fiber and silicon technologies, infrared laser diodes seem to be an ideal light source for realizing practical CRDS systems.

Early attempts demonstrated difficulties in applying LD sources to CRDS: whenever a LD beam is reflected directly back into the laser, as is inevitable in a linear cavity configuration, even under optical isolation, the optical feedback results in phase fluctuations and mode hopping of the LD. In fact, at higher feedback levels, a wide variety of effects ranging from linewidth broadening to complete `coherence collapse' (linewidth > 10 GHz) is often observed and is illustrated in figure 6a[7]. The inherent problem is the formation of `external cavities' by reflective optics with the back facet of the LD that affect both the gain and phase relations of the LD. Thus, whenever back reflection is allowed, the lasing characteristics of the become highly dependent on uncontrollable experimental parameters, most notably the external cavity length.

Several solutions exist to this coupling problem. A LD with a high quality (but expensive) AR coated output facet can function as a gain medium in an external cavity; the feedback from a linear cavity configuration can be completely eliminated by using a ring resonator structure, as will be investigated in the future; or, the external cavity effect can be controlled by placing an acousto-optic modulator (AOM) inside the external cavity, thereby stabilizing the time-averaged behavior of the LD. The last approach, first demonstrated by Martin et al.[7] as a useful scheme for stabilizing LDs in the presence of direct back reflections, was the point of departure for our LD research.

By placing an AOM between the laser diode and the input mirror of the ring-down cavity, the AOM can be used not only to switch the CW beam into and out off the first order diffraction, but simultaneously control LD linewidth. The AOM driving power determines the diffraction efficiency and hence the amount of feedback to the LD. The external cavity length fixes the maximum achievable linewidth for each feedback level (cf. figure 6b)[8]. The first order diffraction feedback drives the LD phase and stabilizes linewidth. Finally, the linewidth can be further enhanced by introducing nonfrequency-shifted that cyclically chirps the LD output through multiple external cavity modes, at twice the AOM driver frequency (cf. figure 6c)[8]. The flexibility in achievable LD linewidth in turn enables many different CRDS applications.

Using the AOM stabilization scheme for a laser diode source, shown in figure 7, we were able to perform CRDS on water vapor present in ambient air or in an evacuated optical cavity[8]. LD linewidth control was performed with feedback from both first and zeroth orders. Spectra of water vapor in room air and at 5 Torr are given in figure 8[8]. Spectra were obtained in one continuous scan. Spectra at ambient pressure used maximum zeroth order feedback (47.6 dB) to achieve the largest possible linewidth (240 - 500 MHz) and cavity coupling. Spectra at low pressures (< 100 Torr) used less zeroth order coupling (58 dB) to achieve a narrower laser linewidth (180 - 240 MHz) and to avoid convolution of the laser line with the absorption line. Scan step size in both cases remained limited to 0.001 nm resolution by the current step resolution (0.1 mA) of the LD driver. No baseline adjustments have been made, and the overall baseline noise results from the excitation of multiple transverse modes in the cavity, which were used to improve light throughput. Nonetheless, our detection limit of 2 x 10^{-8} cm^{-1} remains quite respectable for an inexpensive LD source, especially when compared to pulsed CRDS.

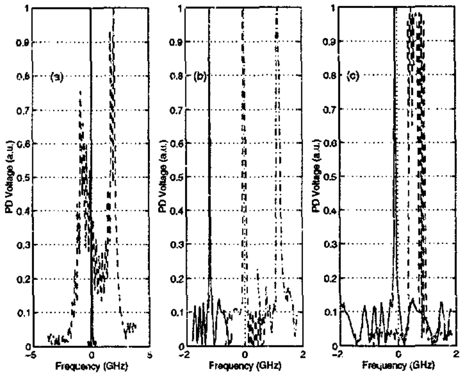


Figure 6: (a) Linewidth for a free-running LD (solid) and for a LD under feedback (dashed). (b) LD linewidth as a function of external cavity length for only first order feedback: L_{ext}=215 cm (solid), L_{ext}=100 cm (dashed), and L_{ext}=215 cm (dash-dotted), (c) LD linewidth for only first order feedback (solid) and both first and zeroth order feedback (dashed).

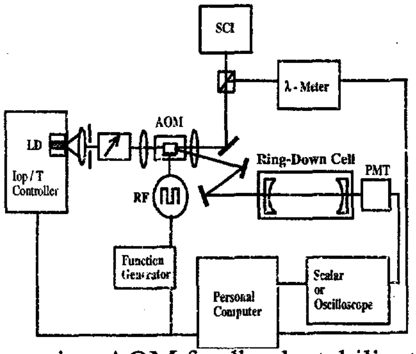


Figure 7: Laser diode CRDS setup using AOM feedback stabilization.

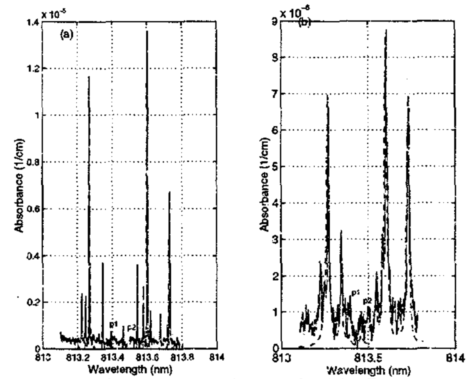


Figure 8: Spectrum of (a) water vapor in room air and (b) 5 Torr water vapor in a cell previously evacuated below 1 mTorr. Spectra based on HITRAN96 are shown as dashed lines.

CONCLUSIONS

CRDS has been applied for quantitative diagnostic study of methyl radicals in a hot-filament reactor used for diamond film synthesis. The methyl radical concentration was found to peak at several mm away from the filament surface, and is attributed to the effect of Soret diffusion. We have extended the diagnostic capabilities of our OPO laser from near-infrared studies of water vapor in harsh environments, such as flames, to mid-infrared studies of the C-H stretch in methane. This will allow us to perform highly sensitive CRDS diagnostics of an arc-jet torch used for diamond synthesis.

Simultaneously, we have demonstrated that it is possible to not only stabilize a free-running laser diode in the presence of strong reflections from a ring-down cavity, but also control the linewidth of the laser diode. The laser diode can also be stabilized to only several MHz, if high resolution is required. We have performed CW-CRDS with ring-down repetition rates of 10-50 kHz, and have achieved a noise level of 2×10^{-8} cm^{-1}, comparable to pulsed CRDS.

ACKNOWLEDGMENTS

The authors would like to acknowledge contributions by the companion research program of Prof. C. H. Kruger, with Dr. T. G. Owano and E. H. Wahl (DOE Grant #DE-FG0388ER13957). This work has been supported by the Office of Basic Energy Research at the Department of Energy (contract #DE-FG03-92ER14303).

REFERENCES

1. P. ZALICKI and R. N. ZARE, "Cavity ring-down spectroscopy for quantitative absorption measurements," *J. Chem. Phys.* 102, 2708 (1995).

2. J. MARTIN, B. A. PALDUS, P. ZALICKI, E. H. WAHL, T. G. OWANO, J. S. HARRIS, C. H. KRUGER, and R. N. ZARE, "Cavity Ring-down Spectroscopy with Fourier-transform-limited Light Pulses," *Chem. Phys. Lett.* 258, 63 (1996).

3. P. ZALICKI, Y. MA, R. N. ZARE, E. H. WAHL, J. R. DADAMIO, T. G. OWANO and C. H. KRUGER, "Methyl radical measurement by cavity ring-down spectroscopy," *Chem. Phys. Lett.* 234, 269 (1995).

4. P. ZALICKI, Y. MA, R. N. ZARE, E. H. WAHL, T. G. OWANO, and C. H. KRUGER, "Measurement of methyl radical concentration profile in a hot-filament reactor," *Appl. Phys. Lett.* 67, 144 (1995).

5. E. H. WAHL, T. G. OWANO, C. H. KRUGER, P. ZALICKI, Y. MA, and R. N. ZARE, "Measurement of absolute CH_3 concentration in a hot-filament reactor using cavity ring-down spectroscopy [diamond CVD]," *Diamond and Related Materials* 5, 373 (1996).

6. J. XIE, J. MARTIN, B. A. PALDUS, E. H. WAHL, M. ZHAO, T. G. OWANO, C. H. KRUGER, and R. N. ZARE, "Cavity Ring-down Spectroscopic Measurements in a Flame," *Appl. Phys. Lett.* (in preparation).

7. J. MARTIN, Y. ZHOA, S. BALLE, K. BERGMANN, and M. P. FEWELL, "Visible-wavelength diode laser with weak frequency-shifted optical feedback," *Optics Communications* 112, 109 (1994).

8. B. A. PALDUS, J. S. HARRIS, J. MARTIN, J. XIE, R. N. ZARE, "Cavity Ring-Down Spectroscopy Using a Frequency-stabilized Laser Diode," *J. Appl. Phys.*, (submitted April 1997).

Final List of Participants

15th Symposium on Energy Engineering Sciences

May 14-15, 1997

Argonne National Laboratory
Argonne, Illinois

S. George Bankoff
Department of Chemical Engineering
Northwestern University
Evanston, IL 60208

Jacob Barhen
Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6355

Tamer Basar
Coordinated Science Laboratory
University of Illinois, Urbana-Champaign
1308 West Main Street
Urbana, IL 61801-2307

Bruce S. Berger
Department of Mechanical Engineering
University of Maryland-College Park
College Park, MD 20742

Harvey W. Blanch
Department of Chemical Engineering
University of California-Berkeley
Berkeley, CA 94720

Ivan Catton
Department of Mechanical and Aerospace Engineering
University of Calfornia-Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

Hermann Fasel
Department of Aerospace and Mechanical
Engineering
University of Arizona
Building 16, Rm. 301
Tucson, AZ 85721

James R. Fincke
Department of Optical and Plasma Physics
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2211

Daniel Frederick
Department of Engineering Science and
Mechanics
Virginia Polytechnic Institute & State University
1410 Highland Circle
Blacksburg, VA 24060

L. B. Freund
Division of Engineering
Brown University
Box D
Providence, RI 02912

Bijoy K. Ghosh
Department of Systems Science and Mathematics
Washington University
One Brookings Drive
St. Louis, MO 63130

Robert Goulard
Division of Engineering and Geosciences, ER-15
U.S. Department of Energy
Office of Basic Energy Sciences
19901 Germantown Road
Germantown, MD 20874

Takashi Hibiki
School of Nuclear Engineering
Purdue University/Kyoto University
1290 Nuclear Engineering Building
West Lafayette, IN 47907-1290

Cynthia D. Holcomb
Physical and Chemical Properties Division
National Institute of Standards & Technology
325 Broadway
Boulder, CO 80303

Mamoru Ishii
Department of Nuclear Engineering
Purdue University
1290 NUCL
West Lafayette, IN 47906

Daniel D. Joseph
Department of Aerospace Engineering
University of Minnesota
110 Union Street
Minneapolis, MN 55455

Allan N. Kaufman
Lawrence Berkeley Laboratory
MS 4/230
Berkeley, CA 94720

Gunol Kojasoy
Department of Mechanical Engineering
University of Wisconsin-Milwaukee
P.O. Box 784
Milwaukee, WI 53201

Erhard Krempl
Mechanical Engineering, Aeronautical
Engineering and Mechanics
Rensselaer Polytechnic Institute
110 8th Street
Troy, NY 12180-3590

Paul A. Libby
Department of Applied Mechanics and
Engineering Science
University of California-San Diego
9500 Gilman Drive
La Jolla, CA 92093-0411

Katja Lindenberg
Department of Chemistry and Biochemistry 0340
University of California-San Diego
9500 Gilman Drive
La Jolla, CA 92093-0340

Mark J. McCready
Department of Chemical Engineering
University of Notre Dame
182 Fitzpatrick Hall
Notre Dame, IN 46556

Francis C. Moon
Department of Mechanical and Aerospace
Engineering
Cornell University
204 Upson Hall
Ithaca, NY 14853

Paul E. Murray
Department of Materials Joining
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2210

Joseph O'Gallagher
Department of Physics
The University of Chicago
Enrico Fermi Institute
5720 South Ellis Avenue
Chicago, IL 60637

Alfonso Ortega
Department of Aerospace and Mechanical
Engineering
University of Arizona
Building 16, Rm. 301
Tucson, AZ 85721

Lynne E. Parker
Computer Science and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6364

Tomio Y. Petrosky
Ilya Prigogine Center for Statistical Mechanics and
Complex Systems
The University of Texas-Austin
RLM 7.220
Austin, TX 78712

Robert E. Price
Division of Engineering and Geosciences, ER-15
U.S. Department of Energy
Office of Basic Energy Sciences
19901 Germantown Road
Germantown, MD 20874

Seth J. Putterman
Department of Physics
University of California-Los Angeles
Los Angeles, CA 90024

Nageswara S. Rao
Center for Engineering Systems and
Advanced Research
Oak Ridge National Laboratory
P.O. Box 2008, MS 6364
Oak Ridge, TN 37831-6364

Walter G. Reuter
Department of Metals and Ceramics
Lockheed Martin Idaho Technology Company
P.O. Box 1625
Idaho Falls, ID 83415-2218

Hermann Riecke
Department of Applied Mathematics
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

Huseyin Sehitoglu
Department of Mechanical and Industrial
Engineering
University of Illinois, Urbana-Champaign
1206 West Green Street
Urbana, IL 61801

Herschel B. Smartt
Department of Materials Joining
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2210

Vladi S. Travkin
Department of Mechanical and Aerospace
Engineering
University of California-Los Angeles
48-121 Engineering IV
Box 951597
Los Angeles, CA 90095-1597

Roland Winston
Department of Physics
University of Chicago
5640 South Ellis Avenue
Chicago, IL 60637

Qiao Wu
Department of Nuclear Engineering
Purdue University
West Lafayette, IN 47907

Ying Xu
Computer Science and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2008, MS 6364
Oak Ridge, TN 37831-6364

Richard N. Zare
Department of Chemistry
Stanford University
Mudd Building
Stanford, CA 94305-5080