Correlation with a spatial light modulator having phase and amplitude cross coupling

Richard D. Juday

In correlation filtering a spatial light modulator is traditionally modeled as affecting only the phase or only the amplitude of light. Usually, however, a single operating parameter affects both phase and amplitude. An integral constraint is developed that is a necessary condition for optimizing a correlation filter having single parameter coupling between phase and amplitude. The phase-only filter is shown to be a special case.

I. Introduction

There are programmable continuously variable spatial light modulators (SLMs) that are driven by a single parameter (voltage or charge in electrically addressed SLMs, intensity in light-addressed SLMs). The Hughes liquid crystal light valve (LCLV) is an example of the light-addressed variety, Texas Instruments deformable mirror device (DMD) is an electrically addressed SLM. An SLM affects both phase and amplitude to varying degrees. The cross coupling between phase and amplitude may be adjusted during the manufacture of the device or in the details of how it is inserted into the stream of the correlation. But given that the SLM is controlled through a single parameter, one cannot achieve any arbitrary combination of phase and amplitude. Once an SLM is installed in a correlator, the cross coupling is not usually adjustable, and there is never pixel-by-pixel independence of the cross coupling.

The situation is illustrated (for one location in the filter plane) in Fig. 1. A computed matched filter value does not lie on the SLM operating curve, but we must choose a point on that curve to represent it. We might choose the indicated matched-phase point (solid square), but a nearby point (open circle) of nearly the same phase has a larger amplitude and thus may be expected to contribute to a larger correlation signal. One rightly suspects that the optimum representation of any one filter element depends on the ensemble of elements in the filter plane. We develop an integral constraint to help in optimizing the ensemble of representations.

Most previous work in simulating correlations has not included an SLM's cross coupling between phase and amplitude, and none known to this author has explicitly included the full cross coupling while optimizing the filter. In strongly related work Horner[1] did show the effect of coupled amplitude and phase, but he did not adjust what we call the phase-mostly filter (PMF) for exact maximum performance in the presence of the coupling. Horner's coupling was also limited as a function to the first few terms in a Taylor series. In the present paper we develop the explicit inclusion of the cross coupling of an arbitrary (but presumed known) form while creating an optimum correlation filter. If the cross coupling can be adjusted, it can be varied to seek a combination of cross coupling and filter signal that further optimize the correlation. Juday *et al.*[2] showed a simulation of optimizing a cross-coupled filter by the brute force relaxation method of Juday and Daiuto.[3] The method given here considerably improves on the brute force method by reducing the amount of computation needed.

The analysis is presented for a 1-D signal, but the generalization to image processing in two dimensions is obvious. We assume that we are presented at the Fourier plane with a known transform $X(\omega)$. We assume further that the correlation performed by the subsequent optics and detection can be modeled as shown, i.e., that the system performs an exact Fourier transform and that all diffraction orders are caught with no windowing.

Following Franks[4] we distinguish between a function and a functional as follows. A function maps from the domain of complex numbers into the range of complex numbers. For a functional, the domain is a set of functions and the range is a set of complex

The author is with NASA Johnson Space Center, Tracking & Communications Division, Houston, Texas 77089.

Received 4 October 1988.

numbers. A functional is to be thought of as a function of a function. The Fourier transform is an example of a functional:

$$X(\omega) = \int_{-\infty}^{\infty} \exp(-j\omega t)x(t)dt$$

carries the full behavior of $x(\cdot)$ into the value of $X(\cdot)$ at each ω. We will also distinguish between a function $x(\cdot)$ and the value of the function $x(\omega)$ at a particular value ω of the argument.

II. Formulation of the Problem

Suppose we wish to filter a signal to recognize the presence of a reference pattern $x(t)$ whose transform is $X(\omega)$:

$$x(t) \rightarrow X(\omega) = A(\omega) \exp[j\phi(\omega)] \tag{1}$$

by using a SLM constrained to the values

$$H(\omega) = f[s(\omega)] \exp\{jg[s(\omega)]\}, \tag{2}$$

where $s(\omega)$ is the control value for the filter at frequency ω, and $f(s)$ and $g(s)$ are, respectively, the modulator's resulting amplitude and phase. Once $s(\cdot)$ is chosen, it is converted into the filter $H(\cdot)$; the usual properties of optical correlation then follow—linearity, shift invariance, etc. The filtered signal produces a field E_s at the correlation output plane:

$$E_s(t) = \int_{-\infty}^{\infty} \exp(-j\omega t)A(\omega) \exp[j\phi(\omega)]f[s(\omega)] \times \exp\{jg[s(\omega)]\}d\omega$$

$$= \int_{-\infty}^{\infty} \exp(-j\omega t)X(\omega)H(\omega)d\omega. \tag{3}$$

Signal E_s is intensity detected. We need a criterion to be maximized by choice of $s(\cdot)$. We concentrate on a simple quantity, the intensity at the center of the correlation plane in response to the centered reference object. This is generally a beneficial thing to do, although other criteria may be of more practical import and will be treated in further work. The stated optimization corresponds to maximizing the magnitude of the overlap integral

$$\int_{-\infty}^{\infty} X(\omega)H[s(\omega)]d\omega$$

by choice of $s(\cdot)$. Generalization to other optimization criteria is conceptually straightforward. Thus we shall determine an integral constraint to be satisfied if I_s is to be maximized, where

$$I_s := E_s^*(t = 0)E_s(t = 0). \tag{4}$$

The symbol := is the asymmetrical identity ($a := b$ means a is defined by b), and E^* is the complex conjugate of E. The problem may now be stated: First, given $A(\omega)$, $\phi(\omega)$, and the functions $f(\cdot)$ and $g(\cdot)$, determine the scalar functions $s(\cdot)$ that produce values E_s whose first-order variation is only in phase. Second, among the finite number of those choices for $s(\cdot)$, select the one whose magnitude of E_s is maximum.

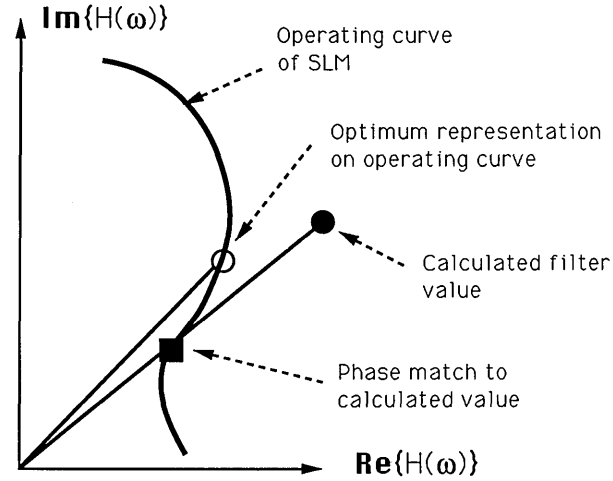

Fig. 1. Representation of the complex value of a filter value by a point lying on the operating curve of a spatial light modulator. The solid dot is a desired value such as the classical matched filter. The square matches the phase of the desired value. The open circle nearly matches the phase and, by having a larger amplitude than the square, transmits more energy into the correlation intensity. The optimum representation of any one value depends on the ensemble of desired values.

III. Formulation of the Solution

Henceforth we implicitly look at only the central value of the filtered signal, since for our present purposes we are optimizing only that value.

We presume reasonable behavior, including invertibility, for $f(s) \exp[jg(s)]$. That is, if we know what phase and amplitude change we wish the SLM to produce (and they lie on its operating curve), we know what signal to apply to the SLM. Particular interest applies to $g(\cdot)$ because phase is often the stronger operational parameter one wishes to control.

We can develop necessary conditions for an optimal $s(\cdot)$ by using the concept of a varied function from the calculus of variations. We loosely follow Arfken.[5]

A necessary condition for maximum central intensity is that the variation in I_s be zero. The only possible sources of variation in I_s are the amplitude and phase of E. With D the (real) amplitude and q the phase of (complex) electric field E, the (real) intensity I is

$$I = E^*E = [D \exp(jq)^*][D \exp(jq)] = D^2. \tag{5}$$

Changing the phase, as is well known, does not affect the intensity I. We shall constrain $s(\cdot)$, then, so that an arbitrary first-order change in it will cause nothing other than a phase change of E.

We form the small variations of $s(\cdot)$ as follows. We introduce an arbitrary function $\mu(\cdot)$; it need not have a small value anywhere, although we wish it to be bounded. $\mu(\cdot)$ gives the shape of the change we induce into $s(\cdot)$. It is the generalization of a vector in an ordinary directional derivative. Now, for an arbitrary scalar α we form the varied function

$$r(\cdot,\alpha) = s(\cdot) + \alpha\mu(\cdot), \tag{6}$$

and E_s changes to E_r, a function of α according to $r(\cdot,\alpha)$:

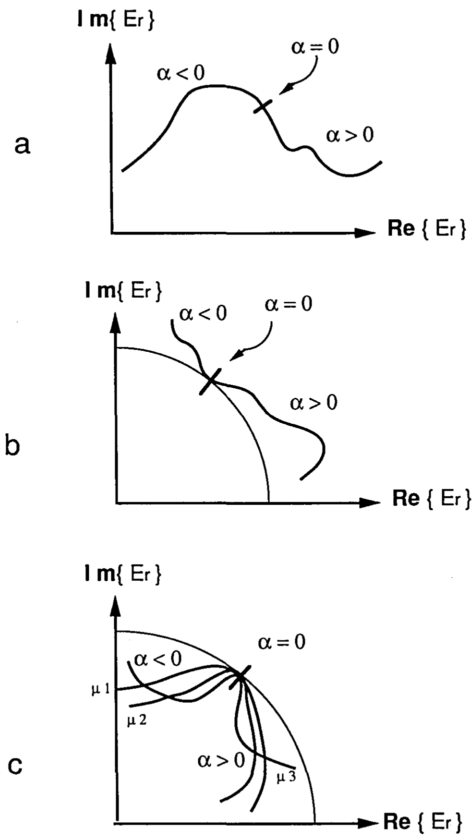

Fig. 2. Hypothetical curves showing the change in the complex value of E_r as α is changed for some choice of $\mu(\cdot)$: (a) $|E_r|$ is not stationary at $\alpha = 0$ for this combination of $s(\cdot)$ and $\mu(\cdot)$. (b) Although $|E_r|$ is stationary at $\alpha = 0$, it is at a minimum. (c) To maximize the correlation strength by choice of $s(\cdot)$ we require that all choices for $\mu(\cdot)$ will produce a maximum in $|E_r|$ for $\alpha = 0$.

$$E_r(\alpha) = \int_{-\infty}^{\infty} A(\omega) \exp[j\phi(\omega)]f[r(\omega)] \exp\{jg[r(\omega)]\}d\omega. \tag{7}$$

Figure 2 shows some possible tracks of E_r as functions of α. We note that $E_s = E_r|_{\alpha=0}$. For differentially small α, we observe the rate at which $E_r(\alpha)$ changes with respect to α. For $\alpha = 0$ that rate of change can be thought of as the component of the gradient of E_s in the direction of the function $\mu(\cdot)$. We can then constrain the function $s(\cdot)$ to produce the specified behavior of the gradient of E_s.

From Eq. (6) we have the partial derivative

$$\frac{\partial r}{\partial \alpha} = \mu. \tag{8}$$

The first-order variation in E_s corresponding to α and μ is δE_s. Similar to the first term of a Maclaurin series expansion,

$$\begin{aligned}\delta E_s &:= \alpha \left\{\frac{\partial}{\partial \alpha} \int_{-\infty}^{\infty} A \exp(j\phi)f(r) \exp[jg(r)]d\omega\right\}\bigg|_{\alpha=0} \\ &= \alpha \left\{\int_{-\infty}^{\infty} A \exp(j\phi) \frac{\partial}{\partial \alpha} \{f(r) \exp[jg(r)]\}d\omega\right\}\bigg|_{\alpha=0} \\ &= \alpha \int_{-\infty}^{\infty} A \exp\{j[\phi + g(s)]\}\mu[f'(s) + jf(s)g'(s)]d\omega, \end{aligned} \tag{9}$$

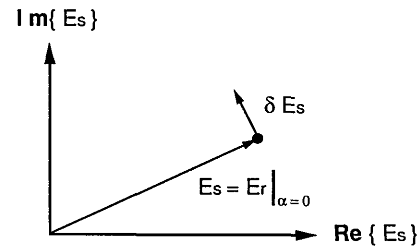

Fig. 3. Variation in E_s, δE_s is a function of $\mu(\cdot)$. If $|E_r|$ is an extremum at $\alpha = 0$, δE_s is perpendicular to E_s for any choice of $\mu(\cdot)$.

where A, ϕ, r, μ, and s are all functions of ω, and partial derivatives become ordinary derivatives for functions of one variable.

For the best filtering function $s(\cdot)$, the point along the track of E_r where its magnitude is a maximum must occur at $\alpha = 0$, for any disturbance function μ. This means that there must exist a real constant [a functional generally dependent on $\mu(\cdot)$ and hence denoted k_μ] so that for differentially small α (see Fig. 3)

$$E_s + \delta E_s \approx \exp(jk_\mu \alpha)E_s, \tag{10}$$

$$\frac{\delta E_s}{E_s} \approx \exp(jk_\mu \alpha) - 1 \approx jk_\mu \alpha. \tag{11}$$

Equation (11) is seen to be a statement of the phase of the integral expression [Eq. (9)] for δE_s inasmuch as E_s is a complex constant and k_μ is real. Being entirely arbitrary, the variation function μ can pick out any localized frequency interval and the result must not change the phase of the integral. Consequently, the phase of the integrand for δE_s must not vary with frequency. Since μ and A are real, from Eqs. (3), (9), and (11) the ratio

$$\frac{\exp\{j[\phi + g(s)]\}[f'(s) + jf(s)g'(s)]}{E_s} \tag{12}$$

must be purely imaginary. The numerator must have constant phase (call it Θ) if $s(\cdot)$ is optimum. For the ratio itself to be purely imaginary Θ must also satisfy the constraint

$$E_s = jD \exp(j\Theta) \tag{13}$$

with some real constant D. Because (within modulo 2π)

$$\arg\{y \exp(jx)\} = x + \arg\{y\}, \tag{14}$$

once Θ is in hand we calculate $s(\omega)$ from

$$\phi(\omega) + g(s) + \arg\{f'(s) + jf(s)g'(s)\} = \Theta. \tag{15}$$

For any value of the SLM control value s, there is a single value of the expression

$$g(s) + \arg\{f'(s) + jf(s)g'(s)\} =: a(s), \tag{16}$$

which we shall call the augmented phase of the SLM [since $g(s)$ is the actual physical phase modulation of the SLM]. The augmented phase is the analog of the phase produced in a phase-only SLM, including additionally the amplitude variation of the cross-coupled SLM as it affects the correlation process. A physical

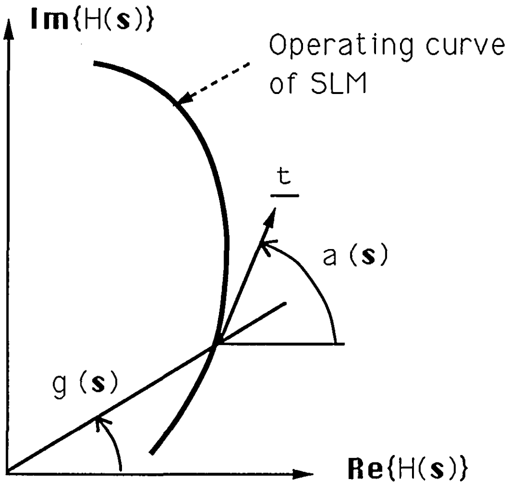

Fig. 4. Physical interpretation of an augmented phase; an SLM operating curve is drawn in the complex plane. For a control value s the SLM phase value is $g(s)$, and the augmented phase is $a(s)$. Vector **t** is the tangent to the operating curve. Since the operating curve of a phase-only SLM is a circle, its augmented phase would be $g(s) + \pi/2$.

interpretation of the augmented phase is shown in Fig. 4. Because we will often move from an augmented phase to the control value that produces it, we introduce a function p. It is the (possibly multivalued) inverse such that $s = p[a(s)]$.

For a stationary value of $|E_s|$, then, a necessary condition is that

$$\phi(\omega) + a[s(\omega)] = \Theta, \tag{17}$$

and the constant Θ is converted to the control signal by

$$s(\omega) = p[\Theta - \phi(\omega)]. \tag{18}$$

Furthermore, however, the constant Θ must satisfy the integral constraint shown here with its full frequency dependence:

$$\int_{-\infty}^{\infty} A(\omega) \exp[j(\phi(\omega) + g\{p[\Theta - \phi(\omega)]\})] \times f\{p[\Theta - \phi(\omega)]\} d\omega = jD \exp(j\Theta), \tag{19}$$

where the only condition on D is that it be real. This is the integral constraint. It must be solved for the constant Θ that is consistent with the optimum correlation against the pattern whose transform is $A(\omega) \exp[j\phi(\omega)]$. Note that Θ appears deeply embedded in the left side of Eq. (19) and more accessibly on the right side. In Eq. (19) the amplitude spectrum $A(\omega)$ is seen explicitly to affect the optimum filter when cross-coupling is taken into account. Previous methods have not considered the amplitude spectrum in the optimization of a filter.

IV. Relationship to the Phase-Only Filter

We can easily particularize this result for the POF by letting $f \equiv 1$. For such an SLM,

$$\arg\{f'(s) + jf(s)g'(s)\} = \pi/2, \tag{20}$$

since $f(\cdot)$ and $g(\cdot)$ are real and $f' \equiv 0$. The augmented phase of a phase-only SLM is $g(s) + \pi/2$. Using Eq. (19) in Eq. (15) gives

$$\phi + g(s) + \pi/2 = \Theta, \tag{21}$$

$$g(s) = \Theta - \phi - \pi/2. \tag{22}$$

Equation (3) becomes

$$E_s = \int_{-\infty}^{\infty} A \exp(j\phi) \exp[j(\Theta - \phi - \pi/2)] d\omega$$

$$= \frac{1}{j} \exp(j\Theta) \int_{-\infty}^{\infty} A d\omega. \tag{23}$$

Comparing with Eq. (19), we identify D with the negative of the integral of the amplitude, a real quantity as required. The integral constraint [Eq. (19)] is satisfied by any value of Θ, a well-known result. The POF evens out the phase of the transformed reference image if that image occurs, properly centered, in the input plane. (Translating the reference image merely adds a phase tilt.) Any value for Θ will do, since adding a constant phase at the filtering plane has no effect on the intensity-detected correlation. Note also that the amplitude of the reference image spectrum does not enter the formulation for the POF, consistent with the POF's inability to alter amplitude.

A distinction between the optimum filter realized in a coupled SLM [Eq. (19)] and in a phase-only SLM [Eq. (23)] is apparent in the integral constraint for the coupled SLM. In the coupled SLM the amplitude spectrum of the reference image transform enters explicitly. If the phase of the coupled filter is changed by a uniform amount at all frequencies, the correlation strength can be affected—a consequence of the filter's changing amplitude while the phase is changing. Such is not the case for the phase-only filter.

V. Solution Method

We now outline two methods for solving the necessary conditions for a given cross-coupled modulator and a pattern to be recognized. In simulations the first has not always converged; the second is slower but more sure. Both methods require that the behavior of the SLM be known.

The form of the integral constraint fairly begs for an iterative solution. Begin with an arbitrary value Θ_1. At each frequency ω, solve the augmented phase equation for $s(\omega)$:

$$\phi(\omega) + g(s) + \arg\{f'(s) + jf(s)g'(s)\} = \Theta_1. \tag{24}$$

Drive signal $s(\omega)$ is put into the correlation integral to yield Θ_2 according to

$$\int_{-\infty}^{\infty} A \exp(j\phi) \exp[jg(s)] f(s) d\omega = \frac{1}{j} D \exp(j\Theta_2). \tag{25}$$

Thus Θ_2 is a function of Θ_1; when $\Theta_1 = \Theta_2$, the required condition is met, and we have a correlation amplitude–stationary value in the function $s(\cdot)$. The iterative generalization is immediately apparent. In simulations, however, it has generally diverged.

The more robust method is to scan Θ_1 in the whole range $[-\pi,\pi]$ in sufficiently fine detail to localize the values where the integral constraint is met, i.e., where $\Theta_1 = \Theta_2$, after Θ_2 is calculated as above. A binary search produces the value of Θ consonant with the accuracy in knowing the operating curves for the SLM phase and amplitude.

The formulations above generate an extremum. The correlation intensity is a scalar function of the scalar Θ_1, and determination of the global maximum intensity is relatively straightforward, given the assistance of the augmented phase function. In any event the solution is easily examined for yielding a maximum or minimum.

VI. Global Nature of the Solution

We have constructed a necessary condition on $s(\cdot)$ for the filter $H(\omega) = f[s(\omega)] \exp\{jg[s(\omega)]\}$ to produce a first-order stationary correlation intensity as $s(\cdot)$ is varied. We have also outlined a solution method for $s(\cdot)$. We now consider sufficient conditions for a calculated solution to be a global maximum. The matter is easily resolved if we have localized all solution values of Θ using the robust search method outlined above, since the correlation intensity for each can be tested to find out if it is a global maximum.

If the SLM is operated over a range of augmented phase equal to 2π, the solution given above is global. If, however, the SLM is able to wind through several 2π rad of augmented phase while its amplitude changes but remains appreciable, the situation is more complicated. In that case, there are generally multiple values of s that produce the augmented filter phase (modulo 2π). Alternative values of $s(\omega)$ produce filters of differing amplitudes without their having a differing augmented phase. The argument of the complex integral for E_s is then (probably) changed, although the argument of the integrand is everywhere left unchanged. An *ad hoc* method of selecting among candidate values of $s(\omega)$ is as follows: For any solution Θ, choose the value of $s(\omega)$ satisfying the augmented phase condition [Eq. (15)] that also maximizes $f(s)$. We choose that value of $s(\omega)$ since it puts the most energy through the filter. Now we can calculate I_s for the solution Θ, and we select the solution Θ that has the largest value of I_s. As a proof this is less than rigorous. It is based on the idea that if some care is taken to even up the phases of the frequencies passed by the filter, it is a good idea to pass as much energy as possible at each frequency.

VII. Discussion

The maximum-intensity optimization criterion is not the most general; in particular, it does not necessarily maximize the SNR. The optimization of the cross-coupled filter in the presence of noise would be dependent on the relative spectral dependencies of noise and signal. Then the choice of $s(\cdot)$ would be the one with the largest ratio of the correlation signal intensity I_s to the noise intensity I_n that passes through the filter produced by $s(\cdot)$. That optimization criterion is thus expressed as a ratio of integrals, a less tractable problem than maximizing a single integral. Maximizing the correlation signal intensity, however, is certainly a step in the right direction. In later work we carry noise considerations further.

We explicitly assumed knowledge of the amplitude response $f(s)$ and the phase response $g(s)$ as functions of the control signal s. We require those functions' derivatives as well. (As a result this method does not readily extend to binary SLMs.) Knowing a function to high accuracy, however, does not guarantee knowledge of its derivative to equivalent accuracy. Direct measurement of the derivatives, as opposed to inference from the form of the functions themselves, is a difficult task. No SLM is known to the author to have been so completely characterized. If an optimum control value can be determined for a set of arbitrary reference signals (e.g., as by the method of Juday and Daiuto[3]), it may be possible to infer the derivatives by beginning with the direct measurements of $f(s)$ and $g(s)$. This method could become a tool in the characterization of SLMs.

VIII. Conclusions

We have developed the necessary condition for one criterion of optimizing an optical correlation filter realized with a one-parameter coupled phase and amplitude SLM filter. The optimum filter explicitly involves the amplitude spectrum of the pattern to be recognized. Details of the solution vary strongly with the form of the coupling. The phase-only filter is a special case. We have outlined solution methods for the filter, and we have indicated areas for further research on coupled filters optimized for signal-to-noise performance.

References

1. J. L. Horner, "SLM Induced Phase Distortion and Its Effect on Optical Correlation," Proc. Soc. Photo-Opt. Instrum. Eng. **754**, 157–161 (1987).
2. R. D. Juday, S. E. Monroe, Jr., and D. A. Gregory, "Optical Correlation with Phase Encoding and Phase Filtering," Proc. Soc. Photo-Opt. Instrum. Eng. **825**, 149–155 (1987).
3. R. D. Juday and B. J. Daiuto, "Relaxation Method of Compensation in an Optical Correlator," Opt. Eng. **26**, 1094–1101 (1987).
4. L. E. Franks, *Signal Theory* (Dowden & Culver, Stroudsburg, PA, 1981).
5. G. Arfken, *Mathematical Methods for Physicists* (Academic, New York, 1985).

Discussions with several individuals were very helpful. Among them are Tom Fischer of Texas A&M University and Joe Horner of Rome Air Development Center. The editors and an anonymous reviewer provided criticism that led to a clearer exposition. The work reported here relates to a memorandum of understanding between NASA and the U. S. Army Missile Command; MICOM's financial support is gratefully acknowledged.