REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-01-

0148

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
04-02-2001	Final Technical Report	12/16/96 - 12/15/99

4. TITLE AND SUBTITLE
Estimation-theoretic methods for unconventional imaging

5a. CONTRACT NUMBER
NA

5b. GRANT NUMBER
F49620-97-1-0053

5c. PROGRAM ELEMENT NUMBER
NA

6. AUTHOR(S)
Timothy J. Schulz

5d. PROJECT NUMBER
NA

5e. TASK NUMBER
2301/FS and 2304/ES

5f. WORK UNIT NUMBER
NA

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Michigan Technological Univ.
1400 Townsend Drive
Houghton, MI 49931

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

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

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMITTAL DTIC. THIS TECHNICAL REPORT
HAS BEEN REVIEWED AND IS APPROVED FOR PUBLIC RELEASE
LAW AFR 190-12. DISTRIBUTION IS UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The objective of this program was to perform basic research on the application of estimation-theoretic processing methods for unconventional imaging systems. Imaging applications involving both incoherent (passive) and coherent (active) illumination were investigated, and results are described in this report for i) lower bounds on the estimation accuracy for multi-channel phase retrieval; and ii) dual-channel signal recovery from auto- and cross-correlation data. As a result of i), insight is gained into the relative merits of a Hartmann sensor for wavefront estimation, and the surprising result is obtained that a conventional image encodes more information about a wavefront than does a Hartmann sensor. As a result of ii), a novel processing method is derived for recovering two object functions from measurements of their auto- and cross-correlation functions.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Timothy J. Schulz
unclassified	unclassified	unclassified	uncl	20	19b. TELEPHONE NUMBER *(include area code)* 906 487-2550

Final Technical Report

Estimation-theoretic methods for unconventional imaging

Sponsored by:

Air Force Office of Scientific Research

Grant F49620-97-1-0053/P02

Principle Investigator:

Timothy J. Schulz
Department of Electrical and Computer Engineering
Michigan Technological University

Contents

Research Objectives

As stated in the research proposal, the objective of this program was to perform basic research on the application of estimation-theoretic processing methods for unconventional imaging systems. Imaging applications involving both incoherent (passive) and coherent (active) illumination were to be investigated, and the four significant thrusts of the program were to be:

1. the development of optimal processing methods for identifying and correcting wave-front errors that arise due to a non-homogeneous propagation medium or sensor-platform instabilities;

2. the utilization of statistical priors and penalties in the image-estimation and system-identification procedures;

3. the utilization of novel and/or auxiliary measurements such as those supplied by a wave-front sensor or pupil mask; and

4. the development of improved methods for image formation from sparse pupil-plane arrays, and an investigation into the optimal number and placement of array elements for coherent active-illumination systems.

Because funding for this project was reduced from three to two years, most of the efforts were restricted to the first three objectives. Collaborative efforts were conducted with Air Force personnel at the Air Force Research Laboratory in Albuquerque, NM and the Air Force Maui Optical Station in Maui, HI.

Research Methods and Results

This section documents the research methods used and results obtained for an investigation of estimation-theoretic methods for unconventional imaging. The methods and results are summarized for two areas:

1. Estimation-theoretic bounds on the estimation accuracy for multi-channel phase retrieval; and

2. Dual-channel signal recovery from auto- and cross-correlation data.

Lower Bounds on Estimation Accuracy for Multi-Channel Phase Retrieval

Wavefield aberrations induced by optical defects in monolithic lenses or mirrors, alignment errors in segmented-aperture systems, or wave propagation through an inhomogeneous medium can severely limit the resolution of an optical imaging-system and result in significantly distorted imagery. Adaptive compensation or post-detection processing can be used to correct for the aberrations and produce undistorted imagery, but these procedures require accurate estimation of the wavefield aberrations – a phase-retrieval problem must be solved.

Hartmann sensors are commonly used to estimate wavefront or optical aberrations in an imaging system. These sensors operate on the principle that the wavefront aberrations are well-approximated by linear tilts over small spatial regions – a Hartmann sensor segments the wavefield into small spatial segments with a lenslet array, and, in principle, the spots formed by the array translate according to the local wavefront slope [1]. The local slope estimates are then processed to estimate the wavefront (or aberrations). The point-spread function from a conventional image can also be used to estimate the wavefront or aberrations, but this is done less frequently. This method was used out of necessity, however, to analyze the aberrations of the Hubble Space Telescope's infamous primary mirror [2]. Another method that can be applied to the wavefront or optical-aberration estimation problem is phase diversity, whereby part of the light from a conventional image is split into a diversity channel where an additional aberration is introduced. This aberration is frequently induced by collecting the diversity data out of the focal plane [3].

Performance comparisons for the wavefront- and aberration-estimation systems mentioned above are typically linked to specific processing schemes. In contrast with this approach, our goal in this project is to provide the fundamental limits on wavefront- or aberration-estimation accuracy that are imposed by the use of various sensing modalities. To accomplish this, the Fisher information and Cramer-Rao bound are used to quantify the information content of various modalities, and simulation results are presented to compare

optimal estimation performance to the bounds. A surprising result of the analysis shows that data recorded by a conventional imaging system contains more information for parameter identification than do those recorded with a Hartmann sensor.

Mathematical Models

For quasi-monochromatic wave propagation through a relatively homogeneous medium, the diffraction-limited incoherent point-spread function for a conventional imaging system or Hartmann wavefront sensor is well-modeled as:

$$h(y) = \left| \int_{\mathcal{P}} T(u) s(y,u) du \right|^2, \tag{1}$$

where y and u are two-dimensional position variables in the sensor and pupil planes, respectively, $T(\cdot)$ is the pupil transmittance function, $\mathcal{P}$ is the pupil region, and $s(\cdot,\cdot)$ is the system transfer function for propagation from the pupil plane to the sensor plane. For the case of an in-focus imaging system in the Fresnel approximation, the pupil transmittance function is modeled as

$$T(u) = A(u) \exp\left(-j\frac{\pi}{\lambda f}|u|^2\right), \tag{2}$$

and the system propagation function is modeled as

$$s(y,u) = \frac{\exp\left(j\frac{2\pi}{\lambda}d_i\right)}{j\lambda d_i} \exp\left(j\frac{\pi|y-u|^2}{\lambda d_i}\right), \tag{3}$$

where λ is the nominal wavelength of the detected radiation, f is the system focal length, $A(\cdot)$ is the system aperture or pupil function, and d_i is the distance from the pupil plane to sensor plane. For a P-element Hartmann type lenslet array, the pupil transmittance function can be modeled as

$$T(u) = \sum_{p=1}^{P} A_l(u-u_p) \exp\left(-j\frac{\pi}{\lambda f_l}|u-u_p|^2\right), \tag{4}$$

where u_p denotes the location of the pth lenslet in the array, $A_l(\cdot)$ is the lenslet aperture function, and f_l is the lenslet focal length. In this situation the Fresnel approximation for propagation from the lenslet plane to the sensor plane may not be adequate, and the more general Huygens-Fresnel propagation function:

$$s(y,u) = \frac{1}{j\lambda} \frac{\exp(j\frac{2\pi}{\lambda}\sqrt{|y-u|^2+d_l^2})}{\sqrt{|y-u|^2+d_l^2}}, \tag{5}$$

should be used, where d_l is the distance from the lenslet plane to the sensor plane.

For situations involving optical aberrations in monolithic lenses or mirrors, alignment errors in segmented systems, or wave propagation through an inhomogeneous medium, the system point-spread function is often modeled with a single phase-screen aberration:

$$h(y;\theta) = \left| \int_{\mathcal{P}} T(u) \exp\left[j\theta(u)\right] s(y,u) du \right|^2, \tag{6}$$

where $\theta(\cdot)$ denotes the phase-error (in radians) caused by the system aberration, and $T(\cdot)$ is the abberation-free pupil transmittance function. The strength of an aberration is often quantified through the average of its square over the pupil region $\mathcal{P}$:

$$\epsilon^2 = \frac{\int_{\mathcal{P}} \theta^2(u) du}{\int_{\mathcal{P}} du}, \tag{7}$$

but care must be used when applying this measure because constant offsets (piston) and linear terms (tilt) will increase ϵ^2, but will not affect the quality of a system's imagery.

For systems that divide the incoming light into multiple channels – such as phase diversity or the simultaneous collection of a conventional image and a Hartmann wavefront-sensor image – the aberrated point-spread functions for each channel can be modeled as:

$$h_k(y;\theta) = \alpha_k \left| \int_{\mathcal{P}} T_k(u) \exp\left[j\theta(u)\right] s_k(y,u) du \right|^2, \; k = 1, 2, \ldots, K, \tag{8}$$

where the channel point-spread functions are indexed by k and each channel is characterized by its pupil transmittance function $T_k(\cdot)$ and propagation kernel $s_k(\cdot,\cdot)$. The scale factor α_k accounts for both the exposure time and the percentage of received light that is diverted to the kth channel. For situations of interest here, the transmittance functions for each channel $T_k(\cdot)$ can be normalized so that $h_k(y;\theta)$ integrates to α_k. Then α_k is a parameter that determines the expected number of photocounts (light level) in the kth channel.

In many applications and analyses it is common to expand the wavefront error in terms of an appropriate set of basis functions:

$$\theta(u;\boldsymbol{a}) = \sum_n a_n \phi_n(u), \tag{9}$$

with the basis set $\{\phi_n(\cdot)\}$ selected in a manner that allows for accurate modeling of the aberration function [1]. When the basis set is orthogonal over the system pupil with

$$\frac{\int_{\mathcal{P}} \phi_n(u)\phi_m(u) du}{\int_{\mathcal{P}} du} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}, \tag{10}$$

then

$$\begin{aligned}\epsilon^2 &= \frac{\int_{\mathcal{P}} \theta^2(u)du}{\int_{\mathcal{P}} du} \\ &= \frac{\sum_n \sum_m a_n a_m \int_{\mathcal{P}} \phi_n(u)\phi_m(u)du}{\int_{\mathcal{P}} du} \\ &= \sum_n a_n^2, \end{aligned} \tag{11}$$

and the average square aberration strength is equal to the sum of the squares of the coefficients $\{a_n\}$.

The Zernike polynomials are probably the most commonly-used basis functions, but other choices are possible. When an appropriate statistical model is available for the wavefront errors, for instance, the Karhunen-Loeve basis functions can be used. Whereas the Karhunen-Loeve expansion has the advantage of being optimal in the sense of compressing aberration energy into low-order terms, the Zernike functions have a convenient mapping to familiar aberration modes such as defocus, tilt, coma, etc. For thin-screen atmospheric turbulence models, however, the compression properties of the Zernike polynomials are very close to those of the optimal Karhunen-Loeve functions [1]. Regardless of which basis expansion is used, the point-spread function for the kth channel can be modeled as:

$$h_k(y;\boldsymbol{a}) = \alpha_k \left| \int_{\mathcal{P}} T_k(u) \exp\left[j\theta(u;\boldsymbol{a})\right] s_k(y,u)du \right|^2, \tag{12}$$

where the notation $h_k(y;\boldsymbol{a})$ now shows the explicit dependence of the point-spread function on the expansion coefficients.

In the remainder of this section we consider lower bounds on the accuracy with which the aberration parameters $\boldsymbol{a}$ can be estimated from noisy measurements of the point-spread functions $\{h_k(y;\boldsymbol{a})\}$. As a result of this study, insight is gained into the following important question:

> *Which modality – conventional image or Hartmann wavefront sensor image – allows for better accuracy when estimating the aberration parameters?*

This question is answered through a detailed analysis of the Cramer-Rao lower bound on estimation accuracy.

Noise Models and the Cramer-Rao Lower Bound

We consider a general scenario in which light from a known point-source is gathered by an imaging instrument and split into K channels. The channels might consist of a conventional imaging channel, a Hartmann wavefront sensor channel, a phase diversity channel, or any combination of these. The detection of optical radiation in each channel is fundamentally a random or stochastic process, and the semi-classical theory of photo-detection provides one means for modeling this randomness [4]. According to the theory, the best possible detector of optical radiation will record photo-events that can be modeled as samples from a conditional Poisson process with a rate function that is proportional to the image-plane intensity. Accordingly, the photo-event locations can be described by the two-dimensional Poisson counting process $\{N_k(y),\ y = (y_1, y_2) \in \Re^2\}$, where $N_k(y)$ denotes the number of photo-events that occur over the spatial region $\{y' : y' \leq y\}$ in the kth frame. The expectation of $N_k(y)$ conditional on $h_k(y; \boldsymbol{a})$ is then:

$$E\left[N_k(y)|h_k(y;\boldsymbol{a})\right] = \int_{y' \leq y} h_k(y';\boldsymbol{a})dy'. \tag{13}$$

Detectors of finite size will integrate these counts over spatial regions to form the data:

$$d_k[p] = \int w_{kp}(y)dN_k(y), \tag{14}$$

where $w_{kp}(y)$ is a normalized (to unit integral) weighting function that models the spatial region-of-integration for the pth detector element in the kth channel, and $dN_k(y)$ can be loosely interpreted as the number of photo-events occurring in the square region defined by y and $y + dy$. This Riemann-Stiltjes integral is interpreted as

$$\int w_{kp}(y)dN_k(y) = \begin{cases} 0, & \mathcal{N}_k = 0 \\ \sum_{l=1}^{\mathcal{N}_k} w_{kn}(y_{kl}), & \mathcal{N}_k > 0 \end{cases}, \tag{15}$$

where

$$\mathcal{N}_k = \int dN_k(y) \tag{16}$$

is the total number of photo-events occuring in the kth channel, and $\{y_{kl}\}$ are the spatial locations for photo-events in the kth channel. The conditional expectation of these data is

$$E\left\{d_k[p]|h_k(y;\boldsymbol{a})\right\} = \int w_{kp}(y)h_k(y;\boldsymbol{a})dy, \tag{17}$$

and, because the detector regions are non-overlapping, the data from different detector elements are statistically independent. Because our goal is to quantify the fundamental ad-

vantages of various sensor modalities for the estimation of wavefront aberration parameters, we consider only the effects of photon noise in our subsequent analysis. If desired, other noise effects such as read noise for charge-coupled device (CCD) cameras can be included to characterize performance limits for specific cameras.

The collected data for all detector elements of all channels are denoted by $\mathcal{D} = \{d_k\}_{k=1}^{K}$, and, consistent with the semiclassical theory for the detection of radiation, the conditional probability mass function for these data is modeled as:

$$\begin{aligned} P(\mathcal{D};\boldsymbol{a}) &= P(d_k;\boldsymbol{a}) \\ &= \prod_{k=1}^{K}\prod_{p} \exp(-h_k[p;\boldsymbol{a}])\,(h_k[p;\boldsymbol{a}])^{d_k[n]}/d_k[n]!, \end{aligned} \tag{18}$$

where

$$h_k[p;\boldsymbol{a}] = \int w_{kp}(y) h_k(y;\boldsymbol{a}) dy. \tag{19}$$

Cramer-Rao Lower Bound for the Estimation of Deterministic Parameters

The Cramer-Rao bound is an information-theoretic limit on the estimation accuracy one can expect when estimating parameters from random measurements [5]. Fundamental to this bound is the Fisher information which quantifies the parameter-specific information contained in the measurements. An important property of the Fisher information is additivity for independent measurements. That is, the Fisher information for the collection of data $\mathcal{D}$ is equal to the sum of the Fisher information for the data from each channel:

$$\boldsymbol{J}(\boldsymbol{a}) = \sum_{k=1}^{K} \boldsymbol{J}_k(\boldsymbol{a}), \tag{20}$$

where $\boldsymbol{J}_k(\boldsymbol{a})$ is an $N \times N$ matrix defined according to

$$[\boldsymbol{J}_k(\boldsymbol{a})]_{nm} = -E\left[\frac{\partial^2 \ln P(d_k;\boldsymbol{a})}{\partial a_n \partial a_m}\right]. \tag{21}$$

Consistent with the Poisson photocount model, the Fisher information for each channel is evaluated as:

$$[\boldsymbol{J}_k(\boldsymbol{a})]_{nm} = \sum_{p} \frac{\partial h_k[p;\boldsymbol{a}]}{\partial a_n} \frac{\partial h_k[p;\boldsymbol{a}]}{\partial a_m} \frac{1}{h_k[p;\boldsymbol{a}]}, \tag{22}$$

where

$$\frac{\partial h_k[p;\boldsymbol{a}]}{\partial a_n} = -2\alpha_k \mathrm{Im}\left\{\int w_{kp}(y) \int_{\mathcal{P}} \phi_n(u) T_k(u) \exp\left(j\sum_n a_n \phi_n(u)\right) s_k(y,u) g_k^*(y;\boldsymbol{a}) du dy\right\}, \tag{23}$$

with

$$g_k(y;\boldsymbol{a}) = \int_{\mathcal{P}} T_k(u) \exp\left(j \sum_n a_n \phi_n(u)\right) s_k(y,u) du. \tag{24}$$

Because both $h_k[p;\boldsymbol{a}]$ and its derivative scale linearly with α_k, the Fisher information also scales linearly with this parameter.

The Cramer-Rao lower bound makes use of the Fisher information to formulate the following bound on the error covariance for any unbiased estimator of the unknown parameter vector $\boldsymbol{a}$:

$$E\left[(\hat{\boldsymbol{a}} - \boldsymbol{a})(\hat{\boldsymbol{a}} - \boldsymbol{a})^T\right] \geq [\boldsymbol{J}(a)]^{-1}, \tag{25}$$

where the inequality implies that the difference between the left-hand matrix and the right-hand matrix is non-negative definite. As a consequence of the Cramer-Rao bound,

$$E\left[\sum_n (\hat{a}_n - a_n)^2\right] \geq \text{trace}\left\{[\boldsymbol{J}(\boldsymbol{a})]^{-1}\right\}, \tag{26}$$

so that the Fisher information can be used to place a lower bound the expected value of the residual aberration strength.

Example

In this section, we present an example in which performance bounds are computed for the estimation of wavefront-error parameters from a conventional image and from a Hartmann sensor. For this study, the Hartmann array contained approximately 8 subapertures across the system pupil as shown in Figure 1. Lower bounds on the estimation error for the estimation of the first 30 coeficients in the Zernike expansion of the wavefront were then computed for a conventional imaging system and a Hartmann sensor system. Because the Zernike coefficient vector $\boldsymbol{a}$ is a random parameter for turbulence-induced aberrations, the lower bound on the residual aberration is also random. Accordingly, we compute the bound for many realizations of the turbulence parameters for various ratios of aperture diameter D to seeing parameter r_0, where r_0 quantifies the turbulence strength [1]. The resulting bounds are averaged over 200 realizations and plotted in Figure 2. The bound values are normalized by the expected number of detected photons, so that the bound for a particular situation can be obtained by dividing our bound by the expected number of photons. Whereas this bound is computed for the estimation of only the first 30 coefficients, emperical analysis has shown the bound to be independent of the number of coefficients that are actually present in the wavefront aberrations. That is, the estimation for the first 30 coefficients when 40 coefficients are actually present results in the same bound on performance. In addition, maximum-likelihood estimation of the parameters produces error variances very close to

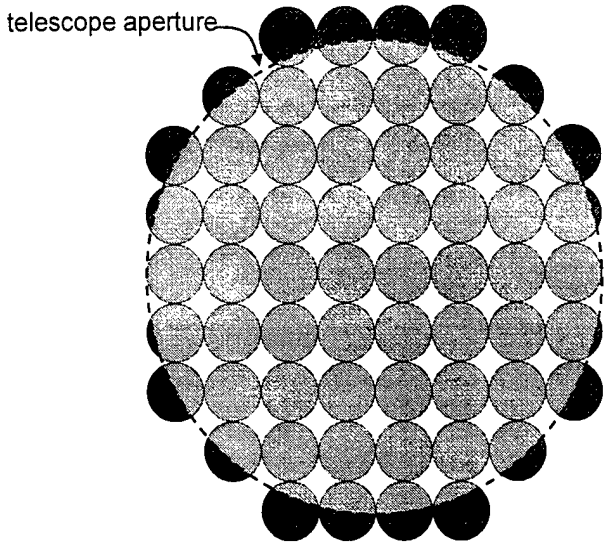

Figure 1: Lenslet array used to characterize the performance of a Hartmann array for phase-error estimation. The telescope aperture is seen to contain roughly 8 lenslet apertures across its diameter.

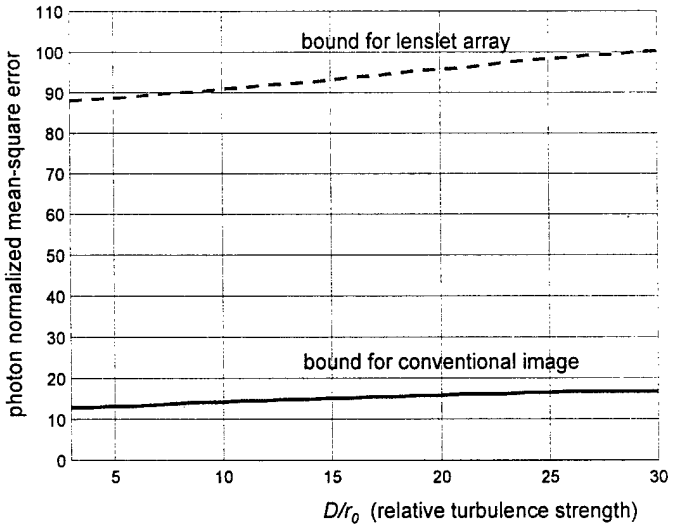

Figure 2: Cramer-Rao lower bounds on the estimation of the first 30 Zernike coefficients for turbulence-induced wavefront aperrations. The bounds are normalized by the expected number of detected photons.

these bounds.

Impact on the Air Force

The methods and results presented in this section have a direct impact on the design of future sensing systems for Air Force space surveillance. Collaborations on and dissemenation of this work has been facilitated through Bruce Stribling at the Air Force Maui Optical Station in Maui.

Dual-Channel Signal Recovery from Auto- and Cross-Correlation Data

In this section we consider the inverse problem of recovering two non-negative signals from measurements of their auto- and cross-correlation functions. Problems that are closely related to this include phase retrieval, wherein an object function must be recovered from a measurement of only its autocorrelation; and blind deconvolution, wherein two object functions must be recovered from a measurement of their convolution (or cross-correlation). Whereas these problems have, in general, been cosidered separately, common themes are seen in the many solutions that have been proposed. In almost all cases, for instance, the solutions are iterative and require the utilization of auxilliary information about the range or domain of the object function.

The general problem we address can arise in a variety of applications ranging from astronomical and space-object imaging to synthetic aperture radar (SAR) surveillance. We propose and discuss an iterative solution wherein a sequence of object estimates is produced having the desirable properties that each estimate in the sequence is non-negative, and that a measure of discrepancy (I-divergence) between the measured correlations and the estimated correlations decreases monotonically with the number of iterations.

Specifically, if the unknown signals are denoted as $o_1(\cdot)$ and $o_2(\cdot)$, then the auto- and cross-correlation functions for these signals are denoted by:

$$r_{11}(y) = \sum_x o_1(x)o_1(x-y), \tag{27}$$

$$r_{22}(y) = \sum_x o_2(x)o_2(x-y), \tag{28}$$

and

$$r_{12}(y) = \sum_x o_1(x)o_2(x-y), \tag{29}$$

where r_{11} is the autocorrelation for o_1, r_{22} is the autocorrelation for o_2, r_{12} is the crosscorrelation for o_1 and o_2, and we have restricted our attention to signals defined on discrete space

(or time).

The Signal Recovery Algorithm

We first state an iterative algorithm for the recovery of o_1 and o_2 from the noisy measurements of r_{11}, r_{22}, r_{12}:

$$o_1^{\text{new}}(x) = o_1^{\text{old}}(x) \frac{\sum_y \left[o_1^{\text{old}}(x-y) + o_1^{\text{old}}(x+y)\right] \frac{d_{11}(y)}{r_{11}^{\text{old}}(y)} + \sum_y o_2^{\text{old}}(x-y) \frac{d_{12}(y)}{r_{12}^{\text{old}}(y)}}{2O_1 + O_2}, \tag{30}$$

and

$$o_2^{\text{new}}(x) = o_2^{\text{old}}(x) \frac{\sum_y \left[o_2^{\text{old}}(x-y) + o_2^{\text{old}}(x+y)\right] \frac{d_{22}(y)}{r_{22}^{\text{old}}(y)} + \sum_y o_1^{\text{old}}(x+y) \frac{d_{12}(y)}{r_{12}^{\text{old}}(y)}}{2O_2 + O_1}, \tag{31}$$

where O_1 and O_2 satisfy

$$O_1 = \frac{2\sum_y d_{11}(y) + \sum_y d_{12}(y)}{2O_1 + O_2}, \tag{32}$$

and

$$O_2 = \frac{2\sum_y d_{22}(y) + \sum_y d_{12}(y)}{2O_2 + O_1}, \tag{33}$$

and the noisy measurments are of r_{11}, r_{22}, and r_{12}, are d_{11}, d_{22}, and d_{12}, respectively. As we will soon show, this algorithm has the desireable properties that o_1^{new} and o_2^{new} are always nonnegative functions (provided that the initial estimates are nonnegative), and that

$$D^{\text{new}} \leq D^{\text{old}}, \tag{34}$$

where D^{new} and D^{old} are the I-divergence discrepancy between the data and estimates (new or old) defined as:

$$\begin{aligned} D^{\text{new}} &= \sum_y \left\{ d_{11}(y) \ln \frac{d_{11}(y)}{r_{11}^{\text{new}}(y)} + [r_{11}^{\text{new}}(y) - d_{11}(y)] \right\} \\ &+ \sum_y \left\{ d_{22}(y) \ln \frac{d_{22}(y)}{r_{22}^{\text{new}}(y)} + [r_{22}^{\text{new}}(y) - d_{22}(y)] \right\} \\ &+ \sum_y \left\{ d_{12}(y) \ln \frac{d_{12}(y)}{r_{12}^{\text{new}}(y)} + [r_{12}^{\text{new}}(y) - d_{12}(y)] \right\}, \end{aligned} \tag{35}$$

where

$$r_{ij}^{\text{new}}(y) = \sum_x o_i^{\text{new}}(x) o_j^{\text{new}}(x-y). \tag{36}$$

To show that the I-divergence is non-increasing, we first note that the measure D is equivalent to the negative log-likelihood function for situations in which the data $\{d_{ij}(y)\}$ are independent Poisson variants, with the expected values $\{r_{ij}(y)\}$. Therefore, maximizing the log-likelihood for Poisson data is equivalent to minimizing D. Next, we show that the iterative algorithm is an instance of the expectation-maximization (EM) method, and as such the sequence of estimates has the property of non-decreasing likelihood – or non-increasing D [6].

EM Method for Poisson Data

Suppose the noisy data are Poisson deviates whose means are the actual auto- and cross-correlations. In this case, the log-likelihood is [7]:

$$\begin{aligned} L &= \sum_y \{d_{11}(y) \ln r_{11}(y) - r_{11}(y)\} \\ &+ \sum_y \{d_{22}(y) \ln r_{22}(y) - r_{22}(y)\} \\ &+ \sum_y \{d_{12}(y) \ln r_{12}(y) - r_{12}(y)\}. \end{aligned} \tag{37}$$

Therefore, maximizing L over o_1 and o_2 is equivalent to minimizing D. To use the EM method for this process, we begin by defining the *complete data* as:

$$\tilde{d}_{ij}(x,y) \sim \mathcal{P}\{o_i(x) o_j(x-y)\}, \tag{38}$$

where the notation is meant to imply that $\tilde{d}_{ij}(x,y)$ is a Poisson random variable with mean $o_i(x)o_j(x-y)$. The many-to-one mapping from the *complete data* to the measured or *incomplete data* is:

$$d_{ij}(y) = \sum_x \tilde{d}_{ij}(x,y). \tag{39}$$

The log-likelihood for these complete data is then:

$$\begin{aligned} L_{\text{cd}} &= \sum_{x,y} \left\{ \tilde{d}_{11}(x,y) \ln [o_1(x) o_1(x-y)] - o_1(x) o_1(x-y) \right\} \\ &+ \sum_{x,y} \left\{ \tilde{d}_{22}(x,y) \ln [o_2(x) o_2(x-y)] - o_2(x) o_2(x-y) \right\} \\ &+ \sum_{x,y} \left\{ \tilde{d}_{12}(x,y) \ln [o_1(x) o_2(x-y)] - o_1(x) o_2(x-y) \right\}. \end{aligned} \tag{40}$$

Following the EM method, we form the conditional **E**xpectation of $L_{\rm cd}$:

$$\begin{aligned}\bar{L}_{\rm cd} &= \sum_{x,y}\left\{\bar{\tilde{d}}_{11}(x,y)\ln\left[o_1(x)o_1(x-y)\right]-o_1(x)o_1(x-y)\right\}\\ &+ \sum_{x,y}\left\{\bar{\tilde{d}}_{22}(x,y)\ln\left[o_2(x)o_2(x-y)\right]-o_2(x)o_2(x-y)\right\}\\ &+ \sum_{x,y}\left\{\bar{\tilde{d}}_{12}(x,y)\ln\left[o_1(x)o_2(x-y)\right]-o_1(x)o_2(x-y)\right\},\end{aligned} \tag{41}$$

where $\bar{\tilde{d}}_{ij}(x,y)$ is the expectation of the complete data conditional on the measured (or incomplete) data, and assuming that the actual parameter values are the current estimates ($o_1^{\rm old}$ and $o_2^{\rm old}$). Next, for the **M**aximization step we select $o_1^{\rm new}$ and $o_2^{\rm new}$ to maximize $\bar{L}_{\rm cd}$. From the Kuhn-Tucker conditions, we find the following conditions for a maximizer:

$$o_1^{\rm new}(x)=\frac{\sum_y\left[\bar{\tilde{d}}_{11}(x,y)+\bar{\tilde{d}}_{11}(x+y,y)\right]+\sum_y\bar{\tilde{d}}_{12}(x,y)}{2\sum_z o_1^{\rm new}(z)+\sum_z o_2^{\rm new}(z)}, \tag{42}$$

and

$$o_2^{\rm new}(x)=\frac{\sum_y\left[\bar{\tilde{d}}_{22}(x,y)+\bar{\tilde{d}}_{22}(x+y,y)\right]+\sum_y\bar{\tilde{d}}_{12}(x+y,y)}{2\sum_z o_2^{\rm new}(z)+\sum_z o_1^{\rm new}(z)}. \tag{43}$$

These expressions are simplified by noting that:

$$\bar{\tilde{d}}_{ij}(x,y)=o_i^{\rm old}(x)o_j^{\rm old}(x-y)\frac{d_{ij}(y)}{r_{ij}^{\rm old}(y)}, \tag{44}$$

so that

$$o_1^{\rm new}(x)=o_1^{\rm old}(x)\frac{\sum_y\left[o_1^{\rm old}(x-y)+o_1^{\rm old}(x+y)\right]\frac{d_{11}(y)}{r_{11}^{\rm old}(y)}+\sum_y o_2^{\rm old}(x-y)\frac{d_{12}(y)}{r_{12}^{\rm old}(y)}}{2\sum_z o_1^{\rm new}(z)+\sum_z o_2^{\rm new}(z)}, \tag{45}$$

and

$$o_2^{\rm new}(x)=o_2^{\rm old}(x)\frac{\sum_y\left[o_2^{\rm old}(x-y)+o_2^{\rm old}(x+y)\right]\frac{d_{22}(y)}{r_{22}^{\rm old}(y)}+\sum_y o_1^{\rm old}(x+y)\frac{d_{12}(y)}{r_{12}^{\rm old}(y)}}{2\sum_z o_2^{\rm new}(z)+\sum_z o_1^{\rm new}(z)}. \tag{46}$$

Now, note that if we sum these expressions over x, we obtain the follwing equations:

$$O_1 = \frac{2\sum_y d_{11}(y) + \sum_y d_{12}(y)}{2O_1 + O_2}, \tag{47}$$

and

$$O_2 = \frac{2\sum_y d_{11}(y) + \sum_y d_{12}(y)}{2O_2 + O_1}, \tag{48}$$

where $O_i = \sum_x o_i^{\text{new}}(x)$ is independent of the iteration number. That is, the sum of the object estimates does not change after the first iteration, and this sum is determined by these equations. Therefore, these equations can be solved and the sums used for the resulting iterative equations:

$$o_1^{\text{new}}(x) = o_1^{\text{old}}(x) \frac{\sum_y \left[o_1^{\text{old}}(x-y) + o_1^{\text{old}}(x+y)\right] \frac{d_{11}(y)}{r_{11}^{\text{old}}(y)} + \sum_y o_2^{\text{old}}(x-y) \frac{d_{12}(y)}{r_{12}^{\text{old}}(y)}}{2O_1 + O_2}, \tag{49}$$

and

$$o_2^{\text{new}}(x) = o_2^{\text{old}}(x) \frac{\sum_y \left[o_2^{\text{old}}(x-y) + o_2^{\text{old}}(x+y)\right] \frac{d_{22}(y)}{r_{22}^{\text{old}}(y)} + \sum_y o_1^{\text{old}}(x+y) \frac{d_{12}(y)}{r_{12}^{\text{old}}(y)}}{2O_2 + O_1}. \tag{50}$$

Because the EM method provides a sequence of estimates that has the property of non-decreasing likelihood, it follows that the resulting algorithm also has the property of non-increasing I-divergence. Nonnegativity of the estimates follows naturally provides the initial object estimates are nonnegative.

Example

Shown in Figure 3 are two object function and their auto- and cross-correlation functions. The object estimates obtained by using the iterative method described in this report with random starting guesses are shown in Figure 4 for 1000 iterations.

Impact on the Air Force

The methods and results presented in this section have a direct impact on the processing of data recorded by sensing systems for Air Force space surveillance. Collaborations on and dissemenation of this work has been facilitated through David Voelz at the Air Force Research Laboratory in Albuquerque, NM.

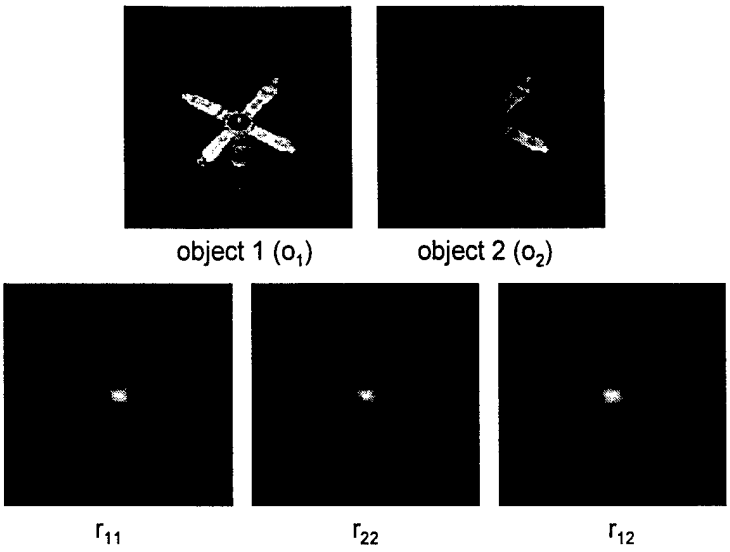

Figure 3: Two object functions, and their auto- and cross-correlation functions.

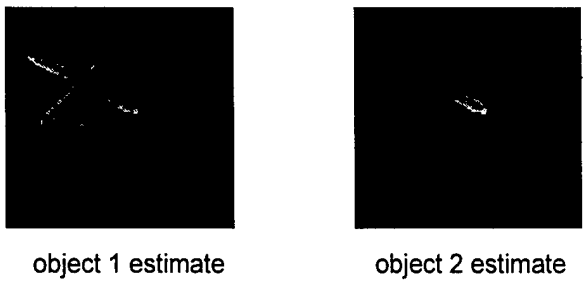

Figure 4: Object estimates after 1000 iterations with random starting guesses.

References

[1] M. C. Roggemann and B. Welsh. *Imaging Through Turbulence.* CRC Press, Inc., 1996.

[2] J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin. Hubble Space Telescope characterized by using phase-retrieval algorithms. *Appl. Opt.*, 32(10):1747–1767, 1993.

[3] R. G. Paxman, T. J. Schulz, and J. R. Fienup. Joint estimation of object and aberrations using phase diversity. *J. Opt. Soc. Am. A*, 9:1072–1085, 1992.

[4] J. W. Goodman. *Statistical Optics.* John Wiley & Sons, New York, 1985.

[5] B. Porat. *Digital Processing of Random Signals: Theory and Methods.* Prentice-Hall, 1993.

[6] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B*, 39:1–37, 1977. (with discussion).

[7] D. L. Snyder and M. I. Miller. *Random Point Processes in Time and Space.* Springer-Verlag, New York, 1991.

Publications

Publications, theses, and manuscripts in preparation or review from this project are documented in this section.

Journal Publications

"Algorithm to increase the largest aberration that can be reconstructed from Hartmann sensor measurements," M. C. Roggemann, and T. J. Schulz, *Appl. Opt.*,.vol. 37, p4321-4329, 1998.

"Joint processing of Hartmann sensor and conventional image measurements to estimate large aberrations: theory and experimental results", M.C. Roggemann, T.J. Schulz, C.W. Ngai, and J.T. Kraft, *Applied Optics*, vol. 38, no. 11, pp. 2249-2255, 1999.

"Multiframe blind deconvolution with real data: imagery of the Hubble Space Telescope", T.J. Schulz, B.E. Stribling, and J.J. Miller, *Optics Express*, vol. 1, no. 11, pp. 355-362, 1997.

"A measurement and data processing approach for estimating the spatial statistics of turbulence-induced index of refraction fluctuations in the upper atmosphere," W. W. Brown, M. C. Roggemann, T. J. Schulz, T. C. Havens, J. T. Beyer, and L. J. Otten, accepted for publication in Applied Optics, December, 2000.

Manuscripts Submitted or in Preparation

"Lower bounds on estimation accuracy for multi-channel phase retrieval", T.J. Schulz, W. Sun, and M. C. Roggemann.

"Dual-channel signal recovery from auto- and cross-correlation data", T.J. Schulz and Y. Zhang.

"Fine-resolution imaging with coherent arrays by the method of penalized maximum-likelihood", T.J. Schulz and F. Li.

Conference Proceedings

"Cramer-Rao bounds for estimation of turbulence-induced wave front aberrations," T. J. Schulz, Wei Sun, and M. C. Roggemann, SPIE paper number 3763-03, Denver, July, 1999.

"Joint processing of Hartmann sensor and conventional image measurements to estimate large aberrations," M. C. Roggemann, T. J. Schulz, C. W. Ngai, and J. T. Kraft, SPIE paper number 3815-21, Denver, July, 1999.

"Joint processing of image plane and wave front sensor information to measure large aberrations in optical systems," M. C. Roggemann, T. J. Schulz, B. M. Welsh, C. W. Ngai, and J. T. Kraft, Proc. SPIE on Novel Optical Systems and Large Aperture Imaging, vol. 3430, 1998.

"A fast algorithm for maximum-likelihood imaging with coherent speckle measurements", T.J. Schulz, Proc. of the 1997 IEEE International Conference on Image Processing, Santa Barbara, CA, pp. 679-682, October 1997.

"Nonlinear models and methods for space-object imaging through atmospheric turbulence," (**invited paper**), T.J. Schulz, 1997 IEEE Nonlinear Signal and Image Processing Workshop, Mackinac Island, Michigan, September, 1997.

Theses and Reports

Cramer-Rao lower bounds for multi-channel phase retrieval, Wei Sun, MS Thesis, Michigan Technological University, 1999.

Fast algorithms for coherent-array imaging using penalized maximum-likelihood estimation, Feng Li, MS Thesis, Michigan Technological University, 1999.

Blind deconvolution with an autocorrelation constraint, Pulsiri Ninkitsaranont, MS Report, Michigan Technological University, 2000.