Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Bayesian Statistical Approaches to Tracking through Turbulence

Ben Fitzpatrick
Department of Mathematics
Loyola Marymount University
and
Tempest Technologies
Los Angeles, CA
bfitzpatrick@lmu.edu

Sarah McCanless
Tempest Technologies
Los Angeles, CA
mccanless@tempest-tech.com

Yun Wang
Tempest Technologies
Los Angeles, CA
wang@tempest-tech.com

***Abstract*— In this paper we describe two Bayesian approaches for tracking through turbulence. The problem considered involves an extended target actively illuminated with several lasers. The returned imagery is used to infer atmospheric tilt. The main application of this technology is the control of a steering mirror which is used to point a laser weapon at the target. Several model-based approaches are examined, including an optical transfer function model and a tilted reference image model.**

I. INTRODUCTION

Proliferation of offensive missile weaponry has become a major concern in today's world. The situation requires the development of efficient and reliable ballistic missile defense systems. More and more countries have ballistic missile capability in the form of systems such as Scud missiles, which can deliver biological, chemical, or even nuclear warheads. Defensive systems against these threats are needed now more than ever. A defensive system of great potential is the Airborne Laser (ABL), which is an aircraft mounted, high-energy laser (HEL) designed to defeat offensive missiles in the boost phase. In order for such a system to be effective, one must be able to maintain a highly accurate track to point the weapon at the missile. The long paths of propagation, as well as the atmospheric fluctuations induced by turbulence, provide major challenges to control systems design.

Compensation for turbulence in laser system design is a fundamentally different kind of "noisy environment" from traditional tracking and control problems, primarily due to the fact that the path the laser follows is distorted by the atmosphere. This distortion, caused by turbulence-induced variation in the index of refraction of the atmosphere, accumulates over the entire travel path of the light, and one would have to know (or estimate) the index of refraction to be able to remove its effect exactly. Like a putt on an uneven green, the steering problem requires pointing in a direction that will eventually bring the light to the right spot on the target after traversing the atmosphere. Unlike the putt, however, we are not allowed to examine the entire path, but only a reduced picture that has essentially compressed the information into a temporal sequence of images at our platform.

In this paper, we explore some model-based tracking algorithms. Our approach relies on Bayesian statistical techniques to extract the underlying atmospheric tilt information from the imagery. The basic relationships center on Kolmogorov's statistical theory of turbulence, the optical transfer function model of imaging, and Bayesian statistical methods of parameter estimation.

II. MODEL BACKGROUND

In this section, we recall the basic ideas of imaging through turbulence. The book by Roggemann and Welsh [1] provides a detailed description of what we briefly outline here.

A. Turbulence Modeling

Propagation of light waves through turbulence produces distortions in both the phase and the amplitude. The problem of tracking is to determine the "linear" component of the phase, referred to as the "tilt." The tilt corresponds to the angle at which we should point so that the outgoing light ends up in the right place. The fundamental relationship for light propagation gives phase shift as a function of the index of refraction perturbations

$$\psi(x,y,t) = \frac{2\pi}{\lambda}\int_P n(x,y,z,t)\,ds$$

in which λ is the wavelength of the light, ψ denotes the phase at the (two dimensional) receiving aperture, n denotes the three dimensional index of refraction of the atmosphere, and P denotes the path the light travels. The integral is a line integral computed along the path traversed by the light. We allow time dependence here, which we will discuss in more detail below. Kolmogorov's theory of turbulence poses a power law structure function model for the index of refraction, given by

$$\Gamma_n(h) = C_n^2|h|^{\frac{2}{3}} = E|n(r+h) - n(r)|^2,$$

in which r denotes the three dimensional position, h denotes the spatial separation in three dimensions, and C_n^2 is a proportionality constant that measures the strength of the turbulence. Using the phase relationship with refractive index and Kolmogorov's model, it is possible to determine the structure function for the phase:

$$\Gamma_\psi(h) = C_\psi^2|h|^{\frac{5}{3}},$$

in which the proportionality constant and the power law exponent are different from that for the index of refraction.

0-7803-8335-4/04/$17.00 ©2004 AACC

The interested reader should consult [1] for a detailed derivation. It is also typically assumed that the phase statistics are Gaussian.

B. Imaging through Turbulence

The process of imaging is generally modeled through a point spread function (PSF) applied to an underlying object. We begin with this standard approach, in which the data z_t at time t arises from a linear operator applied to the reference object: $z_t = A_t(u)$. The operator A_t is a convolution operator, whose kernel is the PSF h_t:

$$z_t(r) = \int h_t(r-s)u(s)\,ds. \tag{1}$$

The turbulence model typically enters the imaging model through the PSF. The specific formulation (as is seen, e.g., in [1]) is often given in terms of the optical transfer function (OTF), which is the Fourier transform of the PSF. The most commonly used form of the OTF is

$$\begin{aligned} H(f) &= \textstyle\int W(r)W(r-\lambda df) \\ &\quad * \exp\{i\psi(r) - i\psi(r-\lambda df)\}\,dr, \end{aligned} \tag{2}$$

in which W is the optical system's pupil weighting function, d is the distance from the pupil to the imaging plane. Thus, the image depends on the phase perturbation in a highly nonlinear way. Even with its apparent complexity, this model is a major simplification of the scenario of interest in the tracking problem, in the following ways. The target is an extended source, not a point source, so that the operator A_t should be modeled by a more general kernel integration and not a convolution. Also, one expects scintillation, or amplitude fluctuations in the light wave, which is not modeled here. Amplitude fluctuations are modeled by an additional factor in the OTF, which is of the form

$$\begin{aligned} H(f) &= \textstyle\int W(r)W(r-\lambda df)\exp\{\xi(r)+\xi(r-\lambda df)\} \\ &\quad \exp\{i\psi(r) - i\psi(r-\lambda df)\}\,dr, \end{aligned} \tag{3}$$

in which ξ denotes the amplitude perturbation at the pupil. We will return to these more accurate but challenging problems, but for the moment we will continue with the simplified model equations (1) and (2) above.

In the tracking problem, the goal is to estimate the line-of-sight (LOS) angle to the target. If the LOS error is 0, then the target is "in the middle" of the image plane. Phase fluctuations, however, can make the target appear to be off-center; in turn, the LOS is apparently in error. To see how this problem occurs, we consider the following very simple case.

III. CORRELATION TRACKING

Suppose that the optical system is of infinite aperture size, with $W = 1$. Suppose further that the phase shift satisfies

$$\psi(x,y) = \psi_0 + (x,y)\cdot\theta,$$

in which ψ_0 is the constant (piston) term, and θ represents the atmospheric tilt. The OTF then becomes

$$H(f) = \exp\{f\cdot\theta\lambda d\},$$

leading to a delta function PSF: $h(x) = \delta_{\theta\lambda d}(x)$. The essential effect of this simple phase function is a translation in the image plane of the target image. It is this shift, $X = \theta\lambda d$, that one seeks to "track" with conventional tracking algorithms. Indeed, this idealization suggests the approach taken in the correlation algorithm.

The most commonly used techniques for estimating tilts are those based on the centroid and the correlation tracker. The centroid tracker, true to its name, computes X by finding the centroid of the observed image z. The feature tracked is in this case the centroid of the true image u. Adjustments must be made due to the fact that the target, in most scenarios of interest, extends outside the image plane. The correlation tracker is developed by posing the estimation problem as a least squares minimization. One estimates by minimizing

$$J(X) = \int |z_t(x) - u(x-X)|^2\,dx.$$

The reason this technique goes by the name "correlation tracker" is that the minimizer can be obtained directly from the peak of the cross-correlation between z and u. Fitts, however, in [2], derived an efficient computational algorithm that does not require explicit maximization of the cross-correlation. Instead, by linearizing the correlation computation, the shifts are determined from a pair of coupled linear equations involving various image integrations.

Again, we emphasize that this algorithm is based on a fairly simple model of the observation process:

$$z_t(x) = u(x-X_t) + \varepsilon_t(x),$$

in which ε is white noise (in t and x). These assumptions are mathematically equivalent to assuming the incoming phase comprises a piston mode (the ψ_0 term, which drops out in the differencing) and tilt modes (the θ vector), plus small fluctuations which are independent of these "main modes." We also remark that the correlation method assumes the reference image function u is known or can be estimated reliably.

IV. A FIRST APPROACH TO BAYESIAN TRACKING

While the correlation algorithm admits fast, efficient real-time implementation, there are several major drawbacks. One, as noted above, is the simplistic modeling of the phase perturbations. Another is the lack of use of previous frame/estimator information. It is our goal here to propose and develop some Bayesian statistical methods in order to use more information, concerning both the nature of the phase and the previous frame data.

A. Bayesian Statistical Inference

The basic idea of Bayesian analysis is to perform inference by determining the conditional distribution of the quantity of interest, given the observations and data at hand. This statement is made quantitative in the following way.

One begins with prior distribution $\pi(\phi)$ on the parameter ϕ of interest. The measurement is modeled with a probability distribution, $p(z|\phi)$, which is the conditional distribution of the data, given the parameter. Bayes theorem is then applied to compute the posterior distribution, which is the conditional distribution of the parameter given the data. One typically estimates parameter values in this paradigm by computing the parameter value which maximizes the posterior density function. Often called MAP (for maximum a posteriori) estimation, this technique is one way of interpreting the statistical framework of Kalman filtering. The basic formula for the posterior distribution of the parameter given the data is

$$\pi(\phi|z) = \frac{p(z|\phi)\pi(\phi)}{\int_\Phi p(z|\phi')\pi(\phi')\,d\phi'},$$

It should be noted here that the demoninator is a normalizing constant (which is independent of the parameter), so that MAP estimation merely maximizes the numerator.

A common connection between estimation methods involves the case in which the observation density is Gaussian. For example, suppose the data is modeled by

$$z = A\phi + \varepsilon,$$

where $\varepsilon \sim N(0, \sigma^2 I)$. Suppose also that we use a prior from the exponential family: $\pi(\phi) \propto \exp\{-N(\phi)\}$.

In this case, MAP estimation is equivalent to minimizing the penalized least squares cost

$$J(\theta) = |z - A\theta|^2 + N(\theta).$$

In the case of a uniform prior, N is constant, and MAP, maximum likelihood, and least squares all produce the same result. In the case of a nonlinear model such as the OTF, we have a nonlinear problem to solve. Below we describe the structure of the phase estimation problem.

B. Bayesian Phase Estimation

In order to incorporate a slightly better turbulence model into the correlation tracking scheme, we pose the following nonlinear least squares problem. We seek to minimize, at each time step,

$$J(\phi) = \int |z_t(x) - h(x,\psi) * u(x)|^2\,dx,$$

in which

$$\psi(x) = \sum_k \phi^k B^k(x)$$

is an expansion of the phase in terms of modes B^k. The most common approximation scheme is the Zernike polynomial expansion. The details of the Zernike ploynomials can be found in [1]. The coefficients ϕ form the parameter vector to be estimated. Note that, by the Plancheral identity, one can equally well formulate this cost functional in the frequency domain:

$$J(\phi) = \int |Z_t(f) - H(f,\psi)U(f)|^2\,df,$$

One can implement any number of iterative methods from stochastic approximation to estimate the phase coefficients ϕ. Of course, this standard least squares approach does not take advantage of the statistical information we have concerning the phase.

The Bayesian form would include the Kolmogorov turbulence model into the observation process, leading to a penalized cost functional of the form

$$J(\phi) = \int |Z_t(f) - H(f,\psi)U(f)|^2\,df + \phi^T C^{-1}\phi$$

in which C is the covariance matrix derived from the structure function for the phase given above.

C. Temporal Filtering

The motion of the target means that the image sequence is temporally correlated. An algorithm for estimating the phase for each frame should integrate this additional information. The type of temporal correlation we expect uses Taylor's frozen flow hypothesis. That model treats the phase sequence as

$$\psi_t(x) = \psi(x - \nu t),$$

in which ν denotes the translational velocity of the target. Then the covariance of $(\psi_t, \psi_{t+h}, \cdots, \psi_{t+kh})$ can be determined directly from the covariance of the process ψ: the temporal phases ψ_t are just spatial samples of ψ in terms of translations in space. Likewise, projecting the phase onto the basis functions B^k, we obtain a covariance matrix for the coefficients ϕ_t^k. Performing a regression analysis to obtain the conditional distribution of ϕ_{t+1} given a past history $\phi_t, \phi_{t-1}, \cdots, \phi_{t-p}$, we develop the phase dynamics with a linear statistical model of the form

$$\phi_{t+1} = \sum_{\ell=0}^{p} A_\ell \phi_{t-\ell} + \eta_t,$$

in which the A_ℓ matrices are determined through the covariance (see, e.g., [3]). For example, with only one step of history, we would have that the covariance

$$E\Big[\phi_t(\phi_t)^T\Big] = E\Big[\phi_{t+1}(\phi_{t+1})^T\Big] = \Sigma_{11},$$

and

$$E\Big[\phi_{t+1}(\phi_t)^T\Big] = E\Big[\phi_{t+1}(\phi_{t+1})^T\Big]^T = \Sigma_{12},$$

leading to the regression equation

$$\phi_{t+1} = \Sigma_{12}\Sigma_{11}^{-1}\phi_t + \eta_t,$$

for some zero mean η. In order to approximate the phase behavior properly, we must take enough terms in the expansion so that the residual term η behaves as white noise.

The measurement process here depends nonlinearly on the phase, even though the regression modeling of the temporal phase correlations produces a linear dynamical system. The computational complexity of the nonlinear filter leads us to consider a more empirically based approach, which is the subject of the next section.

V. A SECOND APPROACH TO BAYESIAN TRACKING

The problem of tilt estimation through nonlinear filtering, with the OTF providing the measurement model, is extremely challenging for frame rates of interest in the ABL system. Thus, we are led to consider a set of filters having a simpler and more efficient implementation. Toward that end, we pose a measurement model of the form

$$y_t(x) = w(x - X_t) + n_t(x),$$

in which $y_t = \log(z_t)$, w is a reference image, X_t is the two dimensional tilt, and n_t is the measurement noise. The multiplicative nature of the scintillation noise leads us to consider a logarithmic model here. Now, the tilt, which as above is modeled as the linear component of the phase, is nearly independent, stochastically, of the rest of the Zernike coefficients of the phase (see [1]), so we consider the model

$$\begin{aligned} X_{t+1} &= X_t + \textstyle\sum_{k=0}^{p} \hat{A}_k X_{t-k} + \delta_t \\ &= X_t + v_t + \delta_t \end{aligned}$$

for the temporal dynamics of the tilt, explicitly separating out the term X_t so that we can include a velocity step. The coefficients $\hat{A}_k$ and the covariance of δ are determined through the regression process described above. Treating δ_t as an uncorrelated sequence is of course an approximation. We plan to examine colored noise problems in future work.

Examining the reference image's behavior under the temporal shift, we expand through the linear term of the Taylor approximation to obtain

$$w(x - X_{t+1}) \approx w(x - X_t) + \nabla w(x - X_t) \cdot (v_t + \delta_t).$$

Thus, we rewrite our observation problem as

$$y_{t+1}(x) = y_t(x) + \nabla w(x - X_t) \cdot (v_t + \delta_t) + \hat{n}_t(x),$$

in which the noise term $\hat{n}$ now is the difference of the original noise terms.

Note that the problem of estimating X_{t+1} becomes a problem of estimating v from this measurement equation. This problem is linear in v, but the dependence on the current estimate of the tilt is nonlinear.

The estimation algorithm we examine is based on a Bayesian formulation. We have a parameter dependent measurement equation, and we have a statistical model of the parameter's behavior. That is, we have that

$$v_t = \sum_{k=0}^{p} \hat{A}_k X_{t-k},$$

and that

$$\begin{aligned} y_{t+1}(x) &= y_t(x) + \nabla w(x - X_t) \cdot (v_t + \delta_t) + \hat{n}_t(x) \\ &= y_t(x) + \nabla w(x - X_t) \cdot v_t + \tilde{n}_t(x) \end{aligned}$$

in which the δ term has been absorbed into the noise term $\tilde{n}$. We assume that the measurement error $\hat{n}$ has the Gaussian distribution with zero mean and covariance R. We also assume that the images are observed as pixel values of intensity, so that the measurement is actually finite dimensional. For notational purposes, we define $Y_t = (y_1, y_2, \ldots, y_t)$ to be the complete observation history vector. Thus, the measurement distribution $p(y_{t+1}|v; Y_t)$ is given by

$$\begin{aligned} &p(y_{t+1}|v; Y_t; X_t) = \\ &\frac{1}{K} \exp\Big\{ -\frac{1}{2} L(y_{t+1}, y_t, v, X_t)^T R^{-1} L(y_{t+1}, y_t, v, X_t) \Big\}, \end{aligned}$$

in which K is the usual Gaussian normalizing constant

$$K = \sqrt{\det(R)} (2\pi)^{N^2/2},$$

and the exponent is

$$L(y_{t+1}, y_t, v, X_t) = y_{t+1} - y_t - \nabla w(x - X_t) \cdot v.$$

It should be noted that the covariance R is dependent on X_t through the inclusion of the δ term above. The"corrector" step of the Bayesian inference is the posterior distribution computation

$$\pi(v|Y_{t+1}) = \frac{p(y_{t+1}|v; Y_t; X_t) \pi(v|Y_t)}{\int_V p(y_{t+1}|v'; Y_t; X_t) \pi(v'|Y_t)\, dv'}.$$

In order to predict forward to the next time, we must not only propagate $X_{t+1} = X_t + v_t + \delta_t$, but we must also predict the posterior so that it will become the next step's prior. The reason for the additional posterior propagation is that the velocity v is not a fixed parameter to be estimated but a dynamic quantity in its own right through the X's. Recall that the velocity v_{t+1} satisfies

$$\begin{aligned} v_{t+1} &= \textstyle\sum_{k=0}^{p} \hat{A}_k X_{t+1-k} \\ &= \hat{A}_0 (v_t + X_t + \delta_t) + \textstyle\sum_{k=1}^{p} \hat{A}_k X_{t+1-k}, \end{aligned}$$

so that the distribution of v_t, conditioned on the past, together with the distribution of δ_t, determine the distribution of v_{t+1}. The computation of Bayesian filter is thus complete.

VI. RESULTS

We have experimented with the Bayesian algorithm of Section V on data supplied by Dr. Bill Brown of the Air Force Research Laboratory at Kirtland AFB. Using a very high resolution wave optics simulation, Dr. Brown generated data for an ABL engagement scenario involving a roughly 200km path of propagation. The turbulence in this scenario is fairly strong, having a Rytov number of 0.70.

The frame rate of the camera simulated is 5000 Hz, and the target is moving at roughly 1.5 km/sec. Dr. Brown's simulation produced 1050 frames of actively illuminated target data, as well as 1050 frames of back-propagated point source data, which provides a "truth" signal. We illustrate the results of 300 frames of analysis in the figure below.

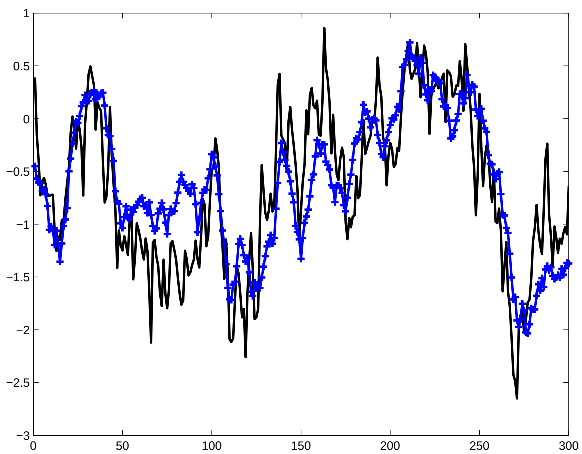

Fig. 1. Longitudinal Tilts, truth and Bayes

In Figure 1, we have the "plus sign" graph denoting the Bayesian tracking scheme following the truth signal very well. This is in distinction with a centroid tracker, which is shown in Figure 2. The centroid signal in Figure 2 is marked with circles on the data points.

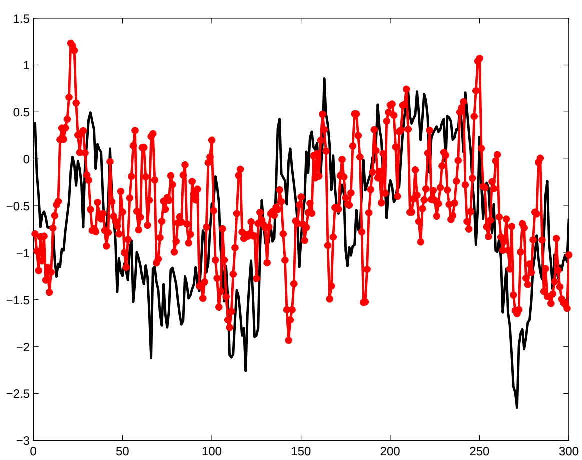

Fig. 2. Longitudinal Tilts, truth and centroid

The Bayesian scheme outperformed the centroid by a wide margin. The rms error over the full 1050 frame set is 0.54 pixels for the Bayesian scheme and 0.94 pixels for the centroid.

VII. Conclusions and Future Work

We have examined in this paper the application of Bayesian statistical methods to the problem of tilt estimation in optical tracking problems. The flexibility of the Bayesian approach allows modeling of many phenomena of interest, particularly the uneven illumination of the target due to the turbulent perturbations of the outgoing beams. We have examined the applicability on open loop simulation data, and the Bayesian approach shows great promise. Future work focuses on closed loop simulation in wave optics software and laboratory settings. Moreover, we plan to reexamine the filter using the more accurate colored noise modeling of the tilt process.

VIII. Acknowledgments

The authors gratefully acknowledge the support of the Joint Technology Office and the Air Force Office of Scientific Research through the Multidisciplinary Research Initiative grant F49620-02-1-0319, for which the University of California at Los Angeles is the prime contractor. We also acknowledge the support of Professor Steve Gibson, the Principal Investigator on F49620-02-1-0319, for both his support of our subcontract and our many discussions on the subjects of this paper. Further, we acknowledge the collaborative support of Dr. Don Washburn, Dr. Russ Butts, and Dr. Bill Brown of AFRL/DES.

References

[1] Roggemann, M., and B. Welsh. *Imaging through Turbulence*, CRC Press, Boca Raton, FL; 1996.

[2] Fitts, *Correlation Tracking via Optimal Weighting Functions*, Report Number P73-240, Hughes Aircraft Co., 1973.

[3] Graybill, F. A. *Theory and Application of the Linear Model*, Duxbury, North Scituate, MA, 1976.