A Bayesian Approach to Sensor Characterization

Doğan A. Timuçin

NASA Ames Research Center, Mail Stop 269-3, Moffett Field, California 94035, USA
(650) 604-1262 • timucin@ptolemy.arc.nasa.gov

***Abstract*— The physical model of a generic electro-optic sensor is derived and incorporated into a Bayesian framework for the estimation of key instrument parameters from calibration data. The sensor characterization thus achieved enables optimal subsequent removal of instrument effects from field data, leading to the highest possible accuracy in the retrieved physical quantities.**

I. Introduction

During the act of measuring a physical signal, a sensor inevitably imparts its own "signature" by altering the signal properties in some unique fashion; in this sense, the measurement device constitutes an integral part of any physical data-collection process. In terrestrial remote sensing, instruments are typically used as relative-measurement devices, for which a complete removal of such "sensor effects" is not crucial: these devices are simply **made** to agree with each other by calibrating them against standard sources. Their common calibrated outputs may, however, still differ appreciably from the true physical input provided by the calibration source.

In order to be able to perform accurate **absolute** measurements, especially of small features in sparse data, it is therefore necessary to develop sophisticated calibration and data-processing algorithms based on a detailed physical model of the sensor. We further maintain that, whenever a parametric model of an observed physical system can be formulated from first principles, the most natural tool for analyzing the collected data is *Bayesian inference*. This is a rich and venerable parameter-estimation technique that is enjoying wide-spread popularity in the scientific data-analysis community on the heels of dramatic recent advances in computational techniques and power. In this paper, we present a Bayesian approach toward characterizing the salient features of a generic electro-optic sensor via fairly rudimentary calibration experiments.

II. Bayesian Inference

At the heart of the Bayesian approach to data analysis is the famous theorem of Bayes (1763) and Laplace (1812), stated here for two continuous-valued random variables x and θ [1]:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)\,\mathrm{d}\theta}, \qquad (1)$$

where $p(x)$ denotes the marginal probability density function (PDF) of x, while $p(x|\theta)$ denotes the conditional PDF of x given a specific value of θ; similarly for $p(\theta)$ and $p(\theta|x)$. In the Bayesian philosophy, these PDFs are viewed as representing our state of knowledge: the sharper, say, $p(\theta)$ is around some value θ^*, the more confident we are that $\theta \simeq \theta^*$ in actuality.

In the context of data analysis, x denotes an experimental measurement, and θ represents an unknown parameter associated with the experiment, the objective being the estimation of this parameter from the observation.[1] In this setting, $p(\theta)$ is referred to as the *a priori* PDF of θ, representing our initial "best guess" and associated uncertainty about θ. Upon observation, this prior is transformed into the *a posteriori* PDF $p(\theta|x)$ via Bayes' theorem (1). This transformation is facilitated through the *likelihood function* $p(x|\theta)$, whose form represents the solution of the underlying modeling problem.

A sharper posterior relative to the prior indicates an improved confidence on our part as to the value of θ after having seen the data; a suitable decision rule may now be used to infer an optimal estimate θ^* (see Fig. 1). The Bayesian approach thus utilizes the solution of the conceptually easier "forward" modeling problem to solve the more difficult, and arguably more interesting, "inverse" inference problem, with domain expertise efficiently brought to bear via the prior and the likelihood. Incidentally, Bayesian inference subsumes the more familiar techniques of *maximum likelihood estimation* and *least-squares fitting* as special cases corresponding to a uniform prior and a Gaussian posterior, respectively [2].

[1]The extension to multiple parameters and data is obvious; *cf.* Sec. IV.

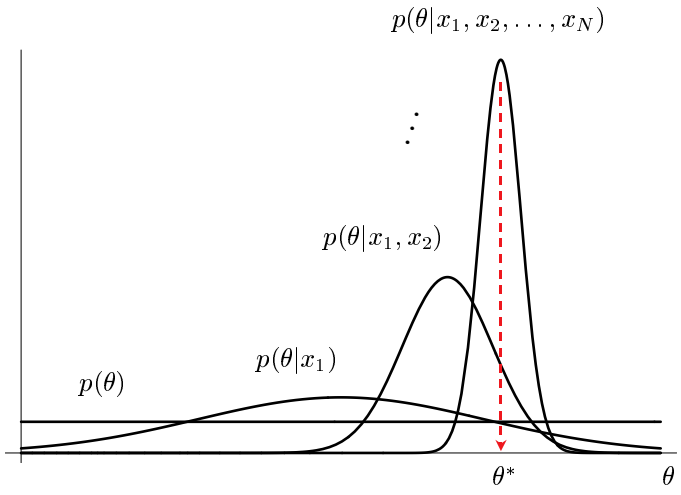

Fig. 1. The classic depiction of the Bayesian learning process. The uniform prior $p(\theta)$ signifies a complete lack of initial knowledge about the unknown parameter θ. As evidenced by the steady sharpening of the posterior, this uncertainty is reduced gradually by applying (1) iteratively on a sequence of observations $\{x_n,\ n = 1, 2, \ldots, N\}$. After the final iteration, the *maximum a posteriori* decision strategy yields the optimal estimate θ^*.

III. Sensor Model

A brute-force, "frontal attack" toward sensor modeling is typically hindered by severe practical difficulties stemming from the complicated designs of most instruments, the details of which may not even be fully known or accessible for characterization by the user, but are **always** subject to random and unpredictable (mechanical, thermal, optical, *etc.*) perturbations during operation. We therefore adopt a *reductionist approach* instead, whereby the sensor is treated as a "black box" intended to provide a robust and economical description of the salient instrument features. With an eye toward Bayesian inference, we derive parametric models for the optical and electrical subsystems of a generic sensor. These subsystems, connected through the (nonlinear) process of photodetection, are separately assumed to be linear and time-invariant, as is appropriate for a sensor intended as a measurement device.

A. Deterministic Analysis

The optical wave incident on the sensor aperture Σ_i serves as the input to the optical subsystem, whose output, in turn, is another optical wave incident on a detector aperture Σ_o terminating the optical train (see Fig. 2(*a*)). The most general relationship between the electric fields associated with these input and output waves may be written in the form

$$\vec{\mathcal{E}}_o(\vec{\rho}_o, t) = \int_{-\infty}^{\infty} \int_{\Sigma_i} \mathsf{T}(\vec{\rho}_o; \vec{\rho}_i; t - t') \cdot \vec{\mathcal{E}}_i(\vec{\rho}_i, t') \, \mathrm{d}^2\vec{\rho}_i \, \mathrm{d}t', \quad (2)$$

where the 3×3 *point-spread matrix* T models the dispersive, space-varying response of the optical subsystem. As is typically the case in practice, this response is assumed to be narrow-band, centered around some frequency ω_c determined, *e.g.*, by a color filter or a grating in the optical train. Note that it suffices to track only the electric fields, since the corresponding magnetic fields may be obtained through Faraday's law as $\vec{\mathcal{H}}(\vec{r}, t) = -\frac{1}{\mu_0} \int_{-\infty}^{t} \vec{\nabla} \times \vec{\mathcal{E}}(\vec{r}, t') \, \mathrm{d}t'$. The connection with the electrical subsystem is established through the total optical power $\mathcal{P}(t)$ incident on the detector (see Fig. 2(*b*)). In terms of the Poynting vector $\vec{\mathcal{S}}(\vec{r}, t) \equiv \vec{\mathcal{E}}(\vec{r}, t) \times \vec{\mathcal{H}}(\vec{r}, t)$ giving the density and direction of propagation of electromagnetic power [3], we have (for a planar detector) $\mathcal{P}(t) = \int_{\Sigma_o} \mathcal{S}_z(\vec{\rho}_o, t) \, \mathrm{d}^2\vec{\rho}_o$.

A portion of this power is absorbed inside the photodiode and leads to the creation of charge carriers at some rate $r(t)$, which, along with "dark" carriers thermally induced at a rate ρ, constitute the photodiode current $i_{pd}(t)$. The dark rate ρ depends chiefly on temperature and may therefore be assumed constant during a single experiment. Meanwhile, the photoelectron generation rate may be put in the form

$$r(t) = \sum_{\alpha,\beta} \iiiint_{-\infty}^{\infty} \Lambda_{\alpha\beta}(\vec{\rho}, t - \tau; \vec{\rho}', t - \tau') \cdot \mathcal{A}_{i\alpha}(\vec{\rho}, \tau) \, \mathcal{A}^*_{i\beta}(\vec{\rho}', \tau') \, \mathrm{d}^2\vec{\rho} \, \mathrm{d}^2\vec{\rho}' \, \mathrm{d}\tau \, \mathrm{d}\tau', \quad (3)$$

$\alpha, \beta = x, y$, where $\mathcal{A}_{i\alpha}$ denotes the complex envelope of $\mathcal{E}_{i\alpha}$. (The key to reaching (3) is the representation of the incident field envelope in terms of its *angular plane-wave spectrum* $\vec{A}_i(\vec{k}, \omega) \equiv \iint_{-\infty}^{\infty} \vec{\mathcal{A}}_i(\vec{\rho}_i, t) \, \mathrm{e}^{-\mathrm{i}(\vec{k} \cdot \vec{\rho}_i - \omega t)} \, \mathrm{d}^2\vec{\rho}_i \, \mathrm{d}t$ [4].)

The complex-valued functions $\Lambda_{\alpha\beta}$ thus introduced involve, in some complicated fashion, the (independent) elements of T, the shapes and sizes of Σ_i and Σ_o, as well as the detector quantum efficiency; in keeping with the reductionist spirit, we do not concern ourselves here with their exact analytic form.

Two electrical outputs are identified. The continuous signal v_t is envisaged as the output voltage of an integrator inside an analog-to-digital converter (ADC). With $f(t)$ denoting the overall baseband *transimpedance impulse response*, we have

$$v_t = \sum_{k=1}^{K} f(t - t_k) + \sum_{l=1}^{L} f(t - t_l) + \xi(t). \quad (4)$$

Here, t_k and t_l respectively denote the time instants of photo- and thermally-induced charge-carrier excitations inside the photodiode, and $\xi(t)$ represents the independent thermal-noise fluctuations of the electronic circuitry. The first term on the right side of (4) is labeled as the "signal" voltage $v_t^{(s)}$, while the two remaining terms together comprise the "noise" voltage $v_t^{(n)}$. The discrete signal s_t is then obtained by passing v_t through a b-bit quantizer inside the ADC, and is assumed to be the actual instrument reading available to the user.[2]

B. Statistical Analysis [5]

We now seek to derive the PDF $p(v)$ of the continuous output signal v_t in (4); it is actually more convenient to deal instead with the corresponding characteristic function (CF)

[2] As a subscript, t indexes epochs of duration T_* – the preset integration time of the sensor – marking the time instants of data read-out.

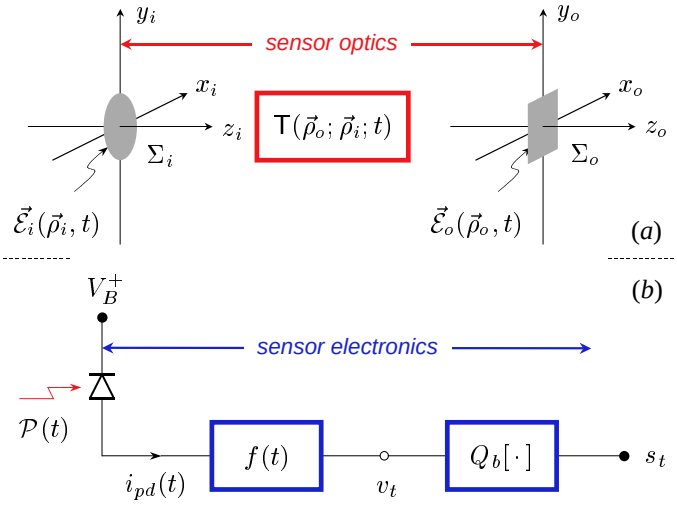

Fig. 2. Sketch of a generic electro-optic sensor. (*a*) $\mathsf{T}(\vec{\rho}_o; \vec{\rho}_i; t)$ represents the entire optical train consisting of lenses, mirrors, waveguides, gratings, *etc.*, residing between some suitably defined input (*i*) and output (*o*) planes. The elements of T connect the various polarization components of $\vec{\mathcal{E}}_i$ and $\vec{\mathcal{E}}_o$. (Note that T is of second rank, since both $\vec{\mathcal{E}}_i$ and $\vec{\mathcal{E}}_o$ are divergence-free by virtue of Gauss's law.) (*b*) $f(t)$ represents the entire electrical baseband consisting of the photodiode, a front-end amplifier, and an integrator. The output signal v_t is quantized to produce s_t, which is then presented to the user as a b-bit digital number. (In an imaging or hyper-spectral instrument, each photodiode constitutes an electro-optic "channel" of this sort, with the diode currents typically multiplexed through a common amplifier–ADC path.)

$C(u) = \int_{-\infty}^{\infty} p(v)\, \mathrm{e}^{\mathrm{i}uv}\, \mathrm{d}v = C^{(s)}(u)\, C^{(n)}(u)$. Thermal noise is customarily modeled as a zero-mean, Gaussian stochastic process, while dark noise constitutes a filtered, homogenenous Poisson point process; *i.e.*, shot noise [1]. On the other hand, due to the inherent randomness of optical wavefields, $r(t)$ is itself a stochastic process, rendering $v_t^{(s)}$ a filtered, doubly-stochastic Poisson point process [6], [7]. Thus,

$$C(u) = \mathrm{e}^{\rho \int_0^{T_*} \left[\mathrm{e}^{\mathrm{i}uf(t')}-1\right] \mathrm{d}t' - \frac{1}{2}\sigma_T^2 u^2} \cdot \int_0^{\infty} \mathrm{e}^{\int_{t-T_*}^{t} r(t') \left[\mathrm{e}^{\mathrm{i}uf(t-t')}-1\right] \mathrm{d}t'}\, p(r)\, \mathrm{d}r, \qquad (5)$$

where σ_T^2 denotes the thermal noise variance, and the PDF $p(r)$ of the rate process $r(t)$ is to be determined subsequently by elucidating the incident field statistics (*cf.* Sec. IV).

For further analytical progress, we now make in (5) the practically justifiable assumption that T_* far exceeds the time constants associated with the photodiode and amplifier impulse responses. This leads to the "photon-counting" approximation $f(t) \simeq \gamma e \equiv \Gamma$, $0 \le t \le T_*$, where γ is the DC gain of the amplifier and e is the electronic charge. With $C_w(u)$ denoting the CF of the *integrated intensity* $w \equiv \int_{t-T_*}^{t} r(t')\, \mathrm{d}t'$, we find

$$C(u) \simeq \mathrm{e}^{\rho T_* \left(\mathrm{e}^{\mathrm{i}\Gamma u}-1\right) - \frac{1}{2}\sigma_T^2 u^2}\, C_w\left[\mathrm{i}\left(1-\mathrm{e}^{\mathrm{i}\Gamma u}\right)\right]. \qquad (6)$$

Incidentally, all moments of v_t can be calculated **exactly** from (5) as $\langle v^p \rangle = \frac{1}{\mathrm{i}^p} \frac{\mathrm{d}^p C(0)}{\mathrm{d}u^p}$. Finally, the PDF of the discrete sensor output is $P\{S_q\} = \int_{V_{q-1}}^{V_q} p(v)\, \mathrm{d}v$, or more conveniently

$$P\{S_q\} = \frac{\delta}{2\pi} \int_{-\infty}^{\infty} C(u)\, \mathrm{sinc}\left(\frac{\delta u}{2\pi}\right)\, \mathrm{e}^{-\mathrm{i}uS_q}\, \mathrm{d}u, \qquad (7)$$

$q = 1, 2, \ldots, 2^b$, where S_q and V_q respectively label the quantization levels and thresholds, ΔV and $\delta = \frac{\Delta V}{2^b}$ are the ADC dynamic range and resolution, and $\mathrm{sinc}(x) \equiv \frac{\sin(\pi x)}{\pi x}$.

IV. Bayesian Sensor Characterization

The statistics of the radiation source must now be considered to complete the problem description. We envisage the use of both thermal and laser sources in order to fully characterize the spectral, angular, and polarization responses of the sensor. For an unpolarized, wide-band thermal source such as an incandescent calibration lamp, the complex envelope of the electric field is modeled as a circular complex Gaussian stochastic process. On the other hand, a linearly-polarized laser source is characterized by a strong coherent wave accompanied by a narrow-band thermal field representing spontaneous-emission noise. An expression covering both cases is [6]

$$C_w(u) = \frac{1}{\left(1 - \mathrm{i}u\frac{\langle \tilde{w} \rangle}{\mathcal{M}}\right)^{\mathcal{M}}}\, \exp\left(\frac{\mathrm{i}u\bar{w}}{1 - \mathrm{i}u\frac{\langle \tilde{w} \rangle}{\mathcal{M}}}\right), \qquad (8)$$

where the integrated intensity has been decomposed into coherent and thermal parts as $w = \bar{w} + \tilde{w}$, and $\mathcal{M}$ denotes the number of spatio-temporal modes of thermal radiation contributing to $\tilde{w}$. For a given source and a specific illumination geometry, one can relate $\bar{w}$, $\langle \tilde{w} \rangle$, and $\mathcal{M}$ to fundamental source characteristics and sensor parameters through (3) [4], [6], [7].

The use of (6) and (8) in (7) leads to an *Edgeworth series* [1]

$$P\{S_q\} = \frac{\delta\, \mathrm{e}^{-\frac{(S_q-\eta)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\, \left[1 + \sum_{m=3}^{\infty} \frac{\kappa_m}{m!}\, H_m\left(\frac{S_q - \eta}{\sqrt{2}\sigma}\right)\right], \qquad (9)$$

where $H_m(\cdot)$ are the Hermite polynomials, and we have

$$\eta = \Gamma\left(\bar{w} + \langle \tilde{w} \rangle + \rho\, T_*\right),$$

$$\sigma^2 = \Gamma\, \eta + \Gamma^2 \left(2\bar{w} + \langle \tilde{w} \rangle\right) \frac{\langle \tilde{w} \rangle}{\mathcal{M}} + \sigma_T^2 + \frac{\delta^2}{12},$$

$$\kappa_3 = \Gamma^2\, \eta + \Gamma^3 \left[3\left(2\bar{w} + \langle \tilde{w} \rangle\right) + 2\left(3\bar{w} + \langle \tilde{w} \rangle\right) \frac{\langle \tilde{w} \rangle}{\mathcal{M}}\right] \frac{\langle \tilde{w} \rangle}{\mathcal{M}}, \ldots$$

We thus see that $\Theta = \{\bar{w}, \langle \tilde{w} \rangle, \mathcal{M}, \rho, \sigma_T^2\}$ constitutes the set of unknown parameters in a calibration experiment where a sequence $\Xi = \{s_n, n = 1, 2, \ldots, N\}$ of measurements is made. Note that $\bar{w} = 0$ for calibration with a thermal lamp, while $w = 0$ for dark calibration. Assuming that T_* is long enough for the samples to become independent, and with $P\{s_n|\Theta\}$ given by (9), we may now rewrite (1) in the form

$$p(\Theta|\Xi) = \frac{p(\Xi|\Theta)p(\Theta)}{p(\Xi)}, \qquad p(\Xi|\Theta) = \prod_{n=1}^{N} P\{s_n|\Theta\},$$

from which elements of Θ may be inferred as in Sec. II.

V. Conclusion

Although many details have been omitted due to limited space, the theory presented here should provide a comprehensive philosophical and analytical framework for the Bayesian characterization of a wide variety of electro-optic sensors from their calibration data. A notable omission is the optical background noise, perhaps most commonly encountered as inter-pixel cross talk in imaging and hyper-spectral instruments, whose characteristics are hard to elucidate with any degree of generality and must be considered carefully for each individual instrument [5]. An important closing remark pertains to the issue of unreliable standards, which may be alleviated by monitoring, and correcting for, the source variations through the long-term time evolution of the inferred parameters.

Acknowledgment

The author would like to thank Dr. K. R. Wheeler for his invitation to participate in this project, Dr. P. A. Pilewskie for technical discussions regarding his instrument, the NASA *Solar Spectral Flux Radiometer*, which motivated the analysis presented here, and Dr. J. C. Coughlan for his financial support through the IS–IDU Project of the NASA CICT Program.

References

[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 1991.
[2] D. S. Sivia, *Data Analysis – A Bayesian Tutorial*, Oxford, UK: Oxford University Press, 1996.
[3] J. D. Jackson, *Classical Electrodynamics*, 3rd ed., New York: Wiley, 1999.
[4] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge, UK: Cambridge University Press, 1995.
[5] D. A. Timuçin, J. F. Walkup, and T. F. Krile, "Accuracy in analog optical processors: statistical analysis," *J. Opt. Soc. Am. A*, vol. 11, pp. 560–571, February 1994.
[6] B. Saleh, *Photoelectron Statistics*, New York: Springer, 1978.
[7] J. W. Goodman, *Statistical Optics*, New York: Wiley, 1985.