Optimization of the contrast in active Stokes images

François Goudail

Optimization of the contrast in active Stokes images

François Goudail

Laboratoire Charles Fabry de l'Institut d'Optique, CNRS, Université Paris-Sud, Campus Polytechnique, RD 128, 91127 Palaiseau, France (francois.goudail@institutoptique.fr)

Received August 26, 2008; revised October 27, 2008; accepted November 20, 2008; posted December 4, 2008 (Doc. ID 100603); published January 7, 2009

Active Stokes imaging consists in illuminating a scene with polarized light and in measuring the Stokes vector of the scattered light. We present a method for determining the polarization state of illumination that maximizes the observed contrast between a target and the background when the scene is partially depolarizing and in the presence of additive Gaussian detection noise. © 2009 Optical Society of America

OCIS codes: 260.5430, 030.4280.

Active Stokes imaging consists in illuminating a scene with polarized light and in measuring the Stokes vector of the scattered light. This provides useful information about the scene and makes it possible to discriminate objects that have different polarization-scattering properties. Many applications are found in biomedicine, remote sensing, or imaging through turbid media. The design of Stokes imagers that minimize and/or equalize the noise power in the different Stokes channels has been widely studied [1]. In this Letter, we seek to optimize the contrast between two regions in the scene that have different polarization properties, such as a target of interest and a background. We will consider that the structure of the Stokes imager is fixed and that the only degree of freedom available to optimize the contrast is the polarization state of the illumination.

The appropriate expression of the contrast in an image depends on the statistical properties of the noise that perturbs it [2]. In this Letter, we will assume that the noise is additive and Gaussian, since it is a simple model, generally adequate for detection noise. We propose a method for determining the polarization state of illumination that maximizes the contrast in the presence of such noise and illustrate it in several examples. These results will be useful for optimizing information extraction from active Stokes imagers in remote sensing and biomedical applications.

A Stokes vector $\vec{S}$ is estimated from N intensity measurements $\vec{I}$ ($N \geqslant 4$) in the following way:

$$\vec{I} = W\vec{S} + \vec{b}, \tag{1}$$

where W is a $N \times 4$ matrix that is determined as part of the calibration process of the polarimeter and $\vec{b}$ is a noise vector that will be assumed independent of $\vec{S}$. Then $\vec{S}$ is retrieved from $\vec{I}$ using the pseudoinverse matrix $W^{\dagger} = (W^T W)^{-1} W^T$ (one has $W^{\dagger} = W^{-1}$ if $N=4$), so that the estimate of $\vec{S}$ is

$$\hat{\vec{S}} = W^{\dagger}\vec{I} = \vec{S} + \vec{n}, \tag{2}$$

where $\vec{n} = W^{\dagger}\vec{b}$. With good approximation, the noise $\vec{b}$ can be considered white ($\langle \vec{b}\vec{b}^T \rangle = \sigma^2 \mathcal{I}$), where T denotes vector transposition, $\langle . \rangle$ ensemble averaging, $\mathcal{I}$ is the 4×4 identity matrix, and σ^2 is the variance of $\vec{b}$. The noise $\vec{n}$ that perturbs the measure of $\vec{S}$ may thus not be white, since its covariance matrix is

$$\Gamma_n = \langle \vec{n}\vec{n}^T \rangle = \sigma^2 (W^T W)^{-1}. \tag{3}$$

Our purpose is to discriminate two regions a and b whose polarimetric responses are described by their Mueller matrices M_a and M_b. We use the Mueller formalism [3], since we are interested in remote sensing or biomedical applications, where scenes are often highly depolarizing and the Jones formalism [3] is not sufficient. The scene is illuminated with purely polarized light that can have any state on the Poincaré sphere and that is represented by its Stokes vector $\vec{S}$ (see Fig. 1). The Stokes vector scattered by region a (b) is $\vec{S}_a = M_a \vec{S}$ ($\vec{S}_b = M_b \vec{S}$). The measures are perturbed by the additive Gaussian noise $\vec{n}$ defined above. The adequate expression of the contrast between the two regions is [4]

$$\mathcal{C}(\vec{S}) = (\vec{S}_a - \vec{S}_b)^T \Gamma_n^{-1} (\vec{S}_a - \vec{S}_b) = \frac{\vec{S}^T G \vec{S}}{\sigma^2}, \tag{4}$$

with

$$G = (M_a - M_b)^T W^T W (M_a - M_b). \tag{5}$$

It can be noticed that G is a symmetric matrix. Equation (4) means that if two scenes are such that the couples $(\vec{S}_a, \vec{S}_b)$ are different, but $\mathcal{C}(\vec{S})$ is identical, an optimal processing algorithm (performing detection or estimation, for example) applied to these two scenes will lead to identical performance [2]. In the sequel of this article, for illustration purpose, we will consider that W is optimal in the way defined in [5]. It has the property that $W^T W$ is diagonal with elements $(1, 1/3, 1/3, 1/3)$, and thus the variance on the

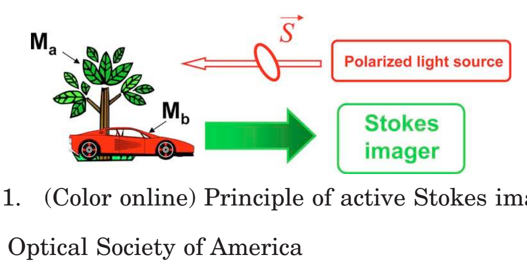

Fig. 1. (Color online) Principle of active Stokes imaging.

Stokes vector estimates [see Eq. (3)] is σ^2 for S_0 and $3\sigma^2$ for all other elements of the Stokes vector (the noise is "equalized").

Our objective is to determine the illumination Stokes vector $\vec{S}$ that maximizes the contrast $\mathcal{C}(\vec{S})$. The constraint on this optimization problem is that the vector $\vec{S}$ represents purely polarized light, that is, it can be written as $\vec{S}^T = I_0(1,\vec{s}^T)$, where $\vec{s}$, the reduced Stokes vector, is a three-dimensional unit-norm vector and I_0 is a scalar that denotes intensity of light. Let us write the matrix G as follows:

$$G = \begin{bmatrix} G_{00} & \vec{m}^T \\ \vec{m} & \tilde{G} \end{bmatrix},$$

where $\tilde{G}$ is a 3×3 symmetric matrix and $\vec{m}$ is a three-dimensional vector. With this notation, the expression of the contrast becomes

$$\mathcal{C}(\vec{s}) = \frac{I_0^2}{\sigma^2}(G_{00} + 2\vec{s}^T\vec{m} + \vec{s}^T\tilde{G}\vec{s}). \qquad (6)$$

It is seen that the contrast is trivially proportional to the intensity signal-to-noise ratio I_0^2/σ^2. Our goal is to find the vector $\vec{s}_{max}$ that maximizes $\mathcal{C}(\vec{s})$ under the constraint that $\|\vec{s}\|^2 = 1$. This optimization problem has already been encountered by Kostinski *et al.* in a different context [6]. The solution of this problem is simple when $\vec{m} = 0$. This happens in particular when the matrices M_a and M_b have the same diattenuation and polarizance vectors [3]. In this case, the function to maximize is simply $F(\vec{s}) = \vec{s}^T\tilde{G}\vec{s}$. The solution to this problem is well known; the vector $\vec{s}_{max}$ that maximizes $F(\vec{s})$ is the eigenvector of $\tilde{G}$ associated to the largest eigenvalue.

As an example of application, let us take two purely depolarizing matrices with different polarization properties. That is, M_a and M_b are diagonal with elements respectively equal to $\vec{D}^a = (1, 0.5, 0.4, 0.1)$ for M_a and $\vec{D}^b = (1, 0.3, 0.3, 0.4)$ for M_b. Since the matrices are diagonal, the solution is trivial, $\vec{s}_{max} = (0, 0, \pm 1)^T$, since these vectors correspond to the largest difference of depolarization between the two regions $|D_3^a - D_3^b| = 0.3$ and lead to the same maximal value of $F = 0.09$. They represent left and right circular states. More generally, it can be noted that orthogonal polarization states corresponding to $\vec{s}$ and $-\vec{s}$ always lead to the same value of $F(\vec{s})$; when $\vec{m} = \vec{0}$, the optimization problem thus has at least two solutions. The minimum value of $F(\vec{s})$ is 0.01 and is obtained for linear polarization at 45°, $\vec{s} = (0, 1, 0)^T$.

Let us now consider the general case where $\vec{m} \neq 0$. The function in Eq. (6) can be simplified by diagonalizing the symmetric matrix $\tilde{G}$, that is, defining $\tilde{G} = X\Lambda X^T$, where Λ is a diagonal matrix with diagonal values λ_i, $i \in [1,3]$ and X is a unitary matrix. Defining the new variables $\vec{u} = X^T\vec{s}$ and $\vec{p} = X^T\vec{m}$, the function to optimize becomes $F(\vec{u}) = 2\vec{u}^T\vec{p} + \vec{u}^T\Lambda\vec{u}$, under the constraint $\|\vec{u}\|^2 = 1$. The associated Lagrange function is

$$\mathcal{F}_\mu(\vec{u}) = 2\sum_{i=1}^{3} p_i u_i + \sum_{i=1}^{3} \lambda_i u_i^2 - \mu \sum_{i=1}^{3} u_i^2,$$

where μ is the Lagrange parameter. Annulling its partial derivatives with respect to u_i, $i \in [1,3]$, one obtains a set of three equations, $u_i(\mu - \lambda_i) = p_i$, $i \in [1,3]$. Substituting these equations into the constraint, that is, $\Sigma_i u_i^2 = 1$, it is easily seen that a necessary condition for μ being an acceptable value of the Lagrange parameter is that it verifies the following equation:

$$p_1 t_2 t_3 + p_2 t_1 t_3 + p_3 t_1 t_2 = t_1 t_2 t_3, \qquad (7)$$

with $t_i = (\mu - \lambda_i)^2$. This is a polynomial equation of order 6 in μ. It thus has six roots μ_k, $k \in [1,6]$ that can be determined numerically by standard methods such as Jenkins–Traub [7]. Only real-valued roots are relevant. The vectors $\vec{u}^k$ associated to each real-valued root μ_k can be determined in the following way: if $\forall i \in [1,3]$, $\mu_k \neq \lambda_i$, the vector $\vec{u}^k$ is given by

$$u_i^k = p_i/(\mu_k - \lambda_i). \qquad (8)$$

If $\exists i$, $\mu_k = \lambda_i$, several cases must be considered. Let us assume, without loss of generality, that the eigenvalues are in ascending order: $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Suppose for example that $\mu_k = \lambda_1$ (this can happen only if $p_1 = 0$). The different cases to consider are

- if $\lambda_1 \neq \lambda_2$: $u_2^k = p_2/(\lambda_1 - \lambda_2)$ and $u_3^k = p_3/(\lambda_1 - \lambda_3)$. This is acceptable if $[u_2^k]^2 + [u_3^k]^2 < 1$. In this case, one has two possible solutions that correspond to $u_1^k = \pm\sqrt{1 - [u_2^k]^2 + [u_3^k]^2}$.
- if $\lambda_1 = \lambda_2$: $u_3^k = p_3/(\lambda_1 - \lambda_3)$. This is acceptable if $[u_3^k]^2 < 1$. In this case, one has an infinity of possible solutions that all verify the relation $[u_1^k]^2 + [u_2^k]^2 = 1 - [u_3^k]^2$.

It can be noticed that if $\lambda_1 = \lambda_2 = \lambda_3$, and if at least one value of p_i is different from zero, then μ cannot be equal to λ, so the two previous cases are the only possible ones. The cases $\mu_k = \lambda_2$ and $\mu_k = \lambda_3$ are treated in the same way.

After all the possible vectors $\vec{u}^k$ that correspond to local extrema have been computed, the one that maximizes $F(\vec{u}^k)$ is determined: $\vec{u}_{max} = \arg\max_{\vec{u}^k}[F(\vec{u}^k)]$. Finally, the reduced Stokes vector that leads to the maximal contrast is $\vec{s}_{max} = X\vec{u}_{max}$. It is also possible to determine the maximum of the function $F(\vec{u})$ with standard numerical optimization software. However, these methods lead to local maxima, whereas the function $F(\vec{u})$ can have more than one maximum. It is thus necessary to solve the problem with a sufficient number of different starting points. A further interest of the above-described method is to be able to determine cases where the optimum is not unique.

Let us take an example of application of this method. Assume that the depolarizing matrix M_a is the same as before, but M_b now has nonzero polarizance such that

$$M_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a \times 0.866 & 0.3 & 0 & 0 \\ a \times 0.5 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 0.4 \end{bmatrix}, \quad (9)$$

where a is a scalar. The normalized polarizance vector $\vec{P}^T = (0.866, 0.5, 0)$, which is the principal polarization state of light obtained when illuminating the material with totally depolarized light [3], corresponds to a linear state of azimuth 15°. It can be shown that $\vec{m} = a(-0.173, -0.05, 0)^T$ [see Eq. (6)]. This vector is collinear to the reduced Stokes vector $\vec{s}_m$ of a linear polarization state of azimuth 98.1°, which is represented in Fig. 2 with a diamond symbol on the Poincaré sphere. We have also plotted in Fig. 2 the trajectory of $\vec{s}_{max}$, determined with the method described above, when a takes ten different values between 0 and 0.35 (it has been checked that for all these values of a, M_b is a physically realizable Mueller matrix [3]).

When a is zero, the optimal state is circular as shown above. When a increases, the ellipticity of $\vec{s}_{max}$ decreases, and its azimuth remains constant at 95.1°, which is close to that of the vector $\vec{s}_m$. Then, when $a > 0.2841$, the ellipticity becomes equal to zero (the optimal state is rigorously linear), and its azimuth varies slowly towards that of $\vec{s}_m$. Indeed, as a increases, the norm of $\vec{m}$ increases and so does the influence of the term $2\vec{s}^T\vec{m}$ in Eq. (6). It is thus normal that $\vec{s}_{max}$ gets more collinear to $\vec{s}_m$ so that this term increases.

We have represented in Fig. 3 the variation with a of the contrast $\mathcal{C}(\vec{s})$ for different illumination states: the optimal state $\vec{s}_{max}$, the left circular state, and the linear state with azimuth 95.1°. When $a = 0$, the circular state is optimal, and the contrast is 2.3 times higher than what it would be for the linear state. On the other hand, when $a > 0.25$, the linear state is quasi optimal, and one gains a factor 1.35 in contrast compared to a circularly polarized illumination. This illustrates the interest of adapting the illumination state of polarization to the characteristics of the regions to discriminate.

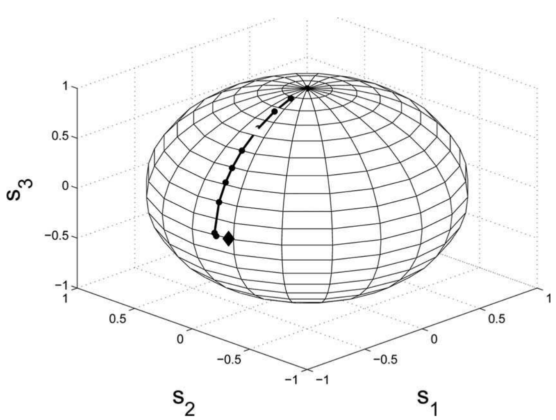

Fig. 2. Solid curve, optimal polarization state on the Poincaré sphere as a varies (solid curve). The matrix M_a is diagonal with elements $\vec{D}^a = (1, 0.5, 0.4, 0.1)$, and the matrix M_b is in Eq. (9). Diamond, reduced Stokes vector $\vec{s}_m$ collinear to $\vec{m}$.

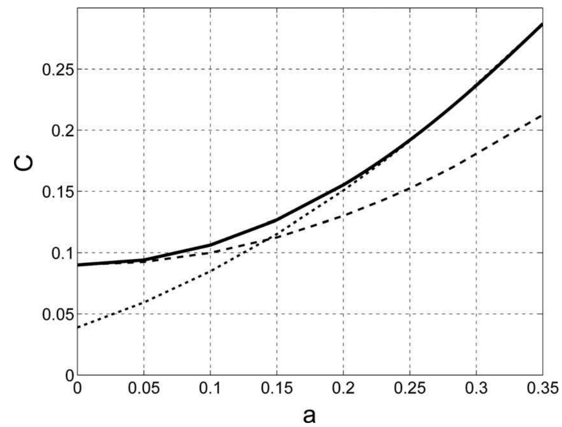

Fig. 3. Variation of the contrast $\mathcal{C}(\vec{s})$ for different states of illumination $\vec{s}$, and $\sigma^2 = 1/3$. Solid curve, optimal value $\vec{s}_{max}$. Dashed curve, left circular; dotted curve, linear with azimuth 95.1°.

In summary, a method has been given to compute the polarization state of illumination that optimizes the contrast for region discrimination in active Stokes images perturbed with additive Gaussian noise. A significant increase of the contrast can be obtained by adapting the state of polarization of illumination to the two regions to discriminate. This method is thus useful to optimize the information content of active Stokes images of partially depolarizing scenes in remote sensing and biomedical optics. The study of contrast optimization in the presence of photon detection noise [8] is an interesting perspective to this work, since the optimal illumination state may be different in this case.

The author wishes to thank Arnaud Bénière for fruitful discussion.

References

1. J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, Appl. Opt. **45**, 5453 (2006).
2. F. Goudail, P. Réfrégier, and G. Delyon, J. Opt. Soc. Am. A **21**, 1231 (2004).
3. D. Goldstein, *Polarized Light*, 2nd ed. (Marcel Dekker, 2003).
4. K. Fukunaga, *Introduction to Statistical Pattern Recognition* (Academic, 1990).
5. J. S. Tyo, Opt. Lett. **25**, 1198 (2000).
6. A. B. Kostinski, B. D. James, and W. M. Boerner, J. Opt. Soc. Am. A **5**, 58 (1988).
7. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++: the Art of Scientific Computing*, 2nd ed. (Cambridge U. Press, 2002).
8. V. L. Gamiz and J. F. Belsher, Opt. Eng. **41**, 973 (2002).