Optical sectioning with Structured Illumination Microscopy for retinal imaging: inverse problem approach

Roberto BAENA-GALLÉ[1], Laurent M. MUGNIER[2], François ORIEUX[3]

[1]Institut de la Vision – Université Pierre et Marie Curie
17 Rue Moreau, 75012 Paris, France

[2]ONERA – The French Aerospace Lab
29 Avenue de la Division Leclerc, 92320 Châtillon, France

[3]Univ. Paris-Sud – Laboratoire des Signaux et Systèmes
3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France

roberto.baena@gmail.com, mugnier@onera.fr

Résumé – La microscopie par illumination structurée (ou SIM en Anglais) est une technique d'imagerie permettant d'obtenir super-résolution et sectionnement optique en microscopie de fluorescence. L'échantillon objet est éclairé par des motifs de franges sinusoïdales avec différentes orientations et déphasages. Ceci a pour effet d'introduire par repliement de spectre des informations à hautes fréquences spatiales de l'objet à l'intérieur du support de la fonction de transfert. L'image résultante est traitée avec des logiciels de reconstruction dédiés qui permettent de récupérer des fréquences spatiales au-delà de la coupure de l'instrument et, simultanément, d'enlever la lumière provenant des tranches défocalisées si l'objet est volumique (ce qu'on appelle sectionnement optique).

Malheureusement, alors que pour les échantillons statiques, les déphasages des sinusoïdes peuvent être définis par l'utilisateur, fournissant ainsi des solutions analytiques, ce n'est pas possible pour des échantillons *in-vivo*, en particulier pour des images rétiniennes, du fait des mouvements oculaires incontrôlés.

L'objectif de cette communication est de démontrer que la SIM peut être appliquée à l'imagerie rétinienne *in-vivo* afin d'obtenir à la fois sectionnement optique et super-résolution à partir d'images grand champ corrigées par optique adaptative. Nous introduisons une nouvelle approche, dans un cadre bayésien, qui permet de manière simple d'atteindre un sectionnement optique maximal. À cette fin, une image conventionnelle est enregistrée, recalée et soustraite de chacune des observations SIM basse résolution, de sorte que les données différentielles résultantes contiennent uniquement des informations sur la tranche focalisée de l'objet. Parce que notre méthode ne modélise pas explicitement une observation 3D mais conserve la simplicité d'un modèle 2D, elle est rapide et facile à mettre en œuvre dans la pratique. Nous montrons des résultats de simulations qui prouvent la validité de notre approche.

Abstract – Structured Illumination Microscopy (SIM) is an imaging technique for obtaining super-resolution and optical sectioning (OS) in wide-field fluorescence microscopy. The object sample is illuminated by sinusoidal fringe patterns at different orientations and phase shifts. This has the effect of introducing high frequency information of the object into the support of the transfer function by aliasing. The resulting image is processed with dedicated reconstruction softwares which allow recovering high frequencies beyond the instrument cut-off and, simultaneously, removing the light coming from the out-of-focus slices of a 3D volume (which is called optical sectioning).

Unfortunately, whereas for static samples the phase shifts of the sinusoids can be set by the user thus providing analytical solutions, this is not possible for *in-vivo* samples, and in particular for retinal images, due to the uncontrolled eye movements.

The aim of this communication is to demonstrate that SIM can be applied to *in-vivo* retinal imaging in order to obtain both OS and super-resolution from flood-illuminated and adaptive-optics corrected observations. We introduce a new approach, within the Bayesian framework, which allows in a simple way to achieve maximal OS. For that purpose, a conventional wide-field image is registered and subtracted with respect to each one of the low resolution SIM observations, hence, the resulting differential data only contains information on the in-focus slice of the object. Because our method does not model explicitly a 3D observation but keeps the simplicity of a 2D model, it is fast and easy to implement in practice. We show results from simulations that prove the validity of our approach.

1 Model of Structured Illumination Microscopy images

Structured illumination microscopy (SIM) is a wide-field fluorescence microscopy method which allows increasing both the lateral [1] and the axial [2] resolution of an observed 3D volume. This is achieved by illuminating the sample of interest with a light pattern, usually a sinusoid characterized by a modulation frequency, which has the effect of shifting the object frequency content. In other words, the key idea of SIM is to inject object high frequencies into the support of the optical transfer function (OTF) of the instrument, below the diffrac-

tion cut-off frequency f_c, by means of amplitude modulation before the convolution with the PSF, i.e., SIM introduces aliasing through modulation. The illumination modulation pattern is expressed as :

$$m(x,y) = 1 + m'(x,y) = 1 + \cos\big(2\pi(k_x x + k_y y + \phi)\big), \quad (1)$$

where (x, y) are spatial 2D coordinates, (k_x, k_y) is the modulation frequency of the sinusoid, and ϕ is the relative phase shift of the sinusoid w.r.t. the sample. Therefore, three deltas in the Fourier domain are introduced by the pattern and placed on the frequencies $(0, 0)$, (k_x, k_y) and $(-k_x, -k_y)$. If a 2D object $o(x, y)$, with continuous Fourier transform $\tilde{o}(f_x, f_y)$, is multiplied by the pattern expressed by Equation (1), and observed with a microscope characterized by a PSF $h(x, y)$, corresponding to an OTF $\tilde{h}(f_x, f_y)$, and if we omit the presence of noise, then an image is formed both in Fourier and direct space as :

$$i(x,y) = h(x,y) * \big(m(x,y) \cdot o(x,y)\big), \quad (2)$$

and

$$\begin{aligned}\tilde{i}(f_x,f_y) = \tilde{h}(f_x,f_y)\Big(&\tilde{o}(f_x,f_y) \\ &+ \tilde{o}(f_x - k_x, f_y - k_y)e^{+2i\pi\phi} \\ &+ \tilde{o}(f_x + k_x, f_y + k_y)e^{-2i\pi\phi}\Big),\end{aligned} \quad (3)$$

where $*$ represents the convolution operator. This is an expression with three unknowns, i.e., the three replicated objects placed on frequencies $(0, 0)$, (k_x, k_y) and $(-k_x, -k_y)$, hence, two more equations are needed to make this system solvable. In the traditional SIM approach, assuming the sample is fixed, this can be achieved by modifying the sinusoid phase ϕ to known values 0, 120 and 240 degrees, yielding an analytical solution. In the case of retinal imaging, the natural motion of the eye provides those relative phase shifts, keeping the pattern position fixed, but more than 3 images (i.e., phases shifts) must be acquired since it is unlikely that the random trembling of the eye produces the aforementioned values. This must be repeated at least for three orientations of the sinusoid at 0, 60 and 120 degrees to have a good coverage of the whole 2D Fourier domain.

Figure 1 shows graphically the concept of SIM. One can see that the first consequence of multiplying the sample by a sinusoidal pattern is shifting object high frequencies, usually beyond the instrument cut-off so they are unobservable, into the OTF support. Therefore, with convenient post-processing techniques they can be disentangled thus allowing us to increase the recovered object resolution up to twice the cut-off frequency.

Equation (2) can be extended along a 3D volume by considering a finite number of thin slices, each of them blurred with PSFs which increase their defocus level with the axial distance :

$$i_l(x,y) = \sum_{z=-Z}^{Z} h_z(x,y) * \Big(m_l(x,y) \cdot \big(o_z(x,y) * s_l(x,y)\big)\Big), \quad (4)$$

where z indexes the slice along the 3D volume, and the term $s_l(x, y)$ has been introduced to model a shift of the object due

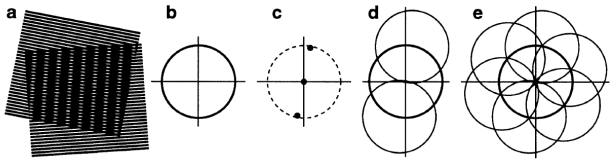

FIGURE 1 – Concept of SIM. (a) if two line patterns are superposed (one from the object, one being the illumination pattern), their product contains coarser fringes. This visible Moiré effect shows the aliasing. (b) Spatial frequency support of OTF. (c) A sinusoidal pattern plus a constant translates into 3 Diracs in the Fourier domain. Spatial frequencies of the original signal will be now centered and replicated around these positions. (d) This means that high frequencies of the object lie now inside the OTF support, superimposed with the original signal. (e) Different orientations of the pattern will allow an isotropic frequency coverage. After Ref. [1].

to, for instance, the natural motion of the eye, i.e., $s_l(x, y) = \delta(x - x_l, y - y_l)$. Therefore, due to $s_l(x, y)$ we can keep the illumination pattern fixed setting $\phi = 0$ in Equation (1). Bearing in mind that $m(x, y) = 1 + m'(x, y)$ and since, for $|z|$ sufficiently large, the corresponding defocused OTF $H_z(f_x, f_y)$ is very small around the modulation frequency (k_x, k_y) of the pattern $m(x, y)$ (as can be seen in Fig. 2), Equation (4) can be approximated as :

$$\begin{aligned} i_l(x,y) &\approx \Big(\sum_{z=-Z}^{Z} h_z(x,y) * o_z(x,y)\Big) * s_l(x,y) \\ &\quad + h_0(x,y) * \Big(m'_l(x,y) \cdot \big(o_0(x,y) * s_l(x,y)\big)\Big) \\ &\triangleq i_{CV}(x,y) * s_l(x,y) + i_{l|OS}(x,y), \end{aligned} \quad (5)$$

where we have defined $i_{CV}(x, y)$ as one conventional widefield observation, which is shifted to a random position by s_l, and $i_{l|OS}$ as the aliased replica; h_0 and o_0 are, respectively, the PSF and the object at the focused slice $z = 0$. Looking at Equation (5) one can see that the conventional image i_{CV} contains contributions from all object slices whereas $i_{l|OS}$ only contains information about the in-focus slice of the object.

Figure 2 shows graphically how optical sectioning (OS) is obtained. Five plots representing the OTF and object spectra, at five different slices along the volume, are depicted. The leftmost panel is in focus while defocus increases from left to right. In that panel, the conventional image is centered at the zero frequency while the two aliased replica, attenuated by the OTF, are placed at half the cut-off frequency. The reader can notice how the two replica of the original signal are canceled by the first zero of the defocused OTF in the fourth panel, thus achieving OS.

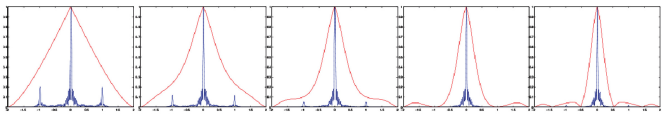

FIGURE 2 – Spectrum of the image of a sinusoidally illuminated object (blue) and corresponding modulation transfer functions of an incoherent imaging system (red) for different axial distances. After Ref. [3].

2 Proposed method for OS with retinal images

In this section we show a new approach, within the Bayesian framework [4], to obtain OS from a set of low resolution (LR) SIM images in the context of retinal observations. Authors in reference [5] propose a 3D image model in order to explicitly remove the out-of-focus content from the plane of interest. This approach has the drawback of increasing the number of unknowns to be estimated since all the $2Z+1$ slices along the 3D volume must be calculated, hence, it is computationally costly and the minimization is prone to instabilities. Unlike [5], our model is still 2D and it is based on the fact that most of the out-of-focus content lies in the conventional image $i_{CV}(x,y)$, whereas the aliased signals $i_{l|OS}(x,y)$ exhibit OS. Our approach uses this physical insight of the image formation process; it consists in performing and modeling the removal of the conventional image from the LR SIM data to eliminate the out-of-focus light. The conventional image $i_{CV}(x,y)$ can be recorded independently of $i_l(x,y)$, it is then only necessary to estimate the shifts $s_l(x,y)$ between them, in order to subtract the correctly shifted version of $i_{CV}(x,y)$ from $i_l(x,y)$ (see Eq. (5)). The recording of $i_{CV}(x,y)$ can easily be done by switching off the modulation over the sample—this is a realistic assumption since the illuminations are usually projected with semiconductor devices that are able to change their pattern within milliseconds.

Therefore, using lexicographic matrix notation, the image model that we propose describes the differential LR SIM data as :

$$\mathbf{i}_{l|OS} = \mathbf{i}_l - \mathbf{i}_{CV} * \mathbf{s}_l = \mathbf{H}_0 \mathbf{M'}_l \mathbf{S}_l \mathbf{o}_0 + \mathbf{n}_l, \qquad (6)$$

where vector $\mathbf{i}_l$ is one LR SIM data image indexed by $l \in [1, L]$, L is the total number of LR SIM images at all fringe orientations and phase shifts (at least 3 images and 3 rotations are needed reaching a minimal number of 9 images), $\mathbf{o}_0$ is the unknown 2D object to be estimated at slice $z = 0$, matrix $\mathbf{H}_0$ is a Toeplitz-Block-Toeplitz matrix representing convolution with the discrete 2-D focused PSF, matrix $\mathbf{M'}_l$ is a diagonal matrix where the diagonal elements correspond with the modulation patterns $m'_l(x,y)$, which are fixed for all the image data at the same pattern rotation, matrix $\mathbf{S}_l$ represents the bidirectional object shift (X and Y coordinates) w.r.t. the pattern for observation l, and $\mathbf{n}_l$ models the noise. In this model, the illumination patterns, at all rotations, are considered to be known whereas the object shifts, represented by $\mathbf{S}_l$, must be estimated.

The cost function to be minimized adopts the following expression when homogeneous white Gaussian noise of the same variance σ^2 is considered for all images at all fringe rotations :

$$F = \sum_{l}^{L} \frac{1}{2\sigma^2} ||\mathbf{i}_{l|OS} - \mathbf{H}_0 \mathbf{M'}_l \mathbf{S}_l \mathbf{o}_0||^2 + R(\mathbf{o}_0). \qquad (7)$$

There is no analytical solution for Equation (7) under positivity constraint, so the minimization must be done with numerical methods. Here, we have used a Variable Metric with Limited Memory and Bounds called VMLM-B [6].

We have assumed that a conventional observation of the sample can be recorded independently removing the illumination pattern so an image $\mathbf{i}_{CV}$ is available. Before subtracting the estimated conventional image to the LR SIM ones they must be registered, with the difficulty that the latter ones are modulated when the conventional image is not. Here, we use the initialization of the method developed by [7] for a similar image model, which begins by exploring all possible shifts on a one-pixel-pitch grid (making use of correlation products in the Fourier domain) in order to avoid getting stuck in local minima. Then we perform a least square estimation w.r.t. the object shifts $\mathbf{S}_l$ in order to improve the estimation at the sub-pixel level. Once the conventional image $\mathbf{i}_{CV}$ is registered with each one of the LR SIM data $\mathbf{i}_l$, the former is subtracted from the latter in order to create the new differential data $\mathbf{i}_{l|OS}$, which contains only the aliased information (see Fig. 3).

Then, the object is estimated by a minimization of the MAP metric of Equation (7) with a regularization model $R(\mathbf{o}_0)$ which assumes a stationary Gaussian prior probability distribution described by a mean object $\mathbf{o}_m$ and a 3-parameter PSD. These 3 parameters can be estimated by maximum likelihood prior to the image reconstruction, as in [8]. Positivity was also imposed.

3 Simulations and results

To test the validity of this approach, ground truth images shown in Figure 3 (top row, left and middle panel) were padded with zeros and shifted a random number of pixels around their central positions. These images were used, in our simulations, as the in- and the out-of-focus slices within a 3D volume. Each one of the resulting images was multiplied by fringe patterns for three different rotations at 0, 60 and 120 degrees. Different numbers of images per rotation were considered, e.g., 9, 18, 36 and 72. The modulation frequency was set to approximately half of the instrument cut-off, *i.e.*, $(k_x, k_y) \sim (f_c/2, f_c/2)$.

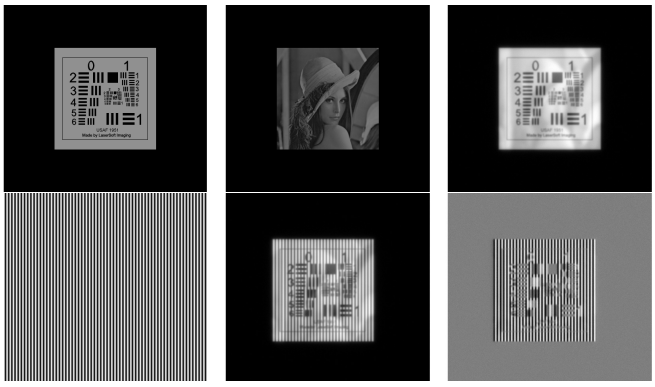

FIGURE 3 – Top row. Left panel : in-focus ground truth object. Middle panel : out-of-focus ground truth object. Right panel : conventional observation $\mathbf{i}_{CV}$. Bottom row. Left panel : one illumination pattern. Middle panel : low resolution SIM image $\mathbf{i}_l$. Right panel : differential data $\mathbf{i}_{l|OS}$.

Then, the in-focus object was degraded by an unaberrated

focused PSF that was Nyquist sampled, while the out-of-focus one was blurred with a defocused PSF with a defocus of 1.15 radian RMS. Finally, the blurred images were summed and corrupted with white Gaussian noise ; two different levels of noise were evaluated with variance corresponding to an average of 10^3 and 10^4 photons per pixel. Figure 3 depicts the ground truth objects, as well as the conventional observation $\mathbf{i}_{CV}$, a LR SIM observation $\mathbf{i}_l$ and the differential data $\mathbf{i}_{l|OS}$.

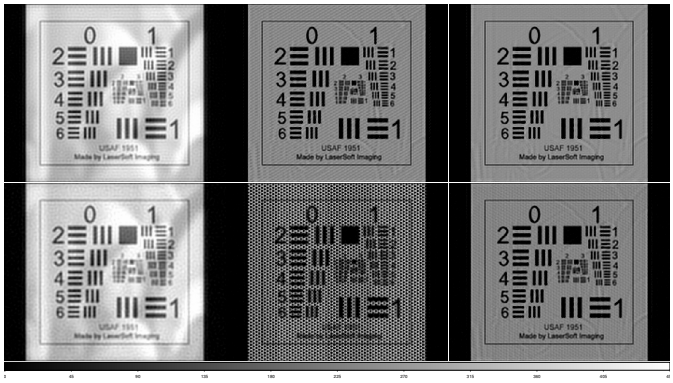

FIGURE 4 – Top row : 10^4 photons per pixel reconstructions ($SNR = 100$). Bottom row : 10^3 photons per pixel ($SNR = 31.6$). From left to right : Wiener reconstruction of the conventional image, optical sectioning reconstructions with 9, 18 images per rotation in the case of high SNR, and 18, 36 in the case of low SNR.

Figure 4 shows results obtained for the two considered levels of noise. One can see that most of the out-of-focus content has been successfully removed in the reconstructions. We have observed in some solutions the appearance of a characteristic dotted pattern that is more severe with the combination of low SNR and an insufficient number of images per rotation (Fig. 4, bottom row, middle panel). These residuals are related in the Fourier domain to the spectra created by the superposition of the three orientated fringes and they can be canceled by increasing the number of images per rotation.

Finally, in Figure 5 one can observe the efficiency in removing the out-of-focus content for several modulation pattern frequencies. Only when this is set close to the optimal value (approximately half of the instrument cut-off frequency or above), is the out-of-focus image mostly removed (Fig. 5, middle panel).

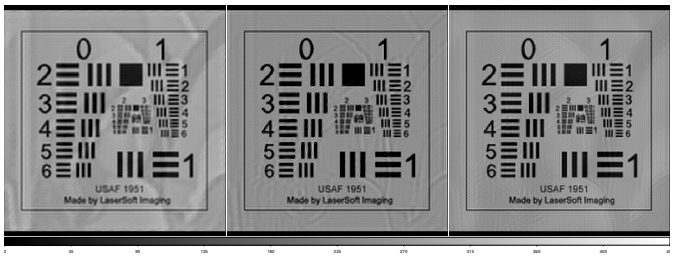

FIGURE 5 – Left panel : reconstruction with illumination pattern frequency at $\sim f_c/4$. Middle panel : $\sim f_c/2$, frequency position to achieve maximal optical sectioning. Right panel : $\sim f_c/1.33$. Reconstructions were made with 36 images per rotation and 10^4 photons per pixel ($SNR = 100$).

4 Conclusions

Structured Illumination Microscopy (SIM) is a powerful tool to achieve optical sectioning (OS) and super-resolution from wide-field fluorescence microscopy. If the characteristics of the modulation patterns are not perfectly known, as *e.g.*, in retinal imaging, where the eye motion prevents knowing the shifts between the object and the fringes, then they must be estimated. In this paper, we have presented a method which keeps the simplicity of a 2D model for estimating both the object and its shifts within the Bayesian framework. This method consists in, firstly, recording and registering a conventional observation and subtracting it from the low resolution SIM images, and secondly, performing a MAP minimization of a cost function to compute the most likely object compatible with the differential data and noise statistics. We have shown that this method is able to perform OS and is computationally efficient.

Acknowledgements

This research has been partly funded by the European Research Council through the SINERGY project HELMHOLTZ, by the ANR RHU project LIGHT4DEAF, and by the CNRS Défi Imag'In project RESSOURCES. We extend our gratitude to Dr. Michel Paques, to Dr. Loïc Denis and to Dr. Serge Meimon.

Références

[1] M.G. Gustaffson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198 :82–87, 2000.

[2] M.A.A. Neil, R. Juskaitis, and T. Wilson. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Opt. Lett.*, 22(24) :1905–1907, 1997.

[3] S. Gruppetta. Structured illumination for in-vivo retinal imaging. In *Frontiers in Optics 2013*, page FW2F.1. Optical Society of America, 2013.

[4] F. Orieux, E. Sepulveda, V. Loriette, B. Dubertret, and J.-C. Olivo-Marin. Bayesian estimation for optimized structured illumination microscopy. *IEEE Transactions on image processing*, 21 :601–614, 2012.

[5] Jost A., Tolstik E., Feldmann P., Wicker K., Sentenac A., and Heintzmann R. Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slick blind-sim reconstruction. *Plos one*, 10(7) :1–10, 2015.

[6] É. Thiébaut. Optimization issues in blind deconvolution algorithms in astronomical data analysis ii. volume 4847, pages 174–183. Proc. Soc. Photo-Opt. Instrum. Eng., 2002.

[7] L. Blanco, L. Mugnier, A. Bonnefois, and M. Paques. Registration and restoration of adaptive-optics corrected retinal images. In *Proceedings of the 2014 International Workshop on Computational Intelligence for Multimedia Understanding. IEEE*, 2014.

[8] L. Blanco and L. M. Mugnier. Marginal blind deconvolution of adaptive optics retinal images. *Opt. Expr.*, 19(23) :23227–23239, November 2011.