A 2-D SPECTRAL ANALYSIS METHOD TO ESTIMATE THE MODULATION PARAMETERS IN STRUCTURED ILLUMINATION MICROSCOPY

Laurent Condat[1*], *Jérôme Boulanger*[2†], *Nelly Pustelnik*[3], *Souleymen Sahnoun*[1], *and Lucie Sengmanivong*[2]

[1]GIPSA-lab, CNRS and University of Grenoble, F-38402 St Martin d'Hères, France.
[2]UMR144 CNRS/Institut Curie, 12 rue Lhomond F-75005 Paris, France
[3]Physics Laboratory, CNRS and ENS Lyon, F-69007 Lyon, France.
Contact: `laurent.condat@gipsa-lab.grenoble-inp.fr`

ABSTRACT

Structured illumination microscopy is a recent imaging technique that aims at going beyond the classical optical resolution limits by reconstructing a high-resolution image from several low-resolution images acquired through modulation of the transfer function of the microscope. A precise knowledge of the sinusoidal modulation parameters is necessary to enable the super-resolution effect expected after reconstruction. In this work, we investigate the retrieval of these parameters directly from the acquired data, using a novel 2-D spectral estimation method.

Index Terms— Structured illumination microscopy, SIM, sinusoidal modulation, frequency estimation, spectral analysis, super-resolution

1. INTRODUCTION

Optical microscopy has become a key research tool in modern biology. However, its resolution, which is the ability to discriminate small and close objects, is intrinsically limited by light diffraction. Because a microscope has a lowpass transfer function, the high-frequency content of the scene is lost, and the acquired image is blurry. Several techniques have been proposed to improve the resolution of microscopy; in this article, we focus on the recently developed approach of *structured illumination microscopy* (SIM) [1, 2]. The idea of SIM is to illuminate the object with a sinusoidal fringe pattern, so that the high frequency content of the imaged scene is aliased as low frequency content, which is not canceled by the transfer function of the microscope. This way, the frequency content of the acquired image is a mixture of the low and high frequency parts of the spectrum of the unknown desired image, which must be separated subsequently. Thus, several images are acquired with different modulation angles and phases, so that the disambiguation process is possible [1, 3–5]. However, the modulation frequencies and phases must be known precisely for the reconstruction to achieve the desired goal. Even a small error on these parameters can have disastrous effects, since not only the super-resolution effect is lost, but also the reconstructed image is corrupted with remaining oscillating artifacts [6]. Moreover, calibration of the device prior to acquisition, to estimate the unknown modulation parameters, is not appropriate, since even a small perturbation can change them; we recall that the unit of space considered here is a few microns. We believe that this is

the sensitivity of SIM to the modulation parameters which has hampered its wide diffusion until now. Therefore, we focus in this paper on robust and precise estimation of these parameters, directly from the acquired set of modulated images.

The paper is organized as follows. In sect. 2, we model the setup of SIM and formulate the considered problem. In sect. 3, we briefly recall the state of the art in 2-D spectral estimation and we present our new method. Finally, we present the estimation results on a real image stack.

2. IMAGING MODEL AND PROBLEM FORMULATION

In SIM, the sample to be acquired is illuminated with an incoherent sinusoidal pattern of light. Neglecting the 3-D aspects and reasoning on the image in the focal plane, the forward acquisition model is the following: $x \in \mathbb{R}^{N_1 \times N_2}$ is the unknown sought-after image, of size $N_1 \times N_2$. The set of acquired images $y_m \in \mathbb{R}^{N_1 \times N_2}$, $m = 1, \ldots M$, are such that $y_m = u_m * h + \varepsilon_m$, where $*$ denotes convolution, h is the 2-D point-spread function of the microscope, ε_m is a nuisance image modeling the random noise and the model imprecisions, and

$$u_m[k_1, k_2] = x[k_1, k_2]\big(1 + \alpha_m \cos(f_{1,m}k_1 + f_{2,m}k_2 + \phi_m)\big), \quad (1)$$

where $k_1 = 1, \ldots, N_1$ and $k_2 = 1, \ldots, N_2$ are the horizontal and vertical pixel indexes, $\alpha_m \in]0,1]$, $f_{1,m} \in [0,\pi[$, $f_{2,m} \in]-\pi,\pi[$ and $\phi_m \in [0, 2\pi[$ are the unknown amplitude, horizontal frequency, vertical frequency, and phase of the modulation pattern used to acquire the m−th image, respectively. Only the modulation angle changes, and not the frequency along the 1-D modulation direction; that is, $f_{1,m}^2 + f_{2,m}^2$ is the same for all m. Note that, in addition to the fact that the modulation parameters are unknown, the model is only approximate: the excitation pattern is never exactly sinusoidal and it can fluctuate along the time, for instance due to temperature variations.

The number M of acquired images is the product of the number of different angles M_{angle} and the number of different phases M_{phase} for each angle. So, for every $m = 1, \ldots, M_{\text{phase}}$, or $m = M_{\text{phase}} + 1, \ldots, 2M_{\text{phase}}$, and so on, the frequencies $f_{1,m}$ and $f_{2,m}$ are the same, only the phases ϕ_m are different. Moreover, theses phases are separated by the integer multiples of $2\pi/M_{\text{phase}}$. Therefore, for every $m' = 1, \ldots, M_{\text{angle}}$, we have

$$\sum_{m=(m'-1)M_{\text{phase}}+1}^{m' M_{\text{phase}}} \cos(f_{1,m}k_1 + f_{2,m}k_2 + \phi_m) = 0. \quad (2)$$

Consequently, we define the image $w_{m'}$ as

$$w_{m'} = \frac{1}{M_{\text{phase}}} \sum_{m=(m'-1)M_{\text{phase}}+1}^{m'M_{\text{phase}}} y_m \tag{3}$$

$$= x * h + \frac{1}{M_{\text{phase}}} \sum_{m=(m'-1)M_{\text{phase}}+1}^{m'M_{\text{phase}}} \varepsilon_m \approx x * h. \tag{4}$$

Thus, averaging the images obtained with same modulation angle but different phases yields approximately the image that would have been obtained with classical wide-field microscopy, without modulation. So, for every angle, we compute this image $w_{m'}$ averaged over the phases and subtract it from the corresponding images y_m: for every $m = 1, \dots, M$, we define the image $v_m = y_m - w_{m'}$, where $m' = \lfloor (m-1)/M_{\text{phase}} \rfloor + 1$. Then, the modulation and convolution do not commute, but in first approximation, we have, for every $k_1 = 1, \dots, N_1, k_2 = 1, \dots, N_2$,

$$v_m[k_1, k_2] \approx w_{m'}[k_1, k_2]\beta_m \cos(f_{1,m}k_1 + f_{2,m}k_2 + \phi_m), \tag{5}$$

for some unknown constant β_m. Thus, given the images v_m and $w_{m'}$, our goal is to find estimates of $f_{1,m}$, $f_{2,m}$ and ϕ_m by weighted spectral estimation. More precisely, we want to solve the nonlinear least-squares problem:

$$\text{Find } (\tilde{\beta}_m, \tilde{f}_{1,m}, \tilde{f}_{2,m}, \tilde{\phi}_m) = \underset{\beta>0, f_1\in[0,\pi[, f_2\in]-\pi,\pi[, \phi\in[0,2\pi[}{\arg\min} \sum_{k_1=1}^{N_1} \sum_{k_2=1}^{N_2}$$

$$\left| v_m[k_1, k_2] - w_{m'}[k_1, k_2]\beta \cos(f_1 k_1 + f_2 k_2 + \phi) \right|^2. \tag{6}$$

This parametric problem is very difficult to solve, because the cost function to minimize is nonconvex and very oscillating, with many local minima [7, 8]. Note that looking for the maximum of the amplitude of the Fourier transform of the image v_m yields estimates of the frequencies $\tilde{f}_{1,m}$ and $\tilde{f}_{2,m}$, but on the Fourier grid: these values will be multiples of $2\pi/N_1$ and $2\pi/N_2$, respectively. Such estimates are too coarse for the reconstruction of the super-resolved image to be effective; that is why we want to use a *high resolution* spectral estimation technique [7], which is able to retrieve the parameters exactly in absence of noise. So, in the next section, we extend to the 2-D case the original 1-D approach developed recently by the first author in [8], which achieves maximum-likelihood estimation even in presence of a significant amount of noise.

3. HIGH RESOLUTION ESTIMATION OF THE MODULATION PARAMETERS

3.1. State of the art

Spectral analysis consists in retrieving the unknown amplitudes, phases and frequencies of a sum of sinusoids or complex exponentials, from noisy samples [7]. In 1-D, there is a vast literature on spectral analysis and well-established *subspace methods* like MUSIC or ESPRIT; they are fast but suboptimal in comparison with maximum-likelihood estimation, which is believed NP-hard in general [8]. Spectral analysis in 2-D has been investigated for many years [9] and some of the 1-D subspace methods have been extended to 2-D [10–12]. In our context, these methods are not applicable, because 1) they do not perform weighted spectral estimation; that is, they do not account for the presence of the weight image $w_{m'}$ in the fit (6); 2) they require storage and SVD computations of very large matrices, of size around $N_1.N_2 \times N_1.N_2$, which is not practicable for $N_1 = N_2 = 512$.

3.2. Proposed method

The proposed approach takes its roots in the *annihilation* property, which dates back to Prony's work in the eighteenth century [13]: consider a sequence $v = \{v_n\}_{n=1}^N$ made of $N \geq 3$ uniform samples of a sinusoid; that is, for every $n = 1, \dots, N$, $v_n = \eta \cos(\xi n + \varphi)$, for some real parameters η, ξ, φ. Then, there exists a symmetric sequence $h = \{h_1, h_2, h_3\}$, with $h_3 = h_1$, called the annihilating filter, such that the convolution of the sequences v and h is identically zero: $\sum_{k=1}^3 h_k v_{n+k-1} = 0$, for every $n = 1, \dots, N-2$. Thus, the proposed method consists in denoising the noisy image v_m, so that it matches the form in the right-hand side of (5), by finding the sinusoidal image $\tilde{c}$, with $\tilde{c}[k_1, k_2] = \cos(\tilde{f}_{1,m}k_1 + \tilde{f}_{2,m}k_2 + \tilde{\phi}_m)$ for every k_1, k_2, solution to (6). Then, we easily find the two annihilating filters $\tilde{h}_1$ and $\tilde{h}_2$, both symmetric and of size three, which annihilate every row and column of $\tilde{c}$, respectively. The two roots of the polynomial $\tilde{h}_{1,1}z^2 + \tilde{h}_{1,2}z + \tilde{h}_{1,3}$ are, with very high probability, on the complex unit circle, and are $e^{+j\tilde{f}_{1,m}}$ and $e^{-j\tilde{f}_{1,m}}$. We retrieve the frequency $\pm \tilde{f}_{2,m}$ from $\tilde{h}_2$, the same way. Finally, using a linear regression, we find the complex amplitude $\tilde{\beta}_m e^{j\tilde{\phi}_m}$ and disambiguate the sign of $\tilde{f}_{2,m}$. We refer to [8] for more details on this procedure, which yields the estimates $\tilde{f}_{1,m}, \tilde{f}_{2,m}, \tilde{\phi}_m, \tilde{\beta}_m$ from the sinusoidal image $\tilde{c}$. In the sequel, we describe the core of our approach, which is how the image $\tilde{c}$ is computed.

First, let us define the orthogonal operator T_1, which maps an image of size $N_1 \times N_2$, with $N_1 \geq 5$, viewed as a matrix of size $N_2 \times N_1$, to a centro-symmetric Toeplitz matrix of size $6N_2 \times N_1 - 2$ as follows: each row of the image is mapped to a Toeplitz matrix of size $3 \times N_1 - 2$ and these matrices are stacked on top of each other. Finally, the obtained matrix is duplicated, the order of the rows and columns is reversed, and this duplicated matrix is stacked under the first one. A small example is the following:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \end{bmatrix} \overset{T_1}{\mapsto} \begin{bmatrix} a_{1,3} & a_{1,4} & a_{1,5} \\ a_{1,2} & a_{1,3} & a_{1,4} \\ a_{1,1} & a_{1,2} & a_{1,3} \\ \hline a_{2,3} & a_{2,4} & a_{2,5} \\ a_{2,2} & a_{2,3} & a_{2,4} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ \hline\hline a_{2,3} & a_{2,2} & a_{2,1} \\ a_{2,4} & a_{2,3} & a_{2,2} \\ a_{2,5} & a_{2,4} & a_{2,3} \\ \hline a_{1,3} & a_{1,2} & a_{1,1} \\ a_{1,4} & a_{1,3} & a_{1,2} \\ a_{1,5} & a_{1,4} & a_{1,3} \end{bmatrix} \tag{7}$$

We also define the matrix p of size $6N_2 \times N_1 - 2$, whose every entry is the inverse of the number of times the entry at same position in the output of T_1 has been repeated. For the small example above, one has

$$p = \begin{bmatrix} 1/6 & 1/4 & 1/2 & 1/6 & 1/4 & 1/2 & 1/6 & 1/4 & 1/2 & 1/6 & 1/4 & 1/2 \\ 1/4 & 1/6 & 1/4 & 1/4 & 1/6 & 1/4 & 1/4 & 1/6 & 1/4 & 1/4 & 1/6 & 1/4 \\ 1/2 & 1/4 & 1/6 & 1/2 & 1/4 & 1/6 & 1/2 & 1/4 & 1/6 & 1/2 & 1/4 & 1/6 \end{bmatrix}^{\mathrm{T}} \tag{8}$$

We also define the operator $T_2 : v \mapsto T_1(v^{\mathrm{T}})$. Then, one can show [7] that if c is an image of a sinusoid, then $T_1(c)$ and $T_2(c)$ are both of rank only two. Hence, our approach consists in computing the matrices $\tilde{t}_1$ and $\tilde{t}_2$ of rank 2, such that $\tilde{t}_1 = T_1(\tilde{c})$ and $\tilde{t}_2 = T_2(\tilde{c})$, where $\tilde{c}$ is the desired sinusoidal image. That is, we we want to solve the *structured low-rank approximation* problem [14]:

Considered matrix approximation problem:
Find the pair of matrices $(t_1, t_2) \in \underset{t_1, t_2 \in \mathbb{R}^{6N_2 \times N_1 - 2}}{\arg\min}$

$$\left\| (\sqrt{p}, \sqrt{p}) \circ \big(T_1(w_{m'}) \circ t_1 - T_1(v_m), T_2(w_{m'}) \circ t_2 - T_2(v_m)\big) \right\|_F^2$$
s.t. $(t_1, t_2) \in \mathcal{T}$, $\mathrm{rank}(t_1) = 2$, $\mathrm{rank}(t_2) = 2$, \hfill (9)

where $\circ$ is the Hadamard entrywise product of matrices, the square root is taken entrywise, $\| \cdot \|_F$ is the Frobenius norm, and we define the set $\mathcal{T} = \{(T_1(c), T_2(c)) \; : \; c \in \mathbb{R}^{N_2 \times N_1}\}$. The orthogonal projection $P_\mathcal{T}$ of a pair of matrices (t_1, t_2) onto $\mathcal{T}$ is simply done by averaging and the orthogonal projection $P_{\mathcal{R}_2}$ of a matrix onto the set of matrices of rank at most 2 is done by SVD truncation; see more details in [8] for these two operations. To solve the problem (9), we alternate between a gradient descent with respect to the cost function, which is the squared Frobenius norm, and the two projections $P_\mathcal{T}$ and $P_{\mathcal{R}_2}$ to enforce the constraints. The iterative algorithm is the following:

Proposed algorithm to solve (9):
Choose the parameters $\mu > 0$, $\gamma \in \,]0,1[$, and the initial estimates $t_1^{(0)}, t_2^{(0)}, s_1^{(0)}, s_2^{(0)}$. Then iterate, for every $l \geq 0$,

$$\left|
\begin{aligned}
t_1^{(l+1)} &= P_{\mathcal{R}_2}\Big(s_1^{(l)} + \gamma(t_1^{(l)} - s_1^{(l)}) - \mu\, p \circ T_1(w_{m'}) \circ \\
&\quad \big(T_1(w_{m'}) \circ t_1^{(l)} - T_1(v_m)\big)\Big), \\
t_2^{(l+1)} &= P_{\mathcal{R}_2}\Big(s_2^{(l)} + \gamma(t_2^{(l)} - s_2^{(l)}) - \mu\, p \circ T_2(w_{m'}) \circ \\
&\quad \big(T_2(w_{m'}) \circ t_2^{(l)} - T_2(v_m)\big)\Big), \\
(s_1^{(l+1)}, s_2^{(l+1)}) &= (s_1^{(l)}, s_2^{(l)}) - (t_1^{(l+1)}, t_2^{(l+1)}) + \\
&\quad P_\mathcal{T}\big(2(t_1^{(l+1)}, t_2^{(l+1)}) - (s_1^{(l)}, s_2^{(l)})\big).
\end{aligned}
\right.$$

The algorithm can be shown to converge when it is applied to solve problems with two convex constraint sets [8], but in our case, the set $\mathcal{R}_2$ is not convex, so that there is no guarantee of convergence. But we found out empirically that the algorithm always yields the expected solution, for an appropriate choice of the parameters μ and γ. For our experiments, we use $\mu = 0.1$, $\gamma = 0.5$, and we initialize the images $t_1^{(0)}, t_2^{(0)}, s_1^{(0)}, s_2^{(0)}$ as identically zero. A few iterations are enough to yield an estimate of the sinusoidal image $\tilde{c}$, but about 1,000 iterations are necessary to achieve convergence within machine precision, which takes several minutes.

We tested our method on a real dataset, see fig. 1 and its caption for illustrations and details. The estimated frequencies are summarized in the table below. The std. dev. of the frequencies around their mean for the corresponding angle is 5e-4. $\tilde{f}_{1,m}^2 + \tilde{f}_{2,m}^2$ is almost the same for every m, around 0.7485 with std. dev. 2e-3.

	$m = 1, \ldots, 5$				
$\tilde{f}_{1,m}$	0.8415	0.8415	0.8407	0.8416	0.8412
$\tilde{f}_{2,m}$	0.2004	0.2005	0.1993	0.2006	0.1997
	$m = 6, \ldots, 10$				
$\tilde{f}_{1,m}$	0.5955	0.5953	0.5955	0.5958	0.5954
$\tilde{f}_{2,m}$	-0.6296	-0.6301	-0.6293	-0.6285	-0.6291
	$m = 11, \ldots, 15$				
$\tilde{f}_{1,m}$	0.2466	0.2465	0.2457	0.2459	0.2458
$\tilde{f}_{2,m}$	0.8289	0.8295	0.8273	0.8285	0.8283

4. CONCLUSION

We proposed a new algorithm for 2-D weighted spectral estimation, based on a annihilation technique robust to noise, and applied it to retrieve the parameters of the sinusoidal modulation in images acquired by structured illumination microscopy. Future work will include theoretical and experimental performance evaluation and comparison with other methods. Moreover, the impact of the estimation quality of the modulation parameters on the quality of the reconstructed super-resolved image, using for instance the method proposed in our companion paper [15], remains to be studied. A Matlab implementation of the method is available on the webpage of the first author.

5. REFERENCES

[1] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.*, vol. 198, no. 12, pp. 82–87, May 2000.

[2] M. G. L. Gustafsson, L. Shao, P. M. Carlton, C. J. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys. J*, vol. 94, no. 12, pp. 4957–4970, June 2008.

[3] J. Conchello, "Image estimation for structured-illumination microscopy," in *Proc. of SPIE*, vol. 5701, 2005.

[4] F. Orieux, E. Sepulveda, V. Loriette, B. Dubertret, and J.-C. Olivo-Marin, "Bayesian estimation for optimized structured illumination microscopy," *IEEE Trans. Image Processing*, vol. 21, no. 2, pp. 601–614, Feb. 2012.

[5] D. Stoker and E. Matlin, "An estimation theoretic framework for structured illumination microscopy," in *Imaging and Applied Optics Technical Papers, Proc. of Optical Fabrication and Testing*, 2012.

[6] L. H. Shaefer, D. Schuster, and J. Schaffer, "Structured illumination microscopy: artefact analysis and reduction utilizing a parameter optimization approach," *Journal of Microscopy*, vol. 216, pp. 165–174, Nov. 2004.

[7] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Prentice Hall, NJ, 2005.

[8] L. Condat and A. Hirabayashi, "Cadzow denoising upgraded: A new projection method for the recovery of Dirac pulses from noisy linear measurements," 2012, preprint hal-00759253.

[9] H. Kimura and Y. Honoki, "A hybrid approach to high resolution two-dimensional spectrum analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, pp. 1024–1036, July 1987.

[10] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Processing*, vol. 40, pp. 2267–2280, Sept. 1992.

[11] S. Rouquette and M. Najim, "Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods," *IEEE Trans. Signal Processing*, vol. 49, no. 1, 2001.

[12] L. Huang, Y. Wu, H. C. So, Y. Zhang, and L. Huang, "Multidimensional sinusoidal frequency estimation using subspace and projection separation approaches," *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5536–5543, 2012.

[13] T. Blu, "The generalized annihilation property—A tool for solving finite rate of innovation problems," in *Proc. of Int. Workshop on Sampling Theory and Appl. (SampTA)*, Marseille, France, May 2009.

[14] I. Markovsky, *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer, 2012.

[15] J. Boulanger, N. Pustelnik, and L. Condat, "Non-smooth convex optimization for an efficient reconstruction in structured illumination microscopy," in *Proc. of ISBI*, Beijing, China, 2014.
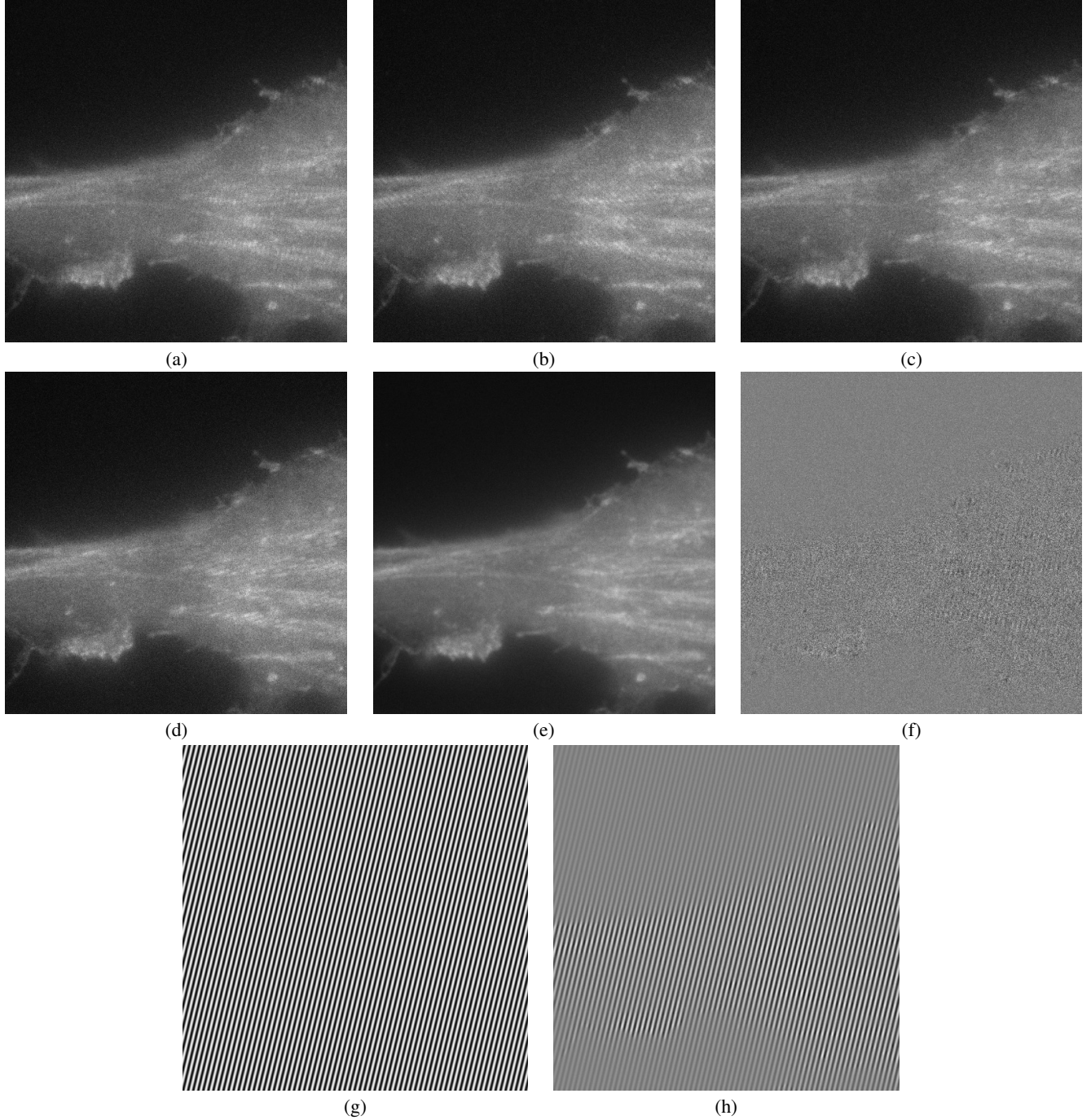
Fig. 1. (a)–(d): the four images y_1, y_6, y_{11}, y_{12} of a data set of fifteen images, with $M_{\mathrm{angle}} = 3$ angles and $M_{\mathrm{phase}} = 5$ phases for each angle. So, the images in (a)–(c) have been acquired with different modulation angles and the image in (d) has been acquired with the same angle as the image in (c), but with a different phase. The microscope is a Nikon N–SIM and the image sample is a LifeAct (mCherry) stably transfected RPE fixed cell. The images are of size 512×512. (e): the image w_1, which is the average of the five images y_1, y_2, y_3, y_4, y_5, acquired with same modulation angle but different phases. (f): the difference image v_1 between the image y_1 in (a) and the average image w_1 in (e). (g): the computed image $\tilde{c}$ solution to (6). (h): the entrywise product $w_1 \circ \tilde{c}$, which can be viewed as a denoised version of the image v_1 in (f), in the sense of the model (5). The images in (f), (g), (h) have been normalized, so that zero values are represented by a middle gray.