# Snap-shot profilometry with the Empirical Mode Decomposition and a 3-layer color sensor

Sébastien Equis,[1,3] Raik Schnabel,[1] and Pierre Jacquot[1,2]

[1]Nanophotonics and Metrology Laboratory, Swiss Federal Institute of Technology Lausanne
EPFL-STI-NAM, Station 11, CH-1015 Lausanne, Switzerland
[2]pierre.jacquot@epfl.ch
[3]sebastien.equis@epfl.ch

**Abstract:** Remote sensing finds more and more applications, from industrial control, to face recognition, not forgetting terrain surveying. This trend is well exemplified by fringe projection techniques, which enjoyed a considerable development in the recent years. In addition of high requirement in terms of measurement accuracy and spatial resolution, the end-users of full-field techniques show a growing interest for dynamic regimes. We report here what we believe to be the use for the first time of a CMOS 3-layer color sensor (Foveon X3) as the key element of a RGB fringe projection system, together with the processing specifically elaborated for this sensor. The 3-layer architecture allows the simultaneous recording of three phase-shifted fringe patterns and features the precious asset of an unambiguous relationship between the physical sensor pixel and the picture pixel and this for each color layer, on the contrary of common color sensor arrays (Bayer mosaic and tri-CCD). Due to the overlapping of the spectral responses of the layers, color transformation is mandatory to achieve the separation of the three phase-shifted RGB projected fringe patterns. In addition, we propose the use of the Empirical Mode Decomposition to equalize the non-uniform responses of the three layers. Although the conversion of the phase into a height is of primary importance in an actual measurement, it is not treated here, the literature being profuse on the central projection model.

©2011 Optical Society of America

## References and links

1. F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," Opt. Eng. **39**(1), 10–22 (2000).
2. E. Stoykova, A. A. Alatan, P. Benzie, N. Grammalidis, S. Malassiotis, J. Ostermann, S. Piekh, V. Sainov, C. Theobalt, T. Thevar, and X. Zabulis, "3-D time varying scene capture technologies – a survey," IEEE Trans. Circ. Syst. Video Tech. **17**(11), 1568–1586 (2007).
3. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," Opt. Laser Technol. **48**, 149–158 (2010).
4. P. S. Huang, C. Zhang, and F.-P. Chiang, "High-speed 3-D shape measurement based on digital fringe projection," Opt. Eng. **42**(1), 163–168 (2003).
5. K. Creath, *Interferogram Analysis* (Institute of Physics Publishing, Bristol, UK, 1993), Chap. 4.
6. P. Groot, "Derivation of algorithms for phase-shifting interferometry using the concept of a data-sampling window," Appl. Opt. **34**(22), 4723–4730 (1995).
7. D. W. Phillion, "General methods for generating phase-shifting interferometry algorithms," Appl. Opt. **36**(31), 8098–8115 (1997).
8. Y. Surrel, "Design of algorithms for phase measurements by the use of phase stepping," Appl. Opt. **35**(1), 51–60 (1996).
9. J. Novak, P. Novak, and A. Miks, "Multi-step phase-shifting algorithms insensitive to linear phase shift errors," Opt. Commun. **281**(21), 5302–5309 (2008).
10. X. Colonna de Lega, and P. Jacquot, "Deformation measurement with object-induced dynamic phase shifting," Appl. Opt. **35**(25), 5115–5121 (1996).
11. C. Wust, and D. W. Capson, "Surface profile measurement using color fringe projection," Mach. Vis. Appl. **4**(3), 193–203 (1991).

12. http://www.foveon.com.
13. P. Tankam, P. Picart, D. Mounier, J. M. Desse, and J. C. Li, "Method of digital holographic recording and reconstruction using a stacked color image sensor," Appl. Opt. **49**(3), 320–328 (2010).
14. http://www.alt-vision.com/.
15. S. Lei, and S. Zhang, "Digital sinusoidal fringe pattern generation: defocusing binary patterns vs focusing sinusoidal patterns," Opt. Laser Technol. **48**, 561–569 (2010).
16. J. A. N. Buytaert, and J. J. J. Dirckx, "Phase-shifting moiré topography using optical demodulation on liquid crystal matrices," Opt. Lasers Eng. **48**(2), 172–181 (2010).
17. C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," J. Appl. Photogr. Eng. **2**(3), 95–99 (1976).
18. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," Proc. R. Soc. Lond. A **454**(1971), 903–995 (1998).
19. P. Flandrin, *http://perso.ens-lyon.fr/patrick.flandrin/emd.html*.
20. S. Equis, "Phase extraction of non-stationary signals produced in dynamic interferometry involving speckle waves," EPFL thesis n° **4514**, Lausanne, (2009). *http://biblion.epfl.ch/EPFL/theses/2009/4514/EPFL_TH4514.pdf*.
21. S. Equis, and P. Jacquot, "Phase extraction in dynamic speckle interferometry by empirical mode decomposition and Hilbert transform," Strain **46**(6), 550–558 (2010).
22. S. Equis, and P. Jacquot, "The empirical mode decomposition: a must-have tool in speckle interferometry?" Opt. Express **17**(2), 611–623 (2009).
23. H. A. Aebischer, and S. Waldner, "Simple and effective method for filtering speckle interferometric phase fringe patterns," Opt. Commun. **162**(4-6), 205–210 (1999).

## 1. Introduction

The intersection of a light plane with a certain observation line uniquely determines a point in a 3D space. Fringe projection techniques [1–4] are exploiting and parallelizing this triangulation principle to eventually give access to the surface topology of an object. The most widespread technique to retrieve unambiguously this surface height information requires at least three frames with corresponding phase-shifted fringe patterns and uses one of the numerous available phase-shifting algorithms [5–9]. Though acknowledged to be the soundest technique in terms of spatial resolution and measurement accuracy, it necessitates the capture of several images at different instants. It thus reduces greatly the temporal bandwidth of the measurement, as the phase change (i.e. the sum of the object phase change and the external phase increment) between consecutive frames must not exceed $\pi$ to fulfill the sampling Nyquist criterion [10]. To avoid complex setups, the idea of using color fringes for surface profile measurement was proposed [11] in the early nineties. The three phase-shifted patterns are not sequentially grabbed on a monochrome camera but coded over three colors and consequently captured by a color sensor.

Foveon [12] has conceived a CMOS sensor which takes advantage of the dependency of the absorption law of silicon with the light wavelength. The higher the wavelength of the incident light, the higher the penetration depth into the stack. Thus, with photodiodes stacked and embedded into a slab of silicon, a tricolor response is thus available while the spatial resolution is preserved and rigorously the same for the three channels. Due to the spectral overlapping between the three layers, a color transformation step is required to correctly separate the three channels. To avoid tedious procedures of calibration which consist in a laborious tuning of the gains of the different projected colors to obtain equal responses on the three channels of the sensor, we propose here the use of the Empirical Mode Decomposition (EMD) as an amplitude equalizing tool so that a phase-shifting algorithm can be meaningfully applied. The surface topology is thus recovered with one single image.

The paper is organized as follows: after some words about the Foveon X3 sensor, the processing steps, prior to the phase extraction strictly speaking, are then detailed. The EMD is then briefly described. Experimental results will be finally presented and discussed.

## 2. Layered CMOS camera

The X3 Foveon sensor stacks red, green and blue pixels vertically, increasing the information density of the recorded image – or in other words, features a clear correspondence between the sensor and the image pixels - while simultaneously eliminating the color sampling artifacts

found for instance in classical color sensor using Bayer mosaic (Fig. 1). This opens up the range of applications of full-field optical methods, e.g [13]. The camera we used is a development kit camera from Alternative Vision [14]. It is able to record single images and sequences as well, in 8 or 12 bits. The resolution sensor is 2652 × 1768 for each layer with a 7.8 µm pitch, and the frame rate of 5 fps in full resolution can be accordingly increased if the area of interest is reduced and/or if pixel binning is used up to 30 fps. As projection device, we used the model EMP 821 3LCD from EPSON.
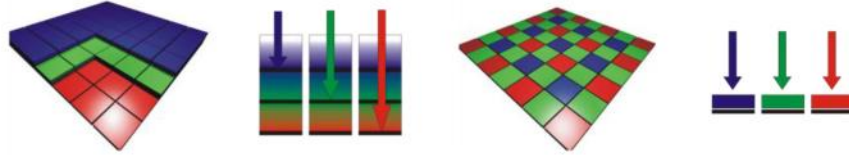


Fig. 1. Layered CMOS sensor vs. conventional Bayer matrix sensor.

## 3. Fringe projection and phase computation

In [15] the authors demonstrate that the phase error is generally lower with a defocused binary pattern than with a focused sinusoidal one. In the case where high sensitivity is sought, small apertures are acceptable and above all with grayscale fringe patterns, this alternative is undoubtedly to be considered. In our case, with rather large aperture of f/4, projection angles in the range of 10°-20°, and colored patterns, the sine profile is to be preferred. Consequently, similarly to what has been recently proposed in [16] in the framework of Moiré topography, the non-linearities of the video projector response have to be corrected − the gamma of such projector is indeed not unitary to compensate for the intrinsic non-linearities of the human eye.

The projected compound color grid is the superimposition of three sinusoidal grids of the same period, shifted by one third of the period to each other, and respectively made of a varying degree of saturation of pure red (0-255,0,0), green (0,0-255,0) and blue (0,0,0-255) RGB colors, as shown in Fig. 2.
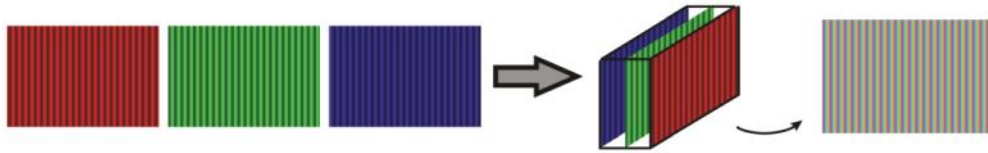


Fig. 2. Colored fringe pattern resulting from the superimposition of three spatially shifted fringe patterns.

Ideally, the observation system would record simultaneously three independent images, delivered by each of the three layers of the color sensor, whose respective intensities, $s_{R,G,B}$, would obey the following generic equations:

$$s_{R,G,B}(\mathbf{r}) = b(\mathbf{r}) + m(\mathbf{r})\cos\varphi_{R,G,B}(\mathbf{r}); \; \varphi_R(\mathbf{r}) + 2\pi/3 = \varphi_G(\mathbf{r}) = \varphi_B(\mathbf{r}) - 2\pi/3 \quad (1)$$

In parallel projection and telecentric observation (see Fig. 3), with fringes parallel to the $y$ axis, projected in a direction making an angle $\alpha$ with the $z$-axis, and of apparent period $p$ in the $z$-observation direction, the relationship between $\varphi$ and $h$ is simply:

$$\varphi(x,y) = \left(2\pi / p\right)\left(x + h(x,y)\,\mathrm{tg}\,\alpha\right) + \varphi_0, \quad (2)$$

where $\varphi_0$ accounts for the position of the projected fringes with respect to the origin of the coordinates.
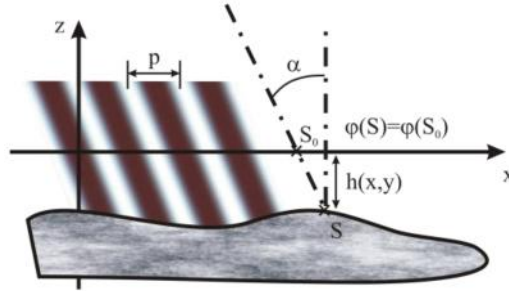
Fig. 3. Principle of the fringe projection method.

The system of Eqs. (1) – now concomitant and not sequential – is typical of a 3-image phase-shifting algorithm and yields:

$$\varphi(x, y) = \arctan\left(\sqrt{3}\ \frac{s_R(x, y) - s_B(x, y)}{2s_G(x, y) - s_R(x, y) - s_B(x, y)}\right) \tag{3}$$

In fact, Eq. (1) can be obtained only after a dedicated processing of the rough signals acquired on each layer, which makes the core of the submission.

## 4. Linearization and color transformation

As for other devices [16], a linearization step is first mandatory. From the three images of a projected linear gray levels chart onto a white diffusing screen, a look-up table (LUT) is built for each layer so as to allocate to each pixel a new value which reinstates the wanted linearity (Fig. 4). The so-built LUTs incorporate the non-linearities of both the video projector and the CMOS camera, and are thus valid only for this pair of devices. They are established once for all and do not depend on the geometric (projection angle, fringe period) and optical parameters (aperture, magnification) of the set-up. Though our experiments involved only objects of nearly uniform color and reflectivity, this step being nothing else but an input-output gray level transformation, does not depend in principle on the nature of the object surface.
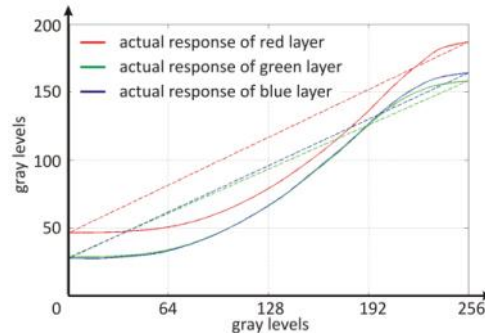


Fig. 4. Sensor layers responses to a linear gradient (plain curves) and linear counterparts (dashed lines).

The architecture of the sensor gives rise to serious overlaps among the spectral response of each layer: while peaking respectively in the red, green and blue parts of the spectrum, their spectral sensitivity curves extend over a band of definite width. The pure red, corresponding to the (1,0,0) point in the RGB space, will induce a response of each layer and not only of the red layer. The overall result is that each layer captures a fringe pattern which is after linearization a linear combination of the original patterns. It results in a slight phase-shift from

the original components and an altered amplitude. Considering an excitation $X_k=(x_R,x_B,x_G)$, the sensor gives a response $R_k=(r_R,r_G,r_B)$, which is related to the stimulus $X_k$ through:

$$R_k = S \cdot X_k ,\qquad(4)$$

where "." denotes the inner product and $S$ the sensor response matrix. The sought matrix to carry out the color transformation is the inverse of $S$. Solving Eq. (4) necessitates at least nine equations, but in practice it is preferable to consider $N$ excitations ($N > 3$) and solve the following over-determined system:

$$T=\begin{pmatrix} R_1 & \ldots & R_N \end{pmatrix}\backslash\begin{pmatrix} X_1 & \ldots & X_N \end{pmatrix},\qquad(5)$$

where "\" designates the solving of the system of $N$ equations in a least square sense (e.g. left divide function in MATLAB) and $T$ the sought-after color transformation matrix.

The GretagMacbeth ColorChecker [17] is a widely used color chart for imaging system calibration and assessment. To solve Eq. (5), we created a digital version of the GretagMacbeth chart where each of the 24 squares has coordinates in standard RGB space. Each square of the GretagMacbeth color chart constitutes itself an excitation $X_k$, and the matrix $T$ of Eq. (5) is thus obtained by solving an over-determined system of 72 equations. The raw image grabbed by the camera is shown in Fig. 5, jointly with its linearized dark-subtracted counterpart and the final image resulting of the color transformation.



Fig. 5. Raw image of the projected digital GretagMacbeth color chart (left); outcome of the linearization and dark frame removal step (middle); result after subsequent color transformation (right).

The matrix solution for converting the linearized, dark subtracted image from the sensor color space to the standard RGB space is given below.

$$T=\begin{pmatrix} 1.00 & -0.82 & 0.09 \\ -0.58 & 1.27 & -0.36 \\ 0.48 & -1.75 & 1.50 \end{pmatrix}\qquad(6)$$

The spectral overlapping of the three layers is illustrated by the large coefficients off-diagonal. The color-transformed images necessitate a final step prior to phase computation due to the background and modulations intensities discrepancies between the RGB signals.

## 5. The Empirical Mode Decomposition

The Empirical Mode Decomposition has been developed a decade ago to address [18], jointly with the Hilbert spectral analysis, the thorny problem of the characterization of signals which are fulfilling neither the stationarity nor the linearity conditions. The EMD technique, fed with any non-stationary signal, yields a sparse set of functions, namely the intrinsic mode functions (IMF) of the signal. The starting idea is to consider the signal $s(t)$ constituted by a detail part (local high frequency) $d(t)$, and a residue part (local low frequency) $m(t)$. The detail part is sifted out from the raw signal by removing the mean envelope, whose computation is based on a cubic spline fitting between the signal extrema. The interpolation with a cubic spline kernel is acknowledged to have the best performances in most cases. The residue is then

considered itself as a signal to process and thus split into a detail and a residue part as well. The final algorithm [19] is represented in Fig. 6:
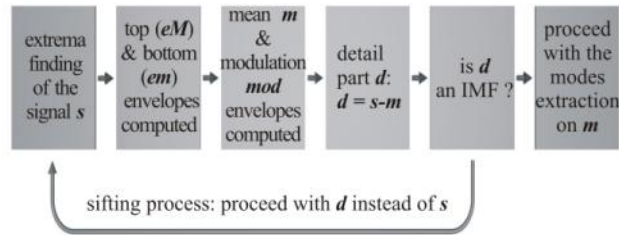


Fig. 6. Classical EMD algorithm.

We get in fine the following final decomposition at the rank K:

$$s(t) = \sum_{k=1}^{K} d_k(t) + m_K(t), \qquad (7)$$

where the $d_k(t)$ are the extracted IMFs and $m_K(t)$ is the final residue. The decomposition can be stopped at any rank K depending on the purpose, or in other words, depending on the frequency band the sought-after information belongs to (detrending, denoising, texture extraction and so on). If a total decomposition is wished, the procedure is ended when the current residue contains less than three extrema.

In [20–22], in the specific context of dynamic speckle interferometry, a dedicated implementation of EMD has been proposed. In one single iteration of the sifting process, the fluctuating background is removed, leaving a zero-mean centered signal, lending itself to a straightforward phase extraction. The sine qua non condition to allow such a fast extraction of the mode is a clear separation of the spectral ranges of the phase on one hand, and of the modulation and the background on the other hand. As discussed in detail in [20], these conditions can be usually fulfilled by a proper choice of the experimental parameters.

In the present framework, by analogy, we chose experimental conditions (projection angle and fringe pitch) so that the part of the spectrum which contains the height information of the object – centered around the carrier frequency – is clearly separated from the one of the reflectivity and colors fluctuations. A plaster model ($40 \times 40$ cm$^2$) with a cushion shape is illuminated with a tricolor fringe pattern. Figure 7 illustrates the way the EMD works on one portion of a row of the area of interest. The considered line is emphasized by a thick red line. The raw profiles from the three layers are shown in the top right graph, while the fully-processed signals – as a reminder, after linearization, dark-frame subtraction, color transformation and EMD – are shown below.
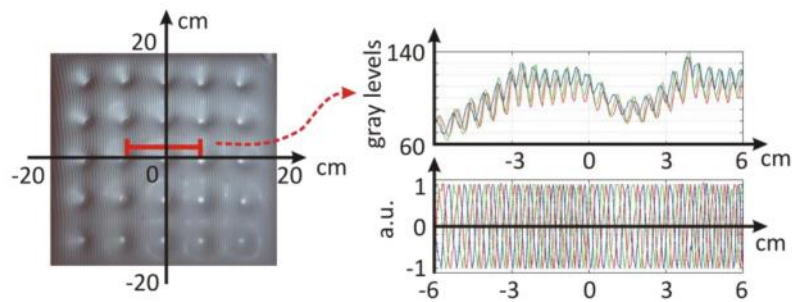


Fig. 7. Plaster model illuminated with color fringe pattern (left); cross-sections of fringe patterns before (top right) and after the full processing (bottom right) of the portion designated by the thick red line.

## 6. Experimental results, discussion and outlooks

The presented processing scheme applied on the fringe pattern shown on the left of Fig. 7 yields the phase map of Fig. 8 (left). On a common computer equipped with a Core2 2.66 GHz processor, the whole procedure starting with the raw image till the extraction of the raw phase, takes 2 minutes for a 8 bits image. For the sake of visibility, the raw phase map is shown with the carrier removed (by a best plane fitting) and rewrapped. Cross-sections of the raw and the filtered phase are also compared. The used filter is the well-known iterative sine cosine filter [23] implemented here with a 5×5 box kernel and iterated 10 times. A 3D view of the filtered phase map is added which reproduces faithfully the topology of the test object.
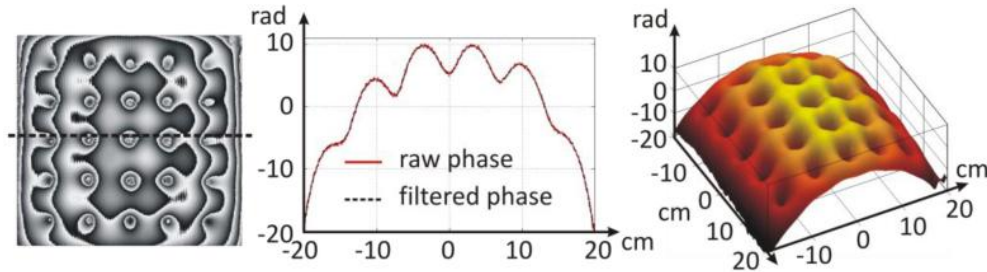


Fig. 8. (left to right) Raw phase map, cross-sections of the unwrapped raw and filtered phases and 3D view of the filtered phase.

We have shown here the feasibility of carrying out snap-shot profilometry with the layered CMOS photosensor from Foveon. So as to address the issues inherent to this new technology, we propose a road map which yields promising results. The 3-frames phase-shifting algorithm is known to be adversely affected by harmonics and inaccurate phase-shifts. A careful calibration of the projection and acquisition chain is thus mandatory (linearization and color-transformation). The EMD proves again its versatility as it allows to normalize the three layers responses accurately, efficiently and last but not least adaptively. Even if the technique should be able, in principle, to cope with chromaticity non-uniformities of the object, the point has to be assessed experimentally, just as a thorough error analysis remains to be done. More advanced and robust algorithms can be used [10] with the constraint of a higher required number of phase-shifted fringe patterns, with still a gain of three in the number of frames to capture, compared to classical fringe projection systems using monochrome cameras.

Further analyses of the sensor performances are currently in progress, along two principal directions: *i)* to know whether, in addition of being three times faster, the 3-layer color sensor and its dedicated processing is more accurate than the standard 3-image phase-shifting procedure based on three time-separated acquisitions, *ii)* to discuss up to which extent the system can cope with color objects. Our first experiments do not show at least any evident discrepancies between the phase maps obtained with the new method and those produced, everything else kept as far as possible identical, by a black&white CCD camera or by the Foveon camera used sequentially and separately in each of its three layers. As for the second concern, we observed at the qualitative stage that color objects can be measured as well. A quite obvious condition must however be fulfilled: the object under analysis must not contain colors with one null RGB coordinate, so that after color-transformation, information remains from each phase-shifted pattern.

### Acknowledgments