# Dynamic speckle analysis using multivariate techniques

**José M López-Alonso[1], Javier Alda[1], Héctor Rabal[2], Eduardo Grumel[2] and Marcelo Trivi[2]**

[1] Applied Optics Complutense Group. University Complutense of Madrid. C/ Arcos de Jalón, 118. E-28037 Madrid. Spain
[2] Centro de Investigaciones Ópticas (CONICET La Plata—CIC), UID Optimo, Dpto. Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata, Casilla de Correo 3, 1897 Gonnet- La Plata, Argentina

E-mail: javier.alda@ucm.es

## Abstract

In this work we use principal components analysis to characterize dynamic speckle patterns. This analysis quantitatively identifies different dynamics that could be associated to physical phenomena occurring in the sample. We also found the contribution explained by each principal component, or by a group of them. The method analyzes the paint drying process over a hidden topography. It can be used for fast screening and identification of different dynamics in biological or industrial samples by means of dynamic speckle interferometry.

Keywords: speckle, image processing, instrumentation, metrology, dynamic speckle

## 1. Introduction

Dynamic speckle patterns inform about the temporal characteristics of the underlying physical phenomena in the sample. Usually, several phenomena concur and contribute to the final observed intensity variations, including scattering, Doppler effect, polarization rotation, etc [1]. Since the same effects can be due to one or more of these origins, it is difficult to identify the actual mechanisms at work, and the quantification of their individual contributions.

Several techniques for analyzing dynamic speckle patterns have been proposed in the literature [1]. Many of them are heuristic in origin and others are specific to some applications. Recently, some artificial intelligence tools have been developed. They include both supervised [2] and non-supervised learning algorithms [3]. They make use, simultaneously, of several measuring techniques for screening some effects. Nevertheless, the problem is not fully solved and the identification and segmentation of regions showing different inner dynamics is of considerable interest, for example, in biology and agronomy. In this work we analyze sequences of images of dynamic speckle using a multivariate technique: the principal components analysis (PCA) [4].

This method can distinguish processes occurring in a video sequence which is characterized by different correlation structures. The computational resources and time necessary to perform PCA are proportional to the data volume (number of frames × number of pixels at each frame) and it is in the range of tens of seconds. This processing time is much longer than the time consumed in evaluating other figures of merit [5]. However, PCA retrieves more information about the dynamic characteristics of the physical process providing temporal parameters beyond the acquisition and processing time duration, along with spatial structures. To show how PCA works in dynamic speckle analysis we apply it to a quite explored experimental situation: the drying of paint on a surface of hidden topography [1, 6]. In this case the sample is a coin. This kind of object has been used in the literature to compare the performance of different algorithms.

To understand the method and its limitations, in section 2 we briefly present the PCA technique and its practical implementation for the analysis of a video sequence. Section 3 explains how the results of the PCA describe the inner dynamic of the physical process that is observed, and how the results can be applied to any other video sequence of
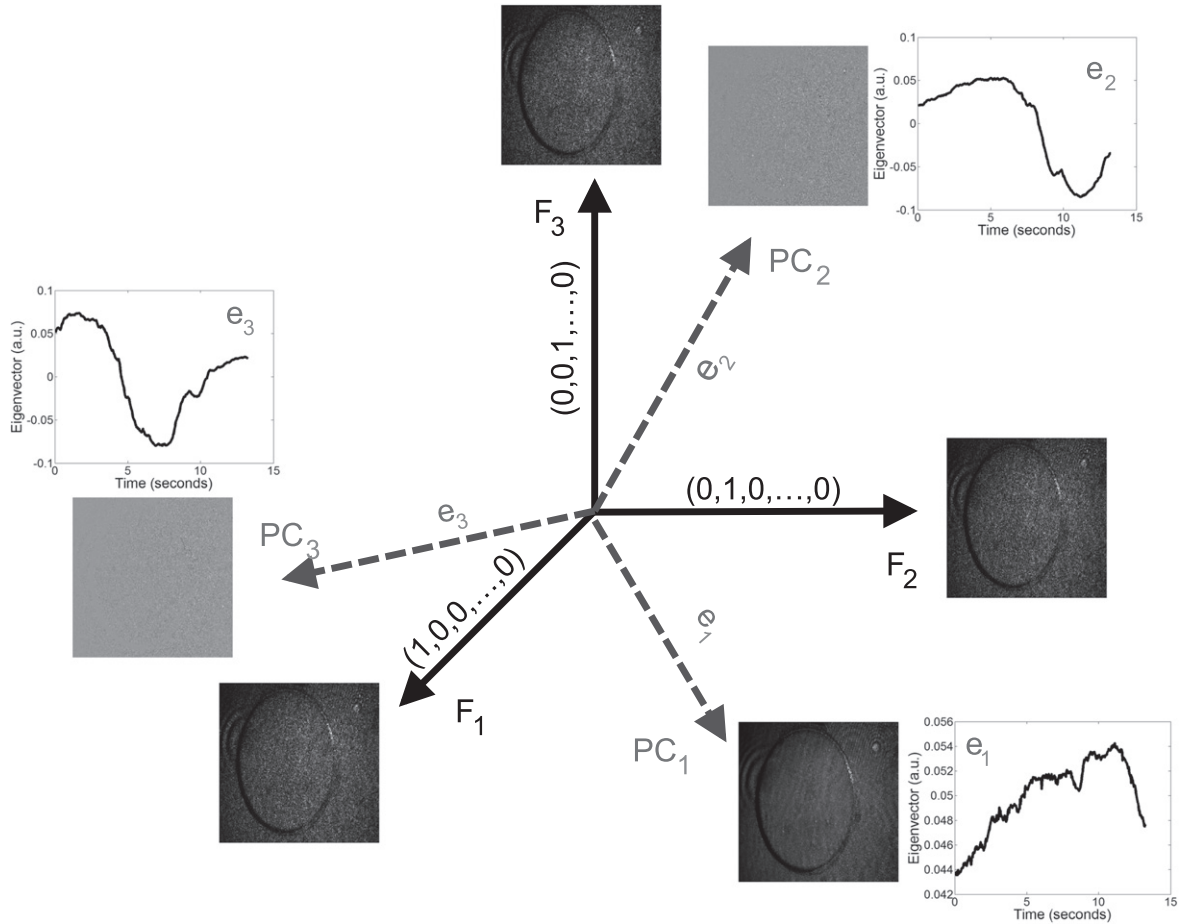
**Figure 1.** Representation of the PCA transformation as a rotation from the frame reference system, $\left\{F_1, F_2, F_3, \ldots, F_N\right\}$, to the principal component reference system, $\left\{PC_1, PC_2, PC_3, \ldots, PC_N\right\}$. To illustrate this rotation we have plotted both the frames and the PC along with the corresponding the eigenvectors, $e_\alpha$. These images are extracted from the analysis made to the dynamic speckle sequence analyzed in this paper. In this figure, we show the first three elements of each set.

dynamic speckle patterns. Finally, in section 4 we summarize the main findings of our contribution.

## 2. The principal component analysis applied to a sequence of frames

PCA, a multivariate technique, analyzes the structure of the variance of a collection of statistical realizations of a set of variables [4]. Among other applications, it reduces the dimension of a set of data by generating a collection of new variables that do not exhibit correlation among them, and selecting the most meaningful. These new variables are linear combinations of the original, and therefore they represent the same magnitude. In this work, the original variables are the collection of $N$ consecutive frames organized as a video sequence, grabbed by a CCD camera from an active sample producing dynamic speckle pattern when illuminated by laser light. In this paper, we name $F$ the following $N$-dimensional variable: $F = \left\{F_1, F_2, \ldots, F_t, \ldots, F_N\right\}$, where $F_t$ represents the $t$ frame in the temporal sequence. To properly account for

inner correlations between frames, the mean of each frame is removed from each one. The PCA transformation provides a new set of frames that we label as $\bar{F}$. We also know that each frame represents a bi-dimensional distribution of irradiance having $R$ rows and $C$ columns. From an statistical point of view, each pixel is a realization of $F$. Then, the number of realizations of the original variable $F$ is $M = R \times C$. If we maintain this bi-dimensional arrangement, we would need to deal with a data cube of consecutive frames. However, it is possible to transform each $R \times C$ frame into a $M$ vector by a simple rearrangement of the data. Both the removal of the mean and the rearrangement of the pixel's values are reversible procedures, so the original frames can always be retrieved.

The frames are connected through the co-variance matrix, $S$, of the zero-mean original data set, $\bar{F}$. As expected, $S$ shows non diagonal elements accounting for the inner correlation among frames. However, it is possible to diagonalize this co-variance matrix by moving to another set of variables that now are uncorrelated. This transformation, that can be interpreted as a rotation within a $N$-dimensional space, is made by the PCA method. This rotation moves from the frame

reference system, $F$, where the images are correlated, to a new reference system, PC, where the correlation between frames is zero. Figure 1 shows the transformation, including the original frames ($F_1$, $F_2$, and $F_3$) that can be interpreted along directions $(1, 0, 0, …, 0)$, $(0, 1, 0, …, 0)$, and $(0, 0, 1, …, 0)$, respectively. The new uncorrelated variables, named as principal components, are related to the original variables by the following linear relation:

$$PC_\alpha = \sum_{t=1}^{N} e_{\alpha,t} \bar{F}_t. \tag{1}$$

The dimensionless coefficients, $e_{\alpha,t}$, of the linear combination describing this transformation are the eigenvectors, $E_\alpha$, of $S$. In figure 1 we interpret these coefficients as the new directions, $e_1$, $e_2$ and $e_3$, along the new PC reference system. In this figure we have also presented the images corresponding to the given principal components. Although the original images—represented as $F_1$, $F_2$, $F_3$—are quite similar among them, the new images—described by the principal components $PC_1$, $PC_2$, and $PC_3$—are already different and show no correlation. The criterion to obtain each direction is that the corresponding principal component should contain the maximum amount of variance of the remaining data set. As a matter of fact, the variance of the resulting component images will be an important result of the analysis, and the method provides these values in decreasing order. It can be demonstrated that the $e_{\alpha,t}$ coefficients obey the following relation:

$$\left(S - \lambda_\alpha I\right)E_\alpha = 0, \tag{2}$$

where $I$ is the $N \times N$ identity matrix. The eigenvalues $\lambda_\alpha$ obtained from the previous equation represent the variance associated with each principal component, $PC_\alpha$. Therefore, the fraction of the total variance explained by each PC, $\Omega_\alpha$, is given by

$$\Omega_\alpha = \frac{\lambda_\alpha}{\sum_{\alpha=1}^{N} \lambda_\alpha}. \tag{3}$$

We may also note that the principal components are organized in decreasing order of their contribution to the total variance of the data, being the sub-index $\alpha = 1$ the one corresponding with the most relevant PC. The eigenvectors, $E_\alpha$, can be arranged as a $N \times N$ matrix. If both the original variables, $\bar{F}_t$, and the principal components, $PC_\alpha$ are organized as two $M \times N$ matrices, the transformation given in equation (1) can be written as: $PC = \bar{F}E$. Using this matrix relation, or equation (1), we obtain the original frames from the principal components as:

$$\bar{F}_t = \sum_{\alpha=1}^{N} e_{\alpha,t} PC_\alpha. \tag{4}$$

This equation reconstruct the original data (except for the mean of each frame that can be easily added to the result of equation (4)) using all the principal components. However, the sum can be carried out by selecting some meaningful principal components to produce a filtered, or rectified, version of the original data.

As a general tool of multivariate analysis, one of the main drawbacks of the PCA method is the accurate interpretation of the results for a given application. However, this statistical origin is also an advantage of the method because its sound foundation. In fact, PCA has been successfully applied to a variety of fields both in optics: spectroscopy, facial recognition, etc; and other areas as econometrics, or computational linguistics [4]. Our case is similar to the analysis of the noise in a sequence of video frames.

## 2.1. Grouping and classification of principal components

As we previously noted, one of the most difficult tasks in analyzing the results from PCA is the identification of the meaningful principal components. Some authors suggest a qualitative criterion: the scree test, based on the trend of the eigenvalues [7]. The cumulative variance contribution may also serve as a quantitative criterion. In this case, the threshold value for this cumulative value is somehow different for each case. An automatic method to identify properly the meaning of the principal component results is based on the calculation of the uncertainties of the variance of a given PC (equation (3)) [8, 9]. If two consecutive eigenvalues are equal within their uncertainties, then they should be considered together because they explain contributions to the variance that are not distinguishable. When this happens, they merge within the same process [8]. With this criterion, two principal components, $PC_\alpha$ and $PC_{\alpha+1}$, with overlapping variances in their uncertainties must be considered as a single unity. Then, it is not meaningful, nor statistically correct, to reconstruct the signals using each of them separately. This reasoning can be extended to a higher number of PCs if the uncertainties of their variances successively overlap. A more detailed explanation of the statistical reasoning of this criterion is given in [8]. This grouping strategy has been successfully applied in several fields [10]. Therefore, single-component and multicomponent processes are defined. The single-component that appear isolated, i.e., whose variance uncertainty does not overlap with any other, can be considered as relevant or independent. Typically, these isolated principal components show a harmonic dependence in their corresponding eigenvectors. We can relate their variance with a given frequency. In this case, the corresponding eigenvalue represents the value of the power spectrum density (PSD) of the underlying process (see [11]).

When we consider the last PCs, i.e., those with a minimum contribution to the variance of the original data, we typically find a process that merges together a large amount of PCs [8]. Often, their contributions to the variance overlap continuously. Besides being linked by their uncertainties, their associated eigenvectors vary very rapidly and cannot be identified with a specific value of frequency but a mix of them. They represent a residual noise that contains, not only the measurement uncertainties, but also a high frequency contribution involving a wide frequency band-width.

The previous classes of principal components (a set of isolated principal components, and a continuous overlapping principal components) have already been analyzed [8, 10, 12–
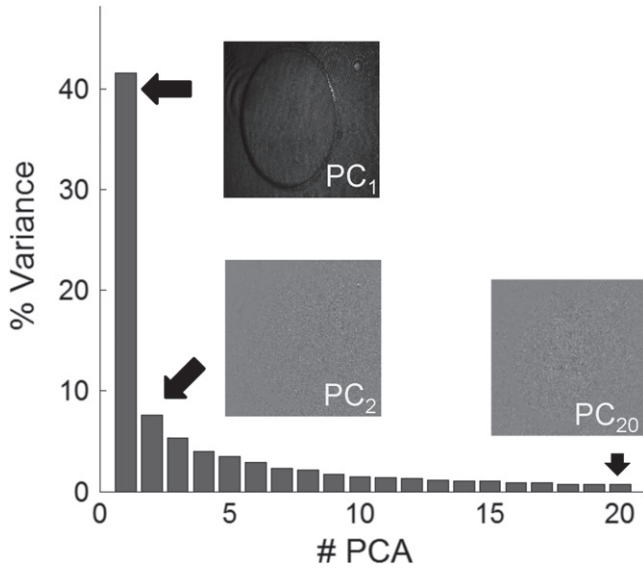
**Figure 2.** Relative contribution of the first 20 principal component to the total variance of the original frames. The images correspond to PC$_1$, PC$_2$, and PC$_{20}$.



**Figure 3.** Relative contributions of the eigenvalues of the principal components analysis with their uncertainties for the coin data set. They are expressed as a percentage of the total variance and correspond with the relative weight given by equation (3). The vertical lines show the limits of the three classes identified in this analysis. The inset shows part of the variance contribution of class P3 where the overlapping between successive eigenvalues is clearly observed.
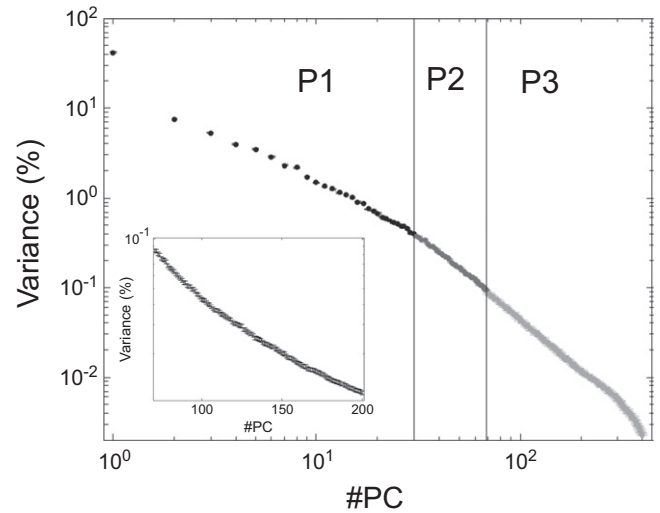
15]. However, some other structures may appear in the data set. Then, grouping and classification will help associate groups of principal components with different physical processes. >From here, we will perform an analysis of a sequence of dynamic speckle of a painted coin.

The application of the PCA method requires the diagonalization of the covariance matrix, the evaluation of different statistics of the data set, and the reconstruction of a filtered set of frames by using a matrix product. For a typical data set containing some few hundreds of frames, and having around $10^5$ pixels, the time consumed by these numeric procedures is proportional to the data volume.

## 3. Application of the PCA method to dynamic speckle sequences

In this contribution PCA is applied to a sequence of 398 images obtained at a frame rate of 30 fps (sampling frequency, $f_{Sampling} = 30\,Hz$) with a size of $512 \times 512 = 262144$ pixels and lasting 13.3 s. The analysis has been made with a personal computer with an Intel-i7 processor, 16 Gb of RAM, under the Matlab environment. The PCA results are obtained after 21 s. If the data follow a normal distribution, the grouping and classification are straightforward and consumes less than 1 s. However, if the normality assumption fails, then it is necessary to calculate fourth order cumulants (see reference [8]) and the grouping process takes around 4 min. Finally, the reconstruction of the frames using a selected set of frames is evaluated in 2 s. The computational time can be reduced when using dedicated and compiled algorithms.

Our sequence of images contains the dynamic speckle produced by a coin covered with paint along the drying of the paint coating [1]. The underlying topography cannot be

perceived during the drying process, but the coating thickness varies accordingly, and thin paint dries before thick ones. So, measures of dynamic activity should show the hidden topography. The method can be straightforwardly adapted to other dynamic processes.

We first represent the relative importance of the PC to the total variance of the original data set (see figure 2). The first principal component explain 41.5% of the total variance of the data, and the following ones are contributing in decreasing order.

Figure 3 shows a log–log plot of the percentage of the eigenvalue contributions obtained after diagonalizing the covariance matrix, $S$. They are represented as a function of the ordinal of the associated principal component. After analyzing the results, three classes of structures can be observed that we label as P1, P2, and P3.

The first one, P1, is formed by the first 30 PCs. This class represents 87.2% of the total variance of the original data. These PCs are isolated and independent because they are single-component processes. This means that their eigenvalues, considering their uncertainty due to the sample, do not overlap. Their associated eigenvectors can be fitted to sinusoidal functions of time (see inset of eigenvector #6 in figure 4). Therefore, for each principal component included in this spatial-temporal structure, the associated variance , i.e., the eigenvalue, is related to a single frequency. This frequency is determined by the time evolution of the associated eigenvector [10, 11]. To prove this, figure 4 shows the Fourier transform of each eigenvector. We can observe that the sinusoidal eigenvectors included in the P1 class present a narrow peak if compared with the rest that are not isolated eigenvectors. The results obtained from the PCA for this first
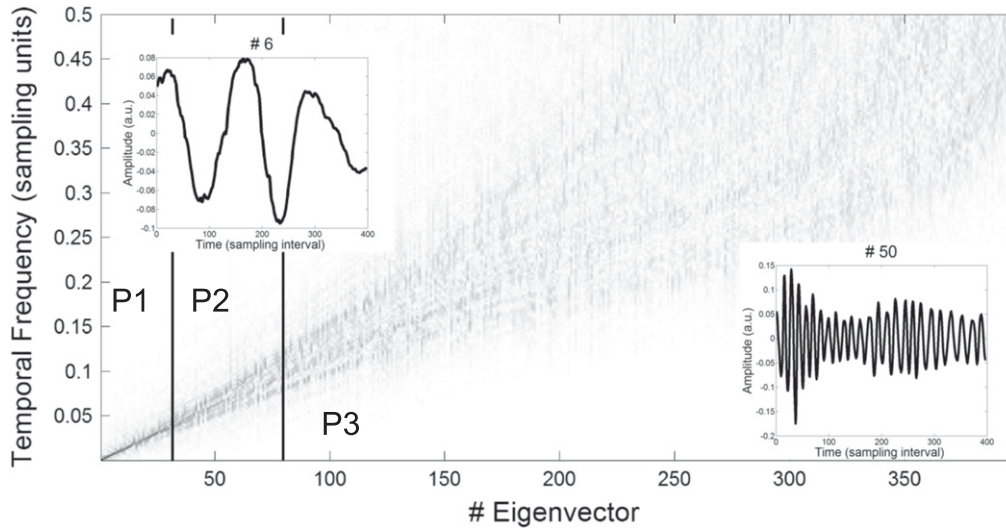
**Figure 4.** Fourier transform of the eigenvectors for the coin data set. The first 30 eigenvectors show a sharp peak in frequency, while the rest have a wider spectrum. Eigenvector #6 (isolated and sinusoidal) and #50 are shown as insets. The vertical lines limit the three classes identified in this analysis. The sampling frequency is $f_{\text{Sampling}} = 30$ Hz.
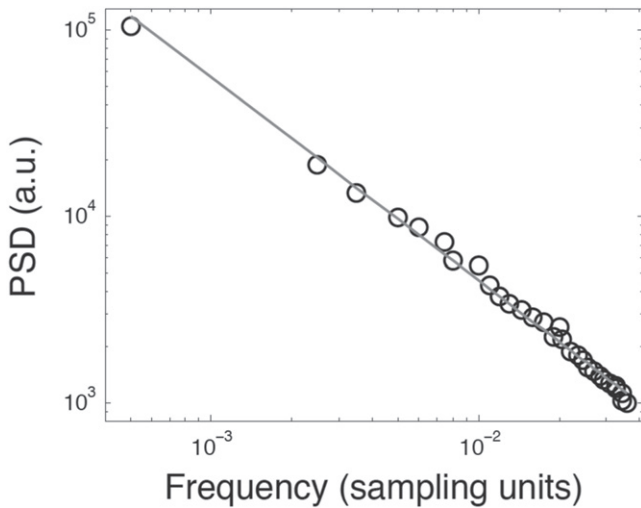


**Figure 5.** PSD adjustment for the sinusoidal isolated eigenvectors. This analysis can be only applied to those eigenvectors included in the P1 class. The sampling frequency is $f_{\text{Sampling}} = 30$ Hz.

structure P1, can be seen as a PSD function [10, 11]. This interpretation is graphically presented in figure 5, that is a log–log plot of the eigenvalues. Versus the frequencies associated with each eigenvector. These data can be linearly fitted showing a temporal evolution of the type $f^{-\alpha}$. Then, the PSD($f$) is described by the following equation:

$$\text{PSD}(f) = \frac{30 \pm 2}{f^{1.09 \pm 0.02}}. \tag{5}$$

This fitting yields a value of $\alpha = 1.09 \pm 0.02$.

The third class, P3, groups together in the same multi-component process principal components ranging from #70 to #398. They represent 5.4% of the total variance of the original data. In figure 4 we can see how the range in temporal frequency components of these PCs is wide and contains high

frequency components until the Nyquist frequency ($f_{\text{Nyquist}} = 0.5 f_{\text{Sampling}} = 15$ Hz).

Finally, the second class, P2, includes those PCs from #31 to #69 that cannot be included in P1 (representing a PSD) nor the ones in the P3 class (representing a uncorrelated high-frequency temporal noise). By analyzing the uncertainties in the data represented in figure 3 and applying the grouping strategy, we conclude that this subset of principal components is not a single process. Nevertheless, the eigenvectors of the P2 class share the feature of being defined by two frequencies or even three (see figure 4 and the inset representing eigenvalue #50), instead of one as it happens in P1 class.

Once the principal components are split according to the aforementioned criteria, we can proceed to the rectification, or filtering, process. Rectification consists in the reconstruction of a sequence of $N$ frames (the same number as in the original sequence) using only the principal components obtained in each one of the defined classes. We select a set of principal components when performing the sum in equation (4) and add the mean value of each frame. We obtain three sequences of frames, one for each class. The importance of each class can be given by the sum of the contributions, $\Omega_\alpha$, of the principal components belonging to each one. The first class, P1, involves 87.2% of the total variance of the original data, P2 represents about 7.5%, and the third class, P3, accounts for remaining 5.4%.

Several types of analysis can be performed on the images to find the reason behind their differences (see for example Fujii's method in [1]). In figure 6 we show the mean and the standard deviation of each rectified frame set. For the three sequences, the mean of each pixel, and its standard deviation are calculated and shown as images. We see that the images of the mean do not reveal useful information. However, the standard deviation images reveal the underlying topography of the object. In this way, a high value in the image implies a point where the dynamic speckle varies with high amplitude

Mean                STD



P1

87%

P2

7.5%

P3

5.4%
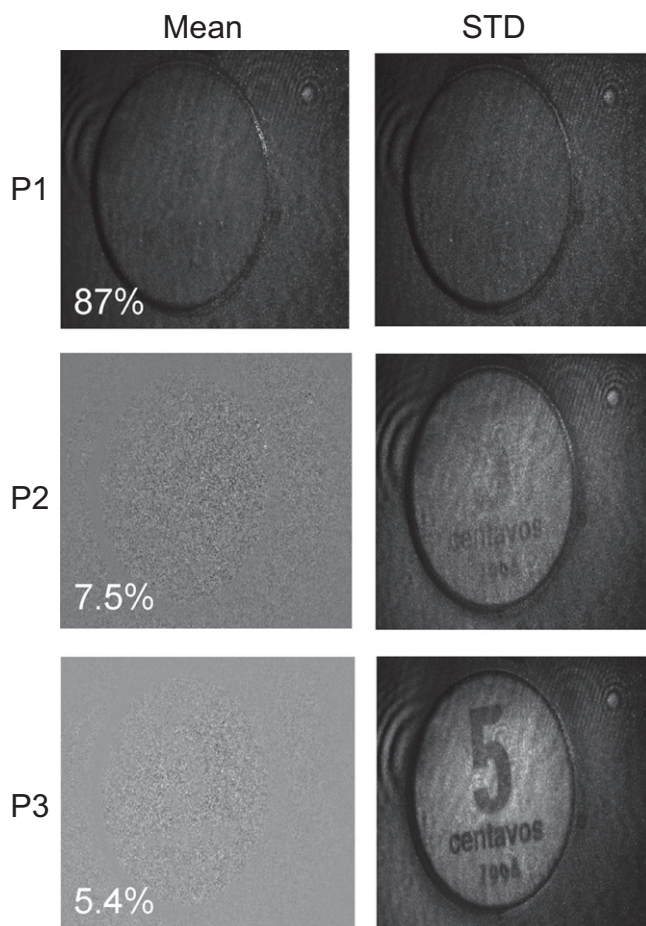
**Figure 6.** View of the mean and the standard deviation of each pixel for the rectified frames for coin data set. The rectification is for P1 class (top), P2 class (middle) and P3 class (bottom). The numbers show the percentage of the total variance explained by each class.

and fast temporal evolution. A quite similar analysis, using the noise instead of the signal as the figure of merit, has been successfully applied to the detection and identification of buried land-mine objects with principal components [16].

The first class, P1, shows a high absolute value of the standard deviation with an almost uniform spatial distribution over the image. In the second one, P2, regional differences appear, which are clearly visible in the third class, P3, where the topography of the coin, letters and numbers, can be distinguished. The region surrounding the digit '5' has the strongest activity. The fundamental difference between the second and the third classes is spatially located around the digit, clearly visible in the third but not so much in the second. The standard deviation decreases from P1 class values to P2 and P3. However, in these last two classes, the spatial distribution of the variance changes more than in P1, revealing the underlying topography. This analysis of the noise is possible after applying the PCA method and the previously explained grouping strategy.

PCA can also assign time scales to these classes. >From figures 3 and 4 we infer that the highest frequency achieved by class P1 corresponds to about 1 s (1 Hz), and about 0.3 s

for class P2. This implies that the drying dynamics that produces P1 involves movements in the speckle in the order of seconds or longer, and P2 between 1 and 0.3 s. The third class, P3, is then, associated to an even shorter time scale.

We can relate topography and time sales with the drying mechanism. The drying of paint is a rather complex process involving several steps including solvent evaporation and refractive index variations, features to which dynamic speckle is sensitive [17]. The initial stage of the process is when the solvent evaporates, and it is called the constant rate period. Afterwards, there is a falling rate period in which the drying rate decreases to zero. The beginning of the falling rate period depends on the film thickness and solvent evaporation [18]. Since the thickness of the coating on the over-relief of the coin is locally variable, the paint can be locally in different drying stages. Thin paint regions should be in later stages compared to thick ones. Therefore, activity in the speckle pattern varies with the hidden coin topography [6]. The results obtained from the PCA method shows that classes P2 and P3, associated with time periods shorter than 1 s, reveal the hidden topography of the coin and can be related with regions at later drying stages. From the results obtained for the P3 class, we observer that the standard deviation is smaller in those regions where the paint coating in thinner. This means that the contribution to the rapid variation of the P3 class is proportional to the thickness of the paint. At the same time, class P1, is identified as having a time period longer than 1 s and following a $1/f$ PSD. We should emphasize that this information has been obtained from a sequence frame lasting 13.3 s. Although this time period is much shorter than the total drying time, the results obtained from PCA already reveal different processes that last longer than acquisition time.

## 4. Conclusions

We presented the application of a multivariate technique, the PCA method, to the analysis of a sequence of images of dynamic speckle interferometry. We obtained quantitative results of different processes that can be related with the inner dynamic of the drying process of paint coating a coin. The classes identified here are related with temporal and spatial distributions at different spatial-temporal scales. Additionally, these results have been obtained from a sequence lasting just 13.3 s. This time duration is much shorter than the typical drying times of paint, allowing fast monitoring of the process. The method provides a quantitative evaluation of the reported classes of spatial-temporal evolutions. Most of the variance of the original data (87.2%) is explained as a $1/f$ PSD with a time scale longer than 1 s. The remaining variance (13% of the total) presents a temporal evolution with time constants faster than 1 s. Furthermore, they already reveal the hidden topography that is associated with later drying states. The capability of the PCA method to section the variance of the data into independent components allows a meaningful analysis relating physical sources to different classes of spatial-temporal evolution.

The computational time duration of the PCA method makes difficult its real-time application with current and standard computational capabilities. In these conditions, PCA can be seen as a fast pre- and post-processing tool that allows filtering of the data and provides parameters with time characteristics longer than the time duration of the analyzed sequence. Also, PCA results are given in terms of PSDs and spatial maps of characterizing parameters. This identification relates the observed classes with the dynamic processes associated to evaporation of solvent, displacement of pigments at different sizes or scales. The method provides new information, allows the trimming of the components of interest, and does not require previous knowledge of the dynamics of the phenomenon under investigation as it adapts itself to the signal.

# References

[1] Rabal H J and Braga R A Jr (ed) 2008 *Dynamic Laser Speckle and Applications* (Boca Raton, FL, London: CRC Press, Taylor and Francis) pp 14–18

[2] Passoni I, Rabal H, Meschino G and Trivi M 2013 Probability mapping images in dynamic speckle classification *Appl. Opt.* **52** 726–33

[3] Passoni L, Dai Pra A L, Scandurra A, Meschino G, Weber C, Guzman M, Rabal H, Trivi M and Estévez P A 2013 Improvements in the visualization of segmented areas of patterns of dynamic laser speckle *Advances in Self-Organizing Maps* (*AISC* vol 198) (Berlin: Springer) pp 163–71

[4] Jolliffe I T 2002 Principal component analysis *Springer Series in Statistics* (Berlin: Springer)

[5] Sendra G H, Dai Pra A L, Passoni I L, Arizaga R, Rabal H and Trivi M 2010 Biospeckle signal descriptor: a performance comparison *Proc. SPIE* **7387** 73871K-1–5

[6] Faccia P A, Pardini O R, Amalvy J I, Cap N, Grumel E E, Arizaga R and Trivi M 2009 Differentiation of the drying time of paints by dynamic speckle interferometry *Prog. Org. Coat.* **64** 350–5

[7] Cattell R B 1966 The scree test for the number of factors *J. Multivariate Behav. Res.* **1** 245–76

[8] López-Alonso J M, Alda J and Bernabeu E 2002 Principal components characterization of noise for infrared images *Appl. Opt.* **41** 320–31

[9] North G R, Bell T L, Cahalan R F and Moeng F J 1982 Sampling errors in the estimation of empirical orthogonal functions *Mon. Weather Rev.* **19** 699–706

[10] López-Alonso J M and Alda J 2005 Characterization of dynamic sea scenarios with infrared imagers *Infrared Phys. Technol.* **46** 355–63

[11] López-Alonso J M and Alda J 2003 Operational parametrization of the 1/*f* noise of a sequence of frames by means of the principal component analysis in focal plane arrays *Opt. Eng.* **42** 1915–22

[12] López-Alonso J M, Rico-Garcia J M and Alda J 2004 Photonic crystal characterization by FDTD and principal component analysis *Opt. Express* **12** 2176–86

[13] López-Alonso J M and Alda J 2004 Correlation in finance: statistical approach *Proc. SPIE* **5471** 311–21

[14] López-Alonso J M and Alda J 2004 Identification of weak faint point sources by using principal component analysis *Proc. SPIE* **5238** 142–52

[15] López-Alonso J M, Monacelli B, Alda J and Boreman G D 2005 Infrared laser beam temporal fluctuations: characterization and filtering *Opt. Eng.* **44** 054203

[16] du Bosq T, López-Alonso J and Boreman G 2006 Millimeter wave imaging system for landmine detection *Appl. Opt.* **45** 5686–92

[17] Blandin H P, David J C, Vergnaud J M, Illien J P and Malizewicz M 1987 Modelling of drying of coatings: effect of the thickness, temperature and concentration of solvent *Prog. Org. Coat.* **15** 163–72

[18] Amalvy J I, Lasquibar C A, Arizaga R, Rabal H and Trivi M 2001 Application of dynamic speckle interferometry to the drying of coatings *Prog. Org. Coat.* **42** 89–99