Analytica Chimica Acta 706 (2011) 229–237

Contents lists available at SciVerse ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

# Development of a technique based on multi-spectral imaging for monitoring the conservation of cultural heritage objects

Emilio Marengo[a,*], Marcello Manfredi[a], Orfeo Zerbinati[a], Elisa Robotti[a], Eleonora Mazzucco[a], Fabio Gosetti[a], Greg Bearman[b], Fenella France[c], Pnina Shor[d]

[a] University of Eastern Piedmont, Department of Environmental and Life Sciences, via T. Michel 11, 15121 Alessandria, Italy
[b] IAA for Imaging Technologies of the Dead Sea Scrolls, ANE Image, Pasadena, CA, USA
[c] Preservation Research and Testing Division, Library of Congress, 101 Independence Avenue, SE, Washington, DC 20540, USA
[d] Director of Dead Sea Scrolls Projects, Israel Antiquity Authority, Rockefeller Museum, Post Office Box 586, Jerusalem, 91004, Israel

ARTICLE INFO

*Article history:*
Received 27 April 2011
Received in revised form 4 August 2011
Accepted 30 August 2011
Available online 7 September 2011

*Keywords:*
Surface artifacts monitoring
Cultural heritage preservation
Multi-spectral imaging
Multivariate analysis

ABSTRACT

A new approach for monitoring the state of conservation of cultural heritage objects surfaces is being developed. The technique utilizes multi-spectral imaging, multivariate analysis and statistical process control theory for the automatic detection of a possible deterioration process, its localization and identification, and the wavelengths most sensitive to detecting this before the human eye can detect the damage or potential degradation changes occur. A series of virtual degradation analyses were performed on images of parchment in order to test the proposed algorithm in controlled conditions. The spectral image of a Dead Sea Scroll (DSS) parchment, IAA (Israel Antiquities Authority) inventory plate # 279, 4Q501 Apocryphal Lamentations B, taken during the 2008 Pilot of the DSS Digitization Project, was chosen for the simulation.



## 1. Introduction

The study of cultural heritage, archaeological, historical, artistic and archive, library or museum objects must be dealt with on an interdisciplinary level, making use of different experiences and skills necessary in achieving a common objective: preservation of the original object, both substrate and the media that contains the information on the object.

Conservation science is consistently striving to find better ways to preserve cultural heritage objects and to obviate damage liable to be caused by environmental or accidental factors. Materials constituting the cultural heritage object are subjected to changes over time, due to the interaction between the object and the physical factors (light, temperature, relative humidity, oxygen, particulates), the chemical factors (atmospheric oxygen, various pollutants) and the biological agents (bacteria, fungi, insects, mold). The monitoring of cultural heritage objects over time is critically important in order to alert the conservator when potentially damaging changes are occurring. The method described in this article follows a new approach in the context of cultural heritage aimed at the automatic and fast detection of a developing deterioration process, and its localization and identification. This requires the direct measurement of the reflectance spectrum of the artifact through a non-invasive method, leaving the object unchanged for successive examinations. While our current work involves texts on parchment and papyrus, the method is extensible to other objects.

The utilization of multispectral imaging and digital image processing can be beneficial for the preservation of cultural heritage objects. Modern imaging technologies have been having a significant impact on cultural heritage and archaeology in the last decade. A number of applications of multispectral imaging in the field of cultural heritage are present in the literature, mainly regarding characterization studies and artifacts analysis [1–4]. Spectral imaging of manuscripts has been used to improve historical documents readability [5–7].

For the case of the Dead Sea Scrolls, mostly on parchment, it is known that the changes in legibility are driven by changes in the parchment reflectance in the visible. Spectral imaging showed that the easy-to-read scrolls show significant differences between the ink and parchment spectra in the visible, while they are very similar for hard-to-read ones. However, for the illegible scrolls, the parchment reflectance increases significantly relative to the ink in the IR, which is why infrared photography of the scrolls, done in the early 1950s was successful [8,9]. These changes in reflectance suggest a natural way to monitor the scrolls for changes, namely monitor the reflectance through repeated imaging, analysis and comparison.

The spectral imaging process obtains a complete spectrum for each pixel of the image. The potential applications are innumer-

* Corresponding author. Tel.: +39 0131 360259; fax: +39 0131 360250.
*E-mail address:* marengoe@tin.it (E. Marengo).

0003-2670/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2011.08.045

able, and have already been applied in the fields of biology and biomedical applications, pollution control, and other disciplines [10–12].

There are two main methods of capturing spectral imaging data: one is to illuminate the sample with broadband light and filter the image detection, separating wavebands by filtering between the object and the camera, while the other method is to filter the illumination or use narrow waveband illumination in conjunction with an unfiltered camera. This second method has conservation advantages due to reduced light levels and heat compared to the former.

Chemometric techniques provide useful tools for extracting the systematic information from large and complex datasets. These methods have already been applied in the field of cultural heritage [13,14] for several purposes like provenance and classification studies, objects and manufacturing techniques characterization, monitoring and preservation.

A powerful tool to investigate surface degradation process of artifacts surfaces can be demonstrated through the application of the statistical process control (SPC) theory [15–18]. This has already been demonstrated in the monitoring of the conservation state of wooden objects and canvas painted with inorganic pigments, analyzed by Raman and IR spectroscopy [19,20]. The application of multivariate statistical process control approaches to image data was first developed for real-time process monitoring and control [21], for industrial [22] and agricultural [23] image-based problems.

A method based on multispectral imaging coupled to multivariate analysis is proposed here for monitoring the state of health of cultural heritage objects surfaces and, in general, for every kind of surface whose conservation state needs to be monitored. Industrial applications of the same technique can be easily envisaged for applications including the control of solid catalysts, ion exchange resins, raw and finite materials.

To test the algorithms a multispectral image of a Dead Sea Scroll (DSS) parchment, IAA (Israel Antiquities Authority) inventory plate # 279, 4Q501 Apocryphal Lamentations B (dated between 50-25 BCE [24]) was used. The applicability of the approach and its limitations were studied by generating virtual images of the parchment containing an artificially degraded region.

Then, the application of this monitoring technique to a real situation is briefly presented to show an example where a parchment was subjected to a real accelerated degradation process. The changes of the surface were investigated using the proposed approach.

## 2. Theory

### 2.1. Multispectral imaging

Multispectral imaging refers to the capture of multiple images of specific wavebands of the spectral region, with each image acquired at a different wavelength [25], obtaining a complete spectrum for each pixel of the image. The resulting dataset is a 3-way data matrix, also called CUBE, where *x* and *y* axes are the coordinates of the pixels of the image and in the third dimension there are the reflectances of the pixels at defined wavelengths. Full application of Multispectral imaging can typically span the wavelength range from 380 nm to 1100 nm, capturing ultraviolet (UV), visible (vis) and near infrared (NIR) spectral regions.

One of the reasons to acquire digital spectral data is the capacity to use image processing software to improve the visible contrast between ink and written or painted substrate and consequently enhance legibility of the text. There are several types of multispectral instruments including spectral scanners, that use electro-optical devices; spatial scanners, which use prisms, gratings or beam splitters to create spectral discrimination; interferometric analyzers, that typically acquire a 2D image and scan optical path differences to obtain a complete interferogram; hybrid instruments, such as computed tomographic imaging spectrometers and polarization-dependent rotogram devices [26].

A system that uses LED illumination to prove the spectral component, coupled with a 39 MP monochrome camera was employed for the real application example. Cold LED illumination eliminates a major objection to spectral imaging that uses broad band lighting and then filters the detection side. This method also lets us acquire high-resolution spectral images ($7216 \times 5412$ pixels for each image, sensor array size of $49 \times 37$ mm), something that is not possible with other methods. Details of the system and its performance for both spectroscopy and color rendering are in the literature [27,28].

### 2.2. Principal component analysis (PCA)

The multivariate approach represents the only possible choice when datasets are characterized by a large number of variables and/or objects. In the present case the dataset is challenging, due to complex correlation patterns related to the use of a spectral description of the surface (obtained from the spectral imaging) [29,30]. A rationalization of the problem can be obtained by means of principal component analysis (PCA). PCA may allow the separation of the systematic information from the experimental noise and the natural fluctuations: this is true when systematic variations overcome the variability due to experimental error. PCA provides a new set of orthogonal variables, a linear combination of the original components, to describe the system under investigation in a very compact and efficient way. This analysis can provide several types of information useful for pattern recognition analyses. The most important are: the *scores* (**T** matrix), namely the projections of the objects onto the space given by the relevant PCs and the *loadings* or *weights* (**L** matrix), the coefficients of each original variable in the linear combination defining each PC.

In the present paper PCA is performed on the characterization data (3-way data arrays describing the sample when no degradation is applied) providing information on the sources of variability that characterize the dataset before the application of a damaging effect. The degradation images are then projected onto the relevant PCs calculated from the original characterization data. The analysis of the scores of the degraded images allows the identification of the presence of relevant changes caused by the degradation process, while the analysis of the loadings (calculated on the characterization dataset) may suggest the potential causes that produced the changes.

Using the subset of relevant PCs it is possible to rebuild the original dataset (matrix $\mathbf{X}'$). In this way it is possible to filter, for example, the experimental noise or the unnecessary information:

$$\mathbf{X}' = \mathbf{T} \cdot \mathbf{L}'$$

The difference between the original data **X** and the re-calculated data $\mathbf{X}'$ from the relevant PCs is called *matrix of the Residuals* (**R**) and contains the information not accounted for by the PCs used to recalculate $\mathbf{X}'$. The residual matrix is expected to contain only noise and random fluctuations when it is calculated from the same data used for calculating the PCs.

In the present case, the residuals matrix was calculated on degraded images after re-projection onto the space given by the relevant PCs calculated on the characterization data. In this particular case, the residuals matrix calculated on the degraded images may contain systematic information connected to the eventual presence of new species that were not present during the characterization step, species originated from the degradation process. The re-projection of the degraded images on the PCs calculated from

the characterization data gives information on degradation mechanisms that involve species and structures that were already present during the characterization. If new species form during the degradation (for example, a superficial coating degrades making it visible the layer below), they cannot be accounted for by the PCs calculated on the characterization data, since these species were not present. They represent new information [31,32] that will be present in the **R** matrix calculated on degraded images. In this case the **R** matrix of the degradations will contain systematic information, beside the expected random noise and natural system fluctuations: this systematic information can be detected by performing a further PCA on the residual matrix of only degraded images.

### 2.3. Multivariate statistical process control

SPC concepts and methods are essential in the monitoring of industrial processes with the contemporary targets of identifying when the process presents anomalous behaviors, of preventing the effect of these anomalous behaviors and of identifying the *Special Cause* [15] that produced the anomalous effect. They can be advantageously applied also to cultural heritage monitoring as showed by Marengo et al. [19,20].

The objective of this technique is the automatic monitoring of the current state of preservation of artifacts and changes over time in order to verify that the object remains in the state of *statistical control*, which means that no change is taking place and that no degradation process is active.

The first step, like when industrial process control must be implemented, is to collect data that permit to assess and characterize the system natural variability. This variability is characteristic of the investigated system and permits to understand when some relevant, not random, changes take place.

Using Shewhart's approach, it is possible to calculate a lower and an upper control limit (LCL and UCL) that when overridden indicate a high probability that a Special Cause, namely an anomalous source of variability, is present. The LCL and UCL are calculated using the system natural variability evaluated from data collected while the system was in statistical control. If the natural variability of the experimental setting is not correctly assessed, false/erroneous alarms or not sufficiently sensitive control charts for the identification of changes would be obtained. In the present case Shewhart control charts were constructed using the scores of the relevant PCs [33,34].

When a sample infringes the natural structure that characterizes the dataset during the "in-control state" it will show anomalous score values or residuals and shall be detected and identified by their analysis.

## 3. Experimental

The multispectral image of a Dead Sea Scroll parchment (4Q501-279, Apocryphal Lamentations B) shown in Fig. 1 was chosen for the simulation.

The parchment imaging was performed at 35 wavelengths from 650 to 990 nm in 10 nm steps. Data was captured using a CRI Nuance Imaging Spectrophotometer (CRI, MA, USA) and normalized by dividing for a white reference cube obtained acquiring a 99% reflecting lambertian reflector (Spectralon®) in order to have reflectance spectra between 0 and 1.

In the case of the real application presented to show the method performance a Eureka Vision LED system from Megavision (CA, USA) was used. The system has a 39-megapixel Kodak CCD monochrome sensor array with 7216 × 5412 pixels and two EurekaLight™ LED illumination panels. The LED emissions were centered on the following wavelengths: 365, 450, 465, 505, 535, 592, 625, 638, 700, 735, 780, 870 and 940 nm. LEDs emit in narrow spectral bands over ranges of wavelengths from the near ultraviolet to the near infrared. The LED bandwidth ranges from about 10 nm in UV to 40 nm in the IR. LEDs 39 MP images can be obtained in about 1 min with no thermal exposure of the object as the minimum necessary light is generated to properly expose the objects to the spectral band of interest: the samples are safely exposed.


**Fig. 1.** Dead Sea Scrolls image at 960 nm.

A new 15 cm × 20 cm calfskin parchment was used (Pergamena Parchment LLC, NY, USA).

All the algorithms necessary for the treatment of the data were developed using MATLAB (R2007b, The MathWork, USA).

## 4. The applied procedure

The procedure applied in this research was devoted to evaluate the applicability and the limitations of spectral imaging by utilizing and manipulating simulated data. The following steps constitute the range of factors under consideration.

(1) *Addition of noise* – In order to be able to characterize what can be considered a natural variability of the images, it is necessary to have several images of the same object, recorded over time while no degradation changes have occurred to the object. These images must account for all the natural variability sources, for example: instrumental noise, small alignment errors, normal hysteresis of the object, etc. In the present case noise was artificially added to the spectral images in order to simulate the natural experimental variability. A white noise was employed to this purpose. Different images were generated with the addition of a random white noise and the images were successively used to evaluate the control limits of the Shewhart charts built on the relevant principal components. In agreement with experimental evidence, the maximum white noise level was established to be 0.01 reflectance units with respect to the spectral intensity measured in each pixel for each different wavelength. This procedure was applied to obtain five different replications of the starting image of the DSS, obtaining at the end five 3-way arrays (image cubes) of the DSS in presence of the white noise. Five 3-way characterization arrays were considered sufficient in this case since a white noise of known entity was added to simulate the natural variability. These 3-way arrays were used for calculating the relevant PCs and the control charts limits. In the present case, two PCs were considered significant.

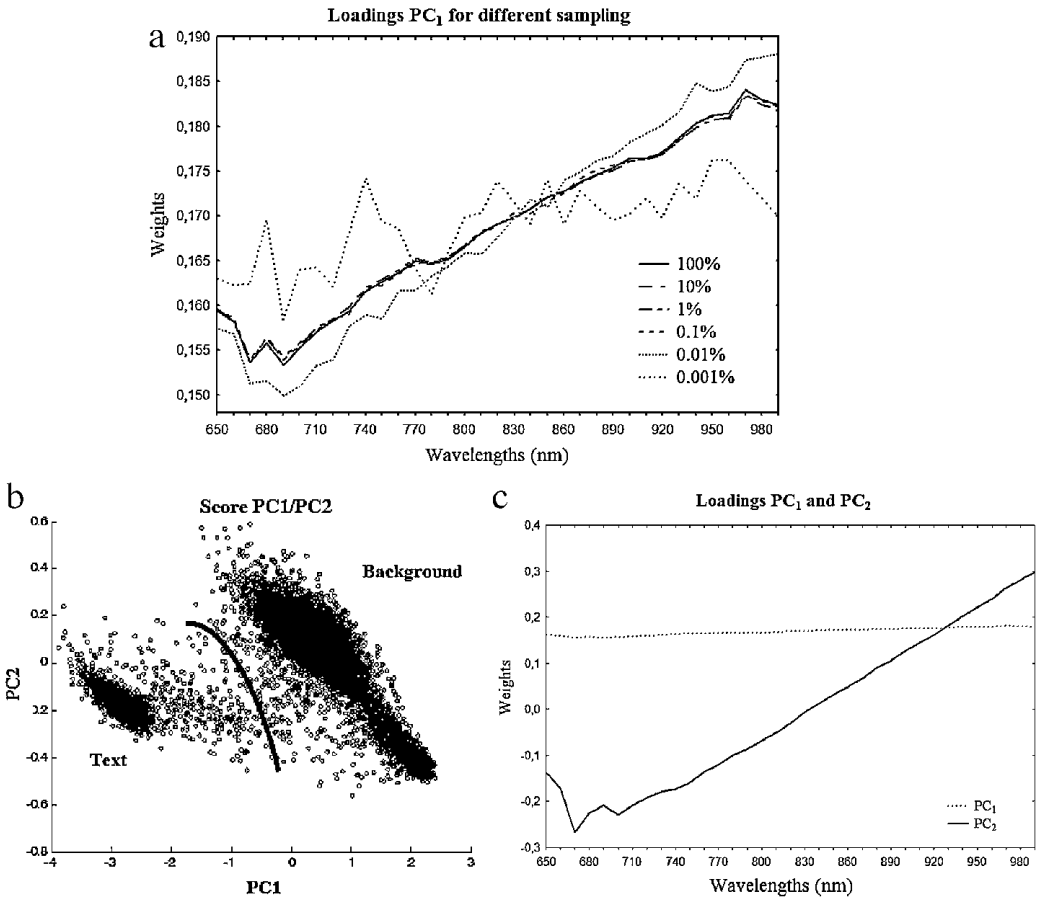


**Fig. 2.** Loadings of $PC_1$ for different samplings (a); score-plot of $PC_1$ and $PC_2$ (b); and loading plot of $PC_1$ and $PC_2$ from IR (left) to UV (right) (c).

(2) *3-way arrays* (*CUBEs*) *unfolding* – The 3-way arrays are unfolded into a bi-dimensional matrix (**X**) where the columns are the 35 wavelengths and the rows are the pixels of the different images. The columns are mean centered prior to any statistical analysis.
(3) *Sampling of the dataset* – A good characterization of the in-control natural variability needs several 3-way arrays to be processed simultaneously, a process that requires a large volume of memory (in the level of the Giga- or Tera-bytes). A random selection of individual pixels that preserves both the macro and micro information present in the 3-way arrays was then tested. This sampling was performed on the rows of the data matrix. Different percentages of the number of pixels present in the images (sampling percentages) were compared to select the best compromise between memory requirements and quality of the results. For every 3-way characterization array, the same random pixels were sampled. A PCA was then performed on the pixels sampled for each sampling percentage considered and the results were compared to the case where PCA is applied to the overall dataset, without pixel selection. The loadings were compared with the aim to identify the minimum percentage of pixels that preserves the information present in the cubes, i.e. providing results comparable to the case where no pixel selection is applied. This procedure allowed the selection of a representative subset of pixels, indirectly preserving both macro and micro information. A percentage of 0.1% proved to be the minimum allowed sampling percentage.

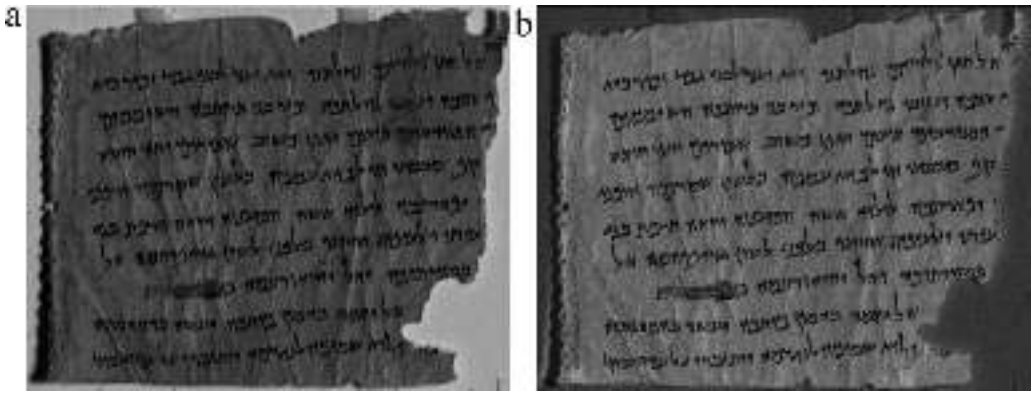


**Fig. 3.** Projection of the scores of $PC_1$ (a) and $PC_2$ (b).

(4) *Principal component analysis* – PCA was then performed on the five 3-way characterization arrays after pixel selection. This step enables the description of the in-control natural variability by means of the relevant PCs that take into account the systematic relationships present in the data. These relationships may include the covariance structure and the different sources of systematic information, like the presence of different figures or different colors. In the present case two PCs were considered relevant.

(5) *Multivariate control chart* – Multivariate Shewhart charts are then calculated using the scores of the relevant PCs obtained in the characterization phase. It is possible from the training set data to calculate the upper and lower control limits (UCL and LCL) that permit to identify the pixels that show anomalous behaviors along time. UCL and LCL are calculated independently for each pixel as the $\pm 3\sigma$ variation around the average score of the corresponding pixel.

(6) *Deterioration of the parchment* – The deterioration of the parchment was simulated by degrading the 3-way arrays that had previously had white noise included. A Gaussian shape modification with predetermined size was added to each multispectral image of the cube. All wavelengths were interested by the degradation: the intensity of the Gaussian shape increased from visible to IR in order to simulate the parchment deterioration spectra described by Bearman et al. [9]. The Gaussian deterioration was located at the pixel with coordinates $x = 400$ and $y = 400$ on the original image: the intensity of the deterioration is therefore maximum in this position and radially decreases from this position in all directions. This particular shape allows to evaluate how the control charts are effective in the identification of degradation mechanisms with spatial diminishing effects. Three degraded 3-way arrays of increasing intensity were built (t1, t2, t3).

(7) *Monitoring of the parchment* – The "virtually aged" 3-way arrays were unfolded and centered using the mean calculated in the training session and the new data were projected onto the PC space previously obtained. The scores of the new image created were then compared with the LCL and UCL, calculated on the characterization data, to identify the pixels whose variation exceeds the control limits. Contribution plots show the contribution of each wavelength to the aged image for a defined pixel. These plots allow the assessment of which wavelengths were most affected by a significant degradation.

(8) *Analysis of the residuals* – As previously indicated, a further principal component analysis of the degraded residuals matrix obtained after re-projection of the degradation images along the relevant PCs calculated from the characterization data, provided information on the eventual presence of new sources of variation, namely the formation of new compounds on the surface.

## 5. Results and discussion

During this work the main assumption was that the experimental noise present in different replicated shots comes uniquely from instrumental noise (white noise).

Before starting the simulation, a preliminary study was performed to identify the optimal percentage of pixels of the images that needed to be sampled in order to obtain stable and reliable results: preserving both the macro and micro information present in the 3-way arrays.

Six different percentages of pixels randomly sampled from the five 3-way arrays, namely 100% (all the pixels are considered: this is used as a reference), 10%, 1%, 0.1%, 0.01% and 0.001% were compared. PCA was performed on the 6 mean centered datasets

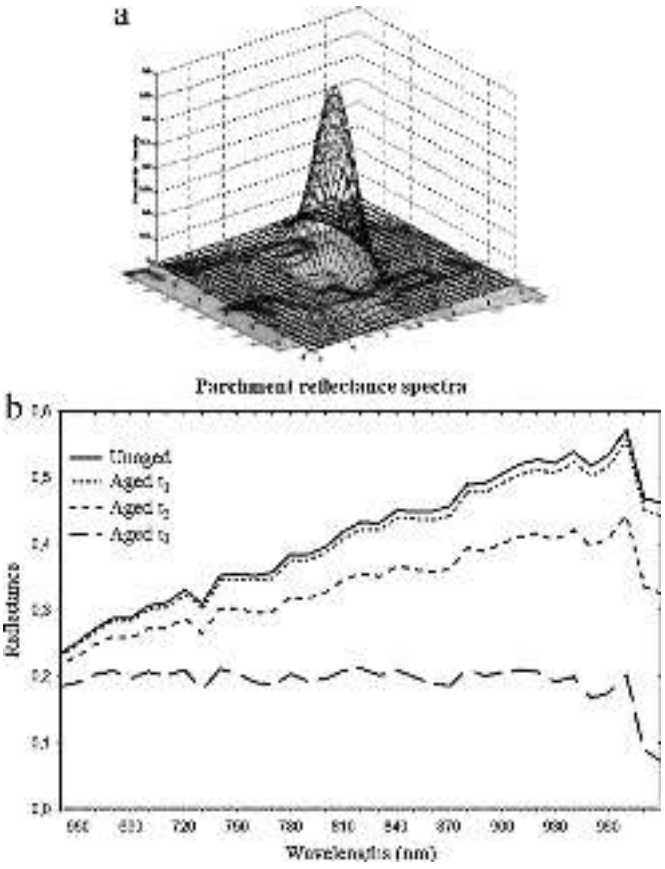


**Fig. 4.** Degradation shape applied at the images surface of the 3-way array (a). Reflectance spectra of un-aged (upper) and aged parchment over time (b); spectra were normalized by dividing for a white reference 3-way array obtained acquiring a 99% reflecting lambertian reflector (Spectralon®) to have reflectance spectra between 0 and 1.

obtained from the different random samplings. Fig. 2 shows the loadings of the first principal component for the different samplings: it is evident that the loadings preserve the information present in the 3-way arrays down to a 0.1% sampling.

The pattern of the 0.01% and 0.001% samplings are different from the others and show that these choices produce results that differ significantly from the raw images (100% sampling). As a consequence of these results we used a 0.1% sampling percentage of the pixel available. This sampling procedure indirectly preserves both macro and micro structures since the results obtained after sampling are comparable to those obtained from the overall dataset.

### 5.1. Training set

The results of PCA performed on the autoscaled simulated characterization images show that the first two PCs account for almost the total amount of the original variance (99.21%).

The score-plot of $PC_1$ and $PC_2$ is showed in Fig. 2b: the pixels (the objects) are well separated along $PC_1$ in two clusters that represent text (black pixels) and background (white pixels). We must conclude that the first principal component explains the differences between text and background.

The differences observed in the score plot can be explained using the loading plot showed in Fig. 2c. The information accounted for by the first component is characterized by an equal contribution of all wavelengths. Instead $PC_2$ has a typical contrast effect since the contribution of the farthest IR and visible wavelengths are opposed: it has a positive contribution particularly at 970, 980 and 990 nm and a negative contribution particularly at 670 nm.

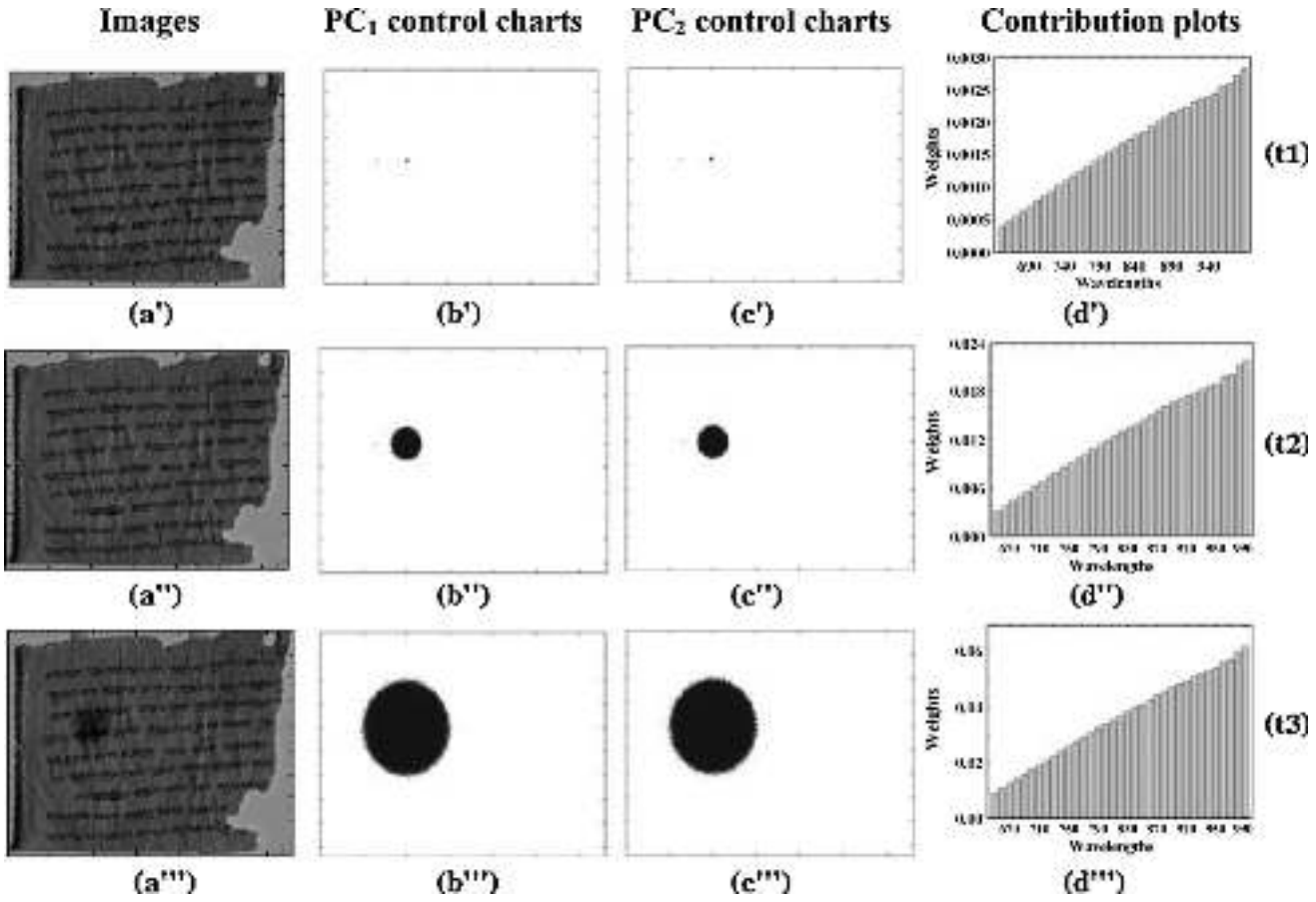


**Fig. 5.** Images of the degraded parchment (t1: a′, t2: a″, t3: a‴); $PC_1$ control charts (t1: b′, t2: b″, t3: b‴); $PC_2$ control charts (t1: c′, t2: c″, t3: c‴); contribution plots of the pixel in the center of the degraded area (t1: d′, t2: d″, t3: d‴).

Each 3-way characterization array was then projected onto the PC space using the loadings of the first and the second PC. The result is a matrix of scores filtered to account for the information corresponding to the first two PCs respectively where the 35 original channels (35 wavelengths) have been compressed into two new orthogonal channels ($PC_1$, $PC_2$). The image obtained from the scores along $PC_1$ (Fig. 3a) is very similar to the original image (Fig. 1) as only differences related to pigmented and clear areas are accounted for. Fig. 3b illustrates the scores of the second component where $PC_2$ accounts for a significant part of the dataset information.

### 5.2. Monitoring the deterioration

The deterioration of the parchment was simulated by adding to the 3-way arrays already modified with the white noise, a Gaussian shape modification with predetermined size. The area of the degradation covered 1600 pixels of the original image and all the wavelengths were interested by the degradation: the intensity of the modification increased from visible to IR (Fig. 4b). The spatial intensity of the degradation is a Gaussian surface whose maximum intensity is located in the pixel of coordinates $x = 400$ and $y = 400$ of the original image as shown in Fig. 4a. Parchment degradation was spatially modulated as a Gaussian surface since we want to focus the degradation within a region of interest, with an effect decreasing homogeneously from the center of the Gaussian surface to the borders; in this way it is possible to evaluate if control charts can detect increasing levels of degradation and how they behave at the border.

In the simulation of the degradation process, the intensity and the area of the Gaussian surface were linearly increased over time, providing three new 3-way arrays of progressively more deteriorated images (t1, t2, t3; Fig. 5a).

Each degraded 3-way array was projected onto the PC space previously calculated from the characterization images and the scores of the PCs were compared with the UCL and LCL limits. The Shewhart control charts for $PC_1$ and $PC_2$ are reported in Fig. 5b and c respectively. These images show how the control chart is able to precisely identify the degradation process, either with respect to $PC_1$ and $PC_2$: black pixels correspond to out-of-control pixels, namely pixels beyond the LCL for $PC_2$ and the UCL for $PC_1$. The highlighted areas in the control charts correspond perfectly to the simulated degraded areas, and the charts show how the degradation correctly increases along time.

There are few out-of-control pixels outside the degraded area. These are false alarms and are caused by the natural variability of the imaging system. Since the number of analyzed pixels in the image is very large ($n = 1{,}240{,}800$), the probability that few false alarms are present is also quite large. Nevertheless the false alarms do not show a systematic behavior and they are typically spread all along the image. This explains the importance of controlling very precisely the acquisition of the images and the matching and characterization steps, that permit an accurate account of all sources of natural variability in the experimental system when it is *in-control* conditions.

The contribution plots in Fig. 5d show which wavelengths are affected by the degradation process of a specific pixel. They can be

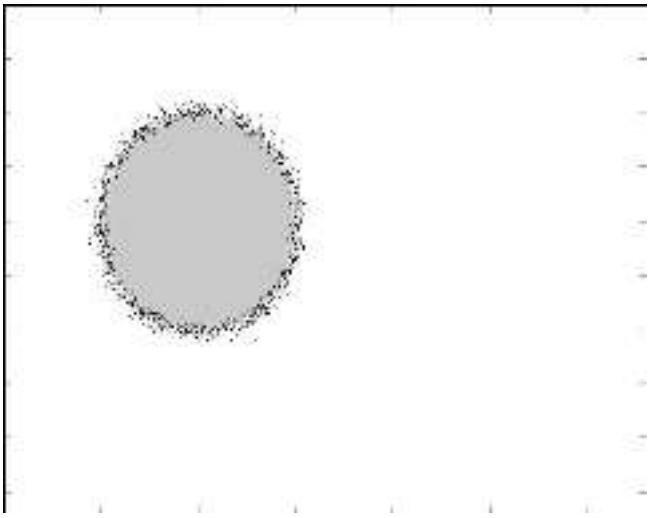

**Fig. 6.** $PC_1$ control chart of the degradation at t3 with the detection of the vanishing of the degradation effect (black pixels).

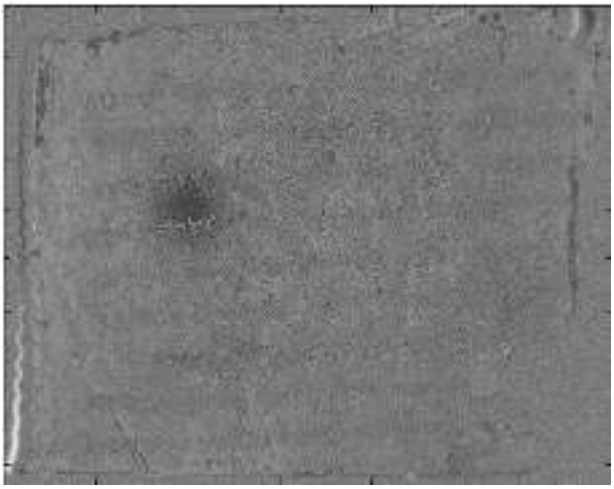

**Fig. 7.** Scores of the first component of the residuals of the degradation at t3.

calculated for each pixel of the image. In the present case, most of the wavelengths contribute to the perceived degradation. This is illustrated in Fig. 5d, which shows the contribution plot for the pixel of coordinates $x = 400$ and $y = 400$, namely the center of the degradation ellipsoid. Observing the contribution plots (Fig. 5d) it is evident that the contribution plot of the degraded pixel has the same trend of the degraded spectrum (Fig. 4b).

As the reflectance spectrum decreases the higher the values observed on the contribution plot.

Control charts also detected the vanishing of the degradation effect at the border of the Gaussian shape used. Fig. 6 reports a particular of $PC_1$ control chart of the degradation at t3: grey pixels correspond to out of control pixels, i.e. beyond UCL, while black pixels represent pixels that are close to the UCL but which are still in control, i.e. beyond $2/3 \times$ UCL.

The out-of-control situations identified by this procedure correspond to pixels showing an effect due to the degradation that can be explained looking at the loadings of the significant PCs calculated from the characterization images: this means that the degradation mechanism involves structures and species that were already present during the characterization.

#### *5.2.1. Principal component analysis of the residuals of the degradation analyses*

With the aim of further investigating the effects of the degradation process, a new PCA was performed on the residuals of the degraded images obtained by subtracting the information accounted for by the first two PCs (calculated from the characterization images) from the degraded images. As previously stated, the residuals of the degraded images can contain the information not accounted by the two PCs that contain the systematic information: if the degradation involves species and structures that were not present during the characterization, this information cannot be accounted for by the relevant PCs that were built just on the characterization data.

The first principal component accounts for 62% of the original variance and was considered the only relevant PC. $PC_1$ is mainly constituted by the IR wavelengths at 990 and 950 nm.

The analysis of the residuals identifies degraded areas only for the third virtual degradation where the aging process was more intense. The first and the second cases do not evidence unusual distribution of the residuals. Fig. 7 illustrates the projection of the scores of the first principal component of the residuals for the degradation image at t3. As it can be observed, the only information present is related to the degradation. The other part of the image does not show structured information and, for example, the symbols present on the original DSS are not represented as their information is accounted for by the first two PCs employed for the monitoring. The elliptical spot present in the figure that corresponds to the degraded area, is due to changes of the spectra that are not accounted for by the two PCs calculated on the characterization images. This area shows that a new shape of the spectra has originated from the simulated degradation. This new shape of the spectra can normally be associated with the formation of new species on the degraded surface. In the present case, the change of the shape corresponds to the spectra of the Gaussian surface that was artificially added to the original image to simulate the degradation.

## 6. Real case application

The developed technique was applied to a new piece of commercially manufactured modern parchment using a LED multispectral imaging to analyze a sample of parchment before and after a degradation treatment. Fig. 8a represents the parchment sample used: some signs were drawn on the surface to help alignment and the evaluation of degradation of both parchment and ink. Thirty-three 3-way arrays of multispectral images at 13 different wavelengths from UV–vis–NIR were collected for the training session. Then a simulation of accelerated deterioration of the parchment was performed under controlled conditions for accelerated aging in a PGC (parameter generation and control) aging chamber at 80 °C and 50% relative humidity for 6 h. One hour after the accelerated aging was completed the parchment was imaged under the same instrumental conditions used for the training. The parchment did not show changes detectable by the human eye. The procedure previously described for the simulated data was applied to build the control charts. The control charts for the first 2 PCs are represented in Fig. 8e and f and blue and red pixels in the charts correspond to out-of control points of the parchment: blue pixels represent regions that exceed the LCL and red pixels are the regions that exceed the UCL. White pixels correspond to in-control points. The large number of red and blue pixels show the evidence that some degradation took place. The wavelengths mainly affected by the degradation can be identified by means of the contribution plot of $PC_1$ (Fig. 8b) that shows the contribution plot for the blue pixel of coordinates $x = 400$ and $y = 600$: the band at 365 nm is the one most

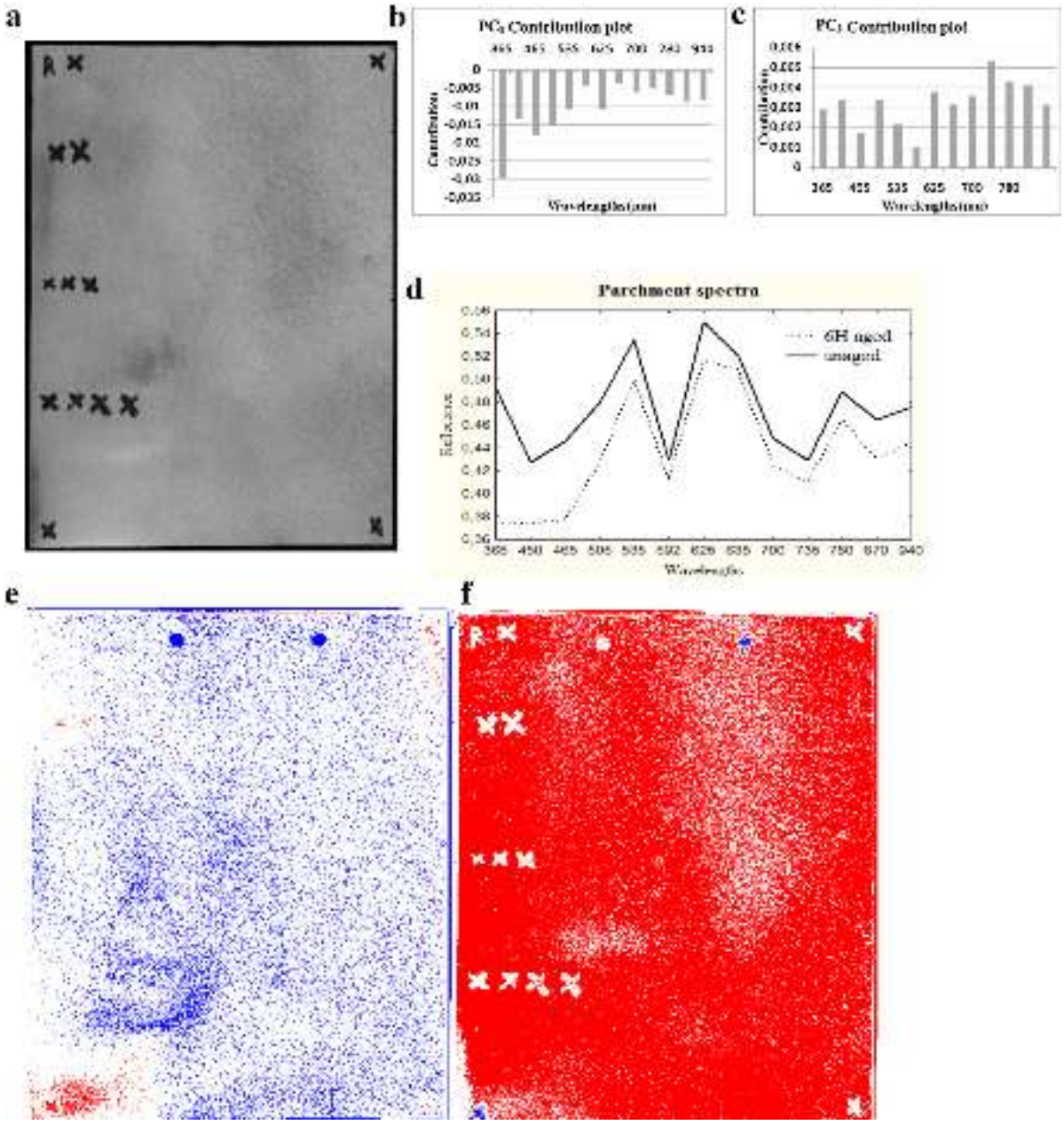


**Fig. 8.** Parchment image at 365 nm (a); un-aged and 6 h aged parchment spectrum for the pixel of coordinates $x = 400$ and $y = 600$ (b); contribution plot of $PC_1$ for the pixel of coordinate $x = 400$ and $y = 600$ (c); contribution plot of $PC_2$ for the pixel of coordinates $x = 400$ and $y = 600$ (d); and control chart of the first (e) and the second component (f).

affected by degradation since it shows the largest negative contribution. Blue pixels in Fig. 8e show a small reflectance at 365 nm, while red pixels show an opposite behavior. Fig. 8d represents the spectra of the same pixel before and after degradation: the contribution plot of $PC_1$ and the degraded spectrum have the same behavior.

The contribution plot of $PC_2$ of the same pixel in Fig. 8c shows why the pixel is out-of-control in the corresponding chart (Fig. 8f). In Fig. 8f wide areas of the image contain red pixels, which exceed the UCL: these pixels are characterized by a larger reflectance towards the IR region (largest positive contribution in the contribution plot of $PC_2$). The two blue circles in the upper side of the chart of $PC_1$ and $PC_2$ represent the holes that were made after the characterization to put the parchment in the oven for the degradation.

## 7. Conclusions

The aim of this project is the development of a non-invasive technique based on multispectral imaging for monitoring the conservation of cultural heritage objects.

The simulation shows that the control charts built on the relevant PCs are able to accurately detect the degradation process. In the first two simulated data scenarios, the degradation was not visible to the human eye, but could be detected by the algorithm. The contribution plots help the analyst to determine which wavelengths are involved in the assessment of the degradation process.

PCA performed on the residuals matrix of the degraded images enabled the identification of effects on the surface not accounted for by the PCs previously calculated on the characterization data, i.e. spectral evidence related to the possible occurrence of structural

changes of the sample surface because of the aging treatment and not just a decrease/increase of the intensity related to the species originally present on the surface.

A crucial step to make this methodology work in a real application is to collect an appropriate training set to measure the natural variability of the system: if the natural variability is not correctly assessed false alarm will be obtained. In the present case 5 characterization 3-way arrays were considered sufficient since the natural variability was simulated by the addition of a white noise of known entity. A more relevant number of replications should be collected during the characterization phase to assess the actual noise structure in real applications.

The results of the simulations and of the real application are encouraging: the proposed approach could be a good strategy for a fast and accurate control of cultural heritage goods preservation state. Today, there are many types of multispectral imagers that are employed in several fields from biomedicine to industry, from biology to cultural heritage conservation: this technique can be a useful tool and can be applied to any kind of surface that needs to be monitored.

## References


[1] W. Christens-Barry, F. France, K. Knox, R. Easton, M. Toth, SPIE, vol. 7249, 2009.
[2] F. France, W. Christens-Barry, M. Toth, K. Boydston, in: D.G. Stork (Ed.), SPIE, vol. 7531, 2010.
[3] R.L. Easton Jr., K.T. Knox, W.A. Christens-Barry, K. Boydston, M. Toth, D. Emery, W. Noel, Proc. SPIE (2010) 7531–7612.
[4] R. Padoan, Th.A.G. Steemers, M.E. Klein, B.J. Aalderink, G. de Bruin, National Archive of The Netherlands, Postbus 90520, 2509 LM The Hague, The Netherlands.
[5] K. Knox, Proc. SPIE 6810 (2008) 681004–681011.
[6] R. Easton, W. Noel, Gazette du Livre Médiéval 45 (2004) 39–49.
[7] G.H. Bearman, S. Spiro, Bibl. Arch. 59 (1996) 56–66.
[8] D.M. Chabries, S.W. Booras, G.H. Bearman, Antiquity 77 (2003) 296.
[9] G. Bearman, W.A. Christens-Barry, J. Palarch's, Arch. Egypt/Egyptology 6 (7) (2009) 1–20.
[10] M.J. Carlotto, M.B. Lazaroff, M.W. Brennan, SPIE, vol. 1819, 1992.
[11] P. Colarusso, L.H. Kidder, I.W. Levin, J.C. Fraser, J.F. Arens, E.N. Lewis, Appl. Spectrosc. 52 (1998) 106A–120A.
[12] V.C. Paquit, K.W. Tobin, J.R. Price, F. Mèriaudeau, Opt. Express 17 (14) (2009) 11360–11365.
[13] J. Bentley, T.J. Schneider, Comput. Stat. Data Anal. 32 (2000) 465–483.
[14] S. Baronti, A. Casini, F. Lotti, S. Porciniai, Chemom. Intell. Lab. Syst. 39 (1997) 103–114.
[15] W.A. Shewart, Economic Control of Quality of Manufactured Product, Van Nostrand, Princeton, NJ, 1931.
[16] D.C. Montgomery, Introduction to Statistical Quality Control, 3rd ed., Wiley, New York, 1991.
[17] W.H. Woodal, D.J. Spitzner, D.C. Montgomery, et al., J. Qual. Technol. 36 (3) (2004) 309–320.
[18] X. Pan, J. Jarret, J. Appl. Stat. 31 (4) (2004) 397–418.
[19] E. Marengo, E. Robotti, M.C. Liparota, et al., Anal. Chem. 75 (20) (2003) 5567–5574.
[20] E. Marengo, et al., Talanta 63 (2004) 987–1002.
[21] J. MacGregor, M. Bharati, Ind. Eng. Chem. Res. 37 (1998) 4715–4724.
[22] J. Prats-Montalbàn, A. Ferrer, J. Chemom. 21 (2007) 10–23.
[23] Dardenne, et al., J. Chemom. 18 (2004) 341–349.
[24] Y. Yadin, The Message of the Scrolls, Simon and Schuster, New York, 1957, pp. 161–162.
[25] S.W. Booras, D.R. Seely, ZPE 165 (2008) 175–185.
[26] T. Vo-Dinh (Ed.), Biomedical Photonics Handbook, CRC Press, 2003.
[27] G. Bearman, W. Christens-Barry, K. Boydston, Proc. of Eikonopoiia, Helsinki, Finland, 2010, pp. 108–114.
[28] W.A. Christens-Barry, K. Boydston, R.L. Easton, Proc. of Eikonopoiia, Helsinki, Finland, 2010, pp. 27–38.
[29] K.C.S. Pillai, in: S. Kotz, N.I. Johnson (Eds.), Encyclopedia of Statistical Science, vol. 3, Wiley, New York, 1983, pp. 668–673.
[30] H. Hotelling, in: Hastay, Wallis (Eds.), Multivariate Quality Control-Techniques of Statistical Analysis, McGraw-Hill, New York, 1947.
[31] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S.D.E. Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam, 1998.
[32] D.L. Massart, B.G.M. Vanderginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.
[33] A.J. Hayter, K.L.J. Tsui, Qual. Technol. 26 (1994) 197–208.
[34] S. Wold, K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst. 2 (1987) 37–52.