BERNARD J. AALDERINK, MARVIN E. KLEIN, ROBERTO PADOAN,
GERRIT DE BRUIN, AND TED A. G. STEEMERS

# Quantitative Hyperspectral Imaging Technique for Condition Assessment and Monitoring of Historical Documents

## ABSTRACT

Quantitative hyperspectral imaging (QHSI) is a non-destructive remote sensing technique which can detect small changes of the optical characteristics of material surfaces before they become visible for the human eye. The *Nationaal Archief* (National Archives of the Netherlands, The Hague) is conducting an applicability study on the use of QHSI for detecting, measuring and visualizing optical changes in historical documents caused by aging process and conservation treatments.

Repeated hyperspectral measurements of a document, taken for example before and after an exhibition, can be used to detect, map and classify subtle changes of the document condition with very high spatial resolution. Due to small differences in the position and deformation of the document during a storage or exhibition period, the measurement data first need to be spatially aligned. Mathematical transformations have to be applied to the hyperspectral image data in order to ensure that any particular pixel coordinate refers in both measurements to exactly the same location on the object. Then the differences between two corresponding spectral images from the first and second measurement can be calculated pixel-by-pixel. Non-zero differences of the pixel values mark optical changes within the document, which can be visualized in grayscale or color-coded images that help the conservator to identify the most critical areas on the document.

By calculating the differences, not only for individual spectral images but for the entire two hyperspectral data cubes, it may be possible to distinguish various degradation effects and provide a detailed statistical description of the spectral changes of the recorded sample. This makes the QHSI technique a valuable tool for an objective assessment of the document condition.

## INTRODUCTION

Most archival institutions regularly assess the condition of their collections in order to determine whether documents are suitable for transport, public exhibition and access by researchers. This procedure typically includes a detailed visual inspection of the documents by an expert, supplemented by a photographic documentation and possibly the application of non-destructive measurement techniques at a few selected locations. However, conventional condition assessments can easily fail to detect subtle changes of the varied materials composing historical documents. There is in fact the risk that small optical changes over large areas or large changes in very small areas are overlooked, because they were not documented with sufficient precision in previous condition reports.

Quantitative hyperspectral imaging (QHSI) is a non-destructive imaging technique that allows one to measure and document the optical characteristics at millions of object points simultaneously and with high accuracy. The *Nationaal Archief* (National Archives of the Netherlands, The Hague) are investigating the potential of this technique to quantify and map optical changes of documents resulting from deterioration processes with high spatial resolution. Samples were manufactured from different materials and artificially aged in several steps to simulate the effects that exhibitions can induce in original documents under controlled conditions. After each aging step or exhibition period the samples were measured with the SEPIA quantitative hyperspectral imager owned by the *Nationaal Archief*.

In addition to these measurements on artificially aged samples, hyperspectral measurements are carried out also on a series of original documents in order to monitor natural aging process and the effect of exhibition environments directly on the original artifacts.

In this article we discuss techniques to compare the hyperspectral data sets obtained from repeated measurements of the same object so that it becomes possible to characterize and visualize local changes in the reflectance spectra.

---

Poster presented at AIC's 38th Annual Meeting, May 11–14, 2010, Milwaukee, Wisconsin.

## MEASUREMENT PRINCIPLE

The SEPIA quantitative hyperspectral imager used in these experiments is based on two wavelength Tunable Light Projectors (TULIPs), which illuminate the document under an angle of 45°. These TULIPs are combined with a monochrome digital camera which records the document from above (fig. 1). The TULIPs subsequently illuminate the document with a series of 70 well-defined optical wavelengths in the ultra-violet, visible and near-infrared wavelength range (365–1100 nm). At each wavelength, a 4 megapixel grayscale image of a document area of 125 mm x 125 mm is recorded, corresponding to a resolution of 60 $\mu$m x 60 $\mu$m per pixel (ca. 400 dpi).

To translate the pixel values into quantitative measurements of the local spectral reflectance of the document, the recorded images at each wavelength band have to be compared to recordings of a reference target. In this case a white reference target (Spectralon target, supplied by Labsphere inc.) with known reflectance is used for this calibration step. After this calibration, the value of each image pixel represents a precise measurement of the fraction of light reflected from the corresponding tiny document area at this particular wavelength and can be regarded as a local quantitative reflectance measurement. This hyperspectral imaging technique is therefore referred to as *quantitative hyperspectral imaging* [Klein et al. 2008]. The entire set of these (calibrated) spectral images is called the *hyperspectral data cube*. It contains for each pixel the entire spectral reflectance curve (fig. 2).

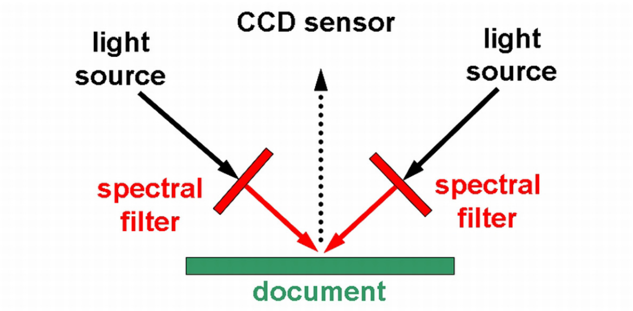

Fig. 1. Schematic over view of the QHSI instrument setup

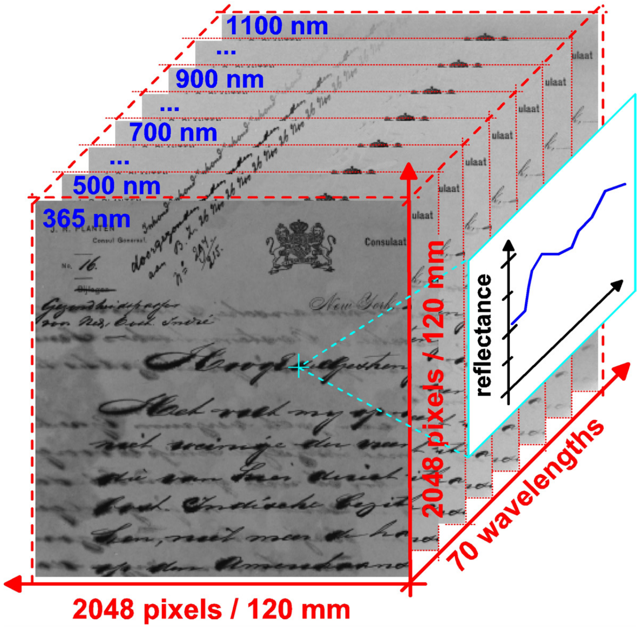

Fig. 2. Schematic representation of the hyperspectral data cube

The spectral information in the hyperspectral data cube can then be used to distinguish different writing materials such as inks and pigments, to enhance the legibility of degraded texts, to determine deterioration effects on a document following an exposition or treatment, and to measure spectral changes caused by aging processes.

## ALIGNING HYPERSPECTRAL MEASUREMENT DATA

In order to develop and test different techniques for analyzing series of hyperspectral recordings a number of discarded documents of the *Nationaal Archief* were artificially aged in 5 steps. During each aging period, the samples were exposed to changes of temperature, relative humidity and to a certain light dose, as described in detail in [Padoan et al. 2009]. In particular, the light dose and the spectral content of the irradiation were set to simulate in each accelerated aging step the effect of 1 month of light exposure received by documents in the exhibition room of the *Nationaal Archief*. The lighting conditions for the documents in this exhibition room[1] are expected to induce only a minimal cumulative light aging effect [Johnston-Feller 2001, Schaeffer 2001, Thomson 2002].

Before, between and after the artificial aging steps hyperspectral imaging measurements of the sample were carried out. These hyperspectral measurements are referred to as recording R0, R1, R2, R3, R4 and R5.

Figure 3 shows color images of one of these documents (a 19th century printed paper cut in a format of 63 mm x 20 mm) that were calculated from the hyperspectral data measured before aging was applied (R0), after 2 aging steps (R2) and after 5 aging steps (R5), respectively. A careful comparison of these calibrated images indicates that the artificial aging has induced a very small intensification (darkening) of the foxing that had developed due to more than 100 years of natural aging before artificial aging was applied.

In order to be able to quantify for all points on the document the spectral changes caused by the artificial aging, the corresponding spectral images of all three hyperspectral measurements have to be aligned relative to each other with pixel-accuracy. Figure 4 illustrates the importance of the alignment of the hyperspectral data prior to the pixel-by-pixel analysis. Figures 4A and 4B show in grayscale the difference of

Fig. 3. Photos of artificially aged original document before aging, after 2 and after 5 aging steps simulating as many months of exhibition in a suitable environment

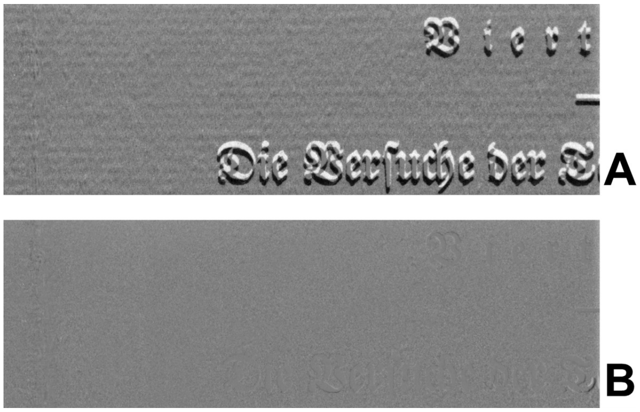

Fig. 4. Difference of calibrated 900 nm spectral images recorded before artificial aging and after 5 aging steps. White=positive values, dark=negative values, medium grey = zero difference; A) without alignment of hyperspectral data cube; B) with alignment of data cube

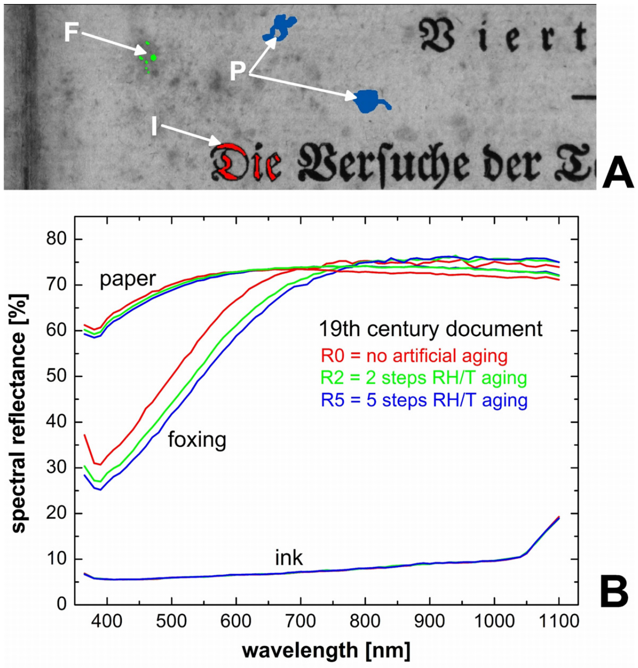

Fig. 5. A) Regions-of-interest (ROIs) that were defined on the artificially aged object sample for foxing (F), paper (P) and ink (I) areas. For all aging steps exactly the same areas were used due to the alignment of the hyperspectral data. B) Spectral reflectance curves of all ROIs for the three aging steps showing changes of several percent for the foxing area

the pixel values of the 900 nm calibrated spectral images taken from the R5 recording and from the R0 recording. Positive difference values are displayed as white pixels, negative values as dark pixels, and zero difference as medium gray pixels. Figure 4A shows the difference image calculated without prior correction for the mainly vertical shift of the recorded area by about 0.7 mm (11 pixels) between the two hyperspectral recordings of the document. Even such a very small misalignment means that inked areas do not overlap well in both spectral images. In fact, the overlap becomes inaccurate in any of the 70 spectral images, because all spectral images within the same hyperspectral measurement are very well aligned. Without prior alignment of the two measurements, any (small) change of the spectral characteristics of the sample induced by the artificial aging will be completely obscured by the large calculation error caused by the spatial misalignment. Figure 4B shows the difference image of the same 900 nm spectral images after a so-called image registration algorithm was applied to correct the shift, rotation and warp of the R5 image relative to the sample area recorded in the R0 image. The homogeneous medium grey proves that the alignment of both 900 nm images has worked very well especially around the text areas. By applying exactly the same spatial correction to all other hyperspectral images of the R5 recording, each of them becomes fully aligned with its counterpart in the R0 recording. In this way it becomes possible to directly compare the entire spectral reflectance curves of even very small sample areas such as thin ink lines or small foxing spots.

## AGING-INDUCED CHANGES OF SPECTRAL CURVES IN REGIONS-OF-INTERESTS

Having aligned the R2 and R5 recording relative to the R0 recording of the artificially aged document sample, any regions-of-interest (ROIs) that are marked in a spectral image of one of the recordings mark exactly the same document areas in all spectral images of all three recordings. Figure 5A shows three ROIs that were manually marked on the 500 nm spectral image of the R5 recording, covering small areas of ink, paper substrate and foxing, respectively. For each of these ROIs, the mean calibrated spectral reflectance curves of their pixels were extracted from the three hyperspectral measurements. Figure 5B shows these spectral curves with a different color for each of the three measurements.

The three curves for the ink area overlap perfectly, which indicates a high reproducibility of the measurements in combination with the spatial alignment. While the artificial aging obviously has not affected measurably the optical characteristics of the ink, this high reproducibility guarantees that any observed spectral changes are indeed significant and not a measurement error.

The curves for the paper ROI show a small, systematic reduction of the spectral reflectance at the shorter visible wavelengths (365–600 nm) with a maximal difference of less then 2% between the R0 recording and the R5 recording. In the infrared, an increase of the spectral curve from the R0 to the R2 recording, but no further change from the R2 to the R5 recording is observed. The maximal difference is less then 2% in the infrared, as well.

The foxing areas show a much stronger reaction to the artificial aging with a considerable drop of the spectral curve over the entire visible range up to about 800 nm. In the green and red spectrum, the absolute change peaks at about 9%. Such a large change of the (visible) reflectance spectrum—especially in comparison with the normal paper substrate itself—could be expected to be clearly visible also in the calibrated color images of figure 3. However, the foxing is not homogeneous and the high-intensity foxing area marked as the ROI is very small, so that such very local changes can be overlooked or underestimated considerably in a visual comparison of the color images.

## AGING EFFECTS AT SELECTED WAVELENGTHS

Comparing the mean spectral curves of ROIs provides valuable quantitative information about the influence of artificial aging on the optical characteristics of different document areas. However, by selecting the ROIs the spatial variation of the optical characteristics within the measured document area is completely neglected. There is the risk that the results obtained for the ROIs are not representative for the entire area, because the manual definition of ROIs is always subjective to some degree. Since all spectral images of all measurements are aligned, it has become possible to compare the document before and after aging on a pixel-by-pixel basis, i.e. with high spatial resolution. From the ROI analysis we know how the spectral curves for foxing areas and normal paper areas change. Based on these results the wavelengths 450 nm, 610 nm and 710 nm were chosen for the pixel-by-pixel analysis. For each of these wavelengths, the grayscale calibrated spectral image of the R0 measurement (no aging applied) was subtracted from the corresponding grayscale spectral images of the R2 and of the R5 measurements. The resulting three difference images for each the R2 and the R5 measurement were assigned to the blue, green and red channels of the two false-color images shown in figure 6A) and 6B), respectively. Neutral gray indicates areas which have no measurable change in reflectance in any of the three wavelengths. If an area has a measurable change in any of the three selected wavelengths the corresponding pixel is colored.

Fig. 6. False-color images showing the spatial variation of the effect of artificial aging on the spectral reflectance at three different wavelengths. Red channel: 710 nm; Green channel: 610 nm; Blue channel 450 nm. A) Difference image between 2 aging steps and no aging. B) Difference between 5 aging steps and no aging

In both false-color images the ink areas have in fact a neutral gray, which means that these areas were not influenced in a measurable way by the artificial aging. As opposed to this, the paper substrate and especially the foxing areas are shown in color already in the R2 image. This color intensifies in the R5 image and additional foxing spots become visible. This visualizes the progression of the spectral changes caused by the repeated artificial aging.

## COMPARING SPECTRAL CURVES WITH PIXEL-RESOLUTION

All spectral images within the data cube of a single hyperspectral measurement are aligned with pixel-accuracy. This means that for the tiny area on the object corresponding to each pixel the entire spectral reflectance curve is available for processing. Due to the alignment of the entire hyperspectral data cubes with respect to each other, it is possible to compare for all pixels the spectral curves measured for the R0, R2 and R5 artificial aging steps.

When comparing two spectral curves of the same object area measured before and after aging it is very convenient to express the similarity of both curves by a single number. In our case one has to use a suitable function to calculate from the 70 pairs of corresponding spectral values this

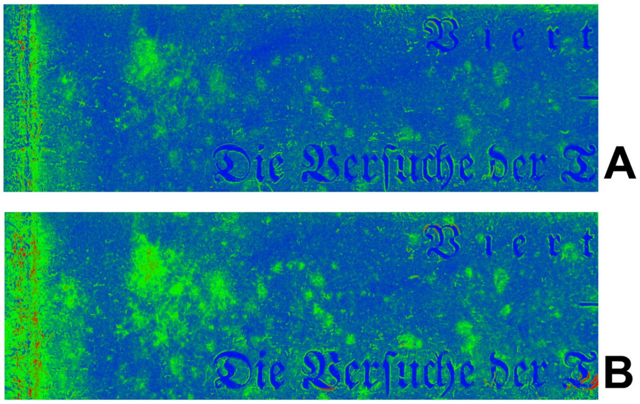

Fig. 7. False-color images showing the increase of the extension and intensity of the foxing due to artificial aging. Red pixels have a high, blue pixels a low SDS values, which indicates large and small changes in the spectral curves (380–750 nm) of the corresponding area. A) SDS image for R2. B) SDS image for R5

single number that measures the similarity of the two curves. There are many different functions that serve the purpose of comparing spectral curves, such as the Spectral Distance Similarity (SDS) and the Modified Spectral Angle Similarity (MSAS) function [Homayouni et al. 2004] which are commonly used in geosciences for the analysis of hyperspectral data of the Earth surface.

To compare the spectral curves of the R2 and of the R5 measurement with the R0 measurement, the SDS values for all image pixels were calculated for the visible part of the spectral curve (380 to 750 nm), in which the ROI analysis had revealed the most significant effects of the artificial aging. Figure 7A and 7B show the color-coded results of the SDS calculation respectively for the R2 and the R5 measurement with the R0 measurement. Pixels that have similar spectral curves before and after artificial aging (high SDS value) are colored blue. This means that if the aging had induced no spectral changes at all on the document, both images would be completely blue. Those pixels whose spectral curves were changed significantly by the aging are shown in green color and object areas suffering the largest effects are indicated by red pixels.

Figure 7 shows that the text area on the document has practically not been affected by the artificial aging, as has already been indicated by the results of the previous analysis steps using selected ROIs and selected wavelengths. The comparison of Figure 7A (2 aging steps) with Figure 3A reveals that the largest changes occur in areas where foxing is already present. In Figure 7B it can be seen how this foxing then further intensifies by the artificial aging up to R5. The comparison of the (calibrated) color images in figure 3 shows only very small changes in the document, which can be easily overlooked. By applying the SDS analysis to the aligned hyperspectral data cubes, such changes can be measured and visualized with high contrast.

The SDS analysis based on the aligned data cubes demonstrates the capability of the QHSI technique to measure small changes of the spectral characteristics of the document, even if induced by only a few months of exhibition in an optimal environment.[1]

## SUMMARY AND CONCLUSIONS

The quantitative hyperspectral imaging technique provides calibrated spectral curves of the measured document area with the high spatial resolution of digital images of several million pixels. In this paper we demonstrate how this technique can be used to investigate aging effects as they occur for example when documents are displayed in museums or other public exhibitions for extended periods of time. By applying artificial aging to a sample document the effects of a few months of exhibition of a historic document in a well-controlled environment were simulated. At different stages of the artificial aging, hyperspectral measurements of the sample document were carried out.

On the hyperspectral images, three regions-of-interest (ROIs) were defined manually. Especially the ROI defined in an area strongly affected by foxing showed changes of several percent of the spectral reflectance values in the visible spectral range for a simulated aging of several months of exhibition, whereas the print areas showed no change at all.

As opposed to comparing individual (spectral) images taken at the different aging steps, such ROI analysis has the advantage of allowing a comparison of the entire spectral curves. However, the manual definition of regions-of-interest is always subjective and it involves a high risk that the obtained spectral curves are not representative for the entire document. This problem can be overcome by using the hyperspectral image data to compare in parallel for all document locations the spectral values measured at the different stages of aging.

In order to fully exploit the high spatial resolution of the technique, image processing has to be applied to align spectral images taken at different aging steps with pixel-accuracy. Then it is possible to calculate for each pixel of the spectral images the changes of the spectral values induced by the aging. By combining the difference images between the artificially aged and not-aged sample document calculated for the wavelengths 450 nm, 610 nm and 710 nm, a false-color image was generated. Such false-color images calculated for two different aging stages visualize how the foxing spreads and intensifies on the paper substrate. They also show that the paper substrate itself shows only little changes, while no change at all is seen in the ink areas.

This type of false-color images maintain the full spatial resolution of the hyperspectral measurements, which allows one to detect with high sensitivity spectral changes at any location on the document for the three select wavelengths. However, especially in the case of historical documents, the pre-selection of only 3 specific wavelengths will in general not suffice to detect and visualize the spectral changes of all materials present on the artifact with the required sensitivity. Due to the alignment of all spectral images it is possible to compare for each pixel the entire spectral curves measured at the different stages of aging. By using the so-called spectral distance similarity (SDS) function, false-color images were generated that show spectral changes on the document at all wavelengths in the range from 380 to 750 nm.

In conclusion, the pixel-wise full-spectrum analysis discussed in this paper is a unique feature of the quantitative hyperspectral imaging technique that enables the detection of spectral changes without a prior, subjective selection of the analyzed area or wavelength. This feature is of particular importance due to the great diversity of materials and degradation effects that can be encountered even on a single document. For studying and documenting optical changes of historical documents during storage or exhibitions quantitative hyperspectral imaging is therefore a very valuable addition to the toolbox of non-destructive measurement techniques that are available to the book and paper conservator to assess and monitor the condition of historical documents.

## ACKNOWLEDGMENTS

The authors would like to thank Dr Michael S. Brown, Department of Computer Science, School of Computing at the National University of Singapore, for fruitful discussions.

## NOTE

1. In the internal exhibition room of the *National Archief* documents are displayed at a temperature of 20 C and a relative humidity of 49%. The illumination is set to 50 lux with UV filtering applied. Documents are displayed for 48h per week over a period of three months, following a rotation procedure that generally allows a rest period of two years for each exhibited document.

## REFERENCES

Homayouni, S., Roux, M., Hyperspectral image analysis for material mapping using spectral matching, ISPRS Congress Proceedings 2004.

Klein, M.E., Aalderink, B.J., Padoan, R., de Bruin, G., Steemers, Th.A.G., Quantitative Hyperspectral Reflectance Imaging, Sensors 8:5576–5618 (2008).

Padoan, R., Klein, M.E., Aalderink, B.J., de Bruin, G., Steemers, Th.A.G., Monitoring aging processes of archival documents by means of quantitative hyperspectral image: A part of the hyperspectral project at the Nationaal Archief (National Archives of the Netherlands), The Book and Paper Group Annual Vol. 28, 2009.

Johnston-Feller, R., Color science in the examination of museum objects: nondestructive procedures. The Getty Conservation Institute, Los Angeles 2001.

Thomson, G. The museum environment. Bodmin: Butterworth-Heinemann. Oxford 2002, pp. 28–31

Schaeffer, T.T., Effects of light on materials in collections: data on photoflash and related sources, The Getty Conservation Institute, Los Angeles 2001.

BERNARD J. AALDERINK
Applied Research Engineer
Art Innovation
Oldenzaal, the Netherlands
benno.aalderink@art-innovation.nl

MARVIN E. KLEIN
Senior Research Engineer
Art Innovation
Oldenzaal, the Netherlands
marvin.klein@art-innovation.nl

ROBERTO PADOAN
Paper Conservator
Nationaal Archief
The Hague, the Netherlands
roberto.padoan@nationaalarchief.nl

GERRIT DE BRUIN
Head of Restoration and Conservation
Nationaal Archief
The Hague, the Netherlands
gerrit.de.bruin@nationaalarchief.nl

TED A. G. STEEMERS
Preservation Policy Officer, National Archivist
Director of the Collections Department, Nationaal Archief
The Hague, the Netherlands
ted.steemers@nationaalarchief.nl