David Howell
Bodleian Libraries, University of Oxford

"RECENT ADVANCES IN THE SCIENTIFIC INVESTIGATION OF SPECIAL COLLECTION MATERIALS IN LIBRARIES"

Abstract

Although advanced imaging and analytical techniques have long been available, it is only over the last three to five years that they have been more generally applied to serious study of library materials. This has been driven by incredible advances in technology which allow for non-invasive, non-sampling, non-contact methods to be designed specifically with the curators' concerns for the safety of the material in mind. At the same time the curators' increasing awareness of what is possible has increased the demand for such services.

For Special Collections, one of the outstanding developments has been hyperspectral imaging. This is a technique where an item is scanned and each 'pixel' represents a full colour spectrum, commonly 400–1000 nm. The large files produced in this way can be interrogated by sophisticated software to reveal erasures and obscured areas of detail invisible to the naked eye. Hyperspectral imaging can also be used for identifying pigments, especially organic materials. This technique is completely safe if applied correctly and with continuous dialogue between curators, conservators and the people carrying out the analysis/imaging.

When applied in combination with other techniques available in the Bodleian libraries including Raman spectroscopy and X-ray fluorescence spectroscopy it has been demonstrated that a great deal of very useful information can be revealed. It is now possible to employ science to date manuscript material, and to infer where items might have been made. We can also use hyperspectral imagery to reveal underlying textual and pictorial content previously obscured from view.

This paper will examine how these techniques have been employed with varying degrees of success on a range of materials including the late medieval Gough Map of Great Britain, and several examples of mediaeval illuminated manuscripts, all items held within the Bodleian Library's collections.

Keywords: hyperspectral imaging, Raman spectroscopy, X-ray fluorescence spectroscopy, Special Collections

BACKGROUND

The Bodleian Libraries have a reputation for a long history of excellence in conservation. The current conservation studio is purpose built within the recently (2015) refurbished Weston Library, the Special Collections Library of the Bodleian Libraries. The conservation section is divided into paper, book and preventive conservation and is staffed by six Institute of Conservation (Icon) accredited conservator/restorers and around ten other conservators and technicians.

A more recent and related addition to the Bodleian's research effort is Heritage Science[1], a specialised service for a number of 'customers' including conservators but also readers, curators and academics. The name Heritage Science was deliberately chosen to reflect this wide customer base and to differentiate the service from 'conservation science' which is usually embedded within a conservation department. The Heritage Science section has quickly grown from just one person, David Howell, who is head of the section, to having Richard Mulholland, a Postdoc researcher funded by the Leverhulme Trust, Ian Maybury, a DPhil (PhD) student funded by the University of Oxford Fell fund and the EPSRC Centre for Doctoral Training in Science and Engineering in Arts, Heritage and Archaeology (SEAHA) and Dáire Browne, a 4th year undergraduate chemist. The section also has access to appropriate analytical equipment for a Special Collection Research library which will be described here.

HYPERSPECTRAL IMAGING (HSI)

In hyperspectral imaging we produce a reflectance spectrum for each point in a scan. Hyperspectral sensors, more commonly known as imaging spectrometers or hyperspectral cameras, collect spectral information across a continuous spectrum by dividing the spectrum into many narrow spectral bands. Hyperspectral imaging spectrometers can have up to several hundred bands with a spectral resolution of 10 nanometres (nm) or narrower. This is much greater spectral resolution than multispectral setups which commonly have around nine to twelve spectral bands and a spectral resolution of around 100 nm (Spectral resolution refers to the width of each band within the captured spectrum). In this way a hyperspectral scan will produce a file which contains very detailed colour information for each pixel, from perhaps 300 to 1000 bands from 380 nm up to to 1000nm although systems are available which scan well into the short-wave infrared (SWIR) region up to 2500 nm. This dense information is really only useful when interrogated using computational methods. Like multispectral imaging, the technique is very useful in revealing hidden texts and other information invisible to the naked eye.

1 http://www.bodleian.ox.ac.uk/our-work/heritage-science

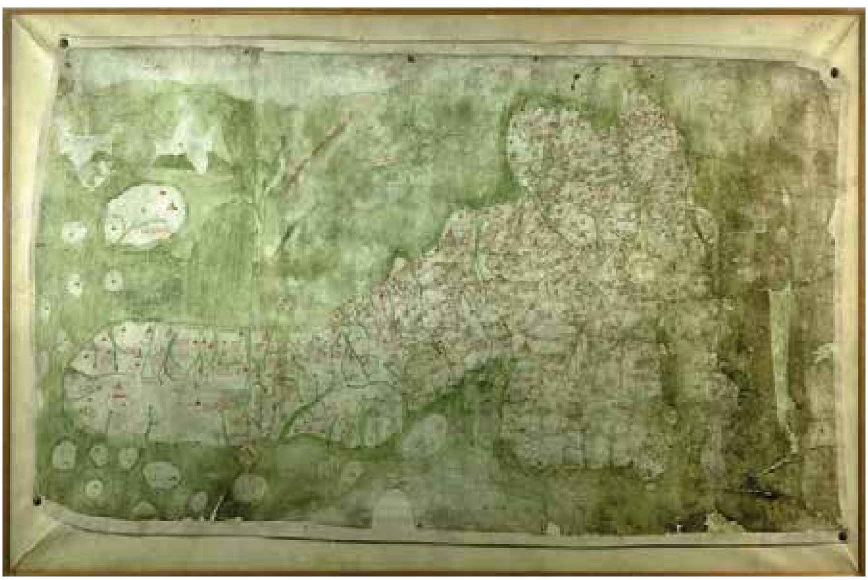
Headwall Hyperspectral Imaging

The Headwall system pictured here is called a push broom setup, a line scanner where the objects moves under the scanner line in the same way that a photocopier scans a page. One of the major advantages of this system is the ease of setup and short period of time and training required to get up and running. The scanning is carried out with bespoke software, Hyperspec® III, written specifically for operating the equipment and is both simple and intuitive. The setup can take perhaps half an hour to complete but once done a scan can be completed in perhaps 10 minutes for an average size 'normal' manuscript or printed text. If there are several similar objects which are approximately the same height and thus same distance from the lens they can simply be scanned sequentially with no adjustment to the setup. This allows this equipment to provide a 'service' rather than only being used for 'projects'.

The Headwall system has 1600 Spatial bands (pixels) and 923 Spectral bands from 380 to 1000 nm. This region of the spectrum is called Visible and Near Infrared (VNIR) and this produces very complete reflectance spectra with high spatial resolution but the resultant files are very large, sometimes creating folders containing files up to 30 GB in size.

There are many software packages for looking at the spectral data (more correctly termed hyperspectral cubes) in these large files. The Bodleian libraries uses ENVI®, a high cost but extremely useful program written specifically for looking at multispectral and hyperspectral files. ENVI® is the choice at Bodleian as the Hyperspec® III software produces files in a format that can be imported into ENVI with no further processing. Incidentally, it is also used by the heritage imaging experts at Rochester Institute of Technology and at The National Gallery of Art, Washington, amongst others in the heritage field. It has built-in algorithms for grouping together spectra that have some similarity. Which algorithms are used for any particular object depends on the nature of the object and the information that is being sought. However, because so many different operations have been automated, it requires very little knowledge to get up and running with simple but effective analysis. Because ENVI® is written for a wide range of users there are several hundreds of automated algorithms, most of which are not useful for the study of library materials. But the use of software is the only way that the vast amount of data can be interrogated. The mathematics are simple. Say a scan is taken of a square object. The Headwall scanner is 1600 pixels across so the total number of pixels in a square will be 2,560,000 pixels. Then each pixel is represented by 923 pieces of spectral data giving a data set of 2,362,880,000 components. And this is a square scan; the scanner can be left to scan a much longer strip giving even more data.

THE GOUGH MAP



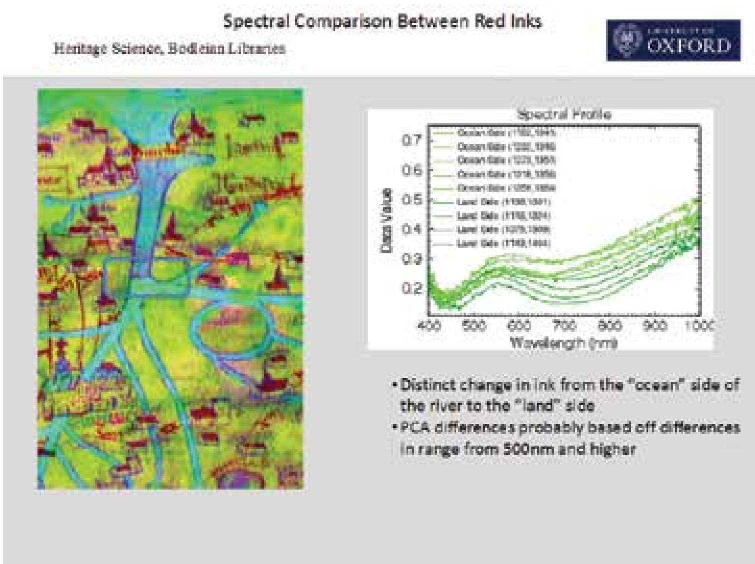
The Gough map

The first object we scanned when the Headwall equipment arrived in Oxford was the famous late medieval Gough map and this has been reported elsewhere.[2] This map is considered to be the earliest surviving map of the island of Great Britain

2 Millea N., Howell D. (2018) Revealing the Past: How Science Is Unlocking Cartographic Secrets. In: Altić M., Demhardt I., Vervust S. (eds) Dissemination of Cartographic Knowledge. Lecture Notes in Geoinformation and Cartography. Cham: Springer.

and consists of around 600 place names and over 200 rivers. Although the map clearly depicts late medieval England and Wales quite accurately (Scotland less so) there remains a number of academic queries. Have certain place names been erased or changed? Is it possible to detect different pigments? Essentially, is there information remaining that is not visible to the naked eye? Because of the status of the object the map was chosen as a priority. Hyperspectral imagery of the map was initially captured in January 2015, but early results were somewhat disappointing even after a great deal of computational analysis of the data.

But the Gough map demonstrates the value of hyperspectral data. The archived Gough files were uploaded to Rochester Institute of Technology where their experts analysed the data and they are beginning to show signs of revealing more information, much to the excitement of the academics involved. It is early days but it is hoped that the data will continue to yield information leading to increased knowledge of the map over time. The Gough map definitely comes into the category of an ongoing project.



Comparison of red inks in the Gough map

BRIEF LIVES BY JOHN AUBREY

Brief Lives is a collection of short biographies written by John Aubrey (1626–1697) in the last decades of the 17th century, the manuscripts of which are held by Bodleian Libraries. We were approached by an expert and publisher on Aubrey, Dr. Kate Bennett, from Magdalen College, Oxford, to scan the front cover of one of the manuscripts. Her question, which she had already investigated with infra-red photography, was could we reveal the title 'Brief Lives' on the page as it is not visible with the naked eye. The scan took only a few moments and it was startlingly quick using the software to reveal an indisputable 'LIVES"! This is an example of using the equipment as a service. An academic or reader comes to us with a clear academic query. We analyse (scan) the objects and carry out computation that will hopefully yield information quickly. We then pass that 'new' information on to the 'customer', who interprets it to produce new knowledge.

'LIVES' revealed on the cover of Aubrey's manuscript

MS. ARCH. SELDEN. A.2 (PREVIOUSLY CODEX SELDON BUT NOW REFERRED TO AS CODEX AÑUTE AFTER THE PLACE WHERE IT WAS MADE)

Although it is technically a colonial document, because it is made in such a precolonial style Codex Añute is considered to belong to the group of less than twenty precolonial Mexican Codices still in existence. For over sixty years it has been known that this manuscript was in fact two books in one, a palimpsest. Underneath the gesso surface in the Codex Añute were remains of an older manuscript. In previous invasive research, which today would be considered totally unacceptable, one blank page had the gesso scraped off to reveal part of the underlying codex. These pre-Columbian codices are pictographic books, the scribe using pictures or 'glyphs' to express meaning rather than letters. Unfortunately, the glyphs revealed in the 1950s were not sufficiently clear to be able to increase knowledge except that to prove they were there.

As part of a larger project some of the pages of Codex Añute were scanned and after extensive analysis, over nearly a year, clearer results were obtained and the work published[3]. The story caught the imagination of an international audience and the work was reported in well over a hundred media outlets. This is proof that heritage science can be used to produce that increasingly important and demanded outcome, impact. Impact is important not only to justify resourcing the work itself but also to enhance the reputation of Bodleian Libraries as a serious research institute carrying out its own pioneering research as well as enabling others to do theirs.

Mexican Palimpsest revealed

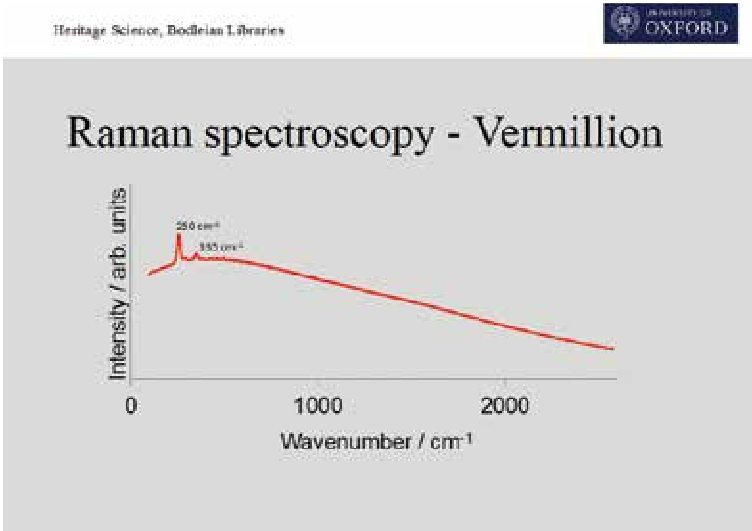
THE SCIENTIFIC ANALYSIS OF PIGMENTS

As well as revealing hidden images and text, hyperspectral imaging has the power to provide some chemical characterization, especially when used in combination with other techniques. Pigment identification is extremely popular because it offers information of use to many 'customers'. Conservators can use pigment iden-

3 Snijders, L., Zaman, T., & Howell, D. (2016). Using Hyperspectral Imaging to reveal a hidden precolonial Mesoamerican codex. Journal of Archaeological Science: Reports, 9, 143–149. doi: 10.1016/j.jasrep.2016.07.019.

tification to help with identifying restorations and interventions, monitor degradation, and to inform conservation strategies. Characterizing pigment palettes can sometimes help inform academics of the provenance and date of objects. Social scientists may use the technical knowledge to understand trade routes, cultural interactions and production techniques

The SEAHA DPhil mentioned earlier is currently involved in the development of pigment analysis. The reflectance spectra produced in the VNIR range tend to have quite broad featureless peaks. This rather sparse data can be enhanced by correlating to more definitive analytical techniques. The techniques being used at Bodleian Libraries in combination with hyperspectral imaging are x-ray fluorescence spectroscopy (XRF) for elemental analysis and Raman spectroscopy for molecular analysis. In libraries, the most common uses of XRF is in the identification of metallic constituents such as gilded illumination and the confirmation of a pigments identity. Pigments containing heavy metals such as Copper (Egyption blue and azurite), Arsenic (orpiment and realgar) and Mercury (vermillion) are particularly suitable for XRF analysis. XRF works when an x-ray beam with enough energy interacts by displacing electrons from the inner orbital shells of the atom. This displacement occurs as a result of the difference in energy between the primary x-ray beam and the binding energy that holds electrons in their proper orbits. Electrons are fixed at specific energies in their positions in an atom, and this determines their orbits. When electrons are knocked out of their orbit they leave behind vacancies making the atom unstable and the atom corrects the instability by filling the vacancies. The vacancies are filled from higher orbits that move down to a lower orbit where a vacancy exits. For example, if an electron is displaced from the innermost shell of the atom (the one closest to the nucleus), an electron from the next shell up can move down to fill the vacancy. This is fluorescence. It is the energy of the fluorescence that gives a 'fingerprint' for each element. Raman spectroscopy is an increasingly popular and powerful tool in the analysis of works of art and manuscripts. It is a non-invasive and non-damaging means of characterizing samples in situ and provides unequivocal evidence for the presence of materials. For applications in libraries it is normally coupled with a microscope which allows detail to be obtained from tiny areas a few tens of micrometres in area down to the identification of single grain of a pigment. It reveals levels of information about the chemical composition of a material that are unmatched by other techniques. The Raman equipment comprises the shining a single wavelength or colour of light generated by a laser onto the sample and then collecting and analysing the spectrum of the scattered light. Most of the scattered light known as the Rayleigh scatter is exactly the same colour as that of the incident light source. But a tiny fraction, about one millionth of the light shone on the sample, loses some of its energy to the material and changes its wavelength. This is known as Raman scattering. Due to the need for the conservation of energy, the difference between the energy of the incident light and the Raman scattered light corresponds to the vibrational energies of the molecules. By analysing the spectral profile of the scattered light, we can obtain the vibrational spectrum of the sample and again we have a different 'fingerprint'.



Example of Raman spectrum

A range of known, mainly readily available, pigment sample 'swatches' were prepared as part of the Leverhulme funded project. These have been invaluable in 'calibrating' and comparing the three techniques. In addition to this the techniques have been used in turn on areas of 'real' special collection objects. Whilst this project is at quite an early stage the results are extremely promising. The advantage of 'imaging' over the spot analysis is the ability to rapidly map where pigments are on an object.

Using hyperspectral imaging to identify pigments using reference spectra

BAKHSHALI MANUSCRIPT

One of the 'treasures' of the Bodleian libraries is the Bakhshali manuscript, an archaeological mathematical text written on birch bark excavated in 1881. It is named after the location in present-day Pakistan, where it was found. It is said to be not only the oldest manuscript in Indian mathematics but also possibly the earliest use of a zero in mathematics. The birch bark has been protected by being mounted in mica, a transparent mineral, but the bark is considerably darkened making it difficult to read. Hyperspectral imaging was carried out on one of the 70 leaves with very good results with the text being much more legible. It is hoped that eventually all of the leaves will undergo hyperspectral imaging.

Interestingly, although the terms 'oldest' and 'earliest' were frequently used to describe the Bakhshali manuscript, the date was actually unknown with estimates based on content, and with the contention that it was a copy of an earlier version. In late 2016 a request came from the Science Museum in London to borrow a leaf from the manuscript for the exhibition 'Illuminating India: 5000 years of science and innovation'. Attached with this request was a further question; could the Bakhshali be radiocarbon (C14) dated? After much discussion, Richard Ovenden, Bodley's Librarian made the decision on the advice of the Head of Heritage Science that the dating could go ahead. In 2017, three samples from the manuscript were carefully removed from 3 different leaves and were shown by radiocarbon dating to come from three different centuries: from AD 224–383, 680–779, and 885–993.

This information has created a great deal of discussion within the academic community both in Mathematics and Sanskrit. Many people disagree with the dates but the profile of the object, the subject of the history of the language and mathematics has been brought to a massive audience who would not in other circumstances been interested in the subjects. This is tremendous impact increasing the status of Bodleian Libraries as not only an enabler of research but also a research institute in its own right.

CONCLUSION

Heritage science can be of enormous benefit to enhance service to readers, scholars and researchers. But on a bigger stage the outputs of heritage science can be used to create impact. Impact is of course critical at a time when so many libraries are under threat, especially in the U.K. In the Internet and digital age libraries need to rebrand and demonstrate that they make a difference. The results of heritage science can help in that process. Routine revealing of hidden texts and routine pigment identification reinforce the library service. Big projects and major discoveries create impact. Both are achievable and necessary.