DIGITAL PAINTING ANALYSIS, AT THE CROSS SECTION OF ENGINEERING, MATHEMATICS AND CULTURE

Bruno Cornelis, Ann Dooms, Jan Cornelis, Frederik Leen, Peter Schelkens

ABSTRACT

Museums started digitizing their collections, mainly for preservation, documentation and dissemination purposes, but one can go further than archival: digital acquisition enables digital analysis of art and in particular paintings. This proves to be a useful, non invasive tool, for many applications such as restoration, conservation, art history, material and structure characterization, authentication, dating and even style analysis. Paintings are complex structures. Analysis of the support and all pictorial layers requires multimodal imaging, high resolution and standardized data acquisition. We situate our own research on canvas characterization, craquelure detection, multispectral data analysis, object extraction and image enhancement within the state of the art.

1. INTRODUCTION

Thanks to advances in the field of image acquisition as well as the wide range of imaging modalities currently available, museums widely started digitizing their collections. This is mainly for the purpose of archiving our vast cultural heritage, but it also enables the scientific analysis of art objects. In this paper, we will focus on the digital analysis of the painting support and its pictorial layer.

New modalities to acquire paintings include extremely high resolution digital photography, multispectral imaging, digital radiography, ultraviolet fluorescence and infrared reflectography. All these have specific advantages and bring additional information to the study of paintings. Multispectral imaging for example enables an objective color measurement, which is more precise than regular RGB, by providing significantly wider gamut information. Next to providing more precision on colors, it also allows to study the painting at wavelengths that lie beyond the visible domain, which can lead to a richer analysis. X-ray imaging penetrates deeper or even goes through the pictorial layer, revealing parts of the painting’s support (which can be made of various types of canvas or wood).

More exotic modalities such as confocal X-ray fluorescence are used to study the pictorial layer at different depths [1]. This is extremely interesting for the study of the painting’s creation process. Synchrotron radiation based X-ray fluorescence made it even possible to uncover a lost painting by Vincent van Gogh who often reused canvas of an older painting and painted a new or modified composition on top [2].

Early projects and initiatives in the context of art involved the design of different image acquisition systems such as high precision colorimetric cameras. A lot of effort was put into finding an objective and scientific way of capturing and archiving color measurements of paintings. Some early examples include the European Commission (EC)-supported VASARI (Visual Arts System for Archiving and Retrieval of Images) project [3] where a colorimetric scanner system for direct digital imaging of paintings was developed and its successor the MARC (Methodology for Art Production in Colour) project [4]. A high resolution and multispectral imaging system was developed as part of the CRISATEL EC project [5], which is dedicated to the digital archiving of paintings. The goal is to obtain an image representation independent of any specific illuminant and be able to provide high fidelity reproductions of the paintings (on display or print). The hardware of the CRISATEL multispectral imaging system has been developed by Lumiere Technology [6] and is composed of an electronically controlled lighting system and a high resolution digital CCD camera equipped with thirteen bandpass filters.

It is clear that handling and visualizing the resulting images from these types of devices can be quite difficult due to their high resolution and data volume (in multispectral imaging for example we obtain one image per spectral wavelength considered). Next to acquisition hardware, adequate image handling and distribution software needs to be developed as well. An example of a free image processing system is VIPS [7], which was developed to handle large images, to work with colour and to perform scientific analysis. Amongst VIPS users we have various universities and museums such as the National Gallery, the MoMA and the Louvre. The Viseum [8] and Acohir EC projects worked on a system to make extremely high-resolution images accessible over the internet and eventually resulted in an open source light-weight client-server system called IIPImage. A more detailed overview of these projects (and others) can be found in [9, 10].

Recently, efforts have been made to acquire art on a larger scale and to make the results available to the research community and the larger public. Within the framework of the CIP-ICT PSP call by the EC, the expert center PACKED is coordinating the project “Digitising Contemporary Art” which aims at digitising contemporary art objects from small to middle-sized museums and other art institutions. Another nice example of art dissemination is ARTstor, a digital library of more than one million images in the areas of art, architecture, the humanities, and social sciences with a set of tools to view, present, and manage images for research and pedagogical purposes. The Digital Library is made available to the arts community as well as to disciplines outside of the arts. Recently, Google released their Art Project which makes it possible to explore several museums and view hundreds of art works at extremely high resolutions, online.

This rapid evolution in acquisition of art also enables the analysis of the art object through its digital image counterpart. Hence, a crossdisciplinary interaction emerges between image analysts, mathematicians and art historians, putting to use recent advances made in the field of image processing. This is materialized through the Digital Painting Analysis (DPA) initiative1, bringing together several research teams from universities and museums to tackle art related questions such as artist authentication, dating, etc. To stimulate the interaction between the art historical world and branches of digital image processing, the DPA initiative organizes several work-

---

1http://www.digitalpaintinganalysis.org/

---

This research was supported by the Fund for Scientific Research Flanders (project G.0206.08, the Ph. D. fellowship of Bruno Cornelis and the postdoctoral fellowship of Peter Schelkens).
shops named IP4AI (or Image Processing for Artist Identification) at the Van Gogh museum in Amsterdam, the MoMA in New York and in Brussels, together with a symposium, to facilitate a dialog between the involved communities. The Van Gogh Museum and the Kröller Müller Museum made it possible for the participating teams to work with high resolution digital images of paintings (mostly van Goghs) from their collections.

Work performed in the digital processing and analysis of paintings includes the development of algorithms in the context of restoration, conservation, art historical matters and style analysis. Some of these algorithms are applied on a whole set of paintings, but it is not uncommon to focus on paintings by a specific artist (cfr. the work performed on van Gogh paintings by several members of the DPA initiative) or even one painting such as the study of the Mona Lisa [11] and our previously published work on a Gauguin painting [12] and the Ghent Altarpiece [13, 14]. The main advantages of digital analysis techniques is that they form an extensive painting [12] and the Ghent Altarpiece [13, 14]. The main advantages of digital analysis techniques is that they form an extensive set of non-invasive forensics tools (indeed, however much insight chemical analysis can yield, it requires the destruction of a sample from the painting and is therefore seldom allowed by conservators) and it can automatize time-consuming or error prone manual processes. We expect that the art community will gradually learn to use and trust these tools; a similar emergence and eventual success took place in the medical world in the mid 80s, with the advent of digital radiology and computed tomography.

Section 2 of this paper describes the study of the support of the painting and the answers it can provide while Section 3 deals with the study of the pictorial layer and its applications. Finally, we end the paper with conclusions and a suggestion for future work.

2. STUDY OF THE PAINTING’S SUPPORT

The support of a painting (either wood or canvas) can provide a surprising amount of information about the painting itself. Panel paintings (such as many Early Netherlandish paintings from the 15th and early 16th century) where the support consists of flat wooden panels, either in a single piece or many pieces joined together, were very popular until the emergence of canvas in the 15th century. Den drochronology (or tree-ring dating) has become important to art historians in the dating of panel paintings while for paintings on canvas, the counting of threads in the weft and warp directions can also provide an indication about a painting’s origin. The standard procedure for manual thread counts is time-consuming, tedious, prone to errors and sometimes beyond the scope of the human eye [15]. As mentioned earlier, X-ray images of paintings detect hidden information behind the first layer of paint such that a hidden painting or drawing can be revealed. However, on this kind of images the canvas becomes visible as well.

Extensive research [16, 17] has been performed on the analysis of these images, where the manual thread counting is replaced by a semi-automatic algorithm that produces thread density maps to obtain a “fingerprint” for the canvas which can then be used for dating and authentication purposes. Next to thread densities, the variations in thread angle are measured as well. This clearly indicates a phenomenon known as cussing which is the stretching of the canvas when priming the canvas cloth or due to nailing of the canvas on a frame. In [18] we extend upon the work mentioned above by studying how to characterize the canvas by extracting global features such as average thread width, average distance between successive threads (i.e. thread density) but also the spatial distribution of primers. These features are then used to construct a generic model of the canvas structure. Secondly, we investigate whether we can identify different pieces of canvas coming from the same bolt (a crucial element for dating, authentication and identification of restorations). Therefore we extract robust local features with techniques generally used for texture analysis, such as the co-occurrence matrix and the grayscale- as well as rotation invariant local binary patterns. Both the global characteristics mentioned earlier and these local properties are used to compare the “fingerprint” of different pieces of cloth coming from the same or different bolts (i.e. large rolls of canvas cloth). We performed this research on digital X-rays (taken by UZ Brussels in DICOM format) of primed canvas samples produced by a Flemish company (Claessens) where canvasses are still primed in a traditional manner. Different types of canvas were considered (from very high thread density to very rough thread density canvas) as well as different types of primers (see Fig. 1). The advantage of working with unpainted canvas is that this dataset constitutes a ground truth. This research is ongoing and the algorithms will be applied on X-rays of actual paintings.

3. STUDY OF THE PICTORIAL LAYER

The applications that belong to this category are numerous and range from the faithful and accurate representation of the pigments used during the painting to the study of the underlying pictorial layers. The presence of false colors (including the ones lying outside the visible domain) and identified different types of restorations that were either not very clear or even invisible to the naked eye. The first one is the application of masticque, a substance that is used to cover cracks. The restoration done with masticque was a bit sloppy and has a slightly different color than the underlying paint. An area of restoration with masticque is shown under label “1” in Fig. 2. The second type of restoration also concerns the concealment of cracks but by applying small brushstrokes of the same color as the background and is therefore very difficult to distinguish by eye.

Figure 1: X-ray patches of different types of primed canvas
3.2 Crack detection and digital inpainting

One of the most common deteriorations of paintings is the breaking of the paint layer (called craquelure or cracks). These manifest themselves in many patterns which, according to [21], are visible records of the physical tensions within the structure of the painting. These tensions are dependent on the choice of support, materials and methods employed by the artist. As for most 15th century Flemish paintings on Baltic oak, fluctuations in relative humidity, acting on the wooden support, are the main cause for crack formation. In the case of canvas paintings, the interaction between the canvas and the stretcher, the environment and physical impacts cause many different types of cracks. The detection of cracks can be used for damage assessment (i.e. determine a painting’s state of conservation) but is also the first step in digital restoration and can be used in judging authenticity. Generally speaking, cracks can be classified into many types, for example in [22], cracks are classified into 4 different categories namely, French, Flemish, Dutch and Italian.

Digital image processing can automatically detect (or even classify) cracks. A popular technique is the use of the top-hat transformation [23, 24]. As often dark cracks on a lighter background need to be detected the black top-hat (or closing top-hat) transform \( TH_B(A) \) is used, which is defined as the difference between the input image \( A \) and its morphological closing(s) \( \phi_B(A) \) using a structuring element \( B \), resulting in a grayscale image with enhanced details. Clearly, the structuring element should be chosen according to the size and nature of the cracks to be detected. Our goal in [18] is to increase the readability of a book painted in one of the panels of the Ghent Altarpiece (1432) through digital restoration (or inpainting). The detection of cracks is complicated by the letters in the book which have similar size and color properties as the accumulated dirt in the cracks. Also, most of the cracks are surrounded by a bright border which is caused by light that is reflected on the elevated paint around the cracks, or by cleaning, which exposes the underlying white preparation layer on some of these ridges. Not only do the cracks need to be detected and inpainted, but these bright borders as well. In order to detect cracks of varying size and avoid detecting letters, a multiscale morphological approach is proposed. A set of crack maps is constructed by applying the previously mentioned top-hat transform with structuring element sizes ranging from \( 2 \times 2 \) to \( 10 \times 10 \) pixels. These are then combined in a novel way; i.e. the crack maps of the two finest scales are added to form our base map which is used as a reference for selecting cracks in crack maps corresponding to coarser scales. This way of working has the major advantage that most unwanted structures, in this case the letters, that are often detected at coarser scales, will not be included in the final map, while cracks continue to grow from fine to coarser scales. The detection of the borders and their treatment as missing regions leads to improved inpainting results as well. Crack inpainting methods considered in the literature so far include order statistics filtering, anisotropic diffusion [25] and interpolation [26] which remove cracks, one pixel at the time. Since these pixel-wise methods do not perform sufficiently well in our case, a patch-based inpainting method [27] was improved by introducing a novel patch selection approach that we called constrained candidate selection (see Fig. 3 for results on a part of the book).

In [12], we apply the top-hat transform to each spectral band separately and the crack maps are superimposed, resulting in a better detection than when using a single image. Preliminary experiments show that in some regions pixels belonging to a crack (which appear very dark or black to the human eye) have a different spectral response than dark/black paint. This means that spectral properties will further allow us to detect cracks in the darker regions of the painting as well, which is an unsolved problem when working in the more traditional RGB or HSV color space.

3.3 Brushstroke analysis - quantifying artistic style

Recently, a plethora of statistical and mathematical approaches were applied in attempts to classify brushstrokes [28] or quantify artistic style [29]. A digital tool for classification and authentication of works of art is presented in [30] by inspecting first- and higher order wavelet statistics of a selection of drawings by Pieter Bruegel the Elder. By looking for (in)consistencies across different drawings or paintings, or, within a single work they address the so-called “problem of many hands” where regions of collaborative work have to be determined. In [29], sparse coding is used, stating that its success in the analysis of natural scenes indicates its possible use for modeling features in drawings and other 2D media. Sparse coding enables the distinction between the same set of drawings by Pieter Bruegel the Elder and a series of imitations. It is believed that, although an imitation may be perceptually similar, the subtle differences in strokes can reveal the presence of an imitation. In [31, 32] it is stated that the drips and swirls of paint by Jackson Pollock create fractal patterns. Analysing the properties or the “fractality” of these patterns offers a promising test for authenticating a Pollock drip painting. Different research teams within the DPA (see Section 1) from the universities of Princeton, Maasstricht and Penn State used statistics derived from different types of wavelets in order to assist in matters such as authentication, dating and style analysis [33]. As part of the Princeton team we give an overview of the tools and general methodology that is being used to tackle the challenges formulated by art historians, hence providing convincing arguments in favor of digital image processing. Detailed results can be found in [34, 35, 36].

For the analysis of paintings it is crucial to extract distinguishing features/statistics that truly characterize the style of an artist. It is obvious that simple image statistics such as mean and variance will not suffice by themselves because they do not describe image structure sufficiently well. More complex models that provide additional information are needed. The analysis approach chosen for the challenges of authentication, dating and style characterization consists of three main steps: transform, modeling and classification.

- **Transform**: A multiresolution transform, the Dual-Tree Complex Wavelet Transform (DTCWT), is performed on patches of
the image: it provides approximate shift invariance and directional selectivity (properties standard wavelet transforms lack). The DTCWT uses two parallel filter banks and produces six subbands of coefficients that allow analysis of changes in the image in six directions.

- **Modeling:** The large number of pixels, and thus also of transform coefficients, and the noise on the pixel values (due to the acquisition process) require robust dimensionality reduction and feature extraction techniques. Hidden Markov Trees (HMT) are used to design a statistical model to represent images. Due to the multiresolution nature of the wavelet transform, the wavelet coefficients can be arranged into a quadtree (one coefficient from a coarser scale corresponds to four wavelet coefficients at the next finer scale). The wavelet coefficients are modeled as samples from a mixture of two Gaussian distributions, one with a large variance for the coefficients corresponding to an edge and one with a small variance for coefficients from a smooth region. HMT models the statistical dependencies between wavelet coefficients at different scales.

- **Classification:** The model parameter vectors extracted in the previous step are used as the input for several types of classification algorithms such as Support Vector Machines, AdaBoost, Decision Stump and Random Forest.

The results obtained for the first three IP4AI workshops in Amsterdam and New York were promising. It is clear however, that these digital techniques on their own are not sufficient to provide conclusive answers to questions of interest to art historians. Nevertheless, they will likely be a worthy addition to the toolbox of art historians and conservators and have the great advantage of not being invasive. There is also still room for improvement in the different steps of the analysis of paintings. It is worth pointing out that in order to apply such techniques, the quality of the acquired dataset (i.e. high resolution images) is of utmost importance. Only images of equal quality can be meaningfully compared with each other.

### 3.4 Computer vision based analysis

Recently, computer vision based techniques were applied to problems concerning the history and interpretation of art. These new computer methods can sometimes be made more perceptive than even a trained artist or art historian in matters such as judging perspective [37] or tracking illumination within a painting [38]. The dewarping of distorted images in curved mirrors, depicted within paintings, provided new views into artists’ studios and engendered art-historical consequences [39]. Computer graphics reconstructions of artists’ studios even allows to investigate “what if” scenarios and thus gain insight into the methodology of an artist [40].

In [14], we focussed on very specific objects within the Ghent Altarpiece, i.e. pearls and beads. Their most important characteristics in paintings are their surface reflectance which makes the use of spatial information essential. Therefore, a set of techniques are proposed to analyze and measure spatial characteristics of the digital images of pearls, all based on the image spatiograms [41], which extend the concept of histograms to the spatial domain. We tested these using both the images of painted pearls (by the Van Eyck brothers (1432), and copyists Jef Van der Veken (1945) and Charlotte Caspers (2010)) and photographed pearls. Additionally, we consider the pearls painted by Hans Memling in his masterpiece Maria Maddalena Baronecci (1470). We used the so called spatiogram similarity metric to quantify the similarity between pearl images. An obvious use case for this metric is the detection of forgeries by comparing “suspicious” pearls to those painted by the original artist. Secondly, a novel set of metrics built around the spatial characteristics of the image data is proposed. Experiments demonstrate good correlation between the new metrics and visually observed image features such as smoothness of the transitions between areas of different lightness, and realism of the visual appearance of the painting (similarity with the photographed pearls). Lastly, a method for matching spatiograms of the images is introduced (see Figure 4). In the domain of pearls, this can be used either to assist the copyists of the pearl paintings, by teaching them about the characteristics of the original painted pearls, or again to assist the art historians in better understanding the differences or similarities between different artists and their styles of painting the pearls, or other jewels.

### 4. CONCLUSIONS

This paper describes the use of mathematics, statistics, image processing and computer vision techniques as an aid for the art community. When we look at the emerging number of publications in this domain and the increasing amount of special sessions in conferences and workshops, it is clear that this branch of research has evolved into a hot topic. The main advantage of all modalities mentioned in Section 1 and the digital analysis techniques that are covered in Sections 2 and 3 is that they form an extensive set of non-invasive forensics tools (in contradiction to chemical analysis which requires the destruction of a sample from the painting). We expect that the art community will gradually learn to use and trust these digital image processing tools. However, it is our belief that if we want to improve even further upon this collaboration there is a need for more open databases and algorithm repositories that allow to compare newly developed algorithms with current state of the art. Unfortunately, open access to data is still a difficult subject for many museums and clear agreements need to be made. As this research track will continue to grow and gain in popularity a need for standardization (e.g. in data as well as metadata formats) will arise. This is much like what the DICOM standard engendered in the medical world by making proprietary formats obsolete and enabling the manufacturer-independent data-pooling into the well-established PACS (Picture Archiving and Communications Systems) databases and the routinely exchange of large volumes of image data between hospitals, radiologists, general practitioners and research institutes.

### REFERENCES


