VIRTUAL UNDERPAINTING RECONSTRUCTION FROM X-RAY FLUORESCENCE IMAGING DATA

Anila Anitha, Andrei Brasoveanu, Marco F. Duarte, Shannon M. Hughes, Joris Dik, Koen Janssens, Matthias Alfeld

ABSTRACT

This paper describes our work on the problem of reconstructing the original visual appearance of underpaintings (paintings that have been painted over and are now covered by a new surface painting) from noninvasive X-ray fluorescence imaging data of their canvases. This recently-developed imaging technique yields data revealing the concentrations of various chemical elements at each spatial location across the canvas. These concentrations in turn result from pigments present in both the surface painting and the underpainting beneath. Reconstructing a visual image of the underpainting from this data involves repairing acquisition artifacts in the dataset, underdetermined source separation into surface and underpainting features, identification and inpainting of areas of information loss, and finally estimation of the original paint colors from the chemical element data. We will describe methods we have developed to address each of these stages of underpainting recovery and show results on lost underpaintings.

1. INTRODUCTION

In recent years, imaging and digital image processing have increasingly been used to aid in art restoration and reconstruction. The primary motivation is that imaging can be applied without harm to a work of art, and digital image processing techniques can then be used on the resulting data in order to enhance features of interest, remove obstructive artifacts, virtually undo the effects of aging, and/or merge data from multiple imaging methods (e.g. from visible, infrared, and ultraviolet light) to create a virtual restoration or reconstruction. One problem for which this strategy is particularly well-suited is that of virtually reconstructing underpaintings, i.e. paintings that have been painted over. Since these underpaintings often lie beneath priceless works of art, they may only be reached via non-invasive imaging. Many such underpaintings exist: for example, a recent X-ray analysis of 130 Van Gogh paintings at the Van Gogh Museum in Amsterdam showed that almost 20 of the 130, roughly 15%, contained some sort of underpainting [5].

A number of non-invasive imaging methods have historically been used to gain information about underlayers of the painting, the most common of which are classic X-ray photography, infrared imaging, and multispectral imaging. X-ray photography produces a single image of the canvas showing the combined X-ray absorbance of all layers in the painting (see Fig. 1(b)). Infrared imaging uses infrared light to penetrate the top layers of the painting and reveal infrared-reflective features beneath the surface (see Fig. 1(c)). It is particularly successful at imaging any preliminary sketches the artist made on the canvas before painting. Multispectral imaging combines images under several different wavelengths of visible, infrared, and perhaps ultraviolet light. However, the data these can provide is limited. As can be seen in Fig. 1, none of these imaging methods provides enough information for a complete visual reconstruction of the underpainting.

Hence, new types of imaging are being developed. In 2008, a team headed by coauthors J. Dik and K. Janssens developed a novel synchrotron-based X-ray fluorescence imaging method [4]. This non-invasive technique, when performed on a painting, gives a set of images showing the spatial distributions of specific chemical elements, including As, Ba, Bl, Cd, Co, Cr, Cu, Fe, Hg, Mn, Pb, Sb, Sr, and Zn. See Figs 1(e-g) for examples. Furthermore, a more recent portable imaging method developed by M. Alfeld et. al. [1] also uses X-ray fluorescence to produce chemical channel images, but eliminates the need for a synchrotron, allowing museums to produce (slightly noisier) images of this type in-house. This information regarding various chemical elements, as compared to other methods, provides the richest available data to use in reconstructing an underpainting. Hence, we will focus on this exciting new type of data when developing our methods.

Although a vast improvement over previous imaging methods, this set of chemical channel images still does not provide for art historians what a visual image of the original painting would. Different features of the painting are split across different chemical channels making it difficult to visualize the original work. Producing a visual reconstruction from this data aids art historical scholarship about the painter and his/her working method, oeuvre, etc. and allows recovery of these important lost pieces of our cultural heritage.

However, the process of producing such a visual reconstruction of the underpainting is non-trivial. First, as in any imaging process, this technique produces its own unique artifacts that must be corrected. Second, features of surface and underpaintings containing the same pigment will be mixed together in the corresponding chemical element’s image and must be separated. An example can be seen in the mercury (Hg) channel of the Van Gogh in Figure 1(f) where both the woman’s lips from the underpainting and a group of pink flowers from the surface painting (upper left) are visible. Third, surface painting features may block signals from the underlayers from reaching the surface to be imaged, so there are typically areas of loss in some chemical element images that must be identified and inpainted. This can be observed in the similar pattern of black spots in each of Figs. 1(e), 1(f), and 1(g). Finally, the various chemical channel images must be combined into a single visual image, including reconstructing the colors from the pigment data. We summarize this in the diagram in Figure 2 illustrating the four main stages of recovery. In this paper, we present methods we have developed for each of these four stages and show results. (Methods from Secs. 4.1 and 5 were previously described in [2].)

1.1 Data

We work with imaging datasets from two paintings. The first is of Vincent van Gogh’s “Patch of Grass”. As shown in Fig. 1, imaging data shows the presence of a woman’s portrait hidden under the surface painting. Of the 14 chemical channels measured, the antimony (Sb) and mercury (Hg) channels, corresponding to Naples yellow and Vermilion red respectively, contain the most information regarding the woman’s portrait. Small traces of it also appear in the arsenic (As), zinc (Zn), and lead (Pb) channels (Paris green, Zinc yellow, and Lead white pigments respectively). Fortunately for our recovery work, Van Gogh has made a major change in palette here between the first and second painting and, as a result, they contain mostly distinct pigments. The second painting is a portrait by the German artist Philipp Otto Runge (see Fig. 3). While there is not a completely distinct underpainting for the Runge canvas, the chemical channel images of the Runge painting suggest that the painting has been altered from its original form. In the cobalt (Co) and mercury (Hg) channels, we see that the woman originally had ribbons in her hair, matching the pigments in, and hence likely the color of, the ribbon on her dress. Also, the antimony (Sb), iron (Fe), and lead (Pb) channels show that she originally had long, flowing, curly hair and a ruffled low-plunging neckline.
with a fairly uniform distribution through the entire picture. The analysis of the portrait under “Patch of Grass” reveals that approximately 70 – 80 of the 698 horizontal lines (∼ 10%) are acquired in one scan, so shifts are identical across them. Examination of the portrait under “Patch of Grass” reveals that approximately 70 – 80 of the 698 horizontal lines (∼ 10%) are damaged, with a fairly uniform distribution through the entire picture.

In each distorted line, the amount of shift increases/decreases monotonically and thus appearing shifted in their respective horizontal lines. This shifting behavior is performed with the aid of a scanning X-ray beam that irradiates the object pixel by pixel, line by line horizontally, switching direction per line above and below. (See Figure 4.) This shifting behavior is the result of a timing problem in the acquisition. The analysis is required that allows up to $d = 2$ discontinuities in each line was enough to capture most of the occurring shift patterns.

For each horizontal line, we model the shift on that line, i.e. the displacement of each pixel as a function of horizontal position, as a piece-wise constant function with at most $d$ discontinuities. The shift function $S_{t,a}(n)$ thus has domain $\{1, \ldots, L\}$ where $L$ is the length of each horizontal line in pixels, and range $\{-M, \ldots, M\}$ where $M$ is the maximum allowable shift (we chose $M = 25$). It has $d$ discontinuities in each line was enough to capture most of the occurring shift patterns. We further restrict the parameter $d$ to $d = 1$ through total variation minimization. The rationale is that introducing artificial horizontal line shifts into an otherwise normal image will reduce vertical continuity in the image and increase the total variation as measured in the vertical direction. We thus minimize the variation, as measured vertically, as a strategy for correcting these artifacts.

Our method has two variants, titled single and dual [3]. In the single method, the parameters are chosen to minimize the total variation of the tested line with respect to the line above. In the dual method, the parameters are chosen to minimize the total variation of the tested line with respect to both the lines above and below.

For the single method, this means that letting $g_i$ and $f_j$ be the values of the $i$th pixel of the line in question and the line above it respectively, we then estimate the shift parameters as:

$$
\text{argmin}_{f_1, \ldots, f_n} \sum_{i=1}^{d+1} \sum_{k=1}^{l_i-1} |f_j - g_j + \sum_{k=1}^{i-1} a_k|^2
$$

where we have set $l_0 = 1$, $l_{d+1} = \min(L, L - \sum_{k=1}^{d+1} a_k)$. We further require that $l_n - 1 + \sum_{k=1}^{n-1} a_k \leq L$ for all $n$ and $l_n - 1 + \sum_{k=1}^{n-1} a_k \geq 1$ for all $n$, so that the sum remains defined. Each line is corrected before moving on to the next. However, to avoid the correction of

![Figure 1: Comparison of different types of non-invasive imaging on Vincent van Gogh’s “Patch of Grass” including (a,d) the surface painting under visible light, (b) traditional X-ray imaging and (c) infrared imaging of portions of the painting, (e)-(g) three different chemical channel images of the painting produced via the synchrotron X-ray fluorescence technique of [4].](image)

**2. RECOVERY STAGE 1: ACQUISITION ARTIFACT CORRECTION**

Chemical element images obtained through the X-ray fluorescence imaging technique may contain horizontal lines or portions thereof whose pixels are shifted horizontally left or right with respect to the lines above and below. (See Figure 4.) This shifting behavior is the result of a timing problem in the acquisition. The analysis is performed with the aid of a scanning X-ray beam that irradiates the object pixel by pixel, line by line horizontally, switching direction with each line [4]. Occasionally, this scanning beam gets delayed, which results in some of the pixels getting captured later than scheduled and thus appearing shifted in their respective horizontal lines. In each distorted line, the amount of shift increases/decreases monotonically, although the direction varies by line. All channel images are acquired in one scan, so shifts are identical across them. Examination of the portrait under “Patch of Grass” reveals that approximately 70 – 80 of the 698 horizontal lines (∼ 10%) are damaged, with a fairly uniform distribution through the entire picture.

![Figure 2: Diagram illustrating the four main subproblems in underpainting recovery.](image)
spurious shifts, only shifts whose pixel size in absolute value $|a_k|$ is greater than the threshold $T$ (we used $T = 5$) are actually corrected. 

For the dual method, we perform a similar minimization, but incorporate both the pixel values of the line above $f_{l-1}^n$ and those of the line below $f_l^n$ into the minimization:

\[
\arg\min_{I_{l-1}, a_1, ..., a_d, n=1} \sum_{j=l-1}^{d+1} \sum_{k=1}^{l-1} \left| f_{j}^n - g_{j} + \Sigma_{i=1}^{l-1} a_n \right|^2 + \left| f_{j}^n - g_{j} + \Sigma_{i=1}^{l-1} a_n \right|^2
\]

where $l_0$ and $l_{d+1}$ are again defined as above. No threshold is used here since the dual method seems to suffer less from spurious small shifts. When using this method, we estimate the correction for each horizontal line based on the lines above and below it, but save this information for later, rather than correcting the line immediately. Only after we finish estimating the corrections for all lines in the image do we proceed to correct all the found shift patterns. However, we may find afterwards that lines are still misaligned since each line’s correction has been based on the original positions of the lines above and below, which may have changed in the meantime. Hence, the method is applied iteratively, with the repaired channel hopefully converging to the true one.

Since the single method assumes the top horizontal line in each image to be correct and then corrects each subsequent line once based on the one above, this method is fast and works well when the erroneously shifted lines are uniformly sparse across the image, but problems can arise when an area of the image has many consecutive lines with shifts, because inexact fixes can propagate to the lines below. Meanwhile, the more intensive iterative dual method seems to be more robust and can be used in areas of high density acquisition errors without worry that errors will propagate from one line to the next. Hence, to produce the result of Fig. 4, we use a combination of the two methods. We use the single method with $T = 5$ to correct large single shifts in all areas except the woman’s forehead which is particularly dense with shifts. In this area, we use the single method with a threshold of $T = 15$, followed by 30 iterations of the dual method. Together these two methods corrected almost all the shift patterns in the image.

3. RECOVERY STAGE 2: SOURCE SEPARATION

The problem of separating surface and underpainting features in each chemical channel image is an underdetermined source separation problem. We wish to split the dozen or so chemical channel images into twice as many sources with only a single RGB color visible light image of the surface painting to aid in the separation. This side information, while potentially useful, is difficult to incorporate since it comes from a different type of imaging than the sources themselves; pigments show up very differently in the chemical element and visible light images. Nevertheless, we would like to separate each chemical element mixture image into two sources, one of which is related to the visible light surface image and one of which is unrelated. Inspired by the metrics for measuring relatedness between images of different modalities that are popular in the multimodal image registration literature (see e.g. [6]), we attempt to do this by minimizing the conditional entropy of one source given the visible light surface image while maximizing the conditional entropy of the other source given this surface image. For each chemical element, this translates into minimizing over $C_{surf}$ and $C_{under}$:

\[
\frac{H(C_{surf}|C_{under})}{H(C_{surf})} - \frac{H(C_{under}|C_{surf})}{H(C_{under})} + \lambda TV(C_{surf}) + \lambda TV(C_{under})
\]

subject to constraints

\[
C_{ch} = C_{surf} + C_{under}; C_{surf}(x,y) \geq 0, C_{under}(x,y) \geq 0 \forall x,y
\]

where $C_{ch}$ is the original chemical element image, $I_{surf}$ is the visible light surface image, $C_{surf}$ and $C_{under}$ are the two sources we are splitting into, $H(I)$ is the pixelwise entropy of the image $I$ (see [6]) and $TV(I)$ is the total variation of the image $I$. We solve this optimization problem via gradient descent, but we perform this gradient descent on the wavelet coefficients of $C_{surf}$ and $C_{under}$, which helps to maintain coherent image features and avoid pixelwise overfitting. We use a multiresolution approach, allowing the wavelet coefficients at the coarsest scales to converge before starting to work on the wavelet coefficients of the next finest scale as well. Figure 5 shows a preliminary result of this procedure on a synthetic example of the classic “peppers” and “baboon” images mixed to create a model chemical image. An RGB color “peppers” image models side information of a different modality. We note that this method seems to produce a good separation of the two underlying sources.

4. RECOVERY STAGE 3: IDENTIFICATION AND INPAINTING OF AREAS OF ATTENUATION

As previously noted, X-ray fluorescence imaging data may contain areas of information loss where a particularly thick or X-ray-absorbent surface feature has blocked signal from the underlayers.
This is clearly seen in Figure 1, in which all three channels of X-ray fluorescence imaging show a similar pattern of black spots, and we observe additional darkened streaks across the images, most notably in the antimony channel. We have developed two methods for identifying the locations of obstructing surface features so that the areas they have obscured can be inpainted. The first is based on identification of obstructing surface hues [2], typically representing pigments that are particularly X-ray absorbent. The second method attempts to identify surface features that are particularly thick using raking light photographs of the surface painting. We describe each of these in the next two subsections.

### 4.1 Hue Identification Method

We wish to identify surface hues that are highly correlated with darkened areas in multiple chemical channels. Thus, for each RGB color $\mathcal{C}$, we use the pixels of this color in the surface painting as a mask applied to the chemical channel images. We then compare the average grayscale value of the masked-off region with the average grayscale value in a small surrounding area. Surface colors for which the masked region is significantly darker than its surroundings are likely attenuating. More precisely, for each color $\mathcal{C}$ and chemical element $\mathcal{E}$, we compute the mean of the masked region,

$$\tau_{\mathcal{C}, \mathcal{E}} = \text{average}(I_{\mathcal{E}}(x, y) | I_{\text{surface}}(x, y) = \mathcal{C}),$$

where $I_{\mathcal{E}}$ is the image for chemical element $\mathcal{E}$ and $I_{\text{surface}}$ the surface image, as well as the mean of the surrounding area

$$\rho_{\mathcal{C}, \mathcal{E}} = \text{average}(I_{\mathcal{E}}(x, y) | I_{\text{surface}}(x, y) \neq \mathcal{C}, \text{ and } \exists x', y' \text{ s.t. } (x' - x)^2 + (y' - y)^2 \leq R, I_{\text{surface}}(x', y') = \mathcal{C})$$

for a given distance threshold $R$ (we used 50 pixels). We then compare $\tau_{\mathcal{C}, \mathcal{E}}$ and $\rho_{\mathcal{C}, \mathcal{E}}$ to determine whether the pixels under this surface color tend to be darker than the surrounding pixels. If the ratio $\tau_{\mathcal{C}, \mathcal{E}} / \rho_{\mathcal{C}, \mathcal{E}}$ exceeds a specified threshold (we used 1.1) and $\rho$ is sufficiently greater than 0, we consider this color $\mathcal{C}$ attenuating.

#### Figure 6: Results of the two attenuated area identification and inpainting procedures.

(a) Original antimony channel image showing attenuated regions. (b) Antimony image with estimated attenuation locations from the attenuating hue identification method marked in green. (c) Result of the attenuating hue identification method: Repaired antimony image with the locations marked in green in (b) (and a small surrounding neighborhood) inpainted. (d) Result of the thickness identification method: Attenuation under thick ridges reversed as described in Sec. 4.2.

An experiment on the Van Gogh selected shades of yellow, dark green, and occasionally pink as those likely to have attenuated the underpainting signal. These first two colors, yellow and dark green, clearly appear in a pattern mimicking that of the attenuated regions. The pink is initially surprising, but on more careful inspection, we see darkened areas in the antimony channel corresponding to the pink flowers in the upper left corner of the surface patch. Using the chosen colors, we then produce a mask for inpainting, by growing the area identified (shown in Figure 6(b)) slightly. Fig. 6(c) shows the result of inpainting using this mask.

### 4.2 Thickness Estimation Method

The second method attempts to estimate the locations and amount of attenuation based on the thickness of the paint layer from the surface painting. The procedure uses raking light photography of the painting, which is commonly used to study the painting’s texture. In raking light photography, a light source is placed at a shallow angle at the border of the painting. This results in a photograph (Fig. 7(a)) in which the ridges present in the painting are highlighted by their brightness; additionally, these ridges cause shadows to appear opposite to the location of the light source.

To identify the locations of ridges, we use the saturation channel of the raking light estimated map shown in Fig. 7(b). The height and estimate of a color is a metric of its light intensity and its distribution across the spectrum of different wavelengths/colors. In the figure, high values of saturation correspond to colors that are diffused in the spectrum, i.e. grayscale colors, while low values of saturation correspond to pure, bright colors. Thus, one can identify ridges that reflect the raking light by masking low saturation values and identify dark shadows by masking high saturation values. Figures 7(c) and 7(d) show masked images where the upper threshold and lower thresholds were set to 0.3 and 0.5 respectively. These ridge and shadow maps have been processed morphologically to remove spurious detections consisting of single pixels in the image.

After identifying ridges and shadows, we perform simple estimation of ridge height by exploiting the shadow information. We measure the length of the shadow in the direction of the raking light and label the corresponding ridge (which is directly adjacent to the shadow in the same direction) with the calculated shadow length, providing us with a ridge height estimate shown in Fig. 7(e).

We then aim to establish a functional approximation for the attenuation seen in the underpainting’s chemical element images. To perform this approximation, we merge information from the antimony layer image (see Fig. 6(a)) with information from the raking light photograph. Therefore, we perform basic registration of the raking light photograph in Fig. 7(a) with respect to the image in Fig. 6(a). Denote by $I_{\mathcal{E}}(x, y)$ the intensity of the antimony image at pixel $(x, y)$. We estimate the amount of attenuation for each pixel in the ridge map as

$$a(x, y) = \max_{x', y'} (x' - x)^2 + (y' - y)^2 \leq n, \text{ and } I_{\mathcal{E}}(x', y') = \mathcal{E}, I_{\text{surface}}(x', y') = \mathcal{C},$$

i.e., the difference between the intensity of the pixel in question and the maximum pixel intensity within a certain $n$-neighborhood. Once this attenuation has been estimated for every pixel in the ridge map, we average all attenuation values for each different height value to create a map between height estimates and attenuation estimates. Figure 7(f) shows the estimated map between ridge height and estimated attenuation, together with a best linear fit to the obtained mapping. The linear functional approximation was used to correct the attenuation, by scaling the attenuated values accordingly, with the results shown in Figure 6(d).

As Fig. 6(d) shows, the attenuation correction is not always accurate. While errors in registration are a possible cause, it is also evident that the correction for many ridges of paint is performed only on their lower halves. This type of error is not as strong for thin ridges corresponding to blades of grass; however, it is more noticeable for blob-like ridges that correspond to flowers. This is likely due to the fact that roughly half of the ridge (the upper half, which is located on the side opposite to the raking light) is not detected by the saturation-based approach. In future work, we intend to combine the two methods, using both attenuating surface hue identification and thickness estimation, and both inpainting and scaling of attenuated pixels, to produce improved results.
of its cobalt and mercury channel values, and the vector for a purple color. We then proceed via k-nearest-neighbour estimation algorithm [6]. We then estimate the mapping from the chemical element channels to colors. For simplicity, we begin by trying to find a mapping from the Hg and Co channels only since these red and blue pigments are the main chemicals that would be responsible for a purple color. We then proceed via k-nearest-neighbour estimation. First, for every pixel \( j \) in the waist ribbon, the vector \( \mathbf{x}_j \in \mathbb{R}^2 \) of its cobalt and mercury channel values, and the vector \( \mathbf{y}_j \in \mathbb{R}^3 \) of its RGB values in the surface painting are stored as training examples. Now, given a new vector \( \mathbf{x} \in \mathbb{R}^2 \) of the cobalt and mercury values for some new pixel, we estimate its RGB values \( \mathbf{y} \in \mathbb{R}^3 \) as

\[
y = \text{average}(\{\mathbf{y}_j | \mathbf{x}_j \text{ is one of the k nearest neighbors of } \mathbf{x}\})
\]

where “nearest” is defined as most similar in terms of Euclidean distance in \( \mathbb{R}^2 \).

Results of this procedure are shown in Figure 8. We see that a reasonable color appears and, like the waist ribbon, the hair ribbons show visually plausible evidence of shading with highlights appearing more blue and shadows appearing more red. Finally, we present a result on which we know ground truth: we predict the right half of the waist ribbon from the left half. We see that the estimated waist ribbon is similar to the true ribbon.

The more general problem of reconstructing colors in the absence of training examples remains to be solved. Still, special circumstances occasionally permit us to construct a reasonable guess of the portrait’s coloring. For example, to color the woman’s portrait under “Patch of Grass,” similar portraits of peasant women by Van Gogh were nonlinearly warped and registered to it. Their chromatic information was then fused with the underpainting’s luminance values (see [3]) to create the reconstructed underpainting shown in Stage 4 in Fig. 2.

6. CONCLUSIONS

This paper has provided an overview of the methods we have developed to address each of the four stages of underpainting recovery from X-ray fluorescence imaging data. While methods for some stages of recovery, such as acquisition artifact correction and inpainting of losses, already perform quite well, methods for other stages such as color reconstruction from the chemical element data are still in the early stages of development. We hope in future work to refine the methods for each individual stage further and to bring these methods for individual stages together into a unified framework for underpainting recovery from X-ray fluorescence data.

REFERENCES


