

Henry Chopp  | Robert R. McCormick School of Engineering and Applied Science, Northwestern University, Evanston, IL 60208 USA
| E-mail: henrychopp2017@u.northwestern.edu

Alicia McGeachy | Robert R. McCormick School of Engineering and Applied Science, Northwestern University, Evanston, IL 60208 USA
| E-mail: alicia-mcgeachy@northwestern.edu

Matthias Alfeld | Materials Science and Engineering, TU Delft, 2628 CD, Delft, The Netherlands | E-mail: m.alfeld@tudelft.nl

Oliver Cossairt | Robert R. McCormick School of Engineering and Applied Science, Northwestern University, Evanston, IL 60208 USA
| E-mail: ollie@eecs.northwestern.edu

Marc Walton | M+, Kowloon, Hong Kong
| E-mail: marc.walton@mpls.org.hk

Aggelos Katsaggelos  | Robert R. McCormick School of Engineering and Applied Science, Northwestern University, Evanston, IL 60208 USA
| E-mail: a-katsaggelos@northwestern.edu

Image Processing Perspectives of X-Ray Fluorescence Data in Cultural Heritage Sciences

Abstract—X-ray fluorescence (XRF) analysis of art objects has rapidly gained popularity since the late 2000s due to its increased accessibility to scientists. This introduced an imaging component whereby the XRF image volume provides clues as to which chemical elements are present and where they are located spatially in the object. However, as is the nature of collecting measurements, there are limitations preventing perfect acquisition; e.g., spatial resolution, signal-to-noise ratio, etc. The field of image processing, in part, aims to overcome these limitations. Image processing applications in XRF imaging are only just starting to arise due to the increased interest and availability in XRF analysis. In this article, we aim to reach readers in XRF imaging or image processing in an effort to call for further research in the field. We review the basics of XRF imaging and analysis that is tailored for those unfamiliar with this imaging modality. We then delve into various publications of image processing methods as applied to XRF data. Throughout this article, we examine (and opine on) the XRF field through a lens of the image processing field.

In 1887, Vincent van Gogh painted a patch of grass, aptly titled, *Patch of Grass*. This painting is shown on the left of Figure 1. Its surface appearance provides only a portion of the work and the artist's history; it is known that van Gogh by this time was transitioning from a darker paint palette to a lighter one. Underneath the visible surface, the painting embodies this transition. Little would van Gogh have known that over a century later, Dik et al. [1] would use x-ray fluorescence (XRF) imaging to analyze it (especially considering x-rays had yet to be discovered by Wilhelm Roentgen until 1895). Their colorized result, shown on the right of Figure 1, uncovered in detail a woman's face that van Gogh painted over. Known to reuse canvases, van Gogh composed the woman and the background with dark pigments prior to overpainting a more vibrant grassy scene. The intersection of XRF imaging and image processing then becomes apparent, as Anitha et al. [2], [3] have shown how to restore and colorize these underpaintings. XRF imaging is powerful in that it is capable of revealing the iterations of a painting before the final coat is applied.

In XRF imaging, a sample is scanned pixel-by-pixel with an x-ray pencil beam. This stimulates the atoms in the sample to emit characteristic fluorescence x-ray radiation. Elemental distribution images are then calculated from the data. While this article is focused on XRF imaging, a number of techniques (e.g., hyperspectral imaging) have been used in object analyses for tasks, such as pigment identification. We refer the reader to these review papers [4], [5], instead noting that XRF is one of several complementary techniques for identifying the elemental and chemical makeup of the object.

2692-4080 © 2022 IEEE

Digital Object Identifier 10.1109/MBITS.2022.3197100

Date of publication 8 August 2022; date of current version 8 December 2022.

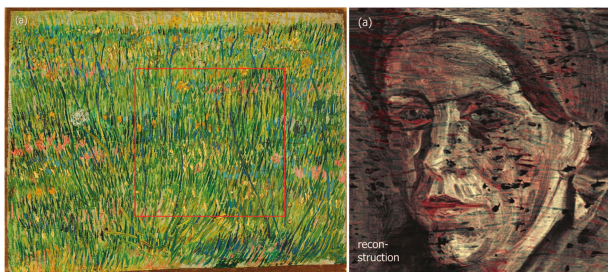


Figure 1

(Left) Vincent van Gogh, *Patch of Grass*, Paris, April–June 1887, oil on canvas, 30×40 cm, Kröller–Müller Museum, Otterlo, The Netherlands (KM 105.264; F583/JH1263). The red frame indicates the field of view of the right image (rotated 90° counter-clockwise). (Right) Approximate color reconstruction using chemical elements Sb and Hg from XRF data. Adapted with permission from Dik et al.'s work [1]. Copyright 2022 American Chemical Society.

The use of XRF in cultural heritage science is not new since acquiring XRF data is non-destructive—certainly an ideal for preservation. What differs is the use of imaging as opposed to spot analysis: prior studies examine select locations of interest to identify pigments and materials. These locations would be chemically understood, but are not necessarily representative of other areas similar in visual appearance.

With the advent of XRF imaging in the late 2000s, a flurry of research incorporated XRF imaging of paintings. Researchers published applications that expose other hidden paintings [6], [7], authenticate paintings [8], or aid in conservation efforts [9], for example. Researchers additionally developed various low-cost, mobile XRF imaging instruments [10], [11], [12], [13] that expanded access to XRF analysis (early experiments used synchrotron sources). Mobility is desirable as it minimizes the risk of damaging the painting by bringing the machinery to the painting, not vice versa.

As with any imaging instrument, however, there are limitations to its capabilities, e.g., signal-to-noise ratio, spatial resolution, and acquisition time. Image enhancement techniques prove to alleviate these concerns in a wide variety of signal domains. A plethora of denoising, super-resolution, and subsampling methods exist in the literature. Oftentimes these algorithms are domain-specific in order to incorporate prior knowledge of the signal. The solutions of these models depend on how the priors are incorporated. Everyone should be cognizant of how these biases affect the results, especially since no algorithm is perfect—which is why these fields are still active. Due to the relative newness of XRF imaging, there are understandably fewer dedicated publications of XRF image enhancement techniques. Perhaps too those spectroscopists instead of imaging scientists first developed the field, XRF volumes are conventionally viewed as a stack of

spectra rather than a stack of *images*. This is not to say that writing this article on XRF image processing techniques is unwarranted; rather, the limited number of existing algorithms surveyed in this article only adds purpose.

In this article, we want to provide an image processing perspective to XRF analysis. For example, many of the papers that address interpretability of XRF data use denoising methods, yet denoising is often nowhere mentioned in these papers. This is by no means a criticism of the works, but hopefully, via this article, we bring an additional viewpoint to the published work. We have two main goals for this article: we aim to reach readers from 1) the XRF community to provide a resource for enhancing XRF imaging and why it should be further researched, and 2) the image processing community to introduce XRF imaging, establish the current state of XRF image processing research, and emphasize the need for further developments in the area. In the discussion to follow, we hope that researchers of either discipline can identify aisles of opportunity for further development in XRF image processing and perhaps foster new interdisciplinary collaborations.

This article is structured as follows: we first introduce the science of XRF imaging and analysis. Second, we review and provide new insights into different areas of XRF image processing, namely denoising, super-resolution and inpainting, and subsampling. Finally, we opine the state of XRF image processing research as well as directions for further research.

XRF Imaging Overview

XRF imaging was introduced as an alternative to other techniques that image art objects beneath their surfaces. X-ray and infrared radiation-based imaging are common ways of viewing internal structures of paintings [14], although practically the entire electromagnetic spectrum has been used for these investigations [4]. These methods are employed to avoid extracting samples of the painting. What distinguishes XRF from other modalities is its ability to elucidate atomic elemental composition; this only further reduces the need to remove paint samples for chemical analysis. We will discuss how the underlying science of XRF imaging is used to identify pigments and materials throughout the layers of paints. For a more in-depth yet gentle introduction to XRF spectrometry than provided here, we refer the reader to Brouwer's work [15], which is tailored toward those new to the field.

Physics of XRF Spectroscopy

To collect XRF data, a source illuminates an object with a continuous spectrum of x-rays. These x-rays are collimated on a small spot. As the sample is exposed to x-rays, some of the x-rays are absorbed by the electrons in the sample. Impacted electrons may be dislodged from the atom if the energy of the incoming x-rays is larger than that of the binding energies. Losing electrons creates energetically unfavorable vacancies

in the atom's electron configuration. To stabilize, electrons in outer orbitals move inward to fill the vacancy.

During this transition, energy is conserved. Outer orbitals have higher energies than inner orbitals, so the transitioning electron loses energy. This loss is realized as a photon emitted by the atom. The photon's energy equals the energy loss.

Not all photon energies are possible, however. Within any type of atom, there are different electron orbitals/shells at unique and distinct energy levels. The lowest energy orbital is the K-shell, which holds two electrons. The L-shell, subdivided into three subshells, has the next three lowest energy levels for eight additional electrons. The M-shell has five subshells, all with greater energy than those of the L-shell; it can hold eighteen electrons. These shells constitute the main transition lines whereby electrons fill the vacancies: electrons typically move from in the following:

- 1) L- to K-shell;
- 2) M- to K-shell; and
- 3) M- to L-shell.

Not all transitions are possible, and some transitions are more likely to occur than others.

The fluorescence photons are emitted in an isotropic manner, and those not absorbed on their path to the detector are recorded by it. In the semiconductor detector, the photons create a charge that enhances the conductivity. This is transferred into the energy dispersive spectrum used in XRF analysis.

The sources of the incoming x-rays are not only from the electrons in atoms of the top surface layer. X-rays have high energy that can penetrate below the surface layer and interact with hidden atoms. Photons from these atoms must pass back through intermediate layers and into the detector to be recorded. These photons are less frequent than those of the same element that lies on the top level, but are still present in large quantities. There are established limits on the penetration depth that depend on a multitude of factors, but paintings are often thin enough to record photons throughout all the layers.

XRF Measurement Challenges

There are some challenges that arise in XRF spectrometry that introduce noise or artifacts into the recorded spectrum. Here, we highlight some problems that are characteristic of XRF spectroscopy.

X-Ray Source

X-ray tubes emit incoherent, polychromatic x-rays by accelerating electrons from a filament toward an anode. Upon

contact, the electrons decelerate, and x-rays are generated. A sizeable portion of these x-rays is inadvertently reflected back into the detector, which records a broad spectrum of signals. This is called the *continuum*, which can be estimated and subsequently subtracted from the spectrum. Some collisions even result in electron vacancies in the anode itself. Photons characteristic of the anode's material are then emitted from the source, which can create a false (or amplified) peak in the XRF spectrum due to this *backscattering*.

Interactions Between X-Rays and the Object

Aside from the source, x-ray interactions with electrons of the object of interest do not always involve electron ejection or x-ray reflection. *Rayleigh scattering* may occur when electrons hit by x-rays instead vibrate at the same frequency as the incident photons. The vibrations cause photons of the same frequency to be released, which contributes toward the continuum.

Compton scattering occurs when the incoming x-ray is backscattered, but loses some of its energy. This scattering phenomenon is more apparent in low- Z elements (i.e., elements of low proton count), but can disappear in high- Z elements.

Detector

One last major origin of error occurs at the detector. Escape peaks occur when incoming photons excite the detector itself. The XRF photon is not reabsorbed but rather escapes the detector. The photon then loses some of its energy before being converted to a voltage, but the energy loss is well-documented based on the material of the detector.

Pileup peaks can also be produced where two photons are incident on the detector in a small time window. This creates seemingly large energy equal to the sum of the individual photon contributions that the postprocessor cannot resolve as two distinct photons. Pileup is enhanced by high count rates.

These sources of error cannot be controlled, but they can be mitigated in XRF analysis. What can be (roughly) controlled is the number of photons recorded by adjusting the scan time. Photons arrive according to a Poisson process with some unknown underlying rate. Since XRF spectrometry is a photon counting measurement, photon peaks can only be detected if there are enough arrivals to distinguish them from both noise and the continuum. The longer the scan time, the more apparent the peaks will be. XRF imaging presents a challenge in that the dwell times per spot cannot be too small such that peaks are lost in the noise, and cannot be too long such that it takes an excessive amount of time to collect the volume.

Collection times are long since XRF systems are single-pixel scanners. In order to generate an XRF image volume, the x-ray source and detector are mounted on a gantry that moves in a (typically) raster scanning motion to cover the desired area. The acquisition process can be lengthy depending on the dwell time and spatial resolution. For example, the XRF volume in Figure 1 reportedly took two days to collect a 17.5×17.5 cm area [1], although being an early paper, care was taken to get good statistics. While XRF systems have since improved, the scan times are still generally on the order of hours or days. The same area that was scanned in *Patch of Grass* can nowadays be scanned in approximately an hour.

Notation

Before understanding how XRF analysis is done, we need to establish a notation. Throughout this article, we use the following rules:

- 1) lowercase lettering denotes scalars;
- 2) uppercase lettering denotes matrices and vectors; and
- 3) boldface uppercase letters denote 3-D tensors.

Let $\mathbf{X} \in \mathbb{Z}_+^{C \times H \times W}$ be the collected XRF data where $\mathbb{Z}_+$ is the set of nonnegative integers. The volume has height H , width W , and channels C . Each channel corresponds to an energy level where the incoming photons are binned. Each entry $\mathbf{X}_{c,h,w}$ contains the number of recorded photons at pixel (h, w) with energy c .

As will become clear, many XRF analysis techniques revolve around dictionary learning or other matrix factorization methods. Thus, we introduce here some additional terms: $D \in \mathbb{R}_+^{C \times M}$ is the dictionary composed of M different spectra, and $\mathbf{A} \in \mathbb{R}_+^{M \times H \times W}$ denotes the abundances of each of the M spectra.

In dictionary learning, D and $\mathbf{A}$ are found such that

$$\mathbf{X} \approx D\mathbf{A} \quad (1)$$

where the matrix and tensor multiplication is carried out via

$$\mathbf{X}_{c,h,w} = \sum_{m=1}^M D_{c,m} \cdot \mathbf{A}_{m,h,w}. \quad (2)$$

This is the basis for many of the techniques surveyed here.

XRF Analysis of Individual Response Lines

The core of XRF analysis is unmixing the resultant spectra: which atomic elements are present, and of the present elements, how much is present? All collected spectra are essentially linear combinations of the spectra of

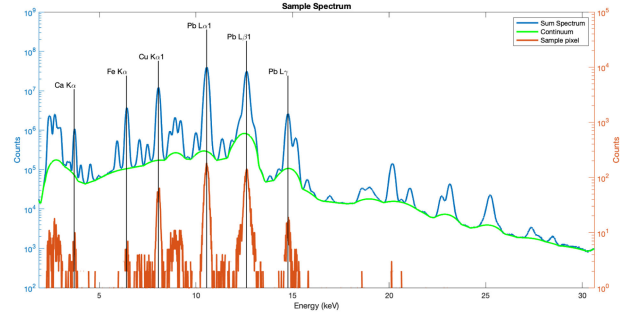


Figure 2

Sample XRF spectra with select peaks labeled. Note the different y-scales for the spectra. (Blue) The sum spectrum. (Green) The continuum. (Red) A sample pixel.

individual elements (along with the continuum, noise, backscattering, etc.). Even within the elemental spectra, they too can further be decomposed into the different emission lines.

The first step in XRF analysis is identifying the peaks that indicate the presence of a certain element. This is done manually or automatically by examining the sum spectrum, i.e., the spatial sum of all the spectra

$$S_c = \sum_{(h,w)} \mathbf{X}_{c,h,w}. \quad (3)$$

This provides the least noisy presentation of which elements lie in the object. Since it is very likely the same elements/compounds exist throughout the painting spatially, the sum of many measurements reduces the noise. Any elements identified are included in a dictionary composed of the elemental responses. Figure 2 shows a sample sum spectrum as well as a single pixel's spectrum. The peaks are easy to notice in the sum spectrum, but are more difficult to identify when analyzing the single pixel—some maxima may be due to noise in the individual pixel. Once the peaks are identified, a table can be used to attribute the peak's energy to an atomic element.

Of the identified elements, the next step is to decompose the XRF signal at each pixel according to the dictionary. Each peak is often modeled as a Gaussian, and each element consists of one or more peaks. These XRF response curves for elements present in the sample are the columns of the dictionary D . The continuum response is oftentimes included in the dictionary. A non-negative least-squares approach is then used to fit the XRF signal at each pixel individually

$$\mathbf{A}^* = \arg \min_{\mathbf{A} \geq 0} \mathcal{L}(\mathbf{X}, D, \mathbf{A}) \quad (4)$$

for some objective (loss) function $\mathcal{L}$, typically the L2 norm. Optimal abundance $\mathbf{A}^*$ conveys the relative amount of each element present. One can then visualize the individual channels of the abundance matrix to see how much of each element are present across the painting—these are the *elemental maps*.

PyMca [16] is a commonly used platform that can carry out the tasks previously in a streamlined approach. It also takes into account other modeling factors beyond the scope of this article.

XRF Analysis Using Different Bases

Since XRF data analysis is fundamentally an unmixing problem, most techniques use some form of dictionary learning and matrix factorization to analyze the data. Whereas the elemental decomposition formulation of (4) only solves for the abundance matrix $\mathbf{A}$, another formulation solves for the dictionary as well

$$D^*, \mathbf{A}^* = \arg \min_{D, \mathbf{A} \geq 0} \mathcal{L}(\mathbf{X}, D, \mathbf{A}). \quad (5)$$

This allows for a more complex representation of the data that can aid in interpreting the data.

In particular, interpretability has posed an issue in terms of identifying trace elements. Trace elements have short peaks that may be lost in the total sum spectrum. *A priori* knowledge is sometimes needed to identify the trace elements in the sum spectrum, and it can be just as difficult to predict the abundance of the trace elements in the individual spectra [17].

In addition to finding trace elements, XRF analysis seeks out correlations in the data to identify not just the individual chemical *elements*, but rather the chemical *compounds* and where they appear spatially. These compounds can be difficult to identify using the individual elemental peaks alone. For example, say an element is present in multiple distinct compounds. It is a challenging task to separate out how much of each compound (as well as which compounds) may be present.

Many classical and some newer methods in data processing are used to analyze the chemical composition of paintings. These analysis techniques can also be used to denoise data in the spectral domain; we will explore this in the next section.

XRF Interpretation Methods as Spectral Denoising Mechanisms

The first attempts to better interpret XRF image data to revolve around the fact that pure elements typically do not exist on their own in paintings, but rather as a compound.

Vermillion, for example, is a red paint that was previously made from an Hg and S compound. In XRF analysis, characteristic peaks of Hg and S would appear should vermilion exist. This perhaps allows for dimensionality reduction that identifies pigments and mixtures of different paints that the artist used.

In XRF imaging, changing the composition of the dictionary D from individual elemental responses to a new basis is how many published analyses are carried out. As will become clear, these methods can often be considered denoising algorithms that are able to smooth the original data.

Data denoising is often overlooked in XRF analysis. Notice the bottom plot of Figure 2 again. The signal is quite noisy compared to the sum spectrum, as the signal is inherently discretized and most peaks occur under 20 photon counts. There are two primary ways to mitigate the effects of noise on the individual spectra of a pixel: 1) increase the dwell time for each pixel, and/or 2) use image processing techniques to denoise the data. The former option is typically not available since experimentalists already set the dwell time to the longest reasonable length. Even the slightest addition of dwell time can have immediate impacts on the total scan time. For example, if the scan area is 500×600 px, each additional millisecond of dwell time per pixel adds five minutes of overall scan time. Instead of increasing the dwell time as a way to denoise the data, denoising techniques in image processing can be applied.

In this section, we first address early methods that use principal component analysis (PCA) or clustering approaches. Then, we go into more popular methods used today to perform XRF analysis. Throughout this section, we note how these techniques can be repurposed as potential denoising methods alongside their use as an analysis method in XRF literature.

PCA and Clustering Methods

Much of the early work in XRF interpretation uses PCA to generate a new XRF basis [18]. PCA is known to be an easy but effective way to extract correlations in data as the components (dictionary members) are the eigenvectors of the covariance matrix. The eigenvectors corresponding to the largest eigenvalues are chosen as they best capture the direction of the highest variance in the data. Oftentimes only a few components are needed to capture a significant amount of the variation. Each pixel is then approximated by a linear combination of the PCA vectors that compose D . The abundances $\mathbf{A}$ are found via (4) sans the non-negativity constraint. XRF interpretation here identifies peaks that correlate with one another.

What is perhaps missed in the XRF analysis is that PCA is also effective at signal denoising [19]. Many denoising algorithms

have been developed around PCA to remove noise in images along the spectral and spatial components that could be of use in XRF analysis. The application of PCA here would be no exception, although perhaps rudimentary by today's image processing standards in terms of complexity.

Later, Vekemans et al. [20] used a combination of PCA and K-means clustering to automatically extract correlations. The goal here was to identify distinct regions of similar XRF response. By doing so, the authors are able to find sum spectra of similar regions. This aided them in the identification of trace elements that may be lost in the overall sum spectrum.

To find these regions, the authors first perform PCA over the channels of the XRF data. This finds the eigenvectors (or equivalently eigenimages) of the covariance matrix. Then, they perform K-means clustering of the pixels of the first n principal eigenimages to find clusters that represent a combination of elemental XRF responses. This method helped them approach the goal of automatic segmentation: by segmenting the images, the sum spectra of each cluster can be used to better detect trace elements. This was one of the earliest works that used a combination of spectral and spatial methods for XRF analysis.

Clustering-based methods have also been shown as a way to denoise the data as well [21]. By establishing representative cluster centroids, the data are mapped into a lower dimension where the noise is mitigated. In this case, spectral denoising is done through K-means clustering while spatial denoising is done via PCA.

Image processing is rich in PCA and clustering methods for denoising, although more common now is the use of neural networks, which are often not an option in XRF analysis due to the lack of publicly available data. That being said, there is a plethora of image and spectral denoising algorithms in signal processing literature more advanced than PCA and K-means clustering. Some of these algorithms may be suitable for XRF analysis perhaps with some changes that incorporate prior XRF domain knowledge.

Nonnegative Matrix Factorization

While PCA-based techniques were popular, the results are not always interpretable since PCA decomposition can contain negative values. What exactly would a negative amount of a spectrum indicate? All raw XRF data consists of nonnegative photon counts, so it is not at all clear how to interpret PCA decomposition in a physical sense. This realization led to the additional constraint that both the dictionary and abundances must be nonnegative to provide feasible results.

Instead of PCA-based interpretation methods, Alfeld et al. [22] proposed using nonnegative matrix factorization (NMF) as a way to analyze the XRF correlations along the spectral

dimension. NMF is a problem that directly addresses the constraints of (5) unlike PCA. Once the basis (dictionary) and abundance matrix are found, the results are more readily interpretable as all the entries are nonnegative. The downside is that the cost function is typically nonconvex. An iterative algorithm converges to some local minimum that depends on the initial values of D and A . Different initializations should be tested to find the best local minimum.

NMF is often solved using an algorithm called fast nonnegative least squares [23]. The method iteratively solves the following two equations until convergence:

$$A^{k+1} \leftarrow \arg \min_{A^{k \geq 0}} \|X - D^k A^k\|_F^2 \quad (6)$$

$$D^{k+1} \leftarrow \arg \min_{D^{k \geq 0}} \|X - D^k A^{k+1}\|_F^2 \quad (7)$$

where the superscript $k \geq 0$ denotes the iteration number. The entire, simple algorithm for solving (6) and (7) individually is provided in the reference.

It was found in Alfeld et al.'s work [22] that NMF provides more interpretable results. There was, however, some difficulty in distinguishing between different pigment groups that contain some of the same elements. This was mitigated by incorporating the additional information that some of the compounds were known *a priori*. NMF here was only done on some of the elemental maps instead of the XRF volume as a whole. Santos et al. [24] also used NMF, but on the entire XRF volume.

Again, just as PCA can smooth the data, so too can NMF. Perhaps one of the most noticeable differences is the nonnegativity constraint imposed on the NMF problem that introduces sparsity to some extent. There is a possibility that in the final spectral decomposition, D and A contain entries of 0. Sparse representations of A are known to be able to effectively denoise when redundancies exist in the data [25]. Intuitively, these zero-valued elements indicate the nonexistence of some XRF response. The absence of certain responses can be just as beneficial in XRF analysis as a response that is present [26].

Other Factorization Methods

NMF provides a nonnegative factorization of the XRF data, but in its native form does not require any other constraints to be imposed. Take, for example, the issue of many solutions for the decomposition. This so-called rotational ambiguity is easily illustrated by choosing some non-singular matrix $Q \in \mathbb{R}^{M \times M}$ [27]. The XRF decomposition of (1) can be rewritten as

$$X \approx (DQ)(Q^{-1}A). \quad (8)$$

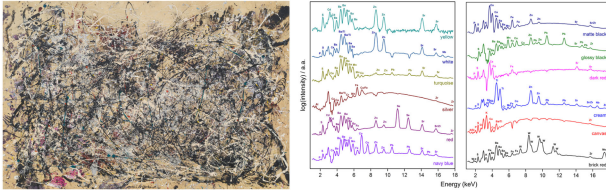


Figure 3

(Left) Pollock Number 1A, 1948 (1948). Oil and household enamel paint on canvas (172.7 × 264.2 cm) The Museum of Modern Art. (Right) Plots of the dictionary after applying MCR-ALS. Adapted from Martins et al.'s work [28].

Any invertible matrix Q can be chosen so long as the entries of DQ and $Q^{-1}\mathbf{A}$ are all nonnegative. Additional constraints can be added in order to further constrain the solution.

Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS)

One constrained NMF algorithm that is employed is called MCR-ALS [27]. While this may not be familiar to those in the image processing community, this algorithm is in essence a NMF framework where certain additional constraints can be added. A popular constraint to include is that the final dictionary is a combination of the individual elemental spectra. As a least squares minimization, we have that

$$B^*, \mathbf{A}^* = \arg \min_{B, \mathbf{A} \geq 0} \|\mathbf{X} - (DB)\mathbf{A}\|_F^2 \quad (9)$$

where $B \in \mathbb{R}_+^{M \times N}$ is the mixing matrix describing the linear combination of the individual elemental spectra D . Note that B need not be a square matrix; in fact, it is the case that $N < M$ to reduce the dimensionality of the dictionary and establish XRF correlations. The new dictionary is $DB \in \mathbb{R}^{C \times N}$. The dimensions of $\mathbf{A}$ are also modified to $\mathbf{A} \in \mathbb{R}_+^{N \times H \times W}$ in order to have a valid matrix multiplication operation.

With B known, dictionary DB is quite easily interpreted as simply a linear combination of the original dictionary. Martins et al. [28] analyzed a painting by Jackson Pollock in this manner, as seen in Figure 3. Their analysis shows that the dominant colors can be identified based on the peaks of each spectrum in the dictionary. It however does not encompass a complete separation of the compounds present in the painting, which would be ideal considering the many overlapping paints that are characteristic of Pollock's work. Still, with only twelve dictionary endmembers, MCR-ALS is able to capture correlations that are present in the XRF spectra. The authors report that their criterion for selecting N is in part based on whether at least 95% of the variance can be explained.

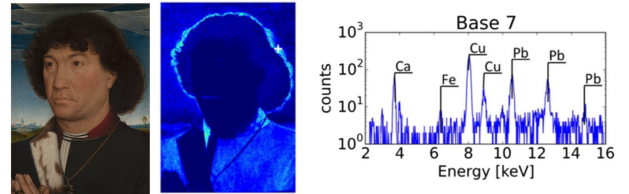


Figure 4

(Left) Portrait of a Man from the Lespinette Family, Hans Memling, 1485–1490, Mauritshuis, The Hague, Oil on panel, 30.1 × 22.3 cm². (Middle) Abundance map of base 7. The base pixel is denoted by a yellow plus in the map. (Right) The spectrum of base 7. Adapted from [29] with permission from Elsevier.

We note that along with the other aforementioned techniques, MCR-ALS can be used to denoise the original data in addition to the XRF interpretation. Representing all pixels as a linear combination of 12 spectra does not allow for much of the noise to be reconstructed without incurring more error in the overall cost of the objective function of (9).

Simplex Volume Maximization (SiVM)

Aside from MCR-ALS, SiVM has also been proposed [29] to overcome computational speed issues of NMF and MCR-ALS. SiVM is also unique in that the dictionary is composed of data points in the XRF volume $\mathbf{X}$. These points are chosen from the set of points that lie on the convex hull of $\mathbf{X}$. Oftentimes the number of vertices of the hull is more than the desired number of endmembers in the dictionary. SiVM chooses the M extreme points such that the volume of the resulting simplex is maximized. These extreme points compose the dictionary, making XRF interpretation perhaps even easier than that of MCR-ALS since the basis endmembers can be readily found in the original data.

The authors reported some overlap in the elements present in the dictionary, but still gleaned new information that can be seen in Figure 4. In particular, they noticed a glow around the hair in one of the bases that was not present in the individual elemental maps that highlights the intersection of paint that compose the hair (Ca) and the sky (Cu and Pb).

This method illustrates another way to select the dictionary, which has been shown to be a possible denoising algorithm [30]. The authors of this article similarly note that hyperspectral unmixing is a form of denoising the data.

Dictionary Denoising of Poisson Data

A final method involves the combination of Poisson noise modeling and dictionary learning explicitly for denoising the

XRF volume. Chopp et al. [31] proposed to use the Poisson negative log-likelihood loss (PNLL) when solving for the dictionary, particularly when the XRF pixels have relatively low photon counts. This provides a better model for the noise instead of the L2 norm, which assumes the data are drawn from a Gaussian model.

The PNLL, derived from the Poisson maximum likelihood estimation, is given by

$$\text{PNLL}(\mathbf{X}, D, \mathbf{A}) = \sum_{i \in (c,h,w)} (D\mathbf{A})_i - \mathbf{X}_i \cdot \ln (D\mathbf{A})_i. \quad (10)$$

When the PNLL is used as the objective function in a minimization problem, it minimizes the PNLL estimation (or equivalently maximizes the Poisson likelihood estimation). The objective function uses the PNLL instead of the L2 norm and is regularized using an adaptive total variation (TV) regularizer as well as an L0 penalty term on the abundance tensor to introduce sparsity, which is known to be an effective modeling strategy in smoothing data with redundancies [25].

They provide experiments comparing different denoising algorithms of fast XRF raster scans where the ground truth is known. The algorithm outperformed MCR-ALS as a denoising method, both in terms of the error in the denoised XRF volume and the denoised elemental maps.

Perspectives

The dictionary-based XRF analysis algorithms reviewed here can be thought of as a method to denoise the data, even when this is not explicitly mentioned or studied in the papers present. Clearly, there is room for more advanced algorithms to arise for both XRF analysis as well as denoising. Perhaps by applying other denoising algorithms in image processing literature (or developing a method on one's own), better XRF analysis can be accomplished. The converse could also hold true whereby developing a new XRF analysis technique can decrease the noise present in the original data volume.

We encourage those that have applied dictionary decomposition methods in XRF imaging to revisit their techniques from a denoising perspective. Using the smoothed $D\mathbf{A}$ volume in place of $\mathbf{X}$, the original elemental maps can be found. Noisier maps with low count rates may appear smoother than before, and it would be an interesting study to see how well each of these methods performs as a denoiser.

We would also like to note that many of the algorithms for XRF interpretation only use spectral denoising techniques. It is well known in image processing literature that pixels in a local area generally have low variance in their values. This knowledge could be applied in the *spatial* domain in the form of a TV regularizer

$$\text{TV}(\mathbf{A}) = \text{TV}_H(\mathbf{A}) + \text{TV}_W(\mathbf{A}) \quad (11)$$

$$\text{TV}_H(\mathbf{A}) = \sum_{c=1}^C \sum_{h=1}^{H-1} \sum_{w=1}^W (\mathbf{A}_{c,h+1,w} - \mathbf{A}_{c,h,w})^2 \quad (12)$$

$$\text{TV}_W(\mathbf{A}) = \sum_{c=1}^C \sum_{h=1}^H \sum_{w=1}^{W-1} (\mathbf{A}_{c,h,w+1} - \mathbf{A}_{c,h,w})^2. \quad (13)$$

This regularizer penalizes large changes in neighboring pixels. Incorporating this term into the objective function, we have

$$D^*, \mathbf{A}^* = \arg \min_{D, \mathbf{A} \geq 0} \mathcal{L}(\mathbf{X}, D, \mathbf{A}) + \lambda_{\text{TV}} \text{TV}(\mathbf{A}) \quad (14)$$

for some scalar $\lambda_{\text{TV}} > 0$. Solving minimization problems of this form could improve how XRF signals are analyzed with the added prior knowledge, particularly in overcoming peak detection limits with weak XRF signals, as seen in Chopp et al.'s work [31]. Using neighboring information could provide a better decomposition of the pixels individually.

Take the Memling painting in Figure 4 which has many spatially smooth regions. It is reasonable to predict that most neighboring signals do not have wildly varying spectra. Adding TV regularization could provide a different optimized dictionary that takes these spatial relations into account. Of course, the TV term is not set in stone, and could be adapted.

An adaptation is likely needed in the Pollock painting of Figure 3, for example, since there is a large amount of high spatial frequency. There are many different thin paint lines that intersect each other, so it is reasonable to assume that neighboring spectra vary quite a bit. The TV regularizer in (11) may inadvertently be detrimental to the results if the spatial resolution of the XRF image is low. Having an adaptive TV regularizer that identifies similar regions could be a way to modify the standard TV equation.

Spatial Deblurring, Super-Resolution, and Inpainting Methods for XRF Volumes

Although XRF interpretation through known spectral denoising techniques is quite common in XRF literature, we have not yet focused on the spatial-based applications of XRF image processing. There have been some published works on deblurring, super-resolution, and inpainting techniques applied to XRF data. The deblurring problem focuses on restoring high-frequency components of the images lost in the data acquisition process. This is mainly an issue in micro-scale XRF imaging (μ -XRF) where the spot size is more of a factor than in the macroscale (MA-XRF) with art objects.

In super-resolution and inpainting problems, some pixels are excluded or missing from measurements typically to reduce

XRF acquisition time. Following hard acquisition time constraints may be a factor that affects the spatial resolution. Instead of capturing the XRF data at the resolution that one desires, time can be saved by decreasing the resolution or sampling select areas.

In this section, we will review the different image processing techniques that have been applied/developed for improving XRF imaging in the spatial domain.

Fourier-Based Deblurring Approaches

General algorithms for image deblurring already exist and can be readily applied to XRF data. Yang et al. [32] conducted an experiment on some elemental maps that compare four different Fourier transform-based super-resolution methods:

- 1) Wiener deconvolution;
- 2) Richardson–Lucy [33];
- 3) fast iterative shrinkage-thresholding algorithm [34]; and
- 4) blind deblurring with L0-regularized intensity and gradient prior (LORIGP) [35].

The first three algorithms are nonblind methods, i.e., the point spread function (PSF) is known. LORIGP is a blind method where the PSF is not known. Since the authors reported that the Richardson–Lucy algorithm performed the best for their μ -XRF datasets, we only review this method and refer the reader to Yang et al.'s work [32].

Let $Z, L \in \mathbb{R}_+^{H \times W}$ be the underlying super-resolved elemental map and the acquired elemental map respectively. With a PSF P , Z can be approximated by

$$Z \approx L * P \quad (15)$$

where $*$ denotes convolution. Richardson–Lucy is an algorithm that can notably be used to deblur images with Poisson noise. The algorithm is iterative and converges at the maximum likelihood solution given the PSF

$$Z^{k+1} = Z^k \cdot \left(\hat{P} * \frac{L}{P * Z^k} \right) \quad (16)$$

where $\hat{P}$ is the “flipped” version of P whereby the rows and columns are reversed. Elementwise multiplication is denoted by $\cdot$, and division is carried out elementwise as well.

The authors tested this method on the Fe–K α , Ca–K α , and Si–K α elemental maps of an iron skarn, which is known to have sharp boundaries that are difficult to capture with μ -XRF imaging. These results are shown in Figure 5. While the Fe–K α and Ca–K α maps have sharper boundaries as they expected, the Si–K α maps show more degradations than

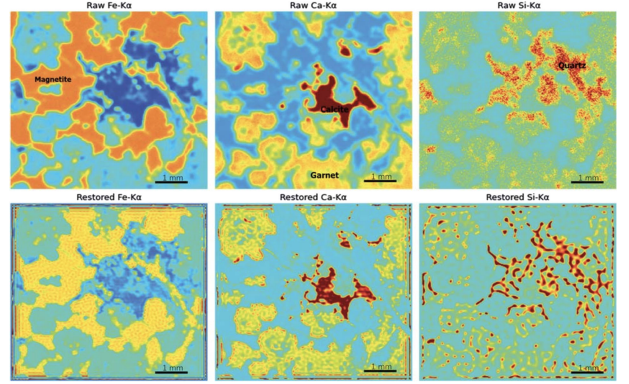


Figure 5

Elemental maps of an iron skarn. (Top row) Raw maps of Fe–K α , Ca–K α , and Si–K α , respectively. (Bottom row) Denoised maps of Fe–K α , Ca–K α , and Si–K α , respectively. Reproduced from Yang et al.'s work [32].

there should be. They attribute this to the high noise in the Si–K α image due to the chemical nature of the quartz (which is composed of Si) as it interacts with x-rays. The authors conclude that the Richardson–Lucy algorithm can be applied to elemental maps that have high XRF responses, but is not as effective with elements that are harder to detect.

Super-Resolution and Inpainting Approaches

The goal of super-resolution is to predict the values of unmeasured pixels. Here, the object is sampled along a uniform rectangular grid, which composes the low-resolution image. This is in contrast to generalized inpainting methods, which are not classically constrained to be sampled in a uniform fashion. Whether due to time constraints or x-ray spot size, the resolution needs to be increased for improved spatial analysis.

Dai et al. [36] proposed a sparse dictionary-based method for super-resolution of the XRF volume, $\mathbf{X} \in \mathbb{R}_+^{C \times H_l \times W_l}$, using information from a high-resolution color image of a painting, $\mathbf{I} \in [0, 1]^{3 \times H_s \times W_s}$. Let $H_l < H_s$ and $W_l < W_h$ to conform with the super-resolution problem. They propose to first separate the low-resolution XRF data $\mathbf{X}$ and the super-resolved XRF image $\mathbf{Y} \in \mathbb{R}^{C \times H_s \times W_s}$ into two components: a visible $(\cdot)_v$ and nonvisible component $(\cdot)_{nv}$ such that

$$\mathbf{X} = \mathbf{X}_v + \mathbf{X}_{nv} \quad (17)$$

$$\mathbf{Y} = \mathbf{Y}_v + \mathbf{Y}_{nv}. \quad (18)$$

The visible component is defined as the portion of the XRF signal attributed to the surface response of the painting. The non-visible component contains the XRF signal that originated from underneath the painting's top layer of paint.

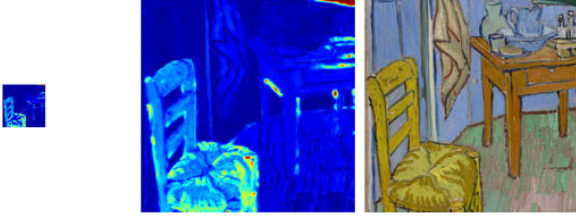


Figure 6

(Left) Low-resolution XRF image. (Middle) The super-resolved XRF image. (Right) The high-resolution RGB image: The Bedroom, Vincent van Gogh, 1853–1890, Arles, October 1888, oil on canvas, $72.4 \times 91.3 \text{ cm}^2$. Adapted from Dai et al.'s work [36].

The purpose of separating the XRF signal is to establish a relationship with the RGB signal, which only images the top layer of paint. They propose using a dictionary decomposition of $\mathbf{Y}_v$, $\mathbf{Y}_{nv}$, and $\mathbf{I}$

$$\mathbf{Y}_v = D_v^{\text{xrf}} \mathbf{A}_v \quad (19)$$

$$\mathbf{Y}_{nv} = D_{nv}^{\text{xrf}} \mathbf{A}_{nv} \quad (20)$$

$$\mathbf{I} = D^{\text{rgb}} \mathbf{A}_v. \quad (21)$$

Notice the abundance $\mathbf{A}_v$ is shared across the XRF and RGB domains. The amount of contributions from the XRF and RGB domains at each pixel is proposed to be equal.

Furthermore, the low- and high-resolution XRF volumes are related by a binary sampling matrix $\mathbf{S} \in \{0, 1\}^{1 \times H_s \times W_s}$ by

$$\mathbf{X}_v = \mathbf{Y}_v \cdot \mathbf{S} = (D_v^{\text{xrf}} \mathbf{A}_v) \cdot \mathbf{S} \quad (22)$$

$$\mathbf{X}_{nv} = \mathbf{Y}_{nv} \cdot \mathbf{S} = (D_{nv}^{\text{xrf}} \mathbf{A}_{nv}) \cdot \mathbf{S} \quad (23)$$

where $\cdot$ is an elementwise multiplication where $\mathbf{S}$ is the multiplier for each channel of the multiplicand.

The optimization problem is then

$$\arg \min_{\Theta \geq 0, \|\mathbf{A}\|_0 \leq s} \|\mathbf{X} - (D_v^{\text{xrf}} \mathbf{A}_v) \cdot \mathbf{S} - (D_{nv}^{\text{xrf}} \mathbf{A}_{nv}) \cdot \mathbf{S}\|_F^2 + \|\mathbf{I} - D^{\text{rgb}} \mathbf{A}_v\|_F^2 + \lambda_{\text{TV}} \text{TV}(D_{nv}^{\text{xrf}} \mathbf{A}_{nv}) \quad (24)$$

where $\Theta = \{D_v^{\text{xrf}}, D_{nv}^{\text{xrf}}, D^{\text{rgb}}, \mathbf{A}_v, \mathbf{A}_{nv}\}$, $\mathbf{A} = \mathbf{A}_v + \mathbf{A}_{nv}$, and $\|\cdot\|_0$ is the L0 pseudonorm. This equation is constrained to have a sparse representation, which has been shown to have smoothing effects when there are redundancies in the data [25]. The first two terms of (24) are fidelity terms, and the third is the TV regularizer from (11) that captures spatial correlations. We refer the reader to the original paper to learn how to minimize this complicated objective function.

Their solution provided better super-resolution results than the other methods designed for hyperspectral images. Figure 6 shows a super-resolved XRF image of *The Bedroom*

by Vincent van Gogh. They were able to capture the XRF response of the curtain in a more accurate manner than the other methods that either filter the curtain out or add artifacts to the map; they report a peak signal-to-noise ratio (PSNR) of 56.19 dB.

As was mentioned earlier, inpainting is closely related to the super-resolution problem since both problems estimate unknown pixel values. Dai et al. [37] pushed their optimization algorithm to be applied to any sampling matrix. The sampling matrix will be discussed in the next section, but the updated optimization algorithm introduces a new penalty term. They postulate that the gradient of the visible component of the XRF volume should be similar to the gradient of the RGB image. This is formulated as a weighted TV regularizer $\text{TV}(D_v^{\text{xrf}} \mathbf{A}_v, \mathbf{I})$ in (24). The new penalty term weighs the TV losses of (12) and (13) based on the RGB image gradient. In low varying areas of the RGB image, they expect the visible XRF component is low varying as well. Thus, a relatively higher penalty in the smoothing is applied on pixels with like neighbors. On the other hand, along the edges of the RGB image, there is high contrast; relatively low penalties are applied here since the high variation is expected that should not be smoothed. We refer the reader to the work [37] for the full optimization and a definition of the adaptive TV regularizer.

These methods provide a good RGB-XRF fusion method for super-resolution and inpainting, although the algorithm was recently improved upon by Su et al. [38]. They make a slight change in the framework from Dai et al. [36]: instead of separating the XRF volume into visible and nonvisible components, they propose to separate the volume by common and unique components. This slight change in thinking introduces a new decomposition scheme

$$\mathbf{X} = \mathbf{X}_c + \mathbf{X}_u = (D_c^{\text{xrf}} \mathbf{A}_c) \cdot \mathbf{S} + (D_u^{\text{xrf}} \mathbf{A}_u^{\text{xrf}}) \cdot \mathbf{S} \quad (25)$$

$$\mathbf{I} = \mathbf{I}_c + \mathbf{I}_u = D_c^{\text{rgb}} \mathbf{A}_c + D_u^{\text{rgb}} \mathbf{A}_u^{\text{rgb}} \quad (26)$$

where $(\cdot)_c$ denotes the common components and $(\cdot)_u$ denotes the unique components. There are also new and updated terms in their objective function

$$\arg \min_{\Theta \geq 0} \|\mathbf{X} - D_c^{\text{xrf}} \mathbf{A}_c \cdot \mathbf{S} - D_u^{\text{xrf}} \mathbf{A}_u^{\text{xrf}} \cdot \mathbf{S}\|_F^2 + \|\mathbf{I} - D_c^{\text{rgb}} \mathbf{A}_c - D_u^{\text{rgb}} \mathbf{A}_u^{\text{rgb}}\|_F^2 + \lambda_{\text{TV}} \text{TV}(D_u^{\text{xrf}} \mathbf{A}_u^{\text{xrf}}) + \lambda_{\text{MI}} \text{MI}(\mathbf{I} \cdot \mathbf{S}, D_u^{\text{xrf}} \mathbf{A}_u^{\text{xrf}} \cdot \mathbf{S}) \quad (27)$$

where $\Theta = \{D_c^{\text{xrf}}, D_u^{\text{xrf}}, D_c^{\text{rgb}}, D_u^{\text{rgb}}, \mathbf{A}_c, \mathbf{A}_u^{\text{xrf}}, \mathbf{A}_u^{\text{rgb}}\}$. The first two terms of (27) are fidelity terms similar to that of (24). The third term is again a TV regularizer term that penalizes high spatial variation in neighboring pixels. The last term $\text{MI}(\cdot)$ is novel in that it penalizes the estimated amount of mutual information shared between the unique components

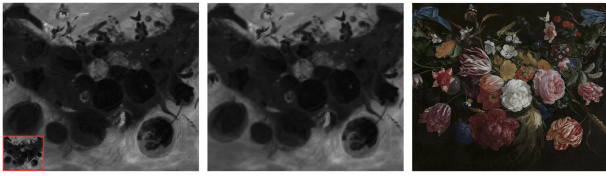


Figure 7

(Left) Low resolution Fe map and high-resolution ground truth Fe map. (Middle) The super-resolved Fe map. (Right) The high resolution RGB image: *Bloemen en Insecten*, Jan Davidsz. de Heem, 49 × 67 cm, Royal Museum of Fine Arts Antwerp, inv. no. 54, oil on canvas. Adapted from Yan et al.'s work [38].

between the low-resolution XRF and RGB data. Notice that the sparsity constraint was dropped as well, which perhaps is replaced by the mutual information loss to ensure the separation of the representations. We refer the reader to Yan et al.'s work [38] and Kraskov et al.'s [39] work on estimating mutual information from which the penalty term was derived.

They performed experiments showing that their method outperforms Dai et al.'s work on the painting *Bloemen en Insecten* by Jan Davidsz. de Heem. Their results are shown in Figure 7. They also report a 2.42 dB increase in the PSNR to reach 47.71 dB.

Super-resolution and inpainting of XRF image volumes using joint dictionary optimization methods prove to be quite effective in estimating an XRF cube in higher dimensions. Whether the dictionaries are found from a sparsity lens or a mutual information standpoint, these early methods are already powerful. Perhaps a joining of these two ideas can reduce the reconstruction error even further.

XRF Subsampling Design

The last category we will address is the design of subsampling algorithms. Subsampling is a staple in signal processing. In the XRF image setting, the goal of this problem is to find a subsampling pattern that reduces the total acquisition time of the XRF data. Since the scan time is quite long in many cases, quickly acquiring this data without sacrificing quality is starting to emerge in the literature.

While not a method to find an optimal sampling pattern, we briefly note the work by Chopp et al. [31] that provided some analysis on subsampling as it relates to the dwell time. Using simulated fast XRF scans of *Bloemen en Insecten*, they tested the limits of how short the dwell time could be without incurring too much error. Scan times that are 20 times as fast as the original scan was reported as a possibility, which would greatly speed up the XRF acquisition process; a mean-squared error of less than 3 was reported for this case.

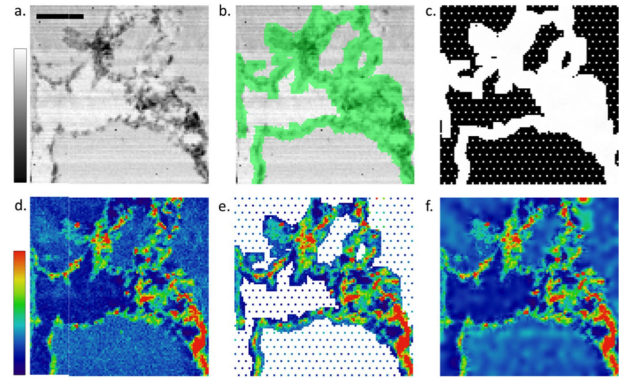


Figure 8

Hand-selected mask algorithm. (a) Quickly acquired STXM map. (b) Hand-selected region of interest. (c) Sampling mask with added sparse uniform sampling. (d) Si map of a full raster scan (6 hour). (e) Si map of the masked scan (2 hours). (f) Inpainted Si map of (e) using biharmonic inpainting. Reproduced from Kourousias et al.'s work [40].

In an effort to break from the raster scan, three existing methods are used to find the optimal sampling pattern in the following:

- 1) manual mask design;
- 2) convolutional neural networks (CNNs); and
- 3) reinforcement learning (RL).

We will cover each in this section.

Manual Mask Design

One simple approach to take is designing the sampling mask on one's own. Kourousias et al. [40] provided a method whereby they reduced the XRF scan time from 6 to 2 hours using a scanning transmission x-ray microscopy (STXM) map. First, the STXM map is rapidly acquired, which is reportedly up to two orders of magnitude faster than the XRF collection. Next, using this image as a guide, the user manually selects the regions of interest for XRF scanning. A sparse, uniform sampling pattern is overlaid on the mask so that the background is not completely ignored. These locations are scanned, then an inpainting algorithm estimates the pixels that were not scanned. The authors provide an example in Figure 8.

This approach is able to better capture the foreground, but it certainly has its limitations—primarily the possibility of human error. For example, a different display method (e.g., linear versus log scale) may highlight details in the first maps that could be undetected by human vision, resulting in foreground exclusion. This is where computational methods come into play. Aside from automating the mask design (and therefore further reducing the total acquisition time), an

algorithmic approach may not only select different areas, but also provide insight into how to allocate dwell times that vary per pixel.

Convolutional Neural Networks

CNN's have grown in popularity in part due to their ability to solve tasks by extracting local correlations in data. When designing a sampling mask, spatial correlations are exploited to find the best pixels to sample. These selected pixels are typically the most difficult ones to estimate should they have been excluded from the set of samples.

Dai et al. [37] used neural networks to find the best sampling mask. To do so, they introduced a CNN NetM whose purpose is to find the optimal sampling mask $\mathbf{S}$ for some given image $\mathbf{I}$ and sampling rate $r \in (0, 1)$. This is tricky to accomplish since there are no ground truth sampling masks available for creating training pairs. To circumvent this issue, another CNN, NetE, was appended to the output of the mask network. NetE requires a subsampled RGB image as input that is inpainted at the output.

The inpainting network was trained first separate from the mask generating network using the L2 norm as a loss function

$$\Theta_E^* = \arg \min_{\Theta_E} \|\mathbf{I} - \text{NetE}(\mathbf{I} \cdot \mathbf{S}; \Theta_E)\|_F^2 \quad (28)$$

where Θ_E are the neural network weights.

Once this network is trained, NetM can be trained with a fixed NetE in a feedforward manner. Instead of providing $\mathbf{S}$, it is instead estimated via NetM, whose loss function is

$$\Theta_M^* = \arg \min_{\Theta_M} \|\mathbf{I} - \text{NetE}(\mathbf{I} \cdot \text{NetM}(\mathbf{I}; r, \Theta_M))\|_F^2 \quad (29)$$

where Θ_M is the neural network weights for NetM. At inference time, the inpainting network NetE is dropped. The output of NetM is not binary; the final mask is instead drawn from a Bernoulli distribution where the pixel values of the mask are the probability of drawing a 1. This binarizes the output with the desired rate. Figure 9 shows a sample random mask and a mask from NetM when *Bloemen en Insecten* is the input (see Figure 7, right, for the painting).

Notice that the background is sparsely sampled due to its uniformity. The remaining samples are concentrated in the foreground where the flowers are located, which is intuitively where one would place the samples. This mask was shown to have improvements for inpainting over other algorithms that solve for an optimal sampling mask achieving a PSNR of 44.10 dB for 20 channels of the XRF volume. The results were consistent over different sampling rates as well.

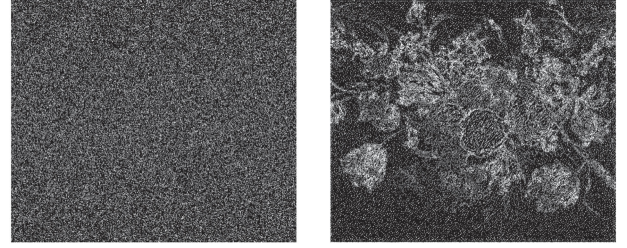


Figure 9

(Left) Random uniform mask with 20% of the pixels selected, denoted with white. (Right) NetM mask of *Bloemen en Insecten* also with 20% of the pixels selected. Reproduced from [37].

Reinforcement Learning

RL, much like CNNs, has gained in popularity for solving tasks related to action taking. The general framework of RL is based on a reward structure. With each action that is taken, a reward (or penalty) is calculated. The larger the reward at the end of an episode (i.e., a complete set of actions), the more the network will learn to take similar steps that previously rewarded it with a high score. RL is typically framed as a maximization problem, but this is not always the case.

Betterton et al. [41] developed a RL algorithm for XRF sampling. They formulated a sampling method using different apertures for the x-ray beam. This is to first capture the general scene in a quick manner, then allot more time to the areas of interest with more focused apertures. The scan times for local areas of the object are optimized with each aperture.

There are two objectives to jointly minimize in their formulation: 1) the main objective, L_Q , which is to provide the best quality image at the end of the scans, and 2) the amount of time spent scanning, L_C , is penalized if it is excessive. A simplified representation of their algorithm for training tries to find a path τ_t for each time step t that minimizes the expected sum of the quality penalty and time penalties

$$\tau^* = \arg \min_{\tau} \mathbb{E} \left[L_Q(\tau_T) + \lambda \sum_{t=1}^T L_C(\tau_t) \right]. \quad (30)$$

The first term only penalizes the deviation of the XRF estimation from the ground truth after the final time step. The second term penalizes the scan time for each of the T trajectories. Hyperparameter λ controls how much the time taken is penalized relative to the XRF fidelity term.

Figure 10 shows results on a cropped region of an XRF scan. The RL approach has both qualitative and quantitative improvements over the raster scanning method. More noise

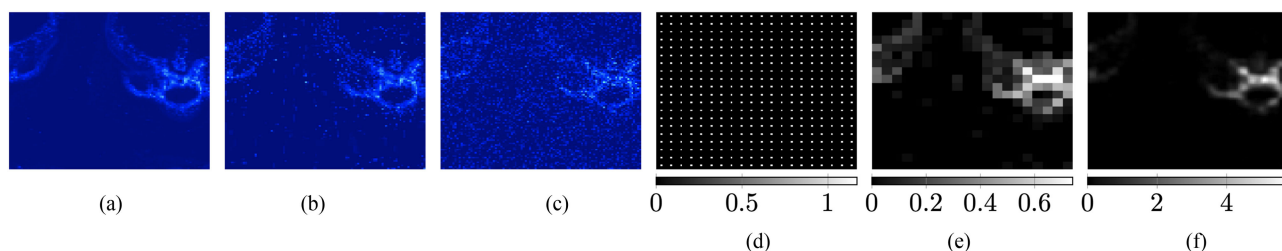


Figure 10

From left to right: (a) ground truth XRF sample. (b) XRF estimation using the method of Betterton et al.'s work [41]. (c) XRF estimation using a raster scan of equal time. (d) Initial scan. (e) Next adaptive scan. (f) Final adaptive scan. The scale in (d)–(f) denotes the scan time. Adapted from Betterton et al.'s work [41].

is present in the raster scan, and the foreground is better estimated as well. This is attributed to how the algorithm allocates more time to the foreground and avoids the background.

These results are impressive, but may be hard to adapt on a large scale. The ground truth resolution is 50×50 px, which is small for XRF volumes. Typical resolutions now are on the order of hundreds of pixels in height and width. RL problems are hard to optimize normally, not to mention the lack of XRF data needed to train large-scale networks.

Conclusion

The field of XRF imaging is quite young. Despite its youth, many paintings have been studied by many groups using many techniques. It is only recently that efforts are being made to join image processing research with XRF research.

We first introduced XRF imaging and analysis particularly for the signal processing community to become acquainted with this new imaging modality—establishing goals and problems that are faced by researchers in the XRF field. We then provided a take of XRF analysis through the lens of image and signal processing particularly as it relates to denoising techniques. Many XRF analysis methods can also be classified as denoising algorithms (mostly denoising in the spectral domain). This aspect, we believe, should see more attention as the metaphorical field is ripe for research. No matter if the goal is to improve how XRF signals are analyzed or denoise the data, we encourage cross-collaboration between the two disciplines to provide perspectives on one another's research.

In addition, we also reviewed some techniques for classical image processing problems, namely deblurring, super-resolution and inpainting, and signal subsampling. We hope that those in the XRF community can see the value of applying these algorithms to their own work, whether out of analytical necessity or just to collect more data in a faster manner.

The lack of published algorithms and XRF data is apparent. Alfeld et al. [42] have published data from the tomb of

Nakhtamun in Egypt available for public use.¹ This data can be used to try the methods reviewed here or develop new algorithms, although cross-collaboration between the image processing and XRF imaging communities is essential to understand the needs and concerns of both communities. We recognize that without a database or publicly available XRF data, it can hinder the advancement of XRF-based image processing algorithms. We, therefore, encourage those with XRF data to make their data publicly available or be open to cross-collaboration.

With time, all paintings naturally degrade. The sooner we, as a collective group, can develop better tools for XRF analysis and acquisition, the more of these timeless pieces we can understand and potentially delay their degradation.

References

- [1] J. Dik et al., "Visualization of a lost painting by Vincent Van Gogh using synchrotron radiation based x-ray fluorescence elemental mapping," *Anal. Chem.*, vol. 80, no. 16, pp. 6436–6442, 2008, doi: [10.1021/ac800965g](https://doi.org/10.1021/ac800965g).
- [2] A. Anitha et al., "Virtual underpainting reconstruction from x-ray fluorescence imaging data," in *Proc. 19th Eur. Signal Process. Conf.*, 2011, pp. 1239–1243.
- [3] A. Anitha et al., "Restoration of x-ray fluorescence images of hidden paintings," *Signal Process.*, vol. 93, no. 3, pp. 592–604, 2013, doi: [10.1016/j.sigpro.2012.09.027](https://doi.org/10.1016/j.sigpro.2012.09.027).
- [4] M. Alfeld et al., "A mobile instrument for in situ scanning macro-XRF investigation of historical paintings," *J. Anal. At. Spectrometry*, vol. 28, no. 5, pp. 760–767, 2013, doi: [10.1039/C3JA30341A](https://doi.org/10.1039/C3JA30341A).
- [5] M. Alfeld and L. De Viguierie, "Recent developments in spectroscopic imaging techniques for historical paintings—A review," *Spectrochim. Acta B: At. Spectrosc.*, vol. 136, pp. 81–105, 2017, doi: [10.1016/j.sab.2017.08.003](https://doi.org/10.1016/j.sab.2017.08.003).

¹Sample data: <https://sourceforge.net/p/datahandlerp/wiki/Tutorial%20Data/>

- [6] E. Pouyet et al., "New insights into Pablo Picasso's La Miséreuse accroupie (Barcelona, 1902) using x-ray fluorescence imaging and reflectance spectroscopies combined with micro-analyses of samples," *SN Appl. Sci.*, vol. 2, no. 8, pp. 1–6, 2020, doi: [10.1007/s42452-020-3130-4](https://doi.org/10.1007/s42452-020-3130-4).
- [7] M. Alfeld et al., "Scanning XRF investigation of a flower still life and its underlying composition from the collection of the Kröller-Müller Museum," *Appl. Phys. A*, vol. 111, no. 1, pp. 165–175, 2013, doi: [10.1007/s00339-012-7526-x](https://doi.org/10.1007/s00339-012-7526-x).
- [8] S. Saverwyns, C. Currie, and E. Lamas-Delgado, "Macro x-ray fluorescence scanning (MA-XRF) as tool in the authentication of paintings," *Microchem. J.*, vol. 137, pp. 139–147, 2018, doi: [10.1016/j.microc.2017.10.008](https://doi.org/10.1016/j.microc.2017.10.008).
- [9] K. Janssens et al., "Rembrandt's 'Saul and David' (c. 1652): Use of multiple types of smalt evidenced by means of non-destructive imaging," *Microchem. J.*, vol. 126, pp. 515–523, 2016, doi: [10.1016/j.microc.2016.01.013](https://doi.org/10.1016/j.microc.2016.01.013).
- [10] M. Alfeld, K. Janssens, J. Dik, W. de Nolf, and G. van der Snickt, "Optimization of mobile scanning macro-XRF systems for the in situ investigation of historical paintings," *J. Anal. At. Spectrometry*, vol. 26, no. 5, pp. 899–909, 2011, doi: [10.1039/C0JA00257G](https://doi.org/10.1039/C0JA00257G).
- [11] F. P. Romano et al., "Real-time elemental imaging of large dimension paintings with a novel mobile macro x-ray fluorescence (MA-XRF) scanning technique," *J. Anal. At. Spectrometry*, vol. 32, no. 4, pp. 773–781, 2017, doi: [10.1039/C6JA00439C](https://doi.org/10.1039/C6JA00439C).
- [12] F. Taccetti et al., "A multipurpose x-ray fluorescence scanner developed for in situ analysis," *Rendiconti Lincei. Scienze Fisiche e Naturali*, vol. 30, no. 2, pp. 307–322, 2019, doi: [10.1007/s12210-018-0756-x](https://doi.org/10.1007/s12210-018-0756-x).
- [13] E. Pouyet et al., "Development of a highly mobile and versatile large MA-XRF scanner for in situ analyses of painted work of arts," *X-Ray Spectrometry*, vol. 50, no. 4, pp. 263–271, 2021, doi: [10.1002/xrs.3173](https://doi.org/10.1002/xrs.3173).
- [14] K. Janssens, J. Dik, M. Cotte, and J. Susini, "Photon-based techniques for nondestructive subsurface analysis of painted cultural heritage artifacts," *Accounts Chem. Res.*, vol. 43, no. 6, pp. 814–825, 2010, doi: [10.1021/ar900248e](https://doi.org/10.1021/ar900248e).
- [15] P. Brouwer, *Theory of XRF*, Almelo, The Netherlands: PANalytical BV, 2006.
- [16] V. Solé, E. Papillon, M. Cotte, P. Walter, and J. Susini, "A multiplatform code for the analysis of energy-dispersive x-ray fluorescence spectra," *Spectrochim. Acta B: At. Spectrosc.*, vol. 62, no. 1, pp. 63–68, 2007, doi: [10.1016/j.sab.2006.12.002](https://doi.org/10.1016/j.sab.2006.12.002).
- [17] M. Alfeld and K. Janssens, "Strategies for processing mega-pixel x-ray fluorescence hyperspectral data: A case study on a version of Caravaggio's Painting Supper at Emmaus," *J. Anal. At. Spectrometry*, vol. 30, no. 3, pp. 777–789, 2015, doi: [10.1039/C4JA00387J](https://doi.org/10.1039/C4JA00387J).
- [18] R. Browning, "Multi-spectral imaging in materials microanalysis," *Surf. Interface Anal.*, vol. 20, no. 6, pp. 495–502, 1993, doi: [10.1002/sia.740200602](https://doi.org/10.1002/sia.740200602).
- [19] G. H. Bakır, J. Weston, and B. Schölkopf, "Learning to find pre-images," *Adv. Neural Inf. Process. Syst.*, vol. 16, pp. 449–456, 2004.
- [20] B. Vekemans, K. Janssens, L. Vincze, A. Aerts, F. Adams, and J. Hertogen, "Automated segmentation of μ -XRF image sets," *X-Ray Spectrometry*, vol. 26, no. 6, pp. 333–346, 1997, doi: [10.1002/\(SICI\)1097-4539\(199711/12\)26:6%3C333::AID-XRS231%3E3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-4539(199711/12)26:6%3C333::AID-XRS231%3E3.0.CO;2-D).
- [21] L. Xu, J. Li, Y. Shu, and J. Peng, "SAR image denoising via clustering-based principal component analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6858–6869, Nov. 2014, doi: [10.1109/TGRS.2014.2304298](https://doi.org/10.1109/TGRS.2014.2304298).
- [22] M. Alfeld, M. Wahabzada, C. Bauckhage, K. Kersting, G. Wellenreuther, and G. Falkenberg, "Non-negative factor analysis supporting the interpretation of elemental distribution images acquired by XRF," *J. Phys.: Conf. Ser.*, Hamburg, Germany, vol. 499, no. 1, 2014, Art. no. 012013, doi: [10.1088/1742-6596/499/1/012013](https://doi.org/10.1088/1742-6596/499/1/012013).
- [23] R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," *J. Chemometrics*, vol. 11, no. 5, pp. 393–401, 1997, doi: [10.1002/\(SICI\)1099-128X\(199709/10\)11:5%3C393::AID-CEM483%3E3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-128X(199709/10)11:5%3C393::AID-CEM483%3E3.0.CO;2-L).
- [24] H. C. Santos et al., "Identification of forgeries in historical enamels by combining the non-destructive scanning XRF imaging and alpha-pixe portable techniques," *Microchem. J.*, vol. 124, pp. 241–246, 2016, doi: [10.1016/j.microc.2015.08.025](https://doi.org/10.1016/j.microc.2015.08.025).
- [25] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, vol. 2, no. 1. New York, NY, USA: Springer, 2010, doi: [10.1007/978-1-4419-7011-4](https://doi.org/10.1007/978-1-4419-7011-4).
- [26] R. Mak et al., "Non-negative matrix analysis for effective feature extraction in x-ray spectromicroscopy," *Faraday Discuss.*, vol. 171, pp. 357–371, 2014, doi: [10.1039/C4FD00023D](https://doi.org/10.1039/C4FD00023D).
- [27] A. De Juan, J. Jaumot, and R. Tauler, "Multivariate curve resolution (MCR). Solving the mixture analysis problem," *Anal. Methods*, vol. 6, no. 14, pp. 4964–4976, 2014, doi: [10.1039/C4AY00571F](https://doi.org/10.1039/C4AY00571F).
- [28] A. Martins et al., "Jackson Pollock's Number 1a, 1948: A non-invasive study using macro-x-ray fluorescence mapping (MA-XRF) and multivariate curve resolution-alternating least squares (MCR-ALS) analysis," *Heritage Sci.*, vol. 4, no. 1, pp. 1–13, 2016, doi: [10.1186/s40494-016-0105-2](https://doi.org/10.1186/s40494-016-0105-2).
- [29] M. Alfeld et al., "Simplex volume maximization (SIVM): A matrix factorization algorithm with non-negative constraints and low computing demands for the interpretation of full spectral x-ray fluorescence imaging data," *Microchem. J.*, vol. 132, pp. 179–184, 2017, doi: [10.1016/j.microc.2017.02.001](https://doi.org/10.1016/j.microc.2017.02.001).

- [30] B. Rasti, B. Koirala, P. Scheunders, and P. Ghamisi, "How hyperspectral image unmixing and denoising can boost each other," *Remote Sens.*, vol. 12, no. 11, 2020, Art. no. 1728, doi: [10.3390/rs12111728](https://doi.org/10.3390/rs12111728).
- [31] H. Chopp, A. McGeachy, M. Alfeld, O. Cossairt, M. Walton, and A. Katsaggelos, "Denoising fast x-ray fluorescence raster scans of paintings," 2022, *arXiv:2206.01740*, doi: [10.48550/arXiv.2206.01740](https://doi.org/10.48550/arXiv.2206.01740).
- [32] J. Yang, Z. Zhang, and Q. Cheng, "Resolution enhancement in micro-XRF using image restoration techniques," *J. Anal. At. Spectrometry*, vol. 37, no. 4, pp. 750–758, 2022, doi: [10.1039/D1JA00425E](https://doi.org/10.1039/D1JA00425E).
- [33] M. Bertero, P. Boccacci, and C. De Mol, *Introduction to Inverse Problems in Imaging*. Boca Raton, FL, USA: CRC Press, 2021, doi: [10.1201/9781003032755](https://doi.org/10.1201/9781003032755).
- [34] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009, doi: [10.1137/080716542](https://doi.org/10.1137/080716542).
- [35] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via l0-regularized intensity and gradient prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2901–2908, doi: [10.1109/CVPR.2014.371](https://doi.org/10.1109/CVPR.2014.371).
- [36] Q. Dai, E. Pouyet, O. Cossairt, M. Walton, and A. K. Katsaggelos, "Spatial-spectral representation for x-ray fluorescence image super-resolution," *IEEE Trans. Comput. Imag.*, vol. 3, no. 3, pp. 432–444, Sep. 2017, doi: [10.1109/TCI.2017.2703987](https://doi.org/10.1109/TCI.2017.2703987).
- [37] Q. Dai, H. Chopp, E. Pouyet, O. Cossairt, M. Walton, and A. K. Katsaggelos, "Adaptive image sampling using deep learning and its application on x-ray fluorescence image reconstruction," *IEEE Trans. Multimedia*, vol. 22, pp. 2564–2578, 2019, doi: [10.1109/TMM.2019.2958760](https://doi.org/10.1109/TMM.2019.2958760).
- [38] S. Yan, H. Verinaz-Jadan, J.-J. Huang, N. Daly, C. Higgitt, and P. Dragotti, "Super-resolution for macro x-ray fluorescence data collected from old master paintings," 2022. [Online]. Available: <https://www.researchgate.net/publication/358827119>
- [39] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, no. 6, 2004, Art. no. 066138, doi: [10.1103/PhysRevE.69.066138](https://doi.org/10.1103/PhysRevE.69.066138).
- [40] G. Kourousias et al., "Compressive sensing for dynamic XRF scanning," *Sci. Rep.*, vol. 10, no. 1, pp. 1–8, 2020, doi: [10.1038/s41598-020-66435-6](https://doi.org/10.1038/s41598-020-66435-6).
- [41] J.-R. Betterton, D. Ratner, S. Webb, and M. Kochenderfer, "Reinforcement learning for adaptive illumination with x-rays," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 328–334, doi: [10.1109/ICRA40945.2020.9196614](https://doi.org/10.1109/ICRA40945.2020.9196614).
- [42] M. Alfeld, S. Pedetti, P. Martinez, and P. Walter, "Joint data treatment for VIS-NIR reflectance imaging spectroscopy and XRF imaging acquired in the theban necropolis in Egypt by data fusion and T-SNE," *Comptes Rendus Physique*, vol. 19, no. 7, pp. 625–635, 2018, doi: [10.1016/j.crhy.2018.08.004](https://doi.org/10.1016/j.crhy.2018.08.004).



schemes to optimize object tracking performance. His research interests include machine learning and optimization techniques for smart image sampling, and restoration/enhancement.



USA. She gained experience in cultural heritage research through an Andrew W. Mellon Foundation fellowship at the Metropolitan Museum of Art (2018–2020).



essor with the TU Delft, Delft, The Netherlands, for "X-rays in art and archaeology," working on the investigation of 3-D cultural heritage objects by dedicated instruments, the development of new data evaluation workflows and contributing to the Smart*Light project, an Inverse Compton Source for X-rays, since November 2018.

Henry Chopp received the M.S. degree in electrical engineering from Northwestern University, Evanston, IL, USA, in 2020, where he is currently working toward the Ph.D. degree in electrical engineering. His work is on the reduction of macro x-ray fluorescence scan and exposure times when imaging cultural heritage paintings. He has previously researched adaptive lossy video acquisition

Alicia McGeachy received the Ph.D. degree in chemistry from Northwestern University, Evanston, IL, USA, where she applied nonlinear optical spectroscopy toward understanding and quantifying the interactions between model cellular membrane systems and nanoscale materials. She is currently an associate scientist with the Metropolitan Museum of Art, New York, NY,

Matthias Alfeld received the Ph.D. degree in the topic of the development of scanning macro-XRF for the investigation of historical paintings, involving the design of instruments and development of data evaluation workflows, from the University of Hamburg, Hamburg, Germany, and University of Antwerp, Antwerp, Belgium, in 2013. He is currently an assistant profes-



Oliver Cossairt received the Ph.D. degree in computer science from Columbia University, New York City, NY, USA. He is currently an associate professor with the Department of Computer Science and the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL, USA. He is the Director of the Computational Photography Laboratory, Northwestern

University. His research interests include optics/photonics, computer graphics, computer vision, machine learning, and image processing. The general goal of the lab is to develop imaging hardware and algorithms that can be applied across a broad range of physical scales, from nanometer to astronomical.



Marc Walton received the Ph.D. degree in archaeological science from the University of Oxford, Oxford, U.K. He is currently the Head of Conservation and Research with M+ Museum for Visual Culture, Hong Kong. Prior to joining M+ in 2022, he was a Senior Scientist with the Northwestern University/Art Institute of Chicago Center for Scientific Studies in the Arts (NU-

ACCESS), which he joined in 2013. At NU-ACCESS, he led numerous collaborative initiatives and research projects to investigate art objects. His research interests include elucidating the trade and manufacture of objects using the physical sciences, and on the development of the use of imaging technologies in the field of conservation science.



Aggelos Katsaggelos received the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA. He is currently a Professor and Joseph Cummings Chair with the ECE Department, Northwestern University, Evanston, IL, USA, and co-director of NU-ACCESS. He runs the Image and Video Processing Laboratory, whose objective is to generate cutting-edge research

results in the fields of multimedia signal processing, multimedia communications, computational imaging, and machine learning.