

PAPER • OPEN ACCESS

## Improvement of the visibility of concealed features in misregistered NIR reflectograms by deep learning

To cite this article: Jan Blažek *et al* 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* **364** 012058

View the [article online](#) for updates and enhancements.

### Related content

- [A technique for detecting and locating polarisation nonuniformities in an anisotropic optical fibre](#)  
V V Burdin, Yuri A Konstantinov, Vladimir P Pervadchuk *et al.*
- [Influence of modulation instability on the operation of phase-sensitive optical time domain reflectometers](#)  
S P Nikitin, P I Ulanovskiy, A I Kuzmenkov *et al.*
- [Use of spectral characteristics of DSLR cameras with Bayer filter sensors](#)  
P A Cheremkhin, V V Lesnichii and N V Petrov

# Improvement of the visibility of concealed features in misregistered NIR reflectograms by deep learning

Jan Blažek<sup>1</sup>, Oldřich Vlašic<sup>2</sup>, Barbara Zitová<sup>1</sup>

<sup>1</sup> Institute of Information Theory and Automation of the CAS, Pod Vodárenskou věží 8, Prague 8, Czech Republic

<sup>2</sup> Faculty of Nuclear Sciences and Physical Engineering, Břehová 7, Prague 1, Czech Republic

E-mail: blazek@utia.cas.cz

**Abstract.** Features of Old Master paintings hidden under the upper layer of a painting are often studied using NIR reflectograms; however their interpretability can be reduced due to the visible content. In our previous work [3] we described the possibility of increasing the visibility of concealed features in NIR reflectograms from the painting surface. The method output, enhanced NIR reflectogram, is produced by extrapolating the VIS data to a NIR range reflectogram and subtracting it from the acquired data in the NIR spectral subband. As a result, separated information from the NIR domain is obtained. This method has a severe limitation, because it requires precise image registration of the VIS and NIR spectral bands. This is often hard to achieve, because DSLR cameras or multiple devices with various optical systems are used for data collection, and the mutual spatial relation of the images is often unknown. Thus, in the original form, the algorithm was applicable only for data acquired using special scanners producing spatially registered images (as in [4]). In this work, we present an extension of the previous algorithm inspired by deep learning. The new concept allows processing of images only partially registered with pixel precision; subpixel accuracy is no longer needed. We suggest an extension of neural network input with neighboring pixels and allocation of extra ANN layers for translation compensation. The results are demonstrated on misregistered images captured by DSLR camera in VIS and NIR.

## 1. Introduction

The study of underdrawings and pentimenti is a standard procedure of artwork investigation. However, the quality of scanning devices is often not sufficient, in the sense of acquisition of spatially registered dataset. For this reason, reflectograms from various spectral bands are in most case studies processed separately. This is a serious limitation because some relevant information can remain unnoticed. The information gain of a spectral subband is often covered with high intensity information which is shared among modalities and in this way remains unnoticed. In our previous research, we suggest an algorithm for separation of information which is shared among the VIS and NIR part of the spectra and the information gain of NIR [3]. The prerequisite for such separation is the accurate registration of all relevant images with subpixel precision.

### 1.1. Registration

Automatic registration of multimodal images of artworks is still an unresolved problem. This is a difficult problem because there is no universal similarity measure for various modalities.



The set of captured modalities contains images (ordered according to their availability) in these bands:

- (i) Visible light (VIS)
- (ii) Near infrared (NIR and also SWIR)
- (iii) Ultraviolet fluorescence (UVF)
- (iv) X-ray
- (v) X-ray fluorescence (XRF)
- (vi) Far infrared or “terahertz” (THz)

While one scanner can be used for both the VIS and NIR bands (with the same optical elements) [4, 6], for THz or X-ray images multiple scanners are necessary. The use of one scanner means that acquired raw data can be spatially registered, or that a registration function is measurable. On the other hand, the use of more than one scanning device requires spatial normalization according to the optical elements used (lenses, prisms, monochromator, etc.), position of the scanned artwork, and physical properties (index of refraction) of the acquired wavelengths. An efficient estimation of the parameters of such a transformation from an acquisition setting is significantly harder than their estimation from the collected dataset.

Precise registration of a multimodal dataset is required for further digital image processing (image comparison [9], false-color images [12, 7], spectral analysis [10, 11], pigment classification, etc.). The registration techniques from the referenced papers are based on control point selection or on minimization of mutual information [1]. As the authors in [1] described, registration of this type is not precise: “..., the attenuation correction is not always accurate. While errors in registration are a possible cause, ...”. The same problem occurs when registration based on control points is used, in spite of the precise registration of control points (precise position of control points is achieved by the application of the similarity measure on roughly placed points [5]). The problem here is that the present transformation function is not polynomial, thus its approximation from a finite number of control points may be impossible. The registration precision varies over the data set in this case.

The most relevant question is “Why is there no relevant similarity measure?”

In our previous research [3] we were looking for the answer. There were also various approaches regarding similarity estimation: intensity difference variance [9], mutual information [2], and coupled dictionaries [8]. All of them worked with a matrix of corresponding intensities (ICM), which contains the frequency of correspondent pixel intensities in two different modalities (2D histogram of intensities). The number of rows and columns corresponds to the intensity levels recognized in each modality and greatly affects the estimated separation. If we use a correspondence matrix as a visualization of mutual information (or entropy), we recognize two extreme patterns:

- The highest entropy/lowest mutual information is reached when the correspondence matrix is uniform with zero total variance.
- The lowest entropy/highest mutual information is reached when there is only one non-zero column for each row, and vice versa.

The first case leads to a very smooth registration error function, which means, that global minima is searched quickly, but often does not reflect the present transformation function, because correct mapping is not a function (multiple values in one modality can correspond to one value in the second one).

The second case leads to a precise definition of global minima without any information gain for each modality.

The similarity of two modalities (without the context of neighborhood pixels) describes well the ICM and, according to the previous description, such a measure is limited to “good” cases. In [3] we quantify the minimal correspondence required for separation of spectral subband information gain, which is about 90%. This threshold was empirically measured on numerous ANNs with a perfectly registered dataset. When these conditions are satisfied, subband information gain can be computed and concealed features are uncovered in the subband reflectogram.

The question is now a little bit shifted. Let’s assume a suitable dataset for VIS/NIR comparison. We would like to use ICM according to [3]. However, “What happens when the dataset is not registered well?”

We try to use the same similarity measure for automatic registration.

## 2. Extension of the comparison algorithm

The use of ANN is advantageous due to its modularity. The main idea is to extend ANN input with the intensities of neighboring pixels. In this case, the ANN can also estimate transformation, necessary for registration of the input and output images of our original ANN. In the following paragraphs we present a proof of this concept on a “dummy architecture”, estimating the image shift.

The architecture of ANN is proposed in Figure 1. The original algorithm contained only white and ochre neurons. These neurons play the role of an extrapolation function from the information content of VIS spectral bands into NIR spectral bands on the registered dataset. The inputs of the original ANN are VIS pixel intensities, and the outputs are intensities taken in NIR. Extrapolation was computed by two hidden layers.

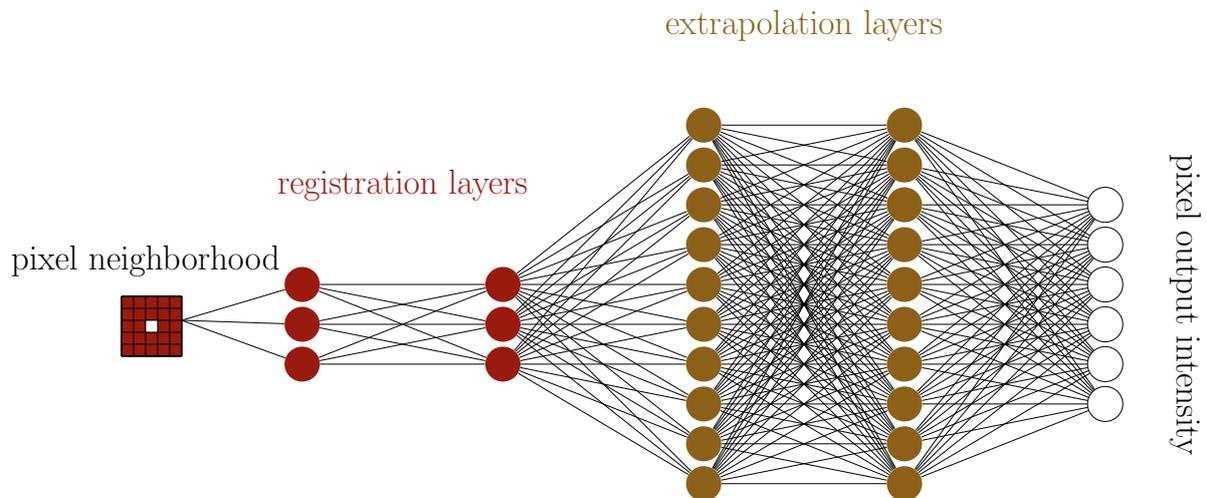
When misregistration is present, we extend the input layer to accept not only one pixel but the whole neighborhood of a specified size. The reason behind this is that we don’t know which pixel corresponds to the one at the output. To prevent the curse of dimensionality, we limit neighborhood size according to the estimated maximal pixel shift. An effective estimation of maximal pixel shift gives us faster computation and more precise results. After the input layer, the registration layers follow. In our case we use two layers with several neurons for shift estimation. We expect that estimation of more complex transformations will require more neurons (and probably more sophisticated architecture), but for pixel shift two layers are more than sufficient. We assume that such architecture of the registration layers (controlled by the extrapolation error function) will identify a pixel from the pixel neighborhood corresponding to the expected output pixel.

The improvement achieved by this new concept compared to the previous work [3] is demonstrated in Figure 2.

## 3. Hypotheses and Experiments

In evaluation of the proposed concept, we track several targets. First of all, we need to confirm or disprove our expectations. Second, we would like to tune ANN architecture for fast training and efficient data processing. We postulate three hypotheses which can disprove the proposed concept or indicate a direction for further research. Moreover, to prevent an over-parametrization of the proposed system, we evaluate effect of each relevant parameter separately. As a result, we would like to define the limits of registration layers.

The atomic operation for evaluating of our hypotheses was the processing of phantoms. These phantoms were created from a precisely registered VIS-NIR dataset (obtained from real artworks — source INO-CNR in Florence) where the NIR bands were slightly shifted. The shift was specific for each phantom and goes from 1px up to 40px. If the shift amount was a non-integer, the NIR intensities were computed by cubic spline approximation.



**Figure 1.** Extension of original algorithm with new registration layers and input pixels (marked with red color).

### 3.1. Hypothesis 1: The absolute amount of misregistration affects performance

We expect that input images must be partially aligned. The reason for this hypothesis is the statistical principle of an ANN training. Reduction of the space of possible transformations can improve the speed and accuracy of the ANN training. And the partial registration decreases the number of relevant input pixels.

For the evaluation, we prepare phantoms with a shift going from 1 to 15 pixels and then measure, on several runs of the algorithm, the error per pixel. Results are shown in Figure 3. The first observation confirms our hypothesis: **An increasing shift amount produces a more erroneous output.** The second observation is the convergence of the histogram curve with the increasing shift amount.

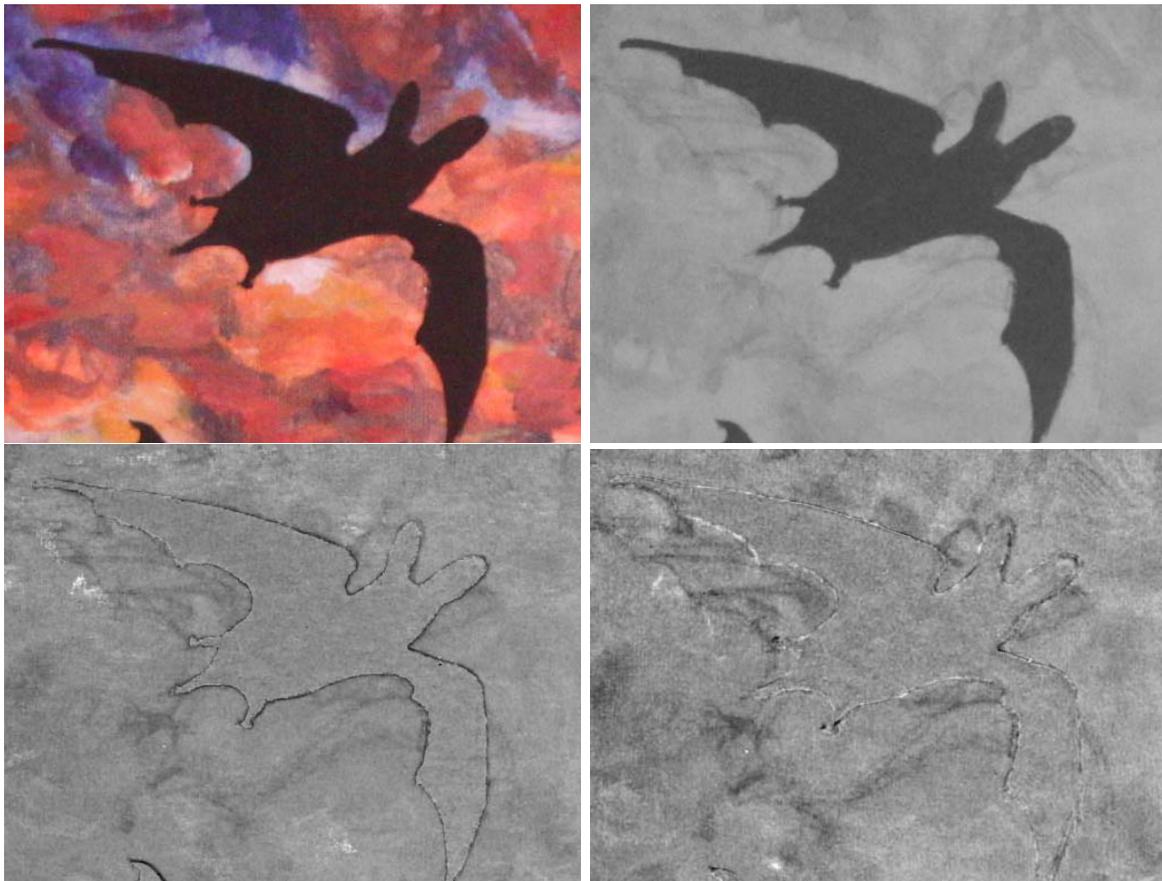
The explanation for the second observation is not clear, but in general we can say that information relevant for the ANN training drops under the ratio where training is efficient.

### 3.2. Hypothesis 2: Layer(s) for estimation of modality misregistration can improve the achieved results

A logical split of the ANN in registration and extrapolation layers is a part of our concept but should be supported by another experiment. There is a possibility that without extra layers, the ANN will be similarly effective. In Figure 4 we present a comparison of the results of ANN with extra layers versus ANN without extra layers. Various layer widths have been tested as well, going from 3 – 3 – 25 – 25 and 5 – 5 – 25 – 25 to 40 – 40 – 25 – 25.

According to our metric (histogram of errors) we see that small amount of neurons in *registration layers* (i.e. 3 – 3 or 5 – 5) is worse than ANN without these layers. This fact poses a question: After the training, are the *registration layers* focused on selecting one of the input pixels? The answer seems to be negative, as observation on trained ANN weights shows that ANN needs extra regularization parameters to achieve this behavior. Consequently, we reject this hypothesis, with the promise of further tests on more suitable ANN architectures.

Here we propose the requirements for *registration layers*. For a shift transformation, the ANN *registration layer* has to identify the most relevant pixel in the input set. This will be represented as the highest weight for the corresponding pixel and decreasing weights with growing distance.



**Figure 2.** Comparison of the result of the original algorithm, working with registered images, with the improved one. (a) VIS subband, (b) misregistered NIR subband, (c) output of original algorithm (without pixel neighborhood), (d) improved algorithm (input contains neighboring pixels). The visibility of underdrawings on image (d) is much better. ANN has efficiently estimated the transform function, while the original algorithm works as an “edge detector”.

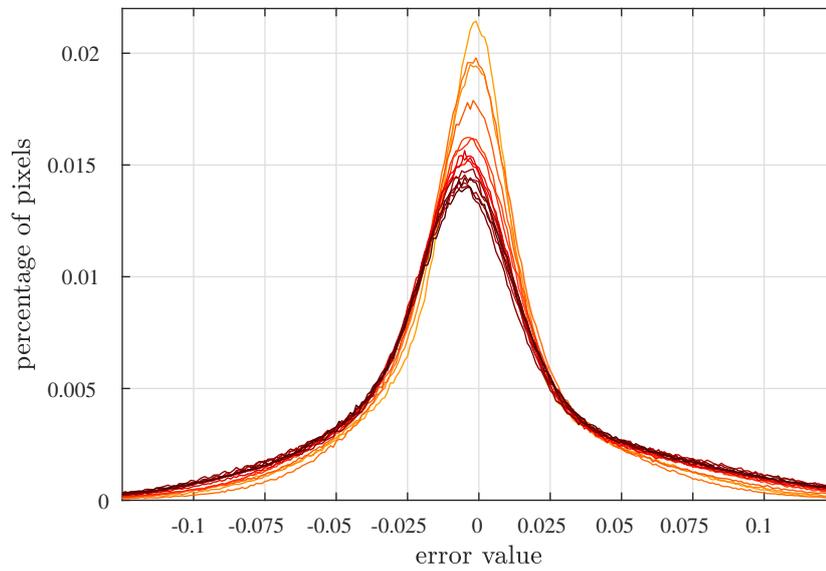
### *3.3. Hypothesis 3: A higher number of output bands processed at one time decreases the performance of a single band*

In practice, it is very time consuming to manually go through numerous images and select the best one for processing. Therefore, the efficiency of “processing simultaneously” should be measured. Such an approach can lead to the failure of ANN training, and therefore it is relevant to ask whether the number of output layers negatively affects the results. Our dataset contains 15 bands in NIR, which allows us to construct an experiment comparing the error achieved for ANNs with from 1 to 15 output neurons, see Figure 5.

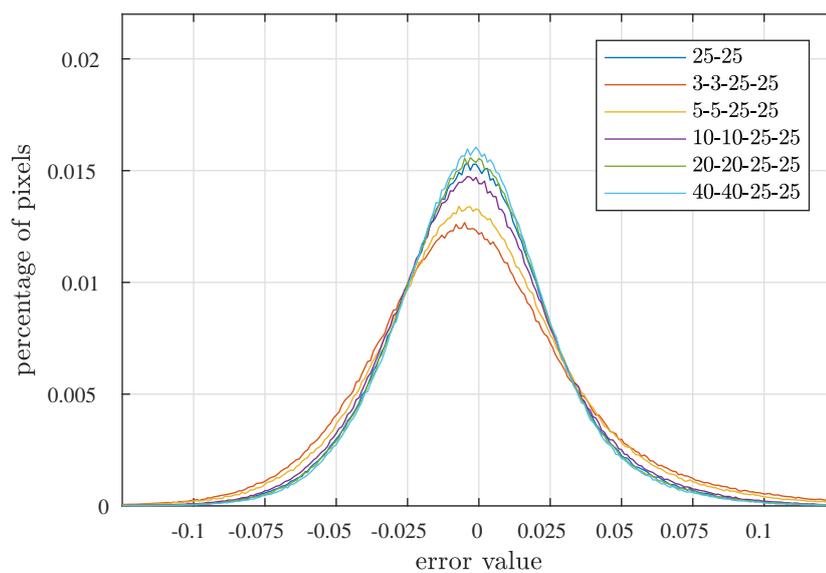
According to Figure 5, the algorithm performs better for a smaller number of bands. However, a deeper analysis of the results shows that for low shifts of modalities (from 0 to 10 pixels), the histograms are almost the same. The hypothesis was therefore neither proved nor disproved. We conclude here that a better metric of *quality* is necessary.

## **4. Conclusion**

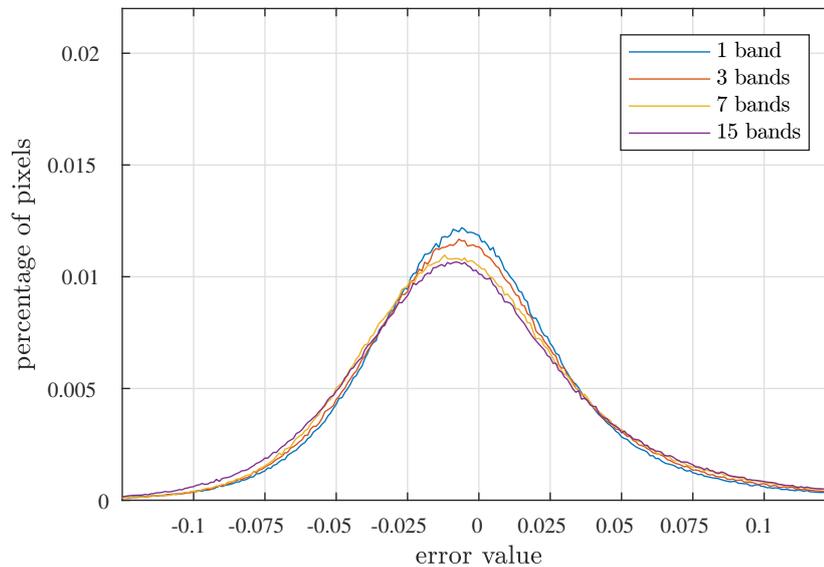
We suggest an improved variant of our algorithm for improvement of the visibility of concealed features in a misregistered NIR modality. We demonstrate the efficiency of the proposed algorithm on a real dataset and demonstrate its limits on manually shifted NIR images. We



**Figure 3.** Histogram of error per pixel. ANN architecture 5 – 5 – 25 – 25. The colors go from the lightest 1px shift to the darkest 15px shift.



**Figure 4.** Comparison of histograms of error per pixel for different ANN architectures. Layer widths are given in the graph legend.



**Figure 5.** Histogram of error per pixel with respect to the number of target NIR subbands, namely 1, 3, 7, and 15. Errors were evaluated over a single band common to all configurations. ANN architecture 3 – 3 – 25 – 25.

identify that a crucial part of the improvement is the addition of neighboring pixels into ANN inputs. We also evaluate several configurations and define limits of the suggested ANN architecture.

Our concept joins two tasks into one. This approach is rare because the complexity of the problem can get out of control. However, we have shown that registration can be joined together with extrapolation using the same similarity measure. A logical future work is the preparation of stronger rules for ANN training to achieve splitting of the registration and extrapolation steps.

This preliminary research raises a lot of questions which should be answered in future work:

- How can we regularize *registration layers* of the ANN to obtain a transformation function?
- Is there an ANN architecture that prevents the curse of dimensionality? And if so, is it possible to process images without a preceding rough registration?

These questions should be answered, and we will focus on them in subsequent work.

### Acknowledgments

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme “Projects of Large Infrastructure for Research, Development, and Innovations” (LM2010005), is greatly appreciated. Without them this research would not be possible.

We would like to acknowledge the Czech Science Foundation, which has supported this work (project no. GA18-05360S).

### References

- [1] Anila Anitha, Andrei Brasoveanu, Marco F Duarte, Shannon M Hughes, Ingrid Daubechies, Joris Dik, Koen Janssens, and Matthias Alfeld. “Virtual underpainting reconstruction from X-ray fluorescence imaging Data”. In: *19th European Signal Processing Conference*. Eusipco. 2011, pp. 1239–1243.

- [2] Anila Anitha, Andrei Brasoveanu, Marco Duarte, Shannon Hughes, Ingrid Daubechies, Joris Dik, Koen Janssens, and Matthias Alfeld. “Restoration of X-ray fluorescence images of hidden paintings”. In: *Signal Processing* 93.3 (2013). Image Processing for Digital Art Work, pp. 592–604. ISSN: 0165-1684. DOI: 10.1016/j.sigpro.2012.09.027. URL: <http://www.sciencedirect.com/science/article/pii/S0165168412003659>.
- [3] Jan Blažek, Jana Striová, Raffaella Fontana, and Barbara Zitová. “Improvement of the visibility of concealed features in artwork NIR reflectograms by information separation”. In: *Digital Signal Processing* 60 (Jan. 2017), pp. 140–151.
- [4] Claudio Bonifazzi, Pierluigi Carcagnì, Raffaella Fontana, Marinella Greco, Maria Mastroianni, Marzia Materazzi, Enrico Pampaloni, Luca Pezzati, and Davide Bencini. “A scanning device for VIS–NIR multispectral imaging of paintings”. In: *Journal of Optics A: Pure and Applied Optics* 10.6 (2008), p. 064011. URL: <http://stacks.iop.org/1464-4258/10/i=6/a=064011>.
- [5] Damon M. Conover, John K. Delaney, and Murray H. Loew. “Automatic registration and mosaicking of technical images of Old Master paintings”. In: *Applied Physics A: Materials Science and Processing* 8790 (2015), pp. 1–8. ISSN: 14320630. DOI: 10.1007/s00339-015-9140-1.
- [6] Costanza Cucci, Andrea Casini, Marcello Picollo, Marco Poggese, and Lorenzo Stefani. “A hyper-spectral scanner for high-quality image spectroscopy: Digital documentation and spectroscopic characterization of polychrome surfaces”. In: *Art '11: 10th International Conference on Non-Destructive Investigations and Microanalysis for the Diagnostics and Conservation of Cultural and Environmental Heritage* (2011), pp. 1–7.
- [7] John K. Delaney, Jason G. Zeibel, Mathieu Thoury, Roy Littleton, Kathryn M. Morales, Michael Palmer, and E. René de la Rie. “Visible and infrared reflectance imaging spectroscopy of paintings: Pigment mapping and improved infrared reflectography”. In: *Proceedings of SPIE*. Ed. by Luca Pezzati and Renzo Salimbeni. Vol. 7391. Springer International Publishing, July 2009, p. 739103. ISBN: 978-0-8194-7674-6. DOI: 10.1117/12.827493.
- [8] N. Deligiannis, J. F. C. Mota, B. Cornelis, M. R. D. Rodrigues, and I. Daubechies. “Multi-modal dictionary learning for image separation with application in art investigation”. In: *IEEE Transactions on Image Processing* PP.99 (2016), pp. 1–12. ISSN: 1057-7149. DOI: 10.1109/TIP.2016.2623484.
- [9] Amy Gooch and Jack Tumblin. “Visualizing pentimenti: Revealing the hidden history of paintings”. In: *Journal of Mathematics and the Arts* 00.00 (2007), pp. 1–11. ISSN: 17513472. DOI: 10.1080/1751347YYxxxxxxx.
- [10] Anna Pelagotti, Luca Pezzati, Natalia Bevilacqua, Veronica Vascotto, Vincent Reillon, and Claudia Daffara. “A study of UV fluorescence emission of painting materials”. In: *Art '05: 8th International Conference on Non-Destructive Investigations and Microanalysis for the Diagnostics and Conservation of Cultural and Environmental Heritage* (2005), A97.
- [11] Paola Ricciardi, John K. Delaney, Michelle Facini, Jason G. Zeibel, Marcello Picollo, Suzanne Lomax, and Murray Loew. “Near infrared reflectance imaging spectroscopy to map paint binders in situ on illuminated manuscripts”. In: *Angewandte Chemie - International Edition* 51.23 (2012), pp. 5607–5610. ISSN: 14337851. DOI: 10.1002/anie.201200840.
- [12] Chad Weiner. *Improved Acquisition of Underdrawings in Oil Paintings Using IR-Reflectography*. Tech. rep. Rochester Institute of Technology, May 1998.