Paint Loss Detection via Kernel Sparse Representation

Shaoguang Huang[1], Laurens Meeus[1], Bruno Cornelis[2], Bart Devolder[3], Maximiliaan Martens[4] and Aleksandra Pižurica[1].

[1]Department of Telecommunications and Information Processing, IPI, imec research group at Ghent University, Belgium

[2]Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel, Belgium

[3]Department of Painting Conservation, Royal Institute for Cultural Heritage (KIK/IRPA), Belgium

[4]Department of Art, music and theatre sciences, Ghent University, Belgium

***Abstract*— Automatic paint loss detection is desired for supporting conservation/restoration treatments of paintings. Firstly, producing condition reports with appropriate damage surveys requires now a lot of manual work from the restorers. Secondly, paint losses have to be accurately detected prior to running virtual restoration. Large variation of paint loss in size, shape, intensity as well as varying and complex background make this problem a challenging task. We develop a multimodal paint loss detection method based on sparse representation, which incorporates the information from multiple imaging modalities in a high-dimensional kernel feature space and makes use of the spatial context. To cope with unreliable labelled data, we introduce a majority voting approach. Experimental results with the data set of the *Ghent Altarpiece* demonstrate the effectiveness of the proposed approach.**

1 Introduction

Digital painting analysis has been a rapidly growing field, attracting a lot of interest recently in the signal processing community [1]. The tasks such as characterization of painting style and forgery detection [2, 3], crack detection [4], authorship identification [5], classification of ancient coins [6], canvases [7] and portraits [8], removal of canvas patterns [9] and inpainting [10, 11] have demonstrated the great potential of digital image processing techniques.

Loss of paint is typically caused by abrasion and mechanical fracture. In old oil paintings, paint losses were often overpainted during various restoration campaigns. Modern conservation treatments typically require not only removal of old varnish, but also removal of old retouches and overpaint, which may reveal paint losses underneath [13]. Detection of such paint loss areas is of great importance to painting conservators for estimating the extent of the damaged area, which needs to be maintained for documenting purposes, but also as a crucial step for virtual inpainting to provide simulations for the actual restoration. Despite the importance of automatic paint loss detection, this problem has received little attention in the literature so far. Nowadays, paintings are typically scanned with a multitude of imaging modalities. During restoration campaigns, additional scans are typically made at various stages of the rest oration treatment. Examples are shown in Fig. 1. (a) - (e). We want to exploit such multi-modal information to detect paint losses more reliably. Our approach will be based on constructing (training) a dictionary of prototypes that can be used to effectively, i.e. sparsely, represent paint loss samples.

Sparse Representation Classification (SRC) [14] proved to be effective in various image classification tasks, especially in computer vision and remote sensing. It assumes that each test sample can be sparsely represented as a linear combination of atoms from a dictionary which is constructed by the selected training samples. Directly applying SRC to our task results in poor performance due to the large variability of paint loss, and complex background. To cope with these challenges, it is necessary to incorporate appropriately both spatial context and inter-modal dependencies. Our previous work employed several spatial features within local patches and achieved a good detection performance [13]. However, hand-crafting such features leaves much choice and would involve ad-hoc choices and a lot of manual tuning. Therefore, in this paper we propose a multimodal paint loss detection method based on sparse representation that directly exploits the information from multiple imaging modalities in the kernel feature space and integrates the spatial information of context into the model.

2 The proposed method

The multiple imaging acquisitions are typically captured via different imaging devices and often have different resolutions. Thus image alignment for all the modalities, which is also called image registration, should be first completed. Here we use a joint photometric and geometric image registration technique [15] to register these images. We concatenate the pixels within a square window in the registered data cube into a vector. By using a kernel function, the vector is projected to a high-dimensional kernel feature space. Next to the two classes: 'paint loss' and 'background', we specify a third class 'crack', which is by art restorers treated differently than larger portions of missing paint called paint loss.

The modified SRC model with respect to sparse coefficients of $\mathbf{x} \in \mathbb{R}^m$ in the projected kernel feature space is

$$\hat{\alpha} = \arg\min_{\alpha} \|\phi(\mathbf{x}) - \phi(\mathbf{D})\alpha\|^2 \;\; s.t. \; \|\alpha\|_0 < K_0, \qquad (1)$$

where $\phi : \mathbb{R}^m \to \mathscr{F} \subset \mathbb{R}^{\hat{m}}$ is an implicit mapping function that projects $\mathbf{x}$ to a higher dimensional space; $\phi(\mathbf{D}) = [\phi(\mathbf{d_1}), \phi(\mathbf{d_2}), ..., \phi(\mathbf{d_N})]$ is the dictionary in the projected space and $\mathbf{d_i} \in \mathbb{R}^m$ $(i = 1, 2, ..., N)$ are the training samples. Once the sparse coefficients are calculated, the class-specific residuals can be computed by

$$\begin{aligned} r_i(\phi(\mathbf{x})) &= \|\phi(\mathbf{x}) - \phi(\mathbf{D_i})\alpha_i\|_2 \\ &= \langle \phi(\mathbf{x}) - \phi(\mathbf{D_i})\alpha_i, \; \phi(\mathbf{x}) - \phi(\mathbf{D_i})\alpha_i \rangle^{1/2} \\ &= (\kappa(\mathbf{x}, \mathbf{x}) - 2\alpha_i^T \mathbf{K_{D_i}} + \alpha_i^T \mathbf{K_{D_i D_i}} \alpha_i)^{1/2}, \qquad (2) \end{aligned}$$

where $\kappa : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is a kernel function defined by $\kappa(\mathbf{x_i}, \mathbf{x_j}) = \langle \phi(\mathbf{x_i}), \phi(\mathbf{x_j}) \rangle$; $\mathbb{K}_{\mathbf{D_i}} \in \mathbb{R}^{N_i}$ is a vector associated with

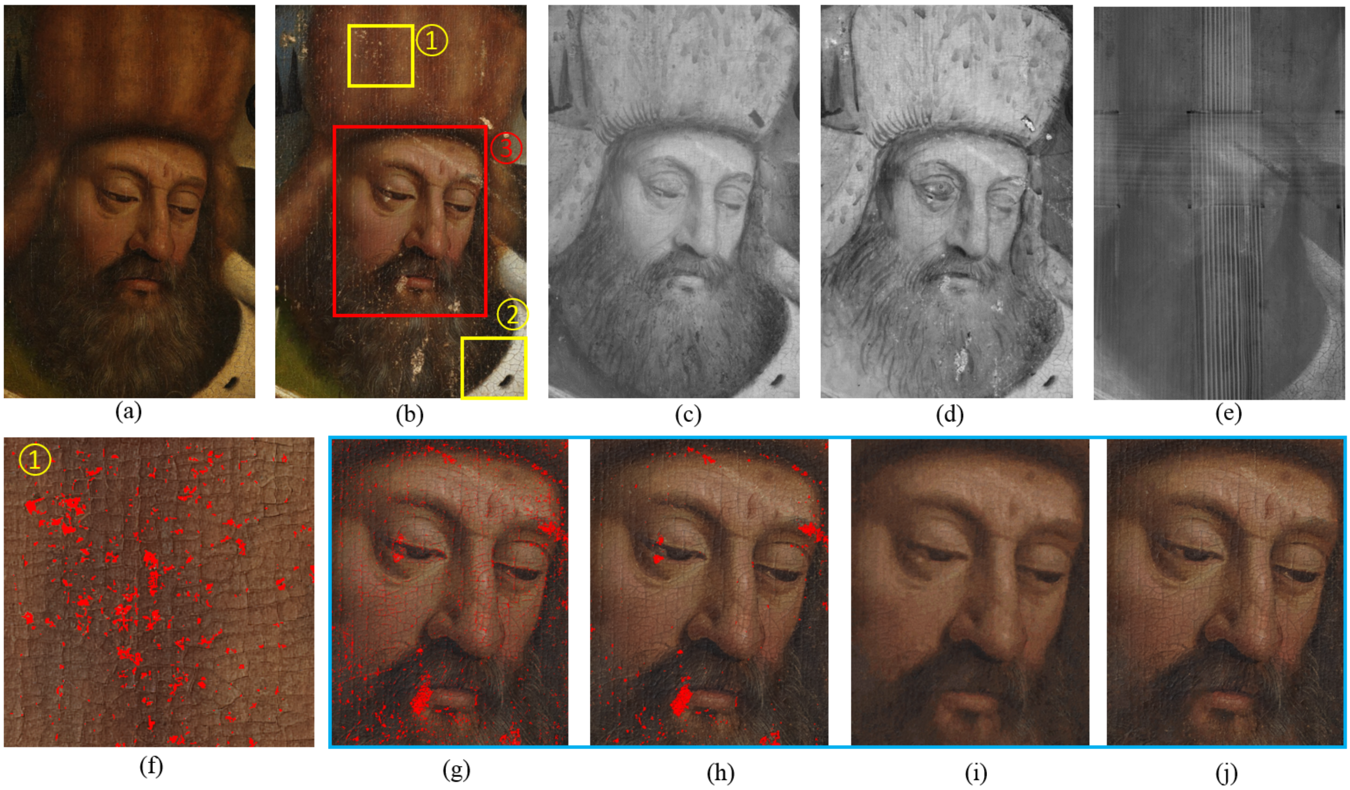

Figure 1: Top row: multiple imaging scans, which include (a) macrophotography before cleaning, (b) macrophotography after cleaning, (c) infrared macrophotography before cleaning, (d) infrared reflectography after cleaning and (e) X-radiography before cleaning. Bottom row: (f) Annotated patch 1 used for training, (g) detection map obtained by applying SRC, (h) detection map obtained by the proposed method, (i) inpainting results using the method of [12] with the SRC map from (g) and (j) inpainting result with the map obtained by our method.

class i in $\mathbb{K}_{\mathbf{D}} \in \mathbb{R}^N = [\kappa(\mathbf{d_i}, \mathbf{x}), \cdots, \kappa(\mathbf{d_N}, \mathbf{x})]^T$; $\mathbb{K}_{\mathbf{D_iD_i}} \in \mathbb{R}^{N_i \times N_i}$ is a matrix corresponding to class i in $\mathbb{K}_{\mathbf{DD}} \in \mathbb{R}^{N \times N}$ with entries $\mathbb{K}_{\mathbf{DD}}(i,j) = \kappa(\mathbf{d_i}, \mathbf{d_j})$ and α_i is a vector associated with class i in α. Then we label the class of a test sample by

$$class(\mathbf{x}) = \arg\min_{i=1,2,3} \; r_i(\phi(\mathbf{x})). \tag{3}$$

We denote by $\mathbf{Map_{crack}}$ the obtained binary crack map. By collecting all the residuals $r_i(\phi(\mathbf{x}_i))$, we form the residual cube. Here we denote by $\mathbf{R} \in \mathbb{R}^{M \times N \times 3}$ the reshaped residual cube, where each layer corresponds to one class.

Typically, paint losses will occupy an area larger than a single pixel. Hence, pixels within a relatively small neighbourhood are likely to belong to the same class and share similar sparse representation coefficients. Therefore we apply a smoothing filter to each layer of the residual cube to make the coefficients of neighbouring pixels similar to each other. In particular, we use for this purpose a weighted least square (WLS) [16] filter. The binary paint loss map, $\mathbf{Map}^{'}$, can be calculated by selecting the smallest smoothed residual. This smoothing has an adverse effect on thin cracks, which tend to be assigned to paint loss (or to background). To solve this, we use the crack map $\mathbf{Map_{crack}}$ generated prior to smoothing, as follows

$$\mathbf{Map} = \mathbf{Map}^{'} \odot \overline{\mathbf{Map_{crack}}}. \tag{4}$$

The training samples in $\mathbf{D}$ of (1) play an important role as they are used to supervise the model to generate the corresponding characteristics of paint loss and background. However, for most cases, compared with the samples of background, the number of paint loss samples is rather small. In addition, accurate annotation on a pixel level is a highly challenging task, which may lead to mislabelled samples. Errors can be caused by blurring in low-resolution images, large transitions and low contrast between target and background, noise, artefacts and so on. To cope with this problem, we suggest a majority voting strategy:

$$identity(\mathbf{x}_j) = \arg\max_{c} \; p_j^c \tag{5}$$

where the fraction $p_j^c = N_j^c / K$ is an empirical probability for the pixel j to belong to the class c. K is the number of simulations and N_j^c the number of times that pixel j was assigned to class $c \in \{Paint\ loss, Other\}$.

3 Results and discussion

We illustrate the detection result on a part of the panel *prophet Zachary*, image patch 3 in Fig. 1 (b). The training samples are from other two image patches in Fig. 1 (b), which were annotated by a painting conservator. Fig. 1 (f) shows one of the annotated image patchs. We set the number of training samples in each class to 80 and K to 10. The imaging modalities in Fig. 1 (a), (b) and (c) are used. Fig. 1 (h) and (j) illustrate paint loss detection results of the proposed approach and virtual inpainting using the detected mask and the inpainting method from [12]. For comparison, we also show the paint loss map in Fig. 1 (g) that is produced by applying the original SRC with multimodal images and majority voting. The corresponding inpainting result is reported in Fig. 1 (i). Obviously the proposed method reduces significantly false detections. Consequently, we avoid previous excessive oversmoothing and undesired removal of cracks during virtual restoration.

References

[1] P. Abry, A. G. Klein, W. A. Sethares, and C. R. Johnson, “Signal processing for art investigation,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 14–16, July 2015.

[2] C. R. Johnson, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang, “Image processing for artist identification,” *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 37–48, July 2008.

[3] L. Platiša, B. Cornelis, T. Ružic, A. Pižurica, A. Dooms, M. Martens, M. D. Mey, and I. Daubechies, “Spatiogram features to characterize pearls in paintings,” in *18th IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 801–804.

[4] B. Cornelis, T. Ružic, E. Gezels, A. Dooms, A. Pižurica, L. Platiša, J. Cornelis, M. Martens, M. D. Mey, and I. Daubechies, “Crack detection and inpainting for virtual restoration of paintings: The case of the ghent altarpiece,” *Signal Processing*, vol. 93, no. 3, pp. 605–619, 2013.

[5] P. Abry, S. G. Roux, H. Wendt, P. Messier, A. G. Klein, N. Tremblay, P. Borgnat, S. Jaffard, B. Vedel, J. Coddington, and L. A. Daffner, “Multiscale anisotropic texture analysis and classification of photographic prints: Art scholarship meets image processing algorithms,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 18–27, 2015.

[6] H. Anwar, S. Zambanini, M. Kampel, and K. Vondrovec, “Ancient coin classification using reverse motif recognition: image-based classification of roman republican coins,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 64–74, 2015.

[7] L. van der Maaten and R. G. Erdmann, “Automatic thread-level canvas analysis: A machine-learning approach to analyzing the canvas of paintings,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 38–45, 2015.

[8] R. Srinivasan, C. Rudolph, and A. K. Roy-Chowdhury, “Computerized face recognition in renaissance portrait art: A quantitative measure for identifying uncertain subjects in ancient portraits,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 85–94, 2015.

[9] B. Cornelis, H. Yang, A. Goodfriend, N. Ocon, J. Lu, and I. Daubechies, “Removal of canvas patterns in digital acquisitions of paintings,” *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 160–171, 2017.

[10] T. Ružic, B. Cornelis, L. Platiša, A. Pižurica, A. Dooms, W. Philips, M. Martens, M. D. Mey, and I. Daubechies, “Virtual restoration of the ghent altarpiece using crack detection and inpainting,” in *Advanced Concepts for Intelligent Vision Systems*, 2011, pp. 417–428.

[11] A. Pižurica, L. Platiša, T. Ružic, B. Cornelis, A. Dooms, M. Martens, H. Dubois, B. Devolder, M. D. Mey, and I. Daubechies, “Digital image processing of the ghent altarpiece: Supporting the painting’s study and conservation treatment,” *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 112–122, 2015.

[12] T. Ružic and A. Pižurica, “Context-aware patch-based image inpainting using markov random field modeling,” *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 444–456, 2015.

[13] S. Huang, W. Liao, H. Zhang, and A. Pižurica, “Paint loss detection in old paintings by sparse representation classification,” in *Proc. ITWIST*, 2016, pp. 62–64.

[14] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, 2009.

[15] H. Q. Luong, B. Goossens, A. Pižurica, and W. Philips, “Joint photometric and geometric image registration in the total least square sense,” *Pattern Recognition Letters*, vol. 32, no. 15, pp. 2061–2067, 2011.

[16] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and detail manipulation,” in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, 2008, p. 67.