Image fusion for art analysis

Barbara Zitová[a] and Miroslav Beneš[a,b] and Jan Blažek[a,b]

[a]Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, Prague, Czech Republic

[b]Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, Prague, Czech Republic

ABSTRACT

Our paper addresses problem of multi-modal data acquisition and the following data visualization for art analysis and interpretation. Various types of modalities for acquisition of digital images are used for art analysis. The data we can obtain using various modalities differ in two ways. The acquired images can differ by their mutual geometry and possibly by their radiometric quality. These are the differences we would like to remove. The group of differences we are interested in are details or characteristics of an artwork, which are apparent just in the certain modality and which bring us new information. The two listed groups represent two categories of image processing methods we have to deal with. The first one is represented by image preprocessing methods such as data enhancement and restoration algorithms, the second class includes effective ways how to combine the acquired information into one image - image fusion. In our paper we present image quality enhancement for microscopic multi-modal data and their segmentation and recent results in data fusion and visualization for art analysis are demonstrated from the second category of methods.

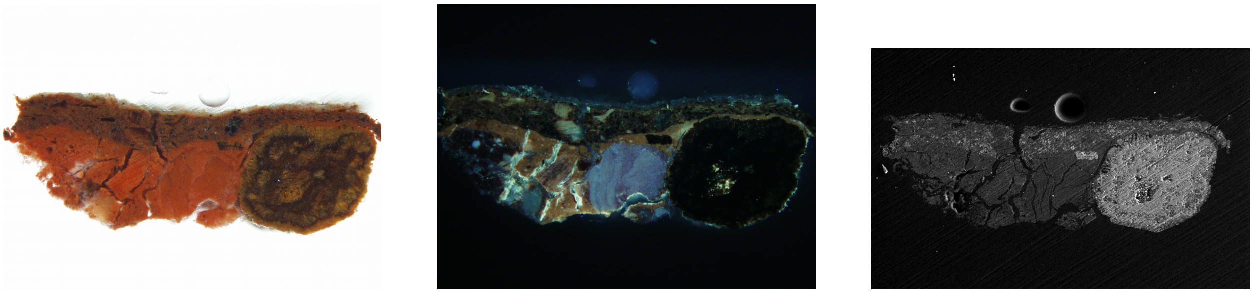
Keywords: Image fusion; visualization; art analysis.

1. INTRODUCTION

Image fusion tries to create an efficient combination of information contained in multiple input images to provide more accurate description of the studied object than any of these individual input images. The one resulting compact image is visualizing all acquired information at once and in this way the end-user has often better possibility to perceive the details not apparent otherwise. The image fusion methods strive to reduce the dimensionality of acquired data while preserving or even stressing details, important for the end-user. Nowadays, the number of possible modalities for image data acquisition is growing (visible spectra, infrared in different wavelength bands, ultraviolet in different wavelength bands, to name the most widely used) and we are able to capture more and more subtle details of objects of interest, of their structure and functionality. The methods for effective combination of revealed information are very important. They enable to perceive paintings which were lost by overpainting (i.e. Patch of Grass by Vincent van Gogh[1]) or see the virtually reconstructed and valuable gems of cultural heritage in their original beauty (Mona Lisa, restored using virtual methods[2]).

Image fusion methods can be found in different application domains such as medical imaging, remote sensing, material science, and surveillance, to name the most important ones. Art analysis can gain a lot from the fusion methodology, combining multispectral data in an effective manner. The multispectral data acquisition itself is widely used for analysis[3] of artworks. Such data sets have to be usually preprocessed to minimize geometrical and radiometrical differences in the acquisition process - different camera setting, different camera viewpoint, different filter sensitivity and noise robustness, etc. We face the problem of multimodal/multitemporal image registration, when data taken within different geometrical coordinates and differing in modality and possibly originating in different times have to be geometrically aligned. The image registration is well studied problem, there are several survey papers[4] on this topic.

Further author information: (Send correspondence to Barbara Zitová)
Barbara Zitová: E-mail: zitova@utia.cas.cz, WWW: zoi.utia.cz
Miroslav Beneš: E-mail: benem9am@utia.cas.cz
Jan Blažek: E-mail: blazek@utia.cas.cz



(a) VIS image (b) UV image (c) SEM image

Figure 1. The cross-section images for artwork material analysis are acquired in three modalities - (a): visible spectrum (VIS); (b): ultraviolet spectrum (UV); (c): scanning electron microscope (SEM)

Removal of radiometrical differences and related image enhancement is discussed in Section 2. Section 3 presents segmentation methodology for microscopic cross-sections, Section 4 demonstrates results of image fusion in three different projects. Finally, Section 5 concludes our paper.

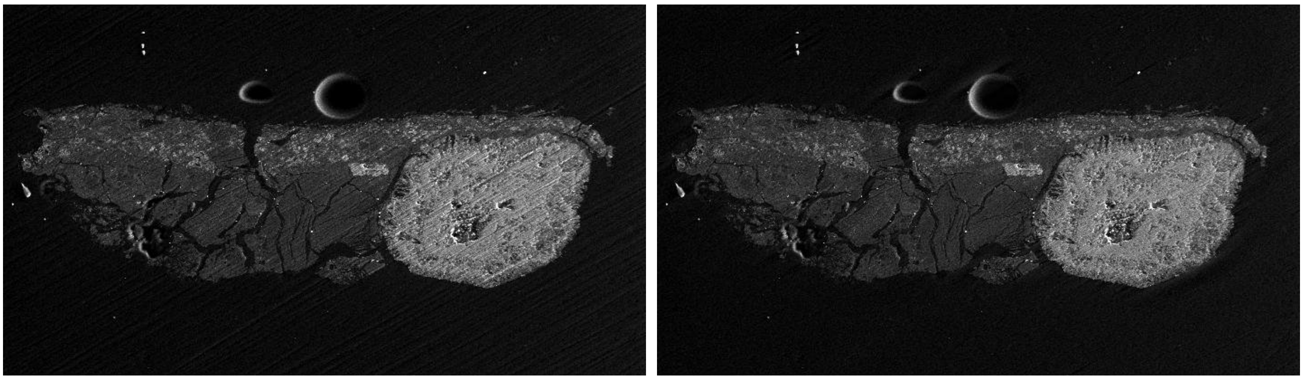
Image enhancement

Image enhancement tries to improve the quality of the processed image so further processing can bring better results. Diminishing of radiometrical differences such as noise or blur, removal of various acquisition artifacts - these algorithms are examples of image enhancement methods. First method introduced here is related to material research of image artwork. Art conservators are taking several minute surface samples from the studied artwork. The samples are taken off at the selected areas to be able to capture material composition of the artwork and they are scanned using several modalities – visible spectrum (VIS), ultraviolet spectrum (UV) and images from scanning electron microscope (SEM) (see Fig.1).

During the preparation for scanning, the samples are embedded in a polyester resin and grounded at a right angle to the surface plane to expose individual painting layers. Grinding of the resin produces noisy artifacts in form of parallel lines (see Fig.2) on the cutting surface due to the mechanical properties of the material. They are affecting image quality of both the background and the studied sample (the foreground) and in this way they are negatively influencing further image processing. Moreover, the cutting pattern cannot be easily removed mechanically by material polishing, because this operation could negatively influence the surface quality of the studied sample due to friction and consecutive temperature increase. Thus, the multimodal input data have to be preprocessed before any higher level image processing can be performed (e.g. image retrieval, material classification).

The proposed method removes the artifacts completely or at least significantly diminishes them and thus improves the performance of image processing. As it was stated, the artifacts are omnipresent and in form of parallel lines. If Fourier transform is applied on the input image, considerable amplitude response can be observed in the direction perpendicular to the artifact lines. Thus, masking the Fourier spectrum and applying inverse Fourier transform can provide required result - removal of the artifacts in image domain.

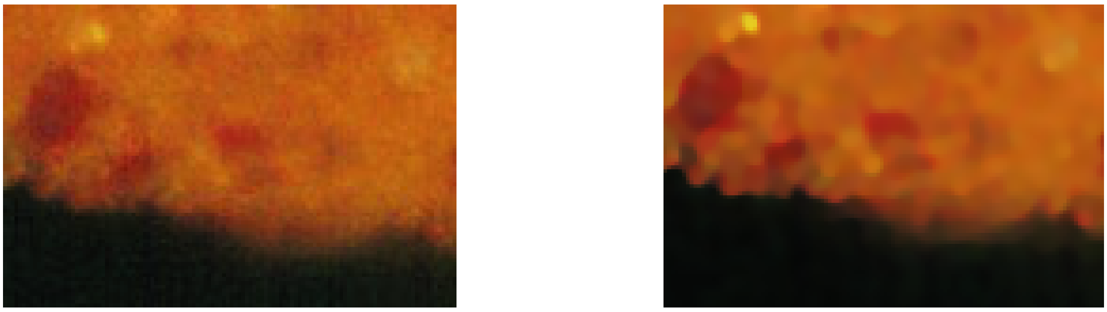
The method is based on the construction of an appropriate directional masking filter corresponding to the direction of these parallel lines. The first step consists of the direction estimation. We exploit the idea that the orientation of the artifact response should corresponds to a distinct peak of the column sum function. This auxiliary function is constructed by consecutive stepwise rotation within the interval $< 0^\circ, 180^\circ >$ of the Fourier spectrum with the step equals to 1°. In each step, the values in columns are summed up and the maximum of the column sums is taken. The sum function is smoothed to inhibit impact of distortions and peaks are found automatically. The directional masking filter is then created in the determined direction and the inverse Fourier transform of the filtered data is computed. The achievable result are shown in Fig.2, where the right image demonstrates the diminishing of the cutting regular pattern. The consequences for further image preprocessing will be shown in the next Section.



(a) SEM image before enhancement (b) SEM image after enhancement

Figure 2. Example of an enhanced SEM image. (a) Original SEM image, the cutting pattern - parallel lines at 45 ° angle - is apparent. (b) Enhanced filtered image - besides the improvement in the background, the texture of cross-section is as well enhanced.

Another frequently present imperfection of the image data is out of focus blur due to the wrong setting of the acquisition camera. This type of radiometrical degradation can be removed by an appropriate deconvolution technique. We expect the model of the present blur to be in the form of Gaussian function and by application of the method described in[5] sufficient improvement for further segmentation or other image processing can be achieved (see Fig.3). Transitions between materials are now much sharper and the boundary of the sample is more distinct, too.



(a) VIS image before enhancement (b) VIS image after enhancement

Figure 3. Example of the deconvolved VIS image. (a) Original VIS image with out of focus blur. (b) Enhanced VIS image with sharper material transitions, more suitable for further segmentation. Consulting M. Sorel

Image segmentation

The aim of image segmentation, another image preprocessing method, is to divide the input image to several regions, which share similar characteristic (e.g. color, texture). One of the subtasks may be image binarization, i.e. division of the input image to two different classes. Therefore one of the two labels is assigned to each pixel in the picture. In our case it is either background or foreground label, because the cross-section has to be isolated from the noisy background, which is formed by the polyester resin.

Due to the distinctive characteristics of the microscopic cross-sections and their polyester resin background the segmentation is difficult. Dozen of binarization algorithm was tested, but only few proved to give satisfactory results. The variance of the algorithms performance can be observed in Fig.4, where binary results of three different algorithms on two different SEM images are presented. Huang[6] algorithm uses fuzzy thresholding method. Intermodes[7] computes the threshold from smoothed histogram and IsoData[8] is an iterative selection method based on ISODATA algorithm. In the top row (sample A) the binary masks are almost identical. On the other hand in the bottom row (sample B) the outputs of the algorithms differ quite radically.

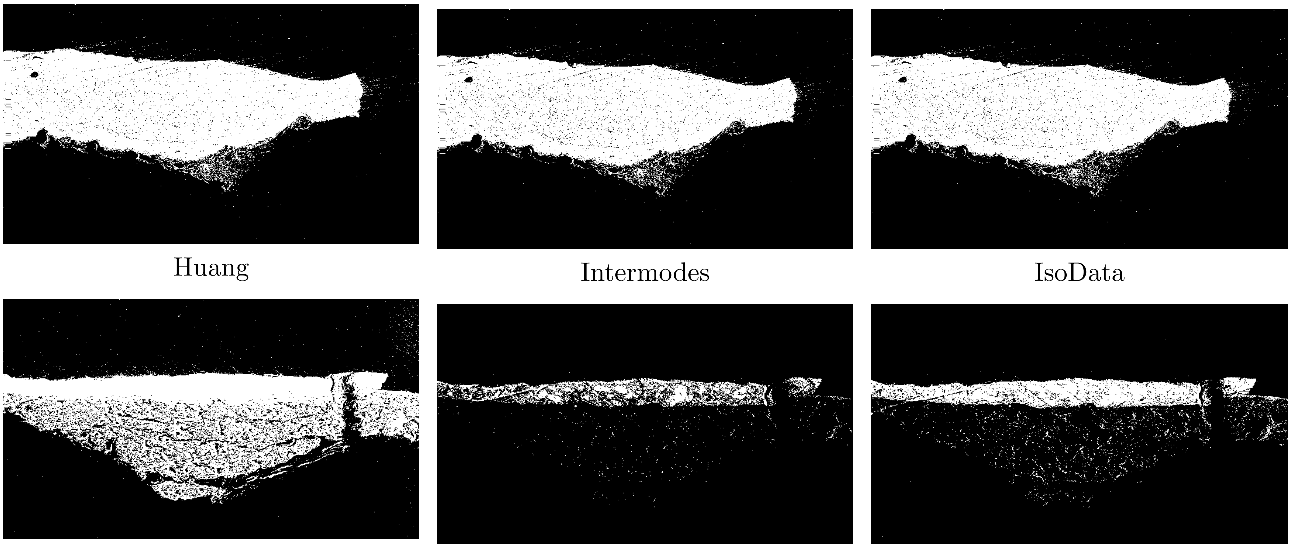


Figure 4. Performance variation of three different binarization algorithms (Huang,[6] Intermodes,[7] and IsoData[8]) on two different SEM images. There are almost identical results in the top row (sample A) and very different results of the algorithms in the bottom row (sample B).

Fig.5 illustrates the impact of the background artifacts on outcome of various algorithms. In this case the method by Triangle[9] was applied. The left image is the original SEM image with the grinding artifacts (parallel lines). In the middle there is a result of binarization algorithm, which was applied to the original SEM image. Finally, the right image shows the result of the binarization algorithm applied to the preprocessed SEM image, after the artifacts were removed using the introduced image enhancement technique. The improvement in the resulting foreground/background segmentation is apparent.

Figure 5. The influence of the background grinding artifacts. The original SEM image (left) is either directly binarized (middle) or the grinding artifacts are firstly diminished using the introduced enhancement method and then binarized (right).

Exhaustive testing of binarization methods was realized in order to determine suitable image segmentation method for each modality (VIS, UV and SEM). In all three cases the Huang algorithm[6] seems to be the most promising. The quality of its output corresponds the most with the needs of the next steps of image processing workflow.

Image fusion

As it was stated the aim of image fusion is to find such combination of input images that all the important and interesting details stay preserved while the dimensionality of the data is reduced. Final output image is much easier to comprehend than the original data stream for human understanding. The fusion algorithms can work

(a) Visible image

(b) Infrared image

(c) Fused image

(a) Fused image - detail

Figure 6. Mary Magdalene painting. The input images for image fusion: (a) visual spectra image; (b) infrared spectra image. The fused image (c) and the detailed version: the triangle structure above the forehead (d).

pixel-wise, based on the mapping of the input pixel values to the output information and not taking into account the image structure and the content (good example is well-known alpha blending). The second category of fusion methods are based on the multiscale decomposition, often using wavelet decomposition applied on input images and followed by particular decision criterion for the output data formation.[10] The low- and high-frequency data drive the fusion in these cases. The last group is based on higher level information about an object of interest, so-called object based image fusion. Detected objects from one data channel are here substituted into the other data channel to form the multimodal representation. The special category of image fusion methods is formed by biologically motivated approaches based on human vision system.

In the art analysis, there are several possibilities how the input data entering the fusion process should be interpreted and as a result what type of image fusion is the most appropriate. In this paper we will address three examples of them on the projects we were involved in.

In the first one, the data set consists of the multispectral images of the Mary Magdalene painting. The images were acquired in the visible spectra and in the infrared spectra IR (an infrared-enabled camera with Hoya R72 filter, 720-1200 nm). Fig.6 shows the intensity images in the visible (a) and in the infrared (b) wavelengths. The painting itself was heavily damaged so both the visual and the computer based evaluation was extremely difficult. Visual inspection uncovered differences in the underdrawing sketches and the surface painting. An

(a) Slovak version (b) Croatien version (c) Fused image

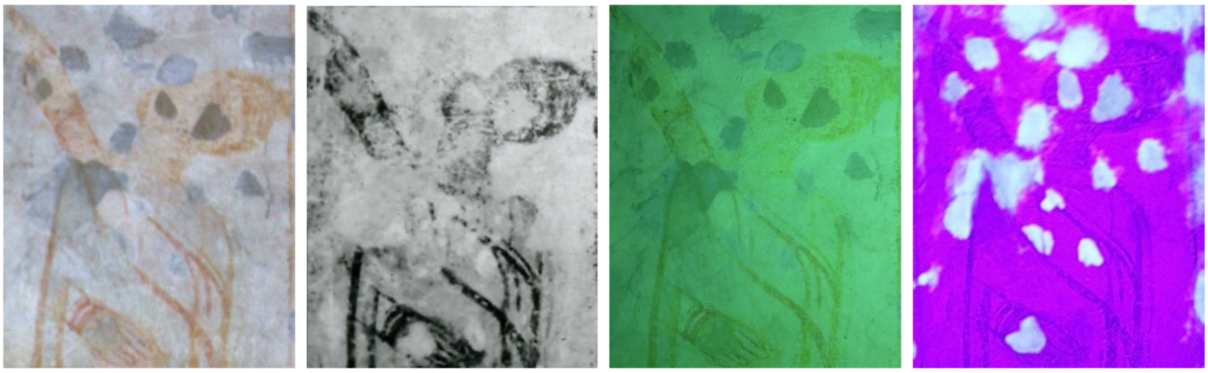
Figure 7. The oil paintings of Marianne Norman Earl Ehrenfels by F. Amerling. Two versions from the Slovak Republic (a) and from the Croatia (b), respectively. The resulting fused image - the more bright intensity represents the more prominent difference.

underdrawing is a preparatory drawing for a painting sketched directly on the ground. It can be found in many artworks, typically outlined with charcoal, but artists have also been known to use chalk, pencil or paint. The underdrawings are later covered with the artists medium. The drafts vary in detail from artist to artist, from simple perspective outlines to detailed sketches. Therefore, the underdrawing can provide much insight into the creative process of an individual artist. Often image acquisition in the infrared modality is used for underdrawing visualization, because the infrared waves are often capable of passing through the paint layer where the visible radiation is reflected by it.

For the image fusion, we applied contrast preserving color fusion,[11] which is a representative of the pixel-based color fusion and tries to preserve intensity changes in the input images equally, independently of the other modality. Here, the 2D rectangular color mapping scheme used for image fusion is created in such a way, that color differences as perceived by the human eye in all points are nearly the same. Fig. 6(c) shows the fused image, created from the visible and IR input images, using the blue-red color space. These part, which are depicted in almost white color, are common in both modalities, while the red and blue parts correspond to the visual and IR modality, respectively.

On the Mary Magdalene painting the IR image shows the difference in the mouth area, detected during visual inspection. While the original IR sketch (Fig. 6(b)) shows the smile on the woman face, the final painting (Fig. 6(a)) much more resembles sad, serious facial expression. However, this finding, done visually, is an example of situation, when the computer assisted approach failed. While people are trained in recognition of human expressions and moods, for automatic image processing the ratio of the difference we are interested in with respect to the painting scratches, lacunas, ageing, noise, etc. speaks for the negative influence of the painting damage, disabling any automatic analysis. We failed to detect automatically this facial difference. On the other hand, there was hidden other detail, where on the contrary the human perception failed.

Fig. 6(d) shows the close-up of the face of the lady. As it was stated, white lines correspond to sketches which are common for both modalities and red and blue color represent differences. The interesting part is the one just above the forehead. There are three line-like structures, formed into the triangle, representing the scarf folding. Such similar structures are apparent in the white color, located just below. The visual inspection is able to perceive these structures in the individual input images (see Fig. 6 (a) and (b)), however it is difficult to judge if they are located at the same place or not. The fused image clearly indicates the difference in geometry of this triangle structure in visible and IR domain.



(a) Input images [VIS, BWS, UVB, UVN]

(b) Fresco

(b) Fused image

Figure 8. The fresco, located in Kostolany pod Tribečom, the Slovak Republic. (a) - the visible spectrum (VS) image; the old black and white photograph (BWS) from the 60s; the ultraviolet broad-band spectrum (UVB) image; the ultraviolet narrow band spectrum (UVN) image. (b) - the fresco in the church. (d) - the fused image with more profound details. Consulting R. Furbach.

The second example how the goal of the image fusion can be seen is based on the comparison of two different paintings.[12] In Fig. 7 you can see two oil paintings created by F. Amerling in 1851. The painting (a) is from the collection of Spis Muzeum in Levoča, Slovak Republic, while the (b) painting belongs to the Osijek Art Gallery in Osijek, Croatia. Both paintings represent Marianne Norman Earl Ehrenfels. The Slovak painting was not signed, the authorship was decided later. In the list of F. Amerling's painting by G. Probszt is mentioned the Croatien version. In both portraits there are the rose in the hair and small gold earring. The identity of the depicted women is very probable, however not proven. They are very similar, but the dresses and background are not identical.

There are theories the paintings were created in different times. One theory by B. Balen claims that the Slovak version is the copy of the Croatien one. However, this would imply that the copy outperforms the original. The Slovak version is more relaxed, its details are more elaborated (the rose, the collar). The second theory by J. Hradilova et al. claims that the Slovak version is the earlier version, even that the time difference was small. After the image registration of the two images we fused them to obtain the difference between the images. The squared difference of the normalized images gives us the resulting image (see Fig. 7(c)), with impressive level of the congruency. The fused image, where the more bright intensity represents the more prominent difference, prove the theory of the single depicted persons and, moreover, the differences encourage the idea of the Slovak original - the differences, visualized in more exact form than by their description, are exactly as expected, referring to the more smooth and not so elaborated and detailed version from the Croatia (the rose, the neck, the nose proportion).

Finally, the last application of image fusion will be presented on the wall fresco data.[13] The aim here is similar as in the first example, however the number of used modalities is higher so the source combination is even more complicated for the human observer. We would like to fuse all data into one detailed image, capturing all important aspects from all information sources.

The fresco (Fig. 8(b)) is situated in the 11th century church in Kostolany pod Tribečom, a small village in the Slovak Republic. The fresco covers large area of the inner church walls and it is heavily damaged. Restoration works have been started recently. The fresco images were acquired in the visible spectrum (VS), in the ultraviolet broad-band spectrum (UVB), and in the ultraviolet narrow band spectrum (UVN) with the spectrum maxima at 368 nm, see Fig.8 (a). Next to this, we have black and white photographs (BWS) of the fresco from the 60s (see Fig.8 (a)), which captures the fresco before aging changes during last fifty years (this channel was not used for the image fusion). All the recent image data were taken by the Czech-Geo company.

The main idea of our approach is to use principle component analysis (PCA) for the normalization of each modality channel. The resulting vectors are then averaged across the adaptively normalized channels. The resulting output image is generated from the averaged vectors by means of the neural network without hidden layers with three input and three output neurons. The neural network is trained on the combination of the input VIS image and resulting output data. In Fig. 8(c) the resulting fused image is presented. The line-like artifacts at the top right corner comes from the image registration (all the input channels had to be geometrically aligned before the fusion itself and not all modality images covers this part of the scene). The quality of the fresco is not improved, we did not used any higher level processing here, however the details from all channels are visible now at once, the details are more apparent and profound (see the cap of the person).

CONCLUSIONS

Our paper addresses problem of multi-modal data acquisition and following data visualization for art analysis and interpretation. The methodology for preprocessing the microscopic cross-section images, taken in visible and ultraviolet spectra and by means of the scanning electron microscope was presented, consisting of the image enhancement using the frequency representation of images, of the image deconvolution, and finally of image segmentation. In the second part of the paper three different image fusion techniques were demonstrated on the multispectral images and on the images acquired from different paintings, respectively. The proposed algorithms were applied on data sets from ALMA, a joint workplace of the Academy of Fine Arts in Prague and the Institute of Inorganic Chemistry of the Academy of Sciences of the Czech Republic, and from the Hamilton Kerr Institute (Dr. Spike Bucklow), University of Cambridge, United Kingdom.

ACKNOWLEDGEMENTS

The work was supported by the Grant Agency of the Czech Republic under the project No. 102/08/1593, by the Academy of Sciences of the Czech Republic under the project M100750901(financial support for international cooperation), and by the the Ministry of Education of the Czech Republic under the project No. 1M0572 (research center).

REFERENCES

[1] Hughes, S., Daubechies, I., Brasoveanu, A., Duarte, M., and Dik, J., "Underpainting Recovery with Xray Synchrotron Imaging and Image Processing," *Proceedings of The Third International Workshop on Image Processing for Art Investigation IP4AI3, USA* , 15–16 (2010).

[2] Elias, M. and Cotte, P., "Multispectral camera and radiative transfer equation used to depict Leonardo's sfumato in Mona Lisa," *Appl. Opt.* **47**, 2146–2154 (2008).

[3] Pelagotti, A., Mastio, A. D., Rosa, A. D., and Piva, A., "Multispectral imaging of paintings," *Signal Processing Magazine IEEE* **25**, 27 – 36 (2008).

[4] Zitova, B. and Flusser, J., "Image registration methods: a survey," *Image and Vision Computing* **21**(11), 977–1000 (2003).

[5] Chambolle, A. and Lions, P., "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**(2), 167188 (1997).

[6] Huang, L.-K. and Wang, M.-J., "Image thresholding by minimizing the measures of fuzziness," *Pattern Recognition* **28**(1), 41–51 (1995).

[7] Prewitt, J. and Mendelsohn, M., "The analysis of cel images," *Annals of the New York Academy of Sciences* **128**, 1035–1053 (1966).

[8] Ridler, T. and Calvard, S., "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems, Man and Cybernetics* **8**, 630–632 (1978).

[9] Zack, G. and Rogers, W., "Automatic measurement of sister chromatid exchange frequency," *Journal of Histochemistry and Cytochemistry* **25**(7), 741–753 (1977).

[10] Li, H., Manjunath, B. S., and Mitra, S. K., "Multisensor image fusion using the wavelet transform," *CVGIP: Graphical Model and Image Processing* **57**(3), 235–245 (1995).

[11] Kamenicky, J. and Zitova, B., "Contrast Preserving Color Fusion," *Proceedings of SPIE IS T/SPIE Electronic Imaging 2011* , 1–6 (2011).

[12] Hradilova, J. and Mislerova, H., "Technika malby portretu 19.stoleti od Friedricha Von Amerlinga," *Acta Artis Academica* , 140–154 (2010).

[13] Blazek, J., Zitova, B., Benes, M., and Hradilova, J., "Fresco restoration: digital image processing approach," *Proceedings of the 17th European Signal Processing Conference EUSIPCO 2009* , 1–5 (2009).