IMAGE PROCESSING IN MATERIAL ANALYSES OF ARTWORKS

Miroslav Beneš
Department of Software Engineering, Charles University, Prague, Czech Republic

Barbara Zitová
Institute of Information Theory and Automation, Academy of Sciences of Czech Republic, Prague, Czech Republic

Janka Hradilová
Academic Laboratory of Materials Research of Paintings, Academy of Fine Arts, Prague, Czech Republic

David Hradil
Institute of Inorganic Chemistry, Academy of Sciences of Czech Republic, Řež, Czech Republic

Keywords: Cultural heritage, Image retrieval, Descriptors, Classification.

Abstract: In this paper we present a system for processing, description and archiving material analyses used during art restoration - *Nephele*. The aim of the material analyses of painting layers is to identify inorganic and organic compounds using microanalytical methods, and to describe stratigraphy and morphology of layers. The results are used to interpret the applied painting technique. The *Nephele* system is a database system for material analysis reports, extended with image preprocessing modules and image retrieval facility. The implemented digital image processing methods are image registration, layers segmentation, and grains segmentation. In the archiving part of the *Nephele*, in addition to traditional database functions we have incorporated image-based retrieval methods into the developed system. They are based on feature descriptions such as the Haralick descriptors of co-occurrence matrices. The presented examples of achieved results show the applicability of the system.

1 INTRODUCTION

The image processing methods play an important role in very distant application areas such as art restoration. These algorithms are a useful tool for restorers due to their ability to improve quality and interpretability of the input data obtained from restored artworks. Painting materials research, which helps make choice of the proper materials for the very restoration, is the field where our proposed system - *Nephele* (Fig. 1) - tries to facilitate the work of restorers.

Each painting materials analysis is precisely described in the form of a report, which contains general information about the artwork as well as description and results of the analyses which were held. Reports database could serve as a knowledge base for further restoration cases. Moreover, based on obtained data and experiences, new analytical methods and description styles can be found and proposed.

Our work is mainly aimed at helping with proper identification of pigments and binders in color layers, where the layer is defined as a consistent and distinguishable part of the painting profile. Such classification gives important information about the age of the used paints and their possible place of origin. The infrared reflectography, which is the most popular method in the area of cultural heritage for color layers identification, is not suitable for the purposes of stratigraphy (learning about layers). Stratigraphy is usually studied in the visible spectrum (VS) (Fig. 2(a)), in the ultraviolet spectrum (UV) (Fig. 2(b)), and by means of the scanning electron microscopy (SEM). These input data can be misaligned due to the changing conditions. Therefore, a method for removal of geometrical differences between the VS and UV images is incorporated in the proposed system. Apart from this, creation of prelim-

Beneš M., Zitová B., Hradilová J. and Hradil D. (2008).
IMAGE PROCESSING IN MATERIAL ANALYSES OF ARTWORKS.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 521-524
DOI: 10.5220/0001079705210524

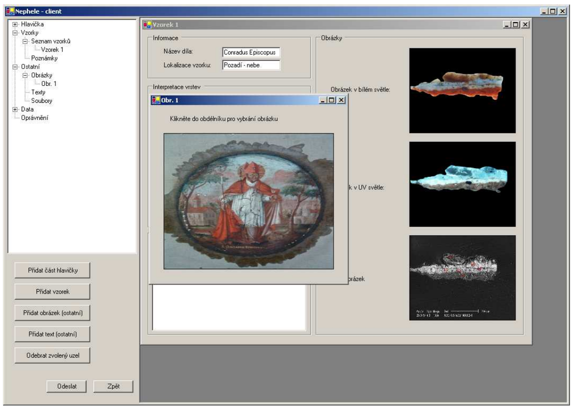
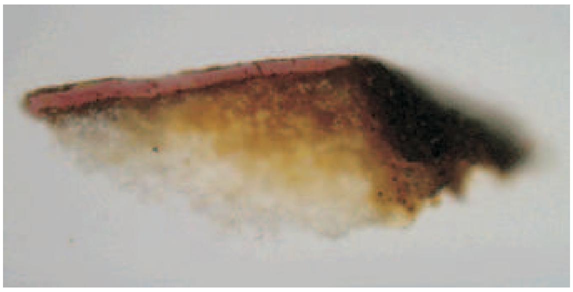
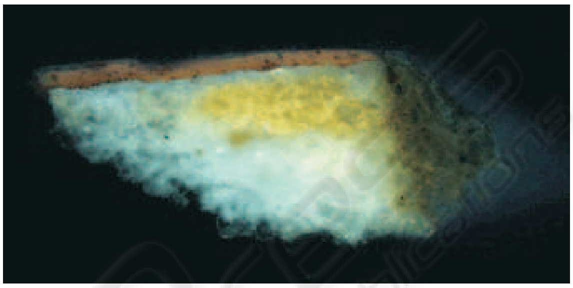


Figure 1: Illustrative example of the main window of the *Nephele.* user interface. The restored artwork is shown, together with the VS, UV and SEM images of one microscopic specimen.



(a)



(b)

Figure 2: The images of the artwork specimen in the visible (a)-top and the ultraviolet (b)-bottom spectra. The single color layers are apparent, especially on the VS (a) image.

inary color layer segmentation and grains segmentation were implemented for further processing.

The second part of our contribution consists of the efficient data retrieval. For such painting materials research reports database, the look-up of the archived reports based only on the text information is often not enough. The ability to fetch reports which describe visually similar specimens/materials can increase the helpfulness of the database. We have incorporated content-based image retrieval methods into the developed system.

The image preprocessing part of the proposed system is described in Section 2, consisting of the image registration module and image segmentation module. Section 3 introduces the content-based image retrieval included in *Nephele*. Section 4 introduces the grains segmentation together with our future plans in this area.

2 IMAGE PREPROCESSING

Stratigraphy, which helps to identify pigments and binders in color layers, is usually studied in the VS, UV, and SEM images (Figs. 2). The analysis works with minute surface samples (0.3mm in diameter) from the selected areas of the artwork. The samples are embedded in a polyester resin and grounded at a right angle to the surface plane to expose the layers. The VS and UV image information is then combined and the final estimate of color layer borders is created, based on the image data and the experience of the experts (possible order and combination of materials for specific artworks, time period, area, etc.). The SEM images can bring even more precise information about the layer structure, however, they are not available for all cases.

During the UV and VS image acquisition process the VS and UV image pairs of the sample are often geometrically misaligned due to the manipulation errors etc. This difference has to be removed before the analysis to be able to compare the corresponding structures in the images. Up to now it used to be done manually by an operator. The proposed image registration module of the system solves the spatial alignment of the image pairs.

Mutual information (MI), originating in the information theory, is a recognized solution for the multimodal registration problem, where images of the same scene are acquired by different sensors. It is a measure of the statistical dependency between two data sets. The main reason for choosing MI was that it does not impose strong limitations on used sensors. One of the first articles proposing this technique is the one by Viola and Wells (Viola and Wells, 1997).

In our approach, we use the speed up of the method, based on the averaging pyramid together with the discrete estimate of histogram. The optimization of the maxima location is a modified version of the method published in (Penney et al., 1998). Moreover, we use the one-channel data, either green channel of the RGB image representation or the first element of the principal component transform (PCT), to reduce the dimensionality of the problem.

After the image rectification, the color layers can be estimated. The segmentation module performs only preliminary segmentation based on both the VS and UV images, because the construction of the full segmentation is a complex task where usually expert knowledge is necessary (certain materials cannot be neighbors, others are always together, etc.). This could be solved by a proper expert system, but this exceeds the topic of our research.

The proposed method is based on cluster analysis using the set of three RGB channels of the VS and three RGB channels of UV specimen images plus spatial information (x and y coordinates included as another two channels). It starts by an iterative *k-means* clustering. The number of classes is set a priori as a maximum expected number of layers by the user. More complex approaches based on texture analysis or other higher level information could bring slightly better results, nevertheless, without expert knowledge the segmentation still remains preliminary.

Often, relatively smooth transitions from one layer to the other produced ragged borders. The first improvement of consistency was achieved by including spatial coordinates to the segmentation feature space. Even better results were obtained after applying morphological operators to detected segmentation and performing a minimum class size check. Further application of segmentation techniques is mentioned in the Section 4.

3 IMAGE RETRIEVAL FACILITIES

The material analysis reports are often used as a knowledge base for consequent restorations. For such usage, it is very important to have effective tools to look-up the relevant reports. One of the possible extensions of the usual database functionality is to make use of the similarity between images contained in reports. The visual image similarity can imply that the used technique/materials on the analyzed artwork is the same/similar as in the archived report or that it can point to the same author, therefore such information can be very relevant. Thus, the image-based data retrieval is often used nowadays beside the traditional text-based search in database systems. The database entries containing images are looked up according to the image similarity to the query image. The so-called *content-based image retrieval* (CBIR) has become very popular recently (Veltkamp and Tanase, 2000) and is used as a part of multimedia systems in art galleries (Addis et al., 2003; Goodall et al., 2004).

In our *Nephele* system, the image-based data querying exploits the VS and UV images of specimens. The similarity of specimens is not based upon specific shape or structure elements which are the results of the random process of sample cut-off. For the considered methods of image retrieval the color and texture characteristics were chosen as the main features.

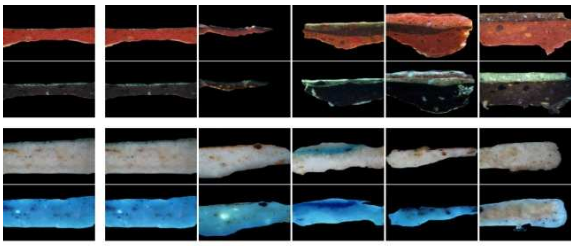


Figure 3: Results of the image retrieval. Left column contains query specimens, next columns in the corresponding rows are results of the retrieval in order of similarity.

The retrieval is based on color features and *co-occurrence matrices* (Haralick et al., 1973). They reflect the joint probability of the occurrence of grey level pairs of two pixels with a defined spatial relationship, formed by a shape operator. The used shape operators were up to two pixels long and all color channels were processed separately. Based on preliminary experiments the four Haralick descriptors were computed from the co-occurrence matrices (Contrast, Inverse difference moment, Entropy, Variance) (Haralick et al., 1973). Apart from them, the color descriptors were included, too, to reflect the main color trends in the data. The image average color and the spectral standard deviation were chosen. Moreover, the R^*-tree indexing structure (Beckmann et al., 1990) with weighted Euclidean metric was implemented to speed-up the retrieval.

The applicability of the method is presented in Fig. 3. There are query images (leftmost column) together with the most similar responses (in the respective rows, in order of similarity from left to right). The visual similarity of the specimens in rows is apparent. The further evaluation of the retrieval would not be statistically significant, because the results are not easy to quantify and the set of samples is rather small.

4 LAYERS DESCRIPTION

The lately included part of the *Nephele* system should lead towards the description module for material characterization. Based on the demand of material scientists, we intend to offer the possibility of the layer description by means of the selected set of features. Such analyses are able to better define used materials,

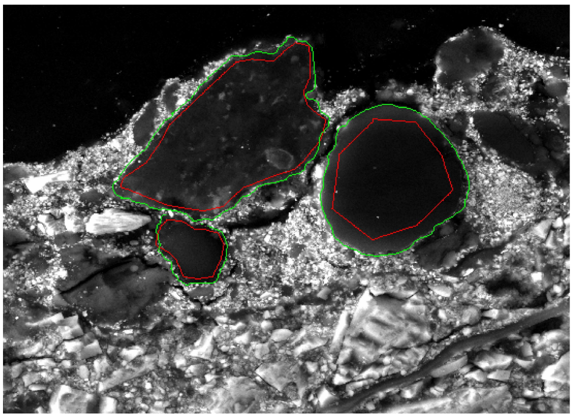


Figure 4: Example of grains segmentation. Green border represents the resulting segmentation.

they will improve the ability to uncover the authors of the artwork by revealing the characteristics of their work with pigments, binders and other components. One of the first necessary step towards is the segmentation of the grains in the single layers. The best data source for such task is the SEM data, where individual grains can be most clearly distinguished. Our algorithmic solution of the grain segmentation is based on the Parametric Snakes (Xu and Prince, 1998). Fig. 4 shows the example of segmented grains in a SEM image.

5 CONCLUSIONS

The proposed system *Nephele* can facilitate the work of material scientists and consequently restorers and offer them a better access to the archived reports they use. To our knowledge, no other similar system has been published up to now. The introduced digital image processing methods enable acquired data preprocessing for further analyses as well as improve of querying above the reports database. The preprocessing of the VS and UV specimen images, used for the identification of pigment and binder present in the artwork, consists of image registration, which makes use of the mutual information approach, and segmentation technique based on the modified k-means clustering. The included image retrieval system is able to provide fetching of reports with visually similar specimen data. The image retrieval is built upon the VS and UV images of the specimens. They are represented using the Haralick descriptors of cooccurrence matrices together with the color descriptors. In the future, single layers will be characterized by means of the selected sets of features for a better definition of the used materials. Recently, the first step - grains segmentation - was implemented, based on the parametric snakes model. The presented examples of achieved results show the applicability of the system.

ACKNOWLEDGEMENTS

This work was partially supported by the Grant Agency of Charles University under the project No. 72507, partially by the Ministry of Education of the Czech Republic under the projects No. 1M0572 (research center) and No. MSM6046144603, and finally by the Grant Agency of the Czech Republic under the project No. 203/07/1324.

REFERENCES

Addis, M., Boniface, M., Goodall, S., Grimwood, P., Kim, S., Lewis, P., Martinez, K., and Stevenson, A. (2003). Integrated image content and metadata search and retrieval across multiple databases. *LNCS*, 2728:91–100.

Beckmann, N., Kriegel, H.-P., Schneider, R., and Seeger, B. (1990). The R^*-tree: An efficient and robust access method for points and rectangles. In *1990 ACM SIGMOD International Conference on Management of Data*, volume 19, pages 322–331.

Goodall, S., Lewis, P. H., Martinez, K., Sinclair, P., Giorgini, F., Addis, M., Boniface, M., Lahanier, C., and Stevenson, J. (2004). Sculpteur: Multimedia retrieval for museums. *LNCS*, 3115:638–646.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621.

Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., and Hawkes, D. J. (1998). A comparison of similarity measures for use in 2D–3D medical image registration. *IEEE Trans. Medical Imaging*, 17:586–595.

Veltkamp, R. C. and Tanase, M. (2000). Content-based image retrieval systems: A survey. *Technical Report*, UU-CS-2000-34.

Viola, P. and Wells, W. M. (1997). Alignment by maximization of mutual information. *Int'l. Journal of Computer Vision*, 24:137–154.

Xu, C. and Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Trans. on Image Processing*, 7(3):359–369.