PHYSICAL OPTICS

Conference materials
UDC 621. 397
DOI: https://doi.org/10.18721/JPM.173.240

Identification of unreadable marginalia by means of hyperspectral imaging: case study of the Ostrog Bible from the Library of Russian Academy of Sciences and Russian National Library

P.S. Baranov[1], V.A. Parfenov[1], S.L. Rongonen[2], N.M. Sokolov[1]✉

[1] St. Petersburg Electrotechnical University "LETI", St. Petersburg, Russia;

[2] Russian Academy of Science Archive, St. Petersburg, Russia

✉ niksokolov2001@mail.ru

Abstract. Experiments on visualization of unreadable manuscript marginalia (fading of ink, spreading of ink, crossed out ink, scraping of ink, gluing with restoration materials) of several copies of the Ostrog Bibles from the National Library of Russia and the Library of Russian Academy of Sciences using hyperspectral imaging have been conducted. The aim of work was to find a way using advanced opto-electronic techniques to read invisible marginalia. In experiments NIR hyperspectral camera operating in the range 400–1100 nm and original self-developed software were used. Marginalia have been visualized by means of hyperspectral imaging that made possible to recognize badly visible texts and bring new information for art historians.

Keywords: hyperspectral imaging, unreadable marginalia, Ostrog bible, manuscript, cultural heritage, opto-electronic techniques

Citation: Baranov P.S., Parfenov V.A., Rongonen S.L., Sokolov N.M., Identification of unreadable marginalia by means of hyperspectral imaging: case study of the Ostrog Bible from the Library of Russian Academy of Sciences and Russian National Library, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (3.2) (2024) 201–206. DOI: https://doi.org/10.18721/JPM.173.240

Материалы конференции
УДК 621. 397
DOI: https://doi.org/10.18721/JPM.173.240

Распознавание нечитаемых маргиналий с помощью гиперспектральной съемки на примере экземпляров Острожской библии из библиотеки Российской Академии Наук и Российской национальной библиотеки

П.С. Баранов[1], В.А. Парфенов[1], С.Л. Ронгонен[2], Н.М. Соколов[1]✉

[1] Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-Петербург, Россия;

[2] Архив Российской академии наук, Санкт-Петербург, Россия

✉niksokolov2001@mail.ru

Аннотация. Представлены результаты экспериментов по визуализации нечитаемых рукописных маргиналий (выцветание, растекание чернил, зачеркивание чернил, соскабливание чернил, заклейка реставрационными материалами) нескольких экземпляров Острожской библии из фондов Российской Национальной библиотеки и Библиотеки Российской Академии наук с использованием гиперспектральной съемки. Целью работы был поиск эффективных способов визуализации нечитаемых маргиналий с помощью инновационных оптико-электронных методов. В экспериментах использован гиперспектральный комплекс с линейным сканированием, работающий в диапазоне 400 – 1100 нм, и оригинальное программное обеспечение собственной разработки.

Ключевые слова: гиперспектральная камера, нечитаемые маргиналии, Острожская библия, рукопись, культурное наследие, оптико-электронные методы

Ссылка при цитировании: Баранов П.С., Парфенов В.А., Ронгонен С.Л., Соколов Н.М. Распознавание нечитаемых маргиналий с помощью гиперспектральной съемки на примере экземпляров Острожской библии из Библиотеки Российской Академии наук и Российской национальной библиотеки // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 3.2. С. 201–206. DOI: https://doi.org/10.18721/JPM.173.240

Introduction

The Ostrog bible is the first printed Bible in the Church Slavonic language, common to the contemporary Russians, Belarusians and Ukrainians. Published in 1581 in the city of Ostrog it became an outstanding example of translation, textual and publishing approaches. The Ostrog bible was a serious factor of the predication of Orthodoxy, the development of theology and printing tradition. Copies of the Ostrog bible have been well studied from different points of view. The diverse studies of the Ostrog bible have formed a large-scale array of interdisciplinary works.

Despite this, there are many unread marginalia on its copies (of which there are about 350 are known). Handwritten marginalia appeared on the pages of the Ostrog bible during the existence of copies and often contain valuable historical information. Most of them have been already introduced into scientific circulation, however, there are also those that, for various reasons, were impossible to read without the use of modern technical means.

Currently, when studying cultural heritage objects, including book and manuscript monuments, optical ones are increasingly being used. In recent years, multispectral analysis [1] and hyperspectral imaging [2] have been widely used for this purpose. Reviews of scientific centres, methods and literature are given in the articles [3, 4]. The study of colourful images in manuscripts [5], palimpsests [6], ink [7, 8], and various surface layers on the works of art [9] is possible using hyperspectral imaging. The main advantages of the method are high spectral resolution and a wide spectral range from near ultraviolet to near infrared. Hyperspectral cameras capture hundreds of narrow spectral bands in each image pixel, allowing to identify subtle differences in materials that cannot be distinguished by multispectral or conventional video cameras.

Also, a great advantage of the method is the non-destructive effect when analysing the object of study, as well as the mobility of the hyperspectral complex, which makes it possible to photograph especially valuable artifacts at the place of their storage.

Materials and Methods

It was important to collect material for an experiment on the object with a rich history of existence. This allowed us to obtain a sufficient number of diverse contexts. In total, we examined 24 copies and identified more than 100 unreadable handwritten marginalia – contexts. 11 copies from the Department of Rare Books of the Library of Russian Academy of Sciences (two of them are in the Department of Manuscripts) are described in the catalogue [10]. 13 copies from the Rare Book Department of Russian National Library, described in the catalogue [11]. Contexts were classified according to the reason for their unreadability. The reasons are diverse: ink fading, ink spreading, ink strike through, ink scraping, gluing with restoration materials.

The fact that all contexts belong to copies of the same edition ensured the same properties of the paper and its age (at least if the marginalia are on the block). As a rule, the copies are currently under the same storage conditions. The choice of different copies of the same edition is more relevant from a historical point of view, however, in this case we are dealing with the same paper, which perhaps makes the experiment easier. From the point of view of historical information content, contexts have different significance - from pen samples to page-by-page owner's inscriptions.

For the hyperspectral camera experiments on visualization 28 contexts were selected, of these: 6 contexts for faded ink; 5 contexts for smeared ink; 3 contexts for gluing with restoration paper; 2 contexts for crossed out with a pen; 1 context for a faded stamp, a dark stain on paper of unknown origin, respectively; 10 contexts for scratched text.

Table

Technical parameters of the FX10 camera

Characteristic	Parameter
Spectral range, nm	400–1000
Spectral resolution/pixel, nm	2.7 when binning 2
Number of spectral channels	224
SNR (maximum)	420:1
Number of readings	1024
Bit ADC	12
Frame rate, Hz	Till 327

Unreadable marginalia were examined by a hyperspectral camera working in the IR and UV bandwidths of the spectrum with further processing using the original software. We use a SPECIM FX10 hyperspectral camera manufactured in Finland. Its technical parameters are presented in the table.

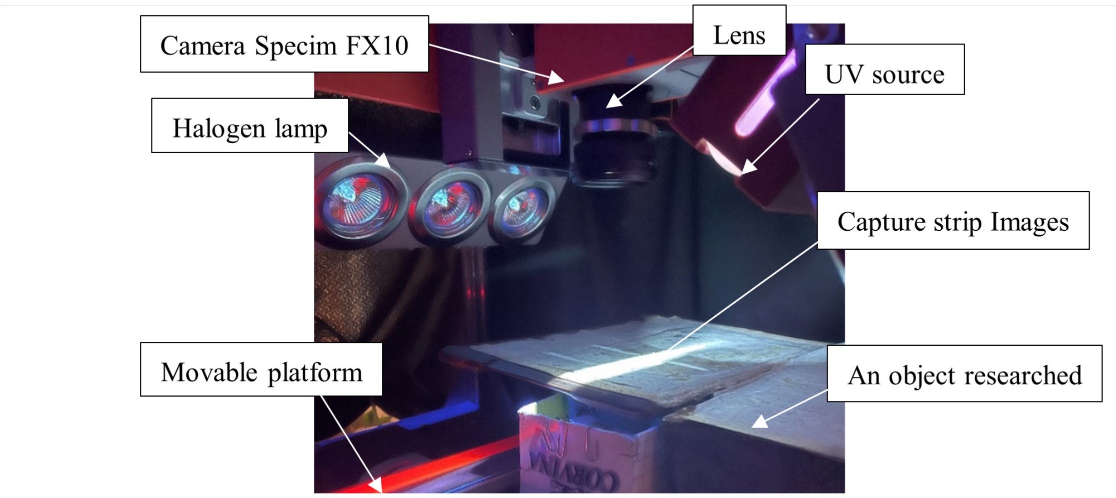

Fig. 1. Scanning an object with a hyperspectral complex

Fig. 1 shows a hyperspectral scanner with an object located inside. The scanned plane of the book was located perpendicularly to the optical axis of the hyperspectral camera. The scanning process used two light sources – a broadband halogen source for reflected light imaging and a narrowband 365 nm UV source for fluorescent imaging. Each light source was located at an angle of 45 degrees to the optical axis of the camera. Processing of the obtained images under two different light sources consisted of the following steps:

- Radiometric calibration [12];
- Independent contrasting by the CLAHE algorithm [13] of each channel for two hyperspectral images;
- Visual identification of the most informative spectral channels;
- Independent processing by the PCA principal component method [14] followed by contrasting with the CLAHE algorithm of each hyperspectral image with two light sources;
- Visual identification of the most informative images of the principal components.

Results and Discussion

Contexts with handwritten marginalia pasted over in restoration process (in our case, it is contemporary restoration paper or thick endpaper pasted over during historical restoration) appear very successfully during hyperspectral photography (Fig. 2). Narrow-spectral images were selected by finding the wavelength with the highest contrast in the region of interest. Marginalia that are not visible due to a dark, unknown spot on the paper cannot be read; other methods should be used.

a)
b)
c)
d)

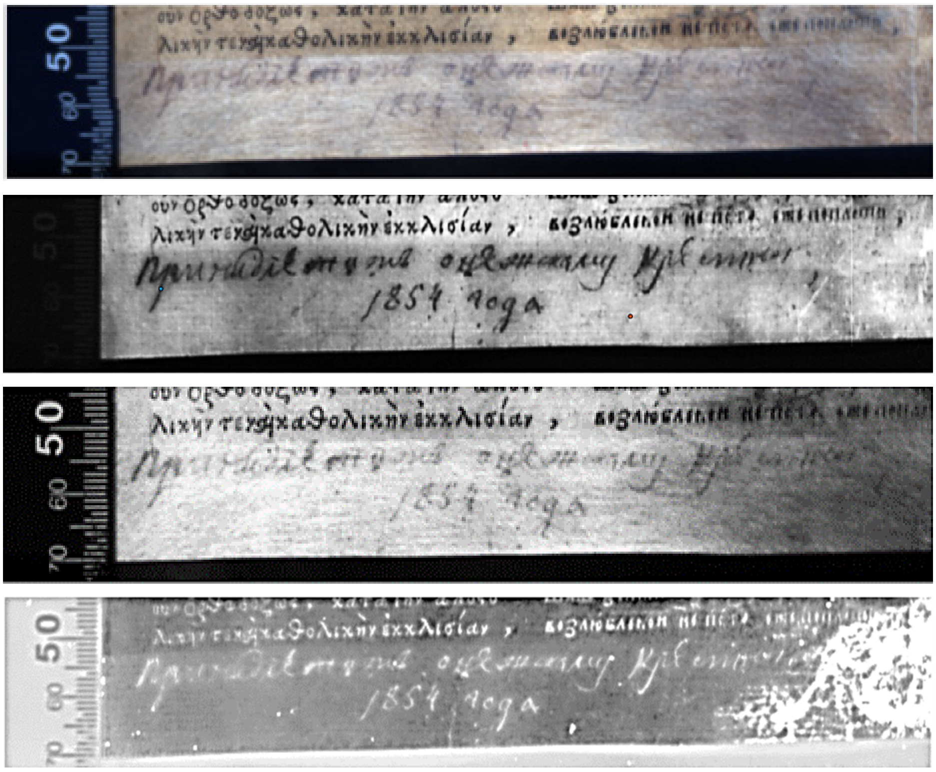

Fig. 2. Unreadable handwritten marginalia covered with restoration paper. Original (*a*); UV light, band 610 nm (*b*); halogen lamp, light band 570 nm (*c*) UV light, PCA-2 (*d)*

Crossed out with (darker) ink using a pen, the lines marked in [10, p. 76] as “crossed out inscription in Polish (?) language” lended themselves well to visualization, and the phrase became readable. Faded library stamp partially appeared during hyperspectral imaging.

The best results were shown on contexts whose “invisibility” was caused by ink fading. The UV spectrum shows an excellent result, while the IR spectrum does not give any result in this case. This is observed consistently in all such contexts.

One of the most common reasons for handwritten marginalia being unreadable is ink smearing. In this case, the use of hyperspectral imaging gives good visualization results (Fig. 3). The ink of the marginalia itself becomes darker, and the smeared layer becomes lighter. Obviously, to further decipher the marginalia, the resulting hyperspectral images should be combined.

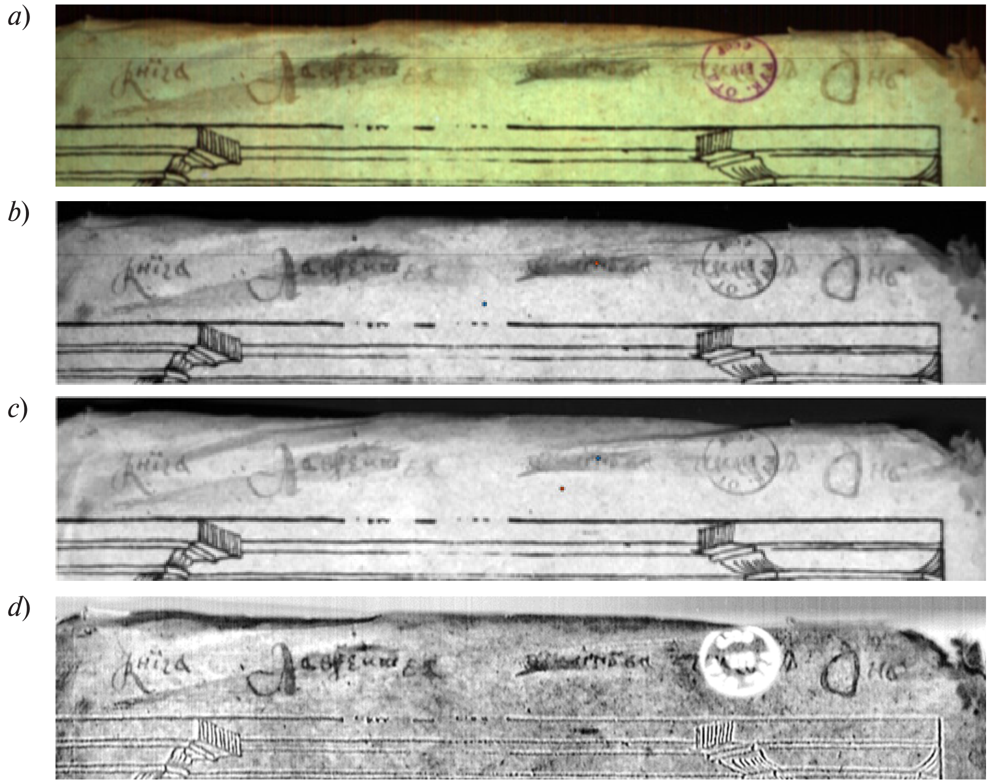

Fig. 3. Unreadable handwritten marginalia erased with ink. Original (*a*); UV light, band 480 nm (*b*); halogen lamp, light band 460 nm (*c*) UV light, PCA-4 (*d*)

The largest number of similar contexts is presented in the case of partially scratched handwritten marginalia. Since the "scratching" was heterogeneous, certain places in the text "succumbed" to hyperspectral photography, but for further reading the resulting visualizations should be combined.

Conclusion

Various reasons for extinction or "invisibility" of handwritten marginalia can be read using hyperspectral imaging. The choice of a successful shooting criterion depends on the characteristics of the "damage" of the marginalia, as well as on the material with which the inscription was made. Let us mention that in some contexts, hyperspectral imaging has revealed traces of bio damage on paper.

Reading these marginalia will add reliable information to the study of the Ostrog bible. Some of the marginalia we studied were read by researchers before (for example, a partially scratched owner's inscription from a copy of Epifanov collection or an inscription covered with restoration silk) [10]. But it was important for us to understand the principle - how hyperspectral imaging works for specific injuries. Further work with visualized contexts will increase the area of the text read.

As a result of image analysis, no universal way to increase the readability of marginalia was identified. The use of UV illuminators, compared to a halogen lamp, provides better results both when processing individual narrow-spectrum channels and in case of the PCA algorithm. The number of the spectral channel and the number of images of the main components are selected individually in each case. As part of further research, it is of interest to jointly process hyperspectral images obtained with different light sources.

At the time of the survey, the degradation of vintage specimens under ultraviolet radiation had not been investigated. Experts from the Library of the Russian Academy of Sciences considered the short-term use of ultraviolet light to be harmless, but further research on the possible effects on the condition of the antique paper and the visibility of the ink is needed.

Acknowledgments

We sincerely thank the staff of the manuscript departments of the Library of Russian Academy of Sciences and the Russian State Library for providing access to Ostrog biblia copies and useful technical discussions.

REFERENCES

1. **Alexopoulou A., Kaminari A., Moutsatsou A., Banou P.,** Imaging Techniques in the Study of Cultural Heritage. Kallipos, Open Academic Editions. (2024) 318.
2. **Cucci C., Casini A.,** Chapter 3.8 - Hyperspectral imaging for artworks investigation. Data Handling in Science and Technology. (32) (2019) 583–604.
3. **Balachenkova A.P., Lyakhovitsky E.A., Tsypkin D.O.,** Problems and prospects for the use of spectrozonal visualization in the study of written monuments. Bulletin of St. Petersburg University. history. 3 (66) (2021) 950–970.
4. **Kalashnikova A.A.,** Application of infrared visualization methods in modern source studies. Bulletin of NSU. Series: History, philology. 19 (8) (2020) 117–127.
5. **Viguerie L. de, Rochut S., Alfeld M., et. al.,** XRF and reflectance hyperspectral imaging on a 15th century illuminated manuscript: combining imaging and quantitative analysis to understand the artist's technique. Heritage Science. 6 (1) (2018) 11.
6. **Rapantzikos K., Balas C.,** Hyperspectral imaging: potential in non-destructive analysis of palimpsests. IEEE International Conference on Image Processing. Genova, Italy. (2005) II-618.
7. **Lyakhovitsky E.A., Tsypkin D.O.,** Infrared visualization of text in the study of monuments of ancient Russian writing. Historical informatics. (4) (2019) 148–156.
8. **Costa Rocha da C., Deborah H., Hardeberg J.Y.,** Ink bleed-through removal of historical manuscripts based on hyperspectral imaging. Lecture Notes in Computer Science. (2018) 473–480.
9. **Sandak J., Sandak A., Legan L., et al.,** Nondestructive evaluation of heritage object coatings with four hyperspectral imaging systems. Coatings. 11(2) (2021) 1–15.
10. Cyrillic editions of the 15th-16th centuries of the Department of Rare Books of the BAN: catalogue. Comp. M.Yu. Gordeeva, A.A. Romanova. SPb, BAN. (2010) 170.
11. Editions of the Cyrillic editions of the 15th-16th centuries. (1491–1600): catalog of the book. from collection GPB. St. Petersburg, 1993. Compiled by Lukyanenko V.I.
12. **Shaikh M.S., Jaferzadeh K., Thörnberg B., Casselgren J.,** Calibration of a Hyper-Spectral Imaging System Using a Low-Cost Reference. Sensors. 21 (11) (2021).
13. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. Journal of VLSI Signal Processing. (38) (2004) 35–44.

THE AUTHORS

BARANOV Pavel S.
psbaranov@etu.ru
ORCID: 0000-0002-8102-3858

RONGONEN Sofia L.
sofiagonobobleva@mail.ru
ORCID: 0000-0002-9050-3239

PARFENOV Vadim A.
vadim_parfenov@mail.ru
ORCID: 0000-0002-2048-4677

SOKOLOV Nikita M.
niksokolov2001@mail.ru

Received 30.07.2024. Approved after reviewing 23.08.2024. Accepted 10.09.2024.