On Hyperspectral Analysis of Water Soluble Writing Inks

1st Tomáš Zemčík
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
zemcikt@vutbr.cz

2nd Karel Horák
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
horak@vutbr.cz

***Abstract*—In this short paper a possibility of using Hyperspectral Imaging based analysis of water soluble inks is investigated. The current research has investigated this method mostly for marker inks, ballpoint pen inks and gel inks, yet curiously the whole water soluble ink class seems underrepresented. Samples were prepared and hyperspectral images of said samples taken, and some preliminary statistical analysis was performed. The data at this stage suggests that at least some samples are visually distinguishable from other samples, while certain other samples are virtually indistinguishable. A method for applying this knowledge in a handwritten text ink identification is briefly outlined and implemented in a simplified way, the results are promising on good input data, but poor on subpar images. The weaknesses of the method are discussed and potential improvement strategies proposed.**

***Index Terms*—hyperspectral imaging, specim IQ, ink analysis, document authentication**

I. Introduction

In the field of document analysis differentiation of various types of ink is an important part of understanding the way said document was created, edited, or perhaps further manipulated over it's lifespan. This understanding is important for a wide field of areas, such as historical document analysis, forensics or restoration [1].

There are numerous established methods for analysing inks. Menżyk et al. in [2] give an exhaustive overview of ink analysis methods in forensic examination of documents, and divide the methods into following categories in order of inquiry:

- **Preliminary examinations** - non-destructive methods requiring little to no specialised equipment. These include visual examination using microscopes or stereomicroscopes, optical analysis using various light sources including in the UV spectra.
- **Chromatographic methods** - methods of chemical analysis that require invasive sample collection from the document, and its further processing by dissolving the ink in appropriate solvent and performing the chromatography. The most common methods are Thin Layer Chromatography (TLC), High-Performance Thin Layer Chromatography (HPTLC), High-Performance Liquid Chromatography (HPLC) and Gas Chromatography (GC).
- **Mass spectrometry** ([2] lists mass spectrometry as part of spectroscopic methods, but for the different characteristics this paper distinguishes it from other spectroscopic methods) - destructive analysis of a sample to determine its precise chemical composition.
- **Spectroscopic methods** - the latest group of methods analysing spectral properties of inks. Spectroscopic methods used in document analysis include Fourier Transform Infrared Spectroscopy (FTIR), Raman spectroscopy and UV-VIS spectroscopy. Hyperspectral Imaging (HSI) is another representative of spectroscopic methods, which on top of the spectral information provides the spatial distribution of the information.

In the current ink analysis state-of-the-art as outlined by [2], and considering the recent developments in HSI techniques, the attention has been mostly concentrated on ballpoint inks, gel inks, printing inks and other kinds of artwork [3], [4]. Water soluble inks, such as drawing inks, calligraphy inks and fountain pen inks have gone mostly unexamined by the HSI methods. In this paper the results of the measurement are presented, and a proof-of-concept application is outlined and discussed to evaluate the feasibility of HSI based analysis on some selected water soluble inks.

II. Materials and methods

A. The sample preparation

Sample cards of the 10 analysed document inks were prepared. The inks were selected to represent some of the most common shades of blue document ink. Deliberately such shades were selected that in some cases they cannot be distinguished by the naked eye, and that depending on the application one could not even discern which of the ink is darker or lighter regardless of the shade. In addition a number of sample cards with marker inks was prepared to compare against the document inks as a reference and control group.

The individual samples were fashioned on a standard 80gsm printing paper representing the most common and available substrate. A square of ink was applied with a brush to supply a sufficient surface for individual pixel spectra analysis. The brush application was selected deliberately despite the fact that it is a rare way to apply ink in practice, it was

selected to create areas of varying thickness of application from heavier to lighter, no attempt was made to create a gradient or any organised sort of pattern, but the nature of the brush application intrinsically results in the desired pattern. In addition a sample text applied with a fine nib is included, and a small square of ink applied very thickly with a nib is also included as a high density application example. All sample coupons were allowed to dry for a period of 24 hours before any measurement. The marker samples were prepared to the identical design, excepting the marker was used, and therefore various patterns in the sample square, and thickness in the sample text occur.

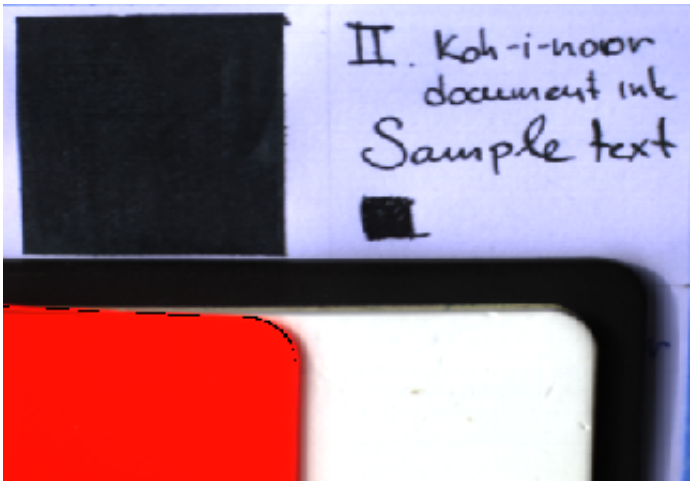

Fig. 1. Sample as prepared and captured (in pseudo RGB colours and cropped)

B. Image acquisition

The described samples were imaged using the Specim IQ hyperspectral camera. The camera is one of the first portable hyperspectral imaging solutions, and was released in 2018 by the Oulu, Finland based Specim company [5]. A multitude of applications have already utilized this camera, such as [6]–[9].

TABLE I
SPECIM IQ BASIC SPECIFICATIONS

Specim IQ basic specifications	
Parameter	**Value**
Spatial resolution	512 x 512 px
Spectral resolution	7 nm
Spectral range	400 - 1000 nm
F/number	1.7
Data format	ENVI compatible
Battery	5200 mAh Li-ion
Battery life	100 captures
Focusing distance	150+ mm
FoV	31° x 31°
Operating temperature	0°C - +40°C

The samples were placed slightly beyond the minimum focus distance of 200 mm from the lens, and illuminated sequentially with sunlight and artificial lighting using a special spectrum-controlled illumination unit tuned to as uniform a spectrum as possible. Calibration reference was included in the scene consisting of a white chip with known uniform reflectance across the spectral range of the camera, and a orange chip of a known optical properties with the reflectance cutoff of 590nm.

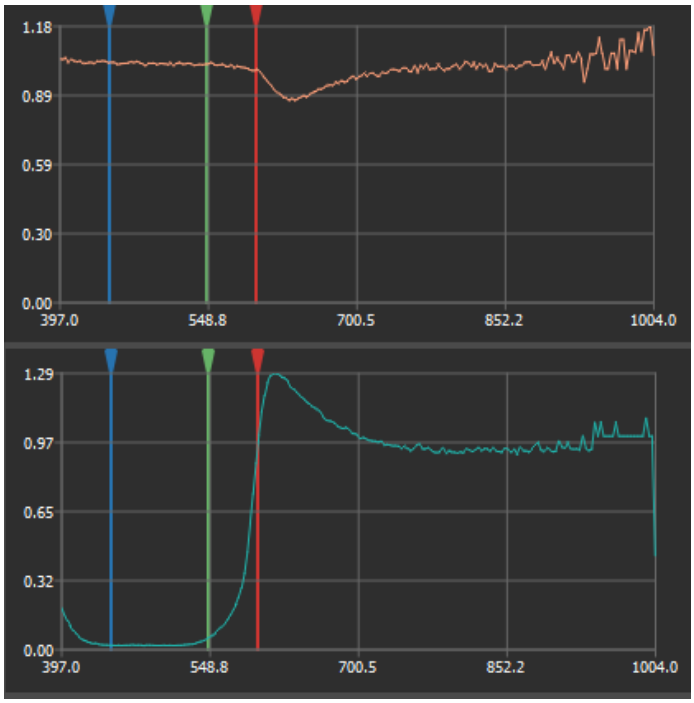

Fig. 2. The measured spectra of the reference chips, the white reference chip on top, the orange reference chip on the bottom

The measured spectra were verified using the OceanOptics USB 2000+ fibre optics spectrometer using the reflective spectrometry method.

The measurements were repeated for all 10 writing ink samples, and 4 control marker ink samples. Captures with multiple samples in a single image were also taken in order to allow for later cross-check between the various samples for variation of illumination (in case of the daylight samples, since the captures and sample changes take a non-negligible amout of time and the lighting conditions may have shifted in the meanwhile).

III. Data analysis

The individual scanned samples were converted to a hypercube image representation of the dimensions 512 x 512 x 204 - representing the 204 spectral slices of 512 px x 512 px each. Within each sample picture the sample ink square of uniform dimensions of 150 px x 150 px was annotated by hand, and a subimage hypercube of the dimensions 150 x 150 x 204 was extracted as a data sample. This represents 22500 individual pixels each containing a full spectrum, where each pixel itself is a sample of that particular place's application of ink. Together they represent a representative sample of ink for further statistical analysis. Such samples may be collected in

a database of inks and be used for classification of unknown samples.

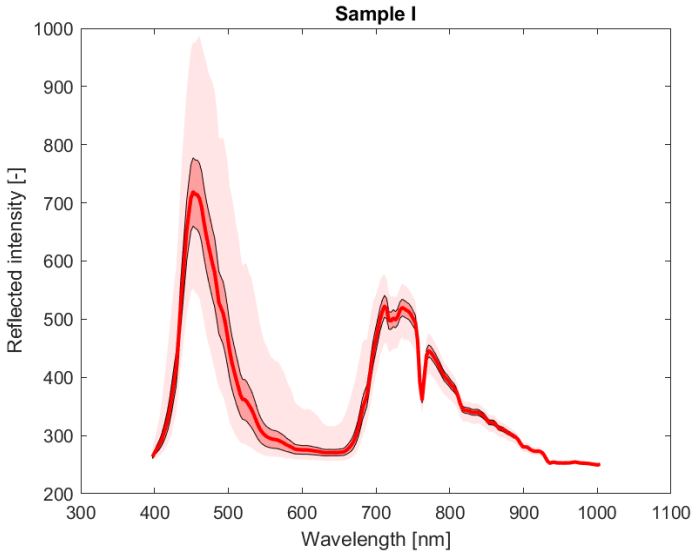

Fig. 3. Sample I

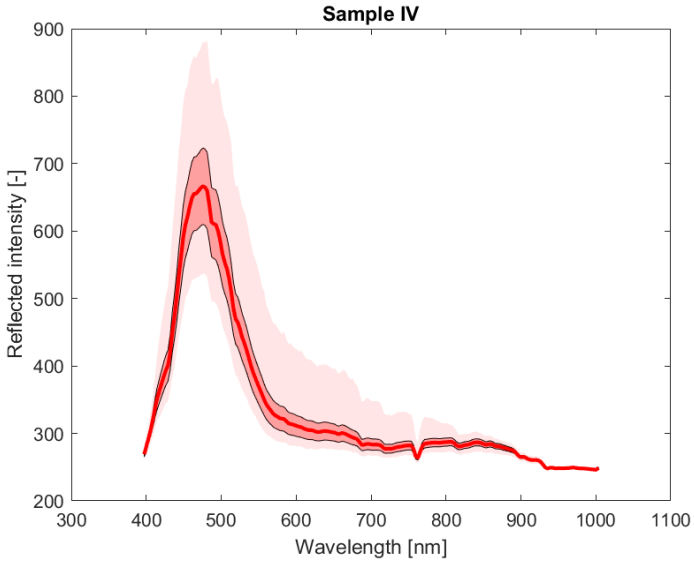

Fig. 4. Sample IV

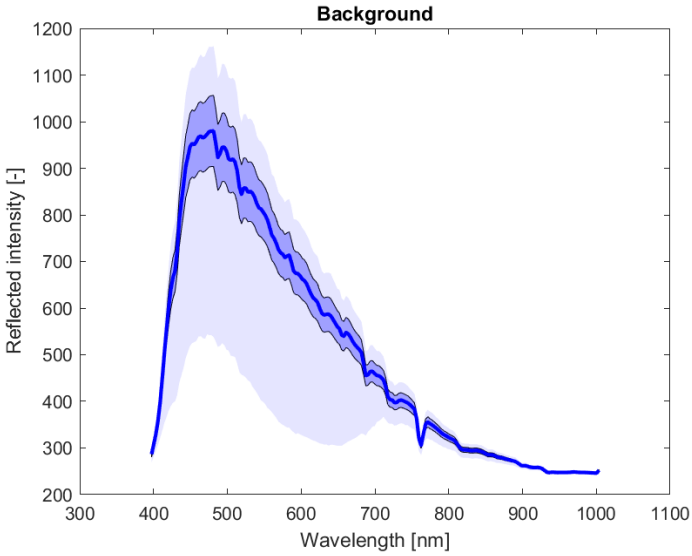

Fig. 5. Background sample

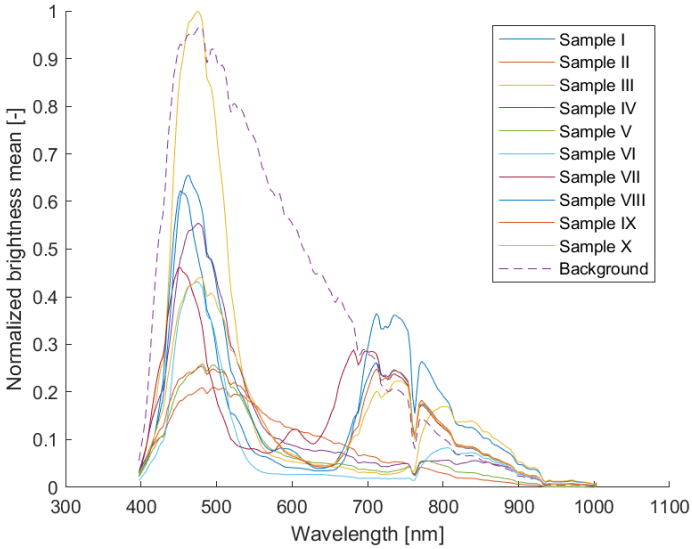

Fig. 6. Sample II

In figures 3 through 5 two ink samples (in red) and a statistical sample of the background (in blue) are represented. The solid line represents the mean value of intensity across the 22500 pixels (a variable higher number of pixels is used for the background sample) for each of 204 spectral bins between circa 400 nm to 1000 nm. The darker shaded area around the mean line represents the standard deviation area across the samples, and the lighter shaded area represents the absolute minimum and maximum for each sample.

In figure 6 the normalised mean spectra of all ten ink samples (in solid lines), and the background (in dashed line) are given. Please note that some of the mean spectra extend above the background, which in reflective spectroscopy would imply the unlikely situation of the inked surface having higher reflectivity than the (white) background. This arises from the slight variance of ambient illumination during image acquisition and from the nature of the statistical analysis used (the background sample was large and included more or less shadowed areas).

Even by visual inspection it is clear that some of the selected samples and the background (as shown above) are easily distinguishable from certain other samples, therefore there is clearly a potential for discrimination between at least some combinations of samples.

IV. The proposed method

The proposed proof-of-concept method is intended to be a simple way to demonstrate that a solution of water soluble ink differentiation exists within the domain of HSI.

For a future usable application, the presumed form of the analysed sample would be a block of handwritten text, or a drawing on some sort of a substrate. During the sample imaging it would be important to ensure a correct acquisition of both the ink application areas, and a sufficient sample of the background for future decomposition. In the current simplified experiment an identical background (a standard copier paper) is assumed.

In figure 7 a workflow of the proposed simple method is outlined, the individual blocks are then described in the paragraphs below:

After **image acquisition** the sample file is collected from the camera and converted into a hypercube representation. Multiple step **preprocessing** is then performed on the image. First a Region of Interest (RoI) is selected, and the hypercube cropped to dimension in order to reduce computational requirements. Then Gaussian blur of appropriate kernel size (depending on the definition of the finest lines) is applied over all the spectral images. This is done in order to avoid having certain spectral slices in good focus, while other slices are our of focus due to the nature of different wavelengths propagating through materials at different velocities, and the imperfect compensation of the quartz lens used. Here a additional performance may be gained by analysing which wavelengths suffer what level of blur and then blurring the other wavelengths accordingly, the distribution will not change significantly with camera settings. And finally known bands of interest may be removed or selected based on a priori knowledge of the materials.

Then **background selection** is performed manually on the indicated image, an area as large as possible of the clear substrate is indicated to the algorithm for a statistical sample to be collected.

Optionally a **dimensionality reduction** step may be included to further reduce the hypercube size. On the analysed samples of water soluble inks 10 principal components have proved to carry well in excess of 90% of the information, therefore this step could offer a significant increase in the system's performance.

Using the previously collected statistical samples of the inks (as described above) and a selected background sample a **Support Vector Machine (SVM) classifier** model is trained to classify individual pixels in the image. Applying this model pixelwise a label image is created with individual pixels representing a class prediction for the corresponding hypercube spectrum.

Because the pixelwise classification is is not perfect on real data, a step of **smoothing and filtering** the resulting label image is taken. By applying a sequence of elementary morphological operations extremely small areas of discrete classes are removed (usually the borderline pixels mixing background and ink spectra or spectra of different inks). Larger areas of single labels are thus smoothed and merged where necessary. The output label image is then produced.

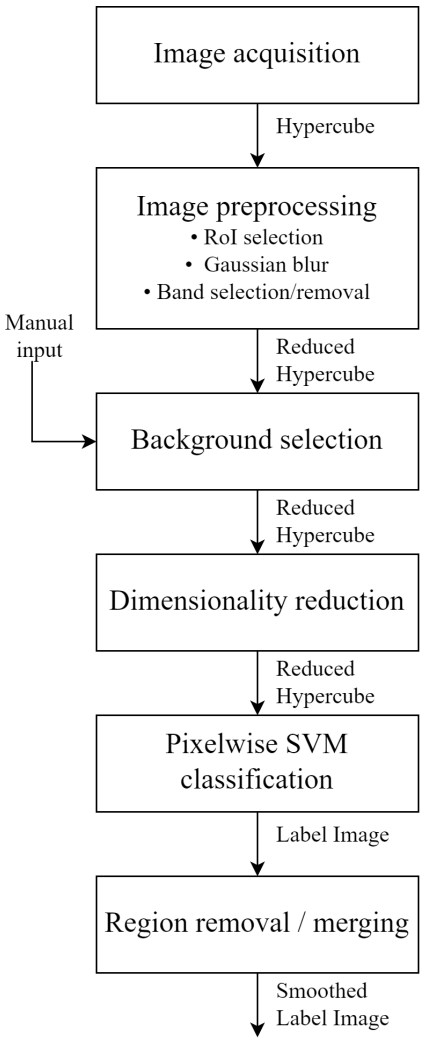

Fig. 7. The simplified proposed method workflow

A. Results and discussion

In figure 8 a successfully produced unsmoothed label image is shown. Note that while Sample III is near perfectly classified, segmented and resolved, the Sample V below is near perfect in the large brush applied area and pen applied square, however the pen applied sample text and description are not well resolved, and missclassified pixels (confused with Sample II in light blue, and Sample IV in red) are also observed. Generally the methods seems to perform fairly well on large resolution images such as this one, some combinations of inks are prone to miss-classification more than others, as expected (subject to future quantitative analysis).

Figure 9 demonstrates a loss of resolving power when an insufficient resolution and lighting image is supplied. The sample VI in blue is clearly being confused with sample III in purple, and Sample IV in red in case of written text. The background spectra, while still correctly classified, have been

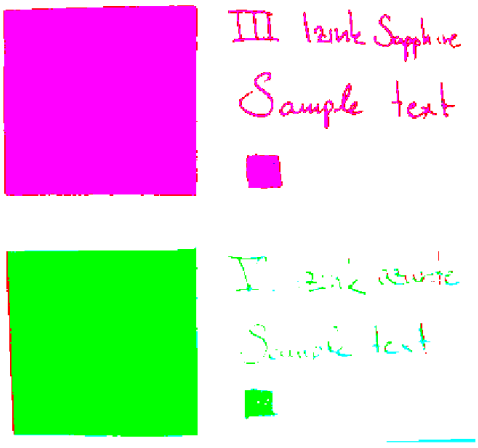

Fig. 8. Unsmoothed label image as a result on good resolution source data (Sample III in purple, Sample V in green and the Background in white)

variable throughout the region of interest due to the insufficient lighting.

Fig. 9. Unsmoothed label image as a result on insuffucuent resolution source data (Sample IV in red, Sample VI in blue and the Background in white)

The method seems to perform fairly poorly when either the resolution of handwritten text, or the lighting conditions are compromised. To remedy the lighting variability an incorporation of the calibration procedure outlined above using the calibration object included in the samples is necessary, experimentation with calibrated images showed improvement in classification quality, however the results have not yet been stable.

The poorer resolution caused problems may to an extent be mitigated by employing hyperspectral endmember extraction methods as outlined in [10]. These methods are designed to decompose spectra of individual pixels into different materials (endmembers) in the image, this helps in resolving the fringe regions where in a single pixels spectra of multiple materials are projected (due to insufficient resolution, blur or other factors). This would be particularly helpful for the thin lined handwritten text.

V. Conclusions

In this short paper a method for preparing water soluble writing ink samples, and capturing their hyperspectral representation using the Specim IQ camera is presented. Examples of the results obtained are shown, including some basic statistical analysis, and a method for automated ink sample recognition is outlined.

Based on the experiments performed the author believes that discrimination between different water soluble inks using a hyperspectral imaging method is (at least in some cases) feasible. In some cases the spectra are virtually identical, and a deeper analysis is required to determine whether or not a solution within the hyperspectral imaging domain (or within the Specim IQ camera spectral range) is possible. Also the selected pixel-wise approach is not perfect for discriminating the ink type for given object in an image, a more sophisticated segmentation and classification approach would be required in the future application.

A method for analysing samples of handwritten text is briefly outlined and discussed. This proof-of-concept method, while being very far from being a usable application, shows it is certainly possible to differentiate certain water soluble inks by way of hyperspectral image analysis - as speculated above. In future work the author intends to incorporate the improvements outlined in the previous chapter (input image calibration and endmember extraction methods), and expand the experimentation on larger sample of handwritten text using more inks in order to perform a quantitative analysis of differentiation ability of this method, it would also be interesting to expand the experimentation into other colours of ink, such as blacks and reds.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 - “Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

References

[1] Morales, Aythami & Ferrer, Miguel & Diaz, Moises & Carmona-Duarte, Cristina & Thomas, Gordon. (2014). “The use of Hyperspectral Analysis for Ink Identification in Handwritten Documents“. Proceedings - International Carnahan Conference on Security Technology. 2014. 10.1109/CCST.2014.6986980.

[2] Menżyk, Alicja & Zadora, Grzegorz & Sajewicz, M.. (2016). “Physicochemical analysis of inks and methods of evidence evaluation - A review“. Problems of Forensic Sciences. 104. 245-278.

[3] Tzeng, Di-yuan & Berns, Roy. (2001). “Spectral-Based Ink Selection for Multiple-Ink Printing I. Colorant Estimation of Original Objects“.

[4] Maria Kubik, Chapter 5 “Hyperspectral Imaging: A New Technique for the Non-Invasive Study of Artworks“, Editor(s): Dudley Creagh, David Bradley, Physical Techniques in the Study of Art, Archaeology and Cultural Heritage, Elsevier, Volume 2, 2007, Pages 199-259, ISSN 1871-1731, ISBN 9780444528568, https://doi.org/10.1016/S1871-1731(07)80007-8.

[5] E. Ikäheimo and J. Jussila, ”Introducing Specim IQ”. Oulu, FInland: Specim, Spectral Imaging Oy, 2018.

[6] J. Behmann et al., “Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for Plant Phenotyping and Disease Detection,” Sensors, vol. 18, no. 2, p. 441, Feb. 2018, doi: 10.3390/s18020441.

[7] Solovchenko, Alexei & Dorokhov, Alexei & Shurygin, Boris & Nikolenko, Alex & Velichko, Vitaly & Smirnov, Igor & Khort, Dmitry & Aksenov, Aleksandr & Kuzin, Andrei. (2021). ”Linking Tissue Damage to Hyperspectral Reflectance for Non-Invasive Monitoring of Apple Fruit in Orchards”. Plants. 10. 310. 10.3390/plants10020310.

[8] Tejada-Casado, Maria & Ghinea, Razvan Ionut & Martínez-Domingo, Miguel & Pérez Gómez, María del Mar & Cardona, Juan & Ruiz-López, Javier & Herrera, Luis. (2022). ”Validation of a Hyperspectral Imaging System for Color Measurement of In-Vivo Dental Structures”. Micromachines. 13. 1929. 10.3390/mi13111929.

[9] Alt, V. & Gurova, Tamara & Elkin, O. & Klimenko, D. & Maximov, L. & Pestunov, I. & Dubrovskaya, O. & M.A., Genaev & Erst, T. & Genaev, K. & Komyshev, Evgenii & Khlestkin, Vadim & Afonnikov, Dmitry. (2020). ”The use of Specim IQ, a hyperspectral camera, for plant analysis”. Vavilov Journal of Genetics and Breeding. 24. 259-266. 10.18699/VJ19.587.

[10] K. V. Kale, M. M. Solankar, and D. B. Nalawade, “Hyperspectral Endmember Extraction Techniques,” Processing and Analysis of Hyperspectral Data, Jan. 2020, doi: 10.5772/intechopen.88910.