Natural Vision Data File Format as a New Spectral Image Format for Biological Applications

Joji Sakamoto[1,*], Jennifer Dumont[2], Laure Fauch[1], Sarita Keski-Saari[2], Lars Granlund[2], Ilkka Porali[2], Joni Orava[1], Jouni Hiltunen[1], Elina Oksanen[2], Markku Keinänen[2], and Markku Hauta-Kasari[1]

[1] School of Computing, University of Eastern Finland, P.O.BOX 111, 80101 Joensuu Finland
[2] Department of Biology, University of Eastern Finland, P.O.BOX 111, 80101 Joensuu Finland
{joji.sakamoto,jennifer.dumont,laure.fauch,sarita.keski-saari, lars.granlund,ilkka.porali,joni.orava,jouni.hiltunen,elina.oksanen, markku.keinanen,markku.hauta-kasari}@uef.fi
https://www.uef.fi

Abstract. Many kinds of spectral image formats are used for various applications, but there is still no existing standard format. Natural Vision data file format is one of the best possible candidates for the standard of spectral image format due to its flexibility to adapt to various kinds of existing image format and capacity to include information needed for each application. In biology, the analysis of huge datasets acquired by various techniques requires the use of specific databases. In order to be able to combine different data, defining a standard spectral image format that includes biological parameters is of prime importance. This paper describes an attempt to use Natural Vision data file format for spectral images related to biology and highlights the merits of Natural Vision data file format as an application oriented spectral image format.

Keywords: Spectral imaging, biological imaging, application oriented, Natural Vision data file format, ENVI data file format, minimum information standard.

1 Introduction

Spectral imaging technology is used in various applications such as remote sensing [1], medical imaging [2] and biology [3]. At present, numerous different file formats are used to store the spectral image data. However, there is still no existing standard format. A need for a standard file format has been discussed by CIE Division 8, TC8-07 of Multispectral imaging such as the MUSP multispectral image format, the Natural Vision data file format, JPEG2000, TIFF, the GeoTIFF and HDF5 [4], [5]. The ENVI data file format is one of the spectral image format used for actual applications of satellite and aircraft remote

A. Elmoataz et al. (Eds.): ICISP 2014, LNCS 8509, pp. 124–132, 2014.

sensing data [6]. The ENVI data file format can include spectral image data, sensor type information and map information for its applications. If the ENVI data file format is used for biological applications to save spectral images and analyze spectral information, the parameters of experimental conditions have to be kept in other files which constitutes a source of error and can lead to difficulty to analyze and compare the images. In addition, in biology, a lot of data is obtained by high-throughput omic technologies and a lot of information is stored in various databases. A standard format should be usable for automated data analysis and data mining could give us new insights of those results but it requires parameters describing unambiguously experimental conditions. Table 1 is an example of parameters linked to one image from an experiment in plant science. The parameters give the main features of the studied sample, the growing conditions of the sample and the treatments applied to the sample. In other fields of biology, the situation will be the same even if the required parameters are not exactly the same as in plant science. Thus, it would be beneficial if the standard spectral image format would allow to store all those parameters within the images with defined controlled vocabulary and ontology.

The Natural Vision data file format is one of the spectral image format used in the Natural Vision project established in 1999 by the National Institute of Information and Communications Technology (NICT) (formerly TAO, Telecommunications Advancement Organization of Japan) [4]. The Natural Vision data file format specified both image and color profile data formats based on the International Color Consortium (ICC) color profile where the profile data is attached to the image file. The Natural Vision data file format has flexibility to attach or include any kinds of image format such as JPEG2000, TIFF and ENVI [4]. This format also offers a possibility to attach application oriented parameters to the image file. This paper presents an attempt to analyze spectral images from biological applications using the Natural Vision data file format by taking advantages of the benefits thereof.

2 Methodology

The current version of the Natural Vision specification is Natural Vision multispectral image metadata format, XML schema (NV-XML) [7]. However, we used the old version of the Natural Vision specification 2.0s (NV2) [8] written in binary form because of large number of sample images available. Natural Vision data file format defines only the metadata of spectral images. The main difference between NV-XML and NV2 is the form of writing metadata. Therefore, using NV2 format does not affect the purpose of this experiment.

Fig. 1 shows the brief structure of NV2 format in case of still image [8]. It is organized in three parts, index, image data and color profile data. The color profile data part can contain the camera information, the illumination spectrum, the statistical information of captured objects, the rendering illumination spectrum and the display information.

As shown in Table 2, future expansion possibilities have been taken into account with a reserved part of 28 bytes at Byte Offset 100-127 of NV2 format.

Table 1. Biological parameters linked to one image from a biology experiment and explanation of those parameters for spectral images in biology

	Parameters	Values	Explanation
Sample information	Species	*Arabidopsis thaliana*	The sample of interest
	Genetic background	Col, Ws, C24, Ler	Common *Arabidopsis* ecotypes
	Age (*day*)	21	The age of the sample when the image has been taken
	Place of imaging	Laboratory	Place where the image has been taken
	Organ	Shoot	This is linked to the scale of the image, it could also be at the tissue scale
Environmental conditions	Light type	halogen lamps	
	Light intensity ($\mu mol.m^{-2}.s^{-1}$)	150	
	Light period (*hour*)	12	
	Day temperature ($°C$)	23	The temperature and the air humidity are usually different in day and night and are important environmental parameters
	Night temperature ($°C$)	19	
	Day air humidity (%)	65	
	Night air humidity (%)	75	
	Growing medium	Agar (0.8%)	Plants can be grown on different media, such as soil, agar or in hydroponic conditions
	Nutrients	½ MS, 0.5 % sucrose	Different kind of nutrients can enhance the growth of plants
	Place	Phytotronic chambers	Phytotronic chambers allow to grow plants in controlled conditions
	Duration (*day*)	7	This is the time needed to have a starting material before any treatments
Treatment information	Type of experiment	Control vs Treatment	
	Type of sample	Treated	Two possibilities : Control or Treated
	Light intensity ($\mu mol.m^{-2}.s^{-1}$)	250	
	Light period (*hours*)	12	
	Light stress duration (*day*)	14	
	Day temperature ($°C$)	10	
	Night temperature ($°C$)	10	
	Cold stress duration (*day*)	14	
	Comments		Free space available if some important details are not mentioned in the previous parameters

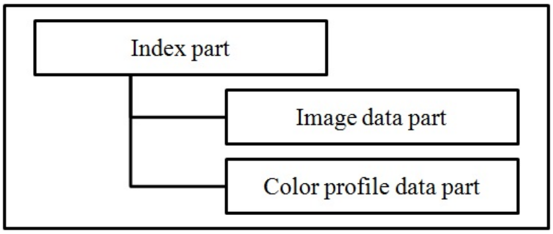

Fig. 1. Brief Structure of NV2 format (Translated from reference [8])

Table 2. Header description of color profile data part in NV2 format (Copied from reference [8])

Byte Offset	Content	Encoded as...
0-3	Profile size	Uint32Number
4-7	CMM Type signature	
8-11	Profile version number	
12-15	Profile/Device Class signature	
16-19	Color space of data	
20-23	Profile Connection space	
24-35	Date and time this profile was first created	dateTimeNumber
36-39	'scsp' profile file signature	
40-43	Primary Platform signature	
44-47	Flags to indicate various options for the CMM such as distributed processing and cashing options	
48-51	Device manufacturer of the device for which this profile is created	
52-55	Device model of the device for which this profile is created	
56-63	Device attributes unique to the particular device setup such as media type	
64-67	Rendering Intent	
68-79	The XYZ values of the illuminant of the Profile Connection Space. This must correspond to D50.	XYZNumber
80-83	Profile Creator signature	
84-99	Profile ID	
100-127	28 bytes reserved for future expansion	

Four bytes of this reserved part are used to distinguish intended application of each spectral image. Here, the development of the new simple software to create Natural Vision file from ENVI file is reported. ENVI images were collected by VNIR (Spectral Camera sCMOS-CL50-V10E-OEM integrated by using Andor Zyla X sCMOS, spectral range: 400-1000nm, spectral resolution: 2.8nm) and SWIR (Spectral Camera SWIR-LVDS-100-N25E integrated by using Spectrograph Inspector N25E, spectral range: 970-2500nm, spectral resolution: 10nm) cameras (Specim, Oulu, Finland). The reserved part of NV2 format will be used to store characters of bio to distinguish that this data is from biology applications. A function has been developed to register the name and value of all the needed parameters describing the experiments (Table 3). All the information related to biology parameters will be attached at the end of NV2 data file. The system is flexible so various parameters can easily be included in the NV2 data file.

Table 3. The structure of application oriented parameters attached at the end of Natural Vision data file format

Byte Size	Parameter name	Parameter value	Encoded as...
4	application name	'bio '	ASCII
4	number of parameter	n	ASCII
30	parameter name 1	species	ASCII
30	parameter value 1	arabidopsis thaliana	ASCII
...	...	...	ASCII
30	parameter name n	chemicals	ASCII
30	parameter value n	control samples	ASCII

3 Experiments

NV2 data file creating tool has been developed from usability point-of-view of our experiments (Fig. 2). Its interface has been kept very simple to enable quick and easy use (Fig. 3). First, the sample image, the white reference image and the dark image are selected among the files on the computer. Next, the type of camera and illumination are selected. Then, the size and the bands number of the sample image are entered. Finally, the original image is displayed on the interface. After that, it is possible to display the extracted area (ROI, Region of Interest) from the original image before registering the biological parameters and save the extracted image in NV2 format with the defined parameters. Fig. 4 shows the flowchart of NV2 data file creating tool.

In this case, hyperspectral images (VNIR: number of bands is 240, SWIR: number of bands is 256) are used as the original images and the spectral reflectance image [R] is computed as follows;

$$[R] = \frac{(sample\ image - dark\ image)}{(white\ image - dark\ image)} \tag{1}$$

The saving part of this NV2 data file creating tool is still under development. The image information of original ENVI file should be extracted automatically from header file of ENVI format (***.hdr). NV2 analyzing tool is also demanded to make its analyzing easily and efficiently.

The spectra of four *Arabidopsis thaliana* plants submitted to different light intensities and temperatures are compared in Fig. 5. Day-time temperature of $23(°C)$ and light intensity of 150 $(\mu mol.m^{-2}.s^{-1})$ are usual control conditions for

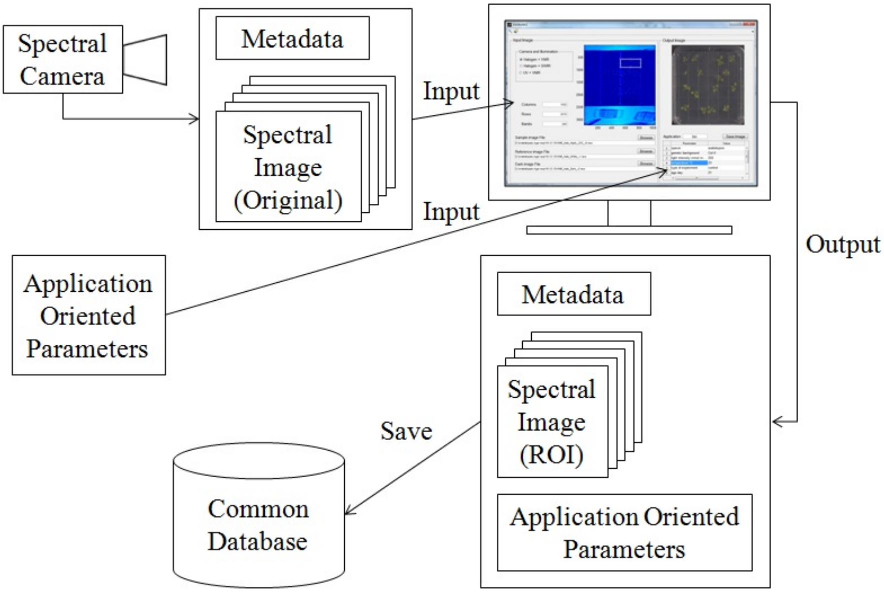

Fig. 2. Schematic Diagram of the NV2 data file creating tool

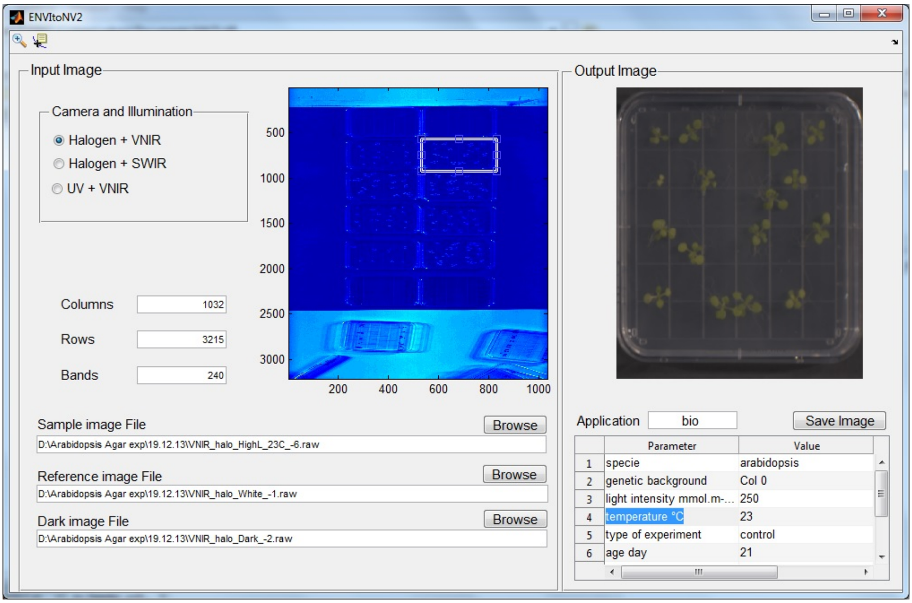

Fig. 3. Interface of the NV2 data file creating tool

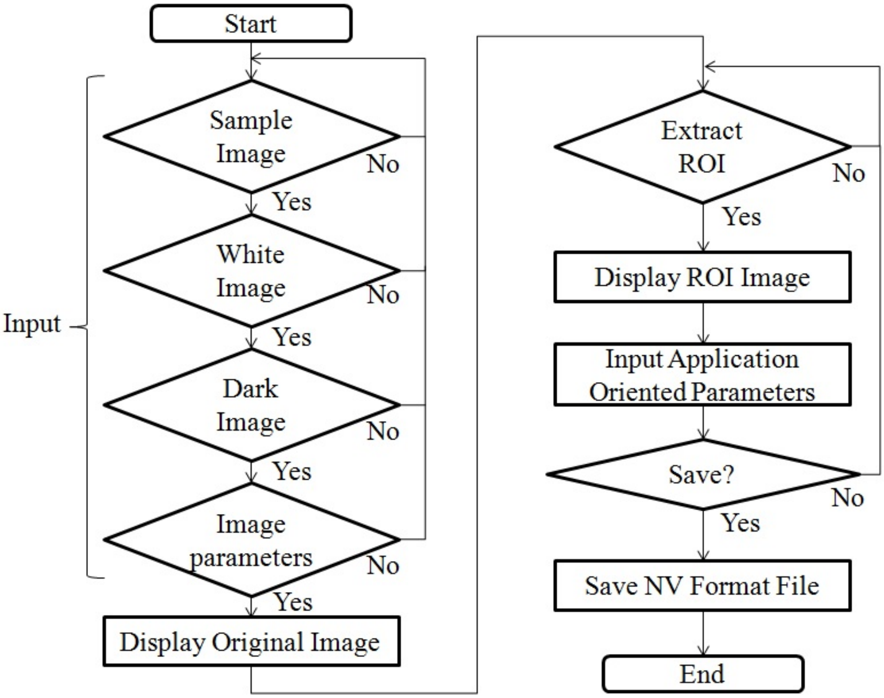

Fig. 4. Flowchart of the NV2 data file creating tool

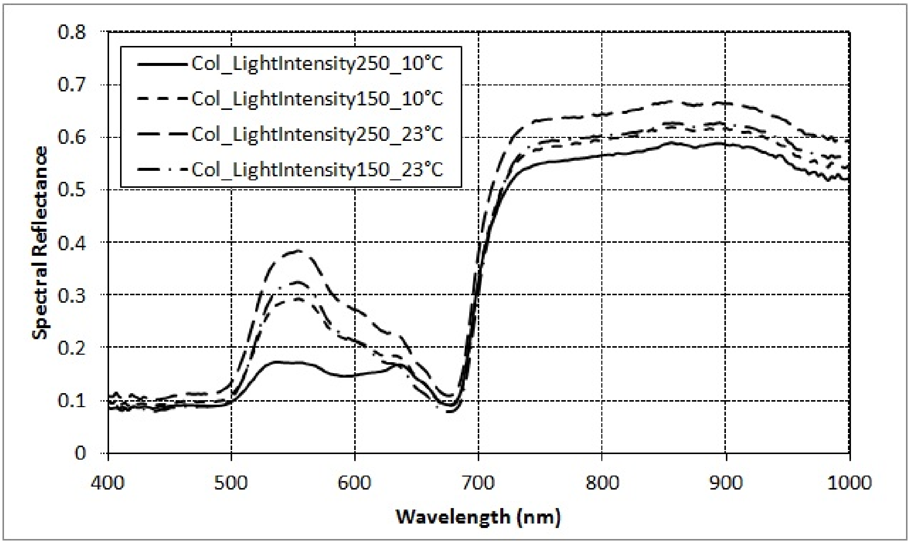

Fig. 5. The spectra of four *Arabidopsis thaliana* plants submitted to different light intensities and temperatures (Col: Genetic background, Light Intensity: $(\mu mol.m^{-2}.s^{-1})$)

these plants. On the other hand, plants subjected to the treatment of low temperature of 10 ($°C$) and high light intensity of 250 ($\mu mol.m^{-2}.s^{-1}$) are stressed. The spectra of the leaves of these plants show different value at 560nm. This difference may be due to the accumulation of protective compounds, such as anthocyanins [9]. Thus, correlating spectral images and their experimental parameters are frequently used in actual applications.

4 Discussion

By using the developed tool, spectral image data and experimental condition parameters can be handled in one file. The tool is flexible, easy to use, helpful to reduce the risk of mistakes since all the information are saved together and gives new possibilities for further analyses linked to databases. Natural Vision data file format can incorporate other image formats. Thus, common Natural Vision data file from the spectral images of various kinds of capturing devices can be made and the images can be compared. imzML is a common data format used for mass spectral imaging which is an emerging new technique [3]. This imzML file is divided into two separate files such as XML file for metadata and binary file for mass spectral data. This structure is quite close to NV-XML format. Thus, the new Natural Vision data file format also has a possibility to include imzML into one image data file. It would be a significant advantage if the spectral imaging data obtained by optical technique could be combined to mass spectral imaging data thanks to similar formats [10]. If the Natural Vision data file format would be used as the standard of spectral image in various applications, this would enable researchers to use a common database to store their spectral images and compare their images to the spectral images from different researchers. Thus, this would serve as a stepping stone for other applications to create new research and exciting possibilities.

5 Conclusion and Future Work

This paper described the usability of Natural Vision data format as a common spectral image format to handle spectral images between researchers targeting biology as one of the applications. One simple software interface is developed to get an original ENVI format file, extract a ROI and input application parameters to create a NV2 file in which those parameters are attached for spectral analyzing in biology. The Natural Vision data file format is a strong candidate to be the standard of spectral imaging due to its capacity and flexibility. Using Natural Vision data file has a possibility to release the researchers from managing an enormous quantity of application oriented parameters related to each spectral image. As future works, the number of the applications is increased to confirm the usability of Natural Vision data file format. Developing Natural Vision tools to adapt various image formats is required. The best way to pack application oriented parameters from various applications into Natural Vision data file should be considered simultaneously.

References

1. Pant, P., Heikkinen, V., Korpela, I., Hauta-Kasari, M., Tokola, T.: Logistic Regression-Based Spectral Band Selection for Tree Species Classification: Effects of Spatial Scale and Balance in Training Samples. IEEE Geoscience and Remote Sensing Letters 11(9), 1604–1608 (2014)
2. Antikainen, J., Von und zu Fraunberg, M., Orava, J., Jääskeläinen, J.E., Hauta-Kasari, M.: Spectral Imaging of Neurosurgical Target Tissues through Operation Microscope. Optical Review 18(6), 458–461 (2011)
3. Schramm, T., Hester, A., Klinkert, I., Both, J.-P., Heeren, R.M.A., Brunelle, A., Laprévote, O., Desbenoit, N., Robbe, M.-F., Stoecklif, M., Spengler, B., Römpp, A.: imzML- A common data format for the flexible exchange and processing of mass spectrometry imaging data. Journal of Proteomics 75, 5106–5110 (2012)
4. Jetsu, T., Herzog, P., Yamaguchi, M., Jääskeläinen, T., Parkkinen, J.: Standardization of Spectral Image Formats. In: AIC Colour 05 10th Congress of the International Colour Association
5. Jetsu, T., Herzog, P., Jääskeläinen, T., Parkkinen, J.: Standardization of Spectral Image Formats. Pattern Recognition and Image Analysis 15(3), 618–620 (2005)
6. ENVI Users Guide, ENVI Version 4.1 September 2004 edn. RSI Research System Inc. (2004)
7. Natural Vision multispectral image meta data format, XML schema (NVXML), Specification Version 1.20, http://nvision.jp/NVXMLformat_ver1.2-fix.pdf
8. Natural Vision Data File Format Specification, Version 2.0s. Akasaka Natural Vision Research Center, National Institute of Information and Communications Technology (2003), http://seika-kokai.nict.go.jp/doc/result/200411002/200411002houkoku-02.pdf
9. Gitelson, A.A., Merzlyak, M.N., Chivkunova, O.B.: Optical properties and nondestructive estimation of anthocyanin content in plant leaves. Photochemistry and Photobiology 74, 38–45 (2001)
10. Taylor, et al.: Promoting coherent minimum reporting guidelines for biological and biomedical investigations: The MIBBI project. Nature Biotechnology 26(8), 889–896 (2008)