



**HAL**  
open science

# Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: theoretical and practical study

Gilles Rabatel, N. Gorretta, S. Labbé

## ► To cite this version:

Gilles Rabatel, N. Gorretta, S. Labbé. Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: theoretical and practical study. *Biosystems Engineering*, 2014, 117 (1), p. 2 - p. 14. 10.1016/j.biosystemseng.2013.06.008 . hal-00947991

**HAL Id: hal-00947991**

**<https://hal.science/hal-00947991>**

Submitted on 17 Feb 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: theoretical and practical study

**Gilles Rabatel**<sup>1</sup>, Nathalie Gorretta<sup>1</sup>, Sylvain Labbé<sup>2</sup>

<sup>1</sup>Irstea, UMR ITAP, centre de Montpellier, 361 rue J.F. Breton - BP5095, F-34196 Montpellier, France

gilles.rabatel@irstea.fr

<sup>2</sup> Irstea, UMR TETIS, centre de Montpellier, F-34196 Montpellier, France

**Abstract.** Multispectral images, including red and near-infrared bands, have proved efficient for vegetation-soil discrimination and agricultural monitoring in remote sensing applications. However, they remain little used in ground-based and unmanned aerial vehicle (UAV) imagery, due to a limited availability of adequate 2D imaging devices. A methodology is proposed to obtain simultaneously the near-infrared and red bands from a standard single RGB camera, after having removed the near-infrared blocking filter inside. Its ability to provide satisfactory NDVI computation for vegetation and soil has been assessed through spectral simulations. Application in field conditions with Canon 500 D and Canon 350D cameras have then been considered, taking into account signal noise and demosaicing concerns. The results obtained have proved the practical usability of this approach, opening new technical possibilities for crop monitoring and agricultural robotics.

**Keywords:** NDVI, aerial imaging, multispectral, near-infrared band

## Nomenclature

UAV	unmanned aerial vehicle
RVI	ratio vegetation index
NDVI	normalised difference vegetation index
R, G, B, NIR	red, green, blue and near-infrared band digital count respectively
CFA	colour filter array
BSOP	band simulation by orthogonal projection
DC	digital count [ ]
$\lambda$	wavelength [nm]

$e(\lambda)$	spectral irradiance at wavelength $\lambda$ [ $\text{W}\cdot\text{m}^{-2}\cdot\text{nm}^{-1}$ ]
$s(\lambda)$	spectral sensitivity at wavelength $\lambda$ [ $\text{m}^2\cdot\text{W}^{-1}$ ]
$E$	irradiance vector for a set of $n$ wavelengths [ $\text{W}\cdot\text{m}^{-2}$ , $n$ dimensions]
$S$	sensitivity vector for a set of $n$ wavelengths [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$V_i$	sensitivity vector of the $i^{\text{th}}$ band of a sensor [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$V_t$	target sensitivity vector [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$V_{ti}$	$i^{\text{th}}$ target sensitivity vector [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$P(V)$	projection of the sensitivity vector $V$ on a given subspace [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$B$	base matrix of a set of $p$ sensitivity vectors [ $\text{m}^2\cdot\text{W}^{-1}$ , $n \times p$ dimensions]
$a_{it}$	$i^{\text{th}}$ coordinate of a projection $P(V_t)$ on the subspace base defined by matrix $B$ [ ]
$A_t$	vector of $a_{it}$ coordinates of a projection $P(V_t)$ [ $p$ dimensions ]
$\text{SAM}(V_1, V_2)$	spectral angle mapper between sensitivity vectors $V_1$ and $V_2$ [rad]
$[V_1^*, \dots, V_p^*]$	set of initial sensitivity vectors of the $p$ channels of a sensor [ $\text{m}^2\cdot\text{W}^{-1}$ , $n \times p$ dimensions]
$F(\lambda)$	transmittance value of an optical filter for the wavelength $\lambda$ [ ]
$F$	transmittance vector of an optical filter for a set of $n$ wavelengths [ $n$ dimensions]
$F_{\lambda_c}$	transmittance vector of a long-wave pass filter with a cutting wavelength $\lambda_c$ [ $n$ dimensions]
$B_{\lambda_c}$	base matrix of a set of $p$ sensitivity vectors modified by a filter $F_{\lambda_c}$ [ $\text{m}^2\cdot\text{W}^{-1}$ , $n \times p$ dimensions]
$P_{\lambda_c}(V)$	projection of the sensitivity vector $V$ on the subspace defined by $B_{\lambda_c}$ [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$Q_P(\lambda_c)$	cost function of the projection of a set of target sensitivity vectors on the subspace defined by $B_{\lambda_c}$ [rad]
$P_N(V)$	normalised projection of the sensitivity vector $V$ on a given subspace [ $\text{m}^2\cdot\text{W}^{-1}$ , $n$ dimensions]
$k$	normalisation factor applied to a projection $P(V)$ to obtain $P_N(V)$ [ ]
$k_i$	normalisation factor applied to the projection $P(V_{ti})$ of a target sensitivity $V_{ti}$ [ ]
$A_{Nt}$	vector of coordinates of a normalised projection $P_N(V_t)$ on a base matrix $B$ [ $p$ dimensions ]
$V_{t1}$	rescaled target sensitivity vector for the red band [ $n$ dimensions ]
$V_{t2}$	rescaled target sensitivity vector for the near-infrared band [ $n$ dimensions ]
$[V_1^*, V_2^*, V_3^*]_{\text{Canon}}$	rescaled initial sensitivity vectors of the Canon camera [ $n \times 3$ dimensions]
$[V_1^*, V_2^*, V_3^*]_{\text{Sigma}}$	rescaled initial sensitivity vectors of the Sigma camera [ $n \times 3$ dimensions]
$F_w$	transmittance vector of a long-wave pass Wratten filter n°25 [ $n$ dimensions]
$B_{\text{Canon}}$	base matrix of the Canon rescaled sensitivity vectors modified by the filter $F_w$ [ $n \times 3$ dimensions]

$\mathbf{B}_{\text{Sigma}}$	base matrix of the Sigma rescaled sensitivity vectors modified by the filter Fw [ $n \times 3$ dimensions]
$[a_1, a_2, a_3]_R$	coordinates of the normalised projection $P(V_{11})$ on the subspace base defined by matrix $\mathbf{B}$ [ ]
$[a_1, a_2, a_3]_{NIR}$	coordinates of the normalised projection $P(V_{12})$ on the subspace base defined by matrix $\mathbf{B}$ [ ]
EGI	excess green index [ ]
NEG	normalised excess green index [ ]
$r(\lambda), g(\lambda), b(\lambda)$	colour matching functions CIE 1931 [ ]
$C_1, C_2, C_3$	digital counts of the three channels of the modified Canon camera [ ]
$b_1, b_2, b_3$	additive noise of the three channels of the modified Canon camera [ ]
M	linear combination applied to the three digital counts of a camera [ ]
S	digital count resulting from the linear combination M [ ]
$b_S$	additive noise of the digital count S [ ]
$SNR_S$	signal noise ratio of the digital count S [ ]
NPI	noise propagation index [ ]
$NPI_R, NPI_{NIR}$	noise propagation index for the simulated digital counts R and NIR [ ]
$g_1, g_2$	digital counts issued from the two green patterns of the Bayer CFA [ ]
w	smoothing filter width [pixels]
$\lambda_w$	Wilks' lambda [ ]

## 1. Introduction

The normalised difference vegetation index, or NDVI, introduced by Rouse, Haas, Schell and Deering (1973), remains today a very popular tool in the remote sensing community that deal with agricultural monitoring. This is mainly due to its remarkable ability to discriminate green vegetation from other material in multispectral satellite images. The reflectance spectrum of any green vegetation material is characterised by a sharp transition between low reflectance in the visible domain (typically 10 to 20%), and high reflectance in the near-infrared domain (typically 50 to 80%). This difference is further accentuated if the red wavelengths in the visible domain due to chlorophyll absorption are considered.

Therefore, any scalar index built to exhibit the contrast between reflectance in red and near-infrared bands, using a multispectral sensor, will be a good candidate for vegetation discrimination, especially if this index is independent from the absolute digital count levels  $R$  (for the red band) and  $NIR$  (for the near-infrared band). The “ratio vegetation index”  $RVI = NIR/R$  is such a candidate. It can take values from zero to infinity, depending of the amount of vegetation in the observed area. The NDVI, expressed as:

$$NDVI = (NIR-R)/(NIR+R) = (RVI-1)/(RVI+1) \quad (1)$$

is a non-linear transformation of the RVI, transposing the same spectral information in the range  $[-1, +1]$ .

According to Jackson, (1991), NDVI is more sensitive to sparse vegetation densities in case of “mixed” pixels, as encountered in remote sensing imagery with low spatial resolution. Other vegetation indices have also been proposed in various soil and vegetation conditions, with respect to their relationship to agronomic indices such as biomass and LAI (leaf area index) (Huete, Liu, Batchily and van Leeuwen, 1997), (Zhengwei, Hu Zhao, Liping Di, Yu, 2009) (Jindong, Dong and Bauer, 2007). Most of these indices (ratio normalized difference vegetation index or RNDVI, soil-adjusted vegetation index or SAVI, transformed normalized difference vegetation index or TNDVI, etc.) rely on the same red and near-infrared bands. It confirms the better ability, already observed by Tucker (1979), of red and near-infrared band combination for crop monitoring, compared to combinations based on green and red bands in the visible domain (e.g.  $R/G$  or  $(R-G)/(R+G)$ , where  $G$  is the digital count for the green band).

In remote sensing, the popularity of NDVI and NDVI-derived vegetation indices has been widely supported by the availability of red and near-infrared channels on most satellite-embedded multispectral sensors (Landsat, SPOT, etc.). Unfortunately, this is not the case for crop monitoring applications at lower spatial scales: most vision systems embedded on ground vehicles or UAV are still based on standard colour cameras, leading to issues of robustness in vegetation detection and characterisation.

This is due to both technical and economic reasons. In satellite technology, image acquisition is based on line-scanning sensors, combined with satellite motion to get 2D images. A wavelength separation can thus easily be obtained through prisms or diffraction gratings associated with several parallel line sensors. However, ground or UAV vision systems generally use 2D acquisition sensors. Two main technical solutions are thus available to get multispectral channels from 2D systems:

- to separate the image into several ones behind the lens, using optical separators. A specific band path filter and CCD sensor can then be allocated to each image (multi-CCD). This solution offers some flexibility in the choice of the channel bands, but it requires accurate mechanical adjustments, leading to limited spatial resolution and high costs.
- to settle a colour filter array (CFA) directly on the CCD (or CMOS) sensor, so that neighbouring pixels are assigned to different channels. The most commonly used CFA is the Bayer (1976). It is widely used on commercial colour cameras, leading to low-cost devices, but it requires a very large production scales. Therefore, it cannot be envisaged for the vegetation monitoring market since it would require the development of specific sensors equipped with NIR micro-filter matrices.

Currently, some camera manufacturers propose multi-CCD devices including a NIR channel, e.g. the MS-4100 (Geospatial Systems Inc., West Henrietta, NY, USA), or the AD-080 (JAI AS, Copenhagen, Denmark).

However, the price of such devices remains prohibitive for most applications in the agricultural sector.

Moreover, their limited spatial resolution is hardly sufficient for UAV image acquisition.

To overcome this situation, some camera end users requiring a NIR channel have developed alternative solutions around standard Bayer matrix RGB cameras, taking advantage of an “undesirable” property of the silicon sensing array: because the colour filters in the Bayer matrix have a filtering action limited to the visible domain, the camera manufacturers are constrained to adding a near-infrared blocking filter to match standard colorimetric requirements. By removing this additional filter, a “modified” camera sensitive to near-infrared wavelengths can be obtained.

A possibility to obtain separate R and NIR channels is to simultaneously use a standard and a modified colour camera, the second being equipped with a near-infrared pass-band filter. However, important issues can arise concerning the pixel alignment of the two images obtained (Dare, 2008).

Another possible approach, which is the focus of this article, is to use a single modified camera associated with a long-wave pass filter, and to build the required R and NIR bands as specific linear combinations of the three resulting channels. This concept can be illustrated by the following.

On the assumption that an “ideal” modified RGB camera is being considered, where the three R,G,B channels deliver digital counts respectively equal to  $R+NIR$ ,  $G+NIR$ ,  $B+NIR$ , if a long-wave pass filter is added in front of the lens that blocks the blue wavelengths, then only the NIR component is received on the blue channel. This component can be subtracted from the other channel digital counts leading to R, G and NIR components.

In practise, the sensitivity of each channel cannot be modelled so simply, and a specific study is necessary to determine the best long-wave pass filter and the best linear combination for a given camera. This approach has been used by the Tetracam company (Tetracam Inc. Chatsworth, CA, USA) in its “agricultural cameras” (ADC series). However, the spatial resolution of these cameras (about 3 M pixels) does not meet the present standard of RGB still cameras ( $> 10$  M pixels).

The purpose of this paper is to investigate how this approach can be extended to the current generation of commercial RGB imaging sensors, in order to combine robust spectral information on vegetation with high spatial resolution. A generic method is proposed to determine the optimal long-wave pass filter and linear combination for virtually any modified RGB camera, provided its sensitivity curves are known. This method, named BSOP (band simulation by orthogonal projection), has been applied through spectral simulation with two commercial imaging devices representative of the recent developments in sensor technology, the Canon 500D (Canon, Tokyo, Japan) and the Sigma SD14 (Sigma, Kawasaki, Japan). In another section, the Canon camera, which gave the best simulation results, is tested in field conditions. Practical issues in terms of image noise and Bayer filter demosaicing are considered, and NDVI computation results are presented.

## **2. The BSOP approach: theory and spectral simulation**

The basic idea developed here is to simulate a desired spectral sensitivity (further referred to as “target sensitivity”) by a linear combination of the real spectral sensitivities available for a given sensor associated with a long-wave pass filter. In the following, this linear combination will be defined and a method to determine the optimal long-wave pass filter associated with the sensor will be proposed.

### **2.1 The BSOP theory**

### 2.1.1 Orthogonal projection of a sensitivity spectrum

The spectral sensitivity of a sensor channel (or band) can be characterised by a function  $s(\lambda)$  of the wavelength

$\lambda$ . It will generate, for a given spectral irradiance  $e(\lambda)$ , a digital count DC :

$$DC = \int_0^{\infty} e(\lambda) \cdot s(\lambda) \cdot d\lambda \quad (2)$$

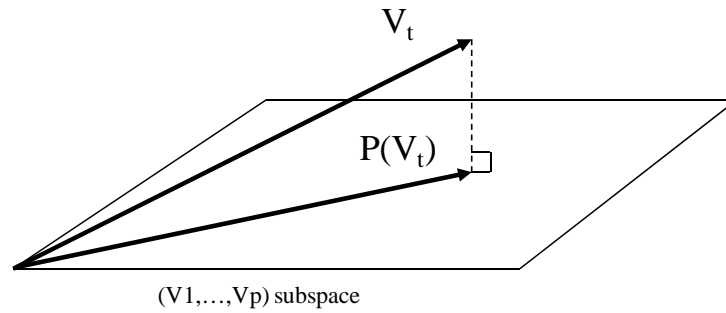
For practical reasons, a limited range of wavelengths and a limited spectral resolution  $\Delta\lambda$  will be considered, allowing any spectral function to be considered as a vector in a spectral space of dimension  $n$ , leading to the discrete expression:

$$DC = \sum_{i=1}^n e(\lambda_i) \cdot \Delta\lambda \cdot s(\lambda_i) = E \cdot S \quad (3)$$

where  $E = [e(\lambda_i) \cdot \Delta\lambda]$  and  $S = [s(\lambda_i)]$  are vectors of dimension  $n$  and  $E \cdot S$  is their scalar product.

Assuming  $V_1, \dots, V_p$  are the actual sensitivities of the  $p$  bands of an imaging sensor, by linear combination of  $V_1, \dots, V_p$ , any virtual sensitivity  $V$  belonging to the subspace generated by the vectorial base  $(V_1, \dots, V_p)$  can be simulated. Considering a given target sensitivity  $V_t$  that requires simulation, in the general case,  $V_t$  will not belong to this subspace, and the better approximation of  $V_t$  will be its orthogonal projection  $P(V_t)$  on the  $(V_1, \dots, V_p)$  subspace (Fig. 1).





**Fig. 1:** Illustration of the orthogonal projection of the target sensitivity  $V_t$

Defining  $\mathbf{B} = [V_1 \dots V_p]$  the matrix whose columns are the  $V_1, \dots, V_p$  vectors, the orthogonal projection of  $V_t$  on  $\mathbf{B}$  can be expressed as:

$$P(V_t) = \mathbf{B} \cdot (\mathbf{B}^T \mathbf{B})^{-1} \cdot \mathbf{B}^T \cdot V_t \quad (4)$$

The coordinates of  $P(V_t)$  in the base  $(V_1 \dots V_p)$ , i.e. the coefficients of the linear combination giving  $P(V_t)$  from  $V_1 \dots V_p$ , are given by the vector  $A_t$  of dimension  $p$ :

$$A_t = [a_{t1} \dots a_{tp}]^T = (\mathbf{B}^T \mathbf{B})^{-1} \cdot \mathbf{B}^T \cdot V_t \quad (5)$$

According to Eqs. (4) and (5):

$$P(V_t) = \mathbf{B} \cdot A_t = [V_1 \dots V_p] \cdot A_t = a_{t1} \cdot V_1 + \dots + a_{tp} \cdot V_p \quad (6)$$

### 2.1.2 Optimal long-wave pass filter

Whatever the set of spectral sensitivities  $V_1 \dots V_p$  and the target sensitivity  $V_t$ , the projection vector  $P(V_t)$  defined above will always exist. However, this does not guarantee a satisfactory result. It has to be verified that  $P(V_t)$  is close to the target vector  $V_t$ . A commonly used measure of the similarity of spectral data is the SAM, or spectral angle mapper (Yuhas, Goetz and Boardman, 1992). This measure evaluates the angle between two vectors  $X_1$  and  $X_2$  of the Euclidian norms  $|X_1|, |X_2|$ , as:

$$\text{SAM}(X1,X2) = \text{acos} (X1.X2/(|X1|.|X2|)) \quad (7)$$

In this case, the value  $\text{SAM}(V_t, P(V_t))$  should be ideally equal to zero, meaning that  $V_t$  belongs to the subspace  $(V1, \dots, Vp)$ . In the real case, where the initial spectral sensitivities of the different channels are a specification of the sensor, there is no *a priori* reason to meet this requirement. Defining  $V1^*, \dots, Vp^*$  as these initial sensitivities, the only degree of freedom to optimise the situation consists in adding an optical filter (e.g. by setting it in front of the lens) which has a spectral transmittance  $F$ , leading to a modified set of sensitivities:

$$(V1, \dots, Vp) = (F.V1^*, \dots, F.Vp^*) \quad (8)$$

Thus, our objective is to determine the optical filter  $F$  that will minimise the SAM between the target  $V_t$  and its projection  $P(V_t)$ , according to Eqs. (4) and (8).

A possible method would be to directly search for an optimal filter  $F$  in the spectral space by minimisation techniques. Such a method would be rather complex to develop and it could lead to non-realisable optical filters, unless constraints of positivity and smoothness were introduced. In the present case, we have chosen to reduce the complexity of the problem by considering only simple long-wave pass filters as  $F$  candidates. The main advantage of this approach is the possibility of implementing the solution with off-the-shelf gelatine or glass filters.

Let us define a candidate long-wave pass filter by its cutting wavelength  $\lambda_c$ :

$$F_{\lambda_c}(\lambda) = 1 \text{ if } \lambda > \lambda_c ; F_{\lambda_c}(\lambda) = 0 \text{ otherwise;} \quad (9)$$

and let us consider  $k$  target sensitivities  $V_{t1}, \dots, V_{tk}$ .

Then for a given wavelength  $\lambda_c$  the following can be computed:

- the subspace matrix  $B_{\lambda_c} = [V1_{\lambda_c}, \dots, Vp_{\lambda_c}] = [V1^*. F_{\lambda_c}, \dots, Vp^*. F_{\lambda_c}]$
- the projected vectors  $P_{\lambda_c}(V_{t1}), \dots, P_{\lambda_c}(V_{tk})$ , according to Eq. (4)
- a global cost function taking into account the similarity between every target and its projection:

$$Q_P(\lambda c) = \text{SAM}(V_{t1}, P_{\lambda c}(V_{t1})) + \dots + \text{SAM}(V_{tk}, P_{\lambda c}(V_{tk})) \quad (10)$$

The optimal long-wave pass filter  $F_{\lambda c}$  will be the one that minimise  $Q_P(\lambda c)$ .

### 2.1.3 Balance of the projected vectors

The SAM criterion above was used to determine a set of vectors  $P(V_{t1}), \dots, P(V_{tp})$  matching as well as possible an initial set  $V_{t1}, \dots, V_{tp}$  of target sensitivities, in terms of spectral shape. Another important point, especially if these vectors are devoted to the computation of ratio indices such as NDVI, is to balance their relative amplitude, so that the ratios of their digital counts remain as close as possible to the original. Though this cannot obviously be obtained for every irradiance spectrum, an approximate solution is to ensure that the projected vector and the target vector have the same L1-norm<sup>1</sup>.

Therefore, the following renormalisation is finally applied to each projected vector:

$$\forall i \in [1, p], \quad P_N(V_{ti}) = P(V_{ti}) \cdot (|V_{ti}|_{L1} / |P(V_{ti})|_{L1}) = P(V_{ti}) \cdot k_i \quad (11)$$

This set of balance correction factors ( $k_i$ ) have also to be applied to the linear coefficients defined in (5):

$$A_{Nti} = k_i \cdot A_{ti} \quad (12)$$

## 2.2 BSOP spectral simulation

In this subsection, the theoretical concepts described above are applied to a set of commercial cameras in order to evaluate the possibility to simulate simultaneously the R and NIR bands. For this purpose, after laboratory measurement of the camera sensitivities, optimal linear combinations are determined, as well as their application to NDVI computation on a set of vegetation and soil spectral data.

<sup>1</sup> The L1-norm of a vector is defined as the sum of the absolute values of its components. If its components are all positive, the L1-norm of a sensitivity vector is equal to its scalar product with a flat spectrum  $E_f = (1, \dots, 1)$ , i.e. is equal to its digital count for a non-colored irradiance source  $E_f$ .

## 2.2.1 Material and methods

### 2.2.1.1 Camera sensitivity measurement

Two commercial still cameras were selected for the BSOP assessment: the Canon 500D (Canon, Tokyo, Japan), representative of recent high resolution Bayer matrix sensors (15 M pixels), and the Sigma SD14 (Sigma, Kawasaki, Japan), for the original technology of its Foveon X3 sensor. The Foveon sensor is made of three separate layers of photodetectors. Since silicon absorbs different wavelengths at different depths, each layer captures a different colour and no Bayer matrix is required, leading to a better spatial resolution. However, there is less control on the spectral sensitivity curve of each channel.

Each camera was opened and its near-infrared blocking filter removed. The spectral sensitivity of each “modified” camera was then measured in the range 440-990 nm, with 10 nm steps. For this purpose, a tunable monochromatic light source of a laboratory spectrometer (V-570, Jasco Inc, Easton, USA) was placed in front of the camera lens via an optical fiber, at a 300 mm distance. Images were taken for each wavelength in the following conditions: no ambient light; focal length 50 mm; sensitivity 100 ISO; Integration time 5 s.

The average level of the light source image was collected for each channel using the raw camera images (see 3.1.3 for more detail on raw images), using home-made image analysis software.

It should be noted that due to the numerous parameters involved (intensity of the light source, camera exposure, etc.), the procedure described here cannot be considered as a radiometric measurement. Sensitivity curves defined in Eq. (3), theoretically expressed as  $\text{m}^2 \cdot \text{W}^{-1}$ , are obtained here up to an arbitrary scaling factor.

### 2.2.1.2 Definition of the target sensitivities

Two target sensitivities  $V_{t1}$ ,  $V_{t2}$  corresponding respectively to the R and NIR bands have been considered. Because no standardised sensitivity data are clearly defined in the literature for the NDVI computation (some authors use the bands TM3 and TM4 of Landsat, others simply use 660 nm and 760 nm wavelengths), the following procedure has been used:

- the R band has been derived from the CIE 1931 colorimetric data (CIE, 2004), by taking the positive part of the standard curve  $\bar{r}(\lambda)$  and applying a 30 nm shift, leading to a bandwidth 600-670 nm at mid-height.

- the NIR band was computed by a 160 nm shift of the R band, leading to a bandwidth 760-830 nm at mid-height.

Both target sensitivity curves have been reported on Fig. 5.

### 2.2.1.3 BSOP computation

In order to be used on real hyperspectral data, all the data defined above were resampled according to the spectral resolution of a hyperspectral camera Hypspx VNIR-1600 (Norsk Elektro Optikk A/S, Norway), i.e. 160 spectral bands from 415 to 993 nm.

According to the notations defined in section 2.1, this led to two sets of vectors  $V1^*$ ,  $V2^*$ ,  $V3^*$  (one for each type of camera) and two vectors  $V_{t1}$ ,  $V_{t2}$  of dimension 160. All related computations were made using Matlab 7 (The MathWorks, Natick, MA, USA).

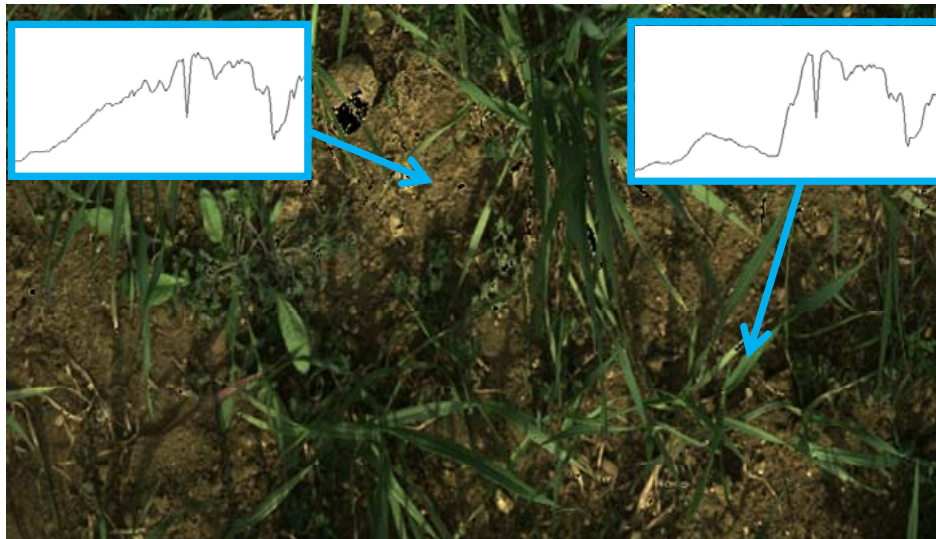
Once the optimal cutting wavelength was determined, the closest existing gelatin Wratten filter was chosen. Its actual transmittance curve was measured with the Jasco V-570 spectrometer, and resampled according to the Hypspx 160 bands, leading to a filter vector  $Fw$ . Finally, for each type of camera, the projected vectors  $P_N(V_{t1})$  and  $P_N(V_{t2})$  on the subspace  $(Fw.V1^*, Fw.V2^*, Fw.V3^*)$ , as well as the corresponding coefficients  $A_{t1}$  and  $A_{t2}$ , were computed and renormalised according to Eqs. (11) and (12).

### 2.2.1.4 Simulation on field hyperspectral images

Hyperspectral images of wheat durum were acquired in experimental fields (INRA, Domaine de Melgueil, France) in March 2011, using the Hypspx VNIR-1600 camera. The camera was set on a motorised translation rail one metre above the ground (Vigneau, Ecarnot, Rabatel and Roumet, 2011). The images included durum wheat at an early stage, as well as various types of dicotyledonous and monocotyledon weeds. A total of 2210 luminance spectra were then collected in the hyperspectral images by manually selecting areas, including wheat, weeds and soil categories (Fig. 2).

For every collected spectrum  $S$ , the digital counts  $R$  and  $NIR$  corresponding to red and near-infrared target sensitivities  $V_{t1}$ ,  $V_{t2}$  were computed using Matlab as  $R = S \cdot V_{t1}$  and  $NIR = S \cdot V_{t2}$ , according to Eq. (3). A reference NDVI value was computed using these  $R$  and  $NIR$  values.

The same operation was made using  $P_N(V_{t1})$  and  $P_N(V_{t2})$  instead of  $(V_{t1}, V_{t2})$  to compute an estimated NDVI for each type of still camera.



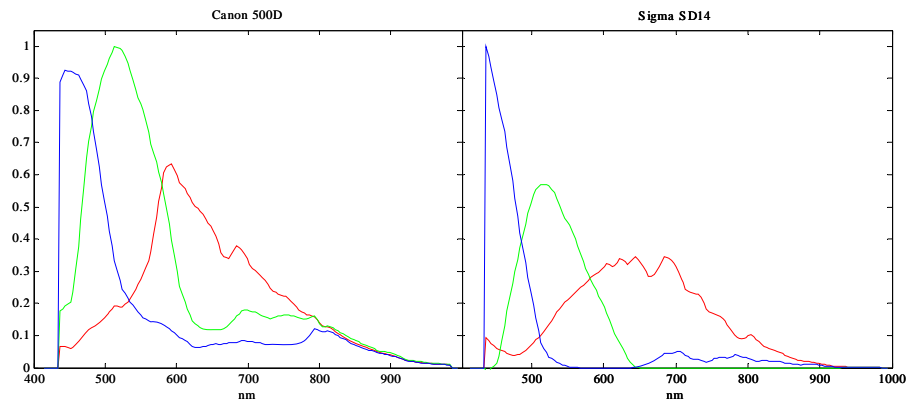
**Fig. 2.** Luminance spectrum examples on a durum wheat crop hyperspectral image

(NB: the atmospheric absorption peaks can be observed on the spectra)

## 2.2.2 Simulation results

### 2.2.2.1 Camera sensitivity measurement

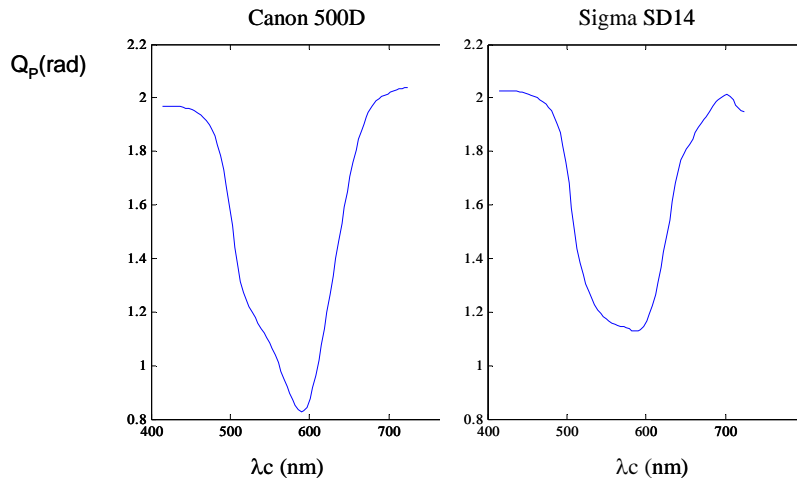
As mentioned in section 2.2.1.1, sensitivity spectra were obtained up to an arbitrary scale factor. They were rescaled so that their maximum value was equal to 1 (initially 1033 and 1632.8 for the Canon and the Sigma cameras, respectively). These rescaled sensitivity curves  $[V_1^*, V_2^*, V_3^*]_{\text{Canon}}$  and  $[V_1^*, V_2^*, V_3^*]_{\text{Sigma}}$ , which were used for further computations, are given in Fig. 3.



**Fig.3.** Spectral sensitivities of the Canon 500D and Sigma SD14 without near-infrared blocking filter  
(sensitivities were rescaled for a maximum value equal to unity)

### 2.2.2.2 Optimal long-wave pass filter

In Fig.4, the BSOP quality curves  $Q_p(\lambda_c)$  according to Eq. (10) is shown. It can be seen that in both cases,  $Q_p(\lambda_c)$  is undefined for  $\lambda_c > 728$  nm, because the projection of  $Vt1$  (i.e. red band) becomes null. The best BSOP quality is obtained for  $\lambda_c \approx 600$  nm, and is much better for the Canon camera.



**Fig 4.** BSOP quality  $Q_p(\lambda_c)$  of Canon and Sigma cameras

### 2.2.2.3 Red and near-infrared bands

According to the results above, a standard Wratten filter Kodak n°25 (gelatin, cutting wavelength: 600 nm) was selected. The actual transmittance Fw was measured with the Jasco spectrometer, and scaled for a maximal value equal to 1, according to Eq. (9). The corresponding projection base for each camera was then computed according to Eq. (8):

$$B_{\text{Canon}} = [V_1 \ V_2 \ V_3]_{\text{Canon}} = [Fw \cdot V_1^*, Fw \cdot V_2^*, Fw \cdot V_3^*]_{\text{Canon}} \quad (13)$$

$$B_{\text{Sigma}} = [V_1 \ V_2 \ V_3]_{\text{Sigma}} = [Fw \cdot V_1^*, Fw \cdot V_2^*, Fw \cdot V_3^*]_{\text{Sigma}} \quad (14)$$

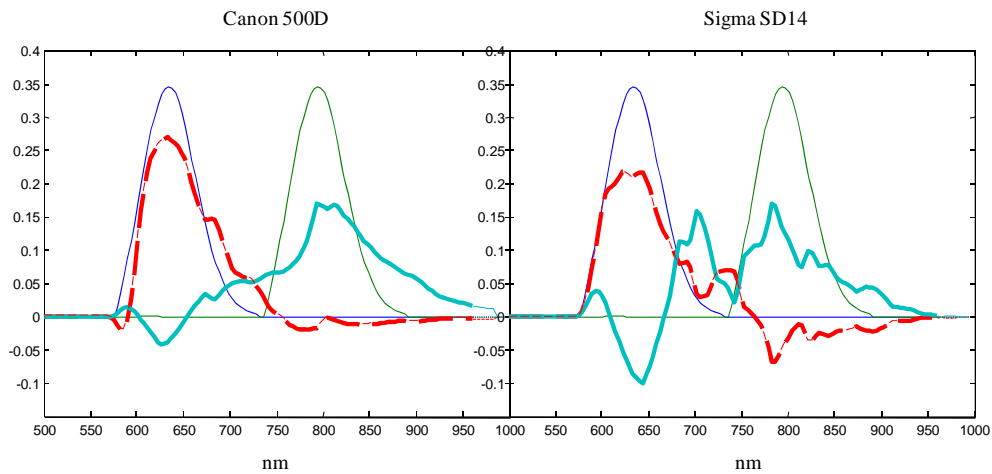
Finally, the BSOP linear coefficients and balance factors were computed according to Eqs. (4), (5) and (12). They are reported in Table 1 along with the SAM values between target sensitivity and projection:



	R					NIR				
	k	SAM (rad)	a1	a2	a3	k	SAM (rad)	a1	a2	a3
Canon 500D	0.876	0.273	<b>0.863</b>	<b>-1.430</b>	<b>0.620</b>	0.743	0.588	<b>-0.333</b>	<b>0.0902</b>	<b>1.745</b>
Sigma SD14	0.774	0.321	<b>0.668</b>	<b>0.008</b>	<b>-3.295</b>	0.858	0.783	<b>-0.324</b>	<b>0.811</b>	<b>5.035</b>

Table 1: Balance correction factors (k), SAM( $V_t$ ,  $P(V_t)$ ) and linear coefficients ( $a_i$ )  
for simulated R and NIR bands with a filter Wratten N°25

It should be noted that the coefficients ( $a_i$ ) given in Table 1 are up to a scale factor, according to the scaling applied on matrices  $B_{Canon}$  and  $B_{Sigma}$  above. However, the corresponding simulated sensitivity curves, computed using Eqs. (4) and (12), do not depend on the  $B_{Canon}$  and  $B_{Sigma}$  scaling. These curves are given in Fig.5, as well as the reference, for comparison. It appears that for both cameras, the simulated NIR band is the less satisfactory since its domain is not limited to wavelengths above 750 nm with wavelengths from 600 to 750 nm also contributing both negatively and positively. This was particularly noticeable for the Sigma camera, where the contribution of lower wavelengths was as high as the contribution of near-infrared wavelengths. These observations confirm the different SAM values reported in Table 1, recalling that the sum of these two values for a given camera corresponded to the minimal value of  $Q_p(\lambda_c)$  in Fig. 4.



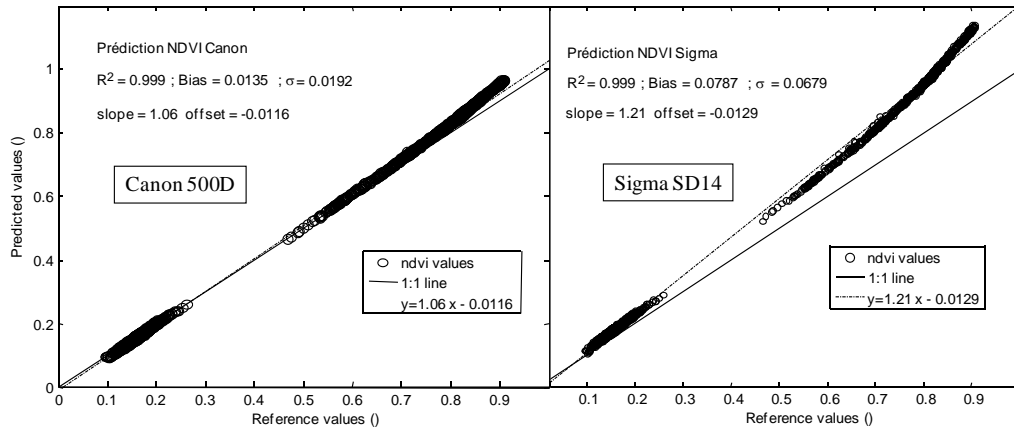
**Fig. 5.** BSOP bands for Canon and Sigma cameras with Wratten filter n°25. Thin lines: target sensitivities  $V_{t1}$ ,  $V_{t2}$ . Thick dash line: BSOP red band; thick solid line: BSOP near-infrared band

#### 2.2.2.4 NDVI computation

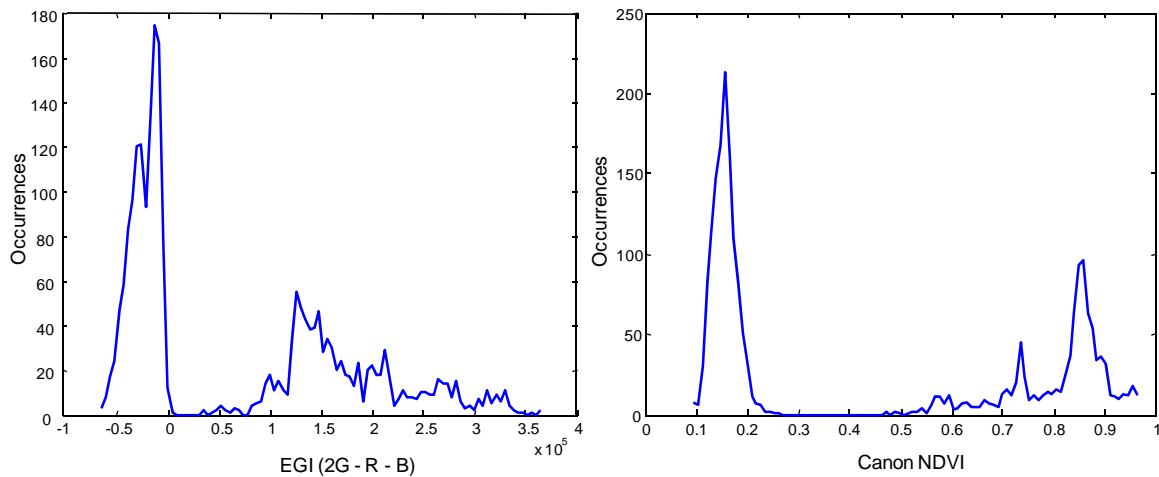
Figure 6 shows the NDVI values computed with the simulated bands of Fig.5, versus the reference NDVI values computed with the target sensitivities  $V_{t1}$ ,  $V_{t2}$ , for the 2210 luminance spectra collected in the durum wheat crop (see section 2.2.1.4).

It can be seen that there are two clearly separated groups of NDVI values. They correspond respectively to soil and vegetation (wheat, monocotyledons and dicotyledons), confirming the remarkable efficiency of NDVI for crop-soil discrimination. For comparison purposes, Fig.7 shows the histograms of EGI or excess green index (Woebbecke, Meyer, Von Bargen & Mortensen, 1995) and NDVI for the whole set of samples, and the better separation between the two groups obtained with NDVI (EGI was simulated for each luminance spectrum using the CIE 1931 standard curves  $\underline{r}(\lambda)$ ,  $\underline{g}(\lambda)$ ,  $\underline{b}(\lambda)$  as mentioned in section 2.2.1.2).

The prediction quality using the Canon camera was significantly better since the NDVI values obtained are nearly equal to the reference except for the highest values ( $NDVI > 0.8$ ) where the error remains less than 10%.



**Fig. 6** Predicted versus reference values of NDVI



**Fig. 7** EGI histogram versus Canon NDVI histogram

### 2.2.2.5 Choice of the Canon camera

From the results above, it appears that the spectral characteristics of the Sigma sensor were not appropriate for a BSOP of R and NIR bands. Concerning the Canon camera, although the shape of simulated bands was not completely satisfactory, the good results in terms of NDVI computation motivated further assessments in real field conditions. This is the purpose of the next section.

### **3. The BSOP approach: practical implementation**

It was shown in the previous section that an efficient simulation of R and NIR bands could be obtained from a Canon 500 D camera after removing its near-infrared blocking filter. The study was based on laboratory measurement of camera characteristics and spectral simulation of image acquisition process. In the present section, the situation under field conditions will be investigated, dealing directly with image digital counts and taking in account practical aspects such as signal-noise ratio and channel demosaicing.

#### **3.1 Material and methods**

##### **3.1.1 Image acquisition**

Various aerial images were acquired using modified cameras (no near-infrared blocking filter). Depending on their availability, either a Canon 500D or a Canon 350D was used, both use similar sensors. A 58 mm red glass filter Hama Rot R8 (25 A) (Hama GmbH & Co, Monheim, Germany) was placed in front of the lens as a long-wave pass filter. Although this filter was specified as being equivalent to the Wratten n°25, its transmittance curve was measured in the laboratory using the Jasco V-570, and the corresponding BSOP coefficients computed again using the Canon 500D sensitivity data. Images were acquired by a private company (AvionJaune, Montferrier, France) using microlight aircraft or UAV, depending on the flight altitude.

For comparison purpose, the same aerial scenes were also acquired using standard RGB cameras. In order to illustrate the plant/soil discrimination ability in both cases, the following procedure has been followed:

- for images acquired with the modified cameras, NDVI values were computed using the R and NIR digital counts obtained from the procedure described in section 2.
- for standard RGB images, the  $EGI = 2G - R - B$  as well as its normalised version  $NEG = EGI / (R + V + B)$  (normalised excess green) were computed. RGB channels were computed from raw images and exported as TIFF files (Tagged Image File Format), in order to avoid JPEG (Joint Photographic Experts Group) compression artefacts.
- NDVI, EGI and NEG images were then computed by scaling their index values in the range 0-255 (8 bits), and thresholded automatically according to their grey level histogram. For this purpose, the Otsu method was used (Otsu, 1979), which minimises the ratio between inter-class inertia and total inertia for the two classes

issued from thresholding. This ratio, known as Wilks' lambda ( $\lambda_w$ ) in multivariate analysis, gives an indication on the thresholding quality and robustness.

In the following examples, NEG results will not be presented, because their  $\lambda_w$  values were always lower than those obtained with EGI. In every case, the focus of the modified camera was set manually. An example of aerial scene is given in Fig. 8.



**Fig. 8** Example of aerial image (Canon 500D, altitude 1000m, vineyard area, September 2011, south of France))

Manual rectangles show selected areas for further detailed analyses

### 3.1.2 Channel linear combination and signal-noise ratio

In section 2, digital counts R and NIR were computed from hyperspectral data according to equation (3). Under field conditions, they have to be computed directly from the digital counts of the camera raw channels, using the linear coefficients  $[a_1, a_2, a_3]_R$  and  $[a_1, a_2, a_3]_{NIR}$  issued from Eqs. (5) and (12).

Defining  $[V_1 V_2 V_3]$  as the raw sensitivity curves of the camera, and  $e(\lambda)$  the irradiance spectrum for a given pixel, the digital count corresponding to a linear combination  $[a_1 a_2 a_3]$  is given by:

$$DC = \int_0^{\infty} e(\lambda) \cdot [a_1 \cdot V_1(\lambda) + a_2 \cdot V_2(\lambda) + a_3 \cdot V_3(\lambda)] d\lambda \quad (15)$$

$$\text{i.e. } DC = a_1 \cdot \int_0^{\infty} e(\lambda) \cdot V_1(\lambda) \cdot d\lambda + a_2 \cdot \int_0^{\infty} e(\lambda) \cdot V_2(\lambda) \cdot d\lambda + a_3 \cdot \int_0^{\infty} e(\lambda) \cdot V_3(\lambda) \cdot d\lambda = a_1 \cdot C_1 + a_2 \cdot C_2 + a_3 \cdot C_3 \quad (16)$$

where  $C_1, C_2, C_3$  are the digital counts of the raw channels for this pixel, after integration by the camera sensor.

As explained in section 2.2.2.3, the coefficients  $[a_1, a_2, a_3]_R$  and  $[a_1, a_2, a_3]_{NIR}$  range up to a scale factor.

However, especially when computing NDVI, the important point is not the absolute value of the NIR and R digital counts, but their signal-noise ratio (SNR) which is linked to the SNR of raw signals and to the way noise is propagated through  $a_1, a_2, a_3$  coefficients.

Let us consider again a linear combination  $M = [a_1 \ a_2 \ a_3]$  to be applied to the raw digital counts  $(C_1, C_2, C_3)$ , and let us call  $S = M \cdot [C_1 \ C_2 \ C_3]^T$  the resulting digital count. Now let us consider that the signals  $C_1, C_2, C_3$  are respectively affected by the additive, centered noises  $b_1, b_2, b_3$ . Then  $S$  will be affected by an additive noise  $b_s$  such that:

$$S + b_s = M \cdot [C_1 + b_1 \ C_2 + b_2 \ C_3 + b_3]^T = S + M \cdot [b_1 \ b_2 \ b_3]^T \quad (17)$$

The variance of  $b_s$  can be computed as:

$$b_s^2 = \langle [b_1 \ b_2 \ b_3] \cdot (M^T \cdot M) \cdot [b_1 \ b_2 \ b_3]^T \rangle = \sum_{i,j} \langle b_i \cdot b_j \rangle \cdot a_i \cdot a_j \quad (18)$$

If we make the hypothesis that  $b_1, b_2, b_3$  are independent noises with the same variance  $b^2$ , it comes:

$$b_s^2 = \sum_{i,j} \langle b_i \cdot b_j \rangle \cdot a_i \cdot a_j = \sum_{i,j} \delta_{ij} b^2 \cdot a_i \cdot a_j = \sum_i b^2 \cdot a_i^2 = b^2 \cdot \text{Trace}(M^T \cdot M) \quad (19)$$

Now let us consider the square of the signal  $S$ :

$$S^2 = [C_1 \ C_2 \ C_3] \cdot (M^T \cdot M) \cdot [C_1 \ C_2 \ C_3]^T = \sum_{i,j} C_i \cdot C_j \cdot a_i \cdot a_j \quad (20)$$

leading to a SNR:

$$SNR_S = \sqrt{\frac{S^2}{b_s^2}} = \sqrt{\frac{\sum_{i,j} C_i \cdot C_j \cdot a_i \cdot a_j}{b^2 \cdot \text{Trace}(M^T \cdot M)}} \quad (21)$$

In the case where  $C_1 = C_2 = C_3 = c$ , we can directly express  $SNR_S$  from the raw signal-noise ratio  $c/b$ :

$$SNR_S = \sqrt{\frac{\sum_{i,j} C_i \cdot C_j \cdot a_i \cdot a_j}{b^2 \cdot \text{Trace}(M^T \cdot M)}} = \sqrt{\frac{c^2 \sum_{i,j} a_i \cdot a_j}{b^2 \cdot \text{Trace}(M^T \cdot M)}} = \frac{c}{b} \sqrt{\frac{\sum_{i,j} a_i \cdot a_j}{\text{Trace}(M^T \cdot M)}} = NPI \cdot \frac{c}{b} \quad (22)$$

with

$$NPI = \sqrt{\frac{\sum_{i,j} a_i \cdot a_j}{\text{Trace}(M^T \cdot M)}} = \sqrt{1 + \frac{\sum_{i \neq j} a_i \cdot a_j}{\text{Trace}(M^T \cdot M)}} \quad (23)$$

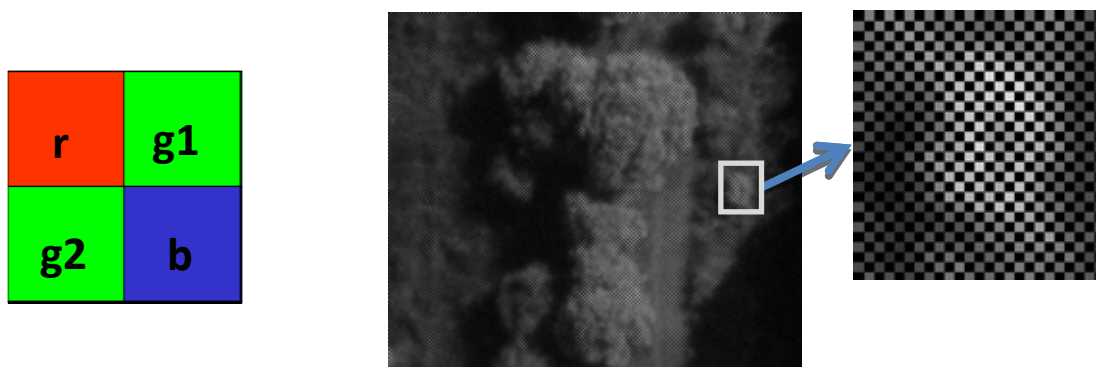
The value  $NPI$  (noise propagation index), which only depends on matrix  $M$ , can be considered as an indicator of the SNR propagation due to the linear combination. Its value can be higher or lower than unity, depending on the sign of the cross-terms  $a_i a_j$  ( $i \neq j$ ). While addition of the raw channels will increase the SNR (making a noise average), subtraction operations reduce the output signal level for a given noise level and thus decrease the SNR.

The  $NPI$  value will be considered for both R and NIR linear combinations in the results section.

### 3.1.3 Raw image importation and demosaicing

Most commercial still cameras provide images in JPEG format. Such a format is not appropriate in our application for several reasons. Radiance signals are coded on 8 bits (256 grey values), limiting the accuracy of

further linear combinations, and internal colorimetric corrections are applied to the each channel. Also, JPEG format includes spatial image compression. However, up market cameras also provide “raw” images, which deliver exactly the sensor output in an uncompressed format. In Canon products, the raw images are available as files that can easily be read using open source software such as DCraw (<http://www.cybercom.net/~dcoffin/dcraw/>). The result is a set of four images corresponding to each position in the Bayer filter pattern (Fig. 9). Pixel values are coded with 16 bits, the effective range depending on the camera (14 bits for Canon 500D). As an option, the dark current level of the CCD sensor can be automatically measured in image margins and subtracted from each pixel’s value.



**Fig. 9:** Bayer filter pattern (left) and image detail for one raw channel (right)

Before using raw channels, a demosaicing step is necessary, which consists in filling black pixels of each channel by interpolating neighbouring active pixels. For this purpose, many interpolation algorithms have been specifically developed (King-Hong & Yuk-Hee, 2006) (Hirakawa & Parks, 2005). However, such algorithms are mainly designed to optimise the picture quality in terms of human visual perception, rather than the absolute pixel value accuracy.

In our case, a demosaicing procedure based on spline interpolation (Thevenaz, Blu, Unser, 2000) was implemented as following:

- each raw channel image was reduced by 2x2 keeping only the active pixels



- the reduced image was smoothed, according to Nyquist–Shannon sampling theorem requirements (Marks, 1991), using a 1<sup>st</sup> order Deriche (1990) recursive filtering. This filter acts like a convolution mask  $g(x,y)$  expressed as:

$$g(x, y) = \alpha \cdot \left(\frac{|x|}{w} + 1\right) \cdot e^{-\frac{|x|}{w}} \cdot \alpha \cdot \left(\frac{|y|}{w} + 1\right) \cdot e^{-\frac{|y|}{w}} \quad (24)$$

where  $(x,y)$  are image coordinates,  $\alpha$  is a normalization factor, and  $w$  is the convolution width. A main interest of this type of filter is that  $w$ , expressed in pixels, may be adjusted to any floating point value with no computational change.

- missing pixel values in the original image were then computed by floating point interpolation in the reduced image
- finally, both processed images issued from  $g_1$  and  $g_2$  were averaged to provide the DC count  $C_2$ , while  $C_1$  and  $C_3$  were directly given by the processed images issued from  $r$  and  $b$  respectively.

It should be noted that after demosaicing, the two images issued from  $g_1$  and  $g_2$  should be identical for every pixel. This provides a way to assess the demosaicing quality, as illustrated further (section 3.2.2).

### 3.2 Results and discussion

#### 3.2.1 Channel linear combination and signal noise ratio

The linear coefficients for the Hama filter as well as the noise propagation indices according to Eq. (23) have been computed, and are given in Table 2.

R				NIR			
a1	a2	a3	NPI <sub>R</sub>	a1	a2	a3	NPI <sub>NIR</sub>
0.9744	-1.7329	0.8477	<b>0.0413</b>	-0.3761	0.0082	2.1522	<b>0.8167</b>

Table 2 : Linear coefficients and noise propagation index for Canon 500D with Hama red filter

It can be seen that the linear coefficients with the glass filter are slightly different from the ones obtained with the gelatine filter in section 2 (indeed, a detailed observation shows significant differences in the shape of their transmittance curves above the 600 nm cutting wavelength). Moreover, it is important to notice that the SNR propagation index  $NPI_R$  is very low. This is due to the fact that the R band is roughly obtained by subtracting the second channel from the sum of the other ones (see coefficients  $a_1, a_2, a_3$ ), leading to a significant signal reduction. However, the NIR digital count was mainly due to the third channel contribution, leading to a better  $NPI_{NIR}$  value.

In the following sections, all R and NIR bands were computed using these coefficients, and their values truncated to zero if negative.

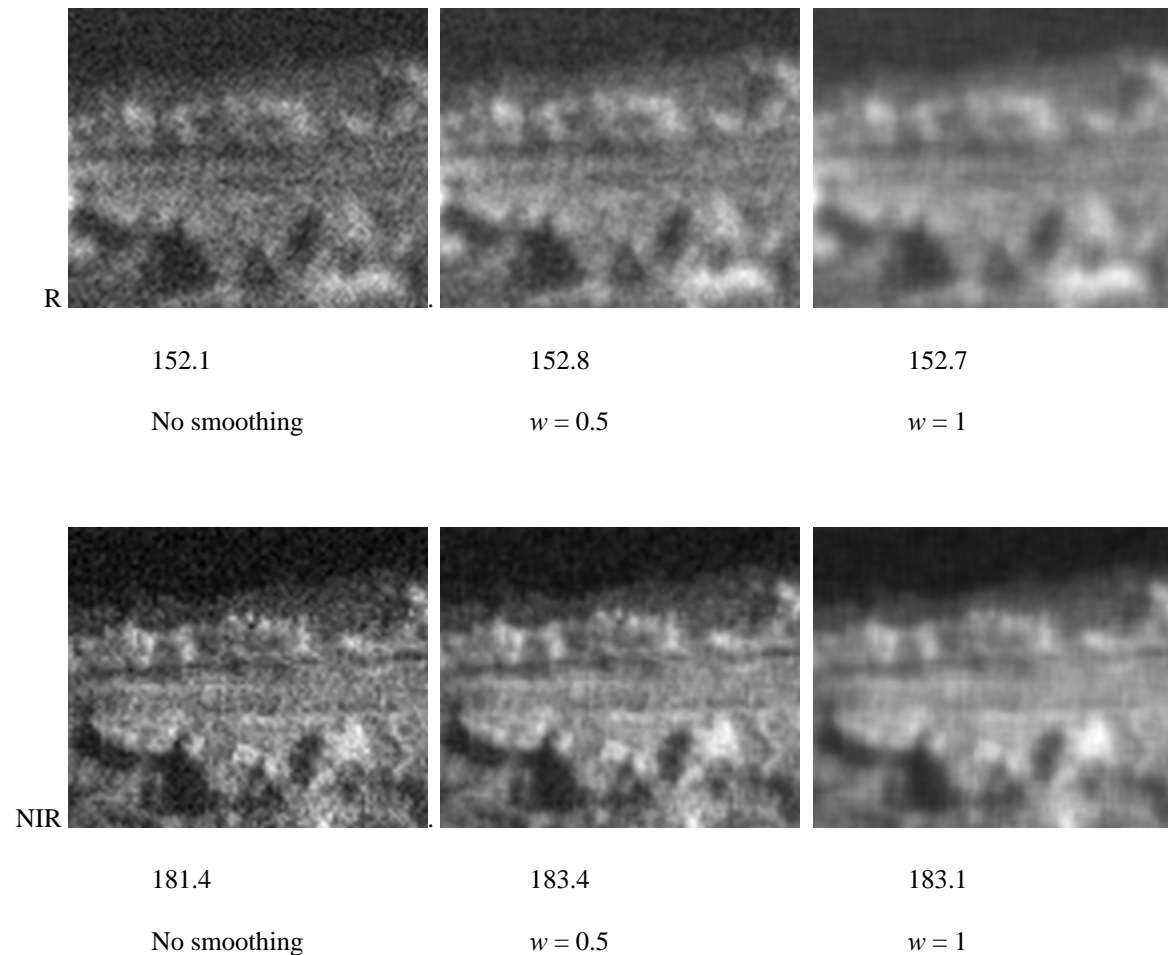
### 3.2.2 Image demosaicing

In order to tune the smoothing step in the demosaicing procedure, a small portion of the 1000 m altitude aerial image (Fig. 7) with significant local variations and similar levels on computed R and NIR bands has been selected. The standard deviation of the difference between the two raw channels  $g_1$  and  $g_2$ , after demosaicing, was reported for various values of the smoothing filter width  $w$ . The results are given in Table 3.

Smoothing	$g_1$ channel average	Standard deviation of ( $g_1-g_2$ )	Ratio (demosaicing SNR)
none	267.69	17.22	15.54529617
$w = 0.1$	270.86	17.32	15.63856813
$w = 0.2$	270.74	16.18	16.73300371
$w = 0.5$	270.08	9.51	28.39957939
$w = 1$	269.84	4.66	57.9055794
$w = 2$	270.15	1.99	135.7537688

Table 3: Standard deviation between raw channels  $g_1$  and  $g_2$  after demosaicing on a textured image area for various smoothing widths, and corresponding SNR

Figure 10 shows some examples of R and NIR computation for this image portion. It confirms that the R band exhibits more demosaicing noise (due to low  $NPI_R$ ), and that a smoothing width of at least 0.5 pixels is necessary.



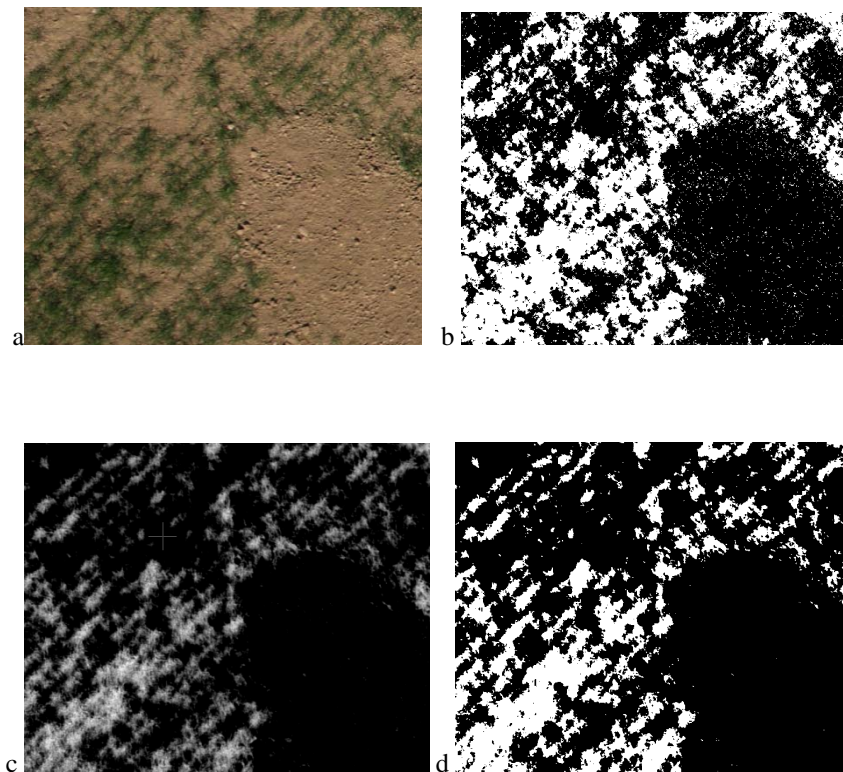
**Fig. 10:** Computed R and NIR bands for various smoothing widths

up: R band ; low: NIR band. Average band level (before 8 bits image scaling) is reported under each image

### 3.2.3 NDVI computation

Figure 11 shows an example of NDVI computation and thresholding using a single modified Canon 350D camera, as well as standard colour acquisition for comparison. Images have been acquired in March 2012 on a

wheat field plot (INRA, Mauguio, France) through two separate UAV flights. The flight altitude was about 30 m, and the spatial resolution was 10 mm. Only portions of the image are shown here.



**Fig. 11:** EGI and NDVI computation on a wheat parcel. a: colour image; b: EGI thresholding ( $\lambda_W=0.66$ ); c : NDVI; d: NDVI thresholding ( $\lambda_W=0.81$ ). Bare soil zone is due to moisture sensor implementations

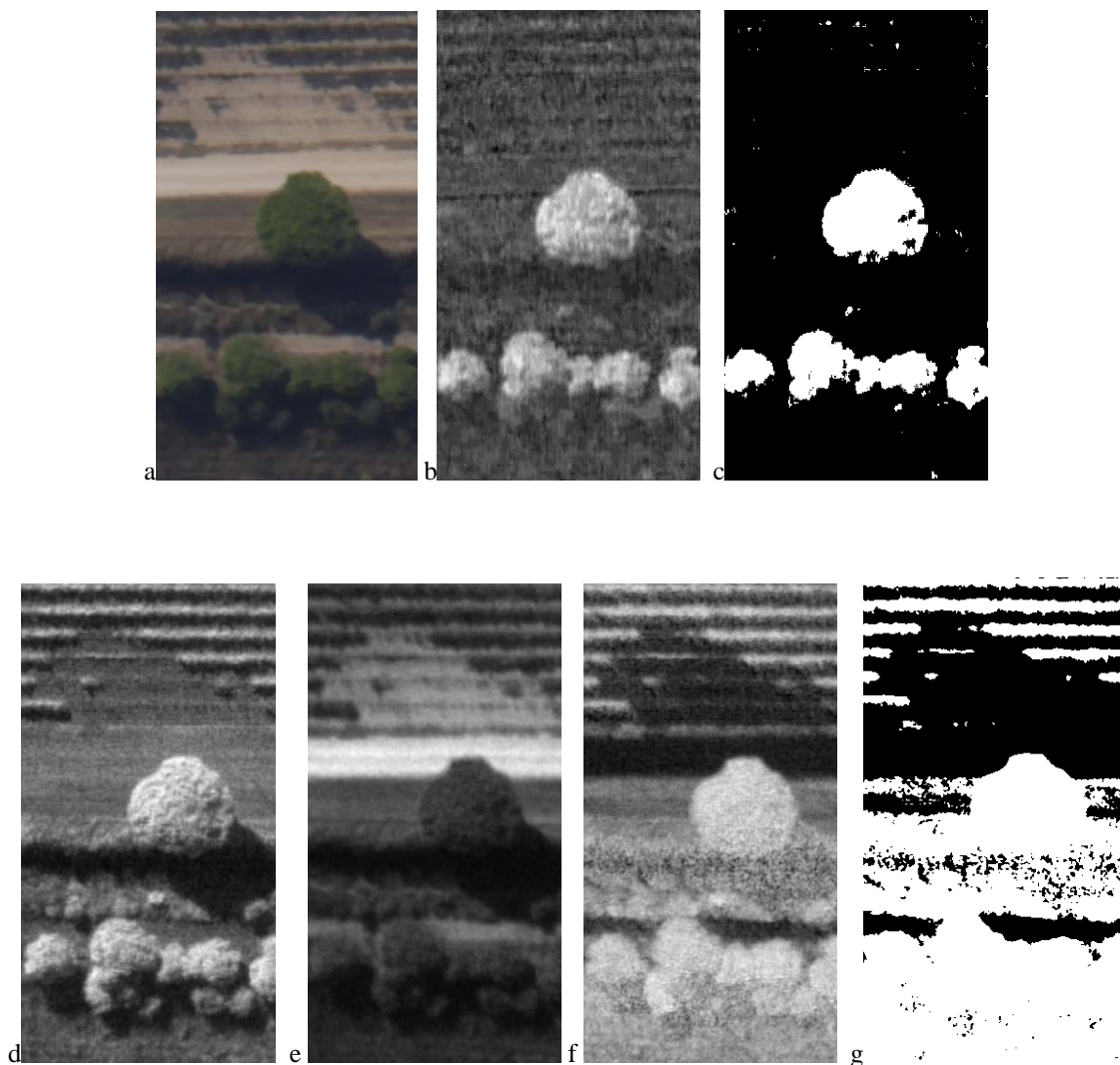
The better accuracy of NDVI thresholding compared to EGI thresholding in Fig.11 is confirmed by the  $\lambda_W$  values (0.81 and 0.66 respectively). Also, it is important to notice that the linear coefficients of Table 2, determined from Canon 500D characteristics, are still suitable for the Canon 350 D.

Figure12 shows a particular case including various types of vegetation. The corresponding images (standard and modified Canon 500D cameras, one flight for each) are issued from the scene of Fig. 8. They were acquired in a vineyard area (Camargue, south of France) in September 2011, with a flight altitude of 1000 m and a 50 mm lens (spatial resolution ~ 100 mm).

According to  $\lambda_W$  values (0.7527 and 0.7531 respectively), EGI and NDVI have similar thresholding quality.

However, with EGI, only the trees are discriminated from the soil. The vine rows (dark green) are only detected using NDVI. In that sense, NDVI appears as more robust for plant detection.

Computed R and NIR bands are also given for illustration.



**Fig. 12:** EGI and NDVI computation on a vine parcel. a: colour image; b,c: EGI computing and thresholding ( $\lambda_W=0.7527$ ); d,e: NIR and R computed bands; f,g: NDVI computing and thresholding ( $\lambda_W=0.7531$ )

#### 4. Discussion and conclusion

We have shown that NIR and R bands for NDVI computation can be obtained from a single standard digital RGB still camera, by removing the near-infrared blocking filter inside the camera and adding an external long-wave pass filter. A generic method, called BSOP, has been proposed to determine the optimal external filter. This method was initially evaluated through simulation using field spectral data, showing the good potential of the Canon 500D camera in this context.

Further tests in field conditions with a Canon 500D and a Canon 350D (using the same linear combination coefficients) confirmed the simulation study, with the following limitations:

- due to the Bayer filter matrix, a slight smoothing step is necessary, reducing the spatial resolution compared to one announced by the camera producer.
- especially when computing the R band, the linear combination amplifies the image noise, because it involves subtractions of the raw signals. Therefore, a particular attention must be paid to get a correct image exposure and take benefits of the depth of the raw signals (14 bits for the Canon 500D).
- NDVI resulting values were assessed on vegetation and soil samples. However, because the R and NIR sensitivities issued from BSOP computation are approximations of the standard values, the results for other materials are not guaranteed.

Due to these limitations, comparisons with alternative approaches have to be considered again at this stage.

Compared with multi-CCD NIR sensors, practical tests have shown that the resulting spatial resolution remains widely higher, at a very lower cost. However, a solution associating a standard and a modified camera, such as in (Dare, 2008), remains attractive for the following reasons:

- it provides a better control of the NIR band shape, by using a near-infrared pass-band filter.
- it has less concerns about SNR
- it allows the standard RGB colour imaging, which can be of a great practical interest, to be kept even if not used for image processing.

As stated above, the main concern associated with using two cameras is about image registration, due to the intrinsic difference of grey-level distributions between NIR and visible images of a same scene. This can be partially overcome when a simple geometric transformation model (typically homography) can be assumed between the two images, i.e. when the imaging elevation is sufficient. Such a problem does not exist in our proposed solution, which remains usable with any camera-scene distance.

A final point to be discussed is the influence of ambient light. NDVI has initially been defined using R and NIR *reflectance* values, independent from the ambient lighting conditions. In remote-sensing, reflectance is obtained from luminance digital counts using sky spectrum and atmospheric correction models. At lower spatial scales, it is usually obtained by means of a reference material such as Spectralon® (Labsphere Inc, North Sutton, NH, USA) placed in the scene. In our case, as with any other alternative approach, it is clear that such reference material should be used in order to balance the R and NIR signal and obtain normalised, comparable NDVI values.

In the section 3, another type of NDVI usage, related to vegetation discrimination, was investigated. In this context, the NDVI was computed directly from luminance signals, and thus the results are formally depending on lighting spectral conditions, as it is the case with most standard RGB discrimination approaches. But a strong underlying hypothesis in the case of NDVI is precisely that the significant gap between vegetation and non-vegetation values (as observed in Fig. 7) allows light spectral variations to be overcome much more efficiently than with RGB. This hypothesis is not specific to our approach.

In conclusion, the approach proposed here can provide an efficient solution for cost-effective, high resolution acquisition of NDVI images. Compared to alternatives, this solution can be particularly interesting in the case of short range imaging for NDVI computation (in association with light measurement reference material) or robust vegetation detection. This opens up new possibilities for crop monitoring, as well as vision for agricultural robotics.

## **Acknowledgements**

The research leading to these results has received funding from the European Union's Seventh Framework Programme [FP7/2007-2013] under grant agreement n°245986.

We acknowledge L'avion jaune (<http://www.lavionjaune.fr>) for having provided the modified cameras and contributed in their spectrometric characterization, as well as for having realized aerial image acquisitions.

## References

Bayer, B. E. (1976). *Color imaging array*. U.S. Patent 3 971 065, July 1976.

CIE (2004). *CIE :15, Technical report – Colorimetry, 3<sup>rd</sup> Edition, 2004*. ISBN 3901906339, 82 pages.

Dare, P.M. (2008) Small format digital sensors for aerial imaging applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B1. Beijing 2008

Deriche, R. (1990). Fast Algorithms for Low-Level Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence archive; Volume 12 Issue 1, January 1990*, 78-87

Hirakawa, K., Parks, T.W. (2005). Adaptive homogeneity-directed demosaicing algorithm. *Image Processing, IEEE Transactions on*, vol.14, no.3, March 2005, 360-369.

Huete, A.R., Liu, H.Q., Batchily, K.L., van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote sensing of environment*, 1997. 59(3), 440-451.

Jindong, W., W. Dong, and M.E. Bauer (2007) Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field crops research*, 2007. 102(1), 33-42.

King-Hong C., Yuk-Hee C. (2006). Color Demosaicing Using Variance of Color Differences Image Processing, *IEEE Transactions on*, vol.15, no.10, Oct. 2006, 2944-2955,



Marks, R.J. (1991). *II: Introduction to Shannon Sampling and Interpolation Theory*, Springer-Verlag, 1991.

Otsu, N. (1979). A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 9, No. 1, 62-66.

Rouse, J. W., Haas, R. H., Schell, J. A. & Deering D. W. (1973). Monitoring vegetation systems in the great plains with ERTS, *Proceedings of the 3rd ERTS Symposium*, Vol. 1, 309-317.

Thevenaz, P.; Blu, T.; Unser, M. (2000). Interpolation revisited [medical images application]. *Medical Imaging, IEEE Transactions on*, vol.19, no.7, July 2000, 739-758,

Tucker, C.J. (1979). Red and Photographic Infrared Linear Combinations for Monitoring Vegetation, *Remote Sensing of Environment*, 8(2), 127-150.

Vigneau, N., Ecarnot, M., Rabatel, G. and Roumet, P. (2011). Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. *Field Crops Research*, 2011. 122(1), 25-31

Woebbecke, D.M., Meyer, G.E., Von Bargen, K. & Mortensen D.A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, Vol. 38, Nr. 1, 259-269.

Yuhas, R.H., Goetz, A.F.H., Boardman, J.W. (1992). Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In *Summaries of 3rd Annual JPL Airborne Geoscience Workshop*, JPL Publication 92-14, vol.1, 147-149

Zhengwei, Y., Hu Zhao, Liping Di, Yu, G. (2009) A Comparison of Vegetation Indices for Corn and Soybean Vegetation Condition Monitoring. in *Geoscience and Remote Sensing Symposium, 2009 IGARSS 2009*, pp IV-801 - IV-804

