

Getting NDVI Spectral Bands from a Single Standard RGB Digital Camera: A Methodological Approach

Gilles Rabatel¹, Nathalie Gorretta¹, and Sylvain Labbé²

¹ Cemagref, UMR ITAP, 361 rue Jean-François Breton, BP 5095, 34196 Montpellier Cedex 5, France

² Cemagref, UMR TETIS, F-34093 Montpellier, France

Abstract. Multispectral images including red and near-infrared bands have proved their efficiency for vegetation-soil discrimination and agricultural monitoring in remote sensing applications. But they remain rarely used in ground and UAV imagery, due to a limited availability of adequate 2D imaging devices. In this paper, a generic methodology is proposed to obtain simultaneously the near-infrared and red bands from a standard RGB camera, after having removed the near-infrared blocking filter inside. This method has been applied with two new generation SLR cameras (Canon 500D and Sigma SD14). NDVI values obtained from these devices have been compared with reference values for a set of soil and vegetation luminance spectra. The quality of the results shows that NDVI bands can now be acquired with high spatial resolution 2D imaging devices, opening new opportunities for crop monitoring applications.

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

1 Introduction

The Normalized Difference Vegetation Index, or NDVI, introduced in the early seventies by [1], remains today a very popular tool in the remote sensing community dealing with agricultural monitoring. This is mainly due to its remarkable ability to discriminate vegetation from other material in multispectral satellite images. Green vegetation is characterized by a high reflectance in the near-infrared domain (typically 50 to 80%), which contrasts with a very low reflectance in the red wavelengths, due to chlorophyll absorption. Let us call R and NIR the digital counts obtained through the red and the near infrared bands of a multispectral sensor. The NDVI, expressed as:

$$\text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}} \quad (1)$$

is a scalar value in the range $[-1, 1]$. The higher is this value, the higher is the probability that it corresponds to green vegetation. By extension, numerous

attempts have been made to directly link NDVI or other derived indexes based on R and NIR to agronomical indices such as biomass and LAI (Leaf Area Index) [2], [3], [4].

The popularity of NDVI in remote sensing has been widely supported by the availability of R and NIR channels on most satellite line-scanning sensors (Landsat, SPOT, etc.). Unfortunately, it is not the case for crop monitoring applications at a lower spatial scale: most vision systems embedded on ground vehicles or UAV (Unmanned Aerial Vehicle), which require 2D sensors, are based on standard colour cameras, leading to robustness issues in vegetation detection.

In most colour cameras, the RGB channels are obtained by setting a mosaic of microfilters (Bayer matrix) directly on the CCD (or CMOS) sensor. This solution requires a very large production scale, and the development of specific near-infrared bayer matrices for the vegetation monitoring market cannot be envisaged. Therefore, for agricultural applications, some camera manufacturers propose multi-CCD devices including a near infrared channel, e.g. the MS-4100 (Geospatial Systems Inc., West Henrietta, NY, USA), or the AD-080 (JAI Ltd, Yokohama, Japan). But their price is often prohibitive. Moreover, their spatial resolution is hardly sufficient for UAV image acquisition.

To face this situation, some camera end users requiring a near infrared channel have developed alternative solutions around standard Bayer matrix RGB cameras, taking benefit of an undesirable property of their 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 add a near-infrared blocking filter to match natural colorimetry requirements. By removing this additional filter, we obtain a modified camera sensitive to near infrared wavelengths. A first possibility to get separate R and NIR bands is thus to use simultaneously a standard and a modified colour camera, the second one being equipped with a near infrared pass-band filter [5]. However, important issues arise concerning the pixel alignment of the two images obtained [6].

Another interesting approach is to use a single modified camera associated with a low-pass filter, and to build the near infrared band as a specific linear combination of the three resulting channels. This concept can be illustrated as following:

Let us assume we have an ideal modified RGB camera, where the three R,G,B channels deliver digital counts respectively equal to $R+NIR$, $G+NIR$, $B+NIR$. If we add a low-pass filter in front of the lens that blocks the blue wavelengths, then we get only the NIR component on the blue channel. This component can be subtracted from the other channel digital counts, leading finally to R, G and NIR components. In the real world, the sensitivity of each channel cannot be modeled so simply, and a specific study is necessary to determine the required low-pass filter and the linear combination for a given camera. Until now, this approach has been used by the company Tetracam (Tetracam Inc. Chatsworth, CA, USA) in its agricultural cameras (ADC series). However, their spatial resolution (~ 3 Mpixels) does not meet the present standard of RGB still cameras (more than 10 Mpixels).

The purpose of the present study is to formalize this last approach and to show how it can be extended to the newest generation of commercial RGB imaging sensors, in order to combine robust spectral information on vegetation with very high spatial resolution. A generic methodology is proposed to determine the optimal low-pass filter and linear combination for virtually any modified RGB camera, provided its sensitivity curves are known. The results obtained with two commercial imaging devices representative of the recent evolutions in sensor technology, the Canon 500D (Canon, Tokyo, Japan) and the Sigma SD14 (Sigma, Kawasaki, Japan) are then presented.

2 The BSOP Theoretical Approach

The basic idea developed here is to simulate a desired spectral sensitivity (further referred as target sensitivity) by a linear combination of the real spectral sensitivities available for a given sensor associated with a low-pass filter. In the following, we will express formally this linear combination, and then propose a method to determine the optimal low-pass filter to be associated with the sensor.

2.1 Band Simulation by Orthogonal Projection (BSOP)

Formally, the spectral sensitivity of a sensor channel (or band) can be characterised by a function $v(\lambda)$ of the wavelength λ . It will generate, for a given irradiance $e(\lambda)$, a digital count:

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

For practical reasons, we will consider in the following a limited range of wavelengths and a limited spectral resolution, allowing us to consider any spectral function as a vector in a spectral space of dimension n , leading to the discrete expression:

$$DC = \sum_{i=1}^n e(\lambda_i).v(\lambda_i) = E.V \quad (3)$$

where E and V are vectors of dimension n and $E.V$ is their scalar product.

According to this formalism, let us call V_1, \dots, V_p the actual sensitivities of the p bands of an imaging sensor. By linear combination of V_1, \dots, V_p , we can simulate any virtual sensitivity V belonging to the subspace generated by the vectorial base (V_1, \dots, V_p) . Now, let us consider a given target sensitivity V_t that we want to simulate. In the general case, V_t will not belong to this subspace, and the better approximation of V_t will be its orthogonal projection on the (V_1, \dots, V_p) subspace (Fig. 1).

Let us call $B = [V_1 \dots V_p]$ the (n, p) matrix whose columns are the V_1, \dots, V_p vectors. The orthogonal projection of V_t can be expressed as:

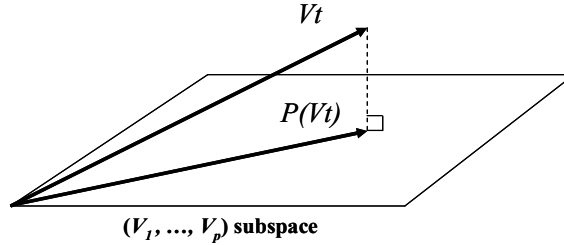


Fig. 1. Illustration of the orthogonal projection of the target sensitivity V_t

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

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 of dimension p :

$$C = B.(B^T B)^{-1}.B^T.V_t \quad (5)$$

According to (4) and (5):

$$P(V_t) = B.C = [V_1 \dots V_p].C = c_1.V_1 + \dots + c_p.V_p \quad (6)$$

2.2 Optimal Low-Pass Filter

Whatever is 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, but this does not guarantee a satisfactory result. We have still to verify 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 [7]. This measure evaluates the angle between two vectors v_1 and v_2 of euclidian norms $\|v_1\|$ and $\|v_2\|$ as:

$$SAM(v_1, v_2) = \text{acos}\left(\frac{v_1.v_2}{\|v_1\|.\|v_2\|}\right) \quad (7)$$

In our case, the value $SAM(V_t, P(V_t))$ should be ideally equal to zero, meaning that V_t belongs to the subspace (V_1, \dots, V_p) . 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. Let us call (V_1^*, \dots, V_p^*) these initial sensitivities. Our only degree of freedom to improve the situation is to associate with the sensor an optical filter (e.g. by setting it in front of the lens) with a spectral transmittance F , leading to a modified set of sensitivities:

$$(V_1, \dots, V_p) = (F.V_1^*, \dots, F.V_p^*) \quad (8)$$

Our objective is thus to determine the optical filter F that will minimize the SAM between the target V_t and its projection $P(V_t)$, according to equations (4) and (8). A first 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, but overall it could lead to non-realisable optical filters, unless constraints of positivity and smoothness are introduced. At the present stage, according to the first considerations described in introduction, we have chosen to reduce the complexity of the problem by considering only simple low-pass filters as F candidates. A main advantage is the further possibility to implement the solution with on-the-shelf gelatine or glass filters.

Let us define a candidate low-pass filter by its cutting wavelength λ_c :

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

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

Then for a given wavelength λ_c we can compute:

- the subspace matrix $B_{\lambda_c} = [V_{1\lambda_c}, \dots, V_{p\lambda_c}] = [V_1 \cdot F_{\lambda_c}, \dots, V_p \cdot F_{\lambda_c}]$
- the projected vectors $P_{\lambda_c}(V_{t1}), \dots, P_{\lambda_c}(V_{tk})$, according to equation (4)
- a global cost function taking into account the similarity between every target and its projection:

$$R(\lambda_c) = SAM(V_{t1}, P_{\lambda_c}(V_{t1})) + \dots + SAM(V_{tk}, P_{\lambda_c}(V_{tk})) \quad (10)$$

The optimal low-pass filter F_{λ_c} will be the one that minimise $R(\lambda_c)$.

2.3 Renormalisation of the Projected Vectors

The SAM criterion above has been used to determine a set of vectors $P_{\lambda_c}(V_{t1}), \dots, P_{\lambda_c}(V_{tk})$ matching as well as possible an initial set V_{t1}, \dots, V_{tk} of target sensitivities, in terms of spectral shape. Another important point, if these vectors are devoted to the computation of agricultural indices like NDVI, is that they provide digital counts as close as possible to the original ones. Though this can obviously not be obtained for every irradiance spectrum, an approximate solution is to ensure that the projected vector and the target vector have the same $L1$ -norm¹.

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

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

¹ 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 an irradiance E_f .

3 Material and Methods

3.1 Camera Sensitivity Measurement

Two commercial still cameras have been selected for the BSOP method assessment: the Canon 500D, representative of recent high resolution Bayer matrix sensors (15 Mpixels), and the Sigma SD14, 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 color. No Bayer matrix is required, leading to a better spatial resolution. In counterpart, there is less control on the spectral sensitivity curve of each channel.

Each camera has been opened and its NIR blocking filter removed (the filter is removable in the Sigma, but not in the Canon, leading to a more delicate operation). The spectral sensitivity of each modified camera has then been measured in the range 440-990 nm, with 10 nm steps. For this purpose, the tunable monochromatic light source of a laboratory spectrometer (V-570, Jasco Inc, Easton, USA) have been remoted in front of the camera lens via an optical fiber, at a 30 cm distance, in a room with no ambient light. Pictures have been taken in raw format for each wavelength in the following conditions:

Focal length: 50 mm; Sensibility: 100 ISO. Integration time: 5 s

The average level of the light source image has then been collected for each channel in the raw camera pictures, using home-made image analysis software.

3.2 Definition of the Target Sensitivities

Two target sensitivities V_{t1} , V_{t2} corresponding respectively to the red and near-infrared bands have been considered. Because no standardized sensitivity data have been found in the literature (some authors use the bands TM3 and TM4 of Landsat, other simply use 660 nm and 760 nm wavelengths), the following procedure has been used:

- the red band has been derived from the XYZ colorimetric data of CIE 1931 (source: <http://www.cvrl.org/database/text/cmfs/ciexyz31.htm>)
- the near-infrared band has been computed by a 160 nm shift of the red band, leading to a bandwidth 760-830 nm at mid-height.

3.3 BSOP Computation

In order to be used further on real hyperspectral data (see section 3.4), all the data defined above (sections 3.1 and 3.2) have been resampled according to the spectral resolution of an hyperspectral camera model Hypspec 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, this has led to two sets of vectors V_1^* , V_2^* , V_3^* (one for each type of camera) and two vectors V_{t1} , V_{t2} of dimension 160. All computations described in section 2 have been made with Matlab

7 (The MathWorks, Natick, MA, USA). Once the optimal cutting wavelength has been determined, the closest existing gelatine Wratten filter has been chosen (see 4.1). Its real transmittance curve has been measured with the Jasco V-570 spectrometer, and resampled according to the Hypspx 160 bands, leading to a filter vector F_w . Finally, for each type of still camera, the projected vectors $P_N(V_{t1})$ and $P_N(V_{t2})$ on the subspace $(F_w.V_1^*, F_w.V_2^*, F_w.V_3^*)$ have been computed and renormalised according to (11).

3.4 NDVI Simulation on Field Hyperspectral Images

Real hyperspectral images of wheat durum have been acquired in experimental fields (INRA, Domaine de Melgueil, France) in march 2011, using the Hypspx VNIR-1600 camera. The camera was set on a motorized translation rail one meter above the ground (see [8] for more details). The scenes include wheat durum at early stage, as well as various types of dicotyledonous and monocotyledon weeds. A total of 2210 luminance spectra have then been collected in the hyperspectral images by manual area selections, including wheat, weeds and soil categories.

For every collected spectrum S , the digital counts R and NIR corresponding to red and near infrared target sensitivities V_{t1} , V_{t2} have been computed as $R = S.V_{t1}$ and $NIR = S.V_{t2}$, according to equation (3). A reference NDVI value has then been computed using this R and NIR values. The same operation has been made using $P_N(V_{t1})$ and $P_N(V_{t2})$ instead of V_{t1} and V_{t2} to compute an estimated NDVI for each type of still camera.

4 Results

4.1 Band Simulation

Fig.2 shows the spectral sensitivities V_1^* , V_2^* , V_3^* that have been measured respectively for the Canon and Sigma cameras without their NIR blocking filter.

In Fig.3, the BSOP quality curves $R(\lambda_c)$ according to equation (10) have been reported. We can notice that in both cases, $R(\lambda_c)$ is undefined for $\lambda_c > 728$ nm, because the projection of V_{t1} (red band) becomes null. The best BSOP quality is obtained for $\lambda_c \approx 600$ nm, and is better for the Canon camera. According to this result, a standard Wratten filter Kodak n^o 25 (reference: 149 7072) with a cutting wavelength of 600 nm has been selected, and its actual transmittance F_w measured with the Jasco spectrometer for further simulations.

The results of BSOP are given in Fig.4. The simulated near infrared band obtained with the Canon camera appears more regular than the one obtained with the Sigma camera, confirming the better results obtained with $R(\lambda_c)$.

4.2 NDVI Computation

Fig. 5 shows the NDVI values computed with the simulated bands of Fig.4, versus the reference NDVI values computed with the target sensitivities V_{t1}, V_{t2} , for the 2210 luminance spectra that have been collected (see section 3.4).

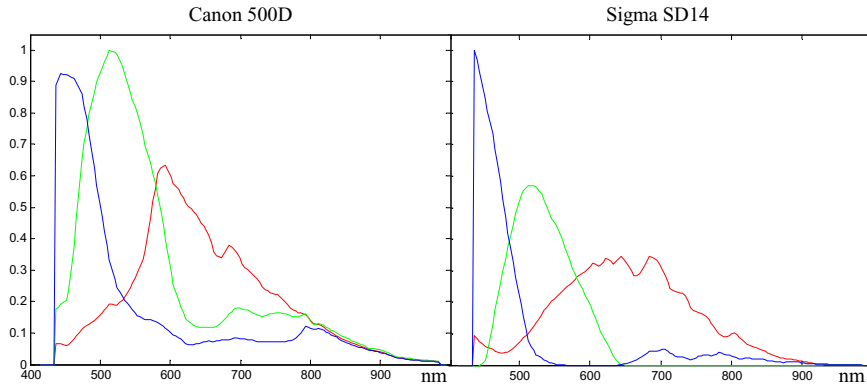


Fig. 2. Spectral sensitivities of the Canon 500D and Sigma SD14 without NIR blocking filter (sensitivities have been rescaled for a maximum value equal to unity)

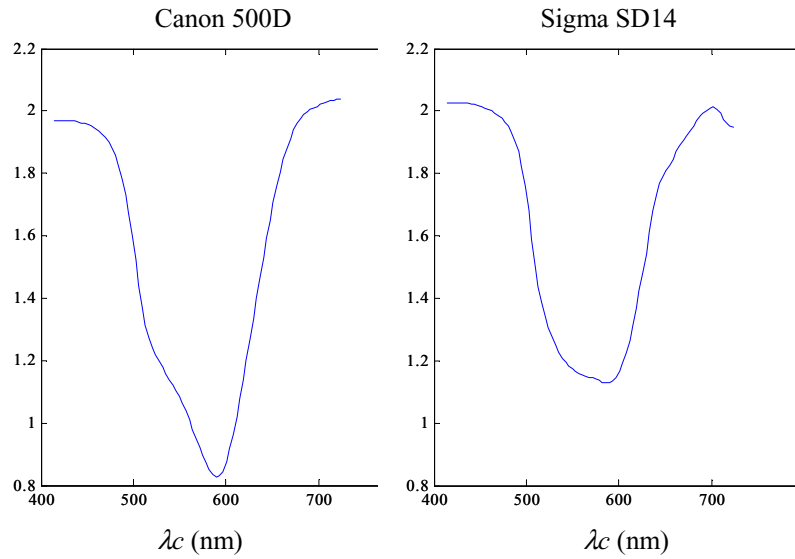


Fig. 3. BSOP quality $R(\lambda_c)$ of Canon and Sigma cameras

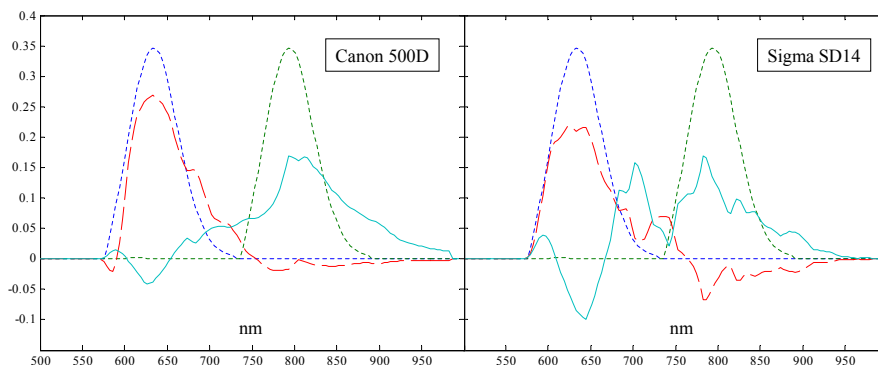


Fig. 4. BSOP bands for Canon and Sigma cameras with Wratten filter n° 25. Dotted lines: target sensitivities V_{t1}, V_{t2} . Dash line: BSOP red band; Solid line: BSOP near-infrared band.

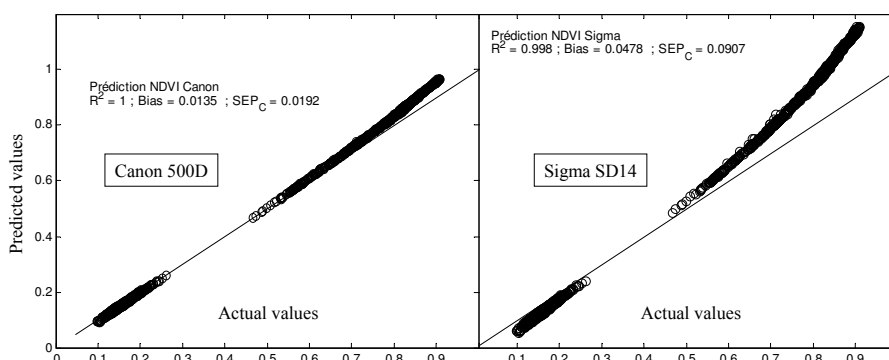


Fig. 5. Predicted versus reference values of NDVI

First, we can notice 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. The prediction quality using the Canon camera is significantly better: the NDVI values obtained are nearly equal to the reference ones, excepted for the highest values (NDVI > 0.8) where the error remains less than 10%.

5 Conclusion

We have shown in this paper that both NIR and R bands for NDVI computation can be obtained simultaneously from a single standard digital RGB still camera, by replacing the near-infrared blocking filter inside by a low-pass filter. A generic method, called BSOP, has been proposed to determine the optimal replacing filter. This method has been applied using real spectrometric data on

two commercial SLR cameras based respectively on a Bayer matrix sensor and a Foveon sensor, associated with a standard Wratten filter. The simulated NDVI values obtained for a large number of luminance spectra of vegetation and soil have been compared with reference NDVI values. The results obtained with the Bayer matrix sensor are particularly satisfactory. The Foveon provides a much less accurate NDVI prediction, but remains a pertinent choice in the context of vegetation/soil discrimination by NDVI thresholding. The results open new possibilities in terms of high-resolution imaging for crop monitoring. Our next step will be the real implementation and test of both types of devices on an UAV, in the frame of a weed monitoring application.

Acknowledgements. The research leading to these results has received funding from the European Unions Seventh Framework Programme [FP7/2007-2013] under grant agreement *n°* 245986. We acknowledge *Lavion jaune* (<http://www.lavionjaune.fr>) for having provided the modified cameras and contributed in their spectrometric characterization.

References

1. Rouse, J.W., et al.: Monitoring vegetation systems in the great plains with ERTS. In: Third ERTS Symposium (1973)
2. Huete, A.R., et al.: A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote sensing of environment* 59(3), 440–451 (1997)
3. Jindong, W., Dong, W., Bauer, M.E.: Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field Crops Research* 102(1), 33–42 (2007)
4. Zhengwei, Y., et al.: A Comparison of Vegetation Indices for Corn and Soybean Vegetation Condition Monitoring. In: *Geoscience and Remote Sensing Symposium, IGARSS 2009, Cape Town* (2009)
5. Lebourgeois, V., et al.: Can Commercial Digital Cameras Be Used as Multispectral Sensors? A Crop Monitoring Test. *Sensors* 8(11), 7300–7322 (2008)
6. Dare, P.M.: Small format digital sensors for aerial imaging applications. In: *XXIst ISPRS Congress, Beijing, China* (2008)
7. Yuhas, R.H., Goetz, A.F.H., Boardman, J.W.: Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In: *Summaries of 3rd Annual JPL Airborne Geoscience Workshop, vol. 1, pp. 147–149. JPL Publication 92-14* (1992)
8. Vigneau, N., et al.: Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. *Field Crops Research* 122(1), 25–31 (2011)