CrossMark

On the value of the Kullback–Leibler divergence for cost-effective spectral imaging of plants by optimal selection of wavebands

Landry Benoit[1] · Romain Benoit[1] · Étienne Belin[1] · Rodolphe Vadaine[2] ·
Didier Demilly[2] · François Chapeau-Blondeau[1] · David Rousseau[3]

Abstract The practical value of a criterion based on statistical information theory is demonstrated for the selection of optimal wavelength and bandwidth of low-cost lighting systems in plant imaging applications. Kullback–Leibler divergence is applied to the problem of spectral band reduction from hyperspectral imaging. The results are illustrated on various plant imaging problems and show similar results to the one obtained with state-of-the-art criteria. A specific interest of the proposed approach is to offer the possibility to integrate technological constraints in the optimization of the spectral bands selected.

Keywords Spectral imaging · Information theory · Plant imaging

1 Introduction

Hyperspectral imaging is an important tool in agriculture and plant sciences to measure, identify, or classify plant tissues [1]. This imaging technique which can produce images of reflectance spectra is commonly used over large geometric scales in remote sensing. It is also gaining more and more interest for observation at smaller geometric scales, in prox-

imal detection [2], especially with its ability to enrich the scalar information produced by single-probe spectrometers with the spatial distribution inherent to images. This trend opens new imaging problems not present in remote sensing nor with single probe spectrometry [3]. Embedding hyperspectral imaging in field requires the development of specific imaging setups [4], also there is a need to couple hyperspectral imaging with 3D imaging systems to take into account the 3D structures of plants [5] in the quantitative interpretation of the acquired spectral informations. In addition, hyperspectral images deliver huge amounts of data, typically some gigabytes per image. This induces challenges on the handling of such big data, especially in the perspective of continuous monitoring or imaging of populations of plants [6]. In a translational research perspective with application to agronomy or high-throughput phenotyping with multiple imaging systems, it is important to promote the use of low-cost setups [7]. Hyperspectral imaging, although not low-cost, can serve to the design of low-cost imaging setups based on a small number of spectral bands which are judged as the most useful part of the spectrum for a given application. This corresponds to a spectral band selection problem. There are multiple criteria in the literature of machine vision for spectral selection based on analysis of high-resolution spectra [8–16] recently reviewed in [17]. The most common in plant sciences (see for a typical example [18]) are transformations of the spectral information contained in every pixel into one or more components that relate directly to sample properties of interest. This includes vegetation indices (see [19,20] for reviews) or linear combination of all the components based on statistical criteria such as done in partial least square analysis PLS (see [21,22] for the statistical principle) or canonical discriminant analysis (CDA) for classification purposes. Both approaches have some intrinsic limitations. Vegetation indices, among which the normalized difference vegetation index (NDVI),

✉ David Rousseau
  david.rousseau@univ-lyon1.fr

[1] Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France

[2] GEVES, Station Nationale d'Essais de Semences (SNES), 25 rue Georges Morel, 49071 Beaucouzé, France

[3] CREATIS; CNRS UMR 5220; INSERM U1044; Université Lyon 1, INSA-Lyon, Université de Lyon, 69621 Villeurbanne, France

are mainly computed with differences and ratios. Such vegetation indices are known to show some saturation noise with pixels randomly giving extreme values although none of the native acquired images are actually saturated. The PLS and CDA methods produce a linear filter combination of all the hyperspectral spectral bands into a single mono-component image. The method requires that all the spectral components are first individually measured by a sophisticated hyperspectral camera, and then optimally combined in a digital computer to form a single monocomponent image. As a result, the approach is difficult to implement physically when coefficients of the combination are negative or vary very sharply along the spectrum. As a useful alternative, one could wish to directly measure one or a few efficient spectral combinations, with specifically designed and physically fast implementable sensors. Such an approach would substitute the use of a general-purpose sophisticated yet slow hyperspectral camera, by low-cost dedicated sensors optimized for specific imaging tasks. In this article, we propose such an approach to perform dimensionality reduction of broadband spectra in the presence of physical constraints. A comparable concern of selecting physically implementable optical filters is addressed in [17]. Filter selection in [17] is addressed as a feature selection problem and tackled by machine learning approaches, with a binary feature produced by thresholding each filter output to implement a classification based on various similarity measures. By contrast here, we test another informational criterion based on the Kullback–Leibler divergence, which does not use threshold binarization for the spectral representation and offers somehow a generic low-level contrast measure free of higher-level assumptions required in machine learning. Our approach is illustrated in several imaging situations, involving different samples of biological interest raising difficult informational tasks to the human eye, and showing the benefit of the optimal spectral band selection from our information–theoretic criterion.

2 Information criterion

In this section, we briefly present the statistical information model of light spectrum recently introduced in [23] that we extend to another informational metric adapted to the imaging situations addressed in the following sections.

Let the light be collected by a sensor incorporating a set of photodetectors of M distinct types labeled by index i. Photodetector i integrates the energy contained in the incident light spectrum $S(\lambda)$ weighted by the spectral sensitivity $f_i(\lambda)$ of the photodetector to output a scalar measurement. We consider that each photon with wavelength λ falling on the sensor has a probability $f_i(\lambda)$ of being detected by a photodetector i. The probability $f_i(\lambda)$ is wavelength-dependent

to account for a wavelength-dependent spectral sensitivity for each photodetector i. When a very large number N of photons are radiated by the source and fall on the sensor, a fraction $NS(\lambda)d\lambda$ is radiated at wavelength λ, among which a fraction $f_i(\lambda)NS(\lambda)d\lambda$ is detected by a photodetector i. By integration over the whole wavelength range $[\lambda_{min}, \lambda_{max}]$ where the input spectrum $S(\lambda)$ contains energy, the total number N_i of photons collected by photodetector i comes out as:

$$N_i = \int_{\lambda_{min}}^{\lambda_{max}} f_i(\lambda)NS(\lambda)d\lambda , \quad i = 1, \ldots M, \tag{1}$$

which matches the macroscopic picture of a weighted integration of the incident light to construct the global response of the photodetector.

A given sensor with spectral resolution capabilities incorporates M distinct photodetector types with M distinct probabilistic spectral sensitivities $f_i(\lambda)$. Consistency of the probabilistic description imposes

$$P_{det}(\lambda) = \sum_{i=1}^{M} f_i(\lambda) \leq 1 \tag{2}$$

for each λ, where $P_{det}(\lambda)$ in Eq. (2) is the global probability for an incident photon at λ to be detected by the sensor, altogether by absorption by one of its M internal photodetectors. In addition,

$$P_{lost}(\lambda) = 1 - P_{det}(\lambda) = 1 - \sum_{i=1}^{M} f_i(\lambda) \tag{3}$$

is the probability that an incident photon at λ is missed by the sensor.

For each photon, we denote by Y the random variable describing the photodetection event occurring in the sensor. A photon emitted by the source is emitted at wavelength λ with the probability $S(\lambda)d\lambda$, and a photon at λ is absorbed by photodetector i with the (conditional) probability $P(Y = i|\lambda) = f_i(\lambda)$. By integration over the whole wavelength range, we obtain the overall probability that a photon emitted by the source is detected by photodetector i as:

$$P(Y = i) = \int_{\lambda_{min}}^{\lambda_{max}} f_i(\lambda)S(\lambda)d\lambda, \quad i = 1, \ldots M, \tag{4}$$

consistent with the global count N_i of Eq. (1). There is also a possibility that the photon at λ is missed that we denote $Y = 0$, occurring according to Eq. (3) with (conditional) probability $P(Y = 0|\lambda) = P_{lost}(\lambda)$, leading to the overall probability of a lost photon as:

$$P(Y = 0) = \int_{\lambda_{min}}^{\lambda_{max}} P_{lost}(\lambda)S(\lambda)d\lambda. \tag{5}$$

In this way, the detection of a photon falling on the sensor is modeled as a random event Y, with $M + 1$ possible outcomes, consisting in a detection by photodetector i with the probability $P(Y = i)$ of Eq. (4) for $i = 1$ to M, or in a lost photon with the probability $P(Y = 0)$ of Eq. (5). In a similar way, the emission of a photon by the light source can be modeled as a random event X which describes the wavelength λ at which this photon is emitted, with the outcome $X \in [\lambda, \lambda + d\lambda[$, or more concisely $X = \lambda$, occurring with probability

$$\Pr\{X \in [\lambda, \lambda + d\lambda[\} = S(\lambda)d\lambda. \tag{6}$$

In [23], the statistical framework of Eqs. (1)–(6) was used to compute the Shannon mutual information $I(X; Y)$, a measure of similarity between the input and output data. In this work, we propose to tackle classification tasks from the output data obtained after spectral reduction into M spectral bands. We will, therefore, consider situations where we want to maximize the dissimilarity of spectra of different classes after the spectral selection. To this end, let Y_A and Y_B be the random variable describing the photodetection event occurring in the sensor for two different classes to be separated. We consider the Kullback–Leibler divergence

$$KL(P(Y_A)\|P(Y_B)) = \sum_{i=0}^{M} P(Y_A(i)) \log_2$$
$$(P(Y_A(i))/P(Y_B(i))), \tag{7}$$

as measure of dissimilarity between Y_A and Y_B. The relevance of Kullback–Leibler divergence, which minimizes the maximum likelihood estimation between an empirical Y_A and a reference distribution Y_B, has been identified for a long time [24] in problems in statistics. In the following, we investigate the relevance of this divergence in the statistical framework of [23] specifically designed for hyperspectral imaging.

3 Biological samples and information task

We present in this section three informational problems of biological interest which are hard to solve with non-optimized approach, i.e., with broad band gray level imaging or RGB imaging and that we will tackle with low-cost imaging systems designed after spectral selection based on the information criterion of the previous section.

The first informational problem considered is the detection of the presence/absence of apple scab at the surface of apple tree leaves. This disease seen as the most serious disease for apple [25] is caused by the fungus *Venturia inaequalis* and requires more than 10 fungicides treatments per year to

Fig. 1 Biological samples considered and informational task associated. *Left* is a leaf of apple tree inoculated on four known areas. The information task consists in the automatic detection of the pixels representing scab-infected leaf tissue. *Middle* represents 64 common wheat grains (*top*) and 64 durum wheat grains (*bottom*). The information task consists in automatically classifying the grains in durum and common. *Right* are three seedlings of fodder beet at a stage of development where there exists a distinction between hypocotyl and radicle. The information task consists in segmenting the two organs

be controlled. The importance of early detection of apple scab has thus triggered the interest for the development of automatic detection from machine vision. Figure 1 shows an apple tree leaf which has been inoculated with scab on four known areas. As illustrated in Fig. 1 (left) and Fig. 2, the scab infection is difficult to perceive from the contrast in an RGB image. Therefore, researchers have turned to other imaging technologies. Thermography [26–28], chlorophyll fluorescence and hyperspectral imaging [29] have been shown to be useful for early detection and quantification of apple scab.

As second informational problem, we considered the binary classification between durum and common wheat grains. The durum wheat grains contain the largest amount of gluten. Their discrimination, possibly fast and automated, is important especially for control and certification purposes. Here again, as shown in Fig. 1 (middle) and Fig. 2, the difference between durum and common wheat is difficult to perceive from the contrast in an RGB image. This classification task has, therefore, also been addressed in the literature with imaging systems including hyperspectral imaging and thermal imaging [30–32].

To further illustrate the interest of our approach, we considered, as third informational problem, the segmentation of organs in seedling. It is important for plant phenotyping to identify hypocotyl radicle and cotyledon early after their formation in the seedlings, and to follow their development during seedling elongation. Such observations carry useful relevance for the better understanding of variations in plant emergence and for prognosis concerning the adult plant [33]. This segmentation task, as shown in Fig. 1 (right) and Fig. 2, is uneasy from RGB images specially the separation

between radicle and hypocotyl. From Fig. 3, we can see that the reflectance intensity of the radicle is systematically lower throughout the spectrum, with a typical difference of 0.2, than the reflectance intensity of the hypocotyl. This is due to the long cells (up to 100 μm) present only in the radicle which increase the absorbing surface for light. The separation of radicle and hypocotyl has, therefore, been investigated from thermal contrasts established in [34] and specific inactinic green light [35] which requires continuous monitoring.

In these three tasks, apple scab detection, durum versus common wheat classification and radicle-hypocotyl segmentation, the imaging modalities used in the literature are rather costly. Given the practical agronomical interest of these informational tasks, it is important to search for lower-cost imaging systems and these, therefore, constitute good candidates to test the information–theoretic approach presented in this study.

For each of the three classification problems, a ground truth is accessible based on prior knowledge. This ground truth is used to construct the ground-truth images shown in Fig. 2, and will serve to assess the automated classification obtained by our information-based image processing protocol.

We demonstrate in the following how the informational criterion of the previous section can efficiently contribute to solve the three informational problems of Fig. 1. To this end, we started with the measurement of high-resolution spectra of the samples presented in Fig. 3. We used an NEO HySpex hyperspectral camera (http://www.hyspex.no/) capable of measuring spectra with 160 equal-width bands over the wavelength range $\lambda \in [\lambda_{min} = 400 \text{ nm}, \lambda_{max} = 1000 \text{ nm}]$ corresponding to a spectral resolution $\Delta\lambda = 3.75$ nm, with a 12-bit quantization. With these hyperspectral images, we computed, as shown in line 3 of Fig. 3, the average high-

resolution spectra of each class of pixels to be separated in the three problems of Fig. 1.

The average of this high-resolution spectra is then used to compute a representation through linear integration weighted by a set of spectral sensitivity functions $f_i(\lambda)$ chosen for the M types of photodetectors. We used a common model of spectral sensitivity [36–38] according to the Gaussian

$$f_i(\lambda) = A_i \exp\left[-\left(\frac{\lambda - \lambda_i}{w_i}\right)^2\right], \tag{8}$$

where λ_i is the central wavelength and w_i the characteristic bandwidth. This Gaussian sensitivity corresponds for
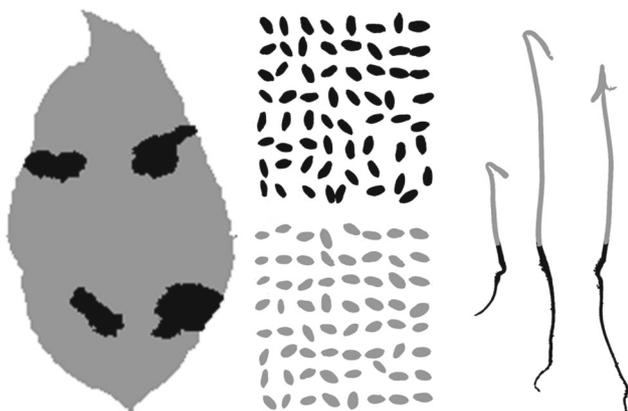

Fig. 2 Ground truth for the information task considered on the biological samples of Fig. 1. *Left* is for the detection of apple scab. *Middle* is for the classification of common versus durum wheat. *Right* is for the segmentation between radicle and hypocotyl organs in seedlings of fodder beet
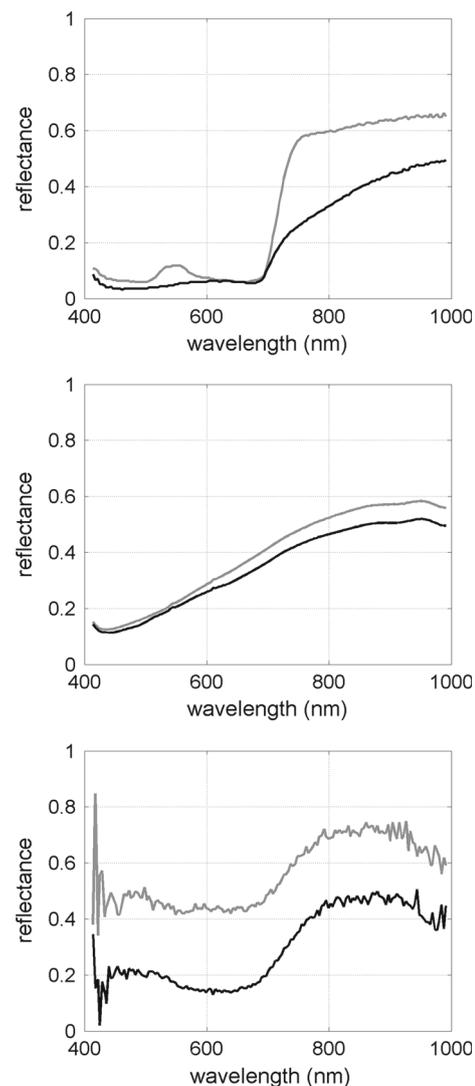

Fig. 3 Reflectance average spectrum, normalized with a white standard, for each class of the information task considered on the biological samples of Fig. 1. *Top* is for the leaf of apple tree with scab. The *gray levels* of the *curve* in the reflectance spectrum correspond to the *gray* levels of the ground truth masks of Fig. 2. *Middle* is for the durum and common wheat grains. *Bottom* is for the seedlings of fodder beet
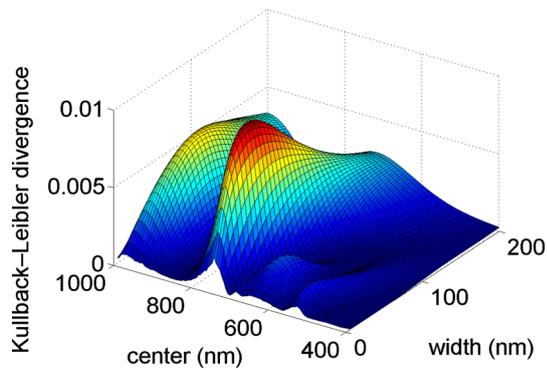
Fig. 4 Optimization landscape of the Kullback–Leibler divergence of Eq. (7) for the apple scab detection problem of Fig. 1 as a function of the parameters (λ_1) center of the spectral Gaussian sensitivity of Eq. 8 and w_1 its bandwidth

instance to a good model for the spectral sensitivity of a broad band photodetector illuminated under large flux of photons (i.e., not in the Poisson regime) by an LED or illuminated by a flat broad band light source filtered by an optical bandpass filter. Spectral selection can then be performed by fixing the number M of allowed photodetectors types and selecting the Gaussian sensitivities (λ_i, w_i) which maximize the Kullback–Leibler divergence of Eq. (7). For the three informational tasks of Fig. 1, we considered without loss of generality, the reference Y_B in Eq. (7), respectively, as the spectrum of the leaf, the spectrum of the durum wheat and the spectrum of the radicle. The optimization landscape obtained for a single photodetector ($M = 1$) in the case of the apple scab detection problem is given for illustration in Fig. 4. By selecting the optimal (λ_i, w_i), we ensure to maximize the dissimilarity, in the Kullback–Leibler sense, between each class of pixel at the output of the M photodetectors. A key point is that since the spectral sensitivities adhere to the Gaussian model of Eq. (8), their optimized versions selected by the informational criterion will usually represent solutions physically realisable (or closely approximated) with current optoelectronic technologies. In the same way, all the results presented in the next section correspond to solutions ready for real-world transfer.

4 Results

We are now in position to present the results of optimal spectral selection based on our information criterion applied to the spectra measured from the three biological samples of Fig. 1.

Table 1 gives, for the apple scab leaf various configurations of spectral reduction including standard references like the red green blue (RGB) sensitivity of the human eye trichromatic response as defined by the International Commission

on Illumination (CIE) [39], the standard RGB to gray conversion of the human eye response according to the linear relation of CIE 1931

$$\text{Gray} = 0.2126R + 0.7152G + 0.0722B, \tag{9}$$

from the human trichromatic RGB responses of CIE, the normalized difference vegetation index

$$\text{NDVI} = \frac{\text{NIR} - R}{\text{NIR} + R}, \tag{10}$$

with NIR a near infrared spectral band and R a red spectral band, both taken here with the Gaussian sensitivity of Eq. (8). We also processed the partial least square approach (PLS), a well-known linear regression adapted to classification tasks. PLS acts as a linear filter on the input data, with coefficients which reflect the covariance between the input data and the classification results. Similarly, we processed the normalized canonical discriminant analysis (CDA) which is also a linear regression adapted to classifications tasks. CDA acts as a linear filter on the input data, with coefficients obtained by maximizing the ratio of the dispersion among classes out of the dispersion within classes. These standard references are compared to the spectral reduction with one, two, three or four spectral bandwidths optimized according to the Kullback–Leibler divergence $K-L$ of Eq. (7). For an identical number of spectral bands, Table 1 expresses the information gain in shannons (Sh) obtained with our approach by comparison with the standard references. Table 1 also quantifies the information gain brought by the increase of the number of spectral bands. The quantitative records in shannons, obtained for the configuration with a single channel, are also found to be in good accordance with the subjective visual inspection of Fig. 5 by comparison to the ground truth of Fig. 2. The visual contrast obtained from the spectral band optimized with the informational criterion clearly outperforms the contrast from CIE 1931 and NDVI while it gives similar results to the PLS and CDA filters. As visible in Fig. 5, the NDVI criterion enhances a small dark region in the leaf which corresponds to only a part of the tissue with pathogen. This is the part with a necrosis which is known to give an important drop in the near infrared reflectance spectrum while we identified that the other areas with the presence of pathogen were not necrotic. We have realized similar analysis on the other samples of Fig. 1 as shown in Table 2 for the durum versus common wheat classification task and in Table 3 for the seedling organs segmentation task. Again the quantitative results expressed in shannons are in accordance with the visual inspection given in Figs. 6, 7 and the ground truth of Fig. 2. Interestingly, one can notice that the single band solution for the segmentation of radicle-hypocotyl in seedling has a larger bandwidth than

Table 1 Comparison of Kullback–Leibler divergence K–L expressed in shannons (Sh) for various configurations of spectral reduction computed from the spectra of the image of apple scab of Fig. 1

	λ (nm)	w (nm)	K–L 10^{-4}
Human eye	X	X	6
RGB to gray	X	X	6
PLS	X	X	0.4
CDA	X	X	0.1
1 with K–L	760	55	92
2 with K–L	750; 950	50; 50	105
NDVI	750; 975	25; 25	89
3 with K–L	750; 850; 950	50; 50; 25	106
4 with K–L	650; 750; 800 ; 950	25; 25; 25; 50	129

Optimization search is made with a step of 5 nm

for the two other classification problems. This is in accordance with the fact that, in the case of the radicle-hypocotyl segmentation, the spectral difference between the reflectance spectrum of the two classes to separate is very broad and almost constant on all the spectrum considered for the study.

This is in accordance with the fact that for the spectral difference between the reflectance spectrum of the two classes to separate in the case of the radicle-hypocotyl segmentation is very broad and almost constant on all the spectrum considered for the study.

To further demonstrate the interest of our information–theoretic approach, we jointly analyze the coefficient of the PLS and CDA filters and the coefficients of the optimal spectral sensitivity $f_1(\lambda)$ for a single bandwidth with the Kullback–Leibler (K–L) divergence. These are plotted in Fig. 8 and it appears that the position of the bandwidth giving the best K–L is positioned in areas where the PLS and CDA filters have rather large positive coefficients. This is consistent with the fact that the visual inspection of the gray level images of Figs. 5, 6 and 7 produced by the K–L criterion, PLS and CDA criteria is visually similar. One can note that for the apple scab problem the spectrum of the two classes spectrum are very close at 700 µm but with contrasted slopes around this wavelength. The coefficient of the K–L method is rather small at this wavelength, since there is no reflectance contrast. Comparatively, the filter coefficient is rather large with positive and negative values for the PLS and CDA methods. The fact that PLS and CDA can use positive and negative coefficients allows to take benefit of the local contrast in the slope of the spectrum between the two classes to be separated. Such contrast cannot be enhanced with the filter obtained from the K–L. However, this points to a specific interest of our approach, since the smooth Gaussian bandwidth given by the K–L criterion is physically implementable while the PLS and CDA filters,
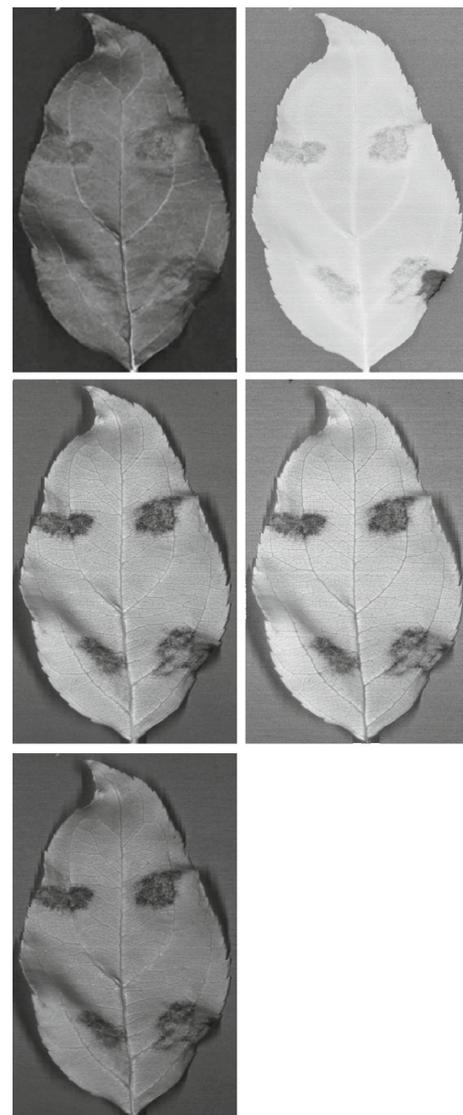
Fig. 5 Visualization for the apple scab problem. *First line left* RGB to *gray* from CIE 1931, *right* NDVI. *Second line* single, *left* bandwidth optimized with Kullback–Leibler divergence, and *right* output of the partial least square filter. *Third line*, output of the canonical discriminant analysis

Table 2 Comparison of Kullback–Leibler divergence K–L expressed in shannons (Sh) for various configurations of spectral reduction computed from the spectra of the image of durum and common wheat of Fig. 1

	λ (nm)	w (nm)	K–L 10^{-5}
Human eye	X	X	6
RGB to gray	X	X	5
PLS	X	X	0.3
CDA	X	X	0.1
1 band with K–L	540	50	6
2 bands with K–L	520; 570	25; 25	7
3 bands with K–L	520; 570; 670	25; 25; 25	7

Optimization search is made with a step of 5 nm

Table 3 Comparison of Kullback–Leibler divergence K–L expressed in shannons (Sh) for various configurations of spectral reduction computed from the spectra of the image of seedlings of Fig. 1

	λ (nm)	w (nm)	K–L 10^{-2}
Human eye	X	X	1.87
RGB to gray	X	X	1.11
PLS	X	X	$9.7 \, 10^{-4}$
CDA	X	X	$2.4 \, 10^{-4}$
1 band with K–L	470	75	2.35
2 bands with K–L	470; 620	75; 75	2.89
3 bands with K–L	450; 570; 670	75; 75; 50	3.15

Optimization search is made with a step of 5 nm
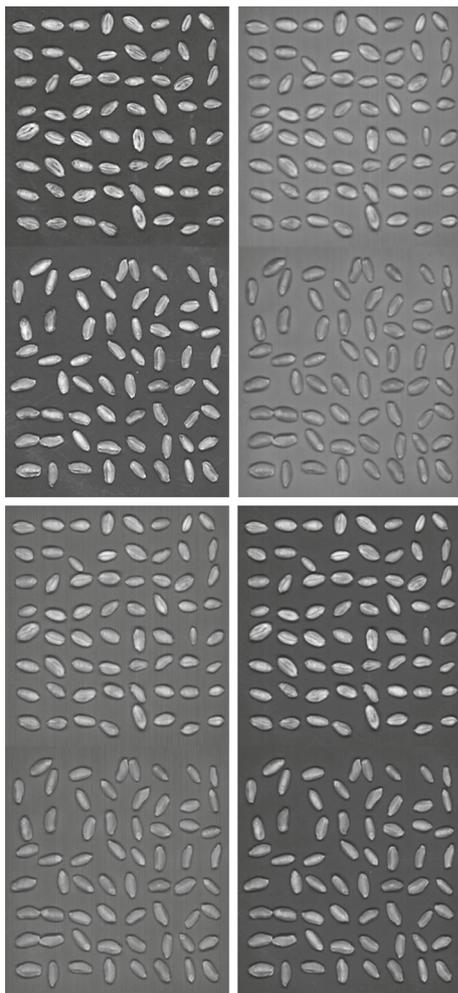
Fig. 6 Visualization for the durum and common wheat grain problem. *First line*, *left* RGB to *gray* from CIE 1931, *right* single bandwidth optimized with Kullback–Leibler divergence. *Second line*, *left* output of the partial least square, *right* output of the canonical discriminant analysis

with negative coefficients and sharp variations, are very hard to implement physically. In this perspective, the optimal couple (wavelength, bandwidth) maximizing K–L is not the only
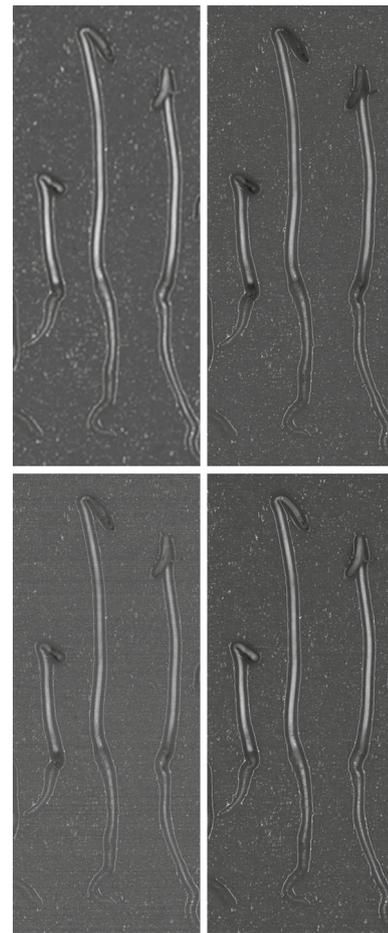

Fig. 7 Visualization for the segmentation of organs of seedlings problem. *First line left* RGB to *gray* from CIE 1931, *right* single bandwidth optimized with Kullback–Leibler divergence, *second line left* output of the partial least square and *right* output of the canonical discriminant analysis

value of practical interest. It is also possible to search for the optimal wavelength for a fixed value of bandwidth corresponding to a technological choice. This is illustrated in Table 4 for the apple scab detection problem with a comparison between (i) the absolute optimal solution,(ii) the optimal solution with a band of 30 nm corresponding to the light-photodetector association of a broad band camera and LED panel and (iii) the optimal solution with a band of 5 nm corresponding to the association of a broad band light with an interferometric filter mounted on a broad band camera. It is, therefore, possible to quantitatively assess the informational gain of a given technological solution here expressed in shannons. The LED panel and interferometric filter solutions appear in shannons, respectively, with a reduction by a factor 0.85 and 1/3 by comparison with the absolute optimal solution which could be implemented by a large band spectrum and a detector mounted with a specifically designed bandpass filter of 55 nm of bandwidth.
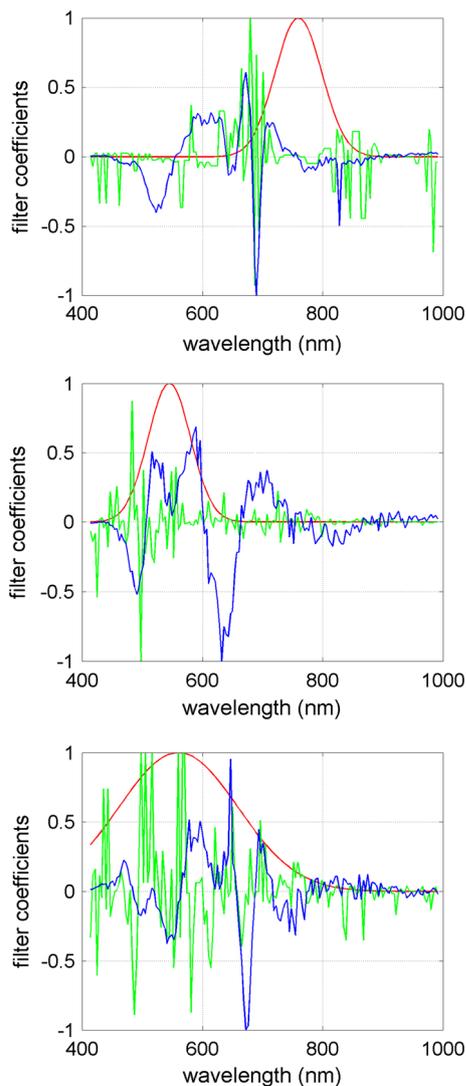
Fig. 8 Comparison of optimal spectral sensitivities selected for a single photodetector by the Kullback–Leibler divergence (*red line*), the coefficients of the PLS filter (*green line*) and the coefficient of the CDA filter (*blue line*). *Top* for the apple scab detection problem, *middle* for the durum and common wheat grain classification and *bottom* with the seedling organs segmentation problem (color figure online)

Table 4 Comparison of informational performance with the Kullback–Leibler divergence of different optimal solutions with a single Gaussian band selection with different technological solutions

1 band with K–L	λ (nm)	w (nm)	K–L 10^{-3}
Optimal	760	55	9.2
LED	750	30	7.9
Interfometric filter	750	10	3.7

Fig. 9 Results of an Otsu threshold applied on the image acquired with a single spectral band optimized with the Kullback–Leibler divergence. This is to be compared with the ground truth of Fig. 2. *Left* detected apple-scab in *dark gray*, *middle* classified durum wheat in *dark gray* and *right* segmented radicle in *dark gray*

To quantitatively assess the informational content of the imaged produce after spectral selection according to the Kullback–Leibler divergence, we now demonstrate that it is very easy to address the informational problems of Fig. 1 with simple hard thresholding detectors. We consider the case of single band selection and we apply a threshold with the automated Otsu method [40]. The detection of apple scab pixels by comparison with the ground truth gives 99.7 % of correct detection in the healthy tissue and 99.8 % in the infected tissue. The classification task of durum versus common wheat grains gives a rate of 98 % of correct classification. The segmentation of radicle and hypocotyl is realized with a rate of 99.2 % of correct classification in the radicle and 98.9 %

in the hypocotyl. Similar results are found when the same Otsu automated threshold method is applied on the gray level images produced by the PLS and CDA method. But, again, the superiority of our waveband selection criterion is in the simplicity of its physical implementation with low-cost optoelectronic devices. The quality of the results, visible in Fig. 9, obtained with a single spectral band low-cost system based on the K–L criterion, has to be appreciated with the fact that these three informational tasks are hardly feasible from a human visual inspection and have previously been tackled in literature with high-cost imaging systems.

5 Conclusion

We presented a general informational methodology based on the Kullback–Leibler divergence for the selection of wavebands used in spectral imaging of plants. We illustrated the value of our approach with three plant imaging problems hard to solve with standard color or gray level imaging. The experimental results are shown to be in agreement with state-

of-the-art spectral reduction approaches like the PLS or the CDA filters. The specific interest of our approach is that the proposed solutions are implementable physically with low-cost LED panels, or filters while the PLS filter requires the use of high-cost hyperspectral imaging systems. Also, the band selection is realized with adjustable bandwidth while usual band selections are realized with fixed bandwidth [41]. This constitutes an extension of the statistical approach recently initiated in [23] with another informational criterion maximizing the information transfer in the sense of Shannon's information theory and which is demonstrated to be efficient also for information detection (Fig. 9).

This opens multiple perspectives. The results obtained produce an interesting practical solution for the design of machine vision applied in proximal detection for plant sciences. The three classification problems addressed here were realized on a small cohort of samples and it would be important to increase the amount of data for the design of a robust waveband selection and include independent validation between training and test samples. However, the principle illustrated in this article remains the same for larger data set. We focused here on binary classification tasks for which a single snapshot with an optimized waveband based on an informational criterion can enhance the contrast and make the classification trivial. This is an important configuration since this corresponds to the lowest possible cost for an imaging system. Also this enabled us to draw comparisons with other generic approaches which also end up with a single component gray level image. Our informational approach is, however, not restricted to a single wavelength selection and can address any spectral reduction problem. In this study, we considered spectra assumed to be fixed with time. When monitoring plants along the development of pathogens, or during the growth process, it is very likely that the spectra of pathogens and healthy tissue or the spectra of different organs of plants will evolve during time. It would, therefore, be interesting to consider the problem of selecting optimal wavelengths to follow a spectral evolution process with a limited number of wavelengths. Also, it would be interesting to specify the wavelengths that would enable to date a spectral evolution signature in time. In this work, we established the best bands from experimental data; it would also be possible to work on simulated hyperspectral reflectance. Such a numerical model has recently been made available online with the website [42] for leaves with possibility to simulate a drop in the major pigments.

References

1. Thenkabail, P.S., Lyon, J.G.: Huete: Hyperspectral Remote Sensing of Vegetation. CRC Press, Boca Raton (2011)
2. Bock, C.H., Poole, G.H., Parker, P.E., Gottwald, T.: Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. Crit. Rev. Plant Sci. **29**, 59–107 (2010)
3. Grahn, H., Geladi, P.: Techniques and Applications of Hyperspectral Image Analysis. Wiley, New York (2007)
4. Vigneau, N., Ecarnot, M., Rabatel, G., Roumet, P.: Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. Field Crops Res. **122**, 25–31 (2011)
5. Behmann, J., Mahlein, A.K., Paulus, S., Kuhlmann, H., Oerke, E. C., Plumer, L.: Generation and application of hyperspectral 3D plant models. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) Computer Vision-ECCV 2014 Workshops. **70**, 117–130. Springer, New York (2014)
6. Rousseau, D., Chéné, Y., Belin, E., Semaan, G., Trigui, G., Boudehri, K., Franconi, F., Chapeau-Blondeau, F.: Multiscale imaging of plants: current approaches and challenges. Plant Methods **11**, 1–6 (2015)
7. Tsaftaris, S.A.: Noutsos: plant phenotyping with low cost digital cameras and image analytics. In: Athanasiadis, I.N., Rizzoli, A.E., Mitkas, P.A., Gómez, M.J. (eds.) Information Technologies in Environmental Engineering, pp. 238–251. Springer, Berlin (2009)
8. Kleynen, O., Leemans, V., Destain, M.-F.: Selection of the most efficient wavelength bands for Jonagold apple sorting. Postharvest Biol. Technol. **30**, 221–232 (2003)
9. Piron, A., Leemans, V., Kleynen, O., Lebeau, F., Destain, M.-F.: Selection of the most efficient wavelength bands for discriminating weeds from crop. Comput. Electron. Agric. **62**, 141–148 (2008)
10. Feyaerts, F., Van Gool, K.: Multi-spectral vision system for weed detection. Pattern Recognit. Lett. **22**, 667–674 (2001)
11. Chao, K., Chen, Y., Hruschka, W., Park, B.: Chicken heart disease characterization by multi-spectral imaging. Appl. Eng. Agric. **17**, 99–106 (2001)
12. Pal, M.: Margin-based feature selection for hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. **11**, 212–220 (2009)
13. Pal, M.: Multinomial logistic regression-based feature selection for hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. **14**, 214–220 (2012)
14. Guo, G., Gunn, S., Damper, R., Nelson, J.: Band selection for hyperspectral image classification using mutual information. IEEE Geosci. Remote Sens. Lett. **3**, 522–526 (2000)
15. De Backer, S., Kempeneers, P., Debruyn, W., Scheunders, P.: A band selection technique for spectral classification. IEEE Geosci. Remote Sens. Lett. **2**, 319–323 (2005)
16. Nakauchi, S., Nishino, K., Yamashita, T.: Selection of optimal combinations of band-pass filters for ice detection by hyperspectral imaging. Opt. Express **20**, 986–1000 (2012)
17. Richter, M., Beyerer, J.: Optical filter selection for automatic visual inspection. In: *IEEE Winter Conference on Applications of Computer Vision (WACV)* **5**, 123–128 (2014)
18. Hansen, P.M., Schjoerring, J.K.: Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. **86**, 542–553 (2003)
19. Thenkabail, P.S., Smith, R.B., De Pauw, E.: Evaluation of narrow-band and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. Photogr. Eng. Remote Sens. **68**, 607–622 (2002)
20. Fiorani, F., Rascher, U., Jahnke, S., Schurr, U.: Imaging plants dynamics in heterogenic environments. Curr. Opin. Biotechnol. **23**, 227–235 (2012)

21. Wold, S., Ruhe, A., Wold, H., Dunn, I.: The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J. Sci. Stat. Comput. **5**, 735–743 (1984)
22. Osborne, S., Kunnemeyer, R., Jordan, R.: Method of wavelength selection for partial least squares. Analyst **122**, 1531–1537 (1997)
23. Benoit, L., Belin, E., Rousseau, D., Chapeau-Blondeau, F.: Information-theoretic modeling of trichromacy coding of light spectrum. Fluct. Noise Lett. **13**, 1–23 (2014)
24. Basseville, M.: Divergence measures for statistical data processing: an annotated bibliography. Signal Process. **93**, 621–633 (2013)
25. Bowen, J.K., Mesarich, C.H., Bus, V.G., Beresford, R.M., Plummer, K.M.: Templeton: *Venturia inaequalis*: the causal agent of apple scab. Mol. Plant Pathol. **12**, 105–122 (2011)
26. Oerke, E.C., Frohling, P., Steiner, U.: Thermographic assessment of scab disease on apple leaves. Precis. Agric. **12**, 699–715 (2011)
27. Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caffier, V., Morel, P., Belin, E., Chapeau-Blondeau, F.: On the use of depth camera for 3D phenotyping of entire plants. Comput. Electron. Agric. **82**, 122–127 (2012)
28. Belin, E., Rousseau, D., Boureau, T., Caffier, V.: Thermography versus chlorophyll fluorescence imaging for detection and quantification of apple scab. Comput. Electron. Agric. **90**, 159–163 (2013)
29. Delalieux, S., Auwerkerken, A., Verstraeten, W.W., Somers, B., Valcke, R., Lhermitte, S., Coppin, P.: Hyperspectral reflectance and fluorescence imaging to detect scab induced stress in apple leaves. Remote Sens. **1**, 858–874 (2009)
30. Mahesh, S., Manickavasagan, A., Jayas, D.S., Paliwal, J., White, N.D.G.: Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. Biosyst. Eng. **101**, 50–57 (2008)
31. Mahesh, S., Jayas, D.S., Paliwal, J., White, N.D.G.: Identification of wheat classes at different moisture levels using near-infrared hyperspectral images of bulk samples. Sen. Instrum. Food Qual. Saf. **5**, 1–9 (2011)
32. Manickavasagan, A., Jayas, D.S., White, N.D.G., Paliwal, J.: Wheat class identification using thermal imaging. Food Bioprocess Technol. **3**, 450–460 (2010)
33. Forcella, F., Arnold, R.L.B., Sanchez, R., Ghersa, C.M.: Modeling seedling emergence. Field Crops Res. **67**, 123–139 (2000)
34. Belin, E., Rousseau, D., Rojas-Varela, J., Demilly, D., Wagner, M.H., Cathala, M.H., Durr, C.: Thermography as non invasive functional imaging for monitoring seedling growth. Comput. Electron. Agric. **70**, 236–240 (2011)
35. Benoit, L., Belin, E., Durr, C., Chapeau-Blondeau, F., Demilly, D., Ducournau, S., Rousseau, D.: Computer vision under inactinic light for hypocotyl radicle separation with a generic gravitropism-based criterion. Comput. Electron. Agric. **111**, 12–17 (2015)
36. Murakami, Y., Obi, T., Yamaguchi, M., Ohyama, N., Komiya, Y.: Spectral reflectance estimation from multi-band image using color chart. Opt. Commun. **188**, 47–54 (2001)
37. Hernández-Andrés, J., Nieves, J.I., Valero, E.M., Romero, J.: Spectral-daylight recovery by use of only a few sensors. J. Opt. Soc. Am. A **21**, 13–23 (2004)
38. Cheung, V., Westland, S., Li, C., Hardeberg, J., Connah, D.: Characterization of trichromatic color cameras by using a new multispectral imaging technique. J. Opt. Soc. Am. A **22**, 1231–1240 (2005)
39. http://www.cie.co.at/
40. Otsu, N.: A threshold selection method from gray-level histograms. Automatica **11**, 23–27 (1975)
41. Piron, A., Leemans, V., Kleynen, O., Lebeau, F., Destain, M.-F.: Selection of the most efficient wavelength bands for discriminating weeds from crop. Comput. Electron. Agric. **2**, 141–148 (2008)
42. http://opticleaf.ipgp.fr/
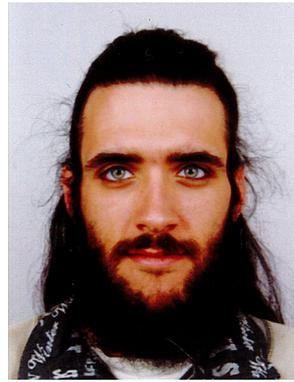
Landry Benoit is a Ph.D. (2012–2015) student in multimodal imaging and image processing for seed and seedling phenotyping at Université d'Angers France.

Romain Benoit is a Ph.D. student (2013–2016), in the field of Robotical analysis through Interval Analysis at Laboratoire Angevin en Ingénierie des Systèmes (LARIS), a research laboratory of the Université d'Angers. He contributed to the present paper during his master internship.

Etienne Belin is an Assistant Professor at Université d'Angers, France. His research activities are at the interface between information (signal and images) processing and physics, with applications in instrumentation and imaging for plant sciences, via the platform PHENOTIC.

Rodolphe Vadaine is an Image Processing Engineer working on multispectral and RGB imaging at GEVES - SNES French National Seed Testing Station - Beaucouzé France.

Didier Demilly is a Research Engineer with INRA, working at GEVES - SNES French National Seed Testing Station - Beaucouzé France. The main topic of his research is the development and use of tools for seed and seedling phenotyping via the platform PHENOTIC.

David Rousseau is professor at Université Lyon 1, France since 2011 with broad interest and skills in imaging and image processing with applications to life sciences including biology and plant sciences. He headed from 2008 to 2012 the PHE-NOTIC platform project with Université d'Angers.

François Chapeau-Blondeau is since 1990 with the University of Angers, France, where he is currently a Professor of Electronic and Information Sciences. His research interests include signal and information processing, imaging, and their applications to life sciences.