



HAL
open science

Réduction de dimensionnalité et saillance pour la visualisation d'images spectrales

Steven Le Moan

► **To cite this version:**

Steven Le Moan. Réduction de dimensionnalité et saillance pour la visualisation d'images spectrales. Autre [cs.OH]. Université de Bourgogne, 2012. Français. NNT : 2012DIJOS034 . tel-00825495

HAL Id: tel-00825495

<https://theses.hal.science/tel-00825495>

Submitted on 23 May 2013

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.



SPIM

Thèse de Doctorat




école doctorale sciences pour l'ingénieur et microtechniques

U N I V E R S I T É D E B O U R G O G N E

DIMENSIONALITY REDUCTION AND SALIENCY
FOR SPECTRAL IMAGE VISUALIZATION

RÉDUCTION DE DIMENSIONALITÉ ET SAILLANCE
POUR LA VISUALISATION D'IMAGES SPECTRALES

 STEVEN LE MOAN

RÉDUCTION DE DIMENSIONALITÉ ET SAILLANCE POUR LA
VISUALISATION D'IMAGES SPECTRALES

DIMENSIONALITY REDUCTION AND SALIENCY FOR
SPECTRAL IMAGE VISUALIZATION

STEVEN LE MOAN

Thèse présentée pour l'obtention du grade de docteur de l'Université de Bourgogne.

Spécialité: INSTRUMENTATION ET INFORMATIQUE DE L'IMAGE

Soutenue le 26 Septembre 2012 devant la commission d'examen :

Prof.	Christine Fernandez-Maloigne	Université de Poitiers	Rapporteur
Prof.	Jean Sequeira	Université d'Aix-Marseille	Rapporteur
Prof.	Jocelyn Chanussot	Université de Grenoble	Examineur
Prof.	Frank Marzani	Université de Bourgogne	Examineur
Prof.	Jon Yngve Hardeberg	Gjøvik University College, Norvège	Co-directeur
Prof.	Yvon Voisin	Université de Bourgogne	Co-directeur
MCF.	Alamin Mansouri	Université de Bourgogne	Encadrant



Bourgogne
Conseil régional

ABSTRACT

Nowadays, digital imaging is mostly based on the paradigm that a combinations of a small number of so-called primary colors is sufficient to represent any visible color. For instance, most cameras use pixels with three dimensions: Red, Green and Blue (RGB). Such low-dimensional technology suffers from several limitations such as a sensitivity to metamerism and a bounded range of wavelengths. Spectral imaging technologies offer the possibility to overcome these downsides by dealing more finely with the electromagnetic spectrum. Multi-, hyper- or ultra-spectral images contain a large number of channels, depicting specific ranges of wavelength, thus allowing to better recover either the radiance or reflectance of the scene. Nevertheless, these large amounts of data require dedicated methods to be properly handled in a variety of applications. This work contributes to defining what is the useful information that must be retained for visualization on a low-dimensional display device. In this context, subjective notions such as appeal and naturalness are to be taken into account, together with objective measures of informative content and dependency. Especially, a novel band selection strategy based on measures derived from Shannon's entropy is presented and the concept of spectral saliency is introduced.

RÉSUMÉ

De nos jours, la plupart des dispositifs numériques d'acquisition et d'affichage d'images utilisent un petit nombre de couleurs dites primaires afin de représenter n'importe quelle couleur visible. Par exemple, la majorité des appareils photos "grand public" quantifient la couleur comme une certaine combinaison de Rouge, Vert et Bleu (RVB). Ce genre de technologie est qualifiée de tri-chromatique et, au même titre que les modèles tetra-chromatiques communs en imprimerie, elle présente un certain nombre d'inconvénients, tels que le métamérisme ou encore la limitation aux longueurs d'onde visibles. Afin de palier à ces limitations, les technologies multi-, hyper, voire ultra-spectrale ont connu un gain notable d'attention depuis plusieurs décennies. Un image spectrale est constituée d'un nombre de bandes (ou canaux) supérieur à 3, représentant des régions spectrales spécifiques et permettant de recouvrer la radiance ou réflectance d'objets avec plus de précision et indépendamment du capteur utilisé. De nombreux travaux de recherche ont fait considérablement progresser les méthodes d'acquisition et d'analyse, mais beaucoup de challenges demeurent, particulièrement en ce qui concerne la visualisation de ce type de données. En effet, si une image contient plusieurs dizaines de canaux comment la représenter sur un écran qui n'en accepte que trois ? Dans cette thèse, nous présentons un certain nombre de méthodes d'extraction d'attributs pour l'analyse d'images spectrales, avec une attention particulière sur la problématique de la visualisation.

ACKNOWLEDGMENTS

During these three years, in Auxerre, Gjøvik, and everywhere else, I received the support and guidance of many people. This is my chance to say thank you.

To begin with, I would like to thank my advisors Alamin Mansouri, Jon Yngve Hardeberg and Yvon Voisin for knowing how and when to push me forward during these years, but also for giving me the opportunity to visit so many places and meet so many people across the world, thank you for everything. And I also thank the Regional Council of Burgundy, who provided the funding for this thesis, as well as the I3Mainz institute and Norsk Elektro Optikk for providing useful data.

Hails and thanks to my colleagues and friends in France and Norway: Gabriele, Daniel, Dibakar, William, Ferdinand, Marius, Jonathan, Cédric, Fabrice, and many others. It has been an honor meeting you all and I am sure that we will keep in touch!

Last but not least, this work would have been nothing without the support of my family. Thank you.

CONTENTS

1	INTRODUCTION	1
1.1	Motivation	1
1.2	Contributions	3
1.2.1	International journals	3
1.2.2	International conferences	4
1.2.3	National conferences	5
1.2.4	Workshops / Symposiums	5
1.3	Dissertation outline	5
1.4	Notation used throughout this document	6
2	DIMENSIONALITY REDUCTION AND DISPLAY OF SPECTRAL IMAGES: A REVIEW	7
2.1	Introduction	7
2.2	A note on color spaces	8
2.3	Fixed basis functions	9
2.3.1	True color	9
2.3.2	Other bases	12
2.4	Principal Component Analysis	13
2.4.1	Definition	14
2.4.2	Alternatives	15
2.5	Independent Component Analysis	17
2.6	Other linear transformations	20
2.6.1	Projection Pursuit	20
2.6.2	Multiresolution analysis	21
2.6.3	Conical mapping	22
2.6.4	Pre-defined correlation	23
2.7	Non-linear dimensionality reduction	24
2.8	Band selection	24
2.8.1	Fixed band selection	25
2.8.2	Statistical methods	25
2.8.3	Information theory	26
2.8.4	Linear Prediction	28
2.8.5	One-bit Transform	30
2.9	Preprocessing and normalization	31
2.9.1	Radiance vs. Reflectance	31
2.9.2	Low SNR and low entropy	31
2.9.3	Invariant representation	32
2.9.4	Scaling and clipping	33
2.10	On the evaluation of trichromatic composites	34
2.10.1	Conveyed information	35
2.10.2	Intrinsic information	35
2.10.3	Appeal	36
2.11	Conclusions	36

3	DIMENSIONALITY REDUCTION AND DISPLAY OF SPECTRAL IMAGES: CONTRIBUTIONS	39
3.1	Introduction	39
3.2	Benchmarking data	39
3.3	New quality measures for dimensionality reduction	40
3.3.1	Naturalness	40
3.3.2	Color entropy	41
3.3.3	Experimental validation	41
3.3.4	Conclusions	44
3.4	A perception-based spectrum segmentation	44
3.4.1	Definition	45
3.4.2	Experimental setup and results	47
3.4.3	Conclusion	48
3.5	An information-theoretic framework for band selection	48
3.5.1	Background	49
3.5.2	Band selection algorithm	56
3.5.3	Experimental setup and results	61
3.5.4	Conclusions	73
3.6	Cluster-Adaptive Projections	73
3.6.1	Statement of the problem	73
3.6.2	CAP	75
3.6.3	Experimental setup and results	76
3.6.4	Conclusion	76
3.7	Conclusions	77
4	SALIENCY AND SPECTRAL IMAGES	79
4.1	Introduction	79
4.2	High Dimensional Saliency	81
4.2.1	Architecture	81
4.2.2	Features and comparison measures	81
4.2.3	Activation, normalization and fusion	84
4.2.4	In color	86
4.2.5	Results	89
4.2.6	Applications	90
4.3	Spectral Saliency Profiles	97
4.3.1	Definition	97
4.3.2	Application to band selection	98
4.4	Conclusions	100
5	CONCLUSIONS AND PERSPECTIVES	103
	BIBLIOGRAPHY	105

INTRODUCTION

1.1 MOTIVATION

When it comes to digital imaging, most of today's acquisition and display hardware such as computer screens, digital cameras, scanners or printers are based on the paradigm that a combination of a small number of primary colors is sufficient to characterize any color that one can see [Grassmann, 1854]. Digital camera for instance, record the color information as a combination of Red, Green and Blue (RGB) and therefore uses a three-dimensional space in which each location corresponds to a different color. Such technology is referred to as trichromatic, and suffers from several drawbacks such as metamerism and the impossibility to represent non-visible wavelengths.

The colors, as perceived by humans, are the result of a complex interaction of a light source, striking and being reflected by objects, and the sensor (camera, eye). Two sensors with different spectral responses can therefore render a scene in two very different ways. Similarly, distinct light sources may produce a different color sensation on a same object. This is for instance the reason why it is recommended to check the color of your clothes outside the shop before buying it. Consequently, when too little spectral information is acquired, pixels become dependent on the illumination and the sensor and can therefore not be considered as invariant features. A good example to illustrate this is metamerism, which is a phenomenon that implies that two objects which appear to have exactly the same color, may look very different under different lighting conditions [Hardeberg, 2001] (see Figure 1). Such objects are metameric color matches, as opposed to spectral color matches that have identical reflectance distributions. In the industry, this "problem", which is actually the very basis of color science, is critical for instance when advertising for products for which color matching and/or tolerance are important, such as clothes or orthodontics. This phenomenon can be dealt with by increasing the precision with which color is measured, which pertains to the spectral resolution of the image. One can then think of an analogy with spatial resolution: the fewer the number of pixels (low resolution), the harder it gets to identify shapes and/or textures.

The second critical problem of low-dimensional imaging is that most cameras capture light solely in the visible wavelengths, which is roughly 400-700nm (see Figure 2). Not all imaging science is dedicated to reproduce what we see, far from it. A lot of very interesting things happen outside the visible range, and not only for scientists as

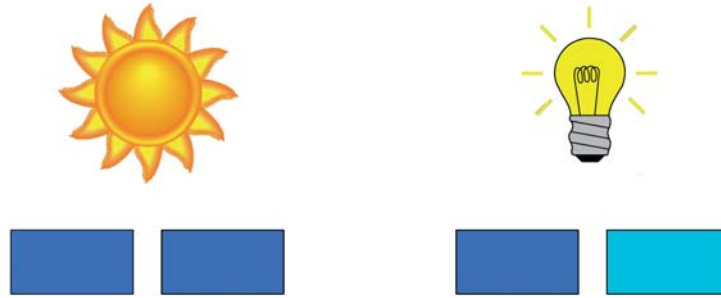


Figure 1: Illustration of metamerism. The two patches appear as similar under sunlight, but different under another illuminant.

there is for instance an increasing interest of near infrared (nIR) photography, because of the unique colors that it allows to render. Classical examples of picturing the non-visible can be found in medical imaging (magnetic resonance, X-ray), surveillance (thermography).

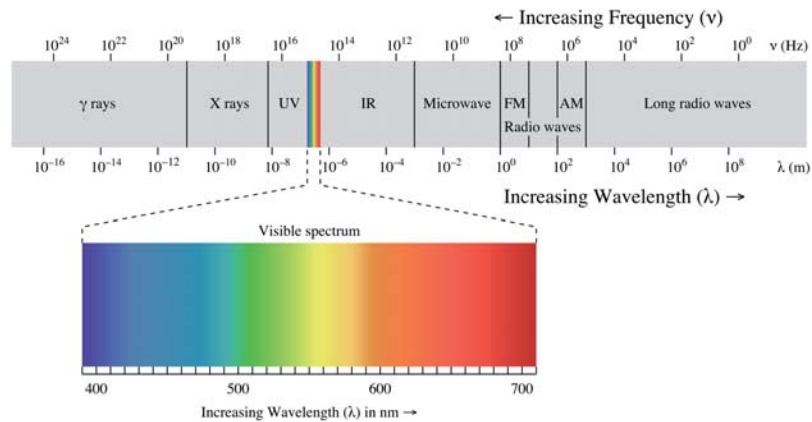


Figure 2: The electro-magnetic spectrum and the visible range (source: wikipedia).

As an answer to these drawbacks, spectral imaging technologies offer the possibility to deal more finely with the electromagnetic spectrum by acquiring light in small ranges of wavelengths, possibly non-visible, and by conveying their respective contents in so-called spectral channels (see Figure 3).

For purposes of readability and genericity, we use in this thesis the expression *spectral* imaging whereas there exists an actual categorization of these images, according to their respective number of channels. Although it is not strict and may vary from one application to another, it is globally recognized that:

- **Multispectral** images contain between 4 and 99 channels and usually range only in the visible wavelengths.

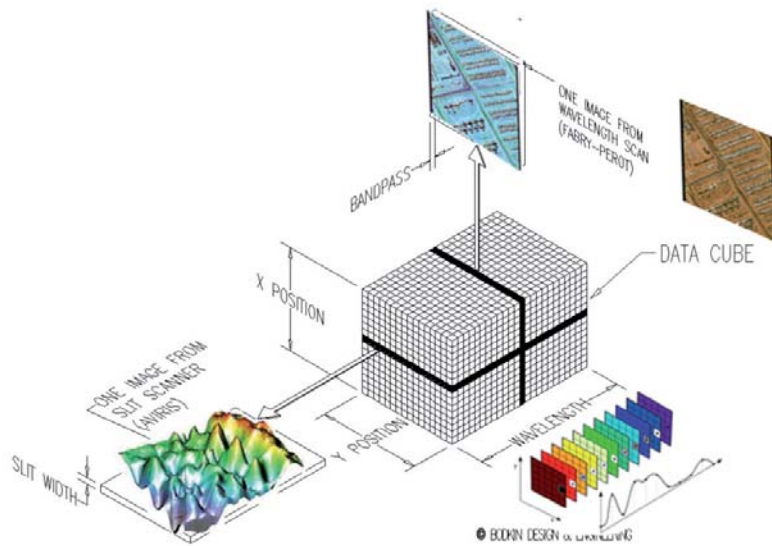


Figure 3: Illustration of a hyperspectral cube (source: hyperspectralimagers.com).

- **Hyperspectral** images contain between 100 and 999 channels and usually range in the visible and near IR (700-1000nm).
- **Ultraspectral** images contain more than 1000 channels and allow to display molecular absorption or emission line features.

Although there is a considerable interest for these technologies on the acquisition side, spectral display devices are yet to become popular. Nevertheless, human interaction with spectral images is crucial to direct or validate automated methods, and to make appropriate decisions and interpretation. But then, how to display an image with dozens of spectral channels on a screen that allows only three?

1.2 CONTRIBUTIONS

The work presented in this thesis led to the publication of several papers. Please find hereafter a list of these contributions.

1.2.1 International journals

- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, "A constrained band selection method based on information measures for spectral image color visualization", *Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 5104–5115, 2011, IEEE.
- (Submitted, August 2012) Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, "Saliency for spectral

image analysis”, to *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE.

- (Submitted, May 2012) Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Visualization of multi/hyperspectral images: a comparative study”, to *International Journal of Remote Sensing*, Taylor & Francis.

1.2.2 International conferences

- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Convex objects recognition and classification using spectral and morphological descriptors”, in *Colour in Graphics, Imaging, and Vision (CGIV)*, 5th European Conference on, Joensuu, Finland, June 2010, pp. 293–299, IS&T.
- Steven Le Moan, Alamin Mansouri, Jon Yngve Hardeberg, and Yvon Voisin, “A class-separability-based method for multi/hyperspectral image color visualization”, in *Image Processing (ICIP)*, 17th International Conference on, Hong Kong, Sep 2010, pp. 1321–1324, IEEE.
- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “An efficient method for the visualization of spectral images based on a perception-oriented spectrum segmentation”, in *Advances in Visual Computing (ISVC)*, 6th International Symposium on, Las Vegas, NV, USA, Nov 2010, vol. 6453 of Lecture Notes in Computer Science, pp. 361–370, Springer.
- Steven Le Moan, Alamin Mansouri, Jon Yngve Hardeberg, and Yvon Voisin, “Saliency in spectral images”, in *Image Analysis (SCIA)*, 17th Scandinavian Conference on, Ystad, Sweden, May 2011, vol. 6688 of Lecture Notes in Computer Science, pp. 114–123, Springer.
- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Spatially variant dimensionality reduction for the visualization of multi/hyperspectral images”, in *Image Analysis and Recognition (ICIAR)*, International Conference on, Burnaby, BC, Canada, June 2011, vol. 6753 of Lecture Notes in Computer Science, pp. 375–384, Springer.
- Steven Le Moan, Alamin Mansouri, Jon Yngve Hardeberg, and Yvon Voisin, “Visualization of spectral images: a comparative study,” in *Proceedings of the 6th Gjøvik Color Imaging Symposium*, September 2011, IEEE.
- Steven Le Moan, Alamin Mansouri, Jon Yngve Hardeberg, and Yvon Voisin, “Saliency-based band selection for spectral image

visualization”, in *Proceedings of the 19th Color Imaging Conference*, San Jose, CA, USA, November 2011, IS&T.

- Steven Le Moan, Ferdinand Deger, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Salient pixels and dimensionality reduction for display of multi/hyperspectral images”, in *Image and Signal Processing, 5th International Conference on*, Agadir, Morocco, June 2012, Lecture Notes in Computer Science, Springer.

1.2.3 National conferences

- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Visualisation d’images spectrales : une méthode basée sur la perception humaine”, in *Proceedings of ORASIS*, Praz-sur-Arly, France, June 2011.
- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Sélection de bandes pour la visualisation d’images spectrales: une approche basée sur l’étude de saillance”, in *Proceedings of the 23rd Colloque GRETSI*, Bordeaux, France, September 2011.

1.2.4 Workshops / Symposiums

- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Hardeberg, “Visualization of three-dimensional spectral data”, in *Colour in Art, Science, Design, Conservation, Research, Printmaking, Digital Technologies, Textiles (CREATE)*, Gjøvik, Norway, June 2010, IS&T.

1.3 DISSERTATION OUTLINE

This thesis is organized as follows. Chapter 2 provides a review of the state-of-the-art in terms of dimensionality reduction for spectral images, with a focus on visualization. In Chapter 3, several contributions to the visualization of spectral images are introduced, including two new quality measures, a perception-based spectrum segmentation, an information-based framework for band selection as well as a brief study on distance-preserving adjustments. Finally, Chapter 4 introduces the concept of non-visual saliency for spectral images and present a variety of applications. Chapter 5 concludes this dissertation and discusses possible future work based on the results reported here.

1.4 NOTATION USED THROUGHOUT THIS DOCUMENT

A low-dimensional representation of a spectral image is referred to as a *composite*.

Vectors are represented in lowercase boldface letters and organized in column, such as:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_L \end{bmatrix}$$

Matrices are represented as uppercase boldface letters:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{L1} & a_{L2} & \cdots & a_{LM} \end{bmatrix}$$

The transpose of \mathbf{A} is noted \mathbf{A}^T .

A spectral channel centered at wavelength λ is represented as the column vector \mathbf{b}_λ , unless spatial coordinates (i, j) are necessary, in which case it is noted as the 2-D matrix $\mathbf{B}_\lambda(i, j)$. For purposes of clarity, we use the following equivalence:

$$\mathbf{b}_1 = \mathbf{b}_{\lambda_1} = \mathbf{b}_{\lambda=\lambda_1}$$

\mathbf{B} denotes a spectral image containing N channels, so that:

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix}$$

The probability mass function of a discrete random variable X is noted equivalently $p_X(x)$ or p_X . The range of possible values of X is denoted $\mathcal{R}(X)$.

In all the definitions of information measures such as entropy, we chose to use a base-2 logarithm to be able to express the results in bits, which we believe to be an intuitive and relevant unit for this study.

All metrics and measures are noted in italics such as $H(X)$.

DIMENSIONALITY REDUCTION AND DISPLAY OF SPECTRAL IMAGES: A REVIEW

2.1 INTRODUCTION

Traditional imaging uses a small number of primaries to represent the full range of perceptible colors. Though sufficient for many applications, it has two critical limitations: metamerism and a bounded range of wavelengths. In trichromacy, pixels can be represented as vectors of values in some three-dimensional space. By extension, pixels from spectral images are high-dimensional vectors, which means that they are defined by a large number of dimensions¹. The number of these, that is the cardinality of the vectors, is also referred to as their **dimensionality**. For example, well-known devices such as the Airborne Visible / Infrared Imaging Spectrometer (AVIRIS)² or HySpex³ contain between 64 and 256 channels. Nevertheless, such high dimensionality engenders a variety of problems and there are therefore several motivations to reduce it:

- **To fit low-dimensional imaging hardware.** Evidently, this is the critical point in this study. Although there is a considerable interest for these technologies on the acquisition side, spectral display devices are yet to become feasible.
- **To emphasize the valuable information.** As high dimensionality implies a tremendous amount of data, spectral images usually contain a lot of redundant information and noise. They also engender some peculiar properties [Jimenez and Landgrebe, 1998] referred to as the *curse of dimensionality*. For example, the volume of a hypercube concentrates in the corners, and the volume of a hypersphere or hyperellipsoid concentrates in an outer shell, which implies that most part of the high-dimensional space is empty. Thus, the meaningfulness of the Euclidean distance is drastically decreased, which may influence distance-based clustering algorithms. Consequently, there is a real challenge as to extract the proper information from spectral images. Moreover, this information may not even be contained in one, but several projections [Parsons et al., 2004].
- **To reduce the size of the data.** Spectral images are usually large datasets and therefore are computationally expensive to manip-

¹ also called attributes, features, components, channels, or bands

² <http://aviris.jpl.nasa.gov/>

³ <http://www.neo.no/hyspex/>

ulate. Reducing the dimensionality of a dataset also reduces its size and therefore can be used as a data compression technique.

There are many different approaches to perform Dimensionality Reduction (DR) and they can be classified according to various criteria. For instance, the DR process can be either **supervised**⁴ or **unsupervised**, depending on whether or not there is a need for *a priori* available data (training set, ground truth,..). DR methods can also project the data in a **linear** or **non-linear** fashion. They can either **transform** the original dimensions (spectral channels) or **select** a subset of them in order to spawn the output reduced feature space. Finally, DR techniques can be ranked according to a **data/vision**-dependency scale. It will be shown in the next parts of this Section that some techniques are indeed purely data-dependent, while some others, inspired by human vision, are completely independent of the data.

After a brief note on the different manners to represent color in trichromacy, the next sections review the state-of-the-art in DR for display of spectral images, with a particular focus on unsupervised an linear techniques, as most of the work presented in this thesis lies under that scope.

2.2 A NOTE ON COLOR SPACES

There are many different approaches to represent color in low dimensionality. In printing industry for instance, it is common to use the four-dimensional **CMYK** (Cyan, Magenta, Yellow and Black) space. In trichromacy, the most widely used representation of color for display and capture devices is by a combination of Red, Green and Blue (**RGB**). There exist several RGB color spaces such as CIERGB, Adobe 1998 RGB, Apple RGB, etc., all of them representing colors in a (slightly) different manner. In 1996, HP and Microsoft designed a standard RGB space (**sRGB**) [Stokes et al., 1996] that was endorsed by many brands such as Intel or Corel.

The RGB spaces are however limited inasmuch as the distance between two locations in such spaces does not correlate with the actual difference of perception. The CIE 1931 Standard **XYZ** Colorimetric system is a theoretical model based on perception only. It introduced the notion of luminance (Y) as a subjective light intensity, independent of chrominance, while X and Y were chosen so that visible colors are expressed with positive values only. Moreover, **XYZ** was the first step towards a perceptually uniform space or, in other words, a space in which that the euclidean distance between two locations is consistent with the difference of perception by human vision. The more recent **CIELAB** color space has such uniformity. It was derived from

⁴ also referred to as assisted

XYZ and characterizes a color by means of an intensity parameter (L^*) which pertains to luminance and two chrominance parameters (a^* and b^*). The Euclidean distance in CIELAB is commonly noted ΔE_{ab}^* .

Finally, other representations may be useful to process the data more intuitively such as Hue-Saturation-Value (**HSV**) (also known as Hue-Saturation-Brightness) and Hue-Saturation-Lightness/Luminance (**HSL**). The difference between these two is that the brightness of a pure color is equal to the brightness of white, while the lightness of a pure color is equal to the lightness of a medium gray.

Figure 4 depicts the geometry of the RGB and HSV color spaces.

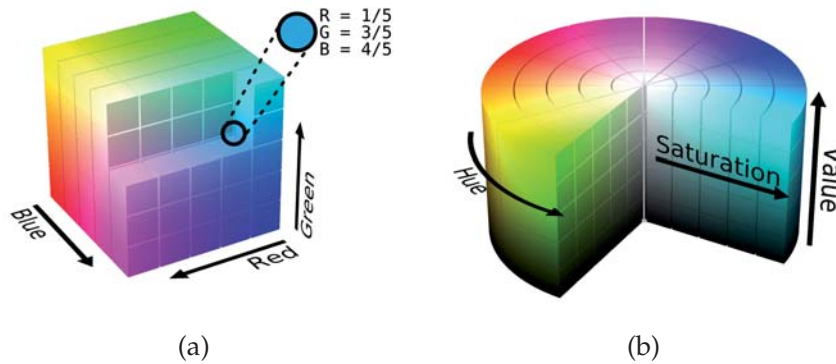


Figure 4: Geometry of color spaces: (a) RGB and (b) HSV (source: wikipedia).

2.3 FIXED BASIS FUNCTIONS

Dimensionality reduction is all about projecting the data (spectral image) on a set of functions which, as stated earlier, can either be data- or vision-dependent. In this section, we tackle the latter case, in which the functions are fixed and known prior to even analyzing the data.

2.3.1 True color

Human perception of colors is known to be a very complex system. The back of the human eye contains photoreceptors (*cones* and *rods*) that react dependently on the wavelength of light and are therefore the cells responsible for the perception of color. The so-called L, M, and S cones are respectively sensitive to the short, medium and long wavelengths (centered around 440, 540 and 575nm, resp.) as depicted in Figure 5. These photoreceptors forward the information to other kinds of cells (*bipolar* and *ganglion*), that eventually convert the signal to neural activity for the brain to understand and process. This color information is "coded" in three dimensions, that is, by means of an

achromatic and two color-opponent channels. The achromatic information pertains to the luminance or overall intensity of light whereas the two other dimensions represent the opposition red/green and blue/yellow. This condition has for instance inspired the design of several color spaces such as CIELAB (see Section 2.2). Needless to say, all people do not perceive color in the same way, as there might be slight discrepancies between two healthy observers.

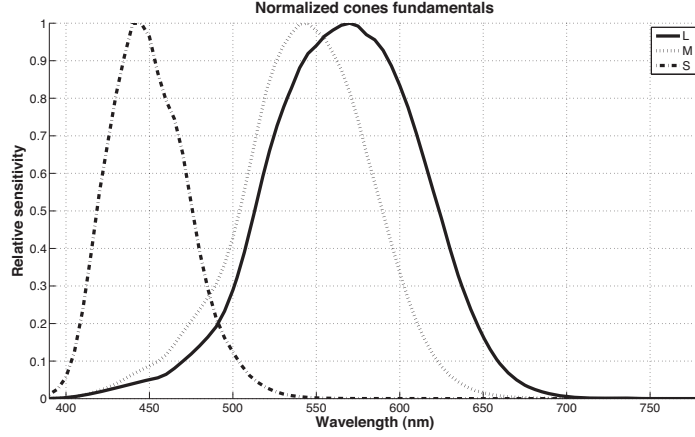


Figure 5: Normalized sensitivities of the L, M and S cones.

In 1931, the *Commission Internationale de l'Eclairage* (CIE) designed a linear computational model to imitate the aforementioned transduction, that is, to turn a specific light spectrum into a corresponding tristimulus and for a *standard observer*. It was obtained through a series of psychophysical experiments in which each viewer had to determine the amount of the primary colors (red, green and blue with primary wavelengths $R_0 = 700$ nm, $G_0 = 546.1$ nm, $B_0 = 435.8$ nm, respectively) that was necessary to have the same color sensation as a monochromatic light at a particular wavelength. The resulting model involves a dimensionality reduction by means of so-called Color Matching Functions (CMF): $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$, which are sketched in Figure 6. For a wavelength-dependent pixel $p(\lambda)$, e.g. in a spectral image, the tri-stimulus values corresponding to an amount of red, green and blue are respectively given by:

$$R_p = \int_{\lambda \in \Lambda_V} p(\lambda) \bar{r}(\lambda) d\lambda, \quad (1)$$

$$G_p = \int_{\lambda \in \Lambda_V} p(\lambda) \bar{g}(\lambda) d\lambda, \quad (2)$$

$$B_p = \int_{\lambda \in \Lambda_V} p(\lambda) \bar{b}(\lambda) d\lambda, \quad (3)$$

with Λ_V representing the visible range of wavelengths. Additionally, the CMF can be defined according to the CIE 1931 Standard XYZ Colorimetric System, so as to have only positive coefficients. The conversion from RGB to XYZ CMF is given by the following linear transform:

$$\begin{bmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.0 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} \bar{r}(\lambda) \\ \bar{g}(\lambda) \\ \bar{b}(\lambda) \end{bmatrix} \quad (4)$$

Note that, in each case, the *equal energy white point* condition must be satisfied, that is, for a flat light spectrum (same energy at each wavelength), the resulting tri-stimulus must be either white or gray, which is consistent with the definition of a perfect reflector. In order for this condition to be satisfied, the following must stand true:

$$\int_0^\infty \bar{r}(\lambda) d\lambda = \int_0^\infty \bar{g}(\lambda) d\lambda = \int_0^\infty \bar{b}(\lambda) d\lambda, \quad (5)$$

and identically for the XYZ functions.

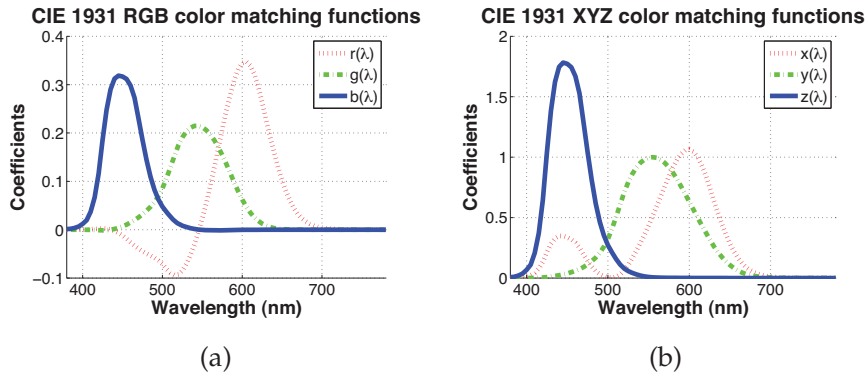


Figure 6: The CIE standard observer color matching functions

When applied to a whole spectral image, a CMF transformation outputs a so-called *true color* composite, as opposed to *false color* or *pseudocolor* in the case of other DR techniques. Such composite results evidently in a very "natural" rendering and, because the spectral weighting envelopes are fixed, the *true color* transformation is also very efficient computationally-wise. Moreover, the association of a color to a light spectrum is data-independent, and thus allows the end user to give a consistent meaning to colors, making interpretation easier. Nevertheless, it belongs to this vision-inspired class of DR methods that do not take the data itself into account at all and therefore, noise, redundancy, and the overall repartition of information along the spectrum are not efficiently handled. Note for instance how the coefficients at the edges of the visible range tend to zero altogether, making the channels in these regions almost neglected in

terms of lightness and saturation. Another downside of using the CMF for display purpose is that, in many applications, spectral images range outside the visible spectrum and actually contain valuable information for instance in the nIR. [Jacobson and Gupta \[2005\]](#) proposed an alternative use of the CMF by stretching them so as to fit the entire range of wavelengths covered by the spectral image. This trick comes evidently with a loss of physical meaning of the output tristimuli but guarantees nonetheless a physically consistent correlation of the red, green and blue channels, which results in natural and intuitive color associations (note that the paper only presents results on hyperpectral images acquired with the AVIRIS sensor).

Figure 7 gives an example of CMF-based visualization on an image ranging to the nIR, with and without stretching.

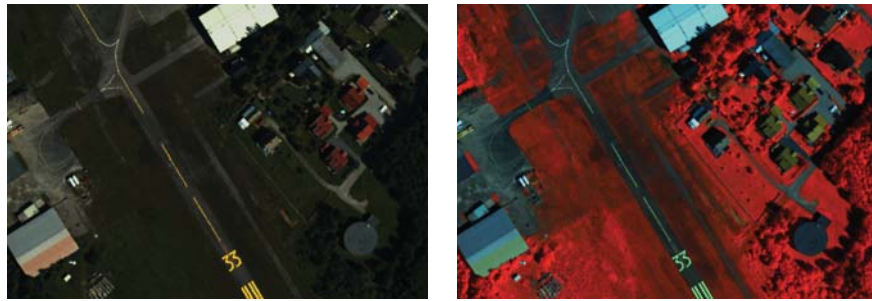


Figure 7: Examples of CMF-based visualization. On the left: the *true color* composite. On the right: result of stretching the CMF to the whole range of wavelengths of the same image. The first one looks more "natural" and allows to easily interpret the scene, while the second one emphasizes the conspicuity of living materials, which are particularly reflective in the nIR.

2.3.2 Other bases

[Jacobson et al. \[2007\]](#) investigated other fixed weighting functions for linear DR and display of spectral and multimodal images. They used three different bases. The constant-luma disc basis (Figure 8a) samples a circular curve in the CIELAB color space, with parameters $L = 50$ and $\sqrt{a^2 + b^2} = 29$ and at a constant ΔE_{ab}^* (Euclidean distance in CIELAB). It allows for an equal luma rendering, to make sure that all channels are given the same emphasis (as opposed to the CMF), as well as an equal hue difference, that is, the difference of rendered hue for a pair of Kronecker deltas (monochromatic lights) is dependent on their respective distance only. Nevertheless, the main downside of this basis is that the endpoint colors (edges of the image spectrum) are almost identical, which drastically decreases the meaningfulness of colors in the composite. Moreover, it does not exactly satisfy the equal energy white point condition as the area under

each curve is slightly different. The unwrapped cosine basis (Figure 8b) is somewhat inspired by the shape of the constant-luma basis and is therefore made of three out-of-phase sinusoids, one of them being modified so as to be monotonically increasing, in order to obtain distinct endpoint colors. This time, all three curves sum to the same value. Finally, the constant-luma border basis (Figure 8c) samples a curve which progresses in sRGB from blue ($R=0.122$, $G=0.122$, $B=1$) to green ($R=0$, $G=0.256$, $B=0$) to red ($R=0.866$, $G=0$, $B=0$) at approximately constant ΔE . Like the first basis, it gives an equal luma rendering all across the spectrum but this time with distinct endpoint colors.

These approaches are very efficient computationally-wise since the weighting envelopes are fixed. However, and although a Signal-to-Noise ratio (SNR) adaptation of the bases is also proposed, they suffer from a lack of adaptivity to the repartition of the valuable information over the image spectrum.

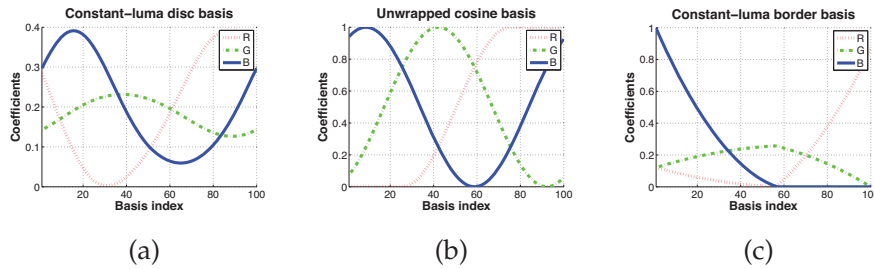


Figure 8: The constant-luma disc (a), unwrapped cosine (b) and constant-luma border (c) bases.

2.4 PRINCIPAL COMPONENT ANALYSIS

Compared to the methods detailed in the previous section, Principal Component Analysis⁵ (PCA) is on the other side of the data/vision-dependent scale. This completely data-adaptive orthogonal linear transformation projects the data to an equally or lower dimensional space of uncorrelated attributes called Principal Components (PCs). It uses variance (energy) as a measure of valuable information and derives the new data so as to keep as much of it as possible (see Figure 9). The next sections will guide you through the different steps of PCA before reviewing some state-of-the-art alternative uses of it.

⁵ also called Principal Component Transform (PCT), Karhunen-Loève Transform (KLT), the Hotelling Transform or Proper Orthogonal Decomposition (POD)

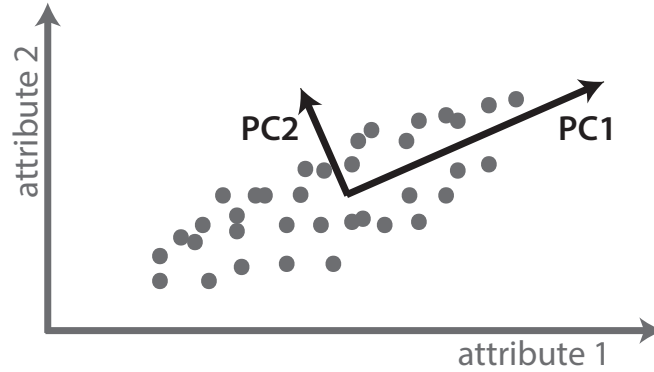


Figure 9: The principle of PCA: the first principal component (PC1) is the direction that contains most of the data's variance.

2.4.1 Definition

Let us consider a spectral image \mathbf{B} with zero mean, obtained by removing the average value across all dimensions.

- Step 1: Covariance matrix.** The covariance of two random variables is defined as a measure of how much they change together. Each column of \mathbf{B} depicts a different random variable, thus, measuring their covariance allows to evaluate the respective similarities/discrepancies between spectral channels. The covariance matrix of data matrix \mathbf{B} is defined as: $\Sigma = \mathbf{B}\mathbf{B}^T$ and contains all covariances of pairwise channels.
- Step 2: Eigendecomposition of the covariance matrix.** This step consists of factorizing Σ into its canonical form, that is, decomposing the matrix in a set of vectors $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$ and corresponding values $\Lambda = [\lambda_1 \ \lambda_2 \ \dots \ \lambda_n]$ that satisfy the following property: $\Sigma\mathbf{v}_i = \lambda_i\mathbf{v}_i, \forall i = 1 \dots n$.
- Step 3: Ranking eigenvectors to compute projection matrix.** This is the core of PCA. Each previously computed eigenvalue λ_i is a measure of how significant its corresponding eigenvector \mathbf{v}_i is, that is, how much of the data's energy is contained in the direction pointed out by \mathbf{v}_i . These eigenvectors are also referred to as Principal Components (PCs) of \mathbf{B} . Most of the information contained in \mathbf{B} can be conveyed by projecting it on the top-ranked eigenvectors, the first PCs. Thus, the projection matrix \mathbf{P} consists of the N' ($\leq N$) first eigenvectors sorted accordingly, column-wise. In most cases, more than 95% of the data's energy is explained by the three first PCs [Tsagaris et al., 2005].

- **Step 4: Deriving the new dataset.** The output dataset \mathbf{B}' is the result of a simple linear operation: $\mathbf{B}' = \mathbf{B}\mathbf{P}^T$.

For a more in-depth tutorial on PCA, refer to [Smith, 2002].

Note that a straightforward mapping of the first three PCs to the sRGB color space is usually not a good strategy for several reasons. First, it implies a ranking of the Red, Green and Blue primaries as all PCs do not contain the same amount of information (energy-wise). Therefore, using the first PC as the red component will result systematically in a reddish rendering. Second, PCA spawns a gamut that might not fit the one of sRGB (negative values, sparse outliers) without proper normalization of the PCs. Not to mention the fact that PCA has the critical downside to be able to confuse high-energy noise for information as variance does not necessarily reflect real SNR in spectral imaging due to unequal noise variances among spectral channels [Green et al., 1988]. Finally, some objects might exist only in subsets of the high-dimensional space [Parsons et al., 2004] and would therefore require a local analysis rather than considering all channels at once. In the next parts, we review some methods to overcome these drawbacks.

2.4.2 Alternatives

2.4.2.1 Vision-consistent PCA

According to a study by Buchsbaum and Gottschalk [1983], our vision works in a PCA-like manner as the achromatic channel (A) holds about 95.7% of the stimulus information, while the red/green (R-G) and blue/yellow (B-Y) opponents channels contain respectively 2.8% and 1.5% of it. On the other hand, the first PC of a spectral image is generally positive and very close to its average radiance and, as their rank decreases, PCs contain less and less spatial information. This condition suggests a convenient mapping of the first three PCs to the A, (R-G) and (B-Y) channels, respectively as was investigated in [Tyo et al., 2003] by means of the HSV color space and in [Zhang et al., 2006] in CIELAB, YCrCb and YUV. Results show consistent colors for regions with similar spectral properties, and good contrast, allowing to coarsely identify pixels classes rapidly and accurately. However, the overall association colors/objects is non-intuitive and the rendering is consequently poor in terms of naturalness.

2.4.2.2 MNF, NAPCA and INAPCA

As stated earlier, variance is generally not a good measure of informative content in spectral imaging inasmuch as noise variance may not be homogenous over the spectral channels. Green et al. [1988] proposed a SNR-adaptive transform called Minimum Noise Fraction

(MNF) which ranks principal components in terms of Signal-to-Noise-Ratio (SNR) rather than variance. It was later re-interpreted as the Noise-Adjusted PCA (NAPCA) [Lee et al., 1990], which consists of normalizing the data's SNR by whitening the data's noise, that is reducing its variance to unity (1) in each channel, prior to the traditional PCA. This model assumes that the data's covariance matrix Σ can be decomposed as a sum of signal and noise components:

$$\Sigma = \Sigma_s + \Sigma_n \quad (6)$$

Let F be a whitening matrix so that:

$$F^T \Sigma_n F = I \quad (7)$$

with I the identity matrix. Then, transforming Σ by F , i.e. $\Sigma_{\text{adj}} = F^T \Sigma F$ yields a noise-adjusted data covariance matrix Σ_{adj} , the eigendecomposition of which thus allows to rank PCs in terms of SNR instead of overall variance.

Chang et al. [1998] introduced the notion of interference as a separate undesirable source, defined as clutter or structured noise and which motivated them to later design the Interference- and Noise-Adjusted PCA (INAPCA) [Chang and Du, 1999], derived in two models. The first one aims to increase the so-called Signal to Noise plus Interference Ratio (SNIR), treating the interference as part of the noise, while the second model tends to remove the interferers prior to performing NAPCA.

The major drawbacks of these adjusted PCA techniques is that they all depend on the estimation of noise (and/or interference) covariance matrices, which is a device-dependent non-trivial task, especially in an unsupervised framework.

2.4.2.3 Segmented PCA

As explained earlier, PCA finds the most informative directions (energy-wise) in the whole high-dimensional feature space, while important objects might exist only in subsets of it [Parsons et al., 2004]. Jia and Richards [1999] introduced a (spectrally) local approach to PCA, based on a partitioning of the image spectrum. In their design, bands are first gathered in several contiguous and compact groups (clusters), so as to maximize the intra-cluster correlation by finding edges in the image correlation matrix. The latter is indeed known to be generally structured in blocks as shown in Figure 10. Eventually, only one PC is extracted from each subgroup to create the lower dimensional composite.

This segmented PCA approach was further investigated by Tsagaris et al. [2005] with different partitioning strategies. The *equal subgroups*

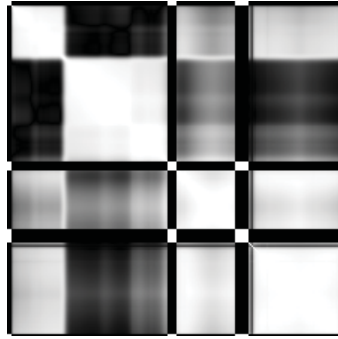


Figure 10: Example of correlation matrix, on a 224-channel hyperspectral image from the AVIRIS sensor. Each location (i, j) represents the correlation between the i -th and j -th channels (white = 1 or -1 and black = 0). It shows sharp edges that can be used for spectrum segmentation.

segmentation (*ES*) divides the spectrum into three equally-sized contiguous groups. The *maximum energy* partitioning (*ME*) selects the size of each group so that the first eigenvalue, corresponding to the first PC of the group, be maximum. Finally, the spectral-signature-based segmentation (*SSB*) is a supervised approach that uses sample reflectances from a set of known materials. A subgroup is formed for each sample and defined as all the channels in which the sample has a higher reflectance than the others. Not only is this last strategy supervised, it requires that each sample curves be more reflective than the others in a reasonable number of bands, which constrains significantly the choice of sample materials.

Although these studies also emphasize the fact that segmenting the PCA decreases its computational complexity, it still requires the computation of a correlation matrix (Jia *et al.*), a succession of eigen-decompositions (*ME*) or a comparison of spectra (*SSB*). More importantly, these criteria are either naive (equal partitioning) or based on the scene's statistics only and therefore do not properly handle the characteristics of the human visual system. In Chapter 3, we introduce a new approach to spectrum segmentation, based on such characteristics.

2.5 INDEPENDENT COMPONENT ANALYSIS

One major issue with PCA-based approaches is that they are based on second order statistics only. The PCs are indeed by definition orthogonal and therefore uncorrelated, while higher order dependencies may exist within spectral datasets [Wang and Chang, 2006]. Moreover, although ranking the PCs is quite straightforward, there is a question as to how many of them should be retained, particularly for classification tasks.

Independent Component Analysis (ICA) was first introduced by [Jutten and Herault \[1991\]](#) as a method for blind source separation, but it is only a decade later, thanks to the work of [Hyvärinen and Oja \[2000\]](#) that ICA started receiving proper attention in many areas of signal processing, and more specifically in spectral image analysis for tasks such as classification [[Dalla Mura et al., 2011](#)], dimensionality reduction [[Wang and Chang, 2006](#); [Moon and Qi, 2012](#)], spectral unmixing [[Xia et al., 2011](#)] or target detection [[Tiwari et al., 2011](#)]. The goal of this technique is to find a linear representation of nongaussian data so that the components are as independent as possible. The general assumption of ICA is that the image \mathbf{B} is a linear mixture, by an unknown mixing matrix \mathbf{A} , of a set of independent random variables in matrix \mathbf{S} :

$$\mathbf{B} = \mathbf{A}\mathbf{S} \quad (8)$$

Note that, for purposes of conveniency and simplification of the ICA algorithm, a common assumption is that \mathbf{B} is centered and whitened, which may quite as well be done as a pre-processing. Then, the goal of ICA is to recover the unmixing matrix \mathbf{W} so that $\mathbf{S} = \mathbf{W}\mathbf{B}$. To do so, ICA seeks to maximize the independence of the sources, the latter being defined as follows. Two random variables X and Y , with respective probability mass functions (pmf) p_X and p_Y and joint pmf $p_{X,Y}$ are independent if, and only if:

$$p_{X,Y} = p_X p_Y \quad (9)$$

That is, two random variables are independent if their joint pmf is equal to the product of the marginal pmf. Therefore and more importantly, the mathematical expectation $E[.]$ of the product of any linear functions $f_1(X)$ and $f_2(Y)$ is also factorizable:

$$E[f_1(X)f_2(Y)] = E[f_1(X)] E[f_2(Y)] \quad (10)$$

Note that two uncorrelated variables do not necessarily satisfy these properties, which proves the usefulness of a higher-order metric to evaluate independence. Most ICA approaches are motivated by the well-known Central Limit Theorem, which states that the distribution of a sum of independent random variables tends towards a gaussian distribution. Therefore, the mixed data \mathbf{B} is usually more gaussian than each of its independent components \mathbf{S} , of which it is assumed that at most one has a gaussian distribution. Consequently, non-gaussianity can be used as a measure of probabilistic independence. Now, unlike in PCA, there are several different manners to implement such a measure in ICA. The most classical approach to

evaluating non-gaussianity is by means of the adjusted fourth standardized moment called kurtosis⁶, which is defined as follows, for a centered and whitened random variable X :

$$Kurt(X) = E[X^4] - 3 \quad (11)$$

This quantity measures the "peakedness" of the probability distribution of a real-valued random variable and the kurtosis of a gaussian variable is equal to zero.

Another measure of non-gaussianity is the negentropy⁷, which estimates the difference in entropy between a given distribution and the gaussian distribution with the same mean and variance. Indeed, it is known that, out of all distributions with a given variance, the gaussian is the least structured one and therefore has the highest entropy [Cover et al., 1991]. The negentropy J of a random variable X is then defined as:

$$J(X) = H(X_{\text{gauss}}) - H(X) \quad (12)$$

where X_{gauss} is a gaussian variable with same mean and variance than X and

$$H(X) = - \sum_{x \in \mathcal{R}(X)} p_X \log_2(p_X) \quad (13)$$

is the entropy of random variable \mathbf{B} , with b being an event of the \mathbf{B} and $p_B(b)$ is the probability density of \mathbf{B} . However, the contrast function J remains quite expensive computationally-wise as it requires the estimation of pmf. In practice, it can be accurately approximated by the maximum entropy principle [Hyvärinen, 1998].

Another way to approaching the problem of ICA is to minimize the mutual information of the sources. The mutual information MI between two random variables X and Y is given by (in the discrete case):

$$MI(X; Y) = \sum_{X, Y \in \mathcal{R}(X, Y)} p_{X, Y} \log_2 \left(\frac{p_{X, Y}}{p_X p_Y} \right) \quad (14)$$

Yet, this principle also lies under the scope of information theory and therefore leads to results similar to the ones obtained using negentropy. More precisely, it is roughly equivalent to finding 1-D subspaces such that the projections in those subspaces have maximum negentropy [Hyvärinen and Oja, 2000]. Section 3.5 in chapter 3 reviews

⁶ also known as *excess kurtosis*

⁷ also referred to as negative entropy or sintropy.

Shannon's information theory [Shannon and Weaver, 1948] further in details.

Having defined independence and ways to measure it, there exist various algorithms to perform ICA such as Infomax [Bell and Sejnowski, 1995] or FastICA [Hyvärinen and Oja, 1997]. In spectral imaging analysis, the latter is commonly used [Lennon et al., 2001a; Zhu et al., 2007], mostly for its computational simplicity.

One major drawback of ICA however, is that there is a question as to how the ICs should be ranked, in order to perform dimensionality reduction. Unlike PCA, it is now irrelevant to use eigenanalysis as ICA is based on higher order statistics. Zhu et al. [2007], proposed a ranking scheme to alleviate this problem. On the one hand, they designed a prioritization process based on correlation and mutual information between each pair of IC and original spectral channel. On the other hand, they proposed a segmented approach to ICA, in a similar way to the segmented PCA presented in Section 2.4.2.3. Compared to PCA, this approach shows good results with a better naturalness, less pre-attentive features, and a higher entropy. The issue of ranking ICs was also addressed by Wang and Chang [2006] using the concept of *virtual dimensionality* [Chang and Du, 2004], which was originally developed for estimating the number of spectrally distinct signatures in hyperspectral images. Results show once again a considerable improvement compared to second-order statistics-based transforms.

In addition to the fact that ICA is overall significantly slower than PCA, its main drawback is that the assumption of independent sources does not always stand true in spectral imaging [Nascimento and Dias, 2005].

2.6 OTHER LINEAR TRANSFORMATIONS

2.6.1 *Projection Pursuit*

Both PCA and ICA can be considered as special cases of a more general method called Projection Pursuit (PP) [Friedman and Tukey, 1974], which seeks "interesting" directions in the feature space. Most PP algorithms measure this "interest" by the deviation from a gaussian distribution [Huber, 1985], that is, the same motivation than in ICA. The difference lies in the fact that PP is a deflative method, which means that it progressively projects the data after each iteration. Basically, PP characterizes a given projection by a numerical index that indicates the amount of structure that is present. Once an interesting structure has been found, it is removed from the data. The data are then examined for further structure, which, if found, is also removed. This process continues until there is no remaining structure detectable within the data. Applications to spectral image analysis

can be found in [Ifarraguerri and Chang, 2000; Renard and Bourenane, 2009].

2.6.2 Multiresolution analysis

Many approaches to multiresolution-based spectral image analysis have been developed, but very few for unsupervised dimensionality reduction.

Wilson et al. [1997] proposed a contrast sensitivity approach that relies on the human visual system to determine which spectral band is most relevant in a neighborhood of a given pixel, and how the features of this selected band are then incorporated into the fused (greyscale) image.

In [Mignotte, 2010], Mignotte introduced a multiresolution-based technique, referred to as M4ICD (multiresolution Markov model for multidimensional imaging color display), that seeks the three bands \hat{L} , \hat{a} and \hat{b} (not necessarily present in the initial spectral image, this is not a band selection approach), that minimize the overall discrepancy of pairwise Euclidean distances between pixels, in the first, second and last third of the image's spectrum, respectively. To do so, it uses a coarse-to-fine approach, initialized at low resolution by a channel from the spectral image and progressively optimized (preservation of distances) and interpolated (by local color averaging) until reaching the fullest resolution. Figure 11 illustrates the coarse-to-fine scheme.

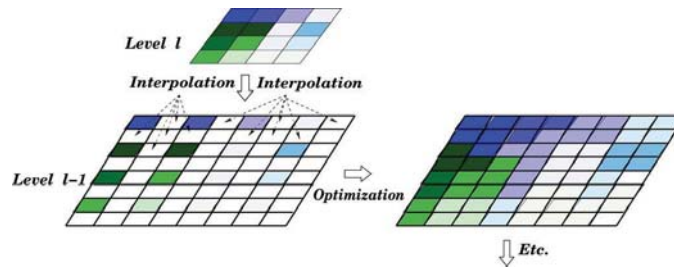


Figure 11: The coarse-to-fine approach to optimizing distance preservation (figure by Mignotte [Mignotte, 2010]).

Eventually, the three components are mapped to CIELAB and converted to a RGB color space for visualization. This approach however has several downsides: not only is it very expensive computationally-wise, it uses the Euclidean distance between pixels spectra as a to-be-preserved criterion, whereas it is known [Parsons et al., 2004] that high-dimensional spaces are sparse, making this measure meaningless. More importantly, it is based on the assumption that the information of luminance and chrominance can be separated in different equally-sized partitions of the image spectrum, which is absolutely counter-intuitive and physically non-consistent.

2.6.3 Conical mapping

Cui et al. [2009] designed a scheme to map high-dimensional pixels to the conical HSV color space, with respect to both the preservation of Euclidean distances between pairs of pixels and the variety of colors in the output composite. The methods features several steps, which are depicted in Figure 12.

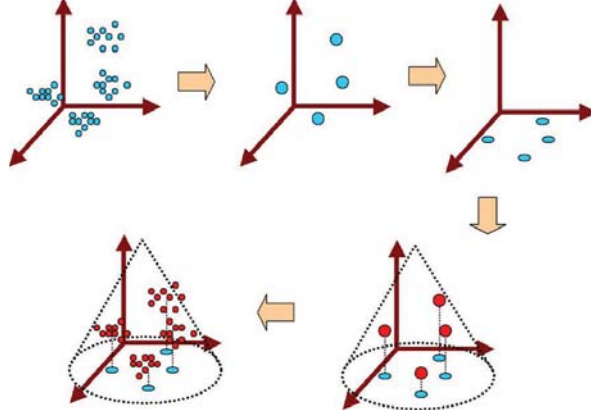


Figure 12: The different steps of the conical mapping approach (figure by Cui et al. [2009]): clustering, projection to Hue-Saturation plane, estimation of luminance and recovering of clustered pixels.

First, pixels are clustered by means of a median cut algorithm (parametrized by M , the number of clusters to be generated) so as to alleviate computational complexity in the next steps, by processing only a few representative points. Then the M cluster centroids are projected to the plane described by their first two PCs, which is assumed to be aligned with the Hue-Saturation subspace. The problem of estimating the Value component (luminance) is then derived as the minimization of the following function:

$$E = \sum_{i=1}^M \sum_{j=i+1}^M \left| D_1(i,j) - \sqrt{(D_N(i,j))^2 - (D_2(i,j))^2} \right|, \quad (15)$$

with $D_1(i,j)$, $D_2(i,j)$ and $D_N(i,j)$ being the Euclidean distance between centroids of classes i and j in terms of luminance, in the Hue-Saturation plane and in the high-dimensional space, respectively. And under the following constrains:

$$\text{if } S(i) \geq S(j), \text{ then } V(i) \geq V(j)$$

$$\text{if } r_i + r_j \leq D_N(i,j), \text{ then } \sqrt{(r_i + r_j)^2 - (D_2(i,j))^2} \leq D_1(i,j)$$

with $S(i)$ and $V(i)$ representing coordinates of class centroid i in HSV and r_i is the average pixel distance to the centroid in cluster i . The

first constraint implies that the data will be contained in a conical hull while the second compels the pixel clusters to be well-separated and therefore aims to occupy as much space as possible in HSV. Clustered pixels are subsequently recovered by interpolation. Eventually, the authors propose a set of interaction tools to focus on particular elements of the image, both spatially and spectrally.

The main advantages of this study is that it tackles the problem of distance preservation in a linear fashion and that it gives an interesting emphasis on interaction. The DR method has however several drawbacks that lower its usefulness in many applications. First, it is sensitive to the initial clustering step. Secondly, it generates an underlying correlation between saturation and luminance. Thirdly, it uses Euclidean distance as a reference for contrast preservation. Finally, it is computationally more complex than most of the other methods from the state of the art.

2.6.4 Pre-defined correlation

Tsagaris and Anastassopoulos [2005] proposed an original methodology to DR for trichromatic visualization by constraining the correlation matrix of the composite, with respect to the statistics of natural scenes in the RGB color space. Their approach is based on the Cholesky factorization that allows to decompose a symmetric positive definite covariance matrix \mathbf{C} by means of an upper triangle matrix \mathbf{Q} such that $\mathbf{C} = \mathbf{Q}^T \mathbf{Q}$. Let \mathbf{Q}_x and \mathbf{Q}_y be such matrices, resulting respectively from the factorization of the input (spectral) and output (trichromatic) images covariance matrices \mathbf{C}_x and \mathbf{C}_y . Then the projection matrix \mathbf{P} can be recovered by the following equation:

$$\mathbf{P} = \mathbf{Q}_x^{-1} \mathbf{Q}_y \quad (16)$$

Note that, if we note $\mathbf{\Sigma}_y$ the output correlation matrix and \mathbf{S}_y the diagonal matrix of the standard deviations of each output's channel, then:

$$\mathbf{C}_y = \mathbf{S}_y \mathbf{\Sigma}_y \mathbf{S}_y^T \quad (17)$$

This latter equation allows us to pre-determine $\mathbf{\Sigma}_y$ so as to impose a certain amount of correlation between the RGB components of the output composite and therefore compute the adequate matrix \mathbf{P} that satisfies this constraint. To do so, the authors averaged the Red/Green, Red/Blue and Green/Blue correlations over a database of 100 color images depicting natural scenes. Finally, the matrix \mathbf{S}_y was built so that the variances of each of the output's channels are equal ($\sigma_R = \sigma_G = \sigma_B$) and their sum is equal to the sum of the variances of the input's channels. Moreover, the authors proposed a

Maximum Energy / Minimum Correlation-based (MEMC) ranking of the original spectral channels in order to convey the largest amount of information possible.

Resulting composites are well-contrasted, show a good natural rendering and a fair preservation of information.

2.7 NON-LINEAR DIMENSIONALITY REDUCTION

Non-linear DR is based on the assumption that the interesting data structures lie on an embedded non-linear manifold (a subset of Euclidean space) within the high-dimensional space. It aims to derive a lower-dimensional coordinate system that better represent such structures.

It has been pointed out by [Bachmann et al. \[2005\]](#) and further demonstrated in [\[Han and Goodenough, 2008\]](#) that spectral datasets contain various sources of non-linearity, such as the properties of the bi-directional reflectance distribution (BRDF), which implies that light is not scattered uniformly from a surface and therefore make the acquired spectra dependent on the scene geometry, or the low spatial resolution of remote sensed images of which pixels actually contain a mixture of "pure" material spectra. Water, particularly, is an important source of non-linearities in such mixtures. Several methods have been proposed for to handle these characteristics for DR of spectral images such as isometric mapping (ISOMAP) [\[Bachmann et al., 2005\]](#), Locally Linear Embedding (LLE) [\[Mohan et al., 2007\]](#), Curvilinear component analysis [\[Lennon et al., 2001b\]](#) or Kernel PCA [\[Fauvel et al., 2006\]](#), which is an alternative to traditional PCA that project the data to a higher dimensional reproducing kernel Hilbert space. Note moreover that data analysis techniques such as detection, classification or quantification can also be thought of as alternative ways to perform non-linear DR as they allow to display, on low dimensional maps, a certain amount of information about the raw spectral image.

Yet, there is no evidence that non-linear approaches are more suitable for DR. A recent extensive comparative review (although not tackling particularly spectral datasets) [\[Van Der Maaten et al., 2007\]](#) shows that, overall, both manners have their own advantages. Particularly, non-linear techniques are usually more time-demanding, since their computational complexity is significantly higher.

2.8 BAND SELECTION

All the previously presented approaches can be referred to as *band transformation* techniques inasmuch as they produce combinations of the original spectral channels to create an enhanced representative triplet. The often mentioned drawback of this kind of approach is the loss of physical meaning attached to a channel. That is, if, ini-

tially, a spectral band is implicitly linked to a narrow range of wavelengths, what can be told about a combination of them? *Band selection* approaches overcome this problem by reducing the dimensionality so that the output (reduced) space is a subset of the input (raw) one, thus preserving the underlying physical meaning of the initial dimensions. This feature is particularly convenient when it comes to interpreting a trichromatic composite while taking full advantage of the fine resolution of the spectral image. In other words, band selection aims to find the best sparse representation of the multidimensional input. We now review the state-of-the-art of such approaches, keeping in mind that only a few of them are designed specifically for visualization purposes.

2.8.1 Fixed band selection

There exist two techniques to perform a selection of spectral channels with respect to their central wavelength only. The most naive one, usually referred to as Uniform Spectral Sampling (USS) is based on the assumption that the information and redundancy are distributed uniformly over the spectral dimension. For example, if the number of channels to be selected is three, USS will go for the first, middle and last one. Needless to say, this is seldom a good strategy. The second "fixed" approach consists of selecting the bands centered in the middle of the red, green and blue ranges (approximately 600, 546 and 436nm) and map them to an RGB color space. Such strategy tends to give good natural rendering but relies on the fact that the corresponding bands contain relevant visual information, which is far from always the case.

2.8.2 Statistical methods

Second-order statistics such as correlation and variance were often investigated for redundancy analysis in spectral images.

Back in 1982, [Chavez et al. \[1982\]](#) developed the so-called Optimum Index Factor (OIF) to select three bands according to their respective energy and correlation. This index is defined as follows:

$$OIF(\mathbf{SI}) = \frac{\sum_{i=1}^3 \sigma_i}{\rho_{1,2} + \rho_{1,3} + \rho_{2,3}} \quad (18)$$

with σ_i representing the variance of band \mathbf{b}_i and $\rho_{i,j}$ is the correlation coefficient between bands \mathbf{b}_i and \mathbf{b}_j . It was recently used for instance in [\[Qaid and Basavarajappa, 2008\]](#) for geological mapping.

The Sheffield index [\[Sheffield, 1985\]](#) is another measure of the information contained by a subset of bands. It is defined as the deter-

minant of the subset's covariance matrix, which turns out indeed to represent the volume spanned by the subset. [Beauchemin and Fung \[2001\]](#) discussed the drawbacks of these two index and proposed a new one, which is based on a minimization of pairwise redundancy only:

$$CI(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \frac{\det(\boldsymbol{\Sigma}_{123})}{\prod_{i=\{1,2,3\}} \sigma_i} \quad (19)$$

where $\boldsymbol{\Sigma}_{123}$ is the covariance matrix of $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$.

[Bajcsy and Groves \[2004\]](#) used the PCA-based ranking, which consists of measuring, for each band, its contribution to the global variance. If λ_j is the eigenvalue corresponding to the j -th PC and v_{ij} is the contribution of band \mathbf{b}_i to the j -th PC (see Section 2.4), then the PCA-rank of \mathbf{b}_i is given as:

$$Rank(\mathbf{b}_i) = \sum_j |\lambda_j v_{ij}| \quad (20)$$

Note that a similar procedure using ICs instead of PCs was used in [\[Du et al., 2003a\]](#). The main inconvenient of this scheme is that it ranks the channels without consideration for inter-band redundancy. Moreover, as previously mentioned, second order statistics are not well suited for spectral image analysis anyway. The use of higher order moments such as skewness and kurtosis was briefly tackled in [\[Du, 2003\]](#) for band prioritization.

2.8.3 Information theory

The most straightforward approach to band selection by information theory is to rank spectral channels according to their entropy [\[Bajcsy and Groves, 2004\]](#). This method is evidently inefficient as it does not consider the possible redundancy of information between channels, which is however known to be usually high. Moreover, without a proper handling of noise prior to measuring entropy, this approach is extremely sensitive to variations of SNR across the spectrum.

[Chang et al. \[1999\]](#) proposed a two-step scheme based on variance and on probabilistic dependence, in order to first rank channels according to their informative content before de-correlating them, thus ensuring a minimal redundancy in the selected subset. Two unsupervised ranking criteria are used, based on variance and SNR, in a similar way to the previously mentioned PCA-ranking. Nevertheless, this ranking is based on intrinsic information only and does not handle n -wise redundancy between channels. Therefore, two channels with high priority (ranking) are both very informative but may be redundant nonetheless. If this redundancy is higher than a certain

threshold, the band with lower priority can then be removed. In order to quantify this redundancy and cope with the problem of band de-correlation, the authors used a symmetric version of the Kullback-Leibler divergence (KLD), that we now define.

Let $KL(X; Y)$ be the KL divergence⁸ of random variables X and Y so that:

$$KL(X; Y) = - \sum_i p_X(i) \log_2 \left(\frac{p_X(i)}{p_Y(i)} \right) \quad (21)$$

Then, the symmetric KL divergence between X and Y can be computed as follows:

$$KL_s(X; Y) = KL(X; Y) + KL(Y; X) \quad (22)$$

This measure has the benefit not to require any second-order pmf estimation, as opposed to mutual information for instance. Nevertheless, the method is sensitive to the divergence threshold under which bands are considered too correlated, which must be adjusted manually. Moreover, the proposed scheme still relies on second order statistics for the band prioritization process, while it is known that higher orders are better suited for spectral image analysis [Wang and Chang, 2006].

Martinez-Uso et al. [2007] proposed a clustering-based band selection algorithm using mutual information (MI) and KLD as measures of redundancy between spectral channels. Band clustering⁹ is but a hierarchical, bottom-up approach to spectrum segmentation (see Section 2.4.2.3). Their method utilizes the so-called Ward's linkage, also referred to as minimum variance clustering, which tends to minimize the intra-cluster redundancy while maximizing the inter-cluster one. Let D be a measure of divergence such as MI or KLD, and C_i , $i = \{1, 2, 3\}$ a set of clusters. Let us consider that $C_1 \cup C_2$ is a candidate to merge into a new cluster. Ward's strategy then defines the "distance" between the candidate couple and C_3 as:

$$D(C_1 \cup C_2; C_3) = a_{13}D(C_1; C_3) + a_{23}D(C_2; C_3) + a_{12}D(C_1; C_2) \quad (23)$$

with

$$a_{ij} = \frac{n_i + n_j}{n_1 + n_2 + n_3} \quad (24)$$

where n_i is the number of bands in C_i . Clusters are chosen so as to maximize D , and the process is repeated until reaching a pre-defined

⁸ also referred to as Kullback-Leibler distance, cross-entropy, relative entropy, information divergence or information gain.

⁹ also referred to as band grouping, subspace identification, spectrum segmentation

number of clusters. Afterwards, one representative band is selected in each cluster, based on its average divergence with the rest of the bands intra-cluster.

More recently, [Cariou et al. \[2010\]](#) proposed a recursive binary band cluster-splitting algorithm based on MI. Given a compact cluster C starting at band b_{\min} and ending at b_{\max} , the method defines the optimal splitting of C at band $b_{\text{opt}} \in [b_{\min}, \dots, b_{\max}]$ so that:

$$b_{\text{opt}} = \operatorname{argmin}_b [MI(\bar{B}_{b_{\min} \rightarrow b}; \bar{B}_{b \rightarrow b_{\max}})] \quad (25)$$

where $\bar{B}_{b_1 \rightarrow b_2}$ is the average of all channels included in the range $\{b_1, \dots, b_2\}$. Once (if) b_{opt} is found, the procedure is applied to the two newly formed clusters and so on, until no more splitting can be done. Eventually, bands in each subgroup are averaged so as to form the reduced set of channels. The inconvenients of this method stem from the unpredictability of its convergence. Either one lets the algorithm converge by itself, in which case, it is not possible to require a specific number of output channels, or one makes the algorithm stop once it reaches a certain level of partitioning and, but then, there is no guarantee that the algorithm does not converge before reaching the desired level.

2.8.4 Linear Prediction

Inspired by algorithms for endmember extraction [[Heinz and Chang, 2001](#); [Ren and Chang, 2003](#)], [Du and Yang \[2008\]](#) proposed an unsupervised band selection scheme based on the error of Linear Prediction (LP) as a measure of independence. The method allows to evaluate the similarity between one and several channels, as explained hereafter.

Let Φ be a matrix of n spectral channels represented column-wise. The LP of a given band \mathbf{b}_{n+1} by Φ is noted \mathbf{b}_{pred} and given by:

$$\mathbf{b}_{\text{pred}} = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}^T \Phi + a_0 \quad (26)$$

with a_i the coefficients that minimize the linear prediction error $e = \|\mathbf{b}_{n+1} - \mathbf{b}_{\text{pred}}\|$. Now, let us note Φ' the concatenation of a column of ones and Φ . Then, the following least square solution allows to find the vector of coefficients $\mathbf{a} = [a_1 \dots a_n]$:

$$\mathbf{a} = (\Phi'^T \Phi')^{-1} \Phi'^T \mathbf{b}_{n+1} \quad (27)$$

The band which maximizes the error of prediction is then considered as the most dissimilar to the ensemble Φ . This process can then be used for any number of bands greater than one in Φ . Note that

this LP-based approach is mathematically equivalent to the Orthogonal Subspace Projection (OSP), also considered in [Du and Yang, 2008], but slightly more expensive computationally-wise.

After pre-processing the data by removing the water absorption and low SNR channels as in [Cai et al., 2007] and by whitening the noise in each channel in order to make the noise component uniform along the spectrum, the LP-based band selection uses a sequential search of optimal bands which progressively adds bands maximizing the dissimilarity with the whole set of already chosen ones. This approach allows to considerably reduce the time requirement of an exhaustive search for pairwise or n -wise optima, and allows to obtain a ranking of the selected components, which is convenient to handle the subsequent mapping to a given color space. The algorithm is initialized by the pair of bands that show the most dissimilarity, which is found by the following procedure:

1. Randomly select a band \mathbf{b}_0
2. Identify \mathbf{b}_1 , the band which is the most dissimilar to \mathbf{b}_0
3. Identify \mathbf{b}_2 , the band which is the most dissimilar to \mathbf{b}_1
4. Proceed until $\mathbf{b}_{i-1} = \mathbf{b}_{i+1}$.

This simple algorithm allows to avoid an exhaustive search while always converging towards the most dissimilar pair of channels \mathbf{b}_i and \mathbf{b}_{i+1} . Note that, even if the selection is performed on noise-whitened channels, the final subset may nevertheless be made out of the original raw ones.

The main downside of LP-based band selection lies in the computational burden engendered by the height of the matrix Φ , which is equal to the number of pixels in the raw image. It may indeed make the computation of the least square solution prohibitive. The authors state however that it is possible to uniformly subsample the image down to 1% of the pixel population to overcome this problem without affecting the result in most cases.

A similar approach tending to minimizing the correlation or dependence of one band with respect to the others was proposed by Chang and Wang [2006]. Two algorithms were presented, based on constrained energy minimization (CEM) and the linearly constrained minimum variance (LCMV) and both are shown to give similar results, the latter being more computationally efficient than the former. Moreover, it was later pointed out [Du et al., 2003b] that, in the case of large SNR ratios, CEM- and LP- based band selection can be considered as equivalent.

2.8.5 One-bit Transform

The One-Bit Transform (1BT) is a tool to analyze the spatial compactness (structure) of a greyscale image, based on a multiband-pass filtering. It generates a binary representation of its input, which makes it a convenient tool for a low-complexity analysis of spectral channels. The 1BT was used by Demir et al. [2009] in a framework that aims to reduce the overall complexity of the band selection process, for efficient hardware implementation.

Let us explain the principle of this transform by first defining a square multiband-pass kernel. Obviously, the size of the latter can be chosen according to the user's need, nevertheless, in Demir's paper, it is fixed to a 17×17 square:

$$\text{kernel}_{1\text{BT}}(i, j) = \begin{cases} 1/25 & \text{if } i, j \in [0, 4, 8, 12, 16] \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

with i and j denoting spatial coordinates. Now let \mathbf{b} be a spectral channel which, once filtered by the multiband-pass kernel, gives \mathbf{b}_F . The 1BT of \mathbf{b} is given by:

$$\mathbf{b}_{1\text{BT}}(i, j) = \begin{cases} 1 & \text{if } \mathbf{b}(i, j) \geq \mathbf{b}_F(i, j) \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

Figure 13 gives an example of a band with its 1BT.

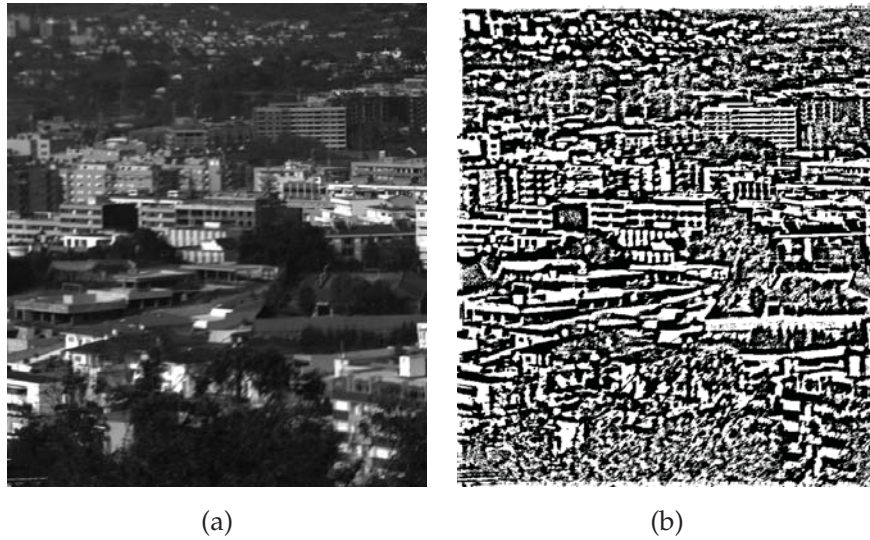


Figure 13: Example of 1BT on a spectral channel. The transform preserves spatial information and emphasizes structure/compactness.

A single-value measure of the informative content of \mathbf{b} , is then obtained by counting the number of transition 0/1 or 1/0 in $\mathbf{b}_{1\text{BT}}$, in

both horizontal and vertical directions. The higher this number is, the less structured is the band, similarly to a measure of entropy. A parametrized local thresholding then allows to make a first coarse band selection by removing the channels whose entropy stands out compared to their close spectral neighborhood. The remaining channels are then compared in terms of correlation, which turns out to be reduced to a simple XOR operation on the binary representations. The two bands with the highest XOR score are first selected as the most dissimilar. The third band is chosen as the one maximizing the sum of scores with the two first ones. Nevertheless, the resulting trichromatic composite show a high correlation between their channels, which implies that this 1BT-based scheme is not well-suited to detect dissimilarities. It gives however a fairly high natural rendering and becomes a very handy technique whenever computational complexity is a critical requirement.

2.9 PREPROCESSING AND NORMALIZATION

2.9.1 *Radiance vs. Reflectance*

By definition, the radiance of an object pertains to its overall appearance under certain viewing conditions while the reflectance is a physical characteristic of its surface, independent of the illumination and of the acquisition sensor. In short-range imaging, converting reflectance to radiance consists simply of a multiplication by a given illuminant such as the ones presented in Figure 14. In remote sensing however, other adjustments are often required to compensate for the effect of the atmosphere [Granahan and Sweet, 2001]. Nevertheless, these radiance/reflectance conversions are mostly linear and have usually no effect on discriminant analysis, as demonstrated in [Hoffbeck and Landgrebe, 1994]. Therefore, we assume that there is very little difference between one and the other for dimensionality reduction. Tyo et al. [2003] argued that radiance data is more intuitive when it comes to visualization, but we assume here that the information to visualize does not necessarily pertain to what can be perceived by our visual system, but rather to the actual scene physical properties. In the light of this assumption, reflectance is a more appropriate feature to work with.

2.9.2 *Low SNR and low entropy*

As already mentioned in Section 2.4.2.2, the SNR is not the same in every raw spectral channel. For instance, some bands may be very dark at the outer limits of the image spectral range, due to the fall-off of sensitivity of the sensor. Moreover, and especially in remote sensing applications, some wavelength ranges are particularly sensi-

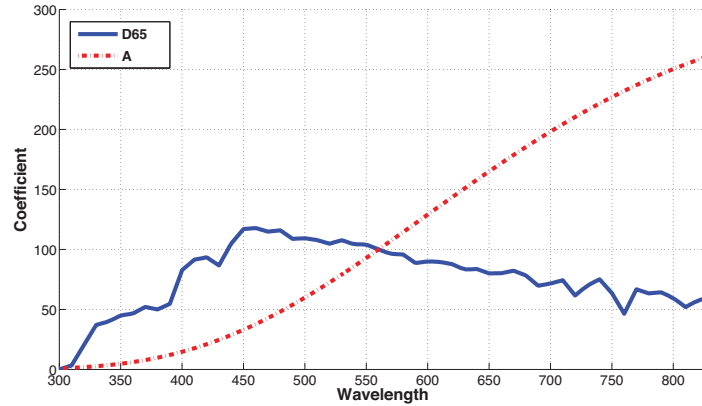


Figure 14: Example of spectral power distributions for two widely used CIE standard illuminants: A and D65. According to the ISO/CIE standard definition, the former is intended to represent typical, domestic, tungsten-filament lighting, while the latter represents average daylight and has a correlated colour temperature of approximately 6500 K.

tive to water absorption¹⁰, which makes them undesirable. In the nIR, there are three noteworthy such regions around, 930, 820, and 730nm. A noise whitening can overcome the first problem and reduce the influence of noise in the DR process, which is for instance the core rationale of the NAPCA. Nevertheless, it requires that the noise covariance matrix be known, which is not a trivial task. [Du et al. \[2003b\]](#) demonstrated that whitening the data instead of just the noise gives very similar improvements, without any knowledge about the noise. Another approach consists of removing those channels whose SNR and/or entropy is low, in order to process only the "good" bands. [Cai et al. \[2007\]](#) assumed that a channel with a low correlation (under 0.8) with its spectral "neighbors" (adjacent bands), has either a poor informative content (low entropy) or a high SNR ratio and therefore can be removed. This unsupervised approach performs indeed well at removing noisy channels, including the ones corresponding to water absorption wavelengths. [Demir et al. \[2009\]](#) employed a similar technique using a measure of local average structure/compactness of the band (see Section 2.8.5) to remove channels that stand out in their spectral neighborhood. In Chapter 3, we introduce an information-theoretic version of the latter technique.

2.9.3 Invariant representation

Since reflectance depends on the scene's geometry, invariant representation of images are used so as to remove the influence of shades

¹⁰ http://en.wikipedia.org/wiki/Electromagnetic_absorption_by_water

and highlights, mostly to enhance segmentation/classification results. Ibrahim et al. [2011] proposed a simple yet effective method to compute an invariant representation of spectral images based on the following processing, applied to all pixels at location (i, j) in band \mathbf{b}_λ :

$$\mathbf{b}'_\lambda(i, j) = \frac{\mathbf{b}_\lambda(i, j) - \min_{n=1 \dots N} (\mathbf{b}_n(i, j))}{\sqrt{\sum_{m=1}^N \left(\mathbf{b}_m(i, j) - \min_{n=1 \dots N} (\mathbf{b}_n(i, j)) \right)^2}} \quad (30)$$

This process gives another meaning to the measure of dissimilarity between two channels. On the one hand, it allows to get rid of some undesirable influences of the scene geometry and make an emphasis on the real physical properties of the materials contained in it. On the other hand, it drastically decreases the meaningfulness of the representation and thus its interpretability, as shown in Figure 15.



Figure 15: Example of true color composite of the (a) raw image, and (b) invariant representation by Ibrahim's method. The invariant composite clearly demonstrates how shadows and highlight have been removed so as to diminish the influence of the scene geometry. Nevertheless, it also decreases the meaningfulness of colors and shapes and makes the subjective interpretation more difficult.

2.9.4 Scaling and clipping

The reflectance of a surface is defined as the ratio of incoming light that is reflected, therefore it must be in the range $[0..1]$, without unit. However, most of the aforementioned projection algorithms tend to distort this gamut as well as the density of the data cloud as shown in Figure 16, which depicts histograms of pixels values in the image before and after PCA. Nevertheless, it is always possible to add a normalization process *a posteriori* so as to enhance the occupation of the color space and thus the "contrast" of the trichromatic composite. The simplest normalization is to uniformly scale the point set so

that it fits into the output gamut. However, the distortion of density induced by the projection might not be affine and introduce outliers. To overcome this problem, it is possible either to clip the outliers to a certain boundary which can be evaluated by subjective assessment of the image’s brightness, or to perform a non-uniform scaling. Moreover, scaling can be done independently in each color channels, although this induces distortions in color distances, particularly when this is done in a perceptually uniform color space such as CIELAB.

Note that band selection does not require to deal with such scaling as it maps channels as-is, which is another advantage of that class of methods.

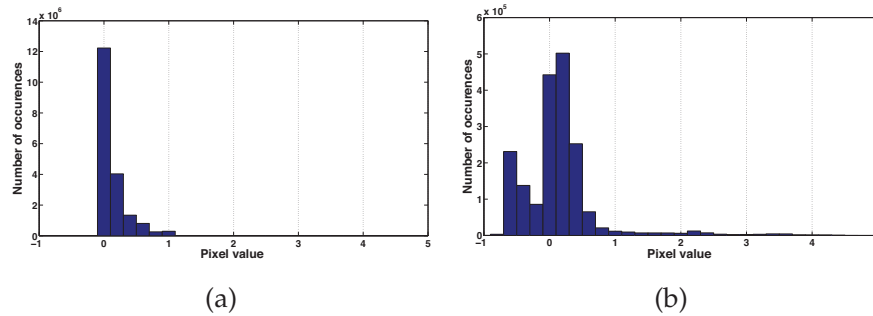


Figure 16: Example of histograms of (a) raw reflectance (b) first 3 principal components. It shows that, although the raw reflectance is bounded in $[0..1]$, the PCA drastically transforms this range, even introducing negative values.

2.10 ON THE EVALUATION OF TRICHROMATIC COMPOSITES

Despite the wide variety of DR methods presented so far, their automatic quality assessment for visualization remains a very challenging and application-dependent task. [Jacobson and Gupta \[2005\]](#) introduced 9 general criteria (so-called *design goals*) to do so: consistent rendering, edge preservation, computational ease, color symbolism, equal energy white point (defined in Section 2.3.1), wavelength shift invariance (all wavelengths are accorded equal weight), smallest effective difference (visual distinctions are no larger than needed to effectively show relative differences), appropriate pre-attentive features and natural palette. Apart from the computational ease (which was used for instance in [[Kaewpijit et al., 2003](#); [Cui et al., 2009](#)]), these design goals can roughly be categorized according to three general criteria: the *conveyed information* from the raw, unreduced image, the *intrinsic information* contained by the composite, and finally the *appeal*, which is related to naturalness, intuitiveness and ease of interpretation. Whereas the two first categories are easily derived into objective and quantitative metrics and measures, the appeal of an image pertains rather to a subjective evaluation, as discussed hereafter.

2.10.1 *Conveyed information*

This first criterion can be equated with the similarity between the high-dimensional image and its corresponding composite. Basically, any criterion used in one projection method can be used to assess the performance of another. For instance, the band-power ratio that was investigated in [Chang et al., 1999], as a criteria for band selection, could very well be used as a performance metric for any other band selection approach. Nevertheless, many metrics and measures have been proposed, especially under the general scope of multimodal image fusion. Liu et al. [2012] pointed out four classes of measures/metrics: information-theory-based [Qu et al., 2002; Tsagaris and Anastassopoulos, 2004; Cvejic et al., 2006; Hossny et al., 2008, 2010], image feature-based [Xydeas and Petrovic, 2000; Zhu et al., 2007; Wang and Liu, 2008], image structural similarity-based [Wang et al., 2004] and human perception-inspired [Chen and Varshney, 2007; Chen and Blum, 2009].

More specifically in the field of spectral image visualization, Du and Yang [2008] also suggested that the result of a classification/clustering in the high-dimensional space can also be used as an estimated "ground truth" reference, in order to evaluate of accurately the discriminative information have been conveyed by the projection. Jacobson and Gupta [2005] used the correlation of spectral angle and hue distances in the (a,b) subspace of CIELAB. The same coefficient was used in [Cui et al., 2009; Mignotte, 2012] with the Euclidean distance instead of spectral angle and in the whole CIELAB space.

2.10.2 *Intrinsic information*

This second criterion pertains to the informative content of the composite. However, in the absence of relevant information about the scene (i.e. unsupervised framework), very few methods have been used under the scope of spectral image visualization. The index presented at the beginning of Section 2.8.2 could serve for evaluation purpose, but are limited to second-order statistics. Zhu et al. [2007] used the marginal entropies of each output channel and one could think also of using the joint entropy. Cui et al. [2009] considered the average perceptual distance between pairs of pixels in the composite as a measure of overall separability of features. Du et al. [2008] used the average inter-class euclidean distance in the (almost) perceptually uniform CIELUV color space. Though the authors used a known ground truth, it is still possible to assume that the result of an unsupervised classification in the high-dimensional space is a good reference for objects identification.

2.10.3 Appeal

Although some of the aforementioned measures and metrics might agree with human visual perception to some extent, most of them cannot predict observer performance for different input images and scenes. It is indeed extremely difficult to design a reliable model of human visual quality assessment [Pedersen and Hardeberg, 2011]. Demir et al. [2009] assumed that the respective correlations of the red, green and blue channels are a good indicator of naturalness, but this idea is rather limited. In [Zhu et al., 2007], the correlation with the reference *true color* (see Section 2.3.1) composite was used, although the use of second order statistics in an RGB color space is not a suitable way to measure perceptual discrepancies. A subjective evaluation was carried out by Cai et al. [2010], on a population of expert subjects, and based on 5 criteria: representation of subtle variations, anomaly pixel detection, estimation of the number of endmembers (objects/materials in the scene), detailed information extraction and overall appearance. Toet et al. [2010] proposed an assisted approach that uses manually (and coarsely) drawn object boundaries, based on the assumption that the overall quality of an image fusion scheme can be characterized by the degree to which the composite represents the objects existing in the unreduced dataset. The DR performance was therefore evaluated according to the accuracy of a certain number of manual segmentations by different users, conveyed into a single reference map, and for three kind of objects: terrain features, living beings and man-made objects. More recently, Lin and Kuo [Lin and Jay Kuo, 2011] conducted an extensive survey on image quality metrics based on perception, according to four main topics: image decomposition, visual feature and artifact detection, just-noticeable distortion modeling, and visual attention map generation. The study shows that a lot of efforts have been directed to better understanding the human visual system. Still, it suggests that most of the state-of-the-art metrics are supervised (so-called Full Reference -FR, as opposed to No-Reference), that is, they require an image of the same dimensionality as the composite as an "original" version of it. Yet, none of the current metrics is able to properly convey subjective notions such as naturalness and intuitiveness in trichromatic imaging.

2.11 CONCLUSIONS

The high dimensionality of spectral images makes them impossible to visualize on tri-chromatic devices such as most of today's computer screens. Reducing the number of channels allows to cope with this problem. In this chapter, we reviewed unsupervised dimensionality reduction for spectral images, with a particular focus on band selection. While traditional approaches such as PCA or ICA seek the best

combinations of spectral channels, band selection techniques aim to find the best subset of them and therefore allow to keep a consistent relation between these channels and the small range of wavelengths which they represent. Several criteria to prioritize bands have been presented, such as second-order statistics or linear prediction, which show considerable downsides compared to information measures such as mutual information. Nevertheless, these measures have not yet been fully exploited for band selection and particularly in terms of three-way interactions. Consequently, a part of the work presented in this thesis aims to further investigate information theory in this context.

Furthermore, we tackled the underlying problems of normalization for gamut mapping and discussed three different criteria to evaluate the quality of a dimensionality reduction for visualization: the conveyed information, the intrinsic information of the composite and its appeal. We determined that there is a growing need for dedicated features to better analyze the information conveyed by DR, in the context of visualization. For these reasons, we present new quality measures (especially one based on naturalness), but more importantly, we introduce the concept of spectral saliency, as a new strategy to measure and compare information in spectral images.

DIMENSIONALITY REDUCTION AND DISPLAY OF SPECTRAL IMAGES: CONTRIBUTIONS

3.1 INTRODUCTION

This chapter introduces the contributions of this thesis made under the scope of dimensionality reduction. They are four-fold: first we present two new measures for quality assessment of tri-chromatic composites, then we present a study on region-based spatially-variant dimensionality reduction. Thirdly, a novel spectral partitioning approach, based on perceptual attributes is explained. Finally, we tackle the topic of information theory-based band selection and present a new constrained method. But first, let us introduce the data that were used throughout these studies.

3.2 BENCHMARKING DATA

In this study, we used a total of 20 spectral images of natural scenes from various sensors:

- Foster's 2002 database¹, used in [Nascimento et al., 2002]. 8 multispectral images containing 31 spectral channels each, covering a part of the visible range of wavelengths (400-700nm).
- Foster's 2004 database², used in [Foster et al., 2004]. 8 multispectral images containing 33 spectral channels each, covering a part of the visible range of wavelengths (400-720nm).
- The Jasper Ridge scene from the AVIRIS sensor³. It contains 224 channels ranging from 400 to 2500nm (sampling rate: 9.3nm).
- The Norway scene from the HySpex sensor⁴. It contains 160 channels ranging from 400 to 1000nm (sampling rate: 3.7nm).

All images contain reflectance data, coded as double precision values (64 bits), and have undergone several pre-processings prior to any of the experiments of this thesis. For white point detection and coarse noise removal, all pixels with value greater than $maxval = m_R + 4\sigma_R$

¹ http://personalpages.manchester.ac.uk/staff/david.foster/...Hyperspectral_images_of_natural_scenes_02.html

² http://personalpages.manchester.ac.uk/staff/david.foster/...Hyperspectral_images_of_natural_scenes_04.html

³ <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>

⁴ <http://www.neo.no/hyspex/>

where m_R and σ_R are respectively the average and standard deviation of the whole raw image, were clipped to this boundary. The result was then divided by *maxval* so that each image has a maximal value of 1, which corresponds to a non-specular white. Moreover, bands with average reflectance values below 2% and those with low correlation (below 0.8) with their neighboring bands were removed, as suggested in [Cai et al., 2007]. Note however that this last step involves solely the "Jasper Ridge" scene, off which we automatically removed the channels at the following indices: 1, 2, 3, 108, 109, 110, 111, 112, 113, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 222, 223, 224.

All experiments were conducted in Matlab 2009b with various tool-boxes. Only the most significant results are presented.

3.3 NEW QUALITY MEASURES FOR DIMENSIONALITY REDUCTION

3.3.1 Naturalness

As discussed in the last chapter, naturalness is a subjective notion of which no computational model have yet been able to grasp the complex meaning. Typically, strongly saturated colors, abnormal hue shifts and/or non-consistent rendering of known objects are interpreted as "non-natural". Evidently, the way humans associate shape with color also implies the role of semantics. It is for example the reason why an image where the sky is blue is more easily interpretable than one where it is green. The *true color* composite of a spectral image is designed so as to be the best match with what one would actually see if one were to be in stead of the camera. Therefore, it can serve as a ground truth for visual information and naturalness.

Zhu et al. [2007] used the correlation between a given composite and its *true color* version in an RGB color space, as a measure of how likely it is that the visualized image is the true color image. This correlation-based measure is however rather difficult to interpret and definitely not suitable to properly evaluate perceptual discrepancies. Let us now consider the Euclidean distance in CIELAB, referred to as ΔE_{ab}^* , and defined as follows, for a pair of pixels (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*) s:

$$\Delta E_{ab}^* = \sqrt{(L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2} \quad (31)$$

As previously mentioned, CIELAB is a perceptually uniform space, which means that ΔE_{ab}^* is proportional to an actual difference of perception by the human visual system. By averaging ΔE_{ab}^* over all pixels of a given scene, and between a composite \mathbf{C} and its corresponding *true color* representation, we obtain what will here be refer to as the natural difference of \mathbf{C} , noted $v(\mathbf{C})$. The smaller the latter, the

more natural the composite. This measure is simple and based on well-established principles, which make it easy to use and interpret.

3.3.2 Color entropy

In order to evaluate the amount of visual information contained in a tri-chromatic image, we define the color entropy H_C as the disorder/unpredictability of perceptible colors. The first step to the computation of H_C is to cluster the pixels in CIELAB so that each cluster represents a single perceptible color. In other words, inside a given cluster C_i , the respective distances between each pair of pixels is under the threshold of Just Noticeable Difference (JND), which means that they are all interpreted as a same color by our visual system. To do so, we propose to use the distance-based K-means classifier. The number of clusters K is defined by a progressive search and the distance to be used is the Eclidean one. In our experiments, we used a starting value of 100 clusters and progressively decreased this value (by steps of 5) until all cluster centroids are separated by at least twice the JND, which we defined to be equal to 3 (see for instance Chapter 2 of [Hardeberg, 2001]), or until an inferior limit of 5 clusters was reached. The resulting class labels are then used to compute the histogram of perceptible colors $\kappa(i)$ (as a discrete estimation of the probability mass function over color clusters $i \in [1 \dots K]$) so as to eventually assess the color entropy H_C :

$$H_C = - \sum_{i=1}^K \kappa(i) \log_2 [\kappa(i)] \quad (32)$$

Seemingly, this process involving several classifications may result in an important computational burden. Nevertheless, the K-means classifier is known to be very efficient and fast on low-dimensional data. During our experiments, we noted execution times under a second on images containing more than 700000 pixels.

3.3.3 Experimental validation

In order to demonstrate the usefulness and reliability of the proposed measures, we show results obtained on 6 composites from two images, one from Foster's 2002 database as well as the "Norway" scene. Moreover, we used two naive metrics used in the literature, for objective comparison. The correlation with the *true color* in RGB [Zhu et al., 2007], $\rho_{TC} = \rho_R \rho_G \rho_B$ as a measure of naturalness, as well as the joint entropy $H(\mathbf{b}_R, \mathbf{b}_G, \mathbf{b}_B)$, which is theoretically better than any combination of the marginal entropies as it takes redundancy into account. Figures 17-18 depict the composites while tables 1-2 give the results obtained.

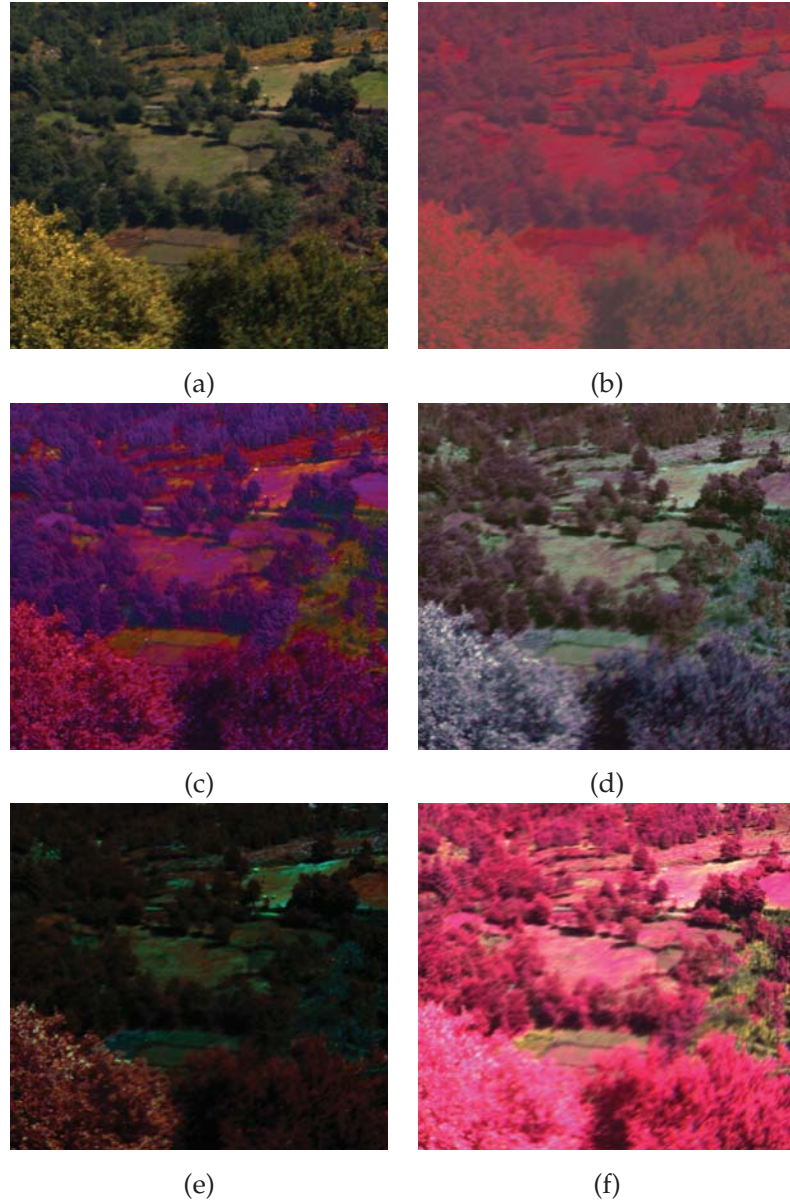


Figure 17: Examples of composites and their respective naturalness according to the proposed measure: scene 4 of Foster's 2002 database. (a) true color - **TC** (b) PCA to sRGB - **PCA_{rgb1}** (c) PCA to sRGB with contrast enhancement - **PCA_{rgb2}** (d) PCA to HSV as in [Tyo et al., 2003] - **PCA_{hsv}** (e) PCA to CIELAB - **PCA_{Lab}** (f) Linear Prediction-based band selection - **LPBS**.

Subjectively, and besides the *true color* composites, we observe that the best natural rendering is produced by the **PCA_{hsv}** technique, mostly because it is the most consistent with the hues of the reference (**TC**). On the other hand, the most "colorful" and somehow less natural is produced by the **PCA_{rgb2}** technique. Now, if we have a look at the

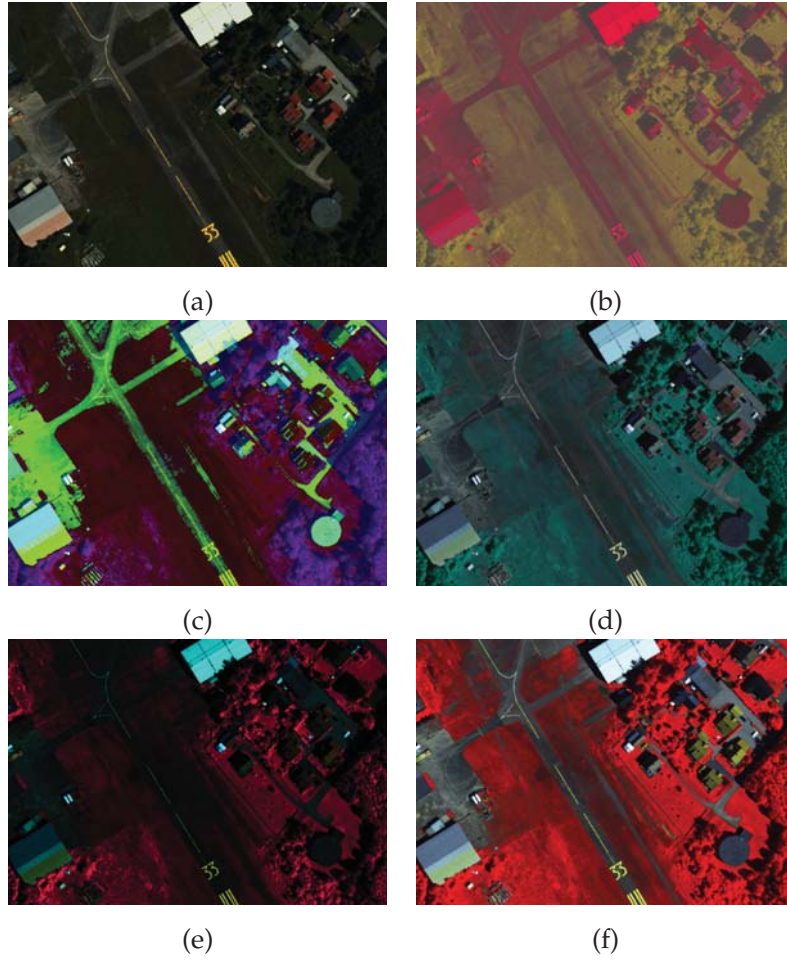


Figure 18: Examples of composites and their respective naturalness according to the proposed measure: scene "Norway". (a) true color - TC (b) PCA to sRGB - PCA_{rgb1} (c) PCA to sRGB with contrast enhancement - PCA_{rgb2} (d) PCA to HSV as in [Tyo et al., 2003] - PCA_{hsv} (e) PCA to CIELAB - PCA_{Lab} (f) Linear Prediction-based band selection - LPBS .

Table 1: Results on the first image

	ρ_{TC}	$H(\mathbf{b}_R, \mathbf{b}_G, \mathbf{b}_B)$	ν	H_C
TC	0.83	16.1	0	4.6
PCA_{rgb1}	0.01	14.9	31.7	2.3
PCA_{rgb2}	0.01	16.1	55.5	6.1
PCA_{hsv}	0.28	16.8	24.3	4.8
PCA_{Lab}	0.23	13.5	33.0	5.4
LPBS	0.24	17.5	47.4	5.2

Table 2: Results on the second image

	ρ_{TC}	$H(\mathbf{b}_R, \mathbf{b}_G, \mathbf{b}_B)$	ν	H_C
TC1	0.82	11.9	0	3.6
PCA_{rgb1}	0.01	12.3	34.2	2.1
PCA_{rgb2}	0.01	11.9	50.8	2.1
PCA_{hsv}	0.46	14.5	22.6	3.0
PCA_{Lab}	0.14	11.6	43.1	4.4
LPBS	0.13	15.6	40.1	4.4

objective results, we note that ρ_{TC} succeeds in identifying the highest natural rendering but fails at discriminating **PCA_{rgb1}** and **PCA_{rgb2}**, although the former performs better at producing naturally-consistent hues and especially contains less of this coarse contrasts and areas of strong saturation which make **PCA_{rgb2}** look very colorful but considerably less intuitive. The proposed metric ν generates results which are much more consistent with the overall appeal of the composites as it agrees with the aforementioned subjective assessment. In terms of entropy now, H_C performs indeed very good at emphasizing colorfulness and variety of colors, unlike the joint entropy which, for instance, gives a high rate to the **LPBS** composite on the first image, while it is clear that this composite is mostly pink and not very "contrasted".

3.3.4 Conclusions

We introduced two new measures for the quality assessment of trichromatic composites: ν and H_C . The first one measures the perceptual difference with a ground truth for naturalness while the second one measures the variety of perceptible colors. Both show results which are consistent with subjective evaluations and outperform more naive approaches that were used in several paper about DR and spectral imaging. Note that we present one more measure, based on saliency, in the next chapter.

3.4 A PERCEPTION-BASED SPECTRUM SEGMENTATION

Most dimensionality reduction (DR) methods compute the optimal projection in the entire high-dimensional feature space, while important objects might exist only in subsets of it [Parsons et al., 2004]. In the previous chapter (Section 2.4.2.3), we introduced several methods [Jia and Richards, 1999; Tsagaris et al., 2005] to partition an image's spectrum in a set of representative subgroups of bands, thus allowing for a local spectral analysis and consequently a lower computational

cost. These methods from the state of the art use criteria which are either naive (equal partitioning) or based only on the scene's statistics (correlation, energy), and therefore do not properly handle the characteristics of the human visual system. In the next Section, we introduce a new spectrum segmentation method which better takes these properties into account.

3.4.1 Definition

The CIE standard observer XYZ Color Matching Functions (CMF) are used to linearly project spectral channels to a tri-stimulus representation corresponding to how a human eye would see the scene. In other words, each spectral channel centered at wavelength $\lambda \in [\lambda_1 \cdots \lambda_N]$ is associated with three positive weighting coefficients $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$, roughly corresponding to its contributions to the perception of the red, green and blue. We propose to interpret this statement as follows: the higher the weighting coefficient $\bar{x}(\lambda)$ the more relevant is the band centered at λ to represent its primary (red for \bar{x} , green for \bar{y} and blue for \bar{z}). Therefore, thresholding the CMF by means of a parameter τ allows to automatically extract the relevant wavelengths to represent a given primary. In addition, segments are compelled not to overlap, that is, one band cannot belong to more than one group, in order to avoid redundancy. As for the choice of these particular CMF, it is motivated by the fact that these are all positive, unlike their RGB version. Note additionally that some problems may rise from using the raw XYZ CMF. First, not all three functions have the same maximal value, which is inconvenient in our framework. There are basically two solutions to cope with this downside: one is to consider the lowest maximum of all three functions as the maximal τ . This would however yield segments of drastically different sizes. Another solution is then to stretch all three functions individually so as the maximum coefficient of each is equal to another. Moreover, this second solution is more consistent with the role of τ , as it turns the CMF into comparable probability functions. A second problem arising from using the raw CMF is that there is a drastic fall-off of the coefficients at the edges of the spectrum, which implies that the channels centered in these wavelengths are not relevant and thus systematically discarded. As a solution to this, we simply compensated these fall-offs by maximizing the blue (resp., red) coefficients at the beginning (resp., end) of the spectrum. If we note the normalized functions with capital letters: $\bar{X}(\lambda)$, $\bar{Y}(\lambda)$ and $\bar{Z}(\lambda)$, respectively, and

the corresponding subgroups Σ_p^τ , $p \in \{X, Y, Z\}$, the following equation summarizes proposed strategy:

$$\mathbf{b}_{\lambda_i} \in \begin{cases} \emptyset & \text{if } m = \tau \\ \Sigma_X^\tau & \text{if } m = \bar{X}(\lambda_i) \\ \Sigma_Y^\tau & \text{if } m = \bar{Y}(\lambda_i) \\ \Sigma_Z^\tau & \text{if } m = \bar{Z}(\lambda_i) \end{cases} \quad (33)$$

with $m = \max[\tau, \bar{X}(\lambda_i), \bar{Y}(\lambda_i), \bar{Z}(\lambda_i)]$ and \mathbf{b}_{λ_i} the band centered at wavelength λ_i . Figure 19 illustrates the role of τ .

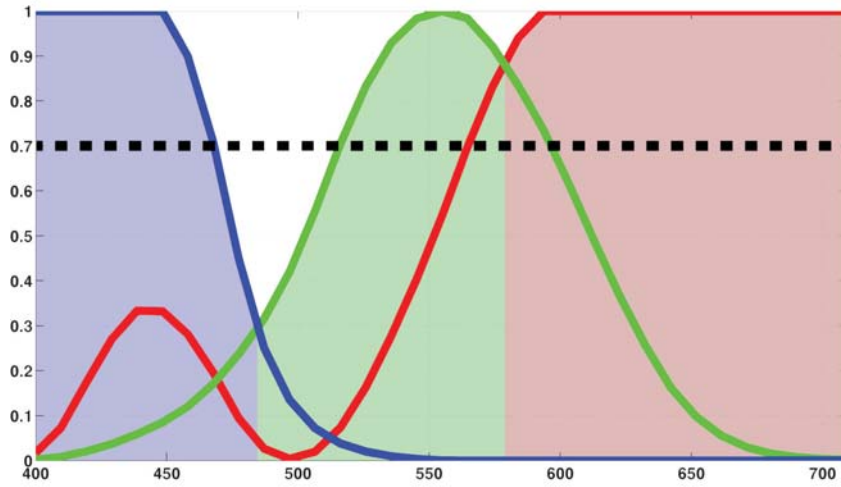


Figure 19: Spectrum segments. The dashed line illustrates an example of thresholding with $\tau = 0.7$, and for which the segments are approximately: $\Sigma_X^{0.7} = 400\text{nm} \rightarrow 470\text{nm}$, $\Sigma_Y^{0.7} = 520\text{nm} \rightarrow 580\text{nm}$, $\Sigma_Z^{0.7} = 580\text{nm} \rightarrow 700\text{nm}$.

If the spectral image contains channels outside the visible wavelengths (400-700nm), the CMF are stretched [Jacobson and Gupta, 2005] so to fit the entire range of the image. Eventually, three segments are obtained: Σ_X^τ , Σ_Y^τ and Σ_Z^τ , the size of which depend on the threshold value:

$$\tau_1 > \tau_2 \rightarrow \Sigma_p^{\tau_2} \in \Sigma_p^{\tau_1}, \forall p \quad (34)$$

The definition of τ makes it a moderation parameter of the following assumption: the higher the weighting coefficient $\bar{x}(\lambda)$ the more relevant is the band centered at λ to represent its primary. If it is set to 0, the assumption is rejected. On the contrary, if $\tau = 1$, the assumption is considered perfectly relevant. Eventually, a representative band from each group may be extracted, concatenated and normalized so as to fit a sRGB gamut and thereby create the trichromatic composite. We will refer to this technique as CMF-based spectrum segmentation.

3.4.2 Experimental setup and results

In this Section, we compare the proposed partitioning approach with the *equal subgroups* and *energy-based* strategies, which were investigated in [Tsagaris et al., 2005]. All three techniques are applied to segmented PCA. The reason why we did not consider the correlation-based approach is that it is very unclear how to segment the correlation matrix, other than manually. Even then, the block structure of the latter seldom allows for a clear segmentation in 3 groups. This experimental setup consists of the 8 images of the Foster's 2002 database and two measures: the naturalness ν , introduced earlier in this chapter and the product of pairwise correlation coefficients of the output channels, noted ρ_{RGB} . The latter is indeed a measure of the redundancy shared by all three channels of the output.

Figure 20 and 21 show the average evolution of both metrics, with respect to the thresholding parameter τ , and over Foster's 2002 database. Note that very similar results were obtained on Foster's 2004 database.

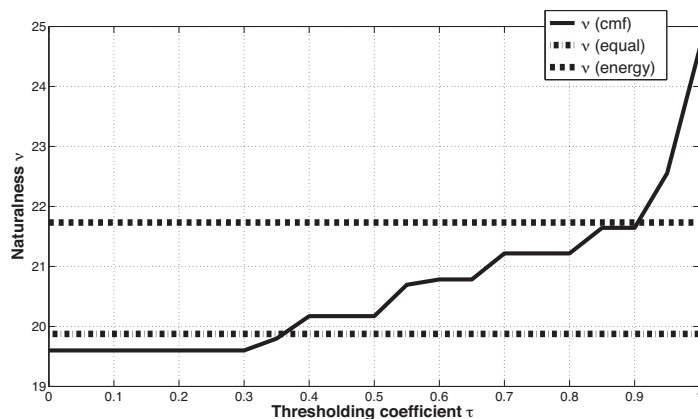


Figure 20: Evolution of naturalness with τ : average values over Foster's 2002 database. In solid line: CMF-based partitioning. In dashed line: equal subgroups-based partitioning. In half-dashed line: energy-based partitioning. When increasing the threshold τ , the resulting composite scores better and better with our measure of naturalness. The best results obtained with our strategy overcome the ones obtained with the benchmark.

Note in Figure 19 that for an increasing $\tau > 0.3$, channels are progressively discarded around 480nm. This threshold value corresponds to the point where the curves in Figure 20 and 21 start "falling". The same remark can be made for $\tau > 0.9$ and around 570nm, which also coincide with a point where the slopes of the aforementioned curves increase. This is due to the fact that less and less channels are taken into account and therefore, a whole range of wavelengths is discarded, which considerably reduces the physical meaning of the transformation and thus naturalness. On the other hand, as the

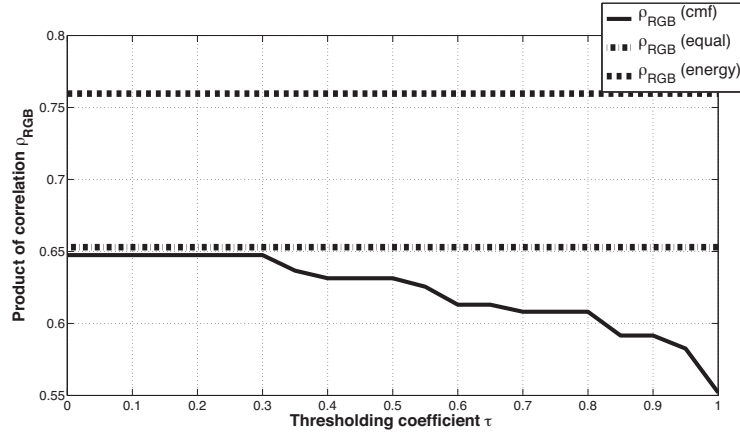


Figure 21: Evolution of channel redundancy with τ : average values over Foster's 2002 database. In solid line: CMF-based partitioning. In dashed line: equal subgroups-based partitioning. In half-dashed line: energy-based partitioning. When increasing the threshold τ , the resulting composite contains less and less correlation between its channels. The best results obtained with our strategy overcome the ones obtained with the benchmark.

band clusters tend to be "separated" by spectral gaps, the probability of redundancy among them is also reduced, hence the increase of ρ_{RGB} . Consequently, we observe that increasing τ simultaneously reduces naturalness and enhances informative content. As for the comparison with the two other segmentation strategies, we note that our approach systematically outperforms the two other in terms of ρ_{RGB} . For $\tau > 0.4$ however, it shows less naturalness than the equal subgroups-based partitioning, but is still better than the energy-based partitioning up to $\tau > 0.9$.

3.4.3 Conclusion

We introduced a new spectrum segmentation strategy which aims to better take into account natural and visual information in DR methods such as PCA or band selection, while alleviating their computational complexity. Results show that proposed strategy performs better than other state-of-the-art approaches and that the threshold value τ can be seen as a parameter to balance between naturalness, informative content and computational efficiency.

3.5 AN INFORMATION-THEORETIC FRAMEWORK FOR BAND SELECTION

As we described in the previous chapter, band selection aims to define the best subset of spectral channels to represent the whole image.

Unlike transformation-based approaches such as PCA or ICA, band selection does not seek for combinations of channels and therefore allows to keep their physical meaning, that is, their underlying association with a small range of wavelengths. We propose a band selection approach, based on information theory, and motivated by two criteria: one pertains to maximizing the information engendered by the subset, and the second criterion consists of minimizing the redundant information between its elements. The spectrum segmentation method presented in the previous Section is used to constrain the selection and ameliorate computational efficiency. This Section is therefore organized as follows: first we recall some elements of information theory such as entropy, mutual information and co-information, then we introduce the different steps of the proposed band selection strategy, before presenting and discussing a variety of results.

3.5.1 Background

Information theory was developed by Claude E. Shannon [Shannon and Weaver, 1948] with the intention of defining a probabilistic framework to measure and quantify information in a set of data. It relies mostly on defining information as disorder or unpredictability. Although, unpredictability can be considered in other domains than time, like in an image for instance, where a group of pixels is informative if it contains unique patterns, which cannot be predicted from the rest of the scene. Information theory is nowadays widely used in fields such as data compression, transmission, or cryptography. Here we give some background information on first, second and third order information measures, in the discrete case. Note that, throughout this Section, we consider a spectral channel \mathbf{b}_λ as a random variable and, whenever necessary, the band's normalized histogram of pixels values serves as an estimation of the probability mass function (pmf) of \mathbf{b}_λ .

3.5.1.1 Entropy

All digital data can be modeled as a series of bits, a unit of information which takes two values, 0 or 1. If the informative content is poor, for instance if all pixels in an image are identical, only a few bits are necessary to properly describe the data without loss. On the other hand, if all pixels are different to each other, a large number of bits will be required not to lose any information. Provided a spectral channel \mathbf{b}_λ , the number of bits necessary to code all of its elements separately is given by the base-2 logarithm of the number of pixels in the band, sometimes referred to as the **information size** of \mathbf{b}_λ :

$$H_0(\mathbf{b}_\lambda) = \log_2 (|\mathbf{b}_\lambda|) \quad (35)$$

with $|\mathbf{b}_\lambda|$ denoting the cardinality of \mathbf{b}_λ . Note that this quantity is not necessarily an integer, and therefore the actual number of bits to code the elements in \mathbf{b}_λ is rather the closest integer greater than $H_0(\mathbf{b}_\lambda)$, but this is irrelevant in this study.

Now that we have quantified the size of information, let us estimate its value. A set of pixels $\mathcal{S} \subset \mathbf{b}_\lambda$ is informative if it cannot properly be predicted by the rest of the pixels in the band, that is, the probability of having \mathcal{S} is poor, considering \mathbf{b}_λ . The **information gain** of \mathcal{S} over \mathbf{b}_λ is then defined as:

$$G(\mathcal{S}) = \log_2 \left(\frac{1}{p_{\mathbf{b}_\lambda}(\mathcal{S})} \right) = -\log_2 (p_{\mathbf{b}_\lambda}(\mathcal{S})) \quad (36)$$

with $p_{\mathbf{b}_\lambda}(\mathcal{S})$ denoting the probability of \mathcal{S} in the spectral channel.

These definitions lead us to the most fundamental element of Shannon's information theory: the **entropy**⁵ associated to a discrete random variables \mathbf{b}_λ , which measures its disorder, uncertainty, or better yet, its unpredictability. It is defined as the sum of information gains of all elements (pixels) of \mathbf{b}_λ , weighted by their own probabilities:

$$H(\mathbf{b}_\lambda) = - \sum_{b \in \mathbf{b}_\lambda} p_{\mathbf{b}_\lambda}(b) \log_2(p_{\mathbf{b}_\lambda}(b)) \quad (37)$$

The entropy of \mathbf{b}_λ is thus a positive value which reaches its maximal value iff all elements in \mathbf{b}_λ have the same probability, i.e., it is possible to gather identical pixels in as many equally-sized groups as there are possible pixel values, which implies that $H(\mathbf{b}_\lambda) \leq \log_2(|\mathbf{b}_\lambda|)$.

It is measured in bits, although it can be alternatively defined with a logarithm in a different base, typically Euler's number e , or 10, in which case its unit is *nat*⁶ or *dit*⁷, respectively.

3.5.1.2 Joint entropy and conditional entropy

Let us now assume that we want to measure the information contained by a set of several random variable $\Phi = [\mathbf{b}_1 \dots \mathbf{b}_k]$. Provided it is possible to compute or estimate their joint pmf p_Φ , the **joint entropy** of the ensemble is:

$$H(\Phi) = - \sum_{\phi \in \Phi} p_\Phi(\phi) \log_2(p_\Phi(\phi)) \quad (38)$$

⁵ also referred to as *self-information*

⁶ *nat* stands for NATural unit ; also referred to as *nit* or *nepit*

⁷ *dit* stands for Decimal digIT ; also referred to as *ban* or *Hart*

It should be noted that this quantity is greater or equal to the maximal marginal entropy from the group and necessarily lower or equal to their sum:

$$\max_i H(\mathbf{b}_i) \leq H(\Phi) \leq \sum_i H(\mathbf{b}_i) \quad (39)$$

It is also possible to estimate the entropy of a variable with respect to another. The entropy of $X \equiv \mathbf{b}_{\lambda_1}$ conditional on $Y \equiv \mathbf{b}_{\lambda_2}$ gives the amount of information provided by the first band, which is not contained by the second one. In other words, it represents the remaining entropy of X given that the value of Y is known. This **conditional entropy** is defined as follows:

$$\begin{aligned} H(X|Y) &= \sum_{y \in \mathcal{R}(Y)} p_Y(y) H(X|Y=y) \\ &= \sum_{x \in \mathcal{R}(X); y \in \mathcal{R}(Y)} p_{X,Y}(x,y) \log_2 \frac{p_X(x)}{p_{X,Y}(x,y)} \end{aligned} \quad (40)$$

Note that this quantity is not symmetric, i.e. $H(X|Y) \neq H(Y|X)$, unless $X = Y$, in which case $H(X|X) = 0$. Moreover, by noticing that $p_{X,Y} = p_X p_{X|Y}$, it can also be written as:

$$H(X|Y) = H(X;Y) - H(Y) \quad (41)$$

Both the joint and conditional entropies can intuitively be represented by means of a diagram (so-called Venn diagram), as depicted in Figure 22.

3.5.1.3 Kullback-Leibler divergence and mutual information

Another way to compare the information contained by random variables is the **Kullback-Leibler divergence** (KLD)⁸, which measures the expected number of extra bits required to code samples from X when using a code based on Y :

$$D(p_X||p_Y) = \sum_{x \in \mathcal{R}(X)} p_X(x) \log_2 \frac{p_X(x)}{p_Y(x)} \quad (42)$$

This value thus quantifies the difference between two probability distributions, over a common support (range of possible values). Should they be identical (i.e. $p_X(x) = p_Y(x)$), their relative entropy is then equal to 0 as $\log(1) = 0$. This provides us with a convenient tool to measure the independence of two variables. Let us recall indeed that X and Y are considered as *independent* iff their joint pmf is equal

⁸ also referred to as *relative entropy*

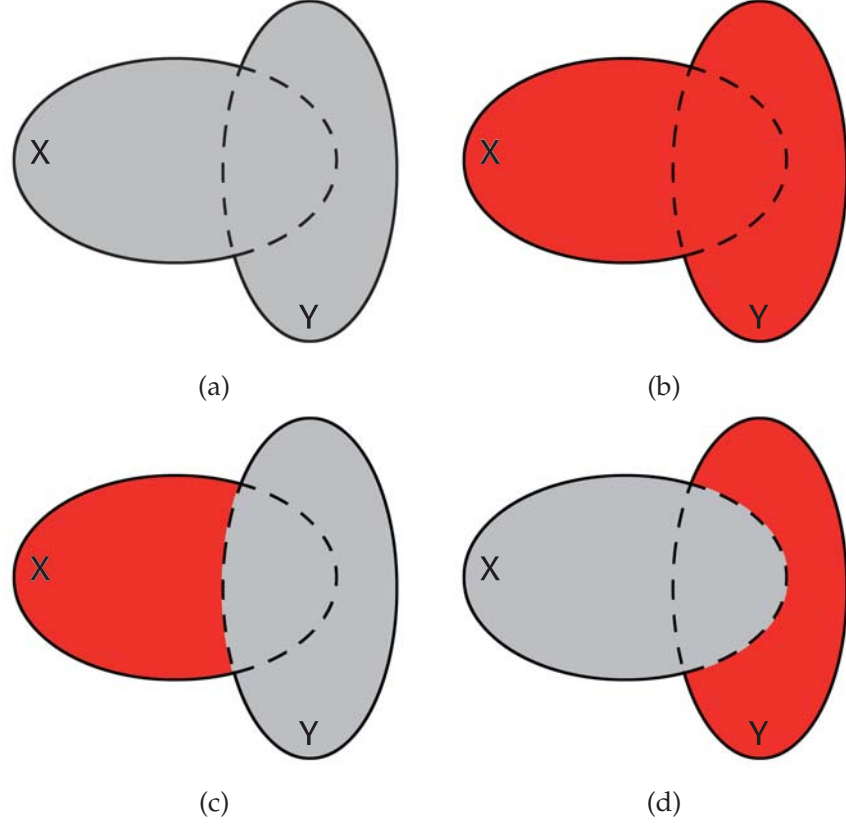


Figure 22: Diagram representation of information measures: (a) each ellipse depicts a marginal entropy, (b) joint entropy $H(X; Y)$, (c) and (d) conditional entropies: $H(X|Y)$ and $H(Y|X)$, respectively.

to the product of their marginal pmf: $p_{X,Y} = p_X p_Y$. Therefore the KLD of $p_{X,Y}$ from $p_X p_Y$, serves as a measure of independence, called **mutual information**:

$$\begin{aligned} I(X; Y) &= D(p_{X,Y} \| p_X p_Y) \\ &= \sum_{x \in \mathcal{R}(X); y \in \mathcal{R}(Y)} p_{X,Y}(x, y) \log_2 \left(\frac{p_{X,Y}(x, y)}{p_X(x) p_Y(y)} \right) \end{aligned} \quad (43)$$

It is nonnegative, symmetric (i.e. $I(X; Y) = I(Y; X)$), and such that $I(X; X) = H(X)$. It is also measured in bits and is equal to zero iff X and Y are independent. It can be expressed by means of entropy terms as:

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X; Y) - H(X|Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X; Y) \end{aligned} \quad (44)$$

Mutual information can as well serve to compare two groups of random variables Φ_1 and Φ_2 of possibly different sizes, following the same definition but using the joint pmf p_{Φ_1}, p_{Φ_2} . For instance, if $\Phi_1 = (X, Z)$ and $\Phi_2 = Y$, then $I(\Phi_1; \Phi_2) = I(X, Z; Y)$ represents the information shared by Y with the dual variable (X, Z) .

Similarly to the conditional entropy, it is possible to define a **conditional mutual information**, measuring the information shared by a couple of variables, given a third one:

$$I(X; Y|Z) = \sum_{x \in \mathcal{R}(X); y \in \mathcal{R}(Y); z \in \mathcal{R}(Z)} p_{X,Y,Z}(x, y, z) \log_2 \left(\frac{p_Z(z) p_{X,Y,Z}(x, y, z)}{p_{X,Z}(x, z) p_{Y,Z}(y, z)} \right) \quad (45)$$

Figure 23 give the diagram representation of mutual information and conditional mutual information.

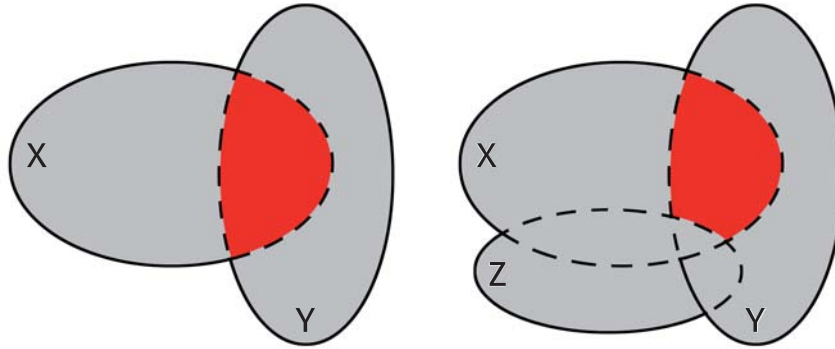


Figure 23: Diagram representation of information measures: (a) mutual information $I(X; Y)$ and (b) conditional mutual information $I(X; Y|Z)$.

Several normalizations of the mutual information are possible to suit various needs. For instance, let us observe that equation 44 implies that:

$$I(X; Y) \leq H(X) + H(Y) \quad (46)$$

which shows that this measure is dependent on the entropy of its inputs. Therefore, two variables with small entropies may be almost identical but will however yield a low value of mutual information. One way to cope with this downside is a normalization which is often referred to as **symmetric uncertainty**, and was used for instance in [Martinez-Uso et al., 2007]. It is defined as:

$$NI_1(X; Y) = \frac{2I(X; Y)}{H(X) + H(Y)} \quad (47)$$

As a number of bits is not always a very intuitive unit to evaluate redundancy, mutual information can also be arranged so as to represent a percentage of shared information. Indeed, we know that $I(X;Y)$ is bounded between zero and $\min[H(X),H(Y)]$, as the highest possible value for $I(X;Y)$ is when X "contains" (information-wise) Y or vice-versa. Thus, we obtain the normalization as suggested for instance in [Kvalseth, 1987], which we will refer to as **true mutual information**:

$$NI_2(X;Y) = \frac{I(X;Y)}{\min[H(X),H(Y)]} \quad (48)$$

Now that we recalled basics of information theory and demonstrated that mutual information is a practical tool to measure the probabilistic dependence of two random variables, what about third-order dependencies? After all, this study being under the scope of trichromacy, there is a real interest as to measure such relations.

3.5.1.4 Multivariate mutual information

Several generalizations of mutual information to higher orders have been proposed in the literature. Watanabe [Watanabe, 1960] introduced the **total correlation**, also known as **multivariate constraint** [Garner, 1962] or **multiinformation** [Studený and Vejnarova, 1998]. It is defined as the difference between the sum of marginal entropies and the joint entropy of the set:

$$TC(X;Y;Z) = \sum_{i \in \{X;Y;Z\}} H(i) - H(X;Y;Z) \quad (49)$$

Total correlation is always positive and equals zero iff $p_{X,Y,Z} = p_X p_Y p_Z$. However, its main drawback lies in the fact that it measures both second and third order indiscriminately, while favoring the third order, as illustrated in Figure 24. We are indeed interested in having a "pure" and therefore more precise measure of third order dependencies.

McGill [McGill, 1954] introduced the **interaction information**, which is defined as:

$$\begin{aligned} A(X;Y;Z) &= I(X;Y|Z) - I(X;Y) \\ &= I(X;Z|Y) - I(X;Z) \\ &= I(Y;Z|X) - I(Y;Z) \end{aligned} \quad (50)$$

and can also be written as a sum of entropies at all orders:

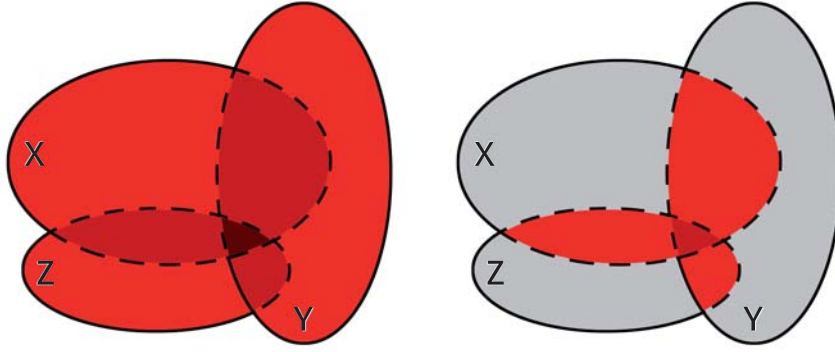


Figure 24: Diagram representation of information measures: (a) sum of marginal entropies $H(X) + H(Y) + H(Z)$, darker areas represent information which are taken twice or three times into account, (b) $TC(X; Y; Z)$. This illustrates why Watanabe's total correlation is actually not an exact third-order dependence measure as it also takes second-order information overlaps into account.

$$\begin{aligned}
 A(X; Y; Z) = & - \sum_{i \in \{X; Y; Z\}} H(i) \\
 & + \sum_{\substack{i \in \{X; Y; Z\} \\ j \in \{X; Y; Z\} \setminus i}} H(i, j) \\
 & - H(X; Y; Z)
 \end{aligned} \tag{51}$$

More recently, Bell [Bell, 2003] re-defined this quantity as the *co-information*, in what appears to us as a more intuitive fashion (see Figure 25):

$$CI(X; Y; Z) = -A(X; Y; Z) \tag{52}$$

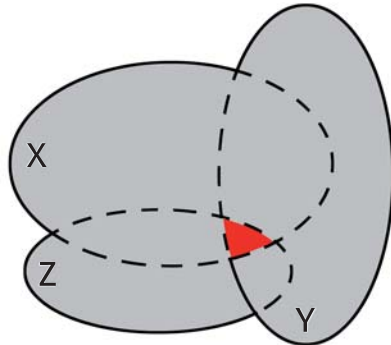


Figure 25: Diagram representation of information measures: co-information

Note that, unlike Watanabe's total correlation, CI equals zero if only a pair of attributes is independent. Moreover, the following relation stands between them:

$$CI(X; Y; Z) = TC(X; Y; Z) - I(X; Y) - I(X; Z|Y) - I(Y; Z|X) \quad (53)$$

A particularly interesting property of co-information is that it can be negative. Although Figure 25 emphasizes what we naturally point out as the third order redundancy area, this diagram representation is not really accurate because of the very existence of the negative case. For the same reason, both McGill and Bell's definitions are correct, nevertheless, we find the latter more suitable and more intuitive for this study.

In the case of positive co-information, we talk about *redundancy*, whereas in the case of negative values, we talk about *synergy*. Redundancies are foreseeable from lower orders while synergies exist only when the random variables are taken together. Co-information is therefore a measure of "hanging togetherness" [Bell, 2003]. If we refer to equation 50, the synergy case exist when, for instance, $I(X; Y|Z) > I(X; Y)$ that is, when the knowledge of Z increases the dependency between X and Y . In order to explain this particular property, a common example is to consider a simple XOR (eXclusive OR) cell with two binary inputs, X and Y and an output $Z = X \oplus Y$. If we consider the inputs as independent, the following stands true: $I(X; Y) = 0$. If we now introduce the knowledge of Z , we also introduce the underlying knowledge of the XOR relation linking the three variables. For instance, if we know that $Z = 0$, we can deduce that $X = Y$, and, by this, we increase the dependency between the inputs so that $I(X; Y|Z) > I(X; Y)$. This whole matter has been further discussed for instance in [Jakulin and Bratko, 2004].

In the case of spectral images, this property also stands true. The knowledge of one channel can increase the mutual information between the two others and, in that case, the smaller the co-information, the higher the shared information. Therefore, in the context of minimizing the dependence of a set of random variables, it is the absolute value of co-information which must be minimized.

3.5.2 Band selection algorithm

In this Section, we introduce the different steps of the proposed strategy, that we will refer to as **IBS** (Information-based Band Selection). It consists of 3 main steps: first, a coarse selection is made based on a comparison of each channel's entropy with the one of its spectral neighborhood. Secondly, the two most dissimilar channels are identified and selected by means of a sequential non-exhaustive search. Thirdly, the last band is chosen so as to minimize the absolute co-

information of the composite. Channels are then concatenated and mapped to sRGB for visualization.

3.5.2.1 First order: coarse selection

We propose to make a first coarse selection allowing for the removal of low informative channels. Spectral images are known to contain high redundancy between neighboring bands and thus, channels that do not meet this definition are considered noisy or poorly informative [Cai et al., 2007]. Either way, these bands can be excluded. Measuring the similarity by means of second-order measures such as correlation or mutual information requires however a high computational burden, since all the pairs of bands must be considered. Another way of coarsely measuring similarity can be done by comparing the intrinsic informative content of a band with that of its neighbors. This can be achieved by means of a moving average as local threshold [Gonzalez et al., 2004; Demir et al., 2009]. We propose to use Shannon's entropy as a measure of informative content. It is computed for each channel, resulting in the solid curve in Figure 26 (for the "Jasper Ridge" image, see description in the next Section).

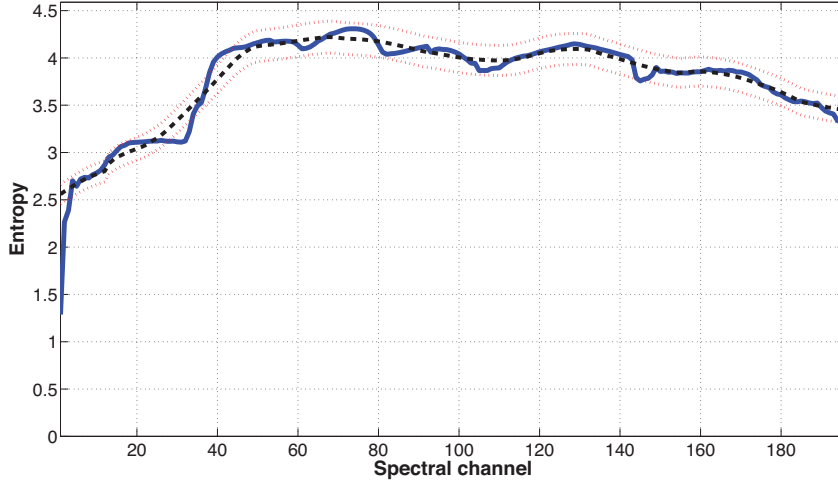


Figure 26: Exclusion of irrelevant bands: entropy, local average and thresholds for the "Jasper Ridge" image. Channels outside the rejection thresholds are excluded ($\delta = 4\%$ and $s = 11$).

The dotted line represents the local average, which is defined for a band \mathbf{b}_i as follows:

$$\bar{H}_s(\mathbf{b}_i) = \frac{1}{s} \sum_{k=-\lfloor s/2 \rfloor}^{\lfloor s/2 \rfloor} H(\mathbf{b}_{i+k}) \quad (54)$$

with s representing the size of the neighborhood. Bands whose entropy is higher (resp. lower) than their local average value moderated

by the thresholding factor δ are then considered as irrelevant. Thus, if a band \mathbf{b}_i reaches the condition in equation 55, it is excluded.

$$H(\mathbf{b}_i) \notin [\bar{H}_s(\mathbf{b}_i) * (100 - \delta); \bar{H}_s(\mathbf{b}_i) * (100 + \delta)] \quad (55)$$

With δ , the thresholding parameter, given in percentage. Both the size of the window and the thresholds have to be set according to the smoothness of the entropy curve. Indeed, the smoother the latter, the lower the probability to have irrelevant bands and, accordingly, the smaller the amount of channels that fall outside the range of relevance. In that case, small value of δ and s are advised in order to gain in precision. On the other hand, a sharp-shaped curve implies strong differences between neighboring channels. In such a case, a large window size is preferred in order to reduce the influence of outliers. This yields a smoother local average curve, and also a less precise analysis. δ allows for adjusting how "strong" the exclusion will be (the lower, the stronger). In the case of the example image "Jasper Ridge", one can notice steep curvature changes, hence our choice to use a large window size ($s = 11$). Moreover, since we first aim at a coarse band selection, we set the threshold to a small value ($\delta = 4\%$).

3.5.2.2 Second and third order

In order to objectively compare the dependences of different pairs of channels, we propose to use the true mutual information, presented in Section 3.5.1. The first two bands to be selected, \mathbf{b}_{λ_1} and \mathbf{b}_{λ_2} , are to be the most independent ones. Though, instead of considering each pair of channels, which would require a considerably high computational cost, we propose to proceed iteratively, by means of an algorithm similar to the one used in [Du and Yang, 2008] (see algorithm 1). The last band \mathbf{b}_{λ_3} is then chosen so as to minimize the absolute value of this quantity:

$$\mathbf{b}_{\lambda_3} = \arg \min_X \frac{|CI(\mathbf{b}_{\lambda_1}; \mathbf{b}_{\lambda_2}; X)|}{H(X)} \quad (56)$$

3.5.2.3 Constraints

Information measures such as NI_2 or CI can be computationally expensive due to the fact that they require the manipulation of 2- and 3-dimensional histograms. Thus, we propose to use the spectrum segmentation approach presented in Section 3.4 in order to restrict and thus alleviate the band selection process. The three channels to be selected are constrained to exist in different segments, so as to reduce the number of possibilities at each step of the algorithm, while imposing a certain distance between channels, so as to take better care of

Algorithm 1 Unconstrained band selection (IBS)

```

i = 0; k = 1
randomly choose j
iterations = 0
while (i != k) and (iterations < 20) do
  i ← j; j ← k
  find k = arg mink [NI2(bj; bk)]
  iterations++
end while
bλ1 ← bj
bλ2 ← bk
find l = arg minl  $\frac{|CI(\mathbf{b}_{\lambda_1}; \mathbf{b}_{\lambda_2}; \mathbf{b}_l)|}{H(\mathbf{b}_l)}$ 
bλ3 ← bl

```

the natural repartition of information along the spectral dimension. Eventually, selected bands are mapped according to their segment of origin (band from the red segment to red channel, etc.). This strategy will be referred to as CIBS (Constrained Information-based Band Selection), and it is detailed in algorithm 2.

Algorithm 2 Constrained band selection (CIBS)

```

i = 0; k = 1
randomly choose j so that bj ∈ ΣXr
iterations = 0
while (i != k) and (iterations < 20) do
  i ← j; j ← k;
  Ξ = argΣ ∈ {ΣXr; ΣYr; ΣZr} bj ∈ Σ
  find k = arg mink [NI2(bj; bk)] with bk ∉ Ξ
  iterations++;
end while
Ξ = argΣ ∈ {ΣXr; ΣYr; ΣZr} (bj ∉ Σ, bk ∉ Σ)
find l = arg minl  $\frac{|CI(\mathbf{b}_{\lambda_1}; \mathbf{b}_{\lambda_2}; \mathbf{b}_l)|}{H(\mathbf{b}_l)}$  with bl ∈ Ξ
bλ1 ← argb ∈ {bj; bk; bl} b ∈ ΣXr
bλ2 ← argb ∈ {bj; bk; bl} b ∈ ΣYr
bλ3 ← argb ∈ {bj; bk; bl} b ∈ ΣZr
map (bλ1, bλ2, bλ3) to (R,G,B)

```

3.5.2.4 Computational considerations

In order to compute the entropy of a random variable, it is necessary to estimate its pmf, which is usually done by histogram analysis. That is, the probability that a pixel has a certain value, or better yet, falls in a certain range of values, is estimated by the proportion of pixels in the same range, in the band. The size of this range is determined by the number of bins, β , used for histogram computation. Let us assume that we manipulate 8-bits long pixels, that is a total of 256 possible values, then a 256-bins histogram would give the best possible resolution for the pmf estimation. Reducing the number of bins β would considerably reduce the computational cost of our approach, as shown in Figure 27, but would also lower the precision of calculations, as shown in Figure 28. Nevertheless, if the entropy of each channel is equivalently modified, it is still possible to accurately compare them to each other. In order to evaluate how entropies are relatively altered by changing β , we computed, on a series of 100 random vectors, entropies at first, second and third order and for a number of bins ranging from 2 to 256. Thus, we computed the correlation of the 100 entropy values obtained for each value of β with the ones obtained with the highest β , and at each order. Results are shown in Figure 29 and indicate a clearly decreasing trend, particularly abrupt for the third order. Consequently, it is inaccurate to reduce β under 256 bins in the present framework.

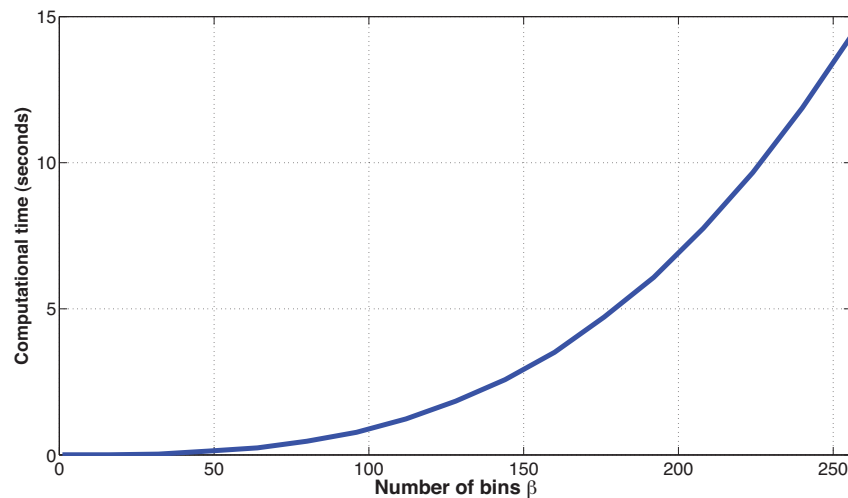


Figure 27: Computational time for third order entropy against number of bins, on Matlab 2009b.

Nevertheless, a simple computational trick can indeed ameliorate the computational cost of our method. Along the steps of the algorithm, many pmf are to be estimated by histogram analysis of single, pairs of and triplets of variables, some of them are required several times and thus can be computed only once and stored until needed

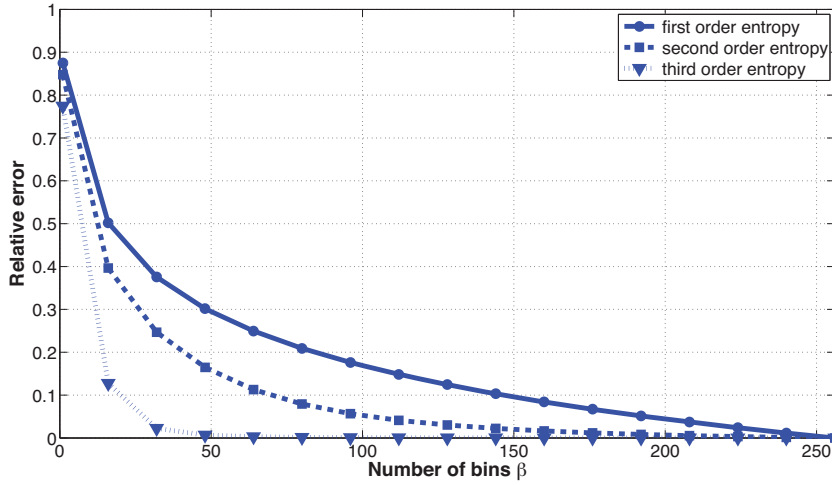


Figure 28: Error of entropy against number of bins.

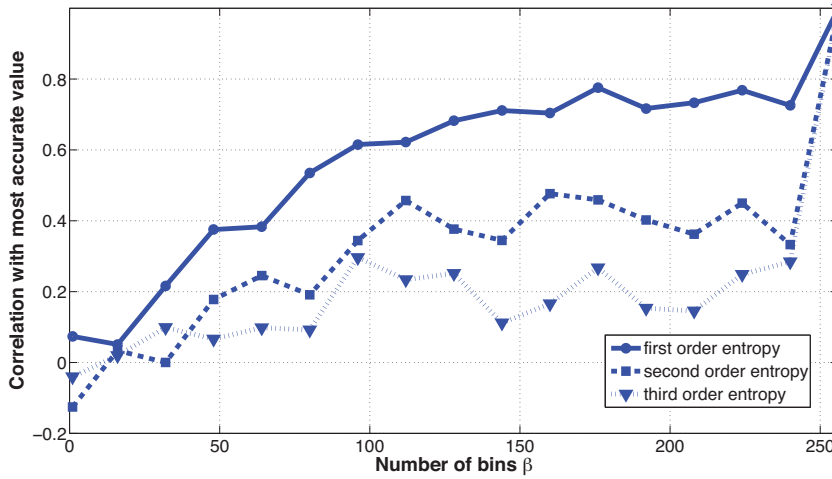


Figure 29: Correlation of entropy against number of bins.

again. For example, the marginal entropy of a band can be required in up to $N - 1$ operations. Computing it only once allows to drastically shrink the time consumption of our method.

3.5.3 Experimental setup and results

The proposed experimental setup aims to extensively evaluate the performance of the proposed method, compared to other state-of-the-art techniques, and based on two criteria: appeal and intrinsic information. The first criterion was measured by means of the aforementioned v while information was evaluated in terms of color entropy and Inter-Class Perceptual Distance (*ICPD*). The latter was computed based on a K-means clustering ($K=5$) on an invariant representation of the spectral image [Ibrahim et al., 2011], and by computing the

average Euclidean distance between clusters, in CIELAB. We remind the reader that we seek for a minimal ν , but a maximal H_C and $ICPD$. For comparison, we used the following methods: *true color* (**TC**), PCA to CIELAB (**PCA_{LAB}**), linear prediction-based band selection (**LPBS**) and One-bit-transform-based band selection (**1BTBS**). Moreover, we also used a naive information-based method, which first prioritizes channels by descending entropy and then de-correlates them by removing the ones with a true mutual information greater than 0.5 (inspired by the strategy presented in [Chang et al., 1999]). This last method will be referred to as *H-prioritization I-decorrelation* **HPID**. Finally, we divided the setup in two Sections: one aims to compare the unconstrained case (**IBS**) to measure the efficiency of the information criteria alone, while the second part tackles the constrained case (**CIBS**).

3.5.3.1 Unconstrained IBS

Tables 3 to 6 give the results obtained for each type of data, and Figures 30 to 33 show examples of composites. As expected, it appears that the PCA-based approach gives very contrasted renderings, with particularly saturated colors and important brightness variations. On the other hand, the band selection techniques seem not to introduce such distortions, thus better conveying the spatial characteristics of the scene, even though their overall rendering is not necessarily appealing. In an attempt to finding an objective definition of naturalness, we observe that **HPID** and **IBS** obtain the best results in terms of ν , which suggests that information-based measures seem to better convey visual information than other methods based on second-order statistics, orthogonality or structure/compactness. During our experiments, we observed particularly that the three channels selected by **IBS** are usually well-spaced along the spectral dimension, which implies that different ranges of wavelengths are represented in the composite, thus making it more physically consistent. Now, in terms of informative content, **IBS** gives very good results, not always better than **PCA_{LAB}**, but still very competitive, and globally better than the other band selection techniques. Only the Norway scene seems less suited for our algorithm as it is outperformed by the **LPBS**, in terms of both color entropy and class-separability, still **IBS** gives very good class-separability compared to the **1BTBS**. We note that this scene is both the largest and most structured one, which is most likely the reason why the **1BTBS**, as well as our approach, tend to extract fine discriminative information such as the texture of the grass and trees, rather than coarse one such as between rooftops and roads.

Table 3: Average on Foster’s 2002 database. The proposed approach (last line) obtains the best naturalness and color entropy. On the other hand, it is the PCA-based strategy that outperforms the others in terms of perceptual distance between classes, followed closely by **IBS**.

	ν	H_C	$ICPD$
TC	/	4.8	11
PCA_{LAB}	40	5.2	17
LPBS	27	5.3	13
1BTBS	30	5.5	11
HPID	25	4.1	10
IBS	23	6.2	16

Table 4: Average on Foster’s 2004 database. Same comments as for Table 3.

	ν	H_C	$ICPD$
TC	/	3.4	10
PCA_{LAB}	30	3.9	19
LPBS	21	4.5	11
1BTBS	24	4.5	8
HPID	19	3.6	10
IBS	15	5.1	13

Table 5: Jasper Ridge scene. Although our approach does not provide the best results on this particular image, it gives a fairly good balance between naturalness and informative content.

	ν	H_C	$ICPD$
TC	/	4	3
PCA_{LAB}	58	4.9	44
LPBS	81	4.4	26
1BTBS	81	4.7	23
HPID	48	4.3	33
IBS	50	4.3	41

Table 6: Norway scene. Again, **IBS** is outperformed in terms of each measure, but gives an overall good compromise.

	ν	H_C	$ICPD$
TC	/	3.6	17
PCA_{LAB}	43	4.7	18
LPBS	40	4.8	15
1BTBS	38	3.8	10
HPID	29	3.7	11
IBS	30	4.1	13

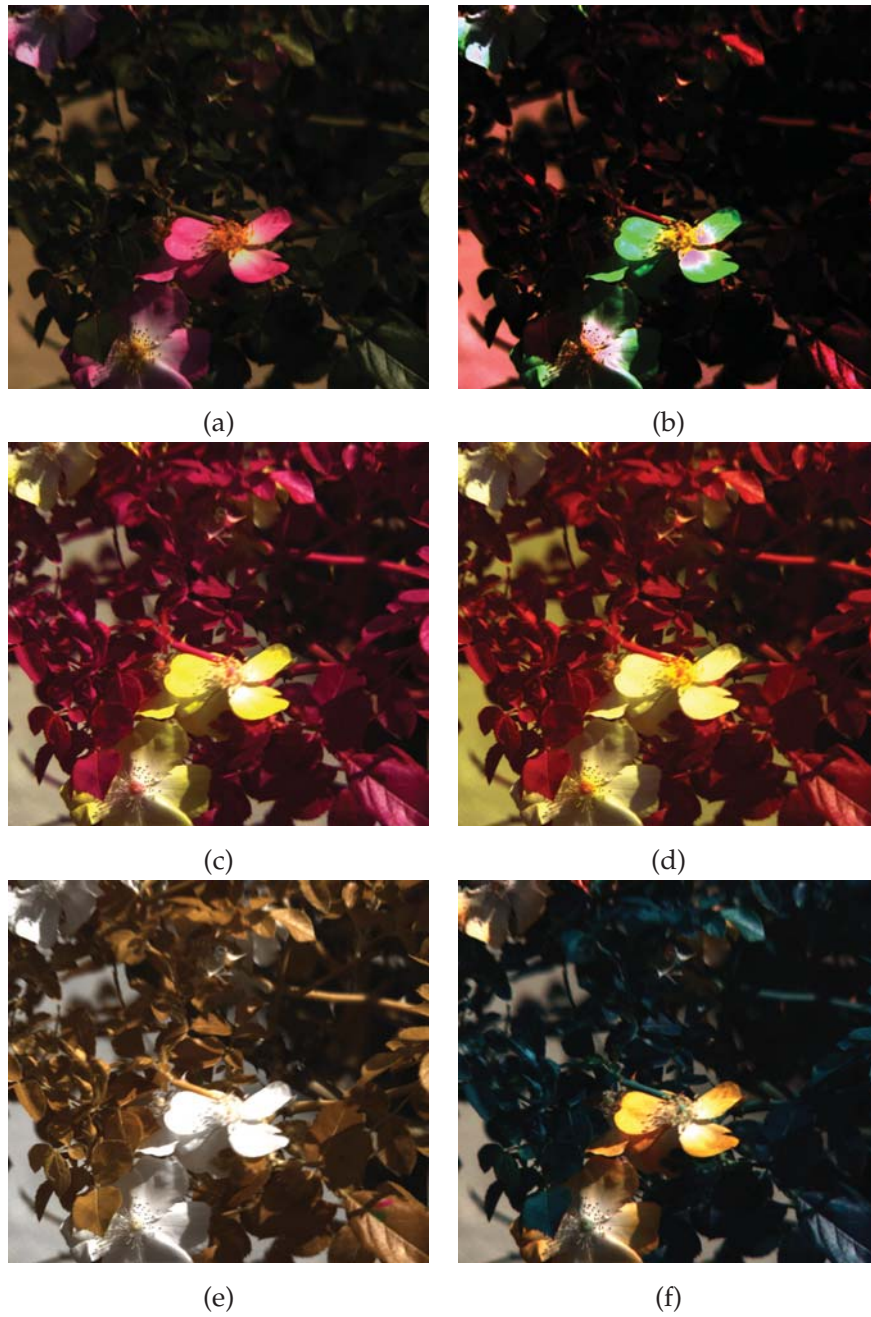


Figure 30: Examples of composites on scene 3 from Foster's 2002 database. (a) **TC** (b) **PCA_{LAB}** (c) **LPBS** (d) **1BTBS** (e) **HPID** (f) **IBS**. Aside from **TC**, our approach is the only one giving a consistent color association as it depicts leaves in greenish hues.

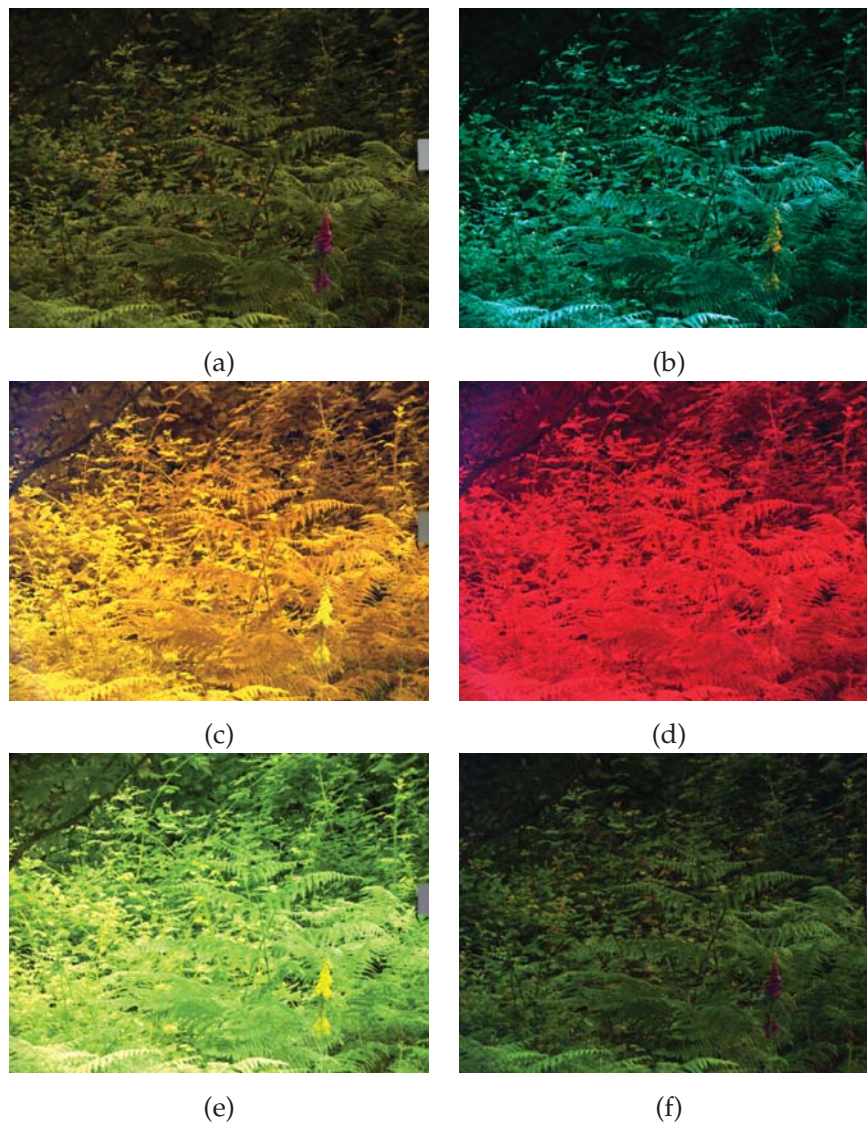


Figure 31: Examples of composites on scene 2 from Foster's 2004 database. (a) **TC** (b) **PCA_{LAB}** (c) **LPBS** (d) **1BTBS** (e) **HPID** (f) **IBS**. Again, the latter gives a consistent and natural color association, with a result very close to the *true color* composite.

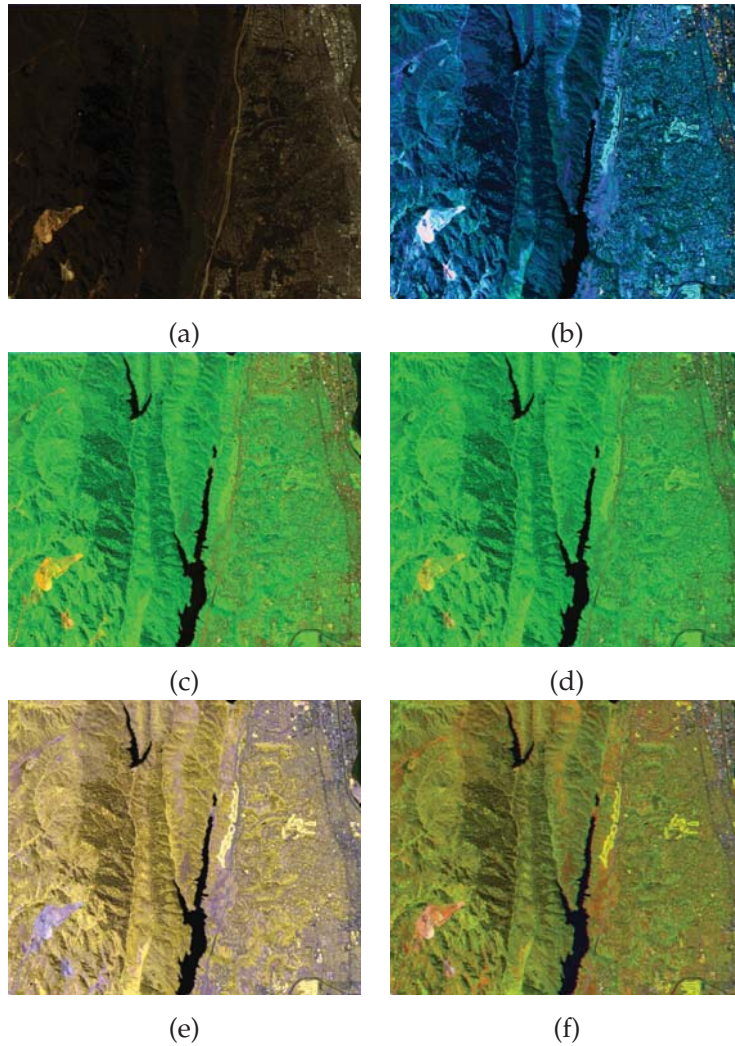


Figure 32: Examples of composites on the Jasper Ridge scene. (a) **TC** (b) **PCA_{LAB}** (c) **LPBS** (d) **1BTBS** (e) **HPID** (f) **IBS**. The PCA-based approach clearly shows the largest amount of information but also exaggerates contrasts and yields non-intuitive color associations. On the other hand, the proposed approach gives a better compromise between showing the important features of the scene and respecting a natural palette of colors.

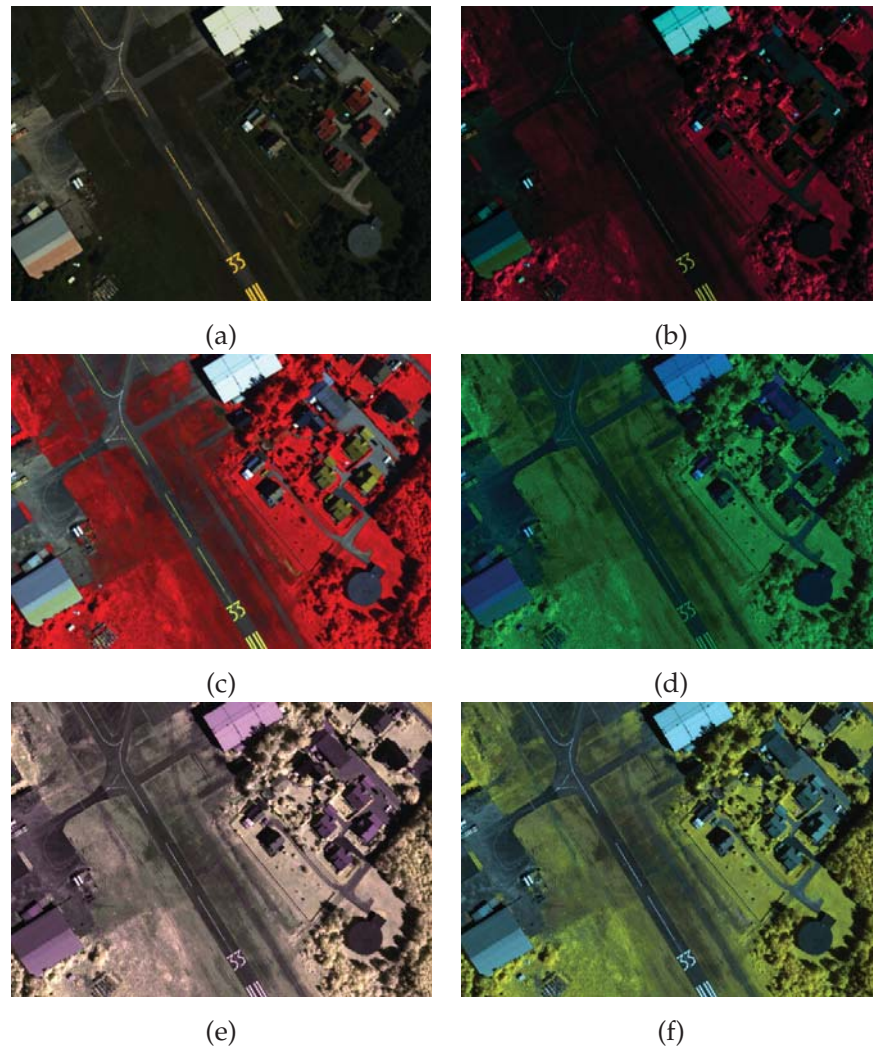


Figure 33: Examples of composites on the Norway scene. (a) **TC** (b) **PCA_{LAB}** (c) **LPBS** (d) **1BTBS** (e) **HPID** (f) **IBS**. Same remarks as for the Jasper Ridge scene.

3.5.3.2 *Constrained IBS*

Tables 7 to 10 give the results for each type of data, and Figures 34 to 35 show examples of obtained composites. For comparison, we used the segmented PCA approach presented in Section 3.4, which we refer to as $\mathbf{sPCA}_{\text{cmf}}$. The reason why this PCA-based strategy looks a little bit washed-out is that PCA creates important distortions of gamut and therefore stretches the gap between the brightest pixel and the darkest, thus increasing the overall average value. Note that we did not use any specific normalization other than a gentle clipping to ten times the overall standard deviation. These results show an improvement of naturalness to even better results than **IBS**. We note that increasing the value of τ augment the informative content of the composite, in accordance with the results obtained in Section 3.4. Besides, the green color plays a crucial role in natural scenes, especially for plants. In the constrained case of our approach, and for $\tau = 1$, the green channel is always selected as the wavelength to which the human perception of green is the most sensitive, which results in very appealing composites.

Table 7: Average on Foster’s 2002 database. In accordance with the results presented in Section 3.4, increasing the threshold τ provides a better informative content but a lower naturalness. Still, the proposed strategies perform better than the segmented PCA.

	ν	H_C	$ICPD$
$\mathbf{sPCA}_{\text{cmf}}$	19	4	7
$\mathbf{CIBS}_{\tau < 0.3}$	17	5.4	7
$\mathbf{CIBS}_{\tau = 1}$	18	5.6	8

Table 8: Average on Foster’s 2004 database.

	ν	H_C	$ICPD$
$\mathbf{sPCA}_{\text{cmf}}$	21	3.6	11
$\mathbf{CIBS}_{\tau < 0.3}$	20	5.6	15
$\mathbf{CIBS}_{\tau = 1}$	21	5.7	19

Table 9: Jasper Ridge scene. Only H_C gives significant results on this scene as such small differences of ν and $ICPD$ are not perceptible.

	ν	H_C	$ICPD$
sPCA_{cmf}	3	0.19	0.25
CIBS_{$\tau < 0.3$}	3	0.26	0.63
CIBS_{$\tau = 1$}	2	0.26	1.6

Table 10: Norway scene. Same remark as for the Jasper Ridge scene.

	ν	H_C	$ICPD$
sPCA_{cmf}	1.65	0.22	0.61
CIBS_{$\tau < 0.3$}	1.48	0.28	0.99
CIBS_{$\tau = 1$}	1.44	0.28	0.83

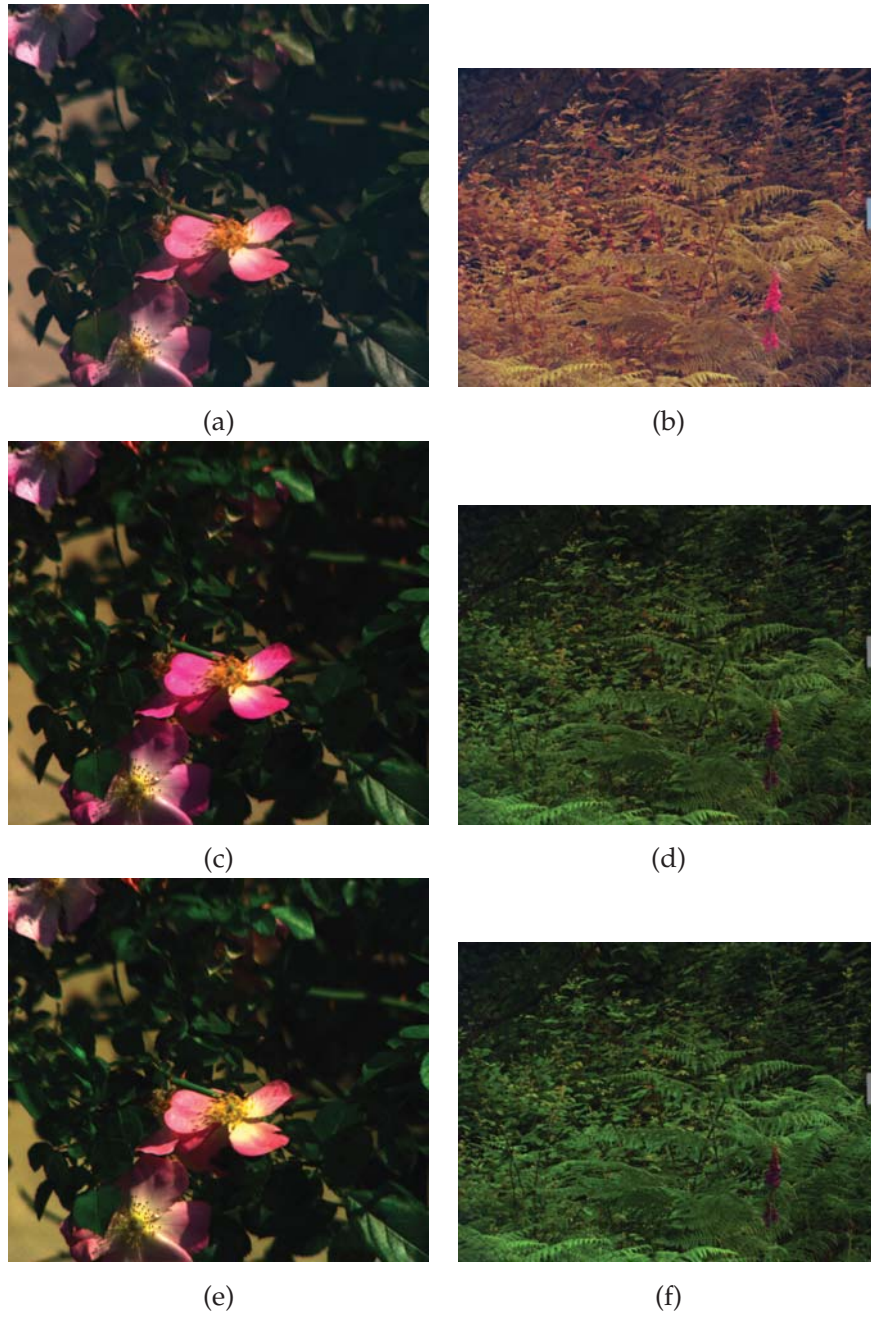


Figure 34: Examples of composites on scene 1 from Foster's 2002 database (left) and scene 2 from Foster's 2004 database (right). First row: $sPCA_{CMF}$, second row: $CIBS_{\tau < 0.3}$, and third row: $CIBS_{\tau = 1}$.

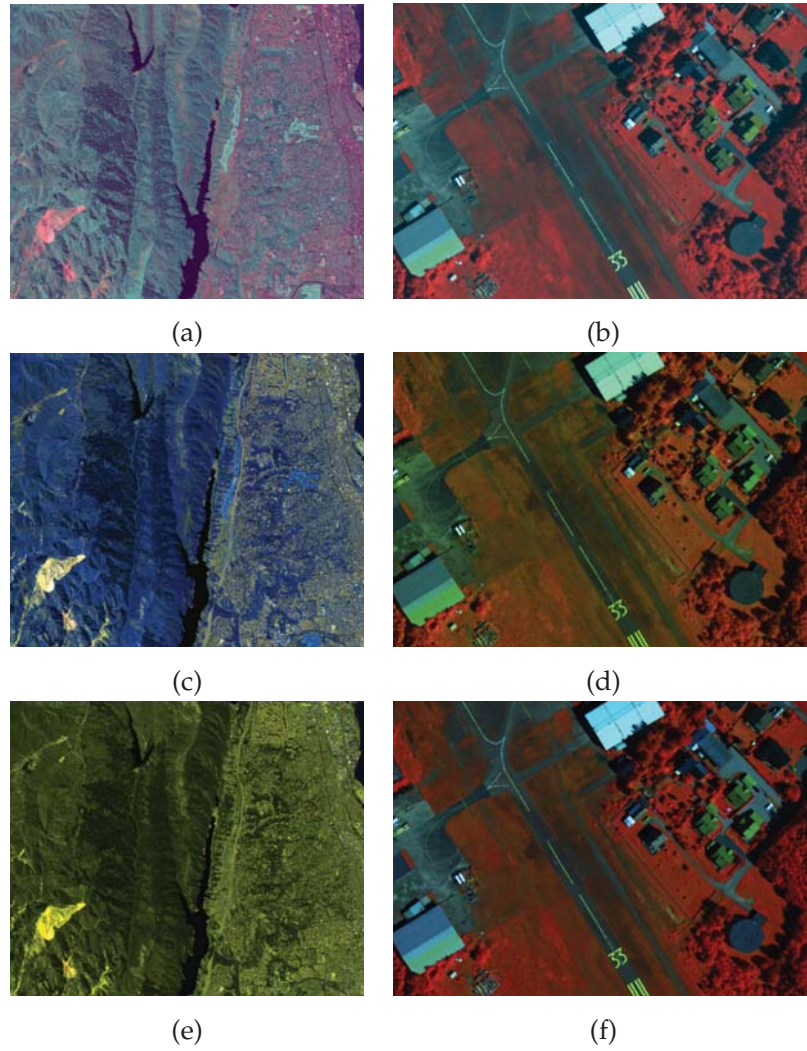


Figure 35: Examples of composites on the Jasper Ridge scene (left) and Norway scene (right). First row: $sPCA_{CMF}$, second row: $CIBS_{\tau < 0.3}$, and third row: $CIBS_{\tau = 1}$.

3.5.4 Conclusions

We developed new information-based band selection strategies called **IBS** and **CIBS** which use measures built upon Shannon's entropy. It is a somehow evolved version of the method presented in our first journal publication [Le Moan et al., 2011b]. It has been established that it is a difficult task to find the reasons why one DR method gives an overall better naturalness than another, without a proper definition of such subjective notion. Although no obvious relation has been found between naturalness and any of the other measures used, these results show that both **IBS** and **CIBS** perform very well at identifying visual information and gives a good compromise between naturalness and informative content.

3.6 CLUSTER-ADAPTIVE PROJECTIONS

3.6.1 Statement of the problem

As explained in Chapter 2, most linear methods of DR for spectral images consider a spectral channel as a whole and therefore perform in a global manner. That is, the computation of the optimal low-dimensional projection is based on the information contained by all pixels together. Nevertheless, as it is the case in the spectral domain, the contribution of some components (pixels) may be undesirable and bias favorably the projection toward the separability of certain classes of pixels, creating distortion of distances between such classes. In the case of data-adaptive methods such as PCA, this distortion is unpredictable as each dataset can render a different set of PCs, as opposed to fixed bases such as the CMF. For that reason, this effect prevents particularly from being able to preserve and/or emphasize the local contrast between specific groups of pixels, a problem which has been addressed for instance in [Cui et al., 2009], where the preservation of pairwise distances between pixels is considered as a criterion to optimize. Nevertheless, some materials and/or objects may require to be easily differentiable on the final composite, be it because they have different spectral characteristics (emphasis on some data properties) or because it is a requirement for a specific task (manual adjustment of the distortion). Let us assume for instance that we have three classes of pixels C_1 , C_2 and C_3 , and that a given DR method (say PCA), stresses the contrast between C_1 and C_3 , while rendering C_2 and C_3 in a same color, because the projection makes these groups overlap in the color space. Now suppose that our application requires however that classes C_2 and C_3 are clearly separable on the composite. Figure 36 shows that using a local PC, computed by means of C_2 and C_3 only, is a better strategy in that case.

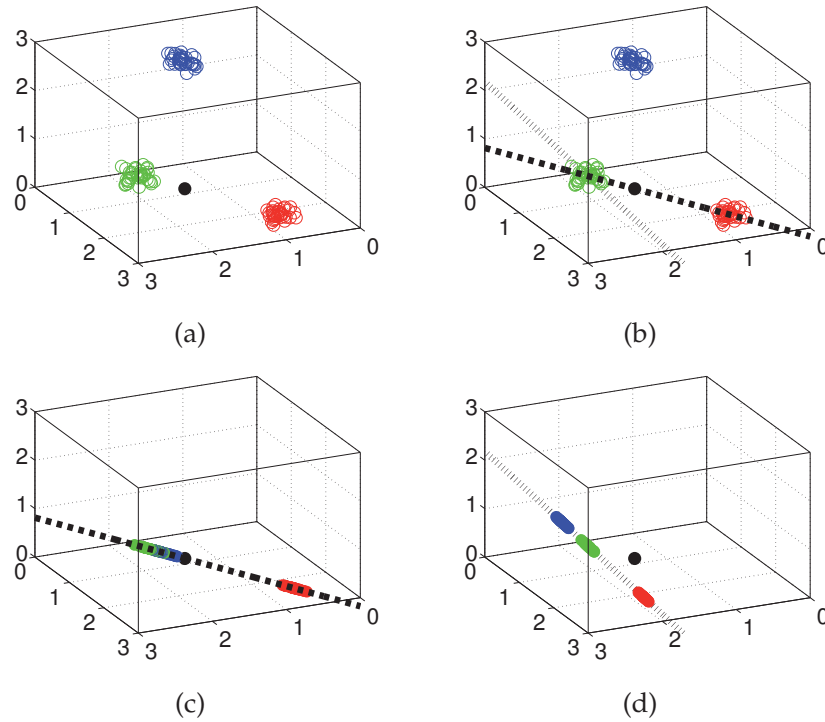


Figure 36: The distortion of distances implied by PCA. (a) Three clusters clearly separated in a given 3D space. The overall average is noted by the black ball. (b) The dashed line denotes the first principal component of the whole dataset (global PC). The pointed line represents the first principal component of the union of the green and blue clusters, neglecting all red points (local PC). (c) Clusters projected on the overall PC. It can be noted that the green and blue clusters overlap. (d) Clusters projected on the local PC. The separation of the green and blue clusters is now much clearer.

Scheunders [Scheunders, 2000] proposed a local mapping approach to DR, based on a segmentation of the image into blocks, which does not efficiently gather pixels with similar properties. Moreover, this technique may produce inconsistent color mappings from one region to another, i.e., using only local mappings does not guarantee that one class of pixels will be given the same color in every region. Nevertheless, global contrast should also be taken into account. In [Le Moan et al., 2011a], we developed a tunable technique to perform DR with traditional approaches such as PCA or band selection locally, so as to emphasize the perceptual distance between specified groups of pixels. Here we propose an updated revision of this strategy, which we will refer to as Cluster-Adaptive Projections (CAP).

3.6.2 CAP

While traditional methods consider each spectral channel as a whole, the core idea of the CAP is to analyze subsets of pixels. For instance, if one desires to enhance the contrast between a couple of specific objects (groups of pixels), one must consider the union of them separately from the rest of the scene, in order to obtain a more dedicated analysis. Therefore, the first step of the proposed technique is to obtain a spatial segmentation of the image so as to identify the different classes/clusters of pixels. This can be achieved either manually or automatically, by means of any classification, clustering and/or segmentation algorithm, preferably distance-based (but not necessarily Euclidean distance-based). In this study, we used only the K-means classifier, in order to demonstrate the efficiency of the proposed technique, but this has to be considered as without loss of generality.

Let I be a spectral image segmented into K classes C_1, \dots, C_K and assume that an application requires to emphasize the difference between a certain number of objects contained in C_γ , with $\gamma \subset \{1, \dots, K\}$. We propose to consider three different mappings:

- A global mapping associated with the whole image (\mathcal{P}_G). Its first PC contains most of the scene's spatial information (which pertains to luminance).
- A local mapping associated with the objects of interest (OoI) on the raw image (\mathcal{P}_γ). In our experiments, we observed that the discriminative information of C_γ was maximized in the first PC. Indeed, the second and third principal components usually contain very little information with less than 10% of the global energy in most cases.
- A local mapping associated with the OoI on an invariant representation of the image [Ibrahim et al., 2011] ($\mathcal{P}_{\gamma,I}$). Its first PC contains highly discriminative information about C_γ , in terms of chroma only.

Since we would like to enhance local contrast without lowering the global one, we propose to use the first PC of all three aforementioned projection in a convenient mapping to the Hue-Saturation-Value color space: the first PC from \mathcal{P}_G to the Value channel, as it contains most of the scene's luminance, the first PC from $\mathcal{P}_{\gamma,I}$ to the Hue channel, as it contains discriminative chromatic information, and finally the first PC from \mathcal{P}_γ to the Saturation channel to complete this intuitive mapping. The resulting composite thus contains valuable discriminative information about the chosen OoI, while preserving the overall structure of the scene in terms of luminance.

3.6.3 Experimental setup and results

In order to demonstrate the efficiency of the proposed model, we imagined a particular application on the "Jasper Ridge" scene from the AVIRIS sensor: we would like to enhance the contrast between the urban area on the right side and its surroundings, thus making it more salient. Figure 37 depicts the classification map as well as the location of both the urban area its surroundings. Figure 38 shows the resulting first principal components obtained from the projection matrices \mathcal{P}_G (global), \mathcal{P}_γ (local - raw) and $\mathcal{P}_{\gamma,I}$ (local - invariant). It shows an improved saliency of the OoI on both local results. Figure 39a depicts the resulting composite of a mapping of the 3 first global PCs to the CIELAB color space while Figure 39b shows the result of our mapping strategy. It can be seen on the composite resulting from a global mapping that the overall contrast is poor as very few colors are depicted. On the other hand, the proposed local adaption renders the urban area in blue-ish hues, which makes it much more salient. Finally, Table 11 depicts the Euclidean distance between the cluster centroids of the two regions that we aim to separate, for all three mapping strategies. It shows that the local mapping yields a better cluster separation on the first PC.

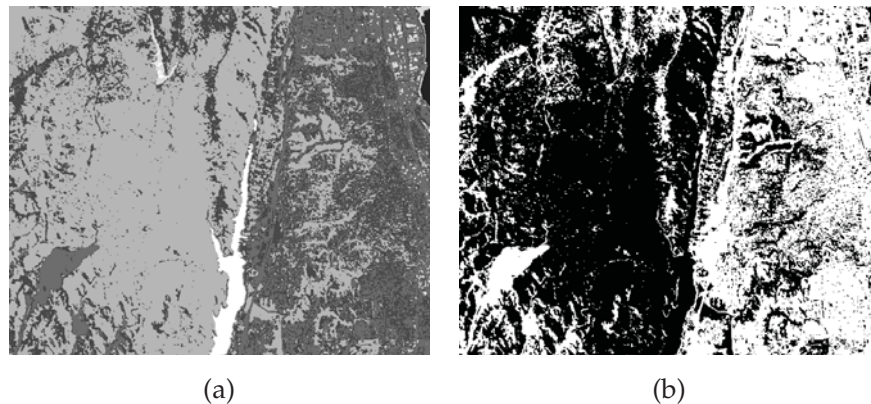


Figure 37: (a) Results from a K-means classification ($K=7$) (b) selected classes (objects of interest).

3.6.4 Conclusion

In this Section, we briefly investigated local DR, based on a spatial segmentation of the scene. We demonstrated that the principal components of a subset of pixels containing objects of interest allows for a better contrast between these objects than the overall PCs. We encourage further research on this topic, especially on the use of different distances, which would for instance better suit the underlying non-linearities of spectral images.

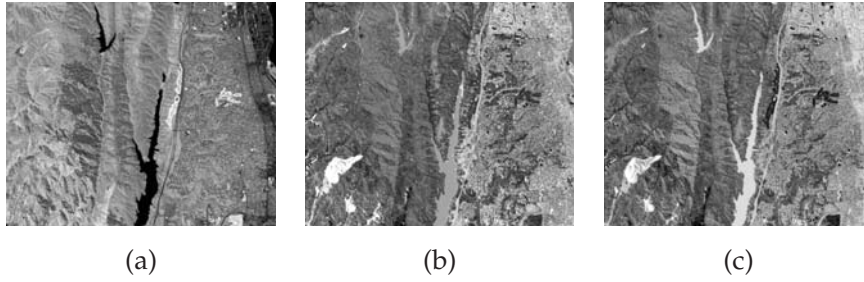


Figure 38: (a) First PC from \mathcal{P}_G (b) First PC from \mathcal{P}_γ (c) First PC from $\mathcal{P}_{\gamma,I}$. Both local strategies show a better contrast on the previously selected objects.

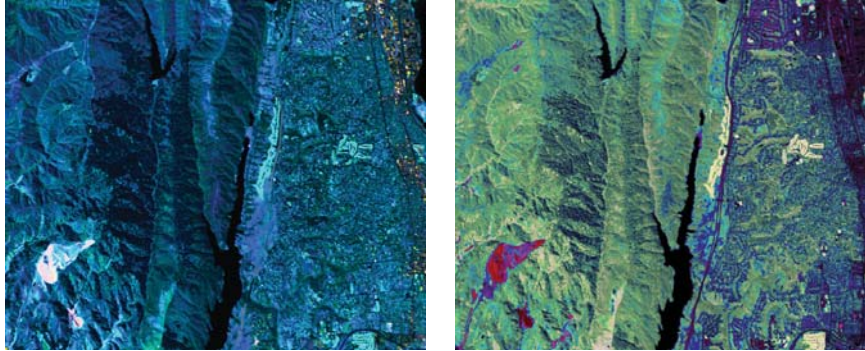


Figure 39: (a) Mapping of three first global PCs to CIELAB (perceptual distance between cluster centroids = 18) (b) Result from the proposed strategy (perceptual distance between cluster centroids = 56). The proposed approach shows a better contrast between the previously chosen objects. For instance, the urban area on the right side stands out much more.

Table 11: Euclidean distances. The global distance is computed using all three PCs.

	\mathcal{P}_G	\mathcal{P}_γ	$\mathcal{P}_{\gamma,I}$
Global	1.34	1.30	1.30
PC1	0.10	0.31	1.30
PC2	1.32	1.30	0.08
PC3	0.18	0.34	0.02

3.7 CONCLUSIONS

In this chapter, we presented a variety of tools for the visualization of multi/hyperspectral images. First, we introduced two new objective measures to assess the naturalness and color entropy of trichromatic composites. We demonstrated that they are able to better grasp visual

information than other commonly used measures. Then, a perception-based spectrum segmentation was introduced, aiming mainly to better take into account natural and visual information in DR methods such as PCA or band selection, while alleviating their computational complexity. Furthermore, we developed an information-based framework for band selection, which allowed us to suggest a possible relation between the subjective notion of naturalness in spectral images and Shannon's entropy. Results show indeed very good naturalness, much better than other state-of-the-art approaches based on second-order statistics, orthogonality and structure/compactness. Composites are both appealing and contain valuable discriminative information. Finally, we briefly tackled the problem of distance distortions in PCA and proposed an adaptive strategy called CAP which is based on a computation of the projection matrices from both global and local topographies as well as a convenient mapping to the Hue-Saturation-Value color space. Results show an increased class-separability on a set of selected objects of interest. Several other techniques for visualization and scene understanding are introduced in the next chapter, which is focused on the notion of spectral saliency.

When it comes to visualizing spectral images, the main challenge lies in defining what needs to be shown, and what needs not. The variety of dimensionality reduction (DR) strategies presented in the previous chapters shows that there are many different manners to do so. In this chapter, we propose a new approach to extracting information for purposes of visualization and interpretation. We introduce the concept of spectral saliency, and define two models to compute highly informative saliency maps, through center-surround analysis of a variety of dedicated features:

- the High Dimensional Saliency (**HDS**) of a pixel p is a single value computed by means of dedicated center-surround comparisons at several scales and represents the prominence of p in terms of several features.
- the Spectral Saliency Profile (**SSP**) $\phi(p, \lambda)$ represents the saliency of p at wavelength λ , based on the computation of individual maps for each spectral channel.

Consequently, the remainder of this chapter is organized as follows: first, we review some important related works on the topics of saliency and visualization of spectral images before presenting and discussing the aforementioned concepts and models in Sections 4.2 and 4.3. Conclusions are drawn in Section 4.4.

4.1 INTRODUCTION

In the context of color and computer sciences, the notion of saliency is closely related to the one of visual attention: "from a given scene, which objects/features will first draw attention and why?". Following early influential work by Treisman *et al.* [Treisman and Gelade, 1980] and Koch *et al.* [Koch and Ullman, 1985], Itti *et al.* [Itti *et al.*, 1998] proposed a general visual attention model to design so-called saliency maps, which purpose is to predict human gaze given a certain picture. This model is based on the extraction of three different features: color, intensity and orientation. The color feature results in two opposition maps: Red/Green and Blue/Yellow. Intensity is computed as the average of the three channels (R, G and B) while orientation feature maps are obtained by means of Gabor filters (to four different angles) on the intensity map. Each feature map is then derived into a gaussian pyramid (multiscale representation) allow-

ing for straightforward center-surround comparisons, which is actually the key idea when it comes to saliency [Gao et al., 2008]. Then, through a series of appropriate combinations, blurring and normalizations, the model outputs a single greyscale saliency map of the same size as the original image. More recently, Harel et al. [2007] proposed to model the feature spaces as Markov chains to better activate and normalize the feature maps. In [Hou and Zhang, 2007], the authors suggested a method using the log-spectrum of the input image in which statistical singularities are assumed to be salient features (or so-called *proto objects*). An information-theoretic approach based on an extraction of independent components of the scene was [Bruce and Tsotsos, 2009].

As opposed to these techniques that seek biological plausibility to relate as much as possible to human vision, some methods have been developed to use high-level features such as face recognition [Cerf et al., 2008; Sharma et al., 2008] or more generally the context of the input image [Goferman et al., 2010]. In [Judd et al., 2009], the authors proposed a supervised approach trained by a large database of eye-tracking data to train a bottom-up, top-down model of saliency based on low, mid and high-level image features. A signature-based approach was introduced by Hou et al. [2011]. A very interesting review by Frintrop et al. [Frintrop et al., 2010] addresses thoroughly the cognitive foundations of many saliency detection techniques from the literature.

Yet, saliency is a much broader concept inasmuch as it can be seen as a way of measuring informative content for any kind of data. Indeed, in the large field of neurosciences, *the salience (also called saliency) of an item – be it an object, a person, a pixel, etc – is the state or quality by which it stands out relative to its neighbors*¹, no matter what kind of feature space is considered. Only a few studies have extended this concept to more than just two-dimensional trichromatic images. Among them, and although not exactly in the scope of this paper, it is worth mentioning the pioneer work by Lee et al. [Lee et al., 2005] on 3D mesh saliency. Even though visual attention has already been used in the context of spectral images for dimensionality reduction purposes [Zhang et al., 2008], an actual saliency map from the entire high-dimensional image has not yet been considered, to our knowledge. Such a model presents however all the advantages of spectral imaging over trichromacy: sensitive to metameric matches, not bounded to visible wavelengths and device-independent.

In this chapter, we present a model of spectral saliency called HDS and we introduce the concept of SSP, which measures saliency locally in the spectral dimension.

¹ <http://en.wikipedia.org/wiki/salience> (neuroscience)

4.2 HIGH DIMENSIONAL SALIENCY

4.2.1 Architecture

The HDS model, whose general architecture is depicted in Figure 40, is based on the comparison of a pixel with its surroundings, in terms of different features and at several scales. Furthermore, the model uses an invariant representation of the spectral image, in order to remove the influence of specularities, shading and surface geometry, which are only relevant in the context of visualization. Indeed, we aim to understand the scene in terms of physical properties more than in terms of perception. We used the approach introduced by Ibrahim et al. [2011], which is derived from the standard dichromatic reflection model for dielectric and the extended dichromatic reflection model for metal. For a spectral channel \mathbf{b}_{λ_i} at wavelength λ_i , the corresponding invariant representation \mathbf{b}'_{λ} is defined as follows:

$$\mathbf{b}'_{\lambda}(i, j) = \frac{\mathbf{b}_{\lambda}(i, j) - \min_{n=1 \dots N} (\mathbf{b}_n(i, j))}{\sqrt{\sum_{m=1}^N \left(\mathbf{b}_m(i, j) - \min_{n=1 \dots N} (\mathbf{b}_n(i, j)) \right)^2}} \quad (57)$$

The HDS model is then structured as follows: first, *feature maps* are extracted, filtered and spatially subsampled into an 8-level gaussian pyramid, so that a vector-pixel at a fine scale represents the center of the corresponding location at a coarser scale (surrounding). As suggested in [Itti et al., 1998], center-surround comparisons are then achieved with centers at scales $c \in \{2, 3, 4\}$ (scale 1 is the original image, scale 8 corresponds to the lowest resolution) and surrounds at scales $s = c + \delta$, with $\delta \in \{3, 4\}$, which yields 6 so-called *conspicuity maps*, for each feature. They are, in turn, normalized and fused into the final map.

The next Sections give more detail about these steps, that is, which features and comparison measures to use and how to normalize the maps.

4.2.2 Features and comparison measures

There exist many features and measures/metrics to compare high-dimensional vectors [Imai et al., 2002]. In this Section, we investigate 6 features combined with 2 spectral match measures, in an attempt to define which ones are the most relevant for the HDS model. Each of the following features is defined for a given spatial location (i, j) and each measure is defined for a pair of feature-pixels.

- **Features**

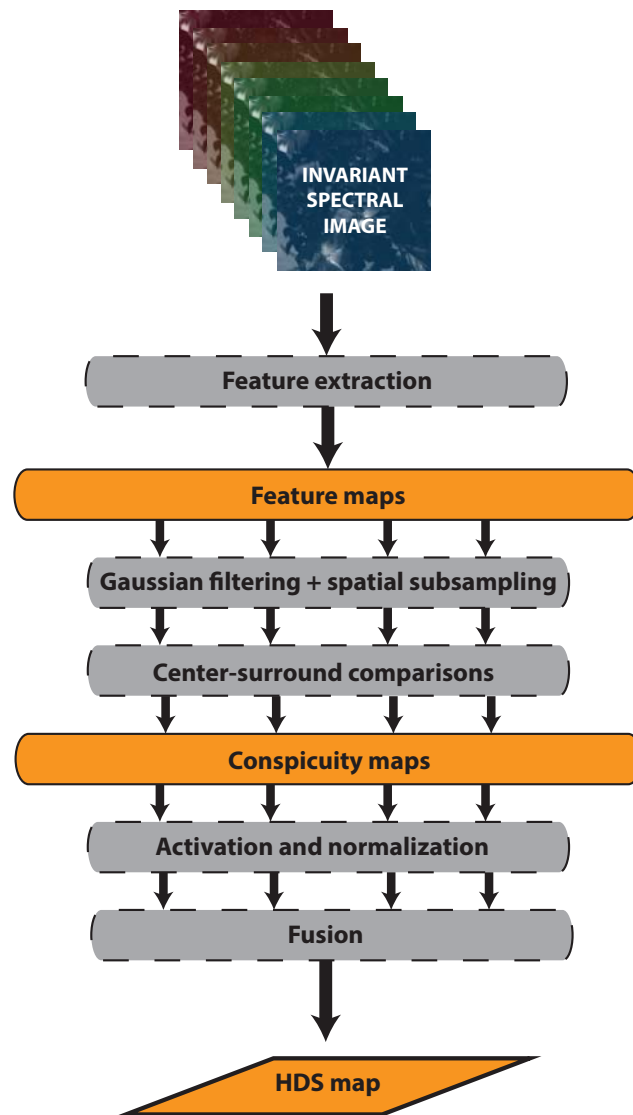


Figure 40: The HDS model.

- **Normalized reflectance:** $ref(i, j)$. High-dimensional pixels, taken directly from the image.
- **CIELAB trichromatic value:** $lab(i, j)$. Reflectance converted to CIELAB by means of the CIE standard observer Color Matching Functions (CMF) and after multiplication by the CIE D65 illuminant.
- **First three principal components** of the image: $pc_3(i, j)$.
- **Average reflectance value:** $avg(i, j)$.
- **Standard deviation of the reflectance curves:** $std(i, j)$. This feature gives an image of how homogenous is the pixel's reflectance along its spectral dimension.

- **Gabor filtering of the first principal component:** $gab(i, j)$. Linear filter for edge detection according to a specified direction. Allows to compare pixels in terms of local orientation (see [Itti et al., 1998]).

- **Measures**

- Euclidean distance (L2-norm):

$$D(\mathbf{f}_1; \mathbf{f}_2) = \sqrt{\sum_n (\mathbf{f}_1(\lambda_n) - \mathbf{f}_2(\lambda_n))^2} \quad (58)$$

- Spectral angle:

$$\theta(\mathbf{f}_1; \mathbf{f}_2) = \cos^{-1} \left(\frac{\mathbf{f}_1 \cdot \mathbf{f}_2}{\|\mathbf{f}_1\| \|\mathbf{f}_2\|} \right) \quad (59)$$

where (i, j) are spatial coordinates and \cdot is the dot product between two vectors. Note that we use the wavelength λ_n to describe the spectral dimension of the feature-pixels. Although, in the case of $pc3(i, j)$, this "spectral" dimension refers to the dimension spanned by the principal components.

Our study considers 6 combinations of the aforementioned features/measures, which we compared in terms of informative content (entropy) and correlation, on the results obtained over a population of 800 pixels randomly chosen from Foster's 2002 database (see Section 3.2 in the previous chapter). A high entropy means that the results are varied and thus informative. We note M_f the combination of measure M with feature f . We did not consider $gab(i, j)$ as they are the only orientation-related features and they will therefore be used in the HDS model in anyway. Results are presented in Tables 12 and 13.

Table 12: Comparison of features/metrics: entropy.

	D_{ref}	θ_{ref}	D_{lab}	D_{pc3}	D_{avg}	D_{std}
Entropy	7.13	7.10	6.73	7.14	7.10	7.03

We observe that the most informative features are the Euclidean distance in the high-dimensional space (D_{ref}) and in the space spanned by the first three PCs (D_{pc3}). They are highly correlated together (0.99), but also to D_{lab} (> 0.8), which implies that one and only one of them should be retained for our model. In the previous chapters, we argued that the Euclidean distance becomes less relevant in high dimensionality due to the *curse of dimensionality*, it can therefore be discarded. Furthermore, we make the assumption that D_{pc3} is more valuable than D_{lab} because it yields a higher entropy, but also because perceptual information is irrelevant in the present framework, which

Table 13: Comparison of features/metrics: correlation.

	D_{ref}	θ_{ref}	D_{lab}	D_{pc3}	D_{avg}	D_{std}
D_{ref}		0.18	0.87	0.99	-0.16	0.03
θ_{ref}			0.24	0.17	-0.15	0.10
D_{lab}				0.86	-0.01	0.16
D_{pc3}					-0.16	0.03
D_{avg}						0.52
D_{std}						

designs a model of spectral saliency. Finally, we note that the combination feature/measure presenting the best overall tradeoff between informative content and correlation is D_{std} .

Based on these remarks, we made the following selection for the HDS model:

- θ_{ref} , as a measure of chromaticity, independent from lightness.
- D_{pc3} , as a PCA-based coarse color difference metric.
- D_{std} , as a measure of grayness or saturation.
- D_{gab} , as a measure of orientations.

4.2.3 Activation, normalization and fusion

After extracting the feature maps and performing the center-surround comparisons, the resulting *conspicuity maps* contain raw saliency information, which need to be normalized and eventually fused into a single and final map. The role of normalization is to *activate* the maps, that is to concentrate brightness into a fewer key locations in order to increase their interpretability [Harel et al., 2007]. We propose to use a simple activation operator, based on mathematical morphology. For a given input intensity image \mathbf{b} , we define $A_r(\mathbf{b})$ as follows:

$$A_r(\mathbf{b}) = \left(\frac{\mathbf{b}}{\max(\mathbf{b})} \right) \bullet \mathcal{B}_r \quad (60)$$

where \bullet is the morphological closing operator [Serra, 1982] and \mathcal{B}_r is ball-shaped structural element of radius r . The greyscale morphological closing spreads out small bright regions so as to enlarge the visibility of salient locations. The amount of spreading is set by the size of the structural element \mathcal{B}_r , which may be adjusted according to the user's need. Small sizes make the activation less effective, but on the other hand, a large structural element may yield a too coarse

result. In this study, we set it to $\frac{1}{20}$ of the image width. Eventually, each feature analysis outputs (in scale 4) a single conspicuity map (the four orientations are averaged) which are then normalized by A_r and averaged for fusion. Averaging the maps allows not to favor one combination feature/measure over the others. Figure 41 shows an example of resulting map, before and after activation.

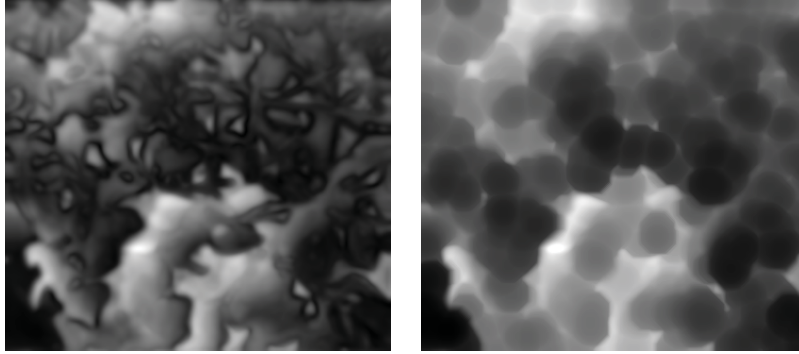


Figure 41: Example of normalization of θ_{ref} conspicuity map obtained from the first scene of Foster's 2002 database. Left: raw ; right: activated. The brightness of a pixel represents its saliency (white pixel: very salient).

The final map is in turn normalized and blurred. It will be referred to as \mathcal{M}_{HDS} (Figure 42 shows an example).

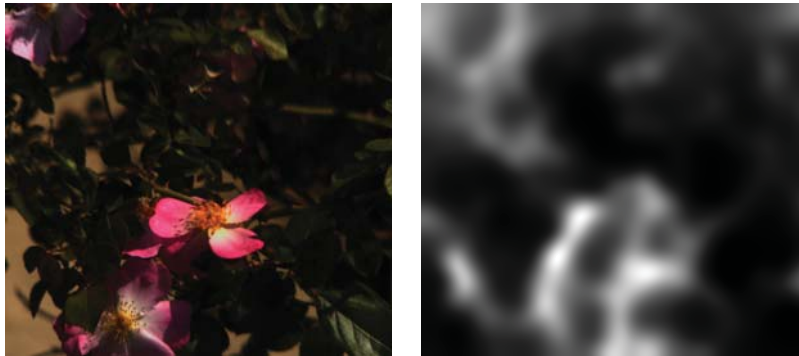


Figure 42: Example of saliency map from the HDS model. Left: *true color* composite. Right: corresponding map.

The output of the HDS model thus represents those pixels whose reflectances stand out, in terms of the presented features. Note that it can be applied regardless of the dimensionality of the image. However for non-spectral images (less than 4 channels), other combinations feature/measures should be more relevant. For trichromatic images, we suggest to use only the Euclidean distance in the device-independent CIELAB color space (D_{LAB}) and in the Gabor maps (D_{gab}). As for greyscale images, we suggest to use D_{ref} (which becomes a simple difference) and D_{gab} .

4.2.4 *In color*

We would like to go further into the concept of spectral saliency and take full advantage of the human visual system, by creating saliency maps in color. We aim indeed to create a single map, highly informative and easy to interpret. Therefore, based on the assumption that several local analyses can yield more information than a single global one, we propose to create color composite maps whose three channels result from a different local analysis each. By local, we mean in the spectral dimension, that is, the image spectrum is to be partitioned into three parts, roughly corresponding to the red, green and blue ranges of wavelengths (see Section 3.4 in the last chapter) within which the HDS model is applied, independently from one part to another. The resulting maps are then concatenated and mapped to sRGB, so as to obtain a highly informative map in color, which we will refer to as $\mathcal{M}_{\text{HDS}}^c$. Figure 43 illustrates the enhanced model, which will be referred to as HDS-Color (HDS-C), and Figure 44 gives an example of result.

On these maps, a bright pixel depicts a high saliency in a range of wavelengths corresponding to its color (for example, red bright: very salient in the last segment, but poorly in the two others). This simple rule of interpretation, makes it possible to adapt the spectrum segmentation strategy to the application. For instance, a visible/near InfraRed (nIR) partitioning would depict in one color the pixels which are salient in the visible and in another color the ones that are salient in the nIR, as illustrated in Figure 45.

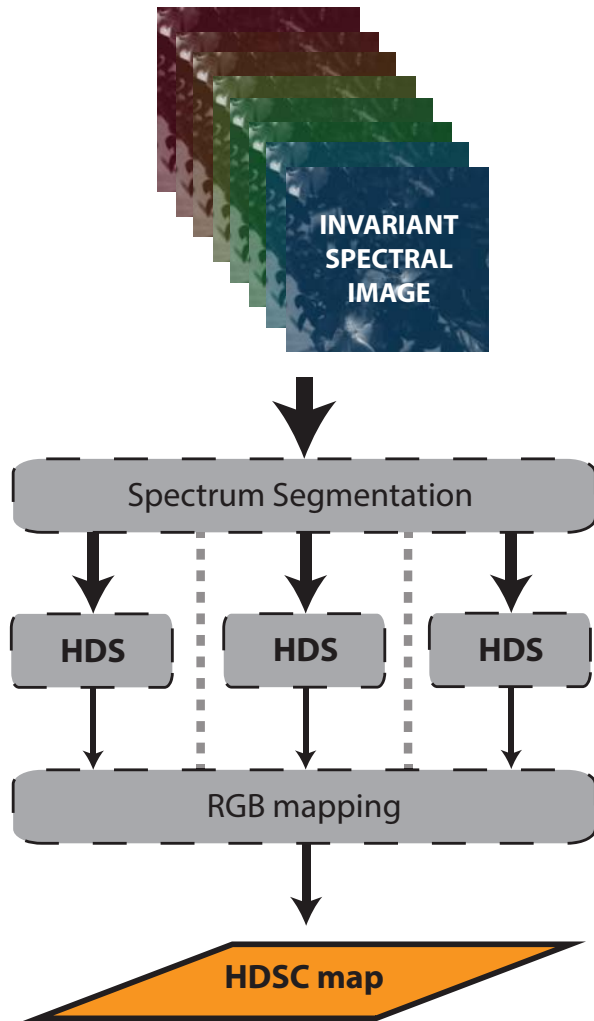


Figure 43: The HDS-C model.

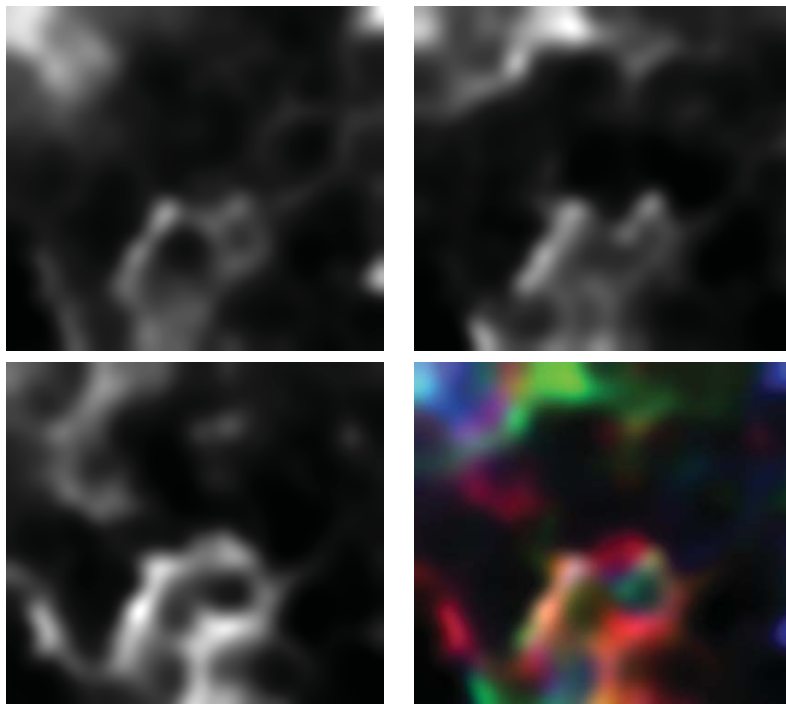


Figure 44: Example of enhanced saliency map obtained from a 31-bands spectral image. From top to bottom and from left to right: Blue map, Green map, Red map, $\mathcal{M}_{\text{HDS}}^c$. The variety of colors shows the sensitivity of the HDS model to the range of wavelengths it is used in. Interpretation is as follows: green pixels depict locations which are salient in the green segment, and respectively for the other colors.

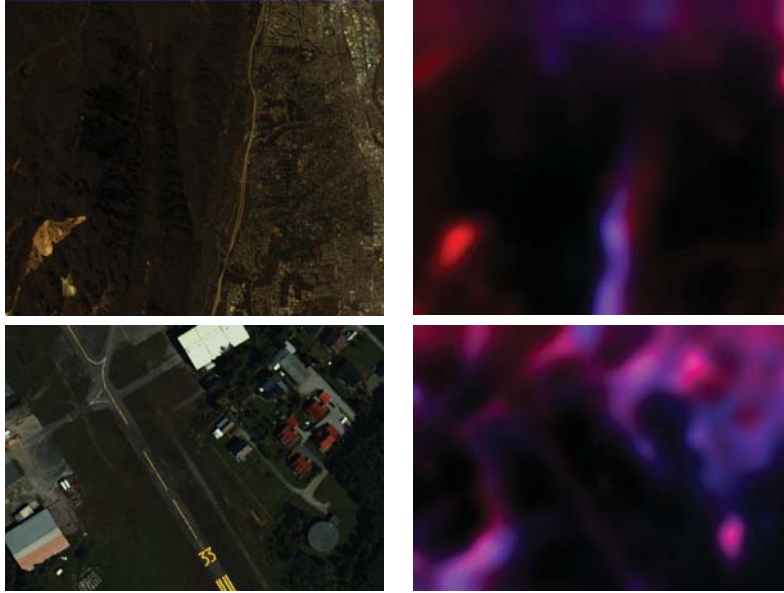


Figure 45: Example of enhanced saliency maps with a visible/nIR partitioning. They show very valuable information such as the river in the first image (surrounded by minerals) and the urban area in the second one (surrounded by vegetation) that stand out mostly in the infrared.

4.2.5 Results

In this section, we present the resulting maps obtained on a variety of spectral images and discuss on the similarity of the HDS with the model presented in [Hou et al., 2011], which we will refer to as Signature-based Visual Saliency (SVS) and which was applied on the *true color* composites. Figure 46 shows the results obtained on four spectral images: scene 1 from Foster's 2002 database ("Flowers"), scene 2 from Foster's 2004 ("Fern"), as well as the Jasper Ridge and Norway scenes (see Section 3.2 in the previous chapter).

It is important to note that these maps depict non-visual saliency, which can be seen as a measure of informative content, as they are computed regardless of the human visual system. Overall, the HDS and SVS models extract different features from the scenes, particularly on the two hyperspectral images. The central river on the "Jasper Ridge" scene is very salient in both visible and nIR, as shown in Figure 45 and therefore shows on the HDS and HDS-C maps. On the "Norway" scene, our models emphasize the urban area on the right side, as it is surrounded with vegetations, creating a contrast in the nIR, whereas the SVS map depicts the red rooftops and road signs as salient features. On the "Fern" image, HDS detects mostly the part with more branches than green leaves, instead of the conspicuous

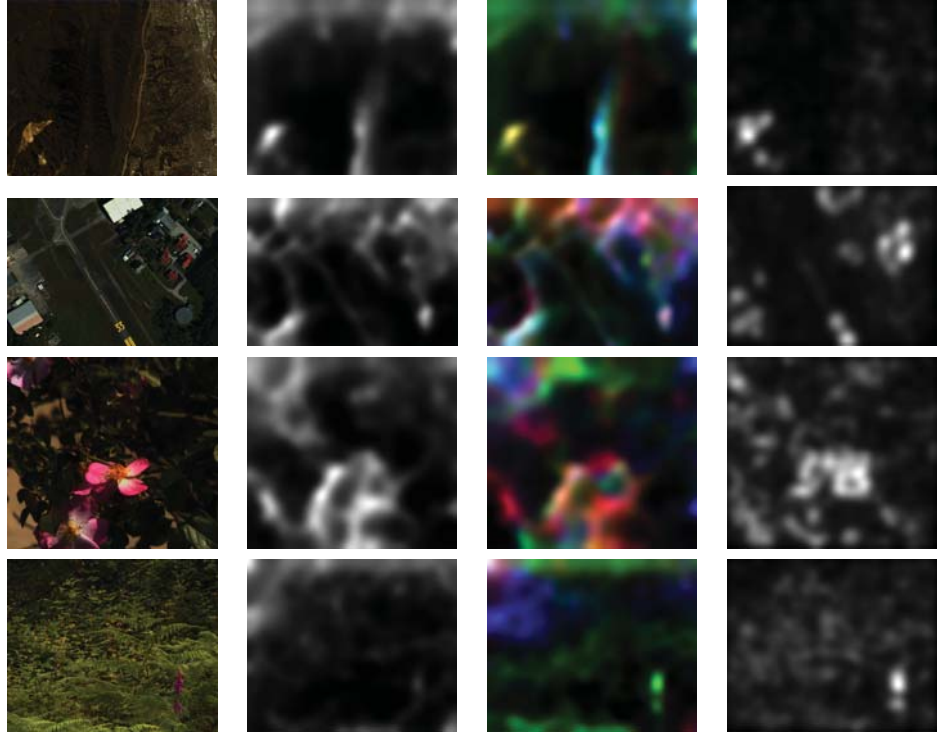


Figure 46: Results obtained on 4 images. Column-wise: *true color* composite, greyscale HDS $\mathcal{M}_{\text{HDS,color}}$, HDS $\mathcal{M}_{\text{HDS}}^c$, SVS map [Hou et al., 2011]. Row-wise: "Jasper Ridge", "Norway", "Fern" and "Flowers" scenes.

purple flower, which shows on the HDS-C map which allow to assess that this flower is salient only in the green wavelengths. As for the "Flowers" scene, it is mostly the central flower and a green area on the top side that stand out according to our models.

4.2.6 Applications

4.2.6.1 Evaluating composites

Since the HDS model can be applied to any dimensionality, it is possible to compare two images with different number of channels. Consequently, it can serve to assess how accurately a dimensionality reduction technique conveys such information into a trichromatic composite. We propose a measure that will be referred to as Mutual Saliency (*MS*) based on the **true mutual information** between the maps in high- and low-dimensionality:

$$MS(\mathcal{M}_1; \mathcal{M}_2) = \frac{I(\mathcal{M}_1; \mathcal{M}_2)}{\min [H(\mathcal{M}_1); H(\mathcal{M}_2)]} \quad (61)$$

where \mathcal{M}_1 and \mathcal{M}_2 are two saliency maps from the HDS model, and of two images of same size but potentially different number of spectral channels. See Section 3.5 in the previous chapter for details on entropy and mutual information.

Figures 47 and 48 depict 6 different composites each, obtained from the Jasper Ridge and Norway scenes, respectively. Tables 14 and 15 give the results obtained with the proposed measure as well as two others, for comparison: ρ_θ is the preservation of pairwise angles as in [Jacobson and Gupta, 2005] and ν is the measure of naturalness introduced in the previous chapter.

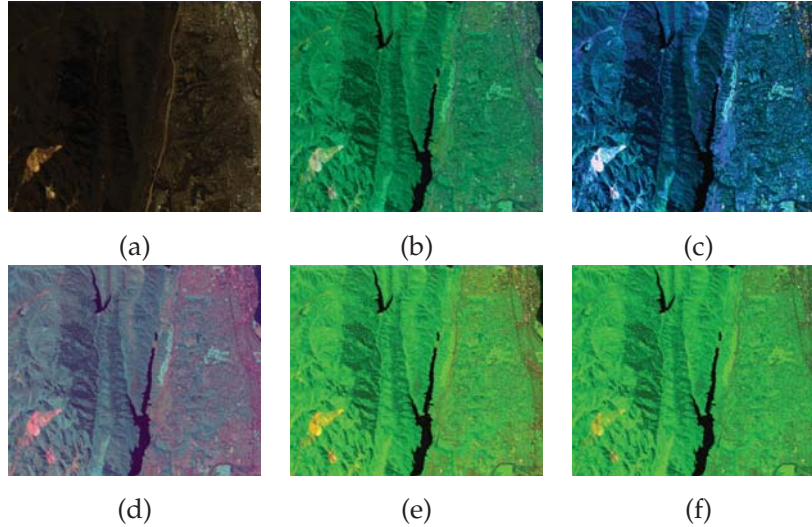


Figure 47: Example of composites from the Jasper Ridge scene. (a) *true color*, (b) *pseudo true color* (stretched CMF), (c) PCA to CIELAB, (d) segmented PCA, (e) linear prediction-based band selection and (f) one bit transform-based band selection.

Table 14: Saliency-based evaluation of dimensionality reduction techniques: Jasper Ridge scene composites.

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6
MS	0.23	0.35	0.24	0.26	0.28	0.28
ρ_θ	0.19	0.93	0.64	0.74	0.72	0.79
ν	0	59	58	50	82	81

We observe that our measure ranks the *pseudo true color* strategy first and the regular *true color* strategy last. Indeed, the latter discards all the channels centered at non-visible wavelengths and therefore fails to properly retain the scene properties. Mutual saliency is not correlated neither with ρ_θ nor with ν , which implies that it conveys

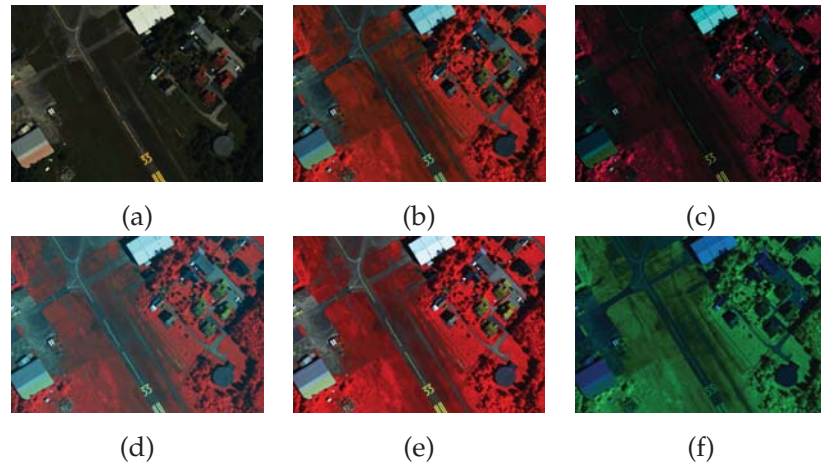


Figure 48: Example of composites from the Norway scene. From left to right and top to bottom: *true color*, *pseudo true color* (stretched CMF), PCA to CIELAB, segmented PCA, linear prediction-based band selection and one bit transform-based band selection.

Table 15: Saliency-based evaluation of dimensionality reduction techniques: Norway scene composites.

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6
MS	0.26	0.46	0.35	0.41	0.41	0.39
ρ_0	0.27	0.83	0.92	0.78	0.87	0.68
ν	0	33	42	30	40	38

a different kind of information. It is therefore a valuable measure in complement with others, to evaluate saliency conveyed during the dimensionality reduction.

4.2.6.2 Dimensionality reduction

The major drawback of most DR methods in the literature is that they are based on the assumption that all the pixels are part of the same population, i.e. they perform a global mapping. Some approaches such as the linear prediction-based band selection [Du and Yang, 2008] require a regular subsampling of the pixel population (down to 1% without noticeable change, according to the authors) in order to alleviate their respective complexity. Scheunders [Scheunders, 2000] proposed to spatially divide the image into square blocks in order to achieve local mappings by means of PCA and Neural Network-based techniques. However, natural scenes are rich and complex, showing large contrasts among their constituents, therefore a more dedicated spatial partitioning would better take care of these properties.

In this section, we investigate the use of the most salient pixels in the scene to compute the projection functions. Three sets are extracted: the salient ones, the surroundings and the background. Only a few pixels are retained for each set, by means of PCA, so that to represent each of the first two sets aforementioned in the dimensionality reduction process. One of the tasks of DR is to convey and/or enhance the relative discrepancies between the various Objects of Interest (OoI), contained in the input data. When it comes to images, it is generally equally weighted over the spatial dimensions, despite the rich and complex properties of natural scenes.

By thresholding an HDS map into three parts, we isolate different sets of pixels according to their respective contribution to the scene:

- The **salient pixels**, Ω_1 , are the pixels whose level of saliency is higher than a threshold T_{up} .
- The **surrounders**, Ω_2 , are the pixels whose level of saliency is lower than T_{up} and higher than T_{down} .
- The **background pixels**, Ω_3 , are all the rest.

Figure 49 shows an example of such segmentation on a natural scene, using different threshold values. The values of the optimal threshold are of course scene-dependent. We recommend to define them according to the separation of objects present in Ω_1 and Ω_2 . For example, the segmentation in Figure 49b would be a more relevant choice than the one in 49c, where the flower petals spread out on both Ω_1 and Ω_2 , which is undesirable. In this study, the optimal thresholds were defined manually for each scene.



Figure 49: Examples of saliency-based thresholding. Left: *true color* composite, Middle: $T_{up} = 0.3$ and $T_{down} = 0.1$, Right: $T_{up} = 0.5$ and $T_{down} = 0.3$. Saturated areas represent the salient pixels while surroundings are shown in grey and background in black.

In order to extract a set of representative pixels from each segment, we used PCA, over the spatial dimensions. During our experiments, we assessed that no more than five principal components are necessary to explain most of the data's energy (more than 95%) and therefore to represent each Ω_1 and Ω_2 (as we disregard the background). Eventually, only 10 pixels are considered to compute the projection matrix.

Moreover, by mastering the number and type of objects present in the input data, one allows the latter algorithm to be more dedicated to conveying the discrepancies between, in our case, objects in Ω_1 and Ω_2 . Considering the relatively high computational complexity of PCA, we performed a random subsampling of 50 pixels in both groups. Moreover, resulting components are then normalized so that to fit the range $[0..1]$. Figure 50 shows an example of the principal components obtained.

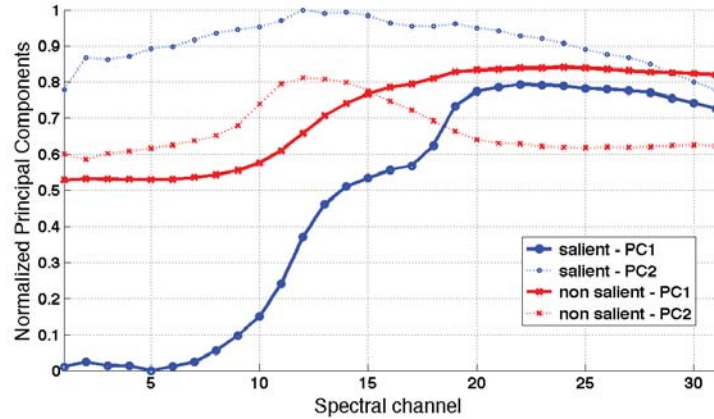


Figure 50: Examples of (first and second) principal components obtained. Disks: representing Ω_1 and Crosses: representing Ω_2 . We can observe for instance that the first PCs (plain lines) are discriminable mostly in the first half of the image's spectrum.

We selected three dimensionality reduction techniques to illustrate the proposed approach.

- Information-based Band Selection (IBS). See Section 3.2 in chapter 3.
- Linear Prediction-based Band Selection LPBS [Du and Yang, 2008] is a state-of-the-art band selection approach which consists of progressively selecting bands by maximizing their respective orthogonality.
- PCA_{HSV} is the traditional Principal Components Analysis of which components are mapped to the HSV color space, according to the normalization used in [Tyo et al., 2003], without shifting the origin of the HSV cone.

Band selection approaches have been implemented in such a way that the bands are eventually sorted by descending wavelength before mapping to sRGB.

Figure 51 shows the *true color* composites of the images used in this study, as well as the corresponding saliency maps. Figures 52 to 54 show the results obtained by means of the different dimensionality reduction techniques, both by considering all the pixels in the image (or a uniform subsampling for **LPBS**) and only a reduced set of pixels.

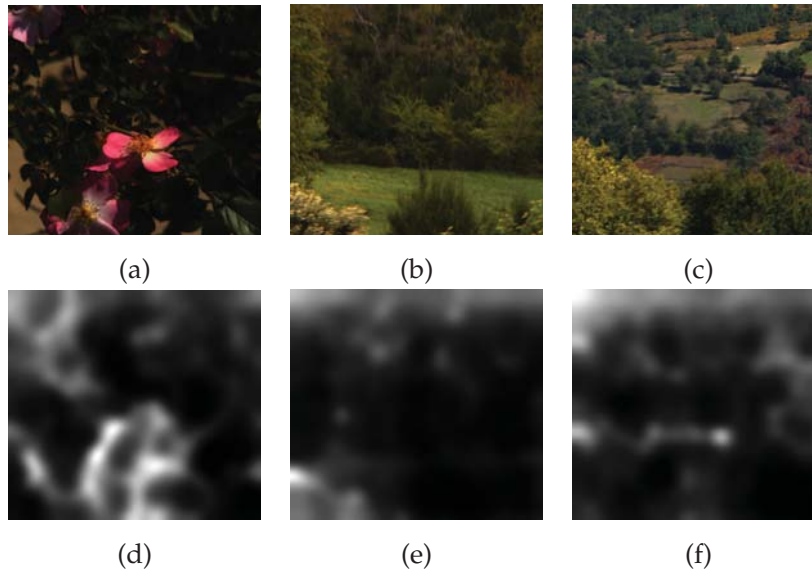


Figure 51: *True color* composites (first row) and the corresponding spectral saliency maps (second row).

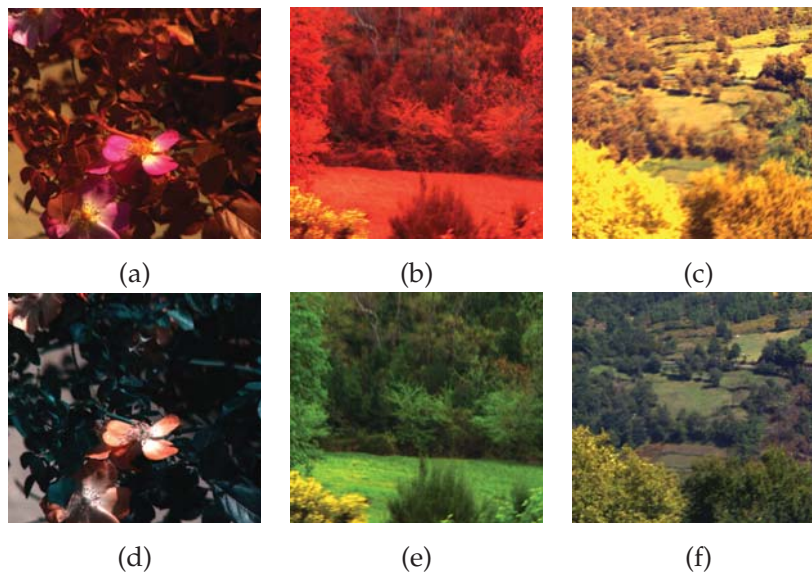


Figure 52: **IBS** approach. First row: using all the pixels in the image. Second row: using a reduced set.

The optimal thresholds for each scene are given in table 16. **LPBS** obtains its best results with a quite low upper threshold (0.3), while

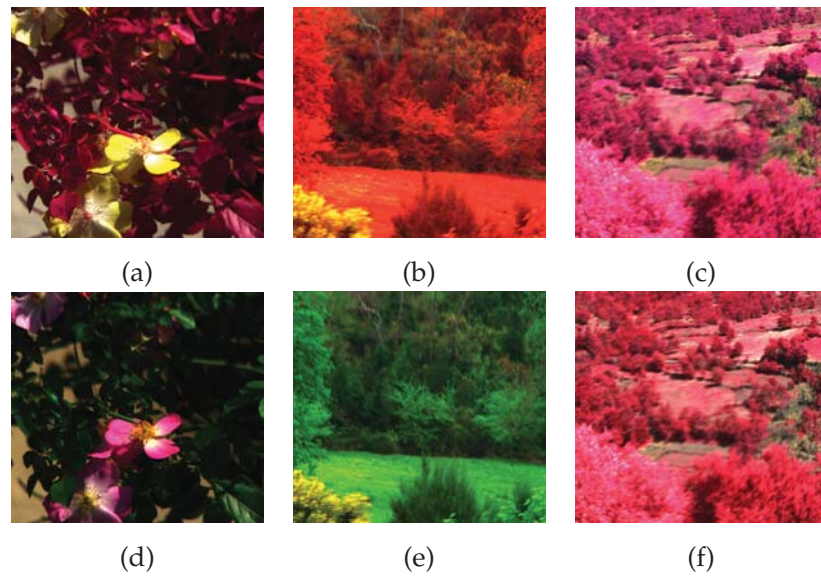


Figure 53: **LPBS** approach. First row: using a uniform subsampling of 1% of the image's pixels. Second row: using a reduced set.

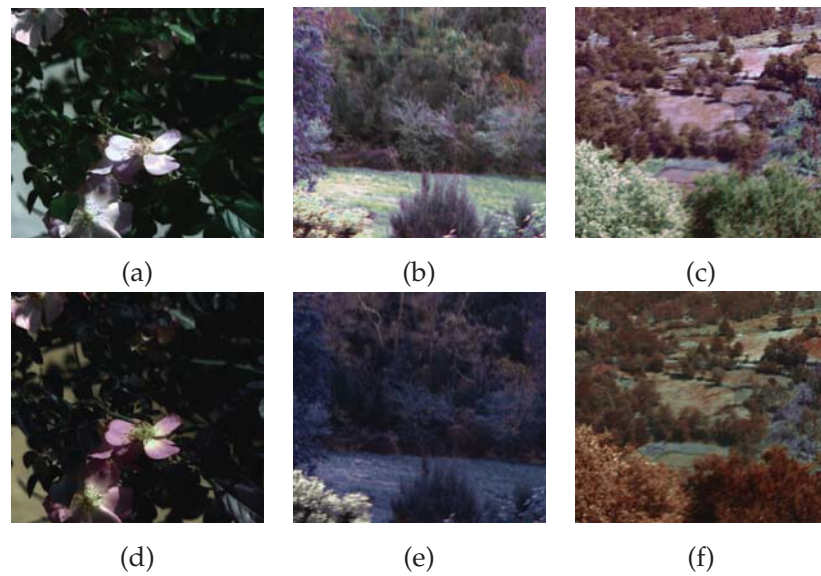


Figure 54: **PCA_{hsv}** approach. First row: using all the pixels in the image. Second row: using a reduced set.

IBS and **PCA_{hsv}** perform better with a very reduced set of pixels. Overall, we observe that the most salient objects are emphasized, mostly because of a darkening or a diminution of contrast of their surroundings.

In order to objectively evaluate the results, we used the color difference metric ΔE^* , which measures the Euclidean distance in the perceptually uniform color space CIELAB. Let ω_1 and ω_2 be two

sets of 20 randomly selected pixels from Ω_1 and Ω_2 , respectively. Now let $\overline{\Delta_{12}}$ be the average color difference between ω_1 and ω_2 , on a composite obtained with considering all pixels or a uniform sub-sampling and let $\overline{\Delta_{12}'}$ have the same definition but on a composite obtained by the proposed approach. We define the improvement of saliency $\delta_s = \overline{\Delta_{12}'} - \overline{\Delta_{12}}$. Table 17 shows the values obtained in this experiment.

Table 16: Optimal thresholds.

	scene 1	scene 2	scene 3
T_{up}	0.3	0.5	0.9
T_{down}	0.1	0.3	0.8

Table 17: Improvements of saliency δ_s . Difference of average Euclidean distance in CIELAB between Ω_1 and Ω_2 , using all the pixels versus using only a subset.

	scene 1	scene 2	scene 3
IBS	18.8	44.9	23.7
LPBS	13.9	32.5	2.1
PCA_{hsv}	20.5	47.3	9.0

Results show that there is an overall increase of conspicuity for the top salient objects. It is not surprising to see that the PCA is more sensitive to the pixel selection as it is more adaptive to the data and has more degrees of freedom than the BS techniques. However, it also shows less contrast in the background areas, due to the fact that these pixels are disregarded during the computation of the projection matrix. Scene 3 shows the best results, mainly because of the well-defined salient region on the bottom left side.

4.3 SPECTRAL SALIENCY PROFILES

4.3.1 Definition

Whereas the HDS model is based on the comparison of high dimensional vectors, the SSP involves the computation of one saliency map per channel. Therefore it engenders, for every pixel, a vector of the same dimensionality (the so-called profile) that depicts its prominence along the spectrum. Since a spectral channel is an intensity image, only two features are necessary here: intensities and orientations. The same scales and normalizations as in the previous Section

were used. Figure 55 gives three examples of SSP, together with their respective reflectance curves.

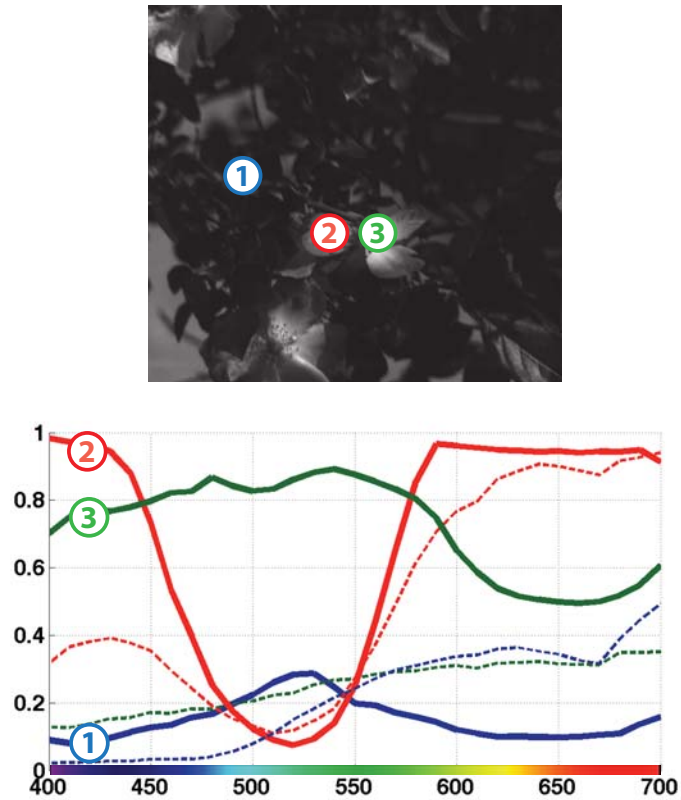


Figure 55: Examples of SSP (plain lines) and their corresponding reflectance curves (dashed). Pixel 1 is fairly salient all over the spectrum, with a slightly worse prominence at the end of the blue range, where its surroundings (green leaves mostly) start to be very reflective. Pixel 2 differs from its neighborhood mainly between 400 and 600 nanometers (yellow/early red), meaning that it is either much more or much less reflective in the blue and green ranges than its surroundings (pixels of the like of pixel 1). Of course, interpretation is easier if a prior knowledge about the scene is available.

4.3.2 Application to band selection

In this section, we use the information-based band selection (**IBS**) strategy presented in the previous chapter, but applied on the SSP instead of on the reflectance data. The goal is to select bands that show different things, that is, with different saliency maps. The approach will be referred to as Saliency based-IBS (**S-IBS**). For our experiments, we used three calibrated multispectral datasets, ranging in the visible spectrum (400-700 nm):

- "MacBeth" is the well-known MacBeth CC color calibration target. It contains 31 channels.
- "Sarcophagus" is a 35 bands (400-740nm) multispectral image representing a portion of a 3rd century sarcophagus from the St Matthias abbey in Trier, Germany [Simon et al., 2010]. It was acquired by means of an 8 channel filter wheel camera ranging only in the visible spectrum (400-740nm). Reflectance was reconstructed by means of a supervised neural-network-based algorithm [Mansouri et al., 2005].
- "Mural" represents a 16th century mural painting from the Brömser Hof in Rudesheim, Germany. It has the same properties as "Sarcophagus".

In order to evaluate the performances of our method, we selected two other dimensionality reduction techniques for comparison. Figure 56 depicts the resulting color visualization of all the images and for the three dimensionality reduction approaches: PCA_{HSV} , LPBS and S-IBS .

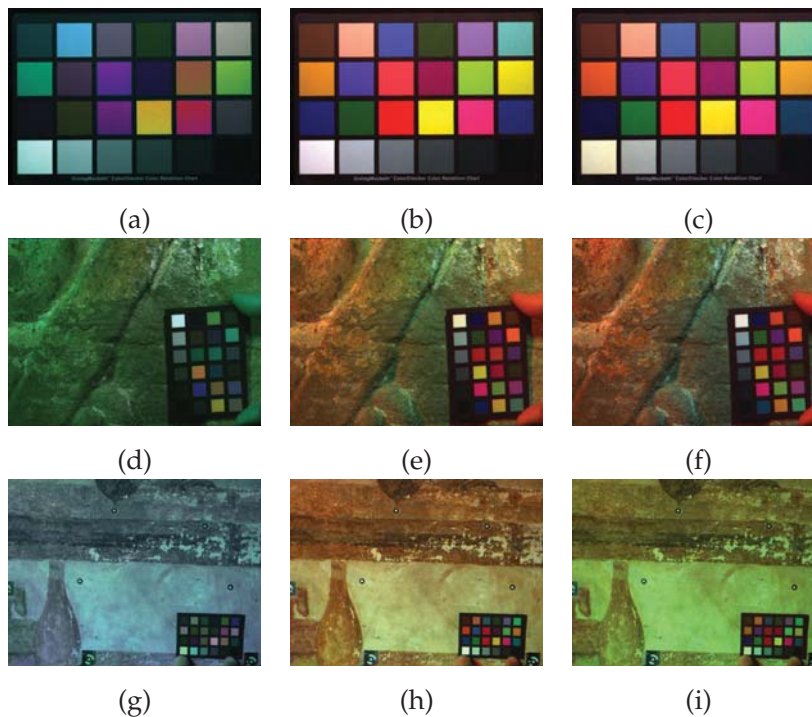


Figure 56: Different representations for each dataset (first column: PCA_{HSV} , second: LPBS and third: S-IBS)

It can be seen that the PCA-based method gives the least appealing results, while LPBS and S-IBS give quite similar and "eye-satisfying" images. On the first dataset, one can notice that the white patch (bottom left) is whiter in the LPBS results, but still very discriminable

from all the others in the **S-IBS**. However, if we now look at the orange-yellow patch (second row, last one on the right), it is much more discriminable from the yellow one in the result by our method. Similar trends on the blue/violet patches allows us to assess that our dimensionality reduction method is the one conveying the more discriminative information (in a perceptual manner).

Furthermore, and in order to objectively compare the results, we chose to use the MacBeth CC target, present in each scene and to compare the CIELAB values of a set of 480 randomly selected pixels (20 by patch) with the "ground truth" ones, provided by *Gretag*, by means of the ΔE_{ab}^* color difference metric. Dynamics of the colorspace components have been set as follows: $L^* \in [0..100]$, $a^* \in [-100..100]$ and $b^* \in [-100..100]$. With this framework, we aim at an assessment of how accurately the dimensionality reduction method can convey the high variety of colors from a high dimensional space to three dimensions. Table 18 gives the minimal, maximal and average perceptual distances in CIELAB between the results and the "ground truth". It can be seen that, even though the **LPBS** gives slightly better minimal and maximal errors in two cases, the proposed approach outperforms it on each dataset, in terms of average ΔE_{ab}^* , and especially on the two last images.

Table 18: Colorimetric errors. We note $\Delta E = \Delta E_{ab}^*$.

		PCA_{hsv}	LPBS	S-IBS
ΔE_{\min}	"MacBeth"	8.55	3.36	3.80
	"Sarcophagus"	8.56	3.43	0.50
	"Mural"	8.38	5.95	2.30
ΔE_{\max}	"MacBeth"	86.47	48.39	45.74
	"Sarcophagus"	80.42	38.16	38.60
	"Mural"	85.63	45.65	35.12
$\overline{\Delta E}$	"MacBeth"	32.30	29.17	28.95
	"Sarcophagus"	35.10	17.31	13.86
	"Mural"	46.76	19.93	15.42

4.4 CONCLUSIONS

We introduced the concept of spectral saliency for multi/hyperspectral images analysis. It presents all the advantages of spectral imaging over trichromacy: sensitive to metameric matches, not bounded to visible wavelengths and device-independent. Two models were presented: the High Dimensional Saliency model and the Spectral Saliency Profiles from which many applications can be derived, and

particularly under the scope of visualization, such as saliency-driven conception, evaluation, adjusting and clustering of trichromatic composites from spectral images. We tackled two applications for the HDS model: evaluation of trichromatic composites and dimensionality reduction over a limited number of relevant pixels, as well as a band selection strategy based on the SSP. Moreover, we introduced a model to compute saliency maps in color, to help figuring out the range of wavelengths in which objects are salient. Results show that spectral saliency provides valuable information, which does not correlate neither with visual saliency, second-order statistics nor with our definition of naturalness, but serves however well for visualization-related applications. Therefore we expect the concept of spectral saliency to be considered for further investigation in the future, and particularly in order to define more precisely its relation with visual saliency.

CONCLUSIONS AND PERSPECTIVES

In this thesis, we presented a variety of methods of feature extraction for spectral image analysis, with a particular emphasis on problems related to display and visualization.

We developed a new band selection strategy called *Information-based Band Selection*, based on measures derived from Shannon's entropy in order to retain a subset of bands containing as much informative content as possible. A spectrum-segmentation method based on a thresholding of color matching functions allows to both alleviate the complexity of the band selection algorithm and enhance the naturalness of the resulting composites. Results show indeed very good naturalness, much better than other state-of-the-art approaches based on second-order statistics, orthogonality and structure/compactness, and thus allow us to suggest a possible relation between this subjective notion of naturalness in spectral images and Shannon's entropy.

Furthermore, two new quality measures were introduced, in order to evaluate trichromatic composites in terms of *naturalness* and entropy of perceptible colors (*color entropy*). We demonstrated that they are able to better grasp visual information than other commonly used measures.

The problem of distance distortion in principal components analysis was briefly tackled and an adaptive strategy called *Cluster-Adaptive Projections* was proposed. It is based on a computation of the projection matrices from both global and local topographies as well as a convenient mapping to the Hue-Saturation-Value color space. Results show an increased class-separability on a set of selected objects of interest.

Finally, we introduced the concept of spectral saliency and provided two models to compute it: the *High Dimensional Saliency* model and the *Spectral Saliency profiles*. The former creates a single saliency map from the whole spectral image whereas the latter depict the prominence of each pixel as a function of the wavelength. An enhancement of the HDS model allowing to create highly informative saliency maps in color was also introduced. Results show that spectral saliency provides valuable information, which does not correlate neither with visual saliency, second-order statistics nor with our definition of naturalness, but serves however well for visualization-related applications. Several applications of interest were considered and successfully tackled by means of these models.

Thus, the present work represents a step further in spectral image analysis, yet, we reckon that the following trails are worth considering for further investigation:

- Although we focused mainly on band selection in the third chapter, we believe that third order information measures can serve other contexts such as independent components analysis. That is, it would be interesting to consider an algorithm that seeks the most independent directions in the high-dimensional feature space, and in terms not only of mutual information, but also of co-information, and possibly higher orders.
- The notion of naturalness is yet to be defined in objective terms. We suggested that Shannon's entropy can convey valuable visual information, but it seems not well suited for all kind of data. It would be interesting to further study this relation, and especially in the context of a specific application.
- We showed that distance preservation during the dimensionality reduction is a criterion that can be tuned by considering solely subsets of pixels. Other classifiers than K-means should be investigated in this context, and not necessarily only distance-based.
- The concept of spectral saliency being brand new, we assume that there are many other applications in which it could potentially be demonstrated as useful. In particular, we think it would be interesting to consider it for object recognition, classification, etc. Our preliminary results showed that the activation/normalization process influences greatly object detection, therefore we believe that this could be a starting point.

The problem of visualizing spectral images, or more generally, large sets of data, is of particular interest as we are facing new challenges due to the recent advent of multimedia technologies and computer science. We hope that this thesis may contribute to elaborating new technologies and give new perspectives on the matter.

BIBLIOGRAPHY

- C.M. Bachmann, T.L. Ainsworth, and R.A. Fusina. Exploiting manifold geometry in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):441–454, 2005.
- P. Bajcsy and P. Groves. Methodology for hyperspectral band selection. *Photogrammetric engineering and remote sensing*, 70:793–802, 2004.
- M. Beauchemin and K.B. Fung. On statistical band selection for image visualization. *Photogrammetric engineering and remote sensing*, 67(5):571–574, 2001.
- A.J. Bell. The co-information lattice. In *Proceedings of the Fifth International Workshop on Independent Component Analysis and Blind Signal Separation*, 2003.
- A.J. Bell and T.J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.
- N.D.B. Bruce and J.K. Tsotsos. Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, 9(3), 2009.
- G. Buchsbaum and A. Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1218):89–113, 1983.
- S. Cai, Q. Du, and R.J. Moorhead. Hyperspectral imagery visualization using double layers. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3028–3036, 2007.
- S. Cai, Q. Du, and R.J. Moorhead. Feature-driven multilayer visualization for remotely sensed hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9):3471–3481, 2010.
- C. Cariou, K. Chehdi, and S. Le Moan. Bandclust: An unsupervised band reduction method for hyperspectral remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 8(9):564–568, 2010.
- M. Cerf, J. Harel, W. Einhäuser, and C. Koch. Predicting human gaze using low-level saliency combined with face detection. *Advances in neural information processing systems*, 20:241–248, 2008.
- C.I. Chang and Q. Du. Interference and noise-adjusted principal components analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2387–2396, 1999.

- C.I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3):608–619, 2004.
- C.I. Chang and S. Wang. Constrained band selection for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6):1575–1585, 2006.
- C.I. Chang, T.L. Sun, and M.L.G. Althouse. Unsupervised interference rejection approach to target detection and classification for hyperspectral imagery. *Optical Engineering*, 37:735, 1998.
- C.I. Chang, Q. Du, T.L. Sun, and M.L.G. Althouse. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(6):2631–2641, 1999.
- PS Chavez, GL Berlin, and LB Sowers. Statistical method for selecting landsat mss ratios. *Journal of Applied Photographic Engineering*, 8(1):23–30, 1982.
- H. Chen and P.K. Varshney. A human perception inspired quality metric for image fusion based on regional information. *Information Fusion*, 8(2):193–207, 2007.
- Y. Chen and R.S. Blum. A new automated quality assessment algorithm for image fusion. *Image and Vision Computing*, 27(10):1421–1432, 2009.
- T.M. Cover, J.A. Thomas, J. Wiley, et al. *Elements of information theory*, volume 6. Wiley Online Library, 1991.
- M. Cui, A. Razdan, J. Hu, and P. Wonka. Interactive hyperspectral image visualization using convex optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 47(6):1673, 2009.
- N. Cvejic, CN Canagarajah, and DR Bull. Image fusion metric based on mutual information and Tsallis entropy. *Electronics Letters*, 42(11):626–627, 2006.
- M. Dalla Mura, A. Villa, J.A. Benediktsson, J. Chanussot, and L. Bruzzone. Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis. *IEEE Transactions on Geoscience and Remote Sensing Letters*, (99):541–545, 2011.
- B. Demir, A. Celebi, and S. Erturk. A low-complexity approach for the color display of hyperspectral remote-sensing images using one-bit-transform-based band selection. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1 Part 1):97–105, 2009.

- H. Du, H. Qi, X. Wang, R. Ramanath, and W.E. Snyder. Band selection using independent component analysis for hyperspectral image processing. In *Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 32nd*, pages 93–98. IEEE, 2003a.
- Q. Du. Band selection and its impact on target detection and classification in hyperspectral image analysis. In *Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 374–377. IEEE, 2003.
- Q. Du and H. Yang. Similarity-based unsupervised band selection for hyperspectral image analysis. *IEEE Geoscience and Remote Sensing Letters*, 5(4):564–568, 2008.
- Q. Du, H. Ren, and C.I. Chang. A comparative study for orthogonal subspace projection and constrained energy minimization. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1525–1529, 2003b.
- Q. Du, N. Raksuntorn, S. Cai, and R.J. Moorhead. Color display for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 46:1858–1866, 2008.
- M. Fauvel, J. Chanussot, and J.A. Benediktsson. Kernel principal component analysis for feature reduction in hyperspectral images analysis. In *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic*, pages 238–241. IEEE, 2006.
- D.H. Foster, S. Nascimento, and K. Amano. Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience*, 21(03):331–336, 2004.
- J.H. Friedman and J.W. Tukey. A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers*, 100(9): 881–890, 1974.
- S. Frintrop, E. Rome, and H.I. Christensen. Computational visual attention systems and their cognitive foundations: A survey. *ACM IEEE Transactions on Applied Perception*, 7(1):6, 2010.
- D. Gao, V. Mahadevan, and N. Vasconcelos. On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of Vision*, 8(7), 2008.
- W.R. Garner. *Uncertainty and structure as psychological concepts*. Wiley, 1962.
- S. Goferman, L. Zelnik-Manor, and A. Tal. Context-aware saliency detection. In *Conference on Computer Vision and Pattern Recognition*, pages 2376–2383. IEEE, 2010.

- R.C. Gonzalez, R.E. Woods, and S.L. Eddins. *Digital image processing using MATLAB*. Prentice Hall, Upper Saddle River, NJ, 2004.
- JC Granahan and JN Sweet. An evaluation of atmospheric correction techniques using the spectral similarity scale. In *IEEE Geoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International*, volume 5, pages 2022–2024. IEEE, 2001.
- H. Grassmann. On the theory of compound colors. *Philosophical Magazine*, 7:254–64, 1854.
- A.A. Green, M. Berman, P. Switzer, and M.D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, 1988.
- T. Han and D.G. Goodenough. Investigation of nonlinearity in hyperspectral imagery using surrogate data methods. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10):2840–2847, 2008.
- J.Y. Hardeberg. *Acquisition and reproduction of color images: colorimetric and multispectral approaches*. Universal publishers, 2001.
- J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. *Advances in neural information processing systems*, 19:545, 2007.
- D.C. Heinz and C.I. Chang. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(3):529–545, 2001.
- J.P. Hoffbeck and D.A. Landgrebe. Effect of radiance-to-reflectance transformation and atmosphere removal on maximum likelihood classification accuracy of high-dimensional remote sensing data. In *IEEE Geoscience and Remote Sensing Symposium, 1994. IGARSS'94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 4, pages 2538–2540. IEEE, 1994.
- M. Hossny, S. Nahavandi, and D. Creighton. Comments on information measure for performance of image fusion'. *Electronics letters*, 44(18):1066–1067, 2008.
- M. Hossny, S. Nahavandi, D. Creighton, and A. Bhatti. Image fusion performance metric based on mutual information and entropy driven quadtree decomposition. *Electronics letters*, 46(18):1266–1268, 2010.
- X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *Conference on Computer Vision and Pattern Recognition (CVPR07)*. IEEE, 2007.

- X. Hou, J. Harel, and C. Koch. Image signature: Highlighting sparse salient regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
- P.J. Huber. Projection pursuit. *The annals of Statistics*, pages 435–475, 1985.
- A. Hyvärinen. New approximations of differential entropy for independent component analysis and projection pursuit. In *Neural Information Processing Systems*, volume 10, pages 273–279, 1998.
- A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural computation*, 9(7):1483–1492, 1997.
- A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.
- A. Ibrahim, S. Tominaga, and T. Horiuchi. Invariant representation for spectral reflectance images and its application. *EURASIP Journal on Image and Video Processing*, 2011(1):2, 2011.
- A. Ifarraguerri and C.I. Chang. Unsupervised hyperspectral image analysis with projection pursuit. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6):2529–2538, 2000.
- F.H. Imai, M.R. Rosen, and R.S. Berns. Comparative study of metrics for spectral match quality. In *Proceedings of the First European Conference on Colour in Graphics, Imaging and Vision*, pages 492–496, 2002.
- L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.
- N.P. Jacobson and M.R. Gupta. Design goals and solutions for display of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(11):2684–2692, 2005.
- N.P. Jacobson, M.R. Gupta, and J.B. Cole. Linear fusion of image sets for display. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3277–3288, 2007.
- A. Jakulin and I. Bratko. Quantifying and visualizing attribute interactions: An approach based on entropy. *Arxiv preprint cs.AI/0308002*, 2004.
- X. Jia and J. Richards. Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):538–542, 1999.

- LO Jimenez and DA Landgrebe. Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 28(1):39–54, 1998.
- T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *International Conference on Computer Vision*, pages 2106–2113. IEEE, 2009.
- C. Jutten and J. Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24(1):1–10, 1991.
- S. Kaewpijit, J. Le Moigne, and T. El-Ghazawi. Automatic reduction of hyperspectral imagery using wavelet spectral analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4):863–871, 2003.
- C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. *Human neurobiology*, 4(4):219–27, 1985.
- T.O. Kvalseth. Entropy and correlation: some comments. *IEEE Transactions on Systems, Man and Cybernetics*, 17(3):517–519, 1987.
- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Harberg. Visualisation d’images spectrales : une méthode basée sur la perception humaine. In *Proceedings of ORASIS, Praz-sur-Arly, France, June 2011a*.
- Steven Le Moan, Alamin Mansouri, Yvon Voisin, and Jon Yngve Harberg. A constrained band selection method based on information measures for spectral image color visualization. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12):5104–5115, 2011b.
- C.H. Lee, A. Varshney, and D.W. Jacobs. Mesh saliency. In *ACM SIGGRAPH 2005 Papers*, page 666. ACM, 2005.
- J.B. Lee, A.S. Woodyatt, and M. Berman. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing*, 28(3):295–304, 1990.
- M. Lennon, G. Mercier, MC Mouchot, and L. Hubert-Moy. Independent component analysis as a tool for the dimensionality reduction and the representation of hyperspectral images. In *IEEE Geoscience and Remote Sensing Symposium, 2001. IGARSS’01. IEEE 2001 International*, volume 6, pages 2893–2895. IEEE, 2001a.
- M. Lennon, G. Mercier, M.C. Mouchot, and L. Hubert-Moy. Curvilinear component analysis for nonlinear dimensionality reduction of hyperspectral images. *Image and Signal Processing for Remote Sensing VII*, 4541:157–168, 2001b.

- W. Lin and C.C. Jay Kuo. Perceptual visual quality metrics: A survey. *Journal of Visual Communication and Image Representation*, 22(4):297–312, 2011.
- Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganière, and W. Wu. Objective assessment of multiresolution fusion algorithms for context enhancement in night vision: a comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (99):1–1, 2012.
- A. Mansouri, FS Marzani, and P. Gouton. Neural networks in two cascade algorithms for spectral reflectance reconstruction. In *IEEE International Conference on Image Processing*, volume 2, pages 718–721. *f*, 2005.
- A. Martinez-Usó, F. Pla, J.M. Sotoca, and P. Garcia-Sevilla. Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):4158–4171, 2007.
- W.J. McGill. Multivariate information transmission. *Psychometrika*, 19(2):97–116, 1954.
- M. Mignotte. A multiresolution markovian fusion model for the color visualization of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(12):4236–4247, 2010.
- M. Mignotte. A bicriteria-optimization-approach-based dimensionality-reduction model for the color display of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(99):1–13, 2012.
- A. Mohan, G. Sapiro, and E. Bosch. Spatially coherent nonlinear dimensionality reduction and segmentation of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing Letters*, 4(2):206–210, 2007.
- S. Moon and H. Qi. Hybrid dimensionality reduction method based on support vector machine and independent component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, (99):1–1, 2012.
- J.M.P. Nascimento and J.M.B. Dias. Does independent component analysis play a role in unmixing hyperspectral data? *IEEE Transactions on Geoscience and Remote Sensing*, 43(1):175–187, 2005.
- S.M.C. Nascimento, F.P. Ferreira, and D.H. Foster. Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A*, 19(8):1484–1490, 2002.
- L. Parsons, E. Haque, and H. Liu. Subspace clustering for high dimensional data: a review. *ACM SIGKDD Explorations Newsletter*, 6(1):90–105, 2004.

- M. Pedersen and J.Y. Hardeberg. Full-reference image quality metrics: Classification and evaluation. *Foundations and Trends® in Computer Graphics and Vision*, 7(1):1–80, 2011.
- A.M. Qaid and HT Basavarajappa. Application of optimum index factor technique to landsat-7 data for geological mapping of north east of Hajjah, Yemen. *American-Eurasian Journal of Scientific Research*, 3(1):84–91, 2008.
- G. Qu, D. Zhang, and P. Yan. Information measure for performance of image fusion. *Electronics Letters*, 38:313, 2002.
- H. Ren and C.I. Chang. Automatic spectral target recognition in hyperspectral imagery. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1232–1249, 2003.
- N. Renard and S. Bourennane. Dimensionality reduction based on tensor modeling for classification methods. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4):1123–1131, 2009.
- P. Scheunders. Multispectral image fusion using local mapping techniques. In *International Conference on Pattern Recognition*, volume 2, pages 311–314. IEEE, 2000.
- J. Serra. Image analysis and mathematical morphology, vol. 1. *Academic Press, London*, page 600, 1982.
- C.E. Shannon and W. Weaver. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 1948.
- Puneet Sharma, Faouzi Alaya Cheikh, and Jon Yngve Hardeberg. Saliency map for human gaze prediction in images. In *Sixteenth Color Imaging Conference*, Portland, Oregon, USA, Nov 2008.
- C. Sheffield. Selecting band combinations from multispectral data. *Photogrammetric Engineering and Remote Sensing*, 51:681–687, 1985.
- C. Simon, U. Huxhagen, A. Mansouri, A. Heritage, F. Boochs, and F.S. Marzani. Integration of high-resolution spatial and spectral data acquisition systems to provide complementary datasets for cultural heritage applications. In *Proceedings of SPIE*, volume 7531, page 75310L, 2010.
- L.I. Smith. A tutorial on principal components analysis. *Cornell University, USA*, 51:52, 2002.
- M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta. A standard default color space for the internet-srgb. *Microsoft and Hewlett-Packard Joint Report*, 1996.

- M. Studeny and J. Vejnarova. The multiinformation function as a tool for measuring stochastic dependence. *Learning in graphical models*, 261, 1998.
- KC Tiwari, MK Arora, and D. Singh. An assessment of independent component analysis for detection of military targets from hyperspectral images. *International Journal of Applied Earth Observation and Geoinformation*, 13(5):730–740, 2011.
- A. Toet, MA Hogervorst, SG Nikolov, JJ Lewis, TD Dixon, DR Bull, and CN Canagarajah. Towards cognitive image fusion. *Information Fusion*, 11(2):95–113, 2010.
- A.M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.
- V. Tsagaris and V. Anastassopoulos. Information measure for assessing pixel-level fusion methods. In *SPIE Conference Series*, volume 5573, pages 64–71. SPIE, 2004.
- V. Tsagaris and V. Anastassopoulos. Multispectral image fusion for improved rgb representation based on perceptual attributes. *International Journal of Remote Sensing*, 26(15):3241–3254, 2005.
- V. Tsagaris, V. Anastassopoulos, and GA Lampropoulos. Fusion of hyperspectral data using segmented PCT for color representation and classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10):2365–2375, 2005.
- J.S. Tyo, A. Konsolakis, D.I. Diersen, and R.C. Olsen. Principal-components-based display strategy for spectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 41(3):708–718, 2003.
- LJP Van Der Maaten, EO Postma, and HJ Van Den Herik. Dimensionality reduction: A comparative review. *Published online*, 10 (February):1–35, 2007.
- J. Wang and C.I. Chang. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6):1586–1600, 2006.
- P. Wang and B. Liu. A novel image fusion metric based on multi-scale analysis. In *International Conference on Signal Processing*, pages 965–968. IEEE, 2008.
- Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- S. Watanabe. Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development*, 4(1):66–82, 1960.

- T.A. Wilson, S.K. Rogers, and M. Kabrisky. Perceptual-based image fusion for hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):1007–1017, 1997.
- W. Xia, X. Liu, B. Wang, and L. Zhang. Independent component analysis for blind unmixing of hyperspectral imagery with additional constraints. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2165–2179, 2011.
- CS Xydeas and V. Petrovic. Objective image fusion performance measure. *Electronics Letters*, 36(4):308–309, 2000.
- H. Zhang, D.W. Messinger, and E.D. Montag. Perceptual display strategies of hyperspectral imagery based on PCA and ICA. In *Proceedings of SPIE*, volume 6233, page 31. SPIE, 2006.
- H. Zhang, H. Peng, M.D. Fairchild, and E.D. Montag. Hyperspectral image visualization based on a human visual model. In *Proceedings of SPIE*, volume 6806, page 68060N. SPIE, 2008.
- Y. Zhu, P.K. Varshney, and H. Chen. Evaluation of ica based fusion of hyperspectral images for color display. In *International Conference on Information Fusion*, pages 1–7. IEEE, 2007.

— SPIM

■ École doctorale SPIM - Université de Bourgogne/UFR ST BP 47870 F - 21078 Dijon cedex
■ tél. +33 (0)3 80 39 59 10 ■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr

