Saliency-Based Band Selection For Spectral Image Visualization

***Steven Le Moan*[1,2]*, Alamin Mansouri*[1]*, Jon Y. Hardeberg*[2]*, Yvon Voisin*[1]**
1 : ***Le2i - Université de Bourgogne, Auxerre, France***
2 : ***The Norwegian Color Research Laboratory - Gjøvik University College, Gjøvik, Norway***

Abstract

In this paper, we introduce a new band selection approach for the color visualization of spectral images. Unlike traditional methods, we propose to make a selection out of a comparison of the saliency maps of the individual spectral channels. This allows to assess how different they are in terms of prominent features. A comparison metric based on Shannon's information theory at the second and third order is presented and results are subjectively and objectively compared to other dimensionality reduction techniques on three datasets, demonstrating the efficiency of the proposed approach.

Introduction

Spectral imagery consists of acquiring a scene at more than three different ranges of wavelengths, usually dozens. Since spectral display devices are yet rare, most of today's popular display hardware is based on the tri-stimulus paradigm [1]. Thus, in order to visualize spectral images, a dimensionality reduction step is required so that only three channels (Red, Green and Blue for example) can contain most of the visual information while easing interpretation by preserving natural colors and contrasts [2].

Tri-stimulus representation of multi/hyperspectral images for visualization is an active field of research that has been thoroughly investigated over the past decades. One of the most common approaches is probably the one referred to as "true color". It can basically be achieved in two different ways: one consists of selecting the bands at 700nm, 546.1nm and 435.8nm (or the closest) and mapping them to the three primaries: R,G and B, respectively. The other one uses the CMF-based band transformation [3] (each primary R,G and B is the result of a specific linear combination of spectral channels in the visible range of wavelengths). Even though it generally yields a natural visual rendering, this approach does not take the data itself into account at all, and thus noise, redundancy, etc. are not accurately handled.

Another very common approach for dimensionality reduction is Principal Components Analysis (PCA), which has been extensively used for visualization purposes. Tyo *et al.* [4], investigated PCA for N-to-3 dimensionality reduction into the HSV color space. An automatic method to find the origin of the HSV cone is also proposed in order to enhance the final color representation. Later, Tsagaris *et al.* [5] suggested to use the fact that the red, green and blue channels, as they are interpreted by the human eye, contain some correlation, which is in contradiction to the underlying decorrelation engendered by PCA. For that reason, the authors proposed a constrained PCA-based technique in which the eigendecomposition of the correlation matrix is forced with non-zero elements in its non-diagonal elements. Several other PCA-based visualization techniques can be found in the literature [6, 7, 8].

In order to alleviate the computational burden of the traditional PCA, Jia et al. [9] proposed a correlation-based spectrum segmentation technique so that principal components are extracted from different segments and then used for visualization. Other segmented PCA approaches are investigated in [10] including equal subgroups, maximum energy and spectral-signature-based partitioning.

In [11], Du *et al.* compared seven feature extraction techniques in terms of class separability, including PCA, Independent Components Analysis (ICA) and Linear Discriminant Analysis (LDA). ICA has also been studied by Zhu *et al.* [12] for spectral image visualization. They used several spectrum segmentation techniques (equal subgroups, correlation coefficients and RGB-based) to extract the first IC in each segment. The use of different color spaces for mapping of the PCs or ICs has been investigated by Zhang *et al.* [13].

In [2, 14], Jacobson *et al.* presented a band transformation method allowing the CMF to be extended to the whole image spectrum, and not only to the visible part. They proposed a series of criteria to assess the quality of a spectral image visualization. Later, Cui *et al.* [15] proposed to derive the dimensionality reduction problem into a simple convex optimization problem. In their paper, class separability is considered and manipulations on the HSV cone allow for color adjustments on the visualization. More recently, we have proposed a method based on class-separability in the CIELAB space for improved spectral image visualization [16].

All the previously presented approaches are *band transformation* techniques inasmuch as they produce combinations of the original spectral channels to create an enhanced representative triplet. As stated earlier, the often mentioned drawback of this kind of approach is the loss of physical meaning attached to a channel. That is, if, initially, a spectral band is implicitly linked to a range of wavelengths, what can we tell about a combination of them ? A particular case of *band transformation* is called *band selection* and consists of linearly combining the channels while constraining the weighting coefficients in the duet

$\{0, 1\}$. In other words, the resulting triplet is a subset of the original dataset. By doing this, one preserves the underlying physical meaning of the spectral channels, thus allowing for an easier interpretation by the human end user.

In [17], Bajcsy investigated several supervised and unsupervised criteria for band selection, including entropy, spectral derivatives, contrast, etc. Many signal processing techniques have been applied to band selection: Constrained Energy Minimization (CEM) and Linear Constrained Minimum Variance (LCMV) [18], Orthogonal Subspace Projection (OSP) [19, 20] or the One-Bit Transform (1BT) [21]. Also information measures based on Shannon's theory of communication [22] have been proven to be very powerful in the identification of redundancy in high-dimensional datasets. Mutual information was first used for band selection by Conese *et al.* [23]. In [24] and [25], two metrics based on mutual information are introduced in the context of image fusion evaluation. They measure how much information is shared by the original and the reduced datasets. In [26], mutual information is used to measure the similarity of each band with an estimated ground truth. Hence, irrelevant bands for classification purpose are removed. In [27], a normalized mutual information is used for hierarchical spectrum segmentation. However, to the best of our knowledge, never has saliency analysis been used as a means for band-selection-based visualization. It is nonetheless a very relevant approach to evaluate the relative informative content of spectral channels and therefore useful in the context of dimensionality reduction.

Visual attention modeling is the study of the human visual interpretation of a given scene. In other words, which objects/features will first draw attention and why. This notion is closely linked to the analysis of saliency. Following early influential work by Treisman *et al.* [28] and Koch & Ullman [29], Itti *et al.* [30] proposed a general visual attention model allowing for the computation of so-called saliency maps, which purpose is to predict human gaze given a certain scene. This model involves center-surround comparisons and combinations of three main feature channels, namely colors, intensity and orientations. More recent work involve for instance the use of spectral residual analysis [31] or information theory [32].

In this paper, we propose a new strategy for the color display of spectral images. Our contributions are based on two main ideas: making use of saliency maps as a means to compare spectral channels as well as measuring third-order redundancy by means of a generalization of Shannon's mutual information called co-information [33]. Consequently, the following is organized as follows: a first section tackles the band selection algorithm by defining a metric called "Normalized Mutual Saliency" and explaining the band selection algorithm, while a second part presents and discusses the results obtained before the conclusion.

Saliency maps

From the literature, one can find several ways of computing a saliency map from a color image but one of the most influential work is the model by Itti *et al.* [30]. It is also one of the simplest method and this is why we have decided to use it in this study. This model is based on the extraction of so-called conspicuity maps, depicting the prominence of every single pixel in terms of three different features: color, intensity and orientation (the latter being analyzed through 4 different angles). In the case of spectral channels, not only color is not involved, but, since each channel describes the same scene, the analysis of orientation conspicuity doesn't require more than one channel to be properly achieved. In the end, this step simply requires the computation of $N + 4$ conspicuity maps for the obtention of N saliency maps (N being the number of spectral channels), each one of them representing the informative content of an individual spectral channel. By depicting the locations of strong center-surround differences, the thusly obtained saliency maps are a powerful means to compare spectral channels and are consequently very suitable for band selection.

Channels comparison

One then has several possibilities to compare saliency maps, the most simple being a summation of pixelwise euclidean distances. In this study, we have focused on the use of information measures for they allow for a statistical comparison of random variables (population of pixels) while taking into account the relative spatial locations of pixels. Thus, we introduce a metric simply derived from Shannon's mutual information that we will refer to as "Normalized Mutual Saliency" (NMS) and which is defined as follows:

$$NMS(im_1; im_2) = \frac{I(s(im_1); s(im_2))}{H(s(im_1) + s(im_2))}$$

with im_1 and im_2 any two images of same spatial dimensions, $s(.)$ an operator computing the saliency map of its input and $H(.)$ and $I(.;.)$ being respectively the entropy and mutual information operators. The normalization is indeed necessary to allow for the metric to be non-relative, as it has been suggested for instance in [27].

Since we need three channels to create the final color composite, we also define the third order NMS, based on the Co-Information, as defined by Bell [33]:

$$CI(X; Y; Z) = H(X) + H(Y) + H(Z) - H(X; Y) - H(X; Z) - H(Y; Z) + H(X; Y; Z)$$

then:

$$NMS(im_1; im_2; im_3) = \frac{|CI(s(im_1); s(im_2); s(im_3))|}{H(s(im_1) + s(im_2) + s(im_3))}$$

Other generalizations of mutual information to higher orders have been proposed in the literature such as Watanabe's *total correlation* [34] which is defined as the difference between the sum of marginal entropies and the joint entropy of the set. However, the main drawback of total correlation is that it measures both second and third order, indiscriminately, while giving more weight to the second order. McGill [35] presented the *interaction information*, which is basically the same as co-information, simply with an opposite sign.

A particularly interesting property of co-information is that it can take both positive and negative values. In the "positive" case, one talks about redundancy, whereas in the case of negative values, one talks about synergy. Redundancies are foreseeable from lower orders while synergies only appear when the set of random variables are taken together. The synergy case appears 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, we consider a simple XOR 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 $I(X;Y|Z) > I(X;Y)$. In the case of spectral images, this principle remains 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, co-information must be as close to zero as possible in order to minimize the superfluous information.

Band selection

The band selection is performed by first finding the most dissimilar couple of channels. Instead of an exhaustive an computationally costly search, we propose to use an algorithm similar to the one used in [20]. A first B_1 channel is selected randomly and the one from which it is the most dissimilar (B_2) is sought among the others. The same procedure is used on B_2 to find B_3, and so on until $B_i = B_{i-2}$. Algorithm 1 describes the procedure for an $N-$bands spectral image.

Algorithm 1 Band selection

```
i = 0; k = 1; iterations = 0;
randomly choose j ∈ [1..N];
while (i != k) and (iterations < 20) do
    find temp = argmin_k[NMS(B_j; B_k)]
    i ← j; j ← k; k ← temp;
    iterations++;
end while
find k = argmin |NSM(B_i; B_j; B_k)|
{R, G, B} ← sort({B_i, B_j, B_k}) by desc. wavelength
```

The maximum number of iteration should be set accordingly to N. For instance, for a 31-bands image, we have assumed that the algorithm can converge within 20 iterations.

Experiments

Datasets

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 3^{rd} century sarcophagus from the St Matthias abbey in Trier, Germany [36]. It was acquired by means of a 8 channels filter wheel camera ranging only in the visible spectrum (400-740nm). Reflectance was reconstructed by means of a supervised neural-network-based algorithm.
- "Mural" is a 35 bands (400-740nm) multispectral image of a 16^{th} century mural painting from the Brömser Hof in Rudesheim, Germany, acquired with a rotating-wheel-based multispectral camera.

As a pre-processing step, bands with average reflectance value below 2% and those with low correlation (below 0.8) with their neighboring bands have been removed, as suggested in [37].

The main reason why we have chosen these scenes is because they all include a MacBeth CC target which facilitates the evaluation of color rendering.

Benchmarking methods

In order to evaluate the performances of our method, we have selected two other dimensionality reduction techniques for comparison.

- PCA_{hsv} is the traditional Principal Components Analysis of which components are mapped to the HSV color space ($PC1 \to V; PC2 \to S; PC3 \to H$).
- LP-based band selection has been proposed by Du *et al.* [20] and consists of progressively selecting bands by maximizing their respective orthogonality. Due to the high memory requirements of this method, a spatial subsampling of the data is necessary. According to Du *et al.*, the subsampling rate can be chosen as high as 1:100 (only 1% of the pixels are kept) without affecting the results. This rate has been applied in this study.

Results

Figure 1 depicts the resulting color visualization of all the images and for the three dimensionality reduction approaches: PCA, LP and NMS.

It can be seen that the PCA-based method gives the least appealing results, while LP and NMS give quite similar and "eye-satisfying" images. On the first dataset, one can notice that the white patch (bottom left) if whiter in the LP result, but still very discriminable from all the others in the NMS-based band selection. 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 L*a*b* values of a set

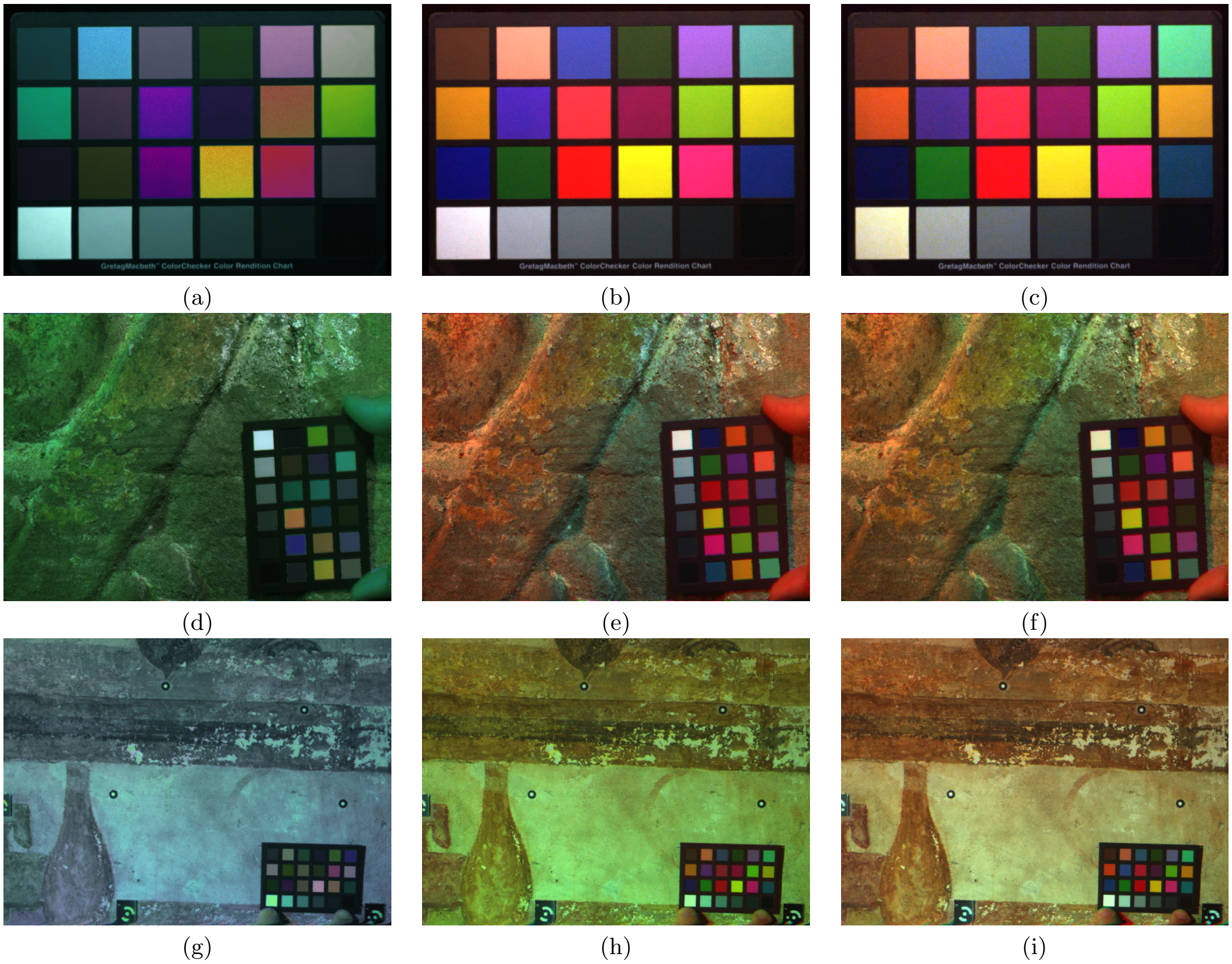

Figure 1. *Different representations for each dataset (first column: PCA, second: LP and third: NMS)*

of 480 manually selected pixels (20 by patch) with the "ground truth" ones, provided by *Gretag*, by means of the CIE76 $\Delta E_{ab}*$ 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 1 gives the minimal, maximal and average perceptual distances in L*a*b* between the results and the "ground truth". It can be seen that, even though the LP-based band selection gives slightly better minimal and maximal errors in two cases, the proposed approach outperforms it on each dataset, in terms of average ΔE, and especially on the two last images.

		PCA_{hsv}	LP	NMS
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**

Colorimetric errors

Conclusion

We proposed a new band selection method based on saliency maps for the dimensionality reduction of spectral image. A simple metric based on information theory and called Normalized Mutual Saliency has been introduced and used as a means to compare spectral channels. Both second and third order versions of this metric have been considered. Results on three different images have been subjectively and objectively assessed, proving the efficiency of the proposed method.

Acknowledgments

The authors would like to thank the Regional Council of Burgundy for supporting this work, as well as the i3Mainz lab in Mainz and the Institut Für Steinkonservierung for providing useful data.

References

[1] H. Grassmann, "On the theory of compound colors," *Phil. Mag*, vol. 7, pp. 254–64, 1854.

[2] N.P. Jacobson and M.R. Gupta, "Design goals and solutions for display of hyperspectral images," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 11, pp. 2684–2692, 2005.

[3] G. Poldera and G.W.A.M. van der Heijdena, "Visualization of spectral images," in *Proc. SPIE*, 2001, vol. 4553, p. 133.

[4] J.S. Tyo, A. Konsolakis, D.I. Diersen, and R.C. Olsen, "Principal-components-based display strategy for spectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, no. 3, pp. 708–718, 2003.

[5] V. Tsagaris and V. Anastassopoulos, "Multispectral image fusion for improved rgb representation based on perceptual attributes," *International Journal of Remote Sensing*, vol. 26, no. 15, pp. 3241–3254, 2005.

[6] JM Durand and YH Kerr, "An improved decorrelation method for the efficient display of multispectral data," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 27, no. 5, pp. 611–619, 1989.

[7] P. Scheunders, "Multispectral image fusion using local mapping techniques," in *International conference on pattern recognition*, 2000, vol. 15, pp. 311–314.

[8] C. Yang, L. Lu, H. Lin, R. Guan, X. Shi, and Y. Liang, "A fuzzy-statistics-based principal component analysis (fs-pca) method for multispectral image enhancement and display," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3937–3947, 2008.

[9] X. Jia and J. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 538–542, 1999.

[10] V. Tsagaris, V. Anastassopoulos, and GA Lampropoulos, "Fusion of hyperspectral data using segmented pct for color representation and classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2365–2375, 2005.

[11] Q. Du, N. Raksuntorn, S. Cai, and R.J. Moorhead, "Color display for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 46, pp. 1858–1866, 2008.

[12] Y. Zhu, P.K. Varshney, and H. Chen, "Evaluation of ica based fusion of hyperspectral images for color display," in *Information Fusion, 2007 10th International Conference on*, 2007, pp. 1–7.

[13] H. Zhang, D.W. Messinger, and E.D. Montag, "Perceptual display strategies of hyperspectral imagery based on pca and ica," in *Proceedings of SPIE*, 2006.

[14] N.P. Jacobson, M.R. Gupta, and J.B. Cole, "Linear fusion of image sets for display," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3277–3288, 2007.

[15] M. Cui, A. Razdan, J. Hu, and P. Wonka, "Interactive hyperspectral image visualization using convex optimization," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 6, pp. 1673, 2009.

[16] S. Le Moan, A. Mansouri, J.Y. Hardeberg, and Y. Voisin, "A class-separability-based method for multi/hyperspectral image color visualization," in *IEEE International Conference on Image Processing*, 2010.

[17] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogrammetric engineering and remote sensing*, vol. 70, pp. 793–802, 2004.

[18] C.I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1575–1585, 2006.

[19] 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 Trans. on Geoscience and Remote Sensing*, vol. 37, no. 6, pp. 2631–2641, 1999.

[20] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 564–568, 2008.

[21] 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 Trans. on Geoscience and Remote Sensing*, vol. 47, no. 1 Part 1, pp. 97–105, 2009.

[22] C.E. Shannon and W. Weaver, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[23] C. Conese and F. Maselli, "Selection of optimum bands from tm scenes through mutual information analysis," *ISPRS journal of photogrammetry and remote sensing*, vol. 48, no. 3, pp. 2–11, 1993.

[24] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, pp. 313, 2002.

[25] V. Tsagaris and V. Anastassopoulos, "An information measure for assessing pixel-level fusion methods," *Image and Signal Processing for Remote Sensing*, 2004.

[26] B. Guo, RI Damper, S.R. Gunn, and JDB Nelson, "A fast separability-based feature-selection method for high-dimensional remotely sensed image classification," *Pattern Recognition*, vol. 41, no. 5, pp. 1670–1679, 2008.

[27] A. Martinez-Uso, F. Pla, J.M. Sotoca, and P. Garcia-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4158–4171, 2007.

[28] A.M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive psychology*, vol. 12, no. 1, pp. 97–136, 1980.

[29] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry.," *Hum Neurobiol*, vol. 4, no. 4, pp. 219–27, 1985.

[30] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*,

vol. 20, no. 11, pp. 1254–1259, 1998.

[31] X. Hou and L. Zhang, “Saliency detection: A spectral residual approach,” in *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*. Ieee, 2007, pp. 1–8.

[32] N.D.B. Bruce and J.K. Tsotsos, “Saliency, attention, and visual search: An information theoretic approach,” *Journal of Vision*, vol. 9, no. 3, 2009.

[33] A.J. Bell, “The co-information lattice,” in *Proceedings of the Fifth International Workshop on Independent Component Analysis and Blind Signal Separation: ICA 2003*, 2003.

[34] S. Watanabe, “Information theoretical analysis of multivariate correlation,” *IBM Journal of research and development*, vol. 4, no. 1, pp. 66–82, 1960.

[35] W.J. McGill, “Multivariate information transmission,” *Psychometrika*, vol. 19, no. 2, pp. 97–116, 1954.

[36] 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*, 2010, vol. 7531, p. 75310L.

[37] S. Cai, Q. Du, and RJ Moorhead, “Hyperspectral imagery visualization using double layers,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3028–3036, 2007.

.

Author Biography

Steven Le Moan *is a PhD student at the Le2i, Université de Bourgogne, France. He received his Master’s degree in Signal Processing and Embedded Systems from the University of Rennes 1 (2009). He is currently working in cooperation with the ColorLab in Gjøvik, Norway in the fields of multivariate analysis of multispectral images for visualization. His technical interests extends also to the study of 3D meshes with multispectral attributes.*

Alamin Mansouri *received his PhD degree in computer vision (particularly in multispectral imaging) from the University of Burgundy, France, in december 2005. Since september 2006 he is associate professor at the same university. His research is focused on color/multispectral image processing and analysis and their applications.*

Jon Y. Hardeberg *is Professor of Color Imaging at Gjøvik University College. He received his Ph.D from Ecole Nationale Supérieure des Télécommunications in Paris, France in 1999 with a dissertation on colour image acquisition and reproduction, using both colorimetric and multispectral approaches. He has more than 10 years experience with industrial and academic colour imaging research and development, and has co-authored over 100 research papers within the field. His research interests include various topics of colour imaging science and technology, such as device characterisation, gamut visualization and mapping, image quality, and multispectral image acquisition and reproduction. He is a member of IS&T, SPIE, and the Norwegian representative to CIE Division 8. He has been with Gjøvik University College since 2001 and is currently head of the Norwegian Color Research Laboratory.*

Yvon Voisin *is a full professor of signal and image processing at the University of Burgundy and is a member of the Image Processing Group at the Le2i. His research interests are 3D reconstruction and motion analysis. He’s also working on the application of artificial vision, especially in biology. Voisin has a PhD in electronic and signal processing from the University of Franche-Comté, France.*