KERNEL STRUCTURAL SIMILARITY ON HYPERSPECTRAL IMAGES

Vicent Talens, Valero Laparra, Jesús Malo and Gustavo Camps-Valls

Image Processing Laboratory (IPL). Universitat de València, Spain.
{vicent.talens, valero.laparra, jesus.malo, gustavo.camps}@uv.es, http://isp.uv.es

ABSTRACT

In this paper, we introduce a non-linear and multidimensional generalization of the Structural SIMilarity index (SSIM) for quality assessment of hyperspectral images. We exploit well-known properties of functional analysis and estimate means, variances, and correlation in proper reproducing kernel Hilbert spaces (rkHs). The so-called Kernel SSIM (KSSIM) is shown to generalize the conventional SSIM and the recently introduced Q_4 and Q_n metrics for remote sensing applications, and naturally works with multidimensional images. For the experimentation, we built a database of different distortions commonly encountered in remote sensing images. KSSIM shows an improved agreement with classification results compared to standard similarity metrics, and high consistency for different noise sources and levels.

Index Terms— Image quality assessment, SSIM, metric, kernel methods.

1. INTRODUCTION

Assessing similarity of an image to a known reference is an active research field [1]. Many applications, such as denoising, coding or multiresolution fusion, need objective measures to evaluate image quality [2]. In general, the standard root-mean-square-error (RMSE) is not appropriate [3], and perceptually meaningful distortion metrics, such as the Structural SIMilarity Index (SSIM) [4] or Visual Information Fidelity (VIF) [5], have become new standards.

The previous measures, however, are designed to fulfill perceptual criteria of similarity and are only applicable to one band (grayscale) images. For color images or video evaluation, SSIM or VIF are applied in each channel and the results are averaged. This, however, does not consider the correlation between channels and frames, respectively, and *ad hoc* procedures have been introduced in the literature. When confronted with multi- and hyperspectral images, a common choice is the *relative dimensionless global error in synthesis* (ERGAS) [6], which is a normalized version of RMSE designed to calculate the spectral distortion. The Q_4 score presented in [7] extends SSIM to the 4-bands case typically encountered in quality assessment of pansharpened multispectral images. In all cases, however, the metrics are not general enough to deal with images of arbitrarily large dimension and nonlinear changes and distortions.

Kernel methods provide a solid theoretical framework to estimate multidimensional data similarities [8]. However, few attention has been payed in the literature to assess image quality with kernels. The Hilbert-Schmidt Independence Criterion (HSIC) is a kernel estimate of statistical dependence between multidimensional random variables, which was exploited in [9] to assess the quality of multiresolution image fusion products. Good results were in general obtained, but the measure was too sensitive to parameter tuning and coping with different distortions simultaneously was difficult. In this paper, we present a kernel version of the familiar SSIM for assessing hyperspectral image quality. The Kernel SSIM (KSSIM) is shown to generalize the conventional SSIM and the Q_4 [7] and Q_n [10] metrics recently introduced for remote sensing applications.

2. STRUCTURAL SIMILARITY IN REPRODUCING KERNEL HILBERT SPACES

The SSIM between two N-sample random variables, $\mathbf{x}$ and $\mathbf{y}$, takes into account three different terms, that stand for *luminance*, *contrast* and *structure*: $\text{SSIM} := l(\mathbf{x},\mathbf{y})^\alpha\ c(\mathbf{x},\mathbf{y})^\beta\ s(\mathbf{x},\mathbf{y})^\gamma$. Computing these terms essentially boils down to estimate the (local) sample means ($\mu_\mathbf{x}$, $\mu_\mathbf{y}$), variances ($\sigma^2_\mathbf{x}$, $\sigma^2_\mathbf{y}$) and correlation ($\sigma_\mathbf{xy}$) terms. A common strategy sets $\alpha=\beta=\gamma=1$, so:

$$\text{SSIM} := \frac{2\mu_\mathbf{x}\mu_\mathbf{y}+c_1}{\mu^2_\mathbf{x}+\mu^2_\mathbf{y}+c_1}\cdot\frac{2\sigma_\mathbf{xy}+c_2}{\sigma^2_\mathbf{x}+\sigma^2_\mathbf{y}+c_2}, \tag{1}$$

where $c_1 = 0.01D$, $c_2 = 0.03D$, and D stands for the dynamic range of the signal. SSIM is typically computed locally following a sliding-window approach. Hereafter, we omit patch indices for the sake of clarity.

2.1. Kernel SSIM (KSSIM)

Kernelization of SSIM requires the estimation of the means, variances and correlation in reproducing kernel Hilbert spaces. This will permit the KSSIM to work with higher-order statistical dependencies and to deal with hyperspectral images of any dimension.

This work was partially supported by the spanish Ministry of Economy and Competitiveness (MINECO) under project TIN2012-38102-C03-01.

Let us first define two feature mappings to corresponding Hilbert spaces $\phi : \mathbf{x}_i \rightarrow \phi(\mathbf{x}_i) \in \mathcal{H}$ and $\psi : \mathbf{y}_i \rightarrow \psi(\mathbf{y}_i) \in \mathcal{F}$, and the three following reproducing kernel functions $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}_j)$, $l(\mathbf{y}_i, \mathbf{y}_j) = \psi(\mathbf{y}_i)^\top \psi(\mathbf{y}_j)$, and $m(\mathbf{x}_i, \mathbf{y}_j) = \phi(\mathbf{x}_i)^\top \psi(\mathbf{y}_j)$. Note that the last kernel $m(\cdot, \cdot)$ exists only for mappings to spaces of the same dimensionality. We then define the KSSIM as a function of the means ($\mu_\mathcal{H}$, $\mu_\mathcal{F}$), variances ($\sigma^2_\mathcal{H}$, $\sigma^2_\mathcal{F}$) and correlation $\sigma_{\mathcal{HF}}$ in the corresponding Hilbert spaces:

$$\text{KSSIM} := \frac{2\mu_\mathcal{H}^\top \mu_\mathcal{F} + C_1}{\mu_\mathcal{H}^\top \mu_\mathcal{H} + \mu_\mathcal{F}^\top \mu_\mathcal{F} + C_1} \cdot \frac{2\sigma_{\mathcal{HF}} + C_2}{\sigma^2_\mathcal{H} + \sigma^2_\mathcal{F} + C_2} \tag{2}$$

The means of finite samples in feature space are defined trivially as $\mu_\mathcal{H} = \frac{1}{N}\sum_{i=1}^N \phi(\mathbf{x}_i)$ and $\mu_\mathcal{F} = \frac{1}{N}\sum_{i=1}^N \psi(\mathbf{y}_i)$. Hence the associated scalar products of the means are given by $\mu_\mathcal{H}^\top \mu_\mathcal{H} = \frac{1}{N^2}\text{Tr}(\mathbf{KI})$, $\mu_\mathcal{F}^\top \mu_\mathcal{F} = \frac{1}{N^2}\text{Tr}(\mathbf{LI})$, and $\mu_\mathcal{H}^\top \mu_\mathcal{F} = \frac{1}{N^2}\text{Tr}(\mathbf{MI})$. Similarly, the variances are given by $\sigma^2_\mathcal{H} = \frac{1}{N}\text{Tr}(\mathbf{K}) - \frac{1}{N^2}\text{Tr}(\mathbf{KI})$ and $\sigma^2_\mathcal{F} = \frac{1}{N}\text{Tr}(\mathbf{L}) - \frac{1}{N^2}\text{Tr}(\mathbf{LI})$. The correlation is finally given by $\sigma_{\mathcal{H},\mathcal{F}} = \frac{1}{N}\text{Tr}(\mathbf{M}) - \frac{1}{N^2}\text{Tr}(\mathbf{MI})$. The operation $\text{Tr}(\cdot)$ computes the trace and $\mathbf{I}$ is a unit matrix, i.e. a $(N \times N)$ matrix of ones. Therefore, all the quantities involved in KSSIM can be estimated without even knowing explicitly the mappings, but just their dot products computed implicitly with kernel functions. In kernel terms, the KSSIM simply reduces to:

$$\begin{aligned}\text{KSSIM} := \quad & \frac{4\,\text{Tr}(\mathbf{MI}) + C_1 N^2}{\text{Tr}((\mathbf{K}+\mathbf{L})\mathbf{I}) + C_1 N^2} \\ & \cdot \frac{N\,\text{Tr}(\mathbf{M}) - \text{Tr}(\mathbf{MI}) + C_2 N}{N\,\text{Tr}(\mathbf{K}+\mathbf{L}) - \text{Tr}((\mathbf{K}+\mathbf{L})\mathbf{I}) + C_2 N},\end{aligned} \tag{3}$$

where $\mathbf{K}$, $\mathbf{M}$ and $\mathbf{L}$ are the kernel matrices obtained using the corresponding kernel functions k, m and l.

2.2. Properties of KSSIM

We next show some properties of the introduced KSSIM.

Property 1 Kernel SSIM generalizes SSIM and Q_n in rkHs.

Proof Note that SSIM is obtained as a particular case of KSSIM if we use linear kernel functions, $k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^\top \mathbf{x}_j$, $l(\mathbf{y}_i, \mathbf{y}_j) = \mathbf{y}_i^\top \mathbf{y}_j$, $l(\mathbf{x}_i, \mathbf{y}_j) = \mathbf{x}_i^\top \mathbf{y}_j$, and set $C_1 = c_1$, $C_2 = c_2$ and $C_3 = c_3$. Kernel matrices then become $\mathbf{K} = \mathbf{X}\mathbf{X}^\top$, $\mathbf{L} = \mathbf{Y}\mathbf{Y}^\top$, and $\mathbf{M} = \mathbf{X}\mathbf{Y}^\top$, whose traces are equal to those of the corresponding (cross-)covariances in input space. Then, one trivially obtains (1). The same proof holds for showing the KSSIM generalizes Q_n, but now computing the kernels with d-band images.

Property 2 KSSIM is bounded.

Proof The Mercer's theorem states that for any continuous symmetric kernel function k with positive integral operator, one can expand $k(\mathbf{x}, \mathbf{z})$ in a uniformly convergent series, $k(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^\infty \phi_i(\mathbf{x})\phi_i(\mathbf{z})$ and, as a consequence, the series $\sum_{i=0}^\infty \|\phi_i\|^2$ is convergent. Therefore the KSSIM is bounded.

The values of an RBF kernel $\mathbf{K}$ for example are upper bounded to 1, so the maximum value of $\text{Tr}(\mathbf{K})$ is N, and the maximum value of $\text{Tr}(\mathbf{KI})$ is N^2. Setting $C_i \ll 1$, $i = 1, \ldots, 3$ the maximum value of KSSIM is 1.

Property 3 KSSIM has a unique maximum

Proof The function $\text{KSSIM}(\mathbf{x}, \mathbf{y}) = 1$ iff $\mathbf{x} = \mathbf{y}$. This can be trivially shown by considering that the kernel functions are symmetric positive definite measures and kernel m only exists if ϕ and ψ actually map to the same rkHs.

Property 4 KSSIM is invariant under isometries for shift-invariant kernels.

Proof When shift-invariant kernels are used, such as the Gaussian or Laplacian kernels, $M(\mathbf{x}, \mathbf{y}) = M(\mathbf{x} - \mathbf{y})$, the function $\text{KSSIM}(\mathbf{x}, \mathbf{y})$ is invariant. If $\mathbf{V}$ is an orthogonal or unitary matrix, then $M(\mathbf{V}\mathbf{x}, \mathbf{V}\mathbf{y}) = M(\mathbf{V}(\mathbf{x} - \mathbf{y})) = M'(\mathbf{x} - \mathbf{y})$, where M' is a p.d. kernel. For example, the RBF kernel $M(\mathbf{V}\mathbf{x}, \mathbf{V}\mathbf{y}) = \exp(-\|\mathbf{V}\mathbf{x} - \mathbf{V}\mathbf{y}\|^2/(2\sigma^2)) = M'(\mathbf{x} - \mathbf{y})$, where kernel M' is an RBF kernel with $\sigma' = \sigma/\|\mathbf{V}\|$.

3. EXPERIMENTAL RESULTS

This section is devoted to the analysis of the capabilities of KSSIM in the evaluation of hyperspectral images under different distortion sources and levels.

3.1. Implementation

Note that the proposed method involves three different kernels ($\mathbf{K}$, $\mathbf{L}$ and $\mathbf{M}$) whose parameter must be tuned. In our experiments we used the Radial Basis Function (RBF) kernel, $k(\mathbf{z}_i, \mathbf{z}_j) = \exp\left(-\|\mathbf{z}_i - \mathbf{z}_j\|^2/2\sigma^2\right)$, $\sigma \in \mathbb{R}^+$ because of its robustness and stability. We used the same σ for the three kernels. Kernel parameters are adjusted to maximize an objective or subjective criterion of image quality. This constitutes a difficult *unsupervised* learning problem unless a database of opinion scores is available, which is not the case for hyperspectral images. In our experiments, we will tune the σ in an indirect way: maximizing the correlation of KSSIM with the classification accuracy of support vector machines (SVM) in images distorted with different noise sources and amounts.

The calculation of KSSIM is very easy as it only involves matrix multiplications (cf. Eq. (3)). However, as for any kernel method, a clear computational problem is observed: KSSIM deals with larger matrices than in the conventional SSIM which need to be computed many times (one *per* pixel in the image). We here propose a simple subsampling strategy. Now the question is how many patches are needed to achieve a consistent estimate. This will obviously depend on the image resolution, number of bands, and noise source and amount.

Figure 1 shows the alignment between KSSIM and the accuracy of classification results of Sec. 3.2 for different σ values and subsampling rates. It is clear that performance is

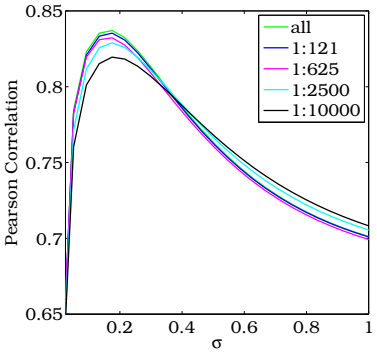

Fig. 1. Agreement between KSSIM and the classification accuracy as a function of σ parameter in the RBF kernels used using different subsampling rates.

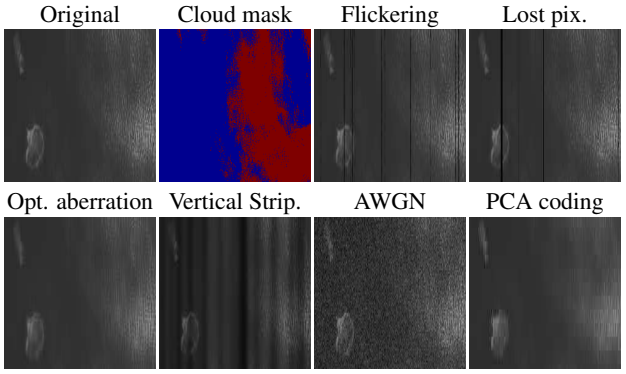

Fig. 2. Examples of the hyperspectral images in the database. Just one band is shown for illustration proposes.

not greatly affected by strong subsampling, and that shows a smooth variation with the sigma value. Therefore, for computational convenience, we will evaluate KSSIM for 1 out of 121 pixels.

3.2. KSSIM assessment in hyperspectral images

A fundamental problem in presenting a multidimensional quality measure has to do with the lack of databases of hyperspectral images with associated quality scores. To circumvent this problem, we generated such a database with common distortions in hyperspectral images: we used 4 MERIS multispectral images and included 6 types of distortions with 5 different amounts. All scenes correspond to MERIS full spatial resolution (FR) images with a pixel size of 260 m across track and 290 m along track and an image size of 321×489 pixels in scenes over Spain, Finland and France.

We included the following distortions: 1) flickering pixels, 2) lost pixels distortion, 3) optical image aberration, 4) vertical striping, 5) additive white Gaussian noise (AWGN) and 6) spatial-spectral distortion produced by image compression. Distortions 1)-5) are related to the acquisition process, while distortion 6) is related to the transmission noise which was here modeled as a simple PCA truncation error. Hence, we generated a database of $4 \times 5 \times 6 = 120$ synthetically distorted MERIS images. See Fig. 2 for illustrative zoomed patches with the considered distortions, along with the original image and a classification (cloud) mask used for evaluation.

We evaluated the performance of the proposed metric and other common measures in the literature in the generated database of distorted hyperspectral images. Essentially, we measure the alignment of the different metrics with the classification accuracy (estimated with the Cohen's kappa statistic) of a support vector machine (SVM) with an RBF kernel trained for cloud classification in every scene. The cloud masking ground truth for classification has been obtained from [11]. Using classification as an indirect measure of image quality has been extensively used in the literature [10,12]. We used 30 images to select the parameter σ in the KSSIM kernels, and tuned the SVM parameters (C, σ) through cross-validation on a trained set formed by 500 randomly picked pixels.

Figure 3 shows the scatter plots and the Pearson's correlation coefficient of different approaches: proposed KSSIM (both with RBF and linear kernel), Q_4 [7], averaged SSIM [4], and ERGAS [7]. Several conclusions can be obtained. First, it is observed that the KSSIM outperforms the rest of the metrics. Second, very poor results are obtained with Q_4, which uses only four bands and is clearly not enough to tackle distortions introduced in all channels. Third, ERGAS also reveals a skewed scatter plot, mainly in non-Gaussian noise sources as expected. Finally, KSSIM with a linear kernel performs similar to the averaged SSIM (in overall Pearson correlation). Finally, note that, even though KSSIM was optimized to treat all kinds of distortions simultaneously, results are also consistent throughout different levels and types of distortions. Table 1 shows the correlation coefficient for each distortion where a consistent gain is obtained by KSSIM except for the case of Gaussian noise. The sensitivity of KSSIM to particular noise sources will be studied in the future.

Table 1. Pearson's correlation *per* distortion.

Distortion	**SSIM**	**KSSIM**
Flickering	0.75	0.84
Lost pixels	0.80	0.86
Opt. aberration	0.05	0.11
Vertical striping	0.73	0.80
AWGN	0.95	0.89
PCA comp.	0.47	0.81
Total	0.79	0.84

3.3. Robustness to noise in high dimensionality images

KSSIM can handle multidimensional datasets in a natural way as it only involves stacking the features before computing the kernels. We use KSSIM to assess the similarity between the classical 220-bands AVIRIS image taken over

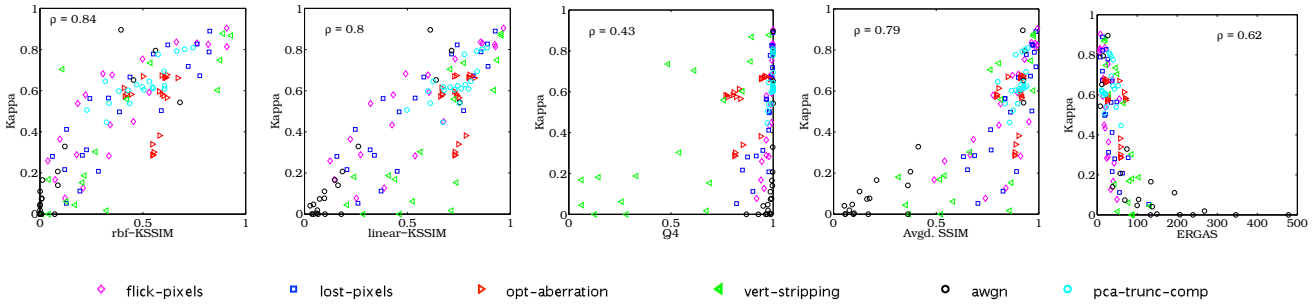

Fig. 3. Scatter plots of SVM Kappa statistic versus different model predictions.

Indiana's Indian Pine and its corrupted version by adding different amounts of Gaussian noise. Figure 4 shows the results obtained by the standard SSIM (applied band-by-band and then averaged) and the KSSIM (just one multidimensional computation is needed), along with their corresponding distortion maps for the range of 0 to 20 dB. The KSSIM values image match the qualitative visual inspection, while the averaged SSIM saturates for low-noise regimes.

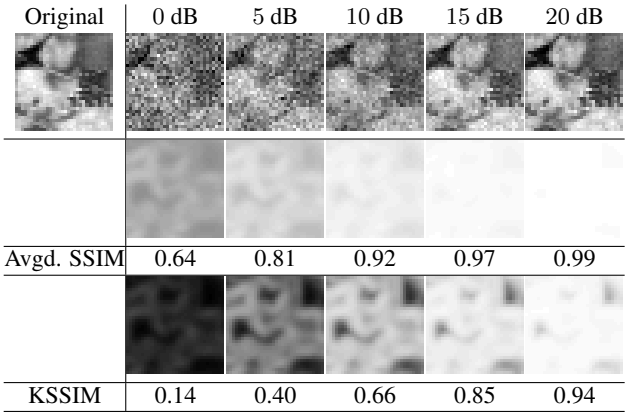

Fig. 4. Example of similarity assessment between a hyperspectral image and its noisy versions. Band 10 (@478.57 nm) is plotted for illustration purposes. From top row to bottom rows: noisy version, SSIM and KSSIM similarity maps (between 0 -black- and 1 -white) and values for the whole image under different amounts of Gaussian noise.

4. CONCLUSIONS

This paper presented a kernel-based generalization of the familiar SSIM in order to estimate similarity between multidimensional images. The method has very good theoretical and practical properties. We illustrated the performance in hyperspectral images and in synthetic products. Ongoing work is related to enlarge the hyperspectral image database for improved evaluation, and the quality assessment of pansharpened products and video sequences.

5. REFERENCES

[1] D. Chandler, "Seven challenges in image quality assessment: Past, present, and future research," *ISRN Signal Processing*, vol. 905685, 2013.

[2] Z. Wang and A. C. Bovik, Eds., *Modern Image Quality Assessment, Synthesis Lectures on Image, Video, and Multimedia Processing*. CO, USA: Morgan & Claypool Publishers, 2006.

[3] V. Laparra, J. Muñoz-Marí, and J. Malo, "Divisive Normalization Image Quality Metric Revisited," *Journal of the Optical Society of America A*, vol. 27, 2010.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Im. Proc.*, vol. 13, no. 4, pp. 600–612, Apr 2004.

[5] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Im. Proc.*, vol. 15, no. 2, pp. 430–444, Feb 2006.

[6] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Fusion of Earth data: merging point measurements, raster maps and remotely sensed images*, Sophia Antipolis, France, 2000, pp. 99–103.

[7] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Rem. Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct 2004.

[8] B. Schölkopf and A. Smola, *Learning with Kernels – SVM, Regularization, Optimization and Beyond*. MIT Press, 2002.

[9] V. Talens, J. Moreno, and G. Camps-Valls, "Kernel image similarity criterion," in *Geosci. Rem. Sens. Symp. (IGARSS), 2011 IEEE Intl.*, Jul 2011, pp. 527–530.

[10] U. Weidner, "Pansharpening relating quantitative quality measures to impact on results of subsequent processing steps," in *ISPRS TC VII Symposium 100 Years ISPRS*, Viena, Austria, 2010, pp. Vol. XXXVIII, Part 7A.

[11] L. Gómez-Chova, G. Camps-Valls, J. Calpe-Maravilla, L. Guanter, and J. Moreno, "Cloud-screening algorithm for ENVISAT/MERIS multispectral images," *IEEE Trans. Geosci. Rem. Sens.*, vol. 45, no. 12-2, pp. 4105–4118, 2007.

[12] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Newark, NJ: Wiley, 2005.