

A Note on the Kernel Spectral Angle Mapper

G. Camps-Valls

This communication introduces a very simple generalization of the familiar spectral angle mapper (SAM) distance. SAM is perhaps the most widely used distance in chemometrics, hyperspectral imaging, and remote sensing applications. We show that a nonlinear version of SAM can be readily obtained by measuring the angle between pairs of vectors in a reproducing kernel Hilbert spaces. The kernel SAM generalizes the angle measure to higher-order statistics, it is a valid reproducing kernel, it is universal, and it has consistent geometrical properties that permit deriving a metric easily. We illustrate its performance in a target detection problem using very high resolution imagery. Excellent results and insensitivity to parameter tuning over competing methods make it a valuable choice for many applications.

Introduction: The Spectral Angle Mapper (SAM) is a method for directly comparing image spectra [1]. Since its seminal introduction, SAM is widely used in chemometrics, astrophysics, imaging, hyperspectral image analysis, industrial engineering, and computer vision applications [2–5], just to name a few. The measure achieved widespread popularity thanks to its implementation in software packages, such as ENVI or ArcGIS. The reason is simple: SAM is invariant to (unknown) multiplicative scalings of spectra due to differences in illumination and angular orientation [6]. SAM is widely used for fast spectral classification, as well as to evaluate the quality of extracted endmembers, and the spectral quality of an image compression algorithm.

The popularity of SAM is mainly due to its simplicity and geometrical interpretability. However, the main limitation of the measure is that it only considers second-order angle dependences between spectra. This paper generalizes the SAM measure to the nonlinear case by means of kernels [7]. This short note aims to characterize KSAM theoretically, and to show its practical convenience over the widespread linear counterpart. The kernel SAM (KSAM) is very easy to compute, and can be used in any kernel machine. KSAM also inherits all properties of the original distance, is a valid reproducing kernel and universal. The induced metric by the kernel will be easily derived by traditional tensor analysis, and tuning the kernel hyperparameter of the used squared exponential kernel function is shown to be stable. The KSAM is here tested in a target detection scenario for illustration purposes. The spatially-explicit detection and metric maps give us useful information for semi-automatic anomaly/target detection. We show experimental evidence of the performance in very high resolution (VHR) satellite image target detection.

Kernel SAM: Given two spectra containing d -bands, $\mathbf{x}, \mathbf{z} \in \mathbb{R}^d$, the spectral angle mapper (SAM) measures the angle, θ , between them

$$\theta = \arccos\left(\frac{\mathbf{x}^\top \mathbf{z}}{\|\mathbf{x}\| \|\mathbf{z}\|}\right), \quad 0 \leq \theta \leq \frac{\pi}{2}. \quad (1)$$

The nonlinear SAM requires the definition of a feature mapping $\phi(\cdot)$ to a Hilbert space \mathcal{H} endorsed with the kernel reproducing property.

Definition 1: Reproducing kernel Hilbert space (RKHS). Given a Hilbert space \mathcal{H} with functions over \mathbb{R}^d , i.e. $f: \mathbb{R}^d \rightarrow \mathbb{R}$, the function $K(\cdot, \cdot): \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is called reproducing kernel of \mathcal{H} if $K(\mathbf{x}, \cdot) \in \mathcal{H}$, and \mathcal{H} is a RKHS.

Property 1: Reproducing property. If $\forall \mathbf{x} \in \mathbb{R}^d$ and $\forall f \in \mathcal{H}$ then $f(\mathbf{x}) = \langle f, K(\mathbf{x}, \cdot) \rangle$ and $\langle K(\cdot, \mathbf{x}), K(\cdot, \mathbf{z}) \rangle_{\mathcal{H}} = K(\mathbf{x}, \mathbf{z})$. This is the reproducing property of the kernel. A function f can thus be represented as a linear function defined by an inner product in the vector space \mathcal{H} .

Now, if we simply map the original spectra to \mathcal{H} with a mapping $\phi: \mathbb{R}^d \rightarrow \mathcal{H}$, the SAM can be expressed in \mathcal{H} as

$$\theta_K = \arccos\left(\frac{\phi(\mathbf{x})^\top \phi(\mathbf{z})}{\|\phi(\mathbf{x})\|_{\mathcal{H}} \|\phi(\mathbf{z})\|_{\mathcal{H}}}\right), \quad 0 \leq \theta_K \leq \frac{\pi}{2}. \quad (2)$$

Having access to the coordinates of the new mapped spectra is not possible unless one explicitly defines the mapping function ϕ .

Nevertheless, thanks to the reproducing property 1, it is possible to compute the new measure implicitly via kernels:

$$\theta_K = \arccos\left(\frac{K(\mathbf{x}, \mathbf{z})}{\sqrt{K(\mathbf{x}, \mathbf{x})} \sqrt{K(\mathbf{z}, \mathbf{z})}}\right), \quad 0 \leq \theta_K \leq \frac{\pi}{2}. \quad (3)$$

Popular examples of reproducing kernels are the linear kernel, $K(\mathbf{x}, \mathbf{z}) = \mathbf{x}^\top \mathbf{z}$, the polynomial $K(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^\top \mathbf{z} + 1)^d$, and the radial basis function (RBF) kernel, $K(\mathbf{x}, \mathbf{z}) = \exp(-\frac{1}{2\sigma^2} \|\mathbf{x} - \mathbf{z}\|^2)$. In the linear kernel, the associated RKHS is the space \mathbb{R}^d itself and KSAM reduces to the standard linear SAM. In polynomial kernels of degree d , KSAM effectively only compares moments up to order d . For the Gaussian kernel, the RKHS is of infinite dimension and KSAM measures higher order spectral dependences. In addition, note that for RBF kernels, self-similarity $K(\mathbf{x}, \mathbf{x}) = 1$, and thus the measure simply reduces to $\theta_K = \arccos(K(\mathbf{x}, \mathbf{z}))$, $0 \leq \theta_K \leq \pi/2$.

The derived KSAM has some additional properties that allow its use in any kernel machine.

Property 2: KSAM is a valid Mercer's kernel. The KSAM is a valid reproducing kernel because the arccos-operation simply expands the argument in an infinite sum of SAM distances weighted by positive scalars α_k , i.e. $\theta_K = \sum_{k=0}^{\infty} \alpha_k \theta^k$.

Property 3: KSAM is universal. The KSAM is a universal kernel on every compact subset of the input domain \mathcal{X} because of the previous series expansion [8].

Geometrical interpretation of kernel SAM: Kernel methods may appear elusive because the mapping ϕ is not explicitly defined, and the vector coordinates in the new feature spaces are not accessible. However, the framework allows to compute distances, angles, displacements, averages, and covariances implicitly in \mathcal{H} from the available data [7]. In addition, and very importantly, we show here that one can compute the metric associated to the used kernel.

For any positive definite kernel, we assume that the mapped data in \mathcal{H} are distributed in a surface \mathcal{S} smooth enough to be considered a Riemannian manifold [9]. The line element of \mathcal{S} can be expressed as

$$ds^2 = g_{ab} d\phi^a(\mathbf{x}) d\phi^b(\mathbf{x}) = g_{\mu\nu} dx^\mu dx^\nu,$$

where superscripts a and b correspond to the vector space \mathcal{H} , $g_{\mu\nu}$ is the induced metric, and the surface \mathcal{S} is parametrized by x^μ . Computing the components of the (symmetric) metric tensor only need the kernel function

$$g_{\mu\nu} = (1/2) \partial_{x^\mu} \partial_{x^\nu} K(\mathbf{x}, \mathbf{x}) - \{\partial_{z^\mu} \partial_{z^\nu} K(\mathbf{x}, \mathbf{z})\}_{\mathbf{z}=\mathbf{x}}. \quad (4)$$

For the RBF kernel with σ parameter, this metric tensor becomes flat, $g_{\mu\nu} = \delta_{\mu\nu}/\sigma^2$, and the squared geodesic distance between $\phi(\mathbf{x})$ and $\phi(\mathbf{z})$ simply becomes

$$\|\phi(\mathbf{x}) - \phi(\mathbf{z})\|_{\mathcal{H}}^2 = 2 \left(1 - \exp\left(-\frac{\|\mathbf{x} - \mathbf{z}\|_{\mathcal{X}}^2}{2\sigma^2}\right) \right) = 2(1 - K(\mathbf{x}, \mathbf{z})). \quad (5)$$

Note that the metric solely depends on the original spectra yet computed implicitly in a higher dimensional feature space \mathcal{H} , whose notion of distance is controlled by the parameter σ : the higher σ the smoother (linear) is the space. Actually, $\sigma \rightarrow \infty$ reduces the RBF kernel to approximately compute the Euclidean distance between vectors, the metric tensor reduces to $g_{\mu\nu} = 1$, and KSAM approximates the SAM solution.

An experimental evidence: A QuickBird image of a residential neighborhood of Zürich, Switzerland is used for illustration purposes. The image size is (329×347) pixels. A total of 40762 pixels were labeled by photointerpretation and assigned to 9 landuse classes (Fig. 2). Four target detectors are compared in the task of detecting the class 'Soil': orthogonal subspace projection (OSP) [10], its kernel counterpart (KOSP) [11], standard SAM [1], and the extension KSAM presented here.

Figure 1 shows the receiver operating characteristic (ROC) curves, and the area under the ROC curves as a function of the kernel lengthscale parameter σ . KSAM shows excellent detection rates, especially remarkable in the inset plot (note the logarithmical scale). Perhaps more importantly, the KSAM method is relatively insensitive to the selection of the kernel parameter compared to the KOSP detector, provided that a large enough value is specified.

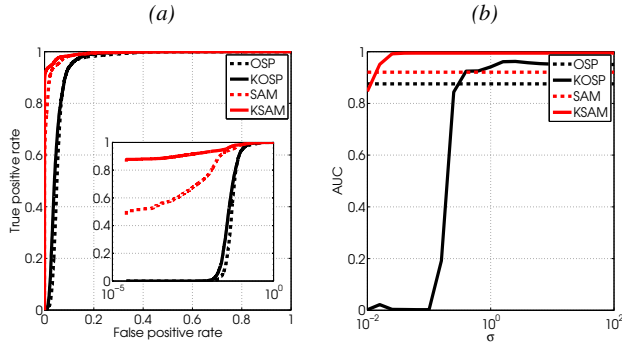


Fig. 1 (a) Receiver operating characteristic (ROC) curves and (b) area under the ROC curves as a function of the kernel lengthscale parameter.

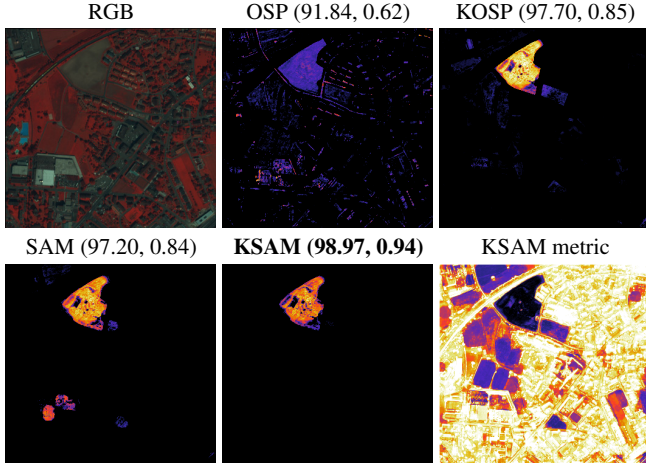


Fig. 2. Detection results for different algorithms (Accuracy, kappa statistic).

The latter experiments allow us to use a traditional prescription to fix the RBF kernel parameter σ for both KOSP and KSAM as the mean distance among all spectra, d_M . Note that after data standardization and proper scaling, this is a reasonable heuristic $\sigma \approx d_M = 1$. The thresholds were optimized for all methods. OSP returns a decision function strongly contaminated by noise, while the KOSP detector results in a correct detection. Fig. 2 shows the detection maps and the metric learned. The (linear) SAM gives rise to very good detection but with strong false alarms in the bottom left side of the image, where the roof of the commercial center saturates the sensor and thus returns a flat spectrum for its pixels in the morphological features. As a consequence, both the target and the roof vectors have flat spectra and their spectral angle is almost null. The proposed KSAM can efficiently cope with these (nonlinear) saturation problems. In addition, the metric space derived from the kernel suggests high discriminative (and spatially localized) capabilities.

Conclusions: A nonlinear generalization of the SAM metric was introduced. The kernel SAM replaces the dot product between spectra by a positive-definite kernel function. It is shown the induced metric by the kernel function and a geometrical intuition for tuning the kernel parameter. Experiments on a real multispectral very high resolution multispectral image illustrated the validity and effectiveness of the proposed method.

Acknowledgment: This paper has been partially supported by the European Research Council (ERC) under the ERC-CoG-2014 project 647423, <http://erc.europa.eu/>

G. Camps-Valls (*Image Processing Lab (IPL), Univ. València, Spain*)

E-mail: gustau.camps@uv.es

References

- 1 F. A. Kruse, A. B. Lefkoff, J. W. Boardman, A. T. Heidebrecht, K. B. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS) – Interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, no. 2/3, pp. 145–163, May/Jun 1993.
- 2 J. E. Ball and L. M. Bruce, "Level set hyperspectral image classification using best band analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3022–3027, Oct 2007.
- 3 C. Hecker, M. van der Meijde, H. van der Werff, and F. D. van der Meer, "Assessing the influence of reference spectra on synthetic sam classification results," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 12, pp. 4162–4172, Dec 2008.
- 4 P. B. Garcia-Allende, O. M. Conde, J. Mirapeix, A. M. Cubillas, and J. M. Lopez-Higuera, "Data processing method applying principal component analysis and spectral angle mapper for imaging spectroscopic sensors," *IEEE Sensors Journal*, vol. 8, no. 7, pp. 1310–1316, July 2008.
- 5 M. A. Cho, P. Debba, R. Mathieu, L. Naidoo, J. van Aardt, and G. P. Asner, "Improving discrimination of savanna tree species through a multiple-endmember spectral angle mapper approach: Canopy-level analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4133–4142, Nov 2010.
- 6 N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geosc. Rem. Sens.*, vol. 42, no. 7, pp. 1552–1565, 2004.
- 7 G. Camps-Valls and L. Bruzzone, Eds., *Kernel methods for Remote Sensing Data Analysis*. UK: Wiley & Sons, Dec 2009.
- 8 C. A. Micchelli, Y. Xu, and H. Zhang, "Universal kernels," *Journal of Machine Learning Research*, vol. 6, pp. 2651–2667, 2006.
- 9 C. J. C. Burges, "Geometry and invariance in kernel based methods," in *Advances in Kernel Methods – Support Vector Learning*, A. S. B. Schölkopf, C.J.C. Burges, Ed. Cambridge, USA: MIT Press, 1999.
- 10 J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Trans. Geosc. Rem. Sens.*, vol. 32, no. 4, pp. 779–785, 1994.
- 11 H. Kwon and N. Nasrabadi, "A comparative analysis of kernel subspace target detectors for hyperspectral imagery," *EURASIP Journal of Advances in Signal Proc.*, vol. 2007, no. 29250, 2007.