UNSUPERVISED AND SUPERVISED APPROACHES TO COLOR SPACE TRANSFORMATION FOR IMAGE CODING

Massimo Minervini, Cristian Rusu, Sotirios A. Tsaftaris

IMT Institute for Advanced Studies
Piazza San Ponziano 6, Lucca, Italy

ABSTRACT

The linear transformation of input (typically RGB) data into a color space is important in image compression. Most schemes adopt fixed transforms to decorrelate the color channels. Energy compaction transforms such as the Karhunen-Loève (KLT) do entail a complexity increase. Here, we propose a new data-dependent transform (aKLT), that achieves compression performance comparable to the KLT, at a fraction of the computational complexity. More important, we also consider an application-aware setting, in which a classifier analyzes reconstructed images at the receiver's end. In this context, KLT-based approaches may not be optimal and transforms that maximize post-compression classifier performance are more suited. Relaxing energy compactness constraints, we propose for the first time a transform which can be found offline optimizing the Fisher discrimination criterion in a supervised fashion. In lieu of channel decorrelation, we obtain spatial decorrelation using the same color transform as a rudimentary classifier to detect objects of interest in the input image without adding any computational cost. We achieve higher savings encoding these regions at a higher quality, when combined with region-of-interest capable encoders, such as JPEG 2000.

Index Terms— Image compression, color space transformation, statistical learning, JPEG 2000.

1. INTRODUCTION

The red-green-blue (RGB) representation is not efficient for coding, due to high correlation between color bands of natural images [1]. To reduce spectral redundancy, image and video compression algorithms operate on luminance/chrominance representations of the color information, achieved through linear transformations of the RGB color space. Each color band is coded independently, therein deploying a variety of techniques to address spatial and, for video, also temporal correlation. A family of such color models is the YC_bC_r [2], adopted by many coding standards. However, due to high variability in source image characteristics, a fixed transform

may easily result in suboptimal performance, thus motivating the adoption in some contexts of a data-dependent one.

The energy-compaction and decorrelation properties of the Karhunen-Loève transform (KLT) make it desirable for color image compression [1, 3–6]. It was shown to be superior to other approaches in a variety of contexts, both for color [7–9] and hyperspectral [10] imagery, and has formed the basis for new fixed transforms [11–13]. However, the computational complexity of calculating the color covariance matrix, limits its applicability in sensing environments with low computational power. A variety of approaches have been proposed to circumvent this bottleneck, either relying on covariance matrix approximations [14], sub-sampling strategies [15, 16], or learning approaches estimating projection directions [16]. In this paper, we propose a new data-dependent color transform, termed aKLT, rooted in the orthogonal Procrustes problem, that preserves energy compaction and performs similar to KLT, but is less computationally complex.

Although KLT and aKLT are designed to match the statistical properties of the image data, they are agnostic to the semantics of the scene (e.g., distinction between foreground and background regions). In present days, image data are often analyzed by computer vision algorithms (e.g., surveillance applications) and their transmission over band-limited channels necessitates their compression. It was shown recently that considering the application and designing data codecs appropriately, that do not maximize fidelity type criteria (such as the mean squared error), but consider how would an analysis algorithm (e.g., a classifier) perform on compressed data, is beneficial from a bit rate perspective [17]. This notion was explored in [17] and [18] with respect to quantization, however, as of now the design of color transforms optimized particularly for classification accuracy has not been considered yet. Thus, given some previously labeled data, we propose a methodology to obtain application-dependent color transforms, that while aiming to retain energy compaction properties, also try to maximize separability of the transformed data.

The rest of this paper is organized as follows. Section 2 details our methodology to learn color transforms from the data. Section 3 demonstrates the proposed approaches, using the JPEG 2000 standard to compress test images. Finally, Section 4 offers concluding remarks.

2. METHODOLOGY

We represent an RGB image as a $3 \times n$ matrix $\mathbf{X} = (\boldsymbol{r}, \boldsymbol{g}, \boldsymbol{b})^{\mathsf{T}}$, where $\boldsymbol{r}$, $\boldsymbol{g}$, and $\boldsymbol{b}$ are the linearized color components, and n is the number of pixels. Prior to lossy coding, $\mathbf{X}$ is projected into a new color space by $\mathbf{T} \in \mathbb{R}^{3 \times 3}$. Each pixel value $\boldsymbol{x}_i = (r_i, g_i, b_i)^{\mathsf{T}}$ in $\mathbf{X}$ is transformed with the linear relation $\boldsymbol{y}_i = \mathbf{T}\boldsymbol{x}_i$. Upon reconstruction, the color transform is inverted, obtaining the approximation $\widetilde{\boldsymbol{x}}_i$ in the RGB domain. To ensure output dynamic range of $\boldsymbol{y}_i$ be the same as $\boldsymbol{x}_i$ (e.g., 0 to 255, for 8-bit unsigned integer representation), we scale the directions (rows) of $\mathbf{T}$ with respect to the ℓ_1 norm [12].

In the following, we address the problem of obtaining data- and application-dependent color transforms. In Section 2.1, based on a heuristic, we derive a new low-complexity transform (aKLT), that adapts to the content using only statistical information from the image. In Section 2.2, we propose a novel approach, that finds color space transformations using supervised learning methods on labeled training data.

2.1. The aKLT: A low-complexity unsupervised data-dependent transform

The KLT produces an orthogonal transformation, $\mathbf{K}$, obtained from the eigendecomposition of the color covariance matrix $\boldsymbol{\Sigma} = \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^{\mathsf{T}}$, where $\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i$ is the mean color vector. The eigenvectors of $\boldsymbol{\Sigma}$, sorted in decreasing order of magnitude of the corresponding eigenvalues, define the directions of $\mathbf{K}$. The KLT achieves complete statistical decorrelation of the color signals and energy compaction in the first channel, thus favoring efficient representation and sub-sampling of the other two channels.

However, estimation of $\boldsymbol{\Sigma}$ can be demanding in memory and computation power, particularly for large images, and its application in resource-constrained sensing devices can be problematic. Thus, we seek to find a transform that is close to the KLT but less computationally complex to obtain.

Let $\overline{\mathbf{X}} \in \mathbb{R}^{3 \times n}$ be the matrix obtained by normalizing each column (pixel) of $\mathbf{X}$ with respect to the ℓ_2 norm. We seek an orthogonal transform $\boldsymbol{\Omega} \in \mathbb{R}^{3 \times 3}$ that maps $\overline{\mathbf{X}}$ into a given reference matrix $\mathbf{W} \in \mathbb{R}^{3 \times n}$, and formulate it as:

$$\underset{\boldsymbol{\Omega}}{\text{minimize}} \ \|\mathbf{W} - \boldsymbol{\Omega}\overline{\mathbf{X}}\|_F \quad \text{subject to} \quad \boldsymbol{\Omega}^{\mathsf{T}}\boldsymbol{\Omega} = \mathbf{I}, \quad (1)$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and $\mathbf{I}$ is the identity matrix. Let $\mathbf{Z} = \mathbf{W}\overline{\mathbf{X}}^{\mathsf{T}}$, and $\mathbf{Z} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}}$ be the singular value decomposition (SVD) of $\mathbf{Z}$. The optimization problem of Eq. (1), known as orthogonal Procrustes problem, admits closed-form solution $\mathbf{U}\mathbf{V}^{\mathsf{T}}$ [19]. In order to concentrate energy in the first direction, we impose structure to $\mathbf{W}$.

$$\mathbf{W} = \begin{pmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} \sum_{i=1}^{n} r_i & \sum_{i=1}^{n} g_i & \sum_{i=1}^{n} b_i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notably, this leads to a simplified form of $\mathbf{Z}$ with only a single direction, $\boldsymbol{a}_1 = \boldsymbol{z}_1^{\mathsf{T}}/\|\boldsymbol{z}_1\|_2$, that corresponds to the principal

Table 1: Comparison of KLT approaches as a function of input size n, where n denotes number of image pixels.

KLT, ACKLT [14]	Penna et al. [15]	IPCA [16]	aKLT
$18n$	$\rho 18n$	$15n$	$11n$

direction, thus making the SVD computation unnecessary.

The vector $\boldsymbol{a}_1$ approximates the principal direction of the KLT. In order to obtain the full transform, we proceed by constructing the 3×3 matrix $\mathbf{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$, where $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ are initialized with random elements (the effect of randomness on performance is explored in Section 3). Subsequently, we use QR factorization to decompose $\mathbf{A}$ into the product $\mathbf{A} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q} \in \mathbb{R}^{3 \times 3}$ has orthogonal columns and $\mathbf{R} \in \mathbb{R}^{3 \times 3}$ is upper triangular. The aKLT transformation matrix, $\widetilde{\mathbf{K}} = \mathbf{Q}^{\mathsf{T}}$, shares relevant properties with regular KLT: (a) orthogonality, and (b) energy compaction capabilities. Although there is no guarantee on sorting and relative amount of energy of second and third channels, this is not of concern from a compression standpoint (e.g., chroma sub-sampling strategies usually treat such components equivalently).

Computation of the KLT is dominated by mean subtraction and calculation of covariance matrix $\boldsymbol{\Sigma}$, requiring $18n$ floating point operations, where n is the number of pixels. We ignore cost of subsequent eigenvalue decomposition of $\boldsymbol{\Sigma}$ to obtain $\mathbf{K}$, as it does not scale with n. Therefore, approaches that speed up this step (e.g., power method, or ACKLT [14]) provide negligible benefit. As shown in Table 1, our proposed aKLT requires only $11n$ operations, approximately 40% reduction in complexity compared to the KLT. The IPCA [16], based on neural networks, achieves an approximation of the principal direction in $15n$ operations, while [15] necessitates to keep a fraction $\rho = 0.6$ of the data to match the aKLT.

2.2. A supervised approach to an application-dependent color transform using labeled pixels

It is known that projecting to principal components is not always optimal from a pattern recognition perspective: clusters of points belonging to semantically different objects in the scene may overlap now in the projected color space. Introducing distortions due to lossy compression may affect this separability further. In this section, we seek to identify a transform that maintains class separation and decorrelation properties.

We assume that this optimization will occur in an offline fashion and we will use a training set (pixels segmented in two classes), thus it is supervised. Compression of newly acquired images at the sensor occurs as before, with the transform now known. The calculation of a new transform is necessary only if the scene conditions change (depending on the process being observed) and if new training data are available.

Let C_1 and C_2 be disjoint sets of pixel values ($C_1 \cap C_2 = \emptyset$) representative of distinct pattern classes. We seek an orthogonal transform $\mathbf{D} \in \mathbb{R}^{3 \times 3}$ that projects data points belonging to distinct classes, $\boldsymbol{x}_1 \in C_1$ and $\boldsymbol{x}_2 \in C_2$, in a domain

where they are maximally separated according to measure $\mathcal{C}$:

$$\underset{\mathbf{D}}{\text{maximize}} \qquad \mathcal{C}\{\mathbf{D}\boldsymbol{x}_1, \mathbf{D}\boldsymbol{x}_2\} \qquad (2a)$$

$$\text{subject to} \qquad \mathbf{D}^\mathsf{T}\mathbf{D} = \mathbf{I}, \qquad (2b)$$

$$\|\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^\mathsf{T} - \boldsymbol{\Lambda}\|_F \leq \varepsilon, \qquad (2c)$$

where $\boldsymbol{\Sigma}$ is the color covariance matrix, $\boldsymbol{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues of $\boldsymbol{\Sigma}$, and $\varepsilon \geq 0$. The last constraint aims to optimize for energy compaction and decorrelation, however, solving this problem is difficult with constraints that could in case appear conflicting. In the remainder, we ignore Eq. (2c) focusing on class separation, but we revisit the complete problem in the next section.

An effective measure of class separability is the Fisher criterion, defined as $J(\boldsymbol{d}) = (\boldsymbol{d}^\mathsf{T}\mathbf{S}_b\boldsymbol{d})/(\boldsymbol{d}^\mathsf{T}\mathbf{S}_w\boldsymbol{d})$, where $\boldsymbol{d} \in \mathbb{R}^3$, $\mathbf{S}_b = \sum_{i=1}^{2}(\boldsymbol{m}_i - \boldsymbol{\mu})(\boldsymbol{m}_i - \boldsymbol{\mu})^\mathsf{T}$ is the between-class scatter matrix, $\mathbf{S}_w = \sum_{i=1}^{2}\sum_{\boldsymbol{x} \in C_i}(\boldsymbol{x} - \boldsymbol{m}_i)(\boldsymbol{x} - \boldsymbol{m}_i)^\mathsf{T}$ is the within-class scatter matrix, $\boldsymbol{\mu} = \sum_{i=1}^{2}P_i\boldsymbol{m}_i$ is the mean sample vector, and $\boldsymbol{m}_i$ and P_i are, respectively, mean and a priori probability of class i. A closed-form solution to finding orthogonal discriminant vectors that maximize the Fisher criterion, can be obtained adopting the Foley-Sammon Transform (FST) [20]. The first direction, $\boldsymbol{d}_1$, corresponds to the eigenvector associated with the largest eigenvalue of the matrix $\mathbf{S}_w^{-1}\mathbf{S}_b$. If $\mathbf{D} = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_r)^\mathsf{T}$ is the set of previously obtained directions, $\boldsymbol{d}_{r+1}$ is found recursively as the eigenvector associated with the largest eigenvalue satisfying $\mathbf{M}\mathbf{S}_b\boldsymbol{d} = \alpha\mathbf{S}_w\boldsymbol{d}$, where $\mathbf{M} = \mathbf{I} - \mathbf{D}^\mathsf{T}(\mathbf{D}\mathbf{S}_w^{-1}\mathbf{D}^\mathsf{T})^{-1}\mathbf{D}\mathbf{S}_w^{-1}$ [21]. The final color transform matrix is defined by $\mathbf{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3)^\mathsf{T}$.

2.2.1. Spatial decorrelation in lieu of spectral decorrelation

Our approach for finding $\mathbf{D}$ above ignored the constraint of Eq. (2c), finding one that only optimizes for separation. We could find a new transform $\mathbf{D}'$ that is close to $\mathbf{D}$ whilst trying to satisfy Eq. (2c), or equivalently, since we know that the (a)KLT optimizes Eq. (2c), we can pose:

$$\underset{\mathbf{D}'}{\text{minimize}} \|\mathbf{D}' - \mathbf{D}\|_F + \lambda\|\mathbf{D}' - \widetilde{\mathbf{K}}\|_F, \qquad (3)$$

thus, finding a transform that is between $\mathbf{D}$ (application-aware, obtained offline) and the aKLT (data-aware, obtained at the sensor). However, while this adapts the supervised transform to unseen data on the sensor and will have decorrelating properties, from a computational perspective it is not attractive, and possibly prohibitive for a low compute-power device. On the other hand, we can exploit the separation properties of $\mathbf{D}$, [20], to obtain spatial decorrelation and recover potential losses in bit rate performance when using solely $\mathbf{D}$.

Spatial decorrelation can be used in an encoder with region of interest (ROI) coding capability (e.g. JPEG 2000 [22]). Thus, we identify potential ROI masks solely on the basis of the transform $\mathbf{D}$. With respect to other approaches obtaining the ROI information from a module external to the encoder [23], this reduces computational overhead.
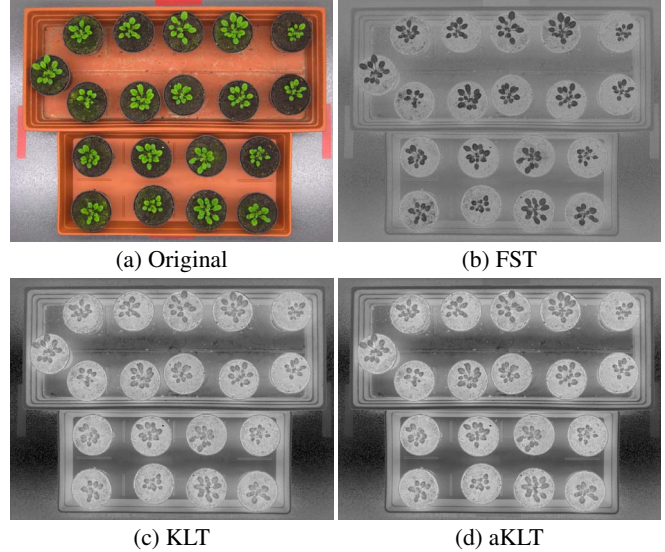

Fig. 1: (a) Test image and projection on the first component of (b) FST, (c) KLT, and (d) aKLT, respectively. Notice the good discrimination capability of the plant objects in (b), and the high similarity between output of KLT (c) and aKLT (d).

The first channel of the FST domain, $y_i^{(1)} = \boldsymbol{d}_1^\mathsf{T}\boldsymbol{x}_i$, corresponds to the projection on Fisher's discriminant vector (see Figure 1b). In an unseen image, to obtain an ROI estimate, $\boldsymbol{\Gamma}(\theta^*) \in \{0,1\}^n$, we decide the class of a pixel (foreground or background) based on a single threshold θ^* on the values of $y_i^{(1)}$. Rather than making any statistical assumptions on the distribution of the data (e.g., Gaussian), for which optimal θ^* have known closed-form solutions, we estimate θ^* from our training set, maximizing the average Dice Similarity Coefficient, $\theta^* = \arg\max_\theta (2 \cdot |\boldsymbol{\Gamma}_{\text{GT}} \cap \boldsymbol{\Gamma}(\theta)|) / (|\boldsymbol{\Gamma}_{\text{GT}}| + |\boldsymbol{\Gamma}(\theta)|)$, between the ground truth of pixels, $\boldsymbol{\Gamma}_{\text{GT}}$, and the classification obtained with $\mathbf{D}$ and threshold θ on the training data.

3. RESULTS AND DISCUSSION

3.1. Experimental settings

The proposed methodology is evaluated on a dataset of 20 images (3108×2324 pixels) from a time-lapse sequence of 19 arabidopsis plant subjects (Figure 1a). These images are ideal to showcase the methodology since they are usually large and due to design requirements they need to be communicated via the Internet to centralized locations for processing [23].

We include in the comparison plain RGB and YC_bC_r (ITU-R BT.601) [2]. KLT and aKLT are computed for each image. FST is estimated on the first image of the sequence, and then applied to all subsequent ones. After color space transformation, images are compressed at various bit rates using the JJ2000 software implementation[1] of JPEG 2000 [22].

The approaches are evaluated according to: (a) reconstruction accuracy, measured with PSNR in image domain,

---

[1] http://code.google.com/p/jj2000/

Table 2: Reconstruction accuracy comparison.

Rate (bpp)	Average PSNR (dB)				
	RGB	YC_bC_r	KLT	aKLT	FST
0.0625	26.75	27.07	27.28	27.33	26.82
0.125	27.86	28.31	28.44	28.45	27.94
0.25	29.09	29.53	29.58	29.58	29.12
0.5	30.53	30.78	30.81	30.80	30.49
1.0	32.39	32.39	32.28	32.26	32.07
2.0	34.86	34.68	34.48	34.36	34.43

and (b) application error. To estimate application error we first build a rudimentary classifier. Similar to the approach in [24], we train a Gaussian mixture model, $\mathcal{M}$, on color features, using labeled foreground (plant) data from the first uncompressed image (excluded from testing). At each bit rate, we calculate the application error, $E_{\mathcal{M}} = \frac{1}{n}\sum_{i=1}^{n}(\mathcal{M}(\boldsymbol{x}_i) - \mathcal{M}(\widetilde{\boldsymbol{x}}_i))^2$, between the probabilistic output of $\mathcal{M}$ evaluated on the n original, $\boldsymbol{x}_i$, and reconstructed, $\widetilde{\boldsymbol{x}}_i$, image pixels.

3.2. Results

In this section, we present the rate-distortion performance of the proposed approaches. We first compare them in terms of overall reconstruction accuracy. Next, we demonstrate the supervised approach in an application-aware context.

Table 2 reports image fidelity results. At low bit rates (< 1 bpp), decorrelating transforms (YC_bC_r, KLT, aKLT) achieve better performance than RGB (0.25 to 0.6 dB improvement in PSNR), with the data-dependent transforms (KLT, aKLT) outperforming the fixed YC_bC_r. Notably, our proposed low-complexity aKLT, $\widetilde{\mathbf{K}}$, exhibits performance very close to regular KLT (on average 0.04 dB difference). As also found by others [25], at higher bit rates RGB representation is the best option, due to noise amplification effects of the transformations and reduced quantization. The supervised FST, $\mathbf{D}$, does not provide any PSNR benefits (performance comparable to RGB with 0.5% relative change, on average across bit rates), probably due to lacking decorrelation capabilities.

On the other hand, the aKLT shows good decorrelating capabilities: in the aKLT domain, on average, second and third channels have linear correlation of 0.05 ± 0.16, and with the first channel of 0.11 ± 0.26 and -0.28 ± 0.05, respectively. In order to assess the sensitivity of the aKLT to the random initialization, we compute 100 different realizations of $\widetilde{\mathbf{K}}$. The behavior of the aKLT oscillates only minimally, with an average (across bit rates) standard deviation of 0.06 dB in PSNR.

Figure 2 compares the approaches from an application standpoint. Color transformation provides up to 1.17 dB improvement in foreground PSNR relative to RGB, with the FST now competitive. When used in a spatial decorrelation context to estimate an ROI, combined with the ROI coding feature of JPEG 2000, the FST + ROI approach obtains a major improvement at all bit rates: 0.8 to 7.7 dB increase in fore-
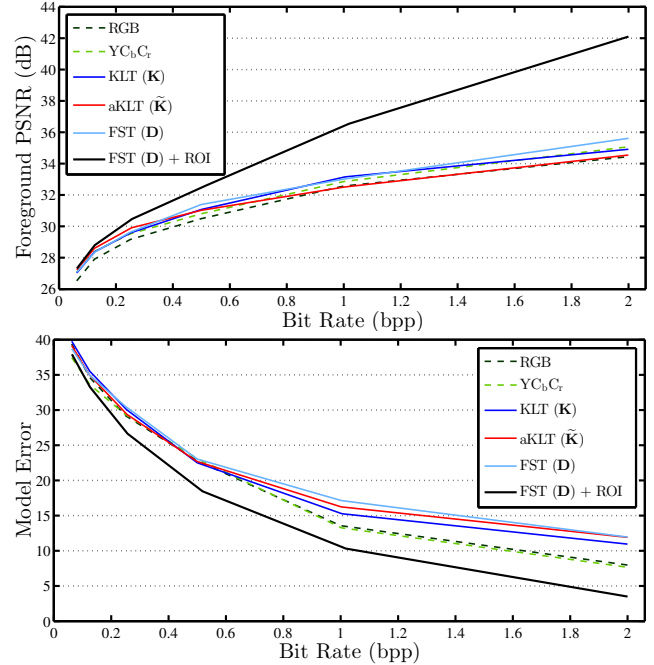

Fig. 2: R-D performance using application-aware metrics: (top) reconstruction accuracy of the objects of interest, and (bottom) model error $E_{\mathcal{M}}$, averaged over all test images.

ground PSNR, and 4 to 56% reduction in application error.

The results envision different use cases for the proposed approaches. The aKLT is general purpose and can be efficiently calculated on a per image basis to target reconstruction accuracy. The supervised approach is suited for application-aware compression and is computed offline. The regularized versions of Eq. (3), will be highly dependent on the free parameter λ and their performance is expected to lie within the bounds of the other two. Therefore, it is best to exploit the classification abilities of the supervised FST to focus bits spatially, which is considerably less computationally demanding.

4. CONCLUSIONS

We address the problem of designing image-adaptive color transforms for coding applications. In recognition of superior performance of the KLT with respect to fixed transforms, we derive a low-complexity approximation, the aKLT, capable of comparable performance, easing adoption on devices. We also formulate a novel supervised approach to obtain color transforms with class separation capabilities, identifying a solution in the FST. We use its classification property to inform the encoder where to focus bit rate, thus improving both reconstruction and application accuracy. When coupled with quantizer design even greater bit rate savings are possible, but that would violate standard compliance. Increased image resolution is expected to emphasize the benefits of the proposed approaches. While we adopt JPEG 2000, our methodology is general and can be adapted to other coding schemes (assuming an ROI-capable encoder for the supervised approach).

5. REFERENCES

[1] W. K. Pratt, "Spatial transform coding of color images," *IEEE Trans. Commun. Technol.*, vol. 19, no. 6, pp. 980–992, Dec. 1971.

[2] *Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios*, ITU-R Recommendation BT.601-5, Oct. 1995.

[3] C. B. Rubinstein and J. O. Limb, "Statistical dependence between components of a differentially quantized color signal," *IEEE Trans. Commun.*, vol. 20, no. 5, pp. 890–899, Oct. 1972.

[4] S. G. Wolf, R. Ginosar, and Y. Y. Zeevi, "Spatio-chromatic model for colour image processing," in *Int. Conf. on Pattern Recognition*, vol. 1, 1994, pp. 599–601.

[5] R. K. Kouassi, J. C. Devaux, P. Gouton, and M. Paindavoine, "Application of the Karhunen-Loeve transform for natural color images analysis," in *Asilomar Conf. on Signals, Systems & Comp.*, vol. 2, 1997, pp. 1740–1744.

[6] E. Gershikov and M. Porat, "Does decorrelation really improve color image compression?" in *Int. Conf. on Systems Theory and Scientific Computation (ISTASC)*, Sep. 2005, pp. 306–309.

[7] S.-E. Han, B. Tao, T. Cooper, and I. Tastl, "Comparison between different color transformations for the JPEG 2000," in *PICS 2000: Image Processing, Image Quality, Image Capture, Systems Conf.*, Mar. 2000, pp. 259–263.

[8] A. A. Kassim and W. S. Lee, "Embedded color image coding using SPIHT with partially linked spatial orientation trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 2, pp. 203–206, Feb. 2003.

[9] Y. Chen, P. Hao, and A. Dang, "Optimal transform in perceptually uniform color space and its application in image coding," in *Int. Conf. on Image Analysis and Recognition (ICIAR)*, Oct. 2004, pp. 269–276.

[10] Q. Du and J. E. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 201–205, Apr. 2007.

[11] P. Hao and Q. Shi, "Comparative study of color transforms for image coding and derivation of integer reversible color transform," in *Int. Conf. on Pattern Recognition (ICPR)*, vol. 3, 2000, pp. 224–227.

[12] H. M. Kim, W.-S. Kim, and D.-S. Cho, "A new color transform for RGB coding," in *Int. Conf. on Image Processing (ICIP)*, Oct. 2004, pp. 107–110.

[13] H. S. Malvar, G. J. Sullivan, and S. Srinivasan, "Lifting-based reversible color transformations for image compression," in *Proceedings of SPIE, Applications of Digital Image Processing XXXI*, vol. 7073, Sep. 2008.

[14] R. Kountchev and R. Kountcheva, "New method for adaptive Karhunen-Loeve color transform," in *Int. Conf. on Telecommun. in Modern Satellite, Cable, and Broadcasting Services (TELSIKS)*, Oct. 2009, pp. 209–216.

[15] B. Penna, T. Tillo, E. Magli, and G. Olmo, "A new low complexity KLT for lossy hyperspectral data compression," in *Int. Geoscience and Remote Sensing Symp. (IGARSS)*, vol. 7, Aug. 2006, pp. 3525–3528.

[16] Q. Du and J. E. Fowler, "Low-complexity principal component analysis for hyperspectral image compression," *Int. J. of High Performance Computing Applications*, vol. 22, no. 4, pp. 438–448, Nov. 2008.

[17] E. Soyak, S. A. Tsaftaris, and A. K. Katsaggelos, "Low-complexity tracking-aware H.264 video compression for transportation surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 10, pp. 1378–1389, Oct. 2011.

[18] B. M. Dogahe and M. N. Murthi, "Quantization for classification accuracy in high-rate quantizers," in *Digital Signal Process. Workshop*, Jan. 2011, pp. 277–282.

[19] P. H. Schönemann, "A generalized solution of the orthogonal Procrustes problem," *Psychometrika*, vol. 31, no. 1, pp. 1–10, Mar. 1966.

[20] D. H. Foley and J. W. Sammon, "An optimal set of discriminant vectors," *IEEE Trans. Comput.*, vol. C-24, no. 3, pp. 281–289, 1975.

[21] Z. Jin, J.-Y. Yang, Z.-S. Hu, and Z. Lou, "Face recognition based on the uncorrelated discriminant transformation," *Pattern Recognition*, vol. 34, no. 7, pp. 1405–1416, Mar. 2001.

[22] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 36–58, 2001.

[23] M. Minervini and S. A. Tsaftaris, "Application-aware image compression for low cost and distributed plant phenotyping," in *Int. Conf. on Digital Signal Processing (DSP)*, Jul. 2013, pp. 1–6.

[24] M. Minervini, M. M. Abdelsamea, and S. A. Tsaftaris, "Image-based plant phenotyping with incremental learning and active contours," *Ecol. Inform.*, 2013.

[25] D. Marpe, H. Kirchhoffer, V. George, P. Kauff, and T. Wiegand, "Macroblock-adaptive residual color space transforms for 4:4:4 video coding," in *Int. Conf. on Image Processing (ICIP)*, Oct. 2006, pp. 3157–3160.