TRANSFORM CODER IDENTIFICATION WITH DOUBLE QUANTIZED DATA

Marco Tagliasacchi* *Marco Visentini-Scarzanella*† *Pier Luigi Dragotti*† *Stefano Tubaro**

* Dipartimento di Elettronica e Informazione, Politecnico di Milano
† EEE Department, Imperial College London

ABSTRACT

The analysis of chains of double transform coders has been recently addressed in the image forensic literature, especially for the case of double JPEG compression. In that case, the transform is assumed to be known a priori (e.g., 2D-DCT), whereas the quantization steps of the first coder need to be determined. In this work, we generalize the analysis to the challenging case in which nothing is known about the first coder, but that the transform is orthonormal. Given a set of vectors observed as output of a chain of two transform coders, we identify both the transform and the quantizer of the first. The key idea is to denoise the observed vectors exploiting the constraints imposed by the first quantizer and then apply our previously proposed method, which successfully performs transform identification in the case of noiseless observations. Experiments on real images validate the proposed approach.

Index Terms— Transform coding, double compression.

1. INTRODUCTION

Transform coding plays a central role in all multimedia communication standards. Hence, the footprints left by transform coding have been successfully exploited in the field of multimedia forensics to reveal, e.g., evidence of tampering [1][2] and resampling [3][4]. Most of the literature focused on single [5][6], double [2], and multiple [7] JPEG image compression, whereas just a few works address the case of video signals, for MPEG-2 [8][9], MPEG-4 [10][11] and H.264/AVC [12], or for the identification of the motion estimation algorithm [13] and the video coding standard [14]. For a survey, see [15] and the references therein. In all these works, the adopted transform is assumed to be known, in terms of both type (e.g., separable 2D-DCT) and size (e.g., 8×8). This simplifies the problem considerably, since the analyst need only to estimate the quantization matrix and, possibly, the shift and/or the resampling factor.

Although earlier standards (e.g., JPEG, MPEG-2 and MPEG-4) adopted the Discrete Cosine Transform (DCT) on 8×8 blocks, more recent coding architectures are more diversified in terms of both the type of transform being used and the block size. For example, H.264/AVC adopts an integer arithmetic approximation of the DCT transform on either 4×4 or 8×8 blocks [16]. The recent HEVC standard under development goes even further, supporting four different transform block sizes in the range from 4×4 to 32×32 [17]. In addition, the adoption of hybrid transforms, which can be obtained by means of a separable combination of DCT and DST (Discrete Sine Transform) is being investigated in [18]. All this indicates that the identification of the actual transform being used might give important clues about the processing history of a digital signal.

Most of the methods appeared in the multimedia forensics literature [15] focus only on a specific type of multimedia signal (e.g., only images or only videos) and are to some extent heuristic. It is therefore natural to try and develop a universal theory of transform coder identification that is independent of the specific application at hand. In our previous work [19][20], we considered the problem of identifying a transform coder when observing its output, by exploiting the footprints left by quantization and leveraging lattice theory notions. However, in many realistic cases, the output is processed further, as in the case of double compression, so that the method in [20] cannot be directly applied. To the best of the authors' knowledge only [21] addresses a related problem, proposing a method to reveal the color compression history, i.e., the colorspace used in JPEG compression. However, the solution proposed in [21] is tailored to work in a 3-dimensional vector space, thus avoiding the challenges that arise in higher dimensional spaces. In parallel to our contribution, a refined version of [21] was proposed in [22] based on the notion of dual lattice.

In this paper we consider a general model that entails a processing chain of two transform coders, which can describe a large variety of practical implementations that are found in lossy coding systems, including those adopted in multimedia communication. In [23] we studied the effect of the above coding chain for the case of a 2-dimensional transform, in order to determine the analytical conditions under which it is possible to navigate back up the signal's history to the first coding stage and determine the first encoder's exact transform parameters. Instead, in this work we address the more challenging case of N-dimensional transforms, with $N \geq 2$. Our main contribution is a method that is able to identify the parameters used by the first transform coder (i.e., the adopted transform and quantizer), when observing the output of the second transform coder, thus generalizing the work in [20].

Section 2 introduces the problem more formally. The proposed algorithm for transform identification is illustrated in Section 3 and validated on real images in Section 4. Section 5 concludes the paper.

2. PROBLEM STATEMENT

Let $\mathbf{x}$ denote a N-dimensional vector and $\mathbf{W}$ a transform matrix, whose rows represent the transform basis functions. Transform coding is performed by applying scalar quantization to the transform coefficients $\mathbf{y} = \mathbf{W}\mathbf{x}$. Let $\mathcal{Q}_i(\cdot)$ denote the quantizer associated to the i-th transform coefficient. We assume that $\mathcal{Q}_i(\cdot)$ is a scalar uniform quantizer with step size Δ_i, $i = 1, \ldots, N$. Therefore, the reconstructed quantized coefficients can be written as $\tilde{\mathbf{y}} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_N]^T$, with $\tilde{y}_i = \mathcal{Q}_i(y_i) = \Delta_i \cdot \text{round}[y_i/\Delta_i]$, $i = 1, \ldots, N$. The reconstructed block in the original domain is given by $\tilde{\mathbf{x}} = \mathbf{W}^{-1}\tilde{\mathbf{y}}$.

The project REWIND acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number:268478.

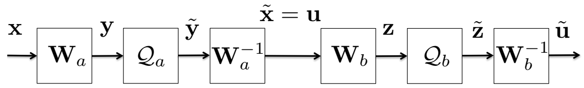


Fig. 1. Block diagram of a chain of two transform coders.

We consider the case in which an input signal is encoded by partitioning it into non-overlapping N-dimensional vectors, which are then processed as illustrated in Figure 1, by cascading two transform coders characterized, respectively, by the transform matrices $\mathbf{W}_a$ and $\mathbf{W}_b$ and quantizers $\mathcal{Q}_{a,i}(\cdot)$, $\mathcal{Q}_{b,i}(\cdot)$, $i = 1, \ldots, N$. We assume that both transform coders work on vectors having the same size N, and that the signal is not shifted or resampled in between the two transforms. Let $\{\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_P\}$ denote a set of P observed N-dimensional vectors, which are the output of the second transform coder. We assume that the second transform coder is completely known. Indeed, whenever this is not the case, it can be identified with the method previously proposed in [20]. Therefore, without loss of generality, we will consider the set of observed vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$, such that $\tilde{\mathbf{z}}_j = \mathbf{W}_b \tilde{\mathbf{u}}_j$.

Due to quantization, the observed vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$ are constrained to belong to a lattice $\mathcal{L}_z$ described by the basis $\mathbf{B}_z = \text{diag}(\Delta_{b,1}, \ldots, \Delta_{b,N})$. Similarly, the unobserved transform coefficients $\{\tilde{\mathbf{y}}_1, \ldots, \tilde{\mathbf{y}}_P\}$ of the first coder belong to a lattice $\mathcal{L}_y$ described by the basis $\mathbf{B}_y = \text{diag}(\Delta_{a,1}, \ldots, \Delta_{a,N})$. Hence, the unobserved vectors $\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_P\} \in \mathcal{L}_x$ with basis $\mathbf{B}_x = \mathbf{W}_a^{-1}\mathbf{B}_y$.

In this paper we study the problem of determining $\mathbf{B}_x$ from a finite set of $P \geq N$ distinct vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$. That is, we seek to determine the parameters of the first transform coder in a chain of two transform coders, when we observe the output of the second one. This is a much more challenging problem than the one addressed in [20], since the observed vectors do not lie on the lattice $\mathbf{B}_x$. As a consequence, a direct application of the method in [20] to the set of vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$ would identify $\mathbf{W}_b$ rather than $\mathbf{W}_a$.

In this paper, we show how to solve the problem in the case the transform matrices are orthonormal, i.e., $\mathbf{W}_a^T\mathbf{W}_a = \mathbf{I}$ and $\mathbf{W}_b^T\mathbf{W}_b = \mathbf{I}$. For the sake of simplifying the notation, we assume that the same step size is used to quantize the transform coefficients, i.e., $\Delta_{a,i} = \Delta_a$ and $\Delta_{b,i} = \Delta_b$, $i = 1, \ldots, N$. However, this condition can be relaxed.

3. AN ALGORITHM FOR TRANSFORM IDENTIFICATION

The key idea of the proposed method is to denoise the observed vectors before applying our previously proposed algorithm in [20]. To this end, we proceed according to the following steps:

1. We denoise a subset of the observed vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$ in such a way that the resulting $D \leq P$ vectors $\{\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}\}$ lie exactly on a lattice whose basis can be expressed as $\mathbf{B}_{\hat{z}} = \mathbf{R}\mathbf{W}_b\mathbf{B}_x$. The orthonormal matrix $\mathbf{R}$ represents the ambiguity introduced by the denoising procedure.

2. We compute an estimate $\hat{\mathbf{R}}$ of $\mathbf{R}$ by formulating an orthogonal Procrustes problem, which seeks the best matching between the set of denoised vectors $\{\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}\}$ and the observed ones $\{\tilde{\mathbf{z}}_{l_1}, \ldots, \tilde{\mathbf{z}}_{l_D}\}$.

3. We adopt our previously proposed method in [20] to the vectors $\{\hat{\mathbf{x}}_{l_1}, \ldots, \hat{\mathbf{x}}_{l_D}\}$, $\hat{\mathbf{x}}_j = \mathbf{W}_b^{-1}\hat{\mathbf{R}}^{-1}\hat{\mathbf{z}}_{l_j}$, $j = 1, \ldots, D$, to obtain an estimate $\hat{\mathbf{B}}_x$ of $\mathbf{B}_x$.

3.1. Exact recovery of vector distances

The denoising operation indicated in Step 1 exploits the orthogonality of the transform $\mathbf{W}_a$ to determine the quantization step size Δ_a and, consequently, to recover inter-vector distances exactly. Indeed, it is possible to express a constraint on the lengths $\tilde{\delta}_j$ of the unobserved vectors $\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_P\}$ as well as on the lengths $\tilde{\delta}_{j_1,j_2}$ of vector differences. That is,

$$\tilde{\delta}_j^2 = \|\tilde{\mathbf{x}}_j\|^2 = a_j\Delta_a^2, \quad a_j \in \mathbb{N}, \quad j = 1, \ldots, P, \tag{1}$$

$$\tilde{\delta}_{j_1,j_2}^2 = \|\tilde{\mathbf{x}}_{j_1} - \tilde{\mathbf{x}}_{j_2}\|^2 = a_{j_1,j_2}\Delta_a^2,\ a_{j_1,j_2} \in \mathbb{N},\ j_1, j_2 = 1, \ldots, P. \tag{2}$$

In practice, both a_j and a_{j_1,j_2} belong to the subset of integer numbers corresponding to those that can be written as the sum of (up to) N squares. However, when $N \geq 4$, this subset coincides with the set of integer numbers, as proven by Lagrange's four square theorem [24].

In order to determine Δ_a, which is unknown, we first note that $\|\tilde{\mathbf{x}}_i\| = \|\mathbf{z}_i\|$, $i = 1, \ldots, P$, since $\mathbf{W}_b$ is orthonormal. Therefore, Δ_a is the square root of the greatest common divisor of $\{\|\mathbf{z}_1\|^2, \ldots, \|\mathbf{z}_P\|^2\}$. However, we have no access to $\mathbf{z}_i$'s, but only to their quantized versions $\tilde{\mathbf{z}}_i$. The second quantization can be seen as a form of noise. Hence, we estimate Δ_a by using the algorithm in [25], which is a generalized Euclid's algorithm for noisy measurements. Given Δ_a, to recover vector distances, we first note that

$$|\tilde{\rho}_j - \tilde{\delta}_j| = |\|\tilde{\mathbf{z}}_j\|_2 - \|\tilde{\mathbf{x}}_j\|_2| \leq \frac{1}{2}\sqrt{N}\Delta_b\,. \tag{3}$$

The quantization error on the length of a vector can be as large as half of the diagonal of the quantization cell of the second quantizer. In case of lengths of vector differences $|\tilde{\rho}_{j_1,j_2} - \tilde{\delta}_{j_1,j_2}|$, the error can be twice as that, since both vectors are quantized.

If the quantization error induced by the second quantizer is sufficiently small, it is possible to recover the exact values of $\tilde{\delta}_j$ and $\tilde{\delta}_{j_1,j_2}$ from the observed vectors. Indeed, we exploit the fact that $\tilde{\delta}_j^2$ is a multiple of Δ_a (a similar argument holds for $\tilde{\delta}_{j_1,j_2}^2$). To this end, we compute an estimate $\hat{\delta}_j$ of $\tilde{\delta}_j$ as follows:

$$\hat{\delta}_j = \sqrt{\mathcal{Q}_{\Delta_a^2}(\tilde{\rho}_j^2)}\,. \tag{4}$$

Note that any value of $\tilde{\rho}_j$ in the interval $[l_j, u_j]$:

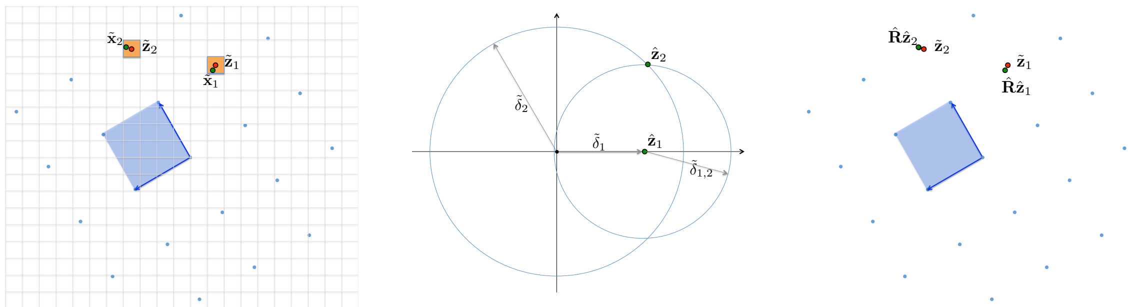
$$\left[\sqrt{\tilde{\delta}_j^2 - \frac{\Delta_a^2}{2}}, \sqrt{\tilde{\delta}_j^2 + \frac{\Delta_a^2}{2}}\right] \tag{5}$$

is quantized to $\tilde{\delta}_j$. Hence, if $|\tilde{\rho}_j - \tilde{\delta}_j| < \min\{u_j - \tilde{\delta}_j, \tilde{\delta}_j - l_j\} = u_j - \tilde{\delta}_j$, it is possible to guarantee that $\hat{\delta}_j = \tilde{\delta}_j$.

Figure 2 illustrates an example in which two N-dimensional (unobserved) vectors $\{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2\}$ are processed by a second transform coder. The coordinate axes are aligned with the basis functions of the known transform $\mathbf{W}_b$. Hence, we display $\{\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2\}$. Figure 3 shows that the corresponding vector lengths, i.e., $\tilde{\rho}_1$ and $\tilde{\rho}_2$, and the distance between vectors, i.e., $\tilde{\rho}_{1,2}$, can be effectively denoised so that $\tilde{\delta}_1$, $\tilde{\delta}_2$ and $\tilde{\rho}_{1,2}$ can be recovered exactly.

We conclude that $\hat{\delta}_j = \tilde{\delta}_j$ whenever the following sufficient condition is satisfied

$$\frac{1}{2}\sqrt{N}\Delta_b^2 < u_j - \tilde{\delta}_j = \sqrt{\tilde{\delta}_j^2 + \frac{\Delta_a^2}{2}} - \tilde{\delta}_j = \tau(\tilde{\delta}_j; \Delta_a). \tag{6}$$



(a) (b) (c)

Fig. 2. A toy example with two unobserved vectors $\{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2\}$ and the corresponding observed vectors $\{\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2\}$, in the coordinate system of the known transform $\mathbf{W}_b$.

3.2. From distances to transform basis functions

Given the denoised values of vector lengths and inter-vector distances, it is possible to write the following set of constraints in the unknown vectors $\{\hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_P\}$

$$\|\hat{\mathbf{z}}_j\|_2 = \sqrt{\mathcal{Q}_{\Delta_a^2}(\tilde{\rho}_j^2)} = \tilde{\delta}_j \quad j \in \mathcal{O}, \tag{7}$$

$$\|\hat{\mathbf{z}}_{j_1} - \hat{\mathbf{z}}_{j_2}\|_2 = \sqrt{\mathcal{Q}_{\Delta_a^2}(\tilde{\rho}_{j_1,j_2}^2)} = \tilde{\delta}_{j_1,j_2} \quad (j_1, j_2) \in \mathcal{D}, \tag{8}$$

where $\mathcal{O}$ denotes the set of indexes of the vectors whose length can be recovered exactly, i.e., those that satisfy (6). $\mathcal{D}$ is similarly defined, denoting the set of indexes of pairs of vectors whose distance can be recovered exactly.

Consider a subset of the unknown vectors $\{\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}\}$ for which the distances from the origin are known, i.e., $l_j \in \mathcal{O}$, and the distances between all pairs are also known, i.e., $(l_{j_1}, l_{j_2}) \in \mathcal{D}$. If $D \geq N$, the position of the vectors can be determined exactly, apart from an ambiguity that can be represented by means of an arbitrary orthonormal transform, which accounts for the rotation around the origin and mirroring with respect to the coordinate axes. Therefore, we proceed in two steps. First, we seek an arbitrary feasible solution $\{\hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_P\}$ of (7). It can be shown that a feasible solution can be found by setting $\hat{\mathbf{z}}_{l_1} = [\tilde{\delta}_{l_1}, 0, \ldots, 0]^T$. Then, remaining vectors are iteratively estimated, by setting $\hat{\mathbf{z}}_{l_j} = [\hat{z}_{l_j,1}, \hat{z}_{l_j,2}, \ldots, \hat{z}_{l_j,j}, \mathbf{0}^T]^T$, $j = 2, \ldots, N$, and finding a solution in a j-dimensional subspace, which can be shown to be formulated as the intersection of a line with a hypersphere centered in the origin and radius $\tilde{\delta}_j$. Details are omitted due to space limitations. Figure 2(b) illustrates a feasible solution corresponding to the toy example in Figure 2(a). The solution is not unique. However, for the problem at hand, it suffices to select arbitrarily a feasible solution.

The resulting vectors $\{\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}\}$ lie exactly on a lattice whose basis can be expressed as $\mathbf{B}_{\hat{z}} = \mathbf{R}\mathbf{W}_b\mathbf{B}_x$. The orthonormal matrix $\mathbf{R}$ represents the ambiguity introduced by the arbitrary choice of the reference system, as well as the arbitrary choice when selecting the feasible solution. In order to solve such ambiguity, we seek an estimate $\hat{\mathbf{R}}$ of the matrix $\mathbf{R}$ by matching the positions of the vectors $\{\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}\}$ to those of the observed vectors $\{\tilde{\mathbf{z}}_{l_1}, \ldots, \tilde{\mathbf{z}}_{l_D}\}$. This can be formulated as the following orthogonal Procrustes problem

$$\begin{aligned} \hat{\mathbf{R}} &= \arg\min_{\mathbf{R}} \|\mathbf{R}\hat{\mathbf{Z}} - \tilde{\mathbf{Z}}\|_F \\ & s.t. \quad \mathbf{R}^T\mathbf{R} = \mathbf{I} \end{aligned} \tag{9}$$

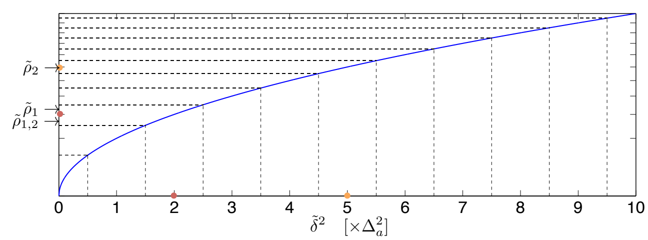


Fig. 3. Square distance values $\tilde{\delta}^2$ are constrained to be integer multiples of Δ_a^2. On the y-axis, it is illustrated the corresponding quantizer applied to denoise observed distances $\tilde{\rho}$.

where $\hat{\mathbf{Z}} = [\hat{\mathbf{z}}_{l_1}, \ldots, \hat{\mathbf{z}}_{l_D}]$ and $\tilde{\mathbf{Z}} = [\tilde{\mathbf{z}}_{l_1}, \ldots, \tilde{\mathbf{z}}_{l_D}]$ and $\|\cdot\|_F$ denotes the Frobenius norm.

Finally, we obtain an estimate of the unobserved vectors as $\hat{\mathbf{x}}_j = \mathbf{W}_b^{-1}\hat{\mathbf{R}}^{-1}\hat{\mathbf{z}}_{l_j}$, $j = 1, \ldots, D$. The set of vectors $\{\hat{\mathbf{x}}_{l_1}, \ldots, \hat{\mathbf{x}}_{l_D}\}$ is guaranteed to lie on a lattice $\hat{\mathbf{B}}_x$, which represents an estimate of the lattice induced by the first transform coder $\mathbf{B}_x$. In our previous work [20] we showed that it is possible to recover $\hat{\mathbf{B}}_x$ with high probability, provided that the number of observed vectors exceeds the dimensionality of the space, that is, $D \geq N + n$. Experimentally, we found that an excess of $n = 6$ vectors is sufficient for the algorithm in [20] to converge to the correct lattice defined by $\hat{\mathbf{B}}_x$, rather than to one of its sub-lattices. The basis functions (row vectors) of the transform $\mathbf{W}_a$ can be obtained as $\hat{\mathbf{w}}_{a,i} = \hat{\mathbf{b}}_{x,i}/\|\hat{\mathbf{b}}_{x,i}\|$, exploiting the orthonormality of the transform.

3.3. Iterative estimation

Given a set of P observed vectors $\{\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_P\}$, the number $D \leq P$ of vectors that can be effectively denoised depends on different factors: i) the statistical distribution of the source from which the original vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_P\}$ are sampled; ii) the quantization step sizes used in the first and second transform coder. A careful analysis of the bound in equation (6) reveals that it is easier to denoise vectors whose length (expressed in Δ_a units) is short, and when $\Delta_b \ll \Delta_a$. In addition, when the dimensionality N increases: i) the average length of vectors increases; ii) the bound in (6) is more stringent, thus enabling to denoise shorter vectors. Therefore, in some cases, it might not be possible to denoise at least $N + n$ vectors at once, since the sufficient condition in (6) may not be satisfied for a large

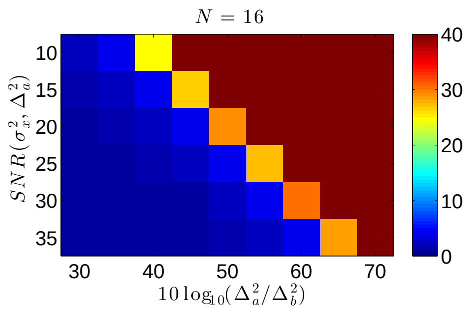


Fig. 4. Number of denoised vectors D.

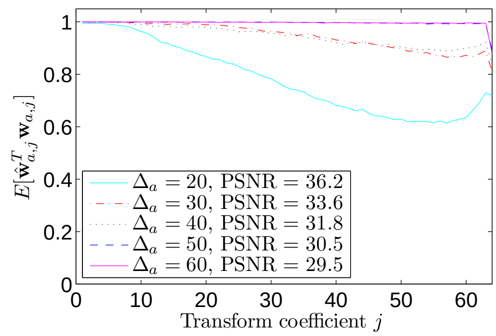


Fig. 5. Estimation accuracy.

enough number of vectors. Hence, we propose to modify the algorithm in such a way that the solution is sought incrementally. The key idea is to start from the largest set of $D^{(0)}$ short vectors that can be efficiently denoised. These are used to estimate a subset of the basis functions of $\mathbf{W}_a$, i.e., $\hat{\mathbf{W}}_a^{(0)} \in \mathbb{R}^{M^{(0)} \times N}$, where $M^{(0)}$ is the dimensionality of the span of the denoised vectors. Then, the remaining vectors are projected in the null-space of $(\hat{\mathbf{W}}_a^{(0)})^T$, and the denosing procedure is applied to the result. The iterative procedure terminates when either $M^{(k)} = N$, or when all vectors are denoised.

4. EXPERIMENTS

We carried out experiments on synthetic data sampled from a N-dimensional i.i.d. Gaussian distribution with variance equal to σ_x^2. This represents the most challenging scenario in our setting, since the isotropic property of the source distribution does not provide any information about the transform being used. For a fixed dimensionality $N = 16$, we sampled $P > N$ vectors and fed them into the processing chain depicted in Figure 1. We varied the signal-to-noise (SNR) ratio due to the quantization of the first quantizer, $SNR \simeq 10\log_{10}\sigma_x^2/(\Delta_a^2/12)$, as well as the ratio between the quantization step sizes of the first and second transform coder. The goal of the experiment is to evaluate in which conditions it is possible to denoise a sufficient number of input vectors. Figure 4 illustrates the number D of vectors which are effectively denoised, when $P = 40$ vectors were observed. It is possible to notice a "cliff" effect, such that the configurations in the top-right part of the figure correspond to cases in which the proposed algorithm is able to find a solution, since $D \geq N + n$. When varying N, we observed that, for a given value of SNR, a smaller value of Δ_b is required when N increases.

We also tested the proposed method on real images. In this case, we considered the 1338 images from the UCID dataset. Each image was compressed with a JPEG-like transform coder. That is, the DCT transform was applied to non-overlapping 8×8 blocks ($N = 64$). Transform coefficients were quantized with a step size in the set $\Delta_a \in \{20, 30, 40, 50, 60\}$. Then, the inverse transform is applied to each block, and the result is rounded to the nearest integer in the pixel-domain. Thus, $\mathbf{W}_b = \mathbf{I}$ and $\Delta_b = 1$. In this case, it was possible to successfully recover the transform by adopting the iterative version of the algorithm illustrated at the end of Section 3, although it was not possible to find at least $D > 64$ vectors that could be denoised at once. Figure 5 shows for each basis function $\mathbf{w}_{a,j}$ associated to one of the DCT coefficients, the quantity $E[\hat{\mathbf{w}}_{a,j}^T \mathbf{w}_{a,j}]$, which indicates the average cosine of the angle between true and estimated basis functions. We observe that when the PSNR is less than 30dB, all basis functions can be estimated with very high accuracy. At higher values of the PSNR, the error increases as more basis functions are estimated. This is due to the fact that, at each iteration, we project in the null-space of the estimated basis functions $(\hat{\mathbf{W}}_a^{(k)})^T$. Hence, the noise in the observed data is due not only to quantization, but also to the non-ideal projection. We also repeated the experiment using a content-dependent KLT transform, estimated independently for each image. The results obtained are very similar to those reported in Figure 5.

5. CONCLUSIONS

In this paper we have studied the challenging problem of estimating both the transform and the quantizer of the first transform coder in a chain of two transform coders. We have shown for the first time that, under specific conditions, the effect of the second transform coder can be removed, so that the problem can be addressed using the algorithm in [20]. Then, we presented an iterative method that successfully solves the problem even when the conditions are not entirely satisfied. Preliminary results on real data show the potential of this new approach also in vector spaces of large dimension.

6. REFERENCES

[1] J. Lukás and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in *Proceedings Digital Forensic Research Workshop*, 2003.

[2] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1003–1017, June 2012.

[3] S. Li, Z. Han, Y. Chen, B. Fu, C. Lu, and X. Yao, "Resampling forgery detection in jpeg-compressed images," in *International Congress on Image and Signal Processing*, October 2010, vol. 3, pp. 1166 –1170.

[4] T. Bianchi and Piva A., "Reverse engineering of double JPEG compression in the presence of image resizing," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, December 2012.

[5] Z. Fan and R. L. de Queiroz, "Identification of bitmap compression history: JPEG detection and quantizer estimation," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 230–235, 2003.

[6] W. Sabrina Lin, Steven K. Tjoa, H. Vicky Zhao, and K. J. Ray Liu, "Digital image source coder forensics via intrinsic fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 460–475, 2009.

[7] S. Milani, M. Tagliasacchi, and S. Tubaro, "Discriminating multiple JPEG compression using first digit features," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012, pp. 2253–2256.

[8] Y. Chen, K. S. Challapali, and M. Balakrishnan, "Extracting coding parameters from pre-coded MPEG-2 video," in *IEEE International Conference on Image Processing (ICIP)*, 1998, pp. 360–364.

[9] H. Li and S. Forchhammer, "MPEG2 video parameter and no reference PSNR estimation," in *Picture Coding Symposium (PCS)*, 2009, pp. 1–4.

[10] W. Luo, M. Wu, and J. Huang, "MPEG recompression detection based on block artifacts," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2008, vol. 6819 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*.

[11] W. Wang and H. Farid, "Exposing digital forgeries in video by detecting double quantization," in *Proceedings of the 11th ACM workshop on Multimedia and security*, New York, NY, USA, 2009, MM&Sec, pp. 39–48, ACM.

[12] M. Tagliasacchi and S. Tubaro, "Blind estimation of the QP parameter in H.264/AVC decoded video," in *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), 2010*, 2010, pp. 1–4.

[13] S. Milani, M. Tagliasacchi, and S. Tubaro, "Identification of the motion estimation strategy using eigenalgorithms," in *IEEE International Conference on Image Processing (ICIP)*, September 2013.

[14] P. Bestagini, A. Allam, S. Milani, M. Tagliasacchi, and S. Tubaro, "Video codec identification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012, pp. 2257–2260.

[15] S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva, M. Tagliasacchi, and S. Tubaro, "An overview on video forensics," *APSIPA Transactions on Signal and Information Processing*, vol. 1, pp. 1–18, 2012.

[16] H.S. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 598–603, July 2003.

[17] M. Winken, P. Helle, D. Marpe, H. Schwarz, and T. Wiegand, "Transform coding in the HEVC test model," in *IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 3693–3696.

[18] C. Yeo, Y. H. Tan, Z. Li, and S. Rahardja, "Mode-dependent transforms for coding directional intra prediction residuals," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 4, pp. 545–554, April 2012.

[19] M. Tagliasacchi, M. Visentini Scarzanella, P. L. Dragotti, and S. Tubaro, "Transform coder identification," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.

[20] M. Tagliasacchi, M. Visentini Scarzanella, P. L. Dragotti, and S. Tubaro, "Transform coder identification based on quantization footprints and lattice theory," *ArXiv e-prints, http://arxiv.org/abs/1211.3869*, November 2012.

[21] R. Neelamani, R. de Queiroz, Z. Fan, S. Dash, and R.G. Baraniuk, "JPEG compression history estimation for color images," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1365–1378, June 2006.

[22] P. Comesaña, F. Pérez-González, and N. Liste, "Quantization lattice estimation for multimedia forensics," in *IEEE International Conference on Image Processing (ICIP)*, September 2013.

[23] M. Visentini Scarzanella, M. Tagliasacchi, and P. L. Dragotti, "Quantisation invariants for transform parameter estimation in coding chains," in *Data Compression Conference*, March 2013.

[24] J. Stillwell, *Elements of number theory / John Stillwell*, Springer, New York :, 2003.

[25] S.D. Casey and B.M. Sadler, "Modifications of the Euclidean algorithm for isolating periodicities from a sparse set of noisy measurements," *IEEE Transactions on Signal Processing*, vol. 44, no. 9, pp. 2260–2272, sep 1996.