

# A MULTI-LAYER IMAGE REPRESENTATION USING REGULARIZED RESIDUAL QUANTIZATION: APPLICATION TO COMPRESSION AND DENOISING

*Sohrab Ferdowsi, Slava Voloshynovskiy, Dimche Kostadinov*

Department of Computer Science, University of Geneva, Switzerland  
{sohrab.ferdowsi, svolos, dimche.kostadinov}@unige.ch

## ABSTRACT

A learning-based framework for representation of domain-specific images is proposed where joint compression and denoising can be done using a VQ-based multi-layer network. While it learns to compress the images from a training set, the compression performance is very well generalized on images from a test set. Moreover, when fed with noisy versions of the test set, since it has priors from clean images, the network also efficiently denoises the test images during the reconstruction. The proposed framework is a regularized version of the Residual Quantization (RQ) where at each stage, the quantization error from the previous stage is further quantized. Instead of codebook learning from the k-means which over-trains for high-dimensional vectors, we show that only generating the codewords from a random, but properly regularized distribution suffices to compress the images globally and without the need to resort to patch-based division of images. The experiments are done on the *CroppedYale-B* set of facial images and the method is compared with the JPEG-2000 codec for compression and BM3D for denoising, showing promising results.

**Index Terms**— dictionary learning, image compression, learning to compress, Vector Quantization, image denoising

## 1. INTRODUCTION

Consider the classical image processing tasks like image compression and denoising. While there exists a wealth of successful methods to address them, the specificity and the intricate optimization in their design hinders their application to more general tasks and setups. For example suppose instead of one single image, we are given a collection of similar-looking images. Can the standard image compression codecs benefit from the shared redundancy to compress the images further? Such a setup is of great practical importance for compression of facial or iris images in biometrics, medical images, or the compression and transmission of very large, but similar-looking images in remote sensing and astronomy. In these cases, the usage of generic codecs like JPEG-2000 whose basis vectors are not adapted to the statistics of images is known to be inefficient.

Take the case of facial images. In spite of the extensive literature in generic image compression, only several learning-based algorithms have studied the compression of facial images. For example, [1] was an early attempt based on VQ. [2] learns the dictionaries based on the K-SVD [3] while [4] uses a tree-based wavelet transform. [5] proposes a codec by using the Iteration Tuned and Aligned Dictionary (ITAD). In spite of their high compression performance, the problem with most of these approaches is that they rely very much on the alignment of images and they are less likely to generalize once the imaging setup is changed a bit. Some of them require the detection of facial features (sometimes manually) and then alignment by geometrical transformation into some canonical form and also a background removal stage.

Similarly for the image denoising tasks, only few methods have benefited from external clean databases of similar images. For example, [6] reports a near  $1dB$  improvement over the BM3D.

On the other hand, one can think of different tasks to be performed jointly. Can more favorable scenarios like the availability of a collection of similar domain-specific images help to compress and denoise images at the same time? As a practical scenario, suppose for example the case where in an object identification system, several exemplar images have been taken with high-quality acquisition systems in the enrollment mode. At query time, however, only low-quality and noisy cameras are available. It is highly desirable to be able to jointly denoise and compress the acquisitions.

The rest of the paper is organized as follows. In section 2, a very brief overview of the general image representation formulation is considered where several relevant cases are quickly reviewed. Section 3 preludes with a review from a problem in rate-distortion theory, namely the reverse water-filling paradigm. This will be used as the core concept behind the proposed Regularized Residual Quantization (RRQ) introduced next. Section 4 conducts experiments on the RRQ algorithm under the image compression and denoising tasks. Finally, section 5 concludes the paper.

## 2. RELATED WORK

Many methods for image representation and dictionary learning can be generalized in the inverse-problem formulation of Eq. 1, where  $X$  contains the data-points (e.g., image patches)  $\mathbf{x}(i)$ 's in its columns. The codebook and the codes can be represented in a matrix form as  $C = [\mathbf{c}_1, \dots, \mathbf{c}_k, \dots, \mathbf{c}_K]^1$  and  $A = [\alpha_1, \dots, \alpha_i, \dots, \alpha_N]$ , respectively, with  $1 \leq i \leq N$  and  $1 \leq k \leq K$ .

$$\begin{aligned} & \underset{C, A}{\text{minimize}} && \|X - CA\|_F^2. \\ & \text{s.t.} && \Omega_C, \Omega_A \end{aligned} \quad (1)$$

$\Omega_C$  and  $\Omega_A$  are a set of constraints on the construction of the codebook and the codes, respectively.

Depending on  $\Omega_C$  and  $\Omega_A$ , the problem of Eq. 1 can be treated in many different ways.<sup>2</sup> For example, under the famous sparsity constraint  $\Omega_A^{\ell_0} : \|\alpha(i)\|_0 \leq s$  or its relaxed version  $\Omega_A^{\ell_1} : \|\alpha(i)\|_1 \leq s$ , the K-SVD algorithm [3] solves it for local minima in an iterative way.

In this work, we follow the VQ-based interpretation of Eq. 1, where, as a general formulation, it is required that:

$$\Omega_A^{\text{VQ}} : \|\alpha_i\|_0 = \|\alpha_i\|_1 = 1.$$

This problem can be solved using the k-means algorithm. However, the lack of structure for this formulation leads to poor generalization performance. To address some of the issues with this simple formulation, Product Quantization (PQ) (e.g., [9] [10]) divides the vectors into several blocks and runs k-means on each of them independently. While PQ can achieve good rate-distortion performance under certain conditions, its lack of design flexibility and the fact that the system should be re-trained for every rate, makes PQ not a suitable solution for image analysis.

As an alternative, RQ is a multi-layer approach that at each layer quantizes the residuals of quantization of the previous layer. While having been extensively studied in the 80's and 90's for different tasks like image coding (e.g., refer to [11], [12] or [9]), its efficiency was limited for more modern applications. In practice, it was not possible to learn codewords for more than a couple of layers.

In this work, we use an approach based on the RQ for which we introduce a pre-processing and an efficient regularization, making it possible to learn arbitrary numbers of layers. Moreover, the introduced regularization makes it possible to go beyond the image patches and work with the high-dimensional image directly. This brings an important advantage for different tasks like image compression. Since the global picture of the image is preserved in the high-dimensional representation, one does not have to encode the relation between similar patches after compression.

<sup>1</sup>Notation: matrix  $X$ , random variable  $X$ , random vector  $\mathbf{X}$  and vector  $\mathbf{x}$ .

<sup>2</sup>See [7] and [8] for detailed reviews and discussions

## 3. PROPOSED FRAMEWORK: RRQ

We first recall a concept from rate-distortion theory which is the quantization of Gaussian independent sources. Although in a slightly different setup than a practical quantization (e.g. being asymptotic), this motivates the core idea behind the RRQ algorithm introduced next in this chapter.

### 3.1. Preliminaries: Quantization of independent sources

The trade-off between the compactness and the fidelity of representation of a signal is classically treated in the Shannon's rate-distortion theory [13]<sup>3</sup>.

A special setup studied in this theory is the rate-distortion for  $n$  independent Gaussian distributed sources,  $X_j$ 's with different variances. Concretely, assume  $X_j \sim \mathcal{N}(0, \sigma_j^2)$ . Define the expected distortion between a random vector and its estimate as  $\mathcal{D} \triangleq \mathbb{E}[d(\mathbf{X}, \hat{\mathbf{X}})]$ , where the distortion between two  $n$ -vectors  $\mathbf{a}$  and  $\mathbf{b}$  is defined as  $d(\mathbf{a}, \mathbf{b}) \triangleq \frac{1}{n} \|\mathbf{a} - \mathbf{b}\|_2^2$ .

Here we ask the question: Given a fixed total distortion  $D$  allowed, i.e.,  $\mathcal{D} \leq D$ , what is the optimal way to divide the distortion (or rate) between these sources such that the overall allocated rate (distortion) is minimized? This can be posed as:

$$\begin{aligned} & \min_{D_j} && \sum_{j=1}^n \max\left[0, \frac{1}{2} \log_2 \frac{\sigma_j^2}{D_j}\right], \\ & \text{s.t.} && \sum_{j=1}^n D_j = D, \end{aligned} \quad (2)$$

where  $D_j$  is the distortion of each source after rate-allocation. The solution to this convex problem is known as the reverse water-filling and is given as:

$$D_j = \begin{cases} \gamma, & \text{if } \sigma_j^2 \geq \gamma \\ \sigma_j^2, & \text{if } \sigma_j^2 < \gamma. \end{cases} \quad (3)$$

where  $\gamma$  is a constant which should be chosen to guarantee that  $\sum_{j=1}^n D_j = D$ .

Denote  $\sigma_{C_j}^2$ , the variance of the codewords for quantization of  $X_j$ . Due to the principle of orthogonality and the independence of dimensions, we have that  $\sigma_{C_j}^2 = \sigma_j^2 - D_j$ . Therefore, according to Eq. 3, the optimal assignment of the codeword variances will be a soft-thresholding of  $\sigma_j^2$  with  $\gamma$ :

$$\sigma_{C_j}^2 = \left(\sigma_j^2 - \gamma\right)^+ = \begin{cases} \sigma_j^2 - \gamma, & \text{if } \sigma_j^2 \geq \gamma \\ 0, & \text{if } \sigma_j^2 < \gamma. \end{cases} \quad (4)$$

This means that the optimal rate-allocation requires that the sources with variances less than  $\gamma$  should not be assigned any rate at all. This, when used in the codebook design, results in sparsity of the codewords which we incorporate in the RRQ algorithm.

<sup>3</sup>Refer to Ch. 10 of [14] for further details of this subsection.

---

**Algorithm 1** Pre-processing

---

**Input:**  $\mathcal{I}_{\text{train}} = \{I(i), \dots, I(N)\}$ , images in the train set.

**Output:**  $X$ , matrix of decorrelated vectors,  $V_1, \dots, V_M$ , rotation matrices for the sub-bands.

```
1: for  $i \in \mathcal{I}_{\text{train}}$  do
2:    $I'(i) \leftarrow \text{2D-DCT}[I(i)]$ 
3:    $\mathbf{x}'(i) \leftarrow \text{zig-zag}[I'(i)]$   $\triangleright$ : zig-zag vectorization
4:   Divide  $\mathbf{x}'(i)$  into  $M$  equal sub-bands:  $\mathbf{x}'(i) = [\mathbf{x}'_1(i), \dots, \mathbf{x}'_M(i)]$ 
5:   for  $m = 1, \dots, M$  do
6:     Stack all  $\mathbf{x}'_m(i)$ 's to get  $X_m$ 
7:   end for
8: end for
9: for  $m = 1, \dots, M$  do
10:  Perform PCA on  $X'_m$  (without dim. reduction) to get  $X_m$  and  $V_m$ , the rotation matrix
11:  Concatenate  $X_m$ 's to get  $X$ 
12: end for
```

---

---

**Algorithm 2** Regularized Residual Quantization

---

**Input:** de-correlated train set  $X$

**Output:** multi-layer codebooks  $C^{(l)}$ 's and index sets  $A^{(l)}$ 's, with  $l = 1, \dots, L$

```
1:  $\hat{X} \leftarrow O$ 
2:  $X^{(0)} \leftarrow X$ 
3:  $[D_1^{(0)}, \dots, D_n^{(0)}] \leftarrow \text{var}[X^{(0)}]$   $\triangleright$  variance per dimension
4: for  $l = 1, \dots, L$  do
5:    $\gamma^* \leftarrow \underset{\gamma}{\text{argmin}} \left( \left| \log_2 K^{(l)} - \sum_{j \in \mathcal{A}_\gamma} \frac{1}{2} \log_2 \frac{D_j^{(l-1)}}{\gamma} \right| \right)$ 
6:   for  $j = 1, \dots, n$  do
7:      $\sigma_{C_j^{(l)}}^2 \leftarrow \left( D_j^{(l-1)} - \gamma^* \right)^+$ 
8:   end for
9:    $S^{(l)} \leftarrow \text{diag}(\sigma_{C_1^{(l)}}^2, \dots, \sigma_{C_n^{(l)}}^2)$ 
10:  for  $k = 1, \dots, K^{(l)}$  do
11:     $\mathbf{c}_k^{(l)} \leftarrow \text{Generate randomly from } \mathcal{N}(\mathbf{0}, S^{(l)})$ 
12:    Concatenate  $\mathbf{c}_k^{(l)}$ 's to get  $C^{(l)}$ 
13:  end for
14:  for  $i = 1, \dots, N$  do
15:     $k^* \leftarrow \text{argmin}_{1 \leq k \leq K^{(l)}} \|\mathbf{x}(i)^{(l-1)} - \mathbf{c}_k^{(l)}\|_2$ 
16:     $\boldsymbol{\alpha}(i)^{(l)} \leftarrow$  all-zero vector with 1 at the  $k^*$  position
17:    Concatenate  $\boldsymbol{\alpha}(i)^{(l)}$ 's to get  $A^{(l)}$ 
18:  end for
19:   $\hat{X}^{(l)} \leftarrow C^{(l)} A^{(l)}$ 
20:   $\hat{X} \leftarrow \hat{X} + X^{(l)}$ 
21:   $X^{(l)} \leftarrow X^{(l-1)} - \hat{X}^{(l)}$ 
22:   $[D_1^{(l)}, \dots, D_n^{(l)}] \leftarrow \text{var}[X^{(l)}]$ 
23: end for
```

---

### 3.2. The RRQ algorithm

Inspired by the setup studied in section 3.1, we argue that after a pre-processing stage, natural images can be represented in a global representation as variance decaying vectors which have independent, or at least uncorrelated dimensions. One might think of the PCA as a simple way to achieve this. However, since the dimensionality of the entire vectorized image is high, apart from the big complexities incurred, there will be too many parameters in the covariance matrix to estimate. Therefore, a global PCA will likely over-fit to the training, largely deviating from the test set. To overcome this issue, we propose the pre-processing in Algorithm 1.

After the PCA rotation matrices are learned from the training set, the same procedure applies to images from the test set. In fact, this pre-processing is a more robust estimation for the global PCA. Instead of  $n^2$  parameters of the direct PCA rotation matrix, with the help of 2D-DCT, this pre-processing has  $m(\frac{n}{m})^2$  parameters to estimate. This is an effective way to trade independence of dimensions for robustness between train and test sets.

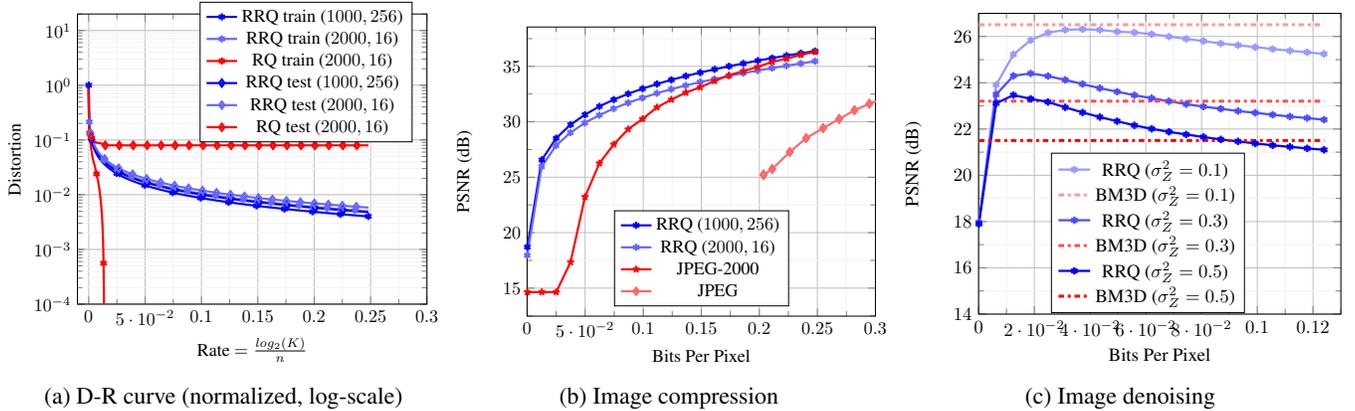
The RRQ framework is introduced in Algorithm 2. For each of the  $L$  layers, given the desired number of codewords,  $K^{(l)}$ , after calculation of the variances of the residuals, the algorithm first finds the optimal  $\gamma^*$  and calculates the optimal variances of the codewords based on Eq. 4 and then randomly generates  $K^{(l)}$  codewords based on these variances. Especially at the first layers, since the data has a very strong decaying character, this makes the codewords very sparse, significantly reducing the complexity and storage cost of the codebooks. The algorithm continues by quantizing the residuals  $\mathbf{x}(i)^{(l-1)}$ 's with the generated codewords and updating the estimations  $\hat{\mathbf{x}}(i)^{(l)}$ 's and the new residuals  $\mathbf{x}(i)^{(l)}$ 's and finishes at the desired  $L$ .

## 4. EXPERIMENTS

We perform the two tasks of image compression and denoising of facial images. For image compression, we compare the performance of our proposed method with the JPEG and JPEG-2000 standards. For denoising, we compare with the BM3D. These are widely considered as baselines for comparison in the literature.

The *CroppedYale-B* set [15] is used which contains 2408 images of size  $192 \times 168$  from 38 subjects. Each subject has between 57 to 64 acquisitions with extreme illumination changes. We choose half of the images for each subject randomly for training and the rest for testing.

We choose two different  $(L, K)$  value-pairs,  $(1000, 256)$  and  $(2000, 16)$ , where  $L$  is the number of layers and  $K$  is the number of codewords per each layer. As described earlier, all codewords are generated randomly according to Eq. 4. Algorithm 1 is used for pre-processing with  $m = 96$  sub-bands. The resulting decorrelated vectors are of size  $n =$



**Fig. 1:** D-R, compression and denoising average performances on the *CroppedYale-B* set.

$192 \times 168 = 32256$  (same as the original images).

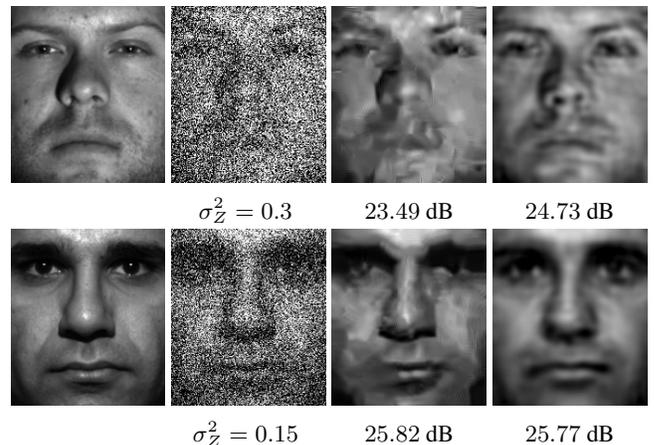
Figure 1a sketches the D-R curve for this set. It is seen that the gap between the training and the test sets for the proposed RRQ is very small, indicating the success of the algorithm in terms of generalization. The non-regularized RQ, on the other hand, while has much lower distortion on the train set, fails to compress the test set at the first several layers.

Fig. 1b shows the results of image compression. These results are averaged over 20 randomly chosen images from the test set. The advantage of the proposed method under this setup over the highly-optimized JPEG-2000 codec is significant, particularly at lower rates. It should be noted that we do not perform any entropy coding over the codebook indices. Further compression improvement can be achieved by entropy coding over the tree-like structure of the codebooks.

The results of image denoising for three different noise levels<sup>4</sup>, averaged over 20 randomly chosen test images are depicted in Fig. 1c. The network is trained over clean images and is exactly the same as the one used for compression. Test images are contaminated with noise and are given as the input to the network for reconstruction. When reconstructing the noisy image, the network uses the priors from the clean images based on which it has been trained. These priors are automatically used in the reconstruction process, serving as a very efficient denoising strategy surpassing (only at highly noisy regimes), the prior-less BM3D.

As the network tries to reconstruct the noisy image with further details, the noise statistics are becoming more present in the reconstructed image, hence degrading the quality. Therefore, depending on the noise variance, the maximum PSNR is somewhere in the middle of the distortion-rate curve. Noisier images have the maximum at lower rates.

Fig. 2 illustrates the denoising quality for two image samples. It is interesting to notice that the BM3D, although producing a smooth image, fails to reconstruct the face contours since it lacks enough priors.



**Fig. 2:** Samples of image denoising. Order of columns: original image, noisy (noise variance), BM3D (PSNR) and RRQ (PSNR).

## 5. CONCLUSIONS

A framework for multi-layer representation of images was proposed where, instead of local patch-based processing, a global high-dimensional vector representation of images is successively quantized within different levels of reconstruction fidelity. As an alternative to the classical RQ framework which is based on k-means, the proposed RRQ along with pre-processing, randomly generates codewords from a regularized and learned distribution. Apart from the many potential advantages of having random codewords, this is shown to lead to efficient quantization with low train-test distortion gaps. The experimental results show interesting promise for different practical scenarios, e.g., when the acquisition devices at the query phase are much noisier than the enrollment cameras. Future works consider using the variance priors to further train the codewords, moreover using entropy coding on the tree of indices for better rate-distortion performance.

<sup>4</sup>Gray values are normalized between 0 and 1.

## 6. REFERENCES

- [1] M. Elad, R. Goldenberg, and R. Kimmel, "Low bit-rate compression of facial images," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2379–2383, Sept 2007.
- [2] Ori Bryt and Michael Elad, "Compression of facial images using the k-svd algorithm," *Journal of Visual Communication and Image Representation*, vol. 19, no. 4, pp. 270 – 282, 2008.
- [3] M. Aharon, M. Elad, and A. Bruckstein, "k -svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov 2006.
- [4] I. Ram, I. Cohen, and M. Elad, "Facial image compression using patch-ordering-based adaptive wavelet transform," *IEEE Signal Processing Letters*, vol. 21, no. 10, pp. 1270–1274, Oct 2014.
- [5] J. Zepeda, C. Guillemot, and E. Kijak, "Image compression using sparse representations and the iteration-tuned and aligned dictionary," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 1061–1073, Sept 2011.
- [6] E. Luo, S. H. Chan, and T. Q. Nguyen, "Adaptive image denoising by targeted databases," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2167–2181, July 2015.
- [7] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, June 2010.
- [8] Julien Mairal, Francis Bach, and Jean Ponce, "Sparse modeling for image and vision processing," *Foundations and Trends in Computer Graphics and Vision*, vol. 8, no. 2-3, pp. 85–283, 2014.
- [9] Allen Gersho and Robert M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Norwell, MA, USA, 1991.
- [10] H. Jegou, M. Douze, and C. Schmid, "Product quantization for nearest neighbor search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 117–128, Jan 2011.
- [11] C. F. Barnes, S. A. Rizvi, and N. M. Nasrabadi, "Advances in residual vector quantization: a review," *IEEE Transactions on Image Processing*, vol. 5, no. 2, pp. 226–262, Feb 1996.
- [12] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: a review," *IEEE Transactions on Communications*, vol. 36, no. 8, pp. 957–971, Aug 1988.
- [13] Claude E Shannon, "Coding theorems for a discrete source with a fidelity criterion," .
- [14] T. Cover and J. Thomas, *Elements of Information Theory 2nd Edition*, Wiley-Interscience, 2 edition, 7 2006.
- [15] A.S. Georghiadis, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.