Scalable Image Coding based on Epitomes

Martin Alain, Christine Guillemot, Dominique Thoreau, Philippe Guillotel

▶ **To cite this version:**

Martin Alain, Christine Guillemot, Dominique Thoreau, Philippe Guillotel. Scalable Image Coding based on Epitomes. IEEE Transactions on Image Processing, 2017, 26 (8), pp.3624-3635. 10.1109/TIP.2017.2702396 . hal-01591504

HAL Id: hal-01591504
https://inria.hal.science/hal-01591504

Submitted on 21 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Scalable image coding based on epitomes

Martin Alain, Christine Guillemot, *Fellow, IEEE,* Dominique Thoreau, and Philippe Guillotel

***Abstract*—In this paper, we propose a novel scheme for scalable image coding based on the concept of epitome. An epitome can be seen as a factorized representation of an image. Focusing on spatial scalability, the enhancement layer of the proposed scheme contains only the epitome of the input image. The pixels of the enhancement layer not contained in the epitome are then restored using two approaches inspired from local learning-based super-resolution methods. In the first method, a locally linear embedding model is learned on base layer patches and then applied to the corresponding epitome patches to reconstruct the enhancement layer. The second approach learns linear mappings between pairs of co-located base layer and epitome patches. Experiments have shown that significant improvement of the rate-distortion performances can be achieved compared to an SHVC reference.**

***Index Terms*—Epitome, super-resolution, scalable image coding, SHVC**

I. Introduction

The latest HEVC standard [1] is among the most efficient codec for image and video compression [2]. However, the ever increasing spatial and/or temporal resolution, bit depth, or color gamut of modern images and videos, coupled with the heterogeneity of the distribution networks, calls for scalable coding solutions. Thus, a scalable extension of HEVC named SHVC was developed [3], [4], which can encode enhancement layers with the scalability features mentioned above by using the appropriate inter-layer processing. Experiments demonstrate that SHVC outperforms simulcast as well as the previous scalable standard SVC [5]. In this paper, we focus on spatial scalability and we propose a novel scalable coding scheme based on the concept of epitome, first introduced in [6], [7]. The epitome in [6] is defined as patch-based appearance and shape probability models learned from the image patches. The authors have shown that these probability models, together with appropriate inference algorithms, are useful for content analysis, inpainting or super-resolution. A second form of epitome has been introduced in [8] which can be seen as a summary of the image. This epitome is constructed by searching for self-similarities within the image using methods such as the KLT tracking algorithm. This type of epitome has been used for still image compression in [9] where the authors propose a rate-distortion optimized epitome construction method. The image is represented by its epitome together with a transformation map as well as a reconstruction residue. A novel image coding architecture has also been described in [10] which, instead of the classical block processing in a raster scan order, inpaints the epitome with in-loop residue coding.

We describe in this paper a novel spatially scalable image coding scheme in which the enhancement layer is only composed of the input image epitome. This factorized representation of the image is then used at the decoder side to reconstruct the missing enhancement layer pixels (*i.e.* not belonging to the epitome) using single-image super-resolution (SR) techniques. Single-image SR methods can be broadly classified into two main categories: the interpolation-based methods [11]–[13] and the example-based methods [14]–[22] which we consider here, focusing on two different techniques based on neighbor embedding [16] and linear mappings [20]. The epitome patches transmitted in the enhancement layer (EL) and the corresponding base layer (BL) patches form a dictionary of pairs of high-resolution and low-resolution patches.

The first method based on neighbor embedding assumes that the BL and EL patches lie on two low and high resolution manifolds which share a similar local geometrical structure. In order to reconstruct an EL patch not belonging to the epitome, a local model of the corresponding BL patch is learned as a weighted combination of its nearest neighbors in the dictionary. The restored EL patch is then obtained by applying this weighted combination to the corresponding EL patches in the dictionary. The second approach based on linear mappings rely on a similar assumption, but directly models a projection function between BL patches and the corresponding EL patches in the dictionary. The projection function is learned using multivariate regression and is then applied to the current BL patch in order to obtain its restored EL version. This super-resolution step reconstructs the full enhancement layer while we only transmit the epitome. The proposed scheme thus allows reaching significant bit-rate reduction compared to traditional scalable coding schemes such as SHVC.

This paper is organized as follows. In section II we review the background on epitomic models. Section III describes the proposed scheme, the epitome generation and encoding at the encoder side, and the epitome-based restoration at the decoder side. Finally, we present in section IV the results compared with SHVC.

II. Background on epitomes

The concept of epitome was first introduced by N. Jojic and V. Cheung in [6], [7]. It is defined as the condensed representation (meaning its size is only a fraction of the original size) of an image signal containing the essence of the textural properties of this image. This original epitomic model is based on a patch-based probabilistic approach. It was shown to be of high "completeness" in [23], but introduces undesired visual artifacts, which is defined as a lack of "coherence". The original epitomic model was also extended into a so-called Image-Signature-Dictionary (ISD) optimized for sparse representations [24].

The aforementioned epitomic models have been successfully applied to segmentation, de-noising, recognition, indexing or texture synthesis. The model of [6], [7] was also used

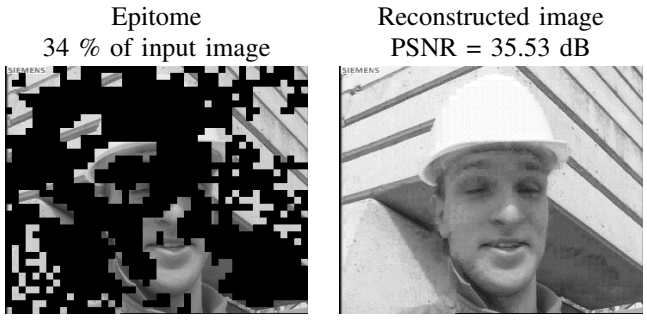


Fig. 1. Epitome of a Foreman frame (left) and the corresponding reconstruction (right).

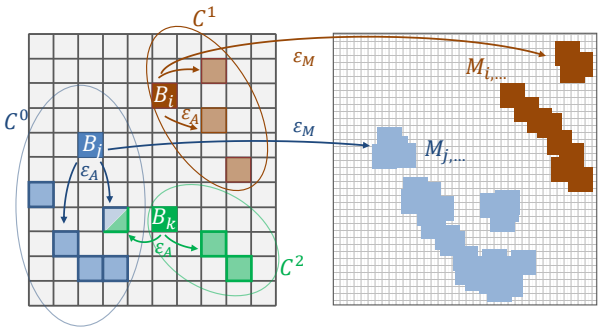


Fig. 2. Two steps clustering-based self-similarities search.

in [25] for intra coding. However, this epitomic model is not designed for image coding applications, and thus have to be coded losslessly, which limits the compression performances.

The epitome construction method used in this paper is thus inspired from the approach introduced in [9]. This epitomic model is dedicated to image coding, and was inspired by the factorized image representation of Wang et al [8]. In this approach the input image I is factored in an epitome E, which is composed of disjoint texture pieces called epitome charts. From this factored representation, a reconstruction I' of the input image is then obtained (see Fig. 1). For that purpose, the input image is divided into a regular grid of non-overlapping blocks B_i (block-grid) and each block is reconstructed from an epitome patch. A so-called assignation map links the patches from the epitome to the reconstructed image blocks. This epitomic model is obtained through a two-step procedure which first searches for the self-similarities within the input image, and then iteratively grows the epitome charts. The second step for creating the epitome charts is notably based on a rate-distortion optimization (RDO) criterion, which minimizes the distortion between the input and the reconstructed image together with the rate of the epitome, evaluated as its number of pixels.

A still image coding scheme based on this epitomic model is also described in [9], where the epitome and its associated assignation map are encoded. The reconstructed image can thus be used as a predictor, and the corresponding prediction residue is further encoded. The results show that the scheme is efficient against H.264 Intra. However, the coding performances of the assignation map are limited, which reduces the overall rate-distortion (RD) gains.

A novel coding scheme was thus proposed in [10], where only the epitome is encoded. The blocks not belonging to the epitome are then predicted in an inpainting fashion, together with an in-loop encoding of the residue. The prediction tools notably include efficient template-based neighbor embedding techniques such as the Locally Linear Embedding (LLE) [26], [27]. The results show that significant bit-rate savings are achieved with respect to H.264 Intra.

In the next section, we describe the proposed scheme, which can be seen as an extension of the latter work to scalable coding.

III. Epitomic enhancement layer for scalable image coding

In this section, we describe a spatially scalable coding scheme in which the enhancement layer only consists in an epitome of the input image. Although the principle could be extended to more layers, for sake of simplicity we describe and validate a two-layer coding scheme in which a low resolution version of the input image is first compressed using HEVC and transmitted as the base layer. The epitome transmitted as the enhancement layer (that we denote epitomic enhancement layer hereafter) is then used at the decoder side to restore the complete image at the enhanced resolution and quality. More precisely, the epitomic EL is encoded using SHVC. No EL residue is transmitted for the pixels not belonging to the epitome, which are then restored using local learning-based super-resolution techniques exploiting pairs of patches from the decoded epitomic EL and the corresponding BL patches. Note that this restoration step is only applied as post-processing after decoding of the epitomic EL, and is not used as an inter-layer prediction tool. The proposed scheme is depicted in Fig. 3. We first describe the epitome construction and encoding method and then explain how the epitome and the base layer are jointly used to reconstruct the image at the full resolution.

A. Epitome generation

The epitome generation method is inspired from the one described in [9] and consists in first searching for patch similarities within the image, the matching patches being then used to construct the so-called epitome charts. Unlike [9], the search for best matching patches is done using a fast clustering-based technique [28], and the criterion used for the chart creation is a simple error minimization instead of a rate-distortion optimization criterion. We observed that the RDO criterion had limited impact on the proposed scheme compression performances.

1) Self-similarities search: The goal of this step is to find for each block $B_i \in I$ a list of matching patches $ML(B_i) = \{M_{i,0}, M_{i,1}, ...\}$, such that the mean square error (MSE) between a block and its matching patches is below a matching threshold ε_M. This parameter eventually determines the size of the epitome, and several values are considered in the experiments (see Table III). In this paper, the lists of

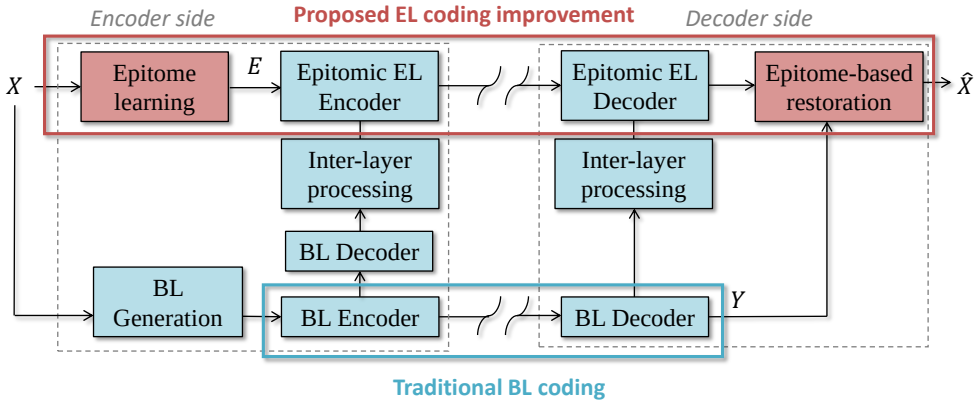


Fig. 3. Proposed scheme for scalable image coding. At the encoder side, an epitome of the input image is generated, and encoded as the enhancement layer. At the decoder side, the enhancement layer patches not contained in the epitome are reconstructed from the base layer and the epitomic enhancement layer.

matching patches are obtained through a two steps clustering-based approach illustrated in Fig. 2. The first step consists in grouping together similar blocks into clusters, so that the distance from a block to the centroid of its cluster is below an assignation threshold ε_A. In practice, this threshold is set to $\varepsilon_A = 0.5 * \varepsilon_M$. In the second step, a list of matching patches is computed for each cluster by finding patches whose MSE is inferior to ε_M with respect to the block closest to the cluster centroid. This list of matching patches is then assigned to all the blocks in the cluster.

From all the lists of matching patches $ML(B_i)$, we can then compute reverse lists $RL(M_j) = \{B_j, B_l, ...\}$ which indicate the set of image blocks that can be represented by each matching patch. Next, we describe how these lists are used to build the epitome charts.

2) Epitome charts creation: The epitome charts are iteratively grown, and both the initialization and the iterative extension of an epitome chart are based on a criterion which minimizes the error between the input image I and the reconstructed image I'.

Formally, we denote $\Delta EC_m, m = 0, ..., M-1$ a set of candidate regions to add to the epitome, where M is the number of candidates. The set of all candidate regions correspond to the set of all the matching patches computed at the previous step which are not yet included in the epitome. When initializing a new epitome chart, a valid candidate region is a matching patch which is not yet in an epitome chart and is spatially disconnected from any existing epitome chart. On the contrary, when extending an epitome chart EC, a valid candidate region is a matching patch which is not yet in an epitome chart and overlaps with EC (see Fig. 4). The actual region added to the epitome ΔEC_{opt} is obtained by minimizing the following criterion:

$$\Delta EC_{opt} = \underset{\Delta EC_m}{\arg\min}(MSE(I, I'_m)) \quad (1)$$

where I'_m is the reconstructed image when the candidate region ΔEC_m is added to the epitome, and the MSE function computes the mean square error between I'_m and the source image I. The reconstructed image I'_m comprises the blocks reconstructed from the existing epitome charts and the new blocks contained in the list $RL(\Delta EC_m)$. During the extension of an epitome chart, additional reconstructed blocks can be obtained by considering the so-called inferred blocks, which are the potential matching patches that can overlap between the current chart EC and the extension ΔEC_m (see Fig. 4). Note that for the pixels of I'_m which are not reconstructed, the MSE can not be computed directly. In our implementation, we assign the maximal MSE value to these pixels. This tends to favor the selection of a candidate region ΔEC_m which reconstructs large regions in I'_m, and thus speed up the epitome chart creation. We illustrate the principle of the epitome chart extension and initialization in Fig. 4.

The extension of an epitome chart stops when no more valid candidate regions can be found. A new epitome chart is then initialized at a new location. The global process stops when the whole image I' is reconstructed.

B. Epitome encoding

The goal being to encode only the epitome as an enhancement layer in a scalable coding scheme, the epitome charts are padded to a fixed block size corresponding to a coding unit (CU) size of the encoder. We use in our experiments the SHVC standard, and the epitome blocks are padded to 8×8 or 16×16 blocks (see Table I). These blocks are then coded using the inter-layer prediction mechanism of SHVC, followed by the SHVC transform and quantization coding tools.

Since the relevant texture information we want to transmit is only contained in the epitome blocks, the blocks not belonging to the epitome are replaced by the up-sampled decoded base layer to create the entire epitomic EL, as shown in Fig. 5. In practice we use the same up-sampling filter as in the SHVC inter-layer processing. When encoding this epitomic EL with SHVC, the bit-rate of most non-epitome blocks is thus negligible thanks to exact inter-layer prediction. This allows to reduce the bit-rate allocated to these blocks without having to modify the SHVC syntax. Note that due to the recursive quad-tree partitioning, some non-epitome blocks may not be exactly predicted if the CU size is different from the epitome block size. We observed in practice that this concerns few blocks, and in fact the epitomic EL bit-rate decreases with the epitome size compared to the baseline SHVC EL (see section IV-B).

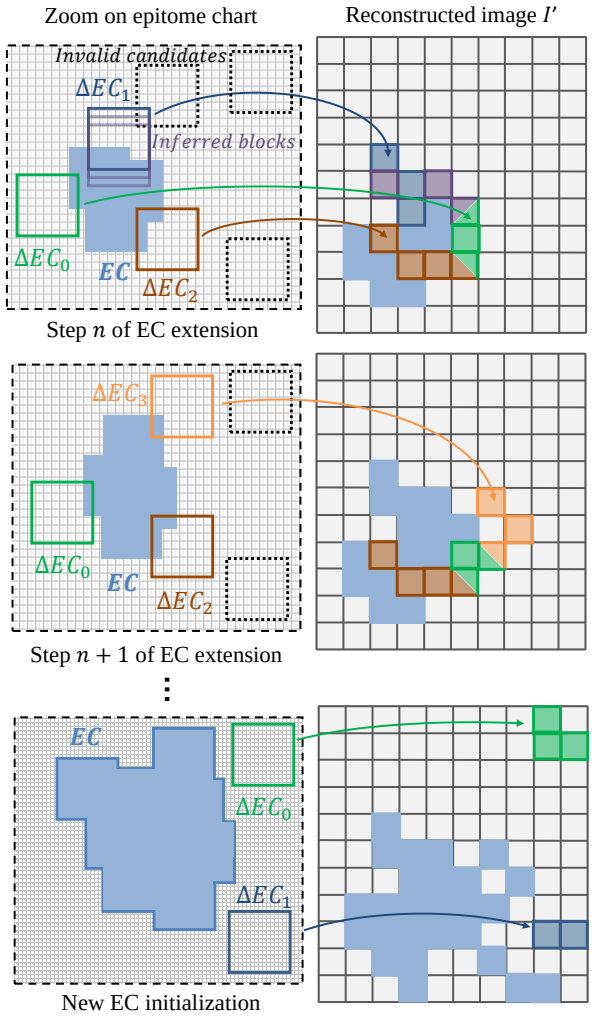


Fig. 4. Example of epitome chart extension. At a step n, we consider 3 valid candidate regions ΔEC_m to extend a current epitome chart EC. These candidate regions correspond to matching patches position found in the original image during the self-similarities search which spatially overlap with EC. Invalid candidates which do not overlap with EC are shown as dashed black blocks. The candidate region minimizing the criterion of Eq. 1 (ΔEC_1 in this example) is added to EC at the step $n+1$. The image blocks reconstructed by ΔEC_1 are added to I'. In this example a new valid candidate appears at step $n+1$. When no more valid candidate region is available to grow the current epitome chart, a new epitome chart is initialized. Candidate regions for the new epitome chart initialization are spatially disconnected from the previously grown epitome chart. The candidate actually selected for the initialization is again obtained by minimizing the criterion of Eq. 1. The process iterates until I' is fully reconstructed.

When decoding the epitomic EL, the non-epitome blocks consisting of up-sampled decoded BL texture need to be restored (see next section). A binary map is sent to the decoder in order to ensure that all the non-epitome blocks are identified prior to the restoration step. We observed in practice that the corresponding bit-rate does not account for a significant part of the overall bit-rate (about 2% in average), and could be even reduced using for example entropy coding.

C. Epitome-based restoration

After decoding the epitomic EL, the non-epitome part of the enhancement layer is processed by considering $N \times N$ overlapping patches, separated by a step of s pixels in both rows and columns. After restoration, when several estimates

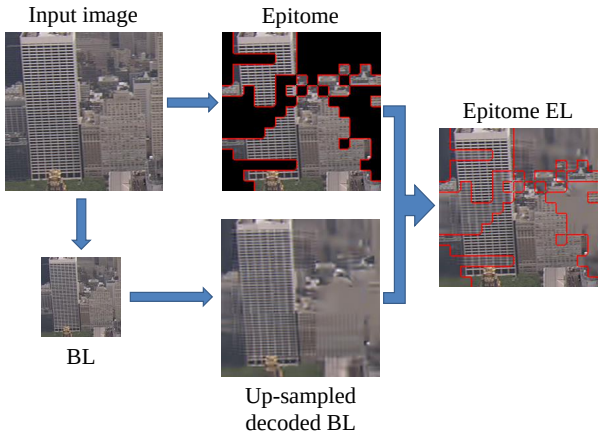


Fig. 5. Epitomic EL creation. The non-epitome blocks of the epitomic EL are replaced by the up-sampled decoded BL (best viewed zoom in). Thanks to exact inter-layer prediction, these blocks are thus encoded at a negligible bit-rate.

are obtained for a pixel, they are averaged in order to obtain the final estimate. Note that before performing the restoration, the BL image is up-sampled to the resolution of the EL using the inter-layer processing filter.

The restoration methods described below are derived from local learning-based SR methods [16], [17], [20]–[22], and can be summarized in the three following steps: K-NN search, learning step, and reconstruction step. These steps are shown in Fig. 6, and described in details below.

1) K-NN search: If we consider the current patch to be processed y, we first search for its K-NN BL patches, within search windows corresponding to the epitome charts locations (see Fig. 6). The K-NN BL patches $y_i, i = 1 \dots K$ are then stored in a matrix $\mathbf{M}_y$ which contains in its columns the vectorized patches. For each neighbor y_i, we have a corresponding EL patch x_i in the epitome, which is stored in a matrix $\mathbf{M}_x$. We thus obtain BL/EL pairs of training patches. In classical SR applications, the pairs of training patches are obtained from a dictionary, which construction is a critical step [14]–[16], [18], [19], [21], [22]. Since here the patches in the epitome are representative of the full image, we can consider that they constitute a suitable dictionary to perform the local learning-based restoration.

Next, we present the learning and reconstruction steps, which exploit the correlation between the pairs of training patches to perform learning-based restoration. We describe two methods to restore the missing EL patches, inspired by SR techniques based on neighbor embedding (NE) [16], [17], [21] and linear mappings [20]–[22], but any other learning-based method could be included in the proposed scheme. Note that many SR methods can be improved using iterative back-projection [29], which enforces the high resolution reconstructed image to be consistent with the input low resolution image. However, this technique will not be considered in the proposed scheme, as it tends to propagate quantization noise from the BL image to the EL reconstruction.

2) Epitome-based Locally Linear Embedding: First, we describe a method relying on LLE, denoted "epitome-based

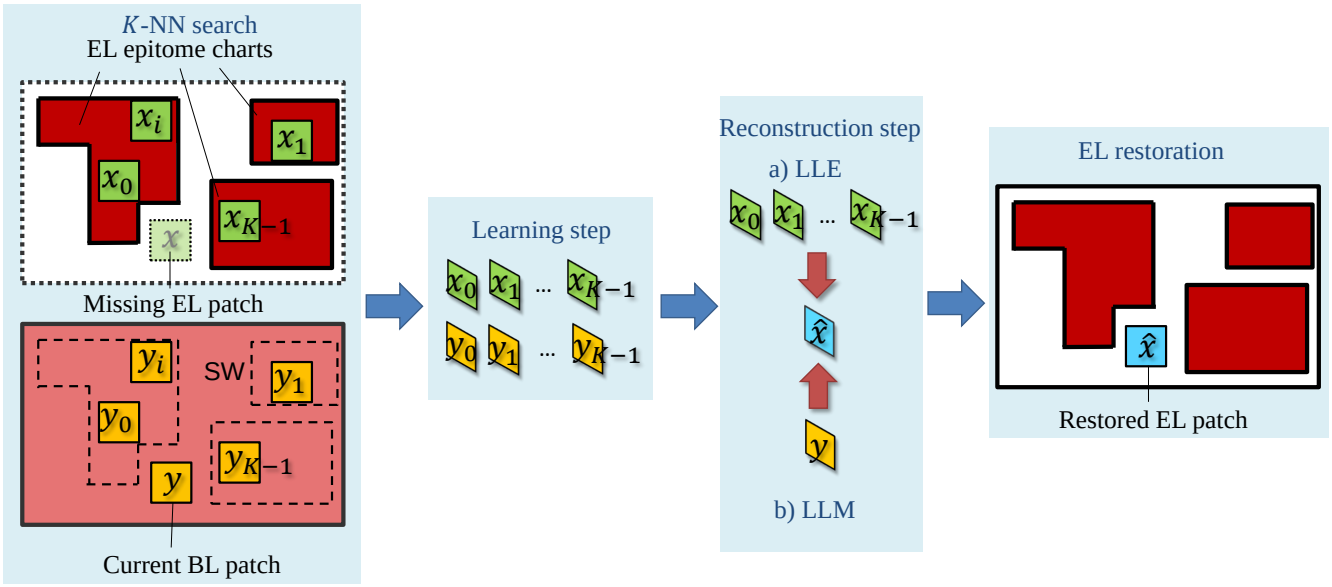
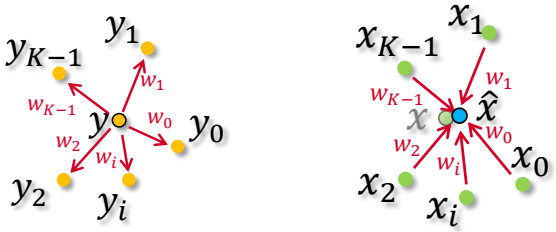


Fig. 6. The K nearest neighbors of the current BL patch are found in search windows (SW) corresponding to the epitome charts. The BL/EL pairs of patches can then be exploited to restore the current missing EL patch.

Locally Linear Embedding" (E-LLE). Similarly to other NE-based restoration techniques, we assume that the local geometry of the manifolds in which lie the BL and EL patches is similar (see Fig. 7). Using LLE, we first learn the linear combination of the K-NN BL patches which best approximate the current patch, and then apply this linear combination to the corresponding EL patches in order to obtain a good estimate of the missing EL patch.



a) Learning: estimate the current BL patch as linear combination of its K-NN BL patches.

b) Reconstruction: Apply the weights of the linear combination learned previously on the K-NN corresponding EL patches to obtain the restored EL patch.

Fig. 7. Two steps local learning-based restoration based on LLE.

Let W be the vector containing the combination weights $w_i, i = 1 \dots K$. The weights are obtained by solving the following equation:

$$\min_W ||y - \mathbf{M}_y W||_2^2 \text{ s.t. } \sum_{i=1}^{K} w_i = 1 \tag{2}$$

The weights vector W is computed as:

$$W = \frac{\mathbf{D}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{1}}. \tag{3}$$

The term $\mathbf{D}$ denotes the local covariance matrix (*i.e.*, in reference to y) of the K-NN stacked in $\mathbf{M}_y$, and $\mathbf{1}$ is the column vector of ones. In practice, instead of an explicit inversion of the matrix $\mathbf{D}$, the linear system of equations $\mathbf{D}W = \mathbf{1}$ is solved, then the weights are rescaled so that they sum to one.

The restored EL patch is finally obtained as:

$$\hat{x} = \mathbf{M}_x W. \tag{4}$$

In practice, several versions of this method can be derived, e.g. by using another NE-based technique such as non-negative matrix factorization [30], or by adapting the weights computation as in the non-local mean algorithm (exponential weights) [31]. However, with such methods the weights are only computed based on the BL patches. In the next section, we propose a method which aims at better exploiting the correlation between the pairs of training patches, based on linear regression.

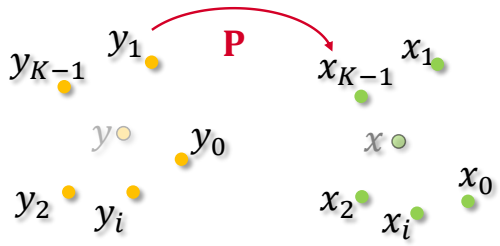
3) Epitome-based Local Linear Mapping: We describe here a method based on linear regression, that we denote "epitome-based Local Linear Mapping (E-LLM)". We want to further exploit the correlation between the pairs of training patches by directly learning a function mapping the BL patches to the corresponding EL patches (see Fig. 8). This function can then be applied on the current patch to restore the EL patch.

The mapping function is learned using multivariate linear regression. The problem is then to search for the function represented by the matrix $\mathbf{P}$ minimizing:

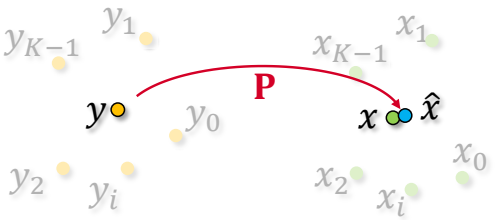
$$E = ||\mathbf{M}_x^T - \mathbf{M}_y^T \mathbf{P}^T||^2 \tag{5}$$

which is of the form $||\mathbf{Y} - \mathbf{XB}||^2$ (corresponding to the linear regression model $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$). The minimization of Eq. (5) gives the least squares estimator

$$\mathbf{P} = \mathbf{M}_x \mathbf{M}_y^T (\mathbf{M}_y \mathbf{M}_y^T)^{-1} \tag{6}$$



a) Learning: a mapping function is learned between the pairs of BL/EL patches using multivariate linear regression.



b) Reconstruction: the function learned previously is applied on the current BL patch in order to obtain the restored EL patch.

Fig. 8. Two steps local learning-based restoration based on linear regression.

We finally obtain the restored EL patch as:

$$\hat{x} = \mathbf{P}y \tag{7}$$

Now that we have formally defined the proposed methods, we study their performances in the next section.

IV. Simulations and Results

A. Experimental conditions

The experiments are performed on the test images listed in Table I, obtained from the HEVC test sequences. This set of images was specifically selected with various characteristics in terms of textures and at different resolutions. The base layer images are obtained by down sampling the input image with a factor 2×2, using the SHVC down-sampling filter available with the SHM software (ver. 9.0) [32]. The BL images are encoded with HEVC, using the HM software (ver. 15.0) [33].

We then use the SHM software (ver. 9.0) [32] to encode the corresponding enhancement layers. Thanks to the hybrid codec scalability feature of SHVC, the decoded BL images are first up-sampled using the separable 8-tap SHVC filter $(-1, 4, -11, 40, 40, -11, 4, -1)$, and directly used as input to the SHM software. Both layers are encoded with the following quantization steps: QP = 22, 27, 32, 37.

In this section we aim to evaluate the performances of the proposed scheme depending on the size of the epitome, which is determined by the matching threshold parameter. Thus for each image, 3 to 4 matching threshold values ε_M are considered in order to yield similar epitome sizes, going from 30% to 90% of the input image size. The selected matching thresholds and corresponding epitome sizes are shown in Table III.

The epitome-based restoration is performed using $N \times N = 8 \times 8$ overlapping patches, with an overlapping step $s = 3$. We set the number of nearest neighbors to $K = 20$.

TABLE I
Test images

Class	Image	Size	Epitome block size
B	BasketballDrive	1920×1056	16×16
B	Cactus	1920×1056	16×16
B	Ducks	1920×1056	16×16
B	Kimono	1920×1056	16×16
B	ParkScene	1920×1056	16×16
B	Tennis	1920×1056	16×16
B	Terrace	1920×1056	16×16
C	BasketballDrill	832×480	8×8
C	Keiba	832×480	8×8
C	Mall	832×480	8×8
C	PartyScene	832×480	8×8
D	BasketballPass	416×240	8×8
D	BlowingBubbles	416×240	8×8
D	RaceHorses	416×240	8×8
D	Square	416×240	8×8
E	City	1280×704	16×16

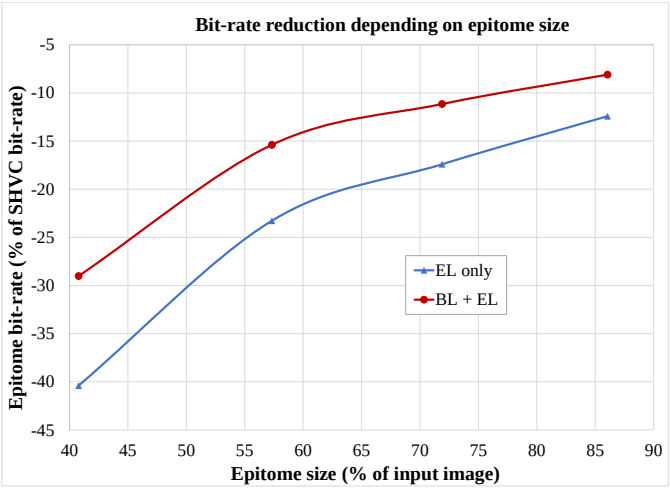


Fig. 9. Average epitome bit-rate gains compared to SHVC depending on the epitome size. Note that these results do not reflect the compression performances as the image quality is not taken into account.

B. Bit-rate reduction

We first evaluate in this section the bit-rate reduction obtained with the proposed scheme compared to the SHVC reference. Note that to assess the actual compression performances the distortion must also be taken into account. Such RD performances are given in the next section.

We give in Fig. 9 the average bit-rates gains (in %) depending on the epitome size, first for the EL only and then for both BL and EL (including the binary map bit-rate). These results show that decreasing the size of the epitome does reduce the bit-rate of the EL, and consequently the overall bit-rate (BL and EL). However, we show in the next section that the bit-rate reductions observed in Fig. 9 are not directly converted into RD performances improvement, as there is a trade-off between the bit-rate reduction and the epitomic EL quality which depends on the epitome size.

C. Rate-distortion performances

We assess in this section the RD performances of the proposed scheme against the SHVC reference.

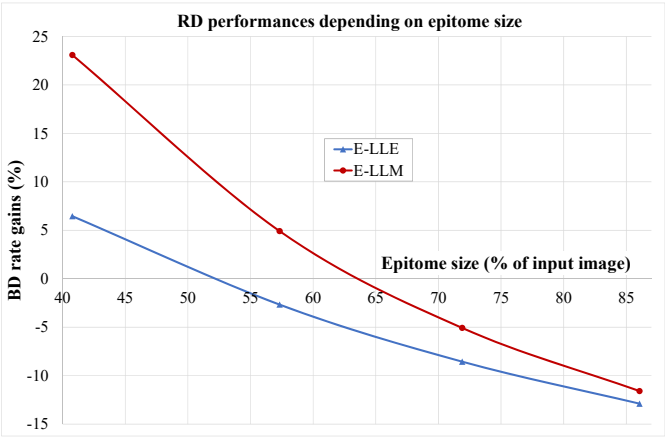


Fig. 10. Average RD performances of the different restoration methods against SHVC depending on the epitome size.

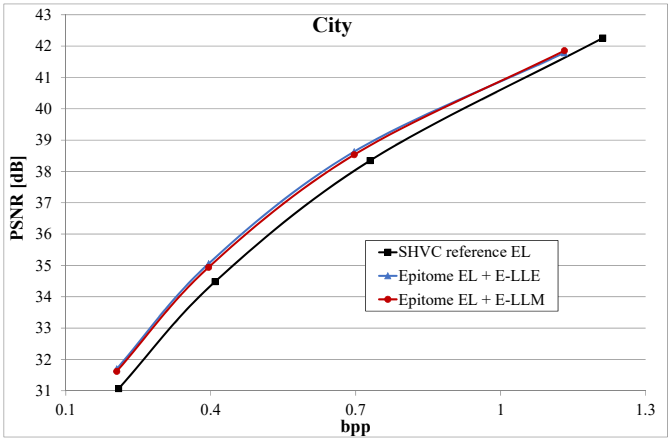


Fig. 11. RD performances of the City image for both E-LLE and E-LLM methods, epitome size = 91.59% of input image.

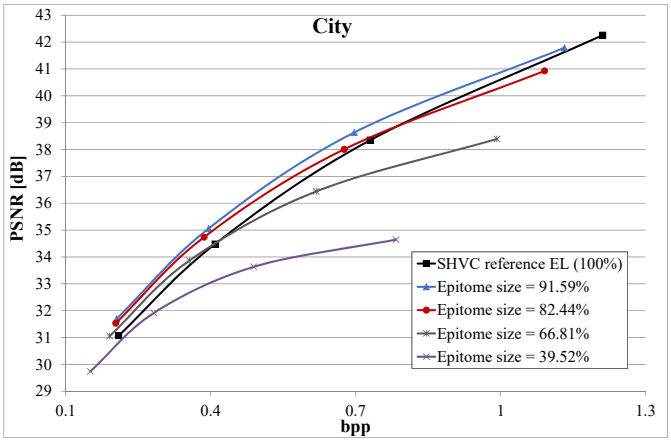


Fig. 12. RD performances of the City image using the E-LLE method, with different epitome sizes.

The distortion is evaluated using the PSNR of the decoded EL, while the rate is calculated as the sum of both BL and EL rates (including the binary map bit-rate). The RD performances are computed using the Bjontegaard rate gains measure (BD-rates) [34] over the four QP values.

We show in Fig. 10 the BD-rates averaged over all sequences depending on the epitome size. The complete results are given in Table III. We can see that significant bit-rate reduction can be achieved compared to SHVC, with more than 10% bit-rate reduction in average at best, and up to about 20% for images like Ducks, Kimono, or Tennis. The E-LLE performs better than the E-LLM method, however, for the best performances, both methods perform similarly.

Note that to obtain improved RD performances against the SHVC reference, a compromise on the epitome size which balances the bit-rate and the epitomic EL quality has to be found. In fact, decreasing the epitome size reduces the bit-rate (as shown in section IV-B), but it also lowers the quality of the epitomic EL compared to SHVC. When the epitome size is too small, this quality loss is not effectively compensated by the restoration step, since only a reduced set of BL/EL patches is provided, while more BL patches need to be processed. This phenomenon is clearly observed in Figs. 10 and 12. Thus, we can see an improvement of the RD performances when the epitome grows. We show in Fig. 10 that RD performances improvement can be obtained for epitome sizes ranging from about 65% to 85% of the input image for the E-LLM method, and about 55% to 85% of the input image for the E-LLE method. However, increasing the epitome size up to 100% of the input image would not lead to better RD performances, as this corresponds to the SHVC EL reference.

In order to better understand the performances of the proposed methods depending on the bit-rate, we show in Fig. 11 the RD curve of the City image for its best performances (biggest epitome), which behavior is representative of the set of test images. We can see that at high bit-rates (QP=22), the bit-rate of the proposed scheme is especially reduced compared to the SHVC reference EL. However, even with the proposed epitome-based restoration, we observe a loss of quality. At low bit-rates (QP=37), the bit-rate of the proposed method is less reduced compared to the SHVC reference EL, but the epitome-based restoration yields a better quality. This behavior explains the overall significant bit-rate reduction we can achieve with the proposed scheme.

In addition, we show in Fig. 12 the RD curves of the LLE-based restoration for the City image depending on the epitome size. We can see that for smaller epitome sizes, at high bit-rates (QP=22), even though the bit-rate is considerably reduced compared to the SHVC reference EL, the quality loss does not allow an improvement in the RD performances, which corroborates our previous analysis.

We give in Figs. 15 and 16 visual examples of the enhancement layers for the City and Cactus images. Note that these examples were chosen for their visual clarity, and do not necessarily correspond to the best RD performances. In order to demonstrate the relevance of the epitome-based restoration step, we show on the top row the epitomic EL before restoration, and on the bottom row the corresponding EL after applying the E-LLE and E-LLM methods. Before restoration, the blocks not belonging to the epitome are particularly visible, as they are directly copied from the up-sampled decoded BL, and clearly lack high frequency details. An obvious improvement of the quality can be observed after restoration for the high-frequency pseudo-periodic textures, such as the building of Fig. 15, or the calendar of Fig. 16.

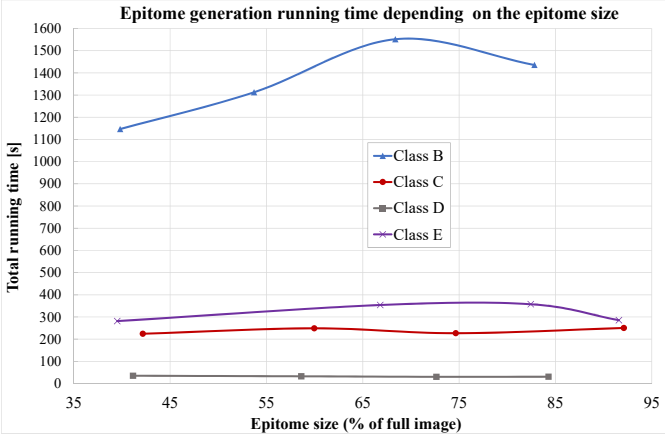


Fig. 13. Epitome generation running time depending on the epitome size for different image classes.

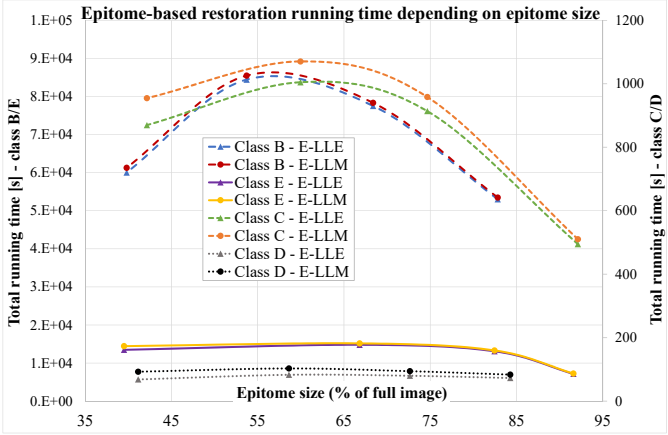


Fig. 14. Running time of the different epitome-based restoration methods depending on the epitome size for different image classes.

Although the E-LLM usually yields lower PSNR than the E-LLE method, we can see that it can perform visually better on high-frequency stochastic textures such as in the highlighted red rectangle of Fig. 16.

D. Elements of complexity

We give in this section some indications about the complexity, evaluated as the running time of the proposed methods. We evaluate the complexity depending on the epitome size and input image size. The results are averaged for each image class (which corresponds to an image size, see Table I). Note that our implementations are not optimized, especially the restoration methods were implemented in Matlab, while the epitome generation algorithm was implemented in C++. However, we give the running times of the SHVC codec for comparison.

We give in Fig. 13 the complexity of the epitome generation. We can see that the epitome generation running time mainly varies depending on the input image size, while the epitome size has a limited impact. For the biggest epitomes, which correspond to the best RD results, we observed that in average 50% to 90% of the complexity is dedicated to the self-similarities search step. As this step is highly parallelizable, the total running time could be reduced by using a parallel implementation, e.g. on GPU. The corresponding running times for the SHVC encoder (BL and EL) are about 20s, 5s, 1s, and 10s for classes B, C, D and E respectively.

We show in Fig. 14 the complexity of the epitome-based restoration step. The processing time is similar for both E-LLE and E-LLM methods, and obviously increases with the size of the image. However, we can observe that overall the complexity is reduced for the biggest epitomes, which interestingly corresponds to the best RD performances. In fact, when transmitting bigger epitomes, less patches not belonging to the epitome have to be processed. The corresponding running times for the SHVC decoder (BL and EL) are about 0.33s, 0.14s, 0.07s, and 0.23s for classes B, C, D and E respectively.

The simulations showed that in average, about 95% of the restoration complexity is dedicated to the K-NN search. The K-NN search was performed with the Matlab *knnsearch* function, based on Kd-tree [35], [36]. In order to reduce the total running time, the complexity of the K-NN search could be further reduced by using more advanced approximate nearest neighbor search [37]–[40] or parallel implementation, possibly on GPU [41].

E. Extension to scalable video coding

The work presented in this paper is dedicated to scalable single image coding, however a straight extension to scalable video coding can be considered by applying the proposed method to each frame of the sequence. Preliminary experiments are conducted on a set of 3 test sequences, consisting of 9 frames with a CIF resolution in order to limit the computation time. The SHVC configuration is set to Random Access. The epitomes are generated using one matching threshold value $\varepsilon_M = 7.0$. In order to exploit the temporal redundancies, the K-NN search step is performed in the epitomes of the two closest frames in addition to the current one.

We show the RD performances measured with the Bjontegaard rate gains in Table II. These preliminary results indicate that the proposed scheme is also expected to bring significant bit-rate reduction when extended to full video sequences. These results are not obvious to predict, since the inter layer prediction is here also competing with inter frame prediction modes, which are much more efficient than the intra prediction modes.

TABLE II
BJONTEGAARD RATE GAINS AGAINST SHVC

Sequence	Epitome size (%) (averaged over all frames)	BD rate gains (%) E-LLE	E-LLM
City	56.66	-26.56	-26.59
Macleans	79.93	-2.24	-2.15
Mobile	82.22	-10.12	-10.20

V. Conclusion and Future Work

We propose in this paper a novel scheme for scalable image coding, based on an epitomic factored representation of the enhancement layer and appropriate restoration methods at the

decoder side. Significant bit-rate reduction is achieved for the spatial scalable application when compared to the SHVC reference EL. These achievements were possible because of the specific epitomic model we used, which provides relevant texture information and is especially suitable for scalable encoding. Note that improvements of the standard tools have been recently proposed, such as the generalized inter-layer residual prediction [42]–[44], or enhanced in-loop prediction mechanism for the EL [45], [46]. The proposed approach is compatible with such improvement of scalable coding schemes, as the coding of the epitome as an enhancement layer would be improved as well.

The proposed scheme could be improved by studying alternative restoration approaches. For instance, different regression could be considered for the E-LLM instead of the direct least-square approach, such as (Kernel) Ridge Regression [21], [47]. Alternatively, the restoration step at the decoder side could be considered as an inpainting problem with prior knowledge on the "holes" to be filled in the form of low resolution patches. Inpainting has been extensively studied over the last decades (see [48] and reference therein for more details), and the exemplar-based multi-scale approaches [49], [50] are well suited in our context.

Future work also includes the adaptation of the proposed scheme for scalable video coding, as preliminary results indicate promising RD performances. The epitomes were here generated separately for each frame. The RD performances would benefit from an epitomic model which takes into account the temporal redundancies. Furthermore, the image self-similarities in the current epitome are found using a single block matching algorithm, while the application at the decoder side is based on multi-patches techniques. New epitomic models have been designed to take into account the epitome application in the generation process. For example, an epitome is proposed in [51] for multi-patches super-resolution, which showed that a more compact representation can be obtained for a similar image reconstruction quality. Such model could thus be considered in the proposed scheme in order to improve the RD performances, at the cost of an increased complexity. In addition, the distortion minimization criterion of Eq. 1 used for the epitome charts creation could be changed into a rate-distortion optimization criterion, as in [9]. Ideally, the distortion would be directly computed on the restored EL instead of the reconstructed image, and the rate directly evaluated as the EL rate.

Finally, the proposed scheme can be extended to other scalable applications, such as color gamut or LDR/HDR scalabilities. Even though we use LLE in this paper for super-resolution, it has been proven efficient for many different applications such as de-noising [52], image prediction [27], or inpainting [48]. We can thus expect the LLE-based restoration methods to be efficient for different scalable applications.

REFERENCES

[1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, 2012.

[3] G. J. Sullivan, J. M. Boyce, Y. Chen, J.-R. Ohm, C. A. Segall, and A. Vetro, "Standardized Extensions of High Efficiency Video Coding (HEVC)," *IEEE J. Sel. Top. Signal Process.*, vol. 7, no. 6, pp. 1001–1016, 2013.

[4] Y. Ye and P. Andrivon, "The Scalable Extensions of HEVC for Ultra-High-Definition Video Delivery," *IEEE Multimed.*, vol. 21, no. 3, pp. 58–64, 2014.

[5] A. Kessentini, T. Damak, M. A. Ben Ayed, and N. Masmoudi, "Scalable high efficiency video coding (SHEVC) performance evaluation," in *World Congr. Inf. Technol. Comput. Appl.*, 2015, pp. 1–4.

[6] N. Jojic, B. Frey, and A. Kannan, "Epitomic analysis of appearance and shape," *IEEE Int. Conf. Comput. Vis.*, pp. 34–, 2003.

[7] V. Cheung, B. J. Frey, and N. Jojic, "Video epitomes," *Int. J. Comput. Vis.*, vol. 76, no. 2, pp. 141–152, 2008.

[8] H. Wang, Y. Wexler, E. Ofek, and H. Hoppe, "Factoring repeated content within and among images," *ACM Trans. Graph.*, vol. 27, p. 1, 2008.

[9] S. Cherigui, C. Guillemot, D. Thoreau, P. Guillotel, and P. Perez, "Epitome-based image compression using translational sub-pel mapping," in *IEEE Int. Work. Multimed. Signal Process.*, 2011, pp. 1–6.

[10] S. Cherigui, M. Alain, C. Guillemot, D. Thoreau, and P. Guillotel, "Epitome inpainting with in-loop residue coding for image compression," in *IEEE Int. Conf. Image Process.*, 2014, pp. 5581–5585.

[11] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, 2001.

[12] M. F. Tappen, B. C. Russell, and W. T. Freeman, "Exploiting the sparse derivative prior for super-resolution and image demosaicing," in *IEEE Int. Workshop Statist. Comput. Theories Vis.*, 2003.

[13] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, p. 95, 2007.

[14] W. T. Freeman, E. C. Pasztor, O. T. Carmichael, S. Hall, and F. Ave, "Learning Low-Level Vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.

[15] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, 2002.

[16] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. 275–282.

[17] W. Fan and D. Y. Yeung, "Image hallucination using neighbor embedding over visual primitive manifolds," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–7.

[18] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," *IEEE Int. Conf. Comput. Vis.*, pp. 349–356, 2009.

[19] J. Yang, J. Wright, T. S. Huang, and M. Yi, "Image Super-Resolution Via Sparse Representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010.

[20] J. Yang, Z. Lin, and S. Cohen, "Fast Image Super-Resolution Based on In-Place Example Regression," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 1059–1066.

[21] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi Morel, "Single-Image Super-Resolution via Linear Mapping of Interpolated Self-Examples," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5334–5347, 2014.

[22] K. Zhang, D. Tao, X. Gao, X. Li, and Z. Xiong, "Learning Multiple Linear Mappings for Efficient Single Image Super-Resolution," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 846–861, 2015.

[23] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.

[24] M. Aharon and M. Elad, "Sparse and Redundant Modeling of Image Content Using an Image-Signature-Dictionary," *SIAM J. Imaging Sci.*, vol. 1, no. 3, pp. 228–247, 2008.

[25] Q. Wang, R. Hu, and Z. Wang, "Intracoding and refresh with compression-oriented video epitomic priors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 5, pp. 714–726, 2012.

[26] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[27] M. Türkan and C. Guillemot, "Image prediction based on neighbor embedding methods." *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1885–98, 2012.

[28] M. Alain, C. Guillemot, D. Thoreau, and P. Guillotel, "Clustering-based Methods for Fast Epitome Generation," in *Eur. Signal Process. Conf.*, 2014, pp. 211–215.

Epitomic EL before restoration

Epitomic EL + E-LLE

Epitomic EL + E-LLM

Fig. 15. City enhancement layer encoded with QP=22. The epitome size is 39.52% of the input image. Before restoration, we can clearly notice in the red rectangle the blocks not belonging to the epitome from their blurry aspect. The quality of these blocks is obviously improved after restoration.

Epitomic EL before restoration

Epitomic EL + E-LLE

Epitomic EL + E-LLM

Fig. 16. Cactus enhancement layer encoded with QP=22. The epitome size is 48.33% of the input image. Before restoration, we can clearly notice in the red rectangle and in the calendar on the bottom right the blocks not belonging to the epitome from their blurry aspect. The E-LLM gives visually superior results compared to the E-LLE for the stochastic texture highlighted in the red rectangle. Both methods improve the quality of the straight lines in the calendar.

TABLE III
EXHAUSTIVE RESULTS FOR EACH TEST IMAGE AND MATCHING THRESHOLD.
(For Bjontegaard rate gains, the best rate saving is indicated in bold.)

Matching Threshold ε_M	Epitome size (% of input image)	Epitome rate gains (%) (rate only)		BD rate gains (%) (RD performance)		Matching Threshold ε_M	Epitome size (% of input image)	Epitome rate gains (%) (rate only)		BD rate gains (%) (RD performance)	
		EL only	EL + BL	E-LLE	E-LLM			EL only	EL + BL	E-LLE	E-LLM
BasketballDrive						BasketballDrill					
9	90.62	-13.38	-8.94	**-17.72**	-17.25	25	87.05	-2.50	-1.82	**-5.69**	-4.66
16	64.10	-21.67	-14.52	-13.34	-10.98	49	59.94	-16.13	-11.51	-1.87	2.07
25	49.33	-29.66	-19.85	-10.89	-5.85	100	42.63	-28.89	-20.54	-0.84	5.24
49	32.34	-43.58	-29.18	-7.12	2.47	225	28.53	-42.39	-30.06	3.72	13.59
Cactus						Keiba					
16	79.85	-15.77	-10.61	**-17.25**	-15.50	16	93.59	-0.39	-0.26	**-6.38**	-6.08
25	71.24	-17.99	-12.10	-16.73	-14.14	49	81.28	-5.50	-3.63	-3.23	-1.52
49	60.66	-22.11	-14.88	-15.38	-11.98	100	63.53	-18.86	-12.43	3.88	8.35
100	48.33	-31.89	-21.41	-10.59	-6.31	225	40.77	-40.86	-26.93	16.38	23.80
Ducks						Mall					
49	89.63	-25.27	-16.75	**-19.33**	-18.87	9	92.95	-8.80	-6.33	**-17.90**	-16.46
100	77.41	-30.53	-20.36	-16.49	-13.97	49	76.28	-3.19	-2.31	-0.09	-1.66
225	48.28	-50.06	-33.54	3.03	10.45	100	66.15	-8.02	-5.79	-3.76	3.48
						225	50.26	-20.05	-14.43	7.14	27.90
Kimono						PartyScene					
9	90.13	-22.22	-10.89	**-19.63**	-19.29	16	94.82	-0.56	-0.43	**-5.29**	-4.13
16	75.53	-58.38	-28.92	-14.29	-10.98	49	81.12	-4.34	-3.34	-0.68	7.75
25	59.36	-38.44	-18.96	-14.59	-12.29	100	67.56	-12.53	-9.65	9.53	26.67
49	35.34	-58.17	-28.78	-13.61	-10.16	225	49.13	-28.36	-21.85	27.89	58.14
ParkScene						BasketballPass					
9	86.58	-12.76	-8.48	**-16.25**	-15.81	9	77.76	-15.78	-11.42	**-13.37**	-10.25
25	73.55	-15.42	-10.26	-14.44	-13.17	16	66.60	-19.22	-13.91	-11.22	-4.20
49	61.99	-21.85	-14.52	-10.01	-6.80	25	56.41	-9.29	-6.78	2.66	13.96
100	47.18	-33.83	-22.45	-3.19	3.58	49	42.31	-18.90	-13.72	8.95	31.40
Tennis						BlowingBubbles					
9	64.49	-15.78	-9.58	**-20.45**	-19.27	25	87.56	-2.40	-1.74	**-4.67**	-1.56
16	50.44	-21.42	-13.00	-19.42	-16.86	49	73.33	-5.04	-3.72	-0.86	5.69
25	43.12	-26.77	-16.24	-17.13	-13.52	100	58.85	-14.45	-10.52	5.50	16.51
49	32.22	-38.19	-23.12	-15.90	-10.67	225	36.92	-35.13	-25.39	18.72	45.71
Terrace						RaceHorses					
9	78.46	-7.93	-5.86	**-12.81**	-12.03	9	91.03	-15.09	-10.14	**-14.65**	-14.23
25	66.39	-10.71	-7.94	-10.93	-9.17	25	79.23	-1.33	-0.96	-2.89	-1.42
100	53.31	-17.84	-13.20	-6.49	-2.66	100	58.14	-15.12	-10.27	8.83	22.80
225	43.50	-26.61	-19.63	-0.17	9.40	225	36.67	-38.80	-26.19	25.35	64.76
City						Square					
25	91.59	-6.02	-4.06	**-9.70**	-8.40	9	80.77	-4.18	-3.24	**-5.32**	-1.79
49	82.44	-9.54	-6.43	-5.80	-1.13	16	71.41	-5.86	-4.54	-4.74	2.36
100	66.81	-20.46	-13.82	3.65	18.10	49	61.09	-5.82	-4.52	-0.80	6.11
225	39.52	-46.98	-31.80	28.20	60.07	225	48.72	-15.73	-12.08	10.92	32.69

[29] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP Graph. Model. Image Process.*, vol. 53, no. 3, pp. 231–239, 1991.

[30] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi Morel, "Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding," in *Br. Mach. Vis. Conf.*, 2012, pp. 1–10.

[31] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, no. 0, 2005, pp. 60–65.

[32] "SHM Software, Ver. 9.0. Available: https://hevc.hhi.fraunhofer.de/trac/shvc/browser," [Accessed: 07- Jun- 2016].

[33] "HM Software, Ver. 15.0. Available: https://hevc.hhi.fraunhofer.de/trac/hevc/browser," [Accessed: 07- Jun- 2016].

[34] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *Doc. VCEG-M33, ITU-T VCEG Meet.*, 2001.

[35] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[36] J. H. Freidman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Softw.*, vol. 3, no. 3, pp. 209–226, 1977.

[37] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein, "The generalized PatchMatch correspondence algorithm," *Lect. Notes Comput. Sci.*, vol. 6313, pp. 29–43, 2010.

[38] N. Dowson and O. Salvado, "Hashed nonlocal means for rapid image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 485–499, 2011.

[39] A. Cherian, S. Sra, V. Morellas, and N. Papanikolopoulos, "Efficient Nearest Neighbors via Robust Sparse Hashing," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3646–3655, 2014.

[40] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "2-D Dictionary Based Video Coding for Screen Contents," *Data Compression Conf.*, pp. 43–52, 2014.

[41] V. Garcia, E. Debreuve, and M. Barlaud, "Fast k nearest neighbor search using GPU," in *Conf. Comput. Vis. Pattern Recognit. Work.*, 2008, pp. 1–6.

[42] X. Li, J. Chen, K. Rapaka, and M. Karczewicz, "Generalized inter-layer residual prediction for scalable extension of HEVC," in *IEEE Int. Conf. Image Process.*, 2013, pp. 1559–1562.

[43] T. Laude, X. Xiu, J. Dong, Y. He, Y. Ye, and J. Ostermann, "Scalable extension of HEVC using enhanced inter-layer prediction," in *IEEE Int. Conf. Image Process.*, 2014, pp. 3739–3743.

[44] A. Aminlou, J. Lainema, K. Ugur, M. M. Hannuksela, and M. Gabbouj, "Differential Coding Using Enhanced Inter-Layer Reference Picture for the Scalable Extension of H.265/HEVC Video Codec," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 11, pp. 1945–1956, 2014.

[45] X. HoangVan, J. Ascenso, and F. Pereira, "Improving enhancement layer merge mode for HEVC scalable extension," in *Pict. Coding Symp.*, 2015, pp. 15–19.

[46] ——, "Improving SHVC Performance with a Joint Layer Coding Mode," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, 2016.

[47] K. Kim and Y. Kwon, "Single Image Super-Resolution Using Sparse Regression and Natural Image Prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, 2010.
[48] C. Guillemot and O. Le Meur, "Image Inpainting : Overview and Recent Advances," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 127–144, 2014.
[49] I. Drori, D. Cohen-Or, and H. Yeshurun, "Fragment-based image completion," *ACM Trans. Graph.*, vol. 22, no. 3, p. 303, 2003.
[50] O. Le Meur, M. Ebdelli, and C. Guillemot, "Hierarchical super-resolution-based inpainting," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3779–3790, 2013.
[51] M. Turkan, M. Alain, D. Thoreau, P. Guillotel, and C. Guillemot, "Epitomic image factorization via neighbor-embedding," in *IEEE Int. Conf. Image Process.*, 2015, pp. 4141–4145.
[52] I.-F. Shen, "Image Denoising through Locally Linear Embedding," *Int. Conf. Comput. Graph. Imaging Vis.*, pp. 147–152, 2005.

Martin Alain received the Master's degree in electrical engineering from the Bordeaux Graduate School of Engineering (ENSEIRB-MATMECA), Bordeaux, France in 2012 and the PhD degree in signal processing and telecommunications from University of Rennes 1, Rennes, France in 2016. As a PhD student working in Technicolor and INRIA (Institut National de Recherche en Informatique et Automatique) in Rennes, France, he explored novel image and video compression algorithms. He is currently a postdoctoral researcher in the V-SENSE project at the School of Computer Science and Statistics in Trinity College Dublin, Ireland. His research interests lie in the broad field of digital signal and image processing including video compression, de-noising and super-resolution, with a focus on light field imaging.

Dominique Thoreau received his PhD degree in image processing and coding from the University of Marseille Saint-Jérôme in 1982. From 1982 to 1984 he worked for GERDSM Labs on underwater acoustic signal and image processing of passive sonar. He joined in 1984 the Rennes Electronic Labs of Thomson CSF and worked successively on sonar image processing, detection/tracking and fusion on visible-infrared videos, and on various projects of video coding. Currently working for Technicolor, his research interests include signal processing, high dynamic range and light field imaging, image and video compression.

Philippe Guillotel received the engineering degree in electrical engineering, telecommunications and computer science from the Ecole Nationale Superieure des Telecommunications (France) in 1988, the M.S. degree in electrical engineering and information theory from University of Rennes (France) in 1986, and the PhD degree from University of Rennes (France) in 2012. He is Distinguished Scientist at Technicolor Research & Innovation, in charge of video processing & user experiences research topics.

He has worked in the Media & Broadcast Industry for more than 25 years covering video processing, video delivery, coding systems, as well as human perception and video quality assessment. He has been involved in many collaborative projects on those technical areas. More recently he started exploration of new topics to improve user experiences in media consumption, in particular to increase immersion and engagement.

Christine Guillemot holds a PhD degree from ENST (Ecole Nationale Superieure des Telecommunications) Paris. From November 1985 to October 1997, she has been with FRANCE TELECOM, where she has been involved in various projects in the area of coding for TV, HDTV and multimedia. From 1990 to mid 1991, she has worked at Bellcore, NJ, USA, as a visiting scientist. Since November 1997, she is 'Director of Research' at Inria, head of a research team dedicated to the design of algorithms for the image and video processing chain, with a focus on analysis, representation, compression, and editing, including for emerging modalities such as high dynamic range imaging and light fields.

She has served as Associate Editor for several IEEE journals (IEEE Trans. on Image Processing, IEEE Trans. on Circuits and Systems for Video Technology, IEEE Trans. on Signal Processing, Eurasip journal on image communication) and she is currently a member of the IEEE IVMSP Committee. She is currently Senior Area Editor of IEEE Trans. On Image Processing (2016-2017) and member of the IEEE- trans. On Multimedia steering committee. She has been a member of the IEEE IMDSP (2002-2007) and IEEE MMSP (2005-2008) technical committees. She is an IEEE fellow since January 2013.