Video enhancement with convex optimization methods

Benoît Boyadjis*†, Andrei Purica*, Béatrice Pesquet-Popescu*, Frédéric Dufaux‡
*LTCI, Telecom Paristech, Universite Paris-Saclay – 46 rue Barrault, Paris, France – first.last@telecom-paristech.fr
†Thales Communications and Security – 4 Av. des Louvresses, Gennevilliers, France – first.last@thalesgroup.com
‡L2S, CNRS, CentraleSupelec, Univ. Paris Sud – 3 rue Joliot Curie, Gif-sur-Yvette, France – first.last@l2s.centralesupelec.fr

***Abstract*—Video enhancement methods enable to optimize the viewing of video content at the end-user side. Most approaches do not consider the compressed nature of the available content. In the present work, we build upon a recently proposed video enhancement approach that explicitly models a compression stage. To apply the enhancement framework on compressed representations requires to extract specific syntax elements during their decoding. This additional information embeds the enhanced result in a domain that closely fits the observation. We evaluate the framework performance in a single source resolution enhancement scenario, and show the method efficiency with respect to state-of-the-art approaches.**

***Index Terms*—video enhancement, super resolution, HEVC, convex optimization**

I. Introduction

Video compression standards, such as the recent High Efficiency Video Coding (HEVC) [1], enable to achieve efficient video compression. Yet, there is a great potential for video enhancement approaches that recover, at the end-user side, the video content with optimal quality. In this work, we particularly focus on the *single source resolution enhancement* issue, often denoted as Single Source Super Resolution (SS-SR). Given the limited number of bands in video broadcast spectrum, SS-SR is of particular interest as it may be used to replace SD and HD transmissions with a single HEVC description that can provide on demand quality adjustments.

Video enhancement is a wide research field, since multiple topics are covered by this denomination (de-noising, deblurring, resolution enhancement, etc.). SS-SR methods are under active exploration [2], [3], [4], [5]. We observe that SS-SR approaches often consider that the observed low resolution image $\overline{x} \in \mathbb{R}^M$ is a down-sampled observation of the image to be found $X \in \mathbb{R}^N$ (where M and N are the number of pixels in the observation and target image, respectively). By denoting $L \in \mathbb{R}^{M \times N}$ the sensing matrix, which may include a low-pass anti-aliasing filter, the following degradation model is obtained:

$$\overline{x} = LX \tag{1}$$

Most recent learning-based SS-SR methods [6], [7], [8], [9], [10], which are arguably the current best performing approaches, rely on this degradation model. Yet, the approach may not adapt well to compressed streams. *Unlearned* compression artifacts may be misleadingly confused with image features. In [11], the authors specifically learn regressors on a compressed dataset. However, in practice,the learned model needs to be related to the compression ratio used.

In a previous work [12], we presented a new theoretical image enhancement framework that handles multiple compressed representations while in [13], we show an extension to multi source compressed video and propose new applications. The framework has the particularity to explicitly model a generic compression stage. A first single source HEVC application of the framework in [14], where the SR potential on Intra frames is discussed. The present work further analyzes the framework behavior with single source HEVC encoding. In particular, predicted frames are now included in the SR framework, and a more thorough experimental evaluation is provided. The method's behavior with respect to one of the best performing SS-SR approaches is discussed in multiple scenarios, showing the high adaptation capability of the proposed framework.

The remaining of this paper is organized as follows. A presentation of the SS-SR framework is proposed in Section II. Then, its application to HEVC content is discussed in Section III. Experimental results are presented in Section IV. Finally, conclusions are drawn in Section V.

II. A compressed video enhancement framework

A. Acquisition model

Let us denote by $X = [X_1, ..., X_K]$, with $\forall \quad i \in [1, K]$, $X_i \in \mathbb{R}^N$ an original video sequence. We model the acquisition process of i-th image by a linear operator $L_i \in \mathbb{R}^{M \times N}$. We consider that the observation $\overline{X} = [\overline{X_1}, ..., \overline{X_K}]$, with $\forall i \in [1, K]$, $\overline{X_i} \in \mathbb{R}^M$ is available in the form of an HEVC bitstream, which relies on the generic hybrid video coding scheme presented in Fig. 1. As many other image/video compression approaches, HEVC relies on block-based transforms and a subsequent quantization to efficiently compress the video frames. Moreover, a prediction step is involved: HEVC does not apply directly the transforms to pixel blocks, but to a residual obtained by differentiating the observation with a prediction. Therefore, the encoded coefficients $z \in \mathbb{R}^M$ contained in HEVC bitstream are the quantized version (quantization operator $\mathcal{Q}$) of the output of the transform T applied to residual pixel blocks, so that the resulting model can be written as follows ($\widetilde{X_i}$ being the HEVC prediction of the i-th frame):

$$\overline{X_i} = T_i^{-1}(z_i),\ z_i = \mathcal{Q}_i\big(\overline{y_i}\big),\ \overline{y_i} = T_i(L_i X_i - \widetilde{X_i}) \tag{2}$$

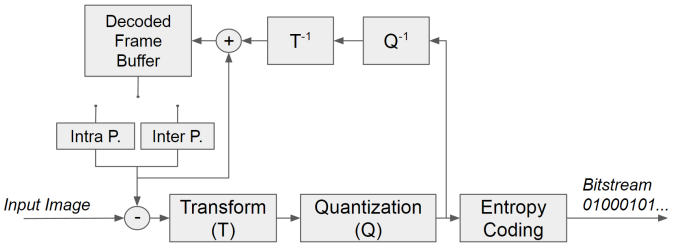

Fig. 1. A generic hybrid video coding scheme.

B. Video enhancement in the compressed domain

For the sake of clarity, this section only gathers a succinct presentation of the video enhancement framework presented in our previous work [12]. We invite the reader to refer to [12] for a more thorough description of the proposed model and additional information on the method used to solve the associated minimization problem.

1) Compression-related clues: Reconstruction levels (z_i) are used as a quality reference in the transform domain. Given an estimate $\hat{X}_i$ of the i-th video frame, we minimize the following data fidelity cost function:

$$J_{\text{DF}}(\hat{X}_i) = \left| T_i(L_i\hat{X}_i - \widetilde{X_i}) - z_i \right|^2 \tag{3}$$

Additionally, note that we know the Quantization Interval (QI) for each quantized coefficient in the bitstream. If we denote by $j_i^{(k)}$ the QI index for the coefficient $z_i^{(k)}$, we can define:

$$C_i = \{ y = (y^{(k)})_{1 \leq k \leq M} \in \mathbb{R}^M \,|\, \forall k,\ y^{(k)} \in QI_{j_i^{(k)}} \} \tag{4}$$

This definitions enable to model the QI validity constraint:

$$\text{Find } \hat{X}_i :\ T_i(L_i\hat{X}_i - \widetilde{X_i}) \in C_i \tag{5}$$

2) Range constraint: Encompassing *a priori* information into the reconstruction problem is a common choice in literature. We first enforce the solution to have pixel values belonging to a specific *range*, typically known given the application domain:

$$\text{Find } \hat{X}_i : \forall k \in [1, N],\ X_i^{\min} \leq \hat{X}_i^{(k)} \leq X_i^{\max} \ . \tag{6}$$

3) TV constraint: Finally, a classical choice is to enforce the smoothness of the solution by limiting its discontinuities according to a suitable metric. We opt here for the classical Total Variation (TV) [15] to measure the discontinuity of the solution. In order to avoid over-smoothing, the TV is introduced as an additional constraint. By denoting X_i^0 the initial estimate for the i-th video frame, the TV boundary used for the enhanced result is derived according to:

$$\text{find } \widehat{X_i} : \quad \text{TV}(\hat{X}_i) \leq \eta_i \quad \text{where} \quad \eta_i = \eta_0 \cdot \text{TV}(X_i^0) \tag{7}$$

In this work, we empirically selected $\eta_0 = 0.95$, which was found to provide a stable and efficient application of the constraint at various compression levels.

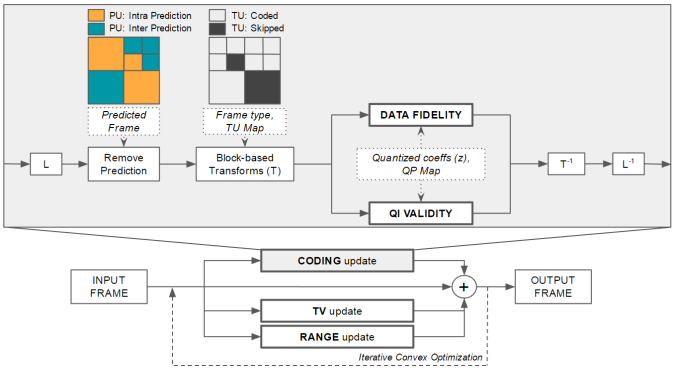

Fig. 2. A schematic view of the usage of HEVC information in the optimization framework.

C. Enhancement model

We now denote ι_C the characteristic function of a closed convex set C, defined by:

$$\iota_C(y) = \begin{cases} 0 & \text{if } y \in C \\ +\infty & \text{otherwise.} \end{cases} \tag{8}$$

The enhancement framework then consists in minimizing, for each video frame, the following criterion :

$$\text{Find } \hat{X}_i \in \underset{\hat{X}_i \in \mathbb{R}^N}{\text{Argmin}} \Big(J_{\text{DF}}(\hat{X}_i) + \iota_{C_i}(T_i(L_i\hat{X}_i - \widetilde{X_i})) + \sum_{s=1}^{S} \iota_{D_{s(i)}}(F_s\hat{X}_i) \Big) \tag{9}$$

where, $\iota_{D_{s(i)}}$ is the closed convex set of constraint s for image i and F_s introduces the range and TV constraints from Eq. (6) and Eq. (7) into the problem formulation ($S = 2$). The range constraint is directly applied to the image, thus, F_1 is the identity function and $D_1(i) = \{X \in \mathbb{R}^M : \forall k,\ X^{(k)} \in [X_{\min}^i, X_{\max}^i]$ with $X_{\min}^i = 0$ and $X_{\max}^i = 255$. The isotropic TV smoothness constraint computes $F_2 = (\nabla_h, \nabla_v)$ (*i.e.* the horizontal and vertical gradients) and D_2 is the closed convex set defined by Eq. 7.

D. A primal-dual solver

We tackle the problem of solving Eq. (9) using the primal-dual algorithm proposed by Combettes *et al.* in [16], known as Monotone Lipschitz Forward-Backward-Forward (M-LFBF) algorithm. This algorithm, unlike other similar methods, assures a lower computational complexity for problems involving linear operators as it does not require any matrix inversion [16]. Furthermore, the block iterative structure of the algorithm allows for efficient parallel implementations on multi-core architectures.

III. Application to HEVC

A schematic view of the optimization process is proposed in Fig. 2. At each step of the iterative structure, the current estimate has to be projected onto the transform bases in order to estimate the data fidelity and QI validity objectives. This

projection relies on information extracted from the HEVC bitstream, which we further detail in the next subsections.

A. Extracting the required HEVC information

We recall that HEVC relies on a hierarchical quadtree structure: the frame is first split into *Coding Tree Units* (CTUs) of fixed size (from 64x64 to 16x16). CTUs are split (potentially recursively) in *Coding Units* (CUs), forming the quadtree structure. Then, *Prediction Units* (PUs) and *Transform Units* (TUs) are rooted at the CU level to gather all the unit information on the prediction and the transform used respectively. PUs and TUs are independent, so that prediction and transform can be made at different sizes inside a CU.

1) Intra and Inter prediction: In HEVC, predicting pixel values rely on Intra or Inter prediction at the PU level. On the one hand, Intra prediction uses previously decoded units of the frame as a reference to predict pixel values for the unit to be encoded. On the other hand, Inter prediction uses motion estimation and compensation from a set of frames amongst previously decoded frames of the current Group Of Pictures (GOP). HEVC always computes a prediction for a PU. We thus obtain the predicted frame $\widetilde{X_i}$ by concatenating all reconstructed PUs. Since $\widetilde{X_i}$ is a static parameter in the proposed framework, PU sizes and types (Intra/Inter) are not required during the enhancement process.

2) HEVC transforms: HEVC transforms residuals at the TU level. TUs are square pixel units that can be recursively subdivided, so different transform sizes are specified (4x4, 8x8, 16x16, 32x32). Due to complexity considerations, HEVC relies on finite approximations of well-known transforms: the Discrete Cosine Transform (DCT) and its inverse (IDCT). Furthermore, a Discrete Sine Transform (DST) is specifically used for 4x4 Intra units [1]. Two elements are extracted from the HEVC bitstream to reproduce the same transforms: the frame type and the TU partitioning (TU map).

3) HEVC quantization: The applicable quantizer is indicated by a Quantization Parameter (QP) ranging from 0 to 51 which serves as an integer index to derive the applicable step size Δ_q [1]. Since HEVC enables QP variation between TUs of the same frame, the QP map (containing the QP of each TU) is extracted from the HEVC bitstream to compute the step-size for each unit.

B. Encoding/decoding configuration

HEVC compliant video streams are generated using the reference software HM 15.0 [17]. The encoding configuration re-uses the default Random Access configuration file, with minor modifications. First, CU-based multi-QP optimization is enabled by setting the parameter MaxDeltaQP at 2. Second, since our SR model has not considered HEVC in-loop filters yet, both the deblocking and sample adaptive offset filters were turned off in the configuration file. Finally, the Group Of Picture (GOP) is set to 8 frames. Decoding is performed with an OpenHEVC decoder [18] which is patched to output the required elements for the optimization approach.

IV. Performance assessment

A. Experimental set up

We evaluate the proposed video enhancement framework in a SS-SR scenario, using a testbed of 6 CIF video sequences. We select the learning-based approach of Timofte *et. al.* [10] as the state-of-the art (SOA) anchor, since it is one of the best performing approaches in the SS-SR literature. In the resolution enhancement scenario, the L_i operator models a re-sampling step performed before the encoding. In particular, the re-sampling consists in a polyphase filtering followed by a decimation. The polyphase filter weights are determined using any popular interpolation method, such as: Lanczos resampling [19], bicubic [20] or filters proposed for SVC [21] or SHVC [22]. The Low Resolution (LR) compressed observation are obtained by applying the re-sampling filter to each video sequence, before HEVC encoding at various QPs. Note that an initial High-Resolution (HR) estimate is used as an initializer for the proposed framework. We showed in [14] that the initialization impacts the framework result, and in general, the best quality initialization provides the best quality result. In this work, unless when specifically signaled, the bicubic version of the polyphase filter is employed (both to generate the LR observation and to propose an HR estimate).

B. HEVC transform skip mode

In HEVC, a prediction is always computed for a CU, but the residual might be entirely skipped. This choice is available for every TU during Rate-Distortion Optimization (RDO) at the encoder side. This scenario where a TU has no residual (and an undefined QP) is not naturally modeled by our framework. In particular, both the data fidelity and the preservation of QI rely on the block QP and the encoded coefficients z_i. On the one hand, we can remove these blocks from the computation of both the data fidelity objective and the QI validity criterion in the optimization framework. On the other hand, we can set the reconstruction levels z_i to zero and pick a default QP for these skipped blocks. Two natural candidates arise for modeling a default QP: a value of 1 will imply strong confidence on pixel values and will limit their potential variation. Another choice is to rely on the maximum encoding QP available during the frame encoding. These scenarios are compared in Table I. The first interesting observation that stems from Table I is that entirely relaxing the compression related constraints for skipped units significantly underperforms the two other methods. This tend to prove that, even for skipped units, it is beneficial to include compression priors. Then, we show that using a QP of 1 for these skipped units is not a reliable choice, particularly at high encoding QPs. Selecting the maximum available QP during encoding offers stable results over all QPs, which is thus the selected solution for the remaining of this work. Note that this result is eventually quite intuitive: if the encoder RDO decides to skip a TU, it is most probably because no information is to be coded at the encoding QP, which is thus a natural candidate for modeling the degree of confidence we can have in the unit pixel values.

		Empirical QP for skipped units				No Crit.	
		1		Max. *			
Sequence	Enc. QP	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Akiyo	1	**35.82**	**0.969**	**35.82**	**0.969**	34.46	0.950
	15	**35.58**	**0.962**	35.38	0.959	33.99	0.954
	30	33.52	0.930	**33.62**	**0.931**	32.93	0.928
Foreman	1	**32.81**	**0.950**	32.80	**0.950**	32.47	**0.950**
	15	30.30	**0.873**	**30.31**	0.873	30.14	0.871
	30	32.51	0.930	**32.59**	**0.952**	31.98	0.916
Bus	1	**27.78**	**0.875**	**27.78**	**0.875**	27.53	0.871
	15	27.41	**0.847**	**27.42**	**0.847**	26.77	0.801
	30	25.38	0.740	**25.62**	**0.743**	25.12	0.722

TABLE I
USING DIFFERENT QP SETTINGS FOR SKIPPED UNITS. (*: MAX. DENOTES THE MAXIMUM FRAME QP)

$\downarrow\uparrow$ method	$\uparrow_H$	$\uparrow_{SOA}$	$\uparrow_{Prop.}(\uparrow_H)$	$\uparrow_{Prop.}(\uparrow_{SOA})$
$\downarrow_L$ (Bic_4)	22.87	22.16	22.93	23.35
$\downarrow_L$ (Bic_8)	22.25	24.17	23.42	24.19
$\downarrow_L$ (SVC)	21.45	20.64	23.75	23.85

TABLE II
COMPARING THE PSNR (DB) OF UP-SAMPLING METHODS $\uparrow_H$, $\uparrow_{SOA}$, $\uparrow_{Prop}(\uparrow_H)$, $\uparrow_{Prop}(\uparrow_{SOA})$ W.R.T DIFFERENT DOWN-SAMPLING FILTERS. (MOBILE, QCIF TO CIF, 1 GOP, QP15)

C. Re-sampling methods

The weights of the down-sampling polyphase filter L (and its up-sampling counterpart H) can be chosen according to various interpolation methods. In order to prove the framework robustness towards these operators, we propose to consider three different re-sampling approaches. First, we test the popular 4-tap Bicubic (Bic_4) interpolation [20]. Then, we also combine the filter with an anti-aliasing effect by stretching the bicubic function to 8 taps (Bic_8). Finally, we evaluate the re-sampling filter proposed in the Scalable Video Coding (SVC) standard [21]. In Table II, we gather the results obtained when performing SR of the first 8 frames of Mobile sequence compressed at QP 15. The behavior of the SOA approach is typical of learning-based approaches: its performance is only optimal in the scenario it has been trained for. Therefore, the SOA approach enables to obtain a 2dB gain when using the Bic_8 filter but fails to improve the quality of other re-sampling schemes. On the opposite, our approach exhibits consistent behavior with respect to the degradation filter.

D. General evaluation

We now conduct a general evaluation of the proposed framework in the SS-SR scenario, and compare our results to the SOA anchor [10]. As detailed in the previous section, a fair comparison of the approaches requires to rely on the Bic_8 re-sampling filter. For each sequence of the testbed, 32 frames are down-sampled and compressed at various QPs (from 1 to 40). The average PSNR and SSIM gains throughout all QPs are measured by the Bjontegaard metric [23], as illustrated in Table III.

The generic behavior described by Table III is consistent over all sequences: when using a bicubic initialization, the proposed approach cannot compete with the results of [10]. However, when using SOA as the initializer, the resulting quality is slightly increased, which proves the efficiency of enforcing the SR result to closely fit the available compressed description. We further illustrate the framework behavior with visual results in Fig. 3, which compares a reference frame obtained with bicubic interpolation (*left*), the SOA result (*middle*), and the refined SOA result using the proposed enhancement framework (*right*). If the bicubic results tend to be over-smoothed, the two other approaches enable a more precise definition of contours and textures.

Scale x2	$\uparrow_{Prop.}(\uparrow_H)$		$\uparrow_{SOA}$		$\uparrow_{Prop.}(\uparrow_{SOA})$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Akiyo	+1.21	+0.008	+2.96	**+0.015**	**+3.05**	**+0.015**
Foreman	+0.60	+0.005	+0.93	**+0.010**	**+1.14**	**+0.010**
Bus	+0.80	+0.021	+1.49	+0.045	**+1.56**	**+0.046**
Mobile	+0.83	+0.05	+1.45	+0.078	**+1.56**	**+0.079**
Flower	+0.34	+0.023	+0.69	+0.041	**+0.76**	**+0.043**
Football	+0.88	+0.012	+1.83	+0.036	**+1.99**	**+0.037**

TABLE III
EVALUATION OF AVERAGE PSNR (DB) AND SSIM GAINS OVER BICUBIC (Bic_4) INTERPOLATION USING THE BJONTEGAARD METRIC [23].

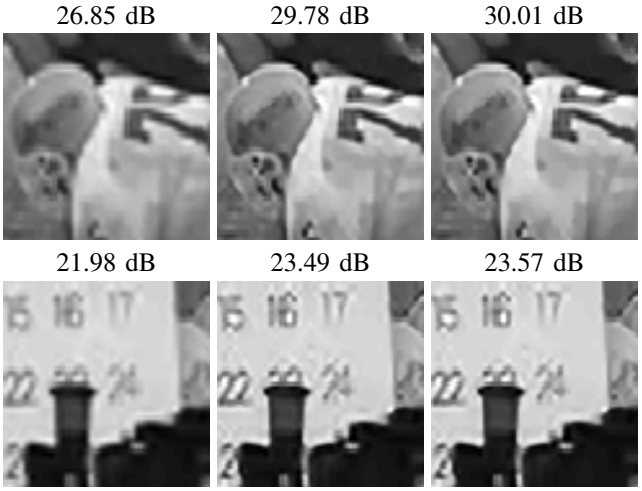

Fig. 3. Details of the up-sampled images using, from *left* to *right*, a Bic_8 upsampling, the SOA approach from [10], and the proposed enhancement framework (initialized with the SOA result). PSNR values are computed on each image patch.

V. CONCLUSION

In this work, we detail the application of a new type of video enhancement approach for HEVC compressed video sequences. The model explicitly uses the available compressed syntax to build a heterogeneous cost function that combines compression-related clues and a priori constraints. Convex optimization enables to efficiently solve the associated minimization problem. The framework thus tends to embed the enhanced result into a domain that closely fits the given compressed observation. Evaluated in a single source resolution enhancement scenario with HEVC encodings, we show that our approach provides interesting results in various configurations, while avoiding typical pitfalls of learning-based approaches. The complexity and real-time capabilities of the proposed framework have still to be more thoroughly analyzed, and short-term research focuses on its implementation and optimization on parallel processing platforms.

REFERENCES

[1] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[2] P. Milanfar, *Super-Resolution Imaging, Digital Imaging and Computer Vision*. Taylor&Francis/CRC Press, 2010.

[3] S. K. Nelson and A. Bhatti, "Performance evaluation of multi-frame super-resolution algorithms," in *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Geelong, Victoria, Australia, 2012.

[4] L. Tian, A. Suzuki, and H. Koike, "Task-oriented evaluation of super-resolution techniques," in *International Conference on Pattern Recognition (ICPR)*, 2010, pp. 493–498.

[5] Y. J. Kim, J. H. Park, G. S. Shin, H.-S. Lee, D.-H. Kim, S. H. Park, and J. Kim, "Evaluating super resolution algorithms," in *SPIE-IS&T Electronic Imaging, Image Quality and System Performance VIII*, vol. 7867, 2011.

[6] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *IEEE International Conference on Computer Vision (ICCV)*, Sept 2009, pp. 349–356.

[7] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*. Springer Berlin Heidelberg, 2012, vol. 6920, pp. 711–730.

[8] C.-Y. Yang and M.-H. Yang, "Fast direct super-resolution by simple functions," in *IEEE International Conference on Computer Vision (ICCV)*, Dec 2013, pp. 561–568.

[9] C. Dong, C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Computer Vision ECCV 2014*, ser. Lecture Notes in Computer Science. Springer International Publishing, 2014, vol. 8692, pp. 184–199.

[10] R. Timofte, V. D. Smet, and L. V. Gool, "A+: Adjusted anchored neighborhood regression for fast superresolution," in *Computer Vision ACCV 2014*, ser. Lecture Notes in Computer Science, vol. 9006. Springer International Publishing, 2015, pp. 111–126.

[11] R. Rothe, R. Timofte, and L. Van Gool, "Efficient regression priors for reducing image compression artifacts," in *IEEE International Conference on Image Processing (ICIP)*, Sept 2015, pp. 1543–1547.

[12] R. Gaetano, B. Pesquet-Popescu, and C. Chaux, "A convex optimization approach for image resolution enhancement from compressed representations," in *18th International Conference on Digital Signal Processing (DSP)*, July 2013, pp. 1–8.

[13] A. Purica, B. Boyadjis, B. Pesquet-Popescu, F. Dufaux, and C. Bergeron, "A convex optimization framework for video quality and resolution enhancement from multiple descriptions," *IEEE Transactions on Image Processing (submitted to)*, 2017.

[14] B. Boyadjis, B. Pesquet-Popescu, F. Dufaux, and C. Bergeron, "Super-resolution of HEVC videos via convex optimization," in *IEEE International Conference on Image Processing (ICIP)*, September 2016.

[15] P. L. Combettes and J.-C. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1213–1222, 2004.

[16] ——, "Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators," *Set-Valued and Variational Analysis*, vol. 20, no. 2, pp. 307–330, 2012.

[17] Joint video team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "H.265/HEVC HM reference software," http://hevc.fraunhofer.de//.

[18] "OpenHEVC Open source HEVC decoder," https:/github.com/openhevc/openhevc/.

[19] K. Turkowski and S. Gabriel, "Filters for common resampling tasks," *Andrew S. Glassner Graphics Gems I, Academic Press*, pp. 147–165, 1990.

[20] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 6, December 1981.

[21] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sept 2007.

[22] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of shvc: Scalable extensions of the high efficiency video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20–34, Jan 2016.

[23] G. Bjontegaard, "Calculation of average PSNR differences between RD curves." in *ITU-T SC16/Q6, 13th VCEG meeting*, Austin, TX, USA, Apr. 2001.