

Seam carving for semantic video coding

M. Décombas^{ab}, F. Capman^a, E. Renan^a, F. Dufaux^b, B. Pesquet-Popescu^b,

^aThales Communications & Security, Laboratoire MMP, 92700 Colombes, France;

^bTélécom ParisTech, Dept. Traitement du Signal et des Images, 75014 Paris, France

{marc.decombas, erwann.renan, francois.capman}@fr.thalesgroup.com

{frederic.dufaux, beatrice.pesquet}@telecom-paristech.fr

ABSTRACT

Compression standards such as H.264/AVC encode video sequences to maximize fidelity at a given bitrate. However, semantic-oriented and content-aware compression remains a challenge. In this paper, we propose a semantic video compression method using seam carving. Seam carving changes the dimension of an image/video with a non-uniform resampling of each row and column while keeping the rectangular shape of the image. Our main contribution is a new approach to identify areas where seams are concentrated. On the one hand, it allows to transmit supplemental seams data at low cost. On the other hand, seams can be synthesized at the decoder in order to recover the original frame size and to preserve the scene geometry. Experiments show that our seam carving method combined with standard H.264/AVC coding results in significant bitrate savings compared with the original H.264/AVC. Reported gains reach 39% at very high bitrates and 22% at very low bitrates. Furthermore, the reconstructed video has the same quality in semantically significant regions.

Keywords: video coding, seam carving, image resizing, saliency map, content-aware semantic compression

1. INTRODUCTION

Video compression aims at efficiently representing image data in order to allow easy storage and transmission. Common solutions include standards such as H.264/AVC [1] with the goal of maximizing fidelity between the original and compressed images at a given bitrate. In this context, fidelity is generally assessed using well-known quality metrics such as Mean Square Error (MSE) and Peak Signal-to-Noise-Ratio (PSNR), or perceptually-based metrics such as Structural SIMilarity Index (SSIM) [2].

In this paper, we consider the problem of video compression under a different angle. More specifically, we aim at preserving the visual quality of the important parts and the semantic content of a video. We target security applications, such as video monitoring using drones, where it is very important to preserve the salient regions but the bitrate available is very low. As the background carries a very low significance, it can be shrunk. The bitrate is then greatly reduced while the quality and position of the important objects in the video are preserved.

In order to achieve this objective and to improve video compression efficiency, we propose to use seam carving. Avidan and Shamir [3] introduced seam carving for content-aware image resizing that preserves semantic content. In particular, this technique has been shown to efficiently adapt the size of an image in order to match a display size. Seam carving is a method based on an energy function and a cumulative energy function. Most past works try to improve the energy function [4][5][6][7] and the cumulative energy function [3][8][9] to avoid that important parts of the image be deformed and to keep temporal coherence for video resizing. A few more recent works use seam carving to compress images and video sequences [10][11][12][13][14][15].

In this paper, we define a new method where seam carving is applied in conjunction with a conventional video coding scheme such as H.264/AVC. As our method is a pre and post processing, the generated bitstream is compatible with all types of video and image standards.

More specifically, the video is split into Groups Of Pictures (GOP) with a predefined size. Next, we use seam carving to spatially reduce the original GOP dimensions in order to discard insignificant information while safeguarding semantically important regions. For each GOP, the number of seams suppressed is defined according to its content and is the same for all frames in the GOP. The motivation is to take into account temporal coherence and to have a more robust

approach. Video compression is then applied on the downscaled video. Given that the quantity of data has been greatly reduced, compression is more effective than when applied on the original GOP. In order to be able to recover the original spatial dimensions without disturbing the objects semantic and the scene geometry, we analyze all the seams positions and encode some key points. With this approach, the amount of data to transmit the seams information is greatly reduced. At the decoder side, these key points control the seam carving up-sampling by modifying the energy function. In this way, the new seams positions are adapted with the content.

We carry out extensive performance evaluation experiments to show that the bitrate is greatly decreased when using the proposed seam carving based video coding. This gain is obtained while maintaining high visual quality on the semantically interesting regions.

This paper is structured as follows. An overview of seam carving is presented in Sec. 2. The proposed semantic video compression method based on seam carving is introduced in Sec. 3. Experimental results are presented in Sec. 4 to assess the performance of the proposed approach. Finally, we draw conclusions and discuss future work directions in Sec. 5.

2. SEAM CARVING REVIEW

Seam carving is a new approach to resize images or video sequences while preserving semantically important content [3]. A seam is an optimal 8-connected path of pixels on a single image from top to bottom or left to right. Formally, let I be an $n \times m$ image, the term *vertical seam* is defined to be the set of points

$$s^X = \{s_i^x\}_{i=1}^n = \{x(i), i\}_{i=1}^n, s. t. \forall i, |x(i) - x(i-1)| \leq 1, \quad (1)$$

with x the horizontal coordinate of the point.

Note that similar to the removal of a row or column from an image, removing the pixels of a seam from an image has only a local effect: all the pixels of the image are shifted left (or up) to compensate for the missing path. The cost of the seam is defined by an energy function and an energy cumulative function. Most of the works have focused on these two functions in order to improve the seam path such that important regions are preserved.

2.1 Energy functions

The first energy function proposed by Avidan and Shamir [3] is the gradient magnitude on the luminance component. In [10], Anh *et al.* use a combination of a saliency map and the magnitude of gradient. Other works propose to use different kinds of saliency maps [4][5][6][7] with or without taking the temporal aspect into account.

2.2 Cumulative energy functions

These cumulative energy functions are dynamic programming algorithms and are used to find the minimum cumulative energy path. In the case of a vertical seam, the first step is to traverse the image from top to bottom and compute the cumulative minimum energy M for all possible connected seams for each entry (i, j) . Avidan and Shamir [3] propose to use backward energy to find the optimal seam path. The drawback of this method is that it does not measure the consequence of the suppression of a seam and can create some artifacts. Therefore, Rubinstein *et al.* propose in [8] to use forward energy defined as:

$$M(i, j) = P(i, j) + \begin{cases} M(i-1, j-1) + C_L(i, j) \\ M(i-1, j) + C_U(i, j) \\ M(i-1, j+1) + C_R(i, j) \end{cases} \quad (2)$$

with

$$C_L(i, j) = |I(i, j+1) - I(i, j-1)| + |I(i-1, j) - I(i, j-1)| \quad (3)$$

$$C_U(i, j) = |I(i, j+1) - I(i, j-1)| \quad (4)$$

$$C_R(i, j) = |I(i, j+1) - I(i, j-1)| + |I(i-1, j) - I(i, j+1)| \quad (5)$$

And $P(i, j)$ is an additional pixel based energy measure, for instance an energy function as defined in Sec. 2.1.

The advantage of this function is that it measures the energy due to newly inserted spurious edges. Indeed, artifacts may appear when a seam is removed and previously non adjacent pixels become neighbors. Recently, Frankovich and Wong

propose an absolute energy which is a combination of forward energy and energy gradient [9]. To take the temporal aspect into account, Rubinstein *et al.* [8] use graphcuts to compute temporal seams.

2.3 Seam carving for image/video compression

Seam carving has been recently applied to image/video compression [10][11][12][13][14][15].

In [10], Anh *et al.* introduce a content-aware multi-size image compression based on seam carving. More precisely, the proposed codec encodes an image into a content-aware progressive bitstream where the seams position and value are encoded. As for the seams position, the absolute coordinate x of the first pixel of the first column is encoded and then the difference between the first pixel of the previous column and the first pixel of the current column. For the other points, the difference between the current pixel and the previous pixel of the same seam is encoded. For the color, the same concept is used by replacing the x coordinate by the absolute R , G , B values. For all the encoded information, an arithmetic coding is used. To compute the seam carving, a saliency map multiplied with the gradient is used as energy function. A Region Of Interest (ROI) bounding box is calculated with the dimensions $w_0 \times h_0$ on the saliency map to control the reduction process and stop it when the reduced image reaches $1.5w_0 \times 1.5h_0$.

Tanaka *et al.* propose an image coding scheme which incorporates seam carving [11]. An image reduced by seam carving is transmitted, along with information about the seams position. As this information is very costly, the seams positions are approximated and the seams R , G , B values are not transmitted. More precisely, the seams positions are estimated as "pillars" of length N , such that the cost of a seam is $(\log_2 3)/N$ bpp. To define the pillar length, a top down approach is applied on a modified cumulative energy function (combination of the forward energy and seams bitrate with a Lagrangian multiplier). The stopping criterion is based on the cumulative energy: if a seam has a cumulative cost superior to a threshold, the reduction process is stopped. Results are evaluated with a SSIM metric [2].

In [12], Wang *et al.* simplify seam carving principle by requiring suppression of straight lines (vertically and horizontally). They combine reduction of the number of frames with the reduction of the spatial dimension by using this method. The energy function is a gradient computed in the same direction as seam carving. A diffusion function is used to avoid seams which are too close to one another, as it may create visual distortions. The evaluation is done with PSNR.

The method proposed in [13] improves upon [11] to limit artifacts created during interpolation. The idea is that the seams should avoid textured areas and the border of objects. If seams cross these areas or are too close to one another, artifacts may appear during reconstruction by linear interpolation. Therefore seams are required to pass through uniform areas even though these may include important regions of the image. For this purpose, a bottom-up approach is used instead of a top-down approach to define the length of an optimal pillar and to update the Lagrangian multiplier.

In [14], the authors apply the method in [11] for video reduction based on the graphcut approach of [8]. The same seam is deleted for the current GOP and an 8-connexity is allowed with the seam from the previous and the next GOP in order to avoid artifacts at the transition between GOPs. To compute the seam for the current GOP, all its frames are used but the intra-frame is given more importance. During the evaluation, only one Quantization Parameter (QP) is used for all the frames of the two test sequences. In [15], a piecewise linear approximation is proposed to find the optimal seam. The novelty is that the pieces of seams can have different directions and lengths.

In [11][13][14][15], the seam computation is modified by combining the forward energy with the seams encoding cost.

2.4 Limitations of current seam carving approaches for image/video compression

When seam carving is applied to image and video compression, the decoder has to synthesize seams data discarded at the encoder in order to recover the original dimensions. The traditional approach for up-sampling by seam carving is to apply the same process as in down-sampling by seam carving, but instead of deleting seams, to rather insert them for instance by duplication or interpolation. Two major difficulties arise. First, this process does not ensure that the inserted seams are in the same position as the deleted seams. Second, pixel duplication or interpolation may introduce artifacts in the synthesized image or video sequence.

The approaches discussed in Sec. 2.3 do not use seam carving up-sampling to synthesize the seams. Instead, the position of each seam is encoded. However, the cost to transmit this information may be prohibitive compared to the saving resulting from the size reduction. Consequently, these approaches only suppress few seams in order to limit seams data overhead and to avoid synthesis problems.

Our main contribution is a new way to represent seams for compression and synthesis. More specifically, a first step is to identify groups of seams and then to select the most relevant ones which are encoded and transmitted. This approach results in very low overhead to transmit seams position. We have also defined a new way to find seams which avoids important areas by modifying the cumulative energy maps. If a seam cross an important region, instead of stopping the process of reduction, the last value of the seam in the cumulative energy map is saturated and the seam carving is restarted. For the synthesis, a modified cumulative energy map allows good reconstruction of the scene geometry and preservation of the salient objects.

3. PROPOSED SEMANTIC VIDEO CODING USING SEAM CARVING

Given the limitations of current seam carving approaches for image/video compression, our objective is to significantly reduce the bitrate while preserving the semantic content. For this purpose, we propose a new approach for semantic video compression with seam carving.

3.1 General approach

A block diagram of the proposed architecture is shown in Figure 1. The different modules will be described in details hereafter. The main idea is to use seam carving to reduce the dimension of the video sequence. Then, the reduced video sequence is encoded with a standard H.264/AVC codec and the seams are encoded with our proposed scheme. After transmission, the videos sequence is reconstructed at the decoder side to recover the original dimension and to preserve scene geometry.

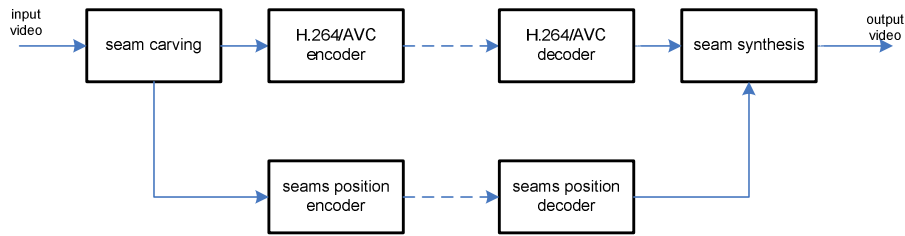


Figure 1. Architecture of proposed semantic video coding using seam carving.

The input video sequence is first divided into GOPs with a predefined size.

Seam carving is next applied to each frame of the GOP as illustrated in Figure 2. As discussed in the state-of-the-art, several methods exist to compute the energy function and the cumulative energy function. In this paper, we choose the saliency map of [17] combined with a gradient as energy function. The proposed seam carving algorithm is detailed in Figure 2 (a). The saliency map of [17] has the advantage to be multiscale, needs no training and is computed in the perceptual color space CIE Lab. To take into account temporal aspects, the current saliency map is linearly weighted with the previous one. To smooth the block effect from the multiscale approach while preserving the important parts of the frame, we then multiply the saliency map with a gradient to form our energy function.

As the forward energy in [8] has proved to be more efficient than the backward energy [3], it is chosen as cumulative energy function.

In parallel, we apply a threshold *Binarisation_Threshold* on the saliency map to separate the important objects from the background and to obtain a saliency binary mask. Next, a median filter is first applied to suppress small artifacts, followed by a dilatation to avoid that seams pass too close to salient regions.

To control the number of seams to be removed, the following process is carried out. Seam carving is first executed until the stopping point when a seam passes inside the saliency binary mask is reached. At this stage, the last position of the seam is saturated. Seam carving can then continue while avoiding salient regions. This saturation process is repeated *Nb_Seam_Jump* times. This approach allows to avoid important areas even though the cumulative energy is low.

As shown in Figure 2 (b), seam carving is applied vertically and horizontally on each frame, using the algorithm of Figure 2 (a). As the standard video coding H.264/AVC uses a 16x16 macro block structure, we suppress seams by multiples of 16.

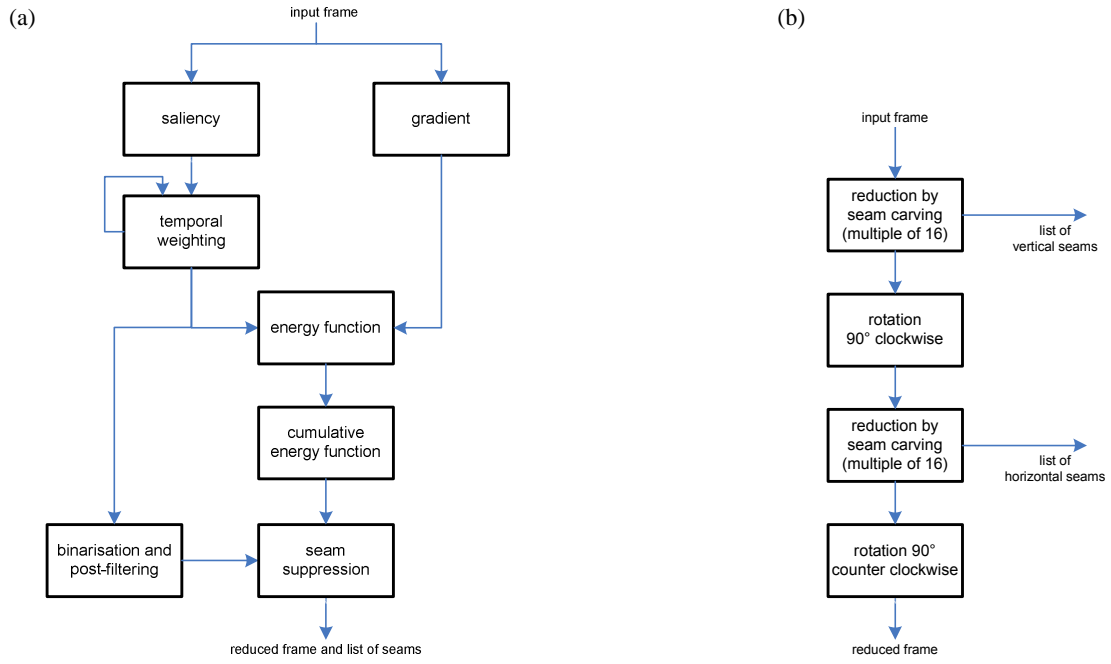


Figure 2. Proposed technique: (a) detailed seam carving algorithm, (b) application to vertical and horizontal seam carving reduction.

The GOP-based seam carving process is illustrated in Figure 3. It is desirable to keep the spatial dimension for the GOP constant and in multiple of 16. In this way, no pixel padding is needed and the effectiveness of H.264/AVC video coding is improved. Moreover, it helps guaranteeing temporally coherent processing. For this purpose, seam carving is applied on each frame of the GOP as described in Figure 2. The number of seams that can be suppressed for each frame of the GOP is obtained. Next, we choose as selection criterion the smallest number of seams removable among the number of seams suppressible for each frame of the GOP. Finally, the reduced GOP frames are encoded using H.264/AVC video coding. In parallel, seams position is encoded as described in Sec. 3.2 and 3.3.

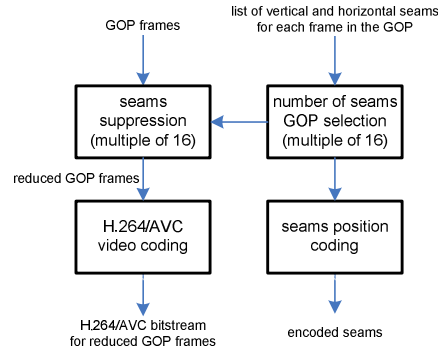


Figure 3. GOP-based seam carving process.

3.2 Seams analysis to define group of seams

Our objective is to encode seams in a lossy way with few bits. In this section, we present a method to identify key points, or group of seams, in order to avoid a large data overhead. It is motivated by the observation that seams avoid important objects in the image and are therefore concentrated in some areas. More specifically, key points are automatically detected and their positions are transmitted. For the sake of simplicity, we describe the method in the case of a vertical seam carving, but it is straightforward to extend it to the horizontal case. Table 1 summarizes this approach as pseudocode.

Table 1. Pseudocode for group of seams detection.

Define Group of seams
For Idx_Line = 1 : Height
For Idx_Seam_In_Line = 1 : Nb_Seam_V
If Idx_Seam_In_Line == 1
Nb_Groups[Idx_Line] = 1
Position_Group_Seam[Idx_Line, Nb_Groups[Idx_Line]] = Seam[Idx_Seam_In_Line]
Nb_Seams_In_Group[Idx_Line, Nb_Groups[Idx_Line]] = 1
Else
If Seam [Idx_Seam_In_Line] - Seam[Idx_Seam_In_Line-1] < T_Gpe
Nb_Seams_In_Group[Idx_Line, Nb_Groups[Idx_Line]]++
Else
Nb_Groups[Idx_Line] ++
Position_Group_Seam[Idx_Line, Nb_Groups[Idx_Line]] = seam[Idx_Seam_In_Line]
Nb_Seams_In_Group[Idx_Line, Nb_Groups[Idx_Line]] = 1
End // if
End // if
End // Idx_Seam_In_Line
End // Idx_Line
Select key lines
For Idx_Key_Line = 1 : Nb_Key_Lines
Current_Line = Idx_Key_Line*Height/Nb_Key_Lines
Size_Biggest_Group = 0
Position_Key_Line[Idx_Key_Line] = Current_Line
For Idx_Oscillate = - $K_{Oscillate}/2$: $K_{Oscillate}/2$
If (Current_Line + Idx_Oscillate) > 0 && (Current_Line + Idx_Oscillate) < Height
For Idx_Group = 1 : Nb_Groups[Current_Line]
If Nb_Seams_In_Group[Current_Line + Idx_Oscillate, Nb_Groups[Current_Line]] > Size_Biggest_Group
Size_Biggest_Group = Nb_Seams_In_Group[Current_Line + Idx_Oscillate, Nb_Groups[Current_Line]]
Position_Key_Line[Idx_Key_Line] = Current_Line
End // if
End // Idx_Group
End // if
End // Idx_Oscillate
End // Idx_Key_Line

First, for each line perpendicular to the seam carving path, we scan the position of the seams and regroup together the positions. More precisely, given a current seam position, if the distance to the next seam position is under a threshold T_Gpe , we add the position to the group, otherwise we create a new group. The process is iterated until the end of the line. The next step is the selection of key groups on a predetermined number of lines Nb_Key_Lines . To initialize the process, we select equidistant reference lines. Then, each line y oscillates in the interval

$$\left\{y - \frac{K_{oscillate}}{2}; y + \frac{K_{oscillate}}{2}\right\} \quad (6)$$

and the line having the largest group is selected. Afterwards, we save the position of the key line and all the group of seams position on this line. We underline that it is not judicious to keep only the position of the largest group of seams on the key lines as it may results in shifting and distorting the frame during seams synthesis.

Figure 4 illustrates our parameters in the case of a vertical seam carving for a 10x5 pixels image. Each color represents a seam. If $T_Gpe=0$, the first line ($y=1$) is composed of one group of 4 seams and a second group of 2 seams. If $T_Gpe=2$, the first line is composed of one group of 6 seams. We continue our illustration by fixing $T_Gpe=1$ and a number of key line $Nb_Key_Lines=1$. The key line is initialized at the line $y=3$. The next step is to find the largest group of seams. If $K_{oscillate}=0$, the key line selected is $y=3$ and the largest group of seam is the first one with 4 seams. If $K_{oscillate}=2$, the key line selected is the line $y=2$ and the largest group of seams have 6 seams. In this case, we save the position of the line $y=2$, the position of the group of seam, $x=3$ and the number of seams inside which is 6.

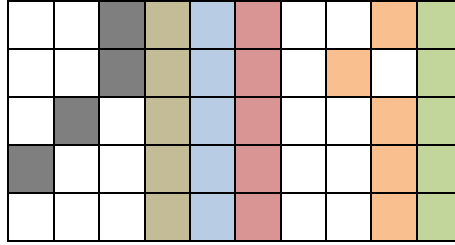


Figure 4. Example of seams analysis to define a group of seams.

3.3 Group of seams encoding

To encode the seams position while limiting the overhead cost, we exploit the characteristics of our data representation. For the sake of simplicity, we describe the encoding for vertical seams, but it is straightforward to extend it to the horizontal case. Table 2 summarizes the procedure as pseudocode.

For each GOP, we first encode the number of vertical seams suppressed, $Nb_Seams_Suppressed_V$, with $(\lceil \log_2(Width) \rceil)$ bits and the number of key lines, $Nb_Key_Lines_V$, with $(\lceil \log_2(Height) \rceil)$ bits.

For each frame of the GOP, we then encode the maximum number of groups of seams, noted Nb_Groups_Max , with $(\lceil \log_2(Nb_Seams_Suppressed_V) \rceil)$ bits, and the maximum number of seams in a group, noted Nb_Seams_Max , with $(\lceil \log_2(Nb_Seams_Suppressed_V) \rceil)$ bits.

For each key line, we first encode the number of groups for this key line, Nb_Groups , with $(\lceil \log_2(Nb_Groups_Max) \rceil)$ bits. Next, we encode the coordinate of the line, $Position_Key_Line$, with $(\lceil \log_2(Height_Reduce - Coordinate_Previous_Line) \rceil)$ bits. For the first line $Coordinate_Previous_Line = 0$ and for the other lines, $Coordinate_Previous_Line$ is the position of the previous line.

Finally, for each group in the key line, we encode the coordinate of the group with $(\lceil \log_2(Width_Reduce - Coordinate_Previous_Group) \rceil)$ bits. For the first group, $Coordinate_Previous_Group=0$ and for the subsequent groups, $Coordinate_Previous_Group$ is the position of the previous group. We encode also the number of seams in the groups with $(\lceil \log_2(Nb_Seams_Max - Nb_Seams_Accumul) \rceil)$ bits. For the first group, $Nb_Seams_Accumul=0$, and for the others groups $Nb_Seams_Accumul$ is the number of seams previously accumulated in the key line.

Table 2. Pseudocode for group of seams encoding for vertical seams; the process is similar for horizontal seams.

	Semantic <i>Number of bits used for coding</i>
For Idx_Gop = 1 : Nb_Gop	
Nb_Seams_Suppressed_V	number of vertical seams suppressed ($\lceil \log_2(\text{Width}) \rceil$)
Nb_Key_Lines_V	number of key lines ($\lceil \log_2(\text{Height}) \rceil$)
For Idx_Frame = 1 : Nb_Frames_In_Gop	
Nb_Groups_Max	maximum number of groups of seams ($\lceil \log_2(\text{Nb_Seams_Suppressed_V}) \rceil$)
Nb_Seams_Max	maximum number of seams in a group ($\lceil \log_2(\text{Nb_Seams_Suppressed_V}) \rceil$)
For Idx_Key_Line = 1 : Nb_Key_Lines	
Nb_Groups[Idx_Key_Line]	number of groups for this key line ($\lceil \log_2(\text{Nb_Groups_Max}) \rceil$)
If Idx_Key_Line == 1	
Position_Key_Line[Idx_Key_Line]	coordinate of the key line ($\lceil \log_2(\text{Height_Reduce}) \rceil$)
Else	
Position_Key_Line[Idx_Key_Line]	coordinate of the key line ($\lceil \log_2(\text{Height_Reduce} - \text{Position_Key_Line}[\text{Idx_Key_Line}-1]) \rceil$)
End	
For Idx_Group = 1 : Nb_Groups[Idx_Key_Line]	
If Idx_Group == 1	
Position_Group_Seam[Position_Key_Line[Idx_Key_Line], Idx_Group]	Coordinate of group of seams ($\lceil \log_2(\text{Width_Reduce}) \rceil$)
Nb_seams_Accumul = 0	Number of seam previously encoded
Nb_Seams_In_Group [Position_Key_Line [Idx_Key_Line], Idx_Group]	Number of seams in the group ($\lceil \log_2(\text{Nb_Seams_Suppressed_V}) \rceil$)
Else	
Position_Group_Seam[Position_Key_Line[Idx_Key_Line], Idx_Group]	Coordinate of group of seams ($\lceil \log_2(\text{Width_Reduce} - \text{Position_Group_Seam}[\text{Position_Key_Line}[\text{Idx_Key_Line}], \text{Idx_Group}-1]) \rceil$)
Nb_seams_Accumul += Nb_Seams_In_Group [Position_Key_Line [Idx_Key_Line], Idx_Group-1]	Number of seam previously encoded
Nb_Seams_In_Group [Position_Key_Line [Idx_Key_Line], Idx_Group]	Number of seams in the group ($\lceil \log_2(\text{Nb_Seams_Suppressed_V} - \text{Nb_seams_Accumul}) \rceil$)
End	
End // Idx_Group	
End // Idx_Key_Line	
End // Idx_Direction	
End // Idx_Frame	
End // Idx_Gop	

3.4 Seam synthesis controlled by group of seams

In this section, we present the modified energy map to require seams to pass through the key points defined during the seams analysis. This approach prevents the seams to pass through important objects and allows the significant regions to find their original position.

To proceed, we follow these different steps. While the frame has not reach its original size, we compute the energy map of the frame at the current dimension. For each key line, we retrieve the positions of the group of seams. Afterwards, we compute the energy cumulative map. Every time we reach a key line, the energy cumulative map is modified by adding a value *Value_Add* to the line excepted at the position of group of seams. We continue by finding the seam that minimizes the cumulative energy map and we add the seam in the frame and compute the pixel value. For the seams pixel value, we can use inpainting as in [16] or a simple linear interpolation as in [3][11][13][14]. The linear interpolation sometimes creates synthesis artifacts.

Figure 5 shows an example of the approach for the sequence Soccer with the seams and key lines visible. In Figure 5 (a), the key lines correspond to the main objects. In Figure 5 (b), the reconstruction seams are well positioned on the key lines. The objects are not deformed and the geometry of the image is preserved.



Figure 5. Example with group of seams visible and synthesized seams:
(a) seams (in orange) with key points (in black) selected with our approach, (b) synthesized seams (in orange).

4. RESULTS

To illustrate our results, we make the choice to use 3 different CIF sequences: Coastguard, Container, Soccer and 1 SIF sequence: Tennis. The sequences are temporally subsampled to 5 fps and we make our tests on 5 GOPs of 3 frames.

The proposed seam carving method adapts the amount of reduction to the content. For instance, we reach average spatial reduction of 50 % for Coastguard, 19.4% for Soccer, 26.55% for Tennis and 27.12% for Container. It is explained by the observation that salient objects are relatively small in Coastguard, whereas Soccer includes larger salient objects with fast motion. Tennis and Container fall in-between.

The different parameters have been heuristically fixed to $T_{Gpe}=1$, $Nb_Key_Lines=4$, $K_{Oscillate}=20$ pixels, $Value_Add=10000$, $Binarisation_Threshold=0.75$, and $Nb_Seam_Jump=16$. We compare our approach with a standard H.264/AVC coding scheme, an approach using seam carving but without transmission of seams and one with a lossless coding of all the seams positions. H.264/AVC is used in intra coding mode. For the approach without transmission of seams, we make the following improvement. Instead of inserting N seams by finding the N first minimum of energy in the cumulative energy map, we use a binary map based on the saliency map to control the insertion. More precisely, the seam carving process finds the minimum of energy until an important area on the binary map is reached. At this moment, X seams are inserted. This process continues until N seams have been inserted and the image reaches its original size.

Overhead information has to be send to synthesis the seams at the decoder side and avoid object distortions and geometry scene errors. The corresponding data rate is independent of the video coding bitrate. Therefore, the cost to transmit the seams becomes more important at high compression ratios.

The reinsertion of the seams creates two kinds of distortion: the first one due to the position of the reinsertion which may create local translation, and the second one due to the padding of the seam. We underline that we are interested to preserve only the important content, whereas the background can be more distorted. As we apply the same quantization for all the methods, we assume that the quality of the important regions remains the same. However, full-reference metrics such as PSNR and SSIM [2] are not appropriate to assess the quality of our proposed method.

Figure 6 illustrates the coded bitrate for different video sequences and for different Quantization Parameters Intra (QPI). These results are also summarized in Table 3. Bitrate saving obtained by seam carving depends strongly on the amount of spatial reduction, but also on the background content. For instance, a suppressed moving texture leads to more significant bitrate saving than the suppression of uniform areas. The proposed technique leads to significant bitrate savings at high bitrates (low QPI), when compared to H.264/AVC. However, these gains decrease along with a decreasing bitrate (high QPI). At mid-range (QPI=24), our approach results in bitrate savings of 31.55%, 10.69%, 7.92%

and 8.68% for Coastguard, Soccer, Tennis and Container respectively. For all sequences, the bitrate obtained with our proposed approach is very close to the one with the method without seams transmission. In other words, the overhead to transmit seams position with the proposed technique is very small. Finally, the transmission of all the seams position results in a very large overhead. Hence, this variant is only viable at very low QPI and performs poorly at high QPI.

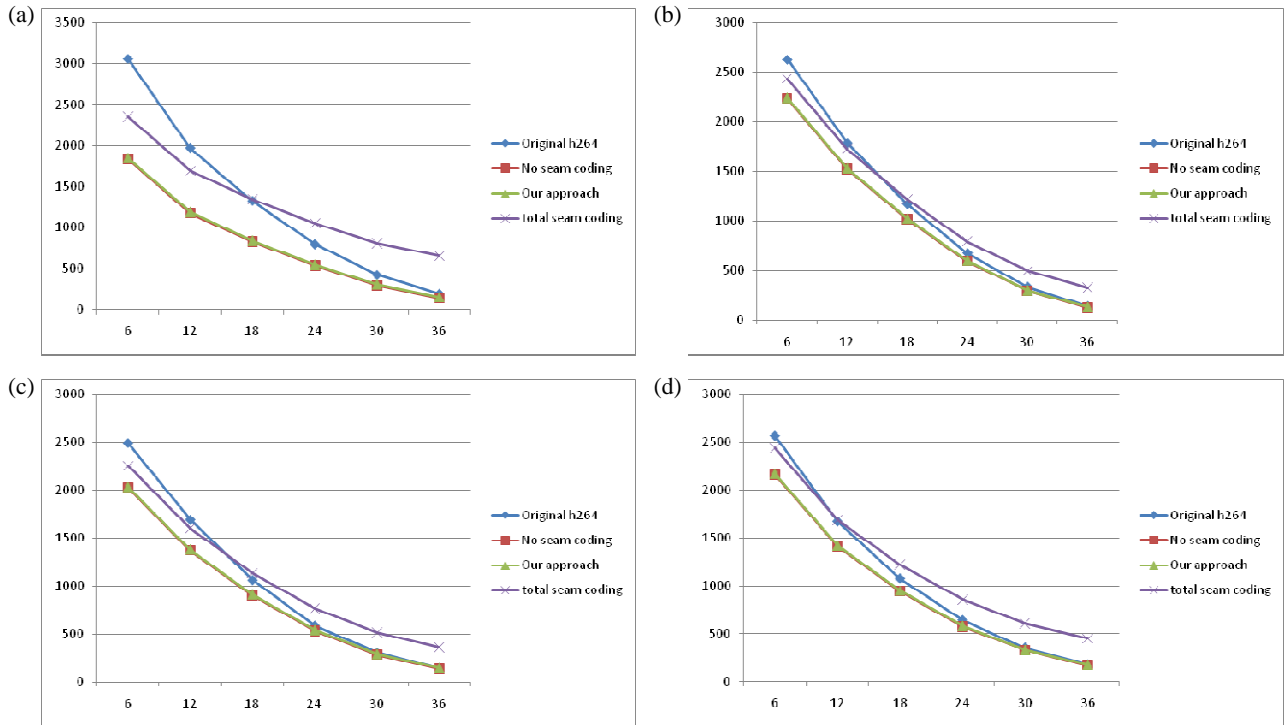


Figure 6. Bitrate (kbps) as a function of quantification parameter (QPI):

(a) Coastguard with 50% spatial reduction for the total sequence, (b) Soccer with 19.4% spatial reduction for the total sequence, (c) Tennis with 26.55% spatial reduction for the total sequence, (d) Container with 27.12% spatial reduction for the total sequence.

Table 3. Percentage of bitrate saved compared to H.264/AVC as a function of QPI:

(1) no seam coding, (2) our proposed approach, (3) total seam coding.

	Coastguard : 50% spatial reduction for the total sequence			Soccer : 19.4 % spatial reduction for the total sequence			Tennis : 26.55% spatial reduction for the total sequence			Container: 27.12% spatial reduction for the total sequence		
QPI	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
6	39.88	39.47	23.03	14.95	14.69	7.35	18.74	18.50	9.60	15.78	15.50	4.93
12	40.35	39.71	14.27	14.62	14.24	3.44	18.90	18.54	5.44	15.46	15.03	-1.20
18	37.41	36.47	-1.40	12.90	12.32	-4.24	14.90	14.33	-6.53	12.28	11.61	-13.63
24	33.11	31.55	-31.22	11.70	10.69	-18.00	8.96	7.92	-29.85	9.81	8.68	-33.77
30	30.34	27.40	-90.8	10.91	8.87	-49.49	6.79	4.79	-67.79	6.93	4.88	-72.20
36	28.75	22.37	-234.3	8.89	3.94	-137.3	7.54	3.34	-149.4	5.24	1.25	-149.0

Figure 7, Figure 8, Figure 9 and Figure 10 show some coded images examples at QPI=24 for the same four approaches under consideration. The synthesis of the seams allows keeping the context such that the visual quality of the important regions is preserved. In Container and Soccer, the synthesis of the seams in the background is very hard to noticed with

our approach and the results are very close to the approach with the transmission of all the seams positions. In Coastguard and Tennis, as we use a linear interpolation and as seams pass through texture areas, artifacts are visible. However, the context remains preserved. Nevertheless, the synthesized sequences give visual results near the approach that transmit all the seams positions. Conversely, results without transmission of seams are rather poor. The player is severely distorted in Tennis. In Coastguard, the bottom of the coastguard boat is disturbed and its position is shifted. In Soccer, the three soccer players are disturbed. The two on the left are bigger and the one on the right is taller.

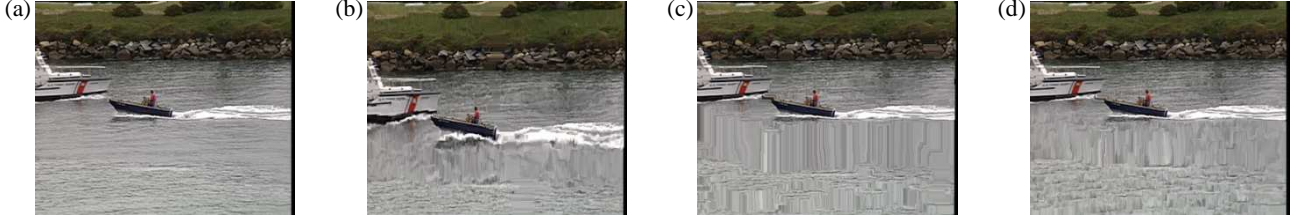


Figure 7. Coastguard: (a) conventional H.264/AVC, (b) no seam coding, (c) our proposed approach, (d) total seam coding.



Figure 8. Soccer: (a) conventional H.264/AVC, (b) no seam coding, (c) our proposed approach, (d) total seam coding.

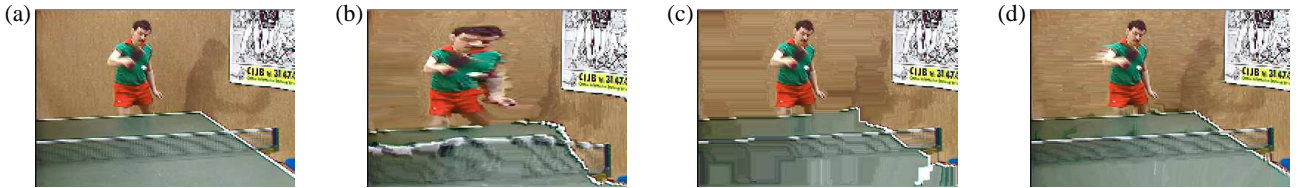


Figure 9. Tennis: (a) conventional H.264/AVC, (b) no seam coding, (c) our proposed approach, (d) total seam coding.

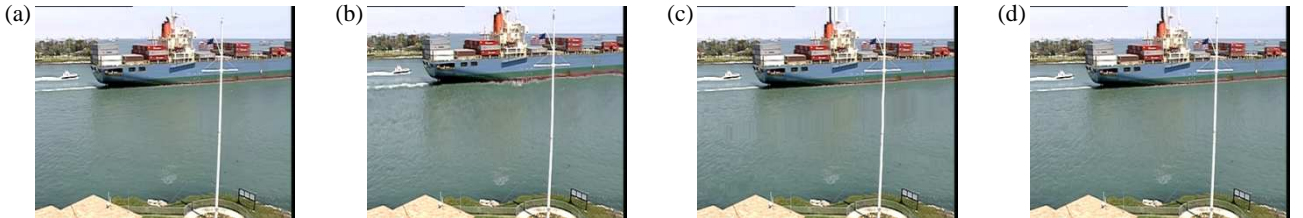


Figure 10. Container: (a) conventional H.264/AVC, (b) no seam coding, (c) our proposed approach, (d) total seam coding.

5. CONCLUSION

The main contribution of this paper is the development of a new semantic content-adaptive video coding scheme. We apply seam carving to reduce the spatial resolution of the video and compress the resulting reduced video sequences with a standard video codec. A new low cost representation of seams is introduced. This information controls the seam carving up-sampling during seams synthesis at the decoder side, in order to preserve the semantic of the video sequences. The salient objects are preserved, and the geometry scene is well reconstructed. Compare to conventional H.264/AVC coding, the proposed technique leads to significant bitrate savings.

Future works will address new metrics to assess seam carving performance, a better handling of temporal aspects, and an improved approach to synthesis the background.

REFERENCES

- [1] Wiegand, T., Sullivan, G. J., Bjøntegaard, G. and Luthra, A., "Overview of the H.264/AVC Video Coding Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, (2003).
- [2] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., "Image quality assessment: From error visibility to structural similarity", IEEE Transactions Image Processing, vol. 13, no. 4, (2004).
- [3] Avidan, S. and Shamir, A., "Seam Carving for Content-Aware Image Resizing", ACM Transactions Graphics, vol. 26, no.10, (2007).
- [4] Hwang, D.-S. and Chien, S.-Y., "Content-aware image resizing using perceptual seam carving with human attention model", IEEE International Conference on Multimedia and Expo, (2008).
- [5] Chen, D.-Y. and Luo, Y.-S., "Content-Aware Video Seam Carving Based on Bag of Visual Cubes", Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), (2010).
- [6] Chiang, C.-K., Wang, S.-F., Chen, Y.-L. and Lai, S.-H., "Fast JND-Based Video Carving With GPU Acceleration for Real-Time Video Retargeting", IEEE Transactions on Circuits and Systems for Video Technology, (2009).
- [7] Achanta, R. and Susstrunk, S., "Saliency detection for content-aware image resizing", 16th IEEE International Conference on Image Processing (ICIP), (2009).
- [8] Rubinstein, M., Shamir, A. and Avidan, S. "Improved seam carving for video retargeting", ACM Transactions on Graphics, (2008).
- [9] Frankovich, M. and Wong, A., "Enhanced Seam Carving via Integration of Energy Gradient Functionals", IEEE Signal Processing Letters, (2011).
- [10] Anh, N., Yang, W. and Cai, J., "Seam carving extension: a compression perspective", in Proc. of the 17th ACM Int. Conf. on Multimedia, (2009).
- [11] Tanaka, Y., Hasegawa, M. and Kato, S., "Image coding using concentration and dilution based on seam carving with hierarchical search", in Proc. of the IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP), (2010).
- [12] Wang, T. and Urahama, K., "Cartesian resizing of image and video for data compression", IEEE Region 10 Conference TENCN 2010, (2010).
- [13] Tanaka, Y., Hasegawa, M. and Kato, S., "Improved image concentration for artifact-free image dilution and its application to image coding", in Proc. of the IEEE 17th International Conference on Image Processing (ICIP), (2010).
- [14] Tanaka, Y., Hasegawa, M. and Kato, S., "Image concentration and dilution for video coding", IEEE Image Electronics and Visual Computing Workshop - IEVC2010, (2010).
- [15] Tanaka, Y., Hasegawa, M. and Kato, S., "Seam carving with rate-dependent seam path information", in Proc. of the IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP), (2011).
- [16] Domingues, D.; Alahi, A. and Vandergheynst, P., "Stream carving: An adaptive seam carving algorithm", 17th IEEE International Conference on Image Processing (ICIP), (2010).
- [17] Rahtu, E. and Heikkilä, J.A., "Simple and efficient saliency detector for background subtraction", IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 1137-1144, (2009).