A MOTION-AUGMENTED SUPER-RESOLUTION SCHEME FOR VERY LOW-BIT-RATE VIDEO ENHANCEMENT

Zoran A. Ivanovski, Lina J. Karam

Department of Electrical Engineering
Arizona State Univesity
Tempe, AZ 85287-5706

Glen P. Abousleman

Compression, Communications & Intelligence Lab
General Dynamics C4 Systems
Scottsdale, AZ 85257

ABSTRACT

This paper presents a video enhancement scheme for very low-bit-rate compressed video. High compression ratios introduce visible compression artifacts and blurring. To enhance the quality of the compressed video, a super-resolution (SR) scheme is employed. However, the effectiveness of the SR process depends heavily on the amount of subpixel motion that exists between frames in a video sequence, and on the accuracy of the motion vectors. To increase the effectiveness of the SR process, we introduce the concept of added motion prior to video compression. The parameters of this added motion are sent to the receiver with the compressed bit-stream. The proposed SR scheme also exploits existing motion between frames, by estimating motion vectors for each pixel using the original (uncompressed) video. However, since transmitting the pixel-based motion vectors is not practical due to the resulting significant increase in side information, the proposed scheme selects and sends only a very small set of motion vectors based upon their significance in the super-resolution enhancement. The results obtained show a significant improvement in the visual quality of the video sequences, as well as improvement in PSNR.

1. INTRODUCTION

The area of digital video enhancement has received considerable attention in the past ten years. Numerous techniques have been proposed with different levels of success. Traditionally, video enhancement techniques are only applied on decompressed frames.

A number of super-resolution (SR) techniques have also been used for video enhancement [1] [2] [3] [4]. These SR techniques make use of multiple video frames for enhancing a single frame, as compared to the non-SR techniques that only make use of the single frame to be enhanced [5]. However, the effectiveness of the SR process depends on the amount of subpixel motion that exist between frames in a video sequence, and on the accuracy of the estimated motion vectors. The amount of the existing motion in different regions of the frame can vary significantly (from no motion to high motion). For example, when a static camera is used, the motion of the background is zero. When the camera is following an object of interest moving with a translational motion, the motion of that object is zero relative to the camera, which renders the SR process unusable for that object. In order to address this issue, we propose an SR scheme in which motion is added to the frames prior to compression.

Furthermore, the SR effectiveness depends on the accurate estimation of the existing motion for each pixel [6] [7]. However, with low-bit-rate video, accurate motion estimation is very difficult to achieve using the decoded frames because of the inherent compression artifacts. On the other hand, performing the motion estimation for each pixel using the original frames at the transmitter, typically results in a large amount of side information if all of the computed motion vectors are transmitted to the decoder. In this paper, we propose an efficient scheme for the prediction of motion vectors at the receiver side based on a limited set of transmitted motion vectors that is sent. For this purpose, the proposed scheme computes the motion vectors using the original (uncompressed) video frames; but only a very small limited set of motion vectors (LSMV) is selected and sent to the receiver, based upon their significance in the super-resolution enhancement. At the receiver, each pixel is assigned a motion vector from the transmitted LSMV to maximize the motion prediction performance.

The proposed SR enhancement algorithm is applied on the decoded frames, based on both the received added motion parameters and the existing motion that is predicted using the received LSMV. The proposed SR enhancement scheme is independent of the applied compression technique, and can work with any standard or non-standard compression scheme. Nevertheless, if a specific knowledge about compression is available, an optimization in terms of quality and computations, is possible.

The paper is organized as follows. Section 2 describes the proposed SR enhancement scheme, including the concept of added motion, the LSMV generation process and the applied SR algorithm. Results from performed experiments on different video sequences are presented in Section 3. A conclusion is given in Section 4.

2. PROPOSED MOTION-AUGMENTED SUPER-RESOLUTION VIDEO ENHANCEMENT SCHEME

A block diagram of the proposed SR video enhancement scheme is shown in Fig. 1. First, a motion estimation procedure is applied to the original (uncompressed) video sequence resulting in motion vectors for each pixel in the frame. The estimated motion vectors are then used in a motion analysis process to determine optimality of the existing motion and to estimate the amount of additional motion that needs to be introduced in the frames in order to optimize the SR enhancement process at the receiver. The resulting frames (containing added motion) are then compressed and sent to the receiver, together with the added motion parameters.

The LSMV generation stage is used for selecting and transmitting only a limited set of the motion vectors based on their signif-

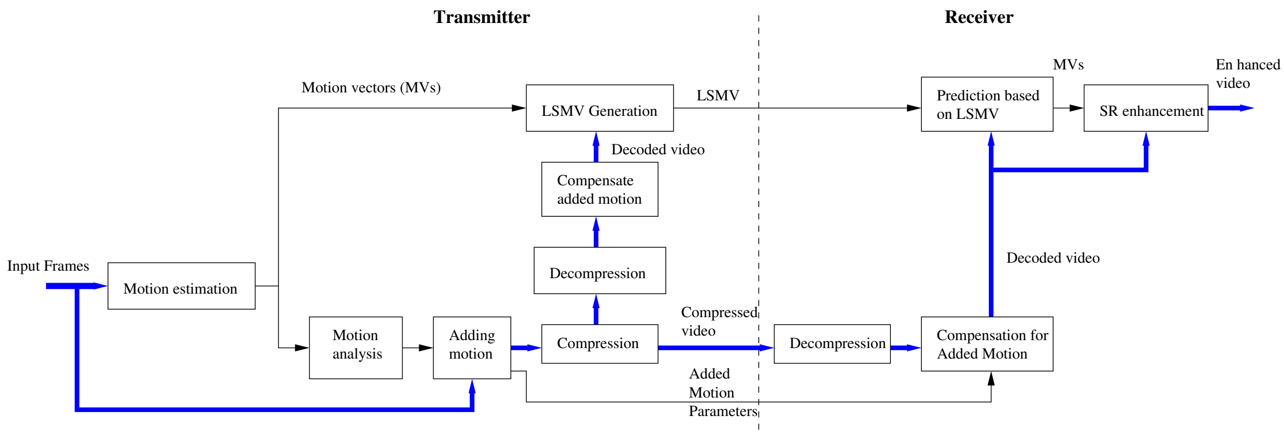

Fig. 1. Block diagram illustrating proposed motion-augmented SR enhancement scheme.

icance for the SR enhancement. For this purpose, it makes use of the decoded video sequence (decompressed and compensated for added motion) and of the full set of computed motion vectors in order to optimize the selection of the LSMV. Details on the SR enhancement procedure, the added motion and the LSMV generation procedures are given in Section 2.1, 2.2 and 2.3, respectively.

At the receiver (Fig. 1), the decompressed frames are motion compensated for the added motion, using the received added motion parameters. Next, a motion prediction is performed using the received LSMV and the decoded video frames in order to generate one motion vector per pixel. These generated motion vectors are then used in the SR algorithm to enhance the decoded frames.

The proposed scheme can be applied independent of the employed compression technique. However, motion vectors estimated from the original frames prior to the compression could also be used in the compression process for prediction, if inter-coding is used.

2.1. SR algorithm

Let Y_n represent the current decoded frame to be enhanced. The proposed SR scheme makes use of one past degraded frame Y_{n-1} and one future decoded frame Y_{n+1} in reconstructing a high-resolution enhanced version, $\widehat{X}_n$, of X_n.

The SR enhancement is performed through the minimization of a norm-L_2-based cost function, given by:

$$\underline{\widehat{X}}_n = \arg\min_{\underline{X}_n} \left[\sum_{k=-1}^{1} \|F_{n,n-k}\underline{X}_n - \underline{Y}_{n-k}\|_2^2 + \lambda \|Q\underline{X}_n\|_2^2 \right], \tag{1}$$

where $\underline{X}_n$ is a column vector representing undegraded high-resolution version of Y_n frame, and $\underline{Y}_{n-k}$ are column vectors representing the decoded frames Y_{n-k}, both in lexicographic order. $F_{n,n-k}$ represents geometrical warping between frames. To estimate this warping, the predicted motion vectors are used. Q is regularization matrix and λ is weighting factor.

Since the proposed scheme is not directly connected to the compression technique used, the regularization used in (1) poses only a smoothing constraint. Some specific regularization terms could be used with specific compression techniques; for example, an appropriate deblocking term could be used with DCT-based compression techniques.

2.2. Added Motion Procedure

As mentioned in Section 1, the amount of the existing motion in different regions of the frame can vary significantly. SR techniques are known to be more effective when a sufficiently high amount of motion is present; these perform poorly in the absence of motion or when only little motion is present. In the proposed scheme additional motion is introduced in the frames X_{n-1}, X_n and X_{n+1} in order to optimize the SR enhancement of the considered frame X_n. However, estimating optimal motion for every pixel or region in the frames is computationally expensive and can generate a large amount of additional motion parameters which should be sent to the receiver as a side information. In addition, applying the additional motion at a subpixel level to the frame requires interpolation which can degrade the high frequency components and can, also, cause overlapping at the borders of the regions with different motion. To address these issues, additional motion is introduced for each original frame X_{n-1}, X_n and X_{n+1} as a global translational shift S, that is applied uniformly to the pixels in a frame. Since, at very low bit-rate, compression causes artifacts and blurring that typically span more than two pixels, we restrict the shift to be of the order of one pixel in the horizontal (left/right) and vertical (up/down) direction. So, under these constraints, the problem of finding the optimal additional motion shift can be formulated as follows:

• for the first frame in the sequence, find the shifts S_1 and S_2, where $S_i \in \{(0,0), (-1,0), (1,0), (0,-1), (0,1), (-1,-1), (-1,1), (1,1)\}, i = 1,2$, such that when applied to the frames X_1 and X_2, respectively, would result in the best SR enhanced version of X_1 by minimizing considered SR cost function (1),

• for all the other frames in the sequence, find the shift S_{n+1}, where $S_{n+1} \in \{(0,0), (-1,0), (1,0), (0,-1), (0,1), (-1,-1), (-1,1), (1,1)\}$, given the shifts S_{n-1} and S_n, such that when applied to the frames X_{n-1}, X_n and X_{n+1}, respectively, would result in the best SR enhanced version of X_n by minimizing considered SR cost function (1).

Adding global integer motion is computationally inexpensive and leaves the frames almost unchanged, except for the very small

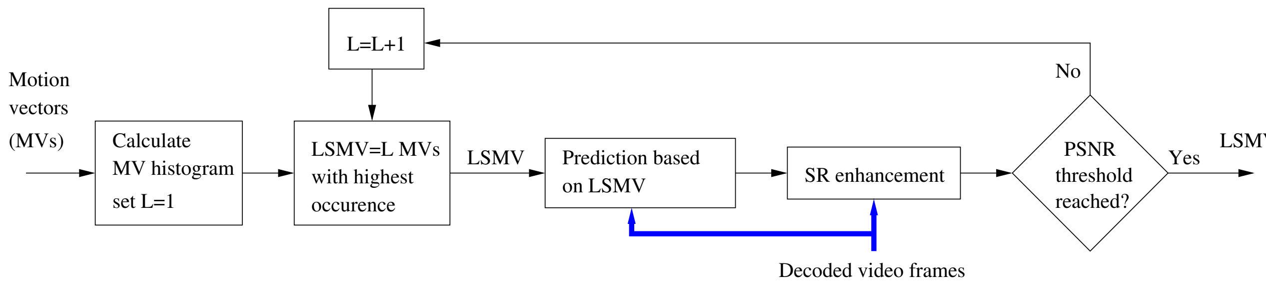

Fig. 2. Block diagram illustrating the process of generation of the limited set of motion vectors.

region near the borders of the frame, thus, avoiding the loss of high frequencies due to interpolation.

2.3. Generation of a Limited Set of Motion Vectors

Fig. 2 shows a block diagram illustrating LSMV generation process at the transmitter side. First, the histogram of the motion vectors estimated from original frames is calculated and sorted in decreasing order. The LSMV is generated in an iterative procedure, starting with one motion vector, the one with the highest histogram value (highest value of occurrence). At each iteration, the generated LSMV is used as part of a motion prediction procedure, in order to estimate predicted motion vectors for every pixel, using decompressed and added motion compensated frames. The resulting predicted motion vectors are then used in the SR procedure to enhance the decoded frame and the PSNR gain is computed. At every iteration, the size of the LSMV is increased by augmenting it with the next motion vector in the sorted list of motion vectors. The procedure terminates when a specified PSNR threshold is achieved or the number of motion vectors has reached a specified limit.

3. EXPERIMENTAL RESULTS

To test the performance of the proposed scheme on sequences with little motion, the QCIF (144×176) Akiyo sequence was used, at a frame rate of 10 frames per second (fps). The sequence was compressed at 50kbps using Motion JPEG2000. Added global motion used in the experiment was ± 1 pixel in the horizontal and/or vertical directions, as discussed in Section 2.2. The size of LSMV generated was approximately 1.3% of the total bit-rate.

Fig. 3 shows the PSNR gain that is obtained by applying the proposed SR scheme with (solid) and without (dashed) added motion. It can be seen that the proposed motion-augmented SR scheme with added motion results in a significantly higher PSNR gain as compared to scheme without added motion (only LSMV). The PSNR gain ranges from 1.1 to 2dB with an average of 1.6dB with the proposed added motion, while the PSNR gain ranges from 0.1 to 0.8dB with an average of 0.4dB without added motion.

The resulting improvement in visual quality is also very significant. When the original decoded sequence and the obtained SR enhanced decoded sequence are played back and compared on a CRT monitor it can be seen that the original decoded sequence suffers from very annoying compression artifacts including blur and ”mosquito” noise, which are significantly removed by the proposed motion-augmented SR scheme. Fig. 4 and 5 show, respectively, the decoded frames 10 and 97 of the Akiyo sequence without and with the proposed motion-augmented SR-enhancement.

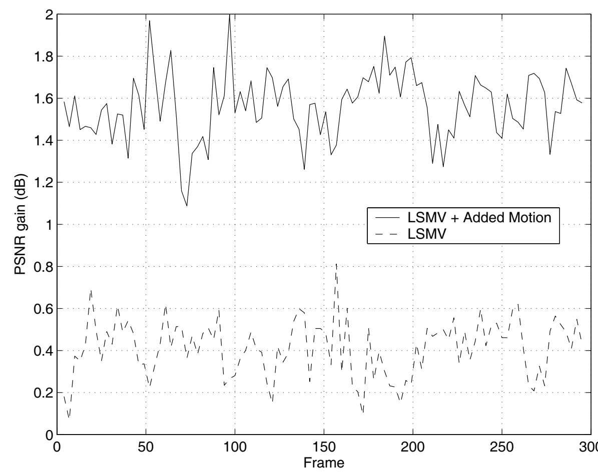

Fig. 3. PSNR gain of the proposed SR estimation for the Akiyo sequence.

4. CONCLUSIONS

A new motion-augmented SR enhancement scheme for very low-bit-rate compressed video is presented. To increase the effectiveness of the SR process, a new concept of added motion is introduced, in which optimality of the motion between frames with respect to SR enhancement is exploited. The existing motion between frames is also exploited using motion prediction procedure based on limited set of motion vectors. Both techniques significantly contribute in effectiveness of the SR enhancement process. The results obtained show improvement in PSNR of up to 2dB, with average of 1.6dB. Also, significant improvement of visual quality of the obtained SR enhanced decoded sequence has been achieved, compared to original decoded sequence.

5. REFERENCES

[1] Y. Altunbasak and A. J. Patti, “A maximum a posteriori estimator for high resolution video reconstruction from mpeg video,” *Proc. 2000 IEEE Conf. Image Processing*, vol. 2, pp. 649–652, 2000.

[2] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, “Multiframe resolution enhancement metods for compressed video,” *IEEE Signal Processing Letters*, vol. 9, no. 6, pp. 170–174, 2002.

(a) decoded frame

(b) SR enhanced decoded frame

Fig. 4. Frame 10 of Akiyo video sequence, compressed at bit-rate of 50kbps using MJPEG2000 without (a) and with proposed SR enhancement (b).

(a) decoded frame

(b) SR enhanced decoded frame

Fig. 5. Frame 97 of Akiyo video sequence, compressed at bit-rate of 50kbps using MJPEG2000 without (a) and with proposed SR enhancement (b).

[3] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian high-resolution reconstruction of low-resolution compressed video," *Proc. 2001 IEEE Conf. Image Processing*, vol. 2, pp. 25–28, 2001.

[4] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Trans. Image Processing*, vol. 13, no. 7, pp. 898 – 911, Jully 2004.

[5] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. Circ. Syst. Video Technology*, vol. 3, no. 6, pp. 421 – 432, December 1993.

[6] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint map registration and high-resolution image estimation using a squence of undersampled images," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1621–1633, December 1997.

[7] Z. Jiang, T.-T. Wong, and H. Bao, "Practical super-resolution from dynamic video sequences," *Proc. 2003 IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 549 – 554, 2003.