

Error concealment algorithm based on sparse optimization

Ting-Lan Lin¹ · Tsai-Ling Ding¹ · Chang-Yi Fan¹ ·
Wen-Chih Chen¹

Received: 14 April 2015 / Revised: 16 September 2015 / Accepted: 30 October 2015 /
Published online: 11 November 2015
© Springer Science+Business Media New York 2015

Abstract In this paper, an application of sparse optimization in the error concealment area is proposed. The spatial and temporal formulations of the pixels in the current frame and reference frame are proposed to solve the problem. Based on the sparse characteristics of nature images, we form sparse optimization problems for both formulations. The optimization problem is solved by the primal-dual interior point method. The solutions are combined for better results. By solving for a limited numbers of significant predictors using the sparse optimization, our algorithm performs subjectively and objectively better for the concealed result; compared to two state-of-the-art spatial-temporal hybrid error concealment methods, the proposed methods can improve by up to 0.19 dB and 1.12 dB in PSNR (Peak Signal-to-Noise Ratio).

Keywords Video communication · Packet loss · Error concealment · Sparse optimization · Primal-dual interior point method

1 Introduction

In video compression technology, the transmission and storage of video signals have become more and more popular. For the network environment, various channel/network errors result in damage or loss of compressed video information during transmission or storage. H.264/AVC is the most popular compression coding standard [16], outperforming the previous standards. It has removed redundancy in the spatial and temporal domain of a video, thus the bitstream is very sensitive to transmission error during the network. Unfortunately, packet errors are unavoidable, which will cause serious display quality degradation.

✉ Ting-Lan Lin
tinglan@cycu.edu.tw

Tsai-Ling Ding
dingtasiling@gmail.com

Chang-Yi Fan
doublechang015@gmail.com

Wen-Chih Chen
hydexiii@gmail.com

¹ Department of Electronic Engineering, Chung Yuan Christian University, Chung-Li, Taiwan

The loss of transmitted packets influences the quality of the received video. For combatting the transmission error, on one hand, error resilience adds redundant information at the encoder with the penalty of decreasing the compression efficiency. On the other hand, error concealment conceals the errors in the post-processing stage after receiving the information at the decoder side without modifying source and channel coding schemes. Since the real time transmission of video stream aims to reduce the delay, it is not possible to retransmit all erroneous or lost packets. Therefore, there is a need for post-processing methods that try to reduce the visual artifacts caused by bitstream error. The simplest error concealment method used in H.264/AVC is the frame copy method. In this method, the lost frame is estimated directly by copying the co-located pixels from the previous frame; this method would result in the frame discontinuity, especially for the high motion sequences. An alternative method called motion copy is also adopted in H.264/AVC, providing another option for error concealment. The concept of the motion copy method is that the motion vectors of the previous frame are used for concealing the current lost frame.

The error concealment algorithms in the literatures can be categorized into three sets: spatial approaches, temporal approaches, and hybrid approaches. Spatial approaches restore the corrupted macroblock by the information of surrounding macroblocks or pixels that are correctly received and decoded. A spatial error concealment method proposed by Wang et al. minimized the first-order derivative-based smoothness measurement [15], and Zhu et al. enhanced it by using second-order derivative that could suppress the induced blurring artifact [22]. In [12], an algorithm based on selective directional interpolation was proposed that could efficiently recover the smooth and edge areas.

Temporal approaches exploit temporal correlation between successive received frames to reconstruct the corrupted blocks. The most important part in temporal approaches is finding the appropriate blocks in previous frames that can be used as information to restore the corrupted blocks. The appropriate blocks are found by estimating lost motion vectors (MV). Much research has been proposed to recover the MVs of the corrupted block. The well-known boundary matching algorithm proposed in [9] used the information between the internal and external boundary of reconstructed blocks to select the MV that minimized the total variation. There are also more sophisticated algorithms [13,18] to obtain better replacements for the corrupted blocks.

Hybrid approaches combine the spatial and temporal approaches. A priority-driven region matching algorithm was proposed to exploit the spatial and temporal information in [4]. Specifically, the work in [19] proposed an auto-regression model that considers spatial and temporal information to recover the lost pixels. The work in [21] first performs a temporal method, called motion vector extrapolation, from the motion vectors in the previous frame to estimate the lost MVs. Then, based on the spatially neighboring available motion vectors around the lost MB, the estimated MVs are improved. The work of Yan et al. [17] considers whole frame loss, and the error concealment algorithm uses the motion extrapolation technique and received depth map to recover the lost frame from the previous frame. A recent work in [10] proposes a dynamic programming method to estimate the missing motion vectors in consecutive blocks, and the block merging procedure is performed for similar motion vectors. A loss frame error concealment method is proposed in [20] for wireless multimedia sensor networks using the inter-view information. In this paper, we propose a novel way to recover the lost pixel by using the theory of sparse optimization based on [19] and considering the sparse characteristics of nature images. Spatial and temporal sparse optimization problems are formulated for the spatial and temporal method. After the problems are solved, the solutions are combined to recover the lost pixels.

The differences between the proposed work and important related works References [12,15,16,22] are spatial approaches for error concealment, and references [9,13,18] are temporal approaches for error concealment, whereas the proposed method is a spatial-temporal approach for the problem. For other recent works, the work in [4] must first pre-process the pixels to find the edges in an image for recovery, whereas the proposed method can directly find lost motion vectors from pixels and the relations among them. The work in [17] performs the error concealment algorithm for color frame based on the depth frame; however, the proposed work does not deal with the transmission of depth frames. Furthermore, the work in [17] solves the problem in the case of whole frame losses, but the proposed work solves the problem for the packet losses in the checkerboard patterns. The algorithms in [10] discuss the lost pattern for consecutive lost blocks, whereas the proposed work focuses on the packet loss in the checkerboard pattern. The work in [20] concentrates on the loss recovery for the multi-view videos, while the proposed work is designed for single-view videos.

The proposed work is most similar with [19] since they both formulate the spatial and temporal relations among neighboring pixels and solve the problem based on matrix computation. Also, in [21], missing motion vectors are estimated using the established mathematical formulations among neighboring motion vectors. These two papers are similar to the proposed paper in that their approaches all depend on the developed mathematical and matrix formulations. Therefore, in this paper, we compare the proposed work against the state-of-the-art works in [19] and [21], due to the similarities among the works, to show the significance of the proposed method.

This paper is organized as follows: Section 2 introduces the theory of sparse optimization. Section 3 discusses the spatial and temporal formulation between pixels in the current and reference frames. The spatial and temporal sparse optimization problems are constructed and solved. Section 4 demonstrates the experimental results compared to two state-of-the-art methods objectively and subjectively. Section 5 concludes the paper.

2 Sparse optimization

The signal processing field proposed an exciting theory, known as “sparse optimization” [1–3,6]. In this new method, the sampling frequency can be greatly reduced and keep the quality of the reconstructed signal. The original signal must be consistent with the sparsity, which means the signal has a more concise expression with the appropriate basis. Considering a 1-D vector $f \in \mathbb{R}^n$ and its orthogonal basis $\Psi = [\psi_1, \psi_2, \dots, \psi_n]$, such as the basis of DCT, it can be expressed as:

$$f(t) = \sum_{i=1}^n z_i \psi_i(t) \quad (1)$$

where $z_i = f(t), \psi_i(t)$. This relationship translated into a matrix expression is shown as such:

$$f = \Psi z \quad (2)$$

where f represents a $n \times 1$ vector, $\Psi = [\psi_1, \psi_2, \dots, \psi_n]$ represents a $n \times 1$ matrix, and z represents a $n \times 1$ vector. We define the sparsity of signal as: when a signal is combined by orthogonal basis Ψ , and when most of the values in z are zero and only have \mathcal{K} non-zero values, we consider f to be \mathcal{K} -sparse. This means that, by restoring signal f , which has n data, we only need sampling \mathcal{K} -point ($\mathcal{K} < n$) in the basis domain, and f could be

reconstructed completely. However, we will not sample in the basis domain but in the signal domain. For example, an n -point signal is transmitted, and only m points are received. In this case, when restoring the signal, matrix theory states that there should be infinite solutions when $m < n$. But what amazes us about sparse optimization is that we could reconstruct the signal perfectly and uniquely. When considering Eq (2) and assuming f is \mathcal{K} -sparse, if we want to solve the unique z , it can be obtained through the following optimization problem:

$$z = \arg \min_{\tilde{z} \in \mathbb{R}^n} \|\tilde{z}\|_0 \text{ s.t. } \Psi\tilde{z} = f \tag{3}$$

where \tilde{z}_0 is \mathcal{L}_0 -norm representing the number of non-zero values in \tilde{z} . Therefore, we can find \tilde{z} , which has the least non-zero values and also conforms to the condition of $\Psi\tilde{z} = f$. However, it is very complicated to solve the \mathcal{L}_0 -norm problem. Fortunately, [3] and [5] proposed an approximated solution to \mathcal{L}_0 -norm:

$$z = \arg \min_{\tilde{z} \in \mathbb{R}^n} \|\tilde{z}\|_1 \text{ s.t. } \Psi\tilde{z} = f \tag{4}$$

Works [1] and [5] prove that the problem has high probability to restore the \mathcal{K} -sparse signal with $m \geq c\mathcal{K} \log(n/\mathcal{K})$ data, where c is a constant. Therefore, the sparser the signal, the easier it can be restored perfectly by the sparse optimization technique. In [14] and [7], the works discuss more methods of sparse optimization to restore the signal and how to restore the information when there is noise present (or when it is near-sparse). The \mathcal{L}_1 -norm optimization problem in Eq (4) can be efficiently solved by the primal-dual interior point method in [8].

3 The Spatial and Temporal Sparse Optimization (STSO)

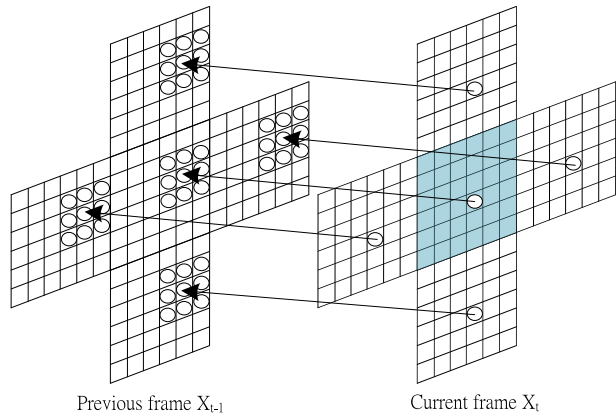
In this section, we collect data from the available pixels surrounding the lost MB as well as pixels in the reference frame, as done similarly in [10]. Pixels in two frames away are also used for the proposed method. The mathematical formulations of these two mechanisms are formulated, and the predicting coefficients are solved by sparse optimization in section 2. We then combine these two mechanisms to form an overall algorithm. We denote this proposed method as STSO (Spatial and Temporal Sparse Optimization).

3.1 Spatial Formulation of Sparse Optimization

For the spatial mechanism, the idea can be explained with Fig. 1. The gray area is the lost block. The initial motion vector is derived by averaging the motion vectors in the neighboring available MVs around the lost MB. The initial motion vector is used by pixel A in the neighboring blocks to point back to reference pixel B in the reference frame. Pixel B is used as a center to form a 3×3 subblock in the reference frame, as shown in Fig. 1. These $3 \times 3 = 9$ pixels are used to predict pixel A by predicting coefficient α :

$$x_t(i, j) = \sum_{k=-1}^1 \sum_{l=-1}^1 \alpha(k, l)x_{t-1}(i + dy + k, j + dx + l) \tag{5}$$

Fig. 1 Nine pixels in the reference frame to predict one available neighboring pixel in the current frame using initial estimated motion vectors



where x_{t-1} is the reference frame pixel, x_t is the current frame pixel, and $\alpha(k,l)$ is the weight on the reference pixel $x_{t-1}(i+dy+k,j+dx+l)$ to the current pixel $x_t(i,j)$. The (dx,dy) is the initial motion vector.

This procedure is repeated for all the pixels in the spatial neighbors in the lost block, and their corresponding 9 reference pixels can be located and collected. With the collected data and Eq (5), we write a formulation in a matrix form for the optimization problem:

$$x_t = X_{t-1} \alpha^T \tag{6}$$

The x_t is a vector containing the neighboring available pixels of the lost MB, and the X_t is the matrix containing corresponding reference pixels according to Eq (5). α^T is a 9×1 vector of predicting coefficients that we aim to determine. Figure 2 is a visual demonstration of the spatial pixel relation in Eq (6). Nine bases (256×1) are multiplied with corresponding 9×1 coefficients α and summed to form the target vector.

The sparse characteristics of nature images are discussed in [11]. *We use this property to recover the lost pixels in nature images in the error concealment application.* We solve for optimal α^T by the sparse optimization technique in section 2: we aim to find a coefficient vector α_{SO} with minimal 1-norm, such that the relation between neighboring available pixels and the reference pixels in Eq (6) holds:

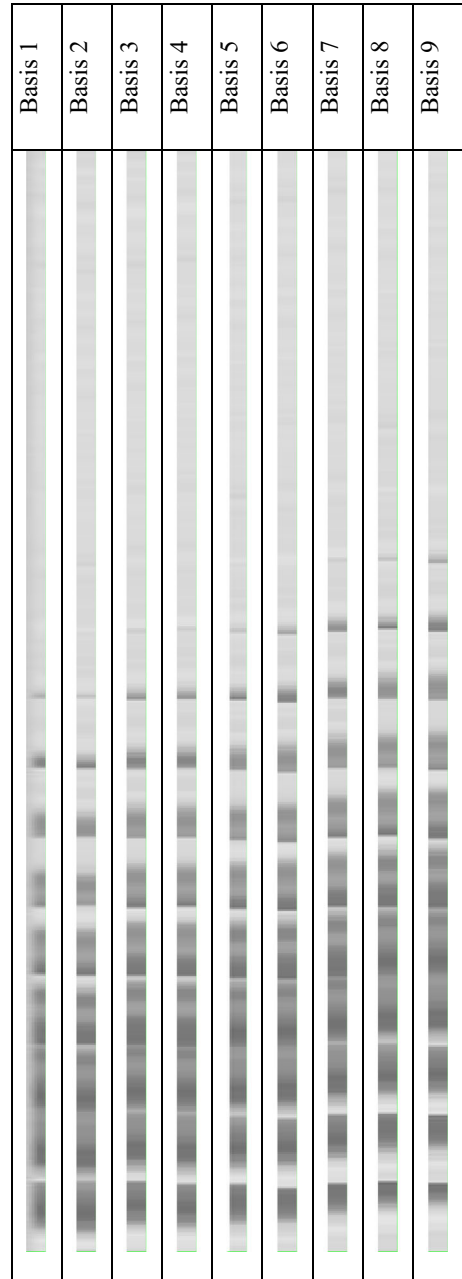
$$\begin{aligned} \alpha_{SO} &= \arg \min_{\alpha \in R^n} \alpha_1 \\ \text{s.t. } x_t &= X_{t-1} \alpha^T \end{aligned} \tag{7}$$

This optimization problem is a perfect fit for the sparse optimization technique introduced in section 2 and can therefore be solved efficiently by the primal-dual interior point method in [8]. After α_{SO} is obtained, the pixels in the lost MB can be found by:

$$x_{t, \text{spatial}}^{\text{lostMB}} = X_{t-1}^{\text{lostMB}} \alpha_{SO}^T \tag{8}$$

where $x_{t, \text{spatial}}^{\text{lostMB}}$ is a $(16 \times 16) \times 1$ vector containing the pixels in the lost MB and recovered by the spatial method. X_{t-1}^{lostMB} is a matrix containing the corresponding pixels in the reference frame according to Eq (5) and the initial motion vector (dx,dy) .

Fig. 2 The pixel illustration of the values in nine columns (each of size 256×1) in \mathbf{X}_{t-1} as nine bases



3.2 Temporal Formulation of Sparse Optimization

In this subsection, the temporal relations between pixels in a different frame are considered, as done similarly in [19]. Specifically, the pixels in one frame before the current frame and two frames before the current frame are considered, as shown in Fig. 3. The initial motion vector

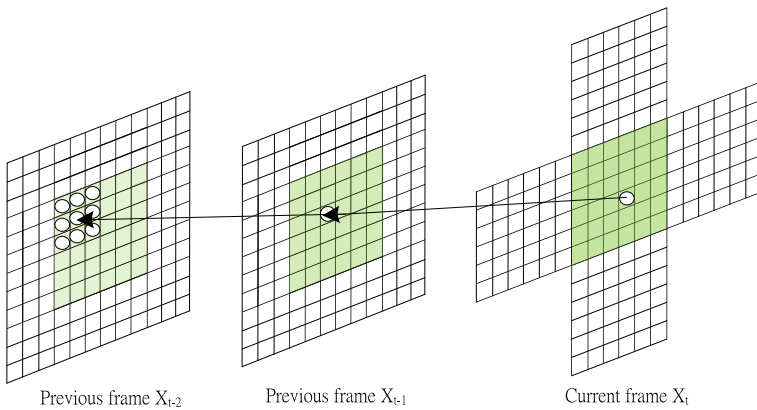


Fig. 3 Temporal relations of pixels

(dx, dy) is used to point from the current pixel to pixel C one frame before the current frame. And the available motion vectors in the previous frame are used to point back to pixel D two frames before the current frame. Centered at D, we collect the surrounding $3 \times 3 = 9$ pixels to predict pixel C. After considering all pixels in the reference frame, the matrix formulation can be written as:

$$\mathbf{x}_{t-1} = \mathbf{X}_{t-2} \boldsymbol{\beta}^T \tag{9}$$

The \mathbf{x}_{t-1} is a vector containing the available pixels one frame away, and the \mathbf{X}_{t-2} is the matrix containing corresponding reference pixels two frames away. $\boldsymbol{\beta}^T$ is a predicting coefficient vector describing the temporal relationship. Figure 4 visually demonstrates the temporal formulation in Eq (9), where nine bases with their corresponding temporal predicting vector $\boldsymbol{\beta}$ are used to form the target vector \mathbf{x}_{t-1} . Again, based on the sparse characteristics of nature images [11], we aim to find a temporal coefficient vector $\boldsymbol{\beta}_{SO}$ with minimal 1-norm, such that the relation among temporally neighboring available pixels in Eq (9) holds:

$$\begin{aligned} \boldsymbol{\beta}_{SO} &= \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^n} \|\boldsymbol{\beta}\|_1 \\ \text{s.t. } \mathbf{x}_{t-1} &= \mathbf{X}_{t-2} \boldsymbol{\beta}^T \end{aligned} \tag{10}$$

The problem is identical to the one in section 2 and can be solved efficiently by [8]. After $\boldsymbol{\beta}_{SO}$ is found, the pixels in the lost MB can be found by:

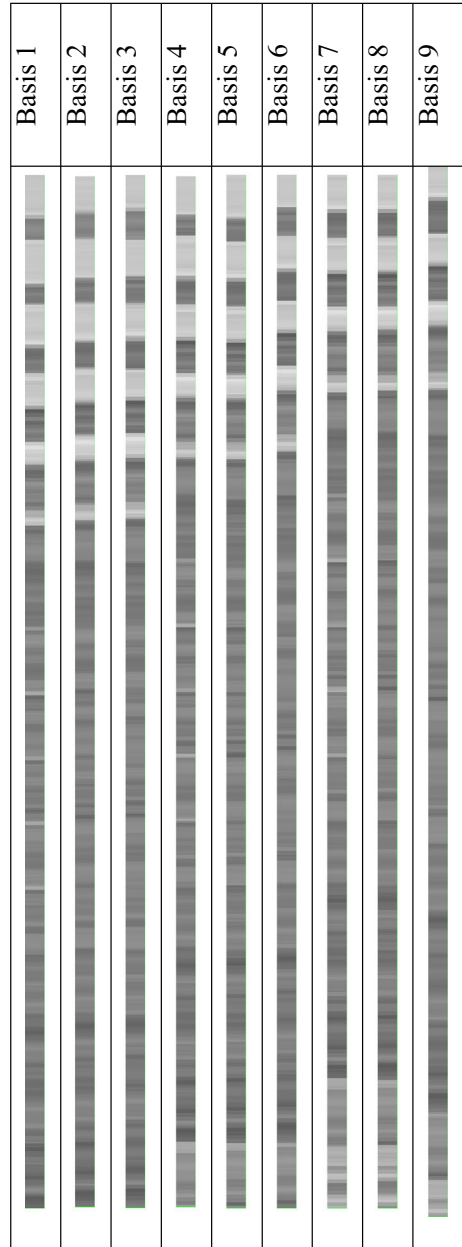
$$\hat{\mathbf{x}}_{t, \text{temporal}}^{\text{lostMB}} = \mathbf{X}_{t-1}^{\text{lostMB}} \boldsymbol{\beta}_{SO}^T \tag{11}$$

where $\hat{\mathbf{x}}_{t, \text{temporal}}^{\text{lostMB}}$ is a vector containing the pixels in the lost MB and recovered by the temporal method. $\mathbf{X}_{t-1}^{\text{lostMB}}$ is the same as in Eq (6).

3.3 Combined algorithm STSO

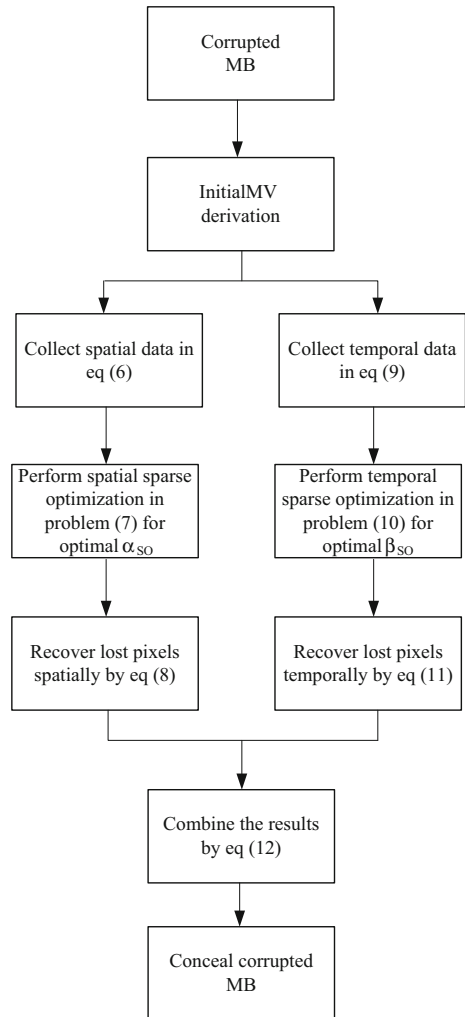
After spatial and temporal sparse optimization are done, we combine the results to obtain the final estimated lost pixels $\hat{\mathbf{x}}_t^{\text{lostMB}}$:

Fig. 4 The pixel illustration of the values in nine columns (each of size 256×1) in \mathbf{X}_{t-2} as nine bases



$$\hat{\mathbf{x}}_t^{lostMB} = \tau \times \hat{\mathbf{x}}_t^{lostMB, \text{spatial}} + (1-\tau) \times \hat{\mathbf{x}}_t^{lostMB, \text{temporal}} \tag{12}$$

The balance τ between the temporal solution and spatial solution is the same as in [19]. The whole STSO algorithm is shown in Fig. 5.

Fig. 5 Proposed STSO method for error concealment

4 Experimentresult

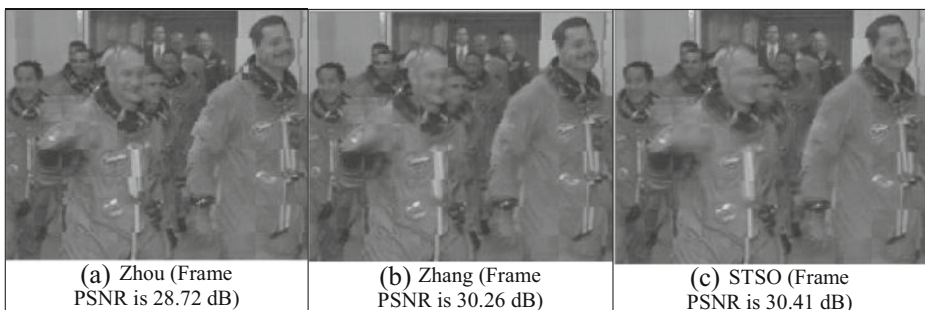
In this section, we discuss the experimental results that compare our method against existing ones. The video encoder is H.264 (JM 18.0). We use IPPPP (I means Intra frame, and P means Predictive frame) as an encoding GOP (Group Of Pictures) structure with a size of 15 frames. We use dispersed mode in FMO (Flexible Macroblock Ordering) so that the packets are encoded/lost in a checkerboard manner. We use various QPs (quantization parameters) from 16 to 38. The videos *crew*, *container*, *md*, *carphone*, *mobile*, *soccer*, *stefan*, *football*, and *foreman* are considered. The video resolution is 176×140 . We compare our proposed STSO method against the state-of-the-art spatial-temporal hybrid Zhang's method in [19] and Zhou's method in [21].

The PSNR (Peak Signal-to-Noise Ratio) is used as a metric to compare the performance. The PSNR is computed between the uncompressed video and the concealed video. Table 1

Table 1 PSNR (dB) performance comparisons for different QPs

Video	QP	STSO	Zhang [19]	Gain	Zhou [21]	Gain
Crew	16	30.41	30.26	0.15	30.03	0.47
	20	29.01	28.91	0.10	28.72	0.43
	24	28.99	28.92	0.07	28.82	0.39
Container	20	40.59	40.49	0.10	40.12	0.27
	22	39.78	39.73	0.05	39.35	0.19
	24	38.76	38.72	0.04	38.37	0.16
Md	24	37.15	36.97	0.18	36.88	0.62
	26	36.66	36.61	0.05	36.47	0.44
	28	35.37	35.30	0.07	35.21	0.39
Carphone	20	30.18	29.99	0.19	29.60	1.12
	24	29.79	29.62	0.17	29.38	0.98
	28	28.92	28.77	0.15	28.55	0.85
Mobile	22	33.78	33.68	0.10	32.55	0.81
	24	32.66	32.58	0.08	31.70	0.76
	28	31.79	31.62	0.17	30.97	0.63
Ssoccer	24	27.56	27.44	0.12	26.81	0.32
	26	26.79	26.71	0.08	26.13	0.17
	32	26.02	25.95	0.07	25.93	0.09
Sstefan	26	28.52	28.37	0.15	28.27	0.38
	28	27.67	27.58	0.09	27.50	0.29
	32	27.03	26.96	0.07	26.95	0.17
Football	32	27.63	27.51	0.12	27.11	0.52
	34	27.11	27.04	0.07	26.66	0.45
	38	26.32	26.30	0.02	25.91	0.41
Foreman	30	28.19	28.07	0.12	27.69	0.50
	34	27.84	27.74	0.10	27.47	0.37
	38	27.47	27.43	0.04	27.18	0.31

The gain is the improvement the proposed STSO method has over the compared methods

**Fig. 6** Error concealment performance comparisons for *crew*, 193th frame in QP=20

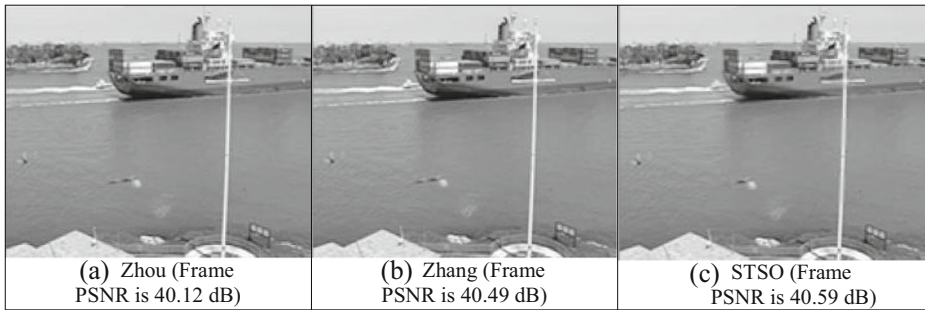


Fig. 7 Error concealment performance comparisons for *container*, 252th frame in QP=20

shows the PSNR comparison among the proposed STSO method, Zhang's method [19], and Zhou's method [21] for different video sequences and different QPs. Due to page limits, we choose representative cases for discussion. First, videos of slow motion (*crew*, *container*, *md*, *carphone*, and *mobile*) are discussed. And the videos of higher motion (*soccer*, *stefan*, *football*, and *foreman*) are discussed later.

For the slower videos, for *crew* video, it can be seen that the STSO performs better than Zhang by 0.15 dB, 0.10 dB, and 0.07 dB with different QPs. STSO works much better than Zhou by 0.47 dB, 0.43 dB, and 0.39 dB. These show outstanding performances of the STSO over the existing works. And the winning margin against Zhou is larger since Zhou does not utilize the information from the past two frames, as done in Zhang and STSO, so Zhou's performances are usually lower. For *container*, the trends remain the same; the STSO is better than Zhang by 0.1 dB, 0.05 dB, and 0.04 dB, and the winning margin is even larger against Zhou by 0.27 dB, 0.19 dB, and 0.16 dB in QP=20, 22, and 24, respectively. Similarly, in *md*, the performance gains of STSO are 0.18 dB, 0.05 dB, and 0.07 dB over Zhang, and they are 0.62 dB, 0.44 dB, and 0.39 dB over Zhou (larger gains as discussed) for different QP settings. For the *carphone* sequences, the improvements of 0.19 dB, 0.17 dB, and 0.15 dB are made over Zhang, and larger improvements of 1.12 dB, 0.98 dB, and 0.85 dB are made over Zhou. When comparing in *mobile*, the winning margin of 0.10 dB, 0.08 dB, and 0.17 dB are achieved against Zhang, and larger margins of 0.81 dB, 0.76 dB, and 0.63 dB are made against Zhou. Therefore, we can first conclude that the STSO is better than Zhang and Zhou for the videos with slower motion.

For the videos of higher motion, for *soccer*, STSO is better than Zhang by 0.12 dB, 0.08 dB, and 0.07 dB, and STSO is further better than Zhou by 0.32 dB, 0.17 dB, and



Fig. 8 Error concealment performance comparisons for *md*, 112th frame in QP=22

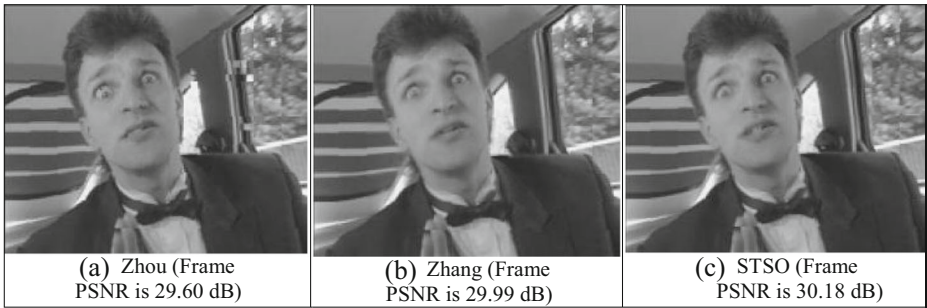


Fig. 9 Error concealment performance comparisons for *carphone*, 191th frame in QP=24

0.09 dB in QP settings being 24, 26, and 32, respectively. In *stefan*, similar situation occurs; by STSO, the improvement of 0.15 dB, 0.09 dB, and 0.07 dB are made over Zhang, and that of 0.38 dB, 0.29 dB, and 0.17 dB are achieved against Zhou in different QPs; again, Zhou has the lowest performances as discussed. In *football*, the performances of STSO are better than Zhang by 0.12 dB, 0.07 dB, and 0.02 dB, and they are also better than Zhou by an even larger margin 0.52 dB, 0.45 dB, and 0.41 dB for QP=32, 34, and 38. Finally, in *foreman*, for QP=30, 34, and 38, the winning margins of STSO over Zhang are 0.12 dB, 0.10 dB, and 0.04 dB, and they are 0.50 dB, 0.37 dB, and 0.31 dB over Zhou. From these discussions, it can be summarized that the STSO works better than the existing methods of Zhang and Zhou, even for fast videos.

To conclude from the data analyses and comparisons in Table 1, it can be understood that the proposed STSO performs better than the state-of-the-art Zhang and Zhou methods for various QP settings and different videos with different textures and motions. Furthermore, as discussed, the winning margin of STSO over Zhang is usually larger than Zhou because the method in Zhou only uses information from the previous frame and does not use information from the previous two frames, which is done by STSO and Zhang. Therefore, the performances of STSO and Zhang are better. Finally, for the largest gain among all the comparisons, STSO improves over Zhang's method by up to 0.19 dB and over Zhou's method by up to 1.12 dB, both for *carphone* sequence at QP=20. Note that both compared methods are state-of-the-art methods, which means that they have the most outstanding performance in the current technology; therefore, it is very difficult to outperform them by a very large margin.

In addition to the objective improvement in PSNR, our algorithm mostly performs better in subjective experience as well. Figure 6 to Fig. 14 show the different concealed frames using three different concealment methods. In Fig. 6, compared to Zhou's method, we can observe that the proposed method reduces the discontinuities in the concealed edges, and compared to

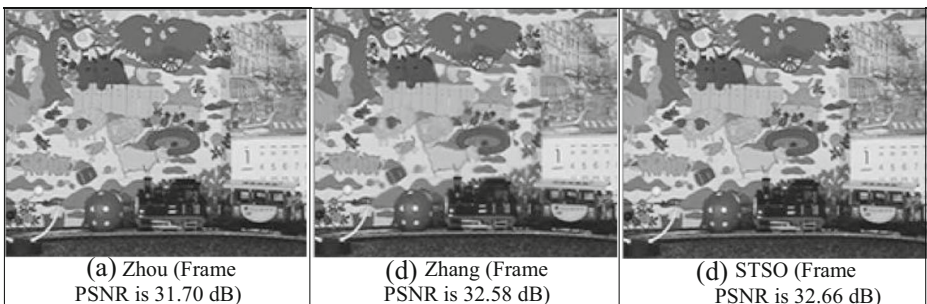


Fig. 10 Error concealment performance comparisons for *mobile*, 201th frame in QP=24

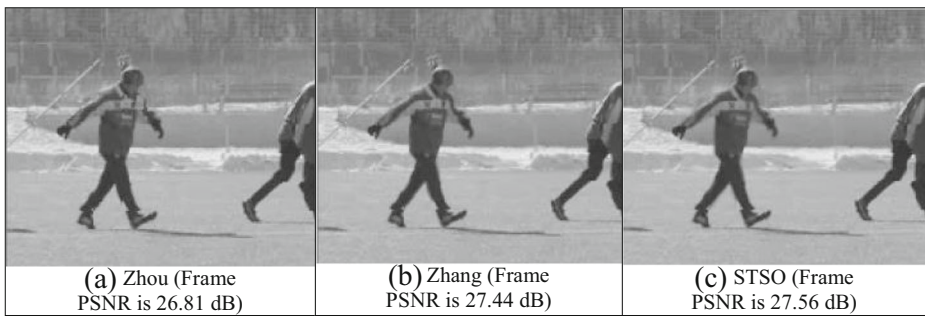


Fig. 11 Error concealment performance comparisons for *soccer*, 96th frame in QP=26

Zhang's method, the proposed algorithm performs better in the detail area; for example, the lines on the clothes. In Fig. 7, since the video is simpler and slower, the performance differences are not obvious. But as the bird flies fast over the frame, the proposed method produces less shadow objects. In Fig. 8, the results of different methods are similar. But for the detail continuity and integrity of the fingers, the proposed method performs better. In Fig. 9, the proposed method has better facial detail and smoother line connection in the background. In Fig. 10, for the faster part in the ball area, we have a more complete result. In Fig. 11, where the video is of higher motion, the proposed method conserves more integrity in the hand and leg area. Figure 12 is also a fast video. The proposed method produces better image integrity of the player and provides more complete English lines in the background. In Fig. 13, in this fast video, we provide a better image integrity in the fast moving player. In Fig. 14, we perform better in the edge areas, especially on the edge of the hat, edge of the face, as well as the detail of the face.

The improved results of the proposed method can be due to the fact that the sparse optimization can always find the solution with a minimal number of important and representative predictors. Therefore, by restricting the solutions to a limited number of important predictors and suppressing the effects of other unimportant predictors, our proposed error concealment algorithm can perform better in objective and subjective measurements.

Complexity analysis We discuss the complexities of the three discussed error concealment algorithms. For Zhou's method [21], it is less complicated since it only needs the generated extrapolated motion vector and the surrounding available motion vectors to perform the motion vector estimation for vertical and horizontal directions. The processing time with Zhou's method for each lost MB is 0.0064 s. For Zhang's method [19] and the proposed

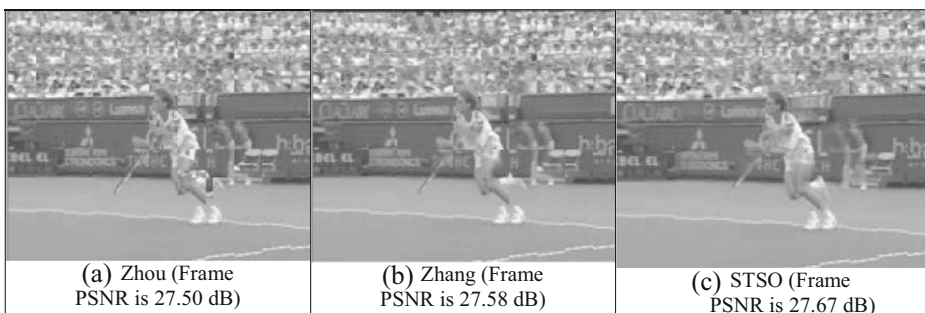


Fig. 12 Error concealment performance comparisons for *stefan*, 236th frame in QP=28

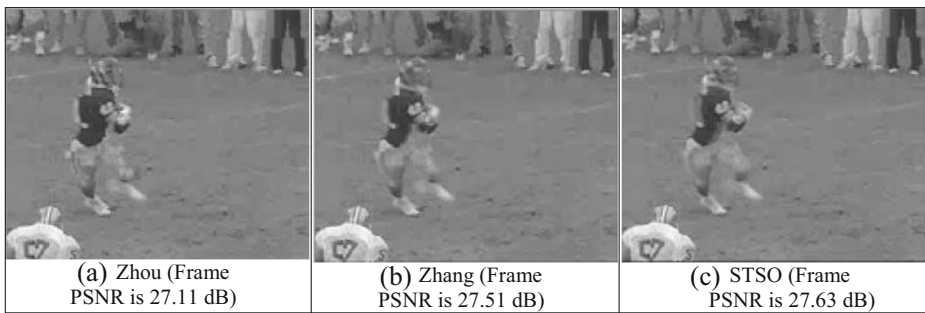


Fig. 13 Error concealment performance comparisons for *football*, 124th frame in QP=32

STSO method, the complexities are comparable since the data collection process and pixel arrangement are similar; for a lost MB, an initial motion vector is first generated by surrounding available motion vectors. With this vector, matrices in Eq (7) are formed, and α 's are solved by Zhang and STSO in different ways. Similarly, with the same initial motion vector and the motion vector in the previous frame, the matrices in Eq (10) are formed, and β 's are solved by Zhang and STSO in different ways. The usages of α and β are combined in Eq (12). The main computational complexities in Zhang and STSO are the optimization processes to find α and β ; thus, the computational time is higher than Zhou's method; for a lost MB, Zhang's method requires 0.038 s and STSO requires 0.043 s. Especially for the STSO, the sparse optimization technique is used; therefore, the processing time is slightly higher with the advantages of higher recovered image quality.

5 Conclusion

In this paper, we use the facts that the nature images has sparse representation. We proposed a spatial-temporal hybrid sparse optimization error concealment STSO method. Our algorithm limits the predictors to numerically useful ones so the performance of the error concealment scheme will not be distorted by other less-useful predictors. As shown in the experimental results, the proposed algorithm improves subjectively and objectively from the two state-of-the-art spatial-temporal hybrid algorithms by up to 0.19 dB and 1.12 dB. These are meaningful results considering these two state-of-the-art algorithms are ones of the most outstanding methods in the existing literatures.

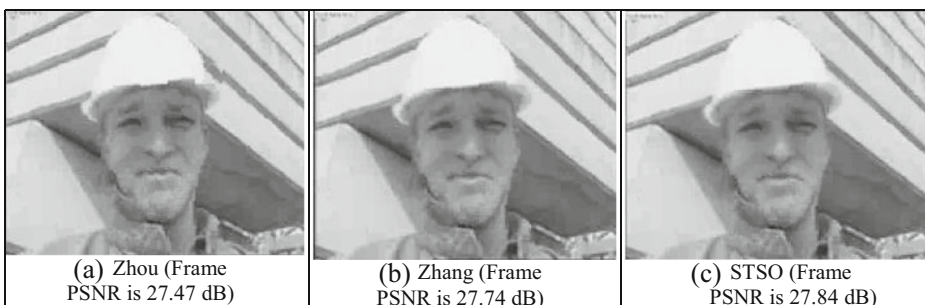


Fig. 14 Error concealment performance comparisons for *foreman*, 52th frame in QP=34

Acknowledgement This research is supported by the National Science Council, Taiwan under Grant NSC 100-2218-E-033-004, NSC 101-2221-E-033-036, NSC 102-2221-E-033-018, and by the Ministry of Science and Technology, Taiwan under Grant MOST 103-2221-E-033-020, MOST 104-2221-E-033-041 and MOST 104-2218-E-033-010.

References

1. Candès EJ, Romberg JK (2006) Quantitative robust uncertainty principles and optimally sparse decompositions. *Found Comput Math* 6:227–254
2. Candès EJ, Romberg JK, Tao T (2006) Stable signal recovery from incomplete and inaccurate. *IEEE Trans Inf Theory* 59:1207–1223
3. Candès EJ, Romberg JK, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52:489–509
4. Chen Y, Sun X, Wu F, Liu Z, Li S (2005) Spatio-temporal video error concealment using priority-ranked region-matching. *IEEE Int Conf Image Proc* 2:1050–1053
5. David L (2006) Donoho, “compressed sensing,”. *IEEE Trans Inf Theory* 52:1289–1306
6. Emmanuel J (2006) Candès and terence tao, “near optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans Inf Theory* 52:5406–5424
7. Haupt J, Nowak R (2006) Signal reconstruction from noisy random projections. *IEEE Trans Inf Theory* 52:4036–4048
8. L1 Magic, [Available] <http://users.ece.gatech.edu/~justin/l1magic/>
9. Lam W-M, Reibman AR, Liu B (1993) Recovery of lost or erroneously received motion vectors. *IEEE Int Conf Acoustics Speech Signal Proc* 3:417–420
10. Lie W-N, Lee C-M, Yeh C-H, Gao Z-W (2014) Motion vector recovery for video error concealment by using iterative dynamic-programming optimization. *IEEE Trans Multimedia* 16(1):216–227
11. Mairal J, Elad M, Sapiro G (2008) Sparse representation for color image restoration. *IEEE Trans Image Process* 17(1):53–69
12. Rane SD, Sapiro G, Bertalmio M (2003) Structure and texture filling in of missing image blocks in wireless transmission and compression. *IEEE Trans Image Process* 12(3):296–303
13. Salama P, Shroff NB, Delp EJ (2000) Error concealment in MPEG video streams over ATM networks. *IEEE J Sel Areas Commun* 18(6):1129–1144
14. Tropp JA, Gilbert AC (2007) Signal recovery from partial information via orthogonal matching pursuit. *IEEE Trans Inf Theory* 53:4655–4666
15. Wang Y, Zhu Q-F, Shaw L (1993) Maximally smooth image recovery in transform coding. *IEEE Trans Commun* 41(10):1544–1551
16. Wiegand T, Sullivan GJ, Bjontegaard G, Luthra A (2003) Overview of the H.264/AVC video coding standard. *IEEE Trans Circuits Syst Video Technol* 13(7):560–576
17. Yan B, Zhou J (2012) Efficient frame concealment for depth image-based 3-D video transmission. *IEEE Trans Multimedia* 14:936–941
18. Yen-Chi Lee, Yucl Altunbasak, and Russell M. Mersereau (2002) “Multiframe error concealment for MPEG-coded video delivery over error-prone networks,” *IEEE Transactions on Image Processing*, vol. 11, no. 11
19. Yongbing Zhang, Xinguang Xiang, Debin Zhao, Siwe Ma and Wen Gao, (2012) “Packet video error concealment with auto regress model,” *IEEE Transaction on Circuit and System for Video Technology*, vol.22, no.1
20. Zhou Y, Xiang W, Wang G (2015) Frame loss concealment for multiview video transmission over wireless multimedia sensor networks. *IEEE Sensors J* 15(3):1892–1901
21. Zhou J, Yan B, Gharavi H (2011) Efficient motion vector interpolation for error concealment of H.264/AVC. *IEEE Trans Broadcast* 57:75–80
22. Zhu W, Wang Y, Zhu Q-F (1998) Second-order derivative-based smoothness measure for error concealment in DCT-based codecs. *IEEE Trans Circuits Syst Video Technol* 8(6):713–718



Ting-Lan Lin received the B.S. and M.S. degrees in electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, in 2001 and 2003, respectively, and the Doctorate degree in electrical and computer engineering from the University of California, San Diego, La Jolla, CA, USA, in 2010. In 2008, he interned in the Display System group at Qualcomm, San Diego, CA, USA. Since 2011, he was an Assistant Professor with the Department of Electronic Engineering, Chung Yuan Christian University, and since August 2015, he has been an Associate Professor with the Department of Electronic Engineering, Chung Yuan Christian University in Taiwan. His current research interests include video compression, video streaming in lossy networks, optimization of packet prioritization, and perceptual video quality.



Tsai-Ling Ding received the B.S. degree in the department of electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan in 2014. He is now pursuing M.S. degree in the department of electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan. His research interests are error concealment in H.264.



Chang-Yi Fan received the B.S. degree in the department of electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan in 2011. He received the M.S. degree in electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, in July 2013. His research interests include error concealment for lossy videos due to packet losses in transmission.



Wen-Chih Chen received the B.S. degree in electronic engineering from the National Changhua University of Education, Changhua, Taiwan. He received the M.S. degree in electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, in July 2013. His research interests include error concealment for lossy videos due to packet losses in transmission. Several related papers have been published. Now he is working at semiconductor industry.