Signal Processing: *Image Communication* 29 (2014) 1121–1137

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Hybrid error concealment method combining exemplar-based image inpainting and spatial interpolation

Byungjin Chung, Changhoon Yim*

Department of Internet and Multimedia Engineering, Konkuk University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 23 January 2014
Received in revised form
12 September 2014
Accepted 14 September 2014
Available online 2 October 2014

Keywords:
Image reconstruction
Error concealment
Image inpainting
Spatial interpolation
Video communication

ABSTRACT

This paper proposes an efficient error concealment method for the reconstruction of pixels that are lost in video communication. The proposed method is developed by combining exemplar-based image inpainting for patch reconstruction and spatial interpolation for pixel reconstruction using adaptive threshold by local complexity. By exemplar-based image inpainting, regions with regular structures are reconstructed. For complex regions with irregular structures, just one pixel is reconstructed using the proposed spatial interpolation method. The proposed spatial interpolation method performs reconstruction by selecting adaptively directional interpolation or neighbor interpolation based on gradient information. Simulation results show that the proposed hybrid method performs reconstruction with significantly improved subjective quality compared with the previous spatial error concealment and image inpainting methods. The proposed method also gives substantial improvements of PSNR compared with the previous methods.

1. Introduction

The applications of video communications are expanding because the transmission speeds of networks become faster and smart mobile devices become more popular. These applications use video coding standards such as MPEG-2 [1], H.264 [2], and high efficiency video coding (HEVC) [3] that were established by ITU-T and ISO/IEC for compression. Compressed video data may not be transmitted in real-time because of queuing and transmission delays when bit-stream packets are transmitted through Internet and wireless networks. When the variation of delay is severe, the video stream would be transmitted with bit errors or packet losses. When the bit error rate (BER) or packet error rate (PER) become larger than some level, it could result in devastating effect on video quality.

To overcome these problems, several techniques have been developed including forward error correction (FEC), automatic retransmission request (ARQ), and the error concealment (EC) [4]. The FEC requires additional bandwidth and ARQ is dependent on the network round trip time and is hard to guarantee real-time transmission. On the other hand, EC could be performed in decoder side and is not dependent on network condition. The EC would be effective to conceal the lost information from packet losses [4,5].

The EC schemes can be categorized as spatial error concealment (SEC) and temporal error concealment (TEC) [6]. The SEC schemes reconstruct the lost macroblock (MB) by spatial interpolation using pixel information of neighboring macroblocks with high spatial correlation. The SEC would give effective reconstruction if the size of the lost area is small. But it can generate blocking artifacts if the size of the lost area is large since the correlations from neighboring pixels become smaller. It is hard to give natural reconstruction as the lost area becomes larger

* Corresponding author. Tel.: +82 2 450 4016; fax: +82 2 458 1997.
E-mail address: cyim@konkuk.ac.kr (C. Yim).

http://dx.doi.org/10.1016/j.image.2014.09.009
0923-5965/

because it performs the reconstruction using only the neighboring MBs that are received correctly. The TEC reconstructs the lost MB as the MB of the same location in the previous frame if the average motion vector of the slice is smaller than some threshold. The TEC performs the reconstruction using boundary matching algorithm (BMA) if the average motion vector is larger than the threshold [6]. The TEC might be dependent on the performance of SEC in the previous frame.

This paper deals with the improvement of the reconstructed video quality by SEC. It adopts image inpainting method to overcome the disadvantage of SEC. Many image inpainting methods are developed to edit or reconstruct damaged regions of digitized images and videos [7–21]. Image inpainting is also applied for image and video compression [22,23].

Diffusion-based image inpainting method reconstructs the scratched area by the extension of texture structure from the boundary region using partial differential equations as in classical fluid dynamics [10,11]. This method is effective to reconstruct a small region but it might show blur and spreading effect when the region is large [14]. Fast marching image inpainting method reconstructs pixels from the boundary region to the internal region [12]. The pixel for reconstruction is replaced by normalized weighted sum of already calculated pixels in one pass.

Exemplar-based image inpainting method [8,9] is developed using the technique of texture synthesis. This method conducts the reconstruction according to the pixel structure having the continuity of edges in isophote direction in patch area. PatchMatch image inpainting method uses the matching (most similar) patch in other parts of image for editing or reconstruction [13]. It searches the matching patch randomly in nearest neighbor field (NNF) using coarse-to-fine approach. Block-based image inpainting method is used to recover lost macroblocks for video error concealment applications [15,16]. It uses Sobel operator to determine the priority of pixels by edge information. Reconstruction is performed in block-based patch using the priority information. Hierarchical super-resolution-based inpainting first performs reconstruction at a coarse resolution, and then high-frequency details are reconstructed on a fine resolution [17]. This method shows impressive performance for image editing with large unknown areas and texture synthesis.

Hybrid method is developed by decomposing image components into structure and texture [24]. After the decomposition, structure image is reconstructed by diffusion-based inpainting, and texture image is reconstructed by exemplar-based texture synthesis. The resulting two images are added to obtain the reconstructed image. Another hybrid method combines different approaches in one unified energy function [25]. The energy function combines self-similarity, diffusion, and coherence terms. The reconstructed image is obtained by minimizing the energy function. Hybrid sparse representation is developed in [26]. Local and non-local sparse representations are combined as Bayesian model averaging. The reconstructed image is obtained by deterministic annealing with iterations. Since these hybrid schemes would require larger amount of computations compared with other image inpainting methods. Hence it would be difficult to be applied for error concealment applications with time constraints.

The existing SEC scheme for H.264 [6] does not consider the structure of neighboring pixel area. Hence internal structure was lost, and the reconstruction looks unnatural. The exemplar-based image inpainting performs the reconstruction in patch area and is effective to reconstruct the structure of patch area with regular structure. However, if there is no similar patch in the image, the selected patch might give noise phenomenon and would not give affordable reconstruction quality.

In this paper, we propose a hybrid error concealment method combining exemplar-based image inpainting and spatial interpolation. When there exists a reliable patch, it performs the reconstruction for the patch by exemplar-based image inpainting. Otherwise, it performs the reconstruction just for the pixel by spatial interpolation. One unique feature of the proposed method is the selective reconstruction by patch (inpainting) or by pixel (interpolation). The simulation results show that the proposed method reduces the blocking artifact, spreading effect, and noise phenomenon significantly and results in natural reconstruction for a variety of images. The proposed method requires less amount of computations compared with previous image inpainting schemes for error concealment, and it can be adopted for applications with time constraints such as video streaming.

The organization of this paper is as follows. Section 2 presents previous related works. In Section 3, we propose the hybrid error concealment method combining exemplar-based image inpainting and spatial interpolation. Section 4 presents simulation results for the proposed method compared with the previous spatial error concealment and image inpainting schemes. In Section 5, we present conclusion and future work.

2. Previous related works

2.1. Exemplar-based image inpainting

The exemplar-based image inpainting conducts the reconstruction process from the boundary region to the internal region gradually until the target (lost) region is fully recovered. The reconstruction order in the target region is determined by the priority of pixels. Once the pixel with the highest priority is determined, the search is performed for the patch which gives minimal sum of squared differences (SSD) inside the source (correctly received) region.

The basic idea of exemplar-based image inpainting is depicted in Fig. 1. Fig. 1(a) shows the target (lost) region Ω, the boundary region $\delta\Omega$ in Ω, and the source (correctly received) region Φ in image $\mathcal{I}$ for inpainting. In Fig. 1(b), the p is the pixel with the highest priority in $\delta\Omega$ and the Ψ_p is the patch including p in center position. The search to find the best matching patch for Ψ_p is performed in source region Φ. The SSD distortion is calculated for the partial patch $\Psi_{\hat{p}}$, where $\Psi_{\hat{p}} = \Psi_p \cap \Phi$ and $\Psi_{\tilde{p}} = \Psi_p \cap \Omega$. In Fig. 1(c), the reconstruction is performed for region $\Psi_{\tilde{p}}$ by pasting the region $\Psi_{\tilde{q}}$, where Ψ_q is the best matching patch and $\Psi_{\tilde{q}}$ is the region in Ψ_q with the same shape as $\Psi_{\tilde{p}}$. This process is performed repeatedly until all the pixels in Ω are filled (reconstructed). The exemplar-based