International Journal of Innovative Computing, Information and Control
Volume **6**, Number **3(B)**, March **2010**

ICIC International ©2010 ISSN 1349-4198
pp. **1241–1255**

OBJECT REMOVAL AND INPAINTING IN MULTI-VIEW VIDEO SEQUENCES

Soon-Young Lee[1], Jun-Hee Heu[1], Chang-Su Kim[2] and Sang-Uk Lee[1]

[1]Signal Processing Lab, INMAC
School of Electrical Engineering and Computer Science
Seoul National University
Seoul, Korea
sylee@ipl.snu.ac.kr; junhee@snu.ac.kr; sanguk@ipl.snu.ac.kr

[2]School of Electrical Engineering
Korea University
Seoul, Korea
changsukim@korea.ac.kr

Received December 2008; revised May 2009

Abstract. An inpainting algorithm for multi-view video sequences, which fills in a mask area without visual artifacts, is proposed in this work. After a user specifies a mask area in a frame, the proposed algorithm clones it to the other frames in the input sequence using temporal and inter-view correlations. The proposed algorithm then selects sites along the mask boundary, over which source blocks are pasted. We develop a cost function, which measures the spatio-temporal consistency of an inpainted result, and minimize the cost function to select the best matching source blocks. Simulation results demonstrate that the proposed algorithm removes and inpaints objects within mask areas efficiently without causing blocking or flickering artifacts.
Keywords: Image inpainting, Video editing, Object removal, Multi-view video

1. **Introduction.** The objective of video inpainting is to fill in a specific area in a video sequence, which a user wants to remove or replace with another object. Since video inpainting can modify parts of objects or backgrounds directly, it has been used in the post-film processing in which undesired objects are removed from films without re-recording. However, video inpainting is a fundamentally ill-posed problem, and conventional video inpainting algorithms require a lot of manual tasks and thus have been applied in limited circumstances only.

Recent advances in video technology facilitate interactive broadcasting and enable us to experience virtual reality. Especially, multi-view video can offer 3-D realism and interaction, since it provides different views of the same scene [1]. However, whereas various algorithms have been proposed for the inpainting of single-view video sequences, there are few inpainting techniques for multi-view video sequences. A multi-view video sequence consists of not only temporally sampled frames but also spatially correlated frames from different viewpoints at the same time. Thus, a multi-view video sequence often requires a huge amount of data, which makes multi-view video inpainting a challenging task. For example, in the pre-processing stage of multi-view video inpainting, more manual interactions are required to specify mask areas to be inpainted and distinguish objects from the background. Moreover, if temporal and inter-view coherence is not taken into account carefully during the inpainting, the resultant sequence may contain visually annoying artifacts.