J Real-Time Image Proc
DOI 10.1007/s11554-013-0375-9

SPECIAL ISSUE PAPER

AVScreen: a real-time video augmentation method

Francisco J. Hernandez–Lopez · Mariano Rivera

Received: 14 May 2013 / Accepted: 29 August 2013

Abstract We present a tool for video augmentation in real-time, which we name the augmentation virtual screen (AVScreen). AVScreen is useful for developing advertisements, commercials, music videos, movies, etc. The main challenges for augmenting videos, in contrast to fixed images, is that moving objects in the foreground may occlude the region to be augmented in the background and that the composition can be affected by camera movements. Therefore, we use a procedure for foreground–background video segmentation in order to deal with such occlusions. Comparisons with foreground–background video segmentation methods of the state of the art in both accuracy and computational efficiency support our choice: we reduce around 70 % of the segmentation error in a popular benchmark database and achieve real-time performance. Moreover, a new stabilization method to augment unstable camera videos is presented. For augmenting video shots, we present an efficient graph-based method for panorama (mosaic) computation. The real-time performance is reached by implementing high computational demanding procedures in GPU. The frame rate of our method is 18 frames per second for a video size of 640 × 480 pixels.

Keywords Video augmentation · Video segmentation · Video stabilization · Mosaic · CUDA GPU

1 Introduction

The large number of available images and videos has motivated new applications for processing them and changing their content. One interesting task is to augment in real-time a region of a given video with an image or another video. Several research projects have focused on creating new augmented video applications for a variety of uses. Relevant works in the literature are the following. The method in [4] generates augmented video sequences by compositing layers such as a natural or synthetic background (BG) and several natural or synthetic foregrounds. A content-based representation of a video shot composed of a background mosaic and one or more foreground (FG) moving objects is proposed in [3]. In [8], a procedure for re-displaying video sequences by giving the user control over the camera's motion and field of view is presented. A method for increasing the complexity of videos in terms of occlusion and crowds in a scalable and controllable manner using virtual agents is reported in [5]. In [2] a technique to fuse in a single scene the FG of different videos, the corresponding motion parts and dynamic BG is proposed.

The core problems in those approaches are the video stabilization, the video segmentation and the mosaic computation. The complexity of the solution depends on video acquisition assumptions, some of these are:

- The camera is static.
- The video is recorded from a single location.

Electronic supplementary material The online version of this article (doi:10.1007/s11554-013-0375-9) contains supplementary material, which is available to authorized users.

F. J. Hernandez–Lopez (✉) · M. Rivera
Centro de Investigacion en Matematicas AC, Guanajuato, Mexico
e-mail: fcoj23@cimat.mx

M. Rivera
e-mail: mrivera@cimat.mx

Published online: 27 September 2013