Overview

We present a tool for image editing that allows us to modify a region of any image or video by another image or video. This application is useful for advertisements, commercials, music videos, movies, etc.

The main difference between editing (augmenting) videos and fixed images is that the occlusions need be managed. Moving objects in foreground may occlude the augmented region in background. So that we use a procedure for Foreground/Background (FgBg) video segmentation, which is implemented in NVIDIA video cards to fulfill the real-time requirement.

Proposal

The general outline of VScreen is shown in Figure 1. In the illustrated case, we want that the video g in (b) appears in the back wall of the video f in (a). The composition is computed using the following steps:

1. **User Interaction I.** First we create a Trimap (c), by marking the interesting region in the frame f where we want put the video g (green scribble), and with another color the uninteresting region (red scribble).
2. **User Interaction II.** In the frames f and g we select 4 points (e) and (f) such that the polygon formed in the frame g is going to be the region that appears in the region selected in the frame f.
3. **Automatic process I.** With the Trimap, we use the QMMF algorithm [1] to segment the interesting region obtaining α1 (d).
4. **Automatic process II.** With the 4 points selected in the frames (e) and (f), we compute the homography H which transforms the points from video g to f. In (g) we apply the transformation H to the video g.
5. **Automatic process III.** We apply a FgBg video segmentation algorithm [2] to the video f, obtaining α2 (i).
6. **Automatic process IV.** Given the segmentation with Trimap α1, the transformed frame Hg and the FgBg video segmentation α2, we apply the composition:

   \[ f_c = α_1(1 - α_2)Hg + (1 - α_1 + α_2)f \quad (1) \]

Note that the selected region in the video f might be occluded by an object or objects in the foreground of the same video (the people walking). Therefore, we apply the FgBg video segmentation method, which is robust to illumination changes, shadows and camouflage situations, in order to segment the people and to obtain a realistic composition. Figure 2 shows some results of VScreen in different indoor and outdoor scenes.

Experiments and Results

VScreen has been implemented using Visual C++ 2008, CUDA and OpenCV libraries. The experiments are run in a PC with Core 2 Quad CPU 2.39 GHz, 2 GB RAM, and Windows XP Operating System (32 bits) using a single core.

The most computationally expensive stage, of Virtual Screen, is the FgBg video segmentation, however this segmentation is only computed in the region to be augmented. So that the computational cost depends on the size of the region to be augmented.

Table 1 shows the computational rate, in frames per second (f/s), of VScreen using different NVIDIA video cards. Those rates were computed by assuming that the augmented region is the whole frame in the background video.

A quantitative evaluation is shown in Table 2. We compare our Fg/Bg video segmentation with a method of the state of the art [3].

Conclusions

We propose VScreen, an interactive and automatic tool for video composition in real-time. The user just selects the region in the original video to be augmented by another new video or image.

VScreen manages the occlusions with the Fg/Bg video segmentation [2] implemented in CUDA in order to compute the composition in real-time.

Experiments and quantitative evaluation demonstrate that VScreen achieves good results and is robust to different scenarios (indoor/outdoor).

For watching the videos, you can visit the website (http://www.cimat.mx/~rivera/screen), where is also possible to download videos (original and modified) with high resolution (1280 x 720).

References

