

VScreen: a real-time augmented video method

Francisco J. Hernandez-Lopez and Mariano Rivera

Centro de Investigación en Matemáticas A.C.
{fcoj23, mrivera}@cimat.mx

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:

- 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).
- 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 .
- Automatic process I.** With the Trimap, we use the QMMF algorithm [1] to segment the interesting region obtaining α_1 (d).
- 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 .
- Automatic process III.** We apply a FgBg video segmentation algorithm [2] to the video f , obtaining α_2 (i).
- Automatic process IV.** Given the segmentation with Trimap α_1 , the transformed frame Hg and the FgBg video segmentation α_2 , we apply the composition:

$$f_c = \alpha_1(1 - \alpha_2)Hg + (1 - \alpha_1 + \alpha_1\alpha_2)f \quad (1)$$

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.

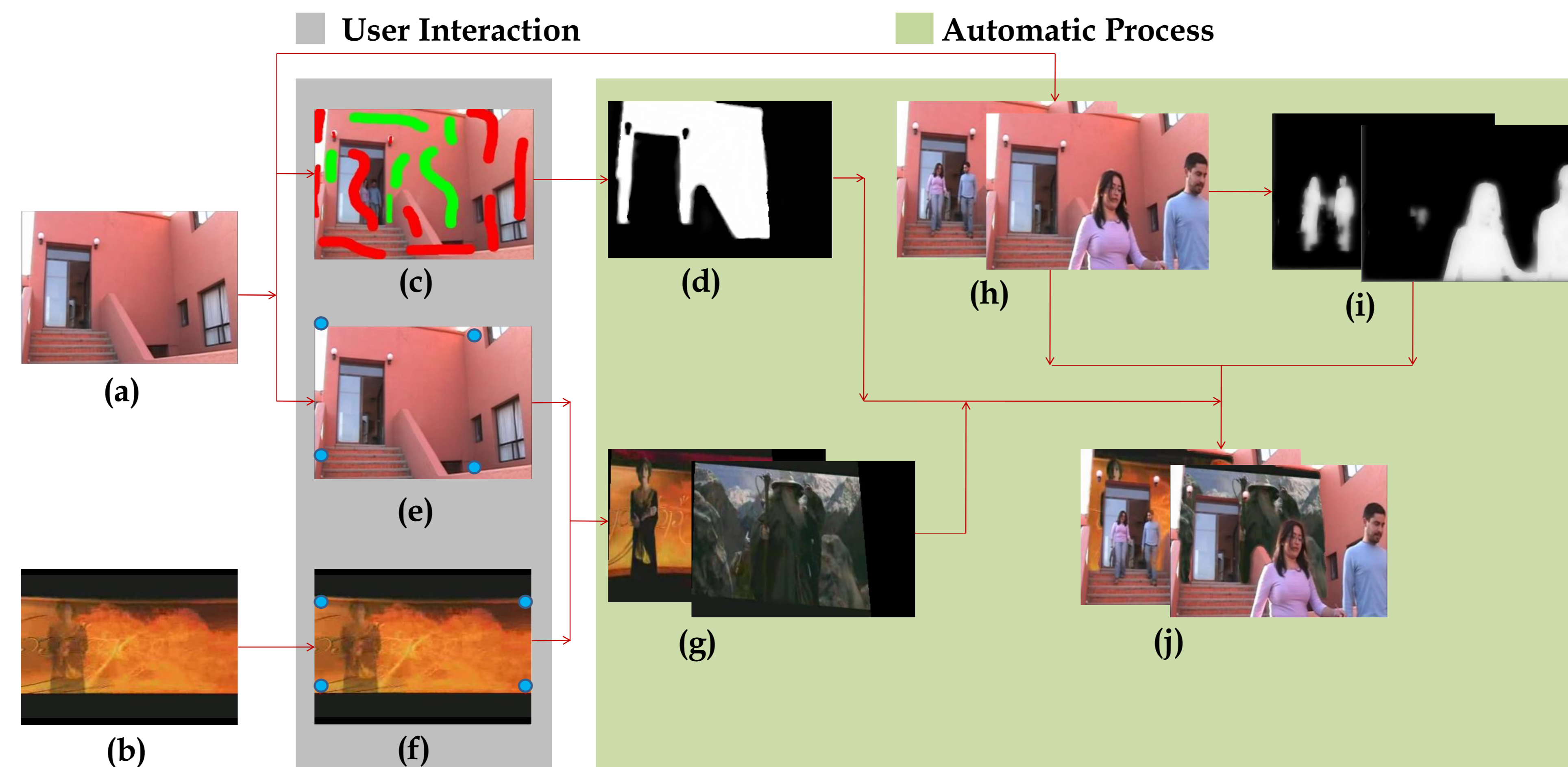


Figure 1. General outline of VScreen. (a), (b) Original videos. (c) Trimap. (d) Applying the QMMF method with Trimap α_1 . (e), (f) Selecting four points for calculating the homography H . (g) Applying the homography to video (b). (h), (i) Applying the FgBg video segmentation α_2 . (j) Composed video.

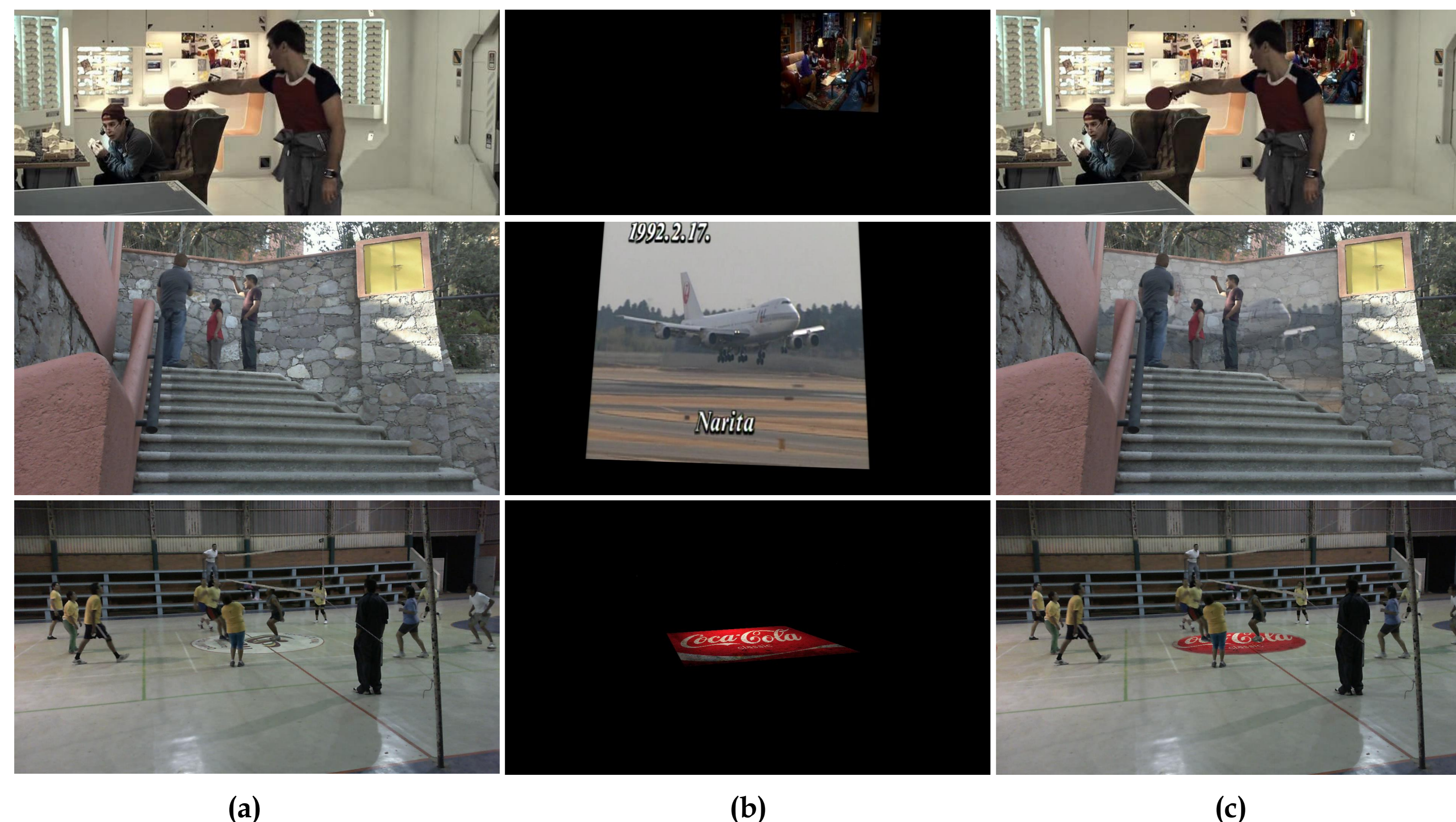


Figure 2. Results of VScreen. (a) Original frame f . (b) Applying the homography to new frame Hg . (c) Composed frame f_c .

Image Size	(f/s)			Total
	GeForce 8800 GT	Tesla C1060	GeForce GTX 480	
320x240	48	63	65	
640x480	12	18	20	

Table 1. Rate (f/s) of VScreen for different NVIDIA video cards assuming the full frame substitution.

Sequence	Method	F. Neg.	F. Pos.	Total
Intelligentroom	[3]	159.1	74.1	233.2
	VScreen	58.0	45.4	103.4
Video2_long	[3]	302.5	59.3	361.8
	VScreen	138.7	18.2	156.9
Video5_long	[3]	1179.3	275.9	1455.2
	VScreen	346.1	24.0	370.1
Video6_long	[3]	1195.0	310.6	1505.6
	VScreen	402.3	46.9	449.2

Table 2. Quantitative evaluation.

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/~mrivera/vscreen>, where is it also possible to download videos (original and modified) with high resolution (1280 x720).

References

- [1] M. Rivera and P. P. Mayorga. Quadratic Markovian probability fields for image binary segmentation. In Proc. ICCV (WICV), pages 1-8. IEEE Computer Society, 2007.
- [2] F.J. Hernandez-Lopez and M.Rivera, "Binary segmentation of video sequences in real time," Proc. MICAI, 163-168, IEEE PRESS, 2010.
- [3] A. García and J. Bescós. Video object segmentation based on feedback schemes guided by a low-level scene ontology. In Proc. ACIVS '08, pages 322-333, 2008.