Proceedings of the 8th European Conference on Visual Media Production, London, 16-17 November 2011, paper number sp-15

VIDEO RE-SAMPLING AND CONTENT RE-TARGETING FOR REALISTIC DRIVING INCIDENT SIMULATION

Alejo M. Heras, Toby P. Breckon, Marko Tirovic
School of Engineering, Cranfield University, UK.

Abstract
We present a video generation and visualization system for integration into a driving incident simulator. We combine the use of video re-sampling, target tracking and object segmentation for content re-targeting to generate realistic video content that can be adapted on demand, based upon simulator inputs to generate a range of driver incident test scenarios.

Keywords: Motion Frame Rate Up-Converters (MFRUC), frame interpolation, object segmentation, feature tracking.

1 Introduction

Driving simulation systems for measuring vehicle driver responses to potential incidents are widely used for performance testing and evaluation with risk limitation. In these systems the perceived realism of the presented simulation is critical to the overall quality and reliability of the results [1].

Traditional approaches for implementing such a system use virtual reality like 3D generated animation models which have a limited level of realism. By contrast in this work to reach the goal of a realistic simulation experience we combine the use of temporal video re-sampling [2], to compensate for varying vehicle speed upon driver demand, and additionally the extraction and re-targeting of *pre-caned* incident objects using a combination of segmentation and feature tracking techniques [3,4].

2 Video Re-sampling

As the original sampled video footage was recorded at a constant frame rate and vehicle speed we require to temporally vary the video sampling in order to maintain a frame rate of 20 fps, within the simulation, in response to varying acceleration and speed demands from the driver. This is crucial in maintaining the perceptual realism of the video based simulation.

The Motion Frame Rate Up-Converter approach of [2] is used to generate the required intermediate frames. This approach extracts the motion information present between each two consecutive frames contained in the original video sequence and performs bidirectional block matching to construct the new intermediate frame samples.

3 Object Extraction

A series of incident objects (e.g. pedestrians/vehicles) are recorded over a set of source videos against a plain background to ease semi-automatic extraction. These objects are automatically extracted using Grabcut segmentation approach [3]. Each object is tracked in each frame using background differencing. Subsequently, the necessary set of background/foreground seed pixels required for grabcut object extraction are automatically set, from the extracted foreground region, and the extraction is finally performed. The sequences of extracted objects are subsequently stored for later insertion, on-demand, into the final composite video output.

4 Object Insertion

Each frame of the up-sampled video sequence is analysed for the presence of a given known marker which indicates the horizontal path where an incident object is to be inserted within the simulator video sequence.

A feature tracking technique, making use of SURF key point matching [4], is used to identify the signature of a given target location present within the video.

Where a target location for insertion is detected, a randomly selected incident object appropriate to that target locale is inserted, re-dimensioned and blended to appear naturally integrated into the rest of the video frame (Figure 1). Re-dimensioning is carried out using a linear re-sizing method. Smoothing is performed on the boundaries of the inserted object and the original scene using mean filtering.

Figure 1: An example incident object (left), identified target position plane (middle), final composite output (right).

5 Results

We present an example sequence in which a pedestrian incident object is dynamically inserted, crossing the road in front of the vehicle path, in an up-sampled video sequence (Figure 2). Note how the pedestrian is dynamically resized upon vehicle approach (in addition to dynamic temporal video re-sampling).

Figure 2: Example pedestrian incident object insertion sequence.

6 Conclusions

Overall realistic video content generation, using a combination of temporal video re-sampling and content re-targeting, is viable for the production of a composite output for driver simulator usage.

Future work could investigate the use of inter object occlusion, object re-lighting and shadow formation within this video compositing task.

[1] Riener, A. (2010), “Simulating On-the-Road Behavior Using a Driving Simulator”, *International Conference on Advances in Computer-Human Interaction*, pp. 25-31.

[2] Yang, Y., Tung, Y. and Wu, J. (2007), "Quality Enhancement of Frame Rate Up-Converted Video by Adaptive Frame Skip and Reliable Motion Extraction", *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 12, pp. 1700-1713.

[3] Rother, C., Kolmogorov, V. and Blake, A. (2004), ""GrabCut" - Interactive foreground extraction using iterated graph cuts", *ACM Transactions on Graphics*, Vol. 23, pp. 309-314.

[4] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008), "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359.