Synthesis of the Spatio-temporally Coherent Line Drawing Animation from Video

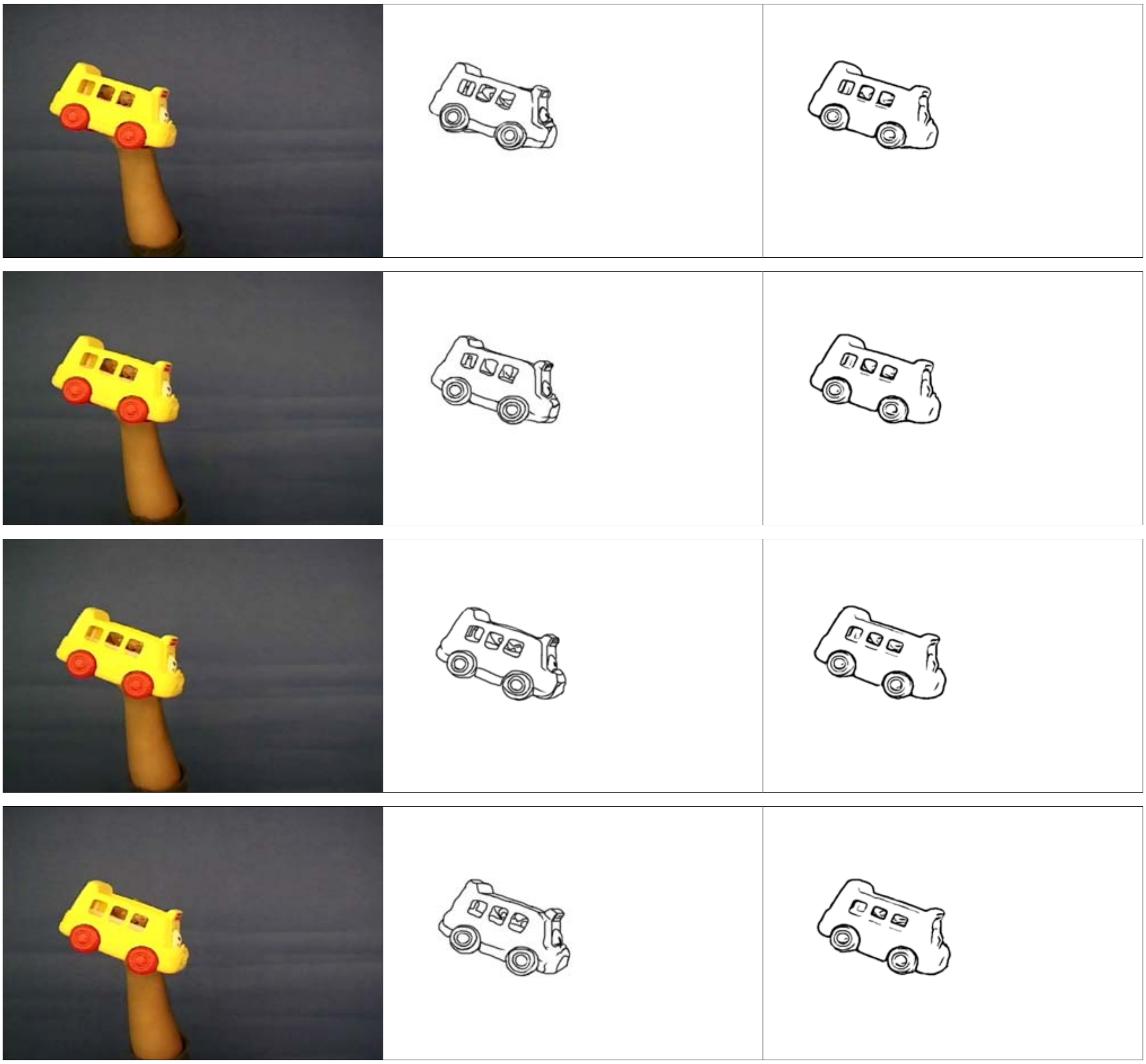

Figure 1: Input video frames (left column), manually synthesized line drawing frames (middle column) and line drawing frames using our method (right column). Note that only the region for the toy bus is considered to be semantically meaningful.

1. Introduction

This paper proposes an non-photorealistic rendering (NPR) method to create a line drawing animation from a real video. NPR is a rendering technique which generates artistic expression or abstraction, such as drawing, painting, cartoons, and technical illustrations. Within NPR, processing a video recorded real scene in artistic styles for use in entertainment industry is highly on demanding. In particular, we focus on stylizing video by line drawing, since line drawing effectively conveys object shapes to the viewer while using minimal amount of data (lines).

To generate a satisfying line drawing animation from real video, temporal coherence between consecutive video frames much be considered. The temporal coherence means that to keep aesthetic completeness of the animation, pixels consisting object shapes should not move too far over time, and new pixels should not suddenly appear in an empty region of the subsequent frame. Without securing the temporal coherence, flickering and popping artifacts may be introduced in the resulting animation.

Figure 2: Input image(left) and the result of [Kang et al. 2007] (right)

Coherent Line Drawing [Kang et al. 2007] (Fig. 2), which is one of the sketchy image stylization methods, extracts a set of spatially coherent, smooth and stylistic lines that well capture important object shapes from a real still image. Our algorithm extends the work to extract such lines from video with temporal coherence.

2. Related Work

Algorithms to render a real still image with sketchy image stylization that have been proposed for years [Salisbury et al. 1997; Kang et al. 2007] cannot be directly applied to individual video frames to achieve a stylized animation from input video due to temporal incoherence. [Litwinowicz et al. 1997; Hertzmann and Perlin 2000] reduce such temporal incoherence by translating strokes from frame to frame in the direction of pixel motion, accounted by an estimated optical flow vector field [Horn and Schunck 1981]. Unfortunately, these methods in practice result in unsatisfactory animations since errors present in the motion estimation quickly accumulate and propagate to subsequent frames. We minimize the use of the optical flow for small bounded regions of frames while easing the error propagation.

[Wang et al. 2004] identifies the input video as spatio-temporal 3D volumes of image data. The user rotoscopes semantically meaningful regions on input video frames and stylizes the regions by fitting the stylization through the semantic surfaces of the volumes. However, since this stylization is a way of vector graphics, we cannot directly employ [Wang et al. 2004] with [Kang et al. 2007] that produces raster output. We modify [Kang et al. 2007] to combine [Wang et al. 2004] and [Kang et al. 2007], and also to preserve temporal coherence.

3. Overview

Our goal is to generate an aesthetic and qualititive line drawing animation from a real video while ensuring spatio-temporal coherence. We adopt *Coherent Line Drawing* [Kang et al. 2007] which well establishes spatial coherence for each video frame, and modify the algorithm both to achieve temporal coherence and to be applicable to [Wang et al. 2004].

Coherent Line Drawing is a flow-based anisotropic difference-of-Gaussian (DoG) algorithm: It firstly constructs an edge tangent flow (ETF) that is a vector field perpendicular to gradient directions, preserving salient edge directions while directing weak edges to follow the neighboring dominant ones. Lines are then constructed by applying an anisotropic DoG filter guided by the local flow recorded in the ETF. Because the algorithm is based on edge directions, spatial coherence between isolated, scattered edge components that clutter the output image is well established by effectively connecting or ignoring those components. Therefore, 'coherent' of the original naming of the algorithm in fact indicates '*spatially* coherent'.

Our semi-automatic algorithm is derived from *Coherent Line Drawing* and [Wang et al. 2004]. Initially, the user outlines some important objects for the whole scene by specifying boundaries for those objects on a few keyframes. After the boundaries are propagated and interpolated in-between frames as stated in [Wang et al. 2004], the input video data is segmented into contiguous space-time volumes, where each volume's slice at each frame time directly corresponds to the subregion of an individual frame segmented by the boundaries. We then apply modified *Coherent Line Drawing* only to each segmented subregion of each frame as follows: an ETF and an optical flow vector field [Horn and Schunck 1981] for the whole region on every frame is calculated first. Between two consecutive frames, the ETF in the segmented subregions on the former frame is translated to the subsequent frame, guided by the estimated optical flow. The translated ETF is adjusted by linearly interpolating the ETF and the original ETF of the subsequent frame. The interpolated ETF in the segmented subregions on the subsequent frame now functions as the initial ETF to be translated to the next subsequent frame. The same process is repeated until reaching the final frame. Finally, the flow-based anisotropic DoG filtering is applied to those adjusted ETFs in the segmented subregions on all frames.

4. References

KANG, H., LEE, S., CHUI, C. K. 2007. Coherent Line Drawing, *In Proc. 5th International Symposium on Non-Photorealistic Animation and Rendering*, 43-50

HERTZMANN, A., AND PERLIN, K. 2000. Painterly rendering for video and interaction. *In Proc. NPAR 2000*, 7–12

HORN, B. K. P., AND SCHUNCH, B. G. 1981. Determining Optical Flow, *Artificial Intelligence* 17, 185-203

LITWINOWICZ, P. 1997. Processing images and video for an impressionist effect. *In Proc. SIGGRAPH 1997*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 151–158

SALISBURY M. P., WONG M. T., HUGHES J. F., AND SALESIN D. H. 1997, Orientable Textures for Image-Based Pen-and-Ink Illustration, *In Proc. SIGGRAPH 1997*, 401-406

WANG, J., XU, Y., SHUM, H.-Y., AND COHEN, M. F. 2004. Video Tooning. *ACM Transactions on Graphics* 23, 3, 574-583