

Temporally Coherent Video De-Anaglyph

Joan Sol Roo, Christian Richardt

▶ **To cite this version:**

Joan Sol Roo, Christian Richardt. Temporally Coherent Video De-Anaglyph. 2014, 10.1145/2614106.2614125 . hal-01060864

HAL Id: hal-01060864
https://inria.hal.science/hal-01060864

Submitted on 4 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Temporally Coherent Video De-Anaglyph

Joan Sol Roo*
Inria, UNICEN

Christian Richardt†
Inria

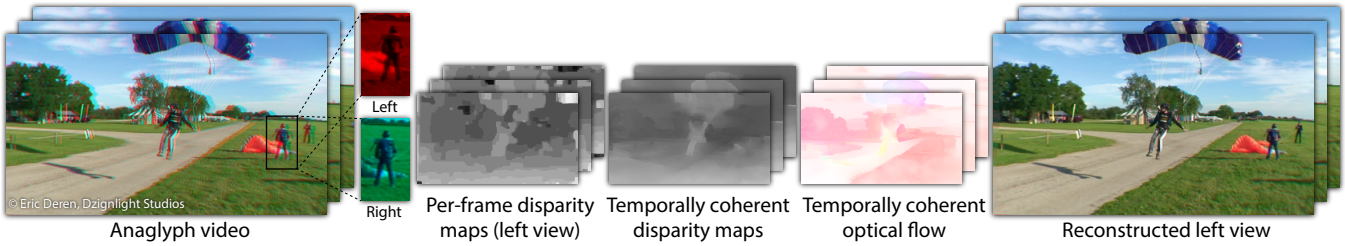


Figure 1: *We convert anaglyph videos (left) to temporally coherent full-color stereo videos (right). Our approach starts with rough, per-frame disparity maps and produces temporally coherent disparity maps and optical flow (center) that are used for reconstructing the stereo views.*

1 Introduction

For a long time, stereoscopic 3D videos were usually encoded and shown in the *anaglyph* format. This format combines the two stereo views into a single color image by splitting its color spectrum and assigning each view to one half of the spectrum, for example red for the left and cyan (blue+green) for the right view. Glasses with matching color filters then separate the color channels again to provide the appropriate view to each eye. This simplicity made anaglyph stereo a popular choice for showing stereoscopic content, as it works with existing screens, projectors and print media. However, modern stereo displays and projectors natively support two full-color views, and avoid the viewing discomfort associated with anaglyph videos.

Our work investigates how to convert existing anaglyph videos to the full-color stereo format used by modern displays. Anaglyph videos only contain half the color information compared to the full-color videos, and the missing color channels need to be reconstructed from the existing ones in a plausible and temporally coherent fashion. Joulin and Kang [2013] propose an approach that works well for images, but their extension to video is limited by the heavy computational complexity of their approach. Other techniques only support single images and when applied to each frame of a video generally produce flickering results.

In our approach, we put the temporal coherence of the stereo results front and center by expressing Joulin and Kang's approach within the practical temporal consistency framework of Lang et al. [2012]. As a result, our approach is both efficient and temporally coherent. In addition, it computes temporally coherent optical flow and disparity maps that can be used for various post-processing tasks.

2 Our Approach

In a nutshell, we first estimate disparity maps between the color channels of the two views to bring them into correspondence. We then fill in missing color information in each view by warping the color channels to fit the other view, using edge-aware colorization to fill in occlusion areas detected using the left-right consistency check.

*e-mail: solroo@gmail.com
†e-mail: christian@richardt.name

Our approach addresses the 4 main problems of temporally coherent anaglyph-to-stereo video conversion:

- **Multimodal Input** – The same color may look very different in the two anaglyph views. The color red, for example, appears bright through the red filter but dark through the cyan filter. This makes it difficult to find good correspondences between views. We use SIFT flow [Liu et al. 2011] to provide robust correspondences despite these inconsistent intensities.
- **Temporal Consistency** – Per-frame disparity maps produce temporally incoherent results. We process per-frame SIFT flow disparity maps using Lang et al.'s framework to obtain temporally coherent disparity maps and hence stereo video results.
- **Channel Alignment** – The two stereo views are two different projections of the same scene. We use disparity-based warping to align the left with the right view, and vice versa.
- **Occlusions** – Some parts of a scene are only visible in one of the two views. In these areas, we use the edge information of the existing color channels to guide the colorization of the other channels from regions that are nearby in space, and along temporal motion trajectories, using the domain transform [Gastal and Oliveira 2011].

Figure 1 shows a few frames from an anaglyph video converted with our approach. The 400×225 pixel video is processed in 6 seconds per frame on a 12-core 2.8 GHz CPU. We also include a supplemental video that shows additional results, also as side-by-side stereo.

We will make our C++ implementations of SIFT flow, the domain transform and Lang et al.'s temporal consistency framework available on our project website http://tempconst.gforge.inria.fr/.

References

GASTAL, E. S. L., AND OLIVEIRA, M. M. 2011. Domain transform for edge-aware image and video processing. *ACM Transactions on Graphics 30*, 4, 69:1–12.

JOULIN, A., AND KANG, S. B. 2013. Recovering stereo pairs from anaglyphs. In *Proceedings of CVPR*, 289–296.

LANG, M., WANG, O., AYDIN, T., SMOLIC, A., AND GROSS, M. 2012. Practical temporal consistency for image-based graphics applications. *ACM Transactions on Graphics 31*, 4, 34:1–8.

LIU, C., YUEN, J., AND TORRALBA, A. 2011. SIFT flow: Dense correspondence across scenes and its applications. *Transactions on Pattern Analysis and Machine Intelligence 33*, 5, 978–994.