A 2D to 3D Video Converter using Optical Flow Information and Least Squares Regression

Hui-Yun Lee
Department of Computer Science and Information Engineering
Chang Gung University
s15253520022001@gmail.com

Jyh-Da Wei
Dept. CSIE, Chang Gung University
Ophthalmology Department, Keelung Chang Gung Memorial Hospital
jdwei@mail.cgu.edu.tw

CCS CONCEPTS

• **Computing methodologies** → **Image processing**; *Parametric curve and surface models*;

KEYWORDS

3d telepresence, optical flow, least squares regression

ACM Reference Format:
Hui-Yun Lee and Jyh-Da Wei. 2018. A 2D to 3D Video Converter using Optical Flow Information and Least Squares Regression. In *Proceedings of HPG '18*. ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

There are some existing methods that can read a video clip and transfer the frame stream into its 3D version. In 2011, Tsai et al. analyzed the line positions to separate the foreground and background parts in an image frame, and produced some successful transfer cases [Tsai et al. 2011]. Jung and Cai compared color cue information with a well-trained color database for assigning depth values to super-pixels in 2015 [Jung and Cai 2015]. Interestingly, we found during our survey that optical flow information was useful to determine the distance between the object and the lens [Horn and Schunck 1981]. This suggested that optical flow can be a promising feature for 2D-to-3D conversion and thus we develop a method that converts an existing 2D video to a 3D version by using optical flow information and least squares regression in this paper.

2 OUR METHOD

Our method includes the following pipelined processes. First, we read the video into frames. We then calculate the optical flow values for each particular frame. In doing this, we can assign six grades, from 0 to 5 with uniform steps, as the depth values to the corresponding pixels. Next, we use the mean-shift technique to partition a given image frame into super-pixels, and assign the maximum depth grade thereof to everywhere of the same mean-shift region. To smooth the change of depth, we also introduce a polynomial function to build a 3D surface according to the depth map. Finally we can render the reorganized 3D video.

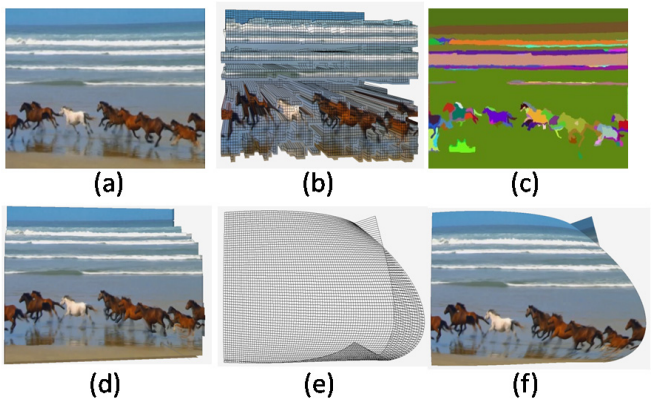

Figure 1: Experimental results of our method.

Figure 1 illustrates these steps, where 1(a) is the original image frame, 1(b) shows the optical flow values, 1(c) displays super-pixels resulting from mean shift process, and 1(d) presents the depth map with depth grades assigned onto mean-shift super-pixels. As Figures 1(e) and 1(f) show the fourth order polynomial surface and the transferred 3D image frame, respectively, we rotate the twisted surface by an angle of 10° to enhance 3D presentation. For further application, we can build 3D models like this and thus we can simulate dual channels for 3D telepresence.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
HPG '18, August 10–12, Vancouver, Canada
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology, Taiwan, R.O.C (grant nos. MOST 103-2221-E-182-049 and MOST 106-2221-E-182-075) and Chang Gung Memorial Hospital (grant no. BMRPB21).

REFERENCES

Berthold K. P. Horn and Brian G. Schunck. 1981. Determining Optical Flow. *Artificial Intelligence* 17 (1981), 185–203.

Cheolkon Jung and Jiji Cai. 2015. Superpixel matching-based depth propagation for 2D-to-3D conversion with joint bilateral filtering. In *Proc. the 2015 IEEE International Conference on Image Processing (ICIP 2015)*. 3515–3519.

Sung-Fang Tsai, Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen. 2011. A real-time 1080p 2D-to-3D video conversion system. *IEEE Transactions on Consumer Electronics* 57, 2 (May 2011), 915–922.