Optimizing Infinite Homography for Bullet-Time Effect

Ding Chen, Ryuuki Sakamoto
Yahoo Japan Corp.*

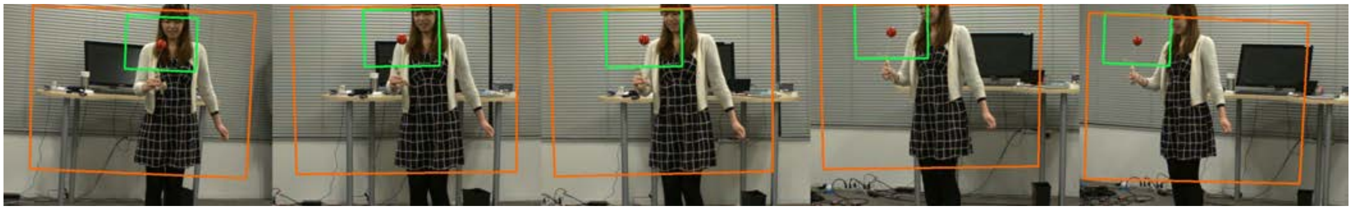

Figure 1: *Cropping areas in the original frames of 5 cameras that supposed to be final output with infinite Homography. The areas indicated by green window present the result by Tomiyama et al.; the orange window presents our result. To make the outcome frames, these areas must be aligned and magnified to the same resolution and aspect ratio as the original frames.*

1. Introduction

Bullet-Time is realized with flipping through frames taken by multi cameras surrounding an object. One of the most famous systems using this effect is the EyeVision [http://www.pvi-inc.com/eyevision] that involves using a robotic pan-tilt head to control the optical axis of the camera. In contrast, Tomiyama et al. [Tomiyama et al. 2006] use tripods to support cameras facing the position of an object approximately, then virtually pan-tilt each camera so that the optical axis towards an user-indicated 3D target point G by applying infinite Homography H_∞ which, however, causes some blank regions distorting the original frame. To eliminate blank regions in the outcome frame, much content is lost as cropping is applied for rectangular displays and as the original aspect ratio is kept. In contrast of conventional method of applying H_∞ which only corresponds to rotation about optical center with invariable camera intrinsic parameters, our algorithm further optimizes the intrinsic parameters in order to retain as much as content as possible for the outcome frames and achieve the smoothness of Bullet-Time effect.

2. Optimizing intrinsic parameters of H_∞

To keep more content, the 2D position g_k (k varies from 1 to N as the number of cameras) of G in original frames is expected to be close to the position p_k where G is mapped in the distorted frame after performing H_∞. p_k is decided by the principal point of intrinsic parameters which defaults to image center. Altering p_k means that move the image surface in parallel with the same camera rotation matrix and optical center. That means, an unusual camera model in which the optical axis does not pass through the image center, but this model effects slightly in the cognitive sense because it's difficult to notice in casual viewing. Then, the distorted frame is expanded until overlapping area of all camera frames meet the required size of crop window without blank regions by manipulate the focal length f_k of intrinsic parameters.

* e-mail: {tchin, ryusakam}@yahoo-corp.jp

SIGGRAPH 2014, August 10 – 14, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-2958-3/14/08

However, optimizing p_k and f_k on individual camera separately will result in a very uncomfortable Bullet-Time effect with no continuity in cognitive when changing a frame from one camera to next. For maintaining the smoothness of the effect as well as retaining the original frame content at the greatest extent possible, we proposed the following algorithm.

Optimizing principal point: The smaller the difference between g_k and p_k is, the lower the content loss is. Therefore the principal point p_k can be set as a fixed value (i.e. as the average of g_k), then p_k is closer to g_k if g_k is away from image center, G becomes immobile in all cameras so that there will be no sense of incongruity. Also, in experience, incongruity is insignificant if p_k of each camera moved smoothly by setting p_k at equal intervals on a straight line that fitted by simple linear regression through the set of g_k. This minimized the difference between g_k and p_k so that the content loss is lower than employing fixed value.

Optimizing focal length: f_k decides the magnitude of the appearance of objects in each camera. In the case that the distance between optical center and G is uneven among all the cameras, f_k can be set as a fixed value in proportion to the shortest distance to make the size of the object consistent in all cameras to keep the continuity in cognitive. Also inspired by varying p_k smoothly, incongruity is insignificant if f_k is changed smoothly by setting f_k on a parabolic curve that fitted through the set of optical center of all cameras. This minimized the expansion of each camera frame so that the content loss is lower than employing fixed value.

We compared our method as the combination of optimizing p_k and f_k with Tomiyama et al. (see **Figure 1**) using loss ratio as an index. The loss ratio is a ratio of the content versus the content in the original frame. In this dataset, our result of loss ratio is 22% while Tomiyama et al. is 79%. By applying on several dataset, our algorithm is proved to be remarkably effective in the case where g_k is away from the image center and the distance between optical center and G is uneven among all the cameras.

References

TOMIYAMA, K., MIYAGAWA, I. AND IWADATE, Y. 2006. Prototyping of HD Multi-Viewpoint Image Generating System: live broadcasting use at gymnastic competition. In *IEIC Technical Report* 106, 429, 43–48.