

A New Video Stabilization Method for Airship Surveillance

Xiaozhou Xu^{1,a}, Jing Yu^{2,b} and Weidong Sun^{1,c}

¹*State Key Lab. of Intelligent Technology & Systems, Tsinghua National Lab. for Information Science & Technology, Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China*

²*College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China*

^a*peter63598@126.com*, ^b*yujing@tsinghua.edu.cn*, ^c*wdsun@tsinghua.edu.cn*

Abstract.

Video Stabilization is one of the most important tasks in video surveillance on boarded of airship platforms. Facing the new problems that airship platforms usually have strong low-frequency vibration caused by wind and low-frequency wobble in video streams could not be effectively reduced by those method designed for casual handheld devices, a new video stabilization method for airship surveillance is proposed. In this method, camera path estimation is carried out by a modified method using GPS and attitude data of the key frames, and a new camera path planning method is given under 3 generalized conditions of airship's motion properties. Quantitative experimental results of some real airship video streams show that, our method can effectively reduce the low-frequency wobble and produce more stable video streams.

Keywords: path reconstruction; path planning; airship surveillance

Introduction

Airships are more and more widely used in both civilian and military fields, and video surveillance is one of the most useful applications of airships which generated much interest in video stabilization for the fact that airships are strongly wiggled by wind.

Generally, as one of the post-processes of digital videos, video stabilization includes three main steps: 1) camera motion estimation; 2) smooth camera path planning; 3) applying per-frame warps to produce stabilized video frames.

For the first step, current approaches employ 2D-transformations (affine or projective transformation) which can't recover camera's real 3D trajectory [1], or use Structure from Motion to estimate camera's path [2,3] which is much more complex and need 3D reconstruction. For the second step, linear motion smoothing model [1], polynomial approximation [4] and spline fitting model [2] has been applied in 2D path smoothing model. As for the last step, motion-inpainting [1] and fixed aspect ratio crop window [4] are widely applied.

But most video stabilization methods only dampen high-frequency jitter which is always caused by motor vibration, and can't be used to remove the low-frequency swing of the video frames caused by wind and the inertia of large airship platforms. To overcome this problem and produce a stable video, a new stabilization method for airship platforms is proposed in this paper, which can make full use of the data from the airships' GPS and attitude sensor.

Camera Path Estimation

In this section, we introduce perspective projection transformation estimation and the notation of inter-frame motion estimation of airship's camera mainly given in [5,6].

Perspective Transformation Estimation. For inter-frame perspective transformation estimation, we track a set of SURF feature points, and apply RANSAC for outlier rejection. To balance the computational complexity and robustness, RANSAC method ends when the matched feature points agree with the estimated model up to a threshold distance. Perspective transformation matrix A_L between frames 1 and 0 is estimated in this way.

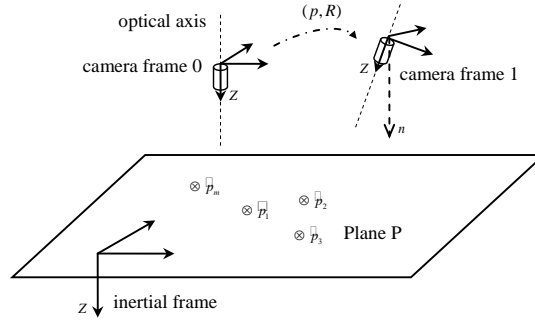


Fig. 1. Geometry of camera frame related to plane

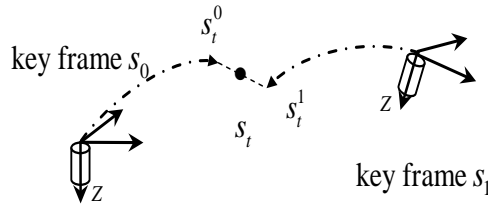


Fig. 2. Path estimation from the key frames

Motion Estimation. We describe the position and orientation of the camera in time t by $(p(t), R(t))$. The orientation of the camera on boarded of airship platforms is denoted by yaw ϕ , pitch θ and roll ψ . Assume we have some coplanar feature points $\{\tilde{q}_i\}_{i=1}^m \in \mathbf{P}$, where \mathbf{P} denotes the plane which means the ground. Fig. 1 shows the geometry between camera frame and inertia frame. If (p, R) are the motion between camera frames 0 and 1, then for $i = 1, \dots, m$, we have:

$$q_i^0 = (R + \frac{1}{d} p n^T) q_i^1, \quad (1)$$

where q_i^0, q_i^1 are the coordinates of the fixed points \tilde{q}_i in camera frames 0 and 1, n denotes the unit normal vector of plane \mathbf{P} in camera frame and d denotes the

distance between camera and the plane P in camera frame of frame 1. More detail process is introduced in [5,6].

It turns out that $A_L = \xi A$ for $\xi \in R$ where $A = R + pn^T / d$, and the middle singular value of the matrix A is equal to 1. We are able to recover the motion parameters $(p, (\theta, \phi, \psi))$ between two frames and structure parameters (n, d) in the second frame using SVD analysis of matrix A [5].

Camera Path Estimation and Accumulated Error Suppression. The accumulated error of path estimation will cause wobble and jitter in stabilized video. Here in this paper, a modified method to get more precise path estimation result and suppress accumulated error using data from GPS and attitude sensor is proposed.

Let $s_0^T = (p_0, \theta_0, \phi_0, \psi_0)$, $s_1^T = (p_1, \theta_1, \phi_1, \psi_1)$ denote the path parameters of camera in key frame 0 and key frame 1. So we can recover the path data of frame t from two key frames. Let $s_t^0 = (p_t^0, \theta_t^0, \phi_t^0, \psi_t^0)^T$, $s_t^1 = (p_t^1, \theta_t^1, \phi_t^1, \psi_t^1)^T$ denote these two results, then we get final path estimation of $s(t) = (p(t), \theta(t), \phi(t), \psi(t))^T$ using weighted sum of s_t^0 and s_t^1 :

$$s(t) = \frac{k-t}{k} s_t^0 + \frac{t}{k} s_t^1, \quad (2)$$

as shown is Fig. 2. For each segment between two key frames, we apply the algorithm above to have a more precise path estimation result.

Camera Path Planning

Considering airship's motion property and the need of video surveillance, our goal is to seek a smooth path satisfies several objectives: 1) the smooth path should be near the original path, so the FOV of the stabilized video remains as large as possible; 2) the smooth path should be smooth enough, most airship's target path is piecewise linear; 3) the smoothed orientation of airship should remain no roll and pitch. Considering the 3 conditions above, a new path planning method is proposed in this paper.

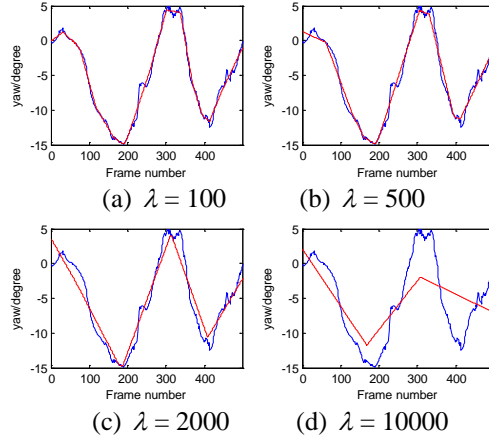


Fig. 3. Path planning result with different λ .

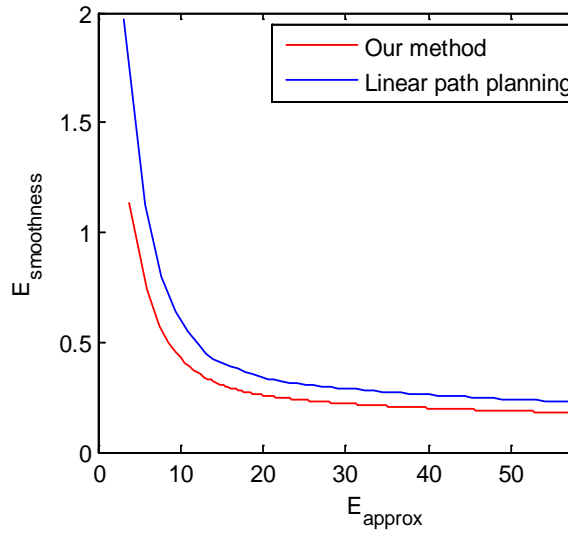


Fig. 4. Relation between FOV loss and video smoothness of two methods.

Orientation Planning. Considering the need of video surveillance, airship's camera should have no pitch and roll in order to have good video quality. So in our smooth path, pitch (around x-axis) and roll (around y-axis) remain 0. As yaw (around z-axis) result in image rotation in video, our goal is reducing both high-frequency jitter and low-frequency swing. Our method is related to L1 trend

filtering [7], which minimize the second derivate in L1 norm, therefore result in a series of linear angles. The objective minimizing function is:

$$\frac{1}{2} \sum_{t=1}^n (\phi(t) - \phi'(t))^2 + \lambda \sum_{t=2}^{n-1} |\phi'(t-1) - 2\phi'(t) + \phi'(t+1)|, \quad (3)$$

where $\phi'(t)$ is the smoothed result of $\phi(t)$, n is the input path length and λ is a nonnegative parameter used to control the trade-off between smoothness and size of the residual. L1 trend filtering is well studied and introduced in [7]. Fig. 3 shows a result of angle path planning of a series of yaw with 500 frames and 4 different λ , where the blue data denotes the original data of yaw path and the red data denotes the path planning result. As λ increases, the output sequence consists of fewer linear pieces and the residual grows. To balance the linearity of path planning and its residual error, we set $\lambda = 2000$ in the rest of this paper.

Position Planning. We disambiguate the motion estimation result of (p, d) using prior GPS and attitude data of key frames. We apply the same method to position series with yaw series. Let $\{p'(t)\}$ denotes the smooth position path of camera and $\{p'(t)\}$ could be get through L1 trend filtering. The objective minimizing function is:

$$\frac{1}{2} \sum_{t=1}^n (p(t) - p'(t))^2 + \lambda \sum_{t=2}^{n-1} |p'(t-1) - 2p'(t) + p'(t+1)|. \quad (4)$$

Video Retargeting

The last step is retargeting video frames with the estimated new camera position and orientation given in the previous section. So clearly the new frame in stabilized video could be rendered by projection model with camera motion $(p'(t) - p(t), \theta'(t) - \theta(t), -\phi(t), -\psi(t))^T$. The method of using motion to produce perspective transformation matrix $A_L(t)$ for frame t is well learned. We apply this model to our task and apply a fixed aspect ratio of 0.9 crop window to the new frames whose FOV is reduced.

Results and Evaluation

We evaluate our method on a set of videos taken by Cannon 5DMarkII on airship platform with the resolution of 576*720. The results from these videos show that our method can effectively reduce both high-frequency jitter and low-frequency wobble of the videos.

Considering the evaluation of stabilized video, two quantitative indexes are given in this paper: FOV loss and video smoothness which are represented by two penalty term that are the smaller the better:

$$E_{approx} = \sum (p(t) - p'(t))^2 + a \sum (\phi(t) - \phi'(t))^2, \quad (5)$$

$$E_{smoothness} = \sum |p'(t-1) - 2p'(t) + p'(t+1)| + b \sum |\phi'(t-1) - 2\phi'(t) + \phi'(t+1)|, \quad (6)$$

where a, b are the scale factor between position and orientation terms. In this paper, we set $a = 10000, b = 100$. These two indexes change while λ in path planning method changes. The curve between E_{approx} and $E_{smoothness}$ shows the stabilization quality of one method, and the result of our experimental video streams are shown in Fig. 4 compared with linear path planning method. Fig. 4 shows that the video smoothness $E_{smoothness}$ of our proposed method is about 20% better than that of linear path planning method [1] with the same FOV loss E_{approx} .

Conclusion

In this paper, a new video stabilization method combined with camera path reconstruction and path planning using L1 trend filter is proposed. This method works as post-process and can effectively stabilize video streams for airship surveillance by dampening both low-frequency wobble and high-frequency jitter. Comparing with traditional linear path planning method, our method performs better regarding to stabilized video's FOV loss and smoothness indexes.

Acknowledgements

This work is supported by the National Science & Technology Pillar Program (No.2012BAH31B01) and National Nature Science Foundation of China (No.61171117).

References

- [1] Matsushita, Y., Ofek, E., Ge, W., Tang, X., & Shum, H. Y. (2006). Full-frame video stabilization with motion inpainting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7), 1150-1163.
- [2] Kopf, J., Cohen, M. F., & Szeliski, R. (2014). First-person hyper-lapse videos. *ACM Transactions on Graphics (TOG)*, 33(4), 78.
- [3] Liu, F., Gleicher, M., Jin, H., & Agarwala, A. (2009, July). Content-preserving warps for 3D video stabilization. *ACM Transactions on Graphics (TOG)*, 28(3), 44.
- [4] Grundmann, M., Kwatra, V., & Essa, I. (2011, June). Auto-directed video stabilization with robust l1 optimal camera paths. *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 225-232.
- [5] Tsai, R.Y., Huang, T.S., Wei-Le Zhu (1982). Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(4), 525-534.
- [6] Shakernia, O., Ma, Y., John Koo, T., John, T., & Sastry, S. (1999). Landing an unmanned air vehicle: vision based motion estimation and nonlinear control. *Asian Journal of Control*.
- [7] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky (2009). l1 trend filtering. *SIAM Review*, problems and techniques section, 51(2), 339–360.