Supplementary Material of Video Stabilization Using IMU-Based Upright Adjustment Prior

Jianxiang Rong, Hua Huang, *Senior Member, IEEE* and Jia Li

I. Introduction

In the supplementary materials, a detailed derivation of the homography matrix decomposition algorithm will be provided.

II. Homography Decomposition

To independently process the different types of motion within a homography transformation, we develop a new homography decomposition algorithm. This algorithm can decompose a 2D homography transformation between two frames into a combination of different fundamental 2D motions. Specifically, the homography matrix is decomposed into translational motion, rotational motion, scaling motion, shear motion, and perspective motion.

A. Motion Components

For a given homography matrix:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}, \tag{1}$$

We will discuss the five 2D motion components contained in the homography matrix. And the calculation of performing this decomposition.

(1) **Translation Component**: The translation component represents the translation of the image along the x and y axes, with 2 degrees of freedom. If the image moves by t_x in the x direction and t_y in the y direction, the matrix form is:

$$\mathbf{H}_{trans} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

(2) **Rotation Component**: The rotation component represents the rotation of the image around a certain center by a specific angle, with 1 degree of freedom. If the center of rotation is (c_x, c_y) and the image rotates clockwise by θ degrees, the matrix form is:

$$\begin{aligned} \mathbf{H}_{rot} = & \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \\ & \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} \\ = & \begin{bmatrix} \cos\theta & -\sin\theta & c_x(1-\cos\theta)+c_y\sin\theta \\ \sin\theta & \cos\theta & c_y(1-\cos\theta)-c_x\sin\theta \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{3}$$

(3) **Scaling Component**: The scaling component represents the scaling of the image along the x and y axes around a certain center, with 2 degrees of freedom. If the center of scaling is (c_x, c_y), the scaling factor in the x direction is λ_x, and the scaling factor in the y direction is λ_y, the matrix form is:

$$\begin{aligned} \mathbf{H}_{sca} = & \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} \\ = & \begin{bmatrix} \lambda_x & 0 & c_x(1-\lambda_x) \\ 0 & \lambda_y & c_y(1-\lambda_y) \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{4}$$

(4) **Shearing Component**: The shearing component represents the shearing of the image along a certain axis direction. The x-axis or y-axis direction can be used as the direction of shear, with 1 degree of freedom. In the case where the x-axis is the direction of shear, if the center of shearing is (c_x, c_y) and the shearing amount in the x-axis direction is s_h, the matrix form is:

$$\begin{aligned} \mathbf{H}_{shear} = & \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} \\ = & \begin{bmatrix} 1 & s_h & -c_y s_h \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{5}$$

In the case where the y-axis is the direction of shear, if the shearing amount in the y-axis direction is s_h, the matrix form is:

J.-X. Rong is with the School of Computer Science, Beijing Institute of Technology, Beijing, China. H. Huang and J. Li are with the School of Artificial Intelligence, Beijing Normal University, Beijing, China.

$$\mathbf{H}_{shear} = \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & -c_x s_h \\ 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

(5) **Perspective Component**: The perspective component represents the flipping of the image along the x and y axes as rotation axes, with 2 degrees of freedom. If the center of flipping is (c_x, c_y) and the flipping amounts along the x and y axes are ϕ_x and ϕ_y respectively, the matrix form is:

$$\mathbf{H}_{perp} = \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \phi_x & \phi_y & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c_x\phi_x + 1 & c_x\phi_y & -c_x(c_x\phi_x + c_y\phi_y) \\ c_y\phi_x & c_y\phi_y + 1 & -c_y(c_x\phi_x + c_y\phi_y) \\ \phi_x & \phi_y & 1 - c_x\phi_x - c_y\phi_y \end{bmatrix}. \tag{7}$$

This algorithm decomposes the homography matrix to obtain the aforementioned five motion components. The sum of the degrees of freedom of these components is 8, which matches the 8 degrees of freedom of the homography matrix. In theory, the decomposition can be performed as follows:

$$\mathbf{H} = \mathbf{H}_{trans} \cdot \mathbf{H}_{rot} \cdot \mathbf{H}_{shear} \cdot \mathbf{H}_{sca} \cdot \mathbf{H}_{perp}. \tag{8}$$

It should be noted that the homography matrix has homogeneity. The result on the right side of Eq.(8) needs to be normalized by dividing by the element in the third row and third column, making this element equal to 1.

The Eq.(8) can be transformed into a system of equations regarding the different motion components of the homography transformation. The center of motion (c_x, c_y) is usually set at the center of the frame, as this configuration aligns better with human visual perception. Given the elements of the homography matrix $\{h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}\}$ and the center of motion (c_x, c_y), the process of solving for motion components $\{t_x, t_y, \theta, \lambda_x, \lambda_y, s_h, \phi_x, \phi_y\}$ can be carried out. The solved motion components can be further converted into their corresponding motion matrices.

B. Computational Process

In this section, we will provide the detailed computational process for homography matrix decomposition. Considering that the computations for shear in the x-axis direction and shear in the y-axis direction are slightly different, we will present the computational processes for both cases.

1) In the case of shear in the x-axis direction: In Eq.(8), first compute the right side of the equation, then normalize it so that the element in the third row and third column becomes 1. Setting it equal to the left side element-wise yields the system of equations:

$$\begin{cases} h_{11} = \dfrac{\lambda_{xc} + \phi_x x_{ct}}{n_\phi}, & \text{(9a)} \\ h_{12} = \dfrac{-\lambda_{ysmc} + \phi_y x_{ct}}{n_\phi}, & \text{(9b)} \\ h_{13} = \dfrac{x_{ct} - c_x h_{11} n_\phi - c_y h_{12} n_\phi}{n_\phi}, & \text{(9c)} \\ h_{21} = \dfrac{\lambda_{xs} + \phi_x y_{ct}}{n_\phi}, & \text{(9d)} \\ h_{22} = \dfrac{-\lambda_{ycps} + \phi_y y_{ct}}{n_\phi}, & \text{(9e)} \\ h_{23} = \dfrac{y_{ct} - c_x h_{21} n_\phi - c_y h_{22} n_\phi}{n_\phi}, & \text{(9f)} \\ h_{31} = \dfrac{\phi_x}{n_\phi}, & \text{(9g)} \\ h_{32} = \dfrac{\phi_y}{n_\phi}. & \text{(9h)} \end{cases}$$

For convenience, we define the following intermediate variables:

$$\begin{cases} n_\phi = 1 - c_x\phi_x - c_y\phi_y, & \text{(10a)} \\ \lambda_{xc} = \lambda_x \cos\theta, & \text{(10b)} \\ \lambda_{xs} = \lambda_x \sin\theta, & \text{(10c)} \\ \lambda_{ysmc} = \lambda_y(\sin\theta - s_h\cos\theta), & \text{(10d)} \\ \lambda_{ycps} = \lambda_y(\cos\theta + s_h\sin\theta), & \text{(10e)} \\ x_{ct} = c_x + t_x, & \text{(10f)} \\ y_{ct} = c_y + t_y. & \text{(10g)} \end{cases}$$

First, from Equations Eq.(9g) and Eq.(9h), we can derive:

$$\phi_x = \frac{h_{31}}{1 + c_x h_{31} + c_y h_{32}}, \tag{11}$$

$$\phi_y = \frac{h_{32}}{1 + c_x h_{31} + c_y h_{32}}, \tag{12}$$

In this way, n_ϕ has become a known element. For the remaining unknowns, first solve for the intermediate variables. From Eq.(9a), Eq.(9b), Eq.(9c), Eq.(9d), Eq.(9e), and Eq.(9f), we can obtain:

$$x_{ct} = n_\phi(h_{13} + c_x h_{11} + c_y h_{12}), \tag{13}$$

$$y_{ct} = n_\phi(h_{23} + c_x h_{21} + c_y h_{22}), \tag{14}$$

$$\lambda_{xc} = n_\phi(h_{11} - \phi_x(h_{13} + c_x h_{11} + c_y h_{12})), \tag{15}$$

$$\lambda_{xs} = n_\phi(h_{21} - \phi_x(h_{23} + c_x h_{21} + c_y h_{22})), \tag{16}$$

$$\lambda_{ysmc} = -n_\phi(h_{12} - \phi_y(h_{13} + c_x h_{11} + c_y h_{12})), \tag{17}$$

$$\lambda_{ycps} = n_\phi(h_{22} - \phi_y(h_{23} + c_x h_{21} + c_y h_{22})). \tag{18}$$

After solving for the above intermediate variables, we can directly obtain from Eq.(10f) and Eq.(10g):

$$t_x = x_{ct} - c_x, \tag{19}$$

$$t_y = y_{ct} - c_y. \tag{20}$$

Then, according to Equations Eq.(10b), Eq.(10c), Eq.(10d) and Eq.(10e), two sets of solutions can be obtained:

$$\begin{cases} s_h = \dfrac{\lambda_{xs}\lambda_{ycps} - \lambda_{xc}\lambda_{ysmc}}{\lambda_{xc}\lambda_{ycps} + \lambda_{xs}\lambda_{ysmc}}, & (21a) \\ \theta = -2\arctan\dfrac{\lambda_{xc} + \sqrt{z}}{\lambda_{xs}}, & (21b) \\ \lambda_x = -\sqrt{z}, & (21c) \\ \lambda_y = -\dfrac{\lambda_{xc}\lambda_{ycps} + \lambda_{xs}\lambda_{ysmc}}{\sqrt{z}}. & (21d) \end{cases}$$

$$\begin{cases} s_h = \dfrac{\lambda_{xs}\lambda_{ycps} - \lambda_{xc}\lambda_{ysmc}}{\lambda_{xc}\lambda_{ycps} + \lambda_{xs}\lambda_{ysmc}}, & (22a) \\ \theta = -2\arctan\dfrac{\lambda_{xc} - \sqrt{z}}{\lambda_{xs}}, & (22b) \\ \lambda_x = \sqrt{z}, & (22c) \\ \lambda_y = \dfrac{\lambda_{xc}\lambda_{ycps} + \lambda_{xs}\lambda_{ysmc}}{\sqrt{z}}. & (22d) \end{cases}$$

Where $z = \lambda_{xc}^2 + \lambda_{xs}^2$. Both sets of solutions satisfy the system of equations and the homography matrix decomposition formula. In practice, the scaling factors between video frames are usually positive numbers; therefore, we adopt the solution from Eq.(22a), Eq.(22b), Eq.(22c) and Eq.(22d) as the result. After calculating all the necessary parameters for the matrix decomposition, we substitute them into each component of the decomposition, thereby completing the homography matrix decomposition.

2) In the case of shear in the y-axis direction: On the other hand, we present the computation process based on shear in the y-axis direction for decomposition in specific cases. Based on the shear matrix in Eq.(6), we can obtain a new system of equations by Eq.(8):

$$\begin{cases} h_{11} = \dfrac{\lambda_{xcms} + \phi_x x_{ct}}{n_\phi}, & (23a) \\ h_{12} = \dfrac{-\lambda_{ys} + \phi_y x_{ct}}{n_\phi}, & (23b) \\ h_{13} = \dfrac{x_{ct} - c_x h_{11} n_\phi - c_y h_{12} n_\phi}{n_\phi}, & (23c) \\ h_{21} = \dfrac{\lambda_{xspc} + \phi_x y_{ct}}{n_\phi}, & (23d) \\ h_{22} = \dfrac{-\lambda_{yc} + \phi_y y_{ct}}{n_\phi}, & (23e) \\ h_{23} = \dfrac{y_{ct} - c_x h_{21} n_\phi - c_y h_{22} n_\phi}{n_\phi}, & (23f) \\ h_{31} = \dfrac{\phi_x}{n_\phi}, & (23g) \\ h_{32} = \dfrac{\phi_y}{n_\phi}. & (23h) \end{cases}$$

For convenience, we define the following intermediate variables:

$$\begin{cases} n_\phi = 1 - c_x\phi_x - c_y\phi_y, & (24a) \\ \lambda_{xcms} = \lambda_x(\cos\theta - s_h\sin\theta), & (24b) \\ \lambda_{xspc} = \lambda_x(\sin\theta + s_h\cos\theta), & (24c) \\ \lambda_{ys} = \lambda_y\sin\theta, & (24d) \\ \lambda_{yc} = \lambda_y\cos\theta, & (24e) \\ x_{ct} = c_x + t_x, & (24f) \\ y_{ct} = c_y + t_y. & (24g) \end{cases}$$

The solutions for ϕ_x and ϕ_y are the same as before and will not be reiterated. From Eq.(23a), Eq.(23b), Eq.(23c), Eq.(23d), Eq.(23e), and Eq.(23f), we can obtain:

$$x_{ct} = n_\phi(h_{13} + c_x h_{11} + c_y h_{12}), \tag{25}$$

$$y_{ct} = n_\phi(h_{23} + c_x h_{21} + c_y h_{22}), \tag{26}$$

$$\lambda_{xcms} = n_\phi(h_{11} - \phi_x(h_{13} + c_x h_{11} + c_y h_{12})), \tag{27}$$

$$\lambda_{xspc} = n_\phi(h_{21} - \phi_x(h_{23} + c_x h_{21} + c_y h_{22})), \tag{28}$$

$$\lambda_{ys} = -n_\phi(h_{12} - \phi_y(h_{13} + c_x h_{11} + c_y h_{12})), \tag{29}$$

$$\lambda_{yc} = n_\phi(h_{22} - \phi_y(h_{23} + c_x h_{21} + c_y h_{22})). \tag{30}$$

The solutions for t_x and t_y are the same as before and will not be reiterated. According to Equations Eq.(10b), Eq.(10c), Eq.(10d) and Eq.(10e), two sets of solutions can be obtained:

$$\begin{cases} s_h = \dfrac{\lambda_{xspc}\lambda_{yc} - \lambda_{xcms}\lambda_{ys}}{\lambda_{xspc}\lambda_{ys} + \lambda_{xcms}\lambda_{yc}}, & (31a) \\ \theta = -2\arctan\dfrac{\lambda_{yc} + \sqrt{z}}{\lambda_{ys}}, & (31b) \\ \lambda_x = -\dfrac{\lambda_{xcms}\lambda_{yc} + \lambda_{xspc}\lambda_{ys}}{\sqrt{z}}, & (31c) \\ \lambda_y = -\sqrt{z}. & (31d) \end{cases}$$

$$\begin{cases} s_h = \dfrac{\lambda_{xspc}\lambda_{yc} - \lambda_{xcms}\lambda_{ys}}{\lambda_{xspc}\lambda_{ys} + \lambda_{xcms}\lambda_{yc}}, & (32a) \\ \theta = -2\arctan\dfrac{\lambda_{yc} - \sqrt{z}}{\lambda_{ys}}, & (32b) \\ \lambda_x = \dfrac{\lambda_{xcms}\lambda_{yc} + \lambda_{xspc}\lambda_{ys}}{\sqrt{z}}, & (32c) \\ \lambda_y = \sqrt{z}. & (32d) \end{cases}$$

For the same reasons mentioned earlier, we select the set of solutions with a positive scaling factor (Eq.(32a), Eq.(32b), Eq.(32c) and Eq.(32d)) as the final result.