1

Visual Transformation

1.1 Image Transform (Warping)

Let us transport the pixel at the location $\mathbf{u}_1 = (u_1, v_1)^\mathsf{T}$ in the first image, $\mathcal{I}_1$ to the new location $\mathbf{u}_2 = (u_2, v_2)^\mathsf{T}$ in the second image under a function, i.e., $\mathbf{u}_2 = f(\mathbf{u}_1)$. The new image $\mathcal{I}_2$ is created by re-arranging pixels from the first image:

$$\begin{aligned} \mathcal{I}_1(\mathbf{u}_1) &= \mathcal{I}_2(\mathbf{u}_2) && (1.1) \\ &= \mathcal{I}_2(f(\mathbf{u}_1)) && (1.2) \end{aligned}$$

where $\mathcal{I}_i(\mathbf{u}_i)$ is the pixel value $(r, g, b)^\mathsf{T}$ at $\mathbf{u}_i$ in the i^{th} image, i.e., the pixel value does not change. This operation of the pixel transport is called *image warping* where the target image is created by re-arranging the pixels in the source under a well-defined function[1].

[1] A function is well-defined if it gives the same result when the representation of the input is changed without changing the value of the input.—Wikipedia

Note that the image warping differs from *image filtering* where the pixel undergoes transformation while the coordinate stays the same, i.e.,

$$\mathcal{I}_2(\mathbf{u}) = g(\mathcal{I}_1(\mathbf{u})). \qquad (1.3)$$

Image blurring, convolution, and gradient computation falls in image filtering as shown in Figure 1.1.

Figure 1.1: Image filtering

(a) Scaling (b) Translation (c) Rotation

Figure 1.2: Translation and rotation

1.1.1 Scaling

Consider resizing an image with s_u and s_v factors along x and y coordinates, respectively.

$$u_2 = s_u u_1 \tag{1.4}$$

$$v_2 = s_v v_1 \tag{1.5}$$

When $s_u = s_v$, it is uniform scaling where the aspect ratio of the image is preserved. where $\mathbf{u} = (u, v)^\mathsf{T}$. Equation (1.6) can be rewritten as a vector linear transform in a homogeneous coordinate:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} s_u & 0 & 0 \\ 0 & s_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.6}$$

where $\mathbf{u} = (u, v)^\mathsf{T}$. Figure 1.2(a) illustrates non-uniform scaling of the left image of Figure 1.1 where the aspect ratio is changed.

1.1.2 Translation

When the origin of the image moves $(t_u, t_v)^\mathsf{T}$, the transformation can be written as:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_u \\ 0 & 1 & t_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix}. \tag{1.7}$$

Figure 1.2(b) illustrates translation of image where $t_u, t_v > 0$.

1.1.3 Rotation

When the image is rotated about the origin (left top corner), the transformation can be written as:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix}. \tag{1.8}$$

Figure 1.2(c) illustrates the rotation of the image about the origin with θ where $\theta > 0$.

1.1.4 Euclidean transformation

A composite of translation and rotation results in *Euclidean transformation*, which can be writte as:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & t_u \\ \sin\theta & \cos\theta & t_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.9}$$

$$= \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.10}$$

where $\mathbf{R} \in SO(2)$ is the 2×2 rotation matrix and $\mathbf{t} \in \mathbb{R}^2$ is the translation vector. Equation (1.10) can be rewritten as:

$$\mathbf{u}_2 = \mathbf{R}\mathbf{u}_1 + \mathbf{t} \tag{1.11}$$

where $\mathbf{u}_1$ is rotated first and then, translated. The transformed image possess the following properties:

- Degree of freedom: 3 = 1 (rotation) + 2 (translation)
- Length is preserved.
- Angle is preserved.
- Area is preserved.
- Parallel lines stay parallel.

Example 1.1 (Rotation about the principal point)**.** *How to represent a transform rotating about the principal point (image center)* $\mathbf{p} = (p_x, p_y)^\mathsf{T}$ *with* θ *as shown in Figure 1.3?*
Solution The rotation about the principal point is equivalent to translate the origin to the principal coordinate and rotate. Consider a new coordinate by translating the origin to principal coordinate $\mathbf{p} = (p_x, p_y)^\mathsf{T}$:

$$\overline{u} = u - p_x \tag{1.12}$$

$$\overline{v} = v - p_y \tag{1.13}$$

where $(\overline{u}, \overline{v})^\mathsf{T}$ is the new coordinate, or

$$\overline{\mathbf{u}} = \mathbf{u} - \mathbf{p}. \tag{1.14}$$

A pure rotation acts on the new coordinate:

$$\overline{\mathbf{u}}_1 = \mathbf{R}\overline{\mathbf{u}}_0 \tag{1.15}$$

By plugging Equation (1.14) into (1.15), the translation $\mathbf{m}$ can be computed as

$$\mathbf{t} = -\mathbf{R}\mathbf{p} + \mathbf{p} \tag{1.16}$$

where $\mathbf{u}_2 = \mathbf{R}\mathbf{u}_1 + \mathbf{t}$.

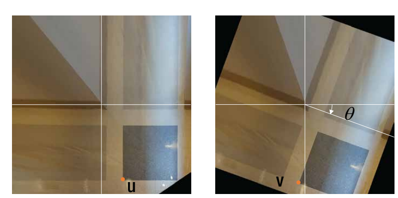

Figure 1.3: Euclidean transform.

1.1.5 Similarity Transform

A composite of uniform scaling and Euclidean transformation results in a similarity transform:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.17}$$

$$= \begin{bmatrix} s\mathbf{R} & s\mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.18}$$

$$= \begin{bmatrix} s\mathbf{R} & \mathbf{t}' \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix}. \tag{1.19}$$

Figure 1.4 illustrates a similarity transform. The transformed image possess the following properties:

- Degree of freedom: 4 = 1 (scale) + 1 (rotation) + 2 (translation)
- Length ratio is preserved.
- Angle is preserved.
- Parallel lines stay parallel.

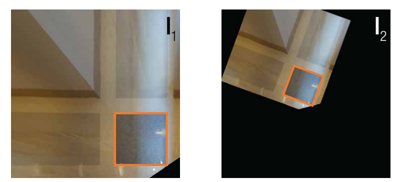

Figure 1.4: Similarity transform.

1.1.6 Affine Transform

The rotation matrix **R** in a similarity matrix is orthogonal, which allow preserving angle and length ratio, and limits the degree of the freedom to 4. When $s\mathbf{R}$ is a general 2×2 matrix, it is called an *affine transformation*:

$$\begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.20}$$

or,

$$u_2 = a_{11}u_1 + a_{12}v_1 + a_{13} \tag{1.21}$$

$$v_2 = a_{21}u_1 + a_{22}v_1 + a_{23}. \tag{1.22}$$

Figure 1.5 illustrates an affine transform. The transformed image possess the following properties:

- Degree of freedom: 6 (a_{11} to a_{23})
- Ratio of areas is preserved.
- Ratio of lengths of parallel line segments is preserved.
- Parallel lines stay parallel.

Figure 1.5: Affine transform.

Example 1.2 (Affine parallelism)**.** *Prove why parallel lines stay parallel under affine transformation.*
Solution

1.1.7 Projective Transform

The most general form of 2D linear transformation is *projective transformation* that generalizes the affine transformation:

$$\lambda \begin{bmatrix} \mathbf{u}_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.23}$$

$$= \mathbf{H} \begin{bmatrix} \mathbf{u}_1 \\ 1 \end{bmatrix} \tag{1.24}$$

where $\mathbf{H} \in \mathbb{R}^3$ is called *homography*. This can be also written as:

$$u_2 = \frac{h_{11}u_1 + h_{12}v_1 + h_{13}}{h_{31}u_1 + h_{32}v_1 + h_{33}} \tag{1.25}$$

$$v_2 = \frac{h_{21}u_1 + h_{22}v_1 + h_{23}}{h_{31}u_1 + h_{32}v_1 + h_{33}}. \tag{1.26}$$

Figure 1.6 illustrates a perspective transform. Note that the homography is defined up to scale, i.e., scalar multiplication results in equivalent homography, $\mathbf{H} \equiv \lambda\mathbf{H}$, or,

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{1.27}$$

$$\equiv \begin{bmatrix} \frac{h_{11}}{h_{33}} & \frac{h_{12}}{h_{33}} & \frac{h_{13}}{h_{33}} \\ \frac{h_{21}}{h_{33}} & \frac{h_{22}}{h_{33}} & \frac{h_{23}}{h_{33}} \\ \frac{h_{31}}{h_{33}} & \frac{h_{32}}{h_{33}} & 1 \end{bmatrix} \tag{1.28}$$

The transformed image possess the following properties:

- Degree of freedom: 8 = 9 (3 $\times$ 3 matrix entries) - 1 (defined up to scale)
- Cross ratio in collinear points is preserved.
- Order of contact is preserved.

Figure 1.6: Perspective transform.

Figure 1.7: Homography

Figure 1.8: Homography mapping