Beyond 2D translation Shi & Tomasi

The above method allows for local translations only, but images can differ from each other in more general ways. A more general model is to allow local scaling, shearing, and rotation. Let's suppose we have two image patches such that, in a neighborhood of $\mathbf{x_0} = (x_0, y_0)$, we have (to first order)

$$I(\mathbf{x_0} + (\mathbf{I} + \mathbf{D})(\mathbf{x} - \mathbf{x_0})) = I(\mathbf{x} + \mathbf{D}(\mathbf{x} - \mathbf{x_0}) + \mathbf{h}) = J(\mathbf{x})$$

where $\mathbf{h} = (h_x h_y)$ is a local translation as above, $\mathbf{x}$ is (x, y), and $\mathbf{D}$ is a general 2×2 matrix that models "deformations" which include shear, rotation, and scaling,

$$\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}.$$

To see how various transformations could be realized, consider that

- For an identity transformation (no rotation, scaling or shear), $D_{11} = D_{22} = 1$ and $D_{12} = D_{21} = 0$.
- For a uniform scaling by a scale factor s, $D_{11} = D_{22} = s$ and $D_{12} = D_{21} = 0$.
- For a shear by a factor λ one could have $D_{11} = 1, D_{12} = 0, D_{21} = \lambda, D_{22} = 1$, causing shearing in the direction of the y axis, or one could have $D_{11} = 1, D_{12} = \lambda, D_{21} = 0, D_{22} = 1$, causing shearing in the direction of the x axis.
- Finally, for a counter-clockwise rotation of θ about (x_0, y_0) $D_{11} = D_{22} = cos(\theta), D_{12} = -sin(\theta), D_{21} = sin(\theta)$.

In all of the above deformations via multiplation by the matrix $\mathbf{D}$ the shifting of $\mathbf{x}$ to $\mathbf{x_0}$ is so that all of these transformations are expressed locally, i.e., with respect to an origin at $\mathbf{x_0}$. We treat these deformations together in the constant matrix $\mathbf{D}$, whose entries we would like to estimate along with those of the vector $\mathbf{h}$, with the property that

$$\sum_{\mathbf{x} \in Ngd(\mathbf{x_0})} (I(\mathbf{x} + \mathbf{D}(\mathbf{x} - \mathbf{x_0}) + \mathbf{h}) - J(\mathbf{x}))^2$$

is as small as possible.

As above, we take a local Taylor series expansion of $I()$ around each $\mathbf{x} = (x, y)$. Letting

$$\Delta\mathbf{x} = \mathbf{x} - \mathbf{x_0} = (\Delta x, \Delta y)$$

we get

$$I(\mathbf{x} + \mathbf{D}\Delta\mathbf{x} + \mathbf{h}) \approx I(\mathbf{x}) + \frac{\partial I}{\partial x}(D_{11}\Delta x + D_{12}\Delta y + h_x) + \frac{\partial I}{\partial y}(D_{21}\Delta x + D_{22}\Delta y + h_y)$$

Notice that the variables here that we wish to solve for are the four D_* and the two h_*.

Substituting into the sum of squares above, we get an order 2 polynomial in a 6D space, which goes to ∞ as any of the six variables go to infinity. Let

$$\mathbf{d} = (D_{11}, D_{12}, D_{21}, D_{22}, h_x, h_y)$$

and let

$$\mathcal{I} = (\frac{\partial I}{\partial x}\Delta x, \frac{\partial I}{\partial x}\Delta y, \frac{\partial I}{\partial y}\Delta x, \frac{\partial I}{\partial y}\Delta y, \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

so we want to minimize

$$\sum (I(x,y) - J(x,y) + \mathcal{I} \cdot \mathbf{d})^2$$

Taking the partial derivatives with respect to the D_* and h_* variables gives six equations:

$$\sum (I(x,y) - J(x,y) + \mathbf{d} \cdot \mathcal{I})\mathcal{I} = \mathbf{0}$$

or

$$(\sum_{(x,y)\in Ngd(x_0,y_0)} \mathcal{I}\mathcal{I}^T)\ \mathbf{d} = \sum_{(x,y)\in Ngd(x_0,y_0)} (J(x,y) - I(x,y))\ \mathcal{I}$$

As we saw earlier in the lecture, one could solve iteratively for the six parameters $\mathbf{d}$.

When can we solve for the $\mathbf{d}$ vector? The 6×6 matrix would need to be invertible, and there is no reason why this is always going to be the case. For example, if $\frac{\partial I}{\partial y} = 0$ for all points in the neighborhood, then we are not going to be able to solve for the h_y variable. Here we are appealing to your intuition of the problem we saw earlier when we discussed the case of pure translation in the Lukas-Kanade registration approach.

A more interesting example would be if the image was rotationally symmetric locally around (x_0, y_0), for example, a "bull's eye" pattern. In this case, the translation component of the transformation could be recovered, but the rotational component of $\mathbf{D}$ could not. i.e. You would be able to rotate the intensity pattern without changing the sum of squared errors. In this case, we would find that there is a zero eigenvalue of the 6×6 matrix.

We are mentioning this 6D problem for a few reasons. First, it is important to appreciate that a simple translation model is not going to be sufficient for all situations. Second, we want to get you to start thinking about higher dimensional linear problems, and what the issues are when we try to solve them. We will see many examples later in the course.