Geometric Transformations

A geometric transformation moves pixels around without changing their grey level values. Let $p(x, y)$ be a picture, and $P(x, y)$ the result of applying a geometric transformation to $p(x, y)$. If the pixel (x, y) in p is mapped to the pixel $(\alpha(x, y), \beta(x, y))$ in P, then:

$$p(x, y) = P(\alpha(x, y), \beta(x, y))$$

and

$$P(x, y) = p(\alpha'(x, y), \beta'(x, y))$$

Where $\alpha'(x, y), \beta'(x, y)$ is the inverse transformation of $(\alpha(x, y), \beta(x, y))$.

Example: if $\alpha(x, y) = x + 2y$, $\beta(x, y) = x$, then in order to determine the inverse transformation we solve the system of equations:

$$X = x + 2y, \qquad Y = x$$

and get:

$$x = Y, \qquad y = \frac{X - Y}{2}$$

so that

$$\alpha'(x, y) = y, \qquad \beta'(x, y) = \frac{x - y}{2}$$

In the discrete case, if the forward transformation is applied to the picture p in order to get P, some pixels in P may not have a source in p, and other pixels may have more than a one source in p. Therefore, when applying a geometric transformation to an image, it is better to apply the inverse transformation. For each pixel (x, y), the value of $P(x, y)$ is determined from: the pixels in the neighborhood of $(\alpha'(x, y), \beta'(x, y))$ by some kind of interpolation. We discussed two interpolation techniques: The Nearest Neighbor, where the values of $(\alpha'(x, y), \beta'(x, y))$ are rounded to the nearest integer, and the Bilinear Interpolation.

Bilinear Interpolation

Let (x, y) be the (real valued) coordinates, obtained by the inverse transformation. Define $\underline{x}$ to be the largest integer smaller than or equal to x, $\bar{x}$ to be the smallest integer larger than or equal to x, $\underline{y}$ to be the largest integer smaller than or equal to y, $\bar{y}$ to be the smallest integer larger than or equal to y. For example, if $x = 3.4, y = 7.3$, then $\underline{x} = 3, \bar{x} = 4, \underline{y} = 7, \bar{y} = 8$. For our purpose we can assume that $\bar{x} = \underline{x} + 1$, and $\bar{y} = \underline{y} + 1$. $\underline{x}, \underline{y}$ are the integer values of x, y respectively. Put:

$$a = x - \underline{x}, \quad b = y - \underline{y}.$$

Notice that: $0 \leq a, b < 1$.

The point (x, y) falls in between the pixels $p(\underline{x}, \underline{y})$, $p(\underline{x}, \bar{y})$, $p(\bar{x}, \underline{y})$, $p(\bar{x}, \bar{y})$. The interpolated value is a weighted sum of these pixels:

$$P(x, y) = (1 - a)(1 - b)p(\underline{x}, \underline{y}) + (1 - a)bp(\underline{x}, \bar{y}) + a(1 - b)p(\bar{x}, \underline{y}) + abp(\bar{x}, \bar{y})$$