DIGITAL IMAGE PROCESSING

Unit 2 – Image Restoration: Geometric Transformations (Gonzalez & Woods)

Topic: Geometric Transformations

1. Purpose and Overview

Geometric transformations are used to correct geometric distortions introduced during image acquisition — for example, distortions caused by non-ideal camera/lens geometry, viewing-angle perspective, or scanning irregularities. Unlike the noise/blur restoration filters discussed earlier, geometric transformation modifies the spatial relationships between pixels rather than their grey-level values directly.

A geometric transformation, also called a rubber-sheet transformation, consists of two independent operations applied in sequence to correct a distorted image g(x,y) and produce a restored image f(x,y):

1. A spatial transformation, which defines the rearrangement of pixels on the image plane (moving pixels to new coordinate locations).
2. Grey-level interpolation, which assigns grey-level values to the spatially transformed pixels, since the new coordinates computed by the transformation are generally not integers.

2. Spatial Transformation of Coordinates

If (x,y) are the pixel coordinates in the original (distorted) image and (x′,y′) are coordinates in the corrected image, the spatial transformation is defined by a pair of mapping functions:

$$x' = r(x, y) \quad , \quad y' = s(x, y)$$

For example, if the distortion is purely a spatial scaling, r(x,y)=x/2 and s(x,y)=y/2. In practice, r and s are almost always unknown and must be approximated. The most widely used model is the bilinear approximation using a set of tie points (control points), giving:

$$x' = c1x + c2y + c3xy + c4 \quad , \quad y' = c5x + c6y + c7xy + c8$$

This model relates the (x,y) coordinates of a pixel in the distorted image to the corresponding (x′,y′) coordinates in the corrected image using eight coefficients c1…c8.

3. Common Elementary Transformations

Transformation	Mapping equations
Translation	$x' = x + t_x$, $y' = y + t_y$
Scaling	$x' = s_x \cdot x$, $y' = s_y \cdot y$
Rotation (by angle θ)	$x' = x \cdot \cos\theta - y \cdot \sin\theta$, $y' = x \cdot \sin\theta + y \cdot \cos\theta$
Shear (horizontal)	$x' = x + sh \cdot y$, $y' = y$

4. Tie Points and the Bilinear Coefficient Model

Since the exact form of the geometric distortion is generally unknown, tie points (control points) whose exact position is known both in the input (distorted) image and the output (corrected) image are used to estimate the transformation coefficients c1…c8. Because 8 unknowns must be solved, a minimum of 4 tie-point pairs (giving 4 equations each for x′ and y′) is required:

$$[x'i] = c1xi + c2yi + c3xiyi + c4 \quad for\ i = 1, 2, 3, 4$$

Solving these simultaneous equations for the four tie points gives the coefficients c1–c4 (and similarly c5–c8 from the y′ equations), which are then used to map every other pixel in the image.

5. Forward Mapping vs Inverse Mapping

The spatial transformation equations can be applied in the forward direction (from input to output) or the inverse direction (from output back to input); the choice has an important practical consequence.

Property	Statement
Forward mapping	Scans the input image pixel-by-pixel and copies each pixel to its computed location (x′,y′) in the output. Since (x′,y′) is generally not an integer, this can leave some output pixels unfilled (“holes”) and others written to more than once (overlaps).
Inverse mapping	Scans the output image pixel-by-pixel, and for each integer output location (x′,y′) computes the corresponding (generally non-integer) input location (x,y) using the inverse mapping functions, then interpolates a grey-level value from the neighbouring input pixels. This guarantees every output pixel gets a value, avoiding holes/overlaps — which is why inverse mapping is used in practice.

6. Grey-Level Interpolation

Because inverse mapping generally produces non-integer input coordinates (x,y), a grey-level value must be interpolated from the surrounding known pixel values. The commonly used interpolation schemes are:

Property	Statement
Nearest-neighbour interpolation	Assigns to (x,y) the grey level of the closest integer-coordinate pixel. Simplest and fastest, but can produce visible pixel-level distortions, especially for images with fine detail or straight edges.
Bilinear interpolation	Uses the four nearest neighbours: v(x,y) = a·x + b·y + c·xy + d, with the four coefficients a,b,c,d determined from the four known neighbouring points. Gives smoother results than nearest-neighbour, at moderate extra computational cost.
Bicubic interpolation	Uses the sixteen nearest neighbours, fitting a surface of the form v(x,y) = ΣΣ aij x^i y^j (i,j = 0..3), with the sixteen coefficients aij determined from the sixteen surrounding points. Produces the sharpest, most accurate results among the three, at the highest computational cost, and is the interpolation of choice in most commercial image-processing software.

7. Image Registration

Image registration is a closely related technique used when the transformation function between a distorted image and a reference (undistorted) image is entirely unknown — typically the case in remote sensing and medical imaging, where the same scene is imaged more than once (e.g. at different times or by different sensors).

The standard procedure uses tie points (also called control points or ground-control points) that can be identified precisely in both images:

1. Identify a set of tie points in the reference (input) image and their corresponding locations in the distorted (sensed) image.
2. Use the tie points to solve for the coefficients of a bilinear (or higher-order polynomial) spatial-transformation model, as in Sec. 4.
3. Apply the resulting transformation (using inverse mapping and grey-level interpolation) to align the sensed image with the reference image.

Registration accuracy depends directly on how precisely and how many tie points are located; more tie points, well distributed over the image, generally give a more accurate model of the underlying geometric distortion.

8. Summary — Choosing an Approach

Requirement	Recommended technique
Distortion function known exactly (e.g. pure scaling/rotation)	Direct application of the known mapping equations (Sec. 2–3)
Distortion function unknown, but tie points available	Bilinear tie-point model (Sec. 4), solved from ≥ 4 point pairs
Avoiding holes / overlaps in the output image	Always use inverse mapping (Sec. 5), never forward mapping
Fast processing, quality not critical	Nearest-neighbour interpolation (Sec. 6)
Balance of speed and visual quality	Bilinear interpolation (Sec. 6)
Highest visual quality (e.g. photographic resizing)	Bicubic interpolation (Sec. 6)
Aligning two images of the same scene (remote sensing / medical)	Full image registration procedure (Sec. 7)