4. GEOMETRIC CORRECTION

Sources of geometric distortion:
- sensor characteristics
  - optical distortion
  - aspect ratio
  - non-linear mirror velocity
  - detector geometry & scanning sequence

- viewing geometry
  - panoramic effect
  - earth curvature

- motions of the aircraft/satellite or target
  - attitude changes (pitch, roll, yaw)
  - position variations (altitude, slew)
  - earth rotation

Distortions appear as:
- changes of scale over the image
- irregularities in the angular relationships among the image elements
- displacement of objects in an image
- occlusion of one image element by another

Correcting the distortions is often costly:
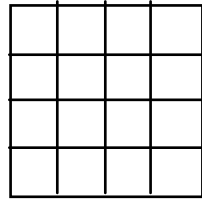- computer & operator time
- affects spatial and radiometric resolution

The nature of the "correction" depends upon the ultimate use of the data:
- area measurements ==> equal area projection
- shape measurements ==>projection which preserves the angular relationships of the
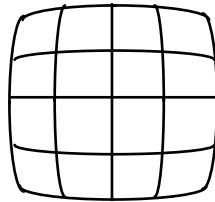   scene

Sources of Distortion
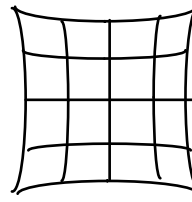### A. Sensor characteristics
    1. Optical distortion: not usually a problem. Most significant in imaging systems that use a large portion of the lens: vidicons, photographic systems
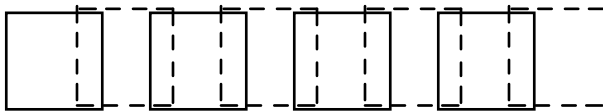

        (a)   (b)   (c)

      a) planimetric map
      b) negative optical distortion (barrel distortion)
      c) positive optical distortion (pincushion distortion)

    2. Aspect ratio: The aspect ratio is the ratio between scales in the vertical direction to that in the horizontal. Since distances are often measured by counting pixels, the shape of the pixel must be defined. A 10-pixel distance may mean one thing in the horizontal direction and something else entirely in the vertical direction.
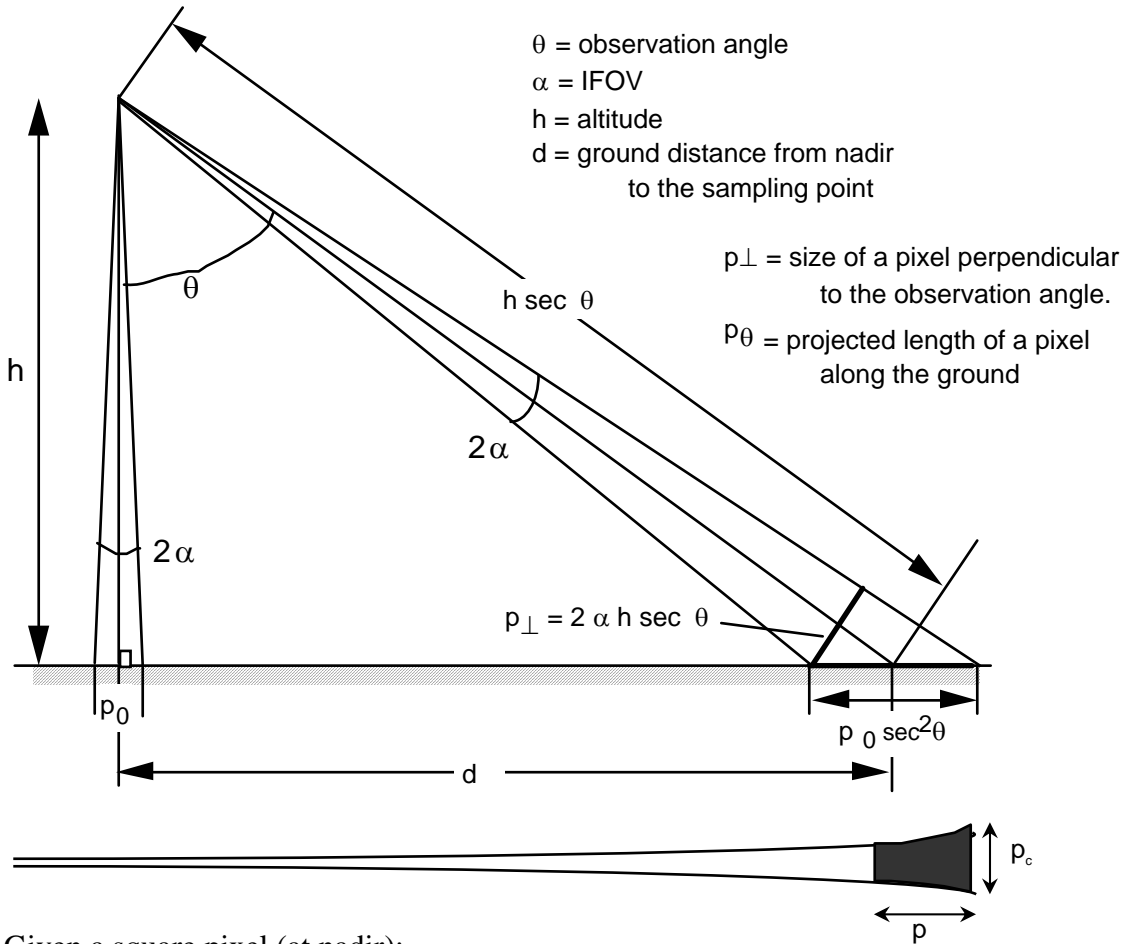

    3. Non-linear mirror velocity: If the velocity of the scanning mirror is constant (the ideal case) the sampling interval is well-defined and the size of the resolution cell, while variable over the scanline, is easily calculated. In many scanning systems, the mirror does not maintain a precise, constant angular velocity. This is notably the case with Landsat MSS which uses an oscillating rather than a rotating mirror. Along-scan distortions are introduced causing pixel compression or pixel stretching at various points along the scan line. This distortion is difficult to correct since it is difficult to measure and may change gradually with time.

    4. Detector geometry & scanning sequence: Any multispectral scanner attempts to sample several spectral bands simultaneously, and many sense several scan lines during one sweep of the scanning mirror. Ideally, all bands fir a single pixel would sense exactly the same point on the earth's surface and adjacent scan lines would line up exactly, pixel for pixel. However, the physical displacement of detectors in the image plane and the fact that only one detector can be sampled at a time introduce some misregistration of different bands for the same pixels, and displacement of adjacent scan lines.
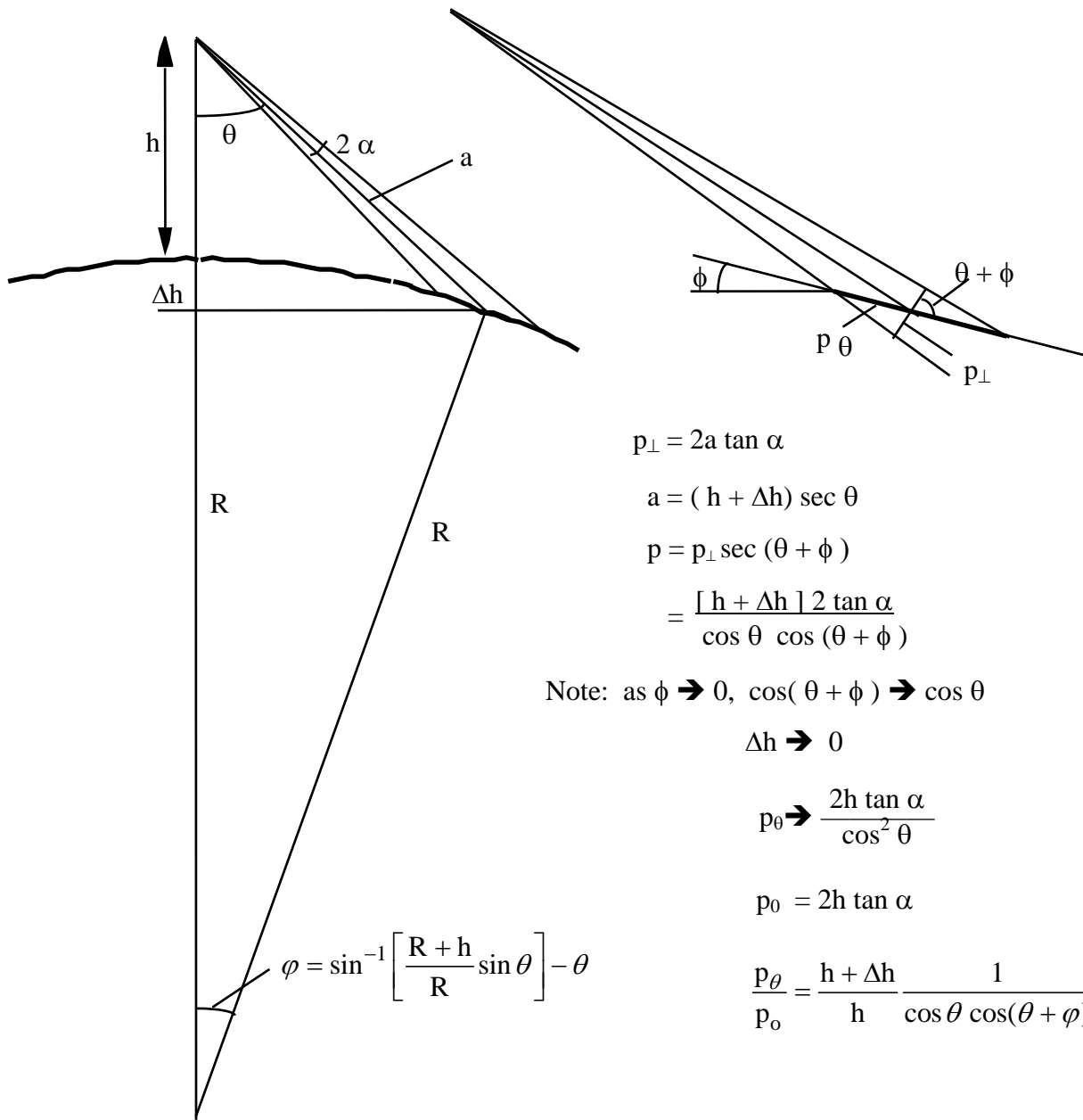
B. Viewing Geometry

Panoramic effect: a scale distortion (an increase in the resolution cell size) for observations
away from nadir due to an increase in the area viewed by the sensor.

a. panoramic effect (flat-earth)

θ = observation angle
α = IFOV
h = altitude
d = ground distance from nadir
        to the sampling point

$p\perp$ = size of a pixel perpendicular
            to the observation angle.
p_θ = projected length of a pixel
            along the ground

Given a square pixel (at nadir):

$p_0 = 2h \tan \alpha = 2h\alpha$ $d = h \tan \theta$
$p_\perp = 2 \alpha h \sec \theta = p_0 \sec \theta$
$p_0 = p_\perp \sec \theta = p_0 \sec^2 \theta$
$p_c = p_0 \sec \theta$

b. panoramic effect (curved earth)

$$p_\perp = 2a \tan \alpha$$

$$a = (h + \Delta h) \sec \theta$$

$$p = p_\perp \sec (\theta + \phi)$$

$$= \frac{[\, h + \Delta h \,]\; 2 \tan \alpha}{\cos \theta\ \cos (\theta + \phi)}$$

Note: as $\phi \rightarrow 0$, $\cos(\theta + \phi) \rightarrow \cos \theta$

$$\Delta h \rightarrow 0$$

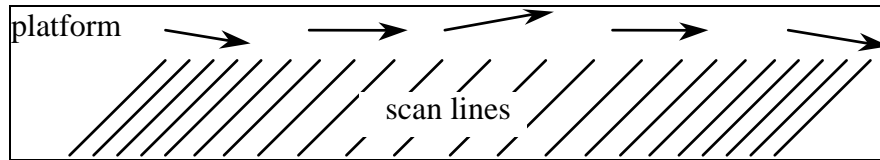$$p_\theta \rightarrow \frac{2h \tan \alpha}{\cos^2 \theta}$$

$$p_0 = 2h \tan \alpha$$

$$\varphi = \sin^{-1}\left[\frac{R + h}{R}\sin \theta\right] - \theta$$

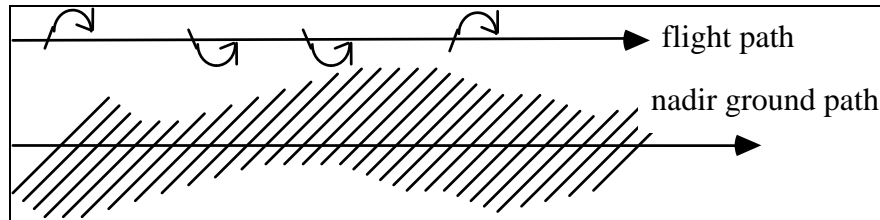$$\frac{p_\theta}{p_o} = \frac{h + \Delta h}{h}\; \frac{1}{\cos \theta\, \cos(\theta + \varphi)}$$

C. Motion of the platform (aircraft/satellite) or target

The motion and orientation of the platform may introduce distortions. This distortion is independent of the distortions caused by the sensor system itself.
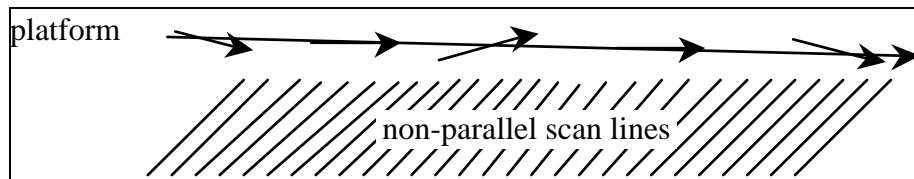
1. Attitude changes: changes in the platform that are significant over the time required to scan a full scene.
   a. Pitch - a rotation of the aircraft (or satellite, or ship) in the direction of motion. Pitch will cause changes in the spacing of the scan lines.
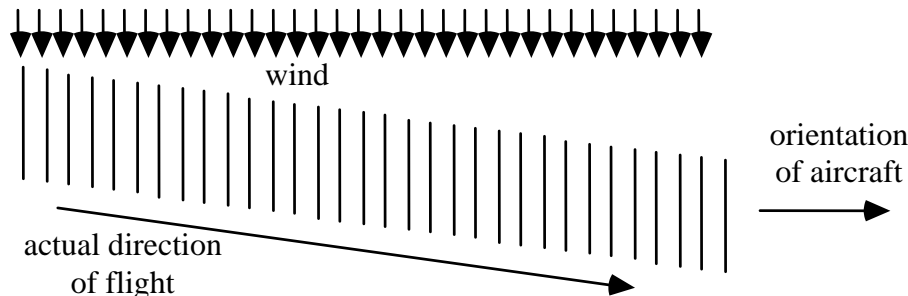

   b. Roll - the rotation about the velocity vector. Roll causes lateral shifts in the scan line position and scale changes in the line direction.

   c. Yaw - rotation of the satellite or aircraft about the local zenith vector, causes rotation and skew distortions.
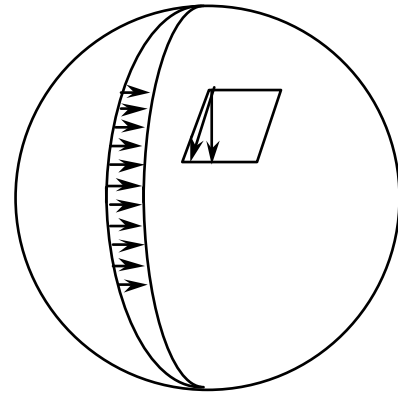

2. Position variations: This category includes gross deviations in the position and or direction of flight of the remote sensing platform. The simplest is a deviation in *altitude* which simply causes a change in scale.

3. Slew: Motion of the platform perpendicular to the intended direction of motion. The problem can occur when there is a cross wind over the flight path. The aircraft is oriented parallel to the intended flight path but the actual motion is at a slight angle to the intended direction. Thus, the scan lines will be perpendicular to the intended path but consecutive scan lines will be shifted in the direction perpendicular to the path.

D. Target effects: Distortions may be introduced into the image by target characteristics, especially by deviations of the target surface from a plane surface and by motions of the target.

Target motion: The type of distortion introduced by motion of the target will depend entirely on the nature of that motion relative to the sampling rate & sequence of the imaging system. For instance, when the motion of the scene is of the same order as the sampling rate. For example, a photo of a *nearby* building taken from the side window of a car moving at 50 mph when the shutter speed is 1/60 second, the image will be blurred. At the other extreme, the rotation of the earth is slow relative to the sampling speed of the Landsat MSS (about 0.4 µs) and it is even slow relative to the scan rate (33 ms/scan). Thus, there is no obvious blur in the final image. However, between the time that the first scan of a Landsat MSS image and the time of the last scan, the earth will have rotated a significant distance relative to the size of a resolution element. The resulting distortions for two different cases of earth rotation relative to satellite motion are illustrated

Geometric Correction

The general process of geometric correction consists of several steps:

1. Select the appropriate projection or reference map (or image).
   a. registration:simple point-too-point match of an image to another image or map,
   b. rectification:correcting an image to a specific map projection.

2. Select a regular grid which fits the desired projection (i.e., determine the spacing and position of the grid points.)

3. Select a set of "ground control points" (GCP's) -- pixels whose locations in the base map can be determined accurately.

4. Define the transformation and compute the positions of the reference grid points in the image coordinate system. The transformation should deform the grid in such a way that the average distance between GCP pixels and their map locations is minimized.

5. Resample the image data in order to assign gray values to each grid point.

Steps 1 and 2 depend on individual applications.

Map Projections

There is no perfect map projection. That is, there is no projection which preserves all the characteristics of a three dimensional surface. Projections have been developed which preserve either area <u>or</u> shape <u>or</u> direction <u>or</u> distance, but never all.

A. <u>Conformal (orthomorphic) projections</u>: Projections that represent **shape** correctly within certain limits. Relatively small portions of the earth's surface are properly represented in shape on a conformal projection. Actually, the conformal principle applies only to each infinitesimal element of the map. Angles at each point are correct, and consequently the local scale in every direction around any one point is constant, so the map user can measure distance and direction between near points with a minimum of difficulty

<u>Ex</u>: Mercator projection:	Lines of constant compass bearings are plotted as straight lines.
Transverse Mercator:	Lines of constant scale are straight lines parallel to the central meridian for the sphere. State Plane coordinates and Universal Transverse Mercator UTM) coordinates are special cases of this projection.
Space Oblique Mercator:	Projection developed (1978) specifically for the continuous mapping of imaging from satellites. The ground track for the satellite is held true to scale, and mapping is made basically conformal. (Because of the relative motion of Earth and satellite, the ground track is curved.)

B. <u>Equal area projections</u>: projections which show area correctly. These projections generally distort shape badly.

Ex: Albers Equal-Area Conic:	No area distortion. Along one or two standard parallels there is no scale or angular distortion.

C. <u>Azimuthal (Zenithal) projections</u>: Projections which correctly portray directions from a specified point.

Ex: Polar Stereographic:	The reference point is the Earth's pole. Lines of constant scale are circles centered at the pole.

D. <u>Equidistant projections</u>: Projections which portray distances correctly from a stated point.

Ex: Azimuthal Equidistant:	Shows both distances and directions correctly from the chosen center of the map. Shape and area are distorted.

There are also projections which preserve none of the four characteristics (area, shape, direction, or distance), but which are useful for other reasons. For small areas of the earth's surface nearly any standard projection will accurately represent the surface accurately. Locally, a simple rectangular coordinate system will often be entirely adequate.

In satellite imagery, the size of the area imaged is generally large enough that the earth's curvature cannot be ignored. The most natural projection in most cases is some form of oblique Mercator projection. For Landsat MSS data, the usual choices include Space Oblique Mercator (SOM), and Hotline Oblique Mercator (HOM), which are both relatively distortion-free for single scenes.

Selection of Ground Control Points
Ground control points are pixels whose positions can be determined accurately on the base map. These are usually easily recognizable features or landmarks in the image.

    Examples: - crossroads
                  - land/water boundaries
                  - cultural features (airstrips, buildings, etc.)

Geometric Transforms
Having chosen the control points, one must determine a transform that will map the base map coordinates to the image data. That is, find a function which relates the base map coordinates, (x, y), to the image coordinates (x', y'):

    $x' = f(x,y);$ $y' = f(x,y)$

The procedure is as follows:
1. Select a transformation (linear, 2nd order, 3rd order, ...)
2. Determine the coefficients for that transformation that will minimize the error in transforming the ground control points. (Usually a least-squares fit procedure.)
3. Analyze the residual errors. (How badly does the transform misplace data points?)
4. If only a few of the control points account for most of the error, then check the outliers to be certain that the coordinates and image positions are correct.
     a. If errors are found in the specification of the control points, rerun the routine that determines the coefficients for the transformation (step 2)
     b. If there is sufficient uncertainty in the location of a ground control point, discard it and rerun the routine (step 2).
4. If the residual errors remain unacceptable, try a different transformation (Usually, this means using a higher order equation.)

The simplest type of transformation is linear, involving translation, rotation, and magnification. Minimally, one needs only two control points to define a linear transform, although many more would be desirable.

1. <u>Linear Transforms</u>: Can be done using only two control points, although that's not a good idea)

    a. Translation: $x' = x + x_o$ $y' = y + y_o$

    b. Rotation: $x' = y \sin \theta + x \cos \theta$ $y' = y \cos \theta - x \sin \theta$

    c. Magnification $x' = mx$ $y' = ny$

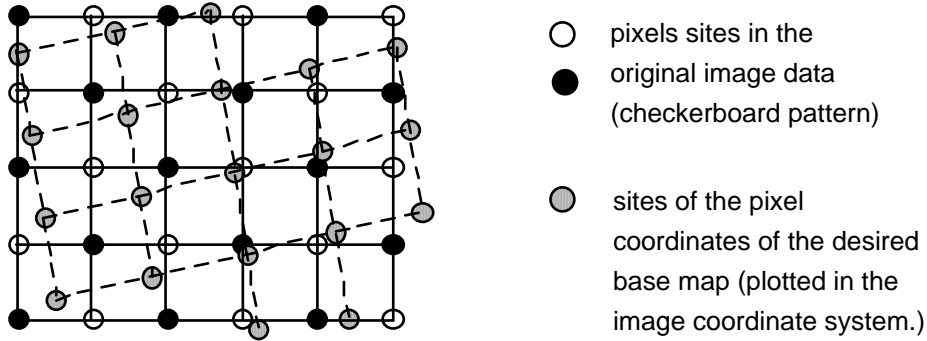   d. General linear transform: $x' = a_o + a_1x + a_2y$ $y' = b_o + b_1x + b_2y$

2. <u>Non-linear transformations</u>

    $x' = a_o + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 + \ldots$
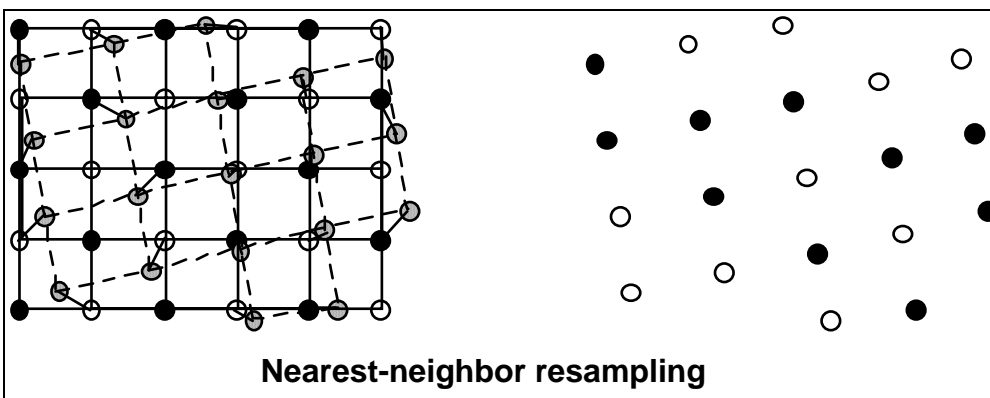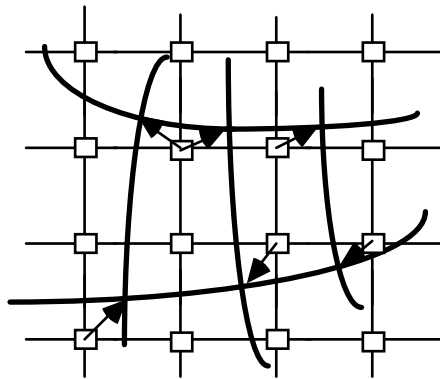 $y' = b_o + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 + \ldots$

Resampling

In resampling, each data point (pixel) in the base map coordinate system is assigned a value (intensity, gray value, etc.) based on the gray values of local image pixels. Consider an example in which an image of a checkerboard pattern is geometrically corrected.

○   pixels sites in the
●   original image data
     (checkerboard pattern)

◍   sites of the pixel
     coordinates of the desired
     base map (plotted in the
     image coordinate system.)
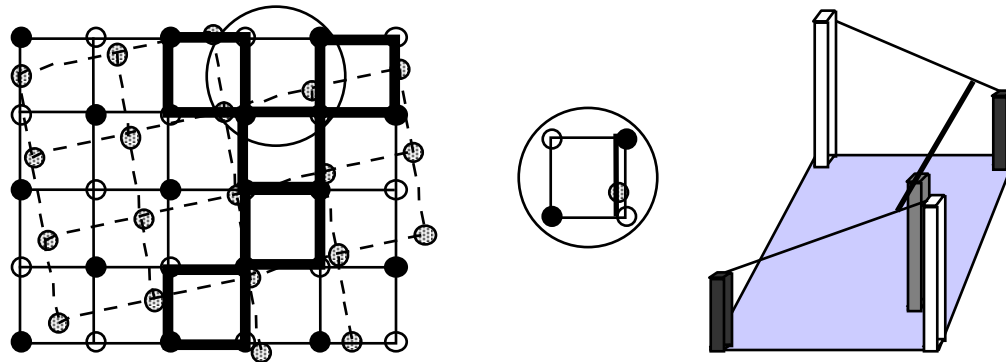
1. Nearest-neighbor resampling

The gray value of the image pixel is assigned to the nearest base map coordinate:
- the base map grid values correspond to actual measured values.
- the value assigned to the grid point will probably not be the same as that which would have been measured at that point.
- fastest and cheapest method of resampling.
- does the least radiometric damage to the data.
- tends to result in a blocky appearance at sharp boundaries.

Nearest-neighbor resampling

2. Bilinear interpolation

   Bilinear interpolation is a two-dimensional version of linear interpolation. It involves finding the four image pixels closest to the grid coordinate and obtaining the value for the grid coordinate by linear interpolation. In effect the variation in gray values is modeled as a series of planar surfaces (Assumes that the original scene varied linearly in the vicinity of the four sample pixels.)

Bilinear Interpolation

   Only three points are needed for a 2-dimensional linear interpolation since three non-collinear points uniquely define a plane. However, bilinear interpolation the 4 nearest neighbors. The approach is to use a sequence of one-dimensional interpolations between pairs of points rather than a single two-dimensional fit. The bilinear procedure is somewhat more efficient than a fully 2-dimensional procedure.

   In this example, the resampled data are undersampled, i.e., there are fewer grid coordinates than image pixels. The dark squares indicate adjacent sets of four pixels for which there is no corresponding grid coordinate. Interpolation is illustrated for the points within the circle.

   Note that the farther away a sample pixel is from the grid coordinate, the less the effect it will have on the value assigned to the coordinate. The image resulting from this resampling will have a smoother, less blocky appearance than one created by nearest-neighbor interpolation. At the same time there will be some loss of radiometric information since the value assigned to the grid pixel represents a weighted average of the four nearest points.