FBSP: Geometric image transformation

Nick Yao Larsen

nylarsen@cfin.au.dk

Spatial transformation

Spatial transformations perform a remapping of pixels

Such transformations include:

- Resizing/scaling/stretching
- Rotation
- Cropping
- Shearing
- Image projections

Affine transformations

Affine transformations:

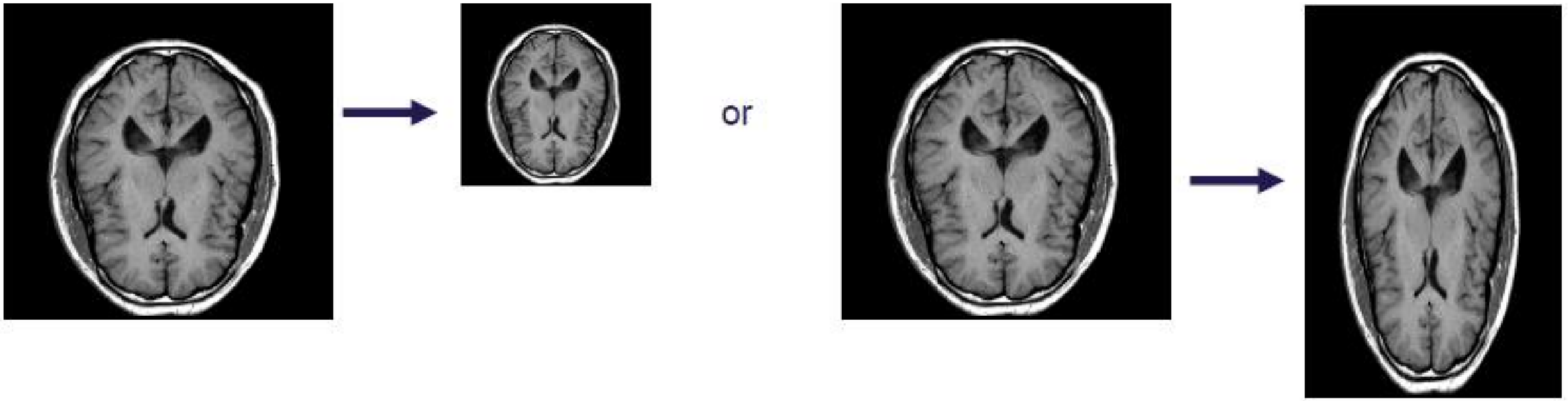
- Straightlines remainstraight
- Parallel lines remainparallel
- Rectangles may become parallelograms
- Three point-pairs are required to calculate transformation

Examples of affine transformations:

1. Scaling/resize/stretch
2. Rotation
3. Cropping
4. Shearing

1. Scaling/resize/stretch

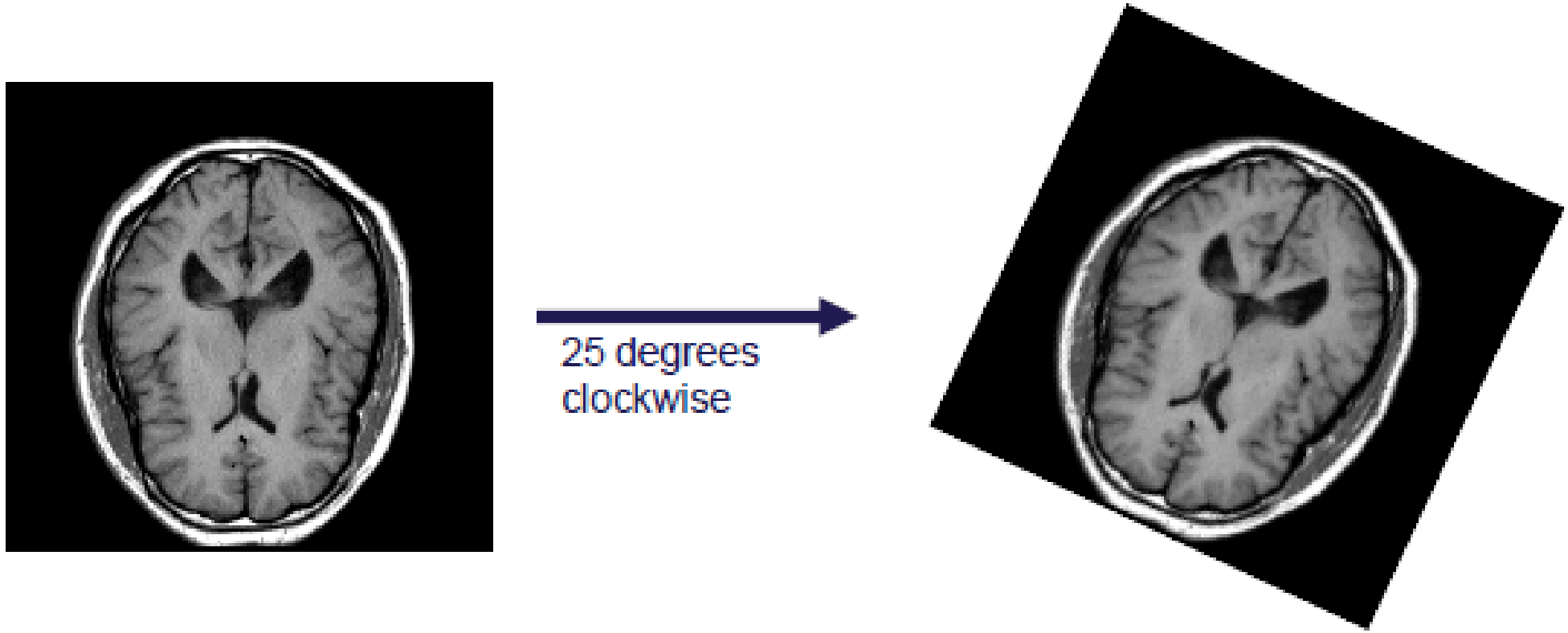
Scaling/resize/stretch the image to change size



2. Rotation

Rotate the image clockwise or counter clockwise

- Usually rotation is defined to rotate around the center point



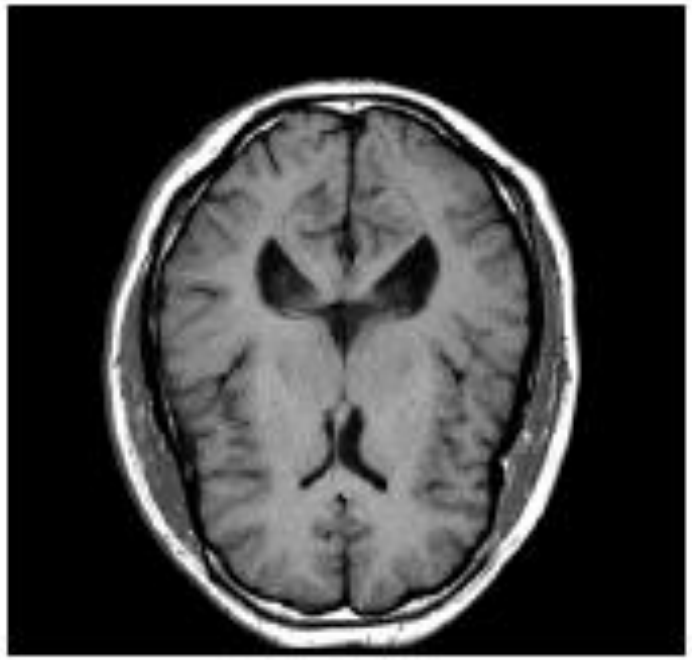
3. Cropping

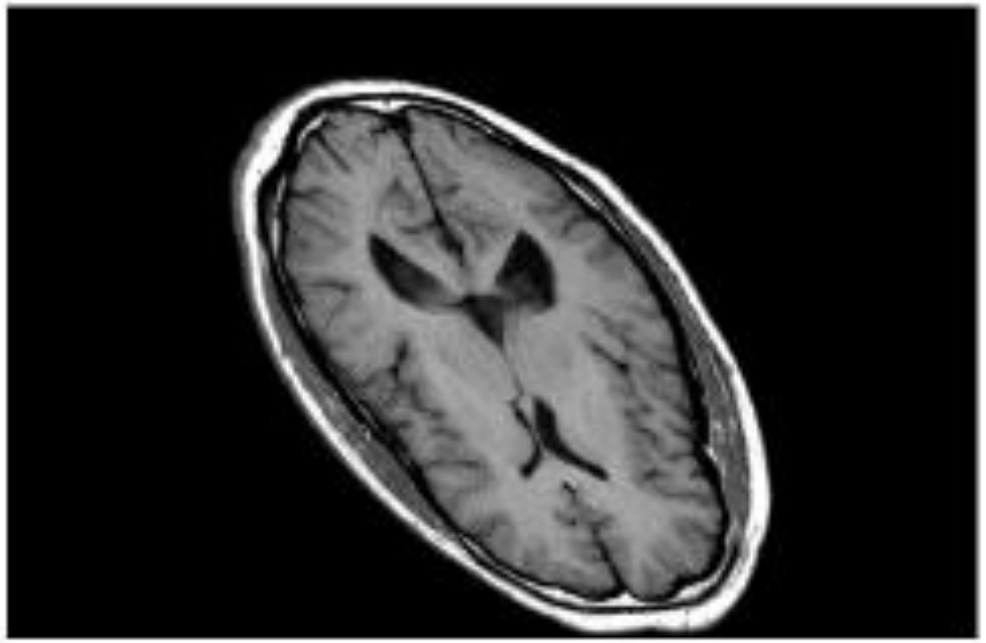
Crop the image to keep only a portion of the original image

- Usually the crop area is defined by a rectangle

4. Shearing

- Shift pixels horizontally or vertically with different amount depending on the position
- Like “pulling” a corner of the image

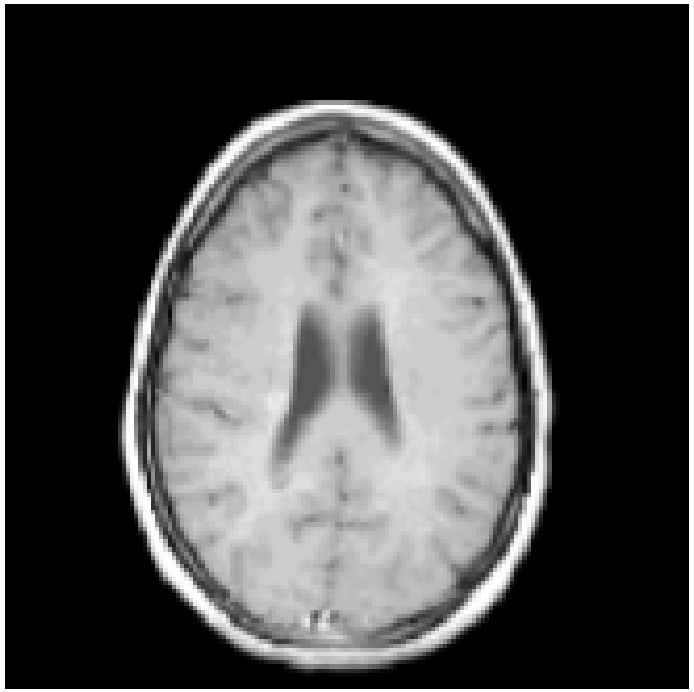


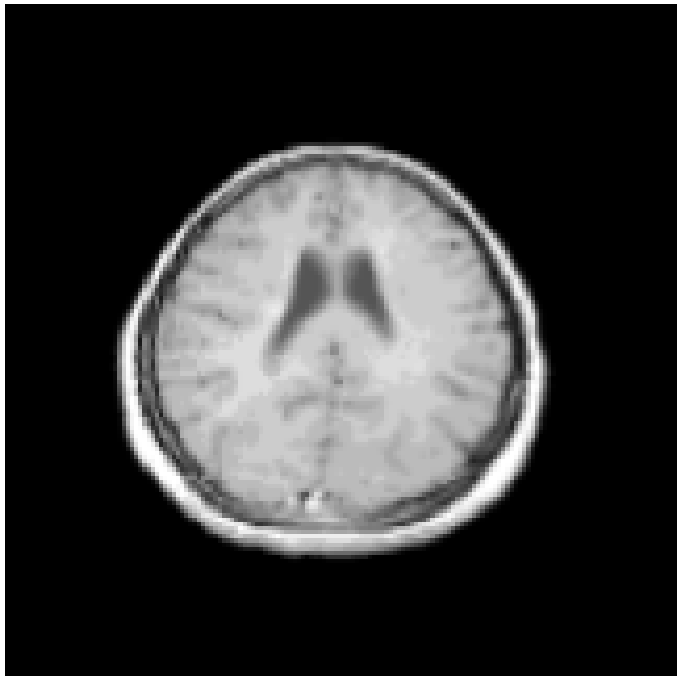


Projective transformations

Projective transformations:

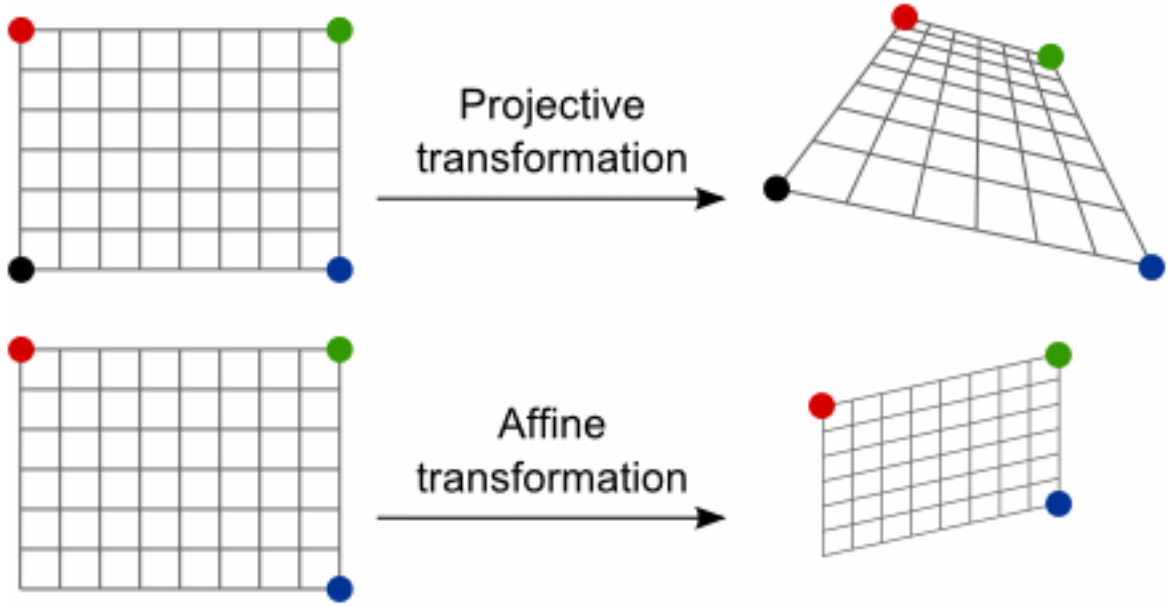
- Straight lines remain straight
- Parallel lines may converge towards ”vanishing points”





Difference between projective and Affine transformation

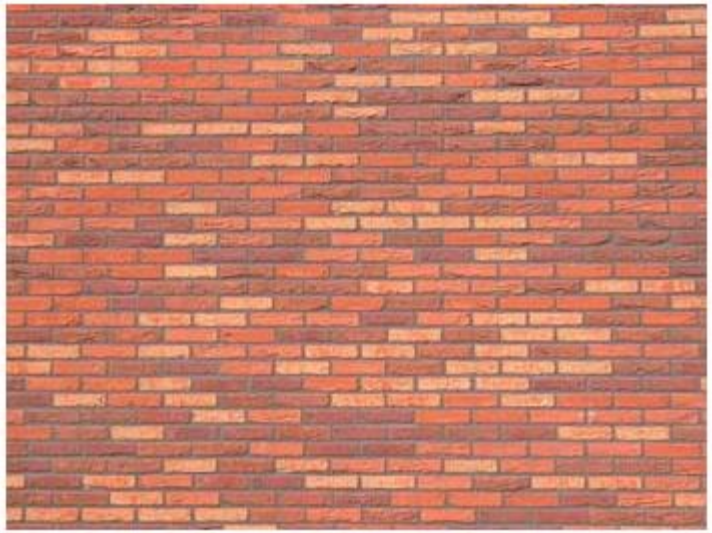
- Projective transformations do not preserve parallelism, length, and angle.
- Affine transformations, unlike the projective ones, preserve parallelism



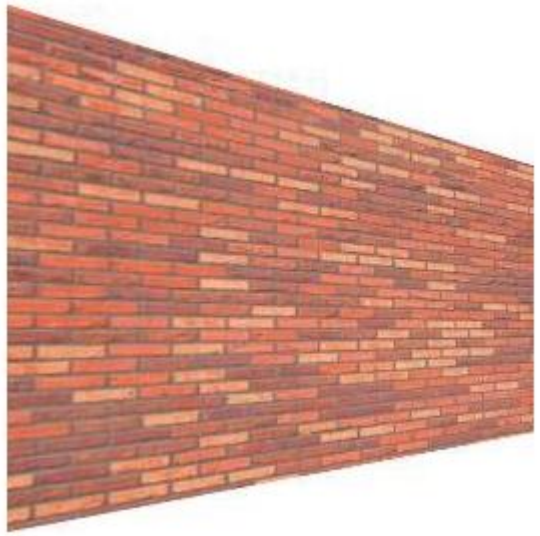
Quiz

Which image is using projective transformation?

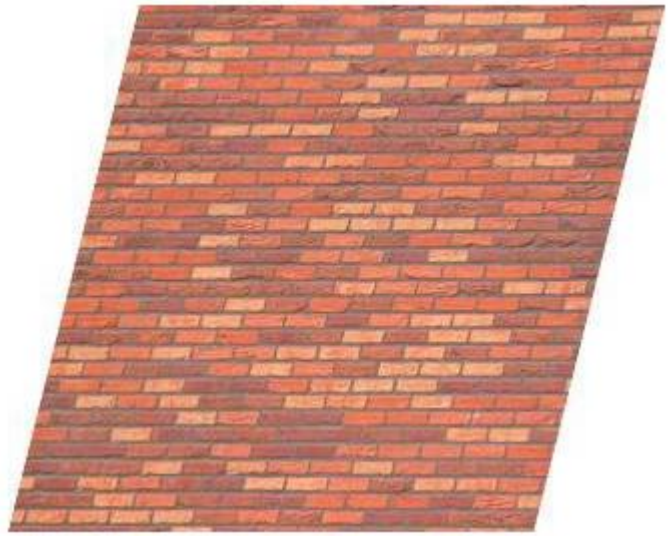
Original



1



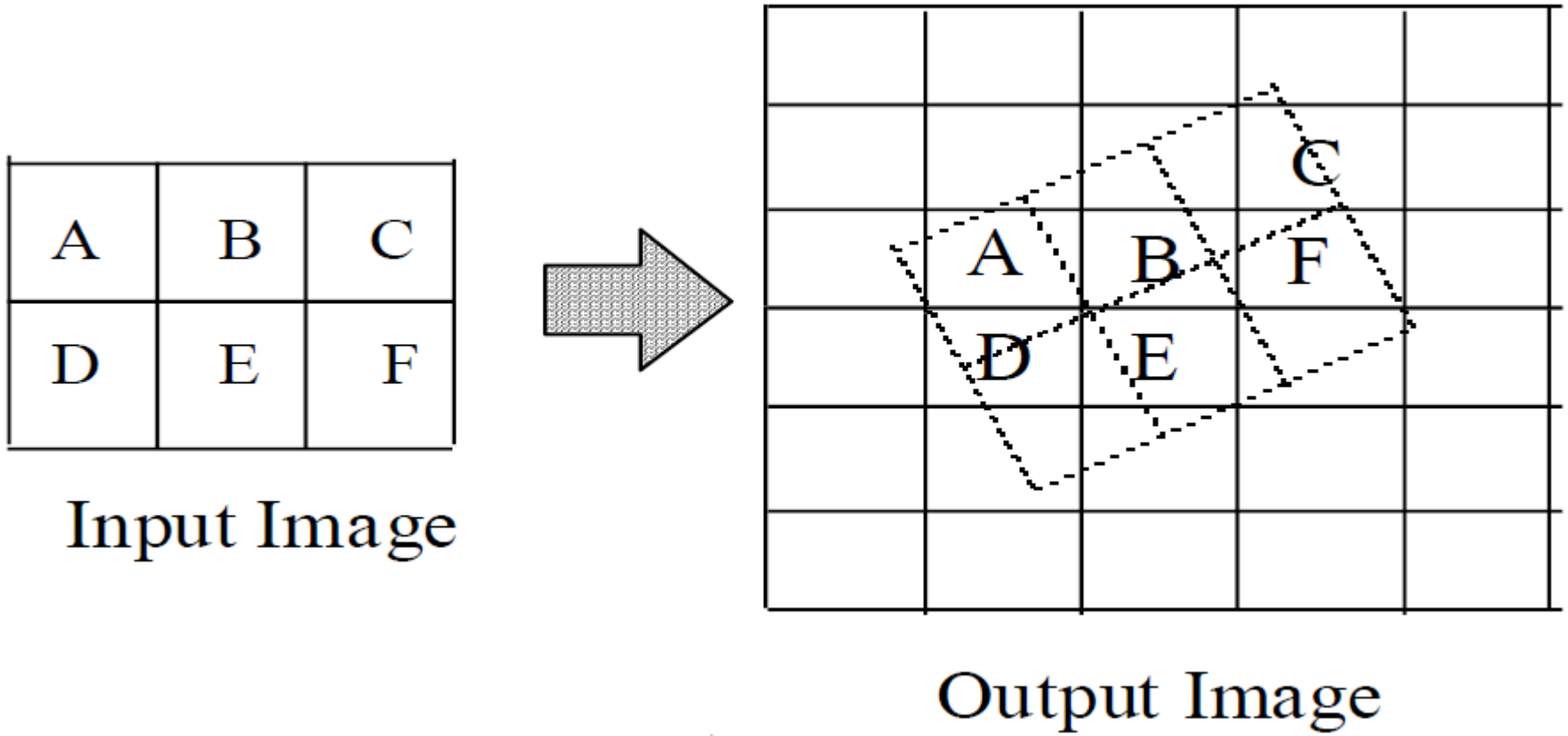
2



Interpolation

- When doing transformations, it is often not possible to map pixels 1 to 1.
- Hence, spatial transformations usually require some form of interpolation in addition to possible anti-aliasing.

- Example:



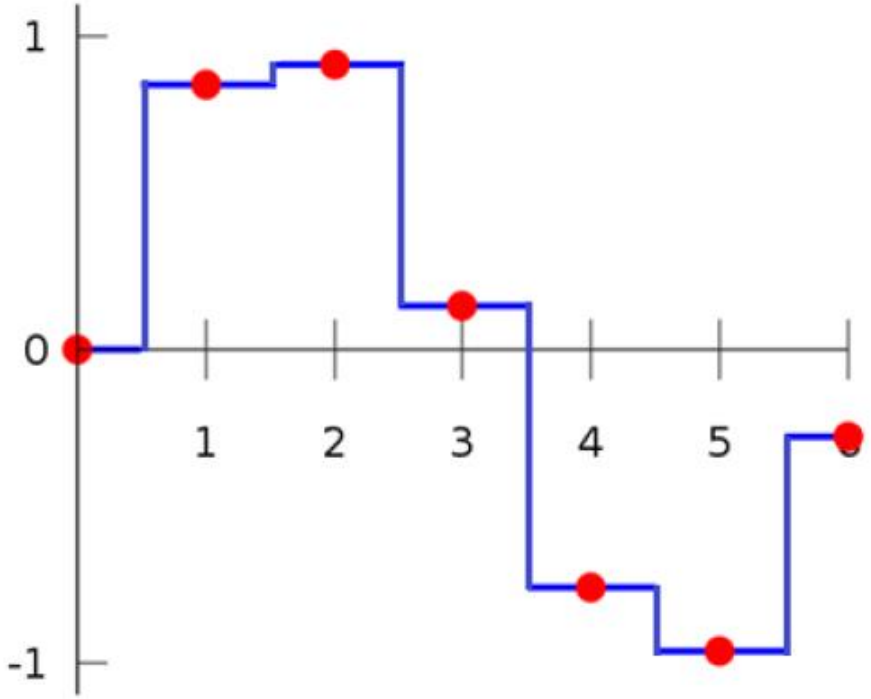
Interpolation

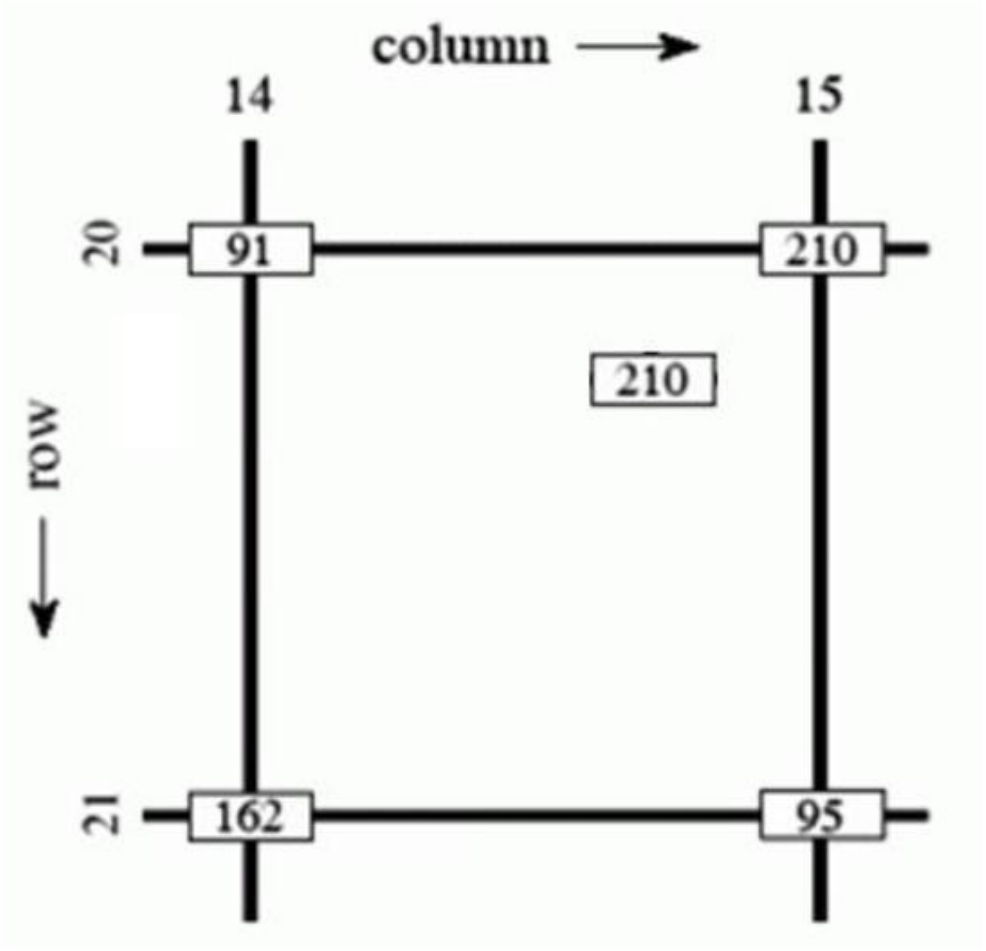
Interpolation methods:

- Nearest Neighbour: The output pixel is assigned the value of the closest pixel in the transformed image. An input pixel may fall into two or more output pixels
- Bilinear interpolation: The out put pixel is the weighted average of the transformed pixels in the nearest 2 by 2 neighborhood
- Bicubic interpolation: The weighted average is taken over a 4 by 4 neighborhood

- Since computation increases with the number of pixels that are considered, there is a tradeoff between quality and computational time

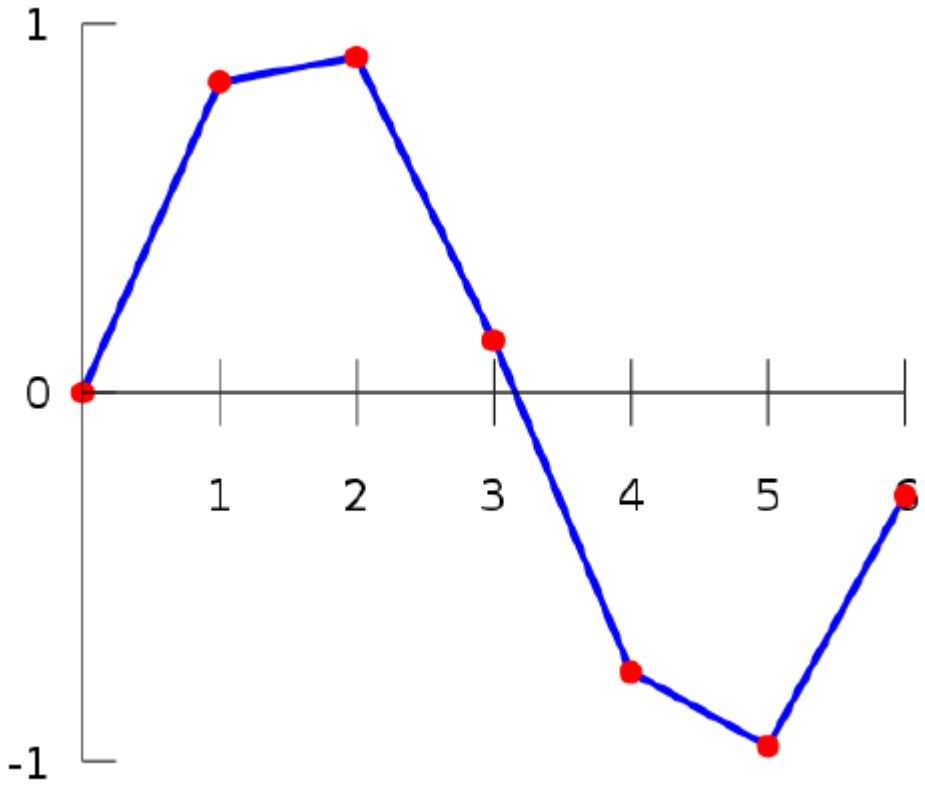
Nearest neighbour interpolation

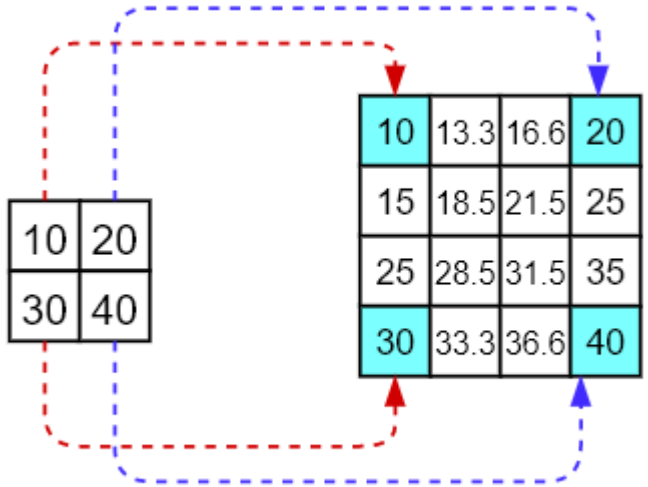


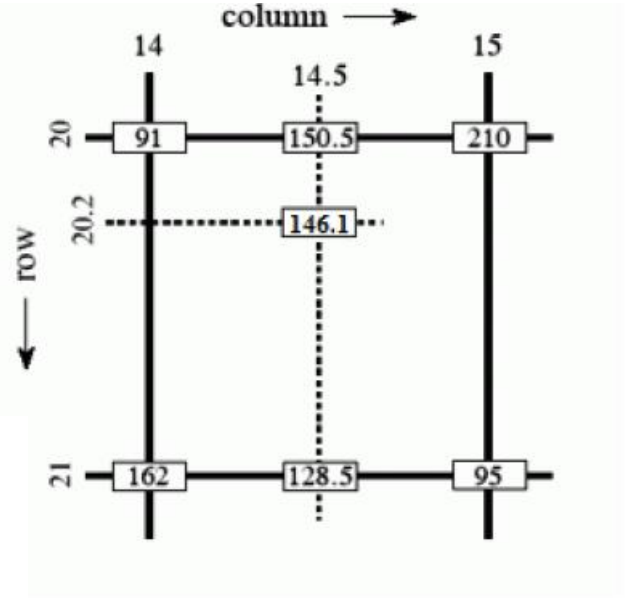


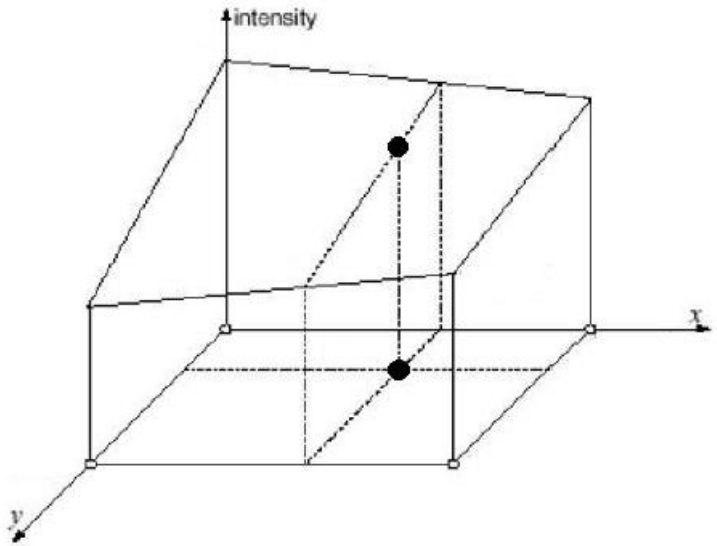
Bilinear interpolation

- In 2D takes 4 pixels (2×2) into account



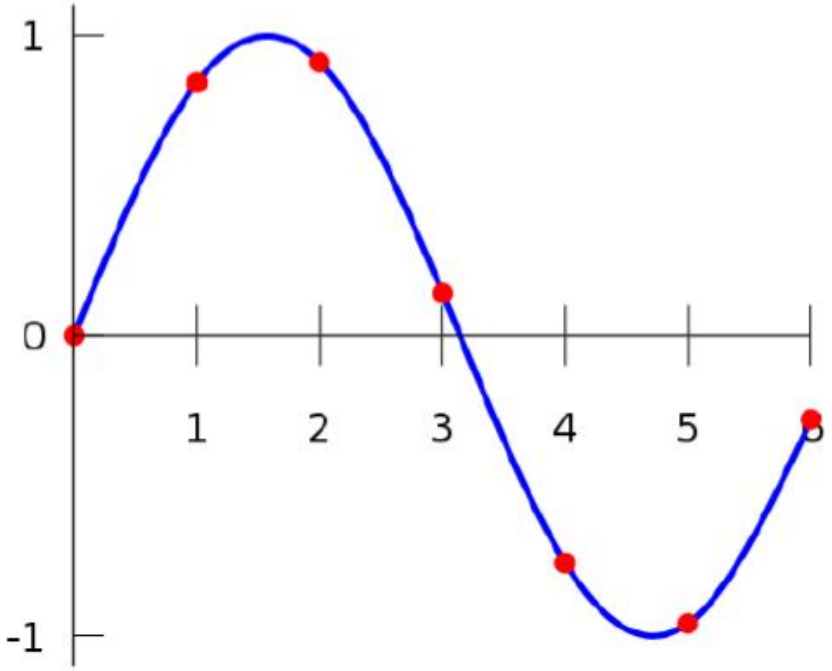


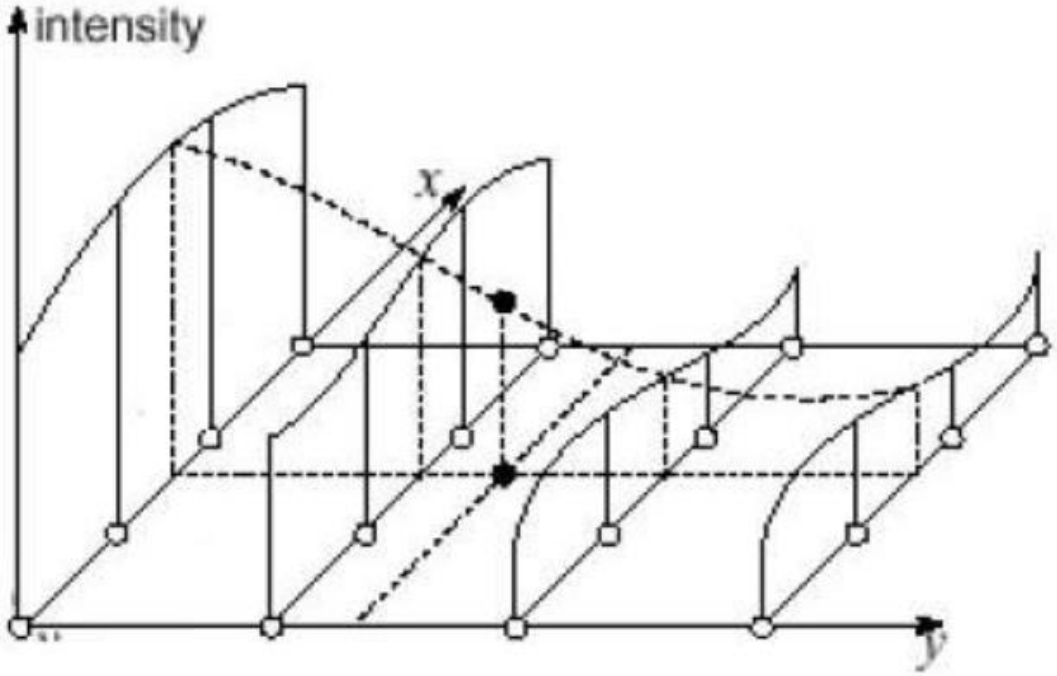




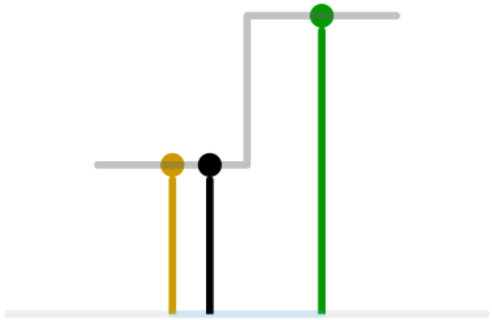
Bicubic interpolation

- In 2D, Bicubic interpolation considers 16 pixels (4×4) output pixels

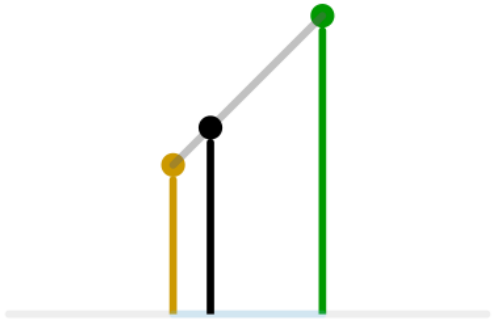




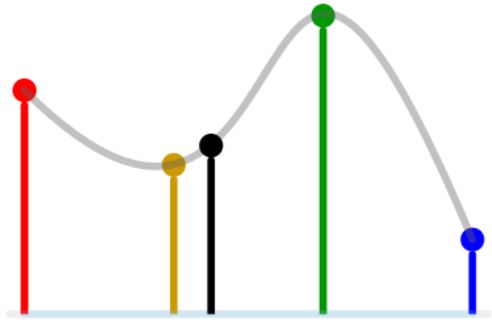
Summary interpolation



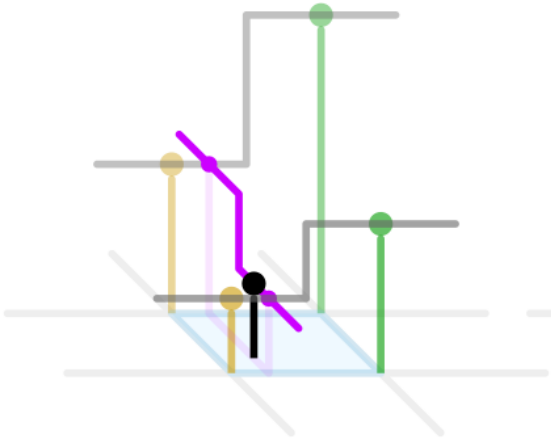
1D nearest-neighbour



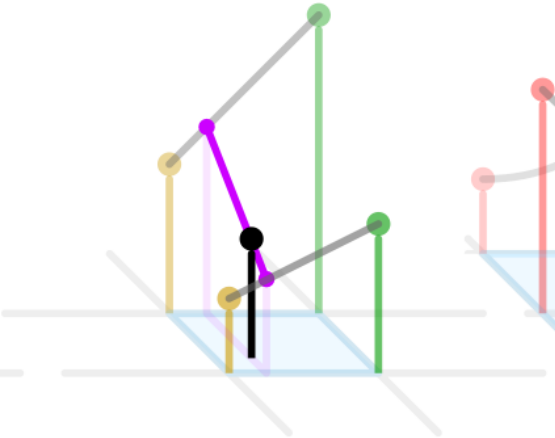
Linear



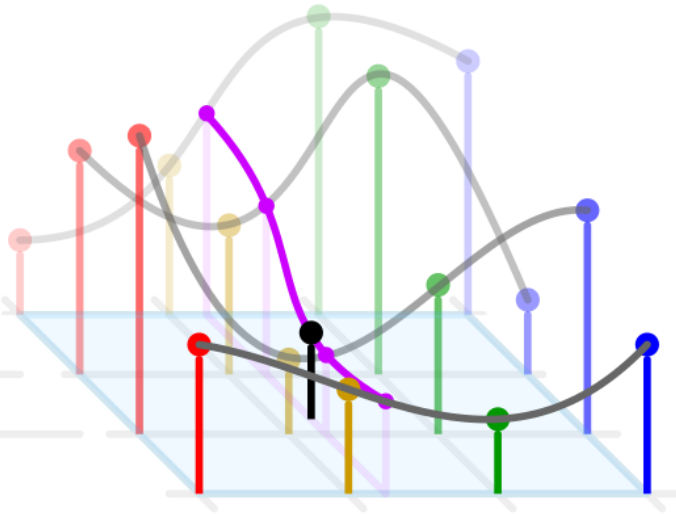
Cubic



2D nearest-neighbour



Bilinear

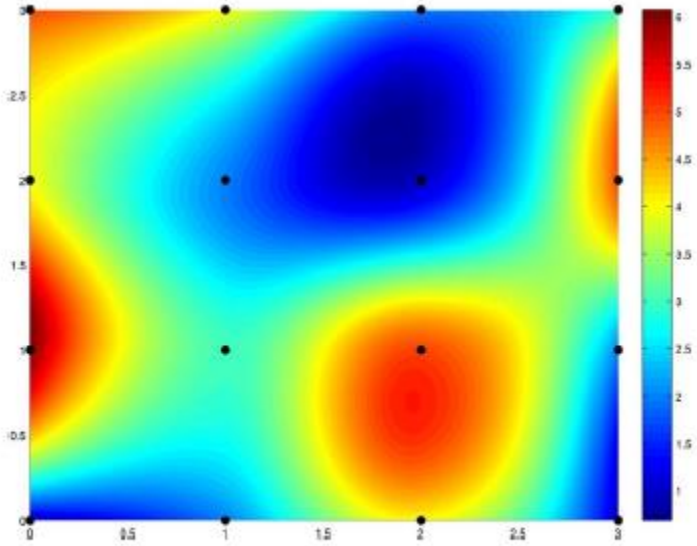


Bicubic

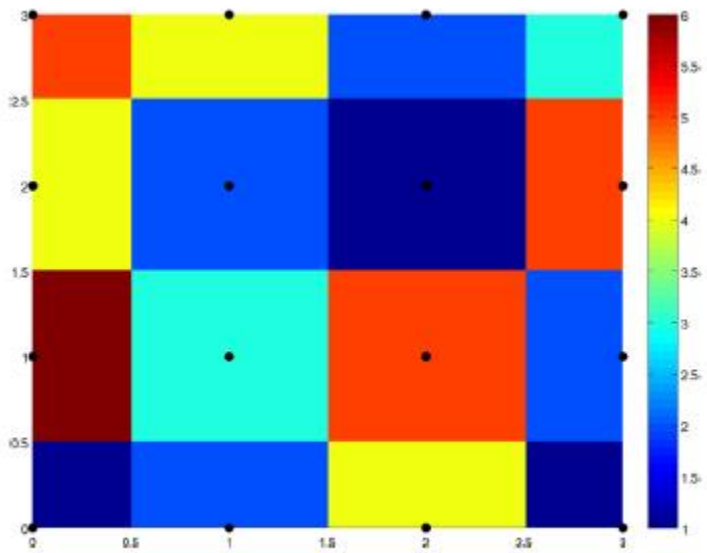
Quiz

Which interpolation methods were used to upscale each of the 3 images below?

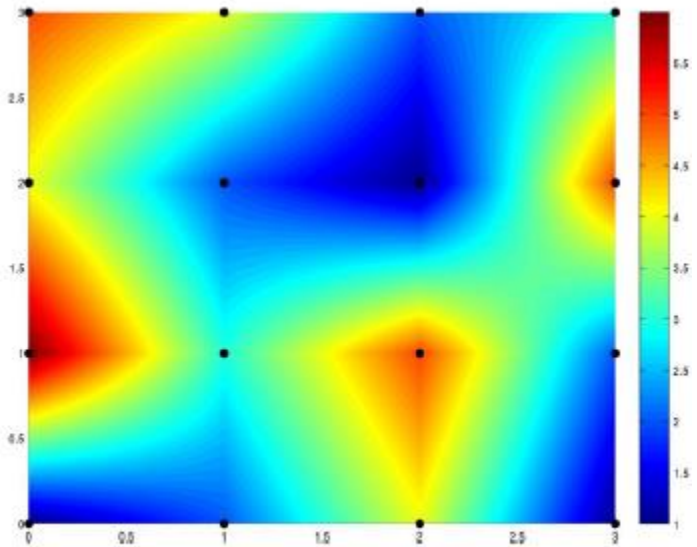
(Original size 4 x 4 px)



A



B



C

A: Nearest-neighbor interpolation
B: Bilinear interpolation
C: Bicubic interpolation

Image registration

- Image registration is the alignment of two or more images so they best superimpose.
- To achieve the best alignment, it may be necessary to transform the images
- Image registration can be quite challenging even when the images are very similar.
- Frequently the images to be aligned are not that similar, perhaps because they have been acquired using different modalities.
- The difficulty in accurately aligning images presents a significant challenge to image registration algorithms, so the task is often aided by a human intervention or the use of embedded markers for reference.

Image registration

- Approaches to image registration can be divided into two broad categories: **unassisted image registration**, and **interactive registration** where a human operator guides or aids the registration process.
- **Unassisted image registration,** relies on an optimization technique to maximize the correlation between the images.
- **Interactive registration,** uses human pattern recognition skills to aid the alignment process, usually by selecting corresponding reference points in the images.

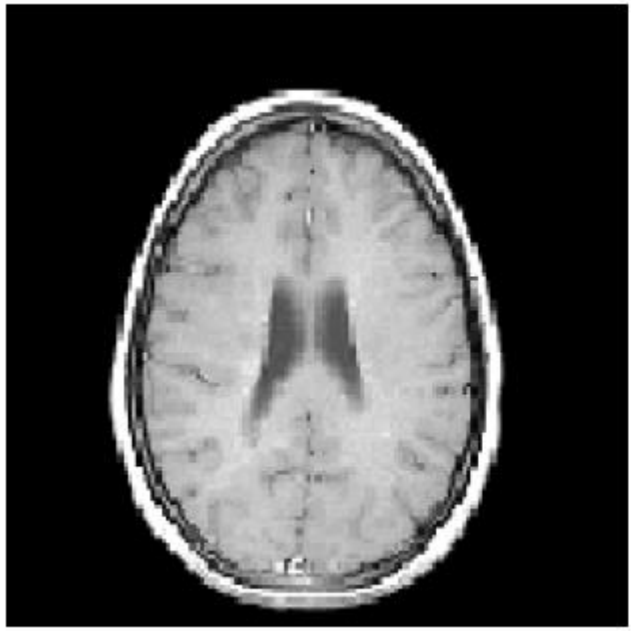
Unadied Image registration

- Unaided image registration involves the application of an optimization algorithm to maximize the correlation, or other measure of similarity, between the images.
- The appropriate transformation is applied to one of the images, termed the “input image,” and a comparison is made between this transformed image and the “reference image” (also termed the “base image”).
- The optimization routine seeks to vary the transformation until the comparison is the best possible.
- The problem with this approach is the same as with all optimization techniques: the optimization process may converge on a sub-optimal solution (a so-called “local maximum”), not the optimal solution (the “global maximum”).

Unadied Image registration

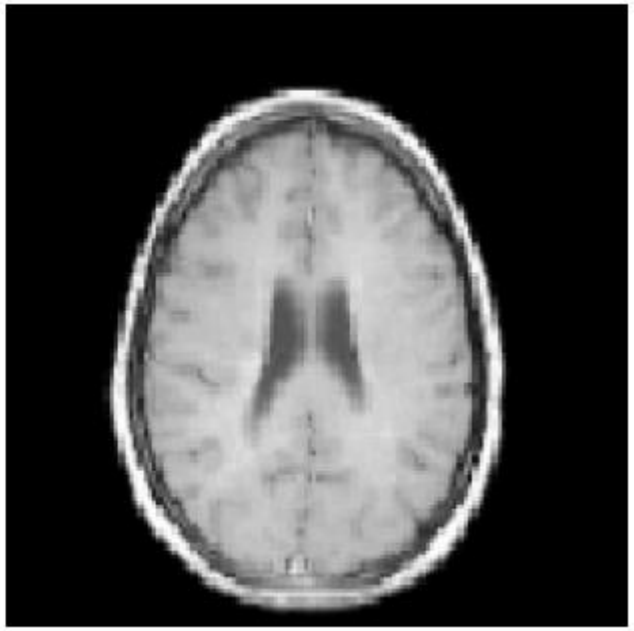
- The optimally realigned image has an alignment quite similar to the original image.

Reference Image



Input Image

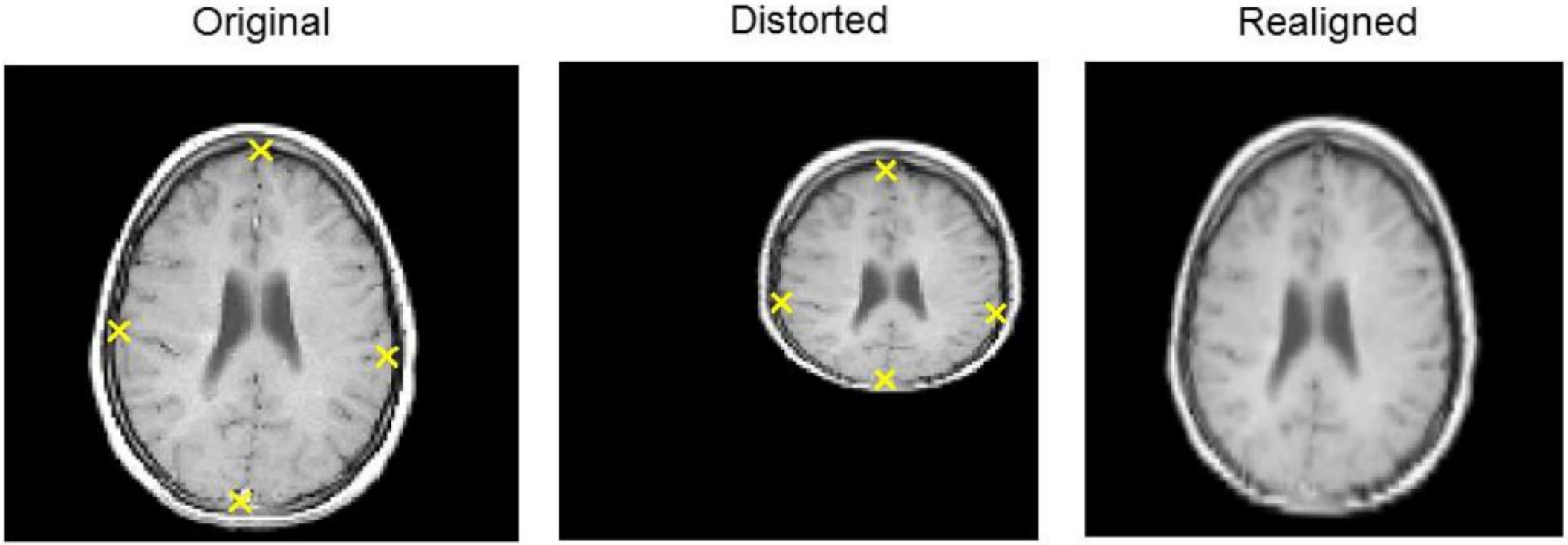
Aligned Image



Interactive Image registration

- The user interactively identifies a number of corresponding features in the reference and input image, and a transform is constructed from these pairs of reference points.
- The number of reference pairs required is the same as the number of variables needed to define a transformation: an affine transformation will require a minimum of three reference points while a projective transformation requires four variables.
- More reference points generally improve the alignment

Interactive Image registration



- Even with interactive alignment using four reference points, the correction is not perfect.
- The resultant transformation of the distorted center image is shown on the right and closely matches the original (correlation 95%). Some reduction in sharpness is seen in the realigned image as a result of information lost in the distortion process.