Medical Image Processing

Module 5

Rendering and Surface Models Registration

Syllabus

Rendering and Surface Models

- Visualization, orthogonal and perspective projection, and the viewpoint, ray casting, surface–based rendering

Registration:

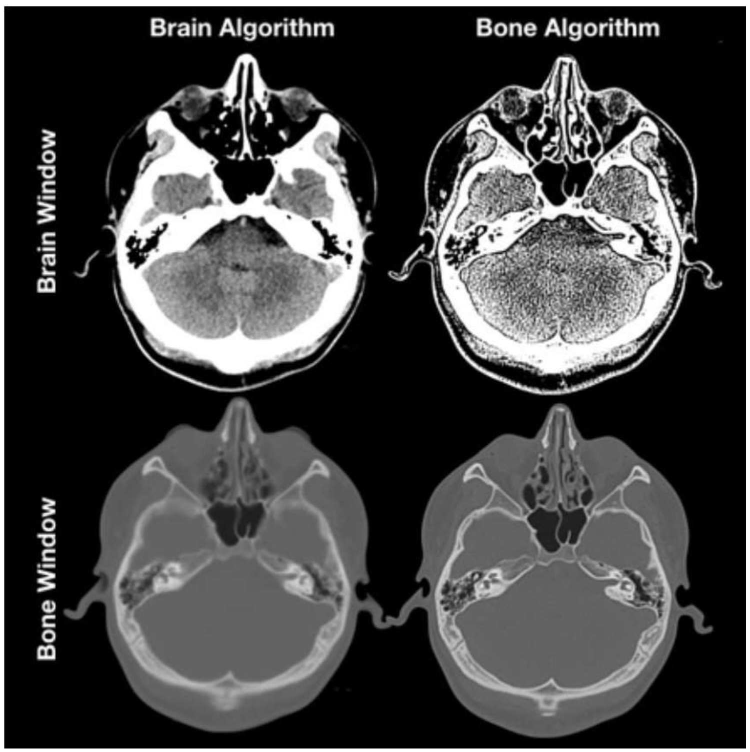
- Fusing information, registration paradigms, merit functions, optimization strategies, some general comments, camera calibration, registration to physical space

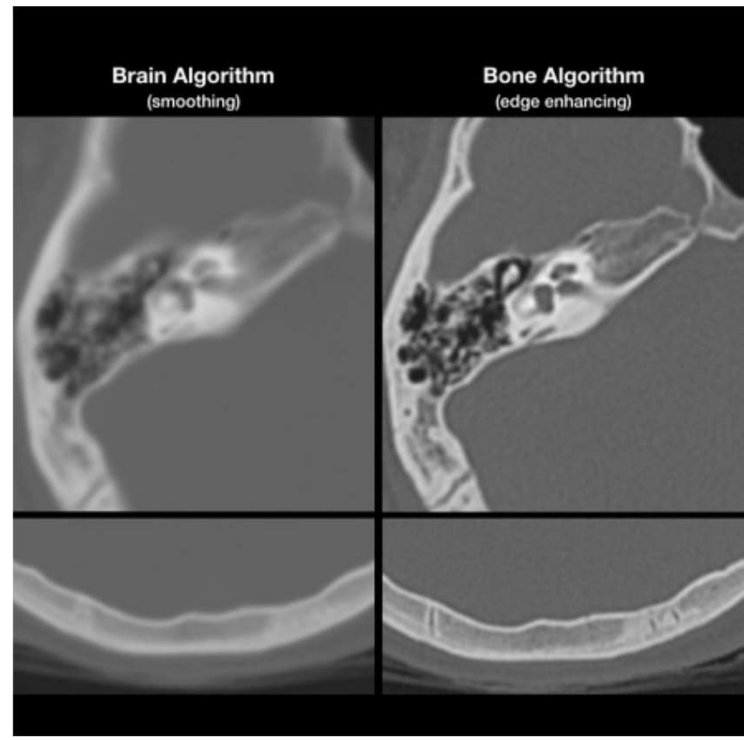
Visualization

- Visualization, generally speaking, is the art and science of conveying information to a human observer by means of an image;
- Visualization transforms data into interactive visual representations to facilitate data understanding and exploration through visual perception and human-computer interaction, and it can present medical images in 3D with high accuracy
- visualization techniques in medical imaging
  - windowing, which makes information hidden in the full intensity depth of the image visible, and
  - reformatting, which allows for exploring a data cube or volume in arbitrary directions.
- Rendering is the core technology for generating these visualizations.
- Medical images are typically generated as 2D projection images or sequences, as in radiography, or as stacks of 2D image slices, as in tomographic imaging.
- To use them for diagnostic or interventional purposes, the image data can be visualized as such, but they can also be shown as re-sliced images or as three dimensional (3D) images.
- Medical images are used not only for diagnostic purposes, but also often serve as the basis for a therapeutic or surgical intervention during which the instruments are guided by and navigate through the image content. Images can be obtained prior to and during surgery.

Visualization

- **Windowing**, also known as **grey-level mapping**, **contrast stretching**, **histogram modification** or **contrast enhancement** is the process in which the CT image greyscale component of an image is manipulated via the CT numbers; doing this will change the appearance of the picture to highlight particular structures. The brightness of the image is adjusted via the window level. The contrast is adjusted via the window width.
- **Window width:** The window width (WW) as the name suggests is the measure of the range of CT numbers that an image contains. A wider window width (2000 HU), therefore, will display a wider range of CT numbers. Consequently, the transition of dark to light structures will occur over a larger transition area to that of a narrow window width (<1000 HU). Accordingly, it is important to note, that a significantly wide window displaying all the CT numbers will result in different attenuations between soft tissues to become obscured





ORTHOGONAL AND PERSPECTIVE PROJECTION, AND THE VIEWPOINT

- The basic idea of rendering is to generate a 2D image from 3D data; every camera and every x-ray machine does this actually.
- The distance of the viewpoint (or x-ray focus) from the 3D scene to be imaged.
- For this purpose we need a matrix that gives us a projection, the projection operator P.
- If one assumes that the origin of all rays that project the 3D scene to the 2D imaging plane lies at infinity we get the following matrix:

$$P_\infty = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This projection images everything onto the x-y plane; the eyepoint is located at infinity in the direction of the z-axis. This is an orthogonal projection: all rays hitting the object are parallel.

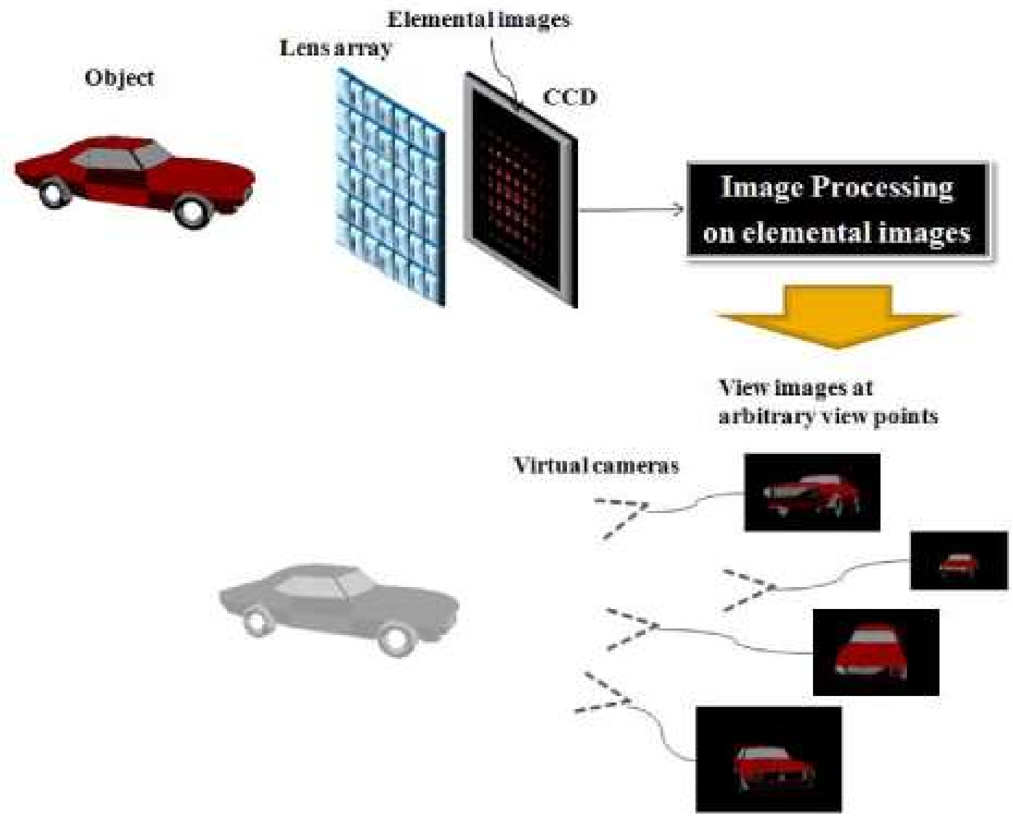
ORTHOGONAL AND PERSPECTIVE PROJECTION, AND THE VIEWPOINT

- Orthogonal projection, which is computationally more efficient than perspective rendering, is the common geometry for rendering.
- To be able to look at our object from different viewpoints.
- When rendering an object, we can do the same. We can either apply a volume transform V on every voxel of our object and apply the projection $PV\vec{x}$ or we can apply another volume transform V ′ to the projection operator: $V'P\vec{x}$.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{pmatrix}$$

where f is the distance of the eyepoint located on the z-axis. Remember that in homogeneous coordinates, the fourth component of the vector $\vec{x} = (x, y, z, 1)^T$ always has to be one. This is not the case when applying P as given in Equation 8.2. Therefore, the resulting vector $\vec{x}' = P\vec{x}$ has to be *renormalized* by an operation $\vec{x}'_{\text{renormalized}} = \vec{x}' * \frac{1}{x'_4}$ so that the fourth element of $\vec{x}'$ becomes one again.

ORTHOGONAL AND PERSPECTIVE PROJECTION, AND THE VIEWPOINT



ORTHOGONAL AND PERSPECTIVE PROJECTION, AND THE VIEWPOINT

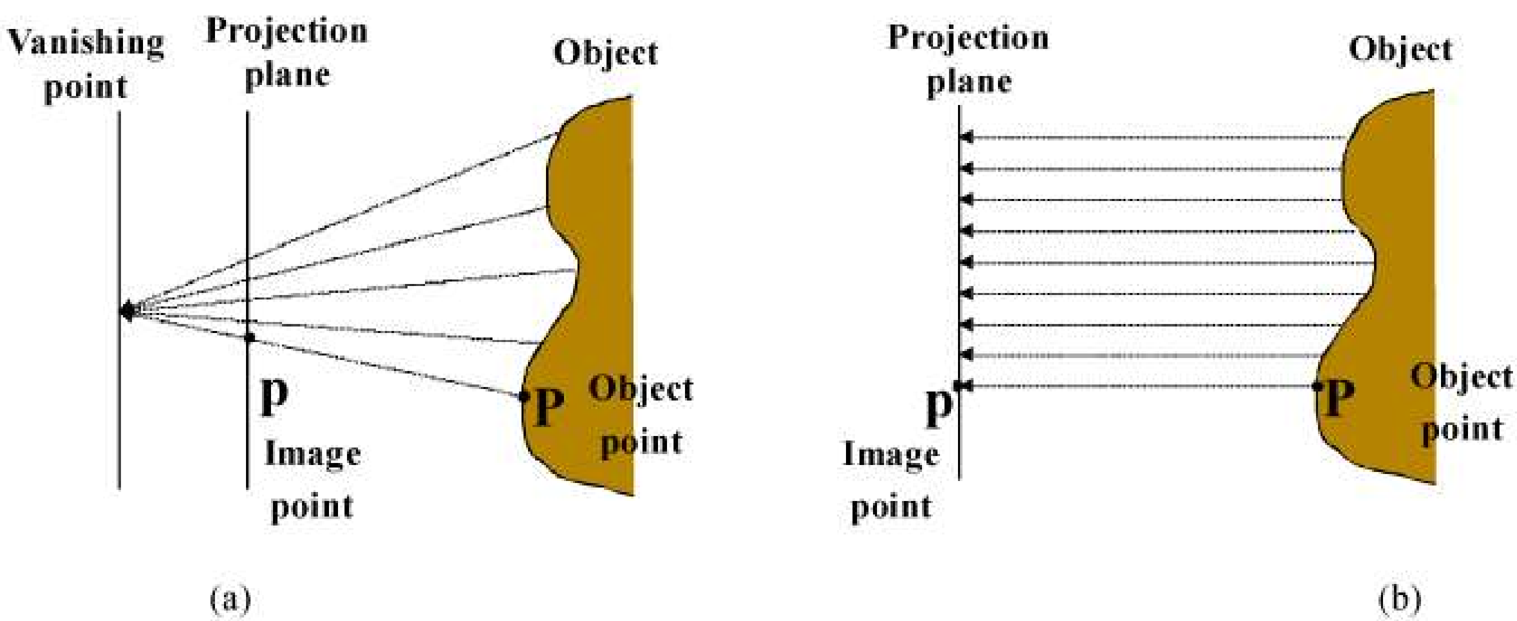
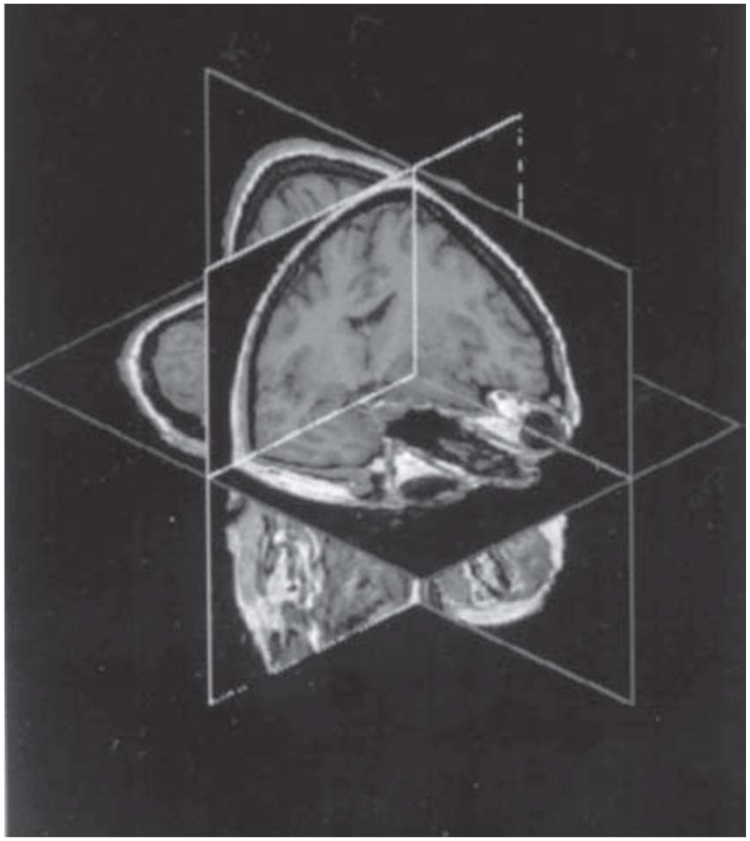
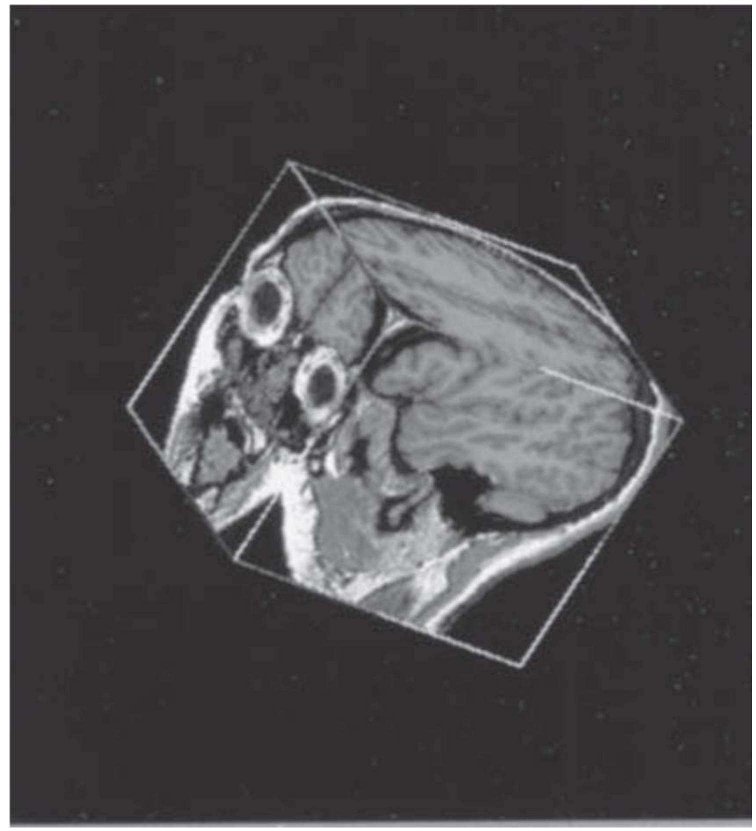


Fig. . Projection geometry (a). Perspective projection geometry (b). Orthographic projection geometry

ORTHOGONAL AND PERSPECTIVE PROJECTION, AND THE VIEWPOINT





Orthogonal sections of a three-dimensional volume image as
(i) intersecting orthogonal planes and (ii) a cubic volume

RAY CASTING

- X-rays emerge from the anode of the x-ray tube and pass through matter.
- In dependence of the density and radio opacity of the object being imaged, the x-ray is attenuated.
- The remaining intensity produces a signal on the detector.
- Ray casting can simulate this behavior if we sum up all gray values ρ in the path of the ray passing through a volume.
- This is, mathematically speaking, a line integral. If we want to simulate a camera, the situation is similar, but a different physical process takes place. Rather than penetrating the object, a ray of electromagnetic radiation is reflected.
- Simulate a ray that terminates when hitting a surface, and which is weakened and reflected; the amount of light hitting the image detector of a camera after reflection is defined by the object's surface properties. These properties are defined by lighting models, shading and textures.

RAY CASTING

- Ray casting is an image-driven technique – each pixel in the imaging plane is being assigned since it is the endpoint of a ray by definition.
- This is a huge advantage of ray casting since round-off artifacts cannot occur.
- Every pixel in the imaging plane has its own dedicated ray. If the voxels are large in comparison to the resolution of the image, discretization artifacts may nevertheless occur.
- In such a case, one can interpolate between voxels along the path of the ray by reducing the increment of the ray.

RAY CASTING

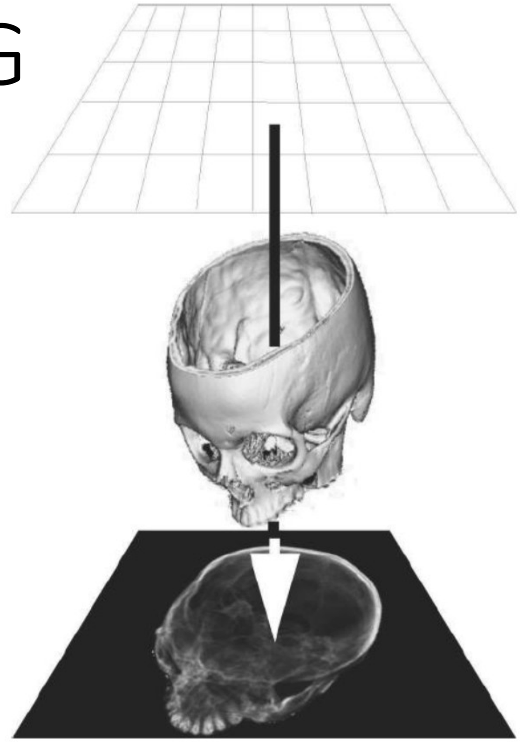


FIGURE: Orthogonal ray casting consists of projecting a beam emanating from a grid parallel to the image plane; the ray determines the appearance of the pixel it aims at on the imaging plane. If a volume rendering algorithm is implemented using ray casting, it integrates the voxel values using a defined transfer function. In the most simple case, this function is a simple max-function; only the most intense voxel is returned and conveyed to the pixel on the rendered image. This rendering technique is called maximum intensity projection (MIP).

RAY CASTING

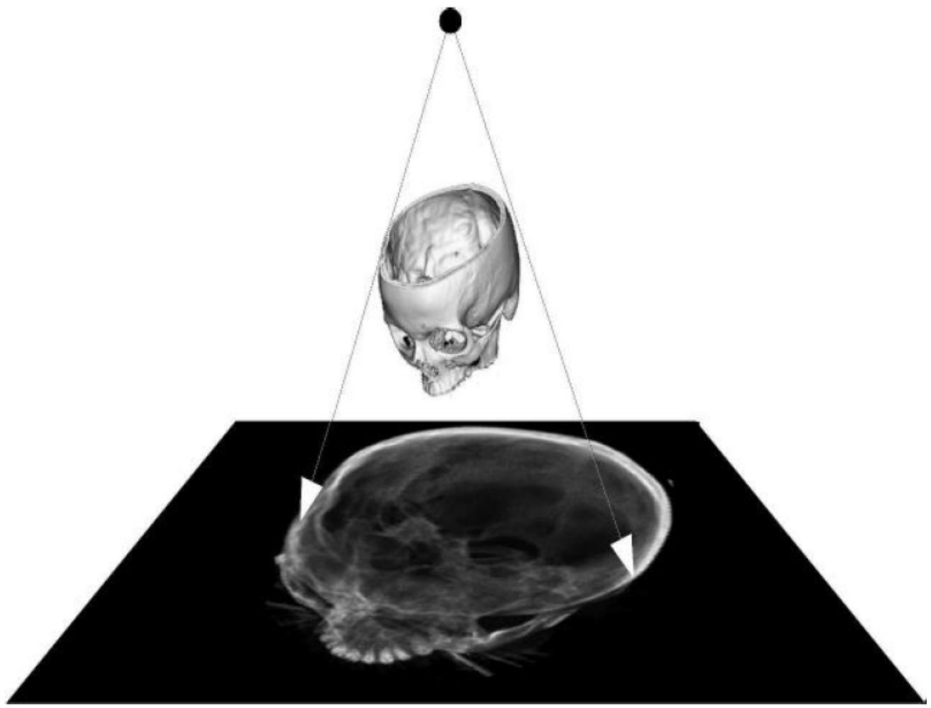
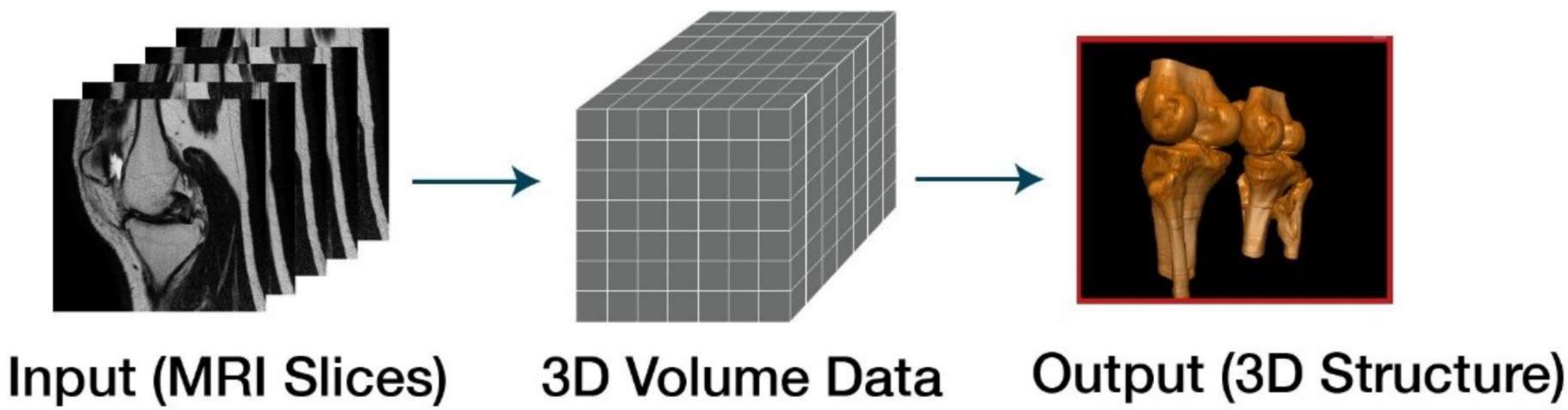


FIGURE: Ray casting can also be used for perspective rendering; here, all rays emanate from an eye point; the remaining steps are similar to orthogonal ray casting.

Visualization

- Visualization of three-dimensional biomedical images is typically performed by either surface rendering or volume rendering techniques.



Properties of rendering algorithms:

- Volume rendering: A ray that passes an object and changes its initial intensity or color during this passage is defined by a volume rendering algorithm. Such an algorithm does not know about surfaces but draws all of its information from the gray values in the volume. It is an intensity-based algorithm.

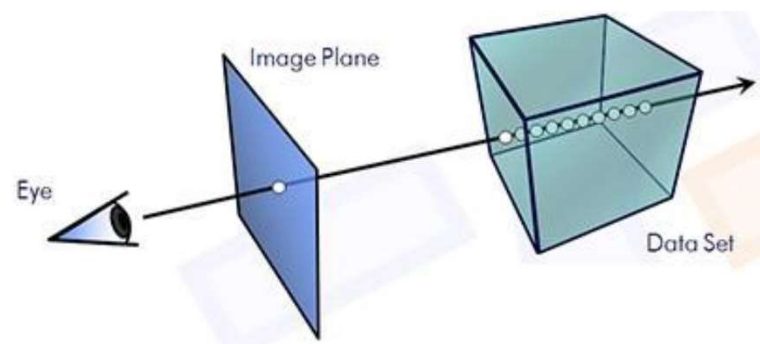
- Surface rendering: If we simulate a ray that terminates when hitting a surface, we are dealing with a surface rendering algorithm. The local gradient in the surrounding of the point where the ray hits the surface determines the shading of the corresponding pixel in the image plane. The drawback lies in the fact that a surface rendering algorithm requires segmentation.

Volume Rendering

- The term volume rendering comprises a set of techniques for rendering discrete three dimensional, i.e volumetric, data sets.
- With respect to medical imaging, such data can be acquired from different sources, such as computed tomography, magnetic resonance imaging, ultrasound, or positron emission tomography.
- The voxels along each ray are weighted according to

$$C_{\text{out}} = C_{\text{in}}(1 - \alpha(i)) + c(i)\alpha(i)$$

- where Cout is the value of the ray as it exits the ith voxel, and Cin is the value as it enters the ith voxel. There are two values associates with each voxel: c(i), a shading or luminance value, which can be based on the voxel value or calculated from a reflection model using the local gradient, and $\alpha(i)$ (i), an opacity derived from the tissue type



Volume Rendering

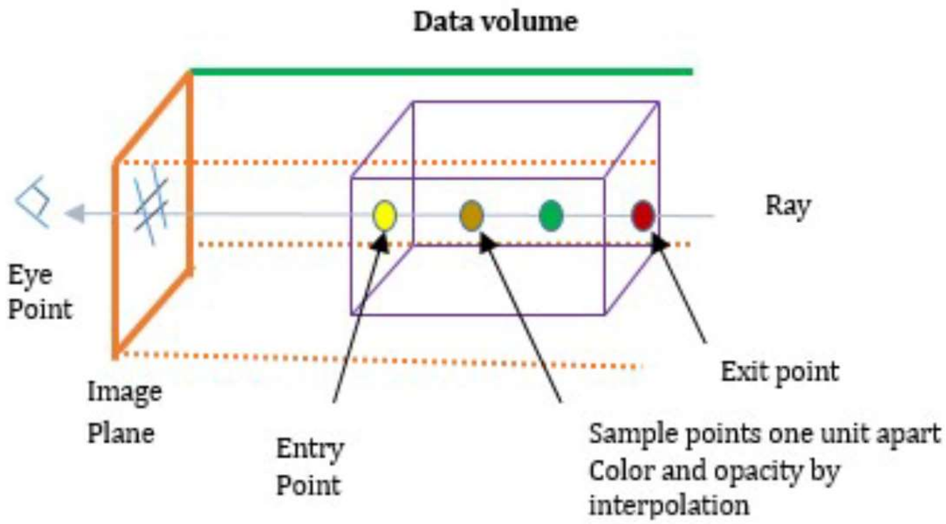
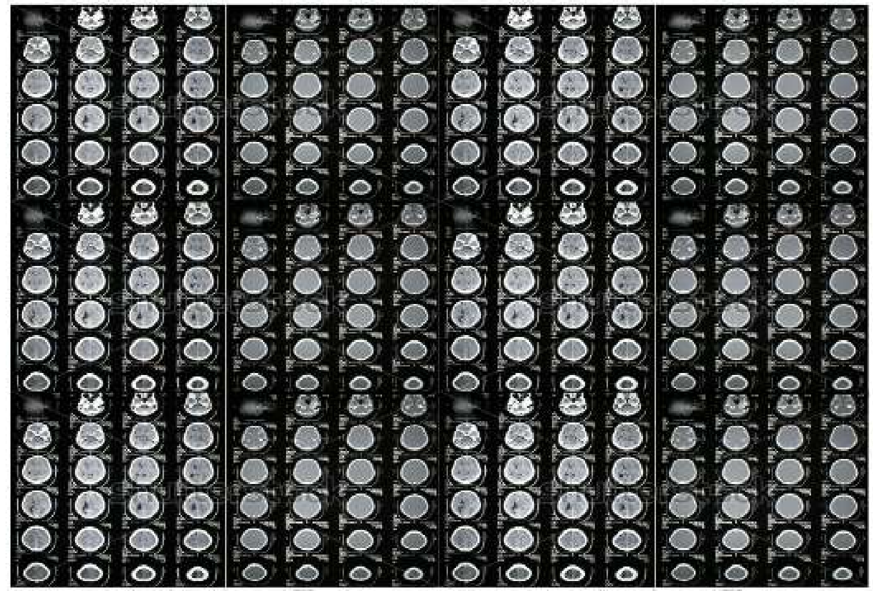
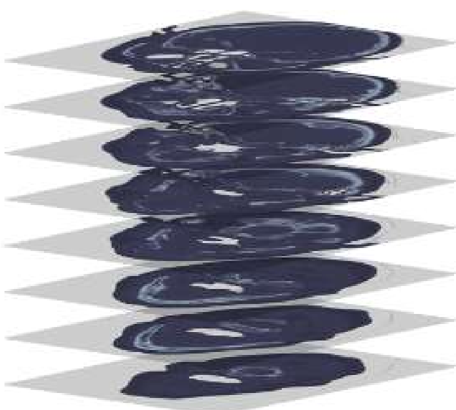


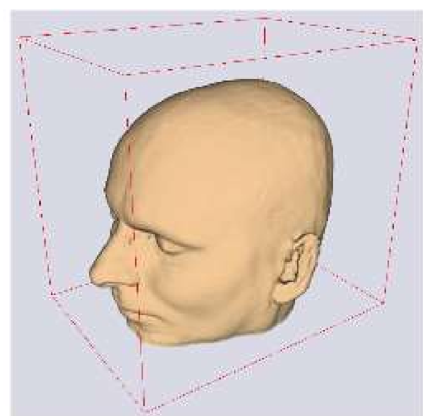
Figure 10: Casting the Rays and Taking Sample

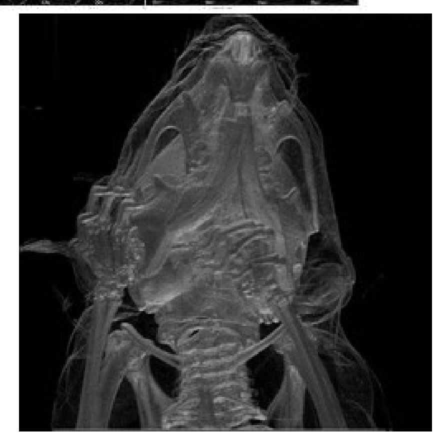
Volume Rendering

- Medical imaging is used for diagnosis as a replacement of surgical investigation.
- Advancements of technology had improved the quality of these scanners. Nowadays, these scanners can generate an accurate representation of anatomical structures of the body and can generate a large number of images for different slices with a very high resolution.
- It became more challenging to handle such amount of data produced by such machines.
- These images are stacked together to make a 3D volume of data.
- Volume rendering is the process of visualization of these 3D data in an interactive manner.









MIP, DRRs and volume rendering

- Let's stick to volume rendering first; if we simulate an x-ray tube and simplify the attenuation model in such a manner that we just project the most intense voxel, we are dealing with *maximum intensity projection* (MIP).
- Figure shows such a rendering.
- While the method sounds extremely simple, it is astonishingly efficient if we want to show high contrast detail in a volume.
- The appearance is somewhat similar to an x-ray with contrast agent.
- DRR: *Digitally rendered radiograph*, a perspective volume rendering from CT simulating an x-ray.

MIP, DRRs and volume rendering

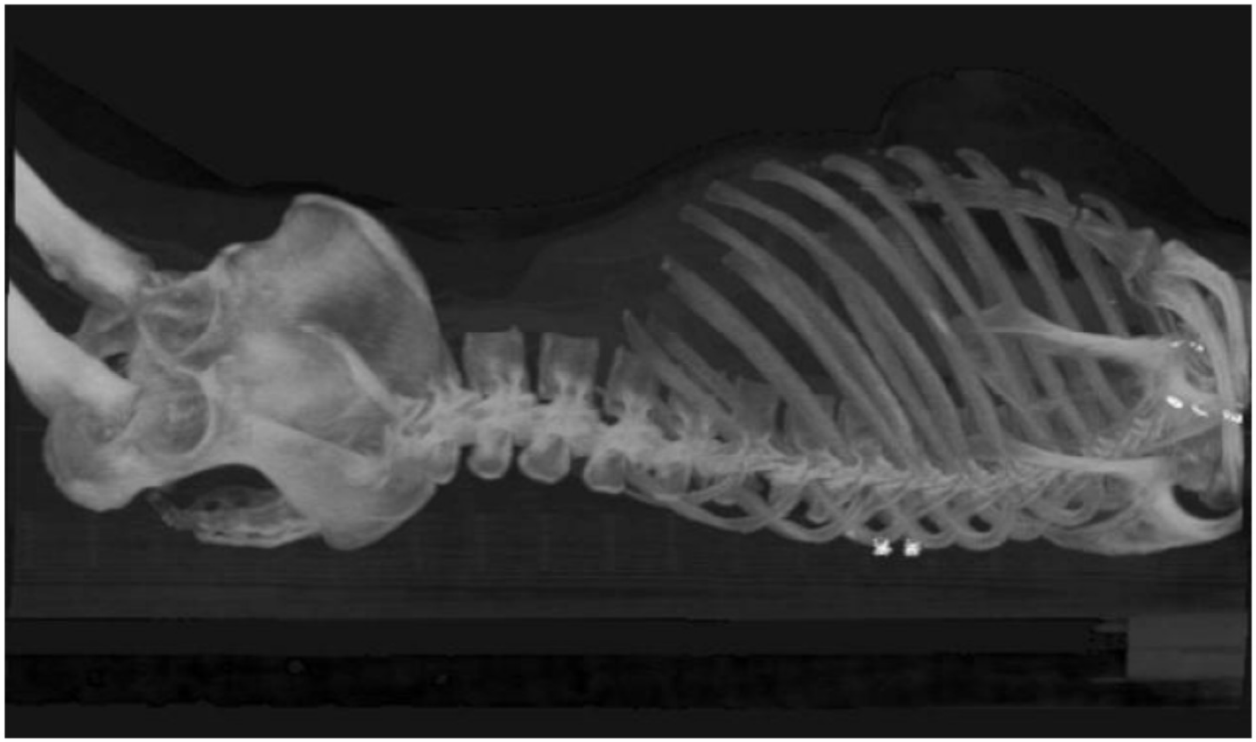
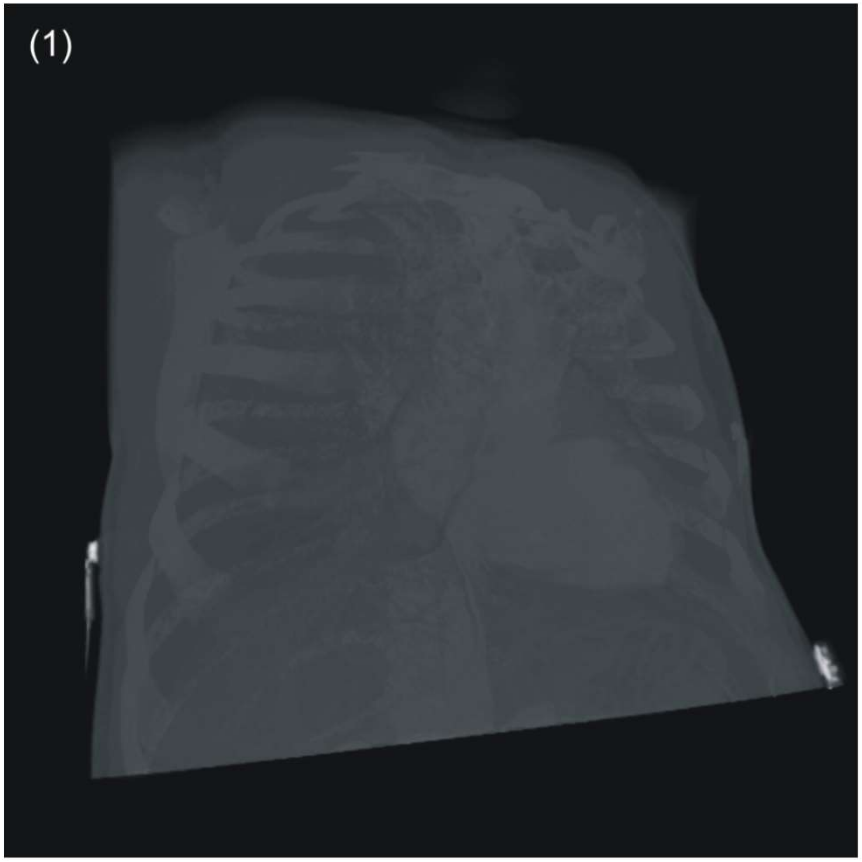


FIGURE: An orthogonal maximum intensity projection (MIP) from a whole-body CT, taken from a lateral viewpoint. In MIP, only the most intense voxel is projected to the imaging plane. MIP is the most primitive form of a volume rendering technique.

MIP, DRRs and volume rendering



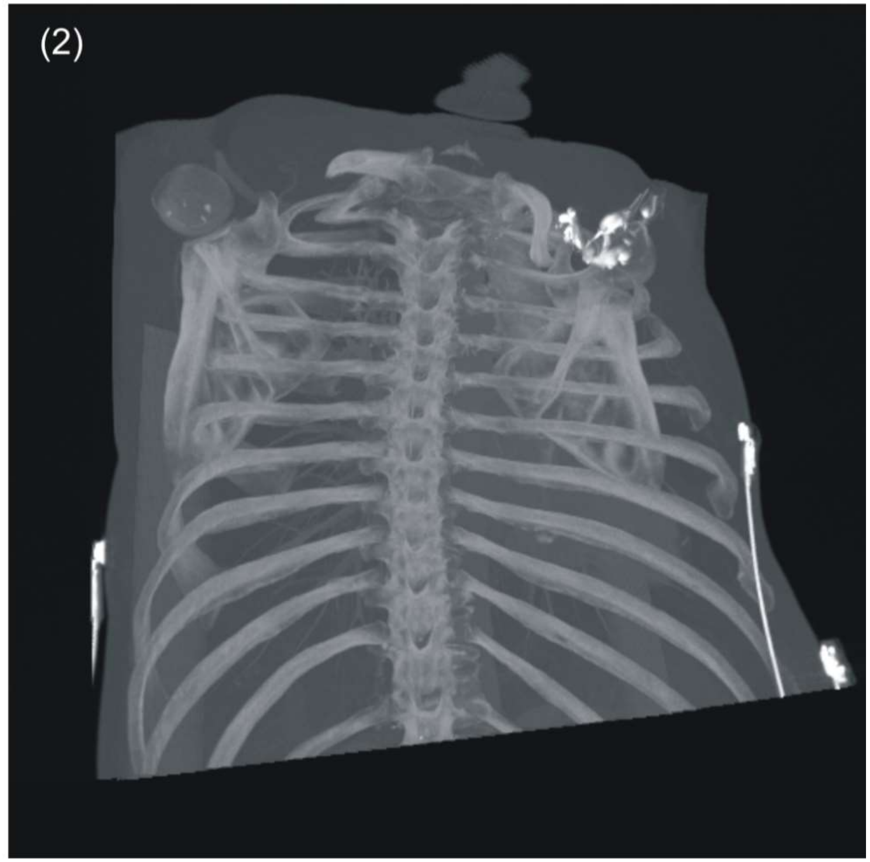
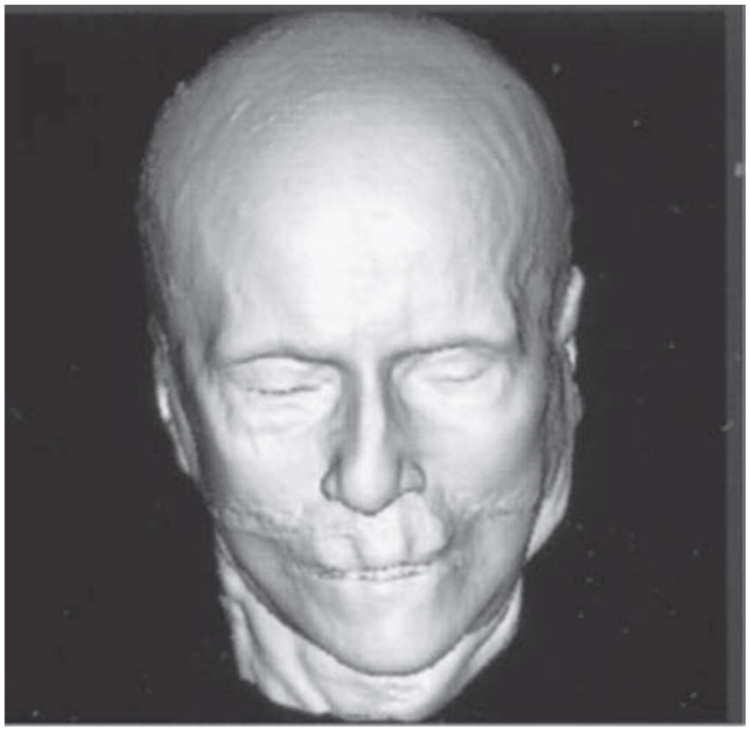


Figure : A comparison of direct volume rendering and Maximum Intensity Projection on the same data set

MIP, DRRs and volume rendering



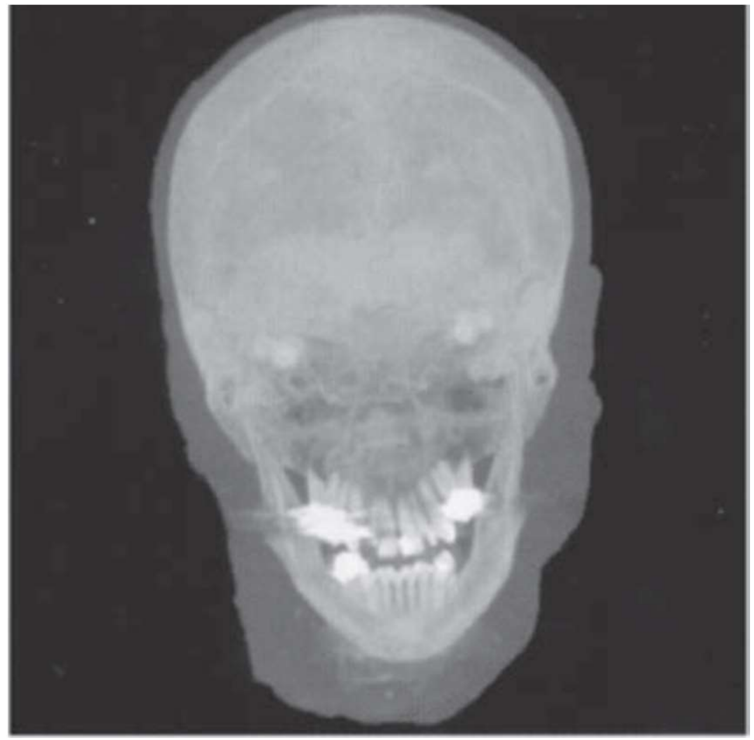


Figure (i) Volume-rendered image using the voxel gradient and (ii) maximum intensity projection image.

MIP, DRRs and volume rendering

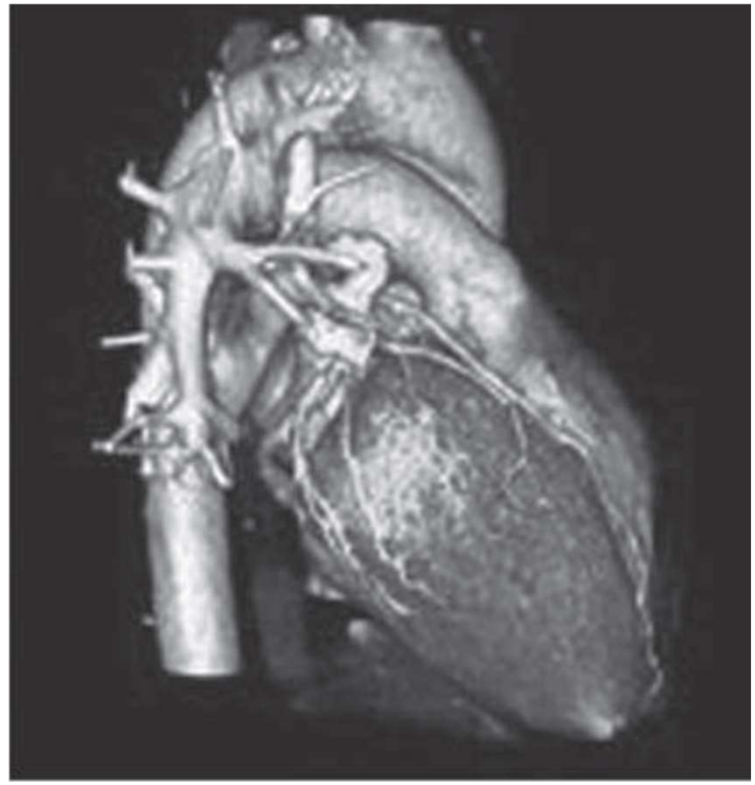


Figure: Volume-rendered CT images of (i) the skull and (ii) the heart

MIP, DRRs and volume rendering

- First, we load a volume that was scaled down and saved with 8 bit depth in order to keep the volume size manageable.
- Next, we introduce a minimum threshold th_{min} for rendering; remember that such a threshold helps to keep the rendered image free from unwanted projections of low-intensity non-zero pixels that may, for instance, stem form the air surrounding the object; a matrix holding the rendered image is allocated, as well as a vector vox.
- This vector holds the actual position of the ray passing through the volume:
- Now start the rendering process.
- The ray, whose actual position is given in the vox vector proceeds to the volume from a starting point above the volume, following the z-axis.
- If it hits the image surface, it terminates. A new gray value rho at the actual voxel position beyond the highest gray value encountered so far is assigned the new maximum gray value maxRho if its value is higher than the rendering threshold th_{min}. Once the ray terminates, the maximum gray value is assigned to the rendered image, and the image is displayed.

MIP, DRRs and volume rendering

- The resolution of the volume to be used is coarse, but it shows the basic principle of ray casting. In this implementation, a straight line parallel to the z-axis is drawn from each pixel on the image plane to the boundary of the volume.
- The most intense pixel in the path of this straight line is finally saved on the image plane.
- Another important component of rendering, which is intensity clipping.
- In order to keep the image tidy from a gray film that stems, for instance, from non-zero pixels of air surrounding the object, it is recommendable to introduce a minimum rendering threshold – if a voxel does not show high intensity, it is omitted.
- Intensity clipping by introducing a rendering threshold must not be mistaken for the segmentation method named thresholding, where a binary volume is constructed by omitting voxels below or above a given threshold.
- Nevertheless, we have already seen the effects of introducing a rendering threshold

MIP, DRRs and volume rendering

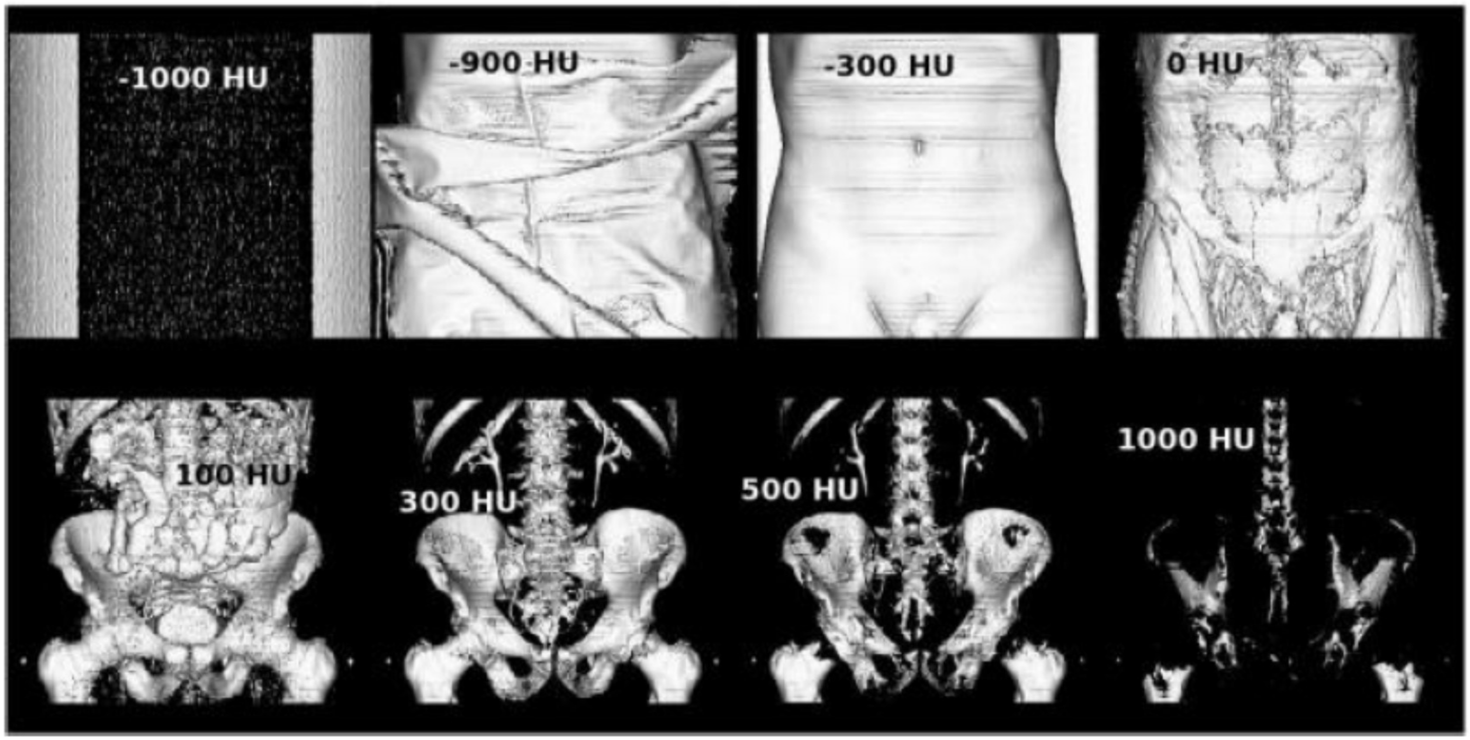


FIGURE: Eight rendered images of an abdominal CT-scan which underwent thresholding at different levels of the minimum image intensity. The first rendering was taken from the volume when setting the threshold to -1000 HU. The renderer displays the surface of the binary volume. Here, this is the shape of a cylinder. The threshold is so low that the air surrounding the patient in the CT was not removed.

MIP, DRRs and volume rendering

- If we simply sum up the voxels encountered by the ray, we end up with a render type that is called summed voxel rendering.
- Figure shows such a summed voxel rendering; it is not exactly a simulation of an x-ray image since the exponential attenuation of the x-ray is not taken into account. It is nevertheless a very good approximation, and besides 2D/3D registration, DRRs are widely used in clinical radiation oncology for the computation of so called simulator images.

MIP, DRRs and volume rendering

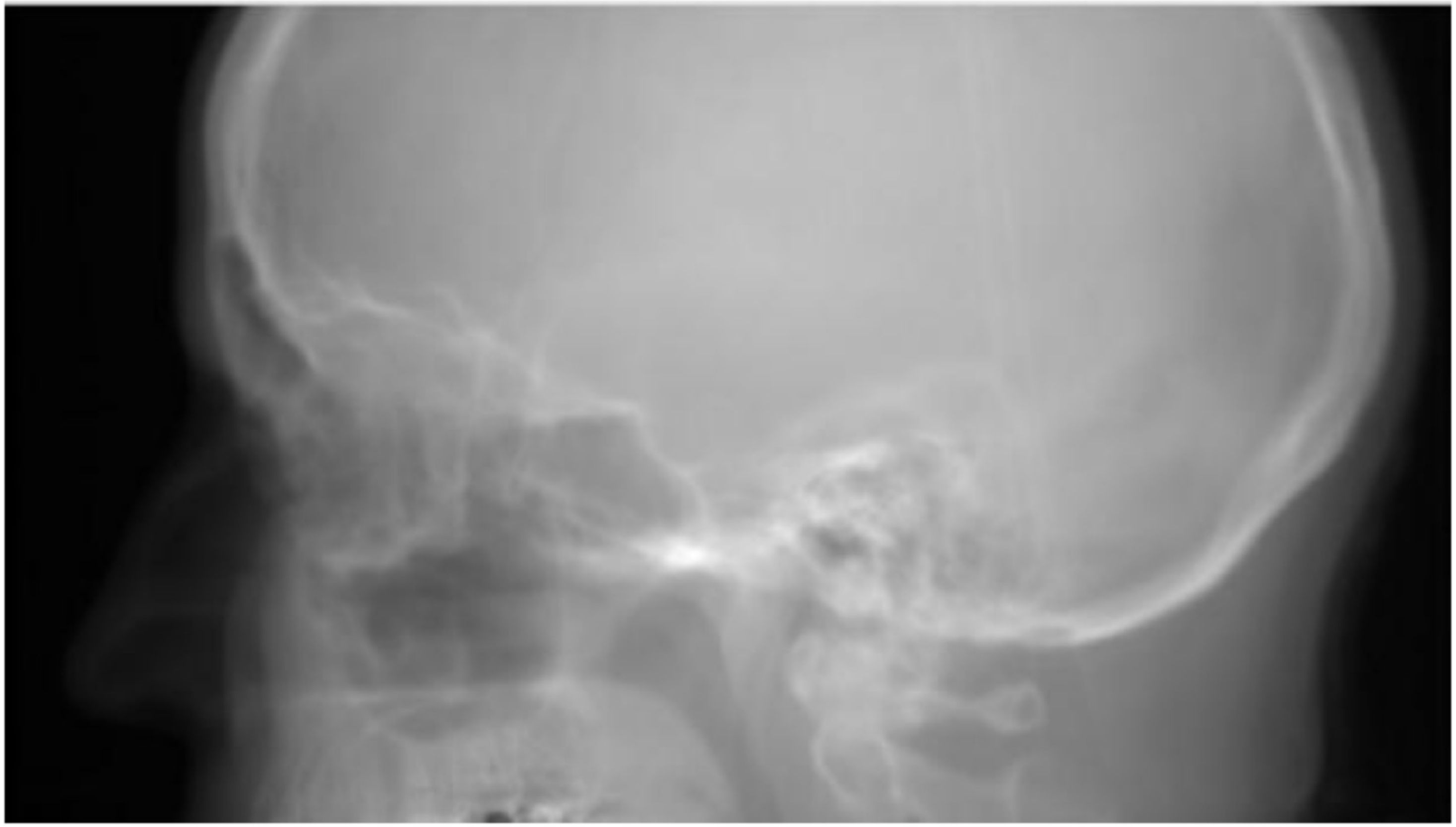


FIGURE: Summed voxel rendering is a simple sum of all voxel intensities in a ray’s path. This is a simplified model of x-ray imaging, where the attenuation of the ray’s energy is not taken into account. Being a volume rendering technique, summed voxel rendering also presents the simplest form of an algorithm for computing digitally rendered radiographs (DRRs)

MIP, DRRs and volume rendering

- The MIP and the DRR are simple volume rendering techniques; in volume rendering, a transfer function determines the final intensity of the rendered pixel.
- The two transfer functions we encountered so far are

$$\begin{aligned} \mathcal{T}_{\mathrm{MIP}} &= \max(\rho) \quad \forall \rho \in \{\vec{x}\} \\ \mathcal{T}_{\mathrm{DRR}} &= \sum \rho \quad \forall \rho \in \{\vec{x}\} \end{aligned}$$

- where ρ is the intensity of voxels located at positions $\{\vec{x}\}$, which is the set of all voxels within the path of the ray. The strength of volume rendering lies in the fact that segmentation is not necessary.
- However, a MIP or a DRR is usually not what one expects from a visualization algorithm for 3D data.
- If one wants to show, for instance, the surface of the body and some internal organs, more sophisticated volume rendering approaches are necessary.

MIP, DRRs and volume rendering

- A transfer function that shows both high-intensity contrast from structures inside the body as well as the surface could look something like this:
- Assign a high value of a given color to the first voxel above the rendering threshold encountered by the ray; one could also introduce some sort of shading here in such a manner that voxels with a greater distance to the origin of the ray appear darker. This very simple type of surface shading is called depth shading. However, the ray is not terminated (or clipped) here.
- If the ray encounters a structure with a gray value within a defined section of the histogram, it may add an additional color and voxel opacity to the pixel to be rendered.
- The definition of ROIs for rendering takes place by choosing the area of the histogram that gets a color assigned.
- A sophisticated volume rendering algorithm creates beautiful images, similar to old anatomical glass models, and circumvents the problems of segmentation. An example of a rendering generated by such a volume compositing technique can be found in Figure

MIP, DRRs and volume rendering

FIGURE: If one chooses a transfer function that assigns opacity and color to gray values, a more sophisticated type of summed voxel rendering is possible. A color version of this image rendered by volume compositing can be found in the JPGs folder on the accompanying CD. This image was created using AnalyzeAVW.

MIP, DRRs and volume rendering

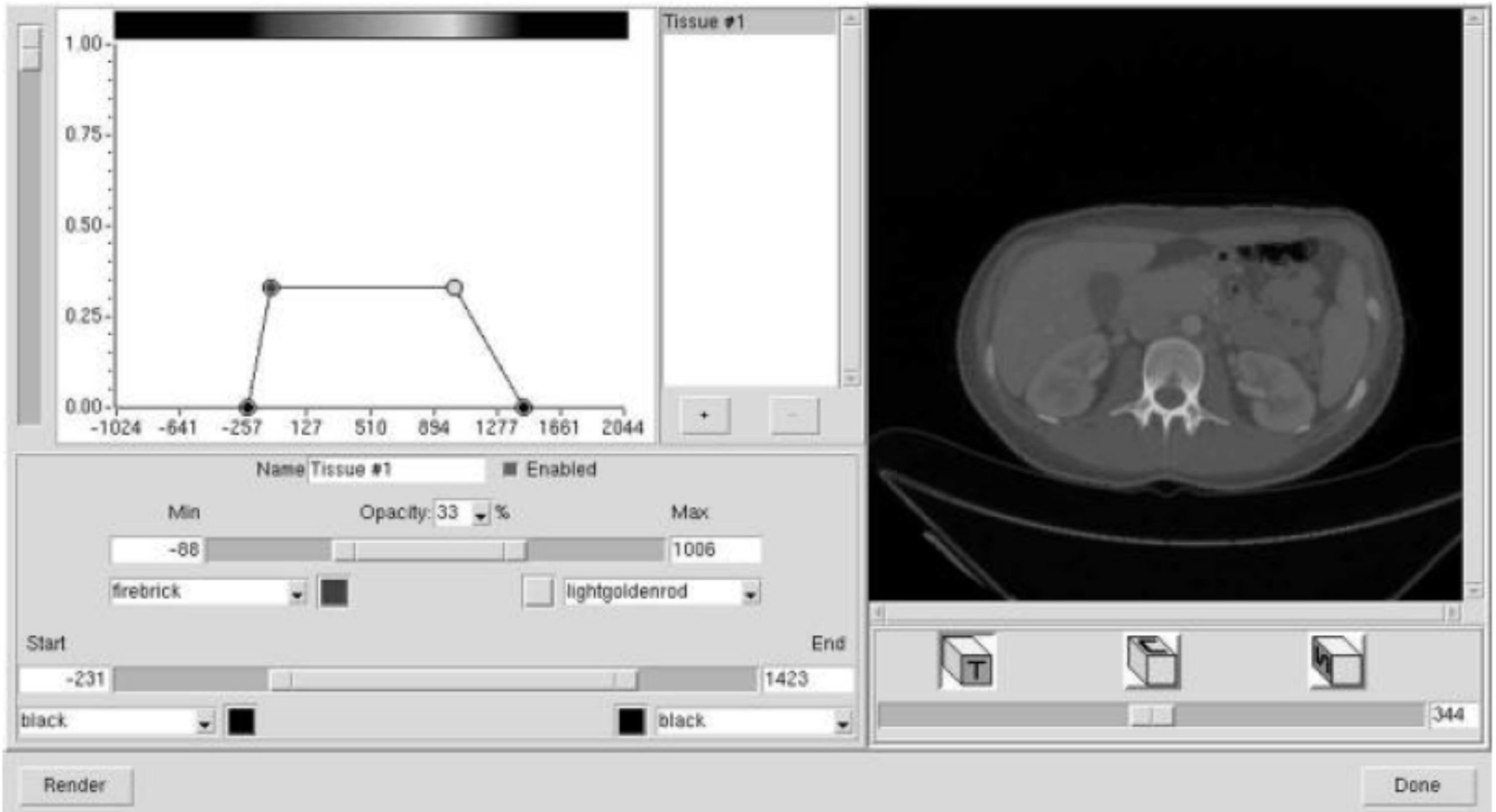


FIGURE: The dialog of AnalyzeAVW for defining a transfer function for the volume compositing process. Opacity and color assigned to voxel gray values are given by manipulating the transfer function.

MIP, DRRs and volume rendering

- Finally, we have to discuss the choice of the viewpoint in the rendering process.
- Changing he view of an object is easy; all we have to do is apply a spatial transform to the volume.
- The only problem here is the considerable confusion if the object leaves the rendering domain – it is hard to debug a code that is formally correct but renders nothing but a void volume.
- The more complex operation is the change of the viewpoint.
- If we transform the position of the observer, we also have to transform the imaging plane accordingly, otherwise the image will be skewed.

Other rendering techniques

- The main problem of ray casting lies in the fact that it is computationally expensive, especially when switching to a perspective view model.
- Various refinements to improve the performance of ray casting do exist, for instance shear-warp rendering.
- But we can also use our old pal, the projection operator P to render images. This method is called splat rendering.
- It is a volume driven rendering method, and implementing a MIP or summed voxel rendering is pretty straightforward – all we have to do is apply the projection operator to the voxel positions and map the gray values to the appropriate pixel positions in the rendering plane.

Other rendering techniques

From figure, we only render the bony part of a volume that has approximately $4.5 * 10^7$ voxels. Approximately 15% belong to bony tissue.

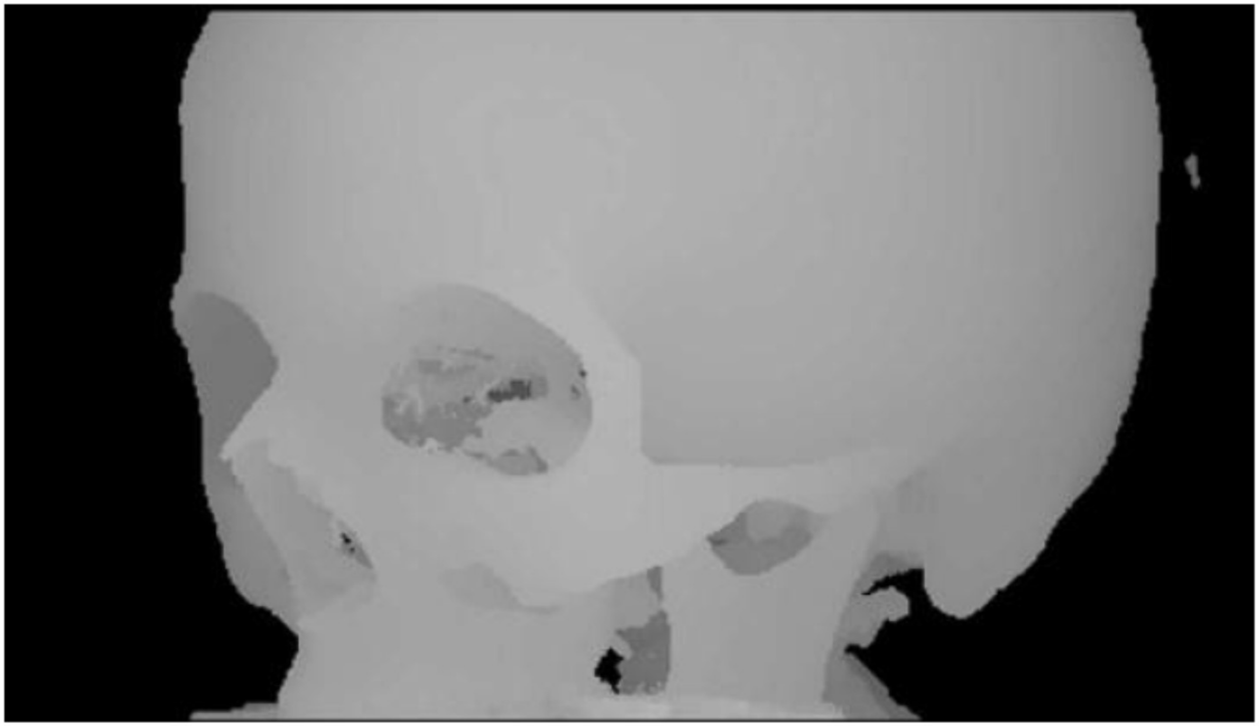


FIGURE: Depth shading is a simple shading technique that assigns a gray value to the pixel to be rendered based upon its distance from the origin of the ray. The further away a voxel lies from the imaging plane, the brighter it gets. Depth shading does not give a photorealistic impression, but it adds a visual clue and can be used if simple structures are to be rendered. This image was created using AnalyzeAVW.

Other rendering techniques

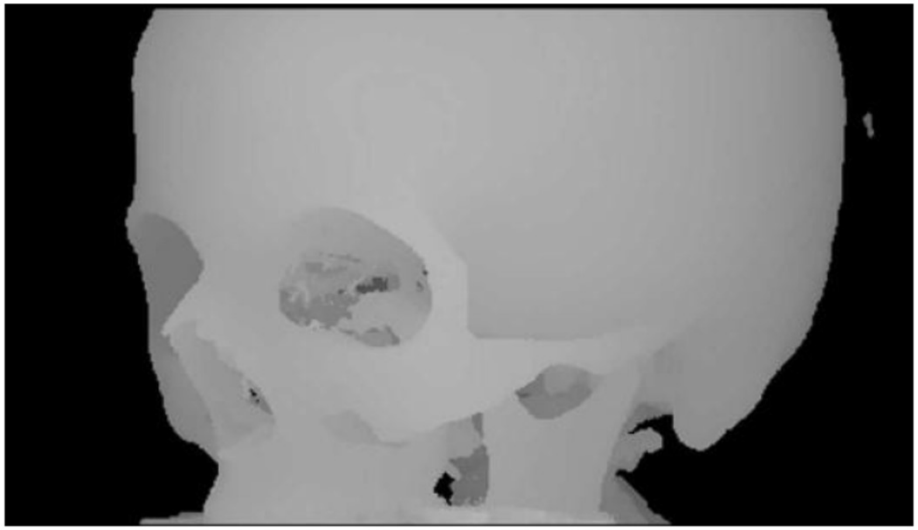


FIGURE: Depth shading is a simple shading technique that assigns a gray value to the pixel to be rendered based upon its distance from the origin of the ray. The further away a voxel lies from the imaging plane, the brighter it gets. Depth shading does not give a photorealistic impression, but it adds a visual clue and can be used if simple structures are to be rendered. This image was created using AnalyzeAVW.

FIGURE : The same volume as in Figure a, rendered by a surface shader. This image was created using AnalyzeAVW.

SURFACE–BASED RENDERING

- In Figure , we make a step from volume rendering to surface rendering.
- The main difference of surface rendering compared to volume rendering methods lies in the fact that we do not use the gray values in the image to assign a gray value or a color to the rendered pixel, but we encode the properties of a surface element to the gray value.
- An example already mentioned is depth shading – the distance of a voxel to the imaging plane determines its gray value.
- An example can be found in Figure. The transfer function for depth shading is given as

$$\mathcal{T}_{\mathrm{DS}} = \max \|\vec{x} - \vec{x}_{\text{Render Plane}}\| \quad \forall \{\vec{x}\}$$

Again, $\{\vec{x}\}$ is the set of voxels lying in the beam's path, and $\vec{x}_{\text{Render Plane}}$ is the end point of the ray. The voxel with the greatest distance to the image plane defines a surface; it is also the first voxel which lies above a rendering threshold encountered by the ray. Its distance gives the gray value of the pixel in the rendering plane the ray aims at.

SURFACE–BASED RENDERING

- A binary dataset containing a segmented volume is absolutely sufficient for surface rendering.
- In order to achieve a more natural view of the object when using surface rendering, we have to place our light sources in the coordinate system used, and we have to employ a lighting model that provides a simulation of optical surface properties.
- The most simple model is actually Lambertian shading; it is based on Lambert's law, which states that the intensity of reflected light from a diffuse surface is proportional to the cosine of the viewing angle.
- If we have a normal vector $\vec{n}$ on our ideal diffuse surface, and the view direction is given by a vector $\vec{v}$, the intensity I of reflected light is given by:

$$I = I_{\max} \frac{\vec{n} \bullet \vec{v}}{||\vec{n}||||\vec{v}||}$$

I_{max} is the maximum intensity, emitted in the direction of the normal vector $\vec{n}$ of the reflecting surface.

SURFACE–BASED RENDERING

- The equation is used to generate a realistic rendering of a surface.
- The normal to a surface voxel can be determined, for instance, by inspecting the neighboring voxels. Consider a 26-connected surrounding of a surface voxel. The ray defines a subset of these 26-connected voxels – the ones in front and directly behind the surface voxel are, for instance, irrelevant.
- The two gradients used express for seven voxels which are 26-connected to a surface voxel hit by one of the rays in the raycasting process.

SURFACE–BASED RENDERING

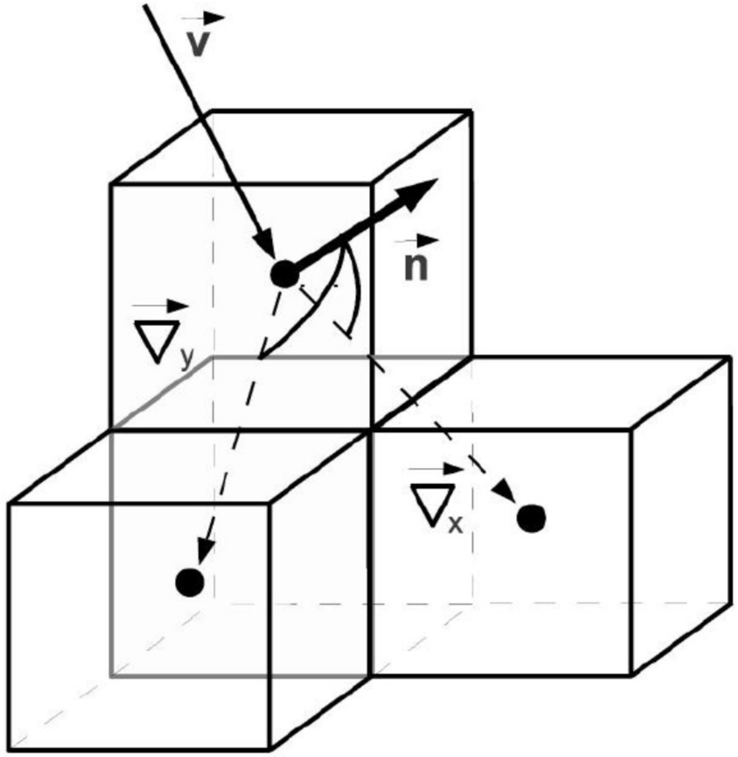


FIGURE: Computing two surface gradients for a surface voxel hit by a ray. The other non-zero voxels in the 26-connected neighborhood are used to compute two gradients ∇_x and ∇_y, which give the slope to the nearest other voxels belonging to the surface.

SURFACE–BASED RENDERING

- If we have the position vector to a surface voxel $\vec{x}_{ijk} = (x_i, y_j, z_k)^T$ at hand and our beam approaches in the z-direction, we have to select those 26-connected non-zero voxels that give two non-collinear gradients; the difference vectors between four neighbors, which can be aligned in x- and y-directions, or which can be determined by the two gradients with the largest norm, give the central differences $\vec{\nabla}_x$ and $\vec{\nabla}_y$ in vector notation.
- These two gradients span a local plane with the surface voxel $\vec{x}_{ijk}$ in the center. The normal vector on this local planar segment is given by the outer or cross product of $\vec{\nabla}_x$ and $\vec{\nabla}_y$. It is computed as

$$\vec{\nabla}_x \times \vec{\nabla}_y = \begin{pmatrix} \nabla_{x_1} \\ \nabla_{x_2} \\ \nabla_{x_3} \end{pmatrix} \times \begin{pmatrix} \nabla_{y_1} \\ \nabla_{y_2} \\ \nabla_{y_3} \end{pmatrix} = \begin{pmatrix} \nabla_{x_2}\nabla_{y_3} - \nabla_{x_3}\nabla_{y_2} \\ \nabla_{x_3}\nabla_{y_1} - \nabla_{x_1}\nabla_{y_3} \\ \nabla_{x_1}\nabla_{y_2} - \nabla_{x_2}\nabla_{y_1} \end{pmatrix}$$

This technique, where a normal vector is assigned to each surface element, is called flat shading. In our case, the surface element is one face of a voxel; a refinement can be achieved by using a finer resolution in voxel space. If we switch to visualization of surface models, where the surface is represented by geometric primitives such as triangles, flat shading may produce an even more coarse surface since a single element can be relatively large

Surface extraction, file formats for surfaces, shading and textures

- The representation of volume image data as data cubes consisting of voxels is, however, not very widespread outside the medical domain.
- 3D graphics as shown in computer games and computer-aided design (CAD) programs rely on the representation of surfaces rather than voxels.
- These surfaces are given as geometric primitives (usually triangles), defined by node points (or vertices) and normal vectors.
- The advantage of this technical surface presentation can be performed directly since the normal vectors and the area to be illuminated are directly defined.
- A computation of gradients and normal vectors is therefore not necessary since this information is already available in the surface presentation.
- Furthermore, modern graphics adapters are optimized for fast computation of renderings from such surface models, and well-developed application programmer interfaces (API) like OpenGL exist for all kinds of rendering tasks.

Surface extraction, file formats for surfaces, shading and textures

- In order to get a nice surface, one has to segment the organ or anatomic structure of interest.
- The great advantage of volume rendering techniques is actually the fact that segmentation is not necessary, although some sort of "soft" segmentation step is introduced in the transfer function.
- Furthermore, we lose all the information on tissue stored in the gray value of every single voxel.
- Another problem lies in the discretization of the surface – compared to the technical structures like screws and mechanical parts displayed in a CAD program, biological structures are of greater complexity, and the segmentation and computation of a mesh of geometric primitives usually lead to a simplification of the surface which may further obfuscate relevant anatomical detail.

Surface extraction, file formats for surfaces, shading and textures

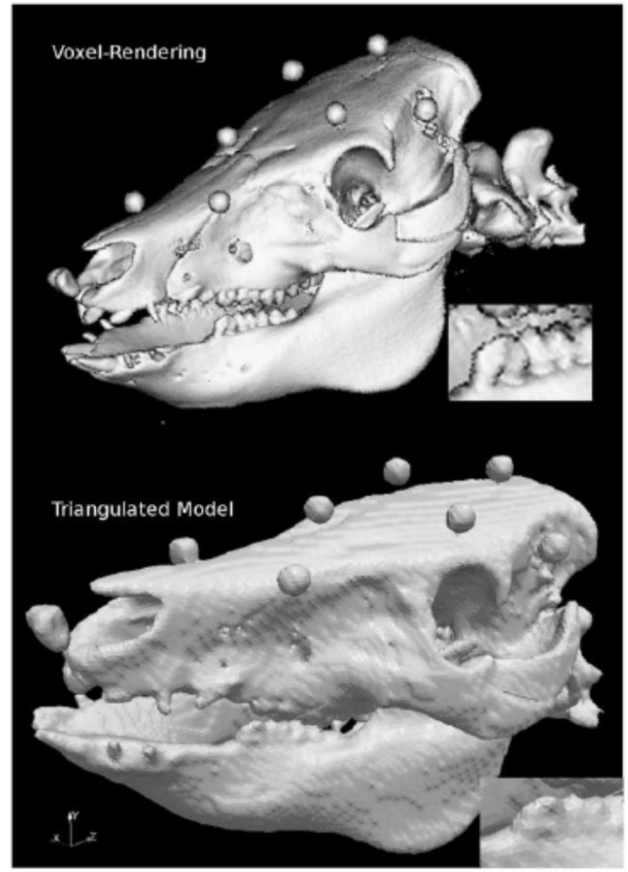


FIGURE: A comparison of voxel and triangulated surface rendering. The upper image of the well-known pig skull was generated from a voxel model of 0.5 mm^3 using AnalyzeAVW. The lower rendering shows the triangulated surface model generated using the surface-extraction module of AnalyzeAVW. The upper volume has a total size of 346.8 MB, whereas the binary file containing the 367061 triangles used for the lower rendering is only 13.1 MB large. A loss in detail in the lower image is evident – just inspect the enlarged section of the mandibular molars.

Surface extraction, file formats for surfaces, shading and textures

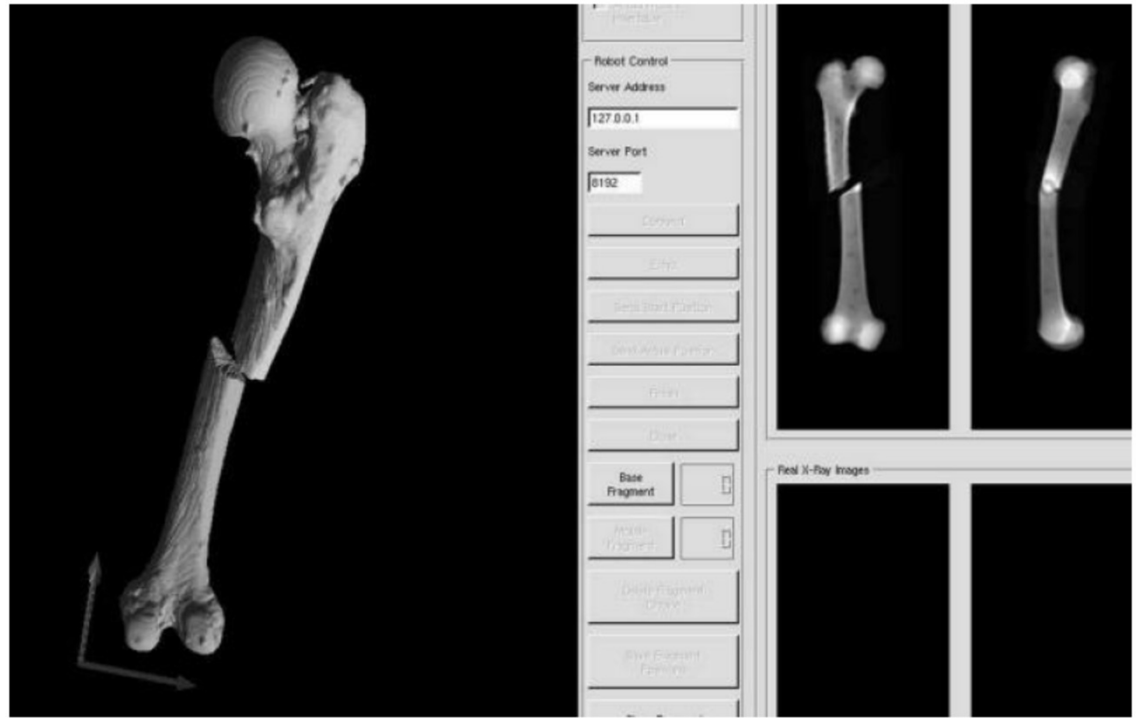


FIGURE: A screenshot of software for simulating long bone fractures. The user can apply an arbitrary fracture to a model of a bone derived from CT. The manipulation of the fragments in 3D is visualized using triangulated mesh models shown in the left part of the image. A collision detection algorithm, which keeps the fragments from intersecting each other, is also implemented. Meanwhile, simulated lateral and anterio-posterior DRRs are rendered as well from a smaller version of the original CT-volume (right hand side). The purpose of this training tool was to provide a simple demonstration for the context of 3D manipulation of bone fragments and the resulting x-ray images.

Surface extraction, file formats for surfaces, shading and textures

- Some fields of application where triangulated anatomical models are extremely helpful:
- *Fast rendering*: Real-time visualization, for instance in augmented reality or surgical simulation, is easily achieved by using appropriate hardware available at a more than reasonable price. Acceleration boards for voxel representations do exist but never gained wide acceptance in the field.

- *Collision detection*: Connected to the problem of simulation is the problem of realtime collision detection. In a rendering environment, usually no feedback is given if two surfaces intersect. Using a voxel-representation, it may be pretty time consuming to detect whether two surfaces collide or not. Collision detection is nevertheless relevant in applications like surgical planning, or in applications where a robot or a simple manipulator is used during an intervention. Using triangulated surfaces allows the use of fast collision detection algorithms.

Surface extraction, file formats for surfaces, shading and textures

- *Finite element modelling*: If we want to simulate the mechanical behavior of tissue, we have to use a simulation software based on the numerical computation of partial differential equations; this technique is generally called finite element modelling (FEM). FEM requires a 3D mesh generated from surface models or voxel volumes. The mechanical properties to be simulated are associated with the mesh elements.

- *Rapid prototyping*: Finally, it is possible to generate 3D plastic models out of volume data by rapid prototyping or a similar technique. The input data for such a 3D printer is usually a triangulated mesh.

Surface extraction, file formats for surfaces, shading and textures

- Therefore, it may be a good idea if we would find a way to transform a segmented, binary voxel volume into a set of geometric primitives forming a surface. In its most basic form, triangulation of a voxel surface takes place by
  - identifying voxels that belong to a surface – these are those voxels that have less than six neighbors.
  - assigning rectangles or triangles that represent the surface.
  - storing the resulting geometric primitives in an adequate form.

This method is generally referred to as a cuberille approach

Surface extraction, file formats for surfaces, shading and textures

- A more sophisticated triangulation algorithm should actually take care of smoothing the sharp edges of the voxel surface.
- Such an algorithm is the marching cubes algorithm.
- Let us illustrate the principle of this method in 2D; if we consider a binary image of a segmented structure, there are four possible elements that form a boundary of the shape when looking at areas of 2 × 2 pixels.
- These are shown in Figure .
- The four elements shown below the actual binary shape are the only elements that can form an outline – if you consider all of their possible positions, you will come to the conclusion that fourteen possible shapes are derived from these elements.
- If one wants to determine the outline of the shape, it is a feasible approach to check whether one of the fourteen shapes fits an arbitrary 2 × 2 sub image. If this is the case, one can assign a line segment representing the outline of the appropriate image element. This is called a marching squares algorithm.

Surface extraction, file formats for surfaces, shading and textures

- The interesting fact is that this method can be generalized to 3D, which results in the aforementioned marching cubes algorithm. Here, the volume is divided in cubic sub volumes of arbitrary size.
- There are 256 different possibilities to populate the eight corners of these cubes.
- If we omit the possibilities which are redundant since they can be generated by rotating a generic shape, we end up with fifteen different configurations.

Surface extraction, file formats for surfaces, shading and textures

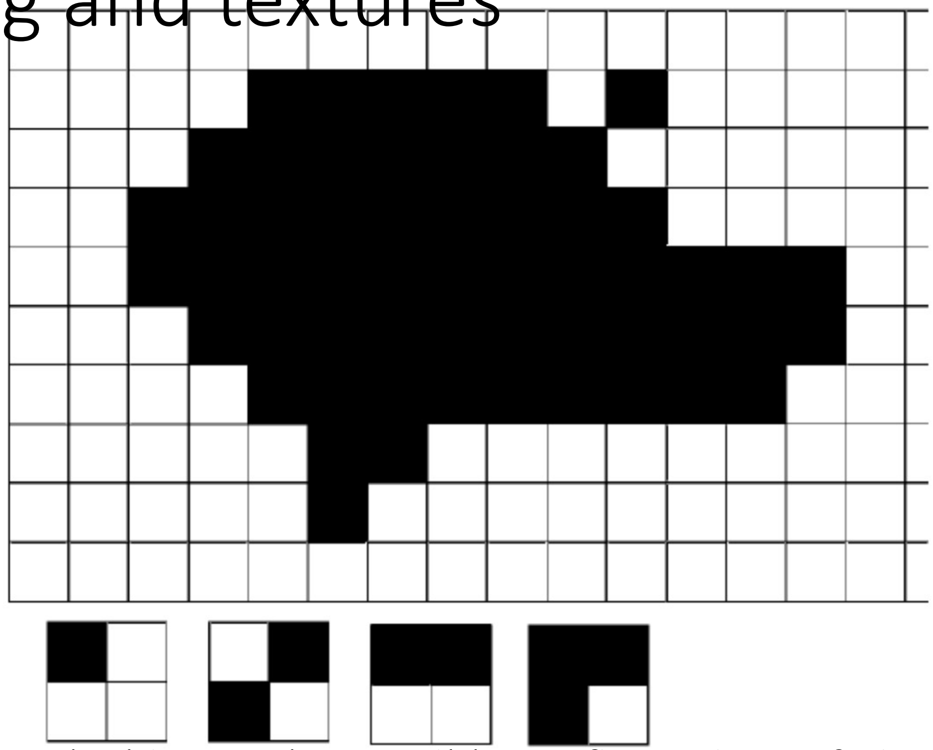


FIGURE : When looking at the possible configurations of pixels forming an outline of a binary 2D shape, we can identify four basic shapes. By rotating these basic shapes, it is possible to identify all straight line segments forming the shape. These line segments form the outline of the shape – the algorithm that does this for a 2D image is called the marching squares algorithm.

Surface extraction, file formats for surfaces, shading and textures

- Fourteen possible 2D configurations can be derived from four generic shapes.
- The advantage of the marching cubes algorithm lies in the fact that it can be easily scaled – choose large cubes as sub volumes, and you will get a coarse grid.
- Small cubes result in a fine grid.
- Two out of fifteen possible configurations with associated triangles are shown in Figure
- An actual implementation of the marching cubes algorithm simply consists of a lookup-table that contains the fifteen basic shapes and a rather straightforward algorithm that assigns the appropriate set of triangles appropriate for the shape encountered in a given cube.

Surface extraction, file formats for surfaces, shading and textures

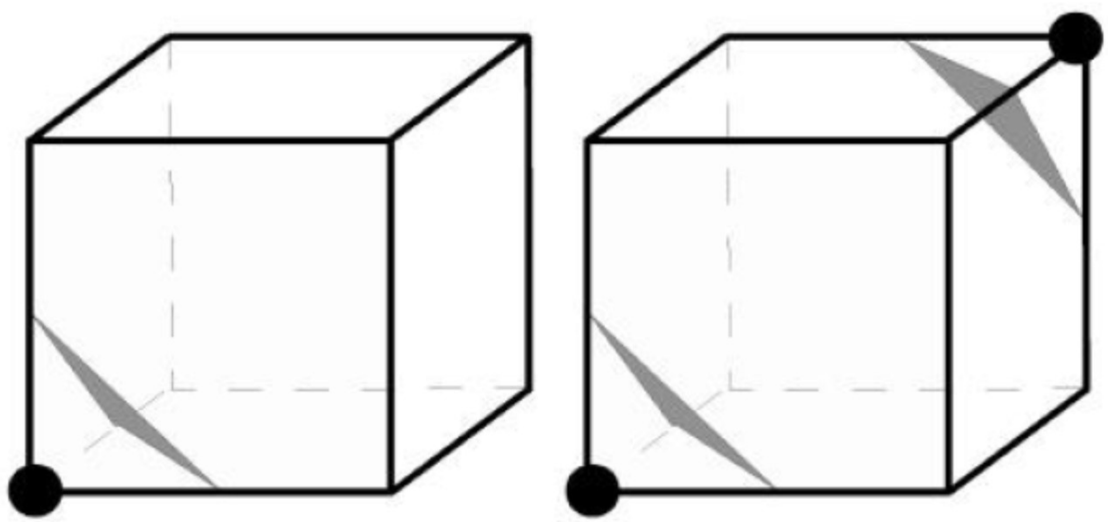


FIGURE: Two out of fifteen possible configurations for the marching cubes algorithm. The black circles denote occupied edge points in the cubic sub volumes. In 3D triangulation algorithm, the volume is divided into cubic sub volumes. Dependent on the number of cube edges populated with non-zero voxels, certain triangle sets are assigned. The result is a smooth triangulated surface.

Surface extraction, file formats for surfaces, shading and textures

- A typical format, which is also used for storing the results of our triangulation effort is the Surface Tessellation Language (STL).
- It is a standard that is readable by most programs for computer-aided design, stereolithography, and surface visualization. It exists in an ASCIIencoded text type and a binary form.
- The general shape of an ASCII-STL file is as follows:
  - solid name
  - facet normal xn yn zn
  - outer loop
  - vertex x1 y1 z1
  - vertex x2 y2 z2
  - vertex x3 y3 z3
  - endloop
  - endfacet
  - …endsolid name
- A binary version of STL therefore exists as well. Other file formats for triangulated surfaces are, for instance, the Initial Graphics Exchange Specification (IGES) or the Standard for the Exchange of Product model data (STEP); they are, however, organized in a similar manner.

Shading models

- Triangles covering large, planar areas may appear rather dull when using a simple flat shading model
- In medical image processing, this is not as big a problem as in technical surfaces s since the surface encountered in the body are usually not very regular, and a fine representation can be achieved by using volumes with a fine voxel grid.
- If one wants to improve the appearance of a model triangulated using a rather coarse grid, it may be useful to interpolate additional normal vectors in dependence of the normal vectors associated with a surface facet.
- Figure illustrates this principle, which is also known as Phong shading.
- In Gouraud shading, interpolates the color of pixels in the image plane between vertices instead of the normal vectors,
- Finally, one may render 2D patterns, so-called textures on the graphic primitives to increase the 3D effect. This is of great importance for instance in computer games, surgical simulation, and other visualization applications.

Shading models

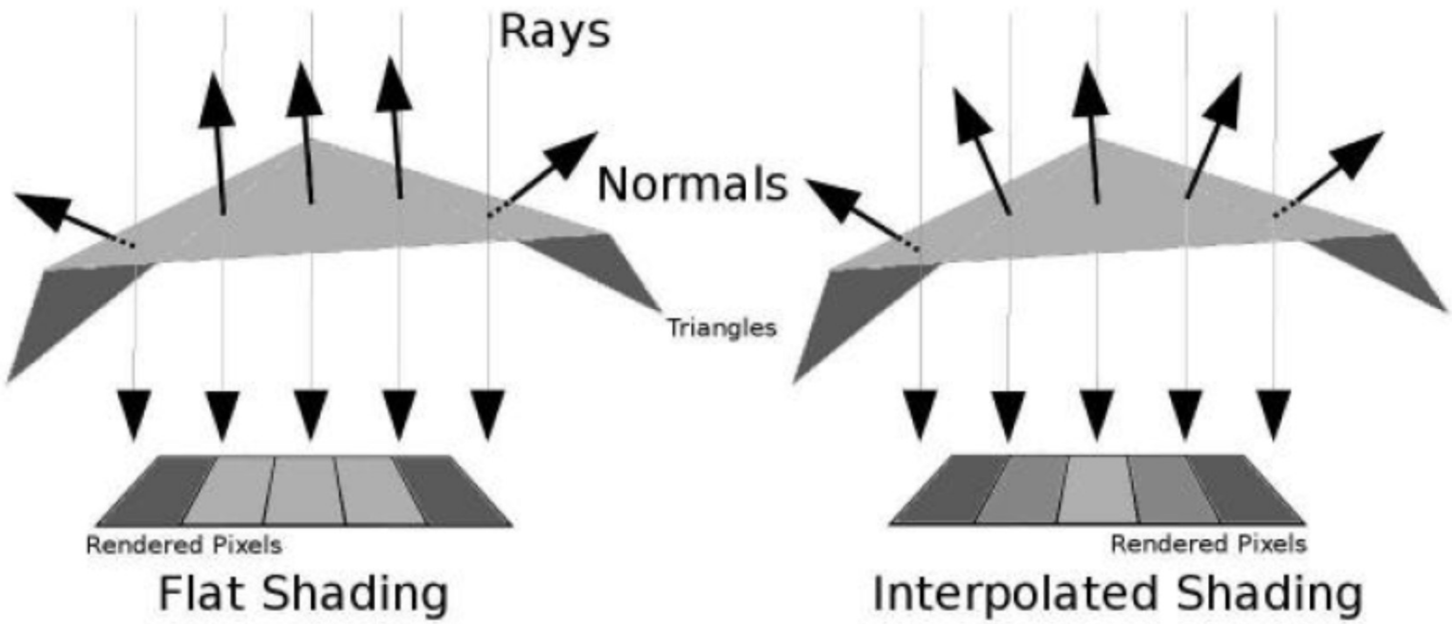


FIGURE: The principle of Gouraud-shading. When rendering a large triangle that is hit by several rays aiming at neighboring pixels in the render plane, we may end up with rather large, dull areas on the rendered image. If the normal vectors for the respective rays are interpolated between the neighboring geometric primitives, we will get a smoother appearance of the rendered surface.

A special application – virtual endoscopy

- An endoscope is an optical instrument with a small wide-angle camera that can be inserted into the body.
- Various types of flexible and rigid endoscopes for neurosurgery, ENT-surgery and, above all, gastrointestinal interventions exist.
- It is also possible to take histological samples or to inject drugs using these devices. Given the ergonomics of an endoscope, it is sometimes difficult to orientate oneself during an endoscopic intervention.
- Therefore it is interesting, for training purposes as well as visualization, to simulate endoscopic views. Since the endoscope is, technically speaking, a wide-angle camera, this can be performed using a perspective surface rendering technique.
- A special problem here is the simulation of barrel distortion, an optical property of wide-angle lenses that is also known as the fisheye effect.
- Figure shows such a virtual endoscopy – here, we take a look in cranial direction through the cervical spine of our pig CT dataset into the calvaria; the rendering threshold is chosen in such a manner that no soft tissue is visible. Virtual endoscopy has gained some importance in the early detection of colon cancer, where it is sometimes used to generate endoscopy-like views from CT data of the lower abdomen.

virtual endoscopy

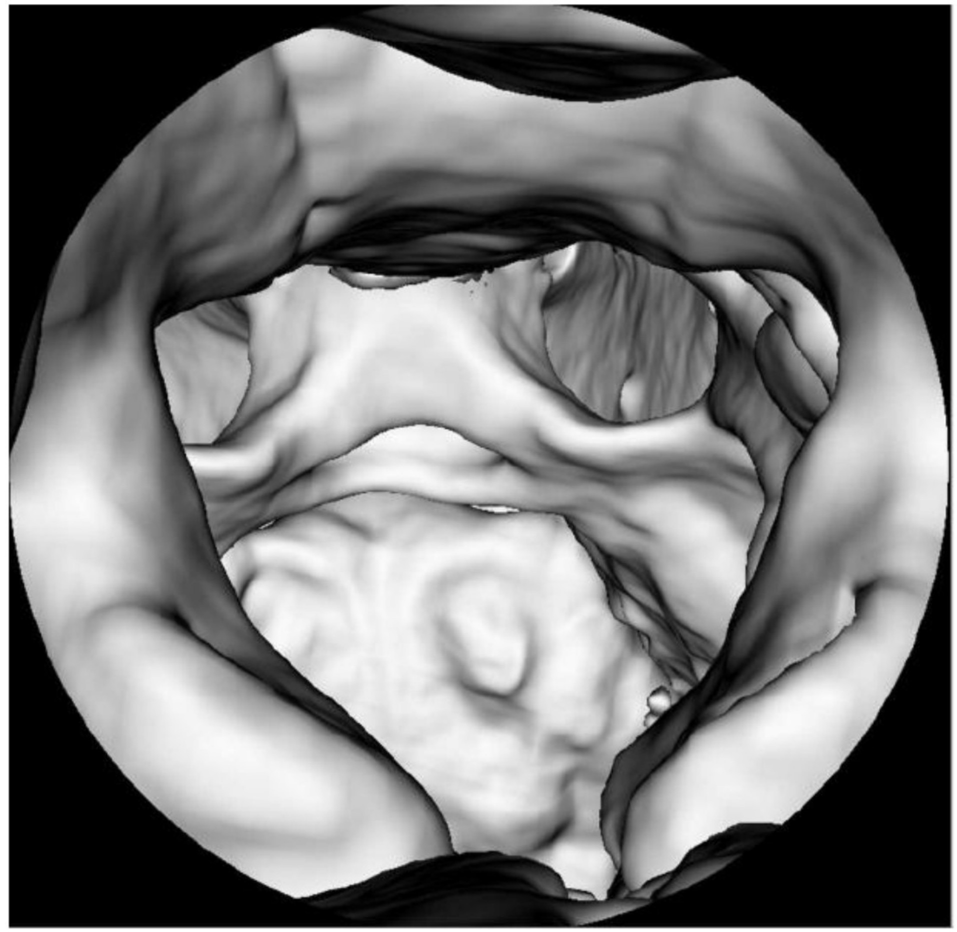


FIGURE: A virtual endoscopy of the spinal canal of the pig dataset, rendered from CT. Basically speaking, this is a surface rendering that, in addition, simulates the barrel distortion induced by the wide angle optics of a conventional endoscope.

Registration

FUSING INFORMATION

- Visualization and interpolation operations provide the core formalism for image registration or image fusion.
- Most modern modalities produce 3D volume image data does not simplify the task of fusing information from different imaging sources.
- This is the domain of registration algorithms, where a common frame of reference for multiple data sources is established.

FUSING INFORMATION

- The main problem lies in the fact that the reference coordinate system of a volume dataset does not reference the patient – it references the scanner.
- Changing patient position between two imaging sessions is, however, necessary in many cases.
- The position of a patient in an MR scanner during a head scan is governed by the shape of the receiver coil and the gradients, whereas patient pose in a CT or PET scanner is governed by the patient's principal axes.
- Combined modalities like PET/CT solve this problem; the poor anatomic image information from PET, especially in body regions other than the brain, renders exact fusion of PET and CT or MR difficult, and the information from the CT scan can even be used to improve volume reconstruction in the PET scan.
- Figure gives a sample of the capabilities of a modern PET/CT system, and PET/MR systems are subject to clinical investigation. However, it is not that simple to combine modalities.
- A CT/MR scanner would be a very clumsy and difficult device, given the fact that the avoidance of ferromagnetic materials in the CT and the appropriate shielding of electromagnetic fields from the CT scanner is a rather delicate task. It would also be uneconomic to use such a system; a CT can handle more patients within a given period of time compared to an MR-tomograph.

FUSING INFORMATION

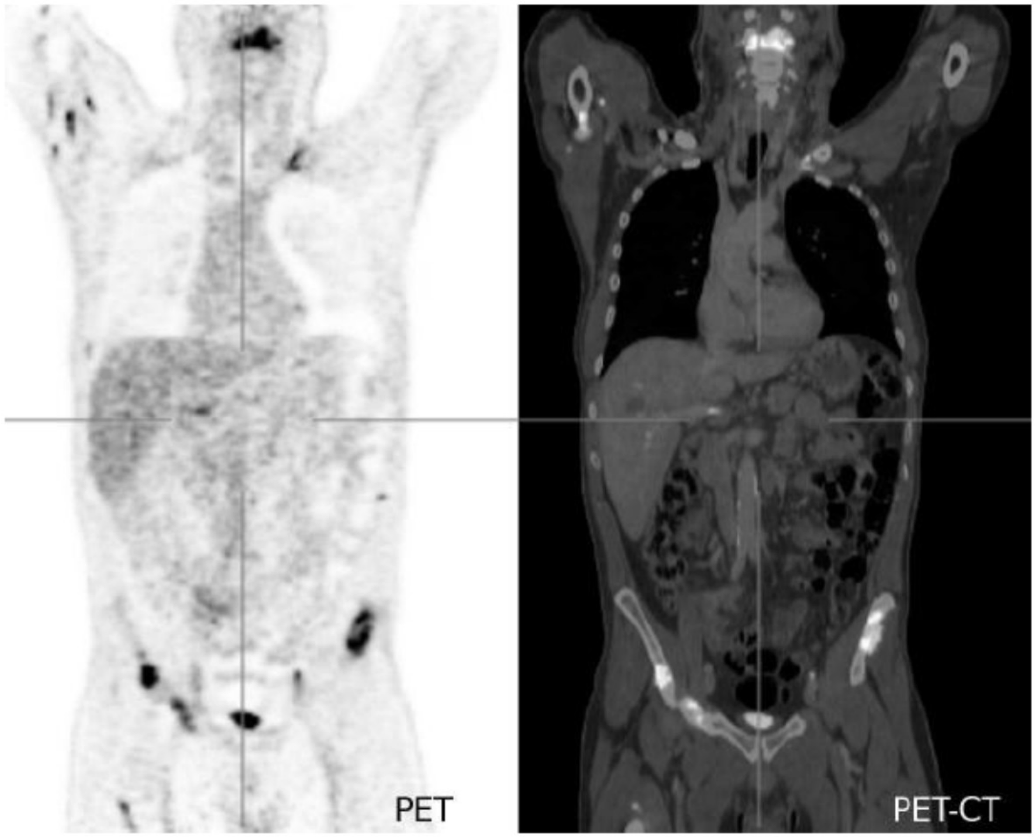


FIGURE: An image sample from a combined PET/CT system. Registration of the PET and the CT volume is automatically achieved by taking both volume datasets in a scanner that is capable of recording both tracer concentration and x-ray attenuation.

FUSING INFORMATION

- Fusion of images – that is, finding a common frame of reference for both images – is therefore a vital task for all kinds of diagnostic imaging.
- Besides fusing images for tasks such as comparison of image series, we can co-register the coordinate system of an external device such as a robot or a LINAC.
- It is also possible to merge abstract information, such as eloquent areas in the brain from a generalized atlas, to patient specific image information.
- In general, all registration algorithms optimize parameters of rigid motion and internal dof (for deformable registration) until a measure that compares images is found to be optimal.
- The choice of this measure, which we will simply call merit function, depends on the specific registration task.

FUSING INFORMATION

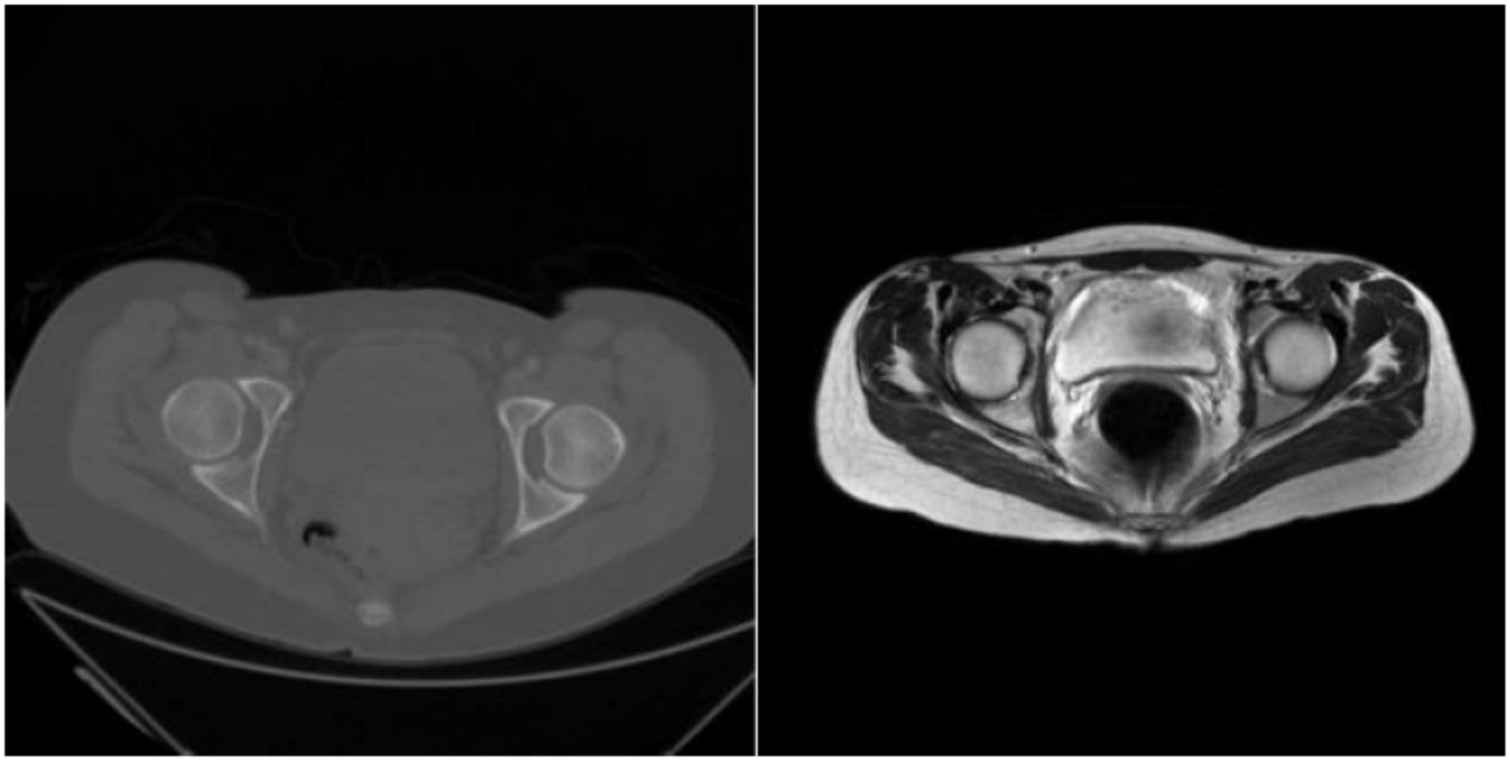
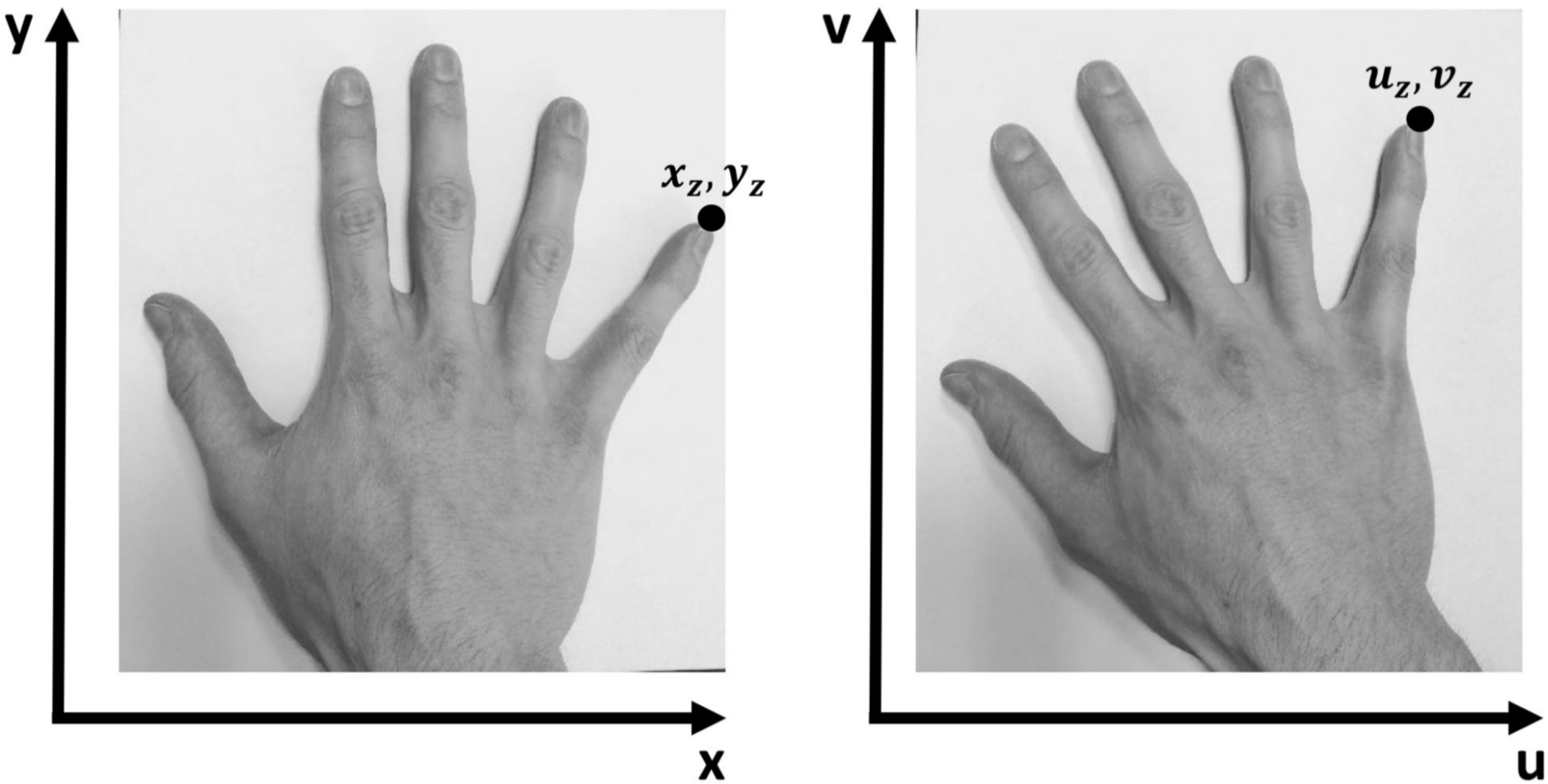


FIGURE: Two slices of the same patient at the same location, the hip joint. The left image shows the CT image, the right image is taken from the MR scan. Differences in the field-of-view and patient orientation are evident. Registration algorithms compensate for these discrepancies by finding a common frame of reference for both volume datasets.

REGISTRATION PARADIGMS

- Image registration is the process of finding accurate spatial correspondence between two or more images.
- Physicians often wish to compare two images of the same anatomical region acquired under different circumstances.
- For instance a pair of images might be of the same patient and the same modality, or different modalities. Also, images of two different patients might be compared.
- In general, registration algorithms determine a volume transformation from identifying common features in two coordinate systems.
- These features can be intrinsic (that is features that are inherent to the image data) and extrinsic.
- Extrinsic features are usually markers attached to the patient; these markers, usually called fiducial markers, are identified in both frames of reference, but cause considerable additional clinical effort and are usually not available for retrospective studies.
- Registration algorithms can also be categorized by the type of image data or frames of reference they operate on.

REGISTRATION



Intra- and intermodal registration

- In intramodal registration, registering images that have the same modality.
- All slices can be viewed with the same orientation.
- In intermodal registration, fusing image data from different modalities.
- Example: A MR/CT fusion using an similarity measure named normalized mutual information (NMI) in AnalyzeAVW can be found in Figure
- In intramodal registration, it may be sufficient to define a measure that compares the voxel gray values ρ at identical positions in the base and match volume.
- In intermodal registration, such an approach is bound to fail since the physical principle for the imaging modalities is different; therefore, there is no reason why a bright voxel in one volume should correspond to a bright voxel in the other volume.

Intermodal registration

- Intermodal (multimodal, crossmodal) registration involves matching images of the same patient acquired from different modalities.
- This category can be further divided on a modality basis into anatomical-anatomical, functional anatomical and functional-functional registration.
- Multi-anatomical registration deals with the registration of imaging studies depicting different aspects of tissue morphology. e.g. CT-MR
- The correlation of functional studies with anatomical studies allows for the anatomical localisation of functional parameters, as functional studies often lack morphological information. e.g. SPECT-MR
- Functional studies correlate differing types of functional information e.g. ^{201}TI or ^{99m}Tc-MIBI SPECT study depicting perfusion with a metabolic 18F-FDG PET study

Intramodal Registration

- Intramodal (monomodal, isomodal) registration involves matching images of one or more patients and the same modality acquired at different times.
- This allows quantitative comparison for longitudinal monitoring of disease progression/recession and postoperative follow up.
- Intramodal registration is well suited to tasks relating to detection: monitoring growth, comparative studies using contrast agents, and subtractive imaging.
- There are three interrelated subcategories of intramodal registration:
  - Inter-study
  - Intra-study
  - Inter-subject
  - Intra-subject

Intramodal Registration: Inter/Intra-Subject Registration

- When all the images involved in registration are acquired from a single patient, this is known as intra-subject registration.
- If the registration is accomplished using images from different patients, this is referred to as inter-subject registration. This implies matching studies taken from different patients but from the same modality.
- *Spatial normalization*: which minimizes shape variability

Intramodal Registration: Interstudy Registration

- Matching images from the same patient and the same modality, but from different studies is known as interstudy registration and is synonymous with temporal or time-series registration.
- Interstudy registration is performed in situations such as quantisation of bone growth (long time interval), monitoring tumor growth (medium to long interval), quantisation of cardiac motion (short interval), or observing the movement of a contrast agent through vasculature (ultra-short interval).

Intramodal Registration: Intrastudy Registration

- Matching images from the same patient and from within the same study is called
- intrastudy registration.
- For example when acquiring a contrast-study, two or more separate images are acquired in the same session. The first of these is known as the precontrast image and is acquired before the administration of a contrast agent (a pharmaceutical used to enhance particular types of tissue). After the contrast-agent has been administered one or more post-contrast images is acquired.

Intra- and intermodal registration

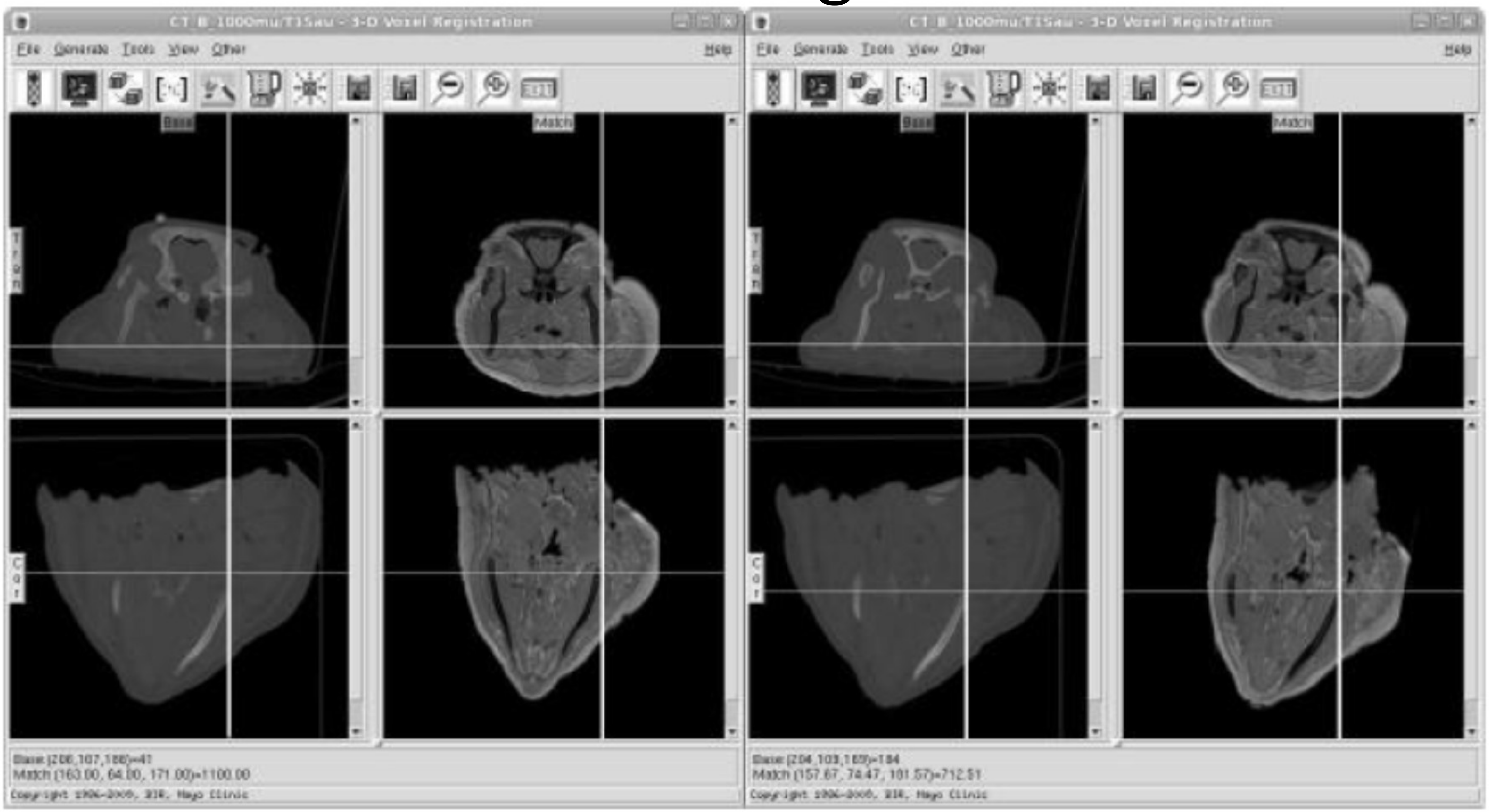


FIGURE : Two screenshots from the registration tool of AnalyzeAVW. On the left side, we see two slices of the MR and the CT scan, unregistered. Gross orientation is similar, but discrepancies are clearly visible. Besides soft tissue deformation, it is also evident that rigid structures such as the calvaria of the well-known pig scan do not coincide. After initiating the registration algorithm (in this case, an implementation of normalized mutual information), the two images match. The resulting volume transformation matrix V can be seen in Figure

Intra- and intermodal registration

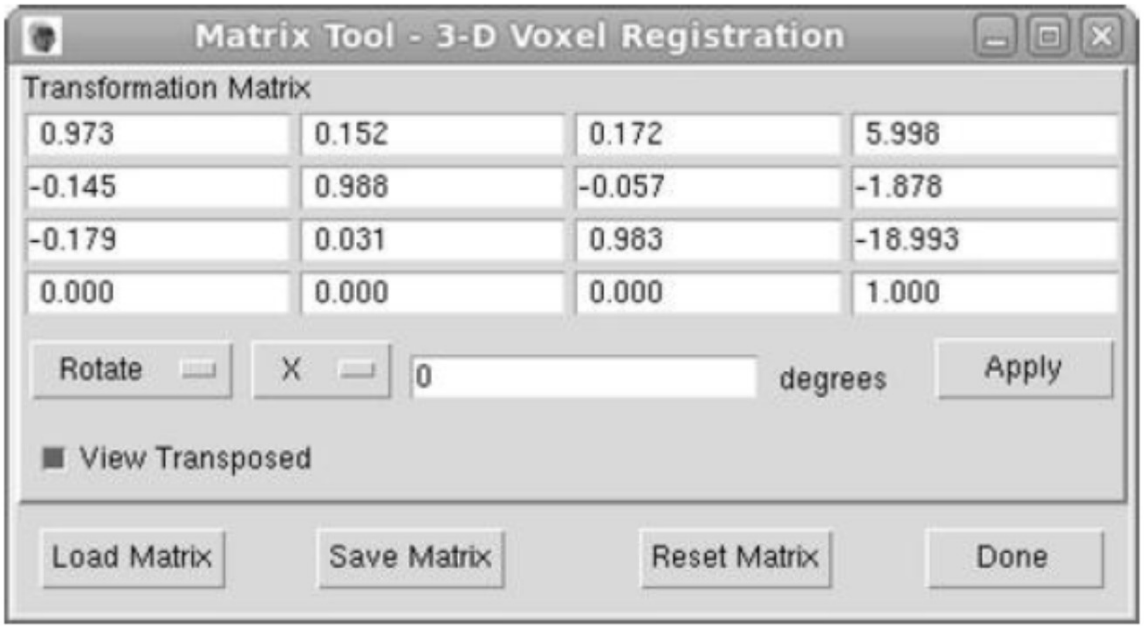


FIGURE: The volume transformation as produced by the AnalyzeAVW registration tool

Rigid and non-rigid registration

- Rigid registration refers to algorithms that are confined to finding an affine transformation in three or six dof;
- The shape of tissue may change and their relative position of inner organs may vary.
- In the registered transversal and coronal slices, we see that bony structures coincide very well.
- The soft tissue does not – this is simply due to the fact that the poor little piglet was decapitated before we scanned it.
- Rigid or affine transformations are suitable for intrasubject registration tasks with little physical deformation or motion
- e.g. head, spine, pelvis

Rigid and non-rigid registration

- In non-rigid registration, the affine transformation is only a first step in alignment of volumes. The fine-tuning is handled by adding additional, internal dof, which handle the deformation as well.
- A non-rigid registration algorithm is therefore not a completely different approach to the registration problem; it relies on the same merit functions and optimization schemes, and it can be intra- or intermodal.
- Need additional assumptions on how image elements are allowed to migrate. Each image element is displaced by a vector that indicates the direction and amount of displacement. The distinctive feature for deformable registration methods is the way of modelling this displacement.
- In general, the mesh defining the displacement field is given, and local differences in the base and the match image are evaluated using a merit function; the local difference in the image define the amount of displacement for a single image element.
- Nonrigid or elastic transformation models are suitable for intersubject registration tasks e.g. anatomical variability across subjects or softtissues that deform over time

Rigid and non-rigid registration

- The algorithms for modelling the displacement field can be put into the following categories:
- Featurelet-based approaches: The most trivial approach is to sub-divide the image to smaller units, which are registered as rigid sub-units. The resulting gaps in the image are to be interpolated. This technique is also called piecewise rigid registration.

Rigid and non-rigid registration

- B-Spline-based approaches: In short, a spline is a numerical model for a smooth function given by a few pivot points.
- A mesh of pivot points is defined on the image, and the merit function chosen produces a force on the B-spline.
- It can model deformation in a smooth manner, but huge local deformation will have a considerable effect on remote areas of the image.

Rigid and non-rigid registration

- Linear elastic models: For small deformations, it is permissible to treat the surface under deformation as an elastic solid.
- Local differences again apply a force on the elastic solid until an equilibrium is reached.
- The linearity assumption is only valid for small deformations, as one might know from real life.
- The advantage of this linear model, however, lies in the fact that non-local effects such as in the case of B-splines should not occur, but on the other hand, huge deformations cannot be compensated.

Rigid and non-rigid registration

- Viscous fluid models: A generalization to the linear elastic model is the viscous fluid model, where strong local deformations are allowed; one can imagine a viscous fluid model as honey at different temperatures – it may be rather malleable or very liquid.
- Such a model allows for strong local deformations at the cost of registration failure if a large deformation occurs because of a flawed choice of internal parameters.

Rigid and non-rigid registration

- Finite-element models: In finite-element modelling (FEM), a complex mesh of the whole image is constructed, with well defined mechanical properties on interfaces of materials with different properties; in such a model, bone can be modelled as rigid, whereas other anatomical structures are handled as viscous fluids or as linear elastic models.
- The power of such an approach is evident since it models the structure of an organism.
- The drawback is the fact that a full segmentation of the volume of interest is necessary, and that constructing the mesh and solving the partial differential equations of a FE-model can be extremely labor and timeintensive.

Applications for Registration in Medicine

- Applications can be broadly grouped into two interrelated categories; clinical (detection and diagnosis) and surgical. For example:
  - Radiation therapy
  - Interventional radiology
  - Diagnostic radiology
  - Image-augmented surgery
  - Preprocedural planning and simulation
  - Minimally invasive procedures

Image Registration Process

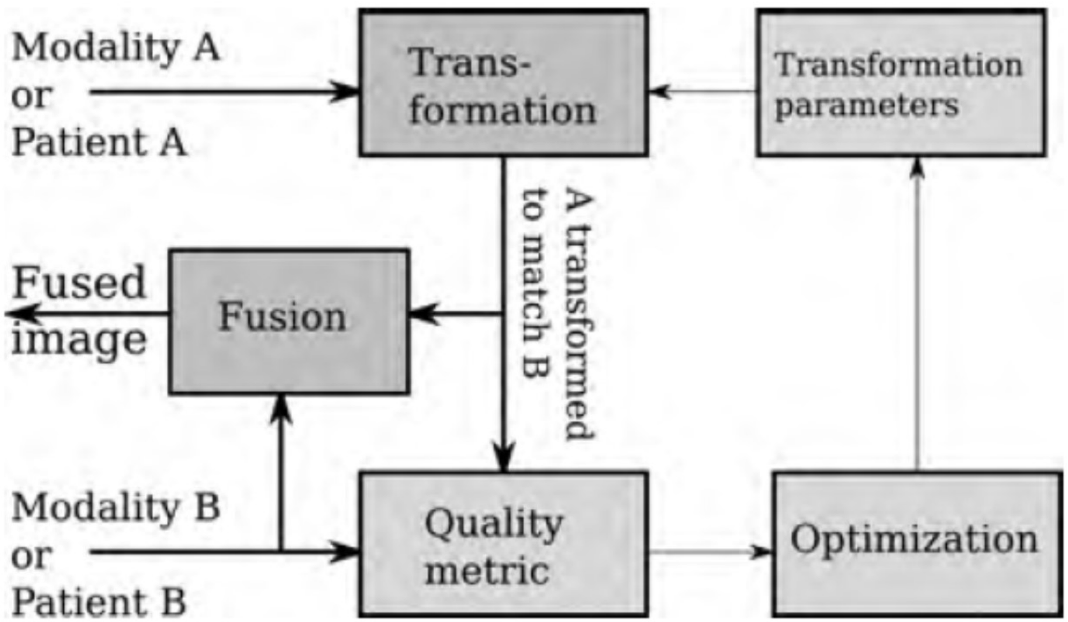
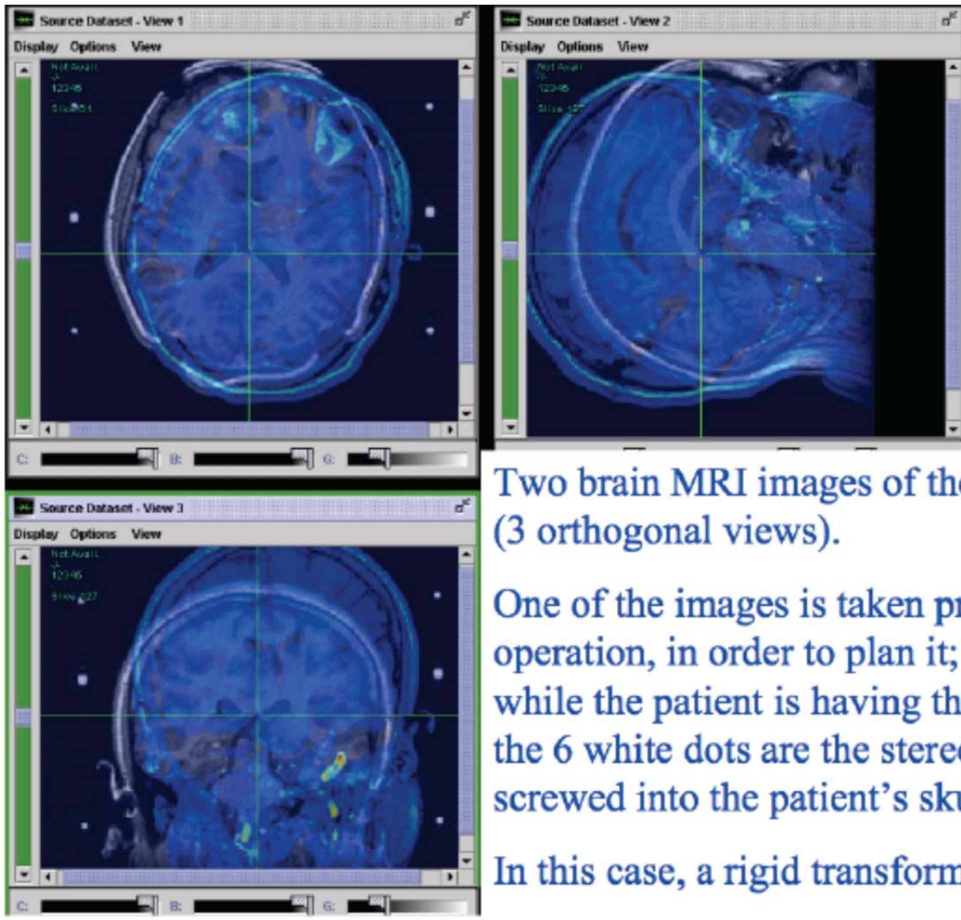


FIGURE: Schematic representation of the registration process. The image from modality or patient A undergoes a transformation to match the reference, that is, modality or patient B. After transformation, A and B are compared and a strategy to optimize the match between the images is applied. On the basis of this strategy, new transformation parameters are computed and applied. The process may be repeated iteratively. Once the registration process has converged, both images may be fused, that is, superimposed for visualization.

Image Registration Process

- The steps needed to spatially superimpose the images are called *image registration*.
- One example is the registration of magnetic resonance imaging (MRI) and nuclear imaging modalities, such as positron emission tomography (PET) or single-photon emission computed tomography (SPECT).
- Both PET and SPECT have a very low spatial resolution.
- MRI provides anatomical details by merit of its high spatial resolution and tissue contrast. Therefore, SPECT or PET images are often superimposed over an MRI scan to allow better visualization of the active sites.
- For an accurate match of the corresponding sites in both scans, the PET or SPECT image needs to be rescaled to match the resolution of the MR image, and translated and rotated to match the spatial locations between the images.
- It is possible to combine several modalities in one gantry to form hybrid imaging systems.

Before Registration

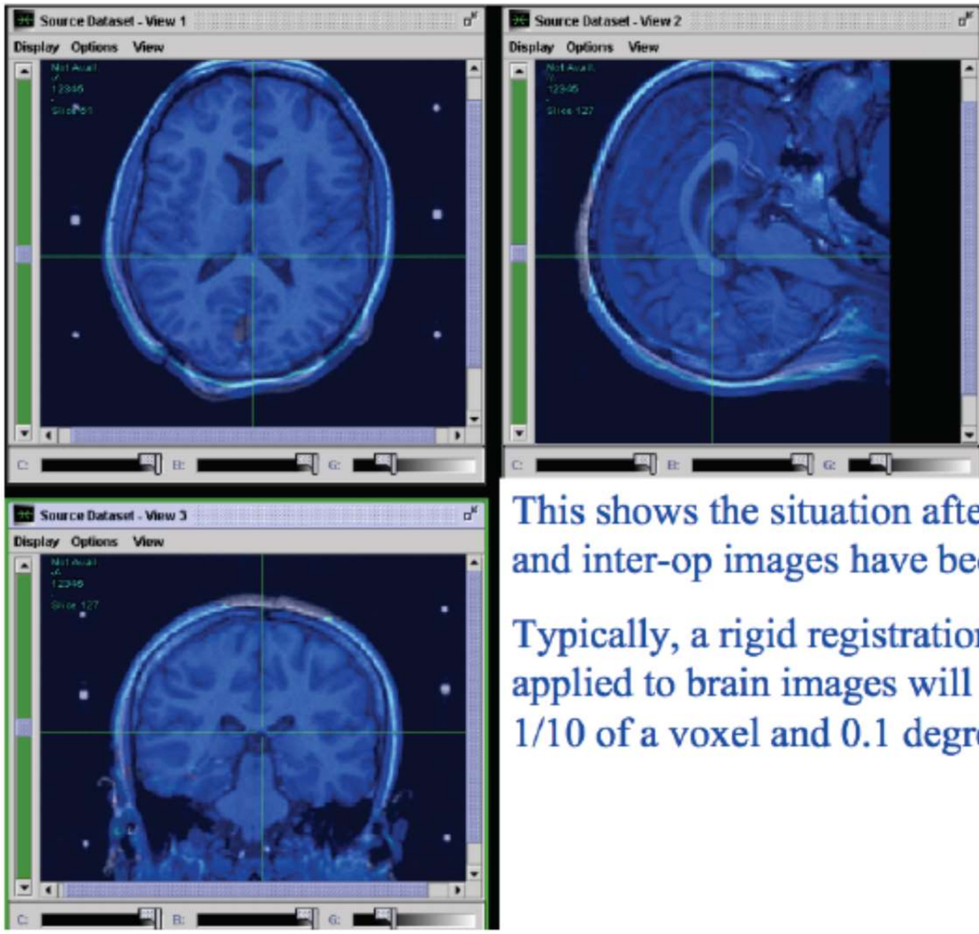


Two brain MRI images of the same patient (3 orthogonal views).

One of the images is taken prior to the operation, in order to plan it; the second while the patient is having the operation: the 6 white dots are the stereotactic frame screwed into the patient's skull.

In this case, a rigid transform suffices

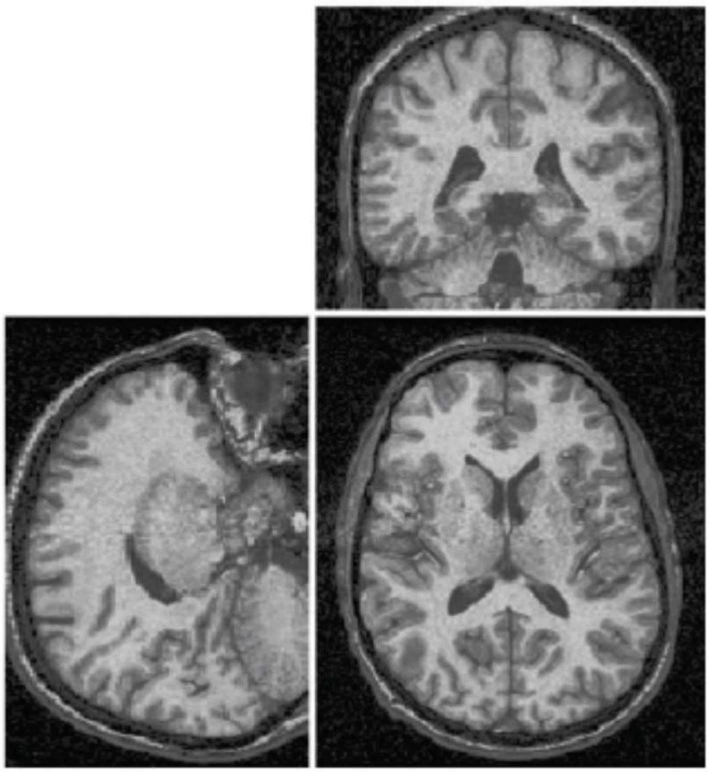
After Registration

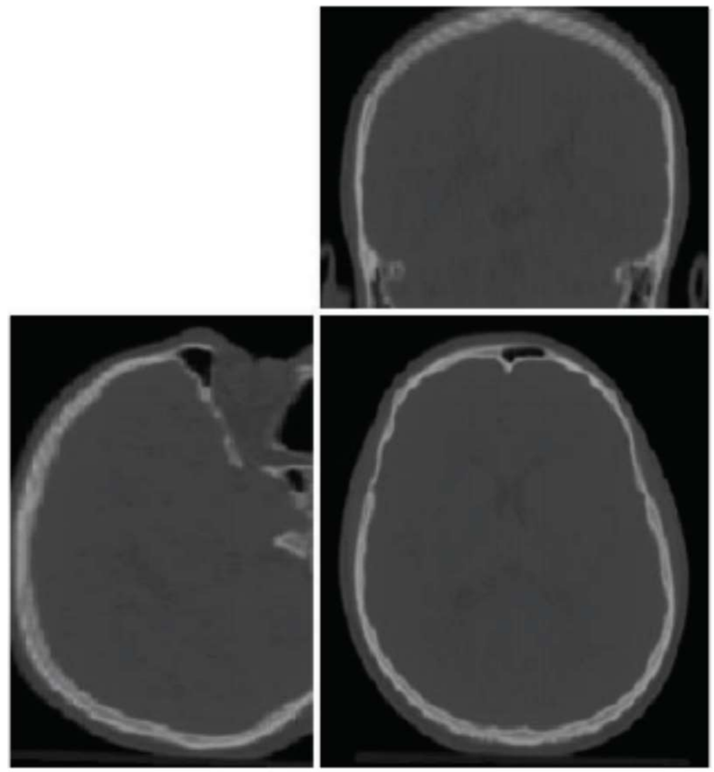


This shows the situation after the pre-op and inter-op images have been aligned.

Typically, a rigid registration algorithm applied to brain images will be accurate to 1/10 of a voxel and 0.1 degrees of rotation

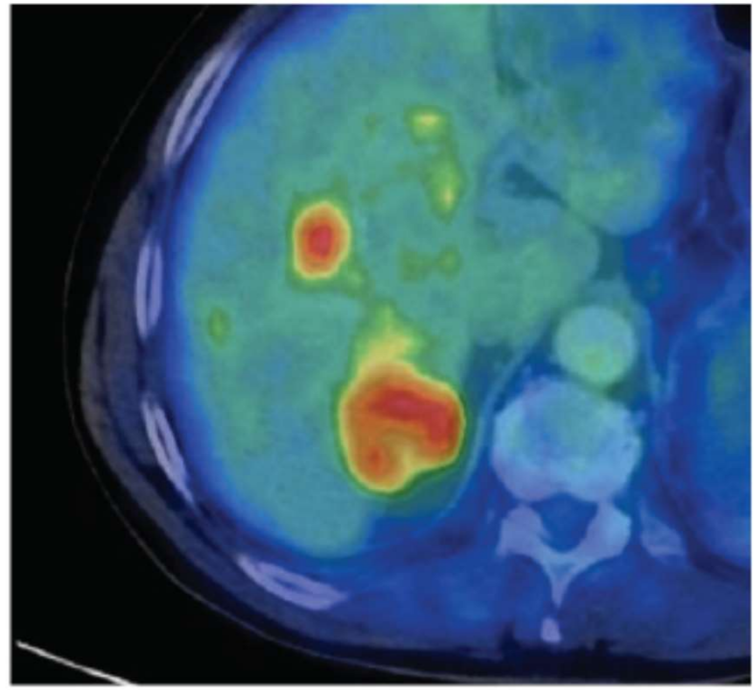
Example: rigid CT/MR registration



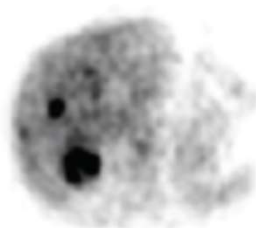


Fusion of information = registration plus combination in a single representation: PET/CT

CT



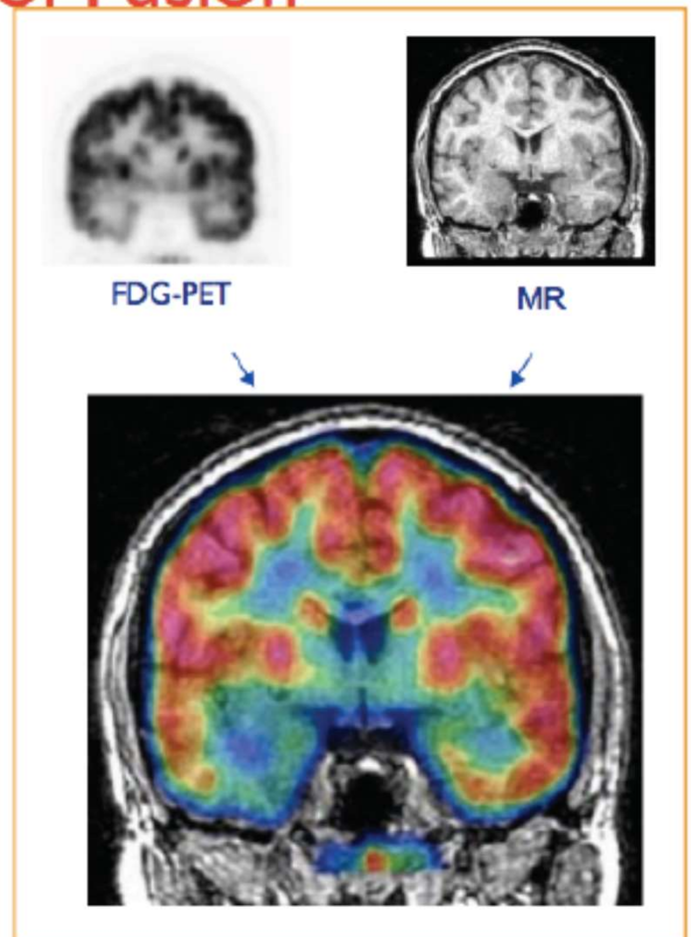
PET



Deformable fusion- PET shows increased metabolism in lesions identified on CT, consistent with active tumour growth rather than necrosis post-radiotherapy

Many Clinical Applications of Fusion

- Cancer staging
- Biopsy planning
- Radiotherapy treatment planning
- Quantitative assessment of treatment response
- Pre-surgical assessment of other conditions e.g. epilepsy
- As an *effective* communication tool when reporting to clinical meetings, referring physicians or to patients
- Whenever multiple data sources may be better assessed together

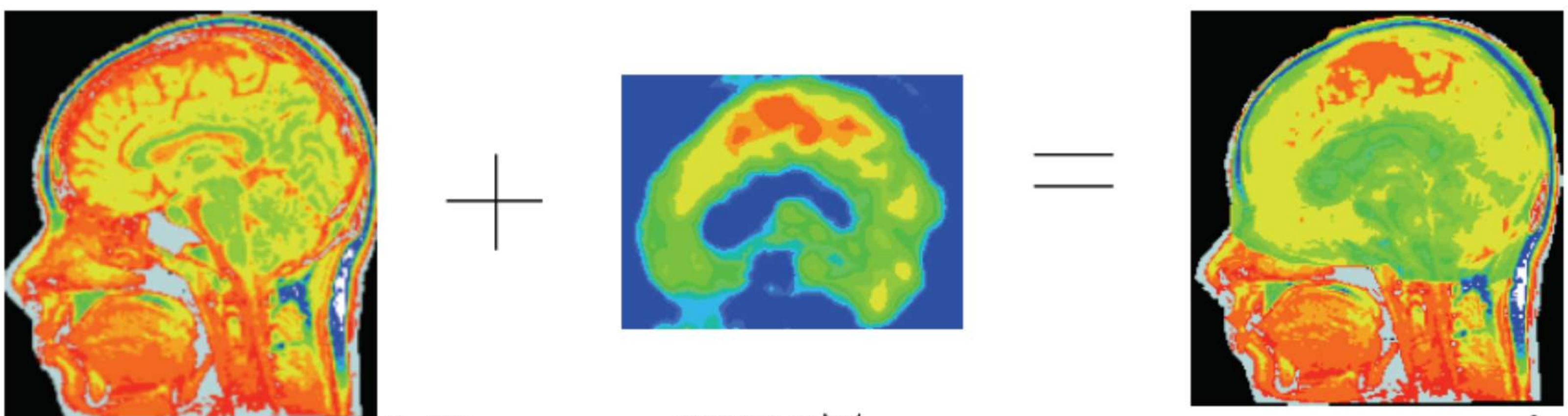


PET data identifies a region of hypometabolism due to epilepsy. Fusion with MR localises the damage to the anterior and medial areas of the right temporal gyrus

Image Registration is a

Spatial transform that maps points from one image to corresponding points in another image

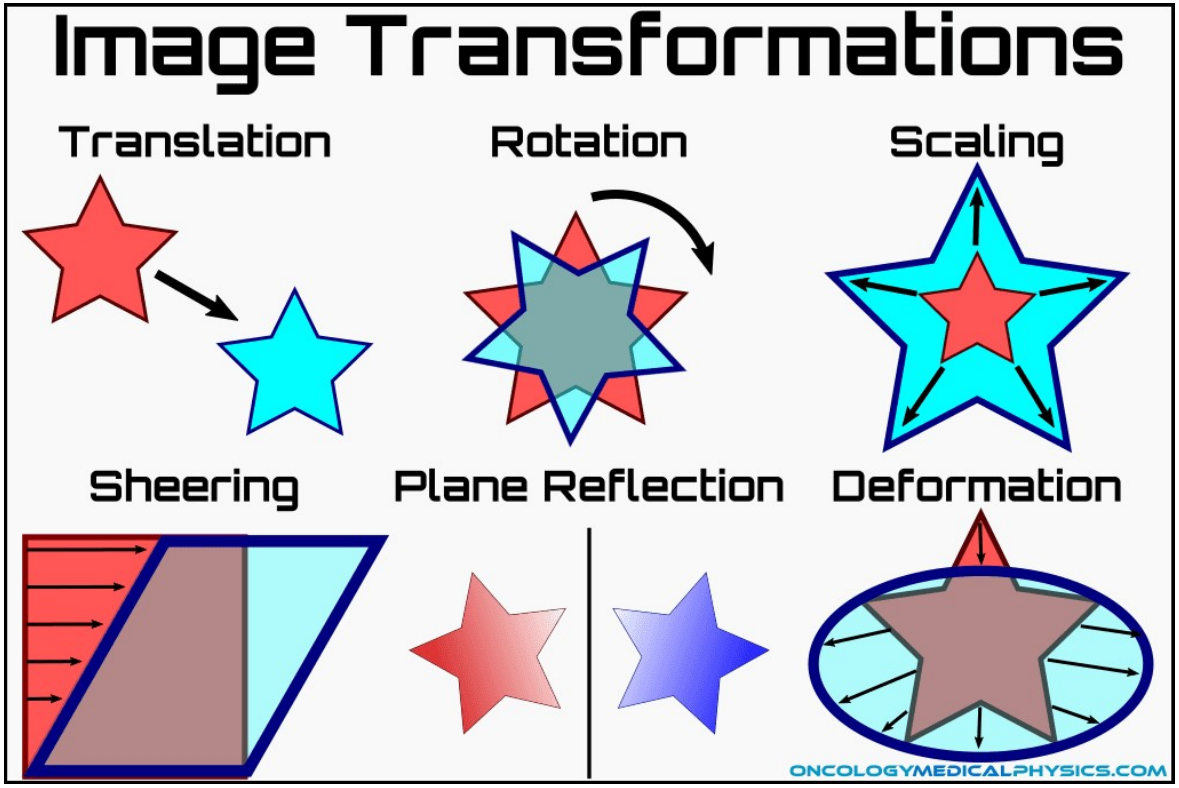
matching two images so that corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged also referred to as image fusion, superimposition, matching or merge



MR SPECT registered

Image Registration is a

- **Spatial transform** that maps points from one image to corresponding points in another image
- – Rigid
  - Rotations and translations
- – Affine
  - Also, skew and scaling
- – Deformable
  - Free-form mapping



Registration Taxonomy

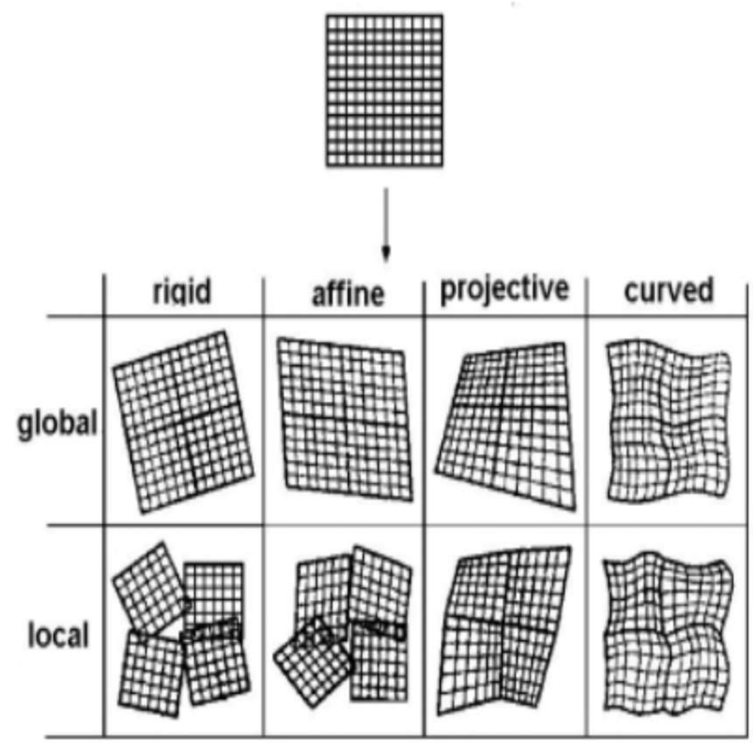
- **Dimensionality**
  - 2D-2D, 3D-3D, 2D-3D
- **Nature of registration basis**
  - Image based
    - Extrinsic, Intrinsic
  - Non-image based
- **Nature of the transformation**
  - Rigid, Affine, Projective, Curved
- **Interaction**
  - Interactive, Semi-automatic, Automatic
- **Modalities involved**
  - Mono-modal, Multi-modal, Modality to model

- **Subject:**
  - Intrasubject
  - Intersubject
  - Atlas
- **Domain of transformation**
  - Local, global
- **Optimization procedure**
- **Object**

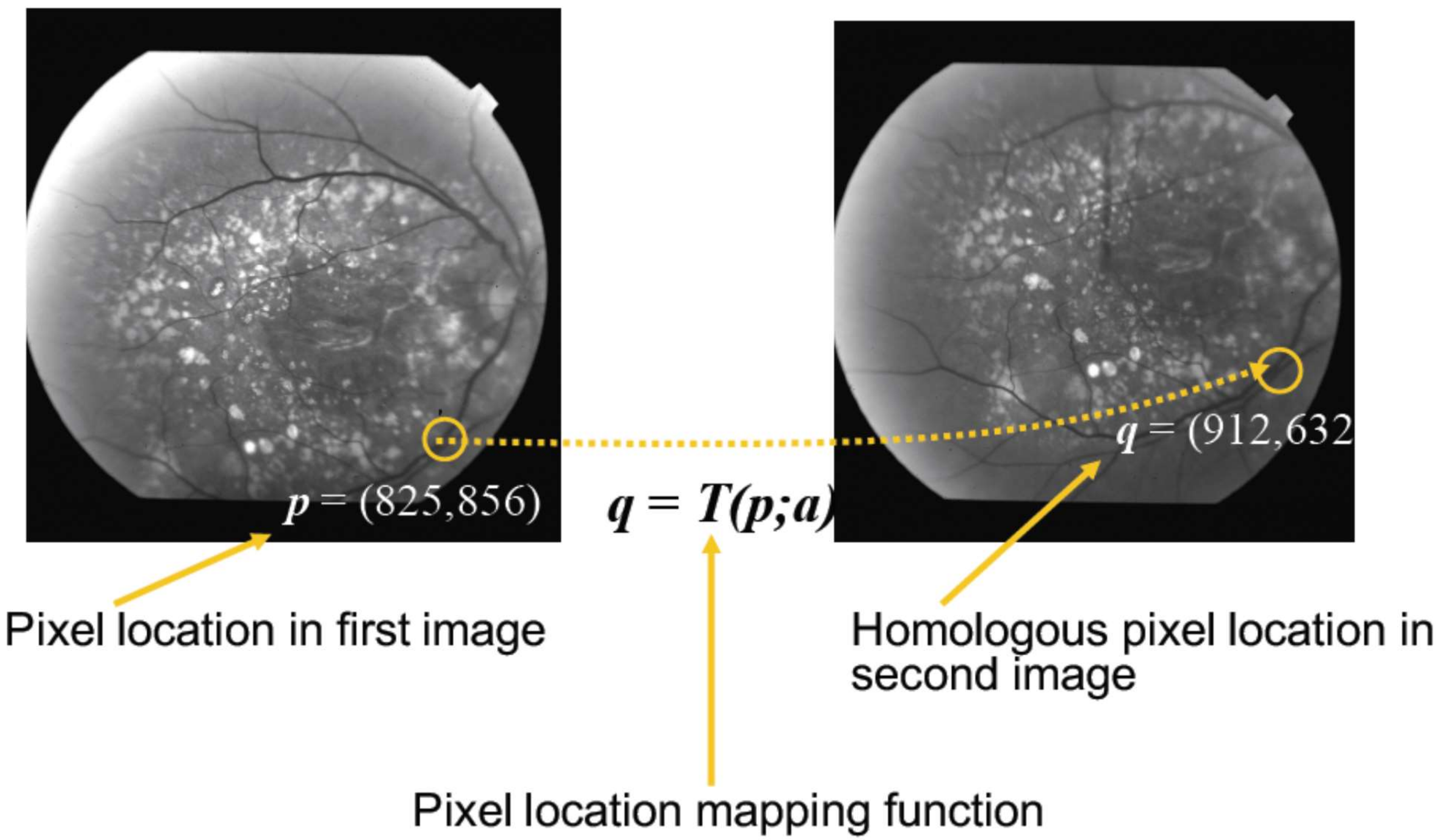
Deformation Models

Method used to find the transformation

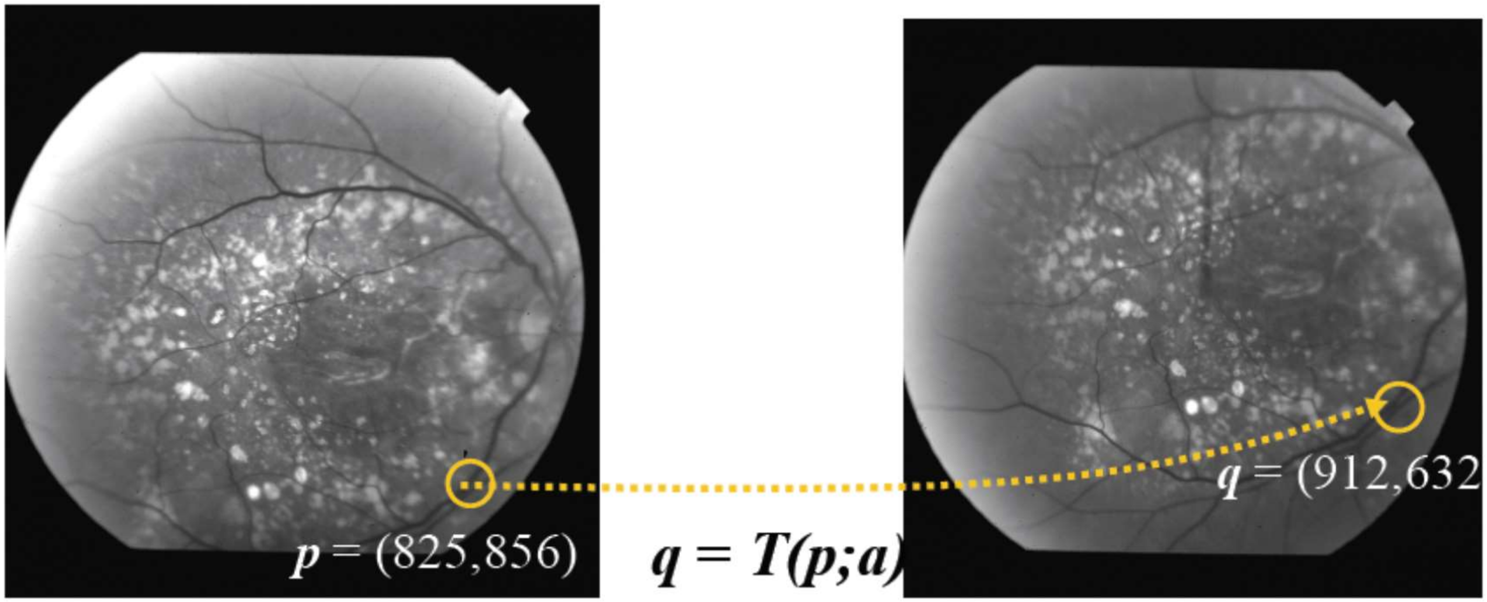
- **Rigid & affine**
  - Landmark based
  - Edge based
  - Voxel intensity based
  - Information theory based
- **Non-rigid**
  - Registration using basis functions
  - Registration using splines
  - Physics based
    - Elastic, Fluid, Optical flow, etc.



Registration is an alignment problem



Registration is an alignment problem



$$\mathbf{p} = (x, y)^T$$

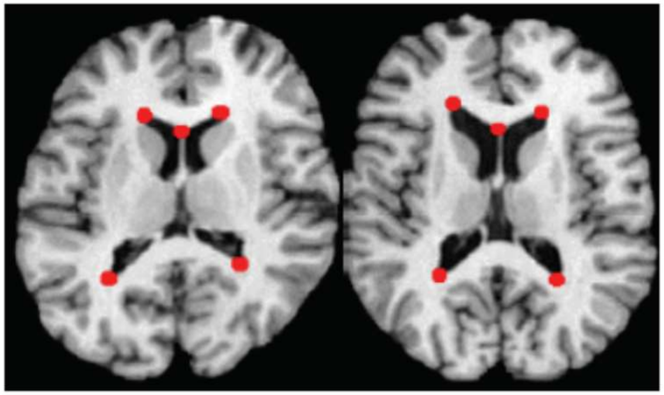
$$\boldsymbol{\Theta} = (s, t_x, t_y)^T$$

$$\mathbf{T}(\mathbf{p}; \boldsymbol{\Theta}) = \begin{pmatrix} sx + t_x \\ sy + t_y \end{pmatrix}$$

Pixel scaling and translation

Landmark Based

- Identifying corresponding points in the images and inferring the image transformation
- Types of landmarks
  - Extrinsic
    - artificial objects attached to the patient
  - Intrinsic
    - internal anatomical structures
- Computing the average or "centroid" of each set of points → translation
- Rotated this point set about the new centroid until the sum of the squared distances between each corresponding point pair is minimized

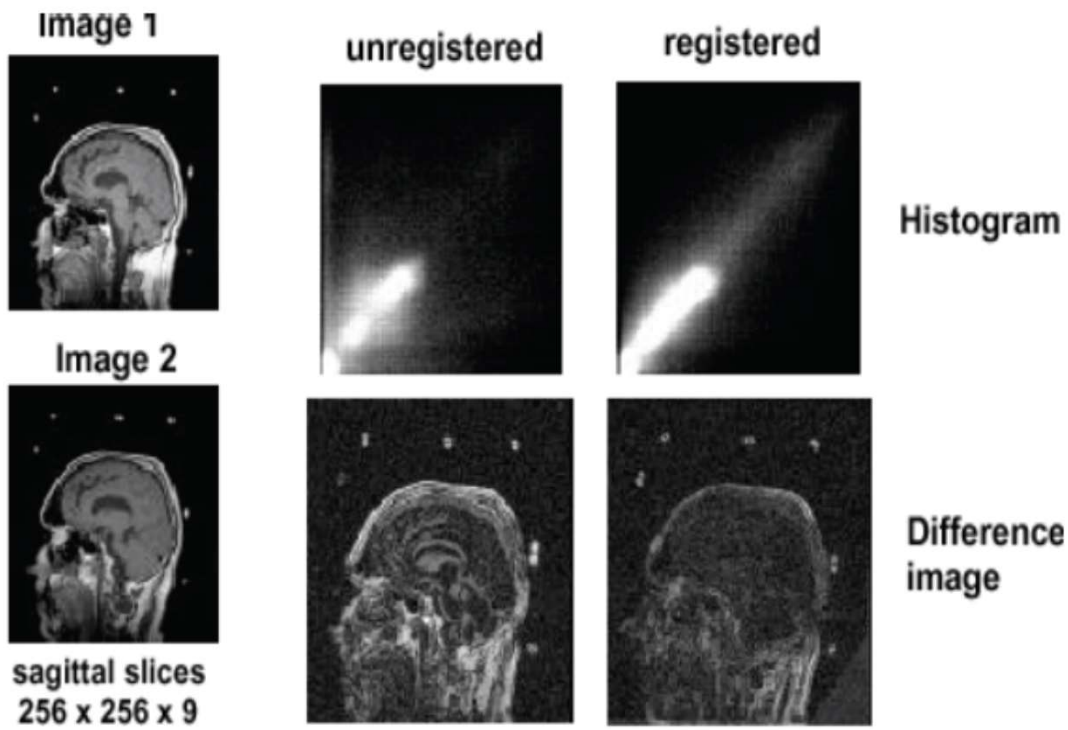


Surface Based

- Method
  - Extracting corresponding surfaces
  - Computing the transformation by minimizing some measure of distance between the two surfaces
- Algorithms used
  - The “Head and Hat” Algorithm
  - The Iterative Closest Point Algorithm
  - *Registration using crest lines*

Intensity Based

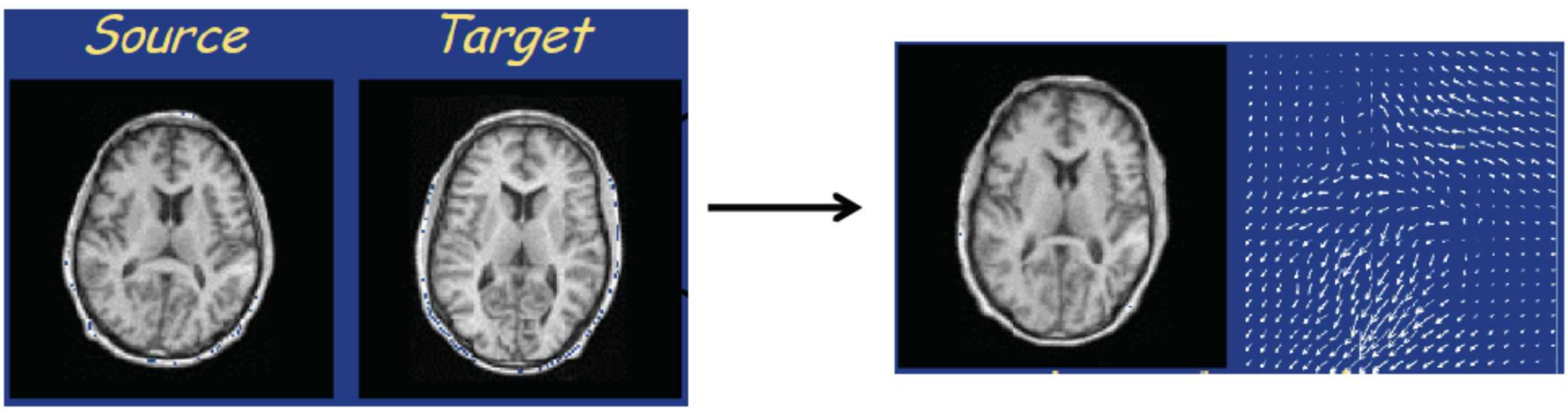
- Method
  - Calculating the registration transformation by optimizing some measure calculated directly from the voxel values in the images
- Algorithms used
  - Registration by minimizing intensity difference
  - Correlation techniques
  - Ratio image uniformity
  - Partitioned Intensity Uniformity



Intensity Based

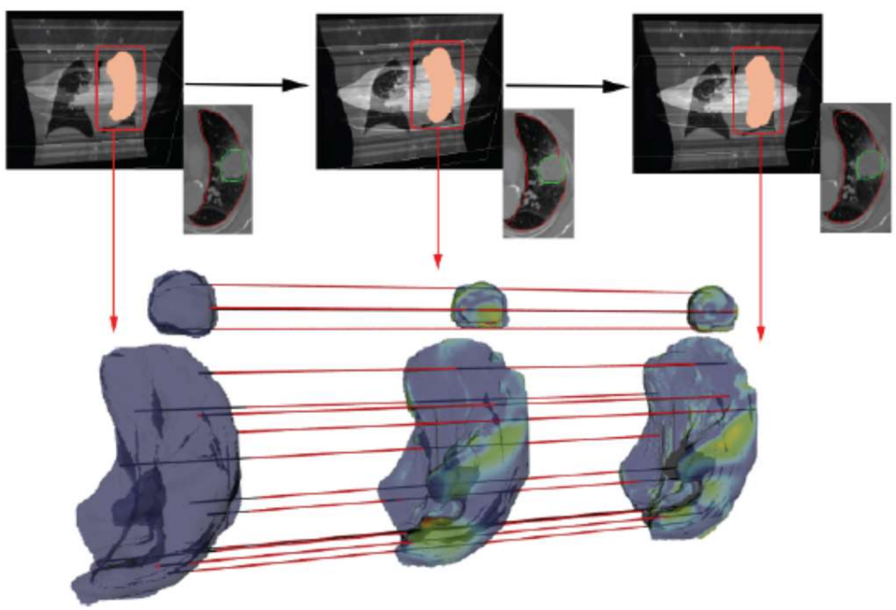
- Intensity-based methods <u>compare intensity patterns in images via correlation metrics</u>

- Sum of Squared Differences
- Normalized Cross-Correlation
- Mutual Information



Feature Based

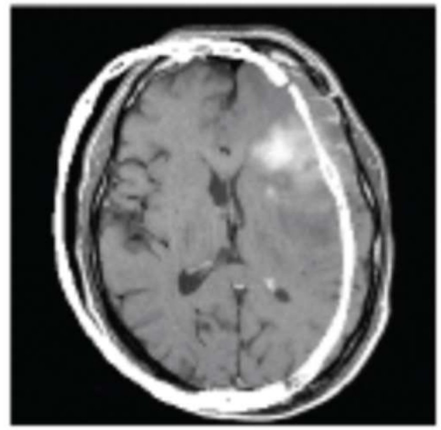
- Feature-based methods find correspondence between image features such as points, lines, and contours.
- Distance between corresponding points
- Similarity metric between feature values
  - e.g. curvature-based registration



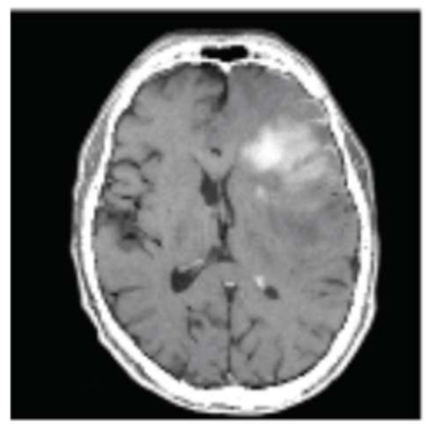
Information Theory Based

- Image registration is considered as to maximize the amount of shared information in two images
  - reducing the amount of information in the combined image
- Algorithms used
  - **Joint entropy**
    - Joint entropy measures the amount of information in the two images combined
  - **Mutual information**
    - A measure of how well one image explains the other, and is maximized at the optimal alignment
  - **Normalized Mutual Information**

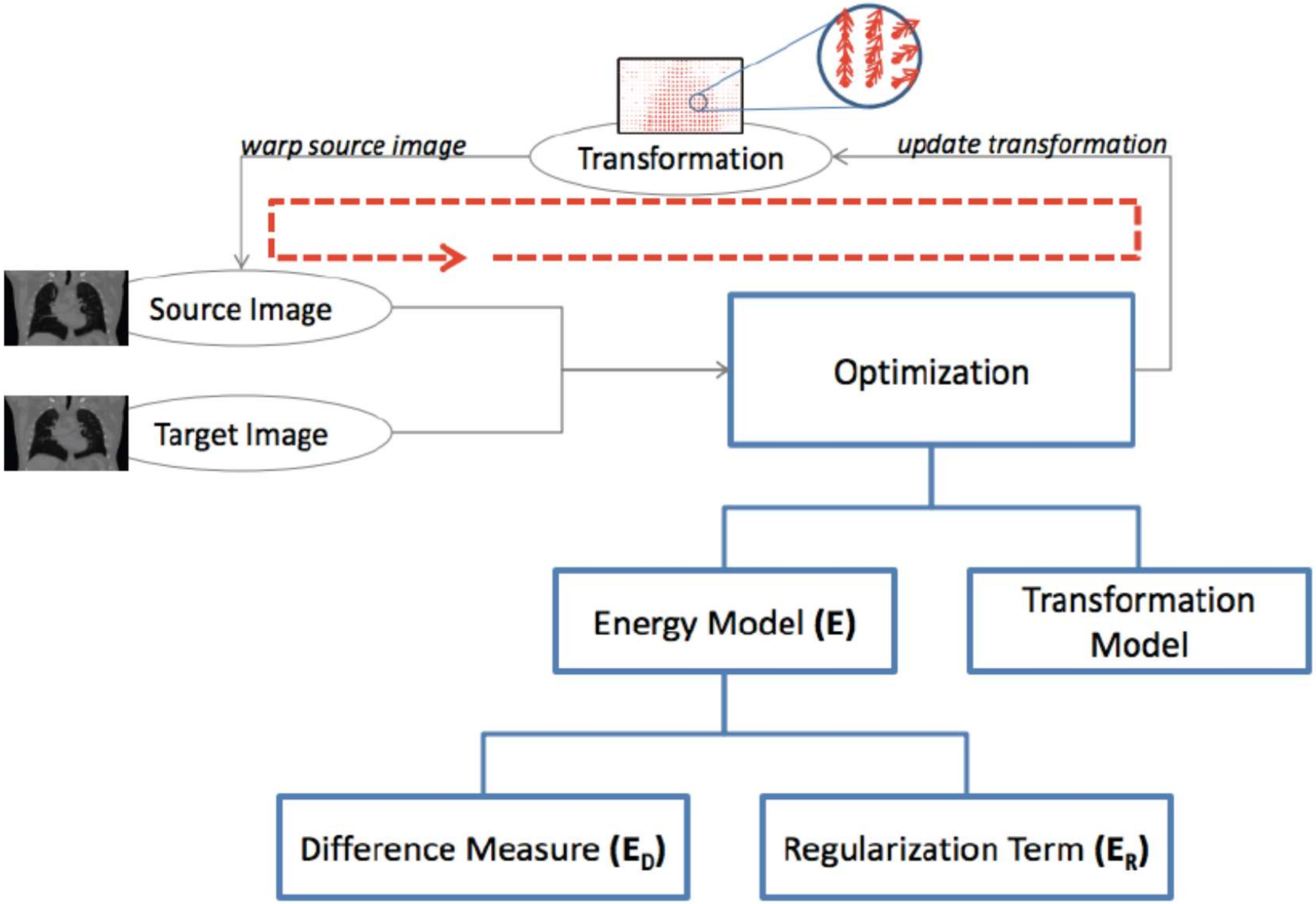
Not registered



Registered

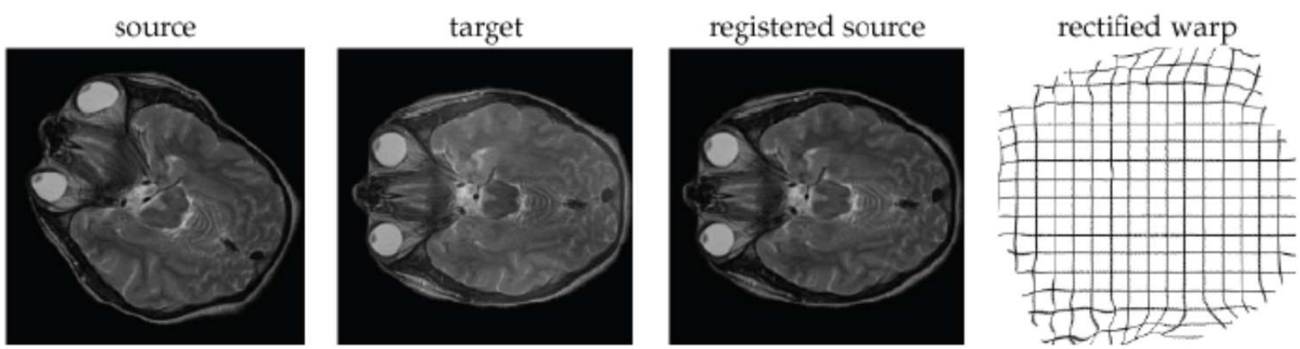


Deformable Registration General Framework

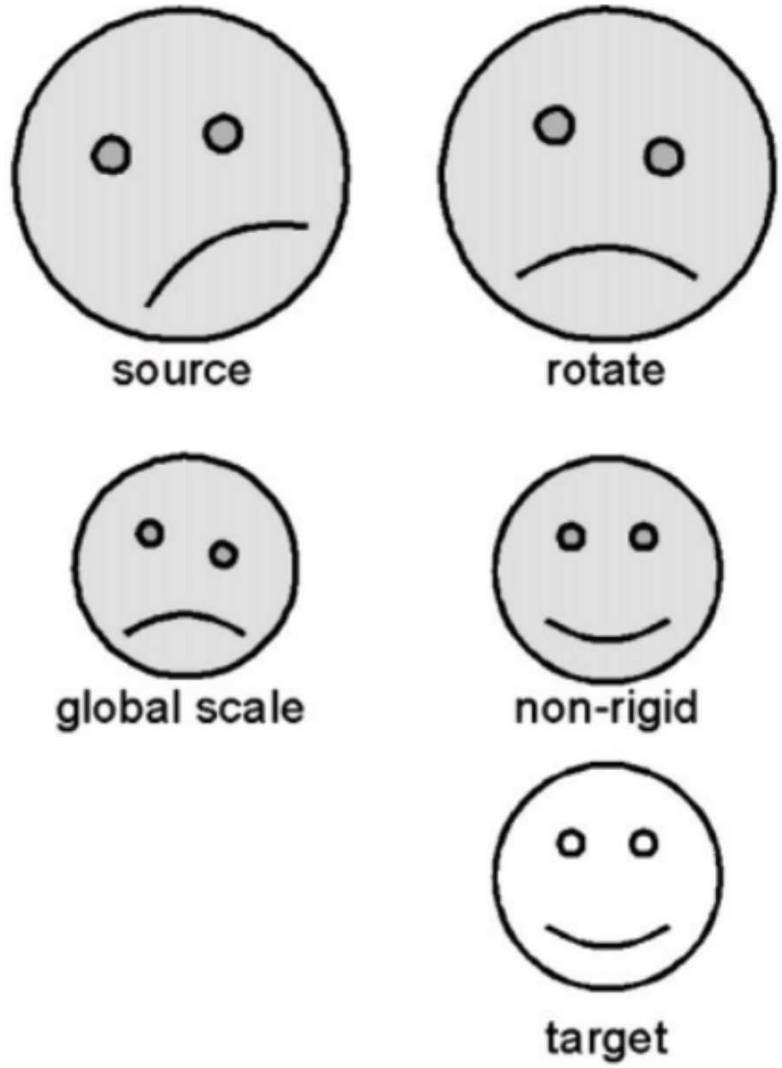


Example: Elastic Registration

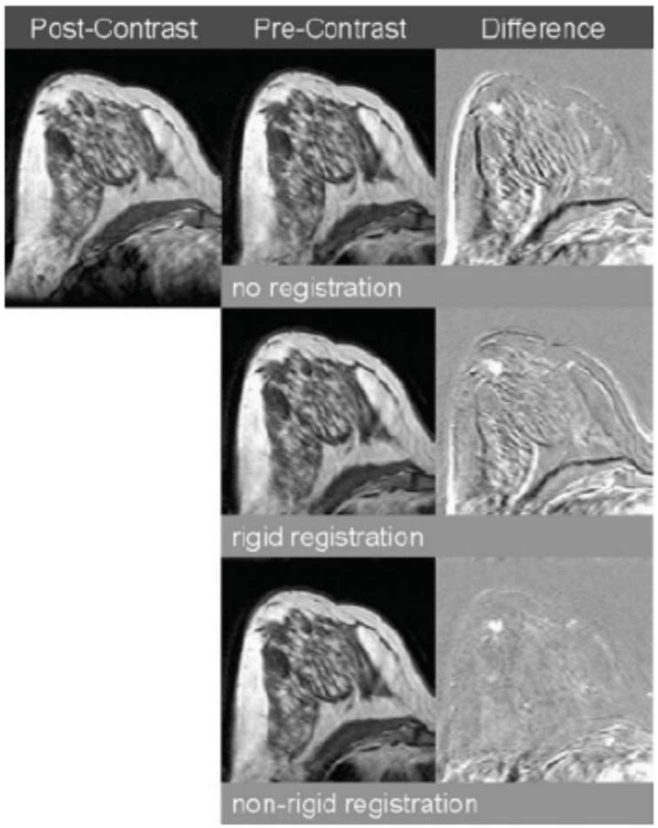
- Model the deformation as a physical process resembling the stretching of an elastic material
  - The physical process is governed by the internal force & external force
  - described by the Navier linear elastic partial differential equation
- The external force drives the registration process
  - The external force can be the gradient of a similarity measure
    - e.g. local correlation measure based on intensities, intensity differences or intensity features such as edge and curvature
  - Or the distance between the curves and surfaces of corresponding anatomical structures.



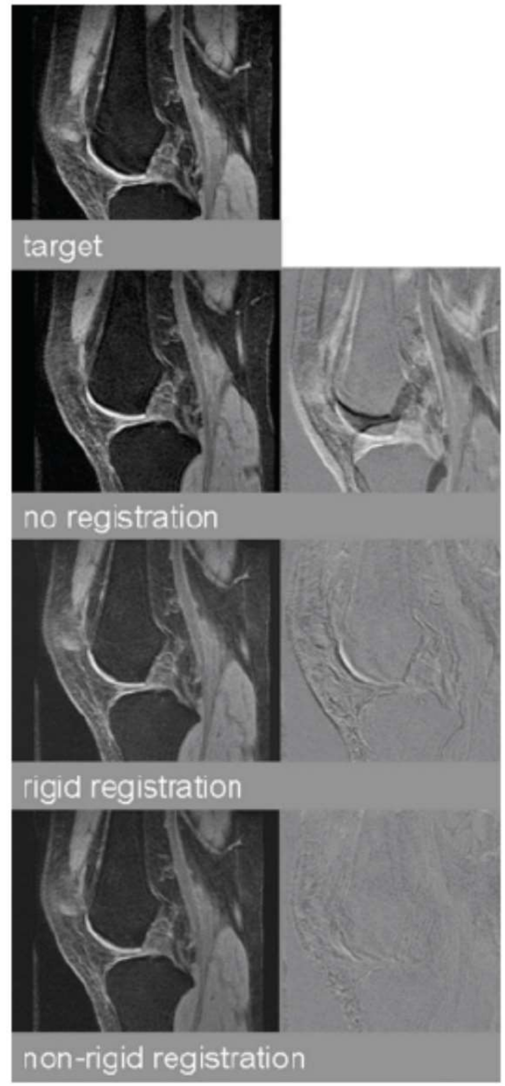
Non-Rigid Registration



Example



Post-Contrast
Pre-Contrast
Difference
no registration
rigid registration
non-rigid registration



target
no registration
rigid registration
non-rigid registration

MERIT FUNCTIONS

- In general, merit functions in registration provide a measure on the similarity of images; therefore, these functions are also referred to as similarity measures or cost functions. In general, these have the following properties:
- They yield an optimum value if two images are aligned in an optimal manner; it is therefore evident that a merit function for intramodal registration may have completely different properties than an merit function for a special intermodal registration problem.
- The capture (or convergence) range of the merit function should be as wide as possible. In other words, the merit function has to be capable of distinguishing images that are only slightly different from those that are completely different, and a well-defined gradient should exist for a range of motion as large as possible.
- Merit functions for image fusion can be defined as intensity-based, or as gradient based.

MERIT FUNCTIONS

- In general, merit functions assume a common frame of reference, and one of the two images "moves" by means of a volume transform; the merit function compares the gray values in the images at the same position in this common frame of reference. Figure illustrates this concept.

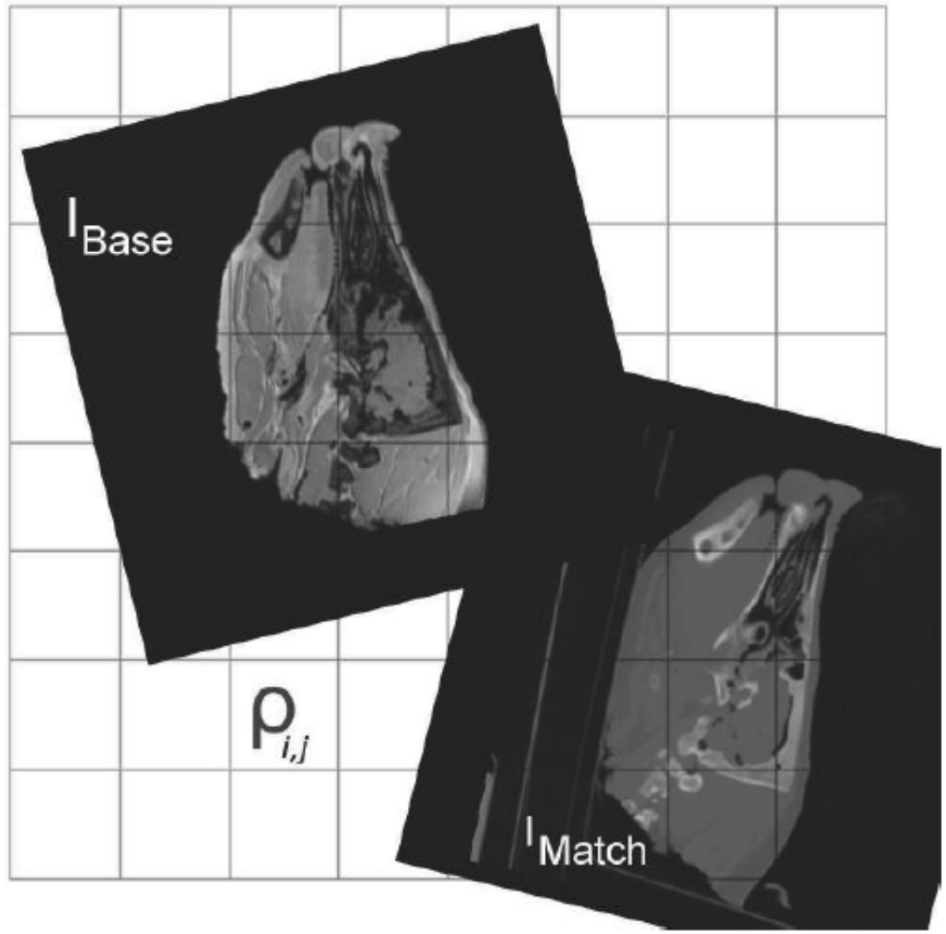


FIGURE: Merit functions compare gray values ρ_{ij} in a common frame of reference. The images I_{Match} and I_{Base} move in this common frame of reference, but the gray value is taken at the same location in the common coordinate system.

MERIT FUNCTIONS

- The intensity-based merit functions are basically statistical measures for giving a measure of mutual dependence between random variables.
- In these cases, the intensity values ρ are considered the random variables, which are inspected at the location of image elements which are assumed to be the same.
- This directly leads to a very straightforward and simple measure, the sum of squared differences:

$$\mathcal{M}_{\mathrm{SSD}} = \frac{1}{N} * \sum^{N} \left(\rho_{\mathrm{Base}}(x, y, z) - \rho_{\mathrm{Match}}(x, y, z)\right)^2$$

- where N is the total number of pixels.
- The inner mechanics of this measure are evident; if two gray values ρ differ, the squared difference will be non-zero, and the larger the difference for all image elements, the larger the sum of all squared differences will be.

MERIT FUNCTIONS

- Another possibility to compare intramodal images is to take a look at the difference image I_{diff} of two images, and to assess the disorder in the result. Such a measure is pattern intensity, given here only for the 2D case:

$$
\begin{aligned}
\mathcal{M}_{\mathrm{PI}} &= \sum_{x,y}^{N} \sum_{d^2 \leq r^2} \frac{\sigma^2}{\sigma^2 + \left(I_{\mathrm{diff}}(x,y) - I_{\mathrm{diff}}(u,v)\right)^2} \\
d &= (x-u)^2 + (y-v)^2 \\
\sigma &= \text{Internal scaling factor}
\end{aligned}
$$

- d is the diameter of the local surrounding of each pixel; if this surrounding is rather homogeneous, a good match is assumed.
- It requires an excellent match of image histogram content; furthermore, it features a narrow convergence range and therefore can only be used if one is already pretty close to the final registration result. The internal scaling factor σ is mainly used to control the gradient provided by MPI.

MERIT FUNCTIONS

- The merit function that can be applied to both inter- and intramodal registration problems is mutual information.
- This measure from information theory gives a measure for the statistical dependence of two random variables which is derived from a probability density function (PDF) rather than the actual gray values ρ given in two images. This sounds rather theoretic, but in practice, this paradigm is extremely powerful. The PDF – we will see later on how it can be obtained – does not care whether a voxel is tinted bright, dark or purple; it just measures how many voxels carry a similar value.
- Therefore, a PDF function can be derived for each image, and the mutual information measure is defined by comparing these functions in dependence of a registration transformation. In general information theory, mutual information (MI) is defined as

$$MI = E\left(P\left(I_{\mathrm{Base}}, I_{\mathrm{Match}}\right)\right) \ln \frac{P\left(I_{\mathrm{Base}}, I_{\mathrm{Match}}\right)}{P\left(I_{\mathrm{Base}}\right) P\left(I_{\mathrm{Match}}\right)}$$

where E (P (I_{Base}, I_{Match})) is the expectation value of the joint PDF – noted as P (I_{Base}, I_{Match}) of the two images IBase and IMatch, and P (IBase) and P (IMatch) are the PDF for the gray values in the single images.

MERIT FUNCTIONS

- PDF is a function that gives the probability to encounter a random variable x in sampling space. For a random variable that follows the normal distribution, the PDF is given as the Gaussian. For an image, a nice approximation of a PDF is the histogram,
- The joint histogram is a 2D function that maps from a 2D-domain, namely the gray values of the images IBase and IMatch, to a statistic that measures the occurrence of a similar gray value at the same pixel (or voxel).
- Note that the gray values need not be the same; the basic idea of mutual information lies in the fact that areas with similar content will have a similar distribution of gray values.
- If we define a measure that becomes optimal if the joint histogram is as well-organized as possible, we can assume an optimum match of the images IMatch and IBase.

MERIT FUNCTIONS

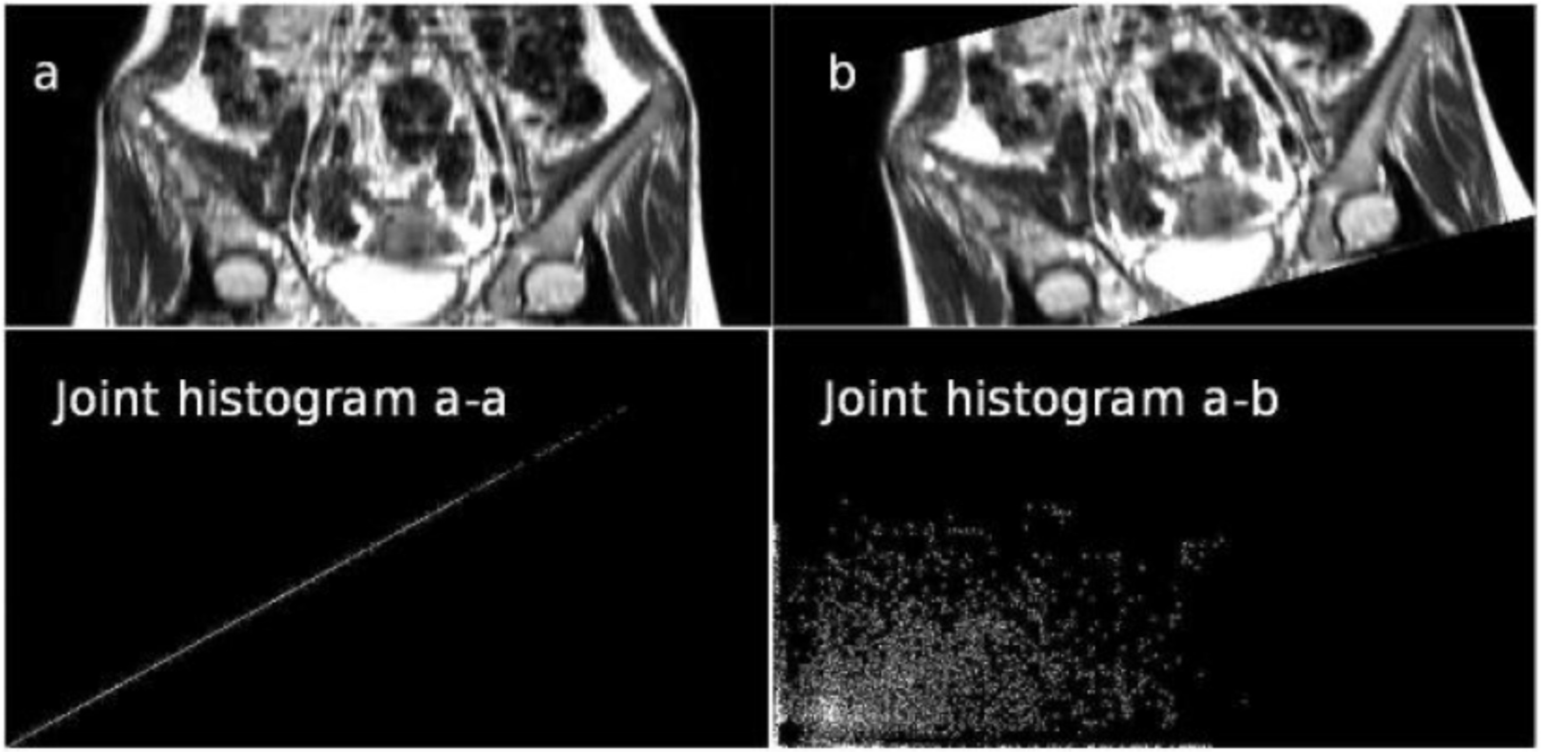


FIGURE: The output from the JointHistogram_9.m script. In the upper row, we see two MR images, which can be found in the LessonData folder and are named OrthoMR (a) and OrthoMRTransf (b). They are the same images, but they differ by a small rotation and translation. The lower images show the output of the JointHistogram_9.m script. 2 – first, the joint histogram of image a and itself is computed (lower left). We see a low level of disorder in the histogram since all gray values are well correlated and aligned nicely on the line of identity in the joint histogram. The joint histogram of images a and b look worse; the common occurrence of similar gray values is scarce, and the cells with high counts in the histogram are more scattered (lower right). Disorder – measured by Shannon’s entropy – is higher here.

MERIT FUNCTIONS

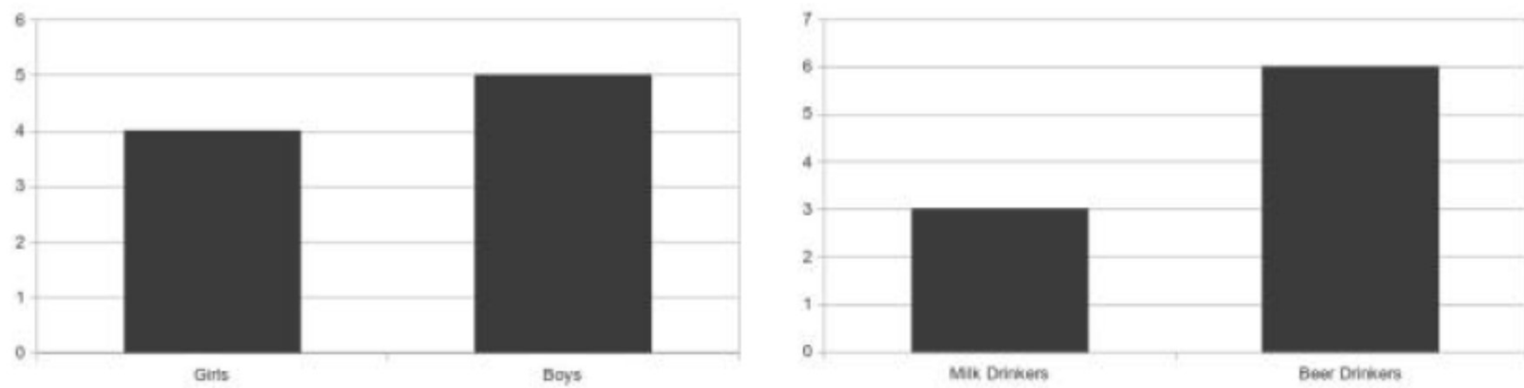


FIGURE : A statistical example for illustrating the joint histogram. Out of a group of nine people, four are female. Furthermore, three group members prefer non-alcoholic beverages.

MERIT FUNCTIONS

A joint histogram of this statistical collective is given in Table

	Girls	Boys	Total
Milk drinkers	2	1	3
Beer drinkers	2	4	6
Total	4	5	9

TABLE: The 2 x 2 table for our simple example of girls and boys, drinking beer or milk. The central part of this table is a joint histogram of the two properties,

MERIT FUNCTIONS

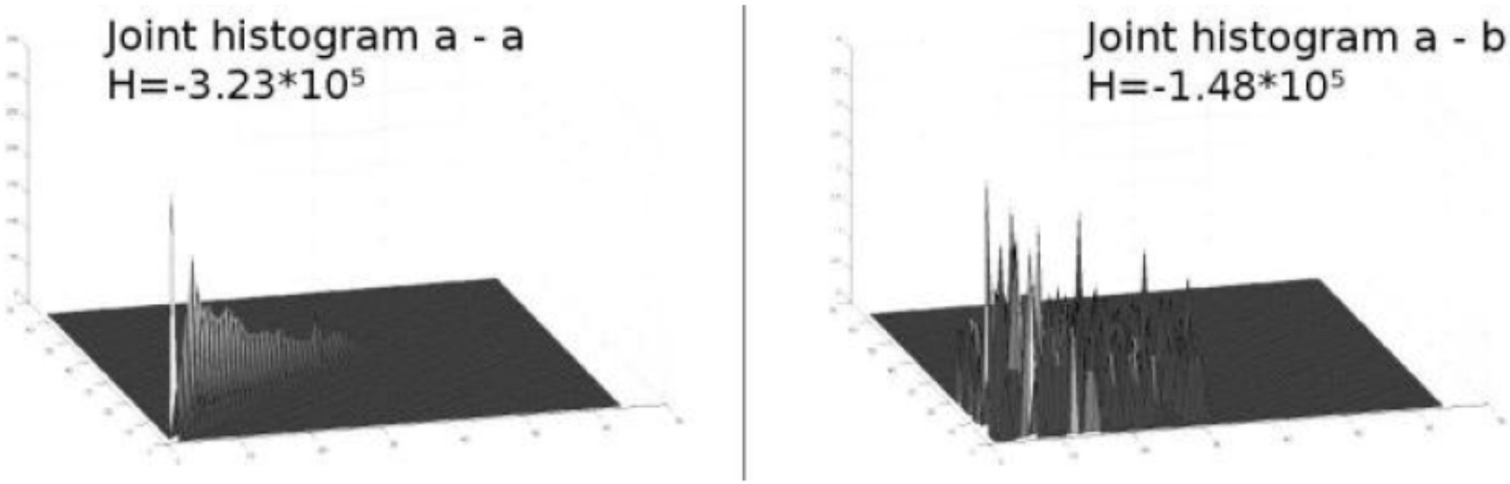


FIGURE: Surface plots of the joint histograms given in Figure 9.6. The first histogram – which is the joint histogram of the MR image a from Figure 9.6 with itself – shows that the histograms are not scatterplots. Rather than that, different bins show different heights. The second surface plot shows the joint histogram of the MR image and its rotated counterpart (image b from Figure 9.6). Apparently, more bins are non-zero, but the overall height is much lower – note that the z-axes have a different scale. The bin height ranges from 0 to 350 in the left plot, whereas the bin height in the right plot does not exceed 4. Shannon's entropy as given in Equation 9.4 was also computed for both histograms. It is considerably lower for the well-ordered left histogram, as expected.

MERIT FUNCTIONS

Application of rotation,

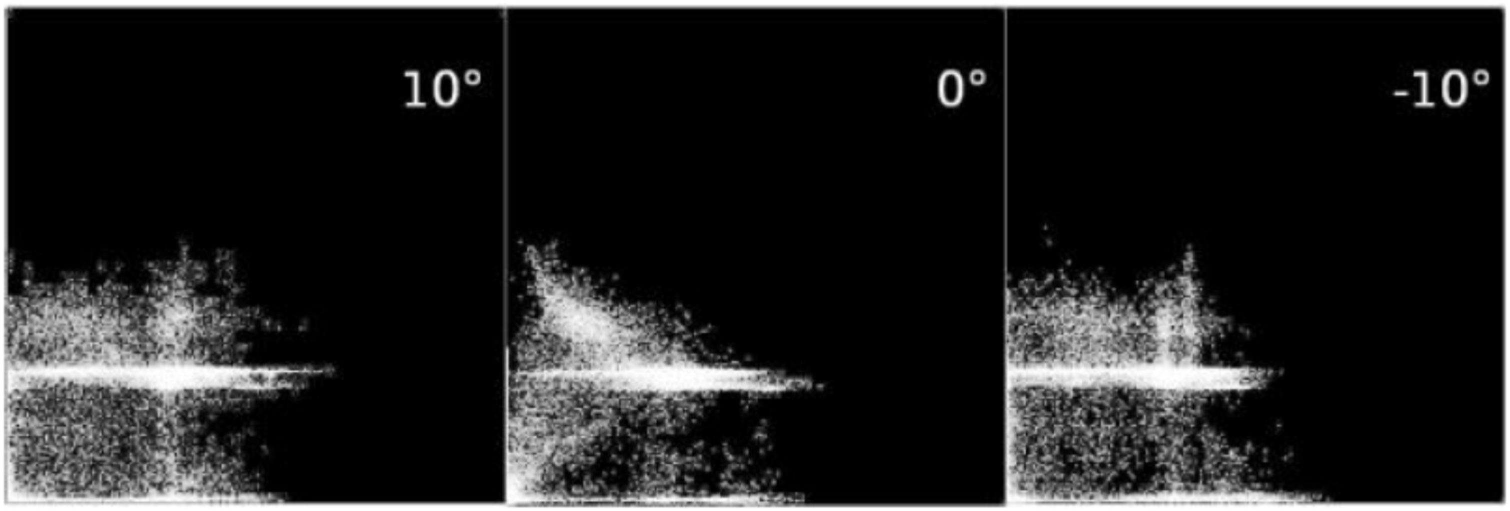


FIGURE: The joint histogram of two registered slices from the LessonData folder, CT.jpg and T1.jpg. The images were rotated against each other with an angle of -10°, 0°, and 10°; while the joint histogram for the optimum case (rotation angle 0°, middle image) does not look like the lower left image in Figure, since the images are taken from different modalities, it is optimal in terms of disorder, as one can see when computing the mutual information as a merit funct

MERIT FUNCTIONS

The joint histograms for overlap and translations in y-direction of -100 and 100 pixels

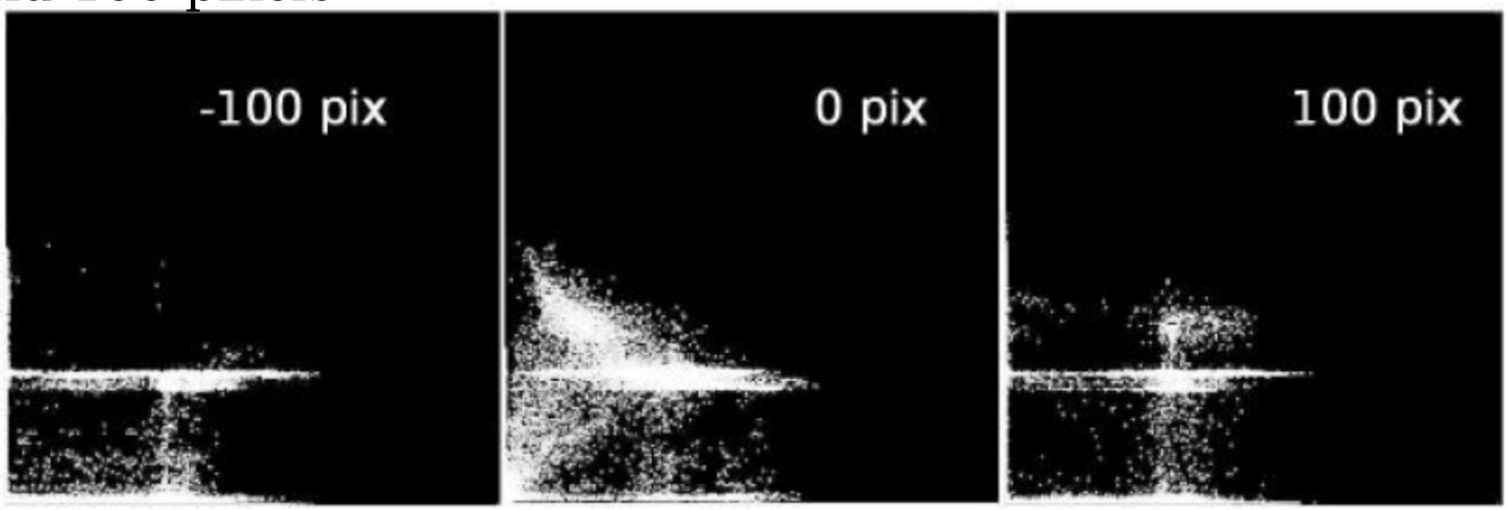


FIGURE: Another example for the joint histogram of two registered slices from the LessonData folder, CT.jpg and T1.jpg. This time, one of the slices was translated by -100 pixels, and by 100 pixels. Since the overlap of the images is less than in the case of image rotation, the disorder in the joint histogram is increased, resulting in visibly less densely populated areas in the joint histogram.

MERIT FUNCTIONS

- a well-ordered histogram is characterized by the fact that most points in the joint histograms are accumulated in a few areas.
- This is independent of the range of gray values – a scaling operations would just stretch the joint histogram, or it would move the densely populated areas to other positions in the joint histogram.
- A measure for this disorder in a PDF P, which we will call entropy2 from now on, is Shannon's entropy

$$H = -\sum_i P_i \ln(P_i)$$

Equation is a good measure for entropy

MERIT FUNCTIONS

- A joint histogram with densely populated bright areas (that is, bins with a high count of occurrences) show a lesser degree of chaos than a sparsely populated joint histogram with a “dark” appearance.
- A high count – that is a bright spot in the joint histogram – increases H, but the logarithm – a function that grows at a very slow rate for higher values.
- The sum of counts in two joint histograms of high and low entropy should be the same.
- As a consequence, the histogram with low entropy has only a few non-zero elements, and the high value of Pi is damped by the term ln (Pi).
- The chaotic histogram, on the other hand, has many summands with a low value Pi and this value is not damped by the logarithmic term.

MERIT FUNCTIONS

For the sample of beer- and milk-drinking boys and girls, we can also compute Shannon's entropy:

$$H = -(2\ln(2) + \ln(1) + 2\ln(2) + 4\ln(4)) = -8.317$$

	Girls	Boys	Total
Milk drinkers	0	3	3
Beer drinkers	4	2	6
Total	4	5	9

TABLE: An alternative 2 x 2 table for beer- or milk-drinking girls and boys. While the total numbers stay the same, this distribution shows lesser disorder since we can make quite likely conclusions about the other property from a given property.

MERIT FUNCTIONS

- define a merit function that utilizes mutual information

$$\mathcal{M}_{\mathrm{MI}} = H\left(I_{\mathrm{Base}}\right) + H\left(I_{\mathrm{Match}}\right) - H\left(I_{\mathrm{Base}}, I_{\mathrm{Match}}\right)$$

- the mutual information merit function (M_{MI})
- Equation is a measure for the common information in two images, which is optimal if the two images are aligned.
- The true power ofM_{MI} lies in the fact that it does not care about gray values; it retrieves all information from the PDF, which is basically given by the histogram. Therefore, it is a truly multimodal merit function.

MERIT FUNCTIONS

- The shape of the merit function can be found in Figure

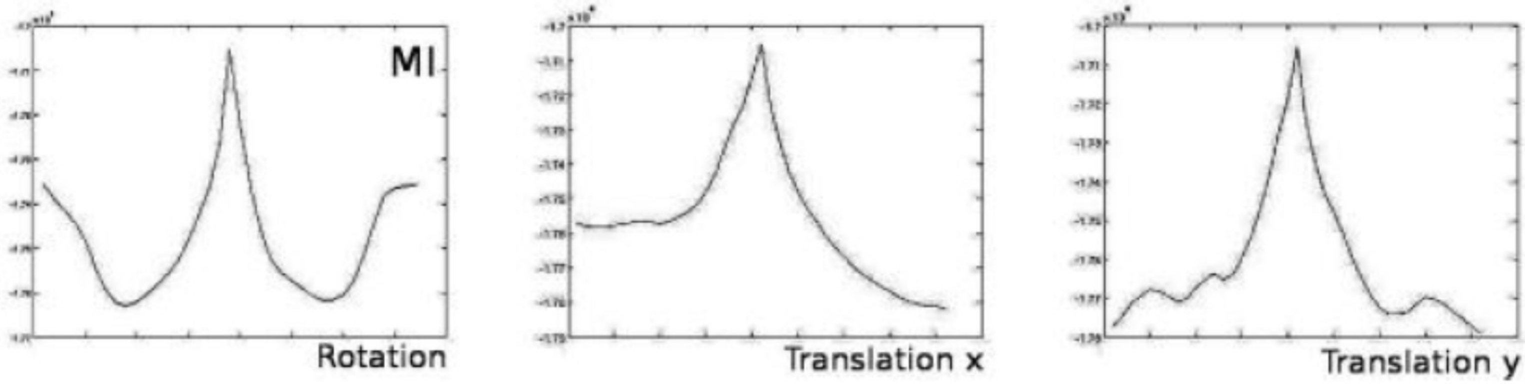


Figure : The shape of M_{MI} is plotted when changing all three dof of rigid motion in 2D separately for the two sample images

MERIT FUNCTIONS

Many more refinements of merit functions for multimodal image registration exist, for instance it is possible to derive the PDF using the so called Parzen-window method.

- The basic idea – minimization of entropy in a joint PDF of the images – however, always stays the same.
- For the sake of cc d the definition of a variation of M_{NMI} mation

$$\mathcal{M}_{\mathrm{NMI}} = \frac{H\left(I_{\mathrm{Base}}\right) + H\left(I_{\mathrm{Match}}\right)}{H\left(I_{\mathrm{Base}}, I_{\mathrm{Match}}\right)}.$$

- Multimodal matching always necessitates that the gray value ρ in corresponding image elements is replaced by a more general measure; the joint entropy in PDFs derived from images is one possibility. Another possibility is the use of gradient information in an image.
- If we can design a merit function based on these gradient images, we may have another merit function for multimodal image fusion at hand. Such a measure is chamfer matching; it uses a binary image for I_{Match} that was made using a Sobel filter and intensity thresholding

MERIT FUNCTIONS

- The chamfer matching merit function MCM is simply defined as:

$$\mathcal{M}_{\text{CM}} = \sum_{x,y} D\left(I_{\text{Base}}(x,y)\right)$$

$$x, y \quad \ldots \quad \text{Pixels in } I_{\text{Match}} \text{ with } \rho > 0$$

$$D\left(I_{\text{Base}}\right) \quad \ldots \quad \text{Distance transform of } I_{\text{Base}}$$

- Equation can be understood as follows; $D(I_{Base})$ – the distance transform – defines grooves in the binary gradient image I_{Base}, and if I_{Match} “drops” into those grooves,MCM becomes optimal.
- This merit function has a severe drawback – it only works well if we are already close to the solution of the registration problem.

MERIT FUNCTIONS

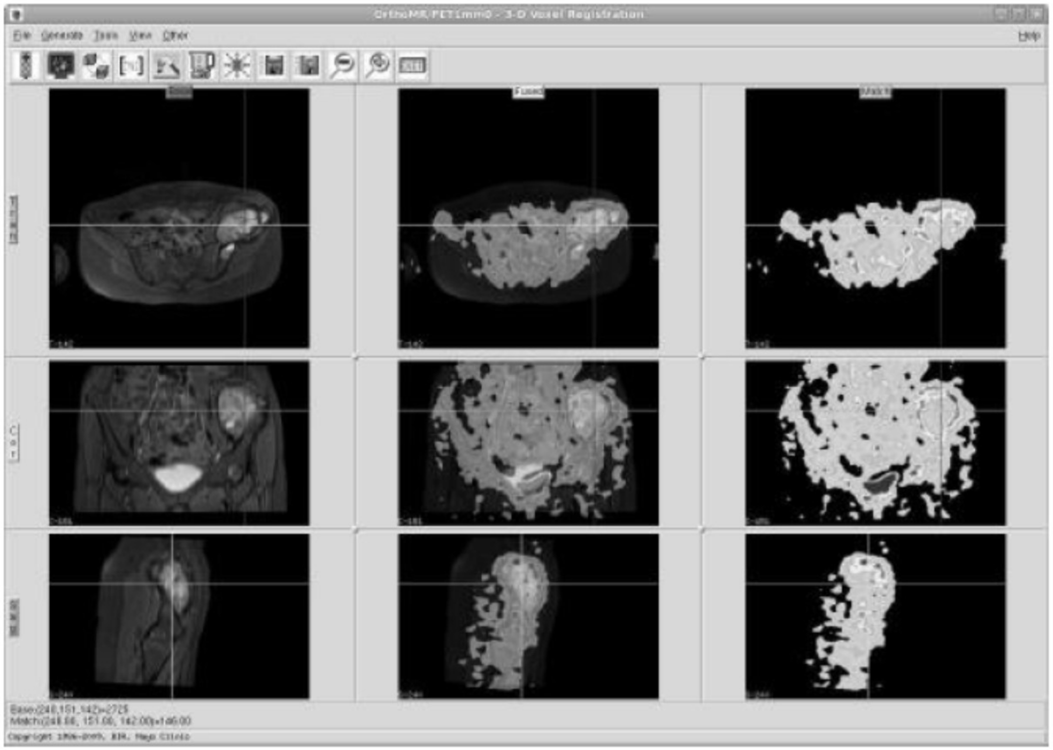


FIGURE: A more complex clinical example of multimodal image registration. Here, a PET and an MR scan of a sarcoma in the pelvis was registered. Since PET provides very little anatomic information, a direct registration of MR and PET data in the abdominal area often fails. However, it is possible to register the MR data with the CT data that came with the PET-scan from a PET-CT scanner using normalized mutual information; in this case, the implementation of AnalyzeAVW was used. The registration matrix obtained is applied to the PET data since these are provided in the same frame of reference as the CT data.

MERIT FUNCTIONS

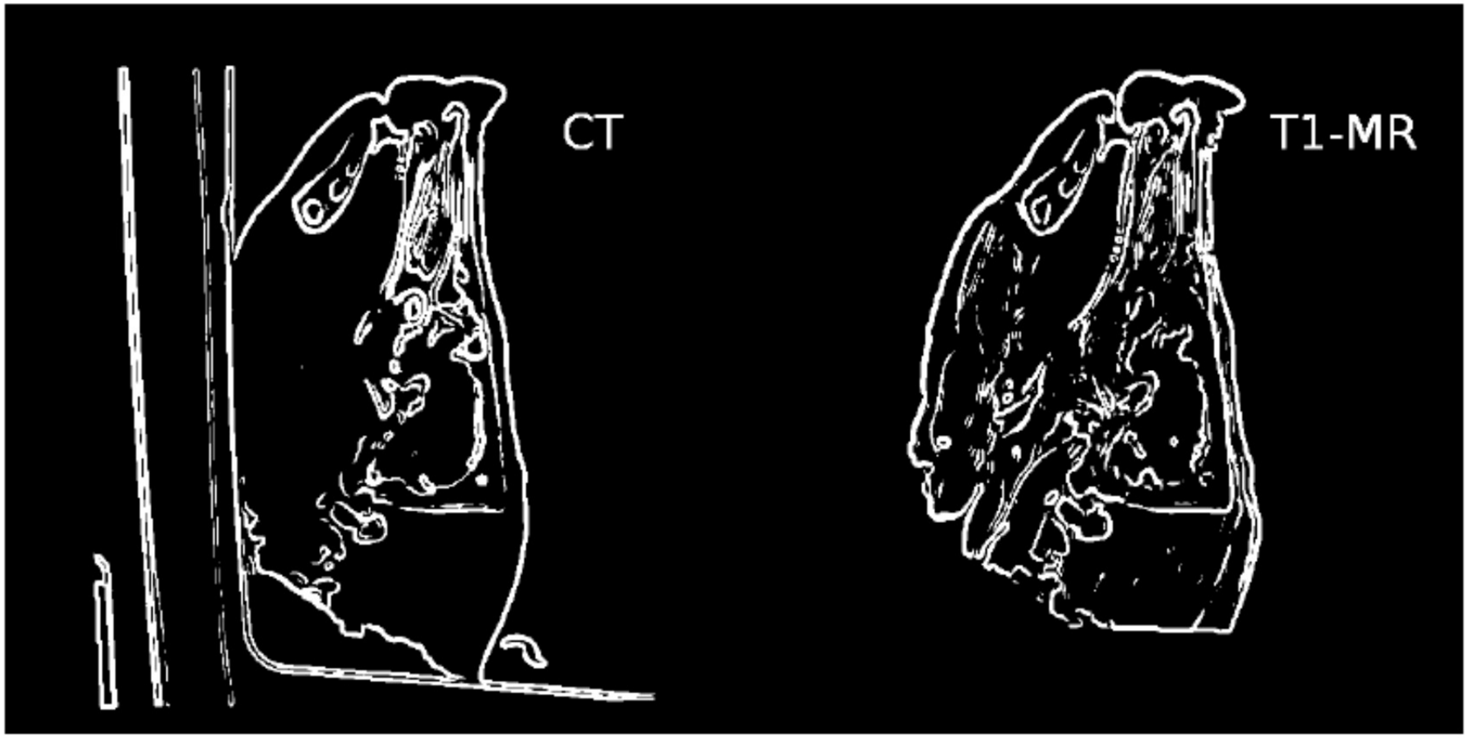


FIGURE: CT.jpg and T1.jpg after Sobel-filtering, optimization of image depth and thresholding. After this process, the differences in gray values in the two images disappear, and only gradient information remains. Applying a gradient-based measure to such an image allows for multi-modal image registration without computing joint histograms or using an intensity based measure.

MERIT FUNCTIONS

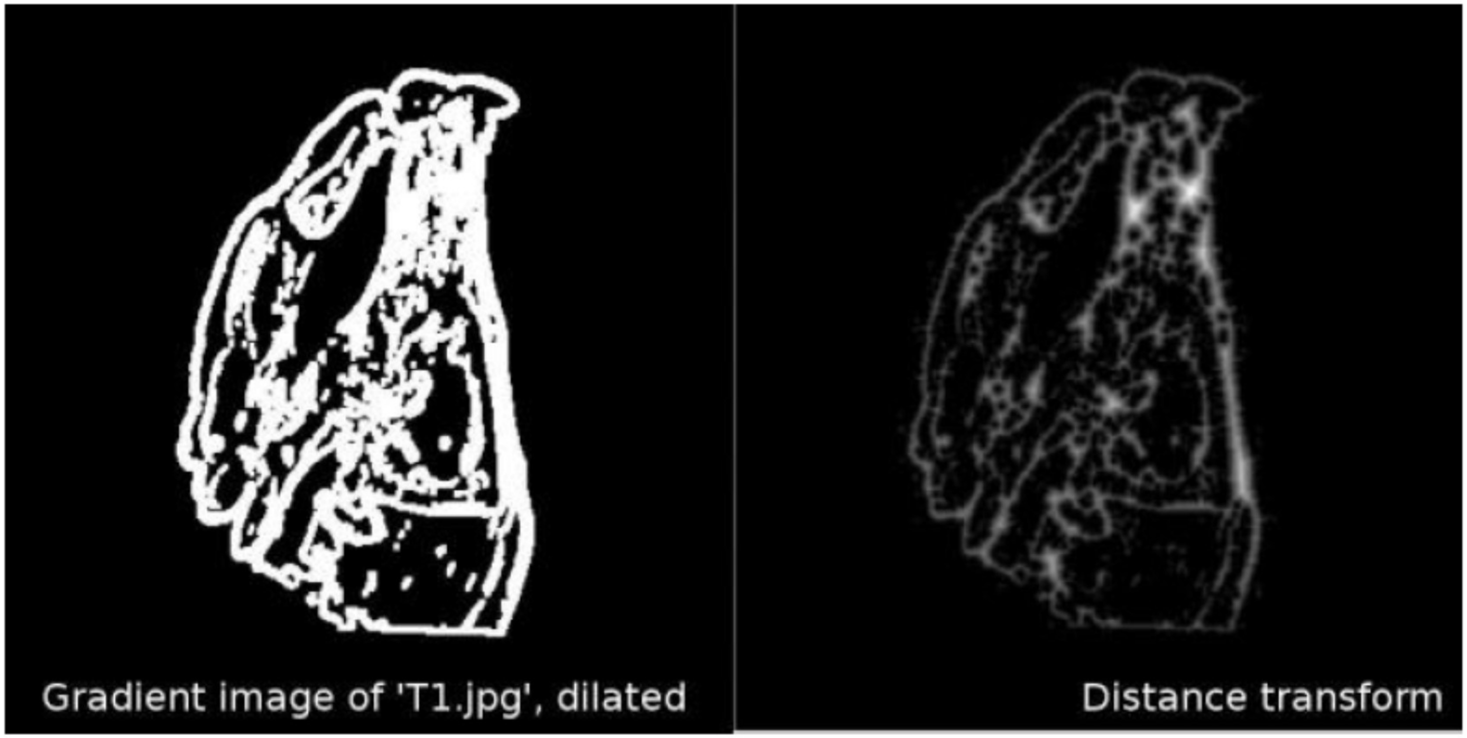


FIGURE: T1.jpg after Sobel-filtering, intensity thresholding, and dilation (left). This image can be found in the LessonData folder as T1_EdgeImageDilated.jpg. The right image shows its distance transform (DT_ChamferMatch.jpg in LessonData). In the chamfer-matching algorithm, the sum of the entries in the distance transform of I_{Base} at the location of non-zero pixels in I_{Match} is computed while I_{Match} moves over the base image.

MERIT FUNCTIONS

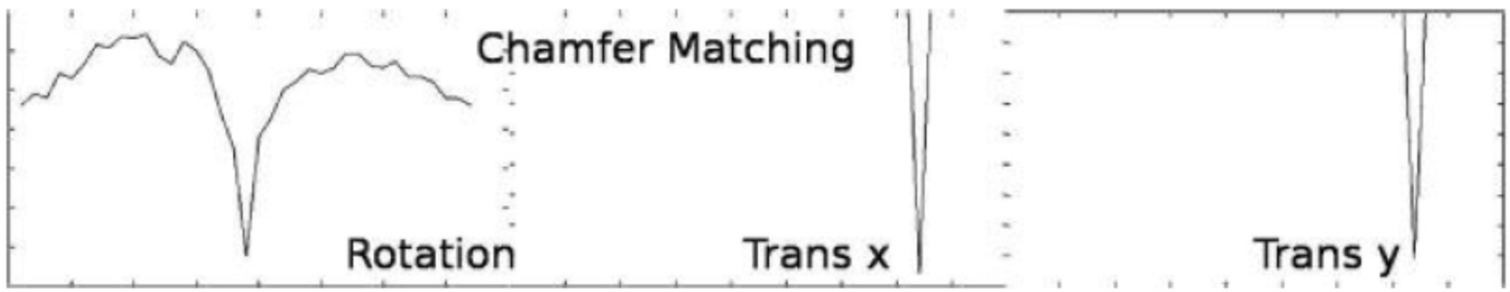


FIGURE: This is a plot of the chamfer matching meritfunction in 3 dof through the solution of the registration problem. As we can see, numerous local minima exist and for translation, there is no wide gradient. Therefore it is necessary to start the optimization algorithm close to the expected solution, since otherwise, finding the minimum is unlikely.

OPTIMIZATION STRATEGIES

- Optimization algorithms can be categorized as local and global algorithms.
- The algorithm starts at a given position and tries to follow the gradient of the merit function until it reaches an optimum.
- Every merit function showing a maximum can be turned over by multiplying it with -1.
- The merit function is only given on discrete pivot points.
- The mathematical counterpart of the saladbowl/marble allegory is the simplex algorithm, also called Nelder-Mead method.
- Here, the marble is replaced by a simplex, which can be considered an n-dimensional triangle.
- If our simplex is one-dimensional and is supposed to find the minimum of a simple one dimensional function $f(x) = y$, it is a segment of a straight line; this simplex crawls down the function $f(x)$, and its step width is governed by the gradient of $f(x)$.
- Once the simplex encounters a minimum, it reduces its stepwidth until it collapses to a point.
- Figure illustrates this algorithm in a humble fashion.

OPTIMIZATION STRATEGIES

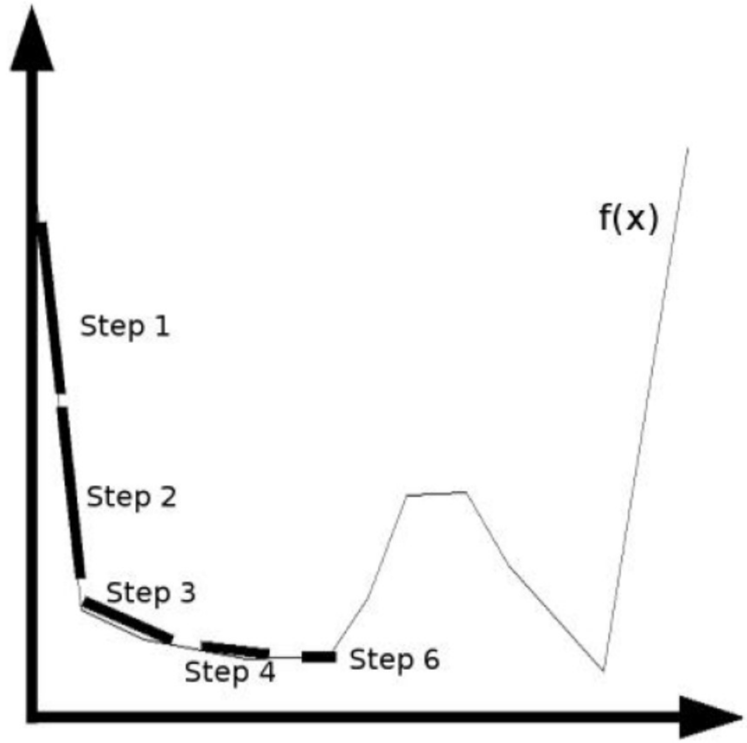


FIGURE: A simple illustration of the Nelder-Mead or simplex algorithm, a local optimization algorithm; a simplex in one dimension is a section of straight line that follows the gradient of a function; as the simplex approaches a minimum, $\frac{d}{dx} f(x) \simeq 0$ true, and the simplex makes smaller steps until it collapses to a point since the gradient to the right is ascending again and the algorithm terminates. It cannot find the global minimum to the right, but it is stuck in the local minimum.

OPTIMIZATION STRATEGIES

- The simplex-algorithm is one of many optimization methods; other algorithms like
- *Powell's method*, *Levenberg-Marquardt, conjugate gradient* and so on are well-documented in numerous textbooks. All these algorithms are *local optimizers* – they are good at finding the obvious nearest minimum, and some of them require the computation of derivative functions, whereas others do not.

OPTIMIZATION STRATEGIES

- Global optimizers try to tackle the problem of finding global optima.
- The simplest type of a global optimization method is a full search in the parameter space with decreasing step width.
- Imagine you have a set of N starting parameters $\vec{p}$, and a gross area you want to explore.
- The area is claimed by defining a starting interval pi;next, one computes all values of the merit function $\mathcal{M}(\vec{p}')$ for all possible combinations of parameters
- $\vec{p}' \in \{(p_1 - \Delta p_1, p_2, p_3, \ldots p_N), \ldots (p_1, \ldots, p_N + \Delta p_N)\}$
- The result is an array of merit function values, and one can easily find the minimum in this array.
- In a subsequent step, the starting parameter set $\vec{p}$ is replaced by the parameters that yield the smallest value, and one continues with a smaller set of interval $\Delta p'$ and so on.

OPTIMIZATION STRATEGIES

- challenge in the evaluation of registration algorithms
  - target registration error (TRE) and fiducial registration error (FRE)
- optimizing for parameters which are available in different units.
  - Rotations are given in radians or degrees, and translations are given in voxels, pixels, or millimeters.

How to achieve good registration results?

- Initialization: it is advisable to start the registration close to expected solution; one may, for instance, position the volumes or images manually. Or it is possible to move the volumes to the same center of gravity – a principal axes transform may also help to obtain an initial guess for the relative rotation of the images.
- Appropriate choice of the merit function: The nature of the registration problem governs the choice of the merit function. In general, one may consider other statistical measures than mutual information if a intramodal registration problem is given. Gradient based methods tend to have a smaller capture range than intensity–based methods.
- Appropriate choice of optimization methods: Global optimizers tend to require lots of computational time while a correct result is not guaranteed. One may try a good implementation of a local algorithm first. Sometimes, a multiscale approach is proposed. Here, the volumes are scaled down and the registration problem is handled on this coarse stage first before proceeding to finer resolution. This approach is, however, not always successful – it may as well lead to local optima and considerable runtime.

How to achieve good registration results?

- Avoid additional degrees-of-freedom: It is of course possible to match images of different scale using an additional scaling transformation, and of course one could always try to carry out a deformable registration in order to avoid errors from changing soft tissue. Don’t do this – additional dof increase the risk of winding up in local optima.
- Restriction to relevant image content: This is an often underestimated aspect. It is very helpful to constrain the area in IBase used for merit function validation. This does not only refer to the spatial domain, but also to intensities. A lower bound in intensity can, for instance, improve robustness by suppressing noise and structures that stem from the imaging modality such as the table in CT. And it can be helpful to exclude parts of the body which moved relative to each other, where it is advisable to exclude the femora since these are mobile relative to the pelvis.

CAMERA CALIBRATION

In camera calibration, the inverse problem is being treated – how can we determine both the 3D-position of an object and the properties of the camera – that is, any projection system including x-ray machines and head-mounted displays – from the projected image?

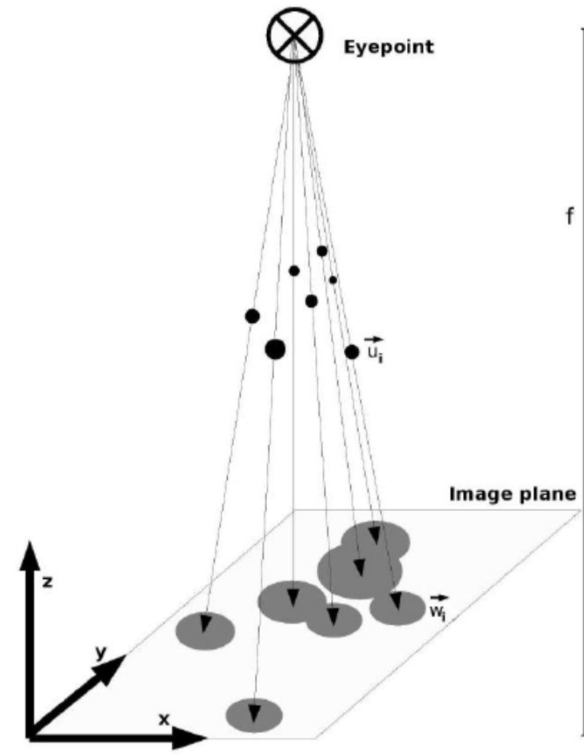


FIGURE An illustration of the camera calibration problem. In this example, the imaging plane is located in the x-y plane of a Cartesian coordinate system. The normal distance of the eyepoint – the source of projection – is located at a distance **f** from the imaging plane. The known coordinates of the fiducial markers $\vec{u}_i$ are projected onto the plane, and their coordinates are given as $\vec{w}_i$. The task is to determine **f** and an affine transformation that locates the eyepoint in the coordinate system.

CAMERA CALIBRATION

In order to tackle the camera calibration problem, we have to remember another matrix, which is the projector P. In mathematics, each matrix that fulfills the requirement $P = P^2$ is called a projection. Let's go to Figure – the effect of $P\vec{u}_i$ is the 2D image of the markers. If we project this image again, nothing will change. A projection matrix that maps 3D positions $\vec{u}_i$ to 2D positions $\vec{w}_i$ to the x-y plane of a Cartesian coordinate system where the eyepoint is located at the the z-axis at a distance **f** is given as:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\mathbf{f}} & 1 \end{pmatrix}$$

One can easily verify that P is a projector indeed. The projector operates on a point $\vec{u}_i$ by multiplication. If we want to apply the volume transformation V on that point prior to projection, we have to compute $PV\vec{u}_i = \vec{w}_i$. You may remember that, in order to change the position of the eyepoint by another volume transformation V', we have to compute $V'PV\vec{u}_i = \vec{w}'$. Since we deal with a projection in homogeneous coordinates here, we have to add a rescaling step to $\vec{w}'$ to get coordinates $\vec{w}$ – remember that the fourth element of vector giving homogeneous coordinates has to be 1, which is not the case when applying an operator like P.

CAMERA CALIBRATION

In the camera calibration problem, we know a set of N planar projected coordinates $\vec{w}_i \in \{\vec{w}_0 \dots \vec{w}_N\}$ and the associated 3D coordinates $\vec{u}_i \in \{\vec{u}_0 \dots \vec{u}_N\}$. We have to find out about V and $\mathbf{f}$. A simple algorithm to achieve this is called the *Direct Linear Transform* (DLT).

$$D\vec{u} = \vec{v}$$

It is, however, necessary to renormalize the resulting vector $\vec{v}$ by computing $\vec{w} = \frac{\vec{v}}{v_4}$. We omit the z-variable since it is of no importance for the projected screen coordinates; therefore we have a resulting pair of screen coordinates:

$$\vec{w}' = \begin{pmatrix} w_1 \\ w_2 \\ 0 \\ 1 \end{pmatrix}.$$

All we have to find out is the twelve components of matrix D; by expanding the matrix

CAMERA CALIBRATION

we can rewrite Equation
$$\begin{aligned} D_{11}u_1 + D_{12}u_2 + D_{13}u_3 + D_{14} &= w_1 * v_4 \\ D_{21}u_1 + D_{22}u_2 + D_{23}u_3 + D_{24} &= w_2 * v_4 \\ D_{41}u_1 + D_{42}u_2 + D_{43}u_3 + D_{44} &= v_4 \end{aligned}$$

twelve unknowns $D_{11}, D_{12}, D_{13}, D_{14}, D_{21}, D_{22}, D_{23}, D_{24}, D_{41}, D_{42}, D_{43},$ and D_{44}. These can be rewritten as a 12×2 matrix:

$$\begin{pmatrix} u_1 & u_2 & u_3 & 1 & 0 & 0 & 0 & 0 & -w_1u_1 & -w_1u_2 & -w_1u_3 & -w_1 \\ 0 & 0 & 0 & 0 & u_1 & u_2 & u_3 & 1 & -w_2u_1 & -w_2u_2 & -w_2u_3 & -w_2 \end{pmatrix} \begin{pmatrix} D_{11} \\ D_{12} \\ D_{13} \\ D_{14} \\ D_{21} \\ D_{22} \\ D_{23} \\ D_{24} \\ D_{41} \\ D_{42} \\ D_{43} \\ D_{44} \end{pmatrix} = \vec{0}$$

REGISTRATION TO PHYSICAL SPACE

- it is necessary to map a coordinate system defined in physical space to an image coordinate system.
- Application examples include the registration of a patient, whose coordinate system is defined by a rigidly attached tracker probe, to an MR or CT-volume.
- Another example includes the registration of a robot's internal frame-of-reference to a patient, who is either fixated by molds or other devices, or who is tracked as well.
- A few techniques for registration to physical space are as follows.

Rigid registration using fiducial markers and surfaces

- If intensity based algorithms fail, one may resort to explicit landmarks, which can either be prominent anatomical features, or explicitly attached so-called fiducial markers.
- This is an extremely important technique, not only in image processing, but also in image-guided therapy, where a match between patient and image data is to be achieved.
- Figure shows the so-called Vogele-Bale-Hohner mouthpiece, a non-invasive device for registration using fiducial markers.
- The mouthpiece is attached to the patient by means of a personalized mold of the maxillary denture.
- A vacuum pump evacuates the mold, therefore the mouthpiece is held in position in a very exact manner.
- The strength of this setup is certainly fusion of anatomical image data and SPECT or PET datasets, where an intensity-based algorithm often fails due to a lack of common information in the images.

Rigid registration using fiducial markers and surfaces

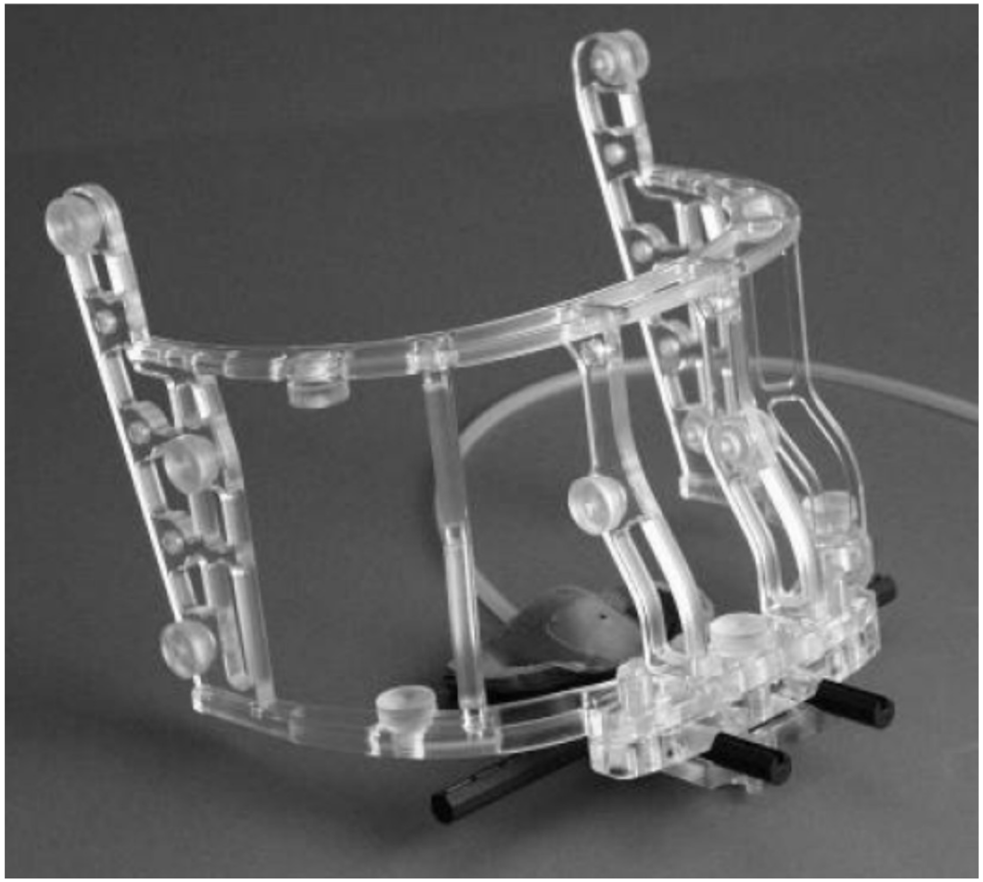


FIGURE : The Vogele-Bale-Hohner mouthpiece, a device for non-invasive registration using fiducial markers. The mouthpiece is personalized by means of a dental mold. A vacuum pump ensures optimum fit of the mouthpiece during imaging.

Rigid registration using fiducial markers and surfaces

- The coordinates of the markers match each other.
- If the transformation matrix is composed as a rotation followed by a translation (V = TR) one can immediately determine three dof of translation by computing the centroid of the marker pairs according to Equation in both frames of reference.
- The difference of the two centroids gives the translation vector.
- After translation of the marker positions so that the centroids of the point pairs $\{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N\}$ and $\{\vec{p}'_1, \vec{p}'_2, \ldots, \vec{p}'_N\}$ match, a rotation matrix R that merges the positions best can be determined if the markers are not collinear – in other words, the points must not lie on a line. And we do need at least three of them.
- Let P_{Base} be the set of coordinates in the base reference frame, and P'_{Match} be the matching set of points in the coordinate system to be registered after they were translated to the centroid. In this setup, the center of rotation is the centroid of the base point set.

Rigid registration using fiducial markers and surfaces

- One of the most widely used algorithms for matching ordered point sets was given by Horn. It matches two ordered point sets P_{Base} and P_{Match} by minimizing

$$\mathcal{M} = \frac{1}{2}\sum_{i=1}^{N} \|\vec{p}_{\mathrm{Base}_i} - V\vec{p}_{\mathrm{match}_i}\|^2$$

- where V is the well-known affine volume transformation matrix.
- After computing the centroids $\vec{\bar{p}}_{\mathrm{Base}}$ and $\vec{\bar{p}}_{\mathrm{Match}}$, one can define a 4 × 4 covariance matrix C of the point sets as

$$C = \frac{1}{N}\sum_{i=1}^{N} \left(\vec{p}_{\mathrm{Match}_i} - \vec{\bar{p}}_{\mathrm{Match}}\right)\left(\vec{p}_{\mathrm{Base}_i} - \vec{\bar{p}}_{\mathrm{Base}}\right)^T$$

Rigid registration using fiducial markers and surfaces

- Next, an anti-symmetric matrix $A = C - C^T$ is formed from C; the vector Δ is given by $\Delta = (A_{23},\ A_{31},\ A_{12})^T$ Using this vector, one can define a 4 × 4 matrix Q:

$$Q = \begin{pmatrix} \operatorname{tr} C & \Delta^T \\ \Delta & C + C^T - \operatorname{tr} C \mathbf{I}_3 \end{pmatrix}$$

- tr C is the trace of matrix C, the sum of all diagonal elements. I_3 is the 3×3 identity matrix.
- If we compute the eigenvectors and eigenvalues of Q and select the eigenvector that belongs to the biggest eigenvalue, we have four quaternions.
- A rotation matrix R can be derived. These are the parameters of the optimum rotation that match the point sets. The translation is derived by computing:

$$\vec{t} = \frac{1}{N} \sum_{i=1}^{N} \vec{p}_{\mathrm{Match}_i} - R\vec{p}_{\mathrm{Base}_i}$$

Rigid registration using fiducial markers and surfaces

- Matching two meshes is a widespread problem, for instance in image-guided therapy.
- If one segments a body surface from a CT- or MR-scan, it is possible to generate a mesh of the resulting surface, consisting of node points.
- Another modality such as a surface scanner, which produces a 3D surface by scanning an object using optical technologies for instance, can produce another mesh of the patient surface in the operating room.
- Surface data can also be collected, for instance, by tracked ultrasound probes, where a 3D point is assigned to a 2D point in the US-image by tracking the scanhead using some sort of position measurement device.
- By matching the surface as scanned and the segmented surface, it is possible to achieve a match of the patients position in the treatment suite to the volume image; in such a case, the node points of the meshes are matched by using a point-to-point registration technique, but the point sets are not ordered.
- This method is called iterative closest point (ICP) algorithm,

Rigid registration using fiducial markers and surfaces

- In the ICP-algorithm, the distance d between a single point $\vec{p}$ and a surface S is defined as $d(\vec{p}, S) = \min_{\vec{x} \in S} \|\vec{x} - \vec{p}\|$
- The result of this search for closest points is an ordered set of corresponding points from S which are closest to the points $\vec{p}$ in our matching surface, which is given as a set of points $\vec{p}$.
- What follows is a point-to-point registration as shown before.
- The next step is to repeat the procedure, with a new set of closest points as the results.
- This procedure can be repeated until a termination criterion is met – usually this is the minimal distance between closest points.
- The ICP-algorithm, which had numerous predecessors and refinements, also lies at the base of surface-based registration algorithms. If one tries to match two intramodal datasets, a feasible way is to segment surfaces in both volumes is to segment these surfaces, create a mesh, and apply an ICP-type of algorithm.
- This is generally referred to as surface registration.

2D/3D registration

- Registration of the patient to image data is not confined to point- or surface-based methods; it can also be achieved by comparing image data taken during or prior to an intervention with pre-interventional data.
- An example of such an algorithm is the registration of a patient to a CT-dataset using x-ray data. If one knows the imaging geometry of an x-ray device, for instance by employing a camera-calibration technique, If the exact position of the x-ray tube focus is known relative to the patient (for instance by using a tracker, or from the internal encoders of an imaging device), one can derive six dof of rigid motion by varying the six parameters for the registration transformation.
- The fact that x-ray is a central projection with a perspective is of the utmost importance in this context since a motion to and from the x-ray tube focus changes the scale of the projection.
- In a 2D/3D registration algorithm, DRRs are produced iteratively, and a merit function is used to compare the DRR to the actual x-ray image.
- Once an optimum match is achieved, the registration parameters are stored and give the optimum transform. This type of registration algorithm is usually more complex and sometimes tedious compared to 3D/3D registration since the optimization problem is not well defined. Multimodal and non-rigid 2D/3D registration are also open challenges.

2D/3D registration

Figure shows a reference x-ray, a DRR before registration and the overlay of the two

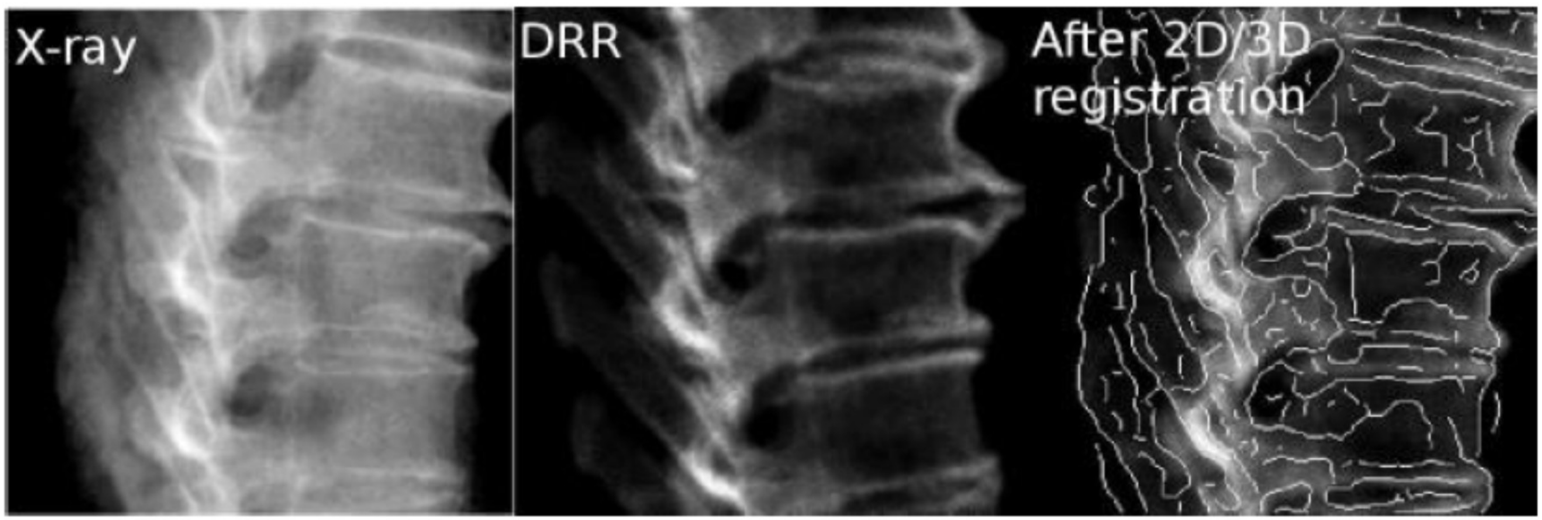


FIGURE : Three images of a spine reference-dataset for 2D/3D image registration. We see the reference x-ray on the left, the DRR generated from the initial guess, and them registration result. An edge image of the x-ray is overlaid over the final DRR after 2D/3D registration.

Thank you