

Mutual Information Based Registration of Medical Images

Pluim et al: Survey

Mattes et al: CT/PET Registration

Background

- Mutual information-based registration was proposed by Viola and Wells (MIT) in 1994-5.
- It has become commonplace in many clinical applications.
- It comes from information theory: the Shannon entropy

$$H = \sum p_i \log (1/p_i) = -\sum p_i \log p_i$$

- The more rare an event, the more meaning is associated with its occurrence

ENTROPY

- Entropy comes from information theory. The higher the entropy the more the information content.

- Entropy =
$$\sum_i -p_i \log_2 p_i$$

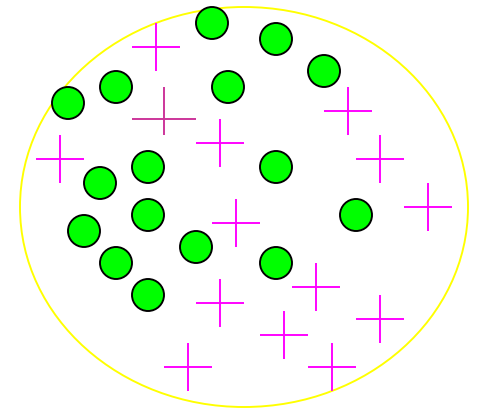
p_i is the probability of event i

Compute it as the proportion of event i in the set.

16/30 are green circles; 14/30 are pink crosses

$\log_2(16/30) = -.9$; $\log_2(14/30) = -1.1$

Entropy = $-(16/30)(-.9) - (14/30)(-1.1) = .99$



2-Class Case:

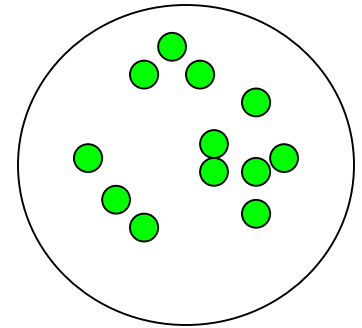
- What is the entropy of a group in which all examples belong to the same class?

- entropy = $-1 \log_2 1 = 0$

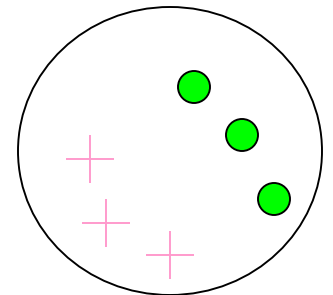
- What is the entropy of a group with 50% in either class?

- entropy = $-0.5 \log_2 0.5 - 0.5 \log_2 0.5 = 1$

**minimum
entropy**



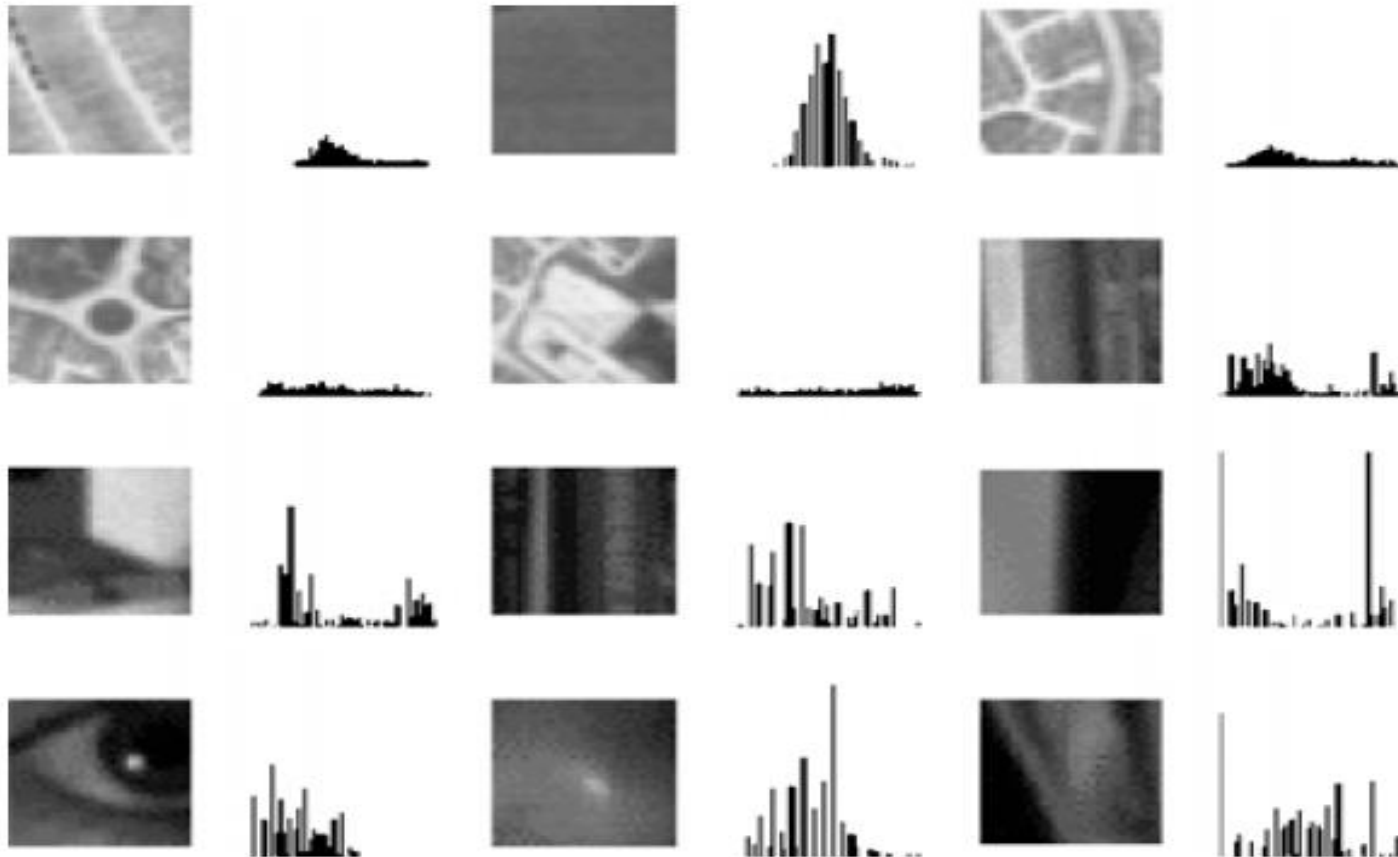
**maximum
entropy**



Entropy for Images

- Shannon entropy can be computed for a gray-tone image.
- It then focuses on the distribution of the gray tones.
- An image consisting of almost a single intensity will have low entropy.
- An image with roughly equal quantities of different gray tones will have high entropy.

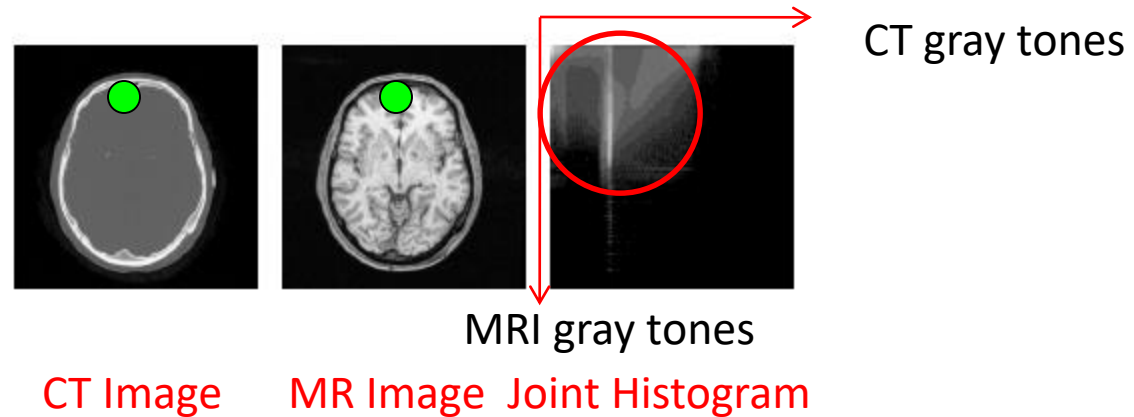
Histograms of Image Intensity



Mutual Information

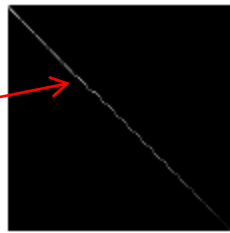
- Woods introduced a **registration measure for multimodality** images in 1992.
- The measure was based on the assumption **that regions of similar tissue** (and similar gray tones) **in one image would correspond to regions in the other image** that also consist of similar gray values (but not the same as in the first image).
- Instead of defining regions of similar tissue in the image, **they defined the regions in a feature space.**
- When the images are correctly registered, the **joint histogram** of the two images will show certain clusters for gray tones of matching structures.

CT/MRI Example



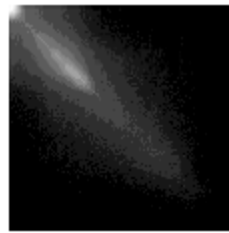
- For each pair of corresponding points (\mathbf{x}, \mathbf{y}) with \mathbf{x} in the CT image and \mathbf{y} in the MR image, there is a **gray tone correspondence** (g_x, g_y) .
- **The joint histogram counts how many times each gray tone correspondence occurs.**

Joint Gray-tone Histograms of an MR Image with itself at Different Rotations



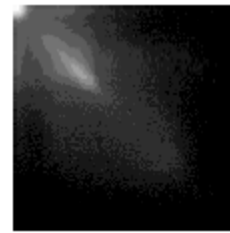
3.82

0 degrees



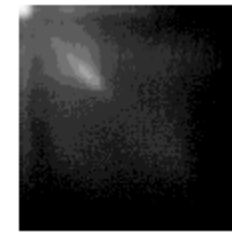
6.79

2 degrees



6.98

5 degrees



7.15

10 degrees

joint entropy

Because the images are identical, all gray-tone correspondences lie on the diagonal of the histogram matrix.

Measures of Mutual Information

- $H = -\sum p_{i,j} \log p_{i,j}$ is the Shannon entropy for a joint distribution; p_{ij} is probability of co-occurrence of i and j .

- Def. 1: $I(A,B) = H(B) - H(B|A)$

$P(b|a)$ is the probability of voxel having value b in B given a in A

- Def. 2: $I(A,B) = H(A) + H(B) - H(A,B)$

Maximizing mutual info is like minimizing entropy.

- Def. 3: Kullback-Leibler distance

$$I(A, B) = \sum_{a,b} p(a, b) \log \frac{p(a, b)}{p(a)p(b)}$$

joint gray values

joint in case of independent images

Different Aspects of Mutual Information Procedures

- **Preprocessing** (ie. filtering)
- **Measures** (entropy measure, normalization measures)
- **Spatial Information** (not just gray tones, but where)
- **Transformation** (applied to register images)
 - rigid
 - affine
 - deformable
- **Implementation**
 - interpolation
 - probability distribution estimation
 - optimization

Modalities

- MR with CT, PET, SPECT, US
- CT with PET, SPECT, other (video, fluoroscopy)
- Microscopy with other

Anatomical Entities

- brain
- thorax/lungs
- spine
- heart
- breast
- abdomen/liver
- pelvis
- tissue
- other

PET-CT Image Registration in the Chest Using Free-Form Deformations

David Mattes, David Haynor, et al
UW Medical Center

- Popular implementation of mutual information registration
- Available in ITK package
- We use it in our research.

Application

- PET-to-CT image registration in the chest
- Fuse images from a modality with high anatomic detail (CT)
- With images from a modality delineating biological function (PET)
- Producing a nonparametric deformation that registers them.

Overall Method

- The PET image has a corresponding **transmission image (TR)**
- The TR image is similar to a CT attenuation map with a higher energy radiation beam, resulting in **less soft-tissue detail and limited resolution**
- Once the TR and CT images are registered, the resulting **transformation can be applied to the emission image** for improved PET image interpretation.
- **GOAL: find a deformation map that aligns the TR image with the CT image and evaluate the accuracy.**

Axial Slice



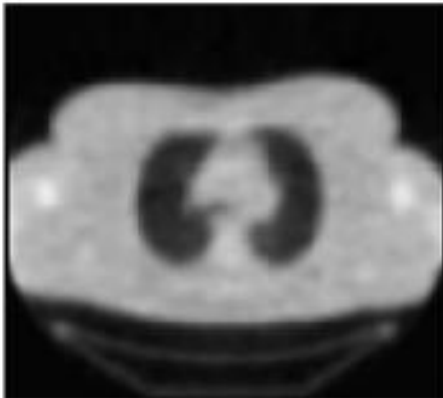
(a)

Coronal Slice

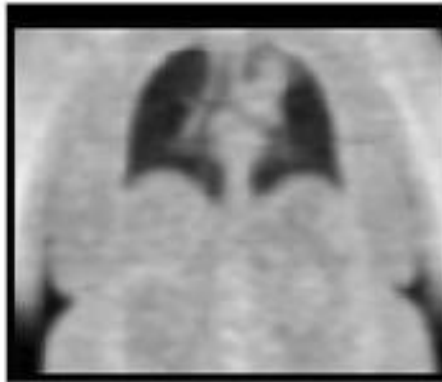


(b)

CT Image



(c)



(d)

TR Image

Methodology

Notation

- $f_T(x)$ is a test image over domain V_T
- $f_R(x)$ is a reference image over domain V_R
- $g(x | \mu)$ is a deformation from V_T to V_R
- μ is the set of parameters of the transformation
- We want to find the set of parameters μ that **minimizes an image discrepancy function S**

$$\mu = \arg \min S(f_R, f_T \circ g(\bullet | \mu))$$

- They hypothesize that the set of transformation parameters that minimizes S brings the transformed test image into best registration with the reference image.

Image Representation

- Optimizing a function requires taking **derivatives**.
- Thus it is easier if the function can be represented in a form that is **explicitly differentiable**.
- This means that both the deformations and the similarity criterion must be differentiable.
- So images are represented using a **B-Spline basis**.
- **Parzen windows** are used instead of simple, bilinear interpolation.

SOME of the Math

- An **image** $f(\mathbf{x})$, coming in as a set of sampled values, is represented by a **cubic spline function** that can be interpolated at any between-pixel position.

$$f(\mathbf{x}) = \sum_i c_i \beta^{(3)}(\mathbf{x} - \mathbf{x}_i)$$

- The spline function is **differentiable**.
- The **smoothed joint histogram** of $(f_R, f_T \circ g(\bullet | \mu))$ is defined as a cross product of the two spline functions.
- Computation of mutual information requires
 - the smoothed joint histogram
 - the marginal smoothed histogram for the test image
 - the marginal probability for the reference image, which is independent of the transformation parameters

- The **image discrepancy measure** is the negative of mutual information S between the reference image and the transformed test image expressed as a function of the transformation parameters μ .

$$S(\mu) = - \sum_{\iota} \sum_{\kappa} p(\iota, \kappa | \mu) \log \frac{p(\iota, \kappa | \mu)}{p_T(\iota | \mu) p_R(\kappa)}$$

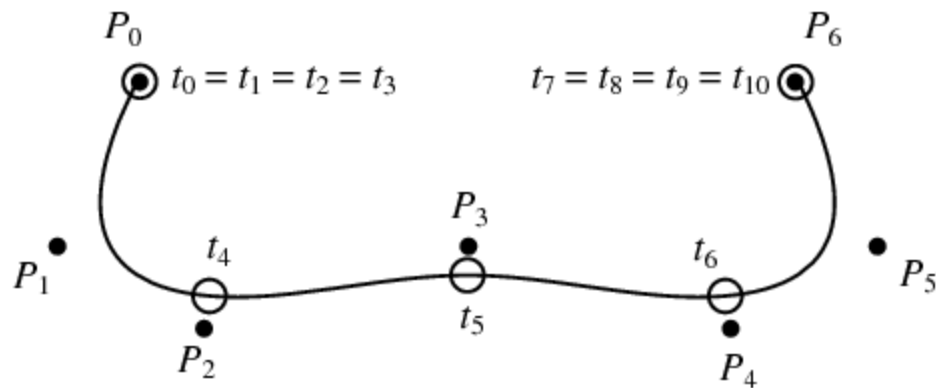
Joint
Marginals

where **p , p_T , and p_R** are the joint, marginal test, and marginal reference **probability distributions**, respectively.

- The variables τ and κ are the histogram bin indexes for the reference and test images, respectively.

Deformations

- Deformations are also modeled as **cubic B-splines**.
- They are defined on a much **coarser grid**.
- A deformation is defined on a sparse, **regular grid** of control points placed over the test image.
- A deformation is varied by defining the **motion $g(\lambda_j)$ at each control point λ_j** .



Transformation

- The transformation of the test image is specified by mapping reference image coordinates according to **a locally perturbed rigid body transformation**.

- The parameters of the transformation are:

$$\mu = \{ \gamma, \theta, \phi, t_x, t_y, t_z; \delta_j \}$$

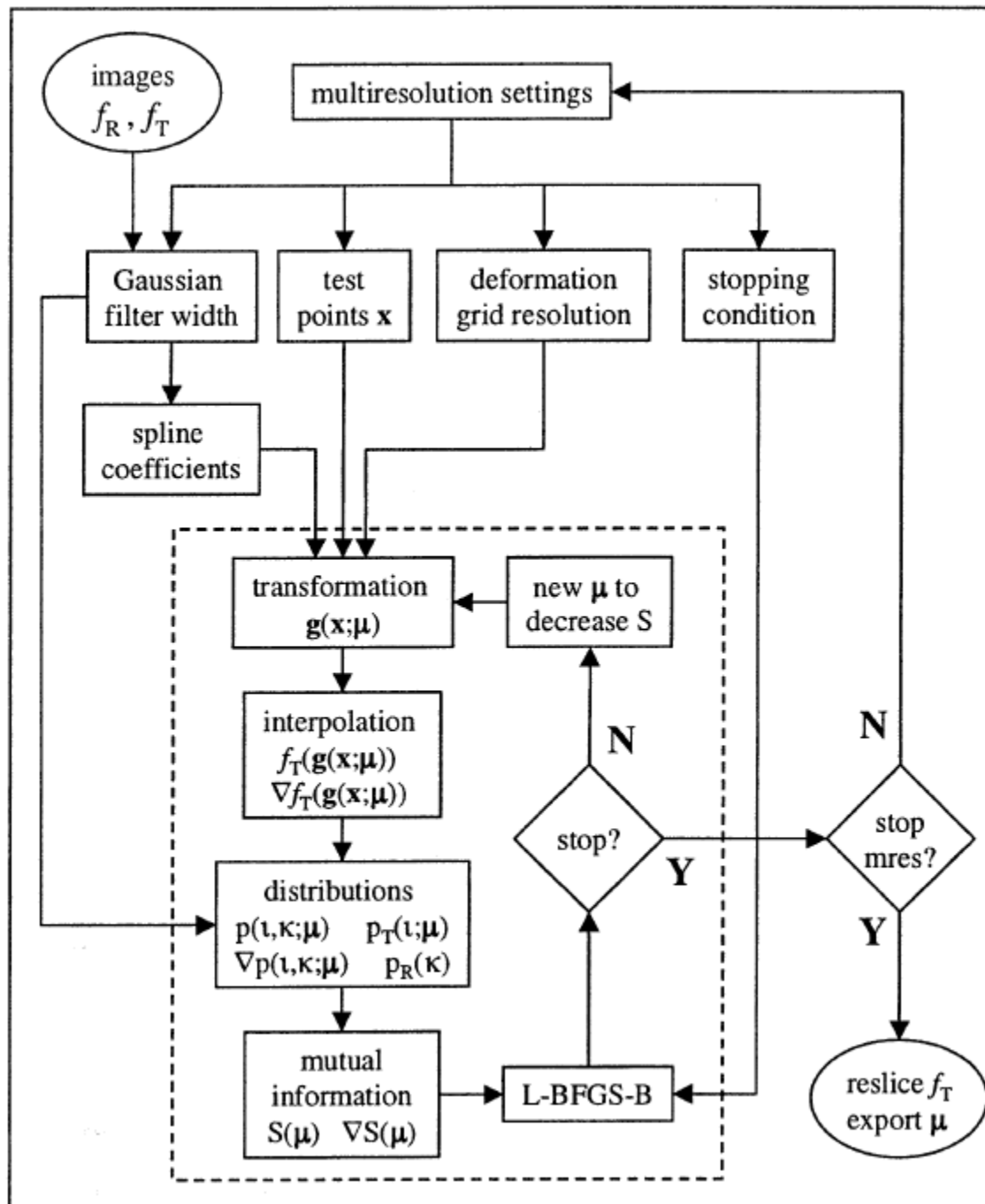
$\{\gamma, \theta, \phi\}$ are the **roll-pitch-yaw** Euler angles,

$[t_x, t_y, t_z]$ is the **translation** vector,

and δ_j is the **set of deformation coefficients** (2200 of them)

Multiresolution Optimization Strategy

- The registration process is automated by varying the deformation in the test image until the discrepancy between the two images is minimized.
- The alignment process is divided into two registrations: one for the rigid body part and one for the deformation
- A limited-memory, quasi-Newton minimization package is used.
- To avoid local minima and decrease computation time, a hierarchical multiresolution optimization scheme is used.



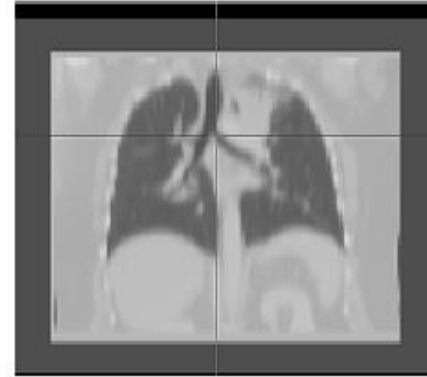
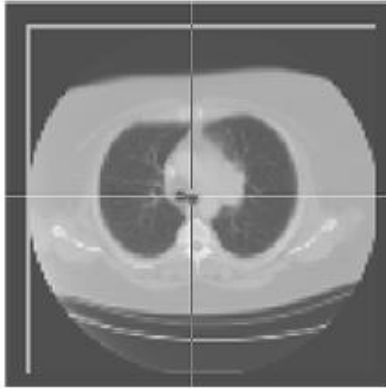
Results

- 28 patients, 27 successful registrations
- 205 slices per image
- average of 100 minutes per registration
- 10 minutes for the rigid body part
- **90 minutes** for the deformable part
- error index of .54, which is in the **0 to 6mm error range**

axial

coronal

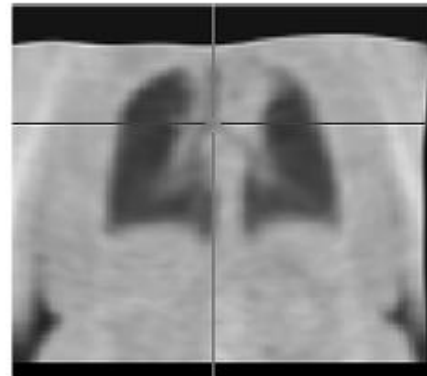
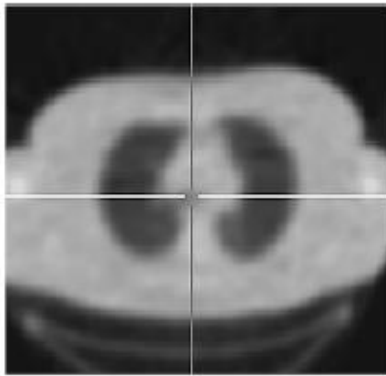
CT



(a)

(b)

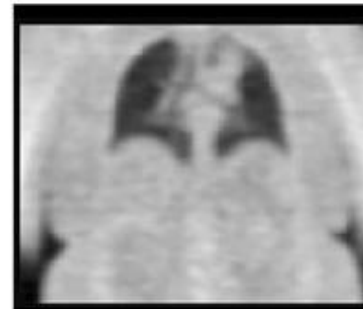
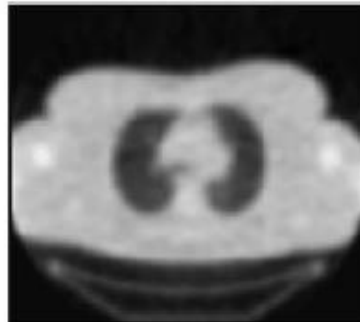
registered TR



(c)

(d)

unregistered TR



(e)

(f)

Sample Images from 7 Anatomic Locations

