Title.  *Motion-compensated Inverse Radon Transformation using Deep Learning: a Medical application*

Presenters. Haoran Chang, Ph.D. Candidate, FIT; and Debasis Mitra, Professor, FIT

Abstract. Radon Transform (RT), invented in 1917 by Johann Radon, has its use in medical imaging, astronomy, Seismic imaging, and physics. RT and its inverse, IRT, transform an image between difference spaces, as Fourier Transform (FT) and its inverse, IFT do. RT and FT are quite inter-related mathematically. Many of the imaging applications of RT-IRT are affected by motions of the imaged object during data acquisition. For example, RT allows non-invasive imaging of heart, but a heart beats too fast during some of the imaging modalities with coarse timing-resolutions. While the issues of noise-removed IRT is a subject of research and practice over a half-century, applications of Convolutional Neural Network (CNN) in IRT is only recently being explored. In this work, we show how such a network may correct for cardiac motion using IRT in image reconstruction that is solely based on CNN. We compare our novel network against some existing paradigms of Deep Learning.

Key Words. Medical image Reconstruction; Motion Correction; Inverse Radon Transformation; Deep Learning; Convolutional Encoder Decoder
Radon Transform

Integral Projection

Angle 0 projection

Angle 90 projection
Radon Transform

Sinogram

Phantom

Radon Transform

Sinogram (Animation)

Horizontal projections through the shape result in an accumulated signal (middle bar). The sinogram on the right is generated by collecting many such projections as the shape rotates. Here, color is used to highlight which object is producing which part of the signal. Note how straight features, when aligned with the projection direction, result in stronger signals.

Inverse Radon Transform (Backprojection)

3 4 3
3 4 3
3 4 3

3 4 3
4 4 4
3 3 3

3 3 3
4 4 4
3 3 3

6 7 6
7 8 7
6 7 6

3 3.5 3
3.5 4 3.5
3 3.5 3

Divided by 2
(# of projections)
Inverse Radon Transform (Backprojection)

(A) Slice used to create projections. (B–G) 1, 3, 4, 16, 32, and 64 projections equally distributed over 360° are used to reconstruct slice using backprojection algorithm.

Figure from "Analytic and Iterative Reconstruction Algorithms in SPECT" by Philippe P. Bruyant
Computed Tomography: Absorption of X-ray Anatomical Imaging

20-40 keV
Emission Tomography:
Functional Imaging

Source: ME Phelps, PNAS, 97(16) 9226–9233, 2000
Problem we are addressing: How to Inverse Motion-added Radon Transform?

Motion types in medical imaging: Beating heart, Respiration, Patient motion, …

Creates blurriness in sinogram, and results in artifacts Reconstructed images
Data Preparation for Training and Validation for Deep Learning

Original Data - Pseudo Heart
Data Preparation for Training and Validation

Augmentation - Randomly select some frames and perform affine transformations
Data Preparation for Training and Validation

Blurring - Affine motion
Neural Network Model for Image Reconstruction

Convolutional Encoder-Decoder with Self-Attention
Neural Network Model

CEDA - Encoder Block

![Diagram of Encoder Block](image)
Neural Network Model

CEDA - Decoder Block
Conventional noise reduction with Deep Learning are performed by U-net
U-net Model for Noise-reduction

U-net (for comparison against ours)
Conventional noise reduction are performed by U-net

*However, U-net maps from image-space to similar image-space*
Result: reconstructions

Image-0
CED 0
CEDA 0
U-net 0
Ground truth 0

Image-1
CED 1
CEDA 1
U-net 1
Ground truth 1

Image-2
CED 2
CEDA 2
U-net 2
Ground truth 2

Image-3
CED 3
CEDA 3
U-net 3
Ground truth 3

Image-4
CED 4
CEDA 4
U-net 4
Ground truth 4

Image-5
CED 5
CEDA 5
U-net 5
Ground truth 5
Result: Comparison Metrics

1. Visual Information Fidelity (VIF)

\[ VIF = \frac{\text{Distorted Image Information}}{\text{Reference Image Information}} = \frac{I(C; F)}{I(C; E)} \]

Where:

- \( I(x; y) \) - Mutual information between \( x \) and \( y \)
- \( C \) - Reference image
- \( E \) - Visual signal of \( C \) at the output of human visual system (HVS)
- \( F \) - Visual signal of distorted \( C \) at the output of HVS
Result

Comparison

- Visual Information Fidelity (VIF)
Result: *Comparison Metrics*

2. Signal-to-noise Ratio (SNR)

\[ SNR = \frac{\mu_{ROI}}{\sigma_{background}} \]

3. Contrast-to-noise Ratio (CNR)

\[ CNR = \frac{\mu_{ROI} - \mu_{background}}{\sigma_{background}} \]

\( \mu \) is mean and sigma is standard deviation

*ROI region of interest (heart)*
Result

- Signal-to-noise Ratio (SNR)

- Contrast-to-noise Ratio (CNR)
Result

Comparison (mean and stdv)

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean and stdv of VIF</th>
<th>Mean and stdv of SNR</th>
<th>Mean and stdv of CNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CED</td>
<td>0.54±0.042</td>
<td>30.44±10.66</td>
<td>20.88±9.72</td>
</tr>
<tr>
<td>CEDA</td>
<td>0.55±0.036</td>
<td>32.12±11.96</td>
<td></td>
</tr>
<tr>
<td>Unet</td>
<td>0.45±0.082</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\textbf{VIF:}  
44 out of 50 reconstructed images CEDA reconstruction is better than U-net denoising,

and 26 reconstructed images with CEDA is better than CED.