Learning to See: perceptual learning and MRI image reconstruction

Matt Rosen
MGH/A.A. Martinos Center

AAPM
San Antonio, TX
16 July 2019

Athinoula A. Martinos Center For Biomedical Imaging

MGH 1811

Learning to See: perceptual learning and MRI image reconstruction

"Bay Zero", Bldg 75, CNY

Athinoula A. Martinos Center For Biomedical Imaging

MGH 1811

Commercial Disclosures

Founder and SAB member:

HYPERFINE

blink·AI

SAB member:

ABQMR

Sponsored research : GE Healthcare

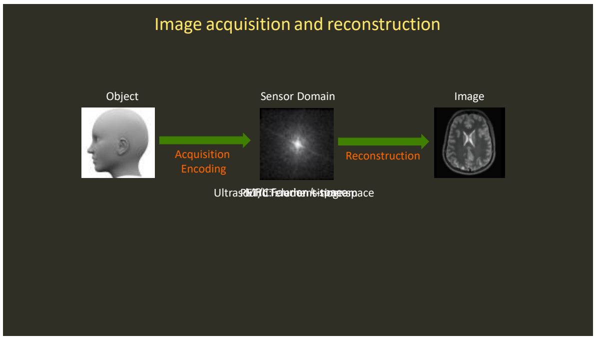
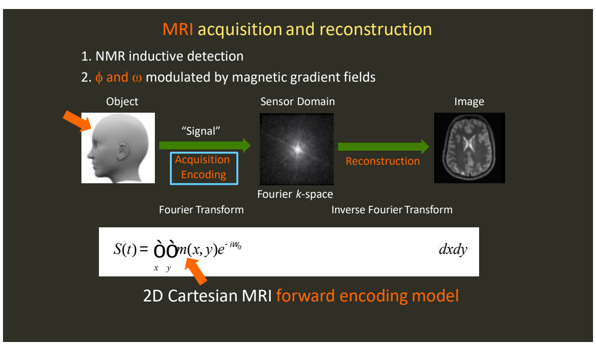
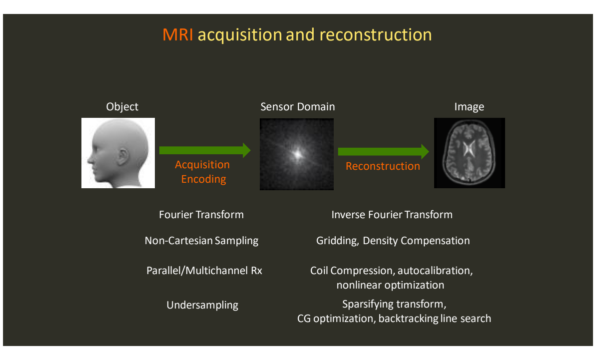
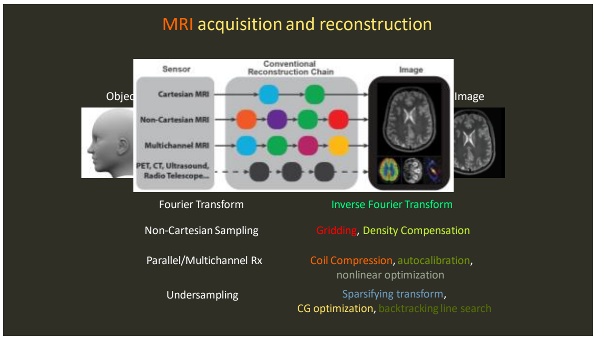
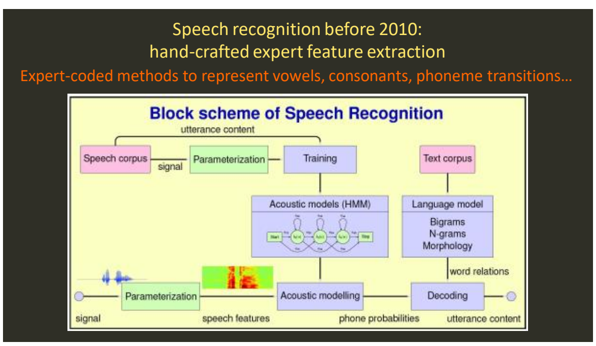


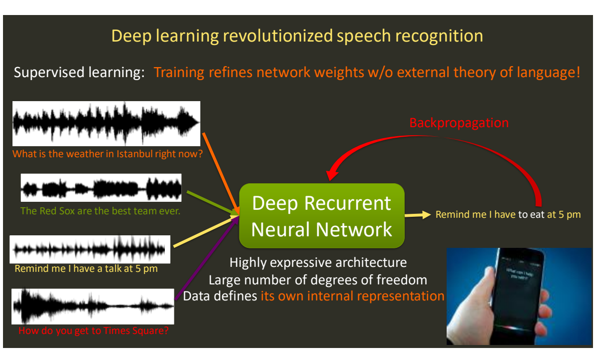
Image acquisition and reconstruction
Object
Sensor Domain
Image
Acquisition
Encoding
Reconstruction


MRI acquisition and reconstruction
1. NMR inductive detection
2. ϕ and ω modulated by magnetic gradient fields
Object
Sensor Domain
Image
"Signal"
Acquisition
Encoding
Reconstruction
Fourier k-space
Fourier Transform
Inverse Fourier Transform
$S(t)=\int_x\int_y m(x,y)e^{-i\omega_0} \quad dxdy$
2D Cartesian MRI forward encoding model


MRI acquisition and reconstruction
Object
Sensor Domain
Image
Acquisition
Encoding
Reconstruction
Fourier Transform
Inverse Fourier Transform
Non-Cartesian Sampling
Gridding, Density Compensation
Parallel/Multichannel Rx
Coil Compression, autocalibration,
nonlinear optimization
Undersampling
Sparsifying transform,
CG optimization, backtracking line search


MRI acquisition and reconstruction
Object
Sensor
Conventional Reconstruction Chain
Image
Cartesian MRI
Non-Cartesian MRI
Multichannel MRI
PET, CT, Ultrasound, Radio Telescope...
Image
Fourier Transform
Inverse Fourier Transform
Non-Cartesian Sampling
Gridding, Density Compensation
Parallel/Multichannel Rx
Coil Compression, autocalibration, nonlinear optimization
Undersampling
Sparsifying transform, CG optimization, backtracking line search


Speech recognition before 2010:
hand-crafted expert feature extraction
Expert-coded methods to represent vowels, consonants, phoneme transitions...
Block scheme of Speech Recognition
utterance content
Speech corpus
signal
Parameterization
Training
Text corpus
Acoustic models (HMM)
Language model
Bigrams
N-grams
Morphology
word relations
Parameterization
Acoustic modelling
Decoding
signal
speech features
phone probabilities
utterance content


Deep learning revolutionized speech recognition
Supervised learning: Training refines network weights w/o external theory of language!
What is the weather in Istanbul right now?
The Red Sox are the best team ever.
Remind me I have a talk at 5 pm
How do you get to Times Square?
Backpropagation
Deep Recurrent Neural Network
Remind me I have to eat at 5 pm
Highly expressive architecture
Large number of degrees of freedom
Data defines its own internal representation

Automated feature extraction: Solving difficult problems

Previously reliant on expert feature engineering:

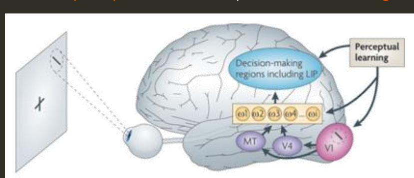
- Speech Recognition
- Natural Language Processing
- AI Gaming (Chess, Go, ATARI)
- Image Classification
- Medical Image Segmentation
- Scene analysis (autonomous vehicles)

Enabled by a technological convergence:

- Algorithms/architectures (Deep+Convolutional Neural Networks)
- Accessibility to training data (Big Data)
- Advanced parallel computing hardware (Multi-GPU)

Perceptual Learning

Refinement of perception based on exposure to and training on stimuli

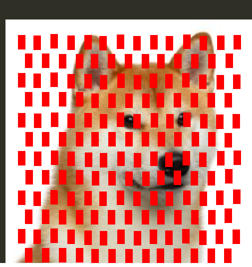


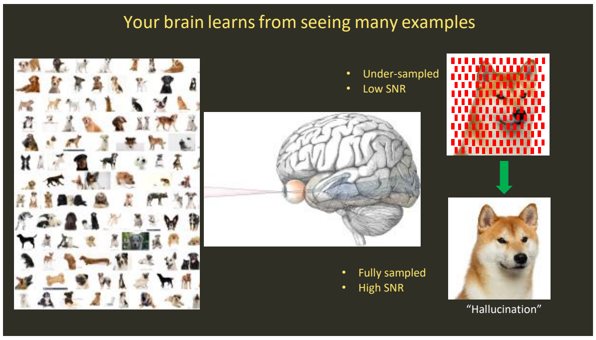
Sasaki, Yuka, Jose E. Nanez, and Takeo Watanabe. "Advances in visual perceptual learning and plasticity." *Nature Reviews Neuroscience* 11.1 (2010): 53-60.

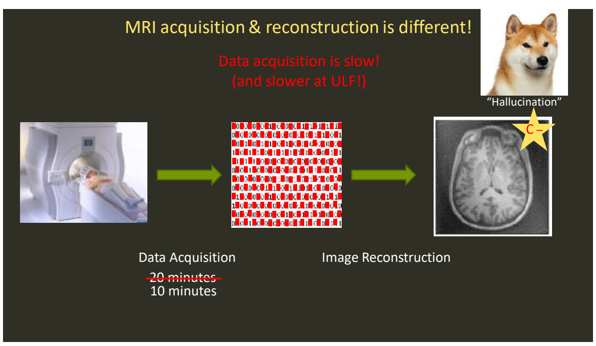
→ Perceptual learning is critical to robust performance in low-SNR settings

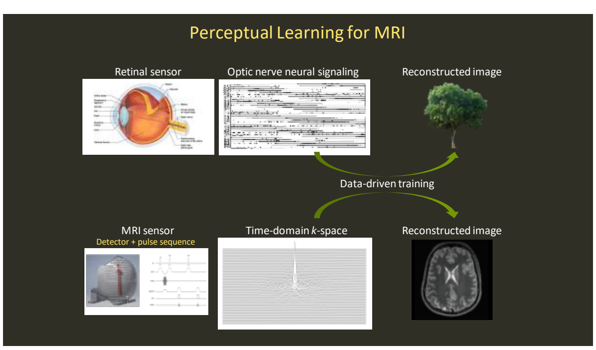
Lu, Z.-L., et al. Visual perceptual learning. *Neurobiology of Learning and Memory* **95**, 145–151 (2011)

What animal is this?



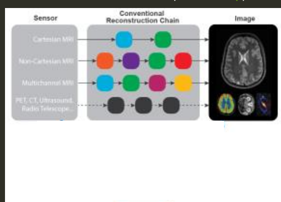

Your brain learns from seeing many examples
Under-sampled
Low SNR
Fully sampled
High SNR
"Hallucination"


MRI acquisition & reconstruction is different!
Data acquisition is slow!
(and slower at ULF!)
"Hallucination"
Data Acquisition
~~20 minutes~~
10 minutes
Image Reconstruction


Perceptual Learning for MRI
Retinal sensor
Optic nerve neural signaling
Reconstructed image
Data-driven training
MRI sensor
Detector + pulse sequence
Time-domain k-space
Reconstructed image

Deep learning for image reconstruction

AUTOMAP: Automated Transform by Manifold Approximation

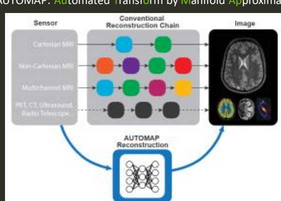


1. Data-driven supervised learning replaces hand-crafted pipelines
2. Perceptual learning biologically-inspired approach improves SNR of noisy data

→ Recast image reconstruction as a supervised learning task

Deep learning for image reconstruction

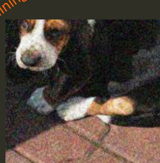
AUTOMAP: Automated Transform by Manifold Approximation



→ Recast image reconstruction as a supervised learning task
- Reconstruction relationship emerges from raw sensor and image data
- Forward model is inverted by learning pairs of examples

Convolutional NN denoiser

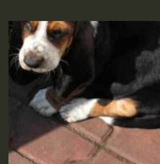
AUTOMAP: not noise training!



Noisy image

CNN

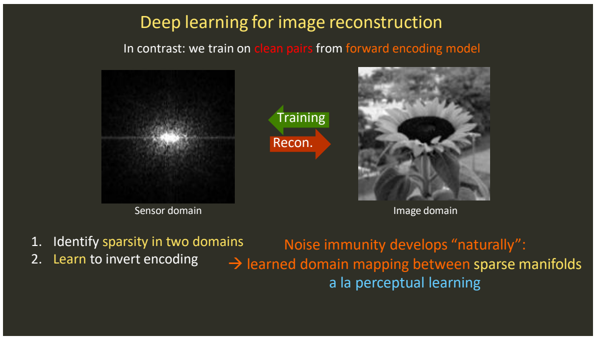
Learned mapping

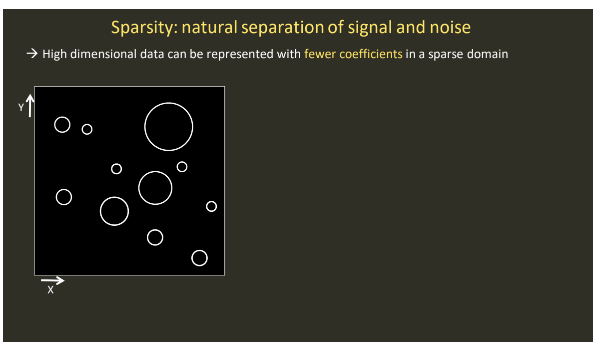


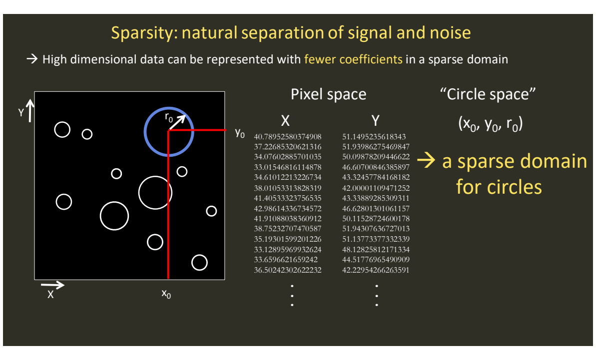
Clean image

Mapping from noisy to clean aka noise training learned from pairs of examples

Images: Jaakko Lehtinen

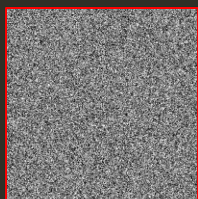

Deep learning for image reconstruction
In contrast: we train on clean pairs from forward encoding model
Training
Recon.
Sensor domain
Image domain
1. Identify sparsity in two domains
2. Learn to invert encoding
Noise immunity develops "naturally":
→ learned domain mapping between sparse manifolds
a la perceptual learning


Sparsity: natural separation of signal and noise
→ High dimensional data can be represented with fewer coefficients in a sparse domain
Y
X


Sparsity: natural separation of signal and noise
→ High dimensional data can be represented with fewer coefficients in a sparse domain
Y
X
r_0
y_0
x_0
Pixel space
X
Y
40.78952580374908 51.1495235618343
37.22685320621316 51.93986275469847
34.07602885701035 50.09878209446622
33.01546816114878 46.60700846385897
34.61012213226734 43.32457784168182
38.01053313828319 42.00001109471252
41.40533323756535 43.33889285309311
42.98614336734572 46.62801301061157
41.91088038360912 50.11528724600178
38.75232707470587 51.94307636727013
35.19301599201226 51.13773377332339
33.12895969932624 48.12825812171334
33.6596621659242 44.51776965490909
36.50242302622232 42.22954266263591
"Circle space"
(x_0, y_0, r_0)
→ a sparse domain for circles

Sparsity: natural separation of signal and noise

Noise can be anything… except sparse!



Possible images: $2^{128 \times 128}$

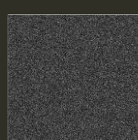
(4,933 digits!)

…we need all those coefficents!

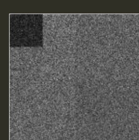
Sparsity: natural separation of signal and noise

Noise is not sparse in any domain!

Not sparse



Fourier domain:
also not sparse



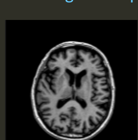
Wavelet domain:
no sparsity here

→ High dimensional data can be represented with fewer coefficients in a sparse domain

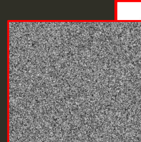
Not surprising given the data dimensionality:
$2^{128 \times 128} >> 10^{80}$ atoms in the universe

Sparsity: natural separation of signal and noise

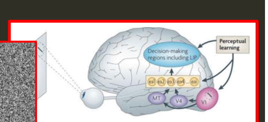
Natural images are special



Not sparse



Fourier domain:
also not sparse

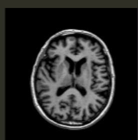


Wavelet domain:
→ sparse

"Brain hallucinates image using learned sparse features"

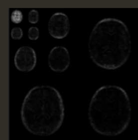
Sparsity: natural separation of signal and noise

→ High dimensional data can be represented with fewer coefficients in a sparse domain



Not sparse

Fourier domain:
also not sparse



Wavelet domain:
→ sparse

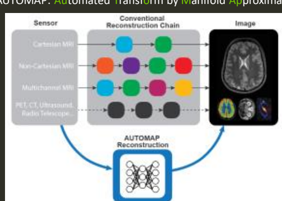
- NN training can encourage efficient internal representation of learned mapping

→ AUTOMAP transform operates between data-defined sparse domains

→ Image is hallucinated from the learned sparse convolutional feature maps

Deep learning for image reconstruction

AUTOMAP: Automated Transform by Manifold Approximation



LETTER nature March 2018

Image reconstruction by domain-transform manifold learning

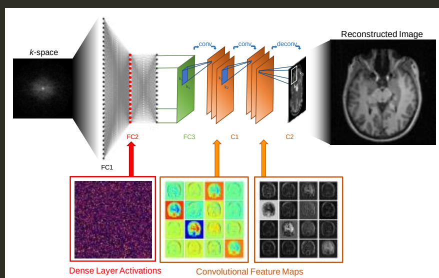
→ Recast image reconstruction as a supervised learning task
- Reconstruction relationship emerges from raw sensor and image data
- Training: conditions joint manifold for sparsity & learns to invert encoding

Neuromorphic approach: AUTOMAP

AUTOMAP: Automated Transform by Manifold Approximation

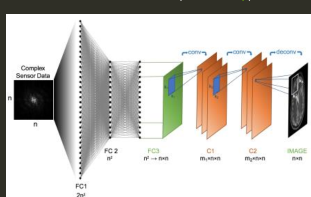
Mathematical transform + sparse properties of natural images

AUTOMAP feed-forward reconstruction



Neuromorphic approach: AUTOMAP

AUTOMAP: Automated Transform by Manifold Approximation

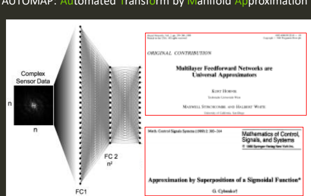


Mathematical transform + sparse properties of natural images

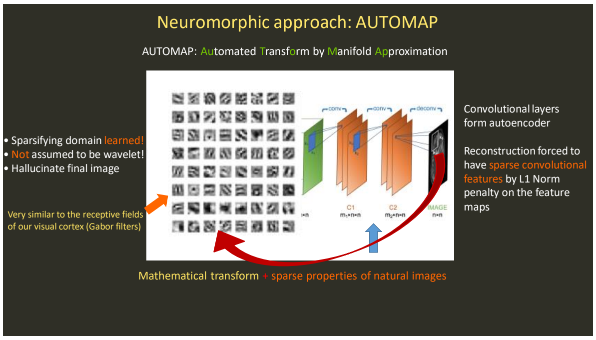
Neuromorphic approach: AUTOMAP

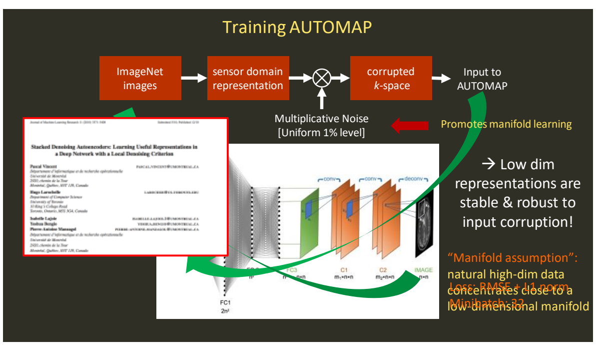
AUTOMAP: Automated Transform by Manifold Approximation

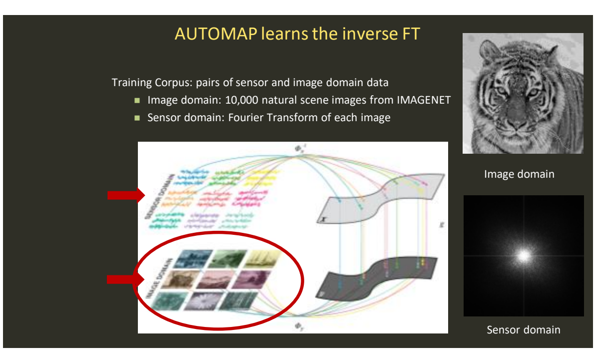
Fully connected layers: universal function approximators that can represent any function on compact set

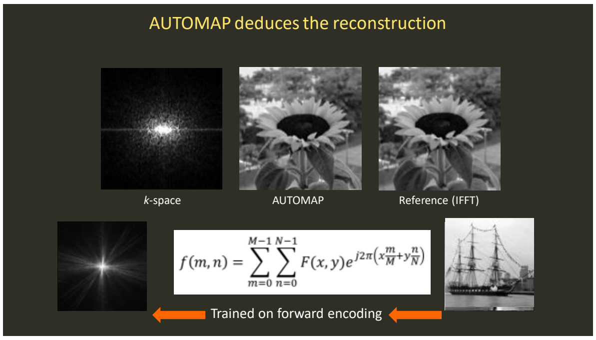


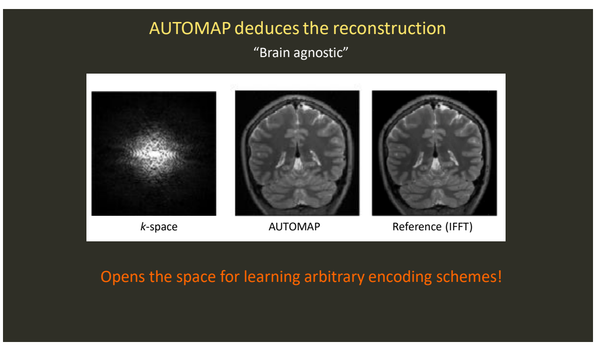
Mathematical transform + sparse properties of natural images

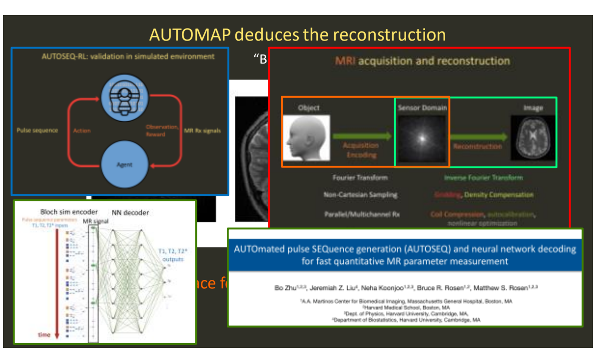

Neuromorphic approach: AUTOMAP
AUTOMAP: Automated Transform by Manifold Approximation
• Sparsifying domain learned!
• Not assumed to be wavelet!
• Hallucinate final image
Very similar to the receptive fields of our visual cortex (Gabor filters)
Convolutional layers form autoencoder
Reconstruction forced to have sparse convolutional features by L1 Norm penalty on the feature maps
Mathematical transform + sparse properties of natural images

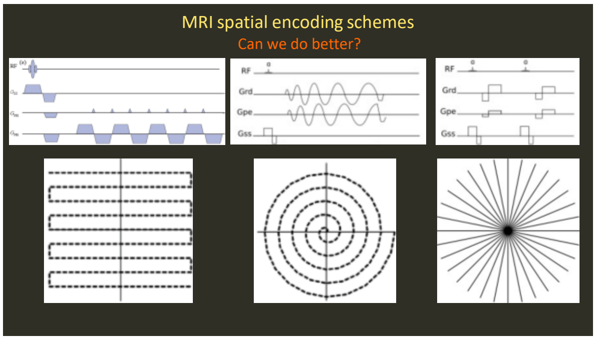

Training AUTOMAP
ImageNet images
sensor domain representation
corrupted k-space
Input to AUTOMAP
Multiplicative Noise [Uniform 1% level]
Promotes manifold learning
→ Low dim representations are stable & robust to input corruption!
"Manifold assumption": natural high-dim data concentrates close to a low-dimensional manifold
Loss: RMSE + L1 norm
Minibatch: 32
Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion

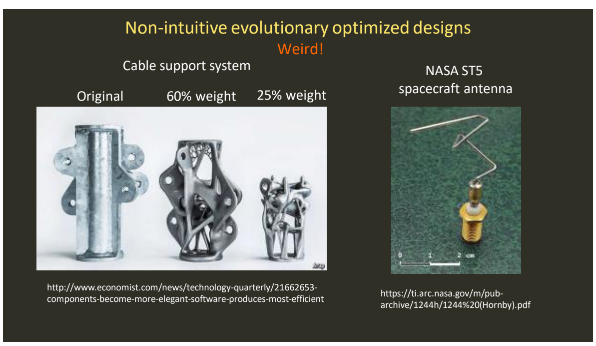

AUTOMAP learns the inverse FT
Training Corpus: pairs of sensor and image domain data
Image domain: 10,000 natural scene images from IMAGENET
Sensor domain: Fourier Transform of each image
Image domain
Sensor domain

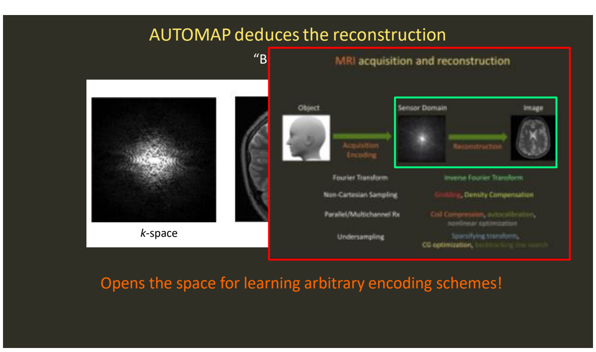

AUTOMAP deduces the reconstruction
k-space
AUTOMAP
Reference (IFFT)
$f(m,n) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(x,y) e^{j2\pi\left(x\frac{m}{M}+y\frac{n}{N}\right)}$
Trained on forward encoding

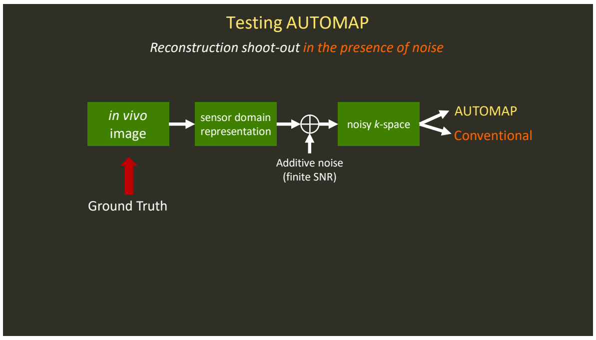

AUTOMAP deduces the reconstruction
"Brain agnostic"
k-space
AUTOMAP
Reference (IFFT)
Opens the space for learning arbitrary encoding schemes!

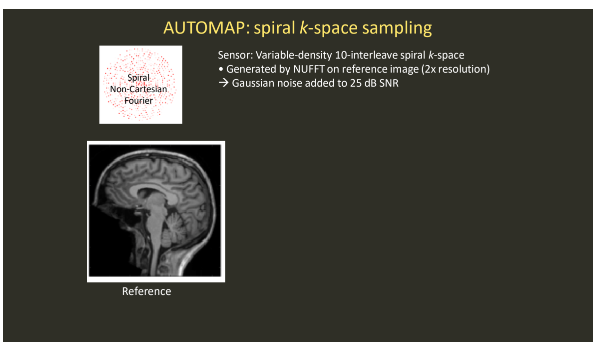

AUTOMAP deduces the reconstruction
AUTOSEQ-RL: validation in simulated environment
Pulse sequence
Action
Observation, Reward
MR Rx signals
Agent
MRI acquisition and reconstruction
Object
Sensor Domain
Image
Acquisition Encoding
Reconstruction
Fourier Transform
Inverse Fourier Transform
Non-Cartesian Sampling
Density Compensation
Parallel/Multichannel Rx
Coil Compression, autocalibration,
nonlinear optimization
Bloch sim encoder
NN decoder
MR signal
T1, T2, T2* outputs
time
AUTOmated pulse SEQuence generation (AUTOSEQ) and neural network decoding for fast quantitative MR parameter measurement
Bo Zhu[1,2,3], Jeremiah Z. Liu[4], Neha Koonjoo[1,2,3], Bruce R. Rosen[1,2], Matthew S. Rosen[1,2,3]
[1]A.A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA
[2]Harvard Medical School, Boston, MA
[3]Dept. of Physics, Harvard University, Cambridge, MA
[4]Department of Biostatistics, Harvard University, Cambridge, MA

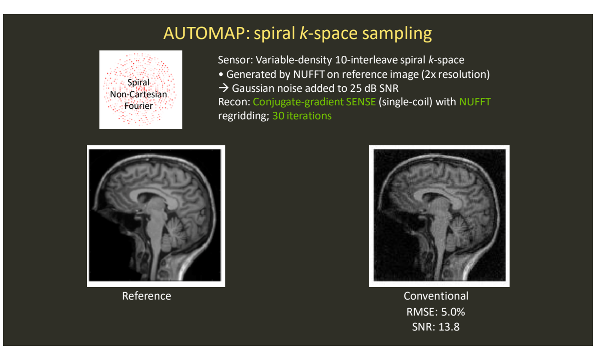

MRI spatial encoding schemes
Can we do better?
RF
Grd
Gpe
Gss
RF
Grd
Gpe
Gss

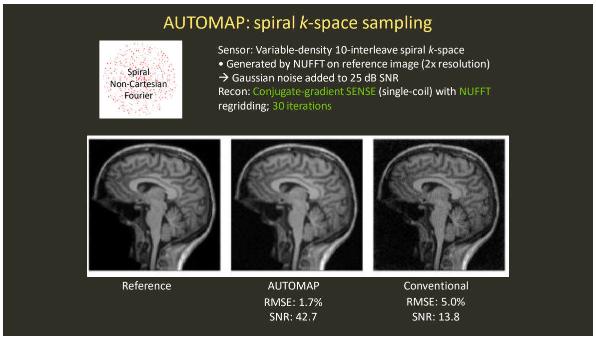

Non-intuitive evolutionary optimized designs
Weird!
Cable support system
Original
60% weight
25% weight
NASA ST5
spacecraft antenna
0 1 2 cm
http://www.economist.com/news/technology-quarterly/21662653-components-become-more-elegant-software-produces-most-efficient
https://ti.arc.nasa.gov/m/pub-archive/1244h/1244%20(Hornby).pdf

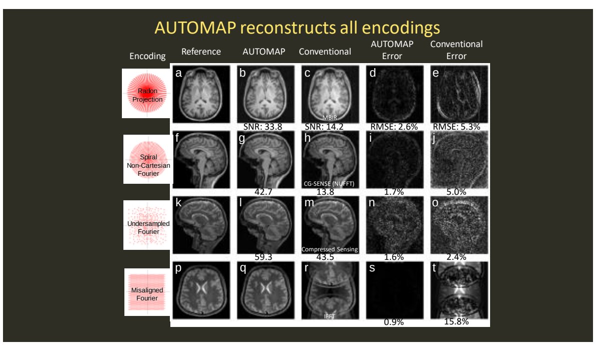

AUTOMAP deduces the reconstruction
"B
MRI acquisition and reconstruction
Object
Sensor Domain
Image
Acquisition Encoding
Reconstruction
Fourier Transform
Inverse Fourier Transform
Non-Cartesian Sampling
Density Compensation
Parallel/Multichannel Rx
Coil Compression,
nonlinear optimization
Undersampling
CG optimization,
k-space
Opens the space for learning arbitrary encoding schemes!

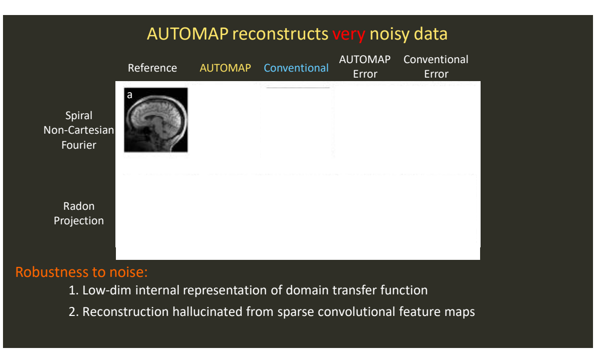

Testing AUTOMAP
Reconstruction shoot-out in the presence of noise
in vivo image
sensor domain representation
Additive noise (finite SNR)
noisy k-space
AUTOMAP
Conventional
Ground Truth

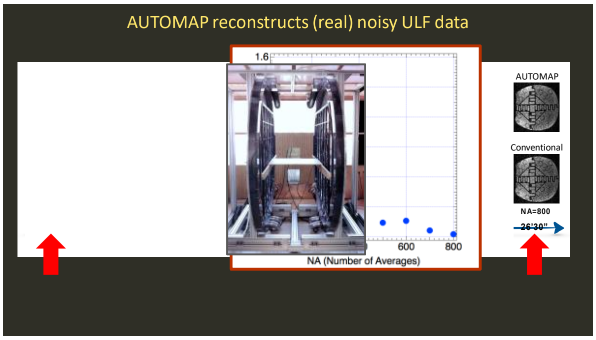

AUTOMAP: spiral k-space sampling
Spiral Non-Cartesian Fourier
Sensor: Variable-density 10-interleave spiral k-space
• Generated by NUFFT on reference image (2x resolution)
→ Gaussian noise added to 25 dB SNR
Reference

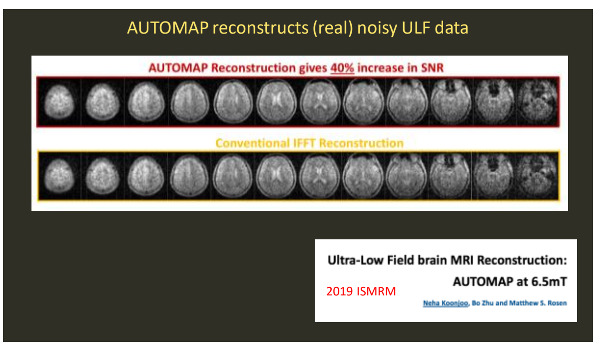

AUTOMAP: spiral k-space sampling
Spiral Non-Cartesian Fourier
Sensor: Variable-density 10-interleave spiral k-space
• Generated by NUFFT on reference image (2x resolution)
→ Gaussian noise added to 25 dB SNR
Recon: Conjugate-gradient SENSE (single-coil) with NUFFT regridding; 30 iterations
Reference
Conventional
RMSE: 5.0%
SNR: 13.8

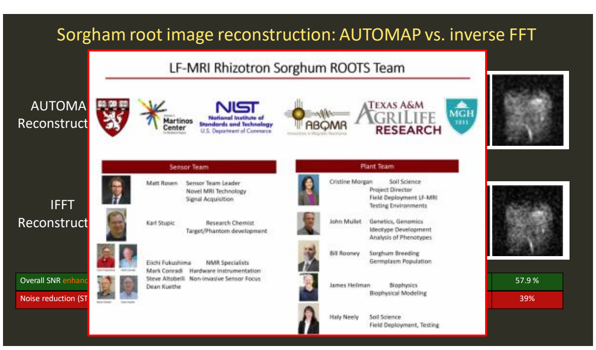

AUTOMAP: spiral k-space sampling
Spiral
Non-Cartesian
Fourier
Sensor: Variable-density 10-interleave spiral k-space
• Generated by NUFFT on reference image (2x resolution)
→ Gaussian noise added to 25 dB SNR
Recon: Conjugate-gradient SENSE (single-coil) with NUFFT
regridding; 30 iterations
Reference
AUTOMAP
RMSE: 1.7%
SNR: 42.7
Conventional
RMSE: 5.0%
SNR: 13.8

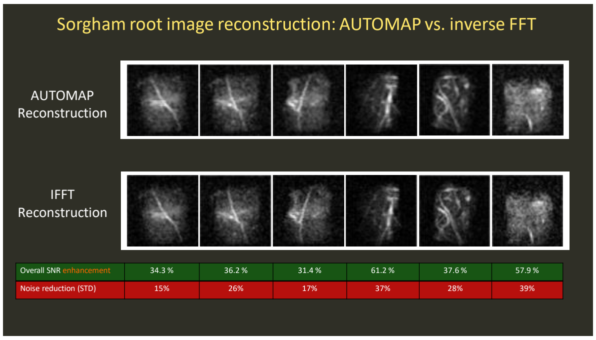

AUTOMAP reconstructs all encodings
Encoding
Reference
AUTOMAP
Conventional
AUTOMAP
Error
Conventional
Error
Radon
Projection
SNR: 33.8
SNR: 14.2
RMSE: 2.6%
RMSE: 5.3%
Spiral
Non-Cartesian
Fourier
42.7
CG-SENSE (NUFFT)
13.8
1.7%
5.0%
Undersampled
Fourier
59.3
Compressed Sensing
43.5
1.6%
2.4%
Misaligned
Fourier
0.9%
15.8%

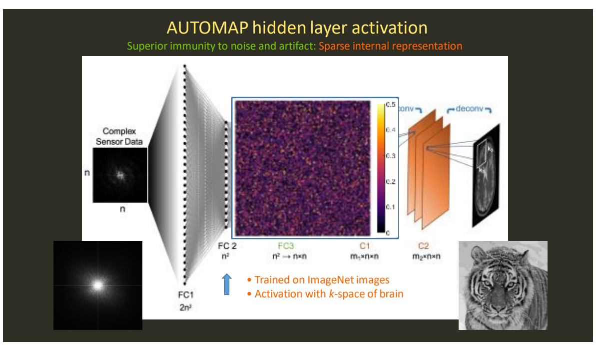

AUTOMAP reconstructs very noisy data
Reference
AUTOMAP
Conventional
AUTOMAP
Error
Conventional
Error
Spiral
Non-Cartesian
Fourier
Radon
Projection
Robustness to noise:
1. Low-dim internal representation of domain transfer function
2. Reconstruction hallucinated from sparse convolutional feature maps

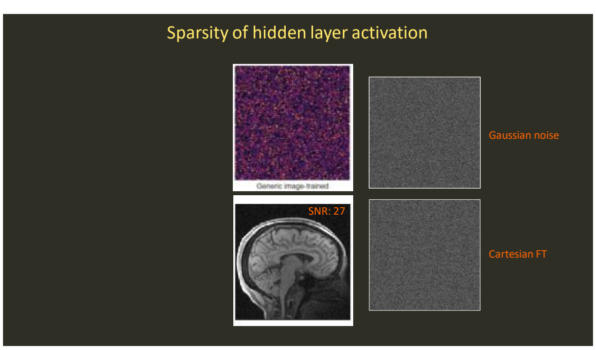

AUTOMAP reconstructs (real) noisy ULF data
1.6
600
800
NA (Number of Averages)
AUTOMAP
Conventional
NA=800
26'30"

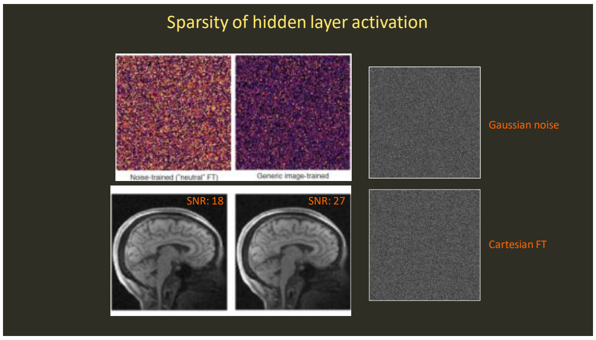

AUTOMAP reconstructs (real) noisy ULF data
AUTOMAP Reconstruction gives 40% increase in SNR
Conventional IFFT Reconstruction
Ultra-Low Field brain MRI Reconstruction:
AUTOMAP at 6.5mT
2019 ISMRM
Neha Koonjoo, Bo Zhu and Matthew S. Rosen

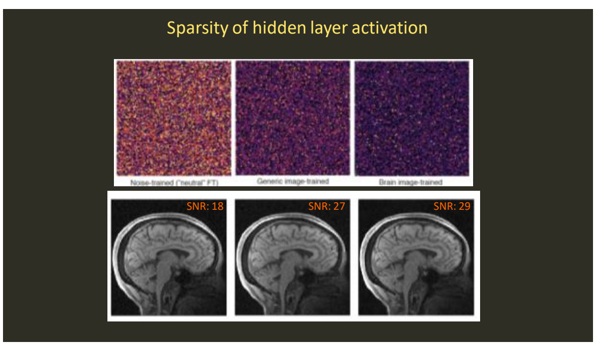

Sorgham root image reconstruction: AUTOMAP vs. inverse FFT
AUTOMA
Reconstruct
IFFT
Reconstruct
Overall SNR enhanc
Noise reduction (ST
57.9 %
39%
LF-MRI Rhizotron Sorghum ROOTS Team
Martinos Center
NIST
National Institute of Standards and Technology
U.S. Department of Commerce
ABQMR
TEXAS A&M AGRILIFE RESEARCH
MGH 1811
Sensor Team
Matt Rosen Sensor Team Leader
Novel MRI Technology
Signal Acquisition
Karl Stupic Research Chemist
Target/Phantom development
Eiichi Fukushima NMR Specialists
Mark Conradi Hardware Instrumentation
Steve Altobelli Non-invasive Sensor Focus
Dean Kuethe
Plant Team
Cristine Morgan Soil Science
Project Director
Field Deployment LF-MRI
Testing Environments
John Mullet Genetics, Genomics
Ideotype Development
Analysis of Phenotypes
Bill Rooney Sorghum Breeding
Germplasm Population
James Heilman Biophysics
Biophysical Modeling
Haly Neely Soil Science
Field Deployment, Testing

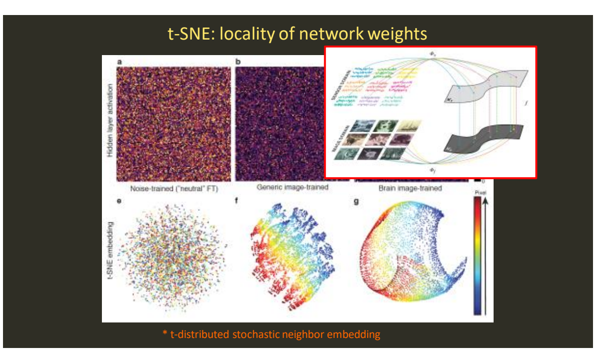

Sorgham root image reconstruction: AUTOMAP vs. inverse FFT
AUTOMAP Reconstruction
IFFT Reconstruction
Overall SNR enhancement | 34.3 % | 36.2 % | 31.4 % | 61.2 % | 37.6 % | 57.9 %
Noise reduction (STD) | 15% | 26% | 17% | 37% | 28% | 39%


AUTOMAP hidden layer activation
Superior immunity to noise and artifact: Sparse internal representation
Complex Sensor Data
n
n
FC1
2n²
FC 2
n²
FC3
n² → n×n
C1
m₁×n×n
C2
m₂×n×n
deconv
• Trained on ImageNet images
• Activation with k-space of brain

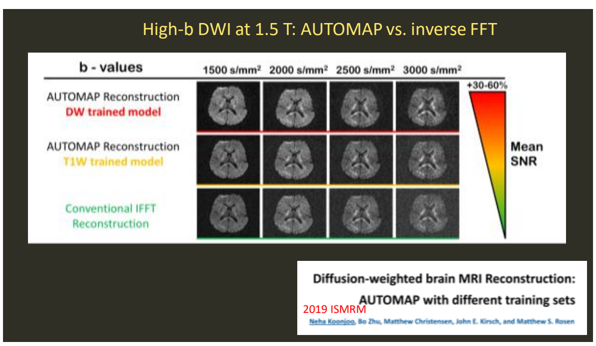

Sparsity of hidden layer activation
Generic image-trained
Gaussian noise
SNR: 27
Cartesian FT

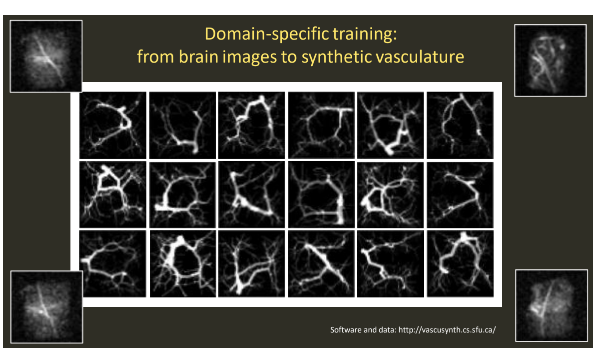

Sparsity of hidden layer activation
Noise-trained ("neutral" FT)
Generic image-trained
Gaussian noise
SNR: 18
SNR: 27
Cartesian FT


Sparsity of hidden layer activation
Noise-trained ("neutral" FT)
Generic image-trained
Brain image-trained
SNR: 18
SNR: 27
SNR: 29

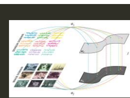

t-SNE: locality of network weights
Hidden layer activation
Noise-trained ("neutral" FT)
Generic image-trained
Brain image-trained
t-SNE embedding
* t-distributed stochastic neighbor embedding


Domain-specific training
Noise-trained ("neutral" FT)
Generic image-trained
Brain image-trained
SNR: 18
SNR: 27
SNR: 29


High-b DWI at 1.5 T: AUTOMAP vs. inverse FFT
b - values
1500 s/mm²
2000 s/mm²
2500 s/mm²
3000 s/mm²
AUTOMAP Reconstruction
DW trained model
AUTOMAP Reconstruction
T1W trained model
Conventional IFFT
Reconstruction
+30-60%
Mean
SNR
Diffusion-weighted brain MRI Reconstruction:
AUTOMAP with different training sets
2019 ISMRM
Neha Koonjoo, Bo Zhu, Matthew Christensen, John E. Kirsch, and Matthew S. Rosen


Domain-specific training:
from brain images to synthetic vasculature
Software and data: http://vascusynth.cs.sfu.ca/

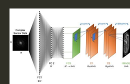
Conclusions



AUTOMAP: unified reconstruction framework

- Universal function approximation
- \+ Manifold learning with deep neural networks

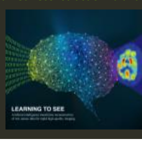
→ Automatically learn optimal reconstruction for arbitrary encodings
→ No imposed expert knowledge

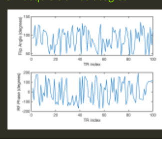


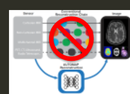
AUTOMAP changes the game:

- Robust immunity to noise
- Faster scan times with less signal averaging (or dose)
- Rapid reconstruction (~1 ms): non-iterative feed-forward computation

→ Generalized reconstruction: brand new acquisition strategies







Brandon Armstrong
Tom Boele
Matt Christensen
Ouri Cohen
Clarissa Cooley
Lina Colucci
Stephen DeVience
Shuning Hwang
Avilash Kalpathy Cramer
Neha Koonjoo
Jeremiah Liu
Cris LaPierre
Najat Salameh
Mathieu Sarracanie
Sheng Shen
Jason Stockman
Bragi Sveinsson
David Waddington
Bo Zhu

@Rosenlab

arpa-e
MGH & BWH CENTER FOR CLINICAL DATA SCIENCE
DARPA
NIH National Institutes of Health

Athinoula A. Martinos Center For Biomedical Imaging

MGH 1811