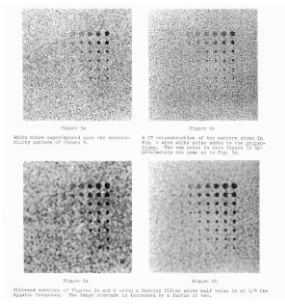
EVALUATION OF NONLINEAR RECONSTRUCTION METHODS

Kyle J. Myers, Ph.D.
Director, Division of Imaging, Diagnostics, and Software Reliability
Office of Science and Engineering Laboratories, CDRH, FDA

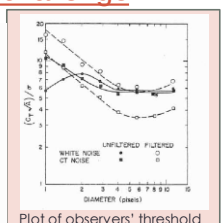
Outline

- Lessons from the early literature
- FDA clearance: evidentiary requirements
  - Reconstruction methods in general
  - Dose-reduction claims in particular
- Where can we go from here?

This is a decades-old challenge



Hanson, SPIE 1977



Plot of observers' threshold contrast vs signal diameter.

Conclusions:

- Visual system has difficulty with CT noise texture.
- Signals must be 2-3 pixels wide.

4

Detection Tasks in Sparse-Scan Images

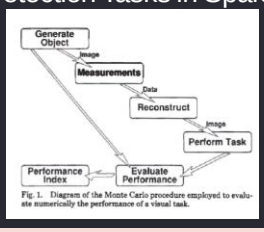
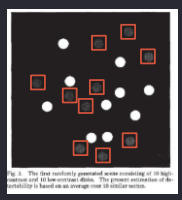


Fig. 1. Diagram of the Monte Carlo procedure employed to evaluate numerically the performance of a visual task.



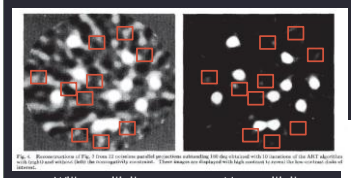
The essence of task-based evaluation:

- Task: Detection of low-contract objects in random scenes
- Imaging system: Axial CT geometries and reconstruction algorithms
- Observer: Humans and various machine readers
- Figure of Merit: Signal detectability

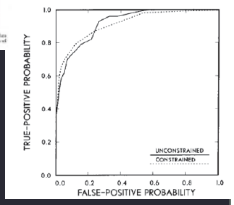
Hanson, JOSA A 1990

5

ART reconstructions from 12 parallel projections over 180°



With positivity constraint | No positivity constraint



Some morals of this story:

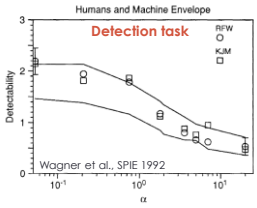
- Need signal-present AND signal-absent locations
- Determination of better algorithm requires multiple images to sample random object locations and noise realizations. (Artifacts are object-dependent.)
- LI and MSE do not necessarily predict observer performance.

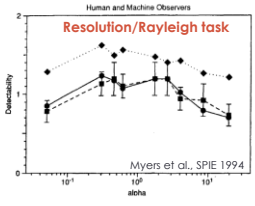
Hanson, JOSA A 1990

6

Human Readers and Model Observers

Maximum Entropy MAP Reconstructions (nonlinear)





More morals of this story:

- Optimal reconstruction algorithm parameters are task-dependent
  - Different trends for detection tasks vs. Rayleigh resolution task.
- Human performance can be predicted by channelized Hotelling observer

7

Dose Reduction Claims in CT

- Sought by industry for iterative reconstruction algorithms
  - IR algorithms already cleared by FDA based on earlier determination of substantial equivalence of IR to FBP for common CT intended use: "to produce cross-sectional images of the body by computer reconstruction of x-ray transmission data taken at different angles and planes…"
  - Low risk: would affect labeling, not product availability
- Established MITA-FDA collaboration to develop framework for evaluation of dose reduction claims, considering possible options including use of:
  - Standard metrics (MTF, NPS, DQE)
  - Computer simulations
  - Physical phantoms
  - Study designs and observers
  - Figures of merit and statistical analysis tools

8

Desired properties for performance evaluation method

- Task-based
  - Detection of objects; discrimination of objects of different sizes; or even an estimation task (claim of same ability to measure a volume, say)
- Objective
  - Figure of merit is lesion detectability; size or shape discriminability; estimation EMSE
- Reliable
  - Error bars are provided to allow meaningful comparisons/conclusions
- Practical in terms of number of images, etc.
- Easily standardized

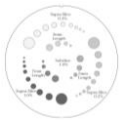
9

MITA-FDA consensus reached

- Standard metrics
  - MTF and NPS are building blocks for ideal observer performance for linear, shift-invariant systems with stationary noise
    - Task-based
  - No consensus in working group on relevance in terms of standard methods for measuring and combining these elements to determine image quality
- Significant literature validating models that predict human performance for simple detection/ discrimination tasks in images with variety of noise textures relevant to CT iterative reconstruction
  - Agreed on use of task-based image quality metrics with either human or model observers, using images of physical phantoms
  - Image based.

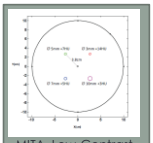
10

Commercial phantoms are not well suited to IQ studies.



Catphan®

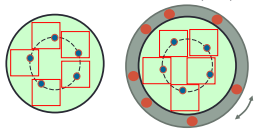
Objects too densely packed to allow search or signal-absent locations.



MITA Low-Contrast Detectability phantom

- Signals with multiple size/contrast combinations.
- Background ROIs from different locations in same slice, different slices in same scan, or different scans.

Courtesy L. Popescu

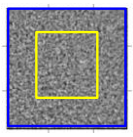


- Alternative concept: multiple identical objects at same radial location per phantom module.
- Randomly placed ROIs (extracted for human experiments) sets up search task.
- Model observer experiment can make use of larger search area; no need for extraction of ROIs.
- No need for background-only ROIs.

11

Advantage of search experiment

- More clinically relevant
  - Need to search, ability to use more realistic signal contrasts, more realistic probing of effect of noise texture
- More efficient use of image "real estate"
  - Fewer images for same statistical power
- Adjustment of ROI size tunes distribution of background-only test statistics
  - Enables post-data-collection flexibility in getting SNR in a useful range
- Applicable to humans or model observers
- Practical data analysis methods are available
  - Popescu, Med Phys 2007, 2011, 2013
  - Wunderlich and Noo, IEEE TMI 2012

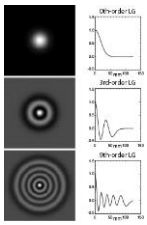


$SNR_1 > SNR_2$

12

Channelized observer models

- 4-5 channels can give reasonable estimates of performance with 10-25 images
  - Will depend on # of signal realizations per image and their detectability
- May need to train model observer for each condition
  - Account for differences in signal and noise properties
- Software is available

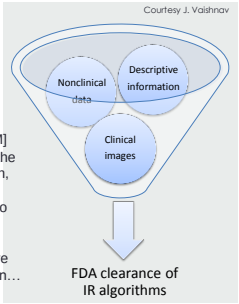


http://radiology.arizona.edu/cgri/image-quality/software/image-quality-toolbox
https://github.com/Barco-VCT/VirtualClinicalTrials

13

CDRH collaboration with MITA

- Codevelopment of phantoms; CDRH dissemination of software for estimation of figures of merit for model observers with confidence intervals
  - http://awunderlich.github.io/IQmodelo/
- Claims accompanied by
  - Tagging info: phantom, task, observer, figure of merit
  - Disclaimer:
  - "In clinical practice, the use of [ALGORITHM] may reduce CT patient dose depending on the clinical task, patient size, anatomical location, and clinical practice. A consultation with a radiologist and a physicist should be made to determine the appropriate dose to obtain diagnostic image quality for the particular clinical task. The following test methods were used to determine the level of dose reduction…



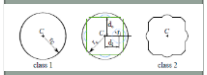
14

Where to go from here?

- Move beyond LCD tests to more complex discrimination and estimation tasks that reflect wider range of clinical applications and information needs of users.
- 3D printing, advanced manufacturing

Courtesy Gavrielides, Li, Petrick

Edge discrimination task



Hahn et al., Fully 3D 2015

https://wiki.cancerimagingarchive.net/display/Public/Phantom+FDA

15

What about use of simulation as regulatory evidence?

- Code for accurate modeling of the imaging physics (radiation transport) is available.
- Statistical analysis packages are available for determining figures of merit and uncertainties.
- Modeling of objects/patients and observers is precompetitive. Conducive to collaboration; need for incentives for code-sharing.
- How to handle black box, which includes IR and hardware?

16

Final remarks

- Evaluation of nonlinear reconstruction methods should be objective and task-based.
  - FDA/CDRH and others make tools available for modeling objects, imaging systems, observers, and for reader study design/analysis.
- NIH/NIBIB's CT U01 review criterion reflects this:
  - "If the application addresses or evaluates CT image quality in the context of the radiation dose reduction research strategy, are the image quality measurements and methods objective and appropriate?"
- Communities of practice and collaboration are essential.
  - Shared development, dissemination, and validation of better phantoms and accurate *in silico* clinical trials tools are key to enabling reliance on them for system evaluation and regulatory decision-making in the future.

THANK YOU!

18

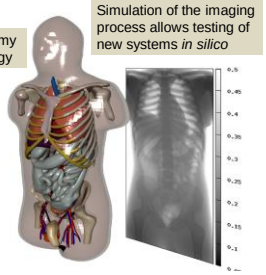
Computational Models for Medical Devices

"Reporting of Computational Modeling Studies in Medical Device Submissions - Draft Guidance for Industry and Food and Drug Administration Staff" (issued January 2014).

- Fluid Dynamics and Mass Transport
- Solid Mechanics
- Electromagnetics and Optics
- Ultrasound
- Heat Transfer

Virtual anatomy and physiology

Simulation of the imaging process allows testing of new systems *in silico*



19

Medical Device Development Tools

- A way for the FDA to qualify tools that medical device sponsors can use in the development and evaluation of medical devices.
- Qualification means that the FDA concurs with evidence that the tool produces scientifically-plausible measurements.
- Draft guidance available at:

www.fda.gov/RegulatoryInformation/Guidances

20

Medical Device Innovation Consortium
501(c)3 Public-Private Partnership
Members include FDA/CDRH, CMS, NIH, and Medical Device Industry

Computer Modeling and Simulation Project Vision

Quick and **Predictable** access for **Patients** to **Innovative** technologies enabled by Computation Modeling and Simulation as **Evidence** of safety and performance

The Future of Evidence

20

MDIC