Penalized Least-Square CT Reconstruction with and without Statistical Weights: Effect on Lesion Detection Performance with Human Observers

Viktor Haase[1,2], April Griffith[3], Zijia Guo[1,3], Mehmet B. Oktay[3], Katharina Hahn[2], Harald Schoendube[2], Karl Stierstorfer[2], and Frederic Noo[3]

[1] Pattern Recognition Lab, FAU Erlangen-Nuremberg, Germany
[2] Siemens Healthcare GmbH, Forchheim, Germany
[3] Department of Radiology and Imaging Sciences, University of Utah, USA

Fully3D – June 19, 2017

Introduction

Penalized Weighted Least-Square CT Reconstruction

Model-Based Iterative Reconstruction

- Model-based iterative reconstruction (MBIR) has been suggested as a robust way to reduce radiation dose while maintaining image quality
- Studies have demonstrated strong clinical potential of MBIR [1]-[3]
- Popular MBIR formulation: penalized least-square reconstruction with statistical weights (PWLS)

[1] Neroladaki et al. "**Computed tomography of the chest with model-based iterative reconstruction using a radiation exposure similar to chest X-ray examination: preliminary observations.**" (2013)
[2] Pickhardt et al. "**Abdominal CT with model-based iterative reconstruction (MBIR): Initial results of a prospective trial comparing ultralow-dose with standard-dose imaging.**" (2012)
[3] Ichikawa et al. "**CT of the chest with model-based, fully iterative reconstruction: comparison with adaptive statistical iterative reconstruction.**" (2013)

PWLS Problem Formulation

- Objective function: $\Phi(\underline{x},\underline{b}) = \Phi_1(\underline{x},\underline{b}) + \beta\,\Phi_2(\underline{x})$
- Desired solution: $\underline{x}^* = \arg\min_{\underline{x}\geq 0}(\Phi(\underline{x},\underline{b}))$
- Data fidelity: $\Phi_1(\underline{x},\underline{b}) = \left\|\mathbf{W}^{-1/2}(\mathbf{A}\underline{x}-\underline{b})\right\|^2$

Data Fidelity

- $\underline{x}$ image pixel values
- $\underline{b}$ CT measurements
- $\mathbf{A}$ forward projection matrix

- $\mathbf{W}$ diagonal matrix of statistical weights

PWLS Problem Formulation (cont'd)

- Objective function: $\Phi(\underline{x}, \underline{b}) = \Phi_1(\underline{x}, \underline{b}) + \beta\, \Phi_2(\underline{x})$
- Penalty term: $\Phi_2(\underline{x}) = \frac{1}{2} \sum_i \sum_j \omega_{ij} \psi(x_i - x_j)$

Regularization

- $\beta > 0$ regularization parameter
- $\omega_{ij} = \begin{cases} 1 & \text{for horizontal and vertical neighbor pixels} \\ 1/\sqrt{2} & \text{for diagonal neighbor pixels} \\ 0 & \text{otherwise} \end{cases}$

i. $\psi(t) = t^2$ quadratic potential

ii. $\psi(t) = \delta \cdot \left[\left| \frac{t}{\delta} \right| - \log\left(1 + \left| \frac{t}{\delta} \right| \right) \right]$ edge-preserving Fair potential, $\delta > 0$

Statistical Weights

- Data fidelity term: $\Phi_1(\underline{x}, \underline{b}) = \left\| \mathbf{W}^{-1/2}(\mathbf{A}\underline{x} - \underline{b}) \right\|^2,$
  with $\mathbf{W}$ as diagonal matrix of statistical weights
- Statistical weights represent the variance of the measurements
- Enable accounting for different noise levels across measurements
- Drawback: wide dynamic range complicates the development of an efficient iterative reconstruction algorithm

➢ What is the image quality improvement given by the statistical weights?

Evaluation Methodology

LROC Analysis with Human Observers: Examination of the Impact of Statistical Weights on Image Quality

LROC Analysis: Experimental Set-Up

- Data simulation:
  - Fan beam data, 3rd generation CT geometry
  - FORBILD head phantom
  - 360-degree circular scan, quarter-detector offset
  - 48 line integrals used to model each ray
  - Detector size: 0.75mm at FOV center
  - Poisson noise
  - No tube current modulation, realistic body-size bowtie filter
  - 40 keV monochromatic beam

➢ Choices made to accentuate the effect of statistical weights

SIEMENS Healthineers | HEALTH UNIVERSITY OF UTAH | FAU FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG SCHOOL OF ENGINEERING

LROC Analysis: Experimental Set-Up (cont'd)

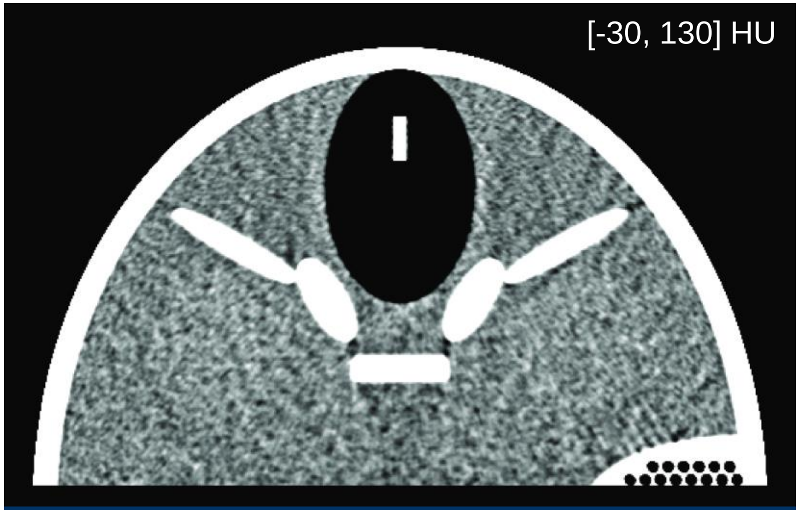
- Image reconstruction:
  - Image pixel size: 0.375mm
  - Same regularization parameter value for all reconstructions
  - Statistical weights exactly known and normalized
  - Distance-driven forward projector
  - Iterative coordinate descent (ICD) method used for reconstruction, iterations stopped when maximum pixel increment falls below 0.1HU

- Image display:
  - Grayscale: [-30,130] HU
  - Grayscale chosen to avoid pixel clipping within gray matter [1]

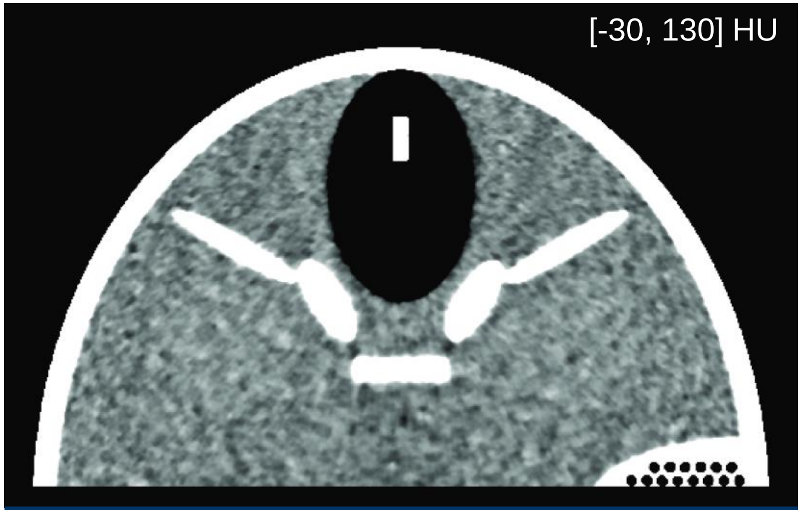
[1] Noo et al. "**Influence of the grayscale on phantom-based image quality assessment in x-ray computed tomography.**" (2015)

Examples of Reconstruction: With Statistical Weights

Computer-simulated fan beam data of FORBILD head phantom



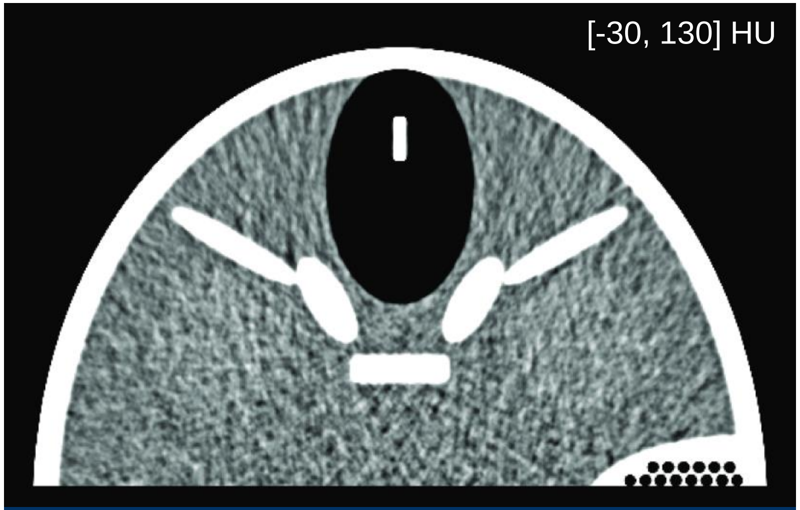
Reconstruction with quadratic potential



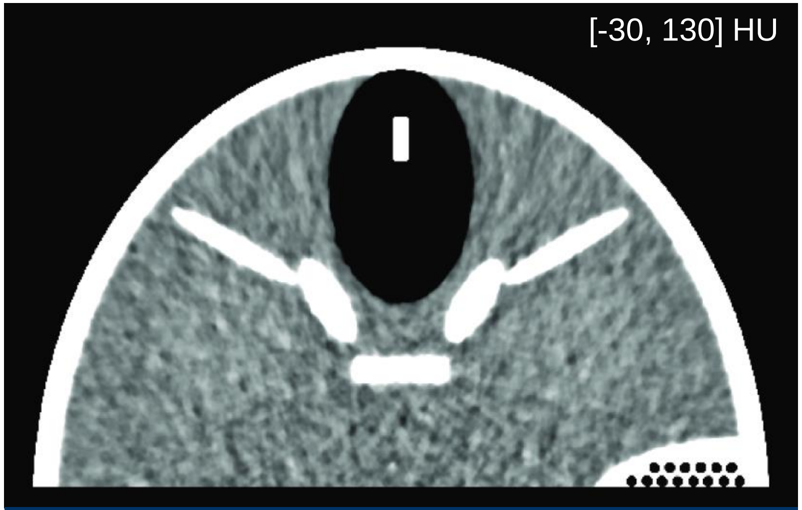
Reconstruction with Fair potential

Examples of Reconstruction: Without Statistical Weights

Computer-simulated fan beam data of FORBILD head phantom



Reconstruction with quadratic potential



Reconstruction with Fair potential

Image Quality Assessment Methodology

- LROC analysis with
  - Human readers
  - Mathematical phantoms

- Task description:
  - Detection of a round lesion within the FORBILD head phantom
  - Fixed lesion size: 7mm diameter
  - Random contrast: [20,30] HU

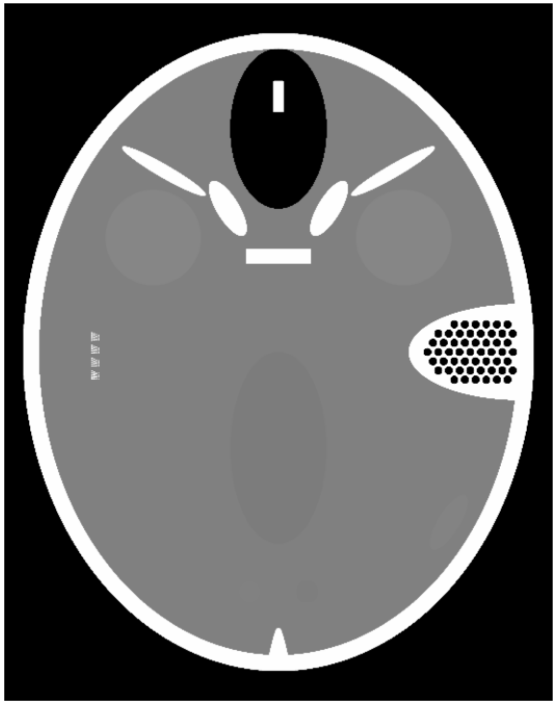


Image Quality Assessment Methodology

- LROC analysis with
  - Human readers
  - Mathematical phantoms
- Task description:
  - Detection of a round lesion within the FORBILD head phantom
  - Fixed lesion size: 7mm diameter
  - Random contrast: [20,30] HU
  - Random location within brain tissue
  - No overlap with bones

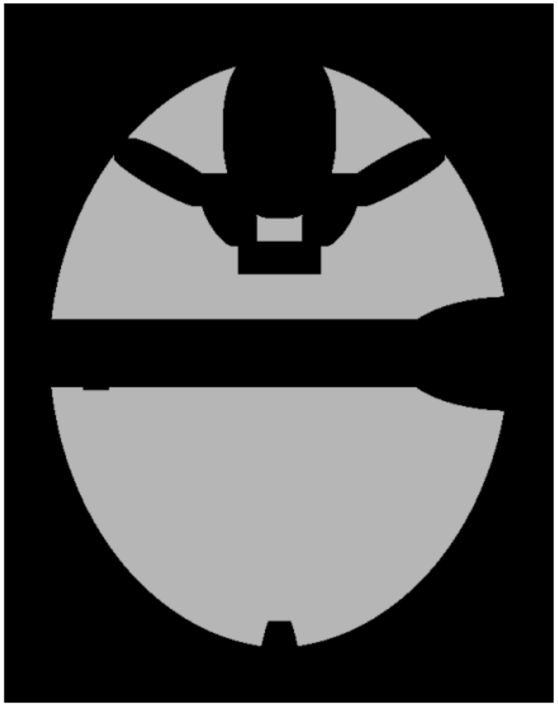
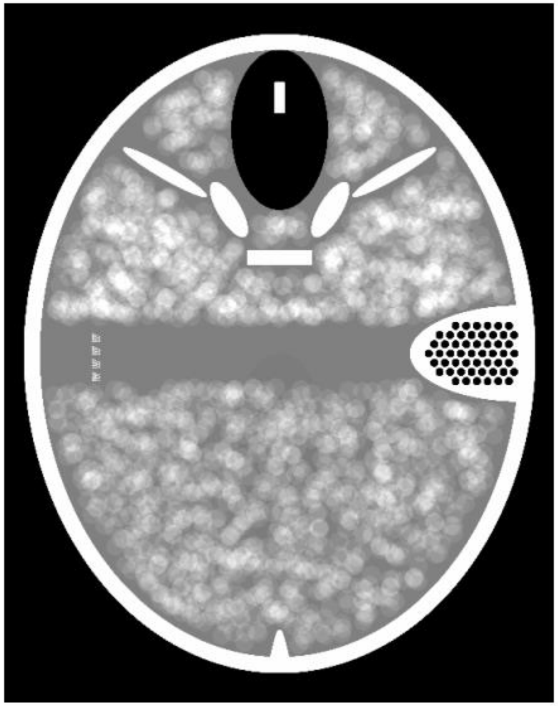


Image Quality Assessment Methodology

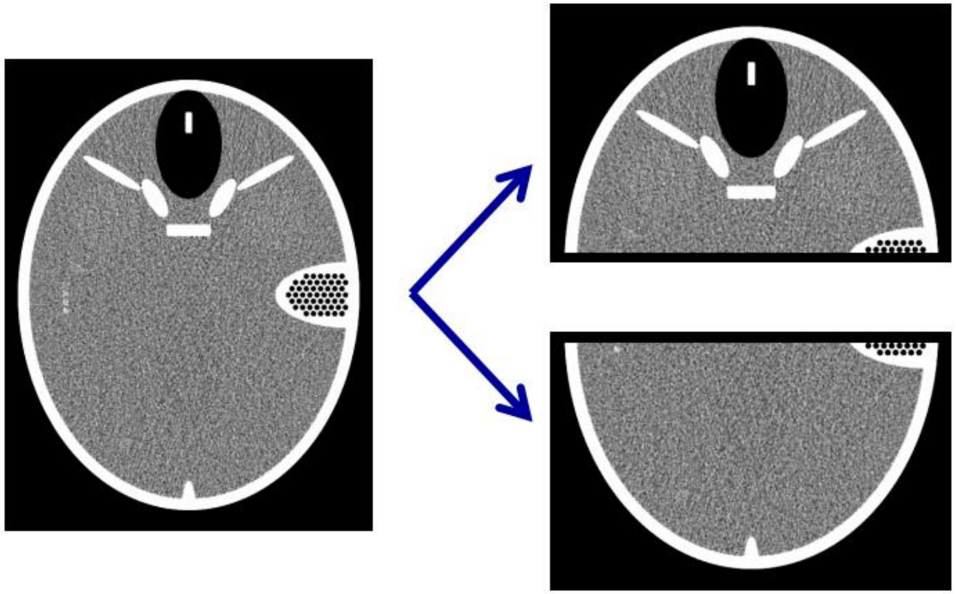
- LROC analysis with
  - Human readers
  - Mathematical phantoms

- Task description:
  - Detection of a round lesion within the FORBILD head phantom
  - Fixed lesion size: 7mm diameter
  - Random contrast: [20,30] HU
  - Random location within brain tissue
  - No overlap with bones



LROC Analysis: Experimental Set-Up (cont'd)

- Generation of cases:
  - 2 test cases per reconstruction
  - Cases statistical independent thanks to gap
  - Population defined by two parameters: upper/lower portion, scan repetition



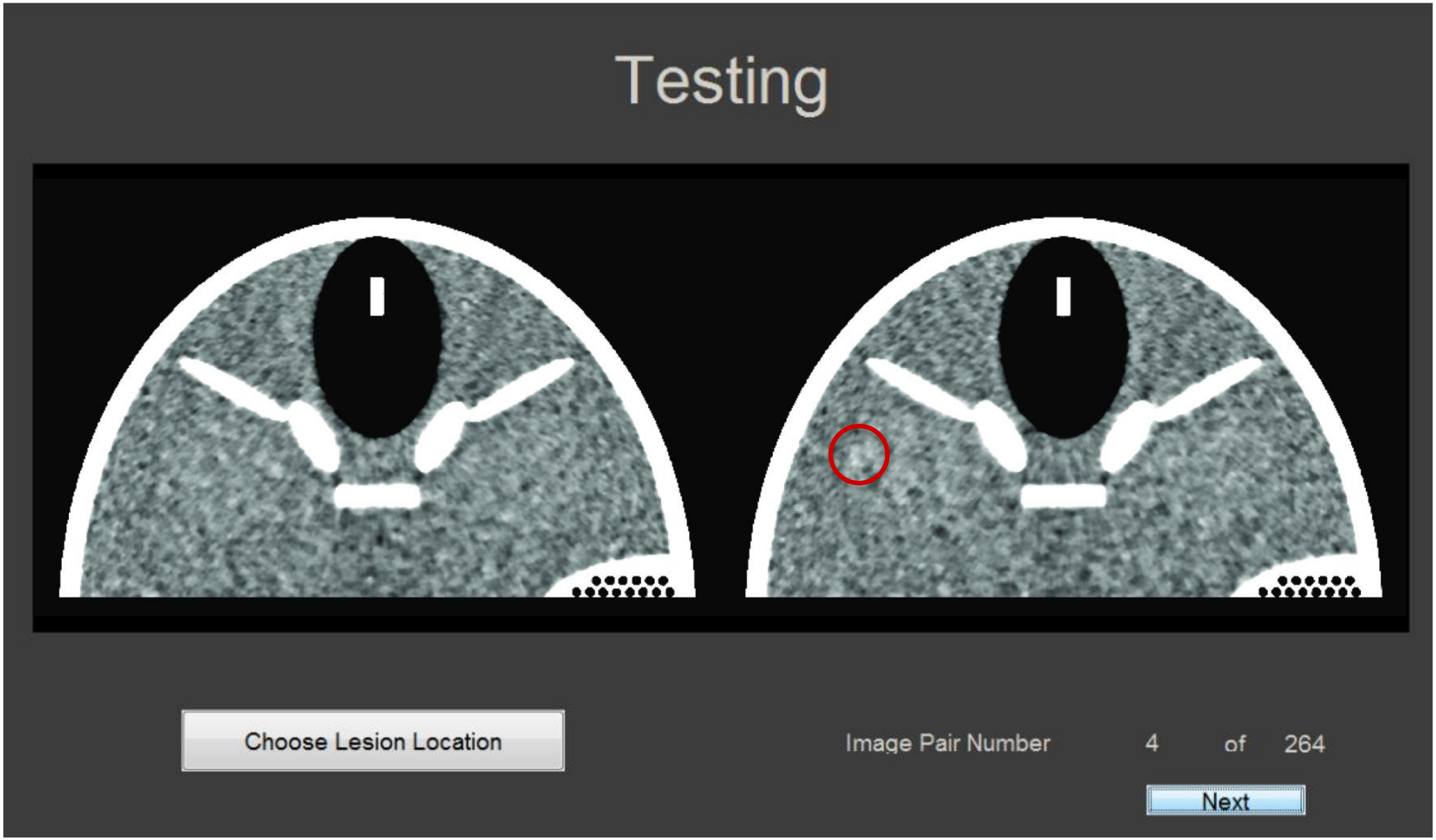
SIEMENS Healthineers | U HEALTH UNIVERSITY OF UTAH | FAU FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG SCHOOL OF ENGINEERING

LROC Analysis: Experimental Set-Up (cont'd)

- Human observers:
  - Reconstruction with quad. potential: 2 readers, both CT scientist
  - With Fair potential: 3 readers, 2 CT scientists and 1 radiology resident
- Two sessions per reader, parameters for one session:
  - 40 training and 160 testing per imaging scenario
  - 80+320 images total per session; reading time: about 90 minutes
  - Random ordering: scenarios and images across sessions and readers
  - Pre-calibrated medical grade monitor, dim room
- Figure-of-merit: area under LROC curve (AUC) is "probability correct" estimated using AFC experiments
- Statistical analysis: fixed-reader effects; data pairing across scenarios, but not readers; correlated linear combinations of proportions [1], [IQmodelo on GitHub]

[1] Noo et al. "**A nonparametric approach for statistical comparison of results from alternative forced choice experiments.**" (2013)

Alternative Forced-Choice Experiment

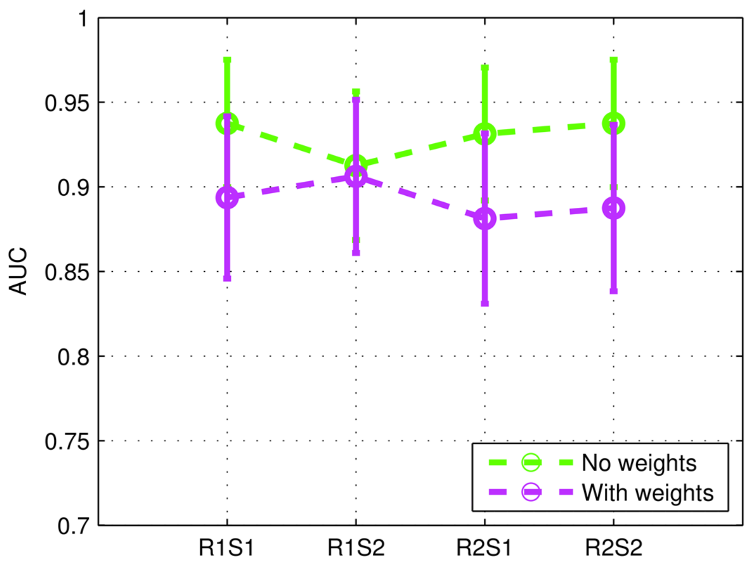


Results and Conclusions

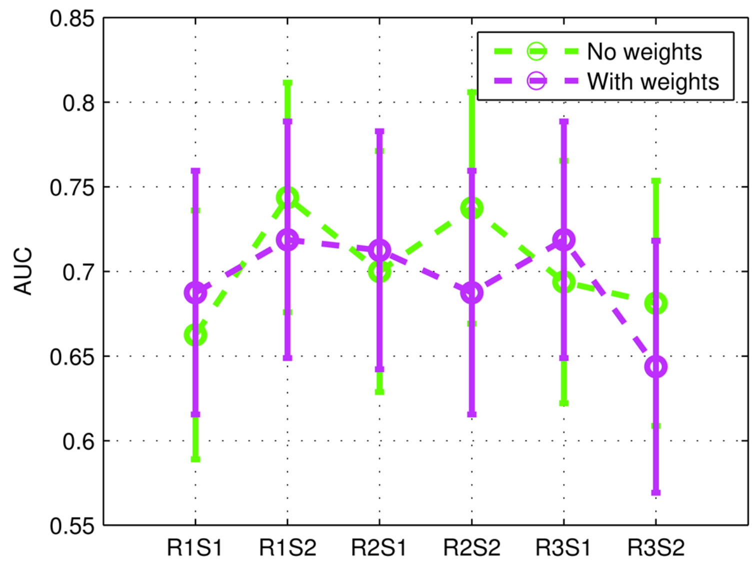
Detectability Performance

Results

Quadratic penalty



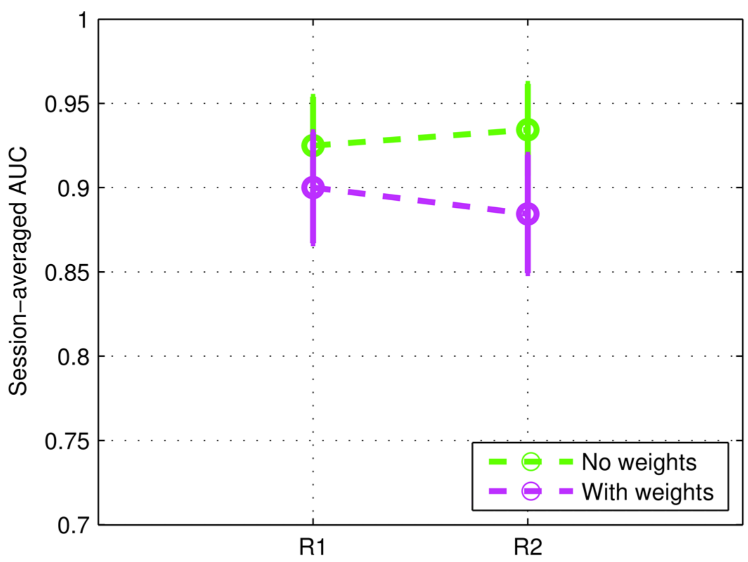
Edge-preserving penalty



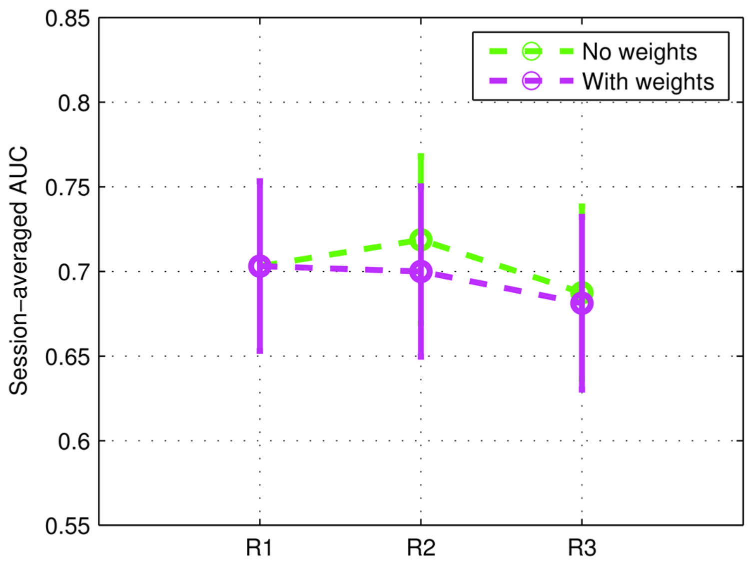
- AUC value obtained for each observer (R) and each session (S)

Results (cont'd)

Quadratic penalty



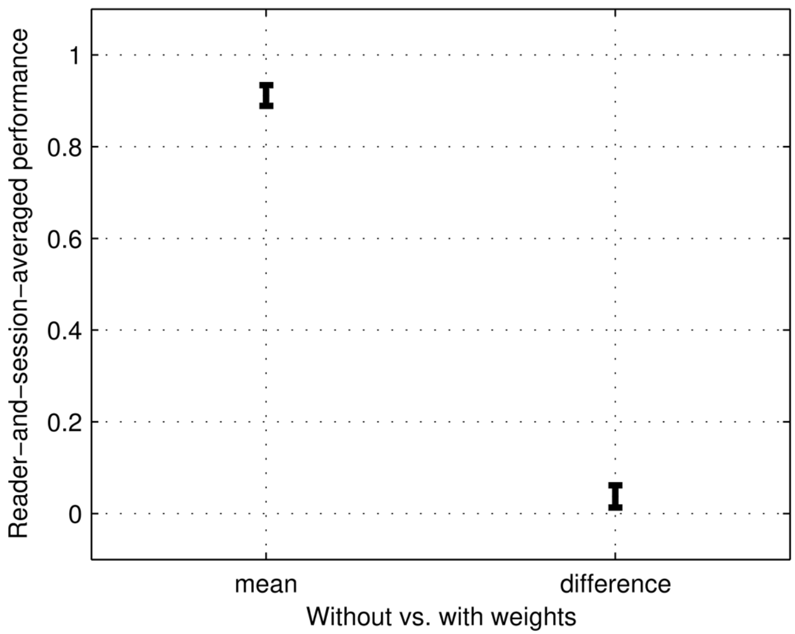
Edge-preserving penalty



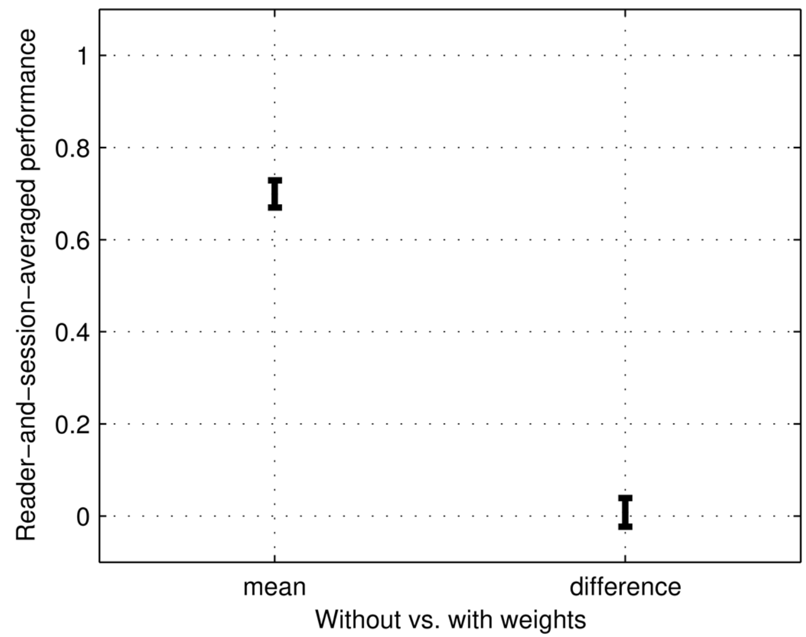
- AUC value obtained for each observer (R) after averaged over sessions

Results (cont'd)

Quadratic penalty



Edge-preserving penalty



- Difference in reader-and-session-averaged AUC reconstruction without weights and reconstruction with weights

Conclusions

- Image quality improvement using statistical weights is not straightforward (impact on computational effort)
- Statistical weights can induce confusing aliasing errors
- Agreement with knowledge that human observers can pre-whiten images to a certain extent [1]
- Results likely dependent on task and image regularization
- Improvements may be possible using smooth weights with weight-based pixel-dependent regularization strength [2]

[1] Abbey et al. "**Classification images for simple detection and discrimination tasks in correlated noise.**" (2007)
[2] Cho et al. "**Regularization designs for uniform spatial resolution and noise properties in statistical image reconstruction for 3-d x-ray CT.**" (2015)

Thank you for your attention!

Any questions?