Challenges and opportunities in assessment of image quality

Ehsan Samei
Department of Radiology
Duke University Medical Center

Disclosures

- Research grant: NIH R01 EB001838
- Research grant: General Electric
- Research grant: Siemens Medical
- Research grant: Carestream Health
Imaging quality and safety

- Quality: Imaging provides a clinical benefit
  - Diagnostic information
  - Image quality: Universally-appreciated
  - Enhancement dependent on illusive criteria
- Safety: Imaging involves a level of “cost”
  - Monitory cost
  - Information excess
  - Radiation cost

Optimization involves 4 steps:
1. Reasonable measures of risk
2. Reasonable measures of image quality
3. Justified balance between the two by targeted adjustment of system parameters
4. Consistent implementation

We cannot optimize imaging without quantification
Outline

1. Quality Characterization
2. Quality assessment
3. Quality implementation
Model-based recons +s

- Decades in the making
- Enabled by high-powered computers
- Significant potential for dose reduction
- Potential for improved image quality

Model-based recons -s

- Speed
- Increased vendor-dependence
- Unconventional image appearance

- Limited utility of prior quality metrics
- Need for nuanced implementation for effective improvement in patient care

ASiR (GE)
Veo (GE)
IRIS (Siemens)
SAFIRE (Siemens)
AIDR (Toshiba)
iDose (Philips)
...

Noise texture?

Images courtesy of Dr. de Mey and Dr. Nieboer, UZ Brussels, Belgium
Iterative Reconstruction

FBP Reconstruction

Images courtesy of Mayo Clinic Rochester, MN

Low Dose CT @ 338 DLP, 1.8 mSv

Images courtesy of Dr. de Mey and Dr. Nieboer, UZ Brussel, Belgium
Qualimetrics 1.0: CNR

- 1st order approximation of image quality
- Related to detectability for constant resolution and noise texture (Rose, 1948)
- Task-generic

\[
\text{CNR} = \frac{Q_I}{Q_B} \frac{Q_I}{Q_B}
\]

Parameters that affect IQ

1. Contrast
2. Lesion size
3. Lesion shape
4. Lesion edge profile
5. Resolution
6. Viewing distance
7. Display
8. Noise magnitude
9. Noise texture
10. Operator noise

Feature of interest
Image details
Distractors

Parameters that are measured by CNR

1. Contrast
2. Lesion size
3. Lesion shape
4. Lesion edge profile
5. Resolution
6. Viewing distance
7. Display
8. Noise magnitude
9. Noise texture
10. Operator noise

Feature of interest
Image details
Distractors
Why CNR is not enough: Noise texture

Resolution and noise, eg 1

Resolution and noise, eg 2
**NPS vs dose**

Nonlinearity


**NPS peak frequency vs dose**


**Resolution’s joint dependence on contrast and dose**

Detectability index

Resolution and contrast transfer
Attributes of image feature of interest
Image noise magnitude and texture

\[
(d_{NPW})^2 = \frac{\left[ \text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) \right]}{\text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) + \text{MTF}^2(u,v)W_{NPS}(u,v)N_i} \]

Chen et al., Relevance of MTF and NPS in quantitative CT: towards developing a predictable model of quantitative... SPIE 2012

Task-based quality index

Fisher-Hotelling observer (FH)

\[
(d_{FH})^2 = \frac{\left[ \text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) \right]^2}{\text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) + \text{MTF}^2(u,v)W_{NPS}(u,v)N_i} \]

Non-prewhitening observer (NPW)

\[
(d_{NPW})^2 = \frac{\left[ \text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) \right]^2}{\text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) + \text{MTF}^2(u,v)W_{NPS}(u,v)N_i} \]

NPW observer with eye filter (NPWE)

\[
(d_{NPWE})^2 = \frac{\left[ \text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) \right]^2}{\text{MTF}^2(u,v)W_{NPS}(u,v)E^2(u,v) + \text{MTF}^2(u,v)W_{NPS}(u,v)N_i} \]

Chen, SPIE 2013

Task function:

\[
C(r) = C_{peak}(1 - \frac{r}{R})^n \]

Task characteristics for detection and estimation

Chen, SPIE 2013
**GE MBIR Dose Reduction Potential**

**Large Feature Detection Task**

Richard, Li, Samei, SPIE 2011

**Small Feature Detection Task**

Richard, Li, Samei, SPIE 2011

**Duke-UUMD-NIST study**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GE Discovery CT 750 HD</th>
<th>Siemens Flash</th>
<th>Philips ICT</th>
</tr>
</thead>
<tbody>
<tr>
<td>kVp</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Rotation time</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>SFOV</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>DFOV</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Recon Algorithm</td>
<td>FBP Standard / ASIRbB</td>
<td>B31f / 13of Safire 3</td>
<td>B / iDoseS</td>
</tr>
<tr>
<td>Recon Mode</td>
<td>Helical</td>
<td>Helical</td>
<td>Helical</td>
</tr>
<tr>
<td>Collimation</td>
<td>0.625</td>
<td>0.6</td>
<td>0.625</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.984</td>
<td>1</td>
<td>0.93</td>
</tr>
<tr>
<td>Slice Thickness</td>
<td>1.25</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Dose levels: 20, 12, 7.2, 4.3, 1.6, 0.9 mGy

For anonymity will be referred to as vendor A, B, and C.
Observer study design

3 scanner models
3 dose levels
2 reconstruction algorithms
10 slices
5 repeated exams
Total of 900 images

12 expert observers from two institutions

How much can IR reduce dose?

Vendor A

Default clinical protocol at 12 mGy
23% dose reduction
How much can IR reduce dose?

**Vendor B**
Default clinical protocol at 12 mGy
6% dose reduction

**Vendor C**
Default clinical protocol at 12 mGy
33% dose reduction

CNR vs observer performance

\[ y = 0.2816x + 0.4569 \]
\[ R^2 = 0.4674 \]
d’ vs observer performance

\[ y = 0.9874x + 1.8585 \]
\[ R^2 = 0.9202 \]

Our contrast detail phantom

CT images of the low-contrast phantom
Acquired CT data

- Siemens SOMATOM Force
- 4 doses (0.7, 1.4, 2.9, 5.8 mGy)
- 2 slice thicknesses (0.6, 5 mm)
- 4 Recons (FBP, ADMIRE 3-5)

TTF and NPS across recons

CT images across dose and recon
Insert Counting Experiment

- Visible groups across:
  - Dose (0.74, 1.4, 2.9, 5.8 mGy)
  - Slice Thickness (0.6, 1.8, 5 mm)
  - Recon (FBP, ADMIRE 3-5)

Total number of discernable objects

IQ metrics vs human performance

CTDIvol (mGy)
Task-Based Detectability
Extended to quadratic PL image reconstruction

Detectability Map d'(x,y)


Justification of Assumptions:
Local Stationarity and Linearity

% difference between Fourier and spatial domain calculation of d'

Ellipse phantom

% difference

Cylinder

Ellipse

Thorax

Similar levels of d' computed via Fourier and spatial domain approach (~5-15% agreement)
Applicable to various objects and imaging tasks.
Local Fourier metrology applicable to quadratic PL reconstruction, gives a useful framework for system design and optimization.


Estimability index (e'):
prediction of the quantification precision capturing the interaction between the imaging system, the segmentation software, and the lesion

\[ e' = \frac{\int \left[ TF^2(u,v,w) \cdot N(u,v,w) \right] dudvdw}{\int \left[ NPS(u,v,w) + N(u,v,w) \right] dudvdw} \]
Empirical precision measurement

Predicted precision matches empirical precision!

\[ PRC = \frac{1.96 \sqrt{\frac{\text{var}(X)}{n}}}{\text{var}(Y) / \sqrt{n}} \]

\[ = 2.77 \sqrt{\sum_{i=1}^{n} \left( \frac{Y_i - \bar{Y}}{\text{var}(Y) / \sqrt{n}} \right)^2} \]

Outline

1. Quality Characterization
2. Quality assessment
3. Quality implementation
Task-based assessment metrology

Mercury Phantom 3.0

- Diameters matching population cohorts
- Depths consistent with cone angles
- Straight-tapered design enabling evaluation of AEC response to discrete and continuous size transitions

Design: Base material

- Polyethylene
  - 80 HU @ 120 KVp
  - Near patient equivalent
  - Affordable
  - Easy to machine

Design: Size

Pediatric representation percentages

<table>
<thead>
<tr>
<th>MP 3.0 section</th>
<th>Water size equivalent</th>
<th>Abdomen Age</th>
<th>Abdomen Percentile</th>
<th>Chest Age</th>
<th>Chest Percentile</th>
<th>Head Age</th>
<th>Head Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>120 mm</td>
<td>112 mm</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>185 mm</td>
<td>177 mm</td>
<td>7</td>
<td>5</td>
<td>5.5</td>
<td>5</td>
<td>50</td>
<td>95</td>
</tr>
<tr>
<td>230 mm</td>
<td>220 mm</td>
<td>15</td>
<td>5</td>
<td>10</td>
<td>50</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td>300 mm</td>
<td>290 mm</td>
<td>21</td>
<td>5</td>
<td>16</td>
<td>50</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>370 mm</td>
<td>355 mm</td>
<td>20</td>
<td>95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Design: Size
Adult representation percentages

<table>
<thead>
<tr>
<th>MP section</th>
<th>Water size equivalent</th>
<th>Abdomen M</th>
<th>F</th>
<th>Chest M</th>
<th>F</th>
<th>Head M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>120 mm</td>
<td>112 mm</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>185 mm</td>
<td>177 mm</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>230 mm</td>
<td>220 mm</td>
<td>0.4</td>
<td>9</td>
<td>0.06</td>
<td>1.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>300 mm</td>
<td>290 mm</td>
<td>27.1</td>
<td>61</td>
<td>14</td>
<td>48</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>370 mm</td>
<td>355 mm</td>
<td>80</td>
<td>90.3</td>
<td>60</td>
<td>87</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Design: Resolution, HU, noise

- Representation of abnormality-relevant HUs
- Sizes large enough for resolution sampling
- Maximum margin for individual assessment
- Iso-radius resolution properties
- Matching uniform section for noise assessment

Non-Stationary Noise and Resolution

Noise and resolution model for Penalized Likelihood (PL) model-based reconstruction.* Predictive framework for NPS, MTF, and detectability index (d’) enables task-based design and optimization of new systems using iterative reconstruction.

**Design: Resolution, HU, noise**

Sections 2-5

Smallest Section 1

---

**imQuest**

*(image quality evaluation software)*

HU, Contrast, Noise, CNR, MTF, NPS, and \( d' \) per patient size, mA modulation profile

---

**Noise texture effect**

IR noise reduction: 65% in uniform bkgd

20% in textured bkgd
Mercury phantom 3.0
Added contrast detail and texture

Textured lung phantom

Textured soft-tissue phantom:
clustered lumpy background*

Textured soft-tissue phantom: clustered lumpy background*

Anatomically informed heterogeneity phantoms

Lung texture phantom
Soft-tissue texture phantom

How was the data analyzed?

\[ I(\hat{x}) = \mu(\hat{x}) + N(\hat{x}) \]

\[ \sigma(\hat{x}) = \text{STD} \]

Noise in the uniform phantom

- FBP
- IR
Noise in the soft-tissue phantom

<table>
<thead>
<tr>
<th>FBP</th>
<th>IR</th>
</tr>
</thead>
</table>

Noise in the lung phantom

<table>
<thead>
<tr>
<th>FBP</th>
<th>IR</th>
</tr>
</thead>
</table>

Comparing noise across textures

<table>
<thead>
<tr>
<th>FBP</th>
<th>IR</th>
</tr>
</thead>
</table>
What about noise texture?

\[ f(x) = \mu(x) + N(x) \]

Lung Texture NPS

- FBP: Red Pixels
- IR: Red Pixels
- IR: Blue Pixels

Outline

1. Quality Characterization
2. Quality assessment
3. Quality implementation

Components of quality assurance

- System performance assessment
  - Quality by inference
- Prospective protocol definition
  - Quality by prescription
- Retrospective performance assessment
  - Quality by outcome
**kV IR optimization**

<table>
<thead>
<tr>
<th>Task</th>
<th>Optimal technique</th>
<th>% dose reduction (wrt 120 kVp FBP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No iodine</td>
<td>80/100 kVp with IRIS</td>
<td>36%</td>
</tr>
<tr>
<td>With iodine</td>
<td>80 kVp with IRIS</td>
<td>43%</td>
</tr>
</tbody>
</table>

Samei, Richard, Med Phys, in press, 2014

---

**Detectability trends**

Smitherman, AAPM, 2014

---

**Protocol optimization**

Setting dose to achieve a targeted task performance for a given size patient

Smitherman, AAPM, 2014
Protocol optimization

- Setting dose to achieve a targeted task performance for a given size patient

Smitherman, AAPM, 2014

Quality-dose dependency

Quantitative volumetry via CT

PRC: Relative difference between any two repeated quantifications of a nodule with 95% confidence

GE

Siemens

Solomon, Samet, Med Phys, 2012
Texture similarity

Solomon, Samei, Med Phys, 2012

CT image quality monitoring

Christianson, AAPM, 2014

Noise per slice

FBP

ASIR
Trends in dose and noise

Abdomen Pelvis Exams (n=2358)

SSDE (mGy) vs. Effective Diameter (cm)

Scanner 1
Scanner 2
Scanner 3

SSDE varies with size

Christianson, AAPM, 2014

Variability in dose and noise

Abdomen Pelvis Exams (n=2358)

SSDE (mGy) vs. GNL (HU)

Scanner 1
Scanner 2
Scanner 3

Target noise level

ACR DIR 25th - 75th

Christianson, AAPM, 2014
Conclusions

• New technologies necessitate an upgrade to performance metrology towards higher degrees of clinical relevance:
  • Physical surrogates of clinical performance
  • “Taskful” metrology and dependencies
  • Incorporation of texture into image quality estimation
  • Extension of quality metrology to quality monitoring and quality analytics

Thank you!