Radiological scores of chest X-rays presented with a green LUT

Poster No.: C-2382
Congress: ECR 2013
Type: Scientific Exhibit
Keywords: Thorax, Computer applications, Conventional radiography, Digital radiography, Computer Applications-Detection, diagnosis, Observer performance, Perception image, Quality assurance
DOI: 10.1594/ecr2013/C-2382

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Purpose

Despite the continuing digitalization in medical imaging, X-ray images are still being displayed using a grey scale, sticking to the conventions from the analogue era. However, with the availability of high quality, medical grade color monitors, there is an increasing interest for applications of color images in radiology (e.g. MRI overlays, PET-CT).

Applying a colored look-up table (LUT) on X-ray images may have some advantage because of the spectral sensitivity of the human eye. The green color is potentially useful as the human eye is most sensitive for green. The use of color combinations represents additional possibilities for optimal visualisation. The superiority of a green LUT for detection studies with a contrast detail phantom has been found in a parallel study by this group (to be published). The present study focuses on the radiologist's appreciation of chest X-rays shown with a green and grey LUT, using a range of exposure settings.
Methods and Materials

Subjects

This study was approved by the Ethical Committee at KU Leuven. Two dead bodies (normal-sized and obese) were used. The bodies were positioned on an 8mm thick PMMA plate and imaged using a bedside X-ray system (XP Hybrid radiography system, Siemens, Erlangen, Germany) and a wireless DRX-1C detector (Carestream, Rochester, NY, USA) that was positioned on a holder under the plate with a gap of 10cm between the body and the detector.

Typical bedside exams of the chest were mimicked (Table 1). Different combinations of tube load, tube voltage, grid resolution (40 or 80 L/cm) and grid positions (aligned or shifted) were included. The total number of chest X-ray images of the normal weight bodies was 65. The obese body was exposed with other exposure settings, compatible with the different body size and the tube load capacity of the X-ray system (Table 2). The total images acquired for the obese body was 14.

Visual grading analysis

Visual grading analysis consisted of scoring the quality of particular anatomical sub-regions as well as the global quality of the image on a 5-point scale (i.e. 1, 3, 5, 7, 9) ranging from poor (1) to excellent (9). In addition, the acceptability (yes/no) of the image for diagnostic work was evaluated. Four anatomical regions were defined: clavicle bone, ribs, lung tissue and pedicle (Figure 1). Two experienced radiologists with over 15 years experience in digital radiology and four younger colleagues with at least 1 year of experience scored all images twice using a grey and green LUT. The reading study was performed using the SARA software platform [6] on a 6 megapixel Barco (Barco, Kortrijk, Belgium) monitor. The ambient light was kept below 10 lux. Total reading time was between 3 and 4h.

Technical evaluation and optimization study of the new technology

The regular commissioning test of the tube and detector included standard measurements of the performance of the tube generator, the signal transfer properties (STP), which relate pixel values in raw images to detector air kerma, and detective quantum efficiency (DQE). The system provides also an exposure indicator (EI) that is obtained with a proprietary algorithm. All dose measurements were performed with a Barracuda dosimeter (RTI, Lund, Sweden).
A technical optimization study has been described elsewhere [10], and we provide a short summary here. The figure of merit (FOM) for the optimization process was inspired by Samei et al. [4], namely:

$$FOM = \frac{SDNR^2}{Dose}$$

Fig. 11

References: Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE

in which $SDNR$ is a clinically relevant signal difference to noise ratio and $Dose$ a measure related with the radiation induced detriment of the patient. We calculated the Energy Imparted [5] as an alternative to Effective Dose. The SDNR was calculated from a piece of Cu (1mm thick and size 1cmx1cm) fixed on top of 5 cm of PMMA with the additional PMMA slabs on top. A large series of exposure settings was performed, using all combinations of PMMA thickness (5, 10, 15, 20 and 30cm), tube voltage (75kV, 90kV and 113kV), EI values (900, 1200, 1500, 1800, 2100) and grids. The effect of grid misalignment on image quality was tested with a similar set-up, shifting the grid from 5cm to up to 30cm and adjusting the X-ray tube to hit the center of the grid.

Images of the dead bodies were linked to corresponding SDNR values by estimating the equivalent amount of PMMA for the two groups of dead bodies. The imparted energy (in mJ) was then calculated for all dead bodies and for all exposure conditions, following the paper by Huda et al [5] from the following parameters: tube voltage, half value layer, PMMA thickness and tube load.

Analysis of visual grading

All reading scores were analyzed pairwise (green vs. grey LUT). Data of normal size and overweight bodies was processed separately. In addition to the readings for all anatomical regions separately, we also calculated the ‘global anatomical score’ as the sum of all four anatomical scores per image.
The distribution of the green and grey LUT scores for the readings of all 4 anatomical regions was studied for all observers separately. The differences between the global anatomical scores for images shown with green and grey LUT were then studied as a function of imparted energy, SDNR and linearized exposure indicator (LEI). The scores for the parameters 'global appreciation of the image' and 'acceptability for diagnostic purposes' were analyzed as a function of imparted energy. Finally, the differences in global anatomical scores were plotted separately for the different tube voltages and for the types of grids used.
**Table 1:** Overview of exposure settings on the standard size dead bodies

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Table 2: Overview of exposure settings on the overweight dead bodies

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE

<table>
<thead>
<tr>
<th>Nr. of images</th>
<th>kVp</th>
<th>mAs</th>
<th>grid</th>
<th>grid shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>75</td>
<td>2.25; 4.5; 9; 14</td>
<td>none</td>
<td>aligned</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>2.8; 11; 22</td>
<td>low-res</td>
<td>aligned</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>2.8; 5.60; 11; 22</td>
<td>high-res</td>
<td>aligned</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>2.8; 22</td>
<td>low-res</td>
<td>10cm</td>
</tr>
<tr>
<td>1</td>
<td>90</td>
<td>22</td>
<td>high-res</td>
<td>10cm</td>
</tr>
</tbody>
</table>

Fig. 1: Overview of the scoring task performed for each image

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Results

Table 3 shows, for each observer, the number of times an image shown with green LUT results in a higher score than the same image shown with grey LUT. There are remarkable differences between observers, with increased green scores ranging between 4.3% and 33.7%. The two observers with the lowest percentage of increased scores for the green LUT had the longest experience. The same trend is seen for the analysis of the overweight cadaver (Table 4). One observer gave higher score to the green LUT display in 42.9% of the images.

The difference in global anatomical score was then plotted as a function of imparted energy in different ways: all the data of all observers and the difference in total anatomical score averaged over all observers (Fig 2). A large cloud of difference scores is seen, with values ranging between -12 and +15. The maximal possible range for this graph is -32 til +32 as all anatomical details are scored on a scale from 1 to 9. Fig 3 show the same graphs for the separate observers. Confirming the values from Table 2, the graphs show obvious differences in scoring between observers.

There is no clear trend in the data that suggests any further trend analysis. The main impression from the data, apart from the variability between observers, is that larger negative difference scores occur for images at lower impacted energy, although this is not the case for each observer. The graphs for the overweight bodies (Fig 4 and Fig 5) confirm these findings.

Fig 6 shows the differences in total anatomical score (for green and grey display) as a function of SDNR for all observers separately. In Fig 7 the results for the overweight body are shown. Image perception differences seem not to be influenced by the SDNR. Fig 8 illustrates the absence of any trend in comparative graphs for the difference in anatomical score as a function of imparted energy, SDNR and LEI. Fig 9 shows the assessment of the image for diagnostic acceptability and general impression as a function of imparted energy. Again, there are no noticeable trends. As there was no clear effect seen from SDNR or imparted energy, there was no effect from the FOM either.

The overall impact of grid and tube voltage on image quality was assessed. These separate analyses were performed for the normal size bodies only as the number of images of the overweight body was insufficient for further categorization of the data. The differences in score for different grid types and for tube voltage seem not affected by any of these parameters. The analysis does suggest that the difference scores are more favorable for the green display for the higher quality images (high SDNR, high resolution grid, high exposure index) than for the lower quality images. Fig 10 illustrates higher
scores for images obtained from the normal size body. Furthermore, scores were higher for images obtained at lower tube voltages, images acquired using a grid, with high-resolution grids resulting in the best scores.

In an interview performed after the reading session, two radiologists (with the poorest scores for the green readings) disliked the green LUT for display, whereas the four other radiologists had experienced some benefit. They all declared explicitly that today they would only accept a grey display for routine work.
### Table 3: Percentage of images with increased score for green display; normal size dead body

<table>
<thead>
<tr>
<th></th>
<th>clavicle</th>
<th>ribs</th>
<th>lung</th>
<th>pedicle</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs 1</td>
<td>0.0</td>
<td>2.9</td>
<td>17.4</td>
<td>15.9</td>
<td>9.1</td>
</tr>
<tr>
<td>obs 2</td>
<td>15.9</td>
<td>18.8</td>
<td>7.2</td>
<td>17.4</td>
<td>14.9</td>
</tr>
<tr>
<td>obs 3</td>
<td>14.5</td>
<td>39.1</td>
<td>21.7</td>
<td>14.5</td>
<td>22.5</td>
</tr>
<tr>
<td>obs 4</td>
<td>26.1</td>
<td>47.8</td>
<td>39.1</td>
<td>21.7</td>
<td>33.7</td>
</tr>
<tr>
<td>obs 5</td>
<td>14.5</td>
<td>13.0</td>
<td>10.1</td>
<td>8.7</td>
<td>11.6</td>
</tr>
<tr>
<td>obs 6</td>
<td>0.0</td>
<td>4.3</td>
<td>10.1</td>
<td>2.9</td>
<td>4.3</td>
</tr>
</tbody>
</table>

### Table 4: Percentage of images with increased score for green display; obese dead body

<table>
<thead>
<tr>
<th></th>
<th>clavicle</th>
<th>ribs</th>
<th>lung</th>
<th>pedicle</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs 1</td>
<td>0.0</td>
<td>7.1</td>
<td>0.0</td>
<td>7.1</td>
<td>3.6</td>
</tr>
<tr>
<td>obs 2</td>
<td>35.7</td>
<td>64.3</td>
<td>71.4</td>
<td>0.0</td>
<td>42.9</td>
</tr>
<tr>
<td>obs 3</td>
<td>14.3</td>
<td>14.3</td>
<td>28.6</td>
<td>7.1</td>
<td>16.1</td>
</tr>
<tr>
<td>obs 4</td>
<td>0.0</td>
<td>14.3</td>
<td>35.7</td>
<td>42.9</td>
<td>23.2</td>
</tr>
<tr>
<td>obs 5</td>
<td>0.0</td>
<td>7.1</td>
<td>21.4</td>
<td>0.0</td>
<td>7.1</td>
</tr>
<tr>
<td>obs 6</td>
<td>7.1</td>
<td>7.1</td>
<td>0.0</td>
<td>0.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>
**Fig. 2:** Difference in scores (Score with green display - score with grey display) of all anatomical regions for all observers (top) and averaged (bottom), for the normal size cadavers

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 3: Difference in scores (Score with green display - score with grey display) of all anatomical regions for all observers separately and for the normal size cadavers, plotted as a function of imparted energy

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 4: Difference in scores (Score with green display - score with grey display) of all anatomical regions for all observers (top) and averaged (bottom), for the overweight cadaver, plotted as a function of imparted energy

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 5: Difference in scores (Score with green display - score with grey display) of all anatomical regions for all observers separately and for the overweight cadaver, plotted as a function of imparted energy

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 6: Difference scores (Score with green display - score with grey display) averaged over all anatomical regions for all observers separately and for the normal size cadavers, plotted as a function of SDNR.

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 7: Difference scores (Score with green display - score with grey display) of averaged over all anatomical regions for all observers separately and for overweight cadaver, plotted as a function of SDNR

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
**Fig. 8:** Average difference scores (Score with green display - score with grey display) of all anatomical regions as averaged over all observers plotted as a function of imparted energy, SDNR and linearized exposure indicator (LEI). All data of standard size cadavers.

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 9: Average difference scores (Score with green display - score with grey display) of scores for 'diagnostic quality' (top) and 'general impression' (bottom) as averaged over all observers and plotted as a function of imparted energy. All data of standard size cadavers.

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Fig. 10: Influence of body size, kV value, grid use and grid resolution on observer scores

© Oral Imaging Center, Department of Dentistry, Oral Pathology and Maxillo-Facial Surgery, Catholic University of Leuven - Leuven/BE
Conclusion

Notwithstanding the better sensitivity of the human eye for green, this was not reflected in an overall positive difference scores in favor of a green LUT. There is therefore no stringent reason to implement green LUTs. This is especially the case for the lower dose images, where the appreciation of the images shown with green LUT tended more different than in higher dose images. It was seen that experienced radiologists judged the quality of a green LUT display worse than younger colleagues do.

The authors want to thank David H. Foos and William J. Sehnert (Carestream, Rochester, NY), Wim Op De Beeck (Carestream, Belgium) and em. Prof. Guy Marchal, MD, PhD for discussing optimization protocols for bedside X-ray imaging. We thank Wim Develter, MD, in the department of forensic medicine for practical help with acquiring the images.

This work was part of the CIMI project, co-funded by IBBT, a research institute founded by the Flemish Government.
References


Personal Information

Hilde Bosmans, PhD
Radiology
Department of Imaging & Pathology
Biomedical Sciences Group
KU Leuven