Does Image Quality Matter?

Author: *Dr. Ralph Schaetzing*

Given a set of images, most of us can generally choose a best one—but do we always need the best one to accomplish a given task? The relatively unchanging accuracy of radiologists interpreting diagnostic radiographs in the face of substantial image quality changes over many decades, would suggest that the answer is "No." Image quality is a complex topic, and some insights into its role in diagnostic imaging were provided in a recent webinar held by Dr. Ralph Schaetzing, Manager, Strategic Standards & Regulatory Affairs at Carestream Health.

Medical imaging is a key element of modern healthcare, and the expectation of many is that advances in the underlying technologies will lead to improvements in image quality, which, in turn, will produce better diagnostic performance from those interpreting the images. However, the data on radiologist performance with chest radiographs over the last half century, a time period over which image quality has certainly not stood still, tell a different story.

Has image quality become irrelevant? Does diagnostic performance depend more on other factors? Are we perhaps evaluating image quality in the wrong place? If so, what is the right place? Where exactly does image quality reside in the medical imaging chain? Such questions are important in understanding the role that image quality plays in the diagnostic process.

The radiographic imaging chain always consists of five basic functions (the first three being the most critical):

1. **Capture** (the detection/recording of an "image in space" produced by the positioning and exposure of a patient)
2. **Process** (most commonly, the optimization of an image for its diagnostic purpose, but also the preprocessing ("clean-up") of the captured raw image, and the preparation for the display device)
3. **Display** (assuming a human viewer is at the end of the imaging chain)
4. **Store**
5. **Distribute**

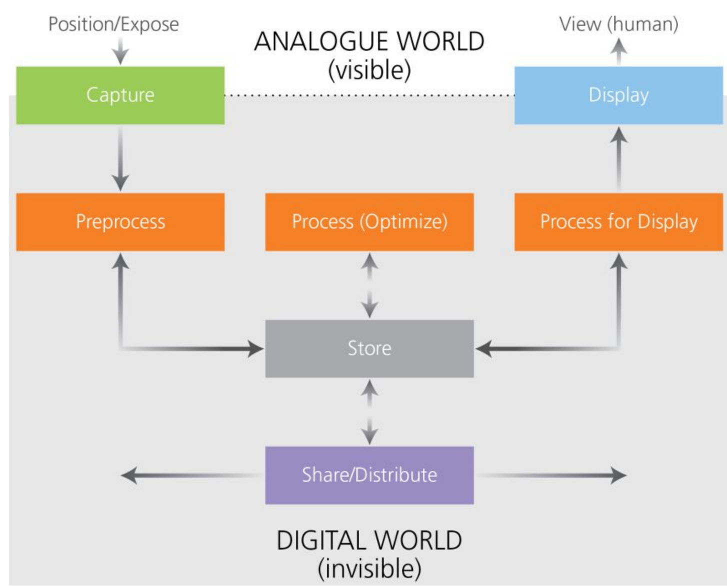

Image quality resides at multiple places along this chain, starting with patient positioning and exposure, through to the eye/brain system of the viewer. A complete description of such a distributed image-quality model actually requires looking at three different "flavors" of image quality:

1. **Objective:** quantitative image properties, measurable without viewer input (e.g., metrics such as contrast, MTF, NPS, DQE[1])

2. **Subjective:** image properties based solely on viewer perception, that is, "in the eye of the beholder"

3. **Performance-based:** image quality based on the ability of the image or system to enable the viewer to complete an interpretation task to some pre-defined level of success

Unfortunately, these three types of image quality, at least at today's quality levels, are only loosely, if at all, correlated. For example, improvements in objective quality have gone unnoticed by viewers. Subjectively preferred images have actually led to *poorer* diagnostic performance. Improved objective image quality may not affect performance at all. These facts make the answer to the title question somewhat more elusive.

Given the many "knobs and buttons" on modern digital-imaging systems (especially the ability to do image processing), it is possible to create from a given input image many different displayed images—a few desirable, most not. So, depending on the parameters selected for the exposure (especially dose) and for image processing, a "good" image detector can produce a "poor" displayed image, and a "poor" image detector can produce a "good" displayed image[2].

Regardless of "good" or "poor," image details not captured on the detector during the exposure step are forever lost. No amount of image processing or sophisticated display technology will bring them back later. This observation provides fuel for the frequent discussion on the trade-offs between detector quality and dose requirements. On the other hand, captured details can also be made more difficult to see (or even to disappear) with improper image processing, and/or a poor display or poor display conditions.

Medical-image processing plays a critical role in the image-quality discussion, and has been an active R&D topic for more than 50 years. A standard tool box has emerged, from which manufacturers can choose basic techniques. Beyond the basics, there is much original, ongoing research focused on automated image analysis and the selection of optimal parameters for each image.

The transition to electronic displays has also driven an increasing emphasis on display calibration (e.g., DICOM Grayscale Standard Display Function, GSDF) and on proper viewing/ambient conditions. Viewing conditions become particularly important for maintaining image quality on electronic displays, since they have a contrast ratio significantly lower than that of film on a viewbox.

The interface between the image display device and the eye/brain system (Human Visual System, HVS) of the observer is a critical junction in the imaging chain. The objective image quality characteristics of the display device are fairly straightforward to measure, while those of the observer are more difficult to characterize, and vary by observer. But, it is exactly the observer's ability to extract the information necessary to make a correct diagnosis from the displayed image that ultimately determines the performance-based image quality of the system.

The (typical) HVS has detection limitations with regard to contrast sensitivity, spatial and temporal resolution, and other factors that must be taken into account when displaying an image (e.g., adaptation, nonlinear luminance response). Every image must land or be made to land in that part of the parameter space where the HVS works best, lest the observer miss relevant clinical details. Among other things, this implies that the imaging system must prepare each image for the output device(s) on which it will be viewed.

However, even a properly prepared image is then processed by the eye/brain system of the viewer, usually a radiologist. There are additional sources of variability in this last, vital link of the imaging chain.

At the outset, it was noted that, at least for chest radiography, there has been little change in diagnostic accuracy since the first such performance measurements were made in the mid-1950s. Have experienced radiologists learned to extract all of the required information contained in radiographs, independent of their image quality? Or have the improvements in imaging systems and image quality since the 1950s not been significant enough to "move the needle" diagnostically?

[1] Modular Transfer Function (spatial resolution); NPS: Noise Power Spectrum (noise); DQE Detective Quantum Efficiency
[2] "Good" and "Poor" must be defined clearly.

Are there perhaps some inherent loss/error mechanisms in the diagnostic process that prevent improvements in objective image quality from leading to improvements in performance? In other words, is the radiologist perhaps the rate-limiting step in the medical imaging chain?

This latter question has been studied extensively over the years. In particular, Kundel[3] provided considerable insight into the sources of reader error in radiography, classifying them into three categories, and quantifying their relative contributions to the errors made:

- Visual Scanning - 15%
- Target Recognition - 37%
- Decision-Making/Interpretation - 48%

Visual scanning is the (systematic) search for potentially clinically relevant targets. Such targets usually generate longer visual "dwell times" during the search, that is, the eyes spend more time looking at the target as they scan across the image (which can be measured using eye-tracking equipment). Target recognition, as its name implies, is the process of accepting a found target as clinically relevant. Finally, decision making classifies the target as relevant to the diagnosis or not.

Objective image quality would be expected to play a role in the first two categories (which generated over half the errors), by ensuring that targets are displayed with sufficient contrast and sharpness to enable a human reader to find and recognize them. Yet the aforementioned performance results seem to indicate that (objective) image quality has not played a significant role.

So, does image quality matter? As a "currency" to trade against dose reduction, the answer is clearly "yes," and this tool is being used actively in modern DR (and CR) systems.[4]

As a method to improve diagnostic performance, the arguments are less convincing. In fact, it appears that the "success" or "failure" of the radiographic imaging chain depends more on the qualities of the viewer at its end than on the qualities of the system used to generate the images to be interpreted. Perhaps further dramatic improvements in image quality (e.g., DQE, image processing) will change this situation in the future, but for now, (objective) image quality should not be relied upon as the final predictor of fitness for purpose.

[3] Kundel H, Visual search, object recognition and reader error in radiology, Proc. SPIE 2004; 5372: 1-11

[4] See other Carestream White Papers on dose reduction- Maximizing Dose Efficiency for Pediatric Patients