

Evaluación de la Calidad de Imagen en Equipos de Radiografía

R. Redondo¹, G. Bueno¹, P. Galán², P.E. Millán³, O. Déniz¹

¹ VISILAB Departamento, Universidad de Castilla La Mancha, Ciudad Real, España, gloria.bueno@uclm.es

² UGC Radiofísica Hospitalaria, Hospital Universitario Regional de Málaga, Málaga. España.

³ Servicio de Física y Protección Radiológica, Hospital Clínico Universitario "Lozano Blesa", Zaragoza, España.

Resumen

Este trabajo evalúa métricas de calidad de imágenes para radiodiagnóstico. Para ello se ha hecho uso de métricas de calidad de imagen tanto matemáticas como perceptuales y han sido comparadas con las evaluaciones psicofísicas arrojadas por expertos sobre las misma base de datos. Se ha observado que algunas métricas como UQI y VIF poseen una alta correlación por encima de 0.82 con los resultados psicofísicos. Una interesante característica de estas métricas es que son generales y por tanto no han necesitado ningún tipo de ajuste previo y podrían ser robustas por ejemplo a fallos de registro.

1. Motivación

Durante la última década, la tecnología de radiografía ha evolucionado de los equipos comunes de pantalla de película de luminiscencia fotoestimulada (radiografía computerizada), dispositivos de carga acoplada, la producción de fotoconductores o recientemente detectores de pantalla plana [1]. En una manera corta y simple estos sistemas irradian a través del paciente rayos de partículas a una superficie sensible capaz de retener las intensidades de los rayos y así permitir después la extracción de la información.

La radiografía computerizada (CR) ha ido continuamente mejorando. Aquí dos grandes avances se han introducido: la lectura de los fósforos por ambos lados o a través de la tecnología de la línea de barrido [5]. Estas dos modalidades de lectura permiten una eficiencia de detección mejorada y una lectura rápida de toda la imagen. Las unidades más avanzadas basadas en fósforos columnares ofrecen resoluciones espaciales no inferiores a los obtenidos con los sistemas de radiografía computerizada convencionales a pesar de sus fósforos más gruesos [1]. Un mayor espesor del fósforo columnar permite una eficiencia mejorada y por lo tanto mejores propiedades frente al ruido. Los sistemas de radiografía computerizada estándar presentan un notable espectro normalizado de potencia de ruido en todas las exposiciones investigadas. El objetivo del trabajo en [8] es lograr una comparación directa entre los 5 sistemas de radiografía computerizada basados en 3 tecnologías diferentes (lectura por un solo lado, lectura de doble cara y escaneado por líneas). Algunos de los sistemas investigados en este trabajo ya fueron analizados por otros estudios en los últimos años. Sin embargo, ninguna comparación directa objetiva se ha hecho con las mismas condiciones de exposición.

Una compromiso común compartido por estas técnicas es el equilibrio entre la radiación vs. diagnóstico, es decir, minimizar la dosis de irradiación sin deteriorar el contraste de la imagen y por lo tanto las tasas de detección de la patología. Sin embargo, la calibración de estos sofisticados equipos requiere docenas de parámetros y en concreto algunos de ellos requieren revisiones periódicas, incluso alguna podría variar en función de las condiciones de visualización. Esto conduce a la necesidad de algoritmos de (semi) control automático de exposición para facilitar la calibración del equipo.

Algunos de los parámetros de calibración dependerán de factores de fabricación que no pueden ser ajustados más tarde, al igual que algunos factores geométricos, las propiedades físicas de los materiales o el ruido intrínseco. Por lo tanto, la calidad se suele medir como una proporción de la señal (intensidad del haz) contra el ruido, donde la composición del ruido depende de varios factores [4]: Detective Quantum Efficiency, un índice de las partículas incidentes eficaces que depende de las propiedades geométricas y físicas del sensor fosforescente (absorción/dispersión), Función de la Modulación de Transformación (MTF en inglés) que depende del material sensor fotofosforescente y el sistema de lectura global (rayos láser, detectores, electrónica ...), Wiener Noise Spectrum debido al proceso óptico (absorción/dispersión) y electrónica (amplificadores, filtros y convertidores A/D).

Sin embargo, otros parámetros, como el enfoque del detector de distancia, campo de irradiación, la dosis o alineaciones deben ser revisados y calibrados periódicamente sólo debido a la degradación y/o a la inestabilidad presente en tales sistemas de radiografía. En cualquier caso, todos los parámetros de calibración afectarán finalmente, en mayor o menor medida a la calidad de las imágenes digitales capturadas.

En este estudio vamos a llevar a cabo un análisis de los parámetros de calibración, que pueden afectar de una manera medible a la calidad de imagen de equipos de radiografía. Para ello vamos a seguir el protocolo español para el control de calidad de radiodiagnóstico [3]. Además también vamos a evaluar los algoritmos basados en métricas de calidad de imagen para facilitar el control automático de la exposición.

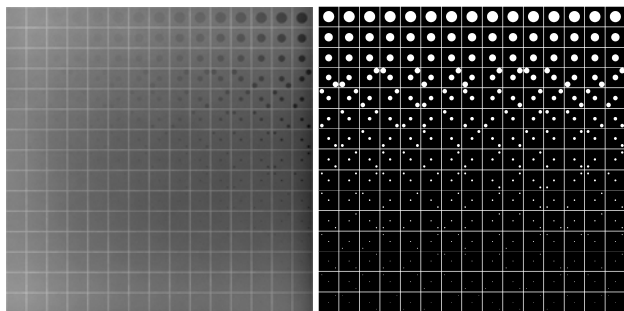


Figura 1. Ejemplo CDRAD registrado e imagen de referencia.

2. Material

Los maniqués CDRAD como el ejemplo de la Figura 1 están hechos de un acrílico de 8 mm de espesor en el que agujeros circulares son perforados en una región cuadrada de unos $15 \times 15 \text{ mm}^2$. El tamaño y la profundidad de los orificios varían logarítmicamente a lo largo de 15 filas y 15 columnas. La pequeña atenuación producida por los objetos implica una relación lineal entre la profundidad y contrastes de modo que tanto la profundidad y el contraste se reducen a la mitad de su valor inicial cada tres regiones cuadradas. Para el presente trabajo una base de datos de 70 imágenes de CDRADs ha sido adquirida con diferentes intensidades de 1-20 mGy. Se utilizaron tres sistemas que utilizan la misma tecnología y configurados para trabajar de manera similar en las mismas condiciones como x-ray Philips Diagnost Digital. Tienen dos detectores, llamados Bucky Mesa y Bucky Pared. La atenuación de los rayos X y dispersión producida en un paciente estándar es simulada por bloques de PMMA de $30 \times 30 \text{ cm}^2$ y 2cm de espesor situados sobre el detector. A menudo los CDRADs son adquiridos rotados y/o desplazados, para ello se ha realizado una etapa de registro de las imágenes que consiste en rotación y posteriormente recorte de la imagen CDRAD.

3. Métodos

Las métricas de calidad de imagen tienen por objetivo capturar la fidelidad entre una imagen distorsionada y una imagen de referencia mejorada. A veces la imagen de referencia no está disponible y la calidad tiene que ser medida en términos de criterios estándar como medida de contraste, contorno, preservación de estructuras o relación señal-ruido (si está disponible también). Estos sistemas utilizan para adquirir imágenes de un maniquí comercial de contraste-detalle (CDRAD, Nuclear Associates [12]) que sirve perfectamente como imagen de referencia. Las métricas con imagen de referencia se pueden agrupar en puramente matemáticas como Error Medio de la Señal (MSE en inglés), la Relación de Pico de Señal a Ruido (PSNR en inglés), L-norma, correlación cruzada (Cross-Correlation), coeficiente de Pearson, que miden diferencias píxel por píxel.

Otras métricas miden las diferencias en términos más complejos de visión por computador como luminancia, contraste, contorno o texturas. Este último grupo se puede agrupar en otras dos categorías: perceptuales y no perceptuales. Las últimas hacen uso de uno o varios términos de visión por computador Mean Structural

Similarity Measure (MSSIM) [13], Information Fidelity Criterion (IFC) [10], Visual Information Fidelity (VIF) [9] or Singular Value Decomposition (SVD) [11]. Las métricas perceptuales objetivas incorporan comportamientos más complejos (no lineales) realizados por el sistema visual humano, tales como la descomposición espacio-frecuencia, la adaptación de luminancia (ley de Weber) o el contraste de enmascaramiento. Entre estas últimas métricas está Visual Differences Predictor (VDP) [2]. Un resumen puede consultarse en [6].

Más allá de las métricas objetivas, ninguna de ellas es capaz de modelar tareas superiores y abstractas de percepción del sistema visual humano como inferencias contextuales, entrenamiento de reconocimiento o áreas de atención. Por lo tanto ninguno de los indicadores citados puede ser adoptado universalmente y tiene que ser contrastado contra métricas subjetivas o psicofísicas. Dos medidas se pueden extraer a partir de experimentos psicofísicos: una derivada de la relación de detectabilidad de los agujeros en el CDRAD y las correspondientes curvas de detección de contraste apuntando el menor de los diámetros observados para cada contraste que se resume en las curvas de Mean Opinion Score (MOS) reportadas por los observadores [7].

Para llevar a cabo los experimentos psicofísicos una interfaz gráfica de usuario ha sido desarrollada con el lenguaje de programación (TCL), siguiendo el procedimiento de estímulo individual de evaluación continua de la calidad (SS-CQE) en la recomendación UIT-R BT.500-11 [7]. Así un total de 10 observadores cualificados han evaluado la calidad de las imágenes CDRAD entre radiofísicos y expertos en procesamiento de imágenes.

4. Resultados

En las gráficas de la figura 2 se muestran los dos mejores resultados de las métricas comparadas con los resultados puntuados por los observadores. Tales distribuciones han sido aproximadas por regresión por un polinomio de tercer grado. Los resultados de correlación y error cometido son mostrados en la Tabla 1. La métrica VIF presenta un alto coeficiente mayor de 0.9 claramente por encima de las demás. También la métrica UQI presenta un coeficiente de correlación alto por encima de 0.8.

METRIC	C.C. Spearman	RMSE
VDP	0.71	0.43
IFC	0.80	0.36
MSSIM	0.35	0.66
SVD	-0.08	0.68
UQI	0.82	0.36
VIF	0.91	0.23
RMSE	-0.43	0.64
Cross Correlation	0.05	0.69
L2-norm	-0.43	0.64
Pearson Coeff.	0.58	0.56

Tabla 1. Coeficientes de correlación de Spearman y error cuadrático (RMSE) de las métricas de calidad comparadas con los resultados psicofísicos (MOS).

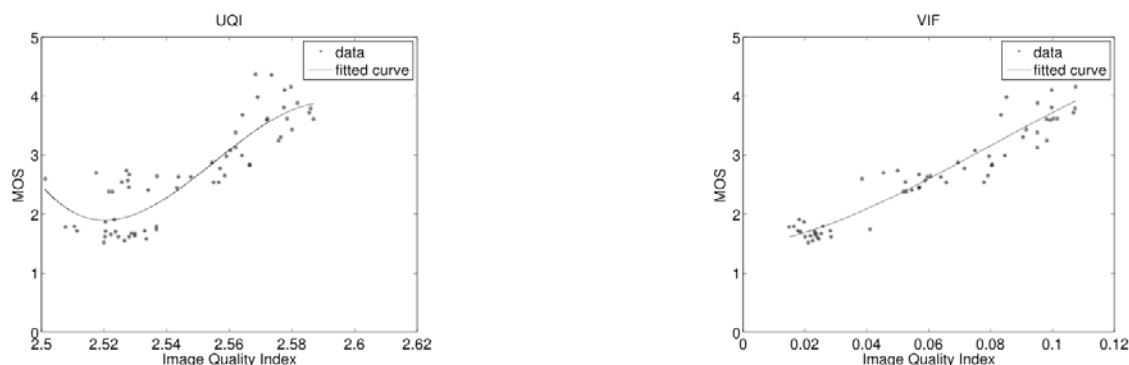


Figura 2. Distribuciones de puntuaciones arrojadas por las métricas de calidad de imagen (IQI) comparadas con la evaluación psicofísica de expertos (MOS). Las distribuciones aproximadas por un polinomio de tercer grado.

5. Conclusiones

Un gran esfuerzo se ha hecho durante la última década para aumentar las técnicas de imagen en radiología computerizada. Las técnicas persiguen incrementar relación señal-ruido en términos de resolución espacial del sensor, sensibilidad de lectura o ruido cuántico y electrónico. Cualquier mejora de estos factores dará lugar a una disminución de riesgo de sobredosis para el paciente. En este respeto el modelado cuántico de los sensores proporciona una manera eficaz de controlar la dosis de exposición, sin embargo, en el último paso se requiere mostrar las imágenes en monitores y ser evaluadas visualmente por los técnicos. Por lo tanto, indicadores de calidad de imagen pueden apoyar modelos cuánticos para estimar el reconocimiento de patologías del radiólogo y a los radiofísicos automatizar la calibración de los equipos para ajustar mejor la dosis de irradiación.

Aquí se han evaluado 10 métricas objetivas entre matemáticas y perceptuales comparadas con la evaluación subjetiva de expertos mediante maniqués de calibración CDRAD. Los resultados mostraron al menos 3 métricas UQI, IFC y VIF con una alta correlación. Entre ellas la última por encima de 0.9 lo que la convierte en una buena candidata como apoyo para la evaluación de dichas imágenes.

Una ventaja de estas métricas es que no requieren de un conocimiento a priori de la imagen de CDRAD. Son por tanto generales y podrían agilizar la calibración de equipos mediante dichos maniqués. Al no necesitar de un conocimiento a priori del maniquí además permite que puedan ser robustas ante fallos de registro.

En futuras investigaciones se extenderá dicho estudio a otras modalidades de radiografía y otros maniqués como CDMAM y ETR1.

Agradecimientos

Los autores agradecen financiación del Consejo de Seguridad Nuclear y de la Unión Europea con el proyecto financiado en el VII FP con el proyecto SP CE241851.

Referencias

[1] Cowen, A., Davies, A., and Kengyelics, S. Advances in computed radiography systems and their physical imaging

characteristics. Clinical radiology 62, 12 (2007), 1132–1141.

- [2] Daly, S. J. Visible differences predictor: an algorithm for the assessment of image fidelity. In SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology, International Society for Optics and Photonics (1992), 2–15.
- [3] De Radiología, S. E. Protocolo español de control de calidad en radiodiagnóstico: Revisión 1 (aspectos técnicos). Madrid: Edicomplet 164 (2002).
- [4] Dobbins III, J. T. Metrics for measuring image quality, 2000.
- [5] Fetterly, K. A., and Schueler, B. A. Performance evaluation of a computed radiography imaging device using a typical front side and novel dual side readout storage phosphors. Medical physics 33 (2006), 290.
- [6] Lin, W., and Jay Kuo, C.-C. Perceptual visual quality metrics: A survey. Journal of Visual Communication and Image Representation 22, 4 (2011), 297–312.
- [7] Recommendation, I. 500-11, methodology for the subjective assessment of the quality of television pictures, recommendation ITU-R BT. 500-11. ITU Telecom. Standardization Sector of ITU (2002).
- [8] Rivetti, S., Lanconelli, N., Bertolini, M., e. Comparison of different computed radiography systems: Physical characterization and contrast detail analysis. Medical physics 37 (2010), 440.
- [9] Sheikh, H. R., and Bovik, A.C. Image information and visual quality. Image Processing, IEEE Trans. on 15, 2 (2006), 430–444.
- [10] Sheikh, H. R., Bovik, A. C., and De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. Image Processing, IEEE Transactions on 14, 12 (2005), 2117–2128.
- [11] Shnayderman, A., Gusev, A., and Eskicioglu, A. M. An SVD-based grayscale image quality measure for local and global assessment. Image Processing, IEEE Transactions on 15, 2 (2006), 422–429.
- [12] Thijssen, M., Bijkerk, K., and van der Burght, R. Manual contrast-detail phantom cdrad type 2.0: project quality assurance in radiology. St Radboud, Netherlands: Department of Radiology, Univ. Hospital Nijmegen (1998).
- [13] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. Image Processing, IEEE Trans. on 13, 4 (2004), 600–612.