

Evaluación Subjetiva/Objetiva de Imágenes Comprimidas con Pérdida

*Lic. Hugo D. Ramón¹
Msc. Claudia C. Russo²
Ing. Armando De Giusti³*

LIDI Laboratorio de Investigación y Desarrollo en Informática ⁴
Facultad de Informática
Universidad Nacional de La Plata

1. Introducción

La compresión de una imagen digital puede facilitar su procesamiento, almacenamiento y transmisión. A medida que los Departamentos de las grandes Organizaciones se vuelven cada vez más digitales y distribuidos, la cantidad de datos multimediales (en particular imágenes) que tienen, obligan a considerar su compresión para su almacenamiento y transmisión.

El objetivo general de la compresión es representar una imagen con la menor cantidad posible de bits con la menor pérdida de calidad posible, acelerando así la transmisión y minimizando los requerimientos de almacenamiento. De manera alternativa, el objetivo es lograr la mejor fidelidad posible para una capacidad limitada disponible de almacenamiento.

En el campo de las comunicaciones visuales se ha destinado mucho trabajo a la codificación de imágenes digitales, dirigido a reducir los requerimientos de la tasa de bits para la transmisión de las imágenes. La experiencia ha mostrado que cada esquema de codificación está sujeto a su propio y único conjunto de causas de pérdida que a menudo son difíciles de caracterizar. Esto se debe a la forma en la que se diseñan los esquemas de codificación para variar selectivamente la precisión de la representación; el observador puede ser insensible a los errores en algunas partes de la imagen pero no en otras. Una compresión considerable se puede lograr sólo mediante algoritmos con pérdida, los que no permiten la recuperación exacta de la imagen original. Esta pérdida de información hace que la compresión y otros algoritmos de procesamiento de imágenes con pérdida tengan problemas de confiabilidad debido a la potencial pérdida de calidad.

Hay una necesidad de mediciones precisas de las pérdidas subjetivas que puedan usarse para predecir la calidad de una imagen. El objetivo de este trabajo es determinar estas mediciones de la distorsión y probar que se pueden usar diferentes medidas objetivas para predecir la calidad subjetiva de una imagen.

Se describirá enfoques para medir la calidad de las imágenes, como ser, razón señal-ruido (SNR), evaluación subjetiva, mediciones del error crudo, entropía, etc.

La necesidad de este tipo de mediciones está particularmente reconocida en el área de la codificación de imágenes digitales. La existencia de mediciones de la distorsión subjetivamente relevantes que reflejen las evaluaciones de la calidad de una imagen que hace un espectador, harían considerablemente más fácil la tarea de diseñar y optimizar los esquemas de codificación.

Estas distorsiones se comparan y contrastan con un conjunto de imágenes representativas de varios dominios de aplicación y se examina cuán buenas son las mediciones de distorsiones que se pueden obtener fácilmente para predecir las evaluaciones subjetivas, que son más caras en tiempo. Los ejemplos son de imágenes tradicionales que se utilizan en el área de procesamiento y compresión de imágenes comprimidas utilizando JPEG estándar [Wallace 1991].

¹ Profesor Adjunto con Dedicación Exclusiva, hramon@lidi.info.unlp.edu.ar

² Profesor Adjunto con Dedicación Exclusiva, crusso@lidi.info.unlp.edu.ar

³ Director LIDI. Profesor Titular con Dedicación Exclusiva. Investigador Principal del CONICET, degiusti@lidi.info.unlp.edu.ar

⁴ calle 50 y 115, La Plata 1900 tel-fax +54 221 22 77 07 <http://lidi.info.unlp.edu.ar>

2. Criterios de fidelidad

La calidad en sí es un atributo con muchas definiciones e interpretaciones posibles, dependiendo del uso que se le dará a las imágenes. Una buena imagen procesada puede ser una que es perceptualmente agradable o útil en una aplicación específica.

2.1. Medidas de la Distorsión

Algo natural para hacer si esta interesado en la fidelidad de una secuencia reconstruida es observar en las diferencias entre los valores originales y los reconstruidos; en otras palabras la distorsión introducida en el proceso de compresión. Dos medidas conocidas de distorsión o diferencia entre las secuencias original y reconstruida es el error cuadrático medio y la de la diferencia absoluta, las cuales son llamadas *medidas de distorsión diferencia*.

Si $\{x_n\}$ es la salida de la fuente y $\{y_n\}$ es la secuencia reconstruida entonces la medida del error cuadrado es la siguiente:

$$d(x, y) = (x - y)^2$$

mientras que la medida de la diferencia absoluta esta dada por:

$$d(x, y) = x - y$$

En general es dificultoso examinar la diferencia en una base termino a termino. Es por eso que un conjunto de medidas del promedio es usado para resumir la información en la secuencia de diferencias.

La medida del promedio mas comúnmente utilizada es la del promedio de los errores cuadrados llamado *error cuadrático medio* (mean squared error)

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N (x_n - y_n)^2$$

Si esta interesado en el tamaño del error relativo a la señal se puede hallar el radio del valor cuadrado promedio de la salida da la fuente y el *mse*, lo que se llama *signal-to-noise* (SNR):

$$SNR(dB) = \frac{\sigma_x^2}{\sigma_d^2}$$

donde es el σ_x^2 cuadrado promedio de la salida de la fuente o señal y σ_d^2 es el *e*. El *SNR* es frecuentemente medido en una escala logarítmica y las unidades de medidas son los decibeles (dB):

$$SNR(dB) = 10 \log_{10} \left(\frac{x^2 \text{ peak}}{\sigma_d^2} \right)$$

Otras veces se tiene interés en el tamaño de error relativo al valor máximo de la señal. Este radio es llamado *peak signal to noise ratio* (PSNR) y esta dado por

$$SNR(dB) = 10 \log_{10} \left(\frac{\sigma_x^2}{\sigma_d^2} \right)$$

Otra medida de distorsión diferencia que es muy usada aunque no tanto como el *mse* es el promedio de la diferencia absoluta:

$$\sum_i^n |x_i - y_i|$$

También suele utilizarse la raíz cuadrada del *mse* si se quiere saber en cuanto difieren en promedio los valores originales de los reconstruidos.

2.2. Calidad Subjetiva

La calidad subjetiva de una imagen reconstruida puede verificarse de varias formas. Un conjunto de imágenes aleatorias puede presentarse a los expertos o usuarios típicos quienes las califican, en una escala de 1 a 5. Siguiendo, se puede realizar un análisis estadístico para sobresaltar promedios, variaciones y otras tendencias en los datos.

Son comunes los tests formales subjetivos en el área de procesamiento del habla y comprensión de audio. Uno de ellos es el Mean Opinion Score (MOS) y el Diagnostic Acceptability Measure (DAM) [Quackenbush 1988].

Estos sistemas de clasificación son comunes en las investigaciones de speech y se están desarrollando algunos esfuerzos para clasificación de videoentrenamientos, pero no existen estandarizaciones para imágenes [Fish 1991].

Existen considerables variaciones en cuanto al rango numérico de las respuestas, en relación de cuando proveer o no de frases descriptivas para cada número y en el intento de que el rango subjetivo, indique la utilidad de la imagen en una aplicación específica. En general se utiliza una clasificación en donde el usuario clasifica de acuerdo al grado de similitud o no similitud entre dos imágenes, una original y otra la procesada, en este caso comprimida. Este último concepto es el que se toma en este trabajo.

Un atributo muy utilizado de una medida objetiva como signal-to-noise (SNR) es la habilidad que tienen para predecir la calidad subjetiva. Esto es importante en algunas áreas como las médicas y videoentrenamiento en donde la calidad subjetiva es de suprema importancia [Cosman 1994].

Existen varios métodos que consideran para cuantificar el grado de correlación entre dos medidas dadas, o la habilidad de dado una predecir otra medida para audio, imágenes y video [Marmolin 1986] [Wang 1992].

3. Trabajos Futuros

Seleccionar el dominio de las imágenes a estudiar para verificar las tendencias supuestas. Las áreas pueden ser de imágenes relacionadas a la medicina, a la astronomía, paisajes, etc.

Definir los modelos estadísticos teniendo uno lineal, como por ejemplo, análisis de varianza mediante un modelo lineal [Montgomery 1991] y uno no lineal que tenga en cuenta las restricciones y características del modelo visual humano, como por ejemplo Multinomial Logistic Regression Model [Agresti 1990].

4. Bibliografía

[Quackenbush 1988] Objective Measures of Speech Quality. S. Quackenbush, T. Barnwell & M. Clements. Prentice Hall Signal Processing Series, 1988.

[Wallace 1991] *The JPEG Still Picture Compression Standard*. G. Wallace. Multimedia Engineering. Digital Equipment Corporation. Maynard, MASS, 1991.

[Fish 1991] A Subjective Visual Quality Comparison of NTSC, VHS and Compressed DS1-compatible Video. R. Fish, T. Judd. Proc. SID (Society for Information Display), vol. 32 nro.2, 1991.

[Wang 1992] An Objective Measure for predicting Subjective Quality of Speech Coders. S. Wang, A. Sekey & A. Gersho. IEEE Journal Sel. Areas Commun. Vol. 10, Junio 1992.

[Santer 1989] *The Statical Analysis of Discrete Data*. T. J. Santer & D. E. Duffy. Springer Verlag, 1989.

[Montgomery 1991] *Diseño y Análisis de Experimentos*. Montgomery. Grupo Editorial Iberoamérica, 1991.