

Traslación de medidas de distorsión a espacios de representación multi-resolución en análisis de imágenes

B. Ortiz-Jaramillo^{§*}, J. García-Álvarez^{*}, G. Castellanos-Domínguez^{*}

**Universidad Nacional de Colombia Sede Manizales.*

Grupo de Control y Procesamiento Digital de Señales

§ e-mail: bortizj@unal.edu.co

(Recibido: Octubre 19 de 2009 –Aceptado: Junio 10 de 2010)

Resumen

Se presenta un método para el cálculo del Índice de Calidad (Quality Index - QI) en imágenes, que consiste en la combinación lineal de tres medidas de distorsión perceptuales, basadas en el Sistema Visual Humano y las cuales son proyectadas a un espacio multi-resolución utilizando operadores en el espacio wavelet; espacio en el cual se realiza la extracción de características, tales como homogeneidades, contornos y bordes. En la validación del método, el QI calculado se contrasta con un puntaje de observación media diferencial, que es calculado experimentalmente sobre un conjunto estándar de imágenes. Los resultados muestran que las técnicas multi-resolución permiten una descripción más adecuada de las métricas perceptuales, además, reduce el tiempo de cómputo, en la medida que las mediciones son prioritarias sobre los detalles. Así mismo, los procesos de medida se calculan sobre el mismo espacio de representación, evitando transformaciones adicionales para extraer características. La estimación propuesta del QI se prueba en una aplicación de detección de rostros, con el fin de determinar las condiciones mínimas de distorsión para efectuar la clasificación de las respectivas imágenes.

Palabras Claves: Árboles cuaternarios, Error de borde, Error de bloque, Imparidad visual, Métrica perceptual

ELECTRONIC ENGINEERING

Translating distortion measures to an multi-resolution spaces for image analysis

Abstract

A method for calculating a Quality Index (Quality Index - QI) in images, which is the linear combination of three measures of perceptual distortion based on Human Visual System is presented. These three measures are projected to a multi - resolution space with operators in wavelet space, in which extraction of features such as homogeneity, contours and edges is carried out. To validate the method, the calculated QI is contrasted with an arithmetical mean observation score, which is determined experimentally on a standard set of images. The results show that the multi - resolution techniques allow a more appropriate description of the perceptual metric, reducing the computation time, as long as measurements are a priority over the details. Likewise, measurement processes are calculated on the same representation space, avoiding additional processing to extract features. This quality index is used in a face detection application to define the minimum distortion condition to classify images.

Keywords: Quadtree, Edge error, Blockiness, Visual impairment, Perceptual metric.

1. Introducción

La compresión de imágenes, que consiste en minimizar la cantidad de datos almacenados, así como la organización de los mismos para su posterior transmisión, tiene alta aplicación en áreas como la tele-medicina, que requiere de esquemas de compresión con alta calidad en la imagen, con requerimientos de tasas de compresión elevadas, y cuyo desarrollo implica el uso de técnicas para medir la calidad de la imagen, como aspecto crítico durante el diagnóstico asistido Ginesu et al. (2006). Luego, cuando el ancho de banda requerido exige tasas de compresión elevadas, las medidas de distorsión son necesarias para definir el nivel de calidad de la imagen y como parámetro de optimización para la tasa de compresión. De acuerdo con la naturaleza de la medición, las medidas de distorsión se dividen en objetivas y perceptuales: Avcibas et al. (2003), Sheikh et al. (2006), García & Castellanos (2008). Como un ejemplo de las primeras, el Error Cuadrático Medio (*Mean Square Error* - MSE) es una medida usada ampliamente debido a la sencillez de su cálculo. Sin embargo, tal medida no analiza aspectos básicos de las imágenes tales como degradación de bordes, cambios de textura u homogeneización de áreas, que en cambio son tenidas en cuenta por las medidas perceptuales. Tales medidas, que hacen posible clasificar y cuantificar la naturaleza de la distorsión, incluyen diferentes etapas: a) el ajuste subjetivo empleando diferentes análisis psico-perceptuales, b) modelos del sistema visual humano (*Human Visual System* - HVS) que aproximen el estímulo visual y c) experimentos realizados sobre un banco de imágenes distorsionadas con la participación de poblaciones amplias de personas Ibarra et al. (2007), Bovik (2009), tal como se establece como estándar en las recomendaciones de la ITU ITU-R (2002).

Sin embargo, la ventaja que presentan las medidas perceptuales de distorsión contrasta con su alto coste computacional de implementación, por cuanto su cálculo requiere que la imagen se transforme de forma adecuada para la extracción de sus características relevantes, a diferencia de las medidas objetivas, e.g., Czekanowski y el MSE, donde el análisis se realiza directamente sobre la representación inicial. En este sentido, se

han propuesto métodos que relacionan una medida objetiva con una perceptual, e.g., mediante coeficientes de regresión lineal Sheikh et al. (2006). Sin embargo, su efectividad está sujeta a las especificidades de la base de representación, a partir de la cual se estima la relación, y el grado del polinomio de regresión, por lo que alguna distorsión de naturaleza distinta no podría ser identificada. En otros casos, se emplean Redes Neuronales como técnica de regresión y ajuste; aunque la selección de la arquitectura adecuada de la red sigue siendo un problema abierto Chen et al. (2009).

En este artículo, se presenta un método de representación que estime el QI de las imágenes médicas, basado en la regresión lineal de tres métricas de distorsión perceptuales completamente referenciadas. El método incluye la simplicidad de la regresión convencional, debido a que utiliza los datos de un único espacio de representación, a pesar de requerir el paso de transformación multi-resolución. Además, presenta mayor flexibilidad por cuanto adquiere la naturaleza funcional de la distorsión, representada por las métricas. Así mismo, las técnicas multi-resolución, en particular, incluyen la descomposición en árboles cuaternarios (*Quadtree Decomposition* - QD) (partición basada en regiones de interés), detectores de bordes basados en wavelet (mapas de bordes más acordes con lo percibido por el HVS), y máscaras de contraste a la sensibilidad sobre los coeficientes wavelet que se ajusten mejor a los esquemas de compresión óptimos.

El artículo se organiza como sigue: en la Sección 2 se presenta la representación básica para el mapeo de la función de sensibilidad al contraste, la selección de los bloques para mapear la degradación de bloques y el deterioro visual al espacio multi-resolución, así como una breve explicación de los detectores de bordes que usan wavelets. En la Sección 3 se presentan los resultados obtenidos de las métricas a evaluar, para luego en la misma Sección presentar una descripción de aplicaciones que se basan en el método puesto a consideración.

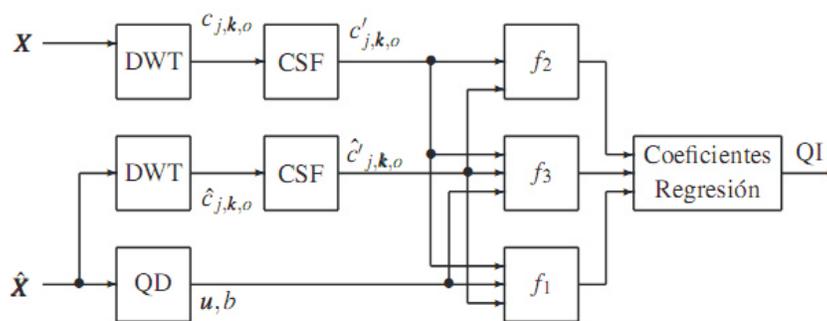


Figura 1. Método para el cálculo del Índice de Calidad - QI.

2. Metodología

En esta sección se presentan las métricas propuestas en Ginesu et al. (2006), así como algunas técnicas García et al. (2009), que permiten su traslado al espacio multi-resolución.

En la Figura 1 se ilustra el método utilizado, el cual en primera medida, consiste en una única transformación de la imagen $X_{m \times n} \in \mathbb{R}^2$ al espacio wavelet. Sin embargo, la QD se realiza sobre la propia imagen. Por lo tanto, se obtienen dos conjuntos de datos:

- Los coeficientes wavelet de la imagen, identificados cada uno por $c_{j,k,o} \in \mathbb{R}^1$ en la escala j posición K y orientación o .

- Los bloques obtenidos de la estructura de cuadros diádicos, generados por el proceso de la QD, e identificados cada uno por $B_{u,b} \subseteq X$ en la posición u de $X_{m \times n}$ de dimensión $b * b$, respectivamente.

Sea $s(w)$ la respuesta en frecuencia de la Función de Sensibilidad al Contraste (*Contrast Sensitivity Function* - CSF) definida como la transformación de cada elemento de la imagen del espacio inicial $X_{m \times n}$, al espacio $X_{m \times n}^{HVS} : X_{m \times n} \rightarrow X_{m \times n}^{HVS}$, $x \rightarrow S, \{x, S\} \in \mathbb{R}$ Gaddipati et al. (1997):

$$S(\omega) = \left(\frac{2}{3} \exp\left(\frac{-\sigma^2 \omega^2}{2}\right) - \exp(-2\sigma^2 \omega^2) \right) \frac{1 + \exp(\beta(\omega - \omega_0)) \cos^2(2\theta)}{1 + \exp(\beta(\omega - \omega_0))} \quad (1)$$

donde w tiene el sentido de frecuencia normalizada, y σ , β y ω_0 son las constantes seleccionadas para cumplir con la condición de observabilidad.

El procedimiento, a partir de los dos conjuntos de datos y la CSF, es el siguiente:

- Calcular los valores resultantes de pasar los coeficientes wavelets por el operador CSF.
- Calcular las medidas de distorsión consideradas (Degradación de bloques, Error de bordes, Deterioro Visual y sus respectivas traslaciones al espacio wavelet).
- Calcular el Índice Global

2.1 Sistema Visual Humano

La transformación (1) actúa como un filtro pasa-banda, debido a que la sensibilidad del HVS tiene justo un máximo de 3-4.5 cicl/grad, decreciendo rápidamente a frecuencias superiores e inferiores. La CSF se traslada al espacio wavelet como se muestra en la Figura 2, Moumquine et al. (2006). Los coeficientes resultantes de la operación son ahora la representación de la imagen convertida al formato HVS, obtenido mediante el uso de pesos sobre los coeficientes wavelet en cada sub-banda.

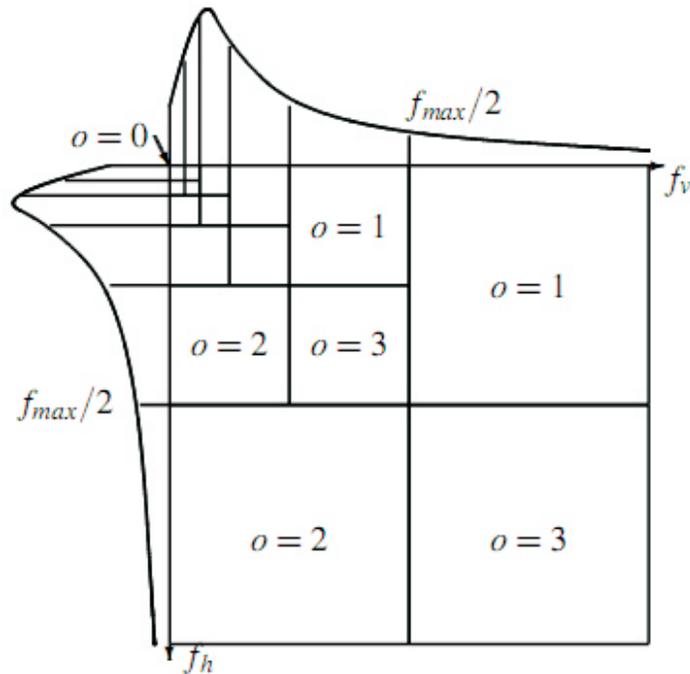


Figura 2. Relación entre CSF y la respectiva descomposición wavelet.

Uno de los métodos propuestos para obtener tales pesos es el de máscara de n -pesos (Beegan (2001) $n = \{6, 11\}$). En este caso, los atributos de baja frecuencia pertenecen a los coeficientes de aproximación wavelet $A_{2^{-j}m \times 2^{-j}n} \in \mathbb{R}^2$, mientras los atributos de alta frecuencia corresponden a los coeficientes de detalle $D_{o, 2^{-j}m \times 2^{-j}n} \in \mathbb{R}^2$, con lo cual, la CSF se puede utilizar en el espacio wavelet, tan solo cambiando los pesos de los coeficientes de cada banda de frecuencia de acuerdo con su importancia perceptual. Concretamente, se obtienen los sub-espacios de detalle W , y de aproximación V , generados al descomponer la CSF en coeficientes wavelet Beegan, (2001). Se define $\{\rho_1, \dots, \rho_5, \gamma_1, \dots, \gamma_5\}$ con ρ_j y $\gamma_j \in \mathbb{R}^1$ como la máscara wavelet, (Contrast Wavelet Mask - CWM), que corresponde a la CSF, donde r_j es el valor pico de $w_j \in W$ y γ_j es el valor pico de $v_j \in V$; aclarando que todos los valores de la CWM deben ser normalizados, tal que $\min\{\rho_1, \dots, \rho_5, \gamma_1, \dots, \gamma_5\} = 1$. Así los coeficientes $D_{o, 2^{-j}m \times 2^{-j}n}$ son multiplicados por su respectivo valor $\sqrt{\gamma_j \rho_j}$ y los coeficientes $A_{2^{-j}m \times 2^{-j}n}$ son multiplicados por los valores γ_1 .

2.2 Degradación en bloques (Blockiness)

La degradación por bloques se evidencia por la reducción de la varianza dentro de algunos bloques $B_{b \times b} \in \mathbb{R}^2$ debido a que desaparecen aquellos elementos que se alejan de la media local μ_B . Otra consecuencia de la degradación es la fuerte cuantización de los bloques vecinos, que producen discontinuidades o transiciones entre los mismos. La métrica propuesta en Ginesu et al. (2006) calcula la degradación en bloques de tamaño regular, analizando la información de la covarianza en los bloques adyacentes $\sigma_{B, B'}$ y la varianza local de cada bloque σ_B y $\sigma_{B'}$, sobre cada transición de bloque. Usando las direcciones únicas definidas por las transiciones de bloque $o=1, 2$, i.e horizontal y vertical, y al definir $e = X - \hat{X}$ como la diferencia pixel por pixel entre la imagen original y la codificada, se puede detectar cuando los errores en dirección ortogonal a la transición provienen del proceso de codificación o son detalles de la imagen original. Entonces, una aproximación a la búsqueda de los bloques en el espacio wavelet, se puede llevar a cabo usando una QD, García et al. (2008). Sean $B_{u_0, b}$ y $B'_{u_0, b}$ el bloque y su

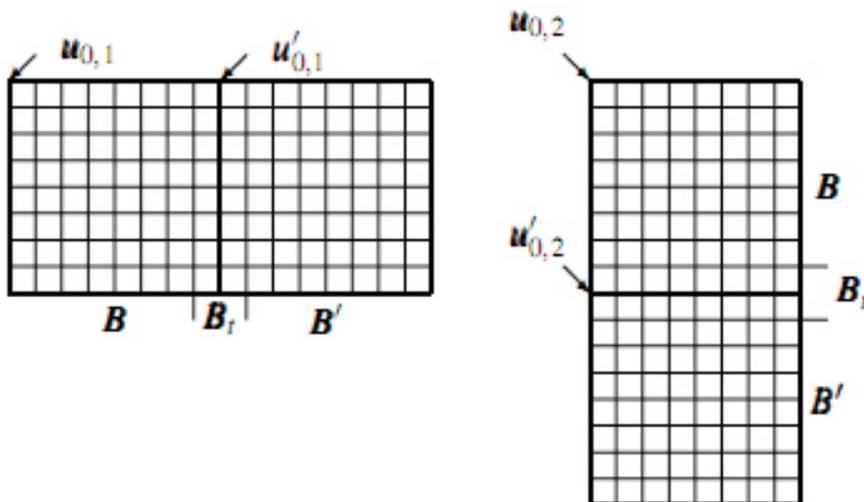


Figura 3. Transiciones a través de los bloques horizontales y verticales.

vecino correspondiente ubicados en la posición \mathbf{u}_o , $\mathbf{u}'_o \in \Omega$, i.e. para una orientación o , $\mathbf{u}_{o,o} = \mathbf{u}'_o - \mathbf{b} \in \mathbb{Z}^2$ y de tamaño $\mathbf{b} = (b_1, b_2) \in \mathbb{N}^2$, generados por la QD con varianza local σ_b en el bloque. Así mismo, sea $\sigma_{t,o}$ la desviación estándar de la transición del bloque $\mathbf{B}_t(o)$ con $o=1,2$ como se muestra en la Figura 3.

Por otro lado, las condiciones de visibilidad de los bloques son: $\sigma_{t,o} > \sigma_B$ y $\sigma_{t,o} > \sigma_{B'}$. Los errores de las transiciones a través de las direcciones ortogonales de los bloques son g_{\perp} y $h_{\perp} \in \mathbb{R}^1$, para las transiciones horizontales y verticales respectivamente.

$$g_{\perp} = \sqrt{\frac{1}{b} \sum_{k=1}^b [\mathbf{e}(\mathbf{u}_{0,1}, \mathbf{u}_{0,2} + k) - \mathbf{e}(\mathbf{u}'_{0,1}, \mathbf{u}_{0,2} + k)]^2} \quad (2a)$$

$$h_{\perp} = \sqrt{\frac{1}{b} \sum_{k=1}^b [\mathbf{e}(\mathbf{u}_{0,1} + k, \mathbf{u}_{0,2}) - \mathbf{e}(\mathbf{u}_{0,1} + k, \mathbf{u}'_{0,2})]^2} \quad (2b)$$

El conjunto de Ecuaciones (2a) y (2b) se compara con la desviación estándar $s_{t(o)}$ de la región de transición en la imagen codificada, verificandose el bloque $\mathbf{B}_{t(o)}$ como: $g_{\perp} > \sigma_{t,1}$ y $h_{\perp} > \sigma_{t,2}$. Entonces, para realizar los cálculos en el espacio multi-resolución debe considerarse la condición sobre la existencia de al menos una transición de bloque, i.e., para el bloque $\mathbf{B}_{\mathbf{u}_o, \mathbf{b}}$ debe existir un bloque disyunto del mismo tamaño $\mathbf{B}_{\mathbf{u}'_o, \mathbf{b}}$ tal que $\mathbf{u}_{o,o} = \mathbf{u}'_o - \mathbf{b}$ para cualquier orientación $o=1,2$ Si

la condición anterior se cumple, la métrica se calcula sobre la estructura QD como sigue:

1. Calcular la QD para la imagen $\hat{\mathbf{X}}$
2. Para el índice $(\mathbf{u}_o, \mathbf{b})$ que define cada bloque \mathbf{B} , i.e., la posición del primer elemento del bloque \mathbf{B} y el tamaño del mismo bloque, encontrar el bloque vecino \mathbf{B}' del mismo tamaño \mathbf{b} para la orientación o .
3. Calcular $\sigma_B, \sigma_{B'}$ y $\sigma_{B, B'}$
4. Calcular g_{\perp} y h_{\perp}
5. Continuar con los bloques restantes hasta encontrar la métrica f_1 definida en Ginesu et al. (2006).

2.3 Error de bordes (Edge error)

En la detección de bordes sobre una imagen, la resolución de detección está directamente relacionada con la escala, en la cual los bordes se detectan. Así, una resolución muy alta implica una escala muy pequeña, lo que puede resultar en la detección del ruido como borde o detección de bordes fragmentados.

A su vez, una baja resolución determina una escala muy alta, haciendo que algunos bordes finos puedan ser no detectables. A diferencia de los detectores de borde clásicos, los basados en wavelets pueden ajustar la escala para evitar el inconveniente anterior Li (2003). Luego, los componentes principales de cada borde se ubican en las escalas de detalle de la transformación wavelet, de tal forma que los componentes de poca relevancia o ruidos se puedan localizar y separar.

Los coeficientes wavelet se pueden interpretar como una sub-muestra de la convolución con una familia wavelet $\Psi_{j,o}$. El hecho que $\Psi_{j,o}$ tenga momentos de desvanecimiento en las direcciones horizontal, vertical o diagonal, permite considerar a $\Psi_{j,o}$ como una derivada parcial de otra función de soporte finito, así, $\Psi_{j,o} = \partial_{2-o}\theta_{j,o}$ donde ∂_1 es la derivada en la primera variable ∂_2 la derivada en la segunda. Entonces,

$$dw_4 = d_{j,o}(k) = \mathfrak{F}_{2-o}(f^* q_{j,o}^*(2^j k)) \quad (3)$$

donde $\theta_{j,o}^*(k) = -\theta_{j,o}(-k)$. La ecuación (3) revela que los coeficientes wavelet son derivadas direccionales de una versión suavizada de la señal, y, por lo tanto, se interpreta éste como un primer paso para la detección de bordes.

El siguiente procedimiento se sugiere para la detección de bordes Li (2003):

1. Seleccionar, dada la escala j , los respectivos coeficientes wavelet $d_{j,o}(\mathbf{k})$, para $o=1,2,3$
2. Calcular, para cada $\mathbf{k} \in \mathbb{Z}^2$

$$\angle d_j(\mathbf{k}) = \begin{cases} \arctan(d_{j,2}(\mathbf{k})/d_{j,1}(\mathbf{k})), & d_{j,1}(\mathbf{k}) \\ \pi - \arctan(d_{j,2}(\mathbf{k})/d_{j,1}(\mathbf{k})), & \text{en otro caso} \end{cases}$$

3. Calcular los módulos máximos, que se define para los coeficientes wavelet como

$$M^j(\mathbf{k}) = M^j(\mathbf{k} \pm \mathbf{v}_j)$$

donde $M^j(\mathbf{k}) = \sqrt{\sum_{i=1}^3 |d_{j,i}(\mathbf{k})|^2}$, $\forall \mathbf{k} \in \mathbb{Z}^2$, es el módulo wavelet. Se escogen los siguientes valores: $\mathbf{v}_j \in \{(1,0), (0,1), (1,1), (1,-1)\}$, de tal manera que sea cercano al vector con coordenadas $(-\sin(\angle d_j(\mathbf{k})), \cos(\angle d_j(\mathbf{k})))$, que es ortogonal a la dirección del borde local.

4. Determinar el punto de borde, $2^j k$ cuando $\mathbf{k} \in \mathbb{Z}^2$ es modulo máximo: $M^j(\mathbf{k})$ es la magnitud y $\mathfrak{D}d_j(\mathbf{k})$ dirección del borde en la escala j , respectivamente.

Algunos criterios adicionales se usan para incrementar la robustez del detector, tales como la selección del filtro wavelet de descomposición o umbralización de los módulos máximos seleccionados. Concretamente, la umbralización debe asegurar que solamente los módulos máximos significantes sean tenidos en cuenta. Así, un criterio puede estar basado en el exponente de Lipschitz para medir la significancia de los bordes Li (2003). De esta manera, se hace el cambio del detector de bordes y la métrica propuesta en Ginesu et al. (2006) puede ser calculada a partir de los valores del modulo máximo como magnitud del borde y $\angle d_j(\mathbf{k})$ como su fase.

2.4 Deterioro Visual (Visual Impairment).

La distorsión perceptual de la imagen depende también del grado de visibilidad local en regiones arbitrarias. Tal grado de visibilidad se puede caracterizar estadísticamente mediante la búsqueda de homogeneidades locales en dichas regiones. El coeficiente de variación, denotado por C_v , es un indicador de la homogeneidad en regiones. Por tanto, el criterio de homogeneidad, basados en García et al. (2008), establece que, si el C_v en un bloque \mathbf{B} generado por la QD, es mayor que el C_v total, la partición en \mathbf{B} debe continuar. A partir de este criterio, la medida de distorsión se realiza mediante el siguiente procedimiento:

1. Descomponer la imagen \hat{X} en un conjunto de bloques \hat{Q}_i , donde $\mathbf{l} = (\mathbf{u}_0, b)$ es un conjunto de índices obtenidos por la QD.

2. Calcular, para cada $\hat{\mathbf{B}} \in \hat{\mathbf{Q}}_1 \subseteq \hat{\mathbf{X}}_{m \times n}$ y $\mathbf{B} \in \hat{\mathbf{Q}}_1 \subseteq \mathbf{X}_{m \times n}$, el error de luminancia, definido como:

$$\varepsilon_l(\mathbf{I}) = \frac{|\mu(\mathbf{I}) - \hat{\mu}(\mathbf{I})|}{\max(\mathbf{X})}$$

donde $\mu(\mathbf{I})$ y $\hat{\mu}(\mathbf{I})$ son los valores medios estimados de \mathbf{B} y $\hat{\mathbf{B}}$ respectivamente. Además, $\max(\mathbf{X})$ es un factor de normalización, usualmente el máximo valor de la escala de grises.

3. Calcular el error de similitud mediante la ecuación:

$$\varepsilon_s(\mathbf{I}) = \frac{\sigma_{\mathbf{B}-\hat{\mathbf{B}}, \mathbf{B}-\hat{\mathbf{B}}}}{\sigma_{\mathbf{B}, \mathbf{B}} + \sigma_{\hat{\mathbf{B}}, \hat{\mathbf{B}}}} \frac{1 - \rho(\mathbf{I})}{2}$$

donde el factor de correlación se estima como,

$$\rho(\mathbf{I}) = \sigma_{\mathbf{B}, \hat{\mathbf{B}}} / \sqrt{\sigma_{\mathbf{B}, \mathbf{B}} \sigma_{\hat{\mathbf{B}}, \hat{\mathbf{B}}}}$$

4. Calcular la contribución del l -ésimo bloque a la métrica, definido mediante la siguiente combinación lineal de ε_l y ε_s :

$$\varepsilon(l) = \frac{\varepsilon_s(\mathbf{I}) + \varepsilon_l(\mathbf{I})}{2}$$

5. Dada n la cantidad de bloques en la imagen, calcular la medida total del error:

$$f_3 = \sqrt{\frac{1}{n} \sum_{\mathbf{I}} \varepsilon^2(\mathbf{I})}$$

Como se dijo antes, los bloques $\hat{\mathbf{Q}}_1$ se calculan a partir de los valores de QD, asegurando que las disparidades visuales relevantes queden identificadas, sin apelar al uso de bloques regulares sobre la imagen completa. Las disparidades visuales relevantes son identificadas en la mayoría de los casos sobre los bordes y regiones no homogéneas.

3. Resultados y discusión

La Figura 4 detalla el proceso de estimación de los coeficientes de regresión y la validación del método propuesto.

3.1 Base de datos

La base de datos empleada es LIVE Image Quality Assessment Database Bovik (2009), que contiene los siguientes tipos de distorsión:

- 174 imágenes degradadas con Fast Fading.
- 174 imágenes degradadas con Blurring.
- 227 imágenes degradadas con compresión JPEG 2000.
- 233 imágenes degradadas con compresión JPEG.
- 174 imágenes degradadas con ruido blanco gaussiano.

En Bovik (2009), se presentan las calificaciones de una valoración subjetiva, Sheikh et al. (2006), en particular de la medida DMOS (Differential

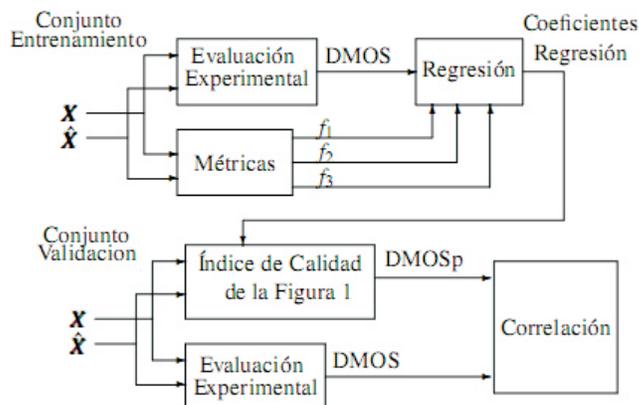


Figura 4. Regresión y validación.

Mean Opinion Score), para cada una de las imágenes de la base de datos analizada. La validación de las métricas se realiza calculando el índice de correlación entre los valores DMOS y los valores resultantes del índice de calidad, en función de las tres métricas analizadas: $\{f_i: i = 1,2,3\} \in \mathbb{R}^1[0,1]$, siendo f_1 la estimación de degradación de bloques, f_2 el error de borde y f_3 el valor de la degradación visual. Los coeficientes del índice se calculan por medio de una regresión lineal sobre un conjunto de entrenamiento seleccionado de los DMOS y las métricas medidas para cada imagen relacionada. De esta manera, el 70% de los valores son para el entrenamiento y el 30% restante se usa para la validación. Así, la regresión lineal queda de la forma $DMOSP = \lambda_0 + \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$, donde $\{\lambda_i: i = 0,1,2,3\} \in \mathbb{R}^1$ denotan los coeficientes de una regresión lineal multivariada, como se valida sobre un MOS diferencial el coeficiente l_0 es despreciable. Así, se calcula el índice de correlación entre los valores DMOS y los índices de calidad calculados por el conjunto de entrenamiento, para cada base de datos relacionada con cada tipo de distorsión y se calcula de nuevo para toda la base de datos. El conjunto de entrenamiento es permutado de nuevo

y se repite el cálculo 10 veces. El resultado es una secuencia de 10 muestras de la correlación para cada distorsión, otra secuencia de 10 muestras de la correlación de la base completa y el promedio y desviación estándar de los 10 intentos. Estos resultados se validan utilizando 2 medidas de error (error máximo $\max(e)$ y error cuadrático medio $RMSE$) y 3 medidas estadísticas (cantidad de valores atípicos N_0 y coeficiente de correlación de Pearson R_p y de Spearman R_s) sobre cada secuencia. Un primer criterio de validación indica que si las medidas de error resultan en valores relativamente grandes, las métricas no podrían ser consistentes con la evaluación subjetiva de la imagen. Un segundo criterio de validación indica que si la magnitud de la correlación se acerca a cero, indica que las métricas no estarían detectando la naturaleza de la distorsión, interpretando las imágenes como datos atípicos, obteniendo un índice de calidad errado sobre la imagen.

3.2 Validación

A continuación se muestran los resultados obtenidos al aplicar el procedimiento descrito anteriormente sobre la base de datos del LIVE Bovik (2009).

Tabla 1. Validación para cada tipo de distorsión.

Distorsión	*Descripción	$\max(e)$	$RMSE$	R_p	R_s	l_1	l_2	l_3	
Fast Fading	Clásico	<i>m</i>	0.31	0.08	0.96	0.96	1.04	1.94	0.13
		<i>s</i>	0.11	0.00	0.00	0.00	0.30	0.31	0.49
	Propuesto	<i>m</i>	0.19	0.07	0.97	0.97	0.55	1.84	1.24
		<i>s</i>	0.02	0.00	0.00	0.00	0.10	0.21	0.28
Blurring	Clásico	<i>m</i>	0.31	0.08	0.94	0.92	0.28	2.39	0.93
		<i>s</i>	0.02	0.00	0.00	0.01	0.30	0.34	0.60
	Propuesto	<i>m</i>	0.14	0.04	0.98	0.97	0.60	1.06	2.26
		<i>s</i>	0.00	0.00	0.00	0.00	0.10	0.20	0.27
JPEG2000	Clásico	<i>m</i>	0.33	0.10	0.94	0.93	0.44	2.38	2.47
		<i>s</i>	0.03	0.00	0.00	0.01	0.82	0.79	0.98
	Propuesto	<i>m</i>	0.25	0.07	0.97	0.96	1.19	0.71	1.60
		<i>s</i>	0.02	0.00	0.00	0.00	0.07	0.14	0.24
JPEG	Clásico	<i>m</i>	0.33	0.08	0.95	0.93	0.65	2.53	0.06
		<i>s</i>	0.04	0.00	0.00	0.00	0.39	0.45	0.59
	Propuesto	<i>m</i>	0.26	0.07	0.96	0.94	1.25	0.33	-0.09
		<i>s</i>	0.02	0.00	0.00	0.00	0.14	0.30	0.22
Ruido blanco	Clásico	<i>m</i>	0.23	0.07	0.97	0.95	-0.03	2.33	3.22
		<i>s</i>	0.03	0.00	0.00	0.00	0.13	0.22	0.61
	Propuesto	<i>m</i>	0.19	0.07	0.97	0.95	0.02	2.27	1.21
		<i>s</i>	0.01	0.00	0.00	0.00	0.10	0.30	0.26

*Método utilizado - variable medida

Tabla 2. Validación para el conjunto completo.

*Descripción		max(e)	N_o	RMSE	R_p	R_s	I_1	I_2	I_3
Clásico	<i>m</i>	0.43	0.40	0.09	0.94	0.93	0.07	3.19	0.73
	<i>s</i>	0.10	0.51	0.00	0.00	0.00	0.05	0.09	0.28
Propuesto	<i>m</i>	0.44	0.00	0.11	0.92	0.92	-0.06	2.75	1.93
	<i>s</i>	0.01	0.00	0.00	0.00	0.00	0.06	0.14	0.26

*Método utilizado - variable medida

En la Tabla 1 se muestran los promedios y la desviación estándar, utilizando el método de validación descrito anteriormente para el conjunto de imágenes degradado mediante cada distorsión, observando que el método propuesto se aproxima más a los valores DMOS para los tipos de distorsión blurring, JPEG2000 y JPEG.

En la Tabla 2 se muestran los promedios y la desviación estándar para el conjunto completo de imágenes, observando que el método clásico se aproxima más a los valores DMOS. Sin embargo, la desviación estándar en el método clásico es mayor que en el propuesto, con lo que, de acuerdo a los criterios de validación, se indica que las métricas pueden ser sensibles a detalles desconocidos que presenta la imagen. Como se observa, el coeficiente correspondiente a f_2 (error de bordes), i.e. λ_2 es mayor que los demás, lo que indica que el detector de bordes provee información relevante sobre la naturaleza de los errores a ser detectados. Luego, la técnica utilizada para f_2 , en el método propuesto, tiene asignado un coeficiente menor que el método clásico, lo que significa que el detector basado en wavelet aproxima mejor los bordes detectados.

Además el coeficiente de f_1 i.e. λ_1 , es bajo en comparación con los demás coeficientes, esto se puede explicar porque ciertas distorsiones se presentan en forma conjunta, i.e., una transición de bloque puede degradar un borde de tal manera que puede ser detectada por la métrica de distorsión de bloque y por la métrica de distorsión de borde, mostrando que esta transición podría ser medida de manera implícita en las otras dos métricas. Todos los resultados presentados están al mismo nivel que los presentados en Sheikh et al. (2006) sin ningún tipo de regresión o de ajuste no-lineal.

En la Tabla 3 se muestra la relación entre la desviación estándar de los coeficientes del método propuesto respecto a los coeficientes del método clásico σ_p / σ_c , mostrando una mayor estabilidad para los casos de ruido blanco y de la base de datos completa en la métrica clásica debido a que la relación es superior a 1.6 ± 0.2 y 1 ± 0.4 , respectivamente. Los valores se normalizan bajo la media del tiempo del cómputo del método propuesto. Como resultado, se observa que la métrica propuesta depende del tipo de imagen que se mida, ya que si se poseen muchos detalles el índice tarda en encontrar las regiones

Tabla 3. Relación desviación estándar coeficientes regresión

Distorsión	I_1	I_2	I_3
Fast Fading	0.33	0.67	0.57
Blurring	0.33	0.58	0.45
JPEG2000	0.08	0.17	0.24
JPEG	0.35	0.66	0.37
Ruido Blanco	0.76	1.36	0.42
Base de datos completa	1.20	1.55	0.92

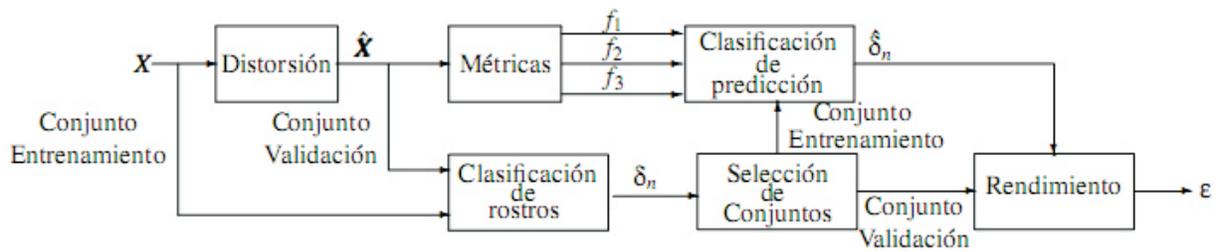


Figura 5. Esquema de Evaluación.



Figura 6. Ejemplos de rostros de la base de datos del CMU.

relativamente homogéneas. En cualquier caso, el método propuesto no supera el tiempo de cómputo del método clásico.

3.3 Caso de aplicación: Imágenes de rostros

Una aplicación del método propuesto consiste en un esquema de identificación de rostros basados en el espacio vectorial (*Principal Component Analysis - PCA*) Turk & Pentland (1991), para la evaluación de las imágenes distorsionadas con JPEG, JPEG2000, *Gaussian blurring*, *Fast Fading* o ruido blanco, para determinar cuáles contienen la información necesaria para ser

identificada y clasificada de manera adecuada. Este esquema se lleva a cabo utilizando las métricas propuestas en la Sección 2.

En la Figura 5 se muestra el procedimiento utilizado para evaluar el esquema propuesto y a continuación se describen cada uno de sus bloques:

- En la variable \mathbf{X} , están contenidas las imágenes de la Carnegie Mellon University (CMU Face Images Data Set) Mitchell (1994), en las cuales la persona está mirando hacia el frente Figura 6.
- En el bloque **Distorsión**, se lleva a cabo la degradación de las imágenes contenidas en \mathbf{X}

utilizando los tipos de ruido ya mencionados obteniendo así el conjunto \hat{X}

- En la etapa de **clasificación de rostros** se utiliza el conjunto en X como conjunto de entrenamiento y el conjunto de validación son las imágenes distorsionadas \hat{X} , para obtener del clasificador una salida definida como $\delta_n : \{0,1\}$, donde un **TRUE** ($\delta_n = 1$) significa que la imagen fue bien clasificada y un **FALSE** ($\delta_n = 0$) es el caso contrario.
- En el bloque **Métricas** se evalúan el conjunto de métricas clásicas y las propuestas, con las imágenes en X como referencia.
- En **clasificador de predicción** se estiman el **TRUE** ($\hat{\delta}_n = 1$) o el **FALSE** ($\hat{\delta}_n = 0$), teniendo como conjunto de entrada las métricas evaluadas en la etapa anterior.
- Por último, se mide el rendimiento del sistema, midiendo el error (ϵ) de predicción según la etiqueta entregada por el clasificador de predicción y la etiqueta real entregada por el clasificador de rostros.
- En la Tabla 4 se muestran los resultados, obtenidos luego de aplicar el esquema de la Figura 5 al Conjunto de Datos de Imágenes de Rostros de la *Carnegie Mellon University Mitchell* (1997). Los mejores resultados de cada conjunto de métricas se resaltan en dicha Tabla.

3.4 Discusión de Resultados

Como se observa en las Tablas 1 y 2, al utilizar QD y detectores de bordes multi-resolución mejora al menos en 2% la predicción objetiva de calidad (R_p y R_s) y reduce a 0 el número de datos atípicos (N_o) en cualquier tipo de distorsión. Un valor agregado es que el método propuesto también evalúa distorsiones diferentes a los producidos por JPEG y JPEG2000, i.e., puede ser evaluadas distorsiones que no presentan de forma directa, e.g., errores de bloque.

La Tabla 3 muestra la gran estabilidad en los coeficientes de la regresión, refiriéndose a estabilidad en el sentido que la desviación estándar es baja comparada con las métricas clásicas, i.e., las constantes de regresión tienen cambios pequeños con el cambio de conjunto de entrenamiento. Así, se proporciona mayor estabilidad en el índice de calidad propuesto, i.e., para casi todos los casos la relación $\sigma_p / \sigma_c < 1$ dando en promedio un incremento de la estabilidad del índice del 38% con respecto al índice clásico. Por otro lado, cuando la imagen es de alto contraste, estas tienen una gran variedad de bordes haciendo que QD requiera demasiado tiempo de cómputo para llegar a esas regiones relativamente homogéneas. Si la imagen es de bajo contraste, estas tienen pocos bordes haciendo que las regiones homogéneas sean fácilmente encontradas por QD, reduciendo el tiempo de cómputo considerablemente. Para la base de datos usada se encuentra en promedio una reducción de tiempo del 37%.

Tabla 4. Resultados de validación del esquema propuesto

Distorsión	Método Clásico						Método Propuesto					
	e		d_1		d_2		e		d_1		d_2	
	m	s	m	s	m	s	m	s	m	s	m	s
Fast Fading	0.07	0.01	4.40	2.31	18.20	1.39	0.08	0.00	0.50	1.08	22.20	2.04
Ruido blanco	0.08	0.00	0.90	1.19	21.90	1.96	0.08	0.01	1.70	1.70	21.20	2.34
JPEG	0.07	0.01	1.10	1.85	21.50	2.41	0.08	0.01	2.30	2.00	21.00	3.52
JP2K	0.08	0.01	3.70	3.30	19.90	3.10	0.08	0.02	3.00	4.50	21.30	2.62
Blurring	0.08	0.00	1.00	2.00	22.20	2.57	0.07	0.00	0.20	0.42	20.50	2.46

La Tabla 4 muestra que los resultados obtenidos al evaluar el sistema de la Figura 5, resaltado se encuentra el mejor rendimiento para cada uno de las técnicas. Así, para el caso del conjunto de métricas clásico se tiene en promedio un $7 \pm 1\%$ de error de clasificación. De las cuales en promedio 4 ± 2 (δ_1) corresponde al número de imágenes que tienen la información suficiente para la identificación pero fueron clasificadas como imágenes con información insuficiente y 18 ± 1 (δ_2) corresponde al número de imágenes que no contenían información suficiente para la identificación pero fueron clasificadas como imágenes con información suficiente. Para el caso del método propuesto se tienen resultados muy similares ya que en promedio solo se obtiene un 7% de error en la clasificación. Adicional se disminuye a 0 (δ_1) que corresponde al número de imágenes que se determinan como insuficientes para la identificación dado que si tienen la información suficiente y aumenta a 20 ± 2 (δ_2) la cantidad de imágenes que son clasificadas como suficientes para la identificación dado que su contenido de información no es suficiente.

4. Conclusiones

El método propuesto de traslación de medidas de distorsión a espacios de representación multi-resolución resulta ser eficiente, por lo menos para los casos de análisis de imágenes naturales y rostros. El método, que incluye medidas de distorsión en un modelo multi-resolución basados en HVS, usando QD y detectores de borde que usan wavelet, es invariante a la traslación y presenta mejoras en la predicción de la distorsión. Además, para el cálculo de f_1 se obtiene una considerable reducción de tiempo de cómputo debido a que el proceso no se hace sobre todos los bloques sino sobre aquellos que se consideran particulares.

El método propuesto, además de presentar resultados comparables con los mostrados en el estado del arte, es más flexible, debido a que la estabilidad en los coeficientes de regresión sugiere que el método puede ser extendible a nuevos tipos de distorsión o bases de datos.

El principal inconveniente del método es la selección de los filtros para el detector de bordes

con la transformada wavelet, ya que todos los filtros no producen el mismo rendimiento en el mapa de detección, de esta manera se puede producir una estimación errónea del índice de calidad. Dejando abierto como trabajo futuro la selección óptima de los filtros wavelet y el diseño de un detector de bordes que no se vea afectado por el cambio de filtro. Así mismo, basados en los resultados en el método propuesto para clasificación de imágenes con pobre calidad, se puede decir que las métricas aunque no sean completamente discriminantes, pueden aportar información relevante para estimar hasta que punto una imagen es válida para una tarea de identificación. Así como trabajo futuro se propone utilizar métricas no-referenciadas para realizar la tarea de clasificación de acuerdo con el contenido de información.

5. Agradecimientos

Los autores desean agradecer al Proyecto financiado por COLCIENCIAS: *Identificación de posturas labiales en pacientes con labio y/o paladar hendido corregido*. Así mismo, se agradece al profesor Hartmut Führ, RWTH Aachen University, por su colaboración y sugerencias.

6. Referencias Bibliográficas

- Avcibas, I., Memon, N., & Sankur, B. (2003). Steganalysis using image quality metrics. *Journal of IEEE Transactions on Image Processing* 12 (7), 221-229
- Beegan, A. (2001). *Wavelet-based Image Compression Using Human Visual System Models*. Doctoral thesis, Virginia Polytechnic Institute and State University.
- Bovik, A (2009). *Laboratory for Image & Video Engineering*. <http://live.ece.utexas.edu/>
- Chen, S., Hsiao, Y., Huang, Y., Kuo, S., Tseng, H., Wu, H., & Chen, D. (2009). Comparative Analysis of Logistic Regression, Support Vector Machine and Artificial Neural Network for the Differential Diagnosis of Benign and Malignant Solid Breast Tumors by the Use of Three-Dimensional Power Doppler Imaging. *Journal of Korean J Radiol* 10(5), 464-471

- Gaddipati, A., Machiraju, R., & Yagel, R. (1997). Steering Image Generation with Wavelet Based Perceptual Metric. *Computer Graphics Forum* 16 (3), 241-251
- García, J., & Castellanos, G. (2008). Estimación de la tasa de compresión utilizando medidas de calidad objetivas completamente referenciadas. *Ingeniería y Competitividad*. 10 (1), 73-83
- García, J., Castellanos, G., & Ortiz, B. (2008). *Image information access using Wedgelet filters*. In Proceedings of IEEE First International Symposium on Applied Sciences in Biomedical and Communication Technologies.
- García, J., Führ, H., Ortiz, B., & Castellanos, G. (2009). *Image Distortion Measure Mappings in Wavelet Representation Space*. In Proceedings of Fourth International Conference on Broadband Communications, Information Technology and Biomedical Applications.
- Ginesu, G., Massidda, F., & Giusto, D. (2006). A Multi-factors approach for image quality assessment based on a human visual system model. *Journal of Signal Processing: Image Communication*. 21 (4), 316-333
- Ibarra, O., Ortiz, B., Reinoso, J., Henao, V., & García, J. (2007). *Evaluación de la Calidad del Servicio en un Sistema Integrado Duo-Play de Ludoteca Multimedia*. In Proceedings of IEEE Congreso Colombiano de Comunicaciones.
- ITU-R Recommendation (2002). *BT.500-11 Methodology for the subjective assessment of the quality of television pictures*.
- Li, J. (2003). *A Wavelet approach to edge detection*. Master's thesis, Sam Houston State University.
- Mitchell, T (1994). *A Neural Network Face Recognition Assignment*.
http://www.cs.cmu.edu/afs/cs.cmu.edu/usr/avrim/www/ML94/face_homework.html
- Mitchell, T (1997). *Machine Learning*. McGraHill
- Moumkine, M., Tamtaoui, A., & Ouahman, A. (2006). *Integration of the Contrast Sensitivity Function into Wavelet Codec*. In Proceedings of Second International Symposium on Communication, Control and Signal Processing.
- Sheikh, H., Sabir, M., & Bovik, A. (2006). A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms. *Journal of IEEE Transactions on Image Processing*. 15 (11), 3440-3451
- Turk, M., & Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*. 3 (1), 71-86