

# Análise de Métricas de Qualidade Visual Aplicadas na Caracterização de Limiares Perceptuais com Variação da Imagem de Referência

Ronaldo de Freitas Zampolo, Diego de Azevedo Gomes e Rui Seara

**Resumo**—Neste trabalho, é apresentado um estudo sobre o comportamento de algumas métricas de qualidade visual aplicadas à caracterização de limiares de percepção de diferenças entre imagens, baseado em um procedimento publicado recentemente pelos mesmos autores. O conjunto de teste utilizado é composto por 360 imagens, com degradações do tipo ruído impulsivo, sal e pimenta, e ruído gaussiano branco aditivo. Tomaram parte na coleta de dados experimentais aproximadamente 160 avaliadores. As métricas usadas são de referência completa de abordagem convencional (SNR, PSNR, e CC) e perceptual (IFC, VIF, C4 e MSSIM). Os resultados obtidos mostram, para cada tipo de degradação, o comportamento médio das métricas utilizadas para diferentes imagens de referência.

**Palavras-Chave**—Comparação entre métricas, JND, limiares de percepção, métricas perceptuais, qualidade visual.

**Abstract**—This work presents a study on the behavior of some visual quality metrics applied to the characterization of image difference perceptual thresholds, based on a procedure published previously by the same authors. The set of test images is composed of 360 images, degraded by either impulsive noise, salt and pepper noise or additive white Gaussian noise. About 160 individuals have taken part in the experimental data acquisition. Full reference metrics of conventional (SNR, PSNR, and CC) and perceptual (IFC, VIF, C4, and MSSIM) approaches are tested. The results show, for each type of degradation, the average behavior of such metrics for different reference images.

**Keywords**—Metric comparison, JND, perception thresholds, perceptual metrics, visual quality.

## I. INTRODUÇÃO

Imagens digitais estão sempre sujeitas a diferentes tipos de distorção, os quais podem ser inseridos pelos processos de aquisição, codificação, armazenamento, transmissão e exibição. Assim, a otimização de sistemas de processamento de imagem, bem como a avaliação de diferentes algoritmos, deve ser feita mediante a adoção de algum critério para se determinar qual ajuste ou estratégia resulta em imagens de melhor qualidade. A abordagem para tal avaliação de qualidade pode ser subjetiva ou objetiva.

A avaliação subjetiva é feita através de testes realizados com alguns avaliadores, os quais emitem suas opiniões acerca da qualidade das imagens que lhes são apresentadas. Apesar de seus resultados retratarem a qualidade visual percebida, o

que é essencial se uma aplicação é concebida para consumo humano, esse método exige muito tempo para ser implementado, sendo de pouca utilidade prática. Por sua vez, as avaliações objetivas pressupõem que métricas de qualidade sejam adotadas para automatizar o processo de avaliação das imagens-teste (imagens sobre as quais se deseja ter um índice de qualidade). As métricas de qualidade visual possuem um custo relativamente baixo em relação à abordagem subjetiva, contudo, ainda não conseguem estimar fielmente a percepção humana. Essas métricas podem ser classificadas em três tipos, de acordo com a quantidade de informação sobre a imagem de referência necessária para o cálculo dos índices de qualidade: (i) referência completa, na qual a imagem de referência deve estar disponível para que se possa medir a qualidade de uma imagem-teste; (ii) referência reduzida, em que é suficiente apenas informação parcial sobre a imagem de referência; e (iii) sem referência, em que não há a necessidade de uma imagem de referência para que a imagem-teste possa ser avaliada. As métricas podem ainda ser divididas em não-perceptuais e perceptuais. As não-perceptuais não levam em conta as características do sistema visual humano (SVH) nas suas formulações. Como exemplo dessas métricas, tem-se o erro quadrático médio (*mean-square error* – MSE) e a razão sinal-ruído de pico (*peak signal-to-noise ratio* – PSNR). As métricas perceptuais são modelos matemáticos derivados a partir de experimentos subjetivos (modelagem de alto nível do SVH) que buscam quantificar, por exemplo, a sensibilidade visual a ruídos por sub-banda, luminância local e mascaramento de contraste [1], [2].

Neste artigo, é efetuada uma avaliação de desempenho de métricas de qualidade visual aplicadas na caracterização de limiares de percepção de diferenças entre imagens com base na metodologia proposta em [3]. Possíveis aplicações de tal caracterização incluem codificação perceptual, inserção de marca d'água digital e redução de ruído. Neste trabalho, há duas diferenças em relação ao apresentado em [3]: o número de tipos de degradação é maior (ruído gaussiano branco aditivo, ruído impulsivo, e ruído *sal e pimenta*; no trabalho anterior, apenas ruído gaussiano branco aditivo foi usado); e todas as imagens do conjunto de teste são usadas como imagens de referência para estimação de limiares de percepção intermediários (em [3], apenas as imagens originais, sem degradação, foram usadas como imagens de referência). Outras características do experimento: as métricas usadas são de referência completa; 160 avaliadores participam das avaliações; e o conjunto de teste é formado por 300 imagens

Ronaldo de Freitas Zampolo e Diego de Azevedo Gomes, Faculdade de Engenharia da Computação, Universidade Federal do Pará, Belém, Brasil, E-mails: zampolo@ufpa.br, diagoges@yahoo.com.br. Rui Seara, Laboratório de Circuitos e Processamento de Sinais - LINSE, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis - SC, Brasil, E-mail: seara@linse.ufsc.br

em tons de cinza derivadas da base IVC [4].

As demais seções são organizadas da seguinte maneira. A Seção II relaciona as métricas de qualidade utilizadas neste artigo. Na Seção III, é mostrado como o processo para coletas de dados foi executado. Na Seção IV, é apresentado o método para calcular a JND (*just noticeable difference*) [5], ferramenta escolhida para identificar os limiares de percepção. A Seção V mostra e discute os resultados obtidos. Finalmente, na Seção VI, são apresentadas as conclusões e as propostas para trabalhos futuros.

## II. MÉTRICAS DE QUALIDADE VISUAL

O *International Telecommunication Union* (ITU) bem como o *Video Quality Experts Group* (VQEG) vêm estabelecendo diretrizes tanto para a criação quanto para a avaliação de desempenho de métricas de qualidade visual [6], [7]. O processo de desenvolvimento de uma métrica para avaliação de qualidade visual pode ser visto como tendo três etapas: (i) formação de um conjunto de imagens-teste avaliadas; (ii) desenvolvimento de uma métrica de qualidade visual; e (iii) avaliação do desempenho da nova métrica, através da comparação desta com outras métricas já existentes.

Na formação do conjunto de imagens-teste, fatores importantes a considerar são os tipos de cena das imagens (ar livre, ambientes fechados, paisagens naturais, gráficos, rostos, etc.); e também o tipo de degradação que as imagens evidenciam (diferentes estratégias de codificação, presença de ruído gaussiano branco aditivo, ruído impulsivo, borramento, dentre outros).

Para a avaliação subjetiva, vários procedimentos podem ser adotados. Por exemplo, uma única imagem de cada vez é apresentada aos observadores que devem classificá-la em uma escala de 0 a 100, na qual 0 indica a pior qualidade possível e 100, a máxima qualidade. Em outra abordagem possível, um par de imagens é apresentado aos observadores que, por sua vez, são orientados a indicar se há ou não diferença entre as imagens do par.

Todo o contexto em que os dados experimentais são obtidos (o que, por sua vez, guarda relação com a aplicação pretendida) tem influência no desempenho de determinada métrica de qualidade visual, fazendo com que métricas de muito bom desempenho em uma dada situação sejam medíocres em outras. A seguir, são apresentadas as métricas utilizadas neste trabalho.

### A. Razão Sinal-Ruído (SNR) e Razão Sinal-Ruído de Pico (PSNR)

A razão sinal-ruído (*signal-to-noise ratio* – SNR) e a razão sinal-ruído de pico (*peak signal-to-noise ratio* – PSNR) são duas métricas de referência completa bastante difundidas. Tais métricas basicamente avaliam a energia da diferença entre o sinal de referência e o sinal-teste. O SNR e o PSNR, em dB, são expressos como segue:

$$SNR = 10 \log_{10} \left\{ \frac{\sum_n x(n)^2}{N \cdot MSE} \right\} \quad (1)$$

$$PSNR = 10 \log_{10} \left\{ \frac{k^2}{MSE} \right\} \quad (2)$$

onde  $N$  é o número de *pixels* da imagem;  $k$  é o maior valor que uma *pixel* pode assumir (para imagens em tons de cinza de profundidade de 8 bits/*pixel*,  $k$  assume o valor de 255); MSE é dado por

$$MSE = \frac{1}{N} \sum_n [x(n) - y(n)]^2 \quad (3)$$

onde  $x(n)$  e  $y(n)$  representam, respectivamente, as imagens de referência e de teste.

Tanto o SNR quanto o PSNR possuem a grande vantagem de exibir complexidade computacional baixa.

### B. Coeficiente de Correlação de Pearson (CC)

Neste artigo, usou-se o coeficiente de correlação de Pearson (CC) calculado entre as imagens de referência e de teste. É importante lembrar que, uma vez adotando o CC como métrica, a avaliação de qualidade torna-se insensível para imagens de teste produzidas por escalamento uniforme de amplitude e variação de nível DC em relação às imagens de referência. O CC é definido como

$$CC = \frac{\sum_n [x(n) - \bar{x}][y(n) - \bar{y}]}{\sqrt{\{\sum_n [x(n) - \bar{x}]^2\} \{\sum_n [y(n) - \bar{y}]^2\}}} \quad (4)$$

onde  $\bar{x}$  e  $\bar{y}$  denotam as médias dos sinais  $x(n)$  e  $y(n)$ , respectivamente.

### C. Similaridade Estrutural Média (MSSIM)

A métrica similaridade estrutural média (*mean structural similarity* – MSSIM) [8], desenvolvida e validada através da base LIVE [9], é considerada uma métrica de baixa complexidade computacional, apesar de pertencer à classe das métricas perceptuais. A MSSIM mede o quanto a estrutura da imagem de teste está distante da estrutura da imagem de referência, associando a similaridade estrutural avaliada com a qualidade percebida. A métrica em questão vem despertando a atenção da comunidade de pesquisadores devido aos bons resultados obtidos na representação da qualidade percebida como também pela sua simplicidade quando comparada com outras métricas psicovisuais. A MSSIM é calculada através de uma função que envolve elementos que buscam caracterizar diferenças entre duas imagens associadas à luminância, o contraste e a estrutura. A MSSIM é dada por

$$MSSIM = \frac{1}{W} \sum_w [l_w(x, y)]^\alpha [c_w(x, y)]^\beta [s_w(x, y)]^\gamma \quad (5)$$

onde  $l_w(x, y)$ ,  $c_w(x, y)$  e  $s_w(x, y)$  são, respectivamente, os componentes de luminância, contraste e estrutura calculadas em uma dada vizinhança  $w$  em torno de um *pixel*; e  $W$  é o número total de vizinhanças processadas. Em geral, a janela que define uma vizinhança desliza pela imagem. Ou seja, o número total de vizinhanças  $W$  é igual ao de *pixels* da imagem. A seguir, têm-se as formulações matemáticas dos componentes da MSSIM.

$$l_w(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (6)$$

$$c_w(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (7)$$

$$s_w(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (8)$$

onde  $\mu_i$  e  $\sigma_i$  denotam a média e o desvio-padrão de uma variável  $i$ , respectivamente; e  $\sigma_{xy}$  é a covariância entre  $x$  e  $y$ . Os parâmetros  $C_1$ ,  $C_2$  e  $C_3$  são constantes incluídas para evitar instabilidades quando os valores dos denominadores se aproximam de zero.

Nas avaliações deste trabalho, os componentes de luminância, contraste e estrutura são também considerados individualmente, sendo representados por  $MSSIM_l$ ,  $MSSIM_c$  e  $MSSIM_s$ , respectivamente. Assim, exceto pelos valores de  $\alpha$ ,  $\beta$  e  $\gamma$ , são usados os valores-padrão definidos pelos autores da MSSIM no código de referência dessa métrica.

#### D. Critério de Fidelidade de Informação (IFC)

A abordagem do critério de fidelidade de informação (*information fidelity criterion* – IFC) parte do pressuposto que o sistema visual humano evoluiu e adaptou-se a partir de estímulos provenientes de cenas naturais, sendo, portanto, sensível às variações das estatísticas que caracterizam esse tipo de imagem [10]. O IFC, então, assume que as estatísticas de imagens naturais podem ser modeladas no domínio das *wavelets* por um modelo tipo GSM (*Gaussian scale mixtures*) e define a qualidade visual como sendo a informação mútua entre uma imagem-teste e uma imagem de referência, considerando todas as sub-bandas envolvidas no processo de análise.

#### E. Fidelidade de Informação Visual (VIF)

A métrica fidelidade de informação visual (*visual information fidelity* – VIF) [11] consiste em uma extensão do IFC. As mesmas hipóteses e modelos são considerados; contudo, nesse caso, a qualidade visual é definida a partir da medida do quanto de informação a imagem de referência possui e o quanto dela pode ser extraída da imagem-teste. As análises desse trabalho apresentam resultados para a VIF, tal qual definida em [11], e para uma versão simplificada, de menor complexidade computacional, denominada VIFp. Essa simplificação é proposta pelos mesmos autores da VIF original.

#### F. C4

Esta métrica é proposta em [12], onde é validada experimentalmente. As etapas para determinar a qualidade visual, em concordância com a métrica C4, consistem em: (a) normalização dinâmica, (b) correção de gama, (c) conversão para um espaço de cores perceptual (ACr1Cr2), (d) filtragem usando função de sensibilidade ao contraste (*contrast sensitivity function* – CSF), (e) decomposição em canais perceptuais, (f) extração

de segmentos orientados, (g) transformação de coordenadas e (h) medida de similaridade. A implementação da métrica C4 usada neste trabalho pode ser obtida em <http://membres.lycos.fr/dcapplications/>. Essa métrica apresenta elevada complexidade computacional, evidenciada pelo número de etapas necessárias para o seu cálculo. Dois tipos de resultados podem ser obtidos pela implementação usada da C4. O primeiro consiste em uma medida de similaridade estrutural (denominada C4sim) e, o segundo, em um mapeamento não-linear dessa para uma medida de qualidade (*quality score*, chamada C4qs). A variação de C4sim é de 0 (pior similaridade estrutural possível) a 1 (identidade estrutural), enquanto a C4qs varia de 0 (qualidade mais baixa) a 5 (excelente qualidade).

### III. METODOLOGIA

Para a formação do conjunto de imagens-teste foram utilizadas as imagens originais (sem distorção) em tons de cinza da base IVC [4], denominadas *Plane*, *Barbara*, *Boats*, *Clown*, *Fruits*, *House*, *Isabel*, *Lena*, *Mandrill* e *Peppers*. Para cada uma dessas 10 imagens, foram geradas outras 22 imagens: 11 degradadas por ruído impulsivo e 11 por ruído do tipo *sal e pimenta*. Para as imagens *Barbara*, *Clown*, *Fruits*, *Isabel* e *Mandrill*, além das duas degradações mencionadas, foram geradas para cada uma 11 imagens degradadas por adição de ruído branco gaussiano. O conjunto de teste foi então organizado em grupos de 12 imagens a partir da imagem original e do tipo de degradação, totalizando 25 grupos. Dentro de um grupo, as imagens são identificadas pelo nome da imagem original do grupo, o tipo de degradação e um número de 0 a 11, em que 0 é atribuído à imagem original e os outros números são atribuídos às imagens de acordo com a intensidade da degradação, de forma que quanto maior o número, maior o nível de degradação. Algumas imagens do grupo *lena\_sp* (imagem original *Lena* e degradação do tipo *sal e pimenta*) são exibidas na Fig. 1.

Os experimentos foram realizados no Laboratório de Processamento de Sinais (LaPS) da Universidade Federal do Pará. Para a automatização dos experimentos, foi desenvolvido um *software* na linguagem de programação *Java* e administrador de banco de dados *MySQL*. O referido *software* possui formulário para cadastro de observadores, em que são solicitadas informações, tais como nome, data de nascimento, alterações na visão, experiência em processamento digital de imagens, dentre outras. Feito o cadastro, 120 pares de imagens selecionados aleatoriamente são, um par por vez, exibidas ao avaliador. Vale ressaltar que as imagens de um par pertencem ao mesmo grupo. Diante de um par de imagens, o observador deve indicar qual imagem apresenta maior nível de degradação. O tempo total de cada experimento é em torno de 10 minutos, bem abaixo do tempo recomendado pelo ITU [6]. O *software* fornece ainda suporte à análise gráfica dos resultados.

São efetuadas cerca de 160 realizações do experimento, utilizando monitores de cristal líquido (*liquid crystal display* – LCD) de 17 polegadas, com resolução de  $1280 \times 1024$  *pixels*. Os observadores são indivíduos de ambos os sexos



Fig. 1. Exemplo de grupo. (a) Imagem original *Lena*. (b)-(d) Três versões da imagem *Lena* degradadas por ruído do tipo *sal e pimenta*.

com idade entre 17 e 60 anos, a maioria sem experiência em processamento digital de imagens.

#### IV. JND E INTERPRETAÇÃO DOS RESULTADOS

Nesse tipo de experimento, em que o observador indica a imagem com maior degradação dentre duas apresentadas, são possíveis apenas duas respostas: as corretas (quando o avaliador aponta a imagem que realmente tem o maior nível de ruído); e as incorretas (quando o observador indica como mais ruidosa a imagem que na verdade apresenta um nível de ruído menor). Tais respostas resultam de duas situações: (a) a diferença entre as duas imagens apresentadas é efetivamente detectada, o que sempre leva a uma resposta correta; (b) a diferença não é percebida, resultando na seleção aleatória de uma das imagens pelo usuário, o que leva a um igual número de respostas corretas e incorretas por simetria [5]. Nessa situação, a probabilidade de resposta correta  $pc$  (probabilidade do observador indicar a imagem que realmente tem mais ruído) é dada por

$$pc = pd + \frac{1 - pd}{2} \quad (9)$$

$$= \frac{pd + 1}{2} \quad (10)$$

onde  $pd$  corresponde à probabilidade da diferença entre as imagens de um determinado par ser efetivamente detectada.

Quando metade dos observadores de fato percebe a diferença entre duas imagens  $pd = 0,5$ , a probabilidade de resposta correta  $pc$  é igual a  $0,75$ . Tal valor é tomado como base para definir uma JND, parâmetro que mede a percepção de diferenças entre imagens. Quanto maior for o valor absoluto de JND, maior será a percepção das diferenças entre as imagens. Dependendo da necessidade, pode-se definir JND para valores diferentes de  $pd$ , os quais então devem ser

indicados explicitamente. Por exemplo, adotando  $pd = 0,60$ , referencia-se a JND associada por JND60 (60%).

Uma definição formal da JND pressupõe que a percepção de diferenças entre pares de imagens obedece a uma dada distribuição de probabilidade. Para se evitar erros decorrentes dos prolongamentos infinitos das extremidades da distribuição normal, adota-se a função de distribuição de probabilidade angular, definida por

$$p_a(z_a) = \begin{cases} \text{sen}^2\left(\frac{z_a}{\sqrt{2\pi}} + \frac{\pi}{4}\right), & |z_a| \leq \sqrt{\frac{\pi^3}{8}} \\ 0, & z_a < -\sqrt{\frac{\pi^3}{8}} \\ 1, & z_a > \sqrt{\frac{\pi^3}{8}} \end{cases} \quad (11)$$

onde  $z_a$  caracteriza o desvio angular.

Invertendo-se (11), obtém-se

$$z_a(p_a) = \sqrt{2\pi} \left[ \text{sen}^{-1}(\sqrt{p_a}) - \frac{\pi}{4} \right]. \quad (12)$$

A JND de uma dada imagem é definida matematicamente (para o experimento particular de indicação de qual dentre duas imagens é a mais degradada) como a razão entre o desvio angular associado à probabilidade de uma imagem ser indicada como tendo mais degradação e o desvio angular associado à probabilidade da resposta correta (13). Assim,

$$JND = \frac{\text{sen}^{-1}(\sqrt{pp}) - \frac{\pi}{4}}{\text{sen}^{-1}(\sqrt{pc}) - \frac{\pi}{4}} \quad (13)$$

onde  $pp$  é a probabilidade da imagem ser indicada como tendo um maior nível de degradação.

No contexto experimental, a probabilidade  $pp$  é aproximada pela frequência relativa de uma determinada imagem ter sido escolhida pelo usuário como tendo maior nível de degradação. Os resultados apresentados neste trabalho consideram  $pd = 0,50$ .

#### V. RESULTADOS

Para auxiliar no entendimento de como os resultados foram obtidos, será apresentado então um exemplo. A Fig. 2 exhibe o gráfico da JND, tendo como referência a imagem 07 (ver Seção III) do grupo *Barbara* com degradação do tipo ruído impulsivo (*barbara\_ri\_07*). No eixo das abscissas, encontram-se os valores de CC para as imagens deste grupo (*barbara\_ri*) e no eixo das ordenadas, as JNDs correspondentes. A região delimitada pelas duas linhas tracejadas localizadas em  $-1$  JND e em  $1$  JND (ou seja, a região em que  $|JND| \leq 1$ ) especifica aquelas imagens que a maioria dos observadores não conseguiu diferenciar da imagem de referência. O primeiro ponto abaixo da reta  $-1$  JND determina o limiar inferior de percepção e o primeiro ponto acima da reta  $1$  JND define o limiar superior de percepção. Na Tabela I, tem-se os limiares inferior e superior medidos pelo CC, com referência às imagens 07 pertencentes ao grupo das degradadas por ruído impulsivo para diferentes imagens originais. Também são apresentados, ainda na Tabela I, os valores médios  $\mu$  e os desvios-padrão ( $\sigma$ ) para os limiares inferior e superior.

A cada limiar inferior (superior) está associado um limiar diferencial  $\Delta$  inferior (superior). A fim de que as comparações sejam justas, antes dos cálculos dos limiares diferenciais, os

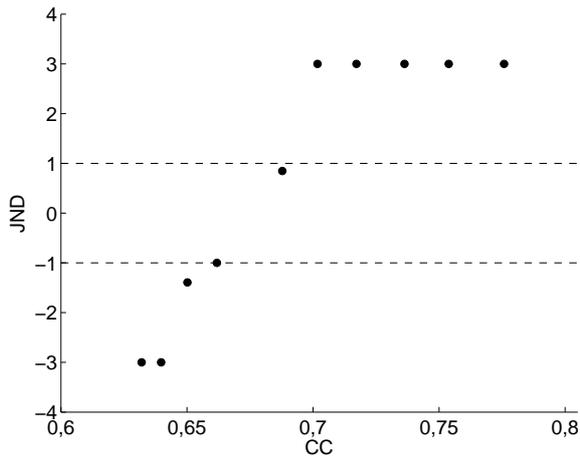


Fig. 2. Gráfico CC x JND para o grupo *barbara\_ri*, com *barbara\_ri\_07* sendo a imagem de referência.

TABELA I

LIMIARES DE PERCEPÇÃO MEDIDOS PELO CC, TENDO COMO REFERÊNCIA AS IMAGENS 07 DEGRADADAS POR RUÍDO IMPULSIVO ( $q_i$ : LIMIAR INFERIOR;  $q_r$ : VALOR DA MÉTRICA PARA IMAGEM 07;  $q_s$ : LIMIAR SUPERIOR;  $\mu$ : MÉDIA;  $\sigma$ : DESVIO-PADRÃO).

Grupo	$q_i$	$q_r$	$q_s$
Avion	0,8565	0,8841	0,9163
Barbara	0,7004	0,7491	0,8036
Boats	0,8114	0,8489	0,8854
Clown	0,7406	0,7634	0,7919
Fruits	0,8068	0,8277	0,8473
House	0,8237	0,8585	0,8719
Isabel	0,8213	0,8385	0,8605
Lena	0,7370	0,7797	0,8030
Mandrill	0,7446	0,7689	0,8233
Peppers	0,7657	0,8056	0,8506
$\mu$	0,7808	0,8125	0,8454
$\sigma$	0,0498	0,0459	0,0401

valores das métricas do conjunto de teste são escalados para permanecerem no intervalo entre 0 e 1. Dessa forma, os limiares diferenciais são definidos por

$$\Delta_i = \hat{q}_i - \hat{q}_r \quad (14)$$

$$\Delta_s = \hat{q}_s - \hat{q}_r \quad (15)$$

onde  $\Delta_i$  e  $\Delta_s$  são os limiares diferenciais inferior e superior, respectivamente;  $\hat{q}_i$  e  $\hat{q}_s$  representam os limiares inferior e superior escalados, respectivamente; e  $\hat{q}_r$  denota o valor escalado da métrica para a imagem de referência (imagem 07, neste exemplo).

O cálculo dos limiares diferenciais é repetido para todas as imagens de referência restantes do grupo das imagens degradadas por ruído impulsivo, resultando em um gráfico semelhante ao da Fig. 3.

Pode-se notar que os limiares diferenciais variam conforme a imagem de referência. Do ponto de vista prático, há algumas possibilidades para a definição de um único valor de limiar diferencial característico com respeito ao tipo de degradação. Por exemplo, se adotada uma política conservadora, considerando o pior cenário, o limiar diferencial inferior (superior)

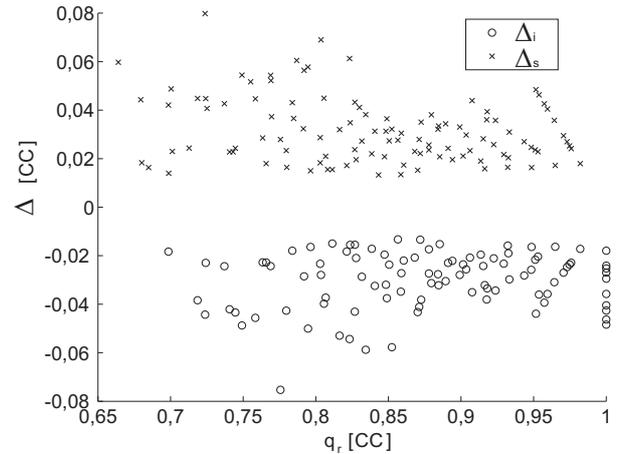


Fig. 3. Limiares diferenciais  $\Delta$  superiores e inferiores (não escalados) do grupo com degradação do tipo ruído impulsivo para diferentes imagens de referência (caracterizadas pelos seus valores de CC,  $q_r$ ).

TABELA II

RESULTADOS PARA O GRUPO DE IMAGENS DEGRADADAS POR RUÍDO IMPULSIVO ( $\mu\Delta_s$ : MÉDIA DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\mu\Delta_i$ : MÉDIA DOS LIMIARES DIFERENCIAIS INFERIORES;  $\sigma\Delta_s$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\sigma\Delta_i$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS INFERIORES;  $\bar{\sigma}$ : MÉDIA DOS DESVIOS-PADRÃO).

Métrica	$\mu\Delta_s$	$\mu\Delta_i$	$\sigma\Delta_s$	$\sigma\Delta_i$	$\bar{\sigma}$
SNR	0,084754	-0,084967	0,082232	0,085713	0,083973
PSNR	0,097876	-0,093721	0,065567	0,067231	0,066399
CC	0,093705	-0,090095	0,039817	0,039676	0,039746
IFC	0,092768	-0,091749	0,116856	0,129136	0,122996
VIF	0,094257	-0,098597	0,163866	0,170302	0,167084
VIFp	0,097555	-0,097195	0,070352	0,073125	0,071739
C4qs	0,057432	-0,100584	0,068670	0,155724	0,112197
C4sim	0,089605	-0,084350	0,118948	0,134379	0,126663
MSSIM	0,101919	-0,096554	0,062208	0,064387	0,063298
MSSIM <sub>l</sub>	0,059501	-0,050373	0,217055	0,199939	0,208497
MSSIM <sub>c</sub>	0,089874	-0,088474	0,044679	0,047422	0,046051
MSSIM <sub>s</sub>	0,094789	-0,093789	0,056462	0,058534	0,057498

poderia ser estabelecido como um valor constante igual ao máximo (mínimo) limiar diferencial inferior (superior) do conjunto de degradação. Neste trabalho, optou-se por calcular a média e o desvio-padrão dos limiares diferenciais ( $\mu\Delta$  e  $\sigma\Delta$ , respectivamente) como recurso para avaliar o comportamento de cada métrica.

Este processo, descrito para o grupo de imagens degradadas por ruído impulsivo, usando o CC para análise, é repetido para os outros tipos de degradação, bem como para todas as métricas testadas. Os resultados finais são apresentados nas Tabelas II, III e IV, para as degradações do tipo ruído impulsivo, *sal e pimenta*, e ruído gaussiano branco aditivo, respectivamente.

Três aspectos principais podem ser discutidos a partir dos dados apresentados nas Tabelas II, III e IV: (a) simetria dos limiares diferenciais inferior e superior; (b) simetria dos desvios-padrão inferior e superior; e (c) comportamento

TABELA III

RESULTADOS PARA O GRUPO DE IMAGENS DEGRADADAS POR RUÍDO DO TIPO *sal e pimenta* ( $\mu\Delta_s$ : MÉDIA DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\mu\Delta_i$ : MÉDIA DOS LIMIARES DIFERENCIAIS INFERIORES;  $\sigma\Delta_s$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\sigma\Delta_i$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS INFERIORES;  $\bar{\sigma}$ : MÉDIA DOS DESVIOS-PADRÃO).

Métrica	$\mu\Delta_s$	$\mu\Delta_i$	$\sigma\Delta_s$	$\sigma\Delta_i$	$\bar{\sigma}$
SNR	0,093686	-0,091974	0,058221	0,058624	0,058423
PSNR	0,107741	-0,104122	0,071683	0,071277	0,071480
CC	0,103830	-0,097412	0,038338	0,032344	0,035341
IFC	0,095335	-0,098891	0,124268	0,129197	0,126732
VIF	0,094264	-0,100028	0,169907	0,177997	0,173952
VIFp	0,100558	-0,100321	0,073552	0,075737	0,074644
C4qs	0,099568	-0,076488	0,114671	0,069852	0,092262
C4sim	0,030192	-0,030087	0,036816	0,038440	0,037628
MSSIM	0,100383	-0,100242	0,067002	0,067862	0,067432
MSSIM <sub>I</sub>	0,048704	-0,043397	0,051413	0,041303	0,046358
MSSIM <sub>C</sub>	0,094583	-0,092674	0,048627	0,047772	0,048200
MSSIM <sub>S</sub>	0,100020	-0,098695	0,058164	0,058399	0,058282

TABELA IV

RESULTADOS PARA O GRUPO DE IMAGENS DEGRADADAS POR RUÍDO GAUSSIANO BRANCO ADITIVO ( $\mu\Delta_s$ : MÉDIA DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\mu\Delta_i$ : MÉDIA DOS LIMIARES DIFERENCIAIS INFERIORES;  $\sigma\Delta_s$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS SUPERIORES;  $\sigma\Delta_i$ : DESVIO-PADRÃO DOS LIMIARES DIFERENCIAIS INFERIORES;  $\bar{\sigma}$ : MÉDIA DOS DESVIOS-PADRÃO).

Métrica	$\mu\Delta_s$	$\mu\Delta_i$	$\sigma\Delta_s$	$\sigma\Delta_i$	$\bar{\sigma}$
SNR	0,166227	-0,262915	0,052407	0,139234	0,095820
PSNR	0,159518	-0,253544	0,047674	0,132825	0,090250
CC	0,126834	-0,077431	0,165440	0,133536	0,149488
IFC	0,165518	-0,313729	0,107736	0,197535	0,152636
VIF	0,157850	-0,197408	0,049199	0,090846	0,070023
VIFp	0,157386	-0,253390	0,046812	0,128224	0,087518
C4qs	0,160708	-0,106389	0,135938	0,122304	0,129121
C4sim	0,152381	-0,157837	0,082059	0,071372	0,076715
MSSIM	0,147362	-0,124834	0,082104	0,072735	0,077419
MSSIM <sub>I</sub>	0,069657	-0,003246	0,140806	0,004726	0,072766
MSSIM <sub>C</sub>	0,131891	-0,087938	0,106525	0,094057	0,100291
MSSIM <sub>S</sub>	0,143559	-0,120995	0,082642	0,071069	0,076855

das métricas para diferentes degradações. Com respeito aos aspectos (a) e (b) mencionados, os dados mostram que há razoável simetria, tanto para o limiar de percepção quanto para os respectivos desvios-padrão, para os conjuntos de teste de imagens degradadas por ruído impulsivo e do tipo *sal e pimenta*. A exceção é a métrica C4qs que apresenta baixa simetria. Contudo, para os dados das imagens degradadas por ruído gaussiano, as simetrias são menores, indicando possivelmente uma não linearidade na percepção de diferenças para esse tipo de degradação. As métricas que apresentam baixa simetria nesse caso são: PSNR, SNR, VIFp e MSSIM<sub>I</sub>.

O comportamento das métricas para as diferentes degradações também é observado. Pode-se notar que, em geral, as métricas estudadas mostram-se regulares. Quatro métricas, no entanto, destoam das demais por apresentarem variações significativas em seus desvios-padrão para as degradações estudadas: CC, VIF, C4sim e MSSIM<sub>C</sub>.

Outro resultado que chamou a atenção decorre da comparação entre o VIF e VIFp. Este último, no tocante às

Tabelas II e III, apresenta melhor caracterização dos limiares diferenciais (menor desvio-padrão). Isso surpreende, pois a VIFp é uma aproximação da VIF, objetivando redução de custo computacional.

## VI. CONCLUSÕES

Neste trabalho, foi realizado um estudo do comportamento de algumas métricas de qualidade visual, aplicadas à caracterização de limiares de detecção de diferenças entre imagens. O conjunto de teste foi composto por diferentes tipos de imagens, degradadas por ruído impulsivo, *sal e pimenta*, e ruído gaussiano branco aditivo. Foram utilizadas métricas de referência completa tanto não perceptuais (SNR, PSNR e CC) quanto perceptuais (MSSIM, IFC, VIF e C4). A partir dos dados experimentais obtidos, foram analisados os aspectos de simetria dos limiares inferior e superior, bem como o comportamento das métricas para os tipos de degradação testados. Como proposta para a continuidade deste trabalho, pretende-se: ampliar o conjunto de degradações testado para incluir resultados referentes a imagens degradadas por má focalização, quantização de coeficientes DCT e quantização de coeficientes *wavelet*; e estimar funções que possam caracterizar melhor a dependência dos limiares inferior e superior em relação à imagem de referência.

## REFERÊNCIAS

- [1] T. N. Pappas, R. J. Safranek, and J. Chen, *Handbook of Image and Video Processing, 2nd ed.* New York: Academic, 2005, ch. Perceptual criteria for image quality evaluation.
- [2] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," *Signal Process.*, vol. 70, pp. 177–200, 1998.
- [3] R. F. Zampolo, D. A. Gomes, and R. Seara, "Avaliação e comparação de métricas de referência completa na caracterização de limiares de detecção em imagens," in *Anais do XXVI Simpósio Brasileiro de Telecomunicações (SBrT 2008)*, Rio de Janeiro, RJ, Set. 2008, artigo no 42803, pp. 1–6.
- [4] P. Le Callet and F. Autrusseau, "Subjective quality assessment ircyn/ivc database," 2005, <http://www.ircyn.ec-nantes.fr/ivcdb/>.
- [5] B. W. Keelan, *Handbook of Image Quality*. Marcel Dekker, 2002.
- [6] ITU-R Recommendation BT.500-11, *Methodology for the Subjective Assessment for the Television Pictures*, Int. Telecomm. Union, Geneva, Switzerland, 2002.
- [7] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment vqeg," on-line, mar 2002. [Online]. Available: [www.vqeg.org](http://www.vqeg.org)
- [8] H. Sheikh, M. Sabir, and A. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- [9] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database," 2006, <http://live.ece.utexas.edu/research/quality>.
- [10] H. Sheikh, A. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.
- [11] H. Sheikh and A. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [12] M. Carnec, P. L. Callet, and D. Barba, "An image quality assessment method based on perception of structural information," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Barcelona, Spain, Sep. 2003, pp. 185–188.