# Aplicativo para caracterização de desempenho de métricas de qualidade visual

Luana Gonçalves, Ronaldo de Freitas Zampolo e Valquíria Gusmão Macedo

***Resumo*— Métricas de avaliação de qualidade visual facilitam a otimização de sistemas de processamento de imagem, se comparadas com o uso exclusivo de avaliação subjetiva. Contudo, tão importante quanto dispor de tais métricas é ser capaz de caracterizar e comparar seu desempenho. Neste trabalho, é apresentado um aplicativo para facilitar a avaliação de desempenho de métricas de qualidade visual, seguindo as recomendações do *International Telecommunication Union* (ITU). Apesar de atualmente ser funcional, o projeto continua em desenvolvimento, devendo em futuro próximo incorporar maior número de métricas, e aprimoramentos na interface.**

***Palavras-Chave*— Qualidade visual, métricas de qualidade visual, avaliação de desempenho de métricas de qualidade visual.**

***Abstract*— Visual quality metrics make easier the optimisation of image processing systems, when comparing with the case when only subjective assessments are taken into account. However, as important as having such metrics available is being able to assess and compare their performances. In this work, we present an application conceived to easily assess the performance of visual quality metrics, according to the International Telecommunication Union (ITU) recommendations. Although the current implementation is already functional, as the project still goes on, more metrics are being integrated as well as some user interface improvements.**

***Keywords*— Visual quality, visual quality metrics, performance assessment of visual quality metrics.**

## I. Introdução

Com a crescente popularização da internet e das câmeras digitais embarcadas em dispositivos móveis, a quantidade de novos vídeos e imagens produzidos diariamente para consumo humano é da ordem de bilhões. Nesse cenário, a busca de estratégias para maximizar a economia de recursos (banda de transmissão e ocupação de memória de armazenamento) com a menor perda de qualidade possível, vem mobilizando os esforços de empresas, bem como de grupos de pesquisa. O desenvolvimento de métricas de qualidade visual (VQM, *visual quality metric*) exerce papel fundamental, pois tais métricas constituem instrumentos de caracterização de qualidade por meio dos quais algoritmos e sistemas de aquisição e processamento podem ser otimizados.

Contudo, tão importante quanto dispor de VQMs, é fundamental ser capaz de avaliar objetivamente o desempenho dessas métricas em diferentes contextos de aplicação.

Luana Gonçalves e Valquíria Gusmão Macedo, Faculdade de Engenharia Elétrica, Universidade Federal do Pará, Belém-PA, Brasil, Emails: luanagoncalves4@live.com, vgmacedo@ufpa.br. Ronaldo de Freitas Zampolo, Faculdade de Engenharia da Computação e Telecomunicações, Universidade Federal do Pará, Belém-PA, Brasil, Email: zampolo@ufpa.br. Este trabalho foi parcialmente financiado pelo CNPq.

Nesse sentido, deve ser mencionado o trabalho do *Video Quality Experts Group* (VQEG), criado em 1997 por um pequeno grupo de especialistas participantes de grupos de estudo dos setores de rádio comunicação e telecomunicações do *International Telecommunication Union* (ITU-R e ITU-T, respectivamente), com o propósito inicial de investigar avaliação visual, tanto no aspecto objetivo (desenvolvimento de métricas) quanto subjetivo (metodologias para coleta de dados experimentais). Atualmente, grande parte dos esforços do VQEG tem sido na validação de novas VQMs com vistas à padronização.

O procedimento padrão para caracterização do desempenho de uma VQM, consolidado atualmente em documentos do VQEG e recomendações do ITU [1], [2], compreende três fases: (a) eliminação de dados estatisticamente raros ou destoantes; (b) definição de função de mapeamento ou de predição; e (c) avaliação do mapeamento ou do preditor. Na primeira etapa, mediante a adoção de algum critério estatístico, são identificados e eliminados das etapas posteriores eventos destoantes do conjunto de dados, os quais poderiam introduzir tendências não representativas do experimento realizado. Na fase de definição de função de mapeamento são otimizados os parâmetros de uma curva de ajuste entre a média do escore de opiniões (MOS, *mean opinion score*) e a VQM sob análise. Tal curva pode ser considerada como um preditor de MOS em função dos valores de VQM para uma imagem ou vídeo em particular. Na fase de avaliação do mapeamento, são considerados os aspectos: acurácia, monotonicidade, e consistência. A acurácia é avaliada pelo cálculo do coeficiente de correlação de Pearson (PCC, *Pearson correlation coefficient*). Por sua vez, a monotonicidade da predição é medida pelo coeficiente de correlação de Spearman (SCC, *Spearman correlation coefficient*). E finalmente, a consistência é dada pela taxa de *outliers*, ou seja razão entre o número de pontos aberrantes e a o número total de pontos associados a cada uma das imagens do banco de dados usadas na avaliação.

O presente trabalho apresenta uma primeira versão de um aplicativo concebido para facilitar a caracterização de desempenho de VQMs, de acordo com as recomendações do ITU. Tal aplicativo pode ser usado tanto para comparar o desempenho de VQMs quanto como recurso de ensino em cursos sobre avaliação de qualidade visual.

O restante deste texto está organizado como segue. Na Seção II, são apresentadas as principais características da versão mais recente do aplicativo proposto, bem como uma breve discussão dos recursos previstos para as próximas versões. Em seguida, na Seção III, são apresentadas as conclusões.

## II. Aplicativo proposto

O aplicativo proposto foi implementado em linguagem Python, visando, sobretudo, rapidez na fase de desenvolvimento, facilidade na manutenção do código fonte e portabilidade. Seu uso é simples, como ilustrado em um possível fluxograma para realizar a avaliação de VQMs pelo aplicativo (Fig. 1). A execução das ações relacionadas no referido fluxograma é facilitada por uma interface gráfica (GUI, *graphical user interface*), onde o usuário especifica os diretórios que armazenam as imagens de referência e de teste, bem como as tabelas de dados obtidas em experimentos subjetivos de avaliação de qualidade visual. É possível também a seleção das VQMs a serem avaliadas, do tipo de mapeamento, e das métricas de avaliação do mapeamento (Fig.2). Pode-se notar que os principais elementos da sistemática de avaliação de VQMs encontram-se presentes na própria concepção da GUI. Na atual versão do aplicativo, são três as VQMs implementadas: PSNR (*peak signal-to-noise ratio*), MSE (*mean square error*) e MSSIM (*mean structural similarity index*) [3], [4].

O aplicativo gera dois tipos de saída: gráficos de dispersão (*scatterplots*) em que as curvas de ajuste entre DMOS (*Difference mean opinion score*) [2], que consiste na diferença entre os MOS das imagens de referência e teste, e VQM são também

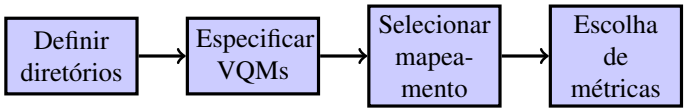


Fig. 1. Possível fluxo de ações para definir uma rodada de avaliação de desempenho de VQMs no aplicativo proposto.

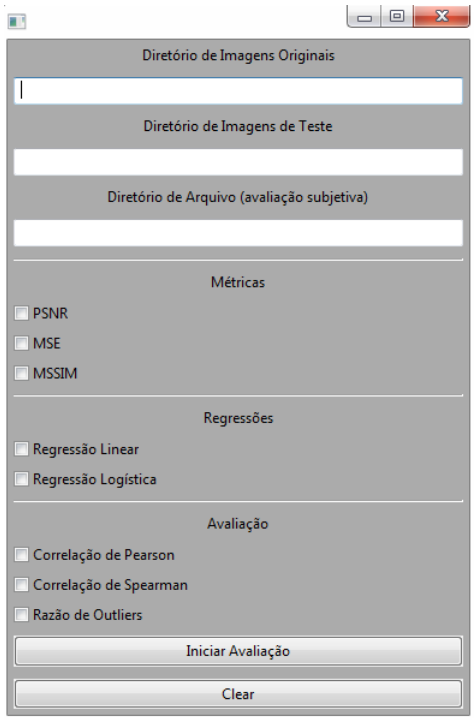


Fig. 2. Interface gráfica do usuário, onde é feita a seleção dos parâmetros da avaliação de desempenho.

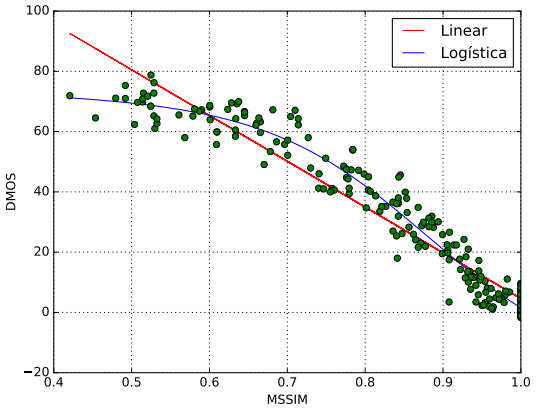


Fig. 3. Exemplo de gráfico de dispersão relacionando DMOS e MSSIM, bem como as curvas de ajuste estimadas.

exibidas (Fig. 3); e relatórios em formato ASCII, contendo as métricas de acurácia, consistência e de pontos aberrantes, bem como os dados de identificação e parâmetros da avaliação realizada.

### A. Pendências e acréscimo de funcionalidades

No sentido de maior aderência às recomendações do ITU, falta implementar a etapa preliminar de eliminação de dados estatisticamente destoantes. Um conjunto maior de VQMs está em processo de integração ao aplicativo a partir do *MeTriX MuX Visual Quality Assessment Package* disponível no endereço *http://foulard.ece.cornell.edu/gaubatz/metrix_mux/*. O projeto é mantido em repositório online[1], onde os interessados poderão obter a versão estável mais recente, contribuir com sugestões ou mesmo colaborar no desenvolvimento do aplicativo.

## III. Conclusões

Este trabalho apresentou a proposição de um aplicativo em Python para facilitar a avaliação de desempenho de métricas de qualidade de visual de acordo com as recomendações do ITU. As principais características do aplicativo foram comentadas, bem como as pendências imediatas e inclusão de funcionalidades atualmente em curso.

## Agradecimentos

Os autores agradecem ao CNPq e à UFPA, que financiaram parcialmente este projeto através do Programa de Iniciação Científica (PIBIC).

## Referências

[1] VQEG. VQEG Final Report of FR-TV Phase II Validation Test. on-line, agosto 2003.
[2] ITU-T. Tutorial – Objective Perceptual Assessment of Video Quality: Full Reference Television, 2004.
[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. 13(4):600–612, April 2004.
[4] Zou Wang and Allan C. Bovik. *Modern Image Quality Assessment*. Morgan & Claypool, New York, 2006.

[1] https://github.com/luaffjk/AQC.GUI