

AVALIAÇÃO DE QUALIDADE DE IMAGEM 3D: UM ESTUDO USANDO ÁRVORE DE DECISÃO

LUCAS SEIDY R. S. IKENOUE¹; NARÚSCI DOS S. BASTOS²; BRUNO ZATT³

¹Universidade Federal de Pelotas - lsrdsikenoue@inf.ufpel.edu.br

²Universidade Federal de Pelotas - naruscibastos@gmail.com

³ Universidade Federal de Pelotas - zatt@inf.ufpel.edu.br

1. INTRODUÇÃO

Nos últimos anos pode-se observar um crescimento significativo do compartilhamento de imagens e vídeos digitais facilitados pelas mídias sociais e serviços de streaming. Neste contexto, as tecnologias de codificação de imagens e vídeos passaram a ser de uso indispensável. Como o processo de codificação e transmissão pode adicionar degradações à qualidade percebida pelo usuário final, vem sendo cada vez mais importante desenvolver métodos e técnicas que viabilizam medir a qualidade percebida pelo usuário. Para isso, a Avaliação de Qualidade de Imagem (IQA) e Vídeo (VQA) permite melhorias na qualidade final por meio de métricas e métodos para medir a qualidade de imagens e vídeos que recebem degradações (REGIS, 2013) e (RODRIGUEZ, 2014). No entanto, o estudo de BANITALEBI-DEHKORDI et al. (2016) observa que a maioria destas métricas são desenvolvidas para conteúdos 2D, deixando um amplo espaço para melhorias sobre a avaliação, precisão e confiabilidade na avaliação de qualidade voltada às imagens e vídeos 3D.

A IQA é disposta em duas classes: subjetiva e objetiva. A avaliação subjetiva é baseada na percepção humana, dada através do julgamento do observador. A métrica de avaliação mais utilizada é o *Mean Opinion Score* (MOS), classificada numa escala de cinco pontuações diferentes (5 - melhor qualidade e 1 - pior qualidade) (RODRIGUEZ, 2014). Já os modelos objetivos utilizam algoritmos com base em modelos estatísticos ou matemáticos, cujo objetivo é obter uma estimativa de qualidade o mais próxima possível da percepção visual humana.

Apesar de ser considerado mais adequado para avaliar a qualidade de imagem e vídeo, modelos subjetivos carecem da disponibilidade de recursos humanos, o que torna a sua aplicação custosa, lenta e, muitas vezes, inviável (WANG, 2004) e (WANG, 2003). Porém, estes modelos são essenciais, pois funcionam como referência para validar a capacidade dos modelos objetivos. Os valores das predições objetivas são comparados em termos de proximidade aos valores dos modelos subjetivos, como do valor MOS, sendo que uma maior correlação entre esses valores indica uma métrica objetiva confiável para predizer a qualidade da imagem (Wang, 2015). Devido a complexidade das informações de imagens 3D, a utilização de paradigmas de Aprendizado de Máquina (AM) podem ser uma boa prática já que de acordo com GASTALDO; REDI (2012), permitem lidar com a tarefa de IQA sobre uma perspectiva diferente das métricas tradicionais, pois têm como objetivo final “imitar” a percepção humana. Para isso comparamos a capacidade do modelo gerado pelo algoritmo J48 de predizer o MOS utilizando métricas de medidas de erro.

2. METODOLOGIA

Este trabalho apresenta um estudo comparativo entre os modelos gerados pelo algoritmo de classificação J48 baseado em AM para diferentes classes (MOS) para a avaliação de qualidade de imagem 3D de referência completa. Sendo assim, adotaram-se as seguintes etapas do processo de AM como metodologia para o desenvolvimento deste trabalho: (1) Aquisição dos dados; (2) Pré-processamento dos dados; (3) Treinamento e Teste; (4) Cálculo de métricas.

1. Aquisição dos dados: os dados utilizados foram extraídos da IVC 3D Image Quality Database (WANG, 2015, WATERLOO, 2015), que contém um conjunto de 790 imagens que apresentam degradações simétricas e assimétricas entre as vistas. Além disso, disponibiliza valores de MOS dos testes subjetivos das mesmas. Foram extraídas e armazenadas através de vetores diferentes características correspondendo a sua imagem e rotulados com o valor do MOS e então utilizadas como entrada para o treinamento do modelo.
2. Pré-processamento: esta etapa segue os passos detalhados a seguir.
 - (a) Limpeza dos dados: foram removidas as informações consideradas irrelevantes para o processo;
 - (b) Transformação dos dados: os valores de MOS fornecidos pela base dados variam em um intervalo de 1 a 100. Para o treino e teste uma faixa de representação foi atribuída para cada variação de classes, correspondendo à:

$$(1) \quad MOS = x \frac{100}{n} - \frac{n}{2}$$

onde (x) representa o MOS subjetivo e (n) o número de classes.

- (c) Separação dos dados: os dados foram divididos em dois conjuntos, sendo 80% para o treinamento e validação (usando validação cruzada) e 20% para o teste;
 - (d) Adequação dos dados: para o formato ARFF (*Attribute Relation File Format*), utilizado no software Weka, ferramenta utilizada para treino e teste dos modelos.
3. Treinamento e teste: Foram usadas 16 imagens estereoscópicas originais para o treinamento e teste, com o total de 790 pares estereoscópicos degradados. Do conjunto de dados, 80% foram separados para o treinamento e validação e 20% para o teste. Aplicou-se o método de validação cruzada k-fold com 10 folds para classificador j48. Além disso, os modelos foram treinados e testados com diferentes classes, sendo elas as classes: MOS 5, 10 e 25.
 4. Cálculo de Métricas de Erro: para a análise dos dados, os resultados obtidos através dos modelos gerados pelo algoritmo j48 foram calculados através das métricas de Erro Médio Absoluto (MAE) e Raiz Quadrada do Erro Médio (RMSE).

$$(2) \quad MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (3) \quad RMSE = \sqrt{\sum_{i=1}^n \frac{(|Y_i - \hat{Y}_i|)^2}{n}}$$

As Equações (1) e (2), correspondem respectivamente as métricas de erros MAE e RMSE, que indicam o valor esperado (Y_i), valor predito ($\hat{Y}_i$) e número de entradas (n). O valor MAE indica que a obtenção de valores próximos de zero, corresponde a uma melhor predição do modelo (ACHARYA, 2021). Já o RMSE, indica a variação média entre um resultado predito (gerado a partir do modelo) e o seu valor real (ACHARYA, 2021).

3. RESULTADOS E DISCUSSÃO

Os resultados obtidos para os modelos gerados pelo algoritmo j48 são apresentados na Tabela I, onde pode-se observar o comportamento dos valores de MAE e RMSE para cada modelo gerado com as diferentes classes.

Tabela I. Valores de MAE e RMSE para os modelos gerados pelo algoritmo j48 com 5,10 e 25 classes.

Classes	MAE	RMSE
5	7.020	8.621
10	6.062	8.495
25	7.407	10.476

O modelo gerado para 5 classes apresenta um MAE de 7.020 e RMSE com valor de 8.621. Já o modelo gerado para 10 classes apresenta o MAE com valor de 6.062 e RMSE de 8.495. O modelo gerado com 10 classes tem o valor da medida de erro um pouco menor do que o modelo com 5 classes. No entanto, modelo gerado para 25 classes, apresenta um MAE de 7.407 e RMSE de 10.476, esses valores são significativamente maiores que os modelos gerados para 5 e 10 classes. Sendo assim, observa-se que a variação de classe leva ao entendimento de que os valores de erro MAE e RMSE diminuem à medida que o número de classes aumenta, esse aumento representa uma porcentagem menor de erros e também uma distância menor entre o MOS predito e MOS real no caso de erro. No entanto, é possível notar que para o modelo gerado com 25 classes os valores de erro aumentam significativamente. Nesta análise pode-se observar que os modelos gerados pelo classificador j48 apresentam um comportamento em forma de "U" quando suas classes são variadas em 5,10 e 25. Este comportamento pode sugerir que os erros tendem a serem maiores quando as classes variam para os extremos.

4. CONCLUSÕES

Este trabalho apresentou um estudo de avaliação de qualidade de imagem 3D de referência completa utilizando o algoritmo J48. Em que foram comparados os resultados das métricas de erros diante a variação de três classes (5,10 e 25) utilizadas para o treinamento e teste do modelo.

A utilização de variações do número de classes para o classificador j48 associado às métricas de erro demonstram que o modelo gerado com 10 classes possui um resultado melhor em termos de MAE e RMSE para a predição do MOS real. Possivelmente porque os erros tendem a diminuir quando as classes são

melhor distribuídas. Ainda, como trabalhos futuros pretende-se realizar análises com técnicas de AM mais robustas e ampliar os estudos para vídeos 3D.

5. REFERÊNCIAS BIBLIOGRÁFICAS

REGIS, C.D.M. **Métrica de Avaliação Objetiva de Vídeo Usando a Informação Espacial, a Temporal ea Disparidade**. 2013. Dissertação (Ph. D.), Universidade Federal de Campina Grande–UFCG.

ZEGARRA, D.R. **Proposta da métrica eVSQM para avaliação de QoE no serviço de streaming de vídeo sobre TCP**. 2014. Dissertação (Ph. D.), Universidade de São Paulo

BANITALEBI-DEHKORDI, A.; POURAZAD, M.T.; NASIOPOULOS, P. An efficient human visual system based quality metric for 3D video. In: **Multimedia Tools and Applications** 75, 2016, p. 4187–4215.

WANG, Z.; LU, L.; BOVIK, A.C. Video quality assessment based on structural distortion measurement. **Signal processing: Image communication** 19, 2 (2004), 121–132.

WANG, Z.; SHEIKH H.R.; BOVIK, A.C. Objective video quality assessment. **The handbook of video databases: design and applications**, 41, 2003, p. 1041–1078.

WANG, J.; REHMAN, A.; WANG, K.Z.; WANG, S.; WANG, Z. Quality prediction of asymmetrically distorted stereoscopic 3D images. IEEE, 2015, **Transactions on Image Processing** vol. 24, no. 11, 2015, p. 3400–3414.

GASTALDO P.; REDI, J.A. **Machine learning solutions for objective visual quality assessment**. 12, 2012.

ACHARYA, S. **What are RMSE and MAE?**. Towards Data Science, 14 mai. 2021. Acessado em 21 ago. 2022. Disponível em: <https://towardsdatascience.com/what-are-rmse-and-mae-e405ce230383>