

Análise de Qualidade de Vídeo - Conclusões

Nome do Teste: teste

Data: 2025-12-29 12:56:17

Análise Detalhada dos Dados de Qualidade de Vídeo

Este relatório apresenta uma análise detalhada dos dados de um teste subjetivo de qualidade de vídeo, combinando métricas objetivas (PSNR, SSIM) com métricas subjetivas (MOS - Mean Opinion Score).

1. Análise Geral dos Resultados

A primeira observação dos dados revela uma inconsistência significativa nos resultados subjetivos (MOS) que impacta diretamente a interpretação das correlações e a qualidade dos modelos de regressão.

- Distribuição do MOS:** A grande maioria dos vídeos (9 de 10) recebeu um MOS de 5.0, enquanto apenas um vídeo (`H265_500kbps.mp4`) obteve um MOS de 7.0. Essa falta de variância no MOS para a maioria dos vídeos é um problema crítico. Em um teste subjetivo bem-sucedido, espera-se uma distribuição mais ampla dos scores, refletindo as diferentes qualidades percebidas. Um MOS constante de 5.0 para vídeos com PSNR e SSIM amplamente diferentes sugere que os avaliadores podem ter tido dificuldade em discernir diferenças de qualidade para a maioria dos vídeos, ou que a escala de pontuação foi utilizada de forma restrita, ou que houve uma "saturação" na percepção da qualidade em um certo nível.
- O Caso do `H265_500kbps.mp4`:** Este vídeo é um ponto fora da curva notável. Apesar de apresentar MOS 7.0 (o mais alto), suas métricas objetivas (PSNR de 27.43 dB e SSIM de 0.856) são de qualidade *mediana ou até inferior* quando comparadas a outros vídeos que receberam MOS 5.0. Por exemplo:
 - `H264_8000kbps.mp4` tem PSNR 38.56 dB e SSIM 0.959 (significativamente melhores objetivamente), mas MOS 5.0.
 - `H264_500kbps.mp4` tem PSNR 22.67 dB e SSIM 0.748 (piores objetivamente), mas também MOS 5.0.

Esta contradição flagrante entre as métricas objetivas e o MOS de 7.0 para este vídeo específico é o principal fator que distorce todas as análises subsequentes.

- **Métricas Objetivas:** As métricas objetivas (PSNR e SSIM) demonstram uma variância esperada, com valores geralmente aumentando com o bitrate e, em alguns casos, com a eficiência do codec (H265 vs H264). O `H265_8000kbps.mp4` apresenta as melhores métricas objetivas (PSNR 40.00 dB, SSIM 0.966), o que é esperado para um vídeo de alta qualidade/bitrate, mas ironicamente recebeu um MOS de 5.0, o mesmo que o `H264_500kbps.mp4` com as piores métricas objetivas (PSNR 22.67 dB, SSIM 0.748).

2. Interpretação das Correlações entre Métricas Objetivas e Subjetivas

As correlações entre as métricas objetivas (PSNR, SSIM) e a métrica subjetiva (MOS) são apresentadas como Pearson e Spearman.

- **PSNR vs MOS:**

- Pearson: -0.335
- Spearman: -0.290

- **SSIM vs MOS:**

- Pearson: -0.236
- Spearman: -0.290

Interpretação:

1. **Sinal Negativo:** Ambas as métricas objetivas apresentam uma correlação *negativa* com o MOS. Isso é altamente problemático e contra-intuitivo. Em testes de qualidade de vídeo, espera-se uma correlação *positiva*, ou seja, quanto maior a qualidade objetiva (PSNR, SSIM), maior a qualidade percebida (MOS). O sinal negativo indica que, conforme as métricas objetivas melhoraram, o MOS tende a diminuir, o que é um resultado inverso ao esperado.

2. **Magnitude Fraca:** Além do sinal incorreto, os valores absolutos das correlações (próximos de 0.3) indicam uma correlação *muito fraca*. Isso significa que há pouca ou nenhuma relação linear (Pearson) ou monotônica (Spearman) entre as métricas objetivas e o MOS neste conjunto de dados.

Causa Provável: A causa mais provável para estas correlações negativas e fracas é a anomalia do MOS 7.0 para o vídeo H265_500kbps.mp4, combinado com a falta de variância nos demais MOS (todos 5.0). O vídeo com MOS 7.0 tem PSNR e SSIM relativamente baixos, enquanto os vídeos com PSNR e SSIM altos recebem MOS 5.0. Este ponto discrepante "puxa" a correlação para o negativo, obscurecendo qualquer relação real que possa existir.

3. Avaliação da Qualidade dos Modelos de Regressão

Os modelos de regressão tentam prever o MOS a partir do PSNR e do SSIM.

- **PSNR → MOS:**

- Linear: `MOS = -0.036 × PSNR + 6.403`

- R^2 Linear: 0.113

- **SSIM → MOS:**

- Linear: `MOS = -2.126 × SSIM + 7.120`

- R^2 Linear: 0.056

Avaliação:

1. **Coeficientes Negativos:** Assim como nas correlações, os coeficientes das variáveis independentes (PSNR e SSIM) nos modelos lineares são negativos. Isso significa que, de acordo com esses modelos, um aumento na qualidade objetiva (PSNR ou SSIM) resultaria em uma *diminuição* no MOS previsto, o que é conceitualmente incorreto para a previsão da qualidade percebida.

2. R^2 Extremamente Baixo:

- R^2 de 0.113 para PSNR significa que apenas 11.3% da variância no MOS pode ser explicada pelo PSNR.
- R^2 de 0.056 para SSIM significa que apenas 5.6% da variância no MOS pode ser explicada pelo SSIM.

Esses valores de R^2 são **extremamente baixos**. Um modelo de regressão útil para previsão da qualidade percebida geralmente busca um R^2 significativamente maior (comumente acima de 0.70 ou 0.80 para uma boa performance preditiva). Valores tão baixos indicam que os modelos lineares são muito ruins em prever o MOS a partir das métricas objetivas neste dataset.

3. **Modelos Polinomiais:** Embora as equações polinomiais de grau 2 sejam fornecidas, o R^2 para estes modelos não foi apresentado. No entanto, dada a fraqueza dos modelos lineares e a natureza dos dados (principalmente o MOS anômalo e a falta de variância), é improvável que os modelos polinomiais ofereçam uma melhoria substancial ou representem uma relação mais precisa. Os gráficos (PSNR vs MOS, SSIM vs MOS) confirmam visualmente a falta de um padrão claro e a presença do outlier, tornando qualquer linha de tendência (linear ou polinomial) um ajuste pobre aos dados.

Em resumo, os modelos de regressão construídos a partir deste dataset são de **péssima qualidade preditiva** e apresentam um relacionamento inverso ao esperado.

4. Conclusões sobre Que Métricas Objetivas Melhor Preveem a Qualidade Percebida

Com base na análise dos dados fornecidos:

- **Nenhuma Métrica Prediz Bem:** Neste teste específico e com os dados disponíveis, **nem o PSNR nem o SSIM demonstraram ser bons preditores da qualidade percebida (MOS)**. Ambas as métricas apresentaram correlações negativas e fracas com o MOS, além de modelos de regressão com R^2 extremamente baixos, indicando que explicam muito pouco da variância

do MOS e predizem uma relação inversa à esperada.

• **Limitação Crítica do Dataset:** É crucial enfatizar que esta conclusão é uma **consequência direta das anomalias e da baixa variância nos dados subjetivos (MOS)**, e não uma afirmação geral sobre a validade do PSNR ou SSIM como métricas de qualidade de vídeo. A presença de um MOS discrepante (7.0 para `H265_500kbps.mp4` com métricas objetivas medianas) e a aglomeração da maioria dos outros vídeos em MOS 5.0 (independentemente de sua qualidade objetiva) tornam impossível para qualquer métrica objetiva correlacionar-se de forma significativa e correta com o MOS.

Recomendações para Futuros Testes:

Para obter insights válidos sobre a relação entre métricas objetivas e subjetivas, seria essencial:

1. Reavaliar o Teste Subjetivo: Investigar a metodologia do teste subjetivo para entender por que a maioria dos vídeos recebeu MOS 5.0 e por que H265_500kbps.mp4 recebeu 7.0. Isso pode envolver:

- Revisar as instruções aos avaliadores.
- Garantir um número suficiente e representativo de avaliadores.
- Verificar a calibração da escala de pontuação (e.g., se 5.0 era o máximo efetivamente usado pela maioria dos participantes para "qualidade boa", e 7.0 sugere uma escala diferente).
- Incluir vídeos de referência (sem compressão ou com compressão mínima) e vídeos com degradação óbvia para garantir uma gama completa de MOS scores.

2. Coletar Mais Dados Subjetivos: Um maior número de avaliações e uma gama mais ampla de condições de vídeo (bitrates, codecs, tipos de conteúdo) ajudariam a obter uma distribuição mais representativa dos scores MOS.

3. Filtrar Outliers: Considerar a análise de outliers nos dados MOS, como o ponto 7.0, para entender sua causa e, se for um erro de medição, removê-lo ou corrigi-lo.

Sem um conjunto de dados MOS mais consistente e com variância adequada, qualquer tentativa de correlacionar métricas objetivas ou construir modelos preditivos será ineficaz.