Lightweight NR Metrics

Kumar Awanish
Quality and Usability Lab (TU Berlin)
MOTIVATION

❖ Increase in popularity of Gaming videos and many vendors such as Twitch.tv, YouTube Gaming, Hitbox.tv.
❖ Due to advancement in hardwares and software, games are getting more complex.
❖ Gaming videos consist of synthetic and artificial content.
❖ More attention for Machine-learning based quality evaluation methods.
Quality Assessment (QA) Metrics

- Full Reference:
  - uses a complete reference signal information.

- Reduced Reference:
  - uses a part of the reference signal.

- No Reference:
  - does not use a reference signal.
Traditional NR metrics like BRISQUE, NIQE failed to predict gaming content. Dataset used: GamingVideoSET.
Existing NR Metrics for Gaming content: NR-GVQM

- Use Frame-level features and model with VMAF.
- Pre-Trained BRISQUE, NIQE score.
- Only GamingVideoSET data for model development.
Existing NR Metrics for Gaming content: NOFU

- Uses MOS score of 90 videos from GamingVideoSET.
- Temporal pooling approach before feeding to ML model.
- Lacks validation set.
Existing NR Metrics for Gaming content: NR-GVSQI

- Uses GamingVideoSET and KUGVD dataset.
- Proper training and validation.
- Uses pre trained BRISQUE, NIQE.
Why need new NR metrics for gaming content!!

- Traditional NR metrics didn’t able to predict the quality of Gaming content with high performance.

- Lack of Training and Validation support.

- Performance of traditional metrics like NIQE, BRISQUE haven’t checked on training for gaming based contents.

- Lack of Lightweight NR gaming metrics.
Proposed Solution

Stage 1

GVSET
Read Images
KUGVD
Feature Extraction
Per Frame values
432000 images with NdNetGaming Score
NDNetGaming Score

Stage 2

Read Videos
180 videos with MOS score (KUGVD+GVSET)
Subjective Score (MOS)

Use of Pretrained Model from Stage 1 for Feature extraction at video level

Feature Extraction
Video level values
Data Preprocessing
Video level features

Feature Selection
Frame Features
ML Model based on NdNetGaming Score.

Feature Selection
Video Features
ML Model based on MOS score.
Proposed Solution: Stage 1

❖ Focus on Spatial aspect of the Video Data.

❖ Feature Extraction at Frame level:
  ➢ BRISQUE Feature:
    ■ Total of 36 features extracted.
    ■ Retain the BRISQUE model on gaming content.
    ■ Find presence of distortion.
  ➢ Histogram of Oriented Gradients (HOG) Features:
    ■ Total of 36 features extracted.
    ■ Metrics for texture descriptor i.e. edge detection.
  ➢ Grey Level Co-occurrence Matrix (GLCM) Features:
    ■ Total of 4 features extracted.
    ■ Metrics for texture analysis.

❖ Data Processing: Finding Outliers.
Feature Selection and Modelling

- **TrainSet**: GamingVideoSET with 351000 frames.
- **TestSet**: KUGVD with 81000 frames.
- **Label**: NdNetGaming.
- **ML Algorithm**: XgBoost Regressor, SVR.
- **Best selected model saved to use in Stage 2.**

<table>
<thead>
<tr>
<th>Features</th>
<th>PLCC</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>0.82848</td>
<td>0.47848</td>
</tr>
<tr>
<td>F2</td>
<td>-0.51723</td>
<td>0.98857</td>
</tr>
<tr>
<td>F3</td>
<td>-0.57449</td>
<td>0.98529</td>
</tr>
<tr>
<td>F1+F2</td>
<td>0.50244</td>
<td>0.73867</td>
</tr>
<tr>
<td>F1+F3</td>
<td>0.90571</td>
<td>0.36214</td>
</tr>
<tr>
<td>F1+F2+F3</td>
<td>0.90764</td>
<td>0.35861</td>
</tr>
</tbody>
</table>

F1: BRISQUE, F2: HOG, F3: GLCM
Proposed Solution: Stage 1

Per Frame Result:
- TestSet: KUGVD data with NdNetgaming Score Per Frame
- SROCC: 0.967
- PLCC is: 0.968
- RMSE: 0.064
Proposed Solution: Stage 1

Video Level Result:

- TestSet: KUGVD data with pooled NdNetgaming Score.
- SROCC: 0.871
- PLCC is: 0.842
- RMSE: 0.321
Proposed Solution: Stage 2

❖ Focus on Temporal aspect of the Video Data.

❖ Feature Extraction at Video level:
  ■ Absolute Motion using block Motion.
  ■ Temporal Information(TI)
  ■ Trained model from Stage1 as an input.

❖ Data Processing: Finding Outliers.
Feature Extraction at Video Level

- Selected based on F score.
- F score is measure for feature selection.
- Features notation:
  - f0: Motion Vector
  - f1: Predicted pooled score
  - f2: TI

![Feature importance chart]

- Feature f1 has the highest F score.
- Feature f2 has the lowest F score.

Note: The chart shows the feature importance with f1 having the highest score (1499) and f2 having the lowest score (111).
Scatter plot of MOS scores

- Trained on MOS values of 90 videos from GVSET.
- Tested on KUGVD data with 90 MOS values.
- All the games that we have subjective results are excluded for training part of stage1.
Scatter plot of MOS scores

- Trained on MOS values of 90 videos from KUGVD.
- Tested on GVSET data with 90 MOS values.
- All the games that we have subjective results are excluded for training part of stage1.
Result

<table>
<thead>
<tr>
<th>NR Metrics</th>
<th>GVSET</th>
<th></th>
<th>KUGVD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCC</td>
<td>SROCC</td>
<td>PCC</td>
<td>SROCC</td>
</tr>
<tr>
<td>BRISQUE</td>
<td>-0.44</td>
<td>-0.46</td>
<td>-0.62</td>
<td>-0.60</td>
</tr>
<tr>
<td>NIQE</td>
<td>-0.72</td>
<td>-0.71</td>
<td>-0.85</td>
<td>-0.84</td>
</tr>
<tr>
<td>NR-GVQM</td>
<td>0.89</td>
<td>0.87</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>NR-GVSQI</td>
<td>0.87</td>
<td>0.86</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>NOFU</td>
<td>0.91</td>
<td>0.91</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LightweightNR</td>
<td>0.93</td>
<td>0.94</td>
<td>0.90</td>
<td>0.91</td>
</tr>
</tbody>
</table>
Conclusion

- Training BRISQUE on gaming content enhances the performance of model.

- Two steps model development helped in robust model.

- Proposed model is lightweight and can be used in real time.

- Designed machine learning based NR metrics have a high correlation with subjective (MOS) score.
Reference


Thank You !!