Video Assessment of TemporaL Artifacts and Stalls (Video ATLAS) Software release.

=================================================================

-----------COPYRIGHT NOTICE STARTS WITH THIS LINE------------

Copyright (c) 2017 The University of Texas at Austin

All rights reserved.

Permission is hereby granted, without written agreement and without license or royalty fees, to use, copy, modify, and distribute this code (the source files) and its documentation for any purpose, provided that the copyright notice in its entirety appear in all copies of this code, and the original source of this code, Laboratory for Image and Video Engineering (LIVE, <http://live.ece.utexas.edu>) and Center for Perceptual Systems (CPS, http://www.cps.utexas.edu) at the University of Texas at Austin (UT Austin, http://www.utexas.edu), is acknowledged in any publication that reports research using this code. The research is to be cited in the bibliography as:

1) C. G. Bampis and A. C. Bovik, "Video ATLAS Software Release"

URL: http://live.ece.utexas.edu/research/quality/VideoATLAS\_release.zip, 2017

2) C. G. Bampis and A. C. Bovik, "Learning to Predict Streaming Video QoE: Distortions, Rebuffering and Memory," Transactions on Image Processing, revised

IN NO EVENT SHALL THE UNIVERSITY OF TEXAS AT AUSTIN BE LIABLE TO ANY PARTY FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OF THIS SOFTWARE AND ITS DOCUMENTATION, EVEN IF THE UNIVERSITY OF TEXAS AT AUSTIN HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE. THE UNIVERSITY OF TEXAS AT AUSTIN SPECIFICALLY DISCLAIMS ANY WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE SOFTWARE PROVIDED HEREUNDER IS ON AN "AS IS" BASIS, AND THE UNIVERSITY OF TEXAS AT AUSTIN HAS NO OBLIGATION TO PROVIDE MAINTENANCE, SUPPORT, UPDATES, ENHANCEMENTS, OR MODIFICATIONS.

-----------COPYRIGHT NOTICE ENDS WITH THIS LINE------------%

Author : Christos Bampis

Version : 3.0

The authors are with the Laboratory for Image and Video Engineering (LIVE), Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX.

Kindly report any suggestions or corrections to cbampis@gmail.com

=================================================================

Please note that the proposed Video ATLAS model uses the following (as described in the paper):

* all 5 features: VQA, R1, R2, M and I
* ST-RRED as the VQA feature
* mean pooling for the VQA feature
* SVR as the learning engine

However, if needed, this software release allows for deploying other Video ATLAS variants, e.g. using a NR method such as NIQE or a Random Forest regressor. For the rest of this readme, it is assumed that Video ATLAS is flexible enough to allow for the use a variety of features, quality models, pooling types and regression engines.

Video ATLAS is a regression engine designed to perform retrospective Quality of Experience (QoE) predictions. Unlike other QoE predictors, Video ATLAS relies on a combination of objective quality models (such as SSIM or STRRED), rebuffering-aware global statistics and memory features to predict QoE. There are several input features for each video sequence:

1. Pooled frame quality scores that characterize the objective video quality when there is no playback interruption. Any pooling scheme (mean, harmonic average, hysteresis, VQpooling) can be used as long as it can collapse the frame quality scores to a single VQA feature for each sequence.

For the rebuffered video sequences, extra care is needed. To compute full reference (FR) VQA, one must first align the two video sequences by removing the time intervals of rebuffering. Otherwise, the reference and the distorted videos are not synchronized and the FR computations will be inaccurate. The LIVE-NFLX Database release contains the “\_no\_stall” 4:2:0 YUV files for the publicly available video sequences.

1. Following the simple notion that the relative amount of time that a video is more heavily distorted is directly related to the overall QoE, Video ATLAS also computes the time (in seconds) that the video is encoded to a value lower than the maximum one. For streaming applications, the bitrate ladder is fixed and known a priori; hence this time interval can be easily measured.
2. The Video ATLAS framework also uses global rebuffering statistics as input. For retrospective quality evaluations, it is has been shown that the number and the duration of rebuffering events are simple and very efficient predictors of the QoE prediction. Video ATLAS is versatile enough to allow for deploying any other such features depending on the application.
3. Memory is probably one of the most important aspects of subjective QoE. In our subjective experiments, we have verified strong recency effects: humans tend to evaluate their QoE based on their latest experiences. Video ATLAS models this memory effect by introducing the “time since last impairment” feature. Assuming that the positions of rebuffering and bitrate changes are known a priori, Video ATLAS computes the time between the last impairment took place and the time the video finishes (and the subject is asked to give an overall QoE score).

Video ATLAS allows the use of any regression model (Support Vector Regression, Random Forest, Ridge Regression) and any objective quality model (Full- or No-Reference) such as NIQE, STRRED, VMAF, MS-SSIM or SSIM. The current release contains pre-computed NR and FR VQA models on the LIVE-Netflix Dataset as well as the subjective data needed for validating the framework.

Video ATLAS is made publicly available as a demo here. The LIVE-Netflix Video QoE Database features are stored and made available through this release. This version contains only the LIVE-Netflix features, the Waterloo DB features will be added in the next version.

This release contains pre-trained models (using the whole LIVE-Netflix Video QoE Database). Each of these models has the following name convention:

[regressor\_type]\_[quality\_model]\_[pooling\_type]

Note that for testing this model the same quality model has to be used. For example, if you want to apply Video ATLAS on your data and have computed SSIM (and the other features), you should pick a pre-trained model that was made using SSIM. Again, the pooling type should match that of Video ATLAS.

It is possible to create your own models by re-training on the LIVE-Netflix Video QoE Database using the “GeneratePreTrainedModels.py” script. You only have to compute the frame quality scores and perform some pooling on those quality scores.

You can also find the pre-computed frame quality scores for several models (SSIM, PSNR, MS-SSIM, NIQE, VMAF, STRRED) in the LIVE\_NFLX\_PublicData\_VideoATLAS\_Release folder, indicated by [quality\_model]\_vec (one for each of the 112 videos in the dataset). See the following description for more details on the files that are available in this release.

***Please note that when testing on the LIVE-Netflix Video QoE Database, the results will be much higher since the pretrained models are trained on the whole dataset.*** Therefore, if you want to train/test on a subset of the data, you should use the train/test splits using the pre-generated train/test indices.

=================================================================

Details about the files in this release:

1. “PretrainedModels” folder: contains ready-to-go models for QoE prediction using Video ATLAS, trained on the LIVE-Netflix Video QoE Database. You need Python to load the pickle files and test the model. Please make sure you load the regressor by doing: regressor.best\_estimator\_.predict(your\_test\_data)
2. “TrainingMatrix\_LIVENetflix\_1000\_trials.mat”: contains 1000 pre-generated random 80% train and 20% test content splits. A value of 1 in the (i,j ) element of the matrix indicatess that the ith video sequence is train on the jth trial (i=1…112, j=1…1000).
3. “TrainingMatrix\_Waterloo\_1000\_trials.mat”: same as above for the Waterloo DB. Useful for benchmarking results in this database too.
4. “VideoATLAS.py”: demo script for Video ATLAS implementation. Can use one of the pre-trained models or train and test on the LIVE-Netflix dataset using the above pre-generated train/test splits.
5. “LIVE\_NFLX\_PublicData\_VideoATLAS\_Release” folder: contains the .mat files for each video. These .mat files have the following naming convention:

“cont\_” [content\_index] “\_seq\_” [sequence\_index]

The content index goes from 1 to 14 (14 contents) and the sequence index from 0 to 7 (8 playout patterns).

These .mat files contain:

-> [quality\_model]\_vec: denotes the frame quality scores for each quality model

-> [quality\_model]\_[pooling\_type] denotes the pooled quality scores using average, hysteresis or VQpooling (indicated by kmeans)

-> Nframes: number of frames

-> vid\_fps: frame rate

-> final\_subj\_score: final (summary) QoE ratings after subject rejection and Z-scoring per viewing session and per subject

-> ns: number of stalls (rebuffering events) for the video

-> ds: duration of the stalls (in seconds)

-> lt: duration of encoding bitrate less than 250 kbps (in seconds)

-> VSQM: VsQM metric extracted for this video

-> tsl: time since last impairment finished (bitrate or rebuffering) measured in seconds. This time interval is measured from the time the last impairment finished until the video finishes (where it is assumed that the subject is asked to give his retrospective QoE evaluation). For patterns 0 and 2 this value is set equal to the video duration. The assumption here is to consider only the adaptive streaming strategies when calculating the memory effect.

When the feature name is followed by “\_norm” (e.g. tsl\_norm) this mean that the corresponding feature has been normalized to the video length (number of frames) it was computed on.