
LIVE Video Quality Database

Kalpana Seshadrinathan - Email: kalpsesh@gmail.com

REFERENCES

1. K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack,
"Study of Subjective and Objective Quality Assessment of Video",
accepted for publication, IEEE Transactions on Image Processing, 2009.

2. K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack,
"A Subjective Study to Evaluate Video Quality Assessment Algorithms",
to appear, SPIE Proceedings Human Vision and Electronic Imaging,
Jan. 2010.

3. URL: http://live.ece.utexas.edu/research/quality/live_video.html

VIDEOS

"videos" provides videos in the LIVE Video Quality Database as
multiple zipped files. MD5 checksums are provided for the zip files,
as well as individual yuv files contained within the zip files, to
verify integrity of your download. NOTE: Each zip file in this
directory is at least 700MB. Download times of an hour or more per
file are typical.

All video files have planar YUV 4:2:0 format and do not contain any
headers. The spatial resolution of all videos is 768x432 pixels. The
first 331776 bytes of each file correspond to the 8-bit Y component of
the first frame, followed by 82944 bytes corresponding to the 8-bit U
component of the first frame, followed by 82944 bytes corresponding to

the 8-bit V component of the first frame. Frames are concatenated to form sequence files.

COMPRESSED_VIDEOS

"compressed_videos" provides the videos in the LIVE Video Quality Database in compressed format as multiple zipped files. MD5 checksums are provided for the zip files to verify integrity of your download.

All H.264 compressed bitstreams in the LIVE Video Quality Database are provided as elementary streams in ".264" format. All MPEG-2 compressed bitstreams are provided as elementary streams in ".m2v" format. H.264 bitstreams suffixed with "2" through "8" (see filenaming convention below) contain RTP packets, while H.264 bitstreams suffixed with "9" through "12" are in Annex B bytestream format.

FILENAMING CONVENTION

The video file naming convention is as follows. Each filename has the following pattern - "xx#_ \$fps.yuv".

"xx" is a pattern that denotes the source/reference video sequence that was used to create the test video. There are ten reference videos in the LIVE VQA Database and the patterns used for each of them are as follows:

"bs" - Blue sky, "mc" - Mobile and Calendar, "pa" - Pedestrian Area, "pr" - Park run, "rb" - Riverbed, "rh" - Rushhour, "sf" - Sunflower, "sh" = Shields, "st" - Station, "tr" - Tractor

"#" is a number that denotes the distortion category that was used to create the test video. There are four distortion categories in the

LIVE VQA database (plus the reference) and the numbers used for them are as follows:

"1" - the original reference video

"2","3","4","5" - Wireless distortions (four test videos per reference)

"6","7","8" - IP distortions (three test videos per reference)

"9","10","11","12" - H.264 compression (four test videos per reference)

"13","14","15","16" - MPEG-2 compression (four test videos per reference)

"\$" is a number that denotes the frame rate of the video sequence in frames per second (fps). All videos in the LIVE VQA Database have frame rates of 25 fps or 50 fps.

As an example, for the "blue sky" sequence which has a frame rate of 25 fps:

1. bs1_25fps.yuv is the reference video.

2. bs2_25fps.yuv,

bs3_25fps.yuv, bs4_25fps.yuv and bs5_25fps.yuv are four test videos obtained from the reference using wireless distortions.

3. bs6_25fps.yuv, bs7_25fps.yuv and bs8_25fps.yuv are three test videos obtained from the reference using IP distortions.

4. bs9_25fps.yuv, bs10_25fps.yuv, bs11_25fps.yuv and bs12_25fps.yuv are four test videos obtained from the reference using H.264 compression.

5. bs13_25fps.yuv, bs14_25fps.yuv, bs15_25fps.yuv and bs16_25fps.yuv are four test videos obtained from the reference using MPEG-2 compression.

"pa", "rb", "rh", "sf", "sh", "st" and "tr" have 250 frames (frame rate of 25fps = 10 seconds of video). "bs" has 217 frames (frame rate of 25fps = 8.68 seconds of video). "mc", "pr" and "sh" have 500 frames

(frame rate of 50fps = 10 seconds of video).

SUBJECTIVE DATA FORMAT

The subjective study was conducted using a single stimulus procedure and the subjects indicated the quality of the video on a continuous scale. Subjects also viewed each of the reference videos to facilitate computation of difference scores using hidden reference removal. Each video was viewed by 38 subjects. Unreliable subjects were discarded using the procedure specified in ITU-R BT 500.11. 9 out of 38 subjects were unreliable in our study and the subjective data provided here is from 29 valid subjects.

The subjective data is provided in two files -

"live_video_quality_seqs.txt" which contains names of the video sequences and "live_video_quality_data.txt" which contains the corresponding subjective scores. Each file has 150 lines corresponding to 150 distorted videos in the LIVE Video Quality Database. Line i of "live_video_quality_data.txt" contains the mean DMOS score followed by the standard deviation of the DMOS scores for the video named in line i of "live_video_quality_seqs.txt".

As an example, the 50th line of live_video_quality_seqs.txt is "tr6_25fps.yuv" and the 50th line of live_video_quality_data.txt is "73.4730 11.2189". This means that the mean DMOS score (averaged across subjects) for the video "tr6_25fps.yuv" is 73.4730 and the standard deviation of the DMOS scores is 11.2189.

ERRATUM TO THE PAPER

Since the publication of the paper, we found that the numbers we

published in the paper for PSNR is slightly incorrect. We provide below accurate results of the performance of PSNR on the LIVE Video Quality Database.

SROCC values for PSNR: Erratum to Table I(a) in the paper.

Wireless: SROCC = 0.657411

IP : SROCC = 0.416685

H264 : SROCC = 0.458537

MPEG2 : SROCC = 0.386228

All : SROCC = 0.539792

LCC values for PSNR: Erratum to Table I(b) in the paper.

Wireless: LCC = 0.668984

IP : LCC = 0.464538

H264 : LCC = 0.549255

MPEG2 : LCC = 0.389127

All : LCC = 0.562126

Variance of the residuals between individual subjective scores and PSNR: Erratum to Table II(a) in the paper:

Wireless residual = 165.043361

IP residual = 169.583086

H264 residual = 180.840682

MPEG2 residual = 185.377176

All Data residual = 182.618213

Variance of the residuals between PSNR and DMOS values: Erratum to Table II(b) in the paper.

Wireless: Residual = 61.529318

IP: Residual = 73.344315

H264: Residual = 85.162100

MPEG2: Residual = 88.272099

All Data: Residual = 82.976602

This causes slight changes to the statistical significance results presented in Table III of the paper. In particular, PSNR is no longer statistically inferior to VSNR on "All data" (entry for M1 vs. M5 in Table III becomes -----) and PSNR is no longer statistically inferior to Spatial MOVIE on "Wireless" (entry for M1 vs. M8 in Table III becomes ----0). Note that corresponding changes occur for entries M5 vs. M1 (-----) and M8 vs. M1 (----1) also. Remaining entries in Table III remain unchanged.