Hybrid Perceptual/Bitstream Validation Test Final Report

Version 2.1

Approved July 10, 2014

Hybrid Co-Chairs:

Chulhee Lee (Yonsei University)	Phone: +82 2 2123 2779	Email: chulhee@yonsei.ac.kr
Silvio Borer (SwissQual)	Phone: +41 32 686 65 65	Email: Silvio.Borer@swissqual.com
Jens Berger (SwissQual)	Phone: +41 32 686 65 65	Email: Jens.Berger@swissqual.com

Editorial History

Version	Date	Nature of the modification
1.0	8 January 2014	Initial Draft, created by Margaret Pinson
1.2	24 January 2014	Executive Summary from Boulder 2014 meeting with preliminary results
1.3	March 2014	Edits by SwissQual, Opticom, Acreo and FUB
1.4	23 April 2014	ILG data analysis inserted into the executive summary
1.5	29 April 2014	Executive summary modified with alternative data presentation
1.6	6 May 2014	Edits by SwissQual. Executive summary modified with alternative data presentation. Change from version 1.5 accepted.
1.7	12 May 2014	More information on subjective experiremnts. Changes from version 1.6 accepted.
1.8	May 2014	Changes to text of executive summary made during the audio call, name of OPTICOM model inserted, table inserted that specifies the division of labor on datasets, plus a few other small changes.
1.9	May 2014	Edits to subjective test descriptions and bug fixes. Full tables added to executive summary, with modified data presentation.
1.10	May 2014	Edits to executive summary data tables. ILG data analysis inserted.
1.12	June 2014	Edits from 1.1 to 1.3 according to audio call of June 12
1.13	June 2014	Add limitations for superset. Accepted changes according audio call. Added text for publication of results. Added sample of aggregation values with/without rebuffering. Other small changes.
1.14	June 2014	Accepted changes according audio call. Edited text for publication of results.
1.15	July 1 2014	ILG data analysis edits and insertions.
1.16	July 2 2014	Plots and secondary analysis.
1.17	July 7	Edits from Stockholm VQEG meeting
1.18	July 10	Edits from Stockholm VQEG meeting. Report approved pending final edits
2.0	July 22	Approved final report, including the final edits approved during the Stockholm VQEG meeting
2.1	August 11	Typographical errors corrected in the Appendix.

Editors:

Margaret Pinson, Silvio Borer

Contributors:

Proponents: Savvas Argyropoulos, Jens Berger, Roland Bitto, Silvio Borer, Marie-Neige Garcia, Chulhee Lee, Peter List, Anna Llagostera, Matthias Obermann, Alexander Raake, Werner Robitza, Christian Schmidmer

ILG: Marcus Barkowsky, Vittorio Baroncini, Kjell Brunnström, Philip Corriveau, Lucjan Janowski, Dragan Kukolj, Patrick Le Callet, Mikołaj Leszczuk, Romuald Pépion, Margaret Pinson, Dejan Sretenovic, Nicolas Staelens, Glenn Van Wallendael, Kun Wang, Vladimir Zlokolica

Other: Martin Kastner

Contents

1 Summary

This document presents results from the hybrid project of the Video Quality Experts Group (VQEG). The goal was to evaluate the performance of hybrid perceptual/bitstream models predicting the perceived video quality based on input consisting of video frames and bitstream information. The Hybrid Test addressed no-reference, reduced reference and full reference hybrid models as well as one no-reference non-hybrid model. This Hybrid Test addresses the following video formats: 1080p at 25 and 29.97 frames-per-second, 1080i at 50 and 59.94 fields-per second, VGA at 25 and 30 frames-per-second, and WVGA at 25 and 30 frames-per-second. The Hybrid Test addresses videos encoded using H.264 and transmitted over RTP/UDP for VGA/WVGA formats, and transmitted over TS/RTP/UDP for HD.

A total of eleven testing laboratories coordinated to perform subjective testing. Ten subjective experiments provided data against which model validation was performed. The experiments were divided as follows: five HD experiments, three VGA experiments, and two WVGA experiments. One of the WVGA experiments was assessed twice, once with rebuffering conditions included and once with rebuffering conditions excluded. The impairments examined were restricted to coding artifacts, packet loss, tandem coding, rebuffering, scaling of the resolution, frame rate reduction, error concealment, slicing, freezing, as well as live and simulated transmission errors.

Two common sets of video sequences were created: one for the HD experiments and another for the VGA/WVGA experiments. These common sets were inserted into each experiment, to anchor the video experiments to one another and assist in comparisons between the subjective experiments. These common sequences were used to map the experiments onto a single scale (called “HD merge”, “VGA merge” in this report).

Models were submitted by four proponents: Deutsche Telekom AG, OPTICOM GmbH, SwissQual AG, and Yonsei University. Different models were submitted for different encryption levels (encrypted and non-encrypted) and reduced reference side channel bitrates (56, 128 and 256 kbits/s). Six hybrid no-reference models, two hybrid reduced reference models, four hybrid full reference models and one no-reference model were submitted. No models were withdrawn. Thirteen models are presented in this final report.

The Hybrid data may not be used as evidence to standardize any other objective video quality model that was not tested within this phase. This comparison would not be fair, because another model could have been trained on the Hybrid data.

1.1 Subjective Datasets

- Hybrid VGA1: This test focuses on live video recording of video streams transmitted over a commercially operated 3G mobile network or transmitted over LAN with simulated network impairments. This dataset has 15 different source videos of VGA resolution at 30fps. Video sequences contain coding impairments and transmission impairments with packet loss, packet delay, and rebuffering due to limited throughput.

- Hybrid VGA2: This VGA experiment included x264 encoding (from 128 Kbps to 1200 Kbps) with simulated transmission errors (burst/random, from 0.5%~1.5%). Down-sampling (to QVGA) followed by up-sampling and error concealments (slicing/freezing) were applied to some sequences. The frame rate was set to 30 fps except for some sequences (10 fps).

- Hybrid VGA3: This database is targeting transmission errors without player rebuffering effects. The database includes simulated transmission distortions, as well as transmission over a commercially operated IP Network. Transcoding, scaling, and error concealment were applied to some sequences as pre- and post-processing. The resolution is VGA (640x480 pixels) and the frame rate is 25 frames per second.

- Hybrid WVGA1: This WVGA experiment includes x264 encoding (from 128 Kbps to 1200 Kbps) with simulated transmission errors (burst/random, from 0.5%~1.5%). Down-sampling (to QWVGA) followed by up-sampling and error concealments (slicing/freezing) were applied to some sequences. The frame rate was set to 30 fps except for some sequences (15 fps).

- HybridWVGA2: This WVGA experiment focuses on simulated rebuffering. Videos were streamed over a local loopback, and changing buffer sizes resulted in packet delay and rebuffering. In addition, the test set contains videos with coding only distortions, and down-sampling before transmission followed by up-sampling at the video decoder. A total of 8 source videos were paired with 11 HRCs, resulting in a total of 88 PVSs each at 25fps.

- HybridWVGA2 no rebuf.: Contains the HRCs from dataset Hybrid WVGA2 that do not contain rebuffering (see section 3.1).

- VGA merge: Datasets VGA1, VGA2, VGA3, WVGA1 and WVGA2 are combined into a single dataset, which provides an estimate of the model's overall VGA/WVGA performance. The algorithm used to combine datasets has some limitations (see section 4.2.3 for details).

- VGA merge no rebuf: Does the same but eliminates rebuffering from dataset HybridWVGA2 (see section 3.1).

- Hybrid HD1: This 1080i 60fps experiment contains x264 encoding / simulated loss (uniform-bursty distributions, low/medium/high packet loss rates) / VLC and T-Labs decoder. Many sequences contained network impairments, which resulted in a cluster of low quality data points (from 1 to 2.5).

- Hybrid HD2: This 1080i 50fps experiment presents typical H264 over UDP streaming scenarios at bit rates from 2 Mbit/s to 15 Mbit/s, with transcoding from lower bit rate to higher bit rate, packet losses (from 5-10 packets up to 0.125%), a relatively short GOP structure (12 or 15 frames in a single GOP) and short IP packets (242 bytes long).

- Hybrid HD3: This 1080p 30fps experiment includes x264-encoded sequences with coding distortions and simulated network errors (uniform and bursty loss), targeting H.264 over UDP streaming scenarios with low (1.5 Mbps) to high (8 Mbps) bitrates and low (0.125%) to high (0.5%) packet loss ratios. Packet loss was concealed, resulting in slicing and freezing artifacts.

- Hybrid HD4: This 1080p25 database consists of sequences containing encoding-only artifacts or degradations caused by packet losses during video streaming over UDP. Furthermore, some more advanced features of H.264 video encoding are used such as Intra-refresh, open GOP structures, and hierarchical B-pictures.

- Hybrid HD5: This database contains 10 different source video sequences (1080i, 60fps). This experiment includes x264 encoding (from 2 Mbps to 14 Mbps) with simulated transmission errors (burst/random, from 0.1%~1.3%). Down-sampling (by a factor 1/2, 1/3) followed by up-sampling and error concealments (slicing/freezing) were applied to some sequences.

- HD merge: Datasets HD1, HD2, HD3, HD4 and HD5 are combined into a single dataset, which provides an estimate of the model's overall HD performance. The algorithm used to combine datasets has some limitations (see section 4.2.3 for details).

1.2 Model Summary

The following seven types of models were analyzed:

- Hybrid-NR Non-encrypted: These models use the decoded video frames, packet headers, and video payload. These models can be deployed in-service but cannot analyze encrypted video.

  - TVM-Hybrid Non-Encrypted — Deutsche Telekom AG (DT)
  - VMon-B — SwissQual AG
  - YHyNR — Yonsei

- Hybrid-NR Encrypted: These models use the decoded video frames and packet headers. These models can be deployed in-service and are suitable for use with encrypted video:

  - TVM-Hybrid Encrypted — Deutsche Telekom AG (DT)
  - VMon — SwissQual AG

- YHyNRe Yonsei

- Hybrid-RR: These models use features extracted from the reference video, the decoded video frames, packet headers, and video payload. These models can be deployed in-service but cannot analyze encrypted video. The 256 kbits/s model is only analyzed on the HD datasets.

  - YHyRR56k Yonsei
  - YHyRR128k Yonsei
  - YHyRR256k Yonsei

- Hybrid-RR Encrypted: These models use features extracted from the reference video, the decoded video frames, and packet headers. These models can be deployed in-service and are suitable for use with encrypted video. The 256 kbits/s model is only analyzed on the HD datasets.

  - YHyRR56ke Yonsei
  - YHyRR128ke Yonsei
  - YHyRR256ke Yonsei

- Hybrid-FR: These models use the reference video, the decoded video frames, packet headers, and video payload. These models cannot be deployed in-service and cannot analyze encrypted video.

  - PEVQ-S OPTICOM GmbH
  - YHyFR Yonsei

- Hybrid-FR Encrypted: These models use the reference video, the decoded video frames, and packet headers. These models cannot be deployed in-service and are suitable for use with encrypted video.

  - PEVQ-Se OPTICOM GmbH
  - YHyFRe Yonsei

- NR: These models use the the decoded video frames only. These models can be deployed in-service and are suitable for use with encrypted video.

  - YNR Yonsei

1.3 Model Performance Summary

The models were evaluated using three statistics that provide insights into model performance: Root-Mean Squared Error (RMSE), Pearson Correlation and Epsilon Independent RMSE. Each model was fitted to each subjective experiment using a 3rd order monotonic polynomial function. RMSE is considered the primary metric for analysis in this report. Thus, RMSE is used to determine whether a model is in the group of top performing models for one video format/resolution (i.e. a group of models that include the top performing model and models that are statistically equivalent to the top performing model).

Tables 1-7 provide RMSE and Pearson Correlation for each type of model. Note that:

- MOS was calculated using a 5-level ACR scale, and thus spans the range [1, 5]. Hybrid-NR and NR models are analyzed using MOS.
- DMOS was calculated on the same [1, 5] scale using the ACR-HR method. Hybrid-FR and Hybrid-RR models are analyzed using DMOS.
- PSNR is computed according to ITU-T Rec. P.340, for comparison purposes.
- Within one table, all RMSE values can be directly compared.
- On the top half of each table (RMSE), the yellow highlights indicate that this model is statistically equivalent to the top performing model on this particular dataset. This statistical equivalence is computed using RMSE.
- On the bottom half of each table (Pearson correlation), the light blue highlights indicate that this model is equivalent to or better than PSNR for this particular dataset. This comparison is made for all models that can be deployed in-service (Hybrid-NR, Hybrid-RR and NR). This statistical equivalence is computed using RMSE.
- On the bottom half of each table (Pearson correlation), the light green highlights indicate that this model is better than PSNR for this particular dataset. This comparison is made for all models that cannot be deployed in-service (Hybrid-FR). This statistical equivalence is computed using RMSE.
- “Mean of VGA” computes the averages for that model over all five VGA and WVGA datasets (i.e., HybridVGA1, HybridVGA2, HybridVGA3, HybridWVGA1, and HybridWVGA2).
- “Mean of VGA no rebuf” does the same but eliminates rebuffering from dataset HybridWVGA2, i.e. computes the average of HybridVGA1, HybridVGA2, HybridVGA3, HybridWVGA1, and HybridWVGA2 no rebuf, (see section 3.1).

- “Mean of HD” computes the average for that model over all five HD datasets (i.e., HybridHD1, HybridHD2, HybridHD3, HybridHD4 and HybridHD5).

Comparisons between different types of models are presented in the body of this report.

Table 1. Hybrid-NR Non-encrypted Model Performance Summary, Using MOS[1]

Statistic	Dataset	TVM-Hybrid Non-Encrypted	VMon-B	YHyNR	PSNR
RMSE	HybridVGA1	0.53	0.54	0.70	0.68
	HybridVGA2	0.60	0.58	0.49	0.62
	HybridVGA3	0.72	0.69	0.52	0.59
	HybridWVGA1	0.73	0.71	0.59	0.62
	HybridWVGA2	0.74	0.74	0.49	0.65
	HybridWVGA2 no rebuf	0.95	0.95	0.42	0.65
	VGA merge/- no rebuf	0.68 / 0.70	0.68 / 0.69	0.59 / 0.59	0.69 / 0.70
	Mean of VGA/ - no rebuf	0.66 / 0.70	0.65 / 0.69	0.56 / 0.55	0.63 / 0.63
RMSE	HybridHD1	0.51	0.53	0.43	0.47
	HybridHD2	0.39	0.43	0.54	0.65
	HybridHD3	0.44	0.42	0.47	0.63
	HybridHD4	0.89	0.95	0.70	0.59
	HybridHD5	0.43	0.42	0.50	0.71
	HD merge	0.60	0.64	0.56	0.63
	Mean of HD	0.53	0.55	0.52	0.61
Pearson Correlation	HybridVGA1	0.84	0.83	0.69	0.72
	HybridVGA2	0.81	0.83	0.88	0.80
	HybridVGA3	0.53	0.59	0.79	0.72
	HybridWVGA1	0.72	0.74	0.83	0.81
	HybridWVGA2	0.67	0.67	0.87	0.76
	HybridWVGA2 no rebuf	0.57	0.57	0.93	0.82
	VGA merge/- no rebuf	0.69 / 0.69	0.70 / 0.70	0.78 / 0.79	0.68 / 0.69
	Mean of VGA/ - no rebuf	0.72 / 0.70	0.73 / 0.71	0.81 / 0.82	0.76 / 0.79
Pearson Correlation	HybridHD1	0.84	0.83	0.89	0.87
	HybridHD2	0.92	0.90	0.84	0.76
	HybridHD3	0.90	0.91	0.88	0.78
	HybridHD4	0.64	0.57	0.80	0.86
	HybridHD5	0.87	0.88	0.82	0.58
	HD merge	0.81	0.78	0.84	0.79
	Mean of HD	0.83	0.82	0.85	0.77

[1] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Blue highlight (bottom) indicates model performs equivalently to or better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (blue) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 2. Hybrid-NR Encrypted Model Performance Summary, Using MOS[2]

Statistic	Dataset(s)	TVM-Hybrid Encrypted	VMon	YHyNRe	PSNR
RMSE	HybridVGA1	0.51	0.53	0.70	0.68
	HybridVGA2	0.59	0.57	0.57	0.62
	HybridVGA3	0.54	0.50	0.60	0.59
	HybridWVGA1	0.71	0.69	0.67	0.62
	HybridWVGA2	0.62	0.64	0.56	0.65
	HybridWVGA2 no rebuf	0.41	0.47	0.47	0.65
	VGA merge/- no rebuf	0.61 / 0.61	0.60 / 0.60	0.66 / 0.66	0.69 / 0.70
	Mean of VGA/ - no rebuf	0.59 / 0.55	0.58 / 0.55	0.62 / 0.60	0.63 / 0.63
RMSE	HybridHD1	0.48	0.48	0.34	0.47
	HybridHD2	0.70	0.69	0.72	0.65
	HybridHD3	0.74	0.75	0.70	0.63
	HybridHD4	0.81	0.80	0.74	0.59
	HybridHD5	0.49	0.42	0.55	0.71
	HD merge	0.69	0.67	0.63	0.63
	Mean of HD	0.64	0.63	0.61	0.61
Pearson Correlation	HybridVGA1	0.85	0.84	0.69	0.72
	HybridVGA2	0.82	0.83	0.83	0.80
	HybridVGA3	0.77	0.81	0.71	0.72
	HybridWVGA1	0.74	0.75	0.77	0.81
	HybridWVGA2	0.78	0.77	0.83	0.76
	HybridWVGA2 no rebuf	0.93	0.91	0.91	0.82
	VGA merge/- no rebuf	0.76 / 0.78	0.77 / 0.78	0.72 / 0.73	0.68 / 0.69
	Mean of VGA/ - no rebuf	0.79 / 0.82	0.80 / 0.83	0.77 / 0.78	0.76 / 0.77
Pearson Correlation	HybridHD1	0.86	0.86	0.93	0.87
	HybridHD2	0.72	0.73	0.70	0.76
	HybridHD3	0.67	0.66	0.71	0.78
	HybridHD4	0.72	0.72	0.77	0.86
	HybridHD5	0.83	0.88	0.78	0.58
	HD merge	0.74	0.76	0.79	0.79
	Mean of HD	0.77	0.77	0.78	0.77

[2] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Blue highlight (bottom) indicates model performs equivalently to or better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (blue) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 3. Hybrid-RR Non-Encrypted Model Performance Summary, Using DMOS[3]

Statistic	Dataset(s)	YHyRR56k	YHyRR128k	YHyRR256k	PSNR
RMSE	HybridVGA1	0.79	0.79	—	0.66
	HybridVGA2	0.49	0.49	—	0.63
	HybridVGA3	0.41	0.41	—	0.56
	HybridWVGA1	0.50	0.50	—	0.59
	HybridWVGA2	0.39	0.39	—	0.60
	HybridWVGA2 no rebuf	0.31	0.30	—	0.59
	VGA merge/- no rebuf	0.57 / 0.58	0.57 / 0.58	—	0.66 / 0.66
	Mean of VGA/ - no rebuf	0.52 / 50	0.51 / 0.50	—	0.61 / 0.61
RMSE	HybridHD1	0.41	0.41	0.41	0.42
	HybridHD2	0.67	0.66	0.66	0.59
	HybridHD3	0.52	0.52	0.52	0.60
	HybridHD4	0.56	0.56	0.56	0.60
	HybridHD5	0.46	0.46	0.46	0.72
	HD merge	0.55	0.55	0.55	0.61
	Mean of HD	0.52	0.52	0.52	0.59
Pearson Correlation	HybridVGA1	0.63	0.63	—	0.77
	HybridVGA2	0.89	0.89	—	0.81
	HybridVGA3	0.88	0.88	—	0.75
	HybridWVGA1	0.88	0.88	—	0.83
	HybridWVGA2	0.92	0.92	—	0.79
	HybridWVGA2 no rebuf	0.96	0.96	—	0.86
	VGA merge/- no rebuf	0.80 / 0.80	0.80 / 0.81	—	0.72 / 0.73
	Mean of VGA/ - no rebuf	0.84 / 0.85	0.84 / 0.85	—	0.79 / 0.80
Pearson Correlation	HybridHD1	0.91	0.91	0.91	0.91
	HybridHD2	0.78	0.79	0.79	0.84
	HybridHD3	0.86	0.86	0.86	0.80
	HybridHD4	0.88	0.87	0.88	0.86
	HybridHD5	0.86	0.86	0.86	0.60
	HD merge	0.86	0.86	0.86	0.83
	Mean of HD	0.86	0.86	0.86	0.80

[3] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Blue highlight (bottom) indicates model performs equivalently to or better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (blue) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 4. Hybrid-RR Encrypted Model Performance Summary, Using DMOS[4]

Statistic	**Dataset(s)**	YHyRR56ke	YHyRR128ke	YHyRR256ke	PSNR
RMSE	HybridVGA1	0.79	0.78	—	0.66
	HybridVGA2	0.49	0.49	—	0.63
	HybridVGA3	0.44	0.44	—	0.56
	HybridWVGA1	0.49	0.49	—	0.59
	HybridWVGA2	0.42	0.41	—	0.60
	HybridWVGA2 no rebuf	0.30	0.30	—	0.59
	VGA merge/- no rebuf	0.58 / 0.59	0.58 / 0.58	—	0.66 / 0.66
	Mean of VGA/ - no rebuf	0.53 / 0.50	0.52 / 0.50	—	0.61 / 0.61
RMSE	HybridHD1	0.38	0.38	0.38	0.42
	HybridHD2	0.74	0.73	0.73	0.59
	HybridHD3	0.64	0.64	0.64	0.60
	HybridHD4	0.60	0.60	0.59	0.60
	HybridHD5	0.39	0.39	0.38	0.72
	HD merge	0.58	0.58	0.58	0.61
	Mean of HD	0.55	0.55	0.55	0.59
Pearson Correlation	HybridVGA1	0.64	0.64	—	0.77
	HybridVGA2	0.89	0.89	—	0.81
	HybridVGA3	0.86	0.86	—	0.75
	HybridWVGA1	0.89	0.89	—	0.83
	HybridWVGA2	0.91	0.91	—	0.79
	HybridWVGA2 no rebuf	0.96	0.96	—	0.86
	VGA merge/- no rebuf	0.79 / 0.80	0.79 / 0.80	—	0.72 / 0.73
	Mean of VGA/ - no rebuf	0.84 / 0.85	0.84 / 0.85	—	0.79 / 0.80
Pearson Correlation	HybridHD1	0.92	0.92	0.92	0.91
	HybridHD2	0.72	0.73	0.73	0.84
	HybridHD3	0.78	0.78	0.78	0.80
	HybridHD4	0.86	0.86	0.86	0.86
	HybridHD5	0.91	0.91	0.91	0.60
	HD merge	0.85	0.85	0.85	0.83
	Mean of HD	0.84	0.84	0.84	0.80

[4] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Blue highlight (bottom) indicates model performs equivalently to or better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (blue) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 5. Hybrid-FR Non-encrypted Model Performance Summary, Using DMOS[5]

Statistic	Dataset(s)	PEVQ-S	Yonsei-hFR	PSNR
RMSE	HybridVGA1	0.65	0.79	0.66
	HybridVGA2	0.51	0.49	0.63
	HybridVGA3	0.52	0.41	0.56
	HybridWVGA1	0.54	0.50	0.59
	HybridWVGA2	0.53	0.39	0.60
	HybridWVGA2 no rebuf	0.51	0.31	0.59
	VGA merge/- no rebuf	0.57 / 0.55	0.57 / 0.58	0.66 / 0.66
	Mean of VGA/ - no rebuf	0.55 / 0.55	0.52 / 0.50	0.61 / 0.61
RMSE	HybridHD1	0.34	0.41	0.42
	HybridHD2	0.51	0.66	0.59
	HybridHD3	0.41	0.52	0.60
	HybridHD4	0.64	0.57	0.60
	HybridHD5	0.50	0.46	0.72
	HD merge	0.51	0.55	0.61
	Mean of HD	0.48	0.52	0.59
Pearson Correlation	HybridVGA1	0.77	0.63	0.77
	HybridVGA2	0.88	0.89	0.81
	HybridVGA3	0.79	0.88	0.75
	HybridWVGA1	0.86	0.88	0.83
	HybridWVGA2	0.84	0.92	0.79
	HybridWVGA2 no rebuf	0.89	0.96	0.86
	VGA merge/- no rebuf	0.81 / 0.83	0.80 / 0.81	0.72 / 0.73
	Mean of VGA/ - no rebuf	0.83 / 0.84	0.84 / 0.85	0.79 / 0.80
Pearson Correlation	HybridHD1	0.94	0.91	0.90
	HybridHD2	0.88	0.79	0.84
	HybridHD3	0.91	0.86	0.80
	HybridHD4	0.83	0.87	0.86
	HybridHD5	0.84	0.86	0.60
	HD merge	0.88	0.86	0.83
	Mean of HD	0.88	0.86	0.80

[5] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Green highlight (bottom) indicates model performs better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (green) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 6. Hybrid-FR Encrypted Model Performance Summary, Using DMOS[6]

Statistic	Dataset(s)	PEVQ-S (pes+rtp)	PEVQ-S (ts+rtp)	YHyFRe	PSNR
RMSE	HybridVGA1	0.65	0.65	0.78	0.66
	HybridVGA2	0.51	0.51	0.49	0.63
	HybridVGA3	0.52	0.52	0.44	0.56
	HybridWVGA1	0.54	0.54	0.49	0.59
	HybridWVGA2	0.53	0.53	0.41	0.60
	HybridWVGA2 no rebuf	0.51	0.51	0.30	0.59
	VGA merge/- no rebuf	0.57 / 0.55	0.57 / 0.55	0.58 / 0.59	0.66 / 0.66
	Mean of VGA/ - no rebuf	0.55 / 0.55	0.55 / 0.55	0.52 / 0.50	0.61 / 0.61
RMSE	HybridHD1	0.34	0.34	0.38	0.42
	HybridHD2	0.51	0.50	0.73	0.59
	HybridHD3	0.41	0.41	0.64	0.60
	HybridHD4	0.64	0.64	0.60	0.60
	HybridHD5	0.50	0.50	0.38	0.72
	HD merge	0.51	0.51	0.58	0.61
	Mean of HD	0.48	0.48	0.55	0.59
Pearson Correlation	HybridVGA1	0.77	0.77	0.64	0.77
	HybridVGA2	0.88	0.88	0.89	0.81
	HybridVGA3	0.79	0.79	0.86	0.75
	HybridWVGA1	0.86	0.86	0.89	0.83
	HybridWVGA2	0.84	0.84	0.91	0.79
	HybridWVGA2 no rebuf	0.89	0.89	0.96	0.86
	VGA merge/- no rebuf	0.81 / 0.83	0.81 / 0.83	0.79 / 0.80	0.72 / 0.73
	Mean of VGA/ - no rebuf	0.83 / 0.84	0.83 / 0.84	0.84 / 0.85	0.79 / 0.80
Pearson Correlation	HybridHD1	0.94	0.94	0.92	0.91
	HybridHD2	0.88	0.88	0.73	0.84
	HybridHD3	0.91	0.91	0.78	0.80
	HybridHD4	0.83	0.83	0.86	0.86
	HybridHD5	0.84	0.84	0.91	0.60
	HD merge	0.88	0.88	0.85	0.83
	Mean of HD	0.88	0.88	0.84	0.80

[6] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Green highlight (bottom) indicates model performs better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (green) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

Table 7. NR Model Performance Summary, Using MOS[7]

Statistic	Dataset(s)	YNR	PSNR
RMSE	HybridVGA1	0.78	0.68
	HybridVGA2	0.87	0.62
	HybridVGA3	0.72	0.59
	HybridWVGA1	0.78	0.62
	HybridWVGA2	0.77	0.65
	HybridWVGA2 no rebuf	0.91	0.65
	VGA merge/- no rebuf	0.79 / 0.81	0.69 / 0.70
	Mean of VGA/ - no rebuf	0.78 / 0.81	0.63 / 0.63
RMSE	HybridHD1	0.58	0.47
	HybridHD2	0.72	0.65
	HybridHD3	0.84	0.63
	HybridHD4	0.83	0.59
	HybridHD5	0.73	0.71
	HD merge	0.76	0.63
	Mean of HD	0.74	0.61
Pearson Correlation	HybridVGA1	0.60	0.72
	HybridVGA2	0.53	0.80
	HybridVGA3	0.54	0.72
	HybridWVGA1	0.67	0.81
	HybridWVGA2	0.63	0.76
	HybridWVGA2 no rebuf	0.61	0.82
	VGA merge/- no rebuf	0.54 / 0.54	0.68 / 0.69
	Mean of VGA/ - no rebuf	0.59 / 0.59	0.76 / 0.77
Pearson Correlation	HybridHD1	0.79	0.87
	HybridHD2	0.69	0.76
	HybridHD3	0.54	0.78
	HybridHD4	0.70	0.86
	HybridHD5	0.55	0.58
	HD merge	0.67	0.79
	Mean of HD	0.65	0.77

[7] Yellow highlight (top) indicates model is either the top performing model for this dataset or has equivalent performance. Blue highlight (bottom) indicates model performs equivalently to or better than PSNR. These statistical comparisons are computed using RMSE for both the top (yellow) and bottom (blue) part. Highlights in the "VGA merge/- no rebuff" rows mark the performance of "VGA merge".

1.4 Conclusions

Based on this analysis VQEG concludes that:

1. The three evaluated no-reference Hybrid Models for quality predictions using non-encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.
2. The three evaluated no-reference Hybrid Models for quality predictions using encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.

3. The three evaluated reduced-reference Hybrid Models for quality predictions using non-encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.
4. The three evaluated reduced-reference Hybrid Models for quality predictions using encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.

5. The two evaluated full-reference Hybrid Models for quality predictions using non-encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.
6. The two evaluated full-reference Hybrid Models for quality predictions using encrypted bitstreams perform well enough to be used for quality assessment and therefore are appropriate to be included in normative sections of Recommendations.

2 Use of the Hybrid Data

Subjective data, objective model validation data, and model analyses are published in this report.

The source and processed video sequences for the eight experiments HybridHD1,…, HybridHD5, HybridVGA1, HybridVGA2, HybridWVGA2 have been approved for redistribution and use in research experiments. Proper approval must be obtained from the copyright holders of the source video sequences. To obtain approval for access to the source video sequences, the Content User Agreement form available from the Consumer Digital Video Library (www.cdvl.org) must be completed. The source and processed video sequences for experiments HybridVGA3, HybridWVGA1 are not available for redistribution.

Appropriate uses for VQEG Hybrid Phase I subjective data, objective data, video material, and analyses include:

- Subjective data and video material may be used to train new objective video quality models.
- The VQEG Hybrid Phase I statistics and analyses may be used to compare models from different proponents provided that the context is maintained. As a first example, if rmse values of a specific model type and video resolution are discussed, then all rmse values of all databases with the same resolution must be presented for all models that are compared. As a second example, if models are compared on subsets of the data, then the publication should mention that the hybrid final report includes additional information.
- Objective data and video material may be used to confirm the performance of a model mentioned in this report.
- Additional experiments may be performed using this video material and subjective data.
- The statistical analysis from this report may be used for promotion of one or some of the models mentioned in this report inside SDOs (Standards Developing Organizations) with the purpose to include the model(s) in standards. Such promotion must however always follow the other guidelines for publications given in this report. Anyone who promotes models from this report to an SDO other than by liaison from VQEG to the SDO needs to inform the VQEG/Hybrid reflector.
- Only proponents of models mentioned in this report as well as ILGs may publish secondary analysis results based on the objective data, unless the models are anonymized.
- Proponents of models mentioned in this report may freely use any statistical data if only their own models are mentioned.

Inappropriate uses for VQEG Hybrid Phase I subjective data, objective data, video material, and analyses include the following:

- Using the data to propose a model not mentioned in this report for standardization is not permitted.
- Use of the video material in a commercial application is not permitted (e.g., product brochure, customer demonstration).
- It is not allowed to claim based upon the use of these data that a model not mentioned in this report has superior performance to the models mentioned in this report.
- Results for these datasets obtained using other models that could have been trained on these datasets must not be compared to results of the models mentioned in this report which were not trained on these dataset and independently validated. Such a comparison is misleading, because the experiments contain mainly source scenes and HRCs that were unknown to the developers of the models presented in this report. Additionally, this comparison is misleading because one dataset has been kept private.

All Publications resulting from any use of the VQEG Hybrid Phase I data, analyses, or video material must:

- Mention the VQEG Hybrid Final Report
- Respect the copyright holders' usage limitations on appropriate uses of the source video
- State clearly that a model was trained on this video material, where appropriate.
- No publication is allowed before the VQEG Hybrid Phase I report has been finalized and published. Paper submissions may be earlier though. An exemption is possible if all proponents agree, in case that this report should not be finalized before the next SG9 and SG12 meetings.
- No one may publish improved models based on the hybrid datasets until ITU-T SG9 has decided models to be included in the Recommendations (i.e., the publication date must be after the SG9 decision; papers may be submitted prior to that).

3 Overview of Validation Process

3.1 Validation Process

See the "Hybrid Perceptual/Bitstream Group Test Plan" for a full description of this validation test's design.

The following changes or clarifications were made to the validation procedure after test plan approval:

- Contrary to section 7.2 of the test plan, no audio was to be included in HD PVSs.
- With respect to section 7.7 it was clarified that it is not allowed to use High-4:2:2-profile for encoding of HD video sequences.
- Contrary to section 9 of the test plan, it was approved that for proponent experiments all PVSs were processed by the same proponent.

The following changes to the validation procedure were approved after model submission:

- Rebuffering simulation of WVGA2 was allowed to be included after submission of models. In addition, it was approved that proponents can withdraw their model from evaluation on rebuffering PVSs of WVGA2.

3.2 Test Laboratories

The independent lab group (ILG) had the role of independent arbitrator for the hybrid perceptual/bitstream test. The ILG performed 1 WVGA and 3 HD subjective tests. For these tests, the ILG was the sole responsible for all aspects related to scene choice, HRC choice, and the design of each subjective test. The ILG also performed all scene selection, validated proponent models and performed the official data analysis. The members of the ILG were:

Independent Lab Group Organization	Website
Institut de Recherche en Communication et Cybernétique de Nantes IRCCyN	http://www.irccyn.ec-nantes.fr http://www.irccyn.ec-nantes.fr/spip.php?rubrique24&lang=en
Acreo Swedish ICT AB	www.acreo.se
RT-RK	http://www.rt-rk.com
Fondazione Ugo Bordoni (FUB)	www.fub.it
Ghent University - iMinds	http://www.iminds.be http://multimedialab.elis.ugent.be http://www.ibcn.intec.ugent.be
Department of Telecommunication of AGH University of Science and Technology	http://kt.agh.edu.pl/en

The proponents submitted one or more models to the ILG for validation. Proponents were responsible for running their own model on all video sequences, submitting the resulting objective data for validation, and coordinating the validation effort. Proponents paid a fee to the ILG laboratories performing the subjective experiments to cover basic costs of those experiments. The list of proponents whose models are included in this report are:

Proponent Organization	Website
Deutsche Telekom AG (DT)	http://www.laboratories.telekom.com
OPTICOM Dipl.-Ing. M. Keyhl GmbH	www.opticom.de
SwissQual AG	http://www.swissqual.com
Yonsei University	http://web.yonsei.ac.kr/hdsp

3.3 Models

Model Name	Model Type	Proponent	Contact Information
TVM-Hybrid Encrypted	No-reference hybrid model for encrypted bitstream	Deutsche Telekom AG (DT)	alexander.raake@telekom.de
TVM-Hybrid Non-Encrypted	No-reference hybrid model for non-encrypted bitstream	Deutsche Telekom AG (DT)	alexander.raake@telekom.de
PEVQ-S	Full-reference hybrid model for encrypted bitstream	OPTICOM GmbH	info@pevq.com
PEVQ-S	Full-reference hybrid model for non-encrypted bit-stream	OPTICOM GmbH	info@pevq.com
VMon	No-reference hybrid model	SwissQual	silvio.borer@swissqual.com
VMon-B	No-reference hybrid model for non-encrypted bitstream	SwissQual	silvio.borer@swissqual.com
YHyFR	Full-reference hybrid model for non-encrypted bitstream	Yonsei Univ.	chulhee@yonsei.ac.kr
YHyFRe	Full-reference hybrid model	Yonsei Univ.	chulhee@yonsei.ac.kr

Model Name	Model Type	Proponent	Contact Information
YHyRR	Reduced-reference hybrid model for non-encrypted bitstream	Yonsei Univ.	chulhee@yonsei.ac.kr
YHyRRe	Reduced-reference hybrid model	Yonsei Univ.	chulhee@yonsei.ac.kr
YHyNR	No-reference hybrid model for non-encrypted bitstream	Yonsei Univ.	chulhee@yonsei.ac.kr
YHyNRe	No-reference hybrid model	Yonsei Univ.	chulhee@yonsei.ac.kr
YNR	No-reference model	Yonsei Univ.	chulhee@yonsei.ac.kr

3.4 Subjective and Objective Data

The subjective data (MOS and DMOS) and objective data for each model presented are available in companion documents as comma separated values (CSV).

3.5 Subjective Test Summary

The subjective tests were conducted using the absolute category scale (ACR) from ITU-T Rec. P.910.

Table 8 shows the organizations responsible for each task within the creation of the ten datasets. A sixth dataset was planned by FUB (HD06, 720 50fps) but not completed due to problems encountered in the PVS generation. An initial selection of scenes was made by NTIA/ITS. The organization identified in the “Scenes” column was responsible for checking the quality of the scenes and replacing as needed. “Design” identifies the organization that made the subjective test design. “PVS & Test Vectors” identifies the organization that created the test vectors and PVSs. “Subjects” identifies the organization that ran subjects through the experiment.

See the executive summary for a brief overview of each experiment. See Annex B for a full description of each dataset.

Table 8. Dataset Creation Duties and Responsibilities

Test	**Format**	**Scenes**	**Design**	**PVS & Test Vectors**	**Subjects**
HD01	1080i 60fps	IRCCyN	DT	RT-RK	IRCCyN
HD02	1080i 50fps	AGH	AGH	RT-RK	AGH
HD03	1080p 30fps	DT	DT	DT	DT
HD04	1080p 25fps	Ghent Univ. - iMinds	Ghent Univ. - iMinds	Ghent Univ. - iMinds	Ghent Univ. - iMinds
HD05	1080i 50fps	Yonsei	Yonsei	Yonsei	Yonsei
VGA01	VGA 30fps	SwissQual	SwissQual	SwissQual	Yonsei
VGA02	VGA 30fps	Yonsei	Yonsei	Yonsei	SwissQual
VGA03	VGA 25fps	OPTICOM	OPTICOM	OPTICOM	Yonsei
WVGA01	WVGA 30fps	Yonsei	Yonsei	Yonsei	OPTICOM
WVGA02	WVGA 25fps	Acreo	Acreo	FUB, RT-RK and Acreo	FUB

3.6 Changes to Models and PVSs After Submission

If a model crashed and was unable to produce a value, the proponent had the choice of either (1) running without the PCAP file or (2) substituting the encrypted model’s value. The missing unencrypted model values were replaced as follows:

Yonsei replaced "-999" for the non-encrypted models (HNR, HRR(56, 128, 256), HFR) for the following three PVSs:

- h01_src01_hrc07

- h01_src07_hrc07
- h01_src09_hrc16

The non-encrypted models crashed and the values produced by the corresponding encrypted models were used.

SwissQual and DT missing values with the value from the encrypted model for the following three PVSs:

- v03_src15_hrc19_h264
- h02_src01_hrc02
- h04_src08_hrc08

Missing values for PVSs that will be eliminated were ignored.

4 Official ILG Data Analysis

4.1 Issues

The Test plan states at chapter 11.1 ”Model Type and Model Requirements“ that “…Model for an encrypted bit-stream MUST provide two modes: One for handling PES payload encryption only and one for handling TS payload encryption.….”

For WVGA/VGA encryption is RTP. For HD, two types of encryption are allowed: TS and PES. Three of the proponents treat the TS and PES encryption identically, while PEVQ-S yields separate answers depending upon the HD encryption types. For this reason, the PEVQ-S encrypted models include encryption type; and two PEVQ-S encrypted models are provided for HD.

4.2 Common Video Clip Set Analysis

The common set analysis is based on the correlation of the common set results obtained for different experiments, similarly to the HDTV analysis. Since two different common sets for HD and VGA/WVGA were used the analysis is repeated twice separately for each common set.

4.2.1 Common Set for HD Experiments

Before the analysis it was revealed that one PVS (HybridHD5_csrc01_hrc12.avi) at the HD4 experiment was incorrect and the subjects have seen a different sequence than the intended PVS. Therefore, this PVS was removed from the common set analysis. Moreover, source number 3 (a sequence with heavy snow) was judged differently in HD4 experiment comparing to other experiments. In the case of experiment HD4 it could be explained by incorrect deinterlacing algorithm, the sequence was interlaced by the experiment was not. Therefore, in case of experiment HD4 the number of PVSes was limited to 17. In all other experiments all sequences were used.

Limiting number of sequences removed sequences with MOS range from 3.5 to 4.5 as presented in Figure 1. Nevertheless, the range of MOSs is not changed therefore the obtained conclusions should be correct. Less points increases probability of having high correlation, but no better solution can be found. Removing all experiment just because of some problems with common set would limit the obtained result without good reason.

The lab-to-lab Pearson correlations are shown in Table 9. The row labeled “average” shows the correlation between one dataset (in the column) to the average MOS calculated across all five HD datasets. Note that common set mapping is always calculated on MOS and includes the original SRCs. Scatter plots are available in Annex C.1.

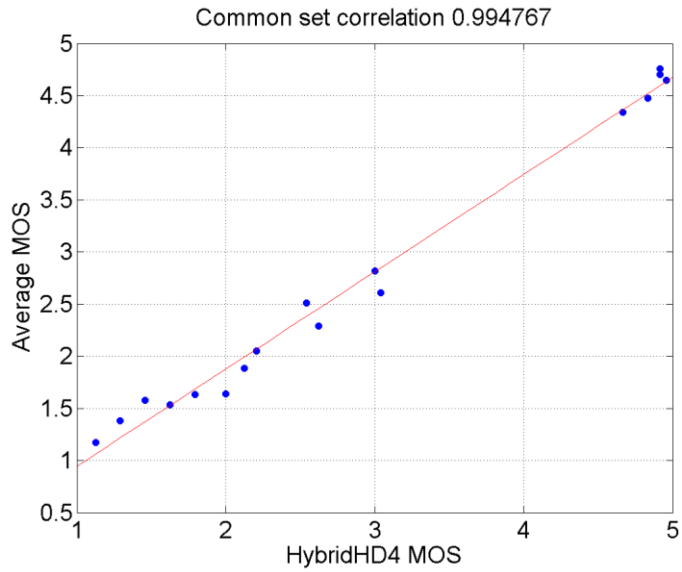


Figure 1. Scatter plot for the limited HD4 common set compared with the mean value obtained for the common set of the all experiments.

Table 9. Lab to lab correlation for HD Experiments

	HybridHD1	HybridHD2	HybridHD3	HybridHD4	HybridHD5
HybridHD1	1.00	0.98	0.98	0.99	0.98
HybridHD2	0.98	1.00	0.97	0.99	0.97
HybridHD3	0.98	0.97	1.00	0.97	0.97
HybridHD4	0.99	0.99	0.97	1.00	0.98
HybridHD5	0.98	0.97	0.97	0.98	1.00
Average	0.99	0.99	0.99	0.99	0.99

The obtained correlations indicate that the supper set analysis can be run on the HD experiments.

4.2.2 Common Set for VGA/WVGA Experiments

In the case of VGA/WVGA experiments no problems with common set PVSs were reported. Therefore, all sequences were used to compute correlation. The lab-to-lab Pearson correlations are shown in Table 10. The row labeled “average” shows the correlation between one dataset (in the column) to the average MOS calculated across all five VGA/WVGA datasets. Note that common set mapping is always calculated on MOS and includes the original SRCs. Scatter plots are available in Annex C.1

Table 10. Lab to lab correlation for VGA/WVGA Experiments

	HybridVGA1	HybridVGA2	HybridVGA3	HybridWVGA1	HybridWVGA2
HybridVGA1	1.00	0.97	0.98	0.96	0.95
HybridVGA2	0.97	1.00	0.97	0.98	0.96
HybridVGA3	0.98	0.97	1.00	0.95	0.93
HybridWVGA1	0.96	0.98	0.95	1.00	0.94
HybridWVGA2	0.95	0.96	0.93	0.94	1.00
Average	0.99	0.99	0.98	0.98	0.97

Similarly to the HD experiments, the obtained results shows high correlation between different experiments and over all mean. The worst correlation is obtained in the case of WVGA2 experiment which was the only one with rebuffering. Nevertheless, even in this case the obtained correlation is high and supper set analysis can be run.

4.2.3 Merged Dataset Procedure and Limitations

The HD dataset PVSs were merged into a single dataset ("HD merge") using a linear map of the common set sequence scores. That is, a linear mapping function was computed using the MOS scores of the 24 common set sequences, including the original sequences. Mapped DMOS scores were calculated after this mapping. Likewise, the VGA and WVGA datasets were combined into a single dataset ("VGA merge").

The advantages of the merged analysis are as follows:

- All available data is examined simultaneously within one dataset.
- The sensitivity of the RMSE significance test increases in response to the increased number of data points.

This mapping procedure and the resulting merged datasets have some limitations:

- The common set sequences were format converted for inclusion into five datasets with disparate formats (e.g., 1080i 50fps for HD02, 1080p30 for HD03, WVGA to VGA). Consequently, subjects viewed and rated slightly different versions of each sequence. The SRC were chosen to minimize the visual impact of these conversions. Still, these visual differences could adversely impact the mapping function in a way that is difficult to detect.
- The mapping function applies a linear fit. The common set is too small to accurately measure and remove nonlinearities.
- The HD common set contains a gap of missing scores between approximately 3.0 (fair) and 4.0 (good), see plots in Figure 1 and section C.1. This resulted from one SRC being eliminated from the HD common set.

Put simply, the merging process introduces some error into the data. The resulting dataset statistics are less accurate than would be a hypothetical subjective dataset conducted on all of those PVSs. Thus, the merged dataset statistics should be considered within the context of the other available data: the description of each dataset, the analyses of individual datasets, and mean computed over all datasets.

4.3 Model Analysis with MOS

This section analyzes Hybrid-NR Encrypted, Hybrid-NR Non-Encrypted and NR models using MOS. Please note the following:

- Dataset names and summaries are available in section 1.1.
- “VGA merge” combines datasets HybridVGA1, HybridVGA2, HybridVGA3, HybridWVGA1 and HybridWVGA2 into a single dataset, which provides an estimate of the model’s overall VGA/WVGA performance. The algorithm used to combine datasets has some limitations (see section 4.2.3 for details).
- “Mean of VGA” computes the averages for that model over all five VGA and WVGA datasets (i.e., HybridVGA1, HybridVGA2, HybridVGA3, HybridWVGA1, and HybridWVGA2).
- “Mean of VGA no rebuf” does the same but eliminates rebuffering from dataset HybridWVGA2 (i.e., computes the average of HybridVGA1, HybridVGA2, HybridVGA3, HybridWVGA1, and HybridWVGA2 no rebuf).
- “HD merge” combines datasets HybridHD1, HybridHD2, HybridHD3, HybridHD4 and HybridHD5 into a single dataset, which provides an estimate of the model’s overall VGA/WVGA performance. The algorithm used to combine datasets has some limitations (see section 4.2.3 for details).
- “Mean of HD” computes the average for that model over all five HD datasets (i.e., HybridHD1, HybridHD2, HybridHD3, HybridHD4 and HybridHD5).

Table 11. RMSE for Hybrid-NR Non-Encrypted, Hybrid-NR Encrypted and NR Models (MOS)

	TVM-Hybrid Non-Encrypted	TVM-Hybrid Encrypted	VMon-B	VMon	YHyNR	YHyNRe	YNR	PSNR
HybridVGA1	0.53	0.51	0.54	0.53	0.70	0.70	0.78	0.68
HybridVGA1	0.53	0.51	0.54	0.53	0.70	0.70	0.78	0.68
HybridVGA2	0.60	0.59	0.58	0.57	0.49	0.57	0.87	0.62
HybridVGA3	0.72	0.54	0.69	0.50	0.52	0.60	0.72	0.59
HybridWVGA1	0.73	0.71	0.71	0.69	0.59	0.67	0.78	0.62
HybridWVGA2	0.74	0.62	0.74	0.64	0.49	0.56	0.77	0.65
HybridWVGA2 no rebuf	0.95	0.41	0.95	0.47	0.42	0.47	0.91	0.65
VGA merge	0.68	0.61	0.68	0.60	0.59	0.66	0.79	0.69
VGA merge no rebuf	0.70	0.61	0.69	0.60	0.59	0.66	0.81	0.70
Mean of VGA	0.66	0.59	0.65	0.58	0.56	0.62	0.78	0.63
Mean of VGA no rebuf	0.70	0.55	0.69	0.55	0.55	0.60	0.81	0.63
HybridHD1	0.52	0.48	0.53	0.48	0.41	0.34	0.58	0.47
HybridHD2	0.39	0.70	0.43	0.69	0.54	0.72	0.72	0.65
HybridHD3	0.44	0.74	0.42	0.75	0.47	0.70	0.84	0.63
HybridHD4	0.89	0.81	0.95	0.80	0.70	0.74	0.83	0.59
HybridHD5	0.43	0.49	0.42	0.42	0.50	0.55	0.73	0.71
HD merge	0.60	0.69	0.64	0.67	0.56	0.63	0.76	0.63
Mean of HD	0.53	0.64	0.55	0.63	0.52	0.61	0.74	0.61

Table 12. Pearson Correlation for Hybrid-NR Non-Encrypted, Hybrid-NR Encrypted and NR Models (MOS)

	TVM-Hybrid Non-Encrypted	**TVM-Hybrid Encrypted**	**VMon-B**	**VMon**	**YHyNR**	**YHyNRe**	**YNR**	**PSNR**
HybridVGA1	0.84	0.85	0.83	0.84	0.69	0.69	0.60	0.72
HybridVGA1	0.84	0.85	0.83	0.84	0.69	0.69	0.60	0.72
HybridVGA2	0.81	0.82	0.83	0.83	0.88	0.83	0.53	0.80
HybridVGA3	0.53	0.77	0.59	0.81	0.79	0.71	0.54	0.72
HybridWVGA1	0.72	0.74	0.74	0.75	0.83	0.77	0.67	0.81
HybridWVGA2	0.67	0.78	0.67	0.77	0.87	0.83	0.63	0.76
HybridWVGA2 no rebuf	0.57	0.93	0.57	0.91	0.93	0.91	0.61	0.82
VGA merge	0.69	0.76	0.70	0.77	0.78	0.72	0.54	0.68
VGA merge no rebuf	0.69	0.78	0.70	0.78	0.79	0.73	0.54	0.69
Mean of VGA	0.72	0.79	0.73	0.80	0.81	0.77	0.59	0.76
Mean of VGA no rebuf	0.70	0.82	0.71	0.83	0.82	0.78	0.59	0.77
HybridHD1	0.84	0.86	0.83	0.86	0.90	0.93	0.79	0.87
HybridHD2	0.92	0.72	0.90	0.73	0.84	0.70	0.69	0.76
HybridHD3	0.90	0.67	0.91	0.66	0.88	0.71	0.54	0.78
HybridHD4	0.64	0.72	0.57	0.72	0.80	0.77	0.70	0.86
HybridHD5	0.87	0.83	0.88	0.88	0.82	0.78	0.55	0.58
HD merge	0.81	0.74	0.78	0.76	0.84	0.79	0.67	0.79
Mean of HD	0.83	0.76	0.82	0.77	0.85	0.78	0.65	0.77

Table 13. Epsilon Independent RMSE (RMSE*) for Hybrid-NR Encrypted, Hybrid-NR Non-Encrypted and NR Models (MOS)

	TVM-Hybrid Non-Encrypted	TVM-Hybrid Encrypted	VMon-B	VMon	YHyNR	YHyNRe	YNR	PSNR
HybridVGA1	0.34	0.28	0.36	0.30	0.50	0.50	0.55	0.47
HybridVGA1	0.34	0.28	0.36	0.30	0.50	0.50	0.55	0.47
HybridVGA2	0.41	0.40	0.39	0.38	0.32	0.39	0.67	0.42
HybridVGA3	0.47	0.31	0.45	0.27	0.29	0.36	0.48	0.36
HybridWVGA1	0.51	0.49	0.50	0.47	0.40	0.47	0.58	0.43
HybridWVGA2	0.57	0.42	0.57	0.43	0.29	0.37	0.58	0.46
HybridWVGA2 no rebuf	0.75	0.24	0.75	0.28	0.24	0.30	0.71	0.45
VGA merge	0.48	0.41	0.47	0.40	0.40	0.46	0.58	0.48
VGA merge no rebuf	0.49	0.40	0.48	0.39	0.40	0.46	0.59	0.48
VGA-WVGA	0.46	0.38	0.45	0.37	0.36	0.42	0.57	0.43
VGA-WVGA-NoRebuffering	0.50	0.34	0.49	0.34	0.35	0.41	0.60	0.42
HybridHD1	0.37	0.28	0.37	0.29	0.22	0.16	0.39	0.26
HybridHD2	0.18	0.50	0.22	0.48	0.33	0.51	0.50	0.43
HybridHD3	0.25	0.52	0.23	0.54	0.27	0.49	0.62	0.41
HybridHD4	0.68	0.61	0.74	0.60	0.50	0.56	0.64	0.39
HybridHD5	0.23	0.30	0.23	0.23	0.31	0.35	0.52	0.51
HD merge	0.42	0.49	0.45	0.47	0.37	0.44	0.56	0.43
Mean of HD	0.34	0.44	0.36	0.43	0.32	0.41	0.54	0.40

These next tables show the group of top performing models for a particular category of model. Every box marked with a one (1) indicates that this model is either the top performing model for that dataset, or is statistically equivalent to that top performing model with 95% confidence. This significance is computed using RMSE. Zero (0) indicates that the model has worse performance than the top performing model for that dataset.

Table 14. Top Performing Models: All Hybrid-NR Non-Encrypted, Hybrid-NR Encrypted and NR Models

	TVM-Hybrid Non-Encrypted	**TVM-Hybrid Encrypted**	**VMon-B**	**VMon**	**YHyNR**	**YHyNRe**	**YNR**	**PSNR**
HybridVGA1	1	1	1	1	0	0	0	0
HybridVGA2	0	0	0	0	1	0	0	0
HybridVGA3	0	1	0	1	1	0	0	0
HybridWVGA1	0	0	0	0	1	1	0	1
HybridWVGA2	0	0	0	0	1	1	0	0
HybridWVGA2 no rebuf	0	1	0	1	1	1	0	0
VGA merge	0	1	0	1	1	0	0	0
HybridHD1	0	0	0	0	0	1	0	0
HybridHD2	1	0	1	0	0	0	0	0
HybridHD3	1	0	1	0	1	0	0	0
HybridHD4	0	0	0	0	0	0	0	1
HybridHD5	1	0	1	1	0	0	0	0
HD merge	0	0	0	0	1	0	0	0

Table 15. Top Performing Models: Hybrid-NR Non-Encrypted Models

	TVM-Hybrid Non-Encrypted	**VMon-B**	**YHyNR**	**PSNR**
HybridVGA1	1	1	0	0
HybridVGA2	0	0	1	0
HybridVGA3	0	0	1	1
HybridWVGA1	0	0	1	1
HybridWVGA2	0	0	1	0
HybridWVGA2 no rebuf	0	0	1	0
VGA merge	0	0	1	0
HybridHD1	0	0	1	1
HybridHD2	1	1	0	0
HybridHD3	1	1	1	0
HybridHD4	0	0	0	1
HybridHD5	1	1	0	0
HD merge	0	0	1	0

Table 16. Top Performing Models: Hybrid-NR Encrypted Models

	TVM-Hybrid Encrypted	**VMon**	**YHyNRe**	**PSNR**
HybridVGA1	1	1	0	0
HybridVGA2	1	1	1	1
HybridVGA3	1	1	0	0
HybridWVGA1	1	1	1	1
HybridWVGA2	1	1	1	1
HybridWVGA2 no rebuf	1	1	1	0
VGA merge	1	1	0	0
HybridHD1	0	0	1	0
HybridHD2	1	1	1	1
HybridHD3	0	0	1	1
HybridHD4	0	0	0	1
HybridHD5	0	1	0	0
HD merge	0	0	1	1

Table 17. Top Performing Models: NR Encrypted Models

	YNR	**PSNR**
HybridVGA1	1	1
HybridVGA2	0	1
HybridVGA3	0	1
HybridWVGA1	0	1
HybridWVGA2	1	1
HybridWVGA2 no rebuf	0	1
VGA merge	0	1
HybridHD1	0	1
HybridHD2	1	1
HybridHD3	0	1
HybridHD4	0	1
HybridHD5	1	1
HD merge	0	1

The next table compares each model's performance with PSNR's performance. Since PSNR is a full reference metric, these models are considered to have superior performance if they are able to match or beat PSNR's performance without access to the original video. Every box marked with a one (1) indicates that this model is statistically equivalent or better than PSNR with 95% confidence. This significance is computed using RMSE. Zero (0) indicates that the model has worse performance than PSNR for that dataset.

Table 18. Equivalent To or Better Than PSNR: All Hybrid-NR Non-Encrypted, Hybrid-NR Encrypted and NR Models

	TVM-Hybrid Non-Encrypted	**TVM-Hybrid Encrypted**	**VMon-B**	**VMon**	**YHyNR**	**YHyNRe**	**YNR**
HybridVGA1	1	1	1	1	1	1	1
HybridVGA2	1	1	1	1	1	1	0
HybridVGA3	0	1	0	1	1	1	0
HybridWVGA1	0	1	0	1	1	1	0
HybridWVGA2	1	1	1	1	1	1	1
HybridWVGA2 no rebuf	0	1	0	1	1	1	0
VGA merge	1	1	1	1	1	1	0
HybridHD1	1	1	1	1	1	1	0
HybridHD2	1	1	1	1	1	1	1
HybridHD3	1	0	1	0	1	1	0
HybridHD4	0	0	0	0	0	0	0
HybridHD5	1	1	1	1	1	1	1
HD merge	1	0	1	0	1	1	0

4.4 Model Analysis with DMOS

This section analyzes Hybrid-FR Encrypted, Hybrid-FR Non-Encrypted, Hybrid-RR Encrypted, Hybrid-RR Non-Encrypted and NR models using DMOS.

Please refer to the notes at the beginning of section 4.3 for information on notation.

Table 19. RMSE for Hybrid-FR Non-Encrypted, Hybrid-FR Encrypted, Hybrid-RR Non-Encrypted, and Hybrid-RR Encrypted models (DMOS)

	PEVQ-S	**PEVQ-S pes+rtp**	**PEVQ-S ts+rtp**	**YHyFR**	**YHyFRe**	**YHyRR 56k**	**YHyRR 56ke**	**YHyRR 128k**	**YHyRR 128ke**	**YHyRR 256k**	**YHyRR 256ke**	**PSNR**
HybridVGA1	0.65	0.65	0.65	0.79	0.78	0.79	0.79	0.79	0.78	—	—	0.66
HybridVGA1	0.65	0.65	0.65	0.79	0.78	0.79	0.79	0.79	0.78	—	—	0.66
HybridVGA2	0.51	0.51	0.51	0.49	0.49	0.49	0.49	0.49	0.49	—	—	0.63
HybridVGA3	0.52	0.52	0.52	0.41	0.44	0.41	0.44	0.41	0.44	—	—	0.56
HybridWVGA1	0.54	0.54	0.54	0.50	0.49	0.50	0.49	0.50	0.49	—	—	0.59
HybridWVGA2	0.53	0.53	0.53	0.39	0.41	0.39	0.42	0.39	0.41	—	—	0.60
HybridWVGA2 no rebuf	0.51	0.51	0.51	0.31	0.30	0.31	0.30	0.30	0.30	—	—	0.59
VGA merge	0.57	0.57	0.57	0.57	0.58	0.57	0.58	0.57	0.58	—	—	0.66
VGA merge no rebuf	0.55	0.55	0.55	0.58	0.59	0.58	0.59	0.58	0.59	—	—	0.66
Mean of VGA	0.55	0.55	0.55	0.52	0.52	0.52	0.53	0.51	0.52	—	—	0.61
Mean of VGA no rebuf	0.55	0.55	0.55	0.50	0.50	0.50	0.50	0.50	0.50	—	—	0.61
HybridHD1	0.34	0.34	0.34	0.41	0.38	0.41	0.38	0.41	0.38	0.41	0.38	0.42
HybridHD2	0.51	0.51	0.50	0.66	0.73	0.67	0.74	0.66	0.73	0.66	0.73	0.59
HybridHD3	0.41	0.41	0.41	0.52	0.64	0.52	0.64	0.52	0.64	0.52	0.64	0.60
HybridHD4	0.64	0.64	0.64	0.57	0.60	0.56	0.60	0.56	0.60	0.56	0.59	0.60
HybridHD5	0.50	0.50	0.50	0.46	0.38	0.46	0.39	0.46	0.39	0.46	0.38	0.72
HD merge	0.51	0.51	0.51	0.55	0.58	0.55	0.58	0.55	0.58	0.55	0.58	0.61
Mean of HD	0.48	0.48	0.48	0.52	0.55	0.52	0.55	0.52	0.55	0.52	0.55	0.59

Table 20. Pearson correlation for Hybrid-FR Non-Encrypted, Hybrid-FR Encrypted, Hybrid-RR Non-Encrypted, and Hybrid-RR Encrypted models (DMOS)

	PEVQ-S	**PEVQ-S pes+rtp**	**PEVQ-S ts+rtp**	**YHyFR**	**YHyFRe**	**YHyRR 56k**	**YHyRR 56ke**	**YHyRR 128k**	**YHyRR 128ke**	**YHyRR 256k**	**YHyRR 256ke**	**PSNR**
HybridVGA1	0.77	0.77	0.77	0.63	0.64	0.63	0.64	0.63	0.64	—	—	0.77
HybridVGA1	0.77	0.77	0.77	0.63	0.64	0.63	0.64	0.63	0.64	—	—	0.77
HybridVGA2	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.89	0.89	—	—	0.81
HybridVGA3	0.79	0.79	0.79	0.88	0.86	0.88	0.86	0.88	0.86	—	—	0.75
HybridWVGA1	0.86	0.86	0.86	0.88	0.89	0.88	0.89	0.88	0.89	—	—	0.83
HybridWVGA2	0.84	0.84	0.84	0.92	0.91	0.92	0.91	0.92	0.91	—	—	0.79
HybridWVGA2 no rebuf	0.89	0.89	0.89	0.96	0.96	0.96	0.96	0.96	0.96	—	—	0.86
VGA merge	0.81	0.81	0.81	0.80	0.79	0.80	0.79	0.80	0.79	—	—	0.72
VGA merge no rebuf	0.83	0.83	0.83	0.81	0.80	0.80	0.80	0.81	0.80	—	—	0.73
Mean of VGA	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.84	—	—	0.79
Mean of VGA no rebuf	0.84	0.84	0.84	0.85	0.85	0.85	0.85	0.85	0.85	—	—	0.80
HybridHD1	0.94	0.94	0.94	0.91	0.92	0.91	0.92	0.91	0.92	0.91	0.92	0.91
HybridHD2	0.88	0.88	0.88	0.79	0.73	0.78	0.72	0.79	0.73	0.79	0.73	0.84
HybridHD3	0.91	0.91	0.91	0.86	0.78	0.86	0.78	0.86	0.78	0.86	0.78	0.80
HybridHD4	0.83	0.83	0.83	0.87	0.86	0.88	0.86	0.87	0.86	0.88	0.86	0.86
HybridHD5	0.84	0.84	0.84	0.86	0.91	0.86	0.91	0.86	0.91	0.86	0.91	0.60
HD merge	0.88	0.88	0.88	0.86	0.85	0.86	0.85	0.86	0.85	0.86	0.85	0.83
Mean of HD	0.88	0.88	0.88	0.86	0.84	0.86	0.84	0.86	0.84	0.86	0.84	0.80

Table 21. Epsilon Independent RMSE (RMSE*) for Hybrid-FR Non-Encrypted, Hybrid-FR Encrypted, Hybrid-RR Non-Encrypted, and Hybrid-RR Encrypted models (DMOS)

	PEVQ-S	**PEVQ-S pes+rtp**	**PEVQ-S ts+rtp**	**YHyFR**	**YHyFRe**	**YHyRR 56k**	**YHyRR 56ke**	**YHyRR 128k**	**YHyRR 128ke**	**YHyRR 256k**	**YHyRR 256ke**	**PSNR**
HybridVGA1	0.37	0.37	0.37	0.49	0.48	0.49	0.48	0.48	0.48	—	—	0.41
HybridVGA1	0.37	0.37	0.37	0.49	0.48	0.49	0.48	0.48	0.48	—	—	0.41
HybridVGA2	0.28	0.28	0.28	0.26	0.26	0.26	0.27	0.26	0.27	—	—	0.39
HybridVGA3	0.29	0.29	0.29	0.20	0.22	0.20	0.22	0.20	0.22	—	—	0.31
HybridWVGA1	0.33	0.33	0.33	0.27	0.28	0.27	0.28	0.27	0.28	—	—	0.35
HybridWVGA2	0.33	0.33	0.33	0.19	0.22	0.19	0.22	0.19	0.22	—	—	0.40
HybridWVGA2 no rebuf	0.29	0.29	0.29	0.10	0.11	0.10	0.11	0.10	0.11	—	—	0.37
VGA merge	0.35	0.35	0.35	0.35	0.36	0.35	0.36	0.35	0.36	—	—	0.41
VGA merge no rebuf	0.33	0.33	0.33	0.35	0.37	0.35	0.37	0.35	0.37	—	—	0.41
VGA-WVGA	0.32	0.32	0.32	0.28	0.29	0.28	0.29	0.28	0.29	—	—	0.37
VGA-WVGA-NoRebuffering	0.31	0.31	0.31	0.26	0.27	0.26	0.27	0.26	0.27	—	—	0.37
HybridHD1	0.12	0.12	0.12	0.19	0.15	0.19	0.15	0.19	0.15	0.19	0.15	0.19
HybridHD2	0.28	0.28	0.28	0.40	0.49	0.41	0.50	0.41	0.50	0.41	0.49	0.36
HybridHD3	0.16	0.16	0.16	0.27	0.38	0.27	0.38	0.27	0.38	0.27	0.38	0.33
HybridHD4	0.42	0.42	0.42	0.35	0.38	0.35	0.38	0.34	0.38	0.34	0.38	0.37
HybridHD5	0.26	0.26	0.26	0.23	0.17	0.23	0.17	0.23	0.17	0.23	0.17	0.47
HD merge	0.29	0.29	0.29	0.32	0.35	0.32	0.35	0.32	0.35	0.32	0.35	0.37
Mean of HD	0.25	0.25	0.25	0.29	0.31	0.29	0.32	0.29	0.32	0.29	0.31	0.34

These next tables show the group of top performing models for a particular category of model. Every box marked with a one (1) indicates that this model is either the top performing model for that dataset, or is statistically equivalent to that top performing model with 95% confidence. This significance is computed using RMSE. Zero (0) indicates that the model has worse performance than the top performing model for that dataset.

Table 22. Top Performing Models: All Hybrid-FR Non-Encrypted and Hybrid-FR Encrypted Models

	PEVQ-S	**PEVQ-S pes+rtp**	**PEVQ-S ts+rtp**	**YHyFR**	**YHyFRe**	**PSNR**
HybridVGA1	1	1	1	0	1	1
HybridVGA2	1	1	1	1	1	0
HybridVGA3	0	0	0	1	1	0
HybridWVGA1	1	1	1	1	1	0
HybridWVGA2	0	0	0	1	1	0
HybridWVGA2 no rebuf	0	0	0	1	1	0
VGA merge	1	1	1	1	1	0
HybridHD1	1	1	1	0	1	0
HybridHD2	1	1	1	0	0	0
HybridHD3	1	1	1	0	0	0
HybridHD4	1	1	1	1	1	1
HybridHD5	0	0	0	0	1	0
HD merge	1	1	1	0	0	0

Table 23. Top Performing Models: Hybrid-FR Encrypted Models

	PEVQ-S pes+rtp	**PEVQ-S ts+rtp**	**YHyFRe**	**PSNR**
HybridVGA1	1	1	1	1
HybridVGA2	1	1	1	0
HybridVGA3	0	0	1	0
HybridWVGA1	1	1	1	0
HybridWVGA2	0	0	1	0
HybridWVGA2 no rebuf	0	0	1	0
VGA merge	1	1	1	0
HybridHD1	1	1	1	0
HybridHD2	1	1	0	0
HybridHD3	1	1	0	0
HybridHD4	1	1	1	1
HybridHD5	0	0	1	0
HD merge	1	1	0	0

Table 24. Top Performing Models: All Hybrid-RR Non-Encrypted and Hybrid-RR Encrypted Models

	YHyRR56k	YHyRR56ke	YHyRR128k	YHyRR128ke	YHyRR256k	YHyRR256ke	PSNR
HybridVGA1	1	1	1	1	—	—	1
HybridVGA2	1	1	1	1	—	—	0
HybridVGA3	1	1	1	1	—	—	0
HybridWVGA1	1	1	1	1	—	—	0
HybridWVGA2	1	1	1	1	—	—	0
HybridWVGA2 no rebuf	1	1	1	1	—	—	0
VGA merge	1	1	1	1	—	—	0
HybridHD1	1	1	1	1	1	1	1
HybridHD2	1	0	1	0	1	0	1
HybridHD3	1	0	1	0	1	0	0
HybridHD4	1	1	1	1	1	1	1
HybridHD5	0	1	0	1	0	1	0
HD merge	1	1	1	1	1	1	0

Table 25. Top Performing Models: Hybrid-RR Encrypted Models

	YHyRR56ke	YHyRR128ke	YHyRR256ke	PSNR
HybridVGA1	1	1	—	1
HybridVGA2	1	1	—	0
HybridVGA3	1	1	—	0
HybridWVGA1	1	1	—	0
HybridWVGA2	1	1	—	0
HybridWVGA2 no rebuf	1	1	—	0
VGA merge	1	1	—	0
HybridHD1	1	1	1	1
HybridHD2	0	0	0	1
HybridHD3	1	1	1	1
HybridHD4	1	1	1	1
HybridHD5	1	1	1	0
HD merge	1	1	1	1

The next table compares each Hybrid-FR model's performance with PSNR's performance. Since PSNR is a full reference metric, these models are considered to have superior performance if they are able to beat PSNR's performance. Every box marked with a one (1) indicates that this model is statistically better than PSNR with 95% confidence. This significance is computed using RMSE. Zero (0) indicates that PSNR is equivalent to or better than the model.

Table 26. Better Than PSNR: All Hybrid-FR Non-Encrypted and Hybrid-FR Encrypted Models

	PEVQ-S	**PEVQ-S pes+rtp**	**PEVQ-S ts+rtp**	**YHyFR**	**YHyFRe**
HybridVGA1	0	0	0	0	0
HybridVGA2	1	1	1	1	1
HybridVGA3	0	0	0	1	1
HybridWVGA1	0	0	0	1	1
HybridWVGA2	0	0	0	1	1
HybridWVGA2 no rebuf	0	0	0	1	1
VGA merge	1	1	1	1	1
HybridHD1	1	1	1	0	0
HybridHD2	1	1	1	0	0
HybridHD3	1	1	1	1	0
HybridHD4	0	0	0	0	0
HybridHD5	1	1	1	1	1
HD merge	1	1	1	1	0

The next table compares each model's performance with PSNR's performance. Since PSNR is a full reference metric, these models are considered to have superior performance if they are able to match or beat PSNR's performance without access to the original video. Every box marked with a one (1) indicates that this model is statistically equivalent or better than PSNR with 95% confidence. This significance is computed using RMSE. Zero (0) indicates that the model has worse performance than PSNR for that dataset.

Table 27. Equivalent To or Better Than PSNR: All Hybrid-RR Non-Encrypted and Hybrid-RR Encrypted Models

	YHyRR56k	**YHyRR56ke**	**YHyRR128k**	**YHyRR128ke**	**YHyRR256k**	**YHyRR256ke**
HybridVGA1	1	1	1	1	—	—
HybridVGA2	1	1	1	1	—	—
HybridVGA3	1	1	1	1	—	—
HybridWVGA1	1	1	1	1	—	—
HybridWVGA2	1	1	1	1	—	—
HybridWVGA2 no rebuf	1	1	1	1	—	—
VGA merge	1	1	1	1	—	—
HybridHD1	1	1	1	1	1	1
HybridHD2	1	0	1	0	1	0
HybridHD3	1	1	1	1	1	1
HybridHD4	1	1	1	1	1	1
HybridHD5	1	1	1	1	1	1
HD merge	1	1	1	1	1	1

4.5 Other Analysis

See Annex C for plots. See Annex D for model fit coefficients, Pearson correlation confidence intervals and RMSE-star confidence intervals.

The ILG data analysis is also available in spreadsheets of comma separated values (CSV). These files include analyses of all models on both MOS and DMOS. This allows comparisons between all types of models.

Appendix 1 Comments by Model Proponents

Note: The proponent comments are not endorsed by VQEG. They are presented in this Appendix to give the Proponents a chance to discuss their results and should not be quoted out of this context.

The proponent can present any additional analyses and refer to any portion of the official ILG data analysis within the proponent comment Appendix, as long as the statements are factually correct and do not exceed two (2) pages per model category per proponent (e.g., 2 pages for one proponent's Hybrid-NR models). Proponents can explain potential performance improvements resulting from model changes such as disallowed bug fixes and algorithm changes. Proponents can mention other advantages not included in the official ILG data analysis, such as computational complexity, outlier analysis, or performance on a subset of the data. Proponents are allowed to question decisions made within the official ILG data analysis.

A.1 Comments on PEVQ-S by OPTICOM

OPTICOM contributed an early version of PEVQ-S to the VQEG Hybrid project in 2012. This hybrid algorithm has been further advanced in the meantime to cover all kinds of novel IP based streaming techniques and is now referred to as PEVQ-S. PEVQ-S is an extension and enhanced version of PEVQ, a full reference model which is already standardized under ITU-T J.247. Due to the standardized core, the image based analysis is very well known, robust and stable. The really new component is the analysis of the bitstream which helps to improve the accuracy in some cases. This said, it should be recognized that the weighting on the final result in PEVQ-S is far greater from the image analysis than that of the bitstream analysis. The optimization of this PEVQ version was clearly targeting HD resolution, which can also be seen from the results. The performance for VGA/WVGA can certainly be further improved by some simple further tuning based on the now available larger dataset for this resolution.

PEVQ-S allows for very efficient implementations and it is already available for many different platforms. A special version which requires virtually no resources on the client side and no access to the reference signal from the client side is also available for HTTP adaptive bitrate streaming.

Our main target for optimization of the model is always balancing the worst case performance with peak performance, which means we try to reduce the severity of outliers. This is especially important for real use cases since it defines how well you can rely on the model predictions. The average performance statistics often mask that and we would not consider a model which varies between excellent and poor performance as reliable enough for standardization.

Our interpretation of the results distinguishes between the performance for VGA/WVGA and the performance for HD.

For VGA/WVGA, PEVQ-S showed excellent results in terms of the mean rmse and mean correlation. Even if not the best, it was close. On the merged superset instead it yielded the best results in all categories, admittedly almost identical to the second best model. The real strength however becomes visible when looking at the worst case behavior. Here it is obvious that of all tested models PEVQ-S has the least severe outliers. In the HD case, PEVQ-S is the clear winner and undoubtedly the best performing model for all statistical metrics.

Looking at the severity of outliers, PEVQ-S outperforms all other models for all tested resolutions and thus provides the most reliable predictions of all tested methods.

A.2 Deutsche Telekom and SwissQual: Comments to the Hybrid Models for Non-Encrypted Case

Both models, Deutsche Telekom's TVM-Hybrid Non-Encrypted and SwissQual's VMon-B, share the same basis and are described below.

Application Scenarios

The models are Hybrid-NR Non-encrypted models. They take as input the video player's decoded video and the transmitted video bitstream. These are no-reference models, no information about the reference video is used. Thus, they are well suited for monitoring solutions at the client side of services like IPTV. Compared to FR models NR models can be applied to video streams, where no knowledge about the source signal is available such as liveTV, live video sharing or live video telephony.

Model Design

The model was designed

- to be accurate and robust.
- to allow for an efficient implementation, as the model's complexity was kept as low as possible.

The model takes as input the transmitted video bitstream and the video player's decoded video. The bitstream is partially parsed, and features are extracted, which are combined with features from the decoded video to estimate an overall video quality.

There are differences between the two models in some of the internal modules. The main difference is in the computation of different features of the decoded video frames for the estimation of slicing degradations.

The models were mainly developed by Savvas Argyropoulos, Peter List, and Anna Llagostera.

Proponent opinion only. This page is not an official VQEG statement.

Model Performance

Table 28: *RMSE values for selected data sets for the models TVM-Hybrid Non-Encrypted and VMon-B.*

	TVM-Hybrid Non-Encrypted	VMon-B
HD1	0.51	0.53
HD2	0.39	0.43
HD3	0.44	0.42
HD5	0.43	0.42

Within the hybrid project of VQEG some datasets have been focused on typical IPTV scenarios deployed in today's networks. For a Non-Encrypted NR model especially applications in an IPTV settop-box or other interfaces are interesting, where a decrypted bitstream is available. The models were mainly trained for HD applications and for the most relevant settings in today's applications, especially GOP sizes of 1-10s. Concentrating on data sets created with such encoding settings, which are frequently used in target applications as IPTV services, the results in RMSE performance are given in Table 28, with an **average RMSE of 0.443** (MOS) for **TVM-Hybrid Non-Encrypted and 0.451** (MOS) **for VMon-B**. This **performance is very close to the performance of the best full reference model** with a RMSE of 0.438 (DMOS) on this data. A Pearson **correlation** coefficient of **84-92%** is reached for **TVM-Hybrid Non-Encrypted** and **83-91%** for **VMon-B**. Thus, these models have the advantage of the wide application range inherent to no-reference models with a small decrease in performance compared to full reference models.

Proponent opinion only. This page is not an official VQEG statement.

A.3 Deutsche Telekom and SwissQual: Comments to the Hybrid Models for Encrypted Case

Both models, Deutsche Telekom's TVM-Hybrid Encrypted and SwissQual's VMon, share the same basis and are described below.

Application Scenarios

The models are Hybrid-NR Encrypted models. They take as input the video player's decoded video and the packet headers of the transmitted video bitstream, and can operate if the payload data is encrypted. These are no-reference models, no information about the reference video is used. Thus, they are well suited for monitoring solutions at the client side of services like IPTV. Compared to FR models NR models can be applied to video streams, where no knowledge about the source signal is available such as liveTV, live video sharing or live video telephony.

It should be noted that an Encrypted NR model is based on header information and requires considerably less computational effort. This type of model can of course also be applied to a non-encrypted video bitstream, it will just restrict to the header information and keep the advantage in computational effort.

Model Design

The model was designed

- to be accurate and robust.
- to allow for an efficient implementation, as the model's complexity was kept as low as possible.
- to be extensible to other video coding algorithms beyond H.264.
- to be extensible to other transmission protocols beyond RTP.

The model takes as input the packet headers of the transmitted video bitstream and the video player's decoded video. The headers are parsed and frame types and frame sizes are estimated. Spatio-temporal features are extracted from the video frames and combined with the information of the bitstream headers, to estimate a coding and a transmission quality. These two are combined to an estimation of the overall video quality.

There are differences between the two models in some of the internal modules. The two biggest differences are in:

- the different strategies for rescaling estimation
- the computation of different features of the decoded video frames for the estimation of slicing degradations.

The two models were mainly developed by Silvio Borer, Anna Llagostera, and Savvas Argyropoulos.

Model Performance

Since a model for an encrypted bitstream does not have access to all details of the encoding settings, the models were trained for the most relevant settings in today's applications: GOP sizes of 1-10s, 1-15 slices, and reasonable MTU sizes.

Applications of an Encrypted NR model are mainly in scenarios where encrypted video-streams are transmitted as in IPTV or in almost all video-on-demand services and in scenarios where a lean model with low computational effort is required. Because of the restriction to header information the Encrypted NR models show considerable advantages in resource consumption.

Table 29: *RMSE values for selected data sets for the models TVM-Hybrid Encrypted and VMon.*

	TVM-Hybrid Encrypted	VMon
VGA1	0.51	0.53
VGA2	0.59	0.57
VGA3	0.54	0.50
HD1	0.47	0.47
HD5	0.49	0.42

Concentrating on data sets created with such encoding settings, which are frequently used in target applications, the results in RMSE performance are given in

Applications of an Encrypted NR model are mainly in scenarios where encrypted video-streams are transmitted as in IPTV or in almost all video-on-demand services and in scenarios where a lean model with low computational effort is required. Because of the restriction to header information the Encrypted NR models show considerable advantages in resource consumption.

Table 29, with an **average RMSE of 0.498** (MOS) **for VMon and 0.521** (MOS) **for TVM-Hybrid Encrypted**. This **VMon performance is almost equal to the best full reference model** (encrypted and non-encrypted) with a RMSE of 0.496 (DMOS) on this data. For TVM-Hybrid Encrypted it is within a 5% increase. A Pearson **correlation** coefficient of **81-91%** for **VMon** and

77-93% is reached for **TVM-Hybrid Encrypted**. Thus, these models have the advantage of the wide application range inherent to no-reference models with a minimal decrease in performance compared to full reference models.

A.4 Yonsei model description

Complexity & application areas

Yonsei hybrid NR models for encrypted bit streams (YHyNRe) mainly use bitrates and packet loss information with additional post-processing. The method is simple and fast enough to be implemented in real time using moderate processors. Yonsei hybrid RR models for encrypted bit streams (YHyRRe) use RR models along with bitrates and packet loss information. Yonsei hybrid FR models for encrypted bit streams (YHyFRe) use FR models along with bitrates and packet loss information. Yonsei hybrid NR models for non-encrypted bit streams (YHyNR) mainly use codec parameters and transmission error information with additional post-processing. The method is simple and fast enough to be implemented in real time using moderate processors. Yonsei hybrid RR models for non-encrypted bit streams (YHyRR) use RR models along with codec parameters and transmission error information. Yonsei hybrid FR models for non-encrypted bit streams (YHyFR) use FR models along with codec parameters and transmission error information.

Yonsei hybrid models are almost transparent to protocols and can be easily modified for any streaming technologies.

Errors and performance improvement

There is an error for the HD models for non-encrypted bit streams and the performance can be improved with the error fixed.

In case of VGA1 (rebuffering case), there was a bug with packet number counting. When packets were retransmitted, it caused minus increments that were erroneously interpreted as a very large number of lost packets. This error produced a large number of outliers in all the models (hybrid NR, hybrid RR, hybrid FR). If this packet counting error is fixed, significant improvement will be obtained for all the models.

Performance evaluation

Yonsei hybrid NR models showed very good performance in all categories. With the packet counting error fixed, the overall performance will be also improved.

Yonsei hybrid RR models showed outstanding performance in all categories. For VGA/WVGA, the RR models are statistically better than PSNR except for VGA1. With the packet counting error fixed, the RR models will be better than PSNR. For HD, the RR models are statistically better than or equivalent to PSNR. Yonsei hybrid RR models showed almost identical performance as the FR models.

Proponent opinion only. This page is not an official VQEG statement.

Yonsei hybrid FR models showed outstanding performance. For VGA/WVGA, the FR models showed the best performance in four sessions (VGA2, VGA2, WVGA1, WVGA2) in terms of RMSE values (correlation: 0.88~0.96) except for VGA1. With the obvious packet counting bug fixed, significant improvement will be obtained for VGA1. For HD, the FR models showed excellent performance (the average correlation of the five sessions: 0.84~0.86). With the HD bug fixed, the overall performance will improve.

Annex A Model Bug Fixes

A.1.1 Deutsche Telekom and SwissQual

For the Hybrid-NR Non-encrypted models the following bugs were fixed:

- In this bug, for every loss event both the slicing and freezing impairments associated with this loss are computed; after the fix, only one type of distortion (slicing or freezing) is evaluated per loss event.
- In this bug, during the decoding of the H.264/AVC bitstream, some macroblocks were wrongly decoded as if they were in SKIP mode; after the fix, the macroblocks are decoded correctly.
- In this bug, the transmission impairments in a GOP extending beyond the end of the edited PVS were not considered; after the fix, the packet-loss related features of the last GOP are aggregated at the end of the edited PVS and the transmission impairment is correctly considered.

A.1.2 OPTICOM

Fix RTP Header parser to be able to deal with RTP Header extensions

The routine for parsing the header fields of every RTP packet does not take the length of the RTP header extension into account correctly. This means the parsing function is out of sync with the bit stream after the first RTP packet. Which means bit stream cannot be analyzed any further and the model returns an error code and no quality estimation.

The fix implements reading the length of RTP header extension and the skipping of the actual extension fields. No additional information is extracted from the bit stream.

Fix for RTP Timestamp Evaluation

The routine for evaluation of the RTP time stamp present in the RTP header was optimized for speed and only evaluates a certain ratio of all received packets. The ratio is fixed value of 20%.

In some rare cases of the newly available bit streams the packet loss pattern in the bit stream does not allow for the RTP time stamp evaluation, e.g. the scheme of the RTP time stamp (DTS or PTS) cannot be evaluated. This leads the model to exit with error code without quality estimation.

The fix increases the ratio of packets to be analyzed to 100%.

The fix does not influence cases where it was possible to validate the time stamp before the fix. The fix does not influence the detection or evaluation of packet loss.

Error in RTP Sequence number pre processing

In order to detect packet loss the RTP sequence number is evaluated. For some HD cases the sequence number wraps around from 65535 to 0 due to the limited word length of the field (16bit).

The function to reorder packets that have been received out of order applies unwrapping of the RTP sequence number as pre-processing step. That means after unwrapping every RTP packet is assigned a strictly monotonically increasing number.

In some rare cases the unwrapping fails which leads to errors while reordering the packets which results in exit of the model at a later step without quality estimation.

The bug fix targets the unwrapping function.

A.1.3 Yonsei

1. **Parsing problems**

   A. **The parsing program (similar to H.264annexbextractor) did not properly handle RTSP cases.** The stand-alone parsing program that produces H.264 Annex.B streams from PCAP files did not properly handle RTSP cases. The bug was fixed by replacing the parsing program and deleting some lines in the main program.

   B. **The submitted models did not properly handle RTP header extension cases.** The header processing routine was modified to handle the extension cases.

   C. **RTP header parsing error**. When the RTP counter is very large, it caused memory error. It was fixed by setting the limit to the RTP counter.

   D. **The submitted models did not work properly when an IP packet has only one TS**. This was fixed by changing a single numerical value (threshold value) for protocol classification (TS/RTP/UDP/IP or RTP/UDP/IP).

E. **Port number mishandling**. This bug was fixed by using the correct port number variable.

2. Memory errors

A. **Memory allocation error**. Since some validation data had IP packets with one TS, the number of packets exceeded the hard limit of memory allocation. It was fixed by increasing the memory assignment.

B. **Memory allocation error in FR module**. In some cases with rebuffering, the PVS length was longer than the corresponding SRC length. Since the FR module allocated memory based on the SRC length, it caused memory error. Since the source programs were not submitted, the FR module was replaced by the RR module for the PVSs with rebuffering.

C. **Initialization without memory allocation for audio data array**: Since audio was not tested in the Hybrid Project, this bug was fixed by not using the audio data array.

D. **Modified ffmpeg crash**. The non-encrypted models had a memory overflow problem for three PCAP files, which could have been fixed by simply checking the upper-bound. Since the source programs were not submitted for the modified ffmpeg, the corresponding encrypted models were used for the three PCAP files.

Annex B Subjective Experiment Designs

B.1HybridHD01

<u>Types of impairments:</u> x264 encoding / simulated loss (uniform-bursty distributions, low/medium/high packet loss rates) / VLC and T-Labs decoder.

The design was done by DT and the PVS were provided by RT-RK.

<u>Video format :</u> 1080i60

<u>Problems :</u> The sequences contained mostly network impairments and were therefore generally of very low quality. In addition, the impairments occurred in a very systematic way which may have biased the observers.

<u>Subjective Testing:</u>

The observers voted with an ACR score of 1 (bad) or 2 (poor) in 55.8% of the cases (for a uniform distribution, it would be expected that 40% of the votes are 1 or 2). The Mean Opinion Score was below 2 in 38.75% and below 3 in 74.38% of the cases (for a uniform distribution, it would be expected that the MOS is below 2 in 25% and below 3 in 50% of the cases).

The subjective test was carried out at the IRCCyN Laboratories in December 2013 with the participation of 24 subjects (12 males and 12 females), a majority being university or high school students and aging from 18 to 45; all were positively screened for visual acuity and color vision.

Training was done using the same video clips used for the test, taking care to select samples over the full range of quality and most of the SRC material used in the actual test.

The test was done using a ClearView Extreme hardware player dedicated to play uncompressed HD video contents in real time. The display was a 40” professional TV Logic LVM401 monitor used in its native HD resolution.

The subjects were seated at a distance of 3H.

The viewing environment respects the ITU-R BT500-11 recommendation with a screen max luminance of 180 cd/m² with a background luminance of 25 cd/m².

Hybrid Validation Test Design			
Lab	Deutsche Telekom		
Resolution	HD - 1920x1080 - interlaced		
Frame Rate	60 fields per second		
Rebuffering	No		
Encoder	x264		
Video Player	Media Player		
Network	LAN (packetization using Sirannon)		
Network impairment	simulated loss (uniform-bursty distributions, low/medium/high packet loss rates)		
Bit-Rates	low (1-4 Mbps)	medium (4-8) Mbps	high (8-30 Mbps)
Decoder	T-Labs Decoder	VLC	
Concealment Type	Slicing	Freezing	default

Notes

10 SRC x (15 HRC + Reference)

Each SRC 14 sec for HRC creation. SRC & PVS edited to 10 sec for subjective testing.

HRC Details (first of two tables)

description	encoder	GOP-type (*)	slices per frame	Network	Timestamp	Network Impairment
reference	NA	NA		NA	NA	NA
No-loss, high bit-rate	x264	M=3, N=30	1	LAN	non-continuous	No
No-loss, medium bit-rate, 720p source	x264	M=3, N=30	1	LAN	non-continuous	No

No-loss, low bit-rate	x264	M=3, N=30	1	LAN	non-continuous	No
High bit-rate, uniform low loss, slicing	x264	M=3, N=30	1	LAN	non-continuous	uniform, 0.125%
High bit-rate, uniform low loss, freezing	x264	M=3, N=30	1	LAN	non-continuous	uniform, 0.125%
High bit-rate, bursty low loss, default	x264	M=4, N=60	1	LAN	continuous	bursty, 0.125%
High bit-rate, uniform medium loss, default	x264	M=4, N=60	1	LAN	continuous	uniform, 0.25%
Medium bit-rate, uniform bursty loss, slicing	x264	M=3, N=30	1	LAN	non-continuous	bursty, 0.5%
Medium bit-rate, uniform low loss, slicing	x264	M=4, N=60	3	LAN	non-continuous	uniform, 0.125%
Medium bit-rate, medium bursty loss, freezing, 720p source	x264	M=4, N=60	3	LAN	continuous	bursty, 0.25%
Medium bit-rate, medium bursty loss, default	x264	M=3, N=30	68	LAN	continuous	bursty, 0.5%
Low-bit rate, uniform low-loss, default	x264	M=4, N=60	68	LAN	non-continuous	uniform, 0.125%
Low-bit rate, bursty medium low-loss, default	x264	M=4, N=60	68	LAN	non-continuous	bursty, 0.25%
Low-bit rate, uniform medium-loss, slicing, 720p source	x264	M=3, N=30	1	LAN	continuous	uniform, 0.25%
Low-bit rate, burst high-loss, freezing	x264	M=3, N=30	1	LAN	continuous	burst, 0.4%

HRC Details (second of two tables)

description	Bit-Rate	Decoder	Concealment Type	player	hrc number	number PVS
reference	NA	NA	NA	NA	HRC01	10
No-loss, high bit-rate	high	VLC	No	Media Player	HRC02	10
No-loss, medium bit-rate, 720p source	medium	VLC	No	Media Player	HRC03	10

No-loss, low bit-rate	low	VLC	No	Media Player	HRC04	10
High bit-rate, uniform low loss, slicing	high	T-Labs Decoder	slicing	Media Player	HRC05	10
High bit-rate, uniform low loss, freezing	high	T-Labs Decoder	freezing	Media Player	HRC06	10
High bit-rate, bursty low loss, default	high	VLC	default	Media Player	HRC07	10
High bit-rate, uniform medium loss, default	high	VLC	default	Media Player	HRC08	10
Medium bit-rate, uniform bursty loss, slicing	medium	T-Labs Decoder	slicing	Media Player	HRC09	10
Medium bit-rate, uniform low loss, slicing	medium	T-Labs Decoder	slicing	Media Player	HRC10	10
Medium bit-rate, medium bursty loss, freezing, 720p source	medium	T-Labs Decoder	freezing	Media Player	HRC11	10
Medium bit-rate, medium bursty loss, default	medium	VLC	default	Media Player	HRC12	10
Low-bit rate, uniform low-loss, default	low	VLC	default	Media Player	HRC13	10
Low-bit rate, bursty medium low-loss, default	low	VLC	default	Media Player	HRC14	10
Low-bit rate, uniform medium-loss, slicing, 720p source	low	T-Labs Decoder	slicing	Media Player	HRC15	10
Low-bit rate, burst high-loss, freezing	low	T-Labs Decoder	freezing	Media Player	HRC16	10

RTP timestamp type	HRC(Loss)	
continuous	6	50%
non-continuous	6	50%

Number of slices per frame	HRC(ALL)	
1	10	67%
68 (1 slice per macroblock row)	3	20%
3	2	13%

Loss type	HRC(ALL)	
No-loss	3	20%
Random	6	40%
Burst	6	40%

	HRC(ALL)	
N/A	3	20%
SLICING	4	27%
FREEZING	3	20%
default	5	33%

(*)GOP-Type: M denotes the distance of the P-frames, N denotes the length of the GOP, e.g. M=3, N=12 would result in a GOP-pattern IbbPbbPbbPbbI
Note that encoders have the possibility to override this setting depending on the content dynamics.
If M=4 reference-B frames can be used for hierarchical B-frame coding.

B.2 HybridHD02

HD2 experiment presents:

- Typical H264 over UDP streaming scenarios.
- Different bit rates, from 2-3Mbit/s to 15 Mbit/s.
- Trans coding from lower bit rate to higher bit rate.
- Packet losses (from 5-10 packets up to 0.125%).
- Relatively short GOP structure (12 or 15 frames in a single GOP).
- Short IP packets (242 bytes long).

RT-RK have generated the PVSs. AGH ran subjective experiment. The RT-RK toolchain used to create PVSs included:

- ffmpeg for encoding/decoding purposes
- Sirannon as a network streamer
- Wireshark as a .pcap capture software
- the Yonsei conversion software (PcapTS)
- the Markov model-based packet loss simulation software from Telchemy

The test material was played from a Blu-ray, therefore deinterlacing was done by the TV screen as it should be. In order to play the material from Blu-ray, it was slightly compressed to fit the Blu-ray specification.

The 24 valid subjects were chosen from a larger pool of subjects provided by hiring a company. The AGH specification contained balanced number of people within three different age groups, both genders and education. Nevertheless, the number of dropped subjects (because of vision test or detection of some answers inconsistency) were not perfectly equally spread for different groups. Finally the average age was 32 years old with minimum of 20 and maximum of 54. A questionnaire run after the experiment showed that the subject's typical way of watching a video content was different, from mobile phone or computer (almost only movies) and not having TV at home to watching a TV programs regularly each day.

Hybrid Validation Test Design	
Lab	AGH
Resolution	HD - 1920x1080 - interlaced
Frame Rate	50 fields per second
Rebuffering	No
Encoder	x264
Video Player	VLC
Network	LAN (packetization using Sirannon)
Network impairment	simulated loss (with network backbone traffic)

Bit-Rates	low (1-4 Mbps), medium (4-8) Mbps
Decoder	ffmpeg
Concealment Type	Default

Notes
10 SRC x (15 HRC + Reference)
Each SRC 14 sec for HRC creation. SRC & PVS edited to 10 sec for subjective testing.

HRC Details (first of two tables)

description	encoder	GOP-type (*)	First bitrate	Second bitrate	Resolution send	slices per frame	Network
reference	NA	NA	NA	NA	NA	NA	NA
No-loss, high bit-rate	ffmpeg	M=2, N=12	15 Mbps	15 Mbps	1920x1080	1	LAN
No-loss, medium bit-rate, 720p source	ffmpeg	M=3, N=15	6 Mbps	12 Mbps	1920x1080	1	LAN
No-loss, low bit-rate	ffmpeg	M=2, N=12	3 Mbps	3 Mbps	960x1080	1	LAN
High bit-rate, uniform low loss, slicing	ffmpeg	M=2, N=12	12 Mbps	12 Mbps	1920x1080	34	LAN
High bit-rate, uniform low loss, freezing	ffmpeg	M=2, N=12	12 Mbps	12 Mbps	1920x1080	8	LAN
High bit-rate, bursty low loss, default	ffmpeg	M=3, N=15	2 Mbps	2 Mbps	1920x1080	8	LAN
High bit-rate, uniform medium loss, default	ffmpeg	M=3, N=15	2 Mbps	2 Mbps	960x1080	34	LAN
Medium bit-rate, uniform bursty loss, slicing	ffmpeg	M=2, N=12	2 Mbps	2 Mbps	1920x1080	8	LAN
Medium bit-rate, uniform low loss, slicing	ffmpeg	M=3, N=15	6 Mbps	6 Mbps	1920x1080	34	LAN
Medium bit-rate, medium bursty loss, freezing, 720p source	ffmpeg	M=3, N=15	6 Mbps	6 Mbps	1920x1080	1	LAN
Medium bit-rate, medium bursty loss, default	ffmpeg	M=2, N=12	6 Mbps	13 Mbps	1920x1080	8	LAN
Low-bit rate, uniform low-loss, default	ffmpeg	M=3, N=15	3 Mbps	3 Mbps	960x1080	8	LAN
Low-bit rate, bursty medium low-	ffmpeg	M=3, N=15	3 Mbps	7 Mbps	1920x1080	34	LAN

loss, default							
Low-bit rate, uniform medium-loss, slicing, 720p source	ffmpeg	M=2, N=12	3 Mbps	3 Mbps	1920x1080	34	LAN
Low-bit rate, burst high-loss, freezing	ffmpeg	M=2, N=12	3 Mbps	3 Mbps	1920x1080	34	LAN

HRC Details (second of two tables)

description	Timestamp	Network Impairment	Bit-Rate	Decoder	Concealment Type	player	hrc number	# PVS
reference	NA	NA	NA	NA	NA	NA	HRC00	10
No-loss, high bit-rate	default	No	high	ffmpeg	No	Media Player	HRC01	10
No-loss, medium bit-rate, 720p source	default	No	medium	ffmpeg	No	Media Player	HRC02	10
No-loss, low bit-rate	default	No	low	ffmpeg	No	Media Player	HRC03	10
High bit-rate, uniform low loss, slicing	default	10-20 packets	high	ffmpeg	default	Media Player	HRC04	10
High bit-rate, uniform low loss, freezing	default	5-10 packets	high	ffmpeg	default	Media Player	HRC05	10
High bit-rate, bursty low loss, default	default	No	low	ffmpeg	default	Media Player	HRC06	10
High bit-rate, uniform medium loss, default	default	10-20 packets	low	ffmpeg	default	Media Player	HRC07	10
Medium bit-rate, uniform bursty loss, slicing	default	5-10 packets	low	ffmpeg	default	Media Player	HRC08	10
Medium bit-rate, uniform low loss, slicing	default	0.125%	medium	ffmpeg	default	Media Player	HRC09	10
Medium bit-rate, medium bursty loss, freezing, 720p source	default	5-10 packets	medium	ffmpeg	default	Media Player	HRC10	10
Medium bit-rate, medium bursty loss, default	default	5-10 packets	medium	ffmpeg	default	Media Player	HRC11	10
Low-bit rate, uniform low-loss, default	default	0.125%	low	ffmpeg	default	Media Player	HRC12	10
Low-bit rate, bursty medium low-loss, default	default	0.125%	low	ffmpeg	default	Media Player	HRC13	10
Low-bit rate, uniform medium-loss, slicing, 720p source	default	10-20 packets	low	ffmpeg	default	Media Player	HRC14	10

Low-bit rate, burst high-loss, freezing	default	0.25%	low	ffmpeg	default	Media Player	HRC15	10

RTP timestamp type	HRC(Loss)	
continuous	0	0%
non-continuous	0	0%

Number of slices per frame	HRC(ALL)	
1	4	27%
8	5	33%
34	6	40%

Loss type	HRC(ALL)	
No-loss	4	27%
Random	0	0%
Burst	12	80%

	HRC(ALL)	
N/A	3	20%
SLICING	0	0%
FREEZING	0	0%
default	12	80%

B.3 HybridHD03

B.3.1Summary

The database HD3 for the subjective experiment at VQEG-Hybrid includes HD-1080p sequences (1920x1080 progressive scan) at a frame rate of 30 frames per second. The impairments include coding distortions and network distortions with different types of packet loss distributions (random and bursty losses). Two different error concealment strategies are applied which result in “slicing” and “freezing” artifacts. As with all HD experiments, the sequences do not include rebuffering degradations, and are encapsulated with the MPEG2-TS/RTP/UDP/IP protocol stack.

B.3.2 Test Design Details

Ten source sequences were selected by the ILGs. The HRCs are listed in the table in the next section. The following properties are the same for all conditions:

Encoder: The x264 encoder (version 1867) was used for the encoding of all sequences.

Decoder: A proprietary H.264 decoder, developed by Deutsche Telekom, was employed for the decoding of the sequences. The decoder can specify whether "slicing" or "freezing" concealment should be applied in case of packet losses. In case of "slicing" concealment, the decoder applies zero-motion previous copy concealment.

Player: The Windows Media Player was used for the playback of the sequences to ensure minimal post-processing of the sequences.

Post-processing: No post-processing was applied.

The table columns are described in the following.

GOP-type: The GOP-type specifies the structure of the Group of Pictures for the encoding of the sequences. The symbol *N* denotes the distance between I-frames and *M* denotes the distance between frames that can be referenced (or equivalently, the number of bi-predictive frames). When *M* is equal to 4 hierarchical B-frame coding was employed and a reference-B was used.

e.g.: (M,N)= (3,30) : IbbPbbP….

(M,N)= (4,30) : IbBbPbBbP…. (where uppercase *B* denotes the reference-B frame)

Slice/frame: Denotes the number of slices per frame: 1 means that the whole frame is encoded in one slice, 3 means that three equally-sized slices were used, and 68 means that each macroblock-row was encoded as a separate slice.

Network: All network degradations were simulated as packet losses which are typical in IP networks. The packetized bitstreams were generated using Sirannon 1.0. As with all HD experiments, the MPEG-2 TS/RTP/UDP/IP protocol stack was employed.

Timestamp (TS): Denotes whether the RTP timestamp is continuous or non-continuous:

- NCTS: non-continuous
- CTS: continuous

Network impairment: Two types of distributions were selected for the selection of packets to be dropped from the bitstream. In the “uniform” pattern, packets to be dropped are selected randomly. In the “bursty” pattern, a 4-state Markov model is employed to select the packets to be dropped.

Three different packet loss ratios were used:

- low: 0.125%
- medium: 0.25%
- high: 0.5%

Bitrate: Denotes the encoding bitrate of the sequences. It must be noted that one-pass encoding with the specified average bitrate was employed, using the rate control of the x264 encoder.

- Low: 1.5 Mbps
- Medium: 4 Mbps
- High: 8 Mbps

Note: The actual bitrate may be slightly different than the specified (input) bitrate based on the rate control of the encoder and the content of the source sequence.

Concealment Type: In “slicing”, the decoder conceals the area which is affected due to the lost packets by neighboring information from the previous frames (zero-motion previous copy concealment). In the “freezing” concealment type, the decoder *freezes* the display output to the last correctly received frame until a new intact frame has been received.

B.3.3 Test Design Summary

description	GOP-type	slices /frame	Network	TS	Network Impairment	Bit-Rate	Concealment Type	HRC number	number PVS
reference	NA	NA	NA	NA	NA	NA	NA	HRC01	10

No-loss, high bit-rate	M=3, N=30	1	LAN	NCTS	No	high	None	HRC02	10
No-loss, medium bit-rate	M=3, N=30	1	LAN	NCTS	No	medium	None	HRC03	10
No-loss, low bit-rate	M=3, N=30	1	LAN	NCTS	No	low	None	HRC04	10
High bit-rate, uniform low loss, slicing	M=3, N=30	1	LAN	NCTS	uniform, 0.125%	high	Slicing	HRC05	10
High bit-rate, uniform low loss, freezing	M=3, N=30	1	LAN	NCTS	uniform, 0.125%	high	Freezing	HRC06	10
High bit-rate, bursty low loss, slicing	M=4, N=60	1	LAN	CTS	bursty, 0.125%	high	Slicing	HRC07	10
High bit-rate, uniform medium loss, slicing	M=4, N=60	1	LAN	CTS	uniform, 0.25%	high	Slicing	HRC08	10
High bit-rate, uniform bursty loss, freezing	M=3, N=30	1	LAN	NCTS	bursty, 0.25%	high	Freezing	HRC09	10
Medium bit-rate, uniform low loss, slicing	M=3, N=30	68	LAN	NCTS	uniform, 0.125%	medium	Slicing	HRC10	10
Medium bit-rate, medium bursty loss, slicing	M=3, N=30	68	LAN	CTS	bursty, 0.25%	medium	Slicing	HRC11	10
Medium bit-rate, medium bursty loss, freezing	M=3, N=30	68	LAN	CTS	bursty, 0.5%	medium	Freezing	HRC12	10
Low-bit rate, uniform low-loss, slicing	M=4, N=60	3	LAN	NCTS	uniform, 0.125%	low	Slicing	HRC13	10
Low-bit rate, bursty medium low-loss, freezing	M=4, N=60	3	LAN	NCTS	bursty, 0.25%	low	Freezing	HRC14	10
Low-bit rate, uniform medium-loss, slicing	M=3, N=30	1	LAN	CTS	uniform, 0.25%	low	Slicing	HRC15	10
Low-bit rate, uniform medium-loss, freezing	M=3, N=30	1	LAN	CTS	uniform, 0.25%	low	Freezing	HRC16	10

B.3.4 Subjective Tests

All playlists were randomly generated to avoid any bias from the presentation of the stimuli to the viewers.

Post-experiment screening of evaluators was applied, according to the Annex IV of the test plan. Five subjects were removed.

B.4 HybridHD04

B.4.1 Summary

The HybridHD4 1080p25 database consists of video sequences containing both encoding-only artifacts and degradations caused by packet losses during video streaming over UDP.

Two different encoders were used: the open-source x264 encoder and the commercially available MainConcept encoder. As part of this database, some more advanced features of H.264 video encoding were used including Intra-refresh, open GOP structures, and hierarchical B-pictures.

For streaming the encoded video sequences and capturing the network trace files, the open source modular multimedia streamer Sirannon was used (http://sirannon.atlantis.ugent.be).

B.4.2 Test Design specifications

Lab:	Ghent University - IBBT
Resolution:	1080p
Frame Rate:	25
Rebuffering:	No
Encoder:	x264 and Mainconcept
Video Player :	NA
Network:	IP
Network impairment:	low, medium, high [gilbert model]
Bit-Rates:	1, 3, 5, 10, 15, QP 37, QP 42
Decoder:	JM
Error Correction:	Frame Copy

B.4.3 Detailed HRC specifications

description	encoder	Video Player	Network	Network Impairment	Bit-Rate	Decoder	Error Correction	player	HRC number	Number PVS
reference	NA	NA	NA	NA	NA	NA	NA	NA	HRC01	10
	x264	NA	IP	NA	5 Mbps	JM	Frame Copy	NA	HRC02	10
	Mainconcept	NA	IP	NA	3 Mbps	JM	Frame Copy	NA	HRC03	10
intra refresh (refresh period=15)	x264	NA	IP	low	15 Mbps	JM	Frame Copy	NA	HRC04	10
intra refresh (refresh period=15)	x264	NA	IP	high	5 Mbps	JM	Frame Copy	NA	HRC05	10
open gop (5 b-frames, refresh period = 18)	x264	NA	IP	low	10 Mbps	JM	Frame Copy	NA	HRC06	10
open gop (5 b-frames, refresh period = 18)	x264	NA	IP	medium	15 Mbps	JM	Frame Copy	NA	HRC07	10
hierarchical B (5 b-frames, refresh period = 36)	x264	NA	IP	high	10 Mbps	JM	Frame Copy	NA	HRC08	10
hierarchical B (5 b-frames, refresh period = 36)	x264	NA	IP	low	3 Mbps	JM	Frame Copy	NA	HRC09	10
IPb (refresh period = 32)	x264	NA	IP	low	5 Mbps	JM	Frame Copy	NA	HRC10	10
IPb (refresh period = 32)	x264	NA	IP	medium	10 Mbps	JM	Frame Copy	NA	HRC11	10
Mainconcept recommended default	Mainconcept	NA	IP	low	3 Mbps	JM	Frame Copy	NA	HRC12	10
Mainconcept recommended default	Mainconcept	NA	IP	NA	1 Mbps	JM	Frame Copy	NA	HRC13	10
IPbbb (refresh period = 20, 2 slices)	x264	NA	IP	low	QP 37	JM	Frame Copy	NA	HRC14	10
IPb (refresh period = 20, 4 slices)	x264	NA	IP	low	QP 42	JM	Frame Copy	NA	HRC15	10
IPbbb (refresh period = 20, 8 slices)	Mainconcept	NA	IP	high	10 Mbps	JM	Frame Copy	NA	HRC16	10

B.5 HybridHD05

The test design, PVS generation and subject test of this test was performed by Yonsei University.

B.5.1 Test design

10 SRCs x 16 HRCs = 160 PVSs (SRC exclude)	
1 Codec	H.264
11 Bitrates (Mbps)	2, 2.5, 3, 3.5, 4, 5, 5.5, 6, 7, 9.5, 14
1 Frame Rate (fps)	30
3 PLRs	Low (0.1%-0.5%), med (0.5%-1.0%), high (1.0%-1.5%)
2 PLCs	slicing, freezing
7 loss ratio (%)	0.1, 0.3, 0.4, 0.55, 0.65, 1.0, 1.3
2 loss types	random, burst
2 Re-size Ratio	1/3, 1/2
Encoder	x264
Server	Sirannon
Decoder	ffmpeg, post-processing(freezing)
Simulated loss	Packet Loss Simulator
Player	VLC Player

HRC	Protocol	Codec	Bit-Rates (kbps)	Frame-Rates	PLR	PLC	Loss rate (%)	Loss type	Slices per frame	Re-scaling
1	MPEG2-TS/RTP/UDP/IP	H.264	14	30	0	0	0	none	68	
2	MPEG2-	H.264	9.5	30	0	0	0	none	68	

	TS/RTP/UDP/IP									
3	MPEG2-TS/RTP/UDP/IP	H.264	7	30	0	0	0	none	68	
4	MPEG2-TS/RTP/UDP/IP	H.264	5.5	30	0	0	0	none	68	
5	MPEG2-TS/RTP/UDP/IP	H.264	3.5	30	0	0	0	none	68	
6	MPEG2-TS/RTP/UDP/IP	H.264	2	30	0	0	0	none	68	
7	MPEG2-TS/RTP/UDP/IP	H.264	2.5	30	0	0	0	none	68	1/3
8	MPEG2-TS/RTP/UDP/IP	H.264	3	30	low	slicing	0.1	burst	68	1/2
9	MPEG2-TS/RTP/UDP/IP	H.264	6	30	low	freezing	0.1	burst	68	
10	MPEG2-TS/RTP/UDP/IP	H.264	2.5	30	low	slicing	0.3	burst	68	
11	MPEG2-TS/RTP/UDP/IP	H.264	7	30	low	slicing	0.4	burst	68	
12	MPEG2-TS/RTP/UDP/IP	H.264	5	30	med	freezing	0.55	random	68	
13	MPEG2-TS/RTP/UDP/IP	H.264	3.5	30	med	slicing	0.65	burst	68	
14	MPEG2-TS/RTP/UDP/IP	H.264	9.5	30	med	slicing	0.65	random	68	
15	MPEG2-TS/RTP/UDP/IP	H.264	4	30	high	slicing	1	burst	68	
16	MPEG2-TS/RTP/UDP/IP	H.264	6	30	high	slicing	1.3	random	68	

B.5.2 Subjective Test

A total of 24 valid subjects, who positively screened for visual acuity and color vision, were chosen in the subjective test. A majority was university students aging from 20 to 32. At the beginning of each experiment, five sample video sequences over the full range of quality were shown to viewers during the training session so that they would be familiar with the assessment procedures before the actual sessions. The 42" professional LCD monitor (SONY LMD-4250W) was used in the tests. The subjects were located at 3H, where H represents the display height.

B.6 HybridVGA01

B.6.1Summary

This test focuses on **live** video recording of video streams transmitted over a commercially operated 3G mobile network or transmitted over LAN with simulated network impairments. 15 different source videos of VGA resolution at 30fps were used. The design and processing were done by SwissQual. The subjective test was run by Yonsei University.

B.6.2 Test design

Encoder	QuickTime, x264	
The following bitrates were used:	High	1000-2000kbit/s
	Medium	512-1000kbits/s
	Low	256-512kbits/s
Streaming	Server Darwin Streaming Server	

Video Player	QuickTime, RealPlayer	
Network Impairment	From live 3G network	
	Simulated using dummynet	For rebuffering by limiting throughput
		For packet loss, by generating 1-3 random intervals of uniform loss

B.6.3 Details

Description	Encoder	max GOP duration (s)	Bitrate	Network	Network degradation	Loss rate	Player	HRC number	Number PVSs
Reference								0	15
LAN - live recording	quicktime	1	high	LAN	none		quicktime	1	5
LAN - live recording	x264	4	high	LAN	none		realplayer	2	5
LAN - live recording	x264	1	medium	LAN	none		realplayer	3	5
LAN - live recording	x264	1	low	LAN	none		realplayer	4	5
LAN - live recording	quicktime	1	high	LAN	loss	high	quicktime	5	5
LAN - live recording	x264	1	high	LAN	loss	high	realplayer	6	5
LAN - live recording	x264	4	high	LAN	loss	low	quicktime	7	5
LAN - live recording	x264	1	high	LAN	loss	low	realplayer	8	5
LAN - live recording	x264	4	medium	LAN	loss	low	realplayer	9	5

LAN - live recording	x264	1	low	LAN	loss	low	realplayer	10	5
LAN - live recording	x264	1	medium to high	LAN	rebuffering		quicktime	11	5
3G network - live recording	x264	1	medium	mobile 3G	good		quicktime	12	5
LAN - live recording	x264	1	medium to high	LAN	rebuffering		quicktime	13	5
3G network - live recording	x264	1	medium	mobile 3G	avg to bad		quicktime	14	5
3G network - live recording	x264	1	medium	mobile 3G	avg to bad		realplayer	15	5

B.7 HybridVGA02

The test design and PVS generation of this test was performed by Yonsei University and the subjective test was run by SwissQual.

B.7.1 Test Design

10 SRCs x 16 HRCs = 160 PVSs (SRC exclude)	
1 Codec	H.264
10 Bitrates (kbps)	128, 256, 300, 320, 448, 500, 512, 704, 1000, 1200
2 Frame Rate (fps)	10, 30
3 PLRs	Low (0.5%-1%), med (1%-1.5%), high (1.5%-)
2 PLCs	slicing, freezing
5 loss ratio (%)	0.5, 0.7, 1, 1.2, 1.5
2 loss types	random, burst
1 Re-scaling size	QVGA
Encoder	x264
Server	Sirannon

Decoder	ffmpeg, post-processing(freezing)
Simulated loss	Packet Loss Simulator
Player	ACREO

HRC	Protocol	Codec	Bit-Rates (kbps)	Frame-Rates	PLR	PLC	Loss rate (%)	Loss type	Slices per frame	Re-scaling
1	RTP/UDP/IP	H.264	1200	30	0	0	0	none	30	
2	RTP/UDP/IP	H.264	1000	30	0	0	0	none	30	
3	RTP/UDP/IP	H.264	704	30	0	0	0	none	30	
4	RTP/UDP/IP	H.264	512	30	0	0	0	none	30	
5	RTP/UDP/IP	H.264	320	30	0	0	0	none	30	
6	RTP/UDP/IP	H.264	256	30	0	0	0	none	30	
7	RTP/UDP/IP	H.264	128	10	0	0	0	none	30	
8	RTP/UDP/IP	H.264	704	30	low	freezing	0.5	burst	30	
9	RTP/UDP/IP	H.264	128	10	low	freezing	0.7	burst	30	
10	RTP/UDP/IP	H.264	448	30	low	slicing	0.7	burst	30	
11	RTP/UDP/IP	H.264	512	30	med	slicing	1	burst	30	
12	RTP/UDP/IP	H.264	1000	30	med	slicing	1	random	30	
13	RTP/UDP/IP	H.264	704	30	med	slicing	1.2	random	30	
14	RTP/UDP/IP	H.264	320	30	high	freezing	1.5	burst	30	
15	RTP/UDP/IP	H.264	500	30	high	slicing	1.5	random	30	
16	RTP/UDP/IP	H.264	300	30	low	slicing	0.7	random	30	QVGA

B.7.2 Subjective Test

The subjective test was carried out by SwissQual.

B.8 HybridVGA03

The test design and PVS generation of this test was performed by OPTICOM GmbH. The subjective test was run by Yonsei University.

B.8.1 Test design

The test was designed for sequences with VGA (640x480) resolution, 25 frames per second. The test was setup as a non-rebuffering experiment, with 16 SRCs and 18 HRCs. A sparse matrix was used to combine SRCs and HRCs, resulting in a total number of 154 PVS. Within the test, simulated network impairments transmitted over LAN, and live video recordings transmitted over a commercially operated IP network where used.

Encoding

X264 was used for video encoding with average bitrate and average quality as restriction. The bitrates used where 250kbits/s, 525kbits/s, 775kbits/s and 1700kbits/s. The quality settings used were 25, 35, and 40.

Streaming

The network streaming of videos was performed using Sirannon and Live555 as streaming servers running under Microsoft Windows OS. On the receiver side VLC Player, or MPlayer were used to control RTP transmission. During transmission the PCAP files were collected using tcpdump. The h264 stream was extracted from the collected pcap files using a tool developed by OPTICOM. Decoding of the h264 stream was carried out using ffmpeg. HRC 2 to HRC 16 were created using the company internal LAN. Within these HRCs the network distortions where introduced artificially by random removal of packets from the recorded pcap file. The live network conditions were recorded using a commercially operated Cable to DSL 6000 connection. No additional distortions were introduced in this case.

Pre- Post processing

HRC14, Tandem Coding 1, was generated by uploading an almost transparent, high bitrate mp4 file to the YouTube video portal. The transcoded mp4 file was then downloaded, the h264 stream extracted, and used as input for the internal transmission chain.

HRC15, Tandem Coding 2, was generated using ffmpeg with different bitrate settings.

For HRC11, Error Concealment One, the ffmpeg internal error concealment was activated, whereas for HRC12, Error Concealment Two, the distorted frames where replaced manually to simulate frame freeze.

For HRC13, Scaling, the VGA input video was reduced in size by a factor of 2 prior to coding. After transmission and decoding the sequence was rescaled to VGA size.

B.8.2 Test Design Summary

HRC Idx	Description	Pre-/Post Proc.	Bit rate	Server	Network	Network Degradation	MTU	# PVS
1	SRC							16
2	Coding only 1		High	Sirannon	LAN	None	Large	8
3	Coding only 2		Medium	Sirannon	LAN	None	Large	8
4	Coding only 3		Low	Sirannon	LAN	None	Large	8
5	Large MTU Size /w PL 1		Medium	Sirannon	LAN	Low PL	Large	8
6	Large MTU Size /w PL 2		Medium	Sirannon	LAN	High PL	Large	8
7	Large MTU Size /w PL 3		Medium	Live555	LAN	Low PL	Large	8
8	Large MTU Size /w PL 4		High	Live555	LAN	Low PL	Large	8
9	Small MTU Size /w PL 1		Medium	Sirannon	LAN	Low PL	Small	8
10	Small MTU Size /w PL 2		Medium	Sirannon	LAN	High PL	Small	8
11	Error concealment 1		High	Live555	LAN	Low PL	Large	8
12	Error concealment 2		High	Live555	LAN	Low PL	Large	8
13	Scaling	Scaling	Medium	Live555	LAN	None / Low PL	Large	8
14	Tandem Coding 1	Tandem 1)	High	Live555	LAN	Low PL	Large	8
15	Tandem Coding 2	Tandem 2)	Low	Live555	LAN	Low PL	Large	8
16	High Bitrate over bad Net		High	Live555	LAN	High PL	Large	8

17	Real life network 1	Low	Live555	Inet	Wild www DSL	Large	8
18	Real life network 2	Medium	Live555	Inet	Wild www DSL	Large	8
19	Real life network 3	High	Live555	Inet	Wild www DSL	Large	8

B.8.3 Subjective Test

The subjective test was carried out at Yonsei University.

B.9 HybridWVGA01

The test design and PVS generation of this test was performed by Yonsei University and the subjective test was run by OPTICOM.

B.9.1 Test Design

10 SRCs x 16 HRCs = 160 PVSs (SRC exclude)	
1 Codec	H.264
8 Bitrates (kbps)	128, 192, 300, 320, 550, 600, 700, 1200
2 Frame Rate (fps)	15, 30
3 PLRs	Low (0.5%-1%), med (1%-1.5%), high (1.5%-)
2 PLCs	slicing, freezing
5 loss ratio (%)	0.5, 0.7, 1, 1.2, 1.5
2 loss types	random, burst
1 Re-scaling size	WQVGA
Encoder	x264
Server	Sirannon
Decoder	ffmpeg, post-processing(freezing)
Simulated loss	Packet Loss Simulator
Player	ACREO

HRC	Protocol	Codec	Bit-Rates (kbps)	Frame-Rates	PLR	PLC	Loss rate (%)	Loss type	Slices per frame	Re-scaling
1	RTP/UDP/IP	H.264	1200	30	0	0	0	none	30	
2	RTP/UDP/IP	H.264	700	30	0	0	0	none	30	
3	RTP/UDP/IP	H.264	600	30	0	0	0	none	30	
4	RTP/UDP/IP	H.264	320	30	0	0	0	none	30	
5	RTP/UDP/IP	H.264	192	15	0	0	0	none	30	
6	RTP/UDP/IP	H.264	128	15	0	0	0	none	30	
7	RTP/UDP/IP	H.264	192	15	low	freezing	0.5	random	30	
8	RTP/UDP/IP	H.264	1200	30	low	freezing	0.5	burst	30	
9	RTP/UDP/IP	H.264	320	30	low	slicing	0.7	burst	30	
10	RTP/UDP/IP	H.264	700	30	low	slicing	0.7	burst	30	
11	RTP/UDP/IP	H.264	128	15	med	freezing	1	random	30	
12	RTP/UDP/IP	H.264	600	30	med	slicing	1	burst	30	
13	RTP/UDP/IP	H.264	128	15	med	slicing	1.2	random	30	
14	RTP/UDP/IP	H.264	320	30	high	slicing	1.5	burst	30	
15	RTP/UDP/IP	H.264	550	30	high	slicing	1.5	random	30	
16	RTP/UDP/IP	H.264	300	30	0	0	0	none	30	WQVGA

B.9.2 Subjective Test

The subjective test was carried out at OPTICOM.

B.10 HybridWVGA02

The test design and PVS generation of this test was performed by Acreo Swedish ICT AB and the subjective test was run by FUB

B.10.1 Test design

The test was designed as a rebuffering experiments in that having 8 SRC and 11 HRCs, giving a total number of PVSs to be 88. The HRCs were a mixture between coding impairments, downsampling and rebuffering. The table summarizes the HRC.

Table 30: test design for WVGA02

HRC	Protocol	Codec	Bit-Rates (kbps)	Frame-Rates	Re-scaling	Rebuffering
1	RTP/UDP/IP	H.264	1200	25		
2	RTP/UDP/IP	H.264	600	25		
3	RTP/UDP/IP	H.264	320	25		
4	RTP/UDP/IP	H.264	192	25		
5	RTP/UDP/IP	H.264	1200	25		x
6	RTP/UDP/IP	H.264	600	25		x
7	RTP/UDP/IP	H.264	320	25		x
8	RTP/UDP/IP	H.264	192	25		x
9	RTP/UDP/IP	H.264	320	25	WQVGA	
10	RTP/UDP/IP	H.264	320	25	WQVGA	x
11	RTP/UDP/IP	H.264	192	25	WQVGA/4	

Encoding

X264 was used for the video encoding with average bitrate as restriction. HRC1-4 were encoded at 4 different average bitrate 1200kbps, 600kbps, 320kbps, and 192kbps.

Streaming simulation and Rebuffering event

The network streaming of the videos were performed by using Gstreamer in a OpenSUSE Linux platform. The encoded H.264 bistreams were transmitted over a local loop network as RTP payloads. In the receiver side, a Gstreamer receiver were configured to receive incoming RTP streams, and get H.264 bitstreams out from RTP packets. Then the H.264 bistreams were decoded. A video player developed by Marcus Barkowsky, called iconvert, were used to record a unique AVI file from the captured playback data and the captured timestamps. The network bandwidth were controlled by token buffer filter (TBF). The rebuffering events were created by adjusting the TBF and buffer size. There are many buffers exists during the whole transmission chain. For example, buffers in TBF, Gstreamer, video player. All these buffers affected the length and position (timing) of the rebuffering events. The network transmission of IP packets were captured by tcpdump tool, which gave the PCAP files.

WVGA/4 and WVGA /16

HRC9-11 are the HRCs that was processed involving resolution reduction. The original WVGA videos were firstly downsampled to ¼ or ⅟₁₆ f the original resolution, 426x240, or 214x120 for HRC11. Then the downsampled videos were encoded by x264 with average bitrate restriction. These bitstreams then transmitted over local loop network, and decoded. At the video playback stage, the player upscaled these videos back to their original WVGA resolution and displayed to the viewers.

B.10.2 Subjective test

The subjective test was carried out at the FUB Laboratories during late December 2013 with the participation of 24 subjects, all of them being university or high school students and aging from 18 to 27; all were positively screened for visual acuity and color vision.

Training was done using the same video clips used for the test, having care to select all the range of quality and all the SRC material used in the actual test.

The test was done using a PC equipped with high speed SSD disk in RAID configuration (four Samsung Pro 256 G) and a high end video board (NVidia GTX 770). The display was a 30" professional EIZO monitor, set to display the picture at HD resolution. The subjects were seated at 2H the active part of the screen. The player was AcrVQWin [1] developed by Acreo.

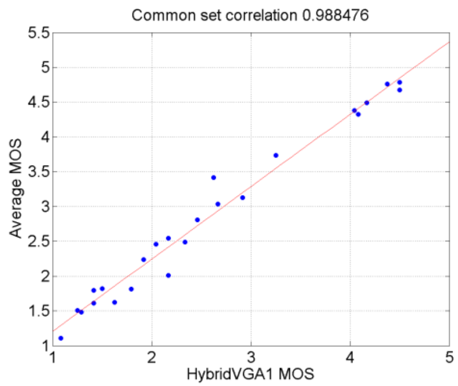
Reference List

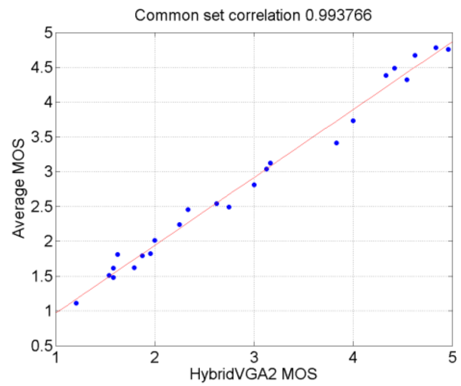
[1] Jonsson, J. and Brunnström, K., "Getting Started With ArcVQWin", Acreo AB, acr022250, (2007)

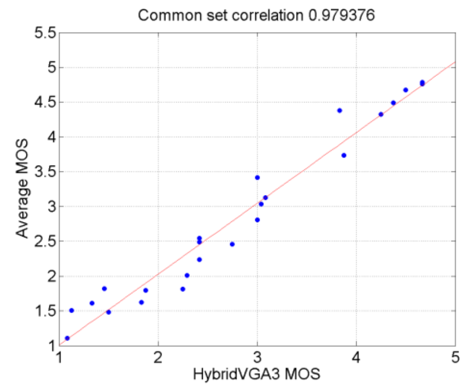
Annex C Plots

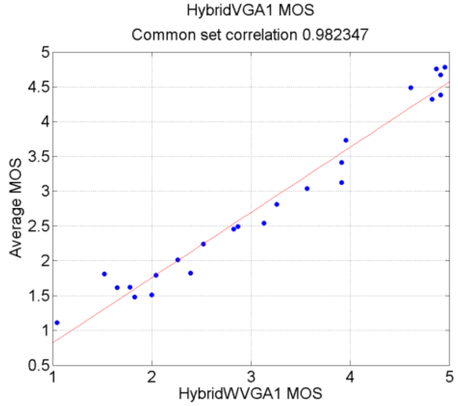
C.1. Common Set Plots

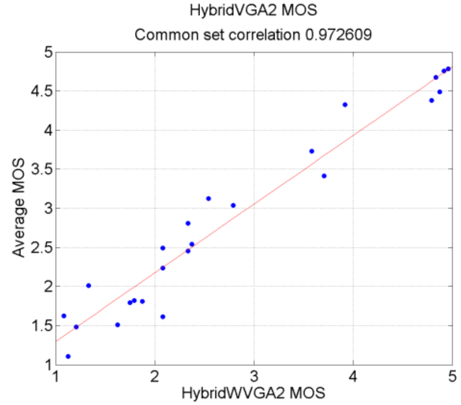
This section contains scatter plots depicting the common set clips only. The x-axis is one dataset's MOS. The y-axis is the average MOS for either all five VGA/WVGA datasets or all five HD datasets, depending upon whether the dataset on the x-axis is VGA/WVGA or HD. The Pearson correlation for the data is printed at the top.

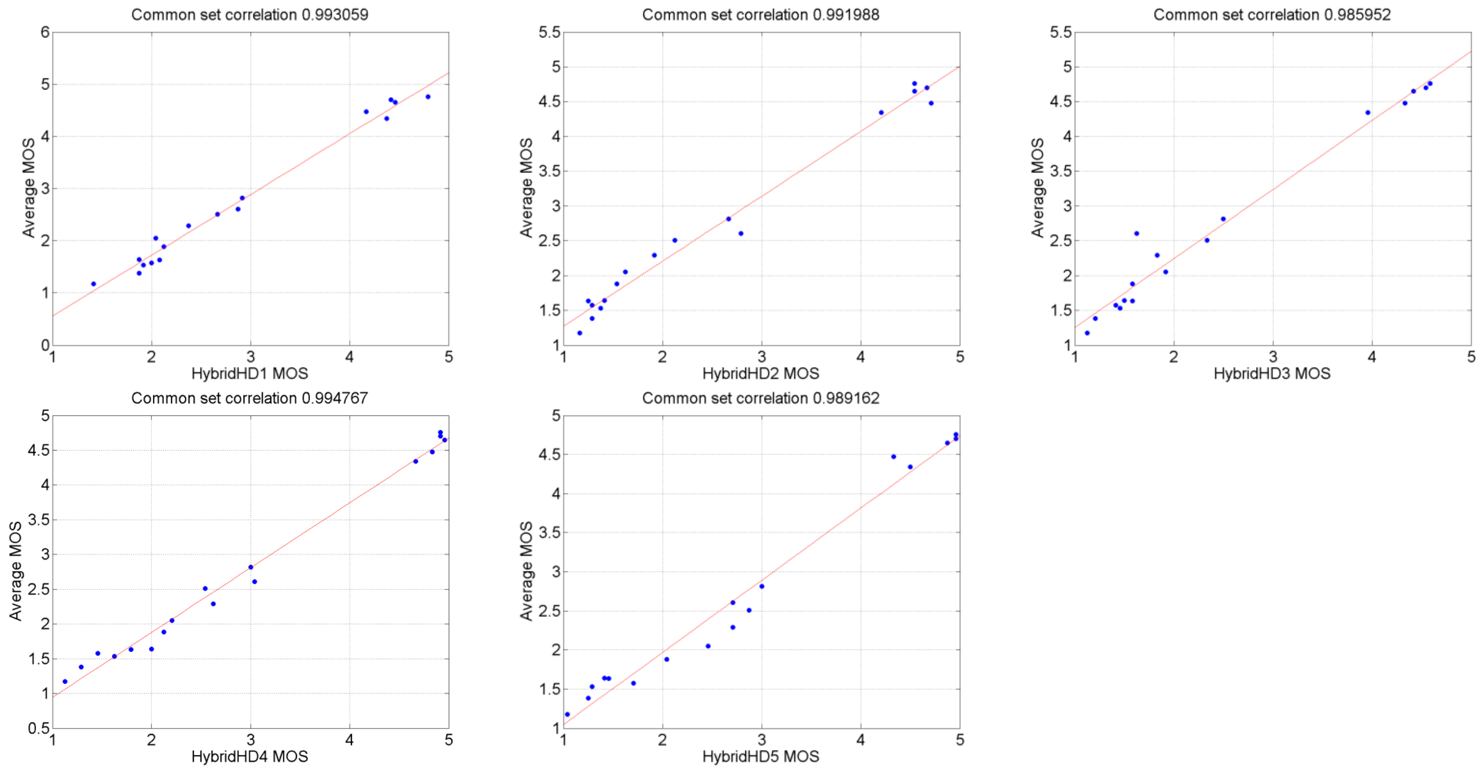








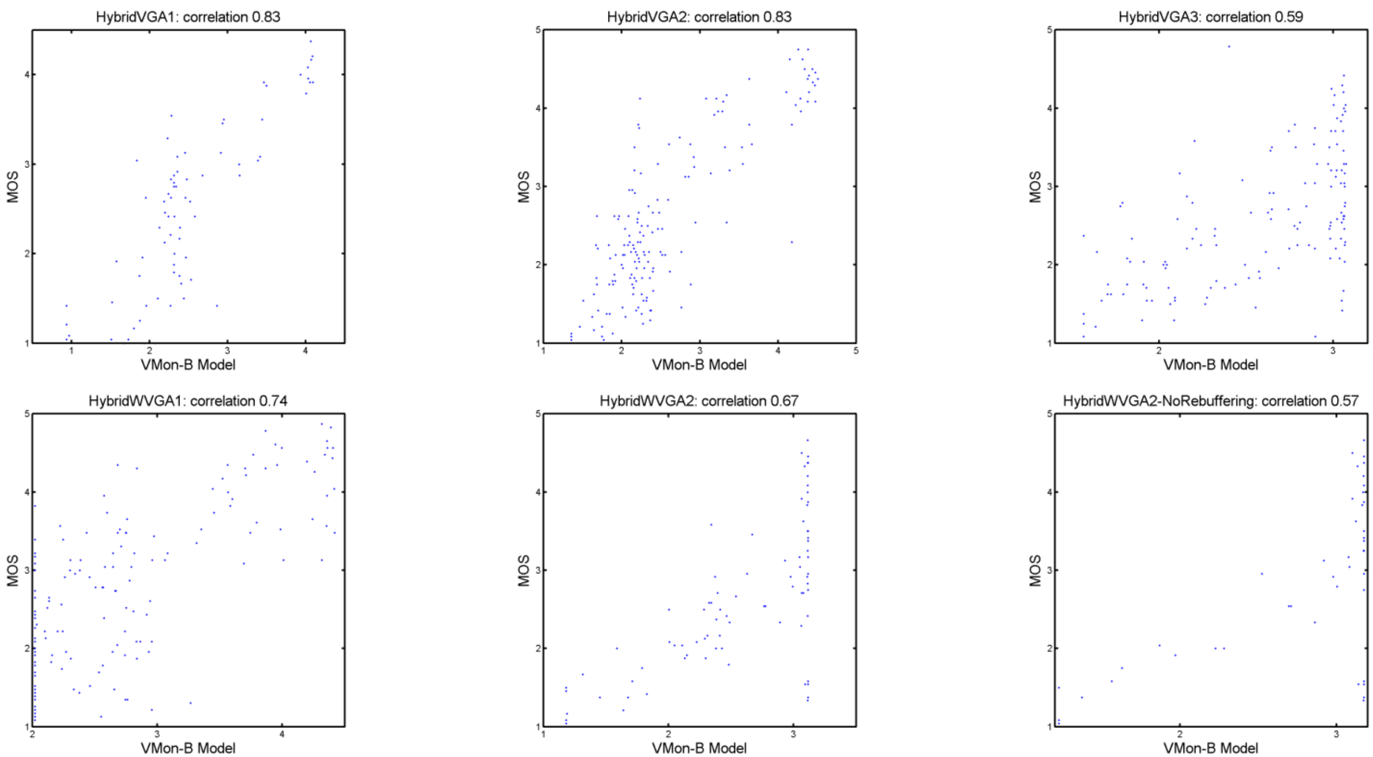


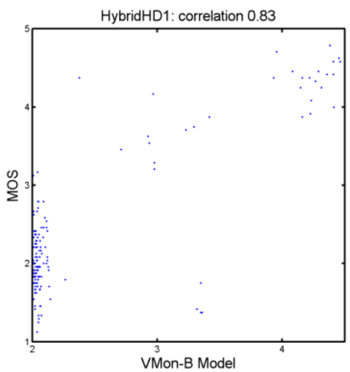

Common set correlation 0.993059
Average MOS
HybridHD1 MOS
Common set correlation 0.991988
Average MOS
HybridHD2 MOS
Common set correlation 0.985952
Average MOS
HybridHD3 MOS
Common set correlation 0.994767
Average MOS
HybridHD4 MOS
Common set correlation 0.989162
Average MOS
HybridHD5 MOS

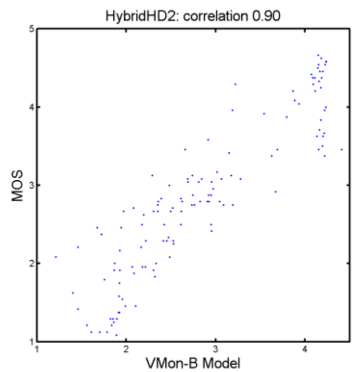
C.2 Plots for Models versus MOS

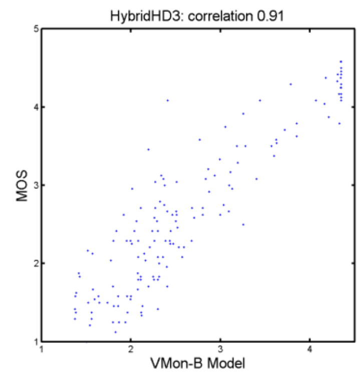
C.2.1. Plots for Model VMon-B versus MOS

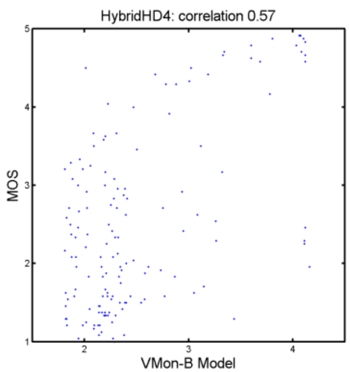
This section contains scatter plots depicting encrypted model VMon-B versus MOS.

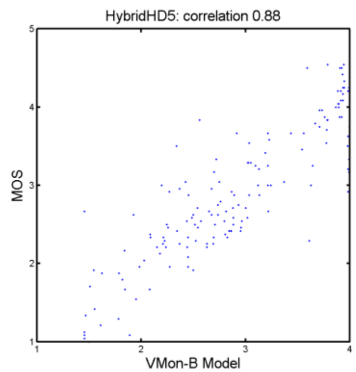


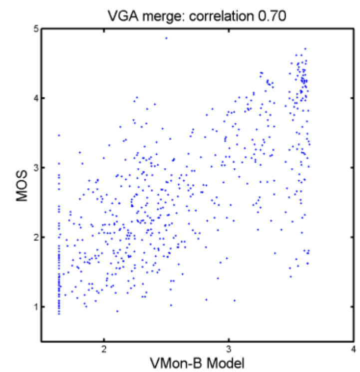

HybridHD1: correlation 0.83
MOS
VMon-B Model

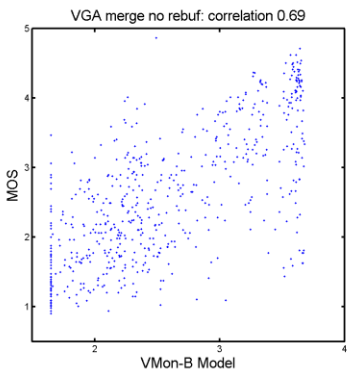

HybridHD2: correlation 0.90
MOS
VMon-B Model

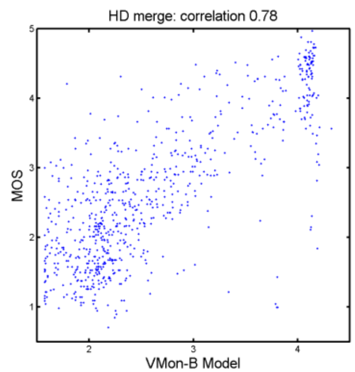

HybridHD3: correlation 0.91
MOS
VMon-B Model


HybridHD4: correlation 0.57
MOS
VMon-B Model


HybridHD5: correlation 0.88
MOS
VMon-B Model

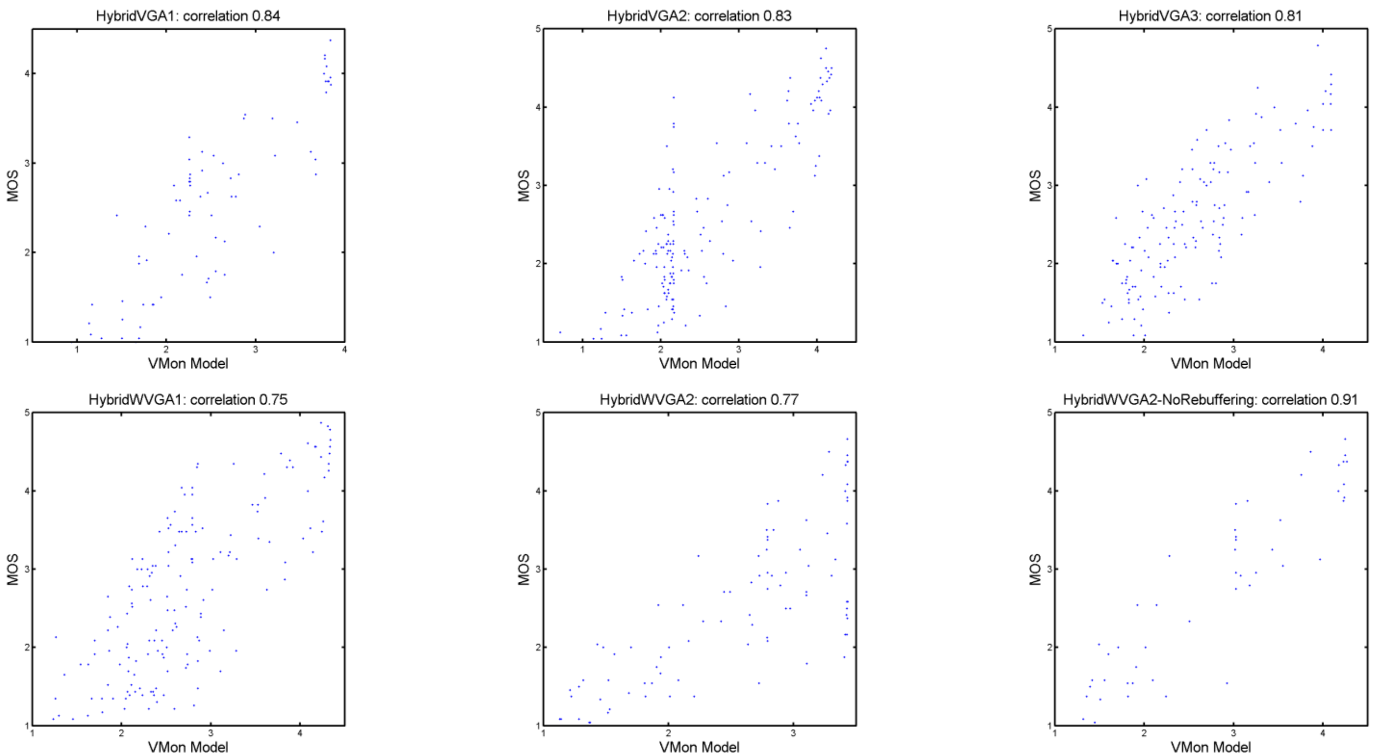

VGA merge: correlation 0.70
MOS
VMon-B Model

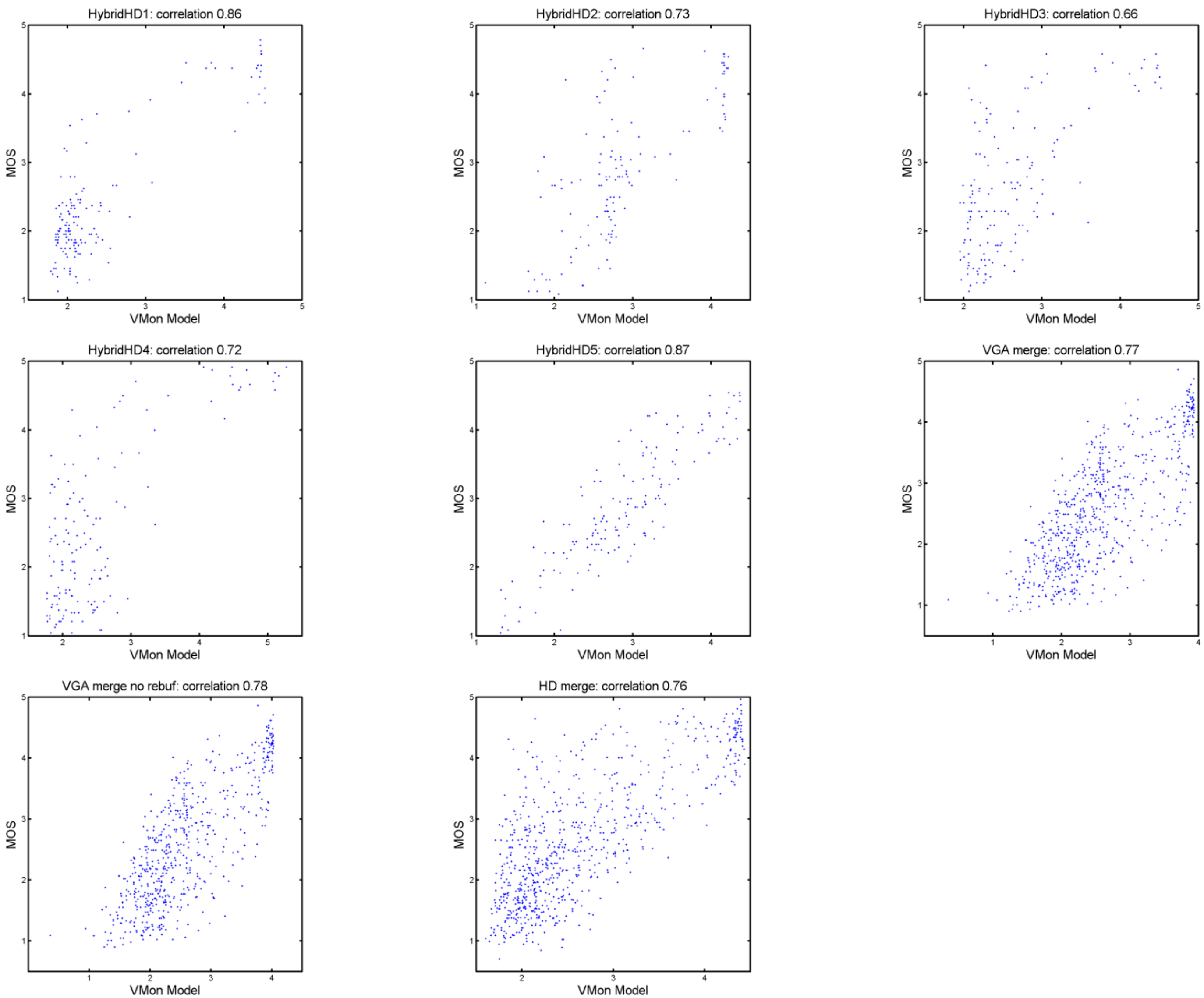

VGA merge no rebuf: correlation 0.69
MOS
VMon-B Model


HD merge: correlation 0.78
MOS
VMon-B Model

C.2.2 Plots for Model VMon versus MOS

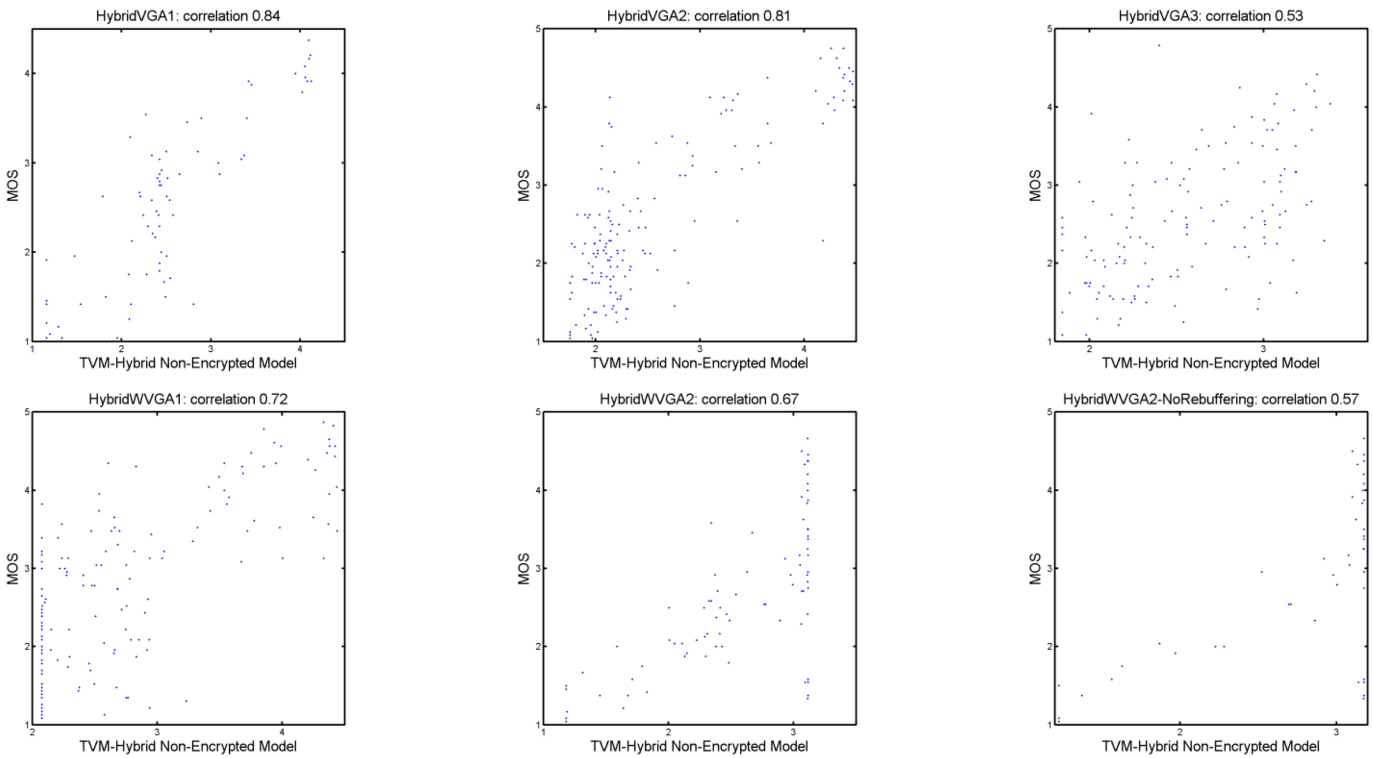
This section contains scatter plots depicting model VMon versus MOS.

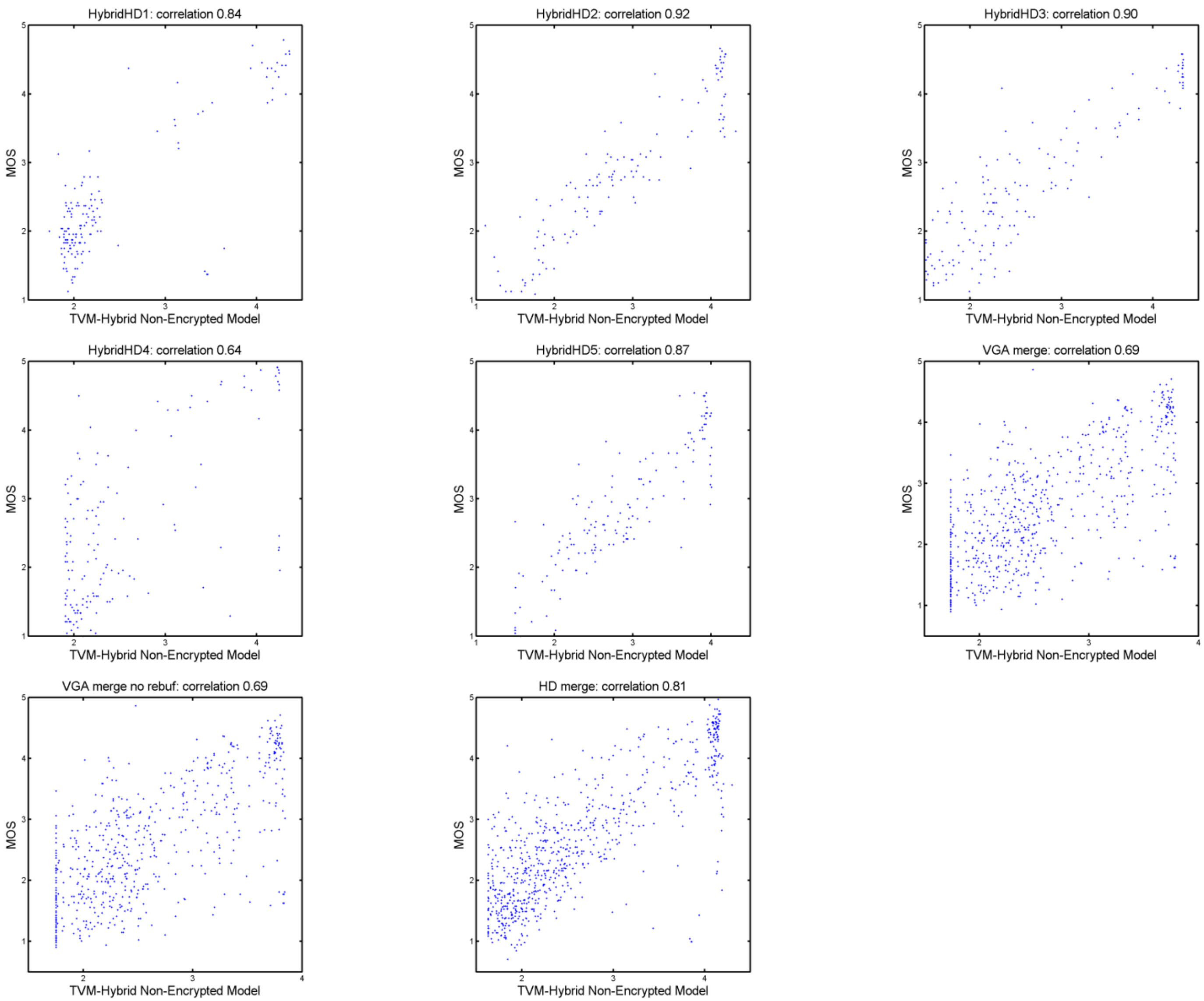



HybridHD1: correlation 0.86
HybridHD2: correlation 0.73
HybridHD3: correlation 0.66
HybridHD4: correlation 0.72
HybridHD5: correlation 0.87
VGA merge: correlation 0.77
VGA merge no rebuf: correlation 0.78
HD merge: correlation 0.76
MOS
VMon Model

C.2.3 Plots for Model TVM-Hybrid Non-Encrypted versus MOS

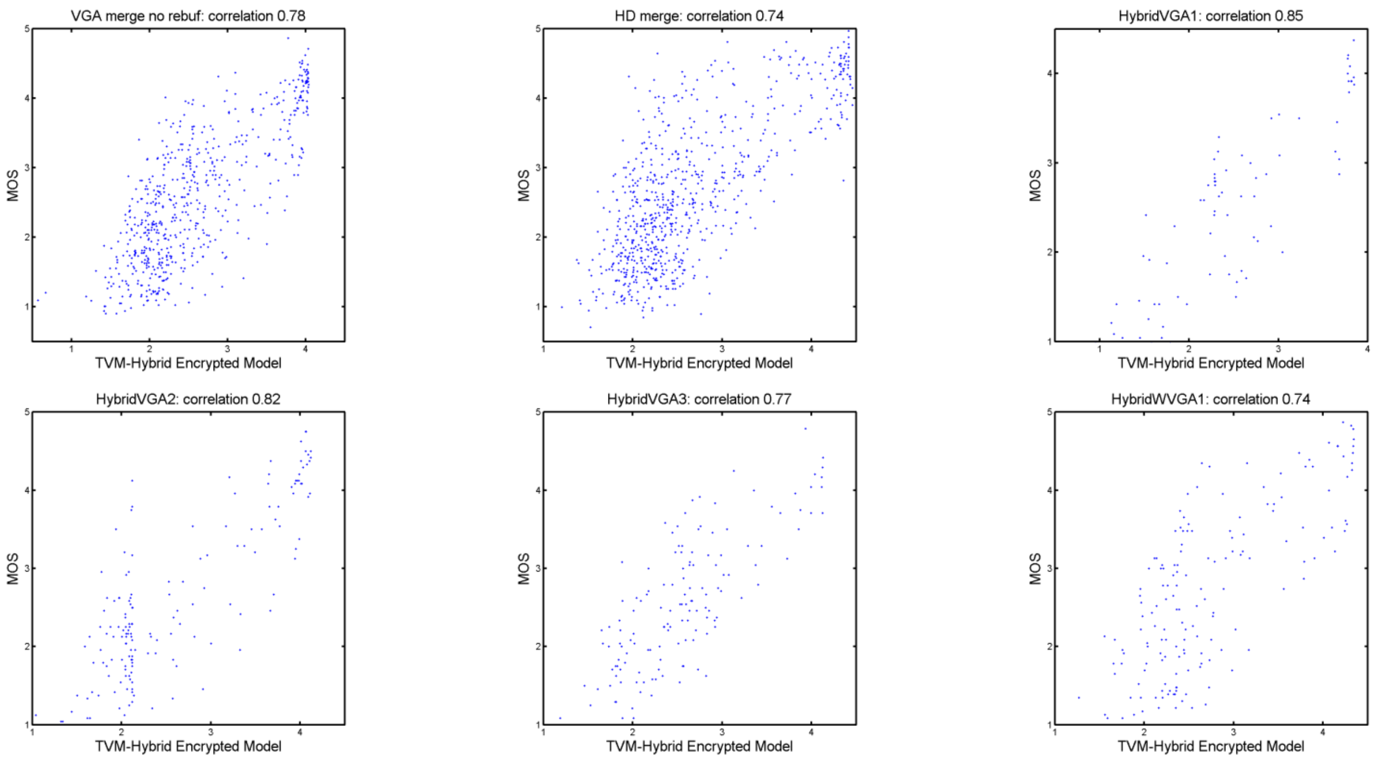
This section contains scatter plots depicting model TVM-Hybrid Non-Encrypted versus MOS.

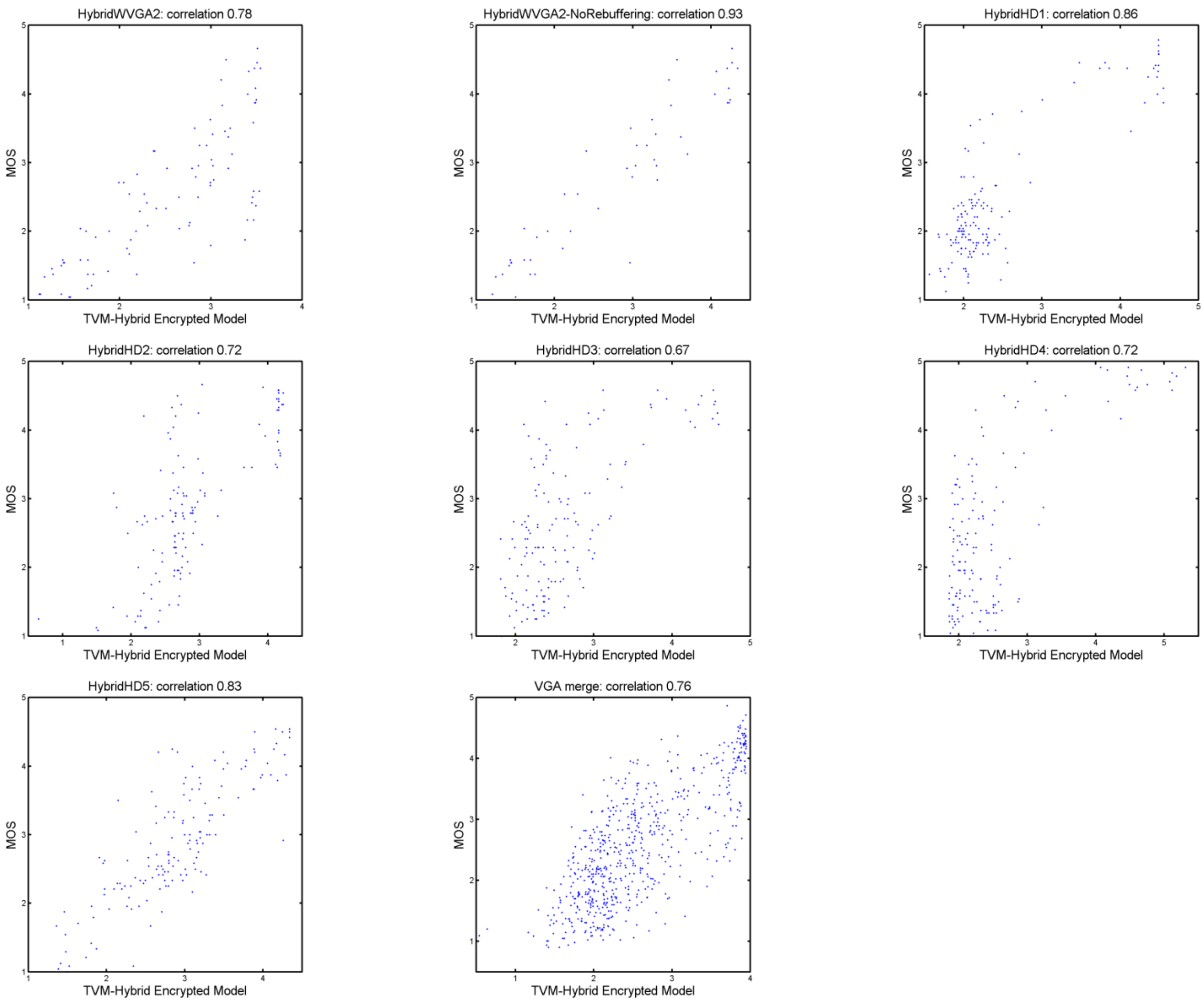



HybridHD1: correlation 0.84
HybridHD2: correlation 0.92
HybridHD3: correlation 0.90
HybridHD4: correlation 0.64
HybridHD5: correlation 0.87
VGA merge: correlation 0.69
VGA merge no rebuf: correlation 0.69
HD merge: correlation 0.81
MOS
TVM-Hybrid Non-Encrypted Model

C.2.4 Plots for Model TVM-Hybrid Encrypted versus MOS

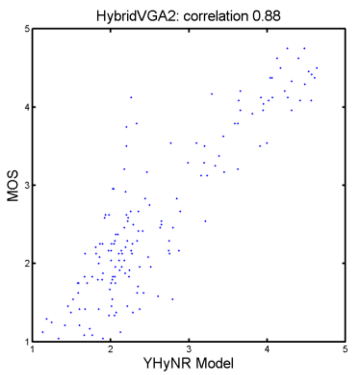
This section contains scatter plots depicting model TVM-Hybrid Encrypted versus MOS.

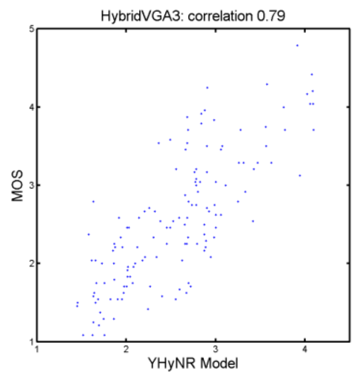


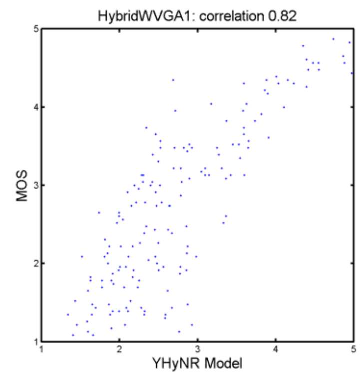

HybridWVGA2: correlation 0.78
HybridWVGA2-NoRebuffering: correlation 0.93
HybridHD1: correlation 0.86
HybridHD2: correlation 0.72
HybridHD3: correlation 0.67
HybridHD4: correlation 0.72
HybridHD5: correlation 0.83
VGA merge: correlation 0.76
MOS
TVM-Hybrid Encrypted Model

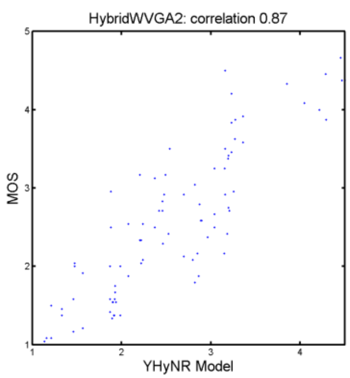
C.2.5 Plots for Model YHyNR versus MOS

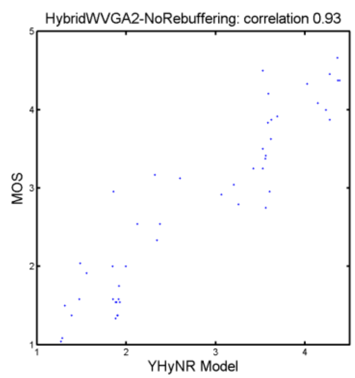
This section contains scatter plots depicting model YHyNR versus MOS.

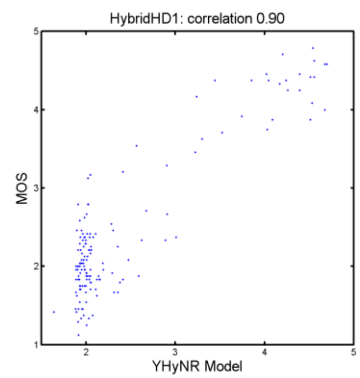


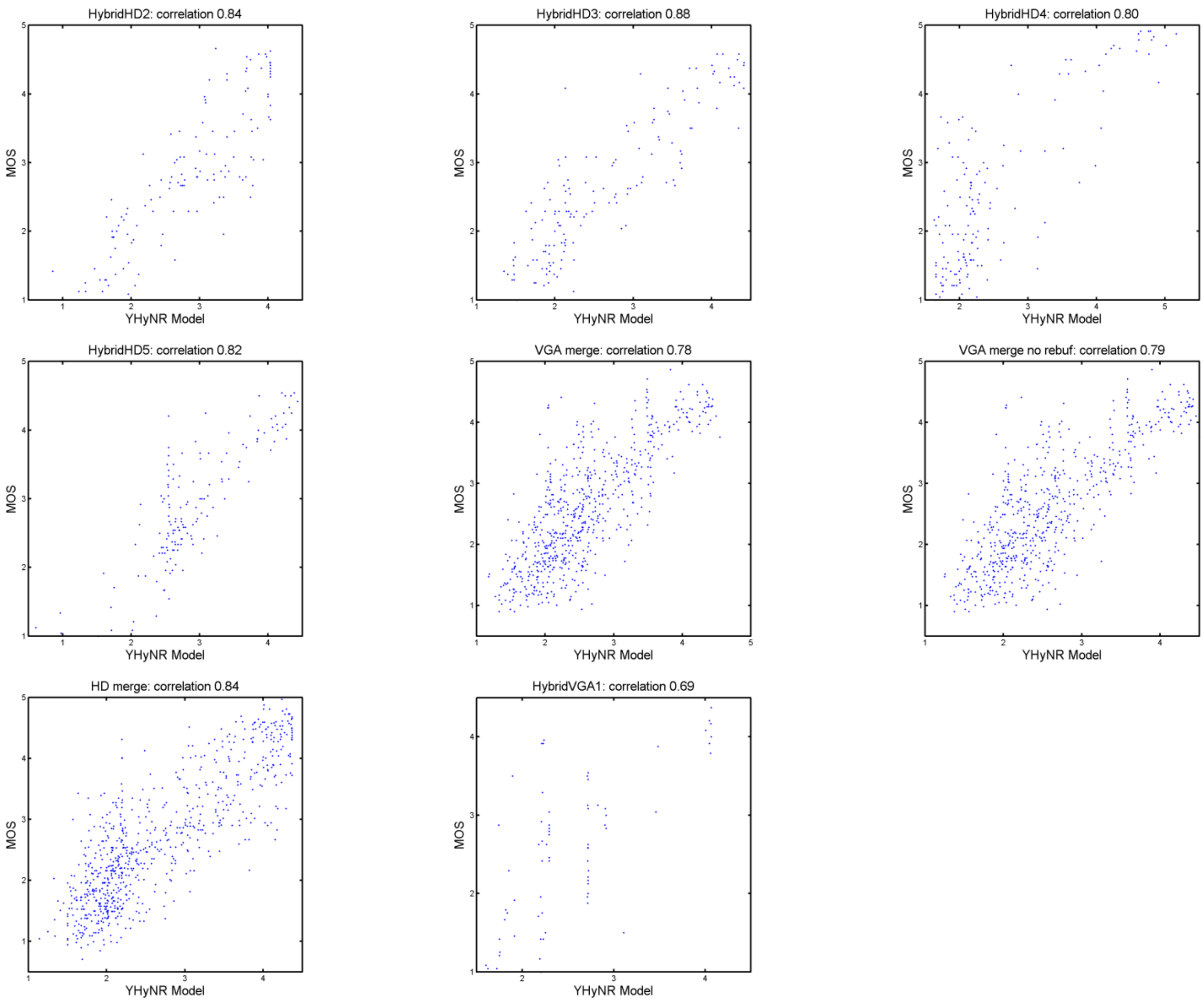






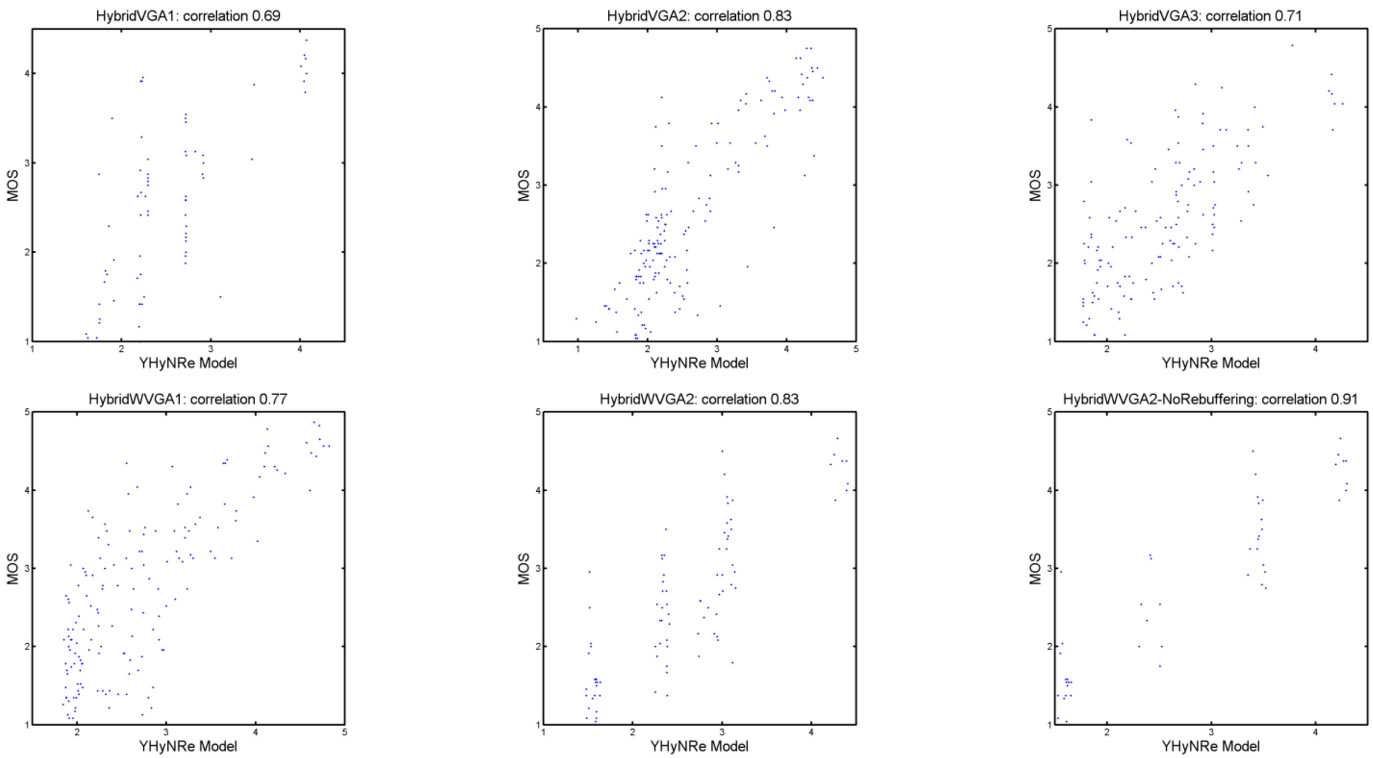


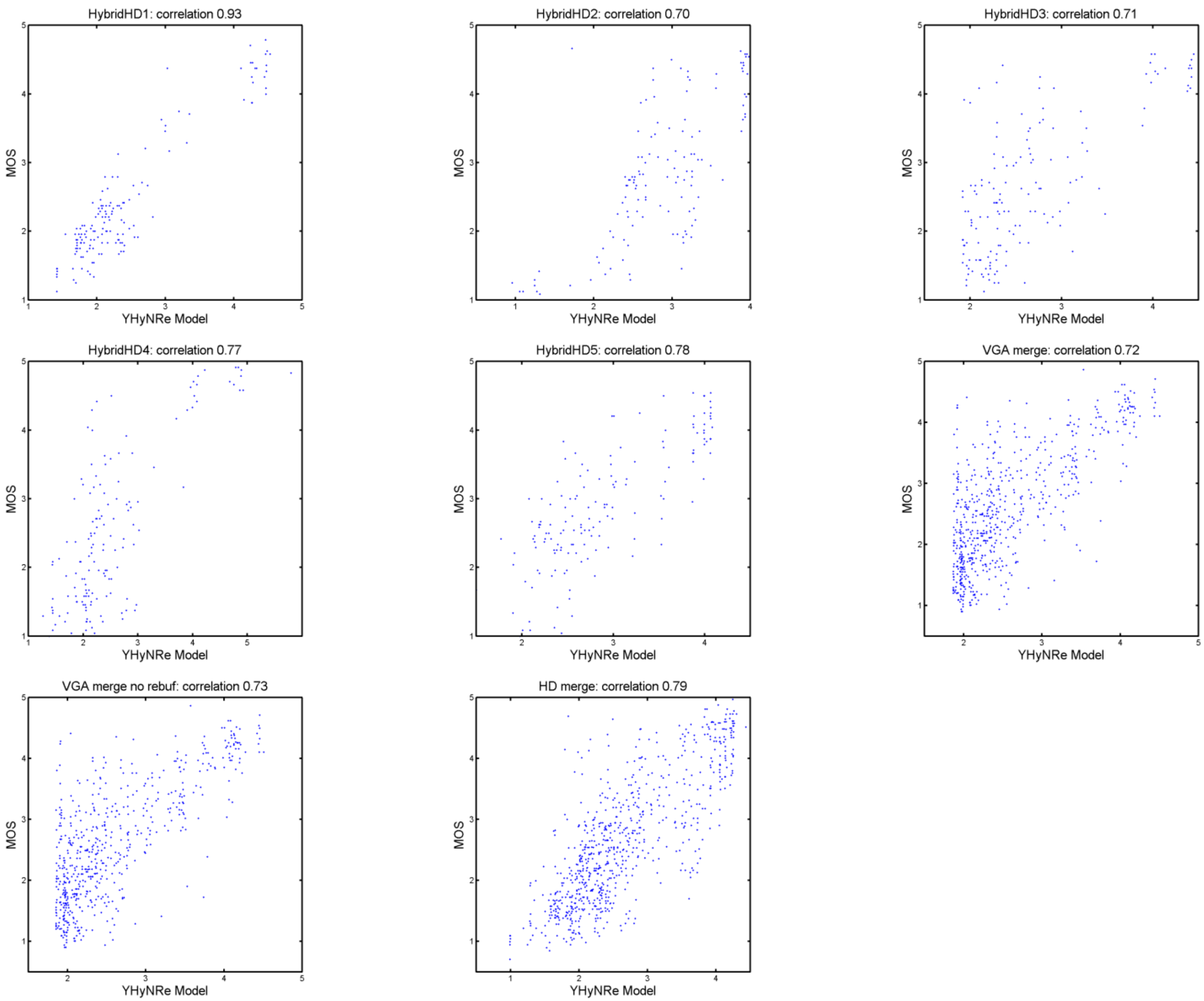



HybridHD2: correlation 0.84
HybridHD3: correlation 0.88
HybridHD4: correlation 0.80
HybridHD5: correlation 0.82
VGA merge: correlation 0.78
VGA merge no rebuf: correlation 0.79
HD merge: correlation 0.84
HybridVGA1: correlation 0.69
MOS
YHyNR Model

C.2.6 Plots for Model YHyNRe versus MOS

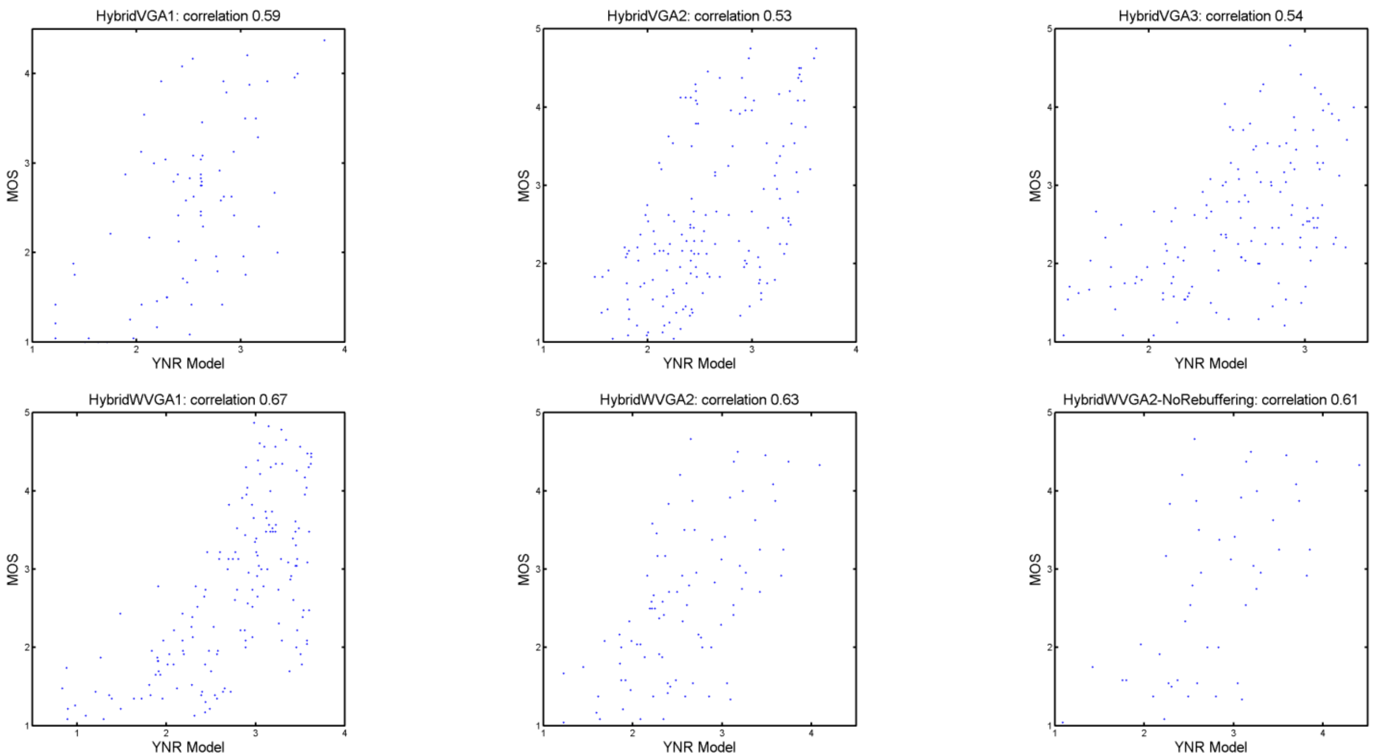
This section contains scatter plots depicting model YHyNRe versus MOS.

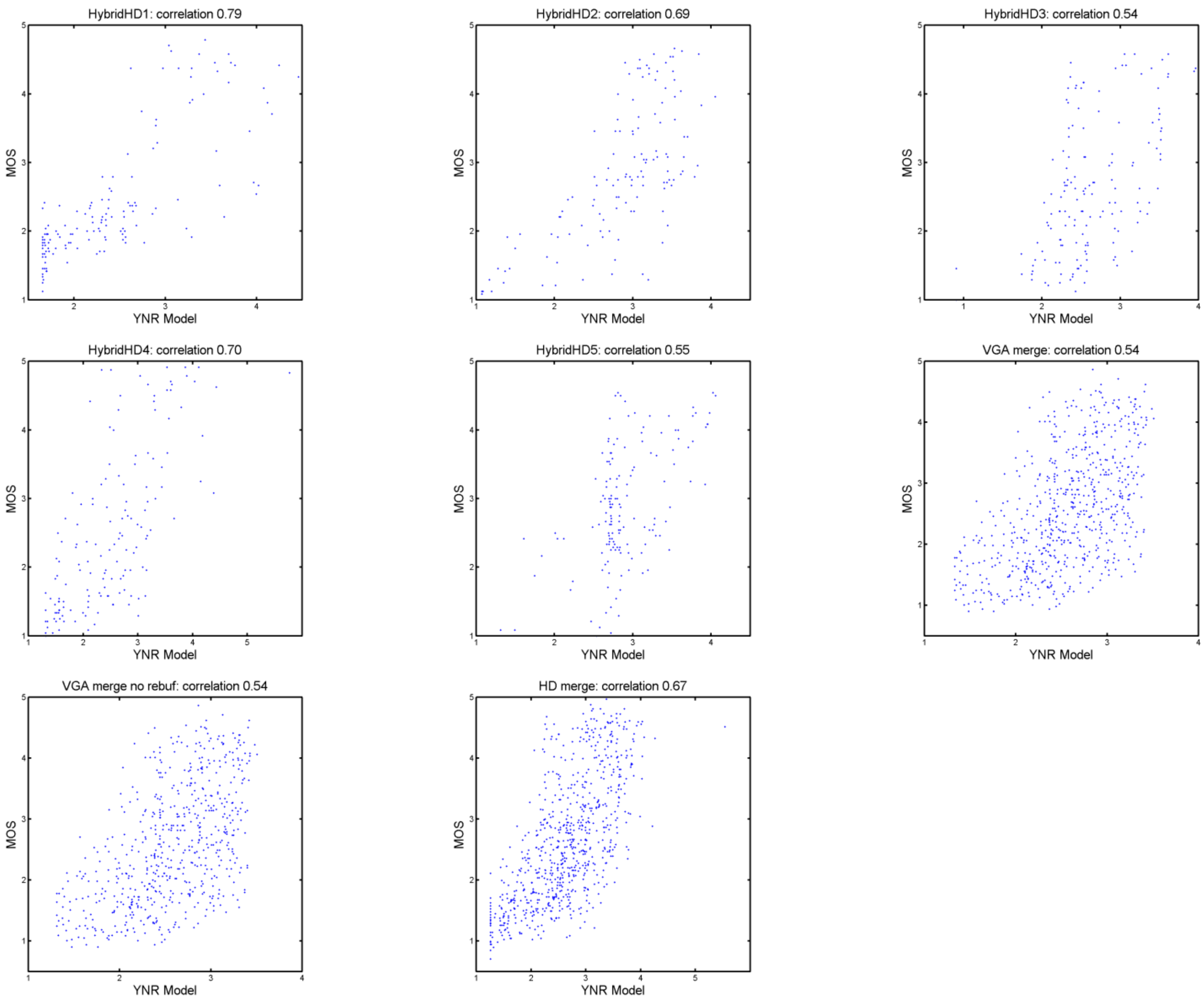



HybridHD1: correlation 0.93
HybridHD2: correlation 0.70
HybridHD3: correlation 0.71
HybridHD4: correlation 0.77
HybridHD5: correlation 0.78
VGA merge: correlation 0.72
VGA merge no rebuf: correlation 0.73
HD merge: correlation 0.79
MOS
YHyNRe Model

C.2.7 Plots for Model YNR versus MOS

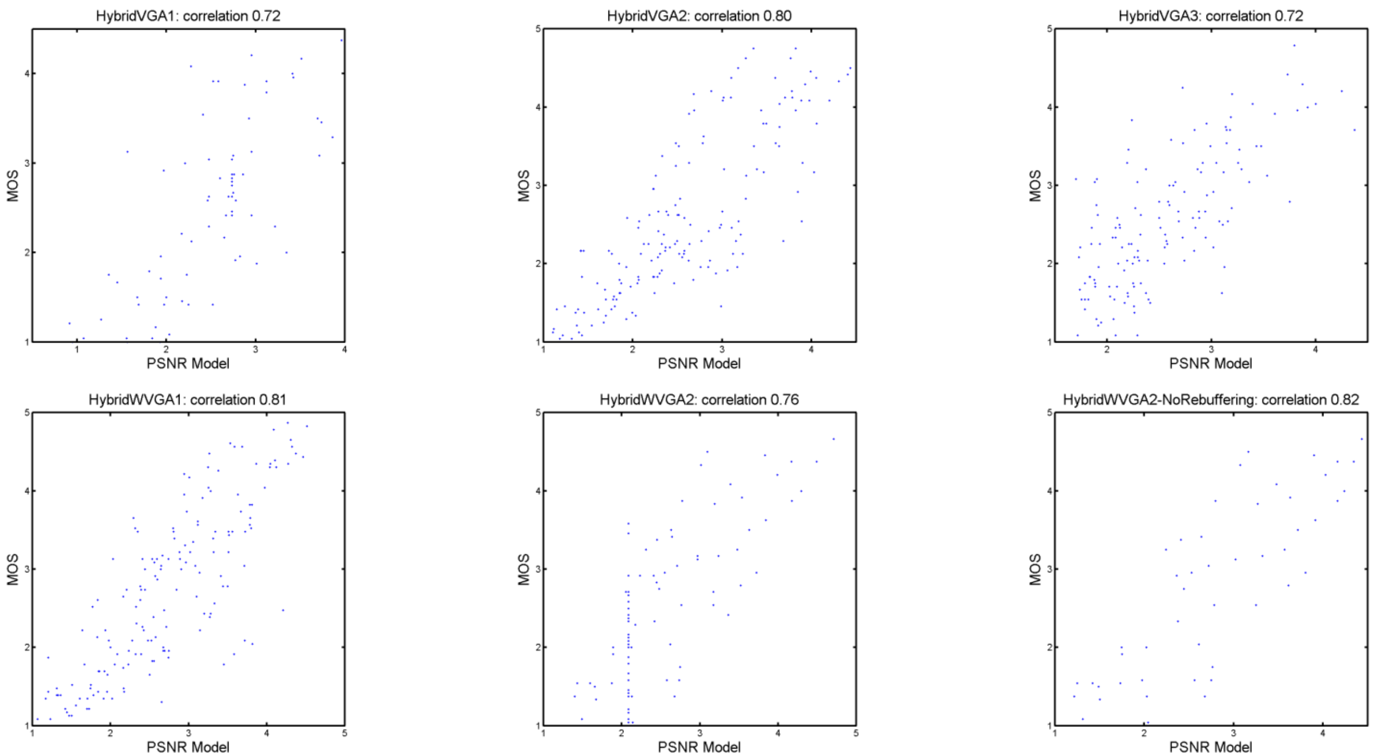
This section contains scatter plots depicting model YNR versus MOS.

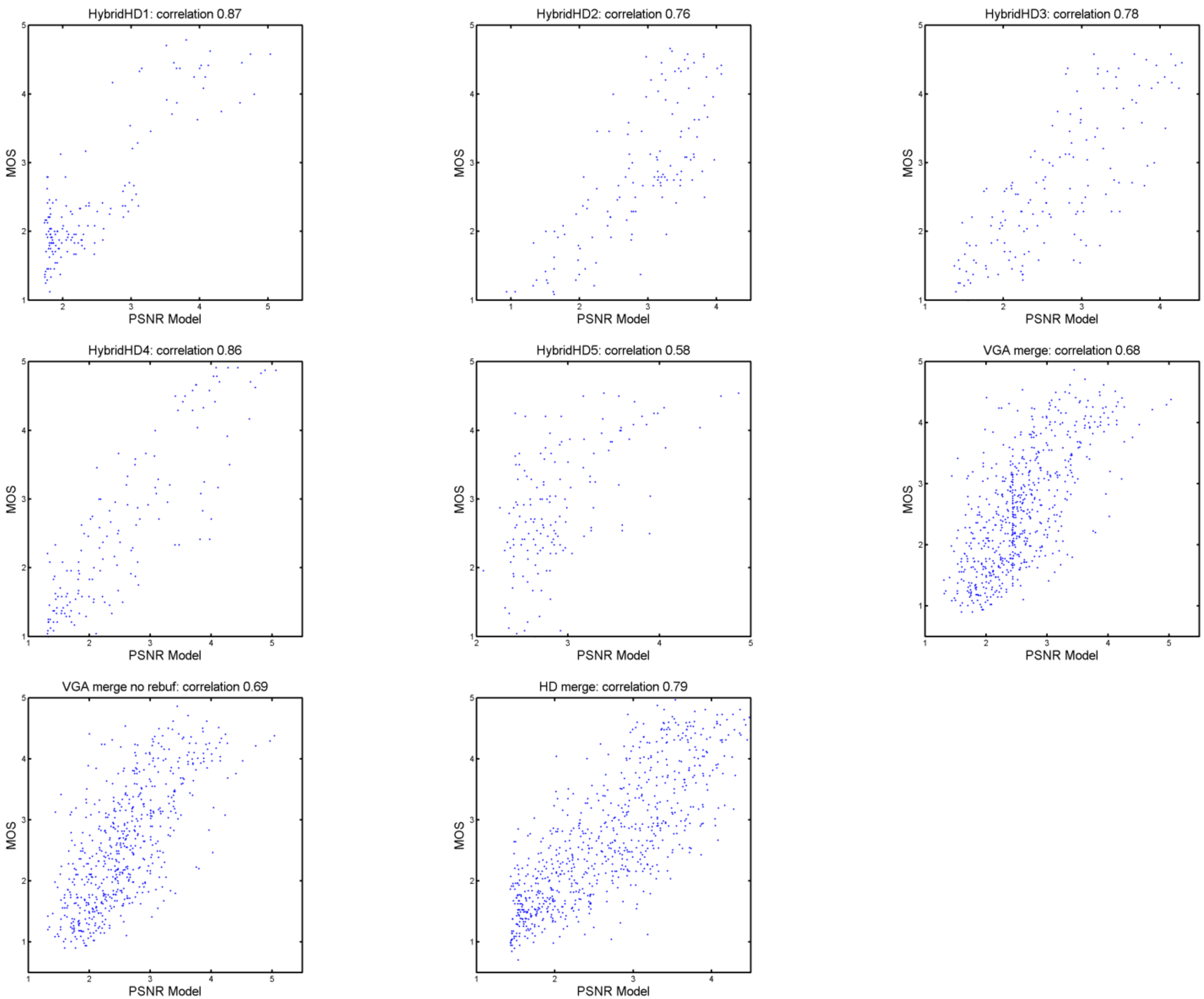



HybridHD1: correlation 0.79
HybridHD2: correlation 0.69
HybridHD3: correlation 0.54
HybridHD4: correlation 0.70
HybridHD5: correlation 0.55
VGA merge: correlation 0.54
VGA merge no rebuf: correlation 0.54
HD merge: correlation 0.67
MOS
YNR Model

C.2.8 Plots for Model PSNR versus MOS

This section contains scatter plots depicting reference model PSNR versus MOS.

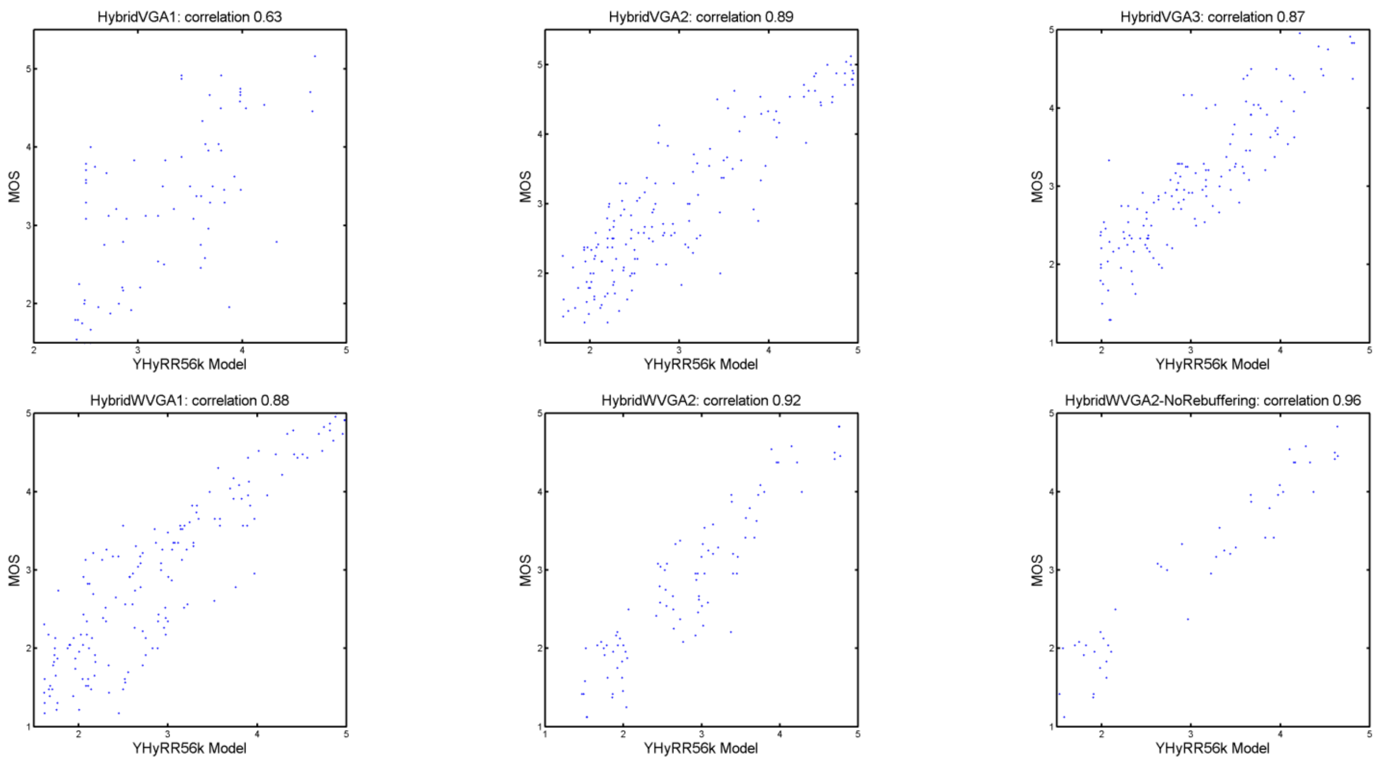


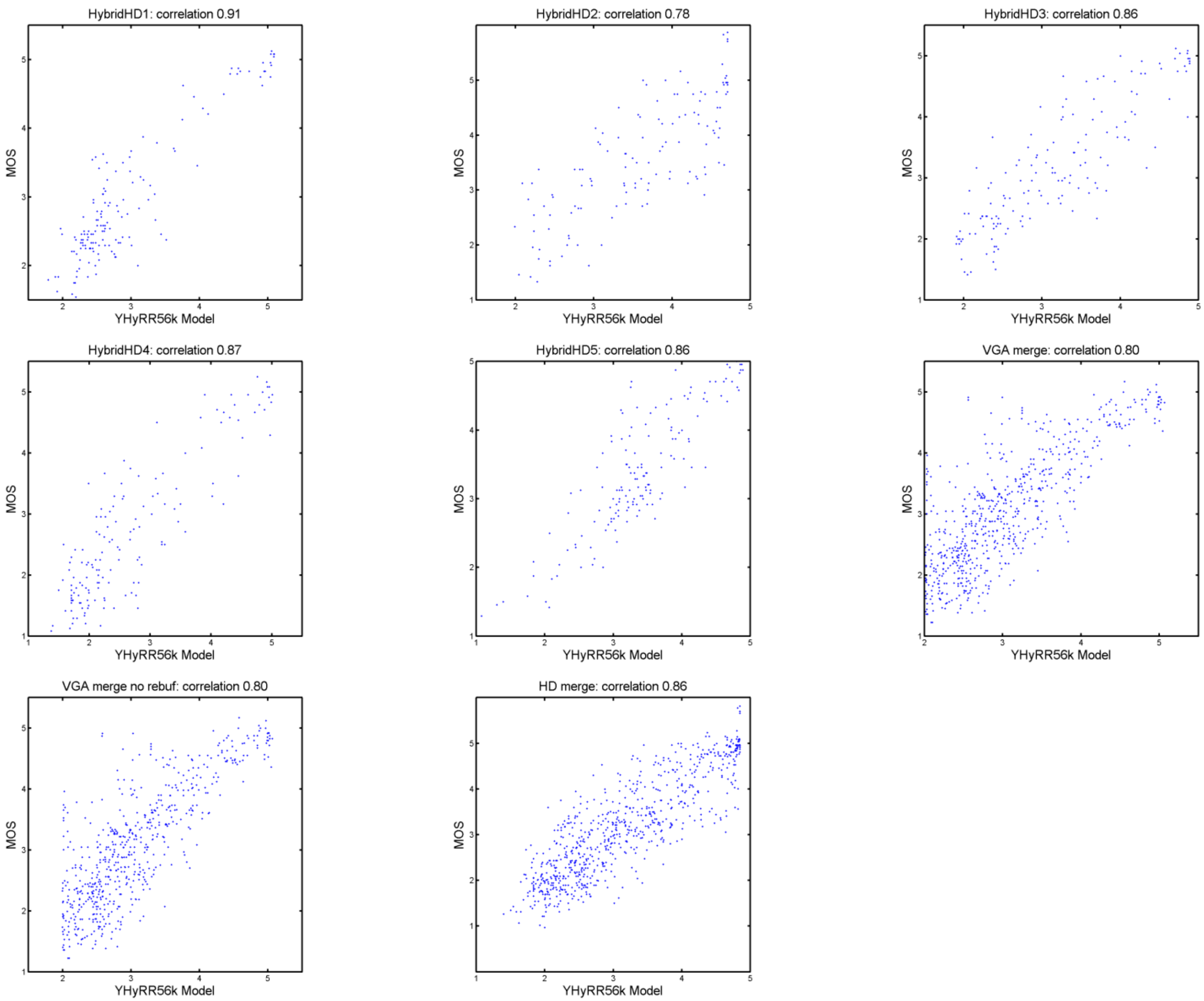

HybridHD1: correlation 0.87
HybridHD2: correlation 0.76
HybridHD3: correlation 0.78
HybridHD4: correlation 0.86
HybridHD5: correlation 0.58
VGA merge: correlation 0.68
VGA merge no rebuf: correlation 0.69
HD merge: correlation 0.79
MOS
PSNR Model

C.3 Plots for Models versus DMOS

C.3.1 Plots for Model YHyRR56k versus DMOS

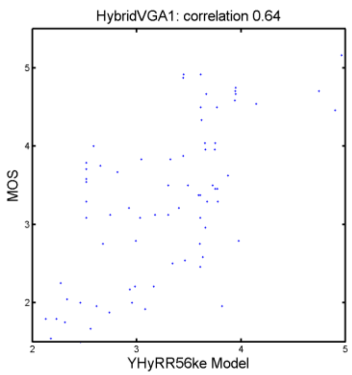
This section contains scatter plots depicting model YHyRR56k versus DMOS.

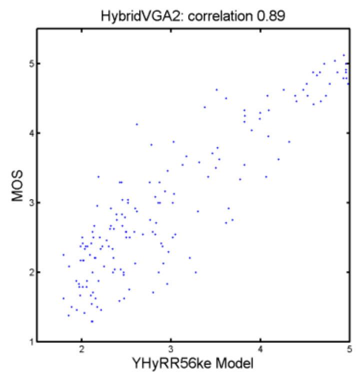


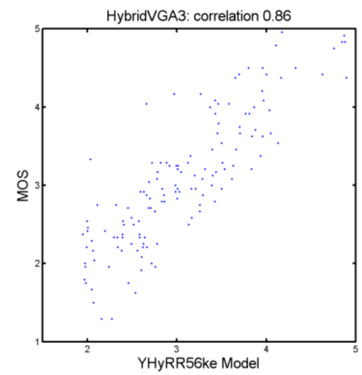

HybridHD1: correlation 0.91
HybridHD2: correlation 0.78
HybridHD3: correlation 0.86
HybridHD4: correlation 0.87
HybridHD5: correlation 0.86
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.86
MOS
YHyRR56k Model

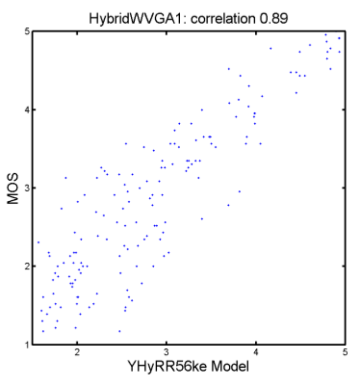
C.3.2 Plots for Model YHyRR56ke versus DMOS

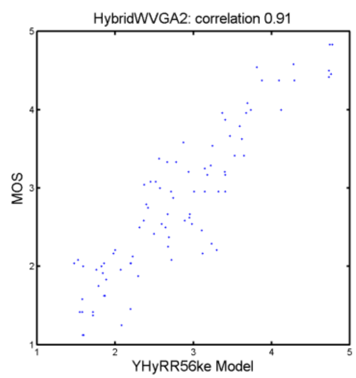
This section contains scatter plots depicting model YHyRR56ke versus DMOS.

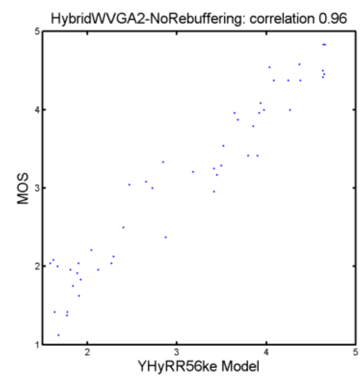


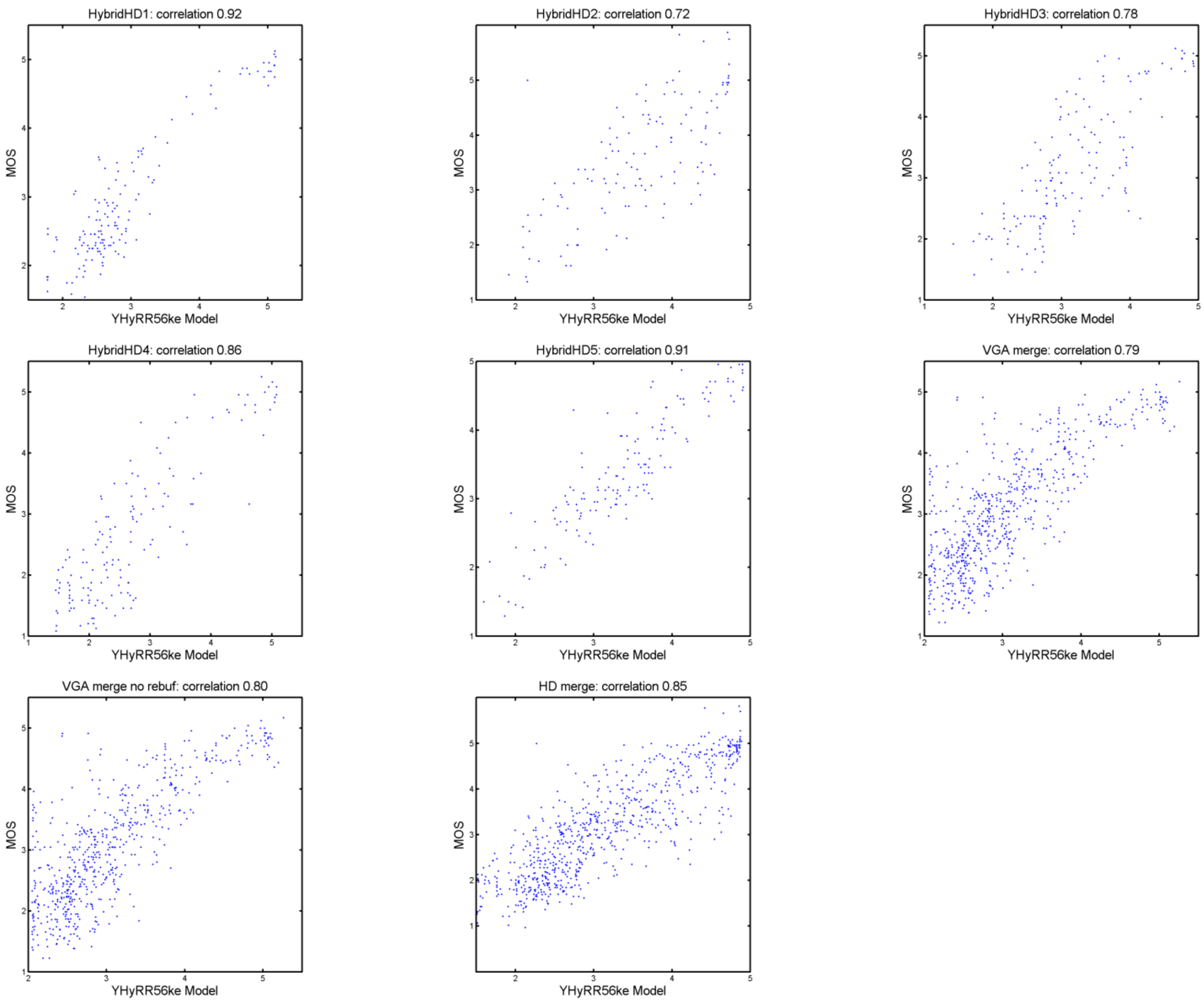






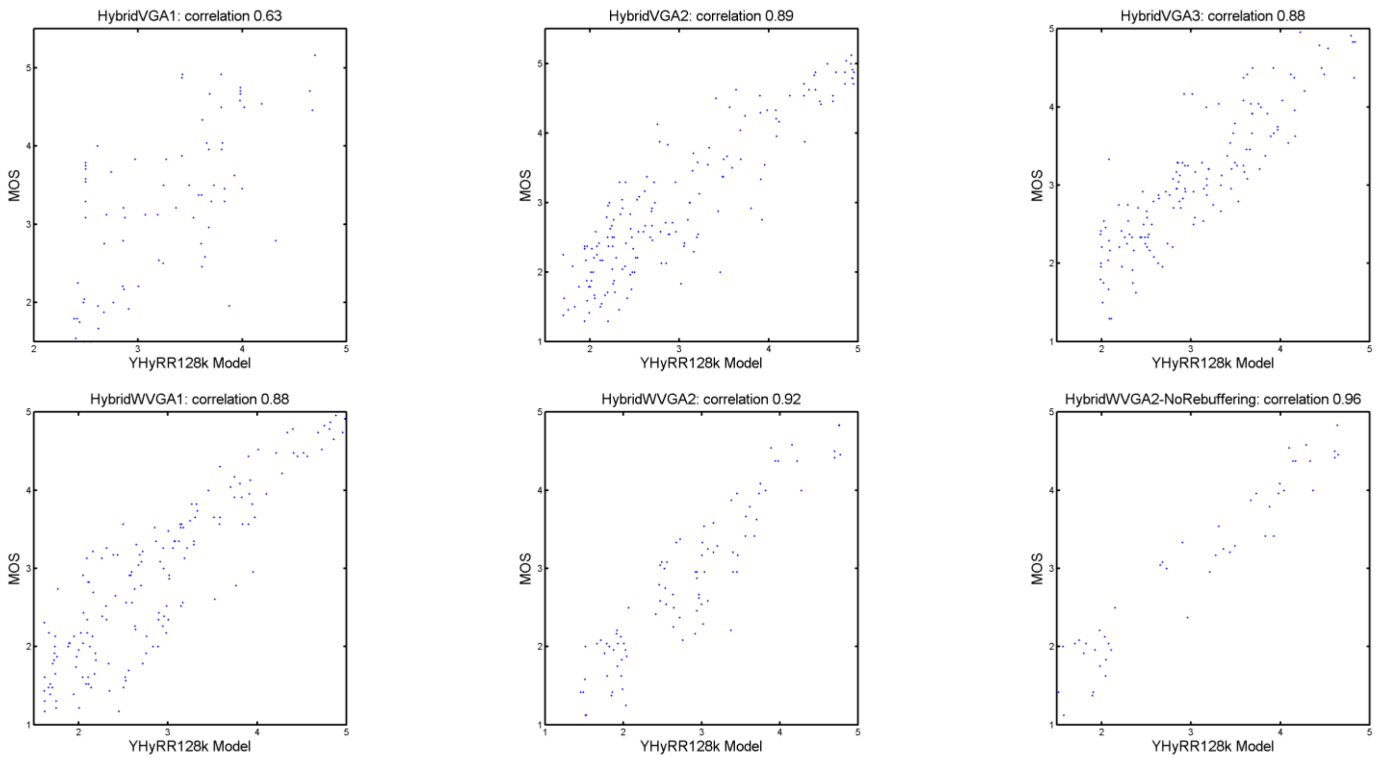


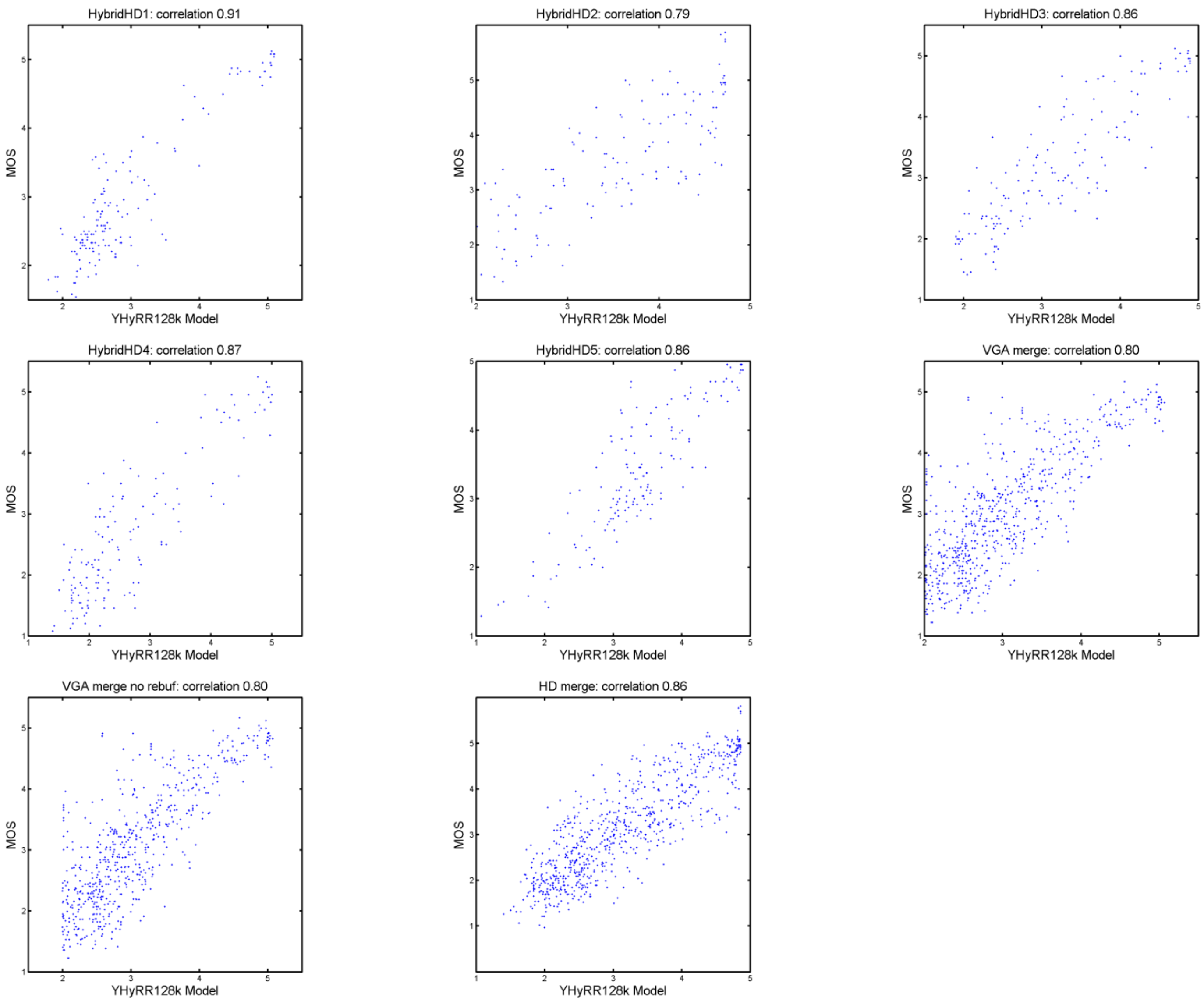



HybridHD1: correlation 0.92
HybridHD2: correlation 0.72
HybridHD3: correlation 0.78
HybridHD4: correlation 0.86
HybridHD5: correlation 0.91
VGA merge: correlation 0.79
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.85
MOS
YHyRR56ke Model

C.3.3 Plots for Model YHyRR128k versus DMOS

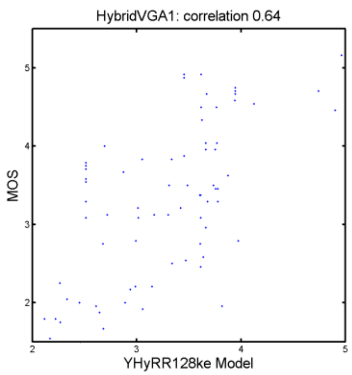
This section contains scatter plots depicting model YHyRR128k versus DMOS.

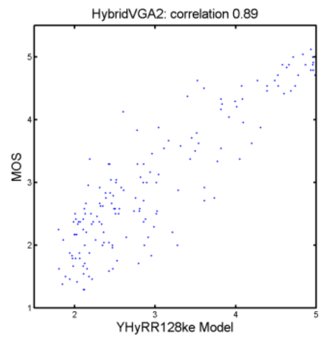


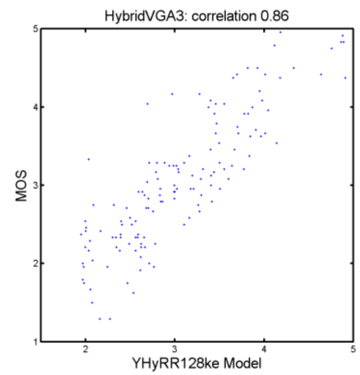

HybridHD1: correlation 0.91
HybridHD2: correlation 0.79
HybridHD3: correlation 0.86
HybridHD4: correlation 0.87
HybridHD5: correlation 0.86
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.86
MOS
YHyRR128k Model

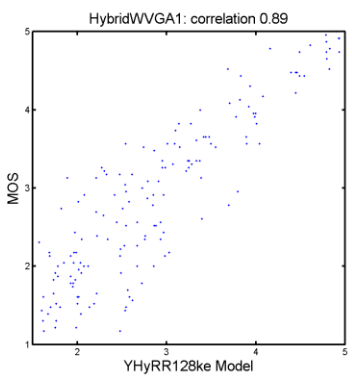
C.3.4 Plots for Model YHyRR128ke versus DMOS

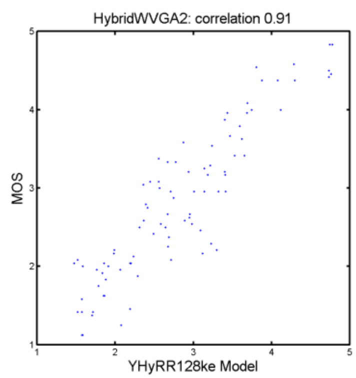
This section contains scatter plots depicting model YHyRR128ke versus DMOS.

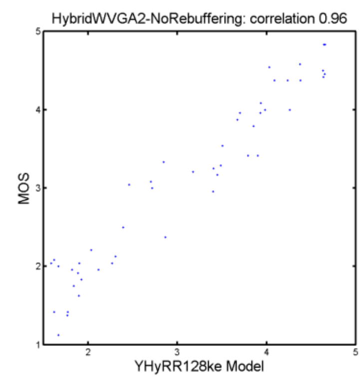


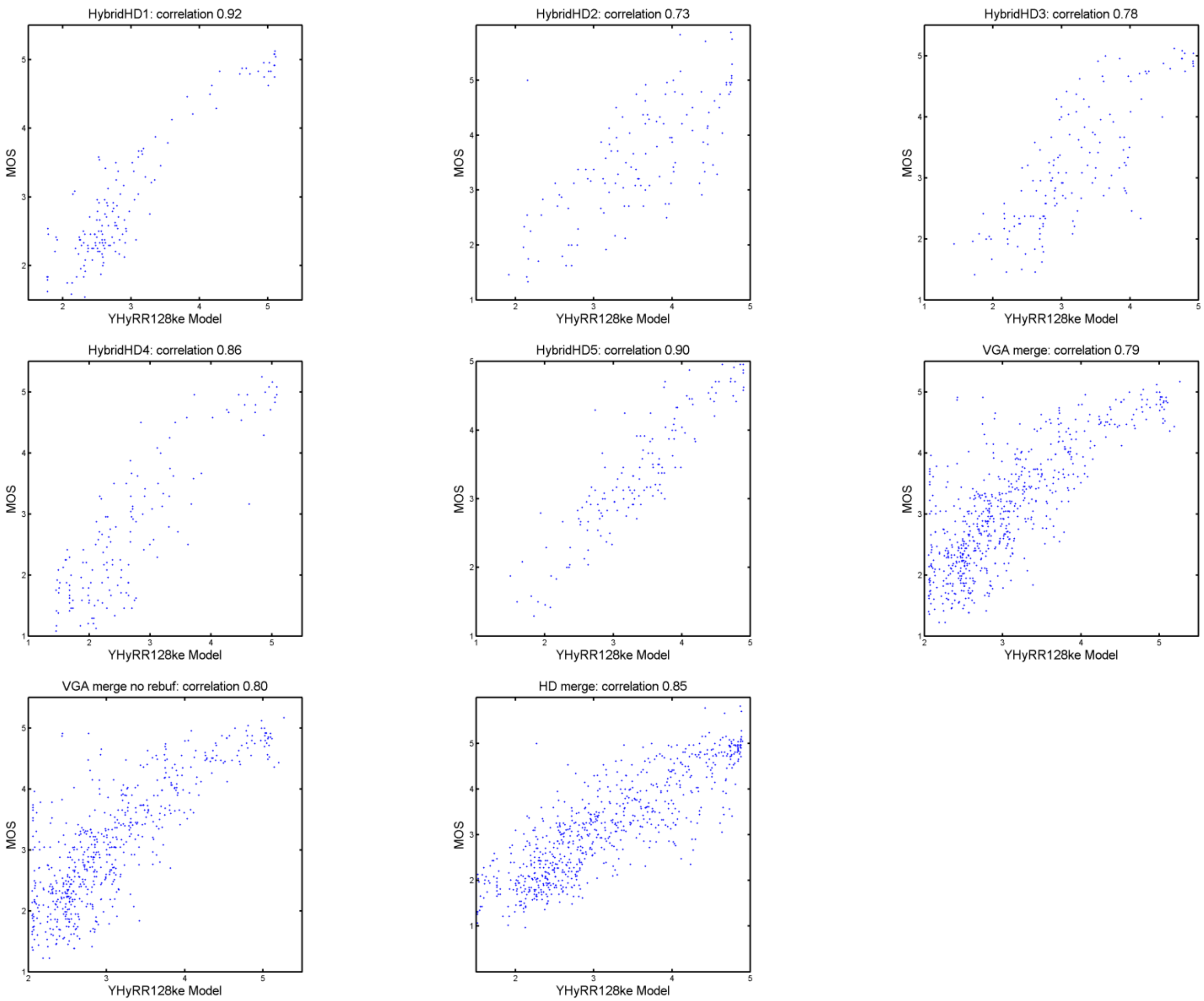






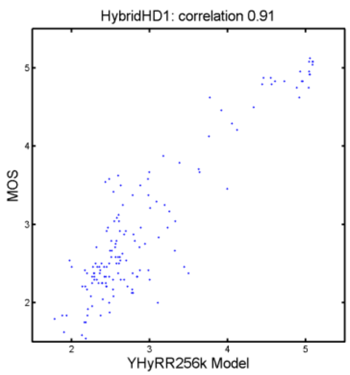


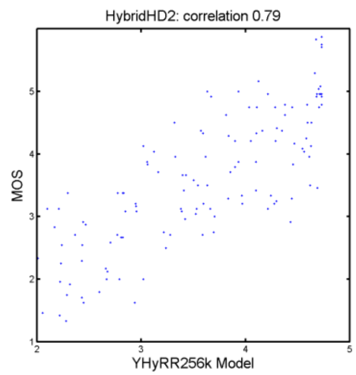


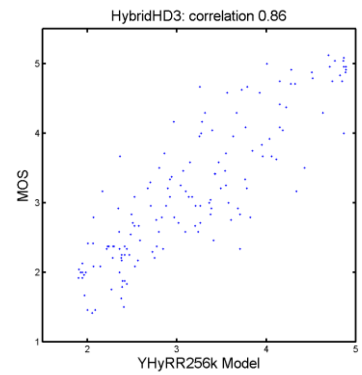

HybridHD1: correlation 0.92
HybridHD2: correlation 0.73
HybridHD3: correlation 0.78
HybridHD4: correlation 0.86
HybridHD5: correlation 0.90
VGA merge: correlation 0.79
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.85
MOS
YHyRR128ke Model

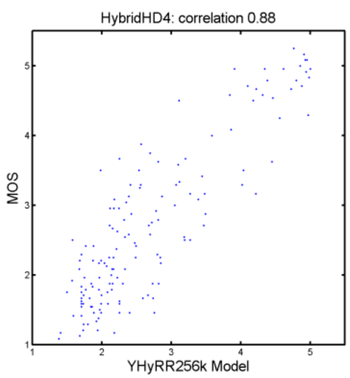
C.3.5 Plots for Model YHyRR256k versus DMOS

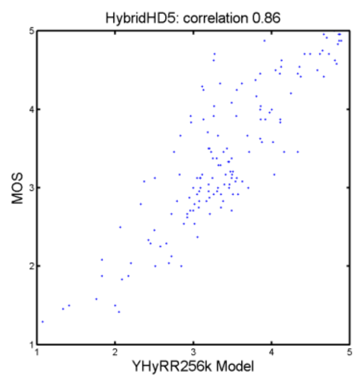
This section contains scatter plots depicting model YHyRR256k versus DMOS.

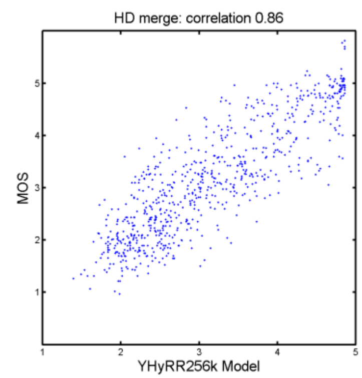






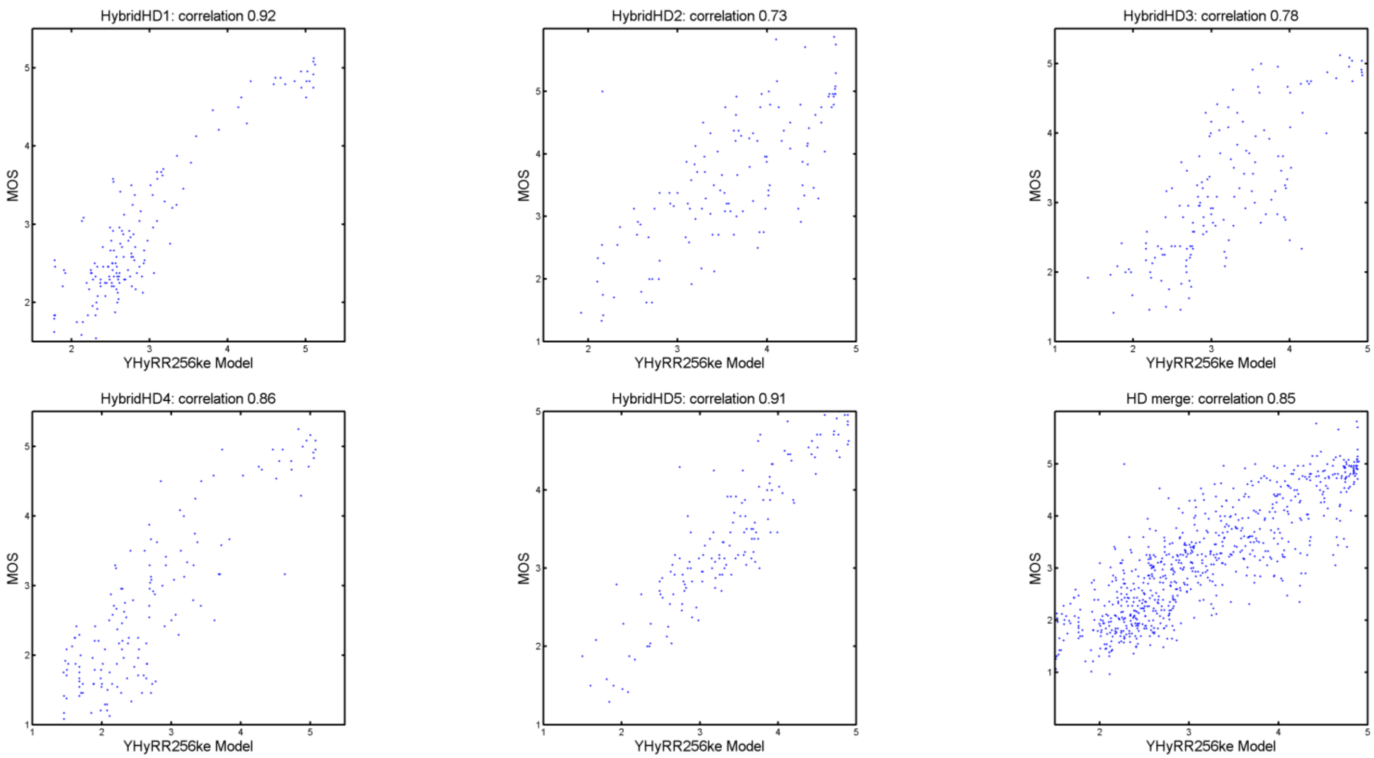






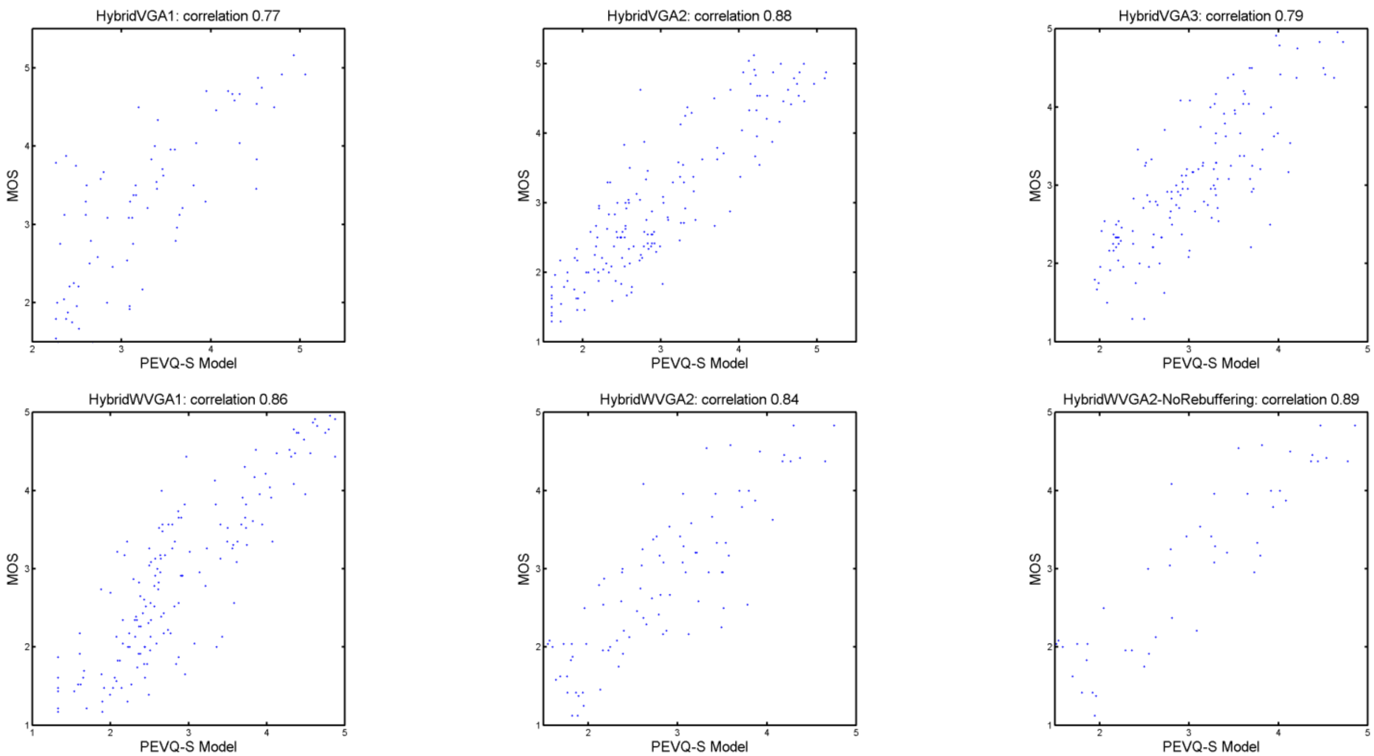
C.3.6 Plots for Model YHyRR256ke versus DMOS

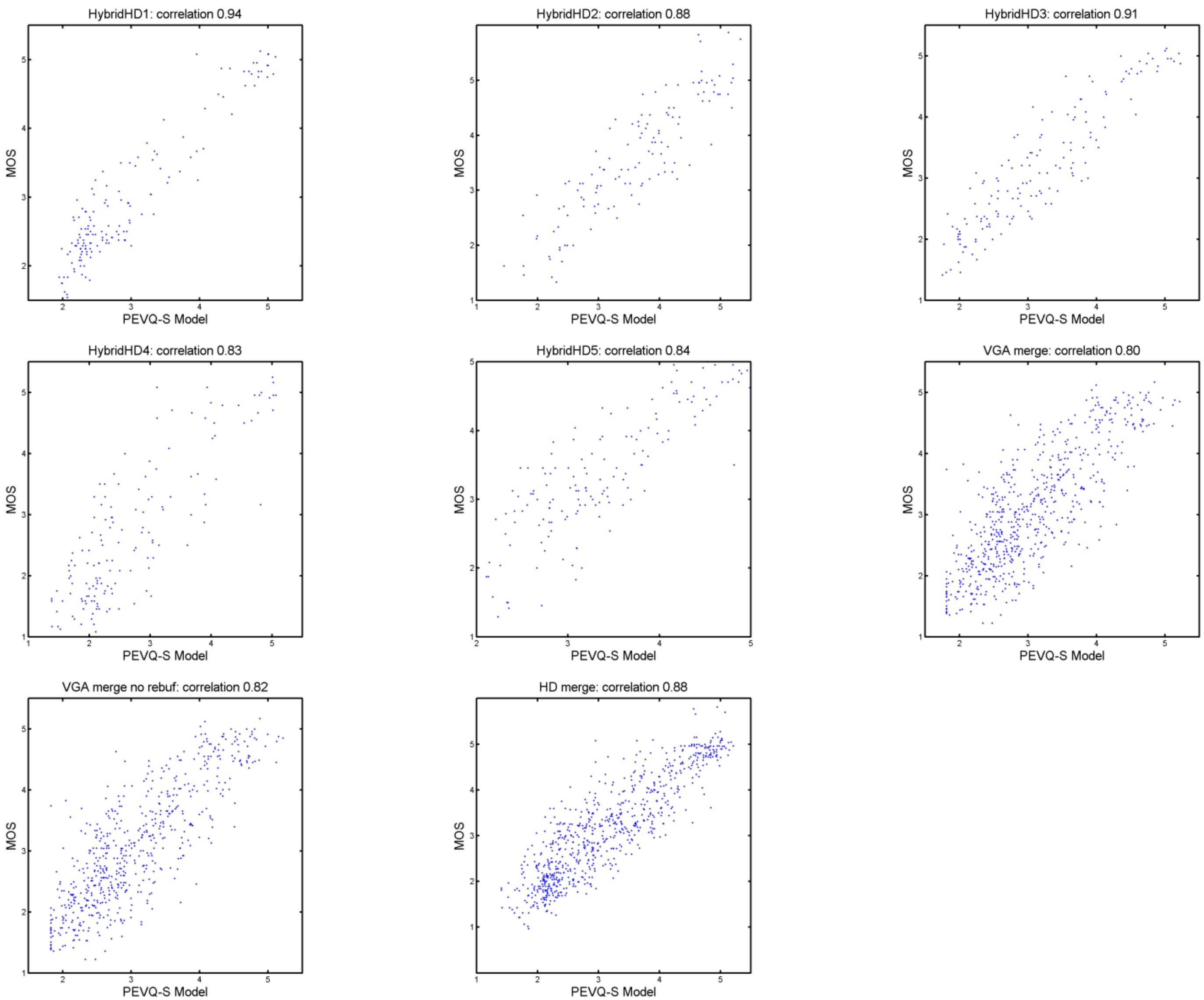
This section contains scatter plots depicting model YHyRR256ke versus DMOS.



C.3.7 Plots for Model PEVQ-S versus DMOS

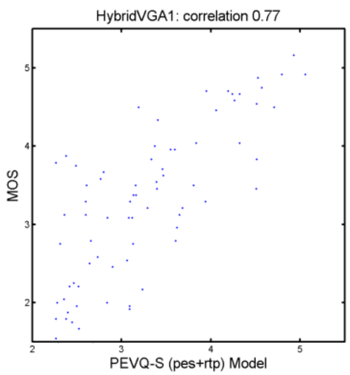
This section contains scatter plots depicting model PEVQ-S versus DMOS.

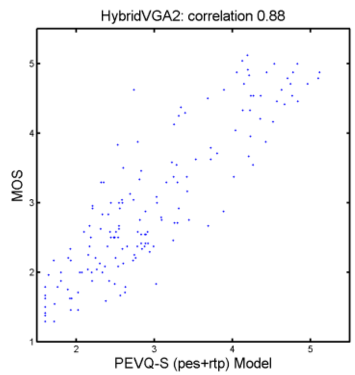


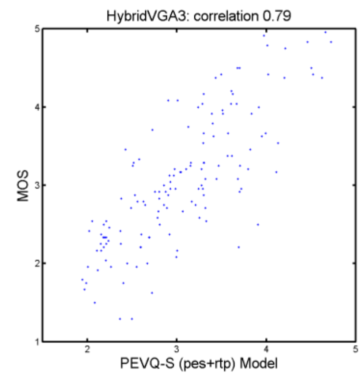

HybridHD1: correlation 0.94
HybridHD2: correlation 0.88
HybridHD3: correlation 0.91
HybridHD4: correlation 0.83
HybridHD5: correlation 0.84
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.82
HD merge: correlation 0.88
MOS
PEVQ-S Model

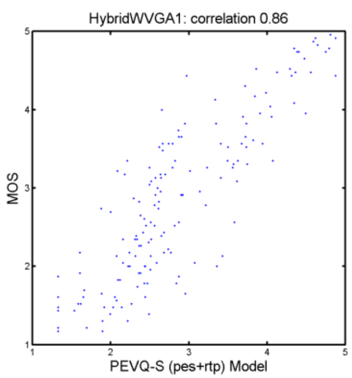
C.3.8 Plots for Model PEVQ-S (pes+rtp) versus DMOS

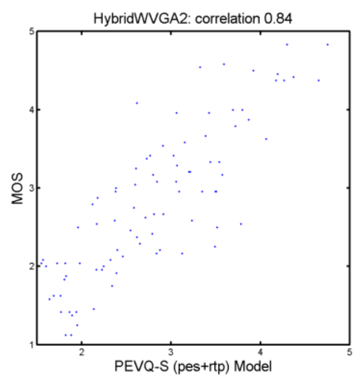
This section contains scatter plots depicting model PEVQ-S (pes+rtp) versus DMOS.

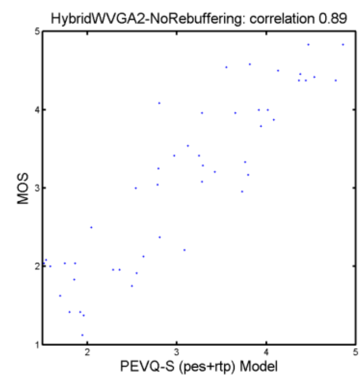


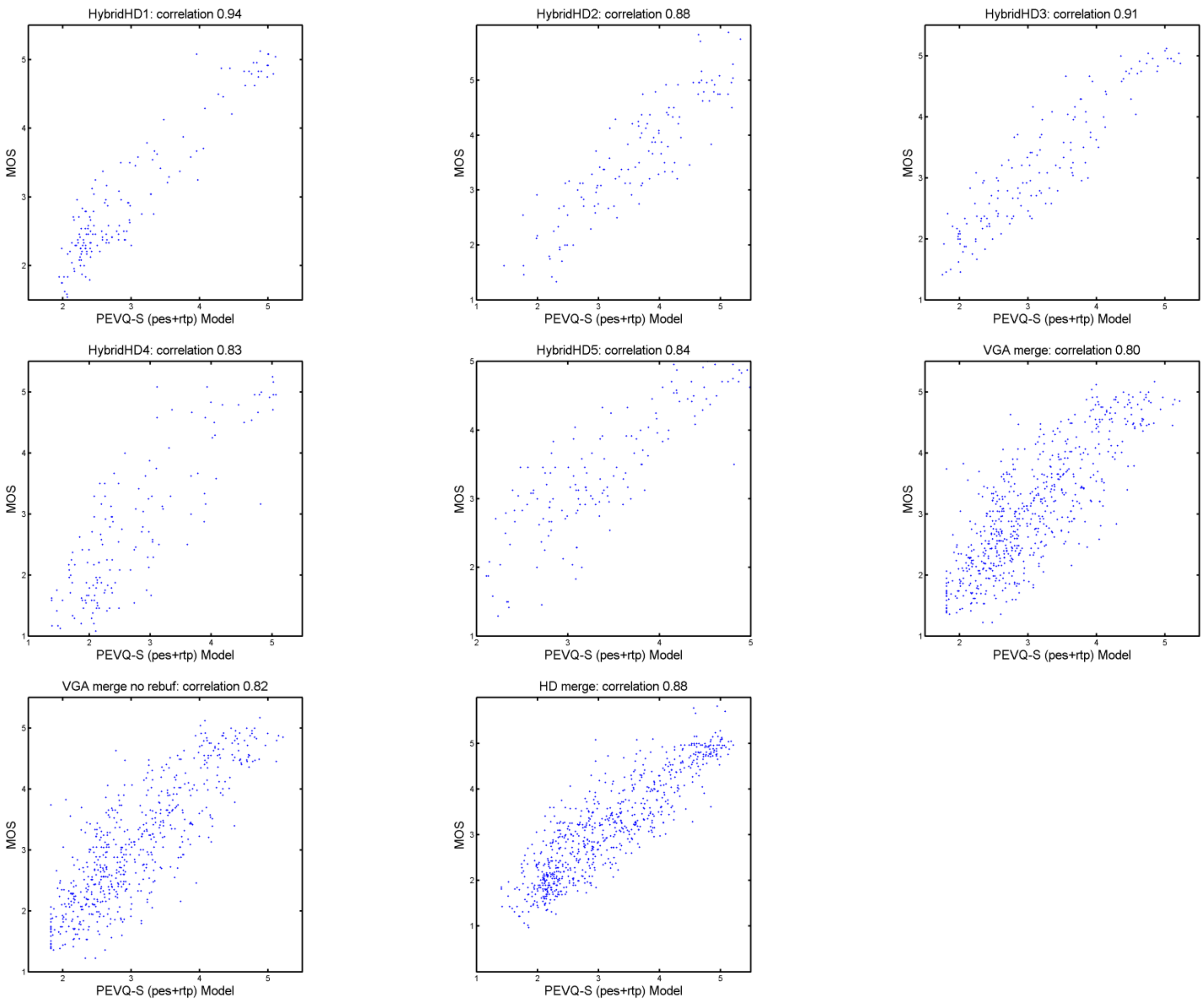






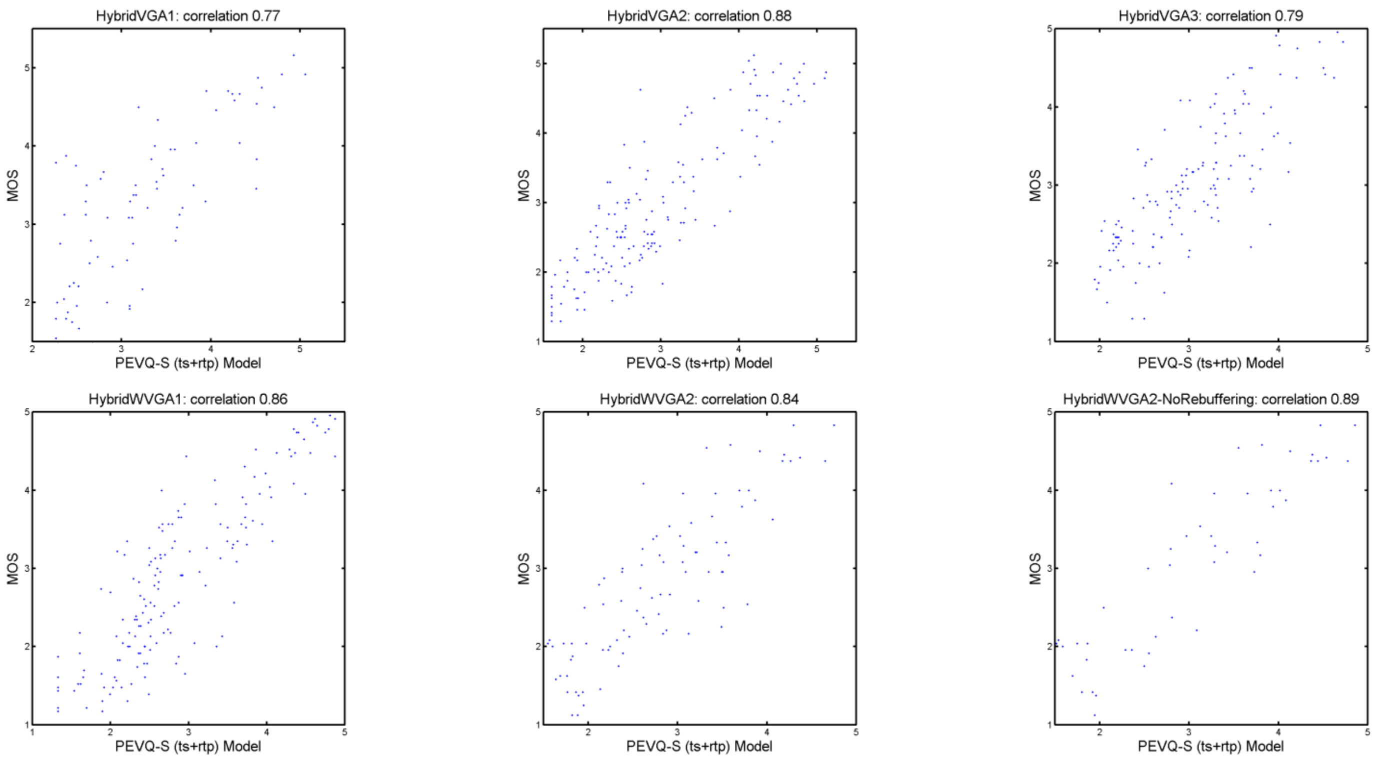


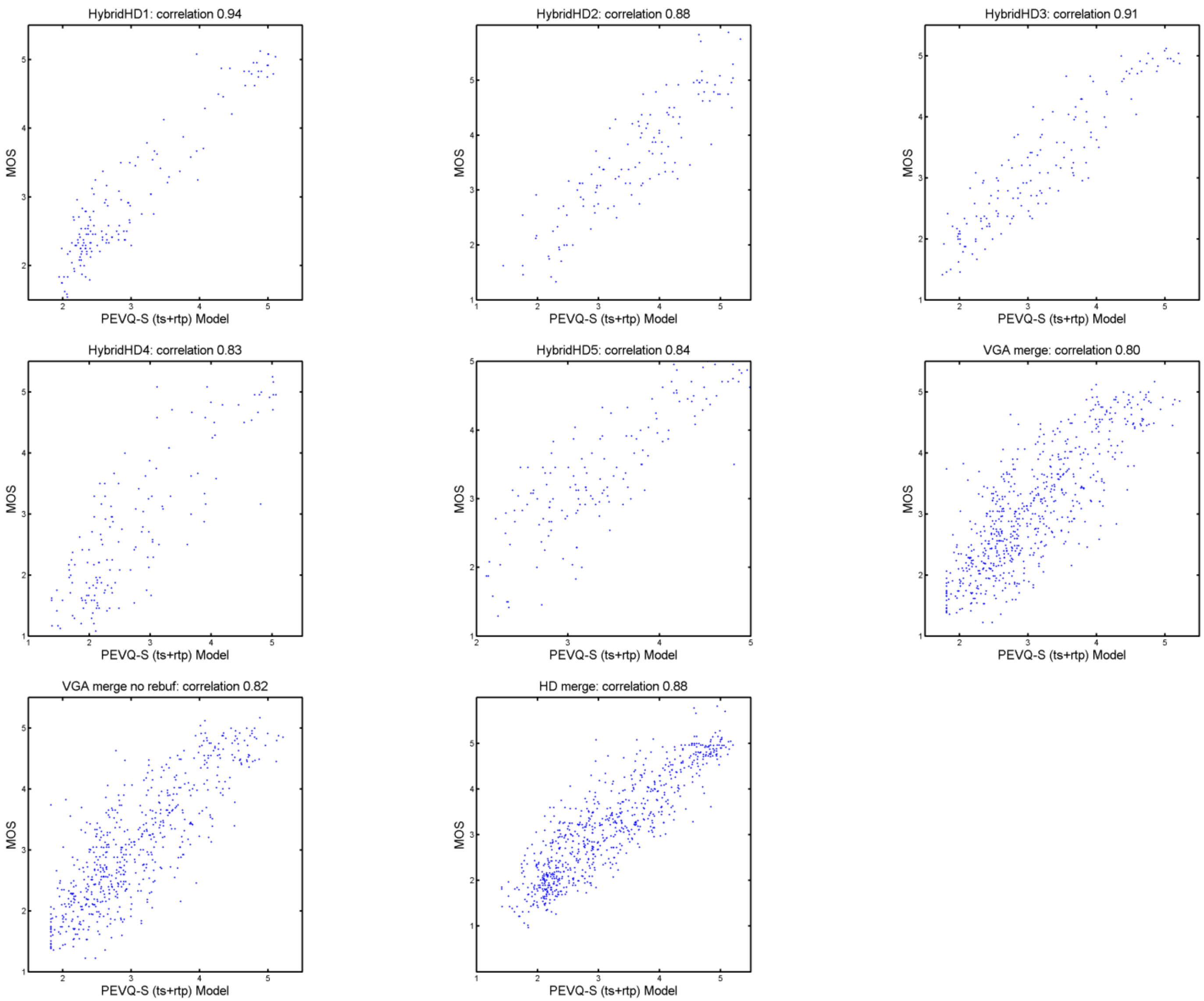



HybridHD1: correlation 0.94
HybridHD2: correlation 0.88
HybridHD3: correlation 0.91
HybridHD4: correlation 0.83
HybridHD5: correlation 0.84
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.82
HD merge: correlation 0.88
MOS
PEVQ-S (pes+rtp) Model

C.3.9Plots for Model PEVQ-S (ts+rtp) versus DMOS

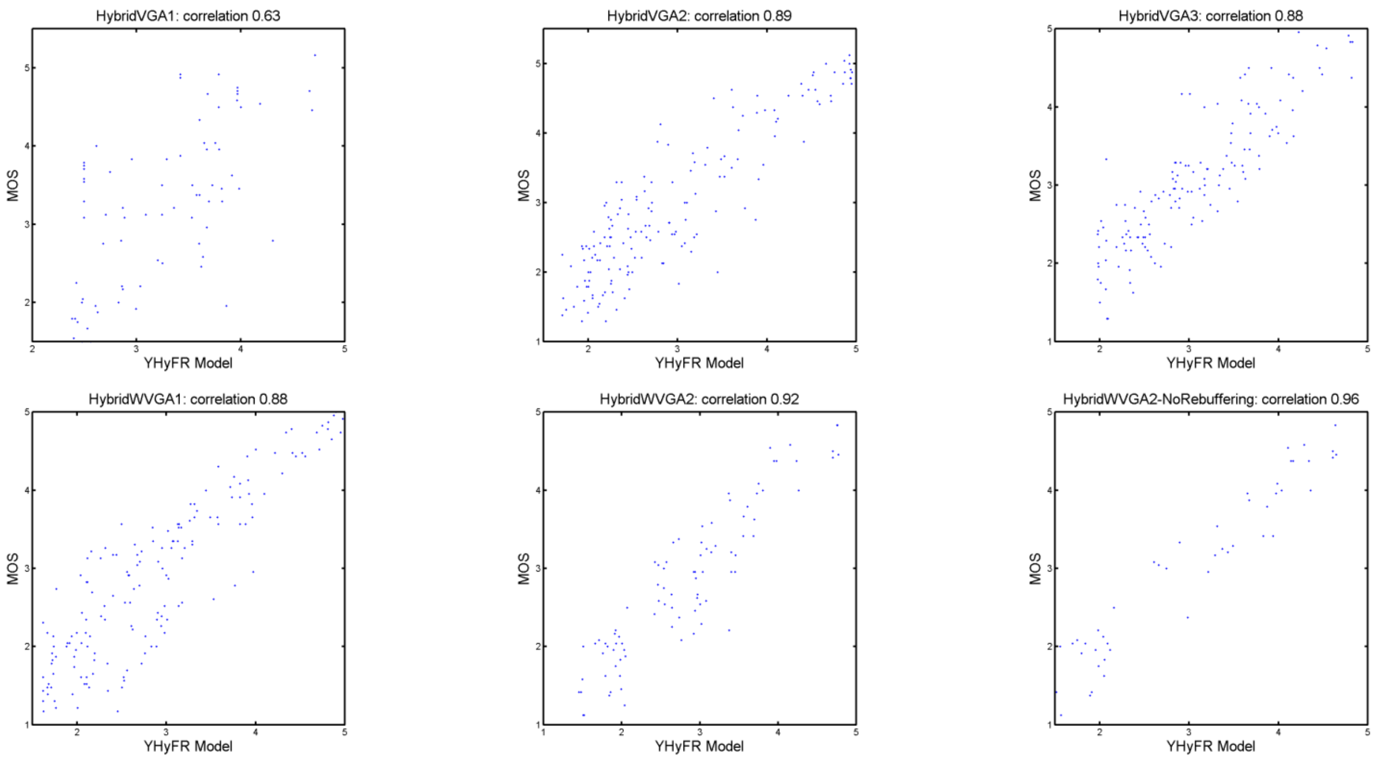
This section contains scatter plots depicting model PEVQ-S (ts+rtp) versus DMOS.

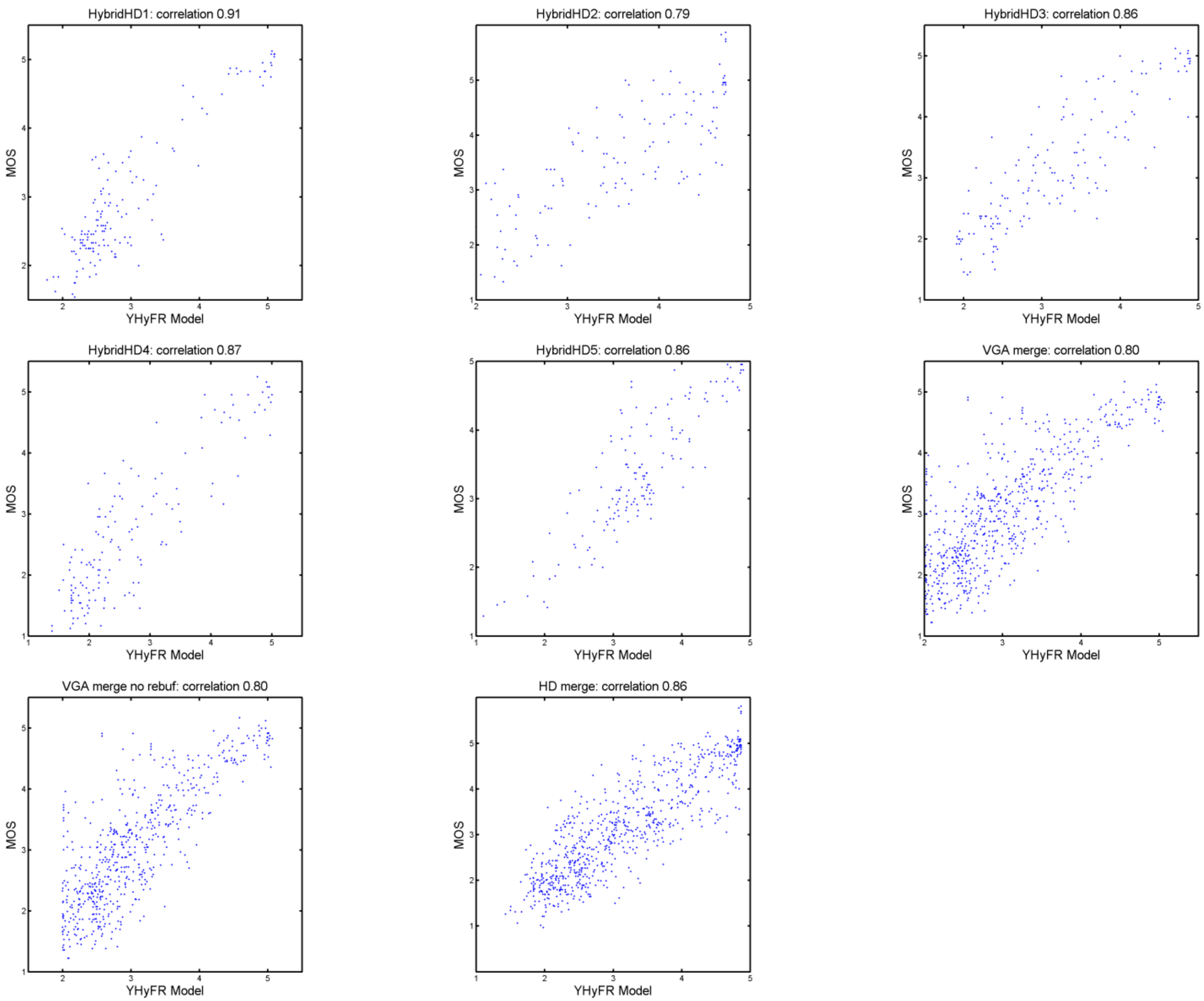



HybridHD1: correlation 0.94
HybridHD2: correlation 0.88
HybridHD3: correlation 0.91
HybridHD4: correlation 0.83
HybridHD5: correlation 0.84
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.82
HD merge: correlation 0.88
MOS
PEVQ-S (ts+rtp) Model

C.3.10 Plots for Model YHyFR versus DMOS

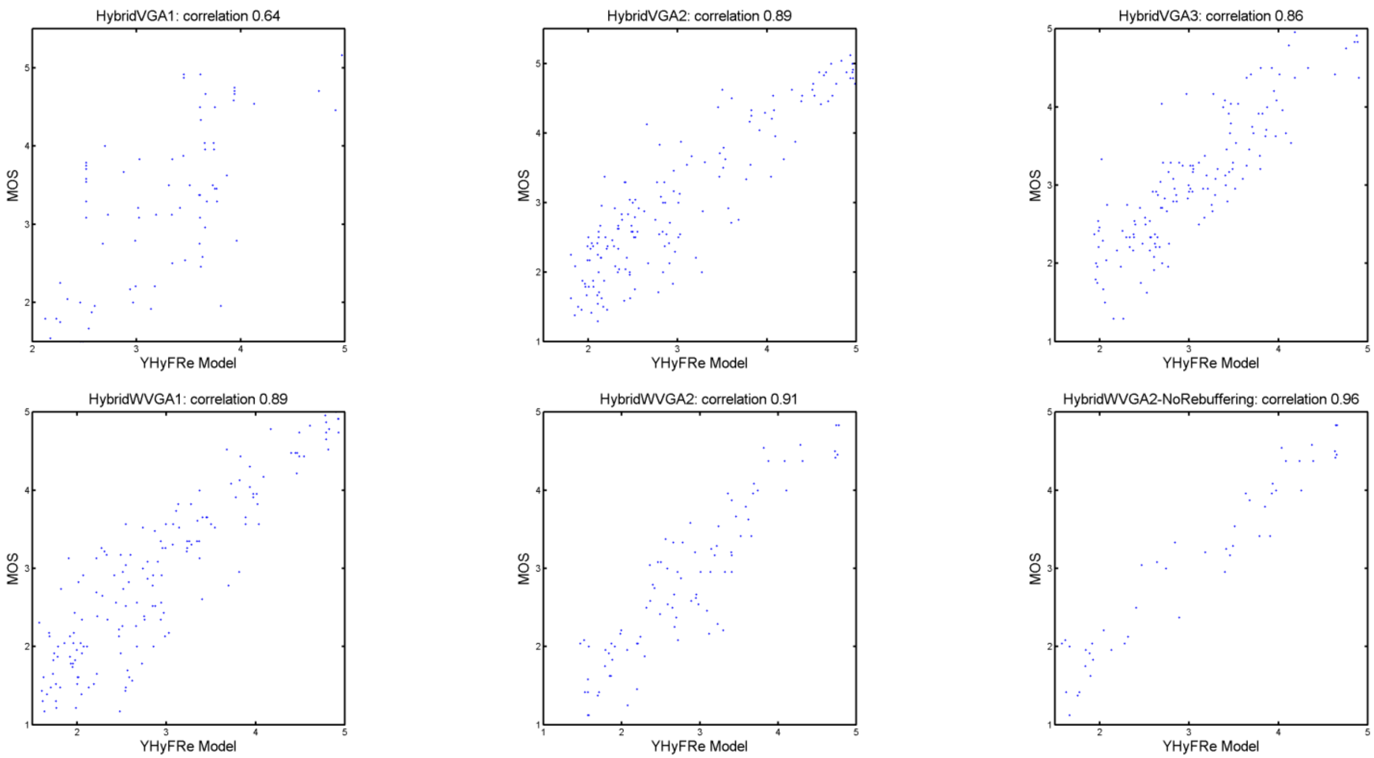
This section contains scatter plots depicting model YHyFR versus DMOS.

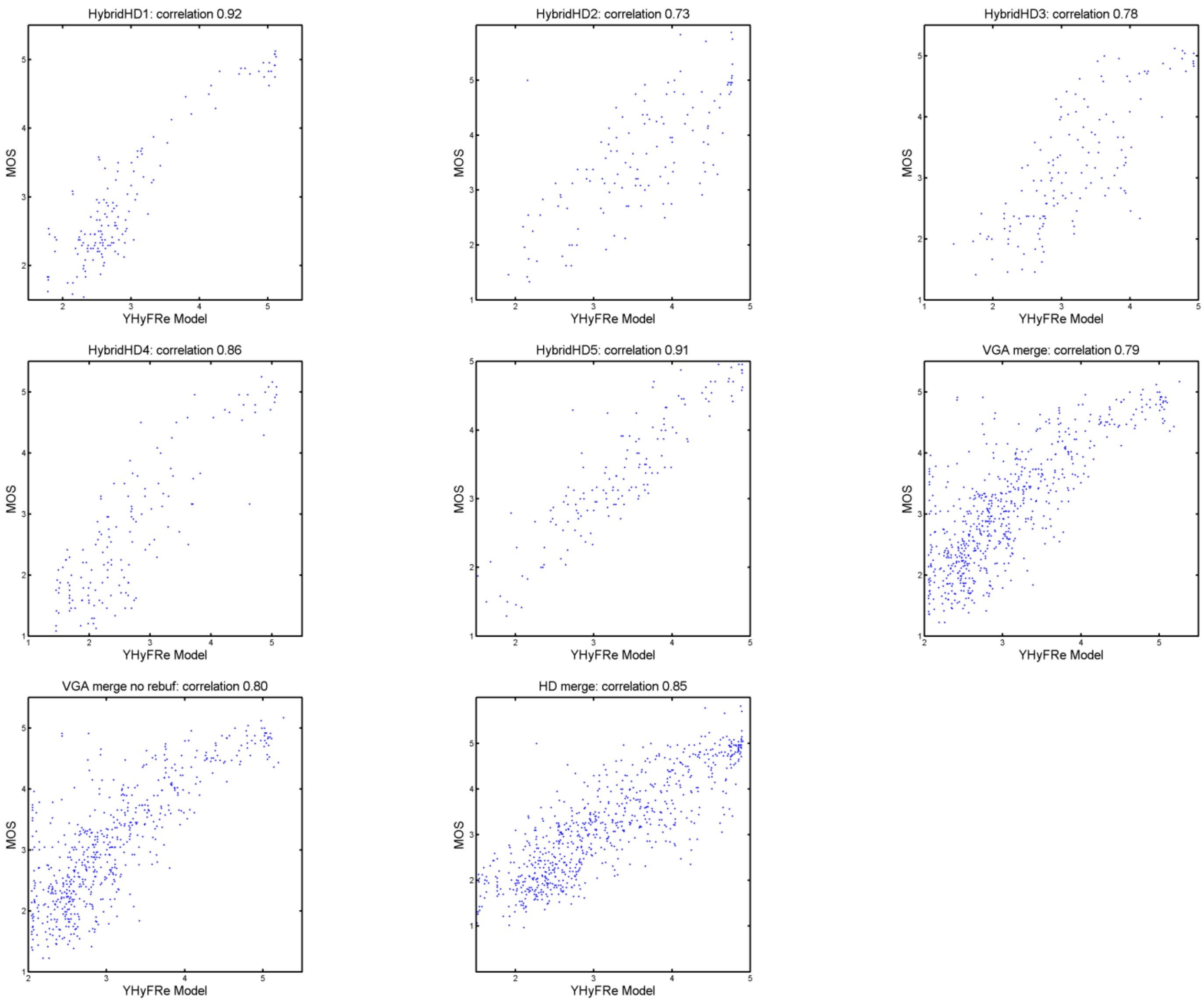



HybridHD1: correlation 0.91
HybridHD2: correlation 0.79
HybridHD3: correlation 0.86
HybridHD4: correlation 0.87
HybridHD5: correlation 0.86
VGA merge: correlation 0.80
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.86
MOS
YHyFR Model

C.3.11 Plots for Model YHyFRe versus DMOS

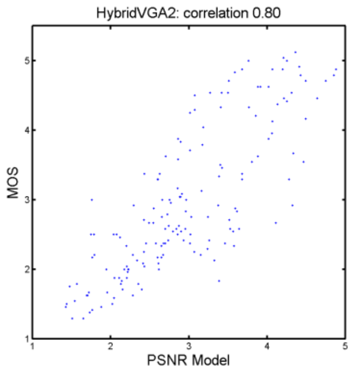
This section contains scatter plots depicting model YHyFRe versus DMOS.

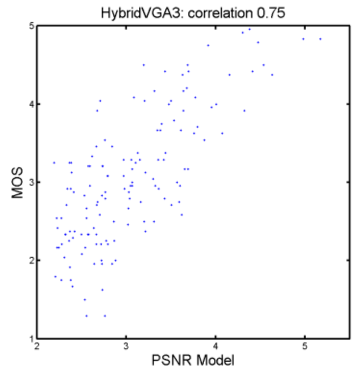


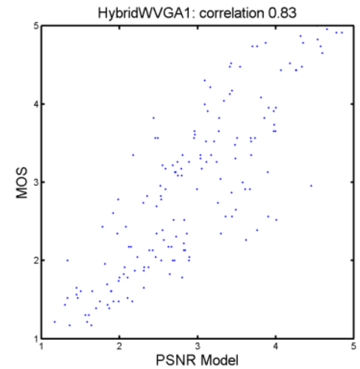

HybridHD1: correlation 0.92
HybridHD2: correlation 0.73
HybridHD3: correlation 0.78
HybridHD4: correlation 0.86
HybridHD5: correlation 0.91
VGA merge: correlation 0.79
VGA merge no rebuf: correlation 0.80
HD merge: correlation 0.85
MOS
YHyFRe Model

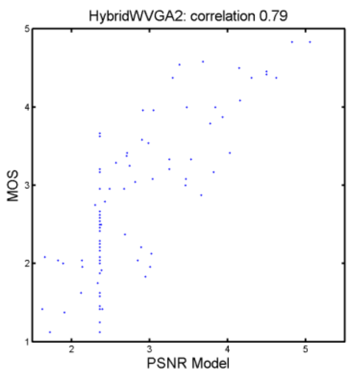
C.3.12 Plots for Reference Model PSNR versus DMOS

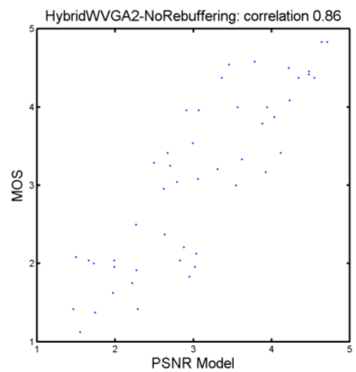
This section contains scatter plots depicting reference model PSNR versus DMOS.

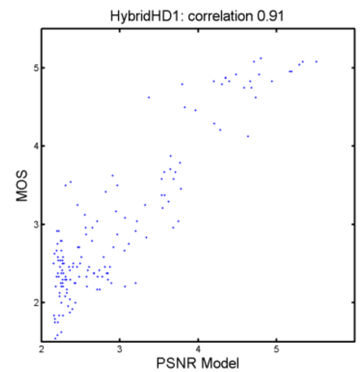


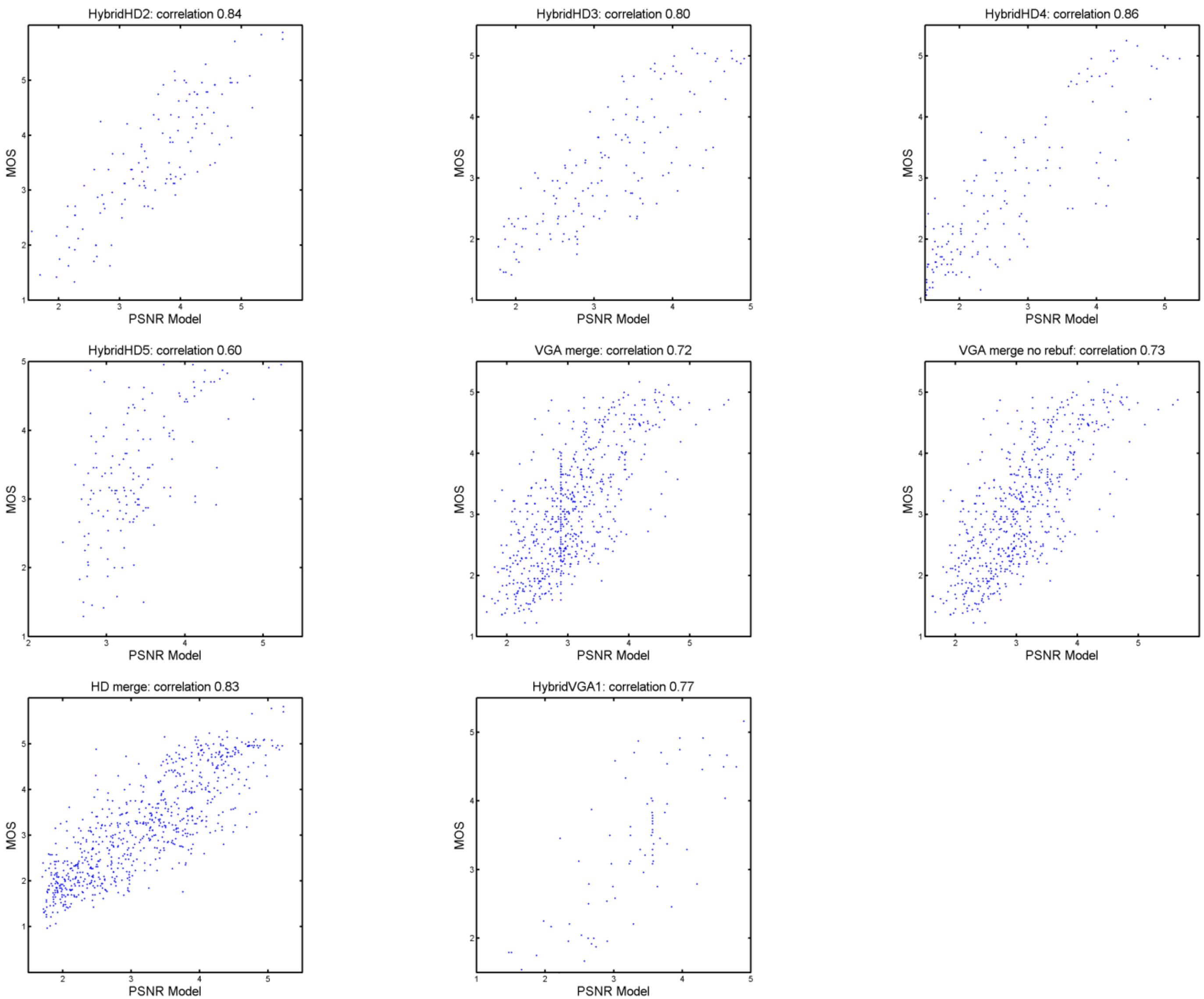











HybridHD2: correlation 0.84
HybridHD3: correlation 0.80
HybridHD4: correlation 0.86
HybridHD5: correlation 0.60
VGA merge: correlation 0.72
VGA merge no rebuf: correlation 0.73
HD merge: correlation 0.83
HybridVGA1: correlation 0.77
MOS
PSNR Model

Annex D Confidence Intervals and Fits

This annex contains the confidence intervals for each metric, and the model fit coefficients. The column labeled "ρ" contains Pearson correlation. The lower and upper confidence intervals are to the right of that statistic.

Table 31. Confidence Intervals and Fits for MOS Analysis

Model TVM-Hybrid Non-Encrypted											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.839	0.756	0.896	0.527	0.63	0.453	0.34	0.258	-2.037	5.557	-2.728
HybridVGA2	0.811	0.751	0.858	0.602	0.678	0.542	0.41	0.196	-1.278	2.919	-0.128
HybridVGA3	0.533	0.403	0.642	0.719	0.816	0.643	0.473	0.055	-0.38	1.168	0.97
HybridWVGA1	0.724	0.641	0.79	0.725	0.816	0.653	0.51	0.115	-0.818	2.306	0.424
HybridWVGA2	0.667	0.531	0.769	0.739	0.871	0.642	0.568	-0.035	0.165	0.559	0.453
HybridWVGA2 no rebuf	0.569	0.339	0.734	0.945	1.193	0.782	0.754	-0.078	0.585	-0.642	1.344
VGA merge	0.689	0.646	0.728	0.683	0.723	0.648	0.479	0.106	-0.772	2.184	0.173
VGA merge no rebuf	0.69	0.646	0.73	0.695	0.737	0.658	0.489	0.113	-0.811	2.24	0.163
HybridHD1	0.837	0.781	0.879	0.515	0.582	0.462	0.369	0.105	-0.567	1.352	0.812
HybridHD2	0.922	0.891	0.945	0.389	0.444	0.346	0.176	-0.066	0.459	0.112	0.457
HybridHD3	0.897	0.86	0.924	0.44	0.497	0.395	0.253	-0.153	1.204	-1.888	2.351
HybridHD4	0.643	0.538	0.728	0.892	1.007	0.8	0.683	-0.135	1.077	-1.79	2.752
HybridHD5	0.872	0.828	0.906	0.427	0.482	0.383	0.233	-0.064	0.473	-0.226	1.288
HD merge	0.809	0.783	0.833	0.603	0.635	0.574	0.42	-0.086	0.722	-1.001	1.986
Model TVM-Hybrid Encrypted											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.852	0.775	0.904	0.507	0.607	0.436	0.282	0.037	-0.117	0.465	0.512
HybridVGA2	0.821	0.763	0.866	0.589	0.662	0.53	0.396	0.403	-3.513	10.205	-7.763
HybridVGA3	0.772	0.695	0.831	0.541	0.613	0.483	0.305	0.054	-0.317	1.168	0.203
HybridWVGA1	0.738	0.658	0.801	0.71	0.798	0.639	0.49	0.252	-2.172	6.589	-4.522
HybridWVGA2	0.783	0.686	0.852	0.617	0.727	0.536	0.423	0.109	-1.234	5.283	-5.23
HybridWVGA2 no rebuf	0.934	0.885	0.963	0.41	0.518	0.339	0.239	0.146	-1.406	5.403	-5.019

VGA merge	0.76	0.725	0.791	0.612	0.648	0.581	0.409	0.183	-1.566	4.887	-3.215
VGA merge no rebuf	0.775	0.741	0.805	0.607	0.643	0.574	0.403	0.191	-1.615	4.952	-3.208
HybridHD1	0.863	0.816	0.899	0.475	0.537	0.426	0.284	0.101	-0.681	1.788	0.313
HybridHD2	0.716	0.62	0.792	0.701	0.801	0.624	0.495	0.537	-5.523	19.013	-19.269
HybridHD3	0.67	0.571	0.75	0.739	0.834	0.663	0.52	0.076	-0.542	1.781	0.085
HybridHD4	0.716	0.627	0.786	0.813	0.918	0.73	0.611	-0.001	0.303	-0.842	2.449
HybridHD5	0.827	0.769	0.872	0.491	0.554	0.44	0.3	0.265	-2.784	10.168	-9.732
HD merge	0.74	0.705	0.771	0.691	0.728	0.657	0.49	0.065	-0.489	1.789	-0.22
Model VMon-B											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.828	0.74	0.888	0.544	0.651	0.467	0.356	0.213	-1.704	4.926	-2.602
HybridVGA2	0.826	0.77	0.87	0.58	0.653	0.522	0.387	0.263	-1.913	4.842	-1.924
HybridVGA3	0.585	0.465	0.685	0.689	0.782	0.616	0.446	-0.031	0.194	0.164	1.214
HybridWVGA1	0.735	0.654	0.799	0.713	0.802	0.642	0.5	0.105	-0.757	2.238	0.383
HybridWVGA2	0.667	0.531	0.769	0.739	0.871	0.642	0.568	-0.035	0.165	0.559	0.453
HybridWVGA2 no rebuf	0.569	0.339	0.734	0.945	1.193	0.782	0.754	-0.078	0.585	-0.642	1.344
VGA merge	0.695	0.653	0.733	0.678	0.717	0.642	0.472	0.056	-0.425	1.515	0.451
VGA merge no rebuf	0.695	0.651	0.734	0.691	0.733	0.654	0.482	0.06	-0.44	1.522	0.47
HybridHD1	0.829	0.771	0.873	0.527	0.595	0.472	0.369	0.131	-0.622	0.983	1.49
HybridHD2	0.904	0.866	0.931	0.431	0.492	0.383	0.219	-0.111	0.946	-1.406	1.679
HybridHD3	0.906	0.872	0.931	0.422	0.477	0.379	0.233	-0.185	1.464	-2.464	2.558
HybridHD4	0.573	0.455	0.672	0.954	1.077	0.856	0.74	-0.102	0.876	-1.502	2.542
HybridHD5	0.876	0.832	0.908	0.422	0.477	0.379	0.228	-0.093	0.732	-0.885	1.687
HD merge	0.782	0.752	0.808	0.64	0.674	0.609	0.452	-0.104	0.916	-1.578	2.342
Model VMon											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.841	0.758	0.897	0.525	0.628	0.451	0.301	0.037	-0.154	0.717	0.151
HybridVGA2	0.834	0.78	0.876	0.568	0.639	0.511	0.376	0.439	-3.886	11.458	-9.097
HybridVGA3	0.809	0.743	0.86	0.5	0.567	0.447	0.265	-0.006	0.151	0.078	1.066
HybridWVGA1	0.753	0.677	0.813	0.692	0.778	0.623	0.471	0.19	-1.709	5.698	-4.246
HybridWVGA2	0.767	0.665	0.841	0.636	0.749	0.553	0.434	-0.065	0.435	0.151	-0.107

HybridWVGA2 no rebuf	0.912	0.848	0.95	0.47	0.594	0.389	0.277	-0.148	1.388	-2.986	3.086
VGA merge	0.769	0.735	0.799	0.603	0.638	0.572	0.399	0.146	-1.296	4.373	-3.015
VGA merge no rebuf	0.782	0.749	0.811	0.599	0.635	0.567	0.394	0.157	-1.371	4.524	-3.096
HybridHD1	0.863	0.816	0.899	0.475	0.537	0.426	0.285	0.042	-0.154	0.368	1.502
HybridHD2	0.728	0.635	0.8	0.689	0.787	0.613	0.479	0.421	-4.23	14.291	-13.537
HybridHD3	0.657	0.556	0.74	0.75	0.847	0.673	0.536	0.025	-0.051	0.268	1.583
HybridHD4	0.724	0.638	0.792	0.803	0.907	0.72	0.601	-0.03	0.515	-1.237	2.579
HybridHD5	0.875	0.831	0.908	0.423	0.478	0.38	0.228	0.107	-1.106	4.537	-3.79
HD merge	0.755	0.722	0.784	0.673	0.709	0.64	0.474	-0.002	0.152	-0.09	1.511
Model YHyNR											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.689	0.548	0.792	0.702	0.841	0.604	0.5	2.717	-21.275	55.536	-45.605
HybridVGA2	0.878	0.837	0.909	0.493	0.554	0.444	0.316	0.09	-0.61	2.617	-1.806
HybridVGA3	0.79	0.718	0.845	0.522	0.592	0.466	0.291	-0.176	1.353	-2.142	2.28
HybridWVGA1	0.825	0.768	0.869	0.594	0.669	0.535	0.4	0.221	-1.694	5.408	-3.637
HybridWVGA2	0.87	0.808	0.913	0.489	0.576	0.425	0.294	0.411	-3.114	8.571	-6.182
HybridWVGA2 no rebuff	0.929	0.877	0.96	0.424	0.535	0.351	0.24	0.03	-0.127	1.252	-0.469
VGA merge	0.778	0.746	0.807	0.592	0.626	0.561	0.395	0.164	-1.266	4.244	-2.737
VGA merge no rebuf	0.788	0.756	0.817	0.591	0.627	0.56	0.395	0.069	-0.511	2.366	-1.256
HybridHD1	0.899	0.863	0.926	0.412	0.465	0.369	0.216	0.16	-0.901	1.964	0.35
HybridHD2	0.842	0.783	0.886	0.542	0.619	0.482	0.325	0.513	-4.901	16.487	-16.98
HybridHD3	0.883	0.842	0.914	0.467	0.528	0.419	0.269	0.02	0.065	0.178	0.904
HybridHD4	0.802	0.737	0.853	0.695	0.785	0.624	0.497	-0.037	0.675	-1.623	2.681
HybridHD5	0.822	0.762	0.868	0.497	0.561	0.446	0.312	0.399	-3.315	9.579	-6.954
HD merge	0.836	0.813	0.856	0.564	0.594	0.536	0.368	0.071	-0.39	1.399	-0.03
Model YHyNRe											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.689	0.548	0.792	0.702	0.84	0.603	0.5	0.34	-2.778	7.579	-4.172
HybridVGA2	0.834	0.78	0.876	0.568	0.639	0.512	0.393	0.327	-2.621	7.812	-5.965
HybridVGA3	0.707	0.613	0.781	0.601	0.682	0.537	0.364	-0.125	1.101	-2.128	2.953
HybridWVGA1	0.77	0.699	0.827	0.67	0.754	0.603	0.467	0.215	-1.659	5.128	-3.162

HybridWVGA2	0.828	0.748	0.884	0.557	0.656	0.484	0.373	0.516	-4.279	12.248	-9.58
HybridWVGA2 no rebuff	0.911	0.846	0.949	0.474	0.599	0.392	0.302	0.2	-1.73	5.861	-4.469
VGA merge	0.717	0.677	0.753	0.657	0.695	0.623	0.459	-0.019	0.39	-0.854	2.362
VGA merge no rebuf	0.728	0.688	0.763	0.659	0.699	0.624	0.461	-0.029	0.464	-0.984	2.415
HybridHD1	0.933	0.909	0.951	0.337	0.381	0.303	0.155	0.023	0.112	-0.236	1.511
HybridHD2	0.699	0.598	0.778	0.719	0.821	0.639	0.509	0.737	-7.216	23.987	-24.527
HybridHD3	0.709	0.619	0.781	0.701	0.792	0.629	0.494	0.082	-0.158	-0.472	2.842
HybridHD4	0.772	0.699	0.83	0.739	0.835	0.664	0.555	0.269	-1.855	4.837	-2.648
HybridHD5	0.781	0.709	0.836	0.546	0.616	0.49	0.354	0.135	-1.239	4.843	-4.531
HD merge	0.787	0.758	0.813	0.633	0.667	0.602	0.443	0.049	-0.194	0.85	0.226
Model YNR											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.595	0.425	0.724	0.779	0.932	0.669	0.55	0.062	-0.046	0.123	1.083
HybridVGA2	0.53	0.409	0.633	0.873	0.982	0.786	0.666	0.059	-0.193	0.517	1.107
HybridVGA3	0.54	0.411	0.648	0.716	0.812	0.64	0.481	-0.114	0.968	-1.824	2.447
HybridWVGA1	0.668	0.572	0.746	0.782	0.879	0.704	0.576	-0.24	2.088	-4.5	3.704
HybridWVGA2	0.631	0.485	0.742	0.77	0.907	0.669	0.577	0.165	-0.651	0.851	0.853
HybridWVGA2 no rebuf	0.61	0.394	0.762	0.91	1.15	0.753	0.707	0.455	-2.912	6.515	-3.744
VGA merge	0.541	0.484	0.594	0.793	0.839	0.752	0.581	-0.038	0.48	-0.846	1.735
VGA merge no rebuf	0.541	0.482	0.596	0.808	0.856	0.764	0.592	-0.047	0.532	-0.924	1.751
HybridHD1	0.787	0.717	0.841	0.581	0.656	0.521	0.392	0.143	-0.607	0.882	1.235
HybridHD2	0.694	0.591	0.774	0.724	0.827	0.644	0.504	-0.026	0.479	-0.879	1.498
HybridHD3	0.54	0.416	0.644	0.837	0.946	0.751	0.616	0.178	-1.246	3.466	-1.493
HybridHD4	0.7	0.608	0.773	0.832	0.939	0.746	0.642	0.06	-0.186	0.584	0.853
HybridHD5	0.546	0.423	0.649	0.732	0.826	0.657	0.524	1.661	-15.315	47.238	-46.052
HD merge	0.671	0.629	0.709	0.761	0.802	0.724	0.555	0.058	-0.186	0.592	0.796
Model PSNR											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.717	0.585	0.812	0.676	0.809	0.581	0.465	0	0.001	0.05	0.412
HybridVGA2	0.799	0.735	0.849	0.62	0.697	0.558	0.421	0	0.014	-0.304	3.087
HybridVGA3	0.718	0.627	0.79	0.592	0.671	0.529	0.357	0	0.014	-0.304	3.682

HybridWVGA1	0.806	0.744	0.854	0.622	0.7	0.56	0.425	0	0.029	-0.738	6.628
HybridWVGA2	0.757	0.65	0.834	0.649	0.764	0.564	0.46	0	0	0.198	-4.184
HybridWVGA2 no rebuf	0.824	0.705	0.898	0.651	0.822	0.539	0.452	-0.001	0.064	-1.754	15.213
VGA merge	0.68	0.636	0.719	0.691	0.732	0.655	0.477	0	0.002	-0.031	1.41
VGA merge no rebuf	0.69	0.646	0.73	0.695	0.737	0.658	0.479	0	0.002	-0.033	1.428
HybridHD1	0.866	0.819	0.901	0.471	0.532	0.422	0.264	0	0.007	-0.171	2.92
HybridHD2	0.764	0.681	0.828	0.648	0.74	0.576	0.426	0	0.009	-0.094	0.34
HybridHD3	0.775	0.702	0.832	0.628	0.71	0.564	0.409	0	0.01	-0.221	2.56
HybridHD4	0.863	0.816	0.899	0.587	0.663	0.527	0.392	0	0.022	-0.49	4.588
HybridHD5	0.582	0.466	0.679	0.71	0.802	0.637	0.512	0	-0.031	0.836	-5.403
HD merge	0.791	0.762	0.816	0.629	0.662	0.598	0.426	0	0.012	-0.279	3.276

Table 32. Confidence Intervals and Fits for DMOS Analysis

Model PEVQ-S											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.774	0.664	0.851	0.648	0.775	0.556	0.369	0.057	-0.303	0.77	1.742
HybridVGA2	0.877	0.835	0.908	0.511	0.575	0.46	0.278	0.012	0.006	0.448	1.14
HybridVGA3	0.793	0.722	0.847	0.519	0.589	0.464	0.291	0.036	-0.184	0.736	1.357
HybridWVGA1	0.858	0.811	0.894	0.541	0.609	0.487	0.326	0.064	-0.468	1.84	-0.109
HybridWVGA2	0.837	0.762	0.891	0.534	0.629	0.464	0.325	-0.018	0.513	-1.888	3.414
HybridWVGA2 no rebuf	0.894	0.818	0.94	0.508	0.642	0.421	0.286	-0.065	0.96	-3.133	4.459
VGA merge	0.805	0.775	0.83	0.565	0.597	0.535	0.347	0.062	-0.414	1.353	0.807
VGA merge no rebuf	0.825	0.798	0.849	0.548	0.581	0.518	0.327	0.047	-0.282	1.043	1.017
HybridHD1	0.938	0.915	0.955	0.339	0.383	0.304	0.123	-0.036	0.387	-0.5	2.025
HybridHD2	0.883	0.837	0.916	0.505	0.577	0.449	0.28	0.045	-0.479	2.474	-1.127
HybridHD3	0.913	0.881	0.936	0.414	0.467	0.371	0.157	0.028	-0.263	1.572	-0.252
HybridHD4	0.832	0.776	0.876	0.643	0.726	0.577	0.421	-0.051	0.449	-0.235	1.216

HybridHD5	0.837	0.782	0.879	0.495	0.559	0.444	0.264	0.038	-0.34	1.752	-0.318
HD merge	0.881	0.864	0.896	0.511	0.539	0.487	0.285	-0.008	0.089	0.554	0.772
Model PEVQ-S (pes+rtp)											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.774	0.664	0.851	0.648	0.775	0.556	0.369	0.057	-0.303	0.77	1.742
HybridVGA2	0.877	0.835	0.908	0.511	0.575	0.46	0.278	0.012	0.006	0.448	1.14
HybridVGA3	0.793	0.722	0.847	0.519	0.589	0.464	0.291	0.036	-0.184	0.736	1.357
HybridWVGA1	0.858	0.811	0.894	0.541	0.609	0.487	0.326	0.064	-0.468	1.84	-0.109
HybridWVGA2	0.837	0.762	0.891	0.534	0.629	0.464	0.325	-0.018	0.513	-1.888	3.414
HybridWVGA2 no rebuf	0.894	0.818	0.94	0.508	0.642	0.421	0.286	-0.065	0.96	-3.133	4.459
VGA merge	0.805	0.775	0.83	0.565	0.597	0.535	0.347	0.062	-0.414	1.353	0.807
VGA merge no rebuf	0.825	0.798	0.849	0.548	0.581	0.518	0.327	0.047	-0.282	1.043	1.017
HybridHD1	0.938	0.915	0.955	0.339	0.383	0.304	0.123	-0.036	0.387	-0.5	2.025
HybridHD2	0.883	0.837	0.916	0.505	0.577	0.449	0.28	0.045	-0.479	2.474	-1.127
HybridHD3	0.913	0.881	0.936	0.414	0.467	0.371	0.157	0.028	-0.263	1.572	-0.252
HybridHD4	0.832	0.776	0.876	0.643	0.726	0.577	0.421	-0.051	0.449	-0.235	1.216
HybridHD5	0.837	0.782	0.879	0.495	0.559	0.444	0.264	0.038	-0.34	1.752	-0.318
HD merge	0.881	0.864	0.896	0.511	0.539	0.487	0.285	-0.008	0.089	0.554	0.772
Model PEVQ-S (ts+rtp)											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.774	0.664	0.851	0.648	0.775	0.556	0.369	0.057	-0.303	0.77	1.742
HybridVGA2	0.877	0.835	0.908	0.511	0.575	0.46	0.278	0.012	0.006	0.448	1.14
HybridVGA3	0.793	0.722	0.847	0.519	0.589	0.464	0.291	0.036	-0.184	0.736	1.357
HybridWVGA1	0.858	0.811	0.894	0.541	0.609	0.487	0.326	0.064	-0.468	1.84	-0.109
HybridWVGA2	0.837	0.762	0.891	0.534	0.629	0.464	0.325	-0.018	0.513	-1.888	3.414
HybridWVGA2 no rebuf	0.894	0.818	0.94	0.508	0.642	0.421	0.286	-0.065	0.96	-3.133	4.459
VGA merge	0.805	0.775	0.83	0.565	0.597	0.535	0.347	0.062	-0.414	1.353	0.807
VGA merge no rebuf	0.825	0.798	0.849	0.548	0.581	0.518	0.327	0.047	-0.282	1.043	1.017
HybridHD1	0.938	0.915	0.955	0.339	0.383	0.304	0.123	-0.036	0.387	-0.5	2.025
HybridHD2	0.883	0.838	0.916	0.504	0.576	0.448	0.28	0.043	-0.465	2.441	-1.115
HybridHD3	0.914	0.883	0.937	0.412	0.465	0.37	0.155	0.026	-0.236	1.496	-0.183

HybridHD4	0.832	0.776	0.876	0.643	0.726	0.577	0.421	-0.051	0.449	-0.235	1.216
HybridHD5	0.837	0.782	0.879	0.495	0.559	0.444	0.264	0.038	-0.34	1.752	-0.318
HD merge	0.882	0.865	0.897	0.511	0.538	0.486	0.285	-0.01	0.099	0.527	0.792
Model YHyFR											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.632	0.472	0.751	0.793	0.949	0.681	0.485	0.013	-0.138	1.251	1.133
HybridVGA2	0.889	0.851	0.918	0.486	0.547	0.438	0.258	-0.271	2.493	-5.915	5.869
HybridVGA3	0.876	0.831	0.91	0.411	0.467	0.367	0.197	-0.088	0.839	-1.465	2.694
HybridWVGA1	0.883	0.844	0.913	0.495	0.556	0.445	0.27	-0.189	1.802	-4.068	4.305
HybridWVGA2	0.918	0.877	0.945	0.388	0.457	0.337	0.187	0.068	-0.358	1.552	-0.469
HybridWVGA2 no rebuf	0.963	0.934	0.979	0.308	0.389	0.255	0.104	-0.116	1.083	-1.933	2.254
VGA merge	0.799	0.769	0.825	0.572	0.605	0.542	0.352	-0.031	0.499	-0.94	2.483
VGA merge no rebuf	0.805	0.775	0.831	0.575	0.61	0.544	0.354	-0.042	0.568	-1.05	2.521
HybridHD1	0.91	0.878	0.934	0.407	0.459	0.365	0.189	0.003	0.177	-0.028	1.643
HybridHD2	0.79	0.715	0.848	0.659	0.752	0.586	0.404	0.205	-1.839	6.483	-5.478
HybridHD3	0.86	0.812	0.897	0.517	0.584	0.464	0.266	-0.107	1.166	-2.808	3.861
HybridHD4	0.873	0.829	0.907	0.565	0.638	0.507	0.346	-0.137	1.146	-1.716	2.078
HybridHD5	0.861	0.813	0.898	0.46	0.519	0.412	0.231	0.235	-2.042	6.717	-4.955
HD merge	0.86	0.841	0.878	0.551	0.581	0.525	0.323	-0.012	0.227	0.091	1.154
Model YHyFRe											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.643	0.487	0.759	0.784	0.938	0.673	0.479	0.165	-1.301	3.702	-0.088
HybridVGA2	0.887	0.848	0.916	0.491	0.553	0.442	0.264	-0.253	2.437	-6.133	6.431
HybridVGA3	0.858	0.806	0.896	0.438	0.497	0.392	0.217	-0.057	0.59	-0.867	2.286
HybridWVGA1	0.886	0.848	0.916	0.488	0.549	0.44	0.277	-0.2	1.877	-4.218	4.357
HybridWVGA2	0.906	0.86	0.938	0.413	0.487	0.359	0.216	0.166	-1.313	4.458	-3.211
HybridWVGA2 no rebuf	0.963	0.935	0.979	0.304	0.384	0.252	0.108	-0.035	0.353	0.106	0.489
VGA merge	0.79	0.759	0.817	0.583	0.617	0.553	0.364	0.008	0.208	-0.347	2.19
VGA merge no rebuf	0.796	0.764	0.823	0.587	0.622	0.556	0.367	0.002	0.24	-0.388	2.197
HybridHD1	0.921	0.892	0.942	0.382	0.431	0.342	0.15	-0.023	0.391	-0.553	1.983
HybridHD2	0.733	0.641	0.804	0.732	0.836	0.651	0.494	0.315	-2.957	10.093	-9.182

HybridHD3	0.777	0.704	0.833	0.638	0.721	0.573	0.378	0.182	-1.704	6.505	-6.098
HybridHD4	0.858	0.809	0.895	0.596	0.674	0.535	0.382	-0.115	1.064	-1.805	2.309
HybridHD5	0.906	0.872	0.931	0.384	0.433	0.344	0.168	0.206	-2.06	8.048	-8.089
HD merge	0.847	0.825	0.866	0.575	0.606	0.547	0.347	-0.039	0.488	-0.674	1.754
Model YHyRR56k											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.631	0.471	0.75	0.794	0.95	0.682	0.486	-0.005	0.003	0.925	1.363
HybridVGA2	0.889	0.851	0.917	0.488	0.548	0.439	0.261	-0.264	2.421	-5.665	5.595
HybridVGA3	0.875	0.829	0.909	0.412	0.468	0.369	0.197	-0.085	0.81	-1.4	2.66
HybridWVGA1	0.883	0.843	0.913	0.495	0.557	0.446	0.27	-0.183	1.75	-3.941	4.205
HybridWVGA2	0.917	0.876	0.945	0.389	0.459	0.338	0.187	0.062	-0.3	1.372	-0.288
HybridWVGA2 no rebuf	0.963	0.934	0.979	0.308	0.389	0.255	0.104	-0.123	1.148	-2.133	2.449
VGA merge	0.798	0.768	0.825	0.573	0.606	0.543	0.352	-0.029	0.479	-0.896	2.457
VGA merge no rebuf	0.804	0.774	0.831	0.576	0.611	0.545	0.354	-0.039	0.547	-1.002	2.493
HybridHD1	0.911	0.878	0.935	0.405	0.458	0.364	0.19	-0.004	0.241	-0.197	1.782
HybridHD2	0.783	0.706	0.842	0.668	0.763	0.595	0.412	0.225	-2.062	7.255	-6.327
HybridHD3	0.859	0.81	0.896	0.519	0.586	0.466	0.266	-0.101	1.112	-2.645	3.711
HybridHD4	0.875	0.831	0.908	0.562	0.635	0.505	0.345	-0.135	1.124	-1.651	2.019
HybridHD5	0.859	0.81	0.896	0.464	0.524	0.416	0.234	0.244	-2.115	6.905	-5.087
HD merge	0.859	0.839	0.877	0.554	0.584	0.527	0.324	-0.014	0.24	0.069	1.165
Model YHyRR56ke											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.641	0.484	0.758	0.785	0.94	0.675	0.48	0.16	-1.264	3.623	-0.042
HybridVGA2	0.885	0.847	0.915	0.494	0.556	0.445	0.268	-0.248	2.389	-5.971	6.259
HybridVGA3	0.856	0.805	0.895	0.44	0.499	0.393	0.219	-0.054	0.567	-0.824	2.275
HybridWVGA1	0.887	0.848	0.916	0.488	0.549	0.439	0.276	-0.195	1.834	-4.109	4.265
HybridWVGA2	0.905	0.859	0.937	0.415	0.489	0.361	0.217	0.162	-1.275	4.337	-3.084
HybridWVGA2 no rebuf	0.964	0.936	0.98	0.304	0.384	0.252	0.108	-0.043	0.429	-0.123	0.71
VGA merge	0.789	0.758	0.817	0.584	0.618	0.553	0.364	0.008	0.203	-0.339	2.193
VGA merge no rebuf	0.795	0.764	0.823	0.588	0.623	0.556	0.367	0.003	0.235	-0.381	2.2
HybridHD1	0.922	0.894	0.943	0.38	0.429	0.341	0.15	-0.027	0.42	-0.609	2.008

HybridHD2	0.724	0.629	0.797	0.742	0.847	0.66	0.502	0.304	-2.879	9.941	-9.109
HybridHD3	0.778	0.705	0.834	0.637	0.719	0.571	0.378	0.186	-1.752	6.705	-6.357
HybridHD4	0.858	0.809	0.895	0.595	0.672	0.534	0.382	-0.114	1.059	-1.793	2.3
HybridHD5	0.905	0.871	0.93	0.385	0.435	0.345	0.17	0.231	-2.292	8.744	-8.754
HD merge	0.846	0.824	0.865	0.577	0.608	0.549	0.349	-0.043	0.519	-0.735	1.786
Model YHyRR128k											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.633	0.474	0.752	0.792	0.947	0.68	0.484	0.001	-0.047	1.055	1.261
HybridVGA2	0.888	0.85	0.917	0.488	0.549	0.439	0.261	-0.264	2.424	-5.687	5.631
HybridVGA3	0.876	0.831	0.91	0.411	0.466	0.367	0.196	-0.085	0.813	-1.409	2.667
HybridWVGA1	0.883	0.844	0.913	0.495	0.556	0.445	0.27	-0.185	1.77	-3.986	4.237
HybridWVGA2	0.919	0.878	0.946	0.386	0.455	0.336	0.185	0.061	-0.292	1.354	-0.281
HybridWVGA2 no rebuf	0.964	0.936	0.98	0.304	0.384	0.251	0.102	-0.127	1.188	-2.262	2.575
VGA merge	0.799	0.769	0.825	0.572	0.605	0.542	0.352	-0.029	0.484	-0.907	2.462
VGA merge no rebuf	0.805	0.775	0.831	0.575	0.61	0.544	0.354	-0.04	0.552	-1.013	2.498
HybridHD1	0.911	0.879	0.935	0.405	0.458	0.363	0.189	-0.006	0.255	-0.227	1.801
HybridHD2	0.788	0.712	0.846	0.661	0.755	0.588	0.406	0.22	-2.001	7.032	-6.072
HybridHD3	0.86	0.811	0.896	0.518	0.585	0.464	0.266	-0.098	1.082	-2.553	3.619
HybridHD4	0.874	0.83	0.907	0.563	0.636	0.505	0.344	-0.143	1.194	-1.841	2.178
HybridHD5	0.859	0.811	0.896	0.463	0.523	0.415	0.232	0.234	-2.024	6.642	-4.854
HD merge	0.86	0.84	0.878	0.552	0.581	0.525	0.323	-0.014	0.248	0.04	1.193
Model YHyRR128ke											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.643	0.487	0.759	0.783	0.937	0.673	0.479	0.162	-1.288	3.685	-0.087
HybridVGA2	0.885	0.846	0.915	0.494	0.556	0.445	0.268	-0.248	2.394	-5.998	6.299
HybridVGA3	0.857	0.806	0.896	0.439	0.498	0.392	0.217	-0.054	0.57	-0.83	2.278
HybridWVGA1	0.886	0.848	0.916	0.488	0.549	0.44	0.277	-0.197	1.847	-4.14	4.291
HybridWVGA2	0.907	0.861	0.938	0.412	0.485	0.358	0.215	0.159	-1.245	4.25	-3.008
HybridWVGA2 no rebuf	0.964	0.937	0.98	0.3	0.379	0.248	0.107	-0.048	0.486	-0.306	0.891
VGA merge	0.79	0.759	0.817	0.583	0.617	0.553	0.364	0.008	0.205	-0.342	2.193
VGA merge no rebuf	0.796	0.764	0.823	0.587	0.623	0.556	0.367	0.003	0.237	-0.384	2.2

HybridHD1	0.922	0.893	0.943	0.38	0.43	0.341	0.15	-0.027	0.417	-0.602	2.004
HybridHD2	0.731	0.639	0.803	0.733	0.838	0.652	0.495	0.327	-3.08	10.519	-9.651
HybridHD3	0.779	0.707	0.835	0.636	0.718	0.57	0.378	0.171	-1.607	6.247	-5.89
HybridHD4	0.857	0.808	0.895	0.597	0.674	0.536	0.383	-0.112	1.046	-1.768	2.287
HybridHD5	0.905	0.871	0.93	0.385	0.435	0.346	0.169	0.218	-2.174	8.39	-8.412
HD merge	0.847	0.826	0.866	0.575	0.606	0.547	0.347	-0.042	0.508	-0.717	1.781
Model YHyRR256k											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridHD1	0.91	0.878	0.934	0.406	0.458	0.364	0.19	-0.005	0.244	-0.197	1.777
HybridHD2	0.789	0.713	0.847	0.661	0.754	0.588	0.406	0.215	-1.949	6.849	-5.866
HybridHD3	0.86	0.811	0.896	0.518	0.585	0.465	0.266	-0.101	1.107	-2.625	3.686
HybridHD4	0.875	0.832	0.908	0.561	0.633	0.503	0.342	-0.14	1.168	-1.757	2.1
HybridHD5	0.861	0.812	0.897	0.461	0.521	0.414	0.231	0.23	-1.991	6.553	-4.778
HD merge	0.861	0.841	0.878	0.551	0.581	0.525	0.322	-0.014	0.244	0.053	1.181
Model YHyRR256ke											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridHD1	0.921	0.893	0.942	0.382	0.431	0.342	0.151	-0.025	0.405	-0.578	1.992
HybridHD2	0.732	0.639	0.803	0.733	0.837	0.652	0.494	0.323	-3.044	10.387	-9.495
HybridHD3	0.778	0.706	0.835	0.636	0.718	0.571	0.377	0.172	-1.619	6.279	-5.914
HybridHD4	0.859	0.81	0.896	0.594	0.671	0.533	0.38	-0.115	1.062	-1.798	2.302
HybridHD5	0.906	0.873	0.931	0.383	0.432	0.343	0.168	0.211	-2.112	8.213	-8.249
HD merge	0.847	0.826	0.867	0.575	0.605	0.547	0.347	-0.041	0.501	-0.698	1.764
Model PSNR											
Dataset	ρ	Lower	Upper	RMSE	Lower	Upper	RMSE*	C3*x^3	C2*x^2	C1*x	C0
HybridVGA1	0.765	0.651	0.845	0.659	0.788	0.566	0.407	0	0.001	0.06	0.879
HybridVGA2	0.805	0.742	0.853	0.631	0.71	0.568	0.392	0	0.014	-0.318	3.501
HybridVGA3	0.753	0.671	0.817	0.561	0.636	0.501	0.308	0	0.012	-0.267	3.955
HybridWVGA1	0.83	0.774	0.872	0.589	0.662	0.53	0.35	0	0.025	-0.62	5.715
HybridWVGA2	0.788	0.693	0.856	0.601	0.708	0.522	0.398	0	0	0.205	-4.142
HybridWVGA2 no rebuf	0.857	0.757	0.918	0.585	0.74	0.485	0.368	-0.001	0.083	-2.365	21.931

VGA merge	0.721	0.682	0.757	0.659	0.697	0.624	0.414	0	0.002	-0.02	1.624
VGA merge no rebuf	0.732	0.692	0.767	0.661	0.701	0.625	0.414	0	0.002	-0.021	1.632
HybridHD1	0.906	0.872	0.931	0.416	0.47	0.373	0.187	0	0.011	-0.265	3.927
HybridHD2	0.835	0.774	0.881	0.591	0.675	0.526	0.359	0	0.001	0.096	-0.623
HybridHD3	0.804	0.739	0.854	0.602	0.68	0.54	0.328	0	0.008	-0.121	1.859
HybridHD4	0.855	0.805	0.893	0.601	0.679	0.539	0.371	0	0.021	-0.476	4.669
HybridHD5	0.603	0.49	0.696	0.722	0.815	0.648	0.473	0	-0.019	0.519	-2.319
HD merge	0.828	0.804	0.85	0.607	0.639	0.577	0.373	0	0.011	-0.214	2.896