

# Perceptual Video Quality Assessment: Spatiotemporal Pooling Strategies for Different Distortions and Visual Maps - Technical Report

A supplementary document for paper 103 in MMSP 2016

**Mohammed A. Aabed and Ghassan AlRegib**

Center for Signal & Information Processing (CSIP)  
School of Electrical and Computer Engineering  
Georgia Institute of Technology  
Atlanta, Georgia 30332-0250  
Email: {maabed,alregib}@gatech.edu

Copyright © 2016. Multimedia & Sensors Lab (MSL), School of Electrical and Computer Engineering, Georgia Institute of Technology (<http://www2.ece.gatech.edu/research/labs/msl/>). All Rights Reserved.

APPENDIX  
**Best Performing Pooling Strategies for the Tested Distortion Maps**

TABLE I

THE BEST POOLING STRATEGIES FOR FULL REFERENCE SQUARED ERROR DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, SROCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	(max,max,0.926,407)	(kurtosis,max,0.926,407)	(max,mean,0.920,407)
	Channel induced (H.264)	40	<b>(max,std,0.896,407)</b>	<b>(kurtosis,std,0.896,407)</b>	(max,max,0.882,407)
	Frame-freezes (stored content)	30	<b>(max,max,0.909,449)</b>	<b>(kurtosis,max,0.909,449)</b>	(max,std,0.889,449)
	Frame-freezes (live feed)*	10	(mean/std,mean/std,0.612,271)	(skewness,median,0.600,271)	(skewness,mean,0.576,271)
	Rate Adaptation	27	(max,std,0.888,407)	(kurtosis,std,0.888,407)	(max,max,0.868,407)
	Temporal Dynamics*	50	(std,norm2,0.828,211)	(std,mean,0.821,211)	(std,norm1,0.821,211)
CSIQ	Compression (H.264)	36	(max,skewness,0.910,380)	(kurtosis,skewness,0.910,380)	(max,kurtosis,0.884,380)
	Compression (HEVC)	36	(max,mean/std,0.842,380)	(kurtosis,mean/std,0.842,380)	(std,median,0.833,380)
	Compression (MJPEG)	36	<b>(max,max,0.829,381)</b>	(kurtosis,max,0.829,381)	(max,median,0.825,381)
	Compression, wavelet based (SNOW)	36	(max,median,0.916,380)	(kurtosis,median,0.916,380)	(max,mean,0.915,380)
	Channel induced (H.264)	36	<b>(mean/std,skewness,0.928,380)</b>	(mean/std,kurtosis,0.925,380)	(mean/std,std,0.904,380)
	Additive white noise (AWGN)	36	(mean,mean,0.903,380)	(norm1,mean,0.903,380)	(norm2,median,0.901,380)
IVP	Compression (H.264)	40	<b>(max,mean,0.866,228)</b>	<b>(kurtosis,mean,0.866,228)</b>	<b>(max,norm2,0.866,228)</b>
	Compression (MPEG2)	30	(mean/std,norm1,0.710,229)	(mean/std,median,0.708,229)	(mean/std,norm2,0.700,229)
	Compression (Dirac)	30	(max,norm1,0.764,229)	(kurtosis,norm1,0.764,229)	(max,mean,0.762,229)
	Channel induced (H.264)	28	(mean/std,mean,0.672,219)	(mean/std,norm1,0.665,219)	(max,mean,0.659,219)
All	Channel induced (H.264)	104	(max,mean,0.721,347)	(kurtosis,mean,0.721,347)	(max,median,0.704,347)
	Compression (H.264)	116	(max,norm2,0.833,337)	(kurtosis,norm2,0.833,337)	(max,mean,0.832,337)

TABLE II

THE BEST POOLING STRATEGIES FOR FULL REFERENCE SQUARED ERROR DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, PLOCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	(max,max,0.884,407)	(kurtosis,max,0.884,407)	(max,mean,0.838,407)
	Channel induced (H.264)	40	<b>(max,std,0.869,407)</b>	<b>(kurtosis,std,0.869,407)</b>	(max,max,0.845,407)
	Frame-freezes (stored content)	30	(max,max,0.871,449)	(kurtosis,max,0.871,449)	(mean,kurtosis,0.862,449)
	Frame-freezes (live feed)*	10	<b>(skewness,kurtosis,0.620,271)</b>	(norm1,std,0.590,271)	(mean,std,0.590,271)
	Rate Adaptation	27	<b>(std,mean/std,0.884,449)</b>	(norm2,mean/std,0.881,449)	(max,mean/std,0.877,407)
	Temporal Dynamics*	50	(max,mean,0.821,211)	(kurtosis,mean,0.821,211)	(max,norm1,0.821,211)
CSIQ	Compression (H.264)	36	<b>(max,skewness,0.863,380)</b>	<b>(kurtosis,skewness,0.863,380)</b>	(std,mean,0.804,380)
	Compression (HEVC)	36	(max,mean/std,0.795,380)	(kurtosis,mean/std,0.795,380)	(std,median,0.767,380)
	Compression (MJPEG)	36	<b>(max,median,0.805,381)</b>	<b>(kurtosis,median,0.805,381)</b>	(max,max,0.804,381)
	Compression, wavelet based (SNOW)	36	(max,std,0.871,380)	(kurtosis,std,0.871,380)	(max,max,0.870,380)
	Channel induced (H.264)	36	<b>(mean/std,std,0.894,380)</b>	<b>(mean/std,skewness,0.894,380)</b>	(max,mean/std,0.820,380)
	Additive white noise (AWGN)	36	<b>(norm1,mean,0.843,380)</b>	<b>(mean,mean,0.843,380)</b>	(norm2,mean,0.837,380)
IVP	Compression (H.264)	40	(max,norm1,0.796,228)	(kurtosis,norm1,0.796,228)	(max,norm2,0.793,228)
	Compression (MPEG2)	30	(max,mean,0.728,229)	(kurtosis,mean,0.728,229)	(max,norm1,0.726,229)
	Compression (Dirac)	30	(max,norm1,0.743,229)	(kurtosis,norm1,0.743,229)	(max,mean,0.738,229)
	Channel induced (H.264)	28	(max,mean,0.670,219)	(kurtosis,mean,0.670,219)	(mean/std,mean,0.663,219)
All	Channel induced (H.264)	104	<b>(max,mean,0.699,347,104)</b>	<b>(kurtosis,mean,0.699,347,104)</b>	(max,median,0.663,347,104)
	Compression (H.264)	116	(std,mean,0.749,337)	(max,norm2,0.742,337)	(kurtosis,norm2,0.742,337)

TABLE III  
THE BEST POOLING STRATEGIES FOR FULL REFERENCE SSIM DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, SROCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	(std,max,0.836,407)	(std,std,0.835,407)	(std,norm2,0.828,407)
	Channel induced (H.264)	40	(std,std,0.854,407)	(std,mean,0.849,407)	(std,norm1,0.849,407)
	Frame-freezes (stored content)	30	(mean/std,mean/std,0.907,449)	(max,max,0.906,449)	(kurtosis,max,0.906,449)
	Frame-freezes (live feed)*	10	(mean/std,kurtosis,0.552,271)	(skewness,norm2,0.442,271)	(skewness,mean,0.430,271)
	Rate Adaptation	27	(std,std,0.812,407)	(std,max,0.795,407)	(std,norm2,0.794,407)
	Temporal Dynamics*	50	<b>(std,mean,0.840,211)</b>	<b>(std,norm1,0.840,211)</b>	(std,norm2,0.837,211)
CSIQ	Compression (H.264)	36	<b>(mean,median,0.939,380)</b>	<b>(norm1,median,0.939,380)</b>	<b>(norm2,median,0.939,380)</b>
	Compression (HEVC)	36	<b>(mean/std,norm2,0.937,380)</b>	(mean/std,max,0.922,380)	(mean,median,0.916,380)
	Compression (MJPEG)	36	(skewness,skewness,0.200,381)	(max,mean/std,0.190,381)	(kurtosis,mean/std,0.190,381)
	Compression, wavelet based (SNOW)	36	<b>(mean/std,norm2,0.924,380)</b>	(mean/std,max,0.922,380)	(mean/std,mean,0.917,380)
	Channel induced (H.264)	36	(mean/std,kurtosis,0.913,380)	(skewness,std,0.912,380)	(mean/std,skewness,0.912,380)
	Additive white noise (AWGN)	36	<b>(max,max,0.906,380)</b>	<b>(kurtosis,max,0.906,380)</b>	(max,median,0.903,380)
IVP	Compression (H.264)	40	(std,median,0.852,228)	(std,norm1,0.845,228)	(std,mean,0.839,228)
	Compression (MPEG2)	30	(mean/std,max,0.843,229)	(mean/std,std,0.812,229)	(mean/std,norm2,0.800,229)
	Compression (Dirac)	30	(mean/std,mean,0.761,229)	(mean,median,0.759,229)	(norm1,median,0.759,229)
	Channel induced (H.264)	28	(mean/std,median,0.660,219)	(mean,median,0.622,219)	(norm1,median,0.622,219)
All	Channel induced (H.264)	104	(std,median,0.665,347)	(mean/std,median,0.624,347)	(std,norm1,0.614,347)
	Compression (H.264)	116	<b>(std,mean,0.862,337)</b>	(std,norm2,0.857,337)	(std,median,0.850,337)

TABLE IV  
THE BEST POOLING STRATEGIES FOR FULL REFERENCE SSIM DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, PLOCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	(std,max,0.831,407)	(std,norm2,0.807,407)	(std,mean,0.807,407)
	Channel induced (H.264)	40	(std,norm2,0.813,407)	(std,mean,0.813,407)	(std,norm1,0.813,407)
	Frame-freezes (stored content)	30	<b>(std,kurtosis,0.898,449)</b>	(mean/std,mean/std,0.865,449)	(norm1,median,0.855,449)
	Frame-freezes (live feed)*	10	(max,skewness,0.492,271)	(kurtosis,skewness,0.492,271)	(max,mean/std,0.491,271)
	Rate Adaptation	27	(mean/std,mean/std,0.816,449)	(std,max,0.776,407)	(std,std,0.774,407)
	Temporal Dynamics*	50	<b>(std,norm2,0.840,211)</b>	(std,norm1,0.838,211)	(std,mean,0.838,211)
CSIQ	Compression (H.264)	36	(mean/std,norm1,0.806,380)	(std,norm2,0.766,380)	(mean/std,median,0.760,380)
	Compression (HEVC)	36	<b>(mean/std,max,0.922,380)</b>	(mean/std,norm2,0.895,380)	(mean/std,std,0.885,380)
	Compression (MJPEG)	36	(max,median,0.231,381)	(kurtosis,median,0.231,381)	(max,max,0.190,381)
	Compression, wavelet based (SNOW)	36	<b>(mean/std,max,0.893,380)</b>	(std,median,0.889,380)	(std,max,0.888,380)
	Channel induced (H.264)	36	(skewness,std,0.874,380)	(mean/std,mean,0.865,380)	(mean/std,mean/std,0.861,380)
	Additive white noise (AWGN)	36	(max,median,0.813,380)	(kurtosis,median,0.813,380)	(max,max,0.774,380)
IVP	Compression (H.264)	40	(std,norm1,0.829,228)	(std,mean,0.819,228)	(std,norm2,0.814,228)
	Compression (MPEG2)	30	(std,norm2,0.787,229)	(std,norm1,0.783,229)	(std,mean,0.775,229)
	Compression (Dirac)	30	(mean/std,mean,0.798,229)	(mean/std,median,0.793,229)	(mean/std,norm2,0.790,229)
	Channel induced (H.264)	28	(mean/std,median,0.646,219)	(std,norm2,0.439,219)	(std,norm1,0.430,219)
All	Channel induced (H.264)	104	(mean/std,median,0.425,347)	(std,norm1,0.397,347)	(std,norm2,0.392,347)
	Compression (H.264)	116	(std,mean,0.729,337)	(std,norm2,0.723,337)	(std,norm1,0.711,337)

TABLE V

THE BEST POOLING STRATEGIES FOR FULL REFERENCE OPTICAL FLOW DIFFERENCE DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, SROCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	<b>(max,max,0.933,449)</b>	<b>(kurtosis,max,0.933,449)</b>	(max,norm2,0.922,449)
	Channel induced (H.264)	40	(max,norm2,0.863,449)	(kurtosis,norm2,0.863,449)	(max,mean,0.862,449)
	Frame-freezes (stored content)	30	(max,max,0.882,449)	(kurtosis,max,0.882,449)	(max,kurtosis,0.882,449)
	Frame-freezes (live feed)*	10	<b>(norm2,kurtosis,0.842,271)</b>	(std,kurtosis,0.782,271)	(norm2,skewness,0.745,271)
	Rate Adaptation	27	<b>(max,kurtosis,0.890,449)</b>	<b>(kurtosis,kurtosis,0.890,449)</b>	(max,std,0.869,449)
CSIQ	Temporal Dynamics*	50	(max,norm2,0.771,211)	(kurtosis,norm2,0.771,211)	(max,mean,0.768,211)
	Compression (H.264)	36	(max,median,0.902,381)	(kurtosis,median,0.902,381)	(max,mean,0.886,381)
	Compression (HEVC)	36	(max,std,0.865,381)	(kurtosis,std,0.865,381)	(max,max,0.855,381)
	Compression (MJPEG)	36	(max,skewness,0.579,381)	(kurtosis,skewness,0.579,381)	(std,mean,0.477,381)
	Compression, wavelet based (SNOW)	36	(max,std,0.832,381)	(kurtosis,std,0.832,381)	(max,max,0.757,381)
	Channel induced (H.264)	36	(mean/std,std,0.894,381)	(mean/std,mean/std,0.883,381)	(max,mean/std,0.869,381)
IVP	Additive white noise (AWGN)	36	(mean,median,0.694,381)	(norm1,median,0.694,381)	(mean,mean,0.676,381)
	Compression (H.264)	40	(max,mean,0.847,253)	(kurtosis,mean,0.847,253)	(max,norm1,0.846,253)
	Compression (MPEG2)	30	<b>(max,median,0.879,253)</b>	<b>(kurtosis,median,0.879,253)</b>	(max,mean,0.869,253)
	Compression (Dirac)	30	<b>(max,median,0.894,253)</b>	<b>(kurtosis,median,0.894,253)</b>	(max,mean,0.887,253)
All	Channel induced (H.264)	104	<b>(max,mean,0.773,373)</b>	<b>(kurtosis,mean,0.773,373)</b>	(max,median,0.768,373)
	Compression (H.264)	116	(max,norm2,0.797,360)	(kurtosis,norm2,0.797,360)	(std,mean,0.794,360)

TABLE VI

THE BEST POOLING STRATEGIES FOR FULL REFERENCE OPTICAL FLOW DIFFERENCE DISTORTION MAPS (SPATIAL POOLING, TEMPORAL POOLING, PLOCC, AVERAGE FRAMES PER SEQUENCE).

Database	Distortion Type	Total Videos	Best	2nd	3rd
LIVE	Compression (H.264)	40	<b>(max,max,0.884,449)</b>	<b>(kurtosis,max,0.884,449)</b>	(max,norm2,0.877,449)
	Channel induced (H.264)	40	(max,norm2,0.840,449)	(kurtosis,norm2,0.840,449)	(max,mean,0.839,449)
	Frame-freezes (stored content)	30	(max,kurtosis,0.882,449)	(kurtosis,kurtosis,0.882,449)	(std,kurtosis,0.869,449)
	Frame-freezes (live feed)*	10	(mean,max,0.470,271)	(norm1,max,0.470,271)	(std,kurtosis,0.463,271)
	Rate Adaptation	27	(max,kurtosis,0.862,449)	(kurtosis,kurtosis,0.862,449)	(norm2,kurtosis,0.828,449)
	Temporal Dynamics*	50	(max,mean,0.790,211)	(kurtosis,mean,0.790,211)	(max,norm1,0.790,211)
CSIQ	Compression (H.264)	36	(max,median,0.859,381)	(kurtosis,median,0.859,381)	(max,mean,0.854,381)
	Compression (HEVC)	36	(max,std,0.836,381)	(kurtosis,std,0.836,381)	(max,max,0.807,381)
	Compression (MJPEG)	36	(max,skewness,0.562,381)	(kurtosis,skewness,0.562,381)	(std,median,0.451,381)
	Compression, wavelet based (SNOW)	36	(max,std,0.833,381)	(kurtosis,std,0.833,381)	(max,max,0.783,381)
	Channel induced (H.264)	36	(mean/std,std,0.873,381)	(skewness,std,0.865,381)	(mean/std,mean/std,0.849,381)
IVP	Additive white noise (AWGN)	36	(norm1,median,0.720,381)	(mean,median,0.720,381)	(norm1,mean,0.694,381)
	Compression (H.264)	40	<b>(max,norm1,0.855,253)</b>	<b>(kurtosis,norm1,0.855,253)</b>	(max,norm2,0.843,253)
	Compression (MPEG2)	30	<b>(max,median,0.865,253)</b>	<b>(kurtosis,median,0.865,253)</b>	<b>(max,mean,0.865,253)</b>
	Compression (Dirac)	30	<b>(max,norm1,0.866,253)</b>	<b>(kurtosis,norm1,0.866,253)</b>	(max,median,0.864,253)
All	Channel induced (H.264)	104	(max,norm2,0.670,373)	(kurtosis,norm2,0.670,373)	(max,mean,0.665,373)
	Compression (H.264)	116	<b>(max,norm2,0.787,360)</b>	<b>(kurtosis,norm2,0.787,360)</b>	(std,mean,0.777,360)