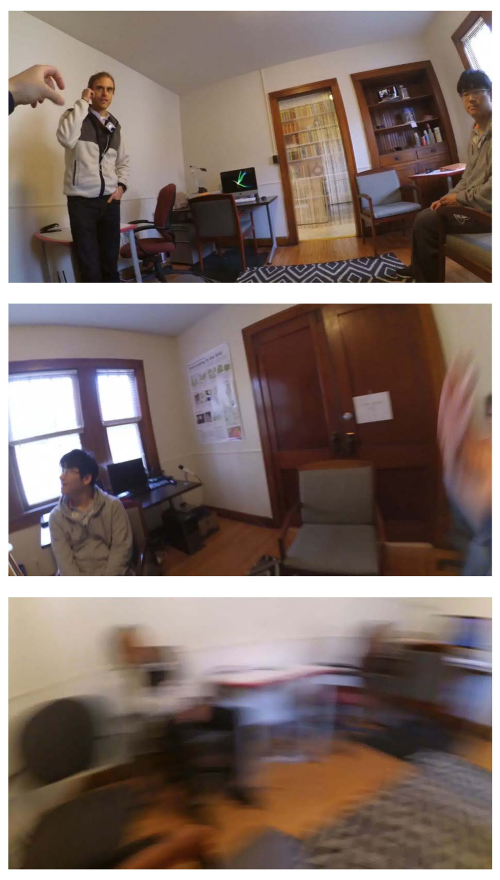
Real-time quality assessment of videos from body-worn cameras

centre for intelligent sensing
Queen Mary University of London

Yuan-Yi Chang, Riccardo Mazzon, Andrea Cavallaro
{y.chang}@se16.qmul.ac.uk, {r.mazzon,a.cavallaro}@qmul.ac.uk

Introduction

- Body-worn camera videos
- Aim: to provide a frame-by-frame quality score of a video → Video quality assessment (**VQA**)
- Challenges
  - scene conditions change abruptly
    - continuous changing of quality
    - score to be calculated quickly
  - uncontrolled scenarios with multiple simultaneous distortions
- Related work
  - full [1], reduced [2] and mutual [3] reference
  - no reference: distortion (e.g. blur) specific [4] and non-distortion specific [5]



Proposed approach: M-BRISQUE

- No-reference and non-distortion specific VQA method with a real-time implementation
- Michelson Contrast (**MC**) to account for distortions of the whole frame - global cue

image max intensity value

$$C_m = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

image min intensity value

- Blind/Referenceless Image Spatial QUality Evaluator (**BRISQUE** [5]) to account for patch-based distortions - local cues
  - Mean Subtracted Contrast Normalised (MSCN) coefficients $\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j)+1}$

    mean of local patch weighted by a Gaussian

    pixel location

    standard deviation of local patch weighted by a Gaussian
    - descriptors of luminance for image patches
    - vary coherently in the presence of a distortion
  - distortion estimation [6] by fitting the histogram of $\hat{I}(i,j)$ with
    - Generalised Gaussian Distribution (GGD) defined by mean (α) and variance (σ_g^2)
    - Asymmetric Generalised Gaussian Distribution (AGGD) defined by mean (η), shape (υ) and variances (σ_l^2, σ_r^2)
- Feature vector to describe the frames

$$\begin{pmatrix} \boxed{C_m}, & (\alpha,\sigma_g^2), & (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_H, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_V, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_{D_1}, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_{D_2}, \\ & (\alpha,\sigma_g^2)^*, & (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_H^*, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_V^*, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_{D_1}^*, (\eta,\upsilon,\sigma_l^2,\sigma_r^2)_{D_2}^* \end{pmatrix}$$

image at half resolution

horizontal orientation, vertical orientation, two diagonal orientations

- Operates in the spatial domain (no Gabor filters, Wavelets or DCT) → computationally efficient

Experimental results

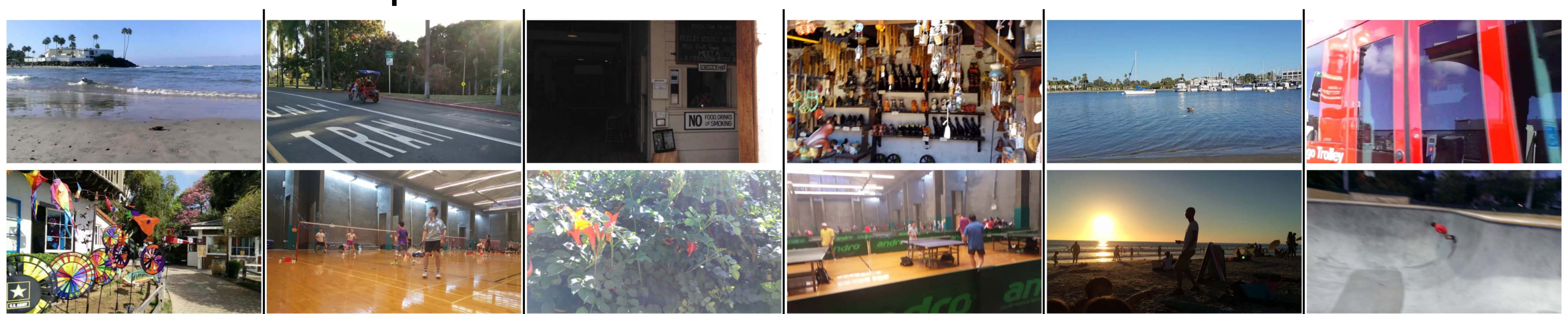
- Spearman's Rank Ordered Correlation Coefficient (**SROCC**) to correlate human judgement ↔ image score

Training

- M-BRISQUE score
  - Support Vector Regression (SVR)
  - Radial Basis Function (RBF) kernel
- Computational and Subjective Image Quality (CSIQ) database [7]
  - 30 original images
  - distortions for each image:
    - JPEG and JPEG2000 compressions
    - global contrast decrements
    - additive pink Gaussian noise
    - additive white Gaussian noise
    - Gaussian blurring

Testing

- LIVE Mobile In-Capture Video Quality Database [8]
- 208 videos captured with 8 hand-held devices

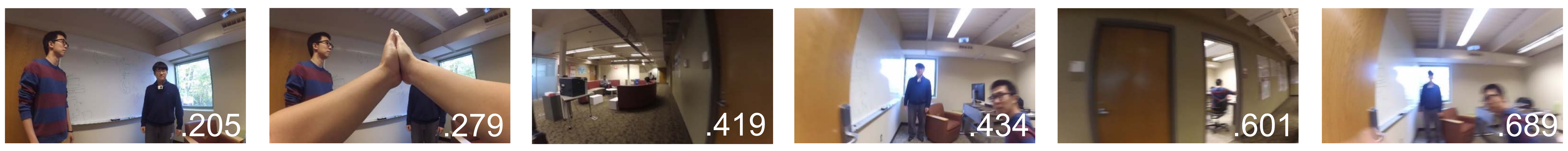


1st best
2nd best

Method \ Distortion	Artifacts	Colour	Exposure	Focus	Sharpness	Stabilisation
RMS	.007	.171	.026	.147	.007	.134
MC	.400	.130	.473	.080	**.580**	.280
BRISQUE	**.601**	.328	.492	.301	.451	**.513**
M-BRISQUE	.558	**.516**	**.601**	**.358**	.544	.508

Superior performance w.r.t. BRISQUE

M-BRISQUE scores (dataset [9])



References

[1] Zhang et al., "Edge strength similarity for image quality assessment," IEEE Signal Processing Letters, 2013
[2] Ma et al., "Reduced-reference image quality assessment using reorganized DCT-based image representation," IEEE TMM, 2011
[3] Bai and Reibman, "Mutual reference frame-quality assessment for first-person videos," IEEE ICIP, 2017
[4] Narvekar and Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," IEEE TIP, 2011
[5] Mittal et al., "No-reference image quality assessment in the spatial domain," IEEE TIP, 2012
[6] Ruderman, "The statistics of natural images," Network: Computation in Neural Systems, 1994
[7] Larson and Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy", Journal of Electronic Imaging, 2010
[8] Ghadiyaram et al., "Subjective and objective quality assessment of mobile videos with in-capture distortions", IEEE ICASSP, 2017
[9] Fan et al., "Identifying first-person camera wearers in third-person videos," IEEE CVPR, 2017

Watch the live demo

Project webpage: http://cis.eecs.qmul.ac.uk/projects/MBRISQUE

Additional info: http://www.eecs.qmul.ac.uk/~andrea