Subjective Assessment of Image Quality Induced Saliency Variation

Lucie Lévêque[1], Wei Zhang[2] and Hantao Liu[3]

[1]Xi'an Jiaotong Liverpool-University, China; [2]Xidian University, China; [3]Cardiff University, United Kingdom

Introduction

- **Visual saliency** allows the effective selection of the most relevant information in a visual scene. It has been widely studied in relation to image quality assessment, but fundamental challenges remain to fully simulate saliency in image quality assessment (IQA).
- In a previous research [1], we conducted an **eye-tracking** experiment to better understand how **distortions** affect the human visual system (HVS). Results showed changes in saliency between an original image and its distorted versions, as illustrated in Fig. 1.

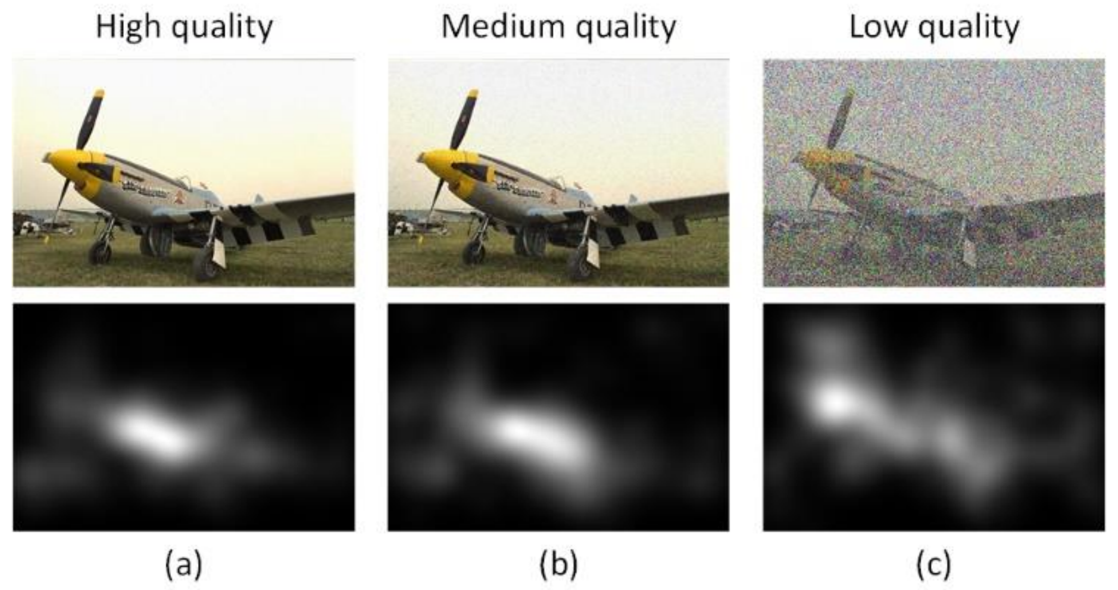

Fig. 1: Illustration of the distortion-induced saliency variation.

- A follow-up study [2] was then undertaken to investigate the relationship between **distortion-induced saliency variation** (DSV) and the perceived image quality, yet reliably measuring DSV is still an open-end question, which we approach here.

Measurement of DSV

- The **SIQ288** database [3] was used in this study. It consists of 288 images of varying quality (eighteen original images, five distortion types at three distorted levels) and their corresponding saliency maps, obtained from an eye-tracking study carried out with 160 observers. Fig. 2 illustrates an example of an original image, its saliency map, and the saliency maps of its distorted versions.
- To measure the DSV, participants were asked to score the **similarity** between the saliency maps of a distorted image and of its original scene using a scale from 0 (i.e., "Bad") to 100 (i.e., "Excellent"). Sixteen experts in **computer vision** from the Visual Computing Research Group of Cardiff University) participated in the experiment.

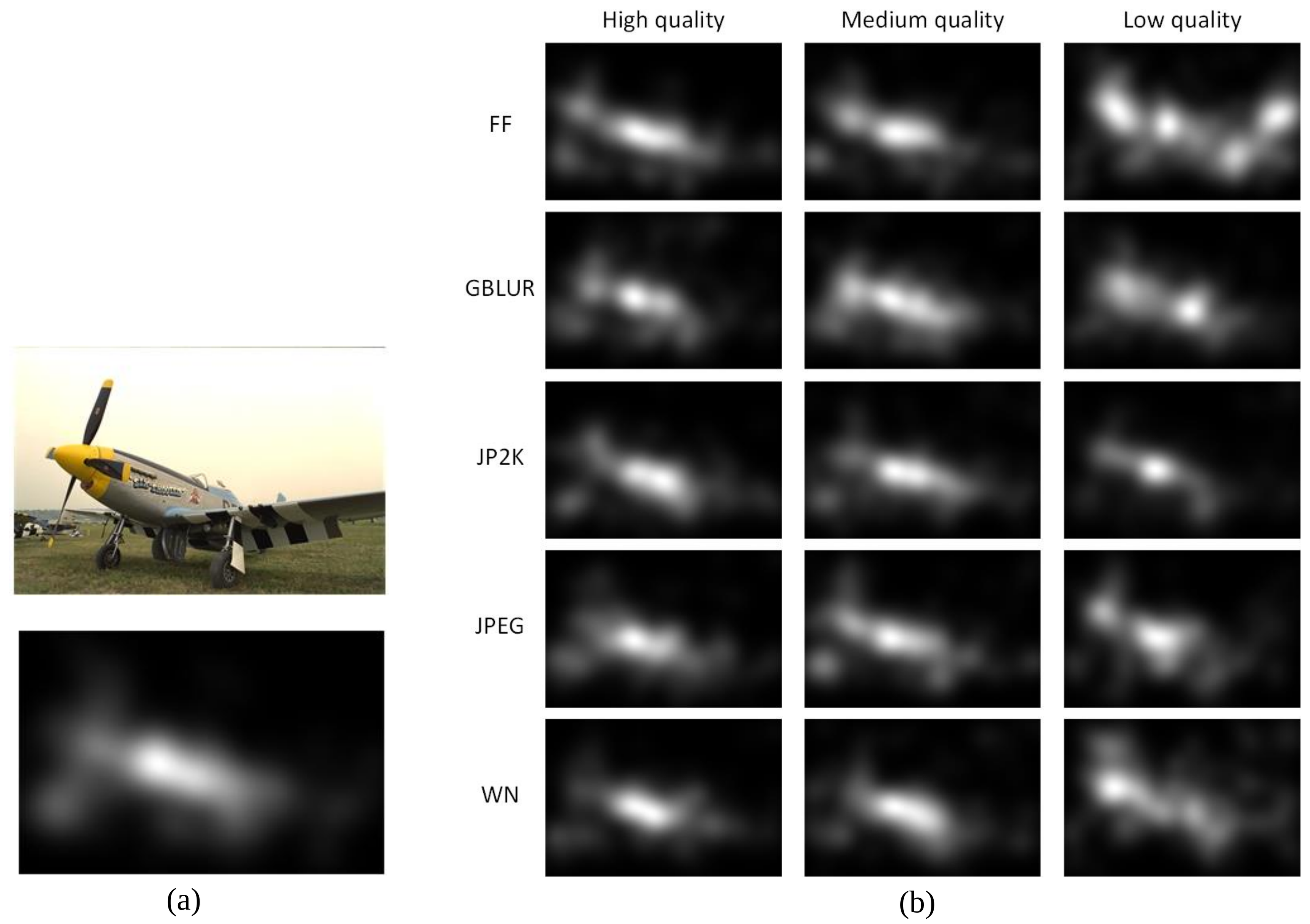

Fig. 2: (a) Illustration of an original image and its saliency map from SIQ288. (b) Illustration of the saliency maps of all distorted images (FF: fast-fading, GBLUR: Gaussian blur, JP2K: JPEG2000 compression, JPEG: JPEG compression, and WN: white noise) originated from the same scene.

Experimental results

- To process the raw data, the subjective scores were first transformed to **difference scores** to discount likely preferences for certain saliency patterns:

$$d_{ij} = s_{ij\_ref} - s_{ij}$$

Score given by subject *i* to reference saliency map *j* (s_{ij_ref}); Score given by subject *i* to test saliency map *j* (s_{ij})

- The difference scores were then converted into **z-scores**:

$$z_{ij} = (d_{ij} - \mu_i)/\sigma_i$$

Mean of all difference scores for subject *i* (μ_i); Standard deviation of all difference scores for subject *i* (σ_i)

- The **difference mean saliency variation score** (DMSS) of each stimulus was finally computed as the mean of the rescaled z-scores over all subjects:

$$DMSS_j = \frac{1}{S}\sum_{i=1}^{S} z'_{ij}$$

Number of subjects (S); Rescaled z-scores (linearly mapped to [0,100]) (z'_{ij})

- An ANOVA was performed to reveal the **behaviours and properties** of human assessment of the saliency variation. Results are presented in Table I.

Factor	df	F	Sig. (p)
Image scene	17	58.298	7.2E-179
Distortion type	4	7.078	0.00001
Distortion level	2	212.382	1.2E-88
Subject	15	0.0001	1
Distortion type * level	8	10.274	2.5E-14

Table I: Results of the ANOVA. z_{ij}' is selected as the dependent var., "image scene", "distortion type", and "distortion level" as fixed independent var., and "subject" as random independent var.

- Results show no statistically significant difference between subjects in scoring. The interaction between "distortion type" and "distortion level" is significant, which implies that **the difference in measurement between the three levels of distortion is not the same for the five distortion types**.
- Fig. 3 illustrates the impact of distortion type. There is a linear relationship between the saliency variation and distortion level for FF, JP2K, and JPEG.

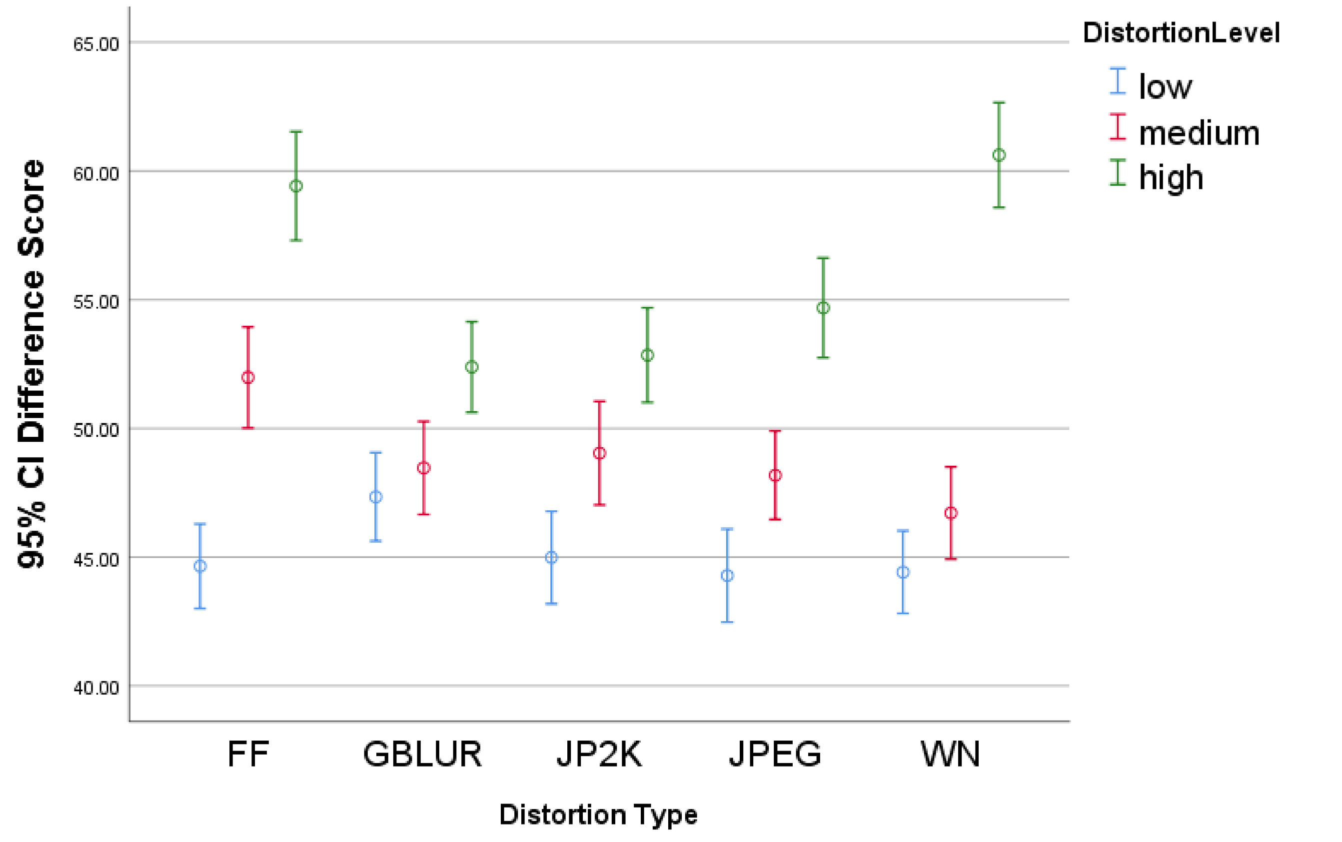

Fig. 3: Illustration of the difference saliency variation score averaged over relevant stimuli for the low, medium, and high distortion levels of each distortion type. Error bars indicate a 95% CI.

Bibliography

[1] W. Zhang, A. Borji, Z. Wang, P. Le Callet, and H. Liu, "The application of visual saliency models in objective image quality assessment: A statistical evaluation", *IEEE TNNLS*, vol. 27, 2016.

[2] W. Zhang, and H. Liu, "Toward a reliable collection of eye-tracking data for image quality research: Challenges, solutions, and applications", *IEEE TIP*, vol. 26, 2017.

[3] W. Zhang, and H. Liu, "SIQ288: A saliency dataset for image quality research", *IEEE MMSP*, 2016.