Subjective Assessment of Image Quality Induced Saliency Variation

Lucie Lévêque1, Wei Zhang2 and Hantao Liu3

1Xi’an Jiaotong Liverpool-University, China; 2Xidian University, China; 3Cardiff 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.](image)

- 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.

![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.](image)

- 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 (Jia, Le, Wu, Zhang) participated in the experiment.

![Fig. 3: Illustration of the difference saliency variation score averaged over relevant stimuli for the three levels of distortion for FF, JP2K, and JPEG.](image)

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

\[
Z_{ij} = \frac{(d_{ij} - \mu_j)}{\sigma_j}
\]

where \(d_{ij}\) is the difference score for subject \(i\) and image \(j\), \(\mu_j\) is the mean of all difference scores for subject \(i\), and \(\sigma_j\) is the standard deviation of all difference scores for image \(j\).

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

\[
DMSS_j = \frac{1}{\text{Number of subjects}} \sum_{i=1}^{\text{Number of subjects}} Z_{ij}
\]

Experimental results

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

\[
d_{ij} = S_{ij,ref} - S_{ij}
\]

where \(S_{ij}\) is the score given by subject \(i\) to the test saliency map \(j\) and \(S_{ij,ref}\) is the score given by subject \(i\) to the reference saliency map \(j\).

- The experimental results showed a significant difference between the distortion types and levels, as indicated by the ANOVA table (Table I).

![Table I: Results of the ANOVA.](image)

Bibliography

