Attention modeling for video quality assessment: balancing global quality and local quality

This paper proposes to evaluate video quality by balancing two quality components: global quality and local quality. The global quality is a result from subjects allocating their attention equally to all regions in a frame and all frames in a video. It is evaluated by image quality metrics (IQM) with averaged spatiotemporal pooling. The local quality is derived from visual attention modeling and quality variations over frames. Saliency, motion, and contrast information are taken into account in modeling visual attention, which is then integrated into IQMs to calculate the local quality of a video frame. The local quality of a video sequence is calculated by pooling local quality values over all frames with a temporal pooling scheme derived from the known relationship between perceived video quality and the frequency of temporal quality variations. The overall quality of a distorted video is a weighted average between the global quality and the local quality. Experimental results demonstrate that the combination of the global quality and local quality outperforms both sole global quality and local quality, as well as other quality models, in video quality assessment. In addition, the proposed video quality modeling algorithm can improve the performance of image quality metrics on video quality assessment compared to the normal averaged spatiotemporal pooling scheme.