Efficient Video Coding Optimization Using a Novel Perceptual Distortion Model

Like Yu[1,2], Feng Dai[1], Yongdong Zhang[1], Shouxun Lin[1]
[1] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2] Graduate University of Chinese Academy of Sciences, 100190 Beijing, China
{yulike, fdai, zhyd, sxlin}@ict.ac.cn

In this paper we propose a video coding optimization strategy for H.264/AVC encoder which uses HVS (Human Visual System) theory to improve the visual quality. It is based on a basic rule that sensitivity of distortion is different according to the image property. Most existing coding strategies which allocate fixed parameters to each unit within a frame are not reasonable because visual redundancy cannot be fully used. Instead, our scheme dynamically changes the coding strategy of different regions. It enhances the quality of regions which is more sensible and slightly decreases the quality of regions which is less sensible. The overall perceptual quality can be improved.

Our work is composed of two parts. Firstly, we establish a new perceptual distortion model which takes the HVS characteristics into consideration. The new model is based on the mask effect about the texture complexity [1] and the mask effect about background stimulus (Weber-Fechner law) [2]. It can be written as: $Perceptual\_Quaity = \varepsilon \cdot MSE$, where $\varepsilon = \alpha \cdot (K \cdot T)^{-2}$. Here K is a function which represents the texture complexity, T is the minimal perception threshold which is related to the background luminance stimulus. We call ε as "mask-effect parameter" and the equation means when the texture complexity and minimal perception threshold are low, the distortion will be enlarged, vice versa. The second part of our work is to optimize important modules in the hybrid video coding framework by using the new distortion model. Here we select the RDO (Rate-Distortion Optimization) module and the quantization parameter allocation module because they have a great impact on coding efficiency. For QP allocation module, an adaptive QP allocation strategy is proposed which modifies the traditional D-Q model by replace "D" (which is often measured in MSE) with the new distortion model so that a new quantization strategy can be achieved, which can be written as: $Q_i = \mu^{1/2} \cdot Q_{Ref}$, where $\mu = \varepsilon_i / \varepsilon_{AVG}$. Q_{Ref} is a reference value provided by rate-control algorithm. ε_i is the mask-effect parameter of the i^{th} macroblock, ε_{AVG} is the average value of all mask-effect parameters within a frame. For RDO module, since it is well known that the value of λ in R-D cost function is the negative slope of R-D curve, we calculated the slope of the R-D function which replaces "D" with the new distortion model. Thus a new value of λ is achieved, which can be written as: $\lambda_i = \mu \times 0.85 \times 2^{(QP-12)/3}$.

The experiment is based on H.264/AVC reference codec JM12.2 and six video sequences are chosen. We use PSNR and SSIM metric for performance evaluation. Our scheme gets lower PSNR than JM (-0.42% on average) because it is not designed for a maximum PSNR result. But in SSIM metric, our scheme gets higher scores than JM (4.73% on average), which means our scheme achieves perceptual quality improvement.

This work is supported by National Basic Research Program of China (973 Program, 2007CB311100), National Nature Science Foundation of China (60802028), Co-building Program of Beijing Municipal Education Commission.

[1] A. Bhat, I. Richardson, S. Kannangara, "A new perceptual quality metric for compressed video," *IEEE International Conference on Acoustics, Speech and Signal Processing,* pp. 933-936, April 2009.
[2] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology.* New York: Appleton-Century, 1942.

1068-0314/11 $26.00 © 2011 IEEE
DOI 10.1109/DCC.2011.90