Anti-Forensics of Lossy Predictive Image Compression (supplement)

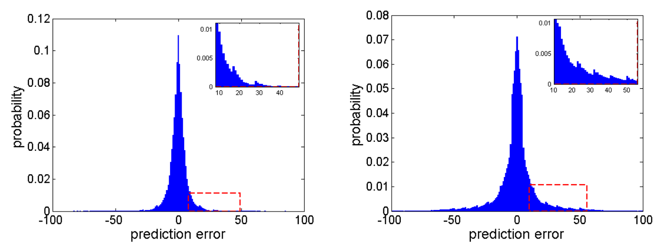

Fig. 1. Un-smooth boundary regions across adjacent quantization bins when universal $\hat{\lambda}$ is used. Left: Lena; Right: Barbara

I. Proposed anti-forensic framework

A. *Analysis of the universal Laplacian assumption*

Fig. 1 shows the results obtained by adopting the universal laplacian assumption, i.e., an unique λ. Obviously, the distribution across different quantization bins is un-smooth. This phenomenon is directly caused by the strict assumption that the distribution of the prediction error e well satisfies universal Laplacian distribution with the estimated parameter $\hat{\lambda}$. For clarity purpose, we call this assumption as universal Laplacian assumption (ULA). Denote N_k as the number of observations satisfying $\breve{e}_i = q_k$, and define function $F(\lambda, k)$ as

$$F(\lambda, k) = \left| \frac{N_k}{N} - \int_{l_k}^{r_k} \frac{\lambda}{2} \exp(-\lambda|e|) \mathrm{d}e \right| \quad (1)$$

Obviously, if the prediction error e satisfies ULA, then $F(\hat{\lambda}, k) \approx 0$ should hold for each k ($k \in [1, K]$) value. The theoretical basis is: under ULA, the empirical possibility that e falls into the kth quantization bin (equal to N_k/N) should be similar with the counterpart w.r.t the modeled distribution (equal to $\int_{l_k}^{r_k} \frac{\lambda}{2} \exp(-\lambda|e|) \mathrm{d}e$). A smaller $F(\hat{\lambda}, k)$ implies that the envelop beyond the kth quantization bin is more accurate being characterized as the corresponding part of Laplacian distribution with parameter $\hat{\lambda}$. While a bigger $F(\hat{\lambda}, k)$ implies that the envelop beyond the kth quantization bin is much deviating from that.

However, For some natural images, ULA is too strict, due to the truth that compared with a continuous, smooth Laplacian distribution, the distribution of prediction error sequence is always with some disturbances. So, For a given image, it is definitely possible that for some values of k, $F(\hat{\lambda}, k) >> 0$, which is in contradiction to ULA. When this happens, it often results in periodic, un-smooth regions across the boundary of adjacent quantization bins, when distributions of d over different quantization bins are forced to be modeled using the same parameter $\hat{\lambda}$. We show this phenomenon experimentally in Fig. 1, from which we could easily see that the reconstruction prediction errors are smoothly distributed in the same quantization bins, while discontinuously across the boundary of adjacent quantization bins.

B. *Dither Generation Based on Piecewise Laplacian Distribution*

In order to erase the periodic and un-smooth artifacts, we relax ULA to the assumption that the prediction error well satisfies a piecewise Laplacian distribution, we call this assumption as PLA. This motivates us to refine the estimated parameter $\hat{\lambda}$ for each quantization bin.

Lemma 1. *Suppose λ_k is the refined PLA parameter for the kth quantization bin, and denote I' as the reconstructed image undergone the proposed anti-forensic operation. Then the distribution of anti-forensically modified prediction error $\breve{e}'$ w.r.t. I' satisfies*

$$f_{\breve{E}'}(\lambda_k, \breve{e}') = \frac{\frac{\lambda_k}{2} exp(-\lambda_k|\breve{e}'|)}{\int_{l_k}^{r_k} \frac{\lambda_k}{2} exp(-\lambda_k|e|) \mathrm{d}e} \times \frac{N_k}{N}, \quad \breve{e}' \in [l_k, r_k] \quad (2)$$

The proof is omitted.

Denote k_0 the index of quantization bin over which any value maps to 0. Thanks to the dither generation method in (6-paper), where a smooth distribution in each quantization bin is intrinsically guaranteed. Under PLA, in order to smoothly connect the boundaries of adjacent quantization bins in the negative side, we refine the λ by using

$$\begin{aligned} \{\lambda_1, ..., \lambda_{k_0}^{neg}\} &= \arg\min \sum_{k=1}^{k=k_0} ||\lambda_k - \hat{\lambda}||_2^2 \\ s.t.\ \alpha_1 J_k^k &\le J_k^{k+1} \le \alpha_2 J_k^k,\ k = 1, ..., k_0 - 1 \end{aligned} \quad (3)$$

where

- $\lambda_{k_0}^{neg}$ represents the refined λ w.r.t. k_0th quantization bin in the negative side;
- J_k^k and J_k^{k+1} are defined by
$$\begin{aligned} J_k^k &= f_{\breve{E}'}(\lambda_k, r_k) - f_{\breve{E}'}(\lambda_k, r_k - 1) \\ J_k^{k+1} &= f_{\breve{E}'}(\lambda_{k+1}, l_{k+1}) - f_{\breve{E}'}(\lambda_k, r_k) \end{aligned} \quad (4)$$
- α_1, α_2 are two scale parameters. Due the truth that the Laplacian distribution obeys an exponential decay in both negative side and positive side, to guarantee smooth transition across the boundaries of adjacent bins, we empirically set $\alpha_1 = 1$, $\alpha_2 = 1.3$.

Accordingly, we refine the λ of PLA in the positive side by using

$$\begin{aligned} \{\lambda_{k_0}^{pos}, ..., \lambda_K\} &= \arg\min \sum_{k=k_0}^{k=K} ||\lambda_k - \hat{\lambda}||_2^2 \\ s.t.\ \alpha_1 J_{k+1}^{k+1} &\le J_{k+1}^k \le \alpha_2 J_{k+1}^{k+1},\ k = k_0, ..., K - 1 \end{aligned} \quad (5)$$

where

- $\lambda_{k_0}^{pos}$ represents the refined λ w.r.t. k_0th quantization bin in the positive side;
- K is the total number of quantization bins;
- J_{k+1}^{k+1} and J_k^{k+1} are defined by

$$\begin{aligned} J_{k+1}^{k+1} &= f_{\breve{E}'}(\lambda_k, r_{k+1}) - f_{\breve{E}'}(\lambda_k, r_{k+1}+1) \\ J_k^{k+1} &= f_{\breve{E}'}(\lambda_k, l_k) - f_{\breve{E}'}(\lambda_{k+1}, r_{k+1}) \end{aligned} \tag{6}$$

The cost functions in (3) and (5) aim to force the refined piecewise Laplacian distribution to be as similar as possible with the universal Laplacian distribution with parameter $\hat{\lambda}$, while these constraint conditions ensure smooth transition across the boundaries of adjacent quantization bins. Finally, the PLA parameters $\{\lambda_1, ..., \lambda_K\}$ can be obtained by setting

$$\lambda_{k_0} = \min(\lambda_{k_0}^{neg}, \lambda_{k_0}^{pos}) \tag{7}$$

where the *min* operator ensures the envelop associated with k_0th bin to be beyond the counterparts of adjacent bins in both negative side and positive side.

However, directly solve (3) and (5) will encounter the following two problems. First, the constraints in (3) and (5) are highly non-linear inequality groups, which bring a huge computational cost when directly solve; Second, the constraint space defined by the inequality groups in (3) and (5) may be *null*, directly solving may get a *null* or an unpredictable solution. To handle above problems, in this paper, we try to solve them in a greedy manner. By considering the symmetric property, here we only focus on solving (3). The main idea is fixing the value of λ_k when refining the λ_{k+1}. Then, according to (3), λ_{k+1} refining process can be cast as

$$\begin{aligned} \lambda_{k+1} = &\arg\min ||\lambda_{k+1} - \hat{\lambda}||_2^2 \\ &s.t. \begin{cases} \alpha_1 J_k^k < J_k^{k+1} < \alpha_2 J_k^k \\ \lambda_{k+1} > 0 \end{cases} \end{aligned} \tag{8}$$

Lemma 2. *(8) is a convex feasibility problem when the following inequality satisfies*

$$(1+\alpha_1) f_{\breve{E}'}(\lambda_k, r_k) - f_{\breve{E}'}(\lambda_k, r_{k-1}) < \frac{1}{r_{k+1} - l_{k+1}} \times \frac{N_{k+1}}{N} \tag{9}$$

In items of Lemma 2, we first check the inequality (9) before refining λ_{k+1}, and two possible cases are considered:

- If (9) satisfies, we directly solve (8) to obtain the refined value;
- If (9) does not satisfy, we just set $\lambda_{k+1} = \hat{\lambda}$. It is worth noting that we experimentally find this happens only in much rare cases.

Due to the symmetric structure of Laplacian distribution, λ_k in the positive side can be refined in a similar way. Eventually, all the estimated λ_k can be plugged into (6-paper) to get the new conditional distribution of the dither d

$$f_D(D = d | \breve{E} = q_k) = \begin{cases} \frac{\lambda_k}{2(1-y)} \exp(-\lambda_k |d|), & \text{if } \breve{e} = 0 \text{ and } d \in [-\tau, \tau] \\ \frac{\lambda_k}{y^{-1} - y} \exp\{-\text{sgn}(\breve{e}) \lambda_k d\}, & \text{if } \breve{e} = n\Delta \text{ and } d \in [-\tau, \tau] \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where $y = \exp(-\lambda_k \Delta/2)$.