

①

* タイトル : Simulated Annealing for JPEG Quantization

* 背景 : JPEG の量子化テーブルは人の視覚特性に基づき高周波成分を重点的に削除しているため、中・低符号化レートでは視覚的なひずみが発生する。改良版として JPEG2000 が提案されたが置き換えに失敗している。

* 目的 : JPEG2000 の二の舞にならないよう、新たな汎用的な Q テーブルを提案することで内容をほとんど変えずに符号化性能を向上させる。

* 手法 : Simulated Annealing でテーブルの最適解を算出。学習画像は 10 枚、誤差評価は FSIM、これを 400 グループ並列処理。

* 結果 : Q テーブルと比較して FSIM 誤差を 10%改善し、20~50%圧縮率を向上可能となった。作成された量子化テーブルの DC はステップ幅が減少 (Q95,75,50,35 すべて) し、ほとんどの低周波成分はステップ幅が増加する傾向が得られた。誤差評価を SSIM に変更した場合でも同様の傾向が得られた。

* 関連性 : 提案手法の結果から得られた「DC のステップ幅が減少し、低周波成分のステップ幅が増加したほうが性能が良くなる」という知見は、自身の研究の「複数の DCT 係数よりも ICA の DC 一個の方が性能が良い」という結果と通じるものがあると感じた。

②

* タイトル : Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff

* 背景 : レートディストーション (R-D) の歪みを最小化したとしても、知覚の面で最小化するとは限らないため、R-D に知覚的品质を組み込んだ理論が提案されている。しかし、効果の特徴づけも、実証もされていない。

* 目的 : レート (R)、歪み (D)、知覚 (P) の 3 つの間にトレードオフが存在することをいくつかの知覚品質評価を用いた実験により証明する。

* 手法 : 音声信号を再構成する時のエントロピー (R)、ハミング距離 (D)、TV 距離 (P) を算出。画像も同様に、エントロピー (R)、MSE (D)、ワッサースタイン距離 (P) を算出し、3 次元、R 対 D、P 対 D、R 対 P の性能を比較。(画像の知覚品質にワッサースタイン距離を用いたのは、知覚品質は出力信号の入力信号との偏差に関連しているため。と解釈)

* 結果 : D の増加に伴い、一般的な R-D は (最小及び、Medium Quality よりも) P が劣化するため、R-D を最小化しても P を最小化できない。また、R の低下に伴い、D と P が劣化。同じ D であれば R を上昇させることで P を向上できる。などから、R、D、P の 3 つの間にトレードオフが存在することを確認。

* 関連性 : レートディストーションを自身の論文でも引用したいと思っていた。ちょうど良いものがあつたため読んだ。先行研究のように手法自体に取り入れるつもりはないが、レートディストーションに知覚品質を加えているため、結果の比較 (通常

の性能比較とは別で) に用いることで知覚の面でどれくらい最適解に近づいているのかを明らかにし, 別の指標からも ICA の有効性を示せるかもしれない?

* 参考文献 :

①

[1] M. Hopkins, M. Mitzenmahcer, S. Wagner-Carena, “Simulated Annealing for JPEG Quantization,” *Data Compression Conference*, pp.412-412, (2018).

②

[2] Y. Blau, T. Michaeli, “Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff,” *The International Conference on Machine Learning*, pmlr:97, pp.675-685, (2019).