

## 進捗報告

亀田ゼミ

M2 中田雄大

## ● 研究概要と振り返り

## ～ 研究概要 ～

画像符号化の目的は画像中の余分な情報を減らすことで画像の効率的な伝送・保存を行なうこと。代表的な方式である離散コサイン変換 (DCT) は画像が必要とする統計的な特徴を抽出・保存できるが人の視覚は考慮されていない課題があるため、人の視覚野が扱う特徴を抽出・保存できる独立成分分析(ICA)と併用することで、両方の特徴を保存することを目指した符号化手法を提案している。

DCT と ICA を組み合わせる際、どのブロックに DCT または ICA を適用すべきなのかを決める必要があり、各ブロックに用いる適切な ICA 基底の種類と個数を模索する中で、最も符号化性能を改善可能である適用ブロックの組み合わせと ICA 基底の種類、個数を明らかにすることで、DCT と ICA を併用した符号化手法の性能改善を行う。

## ～ 前回の振り返り ～

頂いたコメント・質問、今後の予定を基に、正順・逆順・全探索による優先度の画像全体での画質比較と、逆順の正順からの画質改善、提案手法の符号化性能の3つを調査した。

- ・正順・逆順の画像全体の画質は？
  - 正順の方が圧倒的に全探索に近い画質である。
- ・正順よりも逆順の方が効果的なブロックは？
  - 実際に1%弱のブロックで改善が見られた。検討しても良いかも。
- ・提案手法の符号化性能を改めて比較
  - 低レート以外（実用的なレート）で基底選出可能。より正しい結果になったのではないだろうか。

## ～質問・コメント～

- ・逆順で画質が改善するブロックに共通した特徴などは見られる？
- ・領域を絞って基底選出、性能比較を行った方が ICA の有効性を示しやすいのでは？
- ・最終的な選出領域を提示するとより見やすい（分かりやすい）。

- 進捗報告

前回の進捗で「領域を絞って基底選出、性能比較を行った方が ICA の有効性を示しやすいのでは？」という意見を頂いたが、今後の予定である「基底変更による性能向上」と“領域”を決定する必要があるという点で共通している。基底を作成する領域＝性能比較のための領域＝ICA が有効な領域と仮定した時に、どの領域が適しているのかを今回の進捗で実験的に明らかにし、今後の進捗で汎用的に求める手法を検討してこうと思う。

今回の進捗は、基底を変更した時の符号化性能を比較し、符号化性能が向上する領域の特徴を明らかにする。

\*\*\*\*\* 進捗内容 \*\*\*\*\*

1. 基底を変更した時の符号化性能の分析

\*\*\*\*\*

- 1. 基底を変更した時の符号化性能の分析

今回の実験で使用する、基底を作成するための領域は「ICA が有効」ということで正順の基底優先度を適用した時に DCT から性能が改善するブロックを用いている。（この時、各ブロックにおいて、「ICA 基底を〇個～〇個までの範囲なら性能を向上できる」という範囲が決定される。例、0 番のブロックは 2～4 個の範囲でなら性能改善。～1 では画質が低下し、5～ではエントロピーが増加する。）また、今回の実験でレートごとに比較する名称を表 1 に示す。

表 1 名称表

名称	説明
Hybrid	原画像を入力した時の通常の提案手法のこと。ブロックの優先度は正順。
0 含む ICA 領域	入力画像は各レートでの全ての ICA 領域。（入力画像は変動） 最適な基底数の範囲が 0～64 のブロック群。
0 以外の ICA 領域	入力画像は各レートでの基底 0 個の領域を除く ICA 領域。（入力画像は変動） 最適な基底数の範囲が 1～64 のブロック群。
1～3 のみの ICA 領域	最適な基底数の範囲が 1～3 のブロック群。（入力画像は変動）
Q〇で固定 (Q10, Q20, Q50, Q90)	特定のレートの 0 以外の ICA 領域で入力画像を固定。

※表 1 の各レート入力画像は別紙にあります。

表 1 や別紙資料の図 1, 2, 3 にある画像群を入力画像として性能を比較する理由は、

- ✧ 性能を向上可能な領域特徴はレートごとに異なっているのか？
- ✧ 最適な基底数の範囲が抽出する領域特徴は性能の向上に関係あるのか？

等であり、符号化性能の比較により明らかにしていく。

基底を変更した時の符号化性能（PSNR 対 Entropy）を図 1 に示す。

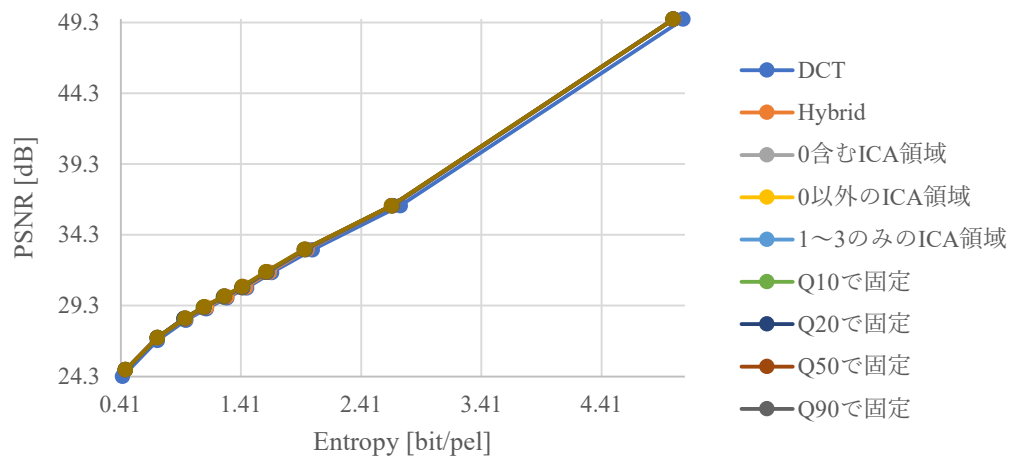


図 1 基底を変更した時の符号化性能（PSNR 対 Entropy）

図 1 は見ても良く分からないため、別紙資料図 4～9 に Q80～Q30 の詳細を示す。また、別紙資料の図 4～9 を見ても良く分からないため、順位付け（主観）を行ったものを図 2 に示す。

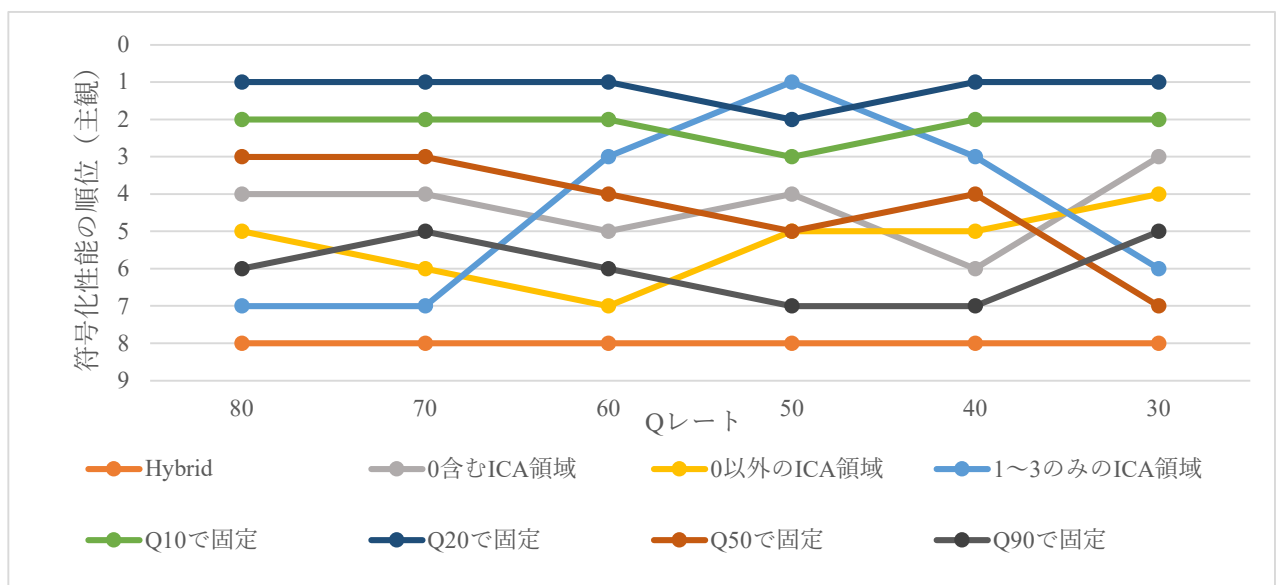


図 2 符号化レートによる性能の推移

図 2 を見るとレートごとに入力画像を変更していくよりも、あるレートに適した領域をすべてのレートで固定する方が符号化性能を向上できることが分かる。このことから、各符号化レートの ICA に適した領域はそこまで変動していないのではないかと推測している。また、固定した入力画像は特徴的なブロックが多い画像 (Q10, Q20)、半々な画像 (Q50)、平坦なブロックが多い画像 (Q90) の 3 パターンとしており、符号化性能は平坦く半々く特徴的という結果であったため、視覚的に特徴的だと思う領域を用いて基底を作成することで符号化性能をより向上できるのではないだろうか。

(Airplane の SSIM 対 Entropy, Barbara の PSNR 対 Entropy, SSIM 対 Entropy は大体同じような結果だったため省略。一応、別紙のおまけに載せてあります。)

図 2 を見ると Q50 付近において「1~3 のみの ICA 領域」の性能が急に向上していることが分かる。ここで、「1~3 のみの ICA 領域」の Q50 で入力画像を固定した場合、どうなるのだろうと気になったため、追加実験を行った。図 2 に対して「1~3 のみの Q50 で固定」の性能を加えたものを図 3 に示す。

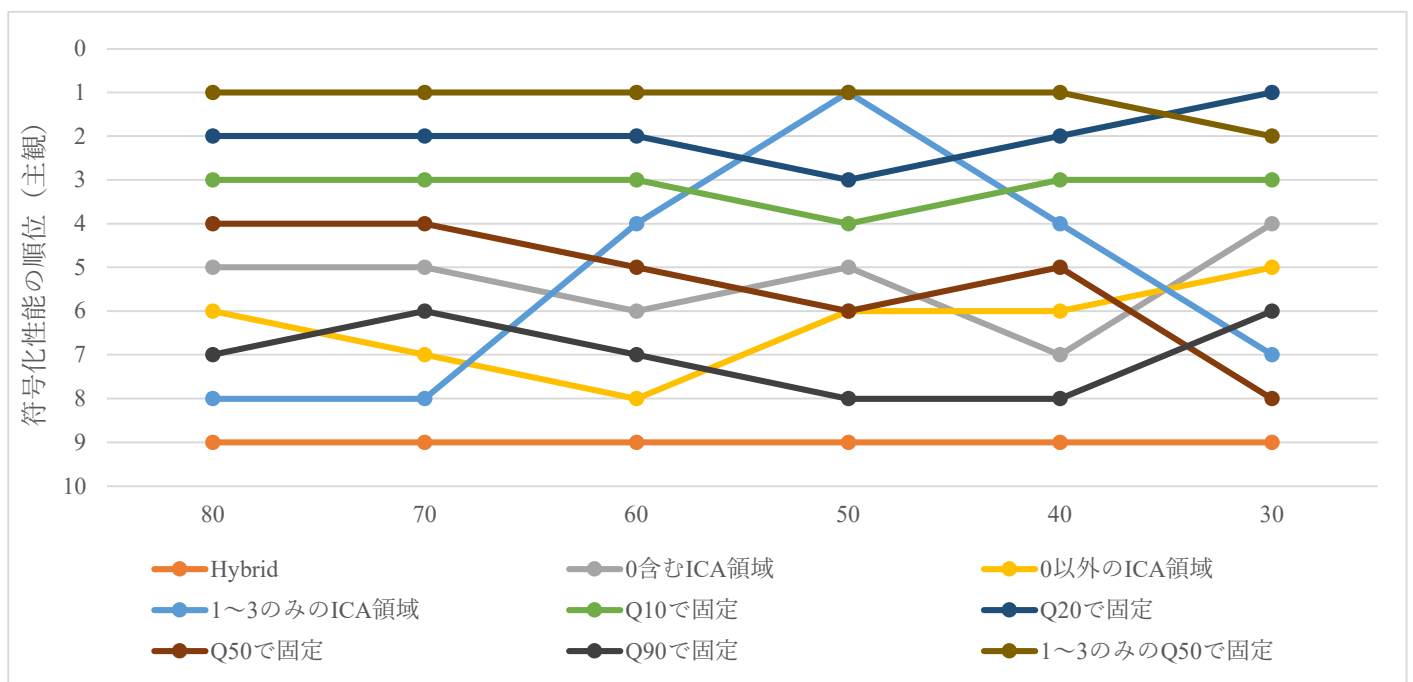


図 3 符号化レートによる性能の推移 (追加バージョン)

また、「Q20 で固定」と「1~3 のみの Q50 で固定」の入力画像の比較を図 4 に示す。

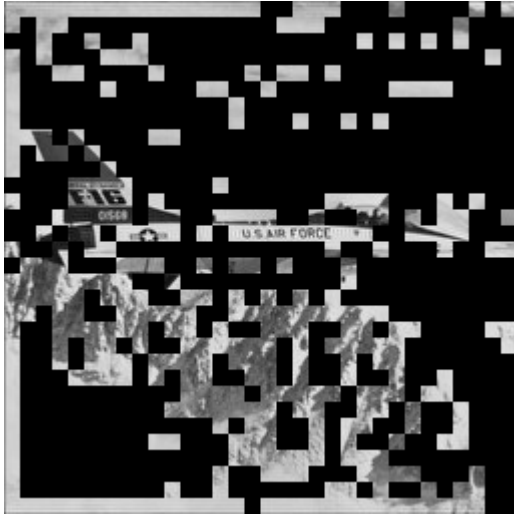


図 4.1 Q20 で固定の入力画像

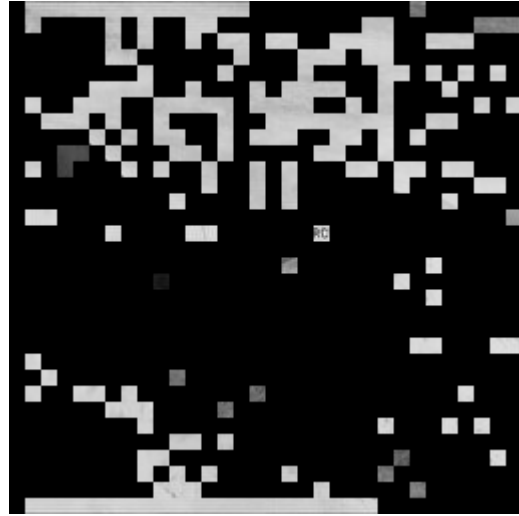


図 4.2 1～3 のみの Q50 で固定の入力画像

図 4 入力画像の比較

図 3, 4 を見ると「1～3 のみの Q50 で固定」の入力画像はほとんど平坦な特徴のブロックであるのにも関わらず、「Q20 で固定」の性能を上回っていることが分かる。選出する基底の数を 3 個までに制限しているため、平坦な特徴はある程度有利であると考えられるが、「1～3 のみの Q50 で固定」は Q10, 20, 50, 90 の性能比較の結果と矛盾している。このことから、平坦な特徴<局所的な特徴<特別な平坦な特徴という関係性があると考えられ、さらに平坦な特徴<局所的な特徴<特別な平坦な特徴<特別な局所的な特徴も存在するのではないかと推測できる。

#### ● 今回の進捗のまとめ

今回の進捗では、前回のコメント・質問・今後の予定を基に、基底を変更した時の符号化性能を比較し、ICA が適した領域はどんな特徴を持っているのかを明らかにした。

・どんな特徴が有効？

→ 視覚的に特徴的だと思う特徴。これをどうやって抽出するのが今後の課題となる（さすがにまだ早い）

今回の実験を通じて、「1～3 のみの Q50 で固定」は選出する基底の数を 3 個までに制限している条件下でのみの話であり、Barbara では確認されていないため、選出する基底数を増やした場合では、今回の結論通り、視覚的に特徴的だと思う特徴の方が

有効なのではないかと推測している。(出力したデータを調べきれてないため、今のところはここまでしか分からない。) 今後として、選出する基底数を増やしたいが、選出基底の評価(現状の3個)では  $64 \times 63 \times 62$  パターンすべての画質を比較しているため、4個以上になるとプログラムの資源的に動かなくなると思われる(多分)。

データを確認したいため全てのパターンの画質を保持し、その後パターンごとに集計(現状保持してるが必要ない)し、上位5パターンの情報も保持している。良い感じの動的配列などの何か良い方法はありませんか? あれば教えていただきたいです。

- 今後の予定

- \* 選出基底数を増やす方法の検討(優先度: 中)
- \* 作成した基底を評価(優先度: 中)
- \* DCT 領域に ICA 基底を加えてみる(優先度: 中)
- \* 領域分類してそれごとに基底を作ってみたい(優先度: 中)