

氏名（本籍）	王 冀（中華人民共和国）
学位の種類	博士（工学）
学位記番号	甲第 245 号
学位授与の日付	令和 3 年 3 月 22 日
学位授与の要件	学位規則第 4 条第 1 項該当
学位論文題目	A Study on Sparse Representation Dictionary Design Method using Multi-class K-SVD and Its Application to Image Coding
論文審査委員	(主査) 教授 八島 由幸 (副査) 教授 熊本 忠彦 教授 今井 順一 准教授 眞部 雄介 東京工芸大学 教授 上倉 一人

## 学位論文の要旨

### A Study on Sparse Representation Dictionary Design Method using Multi-class K-SVD and Its Application to Image Coding

画像の持つ情報量は極めて膨大であり、効率的なネットワーク伝送や、デバイスへの蓄積のためには、圧縮技術が必要不可欠である。現在世の中では、画像圧縮国際標準として、静止画に対する JPEG や、動画に対する H.265/HEVC が、放送・通信・家電等の分野で広く用いられている。しかしながら、近年の、画像の高精細化や SNS の普及により、画像情報がネットワークトラフィックに占める割合は年々著しく増加しており、さらなる高効率圧縮技術の開発が期待されている。

JPEG や H.265/HEVC においては、空間的な相関除去方法として、離散コサイン変換 (DCT) に基づく手法が継続的に採用されている。しかしながら、ある特定の画像を符号化対象として考えると、DCT は必ずしも最適な辞書（すなわち、基底の集合）とはならない。本論文では、スパースコーディングを用いて様々な画像に対して辞書学習を行い、得られた辞書を用いて符号化を行うことで、DCT よりも優れた符号化効率の達成を狙っている。スパースコーディングとは、原信号を、基底の重み付き線形和で表現する際に、非ゼロ係数の個数がある一定の数以下にした時に、原信号と復元信号の二乗誤差が最小になるように、基底及び重み係数を求めるものであり、非ゼロ係数を極めて少数に設定することで大幅な情報圧縮を図ることが期待できる。スパースコーディングを画像圧縮に適用する場合、①画像の局所領域の特徴にどのように適応させるか、②スパースに分布した非ゼロ係数にどのように符号を割り当てて情報量削減するか、③復号画像の主観

画質を向上するためにはスパース係数をどのように量子化すればよいか、を明らかにする必要がある。本論文ではこれらの3つの観点にフォーカスを当てて検討を行い、課題を解決する新手法の提案とともに、実験により提案手法の有効性を確認している。

まず第1は、画像符号化に適したスパース表現可能な辞書設計手法に関する取り組みである。入力データのスパース表現を可能にする辞書を学習するための手法として K-SVD (K-Singular Value Decomposition) を用いる。K-SVD によって設計された辞書の画像表現能力は、学習に用いる画像データの特徴に大きく依存する。従って、入力データを複数のクラスに分類し、クラスごとに辞書を学習する手法を採用する。しかしながら、従来検討されているマルチクラス辞書設計方法は、事前に定められた特徴量を用いた分類に固定されており、画像データの分類段階と辞書学習段階の関係が考慮されていない。このため、本論文では、K-SVD による全学習ベクトルのクラス更新段階と各クラスの辞書更新段階を繰り返し処理によって最適化する新しいマルチクラス辞書設計法を提案する。さらに、設計された辞書を用いた画像符号化システムとして、学習で得られた複数の辞書をエンコーダとデコーダで共有し、画像中の小ブロックごとに最適なクラスの辞書を選択してスパース表現符号化する構成法を提案する。実験の結果、提案手法は、従来の固定的なクラス分けに比較して、BD-bitrate で最大 48%、BD-PSNR で最大 1.6 dB の符号化効率向上を達成できることが明らかとなった。

第2の取り組みは、スパースに分布する非ゼロ係数のエントロピー符号化に関する提案である。本論文では、理論的および実験的観点から重み係数の統計的特性を詳細に分析し、その分析に基づいてスパース係数の効率的なエントロピー符号化方法を提案する。本検討では、非ゼロ係数のレベル値と、非ゼロ係数間のゼロラン長に着目する。詳細な解析により、非ゼロ係数レベル値の分布特性が、ブロック内の非ゼロ係数の個数およびその非ゼロ係数に対応する基底の特徴によって異なることを示し、非ゼロ係数レベル値に対しては、これらに基づくコンテキスト適応符号化が有効であることを明らかにする。さらに、ゼロラン長は、基底の全変動特徴に基づいて基底を並び替えることによって発生確率に偏りを持たせることができ、効率的な符号化が可能であることを示す。実験の結果、提案手法は従来手法と比較して、約 6%の発生ビット量削減を達成できることを明らかにした。

第3の取り組みは、スパース係数の新しい量子化手法に関するものである。スパース表現符号化では、スパース係数を量子化して伝送する。一般的に、JPEG や H.265/HEVC など用いられる DCT 係数に対しては、高周波基底に対する係数の量子化幅を、低周波基底に対する係数の量子化幅よりも大きく設定することで、視覚的な画質を向上できる。一方で、スパース表現を可能とする辞書は、それに含まれる各々の基底が複雑な周波数成分を持ち、DCT 基底や DFT 基底のように規則的な配列を構成していない。そのため、従来の検討では、どの基底に対する係数も同じ量子化幅で量子化する手法が用いられてきた。本論文では、K-SVD によって設計された辞書の各基底の複雑度を定義し、HEVC の DCT 基底の複雑度との類似度に基づいてスパース係数の量子化幅を決定する量子化マトリクスを設計する。この方法により、基底の特徴に応じて適応的に係数量子化を実行できる。主観評価実験により、従来の一様量子化に比較して、同じ圧縮率における 5

段階 MOS (Mean Opinion Score) が約 0.3 ポイント向上し、視覚的画質が改善されることが示された。

以上を総合して、本論文では、スパースコーディングを応用した学習型画像符号化という画像圧縮への新しいアプローチの提案と、それを実現するための3つの必須要素技術、すなわち、辞書設計・係数量子化・符号割り当てについての具体的な方法を明らかにし、従来方式を上回る符号化効率を達成可能であるという知見を得ることができた。

## 審査結果の要旨

画像の持つ情報量は極めて膨大であり、効率的なネットワーク伝送や、デバイスへの蓄積のためには、圧縮技術が必要不可欠である。現在世の中では、画像圧縮国際標準として、静止画に対する JPEG や、動画に対する H.265/HEVC が、放送・通信・家電等の分野で広く用いられている。しかしながら、近年の、画像の高精細化や SNS の普及により、画像情報がネットワークトラフィックに占める割合は年々著しく増加しており、さらなる高効率圧縮技術の開発が期待されている。

JPEG や H.265/HEVC においては、空間的な相関除去方法として、離散コサイン変換 (DCT) に基づく手法が継続的に採用されている。しかしながら、ある特定の画像を符号化対象として考えると、DCT は必ずしも最適な辞書 (すなわち、基底の集合) とはならない。本論文では、スパースコーディングを用いて様々な画像に対して辞書学習を行い、得られた辞書を用いて符号化を行うことで、DCT よりも優れた符号化効率の達成を狙っている。

論文の構成としては、以上述べた研究の背景や意義、従来技術の分析と課題を 1 章および 2 章で述べた後、3 章において、スパースコーディングを画像圧縮に応用する場合に解決すべき 3 つの課題を明確化している。スパースコーディングとは、原信号を、基底の重み付き線形和で表現する際に、非ゼロ係数の個数をある一定の数以下にした時に、原信号と復元信号の二乗誤差が最小になるように、基底及び重み係数を求めるものであり、非ゼロ係数を極めて少数に設定することで大幅な情報圧縮を図ることが期待できる。これを画像圧縮に適用する場合には、①画像の局所領域の特徴にどのように適応させるか、②スパース係数にどのように符号を割り当てて情報量削減するか、③主観画質を向上するためにスパース係数をどのように量子化するか、の 3 つの課題解決が必須であり、以降、4 章、5 章、6 章でこれらの課題①②③のそれぞれに対する新手法を提案している。

第 4 章では、画像符号化に適したスパース表現可能な辞書設計手法を提案している。入力データのスパース表現を可能にする辞書を学習するための手法として K-SVD (K-Singular Value Decomposition) を用いる。K-SVD によって設計された辞書の画像表現能力は、学習に用いる画像データの特徴に大きく依存する。従って、入力データを複数のクラスに分類し、クラスごとに辞書を学習する手法を採用する。しかしながら、従来検討されているマルチクラス辞書設計方法

は、事前に定められた特徴量を用いた分類に固定されており、画像データの分類段階と辞書学習段階の関係が考慮されていない。このため、本論文では、K-SVDによる全学習ベクトルのクラス更新段階と各クラスの辞書更新段階を繰り返し処理によって最適化する新しいマルチクラス辞書設計法を提案した。さらに、設計された辞書を用いた画像符号化システムとして、学習で得られた複数の辞書をエンコーダとデコーダで共有し、画像中の小ブロックごとに最適なクラスの辞書を選択してスパース表現符号化する構成法を提案した。実験により、提案手法は、従来の固定的なクラス分けに比較して、BD-bitrate で最大 48%、BD-PSNR で最大 1.6 dB の符号化効率向上を達成できることが明らかになっている。

第 5 章では、スパース係数の効率的なエントロピー符号化方法を提案している。本検討では、非ゼロ係数のレベル値と、非ゼロ係数間のゼロラン長について、理論的および実験的な解析により、非ゼロ係数レベル値の分布特性が、ブロック内の非ゼロ係数の個数およびその非ゼロ係数に対応する基底の特徴によって異なることを示し、これらに基づくコンテキスト適応符号化が有効であることを明らかにした。さらに、ゼロラン長は、基底の全変動特徴に基づいて基底を並び替えることによって発生確率に偏りを持たせることができ、効率的な符号化が可能であることを示している。実験により、提案手法は従来手法と比較して、約 6% の発生ビット量削減を達成できることが明らかになっている。

第 6 章は、スパース係数の新しい量子化手法を提案している。スパース表現符号化では、スパース係数を量子化して伝送する。一般的に、JPEG や H.265/HEVC など で用いられる DCT 係数に対しては、高周波基底に対する係数の量子化幅を、低周波基底に対する係数の量子化幅よりも大きく設定することで、視覚的な画質を向上できる。一方で、スパース表現を可能とする辞書は、それに含まれる各々の基底が複雑な周波数成分を持ち、DCT 基底や DFT 基底のように規則的な配列を構成していないため、従来の量子化法を適用できない。本章では、K-SVD によって設計された辞書の各基底の複雑度を定義し、HEVC の DCT 基底の複雑度との類似度に基づいてスパース係数ごとに量子化幅を決定できる量子化マトリクスを設計する手法を提案している。この手法により、基底の特徴に応じて適応的に係数量子化を実行できる。主観評価実験により、従来の一様量子化に比較して、同じ圧縮率における 5 段階 MOS (Mean Opinion Score) が約 0.3 ポイント向上し、視覚的な画質が大幅に改善されることが示された。

最後に第 7 章では、提案手法と評価結果をまとめるとともに、今後の研究の方向性にも言及して、論文全体を総括している。

本論文は、スパースコーディングを応用した学習型画像符号化という新しいアプローチの提案と、それを実現するための 3 つの必須要素技術、すなわち、辞書設計・係数量子化・符号割り当てについての具体的な方法を明らかにし、従来方式を上回る符号化効率を達成可能であることを明確にしており、画像圧縮符号化の新しい方向性と性能向上の観点から重要な知見を得たものとして大きな価値がある。従って、学位申請者の王冀は、博士 (工学) の学位を得る資格があると認める。