

科学研究費助成事業 研究成果報告書

令和 3 年 6 月 16 日現在

機関番号：16101

研究種目：研究活動スタート支援

研究期間：2019～2020

課題番号：19K24347

研究課題名(和文)機械学習を応用した次世代動画画像符号化アルゴリズム及びその専用ハードウェアの開発

研究課題名(英文)Next generation video coding algorithm using machine learning and its hardware implementation

研究代表者

片山 貴文(KATAYAMA, Takafumi)

徳島大学・大学院社会産業理工学研究部(理工学域)・助教

研究者番号：70848522

交付決定額(研究期間全体):(直接経費) 2,200,000円

研究成果の概要(和文): AI・IoT環境が普及される中、動画は情報伝達の方法として、今後も大きな役割を担っている。現在も正確かつ精細な動画を提供するために、符号化技術は身近にある様々なデバイスに実装されている。次世代の符号化技術では、並列処理が困難であることから、超高解像度への応用が益々難題となっており、特にハードウェア実装容易度の面から抜本的な解決方法が必要である。本研究を完遂することで、次世代符号化方式と人工知能のそれぞれの技術を統合した専用ICチップが設計可能となる。

研究成果の学術的意義や社会的意義

先行研究では、動画画像符号化と人工知能を組み合わせた手法は提案されていた。しかし、先行研究で提案された方法は、動画画像符号化の効率を向上する半面、人工知能回路がより複雑になることからハードウェアへの実装が困難となっていた。本課題に対して、本研究では人工知能回路の提案と動画画像符号化に対する親和性の評価を実施した。この課題を解決することで、社会で利用される電子デバイスの映像処理が、今後さらに効率的に実現できることが期待される。

研究成果の概要(英文): With the spread of edge environments for AI and IoT, moving images will continue to play a major role as a method of information transmission. Recently, codec technology is implemented to various devices around us in order to provide accurate and detailed moving images. In the next-generation coding technology, parallel processing and application to ultra-high resolution are difficult. Therefore, a drastic solution is required especially from the viewpoint of hardware implementation.

By completing this work, it will be possible to design a dedicated IC chip that combines the technologies of the next-generation coding method and artificial intelligence.

研究分野: Computer science

キーワード: Image processing Versatile video coding Machine learning

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

AI・IoT環境が普及される中、動画は情報伝達の方法として、今後も大きな役割を担っている。現在も正確かつ精細な動画を提供するために、符号化技術は身近にある様々なデバイスに実装されている。次世代の符号化技術では、並列処理が困難であることから、超高解像度への応用が益々難題となっており、特にハードウェア実装容易度の面から抜本的な解決方法が必要である。

2. 研究の目的

本研究は、次世代動画符号化技術に新規機械学習技術を導入することにより、高効率な符号化を達成し、ハードウェア実装時の並列処理困難の問題を解決する。動画符号化技術は、符号化効率の向上を最優先に開発されてきた結果、符号化ユニット間の依存性が高く、並列処理可能な符号化器のハードウェア実装が実現困難となっている。特に高解像度動画の符号化を対象とした、専用 LSI の設計が難題となっている。先行研究では、機械学習技術を用いた一部の符号化パラメータの最適化が行われたが、処理速度と符号化効率の向上を目的とした本質的な改善を実現していない。これらのことから、本研究では、機械学習の最新研究成果を鑑み、符号化構造を抜本的に改造し、動画の分析回路と符号化回路が並列処理可能なアーキテクチャを提案する。

3. 研究の方法

(1) これまでの研究成果[1][2][3]を基に独自の機械学習手法を提案することで動画分析回路を実現する。動画分析回路には、複数の動画パラメータを共通 CNN によって抽出し、動画符号化のリアルタイム処理に適応可能な回路を設計する。CNN にかかる演算量が多く、メモリ使用量も膨大であるため、提案する CNN のアーキテクチャの設計も課題となるが、既に報告された研究成果よりパラメータ削減、データ圧縮、ビット長削減等の手法を併用することで小面積実装を実現する。

(2) 動画符号化回路に必要とされる最低限のパラメータ数と符号化性能の関係性をソフトウェア検証により明らかにする。現在の検証結果では、CNN によるパラメータ抽出は従来手法に比べて遜色のない性能を実現している[3]。従って本符号化構造は、符号化効率と並列処理が両立できる可能性が高い。

4. 研究成果

本報告では、イントラインターの符号化効率を大幅に改善した、動画分析回路を中心に研究成果を報告する。

(1) イントラ符号化では、これまでに提案されている GAN(Generative Adversarial Network)をさらに改善することで、どのような自然画像においても機械学習による圧縮効率の向上が可能となった。図1で示すような、イントラ符号化時の予測画像生成処理に GAN を応用する。これにより、従来では困難であった、予測対象のブロックをより大きなサイズで符号化することが可能となった。図2では予測画像を生成した結果を示す。(a)は入力画像、(b)は従来の方向性予測方法より生成した画像、(c)は提案した GAN から生成した画像である。図2に示すように、従来の方法では困難であったが画像テクスチャのパターンが提案手法により精度良く生成されている。これにより大幅に符号化効率を向上させることに成功した。符号化効率の改善率を先行研究 PS-RNN[4]と比較した結果を表1に示す。平均約 2~3%の符号化効率の向上を達成した。

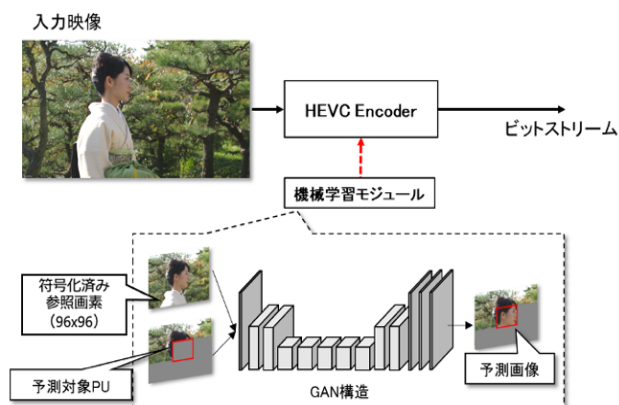


図1 GAN を応用したイントラ符号化方式

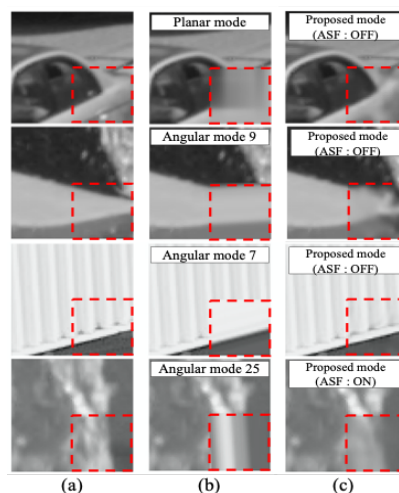


図2 予測結果の比較

表 1 イントラ符号化効率の改善率の比較

Class	Sequence	BD-rate Y		BD-rate U		BD-rate V	
		Proposed	PS-RNN	Proposed	PS-RNN	Proposed	PS-RNN
Class A	Traffic	-3.9	-3.3	-3.7	-2.6	-3.7	-2.4
	PeopleOnStreet	-4.1	-4.0	-3.6	-2.5	-4.4	-2.1
	Class A Average	-4.0	-3.6	-3.7	-2.5	-4.0	-2.3
Class B	Kimono	-6.1	-1.2	-3.4	-0.9	-3.3	-0.9
	ParkScene	-2.2	-2.7	-2.2	-1.6	-2.0	-1.3
	Cactus	-2.7	-2.3	-2.7	-1.5	-1.3	-0.9
	BasketballDrive	-6.3	-1.4	-9.1	-1.2	-8.1	-1.4
	BQTerrace	-5.7	-2.4	-7.5	-0.6	-7.5	-0.5
	Class B Average	-4.6	-2.0	-4.7	-1.1	-4.4	-1.0
Average		-4.4	-2.5	-4.6	-1.6	-4.3	-1.4

(2)インター符号化では、高精度な参照フレームの画像生成手法が難題であることから、機械学習を用いた連続画像の生成手法の提案を行った。本研究課題では、より高精度な参照フレーム生成を実現するため、機械学習処理時の構造にオンライン学習が導入された。インター符号化時の全体概要は図 3 に示す。

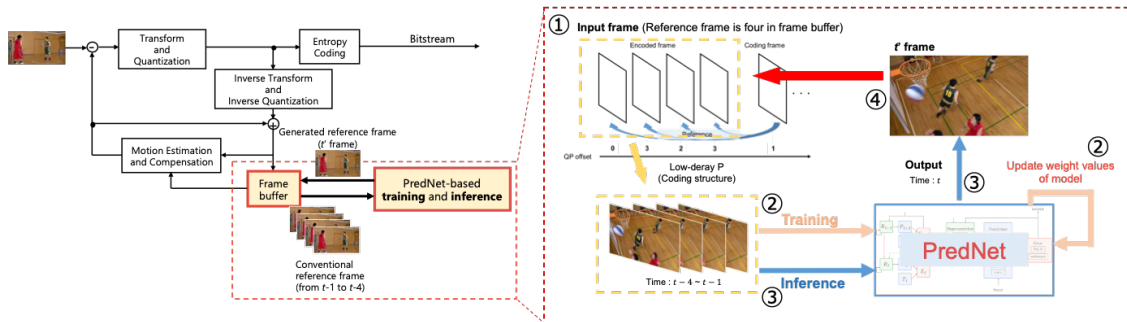


図 3 時系列予測を応用したインター符号化方式

提案したインター符号化方式では、符号化効率が先行研究[5]と比較して、約 2.34%改善されていることを確認した(表 2)。また、提案したインター符号化時の動画画像分析回路から出力されるパラメータと VVC に導入された新規アルゴリズムの親和性を考慮した動画画像符号化回路の設計を行い、国際会議 IEEE International Conference on Consumer Electronics 2021 にて発表を行った[6]。

表 2 インター符号化効率の改善率の比較

Class	Sequence	BD-rate (%)							
		Proposed				Previous Work			
		Y	U	V	Weighted	Y	U	V	Weighted
Class C	BasketballDrill	-0.78	-0.62	0.33	-0.62	3.20	4.35	3.43	3.37
	BQMall	-2.16	-2.31	-2.01	-2.16	0.68	1.04	1.28	0.80
	PartyScene	0.77	0.58	1.15	0.79	1.39	1.69	2.09	1.51
	RaceHorses	-0.16	-4.47	-1.47	-0.86	1.73	1.89	2.43	1.84
	ClassC Average	-0.58	-1.70	-0.49	-0.71	1.75	2.24	2.31	1.88
Class D	BasketballPass	-1.22	-1.01	-3.56	-1.48	0.19	-0.13	0.99	0.25
	BQSquare	2.46	1.26	0.96	2.12	2.50	1.65	1.11	2.22
	BlowingBubbles	0.44	0.30	1.30	0.53	0.82	1.48	1.68	1.01
	RaceHorses	-2.22	-10.19	-8.56	-4.01	1.71	2.64	3.54	2.06
	Class D Average	-0.14	-2.41	-2.47	-0.71	1.30	1.41	1.83	1.38
Average		-0.36	-2.06	-1.48	-0.71	1.53	1.83	2.07	1.63

BD-rate 平均約2.34%改善

[1] 片山 貴文他 : 注目領域検出を用いた符号化アルゴリズムに関する研究, 電気学会論文誌 C, vol. 138, no. 10, pp. 1185-1186, 2018.

[2] T. Katayama et al. : QP Adaptation Algorithm for Low Complexity HEVC based on a CNN-Generated Header Bits Map, in Proc. Int. Conf. Consum. Electron. in Berlin (ICCE-Berlin), 1-5, Berlin, Sep. 2018.

[3] T. Katayama et al. : Fast CU Determination Algorithm Based on Convolutional Neural Network for HEVC, International Journal of Machine Learning and Computing, Vol. 8, No. 5, 442-446, 2018.

[4] Y. Hu et al. : Progressive spatial recurrent neural network for intra prediction, in Proc. of Data Compression Conference, 2017.

[5] T. Laude et al. : HEVC inter coding using deep recurrent neural networks and artificial reference pictures, in Proc. of Picture Coding Symposium (PCS), 2019.

[6] T. Katayama et al. : Reference Frame Generation Algorithm using Dynamical Learning PredNet for VVC, Proceedings of IEEE International Conference on Consumer Electronics (ICCE), Jan. 2021.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 4件/うちオープンアクセス 0件）

1. 著者名 Jiang Xiantao, Song Tian, Katayama Takafumi	4. 巻 12
2. 論文標題 Maximum-Entropy-Model-Enabled Complexity Reduction Algorithm in Modern Video Coding Standards	5. 発行年 2020年
3. 雑誌名 Symmetry	6. 最初と最後の頁 113 ~ 113
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/sym12010113	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Jiang Xiantao, Feng Jie, Song Tian, Katayama Takafumi	4. 巻 19
2. 論文標題 Low-Complexity and Hardware-Friendly H.265/HEVC Encoder for Vehicular Ad-Hoc Networks	5. 発行年 2019年
3. 雑誌名 Sensors	6. 最初と最後の頁 1927 ~ 1927
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s19081927	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Jiang Xiantao, Song Tian, Zhu Daqi, Katayama Takafumi, Wang Lu	4. 巻 21
2. 論文標題 Quality-Oriented Perceptual HEVC Based on the Spatiotemporal Saliency Detection Model	5. 発行年 2019年
3. 雑誌名 Entropy	6. 最初と最後の頁 165 ~ 165
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/e21020165	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Jiang Xiantao, Song Tian, Katayama Takafumi, Leu Jenq-Shiou	4. 巻 11
2. 論文標題 Spatial Correlation-Based Motion-Vector Prediction for Video-Coding Efficiency Improvement	5. 発行年 2019年
3. 雑誌名 Symmetry	6. 最初と最後の頁 129 ~ 129
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/sym11020129	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計12件（うち招待講演 0件 / うち国際学会 7件）

1. 発表者名 Sigalingging K. Xanno, Jules Salzinger, Leu Shiou Jenq, Takafumi Katayama and Tian Song
2. 発表標題 A Study of Smoothness in Neural ODEs
3. 学会等名 The 7th International Forum on Advanced Technologies 2021 (IFAT2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Takafumi Katayama, Tian Song, Takashi Shimamoto and Jiang Xiantao
2. 発表標題 Reference Frame Generation Algorithm using Dynamical Learning PredNet for VVC
3. 学会等名 Proceedings of IEEE International Conference on Consumer Electronics (ICCE) (国際学会)
4. 発表年 2021年

1. 発表者名 Zhang Hongyi, Tian Song, Takafumi Katayama, Takashi Shimamoto and Zhao Liping
2. 発表標題 An Adaptive Selection Method of Screen Content Coding Tools for Educational Video System
3. 学会等名 The 6th International Forum on Advanced Technologies 2020 (IFAT2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Soichiro Iino, Tian Song, Takafumi Katayama and Takashi Shimamoto
2. 発表標題 Low Bitrate VVC-based Drone Video Coding System
3. 学会等名 The 6th International Forum on Advanced Technologies 2020 (IFAT2020) (国際学会)
4. 発表年 2020年

1 . 発表者名 Soichiro Iino, Tian Song, Takafumi Katayama and Takashi Shimamoto
2 . 発表標題 Bitrate Fluctuation Control Method for VVC-based Drone Video Coding System
3 . 学会等名 Proceedings of RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing(NCSP'20)
4 . 発表年 2020年

1 . 発表者名 Takafumi Katayama, Tian Song, Takashi Shimamoto and Xiantao Jiang
2 . 発表標題 GAN-based Color Correction for Underwater Object Detection
3 . 学会等名 OCEANS19 MTS/IEEE Seattle (国際学会)
4 . 発表年 2019年

1 . 発表者名 Lu Wang, Tian Song, Takafumi Katayama and Takashi Shimamoto
2 . 発表標題 Proposal-Aware Visual Saliency Detection with Semantic Attention
3 . 学会等名 International Conference on Intelligence Science and Big Data Engineering (IScIDE 2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 Fumiya Kitawaki, Tian Song, Takafumi Katayama, Xiantao Jiang and Takashi Shimamoto
2 . 発表標題 Improvement of Cross Component Prediction by Multiple Liner Prediction for HEVC
3 . 学会等名 Proceedings of International Technical Conference on Circuits/Systems, Computers and Communications(ITC-CSCC2019) (国際学会)
4 . 発表年 2019年

1. 発表者名 岡久 奈留, 宋 天, 片山 貴文, 島本 隆
2. 発表標題 VVCイントラ予測残差における変換係数予測手法に関する研究
3. 学会等名 電気関係学会四国支部連合大会講演論文集
4. 発表年 2019年

1. 発表者名 中川 靖之, 宋 天, 片山 貴文, 島本 隆
2. 発表標題 オブジェクト検出の高速化アルゴリズムに関する研究
3. 学会等名 電気関係学会四国支部連合大会講演論文集
4. 発表年 2019年

1. 発表者名 山崎 凌佑, 宋 天, 片山 貴文, 島本 隆
2. 発表標題 CNNを用いたHEVCのイントラ符号化効率向上に関する研究
3. 学会等名 電気関係学会四国支部連合大会講演論文集
4. 発表年 2019年

1. 発表者名 長崎 舜, 宋 天, 片山 貴文, 島本 隆
2. 発表標題 VVCの画面間符号化におけるベイジアンネットワークを用いた高速モード決定手法に関する研究
3. 学会等名 電気関係学会四国支部連合大会講演論文集
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------