

令和 5 年 5 月 31 日現在

機関番号：12501

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K11728

研究課題名（和文）誤差許容計算における PCM 書き込み時間削減

研究課題名（英文）Write latency reduction on PCM for approximate computing

研究代表者

難波 一輝（Kazuteru, NAMBA）

千葉大学・大学院工学研究院・准教授

研究者番号：60359594

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究の主目的であった書き込み時間削減については良好な手法を得ることができなかった。しかし、誤差許容計算を用いたシステムにおけるメモリスシステムの消費電力削減技術など関連技術について、複数の手法を明らかにしている。例えば、誤差許容計算を用いたシステムの代表例であるニューラルネットワークシステムにおいては、2種類の電源電圧を使い分ける手法を提案している。重大な認識率の低下無く、消費電力を35%削減することに成功している。

研究成果の学術的意義や社会的意義

本研究においては主目的であった書き込み時間削減については結果を出せていない。しかし、副産物と言える消費電力削減手法などはいずれも実用性の高いものであり、本研究成果の工業的産業的重要性は十分に高いものであったと言える。また、目的であったメモリスシステム書き込み時間についてもいくつかの知見が得られており、研究当初に考えていた問いにもいくらかは答えられた、学術的にも意義がある研究であったと考えている。

研究成果の概要（英文）：This work has not achieved a write time reduction, which was the main objective. However, we have presented several related techniques, such as power consumption reduction on memory systems for an approximate computing system. For example, we have presented a power consumption reduction for a neural network system, a typical example of an approximate computing system. The proposed system uses two different power supply voltages. The proposed method achieves a power consumption reduction of 35%, avoiding a significant reduction in the recognition rate.

研究分野：情報学

キーワード：メモリスシステム 誤差許容計算

1. 研究開始当初の背景

- (1) 現代の情報社会は計算機システムによって支えられている。その中でもメモリシステムは最も重要な部品の1種であり、その性能向上は強く望まれている。従来よく使用されているメモリとして、DRAM や NAND 型フラッシュメモリなどがある。しかし、さらなる発展のため、新しいタイプのメモリシステムが強く望まれている。PCM はこの要望に応えるもっとも有望なメモリの1種と考えられている。しかし、PCM には書き込み時間が長いという欠点もある。そこで、PCM の書き込み時間削減については既に国内外の多くの研究者によって、材料工学、電気電子工学などさまざまな観点から研究がなされている。情報学の観点からは WOM 符号を用いた PCM の書き込み時間削減法が提案されている。符号の符号化器・復号器を用いるというアイデアは工学的に非常に良く、書き込み時間の削減に成功している。さらに、研究代表者らは従来研究においてより高性能な PCM の書き込み時間削減法を提案していた。
- (2) 近年、誤差許容計算に関する研究の重要性が高まりつつある。誤差許容計算機においては、演算処理装置において必ずしも正しい演算結果が得られることを保証しない代わりに、消費電力等のハードウェアオーバーヘッドを削減できるという特徴を有する。誤差許容計算は映像や音声など多少の雑音が加わっても情報が完全には失われないシステムにおいては有用であり、通信システムにおいては既に研究されている。また、誤差許容計算は人工知能の分野においても有用であると言える。

2. 研究の目的

- (1) 本研究はメモリシステムを対象とした誤差許容計算について、PCM の書き込み時間削減法についてその実現と理論を検討することを目的としていた。誤差許容計算においては計算の正しさと、例えば消費電力などハードウェアに関するパラメータがトレードオフの関係にある。誤差許容計算を考えたとき、どのようなトレードオフ関係があるかを理論的・学術的に解析することにより明らかにしようとしていた。

3. 研究の方法

- (1) 本申請課題においてはまず誤差許容計算における具体的な PCM の書き込み時間削減法について検討することとしていた。さらに得られた手法について、シミュレーション実験および実装実験を行い、計算精度、計算時間、消費電力等について測定を行うこととしていた。また、誤差許容を考慮した場合の高信頼化技術についても検討することとしていた。

4. 研究成果

- (1) 本研究の主目的であった書き込み時間削減については良好な手法を得ることができなかった。しかし、誤差許容計算を用いた消費電力削減技術など関連技術について、複数明らかにしており、意義のある研究であったと確信している。本研究成果として例えば以下のものがある。

- (2) メモリシステムの中には図 1 のように電源電圧と信頼性に強い相関関係があるものがある。このようなシステムにおいては電源電圧を下げることで、消費電力を削減することができる一方、信頼性低下のため、システムを使用することが困難になる。しかし、誤差許容計算が許される環境下においては必ずしも高い信頼性を要求されるわけではない。例えば誤差許容計算が許される代表的なシステムであるニューラルネットワークにおいては、図 2 に示すように、語中の特に重要な情報だけ強く保護すれば、メモリシステムの信頼性が低くてもニューラルネットワークシステム全体としては十分に高い信頼性を達成す

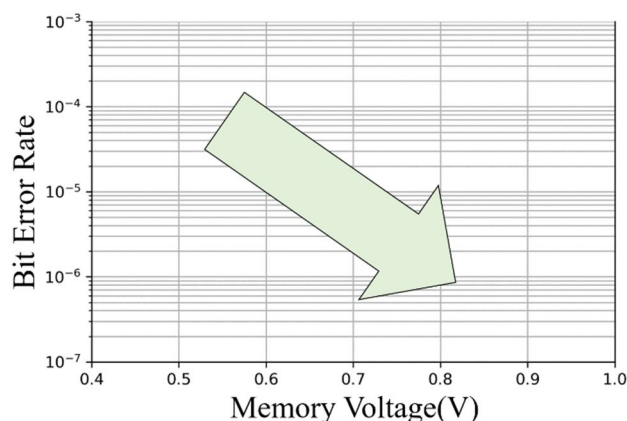


図 1. 電源電圧と信頼性の関係

ることができる。よって、重大なペナルティなしにメモリシステムの消費電力を下げるることができる。具体的に、MNIST を用いた実験では重大な機能低下（認識率低下）をもたらすことなく、消費電力を 35%削減することに成功した。ただし、本研究本来の目的であったメモリアクセス時間については提案手法を用いることで悪化する。

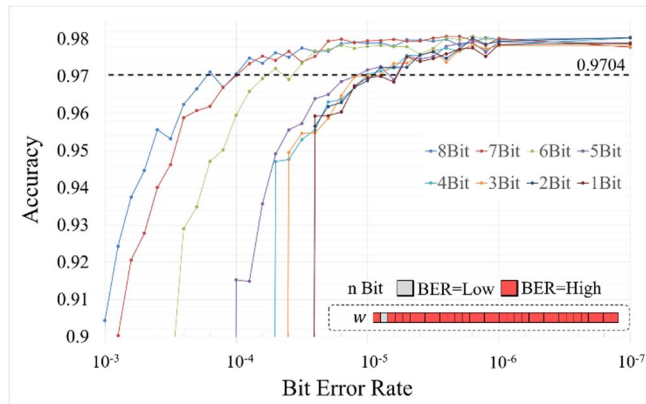


図 2. ニューラルネットワークシステムにおける 2 段階に分けられた語と信頼性の関係

- (3) 先述の複数の電源電圧を扱う手法の応用として、図 3 のような複数のタイプのメモリセルを扱う手法も提案している。本成果については本研究課題の本来のターゲットである PCM など不揮発性メモリにまで応用できていないが、従来型の揮発性メモリシステムである SRAM において良好な結果を得ている。

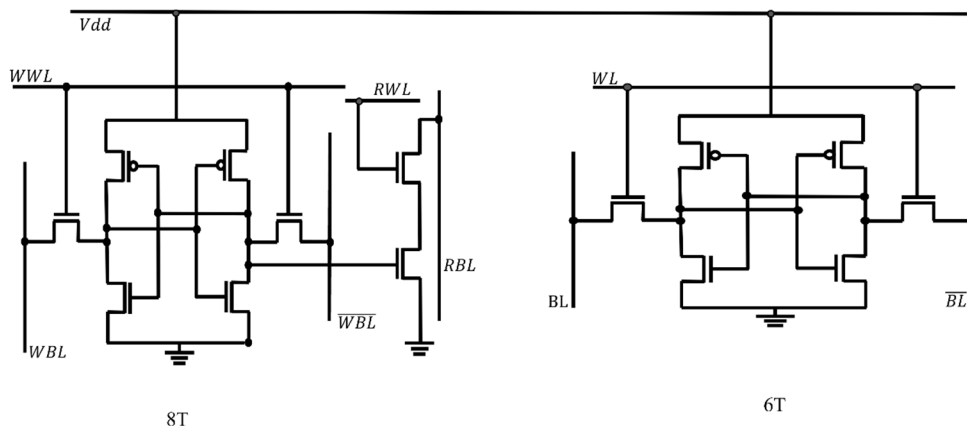


図 3. 2 段階 SRAM システム

- (4) 本研究においては、図 4 に示すように、PCM を用いた連想メモリシステムの構成も明らかにしている。従来の SRAM システムを用いたときと比べ、面積オーバーヘッドの面で良好な結果を得ている。しかしやはり、メモリアクセス速度の面でよくない。このように本研究では、目的の速度向上は達成できなかったが、速度を下げながら、他の利益を得る手法については複数の知見が得られた。

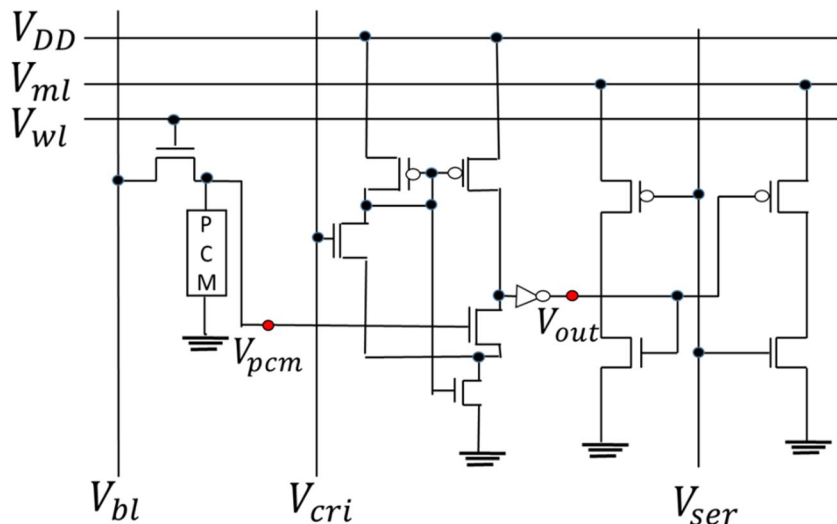


図 4. PCM を用いた連想メモリシステム

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Kozu Keisuke, Tanabe Yuya, Kitakami Masato, Namba Kazuteru	4. 巻 10
2. 論文標題 Low Power Neural Network by Reducing SRAM Operating Voltage	5. 発行年 2022年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 116982 ~ 116986
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2022.3219208	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計10件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 Keisuke Kozu and Kazuteru Namba
2. 発表標題 Relaxing device requirements for non-linearity in Deep Neural Networks accelerators with Phase Change Memory
3. 学会等名 IEEE Int'l Conf. Consum. Electron. Taiwan（国際学会）
4. 発表年 2021年

1. 発表者名 高津 啓佑, 難波 一輝
2. 発表標題 SRAMの動作電圧引き下げによるニューラルネットワークの低電力化
3. 学会等名 信学技報, DC
4. 発表年 2021年

1. 発表者名 高橋 知宏, 難波一輝
2. 発表標題 マルチレベルセル相変化メモリを用いた連想メモリ
3. 学会等名 信学技報, DC
4. 発表年 2021年

1. 発表者名 楊 昊天, 難波 一輝
2. 発表標題 相変化メモリを用いた赤黒木構造の書き込み時間削減
3. 学会等名 信学技報, FIIS
4. 発表年 2020年

1. 発表者名 Tomohiro Ishii, Donghyun Kwon and Kazuteru Namba
2. 発表標題 Stuck-at Fault Tolerance in DNN Using Outliers and Sampling
3. 学会等名 Japan-Korea Joint Workshop on Complex Communication Sciences (国際学会)
4. 発表年 2023年

1. 発表者名 Tomohiro Ishii and Kazuteru Namba
2. 発表標題 Stuck-at fault tolerance in DNN using statistical data
3. 学会等名 IEEE Pacific Rim International Symposium on Dependable Computing (国際学会)
4. 発表年 2022年

1. 発表者名 石井 智大, 難波 一輝
2. 発表標題 外れ値を用いたDNNの縮退故障に対するエラー耐性の向上
3. 学会等名 信学技報, FIIS
4. 発表年 2023年

1. 発表者名 石井 智大, 難波 一輝
2. 発表標題 外れ値と標準化を用いたDNNの縮退故障に対するエラー耐性の向上
3. 学会等名 信学技報, DC
4. 発表年 2023年

1. 発表者名 余 若曦, 難波 一輝
2. 発表標題 動作電圧引き下げによる低消費電力ニューラルネットワークのための6T-8TハイブリッドSRAM
3. 学会等名 信学技報, DC
4. 発表年 2022年

1. 発表者名 Ji Wu, Kazuteru Namba
2. 発表標題 Low power quantized neural network by reducing the operating voltage of SRAM
3. 学会等名 信学技報, DC
4. 発表年 2022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
研究 分 担 者	イン ユウ (Yin You) (10520124)	群馬大学・大学院理工学府・教授 (12301)	

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			
韓国	Pusan National University			