

強誘電体トランジスタを用いた人工知能計算の新方式を開発 ～高い精度での音声認識を実現～

1. 発表者：

名幸 瑛心（東京大学 大学院工学系研究科電気系工学専攻 博士課程 1年）
Kasidit Toprasertpong（東京大学 大学院工学系研究科電気系工学専攻 講師）
中根 了昌（東京大学 大学院工学系研究科電気系工学専攻 特任准教授）
竹中 充（東京大学 大学院工学系研究科電気系工学専攻 教授）
高木 信一（東京大学 大学院工学系研究科電気系工学専攻 教授）

2. 発表のポイント：

- ◆ 強誘電体トランジスタ（FeFET）（注 1、2）を物理リザバー（注 3）とするリザバー・コンピューティング（注 4）と呼ばれる機械学習方式において、音声認識への応用を念頭に、高速計算が可能な並列データ処理により、時系列データを効果的に処理する新しい方式を提案。
- ◆ 0 から 9 までの数字の音声認識に対して、酸化ハフニウム系強誘電体（注 5）からなる FeFET を用いた物理リザバー・コンピューティングにおいて、95.9%の認識精度を実験的に達成。
- ◆ 効率的なオンライン学習や低消費電力の推論の可能性を示すことで、エッジ・コンピューティング（注 6）における人工知能（AI）技術の発展に貢献。

3. 発表概要：

東京大学大学院工学系研究科電気系工学専攻の名幸瑛心 大学院生、Kasidit Toprasertpong（トープラサートボン・カシディット）講師、中根了昌 特任准教授、竹中充 教授、高木信一 教授は、JST 戦略的創造研究推進事業の助成のもと、強誘電体トランジスタを用いたリザバー・コンピューティングという機械学習方式を用いて、高精度の音声認識の実証に成功しました。

IT 技術に欠かせない AI 計算の高効率化と低消費電力化に向けて、AI 計算のハードウェア化と高エネルギー効率の AI 計算方式の実現が強く求められています。この中で、近年、高エネルギー効率で学習ができ、時系列データ処理が得意な物理リザバー・コンピューティングが注目されていますが、実用性の高いデバイスや実用的な応用に対する高効率な計算方式は確立されていません。

本研究では、音声認識への応用を念頭に、半導体製造プロセスに容易に組み込める酸化ハフニウム系強誘電体材料を用いた強誘電体トランジスタ（FeFET）を使った、並列データ処理による物理リザバー・コンピューティングの新しい方式を提案しました。0 から 9 までの数字の英語による音声発話に対する数字認識のタスクに対して、FeFET のドレイン電流、ソース電流、基板電流（注 7）の 3 つの電流成分の時間応答を組み合わせる方式を採用し、応答を測定する時間刻みの最適化、時系列アナログ入力信号の適用、異なる周波数チャンネルの組み合わせなどのさまざまな工夫を盛り込むことにより、認識精度が大幅に向上できることを実験的に示しました。結果として、上記の 0 から 9 までの数字の音声認識において、ソフトウェアを使ったりザバー計算に匹敵する、95.9%の認識精度を実証しました。

本研究成果は、学習負荷が軽く時系列情報処理に向けた物理リザバー・コンピューティングの実用化、特にエッジ・コンピューティング・システムへの応用の方向性を具体的に示してお

り、人工知能活用社会を支える AI ハードウェア・システム技術の革新へ向けた有力な選択肢として、今後の発展が強く期待されます。

本研究成果は、2022年6月12日（ハワイ時間）に国際会議 Symposia on VLSI Technology and Circuits で発行される「Technical Digest」に掲載されます。

なおこの研究成果は、主として、以下の事業・研究領域・研究課題によって得られました。
JST 戦略的創造研究推進事業 チーム型研究（CREST）

研究領域：「情報担体を活用した集積デバイス・システム」（研究総括：平本 俊郎 東京大学 生産技術研究所 教授）

研究課題：「強誘電体分極と電荷の相互作用を利用した新デバイス・システム」

研究代表者：高木 信一（東京大学 大学院工学系研究科 教授）

4. 発表内容：

<研究の背景と経緯>

近年の AI 技術の発展により、AI 技術を活用したさまざまなサービスが急速に展開していますが、現在の AI 計算の多くはソフトウェアにより実現されているため、ハードウェアの点ではメモリ機能を持つ回路と計算機能を持つ回路の間でのデータの移動などに無駄が多く、多大な計算コストや消費エネルギーが必要であり、データセンターなどの大規模な計算機設備での実行が必要です。しかしながら、今後の AI 技術の更なる発展のためには、高速で低消費電力の AI 計算が必須であり、この目的のため AI 計算を効率的に行うことができるハードウェアによる実装を進める必要があります。また、AI 計算の中でも、特に学習機能には多大な計算が必要であることが知られており、低消費電力かつ高速の学習が可能なりザーバー・コンピューティングが、近年、注目されています。以上の観点から、リザーバー機能を物理的ハードウェアで実現した物理リザーバー・コンピューティングという方式が、低消費電力で高速学習が可能でエッジ・コンピューティング向け AI ハードウェアとして、高い関心を集めています。

本研究グループは、新しい物理リザーバー・コンピューティングとして、酸化ハフニウム系強誘電体材料を用いた強誘電体ゲートトランジスタ（FeFET）を活用したリザーバー・コンピューティング方式を提案し、リザーバー・コンピューティングとしての基本動作ができることを実証して、2020年のVLSIシンポジウムで報告しました（図1）。ここで用いた酸化ハフニウム系強誘電体材料は、2011年に初めて強誘電性が報告され、半導体集積回路の製造プロセスでよく用いられている酸化ハフニウムや酸化ジルコニウムを使って実現可能という特徴があることから、現在、半導体業界で活発な研究開発が進んでいる材料であり、Si CMOS（注8）を用いた大規模半導体集積回路と容易に集積化が可能と期待されています。

しかしながら、この FeFET を用いた物理リザーバー・コンピューティングが、どのような応用に有効であるか、また実際の応用に対してどのくらいの性能を発揮できるか、高い AI 計算性能を実現するためにどのようなシステム上の工夫が必要であるか、という点に関しては検証できておらず、そのポテンシャルが十分明らかでないという課題がありました。

<研究の内容>

以上の課題を踏まえて、本研究では、リザーバー・コンピューティングが得意とする時系列データ処理の応用として、音声認識への適用を想定し、高速計算を可能とする並列データ処理による、強誘電体ゲート MOSFET（FeFET）を用いた物理リザーバー・コンピューティングの新

方式を提案し、高い認識率を実現するために必要なさまざまな工夫を導入して、その有効性を実験的に検証しました。

実験には、Si 基板上に 0.7nm の SiO_2 と 10.5nm の膜厚の酸化ハフニウム系強誘電体材料である $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$ (HZO) 膜 (酸化ハフニウム (HfO_2) と酸化ジルコニウム (ZrO_2) の混晶) をゲート絶縁膜として積層した MOSFET を用いています (図 2)。時系列データに対応する入力信号をゲート電圧として印加することで、HZO 膜の分極状態を制御することができます。ここで、物理リザバー・コンピューティングとしては、この分極の記憶特性や分極が時間的に変化する特性を、トランジスタの電流の時間応答特性として読み出し、電流の時間応答のさまざまなパターンを機械学習の手法により分類することで、時系列入力データが含んでいる情報に対する学習や推論を行います。

本研究では、まず、FeFET を用いた物理リザバー・コンピューティングの基礎的性能の向上のため、従来方法のドレイン電流の時間応答による学習・推論に対して、FeFET のドレイン電流、ソース電流、基板電流の 3 つの電流の時間応答を組み合わせることで学習・推論する方式を提案し、その有効性を実験的に示しました。今回応用先とした音声認識を例として、0 から 9 までの数字の英語での音声発話に対して、数字を正しく認識することをタスクとしました。この時、発話の音声データを複数の周波数に分割した時系列データとして扱い、各周波数チャンネルの時系列データをゲート電圧として、各々別の FeFET のゲート電極に入力して並列処理を行い、各 FeFET での推論結果の多数決を取ることで推論を行う方式を提案しました (図 3)。音声発話の特徴的な周波数に分割し並列処理を行うことにより、高速の推論を可能にしました。

この並列処理による音声認識の推論性能を高めるために、時系列データに対する電流応答を読み出す時間刻みの最適化を行うと共に、ゲート電極への入力として従来のデジタル入力からアナログ入力へ変更することで、認識精度が高まることを示しました。更に、異なる周波数チャンネルから得られる電流応答の組み合わせ方を工夫することで、認識精度が向上することを見出し、周波数チャンネルの組み合わせ方法の最適化を進めました。加えて、FeFET 単体としての性能向上が見られたドレイン電流、ソース電流、基板電流の 3 つの電流の時間応答を用いる方法を組み合わせることによって、結果として、音声認識率の実験値として 95.9% を達成しました (図 4)。この高い認識率は、ソフトウェアを使ったりザバー・コンピューティングに匹敵する数値であることから、以上の結果は、FeFET を用いた物理リザバー・コンピューティングが、エッジデバイスにおける時系列データの高効率な情報処理とオンライン学習機能を備えた AI 計算手法として、高い可能性をもっていることを示しています。

<社会的意義と今後の展開>

今回、音声認識応用としての有効性を原理実証した FeFET を用いた物理リザバー・コンピューティングは、従来から期待されていた、学習負荷が軽く時系列情報処理に向けた物理リザバー・コンピューティング方式の実用化、特にエッジ・コンピューティング・システムへの応用の方向性を具体的に示したものとして、今後の発展が強く期待されます。特にここで用いた HZO 膜による FeFET は、現在の集積回路を作製するプロセスとの親和性が高く、比較的容易に集積回路が製造できることから、本方式を用いた大規模集積回路 (LSI) チップも、極めて現実的なソリューションとなり得ると考えられます。この場合、FeFET 物理リザバーと既存の Si CMOS テクノロジーの融合により、今回実証した物理リザバー・コンピューティングの機能を、既存の CMOS システムが提供する信号処理技術や論理演算機能と容易に集積化できるため、高機能の AI 計算用システム LSI への展開も期待できます。以上のことから、これまでの AI 計算方式や AI 計算のためのハードウェアと比べて、極低消費電力性やリアルタイム処理などの

点で、高水準な性能をもっていることが期待できる本提案のデバイスを核とした AI チップは、今後の人工知能活用社会を支える AI ハードウェア技術や AI システム技術の革新へ向けた有力な選択肢となると考えられます。

5. 発表雑誌：

雑誌名：Technical Digest（国際会議 Symposia on VLSI Technology and Circuits で発行）
論文タイトル：Experimental demonstration of novel scheme of HZO/Si FeFET reservoir computing with parallel data processing for speech recognition（音声認識のための並列データ処理による HZO/Si FeFET を用いた新方式リザーバー・コンピューティングの実験的検証）
著者：E. Nako, K. Toprasertpong, R. Nakane, M. Takenaka, and S. Takagi*

6. 問い合わせ先：

<研究に関すること>
東京大学 大学院工学系研究科電気系工学専攻
教授 高木 信一（たかぎ しんいち）
〒113-0032 東京都文京区弥生 2-1-16
Tel：03-5841-0419 Fax：03-5841-0429
E-mail：takagi[at]ee.t.u-tokyo.ac.jp

<報道に関すること>
東京大学 大学院工学系研究科 広報室
Tel：03-5841-0235
E-mail：kouhou[at]pr.t.u-tokyo.ac.jp

科学技術振興機構 広報課
Tel：03-5214-8404 Fax：03-5214-8432
E-mail：jstkoho[at]jst.go.jp

<JST の事業に関すること>
科学技術振興機構 戦略研究推進部 グリーンイノベーショングループ
嶋林 ゆう子（しまばやし ゆうこ）
Tel：03-3512-3531 Fax：03-3222-2066
E-mail：crest[at]jst.go.jp

7. 用語解説：

（注 1）強誘電体トランジスタ（FeFET）

トランジスタとは半導体集積回路において論理演算（計算）などを行う素子のことであり、一般的にはメモリ機能をもたない。強誘電体トランジスタとはトランジスタのゲート絶縁体の部分に強誘電体（注 2）を取り入れ、論理演算の機能と不揮発性メモリの機能を両方備えた素子のことである。近年の酸化ハフニウム系強誘電体膜の発見により、強誘電体トランジスタの研究開発も近年活発に行われている。

(注 2) 強誘電体

誘電体とは電流を流さない絶縁体のうち、電気分極（束縛された正電荷と負電荷の対）が含まれ、外部電界に応じて電気分極の向きや偏りが制御できる物質のことである。通常の誘電体では外部電界を無くすと偏っていた電気分極が元に戻る。一方、強誘電体とは、一度かけた外部電界を無くしても、電気分極が偏ったままで元に戻らない誘電体のことである。また、強誘電体メモリとは強誘電体を用いた記憶素子のことである。

(注 3) 物理リザバー

下記の(注 4)で解説する「リザバー・コンピューティング」において、ソフトウェアではニューラルネットワークで構成されるリザバー部を、高次元の非線形ダイナミクスをもつ材料やデバイスなどのハードウェアで置き換えた方式のことである。

(注 4) リザバー・コンピューティング

リカレントニューラルネットワークの特殊なモデルとして発展した計算方式であり、学習すべき読み出し重みが1層しかないという特徴をもつ、時系列情報処理に適した機械学習手法のひとつである。他のリカレントニューラルネットワークモデルに比べて、学習が極めて高速であり、効率的なオンライン学習が期待できるため、エッジデバイスでの時系列データのリアルタイム処理による音声認識などへの応用が期待されている。

(注 5) 酸化ハフニウム系強誘電体

通常の誘電体膜である酸化ハフニウム膜 (HfO_2) 中に Si、Al、Zr などの種々の元素を混ぜることによって、下記の(注 6)で解説する「強誘電体」特性を発現させた薄膜のこと。2011年に刊行された論文で、初めてその存在が報告された。

(注 6) エッジ・コンピューティング

PC やスマートフォンに代表されるユーザの手元や近くの情報処理機器を用いて、データの処理を行うこと。データセンターの負荷やデータセンターとの通信遅延などの問題を引き起こすことなく、情報処理を行うことができる。

(注 7) ドレイン電流、ソース電流、基板電流

ゲート電極に加えた電圧により半導体側に電子（負の電荷）あるいは正孔（正の電荷）のキャリアを誘起して、電流のオン・オフ動作を行う MOS 型トランジスタ (MOSFET) において、キャリアが半導体側から電極を介して外に流れ出る部分であるドレインを流れる電流をドレイン電流、キャリアが流入する部分であるソースを流れる電流をソース電流、半導体基板を通して流れる電流を基板電流と呼ぶ。

(注 8) CMOS

半導体側に電子を誘起して動作させる MOSFET を n チャネル MOSFET、正孔を誘起して動作させる MOSFET を p チャネル MOSFET と呼び、この n チャネル MOSFET と p チャネル MOSFET を接続して、相補型の動作をさせた半導体構造が CMOS である。現在の半導体集積回路の基本構造となっている。

8. 添付資料：

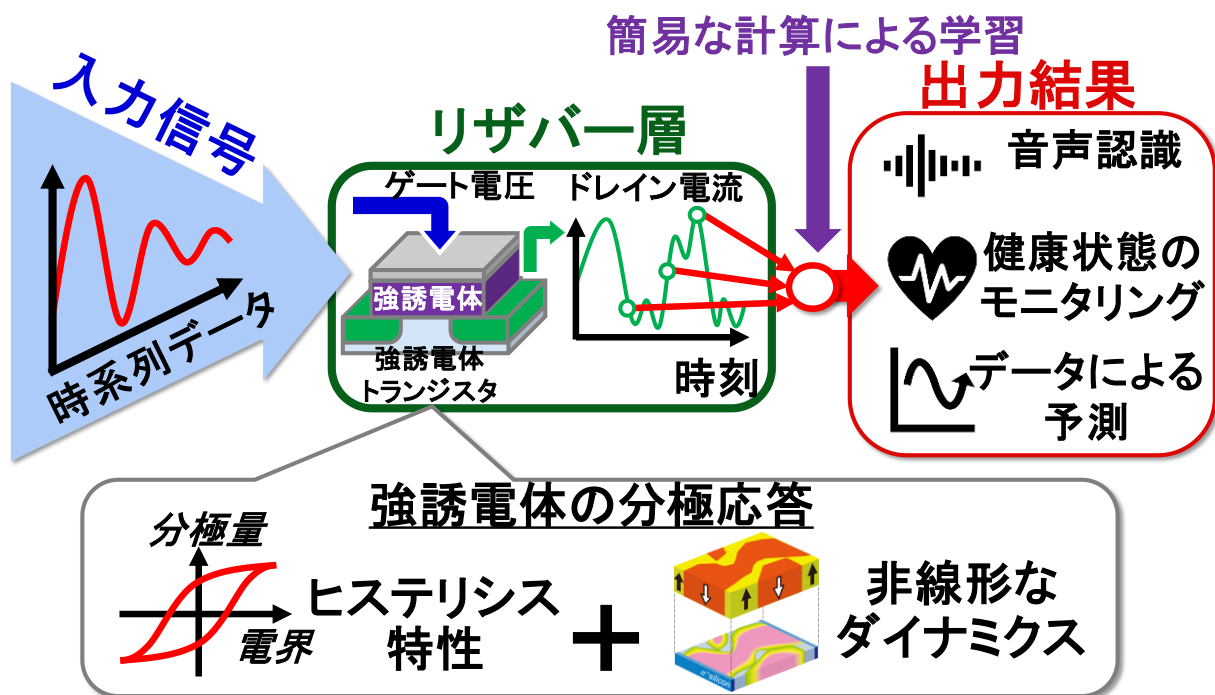


図1 FeFETを用いたリザーバー・コンピューティングの模式図。FeFETによるリザーバー・コンピューティングでは、強誘電体のもつ電気分極の応答特性を利用することで、AI計算が実現できます。リザーバー・コンピューティングは一般に、時系列データに対して判断を下すような応用、例えば、音声認識や健康状態のモニタリング、データによる予測などのタスクが得意とされています。この性質とFeFETによるリザーバー・コンピューティングのもつ低消費電力かつ高速の学習ができる特性を組み合わせることができるため、今回提案した方式は、エッジ・コンピューティングにおいて優れたAI機能を実現できるハードウェアとして期待ができます。本研究では、特に重要なアプリケーションの1つである音声認識をターゲットとしました。

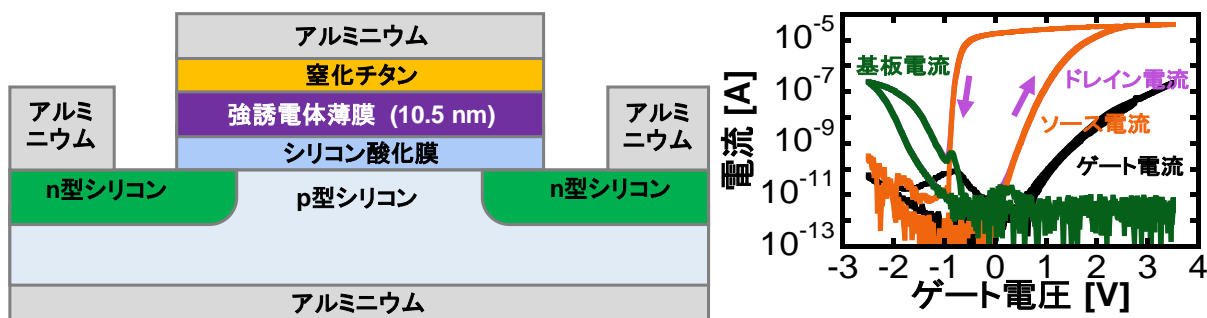


図2 HZO/Si FeFETの(左)素子構造と(右)電流-ゲート電圧特性。FeFETでは、ゲート絶縁膜として、 $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$ (HZO) 膜を用いることで、MOSFETの電気特性にHZO膜中の

電気分極の効果が表れることとなります。ゲート電圧を低い電圧から高い電圧に上げた場合と高い電圧から低い電圧に下げた場合では、ドレイン電流の値が変化しており、これはゲート電圧による HZO 膜中の電気分極の変化により、トランジスタの電流-電圧特性が影響を受けた結果です。

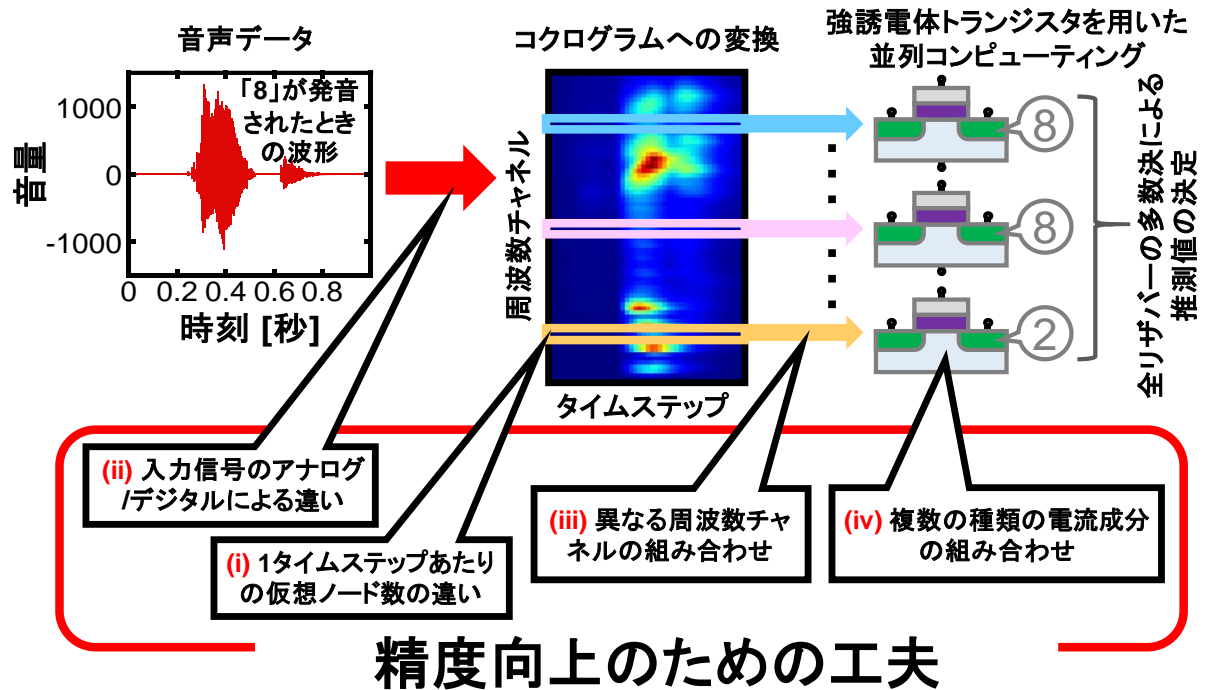


図3 音声数字認識タスクのための FeFET リザーバー・コンピューティングと認識精度向上のためのアプローチの概念図。音声データは、コクログラムと言われる多くの周波数チャネルに分割された時系列データに変換できます。今回提案した FeFET リザーバー・コンピューティングでは、この異なるチャネルのデータを並列に処理することによって、効率的な AI 計算を行います。本研究では、音声数字の認識率を高めるために、図に示した様々な工夫を行いました。

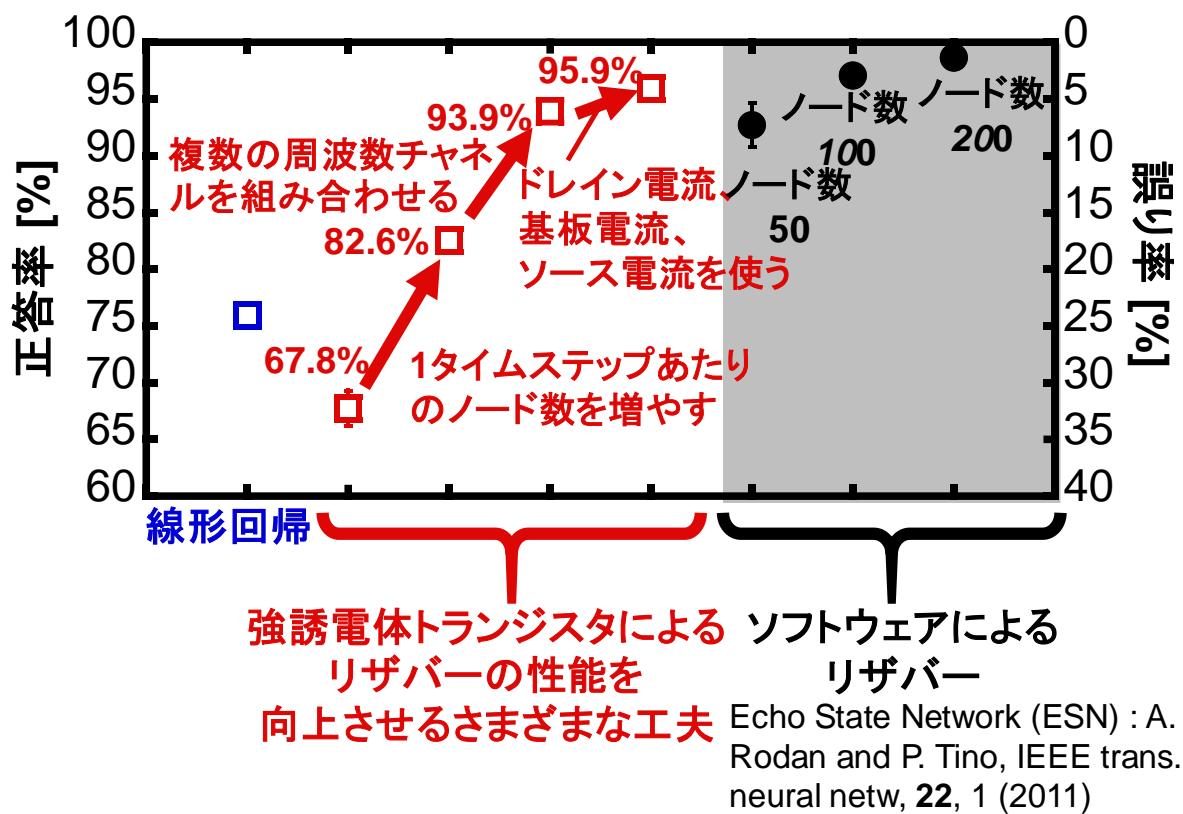


図4 FeFET リザーバ・コンピューティング方式の改良に伴う認識精度の向上と、直接線形回帰およびソフトウェアリザーバの結果との比較。図中に示すさまざまな工夫により、音声数字の認識率が高まっていき、すべての工夫を組み合わせることで、95.9%の高い認識率を実現できました。この認識率は、例えば、50のネットワークサイズをもつリザーバをソフトウェアにより実現したリザーバ計算の認識率を越えており、優れた性能と言えます。