

高速・低消費電力かつスケーラビリティを有する時間領域連想メモリ回路

学生氏名: 安田 雅浩 Email: yasuda-masahiro@hiroshima-u.ac.jp, TEL: 082-424-6265

推薦者: 広島大学大学院 先端物質科学研究科 半導体集積科学専攻 Mattausch Hans Jürgen

研究・開発の目的・狙い

連想メモリはパターン認識等に適用可能な、複数の参照データの中から入力データに最も似ているデータを探し出し、そのデータまたはアドレスを出力する機能メモリであり、画像・音声認識のような高度な認識処理や、画像圧縮などの処理に応用できる。各データの距離（類似度）を表現するためには完全デジタル、デジタル・アナログ混載など様々な手法があるが、本試作チップは小面積・低消費電力かつ微細テクノロジーへのスケーラブルな設計が可能な距離比較手法として、時間領域を利用した新しい手法を用いている。本試作では、リングオシレータベースの時間領域（周波数マッピング）連想メモリの性能とスケーラビリティを確認することを目的としている。

連想メモリによる最小距離検索の実現方法

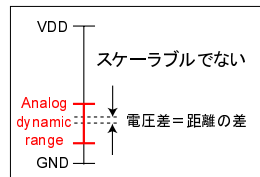
最小距離の参照データを Winner, それに次ぐ類似度をもつデータを Nearest loser, それ以外のデータを Loser と呼ぶ。Winner の検索には検索データとデータベースのそれぞれの参照データの距離の計算・比較という膨大な処理が必要とされる。この処理を完全デジタル回路で行うには回路面積・検索速度・消費電力が大きくなってしまふ。そこで電流・電圧差変換のようなアナログ変換によって高速・低消費電力での検索と小面積での LSI 化を可能にした。この手法では、最終的に距離の差を電圧差という有限の領域に変換して比較を行うため、電源電圧の低下やデータ数の増加に伴い、距離の差を表現できる範囲が狭くなってしまふ。また、トランジスタの微細化に伴う特性ばらつきの影響を完全に打ち消すことは難しく、Winner と Nearest loser を区別できないなど誤検索を生じさせる原因となった。時間領域（周波数マッピング）連想メモリは、距離の差を無限の領域である時間領域に変換することで、ばらつきの影響を軽減し、かつ高速・低消費電力での LSI 化を可能にする。

時間領域（周波数マッピング）連想メモリのコンセプト

提案時間領域（周波数マッピング）連想メモリでは、参照データと検索データの距離の差を信号伝播の遅延時間に変換している。図 1 に示すように、時間差は設計によって無限に拡大できるので、ばらつきの影響を打ち消すこ

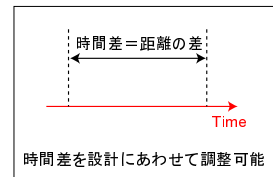
とが可能である。提案回路では、各データの距離に応じて、入力から出力までの遅延回路の総数が変わる仕組みを取り入れている。つまり、距離が小さければ小さいほど、入力から出力までの遅延時間は短くなる。検索開始信号を同時に入力し、最初に出力が出てきた参照データを Winner とすることで最小距離検索を実現する。距離を遅延時間に変換する回路にはリングオシレータ状の遅延回路を採用し、距離 周波数変換を行なった後、周波数 時間変換を行っている。これにより、ばらつきの影響が大きく誤検索が生じてても、設計後に時間差を拡大することができ、検索精度を向上させることができる。この手法による連想メモリを特に周波数マッピング連想メモリと呼ぶこととする。図 2 に周波数マッピング連想メモリのブロック図を示す。

アナログ信号を用いた考え方



距離を表現するための電圧はダイナミックレンジの範囲内でのみ取り扱うことができる

時間領域を用いた考え方



距離の差を時間差で表現することで、設計に合わせて取り扱える範囲を調整可能

図 1: 時間領域マッピングのコンセプト。

リングオシレータベース距離 - 周波数変換回路

距離 - 周波数変換回路には段数可変のリングオシレータベース回路を利用している。図 3 に回路図を示す。インバータチェーンによる遅延回路のあるパスと無いパスを選択可能な 1 ビットステージを各ビットの一つずつ用意し、リングオシレータ制御用の NAND を一つ挿入することでリングオシレータを形成している。各 1 ビットステージは対応するビットの比較結果に応じて、比較結果が一致していれば遅延回路の無いパス (Short path), 不一致ならば遅延回路のあるパス (Long path) を選択する。ここで、全てのパスが Short path だった場合のリングオシレータ 1 周の遅延時間を τ_0 とし、Long path ひとつ分の遅延時間を τ_S とすると、ハミング距離 n_H のリングオシレータの発振周波数 f_{n_H} は

$$f_{n_H} = \frac{1}{2 \cdot (\tau_0 + n_H \cdot \tau_S)} \quad (1)$$

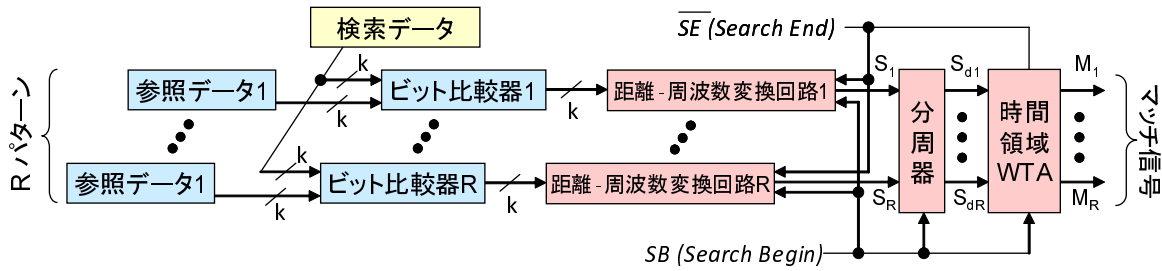


図 2: 周波数マッピング時間領域連想メモリのブロック図

となる．ハミング距離 n_H が小さいほど発振周波数は高くなり，リングオシレータを信号が 1 周する時間は短くなる．Long path の遅延時間 τ_S の大きさはシミュレーションによってばらつきの影響があっても検索結果に影響しないように設計される．

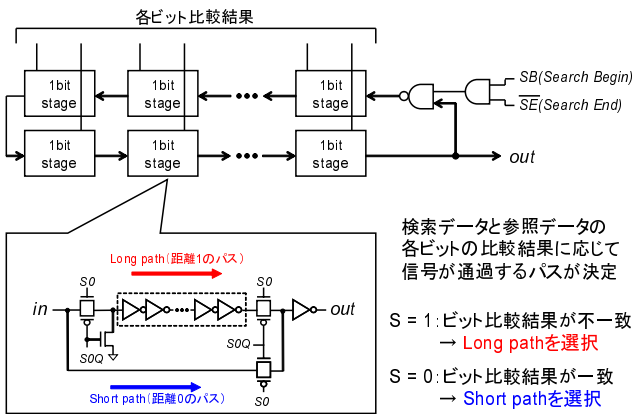


図 3: 段数可変リングオシレータ．

分周器ユニットによる周波数 時間変換

分周器はリングオシレータの何周目のループの出力で時間差を比較するか決定する回路である．分周回数は分周器ユニットへの外部からの制御信号によって指定でき，設計後に検索精度を向上させることができる．分周回数を N とすると，検索開始からリングオシレータの出力が出るまでの時間 τ_{n_H} は，

$$\tau_{n_H} = N \cdot (\tau_0 + n_H \cdot \tau_S) \quad (2)$$

となる．つまり，比較に使う時間差を N 倍に拡大することができる．Winner と Loser を判別するためには次節でのべる時間領域 WTA 回路の分解能以上の時間差が必要となるが，Winner と Nearest loser の距離が 1 しか離れていない τ_{n_H} と τ_{n_H+1} の時間差，つまり τ_S の時間差しかないような厳しい条件では，ばらつき等の影響により，2 つの信号の時間差が時間領域 WTA 回路の分解能を下回る可能性がある．その場合，時間領域 WTA 回路は 2 つの信号のどちらも Winner として出力する．非常に高い検索

精度が必要な場合，分周器を使って時間差を N 倍に拡大することで時間領域 WTA 回路で判別できる時間差まで拡大することができ，最小距離の Winner のみを出力することができる．分周を行うと検索時間はループ回数に比例して大きくなるため，検索時間と検索精度のトレードオフで分周回数を決定する．

時間領域 Winner Take All 回路

時間領域 Winner Take All (WTA) 回路は分周器で指定された回数ループして出力されたリングオシレータの出力信号の中で最初に变化したものを Winner とし，それ以外のデータを Loser と判定するための回路である．提案

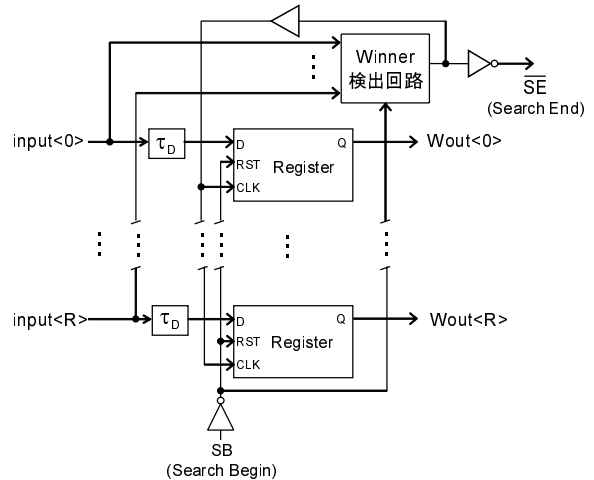


図 4: 時間領域 WTA 回路．

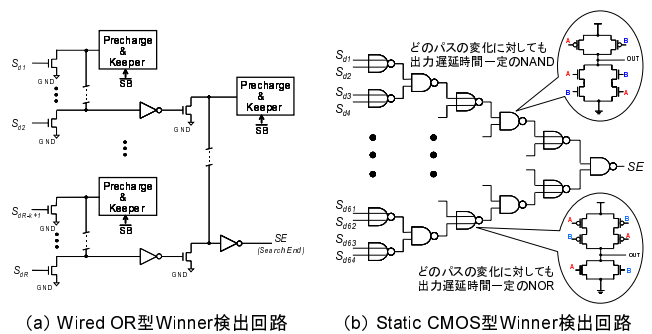


図 5: Winner 検出回路．

連想メモリのために図4に示す高分解能の時間領域WTA回路を開発した。

この回路は各参照データ行に遅延回路とレジスタを一つずつ備え、Winner 検出回路と Search End ツリーを全体に一つ備えている。Winner 検出回路は Winner 行の出力信号の変化を検出し、検出信号を発生させる。Search End ツリーは全行のレジスタのクロック端子にスキューの無いように検出信号を分配する。Winner 行のレジスタのデータ入力だけが変化している時に全行のレジスタをラッチすることで、Winner 行と Loser 行を判別している。ここで、Winner 信号が変化してから、全行のレジスタをラッチするまでに必要な時間を τ_{WD} 、Winner の信号がレジスタにデータ入力される時間を τ_D 、レジスタのセットアップ時間を τ_{ST} とすると、Winner と Loser を判別するために必要な時間差 $\tau_{\Delta var}$ は、

$$\tau_{\Delta var} = \tau_{WD} - \tau_D - \tau_{ST} \quad (3)$$

となる。高分解能を有する為には Winner の信号がレジスタに入力された直後にレジスタをラッチすることが望ましい。そのため、レジスタのデータ入力端子の前には遅延回路を挿入し、 τ_D だけ遅延させている。判別分解能である時間差 $\tau_{\Delta var}$ は τ_{WD} と τ_{ST} のばらつきを考慮して決定される。Winner 検出回路には図5に示すように、Wired OR 型と Static CMOS 型の回路を採用した。これらはこの行が Winner になっても等しい時間で高速に検出信号を出力する特長がある。

2ビットパスエンコーディングによる高速化

以上が周波数マッピング連想メモリの基本構成であるが、更に検索時間を高速化させる手法として、パスエンコーディング手法を開発した。これは図6に示すように、1ビット毎にパスを選択していた従来の処理を複数ビット分まとめて処理することで最短パスのゲート数を減少させる手法である。これにより、1ビットステージと1ピッ

トステージの間に必要だったゲートを削減することができ、M ビットをまとめた場合、基本リングオシレータ遅延時間 τ_0 を約 M 分の 1 に短縮できる。その結果、リングオシレータの出力の遅延時間は、

$$\tau_{nH} = N \cdot \left(\frac{\tau_0}{M} + n_H \cdot \tau_S \right) \quad (4)$$

となり、全体の検索時間の基本周波数に依存した遅延時間の一部を削減することができる。今回の試作では2ビットのパスエンコーディングを行い、その効果と有効性を確認した。

試作結果

周波数マッピング連想メモリの試作結果を表1に示す。試作チップは最小ハミング距離検索連想メモリである。テストチップ1 (TC1) は $64 \times 256\text{bit}$ 、テストチップ2 (TC2) は $128 \times 512\text{bit}$ の連想メモリとしては世界最大規模のデータサイズであり、データサイズに対するスケラビリティを確認できる。また、TC2のみ2ビットパスエンコーディングを取り入れている。実測結果から、最短検索時間はTC1で50ns、TC2で60nsとなった。TC2ではビット数をTC1の2倍に拡張しているため、本来であればTC1の2倍程度の検索時間になるが、2ビットパスエンコーディングによって約40ps程度削減され、ほぼ同等の検索時間となった。最大消費電力に関しても、データサイズの増加に対して、大幅な増加は見られなかった。回路規模はデータサイズに比例し、約4倍となっている。それぞれの結果を規格化し他研究と比較すると、図8に示すように、TC2では飛躍的に向上しており、デジタル手法と比較して、半分以下の検索時間と30分の1以下の消費電力での検索を実現した。また、アナログ手法である電流マッピング手法と比較しても、約半分の検索時間と約3分の1の消費電力となった。

また、アプリケーションで想定される Winner 距離100程度までの検索精度は100%を達成し、Winner 距離128以降では Winner と Nearest loser の差が距離1しかない場合のような、非常に厳しい条件で Nearest loser が Winner として出力される組み合わせがわずかに現れる。この検索の誤差は、例えば Winner の距離が200のときに201の距離を持つ Nearest loser のデータを出力してしまうといった状態であり、アプリケーションに適用した場合は問題ない精度であると考えられる。厳しい条件での Winner と Nearest loser の差を区別する分解能を図9に示す。検索の分解能1の時 Winner と Nearest loser の差1を100%区別でき、分解能2のときは距離2の差を100%区別できる。この結果は、従来のアナログ手法と比較すると非常に優れており、微細テクノロジーに対してスケラブルな手法と言える。

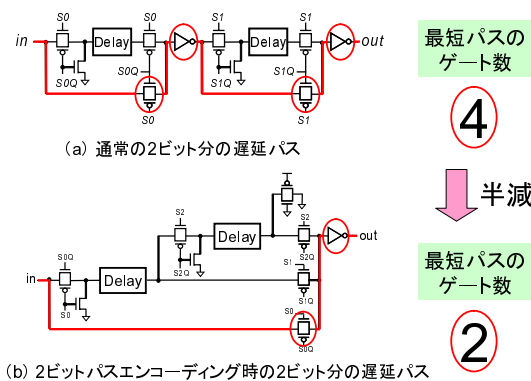


図6: 2ビットパスエンコーディング。

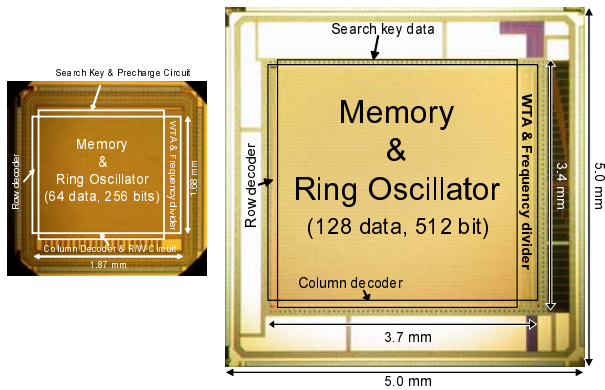


図 7: チップ写真。左は 64 × 256bit , 右は 128 × 512 bit .

表 1: 試作チップの諸元 .

	テストチップ1	テストチップ2
テクノロジー	Rohm 0.18 μm	Rohm 0.18 μm
試作ラン	2009年第1回	2009年第5回
電源電圧	1.8 V	1.8 V
距離指標	ハミング距離	ハミング距離
データサイズ	64-word 256-bit	128-word 512-bit
回路規模	1.9 mm × 1.7 mm	3.7 mm × 3.4 mm
2ビットパスエンコーディング	未使用	使用
WTA回路の構成	Wired OR型	Static CMOS型
分解能 τ_{Avar}	160 ps	154 ps
距離1に対する時間差 τ_s	770 ps	1070 ps
検索時間	50 – 245 ns	60 – 575 ns
	3.05 – 14.95 ps/bit	0.91 – 8.78 ps/bit
最大消費電力	36.54 mW	41.94 mW
	2.23 μW/bit	0.64 μW/bit
面積	全体	3.08 mm ²
	メモリ	0.94 mm ² (31%)
	リングオシレータ	1.63 mm ² (53%)
	WTA	0.11 mm ² (4%)
	12.58 mm ²	0.22 mm ² (2%)

まとめ

リングオシレータベースの周波数マッピング連想メモリを試作し、有効性を検証した。実測の結果、アプリケーションに適用可能な検索精度を有しており、従来手法の半分以下での最短検索時間と、デジタル手法の30分の1以下アナログ手法の3分の1以下の消費電力での最小距離検索を実現した。

発表文献

- [1] M Yasuda, et al., “Low - complexity time - domain winner - take - all circuit with high time-difference resolution limited only by with-in-die variation,” Extended Abs. of SSDM2010, pp. 1164-1165, 2010.
- [2] H. J. Mattausch, et al., “Low-power word-parallel nearest - hamming - distance search circuit based on frequency mapping”, Proc. of ESSCIRC 2010, pp. 538-541, 2010.
- [3] T. Ansari, et al., “Analysis of within - die and die - to - die CMOS - process variation with reconfigurable ring - oscillator arrays”, Extended Abs. of SSDM2010, pp. 802-803, 2010.

参考文献

- [1] H. J. Mattausch et al., “ An architecture for compact associative memories with deca-ns nearest - match capability up to large distances ”, Tech. Digest of ISSCC, pp. 170-171, 2001.
- [2] S. Nakahara et al., “ A digital circuit for a minimum distance search using an asynchronous bubble shift memory ”, Tech. Digest of ISSCC, pp. 504-505, 2004.
- [3] Y. Oike et al., “ A high-speed and low-voltage associative co-processor with Hamming distance ordering using word-parallel and hierarchical search architecture ”, Proc. of CICC, pp. 643-646, 2003.
- [4] H.J. Mattausch et al., “ Fully-Parallel Pattern-Matching Engine with Dynamic Adaptability to Hamming or Manhattan Distance ”, Proc of Symposium on VLSI Circuits, pp. 252-255, 2002.

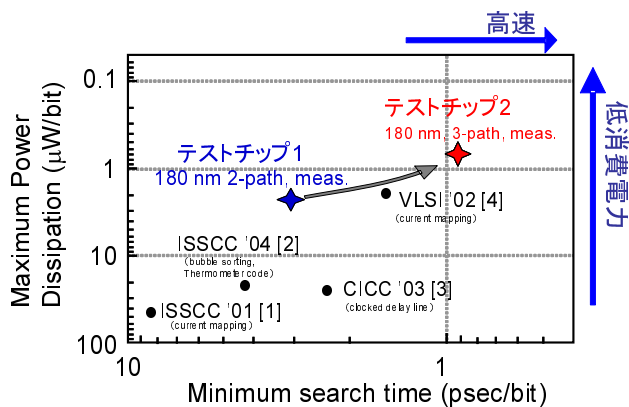


図 8: 最短検索時間と最大消費電力の比較 .

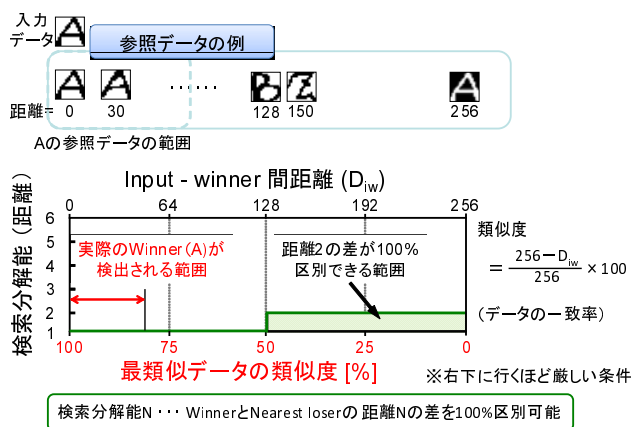


図 9: 検索の分解能 .