

$\Delta\Sigma$ 変調に基づくパルスニューラルネットワークにおける1bit信号同士の乗算の実現

Realization of 1 bit Multiplier for Pulsed Neural Networks Based on Delta-Sigma Modulator

LH-3

村橋 善光[†]道木 慎二[‡]大熊 繁[†]

Yoshimitsu Murahashi Shinji Doki Shigeru Okuma

1 はじめに

ニューラルネットワークの持つ並列演算の特徴を十分に発揮するには、ニューロンをハードウェアに並列実装することが望まれる。ニューロンをハードウェアに実装する場合、デジタル実装とアナログ実装の2つのアプローチから研究されている。アナログ実装の場合、比較的小さな回路規模で実現可能だが、設計が複雑で、外界変化により内部の特性が変化しやすいなど、実際には使いにくいものと思われる。一方、デジタル実装の場合は、ノイズに対してロバストであり、設計が容易であるという利点がある。特に、近年ではFPGAなどのように回路が書き換え可能なデバイスの性能や回路規模も向上し、手軽にニューラルネットワークをハードウェア上に構築することが可能である。

ニューラルネットワークをデジタル回路ハードウェアに実装する場合の問題として、ニューロン間の結線が問題となる。一般的に、デジタル回路で数値を表す場合、32bitなどのようにマルチビットを用いるが、これがデジタル回路の回路規模の増大を招く原因となる。これに対し、近年ではニューロン間の信号伝達を1bitで実現するパルスニューロンモデルが提案されている[1][2]。しかし、これらのパルスニューロンモデルは、生理学的な実験結果、もしくは試行錯誤的な手法に基づいてパルス演算器が実現されている。工学的に有用なニューラルネットワークを構築するためには、1bitのパルス密度信号に対する各種演算器の設計手法を、信号処理の視点に基づき確立する必要がある。

本論文では、以前提案した $\Delta\Sigma$ 変調に基づくパルスニューラルネットワーク[3]における1bitパルス密度信号間の乗算器を、信号処理の観点に基づき提案する。そして、提案手法は従来の確率的な乗算手法と比較して、精度良く1bit同士の乗算を実現できることを示す。

2 $\Delta\Sigma$ 変調に基づくパルスニューラルネットワーク

本論文でパルスニューロンのモデルとして扱う $\Delta\Sigma$ 変調とは、入力信号を1bitのパルス密度信号に変換する手法である。 $\Delta\Sigma$ 変調された信号の出力はただか1bitであるが、原信号よりも十分に高い周波数でサンプリングするオーバーサンプリングにより、その変換精度を向上することが可能である。我々はこれまでに、 $\Delta\Sigma$ 変調の動作とパルスニューロンの動作の類似性に着目し、 $\Delta\Sigma$ 変調に基づくパルス密度型ニューラルネットワークを提案した。

1次の $\Delta\Sigma$ 変調は図1のように構成することができ、その動作は以下の式で記述できる。

$$w(n) = \frac{1}{1-z^{-1}}(x(n) - \tau y(n-1)) \quad (1)$$

$$y(n) = \begin{cases} 1 & w(n) \geq 0 \\ -1 & w(n) < 0 \end{cases} \quad (2)$$

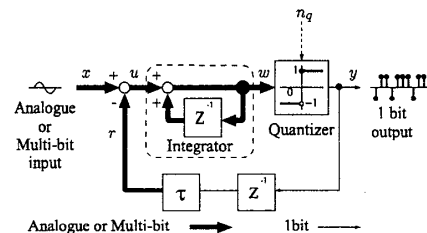
この動作式により $\Delta\Sigma$ 変調された信号の出力 y は以下のように表すことができる。

$$y = \frac{x}{\tau} + (1-z^{-1})n_q \quad (3)$$

x は入力信号、 y は1bitに量子化された $\Delta\Sigma$ 変調の信号出力、 n_q は量子化器で発生する加法的なホワイトノイズ、 τ は $\Delta\Sigma$ 変調器のフィードバックゲインの値である。ここで、右辺の第2項は $\Delta\Sigma$ 変調の量子化ノイズを表している。この項に注目すると、 $1-z^{-1}$ は一種のHPFであり、 $\Delta\Sigma$ 変調の出力の量子化ノイズは高周波数領域に形成され、原信号(低周波数)領域では量子化ノイズが低減されるという性質を持つことになる。これは、ノイズシェーピング効果と呼ばれており、 $\Delta\Sigma$ 変調の重要な特徴の一つである。

従来のパルスニューラルネットワークの多くは、そのパルス生成手法として、M系列等で生成した乱数と入力数値との比較によりパルス信号を生成するRandom Pulse Width Modulation(以後、RPWMと呼ぶ)が使用されてきた。これは、統計的にパルスを扱う場合には優れた性質を持ち合わせているが、量子化ノイズが全周波数領域に均等に分布するため、原信号周波数付近にもノイズが重畳することになる。

一方、 $\Delta\Sigma$ 変調の場合にはニューラルネットワークのように多数の演算を重ねた場合でもノイズシェーピングにより原信号の情報を忠実に伝搬することができる。

図1: $\Delta\Sigma$ 変調器

3 1bit同士の乗算

パルスニューラルネットワークで学習を実現する場合、1bitのパルス密度信号同士の乗算が必要となる場合がある。本論文では、このような乗算器を1bit乗算器(Single-Bit Multiplier: SBM)と呼ぶことにする。

この乗算を実現する手法の一つに、ANDやXNORなどのような簡単な論理演算で実装するモデルが提案されている[1][2]。これらの演算手法は、乗算する互いの信号が無相関で独立であることを仮定し、この場合には理想的な演算結果が得られる。現実的には、演算を重ねるにつれてこのような仮定が成立せず、大規模なニューラルネットワークを構築する場合には、演算誤差が大きくなる。したがって、ノイズに強いデジタル回路で構成しているのにもかかわらず、演算誤差により、大規模なニューラルネットワークの構築が現実的でなくなる。

精度の良いSBMを実現するために、信号処理の観点からこの乗算を考察するとともに、 $\Delta\Sigma$ 変調を用いた1bit乗算器を提案する。

3.1 SBMで発生する演算誤差

サンプリング周波数 $\omega_s = 2\pi f_s$ で標本化する量子化システム(たとえば、 $\Delta\Sigma$ 変調器)に直流信号 x, y を入力するものとする。このとき、角周波数 ω_x, ω_y で発生する出力の量子化誤差をそれぞれ $N_x(\omega_x), N_y(\omega_y)$ とする。 $\omega_x + \omega_y < \omega_s/2$ の場合、量

[†]名古屋大学大学院 工学研究科 電子情報学専攻

[‡]三重大学 工学部 電気電子工学科

子化された2つの値を直接乗算(AND, XNOR)すると、次式を得る。

$$(x + N_x(\omega_x)e^{j\omega_x t})(y + N_y(\omega_y)e^{j\omega_y t}) = xy + N_{xy}(\omega_x, \omega_y) \quad (4)$$

ただし、乗算で発生する誤差を $N_{xy}(\omega_x, \omega_y)$ とする。

$$N_{xy}(\omega_x, \omega_y) = yN_x(\omega_x)e^{j\omega_x t} + xN_y(\omega_y)e^{j\omega_y t} + N_x(\omega_x)N_y(\omega_y)e^{j(\omega_x + \omega_y)t} \quad (5)$$

ここで、 $\omega_x + \omega_y \geq \omega_s/2$ の場合には、第3項のが折り返しノイズが発生することに注意すると、 $N_{xy}(\omega_x, \omega_y)$ は以下の式を得る。

$$N_{xy}(\omega_x, \omega_y) = yN_x(\omega_x)e^{j\omega_x t} + xN_y(\omega_y)e^{j\omega_y t} + N_x(\omega_x)N_y(\omega_y)e^{j(\omega_s - \omega_x - \omega_y)t} \quad (6)$$

(5) (6) 式により求められる $N_{xy}(\omega_x, \omega_y)$ は、乗算における演算誤差となる。各式の第1項および第2項は $\Delta\Sigma$ 変調で発生した量子化誤差成分の定数倍である。これらの成分はノイズシェーピングされた信号であり、原信号付近におけるノイズは低減されているので無視することができる。

しかし、(6) 式の左辺第3項は原信号周波数帯域(低周波帯域)にノイズ成分が加わることを表しており、特に、 $\omega_s - \omega_x - \omega_y = 0$ ($\omega_x = \omega_y = \omega_s/2$: ナイキスト周波数) の場合には直流成分の誤差となって現れ、ひいては学習に悪影響を及ぼすことになる。

3.2 フィルタによるS/N比改善

(5), (6) 式の原信号周波数帯域におけるS/N比を改善するために、 x, y それぞれに以下に記述する $G(z), H(z)$ のフィルタを用いる。

$$G(z) = \frac{K\tau S(z)}{1 + \tau S(z)} \quad (7)$$

$$\text{ただし, } S(z) = \frac{z^{-1}}{1 - z^{-1}}, \tau = f_s/(2\pi f_c), K = \tau$$

$$H'(z) = \frac{1 + z^{-1}}{2}, H(z) = \frac{1}{K} H'(z) G(z) \quad (8)$$

ここで、 f_c はフィルタのカットオフ周波数である。(7) 式は x を K 倍して出力するための一次遅れフィルタである。これにより、(4) 式の左辺第1項の原信号成分の周波数帯域のみを K 倍に増幅することにより、S/N比を改善する。また、(8) 式は移動平均フィルタであり、 y のナイキスト周波数成分の量子化ノイズ $N_y(\omega_s/2)$ を0に抑える伝達関数である。この移動平均フィルタにより、(6) 式の第3項の直流成分付近の誤差を低減することができる。 $H(z)$ には $G(z)$ の伝達関数が従属接続されているが、これは、 x と y の位相を揃えるためであり、この位相の遅れが無視できる場合には省くことができる。

これを実現した回路が図2である。 $1/\tau$ の除算にはフィードバックゲイン τ の $\Delta\Sigma$ 変調器を用いている。

4 実験と考察

従来手法でしばしば用いられる論理ゲート(XNOR)で乗算を実現するSBMと、提案手法のSBMの比較実験を行なった。

実験では、2つの正弦波信号 $A \sin(2\pi f_1 t)$ と $A \sin(2\pi f_2 t)$ それぞれの信号の乗算を行なった。理論的には、この乗算の出力は $0.5A^2 \{\cos(2\pi t(f_1 - f_2)) - \cos(2\pi t(f_1 + f_2))\}$ となり、 $f_1 - f_2$ と $f_1 + f_2$ に信号が現れる。

実験では、2つの正弦波信号を $\Delta\Sigma$ 変調器で1bitのパルス密度信号に変換し、XNORを用いた手法と提案手法のSBMの比較を行なった。実験では、サンプリング周波数を523kHz、

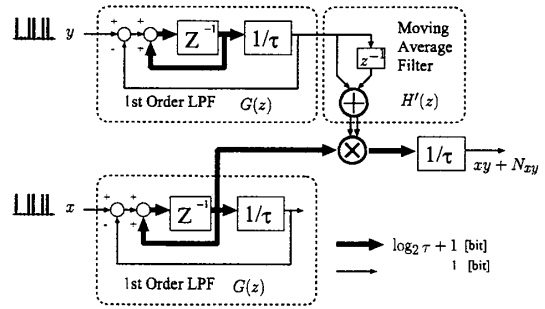


図2: 1bit乗算器の回路

$f_1 = 64 \text{ Hz}, f_2 = 16 \text{ Hz}, A = 0.3$, SBMのフィードバックゲイン $\tau = 128$ としてシミュレーションを行なった。

図3に $\Delta\Sigma$ 変調された信号に対してXNORを用いたSBM(a)と、提案手法のSBM(b)のパワースペクトラムを示す。

XNORによるSBMは $f_1 + f_2, f_1 - f_2$ の成分にスペクトルが現れているが、同時にその周波数領域付近に量子化ノイズ成分も現れている。一方、提案手法では量子化ノイズは低減されており、良好な特性が得られているのが見てとれる。

また、これらの直流成分を求めてみると、XNORのSBM場合で704, 提案手法のSBMの場合で8となり、提案手法は直流成分の誤差も大幅に低減されているのが分かる。ここで完全に直流成分が0とならないのは、 $\Delta\Sigma$ 変調された互いの量子化ノイズ成分のうち、同じ周波数で非直交な成分同士が乗算され、その影響が直流成分に現れるものと考えられる。

5 まとめ

本論文では、 $\Delta\Sigma$ 変調に基づくパルスニューラルネットワークに対して有効なSBMを提案し、シミュレーションにおいて本手法の有効性を検証した。提案手法は、従来のXNORの演算に比べて精度良く乗算を実現し、特に学習に悪影響を及ぼす直流成分の誤差を大幅に低減できることを示した。

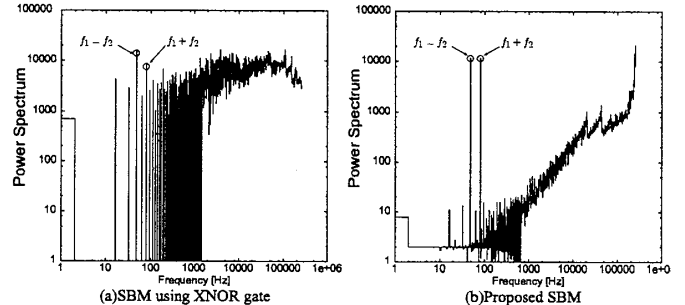


図3: SBMの出力のパワースペクトラム

参考文献

- [1] 江口裕俊, 古田俊之, 堀口浩幸, 榎木杉高: “学習機能をもつパルス密度型ニューロンモデルとそのハードウェア”, 信学論 C-II, Vol. J74, No. 5, pp. 369-376, 1991
- [2] Young-Chul Kim, Michael A. Shanbaltt: “Random Noise Effects in Pulse-Mode Digital Multilayer Neural Networks”, IEEE Transactions on Neural Networks, Vol. 6, No. 1, 1995
- [3] 村橋 善光, 道木 慎二, 大熊 繁: “GHA 学習則による $\Delta\Sigma$ 変調に基づくパルス密度型ニューラルネットワークのハードウェア実装”, 信学技報 NC, Vol. 101, No. 737, pp. 129-136, 2002