República Federativa do Brasil
Ministério da Indústria, Comércio Exterior e Serviços
Instituto Nacional da Propriedade Industrial

(21) BR 102015004646-4 A2

(22) Data do Depósito: 02/03/2015

(43) Data da Publicação: 06/09/2016

(54) Título: GERADOR, CONVERSOR E CODIFICADOR DE POTENCIAIS DE AÇÃO ARTIFICIAIS PARA REDES NEURAIS PULSADAS

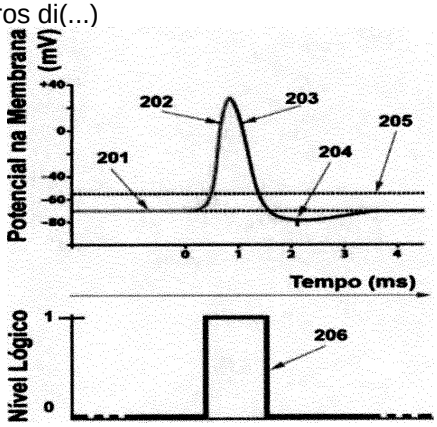
(51) Int. Cl.: G06N 3/02; G06N 3/04; G06N 3/10

(52) CPC: G06N 3/02; G06N 3/049; G06N 3/10

(73) Titular(es): UNIVERSIDADE FEDERAL DE PERNAMBUCO - UFPE

(72) Inventor(es): JOÃO HENRIQUE RANHEL RIBEIRO, JOSÉ RODRIGUES DE OLIVEIRA NETO, JOÃO PAULO CERQUINHO CAJUEIRO

(57) Resumo: Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas descreve um sistema eletrônico que gera potenciais de ação artificiais (spikes) a partir de parâmetros configuráveis e que podem ser alterados em regime contínuo e em tempo real, converte dados digitais em spikes, gerando frequências de disparo em neurônios artificialmente simulados, "codifica" os spikes gerados em três formas de trens de pulsos distintas: pela média de disparo dos neurônios, conhecido em neurociência como "rate coding"; pelo tempo de retardo ou ordem de chegada dos spikes em relação a uma referência, chamado "temporal coding"; ou ainda pelo número de neurônios ativos num intervalo de tempo, conhecido por "population coding", e descreve um sistema eletrônico que serve de interface entre um conjunto de dados de entrada (o mundo físico, sistemas digitais, computadores, etc) e as redes neurais pulsadas (artificiais ou naturais), e faz a conversão de grandezas físicas e/ou informação para um formato apropriado para ser tratado pelas redes neurais artificiais pulsadas e comunica-se com outros di(...)



Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas

[1] A presente invenção descreve um sistema eletrônico que gera potenciais de ação artificiais (*spikes*) a partir de parâmetros configuráveis e que podem ser alterados em regime contínuo. Spikes são pulsos que imitam sinais de saída dos neurônios naturais. A presente invenção, continuamente e em tempo real, converte dados digitais em spikes, gerando frequências de disparo em neurônios artificialmente simulados. Além disso, a invenção "codifica" os spikes gerados em três formas de trens de pulsos distintas: pela média de disparo dos neurônios, conhecido em neurociência como "*rate coding*"; pelo tempo de retardo ou ordem de chegada dos spikes em relação a uma referência, chamado "*temporal coding*"; ou ainda pelo número de neurônios ativos num intervalo de tempo, também conhecido por "*population coding*". A invenção, abreviada pela sigla GCCst, descreve um sistema eletrônico que serve de interface entre um conjunto de dados de entrada (o mundo físico, sistemas digitais, computadores, etc) e as redes neurais pulsadas (artificiais ou naturais). O GCCst faz a conversão de grandezas físicas e/ou informação para um formato apropriado para ser tratado pelas redes neurais artificiais pulsadas. O sistema eletrônico GCCst comunica-se com outros dispositivos pela forma serial ou paralela, tanto na entrada quanto na saída. A invenção se situa nos campos da engenharia elétrica e da ciência da computação

Antecedentes da Invenção

[2] Redes neurais pulsadas (*Spiking Neural Networks - SNN*) é uma área de conhecimento que se encaixa na grande área de estudo da Inteligência Artificial. Diferente de gerações de redes neurais artificiais antecessoras, as

redes pulsadas levam em conta a temporização da informação, imitando mais proximamente os neurônios biológicos. No mundo orgânico, neurônios geram potenciais de ação, que são pulsos elétricos que duram em média um milésimo de segundo (1 *ms*). O potencial de ação é uma forma econômica de codificar informação porque o custo em termos de energia é pequeno. Outro benefício é que neurônios podem compartilhar o tempo em que não estão ativos e executar várias tarefas. Redes neurais pulsadas usam essas vantagens e processam informações baseadas em eventos, em disparos de neurônios artificias que são chamados *spikes*.

[3] Um grande problema das redes pulsadas é que dados apresentados à rede devem ser transformados em *spikes*, ou em série de spikes, conhecidos por trens de pulsos (*spike train*).

[4] Além disso, para serem usados em tempo real, redes neurais precisam ser computadas em sistemas digitais especializados, em computação paralela implementada em FPGAs e SoC. Para alimentar tais sistemas é necessário um conversor em tempo real que converta grandezas do mundo físico e/ou dados de computadores em *spike train* rodando em tempo real. Este é o objetivo primeiro do GCCst, objeto desta patente.

[5] Quando se cria um sistema que converte algo em outra coisa, como o GCCst que converte valores digitais em spikes, está implícito que alguma forma de 'codificação' é utilizada. Na literatura de neurociência três grandes grupos de codificação são descritas: a média de disparo, a codificação temporal, e a codificação por meio da população de células ativas. O GCCst contempla as três formas de codificação. Em outras palavras, um usuário é capaz de

programar o tipo de codificação que o sistema eletrônico descrito nesta patente deverá gerar em resposta aos estímulos ou dados que recebe.

[6] Entre os documentos de patente que descrevem objetos relacionados ao campo da invenção cita-se o documento de patente **US2014317035** - Apparatus and methods for event-based communication in a spiking neuron networks; o documento de patente **WO2014138202** - Generic method for designing spike-timing dependent plasticity (stdp) curves; o documento de patente **US2014074761** - Dynamical event neuron and synapse models for learning spiking neural networks; o documento de patente **US2013204819** - Methods and apparatus for spiking neural computation; o document de patente **WO2007071070** - Spatio-temporal pattern recognition using a spiking neural network and processing thereof on a portable and/or distributed computer

[7] Nenhum dos documentos relacionados reivindica a forma de geração, de conversão ou de codificação descritas nesta invenção, nem os processos e princípios operacionais descritos para o GCCst.

[8] Também não foi encontrado patente que reivindique a forma de uso e os propósitos de utilização explicitados neste documento.

Sumário da Invenção

[9] A presente invenção tem por objetivo resolver problemas constantes no estado da técnica no sentido de gerar, em tempo real e no formato adequado, os potenciais de ações artificiais (*spikes*) para excitar redes neurais que funcionam por meio de processamento de impulsos (spikes), chamadas aqui de redes neurais pulsadas. Além de gerar os *spikes* de forma apropriada, esta invenção cuida de dois outros aspectos importantes, que são a conversão de

dados digitalizados para trens de pulsos, além de codificar os trens de pulsos dentro dos formatos descritos na literatura de neurociência.

[10] É um primeiro objeto da presente invenção um sistema digital implementado em circuitaria digital que *gera* impulsos (spikes) imitando potenciais de ação considerando as temporizações envolvidas nesses sinais.

[11] É um segundo objeto da presente invenção um sistema digital que realiza a *conversão* de grandezas (dados digitais) e parâmetros configuráveis em trem de pulso (*spike train*) apropriado, seguindo restrições descritas na literatura científica de neurociência e de neurocomputação.

[12] É um terceiro objeto da presente invenção um sistema digital que *codifica* os trens de pulsos de acordo com três grandes categorias descritas em neurociência e neurocomputação: (i) codificação pela média de disparos dos neurônios dentro de uma janela de tempo, também conhecida por taxa média de disparos (*rate coding*); (ii) codificação pelo tempo de disparo entre neurônios e um sistema de referências, também conhecida por codificação temporal (*temporal coding*), contemplando neste ao menos duas subclasses que são codificação pela fase *phase-coding* e codificação pela ordem de chegada, ou *ranking-order coding*; (iii) codificação pela população ativa de neurônios, na qual o que conta é o número de neurônios ativos (população ativa) dentro de um intervalo de tempo. A presente invenção é capaz, portanto, não apenas de gerar e converter sinais de estímulos externos ao circuito, como também é capaz de codificá-los de acordo com parâmetros fornecidos pelo usuário.

[13] É, adicionalmente, um quarto objeto da presente invenção o uso desse sistema eletrônico digital como interface entre o mundo físico e as redes neurais pulsadas com a finalidade de:

- transformar grandezas físicas, convertidas para números digitais, em trens de pulsos para redes neurais artificiais pulsadas, ou para redes pulsadas biológicas (desde que devidamente adaptados para os potenciais envolvidos nos neurônios biológicos);

- servir de gerador de trens de pulsos (*spike train*) para redes neurais artificiais;

- alimentar redes neurais com informações de bancos de dados digitais, convertendo dados em trens de pulsos que deverão excitar redes neurais artificiais pulsadas;

- servir de codificador de trens de pulsos para excitar redes neurais artificiais ou naturais usando os tipos de código descritos em neurociência; em que principal característica é que o fluxo de pulsos nos trens de pulsos é *contínuo* e em *tempo real*. Dessa forma, este sistema digital é um gerador, conversor e codificador que garante a alimentação do fluxo contínuo de potenciais de ação para as redes neurais artificiais que também operam em tempo real e contínuo, imitando os sistemas nervosos biológicos.

[14] Ainda, o conceito inventivo comum a todos os contextos de proteção reivindicados se refere à geração de impulsos no formato apropriado, na capacidade de converter uma forma de dados em spikes especificamente adaptados para redes neurais pulsadas, a capacidade de codificação de acordo com os parâmetros definidos pelo usuário, e a generalização desse sistema digital que opera como uma interface genérica entre o mundo físico e/ou bancos de dados e as redes neurais pulsadas.

Breve Descrição das Figuras

[15] Com o intuito de melhor definir e esclarecer o conteúdo do presente pedido de patente, são apresentadas as seguintes figuras:

[16] A figura 1 mostra um esboço de três neurônios artificiais pré-sinápticos [101, 102, 103] interconectados a um neurônio artificial pós-sináptico [104] por meio das conexões sinápticas [105]. O neurônio [104] deve ser entendido como aquele que recebe *spikes* dos pré-sinápticos e executa uma operação de computação neural, gerando como saída um *spike*.

[17] A figura 2 ilustra como é o formato de um potencial de ação genérico (a saída) em um neurônio real [202, 203, 204], e como esse sinal pode ser emulado de forma simplificada em sistemas digitais [206].

[18] A figura 3 ilustra o contexto em que esta invenção se aplica. O mundo externo é representado pela nuvem [301] e nela podemos ter sistemas digitais conversores de grandezas físicas para números, ou podemos ter bancos de dados digitais (nenhum dos dois representados) que serão introduzidos no GCCst por interfaces seriais ou paralelas [302]. A saída do GCCst pode ser por sinais paralelos (cada fio representando um neurônio [304]) ou por sinais serial, por meio de um canal de comunicação [305] com outros sistemas digitais.

[19] A figura 4 mostra um diagrama de blocos do sistema digital objeto desta patente (GCCst). O sistema em si se encontra representado dentro do retângulo tracejado [400]. Externo ao retângulo se encontram as interfaces de entrada [401] e de saída [420] que fazem a comunicação com computadores e outros sistemas digitais. Essas duas interfaces [401, 420] não são objeto desta patente porque é possível implantar o módulo de comunicação com padrões existentes no mercado (I2C, SPI, USB, etc), deixando o sistema mais flexível.

Descrição detalhada da invenção

[20] Como foi dito, a figura 1 ilustra como neurônios pré-sinápticos [101, 102, 103] são interconectados a um neurônio pós-sináptico [104] que processa os sinais que advém dos anteriores. Tomando-se o neurônio [104] como referência, as conexões entre [101, 102, 103] e o neurônio referência são ditas pré-sinápticas [105], uma vez que ocorre entre neurônios que fornecem informação (*spikes*) e a árvore dendrítica do neurônio [104], que neste caso é quem executa a operação computacional. O neurônio [104] também gera como saída um *spike* que será conduzido pelo seu axônio [106] até as conexões pós-sinápticas [107], que são as conexões que o neurônio referência faz em sua saída. A computação realizada por [104] é: somar ponderadamente os spikes que chegam dos neurônios pré-sinápticos e decidir se deve ou não gerar um spike como saída. A soma que o neurônio [104] realiza é uma soma ponderada, porque cada conexão [105] tem um peso, chamado "peso sináptico".

[21] Em um neurônio natural o potencial de equilíbrio (repouso) da membrana [201] gira em torno de -65 mV em relação à referência que é o meio exterior à célula. Neurônios podem sofrer variações neste potencial da membrana no sentido de tornar mais positivo o valor da tensão interna da célula. Quando isso acontece e a tensão da membrana ultrapassa certo limiar [205] uma sucessão de eventos é disparada no interior da célula. Proteínas sensíveis à variação de tensão, presas à membrana da célula, se abrem e um processo de realimentação positiva faz com que canais proteicos se abram, permitindo a entrada de íons com cargas positivas, levando a tensão da célula para um valor positivo de ~+30 mV, como indicado em [202]. Logo em seguida um processo de recuperação da célula acontece, alguns canais proteicos se fecham e a fuga de

íons positivos de dentro da célula faz o neurônio voltar a ter seu potencial de membrana retornando a valores negativos, como indicado em [203]. Na sequência, o neurônio entra em um estado de recuperação do valor da membrana porque no processo de repolarização a tensão da membrana pode ultrapassar o valor de -65 mV de repouso, indicado em [204]. Os períodos indicados em [203] e [204] são chamados refratários, em que o neurônio está se recuperando do disparo. Note-se que o tempo de duração do potencial de ação [202 e 203] é mais ou menos fixo em 1 ms. Dessa forma, admite-se que o potencial de ação em si não contém outra informação exceto se o neurônio (1) disparou ou não (0). Por esse motivo, é possível reproduzir o sinal do spike digitalmente, como indicado em [206]. O período refratário dos neurônios artificiais é simulado pelo modelo matemático do neurônio, e não é necessário ser reproduzido no potencial de ação artificial (*spike*). Ou seja, o período de refração fica ao encargo das redes neurais pulsadas, e não precisa ser reproduzido no circuito que gera o pulso, como o GCCst.

[22] O contexto desta invenção é ilustrado pela Fig.3, pela caixa [303] que internamente converte os dados introduzidos via interface de entrada [302] em saídas. As saídas podem ser fios (sinais) paralelos [304] que emulam o funcionamento de axônios dos neurônios artificiais. Saídas também podem ser feitas por meio de interfaces seriais, representadas pela seta [305]. Assim, as saídas do circuito descrito nesta invenção podem excitar diretamente redes neurais pulsadas em circuito de hardware, representada pela caixa [306], mas também podem ser comunicadas a um sistema de computação digital que simula redes neurais pulsadas, via interface de saída [305]. Como foi dito, a forma da interface de entrada de dados, e a forma da interface de saída de spikes do

circuito não é objeto nesta patente. Assim, mudanças no sistema de entrada ou saída de dados neste sistema digital não configura mudança relevante que descaracterize esta patente. Importante são os circuitos, a lógica, a estrutura e a organização dos circuitos internos de [303], que caracterizam esta invenção, descritos na sequência.

[23] A figura 4 ilustra o diagrama de blocos do GCCst, e a área tracejada da figura equivale ao bloco [303] representado na figura 3. Qualquer que seja o padrão de interface escolhido, a comunicação digital externa é recebida pelo *módulo de recepção* [403], que é uma máquina de estado que identifica quatro tipos de entrada: (i) endereço e (ii) dado da *memória de configuração* [404]; e ainda (iii) endereço e (iv) dados da *memória de limiar* [405]. O módulo de recepção é portanto encarregado de identificar se os sinais de entrada são endereços das memórias [404] ou [405], apontar para um dos endereços nas memórias, e em seguida direcionar dados de entrada para uma das posições dessas duas memórias. O módulo de recepção [403] 'conversa' com as memórias [404, 405] através do barramento de dados e endereços [406].

[24] Além dos dois blocos de memória [404, 405] previamente descritos, há um terceiro módulo de memória [408] que funciona como um conjunto de contadores, denominado *memória contadora*. Os três blocos de memória estão ligados a um *módulo controlador* [412] respectivamente pelos barramentos [410, 411, 409].

[25] Há ainda dois sistemas geradores de *clock* interno. Um gera 500 Hz [414] (com período de 1 ms em nível lógico alto e 1 ms em nível baixo), e o outro gera 128 MHz [413], com *duty cycle* de 50%. Este último pode variar para

além ou aquém de 128 *MHz* sem mudança da característica de funcionamento desse sistema. Ambos alimentam o módulo controlador [412].

[26] O *módulo controlador* [412] endereça e faz leitura nos três blocos de memória mencionados, e a escrita nessas memórias é feita pelo *módulo de recepção* [403]. Os blocos de memórias [404, 405, 408] são implantados em *dual-port RAM* para acelerar o processo de leitura e escrita simultâneas, embora isso não seja requisito que descaracterize a patente. Quando o módulo controlador [412] endereça uma posição na memória de configuração [404] ele também endereça a mesma posição na memória de limiar [405] e na memória contadora [408].

[27] O módulo controlador [412] é responsável por fazer um pareamento dos endereços dos três blocos de memória e por ler dados das três memórias em endereços iguais. Cada endereço ***xx*** nas três memórias [404, 405, 408] contém dados referentes ao neurônio ***xx***, usados pelo módulo controlador [412] para decidir se o neurônio ***xx*** deve gerar um spike ou não.

[28] A cada subida do clock de 500 *Hz*, o módulo controlador [412] aponta para o primeiro endereço de memória e repete até o último endereço possível para as memórias. A cada endereço gerado, o módulo [412] compara se o conteúdo armazenado no endereço ***xx*** da memória contadora [408] é igual ao conteúdo armazenado no endereço ***xx*** na memória de configuração [404]:

- Se os conteúdos forem iguais, o módulo controlador [412] usa o mesmo endereço ***xx*** para apontar a *memória de spikes* [415] e ajusta o valor desta posição de memória para o nível lógico '1' (alto). Isso significa que o

neurônio representado na posição ***xx*** da memória de spikes [415] disparou um potencial de ação artificial (spike).

- Por outro lado, se os conteúdos comparados não forem iguais, o módulo controlador [412] zera a memória de spikes [415] na posição ***xx***, ou seja, esse neurônio não disparou um spike.

[29] Como foi dito, na sequência o módulo controlador [412] incrementa um endereço de memória (***xx+1***) e executa o processo de comparação para outro neurônio, repetindo o processo até o último endereço de memória. Este processo é rápido porque é realizado na frequência de 128 MHz. Dessa forma, o módulo controlador [412] termina seu trabalho de comparação, decisão e atualização rapidamente, e entra em estado de espera da descida do clock do sinal de clock de 500 Hz.

[30] Na descida do clock de 500 Hz o módulo controlador [412] repete o endereçamento de todos os neurônios e zera cada posição da memória de spikes [415]. Isso significa que se algum neurônio dessa memória fora disparado durante o processo de subida do clock de 500 Hz ele ficará em nível lógico '1' apenas por um milésimo de segundo, porque na descida de clock de 500 Hz (~1 ms depois da subida) todos os spikes ativados serão zerados. Também na descida de clock o módulo controlador [412] executa duas operações na memória contadora [408]: se o neurônio ***xx*** fora disparado, o módulo controlador [412] zera o conteúdo do endereço ***xx*** na memória contadora [408]; contudo, se o neurônio ***xx*** não disparou na última subida de clock, o módulo controlador apenas incrementa em uma unidade ao valor armazenado no endereço ***xx*** da memória contadora [408].

[31] Este é o modo de operação default do sistema eletrônico GCCst. Esse procedimento faz com que o sistema rode livremente gerando uma frequência de disparo independente para cada neurônio, que depende do valor gravado na memória de configuração [404]. Cada spike tem a largura do pulso sempre constante e igual a 1 *ms*, como os sistemas biológicos. Este modo de operação livre gera o que em neurociência é chamado de "*rate coding*", porque cada neurônio dispara numa frequência (uma taxa de disparo) que depende apenas do valor gravado na memória de configuração [404].

[32] Existe porém um valor que pode ser gravado em um registrador específico da memória de configuração [404] que modifica toda a regra de operação do sistema. Este conteúdo serve para gerar o código temporal, ou "*temporal coding*". Quando o valor deste registrador for ajustado para um número qualquer, o módulo controlador [412] testa este conteúdo com um conteúdo de mesmo endereço na memória contadora [408].

- Caso os valores comparados sejam iguais, o sistema zera toda a memória contadora [408], gerando assim um reset síncrono em todos os neurônios. Todos os contadores da memória [408] recomeçam do zero ao mesmo tempo. Dessa forma, a partir desse 'reset' a rede neural artificial pulsada pode observar qual neurônio disparará primeiro, qual disparará em seguida, e assim por diante. Em neurociência, essa forma de codificação é chamada "*ranking order*" e é uma das formas de codificação temporal. Outra interpretação que a rede neural artificial pulsada pode ter é que um grupo de neurônios, ou um neurônio específico pode estar codificando algo pelo atraso que carrega em relação ao reset, chamado em neurociência de "*latency coding*"; que ao final é o mesmo fenômeno também descrito como "*phase coding*", em que um ou um

grupo de neurônios codifica algo pelo ângulo de fase em relação a um sinal cíclico.

- Caso o valor armazenado no registrador especial da memória [408] seja zero, o módulo controlador [412] deixa de fazer a comparação, e o sistema opera no modo normal, como descrito anteriormente.

[33] Para operar nos dois modos de codificação anteriormente citados (*rate coding* e *temporal coding*) toda a memória de limiar [405] deve estar zerada. Quando algum valor é carregado na memória de limiar [405] cada neurônio passa a ser testado para verificar se o valor na memória de configuração [404] está abaixo do valor gravado na memória de limiar [405]; isso para cada endereço de memória, ou para todos os neurônios. Por exemplo, no endereço ***xx*** da memória de configuração [404] foi gravado o número "64", que gera uma certa frequência de disparo do neurônio ***xx***. Se no endereço ***xx*** da memória de limiar [405] for gravado o valor "0", o sistema opera nos modos descritos anteriormente. Se no endereço ***xx*** da memória de limiar [405] for gravado o valor "68" então o neurônio ***xx*** na memória de spikes [415] continua disparando normalmente na frequência programada em [404], ou seja, em função do valor 64, porque este é menor que 68. Contudo, se o valor da memória de configuração [404] mudar para "69", então o neurônio ***xx*** na memória de spikes [415] para de disparar, permanecendo zerado.

[34] Temos que lembrar que, em virtude de excitação externa, o endereço ***xx*** da memória de configuração [404] pode ser alterado a qualquer tempo, indicando que o neurônio ***xx*** deve alterar a frequência de disparo. Assim, caso o valor salvo no endereço ***xx*** da memória [404] diminua, a frequência de

spikes aumenta. O contrário também ocorre, se o valor salvo aumentar, a frequência de disparo do neurônio diminui. Portanto, se o valor alterado no endereço ***xx*** de [404] continuar abaixo do valor de limiar (no exemplo, "68") o neurônio continuará disparando, mas se o valor ultrapassar o valor da memória de limiar [405] (ultrapassar "68") então o neurônio ***xx*** deixa de disparar e fica silente.

[35] O efeito do processo descrito acima é que isso permite que uma população de neurônios seja programada a operar a partir de um dado limiar. Ajustando uma população de neurônios para valores de limiar na memória [405] que variam de Y até Z, tal grupo de neurônios começa a operar quando excitado em limiares diferentes que variam desde Y até Z; ou seja, os valores gravados na memória de configuração [404] determinam a frequência que o neurônio vai disparar, e os valores gravados na memória de limiar [405] determinam quando um neurônio começa a disparar. A rede neural artificial pulsada que recebe a saída do GCCst é que vai interpretar esse grupo de neurônios como uma população que cresce ou decresce na medida que mais ou menos neurônios entram na região abaixo dos limiares e começam a disparar. Em neurociência este tipo de codificação é chamado de *population code*, e geralmente reflete se a excitação externa é mais ou menos intensa.

[36] O módulo de saída [417] faz a interface entre a memória de spikes [415] e o mundo exterior. Para melhorar rendimento, a memória de spikes [415] também é implementada em dual-port RAM. Apenas o módulo de saída [417] tem acesso à memória de spikes [415], e ele o faz via um barramento dedicado [416]. Quando um dispositivo de interface externa [420] quer acessar o conteúdo da memória de spikes ele o faz via dois conjuntos de sinais disponibilizados a

partir do módulo de saída [417]. É possível implementar um conjunto de fios paralelos [418] simulando os axônios dos neurônios diretamente a partir do módulo de saída [417]. Facilmente, todo o sistema descrito nesta patente pode implementar 1023 neurônios em um circuito FPGA, utilizando menos de 5% de sua área. Contudo, um FPGA não possui 1023 pinos de saída disponíveis para o usuário. Dessa forma, é provável que o GCCst envie os dados da matriz de spikes [415] para outro sistema computadorizado via uma interface serial, que será implementada por um barramento [419] que se comunica com o módulo de saída [417]. O padrão escolhido para o barramento serial [419] não está definido, podendo ser escolhido qualquer dos padrões usuais de interface serial.

[37] Os objetos da invenção serão imediatamente valorizados pelos versados na arte e pelas empresas com interesses no segmento, e serão descritos em detalhes suficientes para sua reprodução na descrição a seguir.

[38] Uso do Aparato Gerador, Conversos e Codificador de Spike Train (GCCst)

[39] Operacionalmente, o GCCst deve funcionar acoplado a dois sistemas digitais (microcontroladores, computadores), um que programe suas memórias de entrada (Fig. 4: [404, 405]) e outro que receba os sinais de spikes gravados na memória de saída [415]. Essa saída pode ser o disparo de neurônios diretamente em um fio (sinal de saída físico) ou um vetor que indica quais neurônios que dispararam em um instante 't', cuja resolução temporal de 't' é de 1 ms.

[40] Um sistema digital (ex: microcontrolador - μC) deve gravar na memória de configuração [404] um número digital que indica a frequência em spikes de um dado neurônio. Isso é feito por intermédio do módulo de recepção

[403]. Como exemplo, se um μC está monitorando uma temperatura ele deverá obter um valor analógico de um sensor e converter para um número digital. Via interface de entrada este μC informa ao sistema GCCst o valor obtido na medida, gravando no endereço de memória de configuração [404] do neurônio xx o valor da grandeza medida. Feito isso, o sistema GCCst se encarrega de gerar a frequência correspondente para aquele neurônio programado.

[41] Adicionalmente, este μC pode informar um valor de limiar para ser gravado na memória de limiar [405]. Quando a temperatura sobe ou desce, o μC altera os valores nas memórias [404, 405] referentes ao neurônio xx e o GCCst cuida de adaptar e gerar novos padrões para aquele neurônio. Quando a memória de limiar [405] tem algum valor, o neurônio é impedido de disparar em frequências muito baixas. Um μC relativamente simples consegue controlar grande número de neurônios porque tem apenas que reconfigurar o GCCst periodicamente e alterar, em qualquer ordem, os valores memorizados na memória de configuração e limiar [404, 405].

[42] A saída do GCCst é contínua, ou seja, se um neurônio é programado o sistema GCCst começa a gerar a saída de spike correspondente àquele neurônio. Se a saída for em um fio físico da FPGA ou SoC, então aquele fio se comportará como um axônio para um dado neurônio. Nesse caso, colocando-se um osciloscópio ou analisador lógico nesta saída pode-se observar imediatamente a saída desse sinal.

[43] Por outro lado, caso se opte por uma saída serial, um dispositivo externo poderá ler serialmente o conteúdo da memória de spikes [415] do GCCst por intermédio do módulo de saída [417]. Este dispositivo externo pode ler dados

específicos relativo a um ou a um grupo de neurônios, ou poderá buscar todo o conteúdo da memória de spikes [415]. Também é possível programar um pequeno módulo ou inserir um μC que faça a intermediação entre os dados da memória [415] e um computador, via USB ou outro canal de comunicação.

[44] Assim, resumindo o modo de operação desta invenção, o GCCst é um sistema intermediário entre dois sistemas digitais. A entrada é um sistema sensorial ou um conjunto de dados que constantemente escreve nas memórias de configuração e limiar [404, 405] do GCCst os valores correspondentes à frequência de disparo de cada neurônio. O GCCst, objeto dessa invenção, gera spikes na forma de um pulso de 1 ms de duração, e frequência configurada em [404]. O GCCst coloca o resultado (os spikes) na memória de spikes [415], que pode assim ser acessada por outro sistema digital, que pode ser um computador ou uma rede neural artificial pulsada em hardware. Tão ou mais importante que gerar os spikes, o GCCst é capaz de codificar os spikes gerados em uma das três grandes classes de codificação relatadas na literatura de neurociência.

Reivindicações

1. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** revela circuito eletrônico digital implementado em lógica digital, capaz de gerar potenciais de ação artificiais (spikes) para excitar redes neurais artificiais, **caracterizado** por gerar pulsos eletrônicos equiparáveis ao potencial de ação dos neurônios, utilizados para excitar redes neurais pulsadas.

2. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** é sistema eletrônico conversor de grandezas físicas ou dados digitais em spikes, de acordo com a reivindicação 1, **caracterizado** por receber parâmetros (números) digitais e converter automaticamente para trens de pulsos (*spike train*);

3. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** é sistema eletrônico capaz de codificar os trens de pulsos de acordo com a reivindicação 1 ou 2, **caracterizado** por gerar padrões de trens de pulsos codificadas de três formas diferentes: média de disparo (*rate coding*), tempo de disparo de cada neurônio com relação a uma referência sincronizadora (*temporal coding*), e ainda pelo número de neurônios ativos em uma população (*population coding*).

4. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** faz interface do mundo físico e de bancos de dados com redes neurais artificiais pulsadas, de acordo com qualquer uma das reivindicações 1 a 3, que opera em tempo real, **caracterizado** por receber parâmetros e grandezas de entrada e gerar pulsos digitais em fluxo contínuo e em tempo real, e emular potenciais de ação de neurônios naturais.

5. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** é sistema eletrônico digital gerador, conversor e codificador de dados ou informação em spikes, de acordo com as reivindicações de 1 a 4, operando independente da forma usada para suprir informações ao sistema eletrônico, **caracterizado** pelo fato de o circuito poder receber dados seriais por meio de interfaces digitais seriais padrões, e também poder receber vetores de dados paralelos por meio de interfaces paralelas.

6. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** é sistema eletrônico digital gerador, conversor e codificador de dados ou informação em spikes, de acordo com as reivindicações de 1 a 4, operando independente da forma usada para extrair informações do sistema eletrônico, **caracterizado** pelo fato de o circuito poder enviar dados seriais, por meio de interfaces digitais seriais padrões, para outros dispositivos, tipo computadores e redes neurais artificiais, e também poder gerar spikes em uma série de fios (sinais) paralelos que mimetizam axônios, em que cada fio funciona como se fosse o axônio de um neurônio independente.

7. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** é sistema eletrônico digital genérico que serve de interface entre bancos de dados digitais e/ou mundo físico e redes neurais pulsadas, fazendo uma ponte no fornecimento de informações convertidas e codificadas para spikes, de acordo com as reivindicações de 1 a 6, **caracterizado** pelo fato de converter e transformar grandezas físicas ou dados digitais em códigos de disparos neurais, sendo independente do canal usado para introduzir ou extrair spikes desse sistema, podendo ser utilizado como interface entre vários sistemas digitais que fornecem dados e que computam na

forma de redes neurais pulsadas, inclusive em hardware, como FPGAs (*Field Programmable Gate Array*) e/ou SoC (*System on Chips*).

8. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** compreende processo de geração de spike, **caracterizado** pelo circuito de comparação de valores do banco memória de configuração e de limiar com valores do banco de memória contadora (contadores), associados a um sinal de clock de 500 Hz e a um bloco lógico de decisão (módulo controlador), criando um pulso de duração de 1 ms na igualdade dos valores das memórias.

9. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** compreende processo de codificação pela taxa média de disparo (*rate coding*), de acordo com as reivindicações de 1 a 8, **caracterizado** pelo circuito de decisão que aplica 'RESET' na memória contadora (zera a memória contadora) quando o valor dessa atingir o valor programado na memória de configuração.

10. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** compreende processo de codificação temporal (*temporal coding*), de acordo com as reivindicações de 1 a 9, **caracterizado** pelo circuito de decisão (módulo controlador) que aplica 'RESET síncrono' em todo grupo da memória contadora, criando um processo de sincronismo que serve de referência temporal para a codificação de todos os neurônios envolvidos.

11. **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas** compreende processo de codificação por população (*population coding*), de acordo com as reivindicações de 1 a 10

caracterizado pelo circuito que decide se cada neurônio individualmente deve fazer parte da população de neurônios ativa em um dado intervalo de tempo, baseado no parâmetro gravado na memória de limiar de cada neurônio.

12 **Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas**, caracterizado por compreender a comunicação digital externa ser recebida pelo *módulo de recepção* [403], e o *módulo de recepção* [403] identificar quatro tipos de entrada: (i) endereço e (ii) dado da *memória de configuração* [404]; e ainda (iii) endereço e (iv) dados da *memória de limiar* [405], e o *módulo de recepção* [403] ser encarregado de identificar se os sinais de entrada são endereços das memórias [404] ou [405], apontar para um dos endereços nas memórias, e em seguida direcionar dados de entrada para uma das posições dessas duas memórias, e o *módulo de recepção* [403] 'conversar' com as memórias [404, 405] através de barramento de dados e endereços [406], e possuir um terceiro módulo de memória [408] que funciona como um conjunto de contadores, denominado *memória contadora*, e os três blocos de memória estarem ligados a um *módulo controlador* [412] respectivamente por barramentos [410, 411, 409], e o *módulo controlador* [412] ser alimentado por dois sistemas geradores de *clock* interno, sendo que um gera 500 *Hz* [414], com período de 1 *ms* em nível lógico alto e 1 *ms* em nível baixo, e o outro gera 128 *MHz* [413], com *duty cycle* de 50%, e o *módulo controlador* [412] endereçar e fazer a leitura nos três blocos de memória[404, 405, 408] mencionados, e a escrita nessas memórias ser feita pelo *módulo de recepção* [403], e os blocos de memórias [404, 405, 408] serem implantados em *dual-port RAM, e* o módulo controlador [412] ao endereçar uma posição na memória de configuração [404], ele também vai endereçar a mesma posição na memória de

limiar [405] e na memória contadora [408], e o módulo controlador [412] fazer um pareamento dos endereços dos três blocos de memória e ler dados das três memórias em endereços iguais.

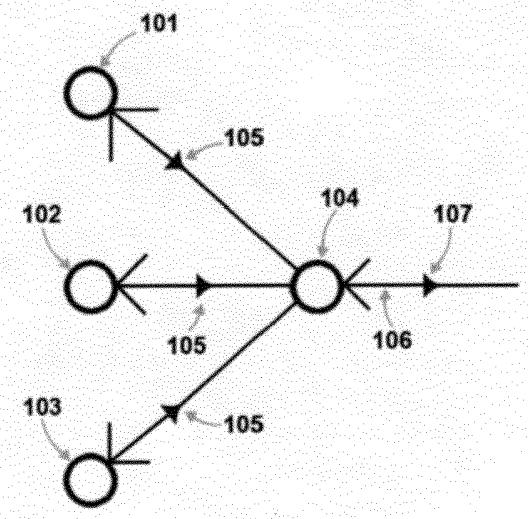


Figura 1

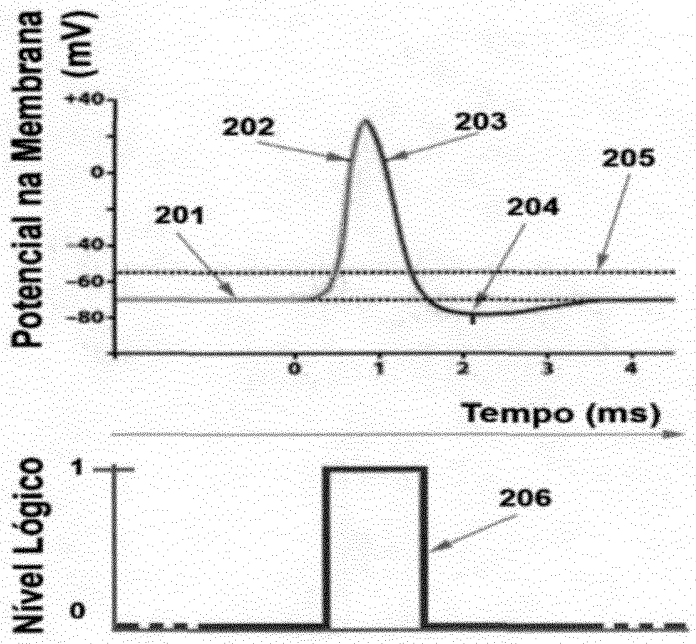


Figura 2

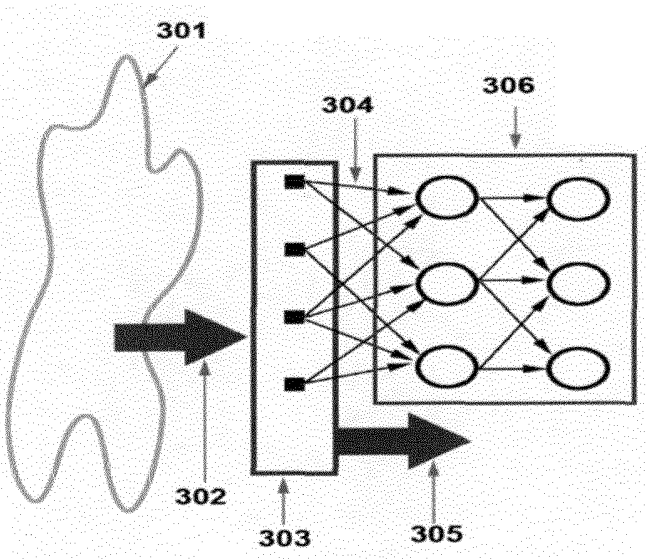


Figura 3

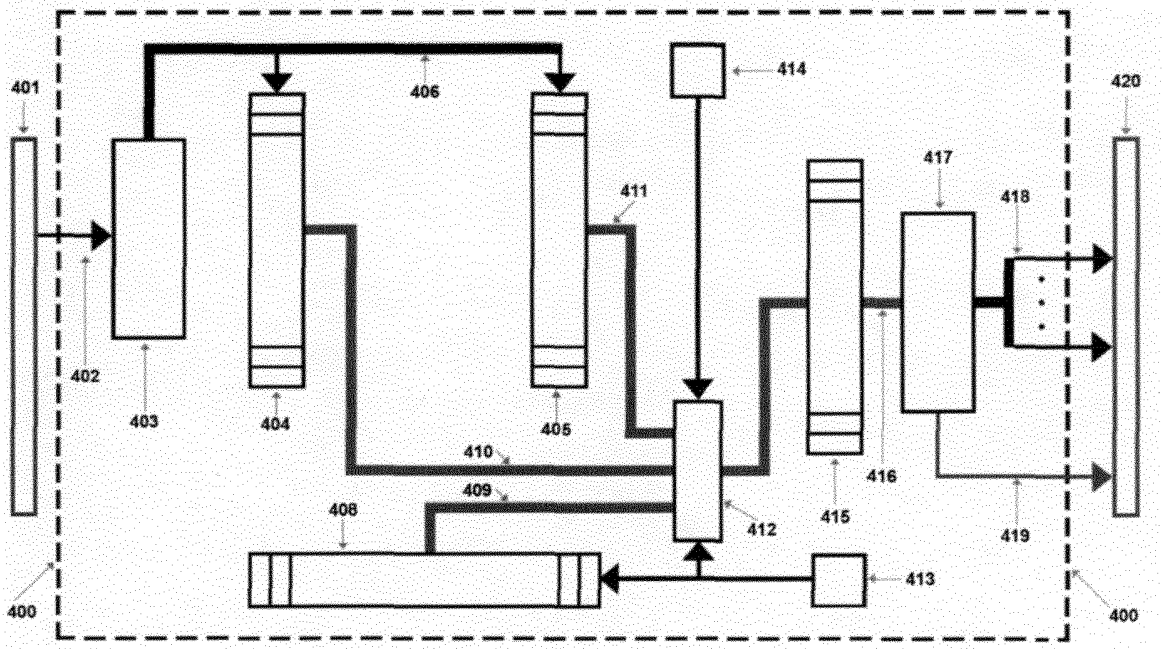


Figura 4

RESUMO

Gerador, Conversor e Codificador de Potenciais de Ação Artificiais para Redes Neurais Pulsadas descreve um sistema eletrônico que gera potenciais de ação artificiais (*spikes*) a partir de parâmetros configuráveis e que podem ser alterados em regime contínuo e em tempo real, converte dados digitais em spikes, gerando frequências de disparo em neurônios artificialmente simulados, "codifica" os spikes gerados em três formas de trens de pulsos distintas: pela média de disparo dos neurônios, conhecido em neurociência como "*rate coding*"; pelo tempo de retardo ou ordem de chegada dos spikes em relação a uma referência, chamado "*temporal coding*"; ou ainda pelo número de neurônios ativos num intervalo de tempo, conhecido por "*population coding*", e descreve um sistema eletrônico que serve de interface entre um conjunto de dados de entrada (o mundo físico, sistemas digitais, computadores, etc) e as redes neurais pulsadas (artificiais ou naturais), e faz a conversão de grandezas físicas e/ou informação para um formato apropriado para ser tratado pelas redes neurais artificiais pulsadas e comunica-se com outros dispositivos pela forma serial ou paralela, tanto na entrada quanto na saída.