

Tarefa de Reconhecimento de Imagens Baseada em uma Memória Endereçada por Conteúdo Mono-Elétron

Diego Côrtes Fagundes, Janaína Gonçalves Guimarães

¹Departamento de Engenharia Elétrica – Universidade de Brasília (UnB), Campus Universitário, Asa Norte, C.P. 4386, Brasília – DF, 70919-970, Brasil

fagundes.dc@gmail.com, janaina@ene.unb.br

Abstract. *This paper presents a Content Addressable Memory structure, formed by memory cells, a comparative circuit and a WTA network. This type of memory is used to address an input parameter based on information contained in its structure. The objective of this work is to develop a circuit that can faithfully compare input with stored image data using single-electron transistors. The proposed architecture was simulated under room temperature (300 Kelvin) for different input images sizing 35 bits each, and compared with 4 images, which were stored in the memory cells. The simulations were performed using the computational tools MATLAB and SIMON.*

Resumo. *Este trabalho apresenta uma estrutura de Memória Endereçada por Conteúdo, formada por células de memória, um circuito comparador e uma rede WTA. Este tipo de memória é usado para endereçar um parâmetro de entrada baseada em informações contidas em sua estrutura. O objetivo deste trabalho é desenvolver um circuito que possa fielmente comparar dados de imagens de entrada com armazenadas usando transistores mono-elétron. A arquitetura proposta foi simulada em temperatura ambiente (300 Kelvin) para diferentes imagens de 35 bits cada, as comparando com 4 imagens, armazenadas nas células de memória. As simulações foram feitas com o uso das ferramentas computacionais MATLAB e SIMON.*

1. Introdução

A era tecnológica na qual nos encontramos está progressivamente encontrando modos de superar as barreiras que tínhamos há alguns anos atrás. A engenharia tem provado ser eficiente se tratando de melhorias tecnológicas, respondendo bem às demandas dessa nova era. O papel de cientistas e engenheiros eletrônicos pode ser resumido ao desenvolvimento de dispositivos com 4 principais objetivos: pequeno dimensionamento, bom gerenciamento de energia, baixa taxa de erros e alta velocidade de processamento. Para inserir essas características em dispositivos eletrônicos, melhores arquiteturas e materiais estão sendo estudados e desenvolvidos. A Nanoeletrônica é um passo muito promissor que está sendo dado no mundo da eletrônica. É provado matematicamente ser mais eficiente que mecanismos microeletrônicos utilizados atualmente, principalmente em termos de área física e gerenciamento de energia.

Um transistor mono-elétron (SET) se difere dos tradicionais transistores de efeito de campo e transistores bipolares de junção na capacidade que sua estrutura possui de armazenar elétrons em dimensões suficientemente pequenas, de modo que a quantização de sua carga e de sua energia são facilmente observáveis, fazendo dos SET's dispositivos essencialmente quantum-mecânicos. Devido às dimensões que podem ser alcançadas por um SET, seu funcionamento passa a ser regido principalmente por fenômenos quânticos como o tunelamento e o bloqueio de Coulomb. Na Figura 1 pode ser visto o modelo de um SET utilizando junções-túnel. Entre as duas junções, encontra-se uma região chamada ilha, a qual é polarizada de forma a adquirir níveis de energia que permitem ou não a passagem de elétrons através da mesma, sendo assim um terminal de controle do transistor. A tensão de polarização aplicada entre os terminais das junções é responsável por gerar o campo elétrico que provocará a corrente através do dispositivo.

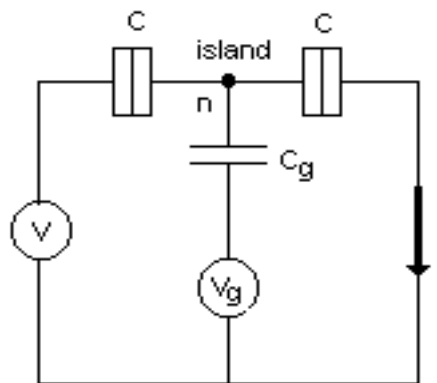


Figura 1. Esquemático de um SET

O funcionamento do transistor de acordo com suas tensões de polarização e de controle (tensão de porta) podem ser observados na figura 2. A curva da corrente por tensão de porta mostra a característica quântica do dispositivo. Pode ser facilmente constatado que existem níveis de energia em que a região da ilha será polarizada para permitir a passagem de elétrons pelas barreiras de potencial por tunelamento. A polarização no dreno, por sua vez, reduz seu potencial eletroquímico. Desse modo, a diferença de potencial entre fonte e dreno fará com que o canal entre eles adote um nível energético que permitirá um fluxo de elétrons através das duas barreiras, ou seja, uma corrente no dispositivo.

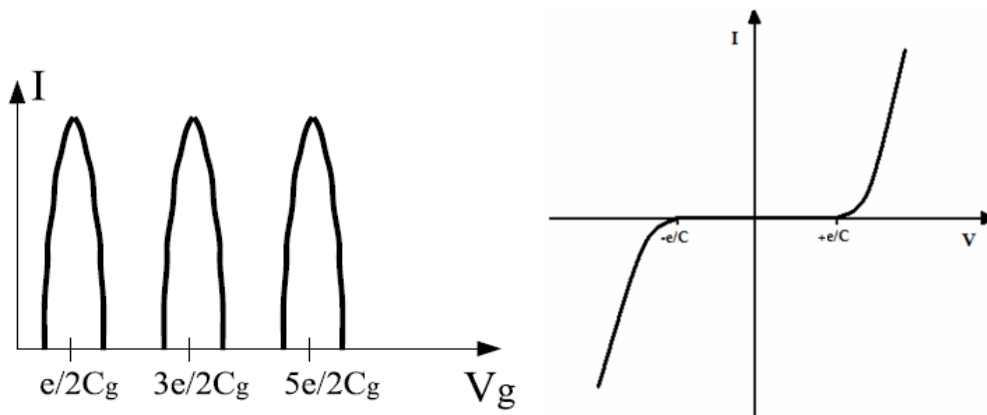


Figura 2. Curvas características do transistor mono-elétron

O funcionamento de circuitos baseados no transistor mono-elétrons pode ser avaliado com o uso de uma ferramenta computacional chamada SIMON, a qual nos permite desenhar, simular e avaliar o desempenho do circuito proposto. Este simulador utiliza o método não determinístico de Monte-Carlo, o qual se baseia em amostras aleatórias (de quantidade prevista pelo operador do simulador) para descrever o comportamento dos elétrons e calcular tensões, correntes e quantidade de carga em determinados locais do circuito. O SIMON faz uso de capacitores e junções túnel para gerar dispositivos nanoeletrônicos (o SET é composto por duas junções e um capacitor de porta). Esta ferramenta é, no entanto, muito ineficiente no projeto de circuitos com muitos componentes, como é o caso deste trabalho. Por esse motivo, o uso da ferramenta MATLAB torna-se essencial. Esta possui várias funções úteis que nos permitem facilmente multiplicar circuitos pequenos desenhados para o SIMON, para então simulá-los em conjunto e nos dar os resultados.

2. Arquitetura Proposta

A arquitetura proposta neste trabalho é uma memória endereçada por conteúdo. O funcionamento do circuito se dá de forma a selecionar qual das palavras binárias armazenadas nas células de memória mais se assemelha à palavra de entrada. Essa seleção é feita a partir de um circuito comparador baseado em portas lógicas NXOR e uma rede WTA. O circuito lógico é capaz de comparar as imagens (no caso deste trabalho) bit a bit, de modo que ele irá procurar quantizar a diferença entre duas imagens em número de bits, o que nos fornece um resultado mais confiável para fins de endereçamento. Na Figura 3 pode ser observado um exemplo que contém os principais blocos do circuito, mostrando como deve se dar seu funcionamento.

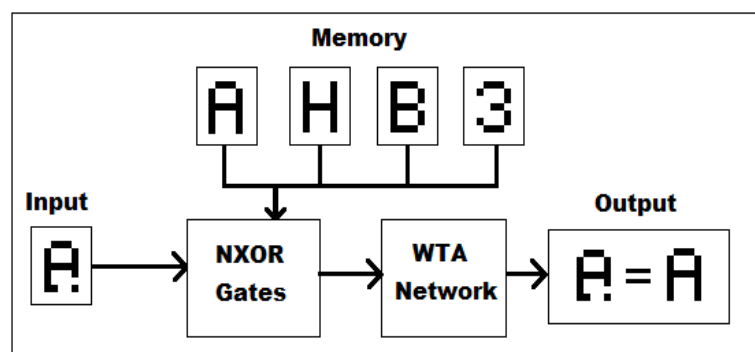


Figura 3. Funcionamento de uma memória endereçada por conteúdo

2.1 Células de Memória

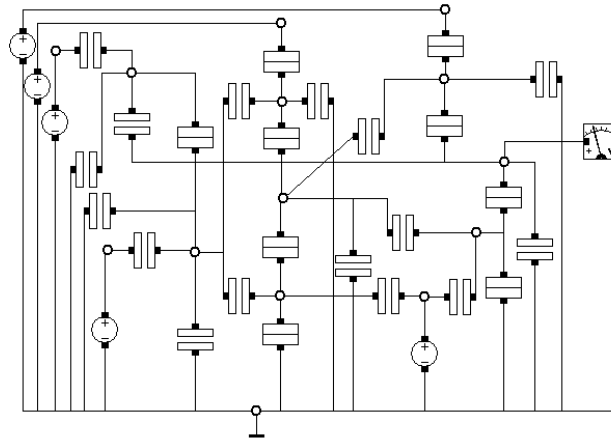


Figura 4. Célula de memória tipo Flip-Flop

Esta parte do circuito contém a informação com a qual queremos comparar a palavra de entrada. A arquitetura usada em cada célula de memória é a de um circuito Flip-Flop. Tem como objetivo armazenar informação e fornecer baixa impedância de saída, característica necessária para manter baixos níveis de tensão para diferentes valores de carga. O circuito proposto para uma célula de memória projetada com SET's é mostrada na Figura 4.

Os parâmetros dos componentes utilizados no circuito (capacitores e junções túnel) foram escolhidos inicialmente a partir de um circuito de memória encontrado na literatura. Após feitos vários testes, chegou-se ao circuito final. Neste, foram utilizadas junções túnel cujas capacitâncias chegam a valores da ordem de zepto Farads (10^{-21} Farads) ou maiores, até 10^{-19} Farads. Valores dessa ordem já têm condições de serem alcançados em processos precisos de fabricação.

As fontes de tensão assumem níveis de 0V e 1.6V para valores lógicos 0 e 1 respectivamente. Foram feitos vários testes no SIMON utilizando o circuito proposto e modificando os parâmetros de seus componentes e valores de tensão aplicados até que as saídas se adequassem às necessidades do sistema.

2.2 Circuito Comparador

A segunda parte do circuito é a estrutura de comparação, formada por operadores lógicos que associam cada bit de entrada com o bit de referência nas células de memória. Isso é feito com o uso da porta NXOR, que fornece 1 para valores iguais de entrada e 0 para diferentes. Um circuito nanoeletrônico capaz de funcionar bem como uma porta NXOR fazendo o uso dos níveis de tensão especificados anteriormente pode ser visto na Figura 5.

Neste circuito, as capacitâncias das junções túnel adotam valores de 10 zepto Farads e, dos capacitores, de 30 zepto Farads. Tais valores já podem ser encontrados com mais abundância em circuitos da literatura atual. A simplicidade do circuito da porta NXOR (apenas 2 transistores e alguns capacitores adicionais) permite seu uso em abundância.

Os níveis de tensão de saída para entradas lógicas 1,1 e 0,0 não são perfeitamente idênticos na porta desenvolvida, sendo diferentes entre si por aproximadamente 0.03V. Embora esse valor não seja significativo para um circuito digital, onde a identificação do nível lógico é bem fácil de ser feita, veremos que, na próxima etapa do circuito, será utilizado um bloco analógico que terá como entradas várias saídas de portas NXOR. Isso fará com que esse pequeno erro seja propagado e dificulte a realização de circuitos muito grandes.

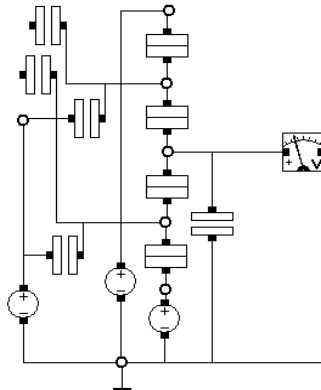


Figura 5. Porta lógica NXOR

2.3 Rede WTA

Como foi dito anteriormente, cada imagem possui um conjunto de portas NXOR, cada uma comparando um de seus bits com um bit da imagem de entrada. Todas as portas desse conjunto têm suas saídas ligadas no mesmo nó, e isso ocorre para cada uma das imagens armazenadas. Logo, teremos (pra 4 imagens) apenas 4 nós de entrada na rede neural WTA (do inglês *Winner Takes All*, ou “O Vencedor Leva Tudo”), a qual nos fornece a palavra (sequencia de bits) mais parecida com aquela que queremos endereçar. A rede WTA utilizada tem quatro nós de entrada, um para cada palavra armazenada no circuito. É importante notar que, dada essa característica de entrada da rede, o circuito passa a não ser mais completamente digital. O trabalho feito pela rede WTA não é capaz de ser realizado por nenhum circuito digital simples, e seu uso aperfeiçoa o funcionamento da memória em termos de velocidade de processamento e consumo de potência. A Figura 6 mostra o circuito proposto para a rede.

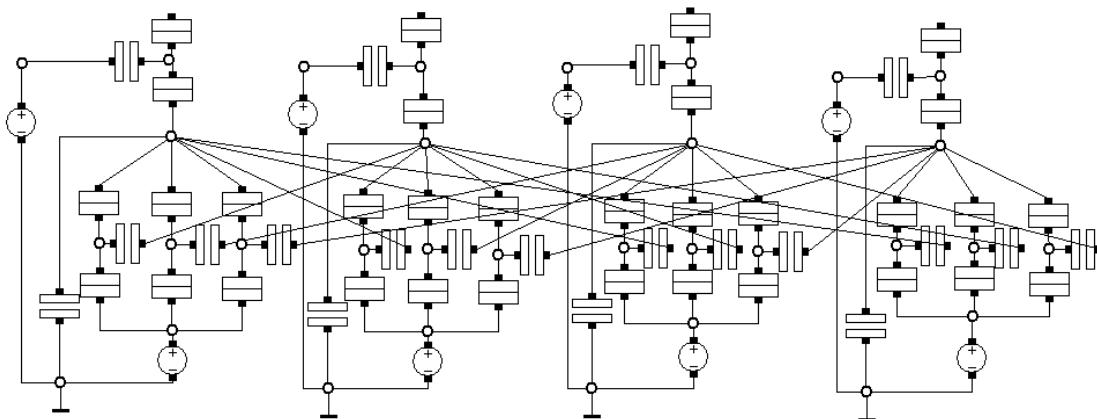


Figura 6. Rede WTA

Podem ser observadas 4 células interconectadas formando a rede. Essa interconexão é responsável pela inibição que a alta tensão de uma das células irá causar nas demais, de modo que, transcorrido o tempo de estabilização, seja possível identificar qual das células possui um maior nível de tensão na entrada.

Os valores de capacitância dos componentes da rede WTA também não foram reduzidos abaixo de 10 zepto Farads. Suas fontes de alimentação, entretanto, forem do padrão estabelecido nas células de memória e nas portas lógicas NXOR (os quais eram de 1.6V), adotando valores incomuns de 0.3V e -0.5V.

2.4 Circuito Completo

Tendo projetado todas as partes da memória endereçada por conteúdo, foi feito um código no MATLAB para juntá-las, multiplicadas em quantas células fossem desejadas. A última variável a ser definida foi o número de bits que teria a imagem com a qual o trabalho foi feito. Estimou-se 35 bits como sendo um tamanho razoável para as imagens (5x7 pixels), dada a facilidade de desenhar letras e números nesta definição. Em virtude das limitações da ferramenta computacional, circuitos maiores se provaram difíceis de serem simulados, levando quase 50 horas para que o SIMON finalizasse uma simulação de imagens de 64 bits e quatro palavras de comparação para um período de simulação de 2ns (2 nano segundos), com um passo de amostragem de 2ps (2 pico segundos).

Na ferramenta computacional MATLAB, foi desenvolvido um código que relaciona o arquivo *.set* gerado a partir de um modelo de circuito construído no SIMON com um bloco lógico presente na ferramenta SIMULINK (interna ao MATLAB). Esse código permite que, a partir de circuitos desenhados no SIMULINK, circuitos nanoeletrônicos no SIMON sejam gerados utilizando os modelos previamente desenhados. Isso permite que circuitos maiores sejam projetados sem que haja necessidade de um tempo exaustivo para que isso seja feito, tornando os processos de design e simulação mais eficientes. Além disso, a ferramenta matemática inicia as simulações e importa os valores do simulador para análises mais precisas dos resultados.

Uma observação importante a ser feita é o número de componentes utilizados no circuito inteiro. Cada célula de memória contém 9 junções túnel e 16 capacitores, cada porta NXOR contém 4 junções e 5 capacitores, e a rede WTA contém 32 junções e 20 capacitores. Como o circuito completo contém 140 células de memória, 140 portas NXOR e uma rede WTA, ele possui 1852 junções túnel e 2960 capacitores.

3. Resultados e Discussões

Para testar o circuito proposto, imagens de 35 pixels foram desenhadas e traduzidas para uma palavra de 35 bits. O melhor par para a imagem de entrada deve ser a imagem que possui a menor distância de Hamming da mesma. A distância de Hamming mostra (em bits) o quão diferente uma palavra é de outra, ou quantos bits de diferença elas possuem entre si. A Figura 7 mostra um exemplo de 9 imagens usadas para comparação, com seus códigos binários correspondentes.

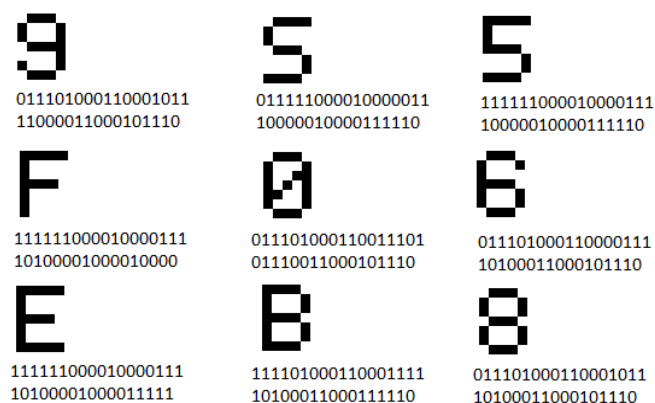


Figura 7. Símbolos de 35 bits e seus respectivos códigos binários

O teste mostrado neste trabalho foi feito com a imagem de entrada “5”, comparada com as imagens “9”, “S”, “F” e “0”, que podem ser vistas na Figura 5. A imagem esperada para apresentar o maior nível de tensão após a comparação era claramente a letra “S”, com a menor distância de Hamming para com o número “5” (apenas 2 bits de diferença). Essas distâncias podem ser vistas na Tabela 1.

Tabela 1. Resultados numéricos

Símbolo de Comparação	Distância de Hamming	Nível de Tensão [V]
9	8	0.39
S	2	0.43
F	7	0.40
0	12	0.32

Na mesma tabela estão descritos os valores médios de tensão alcançados após estabilizados os níveis de saída. Pode ser visto que o maior nível de tensão é atribuído à menor distância de Hamming e vice-versa.

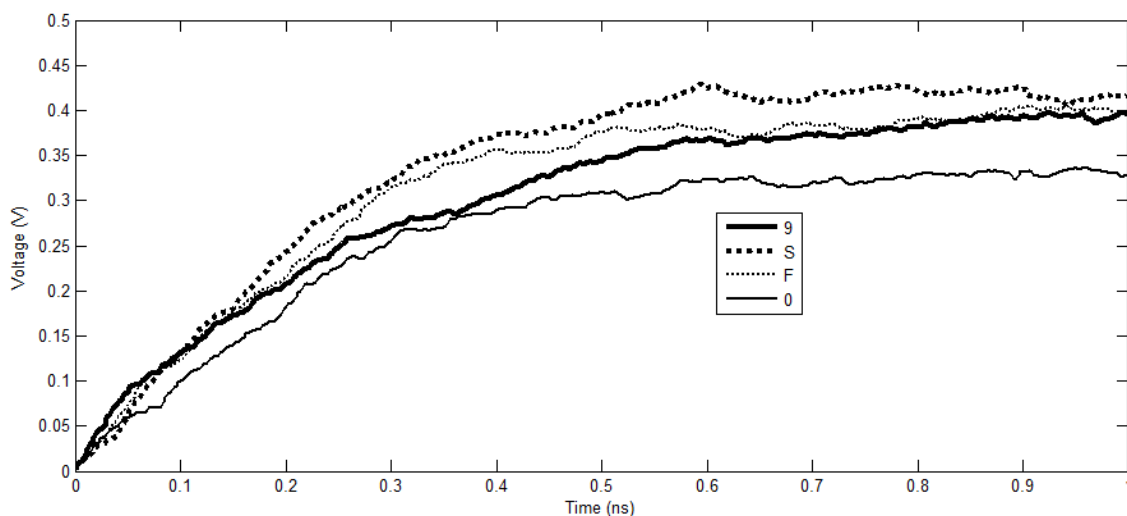


Figura 8. Resultados da simulação

A Figura 8 mostra o resultado da tensão colhida em cada uma das 4 saídas do circuito no primeiro nano segundo de operação. Após esse período, os níveis de tensão já se encontram estáveis e não é observada nenhuma alteração significativa nos mesmos.

Podemos observar que, no período de estabilidade, a tensão de cada palavra segue um padrão de acordo com a distância de Hamming desta à palavra de entrada. A relação entre estas duas variáveis, entretanto, não é linear, dada a complexidade da rede WTA e da forma com que suas saídas dependem umas das outras.

Observa-se também uma equivalência muito grande entre os níveis de tensão das imagens “9” e “F”, visto que as distâncias de Hamming de ambas para com o “5” são muito semelhantes. Isso implica em uma dificuldade cada vez maior de identificar a palavra vencedora quando distâncias próximas são adquiridas. Na medida em que circuitos maiores foram desenhados (para imagens com mais de 35 pixels), notou-se uma diferença decrescente entre os níveis de tensão em palavras com as características descritas, chegando a um ponto em que se tornou impossível a identificação da palavra vencedora. Isso acarretou resultados indesejáveis para circuitos maiores, cuja realização requer não um aumento no número de células, mas parâmetros diferentes para as mesmas.

4. Conclusões

Os resultados obtidos das simulações foram satisfatórios se tratando da identificação da palavra vencedora. Em todos os testes feitos, o circuito respondeu conforme o esperado, relacionando maiores níveis de tensão a menores distâncias de Hamming. Entretanto, avaliando os valores de tensão de saída, observa-se uma diferença muito pequena entre as quatro palavras de comparação, o que torna difícil a identificação da palavra vencedora, principalmente quando suas distâncias de Hamming são muito parecidas. Isso faz com que circuitos de medição muito precisos sejam necessários.

A necessidade de criar circuitos que suportam palavras maiores e mais células de memória faz com que uma maior exatidão nas saídas da porta NXOR seja altamente significativa. Para uso em circuitos de tal magnitude, será necessário modificar os parâmetros de seus componentes ou o nível de tensão em que o circuito trabalha. Para que, no futuro, exista uma facilidade maior no projeto de sistemas digitais mono-elétrons, trabalhos de aperfeiçoamento de todas as portas lógicas nanoeletrônicas estão sendo feitos.

O uso da rede WTA transforma a memória endereçada por conteúdo em um circuito cujas saídas não são digitais. Isso pode ser indesejável dependendo do tipo de carga ou aparelho de medição acoplado à memória. Como adição ao circuito para suprir essa necessidade, pode ser proposto um conversor AD mono-elétron que faça a diferenciação entre os níveis de tensão nas saídas da memória, e que possa completar o endereçamento tornando o circuito digital nos blocos de entrada e de saída.

É importante notar que os resultados foram obtidos a partir de um circuito mono-elétron de larga escala (com quase 5000 componentes) simulado a temperatura ambiente. Essas características são essenciais para futura fabricação de sistemas como este. No entanto para atingir objetivos de substituir mecanismos microeletrônicos existentes, novas arquiteturas e circuitos ainda maiores devem ser propostos.

5. Referências

- Wasshuber C., “Single-Electronics - How It Works. How It’s Used. How It’s Simulated”, Texas Instruments, PO Box 650311, Dallas TX 75265
- Ahmed, H., Nakazato, K., “Single-electron devices”, *Microelectronic Engineering* 32, (1996), 297 – 315
- de Alencar B. M. S. M. and Guimarães J. G., “Single-Electron Content-Address Memory Circuit”
- Guimarães J.G., do Carmo H.C. and da Costa J.C., “Single-Electron Winner-Take-All Network”, *Microelectronics Journal* 35 (2004) 173–178
- Mattausch, H.J., Imafuku, W., Ansari, T., Kawabata, A. and Koide, T., “Low-power word-parallel nearest-Hamming-distance search circuit based on frequency mapping”, *Proceedings of the ESSCIRC*, (2010), 538 – 541
- Mattausch, H.J., Gyohten, T., Soda, Y., and Koide, T., “Compact Associative-Memory Architecture With Fully Parallel Search Capability for the Minimum Hamming Distance”, *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 2, (2002), 218 – 227