- Friday, March 11, 2005
  Kerneltron: Massively parallel mixed-signal VLSI pattern recognition processor

  Recently it has been shown that a simple learning paradigm, the support vector machine (SVM), outperforms some of the most elaborately tuned expert systems and neural networks in object recognition tasks. In run-time, the SVM operates by computing a kernel-based distance between the object's vector at the input and a set of support vectors selected from the training set, and weighting the results to produce the oracle at the output. Real-time SVM recognition of complex objects in streaming video incurs an excessive amount of computation, well beyond even the most powerful digital signal processors available today. This calls for a radically different computational paradigm to efficiently compute kernels in very large dimensions. I will present a massively parallel, fine- grain distributed architecture for real-time kernel ``machines'' and its efficient implementation in mixed-signal VLSI technology. At the core of the externally digital architecture is a high-density, low- power analog array performing binary-binary vector-matrix multiplication, as the elementary operation in computing inner- product based kernels between presented input and stored support vectors. The three-transistor unit cell in the analog array combines a charge injection device (CID) binary multiplier and analog accumulator with embedded dynamic random-access memory (DRAM). I will present various schemes to obtain precise digital results from the internal analog computation in a distributed, parallel fashion, using partial analog-to-digital quantization of partial binary-binary products computed over the array. High output resolution is achieved with low complexity quantizers by oversampling in the input binary representation combined with delta-sigma modulated quantization at the output. In addition, stochastic encoding of the digital inputs relaxes the precision requirements of the quantizers by the square root of the vector dimension owing to the Central Limit in the accumulation of binary terms in the inner product. Our research has resulted in the Kerneltron, the first support vector ``machine'' in silicon. A 3mm by 3mm 0.5 micron CMOS chip features 256 inputs and 128 support vectors, delivering over 1 trillion (1012) multiply-accumulates-per-second for every Watt of power. An integrated bank of 128 delta-sigma modulated algorithmic analog-to- digital converters produce for each output 8 bits of resolution in 32 cycles. A prototype PCI board interfaces a workstation with 4 Kerneltron chips and a Xilinx FPGA to implement and train general kernel-based classification and regression architectures in efficient and reconfigurable manner. Applications of the Kerneltron include artificial vision, automated surveillance, and human-computer interfaces.

  Roman Genov
  Department of Electrical & Computer Engineering U of T