



Doctoral Thesis

Analog CMOS implementation of a multi-layer perceptron with nonlinear synapses

Author(s):

Lont, Jerzy B.

Publication Date:

1993

Permanent Link:

<https://doi.org/10.3929/ethz-a-000943706> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

1 6. Mai 1994

Diss. ETH No. 10244

Analog CMOS Implementation of a Multi-Layer Perceptron with Nonlinear Synapses

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Technical Sciences

presented by
JERZY B. LONT
dipl. El.-Ing. ETH
born March 22, 1959
citizen of Poland

accepted on the recommendation of
Prof. Dr. W. Guggenbühl, examiner
Prof. Dr. G. S. Moschytz, co-examiner

1993

20. April 1994

W. B. Guggenbühl

Abstract

This thesis presents the theory and the hardware implementation of a multi-layer perceptron with nonlinear synapses. A formalization of the error back-propagation algorithm is shown which allows the use of very small nonlinear synapses. The influence of several nonidealities of the implemented circuit on the network performance is analyzed. Nonidealities such as offset voltages, DRAM cell leakage currents, noise induced by switches, and clock feedthrough are discussed. Furthermore, it is shown that the back-propagation algorithm is insensitive to offsets in synapses and somas. The usefulness of large-scale multi-layer perceptrons with nonlinear synapses for real-size applications such as handwritten-character recognition is demonstrated. The design is verified by extensive SPICE simulations and implemented in a 3 μm CMOS process. With this process, two orders of magnitude higher synapse density than with a conventional approach using linear multipliers is achieved. The 1.6 mm by 2.4 mm NEUROCHIP1 contains 18 neurons and 161 synapses in three layers, and provides 16 analog inputs and 4 analog outputs. It is incorporated into a complete neurocomputer consisting of a personal computer for the training phase, an interface card, and the neurochip. Moreover, results obtained with this experimental chip are reported and discussed. Thus, the feasibility of the formalized back-propagation learning algorithm applied to nonlinear synapses is experimentally verified. Due to other nonidealities of the network pretraining of the net on a simulator is necessary. Furthermore, two successful applications of this neurochip, one in emulating a fuzzy controller for an inverted pendulum and the other one in the control of an autonomous vehicle, are demonstrated. Both of them fully exploit the special features of the network such as analog inputs and analog outputs. Finally, the design, the implementation, and experimental results of the NEUROCHIP2 utilizing floating-gate devices for weight storage are

summarized. With this chip, the synaptic density is doubled and some problems originating from the use of DRAM cells are solved. 4.2 millions neurons with as many as 38 millions synapses could be placed on the standard 6 inch wafer using a state-of-the-art 1 μm CMOS process. Such a WSI (wafer scale integration) neural network would have an unprecedented computational power of $4 \cdot 10^{13}$ CPS (connections per second) outperforming modern supercomputers by more than two orders of magnitude.

Keywords. Neural Network, Feedforward Network, Multi-Layer Perceptron, Nonlinear Synapse, Error Back-Propagation Algorithm, Formalized Back-Propagation, Learning Algorithm, Training Algorithm, Neurocomputer, Neurochip, Analog CMOS Integrated Circuit, Wafer Scale Integration, Analog EEPROM, Floating-Gate Device, Nonvolatile Storage, Charge Retention, Fuzzy Logic, Fuzzy Controller, Autonomous Vehicle, Ultrasonic Collision Avoidance System, Handwritten Character Recognition

Kurzfassung

Die vorliegende Arbeit befasst sich mit der Theorie und Hardware-Implementierung eines Mehrschichten-Perzeptrons mit nichtlinearen Synapsen. Dabei wird eine Formalisierung des Error-Back-Propagation-Algorithmus gezeigt, welche den Gebrauch von sehr kleinen, nichtlinearen Synapsen ermöglicht. Es wird u.a. der Einfluss von verschiedenen Nichtidealitäten der implementierten Schaltung auf die Netzwerkeistung analysiert und folgende Nichtidealitäten diskutiert: Offset-Spannungen, Leckströme der DRAM-Zellen, durch die Schalter verursachtes Rauschen und Clock-Feedthrough (Schalter-Übersprechen). Im weiteren wird gezeigt, dass der Back-Propagation-Algorithmus nicht sensitiv ist gegenüber Offsets in Synapsen und Somata. Ferner wird die Tauglichkeit der grossformatigen Mehrschichten-Perzeptre mit nichtlinearen Synapsen für praktische Anwendungen, wie z.B. die Handschrifterkennung, demonstriert. Der Entwurf wird durch umfassende SPICE-Simulationen verifiziert und in einem 3 μm -CMOS-Prozess implementiert. Mit diesem Prozess wird eine um zwei Grössenordnungen höhere Synapsen-Dichte erreicht als bei der Anwendung von konventionellen linearen Multiplizierern. Der 1.6 mm mal 2.4 mm grosse NEUROCHIP1 enthält 18 Neuronen und 161 Synapsen verteilt auf drei Schichten und verfügt über 16 analoge Ein- und 4 analoge Ausgänge. Er ist Bestandteil eines vollständigen Neurocomputers, welcher aus einem Personal Computer für die Trainingsphase, einer Schnittstellenkarte und dem Neurochip selbst besteht. Resultate, die mit diesem experimentellen Chip erzielt wurden, werden aufgeführt und diskutiert. Demgemäss wird die Durchführbarkeit des auf die nichtlinearen Synapsen angewandten formalisierten Back-Propagation-Lernalgorithmus experimentell verifiziert. Infolge anderer Nichtidealitäten des Netzwerks ist ein Vortraining des Netzes auf einem Simulator erforderlich. Zwei erfolgreiche Anwendungen des Neurochips werden

vorgestellt — eine als Emulator eines Fuzzy-Reglers für ein invertiertes Pendel und eine als Regler eines autonomen Fahrzeugs. Beide machen vollen Gebrauch von den speziellen Eigenschaften des Netzwerks im Sinne analoger Ein- und Ausgänge. Schliesslich wird der Entwurf, die Implementierung und die experimentellen Resultate des NEUROCHIP2, der Floating-Gate-Transistoren als Gewichtsspeicher verwendet, zusammengefasst. Mit diesem Chip wird die Synapsen-Dichte verdoppelt und einige Probleme gelöst, die durch die Verwendung von DRAM-Zellen entstehen. Auf einer 6-Zoll-Silizium-Scheibe könnten bei Anwendung eines modernen 1 μm -CMOS-Prozesses 4,2 Millionen Neuronen mit 38 Millionen Synapsen plaziert werden. Ein solches WSI-neuronales Netzwerk (Wafer Scale Integration) hätte eine bisher unerreichte Rechenleistung von $4 \cdot 10^{13}$ CPS (connections per second) und würde die Rechenkapazität moderner Supercomputer um mehr als zwei Grössenordnungen übersteigen.

Stichworte. Neuronales Netzwerk, 'Feedforward' Netzwerk, Mehrschichten-Perzeptron, Nichtlineare Synapse, 'Error Back-Propagation' Algorithmus, Formalisierte 'Back-Propagation,' Lernalgorithmus, Trainingsalgorithmus, Neurocomputer, Neurochip, Analoge CMOS-Integrierte Schaltung, 'Wafer Scale Integration,' Analoges EEPROM, 'Floating-Gate'-Transistor, Nichtflüchtiger Speicher, Speicherverhalten, Fuzzy-Logik, Fuzzy-Regler, Autonomes Fahrzeug, Ultraschall-Kollisionsvermeidungssystem, Handschrifterkennung