Part III  Systems and Signals

Section 1  Computer-Aided Design
Section 2  Digital Signal Processing
Section 1  Computer-Aided Design

Chapter 1  Custom Integrated Circuits

Chapter 2  Computer-Aided Fabrication System Structure
Chapter 1. Custom Integrated Circuits

Academic and Research Staff
Professor Jonathan Allen, Professor John L. Wyatt, Jr., Professor Srinivas Devadas, Professor Jacob K. White, Professor Dimitri A. Antoniadis, Professor Berthold K.P. Horn, Professor Hae-Seung Lee, Professor Tomaso Poggio, Professor Stephen D. Senturia, Professor Charles G. Sodini, Dr. Sami M. Ali, Dr. John G. Harris

Visiting Scientists and Research Affiliates
James H. Kukula

Graduate Students

Technical and Support Staff
Dorothy A. Fleischer, Susan E. Nelson

1.1 Custom Integrated Circuits

Sponsors
Analog Devices
IBM Corporation

Project Staff
Professor Jonathan Allen, Robert C. Armstrong, Donald G. Baltus, Lisa A. Pickelsimer, Mark W. Reichelt, Filip J. Van Aelten

The overall goal of VLSI computer-aided design (CAD) research is to provide the means to produce custom integrated circuits quickly, correctly, and economically. Traditionally, correctness has been checked at several representational levels of abstraction, such as layout (via design rule checking), and circuit and logic representations (both via simulation). These techniques for checking correctness are usually local to the particular representational level involved and, while these techniques are important components of the overall design testing, they do not attempt to provide for alignment and consistency checks between the different representational levels and an input behavioral specification. In addition, they do not characterize the set of possible designs at each representational level corresponding to the initial functional specification in a way that ranges over a variety of performance levels. For this reason, there is an increasing need to provide CAD tools that serve as a framework for design exploration, thus providing the desired performance together with consistently aligned representations at all levels.

This research group studies a variety of research topics with an emphasis on performance-directed synthesis of custom VLSI designs. An overview of the viewpoint that motivates these projects has recently been provided in a major survey paper in which the need for coordinating a design optimization process over the several levels is emphasized. Since design exploration is so central to the production of high-performance designs, emphasis is placed on how performance can be characterized at the several levels of representation and how overall optimal performance can be achieved in an integrated way. In addition to the basic issues of circuit optimization, architectures for digital signal processing have been studied because of the highly parallel nature of the algorithms involved and the need for a very high level of real-time performance in these systems. Emphasis on devel-

---

1 IBM Corporation.

Chapter 1. Custom Integrated Circuits

Noting formally specified systems has been increased so that the design space can be comprehensively searched and verification of the resultant design can be confirmed in terms of the initial behavioral specification. Current projects focus on design methodologies for very high-speed clocked circuits, design of highly optimized array architectures for digital signal processing, and provision of a VLSI design database that provides for incremental consistency maintenance.

There has been a continuing evolution of circuit styles that seek to provide high performance in a minimal space. Thus, in CMOS technology, static designs are utilized, but precharge-evaluate methodologies have also been employed to provide speed, minimize the number of transistors, and increase the utilization of n-channel devices (rather than slower p-channel devices). Recently, examples of a new circuit style called True Single-Phase Clocking (TSPC) have been published.3 This style provides for clock speeds of several hundred megahertz in modest CMOS technologies. While these results have been confirmed, an analysis to explain the speed has not been provided, and no overall design methodology for these circuits has been demonstrated. In her Master's thesis research project, Pickelsimer is providing both the required insight into the cause of the high-speed performance as well as a comprehensive design methodology for these circuits which will cover both combinational and sequential circuits. In addition, comparisons are provided with several CMOS circuit styles which include simulations based on layouts of commonly used circuit modules. These contributions are expected to promote the utilization of TSP clocked circuits for a wide variety of applications and provide an important capability to the circuit aspect of performance-directed synthesis.

In the survey of performance-directed synthesis by Allen, the need for integrated performance optimization across multiple levels of representation was pointed out. In order to gain concrete experience for these techniques,4 Baltus has extended his work on the generation of custom layout from netlist specifications to the systematic exploration of array-like digital signal processing architectures from a high-level functional specification. In this way, designs can be optimized at a high level in a meaningful way with realistic delay models. The input specification for this algorithm is expressed in a language called FLASH, which affords several improvements over the commonly used SILAGE language.5 FLASH designs are then converted to data dependence graphs, and some elements of these graphs are reused and indexed to provide a compact representation that does not grow with the size of the repetitive array aspect of the problem. Techniques to optimize scheduling without commitment to allocation have been provided for array designs in several dimensions, and novel architectures have been obtained for a variety of common signal processing tasks (such as finite impulse response filters). This process of deriving schedules, coupled with the following allocation process, constitutes a comprehensive search of the architectural design space under the given cost function and has already produced many useful but nonintuitive designs. Another novel aspect of this project is the introduction of an affine delay model, which corresponds to the utilization of elements in arbitrary positions within the array by a particular element of the array. By controlling this delay model, it is possible to study several different designs which vary in the amount of communication delay from source modules to the destination module which requires these inputs. This project makes many new contributions to architectural design for digital signal processing systems including scheduling optimization with affine delay models, multidimensional mapping from the original design space to two-dimensional implementation spaces, and strong coupling between layout performance considerations and communication constraints at the architecture level.

Armstrong has been studying the design of an incremental consistency maintenance schema in a multirepresentation VLSI database. Here, the goal is to develop a design that automatically, incrementally, and continuously provides for consistent alignment of all levels of design abstraction. That is, from the start of the design process, all levels are kept in alignment through the introduction of explicit correspondence zones between adjacent representational levels. This project has required the elaboration of complex database schemas, constraint evaluation, and maintenance of broad utility among the overall structural representations.

---

and a substantial implementation that employs over 50,000 lines of code on a contemporary workstation. This project is now being extended in several different ways to include coupling to other design modules, introduction of new schema compilers, and extension of the basic methodology to functional (nonstructural) representations.

There is increasing interest in overall CAD systems where the database is central and surrounded by specific applications for design and verification at a variety of different levels of representation, such as in the database just described. Plans are underway to build a client server architecture that will consist of coordinated workstation, mainframe, and large disk system environments interconnected with high-speed fiber optic switching. This system will serve as a test bed for the combined utilization of high-performance workstations and mainframe-based storage for high-speed, reliable database access. This environment also provides a structure in which distributed parallel algorithms can run in a coherent way within an overall environment that supports other specialized, nonparallel CAD algorithms.

1.2 Analog VLSI Systems for Integrated Image Acquisition and Early Vision Processing

Sponsors
National Science Foundation/Defense Advanced Research Projects Agency
Grant MIP 88-14612

Project Staff
Professor John L. Wyatt, Jr., Professor Berthold K.P. Horn, Professor Hae-Seung Lee, Professor Charles G. Sodini, Professor Tomaso Poggio, Steven J. Decker, Lisa G. Dron, Ibrahim M. Elfadl, Kamyar Eshghi, Mikko Hakkarainen, Dr. John G. Harris, Frederick P. Herrmann, Craig L. Keast, David R. Martin, Ignacio S. McQuirk, Mark N. Seidel, David L. Standley, Clay M. Thompson, Christopher B. Umminger, Paul C. Yu

1.2.1 Project Overview

In real-time machine vision the sheer volume of image data to be acquired, managed and processed leads to communications bottlenecks between imagers, memory and processors, and to very high computational demands. We are designing and testing experimental analog VLSI systems to overcome these problems. The goal is to determine how the advantages of analog VLSI—high speed, low power and small area—can be exploited and its disadvantages—limited accuracy, inflexibility and lack of storage capacity—can be minimized. The work is concentrated on early vision tasks, i.e., tasks that occur early in the signal flow path of animal or machine.

Completed designs include seven different chips for image filtering and edge detection, moment extraction to determine object position and orientation, image smoothing and segmentation, depth determination from stereo image pairs, scene depth determination jointly from stereo depth and surface shading, and camera motion determination. The typical subsystem is physically very small and performs one or more computationally intensive image-processing tasks at hundreds to thousands of frames per second using only tens to hundreds of milliwatts.

There is no single design strategy, but each of the designs has many of the following features:

- sensors (typically on-chip) and processing circuitry tightly coupled,
- parallel computation,
- analog circuits for high throughput, low latency, low power and small area,
- selection of tasks and algorithms requiring low to moderate precision,
- special emphasis on computations that map naturally to physical processes in silicon, e.g., to relaxation processes or to resistive grids,
- emphasis on charge-domain processing, e.g., CCD and switched-capacitors, for maximal layout density and compatibility with CCD sensors,
- sufficiently fast processing that requires no long-term storage circuitry,
careful matching of algorithms, architecture, circuitry and (often custom) fabrication for maximum performance, and

modular design, with a standardized input and output, for compatibility between subsystems.

1.2.2 Recent Progress

This project has shown experimentally that continuous-time CMOS and clocked CMOS/CCD systems, appropriately designed, can perform extremely fast early vision processing at a minimum cost in power and area.

Perhaps the most exciting results are the seven analog vision chip designs described below. All have been fabricated either as complete systems or in the form of test chips containing each of the cells needed in the final array. The image filtering and edge detection chip and the moment chip have been fabricated in system form with on-chip imagers, and testing has been completed. Advanced versions of these two systems are now being developed by their inventors, who have graduated from MIT. The CMOS/CCD and the unclocked CMOS image smoothing and segmentation chips have been fabricated in both test-cell form and as complete systems with on-chip imagers. The former versions have been tested, and testing is now underway for the system versions of these two chips. The depth-from-stereo chip and the depth-from-stereo-and-surface-shading chip are intended as components of larger systems and do not have on-chip imagers. The former has been fabricated with a redesign currently under fabrication, and the latter has been fabricated in a one-dimensional form. Both will be tested soon. The camera motion chip has been fabricated and tested in test-cell form, and a complete system with on-chip imager will be fabricated soon. The tested systems all operate at very high speeds (1,000 to 5,000 frames per second), and all systems were designed for speeds in that same range.

We have also developed a 1.75 micron CCD/CMOS process on campus for vision chip fabrication. The CCD enhancements were developed by this project, and the process is now successfully run by staff so that other designers can use it. The CCD/CMOS image smoothing and segmentation chip was fabricated in this process.

These and other developments are described in more detail below.

1.2.3 Vision Chip Designs

**NMOS/CCD Imager and Parallel, Pipelined Filter for Edge Detection**

Woodward Yang, working with Dr. Alice Chiang at Lincoln Laboratory and Professor Tomaso Poggio, has designed and tested a 64x64 NMOS/CCD imager and analog filter chip. It consists of a high fill-factor CCD imager and a parallel, pipelined processor that convolves the image with a 7x7 Laplacian of a Gaussian mask to enhance edges. All 64 rows of the image are fed in parallel to a fully pipelined analog processor that is pitch-matched to the imager. The entire system operates at 1,000 frames/second on a 10 MHz clock and consumes only 40 mW on-chip. The processor design minimizes area by carrying out delay, division and summing operations entirely in the charge domain and therefore occupies only 15 percent of the total chip area. The chip was conservatively designed using 4 micron design rules, prior to the creation of the MOSIS and campus CMOS/CCD processes. Dr. Yang, now an assistant professor at Harvard, is making a 2 micron redesign for MOSIS fabrication. The entire system would scale to 256x256 pixels in a 1 micron process.\(^7\)

**Continuous-Time CMOS Moment Chip for Object Position and Orientation**

David L. Standley has completed and tested a novel, very specialized chip that rapidly determines the location and orientation of an object in its field of view. The design principle, discovered by Professor Berthold Horn, is entirely new. The location and orientation are most accurately determined to sub-pixel resolution by a least-squares technique using five linear and quadratic grey-scale moments, specifically the intensity-weighted averages of \(x, y, xy, x^2 - y^2\). These functions are all **harmonic**, i.e., their Laplacian vanishes identically, and therefore they can be processed very efficiently by a uniform resistive grid. The circuit exploits this new principle to reduce the order of the calculation for an \(N\times N\) image from \(N^2\) (one addition and multiplication per pixel) to \(4N\) (one addition and multiplication per pixel).\(^7\)

---

operation pair only on each current leaving the boundary of the resistive grid and then to 4 (one operation on the current leaving each corner of a surrounding square made of 4 resistive lines). The chip has aphototransistor array driving a 30x30 resistive grid and was fabricated in a 2 micron MOSIS process. It operates in continuous time at 5,000 frames/second (governed by relaxation time to 1 percent error) on 25 mW. Orientation measurement is less accurate than position, with worst case error of 10 degrees for uncalibrated measurements and 2 degrees for calibrated measurements of a diamond-shaped object. Dr. Standley, presently employed at Rockwell International, is developing a larger 2-chip version of this system.8

We have filed U.S. patent applications for Drs. Yang and Standley’s chips. Both are described in more detail in the conference papers included in the bibliography at the end of this section.

**CMOS/CCD Imager and Fully Parallel Smoothing and Segmentation Chip**

Image segmentation and smoothing systems perform linear smoothing on approximately uniform regions of an image while preserving or enhancing edges by nonlinear operations. When the smoothing operation is quite strong the output image is approximately piecewise constant in intensity with sharp edges, a reduced-bandwidth form that is suitable for rapid subsequent processing by analog or digital systems. Craig Keast, working with Professor Charles Sodini, has designed, simulated and fabricated a CMOS/CCD imager and processor chip to perform rapid image acquisition, smoothing and segmentation. The fabrication was done on campus in the CMOS/CCD process Mr. Keast developed as an enhancement to MIT’s baseline 1.75 micron CMOS process (see CMOS/CCD Fabrication Process section).

The architecture uses standard CCD imaging techniques to transfer image brightness into signal charge. The Gaussian smoothing operation is approximated by a discrete binomial convolution of the image with a controllable support region. The design incorporates segmentation circuits with variable threshold control at each pixel location to enhance edge localization in the image. Once processed, the image can be read out using a standard CCD clocking scheme.

Currently, wafers from the first run are being evaluated using Polaroid’s CCD test system. First order testing of a 4x4 processor array showed functionality of the three system operations: Imaging, Smoothing (spatial low pass filtering), and Smoothing with Segmentation. Detailed sensitivity analysis along with a comparison between measured and simulated data is currently underway. After the characterization of the 4x4 array is complete, the testing focus will shift to the much larger 40x40 imager and processor array that has also been fabricated.8

**Continuous-Time CMOS Imager and Smoothing and Segmentation System with Resistive Fuses**

As an alternative to the clocked system described above, a network consisting of a grid of vertical and horizontal resistors can perform the normally mutually exclusive functions of image smoothing and segmentation simultaneously, when resistive fuses are employed as horizontal resistors. Resistive fuses are nonlinear resistors that act like

---


positive linear resistors for small applied voltages, but their current begins to decrease and eventually becomes quite small for large applied voltages. Thus the fuse "breaks" when the voltage difference between two adjacent pixels is large, inhibiting smoothing at the edges. Since a large number of resistive fuses are needed in a grid, the simplicity and size of the resistive fuse circuit becomes critical.

This project has developed a variety of very compact resistive fuse circuits as described in the section on Special Circuits for Vision Chips. Paul Yu, working with Professor Hae-Seung Lee, has designed a 32×32 array, using one of these fuse types with a 32×32 imager and a read-out system. It has been fabricated through MOSIS, and initial tests seem to indicate that it works correctly. More extensive testing is now underway.¹⁰

**CMOS/CCD Stereo Chip**

Mikko Hakkarainen and Professor Hae-Seung Lee have completed the design of a high-speed stereo vision system in analog VLSI technology. Three stereo algorithms have been considered, and a simulation study examined details of the interaction between algorithm and analog VLSI implementation: the Marr-Poggio-Drumheller algorithm was shown to be best suited for this technology. Particular attention was paid to the fact that analog signal processing accuracy is limited to about 8 bits (1 part in 256). Hakkarainen and Lee have shown that several stages of the algorithm only require about 4 bits of accuracy, and about 6 bits is sufficient throughout.

They have proposed an analog CMOS/CCD system implementation of the Marr-Poggio-Drumheller algorithm. The system consists of six modules: two imagers, two image filters (that perform a spatial band-pass operation), a shift-and-correlate module, and a vote taker. The shift-and-correlate module compares the band-pass filtered left and right images pixel-by-pixel in parallel, and generates match data (degree of matching) for all possible shift (disparity) values for each pixel. This module does most of the processing in the overall stereo system, and thus presents both a speed and functionality bottleneck. For this reason, the shift-and-correlate module was selected for VLSI implementation, the other blocks being initially emulated by a digital computer.

CMOS/CCD signal processing technology is well suited for the shift-and-correlate module due to the ability of a CCD device to function as an analog memory, subtractor, and shifter register. A 40×40 absolute-value-of-difference array, the core processor of the shift-and-correlate module, has been fabricated through MOSIS in a 2 micron CMOS/CCD process. This array performs a pixel-wise comparison between two input arrays (band-pass filtered left and right images). The chip size limitation imposed by MOSIS has kept the number of pixels in this prototype design at 40×40, although a 64×64 array would fit on a full size 1.2 cm × 1.2 cm 2-micron chip.

At a 10 Mhz clock rate for the prototype CCD processor, a remarkable 2,000 image frame pairs can be processed per second for a disparity range of 10 pixels. This experimental array has a serial input structure for interfacing purposes. However, in a full stereo system the input would be column parallel for both left and right images. In such a case, a 64×64 system would reach 9,000 image frame pairs per second for a 10 MHz clock and a 16 pixel disparity range.

A complete test system has been designed, and one 40×40 chip has been fabricated. Errors in the first chip were corrected in a second design, which is now being fabricated while the test system is under construction.¹¹

**Switched-Capacitor System for Merging Depth and Slope Estimates**

Mark Seidel is working with Professor Wyatt and Thomas F. Knight on a switched-capacitor chip to robustly compute the depth of a scene by merging depth estimates (possibly from a stereo algorithm) with surface slope estimates from a shape-from-shading algorithm. This multisensor image fusion design uses the least-squares properties of certain switched-capacitor systems to compute a dense depth map given noisy or sparse


depth and slope input. A 1-D version has been fabricated through the MOSIS system, and testing will begin soon.\textsuperscript{13}

**CMOS/CCD Imager and Focus-of-Expansion Chip for Camera Motion**

Ignacio McQuirk and Professors Berthold Horn and Hae-Seung Lee are developing a chip that determines the direction of camera translation directly from a time-varying image sequence with a real-time analog VLSI system. Professor Horn’s algorithm assumes a fixed world; there is no restriction on the shapes of the surfaces in the environment, only an assumption that the imaged surfaces have some texture, that is, spatial variations in reflectance. It is also assumed that the camera is stabilized so that there is no rotational motion. The focus of expansion (FOE) is the projection of the camera translation vector onto the image plane and gives the direction of camera motion. The FOE is the image of the point towards which the camera is moving and the point from which other image points appear to be receding. Knowledge of the location of the FOE in the case of pure translation also allows coarse calculation of the depth map of the imaged world up to a scale factor ambiguity, as well as the associated time to impact.

Through extensive simulation on synthetic images, various possible algorithms for estimating the FOE were compared in terms of accuracy, robustness, and feasibility for single-chip analog VLSI implementation with on-chip imaging circuitry. Based on these studies, an algorithm was chosen for implementation.

The method chosen for estimating the FOE is based on least squares analysis. It minimizes the sum of squares of the differences at every picture cell between the observed time variation of brightness and the predicted variation given the assumed position of the FOE. The minimization is not straightforward because the relationship between the brightness derivatives depends on distance to the surface being imaged and that distance is not only unknown, but also varies from picture cell to picture cell. Stationary points, where brightness is constant (instantaneously), play a critical role. Ideally, the FOE would occur at the intersection of the tangents to the iso-brightness contours at these stationary points. In practice, image brightness derivatives are hard to estimate accurately given that the image itself is quite noisy. Hence the intersections of tangents from different stationary points may be quite scattered. Reliable results can nevertheless be obtained if the image contains many stationary points, and the point is found that has the least weighted sum of squares of perpendicular distances from the tangents at the stationary points.

A variety of system architectures and circuit designs were explored, and one was chosen for realization in analog VLSI. McQuirk has designed a preliminary CMOS/CCD version of the FOE chip which estimates the FOE, using on-chip CCD imaging circuitry and a row-parallel processing scheme. The chip will accommodate both electrical and optical input. Electrical input makes it possible to distinguish algorithmic performance from chip performance. The full FOE chip will be tested and characterized with electrical and optical input. A control processor, timing generator, and the required CCD clock drivers will be designed to complete the FOE system.

Several components of the FOE chip have been fabricated through MOSIS in a test chip. Agreement between expected and actual circuit behavior has been excellent. The next steps of the project include testing the CCD FOE chip components, revising the design if necessary, and fabricating the preliminary version of the chip. Required support circuitry for the chip will also be designed and constructed. Then full testing of the system with electrical and optical input will begin.\textsuperscript{14}

**1.2.4 CMOS/CCD Fabrication Process for Vision Chips**

Craig Keast and Professor Charles Sodini have developed a four-phase, buried-channel CCD capability to enhance MIT’s 1.75 micron baseline CMOS process. This enhancement allows CCD operations such as image-to-charge conversion, spatial shifting, analog storage, and many other linear signal processing functions necessary for early vision processing. It uses an additional implant step, a second gate oxidation, a second level of polysilicon, and an additional etching step, adding two more lithographic levels to the 13

---


mask double-metal CMOS process. In addition, this process allows the fabrication of high quality poly-to-poly capacitors with low voltage and temperature coefficients that are useful in many analog MOS applications. The process utilizes a 23 nm gate dielectric for high transconductance and a 42 nm poly-to-poly capacitor insulator. The first run processed yielded functional CCD shift registers and CMOS transistors. The CCDs fabricated in this process now perform well, with a transfer efficiency of about 0.99999 and a dark current of 1nA/cm². This technology is currently being used to fabricate some of the designs in this project.15

1.2.5 Special Circuits for Vision Chips

Switched-Capacitor Resistive Grids

Resistor networks can perform several useful computations for machine vision problems. Most IC processes, however, are not designed for implementing large resistor networks on the image plane. The resistivity values are usually too low to make large-valued resistors in a reasonable area. Chris Umminger and Professor Sodini investigated the use of switched capacitor networks as an alternative to resistor networks for performing these computations. They showed that a mapping can be made from any resistor network to an equivalent switched capacitor network that will have the same node voltage solution in steady state. However, it takes several switching cycles for the charge to be distributed in the SC network before steady state is reached. Their results indicate that SC networks can be used in moderate resolution systems (250×250 pixels) operating at video rate without the settling time of the networks becoming a problem. A chip was fabricated and successfully tested which implemented some switched capacitor networks for image smoothing in one dimension. By varying the clocking scheme applied to the network, they were able to adjust the amount of smoothing performed. It was found that under certain conditions these networks are susceptible to charge-pumping; proper design can circumvent this problem.16

Compact Resistive Fuse Circuits

Resistive fuse circuits were described in the Vision Chip Designs section. The original concept and circuit implementation were created by John Harris at the California Institute of Technology, who has now become a postdoctoral student with this project.

Previous implementations require a large number of transistors, at least 33 per fuse. We have developed several much smaller designs that enable the fabrication of much denser grids.

Steven Decker and Professor Wyatt have designed a 4-FET depletion mode fuse that Steve and Professor Sodini have fabricated in an enhancement to MIT's CMOS process. The device has been tested and works essentially as expected.

Although the four-transistor fuse is quite compact, its behavior is not adjustable and the depletion mode FETs require custom processing. Professor Lee has designed two slightly larger fuses that avoid both these problems. One fuse uses seven small FETs, and the other uses 11; both can be fabricated in a standard CMOS process. Furthermore, both the linear resistance and off-voltage parameters can be varied, offering the flexibility of controlling the space constant and the threshold of smoothing independently. Professor Lee and Paul Yu have designed and fabricated a prototype 11-transistor circuit through MOSIS, which was found to be fully functional. The silicon area for each fuse is 75 μm by 105 μm. We have filed a U.S. patent application for these compact fuse designs.17

Simulation of Analog Vision Chips

Andrew Lumsdaine and L. Miguel Silveira, under the supervision of Professor Jacob K. White, developed the CMVSIM program for simulating grid-based analog circuit arrays on the Connection Machine. The program was designed specifically for robotic vision chips and has built-in functionality for handling images as input and output. CMVSIM has two intended uses. The first


is in simulating algorithms, whereby CMVSIM is used to simulate grids of idealized circuit elements so that a designer can tune the idealized network realization of a particular algorithm. The second, and more important, use is in simulating actual VLSI circuits, where CMVSIM is used to simulate a VLSI circuit at an analog level. It is in this arena that CMVSIM's capabilities are so important, because VLSI realizations of vision circuits can easily have hundreds of thousands of devices, and yet the circuit must be simulated in its entirety at an analog level. Such a capability is well beyond that of a conventional circuit simulation program running on a serial computer. Since CMVSIM can simulate very large circuits in a reasonable amount of time, it allows the incorporation of a simulation phase into the design cycle of grid-based analog signal processors.

### 1.2.6 Theory and Algorithms

#### Settling Time of Large Switched-Capacitor Networks

At least some image processing algorithms formulated as minimization problems can be implemented using networks of switches, capacitors, and independent voltage sources (SC networks). In seeking bounds on the settling time of these networks, Mark Seidel and Professor Wyatt have discovered a method for deriving closed-form zero-state step response bounds for discrete-time systems characterized by M-matrices. This new result provides rigorous bounds that can replace the empirical approximations currently used to estimate settling times in switched-capacitor grids.

#### Texture Synthesis and Halftoning Using Markov Random Fields

Using a novel formulation of the Gibbs energy of a Markov random field in terms of generalized graylevel co-occurrences that he calls *aura measures*, Ibrahim Elfadel, working with Professors Yuille, Horn and Wyatt, has been able to gain a deeper understanding of the textural patterns synthesized by using Gibbs random fields (GRF). In particular, with the use of aura measures, he has generalized the “boundary length” optimization principle of first-order binary MRF models to any order and to any number of graylevels. His work has also led to a new insight into the meaning of co-occurrences: they relate the behavior of the texture pattern during the synthesis procedure to physically meaningful phenomena such as mixing, separation, crystallization and precipitation in multiphase fluids.

Underlying the notion of an aura measure is a new set-theoretic concept called the *aura* set. This concept was introduced to give precise meaning to the intuitive notion of one subset of a lattice being in the neighborhood of another subset.

Even more important is Elfadel's discovery that the descriptive “power” of GRF models could be quite limited if the model parameters are not chosen appropriately. When organized in a matrix indexed by the graylevels, the aura measures form an aura matrix that is also a generalization of the co-occurrence matrix. For instance, detailed analysis supported by numerical simulation for the autobinomial GRF model has shown that the ground state of the isotropic attractive case has a tridiagonal aura matrix, while the isotropic repulsive case has an antitridiagonal matrix. In the general anisotropic case, his results mean that there are grayscale patterns that are impossible to model using the ground states of the autobinomial GRF model.

These results indicate, apparently for the first time, a relationship between a texture synthesis model and its co-occurrence matrix. Moreover, the restrictions imposed on the possible ground state patterns should be taken into account before modeling images with Markov/Gibbs random fields.

Within the framework of GRF image models but along another research direction, Elfadel has shown how methods developed in the context of statistical physics can be used to analyze grayscale texture synthesis procedures based on the Monte Carlo probabilistic paradigm. In particular, using the mean-field equations of the texture GRF, the existence of a bifurcation point indicating the presence of a phase transition in the textural pattern has been demonstrated. Using simulations, he has shown that a number of interesting phenomena occur at the bifurcation temperature, e.g., a sudden decrease in energy, sharp peaks in similarity measures between textural patterns, and sudden saturation of the mean-field variables. The mean-field results allows him to argue that for texture synthesis, it is sufficient to simulate the

---

mean-field equations near the bifurcation temperature.\textsuperscript{19}

**Algorithm for Effortless Least-Squares Camera Calibration**

Accurate camera calibration is important for many of the subsystems we propose to build and for other applications. A major part of this task is accurate determination of internal camera parameters, specifically the exact focal length and the center of projection coordinates that describe the lens position precisely in relation to the focal plane. Existing methods are accurate but slow and laborious because they involve constructing calibration arrays with very accurately-known positions and angles relative to the camera.

Professor Wyatt has discovered that one can, in principle, determine the internal camera parameters by processing images of 3-D spheres without requiring precise knowledge of their sizes or of their positions in relation to the camera. Jesus Dominguez has carried out an initial set of simulations under the supervision of Professors Wyatt and Horn that indicate accuracies of better than 1 pixel on a 454×576 pixel image can be obtained using images of only three spheres. Further work is planned with Gideon Sten to explore the possibilities of this method.\textsuperscript{20}

**Multiscale Method for Image Segmentation and Bandwidth Compression**

Lisa Dron, working with Professor Horn, has invented a robust method of image segmentation that appears to work well in computer simulations. The basic idea is to segment an image between two nodes of a grid wherever the difference in intensities at the two nodes exceeds a threshold at all of several different scales of spatial low-pass filtering. (The threshold can vary with the scale.) The method is robust in that local noise spikes in the image are averaged out with filtering and do not contribute false edges, and it is accurate because the unfiltered image plays a role, since threshold must be exceeded there as well. Bandwidth compression is achieved by retaining intensity values only at nodes adjacent to edges. Intensities elsewhere are obtained by spatial averaging using a resistive grid. The method looks promising for analog VLSI implementation, and Ms. Dron is planning a chip based on these two novel ideas.\textsuperscript{21}

**Extremum Principles for Resistive Fuse Networks**

The operation of resistive fuse networks has previously been explained in the literature on an ad-hoc basis. Professor Wyatt has shown that the operation of these networks can now be understood in terms of a network extremum principle involving a scalar function known as the co-content. This insight enables one to understand the operation of these networks in terms of the graduated nonconvexity algorithm in the computer vision literature. Andrew Lumsdaine, working with Professor Wyatt, has developed continuation methods


(i.e., “deterministic annealing”) that drive these networks toward good solutions. 22

1.2.7 Publications of the MIT Analog Vision Chip Project

Journal Articles


Book Chapters


Conference Papers


---

1.3 Mixed Circuit/Device Simulation

**Sponsors**

IBM Corporation
U.S. Navy - Office of Naval Research
Grant N00014-91-J-1698

**Project Staff**

Andrew Lunsdaine, Mark W. Reichelt, Professor Jacob White, Professor Jonathan Allen

The enormous computational expense and growing importance of mixed circuit/device simulation, as well as the increasing availability of parallel computers, suggest that specialized, easily parallelized algorithms can be developed for transient simulation of MOS devices. In earlier work on the WORDS program (Waveform Overrelaxation Device Simulator), the easily parallelized waveform relaxation (WR) algorithm was shown to be a computationally efficient approach to device transient simulation even on a serial machine, though the WR algorithm typically requires hundreds of iterations to achieve an accurate solution.

In order to use WORDS in a mixed circuit/device simulator, we have been investigating ways of making WORDS more robust and efficient. We determined how to compute the terminal currents accurately using different timepoints at different mesh nodes and improved the timestep selection procedure by determining how to refine the timesteps as WR iterations proceed (reducing the total computation by as much as a factor of 2 by using only a few coarse timesteps in early iterations). The more accurate, electric field dependent mobility model was also implemented.

We found experimental evidence that WR using standard overrelaxation acceleration can produce oscillatory results and are investigating methods for eliminating this phenomenon. A frequency-dependent overrelaxation algorithm using lowpass filtering was developed, as well as a waveform conjugate-direction approach.

---


results indicate that both approaches reduce the number of waveform iterations required by more than a factor of seven.

1.4 Simulation Algorithms for Clocked Analog Circuits

**Sponsors**
IBM Corporation  
National Science Foundation  
PYI MIP 88-58764  
U.S. Navy - Office of Naval Research  
Contract N00014-87-K-0825

**Project Staff**
Luis M. Silveira, Steve Leeb, Professor Jacob White

A challenging problem in the area of analog circuits is the simulation of clocked analog circuits like switching filters, switching power supplies, and phase-locked loops. These kinds of circuits are computationally expensive to simulate using conventional techniques because they are all clocked at a frequency whose period is orders of magnitude smaller than the time interval of interest to the designer. To construct such a long time solution, a program like SPICE or ASTAP must calculate the behavior of the circuit for many high frequency clock cycles. The basic approach to making simulation of these circuits more efficient is to exploit only the property that the behavior of such a circuit in a given high frequency clock cycle is similar, but not identical, to the behavior in the preceding and following cycles. Therefore, by accurately computing the solution over a few selected cycles, an accurate long time solution can be constructed. Such approaches are known as “envelope-following” algorithms.26

In our recent work, we have been trying to make the envelope-following algorithm more robust and efficient by exploiting the fact that the envelope of “quasi-algebraic” components in the solution vector need not be computed. An automatic method for determining the quasi-algebraic solution components has been derived, and experimental results demonstrate that this modified method reduces the number of computed clock cycles when applied to simulating closed-loop switching power converters.26

1.5 Numerical Simulation of Short Channel MOS Devices

**Sponsors**
IBM Corporation  
National Science Foundation  
PYI MIP 88-58764

**Project Staff**
Jennifer A. Lloyd, Khalid Rahmat, Professor Dimitri A. Antoniadas, Professor Jacob White

The model used in conventional device simulation programs is based on the drift-diffusion model of electron transport, and this model does not accurately predict the field distribution near the drain in small geometry devices. This is of particular importance for predicting oxide breakdown due to penetration by “hot” electrons. There are two approaches for more accurately computing the electric fields in MOS devices: one is based on adding an energy equation to the drift-diffusion model, and the second is based on particle or Monte-Carlo simulations.

In the first approach, an energy balance equation is solved along with the drift-diffusion equations so that the electron temperatures are computed accurately. This combined system is numerically less tame than the standard approach, and must be solved carefully. We have developed a two-dimensional device simulator in which an energy balance equation is solved for electron temperature along with the usual drift-diffusion equations. The program avoids temperature instabilities produced by previous discretization approaches through the use of a careful application of exponential-fitting to the energy equation. Drain currents for silicon MOSFETs predicted by the simulator, using one set of model parameters, match well with experimental data for devices over a range of channel lengths from 0.90 μm to 0.16 μm. Also, a method to compute substrate current has been derived which uses the electron temperature provided by the simulator. The computed substrate currents match well with measured data, for the regime


above subthreshold, for MOSFETs with channel lengths as short as 0.16 μm.27

In the area of Monte Carlo device simulation, we are focusing on transient calculations with self-consistent electric fields. Specifically, we are trying to apply the recently developed implicit particle methods on very simple three-dimensional structures.

1.6 Efficient Three-Dimensional Interconnect Analysis

Sponsors
Digital Equipment Corporation
IBM Corporation
National Science Foundation
PYI MIP 88-58764
U.S. Navy - Office of Naval Research
Grant N00014-91-J-1698

Project Staff
Keith S. Nabors, Songmin Kim, Dr. Sami M. Ali, Professor Jacob White

We have developed a multipole-accelerated algorithm for computing the capacitance of complicated 3-D geometries and have implemented the algorithm in the program FASTCAP. The method is an acceleration of the boundary-element technique for solving the integral equation associated with the multiconductor capacitance extraction problem. Boundary-element methods become slow when a large number of elements are used because they lead to dense matrix problems which are typically solved with some form of Gaussian elimination. This implies that the computation grows as n³, where n is the number of panels or tiles needed to accurately discretize the conductor surface charges. Our new algorithm, which is a generalized conjugate residual iterative algorithm with a multipole approximation to compute the iterates, reduces the complexity so that accurate multiconductor capacitance calculations grow nearly as nm where m is the number of conductors. For practical problems which require as many as 10,000 panels, FASTCAP is more than two orders of magnitude faster than standard boundary-element based programs.28

Recent work on FASTCAP includes: (1) improvements to the basic multipole-accelerated algorithm29, (2) linking the FASTCAP program to the solid modeler used in the MIT Micro-Electrical-Mechanical CAD (MEMCAD) system, and (3) adding the capability to analyze problems with conductors embedded in arbitrary piecewise-constant dielectric regions.30 With these improvements, the capacitance calculation method gives more accurate results for more intricate problems and retains the linear time complexity of the basic algorithm. The FASTCAP program and manuals are available from MIT.

Recently, it has been suggested that sufficiently accurate integrated circuit cross-talk simulations can be performed by computing the time evolution of electric fields both inside and outside three-dimensional integrated circuit conductors via a finite-difference discretization of Laplace’s equation. We have derived a boundary-element approach to the same calculation, as boundary-element methods require that only conductor surfaces be discretized, and the difficulty constructing volume meshes associated with finite-difference methods is avoided. Two boundary-element approaches have been investigated, and experiments show that the straightforward approach leads to unacceptable discretization errors, and a less intuitive second approach yields good results even with coarse surface meshes. This boundary-

---


element approach makes it possible to perform true
three-dimensional transient interconnect analyses
in less than a minute on a scientific workstation.\textsuperscript{31}

Currently, we are working on development of a
multipole-accelerated 3-D inductance analysis
program.

1.7 Parallel Numerical Algorithms

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-91-J-1698

Project Staff

Andrew Lumsdaine, Mark W. Reichelt, Luis M.
Silveira, Ricardo Telichevsky, Professor William
Dally, Professor Jacob White

We are taking an application-oriented approach to
developing parallel numerical algorithms and are
focusing on circuit and device simulation. Application
programs and techniques are being developed for both massively parallel SIMD machines
like the Connection Machine, or on MIMD machines like the Intel hypercube. In addition, we
are also trying to understand fundamental aspects
of the interaction between architecture and certain
numerical algorithms.

For example, the direct solution of circuit simu-
lation matrices is particularly difficult to parallelize,
in part because methods like parallel nested dis-
section are ineffective due to the difficulty of
finding good separators. For that reason, general
sparse matrix factorization techniques are being studied, and in particular, the interaction between
sparse matrix data structures, computer memory
structure, and multiprocessor communication is
being investigated. To focus this work, a special-
purpose processor for circuit simulation, the
Numerical Engine (NE), is under development.
Preliminary design is complete, and register-
transfer level simulation results indicate that the
specialized processor can achieve up to 80% of its
peak floating-point performance for sparse matrix
factorization and nearly 90% of its peak perform-
ance on model evaluation.

Explicit integration methods avoid matrix solution
and are therefore also interesting algorithms to use
on parallel processors. For this reason, we investig-
ated some of the properties of the recently devel-
oped explicit exponentially-fit integration
algorithms. The results were not very encouraging,
though the theoretical investigation undertaken
yielded several new insights about these
methods.\textsuperscript{32}

Over the last year, CMVSIM, a program for circuit-
level simulation of grid-based analog signal pro-
cessing arrays on a massively parallel processor
has been used to evaluate several signal processor
architectures. CMVSIM uses the trapezoidal rule
to discretize the differential equations that describe
the analog array behavior, Newton’s method to
solve the nonlinear equations generated at each
time-step, and a block conjugate-gradient squared
algorithm to solve the linear equations generated
by Newton’s method. Excellent parallel perform-
ance is achieved through the use of a novel, but
very natural, mapping of the circuit data onto the
massively parallel architecture. The mapping takes
advantage of the underlying computer architecture
and the structure of the analog array problem.
Experimental results demonstrate that a full-size
Connection machine can provide a 1400 times
speedup over a SUN-4/280 workstation.\textsuperscript{33}

An easily parallelized approach to solving large
systems of time-dependent differential equations is
the waveform relaxation algorithm. In such an
approach, the differential equation system is
broken into subsystems which are solved inde-
dependently, each subsystem using guesses about
the behavior of the state variables in other subsys-
tems. Waveforms are then exchanged between the
subsystems, and the systems are resolved with,
hopefully, improved information about the other
subsystems. This process is continued until con-
vergence is achieved. Recent work on theoretical
aspects of these methods has answered several
long-standing questions about multirate stability.\textsuperscript{34}

\textsuperscript{31} D. Ling, S. Kim, and J. White, “A Boundary-Element Approach to Transient Simulation of Three-dimensional In-
grated Circuit Interconnect,” \textit{Proceedings of the 29th Design Automation Conference}, Anaheim, California, June


\textsuperscript{33} A. Lumsdaine, \textit{Theoretical and Practical Aspects of Parallel Numerical Algorithms for Initial Value Problems, with
Also, a new conjugate-direction based acceleration to the waveform relaxation algorithm has been derived and applied to solving the differential-algebraic system generated by spatial discretization of the time-dependent semiconductor device equations. Numerical experiments demonstrate that when used to perform two-dimensional transient simulations of MOS devices, the waveform conjugate-direction method is more than an order of magnitude faster than standard waveform relaxation.35

1.8 Microelectromechanical Computer-Aided Design

Sponsors
Federal Bureau of Investigation
National Science Foundation
Fellowship MIP 88-58764

Project Staff
Mattan Kamon, Professor Martin A. Schmidt, Professor Stephen D. Senturia, Professor Jacob White

High fabrication costs and increasing microsensor complexity is making computer simulation of the realistic geometries necessary, both to investigate design alternatives and to perform verification before fabrication. At MIT, we are developing a MicroElectroMechanical Computer-Aided Design (MEMCAD) system to make it possible for microsensor designers to easily perform realistic simulations. Carefully selected commercial software packages have been linked with specialized database and numerical programs to allow a designer to enter easily a three-dimensional microsensor geometry and perform quickly both mechanical and electrical analysis. The system currently performs electromechanical analyses, such as calculating the capacitance versus pressure (or force) curve for both a square diaphragm deformed by a differential pressure and a rectangular beam deflected by a centrally applied force.36

1.9 Techniques for Logic Synthesis, Formal Verification and Testing

Sponsors
IBM Corporation
Grant/Fellowship
U.S. Navy - Office of Naval Research
Contract N0014-87-K-0825

Project Staff
Professor Srinivas Devadas, James H. Kukula, Kelly S. Bai, Stan Y. Liao, Amelia H. Shen, Filip J. Van Aelten

1.9.1 Introduction

Integrated circuits, whether they are manually designed or automatically synthesized, must meet many different specifications. For example, limits on circuit delay, layout area, and power consumption are typically placed. Reliability restrictions are also placed, for example in the form of noise margins or non-susceptibility to electron-migration-induced failure. The designed circuit also has to be checked for correct functionality and whether it meets its performance goals. The fabricated circuit has to be tested in order to ensure that the fabrication steps are error-free.

At MIT, we are working on various aspects of automated synthesis, as well on design verification and test problems. We are developing logic optimization methods that target not only the area, performance and power consumption of the synthesized design, but also the testability of the designed circuit. We are developing verification mechanisms at the behavioral, logic and circuit


schematic levels that can check whether the circuit implementation has the correct logical functionality and meets the desired performance requirements prior to embarking on the fabrication process. Finally, we are developing methods for efficient test generation at the logic level, so fabricated integrated circuits can be easily tested for correctness.

In the following, we describe our work in the areas of timing, circuit, and behavioral verification, sequential logic synthesis for area and performance, combinational logic synthesis for low power, and test generation.

1.9.2 Timing Analysis and Verification

Over the past year, several projects have been undertaken in the areas of timing analysis and verification. Timing analysis tools are essential to the performance verification problem. The tools we have developed operate on logic-level descriptions of an integrated circuit. The work described in this section has been done in collaboration with Professor Sharad Malik of Princeton University, Dr. Kurt Keutzer and Dr. Albert Wang of Synopsys, Inc.

Timing Simulation

Timing simulation is a very common method for conclusively verifying the timing behavior of a design.

We have shown that in simulating a vector pair in a combinational circuit, the number of events at a gate is bounded above by the number of paths from the primary inputs to that gate. We have discovered a family of circuits, parameterizable by the depth of the circuit, in which the number of events at a gate is precisely the number of paths from the primary inputs to that gate. Thus, the number of events at the output of a circuit can be exponential in either the size or depth of a circuit. This shows that event-driven timing simulation can have exponential complexity, and this fact has been observed in practice on widely-used timing simulators.

We made the observation that to answer the question: “Does an event occur in this circuit at any time at or after τ?,” we do not need a full history of the circuit activity before time τ. In other words, an event-driven timing simulation approach is likely to compute much more information, in the form of events, than is necessary to resolve the question at hand. To fully explore the ramifications of this observation we have developed a calculus of event simulation and applied this calculus to different models of timing behavior. Using these results we can reduce the number of events that need to be evaluated in a circuit simulation. The upper bound on the number of events at a gate g that need to be evaluated to determine if an event occurs after time τ is equal to the number of paths from primary inputs to g of length greater than or equal to τ. For example, on the circuit family mentioned above, the number of events evaluated at a gate is reduced exponentially by our procedure. Over a large class of circuits the number of events required to simulate random vector pairs was reduced by a factor of 2 to 10.

False Path Identification

We have addressed the problem of accurately computing the delay of a combinational logic circuit in the floating mode of operation. (In this mode the state of the circuit is considered to be unknown when a vector is applied at the inputs.) It is well known that using the length of the topologically longest path as an estimate of circuit delay is a poor estimate of circuit delay since this path may be false, i.e., it cannot propagate an event. Thus, the true delay corresponds to the length of the longest true path. This forced us to examine the conditions under which a path is true. We introduced the notion of static co-sensitization of paths which leads to the necessary and sufficient conditions for a path to be true. We have applied these results to develop a delay computation algorithm with the unique feature of being able to determine simultaneously the truth or falsity of entire sets of paths. This algorithm uses conventional stuck-at-fault testing techniques to arrive at a delay computation method that is both correct and computationally practical even for particularly difficult circuits.

Statistical Analysis

An important problem in high performance circuit design is to predict the percentage of fabricated circuits that will achieve a certain performance level (or clock speed). Accurate and efficient means of answering this question have not been developed thus far.

We have developed efficient methods to compute an exact probability distribution of the delay of a combinational circuit, given probability distributions for the gate and wire delays. The derived distribution gives the probability that a combinational circuit will achieve a certain performance, across the possible range. Then this information can be used to predict the expected performance of the entire circuit. We have defined the notion of a correct approximation, based on convex inequality, which never overestimates the percentage
of circuits that will achieve any given performance. We have shown that, given the assumption that all the topologically longest paths are responsible for the delay, our computation technique provides a correct probabilistic measure in the sense given above.

Nevertheless, false paths in a circuit may result in needlessly pessimistic probability distributions. Having defined a notion of falsity of paths when dealing with probability distributions on gate and wire delays, we give methods to identify and to ignore false paths in our probabilistic analysis to obtain correct and accurate answers to the performance prediction question.

Certified Timing Verification

The floating delay of a circuit is the delay under a single-vector static analysis condition that makes conservative assumptions about the state of the circuit before the single vector is applied (see the section above on False Path Identification). The transition delay of a circuit is the delay under a multivector dynamic-analysis condition that makes no assumptions about the state of the circuit before the vector sequence is applied but does allow for the possibility for monotonic speed-up at gates. Recently it has been shown that the transition delay of a circuit is bounded above by the floating delay, and it has been conjectured that this bound is tight in the presence of monotonic speed-ups. We have refuted this conjecture and demonstrated for the first time that the transition delay of a circuit can differ from the floating delay even in the presence of arbitrary monotonic speed-ups in the circuit.

We have used this result to motivate our derivation of a procedure which directly computes the transition delay of a circuit. The result of this procedure is a vector sequence which excites an event along the longest sensitizable path of the circuit under consideration. The most practical application of this result is in certified timing verification. In this scenario, the vector sequence generated by the transition delay calculator is applied by a timing simulator with more accurate timing models to certify the results of static timing verification. Such a procedure promises to give the high accuracy of timing simulation with the computational efficiency and comprehensive path coverage of static timing verification.

1.9.3 Asynchronous Circuit Verification

We have addressed the problem of verifying that the gate-level implementation of an asynchronous circuit with given or extracted bounds on wire and gate delays is equivalent to a specification of the asynchronous circuit behavior described as a classical flow table.

We have developed a procedure that extracts the complete set of possible flow tables from a gate-level description of an asynchronous circuit under the fixed or bounded wire delay models. Given an extracted flow table and the initial flow table specification, we have implemented procedures to construct a product flow table to check for machine equivalence under various modes of operation.

The modes of operation we have considered are fundamental mode and non-fundamental mode. We have considered single output change and multiple output change flow tables, as well as single input change and multiple input change flow tables. Flow table extraction and equivalence checking procedures have been tailored for each mode of operation and each type of asynchronous flow table.

1.9.4 Behavioral Verification

The problem of implementation verification at the behavioral level has proven to be difficult because of the lack of a strong underlying theory similar to automata or switching theory. Pipelined, parallel and serial implementations of a given behavioral description can have dramatically different automata-theoretic terminal behavior.

String Theory

We have used string function theory to develop an efficient and formal methodology for the verification of logic implementations against behavioral specifications.

We have defined six primitive relations between string functions, other than strict automata equivalence, including delay, don’t care times, parallelism, encoding, input don’t care and output don’t care relations. These relations have attributes; for instance, the delay relation has an attribute corresponding to the number of clock cycles of delay. Using string function theory, we have derived associated circuit transformations for each of these primitive relations and formally proven that the correspondence is correct.

During synthesis, these transformations can be applied in sequence to a circuit specification to
produce a logic-level implementation. Given an arbitrarily long sequence of behavioral transformations that correspond to primitive relations occurring in arbitrary order, we have shown that we can always produce a composite relation, wherein each primitive relation occurs at most once in a predetermined order, which also relates the specification and the implementation. We can derive a set of attributes for the composite relation, given the attributes in the transformation sequence. This is possible because we can prove that the primitive relations satisfy precedence and transitivity properties.

We have shown how we can apply the composite transformations with the appropriate attributes to the behavioral specification and logic implementation and verify the resultant behavioral specification against the resultant logic implementation with one automata equivalence check, using any currently available method. If the resultant specification and implementation are strictly automata equivalent, then the implementation is deemed to implement correctly the specification. Preliminary experimental results have been obtained that indicate that the procedure is viable for medium-sized circuits.

**Compositional Verification**

The strategy described in the previous section is only viable for medium-sized circuits because an equivalence check between two automata is required. The number of states in the automata can be astronomically large since both data and control memory elements are treated uniformly and the number of states grows exponentially with the number of memory elements.

We have developed a compositional verification technique which allows for the efficient verification of the specific class of systems with synchronous globally timed control. In such systems, a single controller directs the data through the data path and decides (globally) when to move the data. These systems are verified against a specification in an applicative language such as SILAGE, which can be viewed as describing a maximally parallel synchronous machine. Both implementation and specification have string functions (input/output mappings) associated with them, and correctness is taken to mean that a certain relation holds between these string functions. These relations extend beyond strict input/output equivalence, and provide room for various behavioral transformations.

Sufficient conditions for the validity of these relations have been developed, which can be verified efficiently. They consist of the combinational correctness of the data path modules and the validity of a number of past-tense Computation Tree Logic (CTL) formulae with respect to the controller. The latter can be verified using forward finite state machine (FSM) traversal algorithms. In this way, the verification of the composite machine is reduced to the separate verification of data path modules and controller.

**1.9.5 Sequential Logic Synthesis**

Problems in the area of synthesis of sequential machines from State Transition Graph descriptions and the optimization of logic-level sequential circuits have been addressed. The goal is to obtain logic-level implementations that are minimal in terms of area and/or delay.

**Theoretical Results on State Encoding**

A prime-factorization-based approach to encoding \( N \)-state counters, where \( N \) can be arbitrary, has been developed. First, an analytical expression for the number of cubes in an optimal two-level implementation of a \( 2^k \)-state counter under a radix-2 encoding was derived. A lower bound for the number of cubes in a radix-2 encoding of an arbitrary \( N \)-state counter was also derived.

We have shown that there is a deep and interesting relationship between the prime factorization of a whole number \( N \), and encoding a \( N \)-state counter. Factoring the number \( N \) into prime factors and encoding the smaller submachines separately can result in an encoding that requires substantially fewer cubes than a radix-2 encoding. Analytical expressions for the number of cubes in a prime-factorization-based encoding of a \( N \)-state counter, as well as upper bounds for the ratios of the number of cubes in a prime-factorization-based encoding to a radix-2 encoding for classes of counters have been derived.

Prime-factorization-based encoding algorithms provably perform better than radix-2 encoding for counters, and also lend insight into the algebraic structure of sequential machines.

**Finite State Machine Decomposition**

Current state assignment strategies almost exclusively employ a two-step strategy, where in the first step relationships are derived between states, and in the second step states are encoded taking into account these relationships. We have developed methods for state assignment based on the premise that optimal state assignment corresponds to finding an optimal general decomposition of a finite state machine.
We have arrived at two methods, the first of which performs multiway decomposition on the initial State Transition Graph (STG) based on the notion of transition pairs. The resulting submachines are encoded using constraint satisfaction methods. In the second method, we incrementally construct a state encoding for the given machine, one bit at a time. A $2^n$-state machine is encoded in $n$ steps, by repeated general decomposition of a $2^i$-state machine into a $2^{i-1}$-state submachine and a 2-state submachine. The 2-state submachine is encoded — an extra bit is added to the incremental encoding, and the process continues with the decomposition of the $2^{i-1}$-state submachine.

Experimental results validate the use of decomposition-based techniques to solve the encoding problem.

**Sequential Optimization at the Logic Level**

Many sequential logic circuits are directly designed at the logic level without recourse to State Transition Graph descriptions. For this reason, it is important to be able to operate upon logic-level descriptions of finite state machines and optimize their area or performance characteristics.

We are in the process of investigating use of re-encoding techniques for sequential circuit optimization. In this method, we identify a good encoding for the sequential circuit by using symbolic encoding algorithms that operate on a symbolic description of the circuit which has been extracted from its logic-level implementation. Once this encoding has been identified, we re-encode the circuit, by adding encoding and decoding circuitry. We then re-optimize the circuit using combinational logic optimization algorithms.

The above process may be time-consuming for large circuits. We are developing latch selection algorithms that allow us to focus on the control logic of a design, and extract symbolic information for portions of the design, rather than the entire design.

**1.9.6 Combinational Logic Synthesis for Low Power**

Very little work has been done in the automated synthesis of circuits for low power consumption. While circuit style selection and power supply voltage reduction affect power dissipation considerably, given a specification of desired logical function, different logic-level implementations of the specification can have significantly different power consumptions for a particular circuit style and power supply voltage. In order to arrive at synthesis strategies that minimize power dissipation, we first have to develop efficient power estimation methods.

**Probabilistic Power Dissipation Analysis**

The work described here has been done in collaboration with Professor Jacob White, Dr. Kurt Keutzer of Synopsys, Inc., and Dr. Abhijit Ghosh of Mitsubishi Laboratories.

We have addressed the problem of estimating the average switching activity of VLSI sequential circuits, under random input sequences. This measure can be used to (1) gauge the power dissipation of the sequential circuit and (2) make architectural or design-style decisions during the VLSI synthesis process.

Switching activity is strongly affected by gate delays and, for this reason, we have used a general delay model in estimating switching activity. This model takes into account glitching at gate outputs in the circuit and models the inertial delay of a gate. Our method takes into account correlation caused at internal gates in the circuit due to reconvergence of input signals. In combinational circuits, uncorrelated input patterns with uniform switching rates are typically assumed, simplifying the problem of switching activity estimation. However, in sequential circuits the input sequence applied to the combinational portion of the circuit is highly correlated because some of the inputs to the combinational logic are flip-flop outputs representing the state of the circuit. We have developed exact and approximate methods to probabilistically estimate switching activity in sequential circuits that automatically compute the switching rates and correlations between flip-flop outputs.

**Synthesis for Low Power**

In our experiments with probabilistic power analysis, we noted that power consumption depended quite significantly on the gate-level structure of a combinational or sequential circuit. We then became interested in gauging exactly what logic structures corresponded to low power realizations. With this information we can tailor a logic optimization strategy to produce circuit topologies that consume the least amount of power.

In particular, we have established for dynamic CMOS circuits that are precharged low (high) that implementing AND (OR) gates near the inputs of the circuit and OR (AND) gates near the circuit outputs results in lower power consumption. We are currently precisely characterizing our all obser-
vations and developing a common subexpression extraction technique (for use in multilevel logic minimization) that minimizes the power consumption of a combinational circuit, while attempting to reduce the area as well.

1.9.7 Testing and Synthesis for Testability

Work in the test generation of sequential circuits is being pursued, as well as the synthesis of combinational circuits for easy testability.

Sequential Test Generation

A new approach has been developed for the state justification and differentiation processes in test generation for sequential circuits. In previously proposed algorithms, the justification and differentiation processes employ the enumeration of Boolean product terms or cubes. Cube enumeration, however, is not efficient. This is because when a Boolean function is represented by cubes, the number of cubes may grow unreasonably large. With this representation, only one cube can be enumerated at a time. This results in a depth-first search for justification sequences. Thus, the operations are very time-consuming.

Ordered binary decision diagrams (OBDDs) provide an alternative representation for Boolean functions. It has been shown that many Boolean functions and operations benefit from this representation. The algorithms we have developed use OBDDs for the justification and differentiation processes in sequential test generation. With OBDDs, the search for justification sequences can be done in a breadth-first manner or depth-first manner. In the case of breadth-first search, all cubes of a given Boolean expression can be searched at once. Experimental results demonstrate the superior efficiency of this approach over previous techniques. Speed improvements of over an order of magnitude have been achieved on industrial benchmarks.

Synthesis for Random Pattern Testability

Built-In Self Test (BIST) techniques are becoming increasingly popular. In order to use BIST efficiently, it is important that the design be amenable to random testing. Typically, designs are not very testable by random patterns. Our solution is to guide the logic optimization process to produce circuit topologies that are inherently random pattern testable.

The random pattern testability of a design is influenced by the probabilities that a random vector produces a 1 or a 0 at each wire in the circuit. In order for the design to be randomly testable, it is intuitive that each wire in the circuit be easily controlled to a 0 and easily controlled to a 1. For instance, if the probability that a wire is at a 1 on a randomly applied input vector is strictly zero, it means the wire is redundant for a stuck-at-0 fault. An easily randomly testable design will have close to 0.5 probabilities for each wire in the circuit being a 1/0 on a randomly applied input vector. We are compiling a set of logic transformations that satisfy this condition, and are developing a logic optimization strategy that uses these transformations to improve the random testability of a given design.

1.9.8 Publications

Journal Publications


Ghosh, A., S. Devadas and A.R. Newton. "Test Generation and Verification for Highly Sequen-


Conference Proceedings


