Integrated Circuits and Systems

- *Compression for Wireless Video Transmission*

- *Reconfigurable Architectures for Energy Efficient, Algorithm Agile Cryptography*

- *Low Power DSP for Biomedical Signal Processing*

- *Low Power DCT Core Processor*

- *Low Power Radio for Wireless Sensor Networks*

- *Low Power Beamforming for Wireless Networked Sensors*

- *Vibration-to-Electric Energy Conversion*

- *Energy Aware Routing Protocols for Distributed Microsensor Networks*

- *Energy Efficient Filtering Using Adaptive Precision and Variable Voltage*

- *High Speed, Low Skew Distributed Clocking*

- *Intelligent Transportation Systems*

- *Cost-Effective Hybrid Vision System for Intelligent Cruise Control*

- *Fusion of Machine Vision and Radar Sensors*

- *Recognition of Three-dimensional Images without Decompression*

- *Automatic Brightness Adaptive CMOS Imager System*

- *Application of Pixel-Parralel Image Prioocessing Chip for Intelligent Vehicle*

- *A Programmable, Wide Dynamic Range CMOS Imager with On-Chip Automatic Exposure Control*

- *Characterization of CMOS Photodiodes for Image Sensor Applications*

- *A Differential Passive Pixel Image Sensor*

- *Transistor-level Synthesis of Delta-Sigma Converters*

Integrated Circuits and Systems

continued

- *Superconducting Bandpass Delta-Sigma A/D Converter*

- *Substrate Noise Shaping in Mixed-Signal Systems*

- *Oversampled Pipeline A/D Converters with Mismatch Shaping*

- *A Nyquist-rate Pipelined Oversampling Delta-Sigma A/D Converter*

- *Low Voltage, Low Power CMOS Operational Amplifier Design for Switched Capacitor Circuits*

- *Low Power Reconfigurable Analog-to-Digital Converter*

- *Ultra Low Power Wireless Sensor Project*

- *Low Power 1.8 GHz Frequency Synthesizer Capable of 2.5 Mbit/s Modulation*

- *Automatic Calibration of Modulated Frequenccy Synthesizers*

- *High Data Rate 5.8GHz Wireless Network*

- *Beam Steering for Increased Transmission Power Efficiency in Mobile Devices*

- *Real-time Video Network*

- *Electromigration in Single Crystal Interconnects*

- *Adaptive Body Biasing*

- *Three-Dimensional Integration: Analysis, Modeling and Technology Development*

- *The Effects of Thermal History on the Stress and Reliability of IC Interconnects*

- *A Framework for Collaborative, Distributed Web-based Design*

Compression for Wireless Video Transmission

Personnel
T. Simon (A. P. Chandrakasan)

Sponsorship

Video compression is an integral part of the digital signal processing of the Ultra Low Power Wireless Sensor Project. Image compression, or more generally compression of any source data, is used to minimize system power by trading local computation for transmitted bandwidth. The goal of this project is to optimize the power efficiency of compression at the algorithmic, architecture, and circuit levels.

The image compression work must produce an algorithm/architecture capable of optimizing system wide power for widely differing output bit rates. During, possibly long, periods of no motion in the image source, the system power is determined by the computational cost of determining there is no movement, the operation of the sensor and converter, and the standby losses of all modules and power supply. Besides minimizing standby losses in individual modules through architectural and circuit techniques, system power may be reduced by the image compression module by adaptive control of frame rate, bit resolution, and feedback regarding optimal operating voltages to the power supply.

During periods of motion in the image source, transmission power becomes a significant part of system power. The compression module must achieve good compression performance without using undue computational power. The baseline algorithm uses hierarchical wavelet transformation and zero-tree based significance coding, with compromises designed to localize communication and limit the scope of filtering in the time dimension to reduce frame memory requirements.

The architectural and circuit work is centered around a massively parallel SIMD array of very simple computational/memory cells. The SIMD array is architecturally well suited to running the compression algorithm with great area and power efficiency. The parallel array enables a very low clock rate and hence a low supply voltage.

The low power video encoder has been fabricated and tested. The test shows that the encoder is fully functional, with performance well within simulated bounds. The measured power dissipation is 400-800 µW for 30 frames/sec and 128x128 resolution. The power dissipation varies, as expected, with the requested video quality and the motion content in the video stream.

Reconfigurable Architectures for Energy Efficient, Algorithm Agile Cryptography

Personnel
J. Goodman (A. P. Chandrakasan)

Sponsorship

In the past there have been several standards for implementing various asymmetric techniques such as the ISO, ANSI (X9.*), and PKCS standards. The variety of standards has resulted in a multitude of incompatible systems that are based upon different underlying mathematical problems. The IEEE P1363 Standard for Public Key Cryptography, which is currently under development, recognizes three distinct families of problems upon which to implement asymmetric techniques: integer factorization, discrete logarithms, and elliptic curves. Each family has its advantages and disadvantages (e.g., IF and DL have been around for many years, allowing them to be thoroughly scrutinized for flaws, whereas EC appears to be much more resilient to cryptanalytic attacks but is still relatively new so users should be less willing to trust it).

As a result, system developers have had to either utilize software-based techniques in order to achieve the algorithm agility required to maintain compatibility, or utilized special purpose hardware and restrict themselves to only providing secure communications with compatible systems. However, software-based approaches lead to very computationally intensive implementations that are very energy-inefficient. In the past these inefficiencies could be ignored as the typical user operated from a fixed-location system such as a desktop computer which has no energy constraint. These assumptions break down with the migration to portable, battery-operated nomadic computing terminals, requiring us to re-evaluate the use of a software-based implementation. Hardware-based implementations on the other hand, while being very energy and computationally efficient, are very inflexible and capable of supporting only a single type of asymmetric cryptography.

We attempt to compromise between these two extremes by taking advantage of the fact that the range of required operations is small enough that we can develop domain specific reconfigurable hardware that is capable of implementing the various algorithms. Furthermore, we do this in an energy-efficient manner that enables us to operate in the portable, energy-constrained environments where this algorithm agility is required most of all.

The resulting implementation is known as the Domain Specific Reconfigurable Cryptographic Processor (DSRCP). The DSRCP differs from conventional reconfigurable implementations (e.g., Field Programmable Gate Arrays (FPGA)) in that its reconfigurability is limited to the domain of asymmetric cryptography. This domain requires only a small set of configurations for performing all of the required operations over all possible problem families as defined by P1363. As a result the reconfiguration overhead is small in terms of performance, energy efficiency, and reconfiguration time.

Low Power DSP for Biomedical Signal Processing

Personnel

R. Amirtharajah (A. Chandrakasan)

Portable systems that depend on batteries have a limited operating life and are prone to failure at inconvenient times. We propose a system that uses ambient energy as a power source for a DSP that processes sensor data. The sensor algorithm is power scalable to trade off performance for system power. An example of power scalable processing is a low power DSP for physiological monitoring. The biomedical sensor is a microphone for recording heartbeats, breathing sounds, and voice data. This data will eventually be used to determine the physical condition of the wearer. The first step is detection of the heartbeats, which can be used to determine heart rate as the basis for a physiological assessment.

Evaluation of the spectrogram of the acoustic data indicates that most of the energy from heartbeat sounds lies in the low frequency range, below 200 Hz. We developed a classifier based approach to heartbeat detection that takes advantage of this spectral characteristic to improve detection performance in the presence of speech and other high frequency energy.

A discrete-time matched filter is implemented using the Distributed Arithmetic Unit. Its output is then passed to a nonlinear filtering unit to calculate quantities used in segmentation. The final segmentation, feature extraction, and classification is performed by the programmable microcontroller at the end to produce the class assignment z. The buffer provides a mechanism for synchronization between the front end filtering and the backend processing. This is necessary for power reduction. The filtering front end must be running continuously to process the input samples, which arrive at a fixed rate. However, the back end classification only needs to be performed for every segment, not every input sample. The system operates as follows: first, the front end filters the input and writes important results to the buffer. A small loop is continuously executed in the microcontroller, checking to see if a full segment has been written to the buffer. The filtering units could do

this, but it involves adding circuits that already exist in the ALU of the microcontroller, which is idle anyway while it is trying to detect a segment. To conserve area, we use the microcontroller rather than add complexity to the filter functional units. When a segment is detected, the microcontroller executes the feature extraction and classification code on the data in the buffer that was just written. Architectural simulation of the DSP chip using Verilog shows that 99.8% of the algorithm time is spent executing the matched filtering and other preprocessing functions. These computations therefore dominate both the time and power consumption of the algorithm.

The chip is currently in fabrication and will demonstrate the feasibility of doing practical signal processing using self-power techniques.

Low Power DCT Core Processor

Personnel

T. Xanthopoulos and R. K. Min (A. Chandrakasan)

A DCT (Discrete Cosine Transform) chip targeted to low power video (MPEG2 MP@ML) and still image (JPEG) applications has been designed and fabricated. The chip exhibits two innovative techniques for arithmetic operation reduction in the DCT computation context along with standard voltage scaling techniques such as pipelining and parallelism. The first method exploits the fact that image pixels are typically well correlated and exhibit a certain number of common most significant bits in local areas. These bits constitute a common mode DC offset that only affects the computation of the DC DCT coefficient and is irrelevant for the computation of the higher spectral coefficients. The DCT chip uses adaptive-bitwidth distributed-arithmetic computation units that reject the common most significant bits for all AC coefficient computations, resulting in arithmetic operations with reduced bitwidth operands thus reducing switching activity. We call this method MSB rejection (MSBR).

The second method exploits the fact that in image and video compression applications, DCT is always followed by a quantization step which essentially reduces the precision of the visually insignificant higher frequencies. The DCT chip allows the user to set up to four different classes of precision for each spectral coefficient on a row-by-row basis so that no unnecessary computation is performed if the precision will be lost anyway due to quantization. A row-column peak-to-peak detector classifies each block row and column into one of four classes of computation precision for maximizing image Peak SNR (PSNR) and minimizing the number of arithmetic operations. We call this method Row-Column Classification (RCC).

The chip has been fabricated in a 0.6 um triple metal CMOS process. It is fully functional and dissipates 4.3 mW at 1.5V, 14 MHz.

A digital system showcasing these video coding chips was designed and implemented. This hardware-based system captures and buffers full-motion video data, routes the data through the DCT and inverse DCT devices in real-time, and displays the resulting video stream. The demonstration system is implemented with off-the-shelf components, including an NTSC video decoder, RAM and ROM memories, programmable logic devices, and an LCD display. Control logic written in VHDL handles the flow of real-time video data through the system, coefficient quantization, synchronization signals for the LCD, and an I2C serial bus interface. The system is contained in a single printed circuit board for simple, portable demonstrations.

Low Power Radio for Wireless Sensor Networks

Personnel
S. Cho (A.P. Chandrakasan)

Sponsorship
ARL Advanced Sensors Federated Lab, Center for Integrated Circuits and Systems

Wireless distributed microsensor systems will enable the reliable monitoring of a variety of applications that range from medical and home security to machine diagnosis and chemical/biological detection. The transmission distance of microsensors can be significantly shorter (<10m) than conventional handheld devices. Most sensing applications will also require very low data rates compared to conventional multimedia traffic.

The communication module of a wireless sensor must be designed for low duty cycle activity. For short range transmission (e.g., <10m) at GHz carrier frequencies, the power is dominated by the radio electronics (frequency synthesizer, mixers, etc.) and not the actual transmit power. In order to save power in the radio module, the electronics must be turned off during idle periods.

Unfortunately, frequency synthesizers require a significant overhead in terms of time and energy dissipation to go from the sleep state to the active state. For short packet sizes, the transient energy during the start-up can be significantly higher than the energy required by the electronics during the actual transmission.

We are working on efficient techniques to transmit short packet sizes. We are looking at a variety of approaches that range from frequency synthesizers that have a fast transient response to open loop approaches. Efficient modulation schemes for short range transmission are also being explored. The goal is develop an architecture that provides more than an order of magnitude reduction in power compared to conventional approaches.

Low Power Beamforming for Wireless Networked Sensors

Personnel
A. Wang (A.P. Chandrakasan)

Sponsorship
Lucent Fellowship, Center for Integrated Circuits and Systems

Sensor collaboration in a network of sensors can provide overall energy savings and improved signal quality. One scenario of sensor collaboration is as follows. Within a sensor cluster, individual sensors detect an event, and transmit their data to the local cluster head. At the cluster head, the data is aggregated before transmitting the result to the distant basestation or end-user. In this scenario, transmission energy is conserved since only the cluster head sensor is required to transmit the sensor data large distances. Also, data aggregation can improve signal quality if appropriate signal processing of the sensor data is done at the cluster head. Therefore, there is a trade-off between (1) data aggregation computation at the cluster head and (2) improved signal quality and lowered transmission energy requirements.

The type of data aggregation algorithms we are considering are called "blind beamforming" algorithms. These algorithms are applicable in sensor network scenarios where the sensor locations and source locations are either partially or completely unknown. Also the sources may be farfield or nearfield and their signal signatures may have narrowband or broadband spectra. These algorithms have typically be used in multiple antenna applications in the radar, sonar, and wireless communication areas.

We have benchmarked two data aggregation algorithms on the StrongARM processor. The first is the Maximum Power beamforming algorithm. This algorithm uses the correlation information between the sensor data to amplify the signal with highest peak power spectral density. The beamformer output provides the maximum power solution. The second algorithm is the Least Mean Square (LMS) algorithm for broadband signals. This algorithm combines the sensor data in order to minimize the mean squared error between the desired signal and the beamformer output. Results have shown that the Maximum Power beamforming algorithm consumes 10x more power than the LMS algorithm. Simulations show that both algorithms have similar performance under the same signal-to-noise ratio conditions. We are currently looking at low-power architectures and circuits to implement the LMS algorithm in VLSI.

Vibration-to-Electric Energy Conversion

Personnel
R. Amirtharajah, S. E. Meninger and J. O. Mur-Miranda (J. H. Lang and A. P. Chandrakasan)

We are developing a method for converting ambient vibration energy into electric energy for general purpose consumption. Such energy conversion is useful for powering autonomous electronics such as remote sensors. The extremely low duty cycle of these electronics pushes their power requirements into the microwatt range. Thus, self-powered systems based on harvesting ambient energy become viable alternatives, eliminating the need for batteries and creating low-maintenance environmentally-friendly autonomous systems.

Our method has two key components which are under development as an integrated system. The first component is a MEMS structure comprising a proof mass, its suspension and a variable capacitor. The capacitor is the actual energy converter; by placing charge on the capacitor and then moving it apart, mechanical energy can be converted into electrical energy which can then be stored for general purpose use. The second component is the combined power and control electronics which excite and control the energy conversion process. These electronics are under design following the paradigm of very-low-power electronics since their power consumption is a tax on the energy converter.

The energy conversion system is shown in block diagram form in Figure 5. The mechanical subsystem is modeled as a vibration source which couples into the electrical subsystem through the MEMS variable capacitor. Low-power electronics direct the energy conversion, and supply power to the load. The electronics consist of power electronics which are responsible for exciting the capacitor through its energy conversion cycle, and a digital control core which generates timing pulses to drive the gates of the power FETS in the power electronics.

Last year, we developed models for the system, and used them to design the system for maximum power output per system area. A vibration monitoring application served to guide this optimization. This year, we began fabricating the MEMS structure and the control and power electronics. An initial version of electronics has now been fabricated with a 0.6-micron CMOS technology. The electronics appear to function properly, and consume several microwatts of power in steady state. We expect this will leave approximately ten microwatts of power for general purpose use, which is adequate for the signal processing required by the bearing monitoring application.)
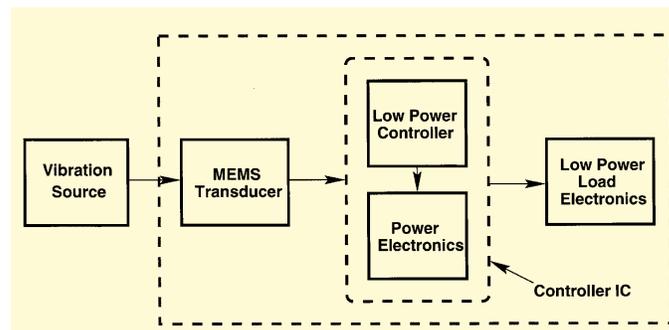

Fig. 5

Energy Aware Routing Protocols for Distributed Microsensor Networks

Personnel

W. Rabiner (A. P. Chandrakasan and H. Balakrishnan)

Wireless distributed microsensor systems will enable reliable monitoring of the environment for a variety of applications including medical and home security, machine diagnosis, and chemical/biological detection. Microsensor networks differ from traditional wireless networks in several important ways. First, microsensor networks typically contain hundreds or thousands of sensing nodes, many more nodes than traditional wireless networks or macrosensors. It is desirable to make these nodes as cheap and energy-efficient as possible and rely on their large numbers to obtain high quality results and achieve fault tolerance in the presence of individual node failure. Second, microsensor networks usually do not require point-to-point communications, so it is not necessary for each node to have a global picture of the network. The information being sensed at each node is only required at a high powered basestation, and thus the problem becomes how to get the global picture of the environment from the sensor nodes to the basestation, which may be very far away, using a minimum amount of energy. Finally, microsensor network applications are often amenable to trade-offs of quality versus resource consumption.

We have exploited these aspects of microsensor networks to develop LEACH (Low Energy Adaptive Clustering Hierarchy), an application-controlled routing protocol that minimizes energy dissipation in sensor networks using localized coordination and control. In addition to localization, LEACH combines the MAC layer communication scheme and the signal processing functions with the routing protocol to achieve energy efficiency. Traditional routing protocols can be energy inefficient due to retransmissions, caused by packet collisions or asynchronization of the sleep states of the transmitter and receiver, and the large amount of data that must be transmitted from the nodes to the basestation. We have designed a MAC protocol specifically suited for the routing protocol in LEACH that maximizes the amount of time the nodes are in the sleep state while reducing the number of wasteful retransmissions. By combining the signal processing functions of the network (e.g., compression, data fusion, classification) with the routing protocol in LEACH, we have also been able to achieve orders of magnitude reduction in the amount of information that must be transmitted to the basestation, hence greatly improving the energy efficiency of the system. We are currently developing a wireless network simulator that will allow us to compare LEACH to conventional approaches such as multi-hop and other energy efficient protocols.

Energy Efficient Filtering Using Adaptive Precision and Variable Voltage

Personnel

A. Sinha (A. P. Chandrakasan)

Sponsorship

NSF

The rapid proliferation of wireless, portable, battery-operated communication devices such as digital cellular phones, has increased the demand for energy efficient implementations of Digital Signal Processing (DSP) systems. Most of these applications have a fixed throughput requirement and doing things any faster is wasteful in terms of power. Finite Impulse Response (FIR) filtering is a very frequently used DSP function. FIR filtering involves an inner product of a fixed finite length vector (i.e., the impulse response h[n]) with shifted samples of the input signal x[n]. Most FIR filters are designed so that they are able to accommodate the maximum precision requirement of the data samples. In general, the datapath scales linearly with the precision requirement and design for the worst case precision is wasteful in terms of energy. Most signal processing applications will only have a few data samples using up the total available precision.

We are investigating a variable bit precision filtering scheme, based on a Distributed Arithmetic (DA) approach, that saves energy in two ways, without loss in any accuracy. First, the precision used is varied based on the requirement for each sample. Let us assume that the maximum precision requirement is M_{max} and the immediate precision requirement is M. The energy per output sample scales down by a factor M/M_{max}. Further, we exploit the fact that lesser precision implies that the same computation can be done faster (i.e. in M cycles instead of M_{max}). We therefore switch down the operating voltage such that we still meet the worst case throughput requirement. The basic architecture is shown in Figure 7. We have demonstrated that 50% to 60% energy savings can easily be obtained in the case of speech data with little hardware overhead to the fixed precision circuit.
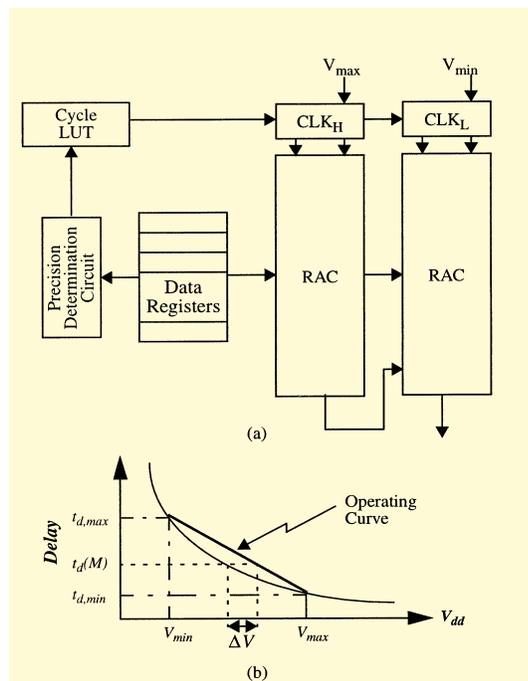

Fig. 7

High Speed, Low Skew Distributed Clocking

Personnel
V. Gutnik (A. Chandrakasan)

Clock distribution in high-speed, high-performance microprocessors takes a significant fraction of the total chip power. As process technology scales feature sizes, however, the skew and jitter continue to increase, and will take an ever-larger fraction of the cycle time.

A distributed clock system may allow faster clock speeds with lower random skew and jitter than possible with traditional clock tree distribution methods. We have fabricated and tested a chip with four 400MHz oscillators in a .6um process. As simulated, it was possible to get the four oscillators phase-locked with no stability problems. A second chip to test a jitter-measurement technique was also fabricated and tested; results there indicate that very accurate on-chip jitter measurement is possible. Design is currently in progress for a robust array of 1GHz oscillators, with the new measurement technique integrated onto the same chip.

Intelligent Transportation Systems

Personnel
B. K. P. Horn, H.-S. Lee, I. Masaki, T. B. Sheridan, C. G. Sodini, J. M. Sussman, and J. L. Wyatt

US citizens are spending, on average, about $1,000 per year for cars, trucks, and roads. Transportation is important not only economically but also socially. The inter-state highway project built a sound infrastructure for our society. What infrastructure do we need for tomorrow? The goal of this project is to develop a technical foundation for tomorrow's transportation systems. Currently we have a number of infrastructures which are independent from each other. Examples include infrastructures for transportation, communication, finance, health care, emergency care, and others. In the next generation, these independent infrastructures will be integrated more closely with advanced information technologies. For example, highway tolls can be charged to driver's bank account automatically with electronic toll gates connected to banks' computers. If a car accident occurs, as another example, the accident can be detected by an air-bag sensor and reported automatically through wireless network to ambulance stations. The ambulance and hospital will have teleconference on the way from the scene to the hospital for a quick care.

The scope of research ranges from small-scale systems to large-scale systems as well as fundamental to application oriented subprojects. Examples of small-scale subprojects are an adaptive dynamic range image acquisition chip, an array processor chip, and a time-to-collision chip. Medium-scale systems include a personal-computer-based real-time three-dimensional machine vision system, a fusion system of machine vision and radar sensors, and an image recognition system for compressed three-dimensional images without decompression. Examples of large-scale systems are a network for real-time image transfer, train control architecture, policies for intermodal systems which consists of cars, trucks, trains, airplanes, and other transportation means.

The research is being carried out at the Intelligent Transportation Research Center in MIT's Microsystems Technology Laboratories. The center is being sponsored by several member companies.

Cost-Effective Hybrid Vision System for Intelligent Cruise Control

Personnel
M. Spaeth (H-S. Lee and I. Masaki)

An essential component of an intelligent cruise control system is a module that computes the distances to objects in the vehicle's field of view. This module must operate in real time at a high frame, so the algorithm used to compute the distance map is simplified using a special trinocular stereo camera geometry and processors specially suited to this application.

The system being designed calculates the distance map using a distance-from-disparity algorithm. Initially, each image is converted into an edge map, so that the edge features may be correlated with between images. Next, the edge positions are refined to sub-pixel accuracy, to increase the resolution of the disparity measurement used to calculate the distances. Finally, the edge positions are correlated, and the distances and disparities are computed. To expedite the search for edge correspondences, the three images are mounted equidistant on a common baseline with aligned optical axes, constraining edge matches to horizontal lines in the other images.

In order to compute the edge map and sub-pixel edge positions efficiently, the ADAP (Analog/Digital Array Processor) MIMD programmable array processor is utilized. Properly programmed the ADAP is fully pipelined and can output data at 1 MIPS, using far less power than a conventional digital processor. The bi-directional edge maps are computed with a 13 sample latency, while the sub-pixel algorithm has an 11 sample latency.

In the current system, sub-pixel edge data is transferred to a PC for the correlation and distance map calculations, but in the future, the early vision processing could be incorporated onto the camera die, and the late processing could be implemented in an ASIC, facilitating a simple stand-alone system.

Fusion of Machine Vision and Radar Sensors

Personnel
Y. Fang and T. Kato (I. Masaki)

Machine vision and radar sensors have their own merits and demerits, respectively. We are trying to achieve a higher performance by fusing these two types of sensors. Currently we are working on two types of fusion systems: one with three-dimensional machine vision and radar sensors and the other with two-dimensional machine vision and radar sensors.

Our three-dimensional vision system consists of two TV cameras and calculates the distance between each object and the camera system from the disparity between the right and left camera images. A time-consuming and difficult part of the distance measurement is to evaluate all possible correspondences between pixels in the right and left camera images. On the other hand, the radar sensor does not have such a correspondence problem but its lateral spatial resolution is much less compared to the machine vision sensor. Let's take a collision warning system for an automobile as an application example. In case there are many cars in front at the different distances, both machine vision and radar sensors will be confused but we will be able to obtain highly reliable data by fusing both sensors.

An integration of two-dimensional machine vision and radar sensors will provide higher reliability and efficiency compared to machine vision only or radar only architecture. For example, if there are two cars at the same distance in front of your car, the radar system will detect that there are two objects at the same measured distance and the machine vision system will calculate whether the lateral distance between those two cars is wide enough for your car to drive through.

Recognition of Three-dimensional Images without Decompression

Personnel
N. S. Love (I. Masaki)

Sponsorship
U.S. DOT

Today's image compression research mostly deals with two-dimensional images like TV images. A feature of this project is to work on three-dimensional images which include distance information between each object and the camera system. Another feature we are working on is a new image compression scheme which does not require decompression for image recognition. Conventionally, image recognition and image compression are two different research areas which are independent from each other, and compressed images need to be decompressed before recognition. With this project, we are developing an image compression method so that the decompression process is not needed for recognition.

First, we have developed a compact three-dimensional image data acquisition system which requires only a personal computer, two plug-in boards, and three TV cameras for real-time operation. The system focuses its computational power on only relevant image regions to achieve its compactness without sacrificing its processing speed. For image compression, we proposed an edge-based method while the main stream of conventional image compression methods use a spatial frequency scheme. Edges are points where the image intensity and/or color change significantly. Images are segmented into small regions by using edge information. The compression system calculates the attributes of each region including distance, color, and intensity. Let's assume, as an recognition example, that we would like to find a red Ford Taurus from a large amount of three-dimensional image database. The image database consists of the following four domains: edge, distance, color, and intensity. The recognition system first uses the color domain to find images which include red color regions. From the corresponding edge and distance domain information, the size of each red region is compared with the expected size. The images with high possibilities of including a red Ford Taurus are decompressed for detailed recognition.

The system will be evaluated as a traffic monitoring video network. The traffic monitoring network consists of a large number of wayside video cameras connected to hierarchical control centers. The centers monitor the number of vehicles per minute and an average vehicle speed, detect vehicle accidents, provide traffic condition information, and control wayside variable message boards.

Automatic Brightness Adaptive CMOS Imager System

Personnel

K.G. Fife, S. Decker and I. Masaki (C.G. Sodini)

Conventional imagers frequently employ a mechanical or electronic shutter to adjust the global integration time for each pixel in an array. While this may eliminate the saturation of bright objects in a scene, the dynamic range captured by the imager remains unchanged. Since image intensities may vary by over six orders of magnitude, details in the darker regions of an image are lost by most conventional imagers.

The brightness adaptive CMOS imager contains a wide dynamic range pixel with a lateral overflow drain. The voltage level on the gate of the reset transistor can be varied during the integration period to control the amount of charge that is accumulated. The imager effectively uses a long integration period for regions of low illumination and a short integration period for regions of high illumination. The imager is capable of capturing scenes with a dynamic range of 1:31,000 (90dB) when a logarithmic function is applied.

 The 256 X 256 array of wide dynamic range pixels has on-chip A/D converters which provide the digital data for the automatic brightness adaptive system. The imaging system either employs a linear or a logarithmic compression scheme. In scenes that do not have a large dynamic range, the linear mode should be used as often as possible because features in an image are more distinguishable when uncompressed. However, when the scene has a large dynamic range, the logarithmic mode is desired. The system decides when to switch between modes by breaking the image into smaller blocks and calculating the average intensity of each region. By comparing the intensities of the individual regions, a wide dynamic range image is targeted.

The brightness adaptation system involves both electronic irising as well as automatic mode switching. Inhibiting the automatic mode switching until the electronic iris feedback loop has settled prevents potential stability problems. The logic required to perform the automatic brightness adaptation has been implemented in VHDL for programmable logic.

Multiple imagers can be used in the system for applications such as stereo vision in intelligent vehicles. In a three-camera setup, all the imagers are synchronized and adjusted based on the center imager's data. This allows for excellent object correlation in a stereo algorithm.

Application of a Pixel-Parallel Image Processing Chip for Intelligent Vehicle Control

Personnel

Z. A. Talib and J. C. Gealow (I. Masaki and C. G. Sodini)

Sponsorship

The implementation of several intelligent vehicle applications requires the real-time performance of multiple low-level image processing tasks. In applications such as lane following, obstacle avoidance, and adaptive cruise control, a large amount of the computation resources are spent on the low-level image processing tasks such as template matching, optical flow, and stereo vision. A processor-per-pixel scheme employing a Single Instruction stream, Multiple Data stream (SIMD) design demonstrates a system fast enough to perform real-time image processing, yet flexible enough to be programmed for a variety of image processing applications.

Typical low-level image processing tasks are performed by applying a uniform set of operations for each pixel in each input image. Thus, they may be efficiently handled by an array of processing elements, one per pixel, sharing instructions issued by a single controller. Using logic pitch-matched to DRAM cells, a single chip provides a 64 x 64 processing element array suitable for real-time applications. Each processing element combines a 128-bit DRAM column with a 256 function one-bit-wide arithmetic logic unit.

The image processing system is comprised of an image data path and a control path for the transmission of pixel data and instructions, respectively. The processing element array receives instructions from the controller, which is managed by the host computer. Analog images from a video camera or other source are converted to digital signals, then reformatted for processing using the processing element array. Output images from the array are converted to a format appropriate for subsequent use.

Fabricated chips are fully functional. operating with a 60 ns clock cycle, the chips dissipate 300 mW. A demonstration system employs four chips to form a 128 x 128 processing element array. Several low-level image processing tasks have been implemented: median filtering, smoothing and segmentation, edge detection, and optical flow computation. All have been successfully performed in real time with input images provided at standard video frame rate. Experimental results are summarized in Table 1.

Current hardware development includes the expansion of the demonstration system from four chips (processing 128 x 128 pixel images) to sixteen chips (processing 256 x 256 pixel images). Current application development includes demonstrating real-time generation of a depth-map using a stereo vision algorithm. Utilizing three cameras separated by a known fixed distance, the stereo vision algorithm determines the absolute distance to an object.

The future plan is to expand and improve the system so that its usefulness can be demonstrated in a real intelligent vehicle control application. While the current system takes raw images as input and returns the full processed images as output, in order to efficiently determine control instructions for a vehicle, it is necessary to extract only the pertinent features (e.g. the location of the edges in the case of edge detection). The low-level information determined by the pixel-parallel processor system would therefore serve as input to a conventional serial microprocessor which would, in turn, determine the appropriate control instructions to deliver to the vehicle.

A Programmable, Wide Dynamic Range CMOS Imager with On-Chip Automatic Exposure Control

Personnel
P. M. Acosta Serafini (C. G. Sodini)

Sponsorship
Center for Integrated Circuits and Systems

Machine vision applications which use visual information typically need an image sensor able to capture natural scenes which may have a dynamic range as high as four orders of magnitude. Reported wide dynamic range imagers may suffer from some or all of these problems: large silicon area, high cost, low spatial resolution, small dynamic range expansion, poor pixel sensitivity, small intensity resolution, etc. The primary focus of the proposed research is to develop a single-chip imager for machine vision applications which addresses these problems, but is still able to provide an ultra wide intensity dynamic range by implementing a novel pixel-by-pixel automatic exposure control. The secondary focus of the research is make the imager programmable, so that its performance (light intensity dynamic range, spatial resolution, light intensity resolution, frame rate, etc.) can be tailored to suit a particular machine vision application.

The imager sensing array has pixels which can be independently read and reset. The proposed brightness adaptive algorithm then predictively scales the voltage in photodiodes that would saturate under normal circumstances based on information gathered in several readout checks. The total integration time is subdivided into several integration times (called integration slots), which are progressively shorter. If it is determined that the pixel will saturate at the end of the current integration slot, then the pixel is reset and it is allowed to once more integrate light, but for a shorter period of time. Each pixel has a small associated memory location needed to store an exponent which identifies the actual integration slot used. This information is used to appropriately scale the digitized pixel output.

Characterization of CMOS Photo-diodes for Image Sensor Applications

Personnel
C-C. Wang (G. Sodini)

Sponsorship
NSF

This research compares the image sensor parameters of dark current and quantum efficiency of N-well/Psub and N^+-diffusion/P-sub photodiodes. Comparisons are made through two separate standard CMOS processes with minimum gate length 2.0 µm and 0.5 µm. The former was a non-silicided process and the latter applied a silicide block mask on active regions. The dimension of the photodiodes is 500 µm by 500 µm to assure current measurement accuracy.

Dark current has large impact on imager quality at low light conditions since it sets the fundamental lower shot noise limit. Over 0-5V reverse bias region, the N-well photodiode presents lower dark current density than the N^+-diffusion diode in both processes. This could be explained by the lightly-doped N-well/P-sub junction, which experienced several thermal cycles to anneal out the lattice damage introduced in well-implantation.

Quantum efficiency is a measurement of the optical-to-electrical conversion efficiency for a detector at a particular wavelength. For wavelengths beyond 600nm, both structures demonstrate comparable response. However, at shorter wavelengths, the N-well diode shows better efficiency. This outcome is attributed to a longer minority carrier diffusion length in the N-well. It could be two orders of magnitude longer than the N^+-diffusion in current processes. Because short-wavelength photons tend to generate carriers near the surface, the carriers generated in the N^+-diffusion layer are more akin to recombine.

According to our measurement results, the N-well photodiode shows lower dark current, and higher quantum efficiency for imager application. A 64x64 pixel array has been implemented and is currently under investigation.

A Differential Passive Pixel Image Sensor

Personnel

I. L. Fujimori (C. G. Sodini)

Over the past decade, CMOS image sensors have received much attention in the electronics industry. Compared to its CCD counterpart, CMOS imagers consume lower power, allow random access and can be integrated with analog and digital functional blocks in a standard CMOS process. The cost for these advantages is a reduction in the image quality, namely an increase in sources of noise, such as dark current and Fixed Pattern Noise (FPN). A large portion of the cost and effort of a CCD fabrication process is dedicated to minimizing the pixel dark current and improving the efficiency of the light-to-voltage conversion. In a standard CMOS process, however, the imager designer has little to no control over the fabrication steps. The challenge of a CMOS imager designer therefore becomes utilizing innovative circuit techniques to achieve CCD quality images in a standard CMOS technology.

A CMOS passive pixel (single transistor) cell is a promising implementation that could potentially reduce the effects of the pixel dark current, and FPN normally observed in a CMOS active pixel. A differential architecture in which the output of a sensing pixel is compared to that of a dummy pixel kept in the dark, is used to reject any common-mode signals such as ground bounce and temperature variations. This differential readout scheme can also be used to subtract the sensing pixel dark current from the dummy cell's dark current, thus reducing the effects of the dark current.

The passive pixel consists of a high-efficiency n-well photodiode and a transistor for row select. The output of the pixel, which is in the form of charge, is converted to a voltage with a sense amplifier at the bottom of every column. Consistent with the low number of transistors per cell, the passive pixel has few sources for fixed pattern noise.

There is one inherent weakness for passive pixels, however. Long wavelength radiation (red and near IR) is absorbed very deep in the substrate of the photodiode. Some of these photogenerated charges will find their way to the depletion region of the photodiode while others may be swept up by the reverse-biased diffusion of the column line. The combined effect of the charge leakage from 256 cells can be significant and will appear as a parasitic current at every column line. Though this parasitic current is also present in active pixels, its effect is more pronounced in passive pixels because charge amplification does not occur within the cell. Fortunately, this signal dependent parasitic current can be removed with correlated double sampling, making the passive pixel a competitive choice in the implementation of CMOS imagers.

While CMOS imagers have not yet achieved the superior imaging quality of CCD's, they offer many advantages for applications where high image quality is not essential, but where low power, low cost and high integration is desired. Some examples include surveillance, biomedical and videoconferencing applications.

Transistor-level Synthesis of Delta-Sigma Converters

Personnel
M. S. Peng (H-S. Lee)

Sponsorship
Heinle Memorial Fund and Center for Integrated Circuits & Systems

With the increasing importance of quick generation of analog circuits, many analog circuit synthesis CAD tools have been designed and are in current development. However, most attempts have been very general, rendering the designs generated highly impractical. Instead of trying to encompass all applications possible, we have chosen to synthesize specific analog circuits for specific applications, thereby decreasing complexity and improving robustness.

The focus of this research is the design and implementation of a general, low-power, oversampling, bandpass, one-bit delta-sigma analog to digital converter for small voltage measurements, particularly MEMs applications. This tool allows a user to enter parameters for the desired design (order of modulator, oversampling ratio, etc.) and a full netlist of capacitors, switches, operational amplifiers, and comparators are generated. The generated design has dynamic range scaling and power optimized operational amplifiers, important characteristics of modern designs.

The current architecture that this tool utilizes is a cascade of integrators/resonators for either low-pass or band-pass delta-sigma architecture. The architecture is implemented as a discrete-time, fully differential design. Switch capacitors are used which makes implementation of the integrators/resonators straightforward.

The behavioral circuit synthesis part of this project is based on the Delta Sigma Converter MATLAB toolbox developed by Dr. Richard Schreier at Analog Devices. The synthesis tool then produces a SPICE net list based on the modulator structure and scalable 2-stage CMOS operational amplifier design. The capacitor values are generated from modulator coefficients and kT/C noise considerations. The SPICE simulations of synthesized modulators agree extremely well with the behavioral simulation results.

Superconducting Bandpass Delta-Sigma A/D Converter

Personnel
J. F. Bulzacchelli (H-S. Lee and M. Ketchen, IBM)

Sponsorship
IBM Cooperative Fellowship

The direct digitization of RF signals in the GHz range is a challenging application for any circuit technology. Traditionally, flash A/D converters have been used to digitize signal frequencies above 1 GHz, but their resolution and linearity are inadequate for most radio systems, which must handle signals with a large dynamic range. Semiconductor bandpass delta-sigma converters are used to digitize IF signals with high resolution, but their performance at microwave frequencies is limited by the speed of semiconductor comparators and the low Q of integrated inductors.

In this program, we present the design of a superconducting bandpass delta-sigma converter for direct A/D conversion of GHz RF signals. The schematic of the circuit is shown in Figure 12. The input signal is capacitively coupled to one end of a superconducting microstrip transmission line, which serves as a high quality resonator (loaded $Q > 5000$). The current flowing out of the other end of the microstrip line is quantized by a clocked comparator comprising two Josephson junctions. If the current is above threshold, the lower junction switches and produces a quantized voltage pulse known as a single flux quantum (SFQ) pulse. If the current is below threshold, the upper junction switches instead. The pattern of voltage pulses generated across the lower Josephson junction represents the digital output code of the delta-sigma modulator. These voltage pulses also inject current back into the microstrip line, providing the necessary "feedback" signal to the resonator. At the quarter-wave resonance of the microstrip line (about 2 GHz in our design), the resonator shunts the lower junction with a very low impedance, the "feedback" current to the resonator is maximized, and the quantization noise is minimized. Because of the high speed of Josephson junctions and the simplicity of the proposed circuit, we expect sampling frequencies in excess of 20 GHz, limited only by the digital circuitry needed to process the output of the delta-sigma modulator.

Circuit performance at a 20 GHz sampling rate has been evaluated with JSIM, a SPICE-like simulator for superconducting circuits. A representative example of the A/D converter's output spectrum is shown in Fig. (ref. no. 2). In this simulation, the A/D converter was driven by a large (-0.8 dBFS) input near 2.13 GHz, just above the frequency band of interest. The minimum in the quantization noise power spectrum is located at 2.05 GHz. In-band noise is -53 dBFS and -57 dBFS over bandwidths of 39 MHz and 19.5 MHz, respectively. In addition to the minimum at 2.05 GHz, there are minima at other frequencies. Near dc, bias inductor L_{bias} shunts the lower Josephson junction of the comparator with a low impedance, and quantization noise is minimized. The other minima correspond to higher-order modes on the microstrip line, including some above 10 GHz which appear in the digital domain as "aliased" modes. In principle, these aliased modes could interfere with the desired noise shaping near 2 GHz. In our design, the sampling frequency and the resonances of the microstrip line have been chosen so that no aliased modes fall near the band of interest. Intermodulation (IM) distortion was also studied with several long JSIM simulations. Over a 39 MHz bandwidth, in-band IM distortion is better than -69 dBFS. Other features of the circuit include unconditional stability and a full-scale input sensitivity of 20 mV (rms).

While a 20 GHz sampling rate improves the performance of delta-sigma converters, the challenges of high speed testing in a cryogenic environment are formidable. Even in the best cryogenic sample holders, the long cables used to connect the superconducting chip to room-temperature electronics have significant losses at frequencies above 10 GHz. In previous reports, we described an optoelectronic clocking technique, which bypasses the bandwidth limitations of conventional electrical testing. In this approach, picosecond optical pulses from a mode-locked laser are delivered (via optical fiber) to an on-chip photodetector, which gener-

ates the clock pulses needed by the Josephson circuitry. We have already demonstrated our optoelectronic clocking system up to 20 GHz and will use it in testing our A/D converter later this year.

While the high speed clocking will be done optoelectronically, the digital outputs of the A/D converter will still be read out over standard coaxial cables. Consequently, direct transfer of the output of the delta-sigma modulator (at 20 Gbits/s) to the room-temperature test equipment is impractical in our setup. Instead, on-chip processing of the data will be used to reduce the bandwidth requirements for readout.

As explained in last year's report, two segments of the modulator's bit stream will be captured with a pair of 128-bit shift registers. The number of clock cycles skipped between acquiring the two segments will be set by an on-chip programmable counter (from 0 to over 8000). Cross-correlation of the two captured segments will be used to provide estimates of the autocorrelation function R[n] of the A/D converter's output, for all values of n up to 8000. Fourier transformation of R[n] will then yield a power spectrum with frequency resolution comparable to an 8K FFT of the original bit stream.

We have just recently completed the design and layout of our entire A/D converter test chip, including the bandpass delta-sigma modulator, the shift registers, the programmable counter, and readout circuitry. The chip is currently in fabrication at HYPRES, Inc., and we expect to be testing the A/D converter within a couple months.

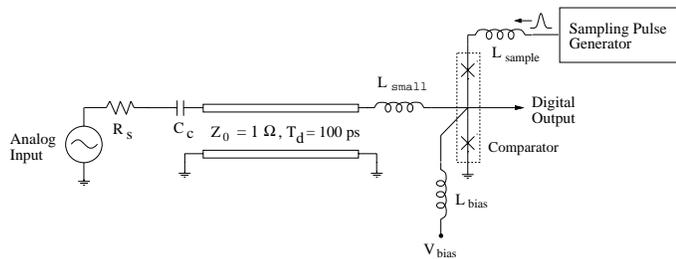# Substrate Noise Shaping in Mixed-Signal Systems

Fig. 12: Superconducting bandpass delta-sigma converter.
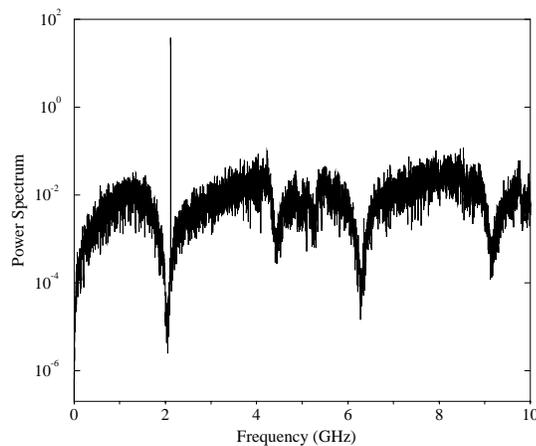
Fig. 13: Simulated output spectrum of proposed delta-sigma converter with large sinusoidal input.

Personnel
M. S. Peng (H-S. Lee)

Sponsorship
Center for Integrated Circuits & Systems

Analog circuits and digital circuits have, for the most of history, been fabricated on separate substrates in electronic systems. However, the drive for higher performance, lower cost, and lower power have pushed for the integration of these two parts onto a single microchip. While plausible and feasible, the marriage of the two is not without problems and difficulties. One of the most insidious problems is the substrate noise coupling from the highly noisy digital circuits to the extremely noise sensitive analog circuits. Digital noise can severely degrade crucial analog performance if not contained properly.

Up to now, most efforts of minimizing digital noise effects on analog circuitry have been to utilize good layout techniques along with computer-aided-design verification. Although this process ensures the mixed-signal system performs as desired, no real effort has been made to design circuits that pointedly address the substrate noise problem.

Therefore, the focus of this research is to characterize and investigate digital noise mitigating circuit techniques, in the digital and analog domains. The benefit of easily integrating analog and digital circuits would be immense and epochal.

Oversampled Pipeline A/D Converters with Mismatch Shaping

Personnel
A. Shabra (H-S. Lee)

In recent years, delta-sigma modulators and pipeline converters have become very popular as analog-to-digital converters. In comparing these converters for wide-band signals, we recognize a few important attributes. Due to the wide bandwidth of the input signal and limited circuit speed, delta-sigma converters afford only low oversampling ratios, which makes high-resolution conversion extremely difficult. The low oversampling ratio generally nullifies the primary advantage of delta-sigma converters; the tolerance to component mismatches. As a practical matter, at input frequencies over 50 MHz, it would be extremely difficult to achieve oversampling ratio greater than 8 with final accuracy over 12 bits with present technologies. At this low oversampling ratio, not only the benefits of delta-sigma modulation disappears, but more importantly many delta-sigma converters are incapable of providing good enough performance.

We believe that a more efficient approach would be to oversample a standard pipeline converter, and shape the mismatch out of band which will be removed by a subsequent digital filter. Since no attempt is made to shape the quantization noise, there is none of the concerns associated with delta- sigma converters with a low oversampling ratio.

This work applies mismatch shaping to a Commutative Feedback Capacitor Scheme (CFCS) pipeline converter. The SNDR improvement is achieve through a combination of oversampling and mismatch shaping, which modulates the distortion energy out-of-band. Simulation results show that at an oversampling ratio of 4 and 64 and component mismatch of 0.1% a 9dB and 35dB improvement in SNDR is achieved respectively, compared to a converter with no mismatch shaping.

To demonstrate the mismatch shaping technique, a test chip is being designed in an advanced CMOS technology. The test chip will target aggressive performance goals, in particular a sampling frequency around 200MHz and a resolution/accuracy of 12bits at an oversampling ratio of only 4.

A Nyquist-rate Pipelined Oversampling Delta-Sigma A/D Converter

Personnel
S. Paul and J. Goodrich (H-S. Lee)

Oversampling and noise-shaping techniques, such as delta-sigma modulation, are widely used in analog-to-digital conversion to achieve resolution that exceeds that of integrated-circuit components. Such converters have an inherent tradeoff between accuracy and speed, whereby resolution in amplitude is achieved at the expense of resolution in time. The data rates of delta-sigma converters are limited, even when their internal circuits are pushed to their highest speed, because their modulators must operate over many clock cycles to produce a single result. Power dissipation is also a concern for these devices since circuits must operate at high speed, and thermal noise sets a lower bound on capacitor sizes. Much attention has been focused on improving the speed and power of delta-sigma analog-to-digital converters through use of higher-order modulators, multi-bit feedback, and multi-bit architectures with single-bit feedback. However data rates remain limited to less than a few MHz and are not easily extended.

A new architecture for oversampling A/D conversion is under development. This architecture, referred to as pipelined oversampling, circumvents the speed-resolution tradeoff of conventional oversampling ADCs by performing spatial, rather than temporal, oversampling. It combines the high-resolution capability of oversampling techniques with the speed advantages of pipelined architectures so that both of these attributes are achievable. Its output data rate equals its internal clock rate, and no higher speed circuits are required. Its oversampling ratio and resolution are determined by its pipeline length. Accuracy and speed for such a device may be independently adjusted within the constraints of a given process technology. Power is improved as a result of reduced circuit speed, which permits voltage scaling and use of low-power technologies, simplified decimation requirements, a charge-domain implementation, and reduced sensitivity to thermal noise. A

Low Voltage, Low Power CMOS Operational Amplifier Design for Switched Capacitor Circuits

Personnel
P. M. Naik (H-S. Lee)

pipelined architecture is also well suited for processing presampled signals because it performs Nyquist sampling.

A pipelined oversampling converter is most practically built using a combination of Charge-Coupled Device (CCD) and CMOS circuits. Although CCDs are not essential to the concept, a CCD/CMOS combination enables devices that would be difficult using either one alone. CCDs provide delay and integration operations that are simple, low power, and highly accurate. Pipelines with hundreds of stages are also easily and compactly achieved. CMOS also plays a vital role by providing digital logic, CCD support circuitry, and additional analog circuit flexibility. Charge domain processing brings additional advantages because fully depleted CCD operations are not subject to thermal noise or voltage coupling from clocks or the substrate.

Two prototype pipelined oversampling converters have been demonstrated. The first achieves 74 dB SNR on an 8 MHz input at an 18 MSPS data rate. Its power dissipation is 324 mW at 18 MSPS. The second prototype achieves 66 dB SNR at a 30 MSPS data rate and has 230 mW power dissipation.

The increasing demand for circuits with low voltage operation and low power consumption are being driven by smaller submicron technologies and the move towards more portable electronic systems. Lower supply rails are necessary due to smaller device breakdown voltages resulting from smaller feature sizes. In addition, low voltage is desirable so fewer batteries are required, thus reducing system size and weight. Low power consumption is necessary to ensure reasonable battery lifetime.

These demands present challenges in circuit design, particularly in the analog domain. Unlike digital circuits where power consumption reduces proportional to the square of the supply voltage, in analog circuits lower voltage actually increases power consumption. Another design challenge that arises is lower drive availability, which decreases device speed. Reduced dynamic range is yet another problem since the supply voltage is reduced while the device threshold voltage remains the same. The objective of this research is to design a low voltage, low power CMOS operational amplifier with high performance for use in switched capacitor circuits.

Table 1 shows some key design goals and Hspice simulation results of the op amp. The power supply is +/-0.9V, the capacitive load is 200fF and the process used was the HP 0.5μm minimum feature size. The design uses a simple two-stage topology with added features to improve performance. These features include cascoding to improve gain; a replica tail to provide constant current which improves common mode input range, supply and common mode rejection ratios; a local switched capacitor common mode feedback circuit; and a bias circuit tracking the opamp. A layout of this chip using a triple metal, single poly n-well process has been completed.

Low Power Reconfigurable Analog-to-Digital Converter

Personnel

K. Gulati (H-S. Lee)

Sponsorship

The objective of this project is to create a low power A/D converter to be employed in a battery powered wireless sensor. The key requirement of the converter is that it be able to cater to a wide range of input data-rate (1sample/s to >1Msample/s), and produce a resolution ranging from 8 bits to as high as 16bits, consuming the lowest possible power consumption. A possible subset of applications envisioned for the A/D converter is depicted in Figure 14. This large size of data-rate versus resolution space warrants the need for several different A/D converter architectures. However, such a converter implementation would require a prohibitively large area. Further, in the event that the total area required is larger than that can be supported by a single chip, the speed of the converter would be dramatically reduced due to off-chip parasitic capacitance.

To be able to cover the entire rate-resolution space, a reconfigurable A/D converter is proposed. Another requirement on the ADC is that it consume the lowest power possible. To achieve this, a host of low power techniques are being employed.

The opamp is the most power-hungry device in the A/D and consequently significant work has been done to reduce the power consumption of this device. Given the observation that SNR, speed and power consumption of an opamp are all interrelated and, specifically, that the power consumption of the opamp is inversely proportional to the square of the SNR (or swing of the amplifier, for a fixed input referred noise), a new operational amplifier topology was created.

A telescopic cascode opamp typically has a higher frequency capability and consumes less power than other topologies. The disadvantage of a telescopic opamp is the severely limited output swing.

The opamp developed in this project (Figure 15) uses the telescopic architecture as its core for reasons of high speed and low power consumption and offers much higher output swing than a conventional telescopic amplifier while maintaining high CMRR and supply rejection (PSRR), and ensuring constant performance parameters. Transistors M7-M8, and M9 are deliberately driven deep into the linear region. In this case, the output swing is improved by 0.7V from a telescopic amplifier and becomes slightly better than that of a folded cascode amplifier. The reduction of gain and CMRR due to the low output resistance in the linear region is compensated by gain enhancement and replica tail feedback, respectively. The gain enhancement employs the well-known differential regulated cascode structure except the control voltage Vncontrol is chosen to bias M7-M8 in the linear region. The gain enhancement amplifier A2' incorporates the replica tail feedback to keep the drain current of M9 constant despite input common-mode voltage variation.

A prototype of the operational amplifier was fabricated (Figure 16) using a single-poly, three-metal 0.8μm HP process. The measured/simulated specifications for the opamp are detailed in Tables 2 and 3.

The fabricated ADC is expected to be able to handle a resolution rangle of 2 bits to 16/18 bits and a input bandwidth of from 1s/s to 50Ms/s. Table 3 details the possible signal types the ADC is expected to able to handle with corresponding power consumption estimates.

The Analog-to-digital converter is currently being layed out. It will be fabricated in a double-poly, triple metal TSMC 0.6μm CMOS process, and characterized, thereafter.

NOTE: SORRY, FIGURES NOT AVAILABLE IN ELECTRONIC FORM

Ultra Low Power Wireless Sensor Project

Personnel
A. Chandrakasan, H.-S. Lee, C. G. Sodini, and T. Barber - Analog Devices

Sponsorship
DARPA

This program is developing a prototype wireless image sensor system capable of transmitting a wide dynamic range of data rates (1 bit/s - 1 Mbit/s) over a wide range of average transmission output power levels (10 μW - 10 mW). The prototype will dissipate approximately 50 mW and is based on research IC's designed at MIT.

The imager is a 256 x 256-pixel sensor being designed in a 0.6 μm CMOS technology. The power dissipation of the imager is reduced by decreasing the power supply voltage and incorporating random-access readout. The key issue in reducing the power supply voltage is the minimization of noise.

To optimize power consumption for different signal frequencies and dynamic ranges, a reconfigurable Analog-to-Digital Converter (ADC) is being designed. The concept is based on the fact that the pipeline and sigma-delta ADCs share the same basic analog building blocks such as switched capacitor integrators, subtractors, and voltage comparators. A switch matrix will configure these components into one of the two converter types, depending on the signal frequency and dynamic range requirements. In addition, the converter will be reconfigured within each type (e. g. # of stages, order,) for each performance range.

The digital signal processing aspects of this system focus on low-power image compression, encryption, and the digital control of the entire system. Efficient power-down techniques are used where only the most significant bits representing the intensity of the image are processed when higher resolution images are not required.

The circuit design of the RF front end is focused on the implementation of narrowband transmit and receive functions that are optimized for low-power operation. The transmitter architecture eliminates the need for power-intensive mixers and D/A converters. The VCO employs bond wire inductors placed on chip and noise-matching techniques. These chips have been used to demonstrate a DECT transmitter capable of 1.25 Mb/s at 1.8 GHz using only 27 mW.

Various techniques are employed to achieve an efficient power supply at low voltage and power levels, including delay-line PWM, discontinuous-mode operation, and shared control for multiple outputs. The circuit is programmable to accommodate adaptive voltages and power levels.

Low Power 1.8 GHz Frequency Synthesizer Capable of 2.5 Mbit/s Modulation

Personnel

D.A. Hitko, D.R. McMahill, N.R. Shnidman and M.H. Perrott (C.G. Sodini)

This research effort focuses on the development of a low power, wireless transmitter capable of digital modulation at data rates greater than 1 Mbit/s at a carrier frequency of 1.8 GHz. Our approach is architectural in nature, with the goal of achieving a design that allows complete integration of the transmitter except for the RF bandpass filter.

To achieve a low power solution, we have chosen a topology having a minimal number of components. We argue that any narrowband transmitter with good spectral characteristics must contain a frequency synthesizer to select the desired output frequency, and a transmit filter to shape the modulation spectrum. Our approach, therefore, is to directly modulate a frequency synthesizer with data that has been shaped by a digital transmit filter. This design removes the need for mixers and I/Q D/A converters found in other transmitter architectures.

The idea of directly modulating a frequency synthesizer is not new, and has been successfully achieved by other researchers using fractional-N synthesizers with noise shaping. Unfortunately, the bandwidth constraints on the modulation data imposed by the synthesizer dynamics are a strong deterrent to achieving data rates above 100 Kbit/s. The contribution of this research is the proposal and verification of a compensation technique that allows over an order of magnitude increase in data rate when directly modulating a frequency synthesizer. This technique is quite simple — the digital transmit filter is modified by convolving it with the inverse of the PLL transfer function seen by the data. To verify the method, a custom 0.6 micron CMOS frequency synthesizer was built and shown to successfully perform GFSK modulation at data rates greater than 2.5 Mbit/s at a 1.8 GHz carrier frequency. Key points to the implementation are that the bandwidth of the synthesizer is 84 kHz, and its power consumption 27 mW.

A demonstration system has been constructed to showcase the CMOS synthesizer IC and the Bipolar VCO IC. The demo system is made up of two self contained systems. The first system is a transmitter which incorporates the two IC's designed at MIT. The transmitter is able to accept an external 1.25 Mbps data stream or generate a Pseudo Random Bit Stream test sequence. In addition an on board oversampled A/D converter provides a means for digitizing an analog signal for transmission. The pulse shaping and precompensation FIR filter has been implemented in a programmable logic device. A pair of digital LCD panel meters provide a real time display of the power consumption of the VCO and synthesizer IC's. The RF output frequency may be tuned to any of the 10 DECT channels in the 1.88-1.9 GHz range.

The other half of the demo system is a basestation which receives the transmitted RF signal and provides a serial data and clock at its output. In addition, an on board oversampled D/A converter is used to drive an audio amplifier and speaker. The audio output portion combined with the A/D in the transmitter allow the demonstration system to be easily portable. A portable CD player is the only external equipment required in addition to the transmitter and the basestation to show the demo.

Automatic Calibration of Modulated Frequency Synthesizers

Personnel
D. McMahill (C.G. Sodini)

The focus of the proposed research is the development of a low power radio frequency transmitter architecture. Specifically, this work will add in service automatic calibration to a modulated PLL frequency synthesizer. Angle modulation is accomplished by modulating the feedback divide value in a phase locked loop frequency synthesizer. A digital precompensation filter is used to extend the modulation bandwidth by canceling the lowpass transfer function of the PLL. The autocal circuit will maintain accurate matching between the digital precompensation filter and the analog PLL transfer function across process and temperature variations. The autocal circuit, which is the main contribution of the proposed research, will be designed to operate while the transmitter is in service. This online calibration eliminates the need for production calibration and periodic down time for calibration cycles. In addition, a higher level of matching will be possible with the calibration circuitry.

Key features of the proposed system include the following:

- RF output carrier center frequency in the 1.88 GHz range

- GMSK/GFSK modulation at 2.5 Mb/s

- No required manufacturing calibration

- Potential for low power operation

- To be implemented in 0.6 micron BiCMOS

High Data Rate 5.8GHz Wireless Network

Personnel
D.A. Hitko (C. G. Sodini)

In a newly opened ISM band centered at 5.8GHz, 150MHz has been allocated for unlicensed operation, which is sufficient bandwidth for implementing wireless networks handling data at rates of up to 1GBit/s. To support these data rates, the wireless modem will need to utilize a multi-level quadrature amplitude modulation scheme, which requires an RF transceiver with high linearity and low noise. In addition, some of the information appliances which reside on this network may be portable, thus power consumption is another important consideration.

To address the power concerns, circuits will be designed such that the power dissipation in them can be dynamically scaled to reflect required levels of transceiver performance. For example, if the transceiver is being used for a moderate data rate signal, a less complex modulation scheme may be used such that a given bit error rate can be achieved for a rather modest Signal-to-Noise Ratio (SNR). In this case, the power dissipation of the components in the RF transceiver would be reduced such that only the required levels of SNR would be produced. Similarly, when data rates are low and a significant portion of the available bandwidth of the network is not being utilized, the linearity requirements are relaxed as more interference could then be tolerated into neighboring, unused channels. However, in cases where a high data rate or extremely low bit error rate is required, SNR and linearity may need to be pushed to the limits of the technology.

Research activities are exploring the use of silicon IC technologies to realize these functions in a manner consistent with the cost targets of commercial applications. The majority of the modem signal processing will be concentrated in the digital domain, allowing the RF transceiver function to be simplified. As novel receiver and transmitter architectures will not be required, the research effort will focus upon circuit design and optimization at the device and component level. Particular

Beam Steering for Increased Transmission Power Efficiency in Mobile Devices

Personnel
N. Shnidman (C. G. Sodini)

emphasis will be placed upon identifying instances where technological improvements will result in better circuit performance. As an example, the effects upon oscillator performance of process enhancements resulting in transistor characteristics such as a lower base resistance, a higher fT, or reduced collector-substrate capacitance, will be explored. The manifestation of these improvements into wireless network system capabilities is also being examined.

This project involves the use of antenna arrays on mobile devices in order to direct transmission power. Phased array antennas allow energy to be constructively added into a beam and electronically steered toward a receiver, the base station.

Some advantages of such a system are reduced transmission power (energy is directed only where it is effective), reduced multi-path (less ambient energy means fewer reflections), and the potential for increased device density (spatial location can be used to distinguish between devices).

The major problem with such a system is that the orientation of the array relative to the base station in not known a priori. The fact that the device can move means that both the spatial location of the device, and the orientation of the device's array relative to the base station can change as the device moves. This ignorance of the device array orientation precludes the base station from being able to determine the transmission direction from the device, as there is no known point of reference by which the base station can direct the beam. To solve this problem the device itself must determine the location of the base station relative to its current position. The location of the base station is determined using the antenna array during reception of transmission from the base station to the device. Each antenna in the array receives an incoming transmission at a slightly different time, and the phase difference between the signals on each antenna can therefore be used to reconstruct the location of the base station relative to the current positioning of the device.

The use of a phased array for a mobile device is facilitated by the planned operating frequency of 5.8GHz. A half wavelength at this frequency is approximately one inch. An electrical length between the elements of one half wavelength is considered the optimal spacing here because placing the elements further apart will result in

28

Real-time Video Network

Personnel

H. Kumazawa (I. Masaki)

grating lobes, and placing them closer will result in excessive coupling between elements. The physical spacing between elements can be reduced even further by making use of high dielectric materials which reduces the physical spacing needed to achieve a given electrical length. This relatively small spacing between elements allows enough array elements on the area of a hand-held device such that a transmission beam can be formed and steered electronically.

Conventionally, we have two types of networks: connection and non-connection types. Telephone networks are examples of connection type while Internet adapts a non-connection scheme. With the connection type, a communication path is established at the beginning of the communication and is kept during the entire communication period. A merit is that a specific communication capacity is guaranteed, but the efficiency is not high. With a telephone network, for example, an assigned communication path cannot be used for other communication during time periods in which no voice information is being transferred. In contrast, the non-connection type has higher efficiency but only "best effort" communication is provided instead of "guaranteed" communication.

We are working on new network architecture which is close to the conventional "guaranteed" network in reliability and also close to the conventional "best effort" network in efficiency. As the first step, we are developing a network which transfer video images in near real-time by using NTCIP (National Transportation Communications for Intelligent Transportation Systems Protocol). The NTCIP is compatible with the protocol for conventional Internet. A number of TV cameras are connected to local control centers, and the local centers are connected to broad-region centers through the NTCIP-based network. The broad-region centers send commands to control TV camera parameters such as camera's viewing direction, zoom/pan, image resolution, and others. Each local center consists of SNMP (Simple Network Management Protocol) agent, real-time middle-ware, MIB (Management Information Base), a camera controller, and an image compression board.

Future systems may include automatic auction system which assign available communication capacity to communication requests in case the demands exceed the capacity of the network. Other research issues are high privacy-protection and security which do not sacrifice the openness of the network significantly.

Electromigration in Single Crystal Interconnects

Personnel
V.T. Srikar (C.V. Thompson)

Electromigration, or current-induced diffusion, leads to mass redistribution and stress evolution in the patterned polycrystalline metal films used to interconnect devices in integrated circuits. Electromigration eventually leads to mechanical failure resulting in open or short circuit failures, so that electromigration is a major concern in the processing and design of reliable circuits. The rate of electromigration is governed by the defect structure, and especially the grain structure, of the interconnect lines. Modern high performance integrated circuits can have up to a kilometer of total length of interconnect lines, the reliability of which is governed by the submicrometer-scale grain structure of the interconnect material. Lines with continuous grain boundary paths along their lengths have much lower reliability than lines with 'bamboo' grain structures in which all grain boundaries are normal to the line axes.

Electromigration-induced failures of interconnects with bamboo or near-bamboo microstructures is affected by transgranular diffusion and failure mechanisms. In order to isolate transgranular failure mechanisms for characterization through accelerated tests, we have developed techniques for producing bi-crystal and single crystal Al films on oxidized silicon wafers. We have used these techniques to experimentally investigate the rates and mechanisms of failure in pure Al single crystal interconnects with different textures and in-plane orientations. Kinetic results are being used in simulations of electromigration-induced failure in bamboo and near-bamboo lines. We have also carried out experiments on Cu-doped single crystal Al interconnects and single crystal Al with Al_3Ti overlayers to investigate the effects of these factors on the kinetics and mechanisms of transgranular electromigration and electromigration-induced failure.

Adaptive Body Biasing

Personnel
J. Kao and S. Narendra (A. Chandrakasan, D.Antoniadis, and V. De (Intel))

As circuit performance and power dissipation continue to improve, supply voltages, threshold voltages, and circuit dimensions all scale accordingly. However, this continued scaling results in increased device parameter variations, which will have a significant impact on overall performance. For example, threshold voltages for small devices are greatly affected by poly CD variation, and as supply voltages shrink, these Vt variation will have a larger impact on circuit performance.

The goal of the adaptive body biasing project is to help reduce the effects of parameter variation on circuit delay by modulating the body bias to tune the threshold voltage. The feedback will actually develop the appropriate bias voltage such that a fixed delay criteria is met. By applying reverse bias to slower paths in the circuit, we can then reduce excessive leakage currents and thus reduce overall power dissipation. This procedure can be performed statically during burn-in to reduce parameter mismatch effects on delay, or can even be performed dynamically during run time to reduce time varying parameter mismatches.

A test chip has been designed to study the effectiveness of adaptive body biasing for reducing circuit delay variations in an aggressive .18μm process. The test chip uses a critical path consisting of a chain of complex inverters that represents the delay of a typical microprocessor. A target frequency is input into the chip, and it's period is then compared to the delay through a line matched to the critical path to be compensated. The feedback mechanism continues to apply reverse bias to the circuits until the target frequency is satisfied. This adaptive biasing scheme can actually be repeated for multiple regions within a single chip, such as for a multiplier block or an ALU, and will allow much tighter control over parameter variations.

Three-Dimensional Integration: Analysis, Modeling and Technology Development

Personnel

A. Rahman and A. Fan (R. Reif)

As the critical dimensions in VLSI circuits continue to diminish, system performance of Integrated Circuits (IC) will be increasingly dominated by interconnect's performance. For the technology generations approaching 100-nm, innovative circuit designs and new interconnect materials and architecture will be required to meet the projected system performance. New interconnect material solutions such as copper and low-k dielectric offers only a limited improvement in system performance. Significant and scalable solutions to interconnect delay problem will require fundamental changes in system architecture, design, and fabrication technologies.

Three-dimensional (3-D) IC can alleviate interconnect delay problem by offering flexibility in system design, placement and routing. In 3-D IC, devices are allowed to exist on more than one active layer, and they can be contacted from both top and bottom active layers. Flexibility to place devices along the third dimension allows higher device density and smaller chip area in 3-D IC. The critical interconnect paths that limit system performance can also be shortened by 3-D integration to achieve faster clock speed. By 3-D integration, active layers fabricated with different front-end processes can be stacked to form systems on a chip.

Analysis and Modeling

A recently derived 2-D stochastic wire-length distribution that has been used to predict interconnect delay and system performance metrics has been extended to 3-D interconnects to determine the trade off of 3-D integration. Using 3-D stochastic wire-length distribution and critical path constraints, figures of merit such as interconnect delay, chip area, complexity (cost) etc. have been estimated. The wire-length distribution is derived using an empirical relation known as Rent's rule which relates the number of input and output terminals of an integrated circuit to the number of logic gates within that circuit. The critical path consists of a set of logic gates, where all the logic gates except one drive average length wires, while one logic gate drives a wire of chip-edge length.

Based on our simulation, 3-D integration results in narrower wire-length distribution, with more local (short) wires and less global (long) wires, than the conventional planar implementation. The length of average and total wires in 3-D integration is also shorter than 2-D integration. Wire-length distribution of 3-D IC with 40 million transistors, consistent with 0.18 µm technology generation. In estimating the wire-length distribution, it is assumed that **i)** Rent's rule can be applied iteratively throughout the system, **ii)** it is equally likely to form vertical interconnects between active layers as horizontal interconnects within active layers, and **iii)** the number of interconnects is conserved in 2-D and 3-D implementation.

Initial case studies suggest that reduction of interconnect delay along the critical path in 3-D IC can be as much as 70%-80% per layer compared to 2-D IC. Similarly, for fixed interconnect delay, total chip (Si) area can be reduced as much as 30%-35% per layer. However, due to the complexity and cost associated with 3-D integration, it may not be practical to integrate more than 4-5 active layers. Currently, case studies are being conducted to identify the applications and design approaches that will be benefited most by 3-D integration.

Fabrication

In direct 3-D integration, active device wafers are bonded together, while all active layers are electrically interconnected using high aspect ratio vias. Prior to wafer bonding, the device wafers are assumed to contain multiple aluminum metal layers and Inter-Level Dielectrics (ILD), thus requiring low-temperature bonding below 450°C to avoid Al degradation.

The Effects of Thermal History on the Stress and Reliability of IC Interconnects

Personnel

V.T. Srikar, A. Gouldstone, M. Kobinsky, M. Gross
(S. Suresh and C.V. Thompson)

To implement a structure with 3-D stacked device layers, a metal-to-metal bonding process has been proposed. Metal (Cu) bumps on both wafers will serve as electrical contacts between via on the top wafer and Al interconnects on the bottom wafer. At the same time, these metal bumps also function as the wafer bonding medium. In addition, dummy metal patterns can be made to increase the bonding surface area. The dummy metal films can also be auxiliary structures such as ground planes or heat conduits for different Si active layers.

Successful wafer bonding was achieved using Cu/Ta (300/50 nm) layers on Si at 400ºC for 30 min (in the EV bonder) and annealed at 400ºC in N_2 for 30 min. The bonding interfaces include blanket-to-blanket or blanket-to-patterned metal films. In the latter case, Cu lines with linewidths ranging from 60-200 µm were wet-etched using H_2O_2:HCl:H_2O (1:1:125 by volume) with an etch rate of 37 Å/sec. The total metal adhesion layer thickness is less than half the usual thickness required for polymer-based wafer bonding (>2 µm). The Cu film does not require special pre-bonding surface preparation, such as metal CMP or ultraviolet light exposures. The bonded pairs at 400ºC exhibited good bonding strength when the razor blade test was applied. Further studies include the etching and filling of inter-wafer vias (8:1 aspect ratio, ~20 µm deep), improve the current wafer-to-wafer alignment tolerance of $\pm$ 5 µm, and creating simple test structures to measure via resistance and demonstrate 3-D feasibility.

The metal films used for integrated circuit interconnects are deposited at elevated temperature, patterned at room temperature and then passivated by encapsulation with glasses deposited at elevated temperature (350 to 400°C), before they are used at temperatures around 100°C. This thermal history affects the stress state of the interconnects and ultimately affects their reliability, as limited by stress or electromigration-induced voiding. The stress state is not only a function of the details of the thermal history, but is also a function of the interconnect material and of the interconnect thickness, width, and spacing, as well as of the thickness and mechanical properties of the passivation layers. We have used a generalized plane strain formulation to predict the stress state as a function of these parameters, allowing for both elastic and plastic accommodation of the thermal strain. We are testing these models through comparisons with experiments in which the radii of curvature of wafers coated with lines of pure Al, Al-alloys and pure Cu interconnects with fixed widths, thicknesses and spacings are measured as the wafers are thermally cycled. Stress measurements are also made using X-ray analyses. Through these experiments we are developing models for the micromechanisms of plastic accommodation of thermal strains as a function of materials selection and structural characteristics, and as a function of interconnect geometry. Results from experimentally validated stress simulations will be incorporated into our electromigration simulation tools.

A Framework for Collaborative, Distributed Web-based Design

Personnel

N. B. G. Konduri (A. Chandrakasan)

The design of future high-performance "system-on-a-chip" will require diverse expertise at various levels of abstraction. The increased complexity and geographical separation of design data, tools and teams has created a need for a collaborative and distributed design environment. We have developed a design environment, CollabTop, that facilitates collaboration among remote users.

The simple user interface and global availability of the World Wide Web makes it an ideal platform for such an infrastructure. The tool is developed as a Java applet that can be used from a browser. The framework has a Java-based hierarchical schematic/block editor front-end that supports collaborative functionality. The schematic edit events (e.g., adding or deleting cells to the design) caused by any designer are shown in the editors of all the other designers in that session. Designers can extract netlist from the framework and can submit it to several remote Web-CAD tools, whose input formats are supported by CollabTop.

CollabTop uses client-server method of communication to achieve this collaboration. The server is a stand-alone Java program, running on the MIT Webserver. The events from all the clients are sent to this central server, which broadcasts them to all designer clients in that session. The interaction between the clients and the server are through the message passing mechanism. The Java serialization mechanisms and sockets are used for implementing the message-passing.

It is critical that the consistency of views in the edit windows be maintained through synchronization. The messages from the server should appear in the same order at each of the clients in order to avoid inconsistent views and the tool uses time-stamping of messages for this purpose. Appropriate hooks are provided to let a user join a session already in progress.

The designers need exclusive control on some critical resources (e.g., mouse while adding wires to the schematic) to ease the design process. Designers can obtain locks on such resources which prevents other designers from using that resource till the lock is released. The server monitors the lock owners periodically, so that the tool can recover the locks in the event of failure of such clients.