PREMIER COLLOQUE IMAGE
Traitement, Synthèse, Technologie et Applications

BIARRITZ — Mai 1984 —



# A LINEAR PROCESSOR ARRAY FOR VIDEO RATE IMAGE PROCESSING AND ANALYSIS
# UN MULTI-PROCESSEUR LINEAIRE POUR LE TRAITEMENT ET L'ANALYSE D'IMAGES EN TEMPS REEL

ERIC SWANE

The Foundation of Scientific and Industrial Research at the Norwegian Institute of Technology
N-7034 Trondheim-NTH, Norway

## RESUME

Un systeme multi-processeur destine principalement au traitement et a l'analyse des image est presente. Il est base sur une procedure nouvelle pour la distribution des pixels et la sortie des resultats. Le systeme comprend un certain nombre de processeurs autonome connectes en une chaine de noeuds. La chaine se comporte comme un registre a decalage transmettant les pixels a la vitesse video. Les transformations consecutives des image et les raffinement des donnees extraites sont executes de la maniere d'un pipeline. Le sous-systeme comprenant un processeur et son noeux de communication se trouve dans un seul circuit VLSI (Very Large Scale Integration). Avec ce circuit comme module, de puissants systemes peut etre construits a prix reasonnables. Le systeme est facilement agrandissable, et le traitement en temps reel peut etre obtenu pour la plus part des taches. Une grande variete d'algorithmes est acceptee, donnant le systeme une bonne capacite d'analyse de l'image. Les communications sont simples et robustes, et la tolerance d'erreurs est propre au concept.

## SUMMARY

A multiprocessor system primarly intended for image processing and analysis is presented. It is based on a novel scheme for pixel distribution and result forwarding. The system consists of a number of selfcontained processors connected in a node chain. The chain functions as a shiftregister conveying pixels at video rate. Consecutive image transforms and refinements of extracted data are executed in a pipeline manner. The subsystem consisting of a processor and its communication node fits into one single very large scale integration circuit. With this circuit as building block, powerful dedicated systems may be built at reasonable prices. The system is easily expandable, and video rate processing may be achieved for most tasks. A great variety of algorithms is supported, giving the system a good image analysis capability. Communications are simple and robust, and fault tolerance is inherent in the concept.

# A LINEAR PROCESSOR ARRAY FOR VIDEO RATE IMAGE PROCESSING AND ANALYSIS
# UN MULTI-PROCESSEUR LINEAIRE POUR LE TRAITEMENT ET L'ANALYSE D'IMAGES EN TEMPS REEL

ERIC SWANE

## 1. INTRODUCTION

Image processing is a challenging field for constructors of multiprocessor systems. The availability of very large scale integration (VLSI) technology has increased the interest for dedicated architectures. The parts of image processing which roughly can be categorised as twodimensional signal processing are especially well structured and suited for VLSI design. It is here that commercially available components can be expected first.

Automatic visual inspection and robot vision are to an increasing extent used in manufacturing. Industrial vision has special requirements concerning image analysis capability, real time capacity, system reliability and operation costs. We have prefered rather to meet these needs and present here a multiprocessor system that is aimed towards industrial applications of image processing and analysis.

The presentation is organized in seven sections. In the second section, following this introduction, the node chain architecture is presented. In section three and four the broadcast and pick principle and the exchange principle forming the base of the concept are explained. In section five the use of a general type processor is advocated. Section six is devoted to some practical aspects of the architecture. In the last section we conclude by stressing the importance of modularity for the rational construction of vision systems.

## 2. THE NODE CHAIN PIPELINE

SINTULF is a network of independent processor subsystems, each consisting of a selfcontained processor with its communication node. The nodes of the subsystems are connected together in a chain. The chain functions as a shiftregister with a node at each stage. Data are shifted from node to node at a high transfer rate. The chain may convey pixels at video rate. The pixel stream may originate directly from an image scanner or from the refresh stream of a raster display, fig. 1. The processor system may produce a corresponding result pixel stream. With the pixel stream concept, complex input-output arrangements are unnecessary.

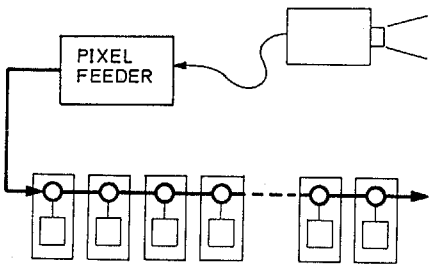


Fig. 1. The node chain pipeline.

The chain is effectively utilized as communication channel by a novel scheme for pixel distribution and result forwarding. Results, either new pixels or descriptive data, are transmitted on the same chain. Groups of subsystems form chain units that function as stages in a processing pipeline. The basic data routing scheme will be explained in the next section.

The computing power is gained by exploiting parallelism primarily along the row axis of the image and secondly by introducing pipelining at the image transform level. The degree of parallelism is freely adjustable up to the limit of one subsystem per image column and pipeline stage. When operating at video rate, the architecture is optimal in the sense that both transfer and computing capacities are fully utilized.

The individual processors are provided with their own instruction memories and program controllers. This categorises SINTULF as a multiple instruction stream multiple data stream (MIMD) machine. The choice of the MIMD structure has several reasons. The data routing scheme is not suited for synchrounous operation of parallel processors. The number of external connections to each subsystem is kept low. It provides the resource allocation freedom required for for flexible expansion and reconfiguration of the chain, for the subdivision in pipeline stages, for the effective exploitation of the resources available, and for the fault tolerance. Finally the processor independency opens for parallel execution of data dependent procedures fundamental to information extraction from images.

In most parallel architectures adding more components usually load the system as a whole either physically or by introducing bottlenecks. Buses or central controllers impose size limitations that prevents real time execution of heavy tasks. In SINTULF the assembly of systems is equivalently easy for any reasonable size. Within the limit of one subsystem per image column there are no functional inconveniences with regard to the number of subsystems included. But with such a quantity of computing power, transforms can be executed at video rate even for very large neighbourhoods. The expandability feature of SINTULF is based on the shiftregister communication channel, the fact that the subsystems are identical and independent, the lack of physical addressing, and the resource allocation facilities. In most cases a linear relation between the amount of hardware and the computing power exploited is obtainable.

When a large number of units are linked together in a chain, the probability of one unit failing and disrupting the chain can not be neglected. However, with a large number of identical and inexpensive subsystems fault tolerance may easily be included in the system. The resource allocation facilities in SINTULF makes program and control parameter loading fast and flexible. An operating system can monitor the system, detect faults and automatically passivate defect units and activate spare ones. Circuit boards with fatal faults directly in their chain channel, can be bypassed by program controlled multiplexers, fig. 2.

Silicon area utilization may seem poor compared to popular approaches like systolic arrays and bit serial SIMD (Single Instruction stream Multiple Data stream) machines. SINTULF has a MIMD structure, but the controllers and the storage for instructions, coeffisients and overlap pixels are in part used redundantly. But the MIMD structure provides a great flexibility and other advantages that for industrial environments do justify the silicon consumption on the individual chips.

A LINEAR PROCESSOR ARRAY FOR VIDEO RATE IMAGE PROCESSING AND ANALYSIS
UN MULTI-PROCESSEUR LINEAIRE POUR LE TRAITEMENT ET L'ANALYSE D'IMAGES EN TEMPS REEL

ERIC SWANE

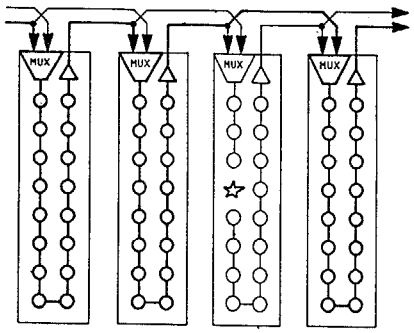


Fig. 2. Bypassing a defect chain board.

## 3. THE BROADCAST AND PICK PRINCIPLE

The two classes of operations based on neighbourhoods are illustrated in fig. 3. The image is either transformed into a new enhanced image, or image information is extracted and condensed into statistics or list structures. In both cases the computations are performed on the data in a neighbourhood around the working position. For neighbourhood transforms a new pixel value is produced. For statistics a describing parameter is computed and the corresponding frequency count updated. The neighbourhoods may be three dimensional and consist of pixel vectors. Any data structure organized in blocks corresponding to the image rows can be handled.

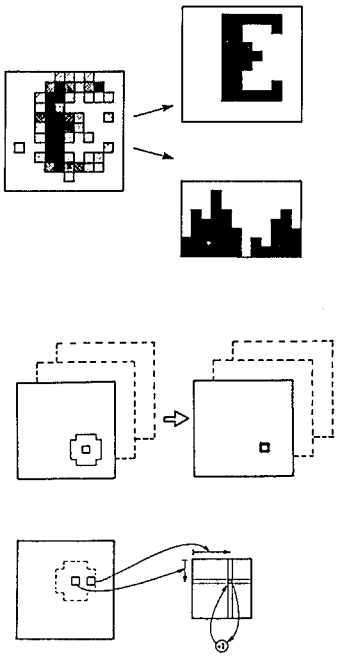

Fig. 3. The two main classes of neighbourhood operations; transforms and statistics.

The basic idea is to broadcast pixels and let the subsystems have the responsibility of selecting and picking the pixels they need. The pixels are transmitted through the node chain in conventional scanning order. A special mark called the head tag is placed in front of each image row. The processors pick pixels from a column interval called a window. The straight forward principle for pixel selection is to count and bypass the pixels between the head tag and the window. The principle allows windows to overlap and overlap pixels to be transferred only once. In practice, the counting principle is replaced by another principle which has the advantage of being independent of image width and chain length and which supports the expandability feature of the node chain concept.

A number of subsystems work in parallel on an image transform. A vertical band in the image is allocated to each subsystem, fig. 4. A neighbourhood covers a small number of rows and columns. Each subsystem may work on a wider window covering several horisontally overlapping neighbourhoods. The subsystems pick the proper pixels from each image row. Only the window rows inside the current neighbourhood need to be available at a given moment. Therefore, data rotate through a ringbuffer as the working position is sliding downwards in the image.

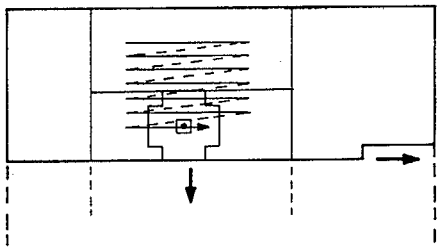

Fig. 4. The working scheme for an individual processor.

Neighbourhood transforms require from a few to thousands of instruction cycles each. All the stages in a pipeline have to work at the same rate. They have to fullfill their task for a certain number of neighbourhoods or data items within the image row time period. When resources have to be fully exploited, each pipeline stage is provided with the appropiate number of subsystems to balance the load of the given task against the row period. The window widths are adjusted correspondingly within each stage.

## 4. THE EXCHANGE PRINCIPLE

When the supply pixels have reached their respective destinations, the corresponding transfer slots on the node chain become free. The results from previously processed rows fit exactly into these vacant slots. Result pixels replace supply pixels in an exchange operation. Exchange takes place in one node at the time. This node is called the exchange point. The point moves along the chain with velocity proportional to the number of subsystems in the pipeline stage. With the broadcast and pick principle and the exchange principle the transfer capacity of the chain is fully utilized at every moment.

The routing scheme may be visualised as the input pixel stream is split into several parallel streams, fig. 5. In the overlap case some pixels are duplicated and directed into several streams. After a certain delay the streams are merged together into a single output stream. The parallel branches represent the processors, each of them working independently on its own pixel stream. The delay, due to the neighbourhood buffering and processing, is approximately equal to the number of row periods corresponding to the neighbourhood height.

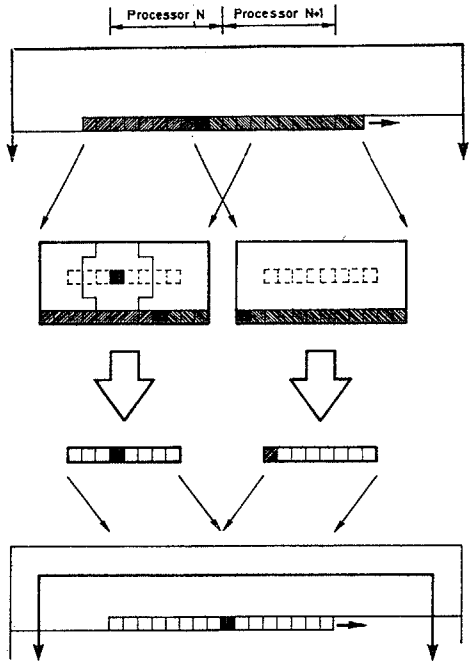


Fig. 5. The pixel routing scheme for two neighbouring processors.

The scheme outlined above applies to one image transform and to subsystems loaded with identical programs. More subsystems may however be linked in as new chain units at the output of the first unit to do further transforms on the same image. The chain units form stages in a processing pipeline. The number of subsystems and the window width is adjusted individually for each group. The final image arrives at the end of the system with a delay approximately equal to the number of row periods corresponding to the sum of the neighbourhood heights involved in all transform stages.

The presented scheme supports the distribution and forwarding of pixels for neighbourhood processing. At the exchange point, which moves along the chain, the supply pixel stream is converted to a result stream. The number of pixels exchanged at each node corresponds to the difference between the window width and the overlapping width, fig. 6. The other main data routing scheme in the node chain concept supports information extraction and refinement. Descriptive data are transferred on a packet basis corresponding to image rows. The process involves data reduction, and the transfer slots on the chain are gradually released to become idle slots. Schemes for program and control parameter loading are based on variants of the same principles.

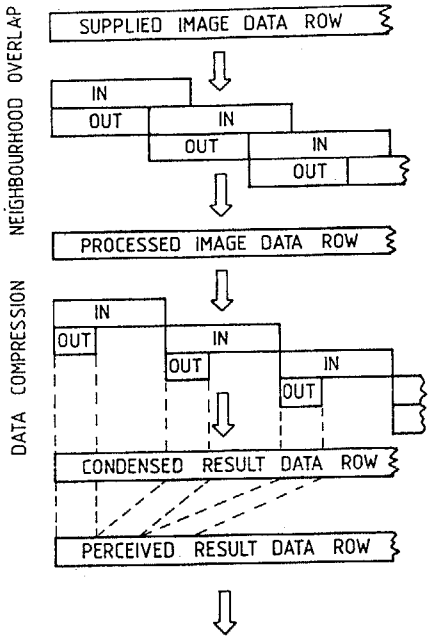


Fig. 6. The main data exchange schemes; neighbourhood processing and data reduction.

## 5. INFORMATION EXTRACTION AND REFINEMENT

The independency of the individual processors is an integral part of the concept. But the design of the individual processors is otherwise not related to the overall architectural solution. For example, they may be specialized towards convolutions or domain transforms. The goal for SINTULF is to support preprocessing, low level and intermediate level image analysis. The node chain in the data reduction mode is appropiate for the stepwize information refinement process in image analysis. As will be shown below by examples, image analysis comprises a great variety of algorithms. These are best implemented by a sequentially working bit parallel processor, or in other words, a conventional processor.

Image processors usually support logical or linear transforms on nine-members-neighbourhoods, but seldom transforms on larger neighbourhoods, data dependent transforms, geometric transforms involving interpolation, and spatially recursive linear and unlinear filters. Some image processors support histogramming as mean for image analysis, but not data dependent navigation, label propagation for object filling and distance mapping, and the enumeration, counting and position registration of events. All these tasks are difficult to implement on multiprocessor architectures, but most of them may be executed effectively by parallely working conventional processors.

Navigation, searches and other procedures of sequential nature are fundamental to information extraction from images. The processors in SINTULF work sequentially and have an instruction set including conditional branch. The processors can operate independently on wide vertical bands overlapping one column. They can track linear features parallely in a number of downsliding windows. The final collection of result lists is coordinated through simple protocols.

## 6. THE SINGLE BUILDING BLOCK

The processors have their own local program memories. They are loaded over the node chain, and common programs are transmitted only once. Algorithms for neighbourhood transforms are usually written as compact instruction loops. The procedures for information refinement can for most tasks be partitioned into smaller functional steps suitable for pipelining. The storage requirement for instructions may therefore be considerated low.

The multiprocessor system is partitioned into a number of identical sybsystems, each consisting of a processor and its communication node. The size of a subsystem allows it, with state of the art technology, to be realized as one single very large scale integrated (VLSI) circuit. With this circuit as basic building block many different processor systems may be constructed. By manufacturing the building block in large series, the component cost is reduced. Few glue components are required, and the price of a complete system becomes reasonable compared to the computational power provided.

The subsystem chip may be bonded into a dual-in-line package. Just a few lines are reqired for the control of the nodes, and no more than 28 pins are required for a bytewide chain. The in and out pins of the chain are placed on either side of the package. The packages are mounted on circuit boards side by side, making the signal paths between the components both simple and short. The shift register approach supports reliability, high transfer rates, and low mounting costs.

## 7. CONCLUSIONS

An effective way of coordinating a great number of conventional processors for real time image processing and analysis has been presented. Topics like performance analysis, subsystem architecture, and host computer interfacing have not been treated.

The field of image processing and analysis is very broad. A lot of systems will be realized in the coming years. Most of them will be specialised towards specific application areas, be flexible within these areas and be easy to operate for nonspecialists. The rational construction of all these systems has to be based on appropriate modules.

The SINTULF building block system is a good hardware base for these systems. Features as flexibility, expandability, capacity, simplicity, robustness, fault tolerance and cost effectiveness makes SINTULF especially apt for industrial environments. The most obvious application areas are automatic visual inspection and robot vision. But the concept of the node chain pipeline is adaptable to most image processing and analysis applications.