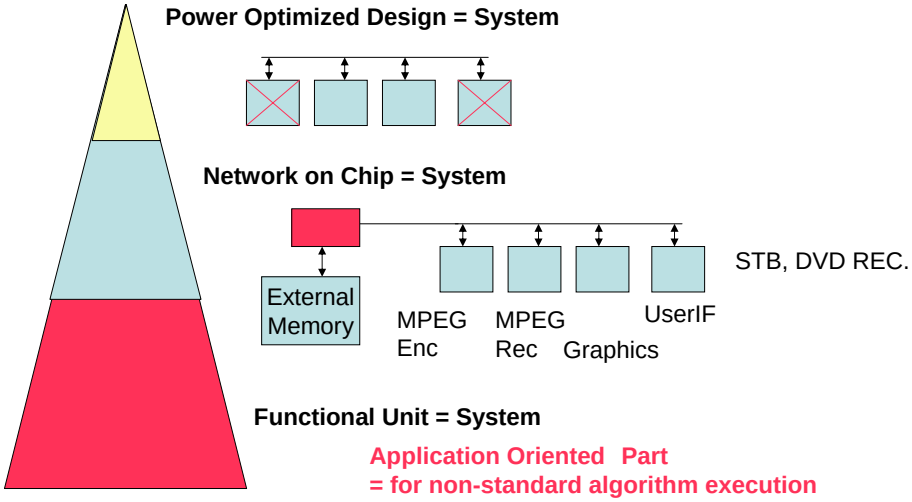
Homogeneous Parallel Processor Architecture for Future Cell Phone Applications

July 8, 2004

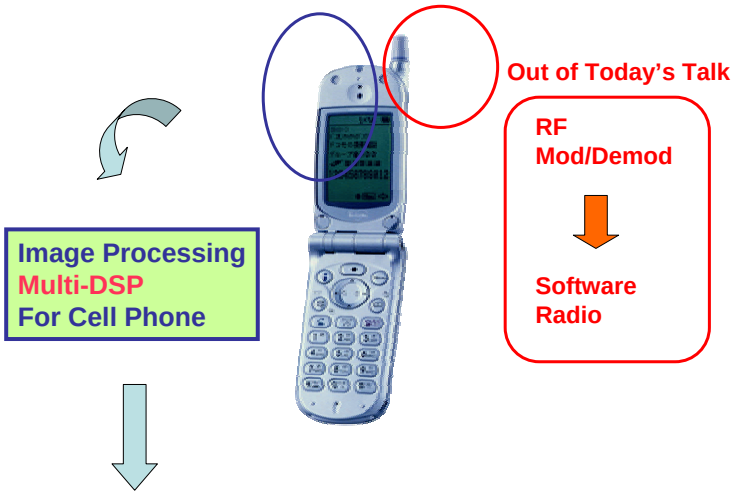
Takao Nishitani
Kochi University of Technology

Talk Point



Application Oriented Part
= for non-standard algorithm execution
which make something different from others

What's MP for Cell Phone SoC

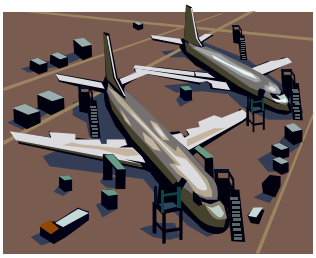


High Processing speed with simple programmability is the most important factor in future cell phone

Outline

- Future Applications
- Requirements of Processing Capability (TOPS)
- TOPS Architecture Candidates
  VSP：SPMD
  IMAP：SIMD
- Evaluation of Power Dissipation
- Conclusion

At Nice Int'l Airport example-1



Immigration

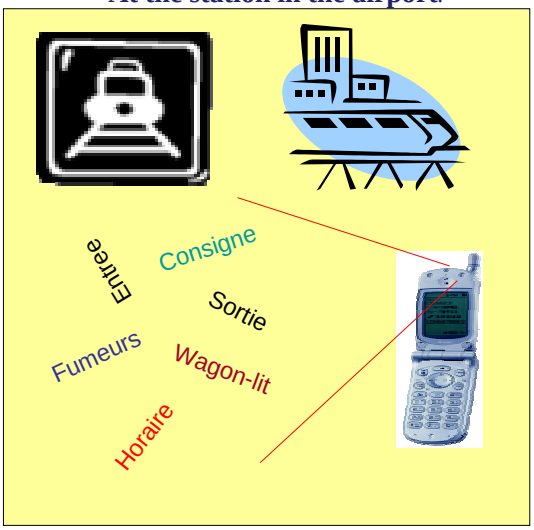
↓

Baggage Claim

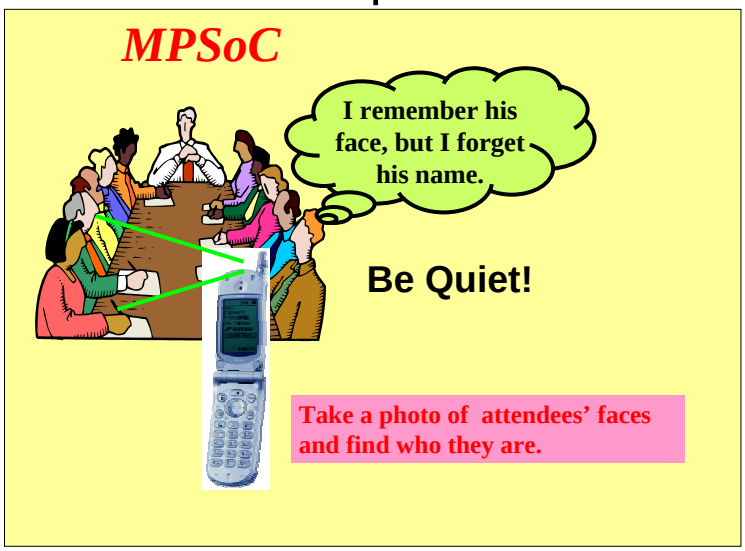
↓

Hurry up to Nice!
However, I cannot read signs in French!

At the station in the airport.



Participating in a Conference example-**1**



Requirements for Cell Phone

✓ **High Speed (TOPS for recognition)**

Camera Resolution reaches more than 3M pixels.
Recognition requires more operations than that of Coding.
Video/Image has potential parallelism in their data.

✓**Low Power (0.5W)**

Long battery operation is required. Leakage Current?

✓**Programmability (Expanding Cell Phone world)**

Ever increasing applications ask for programmability.
A familiar programming language should be better for increasing applications.1

Why so many operations?

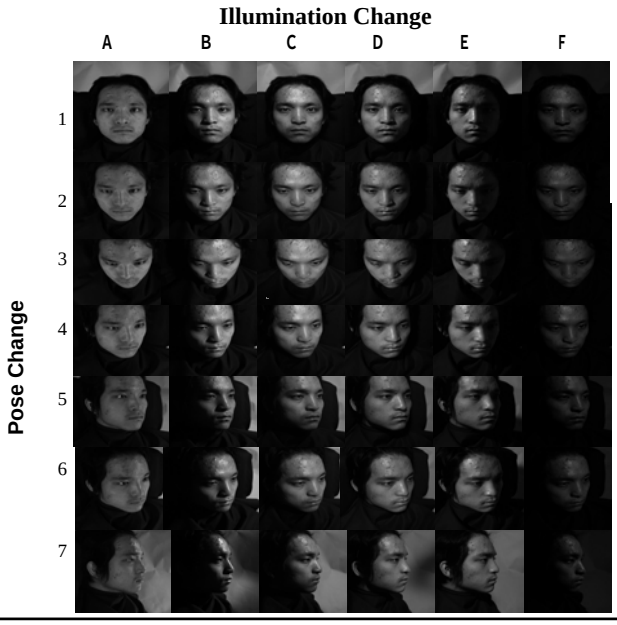
Target:

Object Recognition under Natural Circumstances
Object: Faces and Letters

Difficulty:

Input signals from a camera are generally
Low SNR comes from difference of
(1)Light Condition
(2)Position/Pose
(3)Obstacle Objects Existence

Faces of different poses and illuminations

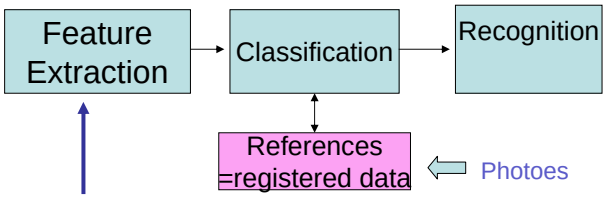


These are photos of the same person

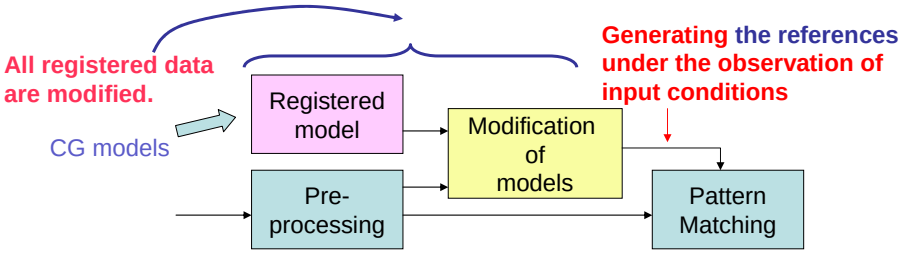
Conventional approaches for these pictures resulted In less than 50% recognition Ratio.

Why TOPS?

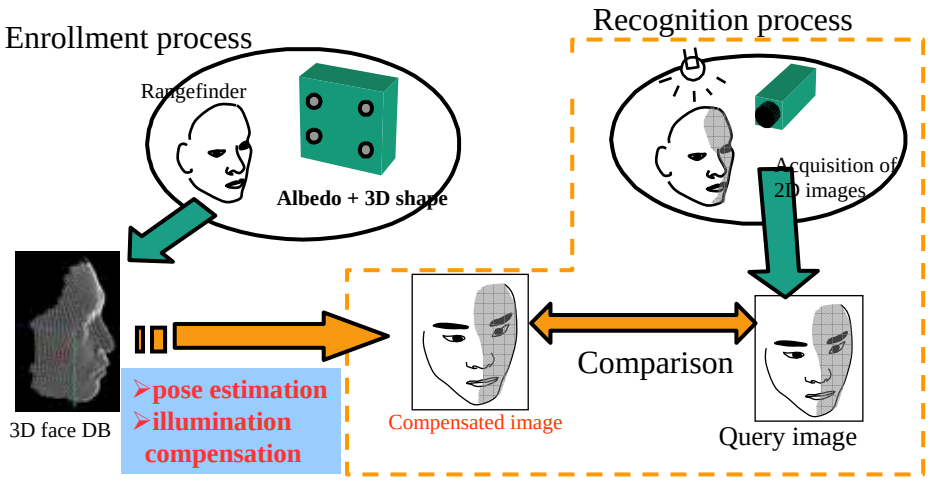
Conventional Approach



Employ DB Modification Approach:



2D(query)-3D(DB) Face Recognition System



Ref: S. Sakamoto, et al., "3D Model Based Face Recognition System with Robustness Against Illumination Changes" NEC R&D 2002 Vol.1.

Results

Conditions:

42 different faces, registered in DB
7 poses and 6 illuminations for each person: 1764 inputs
shown in Page 8.

Results:

Identification Ratio = 96.1%

Face directions in every 5 degree and Matching: 10GO/person

In case of 100 registered people, 1TO is required.

The speed of 1 sec whole processing leads to 1TOPS

Another Example of TOPS DB Modification: Character Recognition
Conventional:50%, DB Modification: 95%, (Postage)

Ref: E. Ishidera, et al., "A Study on Top-down Word Image Generation for Handwritten Word Recognition", ICDAR 2003.

TOPS Cell Phone Candidates

✓ **High Speed Processing (Multiprocessor)**

Parallel Processing by multiple processors
Video/Image domain specific MP is applicable.

✓**Low Power Dissipation (LSI device technology)**

Parallel Processing by conservative LSI process

✓**Programmability (Architecture)**

Homogeneous Processing module for unifying
the control of all programmable portions

Two Approaches

Video/Image Domain Specific Multiprocessor Systems

(1)Video Signal Processors for STV and HDTV (VSP)

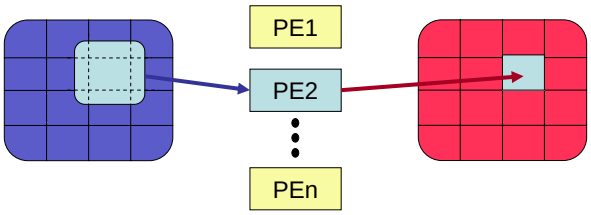
* Based on Overlap-save technique for signal processing
  This means VSP is a loosely coupled Multiprocessor.
* Homogeneous Multiprocessors with Video Rate Busses
* Single Program, Multiple Data Structure
* Loosely coupled processor enables employing standard RISC processor for PE's, resulting C-Language programming.

(2)Integrated Memory Array Processor for Mobile Cruising (IMAP)

* Homogeneous PEs in a linear lines
* Now 128 PEs in a chip
  This means IMAP is a tightly coupled Multiprocessor.
* Expanded C-Language

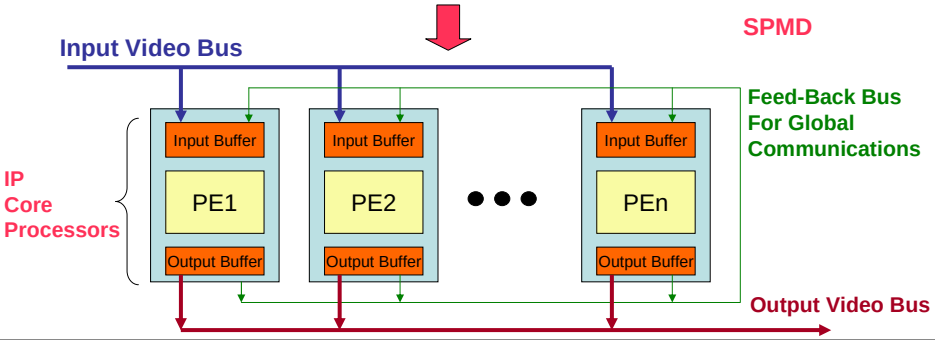
(1)Review of VSP

Every PE fetches larger area than the Assigned Area.



Every PE put the process data at Appropriate timing.

Processing is carried out independently during one frame period.



VSP Summary

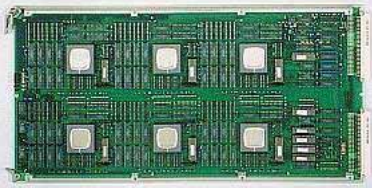
This approach is enough for video encoding and decoding, except total bit rate control, requiring global information of statistics.

This approach was employed in an early version of video codec, shown in a photo.

One possible approach for future Cell Phone MPSOC by employing **standard RISC chips with C-Language.**.

Drawback:

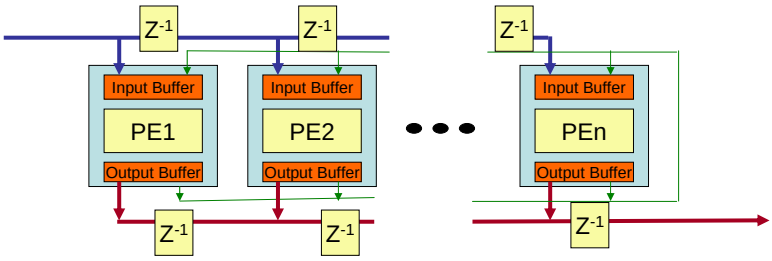
(1) As the input area is usually larger than the processing area, almost 50% of memory is redundant.
(2) MC Engine does not included so as to simplify communications.



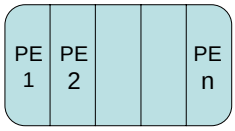
VSP Board for Video Codec

Ref: T. Nishitani, et al., “Parallel Video Signal Processor Configuration based on Overlap-Save Technique and Its VLSI Element: VISP”, Kluwer Academic, J. VLSI Signal Processing, Vol. 1, No. 1, 1988.

IMAP



IMAP is a SIMD Linear Array for column based image Processing,

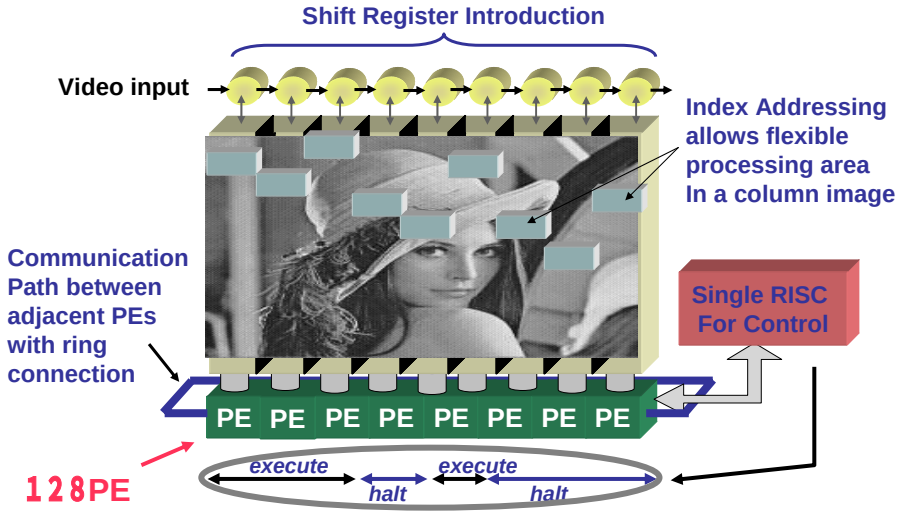


1 Dimensional C Language

and introduces:

(1) connection between adjacent PEs for reducing redundant memories.
(2) Index addressing for keeping arbitrary processing area in PEs in every column.
(3) background I/O operations with shift registers.

IMAP: a SIMD Linear Array Processor



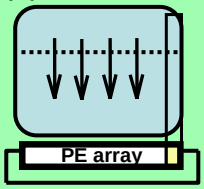
Efficient Selective Instruction

Ref: S. Kyo, et al, "A 51.2GOPS Scalable Video Recognition Processor for Intelligent Cruise Control based on a Linear Array of 128 4-Way VLIW Processing Elements", IEEE J.SSC, Vol.38, Nov. 2003.

Parallel Processing by Index Addressing

The four parallel image processing approaches become possible.

(1)Row-wise

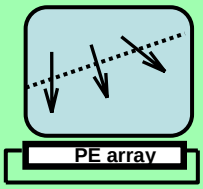


Point/
Local
Neighborhood
Operations:
>2D-Filters

(3)Autonomous

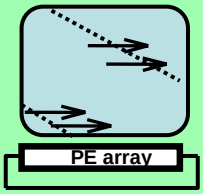
Region Operations,
Such as
Connected
component labeling,
Boundary tracing

(2)Slant-systolic



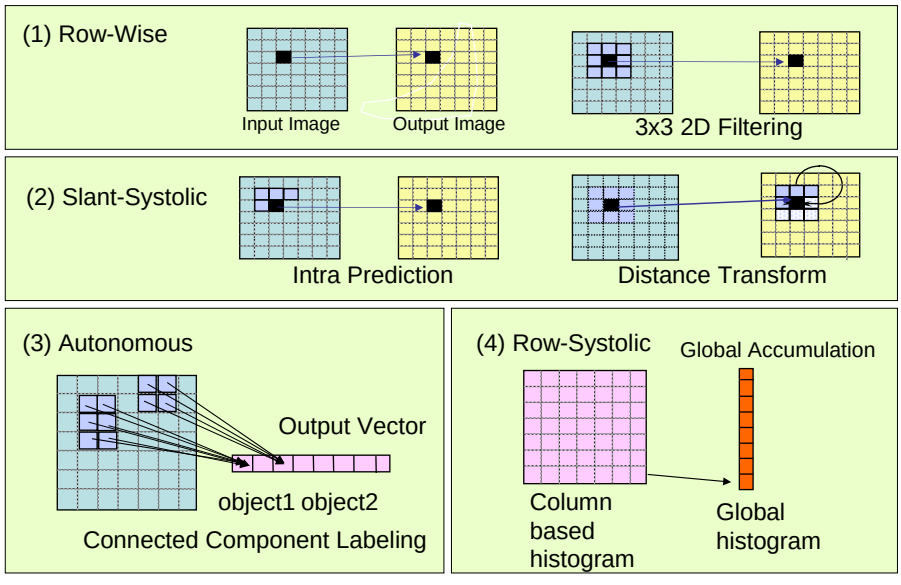
Raster Prediction/
Recursive
Neighborhood
Operation:
>2D Prediction
>Distance Transform

(4)Row-systolic

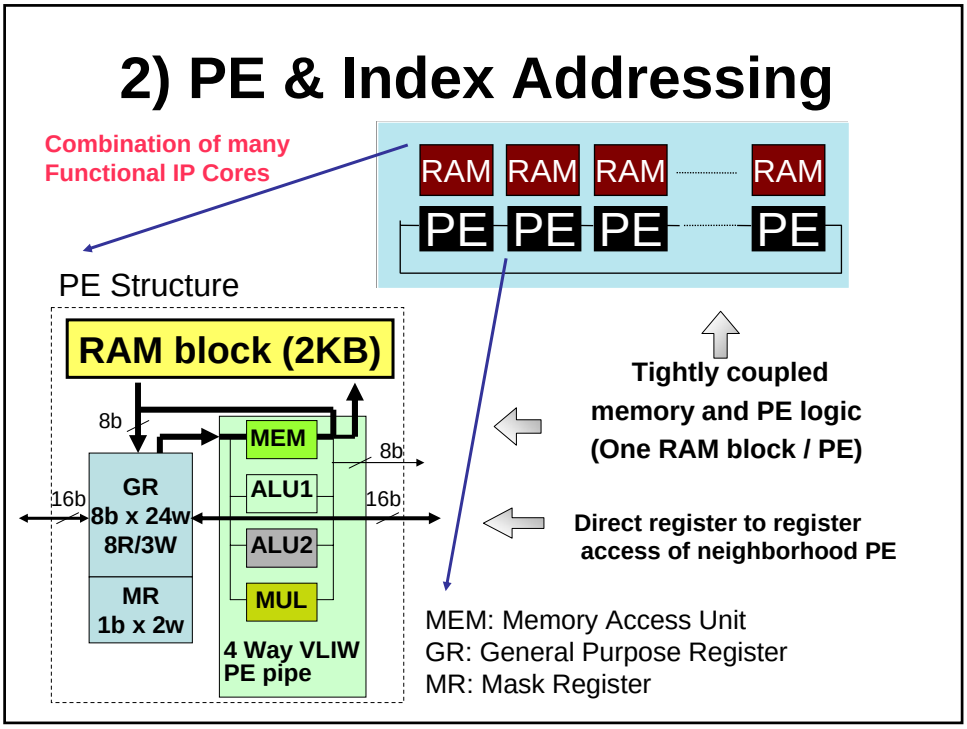


Global Operation/
Accumulation,
Such as Histogram
calculation

Processing Approach Classes



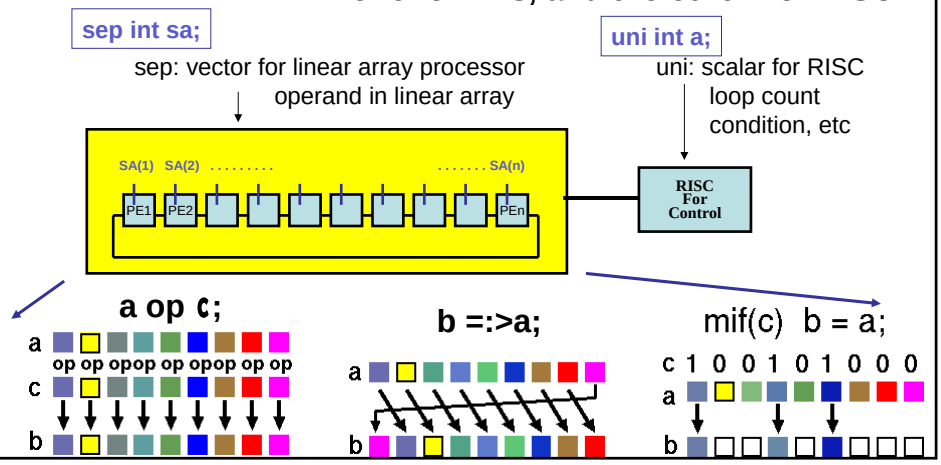
2) PE & Index Addressing



One Dimensional C-Language(1)

One dimensional C-language, expanded from the conventional C, is developed for **explicitly describing** IMAP's parallelism.

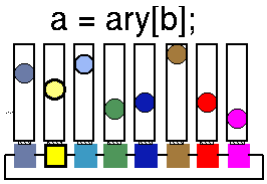
(1)Two different variable class:
one for PEs, and the other for RISC



One Dimensional C-Language(2)

(2) Index Addressing for 2D array makes a vector.

sep char a, b, ary[256];
a = ary[b];

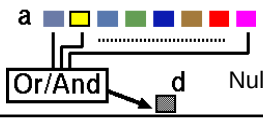


(3) Instructions between sep and uni.

uni int d, e;
Sep char a;

(a) Status of a sep variable

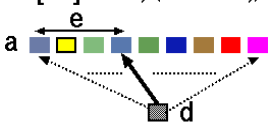
d = :||a; (or :&&a;)



Null Check

(b) Scalar substitution to sep "a"

a:[e:] = d; (or a=d;)



One Dimensional C-Language(3)

(4) Syntax Expansion

(a) Branch based on selected PE portions

mif ~ , melse~

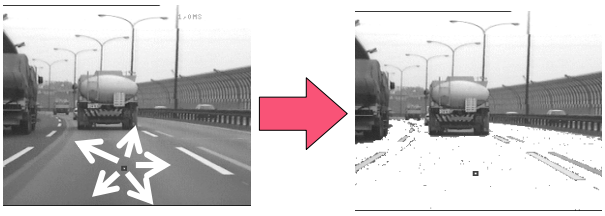
mif (sa==0) ~ PEs specidied by zero elements in SA should do ~.

(b) Loop every PE portions until the condition is satisfied.

mdo~, mwhile~, mfor~

mwhile(sa) ~ loop ~ operations until sa element values on every PEs reach zero.

Real Time Demonstration

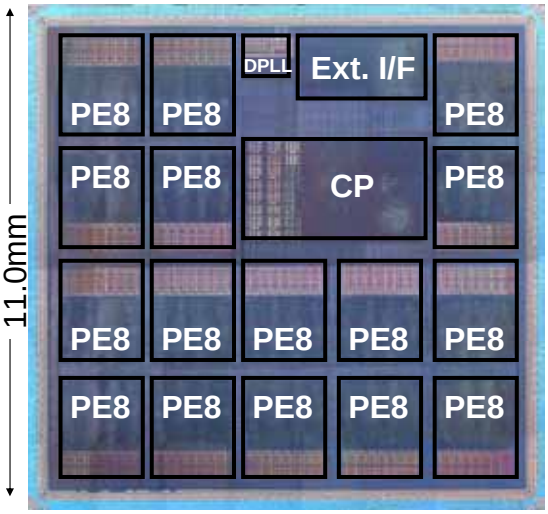


Road area are detected by propagating a label within an area

Front View Surveillance

Back View Surveillance

IMAP Chip Specification
(Future Chip Power dissipation)



(PE8: eight PEs integration block)

50 GOPS Capability

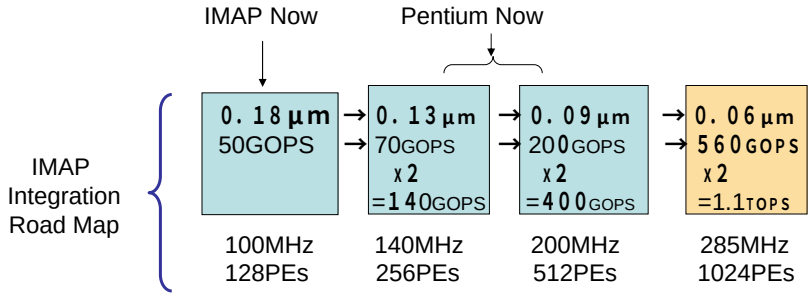
# of processor:	128
# of Tr.	32.7M
Size:	11x11mm
Process:	0.18um
Voltage:	1.8 V
Power:	2.5 – 4.0W
Frequency:	100MHz
On-chip Mem.:	2 Mb
On-chip Mem. Acc. Bw.:	12.8 GB/s
SDRAM I/F:	100MHz (64b)
PCI I/F:	33MHz (32b)

TOPS DSP appears soon

IMAP-CE 50GOPS ➡ Only 20 times more!

Applying Moore's Law:

Physical Parameters reduce to 70% by every process generation



TOPS DSP Power Dissipation

$$P=kCfV^2$$

	C	f	v	Power
IMAP@0.18 (now)	**C1**	100MHz **(f1)**	1.8 volt **(V1)**	**P1=kC1xf1xV1²** **(2.5~4.0W)**
IMAP@0.06	Same die size **(C1)**	285MHz **(2.85f1)**	0.5 volt **(0.28V1)**	P2=k**C1**x**(2.85f1)**x**(0.28V1)²** =0.22P1 **(0.55~0.88W)**

Leak Current: High K + Insulator?

Power management should be employed for Idle PEs.

Homogeneous Multiprocessor will surely open the way to SOC for future cell phone applications!