A 40,000 Pattern/sec Recognition Accelerator with Learning Capability

C. Park, K. Buckmann, J. Diamond, U. Santoni, S.C. The, M. Holler, Intel Corporation
M. Glier, C.L. Scofield, L. Nunez, J. Cole Nestor Inc.

Hot Chips V
August 8-10, 1993

intel Nestor

Outline

- Specialization for Pattern Recognition
- Architecture Overview
- Distance Calculation Units
- Floating Point Pipeline
- Microcontroller/Learning
- I/O
- Specifications/Die Layout
- Summary

Characteristics of Pattern Recognition and Neural Networks

- Pattern Match Measures
  - e.g. inner product, Euclidean, Manhattan, or Hamming Distances
- Low Precision Integer Vector Operations
- Non-linear "decision" functions
- Non-Localized, Unconditional Memory Access
- High Bandwidth Memory Access
  - to Reference pattern memory

Computational Paradigm

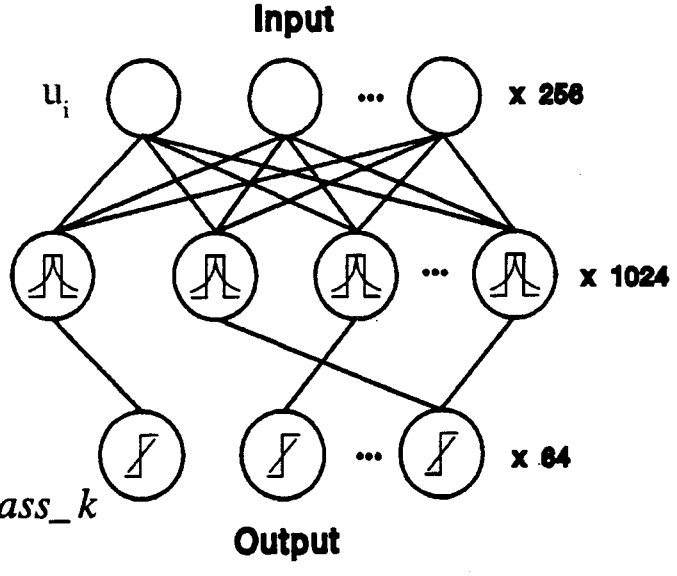
- Probability density functions (pdfs) are estimated, and Bayes' Decision Rule is used to make classifications
- Each output is a mixture of Radial Basis Function

$$D_j = \sum_{j=R_{off}}^{j<Dim+R_{off}} |u_i - W_{ij}|$$

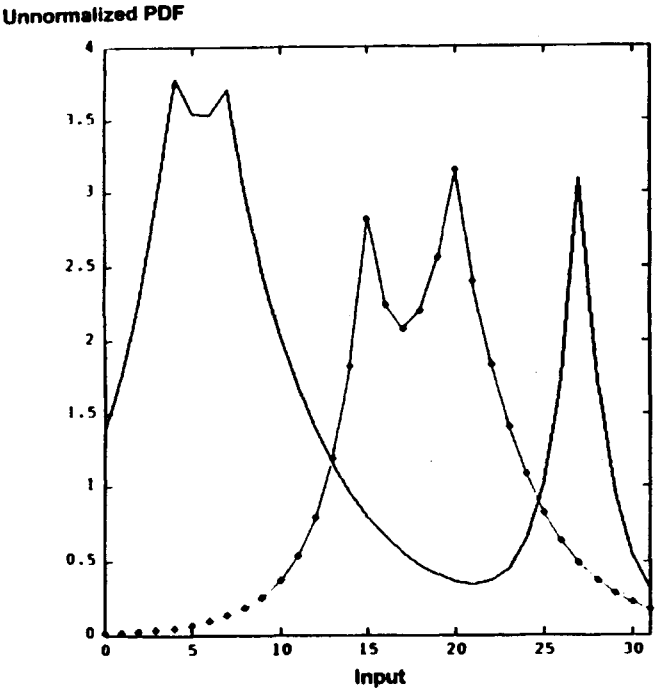
$$rbf_j = C_j 2^{-K_j D_j}$$

$$pdf_k = \sum_{j \in class\_k} rbf_j$$

F_{class_k} = 1 if $D_j < L_j$ for $j \in class\_k$
0 otherwise



PDF Estimation for One Dimensional Case

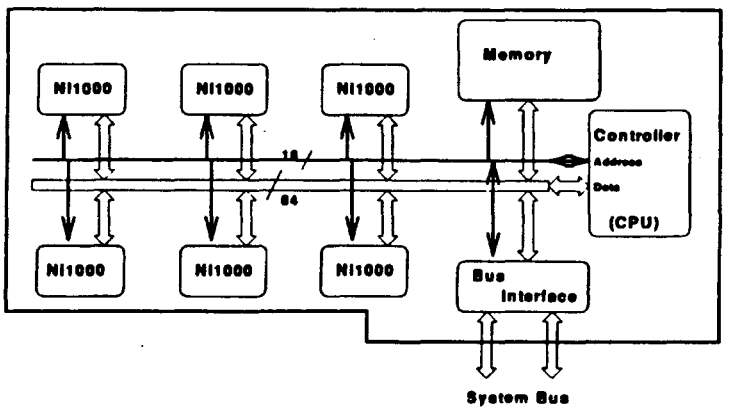


Pattern vector: 5 bit x 256 dim
Distance: 13 bit
Floating point: 16 bit

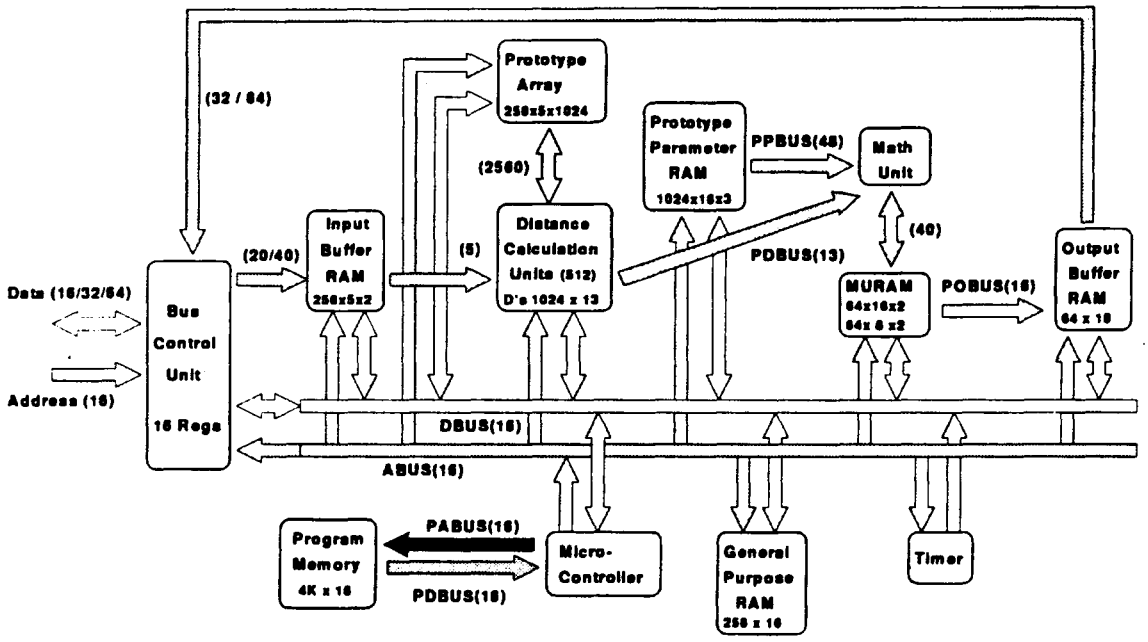
Center	C_j	K_j	Class
4	2	0.5	A
7	3	0.25	A
27	3	0.875	A
15	2	0.75	B
20	3	0.375	B

Typical System Configuration

- Classification accerelator board with single or multi-chips
- Share common data and address buses with local host
- Host controller broadcasts patterns and reads results sequentially using CS and the interrupt features of Ni1000

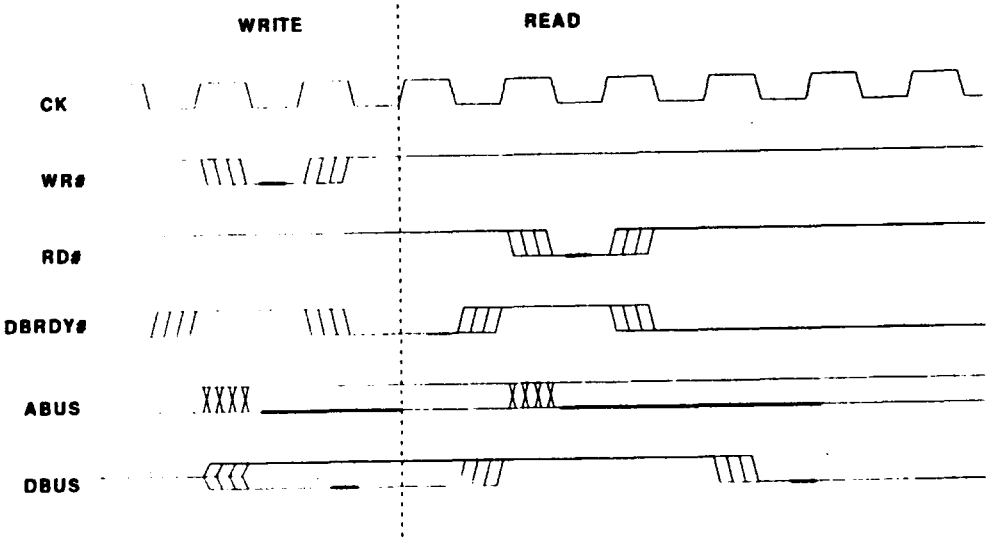


Ni1000 Architecture



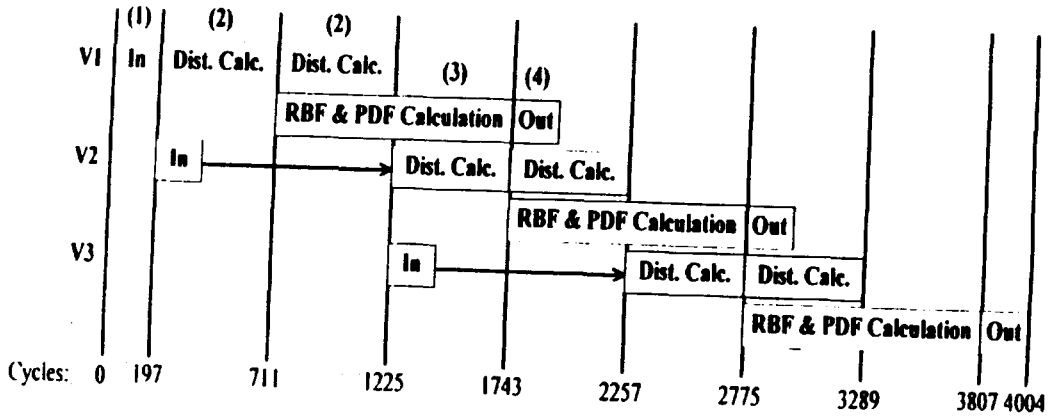
All arrays and registers are mapped with 16 bit address (PA uses column and row addresses)

Internal Bus Timing



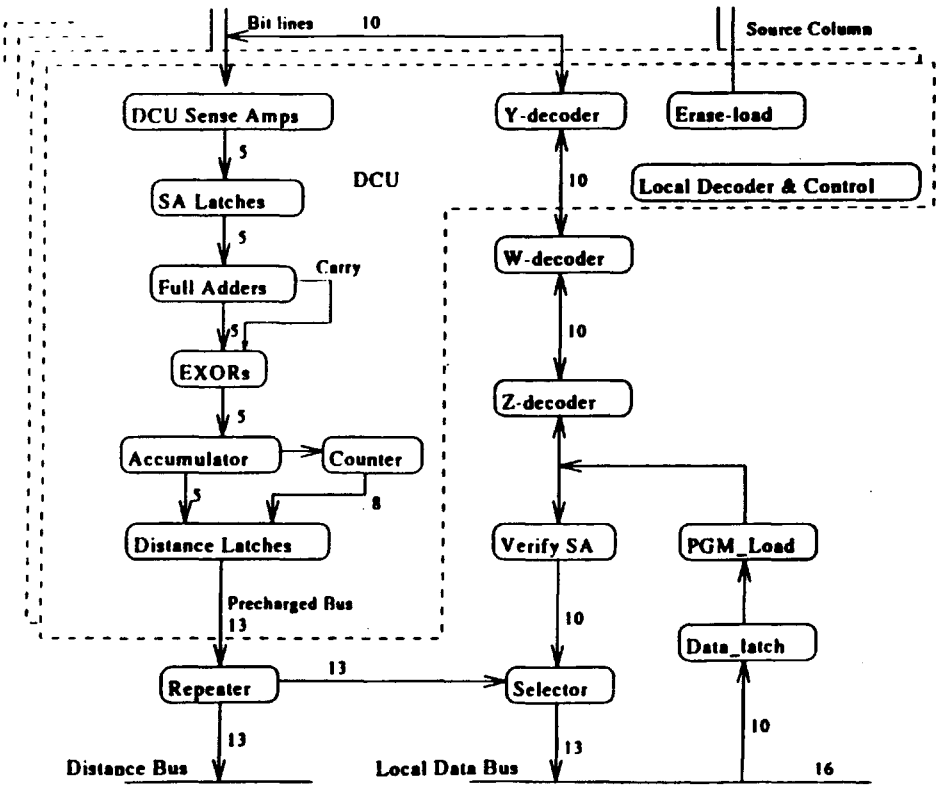
Minimum CPI for write : 2 cycle
read : 3 cycle

Classification Pipeline Timing

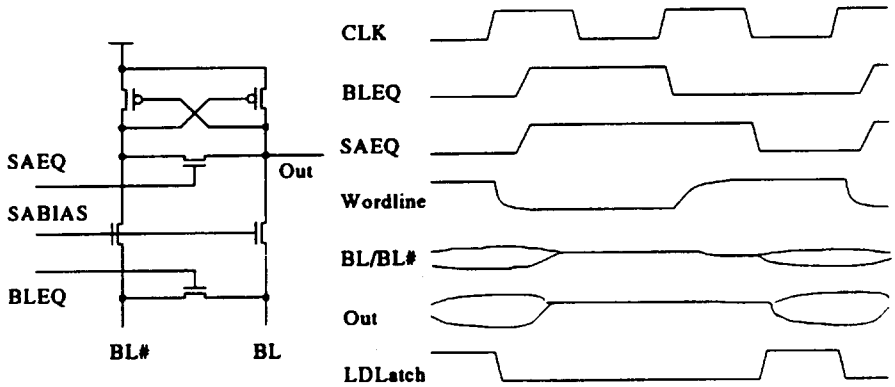


1) 197 Cycles: 256 dim. in 32 bit non-burst mode
2) 514 Cycles: 512 D. Calculation for 256 dim.
3) 1032 Cycles: PDF Cal for 1024 Prototype vectors
4) 197 Cycles: 64 PDFs output in 32 bit non-burst mode

Distance Calculation Unit



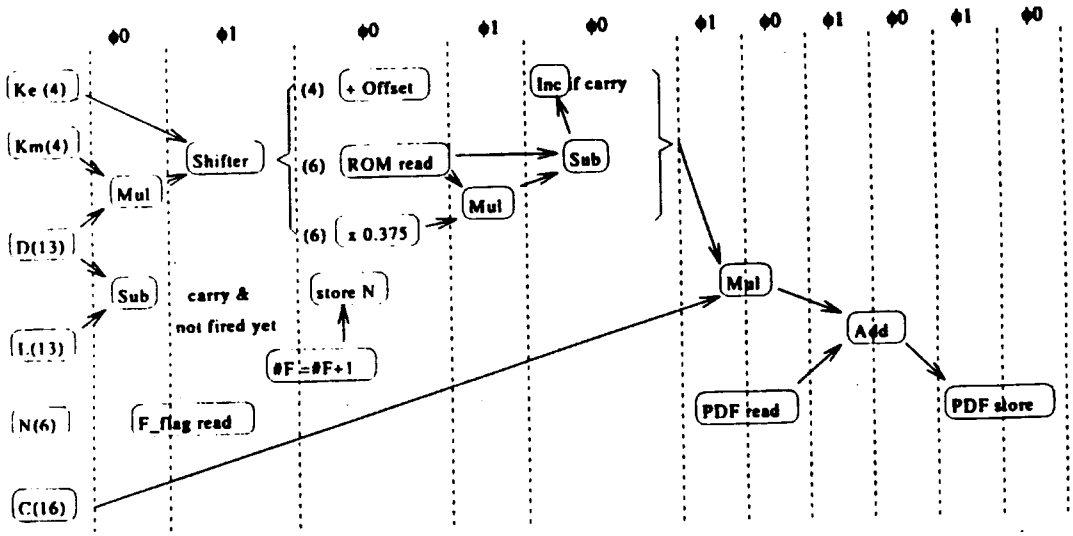
DCU Sense Amp and Timing Diagram



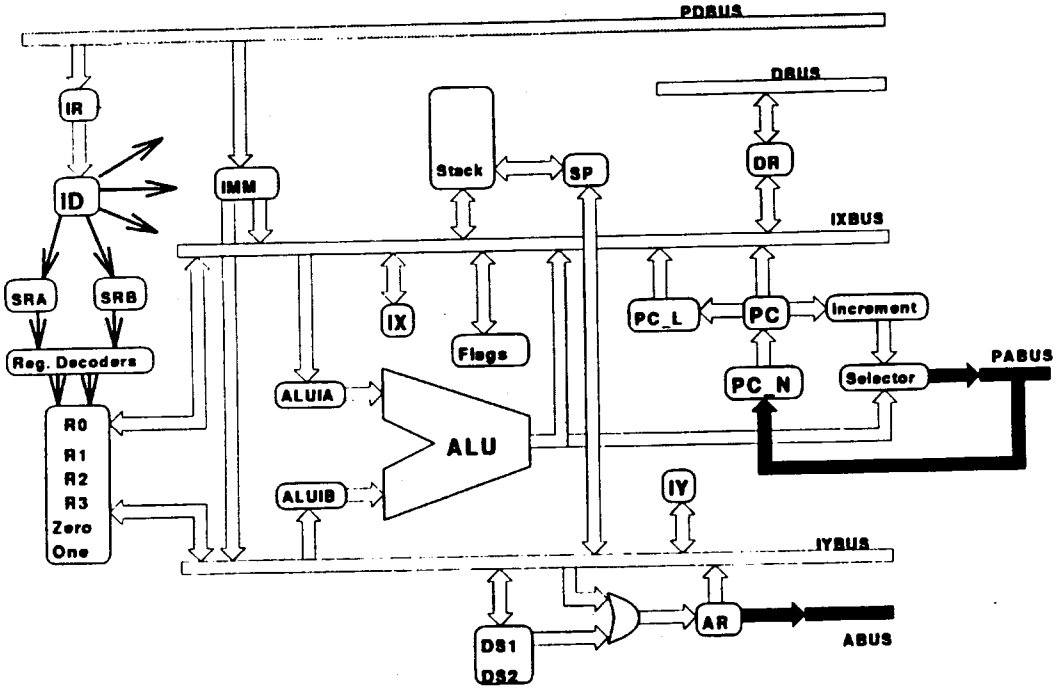
- 2560 DCU Sense Amps in the chip
- 50 nsec Access Cycle

Math Unit Pipeline

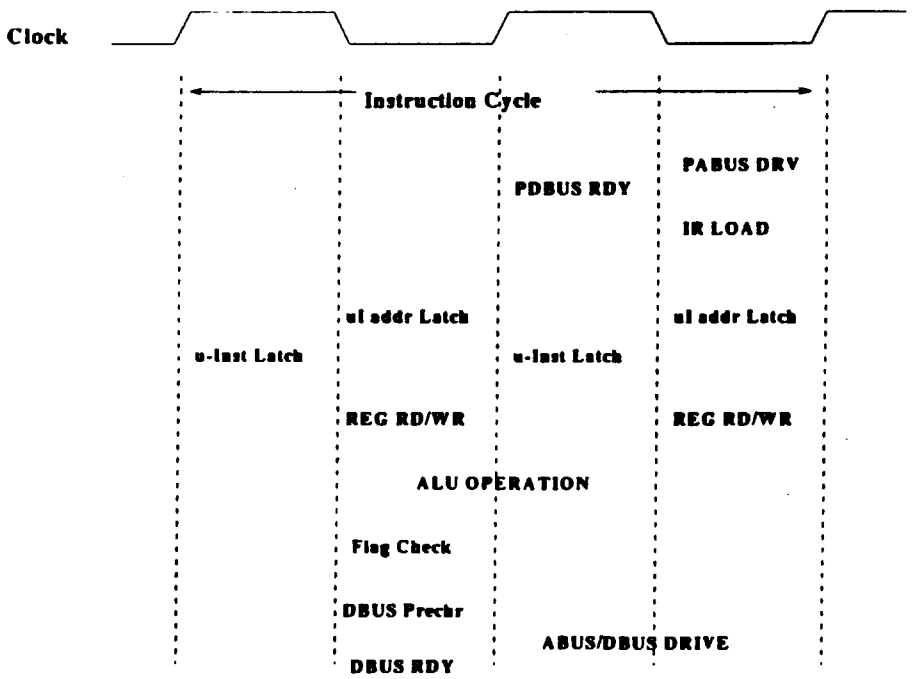
$$PDF_{Class-i} = \sum_{j \in Class-i} C_j 2^{-K_j D_j}$$



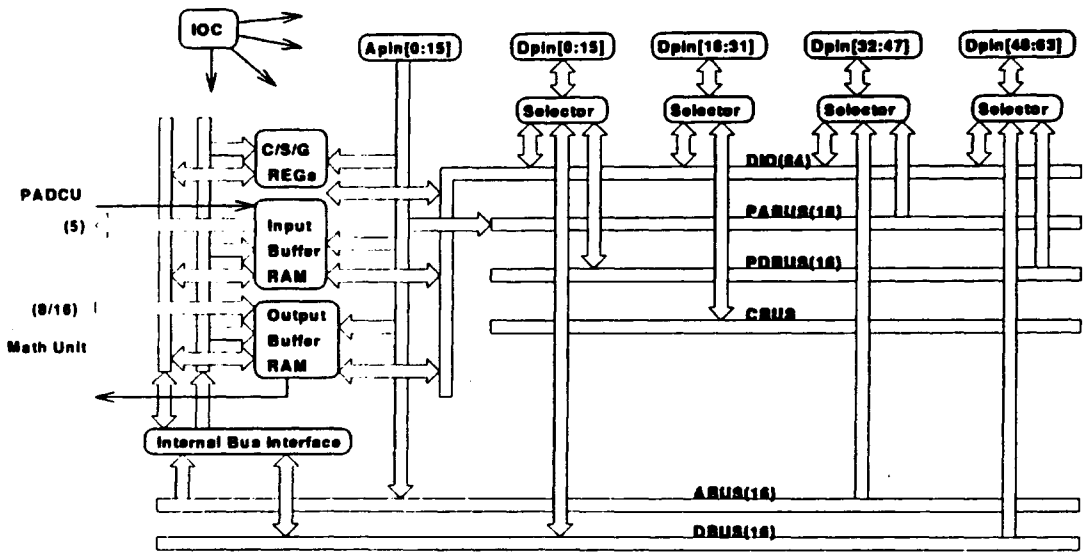
Microcontroller Architecture



Instruction Cycle Timing



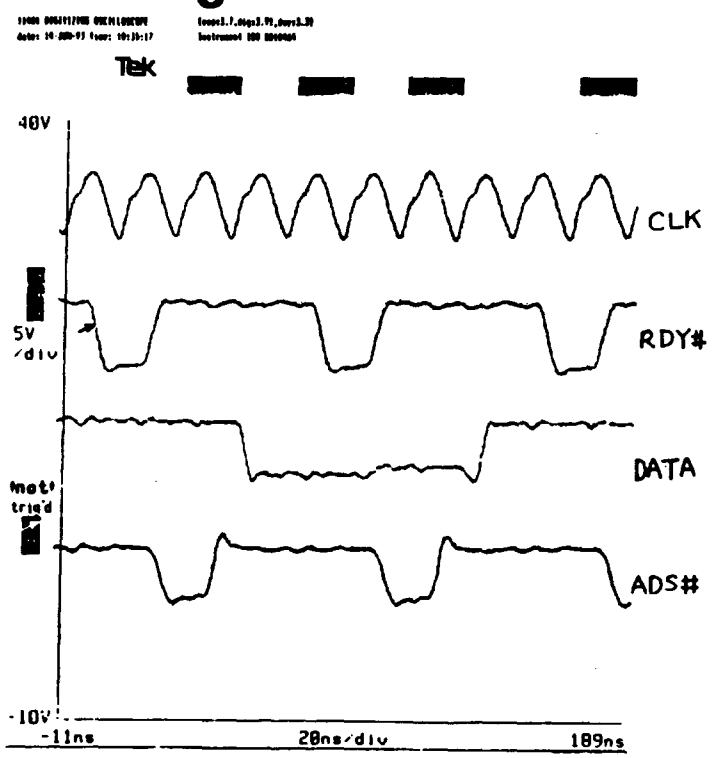
I/O Unit

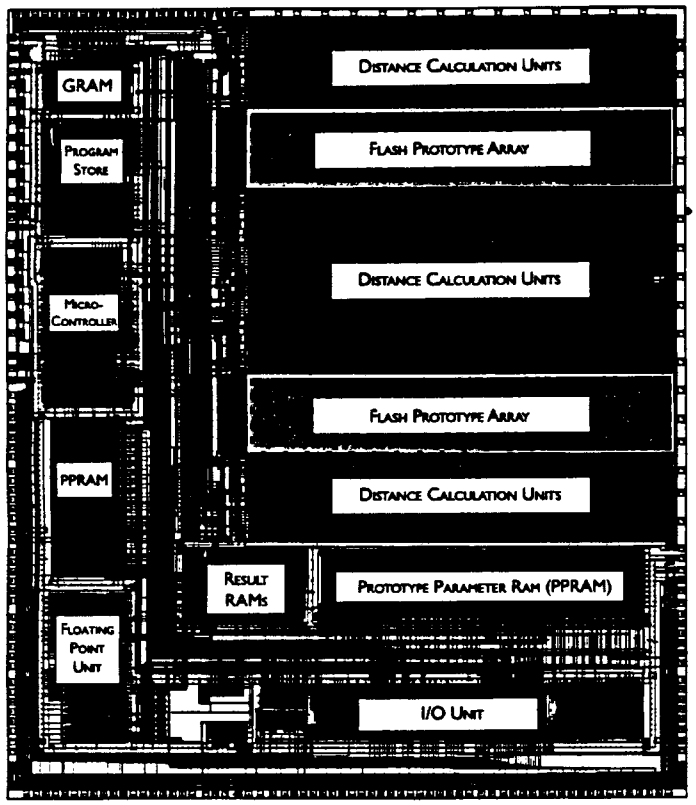


- Non-Burst mode I/O Access needs 4 clocks
- Monitor mode outputs internal bus data (A/D, PA/PD)

intel Nestor

I/O timing Measured Results.





Die Layout

Die size 15.7 x 13.8 mm^2

intel Nestor

Specifications

• Clock frequency	40 MHz
• Performance	
DCU	20 GOPS
FPU	160 MFLOPS
µC	20 MIPS
• Array access BW	10 Giga-word/s (5-bit word)
• Prototype vector size	256 dim x 5 bit
• Number of prototypes	1024 (256 dim) -- 8192 (32 dim)
• Number of transistors	3.75 M trs
• Die size	15.8 mm x 13.7 mm
• Power dissipation	1W typical, 4W peak @ 5V Vcc, 12V Vpp
• Process	.8u CMOS, 2 Metal, 2 Poly, Flash ETOX™
• Package	168-pin PGA

Summary

- A parallel and pipelined architecture has been used to accelerate recognition tasks
- High Density Flash Memory enabled full on chip reference pattern storage.
- It is expected to set a new cost/performance benchmark for pattern recognition computing