Digital Word-Parallel Associative Memory in 180nm CMOS for Nearest Euclidean Distance Search Based on Distance Mapping into Clock-Number Domain

Toshinobu Akazawa, Seiryu Sasaki, and Hans Jürgen Mattausch
Research Institute for Nanodevice and Bio Systems, Hiroshima University, 739-8527, Higashi-Hiroshima, Japan,
E-mail: akazawa-toshinobu@hiroshima-u.ac.jp

Abstract- Word-parallel digital Euclidean Distance (ED) search associative memory based on distance to clock-number mapping is reported. Key enabling concepts are area-efficient sequential square calculation and clock-number minimization per search. An 180nm CMOS test-chip with 32 reference vectors, 16 components, and 8 bit per component was designed for concept verification. At 47MHz and $V_{dd}=1.8V$, fast average search time of 1.19us for code-book-based data compression of images, low power dissipation of 8.75mW and full functionality down to $V_{dd}=1.2V$ were measured. To our best knowledge, practical word-parallel digital architecture for nearest ED search has not been reported previously. The achieved power delay product is even a factor 1.6 (2.5 at $V_{dd}=1.2V$) smaller than with previous digital-analog circuitry.

1. Introduction
Many applications, such as image/video pattern recognition, authentication in security systems or code-book-based data compression require the pattern matching operation [1, 2]. While efficient integrated circuit (IC) solutions for word-parallel Hamming-distance pattern matching based on associative memory are known [3, 4], most practical applications need Euclidean distance (ED) based solutions to obtain highly accurate recognition performance. Unfortunately, practically efficient digital word-parallel implementation architecture for ED-based pattern matching is to our best knowledge unknown today.

2. Algorithm and implementation architecture
ED is the accurate distance of feature vectors with e.g. M-components and N-bit per components. A recognition application shall have R reference vectors. Between the i-th reference vector $R_i$ and an input vector $I$ for recognition processing, ED is defined as

$$ED(M,i) = \sum_{j=1}^{M} (R_{ij} - I_j)^2$$

where $j$ refers to the j-th component. For pattern matching ED$^2$ can replace ED to avoid root calculation [5].

Architecture and processing-flow overviews are given in Fig. 1. Storage cells for reference vectors, distance calculators (DC), distance evaluators (DE), and a bit activator (BA) are the main sub-circuits. For absolute difference (AD) calculation of vector components, an area efficient adder-based circuit [6] is implemented. AD$^2$ of vector components is calculated in each DC in parallel when the SQ signal is asserted. DC results are connected to the corresponding DE units for ED$^2$ mapping into clock-cycle number and word-parallel comparison for all reference patterns. In combination with the bit activator (BA), DE implements a minimization algorithm for the clock number per search. The circuit structures of DC, DE and BA are explained in the block diagrams of Figs. 2-4. Fig.3 shows 1-bit slice of the DC circuit. AD$^2$ calculation is done by serial addition of partial products, reusing the adder for AD calculation and thus maintaining small area. After N clock-cycles final AD$^2$ results are stored in the DFFs.

Direct clock counting [7] achieves short search times when winner-input (WI) distances are small, but worst-case search times become very long, amounting to $N(2N-1)$ clock cycles for ED$^2$ search. Therefore, another key innovation is an efficient search algorithm, which starts from the most significant bit (MSB) and progressively advances to lower-value bits. Our DE solution for 1 feature-vector component is explained in Fig.4. A weighted value counter (WVC), a match detection circuit (MDC) and a MUX are used. First, the comparison between DC outputs and WVC status is restricted to the MSB of all components. Counting starts at the 1st component and advances to the 2nd component when the MDC detects the MSB match (BM$_{1\text{st}}$). MSB-based search continues until at least one pattern issues a match signal (ABM$_{1\text{st}}$) from the MDC of its M$^{\text{th}}$ component. The BA (see Fig. 2) then uses 1 clock cycle for bit-activation-signal (BAS) generation to expand clock counting to the next significant bit and to continue clock counting again from the 1st component. In general, BA expands clock counting from bit k to (k-1) when a match signal for all N-k+1 most significant bits in all components is received from at least one reference pattern. The matching process finishes with the ABM$_{1\text{st}}$ signal of the winner-pattern after expansion to the least significant bit. The described clock-based search algorithm reduces the exponential complexity by sequential counting into clock cycles. The area minimization of the ABM circuit with the 1st search time (SE) is achieved by calculation (AD$^2$) of the ED$^2$ distance calculation for all components. Under these conditions, the next AD$^2$ calculation is performed in parallel for all components.
increase in worst-case ED search time with component-bit number \( N \) to a linear increase according to the equation \( 8N^2 + N - 1 \). Even worst-case clock numbers for ED-based matching become thus quite small.

## 3. Experimental results

The fabricated test chip in 180nm CMOS, implements 32 reference vectors (16 components, each 8-bit). Die photograph and performance data are shown in Fig. 5. Total Si-area is 4.31 mm\(^2\), and BA for clock number minimization uses only 3.9% of this area. Maximum operating frequency is 47MHz (@\( V_{dd} = 1.8V \)), when all 8-bit vector components are used. Fig. 6 shows operating frequency and power consumption down to low supply voltage of 1.2V. Power consumption is reduced from 8.75mW (47MHz, \( V_{dd} = 1.8V \)) to 2.80mW (24MHz, \( V_{dd} = 1.2V \)). The average search-time \( T_{SAV} = n_{WAV} \cdot T_{CLK} \) is determined in applications by the average WI distance \( (n_{WAV}) \) and the clock-cycle time \( (T_{CLK}) \). Code-book-based data compression is used to evaluate the magnitude of \( n_{WAV} \) for a practical case. Fig. 7 shows the clock-number histogram for pattern matching in this implemented minimization algorithm. Direct clock counting leads to \( n_{WAV} = 5249 \) for nearest ED\(^2\)-based search and therefore long \( T_{SAV} \) of 111.68\( \mu \)s. With our minimization algorithm, \( n_{WAV} \) is reduced to only 56 and thus \( T_{SAV} \) becomes about 94 times shorter for this example.

Figs 8, 9 show the comparisons of search time and PD (Power-Delay) product to microprocessor-based approaches (Intel 17-3970x, Atom 330) in advanced 32nm CMOS. It is verified that the reported solution, even when using much less advanced 180nm CMOS, realized more than 4 orders of magnitude higher energy efficiency with in addition much smaller silicon-area consumption (exact micro-processor areas are not known). The comparison of the achieved normalized PD product to previous work, based on a combined digital-analog design [5]: 350nm CMOS, 5 bit, 16 components, 46 reference vectors, is shown in Fig. 10. Besides error-free search, the reported fully-digital ED-based pattern-matching solution achieves also a factor 1.6 (\( V_{dd} = 1.8V \)) and 2.5 (\( V_{dd} = 1.2V \)) smaller normalized power-delay products.

## 4. Conclusion

The reported word-parallel digital ED-based pattern-matching architecture realizes scalable and error-free searching by mapping distances into a clock-cycle number. For minimization of implantation area for square calculation an area-efficient sequential partial product addition is realized and for fast search a clock-number-minimization algorithm is applied. Experimental verification is done in 180nm CMOS, implementing 32 reference vectors with 16 dimensions of 8 bit. Average pattern-matching time of 1.19\( \mu \)s (reduction factor 94 over direct counting) for the application of code-book-based data compression, 8.75mW power dissipation and 5.77\( \mu \)s worst-case matching time are measured at 47MHz and \( V_{dd} = 1.8V \). Correct test-chip operation is verified down to \( V_{dd} = 1.2V \). Improvements over previous high-speed digital-analog architecture [5] are portability to advanced design rules, error-free nearest Euclidean distance search and basically unlimited reference-vector scalability (number and dimension). Additionally, a reduction of the normalized power-delay product per matching operation by a factor 1.6 at \( V_{dd} = 1.8V \) (factor 2.5 at \( V_{dd} = 1.2V \)) is obtained in the test-chip design.

### Acknowledgements

The VLSI chip was fabricated through the chip fabrication program of VDEC, the University of Tokyo in collaboration with, Rohm, Silvaco, Synopsys, and Cadence.

### References