

# Digital Word-Parallel Low-Power Recognition SoC for Mobile Equipment Based on Nearest Euclidean Distance Search and KNN Classification

S. Yamasaki, T. Akazawa, F. An, and H. J. Mattausch

Research Institute for Nano device and Bio Systems, Hiroshima University, Higashi-Hiroshima, 739-8527, Japan

Phone:+81-82-424-6265 E-mail:yamasaki-shogo@hiroshima-u.ac.jp

**Abstract-** An integration architecture for the K nearest neighbor (KNN) recognition algorithm using a nearest Euclidean-distance search associative memory and a classification circuit for the K nearest neighbors with selectable K is reported. The Euclidean-distance search is based on mapping the distances between input data and reference data into an equivalent clock number. An SoC design example of the complete KNN system is implemented in a 180nm CMOS with 32 reference vectors and 2~8 classes. Small silicon area of 3.51 mm<sup>2</sup>, low power consumption of 5.02mW and a maximum operating frequency of 42.9MHz (at Vdd=1.8V) are achieved, which makes the architecture very suitable for mobile devices. An architecture extension for obtaining high classification flexibility is also proposed.

## 1. Introduction

In recent years, a variety of application developments which require pattern matching, such as database searching, handwritten character recognition, face/object recognition in images, or voice recognition have become hot topics [1, 2]. Many users desire these applications in rapidly innovating advanced mobile devices, such as the smart phones.

Conventionally, pattern matching and other computation intensive processing are off-loaded from the mobile device to the cloud, i.e. to data centers with large highly parallel servers. Due to the required data transmission between terminal and cloud, this method involves relatively large time delays. Furthermore, a cloud connection isn't available everywhere and may sometimes be disrupted. Also, total power consumption is very high due to data transmission and external server processing. Therefore, local execution on the mobile device is desirable, even though a software solution is not possible, due to limited power supply from the battery which doesn't allow an on-board CPU with high processing power.

Important classifiers for recognition are based on the K nearest-neighbor (KNN) algorithm. For flexible integrated KNN hardware, which combines high searching speed for the most similar data in a reference-data base with low power consumption, high search reliability and accurate recognition, wide applications in mobile devices can be expected. Furthermore, such a VLSI chip has the potential of becoming an application specific standard product (ASSP) for intelligent systems with recognition capability. Unfortunately, there is no previous example of efficient VLSI integration of the KNN algorithm without algorithm simplification and reduced matching accuracy. This paper reports a flexible low-power recognition SoC (System on Chip) based on the KNN algorithm.

## 2. KNN Realization with searching by clock counting

KNN is a method of statistical classification, based on multiple reference samples closest in the feature space, which is often used in pattern recognition. The number of reference data may be very large and KNN algorithm consistency is fairly reliable. However, computation amount is large because all data must be compared, which is one of the reasons that no efficient VLSI realization of KNN exists. Here, we adopted the distance-search method by clock counting which allows fully-parallel low-power Euclidean-

an-distance search [3]. Fig. 1 shows the block diagram of KNN-algorithm integration with associative memory operating in the clock domain. After calculating the distances between each component of the feature vectors for reference and input samples, these component distances are expressed as a number of clock cycles. During the distance search by clock counting, match signals are received from the reference-data rows in the sequence of their distances to the input sample. In this manner, fully-parallel error-free comparison of data becomes possible. On the other hand, a disadvantage of the clock-counting search is that it takes a long time if the distance of the most similar data is large. The worst-case clock number for the SFCC (Straight Forward Clock Counting) method increases exponentially with the number N of feature-vector component bits. The previously reported CCR (Clock Counting Reduction) method [3] achieves reduction to a linear increase by starting the search from the highest-value bit and including lower value bits in the search sequentially after a match for the higher-value bits has been found.

Match signals are received from the reference-sample rows in the sequence of their distances to the input sample. By summing the match-signal number the K-th nearest neighbor discovery is recognized and the search is terminated by a stop signal. After this, the recognized-class identification by a majority vote is carried-out with an identification circuit that can be freely set to the number K of interest. One evaluated concept identifies the class of each of the K nearest neighbors by the storage location in the associative memory. However, with this concept the reference data has to be sorted and the data number for each class becomes inflexible. Figure 2 shows the preferred architecture for the "KNN clustering circuit", which enables complete flexibility with respect to the number of classes, the number of reference data for each class and the storage locations of the reference data in the associative memory. This flexibility is achieved by a look-up table stored in an SRAM which can be changed for each application and specifies the class information of the reference data in each row of the associative memory. The "Match Signal Storage Registers" on the left are connected to "Match Signal Detecting Circuits" which sequentially identify the row numbers of the K nearest neighbors. For nearest-neighbor row i, an  $act_i$  signal is generated and used to read the correct class information from a look-up table. This class information then controls a de-multiplexer to increase the status of the corresponding class counter by 1. Class identification and counting process are finished when the  $next_R$  signal is asserted at the output of the last "Match Signal Detecting Circuit". Afterwards a comparator at the output of the class counters identifies the recognized class with the highest counting status of the corresponding class counter as the recognition result. The associative memory rows, where the stored data is not among the K nearest neighbors, are skipped in this clustering process. Therefore, the clustering process finishes with the recognition result after K clock cycles.

### 3. Expansion method for feature-vector dimensions

The required number of feature-vector dimensions is different for each application as for example, face recognition, character recognition, fingerprint authentication. Architecture that can handle a different number of dimensions in the same hardware is required for a SoC with standardization potential. A bit slice of the developed DEC (Dimension Extension Circuit) for enabling dimension extension is shown in Fig. 4. With this DEC circuit additional component distances are added sequentially at the time of their input and stored in the lower-right DFF. In this way a flexible feature-vector dimensionality in the range 8~2042 dimensions is achieved.

### 4. Evaluation of fabricated prototype SoC

Fig. 4 shows the layout diagram of the developed KNN classifier designed in 180nm CMOS, as implemented in the recognition SoC. Corresponding specifications are listed in Table 2. The SoC design area is as small as  $3.51\text{mm}^2$  and the part that executes the k-nearest neighbor method (excluding the associative memory) is  $0.445\text{mm}^2$  i.e. only a 12% increased area over the implemented Euclidean-distance-based associative memory with 32 reference vectors. Thus the extension to the KNN algorithm is achieved very efficiently. Maximum measured operating frequency is 42.9MHz (at 1.8V), when all 8-bit vector components are used. Measured power consumption is as low as 5.02mW (at 42.9MHz, 1.8V). These results make the proposed KNN recognition solution very suitable for mobile devices [3, 4]. Fig. 5 shows the number of clocks required for human detection, when SFCC, CRR or CRR with DEC are applied as algorithms for the search by clock counting. Here, DEC is used to reduce the number of dimensions before clock counting, so that the number of clock cycles can be also further reduced.

### 5. Conclusion

The reported word-parallel digital KNN-based pattern-

matching architecture realizes scalable and error-free searching by mapping distances into a clock-cycle number. Area-efficient sequential partial-product addition is used to minimize implantation area for square calculation. For fast search a clock-number-minimization algorithm is applied. Experimental verification is done in 180nm CMOS, implementing 32 reference vectors with 8~2024 dimensions and 8 bit per dimension. A Dimension Extension Circuit (DEC) is used for flexibility of the dimension number. Maximum clock frequency is 42.9MHz (at 1.8V), when all 8-bit vector components are used. Power consumption is as low as 5.02mW (at 42.9MHz, 1.8V), so that the proposed VLSI architecture of KNN-based recognition becomes very suitable for implementation in mobile devices. To our best knowledge efficient VLSI architecture for integration of the KNN recognition algorithm in mobile devices has not been proposed previously. Additionally, because the developed SoC is a fully digital circuit, it is possible to further increase performance and to reduce area consumption by scaling to advanced CMOS processes.

### Acknowledgment

VLSI-chip design was supported by VDEC, the University of Tokyo in collaboration with Semiconductor Technology Academic Research Center (STARC), Rohm CO., LTD., Cadence Designs Systems Inc., Synopsys Inc., Mentor Graphics Co

### References

- [1]D. R. Tvetter, "The Pattern Recognition Basis of Artificial Intelligence," IEEE Press Piscataway,NJ,1997.
- [2]A. Ahmadi et al., Expert Sys. Appl., Vol. 38, No. 4, pp.3499-3513, 2011.
- [3]S. Sasaki et al., "Digital Associative Memory for Word-Parallel Manhattan-Distance-Based Vector Quantization" ESSCIRC'2012, pp.185-188, 2012
- [4]M.A. Abedin et al., "Realization of K-Nearest-Matches Search Capability in Fully-Parallel Associative Memories", IEICE Trans. on Fundamentals, vol. E90-A, No. 6, pp. 1240-1243, 2007

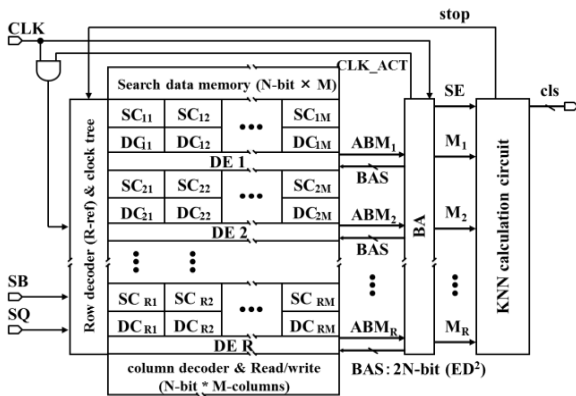


Figure 1: Overall architecture of KNN SoC.

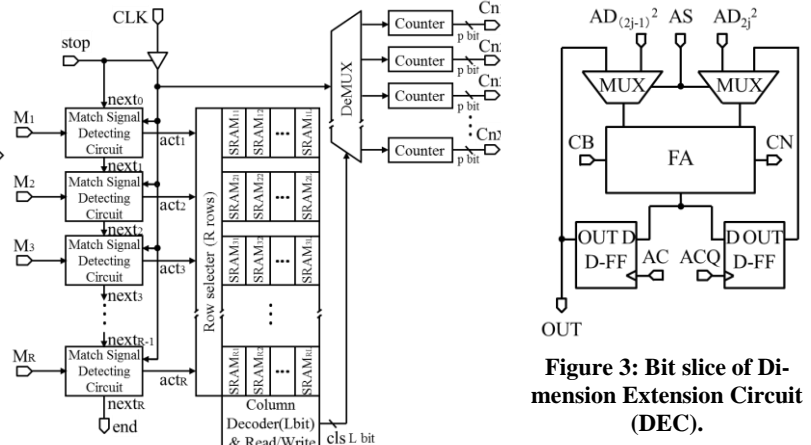


Figure 2: Clustering circuit with large flexibility for implementation of different application.

Figure 3: Bit slice of Dimension Extension Circuit (DEC).

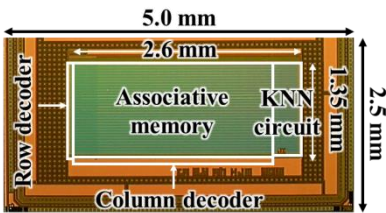


Figure 4: Die photo of the Recognition SoC implementing KNN in 180nm CMOS.

Table 1: Specifications of the recognition SoC.

Technology	180 nm CMOS
Supply voltage	1.8V
Function	1NN search, KNN
Organization	32-word, 8~2048-unit, 8-bit
Area	$3.51\text{mm}^2$
Max operation speed	42.9MHz(@1.8V)
Power consumption	5.02mW(@42.9MHz, 1.8V)

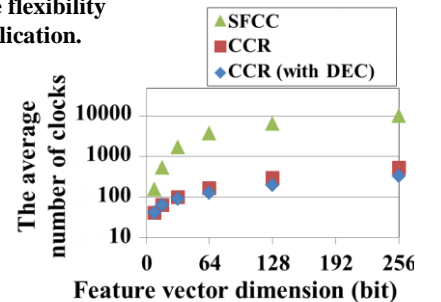


Figure 5: Number of clocks required for human detection.