Enhanced Vote Count Circuit based on NOR Flash Memory for Fast Similarity Search

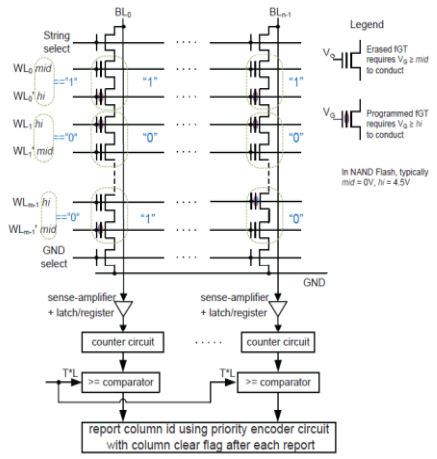
Haiyan Shu, Wenyu Jiang, Xiaoming Bao, Huan Zhou, and Rongshan Yu
Institute for Infocomm Research, A*STAR, Singapore

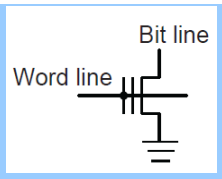
Vote Count Circuits

- **Vote Count concept**
  - Both reference vectors (v_j for j-th vector) and the query vector (q) are converted into *discrete* vectors by some deterministic hashing;
    - let $f_i(v)$ denote a vector v's value at hashed dimension i
    - let c_j denote a counter associated with vector v_j, initialize to 0

for (i = 0; i < L; i++)
$\forall v_j, if\ (f_i(v_j) == f_i(q))\ c_j++;$
$\forall\ v_j$, return v_j as candidates $if\ c_j \geq T;$

- **Vote Count Circuits**
  - Exploiting the fact that when reading a word on a row, its corresponding word-line (WL) will activate all cells on that row; by having a sense-amp at each column or bit-line (BL), the entire row is read out in 1 access cycle;
- **Enhanced Vote Count (EVC)**
  - Comparing m values instead of 1 (bet. query and reference) at a time;
  - This simultaneous m-value comparison made possible by using the *interlocked* design, where 2 cells are used to represent 1 pattern bit. First proposed for NAND Flash, now also proposed for NOR Flash.





NOR flash memory

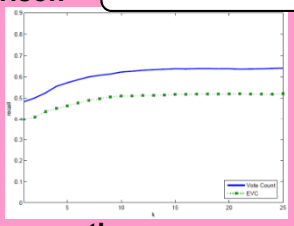
Vote Count Chip and Board

Prototype chip

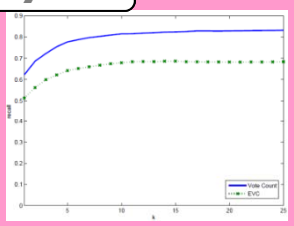
- Customized 0.18um NOR Flash circuit, 1024 word-lines and 1024 bit-lines which can accommodate 1024 reference vectors to be retrieved simultaneously
- Each hashing feature vector may have up to 512 dimensions.
- Retained the mux that shares each sense-amplifier with 32 columns.
- A test PCB is designed to mount the EVC prototype chip, and integrate with an FPGA development board

Performance Analysis

VC and EVC comparison
***L*=512, *m*=8**



*k*NN



*k*NN w. reranking

Speed and power consumption

	DRAM implementation	NAND implementation	NOR implementation
original vote count circuit	$L \cdot \tau_{par}$	$L \cdot \tau_{ser}$	$L \cdot \tau_{par}$
EVC circuit		$\frac{L}{m} \cdot \tau_{ser}$	$\frac{L}{m} \cdot \tau_{par}$

τ_{par}: 25 − 50ns τ_{ser}: 5 − 50μs L: # of hashed dimensions m: # of values simultaneously compared in EVC

Competitive Analysis with Commercial Software and Open Source Toolkits

Performance Metrics (1 Million DB items, find top-100 matches)	EVC		Open Source Toolkits
	low-density	high-density	Multi-index Hashing
Search Speed	$0.139ms$	$6.5\mu s$	$3ms$
Energy per Search	$0.75mJ$	$0.03mJ$	$150mJ$
Physical Space	$50cm^2 \cdot 1cm$	$1-2cm^2 \cdot 1cm$	desktop PC

ICASSP 2016, Mar. 20-25, Shanghai, China