RRAM

資訊工程所

111062510

張又仁

Outline

- Basic introduction
- Why RRAM ?
- Structure of RRAM
- How does RRAM work ?
- Application
- Why RRAM is not in commercial use now ?
- Conclusion

Basic introduction

- Resistive random-access memory (RRAM) is a new type of non-volatile memory.
- Work by changing the resistance across a dielectric solid-state material, often referred to as a memristor.

Why RRAM ?

- As devices are scaling down, DRAM, SRAM and Flash memory are facing the physical limit, which is atrributed to the loss of stored charge at nanoscale and leads to degradation of the performance, reliability, and noise margin.
- Requirements of large refresh dynamic power for DRAM and leakage power for both SRAM and DRAM pose serious challenges for the design.

Why RRAM ?

- **Requirements of new memory**
  - low operating voltage (< 1V)
  - long cycling endurance (> 10^17 cycles)
  - enhanced data retention time (>10 years)
  - low energy consumption (fJ /bit, 1fJ = 10^(-15)J)
  - superior scalability (< 10 nm)

Why RRAM ?

Table 1 Comparison of emerging memory technologies

Memory technology	SRAM	DRAM	NAND Flash	NOR Flash	PCM	STT-MRAM	RRAM
Cell area	$> 100F^2$	$6F^2$	$< 4F^2$(3D)	$10F^2$	$4–20F^2$	$6–20F^2$	$< 4F^2$(3D)
Cell element	6T	1T1C	1T	1T	1T(D)1R	1(2)T1R	1T(D)1R
Voltage	<1 V	<1 V	<10 V	<10 V	<3 V	<2 V	< 3 V
Read time	~1 ns	~10 ns	~ 10 μs	~50 ns	<10 ns	<10 ns	< 10 ns
Write time	~1 ns	~10 ns	100μs–1 ms	10 μs–1 ms	~50 ns	<5 ns	< 10 ns
Write energy (J/bit)	~fJ	~10 fJ	~10 fJ	100 pJ	~10 pJ	~0.1 pJ	~0.1 pJ
Retention	N/A	~64 ms	>10 y	>10 y	>10 y	>10 y	> 10 y
Endurance	$> 10^{16}$	$> 10^{16}$	$> 10^4$	$> 10^5$	$> 10^9$	$> 10^{15}$	$\sim 10^6 – 10^{12}$
Multibit capacity	No	No	Yes	Yes	Yes	Yes	Yes
Non-volatility	No	No	Yes	Yes	Yes	Yes	Yes
Scalability	Yes	Yes	Yes	Yes	Yes	Yes	Yes

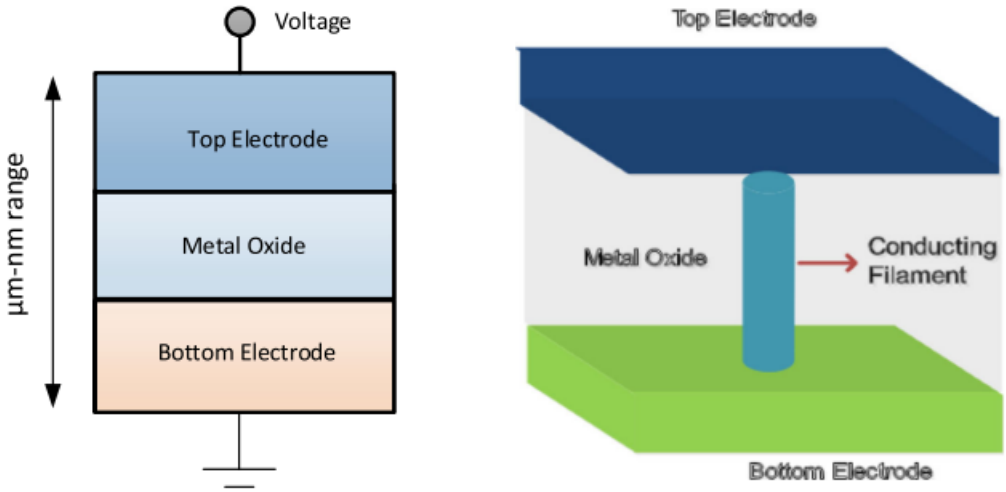
F: Feature size of lithography

RRAM

- **RRAM**
  - Fast read and write spead
  - operate under low voltage and current
  - In-memory computing

Structure of RRAM

- **Metal-Insulator-Metal Structure**
- **The mechanism called Filament Theory.**
  - Dielectric, which is normally insulating, can be made to conduct through a filament or conduction path formed after application of a sufficiently high voltage.
  - Filament or conduction path are arised vacancy or metal defect migration



Structure of RRAM

- They key material is the Insulator, which stores the data.
  - Oxide
  - ions
- In oder to store the Oxygen escape from oxide, the metal parts are plated a layer of Ti/Ta.
- Since the structure and the mechanism do not refer to electron. RRAM has radioresistance, which means it can work under radiation.

Structure of RRAM

- **Typically, an RRAM cell and a transistor form a structure called 1T1R.**
  - The transistor control the RRAM. It set, reset and read the data.
  - It is more easier to integrate RRAM into modern semiconductor fabrication.
  - The RRAM builds above transistors, which means that RRAM will not occupy the precious area on a die.

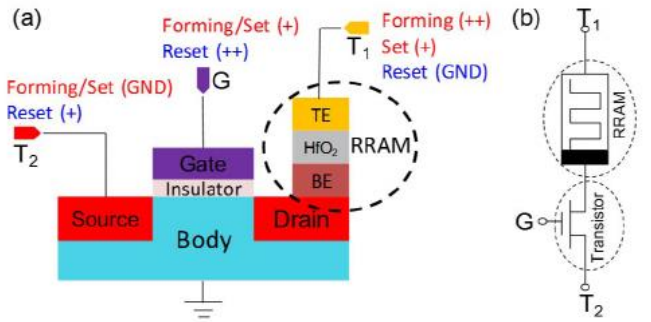


Fig. 1. (a) Schematic of 1T1R RRAM structure and operation conditions during DC sweeping. The RRAM is integrated on the drain-side of an NMOS transistor. Different voltage biases are applied to each terminal during electroforming, set and reset operations. (b) The equivalent circuit of a single 1T1R structure. The top electrode of the RRAM, the source and the gate of the transistor are defined as terminal T_1, T_2, and G, respectively.

Structure of RRAM

- **Bipolar RRAM**
  - The set and reset voltage are opposite.
    - Apply positive voltage to set.
    - Apply negative voltage to reset.
- **Unipolar RRAM**
  - Set or reset voltage are both positive.
  - However, this type of RRAM are not as stable as bipolar one.

Structure of RRAM

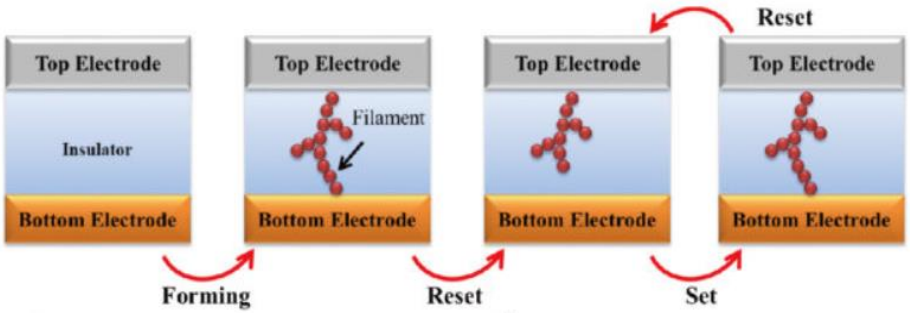
Table 2 Comparison of various RRAM types (*Continued*)

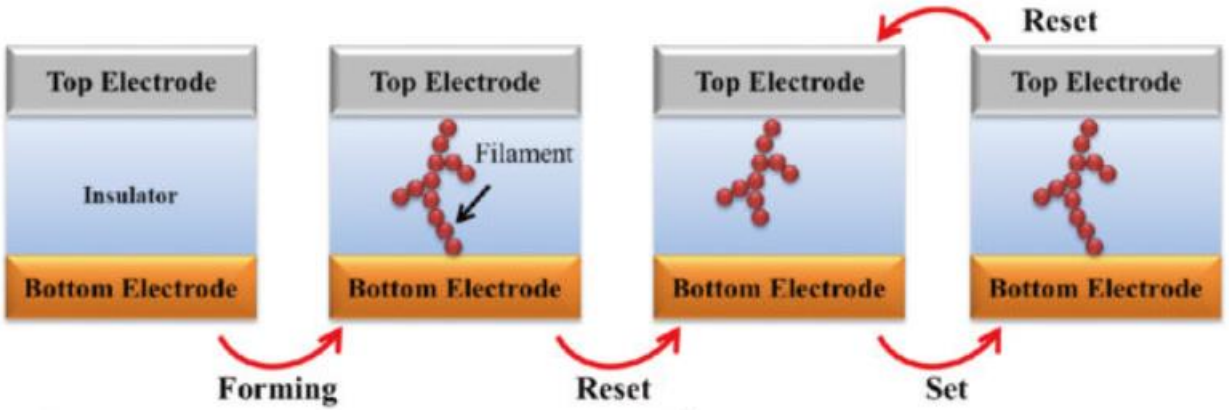
Ref	Year	Top electrode	Oxide material	Bottom electrode	Operation mode	HRS/LRS ratio	Retention	Endurance	V_f	V_{set}	V_{reset}	I_{cc}
			Ag NPs									
[26]	2016	Ti/W	TiO_x/MgO	Ru	Bipolar	< 32	NS	> 10^3 write cycles > 10^9 read cycles	FF	1.4 V	– 1.8 V	8mA
[28]	2016	TiN/Ti	TiO_{2-x}	Au	Bipolar	~ 10^5	NS	10^5 cycles	FF	1 V	– 1 V	2-200μA
[41]	2016	Ag	a-ZnO	Pt	Bipolar	> 10^7	> 10^6s	> 10^2 cycles	FF	0.24 V	– 2 V	0.1-0.5 mA
[87]	2016	W	WO_3/Al_2O_3	TiW/Cu	Bipolar	10^4	NS	~ 300 cycles	3 V	~ 3.5 V	~– 2.5 V	10μA
[27]	2017	TiN/Ti	TiO_{2-x}	Au	Bipolar	< 10^3	> 10^5s	> 50 cycles	FF	<– 0.5 V	<1 V	1-200nA
[88]	2017	Al	HfO_x	Al	Unipolar	~ 10^4	NS	NS	~1V	1.8 V	0.8 V	1 μA-1mA
[89]	2017	Au/Ti	TiO_{2-x}	Au	Bipolar	> 10^4	NS	NS	FF	1 V	– 1 V	5 μA-1mA
[18]	2018	Ti	HfO_2	TiN	Bipolar	>10	10^4s	> 10^7 cycles	FF	0.5 V	– 0.5 V	NS
[31]	2018	ITO	a-TiO_2	Pt	Bipolar	>10	< 10^3s	NS	4.15 V	0.6 V	– 0.5 V	< 200 μA
[47]	2018	ITO	Zn_2TiO_4	Pt	Bipolar	NS	> 10^4s	> 500 cycles	25 V	0.6 V	– 0.6 V	1-10 mA
[72]	2018	Ag	MnO/Ta_2O_5	Pt	Bipolar	10^6	~ 10^4s	100 cycles	FF	0.8 V	– 1.1 V	1 mA
[90]	2018	Pd	HfO_x/ Ag NPs	TiN	Bipolar	10^3	> 10^4s	> 10^8 cycles	5.25 V	2.2 V	– 2.2 V	100 μA

V_f forming voltage, V_{set} set voltage, V_{reset} reset voltage, I_{cc} compliance current, *NS* not specified, *FF* forming free

How does RRAM work ?

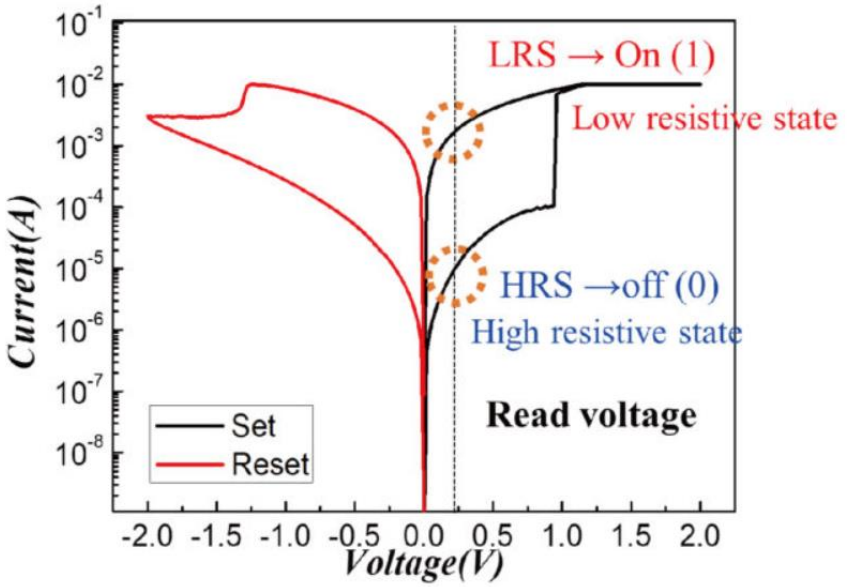
- **Forming: Apply a large current, about 100 µA to create the filament in oxide.**
- **Set : Apply positive voltage so that the oxide turns into filament. Here the RRAM is in Low Resistance State (LRS)**
- **Reset: Apply the opposite and higher voltage so that some of the filament turn into oxide, which is high resistance material. Here the RRAM is in High Resistance State (HRS)**




Reset
Top Electrode
Top Electrode
Top Electrode
Top Electrode
Filament
Insulator
Bottom Electrode
Bottom Electrode
Bottom Electrode
Bottom Electrode
Forming
Reset
Set

How does RRAM work ?

- **Read the data**
  - Apply a low voltage on the RRAM. By measuring the current, we can calculate the resistance and get the data.
  - HRS represent 0
  - LRS represent 1



Application

- **In-Memory Computing**
  - Modern computer is based on Havard architecture or von Neumann architecture.
    - The computing and storage are seperated.
    - The data bandwidth will be the bottleneck of the performance, so-called von Neumann bottleneck.
  - In-Memory Computing is non von Neumann architecture.
    - Save time, consume lower power and high efficiency.

Application

In-Memory Computing

- In-Memory Computing is consider to accelerate matrix operation.
  - Accelerate the AI computation.
  - Up to 10x ~ 1000x speed up.

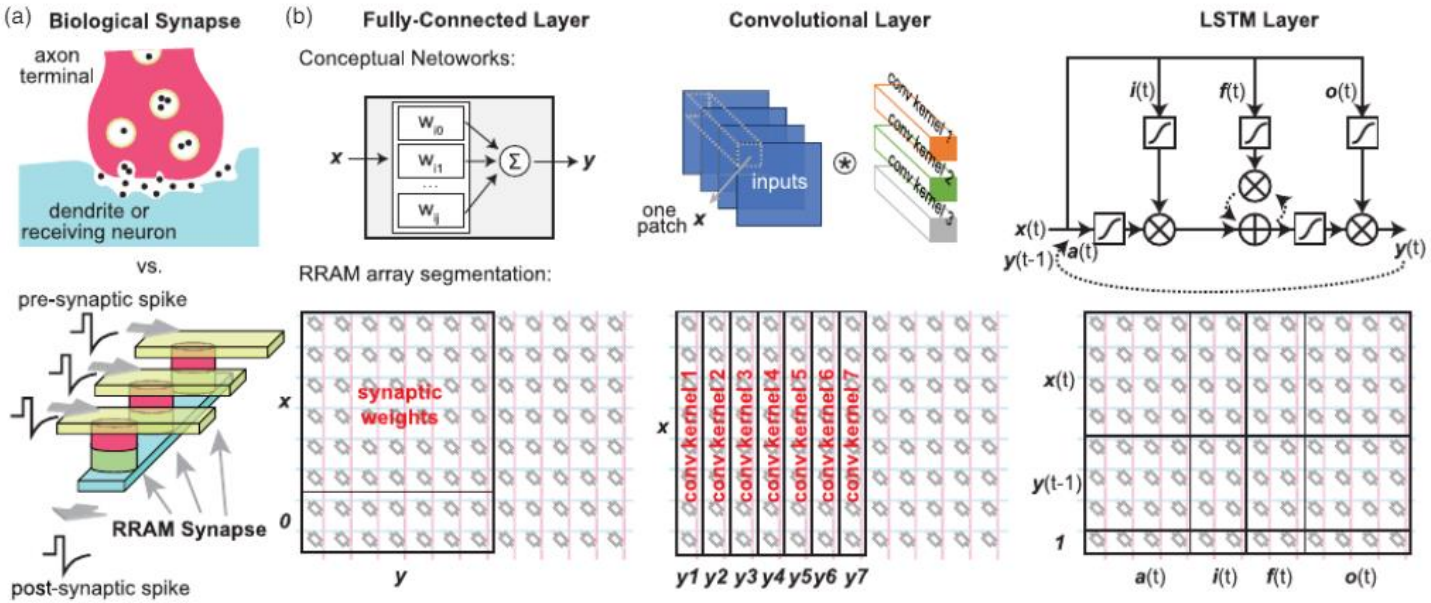


Figure 2. a) Biological synapse versus RRAM synapse.[5] b) Different segmentation schemes for fully connected layers (DNN), convolutional layers (CNN), and LSTM layers (LSTM network), all of which are converted into VMM using the RRAM array.

Application

Synapse operation

- Human brains use analog signal to send information or store information, while computer use digital signals.
- Multi-level Resistance Characteristics of RRAM makes it works more like human brains.

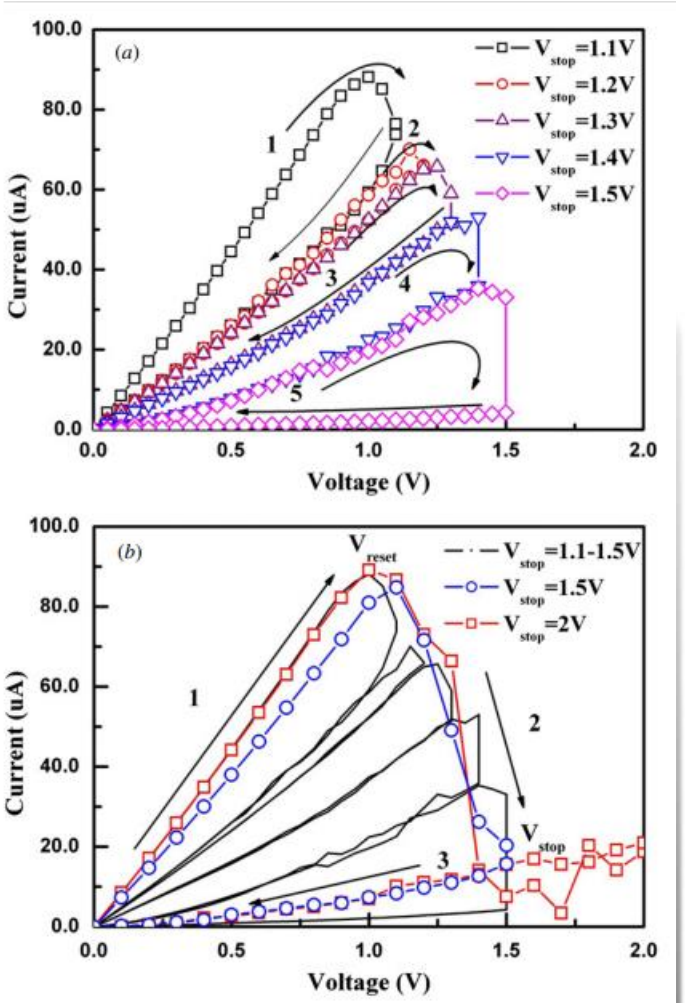
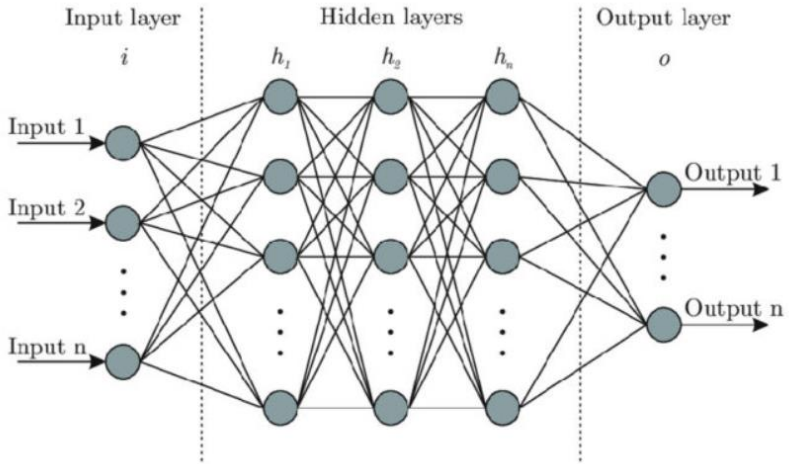


Figure 5. (*a*) Multilevel reset switching behaviors by controlling the V_{stop} ranging from 1.1 to 1.5 V. (*b*) Reset process behavior with the sweeping V_{stop} up to 2 V.

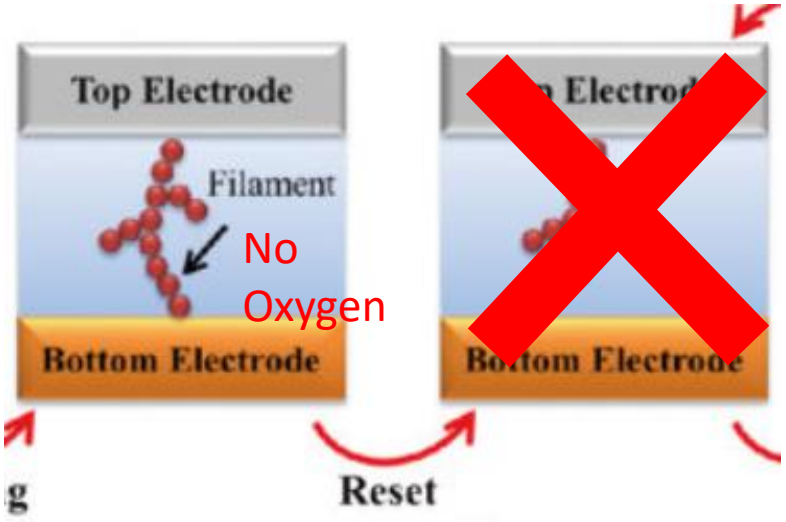
Application

- **Synapse operation**
  - The ANN on RRAM will works more like human brain, which shows great potential on works that human brains are good at, such as classification, speech recognition, graphic recognition etc.



Why RRAM is not in commercial use now ?

- **Endurance is not as good as the research.**
  - The average lifetime of RRAM is about 1000 times set/reset
  - The Oxygen in oxide will escape or go round, which means that oxidation reaction, the reset action, fails and therefore the RRAM stays at LRS.



Why RRAM is not in commercial use now ?

- **Power efficiency is also not good.**
  - Forming, reset require large current.
  - If using smaller current, the data are more likely to loss.
    - Reset : The oxide sometimes turn into filament due to the unstable Oxygen. Therefore the data 0 will become 1.
    - Set : The filaments sometimes turn into oxide. Therefore, the data 1 will become 0.
  - If using larger current (100uA), the filament or oxide structure are stabler.
  - However, the power consumption grows.

Why RRAM is not in commercial use now ?

- **The resistance is not stable.**
  - The filament grows randomly in every set.
  - Resistance LRS will be vary for each set.
    - e.g. 1kΩ this time and 10kΩ next time.
    - The resistance is still lower than HRS.
  - It is difficult to use it in analog representation.

Why RRAM is not in commercial use now ?

- **Speed is not as fast as DRAM.**
  - It takes time to set and reset.
  - RRAM is faster than NOR flash but slower than DRAM in reality.
  - According to the draw back described, RRAM cannot replace DRAM.
    - DRAM is cheaper and the endurance is longer.

Conclusion

- **RRAM is an type of emerging memory.**
- **It is non-volatile and can perform in-memory computing.**
- **It stores data by changing the resistance.**
- **Metal-Insulator-Metal Structure.**
- **Why not in commercial use ?**
  - Low endurance
  - Low energy efficiency
  - Not stable

Reference

[1] https://technews.tw/2022/04/25/ma-tek-popular-science-rram/

[2] https://technews.tw/2022/05/04/ma-tek-popular-science-rram-2/

[3] Wang, Zhuo-Rui, et al. “Functionally complete Boolean logic in 1T1R resistive random access memory." IEEE Electron Device Letters 38.2 (2017): 179-182.

[4] Yan, Bonan, et al. “Resistive Memory-Based In-Memory Computing: From Device and Large-Scale Integration System Perspectives." Advanced Intelligent Systems 1.7 (2019): 1900068.

Reference

[5] Furqan Zahoor, Tun Zainal Azni Zulkifli, et al. “Resistive Random Access Memory (RRAM): an Overview of Materials, Switching Mechanism, Performance, Multilevel Cell (mlc) Storage, Modeling, and Applications.” Nanoscale Reaserch Letters.

[6] 鍾裕隆、陳貞夙 , “離開歐姆定律─電阻式記憶體材料”, 科學發展 2013 年 6 月 486 期 34-38

[7] Ming-Chi Wu, Wen-Yueh Jang, et al. “A study on low-power, nanosecond operation and multilevel bipolar resistance switching in Ti/ZrO2/Pt nonvolatile memory with 1T1R architecture”, Semiconductor Science and Technology, 2012.

[8] Interview with Wen-Yueh Jang, PSMC