

快擦写存储器结构及机理*

常春喜 王保恒

(国防科技大学计算机系 长沙 410073)

摘要 讨论了快擦写存储器典型位元结构及其存储信息原理,详细论述了目前广泛使用的或非与非两种结构的快擦写存储矩阵的结构组成和工作原理。

关键词 存储位元, 存储信息原理, 或非存储矩阵, 与非存储矩阵

分类号 TP333

Structures and Mechanisms of Flash Memory

Chang Chunxi Wang Baoheng

(Department of Computer, NUDT, Changsha, 410073)

Abstract The paper discusses the typical cell structure of flash memory and its principle of information storage. In addition, the organizations and working mechanisms of the NOR and NAND flash memory matrices, which are widely applied at present, are detailed.

Key words memory cell, information storage principle, NOR memory matrix, NAND memory matrix.

快擦写(Flash)存储器亦称快闪存储器或闪速存储器,1983年由Intel公司首先推出,商品化于1988年。Flash存储器技术是在两种非挥发可改写EPROM和EEPROM的基础上发展起来的。它集中了两者的优点,既具有像EPROM一样的单管位元结构,又沿用了传统EPROM热电子隧道效应的编程机制,并具有EEPROM在线、冷电子隧道效应的擦除机制。位元尺寸比EEPROM小10倍。它是目前唯一具有大存储容量、非挥发性、低价格、可在线改写和较高速度等几个特性的存储器。

1 Flash存储器的位元结构及存储信息原理

最简单的位元结构是与EPROM类似的单管叠栅结构。还有两种单管三栅结构,一种是在单管叠栅结构的基础上增加一个专用擦除栅EG,另一种则在单管叠栅结构基础上再增加一个控制栅CG,它位于浮栅FG之下,称为第二CG。除单管结构外还有双管结构、三极管和二极管结构的存储位元。采用多管结构的目的是改善Flash存储器的编程、擦除性能,克服单管叠栅过擦除问题和降低编程擦除电压。但这些结构却使位元结构变得复杂,尺寸加大,制造工艺繁琐,增加了芯片体积和价格。下面只介绍目前广泛使用的单管叠栅结构存储位元。

单管叠栅结构位元依编程擦除方法不同,有多种结构。不论哪一种结构,当FG上驻留了电子,便有附加电场存在,使开启阈值电压高($> 5V$),表示存储“0”信息。当FG上无电子,开启阈值电压低($< 2V$),表示存储“1”信息。编程操作,就是使FG上充上电子,即写“0”。擦除操作,就是放掉FG上的电子,即写“1”(因工厂制作好的位元FG上均无驻留电子,即全为“1”信息)。读出则是在CG上加高阈值电压与低阈值电压的平均值。位元开启通导,读出信息“1”,否则读出信息“0”。

* 1998年3月4日收稿

第一作者:常春喜,男,1966年生,工程师

经典单管叠栅位元结构见图 1 (a), 与 EPROM 的 SIMOS (stacked-gate Injection MOS) 管结构很类似, 故称这种管为 ETOX (EPROM Tunnel OXide) MOS 管。其结构特点是 CG、FG 间氧化层厚度 25nm 左右, FG 与 P-Si 表面间的超薄隧道氧化层质量很高, 能经受多次高压冲击且厚度仅在 10nm 左右, 沟道长度 (栅长) 也很短, 为亚微米级。如片容量 16Mb 的 flash 存储芯片, 其位元 CG、FG 间氧化层厚度 20nm, FG、P-Si 超薄隧道氧化层 10nm 沟道长度 0.8 μm 。

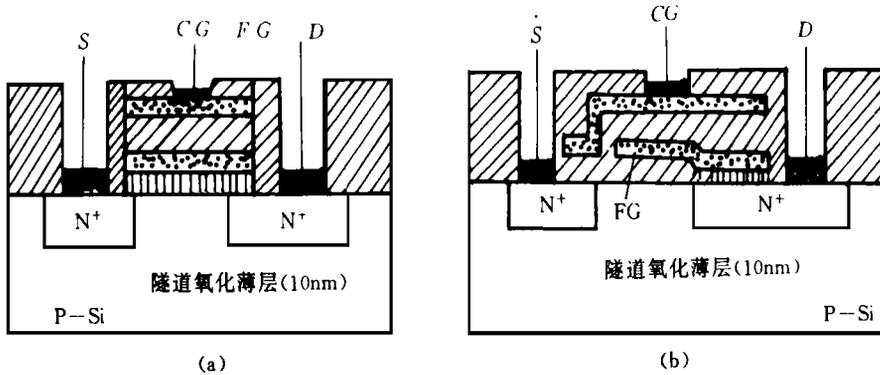


图 1 Flash 存储器单管叠栅位元结构

编程: 即写“0”, S 极接地、CG、D 极接高压 (12~20V), 沟道导电。电子在高漏压作用下, 获得足够的能量, 变成热电子。在 CG 强正电场作用下, 部分电子通过沟道时, 穿过超薄隧道氧化层进入 FG, 称这种物理现象为热电子隧道效应。去掉所加电压, FG 上电子仍然存在, 从而产生一个附加电场, 使开启阈值电压变高 (7V), 变成存“0”位元。

擦除: 即写“1”, S 极加高压 (12~20V), CG 接地, D 极接地或浮空, S 与 FG 极间发生 F-N 隧道效应。FG 上电子被拉到源区, 开启阈值电压变低 (2V 以下), 成为存“1”位元。在强正电场作用下, 称自由电子在源 (或漏) 与浮栅 FG 之间双向穿越超薄隧道氧化层的物理现象, 为 Fowler-Nordheim (F-N) 隧道效应。由于电子穿越超薄隧道氧化层时没有沟道, 电子并未得到足够的能量成为热电子, 所以又称该效应为冷电子隧道效应。

S 极加高压擦除是利用 FG 与源区构成的小电容、分压大、场强高的原理实现的。小电容是由 FG 与源区侧向扩散区面积构成的。因侧向扩散区相对于 FG 的面积很小, 所以电容也很小。之所以利用源极擦除是因为存储矩阵或部分存储矩阵位元的源极都是连接在一起的, 这样可实现整个芯片或分块快速擦除。如果存储位元矩阵或部分存储矩阵的源极不是连接在一起的, 则高压不是加到源极, 而是加到部分位元或全部位元的公共区衬底。如下面将谈到的 NAND 结构存储矩阵就属此种情况, 这样, 也可以实现整个芯片或分块快速擦除。

图 1 (b) 是另一种单管叠栅结构存储位元。其超薄隧道氧化层仅位于漏区上方, 厚度仅为 10nm 左右, FG 不跨越整个沟道。这种结构编程和擦除都是通过漏区超薄隧道氧化层的 F-N 隧道效应。加之 FG 不跨越整个沟道, 使得编程和擦除效率都很高。

编程: CG 加高压 (12~20V), D 极接地, S 极浮空。漏区电子通过 F-N 隧道效应进入 FG, 写入“0”, 实现编程。

擦除: CG 接地, D 极接高压 (12~20V), S 极浮空。FG 电子通过 F-N 隧道效应返回到漏区, 写入“1”, 实现擦除。

单管结构存储位元存在过擦除问题。flash 存储器是整个芯片或分块擦除的, 要擦除的位元有存“0”的, 也有存“1”的, 存“1”位元已被擦除过。擦除时, 对存“0”位元是必须的, 对存“1”位元则是不必要的, 这就造成存“1”位元的过擦除, 甚至使 FG 带正电荷, 使存“1”位元的开启阈值电压很低, 甚至是负电压。为避免这种现象, 在擦除之前, 往往都对所有位元进行一次编程, 即都写入“0”, 然后再擦除, 使编程前所有位元有相同阈值的开启电压。

2 flash 存储器存储矩阵结构及工作原理

flash 存储器存储矩阵结构有“或”阵列和“与”阵列两大类。前者存储位元并接, 呈“或”关系, 包括“或”(OR)和“或非”(NOR)两种。后者存储位元是串接, 呈“与”关系, 包括“与”(AND)和“与非”(NAND)两种。下面以单管叠栅位元为例, 介绍最常见的 NOR 矩阵和 NAND 矩阵结构特点, 并说明其编程、擦除和读出工作原理。

(1) 或非存储矩阵及工作原理

图 2 给出 2 字 \times 8 位 flash 存储器 NOR 结构存储矩阵示意图。WL₀、WL₁ 为字线, DL_i($i = 0, 1, \dots$,

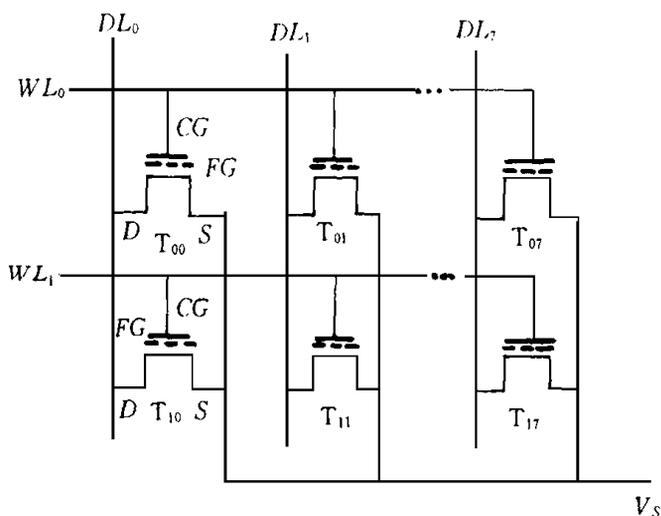


图 2 Flash 存储器 NOR 存储结构示意图

7) 为位线, 读出时接读出放大器 RA, 擦除时浮空, 编程时写“0”的位接高电压+12V, 写“1”的位浮空或接地。

擦除: WL₀、WL₁ 接地, Vs 加+12V, DL₀ ~ DL₇ 浮空。所有位元都发生 F-N 隧道效应, FG 上电子被拉回源区, 即都被写成“1”。

编程: 假定编 0 号单元, 且写入 10001011B 信息。Vs 接地, WL₀ 加+12V, DL₀、DL₄、DL₆ 和 DL₇ 浮空, DL₁、DL₂、DL₃ 和 DL₅ 接+12V, WL₁ 接 0V。存储位元 T₀₁、T₀₂、T₀₃ 和 T₀₅ 产生热电子隧道效应, FG 充上电子, 变为高开启阈值, 写入“0”。存储位元 T₀₀、T₀₄、T₀₆ 和 T₀₇ 不发生热电子隧道效应, 保持擦除时写入的“1”信息。1 号单元因 WL₁ 加 0V 保持原存信息不变。

读出: 此时 Vs 接地, DL₀ ~ DL₇ 接 RA, 电压约为 2V, 假定读 0 号单元, WL₁ 不被选, 接地, 1 号单元的所有位元与 RA 隔离。WL₀ 被选中接+3V, 存“1”位元 T₀₀、T₀₄、T₀₆、T₀₇ 导通, 相应 RA 输出“1”; 存“0”位元 T₀₁、T₀₂、T₀₃ 和 T₀₅ 不导通, 相应 RA 读出“0”, 读出信息 10001011B。应当指出, 对于高集成度大容量 flash 存储芯片, 因其采用了超短沟道工艺, 编程时写“0”位的 DL_i($i = 0, 1, \dots, 7$) 端不是加+12V, 而是加+6V 或更低电压。

(2) 与非 (NAND) 存储矩阵及工作原理

flash 存储器与非结构 8 行、8 列存储矩阵示意如图 3。图中存储位元源、漏串接构成存储矩阵的列。各列同位置处位元的控制栅极 CG 并接构成存储矩阵的行。最上面一行位元的漏极与各自的位线连接, 最下面一行的源极并连接到 Vs。

擦除: WL₀ ~ WL₇ 接 0V, 衬底加高电压+20V, Vs、DL₀ ~ DL₇ 浮空。FG 上电子通过 F-N 隧道效应进入衬底, 实现擦除, 写入“1”。

编程: NAND 存储矩阵不能随机编程, 只能按地址顺序编程。即从 WL₀ 单元(行)开始, 接 WL₀, WL₁, WL₂, ..., WL₇ 顺序编程。编程时, 选中行加高电压+20V, 其它行加+10V, 衬底接地, Vs 浮空,

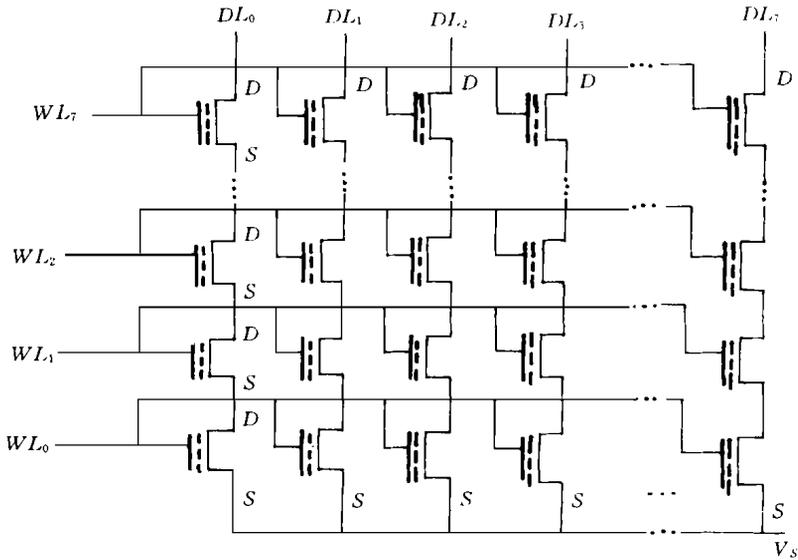


图3 Flash存储器 NAND 存储结构示意图

写“0”和写“1”位元所连位线分别接0V和+10V。这将导致选中行的写“0”位元D极为0V，CG为+20V，产生F-N隧道效应，使FG充上电子，写入“0”。写“1”位元D极为+10V，CG为+20V，不能产生F-N隧道效应，FG未充上电子，维持原擦除时的“1”。假如编程 WL_0 单元，写入信息为11000011B。选中行 WL_0 加+20V，未选中行 $WL_1 \sim WL_7$ 为+10V， DL_0, DL_1, DL_6 和 DL_7 加0V，其它位线加+10V，衬底接地， V_s 浮空，即完成该单元的编程。未选中行因CG均加+10V电压，与选中行写“0”位元连接的位元，其D极为0V，与选中行写“1”位元连接的位元，其D极接+10V，它们都不会发生F-N隧道效应，故保持原存储信息不变。

读出：位线与RA连接， V_s 接地，未选中行加+10V，选中行加+3V。未选中行存储单元的各位元，不论存“1”还是存“0”都通导，不影响被选单元的读出。被选中存储单元的存“1”位元通导，相应RA工作，输出“1”；存“0”位元不通导，相应RA不工作，输出“0”。读出时，可顺序进行，也可随机实现。

构成NOR存储矩阵位元的面积约为NAND矩阵位元面积的3倍，所以NAND结构更紧凑，集成度更高。NOR结构编程、读出都可随机进行，NAND结构编程必须按顺序进行，限制了它在某些方面的使用，但在文件存储系统中常采用页方式进行写入和读出，用NAND结构的flash比NOR结构的Flash有更快的速度。同样编程速度时，NAND编程和擦除电压要比NOR高。

3 结束语

世界上有20多个公司竞相生产和研制Flash，竞争非常激烈，都在大容量、低成本、高速和使用灵活方便上下功夫。Flash是理想的文件存储介质，用它代替硬盘，除速度快、体积小、重量轻、功耗低和不易损坏外，还可就地执行EIP (Execute In Place)。即无须像硬盘那样，先将OS、实用程序和数据文件调入主存，而是可立即执行程序。在日益繁荣的膝上型、掌上型和笔记本型个人计算机中，取代磁盘作为外存，得到越来越广泛的应用。

参考文献

- 1 Betty Price. Semiconductor Memories. 2nd ed. S. L. Wiley, 1996
- 2 Kifung Cheung. 一种故障时性能缓慢下降降级使用的交叉存储系统. 计算机工程与科学, 1988, 10 (3)
- 3 张德清. 一个位独立交叉访问存储器的方案设计. 计算机工程与科学, 1995, 17 (2)