



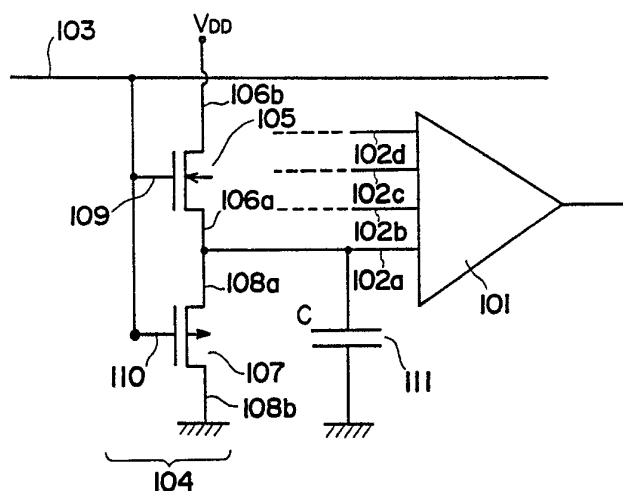
PCT

特許協力条約に基づいて公開された国際出願

(51) 国際特許分類 5 G06G 7/60	A1	(11) 国際公開番号 WO 92/12498 (43) 国際公開日 1992年7月23日 (23. 07. 1992)
(21) 国際出願番号 PCT/JP92/00014 (22) 国際出願日 1992年1月10日 (10. 01. 92) (30) 優先権データ 特願平3/13780 1991年1月12日 (12. 01. 91) JP (71) 出願人; および (72) 発明者 柴田 直 (SHIBATA, Tadashi) [JP/JP] 〒980 宮城県仙台市太白区日本平5番2号 Miyagi, (JP) (72) 発明者; および (75) 発明者/出願人 (米国についてのみ) 大見忠弘 (OHMI, Tadahiro) [JP/JP] 〒980 宮城県仙台市青葉区米ヶ袋2丁目1番17号301 Miyagi, (JP) (74) 代理人 弁理士 福森久夫 (FUKUMORI, Hisao) 〒160 東京都新宿区本塩町12 Tokyo, (JP) (81) 指定国 AT (欧州特許), BE (欧州特許), CH (欧州特許), DE (欧州特許), DK (欧州特許), ES (欧州特許), FR (欧州特許), GB (欧州特許), GR (欧州特許), IT (欧州特許), KR, LU (欧州特許), MC (欧州特許), NL (欧州特許), SE (欧州特許), US		添付公開書類 国際調査報告書

(54) Title: SEMICONDUCTOR DEVICE

(54) 発明の名称 半導体装置



(57) Abstract

A semiconductor device capable of realizing a neuro-computer chip of high integration degree and low power consumption. The device consumes a very small amount of electric power and the combinations of synapses can be realized, using a small number of elements. A first MOS type transistor includes a first gate electrode floating in potential and provided in a region isolating a source region from a drain region via a first insulation film, and plural second gate electrodes capacitively coupled to the first gate electrode via a second insulation film, one of which is in connection with a source electrode. The gate or drain electrode of the MOS type transistor is in connection with a first wiring for transferring signals of high or low level in potential.

(57) 要約

本発明は、消費電力が非常に小さく、かつ少数の素子でシナプス結合が実現でき、高集積度、低消費電力のニューロンコンピュータチップを実現することのできる半導体装置を提供することを目的とする。

ソース及びドレイン領域を隔てる領域に第1の絶縁膜を介して設けられた電位的にフローティング状態にある第1のゲート電極、第1のゲート電極と第2の絶縁膜を介して容量結合する複数の第2のゲート電極、第2のゲート電極の1つにソース電極が接続された第1のMOS型トランジスタ、MOS型トランジスタのゲート電極もしくはドレイン電極が、高レベルもしくは低レベルの2つの電位レベルの信号を伝達する第1の配線に接続されていることを特徴とする

情報としての用途のみ

PCTに基づいて公開される国際出願のパンフレット第1頁にPCT加盟国を同定するために使用されるコード

AT	オーストリア	ES	スペイン	ML	マリ
AU	オーストラリア	FI	フィンランド	MN	モンゴル
BB	バルバドス	FR	フランス	MR	モーリタニア
BE	ベルギー	GA	ガボン	MW	マラウイ
BF	ブルキナファソ	GN	ギニア	NL	オランダ
BG	ブルガリア	GB	イギリス	NO	ノルウェー
BJ	ベナン	GR	ギリシャ	PL	ポーランド
BR	ブラジル	HU	ハンガリー	RO	ルーマニア
CA	カナダ	IT	イタリア	RU	ロシア連邦
CF	中央アフリカ共和国	JP	日本	SD	スーダン
CG	コンゴ	KP	朝鮮民主主義人民共和国	SE	スウェーデン
CH	スイス	KR	大韓民国	SN	セネガル
CI	コートジボワール	LI	リヒテンシュタイン	SU	ソウイェト連邦
CM	カメルーン	LK	スリランカ	TD	チャード
CS	チェコスロバキア	LU	ルクセンブルク	TG	トーゴ
DE	ドイツ	MC	モナコ	US	米国
DK	デンマーク	MG	マダガスカル		

明細書

半導体装置

技術分野

本発明は、半導体装置に係わり、特に神経回路網コンピュータ（ニューロンコンピュータ）を実現するための高機能半導体集積回路装置を提供するものである。

背景技術

半導体は集積回路技術の進展は実に驚くべき速度で進んでおり、例えばダイナミック・メモリを例にとるなら、4メガビットから16メガビットがすでに量産体制にあり、64メガビット以上の容量をもった超々高密度メモリも研究レベルでは実現されつつある。64メガビットメモリは、せいぜい1cm四方のシリコンチップ上に実に約1億2000万個ものMOSトランジスタが集積されている。このような超高集積化技術はメモリ回路ばかりでなく論理回路にも応用され、32ビットから64ビットのCPUをはじめとする、様々な高機能論理集積回路が開発されている。

しかし、これらの論理回路はデジタル信号、すなわち「1」と「0」という2値の信号を用いて演算を行なう方式を採用しており、例えばコンピュータを構成する場合は、ノイマン方式といって、あらかじめ決められたプログラムに従って1つ1つ命令を実行していく方式である。このような方式では単純な数値計算に対しては非常に高速な演算が可能であるが、パターン認識や画像の処理といった演算には膨大な時間を要する。さらに、連想、学習といったいわば人間が最も得意とする情報処理に対しては非常に不得手であり、現在様々なソフトウェア技術の研究が行なわれているが、はかばかしい成果は得られていないのが現状である。

ここで、これらの困難を一挙に解決するため、生物の脳の機能を研究し、その機能を模倣した演算処理を行なえるコンピュータ、すなわち神経回路コンピュータ（ニューロンコンピュータ）を開発しようというまた別の流れの研究がある。

このような研究は、1940年代より始まっているが、ここ数年来非常に活発に研究が展開されるようになった。それはLSI技術の進歩にともない、このよ

うなニューロンコンピュータのハードウェア化が可能となったことによる。

しかしながら、現状の半導体LSI技術を用いてニューロンコンピュータをLSIチップ化するにはまだまだ様々な問題があり、実用化のメドはほとんどたっていないのが実情である。

LSI化における技術的な問題がどこにあるのかを以下に説明する。

人間の脳は極めて複雑な構造を有し、非常に高度な機能を有しているが、その基本的な構成は非常に単純である。すなわち、ニューロンと呼ばれる演算機能をもった神経細胞と、その演算結果を他のニューロンに伝える、いわば配線の役割を担った神経繊維とから構成されている。

この脳の基本単位の構成を簡略化してモデルで描いたのが図9である。901a, 901b, 901cはニューロンであり、902a, 902b, 902cは神経繊維である。903a, 903b, 903cはシナプス結合とよばれ、例えば神経繊維902aを伝わって来た信号に w_a という重みをかけ、ニューロン901aに入力する。ニューロン901aは入力された信号強度の線形和を取り、それらの合計値がある閾値をこえると神経細胞が活性化し、神経繊維902bに信号を出力する。合計値が閾値以下だとニューロンは信号を出力しない。合計値が閾値以上になって、ニューロンが信号を出すことを、そのニューロンが「発火した」と言う。

実際の脳では、これらの演算、信号の伝搬、重みのかけ算等すべて電気化学現象によって行われており、信号は電気信号として伝送・処理されている。人間が学習する過程は、シナプス結合における重みが増減していく過程としてとらえられている。すなわち、様々な入力信号の組合せに対し、正しい出力正しい出力が得られるよう重みが徐々に修正され、最終的に最適の値に落ち着くのである。つまり人間の英知はシナプスの重みとして脳に刻みつけられているのである。

数多くのニューロンがシナプスを介して相互に接続され1つの層を形成している。これらが人間の脳では、6層重ね合わされていることが分かっている。このような構造、機能を半導体デバイスを用いてLSIシステムとして実現することが、ニューロンコンピュータ実現の最も重要な課題である。

図10は、1つの神経細胞、すなわち1個のニューロンの機能を説明する図面

であり、1943年にMcCullockとPitts(Bull:Math. Biophys. Vol. 5, p. 115 (1943))により数学的モデルとして提案されたものである。現在もこのモデルを半導体回路で実現し、ニューロンコンピュータを構成する研究が盛んに進められている。 $V_1, V_2, V_3, \dots, V_n$ は、例えば電圧の大きさとして定義される n 個の入力信号であり、他のニューロンから伝達された信号に相当している。 $w_1, w_2, w_3, \dots, w_n$ はニューロン同士の結合の強さを表す係数で、生物学的にはシナプス結合と呼ばれるものである。ニューロンの機能は各入力 V_i に重み w_i ($i = 1 \sim n$) をかけて線形加算した値 Z が、ある所定の閾値 V_{TH}^* より大となったときに「1」を出力し、また閾値より小のときに「0」を出力するという動作である。これを数式で表せば、

$$Z \equiv \sum_{i=1}^n w_i V_i \quad \dots (1)$$

として、

$$V_{out} = 1 \quad (Z > V_{TH}^*) \quad \dots (2)$$

$$0 \quad (Z < V_{TH}^*) \quad \dots (3)$$

となる。

図10(b)は、 Z と V_{out} の関係を表したものであり、 Z が V_{TH}^* より十分大きいときは1、十分小さいときは0を出力している。

さて、このようなニューロンをトランジスタの組合せで実現しようと思えば、数多くのトランジスタを必要とするばかりか、加算演算を各信号を電流値に変換してこれを足し合わせるにより行うため、多くの電流が流れ多大のパワーを消費することになる。これでは高集積化は不可能である。この問題は、ニューロンMOSFET (ν MOSと略)の発明(発明者:柴田直、大見忠弘、特願平1-141463号)により解決された。

この発明はたった1つのトランジスタでニューロンの働きの主要機能を果たすことができ、しかも電圧信号をそのまま加算演算することができるため、ほとんど電力消費がないという画期的なものである。図11(a)は ν MOS断面構造の一例を簡略化して示したものであり、1101は例えばP型のシリコン基板、1102, 1103は N^+ 拡散層で形成されたソース及びドレイン、1104は

チャネル領域上に設けられたゲート絶縁膜（例えば SiO_2 など）、1106は電氣的に絶縁され電位的にフローティングの状態にあるフローティングゲート、1107は例えば SiO_2 等の絶縁膜、1108 (G_1, G_2, G_3, G_4) は入力ゲートでありニューロンの入力に相当する。

図11 (b) はその動作を説明するためにさらに簡略化した図面である。各入力ゲートとフローティングゲート間の容量結合係数を C_G 、フローティングゲートとシリコン基板間の容量結合係数を C_0 とすると、フローティングゲートの電位 Z は、

$$Z = -w (V_1 + V_2 + V_3 + V_4) \quad \dots (4)$$

$$W \equiv C_G / (C_0 + 4 C_G) \quad \dots (5)$$

とあらわされる。但しここで、 V_1, V_2, V_3, V_4 はそれぞれ入力ゲート G_1, G_2, G_3, G_4 に入力されている電圧であり、シリコン基板の電位は0V、すなわちアースされているとした。

この μMOS はフローティングゲートをゲート電極とみれば通常のNチャネル MOS トランジスタであり、このフローティングゲートからみた閾電圧（基板表面に反転層が形成される電圧）を V_{TH}^* とすると、 $Z > V_{\text{TH}}^*$ で上記 μMOS はオンし、 $Z < V_{\text{TH}}^*$ ではオフする。つまりこの μMOS 1109を1つ用いて例えば同図(c)のようなインバータ回路を組めば簡単に1つのニューロンの機能が表現できるのである。1110, 1111はインバータを構成するための抵抗、1112は NMOS トランジスタである。同図は(d)は、 $V_{\text{OUT1}}, V_{\text{OUT2}}$ を Z の関数として示したものであり、 $Z > V_{\text{TH}}^*$ の入力に対し V_{OUT2} は V_{DD} のハイレベルの電圧を出力している。つまりニューロンが発火した状態を実現している。

(4) 式で示したように、ニューロンへの入力が電圧レベルで加算され、その線形和が閾値以上になるとニューロンが発火するという基本的な動作がたった1つの μMOS によって実現されているのである。電圧モードの加線を行なうので、入力部で流れる電流はコンデンサの充放電電流のみであり、その大きさは非常に小さい。一方、インバータでは、ニューロン発火時に直流電流が流れるが、これは、負荷として、抵抗1110を用いているためであり、前記発明（特願平1-141463号）による CMOS 構成の μMOS ゲートを用いれば、この直

流電流はなくすることができる。

図12は、CMOS構成の一例を示す図面である。図12(a)はCMOSニューロンゲートの断面構造を模式的に表したものであり、1201はP型シリコン基板、1202はn型のウェル、1203a, 1203bはそれぞれ N^+ 型のソース及びドレイン、1204a, 1204bはそれぞれ P^+ 型のソース及びドレイン、1205はフローティングゲート、1206a～dはそれぞれ入力ゲートの電極である。1207, 1208は例えば SiO_2 等の絶縁膜、1209はフィールド酸化膜である。同図(b)は1個のニューロン回路を構成した例であり、1210は同図(a)のCMOSニューロンゲートを記号であらわしたものであり、符号を付した部分は同図(a)の番号と対応している。1211はCMOSのインバータであり、1212, 1213はそれぞれNMOS及びPMOSのトランジスタである。また、1214はニューロンの出力である。

以上の様に、少数の素子で1つのニューロンが構成でき、しかもパワー消費が非常に少ないため、 ν MOSはニューロンコンピュータを実現する上で、不可欠な素子となっているのである。

しかしながら、ニューロンコンピュータを実現するには、ニューロン以外のもう1つ重要な要素、すなわちシナプスも構成する必要がある。図13は、従来技術による ν MOSトランジスタを用いたシナプス結合も含むニューロン回路の基本構成の一例である。

1301は例えば図11(c)に示したようなニューロン回路であり、1302は他のニューロンの出力信号を伝える配線である。1303はシナプス結合回路であり、入力信号に重みを付与するための回路である。NMOSトランジスタ1304のソース1306に負荷抵抗($R+R_x$)の接合されたソースフォロワー回路となっている。従って、NMOSトランジスタのゲート電極1305に発火したニューロンの出力電圧 V_s が印加されると、ソース1306には、 $V_s - V_{TH}$ なる電圧が出てくる(ここで V_{TH} は、NMOSトランジスタ1304の閾電圧である。))。

例えば、 $V_{TH}=0$ のMOSトランジスタを用いたとすると、ソース1306の電位は V_s と等しくなり、この電圧が2つの抵抗 R , R_x で分割されてシナプス結

合回路の出力電圧となり、結線1307によってニューロン1301に伝えられる。この出力電圧は、 $V_S \cdot R_X / (R + R_X)$ となり、 $R_X / (R + R_X)$ なる重みが信号電圧 V_S に掛けられたことになる。 R_X の値を可変にすることにより重みを変更することができる。

図14(a)は可変抵抗の実現方法の一例を示したものである。例えば、1つのMOSトランジスタ1401のゲートに一定の電圧 V_{GG} を印加してやれば、このトランジスタは1つの抵抗の働きをする。 V_{GG} の値を変化させることによりその抵抗値を変化させることができる。

また、同図(b)は、 V_{GG} の値を制御する回路の一例を示したもので、4ビットのバイナリーカウンタ1402とD/Aコンバータ1403とから構成されている。シナプスの結合強さは、4ビットの2進数で表現され、それが、D/Aコンバータ1403によってアナログ電圧に変換され V_{GG} の値として出力される。シナプス結合強度を強めるには、制御信号によりカウンタの値をカウントダウンさせ、 V_{GG} の値を小さくすればよい。逆にシナプス結合強度を弱めるには、カウントアップさせ、 V_{GG} の値を大きくしてやればよい。

さて、図13及び図14に示したようなシナプス結合回路を用いた場合の問題点を次に説明する。

まず、第1の問題点は図13で重みを発生させるのに抵抗による電圧分割を用いている点である。この方式ではこの抵抗に常に電流を流し続けることによって、重みを掛けた出力電圧を保持しているため、常時 $(R + R_X) \cdot V_S^2 / 2$ の電力を消費することになる。ここでは、ニューロン1301における消費電力をCMOSの応用により減少させても回路全体としての消費電力は決して小さくならない。一層が n 個のニューロンからなる2層の神経回路網を考えると、シナプス結合の数は n^2 個となり、ニューロンの数よりシナプスの数の方が圧倒的に多いのである。従って、常時電流を流し続けなければならない。シナプス結合回路を用いる限り、実用的な規模の神経回路網を構成することは消費電力が過大となり、事実上設計不可能となっている。

第2の問題点は、結合の問題の重みを決める、図14(b)に示した回路が多数の素子を必要とし、高集化ができないという事実である。学習機能を有する神

経回路網を構成するためには、各シナプス結合の強さは適宜変更ができ、かつその変更した値を記憶しておく必要がある。同図では、このために4ビットのバイナリーカウンタを用いているが、これだけでも最低30個程度のMOSトランジスタを必要とする。さらにD/Aコンバータを構成するためにも多くの素子を必要とする。さらにこれらの回路が、1つのシナプス結合当り、さらに多くの電力を消費することになり、消費電力の面からも不利となるのである。

要するに、従来知られた技術では、低消費電力化、高集積化のいずれの面からも神経回路網の構成はほとんど不可能と言わざるを得ない。従って、従来の技術ではニューロンコンピュータを実現することはできないのである。

そこで本発明は、このような問題点を解決するためになされたものであり、消費電力が非常に小さく、かつ少数の素子でシナプス結合が実現でき、高集積度、低消費電力のニューロンコンピュータチップを実現することのできる半導体装置を提供するものである。

発明の開示

本発明の半導体装置は、基板上に一導電型の第1の半導体領域を有し、この領域内に設けられた反対導電型の第1のソース及び第1のドレイン領域を有し、前記第1のソース、及び第1のドレイン領域を隔てる領域に第1の絶縁膜を介して設けられた電位的にフローティング状態にある第1のゲート電極を有し、前記第1のゲート電極と第2の絶縁膜を介して容量結合する複数の第2のゲート電極を有し、前記第2のゲート電極の1つにソース電極が接続された第1のMOS型トランジスタを有し、前記MOS型トランジスタのゲート電極もしくはドレイン電極の少なくとも一方が、高いレベルもしくは低いレベルの2つの電位レベルの信号を伝達する第1の配線に接続されていることを特徴としている。

作用

本半導体装置は、少数の素子によってシナプス結合が構成でき、しかも電力消費が非常に少ないため、神経回路網の高集積化、低電力化が可能となり、これによって初めて実用的なレベルのニューロンコンピュータチップを実現することが

できたのである。

図面の簡単な説明

図1 (a) は第1の実施例を示す回路図。図1 (b) は図1 (a) のシナプス結合回路104のみを取り出した回路図。図1 (c) は $i_n - V_{gs}^n$ の関係を示すグラフ。図1 (d) は $i_p - V_{gs}^p$ の関係を示すグラフ。図1 (e) は図1 (b) の回路で、 V_{in} を0 Vから V_{DD} まで変化させたときの V_{out} の変化を示すグラフ。図2 (a) は第2の実施例を示す模式的平面図。図2 (b) は図2 (a) におけるX-X' 断面図。図2 (c) はシナプスとニューロンからなる本発明の半導体装置を記号を用いて表した図。図2 (d) はシナプス結合の主要部分（フローティングゲートを有するNMOS部分）を示す回路図。図2 (e) は図2 (d) のシナプス結合の主要部分をマトリクス状に配置した回路図。図2 (f) は強誘電体薄膜を用いた不揮発性メモリ素子の断面図。図3 (a) はデプレッション型のPMOSFETの特性を示すグラフ。図3 (b) は図1 (b) に示す回路の時間応答特性例を示すグラフ。図3 (c) は図1 (b) に示す回路の時間応答特性例を示すグラフ。図3 (d) は本発明の変形例を示す概念図。図3 (e) は本発明の変形例を示す回路図。図3 (f) は本発明の変形例を示す回路図。図3 (g) は本発明の変形例を示す回路図。図3 (e) はを示す、フローティングゲートを有する構造図。図4 (a) は第3の実施例を示す回路図。図4 (b) は図4 (a) のシナプス結合の部分を取り出して示した概念図。図5 (a) は1つのシナプス結合と1つのニューロン502からなる神経回路網の主要部分のレイアウト例を示す平面図。図5 (b) は図5 (a) の材料を示す図。図5 (c) は図5 (a) におけるX-X' 断面図。図5 (d) は図5 (a) におけるY-Y' 断面図。図6 (a) は第5の実施例を示す回路の構成図。図6 (b) は図6 (a) を簡略化した図。図7 (a) は第6の実施例を示し、図6 (b) の表記法を用いて表現した4層の神経回路網を示す図。図7 (b) は図7 (a) のI (n) 部を示す図。図8 (a) は第7の実施例を示す回路図。図8 (b) は図8 (a) における V_{in} 、 V_{out} の関係例を示したグラフ。図9は脳の神経細胞の基本単位の構成を示すモデル図。図10 (a) は1個

のニューロンの機能を説明するモデル図面。図10(b)は Z と V_{out} の関係を示すグラフ。図11(a)は ν MOSの一例を示す断面構造図。図11(b)は図11(a)の ν MOSの動作を説明するための概念図。図11(c)は図11(a)の ν MOSを用いたインバータ回路図。図11(d)は図11(c)における V_{OUT1} 、 V_{OUT2} を Z の関数として示したグラフ。図12(a)は先行例を示し、CMOSニューロンゲートの模式的断面図。図12(b)は先行例により1個のニューロン回路を構成した例を示す回路図。図13は先行例を示し、MOSトランジスタを用いたシナプス結合を含むニューロン回路の基本構成例を示す回路図。図14(a)は先行例を示し、可変抵抗の実現方法の一例を示す回路図。図14(b)は図14(a)に関する V_{GG} の値を制御する回路の一例を示す回路図

発明を実施するための最良の形態

(第1の実施例)

本発明の第1の実施例を図1を用いて説明する。

図1(a)は、第1の実施例を示す回路図であり、図において101はニューロン回路であり、例えば図11や図12に示した回路を用いればよい。より低消費電力化を求めるならば、図12の回路を用いた方が好ましい。102a～102dは、ニューロン回路の入力端子であり、例えば図12の回路であれば1206a～1206dに対応している。103はニューロンの出力信号を伝える配線であり、例えば図12に示したようなニューロン回路の出力端子1214に接続されており、そのニューロンが発火しているか否かに従って V_{DD} または0Vの電位を保持している。104は1つのニューロンの出力103と1つのニューロン101の入力102aを接続するシナプスの働きをする回路である。105はNMOSトランジスタであり、106a、106bはそれぞれNMOSのソース及びドレインである。

また、107、108a、108bはそれぞれPMOSトランジスタ及びそのソース及びドレインである。NMOS、PMOSのソース電極106a、108aは共にニューロンの入力端子の1つ102aに接続されている。

NMOSのドレイン106b、PMOSのドレイン108bはそれぞれ V_{DD} 及びアースに接続されている。また、NMOS、PMOSのゲート電極109、110は、共に配線103に接続されている。111のコンデンサCは、入力端子102aに付随する容量を、浮遊容量も含めて表したものであり、回路素子として特に付与したものではない。

本回路の動作を説明するために、シナプス結合回路104のみをとり出して図1(b)に示す。NMOS、PMOSのゲート・ソース間電圧をそれぞれ V_{gs}^n 、 V_{gs}^p とし、また、それぞれに流れる電流を i_n 、 i_p として図に示す様に定義したとき、 $i_n - V_{gs}^n$ 及び $i_p - V_{gs}^p$ の関係を示したのが、それぞれ図1(c)、図(d)である。NMOSはエンハンスメント型(閾値： $V_{TH}^n > 0$)であり、PMOSはその閾値 V_{TH}^p が略々0Vに等しく設定されている。

図1(b)の回路で、 V_{in} を0Vから V_{DD} まで変化させたときの V_{out} の変化を図1(e)に示す。 $V_{in} = 0$ Vのときは、 $V_{out} > 0$ であると $V_{gs}^p < 0$ となりPMOSトランジスタ107がONするため、容量Cは放電して $V_{out} = 0$ となる。このとき $V_{gs}^n = 0$ となりNMOSトランジスタ105はOFF状態となる。次に V_{in} を順次大きくして行き $V_{in} > V_{TH}^n$ となると、NMOSトランジスタ105がONして i_n が流れ、容量Cが充電され V_{out} が上昇する。

V_{out} の値は、

$$V_{out} = V_{in} - V_{TH}^n \quad \dots (6)$$

となる。

このときPMOSのゲートには $V_{gs}^p = V_{TH}^n$ と常に正のバイアスがかかるため、トランジスタはOFF状態となり電流は流れない($i_p = 0$)。つまりこのシナプス回路では、直流電流を一切流すことなく V_{out} のレベルを保持することができるのである。 $V_{in} = 0$ 、すなわち前段のニューロンが発火していないときには $V_{out} = 0$ が次段のニューロン101に inputs され、 $V_{in} = V_{DD}$ 、すなわち前段のニューロンが発火しているときには(6)式で決まる電圧が次段に inputs されるのである。

(6)式で $V_{in} = V_{DD}$ とすると、

$$V_{out} = V_{DD} (1 - V_{TH}^n / V_{DD}) \text{ となり、}$$

このシナプスで掛算される重み w は、

$$w = (1 - V_{TH}^n / V_{DD}) \quad \dots (7)$$

と表されるのである。すなわちNMOSトランジスタ105の閾値 V_{TH}^n の値により重み w の値を決定できるのである。

各シナプス毎に、NMOSトランジスタ105の閾値をそれぞれ所定の値に設定しておけば、シナプス結合強度固定の神経回路網が実現できる。各NMOSトランジスタの閾値をかえるには、例えば製造プロセス中でのチャネル部への不純物(B, Pまたは A_S 等)のイオン注入量を1つ1つかえてやる必要がある。これには例えばFIB(Focused Ion Beam)の技術を用い、イオンビームを絞ってトランジスタのチャネル部に照射し、あらかじめ設定されたドーズ量を順次打ち込んで行けばよい。このようにして実現されたシナプス結合強度固定の神経回路網は、学習機能はないが、順回セールスマン問題等の特定の問題に対し、非常に高速に演算処理できることが分かっている。

本発明の第1の実施例で実現した神経回路網は、シナプス結合部では、容量 C の充放電電流が流れるのみであり、従来の抵抗分割形のシナプス結合(図13)にくらべ消費電力をほとんど無視できるくらい小さくできるという大きな特徴がある。また、たった2つのトランジスタで実現されているためシナプス部がチップ上で占める面積も小さく、高集積化に対し有利であるという特徴を有する。また、配線103の電位が0Vから V_{DD} に変化し、容量 C を充電する際、従来例(図13)ではNMOSトランジスタ1304と抵抗 R を通して充電するのに対し、本発明ではNMOSトランジスタ105のみを通して充電するため、実効的な推抗値が小さくなり、それだけ動作を高速化することができる。

図1(a)では、NMOSトランジスタ105のドレイン106bは V_{DD} に接続したが、これは信号線103に接続してもよい。この場合、信号線103は容量 C を充電するための電流も供給する必要があるため、前段のニューロン回路の出力は十分大きな電流供給能力を備える必要がある。この場合は、各シナプス結合に V_{DD} を供給する電源ラインが不要となり、配線数を減少させることができる。

(第2の実施例)

次に本発明の第2の実施例を図2に示す。

この実施例ではシナプス結合強度が可変であり、いわゆる学習機能をもった神経回路網を構成することができる。

図2 (a) は平面図を模式的に示したものであり、同図 (b) はそのX-X'における断面図である。201, 202はそれぞれNMOS及びPMOSトランジスタ、203はフローティングゲートでNMOSのチャネル領域上に例えば厚さ200 Åのゲート酸化膜204を介して設けられており、例えばN⁺型の多結晶シリコン薄膜で形成されている。また、フローティングゲート203は、例えば約50~70 ÅのSiO₂膜205を介してN⁺領域206と対向している。このN⁺領域206は電荷注入用電極である。

207はPMOSトランジスタのゲート電極であると同時にNMOSのトランジスタのフローティングゲートと例えば約200 ÅのSiO₂膜208を介して容量的に結合しており、NMOSトランジスタ201のゲート電極の機能も担っている。この電極は例えばN⁺のポリシリコンでできており、コンタクトホール209を介して例えばA1の配線210に接続されている。このA1配線は、例えば、図11や図12のようなニューロン回路の出力に接続されている。

211は、例えばWSi₂ (タングステンシリサイド) で形成された電極であり、フローティングゲート203と例えば200 ÅのSiO₂膜212を介して容量結合している。213a, 214aはそれぞれNMOS及びPMOSトランジスタのソースであり、ともに次段のニューロンの入力端子の1つに接続されている。213b, 214bはそれぞれNMOS及びPMOSトランジスタのドレインであり、それぞれ電源ライン (V_{DD}) 及びアースライン (V_{SS}) に接続されている。

図2 (a) においては不要な複雑さをさけるため、V_{DD}, V_{SS} 及び次段のニューロン入力への接続配線は省略してある。また、図2 (b) において、215はフィールド酸化膜、216は例えばP型のシリコン基板である。

図2 (c) は、シナプスとニューロンからなる本発明の半導体装置を記号を用いて表したものであり、各部分は、同図 (a), (b) において同一の番号を付した部分と対応している。217は1つのニューロン回路を示しており、

218a～218dはそのニューロン回路の入力端子であり、第1の実施例を示す図1の101及び102a～102dにそれぞれ対応している。

図2(c)と図1(a)を比較すれば明らかであるが、本実施例も第1の実施例と同様NMOS、PMOSを直列につなぐことによりシナプス結合を実現している。異なるのは、NMOSがフローティングゲート203を有していることである。通常の動作時は、例えば電極211、206は0VとしておくことNMOSトランジスタ201はゲート電極207に V_{TH}^n 以上の電圧を与えるとONするエンハンスメント型トランジスタとなり、第1の実施例で説明したのと全く同様の動作をする。

すなわち信号線210の電位が V_{DD} となると218aには $V_{DD} - V_{TH}^n$ なる値が入力され、信号線の電位が0Vになると、218aには0Vが入力される。フローティングゲートを有するNMOSトランジスタ201の働きをより詳しく説明するために、この部分のみをとり出して記号で表したのが図2(d)である。同一番号を付した部分はすべて図2(a)～図2(c)の各部分と対応している。207はNMOSのゲート電極(Gと略記)、211はシナプス強度変更用制御電極(Wと略記)、206は電荷注入用電極(Cと略記)、203はフローティングゲート(FGと略記)である。

今FGの電位をZと表すと、

$$Z = (C_G V_G + C_W V_W + C_C V_C + Q_F) / (C_{OX} + C_C + C_W + C_G) \cdots (8)$$

と表される。

ここで Q_F はフローティングゲート内の電荷の総量であり、 V_G 、 V_W 、 V_C はそれぞれG、W、Cの電位であり、 C_{OX} 、 C_C 、 C_W 、 C_G は図2(c)にて定義した各電極とFG間の容量結合係数である。またここでP型シリコン基板の電位は0Vとした。

前述したとおり、例えば $V_W = 0$ 、 $V_C = 0$ とすると、(8)式より

$$Z = (C_G V_G + Q_F) / C_{TOT} \cdots (9)$$

$$C_{TOT} \equiv C_{OX} + C_C + C_W + C_G$$

FGからみたNMOSトランジスタの閾電圧(NMOSにチャネルが形成されるFG電位の最小値)を V_{TH}^* とすると、(9)式より、

$$(C_G V_G + Q_F) / C_{TOT} > V_{TH}^*$$

のときNMOSがONする。

すなわち

$$V_G > (C_{TOT} / C_G) V_{TH}^* - (Q_F / C_G)$$

となる。

つまり、ゲート電極207からみたNMOSの閾電圧 V_{TH}^n は、

$$V_{TH}^n = (C_{TOT} / C_G) V_{TH}^* - (Q_F / C_G) \dots (10)$$

で与えられることになる。 C_{TOT} 、 C_G 、 V_{TH}^* 等は、デバイスの構造と製造プロセスで決まるため、デバイスが出来上ってしまえば変更できない。従って Q_F の値を変更すれば、 V_{TH}^n を変化でき、その結果(7)式に従ってシナプス結合の重みをかえることができる。通常動作のとき、 V_W 、 V_C ともに0Vであるとしたが、一般に他の値としてももちろんよい。このとき(10)式は、

$$V_{TH}^n = (C_{TOT} / C_G) V_{TH}^* - (Q_F / C_G) - (C_W / C_G) V_W - (C_C / C_G) V_C \dots (11)$$

のように、 V_W と V_C に比例した項がつけ加わるだけである。 V_W 、 V_C を一定の値に保つかぎり V_{TH}^n は Q_F の値のみで決るのである。

Q_F の値を変化させるには、例えば次のようにすればよい。 V_{PP} を V_{DD} よりも大きなプログラム電圧とする。たとえば V_{DD} を5V、 V_{PP} を10Vとしたとする。今、簡単のために、 $C_W = C_G = (C_{OX} / 2) \gg C_C$ にしたと仮定する。この条件は、図2(a)のレイアウトの設計により、近似的に成り立たせることは容易である。 $V_C = 0$ 、 $V_W = V_G = V_{PP}$ とすると、(8)式より $Z = (V_{PP} / 2) = 5(V)$ となる。ここで $Q_F = 0$ であると仮定した。

酸化膜205の厚さを例えば50Åであるとする、この酸化膜には10MV/cmの電界がかかることになり、ファウラー=ノルドハイムトンネル電流が流れ、FG内へはCより電子が注入され、 $Q_F < 0$ となる。

(8)式より $Q_F < 0$ となればZは小さくなるため、トンネル電流は減少し、ある一定量の電子が注入された状態でストップする。もちろん注入される電荷量を正確に制御するためには、 V_W 、 V_G の値をパルス状に変化させてもよい。パルス巾あるいはパルスの高さ、あるいはパルス巾とパルスの高さともに一定でパル

スの数を変化させることにより Q_F の絶対値をコントロールできる。

電子注入は、 V_C をパルス状に変化させることによっても行える。例えば、最初、 V_W 、 V_G 、 V_C をすべて V_{pp} としておき、 V_C のみパルスの的に0 Vにおとしてやってもよい。同様に電子が注入され、その量をコントロールすることができる。以上は、 Q_F が負となる変化であり、(10)式、(11)式より V_{TH}^n の値は大きくなる。すなわち(7)式より、これはシナプスの結合強度 w を小さくする変更操作となる。

次にシナプスの結合強度を強める変更操作について述べる。

これには $V_W = V_G = 0$ (V)とし、 $V_C = V_{pp}$ とすればよい。(8)式より $Z = 0$ Vとなり、酸化膜には今度は逆向きに10 MV/cmの電界が発生し、FGからCへ電子がファウラー=ノルドハイムトンネル現象によって流れる。その結果、FG内には正の電荷が残ることになり $Q_F > 0$ となる。

(10)式、(11)式より V_{TH}^n は小さくなり、(7)式より w は大となるため、これはシナプスの結合強度を大きくする変更操作となる。またこれは、例えば $V_W = V_G = V_C = 0$ Vとしておいて V_C のみパルスの的に V_{pp} に上げてやってもよい。あるいは V_W 、 V_G 、 V_C をすべて V_{pp} としておいて、 V_W 、 V_G のみパルスの的に0 Vに落としてやってもよい。これらのパルス巾、パルス高さあるいはパルス数の制御により、 Q_F の絶対値を自由に制御できることは言うまでもない。

さらにこの方式の重要な点は、所定のシナプスの結合のみ、選択的に、その結合強度 w を変更できることである。図2(d)に示したシナプス結合の主要部分(フローティングゲートを有するNMOS部分)を4個(212、213、214、215)マトリクス状に配置したのが図2(e)である。 G_1 、 G_2 は、信号線210に対応し、2本平行に配置されている。また、 W_1 、 W_2 は信号線211に対応しやはり2本平行に配置されている。例えば212のシナプス結合のみ変更(弱める)しようとする、 G_1 と W_1 のみ V_{pp} の電位を与え、 G_2 と W_2 は0 Vとすればよい。もちろん $V_C = 0$ (V)としておく。図2(d)の説明より、212ではフローティングゲートに電子が注入され、結合強度 w が小さくなる。215のシナプスでは $V_W = V_G = 0$ (V)であり、 $Z = 0$ (V)となるため電子のトンネリングは生じず、 w の値はかわらない。213あるいは214のシ

ナプスでは $V_G = V_{pp}$ 、 $V_W = 0$ あるいは、 $V_G = 0$ 、 $V_W = V_{pp}$ のいずれかであり、 $Z = (1/4) V_{pp} = 2.5$ (V) となる。このとき酸化膜 205 にかかる電界は 5 MV/cm と、212 のシナプスにくらべて半分の大きさとなる。トンネル電流は電界強度に対し指数関数的に変化するため、トンネル電流は流れず、従ってこのシナプスでは Q_F の変化は生じない。

つまり、高電位 V_{pp} を与えた直交する 2 本の信号線の交点に位置するシナプス結合のみ、その結合強度を弱めることができる。すなわち特定のシナプスのみ選択的に結合強度を変更できるのである。結合強度を選択的に強めるためには、まず $V_C = V_{pp}$ とし、 G_1 、 G_2 、 W_1 、 W_2 はすべて V_{pp} につり上げる。次に、例えば G_2 、 W_2 のみを 0 V とすれば、その交点に位置する 215 のシナプスのみその結合強度を強めることができる。

以上の説明は 4 つのシナプス結合についてのみ説明したが、もっと規模の大きなシナプスのマトリクス配置に対しても同様に選択的な強度変更が行えることは明かである。

なお、図 2 (e) は、説明の都合上 NMOS トランジスタのみをマトリクス状に配置して説明を行ったが、各 NMOS トランジスタには、図 2 (a) ~ (c) に示した PMOS 及び関連する配線が付随していることは言うまでもない。

ニューロン回路 217 との接続に関しては第 4 の実施例に関連して詳しく説明する。

本実施例の神経回路網は、1 つのシナプス結合が、図 2 (a) に示した如く 1 個の NMOS トランジスタと 1 個の PMOS トランジスタで構成することができるため、チップ上に占める面積が非常に小さく、シナプス結合を高密度に集積可能である。しかもこのシナプス結合は、フローティングゲートに蓄えられた電荷によってその結合強度を決めることができるため、結合強度を自由に変更し、かつその値を記憶しておくことができる等、生体のシナプスとほぼ同等の機能をたった 2 つの素子で実現しているのである。また数多くマトリクス状に配置したシナプス結合に対し、特定のシナプスのみ選択的に任意の強度に調整することができる機能を有している。さらに、シナプスの出力電位を一定値に保つに際し、直流電流を一切流す必要がないため、電力消費も非常に小さくできる等、超高密度

神経回路網構成に正しく理想的な特性を有しているのである。

従来のシナプス結合（図13、図14）では、電力消費が大きいばかりか、結合強度の記憶に数多くの素子を必要としており、高集積化はほとんど不可能であったのが、本発明によりはじめて高集積度神経回路網の実現が可能となったのである。

以上第2の実施例では、電極203、207に材料として N^+ 型のポリシリコンを、211の材料として WSi_2 を、また、210の材料としてAlを用いた場合について述べたが、これはあくまで一つの例であり、同様の機能を実現できるものであれば他のいかなる材料を用いてもよいことはもちろんである。また、絶縁膜204、205、208、212等の材料、膜厚等に関してもここに例示した材料、膜厚に限定されることがないことは言うまでもない。また、電荷注入用電極206を独立の電極として設けた場合について述べたが、これは例えば、NMOSトランジスタ201のソース電極213aあるいはドレイン電極213bの延在する部分に設けてもよい。こうすれば、電極206に電位を与えるための配線が不要となり、高集積化に対しては有利となる。また、 N^+ 層ではなく、例えばポリシリコンの電極で形成してもよい。

第2の実施例では、選択的な結合強度変更を、2つの電極211(W)、207(G)に加える電圧の組合わせにより行っている。211(W)は強度変更専用設けた電極であるが、207(G)はNMOSのゲート電極を兼ねている。これは、例えば図2(d)に於いてWと同様の強度変更専用電極をもう1つ追加し、これら2つの電極に加える電圧の組合わせでFGに注入される電荷量を制御してもよい。

さらに電荷注入の方法として、ファウラー=ノルドハイムトンネル現象を用いる場合についてのみ説明したが、これはアバランシェ注入やホットエレクトロン注入を用いてもよいことは言うまでもない。

また、第2の実施例ではNMOSトランジスタの閾値を可変とするために、フローティングゲート型のMOSトランジスタを用いているが、これは例えばMNOS型のデバイスでもよい。あるいは、図2(f)に断面構造を示したような、強誘電体薄膜を用いた不揮発性メモリ素子でもよい。

同図に於いて、216はP型のSi基板、217は例えば50ÅのSiO₂膜、218は強誘電体薄膜で、例えば膜厚3000ÅのPZT (Pb(Zr_XTi_{1-X})O₂)を用いる。219は例えばTiの電極である。220a, 220bはN⁺型のソース及びドレインである。ゲート電極219に正又は負のパルスを追加することにより強誘電体薄膜を分極させ、その分極の大きさによってトランジスタの閾値を制御するのである。

以上述べたように、NMOSトランジスタの閾値を可変とするに際し、いかなる手段を用いてもよいことは言うまでもない。図2(c)の構成を用いた場合、フローティングゲート内の電荷量により、閾値 V_{TH}^n を制御できることは既に述べた通りであるが、(11)式より明らかなように、 V_W , V_C に所定の電位を与えることによっても V_{TH}^n を可変とすることができる。シナプス結合をこのように V_W や V_C の電位コントロールによって行ってもよいことは言うまでもない。

図1(a)、及び図2(a)、図2(c)に示したそれぞれ第1及び第2の実施例では、PMOSトランジスタ(107, 202)がNMOS(105, 201)とアース(V_{SS})ラインの間に挿入されている。このPMOSトランジスタの働きを図1(b)を用いて説明する。

$V_{in} = V_{DD}$ のとき $V_{out} = V_{DD} - V_{TH}^n$ となり、容量Cには $C(V_{DD} - V_{TH}^n)$ なる電荷が蓄えられる。次いで V_{in} が0に変化したとするとPMOSトランジスタは、容量Cにたまった電荷をアースに逃がす働きをしている。従ってこのMOSトランジスタは $V_{in} = V_{DD}$ のときはOFFであり $V_{in} = 0V$ のときはONしなければならない。また容量を放電する時間を短くしようと思えばできるだけ大きな電流を流せるよう設計しなければならない。

これには例えば図3(a)に示したような特性をもつデプレッション型のPMOSFETを用いればよい。例えば V_{in} が V_{DD} から0に変化した瞬間のPMOSのゲートにかかる電圧は、 $V_{gs}^p = -(V_{DD} - V_{TH}^n) < 0$ となる。図3(a)と図1(d)を比較すれば、 $V_{TH}^p = 0$ のトランジスタより、 $V_{TH}^p > 0$ のトランジスタの方がより多くの電流を流せることは明かである。

図3(b)に、図1(b)の回路の時間応答特性を示す。立ち上がり時間は、NMOSトランジスタ105のON特性で決まるので変わらないが、立ち下り特

性はデプレッション型のPMOSトランジスタを用いた方が応答が速い。さらに立ち下り時間を短くするためには、 V_{TH}^P の値を大きくしてやればよい。

また一方、 V_{in} が0 Vから V_{DD} へと変化する際、NMOSトランジスタ105、201は $V_{in} \geq V_{TH}^n$ でONするが、このときPMOSトランジスタはOFFしていないと V_{DD} から V_{SS} （アース）への直流電流が流れ電力消費が大きくなる。

これを防ぐには、

$$0 < V_{TH}^P < V_{TH}^n \dots (12)$$

の不等式を満たすようデプレッション型PMOSトランジスタの閾値 V_{TH}^P を設定しなければならない。

直流電流を流さず、かつ容量Cの放電時間をできるだけ小さくして回路動作の高速化を達成するためには $V_{TH}^P < V_{TH}^N$ を満足しつつ、かつ

$$V_{TH}^P \approx V_{TH}^N \dots (13)$$

としてやればよい。

またPMOSトランジスタ（107、202）は $V_{TH}^P < 0$ であるエンハンスメント型PMOSFETを用いてもよい。

この場合、図1（b）の回路の時間応答特性は図3（c）のようになる。立ち上がり特性は、同図（b）と同じであるが、立ち下がり特性は全く異なる。すなわち V_{out} が $|V_{TH}^P|$ に等しくなった時点で出力電圧が一定値となる。なぜならこの時点でPMOSトランジスタがOFFするからである。

しかしこの電圧レベルは時間とともに減少し十分時間が経過すれば0 Vとなる。これはPN接合のリークやPチャネルトランジスタのサブスレッショルド電流等によって容量Cに蓄えられた電荷が逃げて行くことによる。

このようなシナプス結合は、興奮状態をしばらく記憶しておくことができる。つまり、発火したニューロンからの信号が一度入力される（ $V_{in} = V_{DD}$ ）とその後、そのニューロンの出力が低レベルとなっても、 V_{out} には $|V_{TH}^P|$ に等しい電圧が保持されるのである。保持する時間をコントロールするには、例えばPMOSトランジスタ（107、202）と並列に抵抗Rを接続してやればよい。こうすればRCの時定数で出力電圧は0 Vに漸近する。また、持続する出力

電圧のレベルを調整するには、 V_{TH}^P の値を変化させてやればよい。

以上述べたようにPMOSトランジスタ(107, 202)の V_{TH}^P の値を変化させることにより様々な特性をもつシナプス結合が実現できる。 V_{TH}^P は製造プロセス中に、閾値調整用のチャネルイオン注入等によって適当な値に設定してもよい。また、フローティングゲートを有する構造にして、フローティングゲート内の電荷の量を増減させ、 V_{TH}^P を可変としてもよい。あるいは、図3(d)に示した様なフローティングゲートを有する構造として、 V_C の電位あるいは C_C の値により V_{TH}^P の値を変化させてもよい。

V_{TH}^P は、

$$V_{TH}^P = (C_{TOT}/C_G) V_{TH}^* - (Q_F/C_G) - (C_C/C_G) V_C \quad \dots (14)$$

で与えられる。

C_C をかえるには電極301とフローティングゲート302の重なり面積をかえることにより行える。パターン設計の段階で所定の重なり面積にしておけば、イオン注入のドーズ量を打ち分ける必要もなく製造プロセスが簡単となる。また V_C の値によって変化させれば V_{TH}^P を必要に応じてかえることができ、例えば特定のシナプス結合のみ一定時間興奮状態を持続させるようにすることも可能である。

以上、容量Cの電荷を放電するのに、NMOSトランジスタと共通のゲート電極をもつPMOSトランジスタを用いた場合についてのみ述べてきたが、本発明の特徴は、これ以外の負荷素子を用いた場合にも失われることはない。

これらの例を図1(b)に相当する部分回路についてのみ示したのが図3(e)、図3(f)、図3(g)である。PMOSトランジスタのかわりに抵抗303、ゲートをソースにつないだデプレッション型のPMOS304、ゲートをドレインにつないだエンハンスメント型NMOS305等である。これ以外にも、図には示していないが、例えばゲートをドレインにつないだエンハンスメント型のPMOS、ゲートをソースにつないだデプレッション型のNMOS、もしくは一定のゲート電圧を印加したMOSトランジスタ等、負荷抵抗として働くものであれば何を用いてもよい。

いずれの場合も V_{out} のレベルを保持するのに直流電流が流れるため、消費電

力低減の特徴は失われる。しかしこれは負荷の実効的な抵抗値を必要に応じて大きくしてやることで解決できる。しかしこれらの回路を、例えば閾値可変の NMOS トランジスタ 201 と組み合わせて、第 2 の実施例のように用いると、非常に少ない数の素子によって、結合強度も可変でかつその強度の記憶機能を持ったシナプスが得られるのである。これは図 14 (a)、図 14 (b) で示したような従来技術の方法では決して実現することのできない大きな特徴である。

(第 3 の実施例)

次に本発明の第 3 の実施例を図 4 を用いて説明する。

図 4 (a) は、記号で示されているが各記号の意味は、図 2 (a) ~ (d) と同様である。

401、402 はそれぞれ NMOS 及び PMOS のトランジスタであり、ともに共通のフローティングゲート (FG) 403 を持っている。404 は FG と例えば SiO_2 等の絶縁膜を介して容量結合しているゲート電極であり、NMOS (401) 及び PMOS (402) 共通のゲート電極 (G) の働きをしている。

405 は第 2 図 (c) の 211 と同様のシナプス強度変更用制御電極 (W) であり、406 は 206 と同様の電荷注入用電極 (C) である。その他の構成は第 2 の実施例と全く同じであり、407 は全段のニューロンの出力信号を伝える信号線であり、408 はシナプスの出力電圧を次段のニューロン 409 の 1 つの入力端子に接続する配線である。

本実施例の動作を説明するためにシナプス結合の部分を取り出して示したのが図 4 (b) である。各電極 G、W、C に印加する電圧をそれぞれ V_G 、 V_W 、 V_C とし、シリコン基板の電位を 0 V とすると、FG の電位 Z は、

$$Z = (C_G V_G + C_W V_W + C_C V_C + Q_F) / C_{TOT} \quad \dots (15)$$

$$C_{TOT} \equiv C_{OX}^n + C_{OX}^p + C_G + C_W + C_C$$

で与えられる。

FG からみた NMOS、PMOS の反転層のできる閾値をそれぞれ V_{TH}^{n*} 、 V_{TH}^{p*} とすると、NMOS、PMOS 共通のゲート電極 G からみた閾値 V_{TH}^n 、 V_{TH}^p は次式で与えられる。

$$V_{TH}^n = (C_{TOT} / C_G) V_{TH}^{n*} - (Q_F / C_G) - (C_W / C_G) V_W$$

$$- (C_C / C_G) V_C \quad \dots (16)$$

$$V_{TH}^P = (C_{TOT} / C_G) V_{TH}^{P*} - (Q_F / C_G) - (C_W / C_G) V_W \\ - (C_C / C_G) V_C \quad \dots (17)$$

シナプス結合強度変更時以外は、通常 V_W , V_C には一定の電位を供給しておく。例えば $V_W = V_C = 0$ とすると、

$$V_{TH}^n = (C_{TOT} / C_G) V_{TH}^{n*} - (Q_F / C_G) \quad \dots (16')$$

$$V_{TH}^P = (C_{TOT} / C_G) V_{TH}^{P*} - (Q_F / C_G) \quad \dots (17')$$

となり、ともに Q_F の値を変化させることにより閾値を変更でき、このことにより (7) 式に従ってシナプス結合の強度を変更できるのである。

(16') 式、(17') 式より、

$$V_{TH}^n - V_{TH}^P = (C_{TOT} / C_G) (V_{TH}^{n*} - V_{TH}^{P*}) \quad \dots (18)$$

となる。

この式において C_{TOT} , C_G , V_{TH}^{n*} , V_{TH}^{P*} 等は、すべて製造プロセスで決まる値であり、素子ができ上がれば変化しないパラメータである。すなわち $V_{TH}^n - V_{TH}^P$ の値は常に一定であり、シナプス結合の強度をかえても変化しないのである。この事実は次のような大きな意味をもっている。第2の実施例に関し、図3(a)、図3(b)を用いて説明したように、シナプス出力の立ち下り特性を高速化するためにはPMOSトランジスタはデプレッション型 ($V_{TH}^P > 0$) とし、且つ V_{TH}^P の値をできるだけ大きくしてやることが有効であった。

しかし、直流電流が流れることによる消費電力の増大を防ぐためには、

(12) 式の条件、すなわち

$$V_{TH}^P < V_{TH}^n \quad \dots (12)$$

が満足されなければならない。

第2の実施例では、NMOSトランジスタ201のみが閾値可変となっているため、 V_{TH}^n は変化するが V_{TH}^P は一定値のままである。従って常に (12) 式を成り立たせる為には、 V_{TH}^P の値の設定には制限が生じる結果となる。例えば V_{TH}^n の値として0から V_{DD} まで変化させることを考える ((7) 式で $w = 1 \sim 0$ に相当) と、常に (12) 式を満たそうと思えば $V_{TH}^P \approx 0$ としなければならない。しかるに本第3の実施例では (16) 式、(17) 式より明らかなように

V_{TH}^n , V_{TH}^p はともに連動して変化し、その差は常に一定値を保つため、常に(12)式を満足させることが可能である。

つまり製造プロセスによって、

$$V_{TH}^p \cong V_{TH}^n \quad \text{でかつ} \quad V_{TH}^p < V_{TH}^n$$

と設定しておけば、常に最適の状態、すなわち消費電力の増大なく、立下り時間の高速化が実現できるのである。

本第3の実施例は、NMOSとPMOSがフローティングゲートを共有している以外の点では、第2の実施例と同じである。従ってシナプス結合強度の変更に關しても全く同様に行えることは言うまでもない。

(第4の実施例)

次に、本発明の第4の実施例について図5を用いて説明する。図5(a)は、1つのシナプス結合と1つのニューロン502からなる神経回路網の主要部分のレイアウトを描いたものである。

この回路は二層ポリシリコン、二層アルミニウム配線の技術を用いたCMOSプロセスで実現されたものである。各部のパターンがいかなる材料でできているかの凡例を同図(b)に記号で示してある。A1配線はX方向に走っているものが第1層目の配線、Y方向に走っているのが第2層目の配線である。また V_{DD} 、 V_{SS} 、 V_C に関しては、レイアウトを見易くするためにA1配線を省略してあるが、これらは、例えばX方向に走る第1層目のA1配線で電位を供給してやればよい。

図5(c)、図5(d)に同図(a)のX-X'及びY-Y'における断面図を示す。

503、504はそれぞれNMOS及びPMOSトランジスタで、505は両者に共通のフローティングゲート(FG)である。505、506、507はそれぞれNMOS、PMOS共通のゲート電極(G)、シナプス結合変更用制御電極(W)、電荷注入用電極(C)であり、図4(a)と同様の機能をもつ回路を具体的なパターンとして表したものとなっている。

本実施例では、電荷注入用電極508は、ポリシリコン電極505上に成長させた例えば約70Åの熱酸化膜509(図5(c)参照)を介して設けられた、

2層目のポリシリコンにより形成されている。しかしながらこれはもちろん第2の実施例のように N^+ 拡散層206を用いてもよい。

W(507)はコンタクトホール510を介して、第2層A1配線511に接続されている。またG(506)は、信号線512に接続されている。信号線512は図1(a)の103、図2(a)の210と同様にニューロンの出力端子につながっている。

513は絶縁膜514を介して、フローティングゲート(FG)515と容量結合している。515はNMOSトランジスタ516及びPMOSトランジスタ517のフローティングゲートとなっており、CMOSのニューロンゲート519を構成している。これは丁度図12(a)及び同図(b)の1210に相当するデバイスである。その出力はA1配線520によってCMOSインバータ521に入力され、その出力が第1層のA1配線522に出てくる。ここではこの出力はスルーホール523を介して第2層のA1配線524に接続され、さらに512と平行のA1配線525に接続されている。配線524は下方にも伸びている(524b)が、これは次層のニューロンへの出力線となっている(詳しくは第5の実施例で説明する)。

525の信号線はニューロン502の出力を同じニューロンの入力側へフィードバックする信号線を形成しており、501と同様のシナプス回路を介して例えば、ニューロン502の入力に接続される。図において、526はP型Si基板、527はn型のウェル、528はフィールド酸化膜である。

以上のように、本発明の第4の実施例では、フィードバック機能も含め非常にわずかなチップ面積で神経回路網の最も重要な機能が実現されているのが分かる。

以上は1ヶのシナプスと1ヶのニューロンについてのみ説明したが、多数のニューロンからなる神経回路網にも容易に拡張できる。

(第5の実施例)

図6(a)は本発明の第5の実施例を示す回路の構成図面である。 $\psi_I^1, \psi_I^2, \dots, \psi_I^n$ は第1層のn個のニューロン群をあらわしており、各々は図5(a)の502と同様の回路を有している。 $\psi_{II}^1, \psi_{II}^2, \dots, \psi_{II}^n$ 、及び ψ

$I_{III}^1, I_{III}^2, \dots, I_{III}^m$ はそれぞれ第2層のn個のニューロン群、第3層のm個のニューロン群をあらわしている。各ニューロンはすべて502と同様の構造を有している。 $O_I^1, O_I^2, \dots, O_I^n$ は第1層のニューロンからの出力線であり、図5(a)の512の信号線に対応している。

また $I_{II}^1, I_{II}^2, \dots, I_{II}^n$ は第2層のニューロン群への入力線であり、図5(a)の515のフローティングゲートに相当している。 $O_{II}^1, O_{II}^2, O_{II}^3, \dots, O_{II}^n$ は第2層のニューロン群の出力線であり、一部は601a, 601b等の配線を介して第3層のニューロン群の入力線と結合している。また一部は602a, 602b等の配線により、第1層ニューロン群の出力線と平行の配線603a, 603b等に接続されており、フィードバックループを形成している。

601a, 602a, 603aはそれぞれ524b, 524, 525の配線に対応している。604a, 604b等はシナプス結合回路であり、例えば図5(a)の501に対応している。

本実施例では、3層のニューロン群からなる神経回路網のレイアウトを示したものであるが、本発明によると、2次元的に規則正しく配列することにより神経回路を構成することができ、LSI化にとって極めて有利であることが分る。

図6(b)は、同図(a)を簡略化して示したものである。605の $I_I(n)$ は第1層のニューロン群を表しておりnはn個のニューロンからなることを意味している。606の $S(2n, n)$ は2n個の出力(第1層のn個の出力と第2層のn個の出力がフィードバックされたもの)とn個の第2層への入力とを結びつけるマトリクス状に配置されたシナプス群を表しており、シナプスマトリクスと呼ぶ。607の矢印は信号の流れを表している。608の矢印はn個の出力がフィードバックされていることを表している。このような標記法を用いれば、もっと複雑な神経回路網をも簡単に表現することができる。

(第6の実施例)

図7(a)は、図6(b)の標記法を用いて表現した4層の神経回路網であり、本発明の第6の実施例である。ここで $I(n)$ は入力バッファ層であり、同図(b)に示したようにn個の増幅器からなっており、「0」または「1」の

2進信号の入力 $I_1, I_2, \dots, I_n$ をそれぞれ 0 V 及び V_{DD} のレベルとして出力する回路である。図 7 (a) より明らかなように、フィードバックを有する 4 層の神経回路網が規則正しい 2 次元的な配列によって実現されている。人間の脳は 6 層のニューロン群から構成されているといわれているが、同様の配列を平面的にくり返すことにより、何層のニューロンを含む神経回路網でも容易に構成できる。すなわち本発明の半導体装置は、ニューロンコンピュータ実現に非常に有利な特徴を有している。

(第 7 の実施例)

図 8 は、本発明の第 7 の実施例を示す図面である。

NMOS 401 のドレインが 0 V に、PMOS 402 のドレインが $-V_{DD}$ に、また、G が信号線 801 に接続されている以外は、図 4 (a) と同様である。従って図 4 と同じ番号が付してある。

今、404 の電位を V_{in} 、408 の電位を V_{out} とし、 V_{in} 、 V_{out} の関係を示したのが図 8 (b) である。ここで V_{TH}^n は NMOS トランジスタ 401 のゲート電極 404 よりみた閾値である。 $V_{in} = V_{DD}$ のときには $V_{out} = 0$ となり、 $V_{in} = 0$ のときには $V_{out} = -V_{TH}^n$ となる。図より V_{in} は信号線 801 の電位と等しいが、801 の電位はインバータ 802 によって、407 の信号線の電位を反転させた値に等しくなっている。すなわち、前段のニューロンが発火しているときは $V_{in} = 0$ となり、発火していないときは $V_{in} = V_{DD}$ となる。つまりこのシナプスは、前段のシナプスが發火したには $-V_{TH}^n$ なる負の出力を次段のニューロンに伝えるのである。

つまり次段のニューロンが発火するのを抑えるように作用する。これは抑圧性のシナプス結合として働くのである。このように、本発明によれば、生物体に存在する抑圧性シナプスも容易に構成できる。さらに、抑圧レベルは $-V_{TH}^n$ に等しいのであるから、フローティングゲートの電荷量を変化させることにより、このレベルも変化させられるのである。

以上述べたすべての実施例において、103, 210, 407 等の配線は、ニューロンの出力端子につながっている場合について説明したが、必ずしもこれに限定されることはない。例えば図 7 (a) のように入力バッファの出力につな

がっていてもよい。またニューロンの出力データが一度、フリップフロップ等にラッチされ、その値をパストランジスタ等を通してこれらの配線に供給してもよい。また適当な増幅器を介してから供給してもよいことは言うまでもない。

産業上の利用可能性

以上述べたように本発明の半導体装置は、人間の脳に近い基本機能を有した神経回路網をわずかな素子数でしかも低消費電力で実現できる。さらに規則正しい平面的な素子の配列で複雑な神経回路網が構成できるためニューロンコンピュータ実現に最も適した半導体装置である。

請求の範囲

1. 基板上に一導電型の第1の半導体領域を有し、この領域内に設けられた反対導電型の第1のソース及び第1のドレイン領域を有し、前記第1のソース、及び第1のドレイン領域を隔てる領域に第1の絶縁膜を介して設けられた電位的にフローティング状態にある第1のゲート電極を有し、前記第1のゲート電極と第2の絶縁膜を介して容量結合する複数の第2のゲート電極を有し、前記第2のゲート電極の1つにソース電極が接続された第1のMOS型トランジスタを有し、前記MOS型トランジスタのゲート電極もしくはドレイン電極の少なくとも一方が、高レベルもしくは低レベルの2つの電位レベルの信号を伝達する第1の配線に接続されていることを特徴とする半導体装置。
2. 前記第1のMOS型トランジスタにおいて、その閾値が所定の値に変更できるゲート構造を有したことを特徴とする請求項1記載の半導体装置。
3. 前記第1のMOS型トランジスタがフローティングゲートを有するMOS型トランジスタであることを特徴とする請求項2記載の半導体装置。
4. 前記第1のMOS型トランジスタがMNOS（金属-窒化膜-酸化膜-半導体）構造を有するトランジスタであることを特徴とする請求項2記載の半導体装置。
5. 前記第1のMOS型トランジスタがNチャネル型であり、そのソース電極がPチャネル型の第2のMOS型トランジスタを介して、低レベル電圧を供給する電源ラインに接続され、かつ前記第1、第2のMOS型トランジスタのゲート電極が前記第1の配線に接続されていることを特徴とする請求項1ないし請求項5のいずれか1項に記載の半導体装置。
6. 前記第2のMOS型トランジスタがエンハンスメント型、PチャネルMOSトランジスタであることを特徴とする請求項5記載の半導体装置。
7. 前記第2のMOS型トランジスタが略々0Vに等しい閾位を有するPチャネルMOSトランジスタであることを特徴とする請求項5項記載の半導体装置。
8. 前記第2のMOS型トランジスタがデプレッション型PチャネルMOS型トランジスタであることを特徴とする請求項5記載の半導体装置。

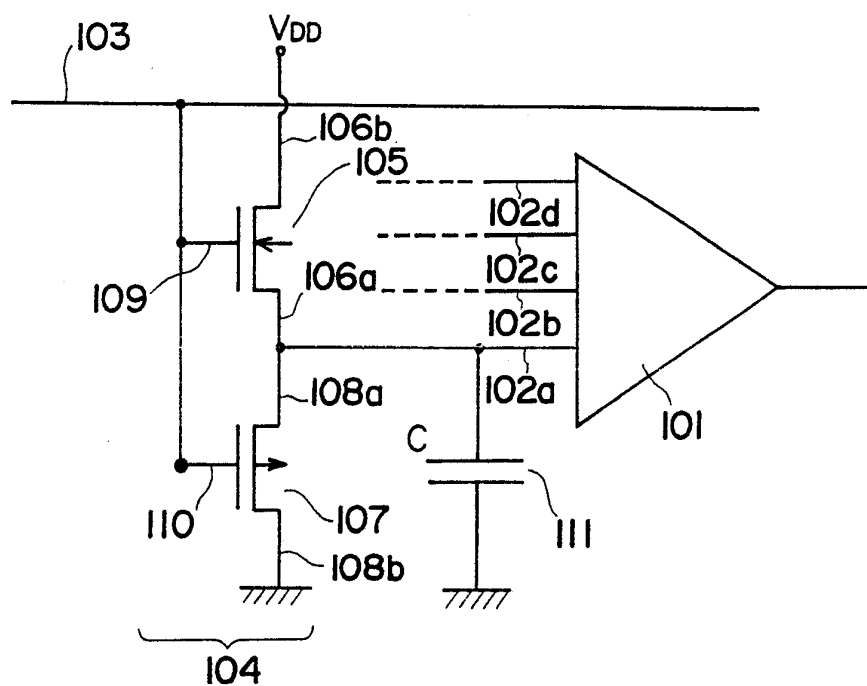
9. 前記第1のMOS型トランジスタの前記フローティングゲートと第3の絶縁膜を介して容量結合する少なくとも1つの第3のゲート電極を有し、前記フローティングゲートと第4の絶縁膜を介して対向する電荷注入用1電極を有し、前記第3のゲート電極並びに前記電荷注入用電極に所定の電位を与えることにより前記フローティングゲートと前記電荷注入用電極の間で前記第4の絶縁膜を通して電荷の授受が行われるよう構成されたことを特徴とする請求項3、請求項5、請求項6、請求項7または請求項8のいずれか1項に記載の半導体装置。
10. 前記第2のMOS型トランジスタがフローティングゲートを有し、かつそのフローティングゲートが前記第1のMOS型トランジスタのフローティングゲートと電氣的に接続され、前記フローティングゲートと第3の絶縁膜を介して容量結合する少なくとも1つの第3のゲート電極を有し、前記フローティングゲートと第4の絶縁膜を介して対向する電荷注入用電極を有し、前記第3のゲート電極並びに前記電荷注入用電極に所定の電位を与えることにより前記フローティングゲートと前記電荷注入用電極の間で前記第4の絶縁膜を通して電荷の授受が行われるよう構成されたことを特徴とする請求項3または請求項5記載の半導体装置。
11. 前記第1の配線が互いに平行に複数本配置され、それと直交する方向に前記第1のゲート電極が互いに平行に複数本配置されたことを特徴とする請求項1ないし請求項10のいずれか1項に記載の半導体装置。
12. 高レベル、中レベル、低レベル3つのレベルの電圧を供給するそれぞれ第1、第2、第3の電源ラインを有し、前記第1の配線が前記第1及び第2の電源ラインの電位をそれぞれ上限、下限とする高・低2つの電位レベルの第1の信号を伝達し、前記第1のMOS型トランジスタがNチャネル型でありそのドレインが前記第2の電源ラインに接続され、前記第1のMOS型トランジスタのゲート電極には、前記第1の信号を逆転した信号が入力されるよう構成したことを特徴とする前記請求項1記載の半導体装置。
13. 前記第1のMOS型トランジスタがフローティングゲートを有するMOSトランジスタであることを特徴とする請求項12記載の半導体装置。

14. 前記第1のMOS型トランジスタのソースがPチャネル型のMOS型トランジスタを介して前記第3の電源ラインに接続されていることを特徴とする前記請求項12ないし請求項14のいずれか1項に記載の半導体装置。

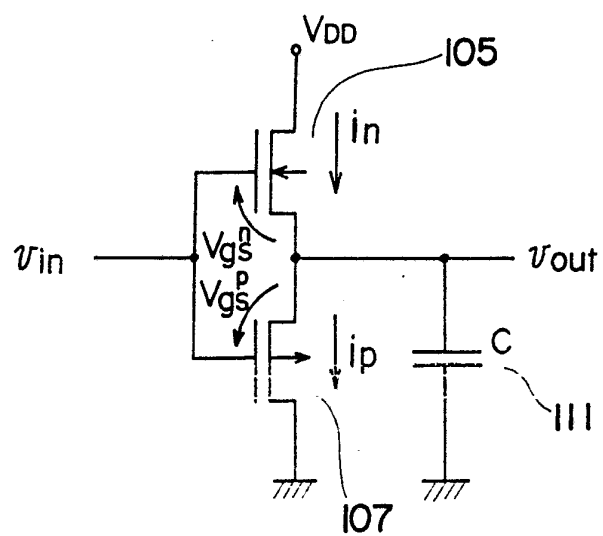
1/31

Fig. 1

(a)

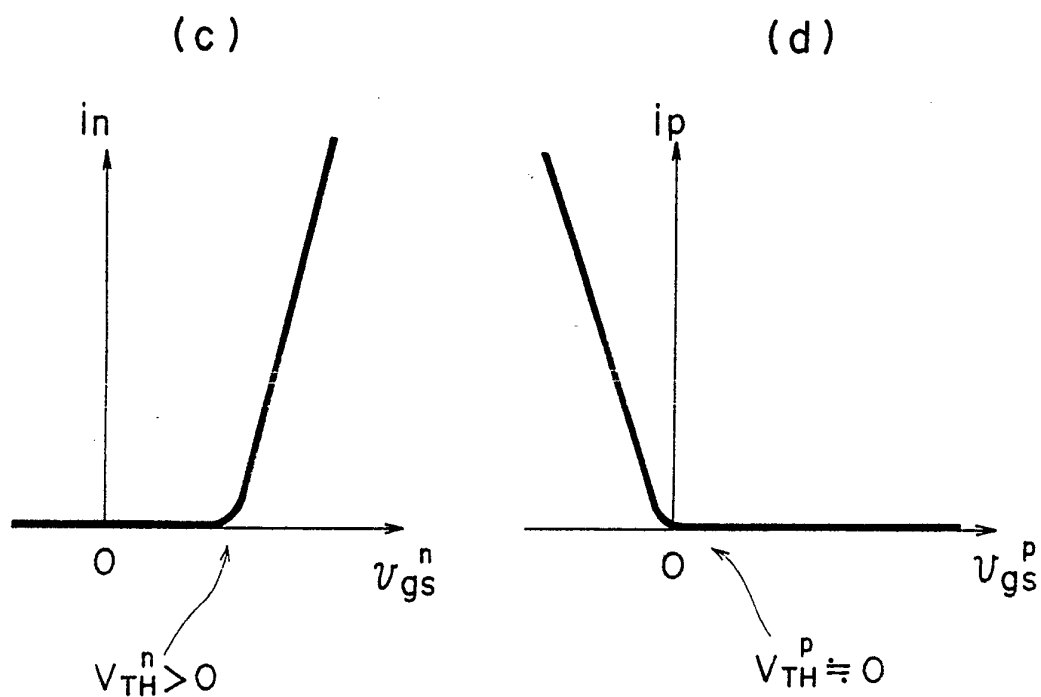


(b)

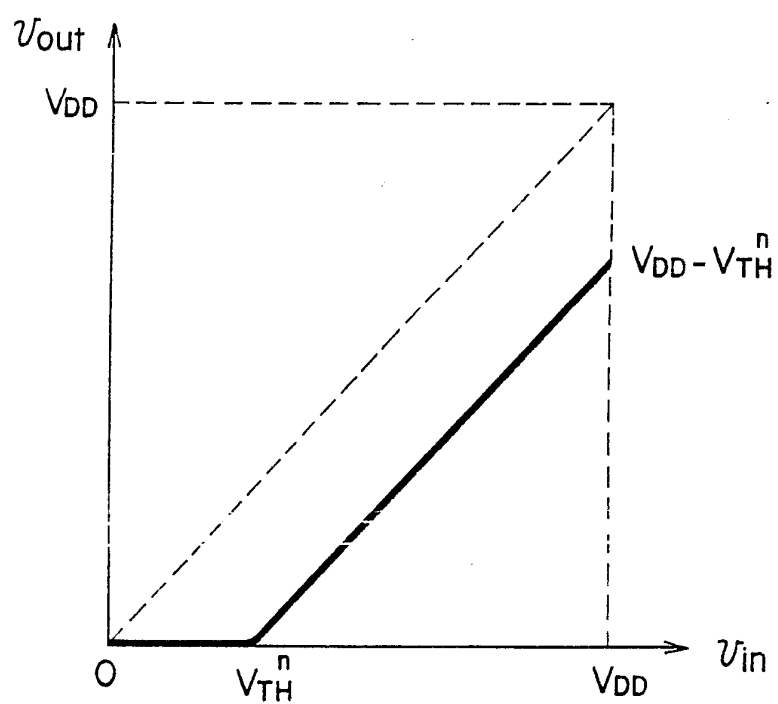


$\frac{2}{31}$

F i g . 1

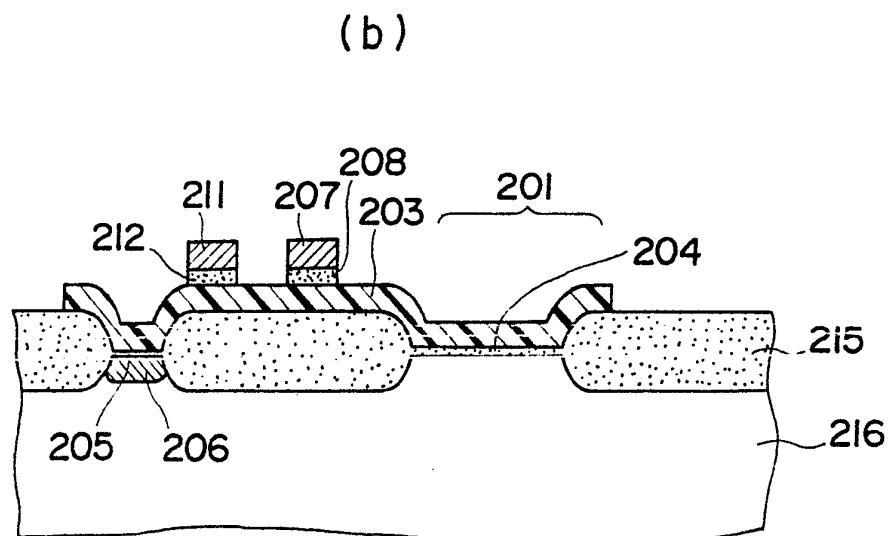
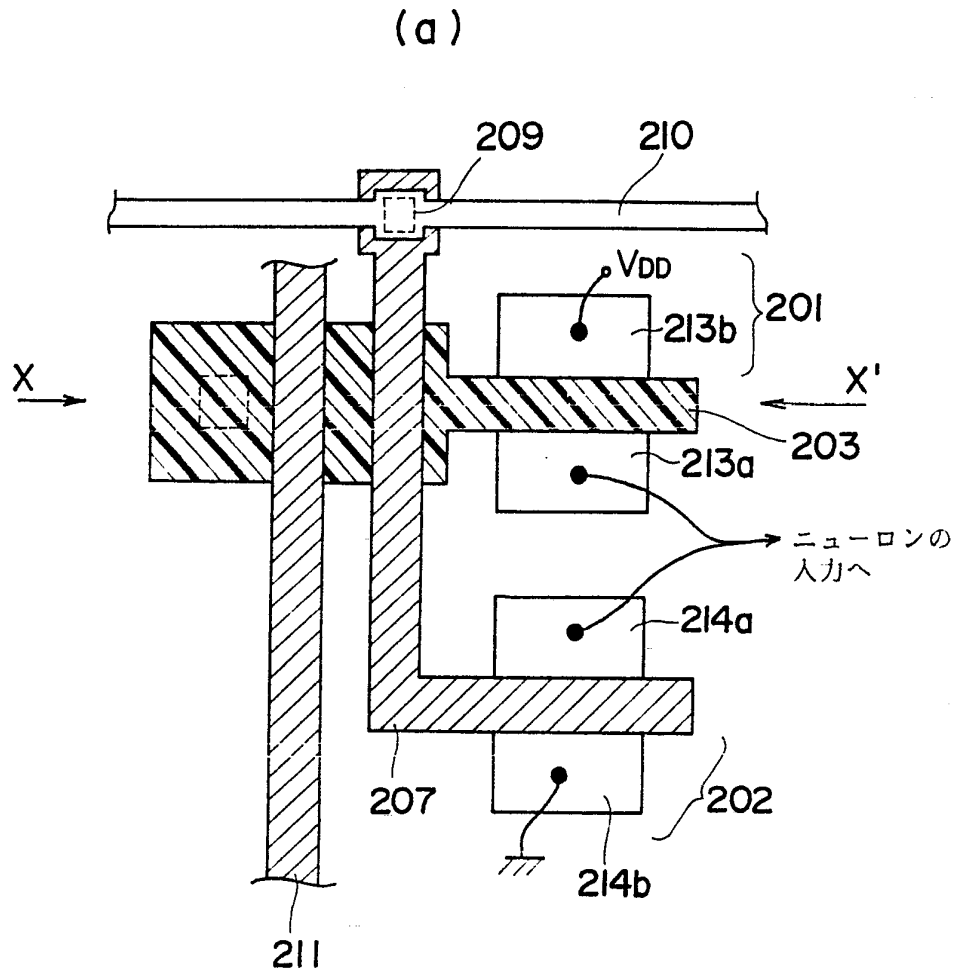


(e)



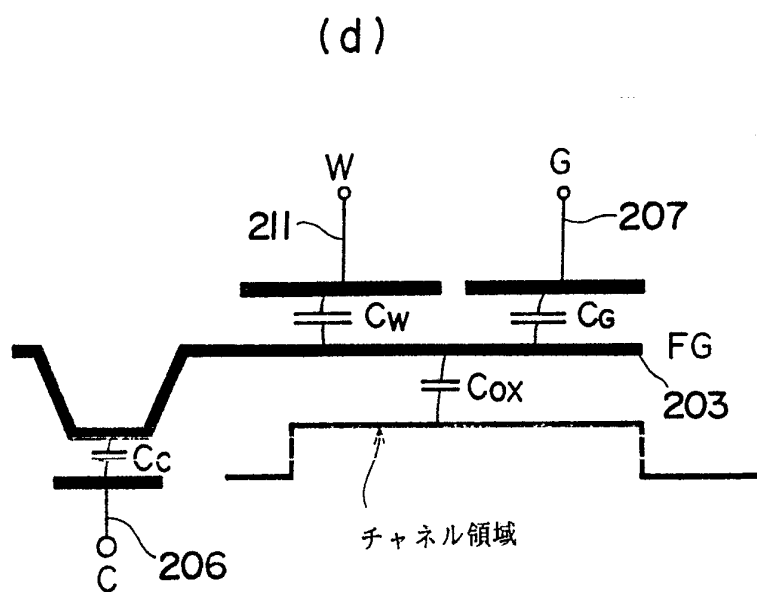
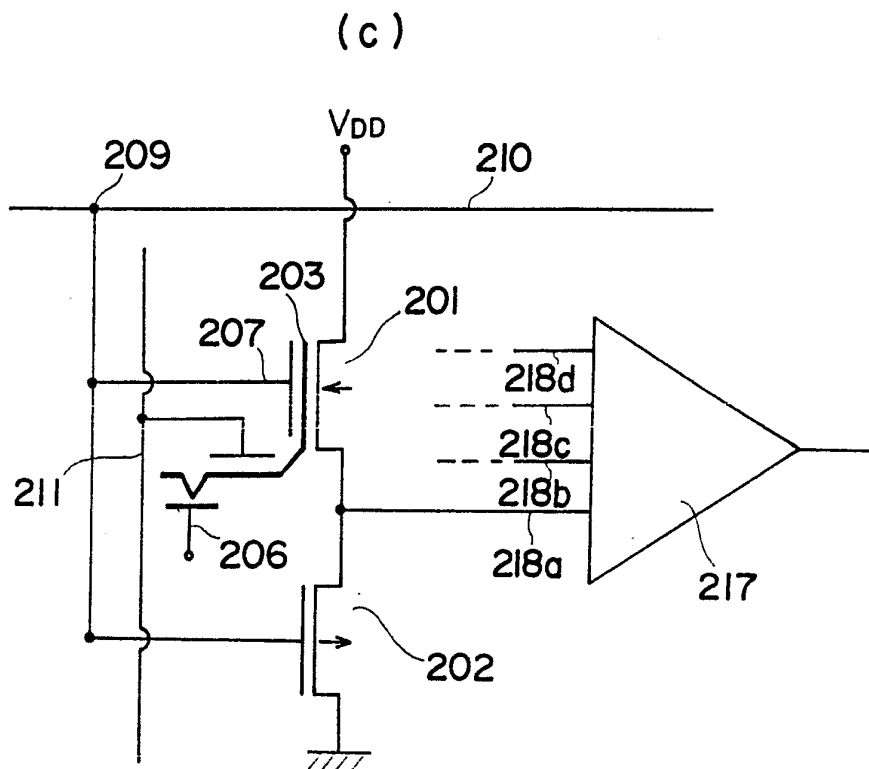
3/31

F i g . 2



4/31

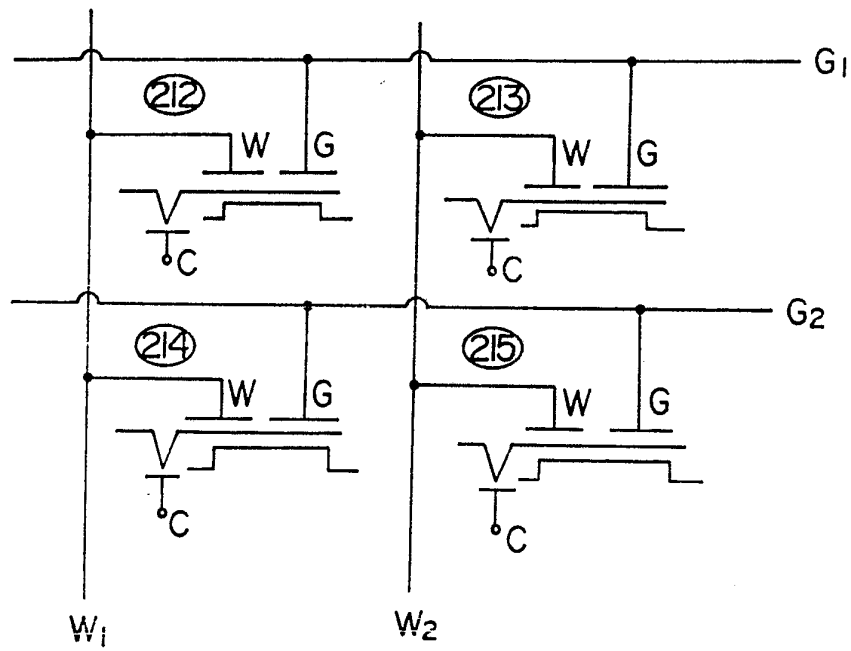
Fig. 2



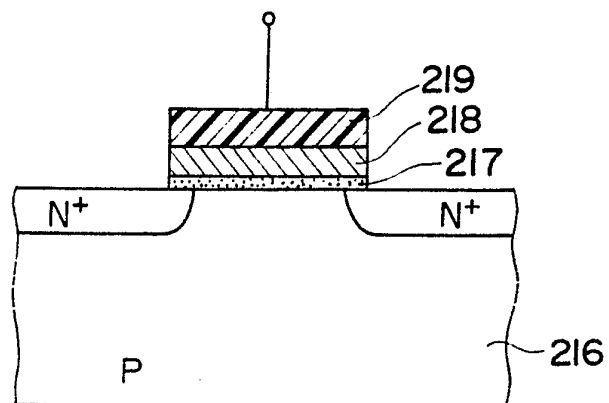
5/31

F i g . 2

(e)



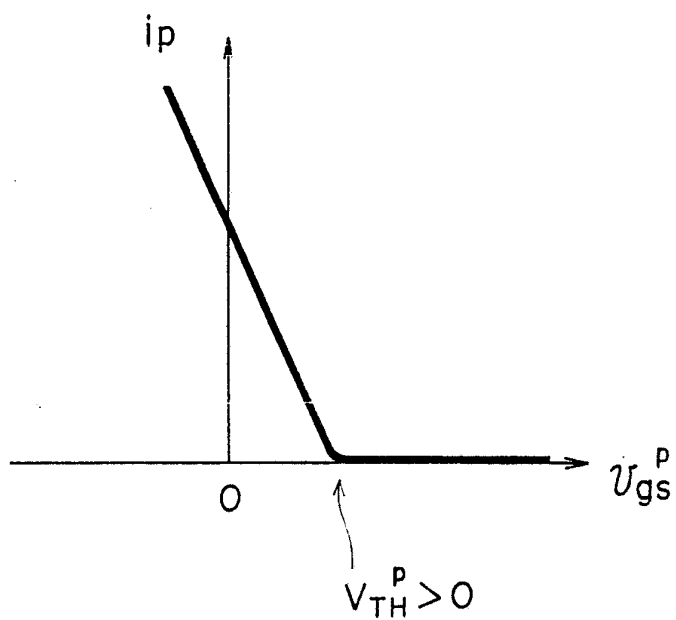
(f)



6/31

F i g . 3

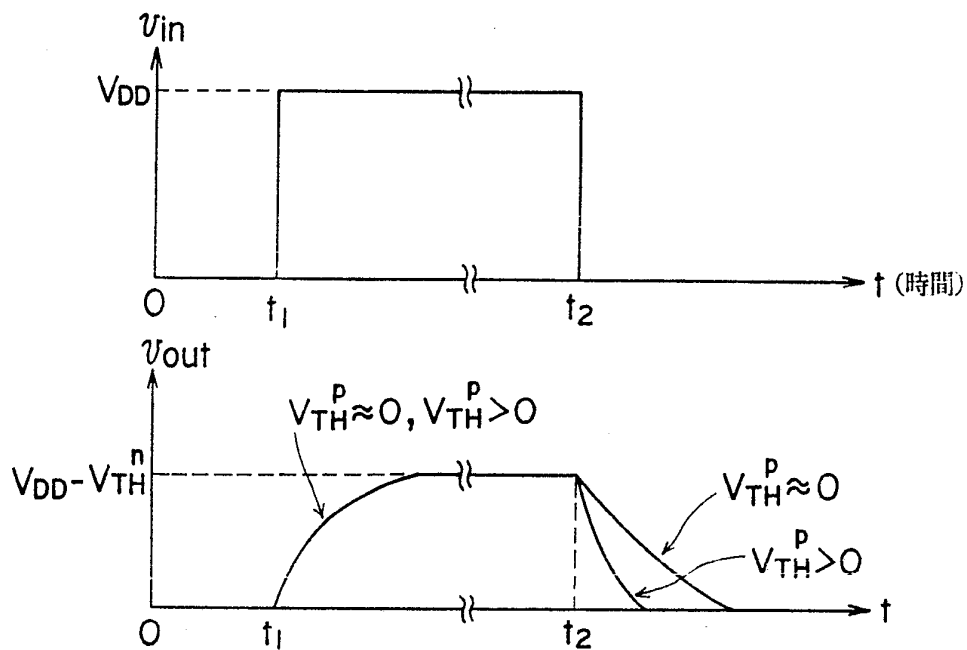
(a)



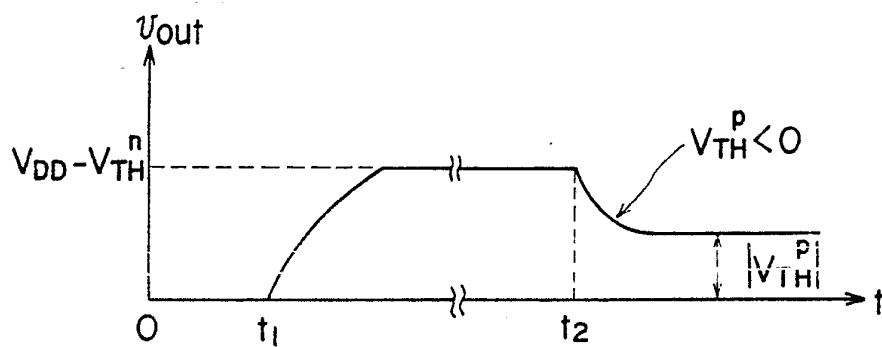
7/31

Fig. 3

(b)



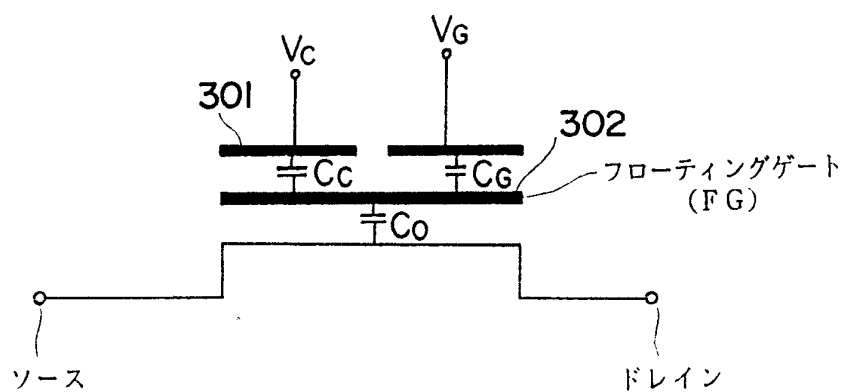
(c)



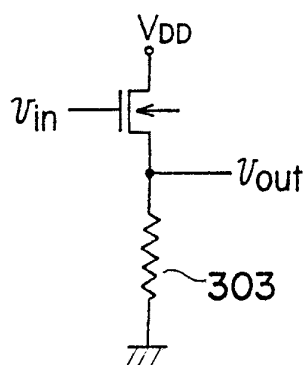
8/31

Fig. 3

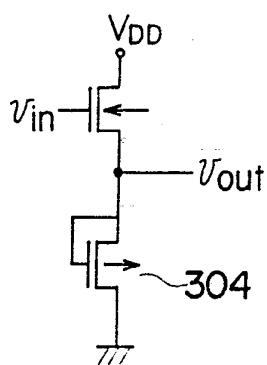
(d)



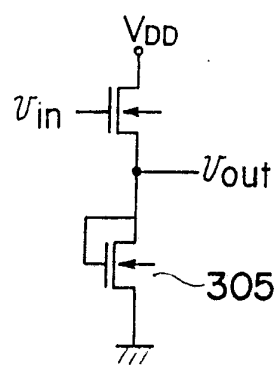
(e)



(f)



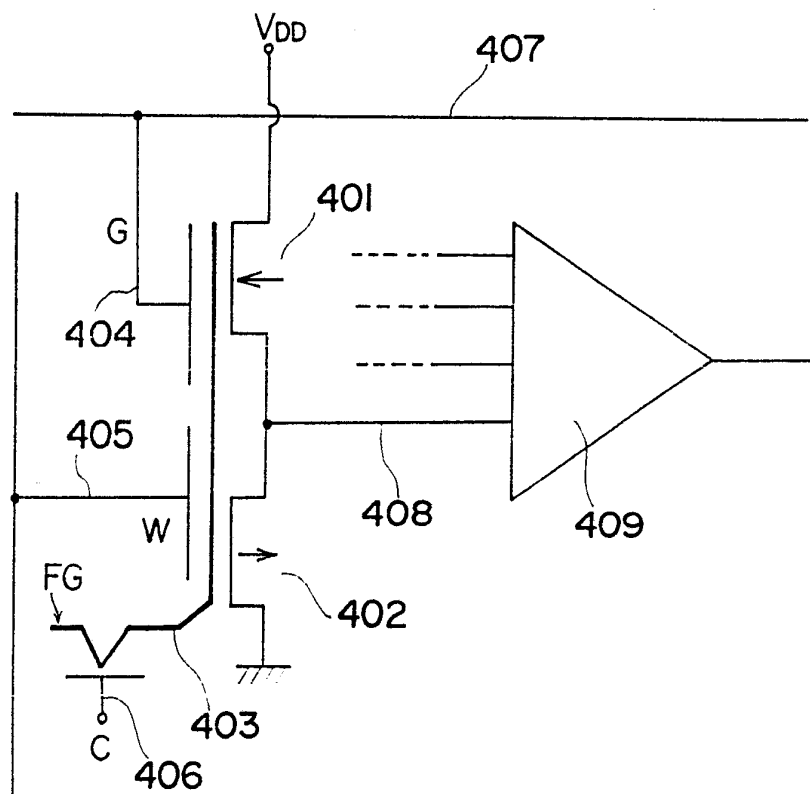
(g)



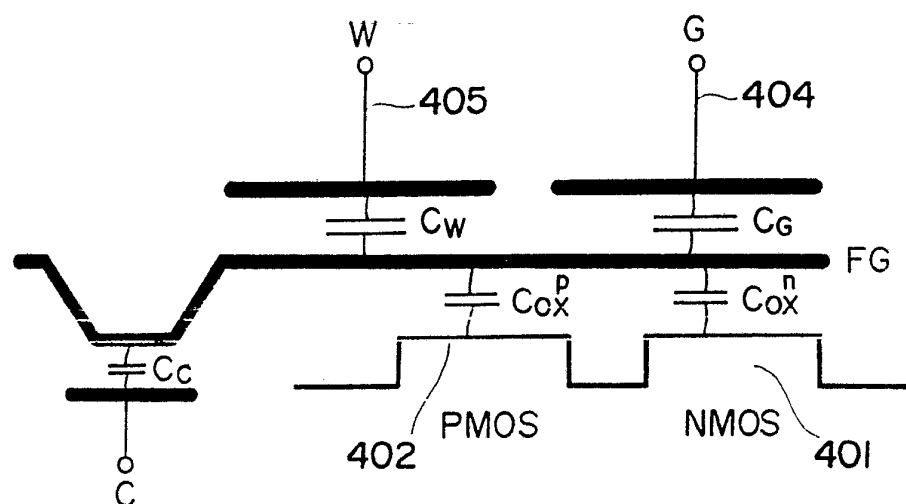
9/31

Fig. 4

(a)



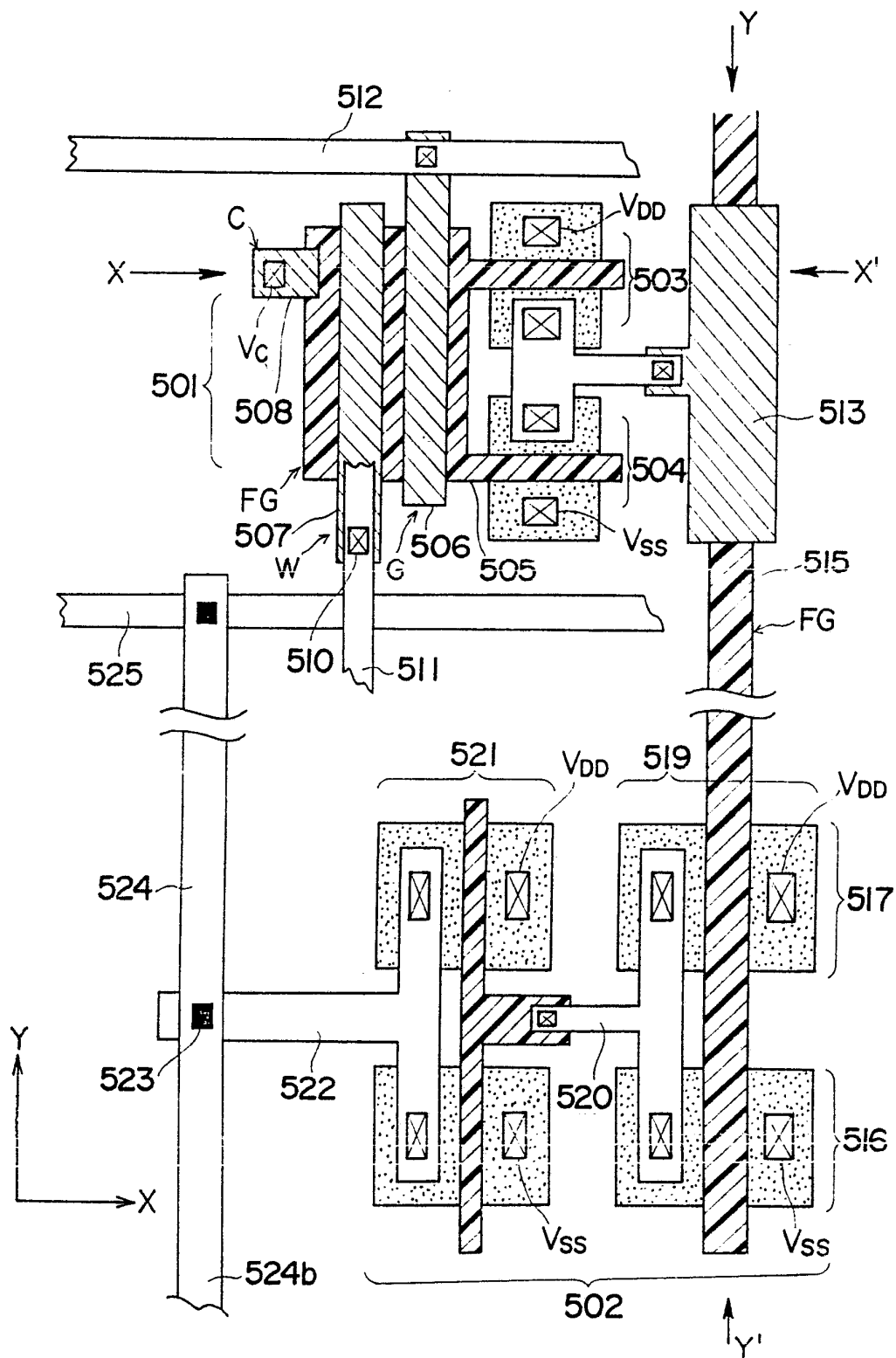
(b)



10/31

Fig. 5

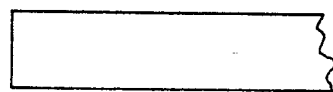
(a)



11/31

Fig. 5

(b)



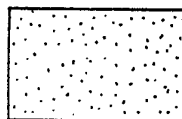
アルミニウム配線



第一層ポリシリコン



第二層ポリシリコン



N⁺ 又は P⁺ 拡散層



コンタクトホール

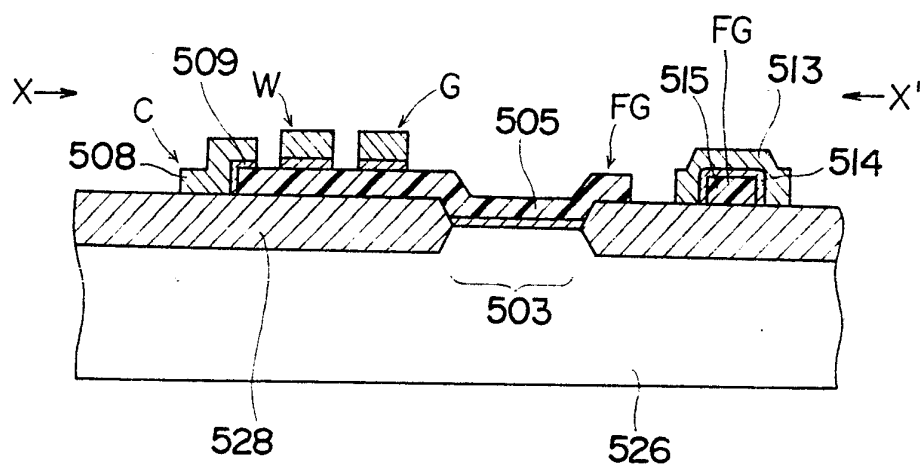


スルーホール

12/31

F i g . 5

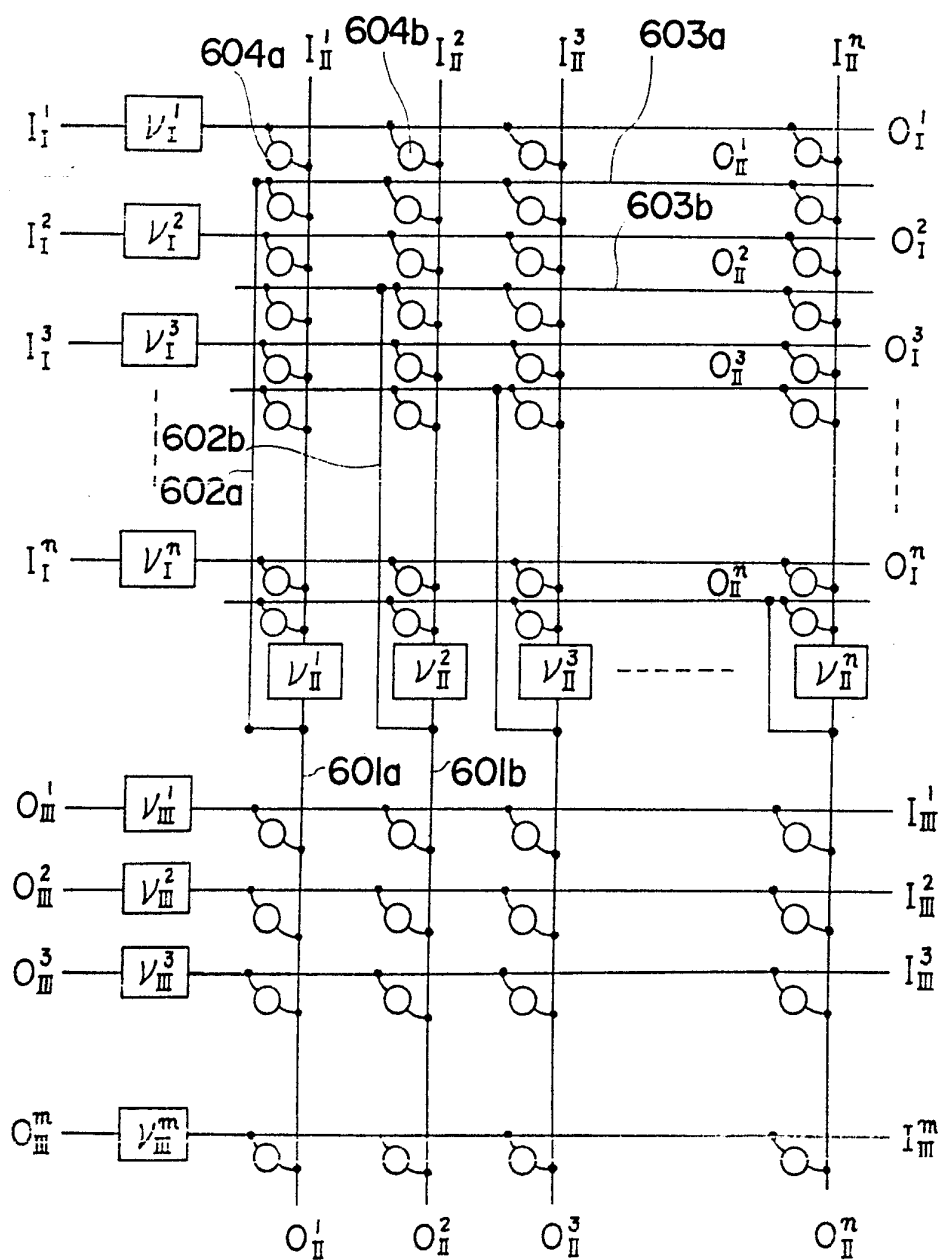
(c)



F i g . 6

14/31

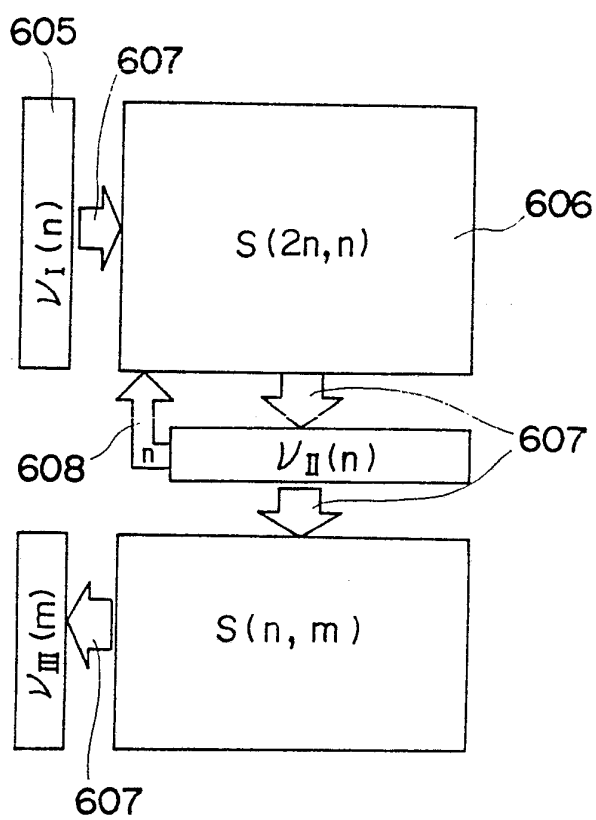
(a)



F i g . 6

15/31

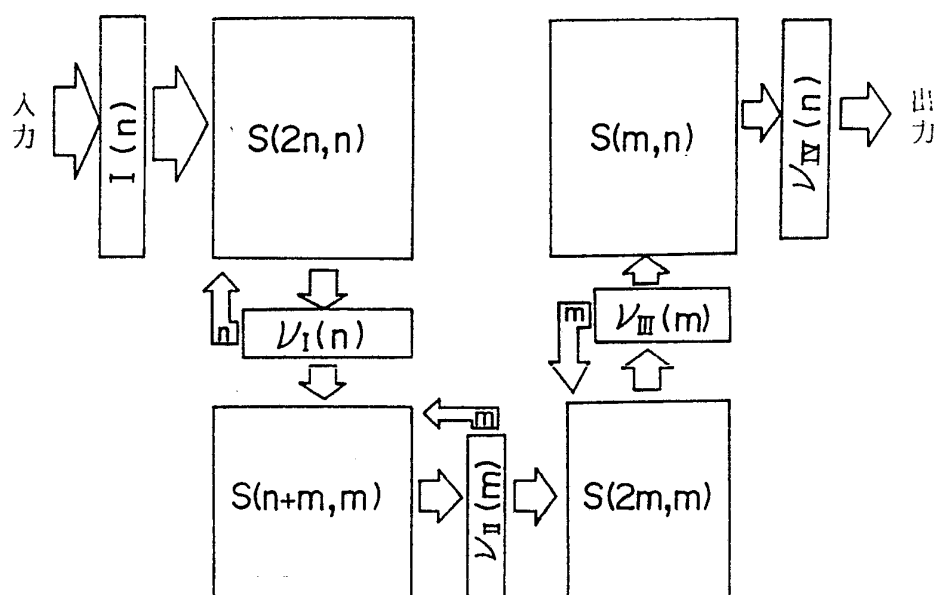
(b)



16/31

F i g . 7

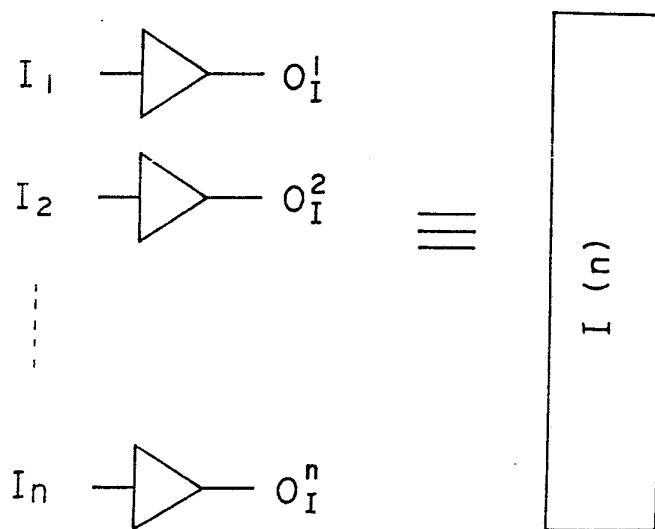
(a)



$17/31$

F i g . 7

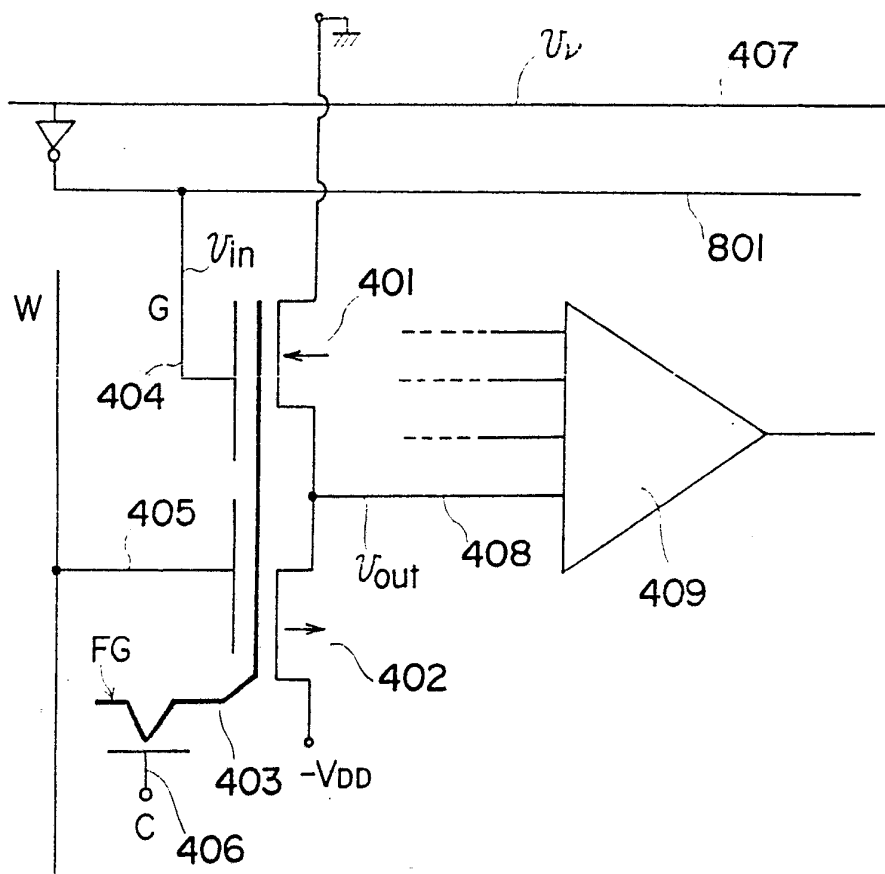
(b)



F i g . 8

18/31

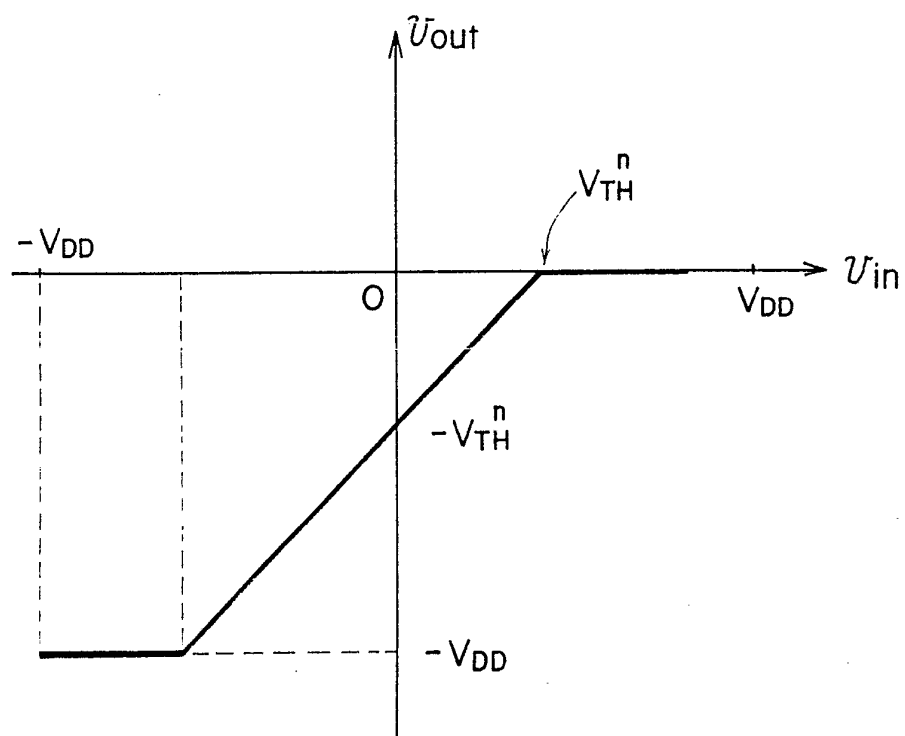
(a)



F i g . 8

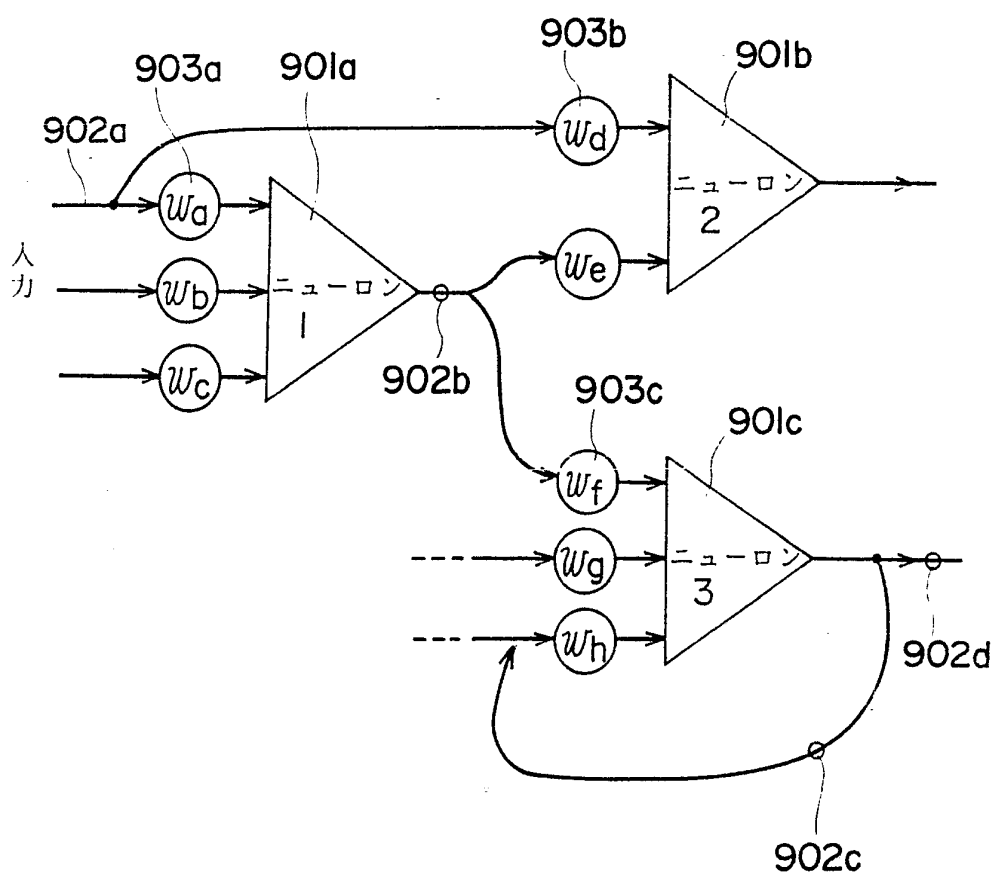
19/31

(b)



20/31

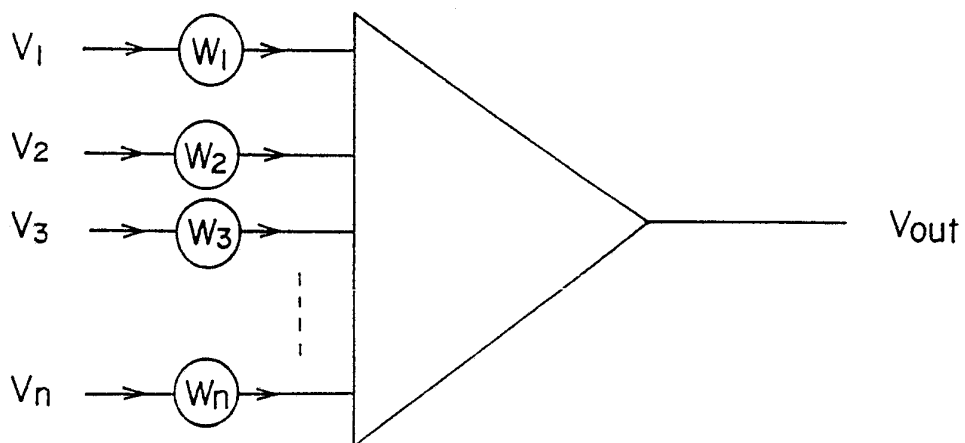
Fig. 9



F i g . 1 0

 $21/31$

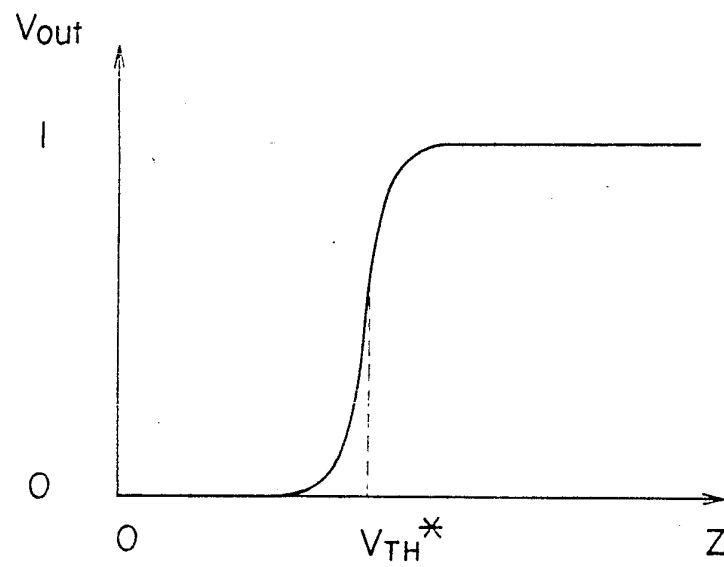
(a)



F i g . 1 0

 $\frac{22}{31}$

(b)

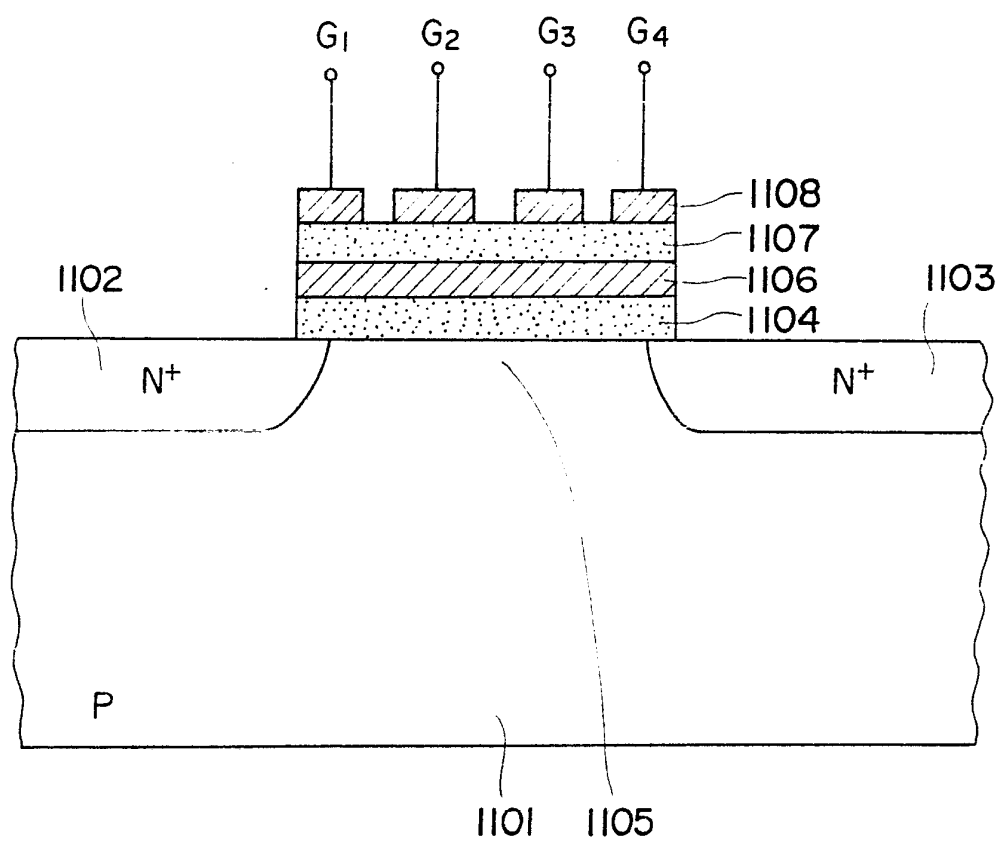


$$Z = \sum_{i=1}^n w_i V_i$$

F i g . 1 1

23/31

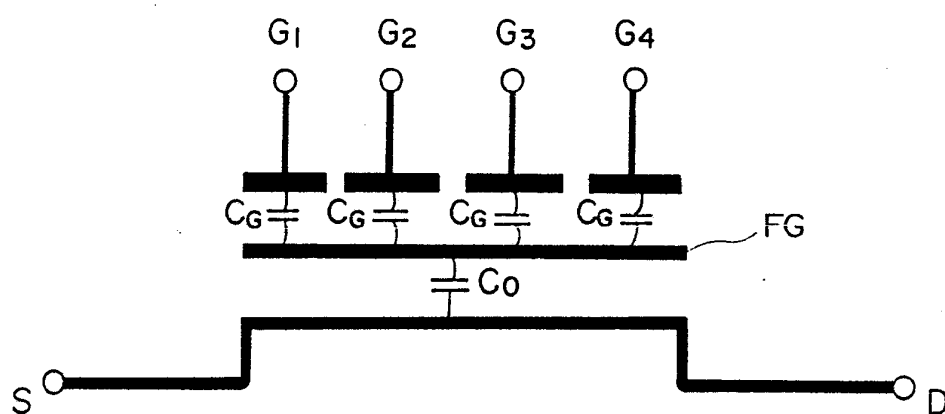
(a)



24/31

Fig. 11

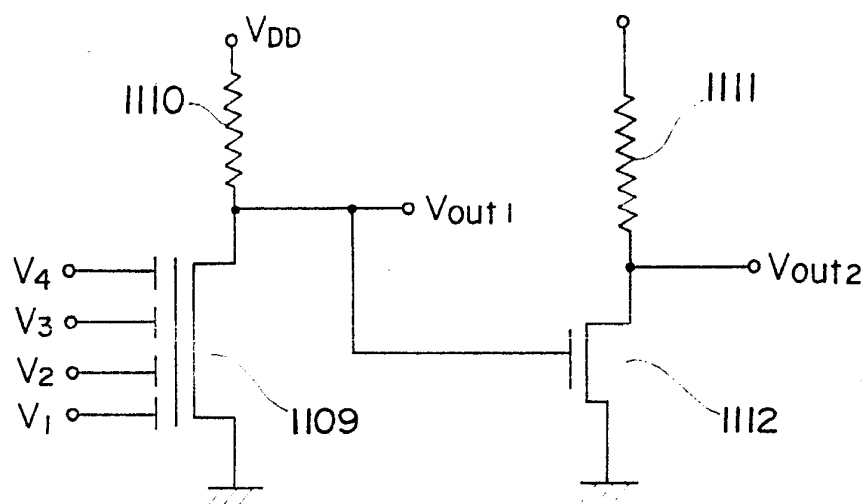
(b)



F i g . 1 1

25/31

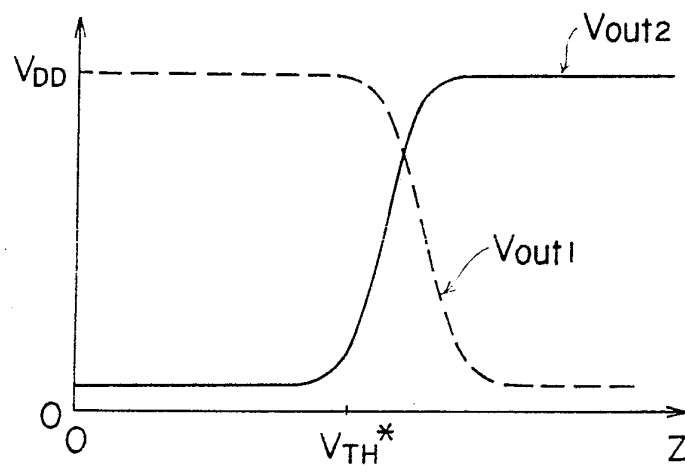
(c)



F i g . 1 1

26/31

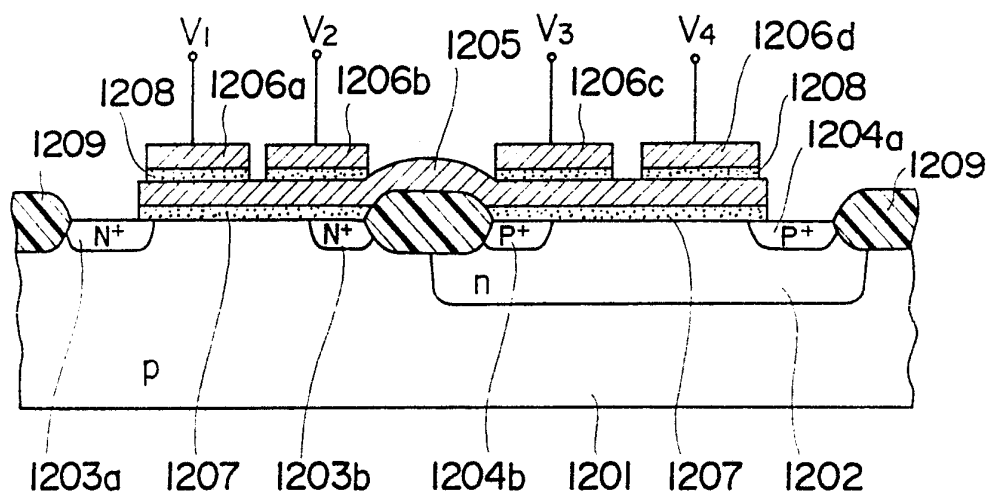
(d)



$$27/31$$

F i g . 1 2

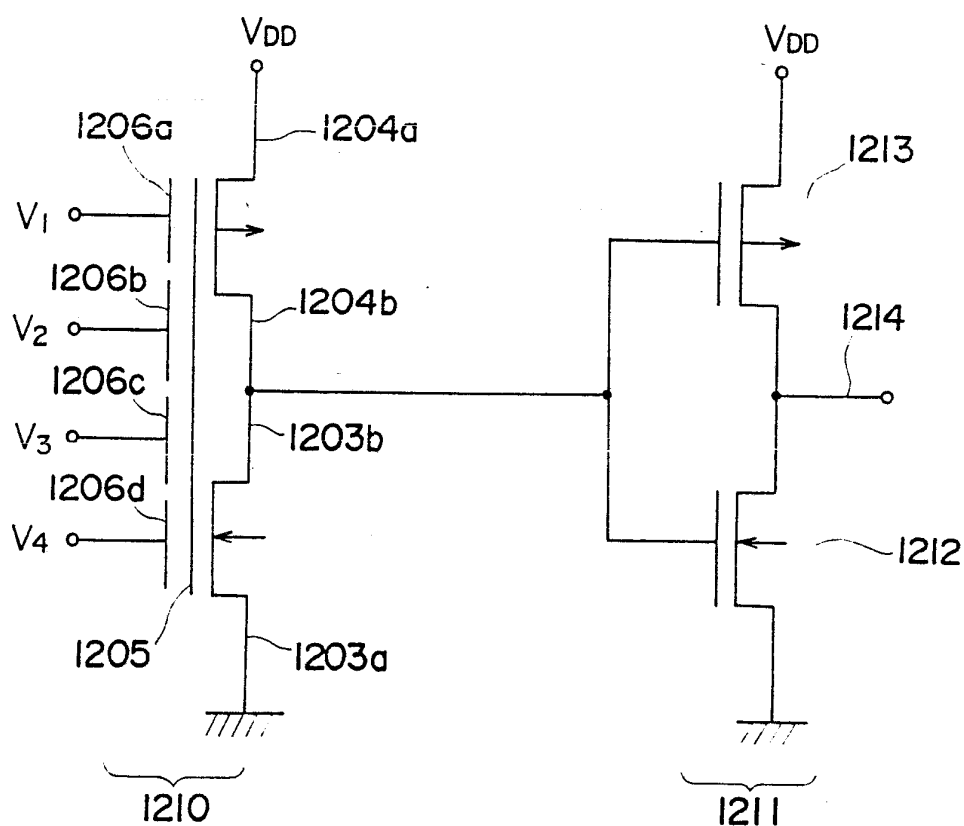
(a)



28/31

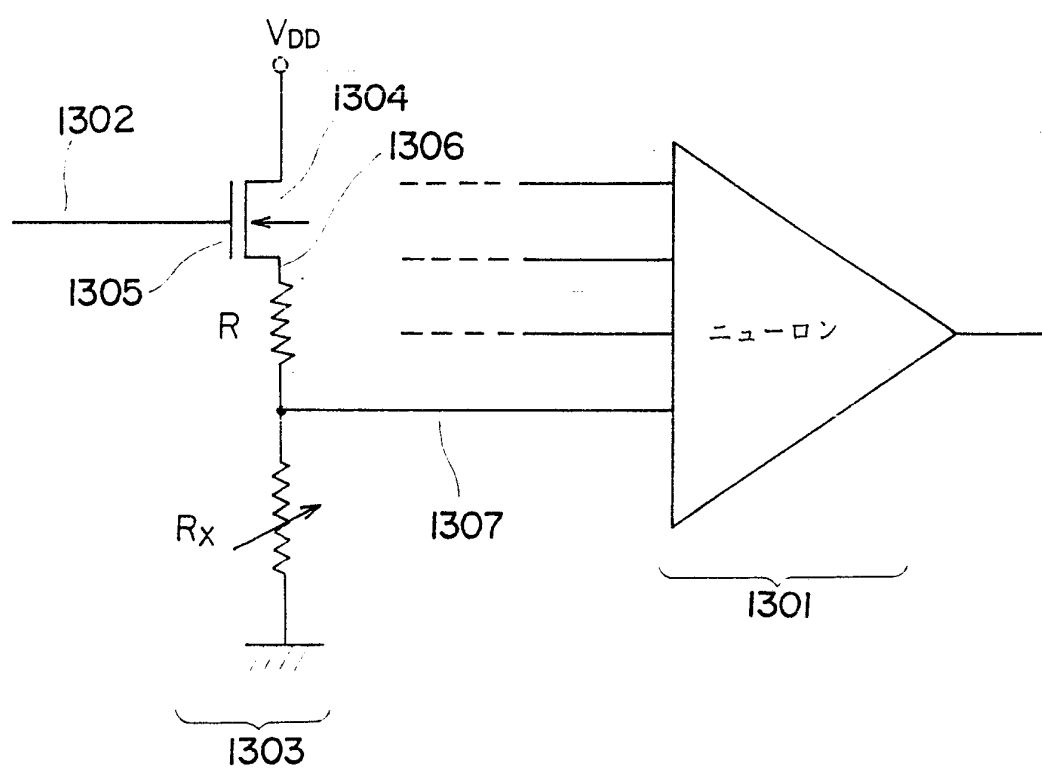
Fig. 12

(b)



29/
31

Fig. 13



30/
31

F i g . 1 4

(a)

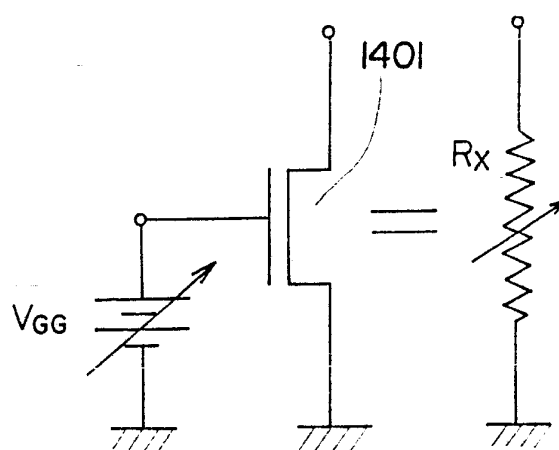
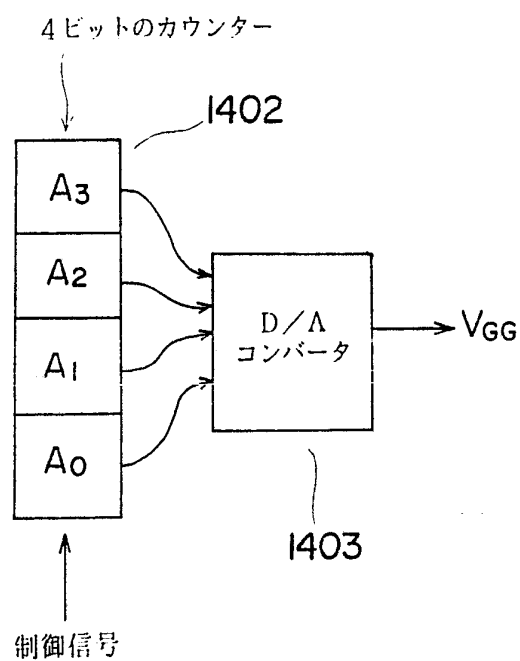


Fig. 14

31/31

(b)



INTERNATIONAL SEARCH REPORT

International Application No PCT/JP92/00014

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) ⁶		
According to International Patent Classification (IPC) or to both National Classification and IPC		
Int. Cl ⁵ G06G7/60		
II. FIELDS SEARCHED		
Minimum Documentation Searched ⁷		
Classification System	Classification Symbols	
IPC	G06G7/60	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched ⁸		
Kokai Jitsuyo Shinan Koho 1972 - 1992		
III. DOCUMENTS CONSIDERED TO BE RELEVANT ⁹		
Category [*]	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No. ¹³
A	JP, A, 2-74053 (Intel Corp.), March 14, 1990 (14. 03. 90), Line 5, lower left column, page 269 to line 18, upper left column, page 271 (Family: none)	1-14
[*] Special categories of cited documents: ¹⁰ "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search		Date of Mailing of this International Search Report
April 7, 1992 (07. 04. 92)		May 12, 1992 (12. 05. 92)
International Searching Authority		Signature of Authorized Officer
Japanese Patent Office		

国 際 調 査 報 告

国際出願番号PCT/JP 9 2 / 0 0 0 1 4

I. 発明の属する分野の分類		
国際特許分類 (IPC) Int. Cl. G 0 6 G 7 / 6 0		
II. 国際調査を行った分野		
調 査 を 行 っ た 最 小 限 資 料		
分 類 体 系	分 類 記 号	
IPC	G 0 6 G 7 / 6 0	
最小限資料以外の資料で調査を行ったもの		
日本国公開実用新案公報 1 9 7 2 - 1 9 9 2 年		
III. 関連する技術に関する文献		
引用文献の カテゴリー※	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	請求の範囲の番号
A	JP, A, 2-74053 (インテル・コーポレーション), 14. 3月. 1990 (14. 03. 90), 第269頁左下欄第5行-第271頁左上欄第18行 (ファミリーなし)	1-14
※引用文献のカテゴリー 「A」特に関連のある文献ではなく、一般的技术水準を示すもの 「E」先行文献ではあるが、国際出願日以後に公表されたもの 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」口頭による開示、使用、展示等に言及する文献 「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 「T」国際出願日又は優先日の後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」同一パテントファミリーの文献		
IV. 認 証		
国際調査を完了した日	国際調査報告の発送日	
07. 04. 92	12.05.92	
国際調査機関	権限のある職員	5 B 7 0 5 2
日本国特許庁 (ISA/JP)	特許庁審査官 森 田 信 一	®