

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
15 September 2011 (15.09.2011)

(10) International Publication Number  
WO 2011/111503 A1

(51) International Patent Classification:

*H01L 21/336* (2006.01) *H01L 27/088* (2006.01)  
*H01L 21/8234* (2006.01) *H01L 29/786* (2006.01)  
*H01L 27/08* (2006.01)

(21) International Application Number:

PCT/JP2011/053589

(22) International Filing Date:

14 February 2011 (14.02.2011)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2010-051008 8 March 2010 (08.03.2010) JP

(71) Applicant (for all designated States except US): SEMICONDUCTOR ENERGY LABORATORY CO., LTD. [JP/JP]; 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP).

(72) Inventors; and

(75) Inventors/Applicants (for US only): YAMAZAKI, Shunpei [JP/JP]; c/o SEMICONDUCTOR ENERGY LABORATORY CO., LTD., 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP). GODO, Hiromichi.

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

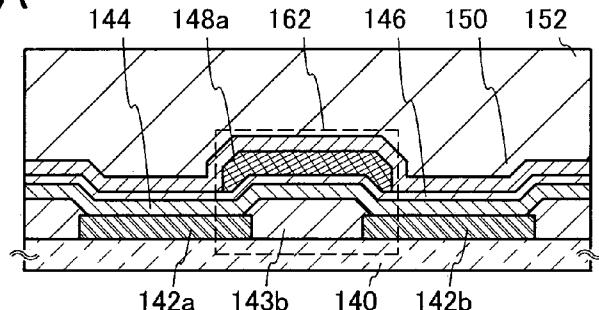
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: SEMICONDUCTOR DEVICE AND METHOD FOR MANUFACTURING SEMICONDUCTOR DEVICE

FIG. 1A



(57) Abstract: Disclosed is a semiconductor device including an insulating layer, a source electrode and a drain electrode embedded in the insulating layer, an oxide semiconductor layer in contact with the insulating layer, the source electrode, and the drain electrode, a gate insulating layer covering the oxide semiconductor layer, and a gate electrode over the gate insulating layer. The upper surface of the surface of the insulating layer, which is in contact with the oxide semiconductor layer, has a root-mean-square (RMS) roughness of 1 nm or less. There is a difference in height between an upper surface of the insulating layer and each of an upper surface of the source electrode and an upper surface of the drain electrode. The difference in height is preferably 5 nm or more. This structure contributes to the suppression of defects of the semiconductor device and enables their miniaturization.

WO 2011/111503 A1

## DESCRIPTION

**SEMICONDUCTOR DEVICE AND METHOD FOR MANUFACTURING  
SEMICONDUCTOR DEVICE**

5

**TECHNICAL FIELD**

[0001]

The present invention relates to a semiconductor device and a manufacturing method thereof. Note that semiconductor devices herein refer to general elements and devices which function by utilizing semiconductor characteristics.

**BACKGROUND ART**

[0002]

There are a wide variety of metal oxides and such metal oxides are used for various applications. Indium oxide is a well-known material and is used as a material for transparent electrodes which are needed for liquid crystal display devices or the like.

[0003]

Some metal oxides have semiconductor characteristics. Examples of such metal oxides having semiconductor characteristics include tungsten oxide, tin oxide, indium oxide, zinc oxide, and the like. A thin film transistor in which a channel formation region is formed using such a metal oxide has been already known (see, for example, Patent Documents 1 to 4, Non-Patent Document 1, and the like).

[0004]

Examples of metal oxides include not only a single-component oxide but also a multi-component oxide. For example,  $\text{InGaO}_3(\text{ZnO})_m$  ( $m$ : natural number) having a homologous phase is known as a multi-component oxide semiconductor including In, Ga, and Zn (see, for example, Non-Patent Documents 2 to 4 and the like).

[0005]

Furthermore, it has been confirmed that an oxide semiconductor including such an In-Ga-Zn-based oxide can also be applied to a channel formation region of a thin film transistor (see, for example, Patent Document 5, Non-Patent Documents 5 and 6, and the like).

## [0006]

In order to achieve high speed operation of a transistor or the like, miniaturization of the transistor is needed. For example, in Patent Document 6, a thin film transistor including an oxide semiconductor used for a channel layer with a thickness of about 10 nm or smaller is disclosed. In Non-Patent Document 7, a thin film transistor including an oxide semiconductor whose channel length is 2  $\mu$ m to 100  $\mu$ m is disclosed.

5

## [References]

## [Patent Documents]

## [0007]

10 [Patent Document 1] Japanese Published Patent Application No. S60-198861

[Patent Document 2] Japanese Published Patent Application No. H8-264794

[Patent Document 3] Japanese Translation of PCT International Application No. H11-505377

[Patent Document 4] Japanese Published Patent Application No. 2000-150900

15 [Patent Document 5] Japanese Published Patent Application No. 2004-103957

[Patent Document 6] Japanese Published Patent Application No. 2010-21170

## [Non-Patent Documents]

## [0008]

[Non-Patent Document 1] M. W. Prins, K. O. Grosse-Holz, G. Muller, J. F. M. Cillessen, J. B. Giesbers, R. P. Weening, and R. M. Wolf, "A ferroelectric transparent thin-film transistor", *Appl. Phys. Lett.*, 17 June 1996, Vol. 68, pp. 3650-3652

[Non-Patent Document 2] M. Nakamura, N. Kimizuka, and T. Mohri, "The Phase Relations in the  $\text{In}_2\text{O}_3$ - $\text{Ga}_2\text{ZnO}_4$ - $\text{ZnO}$  System at 1350 °C", *J. Solid State Chem.*, 1991, Vol. 93, pp. 298-315

25 [Non-Patent Document 3] N. Kimizuka, M. Isobe, and M. Nakamura, "Syntheses and Single-Crystal Data of Homologous Compounds,  $\text{In}_2\text{O}_3(\text{ZnO})_m$  ( $m$  = 3, 4, and 5),  $\text{InGaO}_3(\text{ZnO})_3$ , and  $\text{Ga}_2\text{O}_3(\text{ZnO})_m$  ( $m$  = 7, 8, 9, and 16) in the  $\text{In}_2\text{O}_3$ - $\text{ZnGa}_2\text{O}_4$ - $\text{ZnO}$  System", *J. Solid State Chem.*, 1995, Vol. 116, pp. 170-178

30 [Non-Patent Document 4] M. Nakamura, N. Kimizuka, T. Mohri, and M. Isobe, "Syntheses and crystal structures of new homologous compounds, indium iron zinc oxides ( $\text{InFeO}_3(\text{ZnO})_m$ ) ( $m$ : natural number) and related compounds", *KOTAI BUTSURI*

(*SOLID STATE PHYSICS*), 1993, Vol. 28, No. 5, pp. 317-327

[Non-Patent Document 5] K. Nomura, H. Ohta, K. Ueda, T. Kamiya, M. Hirano, and H. Hosono, "Thin-film transistor fabricated in single-crystalline transparent oxide semiconductor", *SCIENCE*, 2003, Vol. 300, pp. 1269-1272

5 [Non-Patent Document 6] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano, and H. Hosono, "Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors", *NATURE*, 2004, Vol. 432, pp. 488-492

[Non-Patent Document 7] T. Kawamura, H. Uchiyama, S. Saito, H. Wakana, T. Mine, and M. Hatano, "Low-Voltage Operating Amorphous Oxide TFTs", IDW'09, pp. 10 1689-1692

## DISCLOSURE OF INVENTION

[0009]

In the case where a transistor is miniaturized, a defect generated in the manufacturing process becomes a major problem. For example, in a transistor where a semiconductor layer is formed over a wiring functioning as a source or drain electrode, a gate electrode or the like, the wiring has a larger thickness than the semiconductor layer, which causes poor coverage with the semiconductor layer when the thickness of the semiconductor layer is reduced along with miniaturization. As a result, disconnection, 15 defective connection, or the like may occur.

20

[0010]

In the case where a transistor is miniaturized, another problem of a short channel effect arises. The short-channel effect refers to degradation of electrical characteristics which becomes obvious with miniaturization of a transistor (a reduction in channel length 25 ( $L$ )). The short-channel effect results from the effect of an electric field of a drain on a source. Specific examples of the short-channel effect are a decrease in threshold voltage, an increase in S value (subthreshold swing), an increase in leakage current, and the like. The short-channel effect is likely to occur in a transistor including an oxide semiconductor particularly because such a transistor cannot control threshold voltage by 30 doping, unlike a transistor including silicon.

[0011]

In view of this, it is an object of one embodiment of the disclosed invention to

provide a semiconductor device which suppresses a defect and achieves miniaturization. Further, it is another object of one embodiment of the disclosed invention to provide a semiconductor device which maintains favorable characteristics and achieves miniaturization.

5 [0012]

An embodiment of the disclosed invention is a semiconductor device including an insulating layer, a source electrode and a drain electrode embedded in the insulating layer, an oxide semiconductor layer in contact with a part of a surface of the insulating layer, a part of a surface of the source electrode, and a part of a surface of the drain electrode, a gate insulating layer covering the oxide semiconductor layer, and a gate electrode over the gate insulating layer. The part of the surface of the insulating layer, which is in contact with the oxide semiconductor layer, has a root-mean-square (RMS) roughness of 1 nm or less. The difference in height between the part of the surface of the insulating layer and the surface of the source electrode and the difference in height between the part of the surface of the insulating layer and the surface of the drain electrode are 5 nm or more. Note that in the above semiconductor device, the difference in height between the part of the surface of the insulating layer and the surface of the source electrode and the difference in height between the part of the surface of the insulating layer and the surface of the drain electrode may be 20 nm or less.

20 [0013]

Note that in this specification, the root-mean-square (RMS) roughness is obtained by three-dimensionally expanding the RMS roughness of a cross section curve so as to be able to apply it to the measurement surface. The RMS roughness is represented by the square root of the mean value of the square of the deviation from the reference surface to the specific surface, and is obtained by the following formula.

25 [0014]

[Formula 1]

$$R_{ms} = \sqrt{\frac{1}{S_0} \int_{Y_1}^{Y_2} \int_{X_1}^{X_2} \{F(X, Y) - Z_0\}^2 dX dY}$$

[0015]

Note that the measurement surface is a surface which is shown by all the measurement data, and is represented by the following formula.

[0016]

5 [Formula 2]

$$Z = F(X, Y)$$

[0017]

The specific surface is a surface which is an object of roughness measurement, and is a rectangular region which is surrounded by four points represented by the 10 coordinates  $(X_1, Y_1)$ ,  $(X_1, Y_2)$ ,  $(X_2, Y_1)$ , and  $(X_2, Y_2)$ . The area of the specific surface when the specific surface is flat ideally is denoted by  $S_0$ . Note that  $S_0$  can be obtained by the following formula.

[0018]

[Formula 3]

$$S_0 = |X_2 - X_1| \cdot |Y_2 - Y_1|$$

15

[0019]

In addition, the reference surface refers to a surface parallel to an X-Y surface at the average height of the specific surface. In short, when the average value of the height of the specific surface is denoted by  $Z_0$ , the height of the reference surface is also denoted 20 by  $Z_0$ . Note that  $Z_0$  can be obtained by the following formula.

[0020]

[Formula 4]

$$Z_0 = \frac{1}{S_0} \int_{Y_1}^{Y_2} \int_{X_1}^{X_2} F(X, Y) dX dY$$

[0021]

25 Note that in this specification, the root-mean-square (RMS) roughness is calculated in a region of  $10 \text{ nm} \times 10 \text{ nm}$ , preferably  $100 \text{ nm} \times 100 \text{ nm}$ , more preferably  $1 \mu\text{m} \times 1 \mu\text{m}$  from an AFM image obtained using an atomic force microscope (AFM).

## [0022]

Another embodiment of the disclosed invention is a semiconductor device including a first transistor, and a second transistor over the first transistor. The first transistor includes a first channel formation region, a first gate insulating layer provided over the first channel formation region, a first gate electrode provided over the first gate insulating layer so as to overlap with the first channel formation region, and a first source electrode and a first drain electrode electrically connected to the first channel formation region. The second transistor includes a second source electrode and a second drain electrode embedded in an insulating layer, a second channel formation region in contact with a part of a surface of the insulating layer, a part of a surface of the second source electrode, and a part of a surface of the second drain electrode, a second gate insulating layer covering the second channel formation region, and a second gate electrode over the second gate insulating layer. The part of the surface of the insulating layer, which is in contact with the second channel formation region, has a root-mean-square roughness of 1 nm or less. The difference in height between the part of the surface of the insulating layer and the surface of the second source electrode and the difference in height between the part of the surface of the insulating layer and the surface of the second drain electrode are 5 nm or more. Note that in the above semiconductor device, the difference in height between the part of the surface of the insulating layer and the surface of the second source electrode and the difference in height between the part of the surface of the insulating layer and the surface of the second drain electrode may be 20 nm or less.

## [0023]

Another embodiment of the disclosed invention is a method for manufacturing a semiconductor device, including the steps of: forming a source electrode and a drain electrode over a surface with a root-mean-square roughness of 1 nm or less; forming an insulating layer so as to cover the source electrode and the drain electrode; performing planarization treatment of a surface of the insulating layer, thereby forming a planarized insulating layer having a surface with a root-mean-square roughness of 1 nm or less; forming an opening reaching the source electrode and the drain electrode in the planarized insulating layer, so that the difference in height between a part of the surface of the planarized insulating layer and a surface of the source electrode and the

difference in height between the part of the surface of the planarized insulating layer and a surface of the drain electrode are 5 nm or more; forming an oxide semiconductor layer in contact with the part of the surface of the planarized insulating layer, a part of the surface of the source electrode, and a part of the surface of the drain electrode;

5 forming a gate insulating layer so as to cover the oxide semiconductor layer; and forming a gate electrode over the gate insulating layer.

[0024]

Another embodiment is a method for manufacturing a semiconductor device, including the steps of: forming a source electrode and a drain electrode over a surface with a root-mean-square roughness of 1 nm or less; forming an insulating layer so as to cover the source electrode and the drain electrode; performing planarization treatment of a surface of the insulating layer, thereby forming a planarized insulating layer having a surface with a root-mean-square roughness of 1 nm or less and exposing the source electrode and the drain electrode; thinning the source electrode and the drain electrode so that the difference in height between a part of the surface of the planarized insulating layer and a surface of the source electrode and the difference in height between the part of the surface of the planarized insulating layer and a surface of the drain electrode are 5 nm or more; forming an oxide semiconductor layer in contact with the part of the surface of the planarized insulating layer, a part of the surface of the source electrode,

15 and a part of the surface of the drain electrode; forming a gate insulating layer so as to cover the oxide semiconductor layer; and forming a gate electrode over the gate insulating layer. Note that in the above method for manufacturing a semiconductor device, the difference in height between the part of the surface of the planarized insulating layer and the surface of the source electrode and the difference in height

20 between the part of the surface of the planarized insulating layer and the surface of the drain electrode may be 20 nm or less.

25

[0025]

Another embodiment of the disclosed invention is a method for manufacturing a semiconductor device, including the steps of: forming a first transistor including a channel formation region, a first gate insulating layer over the channel formation region, a first gate electrode over the first gate insulating layer, overlapping with the channel formation region, and a first source electrode and a first drain electrode electrically

connected to the channel formation region; forming a first insulating layer having a surface with a root-mean-square roughness of 1 nm or less so as to cover the first transistor; forming a second source electrode and a second drain electrode over the surface of the first insulating layer; forming a second insulating layer so as to cover the second source electrode and the second drain electrode; performing planarization treatment of a surface of the second insulating layer, thereby forming a planarized insulating layer having a surface with a root-mean-square roughness of 1 nm or less; forming an opening reaching the second source electrode and the second drain electrode in the planarized insulating layer, so that the difference in height between a part of the surface of the planarized insulating layer and a surface of the second source electrode and the difference in height between the part of the surface of the planarized insulating layer and a surface of the second drain electrode are 5 nm or more; forming an oxide semiconductor layer in contact with the part of the surface of the planarized insulating layer, a part of the surface of the second source electrode, and a part of the surface of the second drain electrode; forming a second gate insulating layer so as to cover the oxide semiconductor layer; and forming a second gate electrode over the second gate insulating layer.

[0026]

Another embodiment is a method for manufacturing a semiconductor device, including the steps of: forming a first transistor including a channel formation region, a first gate insulating layer over the channel formation region, a first gate electrode over the first gate insulating layer, overlapping with the channel formation region, and a first source electrode and a first drain electrode electrically connected to the channel formation region; forming a first insulating layer having a surface with a root-mean-square roughness of 1 nm or less so as to cover the first transistor; forming a second source electrode and a second drain electrode over the surface of the first insulating layer; forming a second insulating layer so as to cover the second source electrode and the second drain electrode; performing planarization treatment of a surface of the second insulating layer, thereby forming a planarized insulating layer having a surface with a root-mean-square roughness of 1 nm or less and exposing the second source electrode and the second drain electrode; thinning the second source electrode and the second drain electrode so that the difference in height between a part

of the surface of the planarized insulating layer and a surface of the second source electrode and the difference in height between the part of the surface of the planarized insulating layer and a surface of the second drain electrode are 5 nm or more; forming an oxide semiconductor layer in contact with the part of the surface of the planarized insulating layer, a part of the surface of the second source electrode, and a part of the surface of the second drain electrode; forming a second gate insulating layer so as to cover the oxide semiconductor layer; and forming a second gate electrode over the second gate insulating layer. Note that in the above method for manufacturing a semiconductor device, the difference in height between the part of the surface of the planarized insulating layer and the surface of the second source electrode and the difference in height between the part of the surface of the planarized insulating layer and the surface of the second drain electrode may be 20 nm or less.

[0027]

Note that the channel length  $L$  of the transistor is preferably less than 2  $\mu\text{m}$ , further preferably, 10 nm to 350 nm (0.35  $\mu\text{m}$ ). The thickness of the oxide semiconductor layer is in the range of 1 nm to 50 nm, preferably, 2 nm to 20 nm, further preferably, 3 nm to 15 nm. With such a structure, a semiconductor device which operates at high speed and consumes less power can be achieved. For the gate insulating layer, a high dielectric constant material such as hafnium oxide is used. For example, the relative permittivity of hafnium oxide is approximately 15, which is much higher than that of silicon oxide which is 3 to 4. With such a material, a gate insulating layer where the equivalent oxide thickness is less than 15 nm, preferably 2 nm to 10 nm, can be easily formed. In other words, the semiconductor device can be easily miniaturized. Further, as the oxide semiconductor layer, an intrinsic oxide semiconductor which is purified is used. With such an oxide semiconductor, the carrier density of the oxide semiconductor layer can be, for example, less than  $1 \times 10^{12}/\text{cm}^3$ , preferably, less than  $1.45 \times 10^{10}/\text{cm}^3$ , the off-state current of the transistor can be 100  $\text{zA}/\mu\text{m}$  (1  $\text{zA}$  (zeptoampere) is  $1 \times 10^{-21} \text{ A}$ ) or less, preferably, 10  $\text{zA}/\mu\text{m}$  or less, and the S value of the transistor can be 65 mV/dec or less, preferably, less than 63 mV/dec.

When the above structure is employed, the off-state current of the transistor can be  $1 \times 10^{-24} \text{ A}/\mu\text{m}$  to  $1 \times 10^{-30} \text{ A}/\mu\text{m}$  in theory. The gate electrode may be provided to

overlap with the source electrode and the drain electrode, and alternatively, only an end portion of the gate electrode may be provided to overlap with an end portion of the source electrode and an end portion of the drain electrode.

[0028]

5 Note that semiconductor devices herein refer to general devices which function by utilizing semiconductor characteristics. For example, a display device, a memory device, an integrated circuit, and the like are included in the category of the semiconductor device.

[0029]

10 Note that the term such as “over” or “below” in this specification does not necessarily mean that a component is placed “directly on” or “directly under” another component. For example, the expression “a gate electrode over a gate insulating layer” does not exclude the case where a component is placed between the gate insulating layer and the gate electrode.

15 [0030]

In addition, the term such as “electrode” or “wiring” in this specification does not limit a function of a component. For example, an “electrode” can be used as part of a “wiring”, and the “wiring” can be used as part of the “electrode”. Furthermore, the term “electrode” or “wiring” can include the case where a plurality of “electrodes” or 20 “wirings” is formed in an integrated manner.

[0031]

Functions of a “source” and a “drain” are sometimes interchanged with each other when a transistor of opposite polarity is used or when the direction of current flowing is changed in circuit operation, for example. Therefore, the terms “source” and 25 “drain” can be used to denote the drain and the source, respectively, in this specification.

[0032]

Note that the term “electrically connected” in this specification includes the case where components are connected through an “object having any electric function”. There is no particular limitation on an object having any electric function as long as 30 electric signals can be transmitted and received between components that are connected through the object. Examples of an “object having any electric function” are a switching element such as a transistor, a resistor, an inductor, a capacitor, and an element.

with a variety of functions, as well as an electrode and a wiring.

[0033]

According to one embodiment of the disclosed invention, a problem caused by miniaturization can be resolved. As a result, the size of the transistor can be sufficiently reduced. By sufficiently downsizing the transistor, an area of a semiconductor device including the transistor is decreased, and the number of semiconductor devices manufactured with one substrate is increased. Therefore, the manufacturing cost per semiconductor device is decreased. Further, since the semiconductor device is downsized, a semiconductor device with a size similar to that of the conventional semiconductor device can have improved functions. Moreover, advantageous effects such as high-speed operation and low power consumption due to a reduction in channel length can be obtained. That is, according to one embodiment of the disclosed invention, miniaturization of a transistor including an oxide semiconductor is achieved, so that a variety of advantageous effects accompanying therewith can be obtained.

[0034]

As described above, according to one embodiment of the disclosed invention, a semiconductor device which achieves miniaturization as well as suppressing defect or maintaining favorable characteristics can be provided.

## 20 BRIEF DESCRIPTION OF DRAWINGS

[0035]

FIGS. 1A to 1C are cross-sectional views each illustrating an example of a structure of a semiconductor device.

FIGS. 2A to 2G are cross-sectional views illustrating a manufacturing process 25 of a semiconductor device.

FIGS. 3A to 3G are cross-sectional views illustrating a manufacturing process of a semiconductor device.

FIGS. 4A to 4C are a cross-sectional view, a plan view, and a circuit diagram illustrating an example of a structure of a semiconductor device.

30 FIGS. 5A to 5D are cross-sectional views illustrating a manufacturing process of a semiconductor device.

FIGS. 6A to 6C are cross-sectional views illustrating a manufacturing process

of a semiconductor device.

FIGS. 7A-1, 7A-2, and 7B are diagrams illustrating an example of application of a semiconductor device.

FIGS. 8A and 8B are diagrams illustrating an example of application of a semiconductor device.

FIGS. 9A to 9C are diagrams illustrating an example of application of a semiconductor device.

FIG. 10 is a diagram illustrating an example of application of a semiconductor device.

FIGS. 11A and 11B are diagrams illustrating an example of application of a semiconductor device.

FIGS. 12A to 12F are diagrams each illustrating an electronic device including a semiconductor device.

FIGS. 13A and 13B are diagrams each illustrating a model used for calculation.

FIG. 14 is a diagram showing the relationship between gate voltage  $VG$  (V) and drain current  $ID$  (A).

FIG. 15 is a diagram showing the relationship between channel length  $L$  (nm) and threshold voltage  $Vth$  (V).

FIG. 16 is a diagram showing the relationship between channel length  $L$  (nm) and S value (V/dec).

FIGS. 17A and 17B are diagrams each illustrating a model used for calculation.

FIG. 18 is a diagram showing the relationship between gate voltage  $VG$  (V) and drain current  $ID$  (A).

## 25 BEST MODE FOR CARRYING OUT THE INVENTION

### [0036]

Examples of embodiments of the present invention will be described below with reference to the drawings. Note that the present invention is not limited to the following description and it will be readily appreciated by those skilled in the art that the modes and

30 details of the present invention can be modified in various ways without departing from the spirit and scope thereof. Therefore, the present invention should not be interpreted as being limited to the description in the following embodiments.

[0037]

Note that the position, size, range, or the like of each component illustrated in drawings and the like is not accurately represented in some cases for easy understanding. Therefore, the disclosed invention is not necessarily limited to the position, size, range, or  
5 the like as disclosed in the drawings and the like.

[0038]

Note that ordinal numbers such as "first", "second", and "third" in this specification are used in order to avoid confusion among components, and the terms do not limit the components numerically.

10 [0039]

(Embodiment 1)

In this embodiment, a structure and a manufacturing method of a semiconductor device according to one embodiment of the disclosed invention will be described with reference to FIGS. 1A to 1C, FIGS. 2A to 2G, and FIGS. 3A to 3G.

15 [0040]

<Example of Structure of Semiconductor Device>

FIGS. 1A to 1C each illustrate an example of a structure of a semiconductor device. FIG. 1A is a first structural example, and FIG. 1B is a second structural example. FIG. 1C is a modified example of FIG. 1A.

20 [0041]

A transistor 162 in FIG. 1A includes an insulating layer 143b over a substrate 140 having a surface where components are formed, a source electrode 142a and a drain electrode 142b embedded in an insulating layer including the insulating layer 143b, an oxide semiconductor layer 144 in contact with part of an upper surface of the insulating layer 143b, an upper surface of the source electrode 142a, and an upper surface of the drain electrode 142b, a gate insulating layer 146 covering the oxide semiconductor layer 144, and a gate electrode 148a over the gate insulating layer 146.

[0042]

With the use of an oxide semiconductor for an active layer of a transistor as illustrated in FIG. 1A, favorable characteristics can be obtained. For example, the S value of a transistor can be 65 mV/dec or less, preferably less than 63 mV/dec.

[0043]

In addition, part of the upper surface of the insulating layer 143b (particularly referring to a region parallel to the surface where components are formed), which is in contact with the oxide semiconductor layer, has a root-mean-square (RMS) roughness of 1 nm or less (preferably 0.5 nm or less). The difference in height between the part 5 of the upper surface of the insulating layer 143b and the upper surface of the source electrode 142a and the difference in height between the part of the upper surface of the insulating layer 143b and the upper surface of the drain electrode 142b are 5 nm or more.

[0044]

10 As described above, in one embodiment of the disclosed invention, a channel formation region of the transistor 162 is provided over a highly flat region with a root-mean-square (RMS) roughness of 1 nm or less. This makes it possible to prevent a problem such as a short-channel effect even in a situation where the transistor 162 is miniaturized and to provide the transistor 162 with favorable characteristics.

15 [0045]

In addition, the oxide semiconductor layer 144 can have a uniform thickness by improving the planarity of the surface where components are formed, and the transistor 162 can have improved characteristics. Furthermore, a decrease in coverage which may be caused by a large difference in height can be suppressed, and a disconnection or 20 a defective connection of the oxide semiconductor layer 144 can be prevented.

[0046]

Moreover, the path of electric current can be extended by making a small difference in height (e.g., 5 nm to 20 nm) between part of the upper surface of the insulating layer 143b and the upper surface of the source electrode 142a and between 25 the part of the upper surface of the insulating layer 143b and the upper surface of the drain electrode 142b. This makes it possible to alleviate the concentration of an electric field in the transistor 162 and to suppress a short-channel effect.

[0047]

Here, the oxide semiconductor layer 144 is preferably an oxide semiconductor 30 layer which is purified by sufficiently removing an impurity such as hydrogen therefrom or by sufficiently supplying oxygen thereto. Specifically, the hydrogen concentration of the oxide semiconductor layer 144 is  $5 \times 10^{19}$  atoms/cm<sup>3</sup> or less, preferably  $5 \times 10^{18}$

atoms/cm<sup>3</sup> or less, more preferably  $5 \times 10^{17}$  atoms/cm<sup>3</sup> or less, for example. Note that the above hydrogen concentration of the oxide semiconductor layer 144 is measured by secondary ion mass spectrometry (SIMS). The concentration of carriers generated due to a donor such as hydrogen in the oxide semiconductor layer 144, in which hydrogen is 5 reduced to a sufficiently low concentration so that the oxide semiconductor layer is purified and in which defect states in an energy gap due to oxygen deficiency are reduced by sufficiently supplying oxygen as described above, is less than  $1 \times 10^{12}$  /cm<sup>3</sup>, preferably less than  $1 \times 10^{11}$  /cm<sup>3</sup>, more preferably less than  $1.45 \times 10^{10}$  /cm<sup>3</sup>. In addition, for example, the off-state current (per unit channel width (1  $\mu$ m), here) at room temperature 10 (25 °C) is 100 zA (1 zA (zeptoampere) is  $1 \times 10^{-21}$  A) or less, preferably 10 zA or less. In this manner, by using an i-type (intrinsic) or substantially i-type oxide semiconductor, the transistor 162 which has extremely favorable off-state current characteristics can be obtained.

[0048]

15 Note that as disclosed in Non-Patent Document 7 and the like, a relatively large-size transistor whose channel length is 2  $\mu$ m to 100  $\mu$ m can be manufactured with use of an n-type oxide semiconductor having a high carrier density of  $2 \times 10^{19}$  /cm<sup>3</sup>. However, when such a material is applied to a miniaturized transistor whose channel length is less than 2  $\mu$ m, the threshold voltage drastically shifts negatively, and thus it is 20 difficult to realize a normally-off transistor. In other words, the transistor which has a channel length of less than 2  $\mu$ m and is manufactured using such a material does not work in practice. In contrast, an intrinsic or substantially intrinsic oxide semiconductor which is purified has a carrier density of at most less than  $1 \times 10^{14}$  /cm<sup>3</sup>, which does not cause a problem of normally on; thus, with use of such an intrinsic or 25 substantially intrinsic oxide semiconductor, a transistor whose channel length is less than 2  $\mu$ m can be easily realized.

[0049]

A transistor 262 in FIG. 1B has a structure similar to that of the transistor 162. That is, the transistor 262 includes an insulating layer 243b over a substrate 240 having 30 a surface where components are formed, a source electrode 242a and a drain electrode 242b embedded in an insulating layer including the insulating layer 243b, an oxide

semiconductor layer 244 in contact with part of an upper surface of the insulating layer 243b, an upper surface of the source electrode 242a, and an upper surface of the drain electrode 242b, a gate insulating layer 246 covering the oxide semiconductor layer 244, and a gate electrode 248a over the gate insulating layer 246.

5 [0050]

In addition, part of the upper surface of the insulating layer 243b, which is in contact with the oxide semiconductor layer, has a root-mean-square (RMS) roughness of 1 nm or less (preferably 0.5 nm or less). The difference in height between the part of the upper surface of the insulating layer 243b and the upper surface of the source electrode 242a and the difference in height between the part of the upper surface of the insulating layer 243b and the upper surface of the drain electrode 242b are 5 nm or more.

10 [0051]

A difference between the transistor 262 in FIG. 1B and the transistor 162 in FIG. 1A is whether or not part of the insulating layer is provided so as to cover end portions of the source electrode and the drain electrode. In other words, the insulating layer 143b of the transistor 162 in FIG. 1A is provided so as to cover the source electrode 142a and the drain electrode 142b, whereas the insulating layer 243b of the transistor 262 in FIG. 1B does not cover the source electrode 242a and the drain electrode 242b.

15 [0052]

The effects resulting from the structure in FIG. 1B are similar to those obtained in the case of FIG. 1A. In other words, because a channel formation region of the transistor 262 is provided over a highly flat region, a problem such as a short-channel effect can be prevented even in a situation where the transistor 262 is miniaturized, and the transistor 262 with favorable characteristics can be provided. In addition, the oxide semiconductor layer 244 can have a uniform thickness by improving the planarity of the surface where components are formed, and the transistor 262 can have improved characteristics. Furthermore, a decrease in coverage which may be caused by a large difference in height can be suppressed, and a disconnection or a defective connection of the oxide semiconductor layer 244 can be prevented. Moreover, by making a small difference in height between the insulating layer and the source electrode (and the drain electrode) as described above, the concentration of an electric field in the transistor 262

can be alleviated, and a short-channel effect can be suppressed.

[0053]

A transistor 162 in FIG. 1C is a modified example of the transistor 162 in FIG. 1A. Specifically, the transistor 162 in FIG. 1C corresponds to a transistor obtained by 5 modifying the source electrode 142a and the drain electrode 142b in FIG. 1A into a tapered shape. The taper angle can be, for example, greater than or equal to 30° and less than or equal to 60°. Note that the “taper angle” means an angle formed by the side surface and the bottom surface of a layer having a tapered shape (for example, the source electrode 142a) when observed from a direction perpendicular to a cross section thereof 10 (a plane perpendicular to a surface of the substrate 140).

[0054]

<Example of Method for Manufacturing Semiconductor Device>

Next, an example of a method for manufacturing the semiconductor device will be described with reference to FIGS. 2A to 2G and FIGS. 3A to 3G. Here, FIGS. 2A to 15 2G illustrate an example of a method for manufacturing the transistor 162 shown in FIG. 1A, and FIGS. 3A to 3G illustrate an example of a method for manufacturing the transistor 262 shown in FIG. 1B. Note that FIGS. 2A to 2G will first be described below, and then, for FIGS. 3A to 3G, only a main difference from FIGS. 2A to 2G will be described. In addition, the structure shown in FIG. 1C can be manufactured in a 20 manner similar to that of the structure shown in FIG. 1A; thus, the detailed description thereof is omitted.

[0055]

FIGS. 2A to 2G will be described below. First, the source electrode 142a and the drain electrode 142b are formed over the substrate 140 having a surface where 25 components are formed (see FIG. 2A).

[0056]

Although there is no particular limitation on a substrate which can be used as the substrate 140, it is necessary that the substrate 140 has at least heat resistance high enough to withstand heat treatment to be performed later. For example, the substrate may be a 30 glass substrate, a ceramic substrate, a quartz substrate, a sapphire substrate, or the like. Alternatively, the substrate may be a single crystal semiconductor substrate or a

polycrystalline semiconductor substrate of silicon, silicon carbide, or the like, a compound semiconductor substrate of silicon germanium or the like, an SOI substrate, or the like as long as the substrate has an insulating surface. Still alternatively, the substrate may be any of these substrates provided with a semiconductor element. Still 5 alternatively, the substrate 140 may be provided with a base film.

[0057]

Note that the preferred surface of the substrate 140 where components are formed is a sufficiently flat surface. For example, a surface with a root-mean-square roughness (RMS) of 1 nm or less (preferably 0.5 nm or less) is employed. When the 10 transistor 162 is formed over such a surface, the characteristics can be sufficiently improved. In the case where the surface of the substrate 140 has poor flatness, it is desirable that the surface be subjected to chemical mechanical polishing (CMP) treatment, etching treatment, or the like so as to have the above flatness. Note that, for the details of the CMP treatment, the description of CMP treatment for an insulating 15 layer 143 mentioned below can be referred to.

[0058]

The source electrode 142a and the drain electrode 142b can be formed by forming a conductive layer over the substrate 140 having a surface where components are formed and then selectively etching the conductive layer.

[0059]

The above conductive layer can be formed by a PVD method such as a sputtering method, or a CVD method such as a plasma CVD method. As a material of the conductive layer, an element selected from aluminum, chromium, copper, tantalum, 25 titanium, molybdenum, and tungsten, an alloy including any of these elements as a component, or the like can be used. A material including one of manganese, magnesium, zirconium, beryllium, neodymium, or scandium or a combination of a plurality of these elements may be used.

[0060]

The conductive layer may have a single-layer structure or a stacked-layer 30 structure including two or more layers. For example, the conductive layer may have a single-layer structure of a titanium film or a titanium nitride film, a single-layer structure of an aluminum film including silicon, a two-layer structure in which a titanium film is

stacked over an aluminum film, a two-layer structure in which a titanium film is stacked over a titanium nitride film, a three-layer structure in which a titanium film, an aluminum film, and a titanium film are stacked in this order, or the like. Note that the conductive layer having a single-layer structure of a titanium film or a titanium nitride film has an 5 advantage in that it can be easily processed into the source electrode 142a and the drain electrode 142b having a tapered shape.

[0061]

The conductive layer may be formed using a conductive metal oxide. As the conductive metal oxide, indium oxide ( $\text{In}_2\text{O}_3$ ), tin oxide ( $\text{SnO}_2$ ), zinc oxide ( $\text{ZnO}$ ), an 10 indium oxide-tin oxide alloy ( $\text{In}_2\text{O}_3\text{-SnO}_2$ , which is abbreviated to ITO in some cases), an indium oxide-zinc oxide alloy ( $\text{In}_2\text{O}_3\text{-ZnO}$ ), or any of these metal oxide materials including silicon or silicon oxide can be used.

[0062]

Although either dry etching or wet etching may be performed as the etching of 15 the conductive layer, dry etching with high controllability is preferably used for miniaturization. The etching may be performed so that the source electrode 142a and the drain electrode 142b to be formed have a tapered shape. The taper angle can be, for example, greater than or equal to 30° and less than or equal to 60°.

[0063]

20 The channel length ( $L$ ) of the transistor 162 is determined by a distance between upper edge portions of the source electrode 142a and the drain electrode 142b. Note that for light exposure for forming a mask in the case of manufacturing a transistor with a channel length ( $L$ ) of less than 25 nm, light exposure is preferably performed with extreme ultraviolet light whose wavelength is several nanometers to several tens of 25 nanometers, which is extremely short. The resolution of light exposure with extreme ultraviolet light is high and the depth of focus is large. For these reasons, the channel length ( $L$ ) of the transistor to be formed later can be set to less than 2  $\mu\text{m}$ , preferably in the range of 10 nm to 350 nm (0.35  $\mu\text{m}$ ), in which case the circuit can operate at higher speed. In addition, power consumption of the semiconductor device can be reduced by 30 miniaturization.

[0064]

Next, the insulating layer 143 is formed so as to cover the source electrode 142a and the drain electrode 142b (see FIG. 2B).

[0065]

The insulating layer 143 can be formed using an inorganic insulating material 5 such as silicon oxide, silicon oxynitride, silicon nitride, or aluminum oxide. It is particularly preferable that the insulating layer 143 be formed using silicon oxide because the oxide semiconductor layer 144 formed later is in contact with the insulating layer 143. Although there is no particular limitation on the forming method of the insulating layer 143, in consideration of contact with the oxide semiconductor layer 144, 10 a method in which hydrogen is sufficiently reduced is preferably employed. Examples of such a method include a sputtering method and the like. Needless to say, another deposition method such as a plasma CVD method may be used.

[0066]

Next, an insulating layer 143a is formed by thinning the insulating layer 143 by 15 chemical mechanical polishing (CMP) treatment (see FIG. 2C). Here, the CMP treatment is performed under such conditions that the surfaces of the source electrode 142a and the drain electrode 142b remain unexposed. In addition, the CMP treatment is performed under such conditions that the root-mean-square (RMS) roughness of a 20 surface of the insulating layer 143a becomes 1 nm or less (preferably 0.5 nm or less). By the CMP treatment performed under such conditions, the planarity of a surface where the oxide semiconductor layer 144 is formed later can be improved, and the characteristics of the transistor 162 can be improved.

[0067]

The CMP treatment is a method for planarizing a surface of an object to be 25 processed by a combination of chemical and mechanical actions. More specifically, the CMP treatment is a method in which a polishing cloth is attached to a polishing stage, the polishing stage and an object to be processed are each rotated or swung while a slurry (an abrasive) is supplied between the object to be processed and the polishing cloth, and the 30 surface of the object to be processed is polished by a chemical reaction between the slurry and the surface of the object to be processed and by a mechanical polishing action of the polishing cloth on the object to be processed.

[0068]

Note that the CMP treatment may be performed only once or plural times. When the CMP treatment is performed plural times, it is preferable that first polishing be performed at a high polishing rate and final polishing be performed at a low polishing rate. By performing polishing at different polishing rates, the planarity of the surface of the insulating layer 143a can be further improved.

[0069]

Next, the insulating layer 143b having openings reaching the source electrode 142a and the drain electrode 142b is formed by selectively removing regions of the insulating layer 143a, which overlap with the source electrode 142a and the drain electrode 142b (see FIG. 2D). The insulating layer 143a can be selectively removed by a method such as etching.

[0070]

The insulating layer 143a may be etched by either dry etching or wet etching. Further, the insulating layer 143a is preferably etched so that the openings to be formed have an inclined surface. The inclined surface preferably has an angle formed with the surface of the substrate 140 and the inclined surface greater than or equal to 30° and less than or equal to 60° when seen from a direction perpendicular to a cross section (a plane perpendicular to the surface of the substrate 140) of the inclined surface. By etching the insulating layer 143a under such conditions, coverage with the oxide semiconductor layer to be formed later so as to cover the insulating layer 143b is increased, so that a disconnection of the oxide semiconductor layer or the like can be prevented.

[0071]

By the above etching step, a surface having a small difference in height (e.g., 5 nm to 20 nm) is formed by part of an upper surface of the insulating layer 143b, an upper surface of the source electrode 142a, and an upper surface of the drain electrode 142b. In addition, end portions of the source electrode 142a and the drain electrode 142b are covered with the insulating layer 143b. Note that, with a higher patterning accuracy in forming the openings, it is possible to obtain a structure in which end portions of the source electrode 142a and the drain electrode 142b are not covered with the insulating layer 143b.

[0072]

Next, the oxide semiconductor layer 144 covering the above-described surface is formed in contact with part of the source electrode 142a, the drain electrode 142b, and the insulating layer 143b; then, the gate insulating layer 146 is formed so as to cover the oxide semiconductor layer 144 (see FIG. 2E).

5 [0073]

The oxide semiconductor layer 144 contains at least one element selected from In, Ga, Sn, and Zn. For example, a four-component metal oxide such as an In-Sn-Ga-Zn-O-based oxide semiconductor, a three-component metal oxide such as an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an 10 In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor, a two-component metal oxide such as an In-Zn-O-based oxide semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, an 15 In-Mg-O-based oxide semiconductor, or an In-Ga-O-based oxide semiconductor, a single-component metal oxide such as an In-O-based oxide semiconductor, a Sn-O-based oxide semiconductor, or a Zn-O-based oxide semiconductor, or the like can be used. In addition, any of the above oxide semiconductors may contain an element other than In, Ga, Sn, and Zn, for example,  $\text{SiO}_2$ .

20 [0074]

For example, an In-Ga-Zn-O-based oxide semiconductor means an oxide film containing indium (In), gallium (Ga), and zinc (Zn), and there is no limitation on the composition ratio thereof.

[0075]

25 In particular, an In-Ga-Zn-O-based oxide semiconductor material has sufficiently high resistance when there is no electric field and thus off-state current can be sufficiently reduced. In addition, also having high field-effect mobility, the In-Ga-Zn-O-based oxide semiconductor material is suitable for a semiconductor material used in a semiconductor device.

30 [0076]

As a typical example of the In-Ga-Zn-O-based oxide semiconductor material, an oxide semiconductor material represented by  $\text{InGaO}_3(\text{ZnO})_m$  ( $m > 0$ ) is given. Using  $M$

instead of Ga, there is an oxide semiconductor material represented by  $\text{InMO}_3(\text{ZnO})_m$  ( $m > 0$ ). Here,  $M$  denotes one or more metal elements selected from gallium (Ga), aluminum (Al), iron (Fe), nickel (Ni), manganese (Mn), cobalt (Co), or the like. For example,  $M$  may be Ga, Ga and Al, Ga and Fe, Ga and Ni, Ga and Mn, Ga and Co, or the like. Note that the above-described compositions are derived from the crystal structures that the oxide semiconductor material can have, and are mere examples.

5 [0077]

In the case where an In-Zn-O-based material is used as an oxide semiconductor, a target therefore has a composition ratio of  $\text{In:Zn} = 50:1$  to  $1:2$  in an atomic ratio (In<sub>2</sub>O<sub>3</sub>: ZnO = 25:1 to 1:4 in a molar ratio), preferably,  $\text{In:Zn} = 20:1$  to  $1:1$  in an atomic ratio (In<sub>2</sub>O<sub>3</sub>: ZnO = 10:1 to 1:2 in a molar ratio), further preferably,  $\text{In:Zn} = 15:1$  to 1.5:1 in an atomic ratio (In<sub>2</sub>O<sub>3</sub>: ZnO = 15:2 to 3:4 in a molar ratio). For example, in a target used for formation of an In-Zn-O-based oxide semiconductor which has an atomic ratio of  $\text{In:Zn:O} = X:Y:Z$ , the relation of  $Z > 1.5X+Y$  is satisfied.

15 [0078]

As a target used for forming the oxide semiconductor layer 144 by a sputtering method, a target having a composition ratio of  $\text{In:Ga:Zn} = 1:x:y$  ( $x$  is greater than or equal to 0 and  $y$  is greater than or equal to 0.5 and less than or equal to 5) is preferably used. For example, a target having a composition ratio of In<sub>2</sub>O<sub>3</sub>:Ga<sub>2</sub>O<sub>3</sub>:ZnO = 1:1:2 [molar ratio] or the like can be used. Furthermore, a target having a composition ratio of In<sub>2</sub>O<sub>3</sub>:Ga<sub>2</sub>O<sub>3</sub>:ZnO = 1:1:1 [molar ratio], a target having a composition ratio of In<sub>2</sub>O<sub>3</sub>:Ga<sub>2</sub>O<sub>3</sub>:ZnO = 1:1:4 [molar ratio], or a target having a composition ratio of In<sub>2</sub>O<sub>3</sub>:Ga<sub>2</sub>O<sub>3</sub>:ZnO = 1:0:2 [molar ratio] can also be used.

[0079]

25 In this embodiment, the oxide semiconductor layer 144 having an amorphous structure is formed by a sputtering method with the use of an In-Ga-Zn-O-based metal oxide target. The thickness ranges from 1 nm to 50 nm, preferably from 2 nm to 20 nm, more preferably from 3 nm to 15 nm.

[0080]

30 The relative density of the metal oxide in the metal oxide target is 80 % or more, preferably 95 % or more, and more preferably 99.9 % or more. The use of the metal

oxide target having high relative density makes it possible to form an oxide semiconductor layer having a dense structure.

[0081]

The atmosphere in which the oxide semiconductor layer 144 is formed is preferably a rare gas (typically, argon) atmosphere, an oxygen atmosphere, or a mixed atmosphere containing a rare gas (typically, argon) and oxygen. Specifically, it is preferable to use a high-purity gas atmosphere, for example, from which an impurity such as hydrogen, water, a hydroxyl group, or hydride is removed to 1 ppm or less (preferably, 10 ppb or less).

10 [0082]

In forming the oxide semiconductor layer 144, for example, an object to be processed is held in a treatment chamber that is maintained under reduced pressure, and the object to be processed is heated to a temperature higher than or equal to 100 °C and lower than 550 °C, preferably higher than or equal to 200 °C and lower than or equal to 15 400 °C. Alternatively, the temperature of an object to be processed in forming the oxide semiconductor layer 144 may be room temperature (25 °C ± 10 °C (higher than or equal to 15 °C and lower than or equal to 35 °C)). Then, moisture in the treatment chamber is removed, a sputtering gas from which hydrogen, water, or the like has been removed is introduced, and the above-described target is used; thus, the oxide semiconductor layer 20 144 is formed. By forming the oxide semiconductor layer 144 while heating the object to be processed, an impurity in the oxide semiconductor layer 144 can be reduced. Moreover, damage due to sputtering can be reduced. In order to remove the moisture in the treatment chamber, it is preferable to use an entrapment vacuum pump. For example, a cryopump, an ion pump, a titanium sublimation pump, or the like can be used. A turbo 25 pump provided with a cold trap may be used. Since hydrogen, water, or the like can be removed from the treatment chamber evacuated with a cryopump or the like, the concentration of an impurity in the oxide semiconductor layer can be reduced.

[0083]

For example, conditions for forming the oxide semiconductor layer 144 can be 30 set as follows: the distance between the object to be processed and the target is 170 mm, the pressure is 0.4 Pa, the direct current (DC) power is 0.5 kW, and the atmosphere is an

oxygen (100 % oxygen) atmosphere, an argon (100 % argon) atmosphere, or a mixed atmosphere of oxygen and argon. Note that a pulsed direct current (DC) power source is preferably used because dust (powder or flake-like substances formed at the time of the film formation) can be reduced and the film thickness can be made uniform. The 5 thickness of the oxide semiconductor layer 144 is set in the range of 1 nm to 50 nm, preferably 2 nm to 20 nm, more preferably 3 nm to 15 nm. By employing a structure according to the disclosed invention, a short-channel effect due to miniaturization can be suppressed even in the case of using the oxide semiconductor layer 144 having such a thickness. Note that the appropriate thickness of the oxide semiconductor layer differs 10 depending on the oxide semiconductor material used, the intended use of the semiconductor device, or the like; therefore, the thickness can be determined as appropriate in accordance with the material, the intended use, or the like. Note that a surface where the oxide semiconductor layer 144 is formed is sufficiently planarized in one embodiment of the disclosed invention. Therefore, even an oxide semiconductor 15 layer having a small thickness can be favorably formed.

[0084]

Note that before the oxide semiconductor layer 144 is formed by a sputtering method, reverse sputtering in which plasma is generated with an argon gas introduced may be performed so that a material attached to a surface where the oxide semiconductor 20 layer 144 is to be formed (e.g., a surface of the insulating layer 143b) is removed. Here, the reverse sputtering is a method in which ions collide with a surface to be processed so that the surface is modified, in contrast to normal sputtering in which ions collide with a sputtering target. An example of a method for making ions collide with a surface to be processed is a method in which high-frequency voltage is applied to the surface side in an 25 argon atmosphere so that plasma is generated near the object to be processed. Note that an atmosphere of nitrogen, helium, oxygen, or the like may be used instead of an argon atmosphere.

[0085]

After the oxide semiconductor layer 144 is formed, heat treatment (first heat 30 treatment) is preferably performed on the oxide semiconductor layer 144. Through the first heat treatment, excess hydrogen (including water or a hydroxyl group) in the oxide semiconductor layer 144 can be removed, the structure of the oxide semiconductor layer

144 can be ordered, and defect states in an energy gap can be reduced. For example, the temperature of the first heat treatment is set higher than or equal to 300 °C and lower than 550 °C, or higher than or equal to 400 °C and lower than or equal to 500 °C.

[0086]

5 For example, after an object to be processed is introduced into an electric furnace including a resistance heater or the like, the heat treatment can be performed at 450 °C for one hour in a nitrogen atmosphere. The oxide semiconductor layer is not exposed to the air during the heat treatment so that entry of water or hydrogen can be prevented.

[0087]

10 The heat treatment apparatus is not limited to the electric furnace and may be an apparatus for heating an object to be processed by thermal radiation or thermal conduction from a medium such as a heated gas. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp rapid thermal annealing (LRTA) apparatus can be used. The LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc lamp, a carbon arc lamp, a high pressure sodium lamp, or a high pressure mercury lamp. The GRTA apparatus is an apparatus for performing heat treatment using a high-temperature gas. As the gas, an inert gas that does not react with an object to be processed by heat treatment, for example, nitrogen or a rare gas such as argon, is used.

15

20

[0088]

25 For example, as the first heat treatment, GRTA treatment may be performed as follows. The object to be processed is put in a heated inert gas atmosphere, heated for several minutes, and taken out of the inert gas atmosphere. The GRTA treatment enables high-temperature heat treatment in a short time. Moreover, the GRTA treatment can be employed even when the temperature exceeds the upper temperature limit of the object to be processed. Note that the inert gas may be switched to a gas including oxygen during the treatment. This is because defect states in an energy gap caused by oxygen vacancies can be reduced by performing the first heat treatment in an atmosphere including oxygen.

30

[0089]

Note that as the inert gas atmosphere, an atmosphere that contains nitrogen or a rare gas (e.g., helium, neon, or argon) as its main component and does not contain water, hydrogen, or the like is preferably used. For example, the purity of nitrogen or a rare gas such as helium, neon, or argon introduced into a heat treatment apparatus is set to 6N 5 (99.9999 %) or more, preferably 7N (99.99999 %) or more (i.e., the impurity concentration is 1 ppm or less, preferably 0.1 ppm or less).

[0090]

In any case, a transistor with extremely excellent characteristics can be obtained with the use of the oxide semiconductor layer which is an i-type (intrinsic) or 10 substantially i-type oxide semiconductor layer obtained by reducing an impurity through the first heat treatment.

[0091]

The above heat treatment (the first heat treatment) can also be referred to as dehydration treatment, dehydrogenation treatment, or the like because it has the effect of 15 removing hydrogen, water, or the like. The dehydration treatment or the dehydrogenation treatment can be performed after the gate insulating layer 146 is formed, or after a gate electrode is formed. Such dehydration treatment or dehydrogenation treatment may be performed once or plural times.

[0092]

20 After the oxide semiconductor layer 144 is formed, the oxide semiconductor layer 144 may be processed into an island-shaped oxide semiconductor layer. The oxide semiconductor layer 144 can be processed into an island-shaped oxide semiconductor layer by etching, for example. The etching may be performed either before the heat treatment or after the heat treatment. Dry etching is preferably used in 25 terms of element miniaturization, but wet etching may be used. An etching gas or an etchant can be selected as appropriate depending on a material to be etched.

[0093]

The gate insulating layer 146 can be formed by a CVD method, a sputtering method, or the like. The gate insulating layer 146 is preferably formed so as to contain 30 silicon oxide, silicon nitride, silicon oxynitride, aluminum oxide, tantalum oxide, hafnium oxide, yttrium oxide, hafnium silicate ( $\text{HfSi}_x\text{O}_y$  ( $x > 0, y > 0$ )), hafnium silicate

to which nitrogen is added ( $\text{HfSi}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ )), hafnium aluminate to which nitrogen is added ( $\text{HfAl}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ )), or the like. The gate insulating layer 146 may have a single-layer structure or a stacked-layer structure. There is no particular limitation on the thickness of the gate insulating layer 146; the thickness is preferably small in order to ensure the operation of the transistor when the semiconductor device is miniaturized. For example, in the case of using silicon oxide, the thickness can be 1 nm to 100 nm, preferably 10 nm to 50 nm.

5 [0094]

When the gate insulating layer is thin as described above, gate leakage due to a tunnel effect or the like becomes a problem. In order to solve the problem of gate leakage, the gate insulating layer 146 may be formed using a high dielectric constant (high-k) material such as hafnium oxide, tantalum oxide, yttrium oxide, hafnium silicate ( $\text{HfSi}_x\text{O}_y$  ( $x > 0, y > 0$ )), hafnium silicate to which nitrogen is added ( $\text{HfSi}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ )), or hafnium aluminate to which nitrogen is added ( $\text{HfAl}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ ))). The use of a high-k material for the gate insulating layer 146 makes it possible to increase the thickness in order to suppress gate leakage as well as ensuring electrical properties. For example, the relative permittivity of hafnium oxide is approximately 15, which is much higher than that of silicon oxide which is 3 to 4. With such a material, a gate insulating layer where the equivalent oxide thickness is less than 15 nm, preferably 2 nm to 10 nm, can be easily formed. Note that a stacked-layer structure of a film including a high-k material and a film including any of silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, aluminum oxide, and the like may also be employed.

20 [0095]

25 After the gate insulating layer 146 is formed, second heat treatment is preferably performed in an inert gas atmosphere or an oxygen atmosphere. The temperature of the heat treatment is set in the range of 200 °C to 450 °C, preferably 250 °C to 350 °C. For example, the heat treatment may be performed at 250 °C for one hour in a nitrogen atmosphere. By the second heat treatment, variation in electrical characteristics of the 30 transistor can be reduced. In the case where the gate insulating layer 146 contains oxygen, oxygen can be supplied to the oxide semiconductor layer 144 and oxygen

vacancies in the oxide semiconductor layer 144 can be filled; thus, the oxide semiconductor layer 144 which is i-type (intrinsic) or substantially i-type can also be formed.

[0096]

5 Note that the second heat treatment is performed in this embodiment after the gate insulating layer 146 is formed; there is no limitation on the timing of the second heat treatment. For example, the second heat treatment may be performed after the gate electrode is formed. Alternatively, the first heat treatment and the second heat treatment may be performed in succession, or the first heat treatment may double as the second heat treatment, or the second heat treatment may double as the first heat treatment.

10 [0097]

By performing at least one of the first heat treatment and the second heat treatment as described above, the oxide semiconductor layer 144 can be purified so as not to contain impurities other than main components as little as possible.

15 [0098]

Next, the gate electrode 148a is formed over the gate insulating layer 146 (see FIG. 2F). The gate electrode 148a can be formed by forming a conductive layer over the gate insulating layer 146 and then selectively etching the conductive layer. The conductive layer to be the gate electrode 148a can be formed by a PVD method such as a sputtering method, or a CVD method such as a plasma CVD method. The details are similar to those in the case of the source electrode 142a, the drain electrode 142b, or the like; thus, the description thereof can be referred to. Note that although part of the gate electrode 148a overlaps with the source electrode 142a and the drain electrode 142b in the structure employed here, the disclosed invention is not limited to this structure. It is possible to employ a structure in which an end portion of the gate electrode 148a and an end portion of the source electrode 142a overlap with each other, and an end portion of the gate electrode 148a and an end portion of the drain electrode 142b overlap with each other.

[0099]

30 Next, an insulating layer 150 and an insulating layer 152 are formed so as to cover the gate insulating layer 146, the gate electrode 148a, and the like (see FIG. 2G). The insulating layer 150 and the insulating layer 152 can be formed by a PVD method, a

CVD method, or the like. The insulating layer 150 and the insulating layer 152 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon oxynitride, silicon nitride, hafnium oxide, or aluminum oxide.

[0100]

5 Note that the insulating layer 150 and the insulating layer 152 are preferably formed using a low dielectric constant material or a low dielectric constant structure (such as a porous structure). This is because when the insulating layer 150 and the insulating layer 152 have a low dielectric constant, capacitance generated between wirings, electrodes, or the like can be reduced and operation at higher speed can be achieved.

10 [0101]

Note that although a stacked-layer structure of the insulating layer 150 and the insulating layer 152 is used in this embodiment, an embodiment of the disclosed invention is not limited to this example. A single-layer structure or a stacked-layer structure including three or more layers can also be used. Alternatively, a structure in 15 which the insulating layers are not provided is also possible.

[0102]

Note that the insulating layer 152 is desirably formed so as to have a flat surface. This is because when the insulating layer 152 has a flat surface, an electrode, a wiring, or the like can be favorably formed over the insulating layer 152 even in the case where the 20 semiconductor device or the like is miniaturized. Note that the insulating layer 152 can be planarized using a method such as chemical mechanical polishing (CMP).

[0103]

Through the above steps, the transistor 162 including the oxide semiconductor layer 144, which is purified, is completed (see FIG. 2G).

25 [0104]

Note that a variety of wirings, electrodes, or the like may be formed after the above steps. The wirings or the electrodes can be formed by a method such as a so-called damascene method or dual damascene method.

[0105]

30 As described above, in one embodiment of the disclosed invention, a channel formation region of the transistor 162 is provided over a highly flat region with a root-mean-square (RMS) roughness of 1 nm or less. This makes it possible to prevent

a problem such as a short-channel effect even in a situation where the transistor 162 is miniaturized and to obtain the transistor 162 with favorable characteristics.

[0106]

In addition, the oxide semiconductor layer 144 can have a uniform thickness by 5 improving the planarity of the surface where components are formed, and the transistor 162 can have improved characteristics. Furthermore, a decrease in coverage which may be caused by a large difference in height can be suppressed, and a disconnection or a defective connection of the oxide semiconductor layer 144 can be prevented.

[0107]

10 Moreover, the path of electric current can be extended by making a small difference in height (e.g., 5 nm to 20 nm) between part of the upper surface of the insulating layer 143b and the upper surface of the source electrode 142a and between the part of the upper surface of the insulating layer 143b and the upper surface of the drain electrode 142b. This makes it possible to alleviate the concentration of an 15 electric field in the transistor 162 and to suppress a short-channel effect.

[0108]

In the transistor 162 described in this embodiment, the oxide semiconductor layer 144 is purified and thus contains hydrogen at a concentration of  $5 \times 10^{19}$  atoms/cm<sup>3</sup> or less, preferably  $5 \times 10^{18}$  atoms/cm<sup>3</sup> or less, more preferably  $5 \times 10^{17}$  atoms/cm<sup>3</sup> or less.

20 In addition, the density of carriers generated due to a donor such as hydrogen in the oxide semiconductor layer 144 is, for example, less than  $1 \times 10^{12}$  /cm<sup>3</sup>, preferably less than  $1.45 \times 10^{10}$  /cm<sup>3</sup>, which is sufficiently lower than the carrier density of a general silicon wafer (approximately  $1 \times 10^{14}$  /cm<sup>3</sup>). In addition, the off-state current of the transistor 162 is sufficiently small. For example, the off-state current (per unit channel 25 width (1  $\mu$ m), here) of the transistor 162 at room temperature (25 °C) is 100 zA (1 zA (zeptoampere) is  $1 \times 10^{-21}$  A) or less, preferably 10 zA or less. When the above structure is employed, the off-state current of the transistor can be  $1 \times 10^{-24}$  A/ $\mu$ m to  $1 \times 10^{-30}$  A/ $\mu$ m in theory.

[0109]

30 In this manner, by using the oxide semiconductor layer 144 which is purified and is intrinsic, it becomes easy to sufficiently reduce the off-state current of the transistor.

In addition, by using the oxide semiconductor layer 144 which is purified and is intrinsic in this manner, the S value of the transistor can be 65 mV/dec or less, preferably less than 63 mV/dec.

[0110]

5 Next, FIGS. 3A to 3G will be described.

[0111]

In a manner similar to the case of FIGS. 2A to 2G, first, a conductive layer 241a to be a source electrode later and a conductive layer 241b to be a drain electrode later are formed over the substrate 240 having a surface where components are formed 10 (see FIG. 3A). Then, an insulating layer 243 is formed so as to cover the conductive layer 241a and the conductive layer 241b (see FIG. 3B). The details are similar to those in the case of FIGS. 2A and 2B.

[0112]

Next, the insulating layer 243b is formed by thinning the insulating layer 243 by chemical mechanical polishing (CMP) treatment (see FIG. 3C). A difference from 15 the case in FIG. 2C is to perform the CMP treatment under such conditions that surfaces of the conductive layer 241a and the conductive layer 241b become exposed. The other conditions for the CMP treatment are similar to those in the case of FIG. 2C. For details, the description about FIG. 2C can be referred to.

20 [0113]

Next, the source electrode 242a and the drain electrode 242b are formed by thinning the conductive layer 241a and the conductive layer 241b (see FIG. 3D). The conductive layer 241a and the conductive layer 241b can be thinned by etching. For the etching, either dry etching or wet etching may be employed. Note that in order to 25 ensure a sufficient planarity of the surface of the insulating layer 243b, the etching needs to be performed under such conditions that only conductive layers can be selectively etched. An etching gas or an etchant can be selected as appropriate depending on a material to be etched.

[0114]

30 By the above thinning step, a surface having a small difference in height (e.g., 5 nm to 20 nm) is formed by part of an upper surface of the insulating layer 243b, an upper surface of the source electrode 242a, and an upper surface of the drain electrode

242b. In addition, end portions of the source electrode 242a and the drain electrode 242b are not covered with the insulating layer 243b. Note that a side surface of the insulating layer 243b may be inclined with respect to the bottom surface.

[0115]

5 Next, the oxide semiconductor layer 244 covering the above-described surface is formed in contact with part of the source electrode 242a, the drain electrode 242b, and the insulating layer 243b; then, the gate insulating layer 246 is formed so as to cover the oxide semiconductor layer 244 (see FIG. 3E). In addition, the gate electrode 248a is formed over the gate insulating layer 246 (see FIG. 3F). Then, an insulating layer 250 and an insulating layer 252 are formed so as to cover the gate insulating layer 246, the gate electrode 248a, and the like (see FIG. 3G). The details are similar to those 10 in the case of FIGS. 2E to 2G.

[0116]

15 The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0117]

(Embodiment 2)

20 In this embodiment, a structure and a manufacturing method of a semiconductor device according to another embodiment of the disclosed invention will be described with reference to FIGS. 4A to 4C, FIGS. 5A to 5D, and FIGS. 6A to 6C. Note that in some circuit diagrams mentioned below, "OS" is written beside a transistor in order to indicate that the transistor includes an oxide semiconductor.

[0118]

25 <Example of Structure of Semiconductor Device>

FIGS. 4A to 4C illustrate an example of a structure of a semiconductor device. FIG. 4A is a cross-sectional view of the semiconductor device; FIG. 4B is a plan view of the semiconductor device; and FIG. 4C illustrates a circuit configuration of the semiconductor device. Note that a structure of the semiconductor device is mainly 30 described in this embodiment, and operation of the semiconductor device will be described in detail in an embodiment below. Note that the semiconductor device illustrated in FIGS. 4A to 4C is just an example having predetermined functions and

does not represent all semiconductor devices according to the disclosed invention. The semiconductor device according to the disclosed invention can have another function by changing connection relationship of electrodes or the like as appropriate.

[0119]

5 FIG. 4A corresponds to a cross-sectional view along line A1-A2 and line B1-B2 in FIG. 4B. The semiconductor device illustrated in FIGS. 4A and 4B includes the transistor 162 described in the above embodiment, a transistor 160 below the transistor 162, and a capacitor 164.

[0120]

10 Here, a semiconductor material of the transistor 162 and a semiconductor material of the transistor 160 are preferably different materials. For example, the semiconductor material of the transistor 162 may be an oxide semiconductor, and the semiconductor material of the transistor 160 may be a semiconductor material (such as silicon) other than an oxide semiconductor. A transistor including an oxide semiconductor can hold charge for a long time owing to its characteristics. On the other 15 hand, a transistor including a material other than an oxide semiconductor can operate at high speed easily.

[0121]

20 The transistor 160 in FIGS. 4A to 4C includes a channel formation region 116 provided in a substrate 100 including a semiconductor material (such as silicon), impurity regions 120 provided such that the channel formation region 116 is sandwiched therebetween, metal compound regions 124 in contact with the impurity regions 120, a gate insulating layer 108 provided over the channel formation region 116, and a gate electrode 110 provided over the gate insulating layer 108. Note that a transistor whose 25 source electrode and drain electrode are not illustrated in a drawing may also be referred to as a transistor for the sake of convenience. Further, in such a case, in description of a connection of a transistor, a source region and a source electrode may be collectively referred to as a source electrode, and a drain region and a drain electrode may be collectively referred to as a drain electrode. That is, in this specification, the term 30 “source electrode” may include a source region.

[0122]

Further, an element isolation insulating layer 106 is formed over the substrate

100 so as to surround the transistor 160, and an insulating layer 130 is formed to cover the transistor 160. Note that in order to realize higher integration, the transistor 160 preferably has a structure without a sidewall insulating layer as illustrated in FIGS. 4A and 4B. On the other hand, in the case where characteristics of the transistor 160 have 5 priority, a sidewall insulating layer may be provided on a side surface of the gate electrode 110, and the impurity regions 120 may include a region having a different impurity concentration.

[0123]

10 The structure of the transistor 162 in FIGS. 4A to 4C is similar to the structure of the transistor 162 in the above embodiment. Note that in this embodiment, the source electrode 142a (which may be the drain electrode) of the transistor 162 is connected to the gate electrode 110 of the transistor 160. It is needless to say that the transistor 162 can be replaced with the transistor 262.

[0124]

15 The capacitor 164 in FIGS. 4A to 4C includes the source electrode 142a (which may be the drain electrode), the oxide semiconductor layer 144, the gate insulating layer 146, and an electrode 148b. In other words, the source electrode 142a functions as one electrode of the capacitor 164, and the electrode 148b functions as the other electrode of the capacitor 164. Note that the electrode 148b is formed in a process similar to that of 20 the gate electrode 148a of the transistor 162.

[0125]

25 Note that in the capacitor 164 of FIGS. 4A to 4C, the oxide semiconductor layer 144 and the gate insulating layer 146 are stacked, whereby insulation between the source electrode 142a and the electrode 148b can be sufficiently secured. It is needless to say that the capacitor 164 without including the oxide semiconductor layer 144 may be employed in order to secure sufficient capacitance. In addition, in the case where no capacitor is needed, a structure in which the capacitor 164 is not provided is also possible.

[0126]

30 In this embodiment, the transistor 162 and the capacitor 164 are provided so as to overlap with the transistor 160. By employing such a planar layout, higher integration can be realized. For example, given that the minimum feature size is F, the area occupied by the semiconductor device can be 15 F<sup>2</sup> to 25 F<sup>2</sup>.

[0127]

Note that the structure of a semiconductor device according to the disclosed invention is not limited to that illustrated in FIGS. 4A to 4C. Since the technical idea of the disclosed invention is to form a stacked-layer structure with an oxide semiconductor and a material other than an oxide semiconductor, the details of connection relationship of electrodes or the like can be changed as appropriate.

[0128]

<Example of Method for Manufacturing Semiconductor Device>

Next, an example of a method for manufacturing the semiconductor device will be described with reference to FIGS. 5A to 5D and FIGS. 6A to 6C. Note that a method for manufacturing the transistor 162 is similar to that in the above embodiment; thus, a method for manufacturing the transistor 160 will be mainly described here.

[0129]

First, the substrate 100 including a semiconductor material is prepared (see FIG. 5A). A single crystal semiconductor substrate or a polycrystalline semiconductor substrate of silicon, silicon carbide, or the like, a compound semiconductor substrate of silicon germanium or the like, an SOI substrate, or the like can be used as the substrate 100 including a semiconductor material. Here, an example of the case where a single crystal silicon substrate is used as the substrate 100 including a semiconductor material is described. Note that the term “SOI substrate” generally means a substrate where a silicon semiconductor layer is provided over an insulating surface. In this specification, the term “SOI substrate” also means a substrate where a semiconductor layer including a material other than silicon is provided over an insulating surface. That is, a semiconductor layer included in the “SOI substrate” is not limited to a silicon semiconductor layer. Moreover, the SOI substrate can be a substrate having a structure where a semiconductor layer is provided over an insulating substrate such as a glass substrate with an insulating layer interposed therebetween.

[0130]

It is preferable that a single crystal semiconductor substrate of silicon or the like be particularly used as the substrate 100 including a semiconductor material because the speed of reading operation of the semiconductor device can be increased.

[0131]

Note that an impurity element may be added to a region which later functions as the channel formation region 116 of the transistor 160, in order to control the threshold voltage of the transistor. Here, an impurity element imparting conductivity is added so that the threshold voltage of the transistor 160 becomes positive. When the 5 semiconductor material is silicon, the impurity imparting conductivity may be boron, aluminum, gallium, or the like. Note that it is preferable to perform heat treatment after adding an impurity element, in order to activate the impurity element or reduce defects which may be generated during addition of the impurity element.

[0132]

10 Next, a protective layer 102 serving as a mask for forming an element isolation insulating layer is formed over the substrate 100 (see FIG. 5A). As the protective layer 102, an insulating layer formed using a material such as silicon oxide, silicon nitride, silicon oxynitride, or the like can be used, for example.

[0133]

15 Next, part of the substrate 100 in a region not covered with the protective layer 102 (i.e., in an exposed region) is removed by etching using the protective layer 102 as a mask. Thus, a semiconductor region 104 isolated from other semiconductor regions is formed (see FIG. 5B). As the etching, dry etching is preferably performed, but wet etching may be performed. An etching gas or an etchant can be selected as appropriate 20 depending on a material to be etched.

[0134]

25 Then, an insulating layer is formed so as to cover the semiconductor region 104, and the insulating layer in a region overlapping with the semiconductor region 104 is selectively removed; thus, the element isolation insulating layer 106 is formed (see FIG. 5C). The insulating layer is formed using silicon oxide, silicon nitride, silicon oxynitride, or the like. As a method for removing the insulating layer, any of etching treatment, polishing treatment such as chemical mechanical polishing (CMP) treatment, and the like can be employed. Note that the protective layer 102 is removed after the formation of the semiconductor region 104 or after the formation of the element isolation 30 insulating layer 106.

[0135]

Next, an insulating layer is formed over a surface of the semiconductor region

104, and a layer including a conductive material is formed over the insulating layer.

[0136]

The insulating layer is processed into a gate insulating layer later and can be formed by heat treatment (thermal oxidation treatment, thermal nitridation treatment, or the like) of the surface of the semiconductor region 104, for example. Instead of heat treatment, high-density plasma treatment may be employed. The high-density plasma treatment can be performed using, for example, a mixed gas of any of a rare gas such as helium (He), argon (Ar), krypton (Kr), or xenon (Xe), oxygen, nitrogen oxide, ammonia, nitrogen, hydrogen, and the like. It is needless to say that the insulating layer may be formed by a CVD method, a sputtering method, or the like. The insulating layer preferably has a single-layer structure or a stacked-layer structure with a film including silicon oxide, silicon oxynitride, silicon nitride, hafnium oxide, aluminum oxide, tantalum oxide, yttrium oxide, hafnium silicate ( $\text{HfSi}_x\text{O}_y$  ( $x > 0, y > 0$ ))), hafnium silicate to which nitrogen is added ( $\text{HfSi}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ ))), hafnium aluminate to which nitrogen is added ( $\text{HfAl}_x\text{O}_y\text{N}_z$  ( $x > 0, y > 0, z > 0$ ))), or the like. The insulating layer can have a thickness of 1 nm to 100 nm, preferably, 10 nm to 50 nm, for example.

[0137]

The layer including a conductive material can be formed using a metal material such as aluminum, copper, titanium, tantalum, or tungsten. The layer including a conductive material may be formed using a semiconductor material such as polycrystalline silicon. There is no particular limitation on the method for forming the layer including a conductive material, and a variety of film formation methods such as an evaporation method, a CVD method, a sputtering method, or a spin coating method can be employed. Note that this embodiment shows an example of the case where the layer including a conductive material is formed using a metal material.

[0138]

After that, the insulating layer and the layer including a conductive material are selectively etched; thus, the gate insulating layer 108 and the gate electrode 110 are formed (see FIG. 5C).

30 [0139]

Next, phosphorus (P), arsenic (As), or the like is added to the semiconductor

region 104, whereby the channel formation region 116 and the impurity regions 120 are formed (see FIG. 5D). Note that phosphorus or arsenic is added here in order to form an n-type transistor; an impurity element such as boron (B) or aluminum (Al) may be added in the case of forming a p-type transistor. Here, the concentration of the impurity added 5 can be set as appropriate; the concentration is preferably set high when a semiconductor element is highly miniaturized.

[0140]

Note that a sidewall insulating layer may be formed around the gate electrode 110, and impurity regions to which the impurity element is added at a different 10 concentration may be formed.

[0141]

Next, a metal layer 122 is formed so as to cover the gate electrode 110, the impurity regions 120, and the like (see FIG. 6A). The metal layer 122 can be formed by a variety of film formation methods such as a vacuum evaporation method, a sputtering 15 method, and a spin coating method. The metal layer 122 is preferably formed using a metal material which forms a low-resistance metal compound by reacting with the semiconductor material included in the semiconductor region 104. Examples of such metal materials are titanium, tantalum, tungsten, nickel, cobalt, platinum, and the like.

[0142]

20 Next, heat treatment is performed so that the metal layer 122 reacts with the semiconductor material. Thus, the metal compound regions 124 that are in contact with the impurity regions 120 are formed (see FIG. 6A). Note that when the gate electrode 110 is formed using polycrystalline silicon or the like, a metal compound region is also formed in a portion of the gate electrode 110 which is in contact with the metal layer 122.

25 [0143]

As the heat treatment, irradiation with a flash lamp can be employed, for example. Although it is needless to say that another heat treatment method may be used, a method by which heat treatment can be achieved in an extremely short time is 30 preferably used in order to improve the controllability of chemical reaction for formation of the metal compound. Note that the metal compound regions are formed by reaction of the metal material and the semiconductor material and have sufficiently high conductivity. The formation of the metal compound regions can properly reduce the

electric resistance and improve element characteristics. Note that the metal layer 122 is removed after the metal compound regions 124 are formed.

[0144]

Next, the insulating layer 130 is formed so as to cover the components formed in the above steps (see FIG. 6B). The insulating layer 130 can be formed using an inorganic insulating material such as silicon oxide, silicon oxynitride, silicon nitride, or aluminum oxide. It is particularly preferable to use a low dielectric constant (low-k) material for the insulating layer 130 because capacitance due to overlap of electrodes or wirings can be sufficiently reduced. Note that a porous insulating layer with such a material may be employed as the insulating layer 130. The porous insulating layer has a lower dielectric constant than an insulating layer with high density and thus makes it possible to further reduce capacitance due to electrodes or wirings. Alternatively, the insulating layer 130 can be formed using an organic insulating material such as a polyimide or an acrylic resin. Note that although a single-layer structure of the insulating layer 130 is used in this embodiment, an embodiment of the disclosed invention is not limited to this example. A stacked-layer structure with two or more layers may be employed.

[0145]

Through the above steps, the transistor 160 is formed with the use of the substrate 100 including a semiconductor material (see FIG. 6B). A feature of the transistor 160 is that it can operate at high speed. With the use of that transistor as a transistor for reading, data can be read at high speed.

[0146]

After that, as treatment performed before the transistor 162 and the capacitor 164 are formed, CMP treatment of the insulating layer 130 is performed so that an upper surface of the gate electrode 110 is exposed (see FIG. 6C). As treatment for exposing the upper surface of the gate electrode 110, etching treatment or the like can also be employed instead of CMP treatment; in order to improve characteristics of the transistor 162, a surface of the insulating layer 130 is preferably made as flat as possible. For example, the insulating layer 130 is planarized such that its surface has a root-mean-square (RMS) roughness of 1 nm or less. This makes it possible to improve characteristics of a semiconductor device formed over the insulating layer 130.

[0147]

Note that before or after each of the above steps, a step of forming an electrode, a wiring, a semiconductor layer, an insulating layer, or the like may be further performed. For example, when the wiring has a multi-layer structure of a stacked-layer structure including insulating layers and conductive layers, a highly integrated semiconductor device can be realized.

[0148]

After that, the transistor 162 and the capacitor 164 are formed; thus, the semiconductor device is completed.

10 [0149]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0150]

15 (Embodiment 3)

In this embodiment, an example of application of a semiconductor device according to one embodiment of the disclosed invention will be described with reference to FIGS. 7A-1, 7A-2, and 7B. Here, an example of a memory device is described. Note that in some circuit diagrams mentioned below, "OS" is written beside a transistor in 20 order to indicate that the transistor includes an oxide semiconductor.

[0151]

In a semiconductor device which can be used as a memory device, which is illustrated in FIG. 7A-1, a first wiring (1st Line) is electrically connected to a source electrode of a transistor 1000. A second wiring (2nd Line) is electrically connected to a drain electrode of the transistor 1000. A third wiring (3rd Line) is electrically connected to one of a source electrode and a drain electrode of a transistor 1010. A fourth wiring (4th Line) is electrically connected to a gate electrode of the transistor 1010. Furthermore, a gate electrode of the transistor 1000 and the other of the source electrode and the drain electrode of the transistor 1010 are electrically connected to one electrode 25 of a capacitor 1020. A fifth wiring (5th Line) is electrically connected to the other electrode of the capacitor 1020.

[0152]

Here, a transistor including an oxide semiconductor is used as the transistor 1010. Any of the transistors described in the above embodiments can be used as the transistor including an oxide semiconductor. A transistor including an oxide semiconductor has a characteristic of a significantly small off-state current. For that reason, a potential of the 5 gate electrode of the transistor 1000 can be held for an extremely long period even if the transistor 1010 is turned off. Furthermore, with the use of any of the transistors described in the above embodiments, the short-channel effect of the transistor 1010 can be suppressed, and miniaturization can be achieved. By providing the capacitor 1020, holding of charge applied to the gate electrode of the transistor 1000 and reading of data 10 held can be performed more easily. Here, the capacitor described in the above embodiment can be used as the capacitor 1020, for example.

[0153]

A transistor including a semiconductor material other than an oxide semiconductor is used as the transistor 1000. As the semiconductor material other than 15 an oxide semiconductor, for example, silicon, germanium, silicon germanium, silicon carbide, gallium arsenide, or the like can be used, and a single crystal semiconductor is preferably used. Alternatively, an organic semiconductor material or the like may be used. A transistor including such a semiconductor material can operate at high speed. Here, any of the transistors described in the above embodiments can be used as the 20 transistor including a semiconductor material other than an oxide semiconductor, for example.

[0154]

Alternatively, a structure in which the capacitor 1020 is not provided is also possible as illustrated in FIG. 7B.

25 [0155]

The semiconductor device in FIG. 7A-1 utilizes a characteristic in which the potential of the gate electrode of the transistor 1000 can be held, and can thus write, hold, and read data as follows.

[0156]

30 First of all, writing and holding of data will be described. First, the potential of the fourth wiring is set to a potential at which the transistor 1010 is turned on, so that the transistor 1010 is turned on. Accordingly, the potential of the third wiring is supplied to

the gate electrode of the transistor 1000 and to the capacitor 1020. That is, predetermined charge is supplied to the gate electrode of the transistor 1000 (writing). Here, one of two kinds of charges providing different potentials (hereinafter, a charge providing a low potential is referred to as charge  $Q_L$  and a charge providing a high potential is referred to as charge  $Q_H$ ) is applied. Note that three or more kinds of charges providing different potentials may be applied in order to improve storage capacity. After that, the potential of the fourth wiring is set to a potential at which the transistor 1010 is turned off, so that the transistor 1010 is turned off. Thus, the charge supplied to the gate electrode of the transistor 1000 is held (holding).

10 [0157]

Since the off-state current of the transistor 1010 is significantly small, the charge of the gate electrode of the transistor 1000 is held for a long time.

[0158]

Next, reading of data will be described. By supplying an appropriate potential (reading potential) to the fifth wiring while a predetermined potential (a constant potential) is supplied to the first wiring, the potential of the second wiring varies depending on the amount of charge held at the gate electrode of the transistor 1000. This is generally because, when the transistor 1000 is an n-channel transistor, an apparent threshold voltage  $V_{th\_H}$  in the case where  $Q_H$  is supplied to the gate electrode of the transistor 1000 is lower than an apparent threshold voltage  $V_{th\_L}$  in the case where  $Q_L$  is supplied to the gate electrode of the transistor 1000. Here, an apparent threshold voltage refers to the potential of the fifth wiring, which is needed to turn on the transistor 1000. Thus, the potential of the fifth wiring is set to a potential  $V_0$  intermediate between  $V_{th\_H}$  and  $V_{th\_L}$ , whereby charge supplied to the gate electrode of the transistor 1000 can be determined. For example, in the case where  $Q_H$  is supplied in writing, when the potential of the fifth wiring is  $V_0 (> V_{th\_H})$ , the transistor 1000 is turned on. In the case where  $Q_L$  is supplied in writing, even when the potential of the fifth wiring is  $V_0 (< V_{th\_L})$ , the transistor 1000 remains off. Therefore, the data held can be read by measuring the potential of the second wiring.

30 [0159]

Note that in the case where memory cells are arrayed to be used, it is necessary

that only data of a desired memory cell can be read. In order to read data of a predetermined memory cell and not to read data of the other memory cells, in the case where the transistors 1000 are connected in parallel among the memory cells, fifth wirings in memory cells that are not a target for reading are supplied with a potential at 5 which the transistors 1000 are turned off regardless of the state of the gate electrodes, that is, a potential lower than  $V_{th\_H}$ . On the other hand, in the case where the transistors 1000 are connected in series among the memory cells, fifth wirings in memory cells that are not a target for reading are supplied with a potential at which the transistors 1000 are turned on regardless of the state of the gate electrodes, that is, a potential higher than  $V_{th\_L}$ .

10 [0160]

Next, rewriting of data will be described. Rewriting of data is performed in a manner similar to that of the writing and holding of data. That is, the potential of the fourth wiring is set to a potential at which the transistor 1010 is turned on, so that the transistor 1010 is turned on. Accordingly, the potential of the third wiring (a potential 15 for new data) is supplied to the gate electrode of the transistor 1000 and to the capacitor 1020. After that, the potential of the fourth wiring is set to a potential at which the transistor 1010 is turned off, so that the transistor 1010 is turned off. Accordingly, charge for new data is supplied to the gate electrode of the transistor 1000.

[0161]

20 In the semiconductor device according to the disclosed invention, data can be directly rewritten by another writing of data as described above. Therefore, extraction of charge from a floating gate with the use of a high voltage which is necessary for a flash memory or the like is not needed, and thus a decrease in operation speed due to erasing operation can be suppressed. In other words, high-speed operation of the 25 semiconductor device can be realized.

[0162]

Note that the source electrode or the drain electrode of the transistor 1010 is 30 electrically connected to the gate electrode of the transistor 1000 and therefore has a function similar to that of a floating gate of a floating gate transistor used for a nonvolatile memory element. Therefore, in drawings, a portion where the source electrode or the drain electrode of the transistor 1010 is electrically connected to the gate electrode of the

transistor 1000 is called a floating gate portion FG in some cases. When the transistor 1010 is turned off, the floating gate portion FG can be regarded as being embedded in an insulator and thus charge is held in the floating gate portion FG. The off-state current of the transistor 1010 including an oxide semiconductor is smaller than or equal to 1/100000 5 of the off-state current of a transistor including a silicon semiconductor or the like; thus, loss of the charge accumulated in the floating gate portion FG due to leakage of the transistor 1010 is negligible. That is, with the transistor 1010 including an oxide semiconductor, a nonvolatile memory device which can hold data without being supplied with power can be realized.

10 [0163]

For example, when the off-state current of the transistor 1010 at room temperature is 10 zA (1 zA (zeptoampere) is  $1 \times 10^{-21}$  A) or less and the capacitance of the capacitor 1020 is approximately 10 fF, data can be held for  $10^4$  seconds or longer. It is needless to say that the holding time depends on transistor characteristics and 15 capacitance.

[0164]

Further, in that case, the problem of deterioration of a gate insulating film (tunnel insulating film), which is a problem of a conventional floating gate transistor, does not exist. That is, the problem of deterioration of a gate insulating film due to injection of 20 electrons into a floating gate, which is a conventional problem, can be solved. This means that there is no limit on the number of write cycles in principle. Furthermore, a high voltage needed for writing or erasing in a conventional floating gate transistor is not necessary.

[0165]

25 Components such as transistors in the semiconductor device in FIG. 7A-1 can be regarded as including resistors and capacitors as illustrated in FIG. 7A-2. That is, in FIG. 7A-2, the transistor 1000 and the capacitor 1020 are each regarded as including a resistor and a capacitor. R1 and C1 denote the resistance and the capacitance of the capacitor 1020, respectively. The resistance R1 corresponds to the resistance of the insulating 30 layer included in the capacitor 1020. R2 and C2 denote the resistance and the capacitance of the transistor 1000, respectively. The resistance R2 corresponds to the resistance of the gate insulating layer at the time when the transistor 1000 is turned on.

The capacitance C2 corresponds to a so-called gate capacitance (capacitance formed between the gate electrode and the source or drain electrode, and capacitance formed between the gate electrode and the channel formation region).

[0166]

5 A charge holding period (also referred to as a data holding period) is determined mainly by the off-state current of the transistor 1010 under the conditions where the gate leakage current of the transistor 1010 is sufficiently small and R1 and R2 satisfy  $R1 \geq ROS$  ( $R1$  is greater than or equal to  $ROS$ ) and  $R2 \geq ROS$  ( $R2$  is greater than or equal to  $ROS$ ), where  $ROS$  is the resistance (also referred to as effective resistance) between the 10 source electrode and the drain electrode in a state where the transistor 1010 is turned off.

[0167]

15 On the other hand, in the case where the above conditions are not satisfied, it is difficult to secure a sufficient holding period even if the off-state current of the transistor 1010 is sufficiently small. This is because leakage current other than the off-state current of the transistor 1010 (e.g., leakage current generated between the source electrode and the gate electrode) is large. Accordingly, it can be said that the semiconductor device disclosed in this embodiment preferably satisfies the above relationships.

[0168]

20 Meanwhile, it is desirable that C1 and C2 satisfy  $C1 \geq C2$  ( $C1$  is greater than or equal to  $C2$ ). This is because if  $C1$  is larger than or equal to  $C2$ , when the potential of the floating gate portion FG is controlled by the fifth wiring, the potential of the fifth wiring can be efficiently supplied to the floating gate portion FG and the difference between potentials supplied to the fifth wiring (e.g., a reading potential and a non-reading potential) can be kept small.

[0169]

30 When the above relationships are satisfied, a more favorable semiconductor device can be realized. Note that R1 and R2 depend on the gate insulating layer of the transistor 1000 and the insulating layer of the capacitor 1020. The same dependence applies to C1 and C2. Therefore, the material, the thickness, and the like of the gate insulating layer are preferably set as appropriate to satisfy the above relationships.

[0170]

In the semiconductor device described in this embodiment, the floating gate portion FG has a function similar to a floating gate of a floating gate transistor of a flash memory or the like, but the floating gate portion FG of this embodiment has a feature which is essentially different from that of the floating gate of the flash memory or the like. In the case of a flash memory, since a high potential is applied to a control gate, it is necessary to keep a proper distance between cells in order to prevent the potential of the control gate from affecting a floating gate of an adjacent cell. This is one factor inhibiting higher integration of the semiconductor device. The factor is attributed to a basic principle of a flash memory, in which a tunneling current is generated by applying a high electric field.

[0171]

Further, because of the above principle of a flash memory, deterioration of an insulating film proceeds and thus another problem that is the limit on the number of times of rewriting (approximately  $10^4$  to  $10^5$  times) arises.

[0172]

The semiconductor device according to the disclosed invention is operated by switching of a transistor including an oxide semiconductor and does not use the above-described principle of charge injection by a tunneling current. That is, a high electric field for charge injection is not necessary, unlike a flash memory. Accordingly, it is not necessary to consider an influence of a high electric field from a control gate on an adjacent cell, and this facilitates higher integration.

[0173]

Further, charge injection by a tunneling current is not employed, which means that there are no causes for deterioration of a memory cell. In other words, the semiconductor device according to the disclosed invention has higher durability and reliability than a flash memory.

[0174]

In addition, the semiconductor device according to the disclosed invention is advantageous over a flash memory also in that a large step-up circuit or the like is not necessary because a high electric field is not necessary.

[0175]

In the case where the relative permittivity  $\epsilon r1$  of the insulating layer included in the capacitor 1020 is different from the relative permittivity  $\epsilon r2$  of the insulating layer included in the transistor 1000, it is easy to satisfy  $C1 \geq C2$  ( $C1$  is greater than or equal to  $C2$ ) while satisfying  $2 \cdot S2 \geq S1$  ( $2 \cdot S2$  is greater than or equal to  $S1$ ), desirably  $S2 \geq S1$  ( $S2$  is greater than or equal to  $S1$ ), where  $S1$  is the area of the insulating layer included in the capacitor 1020 and  $S2$  is the area of the insulating layer forming a gate capacitor of the transistor 1000. In other words,  $C1$  can easily be made greater than or equal to  $C2$  while the area of the insulating layer included in the capacitor 1020 is small. Specifically, for example, a film including a high-k material such as hafnium oxide or a stack of a film including a high-k material such as hafnium oxide and a film including an oxide semiconductor is used for the insulating layer included in the capacitor 1020 so that  $\epsilon r1$  can be set to 10 or more, preferably 15 or more, and silicon oxide is used for the insulating layer forming the gate capacitor so that  $3 \leq \epsilon r2 \leq 4$  ( $\epsilon r2$  is greater than or equal to 3 and less than or equal to 4).

15 [0176]

A combination of such structures enables further higher integration of the semiconductor device according to the disclosed invention.

[0177]

20 Note that an n-type transistor (n-channel transistor) in which electrons are majority carriers is used in the above description; it is needless to say that a p-type transistor (p-channel transistor) in which holes are majority carriers can be used instead of the n-type transistor.

[0178]

25 As described above, a semiconductor device according to an embodiment of the disclosed invention has a nonvolatile memory cell including a writing transistor where a leakage current (off-state current) between a source and a drain in an off state is small, a reading transistor including a semiconductor material different from that of the writing transistor, and a capacitor.

[0179]

30 With a normal silicon semiconductor, it is difficult to decrease the leakage current (the off-state current) to  $100 \text{ zA}$  ( $1 \times 10^{-19} \text{ A}$ ) or less at ambient temperature

(e.g., 25 °C), whereas this value can be achieved with a transistor including an oxide semiconductor which is processed under appropriate conditions. Therefore, a transistor including an oxide semiconductor is preferably used as the writing transistor.

[0180]

5 In addition, a transistor including an oxide semiconductor has a small subthreshold swing (S value), so that the switching speed can be sufficiently increased even if the mobility is relatively low. Therefore, by using the transistor as the writing transistor, rising of a writing pulse supplied to the floating gate portion FG can be made very sharp. Further, because of such a small off-state current, the amount of charge  
10 required to be held in the floating gate portion FG can be reduced. That is, by using a transistor including an oxide semiconductor as the writing transistor, rewriting of data can be performed at high speed.

[0181]

15 There is no limitation on the off-state current of the reading transistor; it is desirable to use a transistor which operates at high speed in order to increase the reading speed. For example, a transistor with a switching speed of 1 nanosecond or less is preferably used as the reading transistor.

[0182]

20 In this manner, when a transistor including an oxide semiconductor is used as a writing transistor and a transistor including a semiconductor material other than an oxide semiconductor is used as a reading transistor, a semiconductor device capable of holding data for a long time and reading data at high speed, which can be used as a memory device, can be obtained.

[0183]

25 Furthermore, with the use of any of the transistors described in the above embodiments as a writing transistor, the short-channel effect of the writing transistor can be suppressed, and miniaturization can be achieved. Accordingly, a semiconductor device which can be used as a memory device can have higher integration.

[0184]

30 The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0185]

(Embodiment 4)

In this embodiment, an example of application of a semiconductor device according to one embodiment of the disclosed invention will be described with reference to FIGS. 8A and 8B and FIGS. 9A to 9C. Here, an example of a memory device is described. Note that in some circuit diagrams mentioned below, "OS" is written beside a transistor in order to indicate that the transistor includes an oxide semiconductor.

[0186]

FIGS. 8A and 8B are circuit diagrams of semiconductor devices, which can be used as memory devices, each including a plurality of semiconductor devices (hereinafter also referred to as memory cells 1050) illustrated in FIG. 7A-1. FIG. 8A is a circuit diagram of a so-called NAND semiconductor device in which the memory cells 1050 are connected in series, and FIG. 8B is a circuit diagram of a so-called NOR semiconductor device in which the memory cells 1050 are connected in parallel.

[0187]

The semiconductor device in FIG. 8A includes a source line SL, a bit line BL, a first signal line S1,  $m$  second signal lines S2,  $m$  word lines WL, and  $m$  memory cells 1050. In FIG. 8A, one source line SL and one bit line BL are provided in the semiconductor device; however, an embodiment of the disclosed invention is not limited to this structure.

20 A plurality of source lines SL and a plurality of bit lines BL may be provided.

[0188]

In each of the memory cells 1050, the gate electrode of the transistor 1000, one of the source electrode and the drain electrode of the transistor 1010, and one electrode of the capacitor 1020 are electrically connected to one another. The first signal line S1 and the other of the source electrode and the drain electrode of the transistor 1010 are electrically connected to each other, and the second signal line S2 and the gate electrode of the transistor 1010 are electrically connected to each other. The word line WL and the other electrode of the capacitor 1020 are electrically connected to each other.

[0189]

30 Further, the source electrode of the transistor 1000 included in the memory cell 1050 is electrically connected to the drain electrode of the transistor 1000 in the adjacent memory cell 1050. The drain electrode of the transistor 1000 included in the memory

cell 1050 is electrically connected to the source electrode of the transistor 1000 in the adjacent memory cell 1050. Note that the drain electrode of the transistor 1000 included in the memory cell 1050 at one end of the plurality of memory cells connected in series is electrically connected to the bit line BL. The source electrode of the transistor 1000 included in the memory cell 1050 at the other end of the plurality of memory cells connected in series is electrically connected to the source line SL.

5 [0190]

In the semiconductor device in FIG. 8A, writing operation and reading operation are performed for each row. The writing operation is performed as follows. A 10 potential at which the transistor 1010 is turned on is supplied to the second signal line S2 of a row where writing is to be performed, so that the transistor 1010 of the row where writing is to be performed is turned on. Accordingly, a potential of the first signal line S1 is supplied to the gate electrode of the transistor 1000 of the specified row, so that predetermined charge is applied to the gate electrode. Thus, data can be written to the 15 memory cell of the specified row.

15 [0191]

Further, the reading operation is performed as follows. First, a potential at 20 which the transistor 1000 is turned on regardless of charge of the gate electrode thereof is supplied to the word lines WL of the rows other than the row where reading is to be performed, so that the transistors 1000 of the rows other than the row where reading is to be performed are turned on. Then, a potential (reading potential) at which an on state or 25 an off state of the transistor 1000 is determined depending on charge of the gate electrode of the transistor 1000 is supplied to the word line WL of the row where reading is to be performed. After that, a constant potential is supplied to the source line SL so that a reading circuit (not illustrated) connected to the bit line BL is operated. Here, the plurality of transistors 1000 between the source line SL and the bit line BL are turned on except the transistor 1000 of the row where reading is to be performed; therefore, 30 conductance between the source line SL and the bit line BL is determined by the state of the transistor 1000 (whether being turned on or off) of the row where reading is to be performed. Since the conductance of the transistor varies depending on the electric charge in the gate electrode of the transistor 1000 of the row where reading is to be performed, a potential of the bit line BL also varies accordingly. By reading the

potential of the bit line BL with the reading circuit, data can be read from the memory cell of the specified row.

[0192]

The semiconductor device in FIG. 8B includes  $n$  source lines SL,  $n$  bit lines BL,

5  $n$  first signal lines S1,  $m$  second signal lines S2,  $m$  word lines WL, and  $n \times m$  memory cells 1050. A gate electrode of the transistor 1000, one of the source electrode and the drain electrode of the transistor 1010, and one electrode of the capacitor 1020 are electrically connected to one another. The source line SL and the source electrode of the transistor 1000 are electrically connected to each other. The bit line BL and the drain 10 electrode of the transistor 1000 are electrically connected to each other. The first signal line S1 and the other of the source electrode and the drain electrode of the transistor 1010 are electrically connected to each other, and the second signal line S2 and the gate electrode of the transistor 1010 are electrically connected to each other. The word line WL and the other electrode of the capacitor 1020 are electrically connected to each other.

15 [0193]

In the semiconductor device in FIG. 8B, writing operation and reading operation are performed for each row. The writing operation is performed in a manner similar to that of the semiconductor device in FIG. 8A. The reading operation is performed as follows. First, a potential at which the transistor 1000 is turned off regardless of charge 20 of the gate electrode thereof is supplied to the word lines WL of the rows other than the row where reading is to be performed, so that the transistors 1000 of the rows other than the row where reading is to be performed are turned off. Then, a potential (reading potential) at which an on state or an off state of the transistor 1000 is determined depending on charge of the gate electrode of the transistor 1000 is supplied to the word 25 line WL of the row where reading is to be performed. After that, a constant potential is supplied to the source line SL so that a reading circuit (not illustrated) connected to the bit line BL is operated. Here, conductance between the source line SL and the bit line BL is determined by the state of the transistor 1000 (whether being turned on or off) of the row where reading is to be performed. That is, a potential of the bit line BL depends on 30 charge of the gate electrode of the transistor 1000 of the row where reading is to be performed. By reading the potential of the bit line BL with the reading circuit, data can be read from the memory cell of the specified row.

[0194]

Although the amount of data which can be stored in each of the memory cells 1050 is one bit in the above description, the structure of the semiconductor device of this embodiment is not limited to this example. The amount of data which is held in each of the memory cells 1050 may be increased by preparing three or more kinds of potentials to be supplied to the gate electrode of the transistor 1000. For example, in the case where four kinds of potentials are supplied to the gate electrode of the transistor 1000, data of two bits can be held in each of the memory cells.

[0195]

10 Next, an example of a reading circuit which can be used for the semiconductor devices illustrated in FIGS. 8A and 8B and the like will be described with reference to FIGS. 9A to 9C.

[0196]

15 FIG. 9A illustrates an outline of the reading circuit. The reading circuit includes a transistor and a sense amplifier circuit.

[0197]

20 At the time of reading data, a terminal A is connected to a bit line BL to which a memory cell from which data is to be read is connected. Further, a bias potential  $V_{bias}$  is applied to a gate electrode of the transistor so that a potential of the terminal A is controlled.

[0198]

25 The resistance of the memory cell 1050 varies depending on data stored. Specifically, when the transistor 1000 of the memory cell 1050 selected is turned on, the memory cell 1050 has a low resistance, whereas when the transistor 1000 of the memory cell 1050 selected is turned off, the memory cell 1050 has a high resistance.

[0199]

30 When the memory cell has a high resistance, the potential of the terminal A is higher than a reference potential  $V_{ref}$  and the sense amplifier circuit outputs a potential corresponding to the potential of the terminal A. On the other hand, when the memory cell has a low resistance, the potential of the terminal A is lower than the reference potential  $V_{ref}$  and the sense amplifier circuit outputs a potential corresponding to the potential of the terminal A.

[0200]

In this manner, by using the reading circuit, data can be read from the memory cell. Note that the reading circuit of this embodiment is one example. Another circuit may be used. The reading circuit may further include a precharge circuit. Instead of 5 the reference potential Vref, a reference bit line may be connected to the sense amplifier circuit.

[0201]

FIG. 9B illustrates a differential sense amplifier which is an example of sense amplifier circuits. The differential sense amplifier has input terminals  $V_{in}(+)$  and 10  $V_{in}(-)$  and an output terminal  $V_{out}$ , and amplifies a difference between a potential of  $V_{in}(+)$  and a potential of  $V_{in}(-)$ . If the potential of  $V_{in}(+)$  is higher than the potential of  $V_{in}(-)$ , the output from  $V_{out}$  is relatively high, whereas if the potential of  $V_{in}(+)$  is lower than the potential of  $V_{in}(-)$ , the output from  $V_{out}$  is relatively low. In the case 15 where the differential sense amplifier is used for the reading circuit, one of the input terminals  $V_{in}(+)$  and  $V_{in}(-)$  is connected to the terminal A, and the reference potential  $V_{ref}$  is supplied to the other of the input terminals  $V_{in}(+)$  and  $V_{in}(-)$ .

[0202]

FIG. 9C illustrates a latch sense amplifier which is an example of sense amplifier circuits. The latch sense amplifier has input/output terminals  $V_1$  and  $V_2$  and input 20 terminals for control signals  $S_p$  and  $S_n$ . First, the signal  $S_p$  is set high and the signal  $S_n$  is set low, and a power supply potential ( $V_{dd}$ ) is interrupted. Then, potentials to be compared are supplied to  $V_1$  and  $V_2$ . After that, the signal  $S_p$  is set low and the signal  $S_n$  is set high, and the power supply potential ( $V_{dd}$ ) is supplied. If the potentials  $V_{1in}$  25 and  $V_{2in}$  to be compared satisfy  $V_{1in} > V_{2in}$ , the output from  $V_1$  is high and the output from  $V_2$  is low, whereas if the potentials satisfy  $V_{1in} < V_{2in}$ , the output from  $V_1$  is low and the output from  $V_2$  is high. By utilizing such a relationship, the difference between  $V_{1in}$  and  $V_{2in}$  can be amplified. In the case where the latch sense amplifier is used for the reading circuit, one of  $V_1$  and  $V_2$  is connected to the terminal A and an output terminal through a switch, and the reference potential  $V_{ref}$  is supplied to the other of  $V_1$  30 and  $V_2$ .

[0203]

With the use of any of the transistors described in the above embodiments as a writing transistor of a memory cell in the above-described semiconductor device which can be used as a memory device, the short-channel effect of the writing transistor can be suppressed, and miniaturization can be achieved. Accordingly, the semiconductor 5 device which can be used as a memory device can have higher integration.

[0204]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

10 [0205]

(Embodiment 5)

In this embodiment, an example of application of a semiconductor device according to one embodiment of the disclosed invention will be described with reference to FIG. 10. Here, a central processing unit (CPU) is described.

15 [0206]

FIG. 10 illustrates an example of a block diagram of a CPU. A CPU 1101 illustrated in FIG. 10 includes a timing control circuit 1102, an instruction decoder 1103, a register array 1104, an address logic and buffer circuit 1105, a data bus interface 1106, an arithmetic logic unit (ALU) 1107, an instruction register 1108, and the like.

20 [0207]

These circuits are manufactured using any of the transistors described in the above embodiments, an inverter circuit, a resistor, a capacitor, and the like. Because the transistors described in the above embodiments can achieve an extremely small off-state current, a reduction in power consumption of the CPU 1101 can be realized.

25 Furthermore, with the use of any of the transistors described in the above embodiments, the short-channel effect of the transistor can be suppressed, and miniaturization can be achieved.

[0208]

Circuits included in the CPU 1101 will be briefly described below. The 30 timing control circuit 1102 receives instructions from the outside, converts the instructions into information for the inside, and transmits the information to another block. In addition, the timing control circuit 1102 gives directions such as reading and

writing of memory data to the outside, according to internal operation. The instruction decoder 1103 functions to convert instructions from the outside into instructions for the inside. The register array 1104 functions to temporarily store data. The address logic and buffer circuit 1105 functions to specify the address of an external memory. The 5 data bus interface 1106 functions to take data in and out of an external memory or a device such as a printer. The ALU 1107 functions to perform an operation. The instruction register 1108 functions to temporarily store instructions. The CPU includes such a combination of circuits.

[0209]

10 With the use of any of the transistors described in the above embodiments in at least part of the CPU 1101, the short-channel effect of the transistor can be suppressed, and miniaturization can be achieved. Thus, the CPU 1101 can have higher integration.

[0210]

15 The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0211]

(Embodiment 6)

20 In this embodiment, an example of application of a semiconductor device according to one embodiment of the disclosed invention will be described with reference to FIGS. 11A and 11B. Here, an example of a semiconductor device having an image sensor function for reading information of an object will be described. Note that in some circuit diagrams, "OS" is written beside a transistor in order to indicate that the transistor includes an oxide semiconductor.

25 [0212]

FIG. 11A illustrates an example of a semiconductor device having an image sensor function. FIG. 11A is an equivalent circuit diagram of a photosensor, and FIG. 11B is a cross-sectional view of part of the photosensor.

[0213]

30 One electrode of a photodiode 1202 is electrically connected to a photodiode reset signal line 1212, and the other electrode of the photodiode 1202 is electrically connected to a gate electrode of a transistor 1204. One of a source electrode and a drain

electrode of the transistor 1204 is electrically connected to a photosensor reference signal line 1218, and the other of the source electrode and the drain electrode of the transistor 1204 is electrically connected to one of a source electrode and a drain electrode of a transistor 1206. A gate electrode of the transistor 1206 is electrically connected to a 5 gate signal line 1214, and the other of the source electrode and the drain electrode of the transistor 1206 is electrically connected to a photosensor output signal line 1216.

[0214]

Here, transistors including an oxide semiconductor are used as the transistor 1204 and the transistor 1206 illustrated in FIG. 11A. Here, any of the transistors 10 described in the above embodiments can be used as the transistors including an oxide semiconductor. Because the transistors described in the above embodiments can achieve an extremely small leakage current in an off state, the photodetection accuracy of the photosensor can be improved. Furthermore, with the use of any of the transistors described in the above embodiments, the short-channel effect of the transistor 15 can be suppressed, and miniaturization can be achieved. Thus, the area of the photodiode can be increased, and the photodetection accuracy of the photosensor can be improved.

[0215]

FIG. 11B is a cross-sectional view illustrating the photodiode 1202 and the 20 transistor 1204 in the photosensor. The photodiode 1202 functioning as a sensor and the transistor 1204 are provided over a substrate 1222 having an insulating surface (a TFT substrate). A substrate 1224 is provided over the photodiode 1202 and the transistor 1204 using an adhesive layer 1228. An insulating layer 1234, an interlayer insulating layer 1236, and an interlayer insulating layer 1238 are provided over the transistor 1204.

25 [0216]

A gate electrode layer 1240 is provided in the same layer as the gate electrode of the transistor 1204 so as to be electrically connected to the gate electrode. The gate electrode layer 1240 is electrically connected to an electrode layer 1242 provided over the interlayer insulating layer 1236, through an opening formed in the insulating layer 30 1234 and the interlayer insulating layer 1236. Because the photodiode 1202 is formed over the electrode layer 1242, the photodiode 1202 and the transistor 1204 are electrically connected to each other through the gate electrode layer 1240 and the

electrode layer 1242.

[0217]

The photodiode 1202 has a structure in which a first semiconductor layer 1226a, a second semiconductor layer 1226b, and a third semiconductor layer 1226c are stacked in this order over the electrode layer 1242. In other words, the first semiconductor layer 1226a of the photodiode 1202 is electrically connected to the electrode layer 1242. The third semiconductor layer 1226c of the photodiode 1202 is electrically connected to an electrode layer 1244 provided over the interlayer insulating layer 1238.

10 [0218]

Here, a PIN photodiode is given as an example, in which a semiconductor layer having n-type conductivity as the first semiconductor layer 1226a, a high-resistance semiconductor layer (an i-type semiconductor layer) as the second semiconductor layer 1226b, and a semiconductor layer having p-type conductivity as the third semiconductor 15 layer 1226c are stacked.

[0219]

The first semiconductor layer 1226a is an n-type semiconductor layer and is formed with an amorphous silicon film containing an impurity element imparting n-type conductivity. The first semiconductor layer 1226a is formed by a plasma CVD method 20 with the use of a semiconductor source gas containing an impurity element belonging to Group 15 (such as phosphorus (P)). As the semiconductor source gas, silane ( $SiH_4$ ) may be used. Alternatively, disilane ( $Si_2H_6$ ), dichlorosilane ( $SiH_2Cl_2$ ), trichlorosilane ( $SiHCl_3$ ), silicon tetrachloride ( $SiCl_4$ ), silicon tetrafluoride ( $SiF_4$ ), or the like may be used. Alternatively, an amorphous silicon film which does not contain an impurity element may 25 be formed, and then, an impurity element may be introduced into the amorphous silicon film by a diffusion method or an ion implantation method. After the impurity element is introduced by an ion implantation method or the like, heating or the like may be conducted in order to diffuse the impurity element. In this case, as a method for forming the amorphous silicon film, an LPCVD method, a vapor deposition method, a sputtering 30 method, or the like may be used. The first semiconductor layer 1226a is preferably formed so as to have a thickness of 20 nm to 200 nm.

[0220]

The second semiconductor layer 1226b is an i-type semiconductor layer (an intrinsic semiconductor layer) and is formed with an amorphous silicon film. As the second semiconductor layer 1226b, an amorphous silicon film is formed by a plasma CVD method with use of a semiconductor source gas. As the semiconductor source gas, 5 silane (SiH<sub>4</sub>) may be used. Alternatively, Si<sub>2</sub>H<sub>6</sub>, SiH<sub>2</sub>Cl<sub>2</sub>, SiHCl<sub>3</sub>, SiCl<sub>4</sub>, SiF<sub>4</sub>, or the like may be used. The second semiconductor layer 1226b may alternatively be formed by an LPCVD method, a vapor deposition method, a sputtering method, or the like. The second semiconductor layer 1226b is preferably formed so as to have a thickness of 200 nm to 1000 nm.

10 [0221]

The third semiconductor layer 1226c is a p-type semiconductor layer and is formed with an amorphous silicon film containing an impurity element imparting p-type conductivity. The third semiconductor layer 1226c is formed by a plasma CVD method with the use of a semiconductor source gas containing an impurity element belonging to 15 Group 13 (such as boron (B)). As the semiconductor source gas, silane (SiH<sub>4</sub>) may be used. Alternatively, Si<sub>2</sub>H<sub>6</sub>, SiH<sub>2</sub>Cl<sub>2</sub>, SiHCl<sub>3</sub>, SiCl<sub>4</sub>, SiF<sub>4</sub>, or the like may be used. Alternatively, an amorphous silicon film which does not contain an impurity element may be formed, and then, an impurity element may be introduced into the amorphous silicon film by a diffusion method or an ion implantation method. After the impurity element is 20 introduced by an ion implantation method or the like, heating or the like may be performed in order to diffuse the impurity element. In this case, as a method for forming the amorphous silicon film, an LPCVD method, a vapor deposition method, a sputtering method, or the like may be used. The third semiconductor layer 1226c is preferably formed so as to have a thickness of 10 nm to 50 nm.

25 [0222]

The first semiconductor layer 1226a, the second semiconductor layer 1226b, and the third semiconductor layer 1226c are not necessarily formed using an amorphous semiconductor, and they may be formed using a polycrystalline semiconductor or a microcrystalline semiconductor (or a semi-amorphous semiconductor (SAS)).

30 [0223]

The microcrystalline semiconductor belongs to a metastable state which is an intermediate state between an amorphous state and a single crystal state according to

Gibbs free energy. That is, the microcrystalline semiconductor is a semiconductor having a third state which is thermodynamically stable and has a short range order and lattice distortion. In the microcrystalline semiconductor, columnar or needle-like crystals grow in a normal direction with respect to a surface of a substrate. The Raman spectrum of microcrystalline silicon, which is a typical example of the microcrystalline semiconductor, is shifted to a smaller wavenumber region than  $520\text{ cm}^{-1}$  which represents single crystal silicon. That is, the peak of the Raman spectrum of microcrystalline silicon exists between  $520\text{ cm}^{-1}$  which represents single crystal silicon and  $480\text{ cm}^{-1}$  which represents amorphous silicon. The microcrystalline semiconductor includes at least 1 at.% of hydrogen or halogen to terminate a dangling bond. Moreover, a rare gas element such as helium, argon, krypton, or neon may be included to further promote lattice distortion, so that a favorable microcrystalline semiconductor film with enhanced stability can be obtained.

[0224]

This microcrystalline semiconductor film can be formed by a high-frequency plasma CVD method with a frequency of several tens to several hundreds of megahertz or a microwave plasma CVD method with a frequency of 1 GHz or more. Typically, the microcrystalline semiconductor film can be formed by using a gas obtained by diluting a gas including silicon, such as  $\text{SiH}_4$ ,  $\text{Si}_2\text{H}_6$ ,  $\text{SiH}_2\text{Cl}_2$ ,  $\text{SiHCl}_3$ ,  $\text{SiCl}_4$ , or  $\text{SiF}_4$ , with hydrogen. Alternatively, the microcrystalline semiconductor film can be formed by using a gas including silicon which is diluted with hydrogen and one or more rare gas elements selected from helium, argon, krypton, and neon. In this case, the flow rate of hydrogen is set 5 times to 200 times, preferably 50 times to 150 times, more preferably 100 times, as high as that of a gas including silicon. Furthermore, a gas including silicon may be mixed with a hydrocarbon gas such as  $\text{CH}_4$  or  $\text{C}_2\text{H}_6$ , a gas including germanium gas such as  $\text{GeH}_4$  or  $\text{GeF}_4$ ,  $\text{F}_2$ , or the like.

[0225]

The mobility of holes generated by the photoelectric effect is lower than the mobility of electrons. Therefore, a PIN photodiode has better characteristics when a surface on the p-type semiconductor layer side is used as a light-receiving plane. Here, an example where the photodiode 1202 receives incident light 1230 from the substrate 1224 side and converts it into electric signals is described. Further, light from a side on

which the semiconductor layer having a conductivity type opposite to that of the semiconductor layer on the light-receiving plane side is disturbance light; therefore, the electrode layer 1242 is preferably formed using a light-blocking conductive film. Note that the n-type semiconductor layer side may alternatively be a light-receiving plane.

5 [0226]

When the incident light 1230 enters from the substrate 1224 side, the oxide semiconductor layer of the transistor 1204 can be shielded from the incident light 1230 by the gate electrode of the transistor 1204.

[0227]

10 The insulating layer 1234, the interlayer insulating layer 1236, and the interlayer insulating layer 1238 can be formed using an insulating material by a method such as a sputtering method, an SOG method, a spin coating method, a dip-coating method, a spray coating method, a screen printing method, an offset printing method or a droplet discharge method (e.g., an inkjet method) depending on the material.

15 [0228]

The insulating layer 1234 may be a single layer or stacked layers of an inorganic insulating material, with any of oxide insulating layers or nitride insulating layers such as a silicon oxide layer, a silicon oxynitride layer, a silicon nitride layer, a silicon nitride oxide layer, an aluminum oxide layer, an aluminum oxynitride layer, an 20 aluminum nitride layer, or an aluminum nitride oxide layer. In addition, a high-quality insulating layer which is dense and has high withstand voltage can be formed by a high-density plasma CVD method using microwaves (2.45 GHz), which is preferable.

[0229]

For a reduction of the surface roughness, an insulating layer functioning as a 25 planarization insulating film is preferably used as the interlayer insulating layers 1236 and 1238. The interlayer insulating layers 1236 and 1238 can be formed using an organic insulating material having heat resistance such as a polyimide, an acrylic resin, benzocyclobutene-based resin, a polyamide, or an epoxy resin. Other than such organic insulating materials, it is possible to use a single layer or stacked layers of a low dielectric 30 constant material (a low-k material), a siloxane-based resin, phosphosilicate glass (PSG), borophosphosilicate glass (BPSG), or the like.

[0230]

The photodiode 1202 can read information of an object by detecting the incident light 1230. Note that a light source such as a backlight can be used at the time of reading information of an object.

[0231]

5 In the photosensor described above, any of the transistors described in the above embodiments can be used as the transistor including an oxide semiconductor. Because the transistors described in the above embodiments can achieve an extremely small leakage current in an off state, the photodetection accuracy of the photosensor can be improved. Furthermore, with the use of any of the transistors described in the 10 above embodiments, the short-channel effect of the transistor can be suppressed, and miniaturization can be achieved. Thus, the area of the photodiode can be increased, and the photodetection accuracy of the photosensor can be improved.

[0232]

15 The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0233]

(Embodiment 7)

20 In this embodiment, the cases where any of the semiconductor devices described in the above embodiments is applied to electronic devices will be described with reference to FIGS. 12A to 12F. The cases where any of the above-described semiconductor devices is applied to electronic devices such as a computer, a mobile phone set (also referred to as a mobile phone or a mobile phone device), a portable information terminal (including a portable game machine, an audio reproducing device, 25 and the like), a digital camera, a digital video camera, electronic paper, a television set (also referred to as a television or a television receiver), and the like are described in this embodiment.

[0234]

30 FIG. 12A illustrates a notebook personal computer, which includes a housing 701, a housing 702, a display portion 703, a keyboard 704, and the like. At least one of the housings 701 and 702 is provided with any of the semiconductor devices described in the above embodiments. Therefore, a notebook personal computer which operates at high

speed and consumes less power, for example, can be realized.

[0235]

FIG. 12B illustrates a portable information terminal (PDA). A main body 711 is provided with a display portion 713, an external interface 715, operation buttons 714, and the like. Further, a stylus 712 for operation of the portable information terminal, or the like is provided. The main body 711 is provided with any of the semiconductor devices described in the above embodiments. Therefore, a portable information terminal which operates at high speed and consumes less power, for example, can be realized.

10 [0236]

FIG. 12C illustrates an electronic book 720 incorporating electronic paper, which includes two housings, a housing 721 and a housing 723. The housing 721 and the housing 723 include a display portion 725 and a display portion 727, respectively. The housing 721 is connected to the housing 723 by a hinge 737, so that the electronic book 15 720 can be opened and closed using the hinge 737 as an axis. In addition, the housing 721 is provided with a power switch 731, operation keys 733, a speaker 735, and the like. At least one of the housings 721 and 723 is provided with any of the semiconductor devices described in the above embodiments. Therefore, an electronic book which operates at high speed and consumes less power, for example, can be realized.

20 [0237]

FIG. 12D illustrates a mobile phone set, which includes two housings, a housing 740 and a housing 741. The housings 740 and 741 in a state where they are developed as illustrated in FIG. 12D can be slid so that one is lapped over the other. Therefore, the size of the mobile phone set can be reduced, which makes the mobile phone set suitable 25 for being carried around. The housing 741 includes a display panel 742, a speaker 743, a microphone 744, a touch panel 745, a pointing device 746, a camera lens 747, an external connection terminal 748, and the like. The housing 740 includes a solar cell 749 for charging the mobile phone set, an external memory slot 750, and the like. An antenna is incorporated in the housing 741. At least one of the housings 740 and 741 is 30 provided with any of the semiconductor devices described in the above embodiments. Therefore, a mobile phone set which operates at high speed and consumes less power, for example, can be realized.

[0238]

FIG. 12E illustrates a digital camera, which includes a main body 761, a display portion 767, an eyepiece 763, an operation switch 764, a display portion 765, a battery 766, and the like. The main body 761 is provided with any of the semiconductor devices described in the above embodiments. Therefore, a digital camera which operates at high speed and consumes less power, for example, can be realized.

[0239]

FIG. 12F is a television set 770, which includes a housing 771, a display portion 773, a stand 775, and the like. The television set 770 can be operated with a switch included in the housing 771 or with a remote controller 780. The housing 771 and the remote controller 780 are provided with any of the semiconductor devices described in the above embodiments. Therefore, a television set which operates at high speed and consumes less power, for example, can be realized.

[0240]

As described above, the electronic devices described in this embodiment each include any of the semiconductor devices according to the above embodiments. Therefore, an electronic device whose operation speed is increased and power consumption is reduced by miniaturization of the semiconductor device can be obtained.

[Example 1]

[0241]

In this example, results of examination, by computational simulation, of the effect of the disclosed invention in terms of suppression of a short-channel effect will be described. Note that calculations (a first calculation and a second calculation) were carried out using a device simulator “Atlas” produced by Silvaco Data Systems Inc.

[0242]

The first calculation was carried out using models illustrated in FIGS. 13A and 13B. FIG. 13A illustrates a transistor having a structure in which there is a small difference in height between an upper surface of an insulating layer and an upper surface of a source electrode and a drain electrode (hereinafter, Structure A), according to one embodiment of the disclosed invention. FIG. 13B illustrates a transistor having a structure in which there is no such difference in height (hereinafter, Structure B), as a

comparative example. Note that a structure in which an insulating layer 143b does not cover a source electrode 142a and a drain electrode 142b as illustrated in FIG. 13A is employed as Structure A for simplicity of the calculation; there is not much computational difference between the structure and a structure in which the insulating

5 layer 143b covers the source electrode 142a and the drain electrode 142b.

[0243]

In the calculation, the inclination angle  $\theta$  of a side surface of a region where the difference in height is made (hereinafter referred to as a projecting region) and the height  $h$  of the projecting region were variables. Note that components denoted by 10 reference numerals in FIGS. 13A and 13B correspond to the components denoted by the same reference numerals and described in the above embodiment. In addition, the calculation models of this example do not include some components such as the insulating layer 150, but this does not affect the result of calculation.

[0244]

15 Other parameters used for the calculation are as follows.

- Thickness of oxide semiconductor layer: 10 nm
- Material of oxide semiconductor layer: In-Ga-Zn-O-based metal oxide (band gap  $Eg$ : 3.15 eV, electron affinity ( $\chi$ ): 4.3 eV, relative permittivity: 15, electron mobility: 10  $\text{cm}^2/\text{Vs}$ )
- 20 • Thickness of gate insulating layer: 10 nm
- Material of gate insulating layer: hafnium oxide (relative permittivity: 15)
- Material of source and drain electrodes: titanium nitride (work function: 3.9 eV)
- Material of gate electrode: tungsten (work function: 4.9 eV)

25 [0245]

The results of the first calculation are shown in FIG. 14, FIG. 15, and FIG. 16. FIG. 14 shows the relationship between gate voltage  $VG$  (V) and drain current  $ID$  (A); FIG. 15, the relationship between channel length  $L$  (nm) and threshold voltage  $Vth$  (V); and FIG. 16, the relationship between channel length  $L$  (nm) and S value (V/dec).

30 Note that each of FIG. 14, FIG. 15, and FIG. 16 shows results obtained with different inclination angles  $\theta$  of 45°, 60°, and 90° and heights  $h$  of 5 nm, 10 nm, and 20 nm.

[0246]

The results in FIG. 14, FIG. 15, and FIG. 16 show that there is no significant difference between Structure A and Structure B when the channel length  $L$  is more than 100 nm, whereas a negative shift in threshold voltage  $V_{th}$  and an increase in S value are suppressed in Structure A when the channel length  $L$  is 100 nm or less. That is, a short-channel effect can be suppressed in Structure A, in comparison with Structure B.

[0247]

The second calculation was carried out using models illustrated in FIGS. 17A and 17B. FIG. 17A illustrates Structure A, and FIG. 17B illustrates Structure B. A difference between the first and second calculations is that the distance between the source electrode and the drain electrode is defined as the channel length  $L$  in the first calculation, whereas the channel length  $L$  in the second calculation is measured along the projecting shape of the insulating layer. In other words, in FIG. 17A, the channel length  $L$  equals  $L_s + L_c + L_d$ . By defining the channel length  $L$  in FIG. 17A in this manner, the effect of an increase in effective value of the channel length  $L$  can be counteracted, and an effect resulting from the shape can be accurately observed.

[0248]

FIG. 18 shows the relationship between gate voltage  $VG$  (V) and drain current  $ID$  (A) which is obtained from the results of the second calculation. Here, the inclination angle  $\theta$  is fixed to 90°, and the height  $h$  is set to 5 nm, 10 nm, and 20 nm. FIG. 18 shows that a negative shift in threshold voltage  $V_{th}$  is suppressed in the structure in which there is a small difference in height between the upper surface of the insulating layer and the upper surface of the source and drain electrode (Structure A), owing to its shape. In other words, it can be understood that the shape suppresses a short-channel effect.

[0249]

Note that it can be seen from the results of the first and second calculations that a short-channel effect is less easily caused as the height  $h$  increases. However, if there is a large difference in height, a decrease in coverage may cause a disconnection of the oxide semiconductor layer or the like. Thus, the difference in height is set to 30 nm or less, preferably 20 nm or less.

This application is based on Japanese Patent Application serial no. 2010-051008 filed with Japan Patent Office on March 8, 2010, the entire contents of which are hereby incorporated by reference.

## CLAIMS

1. A semiconductor device comprising:

an insulating layer;

5 a source electrode and a drain electrode embedded in the insulating layer;  
an oxide semiconductor layer over and in contact with the insulating layer, the  
source electrode, and the drain electrode;  
a gate insulating layer over the oxide semiconductor layer; and  
a gate electrode over the gate insulating layer,  
10 wherein a difference in height exists between an upper surface of the insulating  
layer and each of an upper surface of the source electrode and an upper surface of the  
drain electrode.

2. The semiconductor device according to claim 1, wherein the upper surface of

15 the insulating layer, which is in contact with the oxide semiconductor layer, has a  
root-mean-square roughness of 1 nm or less.

3. The semiconductor device according to claim 1, wherein the difference in

height between the upper surface of the insulating layer and each of the upper surface of  
20 the source electrode and the upper surface of the drain electrode is 5 nm or more.

4. The semiconductor device according to claim 1, wherein the difference in

height between the upper surface of the insulating layer and each of the upper surface of  
the source electrode and the upper surface of the drain electrode is 20 nm or less.

25

5. A semiconductor device comprising:

a first transistor; and

a second transistor over the first transistor,

wherein the first transistor comprises:

30

a first channel formation region;

a first gate insulating layer over the first channel formation region;

a first gate electrode over the first gate insulating layer so as to overlap

with the first channel formation region; and

5 a first source electrode and a first drain electrode electrically connected to the first channel formation region,

wherein the second transistor comprises:

10 a second source electrode and a second drain electrode embedded in an insulating layer;

a second channel formation region over and in contact with the insulating layer, the second source electrode, and the second drain electrode;

15 a second gate insulating layer over the second channel formation region; and

a second gate electrode over the second gate insulating layer,

wherein a difference in height exists between an upper surface of the insulating layer and each of an upper surface of the second source electrode and an upper surface of the second drain electrode.

15

6. The semiconductor device according to claim 5, wherein the upper surface of the insulating layer, which is in contact with the second channel formation region, has a root-mean-square roughness of 1 nm or less.

20

7. The semiconductor device according to claim 5, wherein the difference in height between the upper surface of the insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 5 nm or more.

25

8. The semiconductor device according to claim 5, wherein the difference in height between the upper surface of the insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 20 nm or less.

30

9. The semiconductor device according to claim 5, wherein the second channel formation region comprises an oxide semiconductor layer.

10. The semiconductor device according to claim 5, wherein the upper surface of the second source electrode and the upper surface of the second drain electrode are partly covered with the insulating layer.

5 11. A memory device including the semiconductor device according to claim 5.

12. A method for manufacturing a semiconductor device, comprising the steps of:

forming a source electrode and a drain electrode over a substrate;

10 forming a first insulating layer so as to cover the source electrode and the drain electrode;

performing planarization treatment on a surface of the first insulating layer, thereby forming a second insulating layer having a planarized surface;

15 forming openings reaching the source electrode and the drain electrode in the second insulating layer, so that a difference in height between the planarized surface of the second insulating layer and each of an upper surface of the source electrode and an upper surface of the drain electrode is formed;

forming an oxide semiconductor layer over and in contact with the second insulating layer, the source electrode, and the drain electrode;

20 forming a gate insulating layer over the oxide semiconductor layer; and

forming a gate electrode over the gate insulating layer.

13. The method for manufacturing a semiconductor device according to claim 12, wherein the planarization treatment is performed so that a root-mean-square 25 roughness of the planarized surface is 1 nm or less.

14. The method for manufacturing a semiconductor device according to claim 12, wherein the difference in height between the planarized surface of the second insulating layer and each of the upper surface of the source electrode and the upper 30 surface of the drain electrode is 5 nm or more.

15. The method for manufacturing a semiconductor device according to claim

12, wherein the difference in height between the planarized surface of the second insulating layer and each of the upper surface of the source electrode and the upper surface of the drain electrode is 20 nm or less.

5        16. The method for manufacturing a semiconductor device according to claim 12, wherein the openings are formed so that the upper surface of the source electrode and the upper surface of the drain electrode are partly covered with the second insulating layer.

10        17. A method for manufacturing a semiconductor device, comprising the steps of:

      forming a source electrode and a drain electrode over a substrate;

      forming a first insulating layer so as to cover the source electrode and the drain electrode;

15        performing planarization treatment on a surface of the first insulating layer, thereby forming a second insulating layer having a planarized surface and exposing the source electrode and the drain electrode;

20        thinning the source electrode and the drain electrode so that a difference in height between the planarized surface of the second insulating layer and each of an upper surface of the source electrode and an upper surface of the drain electrode is formed;

      forming an oxide semiconductor layer over and in contact with the second insulating layer, the source electrode, and the drain electrode;

      forming a gate insulating layer over the oxide semiconductor layer; and

25        forming a gate electrode over the gate insulating layer.

18. The method for manufacturing a semiconductor device according to claim 17, wherein the planarization treatment is performed so that a root-mean-square roughness of the planarized surface is 1 nm or less.

30

19. The method for manufacturing a semiconductor device according to claim 17, wherein the difference in height between the planarized surface of the second

insulating layer and each of the upper surface of the source electrode and the upper surface of the drain electrode is 5 nm or more.

20. The method for manufacturing a semiconductor device according to claim 5 17, wherein the difference in height between the planarized surface of the second insulating layer and each of the upper surface of the source electrode and the upper surface of the drain electrode is 20 nm or less.

21. A method for manufacturing a semiconductor device, comprising the steps 10 of:

forming a first transistor comprising a channel formation region, a first gate insulating layer over the channel formation region, a first gate electrode over the first gate insulating layer, overlapping with the channel formation region, and a first source electrode and a first drain electrode electrically connected to the channel formation 15 region;

forming a first insulating layer over the first transistor;

forming a second source electrode and a second drain electrode over a surface of the first insulating layer;

20 forming a second insulating layer so as to cover the second source electrode and the second drain electrode;

performing planarization treatment on a surface of the second insulating layer, thereby forming a third insulating layer having a planarized surface;

25 forming openings reaching the second source electrode and the second drain electrode in the third insulating layer, so that a difference in height between the planarized surface of the third insulating layer and each of an upper surface of the second source electrode and an upper surface of the second drain electrode is formed;

forming an oxide semiconductor layer over and in contact with the third insulating layer, the second source electrode, and the second drain electrode;

forming a second gate insulating layer over the oxide semiconductor layer; and

30 forming a second gate electrode over the second gate insulating layer.

22. The method for manufacturing a semiconductor device according to claim

21, wherein the planarization treatment is performed so that a root-mean-square roughness of the planarized surface of the third insulating layer is 1 nm or less.

23. The method for manufacturing a semiconductor device according to claim  
5 21, wherein the difference in height between the planarized surface of the third insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 5 nm or more.

24. The method for manufacturing a semiconductor device according to claim  
10 21, wherein the difference in height between the planarized surface of the third insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 20 nm or less.

25. The method for manufacturing a semiconductor device according to claim  
15 21, wherein the openings are formed so that the upper surface of the second source electrode and the upper surface of the second drain electrode are partly covered with the second insulating layer.

26. A method for manufacturing a semiconductor device, comprising the steps  
20 of:

25 forming a first transistor comprising a channel formation region, a first gate insulating layer over the channel formation region, a first gate electrode over the first gate insulating layer, overlapping with the channel formation region, and a first source electrode and a first drain electrode electrically connected to the channel formation region;

forming a first insulating layer over the first transistor;

forming a second source electrode and a second drain electrode over a surface of the first insulating layer;

30 forming a second insulating layer so as to cover the second source electrode and the second drain electrode;

performing planarization treatment on a surface of the second insulating layer, thereby forming a third insulating layer having a planarized surface and exposing the

second source electrode and the second drain electrode;

thinning the second source electrode and the second drain electrode so that a difference in height between the planarized surface of the third insulating layer and each of an upper surface of the second source electrode and an upper surface of the second drain electrode is formed;

forming an oxide semiconductor layer over and in contact with the third insulating layer, the second source electrode, and the second drain electrode;

forming a second gate insulating layer over the oxide semiconductor layer; and forming a second gate electrode over the second gate insulating layer.

10

27. The method for manufacturing a semiconductor device according to claim 26, wherein the planarization treatment is performed so that a root-mean-square roughness of the planarized surface of the third insulating layer is 1 nm or less.

15

28. The method for manufacturing a semiconductor device according to claim 26, wherein the difference in height between the planarized surface of the third insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 5 nm or more.

20

29. The method for manufacturing a semiconductor device according to claim 26, wherein the difference in height between the planarized surface of the third insulating layer and each of the upper surface of the second source electrode and the upper surface of the second drain electrode is 20 nm or less.

1/20

FIG. 1A

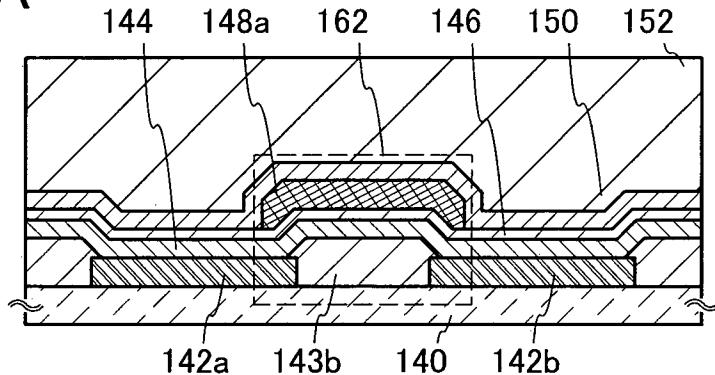


FIG. 1B

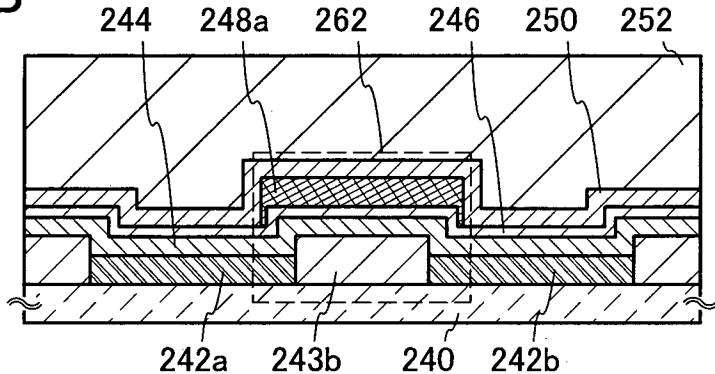
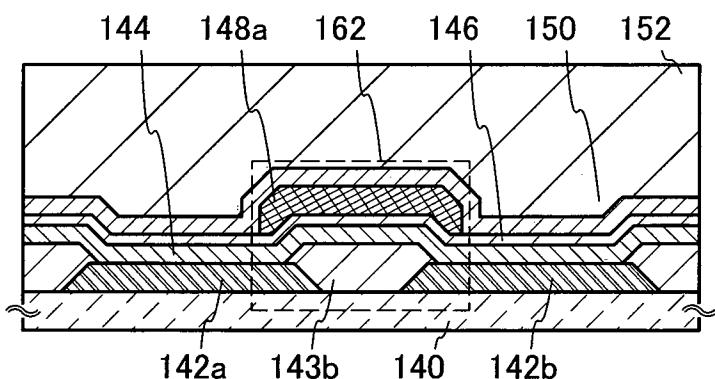
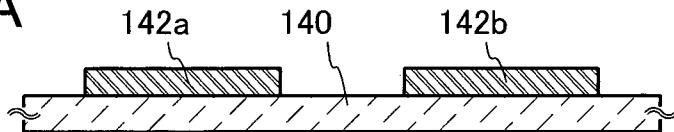
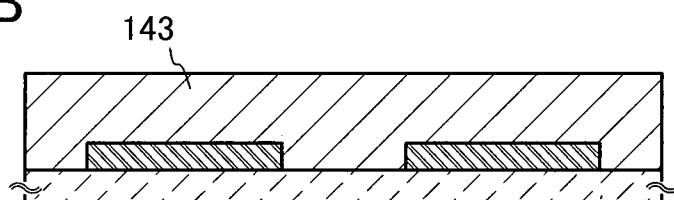
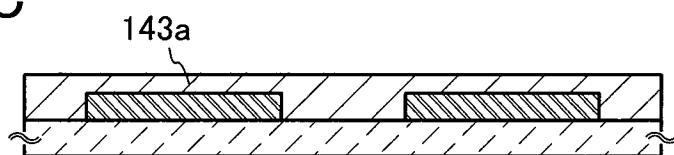
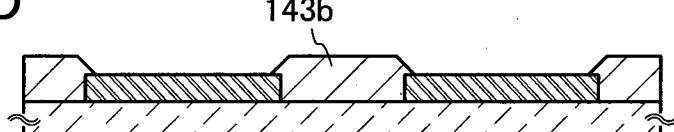
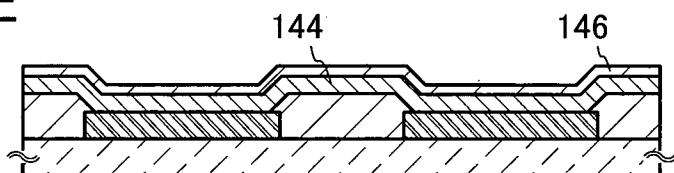
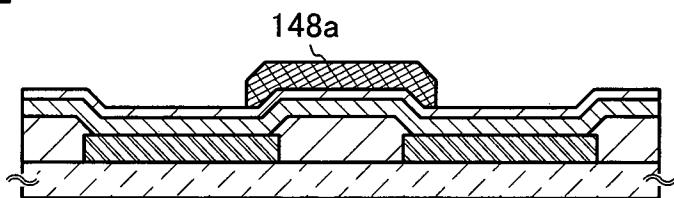
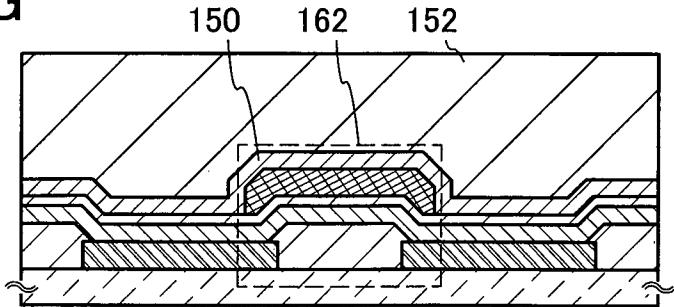


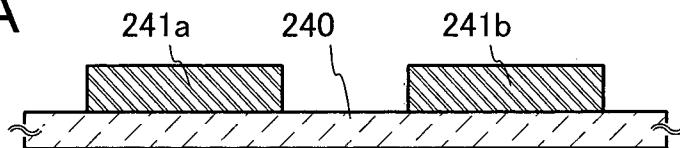
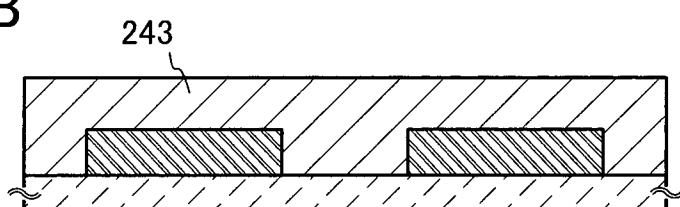
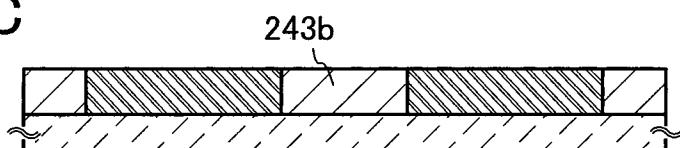
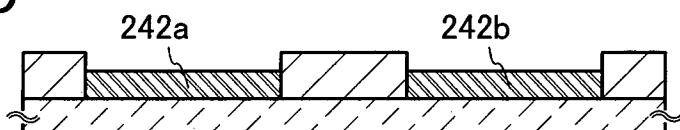
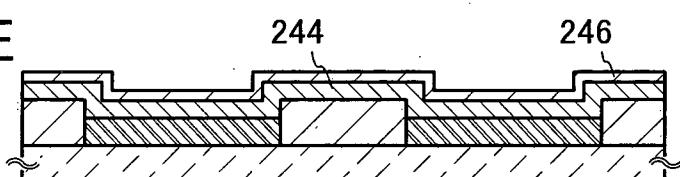
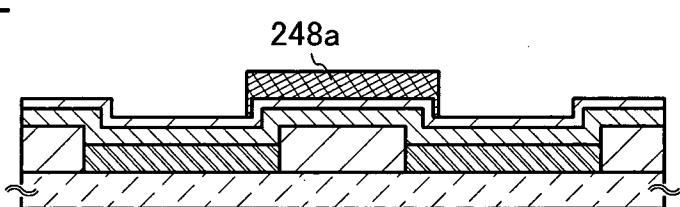
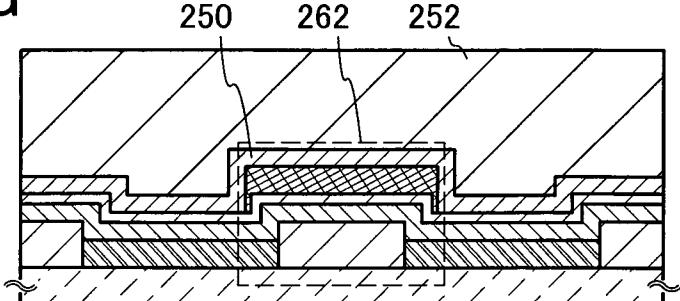
FIG. 1C



2/20

**FIG. 2A****FIG. 2B****FIG. 2C****FIG. 2D****FIG. 2E****FIG. 2F****FIG. 2G**

3/20

**FIG. 3A****FIG. 3B****FIG. 3C****FIG. 3D****FIG. 3E****FIG. 3F****FIG. 3G**

4/20

FIG. 4A

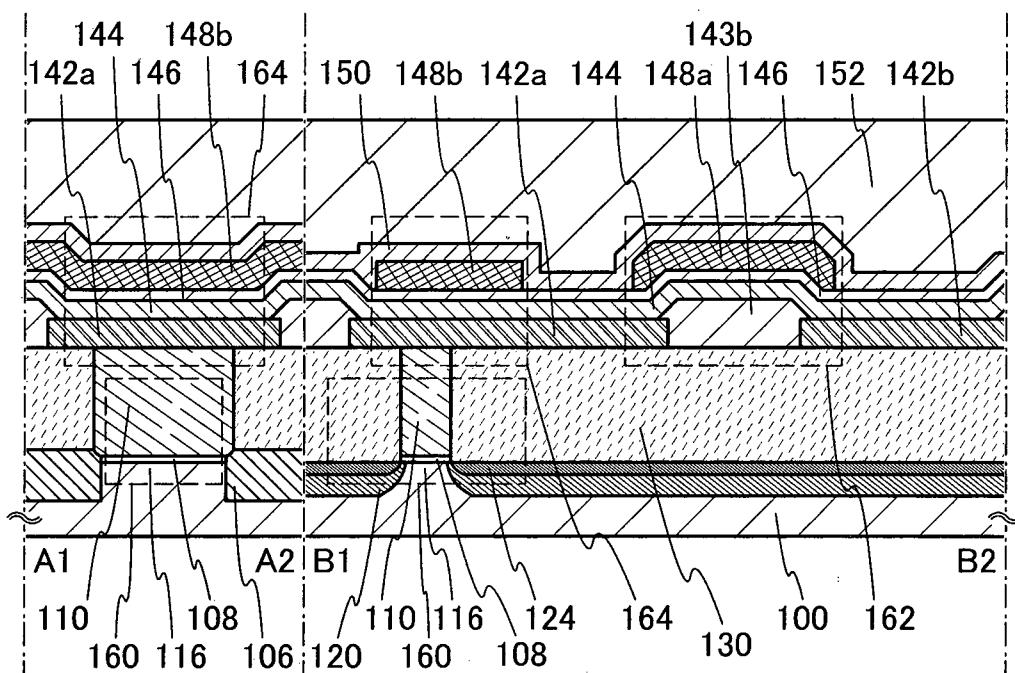


FIG. 4B

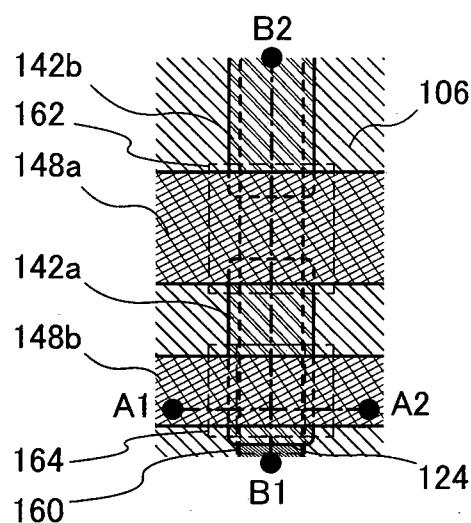
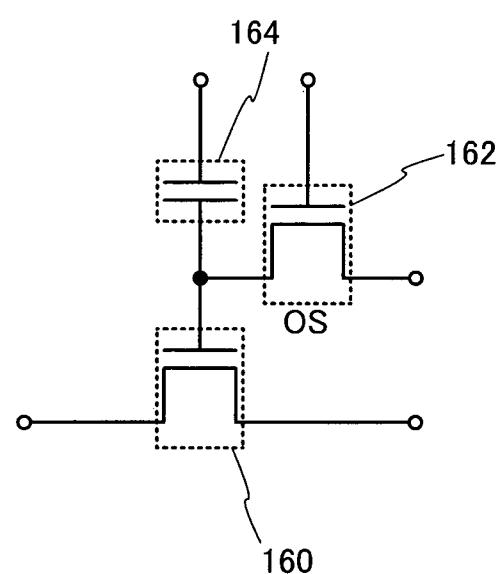


FIG. 4C



5/20

FIG. 5A

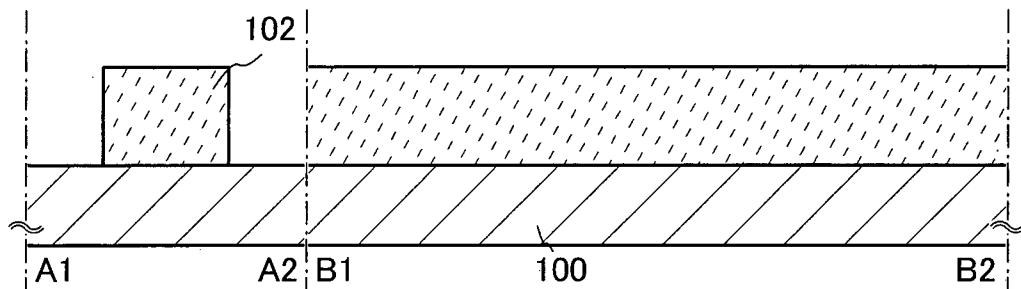


FIG. 5B

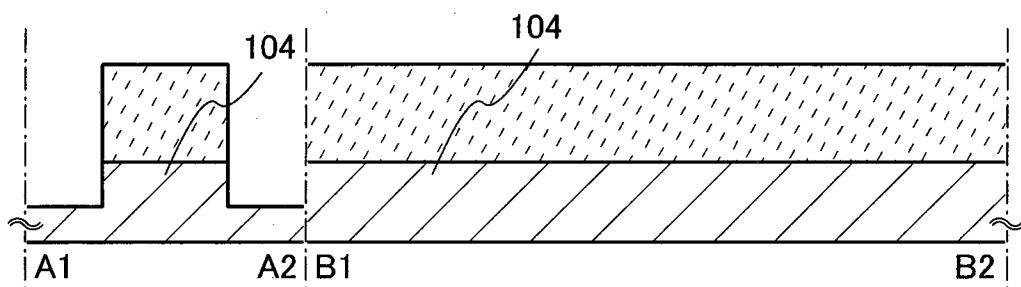


FIG. 5C

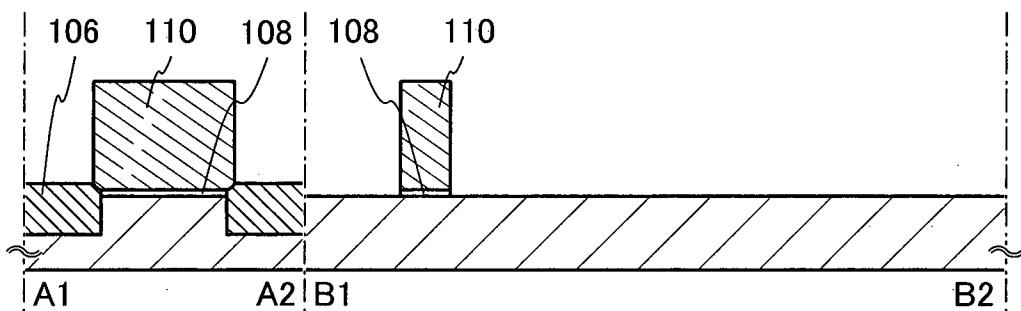
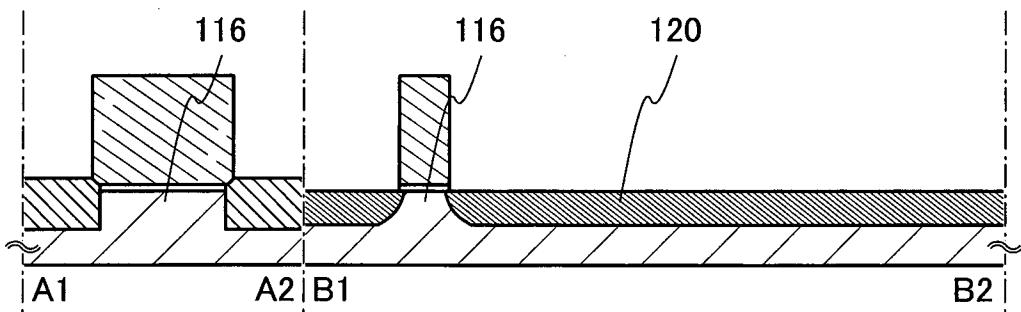


FIG. 5D



6/20

FIG. 6A

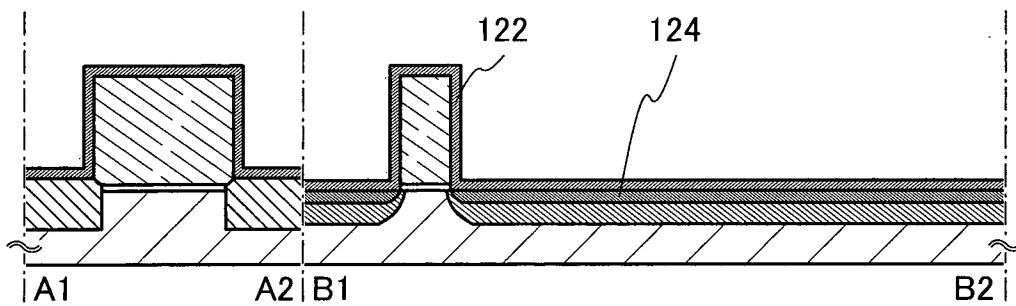


FIG. 6B

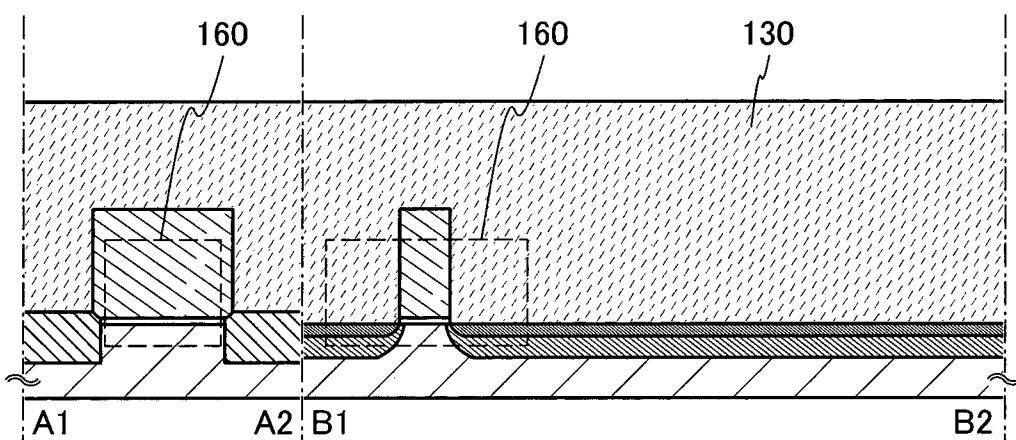
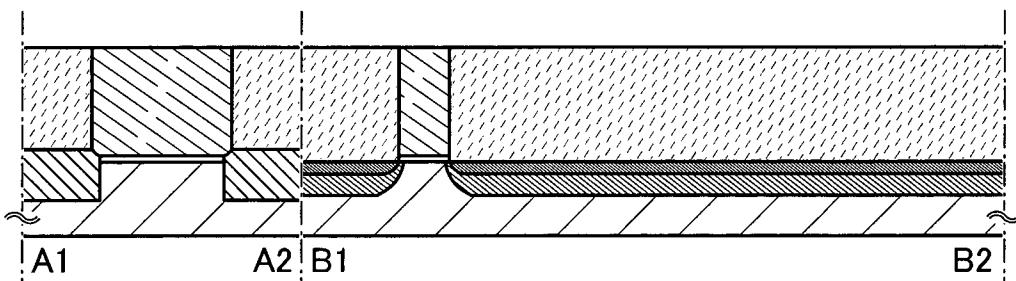


FIG. 6C



7/20

FIG. 7A-1

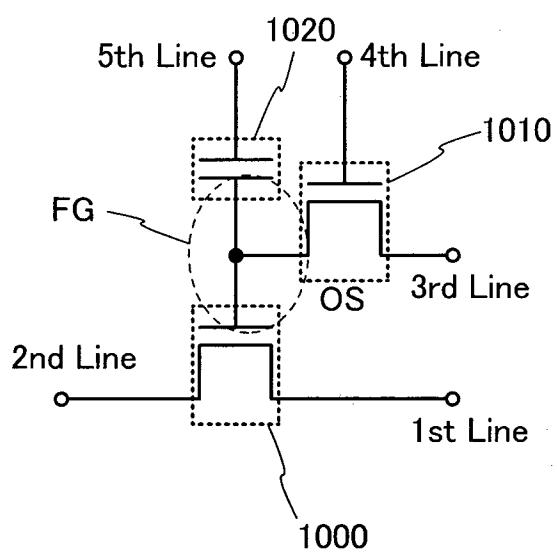


FIG. 7B

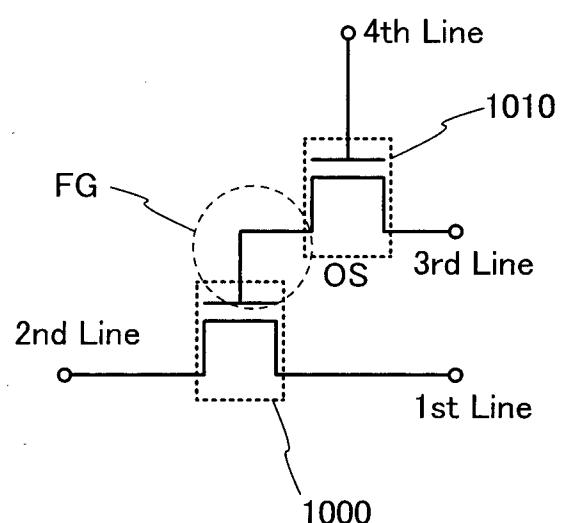
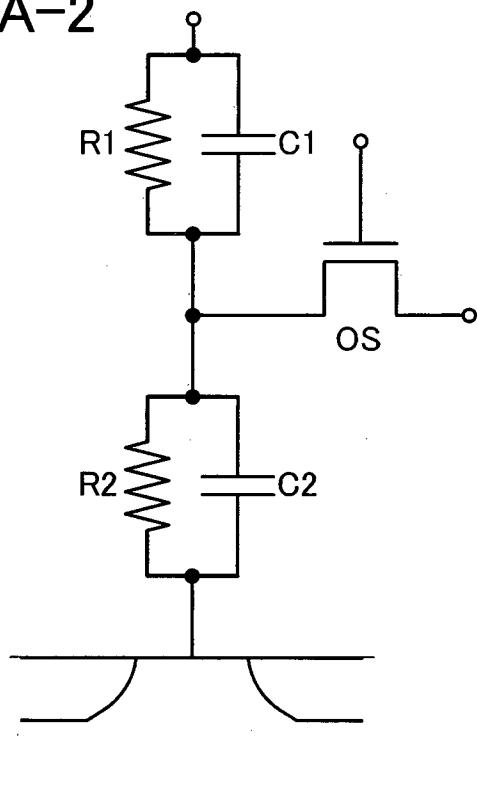


FIG. 7A-2



8/20

FIG. 8A

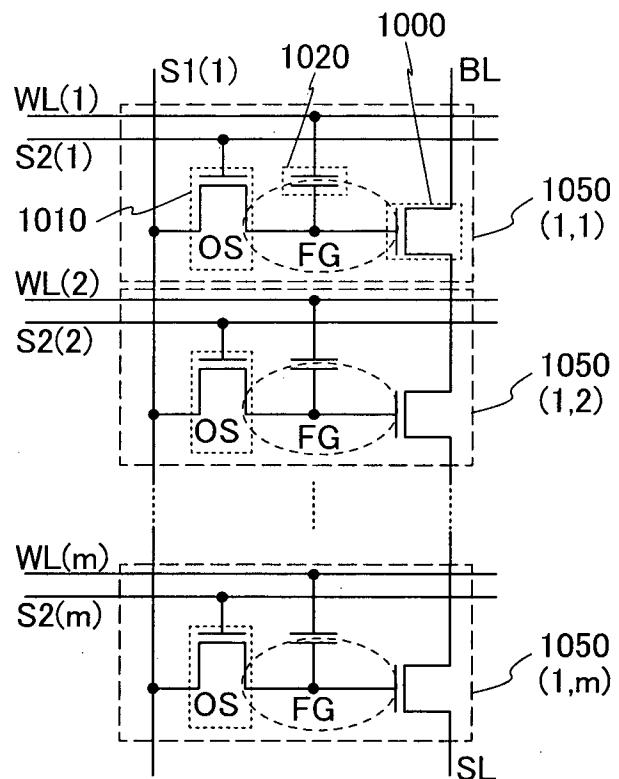
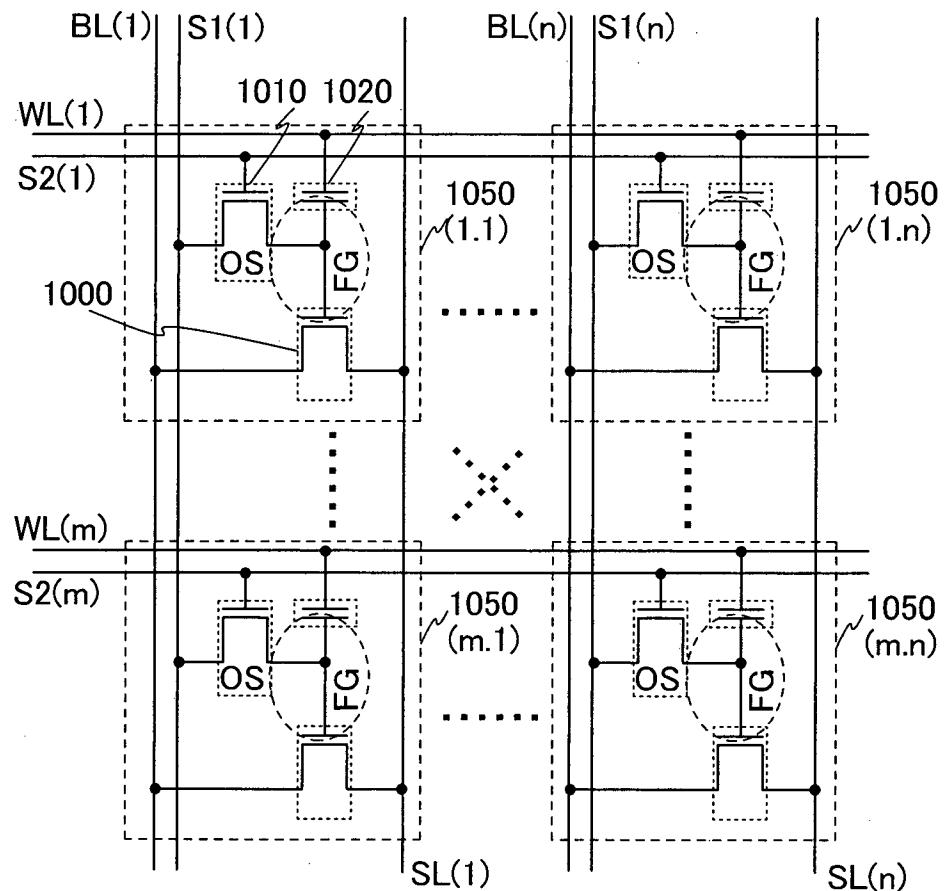


FIG. 8B



9/20

FIG. 9A

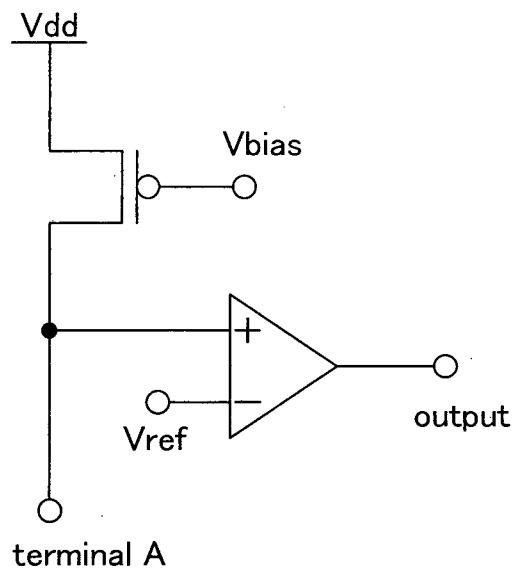


FIG. 9B

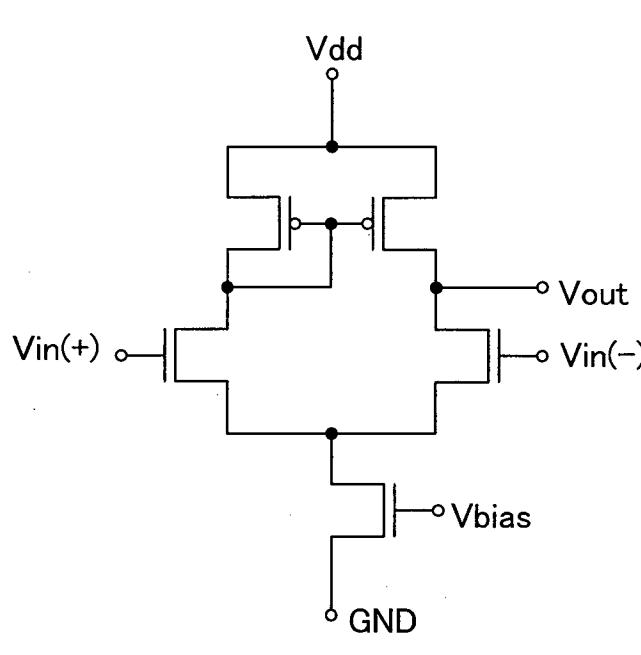
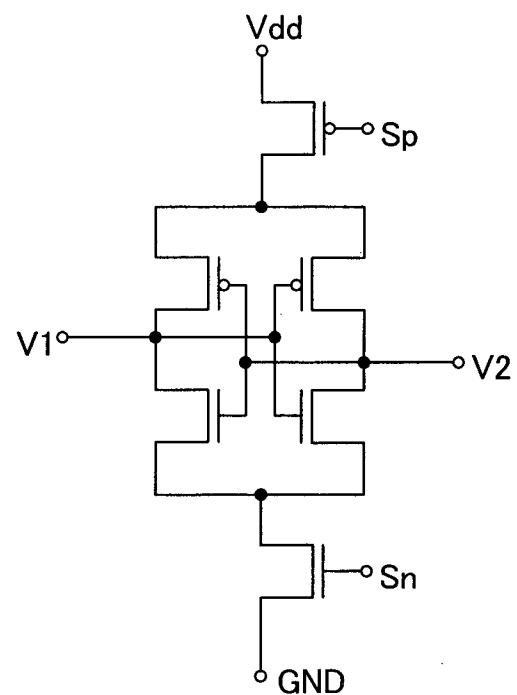


FIG. 9C



10/20

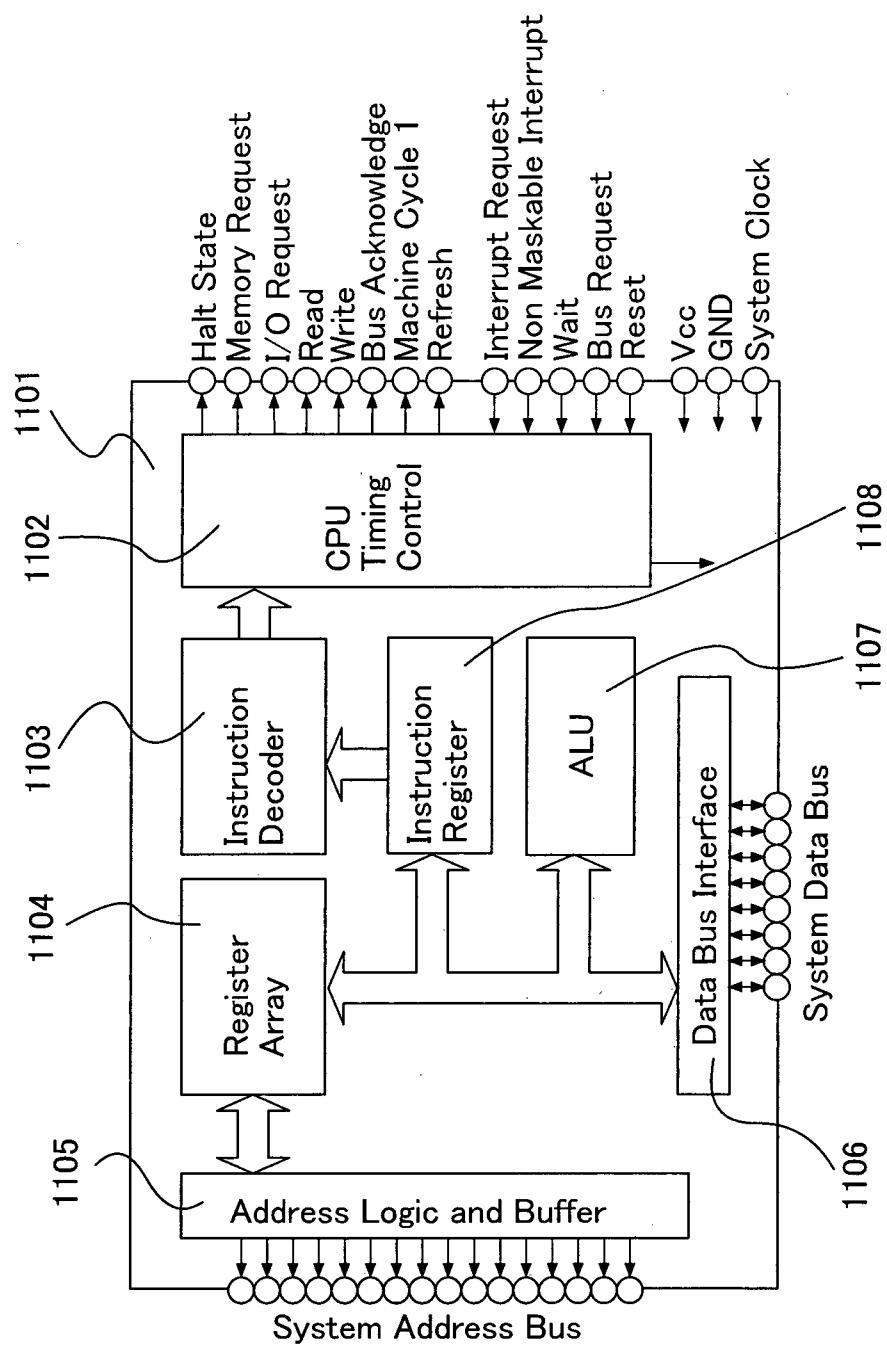


FIG. 10

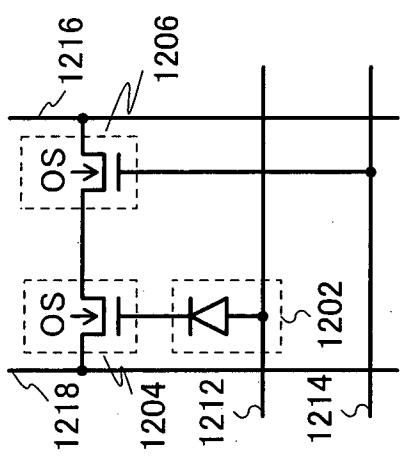


FIG. 11A

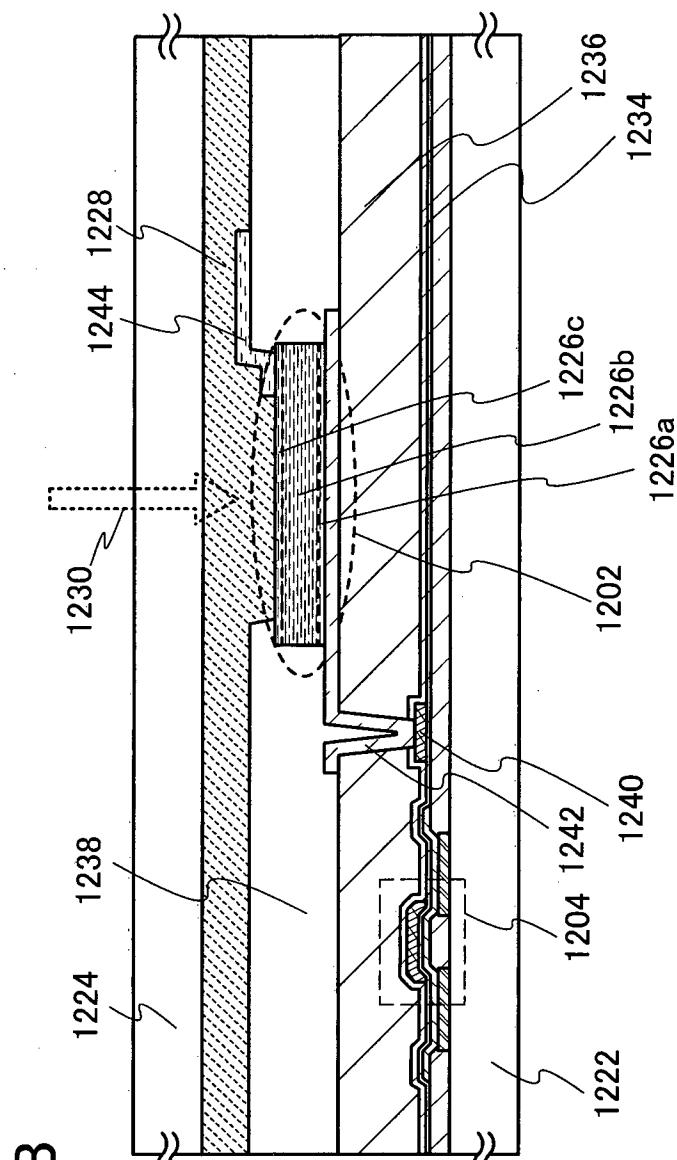


FIG. 11B

12/20

FIG. 12A

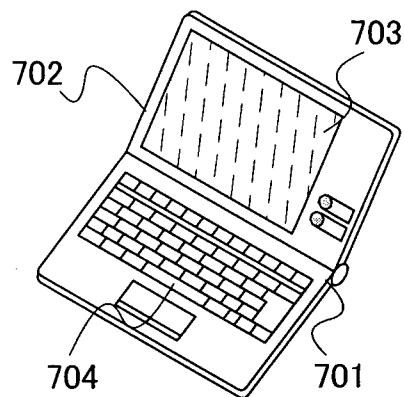


FIG. 12D

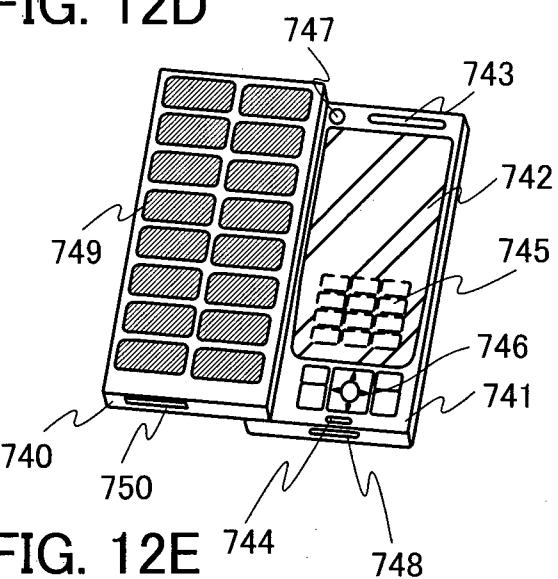


FIG. 12B

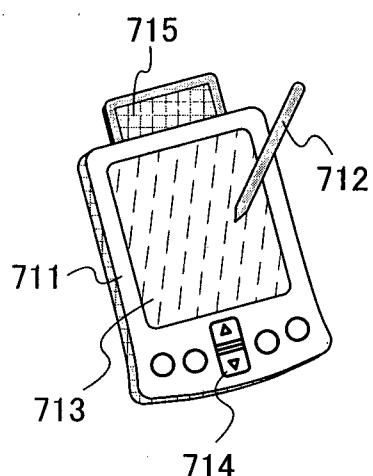


FIG. 12E

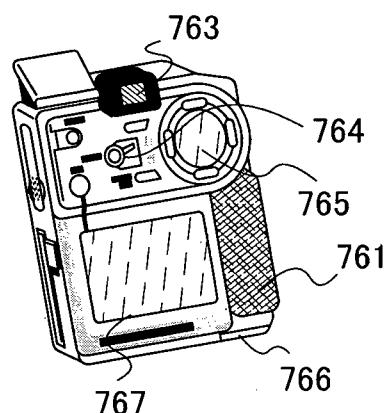


FIG. 12C

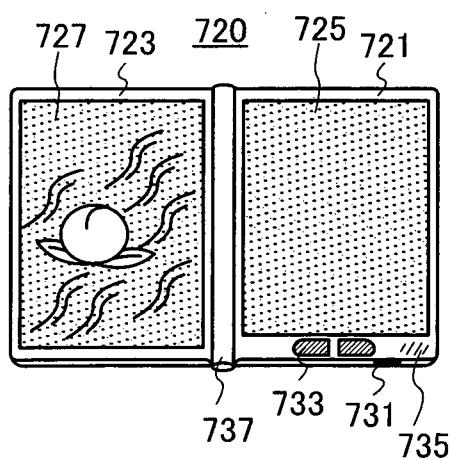
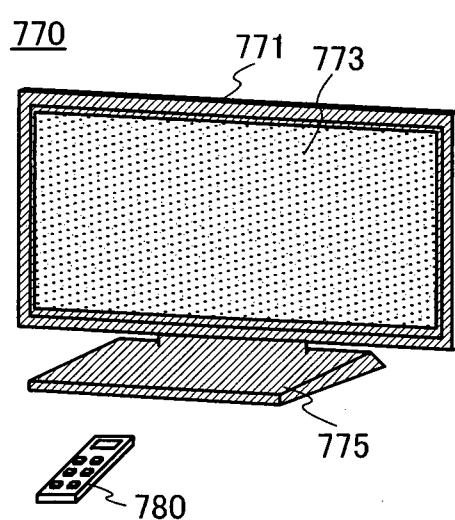


FIG. 12F



13/20

FIG. 13A

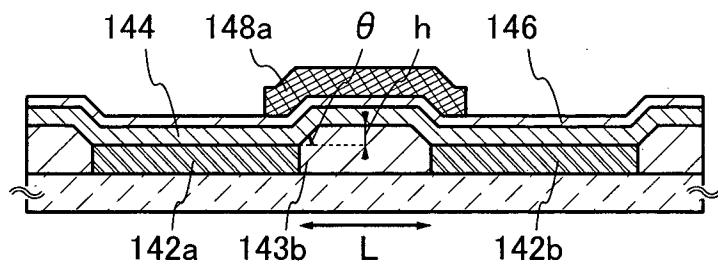
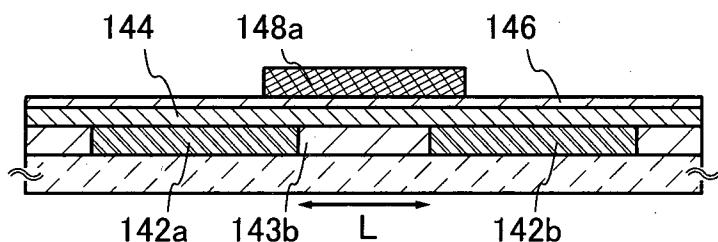
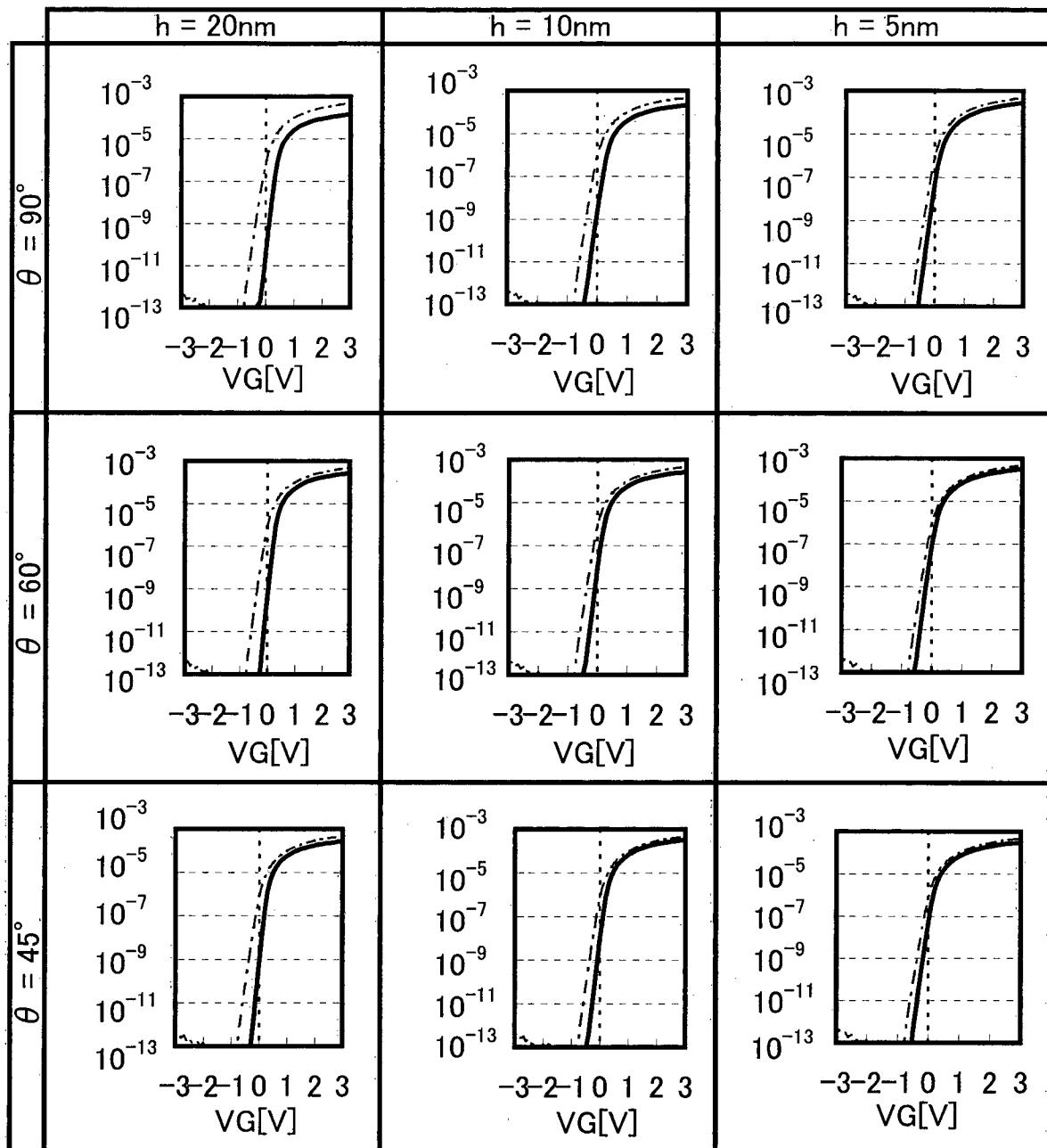


FIG. 13B



14/20

FIG. 14



vertical axis:ID[A]

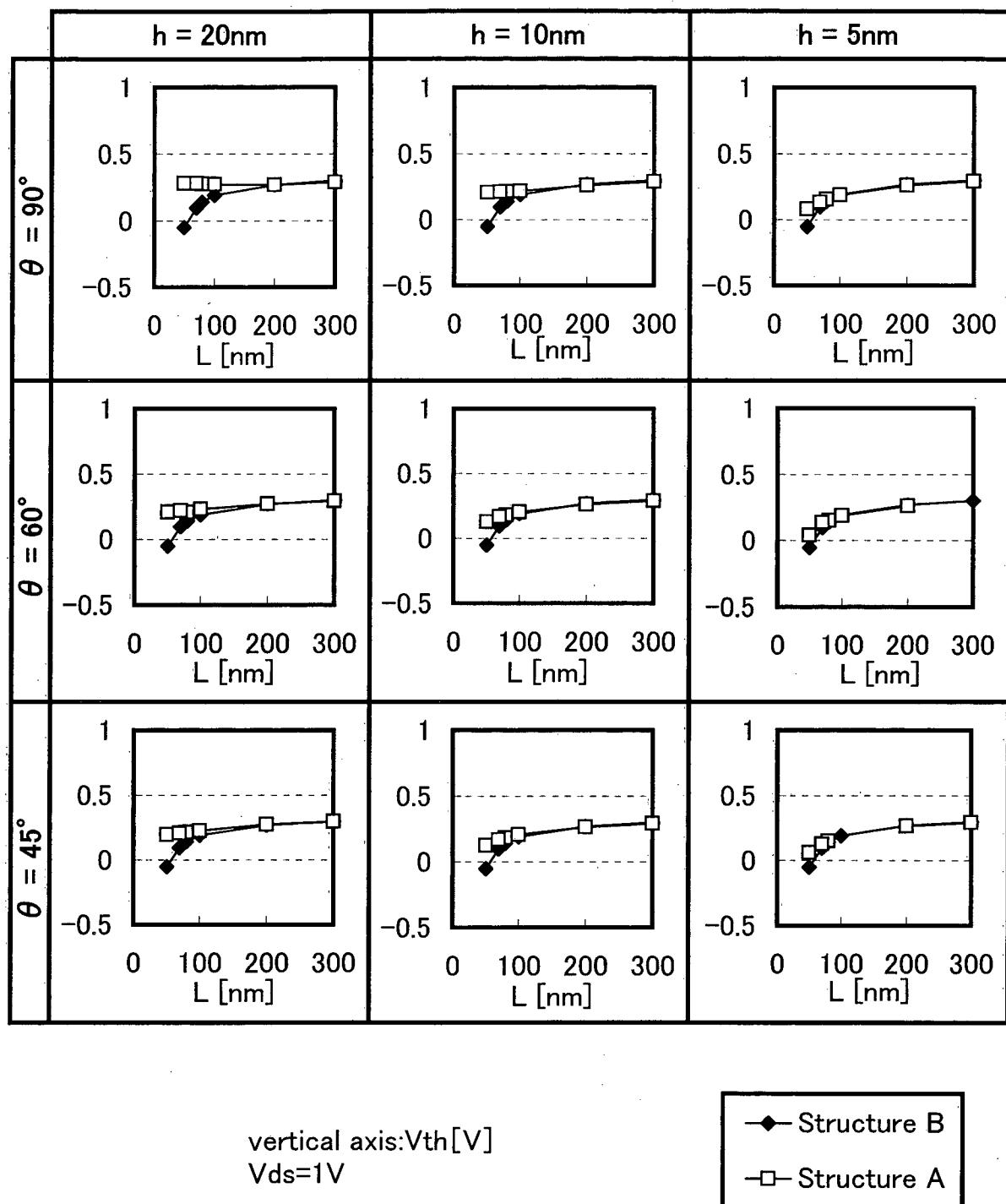
 $W=1\text{ }\mu\text{m}$  $V_{ds}=1\text{V}$  $L=50\text{nm}$ 

--- Structure B

— Structure A

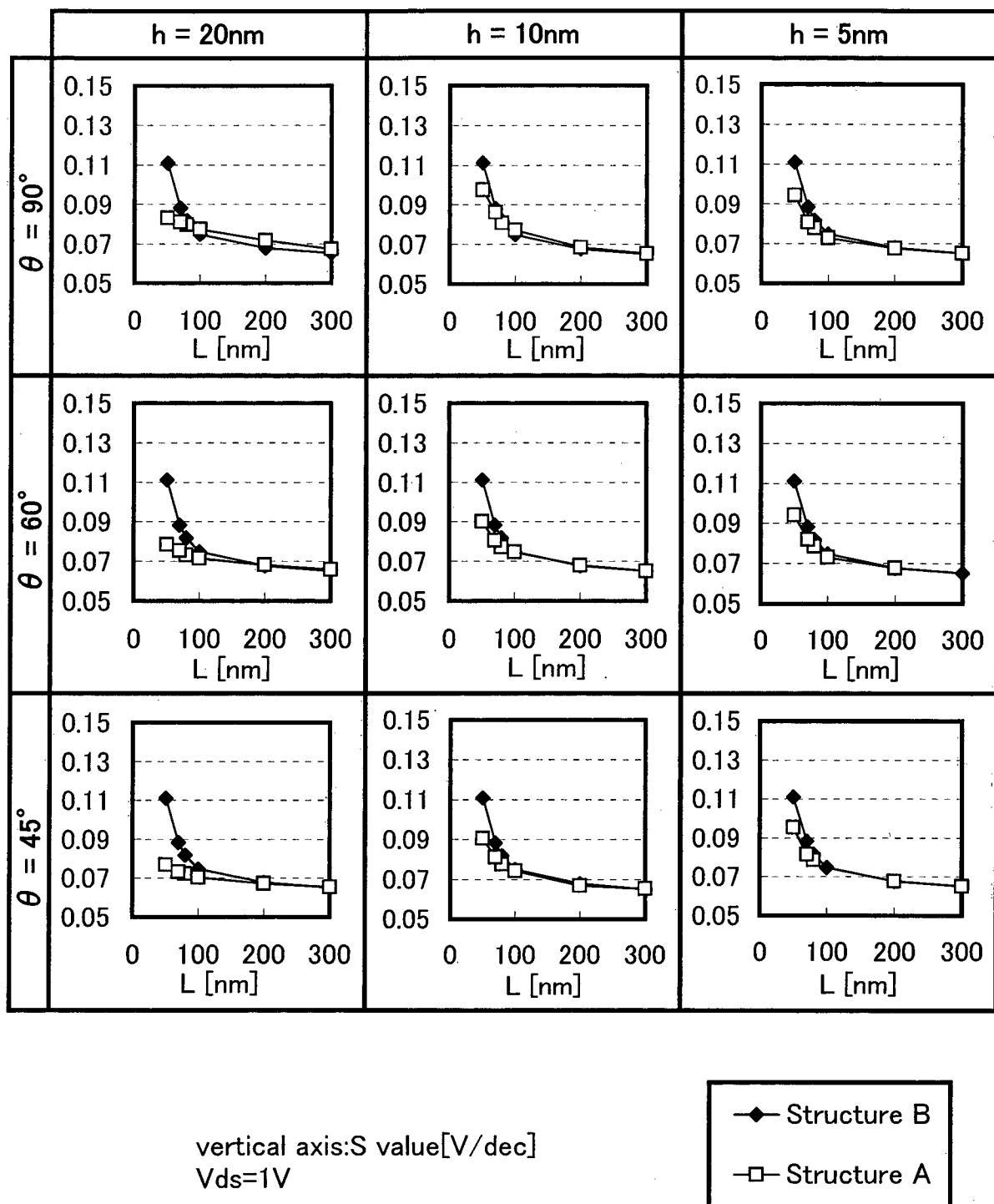
15/20

FIG. 15



16/20

FIG. 16



17/20

FIG. 17A

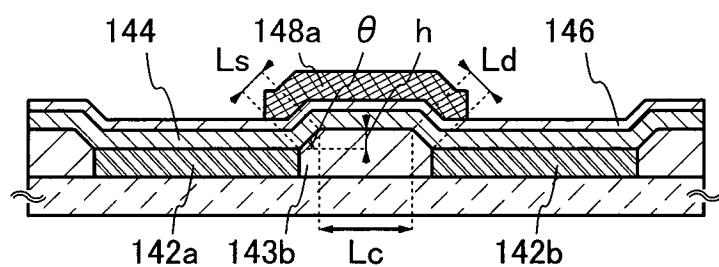
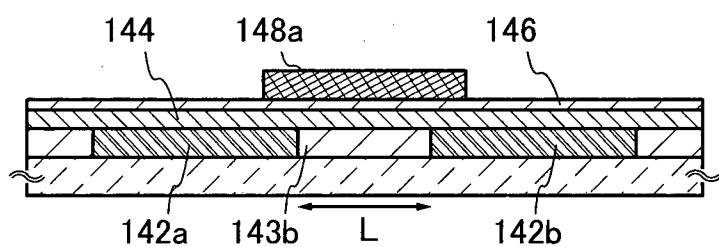
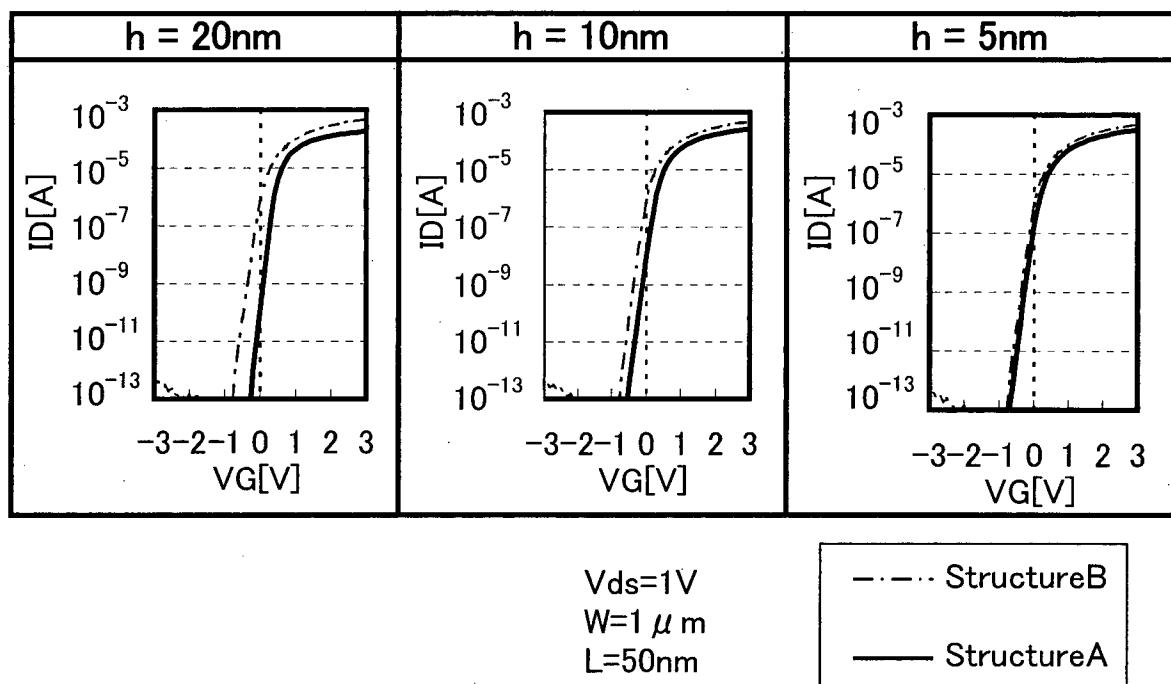


FIG. 17B



18/20

FIG. 18



## EXPLANATION OF REFERENCE

100: substrate, 102: protective layer, 104: semiconductor region, 106: element isolation insulating layer, 108: gate insulating layer, 110: gate electrode, 116: channel formation region, 120: impurity region, 122: metal layer, 124: metal compound region, 130: insulating layer, 140: substrate, 142a: source electrode, 142b: drain electrode, 143: insulating layer, 143a: insulating layer, 143b: insulating layer, 144: oxide semiconductor layer, 146: gate insulating layer, 148a: gate electrode, 148b: electrode, 150: insulating layer, 152: insulating layer, 160: transistor, 162: transistor, 164: capacitor, 240: substrate, 241a: conductive layer, 241b: conductive layer, 242a: source electrode, 242b: drain electrode, 243: insulating layer, 243b: insulating layer, 244: oxide semiconductor layer, 246: gate insulating layer, 248a: gate electrode, 250: insulating layer, 252: insulating layer, 262: transistor, 701: housing, 702: housing, 703: display portion, 704: keyboard, 711: main body, 712: stylus, 713: display portion, 714: operation button, 715: external interface, 720: electronic book, 721: housing, 723: housing, 725: display portion, 727: display portion, 731: power switch, 733: operation key, 735: speaker, 737: hinge, 740: housing, 741: housing, 742: display panel, 743: speaker, 744: microphone, 745: touch panel, 746: pointing device, 747: camera lens, 748: external connection terminal, 749: solar cell, 750: external memory slot, 761: main body, 763: eyepiece, 764: operation switch, 765: display portion, 766: battery, 767: display portion, 770: television set, 771: housing, 773: display portion, 775: stand, 780: remote controller, 1000: transistor, 1010: transistor, 1020: capacitor, 1050: memory cell, 1101: CPU,

20/20

1102: timing control circuit, 1103: instruction decoder, 1104: register array, 1105: address logic and buffer circuit, 1106: data bus interface, 1107: ALU, 1108: instruction register, 1202: photodiode, 1204: transistor, 1206: transistor, 1212: photodiode reset signal line, 1214: gate signal line, 1216: photosensor output signal line, 1218: photosensor reference signal line, 1222: substrate, 1224: substrate, 1226a: first semiconductor layer, 1226b: second semiconductor layer, 1226c: third semiconductor layer, 1228: adhesive layer, 1230: incident light, 1234: insulating layer, 1236: interlayer insulating layer, 1238: interlayer insulating layer, 1240: gate electrode layer, 1242: electrode layer, and 1244: electrode layer.

**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/JP2011/053589

**A. CLASSIFICATION OF SUBJECT MATTER**

Int.Cl. H01L21/336 (2006.01) i, H01L21/8234 (2006.01) i, H01L27/08 (2006.01) i, H01L27/088 (2006.01) i, H01L29/786 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H01L21/336, H01L21/8234, H01L27/08, H01L27/088, H01L29/786

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996  
Published unexamined utility model applications of Japan 1971-2011  
Registered utility model specifications of Japan 1996-2011  
Published registered utility model applications of Japan 1994-2011

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 61-204976 A (Seiko Instruments, Inc.)	1, 4
Y	1986.09.11, Page 3, upper left column, line 1 to Page 4, lower right column, line 1 (No Family)	2, 5, 6, 8-13, 15, 16, 21, 22, 24, 25
A		3, 7, 14, 17-20 , 23, 26-29
Y	US 5712496 A (Seiko Instruments, Inc.) 1998.01.27, column 6, line 13 to column 7, line 63 & JP 08-023101 A	2, 6, 12, 13, 15, 16, 21, 22, 24, 25
Y	JP 2010-021170 A (HITACHI LTD.) 2010.01.28, Par. Nos. [0082] to [0089] (No Family)	5, 6, 8-11, 21, 22, 24, 25

 Further documents are listed in the continuation of Box C.

 See patent family annex.

\* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent 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 when the document is taken alone

“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

Date of the actual completion of the international search

12.05.2011

Date of mailing of the international search report

24.05.2011

Name and mailing address of the ISA/JP

**Japan Patent Office**

3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan

Authorized officer

Takashi WATAHIKI

4M 2934

Telephone No. +81-3-3581-1101 Ext. 3462