US 20070105317A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2007/0105317 A1**

Nakajima (43) **Pub. Date: May 10, 2007**

(54) **METHOD OF MANUFACTURING SEMICONDUCTOR DEVICE**

(76) Inventor: **Kazuaki Nakajima**, Kamakura-shi (JP)

Correspondence Address:
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER LLP
901 NEW YORK AVENUE, NW
WASHINGTON, DC 20001-4413 (US)

(21) Appl. No.: **11/378,254**

(22) Filed: **Mar. 20, 2006**

(30) **Foreign Application Priority Data**

Nov. 9, 2005 (JP) 2005-325024

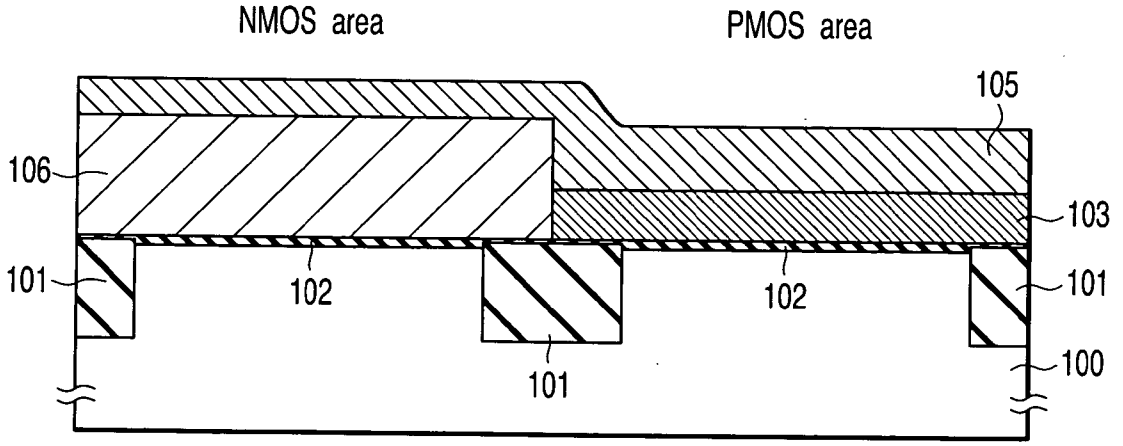
Publication Classification

(51) **Int. Cl.**
H01L 21/336 (2006.01)

(52) **U.S. Cl.** **438/264**

(57) **ABSTRACT**

A method of manufacturing a semiconductor device according to an aspect of the present invention comprises a step of forming a gate insulating layer on a semiconductor substrate, a step of forming a first metal layer on the gate insulating layer, a step of forming a second metal layer including elements for use in work function modulation on the first metal layer, a step of forming a cap layer made of a material having a melting point higher than that of the second metal layer on the second metal layer, and a step of precipitating the elements at an interface between the gate insulating layer and the first metal layer by thermal treatment.



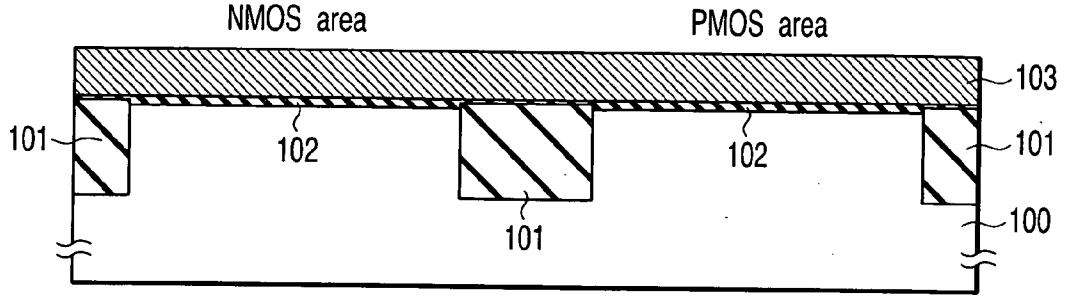


FIG. 1

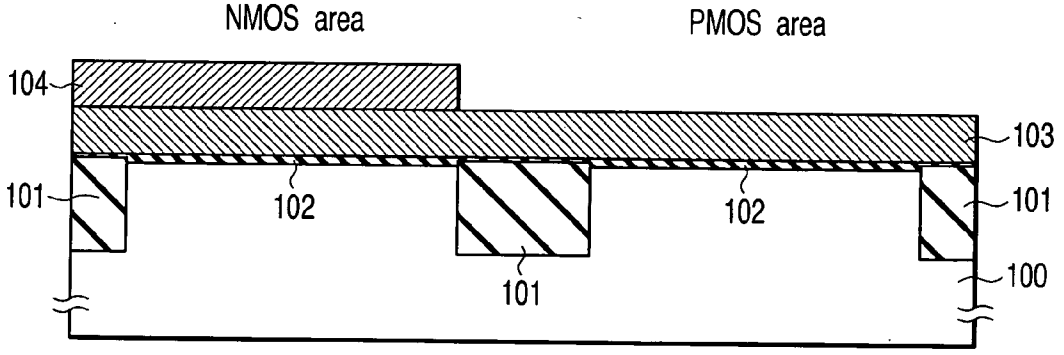


FIG. 2

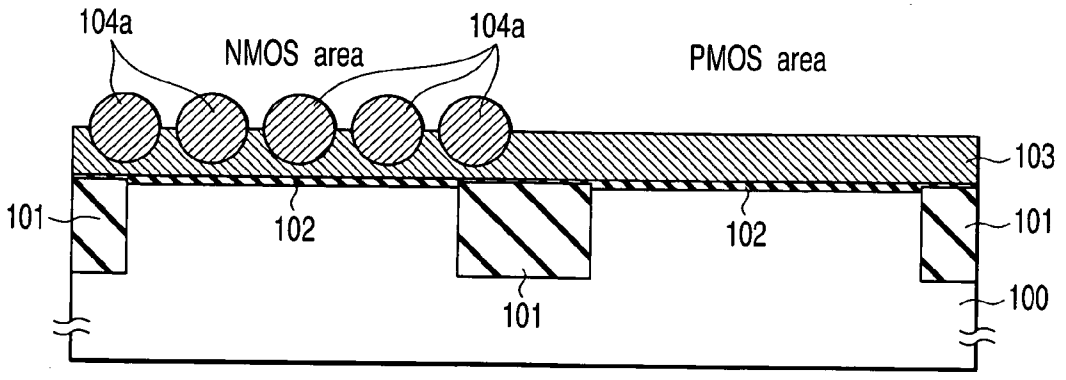


FIG. 3

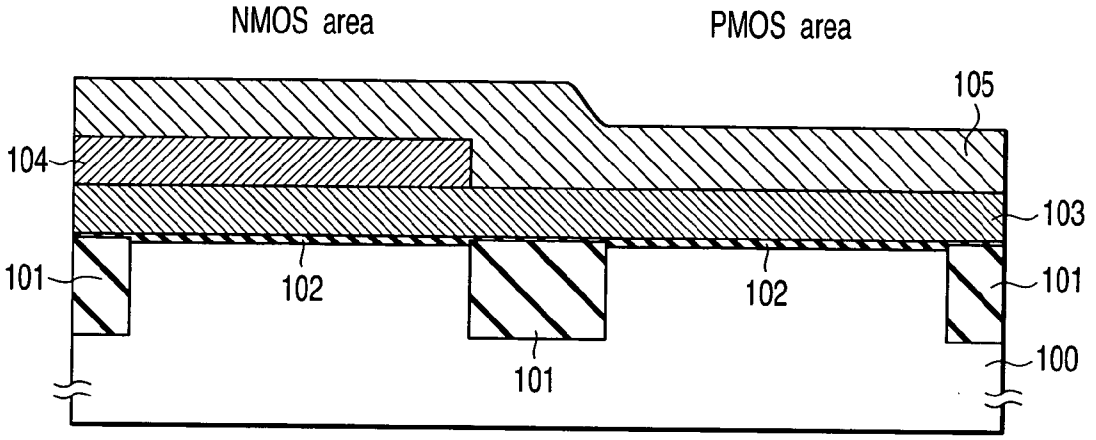

NMOS area
PMOS area
105
104
103
101
102
102
101
101
100

FIG. 4

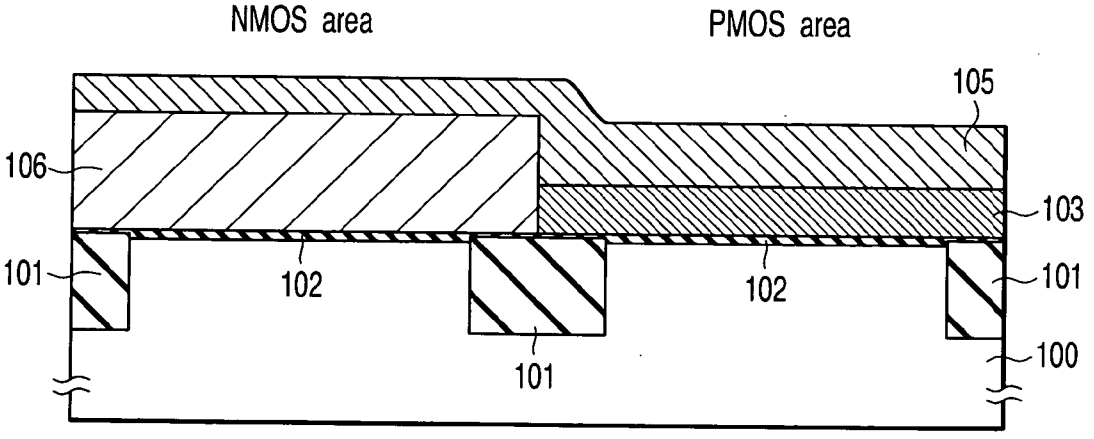

NMOS area
PMOS area
105
106
103
101
102
102
101
101
100

FIG. 5

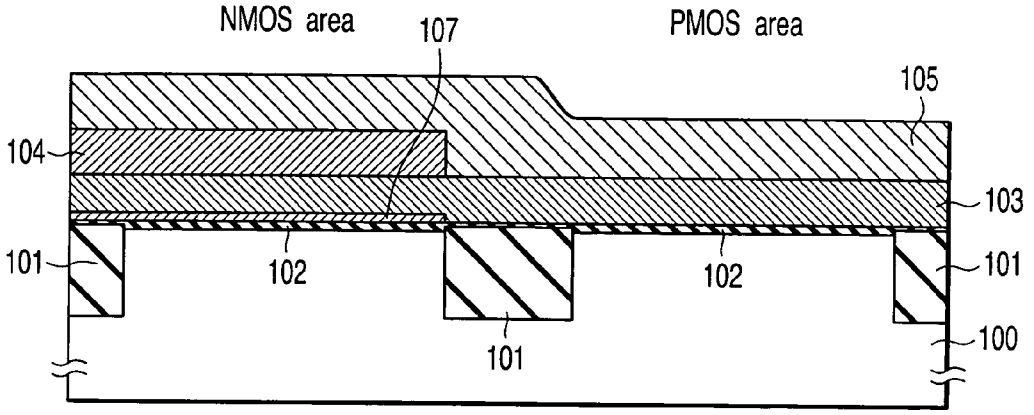


FIG. 6

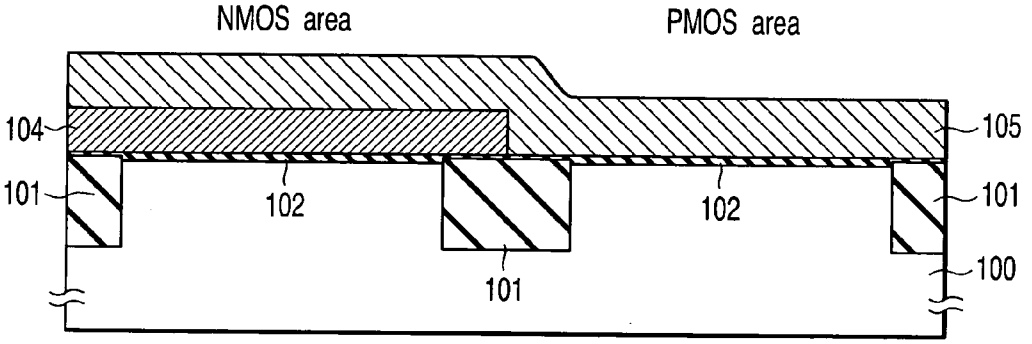


FIG. 7

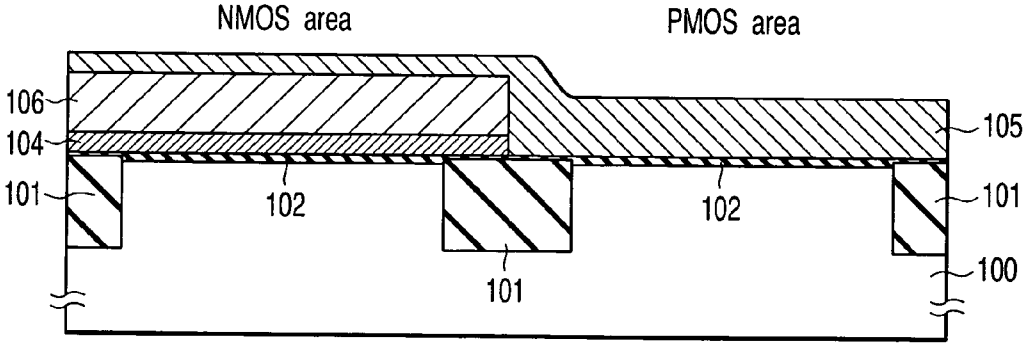


FIG. 8

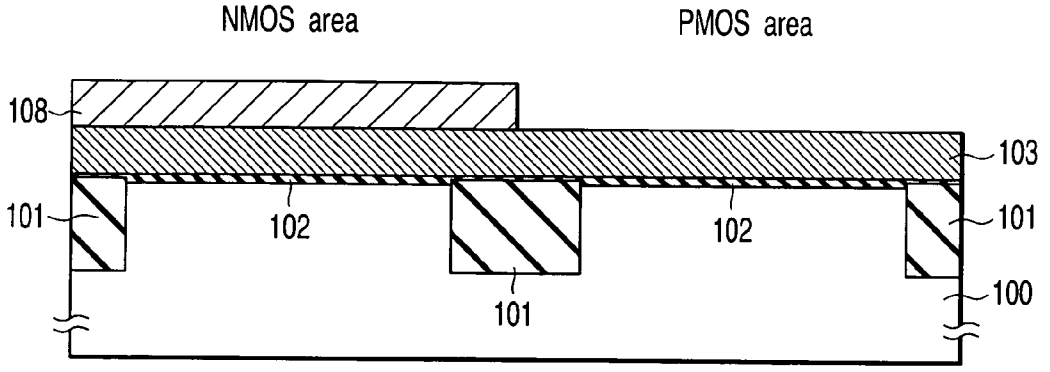


FIG. 9

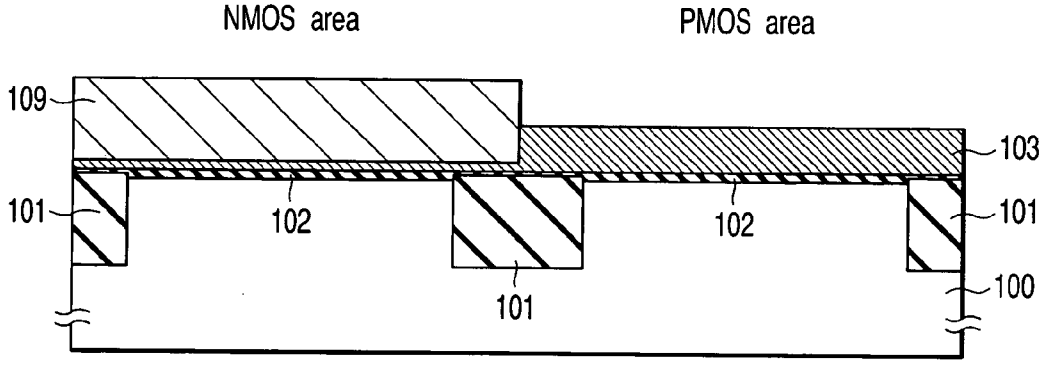


FIG. 10

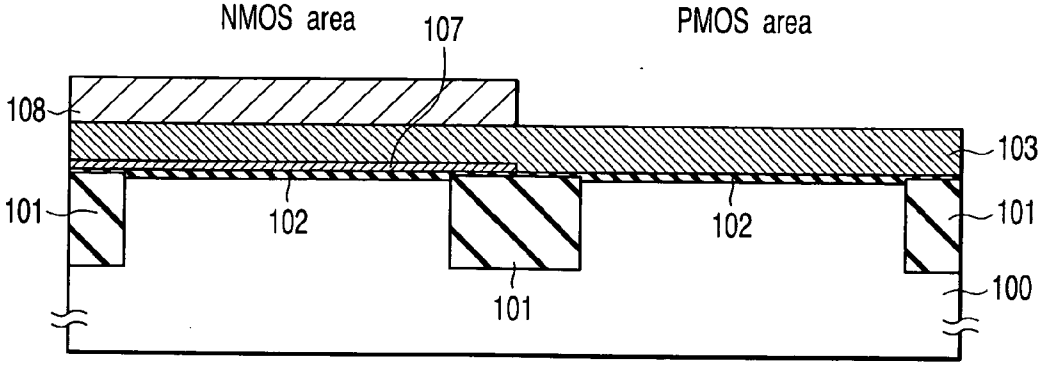


FIG. 11

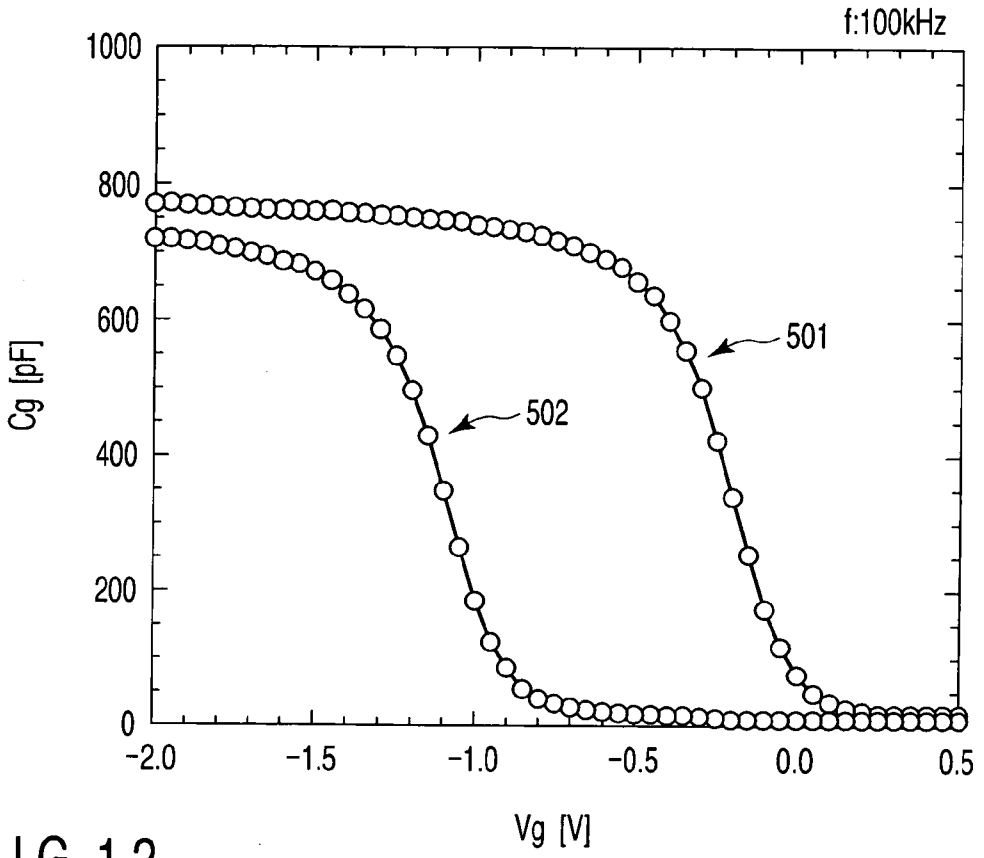

f:100kHz
1000
800
600
400
200
0
Cg [pF]
501
502
-2.0
-1.5
-1.0
-0.5
0.0
0.5
Vg [V]

FIG. 12

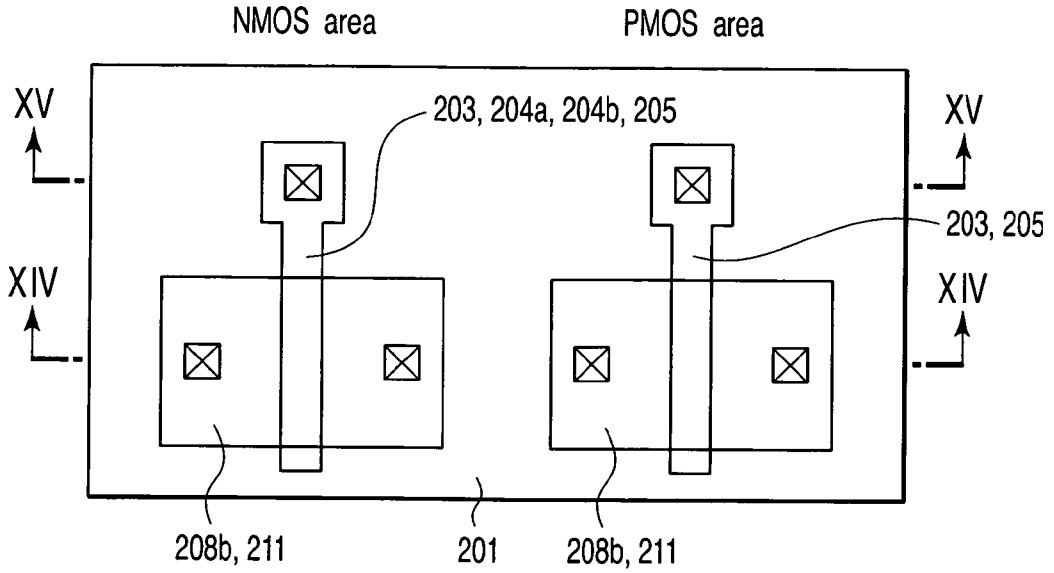

NMOS area
PMOS area
XV
XIV
203, 204a, 204b, 205
203, 205
208b, 211
201
208b, 211

FIG. 13

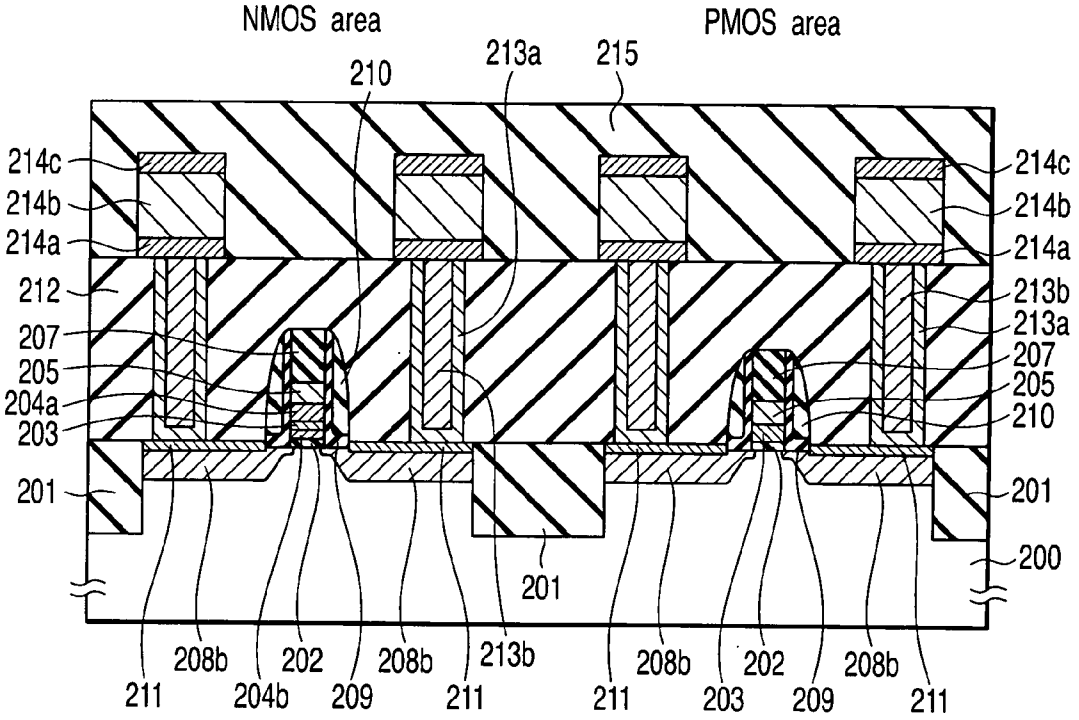

NMOS area
PMOS area
210
213a
215
214c
214b
214a
212
207
205
204a
203
201
213b
200
208b
202
204b
209
211

FIG. 14

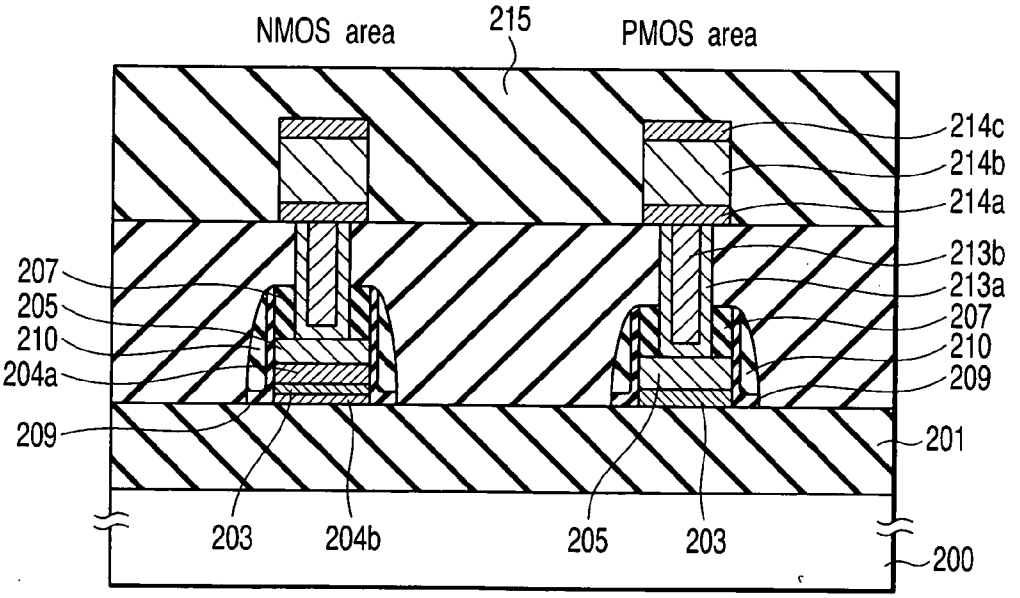

NMOS area
215
PMOS area
214c
214b
214a
213b
213a
207
205
210
204a
209
201
200
203
204b

FIG. 15

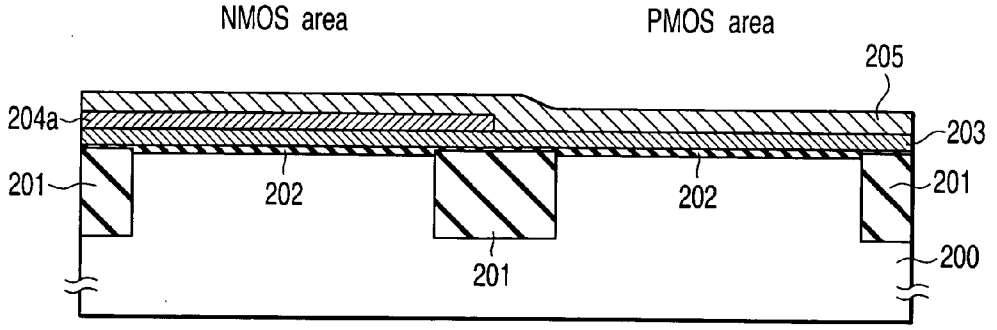


FIG. 16

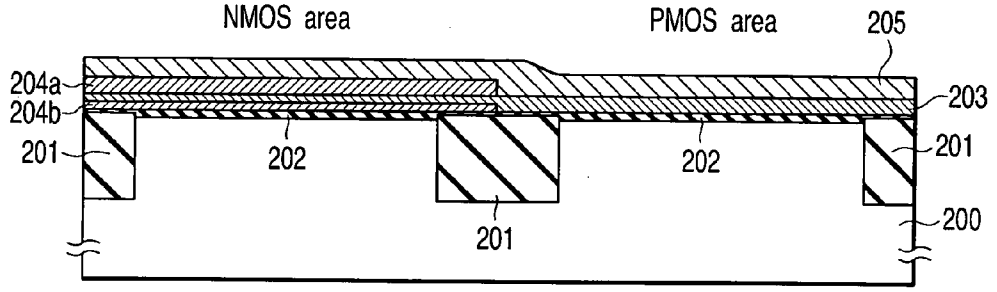


FIG. 17

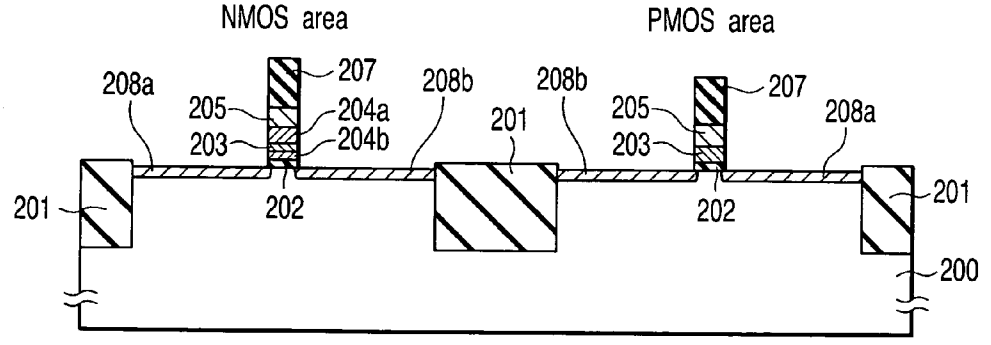


FIG. 18

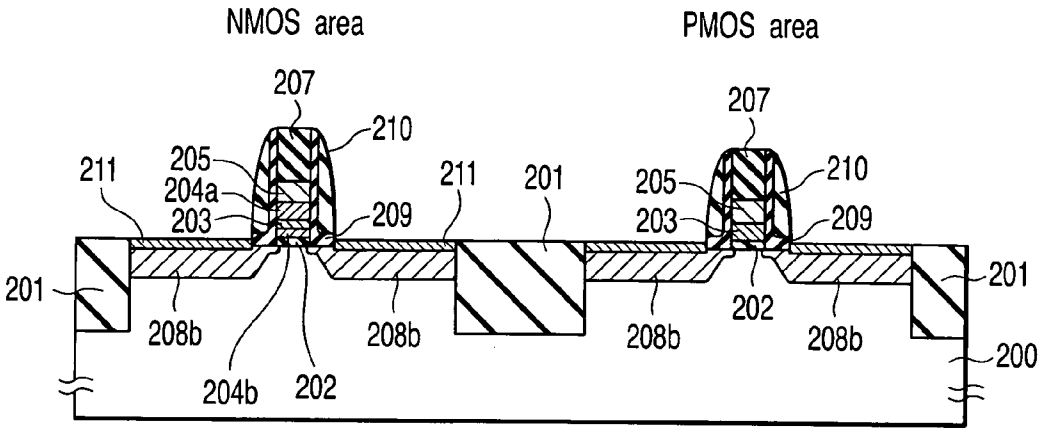

NMOS area
PMOS area
207
210
205
204a
203
209
211
201
208b
204b
202
200

FIG. 19

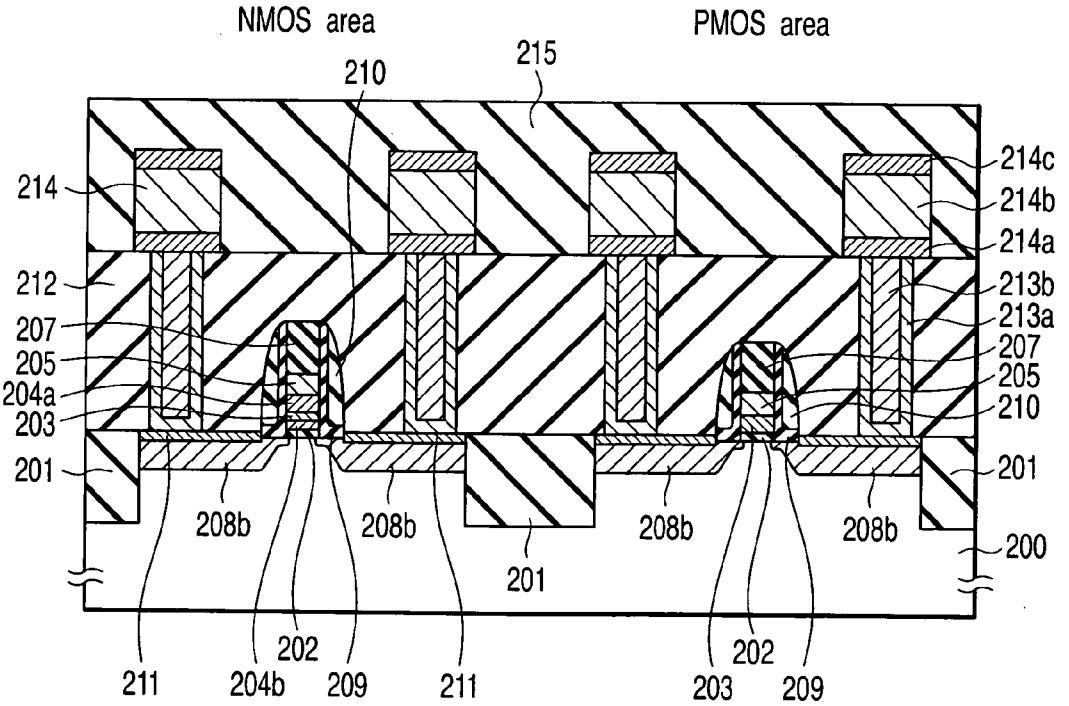

NMOS area
PMOS area
210
215
214
214c
214b
214a
212
213b
213a
207
205
204a
203
201
208b
202
204b
209
211
200

FIG. 20

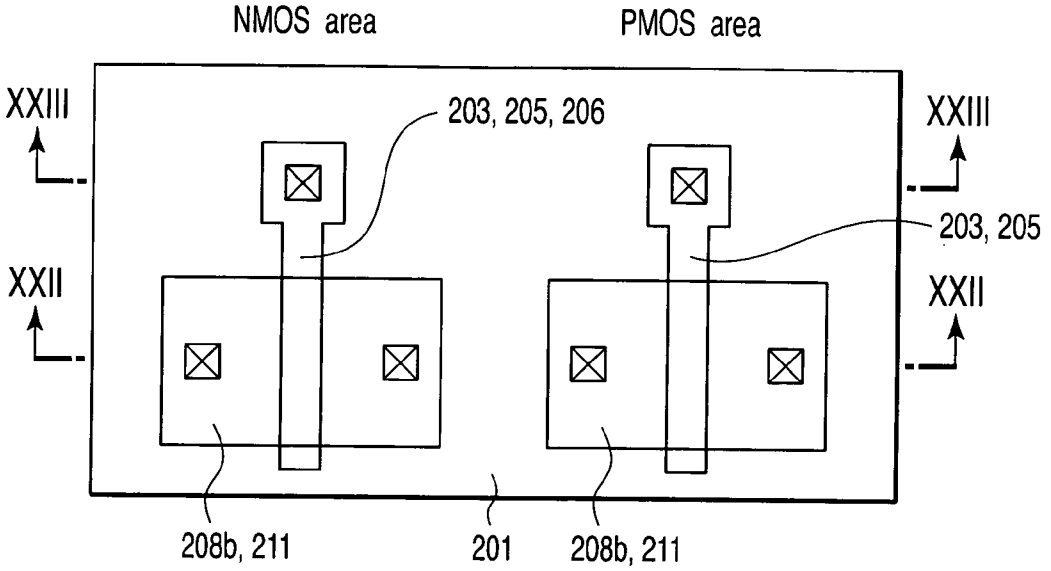

NMOS area
PMOS area
XXIII
XXII
203, 205, 206
203, 205
208b, 211
201
208b, 211

FIG. 21

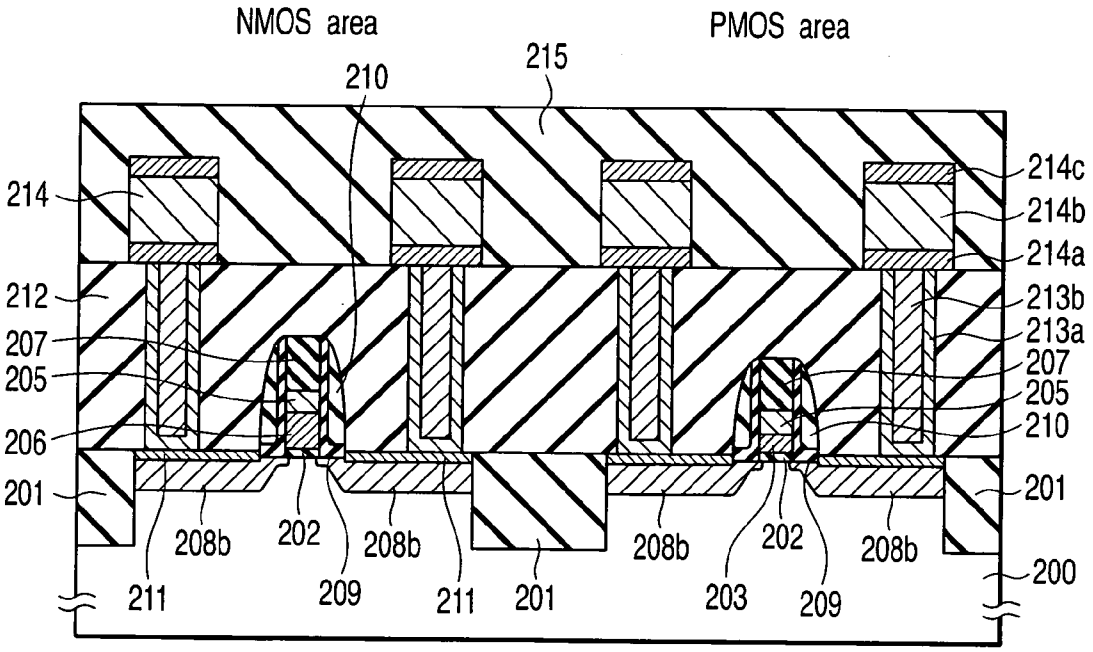

NMOS area
PMOS area
210
215
214
212
207
205
206
201
214c
214b
214a
213b
213a
207
205
210
201
200
208b
202
208b
211
209
211
201
208b
202
208b
203
209

FIG. 22

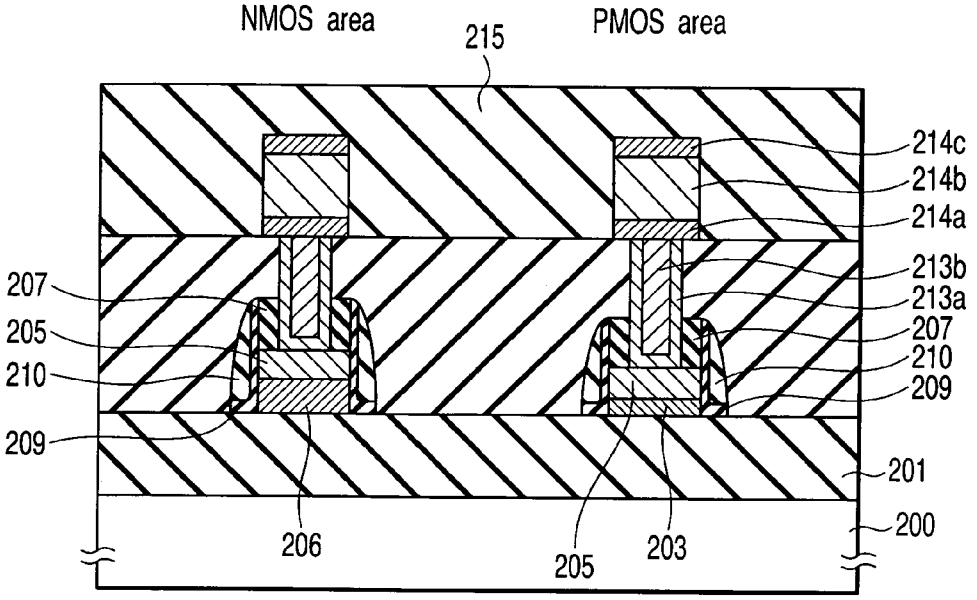

NMOS area
PMOS area
215
214c
214b
214a
213b
213a
207
205
210
209
207
210
209
201
200
206
205
203

FIG. 23

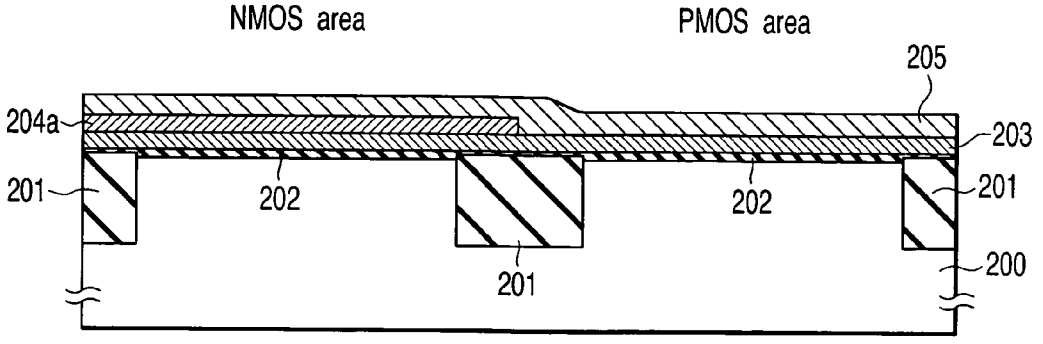

NMOS area
PMOS area
205
204a
203
201
202
202
201
201
200

FIG. 24

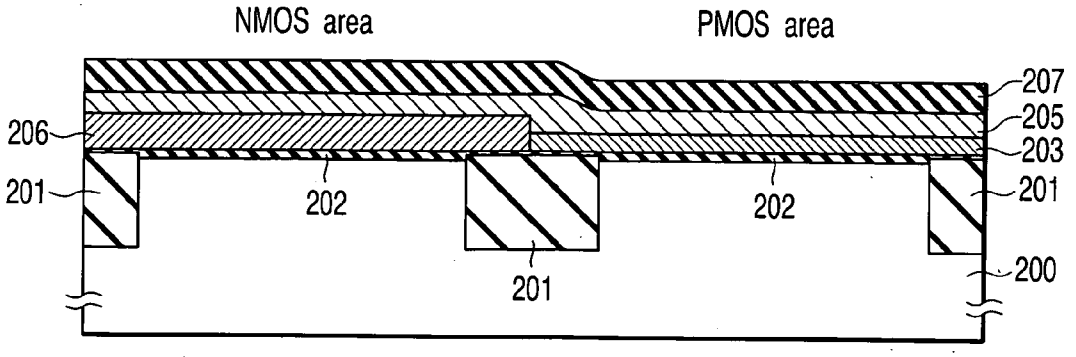


FIG. 25

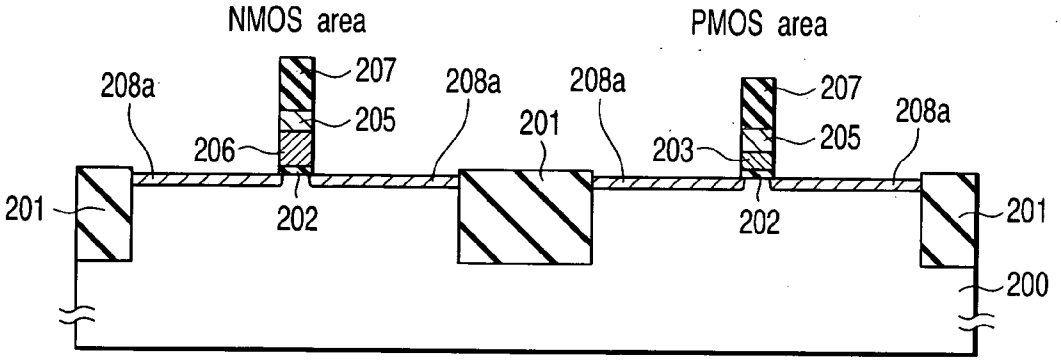


FIG. 26

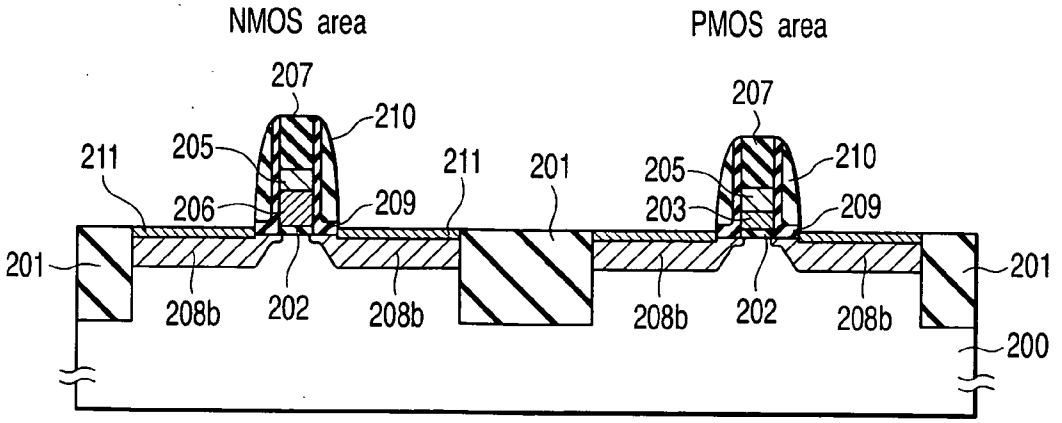


FIG. 27

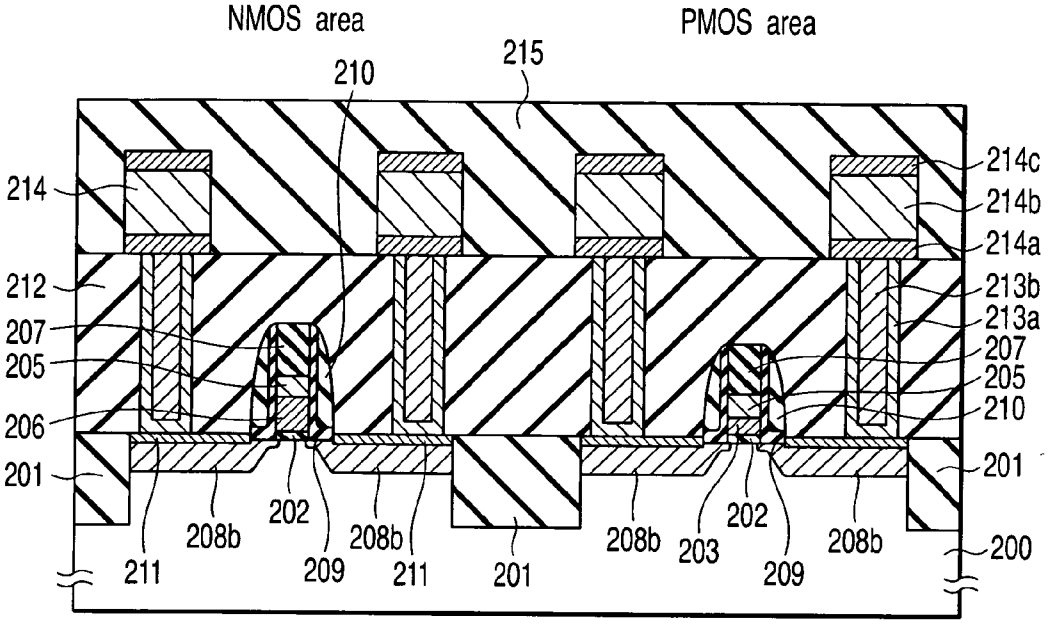

NMOS area
PMOS area
210
215
214
214c
214b
214a
212
213b
213a
207
205
206
201
208b
202
209
211
203

FIG. 28

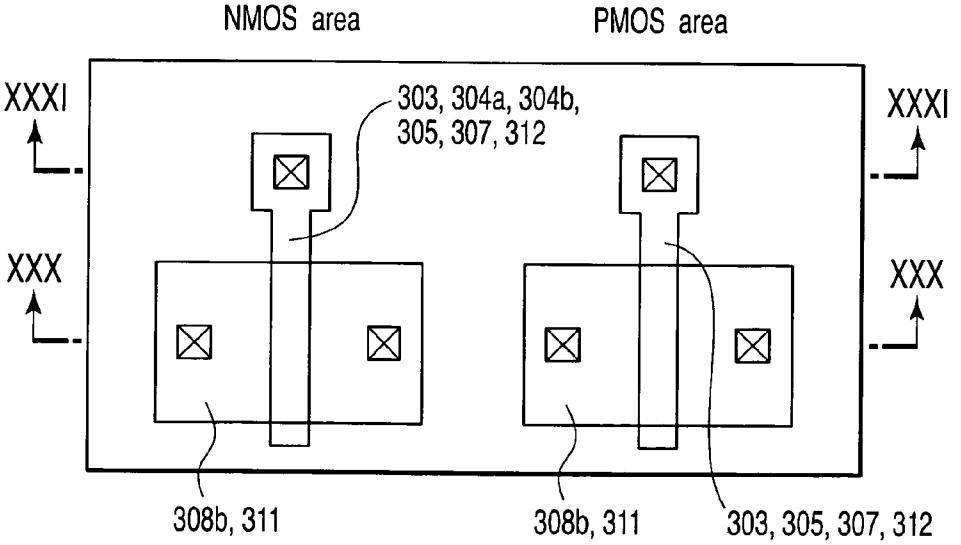

NMOS area
PMOS area
XXXI
XXX
303, 304a, 304b,
305, 307, 312
308b, 311
303, 305, 307, 312

FIG. 29

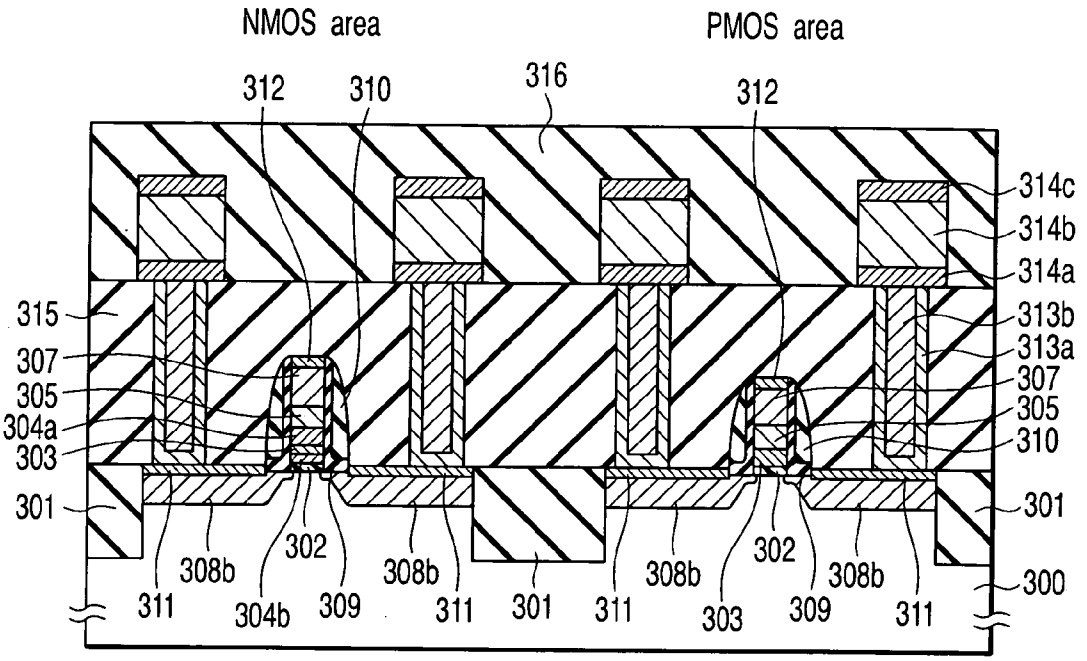

NMOS area
PMOS area
312
310
316
312
314c
314b
314a
315
313b
313a
307
307
305
305
304a
310
303
301
301
308b
302
308b
308b
302
308b
311
304b
309
311
301
311
303
309
311
300

FIG. 30

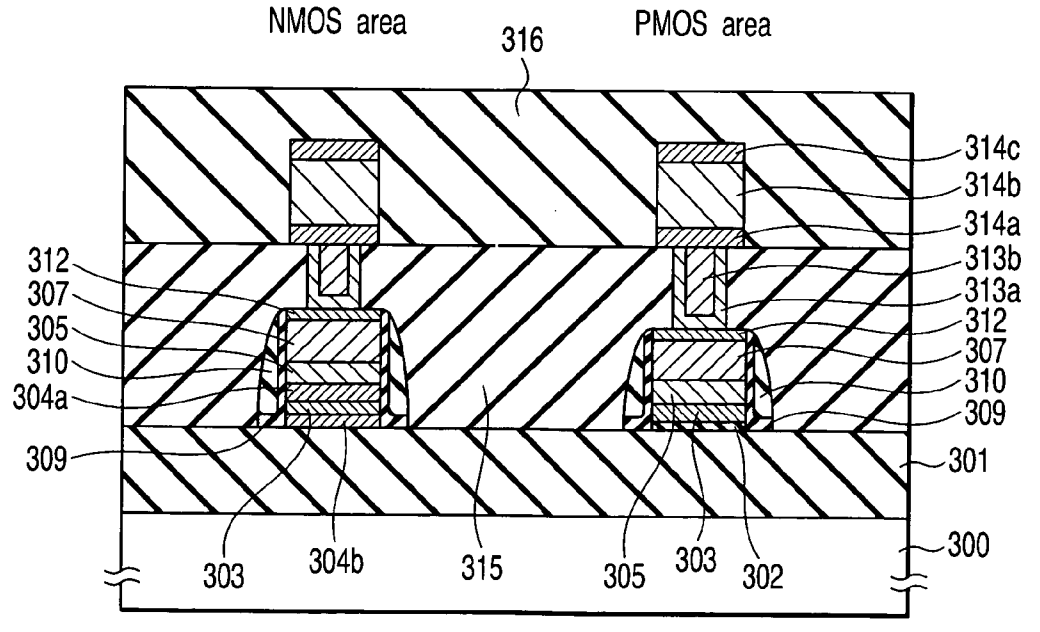

NMOS area
316
PMOS area
314c
314b
314a
312
313b
307
313a
305
312
310
307
304a
310
309
309
301
303
304b
315
305
303
302
300

FIG. 31

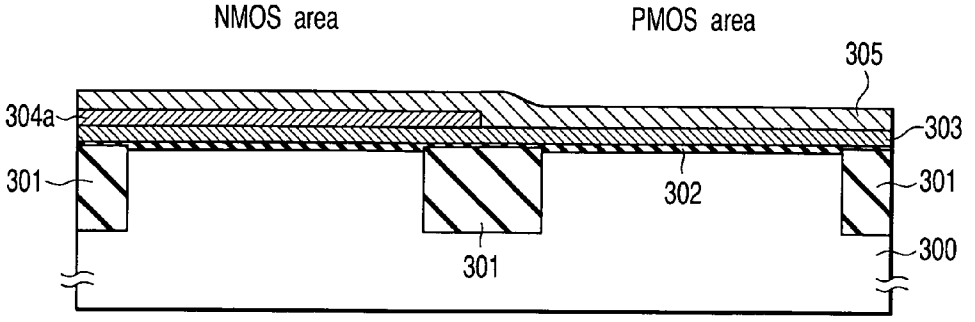


FIG. 32

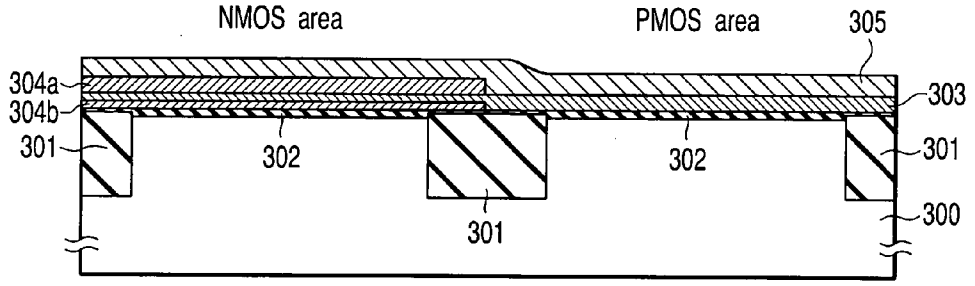


FIG. 33

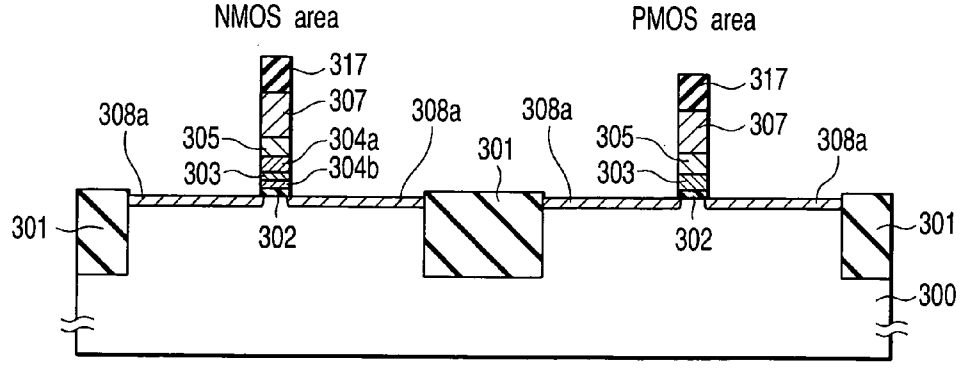


FIG. 34

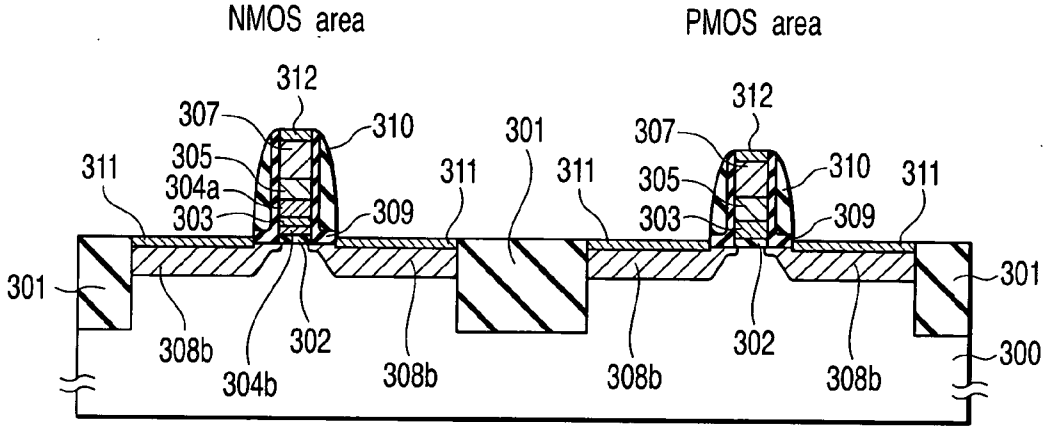

NMOS area
PMOS area
312
307
310
301
305
304a
303
309
311
308b
304b
302
300

FIG. 35

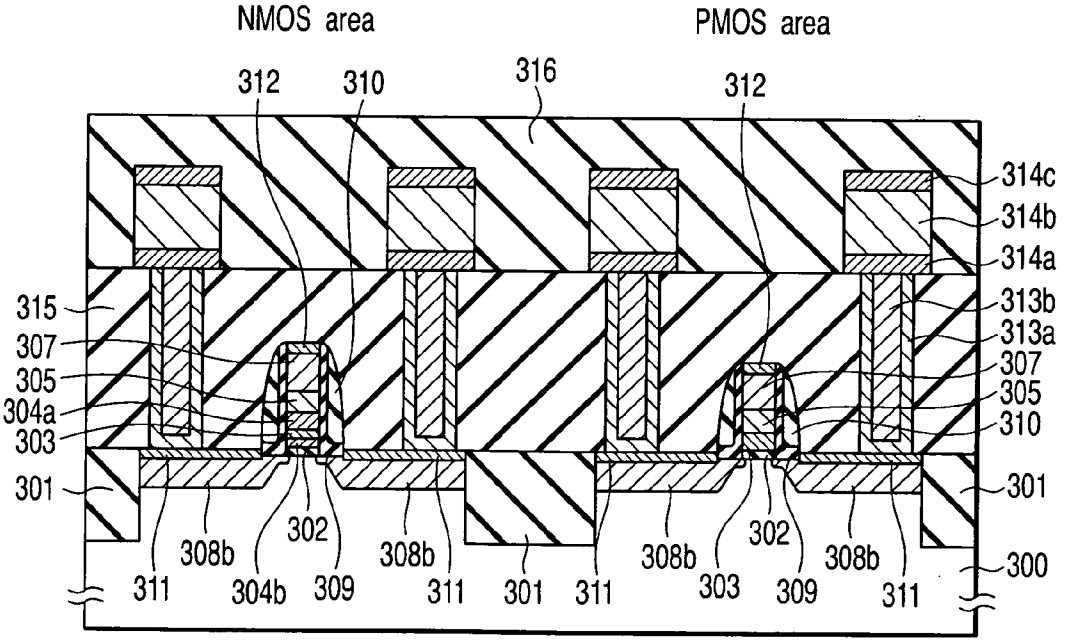

NMOS area
PMOS area
312
310
316
314c
314b
314a
315
307
305
304a
303
313b
313a
309
304b
302
308b
311
301
300

FIG. 36

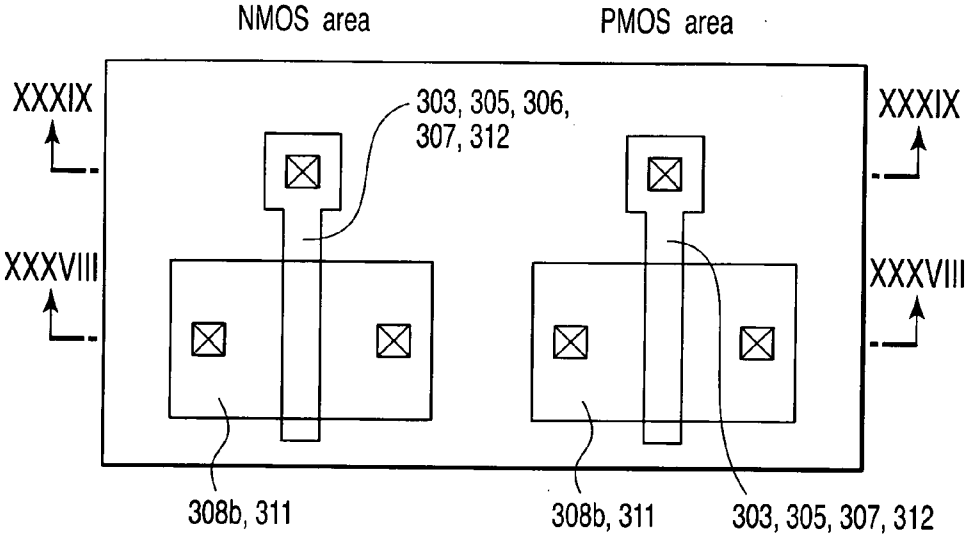

NMOS area
PMOS area
XXXIX
XXXVIII
303, 305, 306,
307, 312
XXXIX
XXXVIII
308b, 311
308b, 311
303, 305, 307, 312

FIG. 37

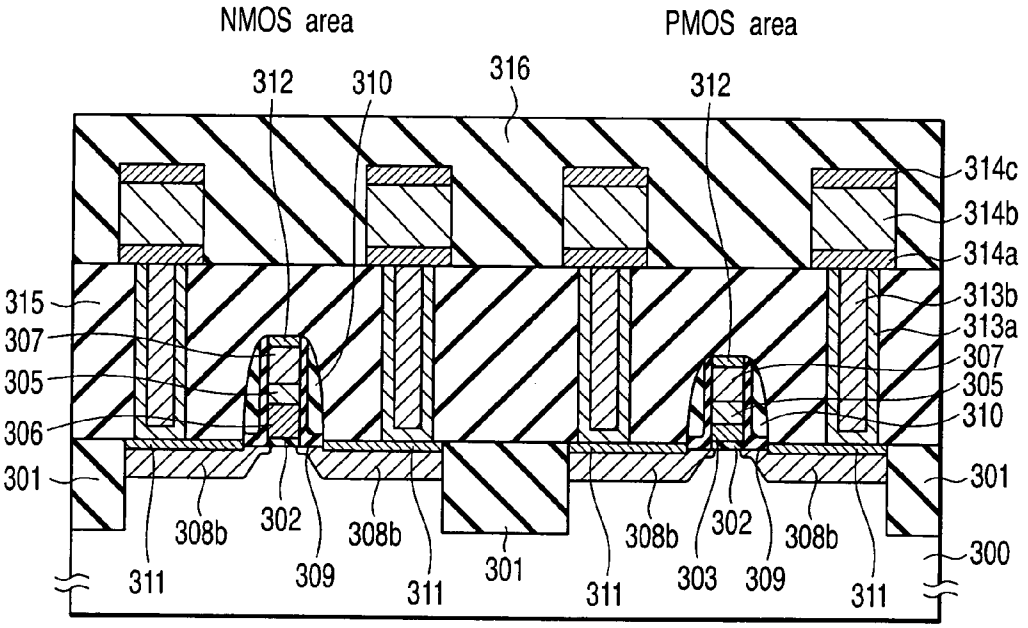

NMOS area
PMOS area
312
310
316
312
314c
314b
314a
315
307
305
306
301
313b
313a
307
305
310
301
308b
302
308b
308b
302
308b
300
311
309
311
301
311
303
309
311

FIG. 38

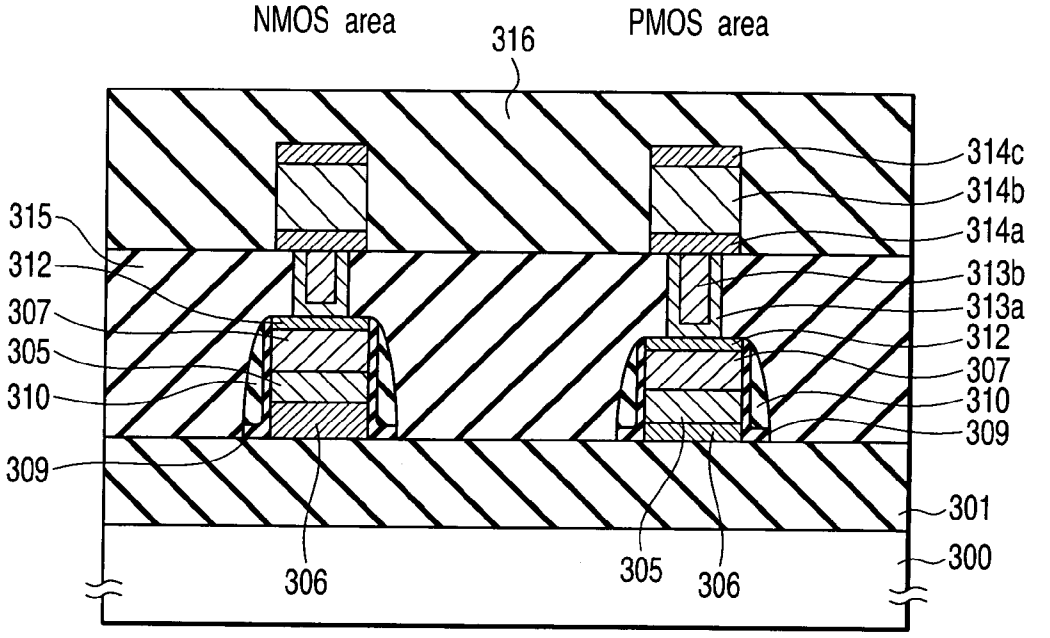

NMOS area
PMOS area
316
314c
314b
314a
315
312
313b
313a
307
312
305
307
310
310
309
309
301
300
306
305
306

FIG. 39

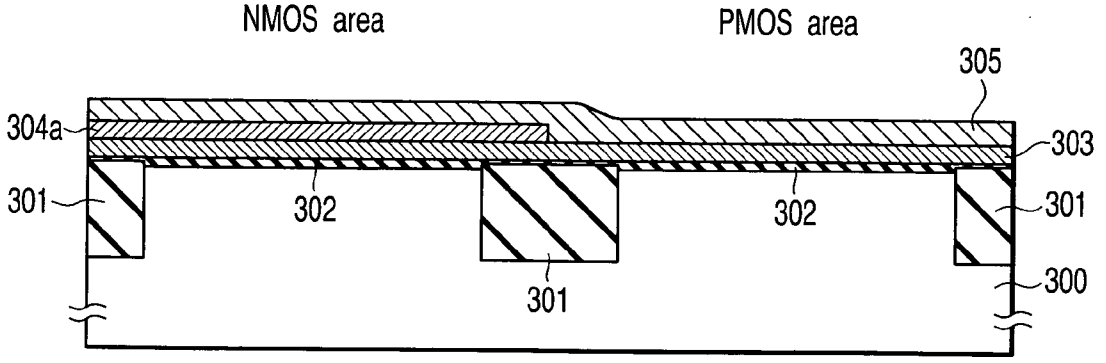

NMOS area
PMOS area
305
304a
303
301
302
302
301
301
300

FIG. 40

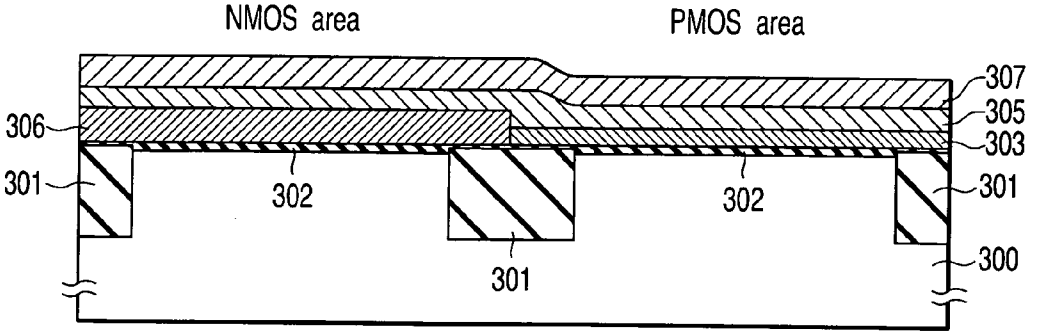


FIG. 41

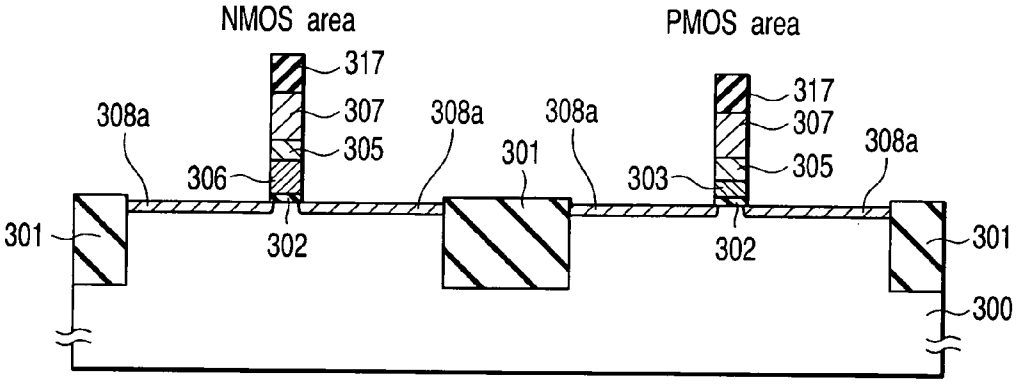


FIG. 42

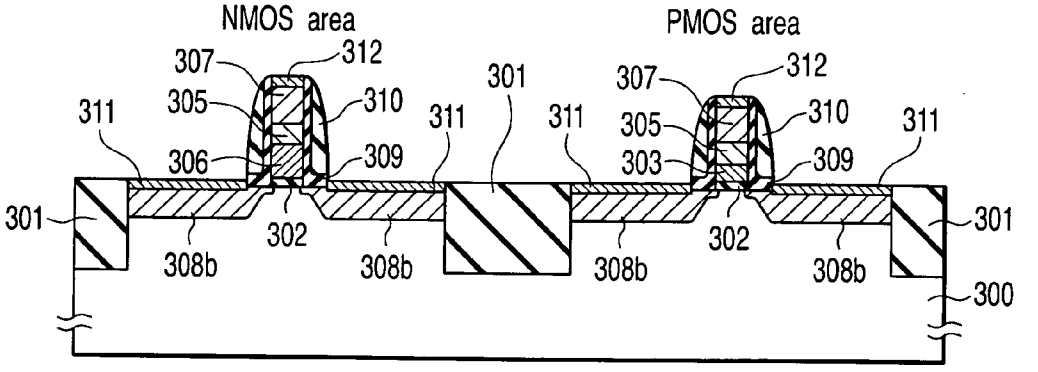


FIG. 43

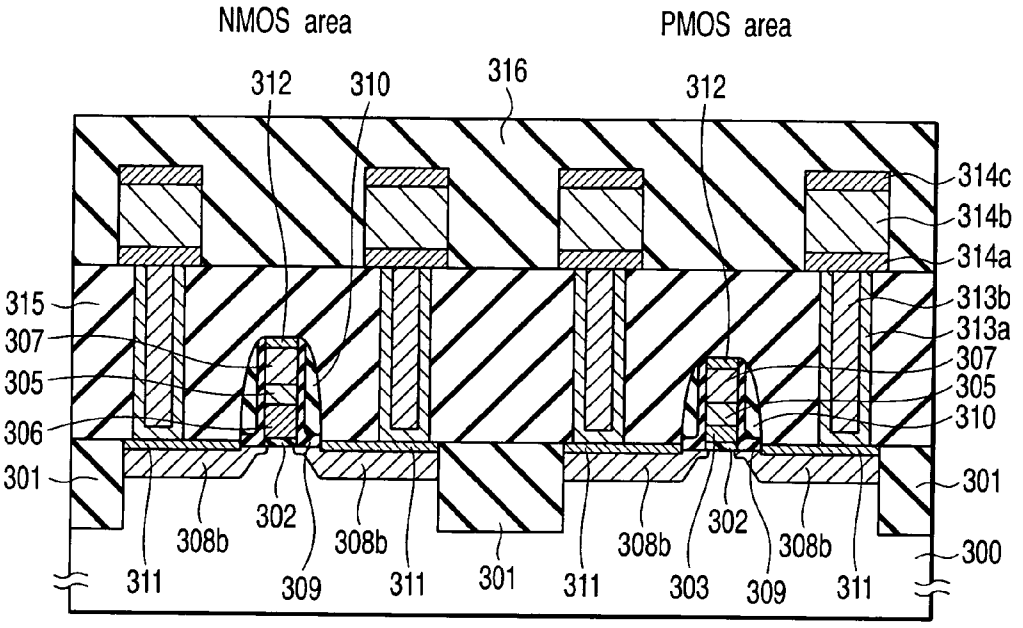

NMOS area
PMOS area
312
310
316
312
314c
314b
314a
313b
313a
307
305
310
315
307
305
306
301
301
308b
302
308b
308b
302
308b
300
311
309
311
301
311
303
309
311

FIG. 44

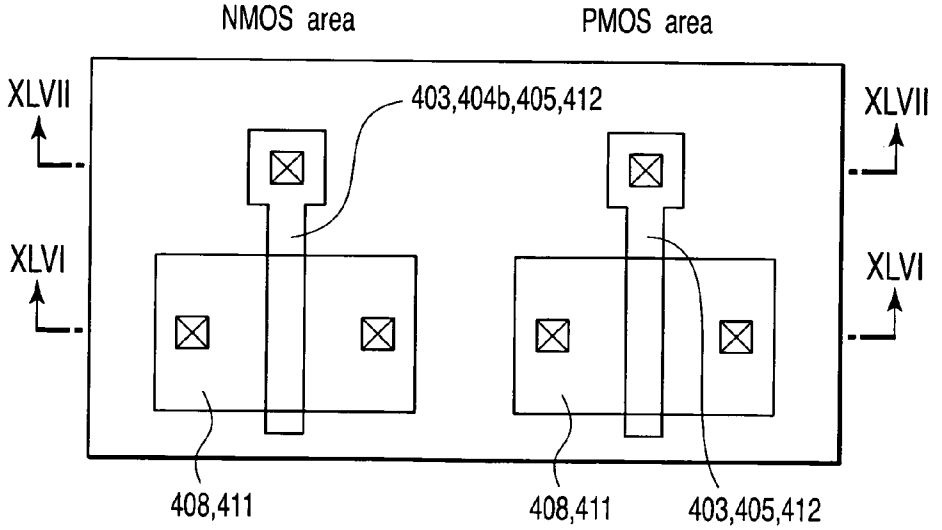

NMOS area
PMOS area
XLVII
XLVI
403,404b,405,412
XLVII
XLVI
408,411
408,411
403,405,412

FIG. 45

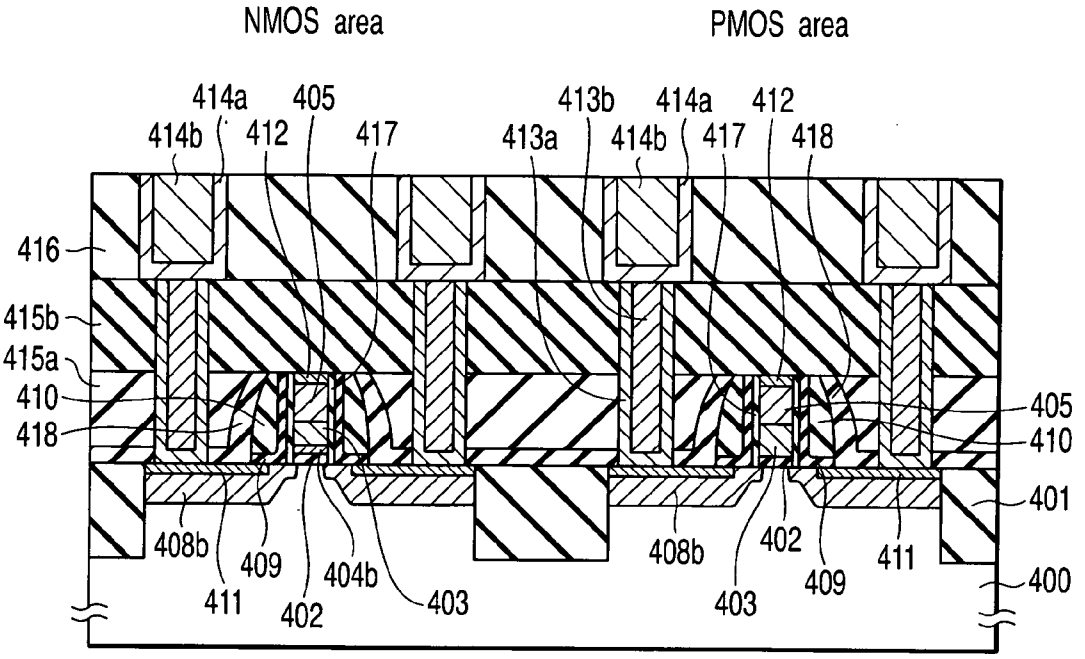

NMOS area
PMOS area
414a
405
414b
412
417
413b
414a
412
413a
414b
417
418
416
415b
415a
410
418
405
410
401
408b
409
404b
411
402
403
408b
402
411
409
403
400

FIG. 46

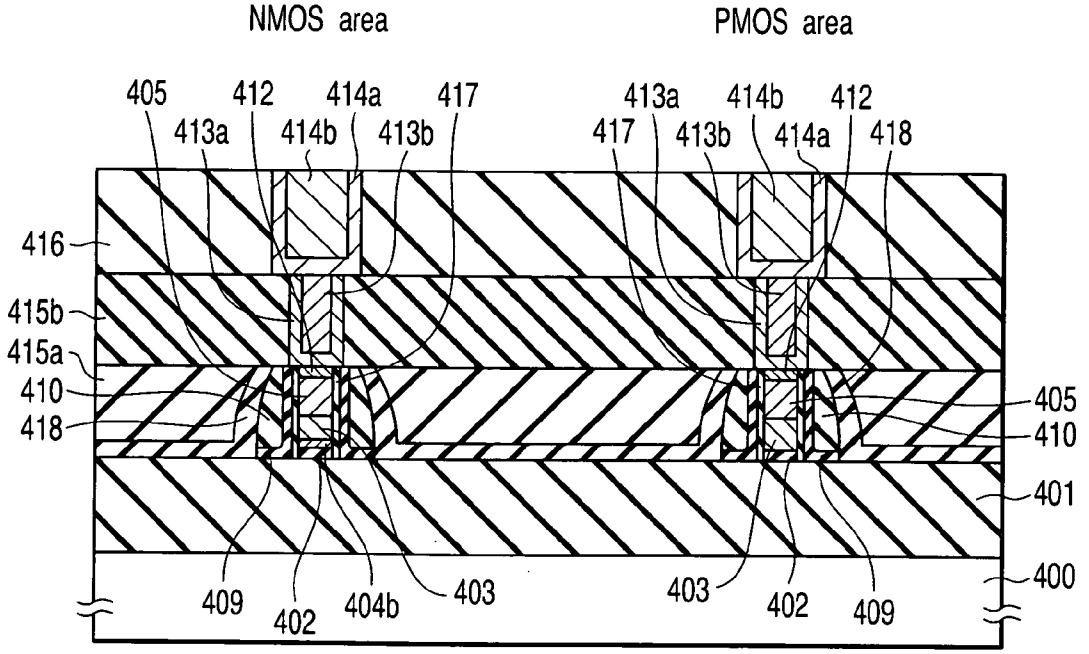

NMOS area
PMOS area
405
412
414a
417
413a
414b
413b
413a
414b
412
417
413b
414a
418
416
415b
415a
410
418
405
410
401
409
402
404b
403
403
402
409
400

FIG. 47

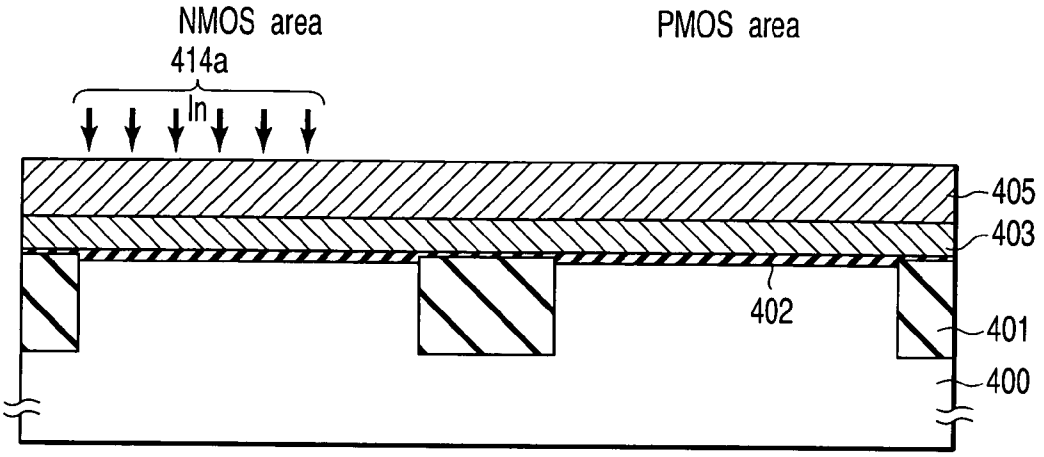


FIG. 48

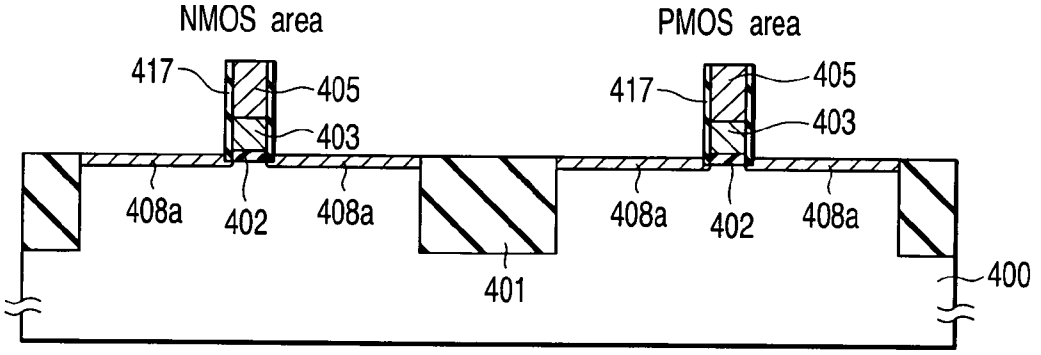


FIG. 49

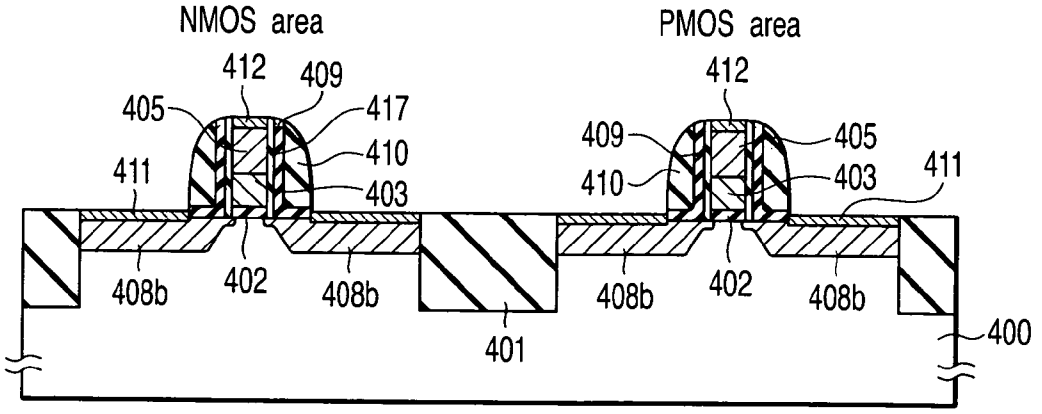


FIG. 50

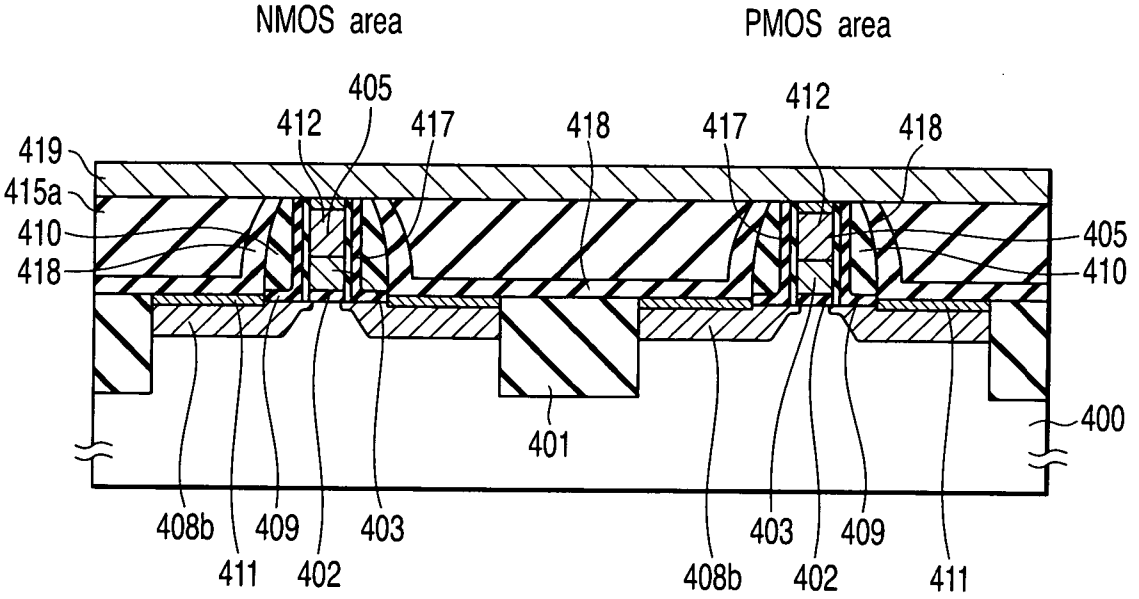

NMOS area
PMOS area
405
412
417
418
417
412
418
419
415a
410
418
405
410
400
401
408b
409
403
411
402
403
409
408b
402
411

FIG. 51

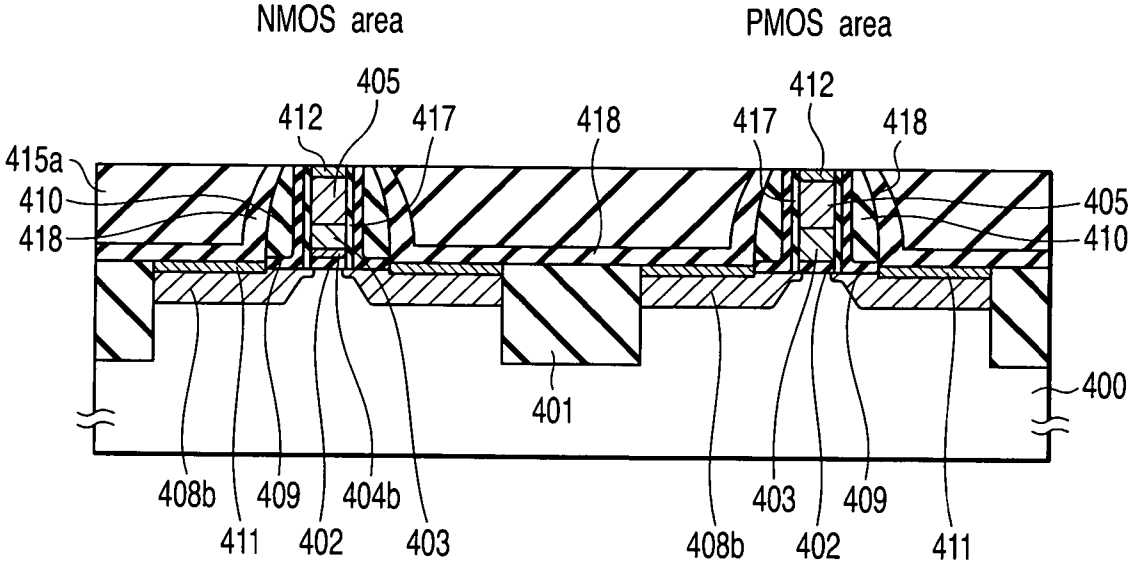

NMOS area
PMOS area
405
412
417
418
417
412
418
415a
410
418
405
410
400
401
408b
409
404b
403
411
402
403
409
408b
402
411

FIG. 52

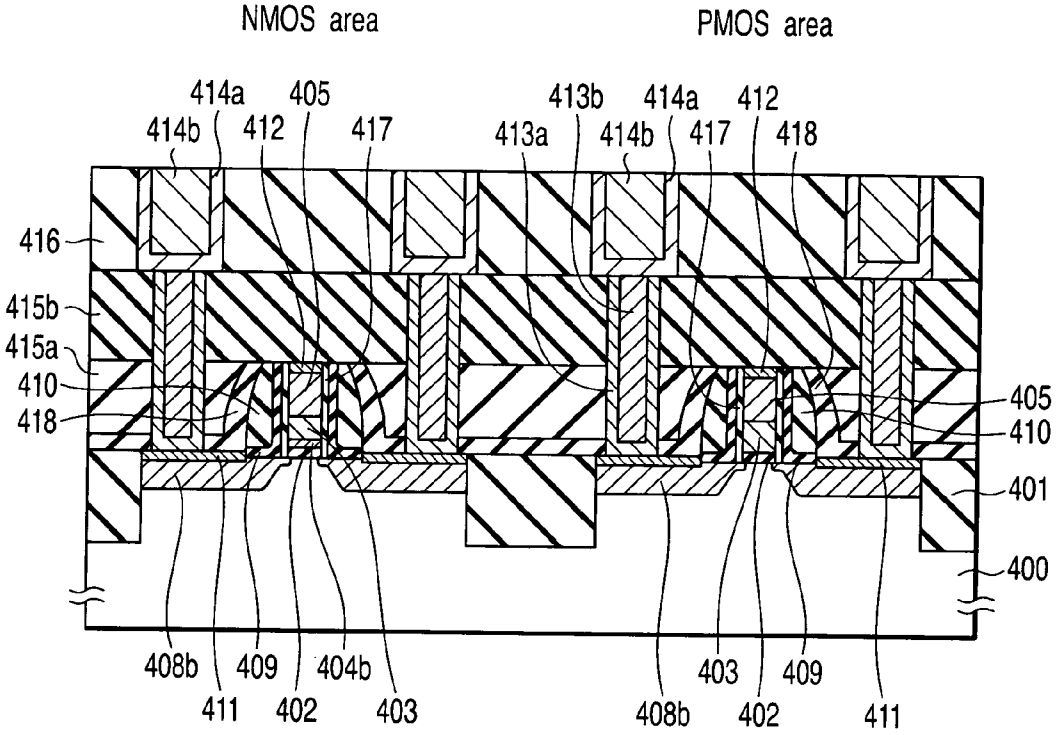

NMOS area
PMOS area
414a
405
414b
412
417
413b
414a
412
413a
414b
417
418
416
415b
415a
410
418
405
410
401
400
408b
409
404b
411
402
403
403
409
408b
402
411

FIG. 53

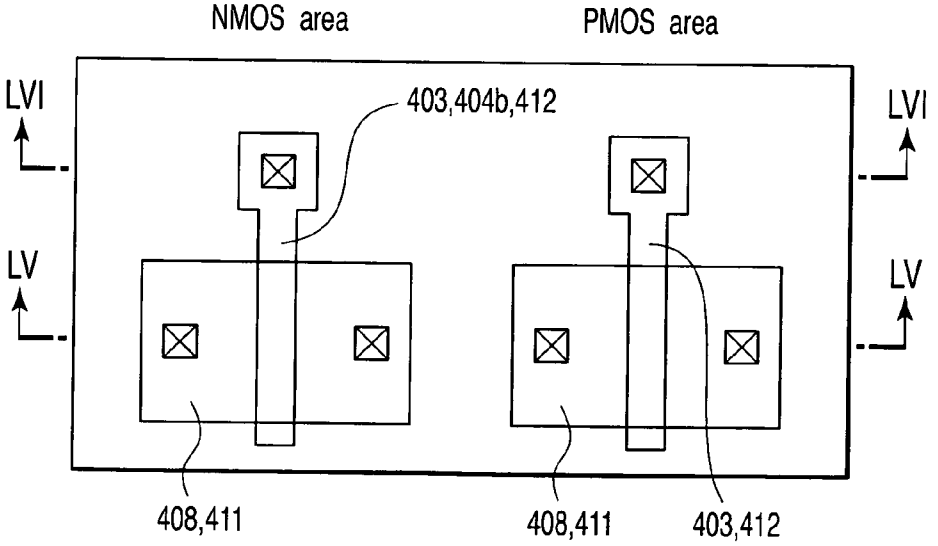

NMOS area
PMOS area
LVI
403,404b,412
LVI
LV
LV
408,411
408,411
403,412

FIG. 54

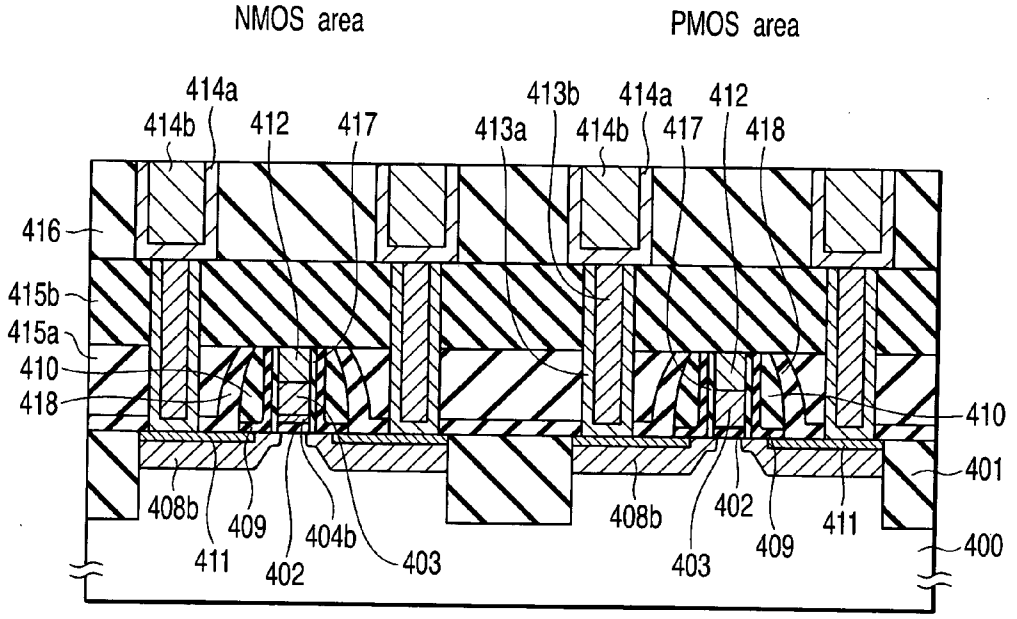

NMOS area
PMOS area
414a
414b
412
417
413b
413a
414b
414a
412
417
418
416
415b
415a
410
418
410
401
408b
409
411
402
404b
403
408b
403
402
409
411
400

FIG. 55

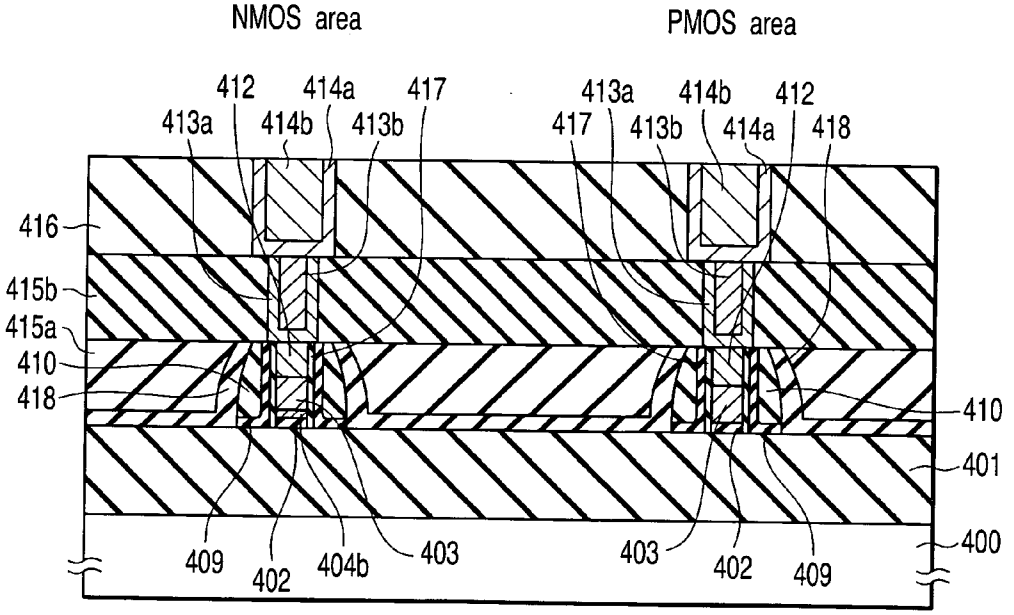

NMOS area
PMOS area
412
414a
417
413a
414b
413b
413a
414b
412
417
413b
414a
418
416
415b
415a
410
418
410
401
409
402
404b
403
403
402
409
400

FIG. 56

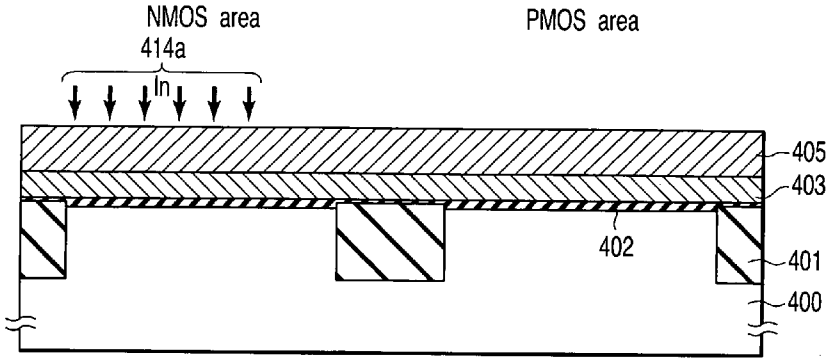


FIG. 57

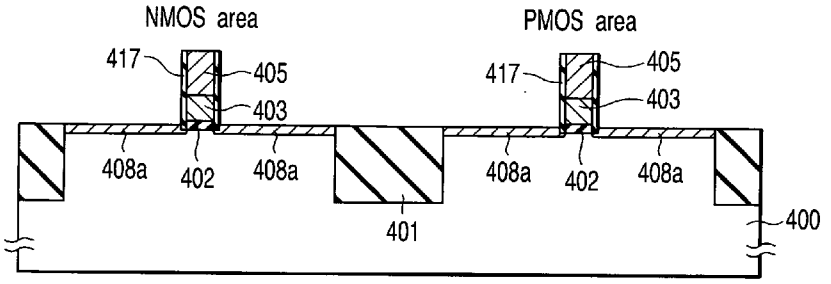


FIG. 58

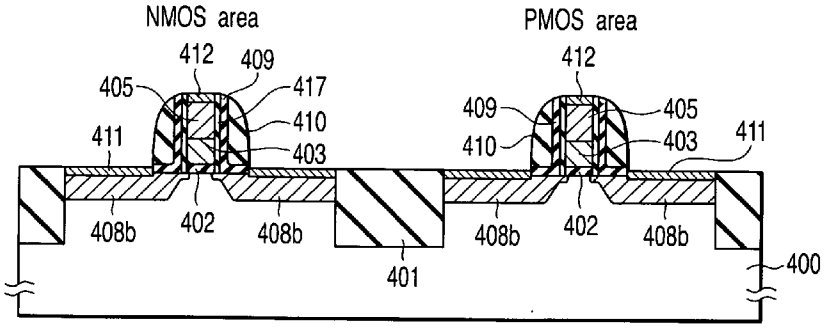


FIG. 59

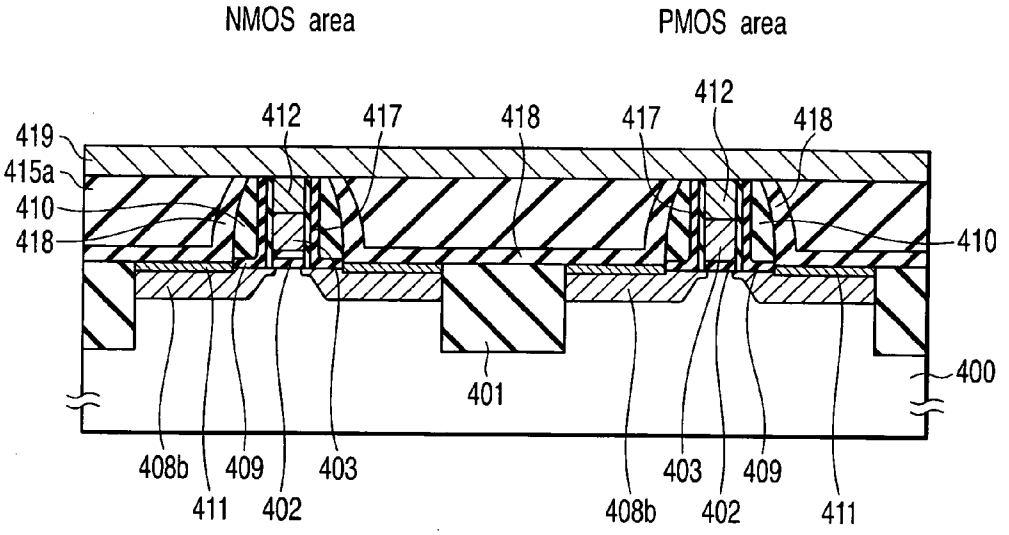

NMOS area
PMOS area
419
415a
410
418
412
417
418
417
412
418
410
400
401
408b
409
403
411
402
403
409
408b
402
411

FIG. 60

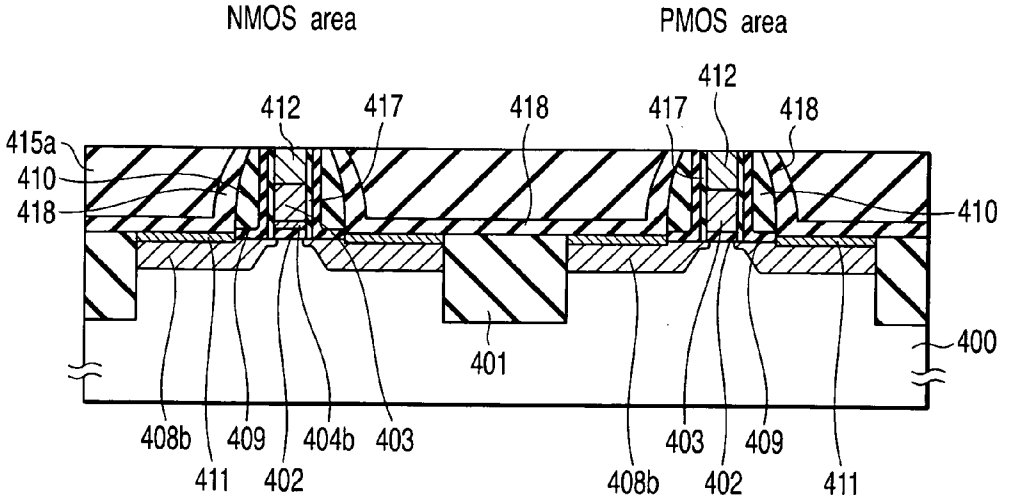

NMOS area
PMOS area
415a
410
418
412
417
418
417
412
418
410
400
401
408b
409
404b
403
411
402
403
409
408b
402
411

FIG. 61

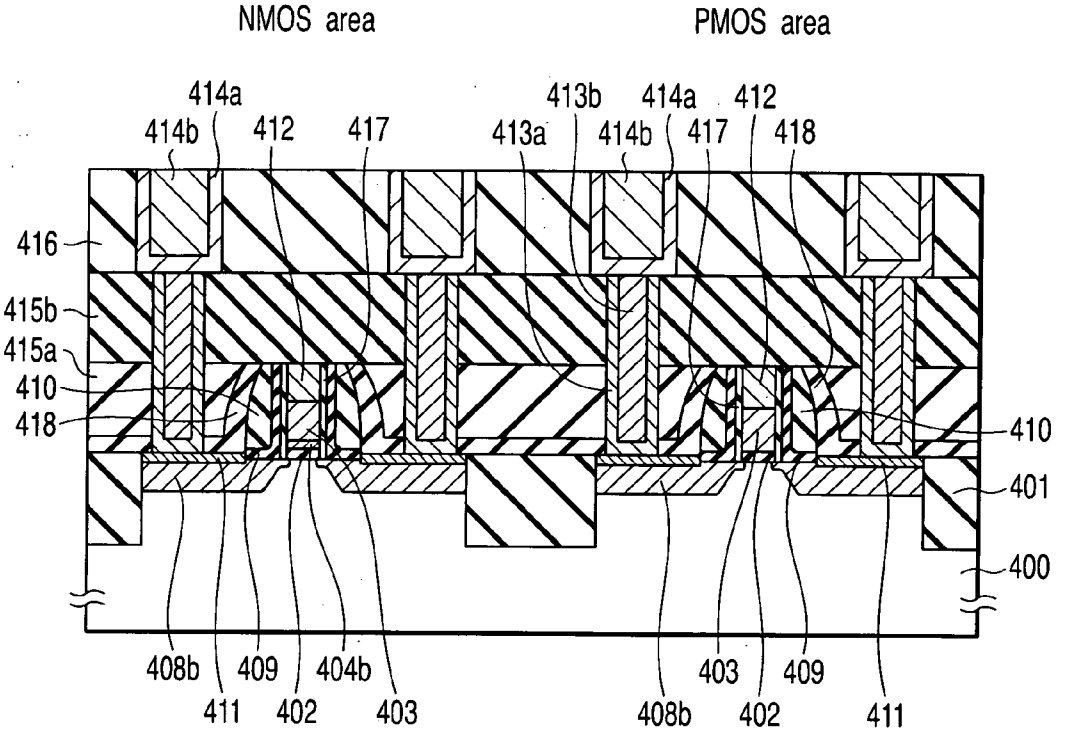

NMOS area
PMOS area
414a
414b
412
417
413b
413a
414a
414b
412
417
418
416
415b
415a
410
418
410
401
400
408b
409
404b
411
402
403
403
409
408b
402
411

FIG. 62

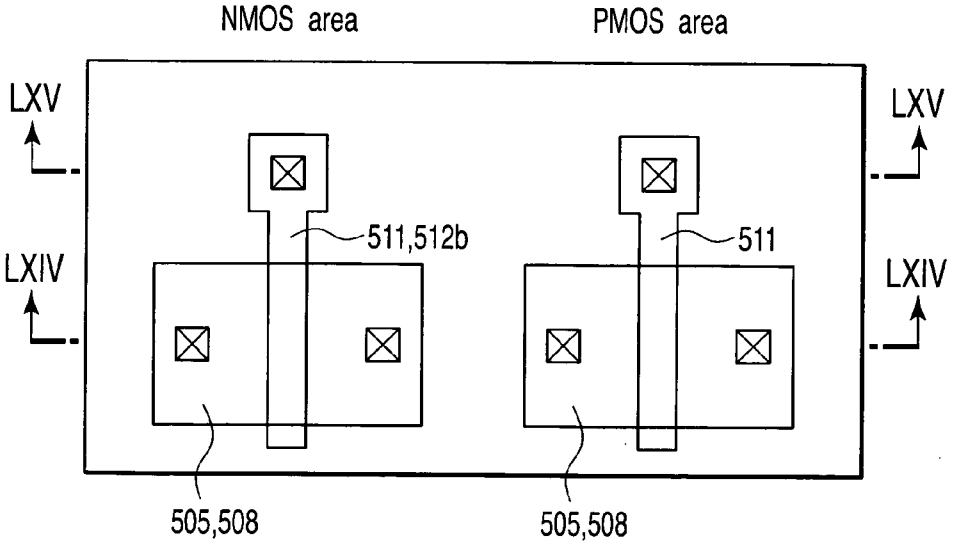

NMOS area
PMOS area
LXV
LXIV
511,512b
511
LXV
LXIV
505,508
505,508

FIG. 63

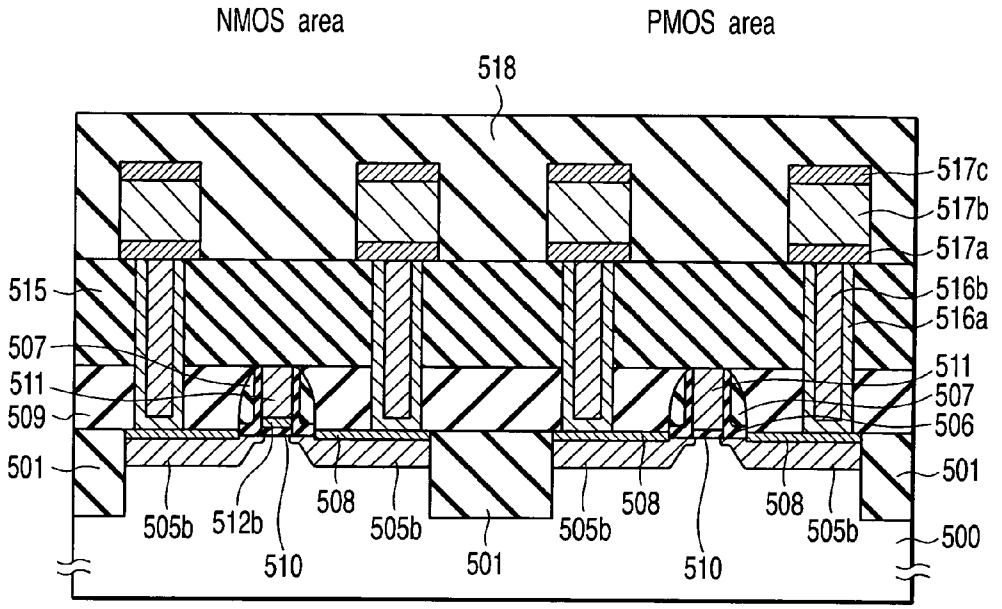

NMOS area
PMOS area
518
517c
517b
517a
515
516b
516a
507
511
509
506
501
505b
512b
508
510
501
500

FIG. 64

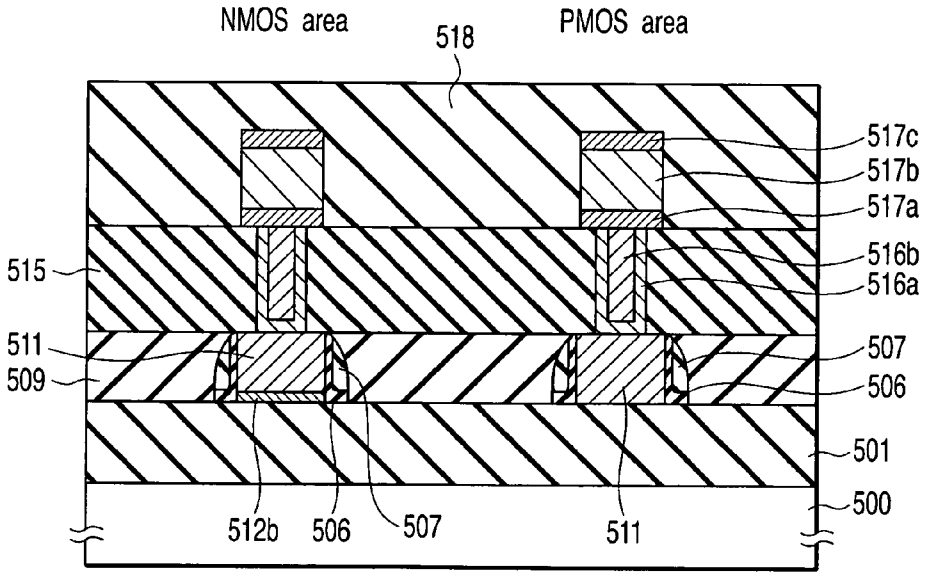

NMOS area
518
PMOS area
517c
517b
517a
515
516b
516a
511
509
507
506
501
500
512b
506
507
511

FIG. 65

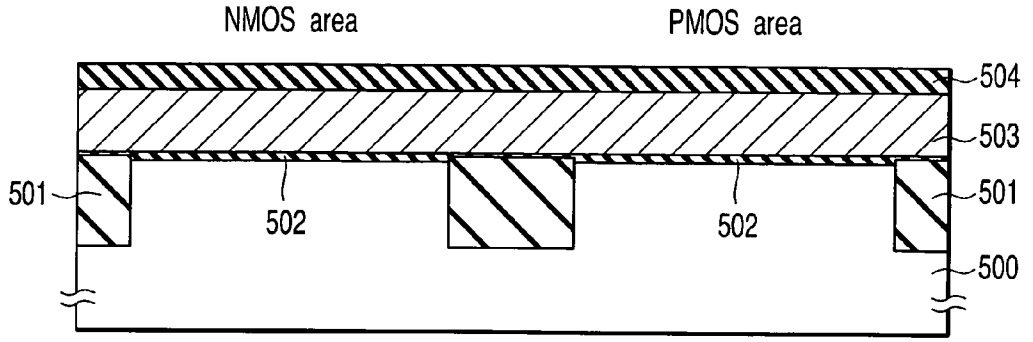


FIG. 66

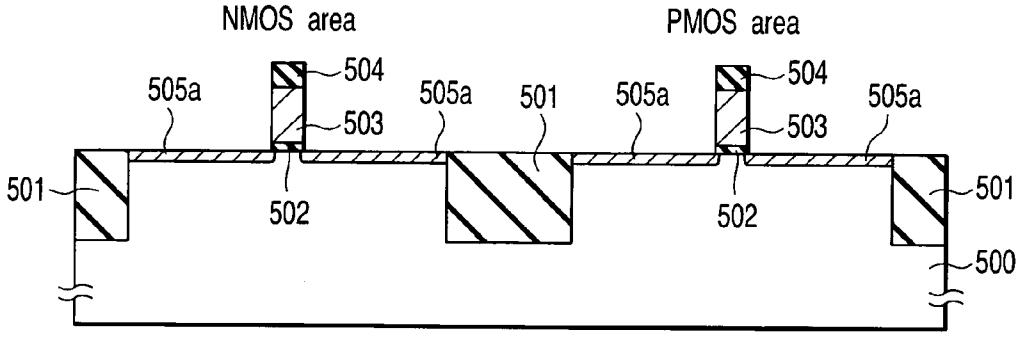


FIG. 67

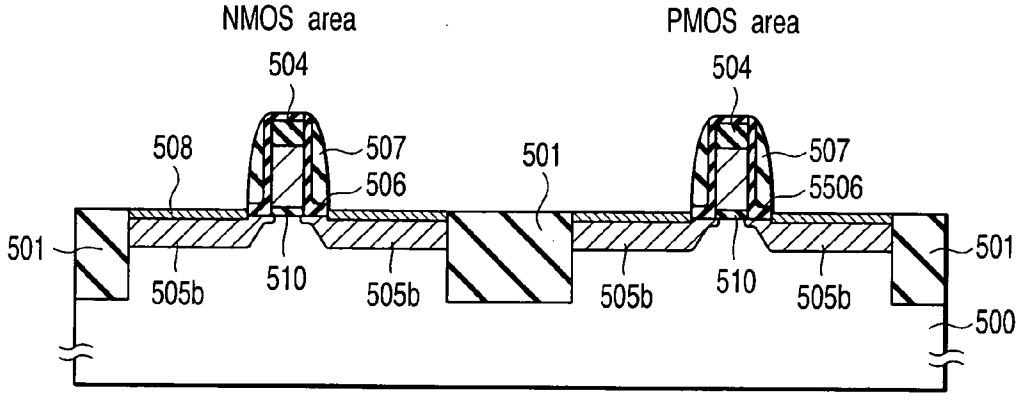


FIG. 68

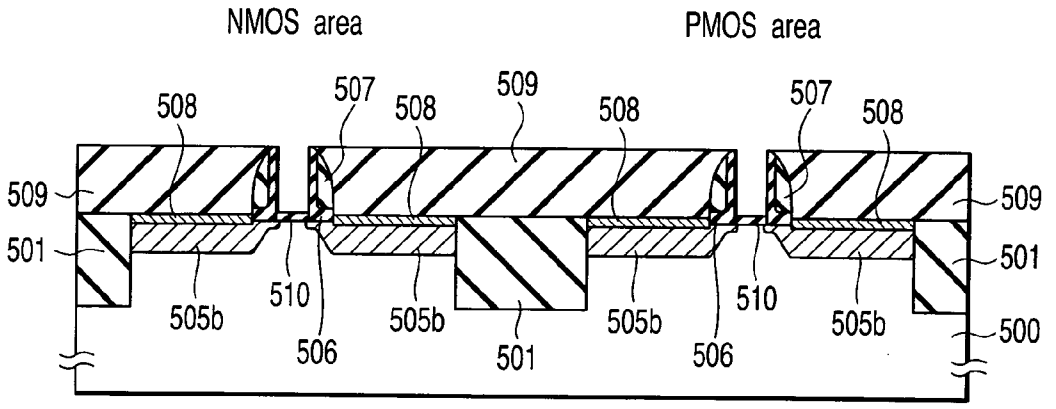

NMOS area
PMOS area
508
507
508
509
508
507
508
509
501
505b
510
506
505b
501
505b
510
505b
509
501
500

FIG. 69

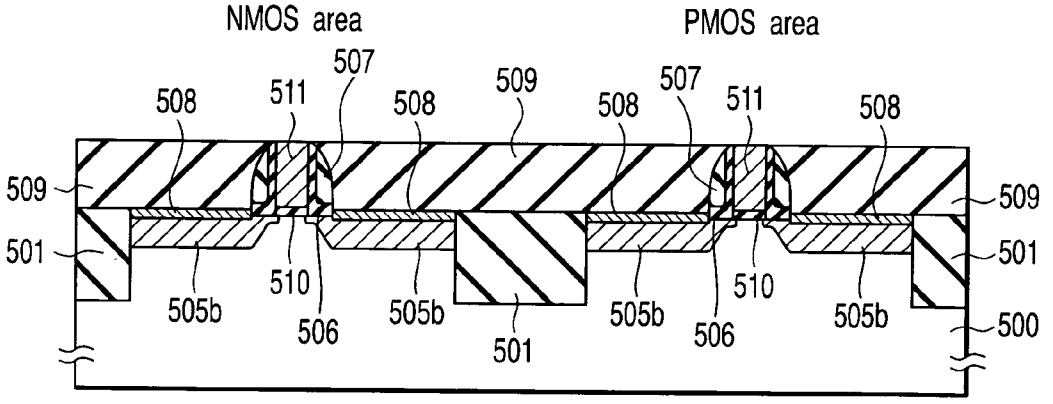

NMOS area
PMOS area
508
511
507
508
509
508
507
511
508
509
501
505b
510
506
505b
501
505b
510
505b
509
501
500

FIG. 70

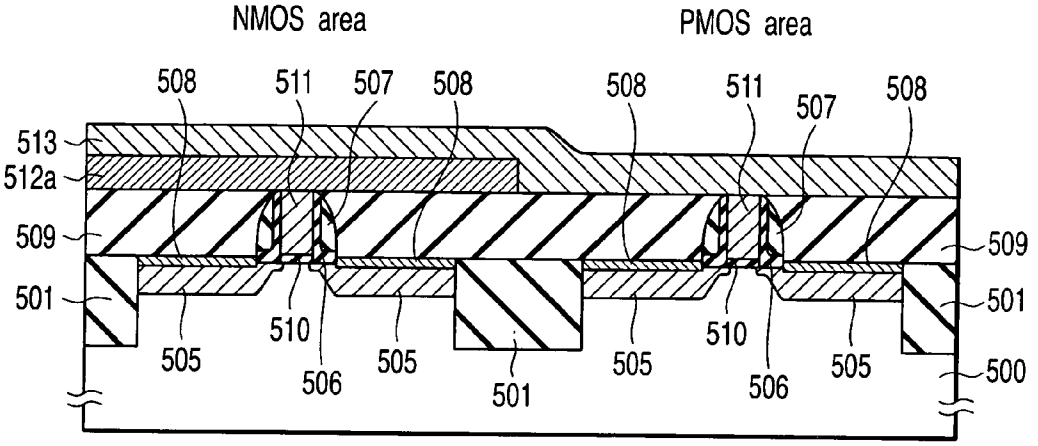

NMOS area
PMOS area
508
511
507
508
508
511
507
508
513
512a
509
501
509
501
505
510
506
505
501
505
510
506
505
500

FIG. 71

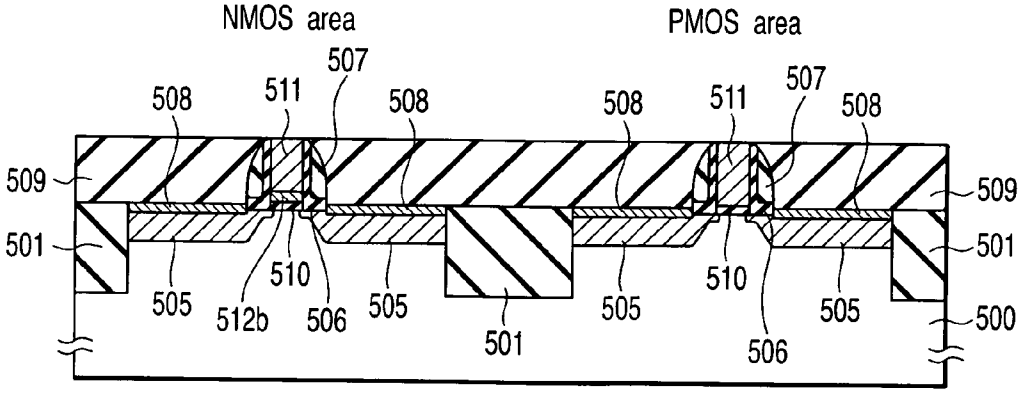

NMOS area
PMOS area
507
508
511
508
508
511
507
508
509
501
509
501
505
512b
510
506
505
501
505
510
506
505
500

FIG. 72

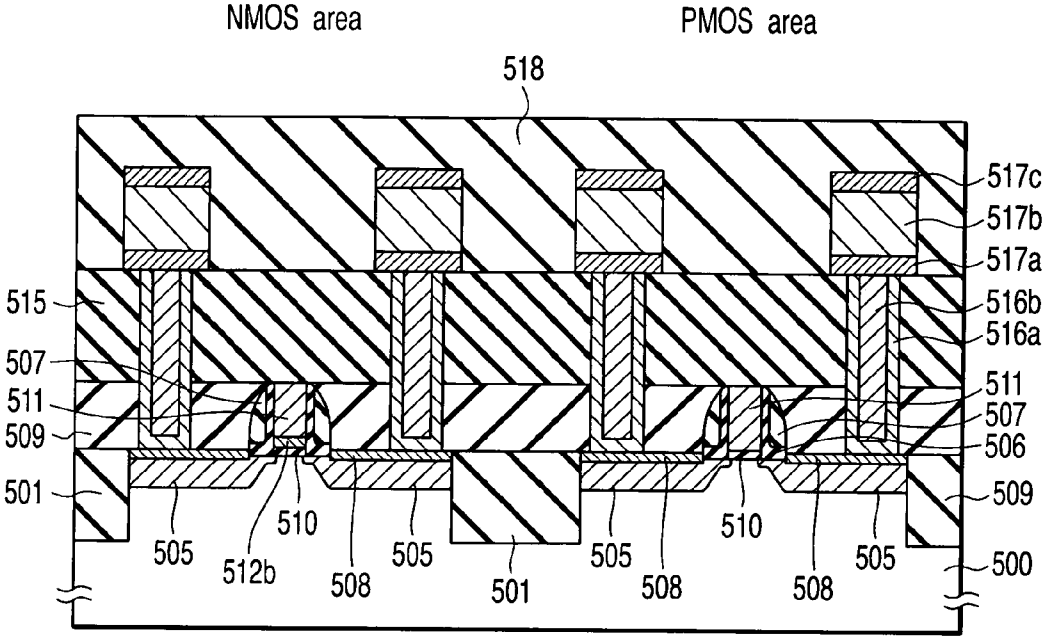

NMOS area
PMOS area
518
517c
517b
517a
515
516b
516a
507
511
509
501
505
512b
510
508
505
501
505
508
510
508
505
509
500

FIG. 73

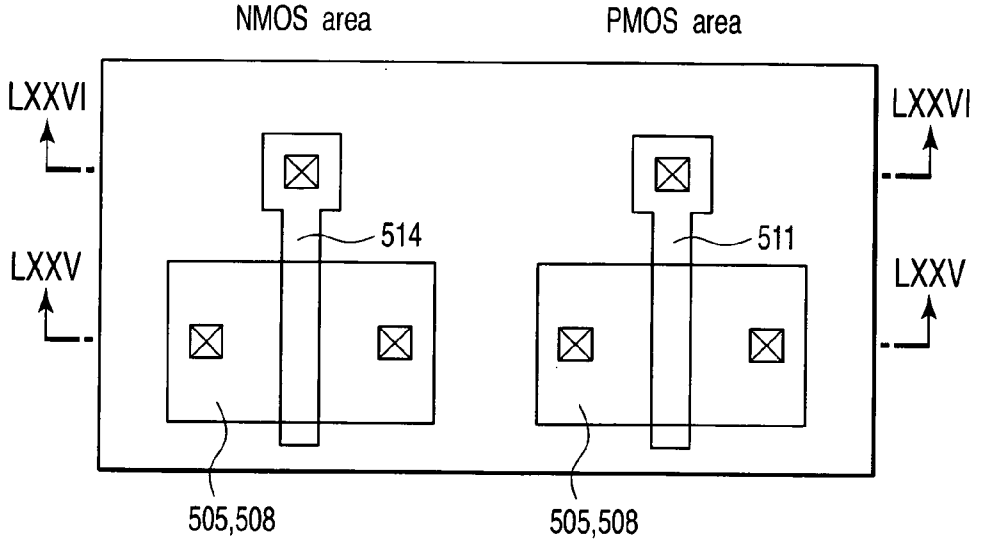

NMOS area
PMOS area
LXXVI
LXXV
514
511
LXXVI
LXXV
505,508
505,508

FIG. 74

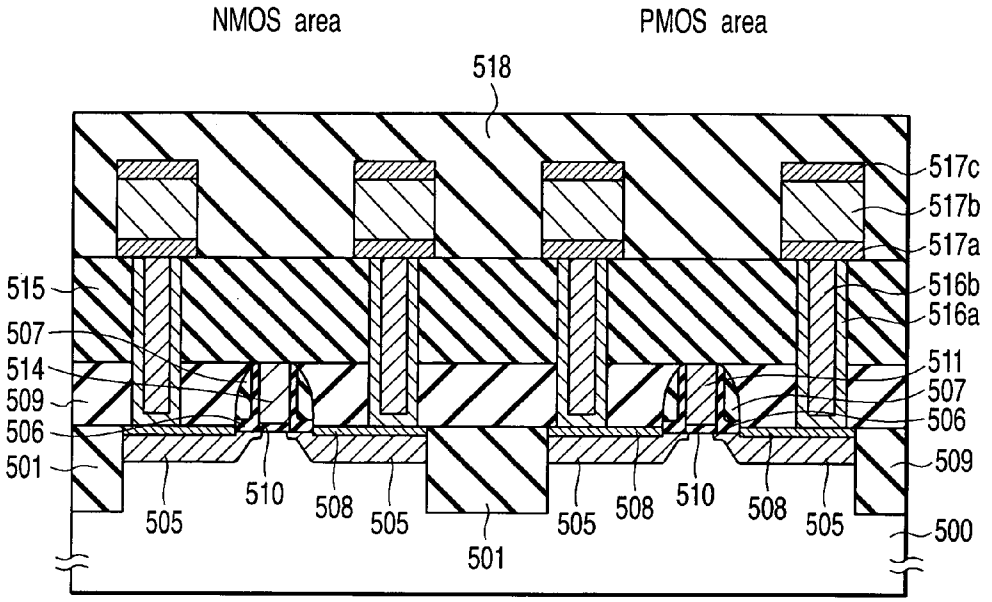

NMOS area
PMOS area
518
517c
517b
517a
516b
516a
515
507
514
509
506
501
511
507
506
509
500
505
510
508
505
501
505
508
510
508
505

FIG. 75

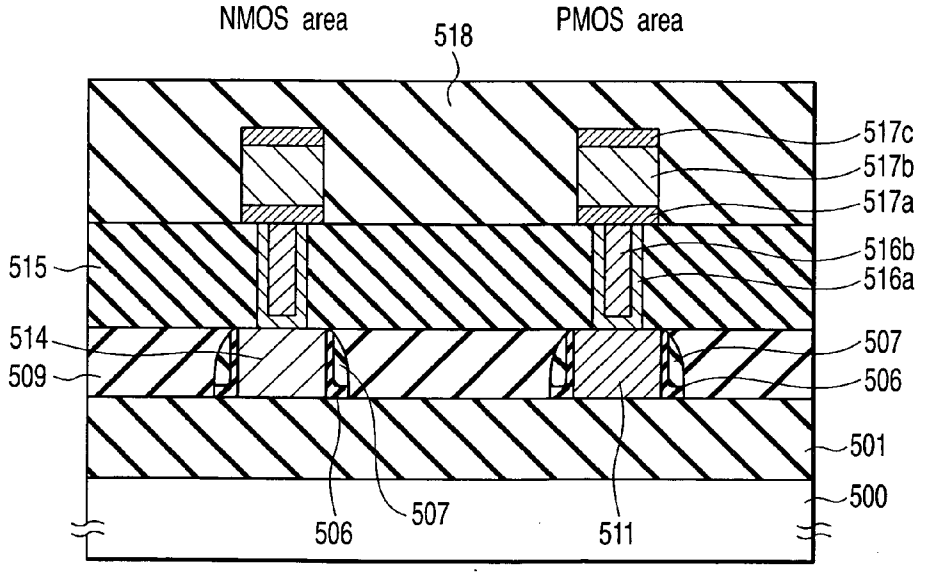

NMOS area
PMOS area
518
517c
517b
517a
516b
516a
515
514
509
507
506
501
500
506
507
511

FIG. 76

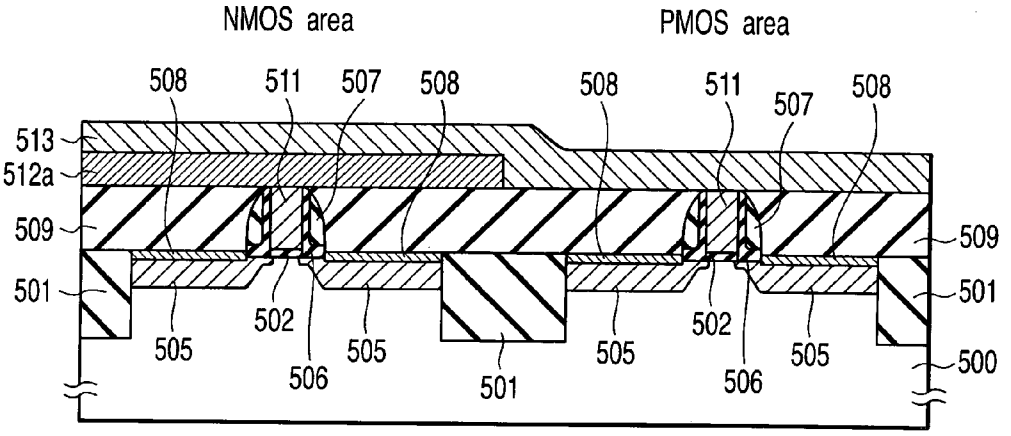

NMOS area
PMOS area
508
511
507
508
508
511
507
508
513
512a
509
509
501
501
505
502
506
505
501
505
502
506
505
500

FIG. 77

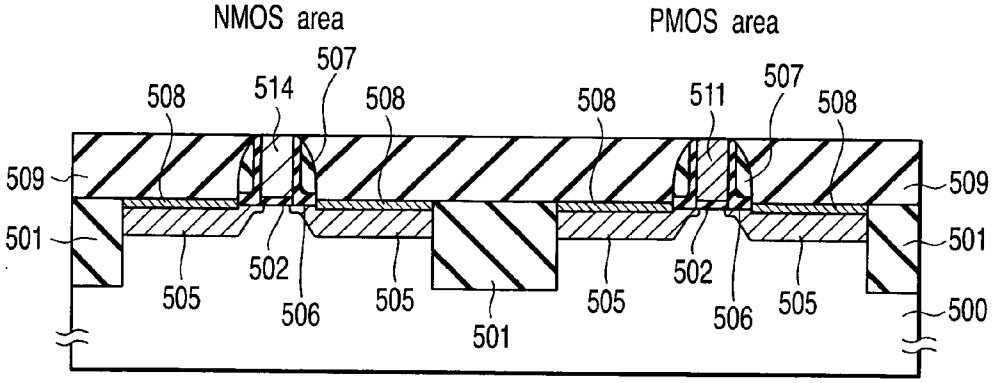

NMOS area
PMOS area
507
508
514
508
508
511
507
508
509
509
501
501
505
502
506
505
501
505
502
506
505
500

FIG. 78

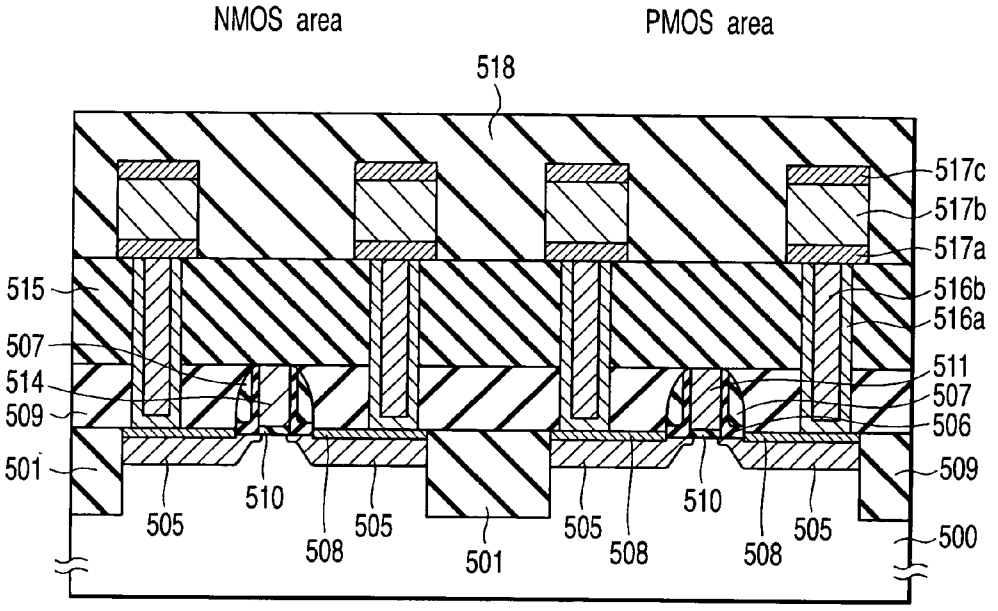

NMOS area
PMOS area
518
517c
517b
517a
515
516b
516a
507
514
509
511
507
506
501
509
505
510
508
505
501
505
508
510
508
505
500

FIG. 79

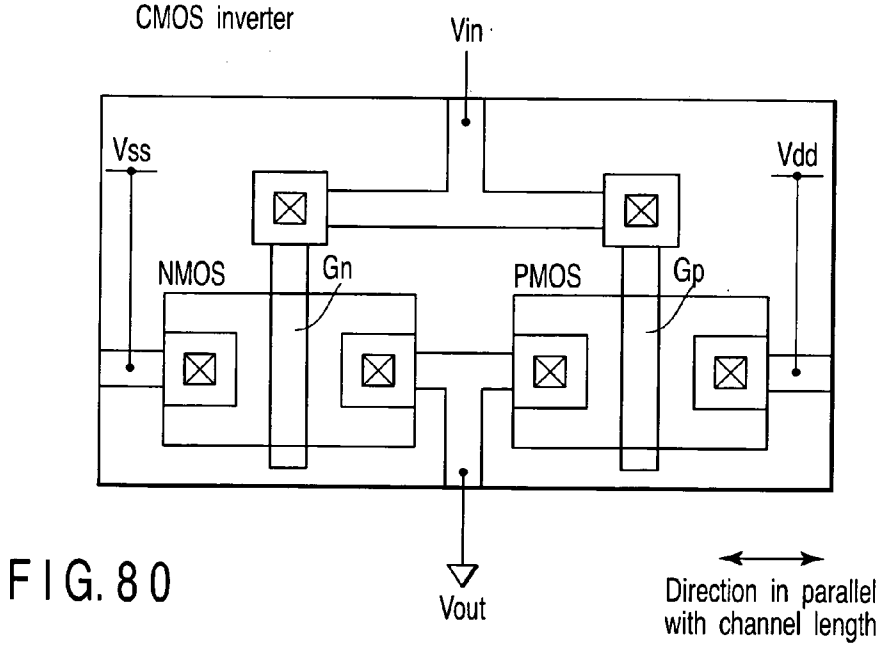

CMOS inverter
Vin
Vss
Vdd
NMOS
Gn
PMOS
Gp
Vout
Direction in parallel with channel length

FIG. 80

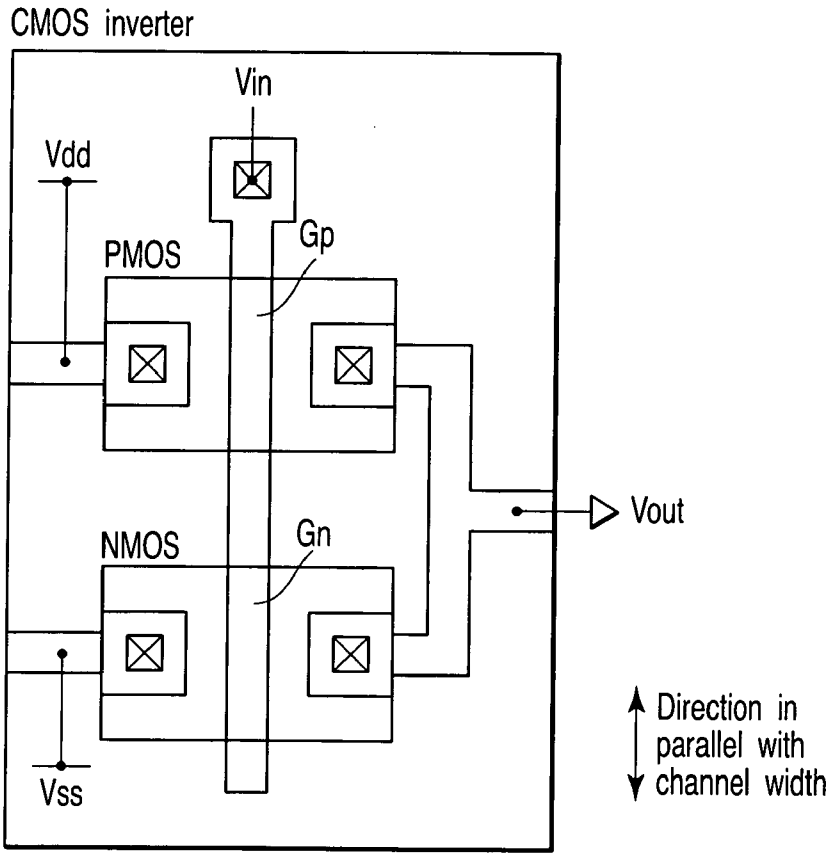

CMOS inverter
Vin
Vdd
PMOS
Gp
Vout
NMOS
Gn
Vss
Direction in parallel with channel width

FIG. 81

METHOD OF MANUFACTURING SEMICONDUCTOR DEVICE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is based upon and claims the benefit of priority from prior Japanese Patent Application No. 2005-325024, filed Nov. 9, 2005, the entire contents of which are incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] 1. Field of the Invention

[0003] The present invention relates to a method of manufacturing a metal gate electrode, in particular, a dual metal gate electrode for use in a CMOS circuit.

[0004] 2. Description of the Related Art

[0005] Conventionally, in order to realize the high performance of a MOSFET subject to miniaturization, a metal gate electrode has been proposed. When a metal gate electrode is used, depletion in a gate electrode, which brings about problems on a polysilicon gate electrode, is not brought about, so that an increase of an effective film thickness of a gate insulating film due to depletion can be prevented, which has an advantage for miniaturization.

[0006] Here, a threshold voltage of the MOSFET having the metal gate electrode is determined by a high impurity concentration in a channel region and a work function of the metal gate electrode. A work function is set to be greater than or equal to 4.8 eV in a p-channel MOSFET, and is set to be less than or equal to 4.3 eV in an n-channel MOSFET

[0007] By the way, existing semiconductor integrated circuits are generally composed of CMOS circuits. Accordingly, in order to form a gate electrode of a CMOS circuit, i.e., a dual metal gate electrode, in a wafer process, it is necessary to establish a film-forming technology for two types of metals, and to have a technology for avoiding the manufacturing processes from being complicated.

[0008] However, because there is a great risk from the standpoint of cost in that the film-forming technology for two types of metals is established, a method of manufacturing a work function modulation dual metal gate electrode in accordance with a W-In alloy process has been proposed (for example, refer to Jpn. Pat. Appln. KOKAI Publication No. 2000-315789).

[0009] There is the problem in this method that, because In is a low melting point material (melting point: 120° C.), In condenses when W and In are alloyed by heat treatment, which does not allow the alloying to progress well.

BRIEF SUMMARY OF THE INVENTION

[0010] A method of manufacturing a semiconductor device according to an aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a first metal layer on the gate insulating layer; a step of forming a second metal layer including elements for use in work function modulation on the first metal layer; a step of forming a cap layer made of a material having a melting point higher than that of the second metal layer on the second metal layer; and a step of precipitating the elements at an interface between the gate insulating layer and the first metal layer by thermal treatment.

[0011] A method of manufacturing a semiconductor device according to another aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a first metal layer on the gate insulating layer; a step of forming a second metal layer including elements for use in work function modulation on the first metal layer; a step of forming a cap layer made of a material having a melting point higher than that of the second metal layer on the second metal layer; and a step of forming an alloy layer of the first metal layer and the second metal layer by thermal treatment.

[0012] A method of manufacturing a semiconductor device according to another aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a metal layer on the gate insulating layer; a step of forming a compound layer including elements for use in work function modulation on the metal layer; and a step of precipitating the elements at an interface between the gate insulating layer and the metal layer by thermal treatment.

[0013] A method of manufacturing a semiconductor device according to another aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a metal layer on the gate insulating layer; a step of forming a compound layer including elements for use in work function modulation on the metal layer; and a step of forming an alloy layer of the metal layer and the compound layer by thermal treatment.

[0014] A method of manufacturing a semiconductor device according to another aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a metal layer on the gate insulating layer; a step of forming a semiconductor layer on the metal layer; a step of injecting elements for use in work function modulation into the semiconductor layer; and a step of precipitating the elements at an interface between the gate insulating layer and the metal layer by thermal treatment.

[0015] A method of manufacturing a semiconductor device according to another aspect of the present invention, comprises: a step of forming a gate insulating layer on a semiconductor substrate; a step of forming a metal layer on the gate insulating layer; a step of forming a semiconductor layer on the metal layer; a step of injecting elements for use in work function modulation into the semiconductor layer; and a step of forming an alloy layer by making the elements react with the metal layer by thermal treatment.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

[0016] FIG. **1** is a cross-sectional view showing a process of a method as a reference example;

[0017] FIG. **2** is a cross-sectional view showing a process of the method as the reference example;

[0018] FIG. **3** is a cross-sectional view showing a process of the method as the reference example;

[0019] FIG. **4** is a cross-sectional view showing a process of a method as a first basic configuration;

[0020] FIG. **5** is a cross-sectional view showing a process of the method as the first basic configuration;

[0021] FIG. **6** is a cross-sectional view showing a process of the method as the first basic configuration;

[0022] FIG. **7** is a cross-sectional view showing a process of a method as a second basic configuration;

[0023] FIG. **8** is a cross-sectional view showing a process of the method as the second basic configuration;

[0024] FIG. **9** is a cross-sectional view showing a process of a method as a third basic configuration;

[0025] FIG. **10** is a cross-sectional view showing a process of the method as the third basic configuration;

[0026] FIG. **11** is a cross-sectional view showing a process of the method as the third basic configuration;

[0027] FIG. **12** is a graph showing the relationship between a gate voltage Vg and a gate capacitance Cg;

[0028] FIG. **13** is a plan view showing a device structure as a first embodiment;

[0029] FIG. **14** is a cross-sectional view taken along the line XIV-XIV of FIG. **13**;

[0030] FIG. **15** is a cross-sectional view taken along the line XV-XV of FIG. **13**;

[0031] FIG. **16** is a cross-sectional view showing a process of a method as the first embodiment;

[0032] FIG. **17** is a cross-sectional view showing a process of the method as the first embodiment;

[0033] FIG. **18** is a cross-sectional view showing a process of the method as the first embodiment;

[0034] FIG. **19** is a cross-sectional view showing a process of the method as the first embodiment;

[0035] FIG. **20** is a cross-sectional view showing a process of the method as the first embodiment;

[0036] FIG. **21** is a plan view showing a device structure as a second embodiment;

[0037] FIG. **22** is a cross-sectional view taken along the line XXII-XXII of FIG. **21**;

[0038] FIG. **23** is a cross-sectional view taken along the line XXIII-XXIII of FIG. **21**;

[0039] FIG. **24** is a cross-sectional view showing a process of a method as the second embodiment;

[0040] FIG. **25** is a cross-sectional view showing a process of the method as the second embodiment;

[0041] FIG. **26** is a cross-sectional view showing a process of the method as the second embodiment;

[0042] FIG. **27** is a cross-sectional view showing a process of the method as the second embodiment;

[0043] FIG. **28** is a cross-sectional view showing a process of the method as the second embodiment;

[0044] FIG. **29** is a cross-sectional view showing a process of a method as a third embodiment;

[0045] FIG. **30** is a cross-sectional view taken along the line XXX-XXX of FIG. **29**;

[0046] FIG. **31** is a cross-sectional view taken along the line XXXI-XXXI of FIG. **29**;

[0047] FIG. **32** is a cross-sectional view showing a process of a method as the third embodiment;

[0048] FIG. **33** is a cross-sectional view showing a process of the method as the third embodiment;

[0049] FIG. **34** is a cross-sectional view showing a process of the method as the third embodiment;

[0050] FIG. **35** is a cross-sectional view showing a process of the method as the third embodiment;

[0051] FIG. **36** is a cross-sectional view showing a process of the method as the third embodiment;

[0052] FIG. **37** is a plan view showing a device structure as a fourth embodiment;

[0053] FIG. **38** is a cross-sectional view taken along the line XXXVIII-XXXVIII of FIG. **37**;

[0054] FIG. **39** is a cross-sectional view taken along the line XXXIX-XXXIX of FIG. **37**;

[0055] FIG. **40** is a cross-sectional view showing a process of a method as the fourth embodiment;

[0056] FIG. **41** is a cross-sectional view showing a process of the method as the fourth embodiment;

[0057] FIG. **42** is a cross-sectional view showing a process of the method as the fourth embodiment;

[0058] FIG. **43** is a cross-sectional view showing a process of the method as the fourth embodiment;

[0059] FIG. **44** is a cross-sectional view showing a process of the method as the fourth embodiment;

[0060] FIG. **45** is a plan view showing a device structure as a fifth embodiment;

[0061] FIG. **46** is a cross-sectional view taken along the line XLVI-XLVI of FIG. **45**;

[0062] FIG. **47** is a cross-sectional view taken along the line XLVII-XLVII of FIG. **45**;

[0063] FIG. **48** is a cross-sectional view showing a process of a method as the fifth embodiment;

[0064] FIG. **49** is a cross-sectional view showing a process of the method as the fifth embodiment;

[0065] FIG. **50** is a cross-sectional view showing a process of the method as the fifth embodiment;

[0066] FIG. **51** is a cross-sectional view showing a process of the method as the fifth embodiment;

[0067] FIG. **52** is a cross-sectional view showing a process of the method as the fifth embodiment;

[0068] FIG. **53** is a cross-sectional view showing a process of the method as the fifth embodiment;

[0069] FIG. **54** is a plan view showing a device structure as a sixth embodiment;

[0070] FIG. **55** is a cross-sectional view taken along the line LV-LV of FIG. **54**;

[0071] FIG. **56** is a cross-sectional view taken along the line LVI-LVI of FIG. **54**;

[0072] FIG. **57** is a cross-sectional view showing a process of a method as the sixth embodiment;

[0073] FIG. **58** is a cross-sectional view showing a process of the method as the sixth embodiment;

[0074] FIG. **59** is a cross-sectional view showing a process of the method as the sixth embodiment;

[0075] FIG. **60** is a cross-sectional view showing a process of the method as the sixth embodiment;

[0076] FIG. **61** is a cross-sectional view showing a process of the method as the sixth embodiment;

[0077] FIG. **62** is a cross-sectional view showing a process of the method as the sixth embodiment;

[0078] FIG. **63** is a plan view showing a device structure as a seventh embodiment;

[0079] FIG. **64** is a cross-sectional view taken along the line LXIV-LXIV of FIG. **63**;

[0080] FIG. **65** is a cross-sectional view taken along the line LXV-LXV of FIG. **63**;

[0081] FIG. **66** is a cross-sectional view showing a process of a method as the seventh embodiment;

[0082] FIG. **67** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0083] FIG. **68** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0084] FIG. **69** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0085] FIG. **70** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0086] FIG. **71** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0087] FIG. **72** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0088] FIG. **73** is a cross-sectional view showing a process of the method as the seventh embodiment;

[0089] FIG. **74** is a plan view showing a device structure as an eighth embodiment;

[0090] FIG. **75** is a cross-sectional view taken along the line LXXV-LXXV of FIG. **74**;

[0091] FIG. **76** is a cross-sectional view taken along the line LXXVI-LXXVI of FIG. **74**;

[0092] FIG. **77** is a cross-sectional view showing a process of a method as the eighth embodiment;

[0093] FIG. **78** is a cross-sectional view showing a process of the method as the eighth embodiment;

[0094] FIG. **79** is a cross-sectional view showing a process of the method as the eighth embodiment;

[0095] FIG. **80** is a plan view showing a CMOS inverter circuit as an application example; and

[0096] FIG. **81** is a plan view showing a CMOS inverter circuit as an application example.

DETAILED DESCRIPTION OF THE INVENTION

[0097] A method of manufacturing a semiconductor device according to aspects of the present invention will be described in detail with reference to the drawings.

1. BASIC CONFIGURATION

[0098] (1) Reference Example

[0099] In a CMOS circuit, work function modulation is carried out with respect to, for example, a gate electrode of an n-channel MOSFET. Namely, after metal used as a gate electrode of a p-channel MOSFET is formed, a work function of a part of the metal is modulated by alloying, and the metal is used as a gate electrode of the n-channel MOSFET.

[0100] Specifically, first, as shown in FIG. **1**, an isolation layer **101** having a shallow trench isolation (STI) structure is formed in a semiconductor substrate **100**, and a gate insulating layer **102** is formed on device regions separated by the isolation layer **101**. Then, a conductive layer **103** made of tungsten (W) whose work function is 4.9 eV is formed so as to have a thickness of about 50 nm on the gate insulating film **102** by, for example, a CVD method.

[0101] Here, in the drawing, an NMOS area denotes an area on which an N-channel MOS transistor is formed, and a PMOS area denotes an area on which a P-channel MOS transistor is formed.

[0102] Next, as shown in FIG. **2**, a conductive layer **104** made of indium (In) whose work function is 3.9 eV is formed so as to have a thickness of about 20 nm on the conductive layer **103** by, for example, sputtering. Further, the conductive layer **104** in the PMOS area is selectively removed by, for example, photo engraving process (PEP) and reactive ion etching (RIE).

[0103] Next, as shown in FIG. **3**, the conductive layers **103** and **104** are made to react by carrying out thermal treatment at about 400° C. for about one hour in the nitrogen atmosphere, for example. Namely, a W-In alloy which can be used as a gate electrode of the n-channel MOS transistor is formed by making an alloy of the tungsten and the indium in the NMOS area.

[0104] There is the problem in this method that, as shown in FIG. **3**, at the time of making an alloy by thermal treatment, indium condenses into a plurality of grains **104***a* before reacting with tungsten because a melting point of indium is low.

[0105] (2) First Basic Configuration

[0106] A first basic configuration relates to a method of preventing a low melting point material having a possibility of condensing from condensing by holding the low melting point material between materials having a higher melting point.

[0107] Specifically, first, as shown in FIG. **4**, an isolation layer **101** having an STI structure is formed in a semiconductor substrate **100**, and a gate insulating layer **102** is formed on device regions separated by the isolation layer **101**. Then, a conductive layer **103** made of tungsten (W) whose work function is 4.9 eV is formed so as to have a thickness of about 50 nm on the gate insulating film **102** by, for example, a CVD method.

[0108] Thereafter, a conductive layer **104** made of indium (In) whose work function is 3.9 eV is formed so as to have a thickness of about 20 nm on the conductive layer **103** by, for example, sputtering. Further, the conductive layer **104** in the PMOS area is selectively removed by, for example, PEP and RIE.

[0109] Further, a cap layer **105** made of tungsten (W) is formed so as to have a thickness of about 100 nm on the conductive layers **103** and **104** by, for example, a CVD method. At this point in time, the NMOS area is made to be a laminated structure of tungsten/indium/tungsten in which the conductive layer **104** is held between the conductive layer **103** and the cap layer **105**.

[0110] Next, as shown in FIG. **5**, the conductive layers **103** and **104** are made to react by carrying out thermal treatment at about 400° C. for about one hour in the nitrogen atmosphere, for example. Namely, a W-In alloy layer **106** which can be used as a gate electrode of the n-channel MOS transistor is formed by making an alloy of the tungsten and the indium in the NMOS area.

[0111] At that time, because the indium is covered with the tungsten serving as the cap layer **105**, the indium does not condense and reacts with the tungsten.

[0112] Note that the work function modulation can be carried out by not only alloying as described above, but also precipitating a material with a low work function at an interface between the gate insulating layer **102** and the conductive layer **103**.

[0113] In this case, as shown in FIG. **6**, a precipitate layer **107** is formed by precipitating the indium configuring the conductive layer **104** in the NMOS area at the interface between the gate insulating layer **102** and the conductive layer **103** via a grain boundary of the tungsten configuring the conductive layer **103** by thermal treatment in the nitrogen atmosphere.

[0114] At this time as well, because the indium is covered with the tungsten serving as the cap layer **105**, the indium does not condense, and is precipitated at the interface between the gate insulating layer **102** and the conductive layer **103**.

[0115] (3) Second Basic Configuration

[0116] A second basic configuration relates to a method of preventing a melting point material having a possibility of condensing from condensing by covering the low melting point material with material having a higher melting point.

[0117] Specifically, first, as shown in FIG. **7**, an isolation layer **101** having an STI structure is formed in a semiconductor substrate **100**, and a gate insulating layer **102** is formed on device regions separated by the isolation layer **101**.

[0118] Further, a conductive layer **104** made of indium (In) whose work function is 3.9 eV is formed so as to have a thickness of about 20 nm on the gate insulating layer **102** by, for example, sputtering. Further, the conductive layer **104** in the PMOS area is selectively removed by, for example, PEP and RIE.

[0119] Thereafter, a cap layer **105** made of tungsten (W) whose work function is 4.9 eV is formed so as to have a thickness of about 100 nm on the gate insulating layer **102** and the conductive layer **104** by, for example, a CVD method. At this point in time, the NMOS area is made to be a laminated structure of tungsten/indium/tungsten in which the conductive layer **104** is covered with the cap layer **105**.

[0120] Next, as shown in FIG. **8**, the conductive layer **104** and the cap layer **105** are made to react with each other by carrying out thermal treatment at about 400° C. for about one hour in the nitrogen atmosphere, for example. Namely, a W-In alloy layer **106** which can be used as a gate electrode of the n-channel MOS transistor is formed by making an alloy of the tungsten and the indium in the NMOS area.

[0121] At this time, because the indium is covered with the tungsten serving as the cap layer **105**, the indium does not condense and reacts with the tungsten.

[0122] Because, in the second basic configuration, the indium directly contacts the gate insulating layer **102**, the conditions thereof are set such that the indium is not peeled from the gate insulating film **102** at the time of thermal treatment.

[0123] (4) Third Basic Configuration

[0124] A third basic configuration relates to a method of preventing a metal from condensing by making a low melting point material having a possibility of condensing into a compound having a higher melting point.

[0125] Specifically, first, as shown in FIG. **9**, an isolation layer **101** having an STI structure is formed in a semiconductor substrate **100**, and a gate insulating layer **102** is formed on device regions separated by the isolation layer **101**. Further, a conductive layer **103** made of tungsten (W) whose work function is 4.9 eV is formed so as to have a thickness of about 50 nm on the gate insulating layer **102**.

[0126] In addition, a conductive layer **108** made of a compound including indium (In) whose work function is 3.9 eV, for example, a compound semiconductor (InP, InSp, or the like) is formed so as to have a thickness of about 20 nm on the conductive layer **103**. Further, the conductive layer **108** in the PMOS area is selectively removed by, for example, PEP and RIE.

[0127] Next, as shown in FIG. **10**, the conductive layers **103** and **108** are made to react with each other by carrying out thermal treatment at about 400° C. for about one hour in the nitrogen atmosphere, for example.

[0128] For example, when InP is used as a compound, the melting point of InP is about 1080° C., which is sufficiently higher than the melting point of In (120° C.). Therefore, a W-In-P alloy layer **109** which can be used as a gate electrode of the n-channel MOS transistor is formed without bringing about condensation.

[0129] For example, when InSb is used as a compound, the melting point of InSb is about 580° C., which is sufficiently higher than the melting point of In. Therefore, the W-In-P alloy layer **109** which can be used as a gate electrode of the n-channel MOS transistor is formed without bringing about condensation.

[0130] Note that, as described in the first basic configuration, work function modulation can be carried out by not only alloying, but also precipitating a material with a low work function at an interface between the gate insulating

layer **102** and the conductive layer **103**. Here, the indium in the conductive layer **108** of FIG. **9** may be precipitated.

[**0131**] In this case, as shown in FIG. **11**, a precipitate layer **107** is formed by precipitating the indium in the conductive layer **108** at the interface between the gate insulating layer **102** and the conductive layer **103** via a grain boundary of the tungsten configuring the conductive layer **103** by thermal treatment in the nitrogen atmosphere.

[**0132**] At this time as well, because the conductive layer **108** is composed of a compound having a high melting point, the conductive layer **108** does not condense and is precipitated at the interface between the gate insulating layer **102** and the conductive layer **103**. The third basic configuration may be combined with the first basic configuration. Namely, a conductive layer serving as a cap layer may be formed on a compound.

2. EXPERIMENTAL RESULTS OF WORK FUNCTION MODULATION

[**0133**] Experimental results of work function modulation will be described.

[**0134**] FIG. **12** shows the relationship between a gate capacitance Cg and a gate voltage Vg of a MOS capacitor.

[**0135**] A curved line **501** is the Cg-Vg characteristic of a MOS capacitor (sample 1) in which the gate electrode is made of tungsten (W), and a curved line **502** is the Cg-Vg characteristic of a MOS capacitor (sample 2) in which a gate electrode is composed of tungsten (W)/indium (In)/tungsten (W).

[**0136**] Note that thermal treatment at bout 400° C. for one hour have been carried out onto both of the samples **1** and **2** in the nitrogen atmosphere at the time of forming the gate electrodes.

[**0137**] A flat band voltage Vfb of the sample 1 is −0.02V, and on the basis thereof, the work function of the gate electrode is calculated to be 4.9 eV. Further, a flat band voltage Vfb of the sample 2 is −0.92V, and on the basis thereof, the work function of the gate electrode of is calculated to be 4.0 eV.

[**0138**] In this way, the main reason for that the work function of the sample 2 is made little is due to the indium being precipitated between the gate insulating layer and the tungsten layer by thermal treatment.

3. MATERIAL EXAMPLES

[**0139**] Material examples which are applied to the methods of manufacturing a metal gate electrode by the work function modulation in accordance with the examples of the present invention will be described.

[**0140**] With respect to the gate electrode of the p-channel MOS transistor, the material is selected from among metals such as W, Pd, Pt, Ni, Co, Rh, Ir, Nb, Mo, Ta, Sb, Bi, Er and Ti, alloys including at least one of the metals, and nitrides, carbides, or silicon nitrides of the metals.

[**0141**] A material for use in work function modulation is selected from among metals such as In, Ga, Tl, Sb and Bi, alloys including at least one of the metals, and the like. The material for use in work function modulation may be In and Ga such as InP, InSb, GaInSb and GaSb, or may be a III-V compound semiconductor including the both of those.

[**0142**] With respect to the cap layer used for preventing the condensation of a metal, the material is selected from among metals such as W, Pd, Pt, Ni, Co, Rh, Ir, Nb, Mo, Ta, Sb, Bi, Er and Ti, alloys including at least one of the metals, and nitrides, carbides, or silicon nitrides of the metals.

[**0143**] The cap layer can be composed of a material having a barrier function which does not contribute to the reaction with a material for use in work function modulation, such as, a titanium nitride, a tantalum nitride, a tungsten nitride, a titanium silicon nitride (TiSiN), a titanium carbide and a tungsten carbide.

[**0144**] The material as a gate electrode of the p-channel MOS transistor and the material as a cap layer may be the same, or may be different from each other.

4. EMBODIMENTS

[**0145**] Next, some embodiments as best modes will be described.

(1) First Embodiment

[**0146**] A first embodiment relates to a method of manufacturing a dual metal gate electrode based on the first basic configuration, and is an example in which work function modulation is carried out by precipitating a material with a low work function directly on a gate insulating layer.

[**0147**] A. Device Structure

[**0148**] FIG. **13** is a plan view of a CMOS circuit which is formed by the method according to the first embodiment. FIG. **14** is a cross-sectional view taken along the line XIV-XIV of FIG. **13**, and FIG. **15** is a cross-sectional view taken along the line XV-XV of FIG. **13**.

[**0149**] An isolation layer **201** having an STI structure is arranged in a semiconductor substrate **200**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[**0150**] The p-channel MOS transistor is configured by source/drain diffusion layers **208***b*, a gate insulating layer **202** on a channel between the source/drain diffusion layers **208***b*, and a gate electrode on the gate insulating layer **202**.

[**0151**] The gate electrode is configured by a ruthenium (Ru) layer **203** on the gate insulating layer **202** and a ruthenium layer **205** on the ruthenium layer **203**. Here, the ruthenium layers **203** and **205** are layers divided from the viewpoint of the process, and the both are unified as a structure.

[**0152**] A silicon nitride layer **207** serving as a mask material is arranged on the ruthenium layer **205**. Further, silicon nitride layers **209** and silicon oxide layers **210** serving as sidewalls are arranged at sidewall portions of the gate electrode.

[**0153**] Nickel silicide layers **211** are formed on the surface regions of the source/drain diffusion layers **208***b*.

[**0154**] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and an aluminum layer **214***b* are connected to the nickel silicide layers **211** and the ruthenium layer **205**

serving as a gate electrode via titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs.

[**0155**] The n-channel MOS transistor is configured by source/drain diffusion layers **208***b*, a gate insulating layer **202** on a channel between the source/drain diffusion layers **208***b*, and a gate electrode on the gate insulating layer **202**.

[**0156**] The gate electrode is composed of an indium (Ru) precipitate layer **204***b* on the gate insulating layer **202**, a ruthenium layer **203** on the indium precipitate layer **204***b*, an InSb layer **204***a* on the ruthenium layer **203**, and a ruthenium layer **205** on the InSb layer **204***a*. Here, the ruthenium layers **203** and **205** are the same as the ruthenium layers **203** and **205** configuring the gate electrode of the p-channel MOS transistor.

[**0157**] A silicon nitride layer **207** serving as a mask material is arranged on the ruthenium layer **205**. Further, silicon nitride layers **209** and silicon oxide layers **210** serving as sidewalls are arranged at sidewall portions of the gate electrode.

[**0158**] Nickel silicide layers **211** are formed on the surface regions of the source/drain diffusion layers **208***b*.

[**0159**] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and a aluminum layer **214***b* are connected to the nickel silicide layers **211** and the ruthenium layer **205** serving as a gate electrode via titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs.

[**0160**] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with the silicon oxide layers **212** and **215** serving as interlayer insulating layers.

[**0161**] B. Method of Manufacturing Semiconductor Device

[**0162**] A method of manufacturing a dual metal gate electrode by work function modulation according to the first embodiment will be described.

[**0163**] First, as shown in FIG. **16**, an isolation layer **201** having an STI structure is formed in a semiconductor substrate **200**, and a gate insulating layer **202** is formed on device regions separated by the isolation layer **201** by a thermal oxidation method. Further, a ruthenium layer **203** is formed so as to have a thickness of about 60 nm on the gate insulating layer **202** by, for example, a CVD method using an organic source.

[**0164**] Thereafter, an InSb layer **204***a* is formed so as to have a thickness of about 20 nm on the ruthenium layer **203** by, for example, sputtering. Further, the InSb layer **204***a* is selectively removed by, for example, PEP and RIE.

[**0165**] In addition, a ruthenium layer **205** is formed so as to have a thickness of about 20 nm on the ruthenium layer **203** and the InSb layer **204***a* by, for example, a CVD method. At this point in time, the NMOS area is made to be a laminated structure of Ru/InSb/Ru in which the InSb layer **204***a* is held between the ruthenium layers **203** and **205**.

[**0166**] Next, as shown in FIG. **17**, thermal treatment at about 500° C. for about thirty seconds is carried out in the nitrogen atmosphere, for example. As a result, In in the InSb layer **204***a* diffuses via a grain boundary of the ruthenium layer **203**, and is precipitated at an interface between the gate insulating layer **202** and the ruthenium layer **203** to become an indium precipitate layer **204***b*.

[**0167**] Next, as shown in FIG. **18**, a silicon nitride layer **207** is formed on the ruthenium layer **205** by, for example, a CVD method. Patterning is carried out the silicon nitride layer **207** by PEP and RIE, and moreover, the ruthenium layers **203** and **205**, the InSb layer **204***a*, the indium precipitate layer **204***b*, and the gate insulating layer **202** are etched by RIE using the silicon nitride layer **207** as a mask.

[**0168**] As a result, a metal gate electrode of a p-channel MOS transistor made of Ru is formed in the PMOS area, and a metal gate electrode of an n-channel MOS transistor composed of a lamination layer of In/Ru/InSb/Ru is formed in the NMOS area.

[**0169**] Thereafter, extension diffusion layers **208***a* are formed, with the metal gate electrodes being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[**0170**] Next, as shown in FIG. **19**, silicon nitride layers **209** and silicon oxide layers **210** are formed by, for example, a CVD method. In addition, the silicon nitride layers **209** and the silicon oxide layers **210** are etched by RIE, and those are made to remain as sidewalls at the sidewall portions of the metal gate electrodes.

[**0171**] Then, source/drain layers **208***b* are formed, with the metal gate electrodes and the sidewalls being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[**0172**] Nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **208***b* by, for example, sputtering. Then, silicon (Si) in the source/drain layers **208***b* and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[**0173**] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and thermal treatment is carried out at about 500° C. for about thirty seconds again, whereby the nickel silicide layers **211** are formed.

[**0174**] Next, as shown in FIG. **20**, a silicon oxide layer **212** covering the p-channel MOS transistor and the n-channel MOS transistor is formed by, for example, a CVD method. Further, the surface of the silicon oxide layer **212** is planarized by, for example, a chemical mechanical grinding (CMP) method.

[**0175**] Then, after contact holes reaching the source/drain diffusion layers **208***b* are formed at the silicon oxide layer **212**, titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs are formed in the contact holes.

[**0176**] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and an aluminum layer **214***b* are formed on

the silicon oxide layer **212**, and a silicon oxide layer **215** covering those electrodes is formed on the silicon oxide layer **212**.

[**0177**] The gate electrode of the p-channel MOS transistor formed by such a method is composed of Ru whose work function is 5.0 eV.

[**0178**] In addition, the gate electrode of the n-channel MOS transistor is composed of a lamination layer of In/Ru/InSb/Ru. Here, a threshold value of the n-channel MOS transistor depends on a work function of In precipitated directly on the gate insulating layer **202**, and the value becomes 3.9 eV, so that work function modulation is appropriately carried out.

[**0179**] Note that, as the gate insulating layer **202**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **202**.

[**0180**] C. Summary

[**0181**] In accordance with the first embodiment, in the method of manufacturing a dual metal gate electrode, a metal is prevented from condensing, so that a material with a low work function is precipitated directly on the gate insulating layer, and work function modulation can be exactly carried out.

(2) Second Embodiment

[**0182**] A second embodiment relates to a method of manufacturing a dual metal gate electrode based on the first basic configuration, and is an example in which work function modulation is carried out by alloying.

[**0183**] A. Device Structure FIG. **21** is a plan view of a CMOS circuit which is formed by the method according to the second embodiment. FIG. **22** is a cross-sectional view taken along the line XXII-XXII of FIG. **21**, and FIG. **23** is a cross-sectional view taken along the line XXIII-XXIII of FIG. **21**.

[**0184**] An isolation layer **201** having an STI structure is arranged in a semiconductor substrate **200**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[**0185**] The p-channel MOS transistor is configured by source/drain diffusion layers **208***b*, a gate insulating layer **202** on a channel between the source/drain diffusion layers **208***b*, and a gate electrode on the gate insulating layer **202**.

[**0186**] The gate electrode is configured by a ruthenium (Ru) layer **203** on the gate insulating layer **202** and a ruthenium layer **205** on the ruthenium layer **203**. Here, both of the ruthenium layers **203** and **205** are unified as a structure in the same way as in the first embodiment.

[**0187**] A silicon nitride layer **207** serving as a mask material is arranged on the ruthenium layer **205**. Further, silicon nitride layers **209** and silicon oxide layers **210** serving as sidewalls are arranged at sidewall portions of the gate electrode.

[**0188**] Nickel silicide layers **211** are formed on the surface regions of the source/drain diffusion layers **208***b*.

[**0189**] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and an aluminum layer **214***b* are connected to the nickel silicide layers **211** and the ruthenium layer **205** serving as a gate electrode titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs.

[**0190**] The n-channel MOS transistor is configured by source/drain diffusion layers **208***b*, a gate insulating layer **202** on a channel between the source/drain diffusion layers **208***b*, and a gate electrode on the gate insulating layer **202**.

[**0191**] The gate electrode is configured by an InRuSb layer **206** serving as an alloy layer on the gate insulating layer **202**, and a ruthenium layer **205** on the InRuSb layer **206**. Here, the ruthenium layer **205** is the same as the ruthenium layer **205** configuring the gate electrode of the p-channel MOS transistor.

[**0192**] A silicon nitride layer **207** serving as a mask material is arranged on the ruthenium layer **205**. Further, silicon nitride layers **209** and silicon oxide layers **210** serving as sidewalls are arranged at the sidewall portions of the gate electrode.

[**0193**] Nickel silicide layers **211** are formed on the surface regions of the source/drain diffusion layers **208***b*.

[**0194**] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and an aluminum layer **214***b* are connected to the nickel silicide layers **211** and the ruthenium layer **205** serving as a gate electrode via titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs.

[**0195**] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with the silicon oxide layers **212** and **215** serving as interlayer insulating layers.

[**0196**] B. Method of Manufacturing Semiconductor Device

[**0197**] A method of manufacturing a dual metal gate electrode by work function modulation according to the second embodiment will be described.

[**0198**] First, as shown in FIG. **24**, an isolation layer **201** having an STI structure is formed in a semiconductor substrate **200**, and a gate insulating layer **202** is formed by a thermal oxidation method on device regions separated by the isolation layer **201**. Further, a ruthenium layer **203** is formed so as to have a thickness of about 60 nm on the gate insulating layer **202** by, for example, a CVD method using an organic source.

[**0199**] Thereafter, an InSb layer **204***a* is formed so as to have a thickness of about 20 nm on the ruthenium layer **203** by, for example, sputtering. Further, the InSb layer **204***a* in the PMOS area is selectively removed by, for example, PEP and RIE.

[**0200**] Further, a ruthenium layer **205** is formed so as to have a thickness of about 20 nm on the ruthenium layer **203** and the InSb layer **204***a* by, for example, a CVD method using an organic source. At this point in time, the NMOS area is made to be a laminated structure of Ru/InSb/Ru in which the InSb layer **204***a* is held between the ruthenium layers **203** and **205**.

[0201] Next, as shown in FIG. **25**, thermal treatment at about 400° C. for about one hour is carried out in the nitrogen atmosphere, for example. As a result, the InSb layer **204***a* chemically reacts with the ruthenium layers **203** and **205** to become an InRuSb layer **206** serving as an alloy layer. Here, the thermal treatment is carried out under the conditions that the InRuSb layer **206** comes down to the gate insulating layer **202**.

[0202] Next, as shown in FIG. **26**, a silicon nitride layer **207** is formed on the ruthenium layer **205** by, for example, a CVD method. Patterning is carried out the silicon nitride layer **207** by PEP and RIE, and moreover, the ruthenium layers **203** and **205**, the InRuSb layer **206**, and the gate insulating layer **202** are etched with the silicon nitride layer **207** being as a mask by RIE.

[0203] As a result, a metal gate electrode of a p-channel MOS transistor made of Ru is formed in the PMOS area, and a metal gate electrode of an n-channel MOS transistor composed of a lamination layer of InRuSb/Ru is formed in the NMOS area.

[0204] Thereafter, extension diffusion layers **208***a* are formed, with the metal gate electrodes being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0205] Next, as shown in FIG. **27**, silicon nitride layers **209** and silicon oxide layers **210** are formed by, for example, a CVD method. Further, the silicon nitride layers **209** and the silicon oxide layers **210** are etched by RIE, and those are made to remain as sidewalls at the sidewall portions of the metal gate electrodes.

[0206] Thereafter, source/drain layers **208***b* are formed, with the metal gate electrodes and the sidewalls being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0207] In addition, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **208***b* by, for example, sputtering. Then, silicon (Si) in the source/drain layers **208***b* and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0208] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **211** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds.

[0209] Next, as shown in FIG. **28**, a silicon oxide layer **212** covering the p-channel MOS transistor and the n-channel MOS transistor is formed by, for example, a CVD method. Further, the surface of the silicon oxide layer **212** is planarized by, for example, a chemical mechanical grinding (CMP) method.

[0210] After contact holes coming down to the source/drain diffusion layers **208***b* are formed at the silicon oxide layer **212**, titanium/titanium nitride **213***a* serving as barrier metals and tungsten **213***b* serving as contact plugs are formed in the contact holes.

[0211] Electrodes composed of titanium nitride layers **214***a* and **214***c*, and an aluminum layer **214***b* are formed on the silicon oxide layer **212**, and a silicon oxide layer **215** covering those electrodes is formed on the silicon oxide layer **212**.

[0212] The gate electrode of the p-channel MOS transistor formed by such a method is composed of Ru whose work function is 5.0 eV.

[0213] Further, the gate electrode of the n-channel MOS transistor is composed of a laminated structure of InRuSb/Ru. Here, a threshold value of the n-channel MOS transistor depends on a work function of InRuSb/Ru directly on the gate insulating layer **202**, and the value becomes 4.1 eV, so that work function modulation can be appropriately carried out.

[0214] Note that, as the gate insulating layer **202**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used in the same way as in the first embodiment. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **202**.

[0215] C. Summary

[0216] In accordance with the second embodiment, in the method of manufacturing a dual metal gate electrode, a metal is prevented from condensing, so that an alloy layer with a low work function is formed directly on the gate insulating layer, and work function modulation can be exactly carried out.

(3) Third Embodiment

[0217] A third embodiment relates to a method of manufacturing a dual metal gate electrode based on the third basic configuration, and is an example in which work function modulation is carried out by precipitating a material with a low work function directly on a gate insulating layer.

[0218] In the third embodiment, a gate structure in which polysilicon is laminated on metal will be described.

[0219] A. Device Structure

[0220] FIG. **29** is a plan view of a CMOS circuit which is formed by the method according to the third embodiment. FIG. **30** is a cross-sectional view taken along the line XXX-XXX of FIG. **29**, and FIG. **31** is a cross-sectional view taken along the line XXXI-XXXI of FIG. **29**.

[0221] An isolation layer **301** having an STI structure is arranged in a semiconductor substrate **300**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0222] The p-channel MOS transistor is configured by source/drain diffusion layers **308***b*, a gate insulating layer **302** on a channel between the source/drain diffusion layers **308***b*, and a gate electrode on the gate insulating layer **302**.

[0223] The gate electrode is composed of a tungsten (W) layer **303** on the gate insulating layer **302**, a titanium nitride (TiN) layer **305** on the tungsten layer **303**, a polysilicon layer **307** on the titanium nitride layer **305**, and a nickel silicide layer **312** on the polysilicon layer **307**.

[0224] Silicon nitride layers **309** and silicon oxide layers **310** serving as sidewalls are arranged at the sidewall portions of the gate electrode.

[0225] Nickel silicide layers **311** are formed on the surface regions of the source/drain diffusion layers **308***b*.

[0226] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are connected to the nickel silicide layers **311** and **312** via titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs.

[0227] The n-channel MOS transistor is configured by source/drain diffusion layers **308***b*, a gate insulating layer **302** on a channel between the source/drain diffusion layers **308***b*, and a gate electrode on the gate insulating layer **302**.

[0228] The gate electrode is composed of a gallium (Ga) precipitate layer **304***b* on the gate insulating layer **302**, a tungsten layer **303** on the gallium precipitate layer **304***b*, a GaSb layer **304***a* on the tungsten layer **303**, a titanium nitride layer **305** on the GaSb layer **304***a*, a polysilicon layer **307** on the titanium nitride layer **305**, and a nickel silicide layer **312** on the polysilicon layer **307**.

[0229] Here, the tungsten layer **303**, the titanium nitride layer **305**, the polysilicon layer **307**, and the nickel silicide layer **312** are the same as the tungsten layer **303**, the titanium nitride layer **305**, the polysilicon layer **307**, and the nickel silicide layer **312** which configure the gate electrode of the p-channel MOS transistor.

[0230] Silicon nitride layers **309** and silicon oxide layers **310** serving as sidewalls are arranged at the sidewall portions of the gate electrode.

[0231] Nickel silicide layers **311** are formed on the surface regions of the source/drain diffusion layers **308***b*.

[0232] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are connected to the nickel silicide layers **311** and **312** via titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs.

[0233] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **315** and **316** serving as interlayer insulating layers.

[0234] B. Method of Manufacturing Semiconductor Device

[0235] A method of manufacturing a dual metal gate electrode by work function modulation according to the third embodiment will be described.

[0236] First, as shown in FIG. **32**, an isolation layer **301** having an STI structure is formed in a semiconductor substrate **300**, and a gate insulating layer **302** is formed by a thermal oxidation method on device regions separated by the isolation layer **301**. Further, a tungsten layer **303** is formed so as to have a thickness of about 20 nm on the gate insulating layer **302** by, for example, a CVD method using an organic source.

[0237] Thereafter, a GaSb layer **304***a* is formed so as to have a thickness of about 20 nm on the tungsten layer **303** by, for example, sputtering. Further, the GaSb layer **304***a* in the PMOS area is selectively removed by, for example, PEP and RIE.

[0238] In addition, a titanium nitride layer **305** is formed so as to have a thickness of about 20 nm on the tungsten layer **303** and the GaSb layer **304***a* by, for example, a CVD method using an organic source. At this point in time, the NMOS area is made to be a laminated structure of W/GaSb/TiN in which the GaSb layer **304***a* is held between the tungsten layer **303** and the titanium nitride layer **305**.

[0239] Next, as shown in FIG. **33**, thermal treatment at about 500° C. for about thirty seconds is carried out in the nitrogen atmosphere, for example. As a result, Ga in the GaSb layer **304***a* diffuses via a grain boundary of the tungsten layer **303**, and is precipitated at an interface between the gate insulating layer **302** and the tungsten layer **303** to become a gallium precipitate layer **304***b*.

[0240] Here, at the time of thermal treatment, the titanium nitride layer **305** functions as a cap layer which prevents the condensation of the GaSb layer **304***a* serving as a compound semiconductor.

[0241] Note that, because the GaSb layer **304***a* has a characteristic that condensation is relatively hard to bring about, the titanium nitride layer **305** can be omitted.

[0242] Next, as shown in FIG. **34**, a polysilicon layers **307** are formed so as to have a thickness of about 90 nm on the titanium nitride layers **305** by, for example, a CVD method. Next, in the PMOS area, a p-type impurity (for example, B) is ion-implanted into the polysilicon layer **307**, and in the NMOS area, an n-type impurity (for example, P) is ion-implanted into the polysilicon layer **307**.

[0243] Thereafter, silicon nitride layers **317** are formed on the polysilicon layers **307** by, for example, a CVD method. Patterning is carried out onto the silicon nitride layers **317** by PEP and RIE, and moreover, the polysilicon layer **307**, the titanium nitride layer **305**, the GaSb layer **304***a*, the tungsten layer **303**, the gallium precipitate layer **304***b*, and the gate insulating layer **302** are successively etched with the silicon nitride layers **317** being as masks by RIE.

[0244] As a result, a metal gate electrode of a p-channel. MOS transistor composed of a lamination layer of W/TiN/poly-Si, and a metal gate electrode of an n-channel MOS transistor composed of a lamination layer of Ga/W/GaSb/TiN/poly-Si is formed in the NMOS area.

[0245] Thereafter, extension diffusion layers **308***a* are formed, with the metal gate electrodes being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0246] Next, as shown in FIG. **35**, silicon nitride layers **309** and silicon oxide layers **310** are formed by, for example, a CVD method. Further, the silicon nitride layers **309** and the silicon oxide layers **310** are etched by RIE, and those are made to remain as sidewalls at the sidewall portions of the metal gate electrode.

[0247] Thereafter, source/drain layers **308***b* are formed, with the metal gate electrodes and the sidewalls being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0248] Further, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **308***b* and the polysilicon layers **307** by, for example, sputtering. Then, silicon (Si) in the source/drain layers **308***b* and the polysilicon layers **307**, and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0249] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **311** and **312** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds.

[0250] Next, as shown in FIG. **36**, a silicon oxide layer **315** covering the p-channel MOS transistor and the n-channel MOS transistor is formed by, for example, a CVD method. Further, the surface of the silicon oxide layer **315** is planarized by, for example, a chemical mechanical grinding (CMP) method.

[0251] After contact holes coming down to the source/drain diffusion layers **308***b* are formed at the silicon oxide layer **315**, titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs are formed in the contact holes.

[0252] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are formed on the silicon oxide layer **315**, and a silicon oxide layer **316** covering those electrodes is formed on the silicon oxide layer **315**.

[0253] The gate electrode of the p-channel MOS transistor formed by such a method is composed of a lamination layer of W/TiN/poly-Si/Ni-silicide, and W whose work function is 4.9 eV is arranged directly on the gate insulating layer **302**.

[0254] In addition, the gate electrode of the n-channel MOS transistor is composed of a laminated structure of Ga/W/GaSb/TiN/poly-Si/Ni-silicide. Here, a threshold value of the n-channel MOS transistor depends on a work function of Ga precipitated directly on the gate insulating layer **202**, and the value becomes 4.2 eV, so that work function modulation can be appropriately carried out.

[0255] Note that, as the gate insulating layer **302**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **302**.

[0256] C. Summary

[0257] In accordance with the third embodiment, in the method of manufacturing a dual metal gate electrode, the condensation of metal is prevented, so that a material with a low work function is precipitated directly on the gate insulating layer, and work function modulation can be exactly carried out.

(4) Fourth Embodiment

[0258] A fourth embodiment relates to a method of manufacturing a dual metal gate electrode based on the third basic configuration, and is an example in which work function modulation is carried out by alloying.

[0259] In the fourth embodiment, a gate structure in which polysilicon is laminated on metal will be described.

[0260] A. Device Structure

[0261] FIG. **37** is a plan view of a CMOS circuit which formed by the method according to the fourth embodiment. FIG. **38** is a cross-sectional view taken along the line XXXVIII-XXXVIII of FIG. **37**, and FIG. **39** is a cross-sectional view taken along the line XXXIX-XXXIX of FIG. **37**.

[0262] An isolation layer **301** having an STI structure is arranged in a semiconductor substrate **300**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0263] The p-channel MOS transistor is configured by source/drain diffusion layers **308***b*, a gate insulating layer **302** on a channel between the source/drain diffusion layers **308***b*, and a gate electrode on the gate insulating layer **302**.

[0264] The gate electrode is configured by a tungsten (W) layer **303** on the gate insulating layer **302**, a titanium nitride (TiN) layer **305** on the tungsten layer **303**, a polysilicon layer **307** on the titanium nitride layer **305**, and a nickel silicide layer **312** on the polysilicon layer **307**.

[0265] Silicon nitride layers **309** and silicon oxide layers **310** serving as sidewalls are arranged at the sidewall portions of the gate electrode.

[0266] Nickel silicide layers **311** are formed on the surface regions of the source/drain diffusion layers **308***b*.

[0267] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are connected to the nickel silicide layers **311** and **312** via titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs.

[0268] The n-channel MOS transistor is configured by source/drain diffusion layers **308***b*, a gate insulating layer **302** on the channel between the source/drain diffusion layers **308***b*, and a gate electrode on the gate insulating layer **302**.

[0269] The gate electrode is composed of a GaWSb layer **306** serving as an alloying layer on the gate insulating layer **302**, a titanium nitride layer **305** on the GaWSb layer **306**, a polysilicon layer **307** on the titanium nitride layer **305**, and a nickel silicide layer **312** on the polysilicon layer **307**.

[0270] Here, the titanium nitride layer **305**, the polysilicon layer **307**, and the nickel silicide layer **312** are the same as the titanium nitride layer **305**, the polysilicon layer **307**, and the nickel silicide layer **312** which configure the gate electrode of the p-channel MOS transistor.

[0271] Silicon nitride layers **309** and silicon oxide layers **310** serving as sidewalls are arranged at the sidewall portions of the gate electrode.

[0272] Nickel silicide layers **311** are formed on the surface regions of the source/drain diffusion layers **308***b*.

[0273] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are connected to the nickel silicide layers **311** and **312** via titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs.

[0274] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **315** and **316** serving as interlayer insulating layers.

[0275] B. Method of Manufacturing Semiconductor Device

[0276] A method of manufacturing a dual metal gate electrode by work function modulation according to the fourth embodiment will be described.

[0277] First, as shown in FIG. **40**, an isolation layer **301** having an STI structure is formed in a semiconductor substrate **300**, and a gate insulating layer **302** is formed by a thermal oxidation method on device regions separated by the isolation layer **301**. Further, a tungsten layer **303** is formed so as to have a thickness of about 20 nm on the gate insulating layer **302** by, for example, a CVD method using an organic source.

[0278] Thereafter, a GaSb layer **304***a* is formed so as to have a thickness of about 20 nm on the tungsten layer **303** by, for example, sputtering. Further, the GaSb layer **304***a* in the PMOS area is selectively removed by, for example, PEP and RIE.

[0279] In addition, a titanium nitride layer **305** is formed so as to have a thickness of about 20 nm on the tungsten layer **303** and the GaSb layer **304***a* by, for example, a CVD method using an organic source. At this point in time, the NMOS area is made to be a laminated structure of W/GaSb/TiN in which the GaSb layer **304***a* is held between the tungsten layer **303** and the titanium nitride layer **305**.

[0280] Next, as shown in FIG. **41**, thermal treatment at about 400° C. for about one hour is carried out in the nitrogen atmosphere, for example. As a result, the tungsten layer **303** and the GaSb layer **304***a* chemically react with each other to become a GaWSb layer **306** serving as an alloy layer. Here, the thermal treatment is carried out under the conditions that the GaWSb layer **306** comes down to the gate insulating layer **302**.

[0281] Note that, at the time of thermal treatment, the titanium nitride layer **305** functions as a cap layer preventing the GaSb layer **304***a* serving as a compound semiconductor from condensing. Further, because the GaSb layer **304***a* has a characteristic that condensation is relatively hard to bring about, the titanium nitride layer **305** can be omitted.

[0282] Next, as shown in FIG. **42**, a polysilicon layer **307** is formed so as to have a thickness of about 90 nm on the titanium nitride layer **305** by, for example, a CVD method. Next, in the PMOS area, a p-type impurity (for example, B) is ion-implanted into the polysilicon layer **307**, and in the NMOS area, an n-type impurity (for example, P) is ion-implanted into the polysilicon layer **307**.

[0283] Thereafter, a silicon nitride layer **317** is formed on the polysilicon layer **307** by, for example, a CVD method. Patterning is carried out onto the silicon nitride layer **317** by PEP and RIE, and moreover, the polysilicon layer **307**, the titanium nitride layer **305**, the GaWSb layer **306**, the tungsten layer **303**, and the gate insulating layer **302** are successively etched with the silicon nitride layer **317** being as a mask by RIE.

[0284] As a result, a metal gate electrode of a p-channel MOS transistor composed of a lamination layer of W/TiN/poly-Si in the PMOS area, and a metal gate electrode of an n-channel MOS transistor composed of a lamination layer of GaWSb/TiN/poly-Si is formed in the NMOS area.

[0285] Thereafter, extension diffusion layers **308***a* are formed, with the metal gate electrodes being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0286] Next, as shown in FIG. **43**, silicon nitride layers **309** and silicon oxide layers **310** are formed by, for example, a CVD method. Further, the silicon nitride layers **309** and the silicon oxide layers **310** are etched by RIE, and those are made to remain as sidewalls at the sidewall portions of the metal gate electrode.

[0287] Thereafter, source/drain layers **308***b* are formed, with the metal gate electrodes and the sidewalls being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0288] Further, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **308***b* and the polysilicon layers **307** by, for example, sputtering. Then, silicon (Si) in the source/drain layers **308***b* and the polysilicon layers **307**, and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0289] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **311** and **312** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds.

[0290] Next, as shown in FIG. **44**, a silicon oxide layer **315** covering the p-channel MOS transistor and the n-channel MOS transistor is formed by, for example, a CVD method. Further, the surface of the silicon oxide layer **315** is planarized by, for example, a CMP method.

[0291] After contact holes coming down to the source/drain diffusion layers **308***b* are formed at the silicon oxide layer **315**, titanium/titanium nitride **313***a* serving as barrier metals and tungsten **313***b* serving as contact plugs are formed in the contact holes.

[0292] Electrodes composed of titanium nitride layers **314***a* and **314***c*, and an aluminum layer **314***b* are formed on the silicon oxide layer **315**, and a silicon oxide layer **316** covering those electrodes is formed on the silicon oxide layer **315**.

[0293] The gate electrode of the p-channel MOS transistor formed by such a method is composed of a lamination layer of W/TiN/poly-Si/Ni-silicide, and W whose work function is 4.9 eV is arranged directly on the gate insulating layer **302**.

[0294] In addition, the gate electrode of the n-channel MOS transistor is composed of a laminated structure of GaWSb/TiN/poly-Si/Ni-silicide. Here, a threshold value of the n-channel MOS transistor depends on a work function of GaWSb directly on the gate insulating layer **202**, and the value becomes 4.1 eV, so that work function modulation can be appropriately carried out.

[0295] Note that, as the gate insulating layer **302**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used in the same way as in the third embodiment. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **302**.

[0296] C. Summary

[0297] In accordance with the fourth embodiment, in the method of manufacturing a dual metal gate electrode, the condensation of metal is prevented, so that an alloy layer with a low work function is formed directly on the gate insulating layer, and work function modulation can be exactly carried out.

(5) Fifth Embodiment

[0298] A fifth embodiment relates to an application example of the first to third basic configurations, and is an example in which elements contributing to work function modulation are injected into a gate electrode by ion implantation, and the elements are precipitated directly on the gate insulating layer by thermal treatment.

[0299] A. Device Structure

[0300] FIG. **45** is a plan view of a CMOS circuit which is formed by the method according to the fifth embodiment. FIG. **46** is a cross-sectional view taken along the line XLVI-XLVI of FIG. **45**, and FIG. **47** is a cross-sectional view taken along the line XLVII-XLVII of FIG. **45**.

[0301] An isolation layer **401** having an STI structure is arranged in a semiconductor substrate **400**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0302] The p-channel MOS transistor is configured by source/drain diffusion layers **408***b*, a gate insulating layer **402** on a channel between the source/drain diffusion layers **408***b*, and a gate electrode on the gate insulating layer **402**.

[0303] The gate electrode is configured by a tungsten nitride (WN) layer **403** on the gate insulating layer **402**, a polysilicon layer **405** on the tungsten nitride layer **403**, and a nickel silicide layer **412** on the polysilicon layer **405**. The polysilicon layer **405** is a conductive material including a p-type impurity (for example, B, BF_2, or the like).

[0304] Silicon nitride layers **417**, silicon oxide layers **409**, and silicon nitride layers **410** serving as sidewalls are arranged at the sidewall portions of the gate electrode. Moreover, silicon nitride layers **418** are arranged so as to cover the silicon nitride layers **410**.

[0305] Nickel silicide layers **411** are formed on the surface regions of the source/drain diffusion layers **408***b*.

[0306] Electrodes composed of a tantalum nitride (TaN) layer **414***a* and a copper (Cu) layer **414***b* are connected to nickel silicide layers **411** and **412** via titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs.

[0307] The n-channel MOS transistor is configured by source/drain diffusion layers **408***b*, a gate insulating layer **402** on a channel between the source/drain diffusion layers **408***b*, and a gate electrode on the gate insulating layer **402**.

[0308] The gate electrode is configured by an indium (In) precipitate layer **404***b* on the gate insulating layer **402**, a tungsten nitride (WN) layer **403** on the indium precipitate layer **404***b*, a polysilicon layer **405** on the tungsten nitride layer **403**, and a nickel silicide layer **412** on the polysilicon layer **405**. The polysilicon layer **405** is a conductive material including an n-type impurity (for example, P, As, or the like).

[0309] Here, the tungsten nitride layer **403**, the polysilicon layer **405**, and the nickel silicide layer **412** are the same as the tungsten nitride layer **403**, the polysilicon layer **405**, and the nickel silicide layer **412** which configure the gate electrode of the p-channel MOS transistor.

[0310] Silicon nitride layers **417**, silicon oxide layers **409**, and silicon nitride layers **410** serving as sidewalls are arranged at the sidewall portions of the gate electrode. Moreover, silicon nitride layers **418** are arranged so as to cover the silicon nitride layers **410**.

[0311] Nickel silicide layers **411** are formed on the surface regions of the source/drain diffusion layers **408***b*.

[0312] Electrodes composed of a tantalum nitride (TaN) layer **414***a* and a copper (Cu) layer **414***b* are connected to nickel silicide layers **411** and **412** via titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs.

[0313] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **415***a*, **415***b*, and **416** serving as interlayer insulating layers.

[0314] B. Method of Manufacturing Semiconductor Device

[0315] A method of manufacturing a dual metal gate electrode by work function modulation according to the fifth embodiment will be described.

[0316] First, as shown in FIG. **48**, an isolation layer **401** having an STI structure is formed in a semiconductor substrate **400**, and a gate insulating layer **402** is formed by a thermal oxidation method on device regions separated by the isolation layer **401**.

[0317] Then, a tungsten nitride layer **403** whose work function is 4.9 eV is formed so as to have a thickness of about 20 nm on the gate insulating layer **402** by, for example, a CVD method using an organic source. Subsequently, a polysilicon layer **405** is formed so as to have a thickness of about 100 nm on the tungsten nitride layer **403** by, for example, a CVD method.

[0318] Thereafter, indium (In) **414***a* is ion-implanted into the polysilicon layer **405** in the NMOS area under the conditions that an accelerating energy is about 40 keV, and a dose amount is about 5×10^{15} cm^{-2}.

[0319] Next, as shown in FIG. **49**, etching is successively carried out onto the polysilicon layer **405**, the tungsten nitride layer **403**, and the gate insulating layer **402** by PEP and RIE. in the PMOS area, a gate electrode of a p-channel MOS transistor whose gate width (channel length) is, for example, 30 nm is formed, and in the NMOS area, a gate electrode of a p-channel MOS transistor whose gate width is, for example, 20 nm is formed.

[0320] Then, silicon nitride layers **417** covering the p-channel MOS transistor and the n-channel MOS transistor are formed so as to have a thickness of about 8 nm, and etch back is carried out thereon, which makes the silicon nitride layers **417** remain at the sidewall portions of the gate electrode.

[0321] Further, with the gate electrodes being as masks, a p-type impurity (for example, B, BF_2, or the like) is ion-implanted into the PMOS area in a self-alignment manner, and an n-type impurity (for example, P, As, or the like) is ion-implanted into the NMOS area in a self-alignment manner, so that extension diffusion layers **408***a* are formed by carrying out thermal treatment at about 800° C. for about five seconds.

[0322] Next, as shown in FIG. **50**, silicon oxide layers **409** and silicon nitride layers **410** are formed by, for example, a CVD method, and etch back is carried out thereon, which makes the silicon oxide layers **409** and the silicon nitride layers **410** remain as sidewalls at the sidewall portions of the gate electrodes.

[0323] Further, with the gate electrodes and the side walls being as masks, a p-type impurity (for example, B, BF_2, or the like) is ion-implanted into the PMOS area in a self-alignment manner, and an n-type impurity (for example, P, As, or the like) is ion-implanted into the NMOS area in a self-alignment manner, so that source/drain layers **408***b* are formed by carrying out thermal treatment at about 1000° C. for about one second.

[0324] Thereafter, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **408***b* and the polysilicon layers **405** by, for example, sputtering. Then, silicon (Si) in the source/drain layers **408***b* and the polysilicon layers **405** and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0325] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **411** and **412** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds again.

[0326] Next, as shown in FIG. **51**, silicon nitride layers **418** covering the p-channel MOS transistor and the n-channel MOS transistor are formed so as to have a thickness of about 30 nm by, for example, a CVD method. Subsequently, a silicon oxide layer **415***a* serving as an interlayer insulating layer is formed so as to have a thickness of about 250 nm on the silicon nitride layers **418** by, for example, a CVD method.

[0327] Then, the silicon nitride layers **418** and the silicon oxide layer **415***a* are planarized by grinding until the top surfaces of the gate electrodes are exposed. At that time, there are cases in which the nickel silicide layers **412** which are the uppermost layers of the gate electrodes are grinded, and some or the entire thereof are removed.

[0328] In such a case, a nickel (Ni) layer **419** is formed again so as to have a thickness of about 10 nm on the planarized silicon nitride layers **418** and silicon oxide layers **415***a* by, for example, sputtering. In addition, silicon (Si) in the polysilicon layers **405** and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0329] Further, as shown in FIG. **52**, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and then, the nickel silicide layers **412** are formed again by carrying out thermal treatment at about 500° C. for about thirty seconds.

[0330] By the way, from the time ion implantation of indium has been carried out into the polysilicon layer **405** up to the present stage, as described above, a plurality of thermal treatment steps are executed. Due to those thermal treatments, indium is precipitated at an interface between the gate insulating layer **402** and the tungsten nitride layer **403** via a grain boundary of the tungsten nitride layer **403** to become an indium precipitate layer **404***b*.

[0331] The work function of the gate electrode of the n-channel MOS transistor becomes 4.1 eV due to the indium precipitate layer **404***b*.

[0332] Note that the nickel layer **419** of FIG. **51** functions so as to prevent the indium in the polysilicon layers **405** from diffusing outward at the time of thermal treatment.

[0333] Next, as shown in FIG. **53**, a silicon oxide layer **415***b* is formed on the silicon oxide layer **415***a* by, for example, a CVD method. Further, the surface of the silicon oxide layer **415***b* is planarized by, for example, a CMP method.

[0334] After contact holes coming down to the source/drain diffusion layers **408***b* are formed at the silicon oxide layers **415***a* and **415***b* and the silicon nitride layers **418**, titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs are formed in the contact holes.

[0335] Further, a silicon oxide layer **416** is formed on the silicon oxide layer **415***b* by, for example, a CVD method. Further, the surface of the silicon oxide layer **416** is planarized by, for example, a CMP method.

[0336] After wiring grooves coming down to the tungsten **413***b* serving as contact plugs are formed at the silicon oxide layer **416**, electrodes composed of tantalum nitride layers **414***a* and copper layers **414***b* are embedded into the wiring grooves.

[0337] A work function of the gate electrode of the p-channel MOS transistor formed by such a method is determined by WN, and the value becomes 4.9 eV. In contrast thereto, a work function of the gate electrode of the n-channel MOS transistor is determined by In, and the value becomes 3.9 eV.

[0338] Note that, with respect to the gate electrode of the p-channel MOS transistor, the material can be selected from among metals such as W, Pd, Pt, Ni, Co, Rb, Ir, Nb, Mo, Ta,

Sb, Bi, Er and Ti, an alloy including at least one of those metals, and a nitride, a carbide, or a silicon nitride of those metals.

[0339] A material for use in work function modulation, here, elements injected into the polysilicon layers **405** by ion implantation can be selected from among In, Ga, Tl, Sb, Bi and the like.

[0340] As the gate electrode layer **202**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used.

[0341] For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **202**.

[0342] In the fifth embodiment, the elements injected into the polysilicon layers **405** are precipitated directly on the gate insulating layers **402**. However, as in the second and fourth embodiments described above, an alloy layer is formed by using the elements injected into the polysilicon layers **405**, and work function modulation may be carried out by the alloy layer.

[0343] In addition, the elements for use in work function modulation are injected into the polysilicon layers by ion implantation. However, those may be injected by another method. Although the layers into which the elements for use in work function modulation are injected are not limited to the polysilicon layers, and are preferably semiconductor layers.

[0344] C. Summary

[0345] In accordance with the fifth embodiment, in the method of manufacturing a dual metal gate electrode, a material with a low work function is precipitated directly on the gate insulating layer due to ion implantation and thermal treatment, so that work function modulation can be exactly carried out.

(6) Sixth Embodiment

[0346] A sixth embodiment relates to an application example of the first to third basic configurations, and is an example in which elements contributing to work function modulation are injected into a gate electrode by ion implantation, and the elements are precipitated directly on the gate insulating layer by thermal treatment.

[0347] A. Device Structure

[0348] FIG. **54** is a plan view of a CMOS circuit which is formed by the method according to the sixth embodiment. FIG. **55** is a cross-sectional view taken along the line LV-LV of FIG. **54**, and FIG. **56** is a cross-sectional view taken along the line LVI-LVI of FIG. **54**.

[0349] An isolation layer **401** having an STI structure is arranged in a semiconductor substrate **400**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0350] The p-channel MOS transistor is configured by source/drain diffusion layers **408***b*, a gate insulating layer **402** on a channel between the source/drain diffusion layers **408***b*, and a gate electrode on the gate insulating layer **402**.

[0351] The gate electrode is configured by a tungsten nitride (WN) layer **403** on the gate insulating layer **402** and a nickel silicide layer **412** on the tungsten nitride layer **403**.

[0352] Silicon nitride layers **417**, silicon oxide layers **409**, and silicon nitride layers **410** serving as sidewalls are arranged at the sidewall portions of the gate electrode. Moreover, silicon nitride layers **418** are arranged so as to cover the silicon nitride layers **410**.

[0353] Nickel silicide layers **411** are formed on the surface regions of the source/drain diffusion layers **408***b*.

[0354] Electrodes composed of a tantalum nitride (TaN) layer **414***a* and a copper (Cu) layer **414***b* are connected to nickel silicide layers **411** and **412** via titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs.

[0355] The n-channel MOS transistor is configured by source/drain diffusion layers **408***b*, a gate insulating layer **402** on a channel between the source/drain diffusion layers **408***b*, and a gate electrode on the gate insulating layer **402**.

[0356] The gate electrode is configured by an indium (In) precipitate layer **404***b* on the gate insulating layer **402**, a tungsten nitride (WN) layer **403** on the indium precipitate layer **404***b*, and a nickel silicide layer **412** on the tungsten nitride layer **403**.

[0357] Here, the tungsten nitride layer **403** and the nickel silicide layer **412** are the same as the tungsten nitride layer **403** and the nickel silicide layer **412** which configure the gate electrode of the p-channel MOS transistor.

[0358] Silicon nitride layers **417**, silicon oxide layers **409**, and silicon nitride layers **410** serving as sidewalls are arranged at the sidewall portions of the gate electrode. Moreover, silicon nitride layers **418** are arranged so as to cover the silicon nitride layers **410**.

[0359] Nickel silicide layers **411** are formed on the surface regions of the source/drain diffusion layers **408***b*.

[0360] Electrodes composed of a tantalum nitride (TaN) layer **414***a* and a copper (Cu) layer **414***b* are connected to nickel silicide layers **411** and **412** via titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs.

[0361] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **415***a*, **415***b*, and **416** serving as interlayer insulating layers.

[0362] B. Method of Manufacturing Semiconductor Device

[0363] A method of manufacturing a dual metal gate electrode by work function modulation according to the sixth embodiment will be described.

[0364] First, as shown in FIG. **57**, an isolation layer **401** having an STI structure is formed in a semiconductor substrate **400**, and a gate insulating layer **402** is formed by a thermal oxidation method on device regions separated by the isolation layer **401**.

[0365] Then, a tungsten nitride layer **403** whose work function is 4.9 eV is formed so as to have a thickness of about 20 nm on the gate insulating layer **402** by, for example, a CVD method using an organic source. Subsequently, a polysilicon layer **405** is formed so as to have a thickness of about 100 nm on the tungsten nitride layer **403** by, for example, a CVD method.

[0366] Thereafter, indium (In) **414***a* is ion-implanted into the polysilicon layer **405** in the NMOS area under the conditions that an accelerating energy is about 40 keV, and a dose amount is about 5×10^{15} cm^{-2}.

[0367] Next, as shown in FIG. **58**, etching is successively carried out onto the polysilicon layer **405**, the tungsten nitride layer **403**, and the gate insulating layer **402** by PEP and RIE. in the PMOS area, a gate electrode of a p-channel MOS transistor whose gate width (channel length) is, for example, 30 nm is formed, and in the NMOS area, a gate electrode of a p-channel MOS transistor whose gate width is, for example, 20 nm is formed.

[0368] Then, silicon nitride layers **417** covering the p-channel MOS transistor and the n-channel MOS transistor are formed so as to have a thickness of about 8 nm, and etch back is carried out thereon, which makes the silicon nitride layers **417** remain at the sidewall portions of the gate electrode.

[0369] Further, with the gate electrodes being as masks, a p-type impurity (for example, B, BF_2, or the like) is ion-implanted into the PMOS area in a self-alignment manner, and an n-type impurity (for example, P, As, or the like) is ion-implanted into the NMOS area in a self-alignment manner, so that extension diffusion layers **408***a* are formed by carrying out thermal treatment at about 800° C. for about five seconds.

[0370] Next, as shown in FIG. **59**, silicon oxide layers **409** and silicon nitride layers **410** are formed by, for example, a CVD method, and etch back is carried out thereon, which makes the silicon oxide layers **409** and the silicon nitride layers **410** remain as sidewalls at the sidewall portions of the gate electrodes.

[0371] Further, with the gate electrodes and the side walls being as masks, a p-type impurity (for example, B, BF_2, or the like) is ion-implanted into the PMOS area in a self-alignment manner, and an n-type impurity (for example, P, As, or the like) is ion-implanted into the NMOS area in a self-alignment manner, so that source/drain layers **408***b* are formed by carrying out thermal treatment at about 1000° C. for about one second.

[0372] Thereafter, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain layers **408***b* and the polysilicon layers **405** (see FIG. **58**) by, for example, sputtering. Then, silicon (Si) in the source/drain layers **408***b* and the polysilicon layers **405** (see FIG. **58**) and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0373] Thereafter, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **411** and **412** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds again.

[0374] Next, as shown in FIG. **60**, silicon nitride layers **418** covering the p-channel MOS transistor and the n-channel MOS transistor are formed so as to have a thickness of about 30 nm by, for example, a CVD method. Subsequently, a silicon oxide layer **415***a* serving as an interlayer insulating layer is formed so as to have a thickness of about 250 nm on the silicon nitride layers **418** by, for example, a CVD method.

[0375] Then, the silicon nitride layers **418** and the silicon oxide layer **415***a* are planarized by grinding until the top surfaces of the gate electrodes are exposed. At that time, there are cases in which the nickel silicide layers **412** which are the uppermost layers of the gate electrodes are grinded, and some or the entire thereof are removed.

[0376] In such a case, a nickel (Ni) layer **419** is formed again so as to have a thickness of about 40 nm on the planarized silicon nitride layers **418** and silicon oxide layers **415***a* by, for example, sputtering. In addition, a thermal treatment at about 350° C. for about thirty seconds is executed.

[0377] Further, as shown in FIG. **61**, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and then, the nickel silicide layers **412** are formed again by carrying out thermal treatment at about 500° C. for about thirty seconds.

[0378] By the way, from the time ion implantation of indium has been carried out into the polysilicon layer **405** (see FIG. **58**) up to the present stage, as described above, a plurality of thermal treatment steps are executed.

[0379] Due to those thermal treatments, by a snowplow effect occurred during silicide reaction, indium is segregated at the interface between silicide and gate insulator. That is, indium is precipitated at an interface between the gate insulating layer **402** and the tungsten nitride layer **403** via a grain boundary of the tungsten nitride layer **403** to become an indium precipitate layer **404***b*.

[0380] The work function of the gate electrode of the n-channel MOS transistor becomes 4.1 eV due to the indium precipitate layer **404***b*.

[0381] Note that the nickel layer **419** of FIG. **60** functions so as to prevent the indium in the polysilicon layers **405** (see FIG. **58**) from diffusing outward at the time of thermal treatment.

[0382] Next, as shown in FIG. **62**, a silicon oxide layer **415***b* is formed on the silicon oxide layer **415***a* by, for example, a CVD method. Further, the surface of the silicon oxide layer **415***b* is planarized by, for example, a CMP method.

[0383] After contact holes coming down to the source/drain diffusion layers **408***b* are formed at the silicon oxide layers **415***a* and **415***b* and the silicon nitride layers **418**, titanium/titanium nitride **413***a* serving as barrier metals and tungsten **413***b* serving as contact plugs are formed in the contact holes.

[0384] Further, a silicon oxide layer **416** is formed on the silicon oxide layer **415***b* by, for example, a CVD method. Further, the surface of the silicon oxide layer **416** is planarized by, for example, a CMP method.

[0385] After wiring grooves coming down to the tungsten **413***b* serving as contact plugs are formed at the silicon oxide layer **416**, electrodes composed of tantalum nitride layers **414***a* and copper layers **414***b* are embedded into the wiring grooves.

[0386] A work function of the gate electrode of the p-channel MOS transistor formed by such a method is determined by WN, and the value becomes 4.9 eV. In contrast thereto, a work function of the gate electrode of the n-channel MOS transistor is determined by In, and the value becomes 3.9 eV.

[0387] Note that, with respect to the gate electrode of the p-channel MOS transistor, the material can be selected from among metals such as W, Pd, Pt, Ni, Co, Rb, Ir, Nb, Mo, Ta, Sb, Bi, Er and Ti, an alloy including at least one of those metals, and a nitride, a carbide, or a silicon nitride of those metals.

[0388] A material for use in work function modulation, here, elements injected into the polysilicon layers **405** by ion implantation can be selected from among In, Ga, Tl, Sb, Bi and the like.

[0389] As the gate electrode layer **202**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used.

[0390] For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **202**.

[0391] In the sixth embodiment, the elements injected into the polysilicon layers **405** are precipitated directly on the gate insulating layers **402**. However, as in the second and fourth embodiments described above, an alloy layer is formed by using the elements injected into the polysilicon layers **405**, and work function modulation may be carried out by the alloy layer.

[0392] In addition, the elements for use in work function modulation are injected into the polysilicon layers by ion implantation. However, those may be injected by another method. Although the layers into which the elements for use in work function modulation are injected are not limited to the polysilicon layers, and are preferably semiconductor layers.

[0393] C. Summary

[0394] In accordance with the sixth embodiment, in the method of manufacturing a dual metal gate electrode, a material with a low work function is precipitated directly on the gate insulating layer due to ion implantation and thermal treatment, so that work function modulation can be exactly carried out.

(7) Seventh Embodiment

[0395] A seventh embodiment relates to an application example of the first embodiment, and is an example in which the method of manufacturing a dual metal gate electrode according to the present invention is applied to a CMOS circuit having a damascene gate structure.

[0396] A. Device Structure

[0397] FIG. **63** is a plan view of a CMOS circuit which is formed by the method according to the seventh embodiment. FIG. **64** is a cross-sectional view taken along the line LXIV-LXIV of FIG. **63**, and FIG. **65** is a cross-sectional view taken along the line LXV-LXV of FIG. **63**.

[0398] An isolation layer **501** having an STI structure is arranged in a semiconductor substrate **500**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0399] The p-channel MOS transistor is configured by source/drain diffusion layers **505***b*, a gate insulating layer **510** on a channel between the source/drain diffusion layers **505***b*, and a gate electrode on the gate insulating layer **510**. The gate electrode is configured by a molybdenum (Mo) layer **511** on the gate insulating layer **510**.

[0400] Silicon oxide layers **506** and silicon nitride layers **507** serving as sidewalls are arranged at the sidewall portions of the molybdenum layer **511**.

[0401] Nickel silicide layers **508** are formed on the surface regions of the source/drain diffusion layers **505***b*.

[0402] Electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are connected to the nickel silicide layers **508** and the molybdenum layer **511** serving as a gate electrode via titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs.

[0403] The n-channel MOS transistor is configured by source/drain diffusion layers **505***b*, a gate insulating layer **510** on a channel between the source/drain diffusion layers **505***b*, and a gate electrode on the gate insulating layer **510**.

[0404] The gate electrode is configured by a gallium (Ga) precipitate layer **512***b* on the gate insulating layer **510**, and a molybdenum layer **511** on the gallium precipitate layer **512***b*. Here, the molybdenum layer **511** is the same as the molybdenum layer **511** configuring the gate electrode of the p-channel MOS transistor.

[0405] Silicon oxide layers **506** and silicon nitride layers **507** serving as sidewalls are arranged at the sidewall portions of the molybdenum layer **511**.

[0406] Nickel silicide layers **508** are formed on the surface regions of the source/drain diffusion layers **505***b*.

[0407] Electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are connected to the nickel silicide layers **508** and the molybdenum layer **511** serving as a gate electrode via titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs.

[0408] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **509**, **515**, and **518** serving as interlayer insulating layers.

[0409] B. Method of Manufacturing Semiconductor Device

[0410] A method of manufacturing a dual metal gate electrode by work function modulation according to the seventh embodiment will be described.

[0411] First, as shown in FIG. **66**, an isolation layer **501** having an STI structure is formed in a semiconductor substrate **500**, and silicon oxide layers **502** are formed by a thermal oxidation method on device regions separated by the isolation layer **501**.

[0412] Then, a polysilicon layer **503** is formed so as to have a thickness of about 100 nm on the silicon oxide layers **502** by, for example, a CVD method. Next, a silicon nitride layer **504** is formed so as to have a thickness of about 50 nm on the polysilicon layer **503**.

[0413] Next, as shown in FIG. **67**, dummy gate electrodes are formed by patterning onto the silicon nitride layer **504** and the polysilicon layer **503** by PEP and RIE.

[0414] In addition, with the dummy gate electrodes being as masks, a p-type impurity (for example, B, BF_2, or the like) is ion-implanted into the PMOS area in a self-alignment manner, an n-type impurity (for example, P, As, or the like) is ion-implanted into the NMOS area in a self-alignment manner, and thermal treatment at about 800° C. for about five seconds is carried out, so that extension diffusion layers **505***a* are formed.

[0415] Next, as shown in FIG. **68**, silicon oxide layers **506** and silicon nitride layers **507** are formed by, for example, a CVD method. Further, the silicon oxide layers **506** and the silicon nitride layers **507** are etched by RIE, which makes those remain as the sidewalls at the sidewall portions of the dummy gate electrodes.

[0416] Then, source/drain diffusion layers **505***b* are formed, with the dummy gate electrodes and the sidewalls being as masks, by carrying out ion implantation of a p-type impurity (for example, B, BF_2, or the like) into the PMOS area in a self-alignment manner, and by carrying out ion implantation of an n-type impurity (for example, P, As, or the like) into the NMOS area in a self-alignment manner.

[0417] Thereafter, nickel (Ni) layers are formed so as to have a thickness of about 10 nm on the source/drain diffusion layers **505***b* by, for example, sputtering. Then, silicon (Si) in the source/drain diffusion layers **505***b* and nickel (Ni) in the nickel layers are made to react with each other by carrying out thermal treatment at about 350° C. for about thirty seconds.

[0418] Subsequently, unreacted nickel layers are removed by using, for example, a compound liquid of sulfuric acid and hydrogen peroxide, and nickel silicide layers **508** are formed by carrying out thermal treatment at about 500° C. for about thirty seconds again.

[0419] Next, as shown in FIG. **69**, silicon oxide layers **509** covering the dummy gate electrodes are formed by, for example, a CVD method. Further, the silicon oxide layers **509** are grinded until the top surfaces of the dummy gate electrodes are exposed, and the surfaces thereof are planarized.

[0420] Then, grooves are formed by selectively removing the dummy gate electrodes and the silicon oxide layers **502** directly under those.

[0421] Subsequently, in the NMOS area, for example, In + ions are ion-implanted into the surface regions in the grooves of the semiconductor substrate **500**, a short time thermal treatment at about 1000° C. (rapid thermal anneal) is carried out, and a channel concentration is controlled, thereby adjusting a threshold voltage of the transistor.

[0422] After the dummy gate electrodes are removed, ultra thin gate insulating layers **510** are formed in the grooves, for example, a plasma oxynitride method.

[0423] Next, as shown in FIG. **70**, molybdenum (Mo) layers **511** with a work function of 4.9 eV are formed so as to have a thickness of about 150 nm by, for example, a CVD method using an organic source. Moreover, the molybdenum layers **511** are grinded by a CMP method, and those are filled into the grooves on the gate insulating layers **510**.

[0424] As a result, gate electrodes configured by the molybdenum layers **511** are formed at both of the PMOS area and the NMOS area.

[0425] Next, as shown in FIG. **71**, a GaP layer **512***a* is formed as a compound layer for work function modulation so as to have a thickness of 20 nm on the silicon oxide layers **509** by, for example, sputtering.

[0426] Then, the GaP layer **512***a* in the PMOS area is selectively removed by, for example, PEP and RIE, which makes the GaP layer **512***a* remain on the gate electrode of the n-channel MOS transistor in the NMOS area.

[0427] Further, a titanium nitride (TiN) layer **513** serving as a cap layer for preventing a metal from condensing is formed so as to have a thickness of about 20 nm on the silicon oxide layers **509** and the GaP layer **512***a*.

[0428] At this point in time, the NMOS area has a structure in which a lamination layer of GaP/TiN is arranged on the gate electrode made of Mo.

[0429] Next, as shown in FIG. **72**, thermal treatment at about 500° C. for about thirty seconds is carried out in the nitrogen atmosphere, for example. As a result, the GaP layer **512***a* does not condense because the GaP layer **512***a* is covered with the titanium nitride layer **513**. In addition, Ga in the GaP layer **512***a* diffuses via a grain boundary of the molybdenum layer **511**, and is precipitated at an interface between the gate insulating layer **510** and the molybdenum layer **511** to become a gallium precipitate layer **512***b*.

[0430] Thereafter, the GaP layer **512***a* and the titanium nitride layer **513** are removed.

[0431] Next, as shown in FIG. **73**, silicon oxide layers **515** covering the p-channel MOS transistor and the n-channel MOS transistor are formed by, for example, a CVD method. In addition, the surfaces of the silicon oxide layers **515** are planarized by, for example, a CMP method.

[0432] After contact holes coming down to the source/drain diffusion layers **505** are formed at the silicon oxide layers **509** and **515**, titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs are formed in the contact holes.

[0433] Electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are formed on the silicon oxide layers **515**, and a silicon oxide layer **518** covering those electrodes is formed on the silicon oxide layers **515**.

[0434] The gate electrode of the p-channel MOS transistor formed by such a method is composed of Mo whose work function is 4.9 eV.

[0435] In addition, the gate electrode of the n-channel MOS transistor is composed of a laminated structure of Ga/Mo. Here, a threshold value of the n-channel MOS transistor depends on a work function of Ga precipitated directly on the gate insulating layer **510**, and the value becomes 3.9 eV, so that work function modulation can be appropriately carried out.

[0436] Note that, as the gate insulating layer **510**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **510**.

[0437] C. Summary

[0438] In accordance with the seventh embodiment, in the method of manufacturing a dual metal gate electrode having a damascene gate structure, the condensation of metal is prevented, so that a material with a low work function is precipitated directly on the gate insulating layer, and work function modulation can be exactly carried out.

(8) Eighth Embodiment

[0439] An eighth embodiment relates to an application example of the second embodiment, and is an example in which the method of manufacturing a dual metal gate electrode according to the present invention is applied to a CMOS circuit having a damascene gate structure.

[0440] A. Device Structure

[0441] FIG. **74** is a plan view of a CMOS circuit which is formed by the method according to the eighth embodiment. FIG. **75** is a cross-sectional view taken along the line LXXV-LXXV of FIG. **74**, and FIG. **76** is a cross-sectional view taken along the line LXXVI-LXXVI of FIG. **74**.

[0442] An isolation layer **501** having an STI structure is arranged in a semiconductor substrate **500**. As a consequence, a PMOS area in which a p-channel MOS transistor is formed and an NMOS area in which an n-channel MOS transistor is formed are formed.

[0443] The p-channel MOS transistor is configured by source/drain diffusion layers **505***b*, a gate insulating layer **510** on a channel between the source/drain diffusion layers **505***b*, and a gate electrode on the gate insulating layer **510**. The gate electrode is configured by a molybdenum (Mo) layer **511** on the gate insulating layer **510**.

[0444] Silicon oxide layers **506** and silicon nitride layers **507** serving as sidewalls are arranged at the sidewall portions of the molybdenum layer **511**.

[0445] Nickel silicide layers **508** are formed on the surface regions of the source/drain diffusion layers **505***b*.

[0446] Electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are connected to the nickel silicide layers **508** and the molybdenum layer **511** serving as a gate electrode via titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs.

[0447] The n-channel MOS transistor is configured by source/drain diffusion layers **505***b*, a gate insulating layer **510** on a channel between the source/drain diffusion layers **505***b*, and a gate electrode on the gate insulating layer **510**. The gate electrode is configured by a GaMoP layer **514** on the gate insulating layer **510**.

[0448] Silicon oxide layers **506** and silicon nitride layers **507** serving as sidewalls are arranged at the sidewall portions of the GaMoP layer **514**.

[0449] Nickel silicide layers **508** are formed on the surface regions of the source/drain diffusion layers **505***b*.

[0450] Electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are connected to the nickel silicide layers **508** and the molybdenum layer **511** serving as a gate electrode via titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs.

[0451] Both of the p-channel MOS transistor and the n-channel MOS transistor are covered with silicon oxide layers **509**, **515**, and **518** serving as interlayer insulating layers.

[0452] B. Method of Manufacturing Semiconductor Device

[0453] A method of manufacturing a dual metal gate electrode by work function modulation according to the eighth embodiment will be described.

[0454] First, as shown in FIG. **77**, processes until the gate electrodes composed of molybdenum (Mo) layers **511** with a work function of 4.9 eV are formed are executed in accordance with the same processes as in the seventh embodiment described above.

[0455] Thereafter, a GaP layer **512***a* is formed as a compound layer for work function modulation so as to have a thickness of about 20 nm on the silicon oxide layers **509** by, for example, sputtering.

[0456] Then, the GaP layer **512***a* in the PMOS area is selectively removed by, for example, PEP and RIE, which makes the GaP layer **512***a* remain on the gate electrode of the n-channel MOS transistor in the NMOS area.

[0457] In addition, a titanium nitride (TiN) layer **513** serving as a cap layer for preventing a metal from condensing is formed so as to have a thickness of about 20 nm on the silicon oxide layers **509** and the GaP layer **512***a*.

[0458] At this point in time, the NMOS area has a structure in which a lamination layer of GaP/TiN is arranged on the gate electrode made of Mo.

[0459] Next, as shown in FIG. **78**, thermal treatment at about 400° C. for about one hour is carried out in the nitrogen atmosphere, for example. As a result, the GaP layer **512***a* does not condense because the GaP layer **512***a* is covered with the titanium nitride layer **513**, and the molybdenum layer **511** chemically reacts with the GaP layer **512***a* to become a GaMoP layer **514**.

[0460] Thereafter, the GaP layer **512***a* and the titanium nitride layer **513** are removed.

[0461] Next, as shown in FIG. **79**, silicon oxide layers **515** covering the p-channel MOS transistor and the n-channel MOS transistor are formed by, for example, a CVD method.

Further, the surfaces of the silicon oxide layers **515** are planarized by, for example, a CMP method.

[**0462**] After contact holes coming down to the source/drain diffusion layers **505** are formed at the silicon oxide layers **509** and **515**, titanium/titanium nitride **516***a* serving as barrier metals and tungsten **516***b* serving as contact plugs are formed in the contact holes.

[**0463**] In addition, electrodes composed of titanium nitride layers **517***a* and **517***c*, and an aluminum layer **517***b* are formed on the silicon oxide layers **515**, and a silicon oxide layer **518** covering those electrodes is formed on the silicon oxide layers **515**.

[**0464**] The gate electrode of the p-channel MOS transistor formed by such a method is composed of Mo whose work function is 4.9 eV. In addition, the gate electrode of the n-channel MOS transistor is composed of GaMoP whose work function is 4.2 eV. In this way, work function modulation of the gate electrode of the n-channel MOS transistor is appropriately carried out.

[**0465**] Note that, as the gate insulating layer **510**, not only silicon oxide, but also an oxide, a nitride, or an oxynitride of Hf, Zr, Ti, Ta, Al, Sr, Y, La and the like, or an oxide, a nitride, or an oxynitride as a compound of those elements and silicon can be used. For example, a high dielectric constant material such as a zirconium oxide and a hafnium oxide is used as the gate insulating layer **510**.

[**0466**] C. Summary

[**0467**] In accordance with the eighth embodiment, in the method of manufacturing a dual metal gate electrode having a damascene gate structure, the condensation of metal is prevented, so that a gate electrode can be transformed into an alloy layer, and work function modulation can be exactly carried out.

5. APPLICATION EXAMPLES

[**0468**] The examples of the present embodiment can be applied to a CMOS inverter circuit as shown in, for example, FIGS. **80** and **81**.

[**0469**] In the example of FIG. **80**, the p-channel MOS transistor PMOS and the n-channel MOS transistor NMOS are arranged side by side in the channel length direction, and the gate electrodes Gp and Gn are independent of each other.

[**0470**] In contrast thereto, in the example of FIG. **81**, the p-channel MOS transistor PMOS and the n-channel MOS transistor NMOS are arranged side by side in the channel width direction, and the gate electrodes Gp and Gn are unified.

[**0471**] In both cases, the transistors can be formed by utilizing the method of manufacturing a metal gate electrode according to the examples of the present invention.

4. OTHERS

[**0472**] In accordance with the examples of the present invention, in the method of a metal gate electrode, the condensation of metal can be prevented, and work function modulation can be exactly carried out.

[**0473**] Note that, in the embodiments described above, the examples of the MOS transistors have been described. However, as the effect that a material for the gate insulating layer is not particularly limited has been described, the examples of the invention can be applied to the general MIS transistors.

[**0474**] In addition, the examples of the invention can be applied to a semiconductor apparatus having a metal-insulator-semiconductor (MIS) structure, for example, an MIS capacitor.

[**0475**] The structure of the isolation layer is not limited to an STI, and may be a field insulating layer formed by, for example, a LOCOS method.

[**0476**] Additional advantages and modifications will readily occur to those skilled in the art. Therefore, the invention in its broader aspects is not limited to the specific details and representative embodiments shown and described herein. Accordingly, various modifications may be made without departing from the spirit or scope of the general inventive concept as defined by the appended claims and their equivalents.

What is claimed is:

1. A method of manufacturing a semiconductor device, comprising:

forming a gate insulating layer on a semiconductor substrate;

forming a first metal layer on the gate insulating layer;

forming a second metal layer including elements for use in work function modulation on the first metal layer;

forming a cap layer made of a material having a melting point higher than that of the second metal layer on the second metal layer; and

precipitating the elements at an interface between the gate insulating layer and the first metal layer by thermal treatment.

2. The method according to claim 1, wherein the second metal layer is formed in an area of an n-channel MOS transistor.

3. The method according to claim 1, wherein the first metal layer is selected from W, Pd, Pt, Ni, Co, Rh, Ir, Nb, Mo, Ta, Sb, Bi, Er, Ti, an alloy including at least one of these metals, and a nitride, a carbide, or a silicon nitride of these metals.

4. The method according to claim 1, wherein the second metal layer is selected from In, Ga, Tl, Sb, Bi, an alloy including at least one of these metals, and a III-V compound semiconductor including at least one of In and Ga.

5. A method of manufacturing a semiconductor device, comprising:

forming a gate insulating layer on a semiconductor substrate;

forming a first metal layer on the gate insulating layer;

forming a second metal layer including elements for use in work function modulation on the first metal layer;

forming a cap layer made of a material having a melting point higher than that of the second metal layer on the second metal layer; and

forming an alloy layer of the first metal layer and the second metal layer by thermal treatment.

6. The method according to claim 5, wherein the second metal layer is formed in an area of an n-channel MOS transistor.

7. The method according to claim 5, wherein the first metal layer is selected from W, Pd, Pt, Ni, Co, Rh, Ir, Nb, Mo, Ta, Sb, Bi, Er, Ti, an alloy including at least one of these metals, and a nitride, a carbide, or a silicon nitride of these metals.

8. The method according to claim 5, wherein the second metal layer is selected from In, Ga, Tl, Sb, Bi, an alloy including at least one of these metals, and a III-V compound semiconductor including at least one of In and Ga.

9. A method of manufacturing a semiconductor device, comprising:

forming a gate insulating layer on a semiconductor substrate;

forming a metal layer on the gate insulating layer;

forming a compound layer including elements for use in work function modulation on the metal layer; and

precipitating the elements at an interface between the gate insulating layer and the metal layer by thermal treatment.

10. The method according to claim 9, wherein the compound layer is formed in an area of an n-channel MOS transistor.

11. The method according to claim 9, wherein the elements are selected from a group consisting of In, Ga, Tl, Sb, and Bi.

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

forming a gate insulating layer on a semiconductor substrate;

forming a metal layer on the gate insulating layer;

forming a compound layer including elements for use in work function modulation on the metal layer; and

forming an alloy layer of the metal layer and the compound layer by thermal treatment.

13. The method according to claim 12, wherein the compound layer is formed in an area of an n-channel MOS transistor.

14. The method according to claim 12, wherein the elements are selected from a group consisting of In, Ga, Tl, Sb, and Bi.

15. A method of manufacturing a semiconductor device, comprising:

forming a gate insulating layer on a semiconductor substrate;

forming a metal layer on the gate insulating layer;

forming a semiconductor layer on the metal layer;

injecting elements for use in work function modulation into the semiconductor layer; and

precipitating the elements at an interface between the gate insulating layer and the metal layer by thermal treatment.

16. The method according to claim 15, wherein the elements are injected into an area of an n-channel MOS transistor.

17. The method according to claim 15, wherein the elements are selected from a group consisting of In, Ga, Tl, Sb, and Bi.

18. A method of manufacturing a semiconductor device, comprising:

forming a gate insulating layer on a semiconductor substrate;

forming a metal layer on the gate insulating layer;

forming a semiconductor layer on the metal layer;

injecting elements for use in work function modulation into the semiconductor layer; and

forming an alloy layer by making the elements react with the metal layer by thermal treatment.

19. The method according to claim 18, wherein the elements are injected into an area of an n-channel MOS transistor.

20. The method according to claim 18, wherein the elements are selected from a group consisting of In, Ga, Tl, Sb, and Bi.

* * * * *