Strained Channel Complementary Field-Effect Transistors and Methods of Manufacture

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application relates to co-pending and commonly assigned patent applications Serial No. 10/423,513 (TSM03-0173), filed April 25, 2003, entitled " Strained Channel Transistor and Methods of Manufacture," and Serial No. 10/379,033 (TSM03-0050), filed February 28, 2003, entitled "Strained-Channel Transistor With a Lattice-Mismatched Zone in the Source/Drain Regions." Both applications are hereby incorporated herein by reference.

TECHNICAL FIELD

[0002] The present invention relates generally to semiconductor devices, and more particularly, the preferred embodiment relates to strained channel complementary field-effect transistors and methods of manufacture.

BACKGROUND

[0003] Size reduction of metal-oxide-semiconductor field-effect transistors (MOSFET), including reduction of the gate length and gate oxide thickness, has enabled the continued improvement in speed performance, density, and cost per unit function of integrated circuits over the past few decades. To enhance transistor performance further, strain may be introduced in the transistor channel for improving carrier mobilities. Therefore, strain-induced mobility enhancement is another way to improve transistor performance in addition to device scaling. There are several existing approaches of introducing strain in the transistor channel region.

[0004] In one conventional approach, as described in a paper by J. Welser et al., published at the December 1992 International Electron Devices Meeting held in San Francisco, CA, pp. 1000-1002 and incorporated herein by reference, a relaxed silicon germanium (SiGe) buffer layer is provided beneath the channel region. Figure 1a shows such an approach. Referring to Figure 1a, a semiconductor device 100 includes a strained silicon layer 110 formed over and abutting a relaxed SiGe layer 112, which is formed over and abutting a graded SiGe buffer layer 114. The graded SiGe buffer layer 114 is formed over and abutting a silicon substrate 116.

[0005] The relaxed SiGe layer 112 has a larger lattice constant compared to relaxed Si, and the thin layer of epitaxial Si 110 grown on the relaxed SiGe 112 will have its lattice stretched in the lateral direction, i.e., it will be under biaxial tensile strain. This result is illustrated in Figures 1b and 1c. Therefore, a transistor 118 formed on the epitaxial strained silicon layer 110 will have a channel region 120 that is under biaxial tensile strain. In this approach, the relaxed SiGe buffer layer 112 can be thought of as a stressor that introduces strain in the channel region 120. The stressor, in this case, is placed below the transistor channel region 120.

[0006] Significant mobility enhancement has been reported for both electrons and holes in bulk transistors using a silicon channel under biaxial tensile strain. In the above-mentioned approach, the epitaxial silicon layer is strained before the formation of the transistor. But there are concerns about the strain relaxation upon subsequent CMOS processing where high temperatures are used. In addition, this approach is very expensive since a SiGe buffer layer with thickness in the order of micrometers has to be grown. Numerous dislocations in the relaxed SiGe buffer layer exist and some of these dislocations propagate to the strained silicon layer, resulting in a substrate with high defect density. Thus, this approach has limitations that are related to cost and fundamental material properties.

[0007] In another approach, strain in the channel is introduced after the transistor is formed. In this approach, a high stress film 132 is formed over a completed transistor structure 130 formed in a silicon substrate 136, as shown in Figure 2. The high stress film or stressor 132 exerts significant influence on the channel 134, modifying the silicon lattice spacing in the channel region, and thus introducing strain in the channel region. In this case, the stressor 132 is placed above the completed transistor structure. This scheme is described in detail in a paper by A. Shimizu et al., entitled "Local mechanical stress control (LMC): a new technique for CMOS performance enhancement," published in pp. 433-436 of the Digest of Technical Papers of the 2001 International Electron Device Meeting, which is incorporated herein by reference.

[0008] The strain contributed by the high stress film is believed to be uniaxial in nature with a direction parallel to the source-to-drain direction. However, uniaxial tensile strain degrades the hole mobility while uniaxial compressive strain degrades the electron mobility. Ion implantation of germanium can be used to selectively relax the strain so that the hole or electron mobility is not degraded, but this is difficult to implement due to the close proximity of the n and p-channel transistors.

SUMMARY OF THE INVENTION

[0009] In prior art schemes of inducing strain in transistors, the introduction of a strain by a stressor may benefit transistors of the first conduction type while degrading transistors of the second conduction type. This and other problems are generally solved or circumvented, and technical advantages are generally achieved, by preferred embodiments of the present invention, which provide a method of manufacturing strained n-channel and p-channel field effect transistors with separately optimized performance.

[009a] In one aspect, there is provided a method of forming a transistor device, the method comprising: providing a substrate comprising a first semiconductor material;
forming a gate dielectric on the substrate;
forming a gate electrode on the gate dielectric;
forming spacers on sidewalls of the gate electrode, the spacers each having a void, wherein said void exposes a portion of said spacer at a lower outer corner thereof;
forming a source region and a drain region within the first semiconductor material adjacent opposite ends of the gate electrode, said source region and drain region comprising a second semiconductor material; and
forming a high-stress film over the gate electrode, the spacers, the source region and the drain region.

[0010] A first embodiment provides a transistor structure for decoupling the effect of a stressor on the channel region of the transistor. Such a structure is useful in the case where carrier mobility may be degraded by the stressor. Disclosed herein is an integration scheme for the optimization of strain in the channel regions of n-channel and p-channel transistors.

[0011] In accordance with a preferred embodiment disclosed herein, a transistor includes a gate dielectric overlying a channel region. A source region and a drain region are located on opposing sides of the channel region. The channel region is formed from a first semiconductor material and the source and drain regions are formed from a second semiconductor material. A gate electrode overlies the gate dielectric. A pair of spacers is formed on sidewalls of the gate electrode. Each of the spacers includes a void adjacent the channel region. A high-stress film can overlie the gate electrode and spacers.

[0012] In accordance with another preferred embodiment disclosed herein, a transistor device is formed on a substrate of a first semiconductor material. A gate dielectric is formed on the substrate and a gate electrode is formed on the gate dielectric. Spacers are formed

on sidewalls of the gate electrode. The spacers each have a void. A source region and a drain region are formed within the first semiconductor material adjacent opposite ends of the gate electrode. A high-stress film can be formed over the gate electrode, the spacers, the source region and the drain region.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] For a more complete understanding of the present invention, and the advantages thereof, reference is now made to the following descriptions taken in conjunction with the accompanying drawing, in which:

[0014] Figure 1a shows a conventional strained silicon transistor with a relaxed SiGe layer as a stressor to induce strain in the top epitaxial strained silicon layer;

[0015] Figures 1b and 1c illustrate the origin of strain in the Si/SiGe heterostructures;

[0016] Figure 2 illustrates another way of introducing strain in the channel is by using a high stress film as a stressor;

[0017] Figures 3a-3g show a first embodiment process flow; and

[0018] Figures 4a-4d show a second embodiment process flow.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0019] The making and using of the presently preferred embodiments are discussed in detail below. It should be appreciated, however, that the present invention provides many applicable inventive concepts that can be embodied in a wide variety of specific contexts. The specific embodiments discussed are merely illustrative of specific ways to make and use the invention, and do not limit the scope of the invention.

[0020] The preferred embodiment of the present invention relates to the field of semiconductor devices, and more specifically, to the manufacture of strained n-channel and p-channel field effect transistors with separately optimized performance.

[0021] With the preferred embodiment it is possible to separately optimize the performance of n- channel and p-channel transistors by engineering the nature and magnitude of the strain in the channel region of the transistors. For example, it is desirable to induce a tensile strain in the channel of the n-channel transistor in the source-to-drain direction. It is also desirable to induce a compressive strain in the channel of the p-channel transistor in the source-to-drain direction.

[0022] In one embodiment, a high-stress film with a tensile stress may be employed to induce tensile strain in the channel region of transistors. Tensile strain in the source-to-drain direction, however, degrades the mobility of p-channel transistors while improving the mobility of n-channel transistors. According to the preferred embodiment of this invention, a stress-decoupling region, e.g. a void, can be employed to decouple the effect of a high-stress film on the channel region if the effect of the high-stress film is to degrade the mobility of carriers in the channel region.

[0023] For example, a high-stress film with a tensile stress exerts a tensile strain in the source-to-drain direction, degrading hole mobility in p-channel transistors. A void may be formed in the spacer region of the p-channel transistor to reduce the effect of the high-stress film. Free surfaces represent a strain free boundary condition and therefore represent regions with very low stress. A void region is defined by or surrounded by a free surface, and the stress in the vicinity of the void is expected to be very low. By placing the void between the high-stress film and the channel region of the p-channel transistor, the negative impact of the tensile stress on the p-channel transistor can be reduced. Therefore, the stress of the high-stress film can be decoupled from selected channel regions by placing voids in the vicinity of the selected channel regions, such as in the spacers of selected transistors. A larger void would result in more decoupling and therefore lower stress levels. For example, the void in the spacers of the p-channel transistor can be larger than the void in the spacers of the n-channel transistor.

[0024] Figures 3a-3g illustrate a first embodiment method for separately optimizing the strain in the channel region for transistors in a semiconductor chip. In the preferred embodiment, the starting material is a silicon substrate, but it is understood that other substrates comprising compound semiconductor substrates, e.g., gallium arsenide, or alloy semiconductor, e.g., silicon-germanium, may be used. The starting material may also be a semiconductor-on-insulator substrate, such as a silicon-on-insulator substrate. The starting material may also include an epitaxially grown semiconductor layer and/or a doped region within a semiconductor layer, e.g., a triple well structure.

[0025] Figure 3a shows the preferred embodiment where the starting material is a bulk silicon substrate 202. Isolation regions 204a, 204b and 204c (collectively 204) are formed in the substrate 202. The isolation regions 204 are preferably shallow trench isolation regions known

and used in the art. It is understood that other isolation structures, such as field oxide (e.g., formed by the local oxidation of silicon) may be used.

[0026] Figure 3a illustrates two active areas, i.e., a first between isolation region 204a and 204b and a second between 204b and 204c. For purposes of illustration, an n-channel transistor will be formed in one of the active areas and a p-channel transistor will be formed in the other active area. As is known in the art, the portion of substrate 202 that includes the n-channel transistor will be doped with p-type dopants and the portion of the substrate 202 that includes the p-channel transistor will be doped with n-type dopants. These doped portions can be accomplished with the use of one or more wells, as know in the art.

[0027] A gate dielectric layer 206 is formed on a top surface of semiconductor region 202. The gate dielectric 206 may comprise silicon oxide, silicon oxynitride, and silicon nitride, or any combination of these materials. Alternatively, the gate dielectric 206 can be formed from a material having a relative permittivity greater than about 5. Examples of such materials include aluminum oxide, lanthanum oxide, hafnium oxide, zirconium oxide, hafnium oxynitride, or combinations thereof.

[0028] A gate electrode material 208 is then deposited. The gate electrode material preferably comprises a conductive material, such as a metal (e.g., tantalum, titanium, molybdenum, tungsten, platinum, aluminum, hafnium, ruthenium), a metal silicide (e.g., titanium silicide, cobalt silicide, nickel silicide, tantalum silicide), a metal nitride (e.g., titanium nitride, tantalum nitride), doped poly-crystalline silicon or poly-crystalline silicon germanium, or any other conductive materials. Combinations of these materials can be used. Any of these conductors can be used along with any of the gate dielectrics listed above. In one example, amorphous silicon is deposited and recrystallized to create poly-crystalline silicon.

[0029] A lithography step is then performed followed by an etching step to form a gate stack as shown in Figure 3a. The gate stack includes gate electrode 208 and underlying gate dielectric 206. In the preferred embodiment, the gate electrode material 208 is poly-crystalline silicon (poly-Si) and the gate dielectric material 206 is silicon oxynitride. The thickness of gate electrode 208 will typically range between about 500 Å and about 2000 Å, preferably less than about 1500 Å. The equivalent silicon oxide thickness of gate dielectric 206 will typically range between about 3 Å and about 50 Å, preferably less than about 20 Å.

[0030] Source and drain extension regions 210 and 212 are formed by ion implantation. In the preferred embodiment, two transistors 200 and 201 with channel regions of different conductivity types are formed. For example, the p-channel extension regions 210 (or 212) can be formed by implanting a dopant such as boron while the n-channel devices are masked. Likewise, the n-channel extension regions 212 (or 210) can be formed by implanting a dopant such as arsenic and/or phosphorus while the p-channel devices are masked.

[0031] Referring now to Figure 3b, spacers 214/216 are formed on the sides of the gate electrode 208. In one example, the spacers may be formed by chemical vapor deposition of a dielectric material, e.g., silicon oxide or silicon nitride, followed by an anisotropic etching of the dielectric material to form simple spacers. In the example of Figure 3b, the spacers are composite spacers. A composite spacer may comprise a dielectric liner 214 and a spacer body 216. The dielectric liner 214 may be formed by the deposition of a dielectric liner material, e.g., silicon oxide, and the spacer body material 216, e.g. silicon nitride, followed by an anisotropic etch using reactive ion etching. In another embodiment, the liner 214 may be an oxide and the spacer body 216 may be a nitride.

[0032] According to the preferred embodiment, a transistor 200 of a first conductivity type in a first active region (the portion of semiconductor region 202 between isolation regions 204a and 204b) is to be formed with a lattice-mismatched zone in the source and drain regions. Referring to Figure 3c, this can be performed by forming a dielectric mask 218 over the second active region (the portion of semiconductor region 202 between isolation regions 204b and 204c), which includes transistor 201. The transistor 200 of the first conductivity type is exposed.

[0033] As shown in Figure 3c, a recess 220 is etched into the source and drain regions of the transistor of the first conductivity type. The etching of the recess 220 may be performed by a dry plasma etching process. In the preferred embodiment, where the substrate 202 is a silicon substrate, the plasma etching process may employ fluorine chemistry. It should be noted that in the etching of the recess 220, the dielectric liner 214 of the composite spacer might be slightly etched to form an undercut 222.

[0034] Referring now to Figure 3d, selective epitaxy is performed to form source and drain regions 224 for the transistor of the first conductivity type. The epitaxially grown material 224 is preferably a semiconductor material having a different lattice constant from that of the substrate 202 material. In the case where the first conductivity type is p-type, the lattice-mismatched material 224 is preferably silicon germanium. In the case where the first conductivity type is n-type, the lattice-mismatched material 224 is preferably silicon carbon or silicon-germanium-carbon ($Si_{1-x-y}Ge_xC_y$) where the mole fraction of carbon y is greater than a tenth of the mole fraction of germanium x. The height of the source and drain region 224 in the transistor of the first conductivity type may be higher than that in the transistor of the second conductivity type.

[0035] It is noted that the process steps of forming the mask 218, etching recess 220, and epitaxial growing lattice-mismatched material 224 can be performed prior to the formation of spacers 214/216 on the sides of the gate electrode 208.

[0036] As shown in Figure 3e, the dielectric mask 218 is removed. In the case where the dielectric mask 218 is silicon oxide, this may be accomplished by an etch in dilute hydrofluoric acid. The removal of the mask 218 may result in an undercut 226 in the dielectric liner 214 of the transistor 201. The undercut 222 may be larger than the undercut 226 because it underwent two such etch processes.

[0037] The source and drain regions for the NMOS transistor 201 are now formed by using ion implantation while covering the PMOS transistors 202. In the preferred embodiment, the dopant is arsenic or phosphorus or a combination of both. The source and drain regions 225 are shown in Figure 3f.

[0038] As also shown in Figure 3f, a high-stress film 228 is formed over the transistors 200 and 201. The high-stress film 228 may be silicon nitride or any other high-stress material. The stress may be compressive or tensile in nature and may have a magnitude in the range of 0.1 to 4 giga-pascals (GPa). The high-stress film 228 is preferably formed by a chemical vapor deposition (CVD) process. The CVD process can be a low-pressure CVD (LPCVD) process or a plasma-enhanced CVD (PECVD) process, as commonly known and used in the art.

[0039] As a result of the undercuts 222 and 226 in the spacers of the transistors, voids 230 and 232 may be formed in the spacers. This is intentional and may bring additional benefits to be described. It is noted that during the formation of the high stress film 228, it is possible to completely fill the undercuts so that no voids will be formed in the final structure. One aspect of this invention teaches the intentional creation of voids in a strained channel transistor to

specifically decouple negative effects of the high stress film on transistors of one conductivity type.

[0040] In the preferred embodiment, the first conductivity type is p-type and the second conductivity type is n-type. Each void 230 and 232 has a lateral extent that can be identified as the length 234 of the void. For purposes of illustration, the length 234 of the void 232 in the spacer of the n-channel transistor 201 is indicated in Figure 3f. The size of a void is the area of the void in the cross-sectional view. A first void is larger than a second void, when the area of the first void in a cross-sectional view is larger than the area of the second void in the same cross-sectional view. A void that has a larger length or a larger size is expected to have a higher decoupling effect than one that has a smaller length or a smaller size.

[0041] In the preferred embodiment, the size of voids 230 and 232 can be measured relative to the width of the spacer 214/216. For example, the void 232 may have a length 234 larger than 5% of the spacer width. In another embodiment, the void 232 has a length that is at least 10% of the width of spacer 214/216.

[0042] If a single high tensile stress film 228 is used to cover both n-channel and p-channel transistors 201 and 200, the channel regions 236 of the transistors will experience a tensile strain. The n-channel transistor 201 will have benefits of mobility enhancement, while the p-channel transistor 200 will have degraded mobility. According to the preferred embodiment, voids 230 may be introduced in the vicinity of the channel region 236 of the p-channel transistor 200 to reduce the negative effects of the tensile stress film 228. According to another embodiment of this invention, the channel region 236 of the p-channel transistor 200 may have a larger void 230 than the void 232 adjacent the channel region 236 of the n-channel transistor 201.

[0043] In another embodiment, the p-channel transistor 200 may have a high-stress film 238 that is compressive in nature and the n-channel transistor 201 may have a high-stress film 228 that is tensile in nature, as shown in Figure 3g. A compressive stress that results in compressive strain in the channel region 236 of the p-channel transistor 200 in the source-to-channel direction will enhance the hole mobility. A tensile stress that result in tensile strain in the channel region 236 of the n-channel transistor 201 in the source-to-channel direction will enhance the electron mobility.

[0044] The structure of Figure 3g may be formed by first forming the structure of Figure 3f and forming a mask (not shown) to cover the tensile high-stress film 228 over the n-channel transistor 201. The high-stress film 228 over the p-channel transistor 200 can then be removed in accordance with the mask. A compressive high-stress film 238 can then be formed over the p-channel transistor 200. For example, the film can be formed over the entire device and then etched to the desired pattern.

[0045] Another method embodiment of the present invention is illustrated in Figures 4a-4d. Items previously discussed with respect to Figures 3a-3g will not be repeated here. Transistors 200 and 201 of two different conductivity types are formed in two different active regions, and the source and drain regions of each of the transistors are recessed using an etch. The etch may employ a dry plasma etch, for example. The resulting structure is shown in Figure 4a, which shows recesses 220 and 240 in transistors 200 and 201, respectively.

[0046] Referring now to Figure 4b, lattice-mismatched regions 242 and 244 are epitaxially grown in the recesses 220 and 240 of the source and drain regions. Different lattice-mismatched materials may be used for the p-channel and the n-channel transistors 200 and 201. This may be accomplished, for example, by forming a first mask (not shown) to cover the n-channel transistor

201 and performing a first epitaxy to selectively grow a first semiconductor material 242 in the source and drain region of the p-channel transistor 200.

[0047] After the first mask is removed, a second mask (not shown) is formed a second mask to cover the p-channel transistor 200. A second epitaxy is then performed to selectively grow a second semiconductor material 244 in the source and drain region of the n-channel transistor 201. The resulting structure is shown in Figure 4b.

[0048] In the preferred embodiment, the first semiconductor material 242 may be silicon-germanium and the second semiconductor material 244 may be silicon-carbon or silicon-germanium-carbon ($Si_{1-x-y}Ge_xC_y$) where the mole fraction of carbon y is greater than a tenth of the mole fraction of germanium x. Since the substrate 202 material is preferably silicon, the presence of a first semiconductor material such as silicon-germanium with a larger lattice constant will result in compressive strain in the channel region 236 of the p-channel transistor 200. Similarly, the presence of a second semiconductor material such as silicon-carbon with a smaller lattice constant than the channel material will results in tensile strain in the channel region 236 of the n-channel transistor 201.

[0049] The spacers 214/216 of the transistors 200 and 201 may be composite spacers, and the dielectric liner 214 of the spacers may be intentionally undercut by exposing the dielectric liner 214 to an etch process. If the dielectric liner 214 is silicon oxide, the etch can be a hydrofluoric acid etch. The resulting structure is shown in Figure 4c.

[0050] A high-stress film 228 (or 238) is then formed over the transistors 200 and 201, as shown in Figure 4d. Although not shown, different films can be formed over the n-channel and the p-channel transistors as discussed previously with respect to Figure 3g. In another

embodiment, a high-stress film is formed over one of the transistors 200 (or 201) but not the other transistor 201 (or 200).

[0051] While this invention has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense. Various modifications and combinations of the illustrative embodiments, as well as other embodiments of the invention, will be apparent to persons skilled in the art upon reference to the description. It is therefore intended that the appended claims encompass any such modifications or embodiments.

CLAIMS

1. A method of forming a transistor device, the method comprising:
providing a substrate comprising a first semiconductor material;
forming a gate dielectric on the substrate;
forming a gate electrode on the gate dielectric;
forming spacers on sidewalls of the gate electrode, the spacers each having a void, wherein said void exposes a portion of said spacer at a lower outer corner thereof;
forming a source region and a drain region within the first semiconductor material adjacent opposite ends of the gate electrode, said source region and drain region comprising a second semiconductor material; and
forming a high-stress film over the gate electrode, the spacers, the source region and the drain region.

2. The method of claim 1 wherein the transistor is a p-channel transistor.

3. The method of claim 2 further comprising forming a recess within the first semiconductor material adjacent opposite ends of the gate electrode prior to forming the source and drain regions.

4. The method of claim 3 further comprising growing a second semiconductor material in the recess.

5. The method of claim 4 wherein the second semiconductor material includes silicon and germanium.

6. The method of claim 4 and further comprising, after the step of forming the source and drain regions, forming a first conductive material on the source and drain regions.

7. The method of claim 1 and further comprising, after the step of forming the gate electrode, forming source and drain extension regions in portions of the first semiconductor region aligned with the gate electrode.

8. The method of claim 1 wherein the transistor is a p-channel transistor and wherein the high-stress film has a compressive stress.

9. The method of claim 1 wherein the high-stress film has a tensile stress.

10. The method of claim 9 and further comprising implanting ions into the high-stress film.

11. The method of claim 1 wherein the transistor is an n-channel transistor.

12. The method of claim 11 wherein the high-stress film has a compressive stress.

13. The method of claim 12 further comprising implanting ions into the high-stress film.

14. The method of claim 11 wherein the high-stress film has a tensile stress.

15. The method of claim 1 wherein the transistor is an n-channel transistor, the method further comprising forming a recess within the first semiconductor material adjacent opposite ends of the gate electrode prior to forming the source and drain regions.

16. The method of claim 15 further comprising growing a second semiconductor material in the recess.

17. The method of claim 16 wherein the second semiconductor material includes silicon and carbon.

18. The method of claim 1 wherein the transistor is an n-channel transistor, the method further comprising, after forming the source and drain regions, forming a second conductive material on the source and drain regions.

Strained Channel Complementary Field-Effect Transistors and Methods of Manufacture

ABSTRACT OF THE DISCLOSURE

A transistor includes a gate dielectric overlying a channel region. A source region and a drain region are located on opposing sides of the channel region. The channel region is formed from a first semiconductor material and the source and drain regions are formed from a second semiconductor material. A gate electrode overlies the gate dielectric. A pair of spacers is formed on sidewalls of the gate electrode. Each of the spacers includes a void adjacent the channel region. A high-stress film can overlie the gate electrode and spacers.

(FIG. 3a)