Hybrid Process for Forming Metal Gates of MOS Devices

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application relates to the following commonly assigned U.S. Patent Application: Application Serial No. 11/552,704, filed October 25, 2006, entitled "Semiconductor Devices with Dual-Metal Gate Structures and Fabrication Methods Thereof," which patent application is incorporated herein by reference.

TECHNICAL FIELD

[0002] This invention relates generally to semiconductor devices, and more particularly to structures of metal-oxide-semiconductor (MOS) devices and manufacturing methods for forming the same.

BACKGROUND

[0003] Metal-oxide-semiconductor (MOS) devices are basic building elements in integrated circuits. A conventional MOS device typically has a gate electrode comprising polysilicon doped with p-type or n-type impurities, using doping operations such as ion implantation or thermal diffusion. It is preferred to adjust the work function of the gate electrode to the band-edge of the silicon; that is: for an NMOS device, adjusting the work function close to the conduction band, and for a PMOS device, adjusting the work function close to the valence band. Adjusting the work function of the polysilicon gate electrode can be achieved by selecting appropriate impurities.

[0004] MOS devices with polysilicon gate electrodes exhibit carrier depletion effect, which is also referred to as a poly depletion effect. The poly depletion effect occurs when applied electrical fields sweep away carriers from regions close to gate dielectrics, forming depletion

layers. In an n-doped polysilicon layer, the depletion layer includes ionized non-mobile donor sites, whereas in a p-doped polysilicon layer, the depletion layer includes ionized non-mobile acceptor sites. The depletion effect results in an increase in the effective gate dielectric thickness, making it more difficult for an inversion layer to be created at the surface of the semiconductor.

[0005] The use of thin gate dielectrics tends to make the carrier depletion effect worse. With thin gate dielectrics, the depletion layer in the polysilicon gate becomes more significant in thickness when compared to the thickness of the thin gate dielectrics, and thus device performance degradation worsens. As a result, the carrier depletion effect in the gate electrodes limits device scalability by imposing a lower bound on how much the effective gate dielectric thickness can be reduced.

[0006] The poly depletion effect was previously solved by forming metal gate electrodes or metal silicide gate electrodes, wherein the metallic gates used in NMOS devices and PMOS devices also preferably have band-edge work functions. Currently, materials suitable for forming gate electrodes of NMOS devices, such as TaC, have been found. However, for PMOS devices, even though metallic materials having band-edge work functions have been found, these materials have poor thermal stability. When exposed to the high temperatures in the front-end-of-line processes, the work functions of these metallic materials shift, for example, toward the mid-gap level. The performance of the resulting PMOS devices is thus adversely affected.

[0007] Existing processes for forming dual-metal complementary MOS (CMOS) devices include two main categories, gate-first approach and gate-last approach. Both approaches have advantageous and disadvantageous features. In a typical gate-first approach, two metal layers having different work functions are separately formed in PMOS and NMOS regions. The metal

layers are then patterned to form gate electrodes. Other components of MOS devices, such as spacers, lightly doped source/drain (LDD) regions, source/drain regions, silicides, and contact etch stop layers are then formed. This process is relatively simple, and the resulting contact etch stop layers are continuous, so that they can effectively apply stresses. However, since the metal gates are formed before the formation and the activation of LDD regions and source/drain regions, they suffer from high thermal budgets, and the work functions of PMOS devices may shift. In addition, if composite metal layers are used, the oxygen in the composite metal layer may be released under the thermal budgets, and cause interfacial layer re-growth. Further, patterning metal layers by etching is relatively difficult, particularly for metals used for the PMOS devices.

[0008] Gate-last approach, on the other hand, typically includes the steps of forming dummy gates for both PMOS and NMOS devices. LDD regions, gate spacers, source/drain regions, and contact etch stop layers are then formed. The dummy gates of PMOS and NMOS devices are then removed, and metals with different work functions are then filled into the openings for PMOS and NMOS devices. In the gate-last approach, metal gates of PMOS devices and NMOS devices both take the advantage of low thermal budgets since they are formed after the formation and activation of LDD regions and source/drain regions. However, the process is complex. In addition, in the cases wherein the formation of high-k dielectrics also uses gate-last approach, the quality of the high-k dielectrics was often not satisfactory. Besides, forming high-k dielectrics on sidewalls of the openings will adversely increase the fringing capacitance between the gate and nearby features, such as source/drain regions and contacts.

[0009] Accordingly, what is needed in the art is a semiconductor structure and respective formation methods that may incorporate dual metal gates thereof to take advantage of the

benefits associated with band-edge work functions while at the same time overcoming the deficiencies of the prior art.

SUMMARY OF THE INVENTION

[0010] In accordance with one aspect of the present invention, a semiconductor structure includes a first MOS device including a first gate, and a second MOS device including a second gate. The first gate includes a first high-k dielectric over a semiconductor substrate; a second high-k dielectric over the first high-k dielectric; a first metal layer over the second high-k dielectric, wherein the first metal layer dominates a work-function of the first MOS device; and a second metal layer over the first metal layer. The second gate includes a third high-k dielectric over the semiconductor substrate, wherein the first and the third high-k dielectrics are formed of same materials, and have substantially a same thickness; a third metal layer over the third high-k dielectric, wherein the third metal layer and the first metal layer are formed of same materials, and have substantially a same thickness; and a fourth metal layer over the third metal layer.

[0011] In accordance with another aspect of the present invention, a semiconductor structure includes a first MOS device including a first gate, and a second MOS device including a second gate. The first gate includes a first high-k dielectric over a semiconductor substrate; a second high-k dielectric over the first high-k dielectric, wherein the first and the second high-k dielectrics are formed of different materials; a first metal layer over the second high-k dielectric, wherein the first metal layer has a thickness great enough for dominating a work-function of the first MOS device; a second metal layer over the first metal layer, wherein the first and the second metal layers are formed of different materials; and a third metal layer over the second metal layer, wherein the third metal layer has a work function close to a valence band of silicon. The second gate includes the first high-k dielectric over the semiconductor substrate; the second metal layer over the first high-k dielectric, wherein the second metal layer in the second gate has

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.