

(12) 发明专利申请

(10) 申请公布号 CN 102016926 A

(43) 申请公布日 2011.04.13

(21) 申请号 200980114125.4

(74) 专利代理机构 北京律盟知识产权代理有限

(22) 申请日 2009.04.21

责任公司 11287

(30) 优先权数据

代理人 宋献涛

12/106,654 2008.04.21 US

(51) Int. Cl.

(85) PCT申请进入国家阶段日

G06T 15/00 (2006.01)

2010.10.21

(86) PCT申请的申请数据

PCT/US2009/041268 2009.04.21

(87) PCT申请的公布数据

W02009/132013 EN 2009.10.29

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 杜云 于春 焦国方

斯蒂芬·莫洛伊

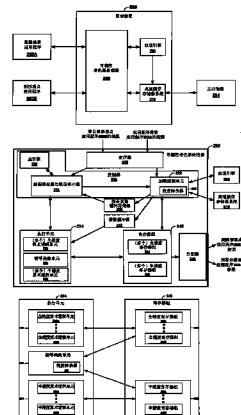
权利要求书 8 页 说明书 14 页 附图 6 页

(54) 发明名称

具有混合精度指令执行的可编程串流处理器

(57) 摘要

本发明涉及一种可编程串流处理器，其能够使用不同执行单元来执行混合精度（例如，全精度、半精度）指令。所述各种执行单元各自能够使用图形数据来在特定精度级别下执行指令。示范性可编程着色器处理器包括一控制器及多个执行单元。所述控制器经配置以接收供执行的指令及接收对所述指令的执行的数据精度的指示。所述控制器还经配置以接收单独转换指令，所述单独转换指令在经执行时将与所述指令相关联的图形数据转换到所述所指示的数据精度。当可操作时，所述控制器基于所述所指示的数据精度来选择所述执行单元中的一者。所述控制器接着使所述选定执行单元使用与所述指令相关联的所述图形数据以所述所指示数据精度执行所述指令。



1. 一种方法，其包含：

接收供在可编程串流处理器内执行的图形指令；

接收对所述图形指令的执行的数据精度的指示；

接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所述所指示数据精度，所述转换指令不同于所述图形指令；

基于所述所指示数据精度来选择所述处理器内的多个执行单元中的一者；以及

使用所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令。

2. 根据权利要求 1 所述的方法，其进一步包含：

接收与所述图形指令相关联的所述图形数据；

在所述选定执行单元执行所述图形指令期间产生具有所述所指示数据精度的计算结果；以及

将所述计算结果作为输出来提供。

3. 根据权利要求 1 所述的方法，其中选择所述多个执行单元中的一者包含：

选择所述处理器内的第一组一个或一个以上执行单元中的一者，所述第一组一个或一个以上执行单元在所述所指示数据精度为第一数据精度时各自使用所述图形数据以所述第一数据精度执行指令；以及

选择所述处理器内的第二组一个或一个以上执行单元中的一者，所述第二组一个或一个以上执行单元在所述所指示数据精度为第二数据精度时各自使用所述图形数据以所述第二数据精度执行指令，所述第二数据精度不同于所述第一数据精度。

4. 根据权利要求 3 所述的方法，其中所述第一数据精度包含全数据精度，且其中所述第二数据精度包含半数据精度。

5. 根据权利要求 1 所述的方法，其中所述执行单元包括所述处理器内的第一组一个或一个以上执行单元，所述第一组一个或一个以上执行单元各自使用所述图形数据以第一数据精度执行指令，且所述执行单元进一步包括所述处理器内的第二组一个或一个以上执行单元，所述第二组一个或一个以上执行单元各自使用所述图形数据以不同于所述第一数据精度的第二数据精度执行指令。

6. 根据权利要求 5 所述的方法，其中：

基于所述所指示数据精度来选择所述处理器内的所述多个执行单元中的一者包含选择所述第一组中的所述执行单元中的一者；且

使用所述选定执行单元来执行所述图形指令包含使用所述第一组中的所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令。

7. 根据权利要求 6 所述的方法，其进一步包含：

接收供在所述处理器内执行的第二图形指令；

接收对所述第二图形指令的执行的所述第二数据精度的指示；

接收第二转换指令，所述第二转换指令在由所述处理器执行时将与所述第二图形指令相关联的图形数据转换到所述所指示的第二数据精度，所述第二转换指令不同于所述第二图形指令；

基于所述所指示的第二数据精度来选择所述第二组中的所述执行单元中的一者；以及

使用所述第二组中的所述选定执行单元来使用与所述第二图形指令相关联的所述图形数据以所述所指示的第二数据精度执行所述第二图形指令。

8. 根据权利要求 1 所述的方法，其中接收对所述图形指令的执行的所述数据精度的所述指示包含对所述图形指令进行解码以确定所述数据精度。

9. 根据权利要求 1 所述的方法，其中使用所述选定执行单元来使用所述图形数据以所述所指示数据精度执行所述图形指令包含：

将所述图形数据转换到所述所指示数据精度；以及

使用所述选定执行单元来使用所述经转换图形数据以所述所指示数据精度执行所述图形指令。

10. 根据权利要求 1 所述的方法，其中与所述图形指令相关联的所述图形数据包含顶点图形数据及像素图形数据中的至少一者。

11. 一种计算机可读媒体，其包含用于使可编程串流处理器进行以下操作的指令：

接收供在所述处理器内执行的图形指令；

接收对所述图形指令的执行的数据精度的指示；

接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所述所指示数据精度，所述转换指令不同于所述图形指令；

基于所述所指示数据精度来选择所述处理器内的多个执行单元中的一者；以及

使用所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令。

12. 根据权利要求 11 所述的计算机可读媒体，其进一步包含用于使所述处理器进行以下操作的指令：

接收与所述图形指令相关联的所述图形数据；

在所述选定执行单元执行所述图形指令期间产生具有所述所指示数据精度的计算结果；以及

将所述计算结果作为输出来提供。

13. 根据权利要求 11 所述的计算机可读媒体，其中所述用于使所述处理器选择所述多个执行单元中的一者的指令包含用于使所述处理器进行以下操作的指令：

选择所述处理器内的第一组一个或一个以上执行单元中的一者，所述第一组一个或一个以上执行单元在所述所指示数据精度为第一数据精度时各自使用所述图形数据以所述第一数据精度执行指令；以及

选择所述处理器内的第二组一个或一个以上执行单元中的一者，所述第二组一个或一个以上执行单元在所述所指示数据精度为第二数据精度时各自使用所述图形数据以所述第二数据精度执行指令，所述第二数据精度不同于所述第一数据精度。

14. 根据权利要求 13 所述的计算机可读媒体，其中所述第一数据精度包含全数据精度，且其中所述第二数据精度包含半数据精度。

15. 根据权利要求 11 所述的计算机可读媒体，其中所述执行单元包括所述处理器内的

第一组一个或一个以上执行单元，所述第一组一个或一个以上执行单元各自使用所

述图形数据以第一数据精度执行指令，且所述执行单元进一步包括所述处理器内的第二组一个或一个以上执行单元，所述第二组一个或一个以上执行单元各自使用所述图形数据以不同于所述第一数据精度的第二数据精度执行指令。

16. 根据权利要求 15 所述的计算机可读媒体，其中：

所述用于使所述处理器基于所述所指示数据精度来选择所述处理器内的所述多个执行单元中的一者的指令包含用于使所述处理器选择所述第一组中的所述执行单元中的一者的指令；且

所述用于使所述处理器使用所述选定执行单元来执行所述指令的指令包含用于使所述处理器使用所述第一组中的所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令的指令。

17. 根据权利要求 16 所述的计算机可读媒体，其进一步包含用于使所述处理器进行以下操作的指令：

接收供在所述处理器内执行的第二图形指令；

接收对所述第二图形指令的执行的所述第二数据精度的指示；

接收第二转换指令，所述第二转换指令在由所述处理器执行时将与所述第二图形指令相关联的图形数据转换到所述所指示的第二数据精度，所述第二转换指令不同于所述第二图形指令；

基于所述所指示的第二数据精度来选择所述第二组中的所述执行单元中的一者；以及

使用所述第二组中的所述选定执行单元来使用与所述第二图形指令相关联的所述图形数据以所述所指示的第二数据精度执行所述第二图形指令。

18. 根据权利要求 11 所述的计算机可读媒体，其中所述用于使所述处理器接收对所述图形指令的执行的所述数据精度的所述指示的指令包含用于使所述处理器对所述图形指令进行解码以确定所述数据精度的指令。

19. 根据权利要求 11 所述的计算机可读媒体，其中所述用于使所述处理器使用所述选定执行单元来使用所述图形数据以所述所指示数据精度执行所述图形指令的指令

包含用于使所述处理器进行以下操作的指令：

将所述图形数据转换到所述所指示数据精度；以及

使用所述选定执行单元来使用所述经转换图形数据以所述所指示数据精度执行所述图形指令。

20. 根据权利要求 11 所述的计算机可读媒体，其中与所述图形指令相关联的所述图形数据包含顶点图形数据及像素图形数据中的至少一者。

21. 一种装置，其包含：

控制器，其经配置以：接收供在可编程串流处理器内执行的图形指令；接收对所述图形指令的执行的数据精度的指示；以及接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所述所指示数据精度，其中所述转换指令不同于所述图形指令；以及

所述处理器内的多个执行单元，

其中所述控制器经配置以基于所述所指示的数据精度来选择所述执行单元中的一者

且使所述选定执行单元使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令。

22. 根据权利要求 21 所述的装置，其中所述多个执行单元包括经配置而以在所述所指示数据精度执行指令的第一执行单元及经配置而以不同于所述所指示数据精度的第二数据精度执行指令的第二执行单元，且其中所述控制器经配置以选择所述第一执行单元来使用所述图形数据以所述所指示数据精度执行所述图形指令。

23. 根据权利要求 21 所述的装置，其中所述多个执行单元包括一个或一个以上全精度执行单元及至少四个半精度执行单元。

24. 根据权利要求 23 所述的装置，其中当所述图形指令的执行的所述所指示数据精度包含全精度时，所述控制器经配置以选择所述全精度执行单元中的一者来使用所述图形数据执行所述图形指令。

25. 根据权利要求 23 所述的装置，其中当所述图形指令的执行的所述所指示数据精度包含半精度时，所述控制器经配置以选择所述半精度执行单元中的一者来使用所述图形数据执行所述图形指令。

26. 根据权利要求 23 所述的装置，其进一步包含：

至少一个全精度寄存器组，其用以在所述至少一个全精度执行单元执行指令时存储计算结果；以及

至少四个半精度寄存器组，其用以在所述至少四个半精度执行单元执行指令时存储计算结果。

27. 根据权利要求 21 所述的装置，其中所述多个执行单元包括至少一个全精度执行单元及至少一个半精度执行单元，且其中当所述图形指令的执行的所述所指示数据精度包含半精度时，所述控制器经配置以切断对所述至少一个全精度执行单元的供电且使所述至少一个半精度执行单元使用所述图形数据来执行所述图形指令。

28. 根据权利要求 21 所述的装置，其中所述处理器包含着色器处理器。

29. 根据权利要求 21 所述的装置，其中所述装置包含无线通信装置手持机。

30. 根据权利要求 21 所述的装置，其中所述装置包含一个或一个以上集成电路装置。

31. 一种装置，其包含：

用于接收供在可编程串流处理器内执行的图形指令的装置；

用于接收对所述图形指令的执行的数据精度的指示的装置；

用于接收转换指令的装置，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所述所指示数据精度，所述转换指令不同于所述图形指令；

用于基于所述所指示数据精度来选择所述处理器内的多个执行单元中的一者的装置；以及

用于使用所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令的装置。

32. 根据权利要求 31 所述的装置，其进一步包含：

用于接收与所述图形指令相关联的所述图形数据的装置；

用于在所述选定执行单元执行所述图形指令期间产生具有所述所指示数据精度的计算结果的装置；以及

用于将所述计算结果作为输出来提供的装置。

33. 根据权利要求 31 所述的装置，其中所述用于选择所述多个执行单元中的一者的装置包含：

用于选择所述处理器内的第一组一个或一个以上执行单元中的一者的装置，所述第一组一个或一个以上执行单元在所述所指示数据精度为第一数据精度时各自使用所述图形数据以所述第一数据精度执行指令；以及

用于选择所述处理器内的第二组一个或一个以上执行单元中的一者的装置，所述第二组一个或一个以上执行单元在所述所指示数据精度为第二数据精度时各自使用所述图形数据以所述第二数据精度执行指令，所述第二数据精度不同于所述第一数据精度。

34. 根据权利要求 33 所述的装置，其中所述第一数据精度包含全数据精度，且其中所述第二数据精度包含半数据精度。

35. 根据权利要求 31 所述的装置，其中所述执行单元包括所述处理器内的第一组一个或一个以上执行单元，所述第一组一个或一个以上执行单元各自使用所述图形数据以第一数据精度执行指令，且所述执行单元进一步包括所述处理器内的第二组一个或一个以上执行单元，所述第二组一个或一个以上执行单元各自使用所述图形数据以不同于所述第一数据精度的第二数据精度执行指令。

36. 根据权利要求 35 所述的装置，其中：

所述用于基于所述所指示数据精度来选择所述处理器内的所述多个执行单元中的一者的装置包含用于选择所述第一组中的所述执行单元中的一者的装置；且

所述用于使用所述选定执行单元来执行所述图形指令的装置包含用于使用所述第一组中的所述选定执行单元来使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令的装置。

37. 根据权利要求 36 所述的装置，其进一步包含：

用于接收供在所述处理器内执行的第二图形指令的装置；

用于接收对所述第二图形指令的执行的所述第二数据精度的指示的装置；

用于接收第二转换指令的装置，所述第二转换指令在由所述处理器执行时将与所述第二图形指令相关联的图形数据转换到所述所指示的第二数据精度，所述第二转换指令不同于所述第二图形指令；

用于基于所述所指示的第二数据精度来选择所述第二组中的所述执行单元中的一者的装置；以及

用于使用所述第二组中的所述选定执行单元来使用与所述第二图形指令相关联的所述图形数据以所述所指示的第二数据精度执行所述第二图形指令的装置。

38. 根据权利要求 31 所述的装置，其中所述用于接收对所述图形指令的执行的所述数据精度的所述指示的装置包含用于对所述图形指令进行解码以确定所述数据精度的装置。

39. 根据权利要求 31 所述的装置，其中所述用于使用所述选定执行单元来使用所述图形数据以所述所指示数据精度执行所述图形指令的装置包含：

用于将所述图形数据转换到所述所指示数据精度的装置；以及

用于使用所述选定执行单元来使用所述经转换图形数据以所述所指示数据精度执行

所述图形指令的装置。

40. 根据权利要求31所述的装置，其中与所述图形指令相关联的所述图形数据包含顶点图形数据及像素图形数据中的至少一者。

41. 一种装置，其包含：

可编程串流处理器；以及

至少一个存储器模块，其耦合到所述处理器，

其中所述处理器包含：

控制器，其经配置以：从所述至少一个存储器模块接收供执行的图形指令；接收对所述图形指令的执行的数据精度的指示；以及接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所述所指示数据精度，其中所述转换指令不同于所述图形指令；以及

多个执行单元，其经配置以执行指令，

其中所述控制器经配置以基于所述所指示的数据精度来选择所述执行单元中的一者，且使所述选定执行单元使用与所述图形指令相关联的所述图形数据以所述所指示数据精度执行所述图形指令。

42. 根据权利要求41所述的装置，其进一步包含耦合到所述处理器的至少一个图形引擎。

43. 根据权利要求41所述的装置，其中所述多个执行单元包括经配置而以所述所指示数据精度执行指令的第一执行单元及经配置而以不同于所述所指示数据精度的第二数据精度执行指令的第二执行单元，且其中所述控制器经配置以选择所述第一执行单元来使用所述图形数据以所述所指示数据精度执行所述图形指令。

44. 根据权利要求41所述的装置，其中所述多个执行单元包括一个或一个以上全精度执行单元及至少四个半精度执行单元。

45. 根据权利要求44所述的装置，其中当所述图形指令的执行的所述所指示数据精度包含全精度时，所述控制器经配置以选择所述全精度执行单元中的一者来使用所述图形数据执行所述图形指令。

46. 根据权利要求44所述的装置，其中当所述图形指令的执行的所述所指示数据精度包含半精度时，所述控制器经配置以选择所述半精度执行单元中的一者来使用所述图形数据执行所述图形指令。

47. 根据权利要求44所述的装置，其中所述处理器进一步包含：

至少一个全精度寄存器组，其用以在所述至少一个全精度执行单元执行指令时存储计算结果；以及

至少四个半精度寄存器组，其用以在所述至少四个半精度执行单元执行指令时存储计算结果。

48. 根据权利要求41所述的装置，其中所述多个执行单元包括至少一个全精度执行单元及至少一个半精度执行单元，且其中当所述图形指令的执行的所述所指示数据精度包含半精度时，所述控制器经配置以切断对所述至少一个全精度执行单元的供电且使所述至少一个半精度执行单元使用所述图形数据来执行所述图形指令。

49. 根据权利要求41所述的装置，其中所述处理器包含着色器处理器。

50. 根据权利要求 41 所述的装置，其中所述装置包含无线通信装置手持机。

51. 根据权利要求 41 所述的装置，其中所述装置包含一个或一个以上集成电路装置。

52. 一种方法，其包含：

分析图形应用程序的多个应用程序指令；

对于指定其执行的第一数据精度级别的每一应用程序指令，产生各自指示其执行的所述第一数据精度级别的一个或一个以上对应经编译指令；以及

产生一个或一个以上转换指令以在所述一个或一个以上经编译指令被执行时将图形数据从第二不同数据精度级别转换到所述第一数据精度级别。

53. 根据权利要求 52 所述的方法，其中所述第一数据精度级别包含全数据精度级别，且其中所述第二数据精度级别包含半数据精度级别。

54. 根据权利要求 52 所述的方法，其中产生所述一个或一个以上经编译指令包含产生在对应应用程序指令指定其执行的全数据精度级别时各自指示所述全数据精度级别的一个或一个以上经编译指令。

55. 根据权利要求 52 所述的方法，其中产生所述一个或一个以上经编译指令包含产生在对应应用程序指令指定其执行的半数据精度级别时各自指示所述半数据精度级别的一个或一个以上经编译指令。

56. 根据权利要求 52 所述的方法，其中所述一个或一个以上经编译指令各自包括预定义字段，所述预定义字段包括在所述对应应用程序指令指定其执行的所述第一数据精度级别时指示所述第一数据精度级别的信息。

57. 根据权利要求 52 所述的方法，其进一步包含将所述一个或一个以上经编译指令存储于存储器中以供后续执行。

58. 一种计算机可读媒体，其包含用于使处理器进行以下操作的指令：

分析图形应用程序的多个应用程序指令；

对于指定其执行的第一数据精度级别的每一应用程序指令，产生各自指示其执行的所述第一数据精度级别的一个或一个以上对应经编译指令；以及

产生一个或一个以上转换指令以在所述一个或一个以上经编译指令被执行时将图形数据从第二不同数据精度级别转换到所述第一数据精度级别。

59. 根据权利要求 58 所述的计算机可读媒体，其中所述第一数据精度级别包含全数据精度级别，且其中所述第二数据精度级别包含半数据精度级别。

60. 根据权利要求 58 所述的计算机可读媒体，其中所述用于使所述处理器产生所述一个或一个以上经编译指令的指令包含用于使所述处理器产生在对应应用程序指令指定其执行的全数据精度级别时各自指示所述全数据精度级别的所述一个或一个以上经编译指令的指令。

61. 根据权利要求 58 所述的计算机可读媒体，其中所述用于使所述处理器产生所述一个或一个以上经编译指令的指令包含用于使所述处理器产生在对应应用程序指令指定其执行的半数据精度级别时各自指示所述半数据精度级别的所述一个或一个以上经编译指令的指令。

62. 根据权利要求 58 所述的计算机可读媒体，其中所述一个或一个以上经编译指令各自包括预定义字段，所述预定义字段包括在所述对应应用程序指令指定其执行的所述第

一数据精度级别时指示所述第一数据精度级别的信息。

63. 根据权利要求 58 所述的计算机可读媒体，其进一步包含用于使所述处理器将所述一个或一个以上经编译指令存储于存储器中以供后续执行的指令。

64. 一种设备，其包含：

用于分析多个图形应用程序指令的装置；

对于指定其执行的第一数据精度级别的每一图形应用程序指令，用于产生各自指示其执行的所述第一数据精度级别的一个或一个以上对应经编译指令的装置；以及

用于产生一个或一个以上转换指令以在所述一个或一个以上经编译指令被执行时将图形数据从第二不同数据精度级别转换到所述第一数据精度级别的装置。

65. 根据权利要求 64 所述的设备，其中所述第一数据精度级别包含全数据精度级别，且其中所述第二数据精度级别包含半数据精度级别。

66. 根据权利要求 64 所述的设备，其中所述用于产生所述一个或一个以上经编译指令的装置包含用于产生在对应图形应用程序指令指定其执行的全数据精度级别时各自指示所述全数据精度级别的所述一个或一个以上经编译指令的装置。

67. 根据权利要求 64 所述的设备，其中所述用于产生所述一个或一个以上经编译指令的装置包含用于产生在对应图形应用程序指令指定其执行的半数据精度级别时各自指示所述半数据精度级别的所述一个或一个以上经编译指令的装置。

68. 根据权利要求 64 所述的设备，其中所述一个或一个以上经编译指令各自包括预定义字段，所述预定义字段包括在所述对应图形应用程序指令指定其执行的所述第一数据精度级别时指示所述第一数据精度级别的信息。

69. 根据权利要求 64 所述的设备，其进一步包含用于将所述一个或一个以上经编译指令存储于存储器中以供后续执行的装置。

70. 一种计算机可读数据存储媒体，其包含：

一个或一个以上第一可执行指令，其在由可编程串流处理器执行时支持图形应用程序的一个或一个以上功能，其中所述第一可执行指令中的每一者指示其执行的第一数据精度级别；

一个或一个以上第二可执行指令，其在由所述处理器执行时支持所述图形应用程序的一个或一个以上功能，其中所述第二可执行指令中的每一者指示其执行的不同于所述第一数据精度级别的第二数据精度级别；以及

一个或一个以上第三可执行指令，其在由所述处理器执行时支持所述图形应用程序的一个或一个以上功能，其中所述第三可执行指令中的每一者在所述一个或一个以上第一可执行指令被执行时将图形数据从所述第二数据精度级别转换到所述第一数据精度级别。

71. 根据权利要求 70 所述的计算机可读数据存储媒体，其中所述第一数据精度级别包含全数据精度级别，且其中所述第二数据精度级别包含半数据精度级别。

## 具有混合精度指令执行的可编程串流处理器

### 技术领域

[0001] 本发明涉及图形处理，且更特定来说，涉及图形处理器架构。

### 背景技术

[0002] 图形装置广泛用于渲染用于例如视频游戏、图形程序、计算机辅助设计 (CAD) 应用程序、模拟及视觉化工具、成像及其类似者等各种应用的二维 (2-D) 及三维 (3-D) 图像。图形装置可执行各种图形操作以渲染图像。所述图形操作可包括光栅化、模板及深度测试、纹理映射、着色 (shading) 及其类似者。3-D 图像可用若干表面来建模，且可用多边形（例如，三角形）来近似每一表面。为渲染目的而用于表示 3-D 图像的三角形的数目视所述表面的复杂性以及所述图像的所要分辨率而定。

[0003] 每一三角形可由三个顶点来界定，且每一顶点与各种属性（例如，空间坐标、色值及纹理坐标）相关联。当图形装置在渲染过程期间使用顶点处理器时，所述顶点处理器可处理所述各个三角形的顶点。每一三角形还由图片元素（像素）构成。当所述图形装置还（或单独地）在渲染过程期间使用像素处理器时，所述像素处理器通过确定每一三角形内的每一像素的分量的值来渲染所述三角形。

[0004] 在许多情况下，图形装置可利用着色器处理器来执行某些图形操作，例如着色。着色为涉及照明及阴影的高复杂的图形操作。当执行渲染时，所述着色器处理器可能需要执行多种不同指令，且通常包括一个或一个以上执行单元以辅助这些指令的执行。举例来说，所述着色器处理器可包括算术逻辑单元 (ALU) 和 / 或初等函数单元 (EFU) 作为执行单元。常常，这些执行单元能够使用全数据精度电路来执行指令。然而，此电路可常常要求更多功率，且所述执行单元可占据由所述图形装置所使用的着色器处理器集成电路内的更多物理空间。

### 发明内容

[0005] 大体上，本发明涉及图形装置的可编程串流处理器，所述可编程串流处理器能够使用不同执行单元来执行混合精度（例如，全精度、半精度）指令。举例来说，所述可编程处理器可包括一个或一个以上全精度执行单元以及一个或一个以上半精度执行单元。在接收到二进制指令及对所述指令的执行的数据精度的指示后，所述处理器即能够选择适当执行单元以用于在所指示的数据精度下执行接收到的指令。所述处理器可包含用于移动图形应用程序的基于指令的适应性串流处理器。

[0006] 通过此做法，所述处理器可避免使用一个执行单元来执行具有各种不同数据精度的指令。结果，可减少或消除不必要的精度提升。另外，应用程序程序员在编写应用程序代码时可具有增大的灵活性。应用程序程序员可为不同应用程序指令指定不同数据精度级别，所述不同数据精度级别接着经编译成由所述处理器处理的一个或一个以上二进制指令。

[0007] 在一个方面中，本发明是针对一种方法，其包括接收供在可编程串流处理器内

执行的图形指令；接收对所述图形指令的执行的数据精度的指示；以及接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所指示的数据精度，其中所述转换指令不同于所述图形指令。所述方法进一步包括基于所述所指示的数据精度来选择所述处理器内的多个执行单元中的一者，以及使用所述选定执行单元以使用与所述图形指令相关联的图形数据来在所述所指示的数据精度下执行所述图形指令。

[0008] 在一个方面中，本发明是针对一种计算机可读媒体，其包括用于使可编程串流处理器进行以下操作的指令：接收供在所述处理器内执行的图形指令；接收对所述图形指令的执行的数据精度的指示；以及接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所指示的数据精度，其中所述转换指令不同于所述图形指令。所述计算机可读媒体进一步包括用于使所述处理器进行以下操作的指令：基于所述所指示的数据精度来选择所述处理器内的多个执行单元中的一者，以及使用所述选定执行单元以使用与所述图形指令相关联的图形数据来在所述所指示的数据精度下执行所述图形指令。

[0009] 在一个方面中，本发明是针对一种可编程串流处理器，其包括一控制器及多个执行单元。所述控制器经配置以接收供执行的图形指令及接收对所述图形指令的执行的数据精度的指示。所述控制器还经配置以接收转换指令，所述转换指令在由所述处理器执行时将与所述图形指令相关联的图形数据转换到所指示数据精度，其中所述转换指令不同于所述图形指令。当可操作时，所述控制器基于所指示的数据精度来选择所述执行单元中的一者。所述控制器接着使所述选定执行单元使用与所述图形指令相关联的图形数据在所指示数据精度下执行所述图形指令。

[0010] 在另一方面中，本发明是针对一种计算机可读媒体，其包括用于使处理器进行以下操作的指令：分析图形应用程序的多个应用程序指令；以及对于指定执行的第一数据精度级别的每一应用程序指令，产生各自指示执行的所述第一数据精度级别的一个或一个以上对应经编译指令。所述计算机可读媒体包括用于使所述处理器进行以下操作的其它指令：产生一个或一个以上转换指令以在所述一个或一个以上经编译指令经执行时将图形数据从第二不同数据精度级别转换到所述第一数据精度级别。

[0011] 在一个方面中，本发明是针对一种计算机可读数据存储媒体，其具有一个或一个以上第一可执行指令，所述一个或一个以上第一可执行指令在由可编程串流处理器执行时支持图形应用程序的一个或一个以上功能，其中所述第一可执行指令中的每一者指示其执行的第一数据精度级别。所述计算机可读数据存储媒体进一步包括一个或一个以上第二可执行指令，所述一个或一个以上第二可执行指令在由所述处理器执行时支持所述图形应用程序的一个或一个以上功能，其中所述第二可执行指令中的每一者指示其执行的不同于所述第一数据精度级别的第二数据精度级别。所述计算机可读数据存储媒体进一步包括一个或一个以上第三可执行指令，所述一个或一个以上第三可执行指令在由所述处理器执行时支持所述图形应用程序的一个或一个以上功能，其中所述第三可执行指令中的每一者在所述一个或一个以上第一可执行指令经执行时将图形数据从所述第二数据精度级别转换到所述第一数据精度级别。

[0012] 在随附图式及以下描述中阐述本发明的一个或一个以上方面的细节。其它特

征、目标及优点将从描述及图式以及从权利要求书而显而易见。

## 附图说明

- [0013] 图 1 为根据本发明的一方面说明可包括于图形处理系统内的各种组件的框图。
- [0014] 图 2A 为根据本发明的一方面说明包括可编程着色器处理器的示范性图形处理系统的框图。
- [0015] 图 2B 为根据本发明的一方面说明图 2A 中所示的着色器处理器的其它细节的框图。
- [0016] 图 2C 为根据本发明的一方面说明图 2B 中所示的执行单元及寄存器组的其它细节的框图。
- [0017] 图 3 为根据本发明的一方面说明可由图 2A 到图 2B 中所示的着色器处理器执行的示范性方法的流程图。
- [0018] 图 4 为根据本发明的一方面说明可用以产生待由图 1 中所示的串流处理器或由图 2A 到图 2B 中所示的着色器处理器执行的图形指令的编译器的框图。

## 具体实施方式

- [0019] 图 1 为根据本发明的一个方面说明可包括于图形处理系统内的各种组件的框图。此图形处理系统可为独立系统或可为例如计算系统或无线通信装置（例如，无线通信装置手持机）等较大系统的部分，或数码相机或其它视频装置的部分。图 1 中所示的示范性系统可包括一个或一个以上图形应用程序 102A 到 102N、图形装置 100，及外部存储器 104。图形装置 100 可通信地耦合到外部存储器 104 及图形应用程序 102A 到 102N 中的每一者。在一个方面中，图形装置 100 可包括于一个或一个以上集成电路或芯片上。
- [0020] 图形应用程序 102A 到 102N 可包括各种不同应用程序，例如视频游戏、视频、相机，或其它图形或串流应用程序。这些图形应用程序 102A 到 102N 可同时运行且各自能够产生执行线程以实现所要结果。线程指示可用一个或一个以上图形指令的序列来执行的特定任务。线程允许图形应用程序 102A 到 102N 具有同时执行的多个任务且共享资源。

- [0021] 图形装置 100 从图形应用程序 102A 到 102N 接收线程且执行由这些线程指示的任务。在图 1 中所示的方面中，图形装置 100 包括可编程串流处理器 106、一个或一个以上图形引擎 108A 到 108N，及一个或一个以上存储器模块 110A 到 110N。处理器 106 可执行各种图形操作（例如，着色），且可计算某些应用程序的超越初等函数。在一个方面中，处理器 106 可包含用于移动图形应用程序的基于指令的适应性串流处理器。图形引擎 108A 到 108N 可执行其它图形操作（例如，纹理映射）。存储器模块 110A 到 110N 可包括一个或一个以上高速缓冲存储器以存储用于处理器 106 及图形引擎 108A 到 108N 的数据及图形指令。

- [0022] 图形引擎 108A 到 108N 可包括执行各种图形操作（例如，三角形设置、光栅化、模板及深度测试、属性设置和 / 或像素内插）的一个或一个以上引擎。外部存储器 104 可相对于存储器模块 110A 到 110N 为大、较缓慢的存储器。在一个方面中，外部存储器 104 位于距图形装置 100 较远处（例如，芯片外）。外部存储器 104 存储可加载到存储器

模块 110A 到 110N 中的一者或一者以上中的数据及图形指令。

[0023] 在一个方面中，处理器 106 能够使用不同执行单元来执行混合精度（例如，全精度、半精度）图形指令，假定不同图形应用程序 102A 到 102N 可具有关于 ALU 精度、性能及输入 / 输出格式的不同要求。作为一实例，处理器 106 可包括一个或一个以上全精度执行单元以及一个或一个以上部分精度执行单元。所述部分精度执行单元可为（例如）半精度执行单元。处理器 106 可使用其执行单元来执行图形应用程序 102A 到 102N 中的一者或一者以上的图形指令。在接收到二进制指令（例如，来自外部存储器 104 或存储器模块 110A 到 110N 中的一者）且还接收到对所述图形指令的执行的数据精度的指示后，处理器 106 即可选择适当执行单元以用于使用图形数据来在所指示的数据精度下执行所接收到的指令。处理器 106 也可接收一单独转换指令，所述单独转换指令在经执行时将与所述图形指令相关联的图形数据转换到所指示的数据精度。在一个方面中，所述转换指令为不同于所述图形指令的单独指令。

[0024] 图形数据可由图形应用程序 102A 到 102N 提供，或可从外部存储器 104 或存储器模块 110A 到 110N 中的一者检索，或可由图形引擎 108A 到 108N 中的一者或一者以上提供。通过基于所指示的数据精度在不同执行单元中选择性地执行指令，处理器 106 可避免使用单一执行单元来执行全精度指令及半精度指令两者。另外，图形应用程序 102A 到 102N 的程序员在编写应用程序代码时可具有增大的灵活性。举例来说，一应用程序程序员可为应用程序指令指定数据精度级别，所述数据精度级别接着经编译成通过处理器 106 处理的一个或一个以上二进制指令。处理器 106 基于与执行单元及二进制指令相关联的数据精度来选择适当执行单元来执行所述二进制指令。另外，必要时，处理器 106 可执行所接收到的转换指令以将与所述指令相关联的图形数据转换到所指示的数据精度。举例来说，如果所提供的图形数据具有不同于所指示的数据精度的数据精度，则处理器 106 可执行所述转换指令以将所述图形数据转换到所指示的数据精度，使得图形指令可通过所选定的执行单元执行。

[0025] 图 2A 为根据一个方面说明包括可编程着色器处理器 206 的示范性图形处理系统的框图。在此方面中，图 2A 中所示的图形处理系统为图 1 中所示的更普通系统的示范性例示。在一个方面中，着色器处理器 206 为串流处理器。在图 2A 中，所述示范性系统包括各自通信地耦合到图形装置 200 的两个图形应用程序 202A 及 202B。在图 2A 的实例中，图形应用程序 202A 为能够处理及管理图形成像像素数据的像素应用程序。在图 2A 的实例中，图形应用程序 202B 为能够处理及管理图形成像顶点数据的顶点应用程序。在一个方面中，图形像素应用程序 202A 包含像素处理应用程序，且图形顶点应用程序 202B 包含顶点处理应用程序。

[0026] 在许多情况下，图形像素应用程序 202A 实施使用较低精度（例如，半精度）数据格式的许多功能，但其可实施使用较高精度（例如，全精度）数据格式的某些功能。图形像素应用程序 202A 也可指定像素数据的指令的以四元组为基础的执行。通常，图形顶点应用程序 202B 实施使用较高精度数据格式的功能，但可能不指定顶点数据的指令的以四元组为基础的执行。因此，不同应用程序（例如，应用程序 202A 及 202B）及到图形装置 200 的对应 API 可指定不同数据精度要求。且，在给定应用程序 202A 或 202B（及对应 API）内，可指定混合精度指令的执行。举例来说，用于图形像素应用程序 202A 的

着色语言可为待由着色器处理器 206 执行的着色器指令提供精度修改器。因此，某些指令可指定执行的一个精度级别，而其它指令可指定另一精度级别。图形装置 200 内的着色器处理器 206 能够以均匀方式来执行混合精度指令。

[0027] 在一个方面中，着色器处理器 206 经由一个或一个以上应用程序接口或 API (未图示) 与图形应用程序 202A 及 202B 相互作用。举例来说，图形像素应用程序 202A 可经由第一 API 与着色器处理器 206 相互作用，且图形顶点应用程序 202B 可经由第二 API 与着色器处理器 206 相互作用。在一个方面中，第一 API 及第二 API 可包含一共同 API。所述 API 可界定由图形应用程序 202A 及 202B 用以使图形装置 200 执行各种图形操作 (包括可由着色器处理器 206 执行的着色操作) 的一个或一个以上标准编程规范。

[0028] 图形装置 200 包括着色器处理器 206。着色器处理器 206 能够执行着色操作。着色器处理器 206 能够与图形像素应用程序 202A 交换像素数据，且进一步能够与图形顶点应用程序 202B 交换顶点数据。

[0029] 在图 2A 的实例中，着色器处理器 206 还与纹理引擎 208 及高速缓存存储器系统 210 通信。纹理引擎 208 能够执行纹理相关操作，且还通信地耦合到高速缓存存储器系统 210。高速缓存存储器系统 210 耦合到主存储器 204。在一方面中，高速缓存存储器系统 210 包括指令高速缓冲存储器及数据高速缓冲存储器两者。指令和 / 或数据可从主存储器 204 加载到高速缓存存储器系统 210 中，接着使指令和 / 或数据可用于纹理引擎 208 及着色器处理器 206。着色器处理器 206 可经由同步或非同步接口与外部装置或组件通信。

[0030] 在一个方面中，着色器处理器 206 能够使用不同执行单元来执行混合精度图形指令。在此方面中，着色器处理器 206 包括一个或一个以上全精度执行单元以及一个或一个以上半精度执行单元。着色器处理器 206 可调用其执行单元以执行图形应用程序 202A 及 202B 中的一者或两者的图形指令。在接收到二进制指令 (例如，来自高速缓存存储器系统 210) 且还接收到对所述指令的执行的数据精度的指示后，着色器处理器 206 即能够选择一适当执行单元以用于使用图形数据来在所指示的数据精度下执行所接收到的指令。图形像素应用程序 202A 可将 (例如) 像素数据提供到着色器处理器 206，且图形顶点应用程序 202B 可将顶点数据提供到着色器处理器 206。

[0031] 着色器处理器也可接收一单独转换指令，所述单独转换指令在经执行时将与所述图形指令相关联的图形数据转换到所指示的数据精度。在一个方面中，所述转换指令为不同于所述图形指令的单独指令。

[0032] 图形数据也可从主存储器 204 或高速缓存存储器系统 210 加载，或可由纹理引擎 208 提供。图形像素应用程序 202A 和 / 或图形顶点应用程序 202B 调用使着色器处理器 206 从高速缓存存储器系统 210 加载供执行的一个或一个以上二进制指令的执行线程。在一个方面中，每一经加载指令指示所述指令的执行的数据精度。另外，必要时，着色器处理器 206 可执行所接收到的转换指令以将与所述指令相关联的图形数据转换到所指示的数据精度。举例来说，如果所提供的图形数据具有不同于所指示的数据精度的数据精度，则着色器处理器 206 可执行所述转换指令以将所述图形数据转换到所指示的数据精度，使得图形指令可通过所选定的执行单元执行。通过基于所指示的数据精度在不同执行单元中选择性地执行指令，着色器处理器 206 可避免使用单一执行单元来执行全精度指令及半精度指令两者。

[0033] 图 2B 为根据一个方面说明图 2A 中所示的着色器处理器 206 的其它细节的框图。在着色器处理器 206 内，定序器 222 从图形应用程序 202A 及 202B 接收线程，且将这些线程提供到线程调度器与现场寄存器 224。在一个方面中，定序器 222 包含一多路复用器 (MUX)。在一个方面中，定序器 222 确定应接受哪些线程，且也可为每一所接受的线程分配多重精度寄存器空间和 / 或其它资源。举例来说，定序器 222 可为半精度指令分配寄存器空间，且也可为全精度指令分配寄存器空间。

[0034] 在一个方面中，从图形像素应用程序 202A 接收到的像素数据包括呈以像素四元组为基础的格式（即，每次四个像素）的属性信息。在此方面中，执行单元 234 每次可处理四个像素。在一个方面中，执行单元 234 可每次一个顶点地处理来自图形顶点应用程序 202B 的数据。

[0035] 线程调度器 224 执行各种功能以对线程的执行进行调度及管理，且可控制线程的执行顺序。根据一个方面，对于每一线程，线程调度器 224 可确定所述线程所需的资源是否准备好，在用于所述线程的任何资源（例如，指令、寄存器堆或纹理读取）未准备好时将所述线程推到一休眠队列中，及在所述资源全部准备好时将所述线程从所述休眠队列移动到一活动队列。线程调度器 224 与加载控制单元 226 介接以便同步用于所述线程的资源。在一个方面中，线程调度器 224 为控制器 225 的部分。图 2B 展示控制器 225 的一实例。控制器 225 可控制与着色器处理器 206 内的指令及数据的处理有关的各种功能。在图 2B 的实例中，控制器 225 包括线程调度器 224、加载控制单元 226 及主引擎 220。在某些方面中，控制器 225 包括主引擎 220、线程调度器 224 及加载控制单元 226 中的至少一者。

[0036] 线程调度器 224 还管理线程的执行。线程调度器 224 从指令高速缓冲存储器 230 取得每一线程的指令、在必要时对每一指令进行解码，及执行对所述线程的流控制。线程调度器 224 选择供执行的活动线程、检查所述选定线程当中的读取 / 写入端口冲突，且在不存在冲突时将用于一个线程的指令发送到执行单元 234，及将用于另一线程的指令发送到加载控制单元 226。线程调度器 224 维持用于每一线程的程序 / 指令计数器，且随着执行指令或更改程序流而更新此计数器。线程调度器 224 还发出从指令高速缓冲存储器 230 取得缺失指令且移除所完成的线程的请求。

[0037] 在一个方面中，线程调度器 224 与主引擎 220 相互作用。在此方面中，线程调度器 224 可将某些职责委托给主引擎 220。在一个方面中，线程调度器 224 可对供执行的指令进行解码，或可维持用于每一线程的程序 / 指令计数器且随着执行指令而更新此计数器。在一个方面中，主引擎 220 设置指令执行的状态，且也可在指令执行期间控制状态更新序列。

[0038] 指令高速缓冲存储器 230 存储用于所述线程的指令。这些指令指示待为每一线程执行的特定操作。每一操作可为（例如）算术运算、初等函数、存储器存取操作或另一形式的指令。在需要时，指令高速缓冲存储器 230 可经由加载控制单元 226 来加载有来自高速缓存存储器系统 210 或主存储器 204（图 2A）的指令。根据一个方面，这些指令为已从图形应用程序代码编译的二进制指令。每一二进制指令指示用于其在着色器处理器 206 内的执行的数据精度。举例来说，与所述指令相关联的指令类型可指示所述指令是全精度指令还是半精度指令。或，根据一个示范性方面，所述指令内的一特定旗标或

字段可指示其是全精度指令还是半精度指令。线程调度器 224 可能能够对指令进行解码且确定每一指令的数据精度（例如，全精度或半精度）。线程调度器 224 可接着将每一指令路由到一能够在所指示数据精度下执行指令的执行单元。此执行单元从常数缓冲器 232 或寄存器组 242 加载指令执行所需的任何图形数据，在下文对常数缓冲器 232 或寄存器组 242 进行更详细描述。

[0039] 在图 2B 中所示的方面中，执行单元 234 包括一个或一个以上全精度 ALU（算术逻辑单元）236、一个或一个以上半精度 ALU 240，及一执行超越初等函数单元 238。ALU 236 及 240 可包括一个或一个以上浮点单元（其实现浮点计算）和 / 或一个或一个以上整数逻辑单元（其实现整数及逻辑运算）。在必要时，执行单元 234 在指令执行期间从常数缓冲器 232 或从寄存器组 242 载入数据（例如，图形数据）。全精度 ALU 236 及半精度 ALU 240 两者能够执行算术运算（例如，加法、减法、乘法、倍增及累积等）且还能够执行逻辑运算（例如，AND、OR、XOR 等）。根据一个方面，每一个 ALU 单元可包含单一四元组 ALU 或四个标量 ALU。当使用四个标量 ALU 时，可通过 ALU 并行地处理四个像素的属性。可使用一四元组 ALU 来并行地处理一像素或一顶点的四个属性。然而，全精度 ALU 236 使用全精度计算来执行指令，而半精度 ALU 240 使用半精度计算来执行指令。

[0040] 初等函数单元 238 可计算超越初等函数（例如，正弦、余弦、倒数、对数、指数、平方根或反平方根），其广泛用于着色器指令中。初等函数单元 238 可通过在比使用简单指令来执行对初等函数的多项式近似所需的时间少得多的时间中计算所述初等函数来改善着色器性能。根据本发明的一个方面，初等函数单元 238 可能能够在全精度下执行指令，但也可能能够将计算结果转换到半精度格式。

[0041] 加载控制单元 226（在图 2B 中所示的示范性方面中，其为控制器 225 的部分）控制着色器处理器 206 内的各种组件的数据及指令的流。在一个方面中，加载控制单元 226 可将着色器处理器 206 的过多内部数据逐出到外部存储器（例如，高速缓存存储器系统 210），且可从纹理引擎 208 和 / 或高速缓存存储器系统 210 取得外部资源（例如，指令、缓冲器或纹理数据）。加载控制单元 226 与高速缓存存储器系统 210 介接且使指令高速缓存存储器 230、常数缓冲器 232（其可存储在图形应用程序 202A 和 / 或 202B 的指令执行期间所使用的均匀数据）及寄存器组 242 加载有来自高速缓存存储器系统 210 的数据及指令。加载控制单元 226 也可将来自寄存器组 242 的输出数据提供到高速缓存存储器系统 210。寄存器组 242 可接收来自一个或一个以上执行单元 234 的输出数据，且可在执行单元 234 当中共享。加载控制单元 226 还与纹理引擎 208 介接。在某些情况下，纹理引擎 208 可经由加载控制单元 226 将数据（例如，纹理数据）提供到着色器处理器 206，且在某些情况下，加载控制单元 226 可将数据（例如，纹理坐标数据）和 / 或指令（例如，取样器 ID 指令）提供到纹理引擎 208。

[0042] 在图 2B 的实例中，加载控制单元 226 还包括精度转换器 228。因为写入到加载控制单元 226 中或读取出加载控制单元 226 的数据可具有不同数据精度（例如，全精度、半精度），所以加载控制单元 226 可能需要在将某些数据路由到不同组件（例如，到寄存器组 242 或到高速缓存存储器系统 210）之前将其转换到不同数据精度级别。精度转换器 228 管理在加载控制单元 226 内的此数据转换。

[0043] 在一个方面中，精度转换器 228 操作以在由着色器处理器 206 执行所接收到的转换指令后即将图形数据从一个精度级别转换到另一精度级别。当经执行时，转换指令将与一接收到的图形指令相关联的图形数据转换到所指示的数据精度。举例来说，所述转换指令可将半精度格式的数据转换到全精度格式，或将全精度格式的数据转换到半精度格式。

[0044] 常数缓冲器 232 可存储在指令执行期间由执行单元 234 使用的常数值。寄存器组 242 存储来自执行单元 234 的所执行线程的临时结果以及最终结果。寄存器组 242 包括一个或一个以上全精度寄存器组 244 及一个或一个以上半精度寄存器组 246。可通过加载控制单元 226 从寄存器组 242 读取最终执行结果。另外，分配器 248 也可从寄存器组 242 接收所执行线程的最终结果且将这些结果分配到图形顶点应用程序 202B 及图形像素应用程序 202A 中的至少一者。

[0045] 图形应用程序（例如，应用程序 202A 及 202B）可能需要使用不同精度级别来处理数据。举例来说，在一个方面中，图形顶点应用程序 202B 使用全精度数据格式来处理顶点数据，而图形像素应用程序 202A 使用半精度格式来处理像素数据。在一个方面中，图形像素应用程序 202A 使用半精度格式来处理某些信息，而使用全精度格式来处理其它信息。在来自图形顶点应用程序 202B 及图形像素应用程序 202A 的线程的执行期间，着色器处理器 206 接收及处理来自指令高速缓冲存储器 230 的针对执行使用不同数据精度级别的指令。

[0046] 因此，在图 2B 中所示的方面中，线程调度器 224 识别所指示的或与从指令高速缓冲存储器 230 加载出的给定指令相关联的数据精度，且将所述指令路由到一适当执行单元。举例来说，如果将所述指令作为全精度指令（例如，经由指令类型或所述指令内所含有的字段 / 标头来指示）来解码，则线程调度器 224 能够将所述指令路由到全精度 ALU 236 中的一者以供执行。来自全精度 ALU 236 的执行结果可存储于全精度寄存器组 244 中的一者或一者以上中且经由分配器 248 而提供回到图形应用程序（例如，图形顶点应用程序 202B）。然而，如果来自指令高速缓冲存储器 230 的指令作为半精度指令由线程调度器 224 来解码，则线程调度器 224 能够将所述指令路由到半精度 ALU 240 中的一者以供执行。来自半精度 ALU 240 的执行结果可存储于半精度寄存器组 246 中的一者或一者以上中且经由分配器 248 而提供回到图形应用程序（例如，图形像素应用程序 202A）。

[0047] 图 2C 为根据一个方面说明图 2B 中所示的执行单元 234 及寄存器组 242 的其它细节的框图。如先前所描述，执行单元 234 包括各种不同类型的执行单元。在图 2C 的实例中，执行单元 234 包括一个或一个以上全精度 ALU 236A 到 236N、一个或一个以上半精度 ALU 240A 到 240N，及一个或一个以上初等函数单元 238。每一全精度 ALU 236A 到 236N 能够使用数据来使用全精度计算来执行指令。在指令执行期间使用的输入数据可从全精度寄存器组 244A 到 244N（在寄存器组 242 内）中的一者或一者以上检索。另外，在由全精度 ALU 236A 到 236N 执行指令期间所产生的计算结果可存储于全精度寄存器组 244A 到 244N 中的一者或一者以上内。

[0048] 类似地，每一半精度 ALU 240A 到 240N 能够使用数据来使用半精度计算来执行指令。在指令执行期间使用的输入数据可从半精度寄存器组 246A 到 246N 中的一者或一者以上检索。另外，在由半精度 ALU 240A 到 240N 执行指令期间所产生的计算结果可存

储于半精度寄存器组 246A 到 246N 中的一者或一者以上内。

[0049] 如先前所描述，初等函数单元 238 能够执行全精度指令，但以半精度格式来存储结果。在一个方面中，初等函数单元 238 能够以全精度格式或半精度格式来存储结果数据。结果，初等函数单元 238 通信地耦合到全精度寄存器组 244A 到 244N，且还通信地耦合到半精度寄存器组 246A 到 246N。根据一个方面，初等函数单元 238 可从寄存器组 242 内的寄存器中的任一者检索中间数据且将最终结果数据存储到所述寄存器中的任一者。

[0050] 另外，初等函数单元 238 包括精度转换器 239。在初等函数单元 238 在全精度数据格式与半精度数据格式之间转换的那些情况下，其可使用精度转换器 239 来执行所述转换。举例来说，单元 238 可从半精度寄存器组 246A 加载输入图形数据且使用所述数据来执行全精度指令。精度转换器 239 可将所述输入数据从半精度格式转换到全精度格式。单元 238 可接着使用经转换的数据来执行全精度指令。如果结果数据将存储回到半精度寄存器组 246A 中，则精度转换器 239 可将结果数据从全精度格式转换到半精度格式，使得其可存储于半精度寄存器组 246A 中。或者，如果结果数据将存储于全精度寄存器组 244A 到 244N 中的一者中，则呈全精度格式的结果数据可直接存储于这些寄存器中的一者中。

[0051] 线程调度器 224(图 2B)能够使二进制指令从指令高速缓冲存储器 230 加载且基于与所述指令相关联的数据精度来在执行单元 234 中的一者中执行。举例来说，线程调度器 224 可将全精度指令路由到全精度 ALU 236A 到 236N 中的一者或一者以上，且可将半精度指令路由到半精度 ALU 240A 到 240N 中的一者或一者以上。线程调度器 224 也可将初等指令路由到初等函数单元 238 以供执行。结果数据可存储于寄存器组 242 内的对应寄存器中。在一个方面中，全精度 ALU 236A 到 236N、初等函数单元 238 与半精度 ALU 240A 到 240N 之间的数据转变经过寄存器组 242。

[0052] 在一个方面中，与每一全精度寄存器组 244A 到 244N 相比，每一半精度寄存器组 246A 到 246N 含有较小寄存器存储空间，且在集成电路上占据较小物理空间。因此，举例来说，与全精度寄存器组 244A 相比，半精度寄存器组 246A 含有较小寄存器存储空间且占据较小物理空间。在一个方面中，一个全精度寄存器组(例如，组 244A)可含有与两个半精度寄存器组(例如，组合的组 246A 与 246B)大体上相同量的寄存器空间，且占据与其大体上相同量的物理空间。

[0053] 类似地，每一全精度 ALU 236A 到 236N 可比每一半精度 ALU 240A 到 240N 占据集成电路内的更多物理空间。另外，每一全精度 ALU 236A 到 236N 通常可使用比每一半精度 ALU 240A 到 240N 多的操作功率。因此，在某些方面中，可能需要限制所使用的全精度 ALU 及全精度寄存器组的数目，且增大所使用的半精度 ALU 及半精度寄存器组的数目，以便最小化集成电路大小且减少功率消耗要求。当着色器处理器 206 为例如移动或无线通信装置(例如，例如移动无线电电话或无线通信装置手持机)或数码相机或视频装置等具有某些功率约束的较小计算装置的部分时，这些方面可尤其适当或有益。

[0054] 因此，在一个方面中，执行单元 234 可仅包括一个全精度 ALU 236A，且寄存器组 242 可仅包括一个全精度寄存器组 244A。在此方面中，执行单元 234 可进一步包括四个半精度 ALU 240A 到 240D，而寄存器组 242 可包括四个半精度寄存器组 246A 到 246D。

结果，执行单元 234 可能能够并行地执行至少一个半精度指令及一个全精度指令。举例来说，所述四个半精度 ALU 240A 到 240D 可每次执行关于四个像素的属性的指令。根据一个方面，因为仅使用一个全精度 ALU 236A，所以 ALU 236A 能够每次执行关于一个顶点的指令。结果，根据一个方面，着色器处理器 206 不需要利用顶点装填缓冲器以装填多个顶点的数据。在此情况下，可直接处理顶点的基于向量的属性数据，而无需将数据转换到标量格式。

[0055] 在另一方面中，执行单元 234 可包括四个全精度 ALU 236A 到 236D，且寄存器组 242 可包括四个全精度寄存器组 244A 到 244D。在此方面中，执行单元 234 可进一步包括八个半精度 ALU 240A 到 240H，而寄存器组 242 可包括八个半精度寄存器组 246A 到 246H。结果，执行单元 234 能够并行地执行（例如）两个四元组上的两个半精度指令及一个四元组上的一个全精度指令。每一四元组或线程为四个像素或四个顶点的群组。

[0056] 在另一方面中，执行单元 234 可包括四个全精度 ALU 236A 到 236D，且寄存器组 242 可包括四个全精度寄存器组 244A 到 244D。在此方面中，执行单元 234 进一步包括四个半精度 ALU 240A 到 240H，而寄存器组 242 包括四个半精度寄存器组 246A 到 246H。可使用全精度 ALU 236A 到 236N、全精度寄存器组 244A 到 244N、半精度 ALU 240A 到 240N 及半精度寄存器组 246A 到 246N 的各种其它组合。

[0057] 在一个方面中，着色器处理器 206 可能能够使用线程调度器 224 以对全精度 ALU 236A 到 236N 中的一者或一者以上及全精度寄存器组 244A 到 244N 中的一者或一者以上选择性地断电或停用。在此方面中，尽管着色器处理器 206 包括一个或一个以上集成电路内的各种全精度组件（例如，全精度 ALU 236A 到 236N 及全精度寄存器组 244A 到 244N），但其可通过在不使用这些全精度组件中的一者或一者以上时对这些全精度组件中的一者或一者以上选择性地断电或停用来节省或减少功率消耗。举例来说，在某些情形下，着色器处理器 206 可确定这些组件中的一者或一者以上未被使用，假定所加载的各种二进制指令将由半精度 ALU 240A 到 240N 中的一者或一者以上执行。因此，在这些类型的情形下，着色器处理器 206 可对所述全精度组件中的一者或一者以上选择性地断电或停用以用于功率节省。以此方式，着色器处理器 206 可依据给定时间处所处理的指令的类型及数目而以动态为基础对一个或一个以上全精度组件选择性地断电或停用。

[0058] 在一个方面中，着色器处理器 206 也可能能够使用线程调度器 224 以对半精度 ALU 240A 到 240N 中的一者或一者以上及半精度寄存器组 246A 到 246N 中的一者或一者以上选择性地断电或停用。在此方面中，着色器处理器 206 可通过在不使用或不需要这些半精度组件中的一者或一者以上时对这些半精度组件中的一者或一者以上选择性地断电或停用以节省或减少功率消耗。

[0059] 着色器处理器 206 可提供各种益处及优点。举例来说，着色器处理器 206 可提供高灵活及适应性接口以满足混合精度指令（例如，全精度指令及半精度指令）的执行的不同要求。着色器处理器 206 可通过在混合精度指令的执行期间避免不必要的精度提升来显著地减少功率消耗。（当着色器处理器 206 将数据从例如半精度格式等较低精度格式动态地转换到例如全精度格式等较高精度格式时，精度提升可发生。精度提升可能在着色器处理器 206 内需要额外电路，且也可使着色器核心处理花费额外时钟循环。）因为线程调度器 224 能够辨识与从指令高速缓冲存储器 230 加载的二进制指令相关联的数据精

度，所以线程调度器 224 能够将所述指令路由到执行单元 234 内的适当执行单元（例如，全精度 ALU 236A 或半精度 ALU 240A）以供执行。

[0060] 着色器处理器 206 也可通过利用较少全精度组件及通过替代地利用更多半精度组件（例如，ALU 及寄存器组）来减小寄存器组 242 中的总寄存器堆大小及执行单元 234 中的 ALU 大小。另外，着色器处理器 206 可通过增大处理能力来增大总系统性能。

[0061] 鉴于与较低功率消耗及增大的性能有关的各种潜在益处，着色器处理器 206 可用于各种不同类型的系统或装置中，例如无线通信装置、数码相机装置、视频记录或显示装置、视频游戏装置，或其它图形及多媒体装置。这些装置可包括一显示器以使用着色器处理器 206 来呈现所产生的图形内容。在一个方面中，由着色器处理器 206 提供的精度灵活性允许其与各种装置（包括多媒体装置）一起使用，此可提供比某些其它图形应用程序低的精度计算或具有比某些其它图形应用程序低的功率要求。

[0062] 图 3 为根据一个方面说明可通过图 2A 到图 2B 中所示的着色器处理器 206 执行的示范性方法的流程图。在此方面中，所述示范性方法包括动作 300、302、303、306、308、310 及 312，且还包括决策点 304。

[0063] 在动作 300 中，着色器处理器 206 接收二进制图形指令及对所述指令的执行的数据精度的指示。举例来说，如先前所描述，线程调度器 224 可从指令高速缓冲存储器 230（图 2B）加载指令。在一个方面中，通过线程调度器 224 对指令进行解码提供关于所述指令的执行的数据精度的信息。举例来说，所述指令可为全精度指令或半精度指令。

[0064] 在动作 302 中，着色器处理器 206 接收与所述二进制指令相关联的图形数据。举例来说，定序器 222 可从图形顶点应用程序 202B 接收顶点数据和 / 或可从图形像素应用程序 202A 接收像素数据。在某些情形下，加载控制单元 226 也可从高速缓存存储器系统 210 加载与所述指令相关联的图形数据。在动作 303 中，着色器处理器 206 进一步接收一转换指令，所述转换指令在经执行时将与所述二进制指令相关联的图形数据转换到所指示的数据精度。

[0065] 在决策点 304 处，着色器处理器 206 确定所述指令是全精度指令还是半精度指令。如上所述，在一个方面中，线程调度器 224 可对所述指令进行解码且确定其是全精度指令还是半精度指令。

[0066] 如果所述指令为全精度指令，则在动作 306 中，在必要时，着色器处理器 206 将任何接收到的图形数据从半精度格式转换到全精度格式。在某些情况下，所接收到的图形数据（如存储于高速缓存存储器系统 210 中或如从图形应用程序 202A 或 202B 处理）可具有半精度格式。在此情况下，将所述图形数据转换到全精度格式，使得其可在全精度指令的执行期间使用。在一个方面中，当接收到的转换指令由着色器处理器 206 执行时，加载控制单元 226 的精度转换器 228 可管理数据格式转换。在动作 308 中，着色器处理器 206 选择一全精度单元（例如，单元 236A（图 2C））以使用所述图形数据来执行所述二进制指令。

[0067] 然而，如果所述指令为半精度指令，则在动作 310 中，在必要时，着色器处理器将任何数据从全精度格式转换到半精度格式。在一个方面中，当接收到的转换指令由着色器处理器 206 执行时，精度转换器 228 可管理数据格式转换。在动作 312 中，着色器处理器 206 接着选择一半精度单元（例如，单元 240A（图 2C））以使用所述图形数据来

执行所述二进制指令。

[0068] 图 4 为根据一个方面说明可用以产生待由图 1 中所示的串流处理器 106 或由图 2A 到图 2B 中所示的着色器处理器 206 执行的指令的编译器 402 的框图。在一个实例方面中，编译器 402 用以产生待由着色器处理器 206 执行的指令。在此方面中，应用程序开发者可使用编译器 402 来产生供着色器处理器 206 执行的二进制指令（代码）。着色器处理器 206 为图形装置 200（图 2A）的部分。应用程序开发者可存取供与图形装置 200 一起使用的应用程序开发平台，且可产生用于图形像素应用程序 202A 和 / 或图形顶点应用程序 202B 的应用级软件。此应用级软件包括图 4 中所示的图形应用程序指令 400。图形应用程序指令 400 可包括通过符合或可转译成 DirectX®、OpenGL®、OpenVG™ 或其它语言的高级着色语言所写的指令。在一个方面中，这些着色语言界定可用于开发编程代码以执行图形操作的一个或一个以上标准 API。

[0069] 编译器 402 可至少部分由处理器所执行的编译器软件支持以接收及处理源代码指令且编译这些指令以产生经编译指令（例如，呈二进制、可执行机器指令的形式）。因此，编译器 402 可由执行与所述编译器软件相关联的计算机可读指令的一个或一个以上处理器形成。在一个方面中，这些一个或一个以上处理器可为应用程序开发者所使用的应用程序开发平台的部分或实施于所述应用程序开发平台中。所述经编译指令可存储于一计算机可读数据存储媒体上以供一个或一个以上处理器（例如，串流处理器 106 或着色器处理器 206）检索及执行。举例来说，本发明涵盖一种计算机可读数据存储媒体，其包括一个或一个以上第一可执行指令、一个或一个以上第二可执行指令及一个或一个以上第三可执行指令。

[0070] 所述第一可执行指令在由处理器执行时可支持一图形应用程序的一个或一个以上功能。另外，所述第一可执行指令中的每一者可指示用于其执行的第一数据精度级别。所述第二可执行指令在由处理器执行时可支持所述图形应用程序的一个或一个以上功能。另外，所述第二可执行指令中的每一者可指示用于其执行的不同于所述第一数据精度级别的第二数据精度级别。所述第三可执行指令在由所述处理器执行时也可支持所述图形应用程序的一个或一个以上功能，其中所述第三可执行指令中的每一者在所述一个或一个以上第一可执行指令经执行时将图形数据从所述第二数据精度级别转换到所述第一数据精度级别。

[0071] 编译器 402 可能能够将图形应用程序指令 400 编译成二进制图形指令 404，二进制图形指令 404 接着能够由着色器处理器 206 执行。着色器处理器 206 可从一数据存储媒体（例如，存储器或数据存储装置）检索这些指令，且执行这些指令以执行计算及其它操作来支持图形应用程序。图形应用程序指令 400 中的若干者可指定执行的特定数据精度级别。举例来说，某些指令可指定其使用全精度或半精度操作或计算。编译器 402 可经配置以应用规则 406 来在编译过程期间分析及剖析图形应用程序指令 400 且产生指示指令 404 的执行的数据精度级别的对应二进制指令图形 404。

[0072] 因此，如果图形应用程序指令 400 中的一者指定全精度操作或计算，则编译器 402 的规则 406 可产生二进制指令 404 中的为全精度指令的一者或一者以上。如果图形应用程序指令 400 中的另一者指定半精度操作或计算，则规则 406 产生二进制指令 404 中的为半精度指令的一者或一者以上。在一个方面中，二进制指令 404 各自可包括一指示所

述指令为全精度指令还是半精度指令的‘操作码’。在一个方面中，二进制指令 404 各自可使用可由着色器处理器 206 进行解码的指令的另一预定义字段、旗标或标头内所含有的信息来指示所述指令的执行的数据精度。在一个方面中，可基于待执行的指令的类型来推断数据精度。

[0073] 编译器 402 还包括能够产生在不同数据精度级别之间转换的二进制转换指令 410 的规则 408。在编译期间，编译器 402 的这些规则 408 可确定此转换在二进制指令 404 的执行期间可为必要的。举例来说，规则 408 可产生转换指令 410 内的将数据从全精度格式转换到半精度格式的一个或一个以上指令。当着色器处理器 206 执行图形指令 404 内的半精度指令时，可能需要此转换。规则 408 也可产生转换指令 410 内的将数据从半精度格式转换到全精度格式的一个或一个以上指令，当着色器处理器 206 执行图形指令 404 内的全精度指令时，可能需要此转换。

[0074] 根据一个方面，当编译器 402 的规则 408 产生转换指令 410 时，着色器处理器 206 可执行这些转换指令 410 以在对应图形指令 404 的执行期间管理数据精度转换。在此方面中，转换指令 410 的执行管理此精度转换，使得着色器处理器 206 未必需要使用某些硬件转换机制来将数据从一个精度级别转换到另一精度级别。转换指令 410 也可允许更有效地将数据传送到使用不同精度级别的 ALU(例如，到全精度 ALU 236 及到半精度 ALU 240)。

[0075] 本文中所描述的组件及技术可以硬件、软件、固件或其任何组合来实施。描述为模块或组件的任何特征可一起实施于集成逻辑装置中或单独地实施为离散但可共同操作的逻辑装置。在各种方面中，这些组件可至少部分形成为一个或一个以上集成电路装置（其可共同地称作集成电路装置），例如集成电路芯片或芯片组。此集成电路装置可用于多种图形应用程序及装置中的任一者中。在一些方面中，举例来说，这些组件可形成例如无线通信装置手持机等移动装置的部分。

[0076] 如果以软件来实施，则可至少部分通过计算机可读媒体来实现所述技术，所述计算机可读媒体包含在由一个或一个以上处理器执行时执行上文所描述的方法中的一者或一者以上的指令。计算机可读媒体可形成可包括封装材料的计算机程序产品的部分。计算机可读媒体可包含例如同步动态随机存取存储器 (SDRAM) 等随机存取存储器 (RAM)、只读存储器 (ROM)、非易失性随机存取存储器 (NVRAM)、电可擦除可编程只读存储器 (EEPROM)、快闪存储器、磁性或光学数据存储媒体。

[0077] 另外或其它，所述技术可至少部分通过计算机可读通信媒体来实现，所述计算机可读通信媒体载运或传达呈指令或数据结构的形式且可由一个或一个以上处理器存取、读取和 / 或执行的代码。可将任何连接恰当地称为计算机可读媒体。举例来说，如果使用同轴电缆、光纤电缆、双绞线、数字订户线 (DSL)，或例如红外、无线电及微波等无线技术从网站、服务器或其它远程源传输软件，则同轴电缆、光纤电缆、双绞线、DSL，或例如红外、无线电及微波等无线技术包括于媒体的定义中。以上各物的组合也应包括在计算机可读媒体的范围内。

[0078] 所利用的任何软件可由一个或一个以上处理器来执行，例如一个或一个以上数字信号处理器 (DSP)、通用微处理器、专用集成电路 (ASIC)、现场可编程门阵列 (FPGA)，或其它等效集成或离散逻辑电路。因此，如本文中所使用的术语“处理器”

或“控制器”可指代上述结构中的任一者或适于实施本文中所描述的技术的任何其它结构。因此，本发明还涵盖包括实施本发明中所描述的技术中的一者或一者以上的电路的多种集成电路装置中的任一者。此电路可提供于单一集成电路芯片装置中或多个可共同操作的集成电路芯片装置中。

[0079] 已描述本发明的各种方面。这些及其它方面在所附权利要求书的范围内。

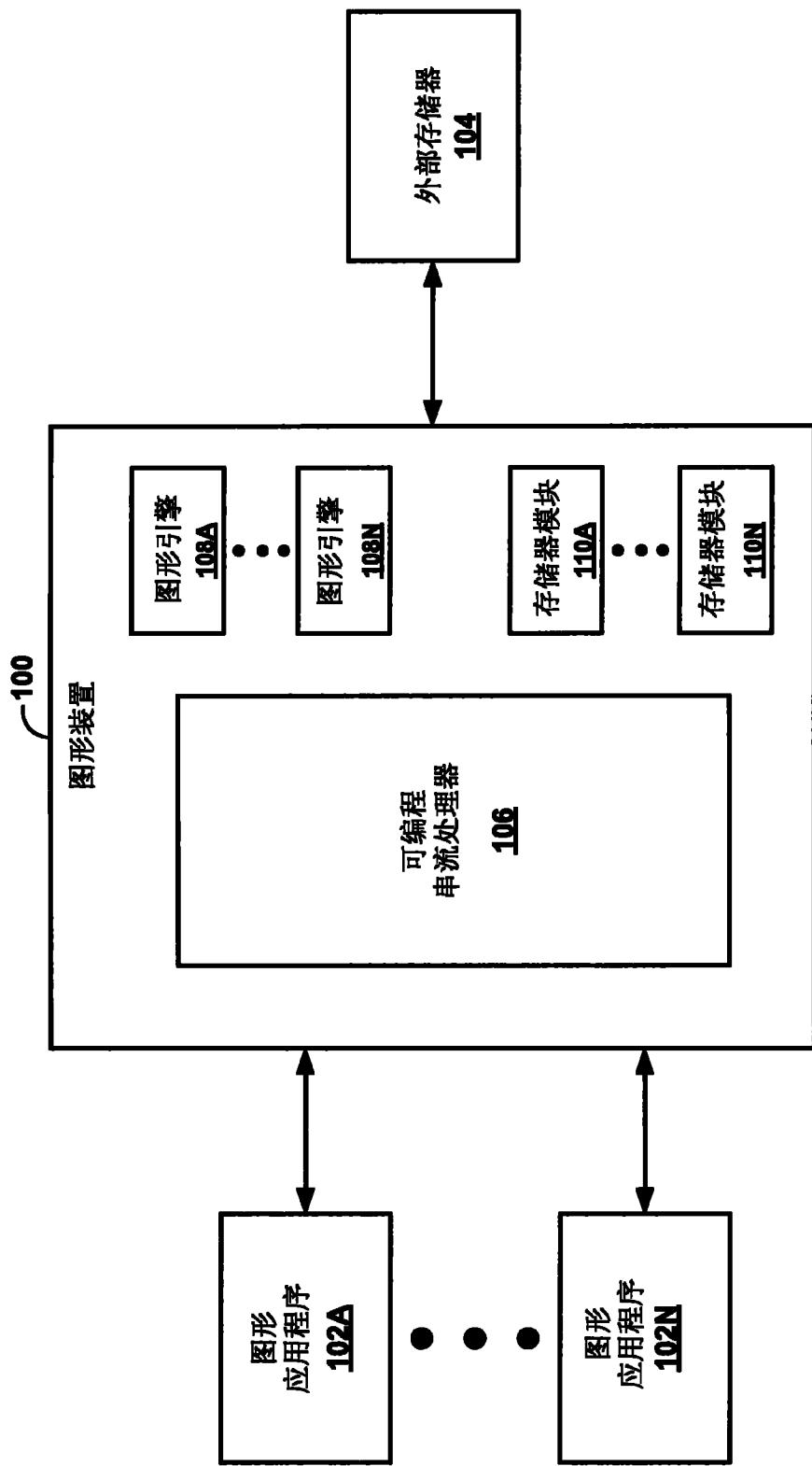


图 1

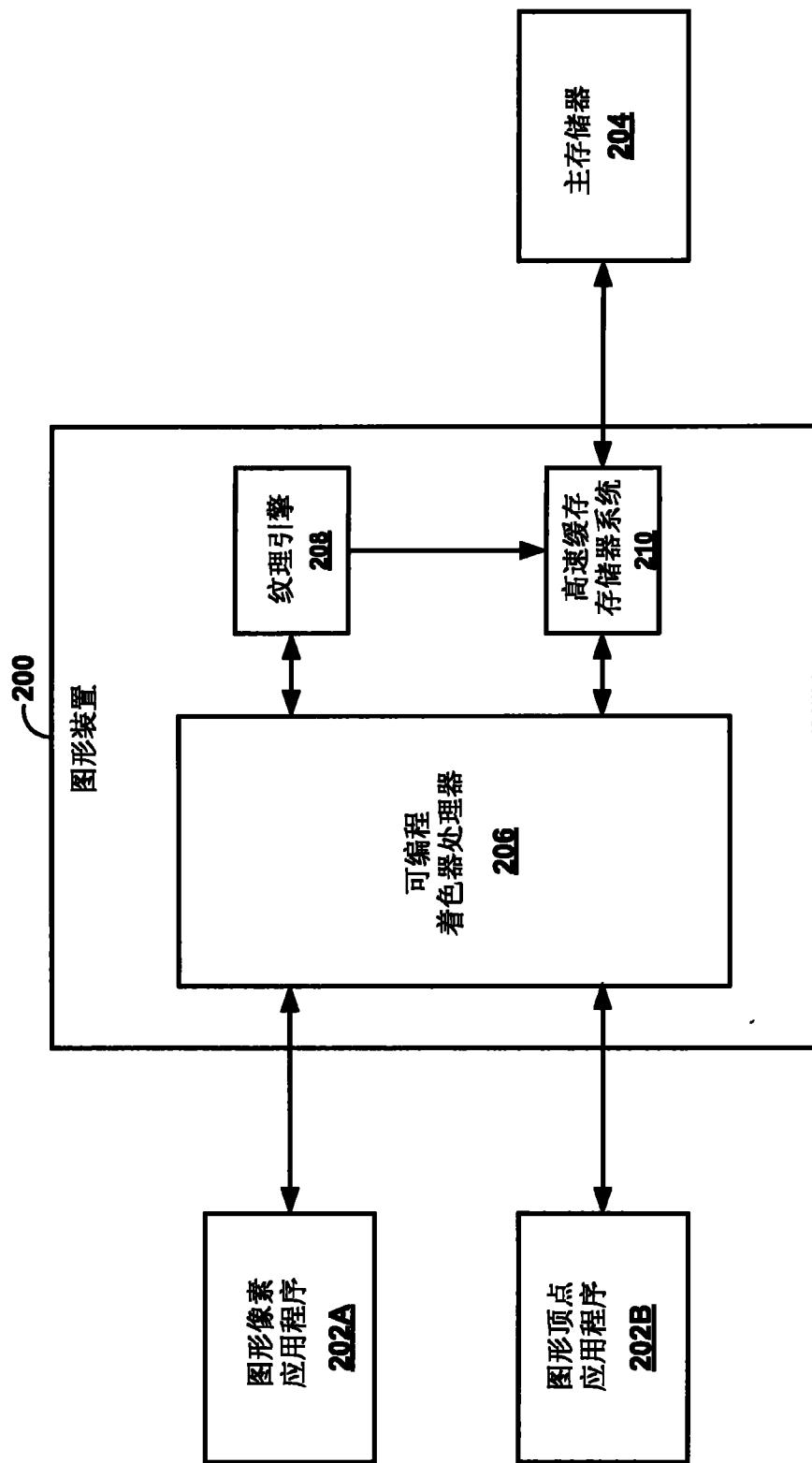


图 2A

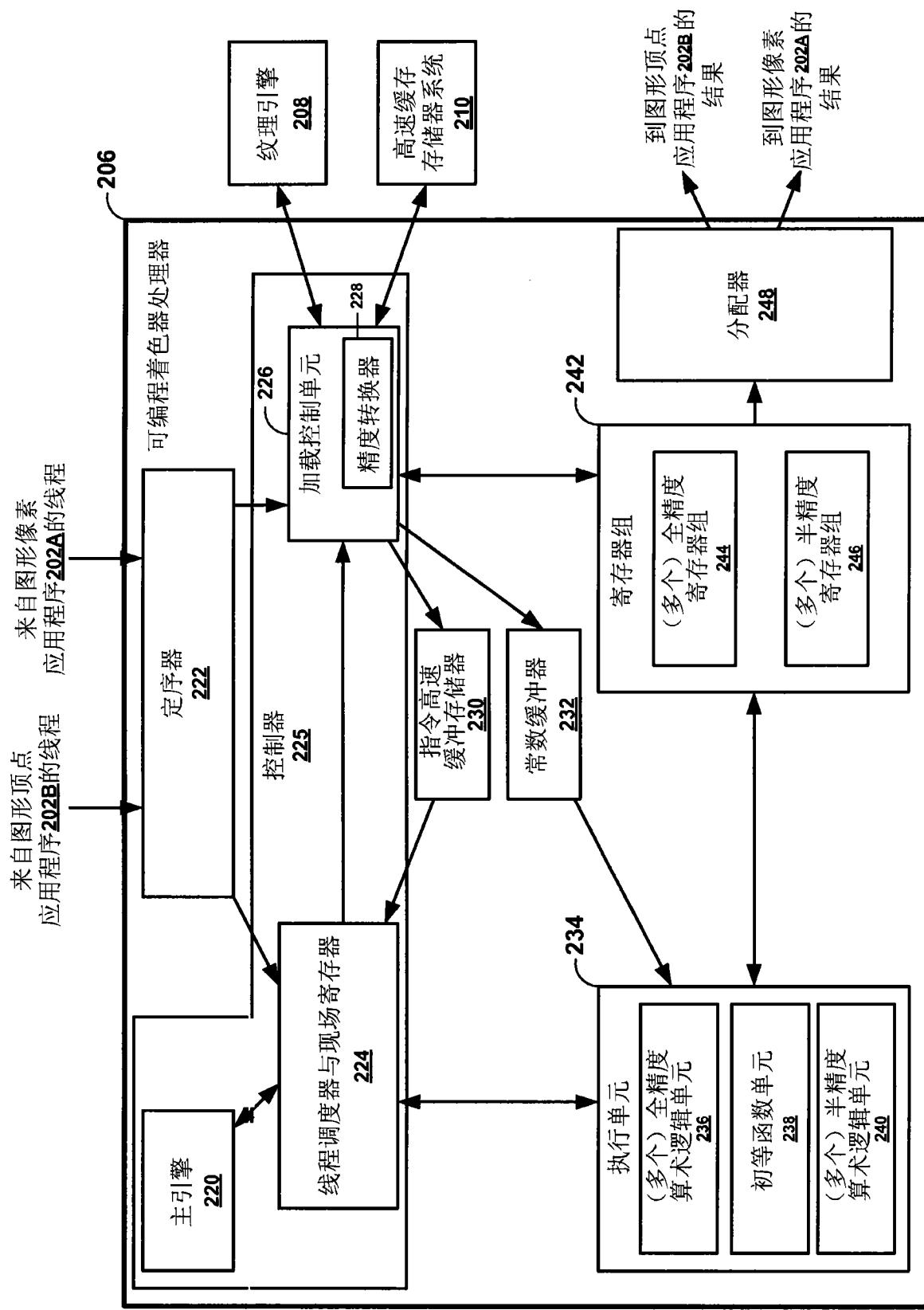


图 2B

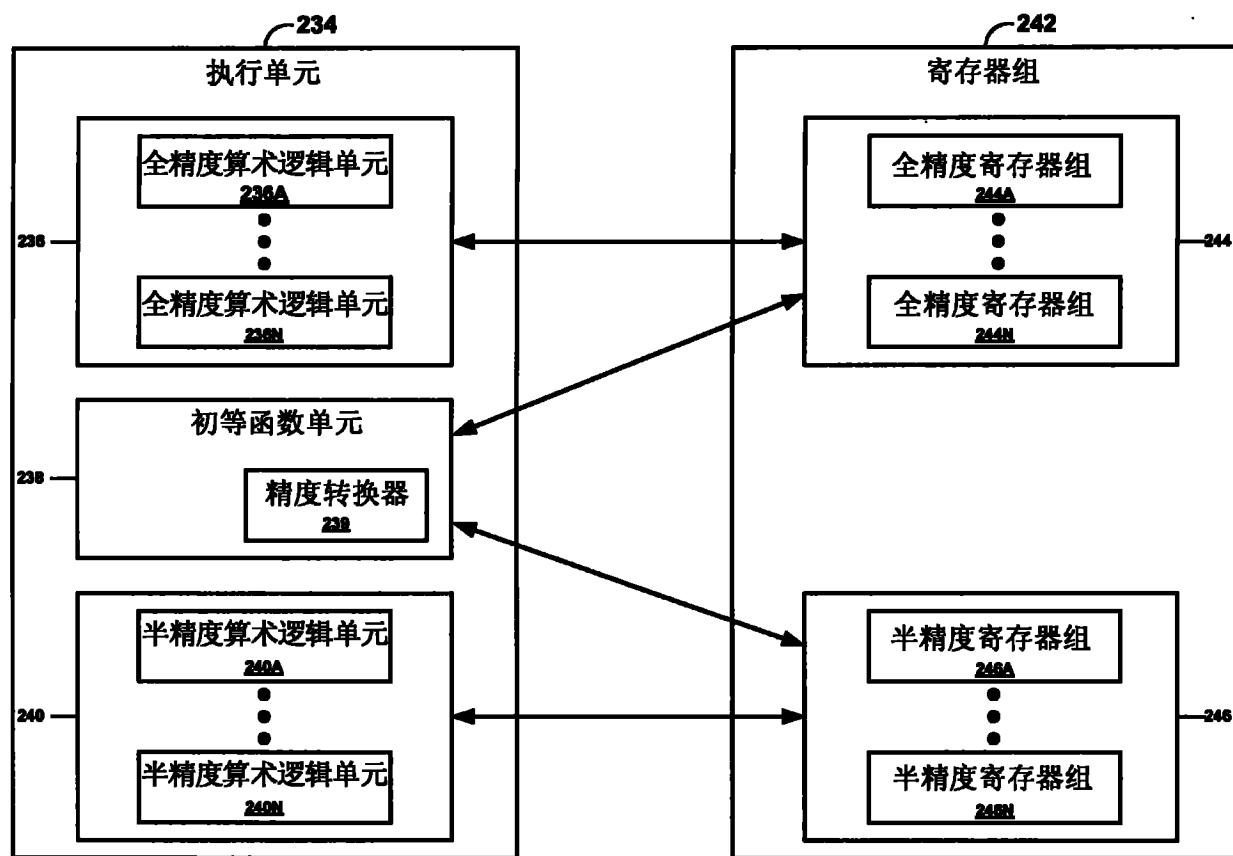


图 2C

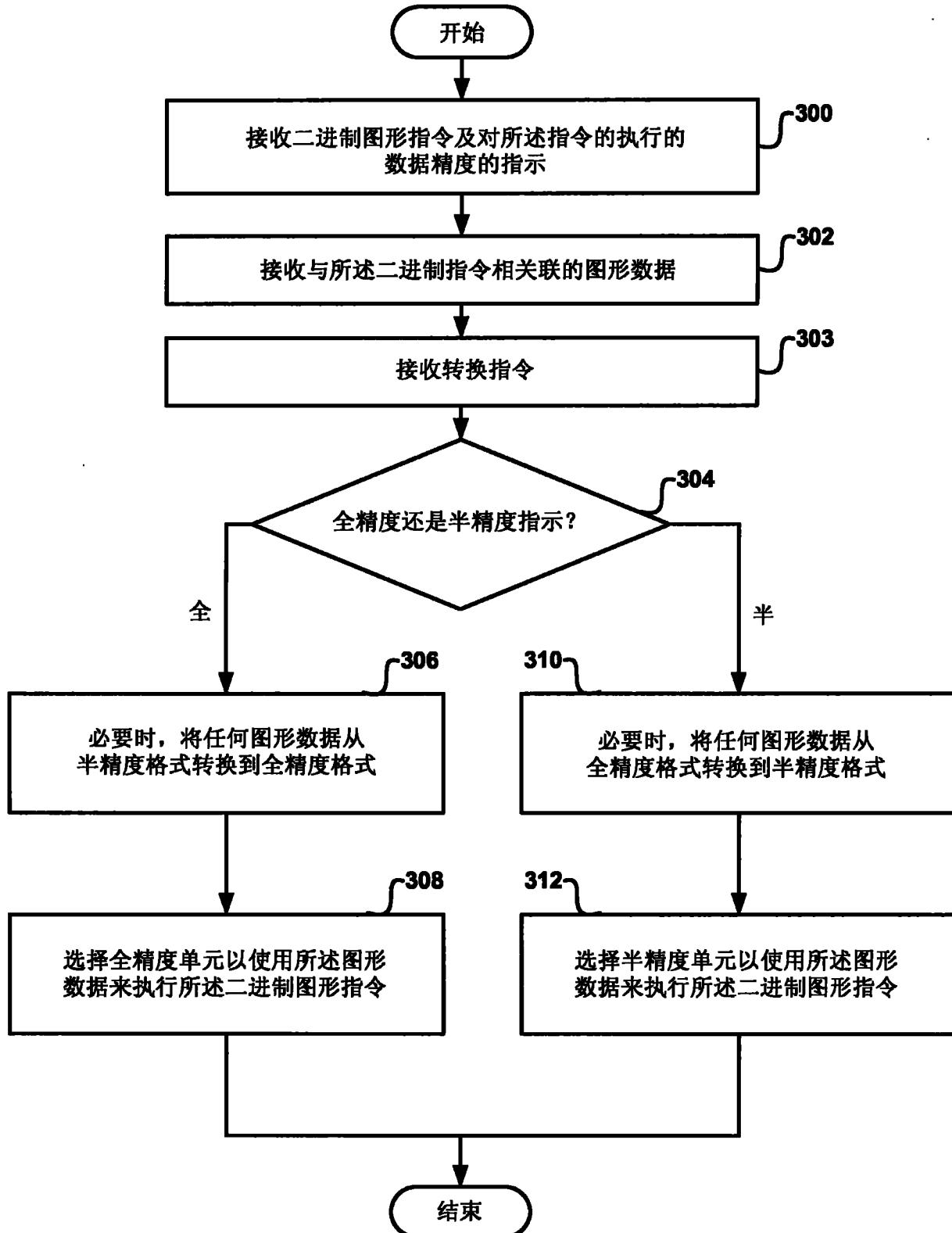


图 3

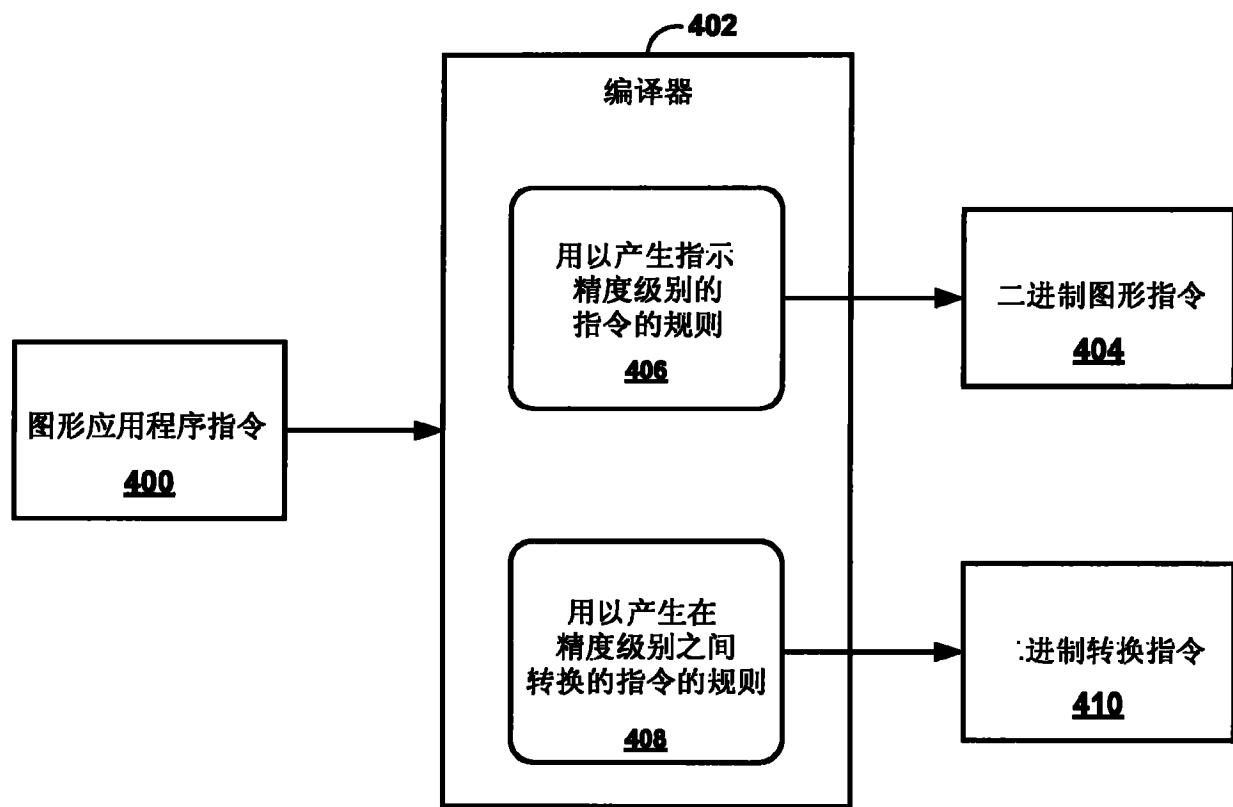


图 4