

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2009-535710

(P2009-535710A)

(43) 公表日 平成21年10月1日(2009.10.1)

(51) Int.Cl.
G06T 15/00 (2006.01)

F I
G06T 15/00 I 00A

テーマコード (参考)
5B080

審査請求 有 予備審査請求 未請求 (全 25 頁)

(21) 出願番号 特願2009-507929 (P2009-507929)
 (86) (22) 出願日 平成19年4月24日 (2007. 4. 24)
 (85) 翻訳文提出日 平成20年12月19日 (2008.12.19)
 (86) 国際出願番号 PCT/US2007/067325
 (87) 国際公開番号 W02007/127745
 (87) 国際公開日 平成19年11月8日 (2007.11.8)
 (31) 優先権主張番号 11/412, 678
 (32) 優先日 平成18年4月26日 (2006. 4. 26)
 (33) 優先権主張国 米国 (US)

(71) 出願人 595020643
 クゥアルコム・インコーポレイテッド
 QUALCOMM INCORPORATED
 アメリカ合衆国、カリフォルニア州 92
 121-1714、サン・ディエゴ、モア
 ハウス・ドライブ 5775
 (74) 代理人 100058479
 弁理士 鈴江 武彦
 (74) 代理人 100108855
 弁理士 蔵田 昌俊
 (74) 代理人 100091351
 弁理士 河野 哲
 (74) 代理人 100088683
 弁理士 中村 誠

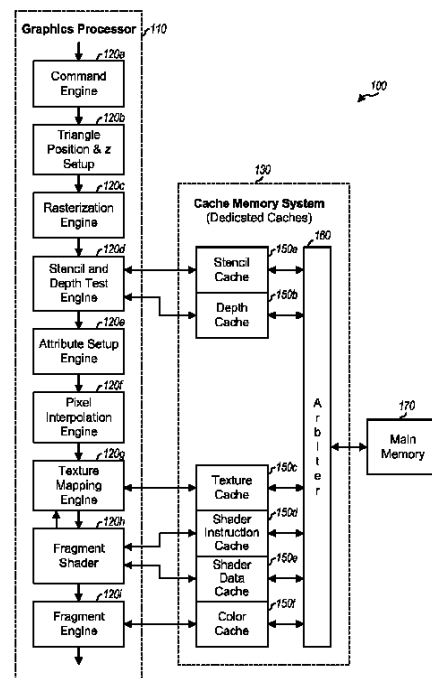
最終頁に続く

(54) 【発明の名称】 構成可能なキャッシュを有するグラフィックスシステム

(57) 【要約】

グラフィックスシステムがグラフィックスプロセッサおよびキャッシュメモリシステムを含んでいる。グラフィックスプロセッサは、グラフィックス画像をレンダリングするための種々のグラフィックス動作を行う処理ユニットを含んでいる。キャッシュメモリシステムは、完全に構成可能なキャッシュ、部分的に構成可能なキャッシュ、または構成可能キャッシュと専用キャッシュの組み合わせを含んでいてもよい。キャッシュメモリシステムは、制御ユニット、クロスバー、およびアービターをさらに含んでいてもよい。制御ユニットは、処理ユニットによるメモリ利用を判定し、そしてメモリ利用に基づいて構成可能なキャッシュを処理ユニットに割り当てうる。これらのキャッシュの良好な利用を実現するためおよびメモリアクセスボトルネックを回避するために、構成可能なキャッシュが割り当てられうる。クロスバーは、処理ユニットをそれらの割り当てられたキャッシュに結合する。アービターは、キャッシュとメインメモリとの間のデータ交換を容易にする。

【選択図】 図5



【特許請求の範囲】**【請求項 1】**

下記を備える装置：

グラフィックス画像をレンダリングするためのグラフィックス動作を行うように構成された複数の処理ユニット；

前記複数の処理ユニットのうちの少なくとも1つの処理ユニットに対するデータを記憶するように構成された複数のキャッシュ；および、

前記複数のキャッシュを前記少なくとも1つの処理ユニットに結合するように構成されたクロスバー。

【請求項 2】

前記複数のキャッシュのそれぞれが前記複数の処理ユニットのうちの任意の1つに割り当て可能である、請求項 1 の装置。

【請求項 3】

前記複数のキャッシュのそれぞれが前記複数の処理ユニットの各部分集合の任意の1つに割り当て可能である、請求項 1 の装置。

【請求項 4】

請求項 1 記載の装置、ここにおいて、

前記複数のキャッシュは、1つまたは複数の処理ユニットに割り当てられる1つまたは複数の専用キャッシュおよび残りの処理ユニットに割り当て可能な少なくとも1つの構成可能なキャッシュを備える。

【請求項 5】

各構成可能なキャッシュは、前記残りの処理ユニットの各部分集合の任意の1つに割り当て可能である、請求項 4 の装置。

【請求項 6】

前記残りの処理ユニットは、デブテストエンジンおよびテクスチャマッピングエンジンを含む、請求項 4 の装置。

【請求項 7】

更に下記を備える請求項 1 記載の装置：

前記複数の処理ユニットによるメモリ利用を確認しかつメモリ利用に基づいて前記複数のキャッシュを前記少なくとも1つの処理ユニットに割り当てるように構成された制御ユニット。

【請求項 8】

請求項 7 記載の装置、ここにおいて、

前記制御ユニットは、先行するグラフィックス画像に対するメモリ利用に基づいて、レンダリングされるべき各グラフィックス画像に対して前記複数のキャッシュを割り当てるように構成されている。

【請求項 9】

請求項 7 記載の装置、ここにおいて、

前記制御ユニットは、前記処理ユニットによるデータ要求、キャッシュヒット/ミス統計、またはそれらの組み合わせに基づいてメモリ利用を確認するように構成されている。

【請求項 10】

請求項 7 記載の装置、ここにおいて、

前記制御ユニットは、画像のレンダリング時に前記複数の処理ユニットによるメモリ利用の変化を検知しかつメモリ利用の検知された変化に基づいて前記複数のキャッシュを再割り当てするように構成されている。

【請求項 11】

更に下記を備える請求項 1 記載の装置：

実行されているグラフィックスアプリケーションによるメモリ利用に基づいて前記複数のキャッシュを前記少なくとも1つの処理ユニットに割り当てるように構成された制御ユニット。

10

20

30

40

50

【請求項 1 2】

請求項 1 記載の装置、ここにおいて、

前記クロスバーは、複数のインタフェースユニットを備え、各インタフェースユニットは、関連付けられた処理ユニットを該処理ユニットに割り当てられたキャッシュの組に結合するように構成されている。

【請求項 1 3】

請求項 1 2 記載の装置、ここにおいて、

各インタフェースユニットは、前記関連付けられた処理ユニットにより要求されたデータが該処理ユニットに割り当てられた前記キャッシュの組の任意の 1 つに記憶されるか否かを判定するように構成された状態機械を備える。

10

【請求項 1 4】

請求項 1 3 記載の装置、ここにおいて、

各インタフェースユニットに対する前記状態機械は、前記複数のキャッシュからのキャッシュヒット/ミスインジケータおよび前記関連付けられた処理ユニットに割り当てられた前記キャッシュの組を示す制御を受信する。

【請求項 1 5】

請求項 1 3 記載の装置、ここにおいて、

各インタフェースユニットに対する前記状態機械は、キャッシュミスが生じた場合に、前記関連付けられた処理ユニットに割り当てられた前記キャッシュの組の 1 つを満たすように構成されている。

20

【請求項 1 6】

前記複数のキャッシュは、キャッシュの少なくとも 2 つのレベルを有する階層構造に配列されている、請求項 1 の装置。

【請求項 1 7】

前記階層構造における少なくとも 1 つのレベルは構成可能な数のキャッシュを有する、請求項 1 6 の装置。

【請求項 1 8】

前記階層構造における少なくとも 1 つのレベルは、構成可能なキャッシュサイズを有する、請求項 1 6 の装置。

【請求項 1 9】

前記複数のキャッシュは、階層構造において構成可能な数のレベルに配列されている、請求項 1 の装置。

30

【請求項 2 0】

前記複数のキャッシュは、構成可能なキャッシュサイズを有する、請求項 1 の装置。

【請求項 2 1】

前記複数のキャッシュに結合されかつ前記複数のキャッシュとメインメモリとの間のデータ交換を容易にするように構成されたアービターをさらに含む、請求項 1 の装置。

【請求項 2 2】

前記複数の処理ユニットは、デブステストエンジンおよびテクスチャマッピングエンジンを含む、請求項 1 の装置。

40

【請求項 2 3】

請求項 2 2 記載の装置、ここにおいて、

前記複数の処理ユニットは、パイプラインに配列されており、かつ前記デブステストエンジンは、前記パイプラインにおいて前記テクスチャマッピングエンジンよりも早期に配置されている。

【請求項 2 4】

下記を備える集積回路：

グラフィックス画像をレンダリングするためのグラフィックス動作を行うように構成された複数の処理ユニット；

前記複数の処理ユニットのうちの少なくとも 1 つの処理ユニットに対するデータを記憶

50

するように構成された複数のキャッシュ；および、

前記複数のキャッシュを前記少なくとも1つの処理ユニットに結合するように構成されたクロスバー。

【請求項25】

前記複数のキャッシュのそれぞれが前記複数の処理ユニットの各部分集合の任意の1つに割り当て可能である、請求項24の集積回路。

【請求項26】

更に下記を備える請求項24記載の集積回路：

前記複数の処理ユニットによるメモリ利用を確認しかつメモリ利用に基づいて前記複数のキャッシュを前記少なくとも1つの処理ユニットに割り当てるように構成された制御ユニット。

10

【請求項27】

下記を備える無線装置：

グラフィックス画像をレンダリングするためのグラフィックス動作を行うように構成された複数の処理ユニットを含むグラフィックスプロセッサ；および、

キャッシュメモリシステム、ここで該キャッシュメモリシステムは下記を備える：

前記複数の処理ユニットのうちの少なくとも1つの処理ユニットに対するデータを記憶するように構成された複数のキャッシュと、

前記複数のキャッシュを前記少なくとも1つの処理ユニットに結合するように構成されたクロスバー。

20

【請求項28】

請求項27記載の無線装置、ここにおいて、

前記キャッシュメモリシステムは、さらに下記を備える：

前記複数のキャッシュに結合されかつ前記複数のキャッシュとメインメモリとの間のデータ交換を容易にするように構成されたアービター。

【請求項29】

下記を備える方法：

グラフィックス画像をレンダリングするためのグラフィックス動作を行うように構成された複数の処理ユニットによるメモリ利用を判定すること；

前記複数の処理ユニットによるメモリ利用に基づいて前記複数の処理ユニットのうちの少なくとも1つの処理ユニットに複数のキャッシュを割り当てること；および、

30

前記少なくとも1つの処理ユニットのそれぞれを該処理ユニットに割り当てられた対応するそれぞれの組のキャッシュに結合することを含む方法。

【請求項30】

1つまたは複数のキャッシュを前記複数の処理ユニットのなかの1つまたは複数の処理ユニットに直接に結合することをさらに含む、請求項29の方法。

【請求項31】

請求項29記載の方法、ここにおいて、

前記複数のキャッシュを割り当てることは、下記を備える：

先行するグラフィックス画像に対するメモリ利用に基づいて、前記複数のキャッシュをレンダリングされるべき各グラフィックス画像に対する少なくとも1つの処理ユニットに割り当てること。

40

【請求項32】

下記を備える装置：

グラフィックス画像をレンダリングするためのグラフィックス動作を行うように構成された複数の処理ユニットによるメモリ利用を判定する手段；

前記複数の処理ユニットによるメモリ利用に基づいて、複数のキャッシュを前記複数の処理ユニットのうちの少なくとも1つの処理ユニットに割り当てる手段；および、

前記少なくとも1つの処理ユニットのそれぞれを該前記処理ユニットに割り当てられた対応するそれぞれの組のキャッシュに結合する手段。

50

【請求項 3 3】

更に下記を備える請求項 3 2 記載の装置：

1 つまたは複数のキャッシュを前記複数の処理ユニットのうちの 1 つまたは複数の処理ユニットに直接に結合する手段。

【請求項 3 4】

請求項 3 2 記載の装置、ここにおいて、

前記複数のキャッシュを割り当てる手段は、下記を備える：

先行するグラフィックス画像に対するメモリ利用に基づいて、前記複数のキャッシュをレンダリングされるべき各グラフィックス画像に対する前記少なくとも 1 つの処理ユニットに割り当てる手段。

10

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は一般的には回路に関し、さらに具体的にはグラフィックスシステムに関する。

【背景技術】**【0002】**

グラフィックスシステムは、ビデオゲーム、グラフィックス、コンピュータ援用設計 (CAD)、シミュレーションおよび視覚化ツール、画像化、等のような種々の用途に対して 2 次元 (2 - D) および 3 次元 (3 - D) 画像をレンダリングする (render) ために広く用いられている。3 - D 画像は、表面でモデル化されることができ、そして各表面は多角形 (典型的には三角形) で近似できる。3 - D 画像を表示するために用いられる三角形の数は、表面の複雑性と画像の所望される解像度によって決まり、極めて大きく例えば何百万であることもありうる。各三角形は、3 つの頂点によって画定され、各頂点は空間座標、カラー値、テクスチャ座標のような種々の属性に関連する。各属性は、4 つまでの要素を有しうる。例えば、空間座標は、典型的には水平および垂直座標 (x および y) とデプス (depth) (z) によって与えられ、カラー値は、典型的には赤、緑、青 (r、g、b) によって与えられ、そしてテクスチャ座標は、典型的には水平および垂直座標 (u および v) によって与えられる。

20

【0003】

グラフィックスシステムにおけるグラフィックスプロセッサは、2 - D または 3 - D 画像をレンダリングするための種々のグラフィックス動作を行いうる。画像は、多数の三角形で構成され、各三角形は画素 (ピクセル) で構成される。グラフィックスプロセッサは、三角形内の各ピクセルの要素値を判定することによって各三角形をレンダリングする。グラフィックス動作は、ラスター化 (rasterization)、ステンシルおよびデプステスト (stencil and depth tests)、テクスチャマッピング (texture mapping)、シェーディング (shading)、等を含みうる。グラフィックスプロセッサは、ピクセルに関連した任意のデータでありうるピクセルデータについて種々のグラフィックス動作を行うので、ピクセルデータに迅速にアクセスできることが、性能に対して大きな影響を有する。

30

【発明の開示】**【発明の概要】**

40

【0004】

構成可能なキャッシュを有しかつ良好な性能を有するグラフィックスシステムがここに記載されている。グラフィックスシステムは、グラフィックスプロセッサおよびキャッシュメモリシステムを含んでいる。グラフィックスプロセッサは、グラフィックス画像をレンダリングするための種々のグラフィックス動作を行う処理ユニットを含んでいる。キャッシュメモリシステムは、完全に構成可能なキャッシュ (fully configurable caches)、部分的に構成可能なキャッシュ (partially configurable caches)、または構成可能なキャッシュ (configurable caches) と専用キャッシュ (dedicated caches) の組み合わせを含むことができる。キャッシュは、キャッシュを割り当てられうる処理ユニットのうちの任意の 1 つに割り当てられうるならば、完全に構成可能である。キャッシュは、そ

50

れが処理ユニットの部分集合のうちの任意の1つに割り当てられうるならば、部分的に構成可能である。キャッシュは、それが特定の処理ユニットに割り当てられるならば、専用である。キャッシュは、処理ユニットに対するデータ（例えば、ピクセルデータおよび/または命令）を記憶する高速メモリである。

【0005】

キャッシュメモリシステムは、制御ユニット、クロスバー（crossbar）、およびアービター（arbiter）をさらに含む。制御ユニットは、処理ユニットによるメモリ利用を判定しかつメモリ利用に基づいて処理ユニットに構成可能なキャッシュを割り当てうる。構成可能なキャッシュは、これらのキャッシュの良好な利用を達成しかつグラフィックスプロセッサ内のあらゆる位置においてメモリアクセスのボトルネック（memory access bottleneck）を回避するために割り当てられうる。クロスバーは、処理ユニットをそれらの割り当てられたキャッシュに結合する。アービターは、キャッシュとメインメモリとの間のデータ交換を容易にする。

10

【0006】

本発明の種々の態様および実施の形態が、以下においてさらに詳細に記述される。

【0007】

本発明の特徴および本質は、同様の参照符号が全体にわたり対応して識別している図面に関連してなされる下記の詳細な説明から更に明白となるであろう。

【発明を実施するための最良の形態】

【0008】

「好例の」（"exemplary"）という語は、本明細書では、「実施例、事例、または例示として適していること」を意味するために用いられる。「好例の」としてここに記述される実施の形態またはデザインは、必ずしも、他の実施の形態またはデザインよりも好ましいあるいは有益であると解釈されるべきものではない。

20

【0009】

図1は、専用キャッシュを有するグラフィックスシステム100のブロック図を示している。グラフィックスシステム100は、グラフィックスプロセッサ110およびキャッシュメモリシステム130を含んでいる。グラフィックスプロセッサ110は、パイプラインで配列されかつ2-Dおよび3-D画像をレンダリングするために用いられる種々の処理ユニットを含んでいる。コマンドエンジン120aは、実行されるべきグラフィックス動作を特定する入来コマンドを受信しかつ復号する。三角形位置およびzセットアップエンジン120bは、後続のラスタライズ処理のための必要なパラメータを演算する。例えば、エンジン120bは、各三角形の3つのエッジのための線形方程式の係数、デプス（depth）（z）勾配のための係数、等を演算する。ラスタライズエンジンは120cは、各三角形またはラインをピクセルに分解し、各ピクセルに対するスクリーン座標を生成する。

30

【0010】

ステンシルおよびデプステストエンジン120dは、イネーブルされていれば、ピクセルが表示されるべきかあるいは捨てられるべきかを判定するために、各ピクセルに対してステンシルテストを行う。ステンシルバッファ（図1には図示なし）は、レンダリングされている画像内の各ピクセル場所に対する現在のステンシル値を記憶する。エンジン120dは、各ピクセルに対する記憶されたステンシル値を基準値と比較し、その比較の結果に基づいてピクセルを保持するかあるいは捨てる。エンジン120dはまた、適用可能であれば、ピクセルが表示されるべきかあるいは捨てられるべきかを判定するために、各ピクセルに対してデプステスト（これはzテストとも呼ばれる）を行う。z-バッファ（図1には図示なし）は、レンダリングされている画像内の各ピクセル場所に対する現在のz値を記憶する。エンジン120dは、各ピクセルのz値（現在のz値）をz-バッファにおける対応したz値（記憶されたz値）と比較し、ピクセルを表示し、そして現在のz値が記憶されたz値よりも近接/接近している（closer/nearer）場合には、z-バッファと場合によりステンシルバッファとを更新し、現在のz値が記憶されたz値

40

50

よりも後にある場合には、そのピクセルを捨てる。

【 0 0 1 1 】

属性セットアップエンジン 1 2 0 e は、ピクセル属性の後続の補間のための必要なパラメータを演算する。例えば、エンジン 1 2 0 e は、属性補間のための線形方程式の係数を演算しうる。ピクセル補間エンジン 1 2 0 f は、ピクセルのスクリーン座標に基づいてかつエンジン 1 2 0 e からの情報を用いて、各三角形内の各ピクセルに対する属性要素値を演算する。テクスチャマッピングエンジン 1 2 0 g は、イネーブルされていれば、各三角形にテクスチャを適用するためにテクスチャマッピングを行う。テクスチャ画像は、テクスチャバッファ（図 1 には図示なし）に記憶されうる。各三角形の 3 つの頂点は、テクスチャ画像における 3 つの (u , v) 座標に関連され、そしてその三角形の各ピクセルは、テクスチャ画像における特定のテクスチャ座標に関連されうる。テクスチャリング (texturing) は、ピクセルのテクスチャ座標によって示された場所におけるテクスチャ画像のカラーで各ピクセルのカラーを修正することによって達成される。

10

【 0 0 1 2 】

各ピクセルは、カラー、デプス、テクスチャ等のような情報に関連づけられる。フラグメントシェーダ 1 2 0 h は、命令のシーケンスで構成されうるソフトウェアプログラムを各フラグメントに適用する。フラグメントシェーダ 1 2 0 h はまた、テクスチャマッピングエンジン 1 2 0 g にテクスチャ要求を送ることもできる。フラグメントエンジン 1 2 0 i は、アルファテスト (alpha test)、フォグ混合 (fog blending)、アルファ混合 (alpha blending)、論理演算 (logic operation)、およびデイザリング動作 (dithering operation) のような機能を各フラグメントに対して行い、そして結果をカラーバッファに与える。

20

【 0 0 1 3 】

図 1 に示されているグラフィックス動作は、技術的に公知である。グラフィックスプロセッサは、オープングラフィックスライブラリ (Open Graphics Library) (オープン G L)、ダイレクト 3 D (Direct 3D) 等のようなソフトウェア・インタフェースを実装しうる。オープン G L は、公表されている 2 0 0 4 年 1 0 月 2 2 日付けの " オープン G L (登録商標) グラフィックス・システム : 仕様書 " バージョン 2 . 0、(" The OpenGL (Registered Trademark) Graphics System : A Specification, " Version 2.0, dated October 22, 2004) という名称の文書に記載されている。

30

【 0 0 1 4 】

一般に、グラフィックスプロセッサは、任意の組のグラフィックス動作のための任意の数の処理ユニットを含みうる。各処理ユニットは、専用ハードウェアで実装されたエンジン、プロセッサ、または両者の組み合わせでありうる。例えば、図 1 に示されているエンジンは専用ハードウェアで実装されうるものであり、他方、フラグメントシェーダはプログラム可能な中央処理ユニット (C P U) で実装されうる。処理ユニットは、所望の最適化に依存して種々の順序で配列されうる。例えば、節電のために、図 1 に示されているように、ステンシルおよびデプステストをパイプラインで早期に行い、目に見えないピクセルが早期に捨てられるようにすることが望ましいであろう。処理ユニットはまた、他の順序で配列されてもよく、例えば、ステンシルおよびデプステストエンジン 1 2 0 d がテクスチャマッピングエンジン 1 2 0 g の後に配置されてもよい。

40

【 0 0 1 5 】

キャッシュメモリシステム 1 3 0 は、グラフィックスプロセッサ 1 1 0 内の処理ユニット 1 2 0 のうちの幾つかのための専用キャッシュ 1 5 0 a ~ 1 5 0 f を含んでいる。図 1 に示されたデザインでは、ステンシルキャッシュ 1 5 0 a は、レンダリングされている画像のある領域内のピクセルに対するステンシル値を記憶し、デプスキャッシュ 1 5 0 b は、同じまたは異なる領域内のピクセルに対するデプス値を記憶する。ステンシルおよびデプステストエンジン 1 2 0 d は、キャッシュ 1 5 0 a に記憶されたステンシル値にアクセスして、これらのステンシル値を用いてステンシルテストを行う。エンジン 1 2 0 d は、キャッシュ 1 5 0 b に記憶された z 値にもアクセスして、これらの z 値を用いてデプステ

50

ストを行う。テクスチャキャッシュ 150c は、三角形にテクスチャをマップするためにテクスチャマッピングエンジン 120g によって用いられるテクスチャマッピングテーブルを記憶する。シェーダ命令キャッシュ 150d は、フラグメントシェーダ 120h に対する命令を記憶し、そしてシェーダデータキャッシュ 150e は、フラグメントシェーダ 120h のためのデータを記憶する。カラーキャッシュ 150f は、ピクセルのためのカラー（例えば、赤、緑、青）値を記憶する。

【0016】

アービター 160 は、キャッシュ 150a ~ 150f とメインメモリ 170 に結合し、かつそれら間のデータ交換を容易にする。処理ユニット 120 がデータのために関連付けられたキャッシュにアクセスして、要求されたデータがそのキャッシュに記憶されていない場合には、キャッシュミスが宣言される。このキャッシュミスにより、必要であれば、キャッシュの現在のコンテンツがメインメモリに戻されて記憶されることになり、そしてデータの新しい領域がメインメモリ 170 からフェッチ (fetched) されて、キャッシュにロードされることになる。要求データがキャッシュから得られて、処理ユニットに与えられうる。アービター 160 は、一度に 1 つのキャッシュがメインメモリ 170 とデータを交換できるようにする。

10

【0017】

図 1 におけるキャッシュメモリシステム 130 は、グラフィックスシステム 100 に対するメモリアクセス性能を向上させる。キャッシュ 150 は、グラフィックスプロセッサ 110 に近接して配置された小さな高速メモリである。高速キャッシュメモリは、処理ユニット 120 がより高速でデータを回収し、かつ結果を記憶することができるから、処理速度を上げる。メインメモリ 170 は、グラフィックスプロセッサ 110 からさらに離れて（例えば、オフチップで）配置された大きい低速メモリである。メインメモリ 170 は、キャッシュ 150 にロードされうるデータおよび命令を記憶する。キャッシュ 150 は、メインメモリ 170 からのフェッチ (fetches) の回数を減少させ、かつ一般にグラフィックスシステム 100 の性能を向上させる。

20

【0018】

図 1 は、専用キャッシュの使用を示す。各キャッシュ 150 は、特定の処理ユニット 120 に対するデータまたは命令を記憶する。各キャッシュ 150 の利用は、実行されているグラフィックスアプリケーションによって決まる。例えば、3-D グラフィックスアプリケーションは、z パスを実行するだけでよく、デプスおよび / またはステンシルテストだけを行わせればよい。このアプリケーションに対しては、ステンシルキャッシュおよび / またはデプスキャッシュだけが利用されればよく、残りのキャッシュ 150c ~ 150f は遊んでいてもよい。他の例としては、単純なシェーダアプリケーションは、シェーダ命令キャッシュ 150d あるいはシェーダデータキャッシュ 150e を利用しなくてもよく、これらのキャッシュは遊んでいてもよい。さらに他の例としては、グラフィックスアプリケーションは、テクスチャマッピングを無効にしてもよく、その場合には、テクスチャキャッシュ 150c は必要とされなくてもよい。これらの例によって分かるように、専用キャッシュを使用すると、ある種のグラフィックスアプリケーションによるキャッシュの利用が低下することになりうる。専用キャッシュを使用すると、異なるグラフィックスアプリケーションに対して異なる場所でメモリアクセスボトルネックが生じることにもなりうる。このボトルネックがグラフィックスシステム 100 の性能を制限する。

30

40

【0019】

図 2 は、完全に構成可能なキャッシュを有するグラフィックスシステム 200 の 1 つの実施の形態を示している。グラフィックスシステム 200 は、グラフィックスプロセッサ 210 およびキャッシュメモリシステム 230 を含んでいる。グラフィックスプロセッサ 210 は、図 1 においてそれぞれ処理ユニット 120a ~ 120i に対して上述したように動作しうる処理ユニット 220a ~ 220i を含んでいる。図 2 に示されているように、処理ユニットのうちの一つかだけが、データおよび / または命令を記憶するためにキャッシュを利用しうる。特に、ステンシルおよびデプステストエンジン 220d に対するス

50

テンシルおよび/またはデプス値、テクスチャマッピングエンジン 220 g に対するテクスチャデータ、フラグメントシェーダ 220 h に対する命令および/またはデータ、およびフラグメントエンジン 220 i に対するカラー値を記憶するためにキャッシュが用いられる。

【0020】

キャッシュメモリシステム 230 は、グラフィックスプロセッサ 210 内の処理ユニット 220 に対するデータおよび/または命令を記憶するために割り当てられうる N 個の完全に構成可能なキャッシュ 250 a ~ 250 n を含んでおり、一般に、N = 1 である。それらのキャッシュは、各キャッシュが、キャッシュを割り当てられうる任意の処理ユニットに割り当てられうるという点で、完全に構成可能である。1つの実施の形態では、各キャッシュ 250 は、1つの処理ユニット 220 だけに割り当てられうるが、所定の処理ユニット 220 はゼロ、1、または複数のキャッシュ 250 を割り当てられうる。制御ユニット 232 は、ステンシルおよびデプスエンジン 220 d、テクスチャマッピングエンジン 220 g、フラグメントシェーダ 220 h、および/またはフラグメントエンジン 220 i に対して、これらの処理ユニットのメモリ利用に基づいて、利用可能なキャッシュ 250 a ~ 250 n を割り当てる。クロスバー 240 は、処理ユニット 220 d、220 g、220 h、および 220 i をそれらの割り当てられたキャッシュ 250 a ~ 250 n に結合する。クロスバー 240 はまた、クロススイッチまたは他の用語で呼ばれてもよい。アービター 260 は、キャッシュ 250 a ~ 250 n をメインメモリ 270 に結合し、それらのキャッシュとメインメモリとの間におけるデータ交換を容易にする。アービター 260 はまた、セクタ、メモリインタフェース、または他の用語で呼ばれてもよい。

10

20

【0021】

一般に、キャッシュメモリシステムは、任意のサイズを有しうる任意の個数のキャッシュを含みうる。キャッシュの個数およびそれらのサイズは、メモリアクセス性能と回路面積との間の妥協に基づいて選択されうる。大きいキャッシュサイズでは、キャッシュヒットレイトが高くなり、かつ要求データがそれらのキャッシュに記憶される可能性が大きくなる。しかし、大きいキャッシュサイズはまた、大きい回路面積を必要とし、それがコストを増加させることになりうる。1つの特定の実施の形態では、キャッシュメモリシステム 230 は、12個のキャッシュ 250 を含み、各キャッシュは 4 キロバイトである。他の実施の形態では、キャッシュメモリシステムに対して、それより少ないあるいは追加のキャッシュならびに他のキャッシュサイズが用いられうる。

30

【0022】

図 3 は、図 2 におけるクロスバー 240 の 1つの実施の形態のブロック図を示す。この実施の形態では、クロスバー 240 は、K 個のインタフェースユニット 310 a ~ 310 k を含んでおり、一般に、K は任意の整数値でありうる。インタフェースユニットはまた、マスターユニットまたは他の用語で呼ばれうる。図 2 に示された実施の形態では、K は 6 に等しく、6 個のインタフェースユニットは、ステンシル、デプス、テクスチャ、シェーダ命令、シェーダデータ、およびカラーに対するものである。所定の処理ユニット 220 は、処理ユニットのデータ要件に依存して、1つまたは複数のインタフェースユニット 310 と相互に作用しうる。クロスバー 240 は、K 個のインタフェースユニット 310 a ~ 310 k に対してそれぞれ K 個のマルチプレクサ (Mux) 340 a ~ 340 k、および N 個のキャッシュ 250 a ~ 250 n に対してそれぞれ N 個のマルチプレクサ 350 a ~ 350 n をさらに含んでいる。

40

【0023】

各インタフェースユニット 310 は、各組のライン 302 を経由して関連付けられた処理ユニット 220 に結合する。各インタフェースユニット 310 は、状態機械 320 および入力/出力 (I/O) ユニット 330 を含んでいる。各インタフェースユニット 310 では、I/O ユニット 330 が関連付けられた処理ユニット 220 からライン 302 を経由してアドレスおよび/または出力データを受信して、そのアドレス/データを N 個のマルチプレクサ 350 a ~ 350 n のそれぞれの 1つの入力に与える。I/O ユニット 33

50

0 はまた、関連マルチプレクサ 3 4 0 から入力データまたは命令を受信して、関連付けられた処理ユニット 2 2 0 にライン 3 0 2 を経由してデータ/命令を与える。

【 0 0 2 4 】

各マルチプレクサ 3 5 0 は、K 個の入力において、すべての K 個の I / O ユニット 3 3 0 a ~ 3 3 0 k からアドレス/データを受信する。各マルチプレクサ 3 5 0 はまた、制御ユニット 2 3 2 から各自の制御 C n を受信して、その制御 C n によって決定される 1 つの入力からアドレス/データを関連付けられたキャッシュ 2 5 0 に与える。各マルチプレクサ 3 5 0 は、このようにして 1 つのインタフェースユニット 3 1 0 を関連付けられたキャッシュ 2 5 0 に結合する。

【 0 0 2 5 】

各マルチプレクサ 3 4 0 は、N 個の入力において、すべての N 個のキャッシュ 2 5 0 a ~ 2 5 0 n から入力データまたは命令を受信する。各マルチプレクサ 3 4 0 はまた、制御ユニット 2 3 2 から各自の制御 P k を受信して、その制御 P k によって決定される 1 つの入力からのデータ/命令を関連 I / O ユニットに与える。各マルチプレクサ 3 4 0 は、このようにして 1 つのキャッシュ 2 5 0 を関連インタフェースユニット 3 1 0 に結合する。

【 0 0 2 6 】

各キャッシュ 2 5 0 は、関連マルチプレクサ 3 5 0 からアドレスを受信して、そのアドレスにおける要求データ/命令がそのキャッシュに記憶されているか否かを判定する。各キャッシュ 2 5 0 は、要求データ/命令がそのキャッシュに記憶されていればキャッシュヒット（例えば、論理ハイ）を示し、あるいは、要求データ/命令がそのキャッシュに記憶されていなければキャッシュミス（例えば、論理ロー）を示すヒット/ミス（h / m）のインジケータ（indicator）を与える。各キャッシュ 2 5 0 は、キャッシュヒットが存在すれば、要求データ/命令を与える。各キャッシュ 2 5 0 は、メインメモリ 2 7 0 からのフェッチ（fetch）を待ち、そしてキャッシュミスが存在する場合には、要求データ/命令を与える。

【 0 0 2 7 】

各インタフェースユニット 3 1 0 では、状態機械 3 2 0 が、すべての N 個のキャッシュ 2 5 0 a ~ 2 5 0 n からのヒット/ミスのインジケータと、どのキャッシュがそのインタフェースユニットに割り当てられているかを示す制御 M k を受信する。状態機械 3 2 0 は、割り当てられたキャッシュに対するヒット/ミスのインジケータに基づいて、要求データ/命令が割り当てられたキャッシュのどれかに記憶されているか否かを、割り当てキャッシュに対するヒット/ミスのインジケータに基づいて、判定する。状態機械 3 2 0 は、要求データ/命令が割り当てられたキャッシュのどれにも記憶されていないならば、フェッチ命令を生成する。

【 0 0 2 8 】

制御ユニット 2 3 2 は、状態機械 3 2 0 a ~ 3 2 0 k それぞれに対する制御 M 1 ~ M K、マルチプレクサ 3 4 0 a ~ 3 4 0 k それぞれに対する制御 P 1 ~ P K、およびマルチプレクサ 3 5 0 a ~ 3 5 0 n それぞれに対する制御 C 1 ~ C N を生成する。これらの制御は、どのキャッシュが各インタフェースユニット 3 1 0 に割り当てられるかを判定する。

【 0 0 2 9 】

図 3 は、完全に構成可能なキャッシュに対するクロスバー 2 4 0 の 1 つの実施の形態を示している。一般に、キャッシュの完全構成可能性は、十分な数の入力を有するマルチプレクサを使用し、可能なキャッシュ割り当てを評価するように状態機械を設計し、そして状態機械およびマルチプレクサに対する適切な制御を生成することによって、実現される。クロスバー 2 4 0 はまた、処理ユニットを利用可能なキャッシュにマップできる他のデザインで実装される。

【 0 0 3 0 】

図 4 は、図 3 におけるキャッシュ 2 5 0 n およびインタフェースユニット 3 1 0 k に対する状態機械 3 2 0 の 1 つの実施の形態のブロック図を示している。キャッシュ 2 5 0 n は、キャッシュコントローラ 4 1 0、タグ RAM 4 2 0、およびデータ RAM 4 3 0 を含ん

10

20

30

40

50

でいる。キャッシュコントローラ 4 1 0 は、キャッシュ 2 5 0 n の各アクセスに対してアドレス照合を行い、正しいデータ/命令がアクセスされているようにする。タグ RAM 4 2 0 は、どのデータ/命令がキャッシュに現在記憶されているかを追跡する。データ RAM 4 3 0 は、データ/命令を記憶する。データ RAM 4 3 0 の各ラインは、キャッシュ 2 5 0 n がデータまたは命令を記憶するように構成されているかどうかに応じて、1 つまたは複数のピクセルあるいは 1 つまたは複数の命令に対するデータを記憶しうる。

【 0 0 3 1 】

キャッシュ 2 5 0 n は、ダイレクトマップキャッシュ構造 (direct map cache structure) または M - ウエイセットアソシエイトキャッシュ構造 (M-way set associate cache structure) で実装されうる。ダイレクトマップキャッシュ構造では、次のレベルにおけるメモリ (例えば、メインメモリ 2 7 0) の各ラインが、データ RAM 4 3 0 の 1 つのラインだけにマップ (mapped) されうる。データ RAM 4 3 0 の各ライン (または各キャッシュライン) は、そのキャッシュラインに記憶されているメインメモリ 2 7 0 の特定のラインを識別するタグ RAM 4 2 0 におけるタグに関連される。1 つの例として、メインメモリ 2 7 0 が 2 5 6 k ラインを有しておりかつデータ RAM 4 3 0 が 4 k ラインを有しているならば、各キャッシュラインは特定の 1 2 ビットアドレスに関連され、そしてメインメモリの 3 2 ラインは各キャッシュラインにマップされる。各キャッシュラインに記憶されているメインメモリ 2 7 0 の特定のラインは、キャッシュラインに対する 1 2 ビットアドレスおよびメインメモリ 2 7 0 の 3 2 ラインのうちどれがキャッシュラインに記憶されているかを示す 5 ビットタグによって識別されうる。M - ウエイセットアソシエイトキャッシュ構造では、メインメモリ 2 7 0 の各ラインは、データ RAM 4 3 0 の M の可能なラインの任意の 1 つにマップされうる。一般に、タグ構造は、メインメモリサイズ、キャッシュサイズ、キャッシュ構造、各キャッシュラインのサイズ、および/または他のファクターで決まりうる。

【 0 0 3 2 】

タグ RAM 4 2 0 は、データ RAM 4 3 0 の各ラインに対するタグを記憶し、データ RAM 4 3 0 における任意のラインが置換されると更新される。各キャッシュラインに対するタグは、そのキャッシュラインに記憶されているメインメモリのラインを示す。制御ユニット 4 1 0 内では、アドレス解析器 (address parser) 4 1 2 が、処理ユニット 2 2 0 によるメモリアクセスのためのアドレスをマルチプレクサ 3 5 0 n から受信し、ターゲットタグおよびキャッシュラインアドレスを得るためにそのアドレスを解析し (parse)、そのキャッシュラインアドレスをタグ RAM 4 2 0 に与え、そしてターゲットタグを比較ロジック 4 1 4 に与える。タグ RAM 4 2 0 は、解析器 4 1 2 からのキャッシュラインアドレスに対する記憶タグを与える。比較ロジック 4 1 4 は、記憶タグをターゲットタグと比較し、それらのタグが等しいならばキャッシュヒットを示し、それらのタグが異なるならばキャッシュミスを示す。タグ RAM 4 2 0 は、キャッシュ 2 5 0 n が M - ウエイセットアソシエイトキャッシュ構造を実装していれば、複数回アクセスされうる。

【 0 0 3 3 】

メモリアクセスユニット 4 1 6 は、キャッシュヒットが存在する場合に、データ RAM 4 3 0 のアクセスに対処する。メモリアクセスが読み出し (read) であれば、キャッシュラインアドレスに記憶されたデータ/命令を読み取るためにアクセスされる。メモリアクセスが書き込み (write) であれば、データ RAM 4 3 0 がキャッシュライン・アドレスに書き込まれ、そのキャッシュラインが書き込まれたことを示すようにキャッシュ「ダーティビット」 ("dirty bit") が更新されて、そのラインは置換される前にメインメモリ 2 7 0 に書き戻されることになる。

【 0 0 3 4 】

インタフェースユニット 3 1 0 k に対する状態機械 3 2 0 は、そのインタフェースユニットに割り当てられたキャッシュに要求データ/命令が記憶されているか否かを判定する。状態機械 3 2 0 内では、マッパー (mapper) 4 5 0 が制御ユニット 2 3 2 から制御 M K を受信し、N 個のキャッシュ 2 5 0 a ~ 2 5 0 n に対する N 個のイネーブル信号を生成す

10

20

30

40

50

る。各イネーブル信号は、関連付けられたキャッシュ250がインタフェースユニット310kに割り当てられるならば論理ハイに設定される。N個のANDゲート452a~452nは、N個のキャッシュ250a~250nそれぞれからN個のイネーブル信号およびN個のヒット/ミスインジケータを受信する。各ANDゲート452は、(1)そのイネーブル信号が、関連付けられたキャッシュ250がインタフェースユニット310kに割り当てられていないことを示す論理ローであれば論理ローを、あるいは、(2)イネーブル信号が、キャッシュがインタフェースユニット319kに割り当てられていることを示す論理ハイであれば関連付けられたキャッシュ250からのヒット/ミスインジケータを与える。ORゲート454は、ANDゲート452a~452nの出力を受信し、インタフェースユニット310kに対するヒットインジケータを生成する。

10

【0035】

キャッシュフィルユニット(cache fill unit)460は、キャッシュミスが存在する場合には、インタフェースユニット310kに割り当てられたキャッシュのうちの1つのキャッシュフィルを指示する。1つの実施の形態では、最近最も使われていない(LRU)、すなわち、最も長期間使われていない割り当てられたキャッシュがキャッシュフィルのために選択される。キャッシュフィルユニット460は、選択されたキャッシュのすべてまたは一部分をメインメモリ270からの新しいデータ/命令で満たしうる。ターゲットキャッシュラインがメインメモリ270から満たされかつ選択されたキャッシュのタグRAMがデータRAMにロードされた新しいデータ/命令を反映するために更新される間、インタフェースユニット310kは機能停止(stalled)になされうる。

20

【0036】

完全に構成可能なキャッシュを有するグラフィックスシステムは、処理ユニットにキャッシュを割り当てることに関して最大の柔軟性を提供する。所定の処理ユニットは、処理ユニットのデータ要件に依存して、利用可能なキャッシュのうちのゼロ、1つ、複数、あるいはすべてを割り当てられうる。利用可能なキャッシュは、キャッシュの良好な利用を実現しかつ実行されているグラフィックスアプリケーションとは関係なしにグラフィックスプロセッサにおける任意の1つの場所におけるメモリアクセスボトルネックを回避する態様で処理ユニットに割り当てられうる。利用可能なキャッシュは、下記のように、処理ユニットに知的に割り当てられうる。

30

【0037】

多くの場合において、部分的に構成可能なキャッシュで高性能が達成されうる。これらのキャッシュは、各キャッシュが処理ユニットの部分集合のうちの1つに割り当てられうる点で部分的に構成可能である。異なるグラフィックス動作に対してキャッシュを使用する利益を確認するために種々のグラフィックスアプリケーションについて研究が行われてきた。これらの研究は、多くのグラフィックスアプリケーションに対してステンシルおよびデプステストとテクスチャマッピングの性能を高めるのにキャッシュが非常に有用であり、カラーおよびシェーダに対してはある程度だけしか有益でないことを示している。

【0038】

したがって、キャッシュメモリシステムは、ステンシルおよびデプステストとテクスチャマッピングに対しては利用可能なキャッシュの多くを構成可能とし、カラーおよびシェーダに対してはキャッシュのうちのほんの僅かを構成可能として、デザインされうる。部分的に構成可能なキャッシュを用いることは、構成可能なキャッシュを割り当てる制御ユニットならびに処理ユニットをそれらの割り当てられたキャッシュに結合するクロスバーのデザインを単純化しうる。

40

【0039】

図5は、構成可能なキャッシュおよび専用キャッシュを有するグラフィックスシステム500の1つの実施の形態を示している。グラフィックスシステム500は、グラフィックスプロセッサ510およびキャッシュメモリシステム530を含んでいる。グラフィックスプロセッサ510は、図1における処理ユニット120a~120iに対して上述したように動作する処理ユニットを含んでいてもよい。具体的には、グラフィックスプロセ

50

ッサ510は、ステンシルおよびデプステストエンジン、テクスチャマッピングエンジン、フラグメントシェーダ、およびキャッシュに記憶されたデータおよび/または命令を利用するフラグメントエンジンを含んでいてもよい。

【0040】

図5に示された実施の形態では、キャッシュメモリシステム530は、R個の部分的に構成可能なキャッシュ550a~550rおよび1つの専用キャッシュ550sを含んでおり、一般にR=1である。表1は、キャッシュ550a~550sのそれぞれの可能な割り当てを示しており、各"X"は有効なキャッシュ割り当てを示している。表1に示されているように、P個の部分的に構成可能なキャッシュ550a~550pのそれぞれは、ステンシルテスト、デプステスト、またはテクスチャマッピングに対して割り当てられる。部分的に構成可能なキャッシュ550qは、シェーダ命令、ステンシルテスト、デプステスト、またはテクスチャマッピングに対して割り当てられる。部分的に構成可能なキャッシュ550rは、シェーダデータ、ステンシルテスト、デプステスト、またはテクスチャマッピングに対して割り当てられる。専用キャッシュ550sは、フラグメントエンジンに対するカラー値を記憶するために用いられる。1つの実施の形態では、P=8であり、キャッシュ550qおよび550rがシェーダ命令およびデータに対しては使用されないと仮定して、10個までのキャッシュがステンシルテスト、デプステスト、および/またはテクスチャマッピングに対して割り当てられる。キャッシュメモリシステム530に対しては、より少数のまたは追加のキャッシュが用いられてもよい。

【表1】

表1

処理ユニット	キャッシュ 550a	...	キャッシュ 550p	キャッシュ 550q	キャッシュ 550r	キャッシュ 550s
ステンシルテスト	X	...	X	X	X	
デプステスト	X	...	X	X	X	
テクスチャマッピング	X	...	X	X	X	
シェーダ命令				X		
シェーダデータ					X	
カラー						X

【0041】

制御ユニット532は、処理ユニットによるメモリ利用に基づいて、構成可能なキャッシュ550a~550rをステンシルおよびデプステストエンジン、テクスチャエンジン、および/またはフラグメントシェーダに割り当てる。クロスバー540は、処理ユニットをそれらの割り当てられたキャッシュ550a~550rに結合する。アービター560は、キャッシュ550a~550sをメインメモリ570に結合する。

【0042】

図5はまた、クロスバー540の実施の形態も示している。この実施の形態では、クロスバー540は、ステンシルテスト、デプステスト、テクスチャマッピング、シェーダ命令、およびシェーダデータのそれぞれに対するインタフェースユニット580a~580eを含んでいる。インタフェースユニット580aは、ステンシルおよびデプステストエンジンをキャッシュ580a~580rのうちのステンシルテストに対して割り当てられたキャッシュに結合する。インタフェースユニット580bは、ステンシルおよびデプステストエンジンをキャッシュ580a~580rのうちのデプステストに対して割り当て

られたキャッシュに結合する。インタフェースユニット580cは、テクスチャマッピングエンジンをキャッシュ580a~580rのうちの割り当てられたキャッシュに結合する。インタフェースユニット580dは、キャッシュ550qが割り当てられておれば、フラグメントシェーダをこのキャッシュ550qに結合する。インタフェースユニット580eは、キャッシュ550rが割り当てられておれば、フラグメントシェーダをこのキャッシュ550rに結合する。各インタフェースユニット580は、例えば図3に示されているように、状態機械およびI/Oユニットを含んでいてもよい。簡単のために、インタフェースユニットに対するマルチプレクサおよびキャッシュに対するマルチプレクサは図5には図示されていない。

【0043】

10

図5は、構成可能なキャッシュおよび専用キャッシュを有するキャッシュメモリシステムの特典の実施の形態を示している。一般に、キャッシュメモリシステムは、任意の個数の構成可能なキャッシュおよび任意の個数の専用キャッシュを含みうる。構成可能なキャッシュは、任意の組のグラフィックス動作に対して使用され、専用キャッシュも任意の組のグラフィックス動作に対して使用されうる。構成可能なキャッシュおよび専用キャッシュの両方を使用すると、制御ユニットおよびクロスバーのデザインを単純化できる。

【0044】

20

図2および5に示された実施の形態では、制御ユニットは、種々の態様で構成可能なキャッシュを処理ユニットに割り当てることができる。1つの実施の形態では、制御ユニットは、先行する画像/フレームにおけるメモリ利用に基づいて、レンダリングされるべき各グラフィックス画像またはフレームに対するキャッシュを割り当てる。制御ユニットは、各処理ユニットによってなされたメモリアクセスの数、各キャッシュに対するキャッシュヒットの数、各キャッシュに対するキャッシュミス数、等をカウントすることによって、メモリ利用を確認しうる。制御ユニットは、つぎにメモリ利用の高い処理ユニットに対してより多くのキャッシュを割り当てることができ、そしてメモリ利用の低い処理ユニットに対してはより少ないまたはゼロのキャッシュを割り当てることができる。

【0045】

30

他の実施の形態では、制御ユニットは、グラフィックス画像/フレームのコヒレンシ(coherency)に基づいて、構成可能なキャッシュを処理ユニットに割り当てる。コヒレンシは、連続した2D/3Dフレームにおける変化の量に関連する。フレームが同様のコンテンツをレンダリングする可能性が大きい場合には、より高いコヒレンシに対してはより少ないキャッシュが割り当てられ、より低いコヒレンシに対してはより多いキャッシュが割り当てられうる。

【0046】

40

さらに他の実施の形態では、制御ユニットは、実行されているグラフィックスアプリケーションの特徴に基づいて、構成可能なキャッシュを処理ユニットに割り当てる。例えば、グラフィックスシステムが2-Dグラフィックスアプリケーションを実行しているならば、デブテストは必要とされなくてもよく、デブテストに対してはキャッシュは割り当てられなくてもよい。対極的に、3-DグラフィックスアプリケーションがZパスだけを用いるならば、すべての構成可能なキャッシュがデブテストに割り当てられうる。他の例としては、単純なシェーダプログラムが実行されておりかつシェーダ命令のすべてがフラグメントシェーダ内に記憶されうるならば、フラグメントシェーダにはキャッシュは割り当てられなくてもよい。

【0047】

さらに他の実施の形態では、制御ユニットは、構成可能なキャッシュを動的に割り当てる。例えば、制御ユニットは、データ要求が処理ユニットに送られる場合には、処理ユニットに1つまたは複数の処理ユニットを割り当てることができる。制御ユニットは、要求の数、キャッシュのヒット/ミスの統計、および/または即時に判定されうる他の要因に基づいて、処理ユニットに割り当てられるキャッシュの数を調節しうる。1つの処理ユニットから他の処理ユニットに切り替えられたまたは再割り当てされたキャッシュに対して

50

、フラッシュ動作および無効化動作 (flush and invalidate operations) がすぐさま行われうる。

【0048】

さらに他の実施の形態では、制御ユニットは、静的および動的割り当ての組み合わせを用いて、キャッシュを割り当てる。例えば、制御ユニットは、例えば、処理ユニットによるメモリ利用についての現在の統計に基づいて、フレーム、画像、またはバッチのレンダリングの開始において、所定の処理ユニットに1つまたは複数のキャッシュを再割り当てすることができる。制御ユニットは、フレーム、画像、またはバッチのレンダリング時に、この処理ユニットに対するキャッシュ割り当てを、例えば周期的に調節することができる。各レンダリング周期における新しいキャッシュ割り当ては、先行するレンダリング周期の間に得られた統計に基づいて決定されうる。

10

【0049】

制御ユニットは、他の基準に基づいて、構成可能なキャッシュを割り当てることもできる。制御ユニットは、グラフィックスアプリケーションが最初に行われる場合に、および/または他の時点で、メモリ利用の変化が検知されると、各画像/フレーム内でキャッシュを割り当てることができる。

【0050】

簡単のために、図1~5は、キャッシュのバンクを含むキャッシュメモリシステムを示している。キャッシュはメモリの1つのブロックで実装されてもよい。キャッシュまたはキャッシュメモリシステムはまた、例えば、レベル1、レベル2、レベル3、等の複数のレベルを有する階層構造で実装されうる。低いレベル(例えば、レベル1)のキャッシュは、高いレベル(例えば、レベル2)のキャッシュより高速であるが小さくなる傾向がある。各レベルにおけるキャッシュは、キャッシュミスが生じると、次に高いレベルのキャッシュで満たされうる。レベルの数、各レベルにおけるキャッシュの数、および異なるレベルに対するキャッシュサイズは、固定であるかあるいは構成可能でありうる。例えば、レベルの数、各レベルのキャッシュの数、および/またはキャッシュサイズは、良好な性能を実現するために選択されうるものであり、かつグラフィックスアプリケーションの特徴に基づいて構成可能でありうる。

20

【0051】

図6は、構成可能なキャッシュを有するグラフィックスシステムを動作させるためのプロセス600の1つの実施の形態を示す。グラフィックス画像をレンダリングするためにグラフィックス動作を行うように構成された複数の処理ユニットによるメモリ利用が判定される(ブロック612)。メモリ利用は、処理ユニットによってなされたメモリアクセスをモニタすることにより、実行されているグラフィックスアプリケーションの特徴を確認することにより、そして/または他の態様で判定されうる。複数のキャッシュが、メモリ利用に基づいて、複数の処理ユニットのうち少なくとも1つの処理ユニットに割り当てられる(ブロック614)。各処理ユニットは、(1)その処理ユニットならびに他の処理ユニットによるメモリ利用および(2)その処理ユニットの割り当てのために利用可能なキャッシュに依存して、キャッシュのうちゼロ、1つ、複数、またはすべてを割り当てられうる。メモリ利用は、例えば、処理ユニットによるデータ要求、キャッシュヒット/ミス統計、等のような種々の統計に基づいて確認されうる。キャッシュ割り当ては、メモリ利用の変化が検知された場合等に、周期的に(例えば、グラフィックス画像/フレームごとに)行われうる。キャッシュはまた、メモリ利用の検知された変化に基づいて、画像/フレームのレンダリング時に再割り当てされうる。少なくとも1つの処理ユニットのそれぞれは、その処理ユニットに割り当てられたキャッシュの各組に、各インタフェースユニットを経由して結合される(ブロック616)。

30

40

【0052】

ここに記載されたグラフィックスシステムおよび構成可能なキャッシュは、ワイヤレスコミュニケーション、コンピューティング、ネットワークング、パーソナルエレクトロニクス、等のために使用されうる。ワイヤレスコミュニケーションに対する、構成可能なキ

50

キャッシュを有するグラフィックスシステムの好例の用途について以下で説明する。

【 0 0 5 3 】

図 7 は、ワイヤレスコミュニケーションシステムにおけるワイヤレスデバイス 7 0 0 の実施の形態のブロック図を示している。ワイヤレスデバイス 7 0 0 は、携帯電話、端末装置、ハンドセット、パーソナルデジタルアシスタント (P D A)、あるいは他のデバイスでありうる。ワイヤレスコミュニケーションシステムは、コード分割多重アクセス (C D M A) システム、グローバルシステムフォアモバイルコミュニケーションズ (Global System for Mobile Communications) (G S M) システム、あるいは他のシステムでありうる。

【 0 0 5 4 】

ワイヤレスデバイス 7 0 0 は、受信パスと送信パスを經由して双方向コミュニケーションを提供することができる。受信パスでは、ベースステーションによって送信された信号がアンテナ 7 1 2 によって受信され、レシーバ (R C V R) 7 1 4 に与えられる。レシーバ 7 1 4 は、受信信号を調整しかつデジタル化し、さらなる処理のためにサンプリングをデジタルセクション 7 2 0 に与える。送信パスでは、トランスミッタ (T M T R) 7 1 6 は、デジタルセクション 7 2 0 から送信さるべきデータを受信し、そのデータを処理しかつ調整し、アンテナを經由してベースステーションに送信される変調信号を生成する。

【 0 0 5 5 】

デジタルセクション 7 2 0 は、例えば、モデムプロセッサ 7 2 2、ビデオプロセッサ 7 2 4、アプリケーションプロセッサ 7 2 6、ディスプレイプロセッサ 7 2 8、コントローラ / プロセッサ 7 3 0、グラフィックスプロセッサ 7 4 0、および外部バスインタフェース (E B I) 7 6 0 のような種々の処理ユニットおよびインタフェースユニットを含んでいる。モデムプロセッサ 7 2 2 は、データ送信および受信のための処理 (例えば、符号化、変調、復調、および復号化) を行う。ビデオプロセッサ 7 2 4 は、カムコーダ、ビデオプレイバック、およびビデオカンファレンスのようなビデオアプリケーションのためのビデオコンテンツ (例えば、スチル画像、ムービングビデオ、およびムービングテキスト) に対する処理を行う。アプリケーションプロセッサ 7 2 6 は、マルチウェイコール (multi-way calls)、ウェブブラウジング (web-browsing)、メディアプレイヤー (media player)、およびユーザインタフェース (user interface) のような種々のアプリケーションに対する処理を行う。ディスプレイプロセッサ 7 2 8 は、ディスプレイユニット 7 8 0 におけるビデオ、グラフィックスおよびテキストの表示を容易にするための処理を行う。コントローラ / プロセッサ 7 3 0 は、デジタルセクション 7 2 0 内の種々の処理ユニットおよびインタフェースユニットの動作を指示する。

【 0 0 5 6 】

グラフィックスプロセッサ 7 4 0 は、グラフィックスアプリケーションに対する処理を行い、かつ下記のように実装されうる。キャッシュメモリシステム 7 5 0 は、グラフィックスプロセッサ 7 4 0 に対するデータおよび / または命令を記憶し、かつ構成可能なキャッシュおよび場合により専用のキャッシュで実装されうる。キャッシュメモリシステム 7 5 0 は、グラフィックスプロセッサ 7 4 0 内の処理ユニットに構成可能なキャッシュを結合するクロスバーと、バス 7 3 2 および E B I 7 6 0 を經由してメインメモリ 7 7 0 にキャッシュを結合するアービターをさらに含む。E B I 7 6 0 は、デジタルセクション 7 2 0 (例えば、キャッシュ) とメインメモリ 7 7 0 との間のデータの転送を容易にする。

【 0 0 5 7 】

デジタルセクション 7 2 0 は、1つまたは複数のデジタル信号プロセッサ (D S P s)、マイクロプロセッサ、縮小命令セットコンピュータ (R I S C s)、等で実装されうる。デジタルセクション 7 2 0 はまた、1つまたは複数の用途特定集積回路 (A S I C s) または他のタイプの集積回路 (I C s) 上に作成されてもよい。

【 0 0 5 8 】

10

20

30

40

50

ここに記載されたグラフィックスシステムおよび構成可能なキャッシュは、種々のハードウェアユニットで実装されうる。例えば、グラフィックスシステムおよび構成可能なキャッシュは、ASICs、デジタル信号処理デバイス(DSPDs)、プログラマブルロジックデバイス(PLDs)、フィールドプログラマブルゲートアレイ(FPGAs)、プロセッサ、コントローラ、マイクロコントローラ、マイクロプロセッサ、および他の電子ユニットで実装されうる。

【0059】

グラフィックスシステムのある部分は、ファームウェアおよび/またはソフトウェアで実装されうる。例えば、制御ユニットは、ここで説明された機能を行うファームウェアおよび/またはソフトウェアモジュール(例えば、手順、機能、等)で実装されうる。ファームウェアおよび/またはソフトウェアコードは、メモリ(例えば、図7におけるメモリ770)に記憶され、プロセッサ(例えば、プロセッサ730)によって実行されうる。メモリは、プロセッサの内部またはプロセッサの外部で実装されうる。

10

【0060】

開示された実施の形態についての前記の記述は、当業者が本発明を実施または利用できるようにするために提示された。これらの実施の形態に対する種々の修正が当業者には容易に明らかとなるであろう、そしてここに明示された一般的原理は、本発明の精神または範囲から逸脱することなしに他の実施の形態に適用されうる。したがって、本発明は、ここに示された実施の形態に限定されるべく意図されたものではなく、ここに開示された原理および新規な特徴に整合した最も広い範囲を与えられるべきである。

20

【図面の簡単な説明】

【0061】

【図1】専用キャッシュを有するグラフィックスシステムを示す。

【図2】完全に構成可能なキャッシュを有するグラフィックスシステムを示す。

【図3】図2のグラフィックスシステムにおけるクロスバーを示す。

【図4】処理ユニットのためのキャッシュおよび状態機械を示す。

【図5】構成可能なキャッシュおよび専用キャッシュを有するグラフィックスシステムを示す。

【図6】構成可能なキャッシュを有するグラフィックスシステムを動作させるためのプロセスを示す。

30

【図7】ワイヤレスコミュニケーションシステムにおけるワイヤレスデバイスを示す。

【符号の説明】

【0062】

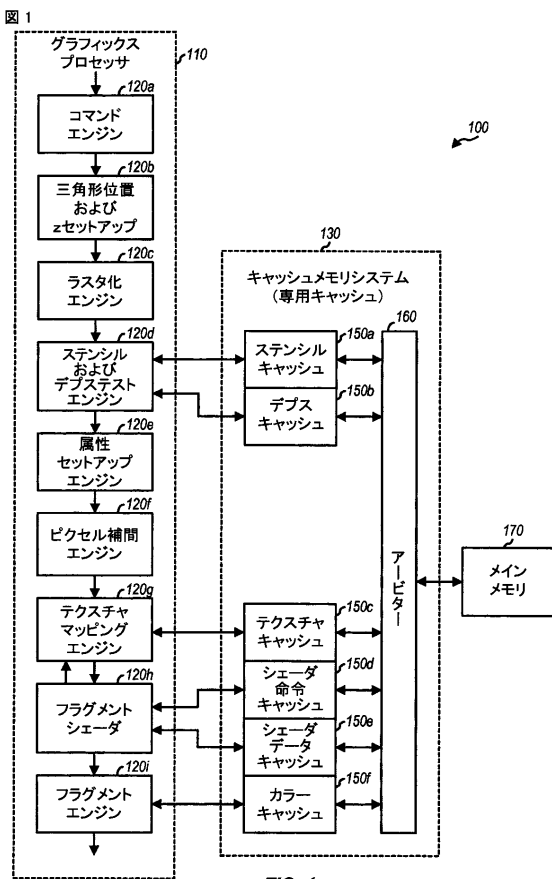
- 100 グラフィックスシステム
- 110 グラフィックスプロセッサ
- 130 キャッシュメモリシステム
- 170 メインメモリ
- 200 グラフィックスシステム
- 210 グラフィックスプロセッサ
- 230 キャッシュメモリシステム
- 270 メインメモリ
- 320 状態機械
- 232 制御ユニット
- 240 クロスバー
- 500 グラフィックスシステム
- 510 グラフィックスプロセッサ
- 530 キャッシュメモリシステム
- 570 メインメモリ
- 600 プロセス
- 700 ワイヤレスアンテナ

40

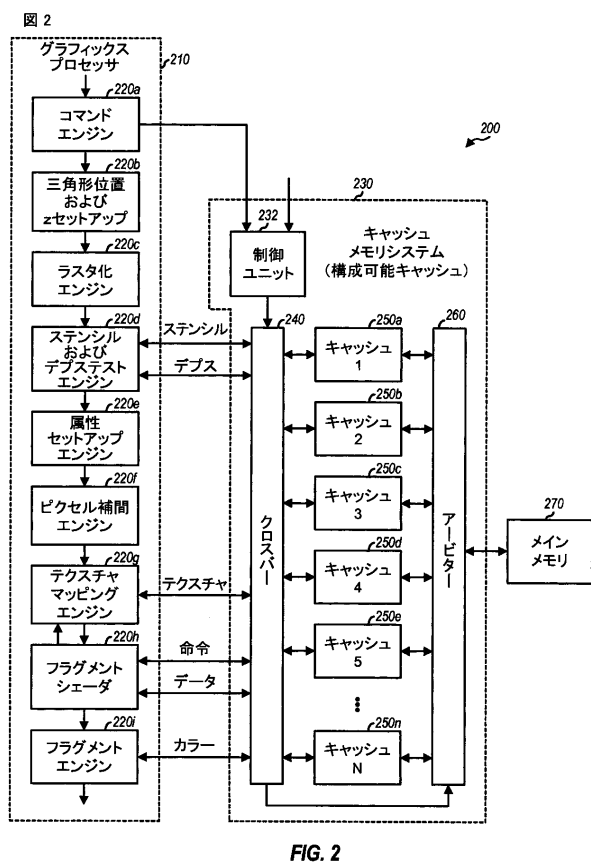
50

- 7 1 2 アンテナ
- 7 1 4 レシーバ
- 7 1 6 トランスミッタ
- 7 2 0 デジタルセクション
- 7 7 0 メインメモリ
- 7 8 0 デ스플레이ユニット

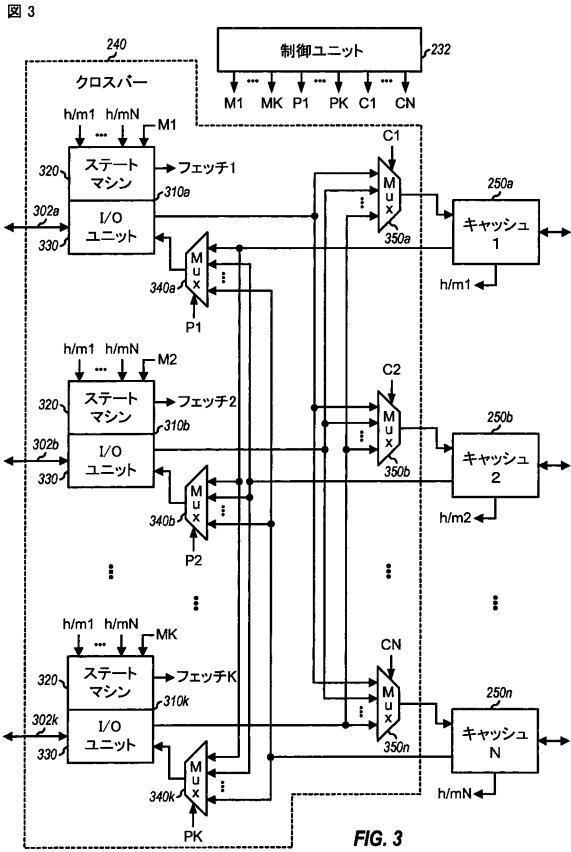
【 図 1 】



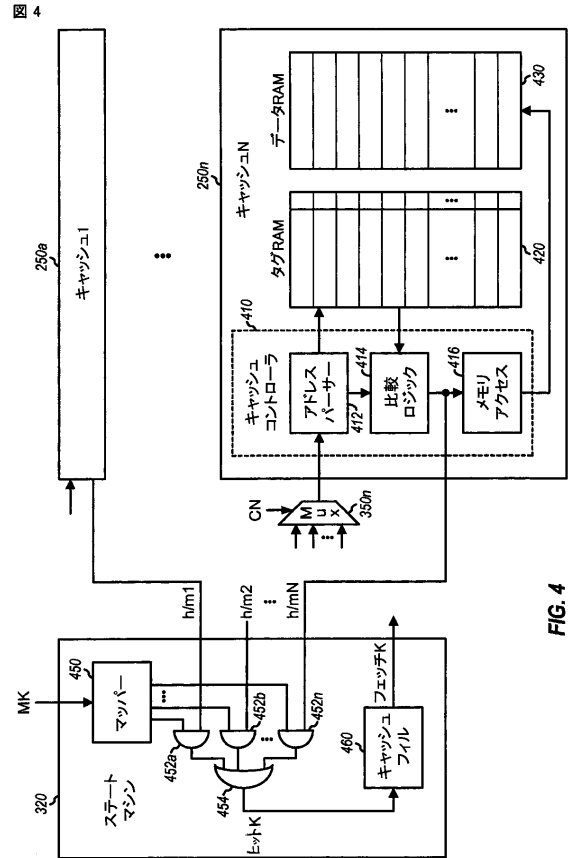
【 図 2 】



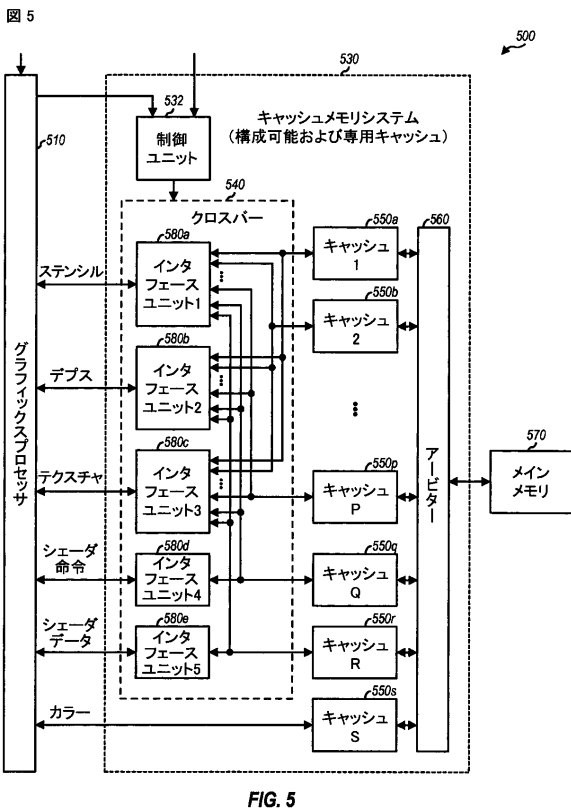
【 図 3 】



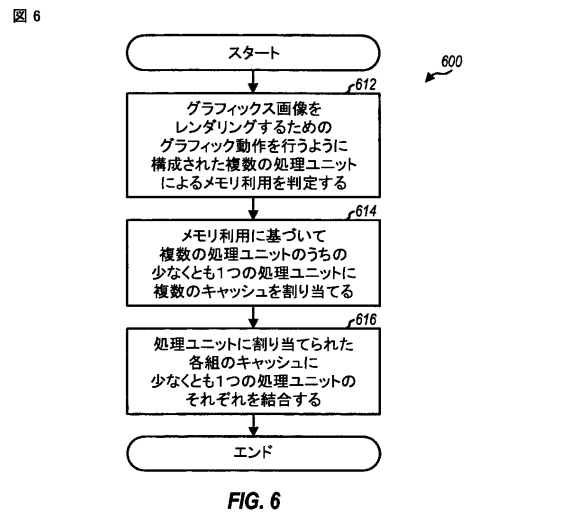
【 図 4 】



【 図 5 】



【 図 6 】



【 図 7 】

図 7

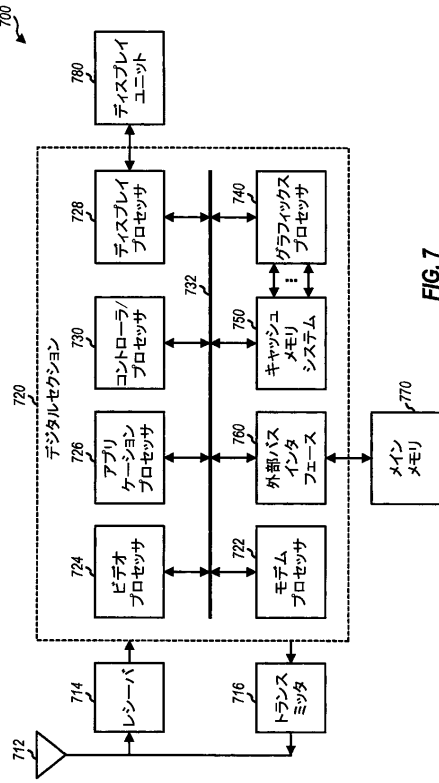


FIG. 7

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/067325

A. CLASSIFICATION OF SUBJECT MATTER INV. G06T11/60 G06T15/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06T		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BALMER K ET AL: "A single chip multimedia video processor" CUSTOM INTEGRATED CIRCUITS CONFERENCE, 1994., PROCEEDINGS OF THE IEEE 1994 SAN DIEGO, CA, USA 1-4 MAY 1994, NEW YORK, NY, USA, IEEE, 1 May 1994 (1994-05-01), pages 91-94, XP010129831 ISBN: 0-7803-1886-2 abstract; figure 1 paragraphs [0003], [0006], [0008]	1-34
X	US 2003/167379 A1 (SOLTIS DONALD CHARLES [US]) 4 September 2003 (2003-09-04) abstract; figures 3,7 paragraphs [0001], [0003] - [0005], [0021], [0026], [0028]	1-34
-/--		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> See patent family annex.
* Special categories of cited documents :		
A document defining the general state of the art which is not considered to be of particular relevance	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
E earlier document but published on or after the international filing date	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
O document referring to an oral disclosure, use, exhibition or other means	*Z* document member of the same patent family	
P document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 7 September 2007	Date of mailing of the international search report 18/09/2007	
Name and mailing address of the ISA/ European Patent Office, P.B. 5618 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Casteller, Maurizio	

INTERNATIONAL SEARCH REPORT

International application No PCT/US2007/067325

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 480 941 B1 (FRANKE HUBERTUS [US] ET AL) 12 November 2002 (2002-11-12) abstract; claim 1 column 1, line 39 - column 2, line 6 column 4, lines 33-46,64 - column 5, line 8 column 6, lines 26-34	1-34
A	EP 0 917 056 A (DIGITAL EQUIPMENT CORP [US] COMPAQ COMPUTER CORP [US]) 19 May 1999 (1999-05-19) abstract paragraphs [0210] - [0212]	1-34

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No
PCT/US2007/067325

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2003167379 A1	04-09-2003	FR 2836732 A1	05-09-2003
US 6480941 B1	12-11-2002	NONE	
EP 0917056 A	19-05-1999	JP 2000132530 A	12-05-2000

フロントページの続き

(81) 指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

1. GSM

- (74) 代理人 100109830
弁理士 福原 淑弘
- (74) 代理人 100075672
弁理士 峰 隆司
- (74) 代理人 100095441
弁理士 白根 俊郎
- (74) 代理人 100084618
弁理士 村松 貞男
- (74) 代理人 100103034
弁理士 野河 信久
- (74) 代理人 100119976
弁理士 幸長 保次郎
- (74) 代理人 100153051
弁理士 河野 直樹
- (74) 代理人 100140176
弁理士 砂川 克
- (74) 代理人 100100952
弁理士 風間 鉄也
- (74) 代理人 100101812
弁理士 勝村 紘
- (74) 代理人 100070437
弁理士 河井 将次
- (74) 代理人 100124394
弁理士 佐藤 立志
- (74) 代理人 100112807
弁理士 岡田 貴志
- (74) 代理人 100111073
弁理士 堀内 美保子
- (74) 代理人 100134290
弁理士 竹内 将訓
- (74) 代理人 100127144
弁理士 市原 卓三
- (74) 代理人 100141933
弁理士 山下 元
- (72) 発明者 ユ、チュン
アメリカ合衆国、カリフォルニア州 92131、サン・ディエゴ、サイプレス・ウッズ・ドライブ 11496
- (72) 発明者 ジャオ、グオファン

アメリカ合衆国、カリフォルニア州 9 2 1 3 0、サン・ディエゴ、ハンターズ・グレン・ドライ
ブ 1 0 6 8 0

(72)発明者 ドウ、ユン

アメリカ合衆国、カリフォルニア州 9 2 1 2 9、サン・ディエゴ、カティデイド・サークル 1
2 3 4 1

Fターム(参考) 5B080 CA08 GA02 GA11 GA22