

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4416694号
(P4416694)

(45) 発行日 平成22年2月17日(2010.2.17)

(24) 登録日 平成21年12月4日(2009.12.4)

(51) Int.Cl. F I
G06F 12/06 (2006.01) G O 6 F 12/06 5 5 O A
G06F 13/28 (2006.01) G O 6 F 13/28 3 1 O B

請求項の数 12 (全 17 頁)

(21) 出願番号	特願2005-139666 (P2005-139666)	(73) 特許権者	395015319
(22) 出願日	平成17年5月12日(2005.5.12)		株式会社ソニー・コンピュータエンタテインメント
(65) 公開番号	特開2006-318178 (P2006-318178A)		東京都港区南青山二丁目6番21号
(43) 公開日	平成18年11月24日(2006.11.24)	(74) 代理人	100105924
審査請求日	平成18年5月11日(2006.5.11)		弁理士 森下 賢樹
		(72) 発明者	大塚 活志
			東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内
		(72) 発明者	佐々木 伸夫
			東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

最終頁に続く

(54) 【発明の名称】 データ転送調停装置およびデータ転送調停方法

(57) 【特許請求の範囲】

【請求項1】

複数バンクを有するメモリにデータを記録するメモリコントローラに対し、記録されるべきデータを順次転送するための装置であって、

データ転送要求主体に対する転送サービスの優先順位とは関係なく、複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択する選択部と、

前記選択されたデータ転送要求主体により転送要求されているデータを前記メモリコントローラに送信する送信部と、を備え、

前記選択部は、同一のデータ転送要求主体に対する転送サービスが連続的に実行されるように前記データ転送要求主体を連続的に選択し、かつ、その複数回の転送サービスによって前記メモリのバンクをまたぐ転送が発生するようにその連続選択回数を定め、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記選択部は、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停装置。

【請求項2】

複数バンクを有するメモリにデータを記録するメモリコントローラに対し、記録されるべきデータを順次転送するための装置であって、

複数のデータ転送要求主体について定められた転送サービスの優先順位にしたがって、前記複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択する選択部と、

前記選択されたデータ転送要求主体により転送要求されているデータを前記メモリコントローラに送信する送信部と、を備え、

前記選択部は、転送サービスを受けたあとのデータ転送要求主体が更なるデータの転送を要求している場合には、前記優先順位にしたがって次に転送サービスを受けるべきデータ転送要求主体が存在しても前記転送サービスを受けたあとのデータ転送要求主体を連続的に選択し、その複数回の転送サービスによって前記メモリのバンクをまたぐ転送が発生するようにその連続選択回数を定め、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記選択部は、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停装置。

【請求項3】

前記バンクのエントリビット数と前記データ転送要求主体に対する1回の転送サービスにより送信されるデータのビット数は同一であることを特徴とする請求項1または2に記載のデータ転送調停装置。

【請求項4】

前記選択部は、ラウンドロビン方式、LRU(Least Recently Used)方式およびランダム方式の少なくともいずれかの選択方式にて、データ転送要求主体を選択することを特徴とする請求項1に記載のデータ転送調停装置。

【請求項5】

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスする方法であって、

データ転送要求主体に対する転送サービスの優先順位とは関係なく、同一のデータ転送要求主体に対する転送サービスを所定の回数連続して行い、かつ、その複数の転送サービスによって前記メモリのバンクをまたぐ転送が発生するよう前記所定の回数が定められ、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停方法。

【請求項6】

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスする方法であって、

前記複数のデータ転送要求主体について定められた転送サービスの優先順位にしたがって、前記複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択して転送サービスを行い、かつ、その転送サービスを受けたあとのデータ転送要求主体が更なるデータの転送を要求している場合には、前記優先順位にしたがって次に転送サービスを受けるべきデータ転送要求主体が存在しても前記転送サービスを受けたあとのデータ転送要求主体に対する転送サービスを連続して行い、その複数回の転送サービスによって前記メモリのバンクをまたぐ転送が発生するようにその連続選択回数が定められ、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記デー

10

20

30

40

50

タ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停方法。

【請求項7】

請求項5または6に記載の方法において、前記バンクのエントリビット数と1回の転送サービスで転送されるデータのビット数は同一であり、同一のデータ転送要求主体に対する転送サービスを2回以上連続して行うことを特徴とするデータ転送調停方法。

【請求項8】

請求項5に記載の方法において、転送サービスの対象となるデータ転送要求主体は、ラウンドロビン方式、LRU方式およびランダム方式の少なくともいずれかの選択方式にて選択されることを特徴とするデータ転送調停方法。

【請求項9】

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスするためのコンピュータプログラムであって、

データ転送要求主体に対する転送サービスの優先順位とは関係なく、同一のデータ転送要求主体に対する転送サービスを所定の回数連続して行い、かつ、その複数の転送サービスによって前記メモリのバンクをまたぐ転送が発生するよう前記所定の回数が定められ、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停プログラム。

【請求項10】

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスするためのコンピュータプログラムであって、

前記複数のデータ転送要求主体について定められた転送サービスの優先順位にしたがって、前記複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択して転送サービスを行い、かつ、その転送サービスを受けたあとのデータ転送要求主体が更なるデータの転送を要求している場合には、前記優先順位にしたがって次に転送サービスを受けべきデータ転送要求主体が存在しても前記転送サービスを受けたあとのデータ転送要求主体に対する転送サービスを連続して行い、その複数回の転送サービスによって前記メモリのバンクをまたぐ転送が発生するようその連続選択回数が定められ、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停プログラム。

【請求項11】

コンピュータにて読み取り可能な記録媒体であって、

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスするためのコンピュータプログラムであって、データ転送要求主体に対する転送サービスの優先順位とは関係なく、同一のデータ転送要求主体に対する転送サービスを所定の回数連続して行い、かつ、その複数の転送サービスによって前記メモリのバンクをまたぐ転送が発生するよう前記所定の回数が定められ、

10

20

30

40

50

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停プログラム

を格納した記録媒体。

【請求項12】

コンピュータにて読み取り可能な記録媒体であって、

複数バンクを有するメモリに対し、複数のデータ転送要求主体から並行してデータ転送要求が発行されたとき、それらデータ転送要求主体に対して順次転送をサービスするためのコンピュータプログラムであって、前記複数のデータ転送要求主体について定められた転送サービスの優先順位にしたがって、前記複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択して転送サービスを行い、かつ、その転送サービスを受けたあとのデータ転送要求主体が更なるデータの転送を要求している場合には、前記優先順位にしたがって次に転送サービスを受けるべきデータ転送要求主体が存在しても前記転送サービスを受けたあとのデータ転送要求主体に対する転送サービスを連続して行い、その複数回の転送サービスによって前記メモリのバンクをまたぐ転送が発生するようにその連続選択回数が定められ、

1回の転送サービスにて転送されるデータのビット数と前記連続選択回数との積が、予め定められたストライド長以下となるよう前記連続選択回数は定められており、前記データ転送要求主体によるデータ転送先アドレスを先読みし、最大ストライド長をバンクの個数とエントリビット数の積とした場合に、先読みされたデータ転送先アドレスが全バンク1周分のアドレスの中で分散している場合は、ストライド長を最大ストライド長よりも短く設定することを特徴とするデータ転送調停プログラム

を格納した記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、メモリにデータを転送するための技術、特に、複数バンクを有するメモリにデータを転送するための技術、に関する。

【背景技術】

【0002】

近年のコンピュータゲームやデジタル放送などの分野に利用されるコンピュータグラフィックス技術や画像処理技術の著しい進歩に伴い、コンピュータ、ゲーム機器、テレビなどの情報処理装置はより高精細の画像データをより高速に処理する能力を求められている。そのためには、演算処理自体の高速化が必要であることは勿論であるが、演算結果や演算のためのデータのメモリアクセスを効率的に実行することも同様に重要である。

【0003】

メモリアクセスを効率よく実行するための方法としてメモリアンタリーブ (Memory Interleaving) とよばれる技術がある。メモリアンタリーブ方式の前提として、メモリはメモリコントローラによる管理単位であるバンクを複数個備える。通常、バンクへアクセスしたときには、実際に読み書きするための時間であるアクセス時間の後に、次のアクセスが可能になるまでの遅延時間が発生する。この遅延時間は、DRAM (Dynamic Random Access Memory) 特有のプリチャージ等の制約により発生する。

メモリアンタリーブ方式では、あるバンクの遅延時間中に次のバンクへのアクセスを開始するので、遅延時間の影響が緩和され、結果としてメモリアクセスが高速化されるというメリットがある。

【発明の開示】

【発明が解決しようとする課題】

10

20

30

40

50

【0004】

しかし、従来のメモリアクセスマニピュレーション方式においても、特定のバンクにアクセス要求が集中することがある。このときには、アクセス要求が集中したバンクの遅延時間の影響によって、情報処理装置全体としての処理パフォーマンスが低下することになる。

【0005】

本発明は、このような課題に鑑みてなされたものであり、その目的は、メモリへのアクセスを効率的に実行するための技術、を提供することにある。

【課題を解決するための手段】

【0006】

本発明のある態様は、複数バンクを有するメモリにデータを記録するメモリコントローラに対し、記録されるべきデータを順次転送するためのデータ転送調停装置である。 10

この装置は、データ転送要求主体に対する転送サービスの優先順位とは関係なく、複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択して、そのデータ転送要求主体により転送要求されているデータをメモリコントローラに送信する。このとき、同一のデータ転送要求主体に対する転送サービスが連続的に実行されるようにデータ転送要求主体を連続的に選択し、かつ、その複数回の転送サービスによってメモリのバンクをまたぐ転送が発生するようにその連続選択回数を定める。

【0007】

ここでいう「優先順位とは関係なく」とは、他のデータ転送要求主体よりも特定のデータ転送要求主体に対して転送サービスを優先的に提供するように予め設定するのではなく、各データ転送要求主体に対して転送サービスを提供する機会が結果として均等化されるものであればよい。 20

【0008】

また、本発明の別の態様も、複数バンクを有するメモリにデータを記録するメモリコントローラに対し、記録されるべきデータを順次転送するためのデータ転送調停装置である。

この装置は、複数のデータ転送要求主体について定められた転送サービスの優先順位にしたがって、それら複数のデータ転送要求主体の中からいずれかのデータ転送要求主体を選択して、そのデータ転送要求主体により転送要求されているデータをメモリコントローラに送信する。このとき、転送サービスを受けたあとのデータ転送要求主体が更なるデータの転送を要求している場合には、優先順位にしたがって次に転送サービスを受けるべきデータ転送要求主体が存在してもその転送サービスを受けたあとのデータ転送要求主体を連続的に選択し、その複数回の転送サービスによってメモリのバンクをまたぐ転送が発生するようにその連続選択回数を定める。 30

【0009】

ここでいう「優先順位」は、データ転送処理の実行中において複数のデータ転送要求主体間において可動であってもよい。

【0010】

なお、本発明を方法、システム、記録媒体、コンピュータプログラムにより表現したのもまた、本発明の態様として有効である。 40

【発明の効果】

【0011】

本発明によれば、メモリへのアクセスを効率的に実行できる。

【発明を実施するための最良の形態】

【0012】

まず、本実施例におけるデータ転送が実現されるべき環境である情報処理装置全体の概要を述べた後、その情報処理装置におけるデータ転送に関連する機能について述べる。

【0013】

図1は、情報処理装置の機能ブロック図である。

情報処理装置100は、統括制御部110、画像処理部120およびメインメモリ14 50

0を含む。情報処理装置100は表示装置150と接続されている。表示装置150は、統括制御部110および画像処理部120の処理の結果得られた画像、映像を出力する。

【0014】

図1などにおいて、様々な処理を行う機能ブロックとして記載される各要素は、ハードウェア的には、CPU、メモリ、その他のLSIで構成することができ、ソフトウェア的には、メモリにロードされた予約管理機能のあるプログラムなどによって実現される。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは当業者には理解されるところであり、いずれかに限定されるものではない。

また、この情報処理装置100では、情報処理装置100を効率よく使用するための機能、環境を提供し、装置全体を統括的に制御するオペレーティングシステム(以下、単に「OS」とよぶ)が実行される。OS上で複数のアプリケーションソフトウェアが実行される。

【0015】

統括制御部110は、1つの主制御装置112と複数の副制御装置116および制御側転送部114を含む。副制御装置116、主制御装置112および制御側転送部114はバス118により相互通信可能である。主制御装置112は、各アプリケーションにおける基本処理単位としてのタスクを各副制御装置116に割り当てる。あるいは、主制御装置112自身がタスクを実行してもよい。副制御装置116がそれぞれ割り当てられたタスクを実行することにより、複数のタスクが並列処理される。

以下、タスクの割り当て処理を含む主制御装置112により実行される処理のことを「メインプロセス」、副制御装置116により実行される処理のことを「サブプロセス」とよぶ。主制御装置112は、ユーザインタフェースに関する処理のように比較的優先度が高い情報処理装置100全体を統括する処理を実行する。これに対して、副制御装置116は、比較的優先度が低いバックグラウンドで実行される計算のようにメインプロセスの下請け的な処理を実行する。

【0016】

制御側転送部114は、メインメモリ140やグラフィックスメモリ128との間のデータ転送、データ退避などを主制御装置112または副制御装置116からの命令、あるいは、画像処理部120からの要求によって制御する。

【0017】

メインメモリ140は、統括制御部110や画像処理部120によって使用される記憶領域である。メインメモリ140には、タスクの実行状態に関連するデータが格納される。メインメモリ140には、たとえば、統括制御部110によりコンピュータグラフィックスに関するタスクが処理されて得られたモデリングデータなどが一時的に格納される。また、このメインメモリ140には、画像処理部120により生成されたデータが退避される場合もある。

【0018】

画像処理部120は、画像処理を専用に行うユニットであり、たとえば、レンダリング処理を実行する。画像処理部120は、統括制御部110からの指示により画像処理を実行する。画像処理部120は、統括制御部110により処理されるそれぞれのタスクに関連する画像処理を行い、生成した画像、映像を表示装置150に出力する。画像処理部120は、複数の画像処理を時分割して並列的に実行してもよい。

【0019】

画像処理部120は、グラフィックスメモリ128、演算ユニット130、ディスプレイコントローラ126、制御ブロック124および画像処理側転送部122を含む。これらのユニット同士も、バス118で接続されており、各ユニット間で相互通信可能である。

【0020】

グラフィックスメモリ128は、画像処理部120により使用、管理されるグラフィッ

10

20

30

40

50

クスデータを記憶するためのメモリ領域である。グラフィックメモリ128には、画像フレームデータが格納されるフレームバッファやZバッファに加えて、画像フレームデータを描画する際に参照される基本データである頂点データ、テクスチャデータ、カラーlookupアップテーブルなどのデータに対応する領域が用意されている。

【0021】

制御ブロック124は、画像処理部120全体を制御するブロックである。制御ブロック124は、演算ユニット130、グラフィックメモリ128、ディスプレイコントローラ126を統括的に制御し、各ブロック間のデータ転送の同期管理やタイマー管理等を行う。

【0022】

画像処理側転送部122は、統括制御部110やメインメモリ140とグラフィックメモリ128の間のデータ転送、データ退避などを制御ブロック124からの命令によって制御する。

【0023】

ディスプレイコントローラ126は、水平および垂直同期信号を生成し、表示装置150の表示タイミングにしたがって、グラフィックメモリ128に格納されるフレームバッファから画像フレームデータのピクセルデータをライン状に順次読み込んでいく。さらにディスプレイコントローラ126は、ライン状に読み込まれたピクセルデータを、RGBのカラー値からなるデジタルデータから表示装置150に対応したフォーマットに変換して出力する。

【0024】

演算ユニット130は、制御ブロック124からの命令にしたがって、グラフィックに関する様々な演算処理を行う。その処理の一例としては、3次元モデリングデータをもとに座標変換、陰面消去、シェーディングを行って画像フレームデータを作成し、フレームバッファに書き込む一連のレンダリング処理などが挙げられる。

演算ユニット130は、特に3次元グラフィックスに関する処理を高速に行うために、ラスタライザ132、シェーダユニット134、テクスチャユニット136などの機能ブロックを含む。

【0025】

ラスタライザ132は、描画する基本物体（以下、「プリミティブ」とよぶ）の頂点データを統括制御部110から受け取り、3次元空間上のプリミティブを投影変換により描画平面上の図形に変換するビュー変換を行う。さらに、描画平面上の図形を、描画平面の水平方向に沿ってスキャンしながら、一列ごとに量子化されたピクセルに変換するラスタ処理を行う。このラスタライザ132によってプリミティブがピクセル展開されて、各ピクセルごとにピクセル情報を算出する。このピクセル情報には、RGBカラー値、透明度を表す値、視点からの奥行きを表すZ値が含まれる。

【0026】

ラスタライザ132は、スキャンラインに沿って所定の大きさのピクセル領域を生成し、シェーダユニット134、テクスチャユニット136へと出力する。ラスタライザ132から出力されるピクセル領域は、一度キューにスタックされ、シェーダユニット134はスタックされたピクセル領域を順に処理していく。

【0027】

シェーダユニット134は、ラスタライザ132により算出されたピクセル情報をもとにシェーディング処理を行い、テクスチャユニット136により得られたテクセル情報をもとに、テクスチャマッピング後のピクセル色を決定し、グラフィックメモリ128内のフレームバッファにシェーディング処理後の画像フレームデータを書き込む。さらにシェーダユニット134はフレームバッファに書き込まれた画像フレームデータに対してフォギング、アルファブレンディング等の処理を行い最終的な描画色を決定してフレームバッファの画像フレームデータを更新する。

【0028】

10

20

30

40

50

テクスチャユニット 136 は、シェーダユニット 134 からテクスチャデータを指定するパラメータを受け取り、要求されたテクスチャデータをグラフィックスメモリ 128 内のテクスチャバッファを読み出し、所定の処理を行った後にシェーダユニット 134 に対して出力する。

【0029】

画像処理部 120 はプリミティブの頂点データなど画像生成のために必要な基本情報や画像生成の開始指示を統括制御部 110 から与えられると、統括制御部 110 とは独立して画像処理を実行する。画像処理部 120 により生成されたデータはグラフィックスメモリ 128 やメインメモリ 140 に転送される。また、統括制御部 110 からの要求、あるいは、画像処理部 120 からの指示によりグラフィックスメモリ 128 に記憶されているデータがメインメモリ 140 に転送されることもある。画像処理部 120 が保持するデータをメインメモリ 140 に転送するときには、画像処理側転送部 122 は制御側転送部 114 に対してデータを転送する。そして、制御側転送部 114 は、受け取ったデータをメインメモリ 140 へ記録する。

【0030】

図 2 は、画像処理部からメインメモリへのデータ転送に関連する機能ブロック図である。

画像処理側転送部 122 は複数のデータ転送要求主体となる D M A C (Direct Memory Access) 170 と、データ転送調停装置 172 を含む。画像処理部 120 においては複数系統からのデータ転送要求がこれらの D M A C 170 を介してデータ転送調停装置 172 に対して発行される。各 D M A C 170 によるデータ転送要求経路をチャンネル(以下、「C H」とも表記する)とよぶ。以下、これらのチャンネルは、C H 0 ~ C H 3 の 4 系統であるとして説明する。この場合、D M A C 170 は 4 個である。

【0031】

データ転送調停装置 172 の選択部 174 は、C H 0 ~ C H 3 のうちのいずれかを選択する。言い換えれば、選択部 174 は C H 0 ~ C H 3 のいずれかに対応づけられた D M A C 170 を選択する。選択方法については後に詳述する。送信部 176 は選択された D M A C 170 から転送を要求されている所定長のデータをバス 118 を介して制御側転送部 114 に転送する。以下、このとき転送されるデータのサイズのことを「基本転送サイズ」とよぶ。仮に、基本転送サイズは 128 バイトであるとして説明する。転送されるデータには、メインメモリ 140 における書き込み先を特定するためのアドレス情報が含まれる。

【0032】

制御側転送部 114 のメモリコントローラ 160 は、データ転送調停装置 172 から転送されたデータをメインメモリ 140 に書き込むが、そのデータ書込命令はメモリコントローラ 160 の書込命令記憶部 162 に一時的に記憶される。書込命令記憶部 162 はキャッシュメモリとして構成されてもよい。また、データ転送調停装置 172 から転送されたデータはメモリコントローラ 160 にて一時的に記憶される。メモリコントローラ 160 は、F I F O (First-In First-Out) 方式にて書込命令記憶部 162 からデータ書込命令から順次取り出して、メインメモリ 140 の該当アドレスに転送されてきたデータを記録する。なお、メモリコントローラ 160 は、F I F O 方式を基本としつつも、書込命令記憶部 162 からデータ書込命令の取り出し順序を適宜変更することができるが、それについては「読み飛ばし方式」として後述する。

【0033】

メインメモリ 140 は、複数のバンクを含む。各バンクは、ここでは、物理的に別個のメモリセルである。あるバンクの遅延時間中に別のバンクへアクセス要求を発行することは可能である。以下、各バンクのエントリビット数は 128 バイトであるとして説明する。各バンクは 128 バイトごとにアドレスが割り振られている。各バンクについてのアドレステーブル 180 については、次の図 3 に関連して説明する。

【0034】

10

20

30

40

50

図3は、メインメモリにおけるアドレステーブルのデータ構造図である。

本実施例においては、メインメモリ140は16個のバンクを有する。以下、これら16個のバンクを識別するために、バンク0、バンク1、・・・、バンク15ともよぶ。

【0035】

メモリのアドレス0～127の128バイト分はバンク0、アドレス128～255の128バイト分はバンク1、・・・というように128バイト単位でバンクが切り替わるようにアドレスが割り振られている。そして、128バイト×16バンクの2048バイト単位ごとに、ふたたび、バンク0からアドレスが割り振られている。すなわち、アドレス2048～2715の128バイト分はバンク0、アドレス2715～2843の128バイト分はバンク1という具合である。

10

【0036】

あるDMAC170が転送要求するデータのメインメモリ140におけるアドレスは連続することが多い。このような、特性に鑑みて、図3のアドレステーブル180に示したようにアドレスを割り振っておくことにより、各バンクへのアクセスが分散されやすくなる。たとえば、バンク0を書き込み先とするデータが転送されてきた後には、バンク1、バンク2という順にデータが転送されてくる可能性が高い。このような場合、バンク0が遅延時間中であっても、バンク1への書き込みを開始できるため、メモリアクセスを効率的に実行できる。

次に、一般的なメモリアクセスの方法を図4および図5に関連して説明し、本実施例におけるメモリアクセスの方法について図6および図7に関連して説明する。

20

【0037】

図4は、データ転送調停装置がデータを転送する一般的な過程を示すフローチャートである。ここに示すS10～S14までの処理は繰り返し実行されるループ処理である。

まず、データ転送調停装置172の選択部174は複数のチャンネルのうちのいずれかを選択する(S10)。たとえば、ラウンドロビン方式で選択するときには、選択部174はCH0～CH3を順番に選択する。選択されたチャンネルのDMAC170が転送すべきデータを持っていれば(S12のY)、データ転送調停装置172の送信部176はそのデータに対するデータ書込命令を制御側転送部114に送信する(S14)。送信後、処理はふたたびS10に戻り、次のチャンネルが選択される。選択されたチャンネルのDMAC170が転送すべきデータを持っていないときには(S12のN)、送信処理を

30

することなく処理はS10に戻り、次のチャンネルが選択される。このように、1回の転送サービスにおいて送信されるべきデータ量である基本転送サイズのデータがいったん送信された後は、次のチャンネルが選択される。なお、選択されたチャンネルにおいて転送要求されているデータが基本転送サイズに満たない場合には、転送要求されているデータはすべて転送される。

【0038】

図5は、図4に示した処理により書込命令記憶部に記憶されるデータ書込命令の態様を示すための模式図である。

ここでは、CH0からはバンク0、1、2に対する転送要求、CH1からもバンク0、1、2、CH2からバンク0、1に対する転送要求が発行されているとする。以下、CH1がバンク0を書き込み先として転送要求したデータ書込命令のことを「CH1：バンク0」というように送信元のチャンネルと書き込み先のバンク名を並べて表記する。

40

図4に示したフローチャートによればデータ転送調停装置172の選択部174は、CH0、CH1、CH2の順に選択するため、「CH0：バンク0」、「CH1：バンク0」、「CH2：バンク0」、「CH0：バンク1」、・・・という順にデータが転送されることになる。書込命令記憶部162は、この転送順にデータ書込命令を記憶する。

【0039】

以下、「書き込み処理」とは、バンクにデータの書き込みを開始してから次の書き込みが可能となるまでの処理であるとして説明する。すなわち、書き込み処理の実行時間は、アクセス時間と遅延時間を合わせた時間である。

50

メモリコントローラ160は、FIFO方式にて一番古い「CH0：バンク0」のデータ書込命令を取り出して、メインメモリ140への書き込み処理を開始する。次にメモリコントローラ160は、「CH1：バンク0」のデータ書込命令を取り出す。このとき「CH0：バンク0」のデータ書込命令についての書き込み処理が完了し、次の書き込み開始が可能な状況になれば、メモリコントローラ160は「CH1：バンク0」のデータ書き込み処理を開始できない。そのため、「CH1：バンク0」のデータの書き込み処理は、「CH0：バンク0」のデータの書き込み処理が完了するまで待たされることになる。さらに、次の「CH2：バンク0」のデータ書込命令もバンク0に対する書き込み要求であるため、これも待たされることになる。

このように、一般的な複数チャンネルにおいてメモリアクセスする場合、特定のバンクにアクセスが集中してしまい、メモリコントローラ160の処理効率が低下してしまうことがある。

【0040】

このような状況への対処方法として、「CH0：バンク0」のデータ書込命令の読み出し後、「CH1：バンク0」や「CH2：バンク0」に優先して「CH0：バンク1」を先に読み出す方法が考えられる。以下、このような処理方式のことを「読み飛ばし方式」とよぶ。読み飛ばし方式によれば、バンク0の遅延時間中にバンク1への書き込み処理を開始できる。また、さらに読み出し順序を変えて、「CH0：バンク1」の次には「CH0：バンク2」を読み出せば、バンク1が遅延時間中であっても、続けてバンク2への書き込み処理を開始できる。そして、最初の「CH0：バンク0」への書き込み処理が完了していれば、「CH1：バンク0」のデータの書き込み処理を開始する。

メモリアクセス時における遅延時間の影響を抑えるために、FIFO方式を基本としつつも、適宜、メモリコントローラ160がデータ記憶部162からデータを読み出す順序をこのように変更してもよい。

【0041】

読み飛ばし方式の場合、メモリコントローラ160により読み飛ばしの実行可否を判断する必要があるが、このために、メモリコントローラ160に過度の処理負荷がかかることは転送スループットの面からみて望ましくない。

【0042】

図6は、本実施例におけるメモリアクセス方式においてデータ転送調停装置がデータを転送する過程を示すフローチャートである。ここに示すフローS20～S24までの処理も繰り返し実行されるループ処理である。以下、図5に関連して説明する方式を「本方式」とよぶ。

まず、選択部174はいずれかのチャンネルを選択する(S20)。このときにもたとえば、ラウンドロビン方式のように既知の選択方法にて選択してもよい。そのほかには、過去にデータ転送を要求する頻度が高い順にDMAC170を選択するLRU(Least Recently Used)方式や、ランダムにいずれかのDMAC170を選択するランダム方式にてDMAC170を選択してもよい。

【0043】

選択されたチャンネルのDMAC170にデータ転送要求があれば(S22のY)、送信部176はそのデータ書込命令を制御側転送部114に転送する(S24)。ここで、選択部174は、そのDMAC170から連続的にデータが送信されている場合、その送信されたデータの合計サイズが、予め定められたストライド長に達しているか否かを判定する(S26)。ストライド長に達していないときには(S26のN)、処理はS22に戻る。そして、先に選択されていたチャンネルのDMAC170がさらにデータ転送要求があれば(S22のY)、再度、同じDMAC170から基本転送サイズのデータが転送される(S24)。

【0044】

一方、ストライド長に達していれば(S26のY)、S20の処理に戻って次のDMAC170が選択される。すなわち、選択されたチャンネルのDMAC170からは、スト

10

20

30

40

50

ライド長を限度として続けてデータが転送され得る。ストライド長が2048バイトであるとすれば、基本転送サイズは128バイトであるので、 $2048 \div 128$ で計16回、同一のDMAC170から連続的にデータが転送され得る。一方、選択されたチャンネルのDMAC170がデータ転送要求をしていないときには(S22のN)、そのまま処理はS22に戻り、次のチャンネルが選択される(S20)。

【0045】

図7は、図6に示した処理により書込命令記憶部に記憶される転送データの態様を示すための模式図である。

図6に示した転送方式によれば、あるチャンネルが選択されたあとに別のチャンネルが選択される場合は2種類ある。一つ目は、選択されたチャンネルからもはや転送すべきデータがなくなったときである。このときには、別のチャンネルが選択される。もう一つは、選択されたチャンネルから連続的に転送されたデータの合計量がストライド長に達したときである。このときにも、別のチャンネルが選択される。図5と同じく、ここでも、CH0からはバンク0、1、2に対する転送要求、CH1からもバンク0、1、2、CH2からバンク0、1に対する転送要求が発行されているとする。図6に示したフローチャートによれば、選択部174はCH0を3回連続で選択した後にCH1を選択することになるので、「CH0：バンク0」、「CH0：バンク1」、「CH0：バンク2」、「CH1：バンク0」、・・・という順にデータが転送されることになる。データ記憶部162は、この転送順にデータを記憶する。

【0046】

メモリコントローラ160は、一番古い「CH0：バンク0」のデータを取り出して、メインメモリ140への書き込み処理を開始する。バンク0への書き込み処理において遅延時間に入ると、「CH0：バンク1」のデータを取り出してバンク1への書き込み処理を開始する。次に「CH0：バンク2」のデータを取り出してバンク2への書き込み処理を開始する。そのため、FIFO方式で読み出しながらも、各バンクの遅延時間の影響を緩和しやすくなる。

特定のチャンネルを介して転送要求されるデータのアドレスは連続することが多いので、このような転送方法によれば、各バンクへのアクセスを分散させやすくなる。

【0047】

なお、各DMAC170の間において選択部174の選択順位として優先順位を設けない方式であるラウンドロビン方式、ランダム方式、LRU方式のほかに、優先順位を設ける場合についても付言しておく。

たとえば、CH0、CH1、CH2、CH3の順に優先順位がつけられているとする。この場合、選択部174は、CH0のDMAC170においてデータの転送要求がなされていれば、CH0を最優先で選択する。一方、CH0のDMAC170においてデータの転送要求がなされていないかチェックする。また、一旦、CH0が選択されると、その選択されたCH0の順位は一時的に下げられ、優先順位はCH1、CH2、CH3、CH0の順に変更される。このとき、CH0以外が選択されると、再び、優先順位はCH0、CH1、CH2、CH3の順に戻される。このように、固定的に優先順位を設定しつつも、選択状況に応じて一時的に優先順位が変更される方式(以下、「可変優先順位方式」とよぶ)の場合であっても、本方式を適用可能である。図4に関連して説明した一般的な方式の場合、優先順位がCH1、CH2、CH3、CH0となっているときにCH1が選択されると、優先順位はCH0、CH1、CH2、CH3となるので次に選択されるのはCH0である。一方、本方式の場合、CH1の選択後においてCH1が更にデータ転送を要求している場合には、再度CH1が選択されることになる。このため、結果的に各バンクへのアクセスを分散させやすくなり、各バンクの遅延時間の影響を緩和しやすくなる。

【0048】

さらに、本方式に先述した「読み飛ばし方式」を併用してもよい。本方式によれば、書込命令記憶部162において複数バンクへの書き込みを間断なく実行しやすい順にデータが

10

20

30

40

50

蓄積されている。データ転送調停装置 172 により、各バンクに対するアクセスが分散するようにデータが転送されるからである。そのため、メモリコントローラ 160 が実行する読み飛ばし処理の実行回数は、図 4 や図 5 に関連して説明した一般的な処理方式に比べて抑制される。

書込命令記憶部 162 のサイズは有限であるため、書込命令記憶部 162 において特定のバンクに対するデータ書込命令ばかりが蓄積された場合には、たとえ読み飛ばし方式によっても各バンクへのアクセスを分散させることができない。これに対して、本方式によればデータ転送調停装置 172 が各バンクへのアクセスが分散するようにデータ書込命令を転送することになるので、データ書込命令の調停が、送信側と受信側、すなわち、データ転送調停装置 172 とメモリコントローラ 160 の両方で効率的に実行されることになり、データのメモリアccessを更に効率化することができる。

10

【0049】

本実施例では、基本転送サイズと、各バンクのエントリビット数を共に 128 バイトとしている。このため、画像処理側転送部 122 と制御側転送部 114 は、128 バイトのデータブロックを 1 単位としてデータを扱えるため、制御側転送部 114 と画像処理側転送部 122 の間でデータをスムーズに取り扱うことができる。ストライド長は、バンクの個数とエントリビット数の積として定められてもよい。本実施例の場合、バンクの個数は 16 でエントリビット数は 128 バイトであるから、ストライド長は 2048 バイトとなる。

【0050】

20

ストライド長は、少なくともメモリコントローラ 160 における書込命令記憶部 162 に格納され得るデータ書込命令の最大個数と基本転送サイズの積以下であることが望ましい。また、バンク数も、書込命令記憶部 162 に格納され得るデータ書込命令の最大個数以下であることが望ましい。

また、DMAC 170 ごとのデータ転送のレイテンシを軽減させることを優先するときには、ストライド長を短くすることにより各 DMAC 170 に対して比較的平等に転送サービスの提供機会を割り当ててもよい。反対に、複数の DMAC 170 全体としてのスループットを重視するときには、ストライド長を長く設定することにより各バンクへのアクセスが分散しやすくなるように処置してもよい。

【0051】

30

各チャンネルについて、それらのデータ書込先のアドレスが全バンク 1 周分のアドレスの中で分散しているときには、ランタイムにストライド長を長く設定してもよい。たとえば、図 3 の場合であれば、2048 バイト分のアドレスの中で分散しているときである。このような場合には、ストライド長を長く設定することにより特定のチャンネルが連続的に選択されやすくなるように処置しなくとも、もともと各バンクへのアクセスが分散されやすい状況にあるといえるからである。データ転送調停装置 172 は、各 DMAC 170 からのデータ書込先アドレスを先読みして、上記指針にてストライド長を増減させてもよい。

また、データ転送を要求している DMAC 170 が少ない場合には、ランタイムにストライド長を長く設定してもよい。このような場合には、ある DMAC 170 を選択して別の DMAC 170 を選択した場合であっても、先に選択された DMAC 170 が次に選択されるまで待たされる時間が短くて済むからである。すなわち、特定の DMAC 170 が連続的に選択されても、他の DMAC 170 が転送サービスを受けるまでに過度に長く待たされるといった状態が生じにくいからである。

40

【0052】

なお、バンク競合が発生したとき、すなわち、特定のバンクに対して連続的にデータ書込要求が来たときに先の書込終了から次の書込開始までの時間として発生する遅延時間と、あるバンクにアクセスしたあとに別のバンクにアクセスするときに発生する切り替えに伴って発生するレイテンシに基づいてストライド長の最低値が決定されてもよい。たとえば、バンク競合が発生したときの遅延時間が 100 ナノ秒であったとする。これに対して

50

、別のバンクにアクセスするときに発生するレイテンシが20ナノ秒であったとする。この場合、 $100 \div 20 = 5$ より、基本転送サイズ×5回を最低ストライド長として設定してもよい。

【0053】

以上、本実施例によれば、メモリインタリーブ方式にてメインメモリ140にデータを書き込み処理するとき、特定のバンクへアクセスが集中するのを防ぎやすくなる。そのため、バンクに対する書き込み処理における遅延時間の影響が抑制され、メモリアクセスがより効率的に実行される。また、特定のDMAC170から多くのデータ書込命令を連続的に発行させるように処置できるが、その一方で、ストライド長によって特定のチャンネルが連続的に選択される回数の上限を設けることにより、複数のDMAC170に対し

10

て転送サービスの単位時間あたりの提供機会が偏らないように処置できる。
また、DRAMの特性として、メインメモリ140に読み出しと書き込みを交互に実行する場合には、そのリードライト切り替えのための処理がオーバーヘッドとなりメモリアクセス効率を悪化させる要因となる。一方、特定のDMAC170からはデータ書込命令かデータ読出命令が連続的に送信される傾向にある。本方式によれば、特定のDMAC170から連続的にデータ書込命令やデータ読出命令を転送させるように処置できるので、このようなリードライト切り替えに伴うオーバーヘッドを軽減できる。

なお、本実施例においては、DMAC170から制御側転送部114に対してデータの書き込みを指示するための命令を転送する処理を中心として説明した。もちろん、本発明の範囲はこれに限られるものではない。DMAC170から制御側転送部114に対して

20

【0054】

以上、本発明を実施例をもとに説明した。この実施例はあくまで例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能で、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【0055】

なお、請求項に記載のデータ転送要求主体の機能は、本実施例においては主としてDMAC170により実現される。これら請求項に記載の各構成要件が果たすべき機能は、本

30

【図面の簡単な説明】

【0056】

【図1】情報処理装置の機能ブロック図である。

【図2】画像処理部からメインメモリへのデータ転送に関連する機能ブロック図である。

【図3】メインメモリにおけるアドレステーブルのデータ構造図である。

【図4】データ転送調停装置がデータを転送する一般的な過程を示すフローチャートである。

【図5】図4に示した処理により書込命令記憶部に記憶されるデータ書込命令の態様を示すための模式図である。

40

【図6】本実施例におけるメモリアクセス方式においてデータ転送調停装置がデータを転送する過程を示すフローチャートである。

【図7】図6に示した処理により書込命令記憶部に記憶される転送データの態様を示すための模式図である。

【符号の説明】

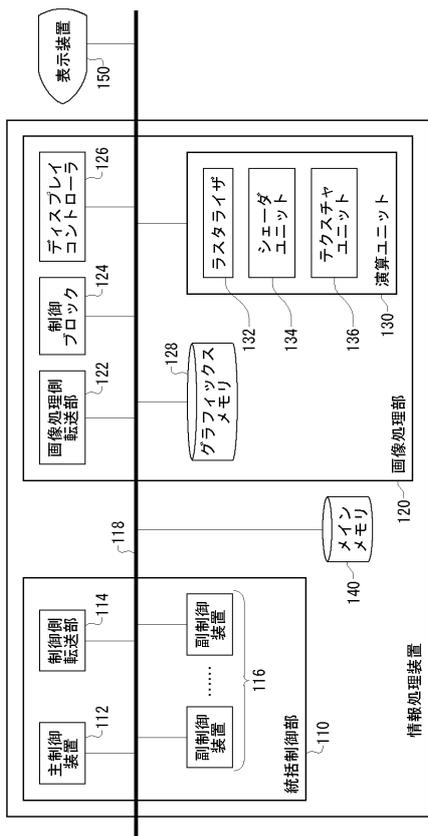
【0057】

100 情報処理装置、 110 統括制御部、 112 主制御装置、 114 制御側転送部、 116 副制御装置、 118 バス、 120 画像処理部、 122 画像処理側転送部、 124 制御ブロック、 126 ディスプレイコントローラ、

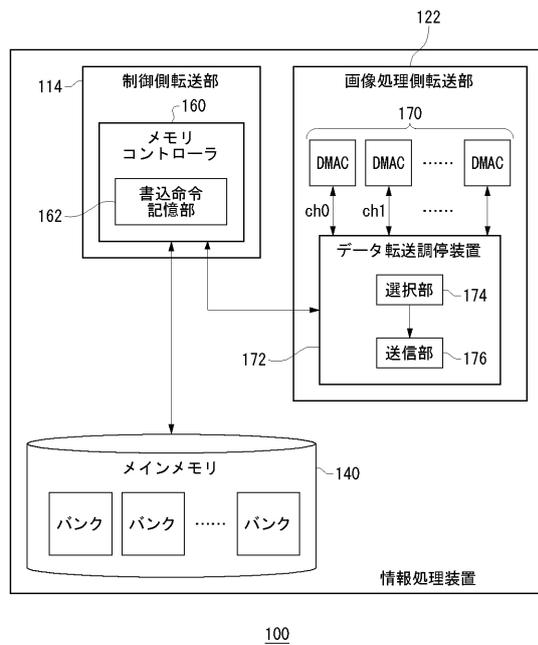
50

128 グラフィックメモリ、 130 演算ユニット、 132 ラスタライザ、
 134 シェーダユニット、 136 テクスチャユニット、 140 メインメモリ、
 150 表示装置、 160 メモリコントローラ、 162 書込命令記憶部、 1
 70 DMAC、 172 データ転送調停装置、 174 選択部、 176 送信部

【図1】



【図2】

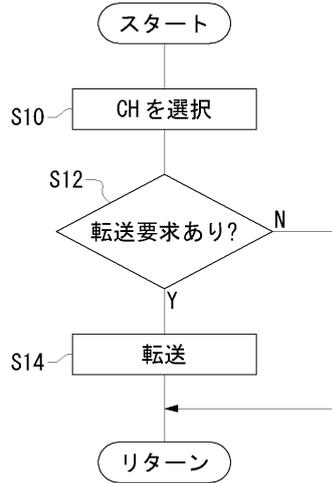


【図3】

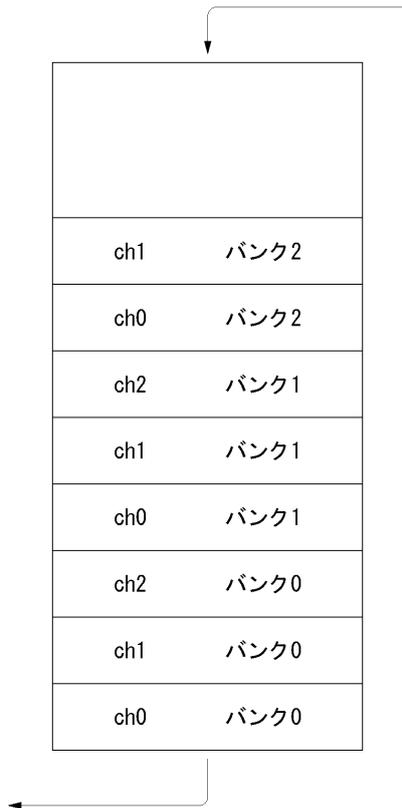
バンク15	1920~2047	3968~4095
.....
バンク3	384~511	2972~3099
バンク2	256~383	2844~2971
バンク1	128~255	2716~2843
バンク0	0~127	2048~2715

180

【図4】

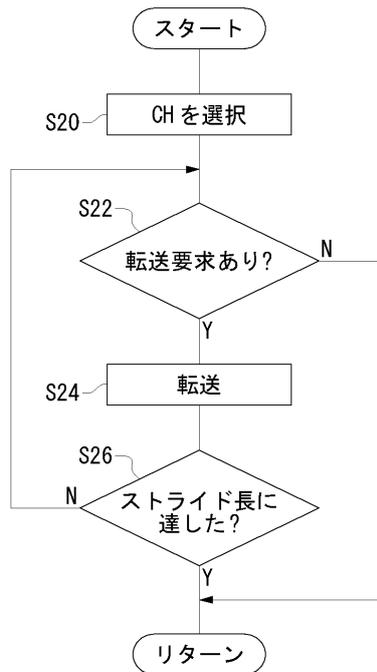


【図5】

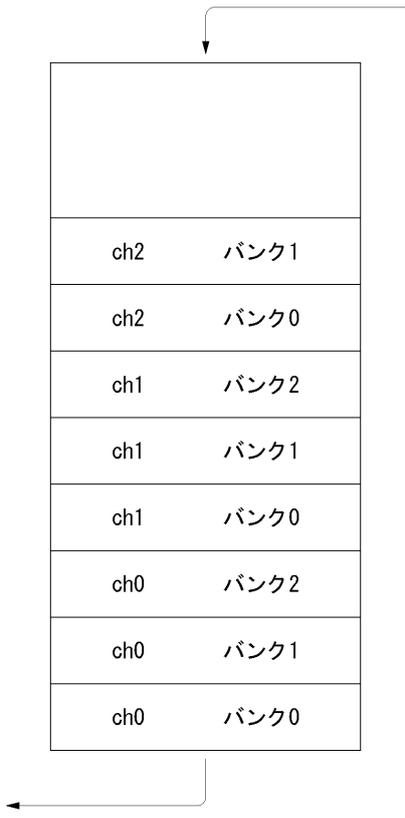


162

【図6】



【 図 7 】



フロントページの続き

審査官 多賀 実

- (56)参考文献 特開昭59-068069(JP,A)
特開2000-010759(JP,A)
特開平02-245858(JP,A)
特開昭58-189752(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00 - 12/06
G06F 13/16 - 13/18
G06F 13/28