



(12) 发明专利申请

(10) 申请公布号 CN 102938653 A

(43) 申请公布日 2013. 02. 20

(21) 申请号 201210453928. 6

(22) 申请日 2012. 11. 13

(71) 申请人 航天恒星科技有限公司

地址 100086 北京市海淀区知春路 82 号院

(72) 发明人 魏耀都 张拯宁 朱翔宇 战勇杰

(74) 专利代理机构 中国航天科技专利中心

11009

代理人 安丽

(51) Int. Cl.

H03M 13/15(2006. 01)

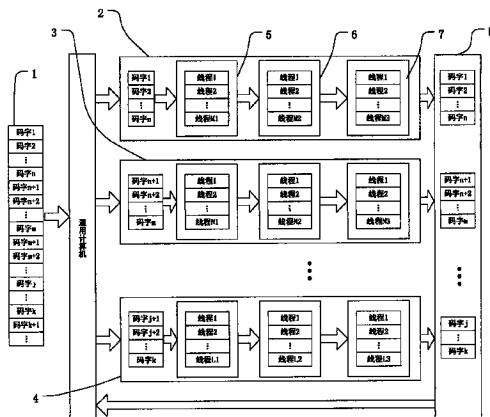
权利要求书 1 页 说明书 5 页 附图 1 页

(54) 发明名称

一种利用图形处理器 GPU 实现的并行 RS 译码方法

(57) 摘要

一种利用图形处理器 GPU 实现的 RS 并行译码方法，以连接有一个或多个 GPU 的通用计算机作为硬件平台，将 RS 译码分为数据读写与初始化、并行译码两部分，利用 GPU 内具有大量计算核心的特点，同时对多个 RS 码字进行译码，实现并行译码。并行译码被分解为 4 个步骤，对每个步骤均设计并行处理方法，使整个译码过程实现并行化，提高 RS 译码的速率。当译码速率上限不足时可以通过增加 GPU 计算核心数量实现经济的升级。本发明根据卫星地面接收系统的要求，采用连接有 GPU 的通用计算机作为硬件平台，通过 GPU 内部计算核心的两层并行结构实现高速并行 RS 译码，在实现高速并行 RS 译码的前提下经济的实现译码速率上限的升级，成本低，易于实现。



1. 一种利用图形处理器 GPU 实现的并行 RS 译码方法,其特征在于实现步骤如下:

(1) 通用计算机首先根据编码参数生成有限域与二进制数据转换查找表、有限域加法查找表、有限域乘法查找表、有限域对偶基和自然基查找表,然后将上述各查找表写入 GPU 纹理内存;

(2) 通用计算机检测已连接的 GPU 个数及各个 GPU 的最大并行规模,根据已连接的所有 GPU 的最大并行规模读入多个待解码的 RS 码字;

(3) 通用计算机将读入的码字写入 GPU 内部的全局存储器,启动 GPU 解码程序进行译码, GPU 中设计多个线程,线程数量远大于读入码字的数量,利用线程实现多个码字的并行译码;

(4) 通用计算机边读取后续码字,边等待 GPU 译码结果;

(5) GPU 译码和后续数据读取都结束后,通用计算机从 GPU 内部的全局存储器中读出译码后的数据,将后续码字写入 GPU 内部的全局存储器,继续进行后续码字的译码,直到所有数据译码完成。

2. 根据权利要求 1 所述的一种利用图形处理器 GPU 实现的并行 RS 译码方法,其特征在于:所述步骤(2)中通用计算机读入多个待解码的 RS 码字的步骤为:

(2. 1) 通过 GPU 编程语言的指令查询所连接的 GPU 数量 n;

(2. 2) 通过 GPU 编程语言的指令逐个查询 n 个 GPU 所支持的最大块数量 B_1, B_2, \dots, B_n 以及每个块支持的最大线程数量 T_1, T_2, \dots, T_n ,根据所处理的 RS 码字的码长 N,确定一次性读入的 RS 码字数量上限为 $(B_1T_1+B_2T_2+\dots+B_nT_n)/(N+1)$;

(2. 3) 通过通用编程语言读入不超过 $(B_1T_1+B_2T_2+\dots+B_nT_n)/(N+1)$ 个 RS 码字。

3. 根据权利要求 1 所述的一种利用图形处理器 GPU 实现的并行 RS 译码方法,其特征在于:所述步骤(3)中启动 GPU 解码程序进行译码的步骤为:

(3. 1) 各 GPU 从全局存储器中读取待译码的码字段,用每个线程处理一个 RS 码字,各线程首先利用查找表将码字从其所使用的有限域基低变换至自然基,然后再次利用查找表将数据变换至有限域;在有限域中将各码字分别乘以相应的系数后,利用归约求和方法将各个块内的线程计算结果相加,得到多个码字的伴随式,若伴随式全为 0 则结束译码,否则继续进行以下步骤;

(3. 2) GPU 使用多个块对多个码字进行关键方程的计算,计算方法为无需求逆的 BM 迭代算法,即 RiBM 算法;GPU 利用多个线程进行计算,每个线程均计算一个迭代参数,计算后得到的关键方程系数存入全局存储器;

(3. 3) GPU 中使用多个块对误码进行纠正,纠正误码的方法为钱搜索和福尼算法;每个块通过多个线程读入关键方程系数,然后各线程分别对所读入的关键方程系数进行乘积运算,随后各线程的运算结果通过归约求和的方法进行求和,根据和值判断当前块所对应的符号是否为错误符号;若当前块所对应符号存在错误,则根据福尼算法计算错误值,并利用块中某一线程执行纠正错误值的程序,得到译码结果;译码结果存储在全局存储器中。

一种利用图形处理器 GPU 实现的并行 RS 译码方法

技术领域

[0001] 本发明涉及一种利用图形处理器 GPU 实现的并行 RS 译码方法,适用于卫星地面接收系统高速数据通信中海量信息的实时处理。

背景技术

[0002] 卫星地面接收系统需要接收通过星地通信链路下传的大量数据信息进行采集、传输、综合分析、计算等处理。星地通信链路是一种不稳定的链路,在该链路上传输信息时存在大量的误码。为了抵抗星地通信链路的误码对数据的影响,卫星地面接收系统必须采用纠错码技术。

[0003] 纠错码技术是卫星地面接收系统抵抗信道误码的关键手段,是数据能够被正确、及时处理的前提,在接收系统中具有较高的处理复杂度。RS 码作为常用的纠错码,其译码速度直接影响卫星地面接收系统的数据处理速度,进而决定了地面接收系统的最大接收能力。

[0004] RS 码是一类具有很强纠错能力的多进制 BCH 码,也是一类典型的代数几何码。它首先由 Reed 和 Solomon 应用 MS 多项式于 1960 年构造出来。用 MS 多项式构造的 RS 码是非系统码,而用 BCH 码构造方法能产生系统码。由于线性码的最大可能的最小距离是校验元的个数加 1,而 RS 码恰好做到这一点,因此,称 RS 码为极大最小距离可分码 (MDS 码)。RS 码采用了 q 进制,所以它是多进制调制时的自然和方便的编码手段;RS 码特别适合于在衰落信道中使用,以克服突发性差错。正由于其优越的性能,RS 码在深空通信、数字音视频通信、磁记录介质等领域取得广泛应用,它是目前应用最为广泛的纠错编码之一。根据对纠错能力的要求不同,可以采用具有不同参数的 RS 码。无论具有何种参数,RS 码所采用的译码方法均相同。例如空间通信常用的 RS 码为空间数据系统咨询委员会 (CCSDS) 推荐的 RS(255,223) 或 RS(255,239) 码,这两种码的译码方法相同,只是纠错能力参数不同。

[0005] RS 译码器的译码过程通常包括三个计算步骤:错误检测、关键方程计算、错误纠正。三个计算步骤均需要进行大量的有限域运算,其中乘法、加法运算的数量庞大,对译码速度的影响也最大。

[0006] 通用计算机只能顺序的执行乘法或加法,因此采用通用计算机及配套软件进行 RS 译码的处理速度较低,无法满足海量卫星数据处理的实时性要求。为了达到实时处理的要求,常用方案是设计专用 RS 译码板卡,通过 FPGA 芯片实现解码过程的高速处理。专用板卡的硬件设计、制造、维护成本高昂,软件不具有通用性。随着卫星下传数据速率的飞速增加,对 RS 译码的速率要求也在不断提高。采用专用处理板卡可以满足现有数据处理速度,但在处理速度需求超出当前板卡性能上限或者所使用的 RS 编码参数变更时只能重新进行板卡硬件及其专用软件的设计,造成设备升级难度大、成本高。

发明内容

[0007] 本发明解决的技术问题是:克服现有技术的不足,提供一种利用图形处理器 GPU

实现的并行 RS 译码方法,根据卫星地面接收系统的要求,采用连接有 GPU 的通用计算机作为硬件平台,通过 GPU 内部计算核心的两层并行结构实现高速并行 RS 译码,在实现高速并行 RS 译码的前提下经济的实现译码速率上限的升级,成本低,易于实现。

[0008] 本发明的技术方案是:一种利用图形处理器 GPU 实现的并行 RS 译码方法,实现步骤如下:

[0009] (1) 通用计算机首先根据编码参数生成有限域与二进制数据转换查找表、有限域加法查找表、有限域乘法查找表、有限域对偶基和自然基查找表,然后将上述各查找表写入 GPU 纹理内存;

[0010] (2) 通用计算机检测已连接的 GPU 个数及各个 GPU 的最大并行规模,根据已连接的所有 GPU 的最大并行规模读入多个待解码的 RS 码字;

[0011] (3) 通用计算机将读入的码字写入 GPU 内部的全局存储器,启动 GPU 解码程序进行译码,GPU 中设计多个线程,线程数量远大于读入码字的数量,利用线程实现多个码字的并行译码;

[0012] (4) 通用计算机边读取后续码字,边等待 GPU 译码结果;

[0013] (5) GPU 译码和后续数据读取都结束后,通用计算机从 GPU 内部的全局存储器中读出译码后的数据,将后续码字写入 GPU 内部的全局存储器,继续进行后续码字的译码,直到所有数据译码完成。

[0014] 所述步骤(2)中通用计算机读入多个待解码的 RS 码字的步骤为:

[0015] (2.1) 通过 GPU 编程语言的指令查询所连接的 GPU 数量 n;

[0016] (2.2) 通过 GPU 编程语言的指令逐个查询 n 个 GPU 所支持的最大块数量 B_1, B_2, \dots, B_n 以及每个块支持的最大线程数量 T_1, T_2, \dots, T_n ,根据所处理的 RS 码字的码长 N,确定一次性读入的 RS 码字数量上限为 $(B_1T_1+B_2T_2+\dots+B_nT_n)/(N+1)$;

[0017] (2.3) 通过通用编程语言读入不超过 $(B_1T_1+B_2T_2+\dots+B_nT_n)/(N+1)$ 个 RS 码字。

[0018] 所述步骤(3)中启动 GPU 解码程序进行译码的步骤为:

[0019] (3.1) 各 GPU 从全局存储器中读取待译码的码字段,用每个线程处理一个 RS 码字,各线程首先利用查找表将码字从其所使用的有限域基低变换至自然基,然后再次利用查找表将数据变换至有限域;在有限域中将各码字分别乘以相应的系数后,利用归约求和方法将各个块内的线程计算结果相加,得到多个码字的伴随式,若伴随式全为 0 则结束译码,否则继续进行以下步骤;

[0020] (3.2) GPU 使用多个块对多个码字进行关键方程的计算,计算方法为无需求逆的 BM 迭代算法,即 RiBM 算法;GPU 利用多个线程进行计算,每个线程均计算一个迭代参数,计算后得到的关键方程系数存入全局存储器;

[0021] (3.3) GPU 中使用多个块对误码进行纠正,纠正误码的方法为钱搜索和福尼算法;每个块通过多个线程读入关键方程系数,然后各线程分别对所读入的关键方程系数进行乘积运算,随后各线程的运算结果通过归约求和的方法进行求和,根据和值判断当前块所对应的符号是否为错误符号;若当前块所对应符号存在错误,则根据福尼算法计算错误值,并利用块中某一线程执行纠正错误值的程序,得到译码结果;译码结果存储在全局存储器中。

[0022] 本发明与现有技术相比具有如下优点:

[0023] (1) 传统的地面接收系统采用专用处理板卡实现实时译码,然而专用处理板卡本

身不具备最大译码速率可升级的能力以及自适应 RS 码参数变化的能力,若需要提升最大译码速率或更改 RS 码参数,则必须重新设计专用处理板卡的软、硬件,设计维护成本高昂。本发明对多个 RS 码字进行并行译码处理,可以在插有少量 GPU 的通用计算机上达到卫星地面接收系统所要求的译码速率。用 GPU 实现错误检测、关键方程计算、错误位置计算和错误值计算,使通用计算机可以将数据的读入与 RS 译码处理并行进行,可以提高卫星地面接收系统的译码速率。

[0024] (2) 与使用专用译码板卡的传统方法相比,本发明的可扩展性更好。当系统性能需要扩展时简单的增加 GPU 个数即可,而专用硬件板卡不具有此性质。

[0025] (3) 与使用专用译码板卡的传统方法相比,本发明的软件设计更简便。专用译码板卡的软件需要烧入芯片,代码修改和调试过程复杂;GPU 所用软件无需烧入处理,修改、调试方法与通用计算机软件一样简便。

附图说明

[0026] 图 1 为利用 GPU 实现的并行 RS 译码示意图。

具体实施方式

[0027] 下面结合附图 1 对本发明作进一步说明。

[0028] GPU 是一种可以连接在通用计算机上的并行处理设备,具有大量的并行计算核心,通常具有独立的物理存储器和高速存储器读取机制,可以在通用计算机上使大量数据的加法、乘法运算速度得到数十甚至数百倍的提升。

[0029] 在 GPU 中,软件以程序核为单位顺序执行,每个程序核均对数据进行并行处理。每个程序核的并行规模可以独立设置。程序核内的并行结构分为两个层次:块和线程。每个程序核软件可以分为多个块并行执行,每个块独立执行程序指令,块与块之间不能进行数据交互。每个块又可以分为多个线程,每个线程独立执行所在的块收到的程序指令,线程之间可以通过每个块内的共享内存进行通信。同时 GPU 的软件无需烧入芯片等处理,调试开发简便。

[0030] 在硬件结构上,GPUs 内部具有上百个硬件计算核心,各个计算核心均可以独立进行运算。通常用一个硬件计算核心运行一个软件线程,从而实现多个软件线程的并行运行。GPU 硬件计算核心与软件线程的映射关系对软件透明,因此同样的软件可以运行在具有不同指标的 GPU 硬件上。在存储器结构上,GPUs 具有全局存储器、共享存储器和纹理存储器,其中纹理存储器可以通过高速缓存实现快速查表功能。

[0031] 利用 GPU 软硬件的并行结构,可以实现并行 RS 译码。由于 GPU 对数据的并行处理速率与其中包含的计算核心数量以及处理主频约成正比,同样的软件可以通过增加硬件计算核心数量实现数据处理速率的提升。因此在性能需求提高时,仅需要增加 GPU 个数或者更换计算核心数量更多、处理主频更高的 GPU 即可实现,无需对软件进行更改。升级难度和成本远低于使用专用译码板卡。

[0032] 本发明采用通用计算机及 GPU 作为 RS 译码的硬件平台,通用计算机通过 PCI-E 接口连接一块或多块 GPU;将 RS 译码过程分为数据初始化、并行译码两部分;通用计算机进行各块 GPU 的任务分配、待译码数据的读取、译码参数初始化、GPU 内部存储器中数据的读写

和数据整理, GPU 进行待译码数据的并行译码处理。GPU 将多个码字的译码任务分配给多个硬件计算核心, 通过多个计算核心的同时译码实现并行译码。

[0033] 通用计算机首先根据编码参数生成有限域与二进制数据转换查找表、有限域加法查找表、有限域乘法查找表、有限域对偶基和自然基查找表, 将各查找表存入 GPU 纹理内存; 然后检测已连接的 GPU 个数及各 GPU 的最大并行规模, 根据已连接的所有 GPU 的最大并行规模读入多个待解码的 RS 码字; 将读入的码字写入 GPU 内部的全局存储器, 启动 GPU 解码程序进行译码, GPU 软件中设计多个线程, 线程数量远大于读入码字的数量, 利用线程实现多个码字的并行译码; 随后通用计算机一边读取后续码字, 一边等待 GPU 译码结果; 待 GPU 译码和后续数据读取都结束后, 从 GPU 内部的全局存储器中读出译码后的数据; 最后将后续码字写入 GPU 内部的全局存储器, 继续进行后续码字的译码, 直到所有数据译码完成。

[0034] GPU 将 RS 译码按顺序分解为检错、关键方程计算、错误纠正三个步骤; 为每个步骤设计一个程序核, 各个程序核具有不同的块规模与线程规模, 通过在 GPU 内的多个硬件计算核心上同时运行多个线程实现并行译码; 三个程序核顺序依次执行, 程序核之间通过全局存储器进行数据传递; 检错程序核通过多个块和线程实现并行, 块数量等于单个 RS 码字的伴随式数量与 RS 码字个数之积, 每个块内的线程数量与 RS 码字长度相同; 关键方程计算程序核通过多个块和线程实现并行, 块数量为 RS 码字个数, 每个块内的线程数量为关键方程系数个数; 错误纠正程序核通过多个块和线程实现并行, 块数量为 RS 码字长度与 RS 码字个数之积, 每个块内线程数量为 RS 码字能够纠正的最大错误字节数量。

[0035] 本发明具体实现步骤如下:

[0036] (1) 通用计算机首先确定 RS 码的码字长度、纠错能力、码生成多项式、有限域生成多项式、本原元。随后生成有限域与二进制数据转换查找表、有限域加法查找表、有限域乘法查找表、有限域对偶基和自然基查找表, 存储在计算机内存中。各查找表根据有限域生成多项式进行计算, 由程序自动生成。

[0037] (2) 通用计算机通过 GPU 编程语言的指令对其所连接的 GPU 进行检测, 检测到 2、3、4 等 n 个 GPU。随后通用计算机逐个查询 n 个 GPU 所支持的最大块数量 B₁, B₂, …, B_n 以及每个块支持的最大线程数量 T₁, T₂, …, T_n, 根据所处理的 RS 码字的码长 N, 确定一次性读入的 RS 码字数量上限为 $(B_1T_1+B_2T_2+\dots+B_nT_n)/(N+1)$ 。根据读入码字数量的上限选取同时译码的 RS 码字数量 k。

[0038] (3) 从待译码的码字 1 中连续读入 k 个码字至通用计算机内存。根据 2、3、4 等 GPU 所支持的并行规模将 k 个码字分为多个码字段, 码字段个数与 GPU 个数相同。通用计算机将各段码字分别传输至各 GPU 的全局存储器。随后通用计算机读入下面 k 个连续的码字。

[0039] (4) 各 GPU 从全局存储器中读取待译码的码字段, 用程序核 5 对所有码字进行检错。程序核 5 具有多个块, 块数量等于读入的码字段数量与单个码字的伴随式个数之积, 即每个块可以求得一个伴随式。每个块内的线程数量等于码字长度, 即每个线程对一个码字进行处理。各线程首先利用查找表将各码字从其所使用的有限域基低变换至自然基, 然后再次利用查找表将数据变换至有限域。在有限域中将各码字分别乘以相应的系数后, 利用归约求和方法将各个块内的线程计算结果相加, 得到多个码字的伴随式, 将其存入全局存储器并检验伴随式中是否存在非零值, 若存在非零值, 则非零标志位置位。

[0040] (5) 通用计算机对非零标志位进行检测, 若置位则运行程序核 5 和程序核 6, 否则

读入的数据中不存在错误,读入的码字即为译码结果,继续进行后续码字的译码处理。

[0041] (6) 若非零标志位置位,则各 GPU 中程序核 6 从全局存储器中读入程序核 5 的计算结果进行关键方程计算。程序核 6 使用多个块,块数量等于 RS 码字个数,即每个块对一个码字进行关键方程的计算,计算方法为无需求逆的 BM 迭代算法。GPU 中每个块内设计多个线程,线程数量与所使用的计算方法中迭代参数的个数相同,即每个线程对一个迭代参数进行计算。计算后得到的关键方程系数存入全局存储器。

[0042] (7) 各 GPU 中程序核 7 从全局存储器中读入程序核 6 的计算结果,同时读入待纠错的码字,利用钱搜索和福尼算法对误码进行纠正。程序核 7 使用多个块,块数量等于 RS 码字个数与单个 RS 码字的伴随式个数之积,即每个块对码字中的一个符号进行错误纠正。每个块使用多个线程,线程数量等于 RS 码字所能纠正的最大错误符号数量。每个块首先通过多个线程读入关键方程系数,然后各线程分别对所读入的关键方程系数进行乘积运算,随后各线程的运算结果通过归约求和的方法进行求和,根据和值判断当前块所对应的符号是否为错误符号。若当前块所对应符号存在错误,则根据福尼算法计算错误值,并利用块中某一线程执行纠正错误值的程序,得到译码结果。译码结果存储在全局存储器中。

[0043] (8) 各 GPU 译码后的码字段 8 分别存储在各 GPU 的全局存储器中。通用计算机在各 GPU 译码结束后从各全局存储器中读取译码后的 RS 码字段,在内存中将各码字段按顺序进行拼接,得到译码结果。

[0044] (9) 通用计算机将读入的下 k 个码字传输至 GPU,开始下 k 个码字的译码,直到所有待译码码字均被译码完毕。

[0045] 本发明未详细阐述部分属于本领域公知技术。

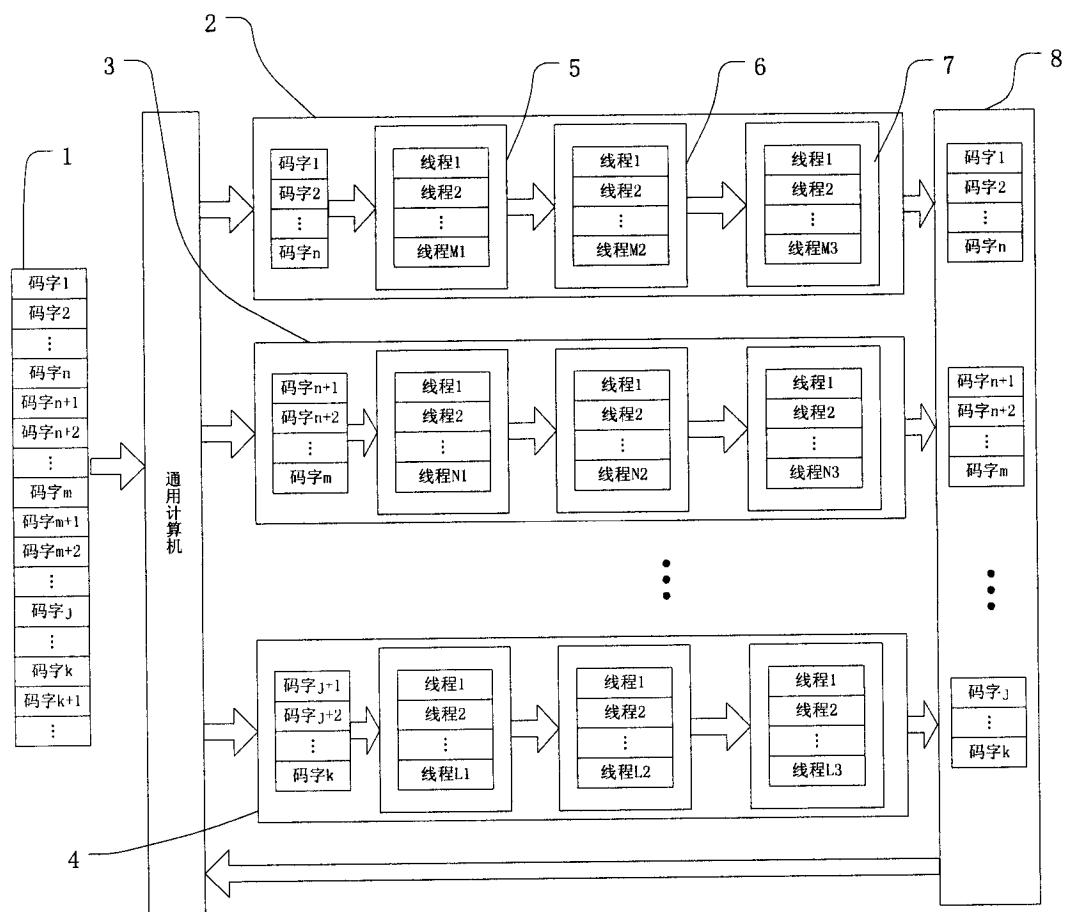


图 1