

云计算环境下不可靠数据恢复方法研究

刘 凯, 王三虎, 薛艳锋, 武彩红

(吕梁学院 计算机科学与技术系, 山西 吕梁 033000)

摘要: 由于云计算环境下不可靠数据中包含有利用价值的信息, 但直接使用可能会给计算机带来负担, 所以需要云计算环境下不可靠数据进行恢复; 当前大多数方法对不可靠数据进行复写时, 利用地址映射层分配新的写入地址, 因此不可靠的数据在一段时间内仍存在于云存储中, 在这样的系统上实现对其的保护, 不需额外保存数据的更新信息; 通过 FLASH 的带外区记录时间戳, 在 FTL 的映射项中增加时间戳信息, 在不跟踪映射表每次更新的情况下, 达到数据快速恢复的目的; 但这种方法对云存储系统性能产生不利影响; 为此, 提出一种基于张量 Tucker 阈值的云计算环境下不可靠数据恢复方法, 首先利用云计算环境下节点自身的随机密钥生成器产生随机会话密钥, 并对不可靠数据 HMAC 报文鉴别码进行计算, 从而实现保护和重构; 在此基础上, 将不可靠数据的阈值分解过程与奇异阈值方法相结合, 从而得到 Tucker 阈值算子, 实现动态的数据恢复, 恢复过程中利用 Tucker 阈值算子与增广拉格朗日乘子方法相结合的方式选择 $n-1$ 秩相似张量, 提出基于增广拉格朗日乘子方法的不可靠数据 Tucker 阈值恢复方法, 完成云计算环境下不可靠数据恢复; 实验证明, 所提方法能够有效提高不可靠数据恢复的准确性, 降低数据恢复的能耗和时间, 具有较强的可行性, 为该课题的应用研究提供理论依据。

关键词: 云计算; 不可靠数据; 数据恢复

Research on Unreliable Data Recovery Method in Cloud Computing Environment

Liu Kai, Wang Sanhu, Xue Yanfeng, Wu Caihong

(Department of Computer Science and Technology, Lüliang University, Lüliang 033000, China)

Abstract: Because the cloud computing environment is unreliable and the data contain valuable information, the direct use may lead to the burden on the computer, so it is necessary to recover the unreliable data in the cloud computing environment. Most of the current methods of copying unreliable data, use the address mapping layer assignment write new address, so unreliable data in a period of time still exists in the cloud storage, the protection is achieved in such a system, do not need to update the channel additional data storage information. The timestamp information is added to the mapping term of FTL through the outer zone record timestamp of FLASH, and the fast recovery of data is achieved without updating the mapping table each time. But this approach has a detrimental effect on the performance of cloud storage systems. Therefore, this paper proposes a method to calculate environment unreliable data recovery method based on tensor Tucker threshold cloud, first calculate the random secret key generator node environment generates a random session key using the cloud, and the unreliable data HMAC message authentication code is calculated, so as to realize the protection and reconstruction. On this basis, combining the unreliable data threshold decomposition process and singular threshold method, in order to get the threshold of Tucker operator, to achieve dynamic data recovery, similar $n-1$ rank tensor combined with Tucker threshold operator and the recovery process using augmented Lagrange multiplier method is chosen, no reliable data recovery method of Tucker threshold augmented Lagrange multiplier based on the proposed, completed under the cloud computing environment is not reliable data recovery. Experimental results show that the proposed method can effectively improve the accuracy of unreliable data recovery, reduce the energy consumption and time of data restoration, and has a strong feasibility. It provides a theoretical basis for the application of this topic.

Keywords: cloud computing; unreliable data; data recovery

0 引言

随着云计算技术的不断发展产生了大量数据, 称为现今社会非常重要的信息资源。然而云计算环境下, 一些数据由于木马病毒等原因, 变成不可靠数据^[1]。但这些数据中可能蕴含非常重要的信息, 直接使用会影响使用效率^[2]。不可靠数据恢复包括数据保护、数据分解等内容, 针对不可靠数据存储位置的不同, 导致不可靠数据恢复方法的不同^[3]。云环境下不可靠数

据恢复的方法需要改进和创新, 以满足当前环境下数据恢复的要求^[4]。但目前大多数数据恢复方法利用手工重建数据链恢复不可靠数据, 通过文件系统转换不可靠数据类型, 完成 RBM 重建, 根据重建后的 RBM, 对不可靠数据位置进行描述, 完成原不可靠数据内容表的重建, 从而完成不可靠数据重建^[5], 这种方法进行不可靠数据恢复能耗较少, 称为当前不可靠数据恢复的重要方法, 随着信息化和网络化的发展与应用^[6], 不可靠数据恢复成为当前研究的热点问题, 随着研究内容的深入, 产生了大量的研究成果^[7]。

文献 [8] 以 SQLite3 文件格式为恢复依据, 提出了一种云计算环境下不可靠数据恢复方法。通过对 SQLite3 数据库的文件格式的分析, 分析不可靠数据记录的存储机制。通过识别 SQLite3 文件格式, 实现不可靠数据页中的自由块和未使用空

收稿日期: 2017-01-05; 修回日期: 2017-02-11。

基金项目: 吕梁学院校内自然科学基金项目 (ZRKN201506); 吕梁学院校级重点大学生创新创业训练项目 (CXCYZD201625)。

作者简介: 刘 凯 (1985-), 男, 山东临沂人, 硕士研究生, 助教, 主要从事计算机软件、数据挖掘、人工智能方向的研究。

间进行寻找和收集, 从而完成不可靠数据恢复方法。但这种方法仅适用于 SQLite3 文件格式, 使用范围受到限制, 难以推广使用。文献 [9] 提出一种基于相干函数无偏估计的不可靠数据恢复技术。首先对云计算环境下不可靠数据结构模型进行构建, 针对需要进行数据恢复的不可靠数据结构进行特征分解和信息模型构建, 采用相干函数特征分解算法对待恢复的不可靠数据进行筛分和残差滤除, 最后采用相干函数无偏估计算法提高不可靠数据恢复的准确度, 这种方法对对数据进行恢复时, 花费时间较长。文献 [10] 提出一种基于模式识别的云环境下不可靠数据恢复方法。通过对不可靠数据进行特征量提取, 并与预先设置的时序相对变化量变化趋势矩阵和时序数值矩阵进行模式匹配, 实现不可靠数据的实时辨识, 并使用错误出现最近时刻的正常数据进行数据快速恢复。

针对传统方法存在的不足, 本文提出一种基于张量 Tucker 阈值的云计算环境下不可靠数据恢复方法, 首先, 利用节点自身的随机密钥生成器生成随机会话密钥, 并完成不可靠数据 HMAC 报文鉴别码计算, 从而实现不可靠数据的保护和重构。采用阈值分解过程与奇异值阈值方法相结合, 实现对不可靠数据的 Tucker 阈值算子的计算, 实现不可靠数据恢复过程中动态地选择 $n-1$ 秩相似张量, 在此基础上, 以基于增广拉格朗日乘子方法为依据, 结合得到的算子, 实现基于张量 Tucker 阈值的云计算环境下不可靠数据恢复。实验证明, 本文所提方法是一种准确度较高、可行性较强的方法, 提高不可靠数据恢复效率, 为该课题的应用研究提供理论依据。

1 基于距离的云计算环境下不可靠数据恢复

在云计算环境下, 不可靠数据对象之间的相关性可以表示成以下几种:

- 1) 在云计算环境中, 与不可靠数据对象相关联的对象一直在一起。
- 2) 与不可靠数据相关联的数据对象在开始几个连续窗口中在一起, 然后, 它们分离, 不在一起。
- 3) 与不可靠数据相关联的数据对象在连续窗口中, 有时在一起, 有时分离。

设在云计算环境中当前数据窗口的前一窗口 $[t-2w+1, t-w]$ 中, 项集 $X = \{T_1, T_2, \dots, T_m\}$ 的关联度为 $\varphi_{t-w}(X)$, 则在当前数据窗口 W_c 中, X 关联度可以表示为:

$$\varphi(X) = \begin{cases} \varphi_{t-w}(X) \exp_{\text{sup}}(X, W_c) & \text{当 } X \text{ 中的数据被当前窗口相同阅读器探测到} \\ \varphi_{t-w}(X) \cdot f & \text{当 } X \text{ 中的数据没有被当前窗口相同阅读器探测到} \end{cases} \quad (1)$$

式中, $\exp_{\text{sup}}(X, W_c)$ 表示在但钱数据窗口 W_c 项集 X 的期望支持度, f 表示该窗口的衰减因子。设定 T_m 和 T_n 表示两个数据对象, 则在窗口 $[t-w+1, t]$ 中, 存在 $\varphi_i(T_m, T_n) = \varphi_i(T_n, T_m)$ 。

设 η 表示云计算环境下不可靠数据的不关联阈值, $\varphi_i(X)$ 表示不可靠数据项集 X 在云计算窗口 W_c 中的关联度, 当满足条件: $\varphi_i(X) < \eta$ 时, 则称云计算环境下不可靠数据的项集树所包含不可靠数据对象不具有关联性。

设定一个窗口 W_c 中, $T_i, T_j \in X$, 当 T_j 满足公式 (2), 则称 T_j 为 T_i 的最大关联标签。

$$\hat{\varphi}_i(T_i, T_j) = \max\{\forall n, \varphi_i(T_i, T_j)\}, i \neq j \quad (2)$$

在云计算窗口 $W[m, n]$ 内, 以不可靠数据 T_i 的观测事务为元素, 以观测时序为顺序惊醒汇集组成集合, 将这个集合称为不可靠 T_i 的概率运动轨迹, 表示为:

$$W_i = \{\langle R_j, T_i, P_j, t_k \rangle \mid 1 \leq j \leq s, m \leq k \leq n\} \quad (3)$$

在云计算窗口 $W[m, n]$ 内, 将由 N 个不可靠数据标签的概率运动轨迹汇集组成的集合称为不可靠数据概率运动轨迹集, 表示为 $MS = \{M_1, M_2, \dots, M_N\}$ 。

分别用 M_i 和 M_j 对数据对象 T_i 和 T_j 的概率运动轨迹进行表述, 且满足条件 $M_i, M_j \in MS$ 。当 M_i 和 M_j 的前 l 项都表示对应的不可靠数据概率相似运动轨迹点, 则称 M_i 和 M_j 为 l -概率相似运动轨迹, l 表示不可靠数据特征的相似度, 记为 $l-PS(M_i, M_j)$ 。

在不可靠数据流中, 不可靠数据 T_i 的最相似运动轨迹表示与不可靠数据 T_i 概率相似运动轨迹最相似的运动轨迹。并将不可靠数据 T_i 的最相似运动轨迹的标签称为最相似运动轨迹标签。

多远统计分析方法是以数据对象所具有的多个观测指标规律为基础进行分析统计, 从而发现数据对象指标存在的相互关联。由于云计算环境下不可靠数据对象存在多方面关联信息, 为多元统计创造了条件。

利用多元统计的判别方法对计算环境下不可靠数据进行恢复, 主要技术有两点:

- 1) 确定数据恢复的指标, 即预报因子;
- 2) 各项指标在恢复过程中的权重

设定一个数据窗口 W_c , 标签 T_i 的最相关标签为 T_{MAX_r} , 且最相似迹标签为 T_{MAX_s} , 则 T_i 在云计算环境 $R_x \in \mathbf{RS}_i$ 的各项指标向量可以表示为:

$$\vec{T}_{R_x} = \{\varphi_1 \exp_{\text{sup}}(T_i, W_c), \varphi_2 p_{iv}, \varphi_3 \exp_{\text{sup}}(T_{MAX_r}, W_c), \varphi_4 \exp_{\text{sup}}(T_{MAX_s}, W_c)\} \quad (4)$$

式中, $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ 分别表示指标所占的权重, p_{iv} 表示 T_i 的先验概率。其欧氏距离可以表示为:

$$D(T_i, R_x) = I_{R_x} I_{R_x}^T \quad (5)$$

根据上式, 得到基于多元统计的判别方法判定规则:

在 $D(T_i, R_x) \geq \max\{\forall j, D(T_i, R_j)\}$ 情况下, 其中 $1 \leq j \leq m$, 则可以说明标签 T_i 位于云计算系统中, 且其后验概率可以计算为:

$$P(T_i, R_x) = \frac{D(T_i, R_x)}{\sum_{j=1}^m D(T_i, R_j)} \quad (6)$$

通过上述论述, 基于距离的云计算环境下不可靠数据恢复方法的具体步骤: 对不可靠数据进行频繁模式挖掘, 从而确定不可靠数据的最大关联标签以及最相似轨迹标签。

利用公式 (5) 对不可靠数据对应的欧氏距离进行计算, 并利用公式 (6) 实现对不可靠数据的后验概率进行计算, 并对后验概率表进行更新, 从而确定不可靠数据标签之间的关联度。

2 基于张量 Tucker 阈值的不可靠数据恢复方法

2.1 数据保护与恢复重构

为实现数据保护, 云计算环境中感知节点 v_i 对云计算环境下不可靠数据 d_i 进行采集, 然后按照以下方式对不可靠数据 d_i 进行操作。以保护数据的完整性和机密性。

首先利用云计算环境下节点自身的随机密钥生成器产生随机会话密钥 K_i ，来保护不可靠数据的完整性和私密性。

然后使用节点密钥 K_i 对不可靠数据 d_i 的 HMAC 报文鉴别码进行计算： $MAC = H_{K_i}(d_i)$ ，以保护不可靠数据的完整性，消除不可靠数据中攻击对其的篡改。

最后使用最开始生成的随机密钥对不可靠数据和鉴别码的联合信息进行对称加密，生成密文，保护数据的机密性。

2.2 基于张量 Tucker 阈值的不可靠数据恢复方法

为了便于计算和保证不可靠数据回复的精度，引入不可靠数据特征的初始值与不可靠数据张量相等的辅助张量 M ，满足辅助张量初始值与张量 X 相等，从而在对不可靠数据恢复进行恢复计算过程中，尽可能保留不可靠数据张量迭代的多样性，从而提高云计算环境下不可靠数据恢复精度。

$$\begin{aligned} & \min \|M\|_* \\ & s. t. \quad M = X \\ & \quad X_{\Omega} = T_{\Omega} \end{aligned} \quad (7)$$

公式 (7) 优化目标函数所要解决的不可靠数据恢复问题通过随机给定部分张量元素下标的子集，根据这些已知元素恢复缺失元素值，使得恢复后的张量为所有可能的恢复张量中的最低秩近似张量。

(矩阵奇异值分解) 秩为 r 的矩阵 $X \in \mathbf{R}^{n_1 \times n_2}$ 的奇异值分解表示为：

$$X = U \sum V^*, \quad \sum = \text{diag}(\{\sigma_i\}_{1 \leq i \leq r}) \quad (8)$$

其中： U 、 V 分别表示 $n_1 \times r$ 、 $n_2 \times r$ 大小的正交矩阵， σ_i 表示不可靠数据奇异值，且 $\sigma_i > 0$ 。

(奇异值收缩算子) 云环境下不可靠数据矩阵 $X \in \mathbf{R}^{n_1 \times n_2}$ 的奇异值收缩算子或奇异值阈值算子可以表示为：

$$D_{\tau}(X) = U D_{\tau}(\sum) V^*, \quad D_{\tau}(\sum) = \text{diag}(\{\max(\sigma_i - \tau, 0)\}) \quad (9)$$

通过公式 (9) 可知，当云环境下不可靠数据矩阵 X 的奇异值较数据阈值 τ 小时，则可认为 $D_{\tau}(X)$ 的秩要比 X 小。

对任意的 $\tau \geq 0$ 和不可靠数据矩阵 $Y \in \mathbf{R}^{n_1 \times n_2}$ ，奇异值收缩算子需要满足：

$$D_{\tau}(Y) = \arg \min_X \frac{1}{2} \|X - Y\|_F^2 + \tau \|X\|_* \quad (10)$$

根据对云计算环境下不可靠数据矩阵奇异值收缩算子的论述，本文提出了一种基于张量分解的张量恢复算子。

张量 $X \in \mathbf{R}^{n_1 \times n_2 \times \dots \times n_N}$ 的 Tucker 阈值算子 $TD_{\tau}(X)$ 定义如下：

$$X \approx \Gamma(1:r_1, 1:r_2, \dots, 1:r_N) \times_1 A^{(1)}(:, 1:r_1) \times_2 A^{(2)}(:, 1:r_2) \dots \times_N A^{(N)}(:, 1:r_N) \quad (11)$$

$$r_i = \text{sum}(D_{\tau}(\sum_i)) = \text{sum}(\text{diag}(\{\max(\sigma_i - \tau, 0)\})) \quad (12)$$

式中， \sum_i 表示张量的 mode- n 矩阵 $A^{(i)}$ 奇异值分解的奇异值。

对任意的 $\tau \geq 0$ 和不可靠数据矩阵 $Y \in \mathbf{R}^{n_1 \times n_2 \times \dots \times n_N}$ ，奇异值收缩算子需要满足：

$$TD_{\tau}(Y) = \arg \min_X \frac{1}{2} \|X - Y\|_F^2 + \tau \|X\|_* \quad (13)$$

增广拉格朗日乘子方法主要对如下限制优化问题求解。

$$\min f(X), s. t. h(X) = 0 \quad (14)$$

将上述优化问题通过拉格朗日函数形式进行表示为：

$$L(X, Y, \mu) = f(X) + \langle Y, h(X) \rangle + \frac{\mu}{2} \|h(X)\|_F^2 \quad (15)$$

式中， Y 表示不可靠数据的拉格朗日乘子， μ 表示正实数。利用交替方向乘子方法对不可靠数据进行分裂并迭代更新从而对不可靠数据进行求解。

$$\begin{cases} X^{k+1} = \arg \min_X L(X^k, Y^k, \mu^k) \\ Y^{k+1} = Y^k + \mu^k h(X^{k+1}) \end{cases} \quad (16)$$

为解决云计算环境下不可靠数据恢复问题，本文通过增广拉格朗日乘子方法确定不可靠数据的优化目标，并将优化目标函数利用拉格朗日函数进行表示为：

$$\begin{aligned} L(M, X, Y, \beta) &= \|M\|_* - \langle Y, X - M \rangle + \frac{\beta}{2} \|X - M\|_F^2 \\ & s. t. X_{\Omega} = T_{\Omega} \end{aligned} \quad (17)$$

其中： β 为一正常数，为了加快本文方法的收敛速度，在迭代中令 $\beta^{t+1} = t\beta^t$ ，其中 $t \in [1, 2]$ 。根据增广拉格朗日乘子方法，得到如下三个子问题的求解形式：

$$\begin{cases} M = \arg \min_M L(M, X, Y, \beta) \\ X = \begin{cases} \arg \min_X L(M, X, Y, \beta) \\ s. t. X_{\Omega} = T_{\Omega} \end{cases} \\ Y = Y + \beta(X - M) \end{cases} \quad (18)$$

利用公式 (15)，对 X 和 Y 进行求解，则有：

$$X = P_{\Omega} \left(\frac{Y - \beta M}{\beta} \right) + P_{\Omega}(T) \quad (19)$$

式中， $\bar{\Omega}$ 表示 Ω 的补集， $P_{\Omega}(\cdot)$ 表示在 Ω 上的投影操作。

可利用张量 Tucker 阈值方法对不可靠数据 M 进行求解，实现不可靠数据恢复，需要对公式 (17) 进行等价变形：

$$L(M, X, Y, \beta) = \|M\|_* + \frac{\beta}{2} \left\| M - \left(X - \frac{Y}{\beta} \right) \right\|_F^2 \quad (20)$$

设定 $N = X(1/\beta)Y$ ，则：

$$\begin{aligned} M &= \arg \min_M \left(\|M\|_* + \frac{\beta}{2} \|M - N\|_F^2 \right) = \\ & (1/\beta) \|M\|_* + \frac{1}{2} \|M - N\|_F^2 \end{aligned} \quad (21)$$

通过上述论述，完成云计算环境下不可靠数据恢复。

3 实验结果与分析

为证明本文提出的基于张量 Tucker 阈值的云计算环境下不可靠数据恢复方法的有效性和实用性，以 Intel P4 2G 处理器为硬件环境，Matlab2008a 为平台，以云计算环境下不可靠数据为实验数据，运用对比法将本文提出的基于张量 Tucker 阈值的云计算环境下不可靠数据恢复方法与文献 [8] 和文献 [9] 所提云计算环境下不可靠数据恢复方法进行比较，完成本次实验。

首先对不可靠数据恢复的准确度进行实验，实验通过在云环境下不可靠数据中随机抽取 50 组不可靠数据，观察恢复后数据的原始性。通过实验，得到三种不可靠数据恢复方法的恢复准确度对比，对比结果如图 1 所示。

通过图 1 可以看出，本文所提不可靠数据恢复方法数据恢复的准确度较高，且准确度较稳定，由于本文在进行数据恢复过程中，首先对数据进行保护和恢复重构，保证了数据的完整性，从而提高不可靠数据恢复的准确度。

表 1 是云计算环境下不可靠数据恢复方法的恢复时间 (s)

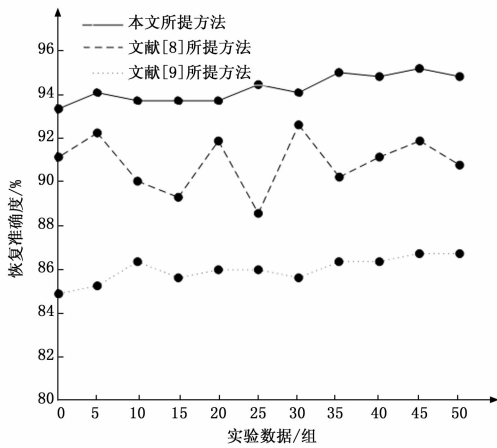


图 1 三种恢复方法的恢复准确度对比

对比。实验分别次利用三种不可靠数据恢复方法对相同数据进行恢复, 得到恢复时间对比, 对比结果如表 1 所示。

表 1 三种不可靠数据恢复方法恢复时间对比

实验次数/次	文献[8]方法 所用时间/s	文献[9]方法 所用时间/s	本文方法 所用时间/s
1	1.0	0.8	0.4
2	1.3	1.1	0.6
3	1.2	0.9	0.7
4	1.3	1.2	0.8
5	1.4	1.4	1.1
6	1.7	1.5	1.3
7	2.1	1.8	1.5
8	1.8	1.7	1.4
9	2.1	1.9	1.6
10	2.2	1.9	1.8

通过表 1 可知, 本文所提不可靠数据恢复方法的恢复耗时较短, 由于张量 Tucker 分解算法复杂度较低, 运行较快, 从而提高不可靠数据恢复方法的恢复速度。进行三种不可靠数据恢复方法进行数据恢复时所占存储空间 (MB) 对比, 实验通过对不同大小不可靠数据进行恢复, 对比恢复后数据大小, 对比结果如表 2 所示。

表 2 三种方法进行不可靠数据恢复所占存储空间对比

数据大小/KB	文献[8]方法 数据恢复所占空间/KB	文献[9]方法 数据恢复所占空间/KB	本文方法数 据恢复所占空间/KB
5	5.7	5.4	5.0
10	11.3	10.8	10.1
15	16.6	15.6	15.1
20	21.3	20.8	20.0
25	26.8	26.1	25.0
30	31.7	31.3	30.1
35	37.2	36.7	35.1

通过表 2 可知, 本文所提方法进行数据恢复后, 数据的大小变化较小, 由于本文针对张量 Tucker 分解算法存在的不足, 将 Tucker 分解算法与奇异值阈值函数相结合, 降低了不可靠数据分解的存储空间, 从而降低不可靠数据恢复所占空间。

对比三种云计算环境下不可靠数据恢复方法的可行性, 设可行性单位为%, 实验通过对不同事件类型的不可靠数据进行恢复, 对比恢复的准确度, 为保证实验的准确度, 对每种事件类型分别进行 500 次实验, 计算每种事件类型不可靠数据恢复的平均准确度作为可行性, 从而得到三种方法的可行性对比, 对比结果如表 3 所示。

表 3 三种方法的可行性对比

事件类型	文献[8]方法 数据恢复 可行性/%	文献[9]方法 数据恢复 可行性/%	文本方法数据 恢复可行 度/%
Normal	91.5	90.5	95.8
DOS/DDOS	87.5	89.6	93.4
Probc	90.8	91.3	96.7
Scan	88.2	88.6	94.2

通过表 3 可知, 本文所提不可靠数据恢复方法对不同事件类型的不可靠数据进行恢复, 可行性都达到 93% 以上, 说明本文所提方法适用于不同事件类型的不可靠数据恢复, 具有较强的可行性。

最后对三种云计算环境下不可靠数据恢复方法的能耗 (焦) 进行对比, 通过实验, 得到三种云计算环境下不可靠数据恢复方法的运行能耗对比, 对比结果如图 2 所示。

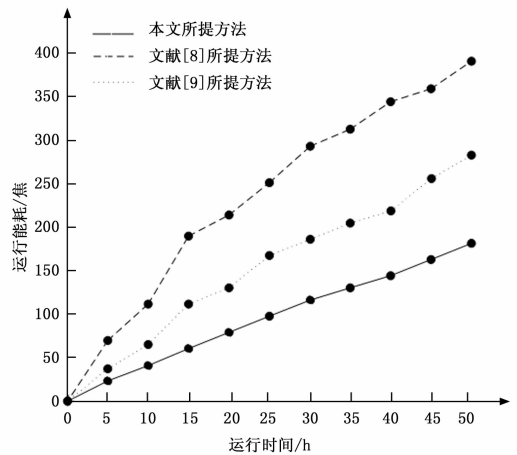


图 2 三种方法的运行能耗对比

通过图 2 可知, 本文所提不可靠数据恢复方法的运行能耗较少, 且本文所提方法能耗较均匀说明本文所提方法运行较为稳定。

综上所述, 本文所提云计算环境下不可靠数据恢复方法能够有效提高不可靠数据恢复的准确度, 降低不可靠数据恢复时间和能耗, 且数据恢复占用的内存较少, 具有良好的使用价值。

4 结束语

云计算环境下存在大量的不可靠数据, 但这些数据中可能含有重要信息, 不可以删除, 这就需要云计算环境下不可靠数据进行恢复。采用当前的不可靠数据恢复方法进行数据恢复时过程较复杂, 并且难以保证恢复后数据的准确度。提出一种基于张量 Tucker 阈值的云计算环境下不可靠数据恢复方法, 通过实验证明, 所提方法的准确度较高, 具有较强的可行性。

(下转第 35 页)

表 1 对比测试结果

mm

组别	1	2	3	4	5	6	7	8
平台位移	0.300	0.800	1.300	1.800	2.800	5.300	8.800	10.000
实验 1	0.302	0.799	1.302	1.801	2.798	5.300	8.798	9.998
实验 2	0.302	0.798	1.298	1.800	2.802	5.302	8.800	9.998
实验 3	0.301	0.798	1.298	1.802	2.800	5.298	8.802	10.002
实验 4	0.298	0.802	1.301	1.802	2.802	5.298	8.802	9.998
实验 5	0.299	0.800	1.299	1.802	2.799	5.301	8.798	10.002
实验 6	0.300	0.802	1.302	1.801	2.802	5.300	8.799	9.999
平均值	0.3003	0.7998	1.3000	1.8013	2.8005	5.2998	8.7998	9.9995
平均误差	0.0013	0.0015	0.0017	0.0013	0.0015	0.0011	0.0015	0.0019
标准差	0.0016	0.0018	0.0019	0.0008	0.0018	0.0016	0.0018	0.0020
重复性	0.0054	0.0023	0.0015	0.0005	0.0006	0.0003	0.0002	0.0002

5 结束语

基于最大长度序列靶标方案, 采用机器视觉, 实现了高精度一维绝对位移检测。采用二值化和灰度叠加的方法对靶标图像进行处理和绝对位置编码识别, 并根据拟合相位提取靶标图像精密位移信息实现了靶标识别和精密位移的检测。一维平移精密检测系统稳定可靠, 获得了满意的测量结果和精度, 检测精度达 $\pm 2 \mu\text{m}$, 并具有可重复性。本方法采用单相机测量, 相机无需标定, 检测系统结构简单, 工作效率高, 且环境适应性和鲁棒性良好, 具有广阔的应用前景。为进一步提高精度, 关于相机失焦、环境温度和光照均匀性等方面的影响值得进一步深入研究。

参考文献:

[1] 于乃功, 黄 灿, 林 佳. 基于单目视觉的机器人目标定位测距方法研究 [J]. 计算机测量与控制, 2012, (10): 2654 - 2656, 2660.

[2] 王炜辰. 三维激光测量仪的关键技术研究 [D]. 长春: 吉林大学, 2015.

[3] Brent Carlson. Geometry Study and Simulation of a Laser-Free Surface Measurement Technique for the Large Adaptive Reflector [R]. 1998.

[4] Orozco B V, Garduño V G. Structured laser light and coordinate measuring systems integration for 3d metrology [A]. Proc. of SPIE [C]. 2005, 5776: 618 - 629.

[5] 张淑平. 基于视觉的并联机器人位姿检测方法研究 [D]. 上海: 东华大学, 2010.

(上接第 31 页)

参考文献:

[1] 张 露, 尚艳玲. 云计算环境下资源调度系统设计与实现 [J]. 计算机测量与控制, 2017, 25 (1): 131 - 134.

[2] 周 虹, 陈 锋. 一种云计算共享数据完整性公开审计方案 [J]. 电子设计工程, 2016, 24 (9): 60 - 62.

[3] 崔立鲁, 朱贵发. 利用重力卫星数据恢复地球质量迁移方法的研究 [J]. 科学技术与工程, 2015, 15 (14): 106 - 109.

[4] 向劲松, 王 举, 吴 涛, 等. 基于脉冲展宽波形的光脉冲位置调制异步采样信号的数据恢复技术 [J]. 光学学报, 2017 (1): 74 - 81.

[5] 韩利华, 程 聪. 遮挡物体的图像恢复方法研究与仿真 [J]. 计算机仿真, 2016, 33 (11): 345 - 348.

[6] 陈若珠, 赵 珊. 基于机器视觉技术的振动台试验变形位移测量方法研究 [J]. 计算机测量与控制, 2016, 24 (6): 56 - 59.

[7] 张子森, 王 鹏, 孙长库. 单目视觉位姿测量方法及数字仿真 [J]. 天津大学学报, 2011, (05): 440 - 444.

[8] Hirata Seichironakajima Nobuuida Shogohojo Harumas. Ultra-sensitive system for measuring distance or position [P]. 国外专利: WO2011043091, 2011-04-14.

[9] 唐恒博. 轨道机器人的高鲁棒定位方法 [D]. 上海: 上海交通大学, 2013.

[10] 王 阳. 基于单码道绝对位置编码的光电式轴角编码器 [D]. 北京: 北京交通大学, 2007.

[11] 杨德华, 费 飞, 李开宇, 等. 一种基于视觉的位置检测编码靶标及系统 [P]. 中国专利: 201510201768. X, 2015-04-24.

[12] Johnson N, Mohan K J, Janson K E, et al. Optimization of incremental optical encoder pulse processing [A]. 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s) [C]. 2013: 769 - 773.

[13] Zago L, Mirsad Sarajlic, Chevalleya F, et al. Metrology system for the calibration of multi-dof precision mechanisms [A]. Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II [C]. 2012.

[14] 周文举. 基于机器视觉的在线高速检测与精确控制研究及应用 [D]. 上海: 上海大学, 2014.

[15] 张浩鹏. 视觉检测系统的若干关键问题研究 [D]. 哈尔滨: 哈尔滨工程大学, 2013.

[16] 李 徽, 杨德华, 翟 超. 六自由度机构位姿的单相机照相测量研究. [J] 光学技术, 2010, 36 (3): 344 - 349.

[6] 刘金明, 谢秋菊, 王 雪, 等. 基于 GSA-SVM 的畜禽舍废气监测缺失数据恢复方法 [J]. 东北农业大学学报, 2015, 46 (5): 95 - 101.

[7] 杨立东, 王 晶, 谢 湘, 等. 基于低秩张量补全的多声道音频信号恢复方法 [J]. 电子与信息学报, 2016, 38 (2): 394 - 399.

[8] 白晋国, 孙红胜, 胡泽明. 一种基于 SQLite3 文件格式的删除数据恢复方法 [J]. 小型微型计算机系统, 2016, 37 (3): 505 - 509.

[9] 陈红玉, 孟彩霞. 基于相干函数无偏估计的数据恢复技术 [J]. 科技通报, 2016, 32 (3): 105 - 108.

[10] 陈亦平, 陈伟彪, 姚 伟, 等. WAMS 错误数据的快速辨识及恢复方法 [J]. 电力自动化设备, 2016, 36 (12): 95 - 101.