

一种快速的三维扫描数据自动配准方法*

杨 琴⁺, 齐 越, 沈旭昆, 赵沁平

(北京航空航天大学 虚拟现实技术与系统国家重点实验室, 北京 100191)

Rapid and Automatic Method for 3D Scanned Data Registration

YANG Shen⁺, QI Yue, SHEN Xu-Kun, ZHAO Qin-Ping

(State Key Laboratory of Virtual Reality Technology and Systems, BeiHang University, Beijing 100191, China)

+ Corresponding author: E-mail: sheen.young@gmail.com, http://vrlab.buaa.edu.cn

Yang S, Qi Y, Shen XK, Zhao QP. Rapid and automatic method for 3D scanned data registration. *Journal of Software*, 2010,21(6):1338–1450. <http://www.jos.org.cn/1000-9825/3543.htm>

Abstract: This paper presents a rapid method to align large sets of 3D scanned data automatically. The method incorporates the technique of image registration into the pair-wise registration. Firstly, it retrieves two texture images from the scanned data to align. A method is proposed to generate the texture image from the range image when scanned data do not contain the texture information. Secondly, it detects the features using SIFT (scale-invariant feature transform) on texture images, and a set of potential corresponding pixels is selected by means of pre-filter and cross validation. Then a matching algorithm, based on RANSAC (random sample consensus) algorithm, is applied to specify the matching pixel pairs between two images. All matches obtained are mapped to 3D space and used to estimate the rigid transformation. Finally, a modified ICP (iterative closest point) algorithm is applied to refine the result. The paper also presents a method to create model graph rapidly for multi-view registration which avoids aligning all pairs of range images. This reconstruction technique achieves a robust and high performance in the application of automatic rebuilding 3D models of culture heritages.

Key words: 3D scanned data; range image; texture image; coarse registration; multi-view registration; image registration

摘 要: 研究了两幅和多幅深度图像的自动配准问题.在配准两幅深度图像时,结合二维纹理图像配准深度图像,具体过程是:首先,从扫描数据中提取纹理图像,特别地,针对不包含纹理图像的扫描数据提出了一种根据深度图像直接生成纹理图像的方法;然后,基于 SIFT(scale-invariant feature transform)特征提取纹理图像中的兴趣像素,并通过预过滤和交叉检验兴趣像素等方法从中找出匹配像素对的候选集;之后,使用 RANSAC(random sample consensus)算法,根据三维几何信息的约束找出候选集中正确的匹配像素对和相对应的匹配顶点对,并根据这些匹配顶点对计算出两幅深度图像间的刚体置换矩阵;最后,使用改进的 ICP(iterative closest point)算法优化这一结果.在配准多幅深度

* Supported by the National Natural Science Foundation of China under Grant Nos.60533070, 60773153 (国家自然科学基金); the Key Grant Project of the Ministry of Education of China under Grant No.308004 (国家教育部科学技术研究重大项目); the National High-Tech Research and Development Plan of China under Grant No.2009AA012102 (国家高技术研究发展计划(863)); the Beijing Municipal Natural Science Foundation of China under Grant No.4102037 (北京市自然科学基金)

Received 2008-09-17; Revised 2008-11-10; Accepted 2008-12-10; Published online 2010-05-06

图像时,提出了一种快速构建模型图的方法,可以避免对任意两幅深度图像作配准,提高了配准速度.该方法已成功应用于多种文物的三维逼真建模.

关键词: 三维扫描数据;深度图像;纹理图像;粗略配准;多幅图像配准;图像配准

中图法分类号: TP391 文献标识码: A

快速、逼真地重建真实物体和场景是虚拟现实和计算机图形学研究的热点问题.随着三维数据采集技术的发展,人们可以方便地使用三维扫描仪采集数据并完成模型重建.许多三维扫描仪不但能够获取模型表面顶点的深度信息,而且还能够获取它们的颜色信息,即可以同时采集到等分辨率的深度图像(range image)和纹理图像(texture image),深度图像也称作点云数据(point cloud data).为了获取物体表面的完整信息,需要在多个不同的视点扫描,再将各次采集的深度图像拼接起来,从而重建出完整的三维模型.因为不同视点处采集的深度图像处于各自独立的坐标系中,所以需要找出这些坐标系间的刚体置换矩阵,然后将所有的深度图像变换到同一个坐标系中,这一过程就是深度图像的配准(registration).目前,深度图像的配准一般需要人工参与完成,当处理大规模数据时,这样的配准过程是非常枯燥和耗时的.因此,研究者们开展了自动配准方法的研究.

深度图像配准方法一般分为两类:两幅图像配准(pair-wise registration)和多幅图像配准(multi-view registration).两幅图像配准算法的目标是找出存在重叠区域的两幅深度图像间最优的刚体置换矩阵,在匹配顶点的误差和最小的情况下,将其中的一幅深度图像拼接到另一幅上.两幅图像的配准首先要做粗略配准(coarse registration),计算出两幅深度图像间刚体置换矩阵的估计值;然后做精细配准(accurate registration),通过最小化预定义的误差函数优化粗略配准的结果.多幅图像配准算法的目标是拼接输入的多幅深度图像,首先利用两幅图像配准找出这些深度图像间的邻接关系,即哪些深度图像间存在重叠区域,然后再根据这些邻接关系将所有深度图像拼接起来,并最小化预定义的全局误差函数.

自动深度图像配准方法的难点是自动的两幅图像配准,而自动的两幅图像配准的难点又在于自动的粗略配准.大多数自动粗略配准方法^[1-5]都是基于局部几何特征匹配的,其主要做法是:1) 定义一种局部几何特征和它的计算方法;2) 基于每个顶点及其邻域中顶点的几何信息计算出所有顶点的特征值,再根据每个顶点的特征值,在待配准的两幅深度图像上分别找出具有显著特征的顶点;3) 通过比较这些顶点的特征值建立它们之间的对应关系;4) 根据正确的顶点对应关系计算出两幅深度图像间的刚体置换矩阵.一般地,低维特征的计算速度相对较快,但因为其蕴含的信息量少,所以这类特征对顶点的区分度较差,且对噪声敏感.高维特征包含了顶点及邻域内顶点的多维信息,但是这类特征的计算过程复杂并且耗时.因此,定义一种高效并可靠的局部几何特征是困难的.另外,由于环境光照和自遮挡等因素的影响,深度图像中普遍存在噪声和几何缺陷,如裂缝和空洞等(如图 1 所示).对此,较难单纯地使用基于局部几何特征的方法完成自动配准.

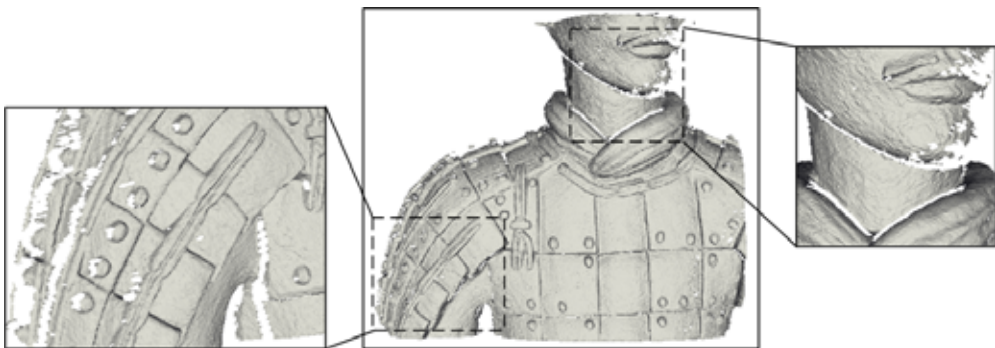


Fig.1 Range image which contains plenty of the geometry flaws

图 1 存在大量几何缺陷的深度图像

考虑到三维扫描仪可以同时采集到深度图像对应的纹理图像,研究者们也尝试结合纹理图像解决两幅深

度图像配准的问题^[6-9].这类配准算法的一般流程是:

- 1) 使用图像处理的方法提取两幅纹理图像上的若干兴趣像素;
- 2) 根据这些兴趣像素的特征找出它们之间的对应关系;
- 3) 根据匹配像素对在深度图像中对应的匹配顶点对完成配准.

这类方法的优点是加入了顶点的纹理信息,增加了配准时的信息量,一般可以更准确快捷地找出两幅深度图像间正确匹配的顶点.同时,由于纹理图像一般不受几何噪声影响,这类方法可以在一定程度上减小几何噪声对配准结果的影响.但是,由于扫描仪的功能、扫描对象的表面材质和环境光照等因素的影响,纹理图像可能不存在或者质量很差,也就无法找出两幅纹理图像间的匹配像素.此时,这类方法就无能为力了.针对这一问题,研究者们又提出了基于最小二乘共形映射(least squares conformal maps,简称 LSCM)将三维模型映射到二维空间再进行匹配的方法^[10].但是,这一方法要求被映射的三维模型同胚于盘,故其局限性较大,很难用于一般的存在几何缺陷的深度图像(如图 1 所示)的配准.

在配准多幅深度图像时,一般的配准方法都会基于两幅深度图像的配准构造出一个表示完整模型的多幅深度图像间拓扑关系的有向图,称为模型图(model graph),然后根据模型图完成所有数据的配准.绝大多数关于多幅图像配准的研究工作都在关注当模型图已知时如何减小全局误差和避免误差积累,而关于如何有效构造模型图的文献却极少.一般地,人们认为构造模型图需要对任意两幅深度图像都作一次配准.对于大规模三维数据的自动配准过程,这种方法是非常耗时的.

针对大规模三维扫描数据,本文对自动配准方法的研究同样着眼于这两类问题.对于输入的两幅深度图像,若它们对应的纹理图像已知,本文方法先基于图像的 SIFT(scale-invariant feature transform)^[11]特征,找出两幅纹理图像间所有可能正确的匹配像素对作为候选集,再利用深度图像中的几何信息从候选集中找出正确匹配的像素对,从而计算出这两幅深度图像间的刚体置换矩阵.当纹理图像未知或无法满足配准要求时,本文提出了一种快速地根据深度图像中顶点的几何特征生成纹理图像的方法.与基于局部几何特征匹配的方法^[1-5]相比,当纹理图像已知时,本文方法不需要计算每个顶点的特征值,且对存在噪声和几何缺陷的数据有极好的鲁棒性.当纹理图像未知时,本文方法将三维空间中顶点的低维几何特征映射到二维空间,然后利用图像配准方法和顶点的几何信息找出匹配的顶点,当处理复杂模型的扫描数据时,本文方法具有明显的速度优势.与已有的结合纹理图像的配准方法^[6-9]相比,本文方法引入了对兴趣像素预过滤和交叉检验的过程,并结合深度图像找出正确匹配的像素,提高了配准结果的准确性.同时,本文方法可以处理没有纹理信息的扫描数据,明显更具有通用性.与基于 LSCM 将三维模型映射到二维图像的方法^[10]相比,本文提出的映射方法不需要原始三维模型同胚于盘,适用性更广,而且由于不需要计算顶点的高维几何特征,其速度更快.对于输入的多幅深度图像,本文方法可以快速生成模型图,避免了对任意两幅深度图像都要做配准,因此显著地提高了多幅图像配准的速度.

1 相关工作

1.1 两幅图像配准方法

1.1.1 精细配准算法

在两幅深度图像配准的研究中,重要的成果有 Besl^[12]和 Chen^[13]提出的迭代最近点(iterative closest point,简称 ICP)算法.ICP 算法以两幅深度图像的起始位置为初始状态,通过一个迭代过程不断减小两幅深度图像间所有匹配顶点的距离和,一直达到某个阈值,最后计算出它们之间的刚体置换矩阵.由于 ICP 算法及其变种普遍需要有一个较好的初始状态以保证结果收敛,所以它们只能用于精细配准过程.在应用 ICP 算法之前,一般都需要使用粗略配准算法计算出待配准的两幅深度图像间的刚体置换矩阵的估计值,并将在此刚体置换下的两幅深度图像作为 ICP 算法的初始状态.

1.1.2 粗略配准算法

不同的基于局部几何特征匹配的粗略配准算法间的差异主要在于局部几何特征的定义和计算方法,如基于旋转图像(spin images)^[1]的方法.它的主要思想是,对于深度图像中的每一顶点,用一幅二维图像表示它及其

邻域顶点的几何信息,然后通过匹配这些二维图像完成配准.旋转图像属于高维特征,其最大缺点是计算过程的时间复杂度和空间复杂度都很高,且不易比较.当处理大规模数据时,使用高维几何特征的自动配准算法的效率都非常低.常用的局部几何特征还有基于物体表面微分特性的描述符^[2]、基于积分不变量的描述符^[3,4]和基于尺度空间的描述符^[5]等.这些局部几何特征的特征表示在刚体变换下保持不变甚至在一定噪声条件下也具有鲁棒性,但是由于低维特征和高维特征在配准速度和稳定性方面的固有矛盾,以及深度图像中存在的噪声和大量几何缺陷,对于大规模数据的配准,单纯使用局部几何特征难以取得较好效果.另外,对于任何一种局部特征描述符,每个顶点的特征值都是根据其邻域内顶点计算出来的,因此对于物体表面不同的顶点,如果它们邻域的几何特征相似,那么在上述顶点处将取得相近的特征值,这使得在两幅深度图像间寻找匹配的特征点成为一件困难的事情.加入额外的约束条件进行剪枝^[3]或者对特征做聚合^[4]都是解决这一问题的有效方法,但它们极大地增加了特征点匹配算法的时间复杂性.另一类粗略配准方法是基于随机抽样一致性(random sample consensus, 简称 RANSAC)算法^[14]的,其特点是速度较快但是缺乏稳定性.Chen^[15]通过在两幅深度图像上预先选取一些参考点和使用空间约束条件等策略,试图在配准速度和结果的稳定性间找到较好的平衡点.

在结合纹理图像的深度图像配准研究方面,Roth^[6]通过匹配兴趣像素对应的顶点组成的三角形寻找正确的匹配关系,这种方法的效率和准确性都不高.Bendels^[7]先利用 SIFT 特征找出匹配像素和对应的匹配顶点,再利用 RANSAC 方法过滤不正确的匹配关系,最终完成配准.Bendels 给出了结合纹理图像配准深度图像的基本框架.Seo^[8]在 Bendels 方法的基础上利用光学校正技术处理纹理图像间存在的三维旋转和投影扭曲问题,但是只能校正平面或近似平面区域的纹理,所以很难用于表面几何形状复杂的物体.刘晓利^[9]使用 Sobel 算子提取纹理图像上的特征像素,然后通过归一化的相关系数找出匹配的特征像素,最后通过 Hausdorff 距离检验这些匹配像素的有效性.这种方法单纯依赖纹理图像的信息找出匹配像素,且需要人机交互选出重叠区域,因此只适用于简单的、拍摄角度变化不大的纹理图像的配准.以上方法^[6-9]最大的局限性是均不适用于纹理图像不存在的情形.Wang^[10]提出了一种利用 LSCM 先将两个三维模型映射到二维空间再作匹配的方法:首先根据曲率较大的顶点的旋转图像找出两个三维模型间的若干对应顶点,再以这些对应顶点作为约束条件完成 LSCM,最后利用归一化的相关系数(normalized correlation coefficient)匹配映射得到的二维图像.Wang^[10]的方法不但要求待匹配的三维模型同胚于盘,而且在映射过程中需要计算每一顶点的曲率和曲率较大顶点的旋转图像并求解一个大型的稀疏线性系统,从通用性和速度两方面考虑,这种方法都较难应用到一般的大数据量深度图像的自动配准中.

1.2 多幅图像配准方法

Pingi^[16]通过预先定义一些扫描策略简化构造模型图时普遍使用的穷举算法.Pingi 认为,用户在扫描一个物体时应该以条带状的路线完成扫描.例如,用户会按照由上到下、由左到右或者以环绕物体的方式进行扫描,并且每两个扫描条带间首尾相接,即前一个条带的最后一幅深度图像和下一个条带的第 1 幅深度图像存在重叠区域.如果完全依照这种策略完成扫描,那么在最终的深度图像序列中,相邻数据间均应有重叠区域,这样就能在线性时间复杂度内构造出模型图.然而,这种面向条带的扫描和配准策略过于理想化,在实际应用中很难保证深度图像序列中任意相邻的数据间均存在重叠区域.

2 两幅深度图像的配准

本文的两幅深度图像配准算法同样包括两个步骤:粗略配准和精细配准.粗略配准分 4 步完成:

- 1) 若扫描数据中已包含顶点纹理信息,则直接提取出等分辨率的纹理图像和深度图像;否则,根据深度图像生成一幅纹理图像;
- 2) 分别从两幅纹理图像中提取显著的 SIFT 特征,并根据这些特征的 SIFT 特征向量找出两幅纹理图像间可能正确的匹配像素对作为候选集;
- 3) 根据候选集中像素对应的顶点的几何信息,使用 RANSAC 算法找出其中正确匹配的像素对;
- 4) 找出正确匹配像素对在深度图像中对应的顶点对,并使用基于四元数的方法^[17]计算出两幅深度图像

间刚体置换矩阵的估计值.

在精细配准阶段,在此刚体置换下的两幅深度图像作为初始状态,使用改进的 ICP 算法优化这一结果.图 2 是两幅深度图像配准过程中一些关键步骤的结果.

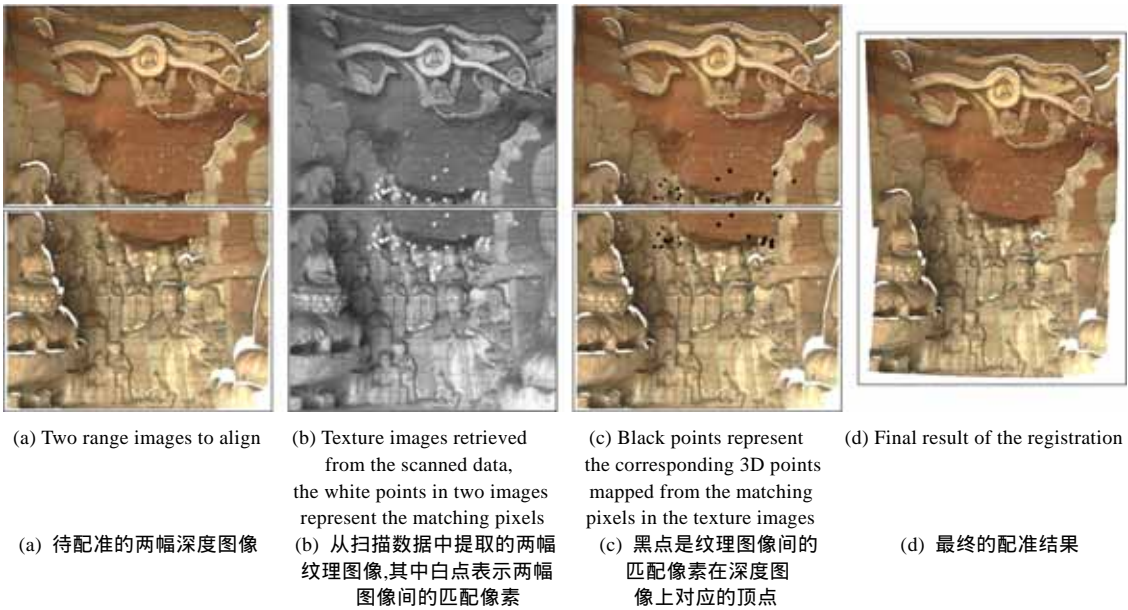


Fig.2 Registration procedure for two range images

图 2 两幅深度图像的配准过程

2.1 粗略配准算法

2.1.1 算法流程总述

对于三维扫描仪每次扫描获得的数据集,其中必然包含一幅深度图像,并可能包含一幅和深度图像对应的纹理图像.当纹理图像不存在时,本文算法先根据深度图像生成一幅纹理图像,故本文算法中总是认为,从一次扫描数据集中可以获取一幅深度图像和一幅纹理图像.设 P 和 Q 是在两个不同视点处扫描获得的数据集,用 R_p 和 I_p 分别表示从 P 中获取的深度图像和纹理图像,用 R_q 和 I_q 分别表示从 Q 中获取的深度图像和纹理图像,以下结合图 3 介绍本文粗略配准算法的步骤:

- 1) 从 P 和 Q 中提取 R_p 和 R_q ,若 I_p 和 I_q 存在,则直接提取 I_p 和 I_q ;否则,由 R_p 和 R_q 生成 I_p 和 I_q (如图 3(a)所示);
- 2) 使用 SIFT 算法分别找出 I_p 和 I_q 中的特征像素(如图 3(b)所示),再从这些特征像素中找出 m 对可能正确的匹配像素对作为候选集 $C = \{(i_k^p, i_k^q) | i_k^p \in I_p, i_k^q \in I_q, k = 1, \dots, m\}$ (如图 3(c)所示);
- 3) 从 C 中选出正确匹配的像素对作为结果集 $C_r = \{(i_k^p, i_k^q) | i_k^p \in I_p, i_k^q \in I_q\}$, $C_r \subseteq C, |C_r| \geq 3$ (如图 3(d)所示);
- 4) 将 C_r 中的元素映射到深度图像,得到 R_p 和 R_q 间的匹配顶点集合,并根据这些匹配顶点计算出 R_p 和 R_q 间的刚体置换矩阵.

在深度图像和纹理图像间存在映射 $F_p:R_p \rightarrow I_p$ 和 $F_q:R_q \rightarrow I_q$,由于本文算法在第 2 步、第 3 步中可以保证结果集 C_r 中的像素对在深度图像中都能找到其对应的顶点对,即 $\forall (i_k^p, i_k^q) \in C_r, \exists v_k^p \in R_p, v_k^q \in R_q$ 使得 $F_p(v_k^p) = i_k^p, F_q(v_k^q) = i_k^q$,故一旦找出 C_r ,即可根据 F_p 和 F_q 的逆映射找出 R_p 和 R_q 间的匹配顶点集合,再用基于四元数的方法^[17]求解刚体置换矩阵.下面重点说明前 3 步的具体实现.

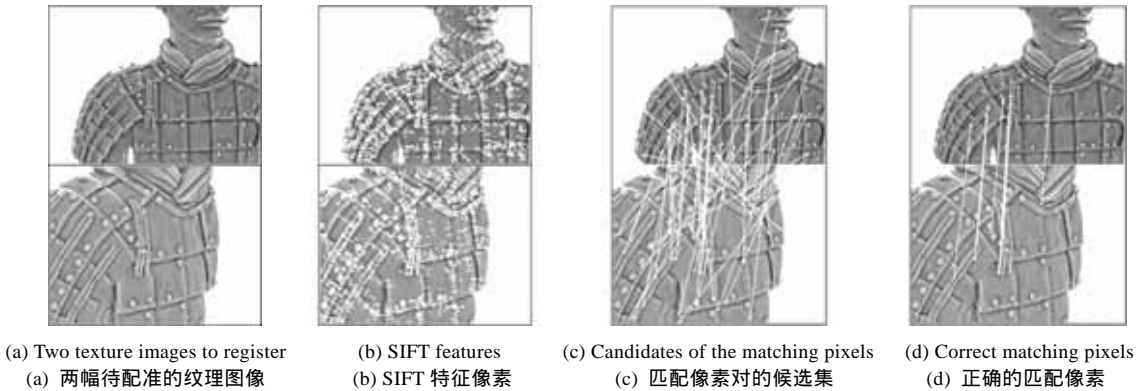


Fig.3 Registration procedure for two texture images

图 3 两幅纹理图像的配准过程

2.1.2 提取或生成纹理图像

若 P 和 Q 中已包含纹理图像,则直接提取它们并转换成灰度图像作为 I_p 和 I_q ;若 P 和 Q 中不包含纹理图像,或者包含的纹理图像无法配准时,则首先根据深度图像生成纹理图像.

给定一幅深度图像 R ,则 $R(x, y) = \begin{cases} v_{x,y} \in \mathcal{R}^3, & (x, y) \text{处有顶点} \\ \emptyset, & (x, y) \text{处没有顶点} \end{cases}$, 设 f 是一种几何特征描述符,对于顶点 $v, f(v)$

表示 v 处的特征值,可以对所有顶点的特征值做线性变换,使得它们的值为 $[0,1]$.那么定义由 f 生成的纹理图像为 $I_f(x, y) = \begin{cases} f(R(x, y)), & R(x, y) \neq \emptyset \\ \emptyset, & R(x, y) = \emptyset \end{cases}$, 这样即可根据三维空间的几何特征 f 生成一幅二维纹理图像.理论上,任何一种几何特征描述符都可以被用来生成纹理图像,但考虑到算法的效率和结果的准确性,则希望 f 具有很好的区分度,并且易于计算.

本文方法生成纹理图像的基本思想是,设扫描数据中包含的深度图像是 $R^0 = \{v_i^0 \mid v_i^0 \in R^3, i = 1, \dots, n\}$, 而 $R^k = \{v_i^k \mid v_i^k \in R^3, i = 1, \dots, n\}$ 是对 R^0 做 k 次平滑后得到的基模型,其中, $v_i^k \in R^k$ 是 $v_i^0 \in R^0$ 经平滑后的对应点,那么定义 $f_\delta(v_i^0) = \text{sign}(v_i^0) \cdot \|v_i^0 - v_i^k\|, i = 1, \dots, n$, 其中, $\text{sign}(v_i^0) = \begin{cases} 1, & (v_i^0 - v_i^k) \cdot \text{norm}(v_i^0) \geq 0 \\ -1, & (v_i^0 - v_i^k) \cdot \text{norm}(v_i^0) < 0 \end{cases}$, $\text{norm}(v_i^0)$ 是 v_i^0 处的法向量.

具体实现中,本文使用基于网格的 Laplacian 平滑得到基模型.由于深度图像 R^0 中已蕴含点云的二维结构信息,故网格化 R^0 是简单而快速的^[18].在平滑过程中,顶点间拓扑关系保持不变,设 $M^k = (V^k, E)$ 为 R^k 对应的三角网格,

其中, V^k 和 E 分别代表顶点和边的集合, ∂M^k 是 M^k 的边界. $\forall v_i^k = (x_i^k, y_i^k, z_i^k) \in M^k$, 则 $v_i^{k+1} = \begin{cases} \frac{1}{d_i} \sum_{j \in N(i)} v_j^k, & v_i^k \notin \partial M^k \\ v_i^k, & v_i^k \in \partial M^k \end{cases}$,

其中, $N(i) = \{j \mid (i, j) \in E\}, d_i = |N(i)|$.一般地,平滑次数 k 越大,深度图像中的噪声在生成的纹理图像中的影响越小;同时,较大的 k 值更有利于 SIFT 算法在纹理图像中检测出原始深度图像中显著的几何特征.但是 k 值越大,生成纹理图像的效率越低.在实验中,我们一般取 $k=64$,这时,可以在配准效率和准确性间取得一个较好的平衡点.图 4 和图 5 分别是西安兵马俑模型和斯坦福大学 Buddha 模型的深度图像经过不同次数平滑后根据原始模型和基模型生成的纹理图.

与基于 LSCM^[10]的方法相比,在将由深度图像生成的网格模型映射到纹理图像的过程中,本文方法不要求原始的三维模型同胚于盘,明显地更具有通用性.同时,本文方法不需要计算每一顶点的高维几何特征和求解大型稀疏线性系统,因此速度更快.

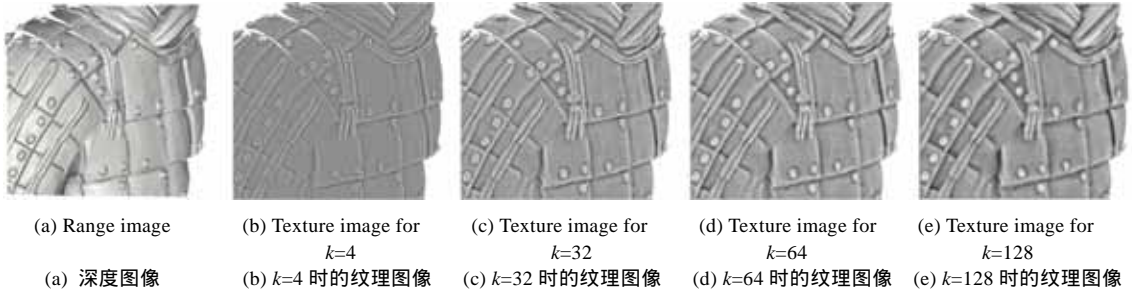


Fig.4 Range image of Terra-cotta Warrior model and the texture images generated from it
图 4 兵马俑模型的深度图像和由它生成的纹理图像

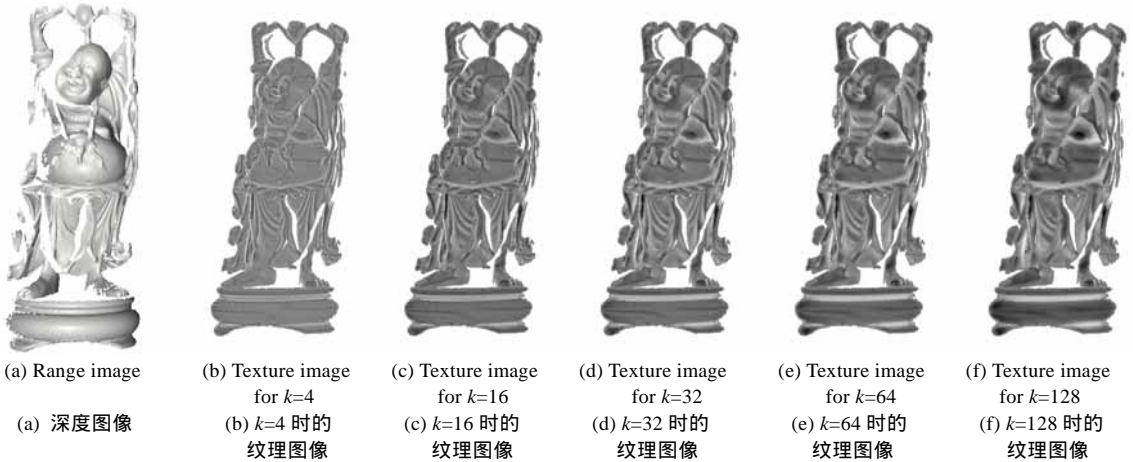


Fig.5 Range image of Buddha model and the texture images generated from it
图 5 Buddha 模型的深度图像和由它生成的纹理图像

2.1.3 找出纹理图像间匹配像素的候选集

考虑到 SIFT 特征不但对图像的缩放和旋转具有不变性,同时在光照变化、噪声、遮挡和较小的视点变化等条件下也有较好的鲁棒性,本文首先使用 SIFT 算法分别找出 I_p 和 I_q 中的兴趣像素集(如图 3(b)所示),然后在 I_p 和 I_q 的兴趣像素集之间寻找匹配像素.对于 SIFT 算法检测出的每一个兴趣像素 i ,用 $sift(i)$ 表示 i 的 SIFT 特征向量,则 $d(i,j)=|sift(i),sift(j)|$ 可以表示两个兴趣像素 i 和 j 之间的相似程度.这个距离越小,则这两个像素越相似;反之,距离越大,它们的差别越大.一般地,使用 SIFT 算法会找出较多的兴趣像素.同时,由于深度图像间可能存在三维角度变化以及纹理图像自身的复杂性,简单使用 SIFT 特征向量的欧氏距离作为评价标准来寻找匹配像素经常会引入大量错误的匹配.因此,本文采用以下步骤找出两幅纹理图像间匹配像素的候选集:

- 1) 对于纹理图像中的兴趣像素 i ,如果 i 在深度图像中对应的顶点不存在,则删除 i ;
- 2) 对于纹理图像中的兴趣像素 i ,如果 i 在深度图像中对应的顶点位于深度图像生成网格的边界或边界顶点的 δ 环邻域上,则删除 i ,一般取 $1 \leq \delta \leq 3$;

3) 通过交叉检验的方法找出 I_p 和 I_q 之间匹配像素的候选集.对于 I_p 中的每一个兴趣像素 i^p ,在 I_q 中找出使 $d(i^p, i_1^q)$ 最小的两个兴趣像素 i_1^q 和 i_2^q ,不妨设 $d(i^p, i_1^q) \leq d(i^p, i_2^q)$,定义 i_1^q 和 i_2^q 之间的差异度为 $diff_{i^p}(i_1^q, i_2^q) = \frac{d(i^p, i_2^q) - d(i^p, i_1^q)}{d(i^p, i_1^q)}$,并预先定义差异度阈值 δ_{diff} .如果 $diff_{i^p}(i_1^q, i_2^q) > \delta_{diff}$,则认为 i_1^q 和 i_2^q 之间有明显差异, i_1^q 可作为 i^p 的对应像素;否则认为 i^p 在 I_q 中有多重对应,应舍弃 i^p .对于前者,再以同样的方法在 I_p 中找出像素 i_1^q 的对应像素 \bar{i}^p ,如果 \bar{i}^p 和 i^p 为同一个像素,则认为 i^p 和 i_1^q 是 I_p 和 I_q 之间的一对候选匹配像素,将 (i^p, i_1^q) 加入候选集 C ;否

则,仍认为 i^p 在 I_q 中没有对应像素.在实验中,一般取 $\delta_{diff}=0.2$.

2.1.4 根据深度图像选取正确匹配像素的集合

单纯使用 SIFT 特征向量找出的对应像素候选集 C 中仍包含错误的匹配,需要引入其他的方法找出 C 中正确匹配的像素对.这是计算机视觉的经典问题,传统方法一般通过比较两个像素的邻域的 NCC(normalized cross-correlation)值、加入极线约束条件以及使用 Hausdorff 距离^[9]等方法来检验它们是否正确匹配.另外, Lhuillier 提出的基于准密繁殖(quasi-dense propagation)的匹配算法^[19]也可被用来找出正确匹配的像素对,且稳定性较高.但是,这些都是完全基于纹理信息的匹配方法,在像素对应的顶点的三维信息已知的情况下,完全可以用三维空间中的约束条件来判断匹配的正确性.

本文结合顶点的三维信息,使用一种自适应的 RANSAC 算法,从候选集 C 中找出正确匹配的像素对.首先对 C 中所有元素 (i^p, i^q) 根据 $d(i^p, i^q)$ 以升序进行排序,并将结果放入一个队列 L 中.由于恢复一个刚体置换矩阵最少只需要三维空间中不共线的 3 对匹配顶点,因此,本文只检验 L 中前 l 个元素是否是正确的匹配.实验中一般取 $l=50$.具体的检验方法是:

- 1) 从 L 的前 l 个元素中随机选取 3 个元素作为一个样本;
- 2) 根据这 3 对匹配像素在深度图像中对应的匹配顶点的几何信息计算出刚体置换 $T^{[17]}$;
- 3) 根据对应的顶点的几何信息,判断其余 $l-3$ 个元素在此刚体置换下是否为正确的匹配: $\forall (i^p, i^q) \in C$, 设 v^p 和 v^q 分别是 i^p 和 i^q 对应的顶点,如果 $\|T(v^p), v^q\|$ 小于预定义的阈值,则认为 (i^p, i^q) 是正确匹配的元素;否则, (i^p, i^q) 是错误的匹配.正确匹配的元素记为内点(inlier),错误的记为外点(outlier),所有内点组成本次采样的一致集(consensus set),然后根据内点的数目更新采样次数上限;
- 4) 重复以上 3 步,直到循环次数达到采样次数上限,最后将内点数目最多的采样的一致集作为结果集.

2.2 精细配准算法

一旦计算出两幅深度图像间刚体置换矩阵的估计值,就可以使用 ICP 算法或其变种以在此刚体置换下的两幅深度图像为初始状态,完成迭代优化.在精细配准过程中,本文改进了经典的 ICP 算法:首先,在迭代过程中不考虑位于深度图像边界上和临近边界的顶点,这样可以取得更好的准确性;其次,本文方法可依照每次迭代时匹配顶点的距离和动态地调整距离阈值,这样在每次迭代过程中,只有阈值范围内的匹配顶点用于计算新的变换矩阵和最小化匹配顶点的距离和.

在精细配准过程中,如果匹配顶点的距离和不收敛,或者收敛于一个超过了预先定义阈值的值,则认为无法配准这两幅深度图像,这表明两幅深度图像间粗略配准的结果不正确或根本不存在重叠区域.在多幅深度图像配准的过程中,这一准则被用来判别两幅深度图像是否有重叠区域.

3 多幅深度图像的配准

配准多幅深度图像时,设表示深度图像间拓扑关系的模型图为 $G(V, E)$.图 G 的每个顶点 $v_i \in V$ 均表示一幅深度图像,图 G 的每一条边 $e_{i,j}$ 表示顶点 v_i 和 v_j 代表的两幅深度图像间有重叠区域,且这条边的权值是从 v_i 到 v_j 的刚体置换矩阵.易知,若 $e_{i,j}$ 存在则 $e_{j,i}$ 也存在,且其权值是 $e_{i,j}$ 权值的逆矩阵.图 G 的生成树表示由这些深度图像构建完整模型的最小元素集,因此一旦构建出图 G ,就可以通过寻找 G 的生成树来重建完整的模型.

设 $S=(s_1, s_2, \dots, s_n)$ 是待配准的 n 幅深度图像的输入序列,当配准 S 时,本文方法同样需要建立模型图 $G(V, E)$.尽管可以对 S 中任意两幅深度图像作一次配准以判断它们是否有重叠区域,若有则计算出相关的刚体置换矩阵,从而构造出图 G 中所有的边,然而当处理大规模数据时,这是一项非常耗时的任务.在 Pingi^[16]的扫描策略下, $\forall s_i \in S(1 < i < n)$, s_i 与 s_{i-1} 及 s_{i+1} 间均存在重叠区域,这样就可以依次对相邻的两幅深度图像 $s_i, s_{i+1} \in S(i=1, \dots, n-1)$ 作配准,构造出图 G 中的边,并保证图 G 是连通的.然而, Pingi 的这种扫描和配准策略过于理想化,在实际中很难保证 S 中相邻的两幅数据 $s_i, s_{i+1} \in S(i=1, \dots, n-1)$ 间均存在重叠区域.

不失一般性,本文假定用户在扫描时会按照一定顺序,即以条带状的方式完成扫描,但是每两个扫描条带间不一定连续,即前一条带的最后一幅深度图像和下一条带的第 1 幅深度图像间不一定存在重叠区域.在此假设

下,序列 S 可以被分成 k 个条带: $S=(W_1, W_2, \dots, W_k), \forall W_i \in S$ 均满足 Pingi 的假设,即 W_i 中任意两幅相邻的深度图像都有重叠区域.

本文采用下面的方法构造出图 G 的边.

首先,将 S 分成若干条带组成的序列:1) 将 s_1 加入第 1 个条带 W_1 中;2) 依次配准 S 中所有相邻的深度图像 $s_i, s_{i+1} \in S(i=1, \dots, n-1)$, 如果 s_i 与 s_{i+1} 有重叠区域,则将 s_{i+1} 加入 s_i 所在的条带中,并将边 $e_{i,i+1}$ 和 $e_{i+1,i}$ 加入图 G ; 否则,将 s_{i+1} 加入一个新条带.

然后,合并条带序列:对条带 $W_i(i=2, \dots, k)$ 中的每幅深度图像,依次检索条带 $W_j(j=i-1, \dots, 1)$ 中是否存在深度图像和 W_i 中的深度图像有重叠区域.如果存在这样的深度图像,那么就可以将 W_i 和 W_j 合并,即将 W_j 中的各元素插入到 W_i 的首元素之前,同时在图 G 中加入相关的边.这一合并条带的过程不断循环,直到只剩余一个条带,或者在任意两个剩余条带间都不存在重叠区域.前者表明图 G 已是连通的,即可以找出它的一棵生成树并构造出完整的模型;而在后一情形下,每个剩余的条带均代表模型的一个部分,但是使用本文的自动配准算法无法将这些部分再拼接起来.

最后,根据图 G 的生成树重建出整个模型.

图 6 以实例说明了多幅深度图像的配准过程:对于输入的 13 幅深度图像,先对所有相邻的两幅数据做配准,这样可以将整个扫描序列分成 5 个条带,然后合并这 5 个条带,最终生成一个完整的模型.与传统的需要配准任意两幅深度图像的多幅配准方法相比,采用本文的方法,同一条带中的深度图像间不会再做配准,因此显著地减小了整个多幅配准过程的时间.易知, Pingi 提出的扫描策略可以视为本文方法的一个特例.

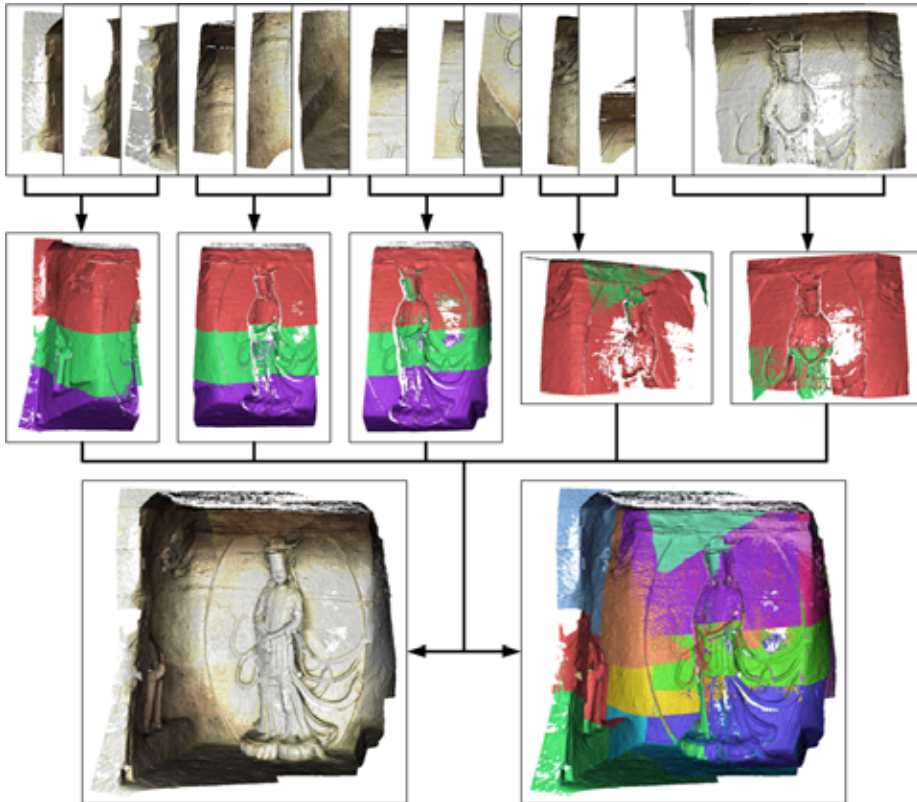


Fig.6 Procedure of multi-view registration

图 6 多幅深度图像的配准过程

4 实验结果和讨论

我们用 C++实现了本文的三维扫描数据的自动配准算法,并将其应用到斯坦福大学的扫描数据以及我们用 Minolta VIVID 910 三维扫描仪采集的重庆大足石刻和西安兵马俑扫描数据的自动配准中.其中,我们采集的扫描数据的分辨率是 640×480,每幅深度图像中包含了约 20~30 万左右的顶点,每幅纹理图像包含超过 30 万的颜色值.对于这样的数据,在一台 2.2GHz Core2 2G 内存的普通 PC 上,当扫描数据中包含的纹理图像可以配准时,两幅深度图像粗略配准的时间一般不超过 3s.这一结果明显快于传统的基于局部几何特征描述符的方法^[2,3,5],实验数据见表 1.图 6~图 8 分别采集自重庆大足石刻的数珠观音、Dazu(I)和 Dazu(II)模型的配准结果.当纹理图像无法配准(西安兵马俑和大足石刻龙柱模型数据)或没有纹理数据(Buddha 和 Armadillo 模型数据)时,使用深度图像生成纹理图像的时间约为 3s~10s.这一时间是由深度图像生成网格的复杂程度和做 Laplacian 平滑的次数决定的,实验数据参见表 2.图 9 是 Buddha 和 Armadillo 模型的配准结果,图 10 是西安兵马俑模型和大足石刻龙柱模型的配准结果.

在多幅数据配准过程中,因为纹理图像和其 SIFT 兴趣像素在配准过程中保持不变,所以使用深度图像生成纹理图像以及计算纹理图像的 SIFT 特征都只需做一次.这样在合并条带的过程中,两幅数据的配准速度会有较大提升.在我们的测试数据中,规模最大、几何信息最复杂的是 Dazu(II)洞窟,其原高 2.72m,宽 1.95m,深 2.40m.它的扫描数据包含 99 幅深度图像,总共 2 700 万左右的顶点,并且深度图像序列对应的扫描路线并不规则,存在多次跳跃的情况.对于这样的数据,使用本文方法仅需要约 30 分钟就能完成全部扫描数据的自动配准.

Table 1 Result of registration with the texture images

表 1 包含纹理图像的扫描数据的配准结果

Dataset	Total range images	Total points	Total strips	Total time (s)
Kwan-yin	13	3 127 391	5	32
Dazu (I)	45	12 161 233	8	220
Dazu (II)	99	27 173 866	23	1 794

Table 2 Result of registration with the texture images generated from range images

表 2 用深度图像生成纹理图像的扫描数据的配准结果

Dataset	Total range images	Total points	Total strips	Total time (s)
Buddha	15	1 099 005	1	46
Armadillo	23	578 375	7	47
Terra-cotta Warrior	54	10 932 156	10	602
Dragon Pillar	49	9 117 437	13	491

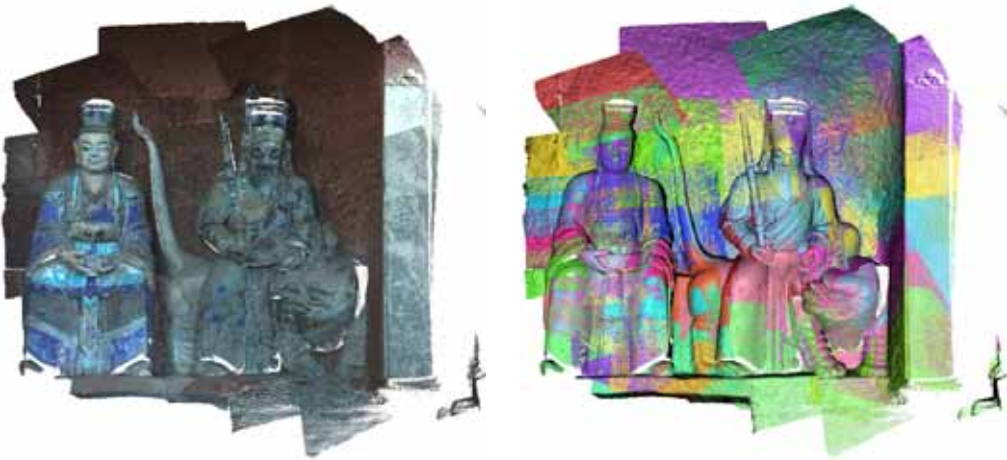


Fig.7 Registration result of Dazu (I) model

图 7 Dazu(I)模型的配准结果



Fig.8 Registration result of Dazu (II) model

图 8 Dazu(II)模型的配准结果

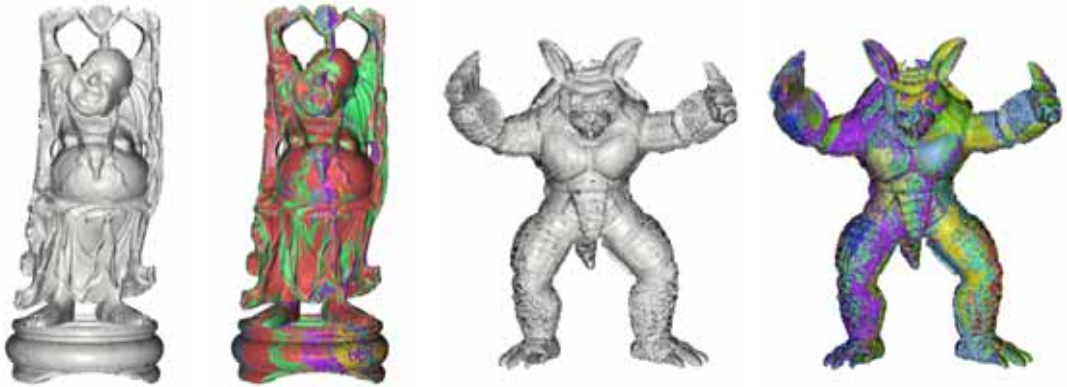


Fig.9 Registration result of Buddha and Armadillo models

图 9 Buddha 和 Armadillo 模型的配准结果

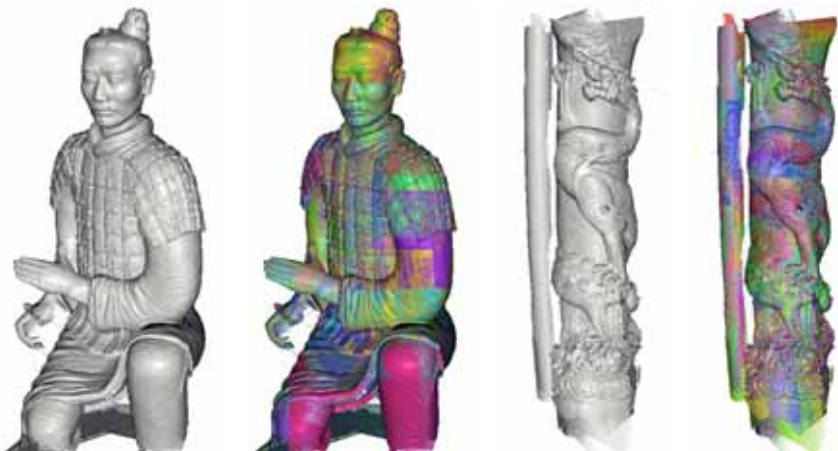


Fig.10 Registration result of Terra-cotta Warrior and Dragon Pillar models

图 10 兵马俑和龙柱模型的配准结果

另外,本文的多幅深度图像的配准方法非常适用于以条带状或近似条带状路线扫描的大规模数据的配准,每一条带中包含的数据越多,算法的效率就越高.例如,对于数珠观音模型(如图6所示),使用穷举算法需要做78次两幅图像配准,而本文方法只需要做16次就能完成所有数据的配准.如果输入的深度图像是随机顺序的,则本文方法会退化到和穷举算法近似的时间复杂度.

在实现中,我们使用 kd -tree来寻找与给定兴趣像素欧氏距离最近的像素.在处理大数据集时,由于无法将所有的深度图像同时载入内存,因此数据一般在使用时才被载入,不使用时则被从内存中移除.由于载入深度图像通常会花费一定时间,因此,我们使用一个LRU(least recently used)高速缓存来管理深度图像的加载和移除.

5 结束语

针对大规模的三维扫描数据,本文研究了两幅和多幅深度图像自动配准的问题.在粗略配准阶段,本文方法结合图像配准技术,利用从扫描数据中提取或由深度图像生成的纹理图像和深度图像中的三维信息找出两幅深度图像间匹配的特征点,再根据这些匹配特征点计算出两幅深度图像间刚体置换矩阵的估计值.在多幅数据配准阶段,本文方法可以快速生成模型图,并完成整个模型的重建.本文方法已成功应用于多种文物的三维逼真建模中,实践证明,对于大规模三维扫描数据,本文的自动配准方法具有很高的效率和很好的准确性.

当然,本文方法也存在局限性.由于本文方法利用纹理图像的配准来寻找两幅深度图像间的匹配顶点,这就要求纹理图像具有较高的分辨率且其像素间有明显亮度差异.对于不包含纹理图像的扫描数据,若深度图像分辨率过低,由其生成的纹理图像则无法配准,从而导致深度图像无法配准.因此,本文算法更适用于配准包含了较丰富纹理信息的扫描数据,或不包含纹理信息但包含了分辨率较高且具有显著几何特征的深度图像.

References:

- [1] Johnson A. Spin-Images: A representation for 3-D surface matching [Ph.D. Thesis]. Pittsburgh: Carnegie Mellon University, 1997.
- [2] Gal R, Cohen-Or D. Salient geometric features for partial shape matching and similarity. *ACM Trans. on Graphics*, 2006,25(1):130–150. [doi: 10.1145/1122501.1122507]
- [3] Gelfand N, Mitra, NJ, Guibas LJ, Pottmann H. Robust global registration. In: Desbrun M, Pottmann H, eds. *Proc. of the 3rd Eurographics Symp. on Geometry Processing (SGP 2005)*. Vienna: Eurographics Association, 2005. 197–206.
- [4] Huang QX, Flory S, Gelfand N, Hofer M, Pottmann H. Reassembling fractured objects by geometric matching. *ACM Trans. on Graphics*, 2006,25(3):569–578. [doi: 10.1145/1141911.1141925]
- [5] Li X, Guskov I. Multi-Scale features for approximate alignment of point-based surfaces. In: Desbrun M, Pottmann H, eds. *Proc. of the 3rd Eurographics Symp. on Geometry Processing (SGP 2005)*. Vienna: Eurographics Association, 2005. 217–228.
- [6] Roth G. Registering two overlapping range images. In: *Proc. of the 2nd Int'l Conf. on Recent Advances in 3-D Digital Imaging and Modeling (3DIM 1999)*. Ottawa: IEEE Computer Society Press, 1999. 191–200.
- [7] Bendels GH, Degener P, Wahl R, Kortgen M, Klein R. Image-Based registration of 3D-range data using feature surface elements. In: Fellner DW, Spencer SN, eds. *Proc. of the 5th Int'l Symp. on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST 2004)*. Brussels: Eurographics Association, 2004. 115–124.
- [8] Seo JK, Sharp GC, Lee SW. Range data registration using photometric features. In: *Proc. of the 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2005)*. San Diego: IEEE Computer Society Press, 2005. 1140–1145.
- [9] Liu XL, Peng X, Li AM, Gao PD. Range images registration combined with texture information. *Journal of Computer-Aided Design & Computer Graphics*, 2007,19(3):340–345 (in Chinese with English abstract).
- [10] Wang S, Wang Y, Jin M, Gu X, Samarasinghe D. Conformal geometry and its applications on 3D shape matching, recognition, and stitching. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2007,29(1):1209–1220. [doi: 10.1109/TPAMI.2007.1050]
- [11] Lowe DG. Distinctive image features from scale-invariant keypoints. *Int'l Journal of Computer Vision*, 2004,60(2):91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
- [12] Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1992,14(2):239–256. [doi: 10.1109/34.121791]

- [13] Chen Y, Medioni G. Object modeling by registration of multiple range images. *Image and Vision Computing*, 1992,10(3):145–155. [doi: 10.1016/0262-8856(92)90066-C]
- [14] Fischler MA, Bolles RC. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981,24(6):381–395. [doi: 10.1145/358669.358692]
- [15] Chen CS, Hung YP, Cheng JB. Ransac-Based darces: A new approach to fast automatic registration of partially overlapping range images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1999,21(11):1229–1234. [doi: 10.1109/34.809117]
- [16] Pingi P, Fasano A, Cignoni P, Montani C, Scopigno R. Exploiting the scanning sequence for automatic registration of large sets of range maps. *Computer Graphics Forum*, 2005,24(3):517–526. [doi: 10.1111/j.1467-8659.2005.00877.x]
- [17] Horn B. Closed-Form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, 1987,4(4): 629–642.
- [18] Turk G, Levoy M. Zippered polygon meshes from range images. In: *Proc. of the 21st Annual Conf. on Computer Graphics and Interactive Techniques (SIGGRAPH 1994)*. New York: ACM Press, 1994. 311–318. <http://portal.acm.org/citation.cfm?id=192241>
- [19] Lhuillier M, Quan L. Match propagation for image-based modeling and rendering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,24(8):1140–1146. [doi: 10.1109/TPAMI.2002.1023810]

附中文参考文献:

- [9] 刘晓利,彭翔,李阿蒙,高鹏东.结合纹理信息的深度像匹配. *计算机辅助设计与图形学学报*,2007,19(3):340–345.



杨琴(1980 -),男,甘肃兰州人,博士生,主要研究领域为虚拟现实,计算机图形学.



沈旭昆(1965 -),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为虚拟现实与可视化.



齐越(1969 -),男,博士,副教授,CCF 高级会员,主要研究领域为虚拟现实.



赵沁平(1948 -),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为虚拟现实,分布式应用系统.