Local Thin Plate Spline Descriptor for Aiding Marker-less Cerebral Artery Navigation Based on 3D Point Cloud Registration

Qiuying Wang, Chang Liu*, and Hao Tang

***Abstract*—Three-dimensional feature descriptors are essential for extracting point correspondences in point cloud data, crucial for 3D computer vision. However, the unordered, incomplete, and degraded nature of point clouds challenges existing methods, leading to weak robustness, poor descriptiveness, and high computational complexity. This paper introduces a novel local surface descriptor thin plate spline feature (TPSF) using thin plate spline for medical navigation scenarios that require high precision tracking of patient faces. The method converts local 3D features from MRI and depth cameras into TPS function coefficients, capturing detailed face surfaces. Similarity between MRI models and depth camera scenes is measured by mean square error (MSE). TPS coefficients from MRI keypoints reconstruct height values of scene keypoints, and similarity is defined as MSE between reconstructed and true heights. To avoid local optimal matching, a novel mismatch filtering algorithm is proposed, using the fully connected conditional random field (CRF) for optimal feature matching and a distance adjacency matrix for mismatch rejection. Experiments on several public datasets with increasing gaussian, shot noise and varying resolutions show TPSF's robustness, descriptiveness, and compactness. Validation on BIWI dataset and private datasets with RealSense depth camera demonstrates its potential in 3D visual applications with flat, feature-scarce environments.**

***Index Terms*—Point cloud registration, Local surface feature, 3D surface description, Feature transformation, Medical navigation**

I. INTRODUCTION

…In short, we give consideration to the deformation of the data caused by different imaging modes in the patient space and the model space, as well as the smooth and flat characteristics of the human surface. In this paper, we present a novel local invariant feature-based methods approach for point cloud registration. The descriptor we designed could describe fine surfaces, resists noise occlusion and clutter, and is compact. Specifically, they are a set of coefficients compute from thin plate spline function. We perform descriptor calculations on the grid sampling points on the point cloud converted from the MRI images, as they effectively reflect the approximate structure of the face. Despite its compactness, the matching strategies we introduce is effective at finding optimal matches. In fact, the correspondence established by the previous methods only based on the similarity between descriptors is unreliable. Different from other false correspondence removal algorithm, we need to avoid the wrong correspondence and form the correspondences establishment as a set of pure strategies in the process of feature matching. This feature matching algorithm for point cloud registration utilizes a fully connected Conditional Random Field (CRF) to optimize the correspondence between scene and model keypoints. By incorporating additional high-penalty states in the fitness and state matrices, the algorithm effectively handles scenarios with no suitable matches, defaulting to a non-match state when necessary. This approach ensures robust and accurate matching by considering both residuals and geometric constraints. Finally, the coarse registration is refined with the ICP algorithm. In this paper, the major contributions are summarized as follows:

1. A robust local reference frame (LRF) with scaling strategy.
2. A local invariant point cloud descriptor is proposed based on thin plate spline.
3. A feature matching algorithm is improved based on fully connected CRF with geometric constraints (GC).
4. The proposed descriptor and registration algorithm are applied to stimulation and real-world scenarios captured by depth camera and MRI images.

The structure of this paper is as follows. Section 2 presents a brief review of related work on descriptor-based point cloud registration algorithms. Section 3,4,5 delineates our proposed method for point cloud local surface representation using TPSF and the improved feature matching algorithm. The experimental evaluation results are analysed in Section 6. Section 7 applies the point cloud registration algorithm based on TPSF features to the problem of 3D medical image registration and tracking the scene video to validate its effectiveness. Section 8 concludes the works.

II. RELATED WORK

A. 3D Local Descriptor

B. Feature Matching

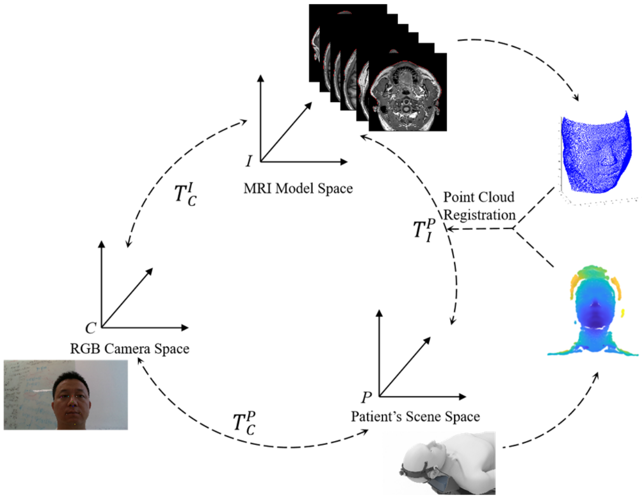

Fig. 1. Motion representation of cerebral artery navigation system.

III. OVERVIEW

A. Motion Representation

As shown in Fig.1, our designed cerebral artery navigation system is formed by three subsystems, namely, RGB camera space(O), MRI model space(I) and patient's scene space(P). MRI model space (I) for processing and visualization of patient MRI data and 3D models. We adopt a deformable-curve-based edge tracking algorithm, which can automatically track the contours of the patient's face in MRI sequences and reconstruct the 3D face model M. In RGB camera space(O), we extracted 10 remarkable points such as the corners of the eyes, corners of the mouth, nose tip and chin in the first RGB frame of the tracking video as biomarkers. At the same time, in patient's scene space(P), the relative point cloud representation P_t is introduced based on the depth camera. We adopt Lie group and Lie algebra in SO(3) to formulate rigid transformation in homogeneous coordinate system since it can represent the motion without ambiguity. The transformation relationship between coordinate systems of each space could be solving as follow:

1. ${}_P^I T$ is the homogeneous transformation matrix of point cloud from space I to P, which could be translated into solving the point cloud registration problem.

$$ {}_P^I T = \begin{bmatrix} R_{3\times3} & t_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix}, \tag{1} $$

where, R and t represent the rotation matrix and the translation vector of I with respect to P.

2. ${}_P^O T$ represents the homogeneous transformation matrix of O with respect to P from extrinsic matrix. Given the RGB camera intrinsic matrix K_c, the biomarkers on RGB images could be mapped to the scene space.

3. ${}_O^I T$ is obtained through the following relationship,

$$ {}_O^I T = \left({}_P^O T\right)^{-1} {}_P^I T. \tag{2} $$

Therefore, the navigation system's algorithm must address the registration problem between the dense target model and the sparse scene, while also accounting for the significant geometric discrepancies.

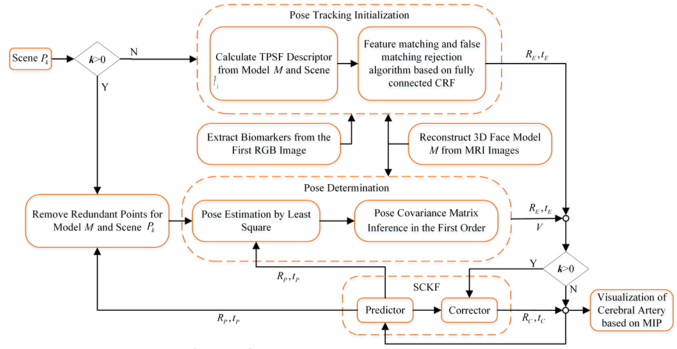

Fig. 2. The workflow of the point cloud registration and pose tracking system.

B. Algorithm Overview

The proposed algorithm registers point clouds from space I and P by presenting a set of TPSF descriptor matching strategies based on fully connected CRFs, and tracks the 3D face pose in real time. The workflow of the point cloud registration and pose tracking system is depicted in Fig.2, where k denotes the time frame. P_k is the scene point cloud taken by a depth camera at time frame of k; M is the model point cloud reconstructed from MRI sequences using deformable-curve-based edge tracking algorithm (beyond the scope of this article). At the beginning of tracking ($k=0$), the face pose tracking system is initialized by calculating TPSF descriptors (see section IV) for the keypoints of model M and scene $P_{k=0}$, and solving a dense CRFs (see section V) for feature matching. We estimate the initial pose of the model R_E, t_E according to the correspondence of key points. At the time frame of $k>0$, we utilize R_P, t_P (provided by predictor at $\mathrm{k}-1$) to remove redundant points for model M and scene P_t. Next, we optimize the same pose R_P, t_P in real time by minimizing the distance between P_k and transformed M with ICP. Then, the covariance of the pose is estimated via the first order optimality condition of distance minimization (see section VI). Based on the covariance, the SCKF generates the final pose estimate R_C, t_C and predicts the R_P, t_P for the next time frame. In the following sections, we will detail our work as outlined in the workflow.

IV. TPS-BASED LOCAL SURFACE DESCRIPTOR

A. Key Point Extraction

B. LRF Construction

C. Feature Point Descriptor

V. FEATURE MATCHING BASED ON FULLY CONNECTED CRF

A. The Similarity of the Corresponding Keypoints

B. Feature Matching Method

VI. EXPERIMENTS AND RESULTS

A. TPSF Descriptor Evaluation

B. Robust Head Pose Estimation

C. Face Tracking on BIWI Dataset

D. Application in Realsense

VII. CONCLUSION

ACKNOWLEDGMENT

REFERENCES