Fourier Analysis Techniques Applied in Data Registration: A Survey

Bo Sun, Weiwei Kong, Liwei Zhang and Jianwei Zhang

Abstract— **The data registration transforms multiple datasets into the same coordinate system, and in this way overlapping components of these sets are aligned. The registration is the fundamental of many advanced tasks, such as data fusion. There are plenty of solutions to data registration, and among them the algorithms based on Fourier analysis are efficient for large offsets but intricate. The surveys about data registration seldom focus on the approaches based on Fourier analysis. In this survey, we express the profound principles of Fourier analysis behind the obscure mathematical equations in plain language, explain the reason why the Fourier analysis techniques could be used in data registration and present how it is used in existing literatures. Transforming datasets into the Fourier domain decouples the rotation from the translation. For the rotation estimation, the algorithms could be categorized into the ones based on re-sampling of the Cartesian Fourier spectral and the ones adopting Spherical Fourier Transform. As to the translation recovery, the Phase Only Matched Filtering (POMF) is widely used and scarcely has a rival. In the end, we discuss the challenge of the registration methods relying on Fourier analysis and point out the future direction.**

I. INTRODUCTION

Data registration is the process of spatially aligning two or more datasets of an object or scene. This basic capability is one of the fundamental tasks within many advanced data processing techniques. The alignment process could determine the correspondences between points in the datasets, enable the fusion of information and estimate the motion of sensors. Besides, if identities of objects in one of the datasets are available, identities of objects and their locations in another dataset can be determined by registering the two datasets.

The task of data registration is to find an optimal geometric transformation between corresponding datasets. The data might be 2D images, 3D point clouds (textured or not) or 3D volumes and so on, which are taken at different times, using various sensors, or from diverse viewpoints. As to the 2D rigid case, the optimal geometric transformation is represented by three parameters: one for 2D rotation and two for 2D translation. For the 3D rigid case, the optimal geometric transformation has 6 Degree of Freedom (DOF): three for 3D rotation and three for 3D translation. So the data registration could be interpreted as a box taken two or more datasets as inputs and output three or six parameters describing the geometric transformation between the datasets[1].

B. Sun, W. Kong and J. Zhang are with TAMS, the Department of Informatics, University of Hamburg, Hamburg, Germany.

L. Zhang is with TAMS, the Department of Informatics, University of Hamburg, Hamburg, Germany; Shenzhen Institute of Advanced Integration Technology, CAS, Shenzhen, China; The Chinese University of Hong Kong, Hong Kong, China

[1]This paper focuses on the rigid registration, and the deformable object registration is not considered here.

The data registration is an important and classical problem, and there are plenty of solutions in existing literatures. The algorithms could be classified into *spatial domain methods* and *frequency domain methods* according to whether the registration is operated in image domain or in frequency domain.

The spatial domain methods could be categorized into *local methods* and *global methods*. The local methods usually depend on an iterative procedure aiming to find the optimal transformation parameters which maximize the similarity measures between the input datasets. The local methods require good initial guesses and are easy to get trapped in local minima due to the iterative local optimization procedure. Further, the runtime of local methods would vary considerably even for the input datasets with the same size. Considering the shortcomings of local methods, the global registration methods take the overall appearance of input data into account. The majority of global methods are feature-based techniques, which make use of the explicit feature correspondences in the datasets. The common flowchart of feature-based registration methods includes: key-point extraction, feature description, feature matching, transformation recovery and refinement. The feature-based strategies could deal with the input data with partial overlaps and large offsets, but confront challenges about how to eliminate the mismatches, especially in the large scale scenes contain numerous similar patches. Besides, the feature-based registration methods rely on the proper crucial parameters heavily, so specific parameters should be set up for different datasets carefully which reduces the intelligence of the algorithms.

Compared with the spatial domain registration methods, the frequency domain methods attract less attention so far but are potentially useful. In a sense, the registration algorithms in frequency domain could be regarded as the global feature-less methods. In global feature-less registration algorithms, the original input data is depicted by a global descriptor, and the transformation between the original data could be solved by aligning the corresponding global descriptors. The Fourier spectrum of the original data could be regarded as this kind of global descriptor.

The frequency based registration techniques have four major advantages:

- ***robustness to noise***: they use all the information available rather than just the information associated with sets of features, so they are more stable than the feature-based methods;
- ***dissociation of rotation and translation***: the separability of rotation and translation is intrinsic in the structure of the Fourier representation of signals, which reduces

the amount of calculation dramatically;

- ***insensitivity to parameters***: compared to the feature-based methods, they are less sensitive to the specific crucial parameters. In other words, the parameters used in frequency based registration techniques just determine the runtime and precision, do not determine the registration succeed or fail.
- ***validity to partially overlapped data***: the phase correlation techniques, such as the Phase-Only Matched Filter (POMF), are adopted in frequency based registration algorithms to deal with the partially overlapped datasets.

The Fourier Transform allows the decoupling of the estimation of rotation parameters from the determination of translation parameters. For the rotation estimation, the Spherical Harmonics based algorithms could give a precise analytical solution but are computational expensive, while the algorithms based on the re-sample structure of the Cartesian Fourier Transform need less computation but do not lead to an exact solution and often suffer from the sensitive interpolation procedure. Considering the re-sampling of Fourier frequency values is quite sensitive to interpolation, the Pseudo-polar Fourier Transform (PPFT) based on Fractional Fourier Transform (FrFT) is introduced to Cartesian Fourier Transform based registration algorithms. Both of two categories of methods are investigated detailedly in this survey. For the translation recovery, the Phase Only Matched Filtering (POMF) is widely accepted. The POMF decouples the local signal energy from the signal structures based on the fact that two shifted signals carry the shift information within the phase of their Fourier spectrum.

Among all the existing surveys [1][2][3][4][5][6][7], there is no one focusing on registration methods based on Fourier analysis. This survey tries to express the principles of the Fourier Transform in the plain language, explain the reasons why the Fourier analysis techniques could be adopted for the registration purpose, and demonstrate how they are used in existing literatures.

This paper is organized as follows: Section II describes the expression of normal Fourier Transform, its principles and philosophy; explains why it could decompose the rotation and translation of datasets. The rotation estimation techniques based on Cartesian Fourier spectral are investigated in section III. Section IV presents the basis theory of Spherical Fourier transform and the Generalized Convolution Theorem, which explain why the Spherical Harmonics could be used to estimate the rotation parameters. The rotation determination techniques based on the Spherical Harmonics are analysed in section V. The translation recovery based on phase-only correlation is discussed in section VI. And finally section VII concludes the challenges and points out the future direction of registration approaches based on Fourier spectrum.

II. PRELIMINARY OF FOURIER ANALYSIS

A. Basic Principles of Fourier Transform

Fourier analysis techniques are extremely significant in signal processing and pattern recognition, since they decompose the function into a linear combination of sinusoidal basis functions. Each of these basis functions is a complex exponential of a different frequency. In other words, the Fourier analysis techniques map a function into a set of coefficients of basis functions, and the coefficient of the basis function with frequency f gives how much power the function contains at the frequency f. That is, the Fourier analysis techniques give us another way to represent a waveform. Admittedly, there are infinite ways to decompose the signals. But the goal of decomposition is to get something easier to deal with than the original signals, and the reason why sinusoids are adopted is that they are the eigenfunctions of the Laplacian operator, hence they maintain fidelity to most real systems. The basis functions of normal Fourier analysis techniques are induced by the Laplacian operator in Cartesian coordinate system. By the same token, the Laplacian operator also has effective forms in other coordinate systems, e.g. polar and spherical coordinate system. The Polar/Spherical Fourier analysis techniques are connected with Cartesian Fourier analysis techniques by the ***Laplacian operator***.

There are four types of Fourier analysis techniques: Fourier Series (FS), Fourier Transform (FT), Discrete-time Fourier Transform (DTFT) and Discrete Fourier Transform (DFT). The FS breaks down a periodic continuous function into the sum of infinite sinusoidal functions. The FT extend the idea of FS to continuous aperiodic functions. The DTFT is the spectral representation for aperiodic discrete signals, and normally the discrete inputs are acquired by digitally sampling the continuous function. It is interesting that the DTFT frequency representation is always a periodic function. So though the result of DTFT is an infinite summation, sometimes it is convenient to regard the DTFT as a transform to a 'finite' frequency representation (the length of one period). The DFT converts a finite list of equally spaced samples of a function into the list of coefficients of a finite combination of complex sinusoids, ordered by their frequencies. Since digital computer can only work with discrete and finite signals, the only type of Fourier analysis technique could be used in computer software is DFT. Please note that the Fourier Transform is usually used as the generic term of the Fourier analysis techniques in the literatures.

The sequence of N complex numbers $x_0, x_1, \ldots, x_{N-1}$ is transformed into an N-periodic sequence of complex numbers $X_0, X_1, \ldots, X_{N-1}$ according to the DFT formula:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi kn/N} \qquad (1)$$

Equation (1) could be interpreted as the cross correlation of the input sequence, x_n, and a complex exponential at frequency k/N. Thus it acts like a *matched filter*, and X_k is the Fourier coefficient with that frequency, which represents how much power contained in the original sequence at that frequency. Based on the coefficients, the original complex data x_n could be expressed as:

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot e^{i2\pi kn/N} \quad (2)$$

The kernel of Fourier analysis technique is the representation of data in another linear space. The Equation (2) could be regarded as the representation of the original sequence in frequency domain.

B. Decoupling of Rotation and Translation

In the data registration application, the *Fourier Rotation Theorem* and *Fourier Shift Theorem* are used to decouple the rotation and translation information between two datasets. The Fourier Shift Theorem indicates that a shift in position in one domain gives rise to a phase change in another domain. The Fourier Rotation Theorem says that a rotation of a dataset by an angle θ implies that its Fourier spectrum is also rotated by the same angle along the same direction. These two theorems are the theoretical bases explaining why the Cartesian Fourier transform could be used in data registration.

Let $v_1(x)$ and $v_2(x)$ be two datasets which are related through a rigid roto-translational motion: $v_2(x) = v_1(R^{-1}x - t)$; according to which, $v_2(x)$ is obtained from $v_1(x)$ by first translating each point of it by the vector t and then rotating the result of this operation by the rotation matrix R. Denote the corresponding Fourier Transform of $v_1(x)$ and $v_2(x)$ by $V_1(s)$ and $V_2(s)$. According to the *Fourier Rotation Theorem* and *Fourier Shift Theorem*, the relationship between $V_1(s)$ and $V_2(s)$ is:

$$V_2(s) = V_1(R^{-1}s)e^{-i2\pi(s^T Rt)} \quad (3)$$

Equation (3) could be simplified in terms of magnitudes:

$$|V_2(s)| = |V_1(R^{-1}s)| \quad (4)$$

From equation (3) and (4), it can be obtained that, in the frequency domain, the estimation of rotation matrix R could be decoupled from the recovery of translation t. Furthermore, we can notice that the phase information in equation (4) is related to both the rotation matrix R and the translation vector t, which means we have to determine the rotation matrix firstly based on equation (4). After that, rerotate the datasets according to the obtained rotation matrix, in this way there is the pure translation offset left. Then the Fourier Shift Theorem could be adopted to calculate the translation parameters. And this is the standard flowchart of the registration techniques based on Fourier analysis.

III. ROTATION RECOVERY BASED ON CARTESIAN FOURIER SPECTRUM

According to the Fourier Rotation Theorem, the estimation of rotation matrix between two datasets could be converted to the determination of the rotation information between the magnitudes of the corresponding Fourier spectra. With regard to the rotation recovery algorithms in existing literatures, they could be classified into two main categories: the methods based on the re-sampling of the Cartesian Fourier spectrum and the methods based on the Spherical Harmonics.

The early try of registration of translated and rotated images using Fourier analysis techniques is presented in [8]. The idea of algorithm in [8] is quite naive. Since there is only phase drift between the Fourier spectra of datasets if they are pure translational duplicate version to each other, the authors define a function that is the quotient of the Fourier spectra and try all tentative values of rotation angle. The true rotation angle approaches the quotient to the exponential form. In other words, the fully automatic directed search strategy is used to determinate the rotation angle. Obviously this method is computational expensive. Further, the authors of [9] point out that the method rests on the observation that if the rotation angle is rather small (not exceeding ± 2 degree), the peak of the quotient may be still spotted out although considerably lower than that associated with a pure translation. This represents a serious problem and makes such technique suitable only for applications where rotations are small and certainly not for general purpose utilizations. Besides, because the rotation angle is applied to magnitude of Fourier spectra not the original data, and the phase drift is related to both translation and rotation, the phase drift caused by rotation could not be compensated, which also reduces its robustness.

The more official algorithms of data registration based on Fourier spectrum are proposed in [10][11]. The essential ideas in these two paper are quite similar. They adopt the well known invariant function named *Fourier-Mellin transform*, which is translation invariant and represents rotation and scale as translations along the corresponding axes in homologous parameter spaces. The Fourier-Mellin Invariant (FMI) descriptor of an image, named Fourier log-magnitude spectra in [11], is obtained by re-sampling the spectral magnitude of this image to polar coordinates and then re-sampling along radial coordinate with a logarithmic function. And the FMI descriptor could also be achieved by re-sampling the spectral magnitude directly onto a rectangular polar-logarithmic coordinate in one step, like what is done in [10][11]. In the polar-logarithmic representation of the spectral magnitude, both the rotation and scale are transformed to translation. Actually the FMI descriptor is not firstly proposed in [10], but all previous methods match the FMI descriptors using cross-correlation, or variants of cross-correlation. Since the FMI descriptor is based on the magnitude of Fourier transform, the cross-correlation of the FMI descriptors generally yields a very broad maximum, which leads to the FMI descriptor based registration methods are unreliable and may give rise to wrong estimates of correlation peaks. The Symmetric Phase-Only Matched Filtering (SPOMF) is introduced to match the FMI descriptors in [10], and this is the main difference from the algorithm in [11]. In this way, the advantages of the SPOMF, which are sharpness of the correlation peaks and robustness in the presence of noise, and the decoupling the rotation, scale and translation acquired by the FMI descriptor are combined together. The problem here is that the limited scale range can be estimated since large scales would alter the frequency beyond recognition. It should be noted that the maximum scale recovered by [10] is 2.0 and the maximum scale recovered by [11] is 1.8. Besides, since the re-sampling of Cartesian frequency values on a polar grid is very sensitive to interpolation, the accuracy of registration algorithms is severely degraded by the approximation errors inherited in the computation of the polar and log-polar Fourier spectra.

A novel frequency domain technique which works in Cartesian coordinates and bypasses the need to transform data from the Cartesian to the polar domain is presented in [9][12][13]. It is an important advance because it is well known that the Cartesian-to-polar coordinate transformation is a numerically sensitive operation, especially when it is dependent on the interpolation of the Fourier spectrum. The fundamental idea of this novel method rests on the property of the Fourier transform magnitudes of the images: For two roto-translated images, the difference between the Fourier transform magnitude of one image and the mirrored replica of the Fourier transform magnitude of the other have a pair of orthogonal zero-crossing lines. These two lines are rotated with respect to the frequency axes by an angle that is half the rotational angle. In [9], the difference function of normalized Fourier transform magnitudes is defined, and the zero-crossing lines of the difference function is used to determine the rotation angle. Therefore, the estimate of the rotation angle is transformed to the detection of two zero-crossing lines. Subsequently, the phase correlation technique is applied to solve the translation parameters.

The research direction of determining rotation angle based on the difference function of normalized Fourier transform magnitudes is extended to the case of 3D rigid motion in [14][15]. This kind of 3D data registration method contains three procedures: rotation axis determination, estimate of the rotation angle around the axis and translation calculation. The rotation axis is determined in [15] by searching a minimum of radial projection of the polar re-sampled differences of the magnitude of Fourier spectra. After obtaining the rotation axis, the rotation angle estimation is a planar rotation problem, which is solved by a combination of the a 1D Fourier transform and a 2D polar Fourier transform defined as a cylindrical Fourier transform. This algorithm involves the interpolation operation in the re-sampling of the difference function in rotation axis estimation step, which increase the uncertainty of the registration result. Another main drawback is that this method requires the common region between two input data to be known, which means this method could not handle with the general partial overlapped datasets and constraints its application.

The methods presented in [16][17] inherit the three-step framework, but apply the 3D Pseudo-polar Fast Fourier Transform (FFT) to avoid the interpolation operation. The Pseudo-polar FFT is firstly proposed in [18], and it could be used to compute the Discrete Fourier Transform (DFT) on pseudo-polar grids without interpolation of Cartesian Fourier spectra. The Pseudo-polar FFT is based on the Fractional Fourier Transform (FrFT) [19] [20], which is a generalization of the conventional Fourier transform. The FrFT depends on a parameters α and can be interpreted as a rotation by an angle α in the time-frequency plane with respect to the conventional Fourier transform. An FrFT with $\alpha = \pi/2$ corresponds to the conventional Fourier transform. Essentially, the α-order FrFT shares the same eigenfunctions as the conventional Fourier transform, but its eigenvalues are the αth power of the eigenvalues of the conventional Fourier transform. Concretely, the FrFT samples the spectrum of a vector with length N at the frequencies: $\omega_k = \alpha l/N, l = -N/2, \ldots, N/2$. In other words, the FrFT could compute the Fourier spectra with arbitrary frequency resolution. Furthermore, the FrFT of a vector of length N can be computed in $O(NlogN)$ operations for any α.

The 3D pseudopolar grid is given by three sets of samples:

$$\begin{aligned} P_1 &\triangleq \left\{ \left(k, -\frac{2l}{N}k, -\frac{2j}{N}k \right) \right\} \\ P_2 &\triangleq \left\{ \left(-\frac{2l}{N}k, k, -\frac{2j}{N}k \right) \right\} \\ P_3 &\triangleq \left\{ \left(-\frac{2l}{N}k, -\frac{2j}{N}k, k \right) \right\} \end{aligned} \tag{5}$$

where $l, j \in \{-\frac{N}{2}, \ldots, \frac{N}{2}\}$, $k \in \{-\frac{3N}{2}, \ldots, \frac{3N}{2}\}$. The pseudo-polar Fourier transform is defined by sampling the Cartesian Fourier transform on the pseudo-polar grids. However, instead of the interpolation of Cartesian Fourier spectra, the pseudo-polar Fourier transform could be calculated by FrFT directly. For fixed l, j, the samples of the 3D pseudo-polar grid are equally spaced in the radial direction. But this space is different for different l, j. And the grid has equally spaced slopes. The detailed calculation of 3D Pseudo-polar FFT could be found in [18].

Three important properties of the 3D Pseudo-polar Fourier transform fit it to be applied in data registration:

1) invertible, and both the forward and inverse transform could be implemented using fast algorithms;
2) do not require re-gridding or interpolation;
3) the pseudo-polar grids could be used as an approximation to the polar/spherical grids.

The estimation of the rotation axis is shown to be algebraically accurate by using the 3D pseudo-polar FFT in [16][17]. But the determination of the rotation axis is still an error-prone part in case that the scan data has interference and occlusion. The pseudo-polar FFT is also applied in 2D image registration in [21].

The most recent registration method based on Cartesian Fourier spectra is presented in [22][23], named Spectral Registration with Multilayer Resampling (SRMR). The SRMR re-samples the spectral magnitude of 3D FFT calculated on discrete Cartesian grids of the 3D data to decouple the 3D rotation and 3D translation, just like the previous techniques. Further, the SRMR also tries to transform the rotation parameter to the translation estimation problem and adopt the phase correlation techniques to figure out it, and this main idea is also inherited from the previous techniques. The most remarkable feature of SRMR is that it uses the spectral structure at a complete stack of layers instead of only one spherical layer. And it does not rely on finding minima indicating the main rotation axes, which makes it extremely robust for the partial overlapped datasets. But the strong point is bought at the cost of working only in a limited range of roll and pitch offsets between input datasets. So the SRMR is only applicable in robotic mapping scenarios, where there is little roll and pitch changes. Furthermore, please note that the SRMR also suffers from the sensitive interpolation of spectral magnitudes.

IV. Basic Theory About Spherical Harmonics

A. Spherical Fourier Transform

The angular part of spherical Laplacian operator's eigenfunctions are named *spherical harmonic* functions $Y_l^m : S^2 \mapsto \mathbb{C}$, where S^2 stands for the unit 2D sphere and $\mathbb{C}$ symbolizes the set of complex number.

$$Y_l^m(\vartheta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(cos\vartheta)e^{im\varphi} \quad (6)$$

where (ϑ, φ) are the spherical coordinates; P_l^m is an associated Legendre polynomial; l, m are integers, $l > 0, |m| < l$. The l is called the degree of spherical harmonics. For the l-th degree, there are $2l + 1$ spherical harmonics basis functions indexed in the range of $-l \leq m \leq l$. Similar to the Cartesian Fourier basis, spherical harmonics represent the different frequency components of spherical functions.

Let $L^2(S^2)$ denote the space of square integrate functions defined on S^2. The spherical harmonics of degree l constitute a $(2l + 1)$ dimensional subspace of $L^2(S^2)$. Spherical harmonics of different degrees are orthogonal to each other. Furthermore, the spherical harmonic functions provide an complete orthonormal basis for $L^2(S^2)$. In other words, any function $f(\vartheta, \varphi) \in L^2(S^2)$ could be expanded as a linear combination of spherical harmonics:

$$f(\vartheta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_l^m Y_l^m(\vartheta,\varphi) \quad (7)$$

$$f_l^m = \int_{S^2} f(\Omega)\overline{Y_l^m(\Omega)}d\Omega \quad (8)$$

where the overline stands for the complex conjugate. Equation (7) is named Spherical Fourier expansion or inverse Spherical Fourier Transform, and f_l^m is commonly called the spherical harmonic coefficients of $f(\vartheta, \varphi)$.

It is obvious that *the Spherical Fourier Transform has similar formula with the Cartesian Fourier Transform*. Moreover, the Spherical Fourier Transform is rotation friendly and plays a vital role in matching patterns on S^2.

The well known property of the Cartesian Fourier transform is that the magnitudes of the Fourier spectra are invariant under translation, since the translation in the signal manifests itself as a phase drift in the Fourier spectra. As to the Fourier transform on the sphere, the action is rotation instead of the translation. The magnitude of the Spherical Fourier spectra is rotational invariant. So the Spherical Harmonics are widely used to constitute the rotation invariant descriptor [24][25][26] and 3D shape retrieval [27].

B. Generalized Convolution Theorem

It is well known that it is possible to convert the rotation of data to translation, and the phase correlation techniques could be used to calculate the translation parameters. And this correlation could be converted to point-wise product via Fourier Transform. In other words, the correlation in time domain equals point-wise multiplication in frequency domain, then the original problem could be solved in frequency domain much more efficiently. Furthermore, this convolution theorem could be generalized to the functions defined on S^2. *Efficient spherical convolution, aided by a fast Spherical Fourier transform and its inverse, contributes to the registration of graphs on* S^2.

Following the tradition, let $SO(3)$ denote the rotation group in 3D space, represented by the 3×3 matrices with determinant one. Given a function f_1 on the sphere, and its rotated replica f_2 for a rotation $g \in SO(3)$: $f_2 = \wedge(g) \cdot f_1$. Registration of the two functions could be achieved by correlating functions:

$$C(g) = \int_{S^2} f_2(\Omega) \bullet \overline{\wedge(g) \cdot f_1(\Omega)} d\Omega \quad (9)$$

and the g maximizing the integral (9) is the rotation between two functions. However, evaluating $C(g)$ for all possible rotations is a terrific time-consuming task.

Spherical Fourier Transform could be adopted to determine the maximum g efficiently based on rotation invariant characteristic of spherical harmonic:

$$\wedge(g)Y_l^m(\Omega) = \sum_{|k|\leq l} D_{km}^l(g)Y_l^k(\Omega) \quad (10)$$

where $D_{km}^l(g)$ are called *Wigner-D function*, and they are the irreducible unitary representations of $SO(3)$. In some sense, the $D_{km}^l(g)$ could be interpreted as the k-th component of $\wedge(g)$ acting on $Y_l^m(\Omega)$. Formula (10) signifies that *the rotated cousins of a spherical harmonic could be expressed as a linear combination of spherical harmonics with the same degree*. According to equation (7), the Spherical Fourier expansions of f_1 and f_2 are:

$$f_1(\Omega) = \sum_{l}\sum_{|m|\leq l} a_l^m Y_l^m(\Omega) \quad (11)$$

$$f_2(\Omega) = \sum_{l'}\sum_{|m'|\leq l'} b_{l'}^{m'} Y_{l'}^{m'}(\Omega) \quad (12)$$

Substituting equations (11) and (12) into equation (9), then utilizing the *"Separation of Variables"* technique and the orthogonality between spherical harmonics, the correlation function could be rewritten as:

$$C(g) = \sum_{l}\sum_{|m|\leq l}\sum_{|m'|\leq l} a_l^m \overline{b_l^{m'}} \bullet \int_{S^2} Y_l^{m'}(\Omega)\, \overline{\wedge(g)Y_l^m(\Omega)} d\Omega \quad (13)$$

Recall the rotation invariant property of the spherical harmonic expressed by equation (10), $\overline{\wedge(g)Y_l^m(\Omega)}$ could be replaced:

$$\begin{aligned} C(g) =& \sum_l \sum_{|m|\leq l} \sum_{|m'|\leq l} a_l^m \overline{b_l^{m'}} \bullet \\ & \int_{S^2} Y_l^{m'}(\Omega) \overline{\sum_{|k|\leq l} D_{km}^l(g) Y_l^k(\Omega)} d\Omega \\ =& \sum_l \sum_{|m|\leq l} \sum_{|m'|\leq l} a_l^m \overline{b_l^{m'}} \bullet \\ & \overline{\sum_{|k|\leq l} D_{km}^l(g)} \int_{S^2} Y_l^{m'}(\Omega)\ \overline{Y_l^k(\Omega)} d\Omega \end{aligned} \tag{14}$$

All over again, based on the orthogonality of spherical harmonics, the integral and the summation on k in equation (14) could be zapped:

$$C(g) = \sum_l \sum_{|m|\leq l} \sum_{|m'|\leq l} a_l^m \overline{b_l^{m'}} \cdot \overline{D_{m'm}^l(g)} \tag{15}$$

The detailed deduction could be found in [28][29]. It is mainly on the strength of the *"Separation of Variables"* technique and orthogonality of spherical harmonics, and makes full use of the two criteria over and over again. As prescribed in equation (15), the convolution in equation (9) is converted to point-wise multiplication, and the correlation $C(g)$ concerning the whole series of g could be evaluated uniformly and efficiently based on the spherical Fourier coefficients. In this way, the rotation g maximizing $C(g)$ could be easily found.

V. ROTATION RECOVERY BASED ON SPHERICAL HARMONICS

The data registration methods based on the Spherical Harmonics attracts scarce attention in the image analysis society. An example of pattern matching on S^2 is introduced in [29] when an efficient algorithm for the numerical calculation of the Spherical Fourier transform of functions defined on the rotation group is proposed.

The Spherical Harmonics and Generalized Shift Transform is officially adopted to recover the rotation of spherical images in [30][31][32]. The images arising from a omnidirectional camera could be mapped onto the sphere, and the underlying mappings of the sphere can reflect a rotational camera motion. The problem of rotation estimation directly from the images defined on the sphere could be addressed based on the Spherical Harmonics without correspondences. The resolution of the rotation space depends on the bandwidth of the harmonic expansion. The experiments show that the methods are suitable for large rotations, and are quite resistant to small translations of a camera. Moreover, it is attractive that [32] presents a novel decoupling of the shift theorem with respect to the Euler angles and exploits it in an iterative scheme to refine the initial rotation estimates.

In order to determine the rotation matrix between 3D data, the first problem is how to describe the structure of the 3D data using a spherical image. The registration method in [33] explores the global representation of range scans named Extended Gaussian Image (EGI), which is built by a spherical histogram of surface orientations. This technique is based on the correlation of the EGIs in the Spherical Fourier domain and makes use of the Generalized Shift Transform to compute the rotation between the EGIs, which is also the rotation between the original scans. The EGI is translational invariant but it has three shortcomings: 1) EGIs only uniquely define convex objects; 2) EGI is only feasible to smooth surface, but causes problems at discontinuities; 3) it fails to deal with the scans contain spherical objects, which would lead to constant histograms and less informative EGIs.

We developed another 3D scan registration technique based on Spherical Harmonic. We propose a new structure representation of scans, named Spherical Entropy Image (SEI). SEI divides the 3D data into several patches according to the polar angle and azimuth angle of points. For each patch, considering the depth of points as the observations of a random variable, build the histogram and compute the entropy of the variable. The SEI is achieved by computing the entropy of patches in a dense manner. The SEIs maintain the visual intuition of the original point clouds. But the SEI is not translation invariant due to the fact that the way to divide scans is translation dependent, so the translation normalization of the scans is essential before calculating SEI. Further, the SEI could be translation invariant if the SEI is computed on the magnitude of the scans, but in this way, the SEI becomes more computational expensive.

VI. TRANSLATION DETERMINATION

Supposing the determined rotation R is correct, there is only translation offset between the datasets after applying the determined R to them.

Due to its ability to deal with the partially overlapped signals, the cross-correlation technique is suitable to determine translation parameters. For all the registration techniques based on the Fourier analysis in the existing literatures, the cross-correlation technique is adopted to recovery the translation information without any exception. However, the value of the standard cross-correlation method is heavily dependent on the energy of underlying signals rather than on the spatial structures, so it often fails to discriminate the signals which are of different shapes but similar energy. Furthermore, the correlation peaks could be relatively broad depending on the signal structures, which makes it difficult and unreliable to locate the correct displacement between noisy signals. In order to decrease the impact of disproportionally large value points which are not present in the overlapped area and achieve more distinct sharp peaks, the POMF algorithm is used to resolve the translation recovery problem by almost all the registration methods based on Fourier analysis techniques.

The POMF decouples the local signal energy from the signal structures based on the fact that two shifted signals carry the shift information within the phase of their Fourier spectrum. Let $f_1(x, y, z)$ and $f_2(x, y, z)$ be two shifted signals, and $\mathcal{F}_1(u,v,k)$ and $\mathcal{F}_2(u,v,k)$ be their corresponding Fourier spectra. The shift between these two translated signals could be solved by the following equations:

$$S(u,v,k) = \frac{\overline{\mathcal{F}_1(u,v,k)}}{|\mathcal{F}_1(u,v,k)|} \bullet \frac{\mathcal{F}_2(u,v,k)}{|\mathcal{F}_2(u,v,k)|} \quad (16)$$

$$s(x,y,z) = \mathcal{F}^{-1}\{S(u,v,k)\} \quad (17)$$

$$(x_p,y_p,z_p) = \arg\max_{(x,y,z)} s(x,y,z) \quad (18)$$

where the overline indicates the complex conjugate, and (x_p,y_p,z_p) is the displacement between the two signals. In theory, it could be used in arbitrary dimensional signal registration problems. Ideally, the $s(x,y,z)$ contains a Dirac peak, but the Dirac pulse deteriorates in practical due to the noise. Please note that the Fourier analysis technique used in POMF is the Cartesian Fourier transform.

The phase-only and amplitude-only matched filters are compared to the classical matched filter using the criteria of discrimination, correlation peak, and optical efficiency in [35], and the pure phase correlation filter outperform other filters remarkably. But the original phase correlation method is claimed to identify integer pixel displacements, which inspires the development of numerous subpixel alternatives. A low complexity subspace identification technique is presented in [36]. The horizontal and vertical components of the translation is separated by a Singular Value Decomposition (SVD) factorization of the phase correlation matrix. Then perform a linear fit to the phase of these separate components in order to identify the magnitude of the translation. This approach identifies the non-integer pixel displacement without interpolation, further this approach is robust to noise and needs limited computational complexity, which makes this method a quite attractive option for phase correlation problem. However, this approach relies on the quality of the linear fit rather than gives the analytic expressions. A closed-form solution to subpixel translation estimation is provided in [37]. This model is based on the philosophy that images with subpixel shifts were in fact originally displaced by integer values, and then the shifts are reduced to subpixel values due to down-sampling. Experiments confirm that high accuracy can be acquired using this approach, and it is available to perform accurate error analysis and evaluate the performance due to the existence of the analytic expression. [38] extends the phase correlation method to N-dimensional datasets. The motion model for translational offsets between N-dimensional images could be represented by a rank-one tensor based on the Fourier shift theorem. The adoption of a high-order SVD could decompose the phase correlation between two N-dimensional datasets to independently translational displacements along each dimension.

The phase correlation techniques are widely used to determine the translation parameters in data registration, and the POMF technique is the best choice without any rivals.

VII. CHALLENGES AND FUTURE DIRECTION

In general, the data registration approaches based on Fourier analysis techniques could handle with the large offsets but the registration results have lower precision. Hence the registration methods based on Fourier analysis techniques are usually used in the coarse registration step, whose results are refined by the local methods, such as Iterative Closest Point (ICP) and Normal Distributions Transform (NDT). The runtime of most local methods are initial estimates dependent, which means the more precise the initial guesses are, the less runtime the local methods require. So overall the usage of coarse registration method (e.g., the techniques based on the Fourier analysis) not only makes the registration automatically but also improves the efficiency.

But for now, there are plenty of challenges for the registration methods based on Fourier spectrum. The Fourier spectrum of 2D image and 3D volume could be calculated directly, but for the 3D points cloud, the original 3D surfaces are rasterized into volume grids. Generally the way to rasterize the 3D surface is assigning a voxel the value of 1 if it is occupied by the surface, otherwise its value is set to be 0. In this way, the final volumes only contain two values: 1 and 0, which makes the Fourier spectra of them less informative. One of the solutions could be estimating the Gaussian curvature of the 3D surface and assigning a voxel the value of Gaussian curvature on the point it contains. In such a way, the volume grids are more informative and the richly structured part of the surface plays a more important role in registration procedure.

With regard to the rotation recovery techniques based on the Cartesian Fourier spectrum, the most existing techniques suffer from the interference and occlusion. Though the recent multilayer method SRMR is quite robust to the partially overlapped dataset, it depends on the sensitive interpolation of spectral magnitudes. A novel multilayer re-sample strategy which uses the FrFT to avoid the interpolation is expected in the future; this kind of algorithm would not only be capable of dealing with the partially overlapped datasets but also be algebraically accurate.

As far as the rotation estimation approaches based on the Spherical Harmonics, the biggest challenge is the amount of calculation. Even a fast algorithm is proposed in [29], the registration algorithms involving Spherical Harmonics need more runtime than the ones based on Cartesian Fourier spectrum. But this disparity between them is shorten by the high-speed arithmetic hardware, and considering the robustness, theoretical completeness and algebraically accuracy, the methods based on Spherical Harmonics are vigorous in the future. Just like the multilayer strategy based on Cartesian Fourier spectrum, the multilayer strategy could also be adopted by the methods based on Spherical Harmonics. As we said, the Spherical Harmonic is the angular part of spherical Laplacian operator's eigenfunctions. The truth is that the the eigenfunction of spherical Laplacian operator contains both the radial and angular structures. So it is more natural for the Spherical Harmonics based registration methods to adopt the multilayer strategy. But unfortunately, there is no fast algorithm for the radial transform, and whether fast algorithms exist is still a question to be answered. That is the reason why the existing methods only consider the angular structure.

In a word, the registration techniques dependent on Fourier spectrum are immature but energetic, and they deserve more attention from image analysis society.

ACKNOWLEDGMENT

This work was supported in part by China Scholarship Council (CSC), the Natural Science Foundation of China (Project No.:61105130) and the project European and Chinese Platform for Robotics and Applications (ECROBOT).

REFERENCES

[1] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Medical Image Analysis*, vol. 4, no. 3, pp. 201–217, 2000.

[2] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, may 2001, pp. 145–152.

[3] M. Rodrigues, R. Fisher, and Y. Liu, "Special issue on registration and fusion of range images," *Comput. Vis. Image Underst.*, vol. 87, no. 1-3, pp. 1–7, July 2002.

[4] H. Pottmann, Q.-X. Huang, Y.-L. Yang, and S.-M. Hu, "Geometry and convergence analysis of algorithms for registration of 3D shapes," *International Journal of Computer Vision*, vol. 67, no. 3, pp. 277–296, 2006.

[5] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image Vision Comput.*, vol. 25, no. 5, pp. 578–596, May 2007.

[6] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A survey on shape correspondence," *Computer Graphics Forum*, vol. 30, no. 6, pp. 1681–1707, 2011.

[7] G. Tam, Z.-Q. Cheng, Y.-K. Lai, F. Langbein, Y. Liu, D. Marshall, R. Martin, X.-F. Sun, and P. Rosin, "Registration of 3D point clouds and meshes: A survey from rigid to nonrigid," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, no. 7, pp. 1199–1217, July 2013.

[8] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-9, no. 5, pp. 700–703, Sept 1987.

[9] L. Lucchese and G. Cortelazzo, "A noise-robust frequency domain technique for estimating planar roto-translations," *Signal Processing, IEEE Transactions on*, vol. 48, no. 6, pp. 1769–1786, Jun 2000.

[10] Q.-S. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 12, pp. 1156–1168, Dec 1994.

[11] B. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *Image Processing, IEEE Transactions on*, vol. 5, no. 8, pp. 1266–1271, Aug 1996.

[12] L. Lucchese, G. Cortelazzo, and C. Monti, "A frequency domain technique for estimating rigid planar rotations," in *Circuits and Systems, 1996. ISCAS '96., Connecting the World., 1996 IEEE International Symposium on*, vol. 2, May 1996, pp. 774–777.

[13] L. Lucchese and G. Cortelazzo, "Noise-robust estimation of planar roto-translations with high precision," in *Image Processing, 1997. Proceedings., International Conference on*, vol. 1, Oct 1997, pp. 699–702.

[14] L. Lucchese, G. Doretto, and G. Cortelazzo, "Frequency domain estimation of 3-D rigid motion based on range and intensity data," in *3-D Digital Imaging and Modeling, 1997. Proceedings., International Conference on Recent Advances in*, May 1997, pp. 107–112.

[15] ——, "A frequency domain technique for range data registration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 11, pp. 1468–1484, Nov 2002.

[16] Y. Keller, A. Averbuch, and Y. Shkolnisky, "Algebraically accurate volume registration using euler's theorem and the 3D pseudo-polar FFT," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, June 2005, pp. 795–800.

[17] Y. Keller, Y. Shkolnisky, and A. Averbuch, "Volume registration using the 3-D pseudopolar fourier transform," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4323–4331, Nov 2006.

[18] A. Averbuch and Y. Shkolnisky, "3D Fourier based discrete Radon transform," *Applied and Computational Harmonic Analysis*, vol. 15, pp. 33–69, 2003.

[19] D. Bailey and P. Swarztrauber, "The fractional fourier transform and applications," *SIAM Review*, vol. 33, no. 3, pp. 389–404, 1991.

[20] L. Almeida, "The fractional fourier transform and time-frequency representations," *Signal Processing, IEEE Transactions on*, vol. 42, no. 11, pp. 3084–3091, Nov 1994.

[21] Y. Keller, Y. Shkolnisky, and A. Averbuch, "The angular difference function and its application to image registration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 6, pp. 969–976, June 2005.

[22] H. Blow and A. Birk, "Spectral registration of noisy sonar data for underwater 3d mapping," *Autonomous Robots*, vol. 30, no. 3, pp. 307–331, 2011.

[23] H. Bulow and A. Birk, "Spectral 6DOF registration of noisy 3d range data with partial overlap," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 4, pp. 954–969, April 2013.

[24] Q. Wang, O. Ronneberger, and H. Burkhardt, "Rotational invariance based on fourier analysis in polar and spherical coordinates," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 9, pp. 1715–1722, Sept 2009.

[25] H. Skibbe, M. Reisert, T. Schmidt, T. Brox, O. Ronneberger, and H. Burkhardt, "Fast rotation invariant 3D feature computation utilizing efficient local neighborhood operators," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 8, pp. 1563–1575, Aug 2012.

[26] Q. Wang, O. Ronneberger, and H. Burkhardt, "Rotational invariance based on fourier analysis in polar and spherical coordinates," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 9, pp. 1715–1722, Sept 2009.

[27] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3D shape descriptors," in *Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, ser. SGP '03. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2003, pp. 156–164.

[28] D.M.Healy, Jr., D.N.Rockmore, P.J.Kostelec, and S.Moore, "FFTs for the 2-sphere-improvements and variations," *The Journal of Fourier Analysis and Applications*, vol. 9, no. 4, pp. 341–385, 2003.

[29] P. J.Kostelec and D. N.Rockmore, "FFTs on the rotation group," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 145–179, 2008.

[30] A. Makadia and K. Daniilidis, "Direct 3D-rotation estimation from spherical images via a generalized shift theorem," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2, June 2003, pp. II–217–24.

[31] A. Makadia, L. Sorgi, and K. Daniilidis, "Rotation estimation from spherical images," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 3, Aug 2004, pp. 590–593.

[32] A. Makadia and K. Daniilidis, "Rotation recovery from spherical images without correspondences," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 7, pp. 1170–1175, July 2006.

[33] A. Makadia, A. Patterson, and K. Daniilidis, "Fully automatic registration of 3D point clouds," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1, June 2006, pp. 1297–1304.

[34] J. L.Horner and P. D.Gianino, "Phae-only matched filtering," *Applied Optics*, vol. 23, no. 6, pp. 812–816, March 1984.

[35] W. Hoge, "A subspace identification extension to the phase correlation method [mri application]," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 2, pp. 277–280, Feb 2003.

[36] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *Image Processing, IEEE Transactions on*, vol. 11, no. 3, pp. 188–200, Mar 2002.

[37] W. Hoge and C. F. Westin, "Identification of translational displacements between N-dimensional data sets using the high-order SVD and phase correlation," *Image Processing, IEEE Transactions on*, vol. 14, no. 7, pp. 884–889, July 2005.