Extracting 3D rotation invariant features from a density map is the foundation for designing efficient tools for template-free similarity detection. Rotation invariant features have been used in the fields of image analysis and computer vision for object recognition during recent years (for examples [5], [6]). These features can characterize the local density distribution around a given location in a density map. However, rotation invariant features face two challenges when applied to molecular density maps, namely, the low resolution and the high level of noise in cryo-EM density maps. So far, little is known about the performance of rotation invariant features in the characterization of molecular density maps that are subject to noise.

In this paper, we propose and study the performance of three rotation invariant features for their ability to characterize molecular density maps. These features include moments based features, histogram features, and spherical harmonics features. The paper is organized as follow: First, we describe the definition of these features and the construction of corresponding feature vectors (section II). In particular, we propose Fast Fourier Transform based methods to speed up calculations of the first two types of features. We then test the performance of these features for classifying molecular density maps. Section III compares the classification performance of these features under different settings, and section IV provides concluding remarks.

II. METHODS

A. General idea

A molecular density map is a 3D image defined by a equally spaced 3D grid of voxels that are associated with values of density intensities. This density grid can be represented as a real function $f(x)$, where $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ is the discrete spatial location of a voxel.

Given the density map $f$, our aim is to calculate a rotation invariant feature vector $P(x)$ that characterizes the density distribution in the proximity of a given voxel location $x$. This feature vector represents the characteristic signature of the density distribution defined by all voxels in the neighborhood of $x$ in $\mathbb{R}^3$. We first describe how we define the neighborhood volume for a voxel $x$, before we further
introduce three methods to calculate a feature vector \( P(x) \) from the intensities of voxels in the neighborhood volume.

1) Definition of Neighborhood Volumes: For each voxel \( x \), a neighborhood volume \( C_{r,s}(x) \) is defined as the group of voxels that fall into in a concentric shell centered in \( x \) and defined by a lower radius of \( r \) and an upper radius of \( s \) (with \( r < s \)). The proximity of each voxel \( x \) can then be represented by a series of concentric shells \( C_{r_{i-1},r_i}(x) \) with \( r_0 < r_1 < \ldots < r_N \) (Figure 1). For a given voxel location \( x \), we define \( N \) concentric shells as \( \{C_{r_{i-1},r_i}(x) : i = 1, \ldots, N\} \) with \( R \) as the largest radius and \( r_i = iR/N \). For simplicity we denote a neighborhood shell \( C_{r_{i-1},r_i}(x) \) from hereon as \( C_i(x) \), representing the \( i \)th shell at voxel location \( x \), which is defined by the two radii \( r_{i-1} \) and \( r_i \).

![Figure 1. Neighborhood volumes defined as a series of concentric shells around voxel location x. Schematic view of a 2D grid with individual voxels shown as dark grey dots. Concentric shells are constructed that are centered at x. The largest radius is defined as R. All radii are defined as \( r_i = iR/N \), with \( N \) the maximal number of shells. A neighborhood volume \( C_{r_{i-1},r_i}(x) \) is defined as all voxels that fall into a concentric shell defined by two radii, \( r_{i-1} \) and \( r_i \) with \( r_{i-1} < r_i \). As an example the neighborhood shell \( C_{10}(x) \) is shown in light red, defined as the set of voxels located between radii \( r_9 \) and \( r_{10} \).](image)

2) Definition of Feature Vectors: A feature vector characterizes the density distribution in the proximity of a given location \( x \) and is constructed by a series of feature values that are derived from the density. The derivation of these feature values must be invariant to rotations of the density map around \( x \). Formally, we define a feature operator \( \hat{P} \), which takes the density values for all voxels in \( C_i(x) \) and generates one or several feature values. For a given voxel location \( x \) a feature vector \( P(x) \) can be defined as the ordered sequence of the feature values for all of its concentric shells \( \{C_i(x) : i = 1, \ldots, N\} \). Therefore,

\[
P(x) = (\hat{P}(C_1(x)), \hat{P}(C_2(x)), \ldots, \hat{P}(C_N(x))).
\]

We will define three types of rotation invariant feature operators, which are described in detail in the following sections.

B. Moment based features

Feature operator \( \hat{P}_{MO} \) calculates the mean and standard deviation that are based on the first and second statistical moments of the voxel intensities in \( C_i(x) \), and \( \hat{P}_{MO}(C_i(x)) = (\mu_i(x), \sigma_i(x)) \),

The average voxel intensity for \( C_i(x) \) is calculated as:

\[
\mu_i(x) = \frac{1}{|C_i(x)|} \int_{y \in C_i(x)} f(y)dy
\]

where \( |C_i(x)| \) is the total number of voxels in \( C_i(x) \). Calculating \( \mu_i(x) \) by summing over all voxel intensities is computationally very expensive, because \( \mu_i(x) \) must be calculated for all \( N \) neighborhood shells \( \{C_i(x) : i = 1, \ldots, N\} \) at all possible locations \( x \). Since density maps are typically large this direct approach is not feasible and it is necessary to speed up the calculations. Here, we propose an approach based on the Fast Fourier Transform (FFT).

Equation 1 can be written as the convolution between a characteristic function \( \chi \) and the density function \( f \). The characteristic function \( \chi \) can be defined for a given set of voxels \( C_i(x) \) as:

\[
\chi_{C_i(x)}(y) = \begin{cases} 
1 & \text{if } y \in C_i(x), \\
0 & \text{if } y \notin C_i(x).
\end{cases}
\]

and

\[
\mu_i(x) = \frac{1}{|C_i(x)|} \int \chi_{C_i(x)}(y) f(y)dy
\]

Because the density map is defined on an equally spaced grid, for all \( x, \ y \in \mathbb{R}^3, \ i = 1, \ldots, N \) we can set \( \chi_{C_i(x)}(y) = \chi_{C_i(0)}(y - x) \) and it is clear that \( |C_i(x)| = |C_i(0)| \). Moreover we can define a function \( \chi_{C_i(0)}(y) = \chi_{C_i(x)}(-y) \), so that

\[
\mu_i(x) = \frac{1}{|C_i(0)|} \int \chi_{C_i(0)}(y - x) f(y)dy
\]

\[
= \frac{1}{|C_i(0)|} \int \chi_{C_i(0)}(x - y) f(y)dy
\]

\[
= \frac{1}{|C_i(0)|} (\chi_{C_i(0)} * f)(x)
\]

(4)

Based on the convolution theorem, \( (\chi_{C_i(0)} * f)(x) \) can be calculated from the pointwise product of the Fourier transform of \( f \) and the complex conjugate of the Fourier transform of \( \chi \). Specifically, let \( \mathcal{F} \) be the Fourier transform operator that can be applied to \( \chi_{C_i(0)} \), \( \chi'_{C_i(0)} \) and \( f \). The convolution theorem states, \( \forall \xi \in \mathbb{R} \),

\[
\mathcal{F}(\chi_{C_i(0)} * f)(\xi) = \mathcal{F}(\chi_{C_i(0)})(\xi) \mathcal{F}(f)(\xi)
\]

(5)
\( \mu_i \) can therefore be directly computed from the inverse Fourier transform of \( \mathcal{F}(\chi_{C_i}(0) * f) \), where \( \chi_{C_i}(0) \) only needs to be calculated once for every given shell \( C_i(0) \) with \( i = 1, \ldots, N \), therefore

\[
\mu_i = \frac{1}{|C_i(0)|} \mathcal{F}^{-1} (\mathcal{F}(\chi_{C_i}(0)) \cdot \mathcal{F}(f))
\]  

(6)

where \( \cdot \) denotes point-wise multiplication.

The standard deviation \( \sigma_i(x) \) of the voxel intensities in \( C_i(x) \) can also be calculated using FFT. Let \( g_{i,x}(y) = (f(y) - \mu_i(x))^2 \). We can formulate:

\[
\sigma_i(x) = \sqrt{\frac{1}{|C_i(x)|} \int \chi_{C_i(x)}(y) g_{i,x}(y) dy}
\]

\[
= \sqrt{\frac{1}{|C_i(0)|} \int (\chi_{C_i(0)} * g_{i,x})(x)}
\]

(7)

According to the convolution theorem, we have

\[
\sigma_i = \sqrt{\frac{1}{|C_i(0)|} \mathcal{F}^{-1} (\mathcal{F}(\chi_{C_i(0)}) \cdot \mathcal{F}(g_{i,x}))}
\]

(8)

C. Histogram features

The feature operator \( \hat{P}_{HI} \) approximates the statistical distribution of voxel intensities in \( C_i(x) \) by a histogram that is calculated as the fractions of voxels in \( C_i(x) \) whose intensities fall into bins with defined range of intensities. Specifically, given \( k \) bins defined by \( k + 1 \) intensity values \( b_j \) with \( b_0 < b_1 < \ldots < b_k \). Let \( H_j \) be the set of locations whose corresponding voxel intensity fall into the \( j \)th bin defined by \( b_{j-1} \) and \( b_j \) cutoff values, i.e. \( H_j = \{ y: b_{j-1} < f(y) \leq b_j \} \), then the fraction of voxels in \( C_i(x) \) with intensities that fall into the \( j \)th bin is calculated as

\[
h_{ij}(x) = \frac{|C_i(x) \cap H_j|}{|C_i(x)|}
\]

(9)

Let the corresponding characteristic function \( \chi_{H_j} \) be defined as:

\[
\chi_{H_j}(y) = \begin{cases} 
1 & \text{if } y \in H_j \\
0 & \text{if } y \notin H_j
\end{cases}
\]

(10)

Then we can express \( h_{ij}(x) \) in the following form:

\[
h_{ij}(x) = \frac{1}{|C_i(0)|} \int \chi_{H_j}(y) \chi_{C_i(x)}(y) dy
\]

\[
= \frac{1}{|C_i(0)|} \int (\chi_{C_i(0)} * \chi_{H_j})(x)
\]

(11)

Based on the convolution theorem, we formulate

\[
h_{ij} = \frac{1}{|C_i(0)|} \mathcal{F}^{-1} (\mathcal{F}(\chi_{C_i(0)}) \cdot \mathcal{F}(\chi_{H_j}))
\]

(12)

and can accelerate the calculation of \( h_{ij}(x) \) using FFT. We obtain a set of rotation invariant histogram features:

\[
\hat{P}_{HIST}(C_i(x)) = (h_{i1}(x), \ldots, h_{ik}(x))
\]

(13)

D. Spherical harmonics features

If a cocentric shell \( C_i(x) \) is thin, i.e. \( r_{i+1} - r_i \approx \) voxel length, then the voxel intensities \( f(y) \) with \( y \in C_i(x) \) can be approximated by a spherical function \( g \) that is defined on the surface of a sphere in spherical coordinates.

\[
f(y)|_{C_i(x)} \approx f(r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \phi)
\]

(14)

where \( r = \frac{r_i + r_{i+1}}{2} \), and \( \theta \) and \( \phi \) are azimuthal and longitudinal coordinate, respectively. In such case, a set of rotation invariant SH features for \( C_i(x) \) can be constructed from \( g \).

\( g \) can be decomposed as an infinite sum of its spherical harmonics [7]:

\[
g(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}^{m}(\theta, \phi)
\]

(15)

where \( a_{lm} \) is a coefficient associated with the complex spherical harmonics function \( Y_{lm}^{m} \) that is independent to \( g \). Based on such a decomposition, a set of rotation invariant features \( s_{il}(x) \) can be constructed from \( g \) as follows [6]:

\[
\hat{P}_{SH}(C_i(x)) = (s_{i1}(x), s_{i2}(x), \ldots)
\]

(16)

where

\[
s_{il}(x) = ||g_l|| = \sqrt{\int \int |g_l(\theta, \phi)|^2 d\theta d\phi}
\]

(17)

with

\[
g_l(\theta, \phi) = \sum_{m=-l}^{l} a_{lm} Y_{lm}^{m}(\theta, \phi)
\]

(18)

For each shell \( C_i(x) \), we calculate \( L \) spherical harmonics features \( \hat{P}_{SH}(C_i(x)) = (s_{i1}(x), \ldots, s_{iL}(x)) \), where \( L \) is a defined maximum cutoff value for the spherical harmonics series.

E. Experimental design

Our goal is to test the performance of the rotation invariant density features to discriminate the density maps of different proteins. This task can be treated as a binary classification problem.

1) Test set: Protein classification is performed with two randomly chosen proteins from different protein families: the biotin carboxylase (PDB id 1dv2) and Sufi (PDB id 2uxt).

For each protein, we generate a simulated density map at 5 Å resolution from the atomic structure, following a procedure in [8]. We then generate density maps of each protein at different noise levels by adding uniformly distributed noise to the maps.
2) Classification procedure: We compare the performance of the feature vectors by measuring how often a single feature vector from a randomly selected protein at a given density position can correctly predict the identity of the protein. The classification procedure is defined by the following steps:

For simplification we denote the two protein as $m_A$ and $m_B$ respectively, and treat $m_A$ and $m_B$ as positive and negative classes respectively. First, randomly select one protein $m$ from the set of two proteins $m_A$ and $m_B$. Then, randomly select a voxel $v$ (at a location $x$) in the density map of $m$ and determine the corresponding feature vector $P(x)$. Next, calculate the difference between $P(x)$ and feature vectors of all other voxels in the density maps of both $m_A$ and $m_B$. The difference between two feature vectors is measured by their Euclidian distance, denoted as $d(P(x), P(y))$.

Given a predefined cutoff value $c_{\text{cut}}$, calculate the number $N_A = |\{y : d(P(x), P(y)) \leq c_{\text{cut}}, y \in m_A\}|$, where $y \in m_A$ corresponds to a voxel location in $m_A$. $N_A$ represents the number of times a feature vector $P(y)$ is detected in $m_A$ whose distance to the target feature vector $P(x)$ is smaller than the cutoff $c_{\text{cut}}$. Accordingly, the number $N_B = |\{y : d(P(x), P(y)) \leq c_{\text{cut}}, y \in m_B\}|$ is calculated from all voxel locations in $m_B$. A classification score is then defined as the difference between the two values $s = N_A - N_B$. Let $d_{\text{dec}}$ be a predefined decision cutoff, if $s > d_{\text{dec}}$, then the voxel $v$ is predicted to be part of protein $m_A$, and if $s < d_{\text{dec}}$ the voxel $v$ is predicted to be part of protein $m_B$. This classification scheme is repeated in 10,000 independent runs, each time starting from a random protein and random voxel location $x$. By comparing predictions with ground truth from the 10,000 classifications the true and false positive rates (denoted as $TP$ and $FP$ respectively), as well as the true and false negative rates (denoted as $TN$ and $FN$ respectively) of the classification are estimated. The cutoff $c_{\text{cut}}$ is selected by the following procedure: We have randomly sampled 100,000 pairs of voxels chosen from the set of all voxels in both proteins. For each voxel location the corresponding feature vector is generated and the Euclidian distance between the pair of feature vectors is calculated. The difference cutoff $c_{\text{cut}}$ is set to a distance value, which separates the 1% smallest distances in the distribution of feature vector distances. Smaller values would improve the classification performance for all feature types. However, the aim of this paper is to compare the performance between the three feature types and the chosen $c_{\text{cut}}$ value allows a clearer feature discrimination. The receiver operator curves (ROC) are generated by calculating the true positive rates and false positive rates of the classification for different decision cutoffs $d_{\text{dec}}$.

### III. Results

The classification performance of the three types of features are tested on molecular density maps with a maximal shell radius $R = 10$ voxel lengths and $N = 10$ shells. For computing histogram features, we set $k = 5$ bins, $b_0$ and $b_5$ be the minimum and maximum intensity respectively, and $b_1, \ldots, b_4$ are set to be equally spaced in the interval $(b_0, b_5)$. For computing spherical harmonics features we set the maximum cutoff value $L = 5$.

For density maps without noise the classification performance of the three types of features are very similar (Figure I top panel). Their classification performance is very good with an area under the ROC curve (AUC) above 0.8 (Table I). This performance indicates that feature vectors capture local similarities in density maps and therefore the feature types appear well suited for template-free detection of similar regions in molecular density maps.

We next test the classification performance for features using noisy density maps. For this purpose, noise is added to the original density maps by adding random numbers following a uniform distribution $U(0, l_n \mu_{\text{int}})$ where $\mu_{\text{int}}$ is the mean intensity of maps and $l_n$ a noise level factor. We compare the feature performance for maps with $l_n = 0, 2$ and $4$ (Figure 2). With increasing noise level one can expect that the classification performance decreases for all of the calculated features types. Interestingly, the SH feature is clearly more robust to noise as can be seen by the AUC values of the ROC curves (Figure 2 and Table I). In fact the classification performance decreases only slightly at high noise levels (Figure 2). This observation is in great contrast to the moments based as well as histogram features, which both show significantly reduced performance with increasing noise levels (Table I). This observed robustness of the SH feature types in density map classification is of great importance since experimental 3D density maps of biological samples often contain high level of noise.

### IV. Conclusion

3D rotation invariant feature extraction is a key step for providing template-free comparisons of noisy molecular density maps. In this paper, we test the power of three 3D rotation invariant features for the structural characterization of molecular density maps at different noise levels. Moreover, we propose FFT based methods for the fast calculation of moments and histogram feature vectors. The FFT based methods accelerate the calculation of the feature vectors for our density maps by a factor of more than a hundred in

---

### Table I

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Noise level $l_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moment</td>
<td>0.83 0.73 0.66</td>
</tr>
<tr>
<td>Histogram</td>
<td>0.82 0.73 0.71</td>
</tr>
<tr>
<td>SH</td>
<td>0.85 0.85 0.78</td>
</tr>
</tbody>
</table>

**AREA UNDER CURVE (AUC) FOR THE CLASSIFICATION OF DENSITY MAPS AT DIFFERENT NOISE LEVELS**
Our current results show that at low noise levels, all three features perform equally well. Because the calculation of histogram based features can be accelerated by FFT, and are more robust than the moment based features, they represent the method of choice for characterizing density maps at low noise level. Because SH features take into account the spectral information, they are significantly more robust in comparison to moments based features when the noise level in density maps is high. For studying highly noisy experimental density maps it is expected that SH features will outperform moments-based features and although the computational time for feature calculation is higher, SH features are recommended.

Although this study focuses on molecular density maps, our work can also provide guidance for constructing rotation invariant features for other types of noisy 3D biological images, such as cellular protein fluorescent images. We are currently working on including other types of 3D rotation invariant features and testing a wider range of settings.

ACKNOWLEDGMENT

We would like to thank Prof. Xianghong Jasmine Zhou for her contributions and valuable suggestions and Prof. Z. Hong Zhou for valuable discussion and the problem formulation. We also acknowledge financial support from the Alfred P. Sloan foundation, the PEW charitable trust and the Human Frontier Science Program (RGY0079/2009-C).

REFERENCES