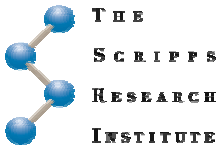


3D and 6D Fast Rotational Matching

Julio Kovacs, Ph.D.

Department of Molecular Biology
The Scripps Research Institute
10550 N. Torrey Pines Road, Mail TPC6
La Jolla, California, 92037

Situs Modeling Workshop, San Diego, CA, Feb. 3-5, 2003

Highlights

- Computationally efficient method to perform rigid body docking.
- Combines spherical harmonics with new set of parameters for rotations, allowing us to deal with the *rotational correlation function* by means of Fourier techniques.
- Applications to:
  - Docking of atomic structures into EM density maps.
  - The "Molecular Replacement Method" in crystallography.
  - Multiple molecule ("interior") docking.

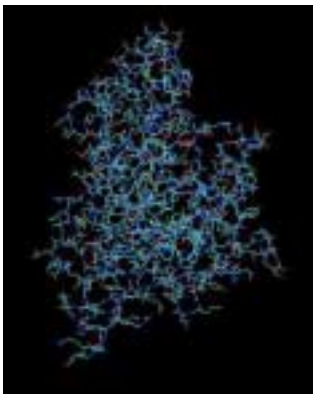
Versions of FRM

- FRMr: radial version —the starting point.
- FRM3D: volumetric version. Suitable for complicated maps such as Patterson maps in crystallography.
- FRM6D: for multiple molecule docking.
- FRM2D: analog of FRM6D for 2-D images. Has applications to reconstructions of EM maps (Yao Cong's lecture).

Radial FRM

CSB Group

Description of the method

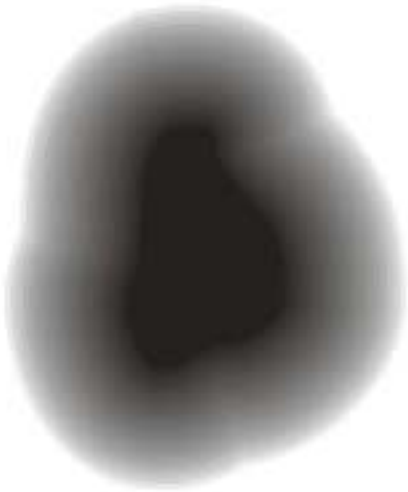


original map

radialized map

Radialization procedure

CSB Group

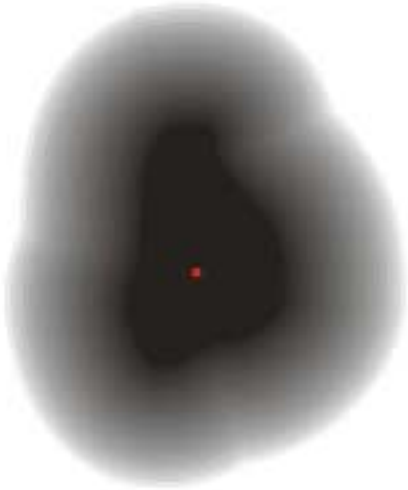


2-dimensional slice of the density map.

(Electron microscopy map or ‘blurred’ atomic structure)

Radialization procedure

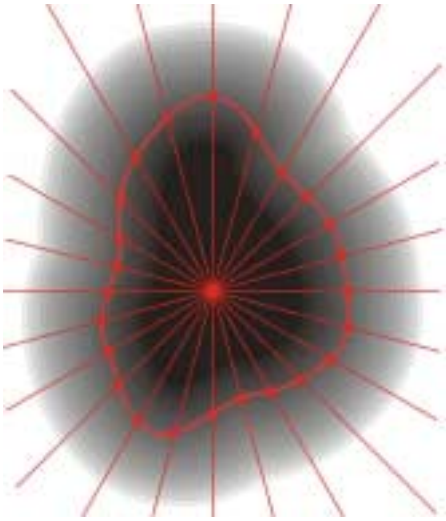
CSB Group



First, the center of mass of each density map is computed

Radialization procedure

CSB Group



Then the desired level surface is determined on an equiangular grid, giving the radialized map

What do we do with the radialized maps?

Radialized map f → $f = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{f}_{lm} Y_{lm}$

Y_{lm} are the *spherical harmonic functions*

Radialized map g → $g = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{g}_{lm} Y_{lm}$

CSB Group

f and g represent the same object

↓

There is a scaling factor λ and a rotation R such that $\lambda \cdot g(R) \approx f$, where $g(R)$ denotes rotation of g by R.

↓

Criterion is to maximize the *rotational correlation function*:

$$c(R) = \int_{S^2} f \cdot g(R)$$

Euler angles ϕ, θ, ψ → change to → $\xi = \phi - \pi/2$, $\eta = \pi - \theta$, $\omega = \psi - \pi/2$

$$c(R) = T(\xi, \eta, \omega)$$

The Fourier transform of T turns out to be:

$$\hat{T}(m,h,m') = \sum_l d^l_{mh}(\tfrac{\pi}{2})\, d^l_{hm'}(\tfrac{\pi}{2})\, \hat{f}_{lm} \bar{\hat{g}}_{lm'}$$

CSB Group

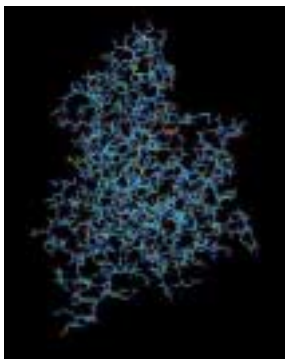
The quantities $d^l_{mn}(\frac{\pi}{2})$ (which come from rotation group theory) are precomputed using a recursive procedure, since the explicit formula is quite involved

$\hat{T}(m,h,m')$ → inverse FFT → $T(\xi, \eta, \omega)$

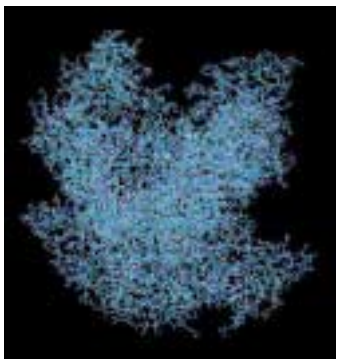
Performance of FRMr

CSB Group

Timing test:



ncd
(2629 atoms)



RNA polymerase
(24837 atoms)

Timings shown correspond to a 1GHz Pentium III Linux PC

Times for FRMr and FTM

CSB Group

(6° angular sampling)

	ncd		RNAp	
main processes	FRMr	FTM	FRMr	FTM
precomputations	0.060		0.060	
centers of mass	0.005		0.040	
sampling	0.049		0.181	
spherical coefficients	0.004		0.004	
$\hat{T}$	0.124		0.124	
inverse FFT	0.146		0.146	
subtotal	**0.39^{s}**	**13^{m}**	**0.56^{s}**	**7^{h}**
side processes	0.44		2.9	

FTM = Fast Translational Matching

Robustness test of FRMr

LYSOZYME (1294 atoms)

BARNASE (2624 atoms)

RIBONUCLEASE INHIBITOR (3410 atoms)

HUMAN SERUM ALBUMIN (4475 atoms)

LETHAL FACTOR PRECURSOR (5947 atoms)

HALOPEROXIDASE (4567 atoms)

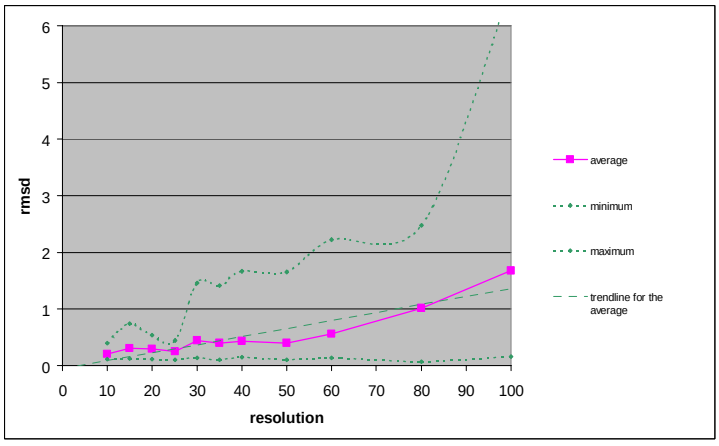
UREASE γ SUBUNIT (6040 atoms)

CHOLERA TOXIN (4098 atoms)

RNA polymerase (24837 atoms)

rmsd vs. resolution

(6° angular sampling)



Volumetric FRM (FRM3D)

Here we handle the original 3D density functions instead of their radialized versions. The correlation function is:

$$c(R) = \int_{\mathbf{R}^3} f \cdot g(R) = T(\xi, \eta, \omega)$$

Expanding in spherical harmonics:

$$f(ru) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{f}_{lm}(r) Y_{lm}(u) \qquad g(ru) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{g}_{lm}(r) Y_{lm}(u)$$

We arrive at: $\hat{T}(m, h, m') = \sum_{l} d^{l}_{mh}(\tfrac{\pi}{2})\, d^{l}_{hm'}(\tfrac{\pi}{2})\, I^{l}_{mm'}$

where: $I^{l}_{mm'} = \int_{0}^{\infty} \hat{f}_{lm}(r)\, \overline{\hat{g}_{lm'}(r)}\, r^2 dr$

Crowther's Fast Rotation Function

This classical method (1972) uses the regular Euler angles as parameters of the correlation function:

$$c(R) = \int_{\mathbf{R}^3} f \cdot g(R) = T(\phi, \psi; \theta)$$

By doing this, only ϕ and ψ can be Fourier-transformed:

$$\hat{T}(m, m'; \theta) = \sum_{l} d^{l}_{mm'}(\theta) I^{l}_{mm'}$$

while θ remains as a parameter that needs to be scanned sequentially.

FRM3D vs. Crowther

CSB Group

$$c(R) = T(\xi, \eta, \omega)$$

$$\hat{T}(m, h, m') = \sum_l d^l_{mh}(\tfrac{\pi}{2}) d^l_{hm'}(\tfrac{\pi}{2}) I^l_{mm'}$$

FRM3D

- Precomputations:
  - FST
  - $d^l_{mn}(\frac{\pi}{2})$
  - FFTW plan
- Reading maps
  - COMs & sampling
  - $\hat{f}_{lm}(r),\ \hat{g}_{lm'}(r),\ I^l_{mm'}$
- $\hat{T}(m, h, m')$
  - $T(\xi, \eta, \omega) = FFT^{-1}_{3D}(\hat{T})$
- postprocessing & saving

$$c(R) = T(\phi, \psi; \theta)$$

$$\hat{T}(m, m'; \theta) = \sum_l d^l_{mm'}(\theta) I^l_{mm'}$$

Crowther

- Precomputations:
  - FST
  - FFTW plan
- Reading maps
  - COMs & sampling
  - $\hat{f}_{lm}(r),\ \hat{g}_{lm'}(r),\ I^l_{mm'}$
- $k = 0, \ldots, B$
  - $\theta_k = \frac{k\pi}{B}, \quad d^l_{mm'}(\theta_k)$
  - $\hat{T}(m, m'; \theta_k)$
  - $T(\phi, \psi; \theta_k) = FFT^{-1}_{2D}(\hat{T})$
- postprocessing & saving

Timings (in seconds) of Crowther and FRM3D

CSB Group

B	Angular sampling	Crowther	FRM3D
32	6°	1.66	1.0
64	3°	19.3	4.0
128	1½°	337.	40.

FRM6D

CSB Group

Description of the method

The given density objects are expanded into spherical harmonics:

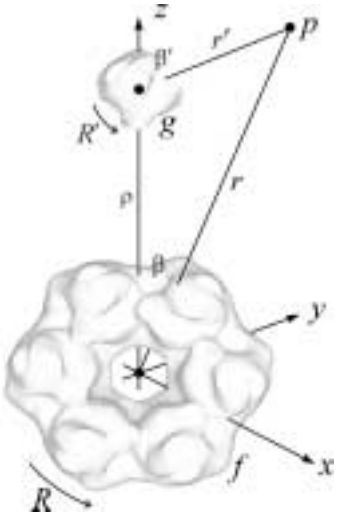
$$f(ru) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{f}_{lm}(r) Y_{lm}(u)$$

$$g(ru) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{g}_{lm}(r) Y_{lm}(u)$$

The idea is to rotate *both* objects while translating one of them along the positive *z* axis only. In this way, the 6D search is performed over 5 angular and 1 linear parameters.

6 DOF matching setup

CSB Group



The correlation is therefore a function of 2 rotations and a distance:

$$c(R, R'; \rho) = \int_{\mathbf{R}^3} f(R) \cdot g(R'; \rho).$$

Euler angles

$R(\phi, \theta, \psi)$ $\quad$ $R'(\phi', \theta', \psi')$

$$\xi = \phi - \pi/2 \qquad \xi' = \phi' - \pi/2$$
$$\eta = \pi - \theta \qquad \eta' = \pi - \theta'$$
$$\omega = \psi - \pi/2 \qquad \omega' = \psi' - \pi/2$$

Using these parameters the correlation function is written as:

$$c(R,R';\rho) = T(\xi - \xi', \eta, \omega, \eta', \omega'; \rho),$$

and the Fourier transform of T turns out to be:

$$\hat{T}(n,h,m,h',m';\rho) = (-1)^n \sum_{l,l'} d^{l}_{nh} d^{l}_{hm} d^{l'}_{-nh'} d^{l'}_{h'm'} I^{ll'}_{mnm'}(\rho).$$

The quantities $I^{ll'}_{mnm'}(\rho)$ are the so called *two-center integrals*, corresponding to the spherical harmonic transforms of the two maps, at a distance ρ of one another:

$$I^{ll'}_{mnm'}(\rho) = \sqrt{(l+\tfrac{1}{2})(l'+\tfrac{1}{2})} \int_0^{\pi} \left[\int_0^{\infty} \overline{\hat{f}_{lm}(r) \hat{g}_{l'm'}(r')}\, d^{l'}_{n0}(\beta') r^2 dr \right] d^{l}_{n0}(\beta) \sin\beta \, d\beta$$

The Scripps Research Institute
CSB Group

To compute these integrals more efficiently, we make the following change of variables:

$$z = r\cos\beta$$
$$s = r\sin\beta$$

whereby the integrals become:

$$I^{ll'}_{mnm'}(\rho) = \sqrt{(l+\tfrac{1}{2})(l'+\tfrac{1}{2})} \int_0^{\infty} \left[\int_{-\infty}^{\infty} \overline{\hat{f}_{lm}(r) \hat{g}_{l'm'}(r')}\, d^{l}_{n0}(\beta) d^{l'}_{n0}(\beta') dz \right] s\, ds$$

In order to do the numerical evaluation, we discretize z and s:

$$z = \bar{z}h$$
$$s = \bar{s}h$$

where the "barred" variables are integers and *h* is the step size. Introducing these we get:

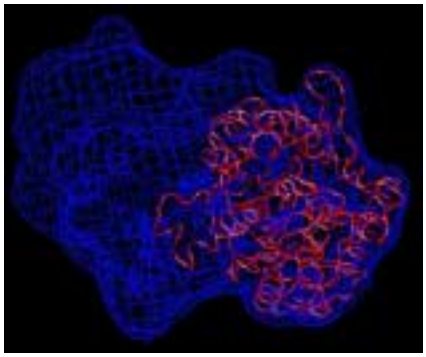
$$I_{mnm'}^{ll'}(\rho) \approx h^3 \sqrt{(l+\tfrac{1}{2})(l'+\tfrac{1}{2})} \cdot$$

$$\cdot \sum_{\bar{s}=1}^{B} \bar{s} \sum_{\bar{z}=\bar{z}_1}^{\bar{z}_2} \overline{\hat{f}_{lm}[[h\sqrt{\bar{s}^2+\bar{z}^2}\,]]\hat{g}_{l'm'}[[h\sqrt{\bar{s}^2+\bar{z}^2-2\bar{\rho}\bar{z}+\bar{\rho}^2}\,]]} \cdot$$

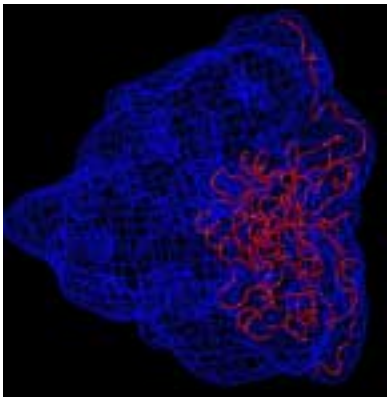
$$\cdot d_{n0}^{l}[[\tfrac{\bar{z}}{\bar{r}} B_x]] d_{n0}^{l'}[[\tfrac{\bar{z}-\bar{\rho}}{\bar{r}'} B_x]]$$

This is the actual expression used in the code.

Test cases



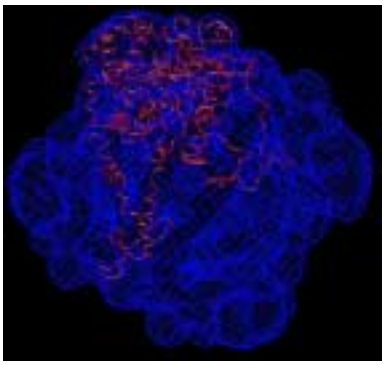
1afw
(peroxisomal thiolase)



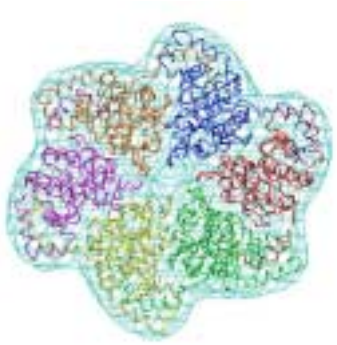
1nic (copper-
nitrite reductase)

Test cases

CSB Group



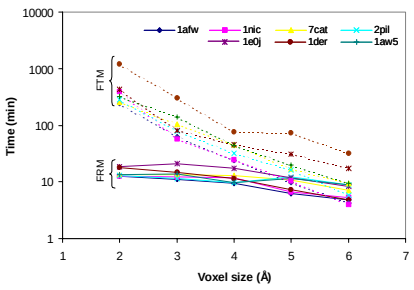
7cat
(catalase)



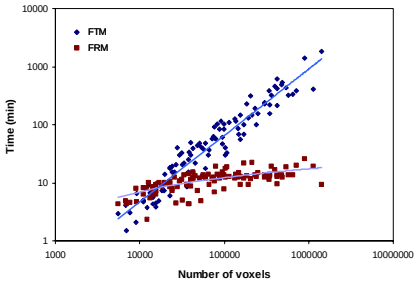
1e0j
(Gp4D helicase)

Comparison timings and trends

CSB Group



Times for various molecules versus the voxel size of the low resolution maps.



Timing trends versus the number of voxels of the low resolution maps.

Notes: the figure on the left shows average times for 3 values of resolutions (10, 15 and 20Å). In all cases, B=16, or 11° sampling. All tests were done applying the Laplacian filter prior to processing.

Conclusions

- Times of few seconds make FRMr and FRM3D suitable for interactive docking sessions.
- FRMr is scale-independent and surface-based.
- FRM3D and FRM6D can be made surface-based by Laplace-filtering the original maps.
- Surface filters, such as Laplacian, are needed in order to avoid spurious fitting results when using FRM6D.
- Will be included in Situs 2.2