Computerized Classification of Surface Spikes in Three-Dimensional Electron Microscopic Reconstructions of Viruses

Younes Benkarroum

The Graduate Center of City University of New York

08/26/2014
• Influenza is a rapidly changing virus that appears seasonally in the human population.

• Every year a new strain of the influenza virus appears with the potential to cause a serious global pandemic.

• During mixed infections, RNPs are reassorted resulting in new combination of HA and NA.

• Knowledge of the structure and density of the surface proteins is of critical importance in a vaccine candidate.

• ART with blobs provides 3D reconstructions of viruses from tomographic tilt series that allow the desired reliable quantification of the surface proteins.
Series Expansion Methods

- Assume that the reconstructed object can be described by a linear combination of finite set of fixed basis functions

\[ f^*(x, y, z) = \sum_{j=1}^{N} c_j b(x - x_j, y - y_j, z - z_j) \]

- Estimate the unknown coefficients of the linear combination based on projection images

\[ p_i^* = \sum_{j=1}^{N} a_{i,j} c_j \]

where \( a_{i,j} \) is the line integral, along the line \( i \), of the shifted basis function centered at \((x_j, y_j, z_j)\)
Algebraic Reconstruction Technique

- ART calculates and uses the difference between the measured data and the calculated forward projection to update the image

\[ c_j^{(k+1)} = c_j^{(k)} + \lambda \frac{p_j^{(k)} - \sum_h a_{i^{(k)},h} c_h^{(k)}}{\sum_h a_{i^{(k)},h}^2} a_{i^{(k)},j} \]

\( \lambda \) is the relaxation parameter and \( i^{(k)} = (k \mod M) + 1 \)

- The algorithm attempts to find a vector \( c \) that is an approximate solution to the linear system \( p = Ac \), where \( p \) is the data vector
Voxel Basis Function

• The choice of the set of basis functions greatly influences the result of the reconstruction algorithm.

• The conventional choice is the voxel basis function which has the value 1 inside the $j^{th}$ voxel and the value 0 otherwise.

• In such case, the coefficient $c_j$ becomes the average value of $f^*$ inside the $j^{th}$ voxel.

• Reconstructions using cubic voxels have undesirable artificial sharp edges.
Blob Basis Function

- Basis functions with spherical symmetry and a smooth transition from one to zero.

\[
b(r) = \begin{cases} 
\frac{l_2\left(\alpha \sqrt{1 - \left(\frac{r}{a}\right)^2}\right)}{l_2(\alpha)} \left(1 - \left(\frac{r}{a}\right)^2\right), & \text{if } 0 \leq x \leq a, \\
0, & \text{otherwise.}
\end{cases}
\]

where \( l_2 \) denotes the modified Bessel function of order 2, \( a \) determines the support of the blob (radius), and \( \alpha \) is a parameter controlling the blob shape.

\( \alpha = 6, \ a = 1 \)
$b(x - x_j)$
Approximation of Constant-Valued Functions Using Blobs

\[ f^* (x) = \sum_{j=1}^{N} c_j b(x - x_j), \quad \text{where } c_j = f(x_j). \]
Approximation of Constant-Valued Functions Using Blobs

\[ f^*(x) = \sum_{j=1}^{N} c_j b(x - x_j), \quad \text{where } c_j = f(x_j). \]

Mathematically, the function \( f^* \) can be written as:

\[ f^* = (f \times \mathbb{1}_{G_\Delta}) * b. \]
Approximation of Constant-Valued Functions Using Blobs

\[ f^* (x) = \sum_{j=1}^{N} c_j b(x - x_j), \quad \text{where} \ c_j = f(x_j). \]

Mathematically, the function \( f^* \) can be written as: \( f^* = (f \times \Pi_{G_\Delta}) * b \).

\[
f(x) = \begin{cases} 
1, & \text{if} \ |x| \leq \rho, \\
0, & \text{otherwise}.
\end{cases}
\]
**Approximation of Constant-Valued Functions Using Blobs**

\[
f^*(x) = \sum_{j=1}^{N} c_j b(x - x_j), \quad \text{where } c_j = f(x_j).
\]

Mathematically, the function \( f^* \) can be written as: \( f^* = (f \times \Pi G_\Delta) \ast b \).

\[
f(x) = \begin{cases} 
1, & \text{if } |x| \leq \rho, \\
0, & \text{otherwise}.
\end{cases}
\]

\[
\Pi_{G_\Delta}(x) = \sum_{p \in G_\Delta} \delta(x - p),
\]

where \( G_\Delta \) is the set of integers \( \times \Delta \).
Approximation of Constant-Valued Functions Using Blobs

\[ f^*(x) = \sum_{j=1}^{N} c_j b(x - x_j), \quad \text{where} \quad c_j = f(x_j). \]

Mathematically, the function \( f^* \) can be written as:

\[ f^* = (f \times \mathbb{I}_{G_{\Delta}}) \ast b. \]

\[ f(x) = \begin{cases} 
1, & \text{if } |x| \leq \rho, \\
0, & \text{otherwise.} 
\end{cases} \]

\[ \mathbb{I}_{G_{\Delta}}(x) = \sum_{p \in G_{\Delta}} \delta(x - p), \quad b(x) = \begin{cases} 
\frac{l_2\left(\alpha\sqrt{1 - \left(\frac{x}{a}\right)^2}\right)}{l_2(\alpha)} \left(1 - \left(\frac{x}{a}\right)^2\right), & \text{if } 0 \leq x \leq a, \\
0, & \text{otherwise.} 
\end{cases} \]

where \( G_{\Delta} \) is the set of integers \( \times \Delta \).
\[(f \times \mathbb{I}_G) \ast b\]

\[\Delta = 0.5\]
\[(f \times \text{III}_{G_\Delta}) \ast b\]
\((f \times \Pi G_\Delta) \ast b\)
\[ f^* = (f \times \mathbb{I}_G) * b \]
$f$ vs $f^*$
Fourier Space Analysis

\[ f^* = \left( f \times \mathbb{I}_{G_\Delta} \right) \ast b \iff \hat{f}^* = \left( \hat{f} \ast \mathbb{I}_{G_\Delta} \right) \times \hat{b} \]
Fourier Space Analysis

\[ f^\ast = (f \times \Pi_{\Gamma_\Delta}) \ast b \iff \hat{f}^\ast = (\hat{f} \ast \Pi_{\Gamma_\Delta}) \times \hat{b} \]

\[ \hat{f}(X) = 2\rho \frac{\sin(\rho X)}{\rho X} \]
Fourier Space Analysis

\[ f^* = (f \times \mathbb{I}_{G_{\Delta}}) \ast b \quad \iff \quad \hat{f}^* = (\hat{f} \ast \mathbb{I}_{G_{\Delta}}) \times \hat{b} \]

\[ \hat{f}(X) = 2\rho \frac{\sin(\rho X)}{\rho X} \]

\[ \mathbb{I}_{G_{\Delta}} = \frac{1}{\Delta^3} \mathbb{I}_{G_{1/\Delta}} \]
Fourier Space Analysis

\[ f^* = (f \ast \sum_{\Delta} G) \ast b \iff \hat{f}^* = (\hat{f} \ast \sum_{\Delta} G) \times \hat{b} \]

\[ \hat{f}(X) = 2\rho \frac{\sin(\rho X)}{\rho X} \]

\[ \sum_{\Delta} G = \frac{1}{\Delta^3} \sum_{\Delta} G_{1/\Delta} \]

\[ \hat{b}(X) = \begin{cases} 
(2\pi)^{\frac{3}{2}} a^3 \alpha^2 \frac{l_2(\alpha)}{l_2(\alpha)} \left( \frac{\sqrt{\alpha^2 - (2\pi a X)^2}}{\sqrt{(2\pi a X)^2 - \alpha^2}} \right)^{\frac{7}{2}}, & \text{if } 2\pi a X \leq \alpha, \\
(2\pi)^{\frac{3}{2}} a^3 \alpha^2 \frac{J_7(\sqrt{(2\pi a X)^2 - \alpha^2})}{l_2(\alpha)} \left( \frac{\sqrt{\alpha^2 - (2\pi a X)^2}}{\sqrt{(2\pi a X)^2 - \alpha^2}} \right)^{\frac{7}{2}}, & \text{otherwise.} 
\end{cases} \]
\((\hat{f} \cdot \mathbb{I}_{G_\Delta}) \times \hat{b}\)
\[(\hat{f} * \widehat{\mathcal{W}_{G_\Delta}}) \times \hat{b}\]
\( \hat{b} \) with random parameters:

\[ a = 0.8 \text{ and } \alpha = 11 \]
\( (\hat{f} * \hat{G}_\Delta) \times \hat{b} \)

\( \hat{b} \) with random parameters:

\( a = 0.8 \) and \( \alpha = 11 \)
\( \widehat{f} \ast \overline{\Pi_{G_{\Delta}}} \times \widehat{b} \)

\( \widehat{b} \) with random parameters:

\( a = 0.8 \) and \( \alpha = 11 \)
\( \hat{b} \) with random parameters:

\[ a = 0.8 \text{ and } \alpha = 11 \]
\( \hat{b} \) crossing \( N_1 \):

\[ a = 0.6 \text{ and } \alpha = 4.861220 \]
\[(\hat{f} * \mathbb{E}_{G\Delta}) \times \hat{b}\]
\[
\left( f * \overline{\text{III}_{G\Delta}} \right) \times \hat{b}
\]

\(\hat{b}\) crossing \(N_1:\)

\(a = 0.6\) and \(\alpha = 4.861220\)
\[ \hat{f}^* = (\hat{f} * \overline{G_\Delta}) \times \hat{b} \]
\[ f^* = (f \times \pi G_\Delta) \ast b \]

\[ \hat{b} \text{ crossing } N_1: \]
\[ a = 0.6 \text{ and } \alpha = 4.861220 \]
\( \hat{b} \text{ crossing } N_1: \)

\( a = 0.6 \text{ and } \alpha = 4.861220 \)
$f \text{ vs } f^*$

Random

Crossing $N_1$
\hat{b} crossing \( N_1 \) and \( N_2 \):

\( a = 0.500427 \) and \( \alpha = 2.515620 \)
Blob Parameters

\[
\hat{b}(X) = \begin{cases} 
\frac{(2\pi)^{3/2} a^3 \alpha^2}{l_2(\alpha)} & \frac{l_7/2}{J_7/2} \left( \frac{\sqrt{\alpha^2 - (2\pi aX)^2}}{\sqrt{(2\pi aX)^2 - \alpha^2}} \right)^{7/2}, \\
\text{if } 2\pi aX \leq \alpha, \\
\frac{(2\pi)^{3/2} a^3 \alpha^2}{l_2(\alpha)} & \frac{J_7/2}{J_7/2} \left( \frac{\sqrt{(2\pi aX)^2 - \alpha^2}}{\sqrt{(2\pi aX)^2 - \alpha^2}} \right)^{7/2}, \\
\text{otherwise.}
\end{cases}
\]
Blob Parameters

\[
\hat{b}(X) = \begin{cases} 
\frac{(2\pi)^{\frac{3}{2}} a^3 \alpha^2}{l_2(\alpha)} \frac{l_7/2}{\left(\sqrt{\alpha^2 - (2\pi a X)^2}\right)^{7/2}}, & \text{if } 2\pi a X \leq \alpha, \\
\frac{(2\pi)^{\frac{3}{2}} a^3 \alpha^2}{l_2(\alpha)} \frac{J_{7/2}}{\left(\sqrt{(2\pi a X)^2 - \alpha^2}\right)^{7/2}}, & \text{otherwise.} 
\end{cases}
\]

\[
\sqrt{(2\pi a X_1)^2 - \alpha^2} = j_1 \\
\sqrt{(2\pi a X_2)^2 - \alpha^2} = j_2
\]
Blob Parameters

\[
\hat{b}(X) = \begin{cases} 
\frac{(2\pi)^{3/2} a^3 \alpha^2}{l_2(\alpha)} \frac{l_{7/2}(\sqrt{\alpha^2 - (2\pi a X)^2})}{\left(\sqrt{\alpha^2 - (2\pi a X)^2}\right)^{7/2}}, & \text{if } 2\pi a X \leq \alpha, \\
\frac{(2\pi)^{3/2} a^3 \alpha^2}{l_2(\alpha)} \frac{J_{7/2}(\sqrt{(2\pi a X)^2 - \alpha^2})}{\left(\sqrt{(2\pi a X)^2 - \alpha^2}\right)^{7/2}}, & \text{otherwise.}
\end{cases}
\]

\[
\sqrt{(2\pi a X_1)^2 - \alpha^2} = j_1 \\
\sqrt{(2\pi a X_2)^2 - \alpha^2} = j_2
\]

\[
a = \frac{1}{2\pi} \sqrt{\frac{j_2^2 - j_1^2}{X_2^2 - X_1^2}} \\
\alpha = \sqrt{\frac{X_2^2 j_2^2 - X_1^2 j_1^2}{X_2^2 - X_1^2}}
\]
\[ \hat{f}^* = (\hat{f} \ast \text{III}_{G\Delta}) \times \hat{b} \]

\( \hat{b} \) crossing \( N_1 \) and \( N_2 \):
\[ a = 0.500427 \text{ and } \alpha = 2.515620 \]
\[ \hat{f}^* = (\hat{f} * \overline{G_\Delta}) \times \hat{b} \]
\[ \hat{f}^* = (f * G_{\Delta}) \times \hat{b} \]
\[ f^* = (f \times \mathbf{III}_{G_\Delta}) \ast b \]

\[ \hat{b} \] crossing \( N_1 \) and \( N_2 \):
\[ a = 0.500427 \text{ and } \alpha = 2.515620 \]
$\hat{b}$ crossing $N_1$ and $N_2$:

$a = 0.500427$ and $\alpha = 2.515620$
$f$ vs $f^*$

Crossing $N_1$

Crossing $N_1 \& N_2$
\( \hat{b} \) with 3\(^{rd} \) zero crossing at \( N_2 \)

Blob parameters:
\( a = 0.500427 \) and \( \alpha = 2.515620 \)
\( \hat{b} \) with 4\(^{th} \) zero crossing at \( N_2 \)

**Blob parameters:**

\( a = 0.661795 \) and \( \alpha = 5.995361 \)
\( \hat{b} \) with 9\(^{th}\) zero crossing at \( N_2 \)

Blob parameters:
\( a = 1.414401 \) and \( \alpha = 16.813501 \)
Bessel Function

\[ \sqrt{(2\pi aX_1)^2 - \alpha^2} = j_1 \]

\[ \sqrt{(2\pi aX_2)^2 - \alpha^2} = j_2 \]
\( \hat{b} \) with 9\(^{th}\) zero crossing at \( N_2 \):
\[ a = 1.414401 \] and \( \alpha = 16.813501 \)
$f \text{ vs } f^*$

$3^{rd}$ zero crossing at $N_2$

$9^{th}$ zero crossing at $N_2$
$f$ vs $f^*$

$3^{rd}$ zero crossing at $N_2$

$9^{th}$ zero crossing at $N_2$
\( \hat{f} \text{ vs } (\hat{b}_3 \text{ & } \hat{b}_9) \)
\( \hat{f} \times (\hat{b}_3 \& \hat{b}_9) \)
3D Grids

\[ \overline{III}_{B_{\beta}} = \frac{1}{4\beta^3} \overline{III}F_{1/2\beta} \]
Approximation of Constant-Valued Functions Using Blobs

\[ f^* = (f \times \text{III}_{B_\beta}) * b \]

\[ f(\vec{x}) = \begin{cases} 
  1, & \text{if } \|\vec{x}\| \leq \rho, \\
  0, & \text{otherwise}. 
\end{cases} \]

\[ b(\vec{x}) = \begin{cases} 
  \frac{l_2(\alpha \sqrt{1 - (\frac{r}{a})^2})}{l_2(a)} \left(1 - \left(\frac{r}{a}\right)^2\right), & \text{if } 0 \leq r \leq a, \\
  0, & \text{otherwise}, 
\end{cases} \]

where \( r \) denotes the norm \( \|\vec{x}\| \) of the vector \( \vec{x} \).
Fourier Space Analysis

\[ f^* = (f \times \Pi_{B_\beta}) \ast b \iff \hat{f}^* = (\hat{f} \ast \Pi_{\hat{B}_\beta}) \times \hat{b} \]
Fourier Space Analysis

\[ f^* = (f \times \mathbb{I}_{B_\beta}) \ast b \iff \hat{f}^* = (\hat{f} \ast \mathbb{I}_{\mathbb{I}_{B_\beta}}) \times \hat{b} \]

\[ \hat{f}(R) = \frac{\sin(2\pi \rho R) - 2\pi \rho \cos(2\pi \rho R)}{2\pi^2 R^3} \]

\[ \mathbb{I}_{B_\beta} = \frac{1}{4\beta^3} \mathbb{I}_{F_{1/2\beta}} \]

\[ \hat{b}(X) = \begin{cases} 
\left(\frac{2\pi}{l_2(\alpha)}\right)^3 \alpha^2 \frac{J_{3/2}(\sqrt{(2\pi aX)^2 - \alpha^2})}{\left(\sqrt{(2\pi aX)^2 - \alpha^2}\right)^{7/2}}, & \text{if } 2\pi aX \leq \alpha, \\
\left(\frac{2\pi}{l_2(\alpha)}\right)^3 \alpha^2 \frac{I_{7/2}(\sqrt{\alpha^2 - (2\pi aX)^2})}{\left(\sqrt{\alpha^2 - (2\pi aX)^2}\right)^{7/2}}, & \text{otherwise.} 
\end{cases} \]
\((\hat{f} \ast \overline{\beta} \cdot \beta) \times \hat{b}\)
\( \hat{f} \ vs \ (\hat{b}_4 \ & \hat{b}_9) \)
$\hat{f}$ vs $(\hat{b}_4 \& \hat{b}_9)$
\((f \ast \mathcal{H}_{B\beta}) \times \hat{b}\)
$\left( f \ast \Pi_{B_{\beta}} \right) \times \hat{b}$
\[(\hat{f} \ast \overline{\overline{B}}_\beta) \times \hat{b}\]
\((\hat{f} \ast \mathbb{I}_{B_{\beta}}) \times \hat{b}\)
Recommended Blob Parameters

\[ a = \frac{\beta}{\pi \sqrt{2}} \sqrt{j_y^2 - j_z^2} \quad \alpha = \sqrt{j_y^2 - 2j_z^2} \]

<table>
<thead>
<tr>
<th>( j_y, j_z )</th>
<th>( a )</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( j_1, j_2 )</td>
<td>( \beta \times 1.738875 )</td>
<td>3.294537</td>
</tr>
<tr>
<td>( j_1, j_3 )</td>
<td>( \beta \times 2.651778 )</td>
<td>9.485434</td>
</tr>
<tr>
<td>( j_1, j_4 )</td>
<td>( \beta \times 3.469269 )</td>
<td>13.738507</td>
</tr>
<tr>
<td>( j_1, j_5 )</td>
<td>( \beta \times 4.247117 )</td>
<td>17.527826</td>
</tr>
<tr>
<td>( j_1, j_6 )</td>
<td>( \beta \times 5.003932 )</td>
<td>21.105107</td>
</tr>
<tr>
<td>( j_1, j_7 )</td>
<td>( \beta \times 5.748062 )</td>
<td>24.563319</td>
</tr>
<tr>
<td>( j_1, j_8 )</td>
<td>( \beta \times 6.483890 )</td>
<td>27.946764</td>
</tr>
<tr>
<td>( j_1, j_9 )</td>
<td>( \beta \times 7.213964 )</td>
<td>31.279745</td>
</tr>
</tbody>
</table>
Central slice of the phantom and its blob representations. Display window: 0.999 - 1.001
Plot of the ball phantom (blue) and the blob representations using $z = 4$ (red) and $z = 9$ (black) along the line $(x, 0, 0)$. 
Errors

\[ E_{\beta,r_1,r_2} = \int_{|\tilde{X}|<\frac{1}{2}\sqrt{2}\beta} \left( 1 - \frac{\hat{b}_{r_1,r_2}(\tilde{X})}{\hat{b}_{r_1,r_2}(\bar{0})} \right)^2 d\tilde{X} + \int_{|\tilde{X}|>\frac{1}{2}\sqrt{2}\beta} \left( \frac{\hat{b}_{r_1,r_2}(\tilde{X})}{\hat{b}_{r_1,r_2}(\bar{0})} \right)^2 d\tilde{X}. \]
## Errors

<table>
<thead>
<tr>
<th>$j_{r_1}$ \ $j_{r_2}$</th>
<th>$j_4$</th>
<th>$j_5$</th>
<th>$j_6$</th>
<th>$j_7$</th>
<th>$j_8$</th>
<th>$j_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j_1$</td>
<td>0.211697</td>
<td>0.255772</td>
<td>0.294584</td>
<td>0.327072</td>
<td>0.353848</td>
<td>0.375856</td>
</tr>
<tr>
<td>$j_2$</td>
<td>0.218827</td>
<td>0.257221</td>
<td>0.294687</td>
<td>0.326843</td>
<td>0.353571</td>
<td>0.375609</td>
</tr>
<tr>
<td>$j_3$</td>
<td></td>
<td>0.313518</td>
<td>0.307784</td>
<td>0.331719</td>
<td>0.355704</td>
<td>0.376623</td>
</tr>
<tr>
<td>$j_4$</td>
<td></td>
<td></td>
<td>0.362973</td>
<td>0.366868</td>
<td>0.381674</td>
<td></td>
</tr>
<tr>
<td>$j_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.424956</td>
<td>0.400413</td>
</tr>
</tbody>
</table>

Values of the error $E_{\beta,r_1,r_2}$ for the $(r_1, r_2)$ pairs $(1 \leq r_1 < r_2 \leq 9)$ with $\beta = 1/\sqrt{2}$
### Recommended Blob Parameters

<table>
<thead>
<tr>
<th>$j_1, j_4$</th>
<th>$\alpha$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta \times 3.469269$</td>
<td>13.738507</td>
<td>13.738507</td>
</tr>
</tbody>
</table>
Recommended Blob Parameters

| $j_1, j_4$ | $\alpha$ | $\beta \times 3.469269$ | 13.738507 |

Standard Blob Parameters

$$\alpha = \sqrt{2\pi^2 \left( \frac{a}{\beta} \right)^2 - 6.987932^2}$$
Recommended Blob Parameters

| $j_1, j_4$  | $\beta \times 3.469269$ | 13.738507 |

$\beta = 1/\sqrt{2}$  
$a = 2.453144$  
$\alpha = 13.738507$

Standard Blob Parameters

$$\alpha = \sqrt{2\pi^2 \left(\frac{a}{\beta}\right)^2 - 6.987932^2}$$

$\beta = 1/\sqrt{2}$  
$a = 2.0$  
$\alpha = 10.4$
Approximations Using Blobs

Display Window: 0.999 - 1.001

Recommended Parameters
\[a = 2.453144\]
\[\alpha = 13.738507\]

Standard Parameters
\[a = 2.0\]
\[\alpha = 10.4\]

Modified Standard Parameters
\[a = 2.0\]
\[\alpha = 10.444256\]
Approximation Using Blobs

Display Window: 0.0 - 1.5

Central slice of the FORBILD abdomen phantom and its blob representation
Approximation Using Blobs

Display Window: 0.18 - 0.3175

Central slice of the FORBILD thorax phantom and its blob representation
Reconstructions Using Blobs

Recommended Parameters
\( a = 2.453144 \)
\( \alpha = 13.738507 \)

Standard Parameters
\( a = 2.0 \)
\( \alpha = 10.4 \)
ART Reconstruction

Display Window: 0.0 - 1.5

(a) blob reconstruction
(b) voxel reconstruction

Central slice of the reconstruction of the FORBILD abdomen phantom
ART Reconstruction

Display Window: 0.18 - 0.3175

Central slice of the reconstruction of the FORBILD thorax phantom
Projection Images
Data Processing

Theory requires that the object to be reconstructed can be represented by a function of finite support.

\[
p_i = \int_a^b \gamma \, dl + \int_b^c \nu(l) \, dl + \int_c^d \gamma \, dl,
\]

\[
p_i' = \int_{a'}^{d'} \gamma \, dl = \int_a^d \gamma \, dl.
\]
Data Processing

Theory requires that the object to be reconstructed can be represented by a function of finite support.

\[ p_i = \int_a^b \gamma \, dl + \int_b^c \nu(l) \, dl + \int_c^d \gamma \, dl, \]

\[ p_i' = \int_{a'}^{d'} \gamma \, dl = \int_a^d \gamma \, dl. \]

\[ p_i - p_i' = \int_b^c [\nu(l) - \gamma] \, dl. \]

\[ \int_{-\infty}^{\infty} f(l) \, dl = \int_b^c [\nu(l) - \gamma] \, dl. \]
Illustration of the projection of two viruses: (left) tilt angle = 0°, (right) tilt angle = 60°.
\[ C_{1u} = \sqrt{x_1^2 + x_3^2} \times \cos(\alpha_u + \arctan\left(\frac{x_3}{x_1}\right)) , \]
Least Squares Estimator

\[ C_{1u}(x_1, x_3) = \sqrt{x_1^2 + x_3^2} \times \cos \left( \alpha_u + \arctan \left( \frac{x_3}{x_1} \right) \right), \]

Let 
\[ d(x_1, x_3) = \sum_u \left( C_{1u}(x_1, x_3) - C'_{1u} \right)^2, \]

\( C'_{1u} \)s are the manually determined values.
Least Squares Estimator

\[ C_{1u}(x_1, x_3) = \sqrt{x_1^2 + x_3^2} \times \cos \left( \alpha_u + \arctan \left( \frac{x_3}{x_1} \right) \right), \]

Let
\[ d(x_1, x_3) = \sum_u \left( C_{1u}(x_1, x_3) - C'_{1u} \right)^2, \]

\( C'_{1u} \)s are the manually determined values.

\[ d(x_{1d}, x_{3d}) = \min_{x_1, x_3} (d(x_1, x_3)), \]

\[ C_{1u} = \sqrt{x_{1d}^2 + x_{3d}^2} \times \cos \left( \alpha_u + \arctan \left( \frac{x_{3d}}{x_{1d}} \right) \right). \]
Least Squares Estimator

\[ C_{1u}(x_1, x_3) = \sqrt{x_1^2 + x_3^2} \times \cos \left( \alpha_u + \arctan \left( \frac{x_3}{x_1} \right) \right), \]

Let
\[ d(x_1, x_3) = \sum_u \left( C_{1u}(x_1, x_3) - C'_{1u} \right)^2, \]

\( C'_{1u} \)s are the manually determined values.

\[ d(x_{1d}, x_{3d}) = \min_{x_1, x_3} \left( d(x_1, x_3) \right), \]

\[ C_{1u} = \sqrt{x_{1d}^2 + x_{3d}^2} \times \cos \left( \alpha_u + \arctan \left( \frac{x_{3d}}{x_{1d}} \right) \right). \]
Data Processing
Processed Data

Three projections of the isolated virus after processing
ART Reconstruction

Three different near-central cross-sections
Recommended
Docked atomic models of HA (yellow) and NA (red)

Fig. 1.12: (Left) Central tomogram slice of influenza B/Lee/40. Several HA (yellow) and NA (red) are shown. RNPs are outlined in blue. (Right) Enlarged view of slice with atomic models of HA and NA docked in the image. Bar is 25 nm
Identification of a Region that Contains All Spikes
Segmentation

- Thresholding Segmentation.

   Fisher’s linear discriminant:
   \[ J = \frac{|m_1 - m_2|^2}{s_1^2 + s_2^2}, \]

   where \( m_i \) and \( s_i^2 \) represent the mean and the variance of the class \( i \).

- Simultaneous fuzzy segmentation into multiple objects.

- Segmentation based on the use of component trees.
Feature Extraction & Classification

a) HA, HA-NA, NA

b, c, d) Image of proteins and virus reconstructions
Evaluation Methodology

• Create 3D phantoms of viruses having HA and NA spikes.

• Generate projection images by simulating the behavior of cryogenic electron microscopy.

• Calculate 3D reconstructions from projection images.

• Apply the full classification procedure to the outputs of the reconstructions.

• Test the classifier massively on many test phantoms with different HA and NA conformations.

• Evaluate the performance of the classifier using a simple and transparent figure of merit called classification purity.
Evaluation Methodology

<table>
<thead>
<tr>
<th></th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA</td>
<td>890</td>
<td>110</td>
</tr>
<tr>
<td>NA</td>
<td>70</td>
<td>930</td>
</tr>
</tbody>
</table>

Table: Example of an array created to evaluate the classification purity.

- The classification purity is 91%

\[
(100 \times (890 + 930)) / (890 + 110 + 70 + 930) \%
\]

- An efficacious classification procedure should result in a high value of classification purity (above 95 %).
## Timeline

<table>
<thead>
<tr>
<th>Item</th>
<th>Execution time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical Development</td>
<td>Sep 2014 - Feb 2015</td>
</tr>
<tr>
<td>Experiments with simulated data</td>
<td>Mar 2015 - Apr 2015</td>
</tr>
<tr>
<td>Dissertation writing</td>
<td>May 2015 - Jul 2015</td>
</tr>
<tr>
<td>Review by the committee and defense</td>
<td>Aug 2015</td>
</tr>
</tbody>
</table>