Visualization Tool
for
Flow Cytometry Data Standards Project

Evgeny Maksakov
maksakov@cs.ubc.ca

CS533C
Department of Computer Science, UBC
in collaboration with
Terry Fox Laboratory, BC Cancer Agency
(Prof. Ryan Brinkman & Dr. Josef Spidlen)
Today

• Flow Cytometry Overview
  – Dataset description
• Existing Visualizations Overview
• Data analysis
  – Current (FlowJo)
  – Proposed
• Prototype Progress
• Future Work
Flow Cytometry

Cell  Measure

Measuring properties of cells in a fluid stream
## List of Flow Cytometry Application Fields

<table>
<thead>
<tr>
<th>Chromatin structure</th>
<th>Immunophenotyping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total protein</td>
<td>DNA cell cycle/tumor ploidy</td>
</tr>
<tr>
<td>Lipids</td>
<td>Membrane potential</td>
</tr>
<tr>
<td>Surface charge</td>
<td>Ion flux</td>
</tr>
<tr>
<td>Membrane fusion/runover</td>
<td>Cell viability</td>
</tr>
<tr>
<td>Enzyme activity</td>
<td>Intracellular protein staining</td>
</tr>
<tr>
<td>Oxidative metabolism</td>
<td>pH changes</td>
</tr>
<tr>
<td>Sulphydryl groups/glutathione</td>
<td>Cell tracking and proliferation</td>
</tr>
<tr>
<td>DNA synthesis</td>
<td>Sorting</td>
</tr>
<tr>
<td>DNA degradation</td>
<td>Redox state</td>
</tr>
</tbody>
</table>

The list is taken from [http://www.basic.northwestern.edu/sharedresources/flowcytometry/](http://www.basic.northwestern.edu/sharedresources/flowcytometry/)
Flow Cytometry (FCM)

FACSCalibur Optical Layout

- **488 nm Blue Laser**
- **Red Diode Laser ~635 nm**
- **Flow Cell**
- **Beam Combiner**
- **FSC Diode**
- **90/10 Beam Splitter**
- **DM 560SP**
- **DM 640LP**
- **Half-Mirror**
- **Fluorescence Collection Lens**
- **FL1 530/30**
- **FL2 585/42**
- **FL3**
- **FL4 661/16**
- **670 LP**
- **661/16**
- **Cy3, DsRed PE PI**
- **Alexa 635 APC Cy5 TOPRO.3**
- **7-AAD PE-Cy5 PerCP PI**

**Labels:**
- **Alexa 488 FITC FDG GFP YFP**
Dataset Properties

Typically for research at the TFL:
- 100,000+ events
- 5-10 dimensions

Capability:
- 1,000,000 events (cells going through the laser beam) per dataset
- Up to 20 dimensions
Dimensions (2 basic dimensions)

side scatter detector (Granularity)

light source (Laser)

forward scatter detector (Size)
Dimensions (GFP intensity & PI)

Green Fluorescent Protein intensity
measures gene expression

Aequorea Victoria (natural owner of GFP)

Mice glow green under ultraviolet light

PI (Propidium Iodide) dye intensity
measures cells’ viability (life cells expunge the dye)

Dimensions (16 fluorescence intensities)

<table>
<thead>
<tr>
<th>Fluorochrome</th>
<th>Fluorescence Emission Color</th>
<th>Ex-Max (nm)</th>
<th>Excitation Laser Line (nm)</th>
<th>Em-Max (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexa Fluor® 405</td>
<td>Blue</td>
<td>401</td>
<td>360, 405, 407</td>
<td>421</td>
</tr>
<tr>
<td>Pacific Blue®</td>
<td>Blue</td>
<td>405</td>
<td>360, 405, 407</td>
<td>455</td>
</tr>
<tr>
<td>AmCyan</td>
<td>Green</td>
<td>457</td>
<td>405, 407</td>
<td>491</td>
</tr>
<tr>
<td>Alexa Fluor® 488</td>
<td>Green</td>
<td>495</td>
<td>488</td>
<td>519</td>
</tr>
<tr>
<td>FITC</td>
<td>Green</td>
<td>494</td>
<td>488</td>
<td>519</td>
</tr>
<tr>
<td>PE</td>
<td>Yellow</td>
<td>496, 564</td>
<td>488, 532</td>
<td>578</td>
</tr>
<tr>
<td>PE-Texas Red®</td>
<td>Orange</td>
<td>496, 564</td>
<td>488, 532</td>
<td>615</td>
</tr>
<tr>
<td>Texas Red®**</td>
<td>Orange</td>
<td>595</td>
<td>595</td>
<td>615</td>
</tr>
<tr>
<td>APC*</td>
<td>Red</td>
<td>650</td>
<td>595, 633, 635, 647</td>
<td>660</td>
</tr>
<tr>
<td>Alexa Fluor® 647</td>
<td>Red</td>
<td>650</td>
<td>595, 633, 635, 647</td>
<td>668</td>
</tr>
<tr>
<td>PE-Cy5*</td>
<td>Red</td>
<td>496, 564</td>
<td>488, 532</td>
<td>667</td>
</tr>
<tr>
<td>PerCP</td>
<td>Red</td>
<td>482</td>
<td>488, 532</td>
<td>678</td>
</tr>
<tr>
<td>PerCP-Cy5.5</td>
<td>Far Red</td>
<td>482</td>
<td>488, 532</td>
<td>695</td>
</tr>
<tr>
<td>Alexa Fluor® 700***</td>
<td>Far Red</td>
<td>696</td>
<td>633, 635</td>
<td>719</td>
</tr>
<tr>
<td>PE-Cy7</td>
<td>InfraRed®</td>
<td>496, 564</td>
<td>488, 532</td>
<td>785</td>
</tr>
<tr>
<td>APC-Cy7</td>
<td>InfraRed®</td>
<td>650</td>
<td>595, 633, 635, 647</td>
<td>785</td>
</tr>
</tbody>
</table>

Attaching markers to cells

- Fluorochrome-labeled antibodies
- Antigenic surface marker

Steps:
1. Incubate
2. Analyze
Current Visualization Solutions

Made deliberately for FCM:

• **FlowJo** (scatterplots, histograms, contour diagrams)

• **FACSDiva** (scatterplots, histograms, contour diagrams)
Current Visualization Solutions

Universal data visualization tool:
• GGobi
  – Draw dotplots and scatterplots, barcharts, spineplots and histograms, parallel coordinate plots, scatterplot matrices
  – Link data points and lines between plots using brushing and identification
  – Pan and zoom
  – Rotate data in 3D and tour high-dimensional data using sequences of 1D, 2D and 2x1D projections
  – Uses R language for data manipulation
Data Analysis Process (FlowJo)

Negative control
(each scatterplot is a new window)

Event Count: 28988
Event Count: 18229
Event Count: 17755

*Event Count* is a total number of cells passed through the laser beam

Important note: sequence of actions is the same all the time for negative control!
Data Analysis Process (FlowJo)

Looking for result

Event Count: 16061

Important note: Same gates as in neg. control apply automatically on the positive set!
Other forms of result visualization (FlowJo)
Proposal

User requirements (based on user studies):
1. See all dimensions at once
2. Improve analysis sequence
3. Leave scatterplots and histograms (scientists used to them)
4. Gating/Filtering feature
5. Provide better usability than FlowJo

Solutions:
1. Use Parallel Coordinates with Gating/Filtering
2. Implement data clustering throughout dimensions
3. Include scatterplots and histograms in the interface
4. Make effective, convenient and interactive interface
Interface for FCM Data Analysis
Prototype progress

Highlighting of the gate. Random set, 3000 points, 7 dimensions.
Prototype progress

Filtering. Random set, 100 000 points, 7 dimensions. Full scale rendering takes ~1min.
Prototype progress

Interaction results. Random set, 3000 points, 7 dimensions.
Future Work

• Visualization of the real data
• Clustering
• Optimization
• User evaluation
3D Parallel Coordinate System for FCM

Marc Streit at al. (2006)
3D Parallel Coordinate System for FCM

- Does not provide any new information about dataset
- Introduces visual occlusions
- Have to rotate to see all data
- Unavailable

Picture from Marc Streit at al. (2006)
Questions...