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

Measuring properties of cells in a fluid stream

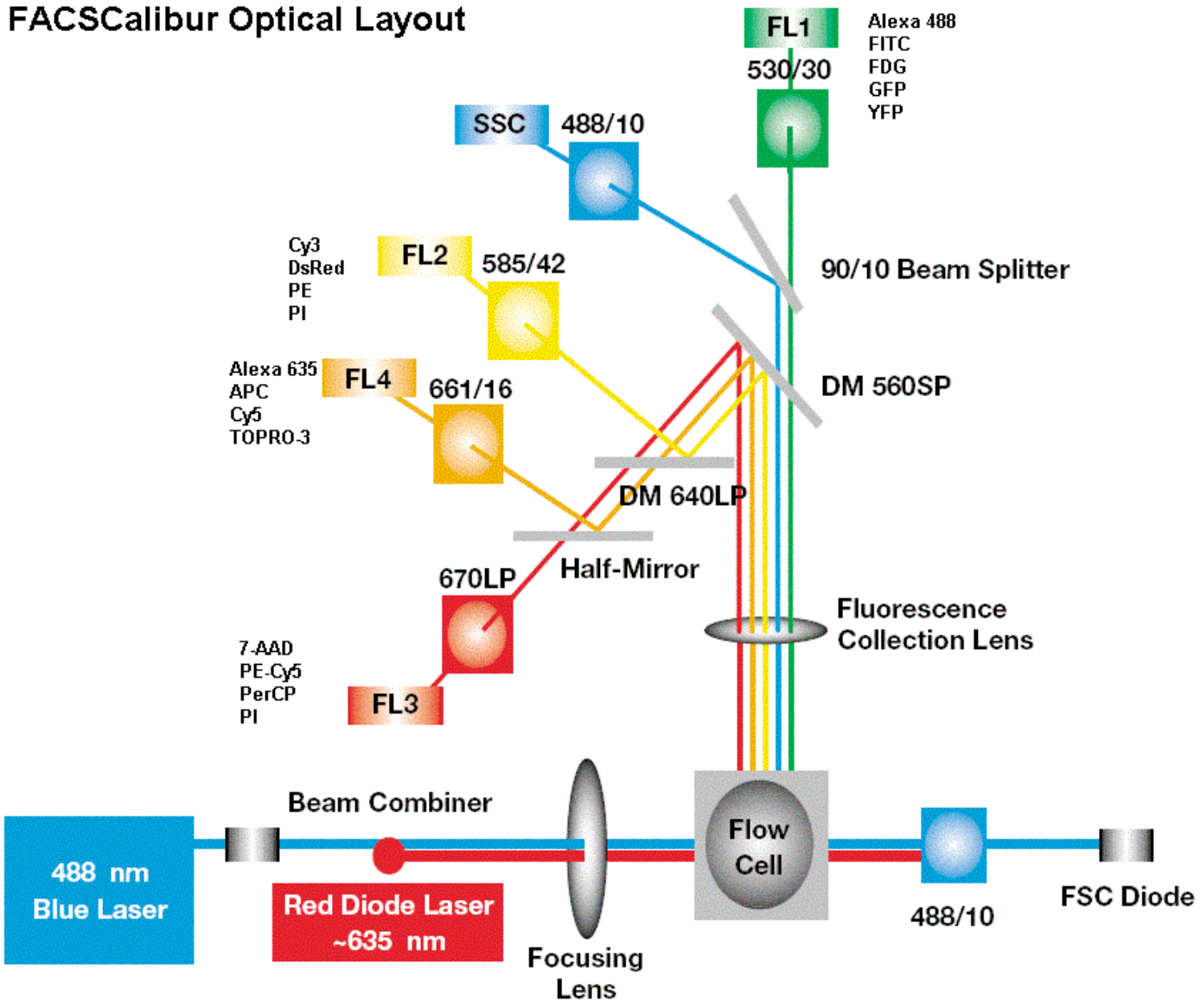
List of Flow Cytometry Application Fields

Immunophenotyping
DNA cell cycle/tumor ploidy
Membrane potential
Ion flux
Cell viability
Intracellular protein staining
pH changes
Cell tracking and proliferation
Sorting
Redox state

Chromatin structure
Total protein
Lipids
Surface charge
Membrane fusion/runover
Enzyme activity
Oxidative metabolism
Sulfhydryl groups/glutathione
DNA synthesis
DNA degradation
Gene expression

The list is taken from http://www.basic.northwestern.edu/sharedresources/flowcytometry/

Flow Cytometry (FCM)



Dataset Properties

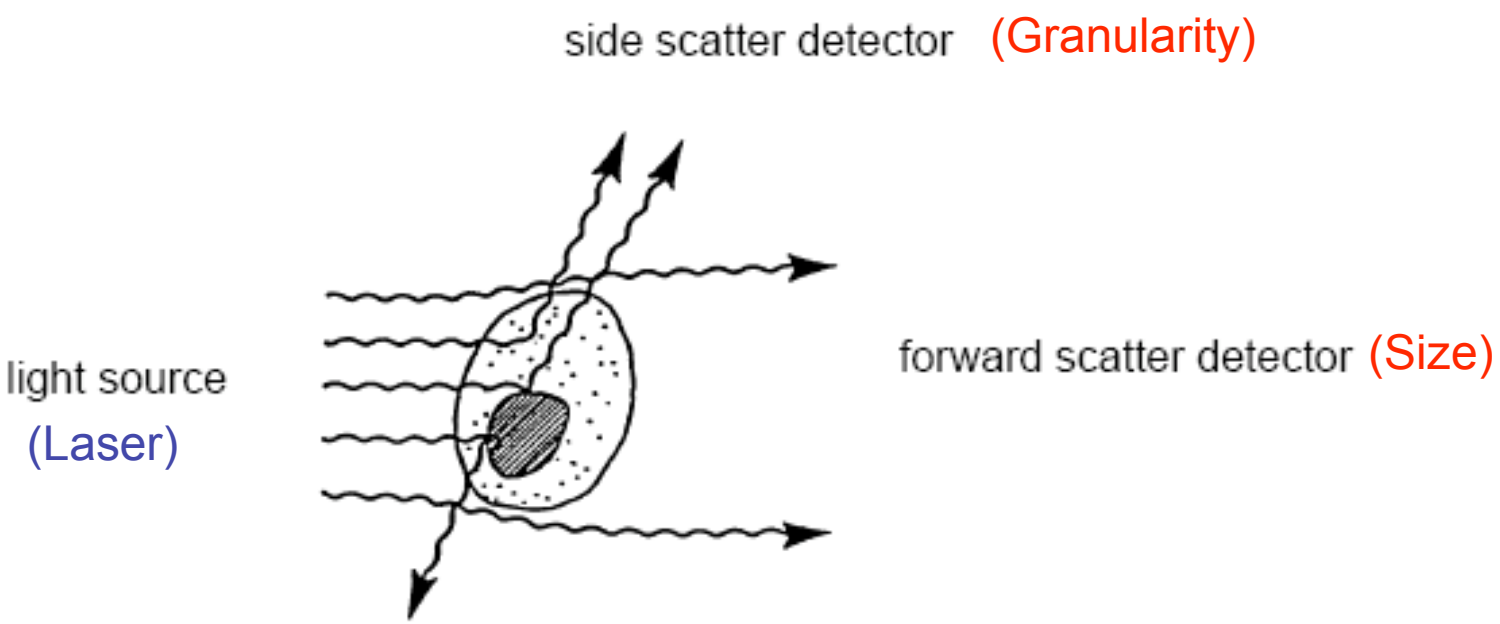
Typically for research at the TFL:

- 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)



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)

Pictures is taken from http://en.wikipedia.org/wiki/Image:Aequorea_victoria.jpg & http://www.upenn.edu/pennnews/photos/

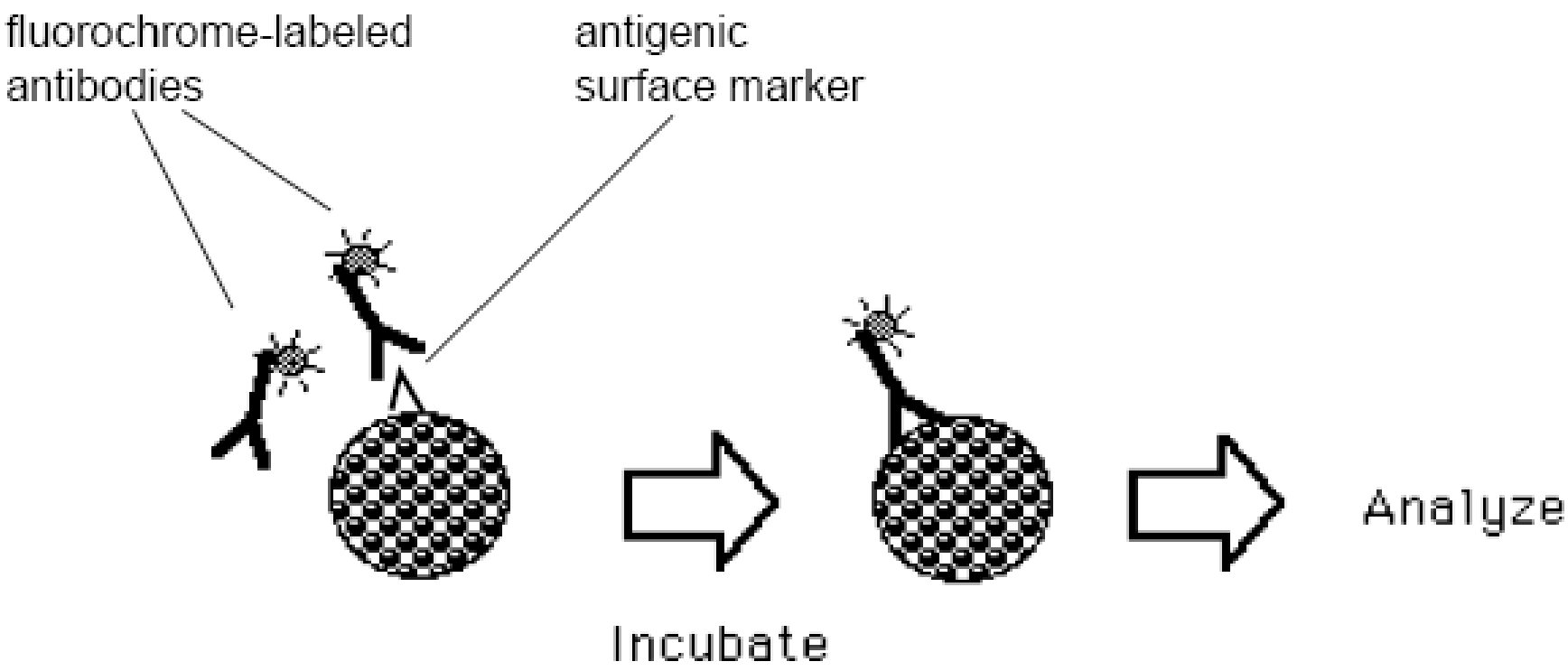
Dimensions (16 fluorescence intensities)

Fluorochrome Specifications

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

Picture from: http://www.bdbiosciences.com/image_library/

Attaching markers to cells



Current Visualization Solutions

Made deliberately for FCM:

- FlowJo (scatterplots, histograms, contour diagrams)
- FACSDiva (scatterplots, histograms, contour diagrams)

Current Visualization Solutions

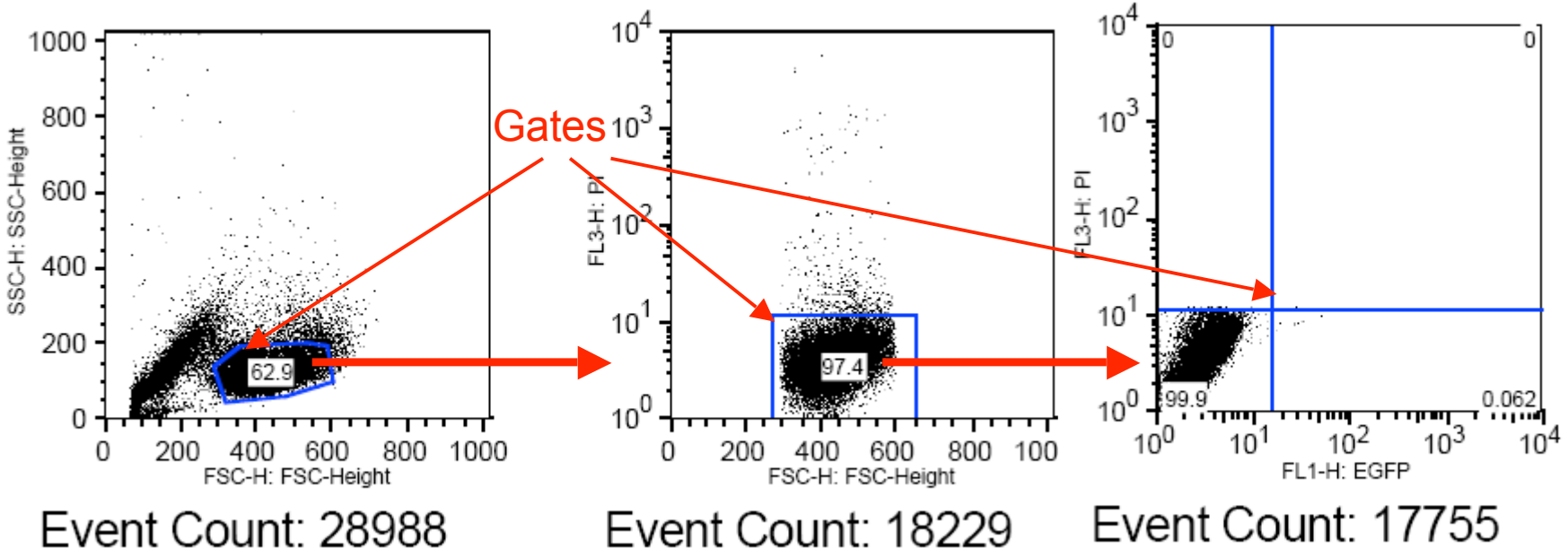
Universal data visualization tool:

- 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)

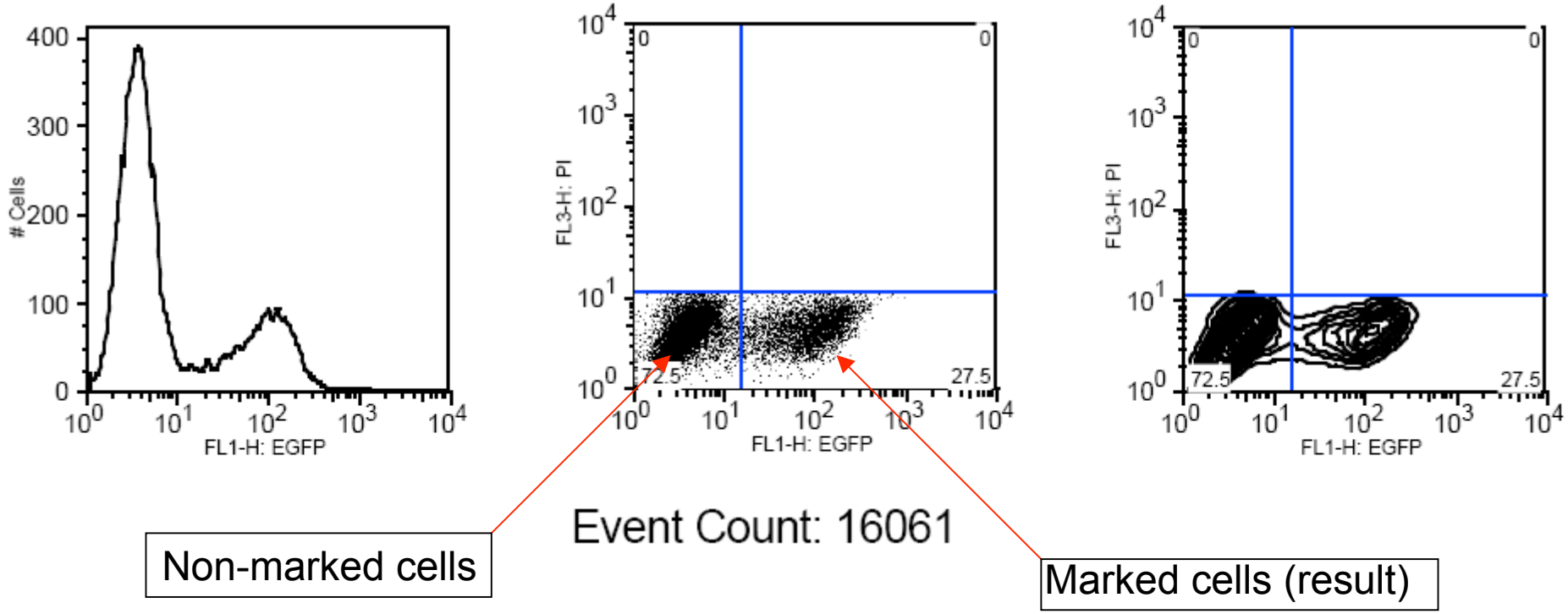


Event Count is a total number of cells passed through the laser beam

Important note: sequence of actions is the same all the time for negative control!

Data Analysis Process (FlowJo)

Looking for result



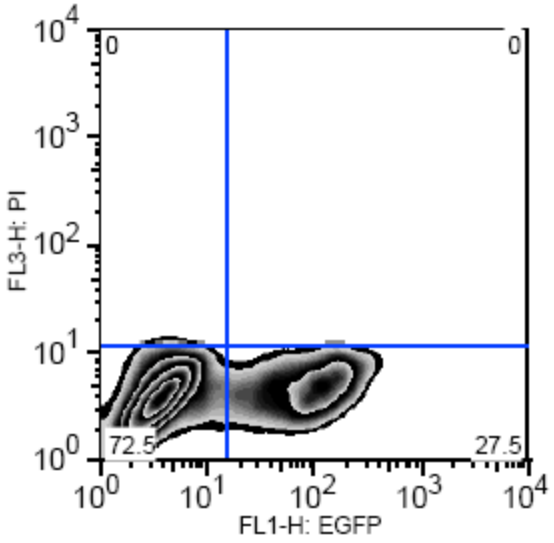
Event Count: 16061

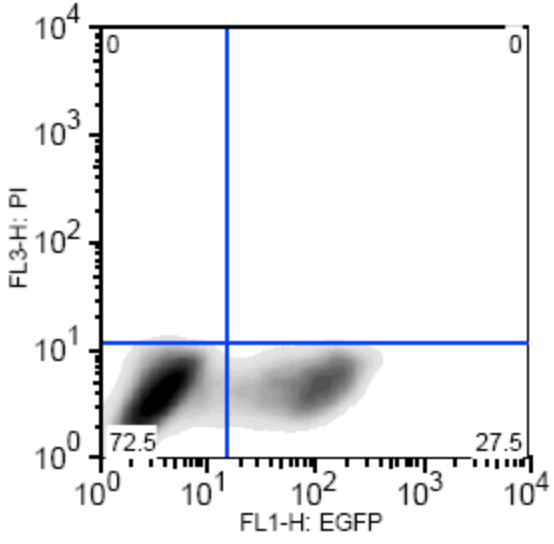
Non-marked cells

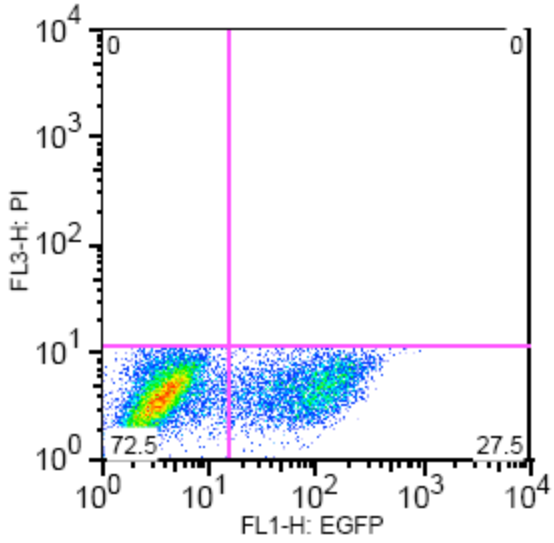
Marked cells (result)

Important note: Same gates as in neg. control apply automatically on the positive set!

Other forms of result visualization (FlowJo)







Proposal

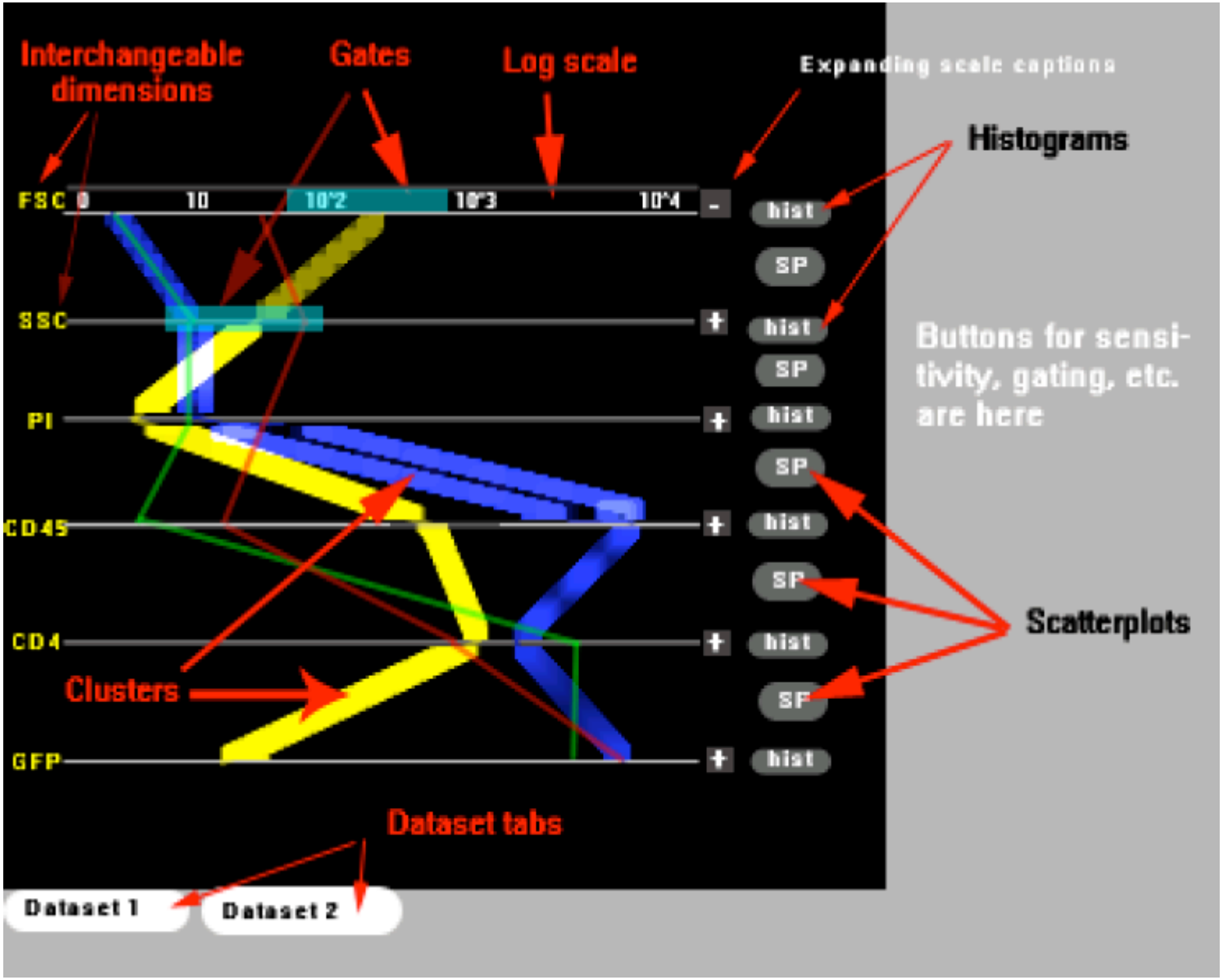
User requirements (based on user studies):

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

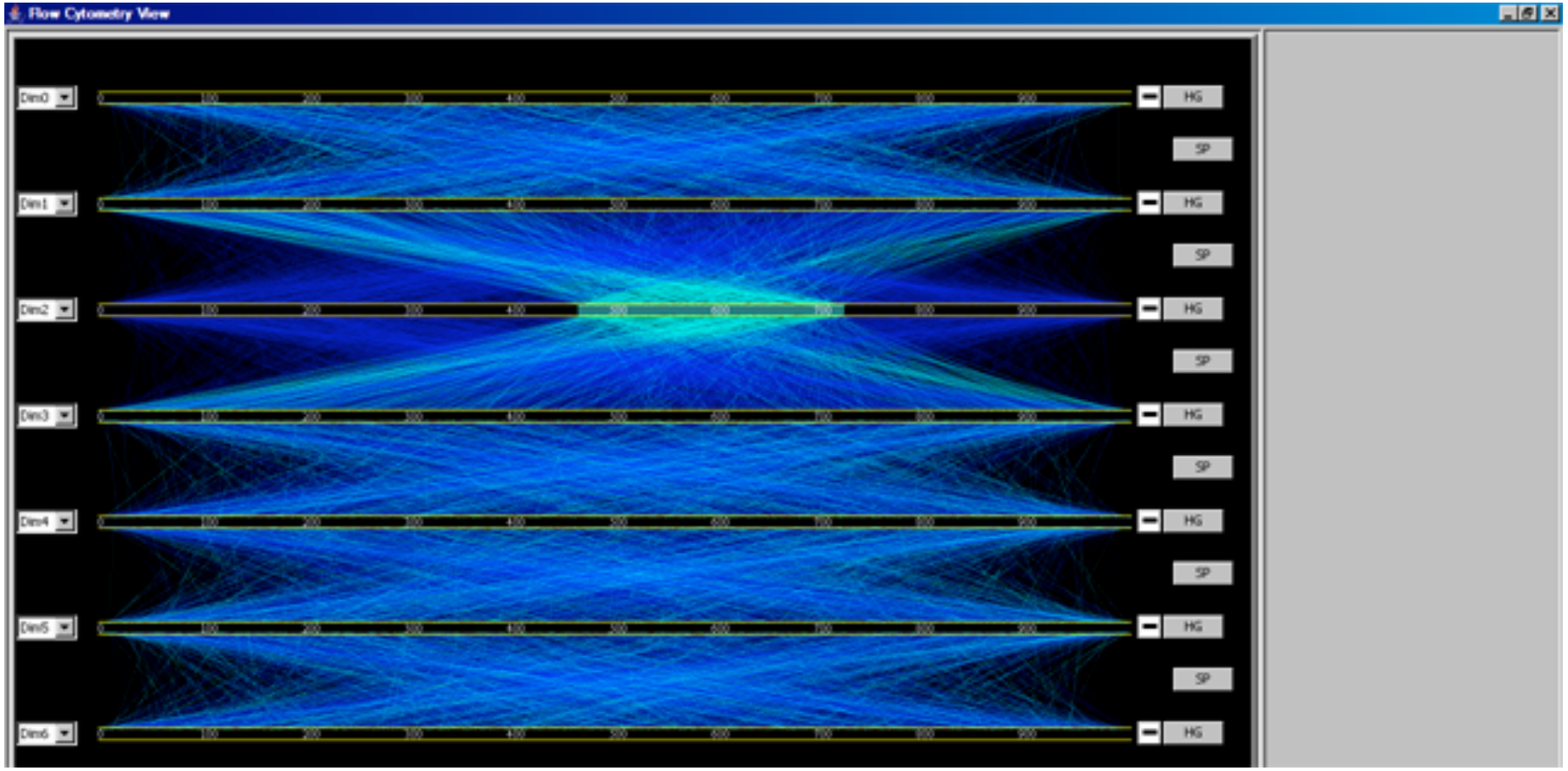
Solutions:

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

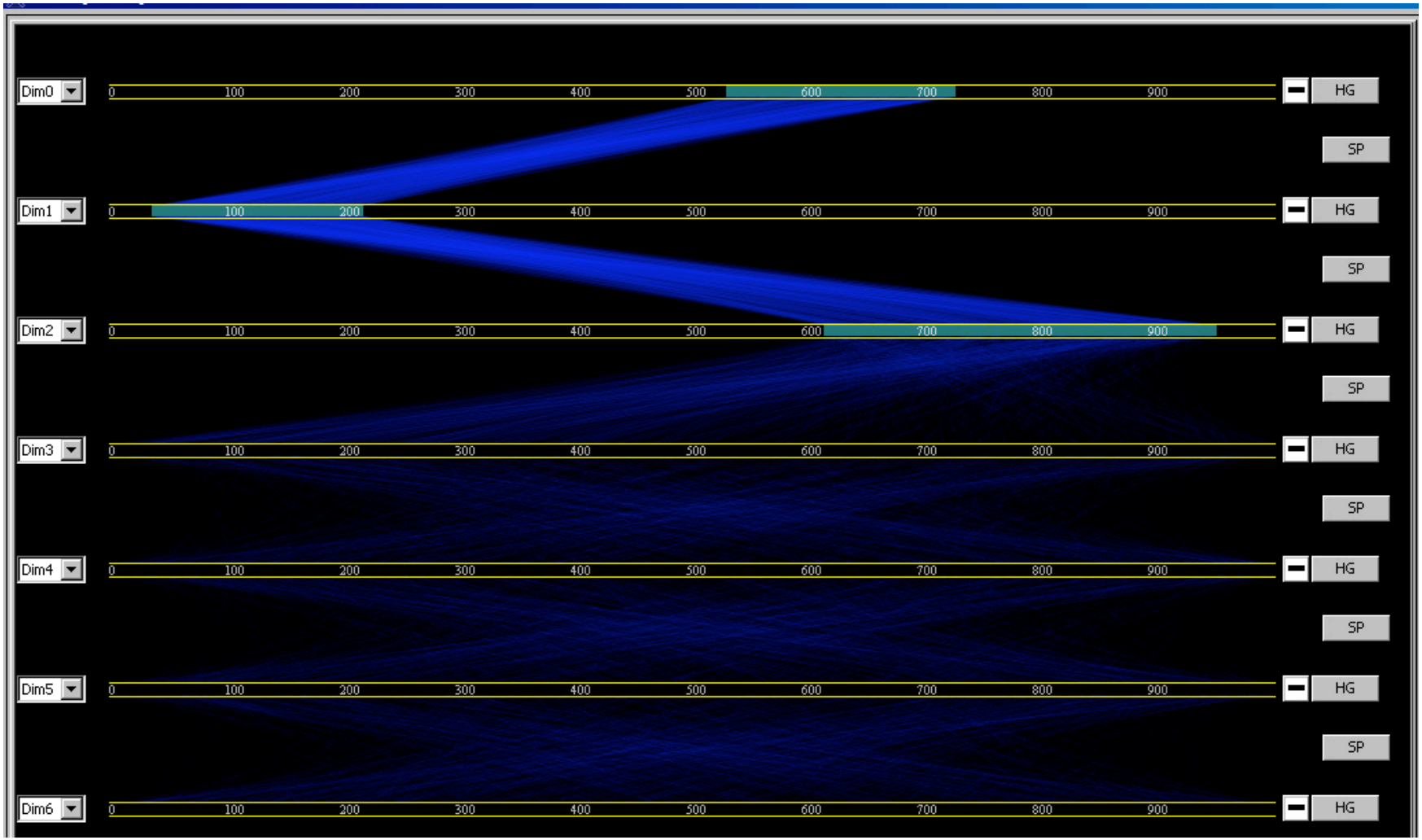


Prototype progress



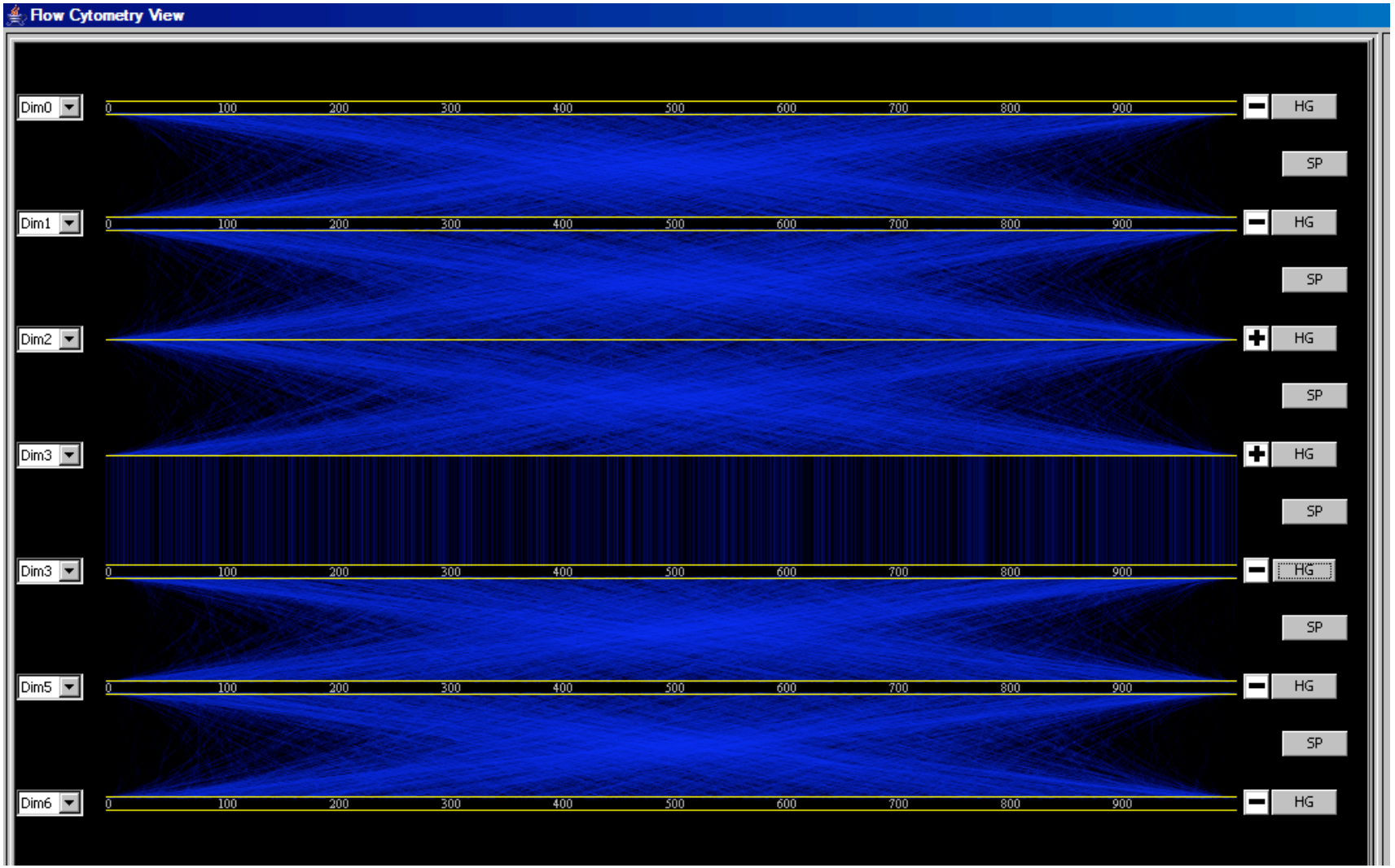
Highlighting of the gate. Random set, 3000 points, 7 dimensions.

Prototype progress



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

Prototype progress



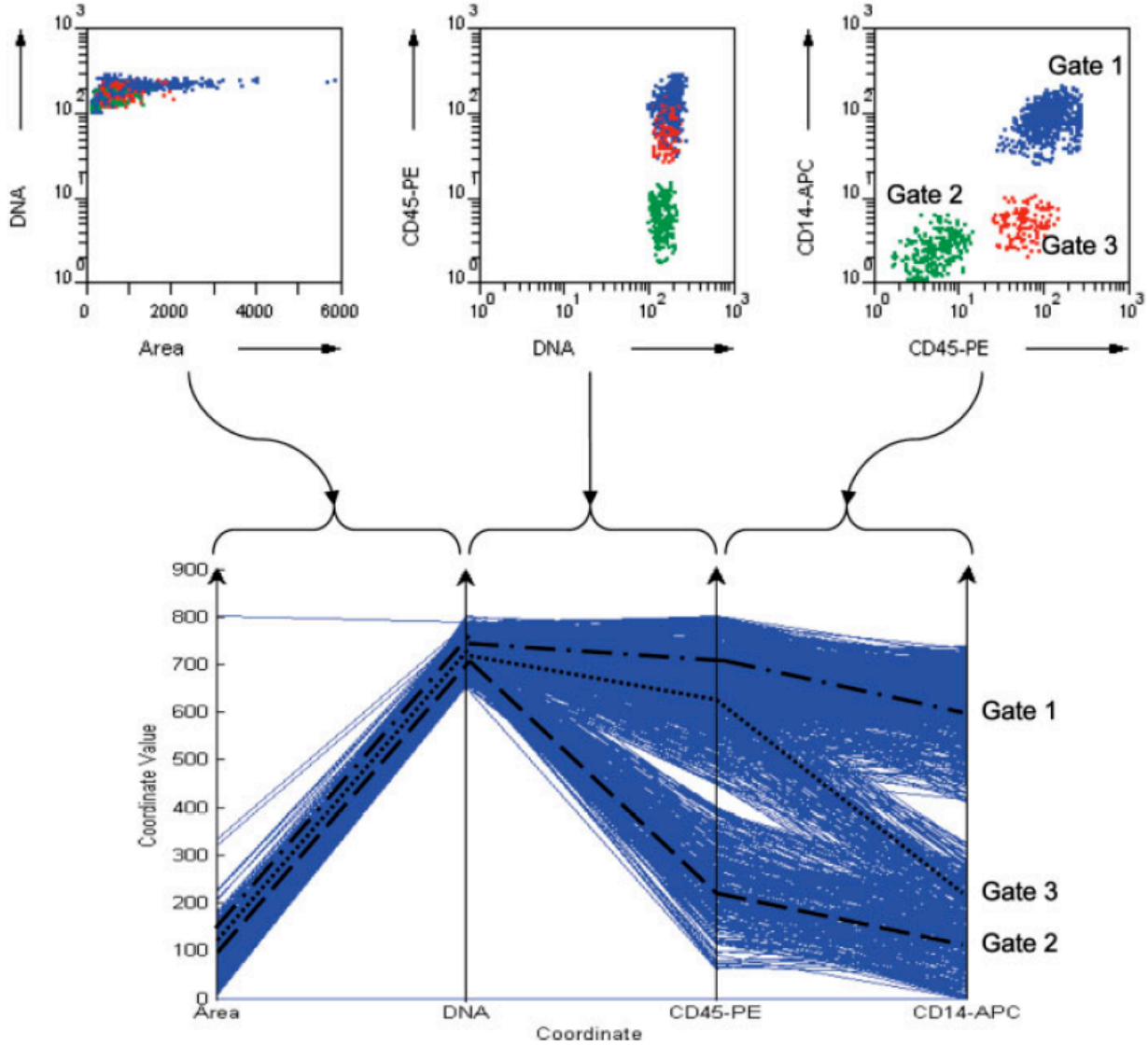
Interaction results. Random set, 3000 points, 7 dimensions.

Future Work

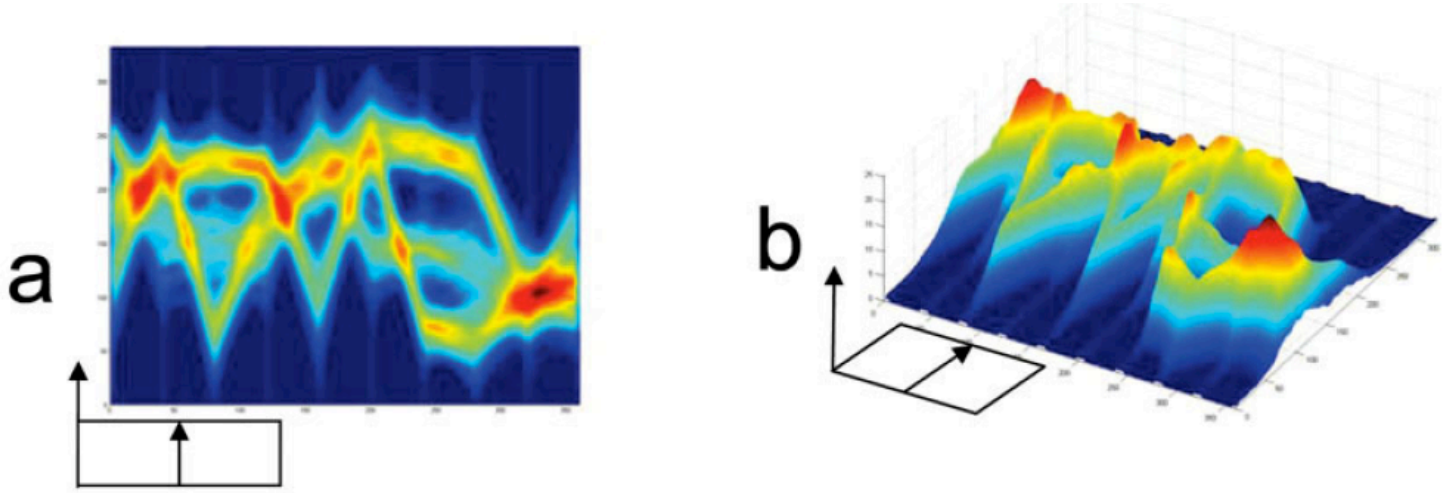
- Visualization of the real data
- Clustering
- Optimization
- User evaluation

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…