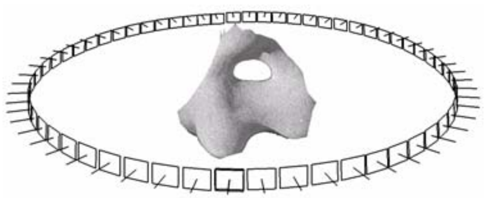
Recognition and 3D Reconstruction from Video

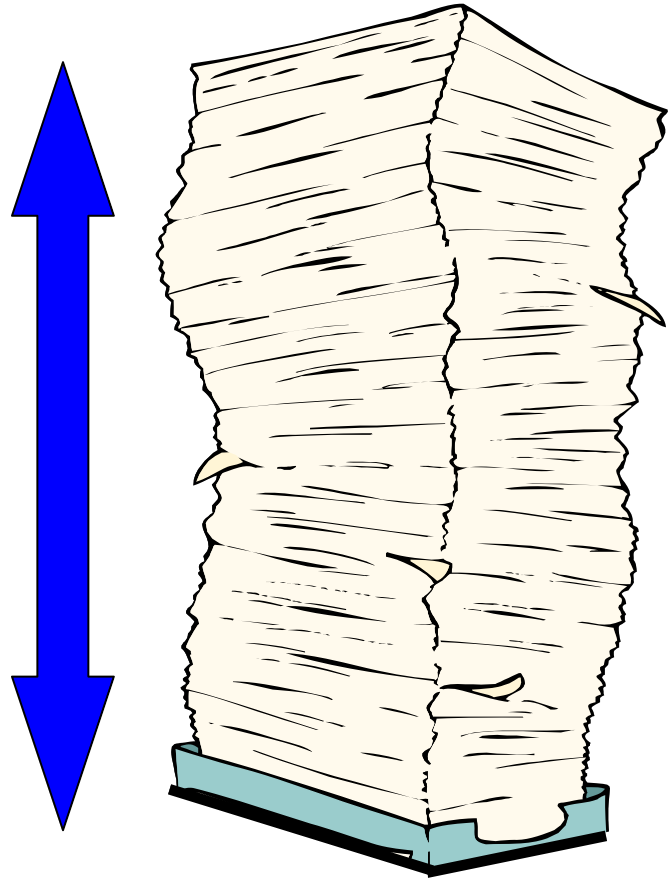
David Nistér



Center for Visualization & Virtual Environments

50 Thousand Images

4m



110,000,000 Images in 5.8 Seconds

UK

Center for Visualization & Virtual Environments

UK

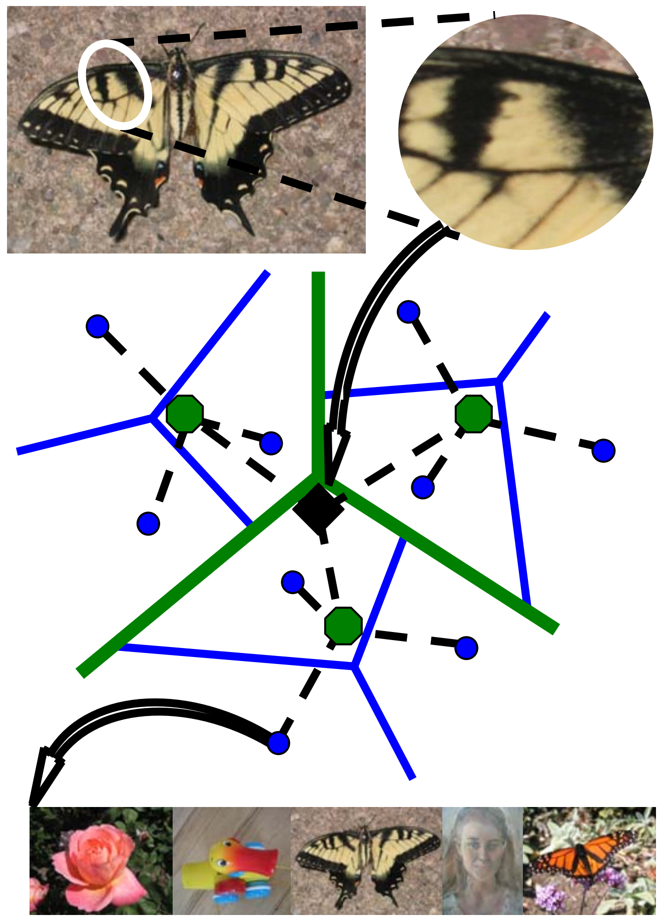
Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments

Scalable Recognition with a Vocabulary Tree

David Nistér, Henrik Stewénius



UK

Center for Visualization & Virtual Environments

New York City CVPR 2006

Towards Urban 3D Reconstruction From Video

A. Akbarzadeh, J.-M. Frahm, P. Mordohai, B. Clipp, C. Engels, D. Gallup, P. Merrell, M. Phelps, S. Sinha, B. Talton, L. Wang, Q. Yáng, H. Stewénius, R. Yang, G. Welch, H. Towles, D. Nistér and M. Pollefeys

UK

Center for Visualization & Virtual Environments

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UK

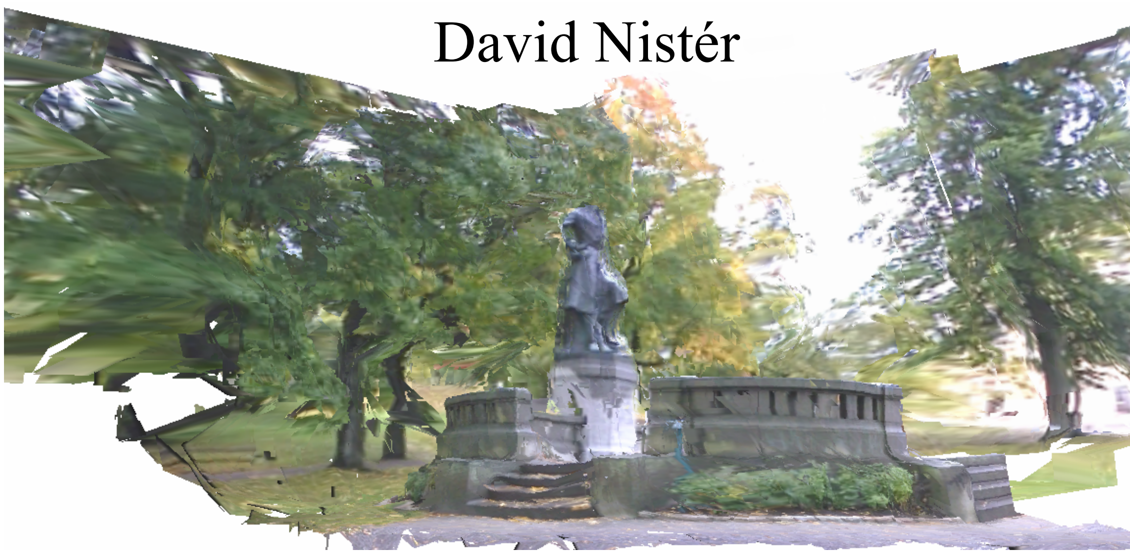
Center for Visualization & Virtual Environments

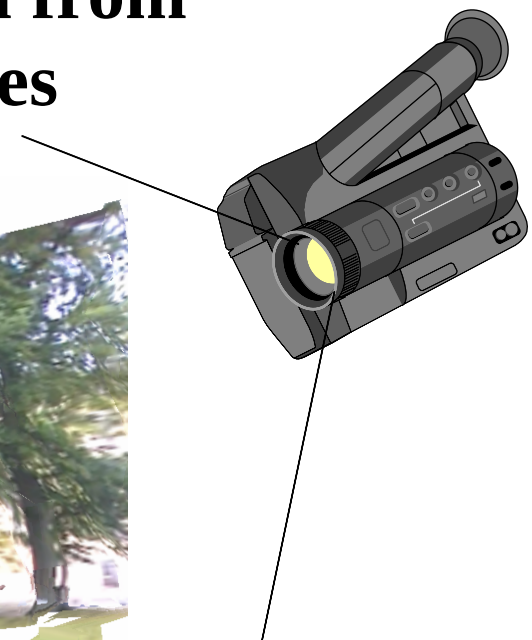
Kungl Tekniska Högskolan

Stockholms Universitet

Automatic Dense Reconstruction from Uncalibrated Video Sequences

David Nistér





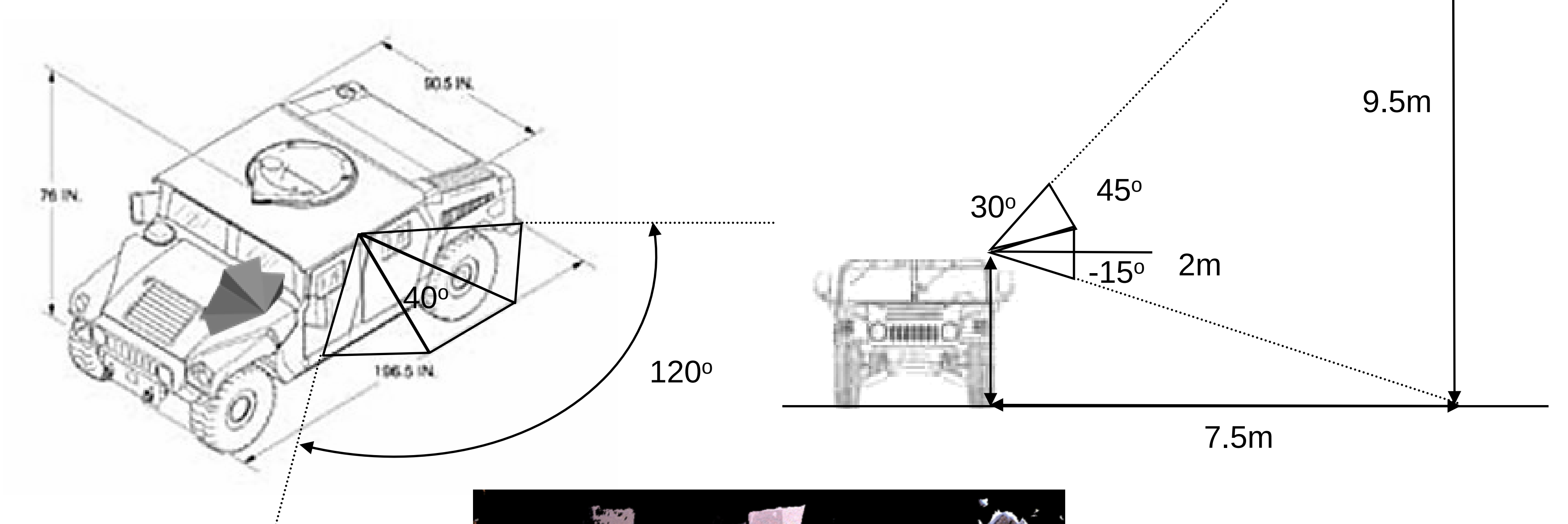
ERICSSON

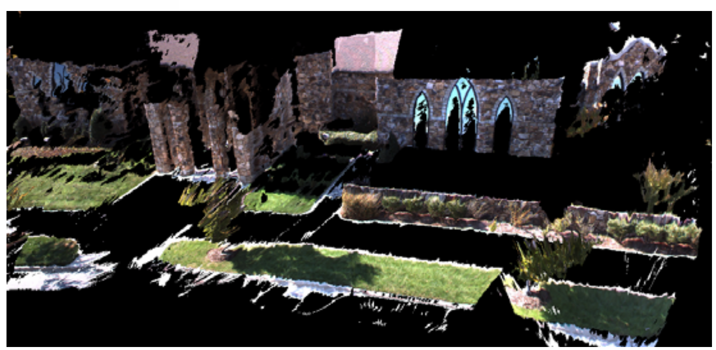
UK

Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments

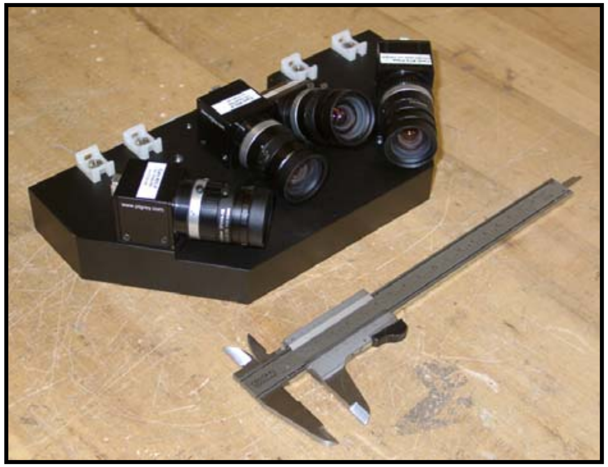




UK

Center for Visualization & Virtual Environments

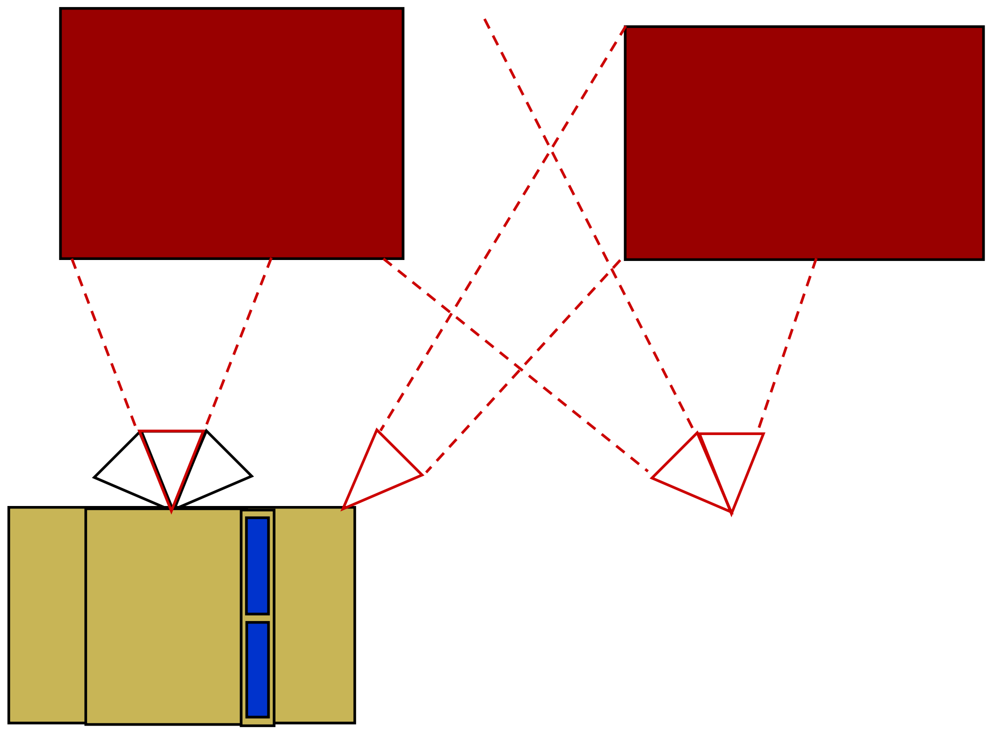
Video collection



2x4 cameras

1024x768@30Hz

UK Center for Visualization & Virtual Environments



Video Data

UK

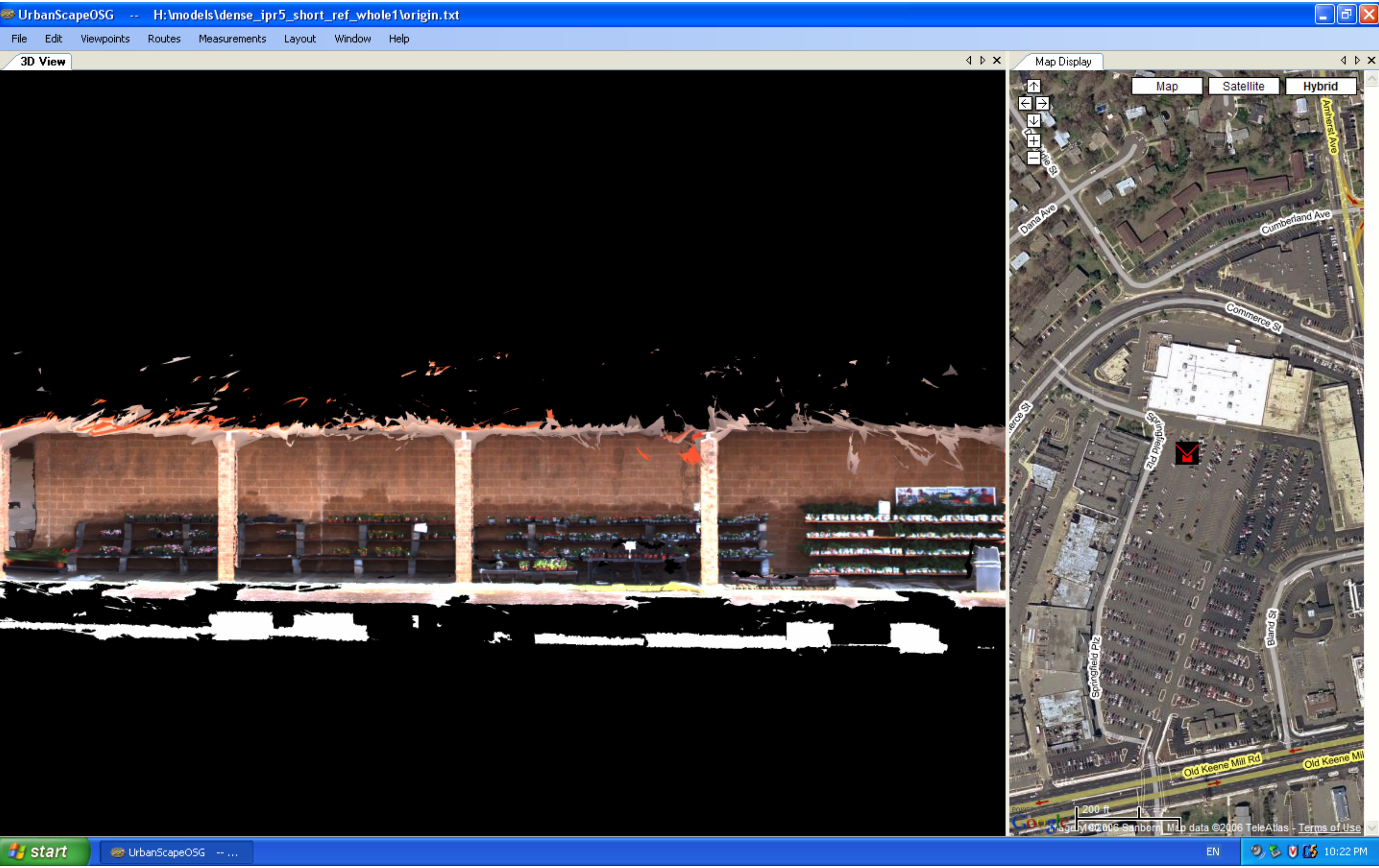
Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments



UK

Center for Visualization & Virtual Environments

Outline

- Feature Extraction and Description
- Matching, Tracking and Indexing
- Geometry
- Surface Reconstruction

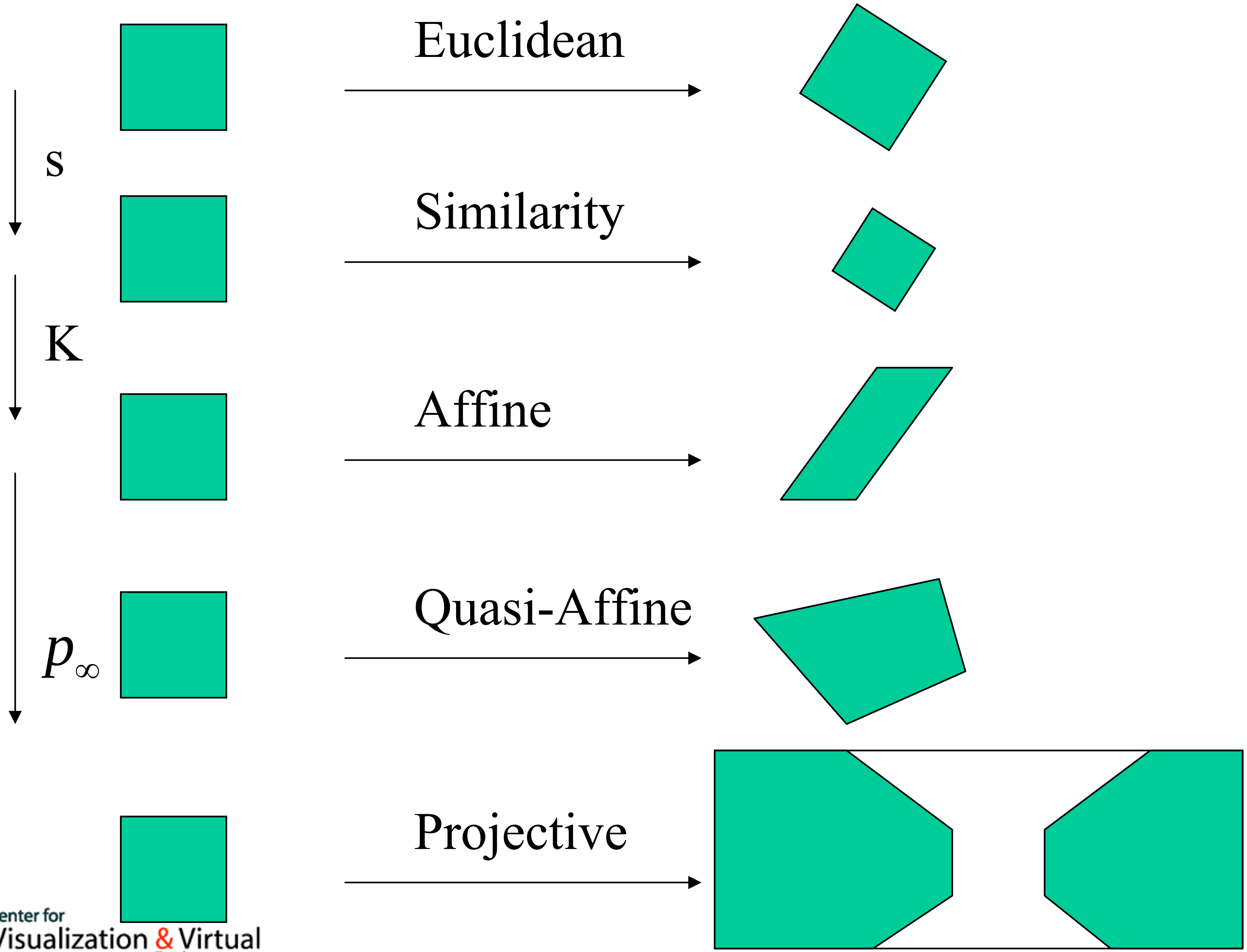
The transformation hierarchy

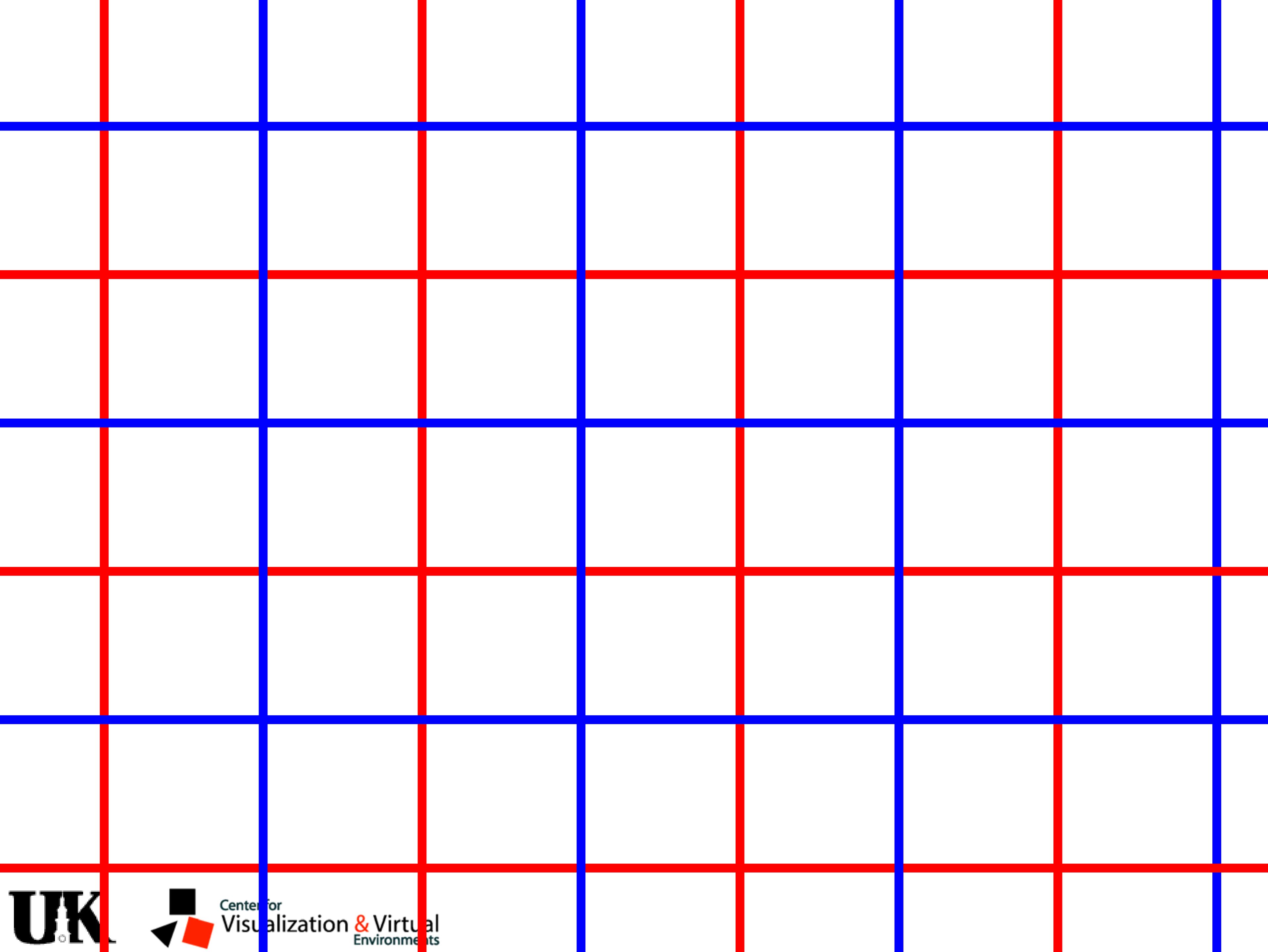
$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

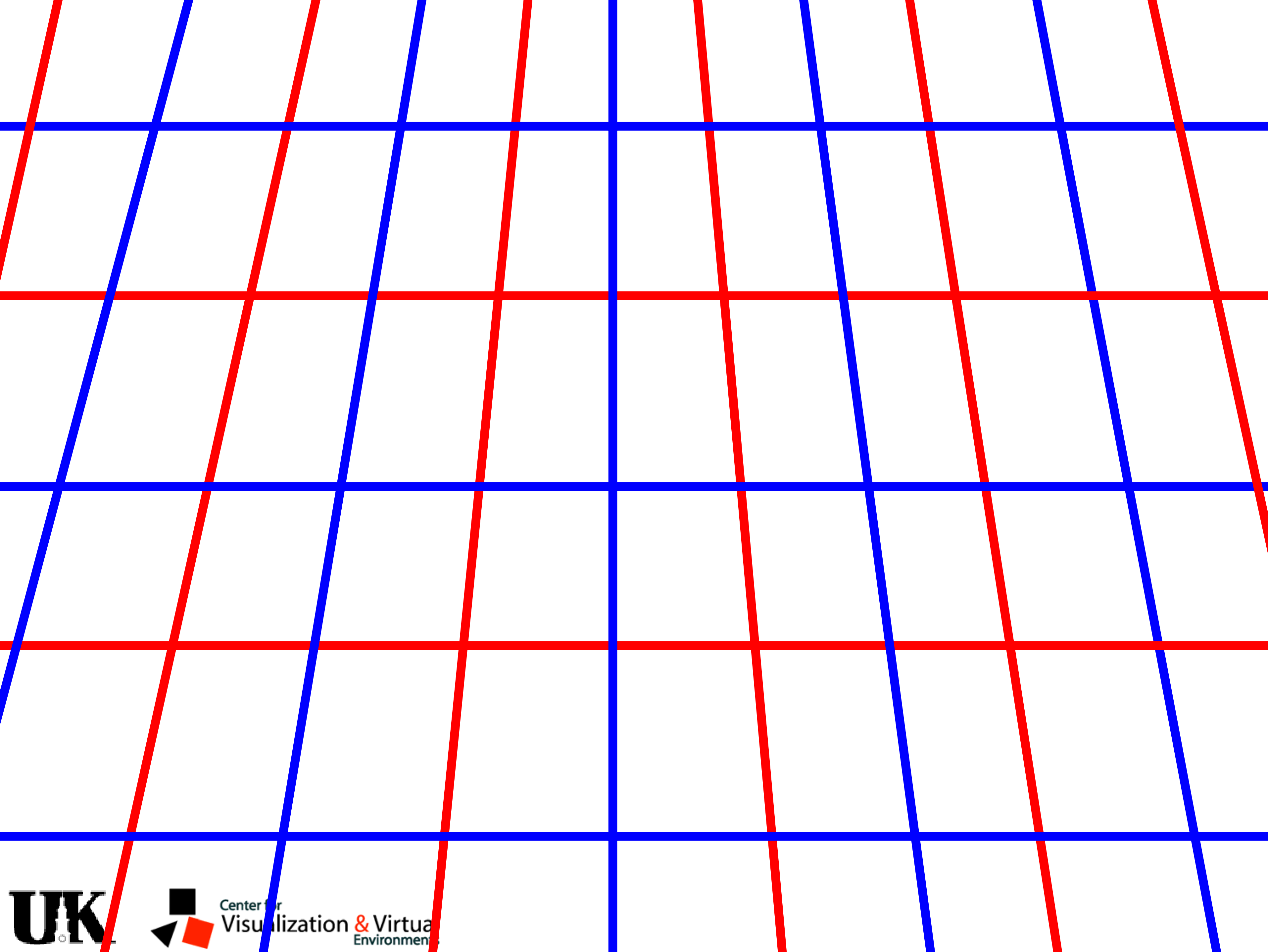
$$\begin{bmatrix} sR & t \\ 0 & 1 \end{bmatrix}$$

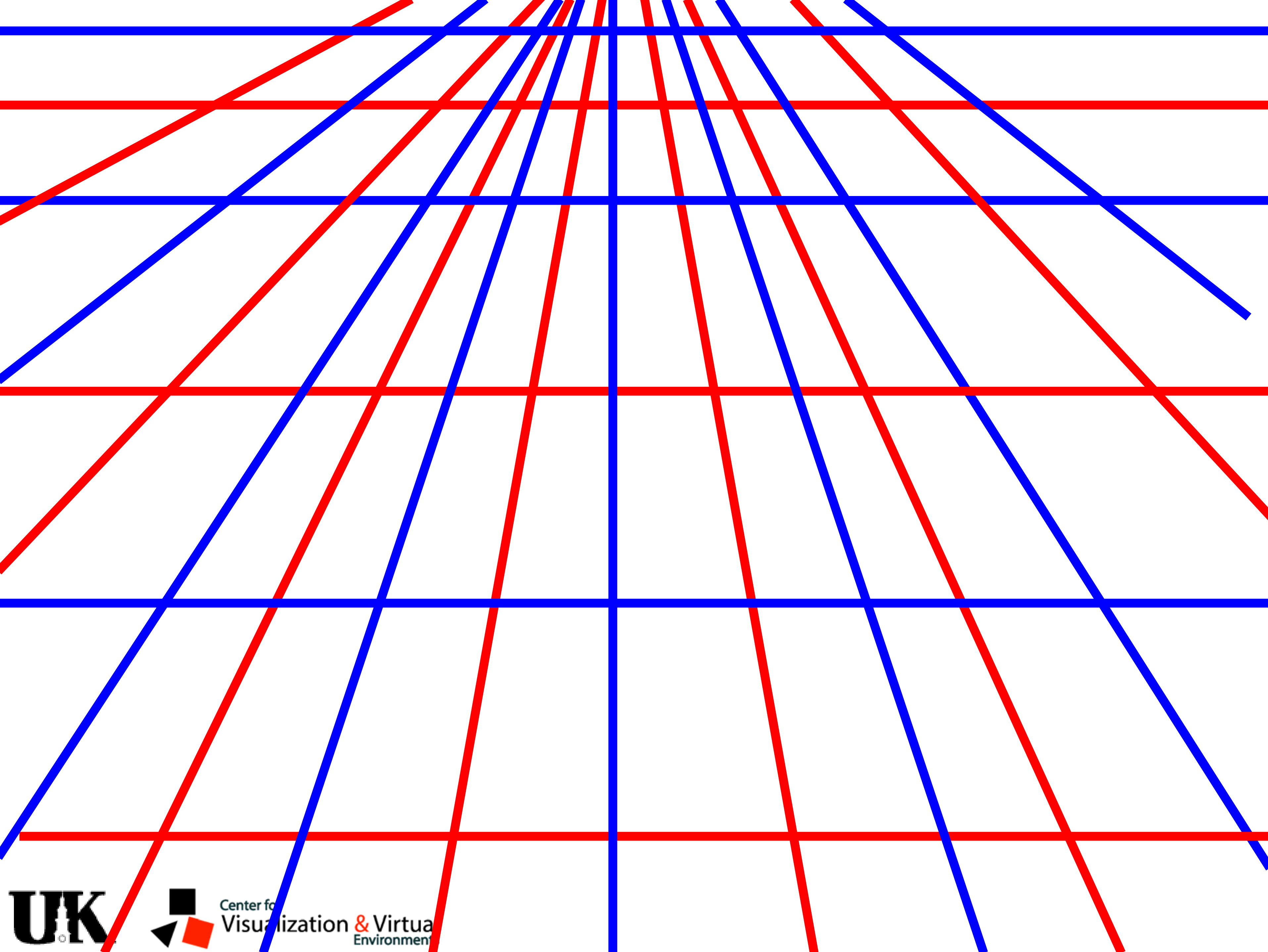
$$\begin{bmatrix} sKR & t \\ 0 & 1 \end{bmatrix}$$

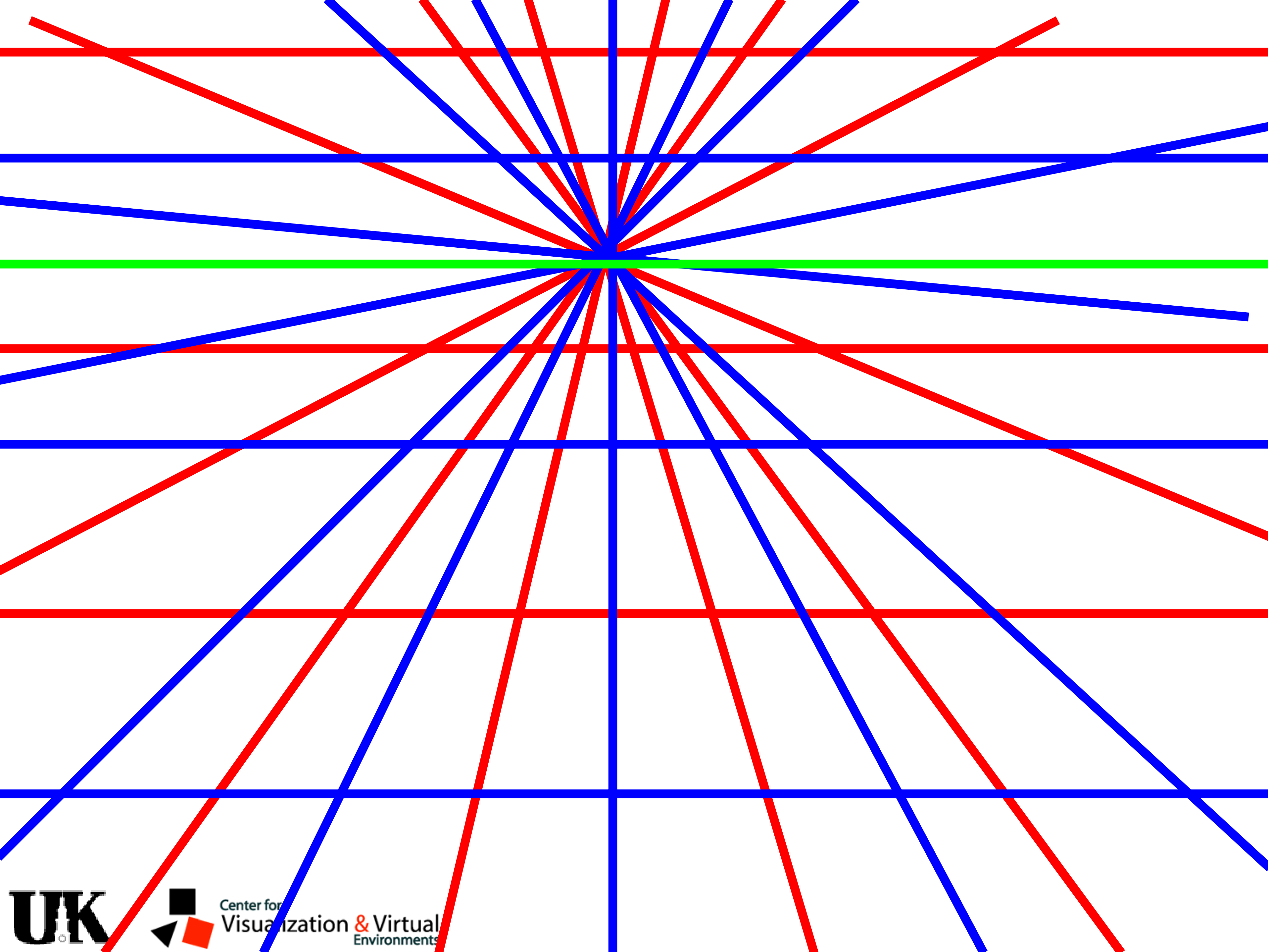
$$\begin{bmatrix} sKR & t \\ v & 1 \end{bmatrix}$$

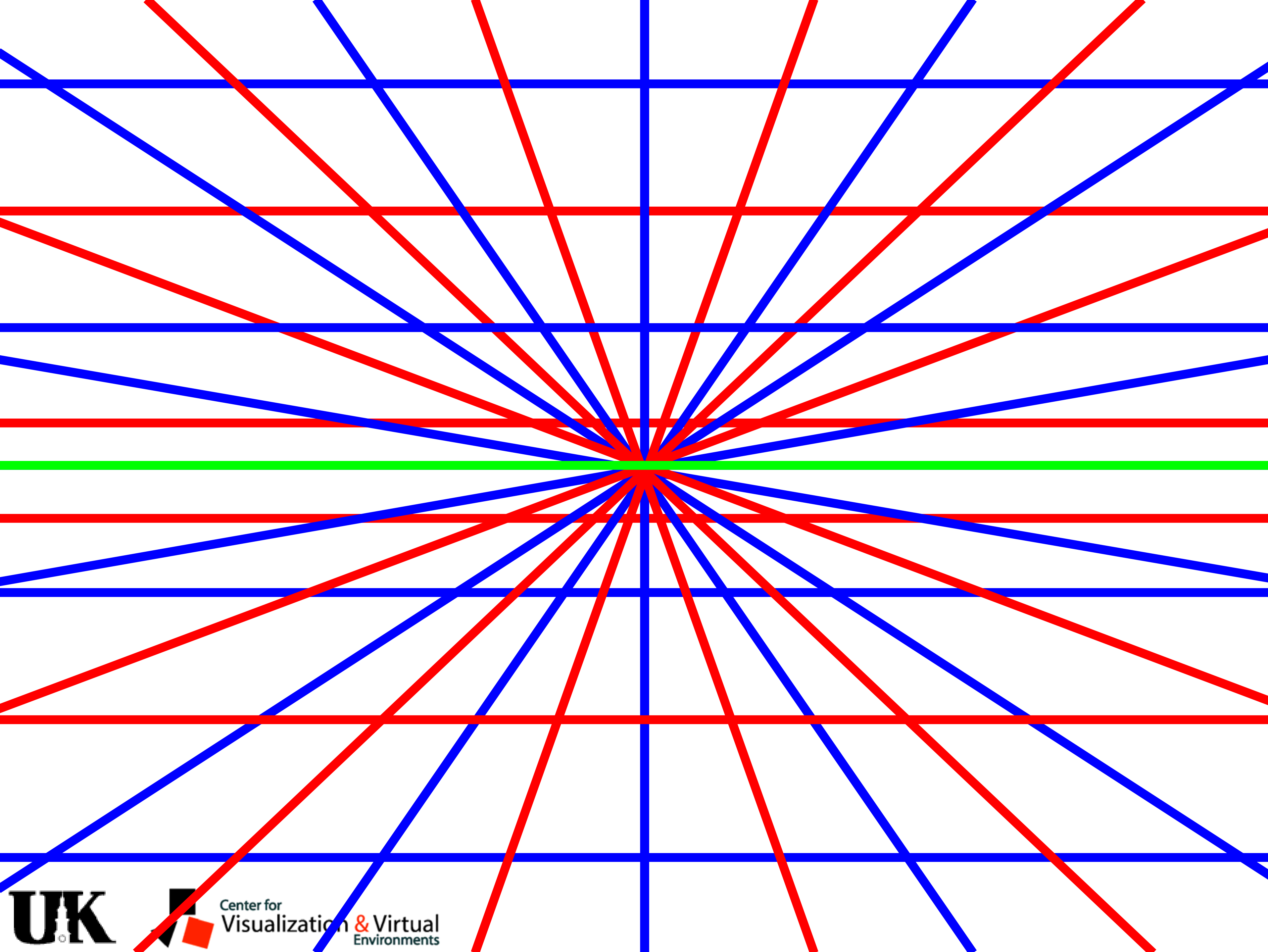


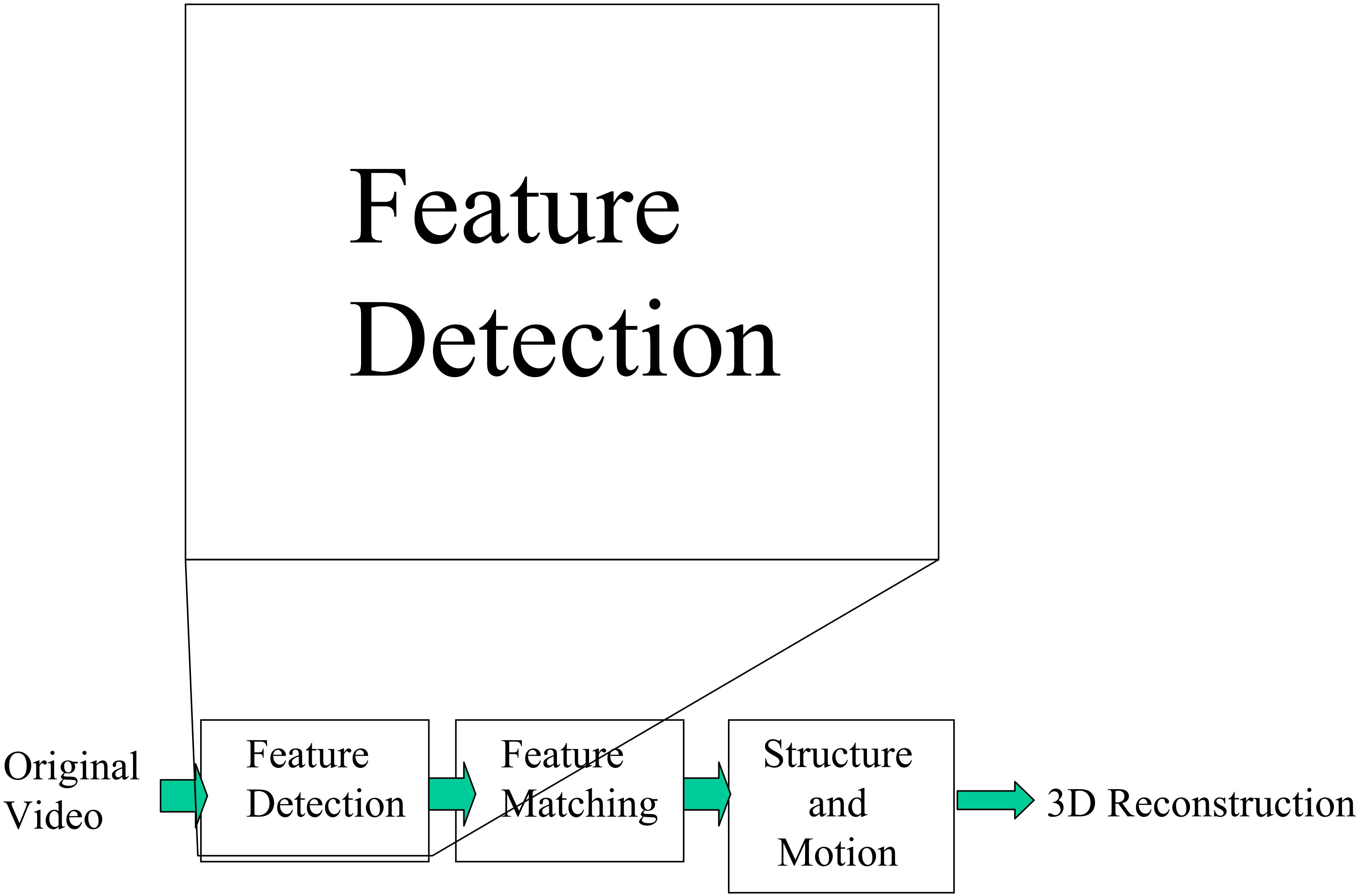

UK
Center for
Visualization & Virtual
Environments

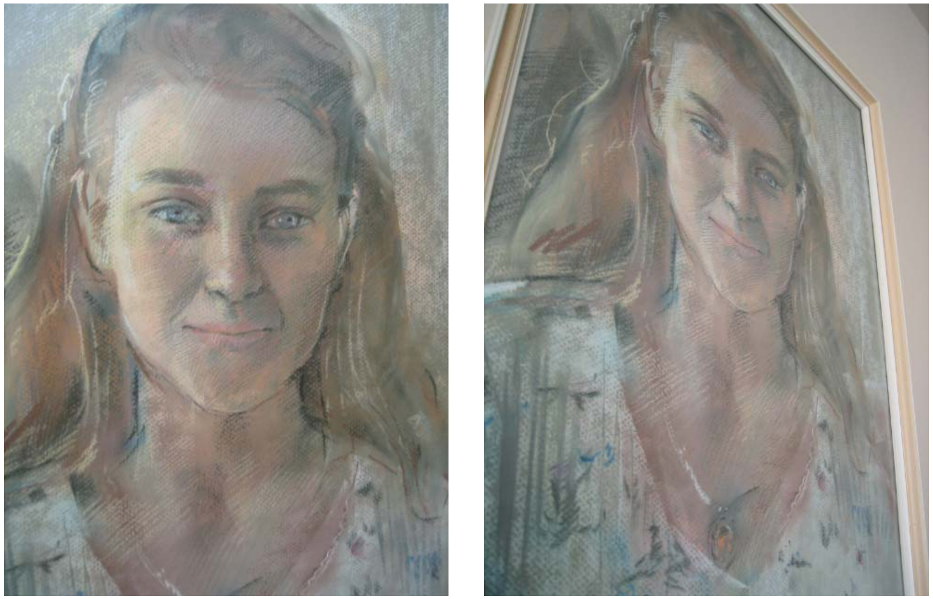

UK
Center f r
Visu lization & Virtua
Environmen s





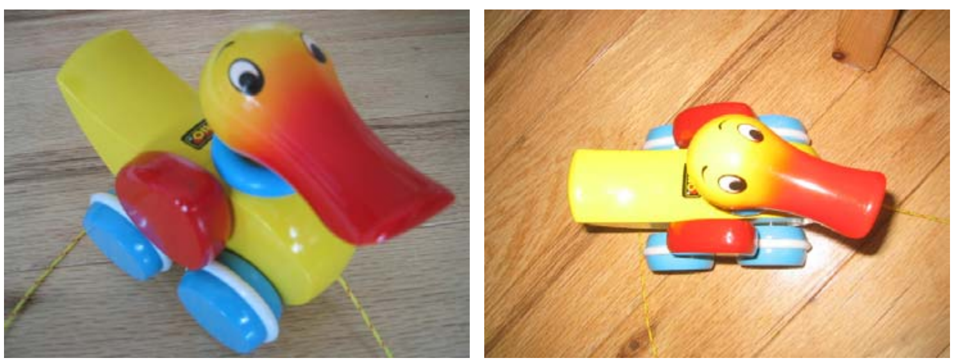

UK
Center for
Visualization & Virtual
Environments


Feature
Detection
Original
Video
Feature
Detection
Feature
Matching
Structure
and
Motion
3D Reconstruction



- Viewpoint Change

- Rotation



- Lighting Variation

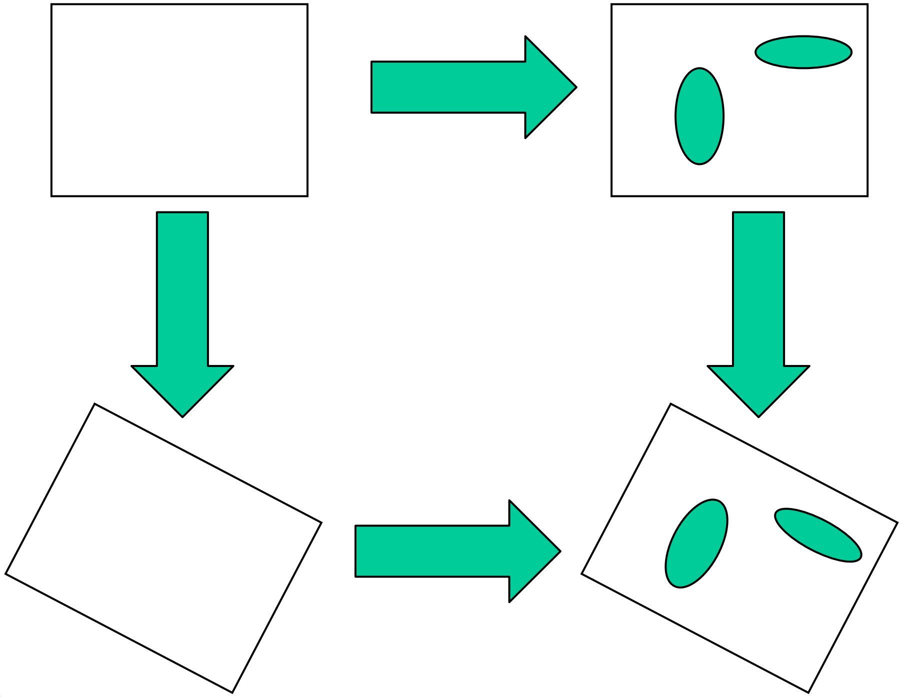
- Scale Change

UK Center for Visualization & Virtual Environments

Invariance or Covariance

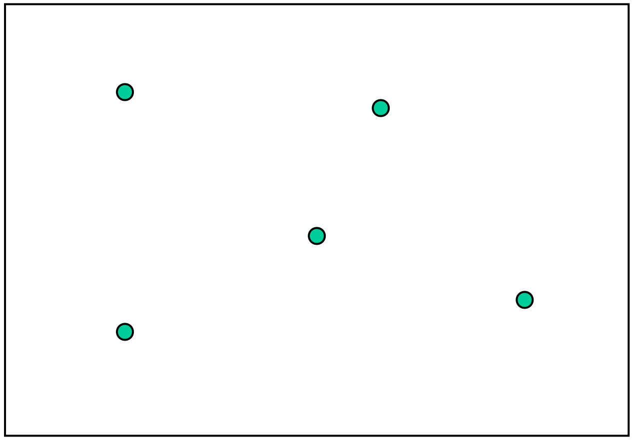
- Detection and image transformation commutes

Detect (Transform(I))=Transform(Detect(I))



Rotation-Invariant Detection

- Moravec
- Förstner
- Harris



Center for Visualization & Virtual Environments

Feature Detection

Harris Corners

Original Video → Feature Detection → Feature Matching → Structure and Motion → 3D Reconstruction

UK

Center for Visualization & Virtual Environments

Feature Detection

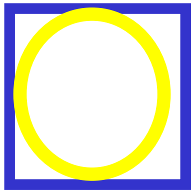
Harris Corners

Original Video → Feature Detection → Feature Matching → Structure and Motion → 3D Reconstruction

UK Center for Visualization & Virtual Environments

Feature Detection

Harris Corners

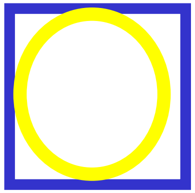


Autocorrelation

$$A(d) = \sum_{x} \left(I(x) - I(x+d) \right)^2$$

Feature Detection

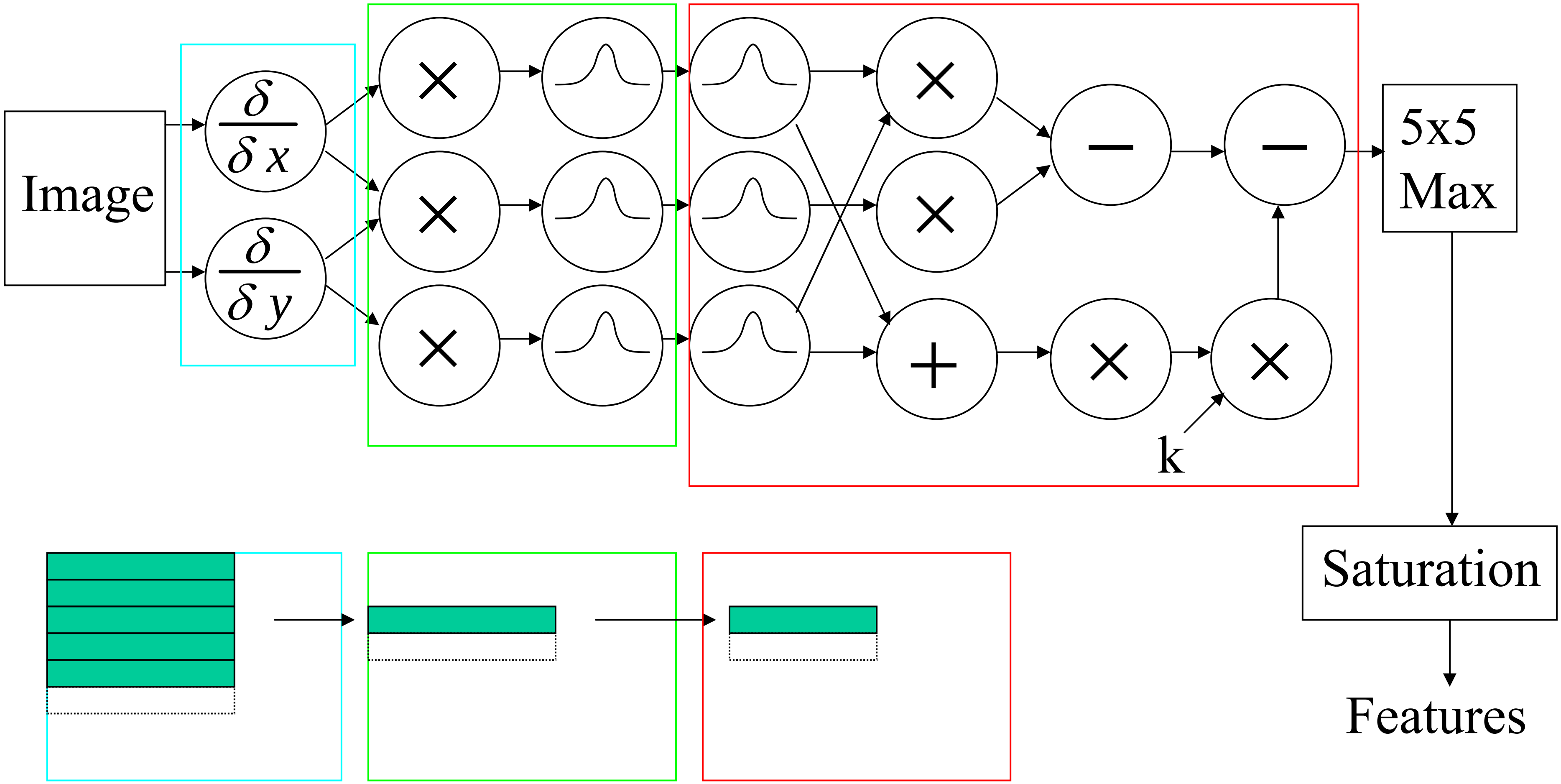
Harris Corners



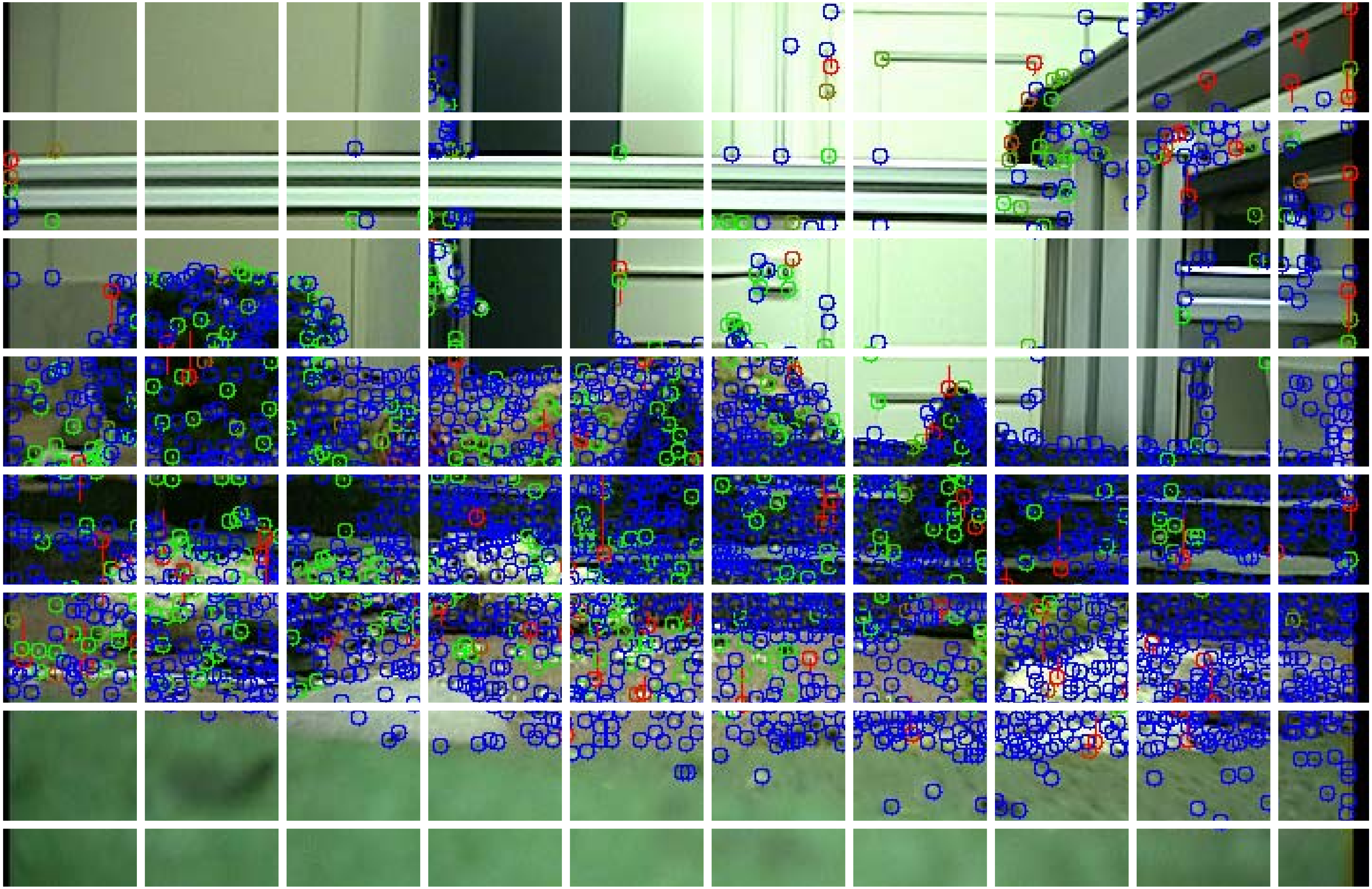
Second Moment Matrix

$$M = \sum_x \begin{bmatrix} \dfrac{\partial^2 I}{\partial x^2} & \dfrac{\partial^2 I}{\partial x \partial y} \\ \dfrac{\partial^2 I}{\partial x \partial y} & \dfrac{\partial^2 I}{\partial y^2} \end{bmatrix}$$

Feature Detection

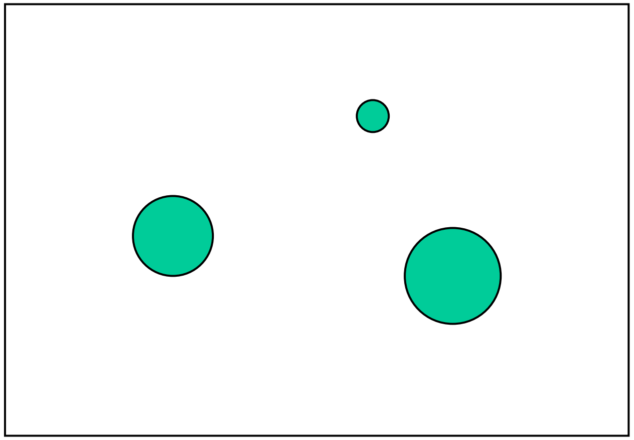


Feature Detection

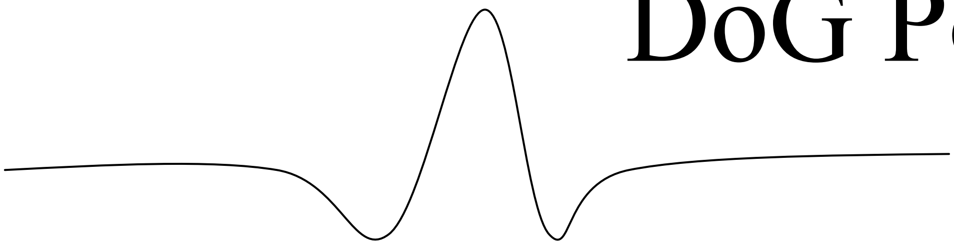


Rotation+Scale Invariant Detection

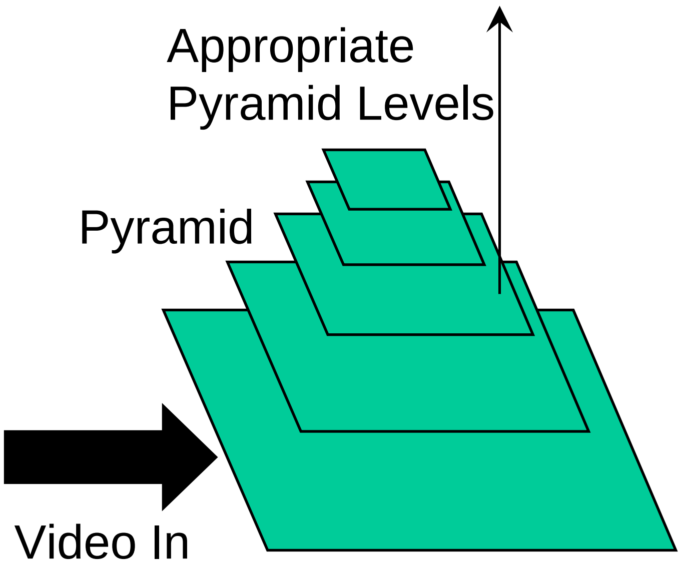
- DoG Points
- Lindeberg, Schmid & Mohr, Lowe

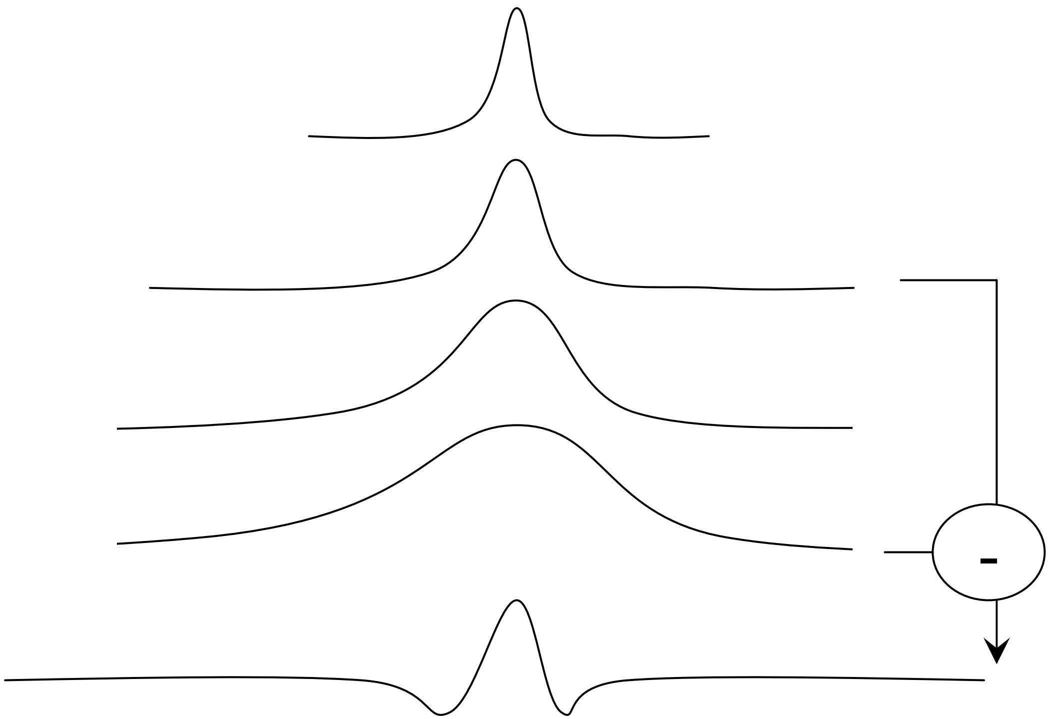


DoG Points



- ‘Blob’ detector

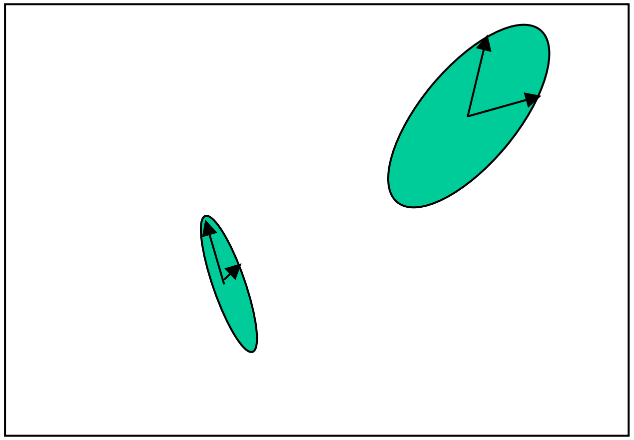




Center for Visualization & Virtual Environments

Affine Invariant Regions

- Tuytelaars & Van Gool
- Mikolajczyk and Schmid
- Matas et al.



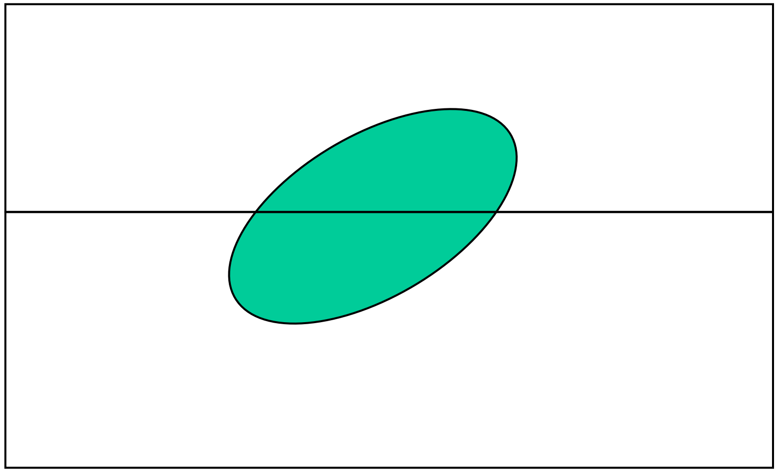
Center for Visualization & Virtual Environments

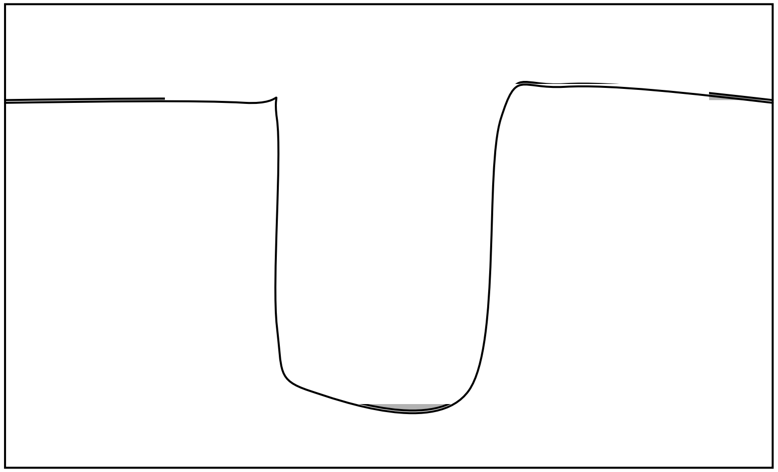
Harris and Hessian Affine

- Mikolajczyk and Schmid

MSER

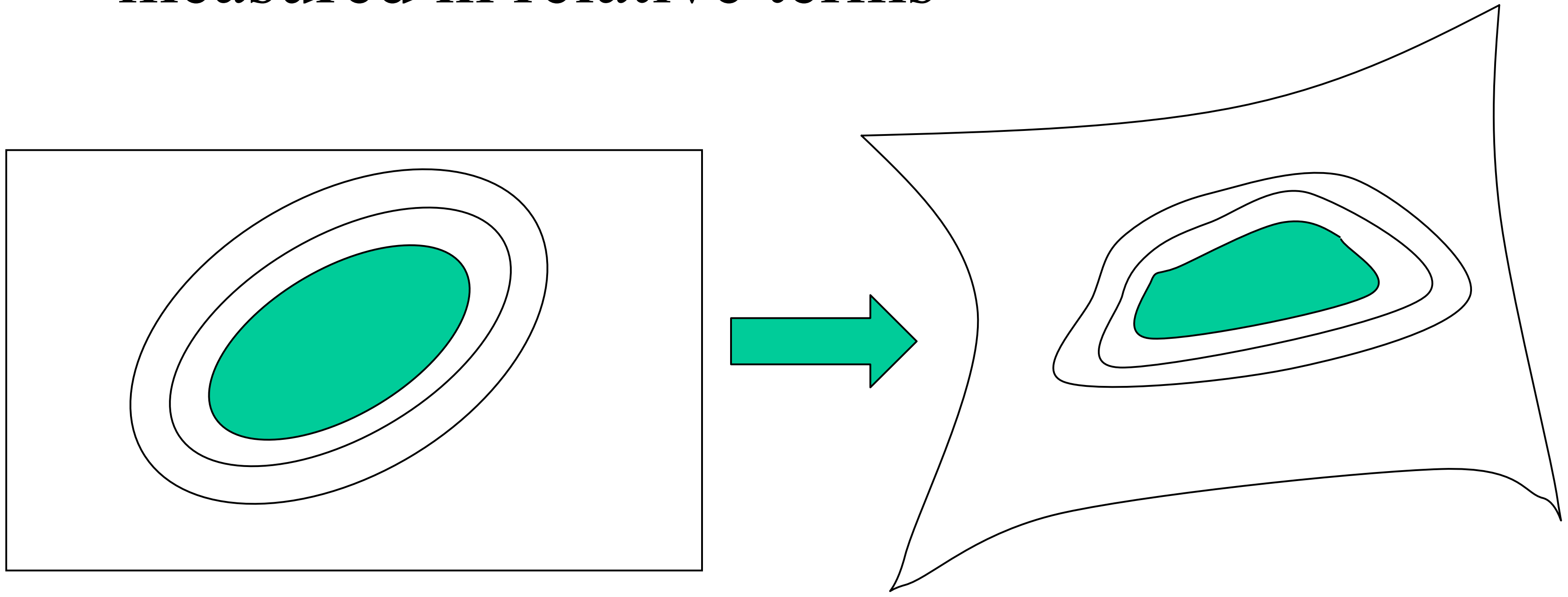
- Matas et al.
- Similar to watershed, but thresholded at minimal change rather than segmented when watersheds join



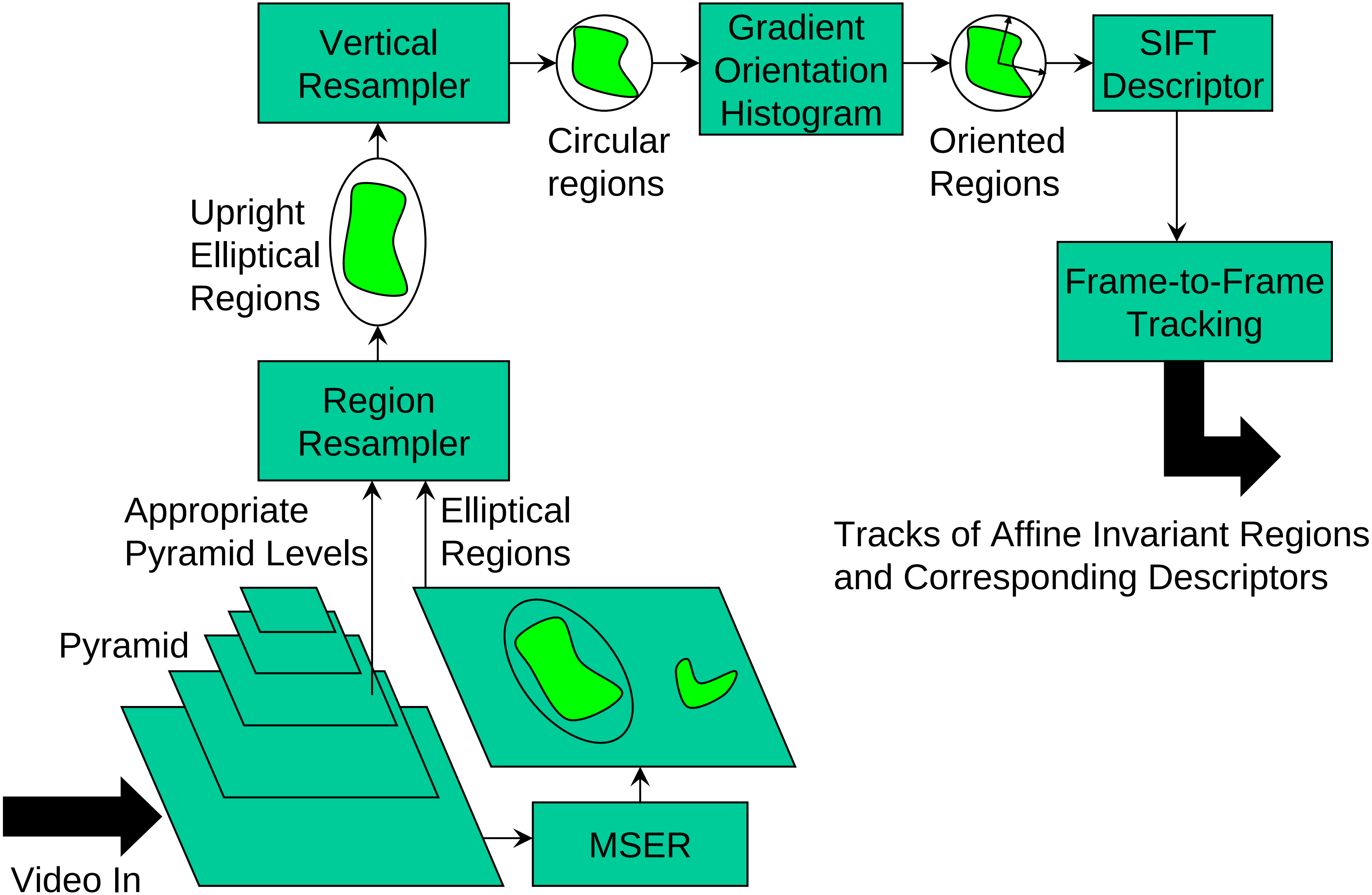


MSER

- Extremal regions are 'continuous-invariant'
- MSER's are affine invariant if growth is measured in relative terms



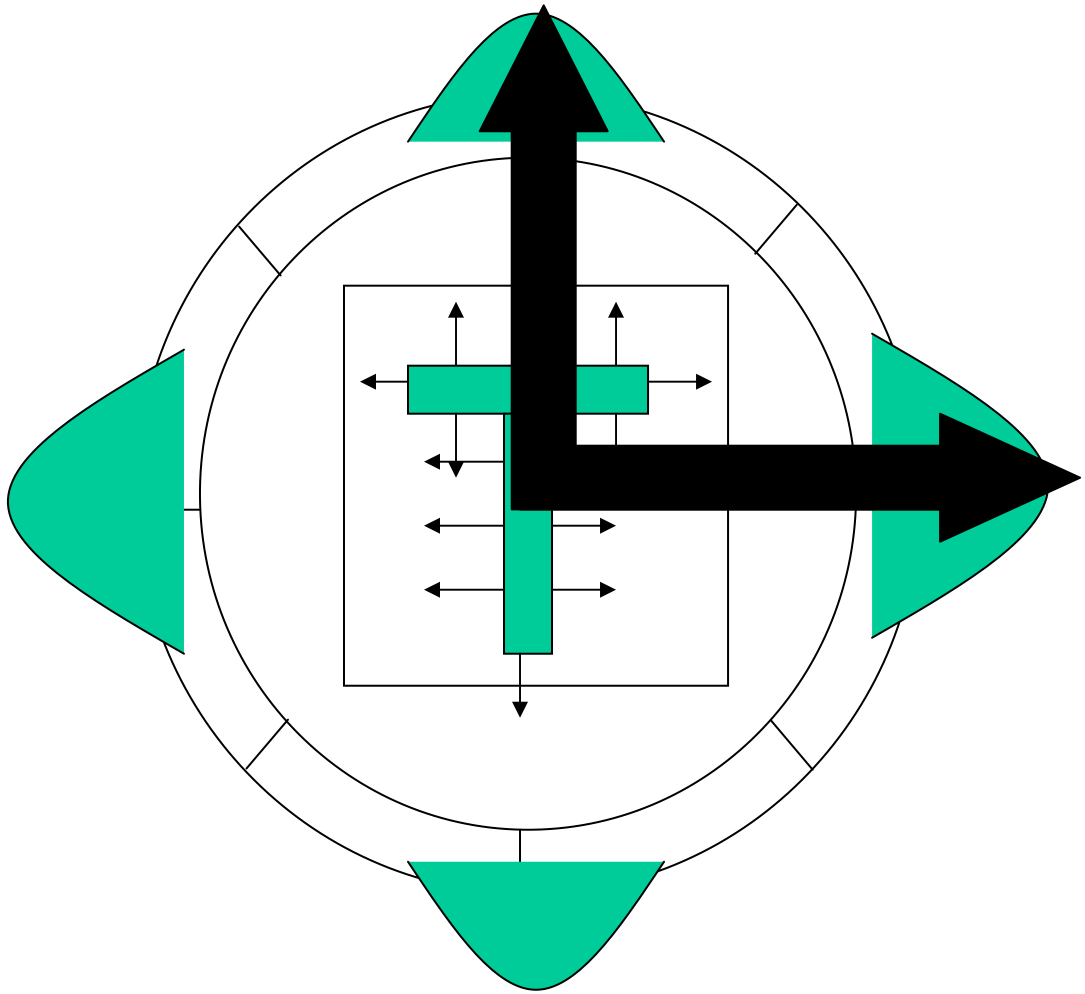
Demonstration of live feature tracking and MSER's



UK

Center for Visualization & Virtual Environments

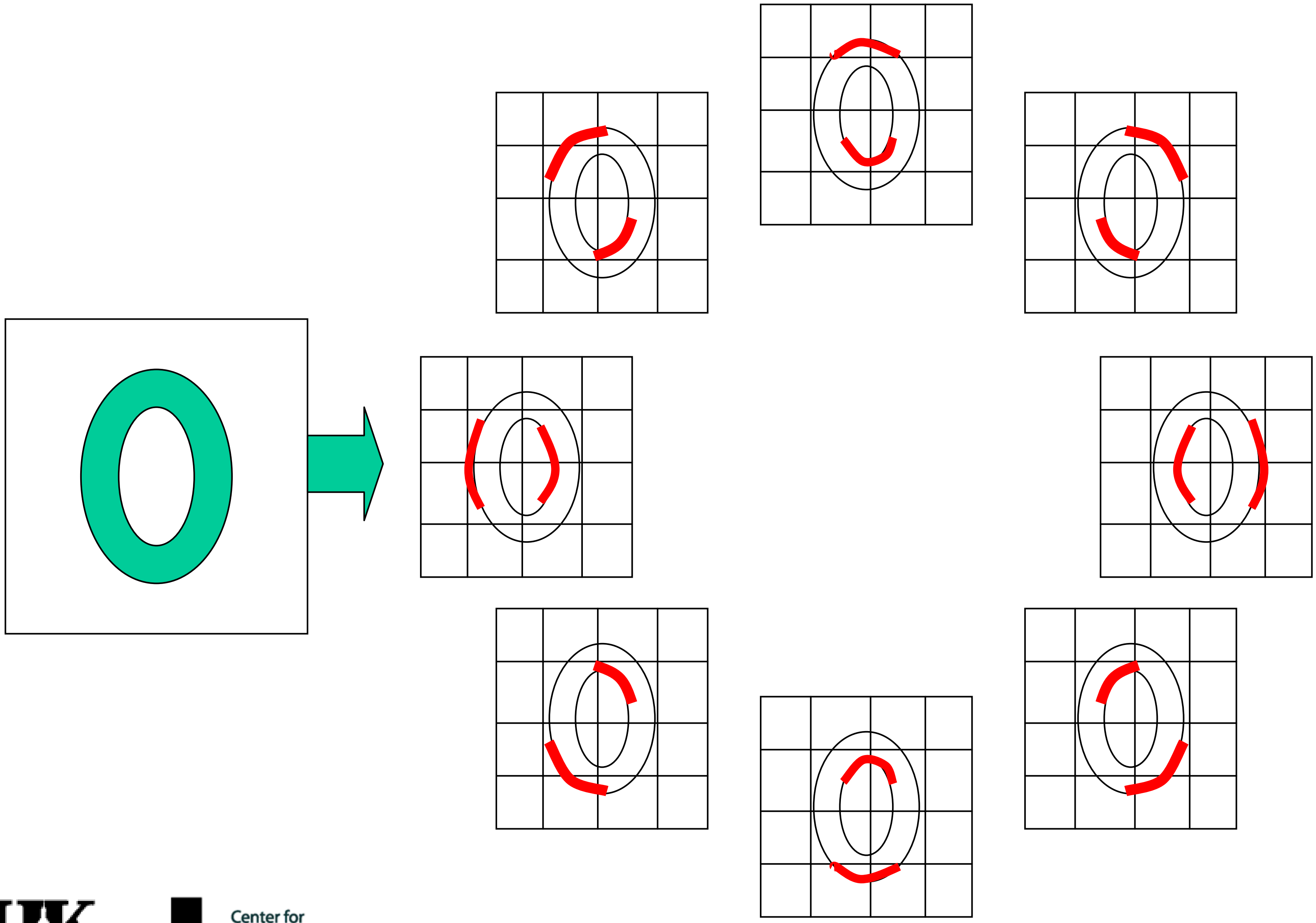
Selecting a coordinate system



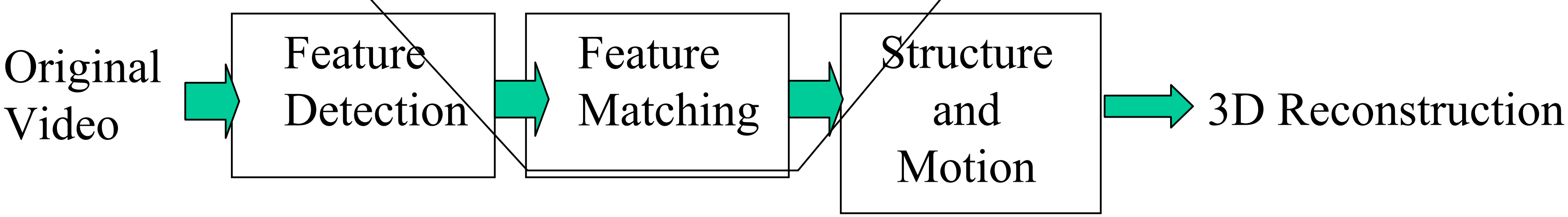
Region Description

- Image Patch
- Normalized Image Patch
- SIFT Descriptor
- DCT Descriptors
- Wavelets

SIFT Descriptor



Feature Matching



2D Tracking

KLT

Harris

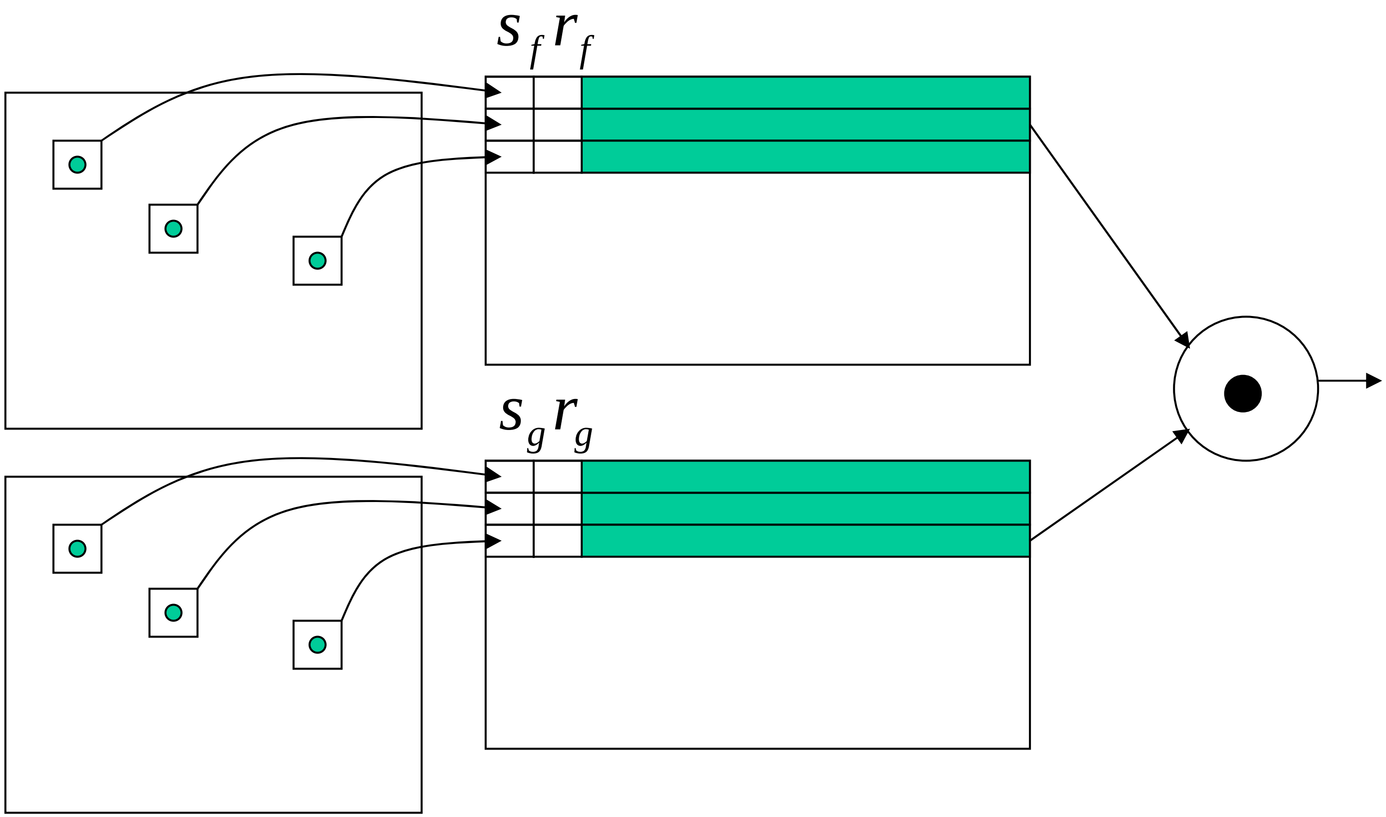
HC

UK Center for Visualization & Virtual Environments

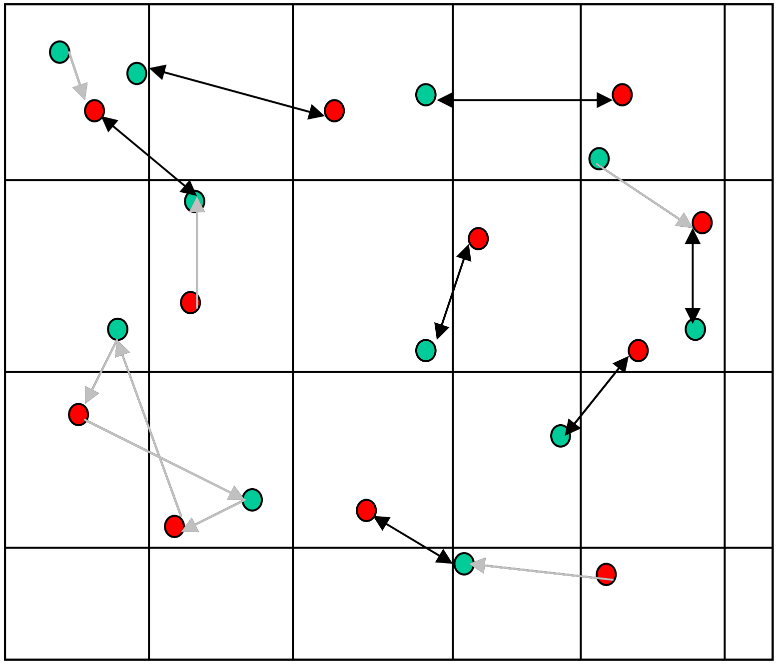
Feature Matching/Tracking

Normalized Correlation

$$\frac{\sum fg - \sum f \sum g}{\sqrt{\sum f^2 - \left(\sum f\right)^2}\sqrt{\sum g^2 - \left(\sum g\right)^2}} = \left(\sum fg - s_f s_g\right) * r_f r_g$$

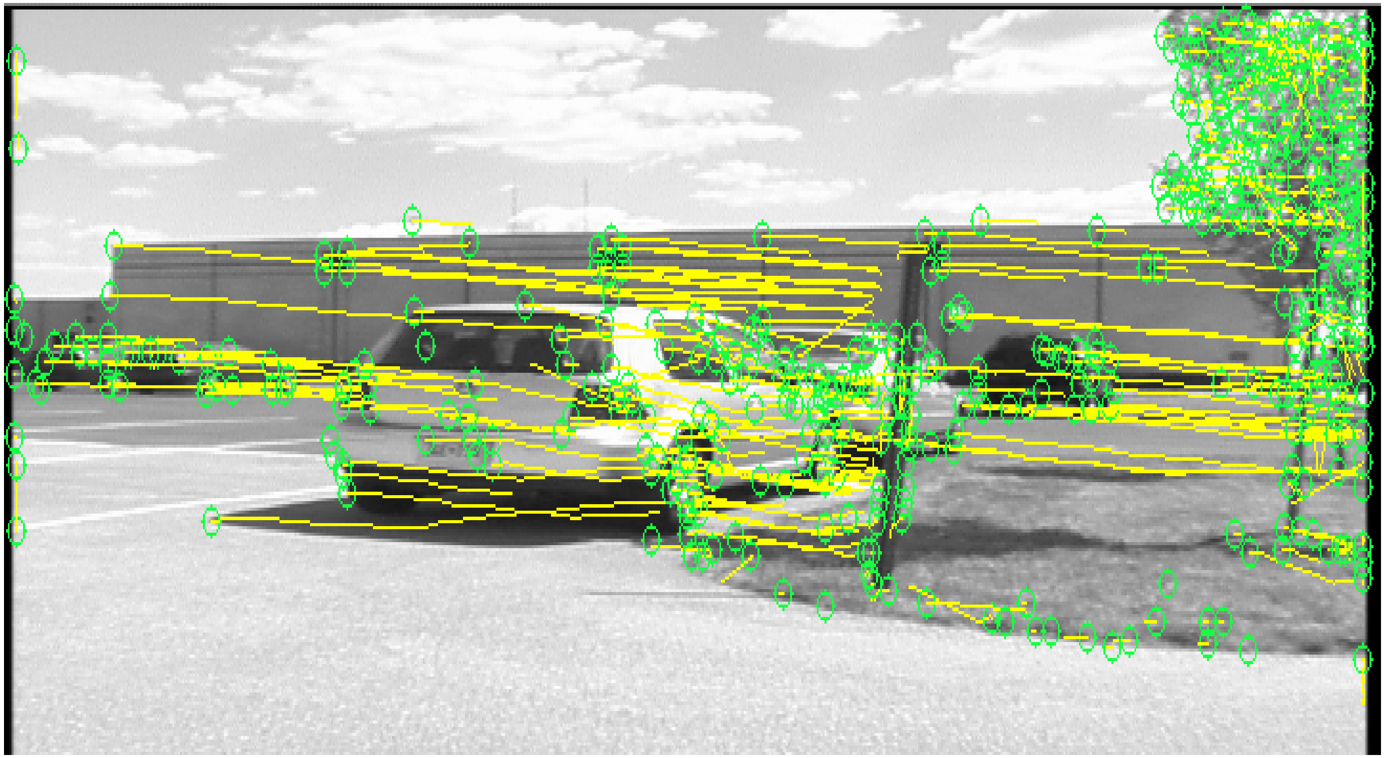


Feature Matching/Tracking



Only retain bidirectional matches
No loops because of symmetry d(a,b)=d(b,a)

Feature Matching/Tracking



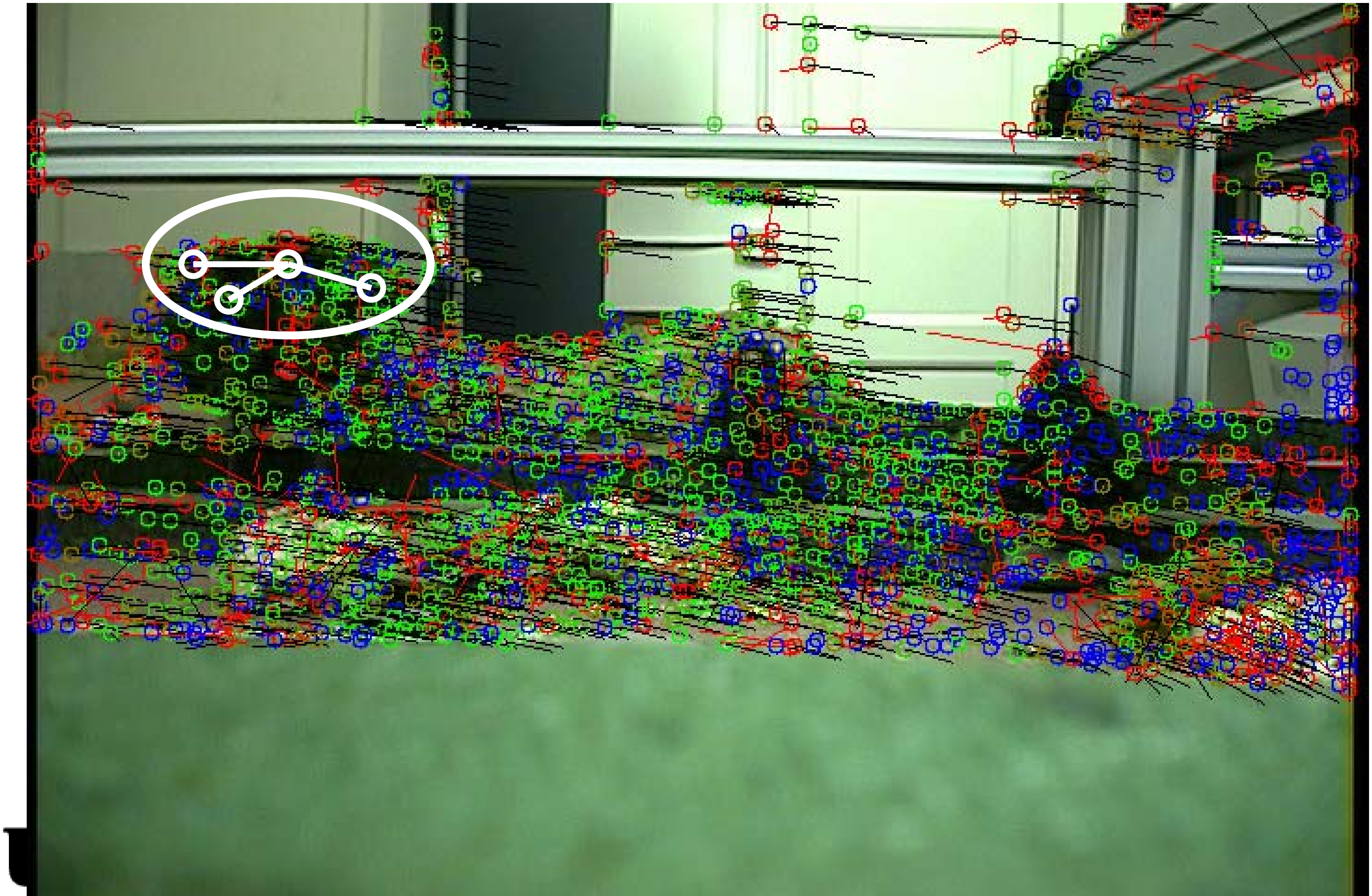
Original Video → Feature Detection → Feature Matching → Structure and Motion → 3D Reconstruction

Center for Visualization & Virtual Environments

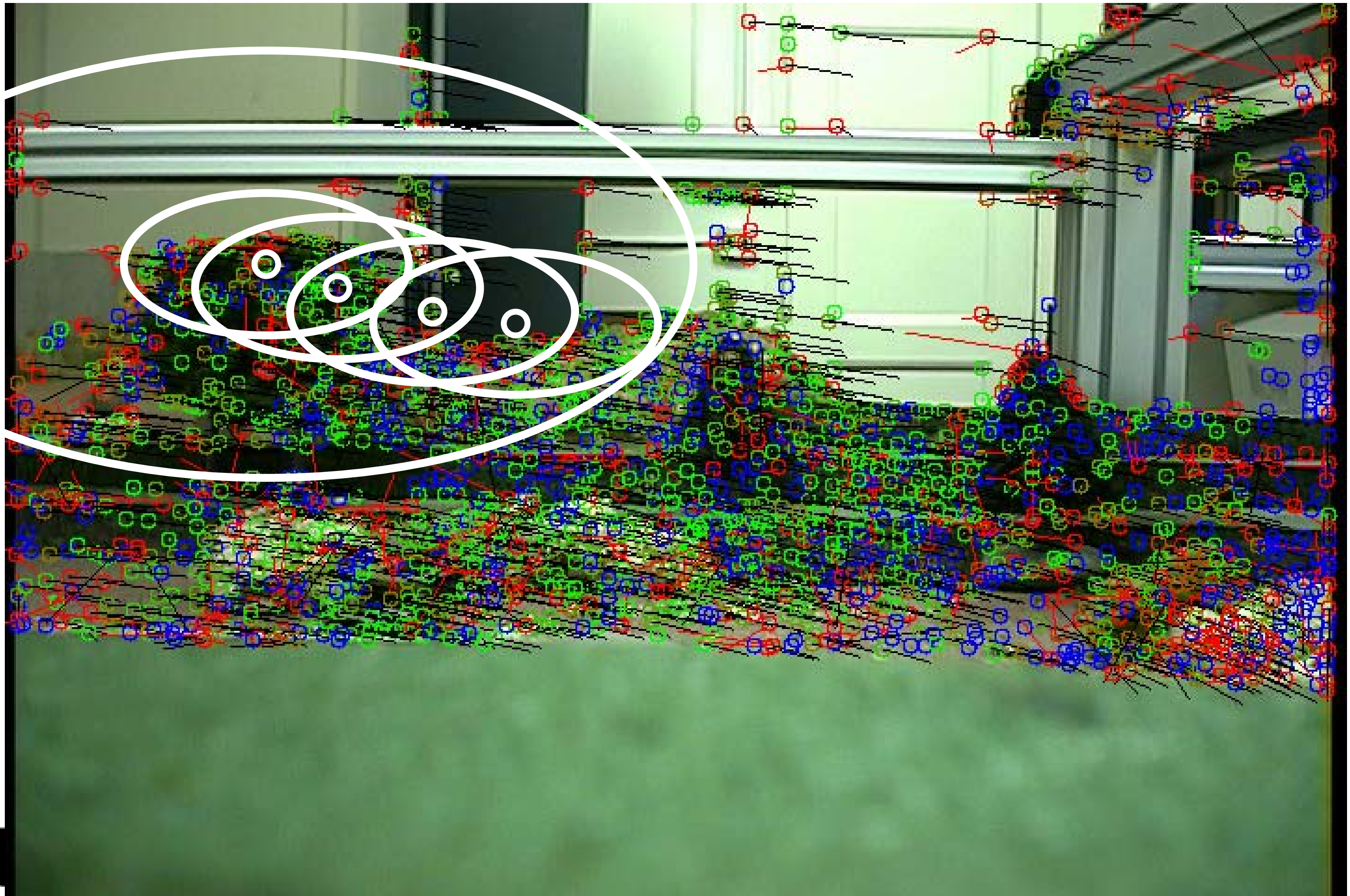
Feature Matching/Tracking

Center for Visualization & Virtual Environments

Feature Matching/Tracking



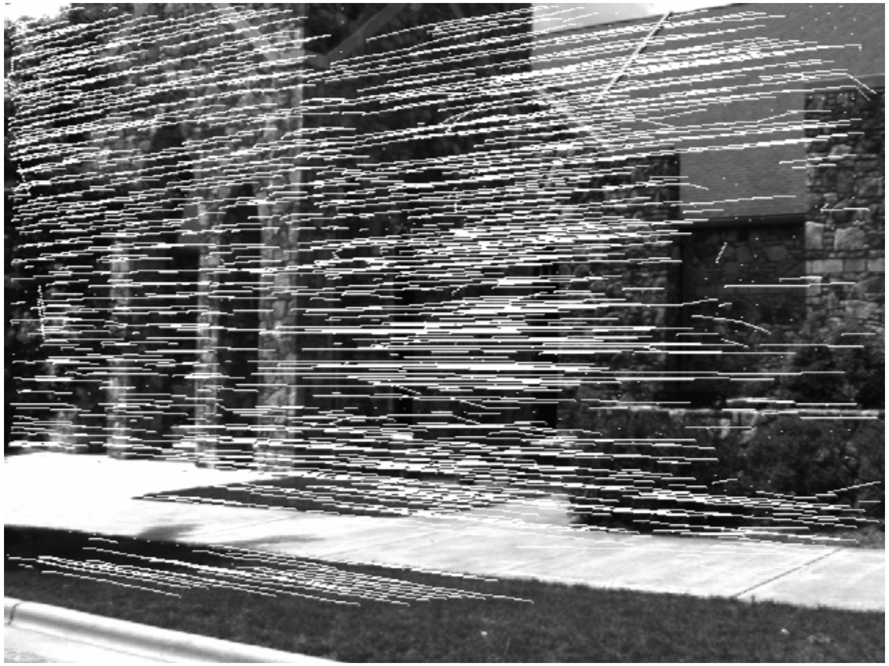
Feature Matching/Tracking



Matching vs Tracking

- Detection, while a tremendous strength in terms of scalability, is a weakness for repeatability

KLT Tracker



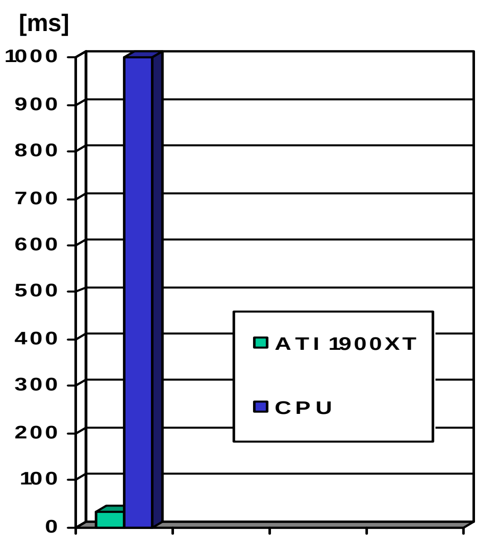
Harris Tracker

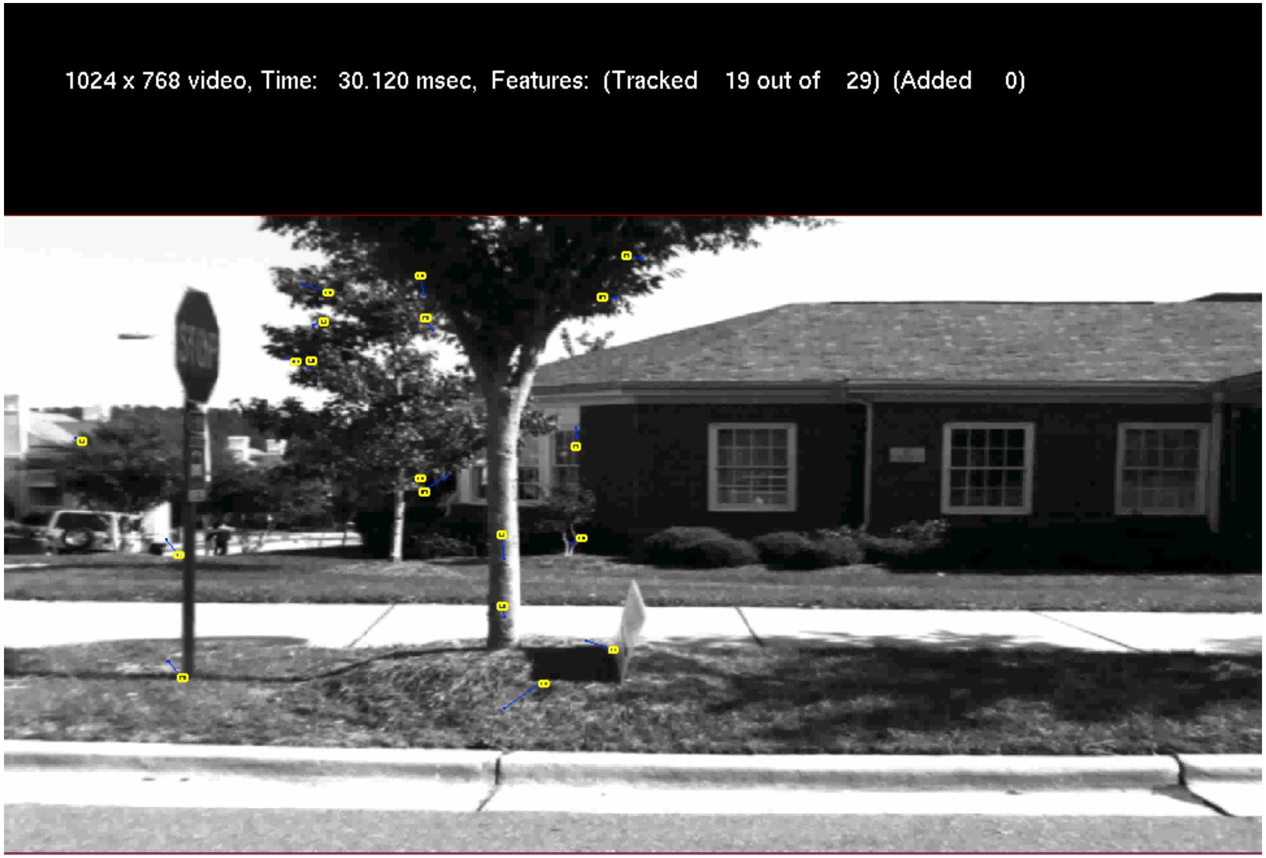
GPU KLT

work of Sudipta Sinha

Image 1024 x 768

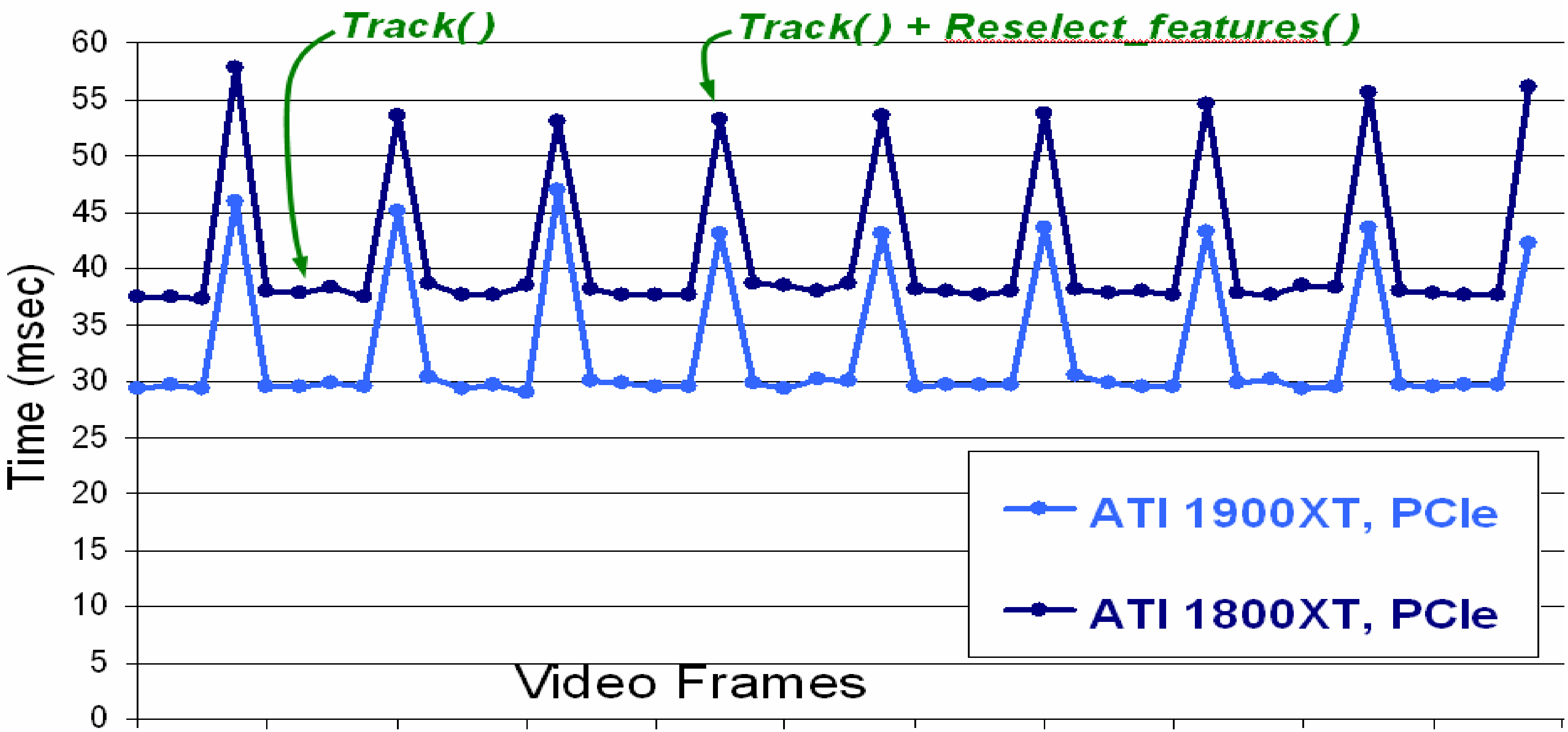
1000 features





GPU-KLT

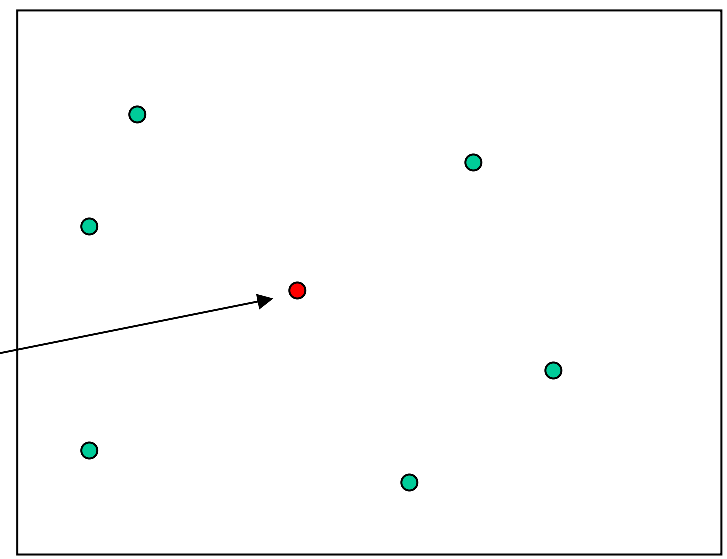
GPU-KLT Timings: **1024 x 768 video, 1000 features.**



Indexing

- Fighting the curse of dimensionality
- Locality Sensitive Hashing (LSH)
- K-d tree
- Vocabulary Tree

Find nearest neighbor



tf-idf

- Term Frequency Inverse Document Frequency
- Is a weighting of words in a document

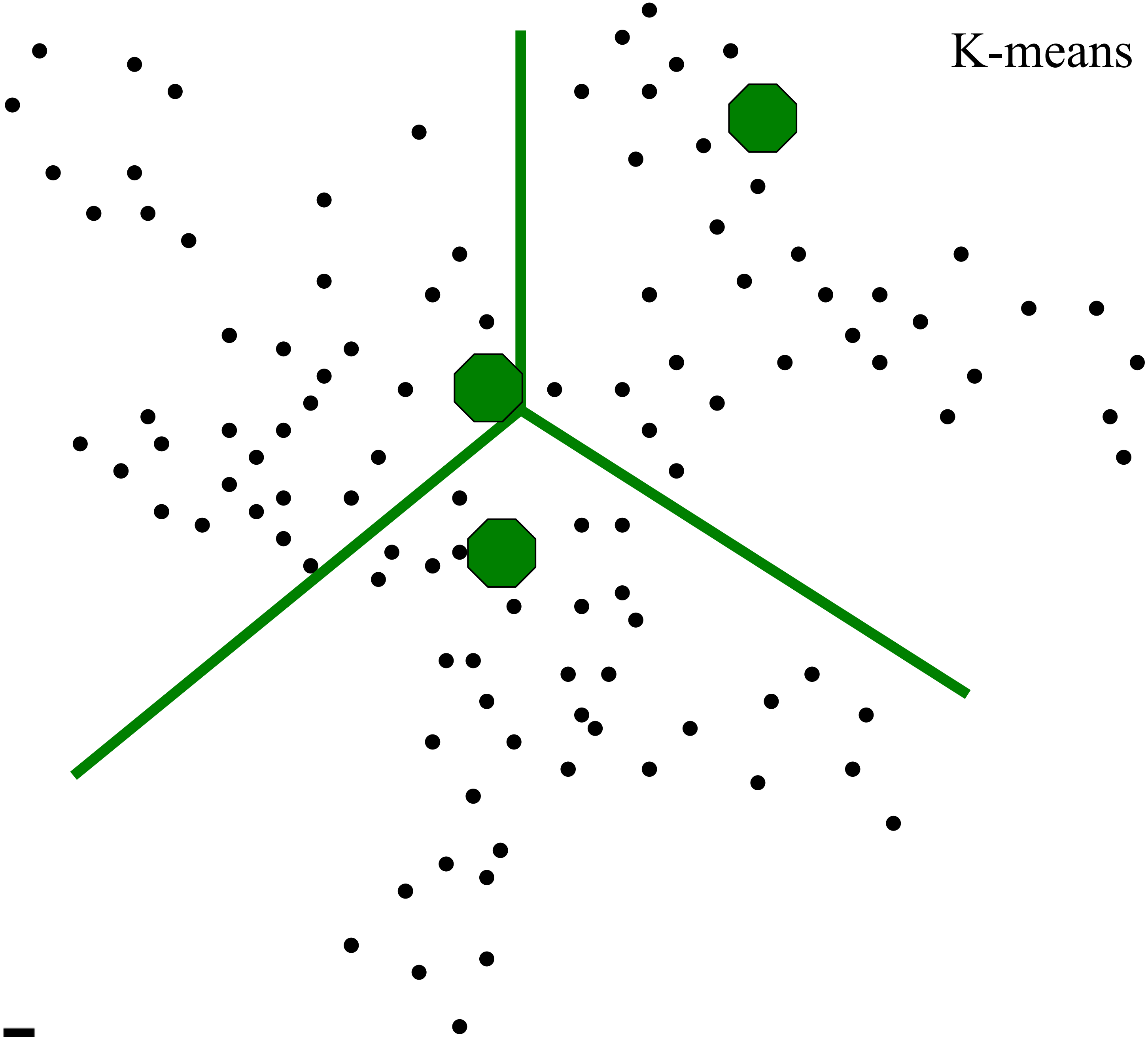
$$(n/N) \log (D/d)$$

Center for Visualization & Virtual Environments

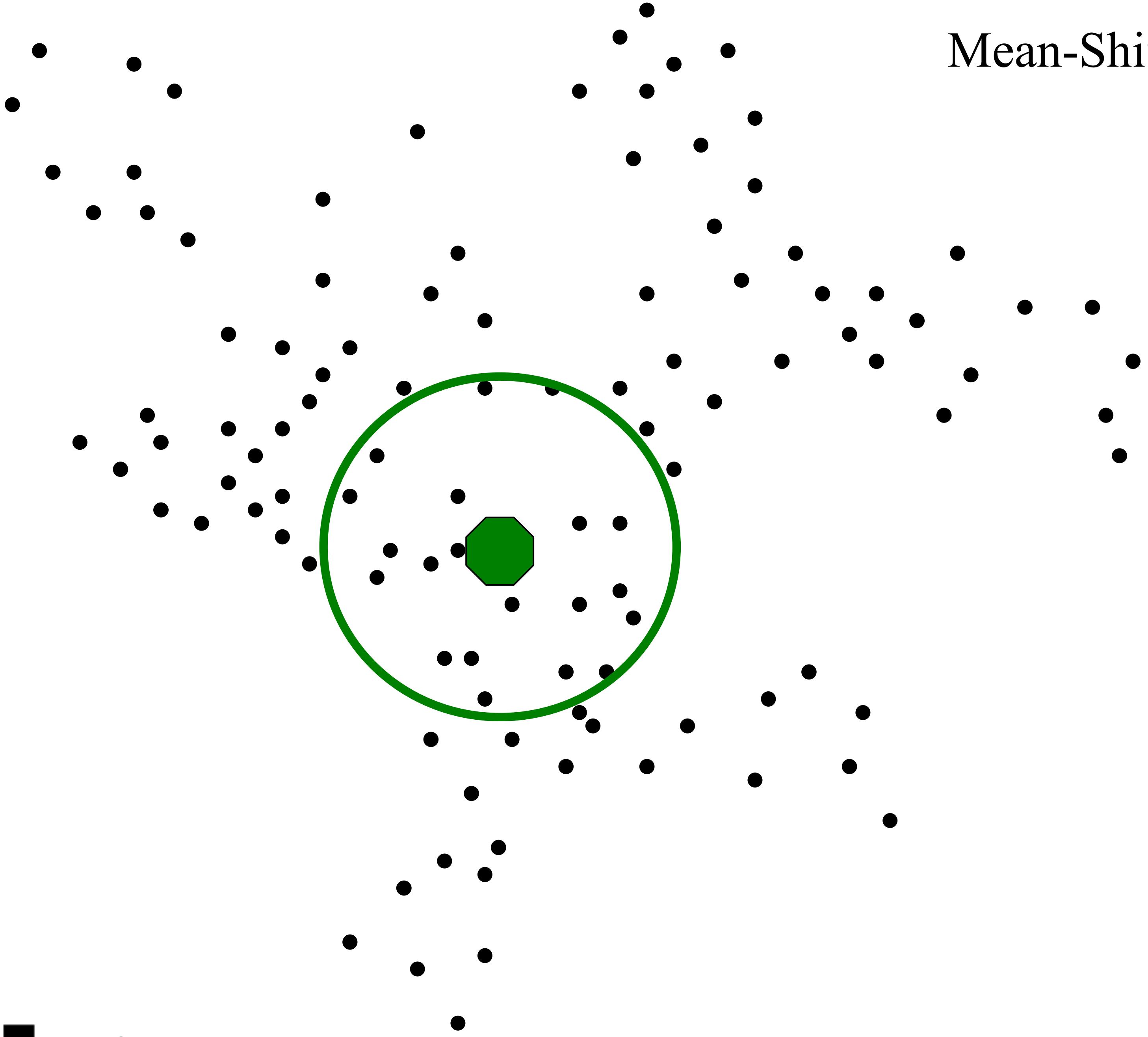
Clustering

- K-Means
- K-Medioids
- Mean-Shift
- Spectral Clustering
- Graph-Cuts

K-means



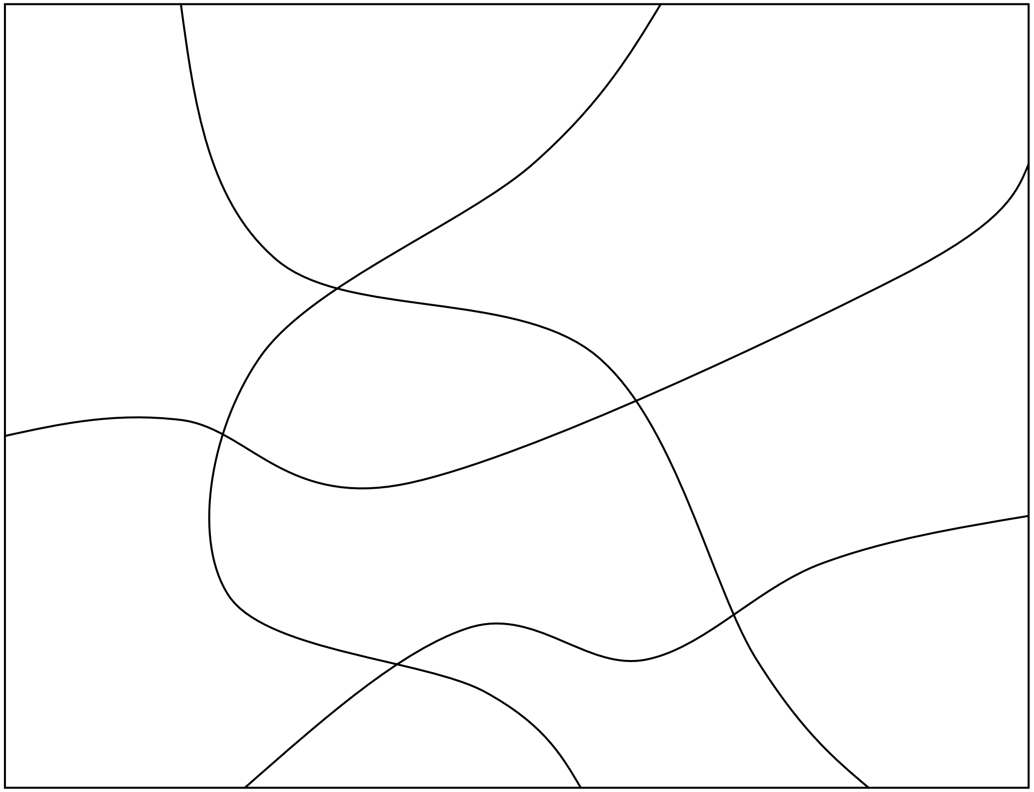
Mean-Shift



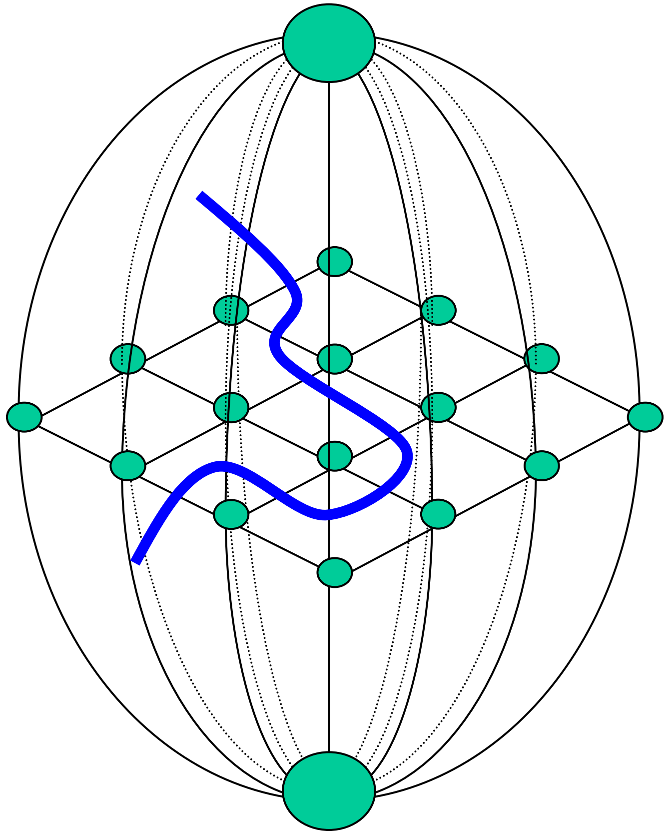
UK

Center for Visualization & Virtual Environments

Break into eigen-modes



v_-



v_+

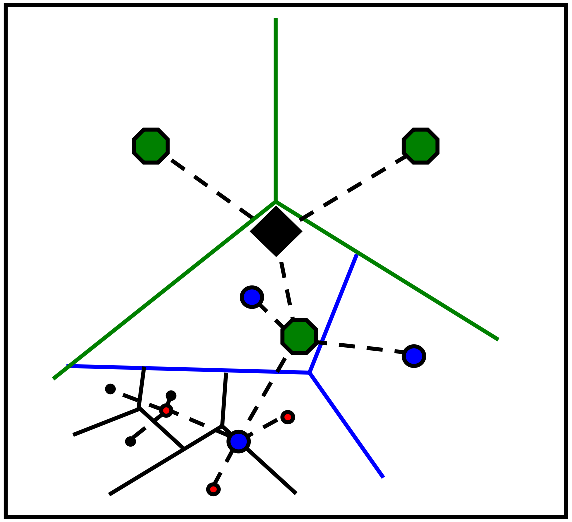
Machine Learning

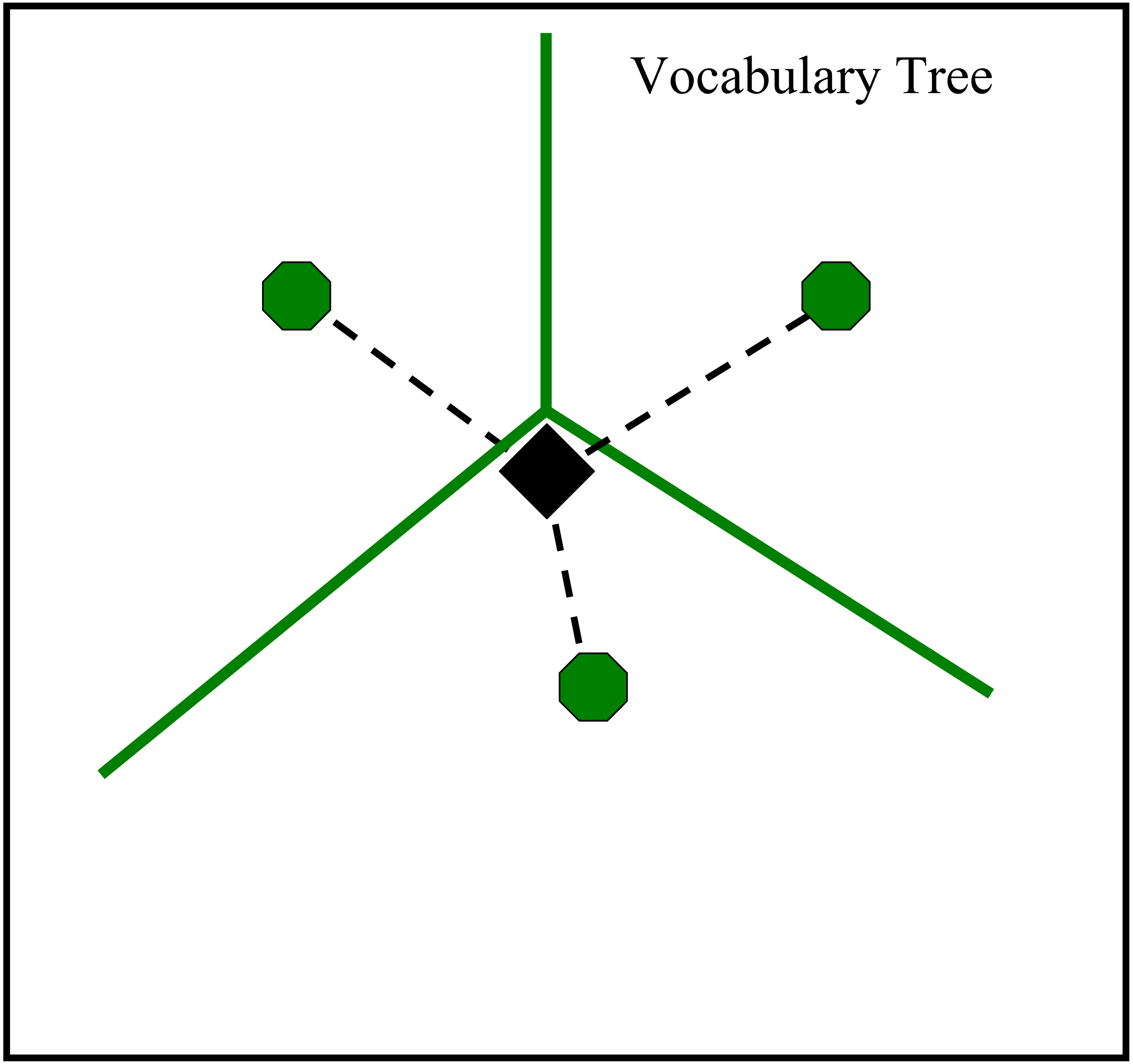
- When parametric invariance is insufficient
- Supervised,Unsupervised,Semisupervised
- Support Vector Machines (SVM's)
- Boosting
- Neural Nets

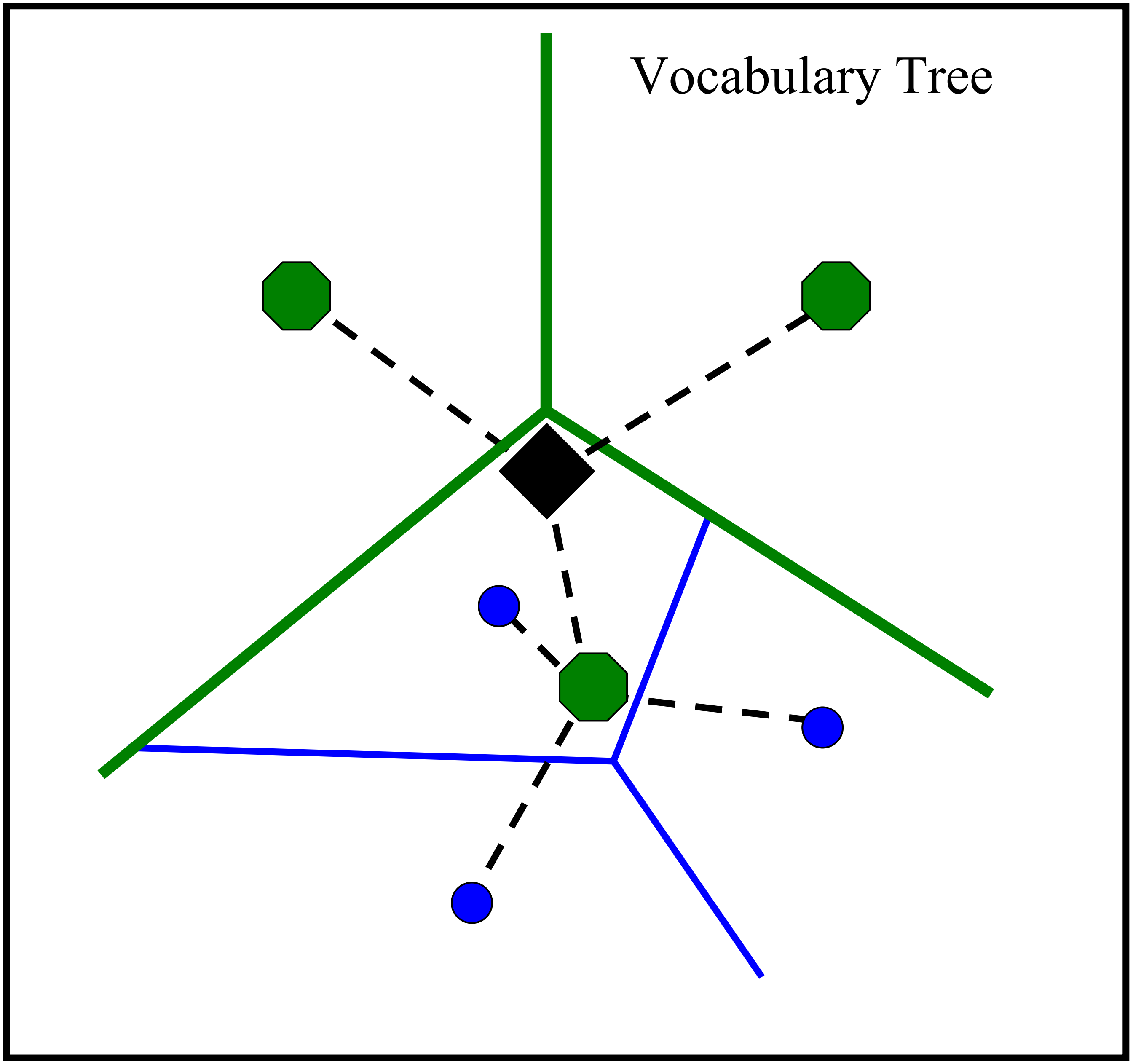
Scalability

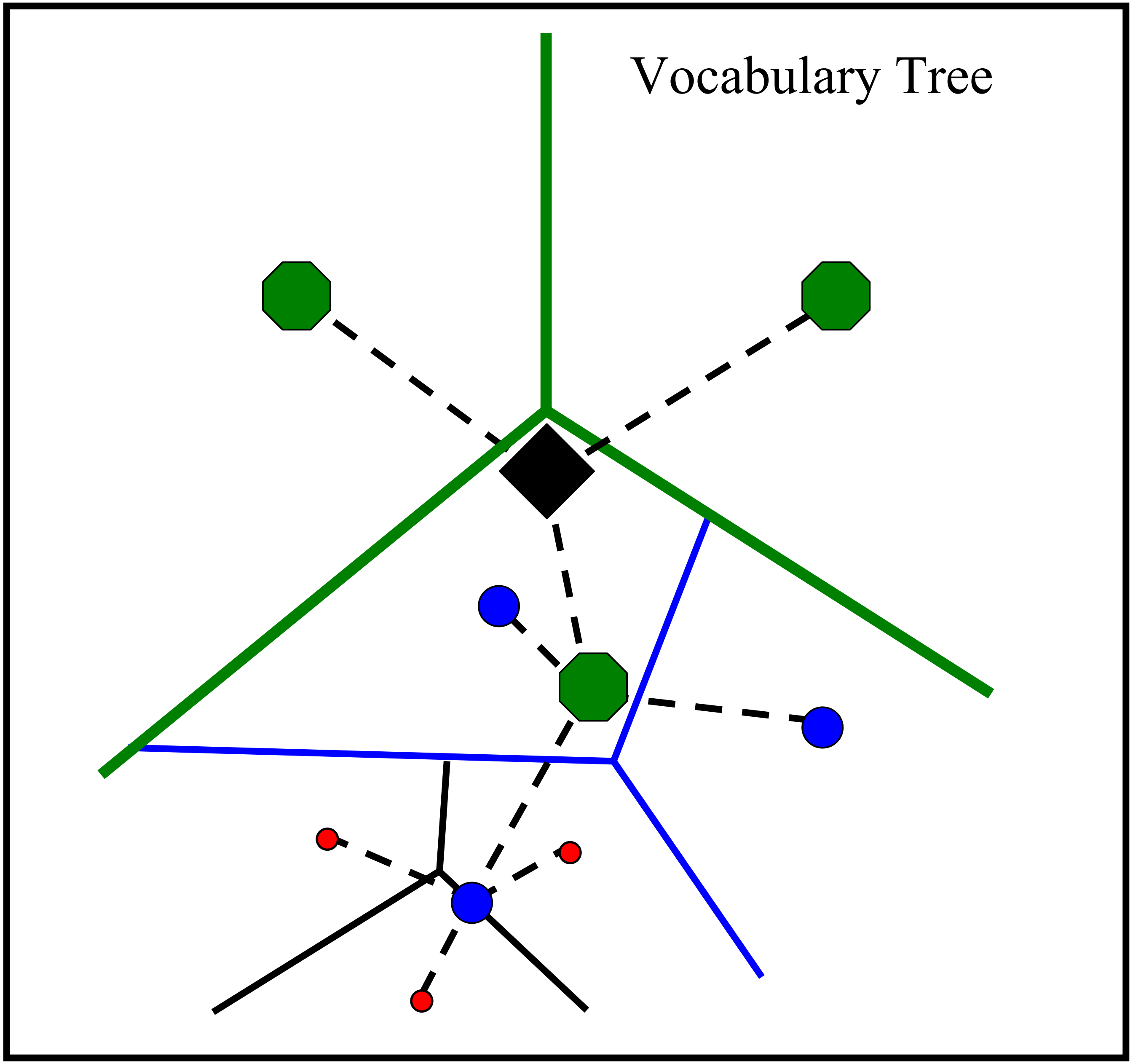
If we can get repeatable, discriminative features,

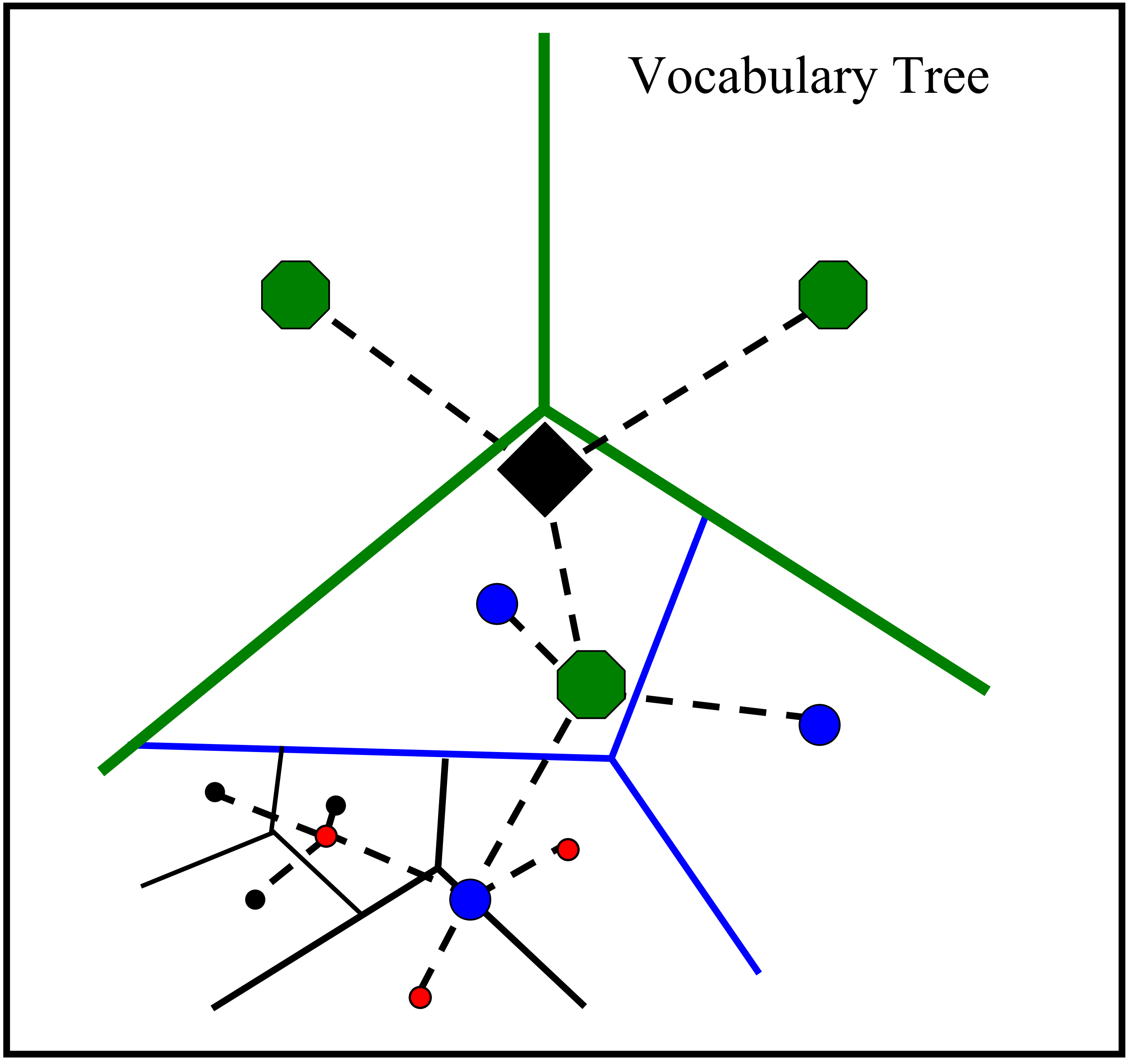
then recognition can scale to very large databases using the vocabulary tree and indexing approach described in Nistér & Stewénius CVPR 2006.




Vocabulary Tree

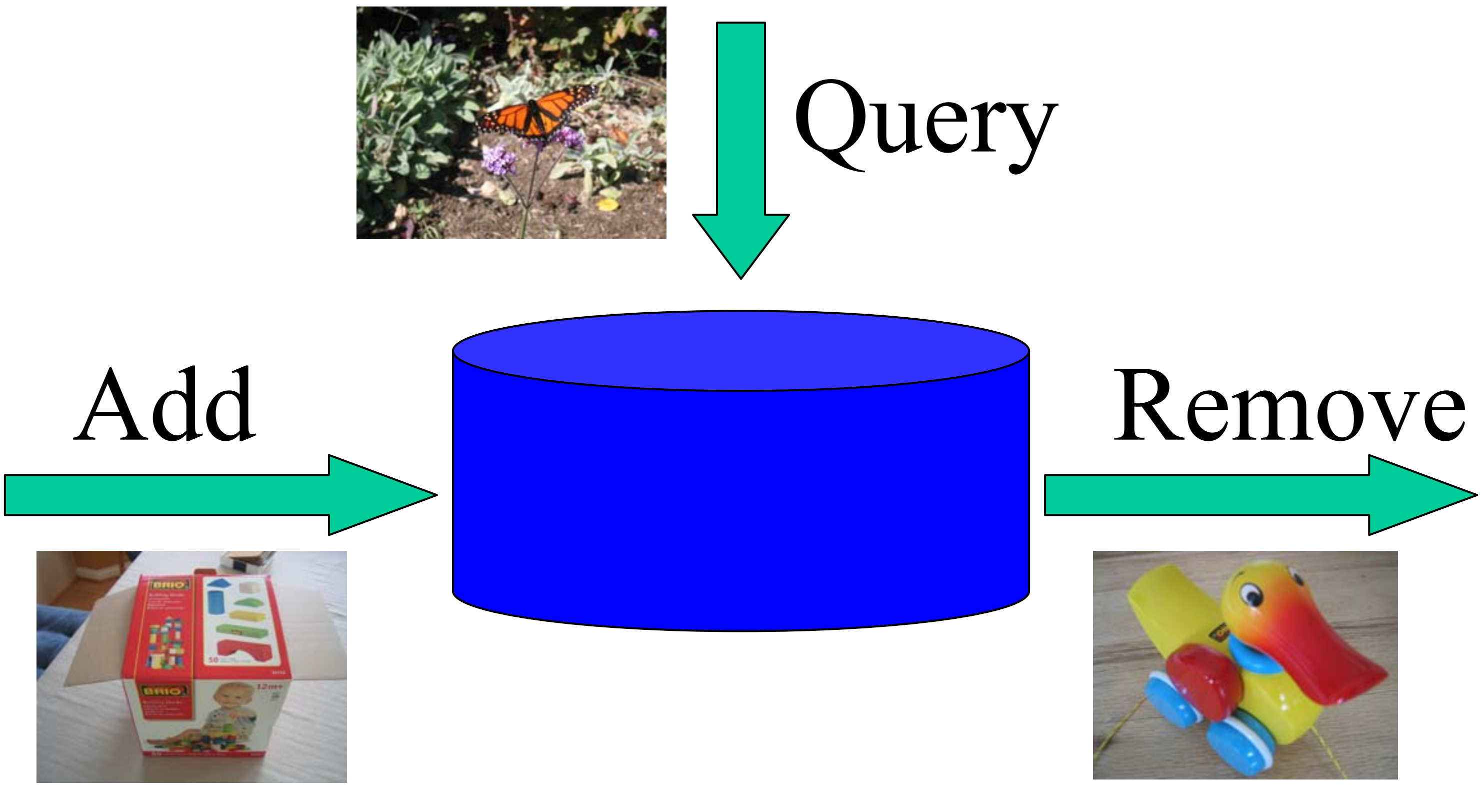

Vocabulary Tree


Vocabulary Tree



Center for Visualization & Virtual Environments

Adding, Querying and Removing Images at full speed

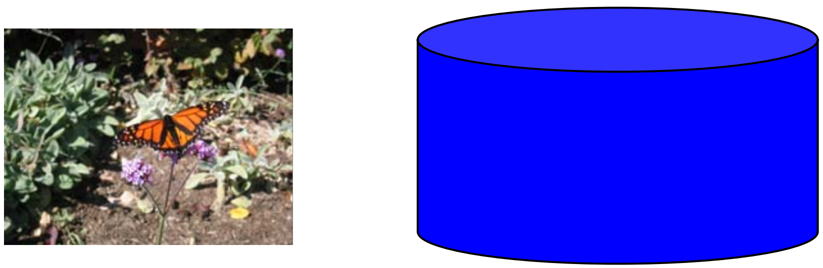


UK

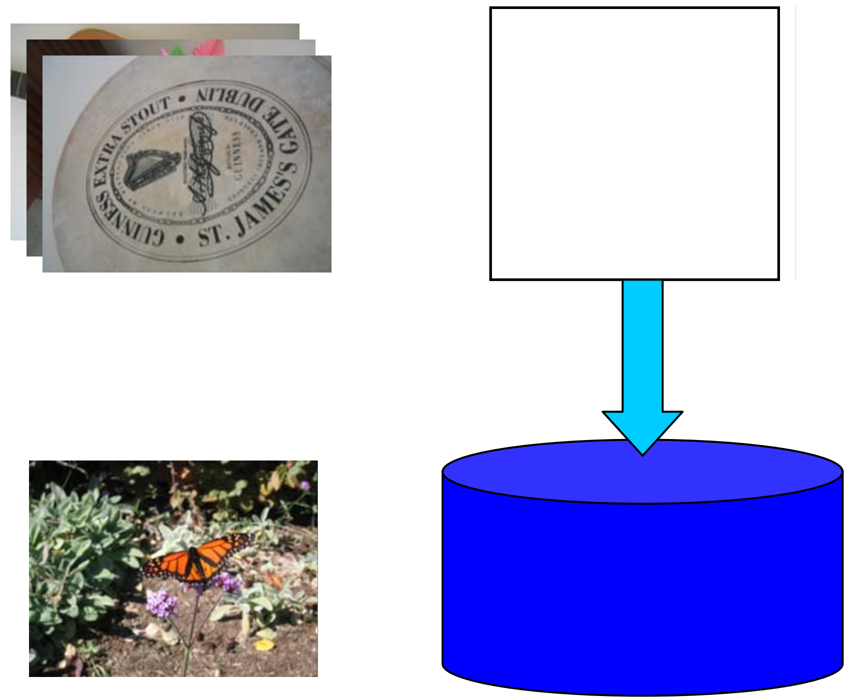
Center for Visualization & Virtual Environments

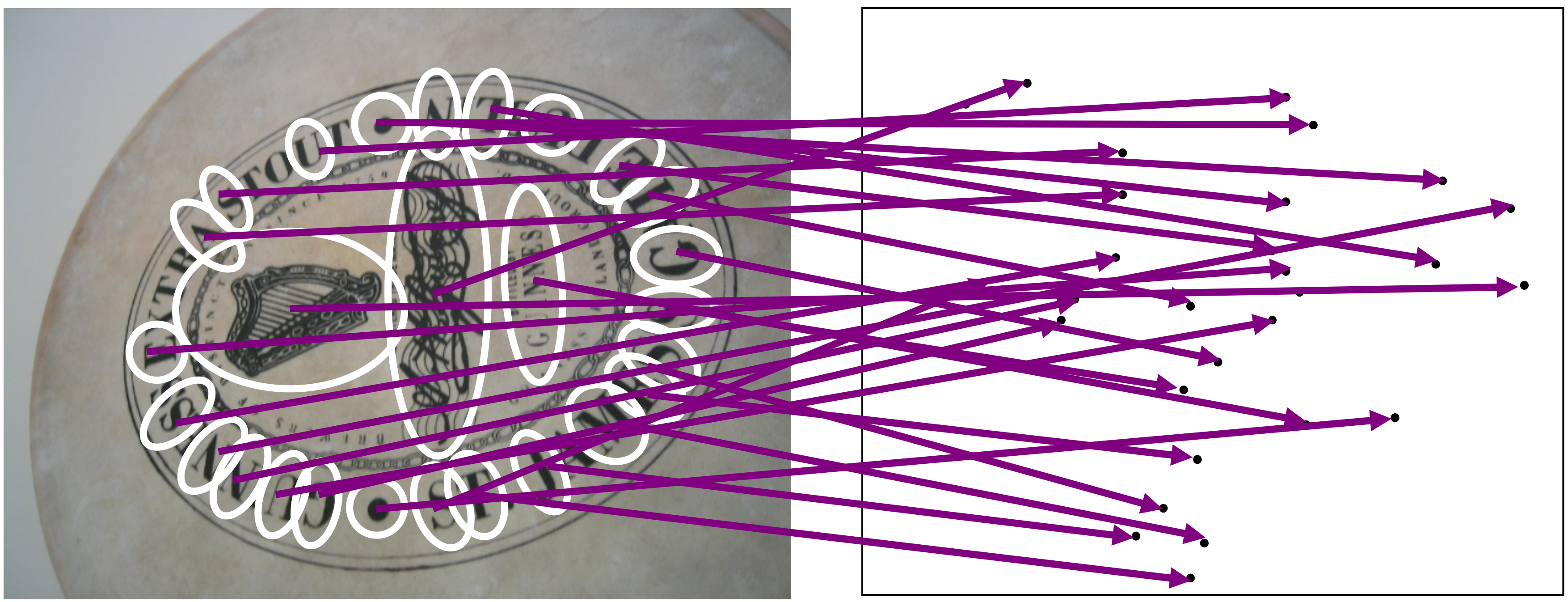
Training and Addition are Separate

Common Approach

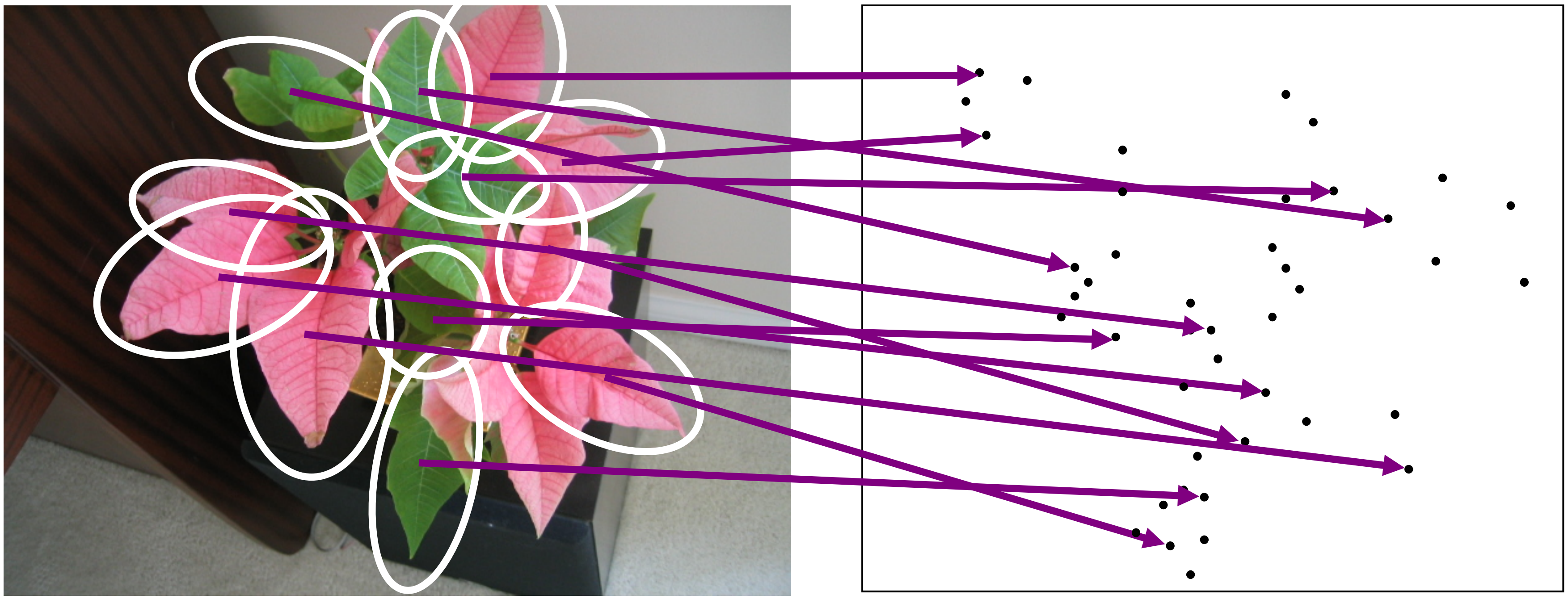


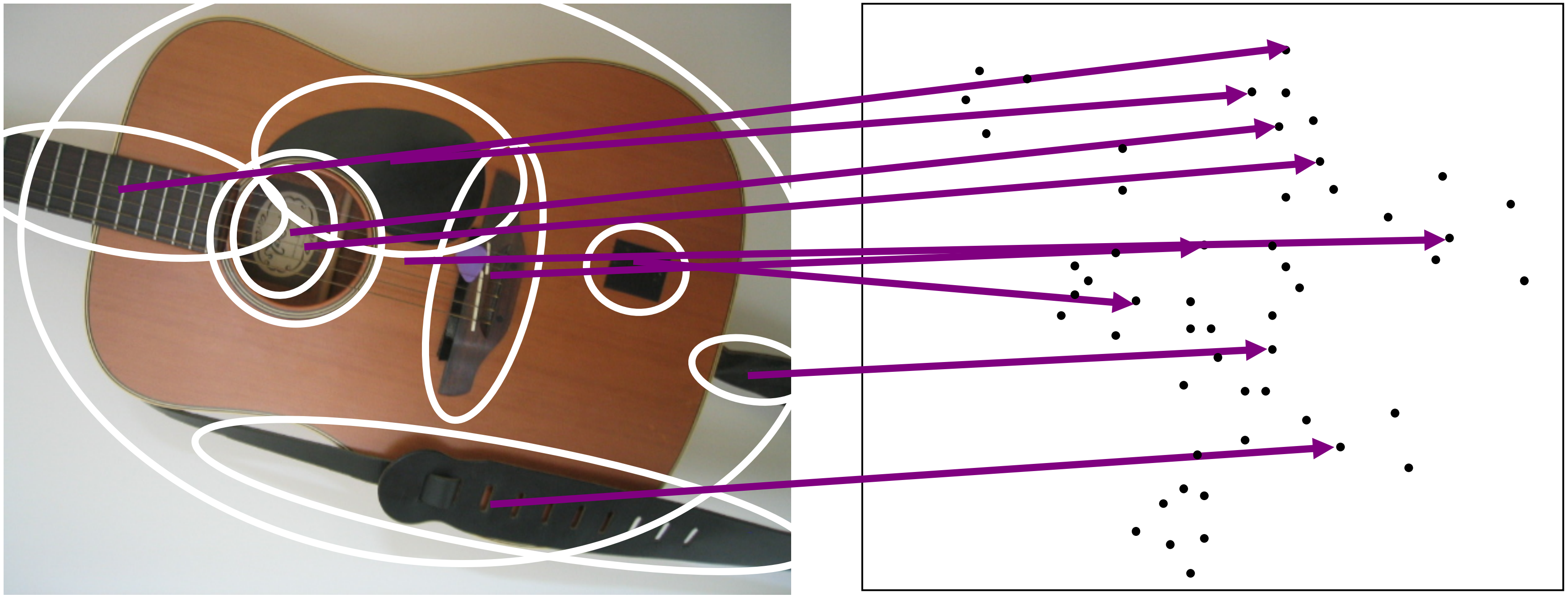
Our approach

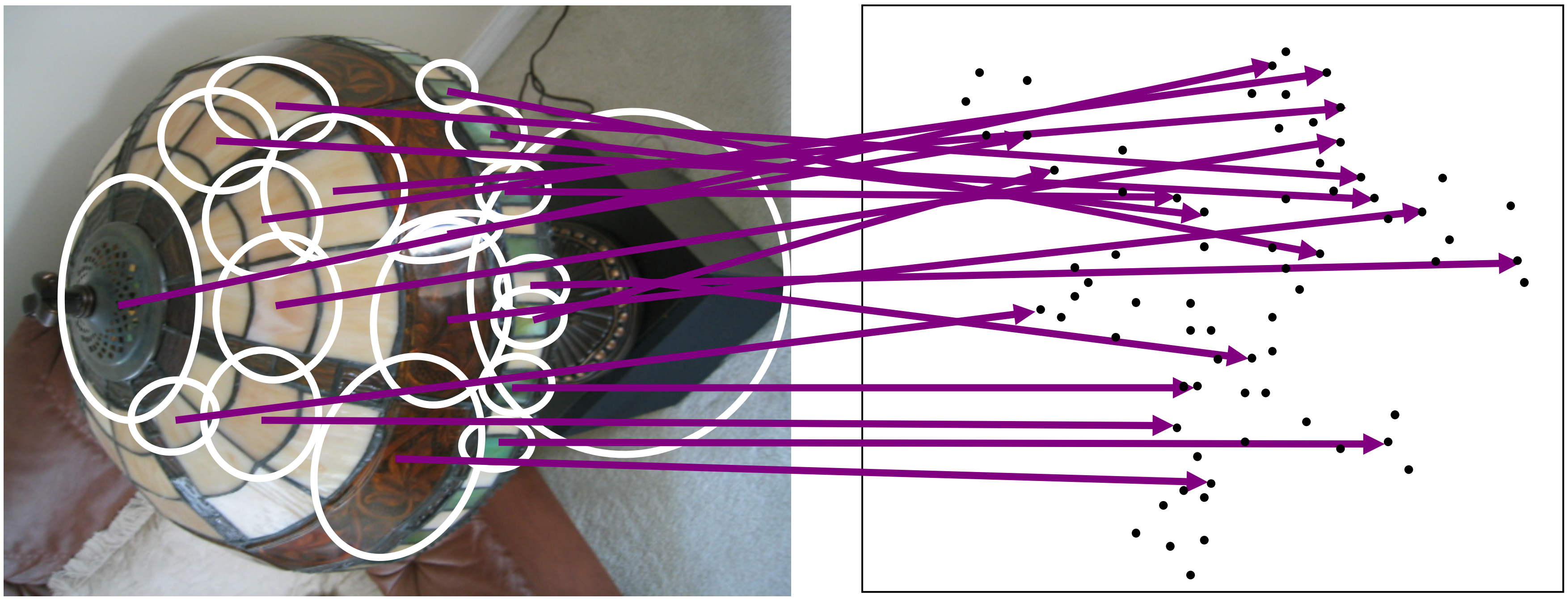




UK Center for Visualization & Virtual Environments

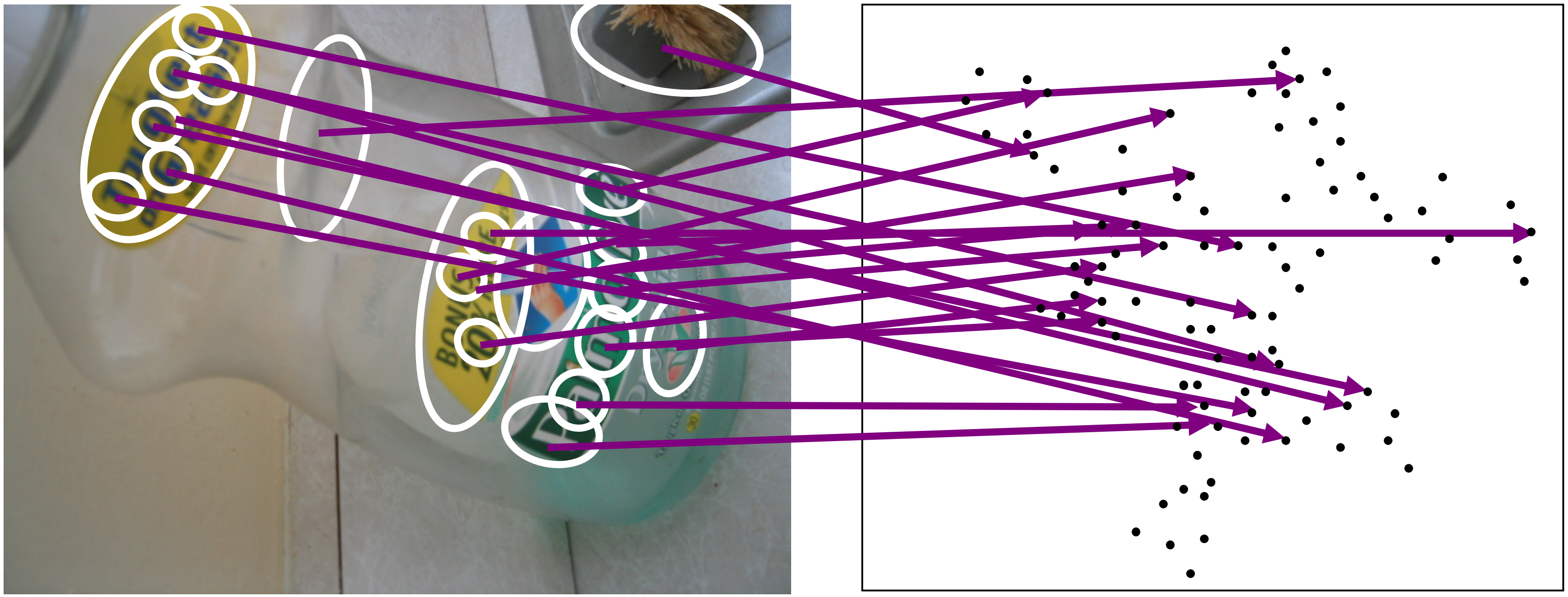


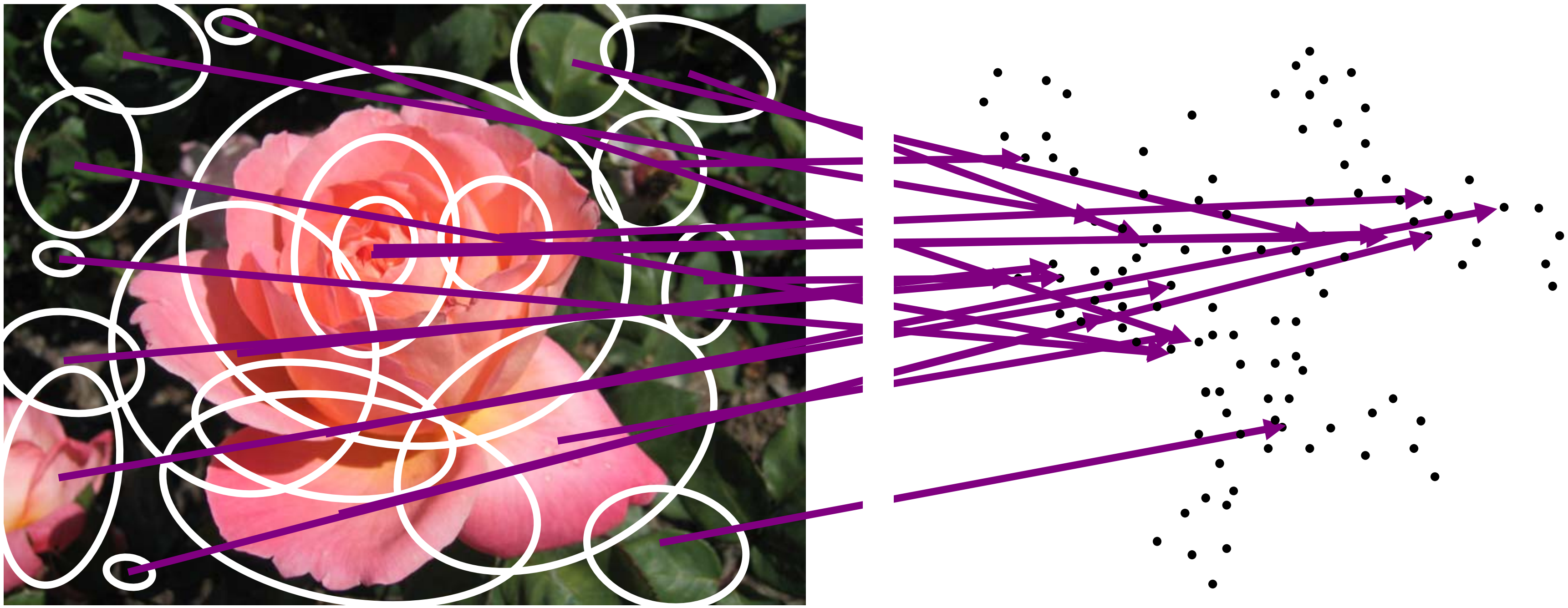




UK

Center for Visualization & Virtual Environments

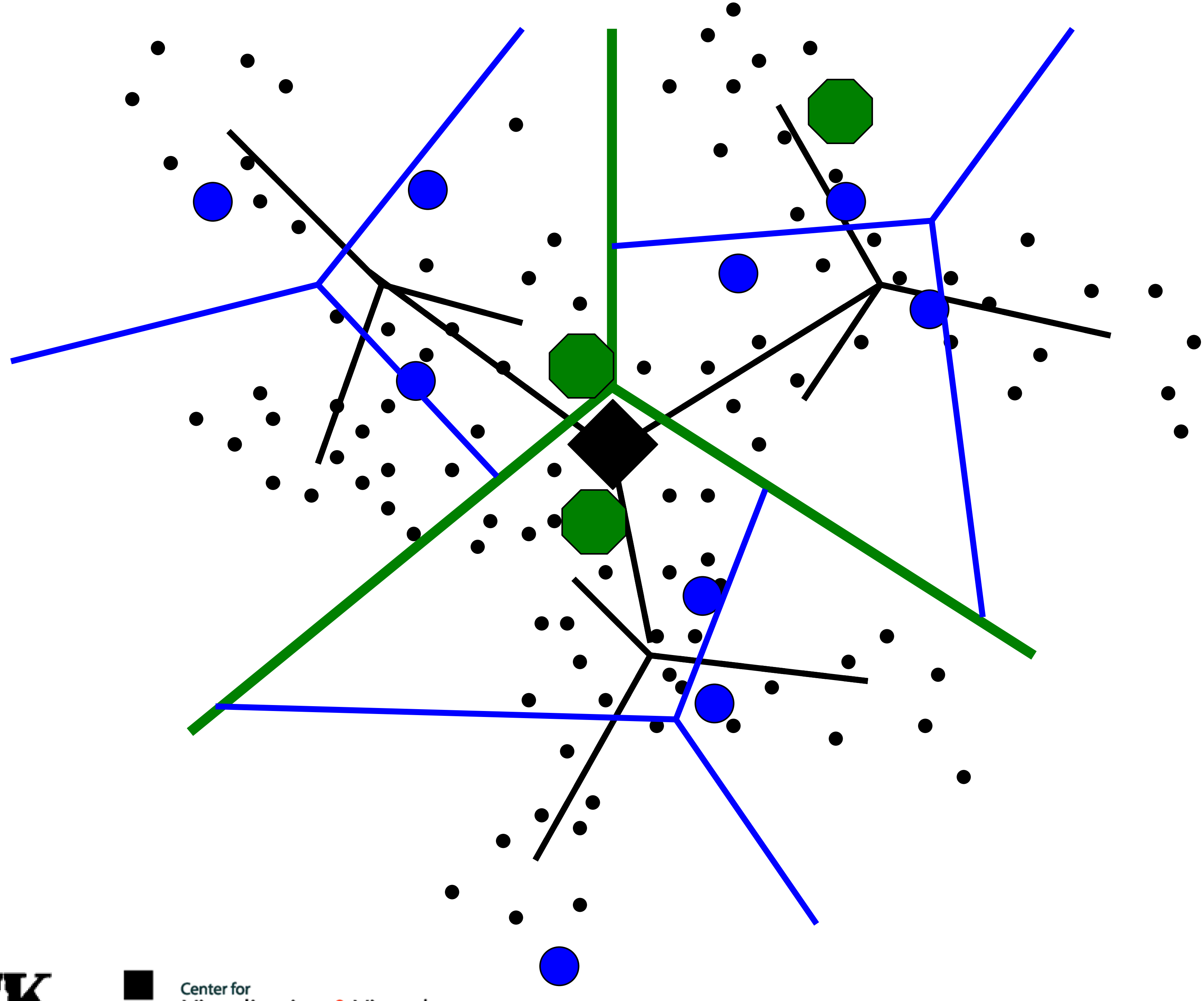




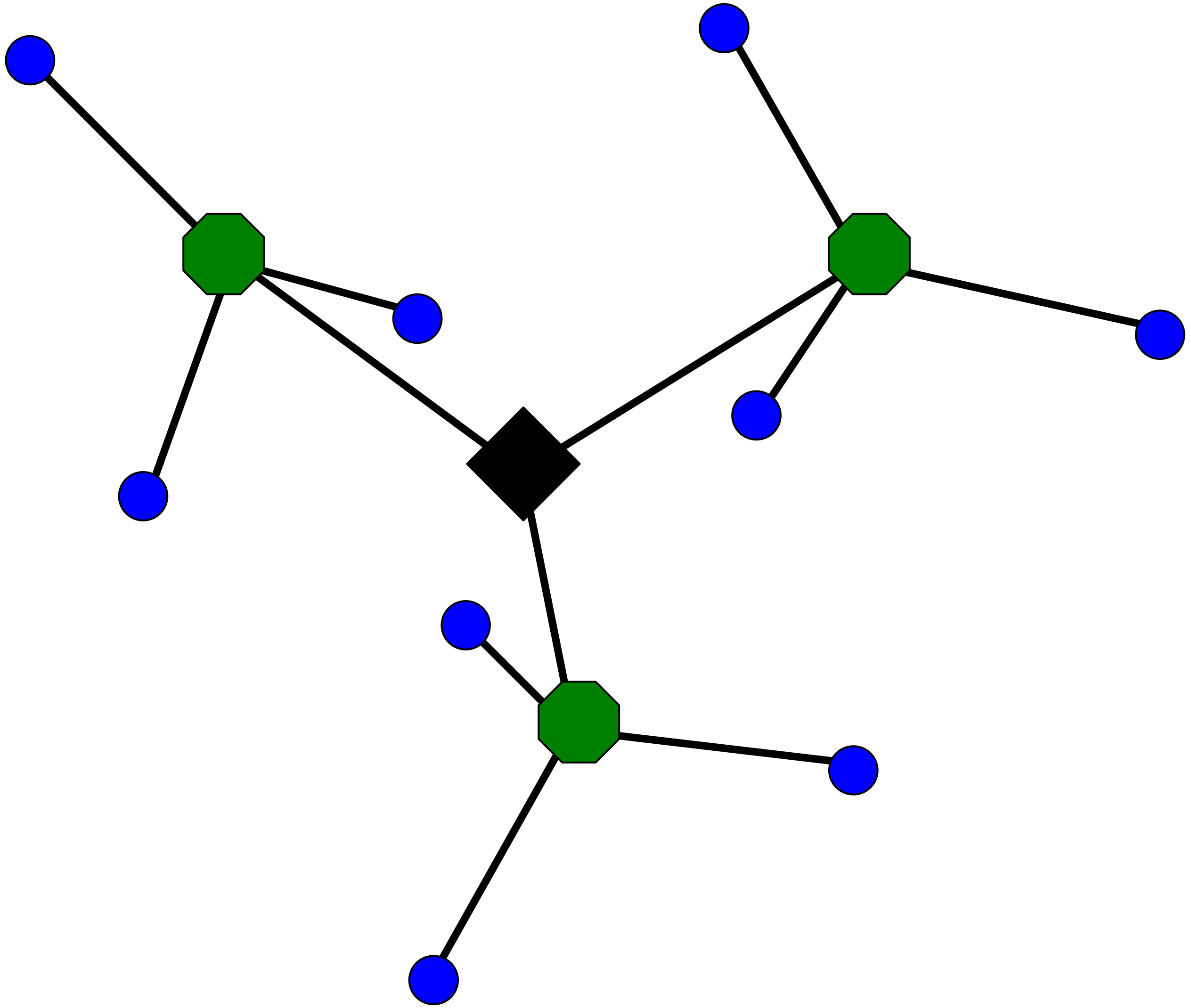
UK Center for Visualization & Virtual Environments

UK

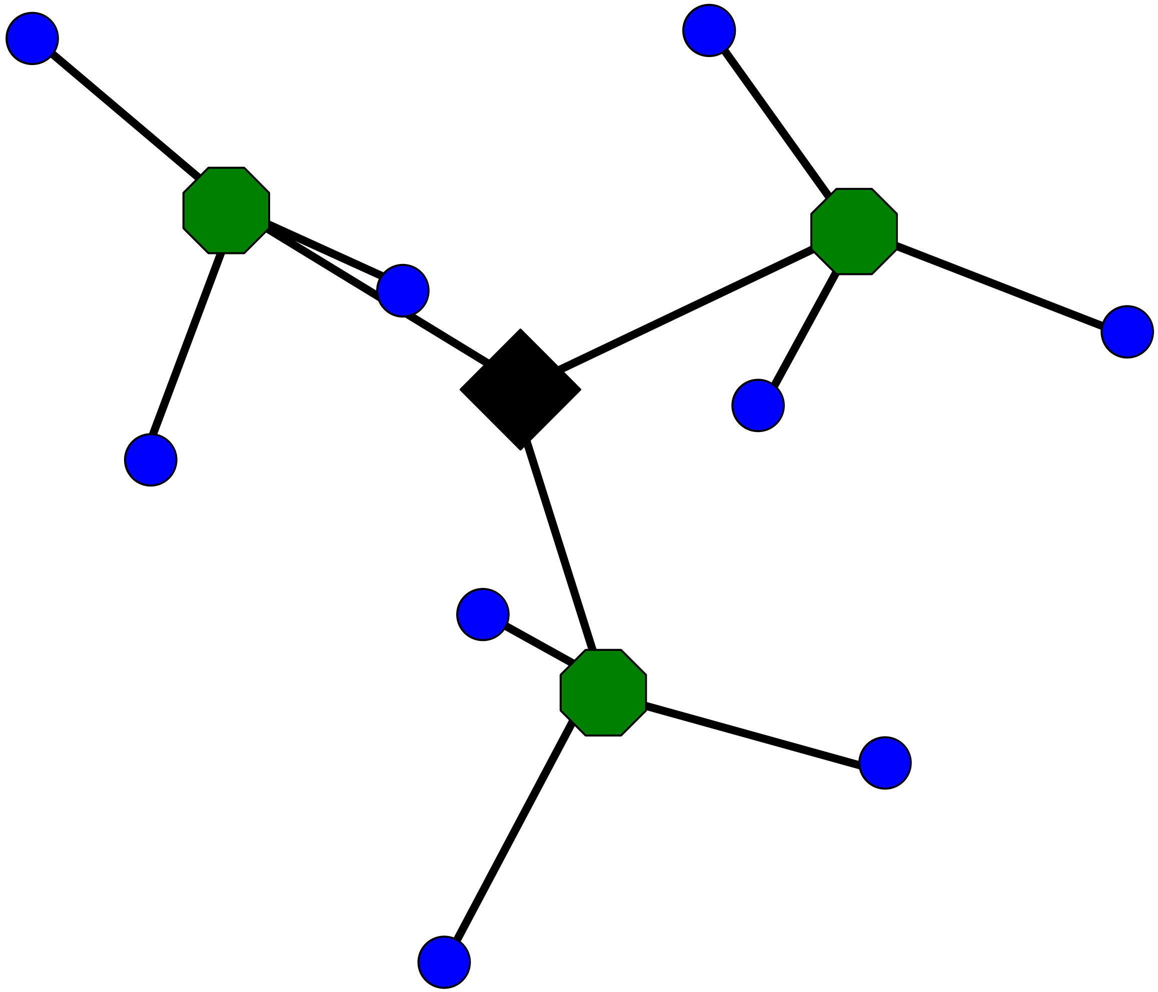
Center for Visualization & Virtual Environments

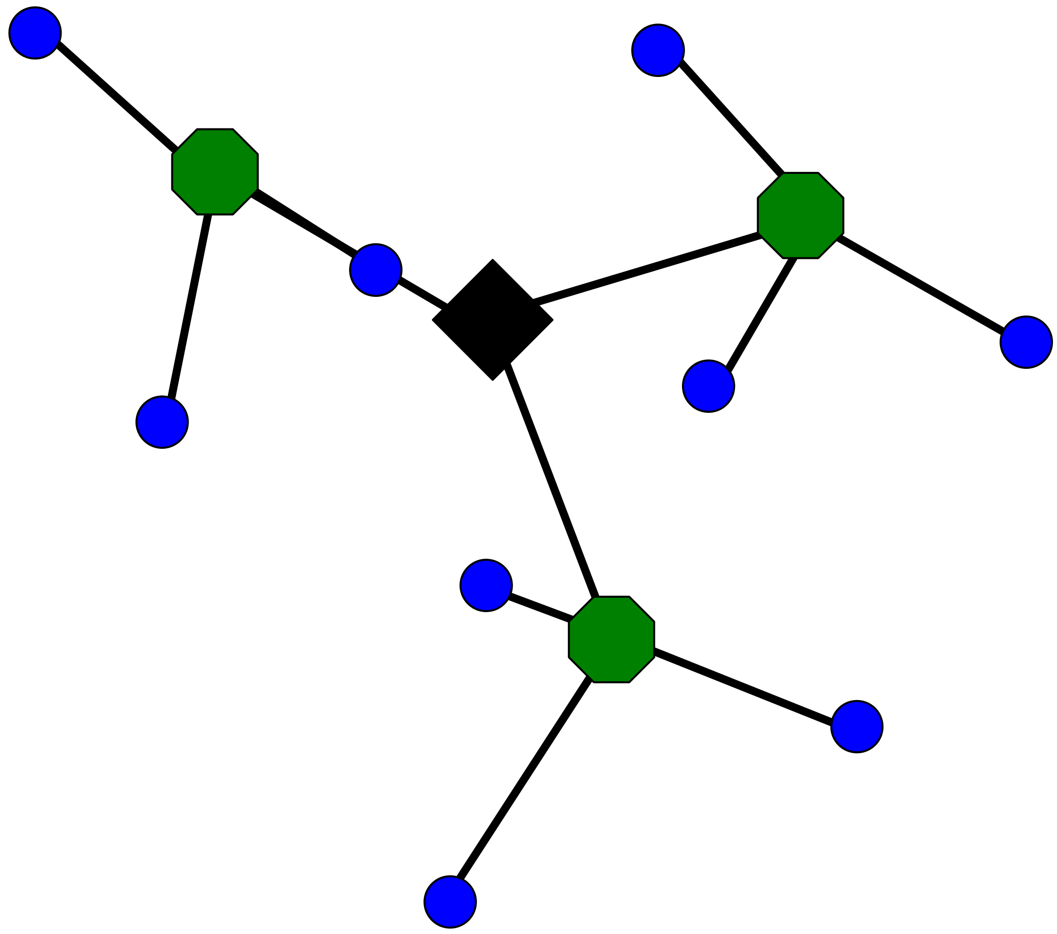


UK Center for Visualization & Virtual Environments

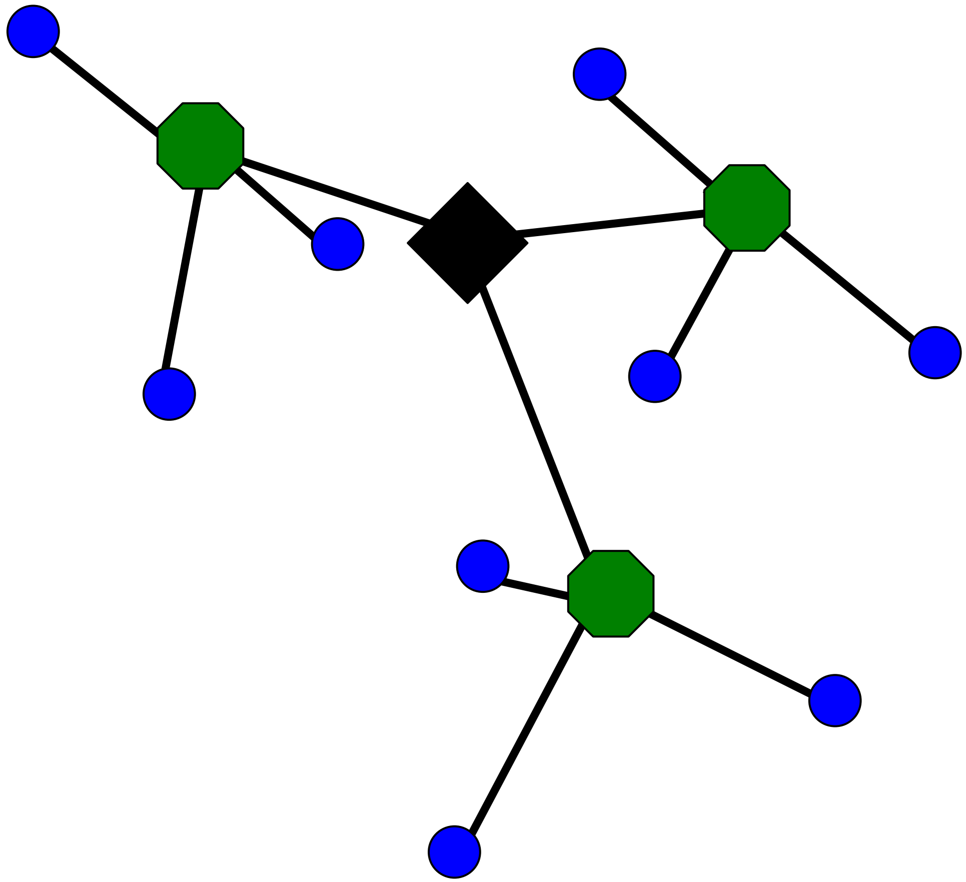


Center for Visualization & Virtual Environments

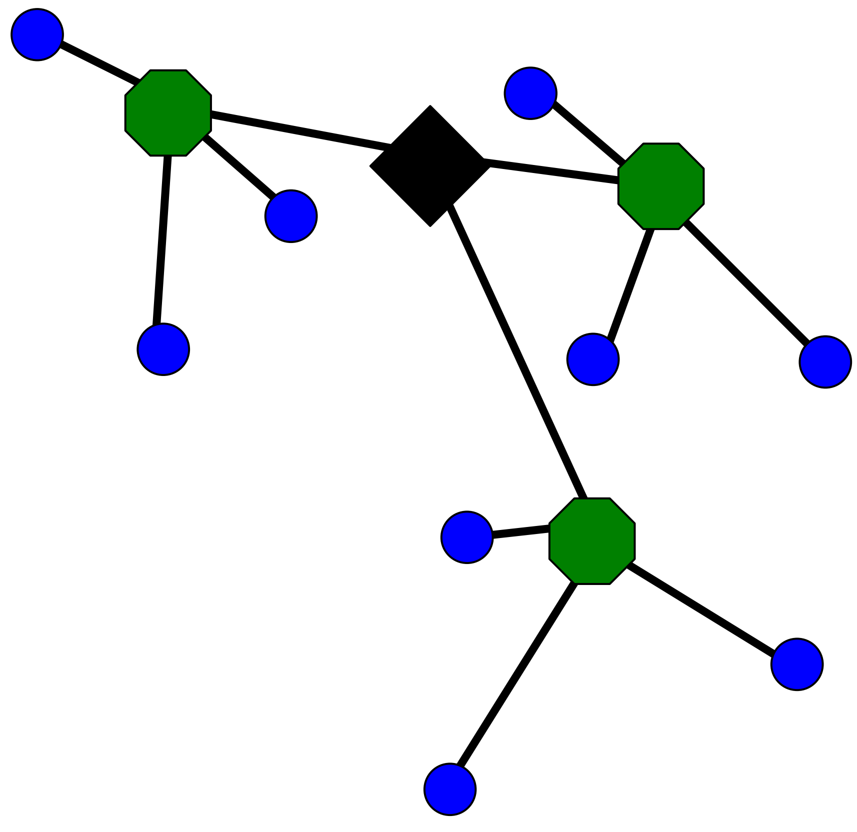


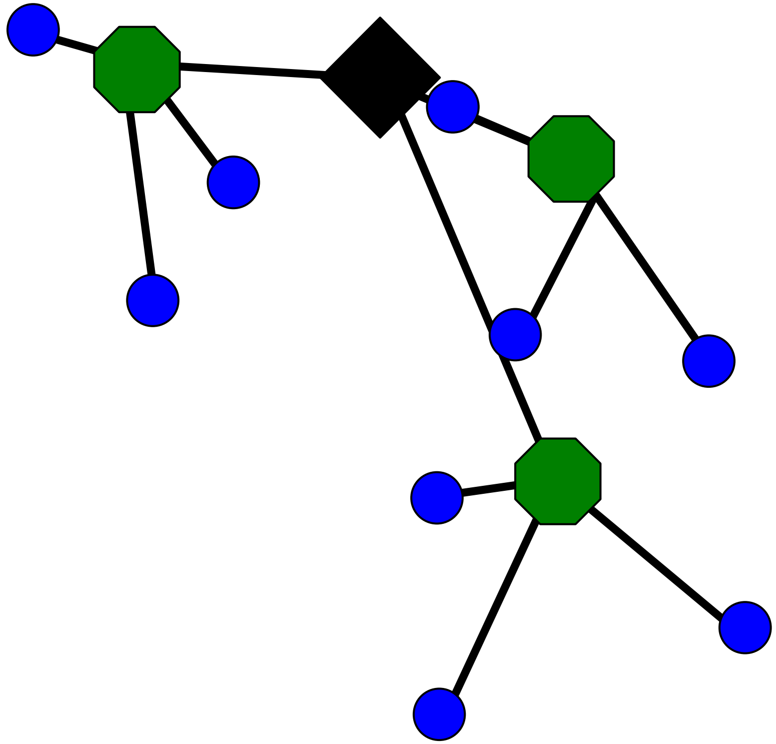


Center for Visualization & Virtual Environments

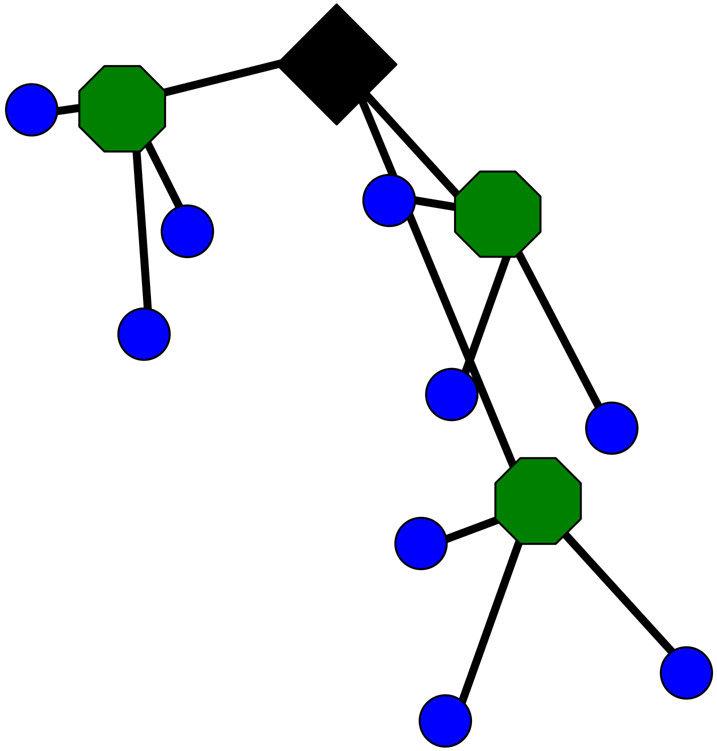


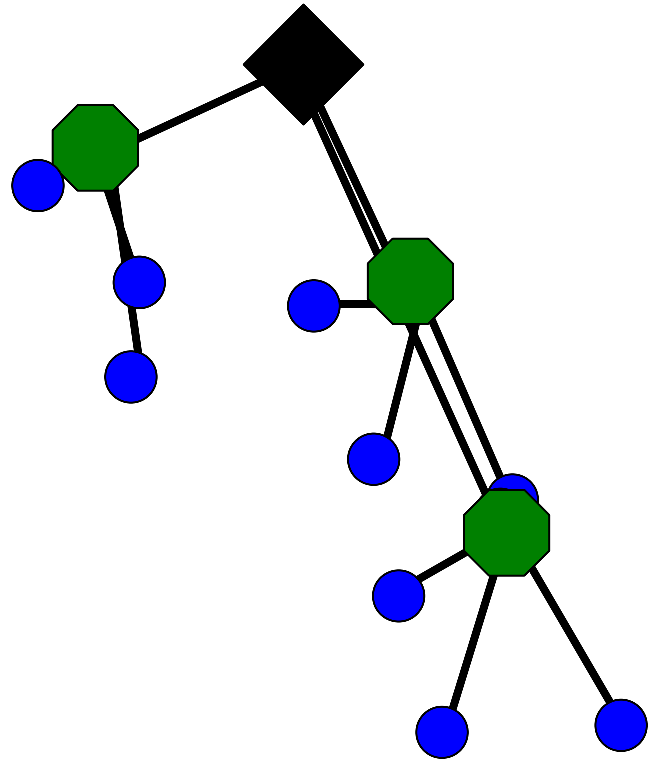
UK Center for Visualization & Virtual Environments

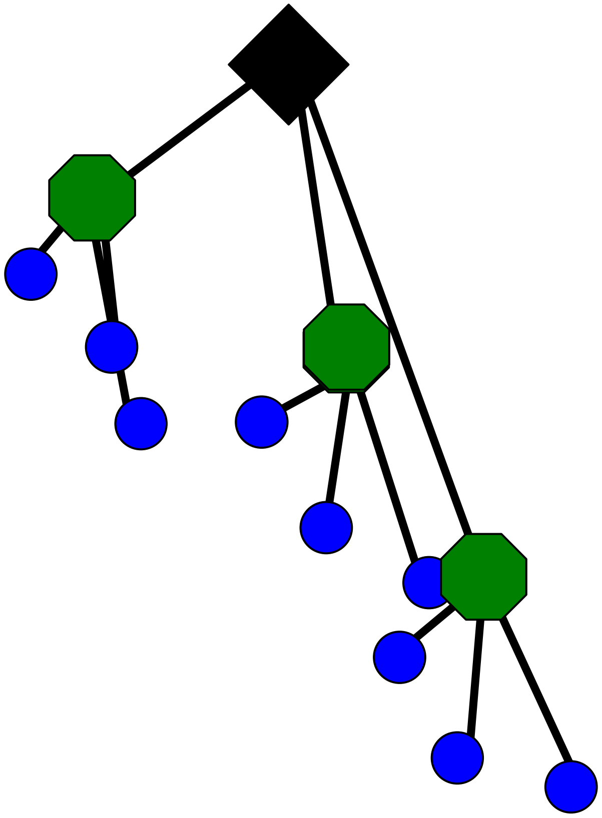


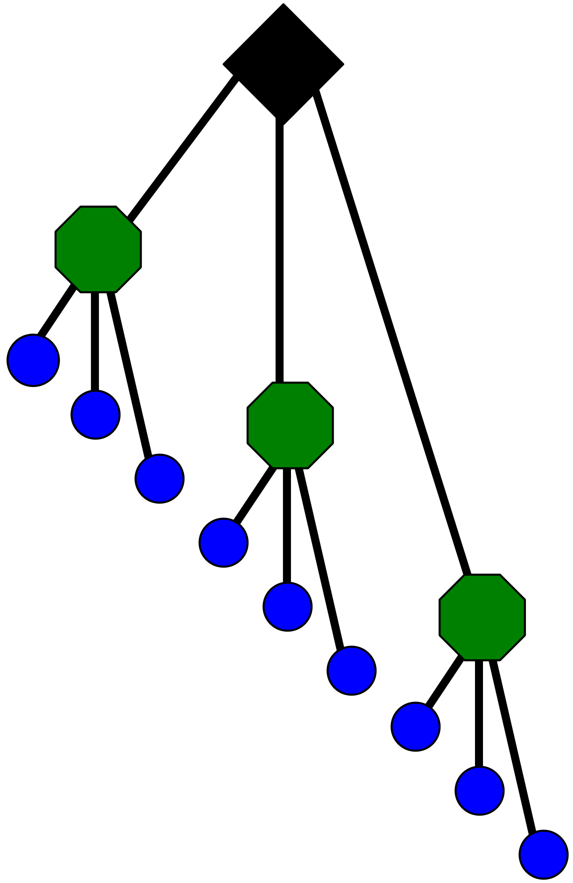


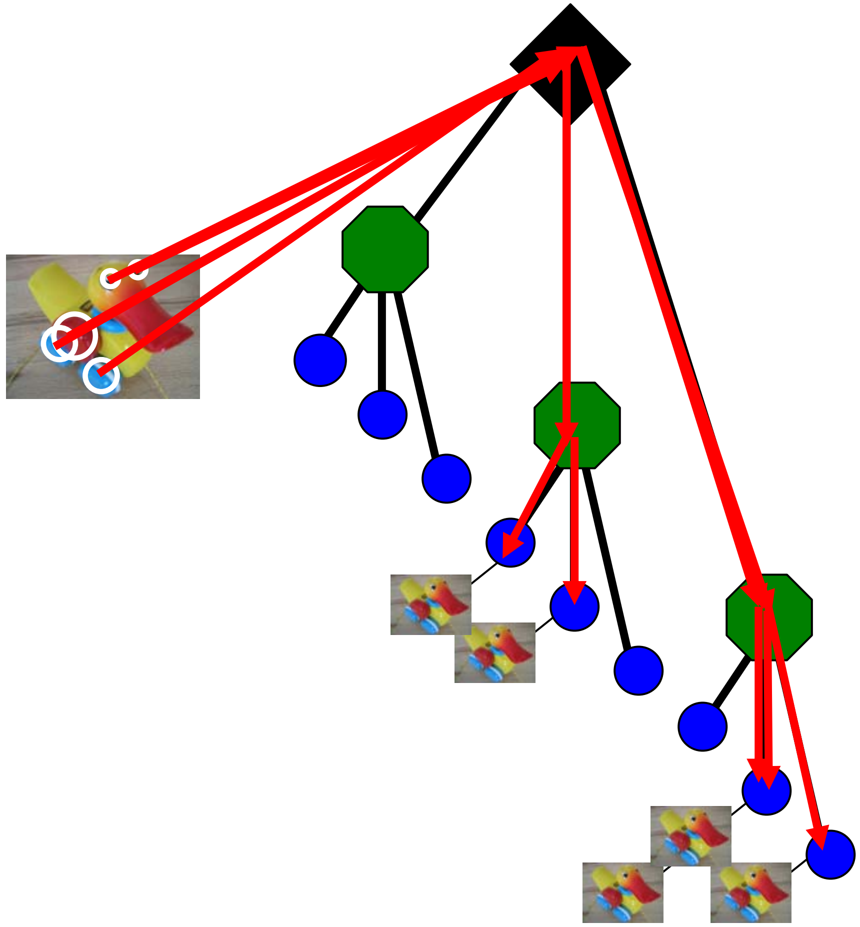
Center for Visualization & Virtual Environments





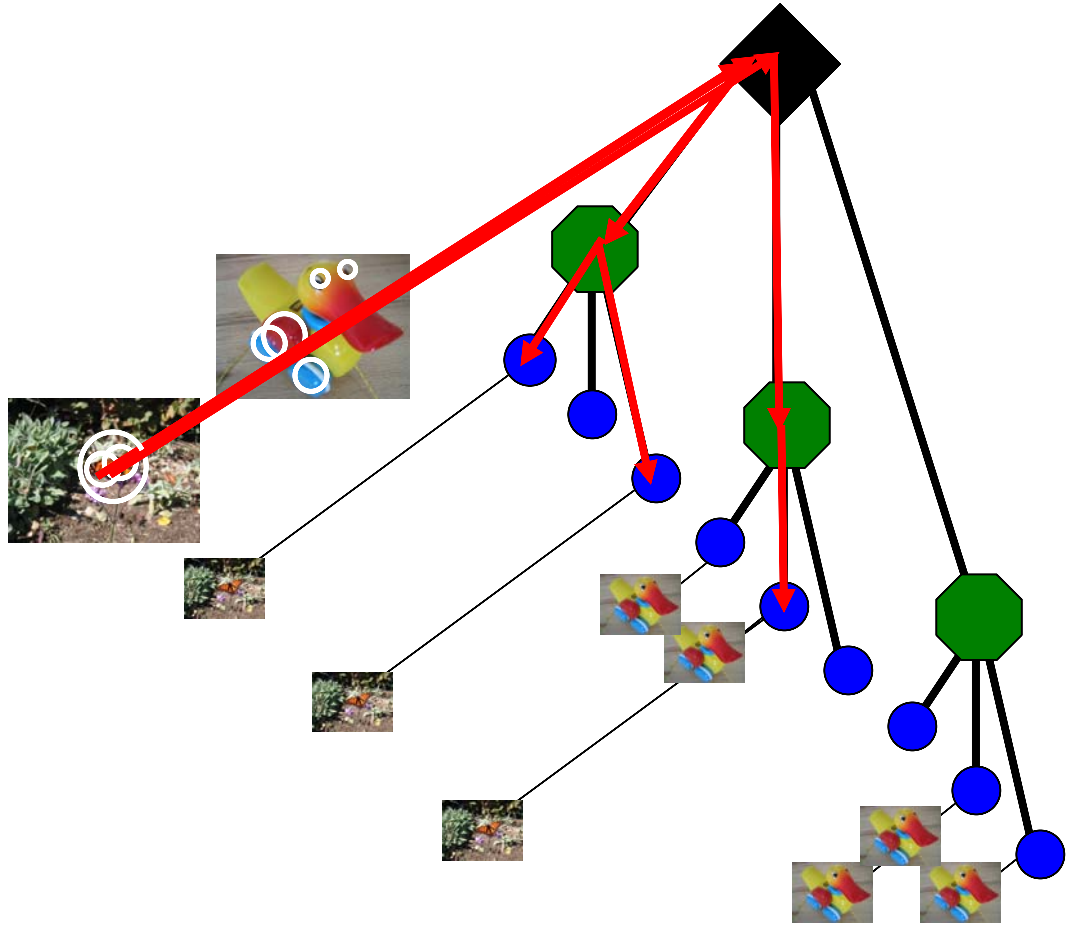


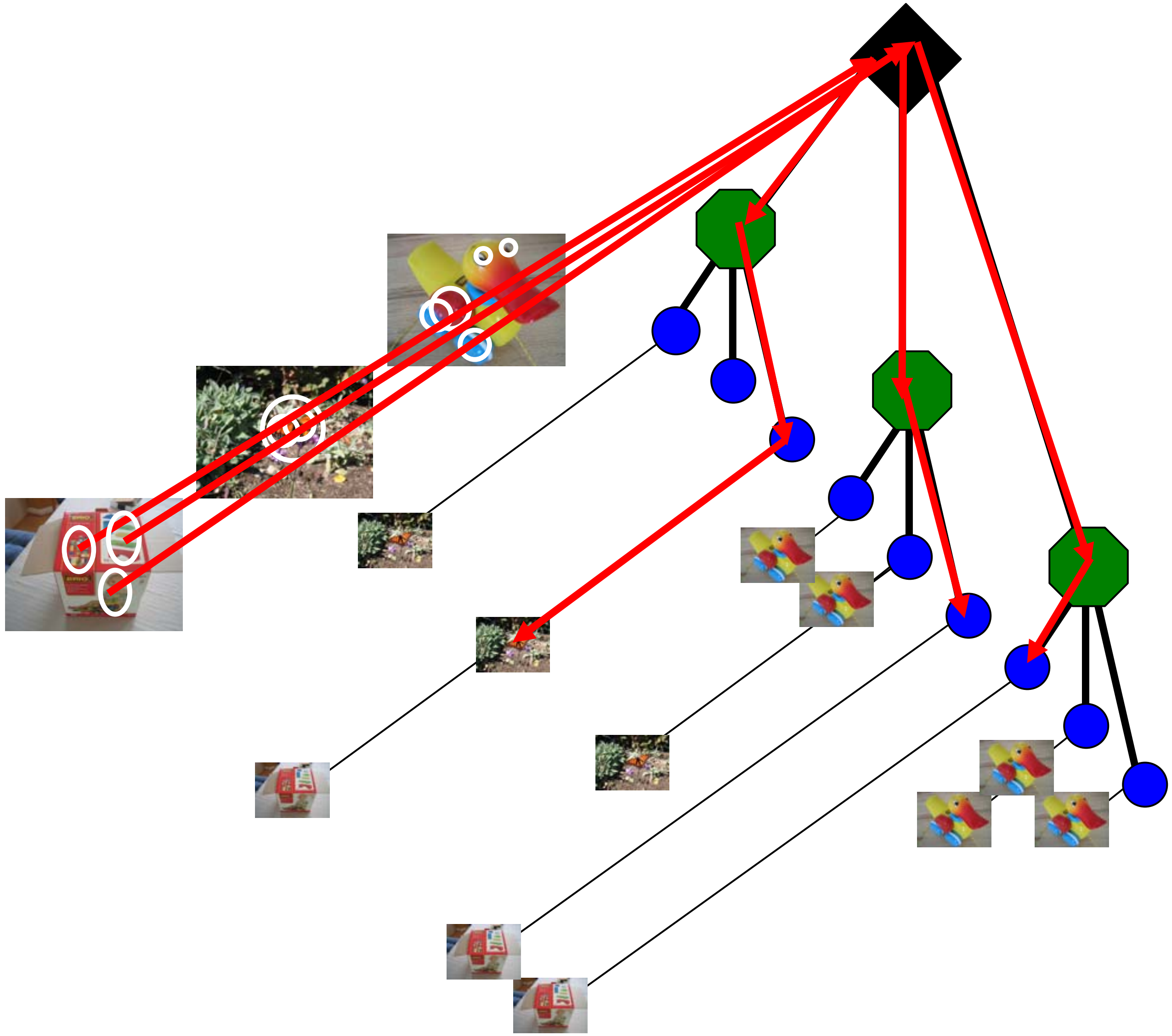


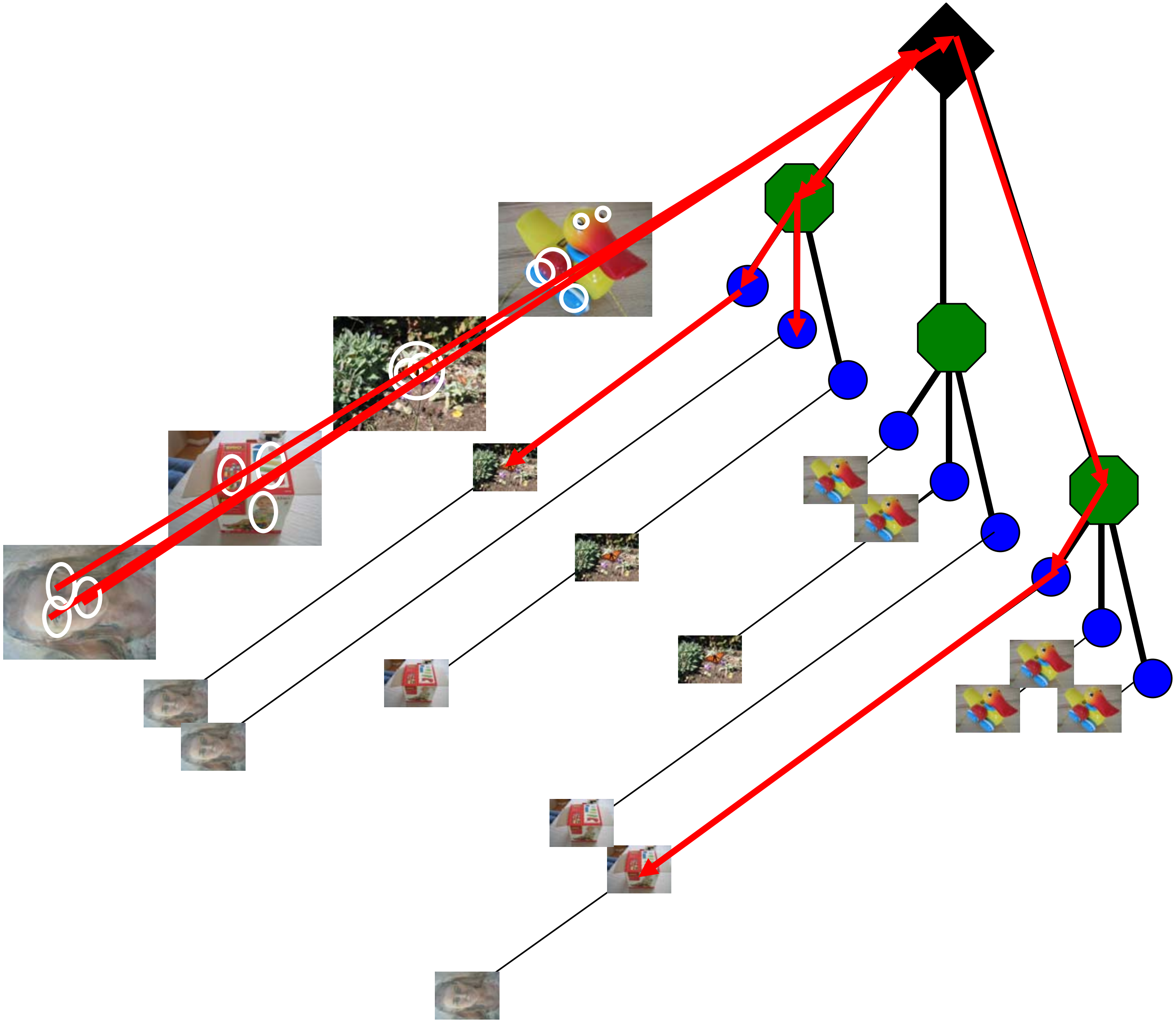


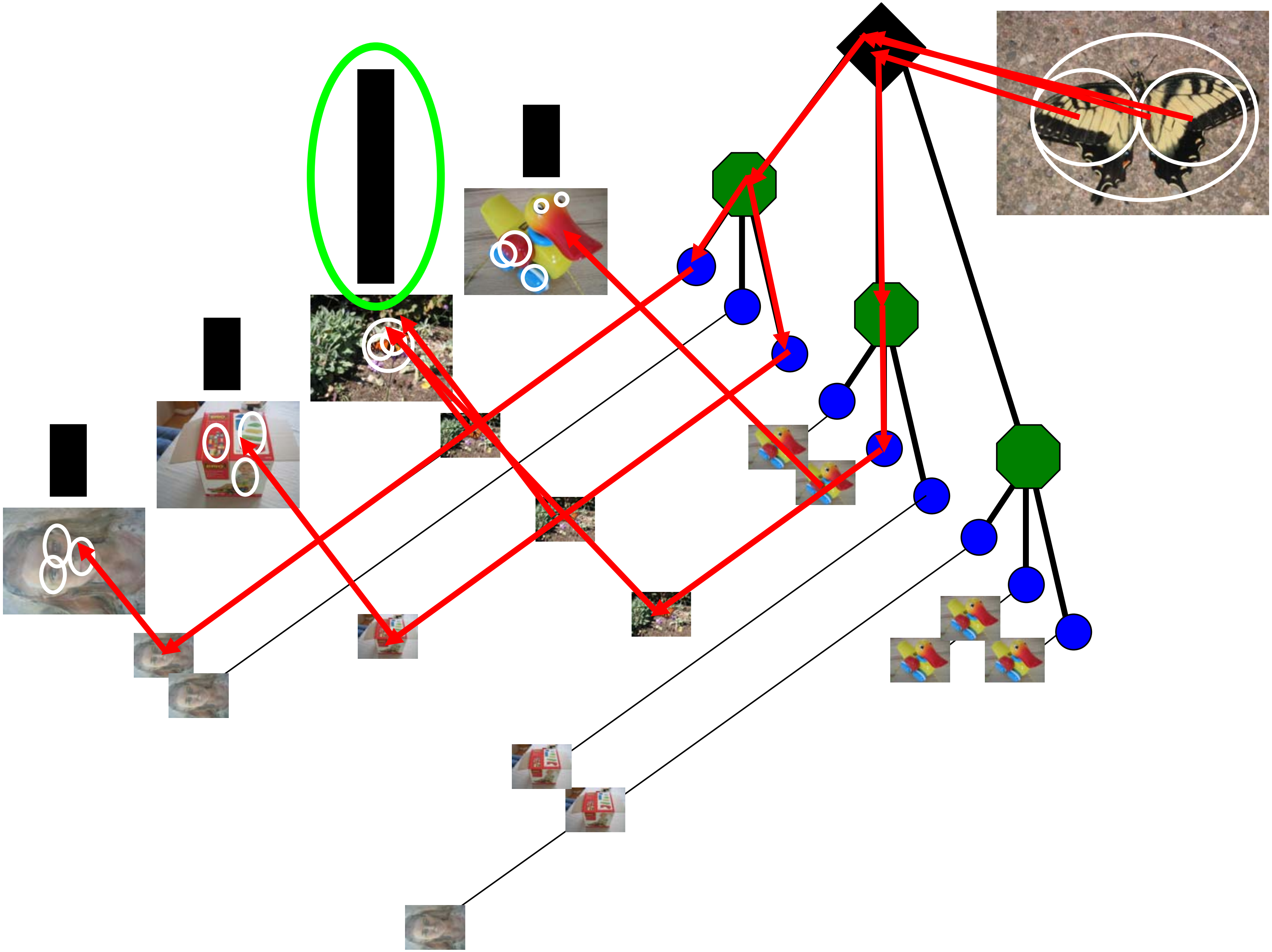
UK

Center for Visualization & Virtual Environments

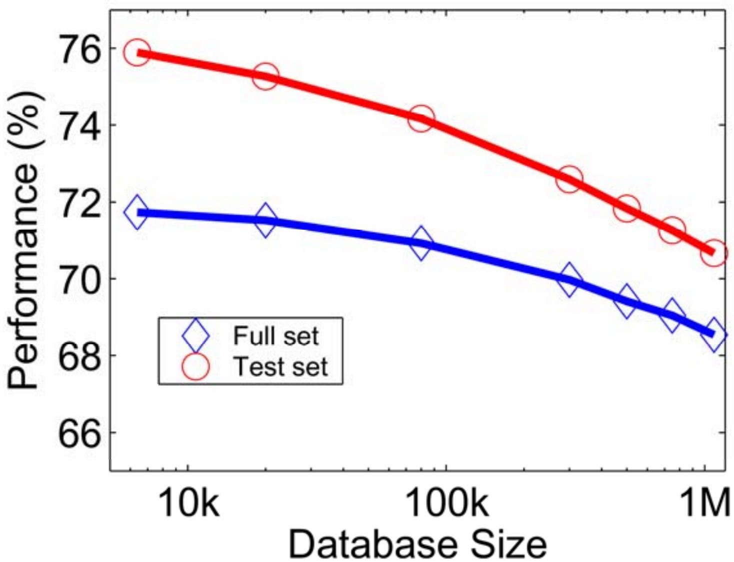


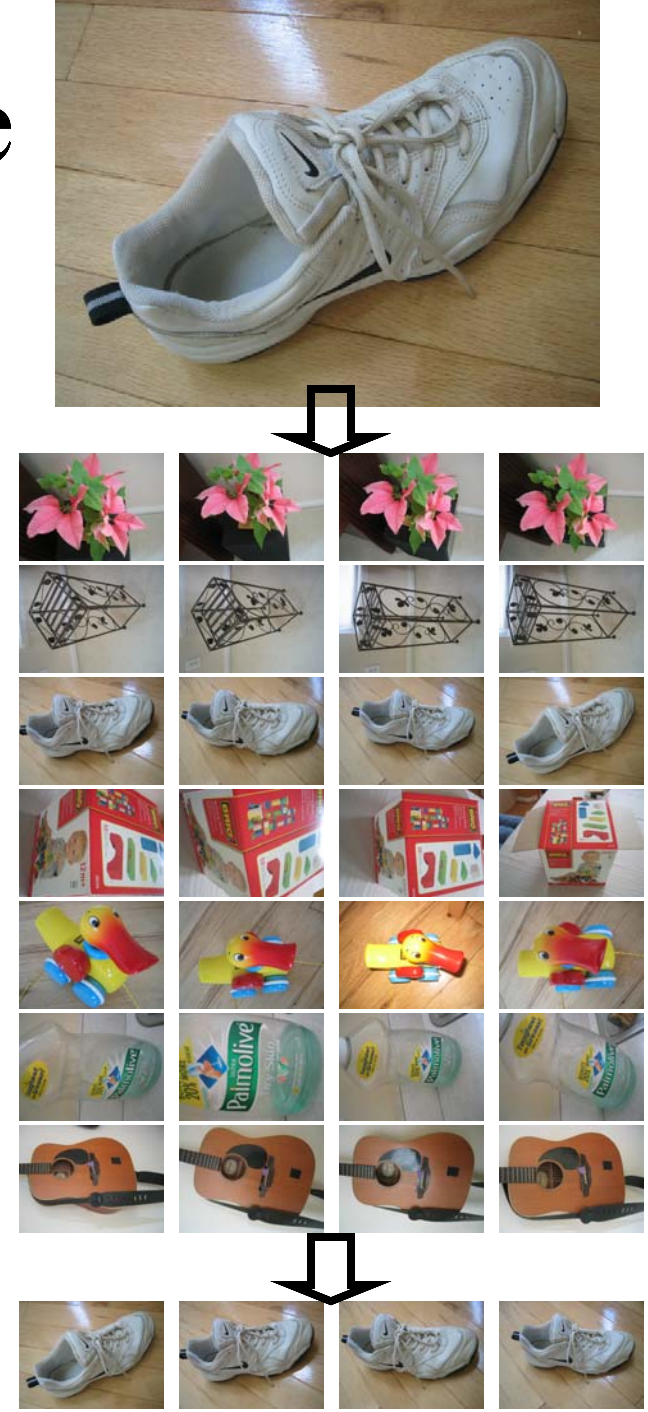






Performance





UK Center for Visualization & Virtual Environments

Recognition Benchmark Images

Henrik Stewénius and David Nistér

The set consists of 2604 groups of 4 images each for a total of 10416 images. All the images are 640x480.

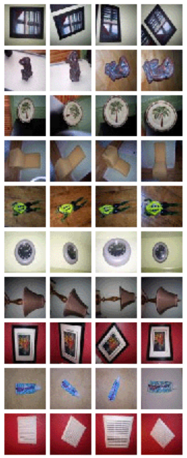
If you use the dataset, please refer to:

- D. Nistér and H. Stewénius, Scalable Recognition with a Vocabulary Tree, CVPR 2006. PDF

Subsets

For users of subsets of the database please note that the difficulty is dependent on the chosen subset. Important factors are:

1. Difficulty of the objects themselves. CD-covers are much easier than flowers. See performance curve below.
2. Sharpness of the images. Many of the indoor images are somewhat blurry and this can affect some algorithms.
3. Similar or identical objects. All the pictures where taken by CS students/faculty/staff and thus keyboards and computer equipment are popular motives. So is computer vision literature.



Download

Please note BEFORE starting your download that the file is almost **2GB**. Please save a local copy in order to save bandwidth at our server.

- Zipped File.

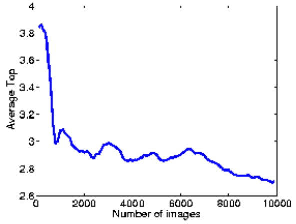
Performance

In the paper we give results either for a subset of 6376 images (all we had at that time) or a smaller subset of 1400 images. The smaller set was used when we did not have an efficient enough implementation in order to handle the larger set.

Performance Measures

- Our simplest measure of performance is to count how many of the 4 images which are top-4 when using a query image from that set of four images.

A matlab implementation which computes this measure: Download.

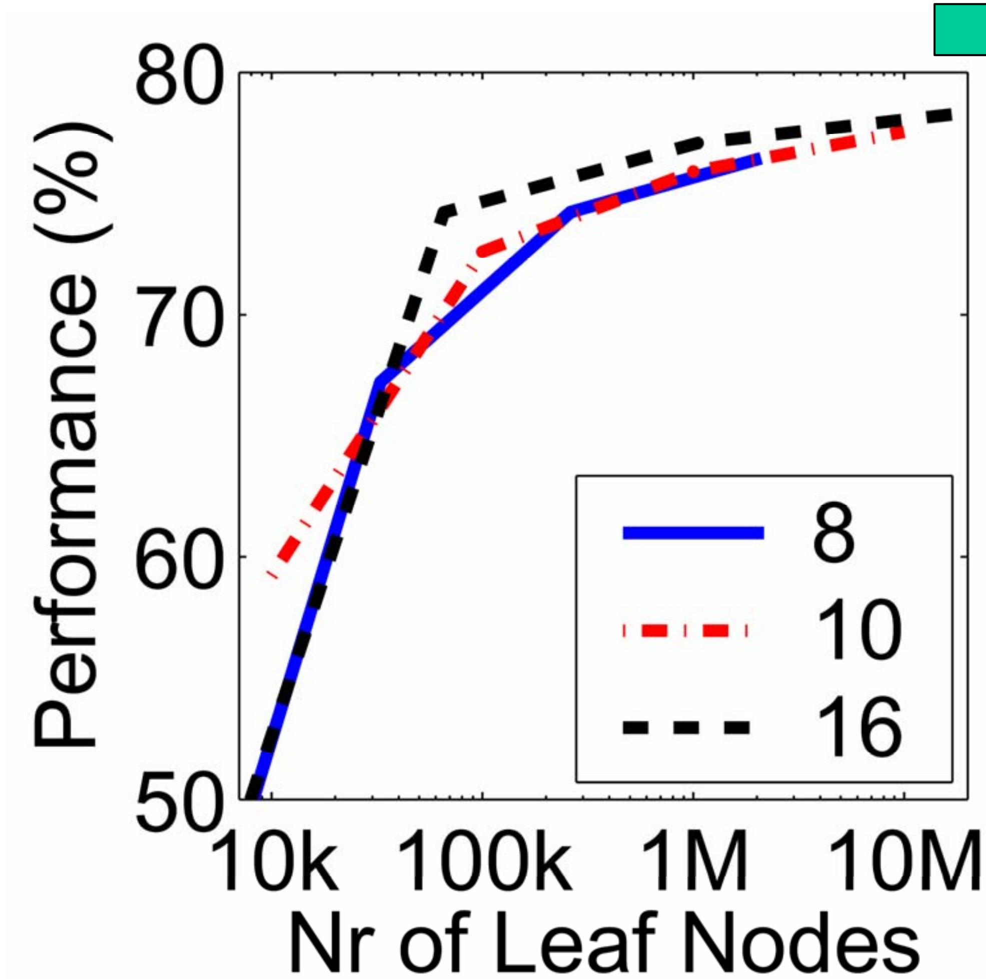


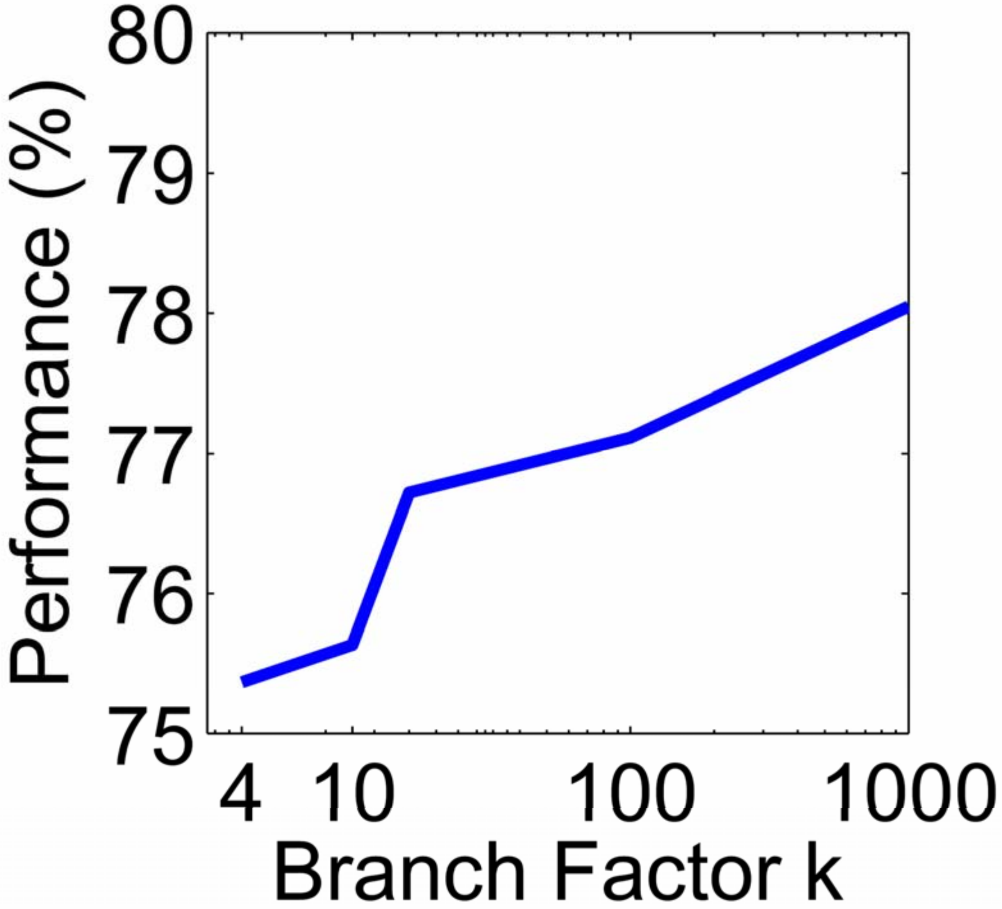
How our performance varies when taking subsets 0:n from the set. These results were run with settings optimized for speed.

Size Matters

Improves Retrieval

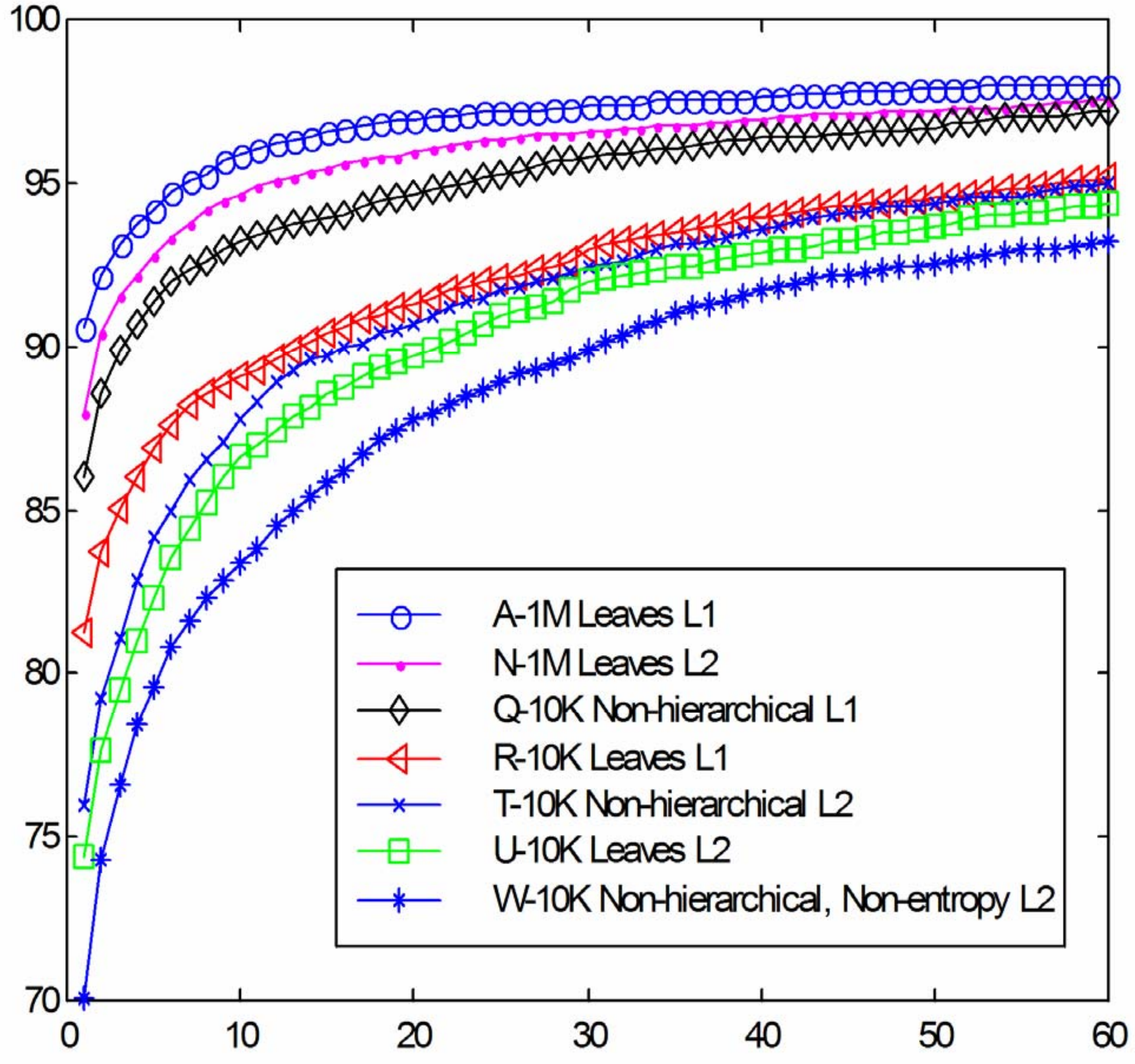
Improves Speed



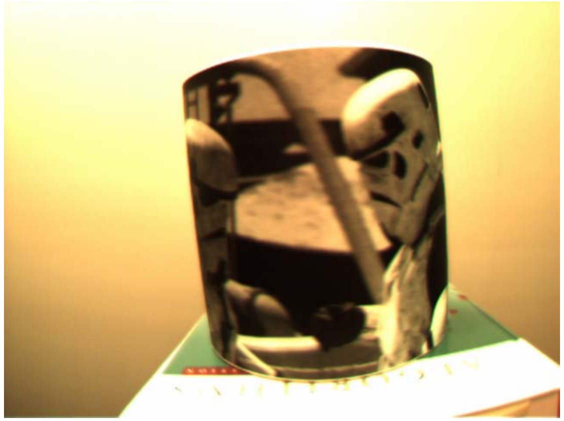


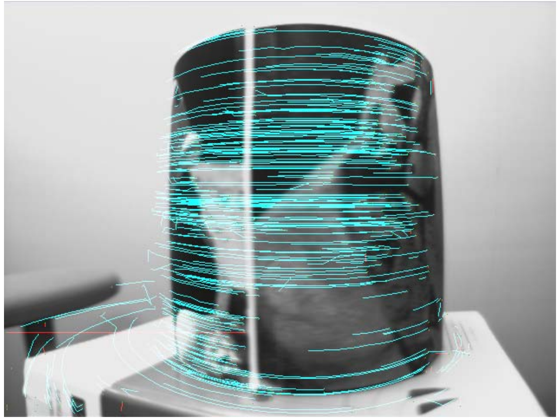
UK

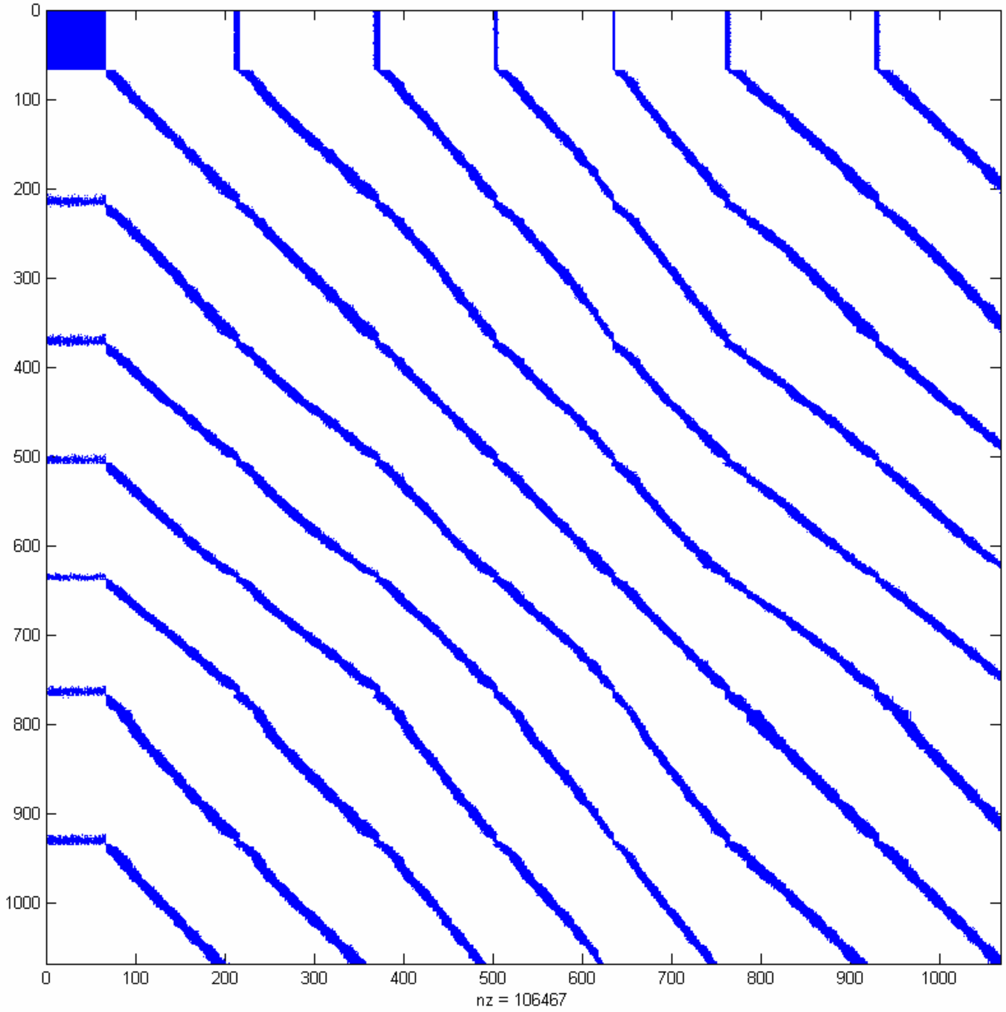
Center for Visualization & Virtual Environments


100
95
90
85
80
75
70
0
10
20
30
40
50
60
A-1M Leaves L1
N-1M Leaves L2
Q-10K Non-hierarchical L1
R-10K Leaves L1
T-10K Non-hierarchical L2
U-10K Leaves L2
W-10K Non-hierarchical, Non-entropy L2

Geometric Verification

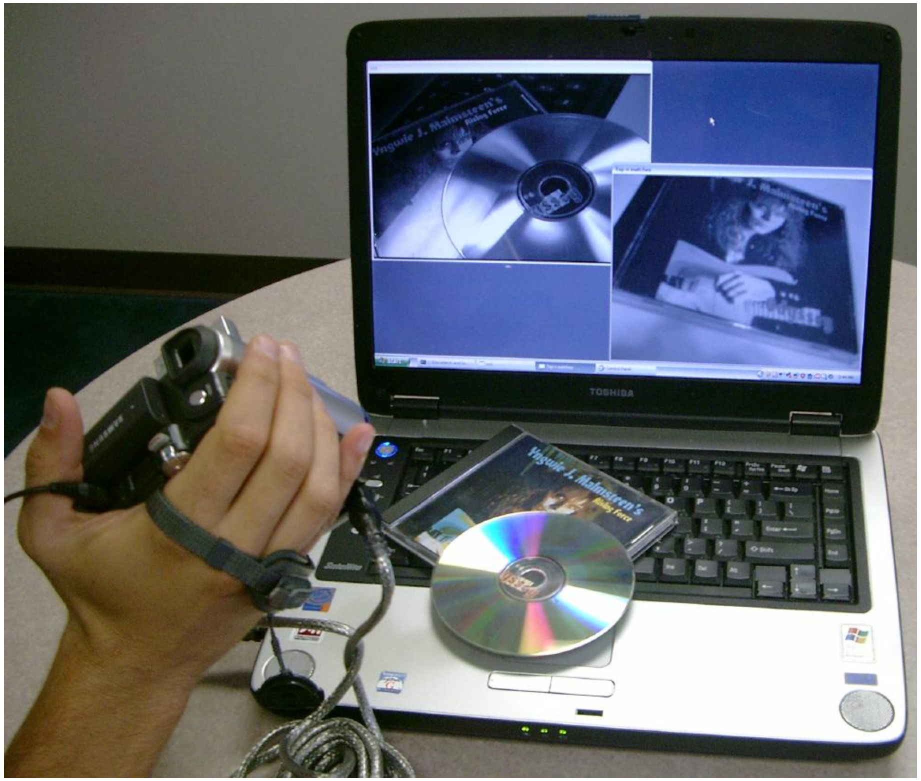






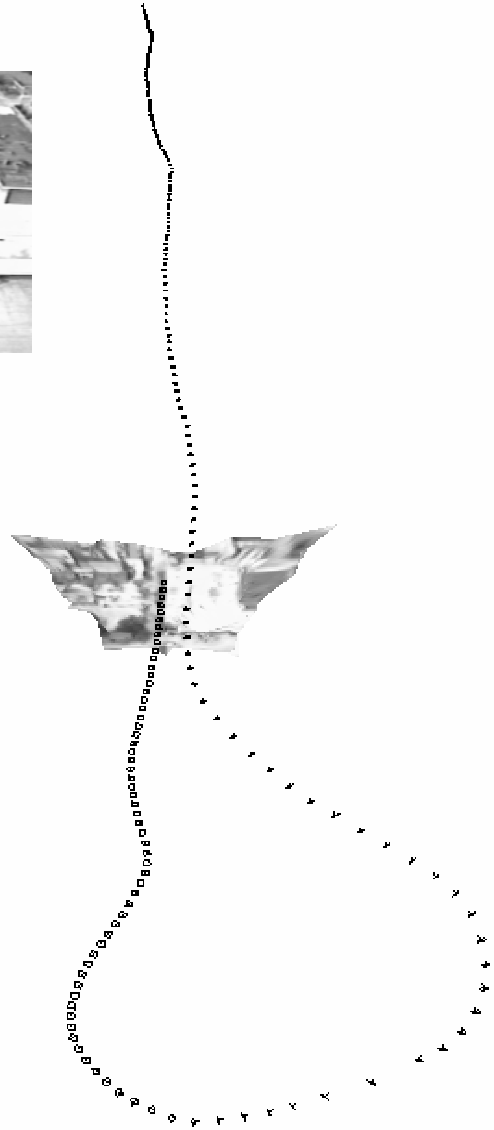
Robust to Clutter and Occlusion

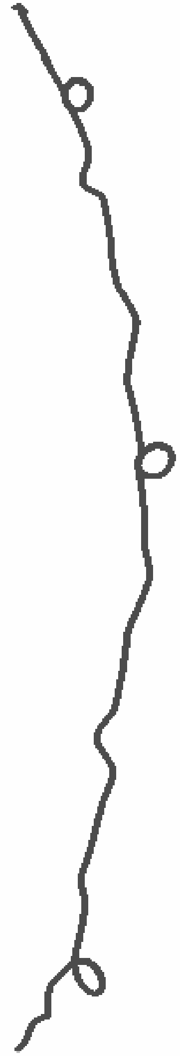
- Local Regions
- Like Web-search



Geometry

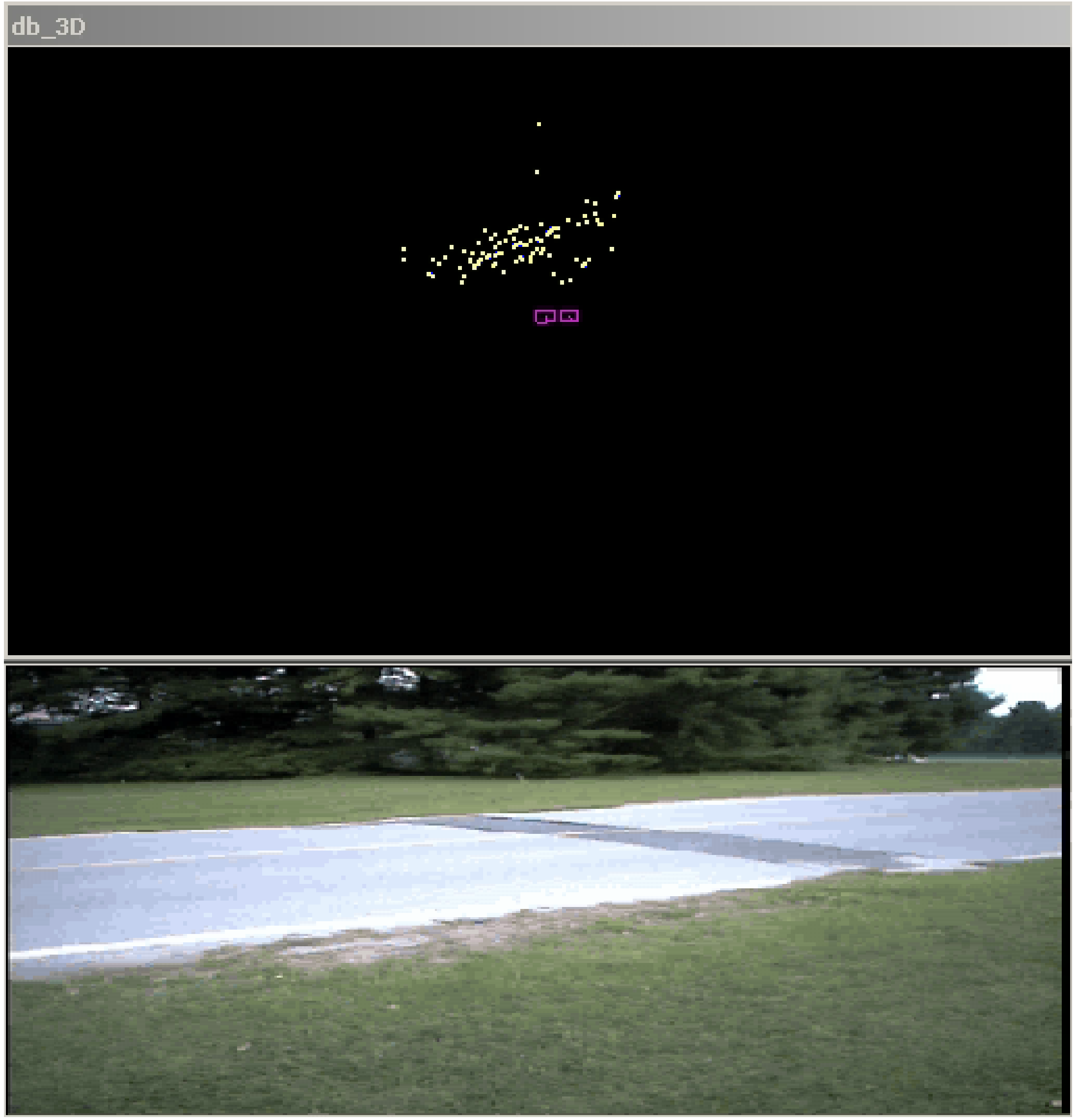
- Demonstration of real-time camera tracking





Visual Odometry

work with Oleg Naroditsky and Jim Bergen

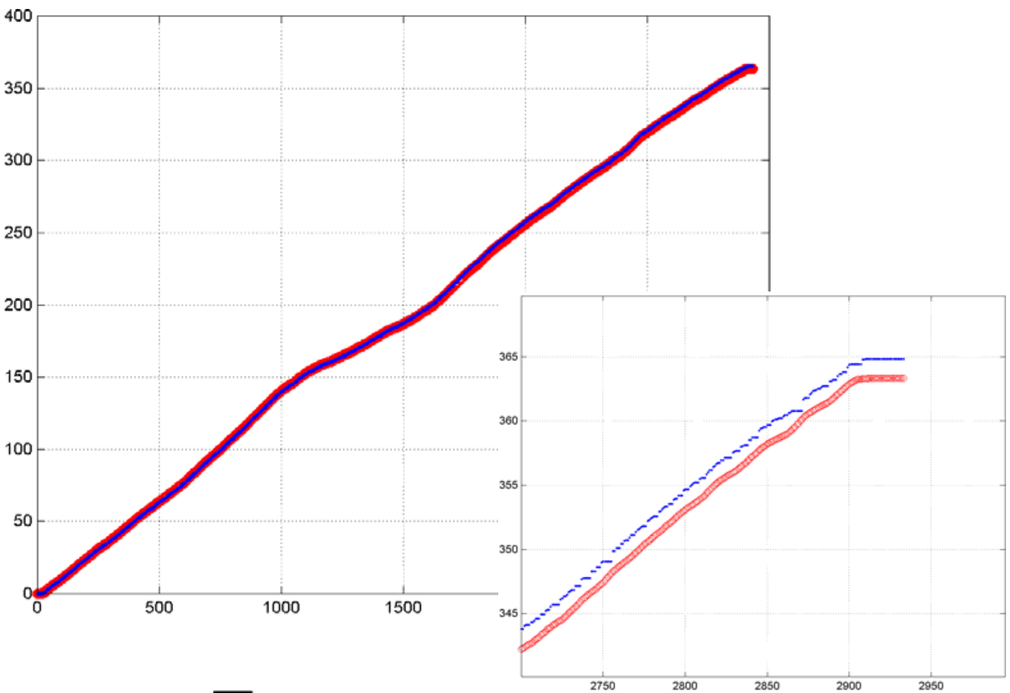


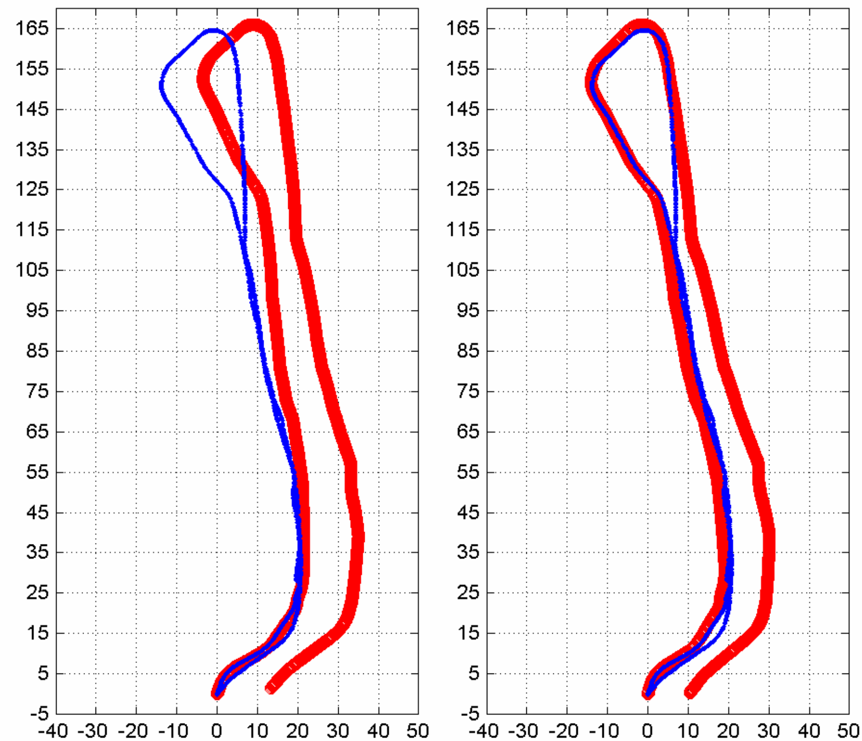
Center for Visualization & Virtual Environments

Visual Odometry

work with Oleg Naroditsky and Jim Bergen

- 365 m without loss of tracking
- 350 m (~ 3.5 minutes) without GPS
- Error in distance traveled ~ 1%
- Accumulated error in position ~ 3-5%
  - e.g. ~10m over ~350m





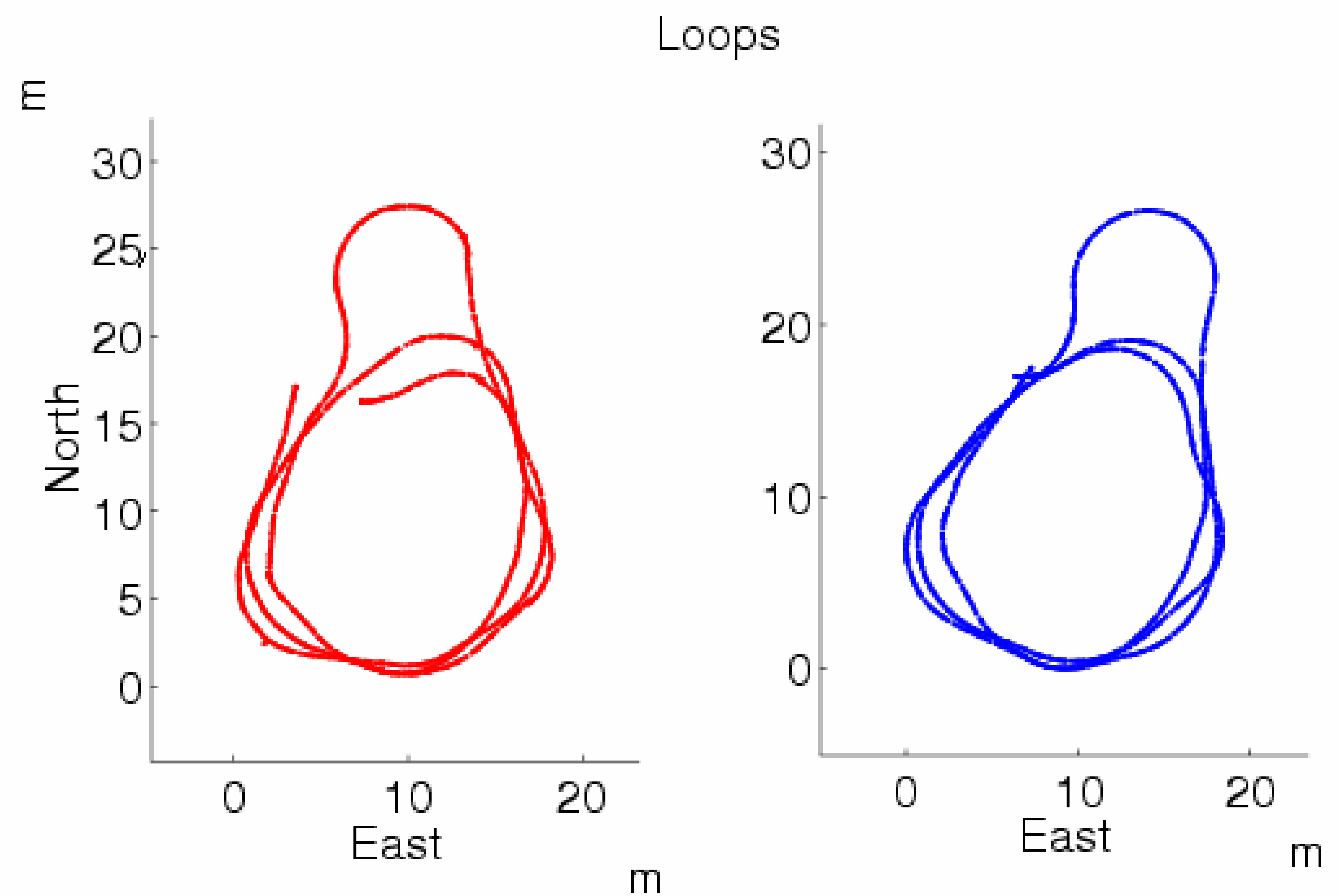
East

UK

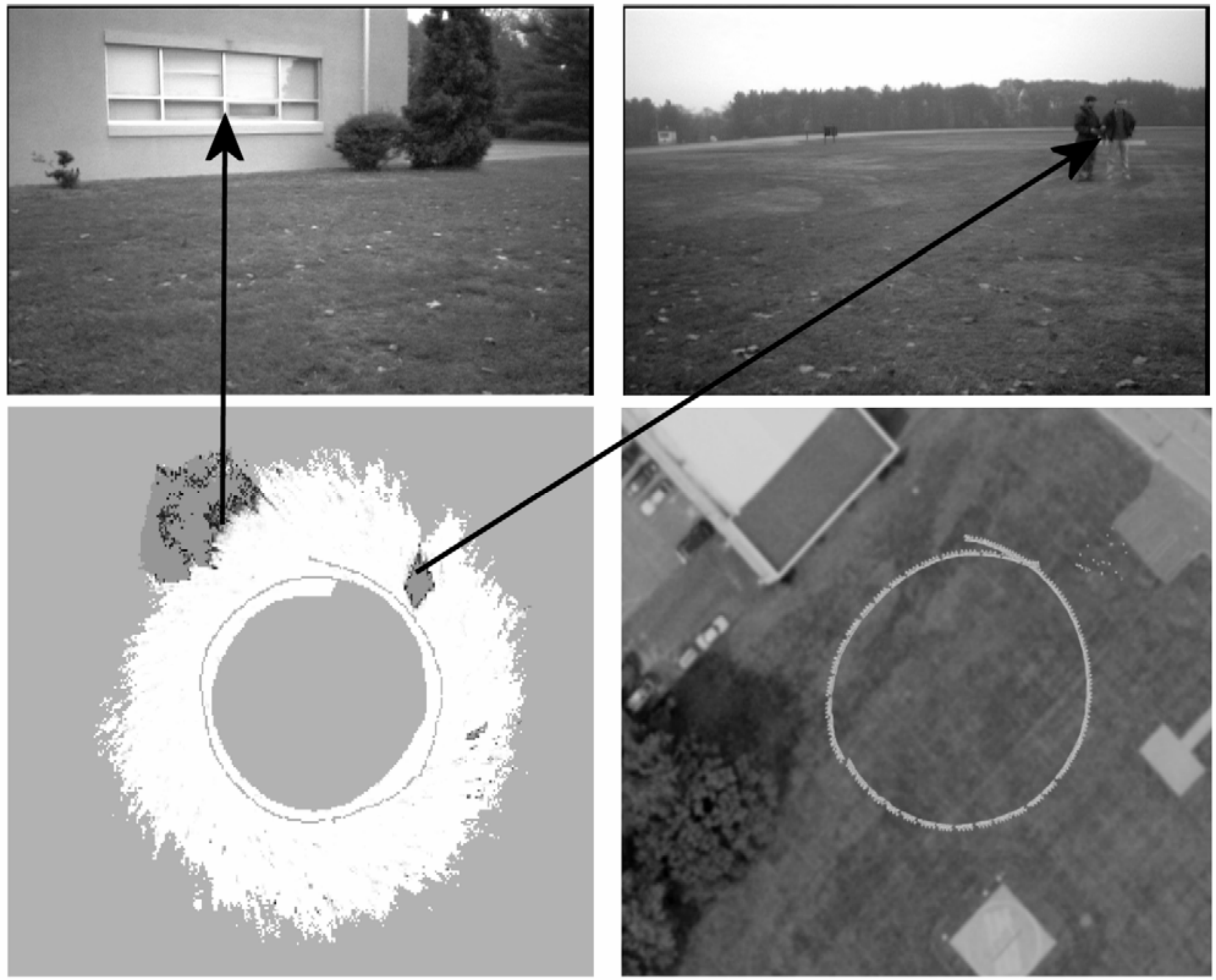
Center for Visualization & Virtual Environments

Visual Odometry

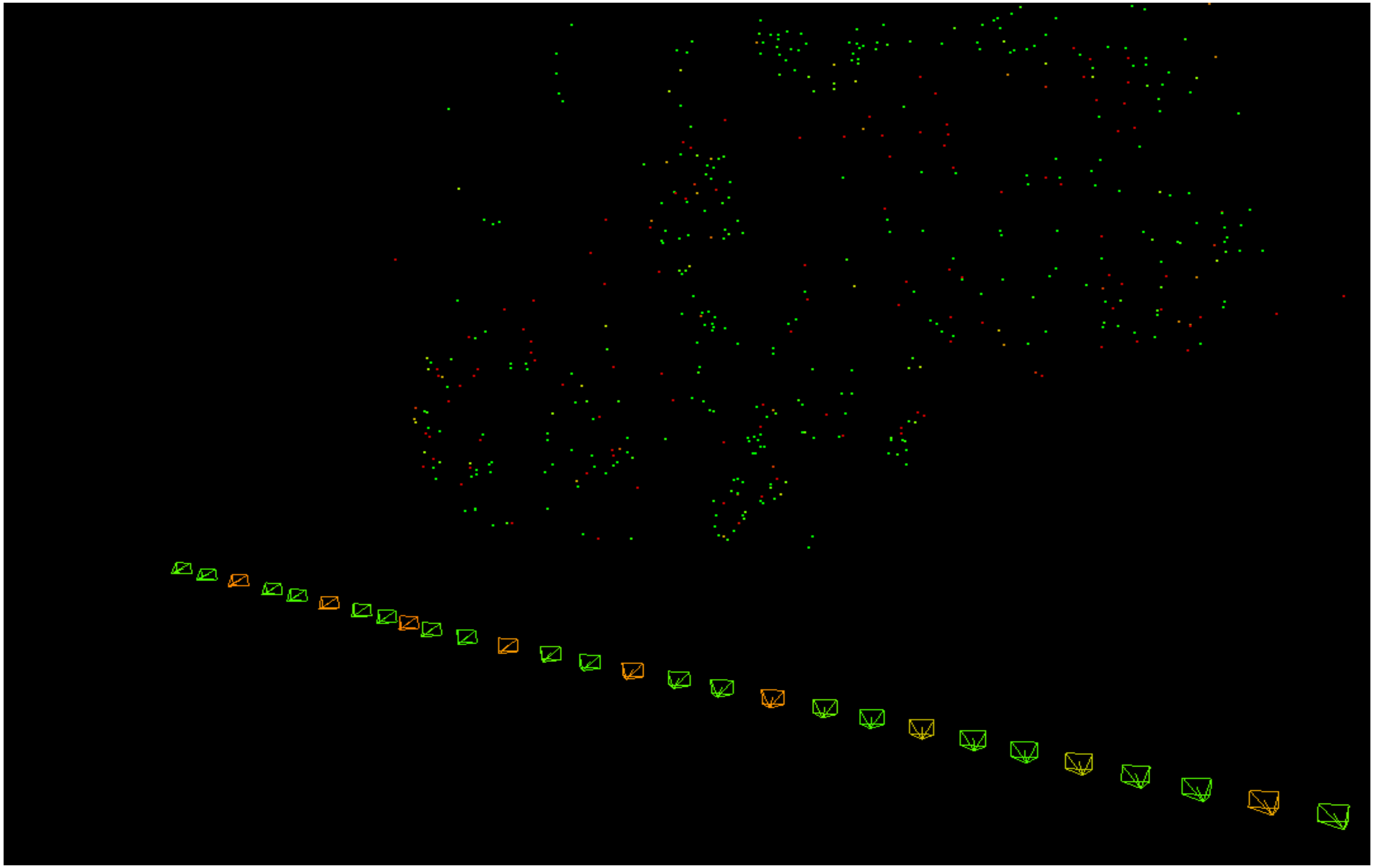
work with Oleg Naroditsky and Jim Bergen



Visual Odometry
work with Oleg Naroditsky and Jim Bergen

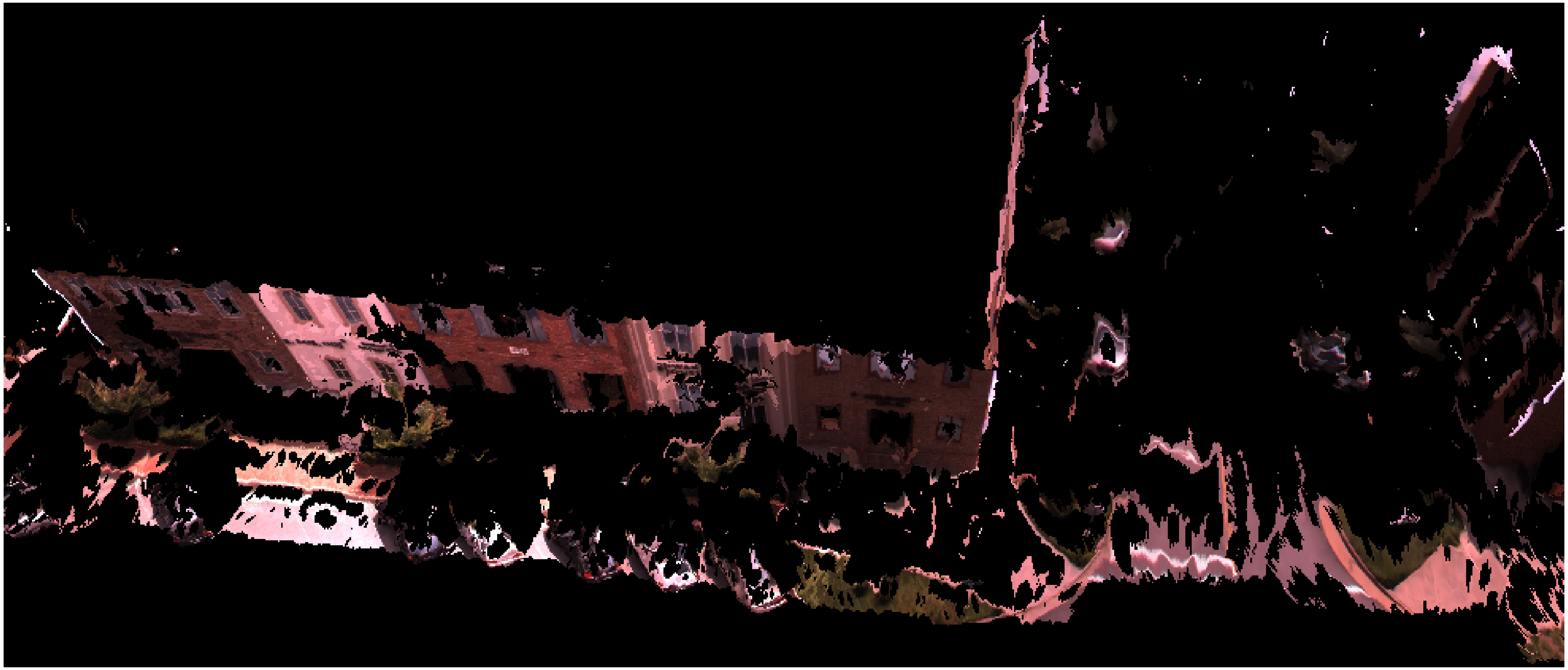


3D Tracker

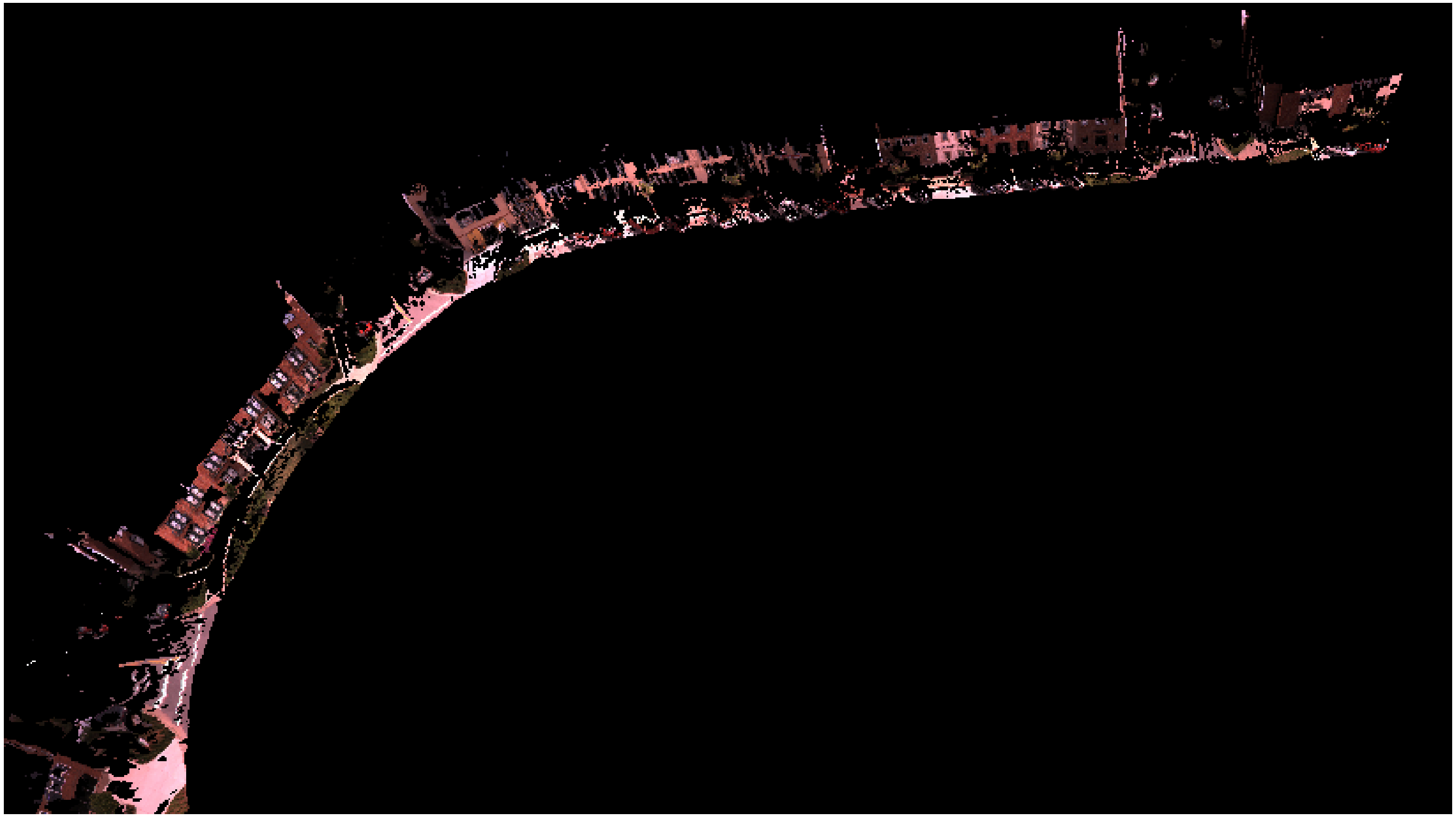


UK

Center for Visualization & Virtual Environments



Center for Visualization & Virtual Environments



UK Center for Visualization & Virtual Environments

UK Center for Visualization & Virtual Environments

Geo Registered Cameras (With INS Data)

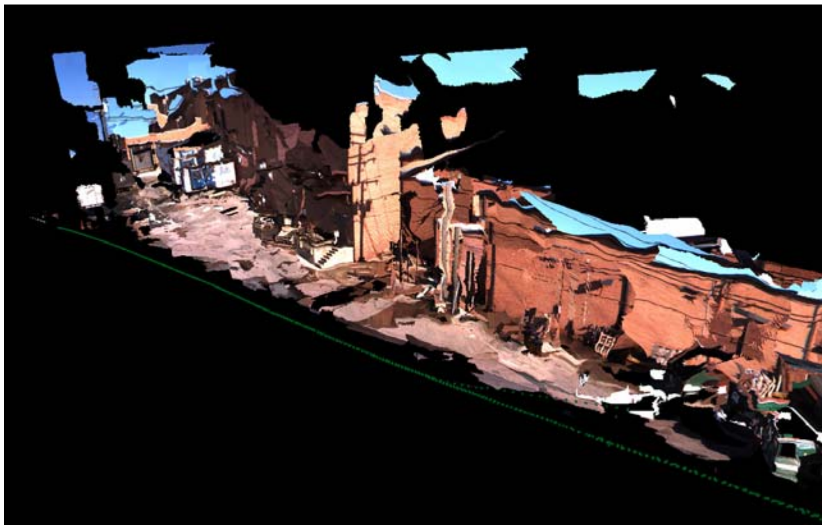
Center for Visualization & Virtual Environments

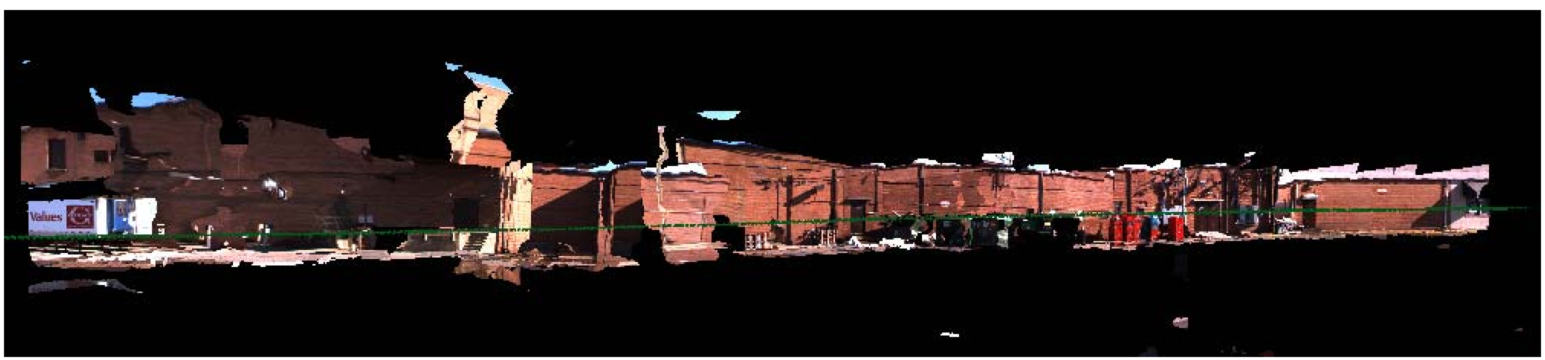
GPS Data Gathering

- Garmin GPS16
  - $200 unit
  - 1Hz updates
- Records
  - Latitude-Longitude
  - Pseudo-range up to 12 satellites
  - Satellite's clock

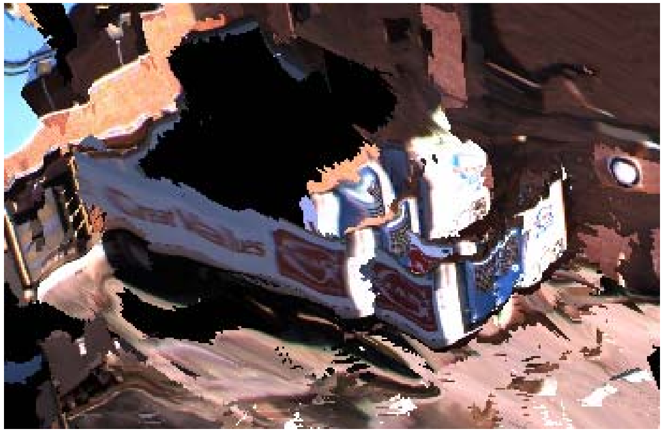
UK Center for Visualization & Virtual Environments

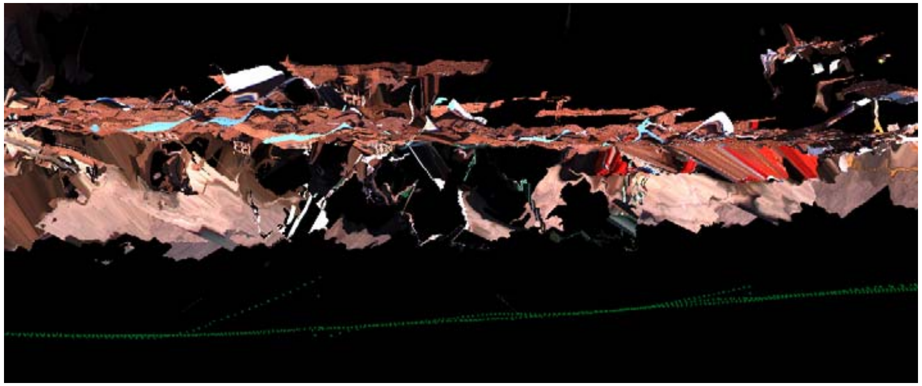
3D Tracking and Geo-registration

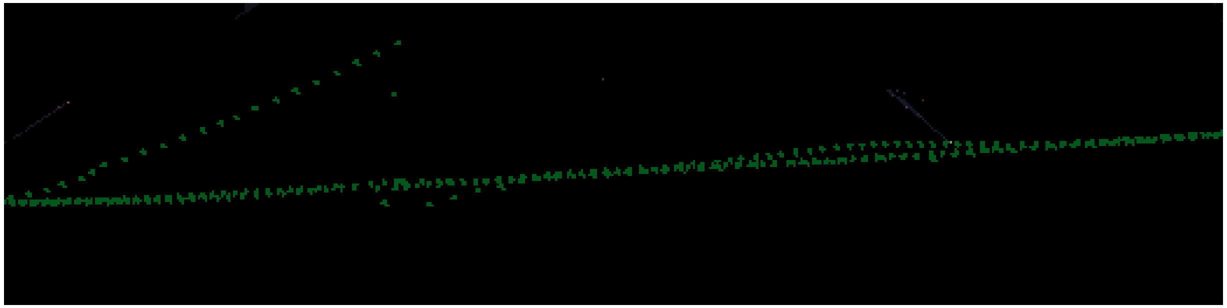




3D Tracking and Geo-registration





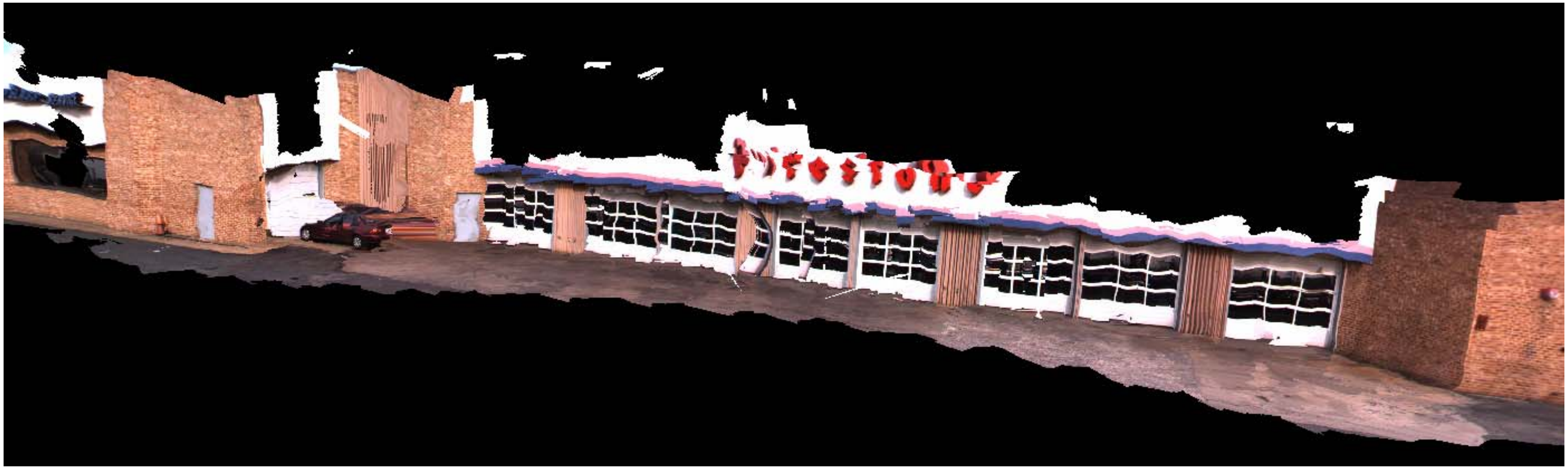


Lever arm calibration

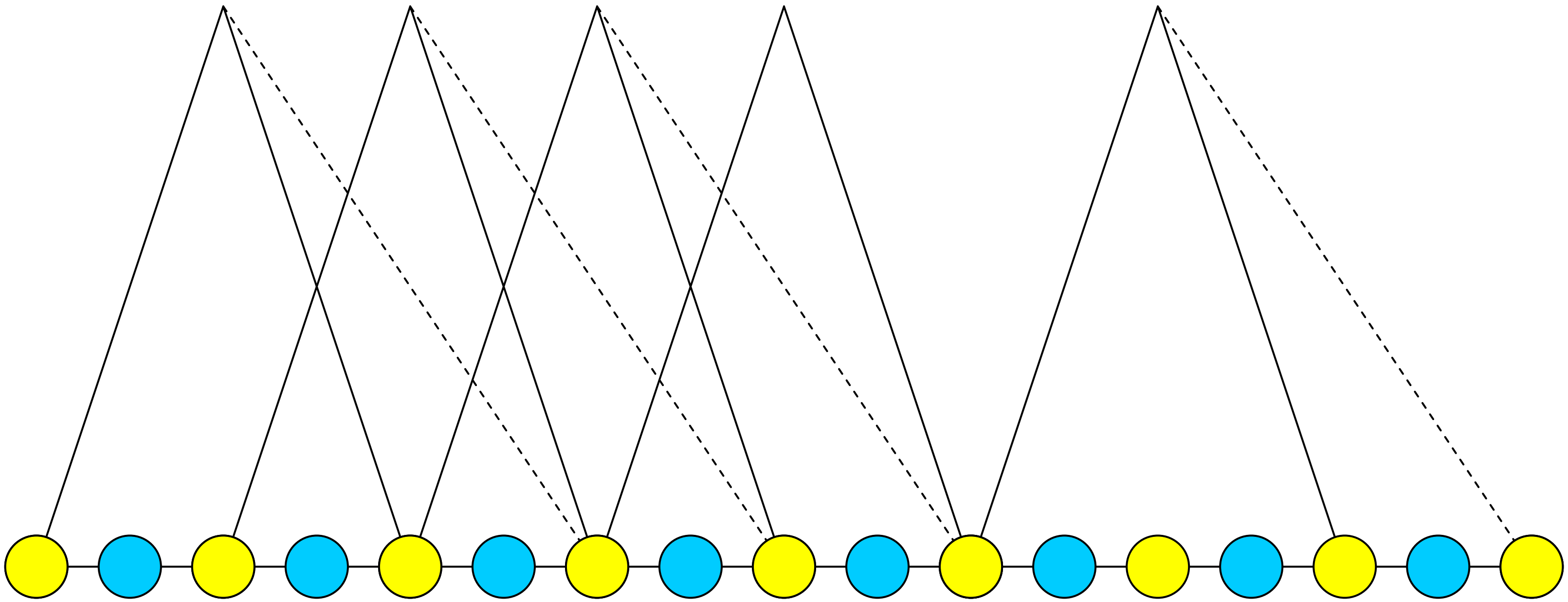
lever arm from drawings

refined lever arm

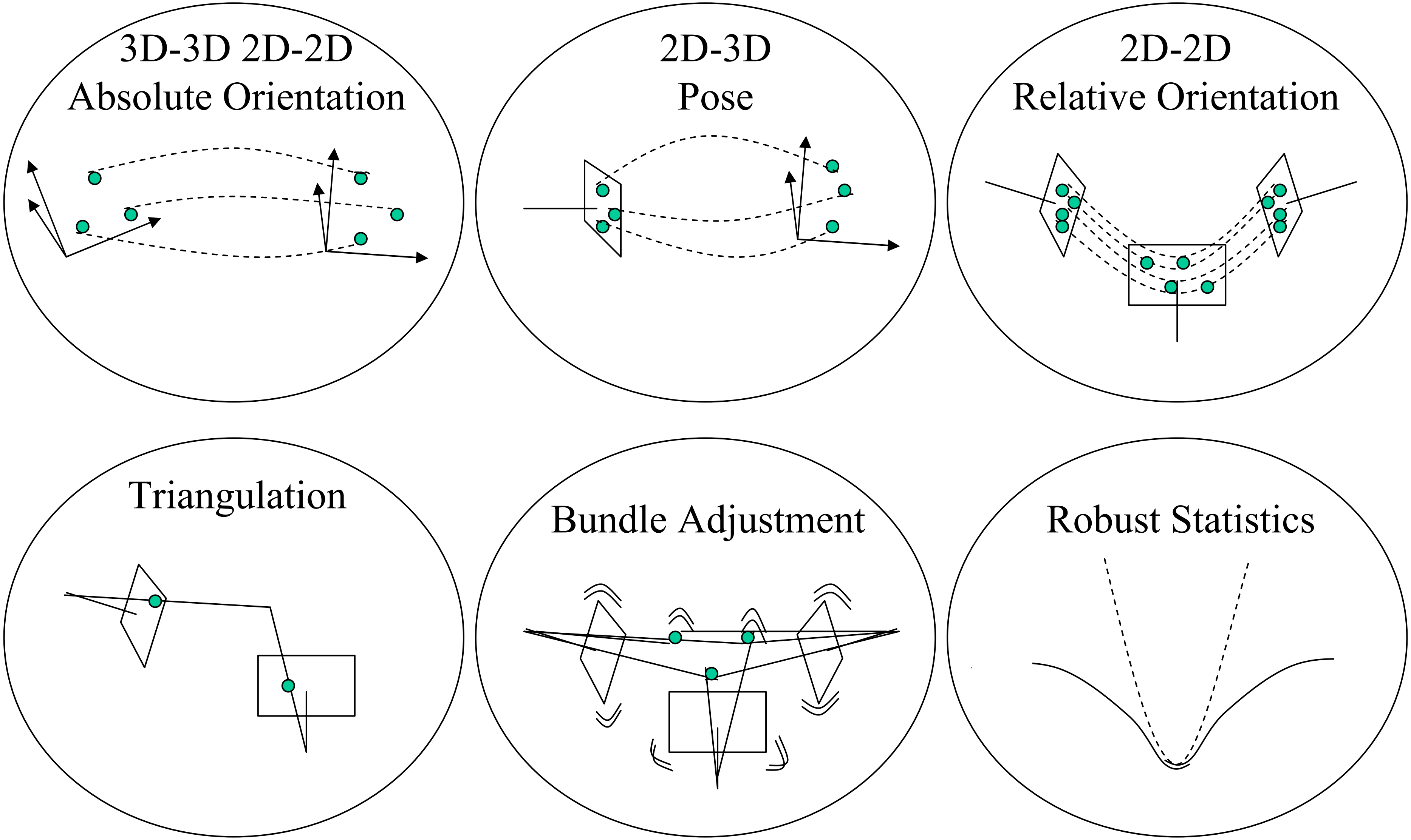
Lever arm calibration



UK Center for Visualization & Virtual Environments



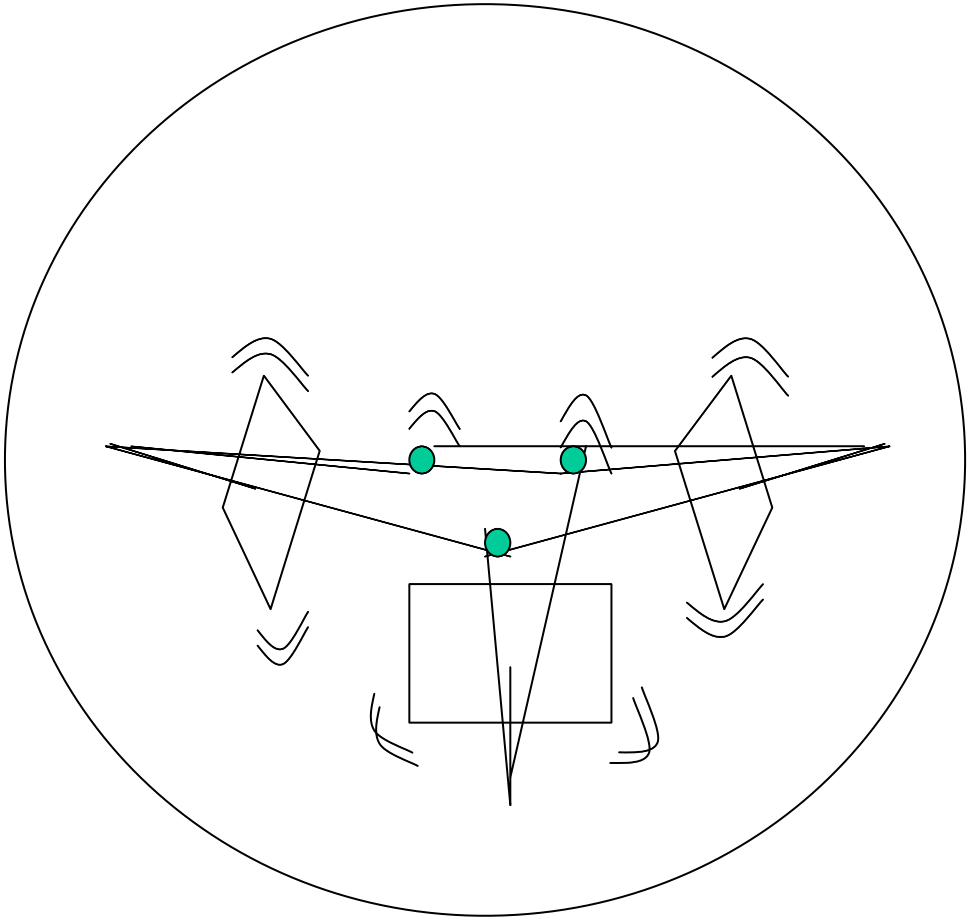
Geometry Tools



UK

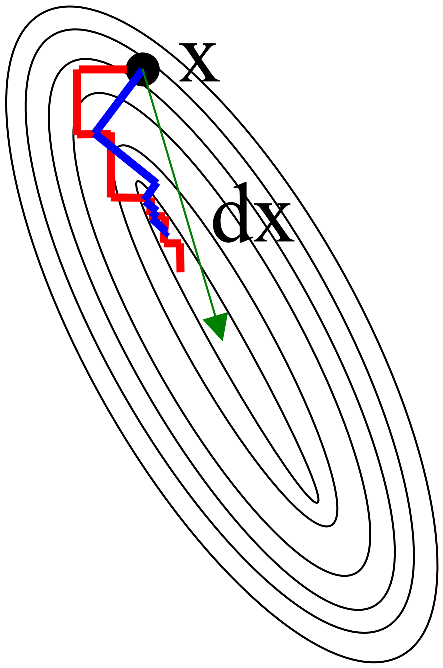
Center for Visualization & Virtual Environments

Bundle Adjustment



Trust Region Methods

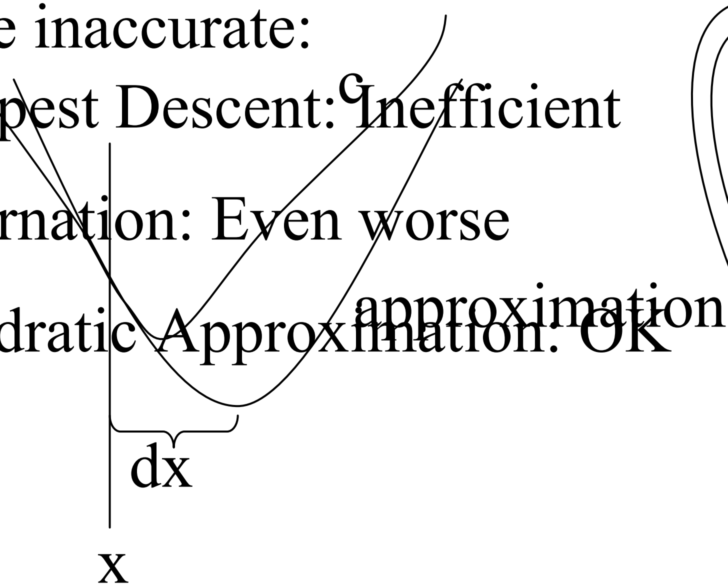
- Steepest Descent: Inefficient
- Alternation: Even worse
- Quadratic Approximation: OK

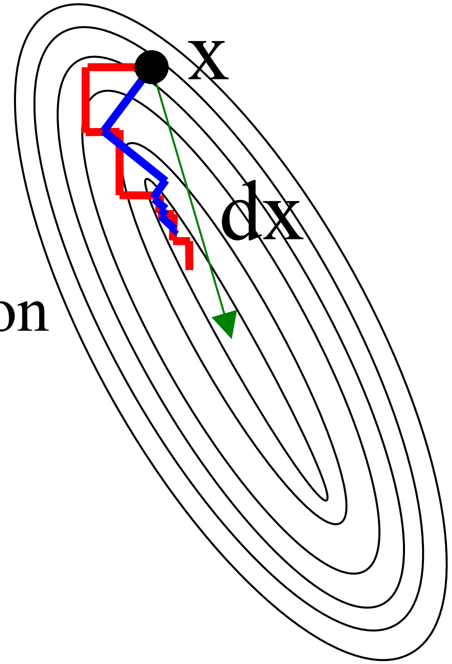


Trust Region Methods

Can be inaccurate:

- Steepest Descent: Inefficient
- Alternation: Even worse
- Quadratic Approximation: OK





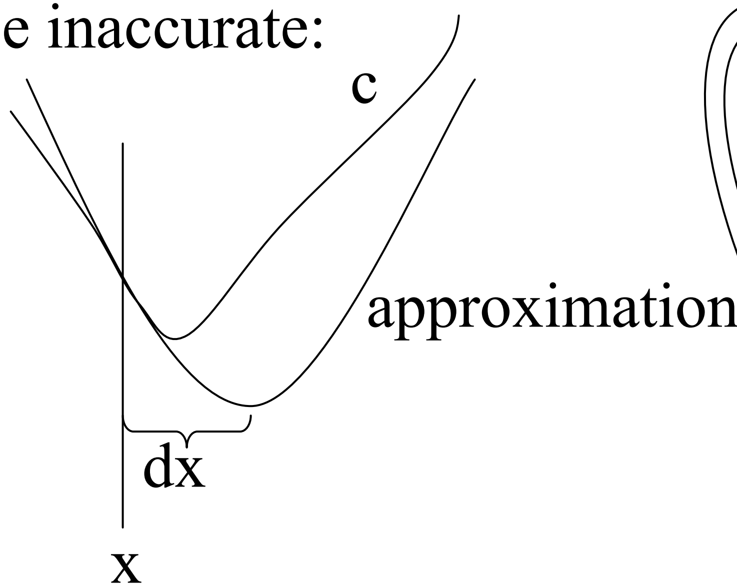
Quadratic approximation:

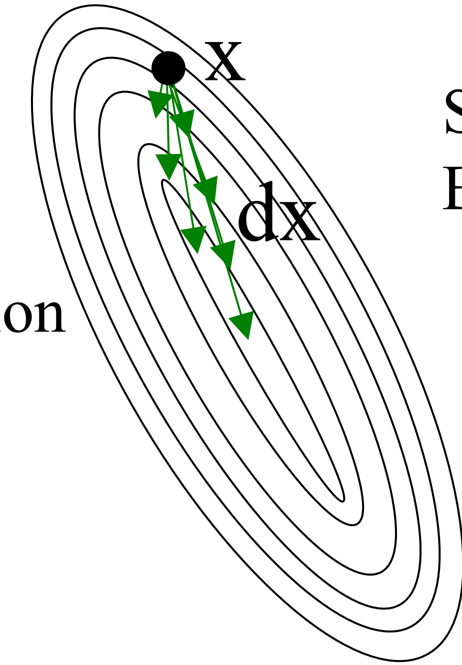
$$c(x+dx) \approx c(x) + \nabla c^{T}(x)dx + dx^{T}H_{c}(x)dx$$

If accurate, then $H_{c}(x)dx = -\frac{1}{2}\nabla c(x)$ at minimum.

Trust Region Methods

Can be inaccurate:





Solution:
Back down dx

Quadratic approximation:

$$c(x+dx) \approx c(x) + \nabla c^{T}(x)dx + dx^{T}H_{c}(x)dx$$

If accurate, then $H_{c}(x)dx = -\frac{1}{2}\nabla c(x)$ at minimum.

Quadratic approximation:

$$c(x+dx) \approx c(x) + \nabla c^{\mathrm{T}}(x)dx + dx^{\mathrm{T}} H_c(x)dx$$

If accurate, then $H_c(x)dx = -\frac{1}{2}\nabla c(x)$ at minimum.

Quadratic approximation:

$$c(x+dx) \approx c(x) + \nabla c^{\mathrm{T}}(x)dx + dx^{\mathrm{T}} H_c(x) dx$$

If accurate, then $H_c(x)dx = -\frac{1}{2}\nabla c(x)$ at minimum.

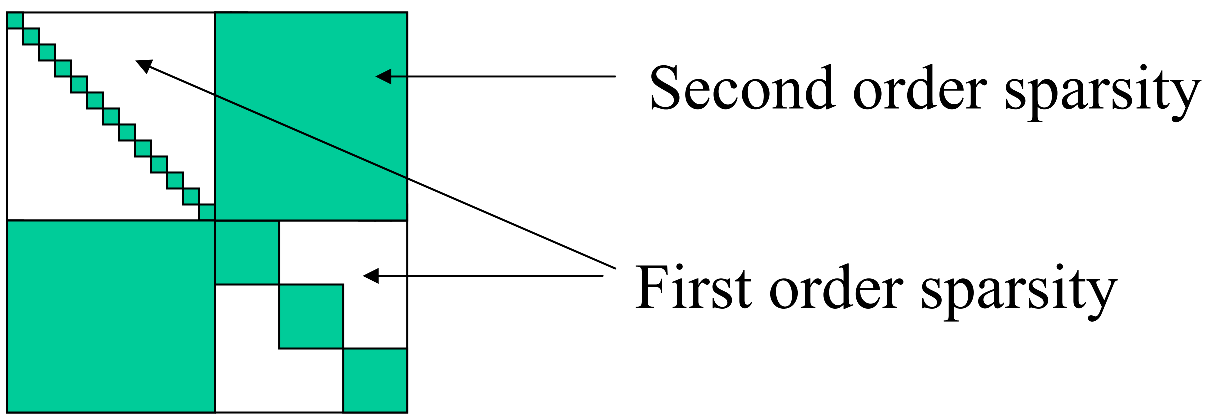
Bundle Adjustment

$$H_c(x)dx = -\tfrac{1}{2}\nabla c(x)$$

Block LU factorization:

Multiply by $\begin{bmatrix} H_{SS}^{-1} & 0 \\ 0 & I \end{bmatrix}$ Multiply by $\begin{bmatrix} I & 0 \\ -H_{CS} & I \end{bmatrix}$

$$\begin{bmatrix} H_{SS} & H_{SC} \\ H_{CS} & H_{CC} \end{bmatrix}\begin{bmatrix} dx_S \\ dx_C \end{bmatrix} = \begin{bmatrix} g_S \\ g_C \end{bmatrix}$$



Bundle Adjustment

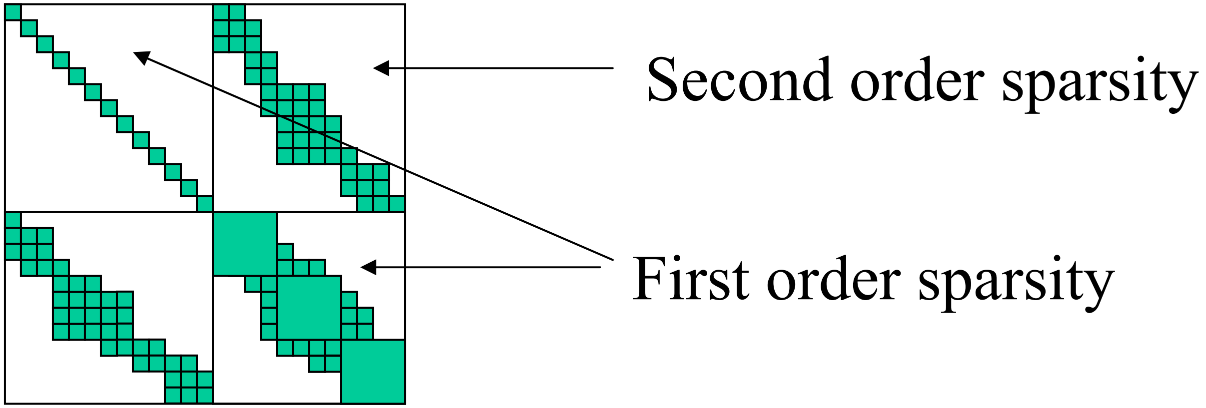
$$H_c(x)dx = -\tfrac{1}{2}\nabla c(x)$$

Block LU factorization:

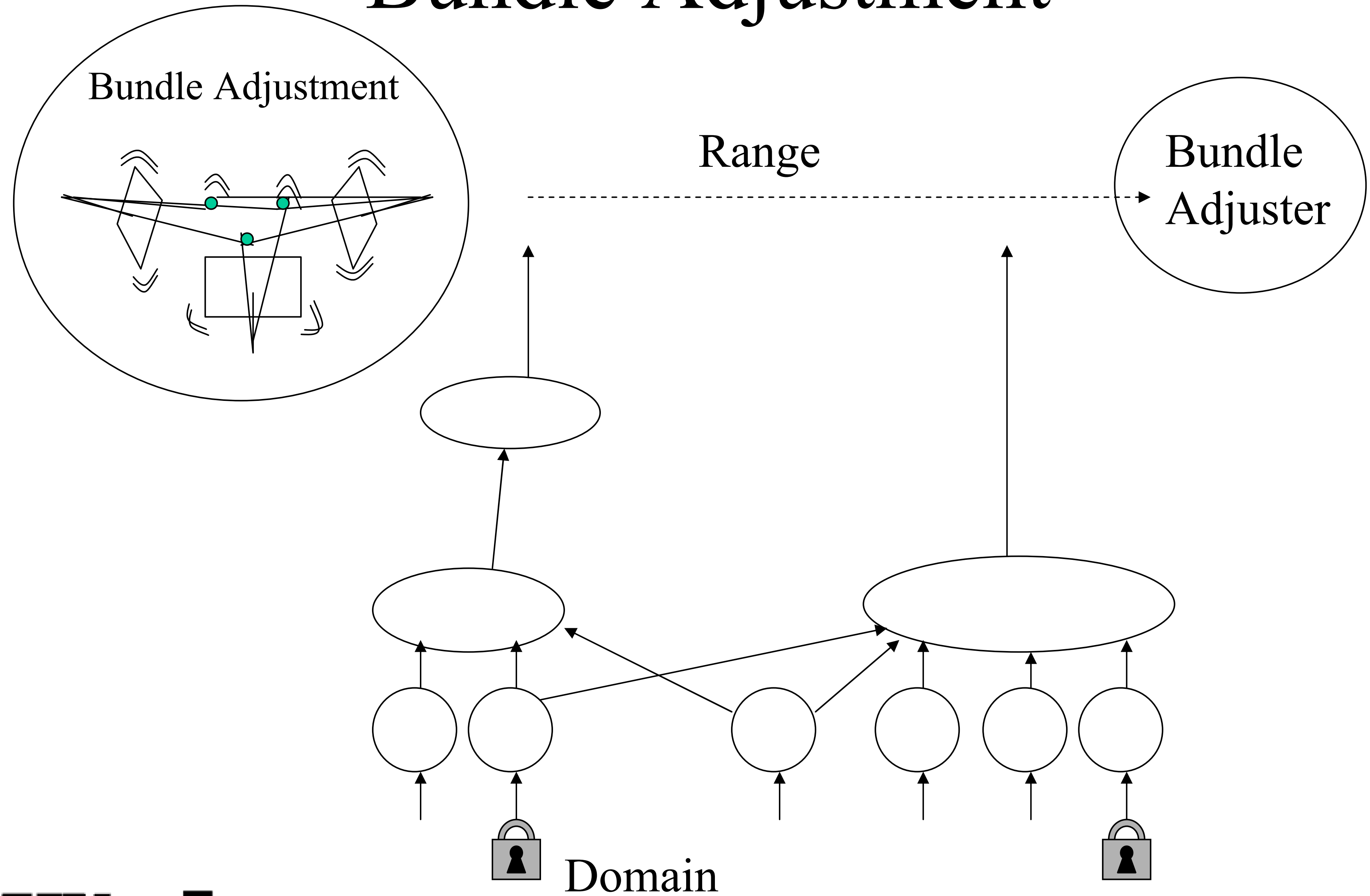
Multiply by

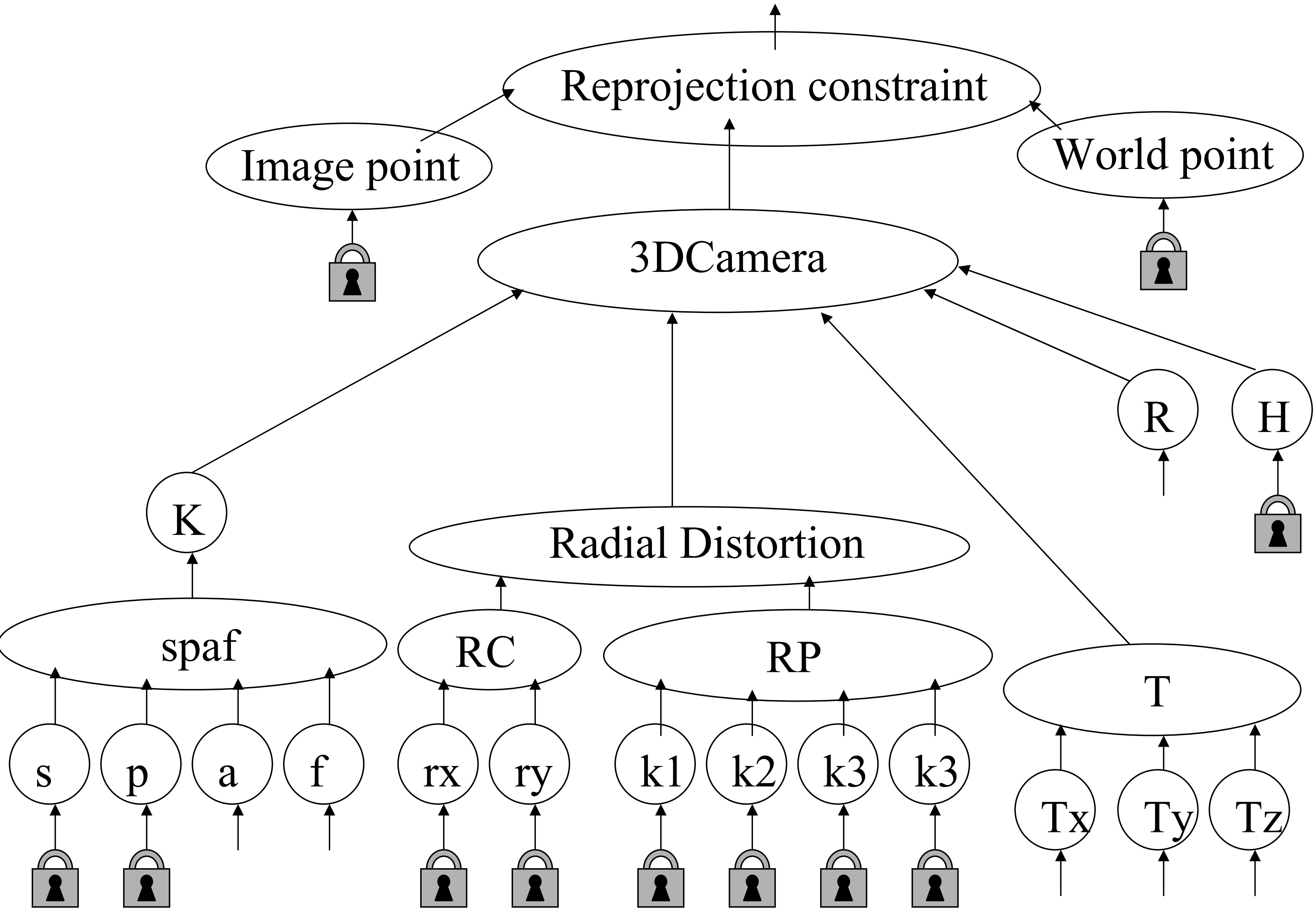
Multiply by

$$\begin{bmatrix} I & H_{SS}^{-1}H_{SC} \\ 0 & H_{CC} - H_{CS}H_{SS}^{-1}H_{SC} \end{bmatrix} \begin{bmatrix} dx_S \\ dx_C \end{bmatrix} = \begin{bmatrix} H_{SS}^{-1}g_S \\ g_C - H_{CS}H_{SS}^{-1}g_S \end{bmatrix}$$



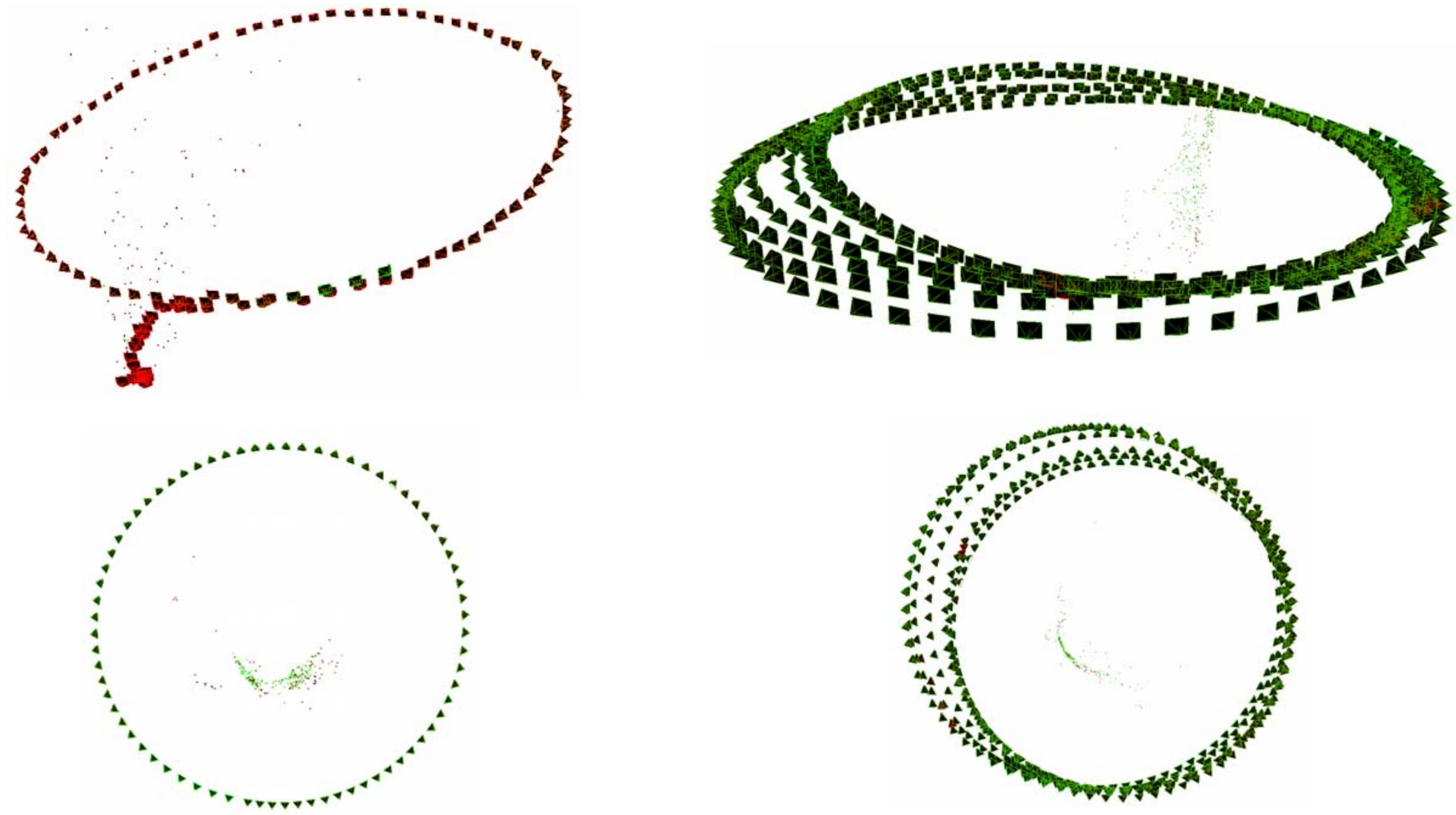
Bundle Adjustment





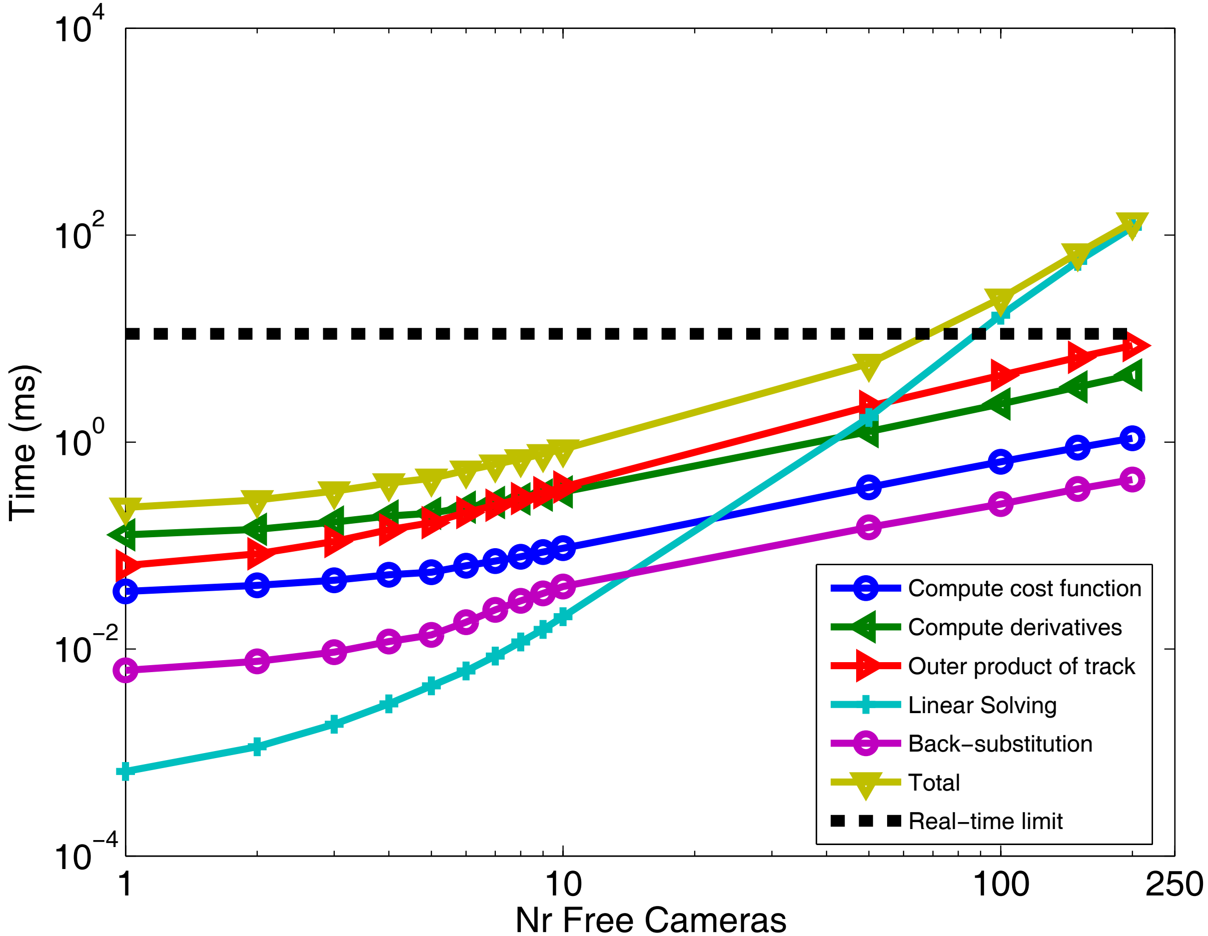
UK Center for Visualization & Virtual Environments

Bundle Adjustment



Center for Visualization & Virtual Environments

Bundle Adjustment



UK

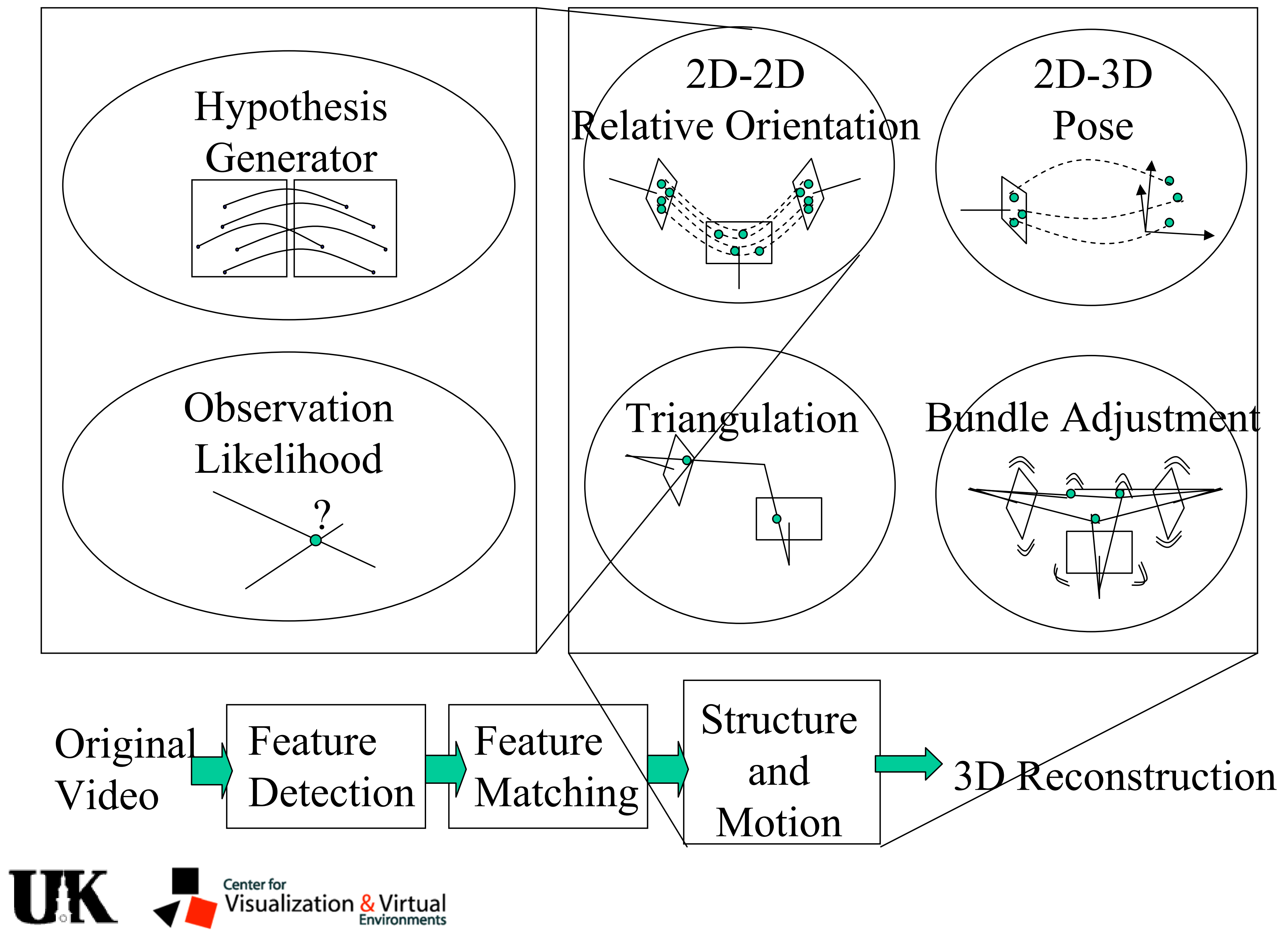
Center for Visualization & Virtual Environments

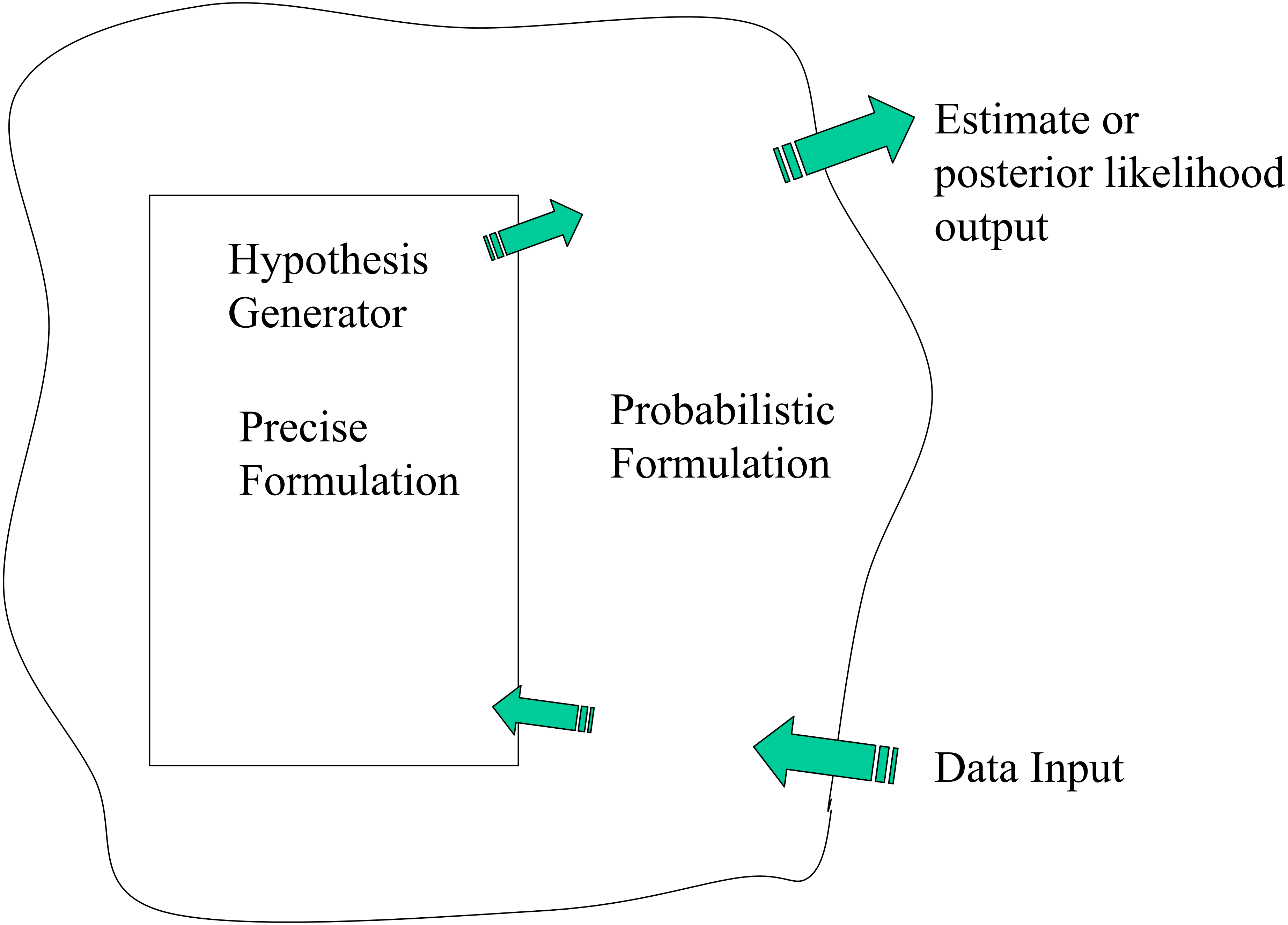
3D Tracking

SBET Only

Bundled

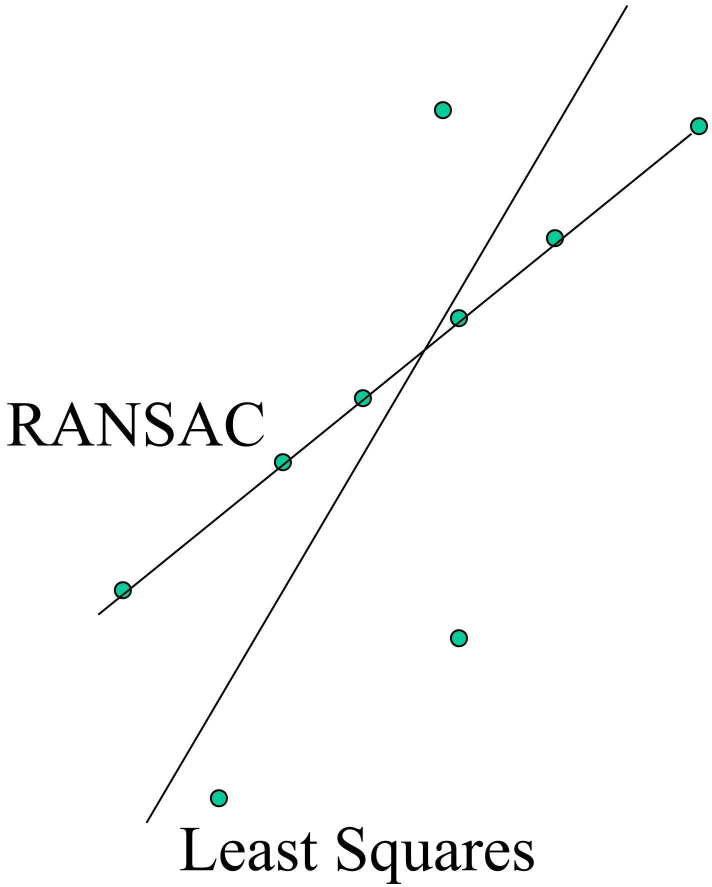
Center for Visualization & Virtual Environments

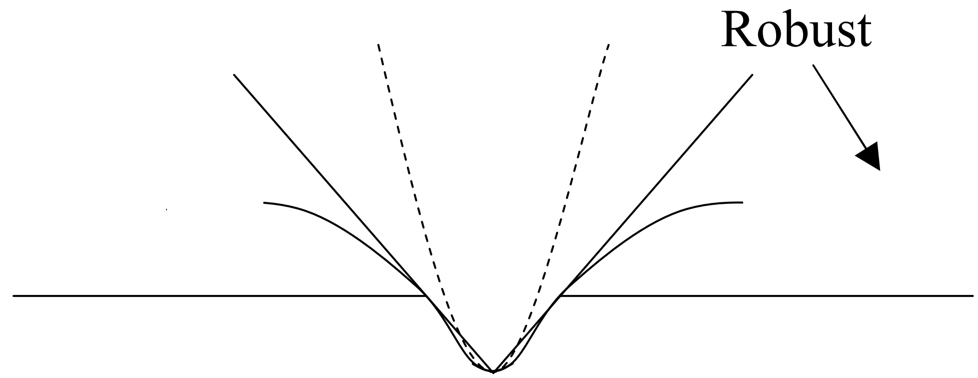

Hypothesis
Generator
Observation
Likelihood
?
2D-2D
Relative Orientation
2D-3D
Pose
Triangulation
Bundle Adjustment
Original
Video
Feature
Detection
Feature
Matching
Structure
and
Motion
3D Reconstruction
Center for
Visualization & Virtual
Environments



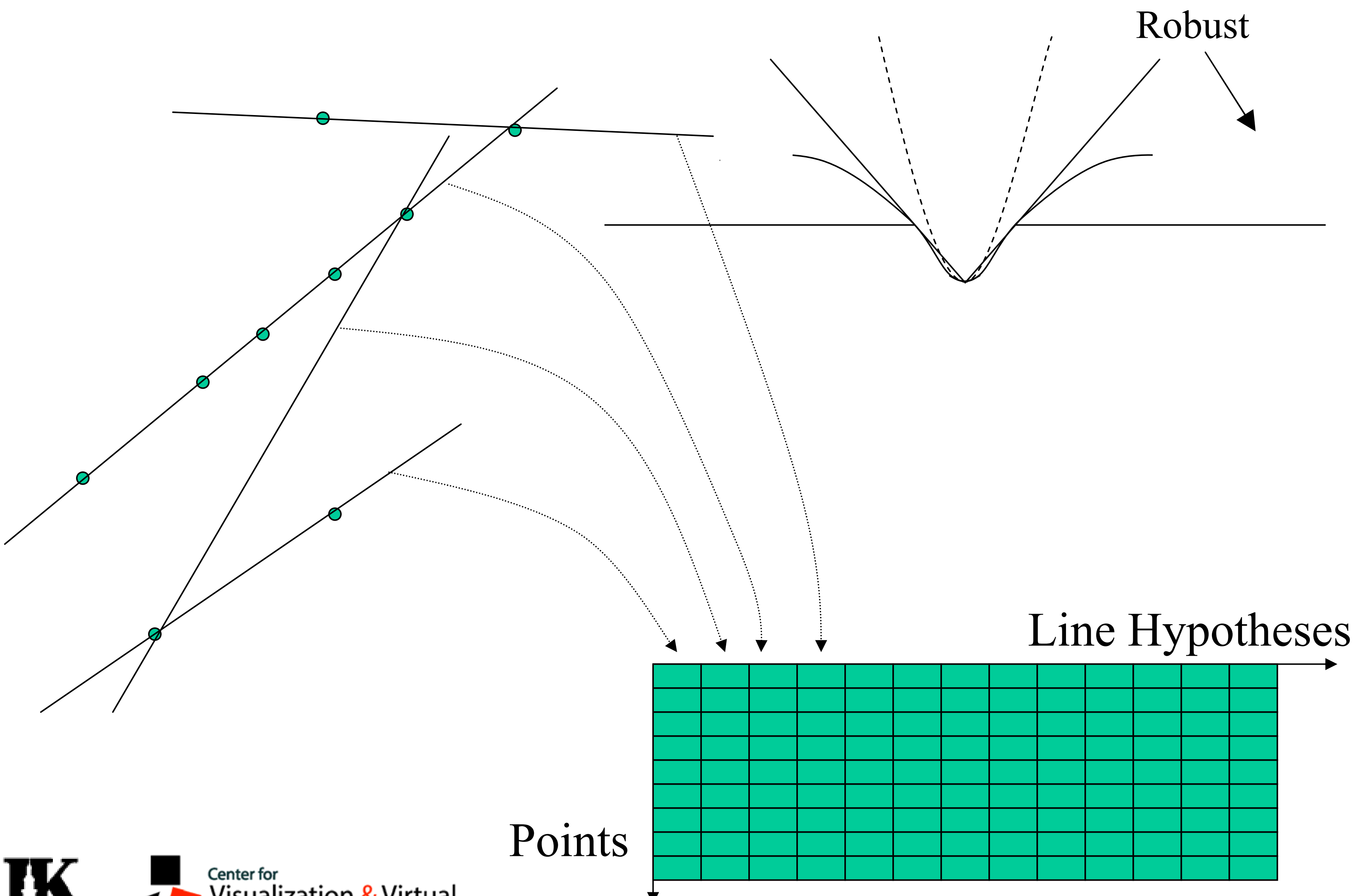
UK Center for Visualization & Virtual Environments

RANSAC- Random Sample Consensus

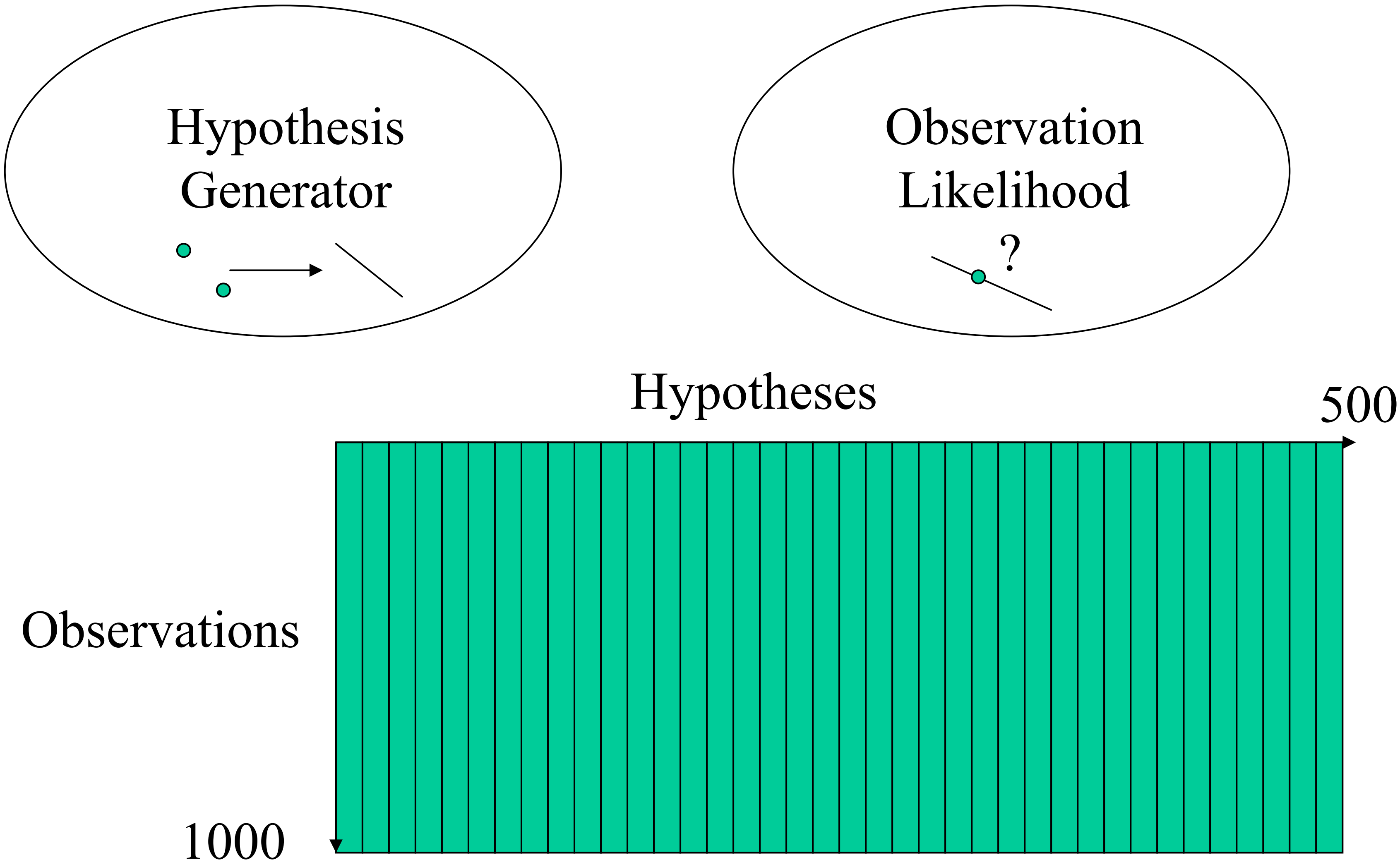




RANSAC- Random Sample Consensus



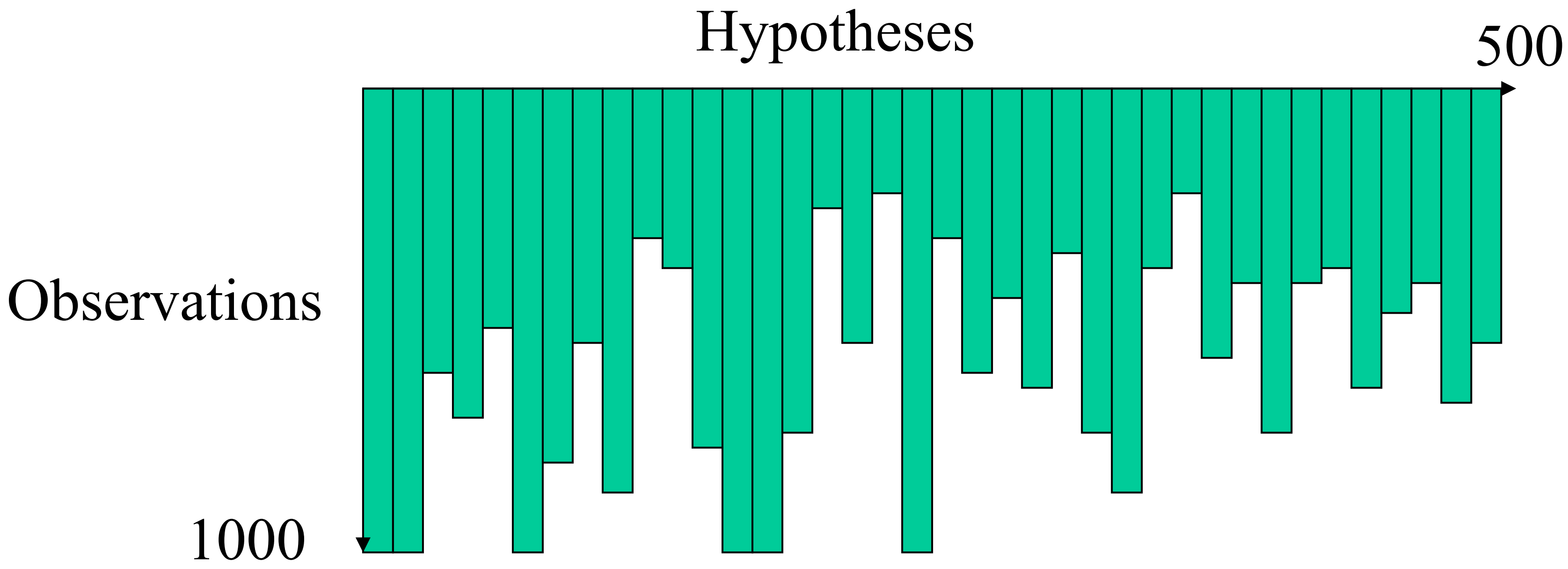
RANSAC



500 x 1000 = 500.000

Preemptive RANSAC

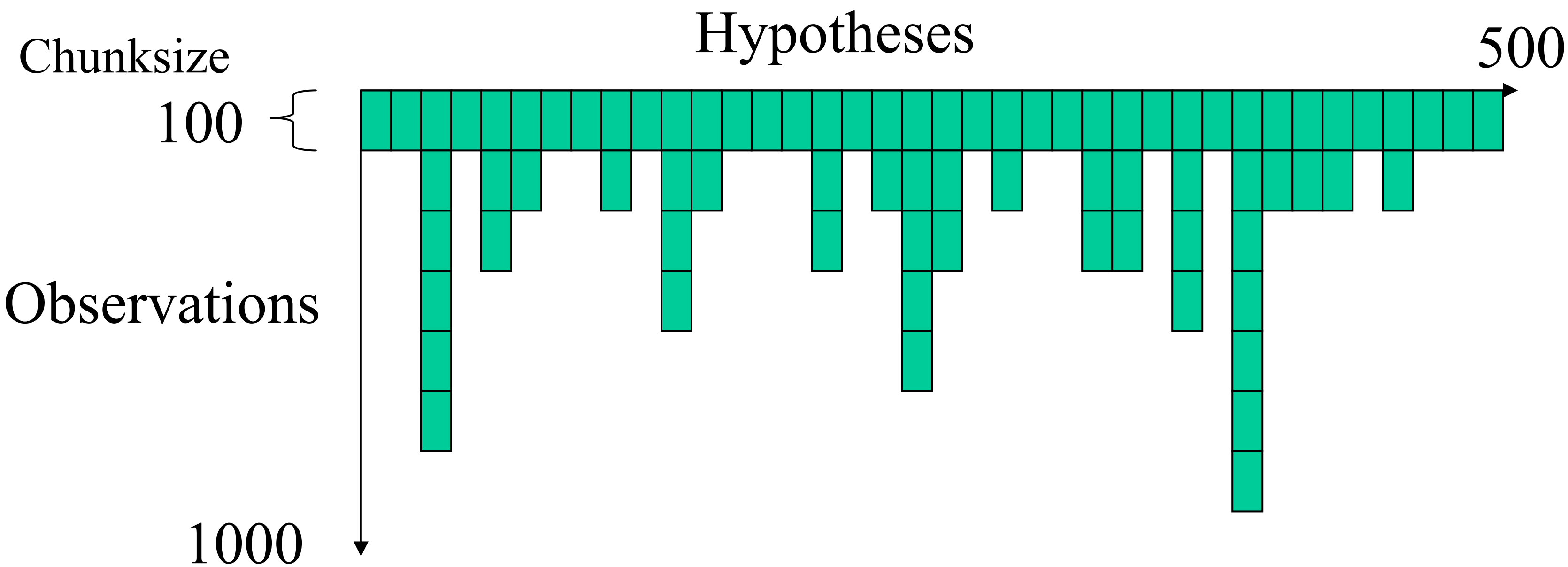
Depth-first Preemption



500 x ???? = ???????

Preemptive RANSAC

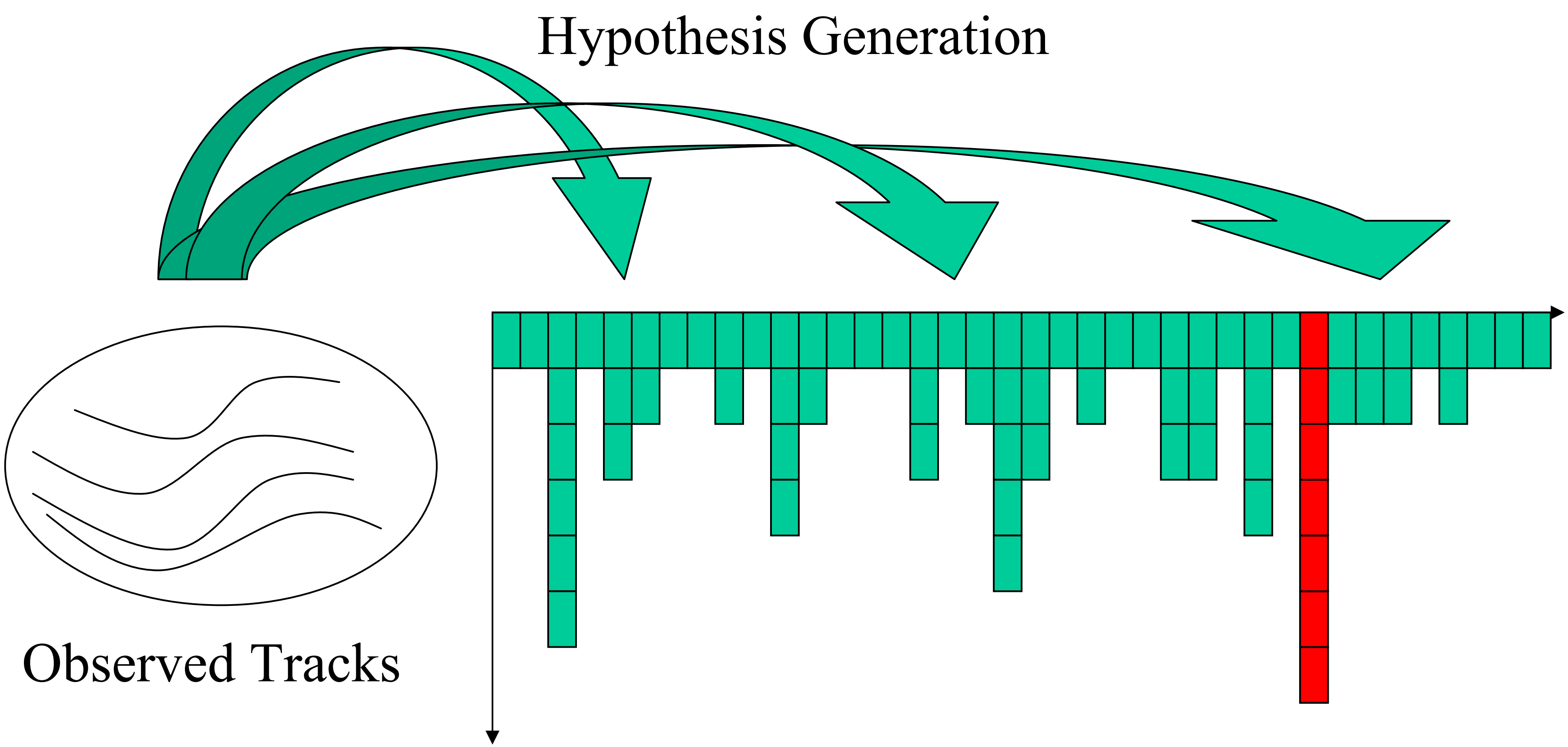
Breadth-first Preemption



500 x 200 = 100.000

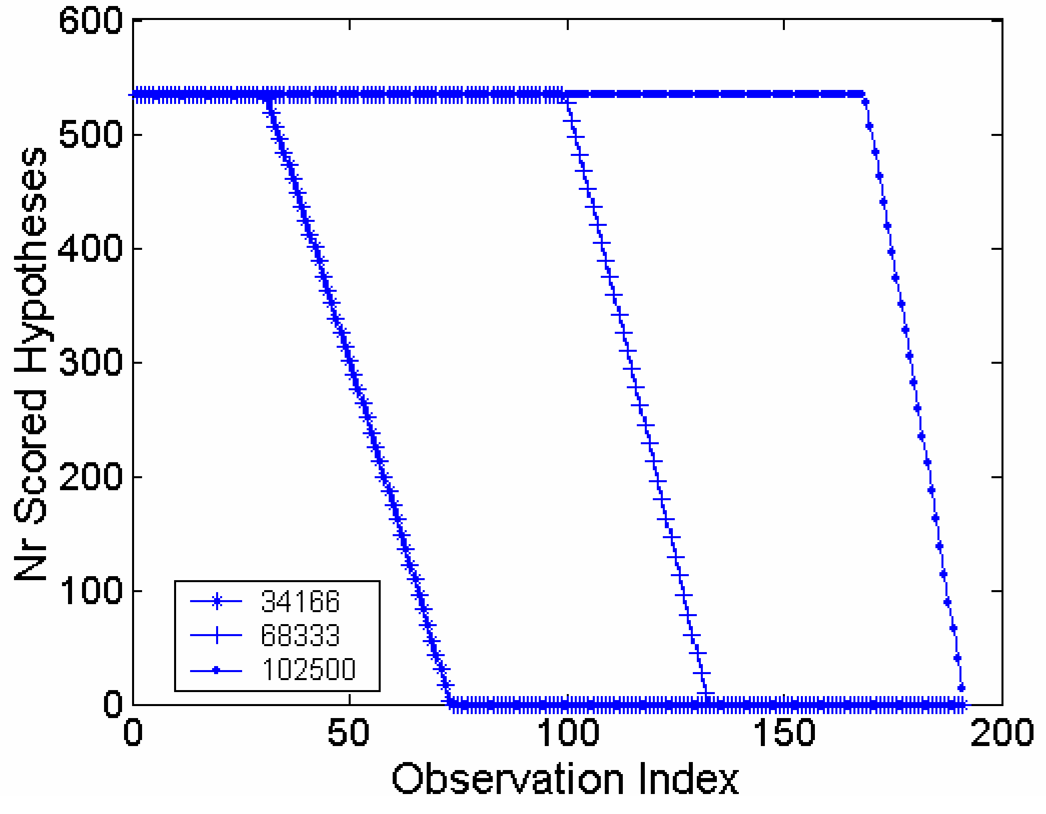
Overhead ~100 microseconds

Preemptive RANSAC



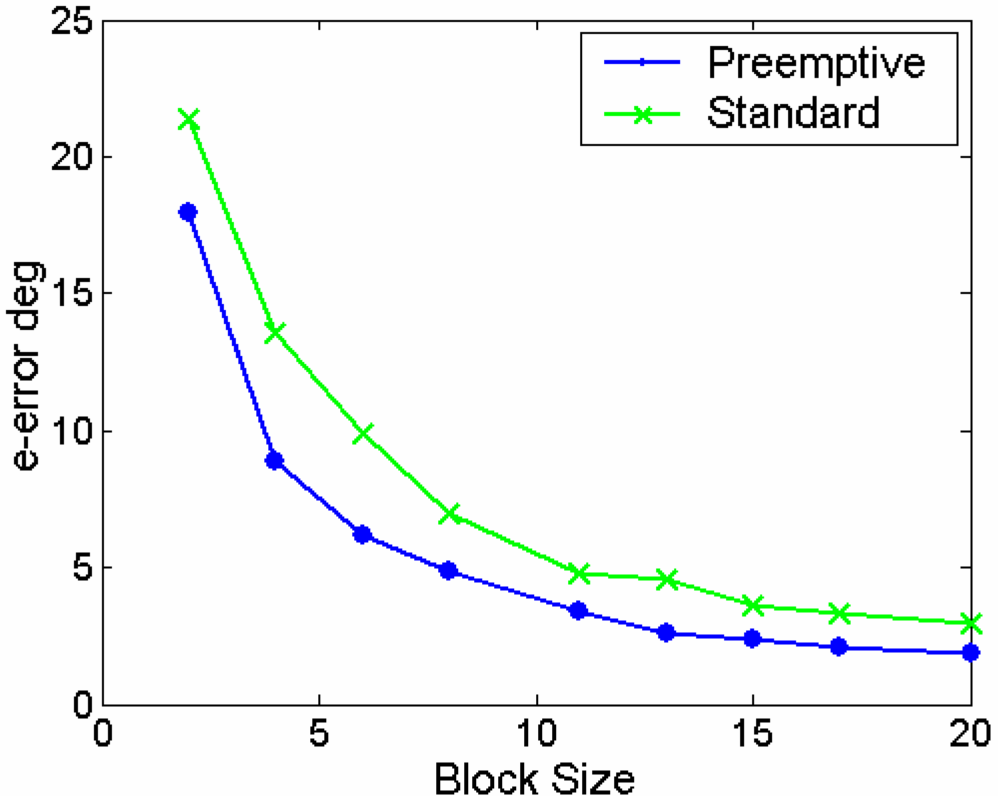
UK Center for Visualization & Virtual Environments

Preemptive RANSAC



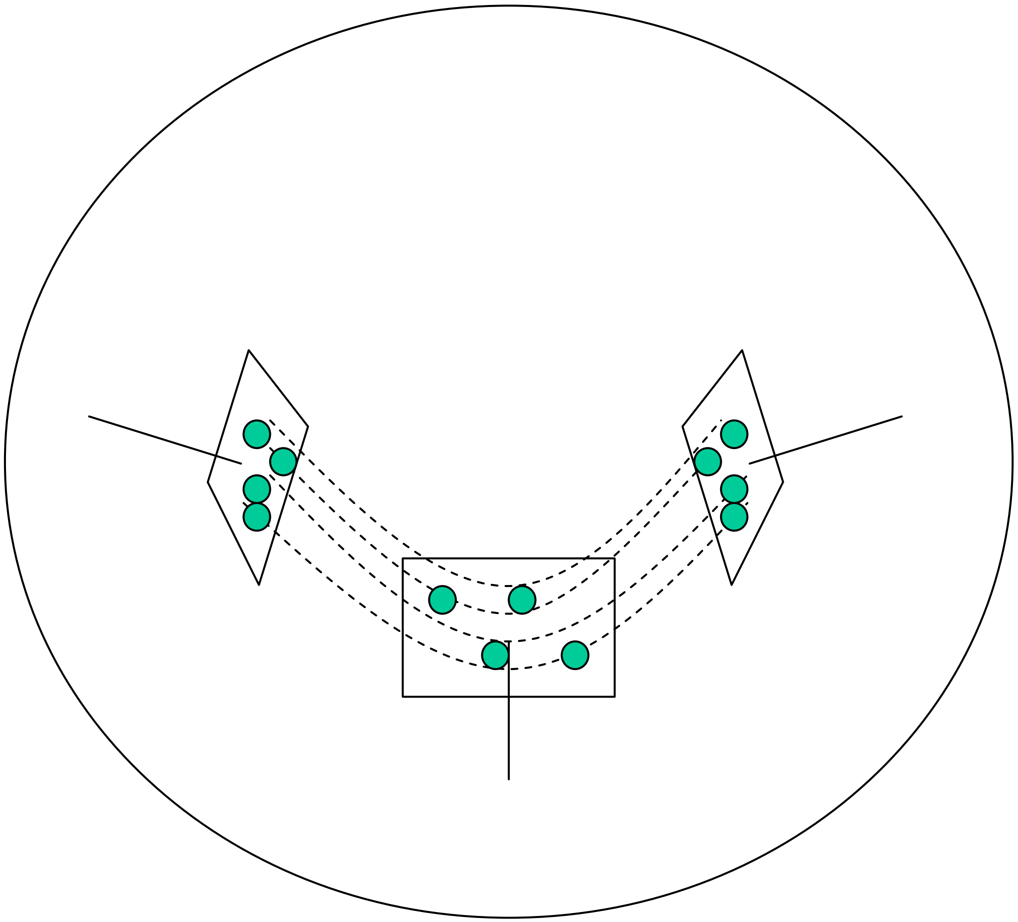
Center for Visualization & Virtual Environments

Preemptive RANSAC

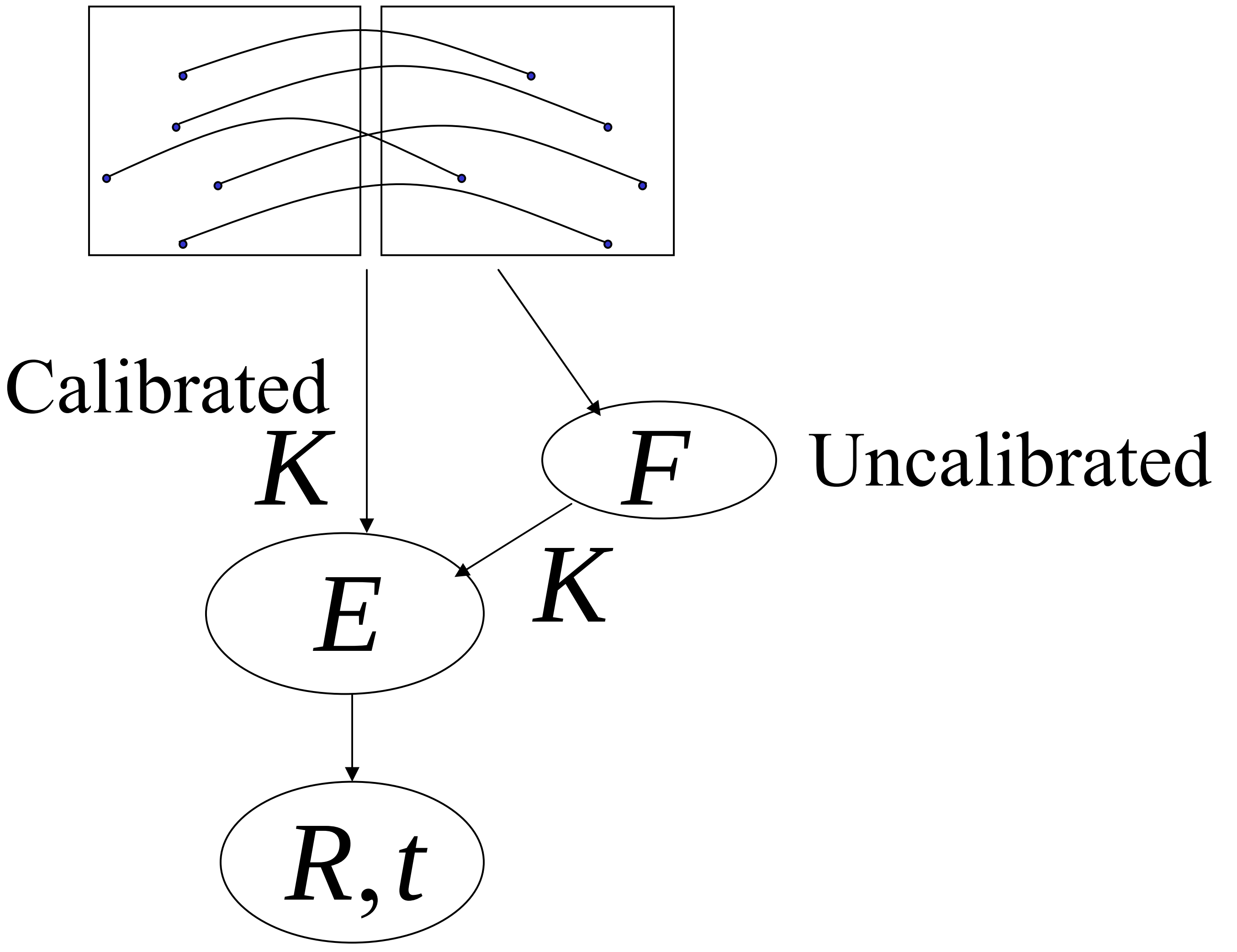


Center for Visualization & Virtual Environments

Relative Orientation



Calibrated vs Uncalibrated



Constraints

$$x'^{\mathrm{T}} F x = 0$$

$$\begin{bmatrix} x_1x_1' & x_2x_1' & x_3x_1' & x_1x_2' & x_2x_2' & x_3x_2' & x_1x_3' & x_2x_3' & x_3x_3' \\ x_1x_1' & x_2x_1' & x_3x_1' & x_1x_2' & x_2x_2' & x_3x_2' & x_1x_3' & x_2x_3' & x_3x_3' \\ x_1x_1' & x_2x_1' & x_3x_1' & x_1x_2' & x_2x_2' & x_3x_2' & x_1x_3' & x_2x_3' & x_3x_3' \\ x_1x_1' & x_2x_1' & x_3x_1' & x_1x_2' & x_2x_2' & x_3x_2' & x_1x_3' & x_2x_3' & x_3x_3' \\ x_1x_1' & x_2x_1' & x_3x_1' & x_1x_2' & x_2x_2' & x_3x_2' & x_1x_3' & x_2x_3' & x_3x_3' \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = 0$$

Constraints

$$\text{SingularValues}(F) = [\sigma_1 \quad \sigma_2 \quad \sigma_3]$$

Uncalibrated: $\sigma_3 = 0 \Rightarrow \det F = 0$

Calibrated: $\sigma_3 = 0 \qquad \sigma_1 = \sigma_2$

$$\Downarrow$$

$$2EE^{\mathrm{T}}E - trace(EE^{\mathrm{T}})E = 0$$

2 Views

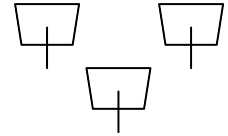
8p
von Sanden, 1908
Longuet-Higgins, 1981

7p
R. Sturm, 1869

6p
Philip, 1996

5p
Kruppa 1913
Nister 2003

3 Views

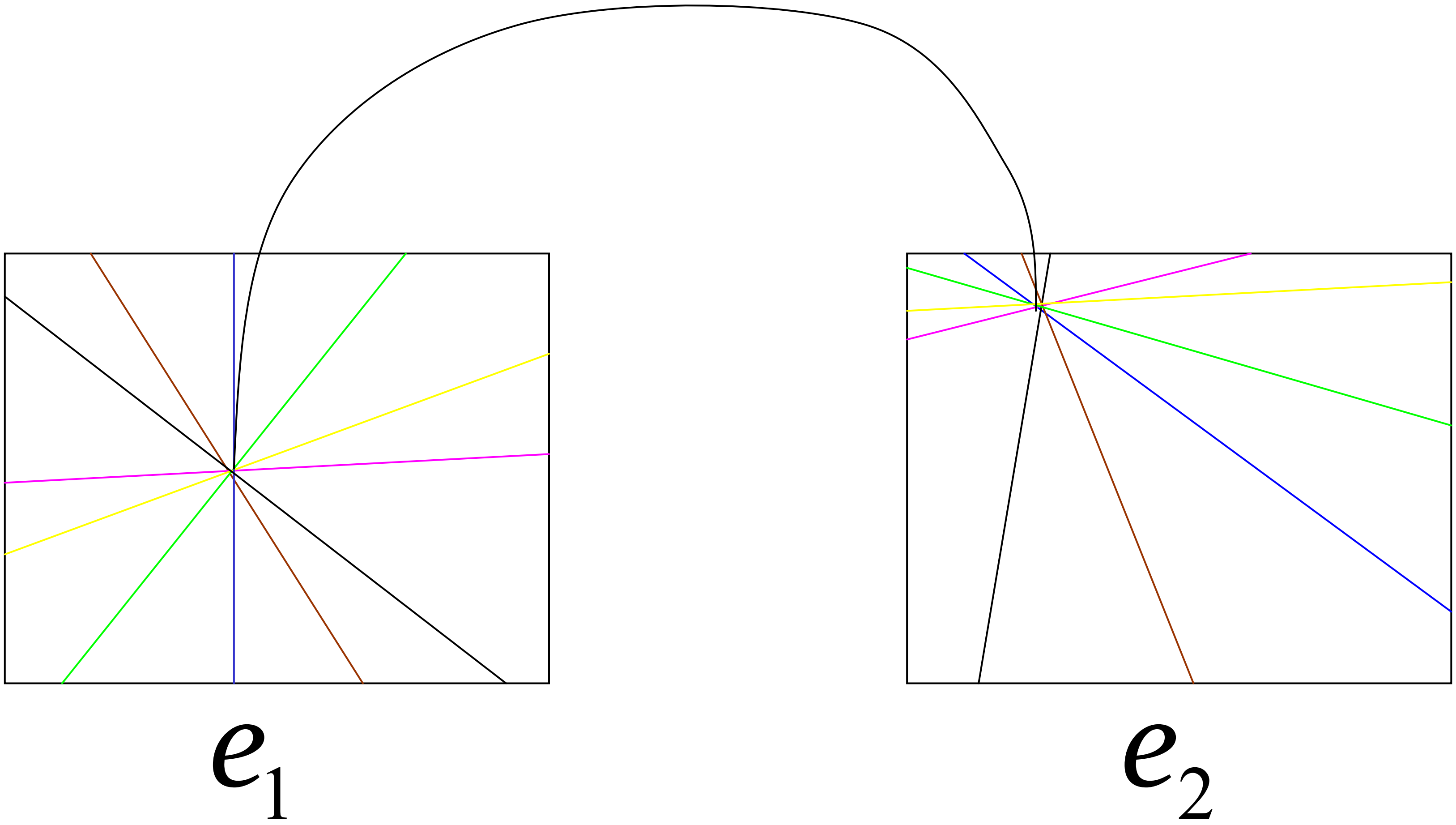


6p
Quan, 1994

4p
Nister, Schaffalitzky, 2004

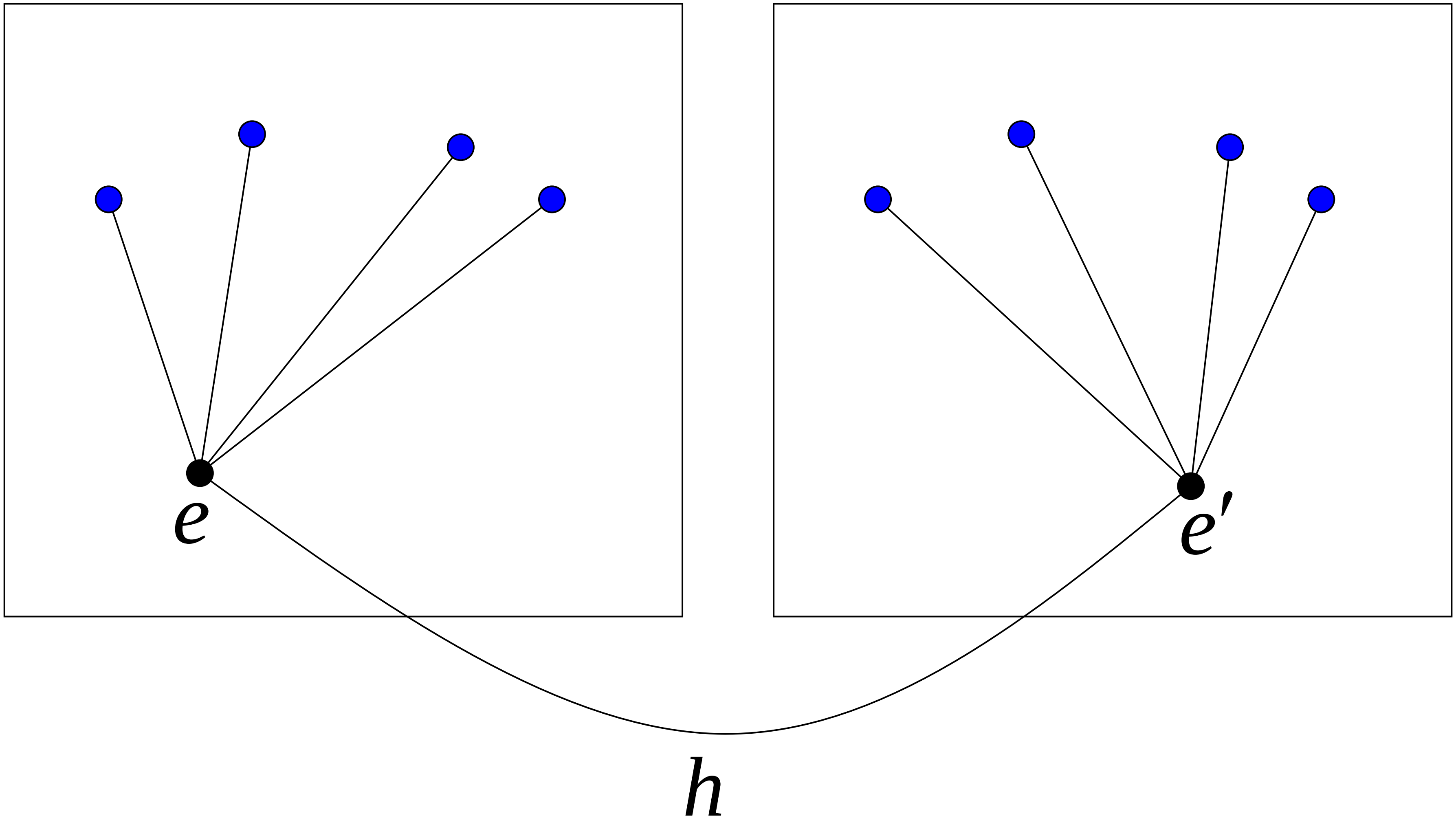
5p
Nister, 2003

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \leftarrow \text{Pure Rotation}$$



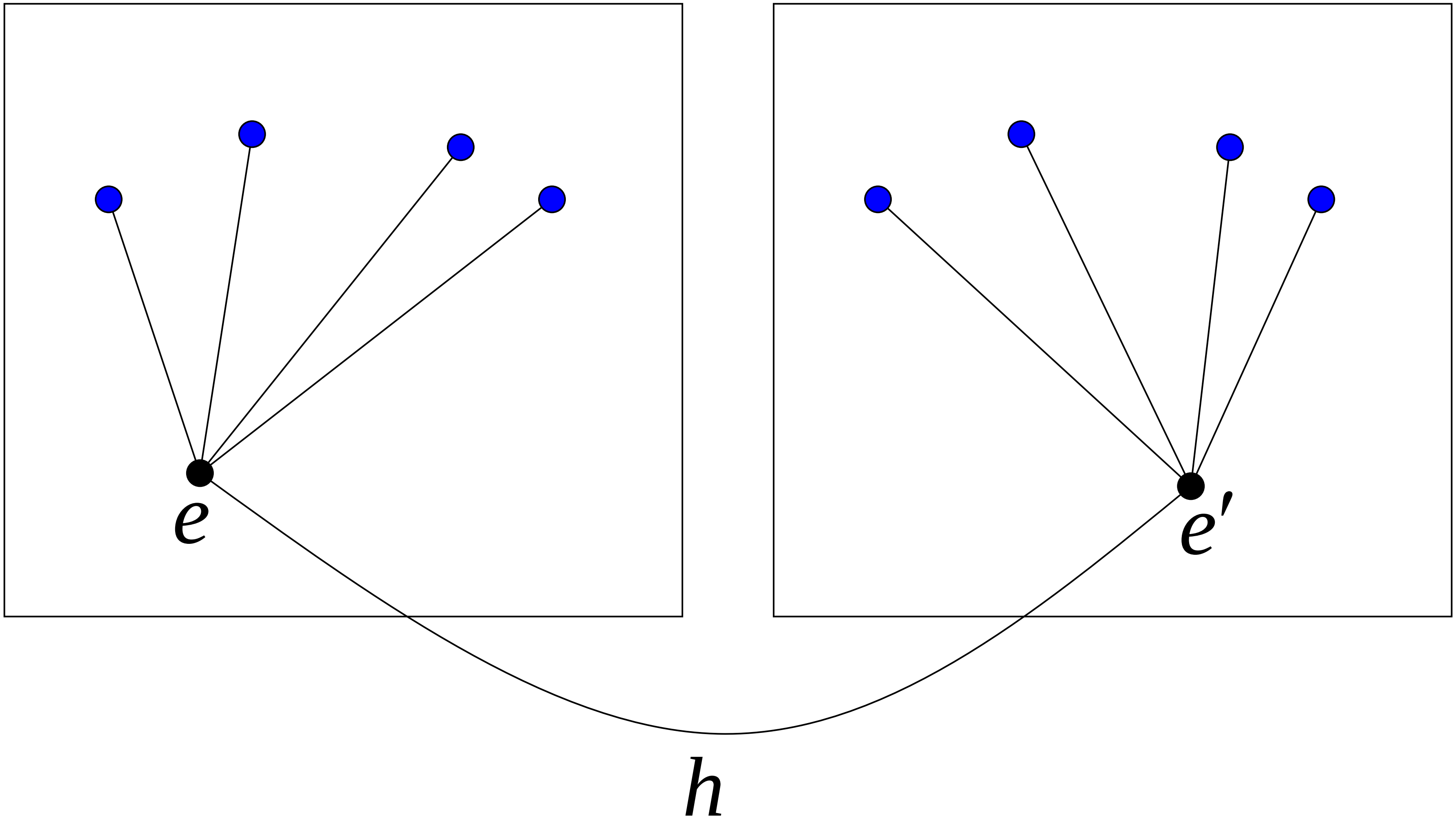
UK Center for Visualization & Virtual Environments

The Epipoles and the Epipolar Line Homography

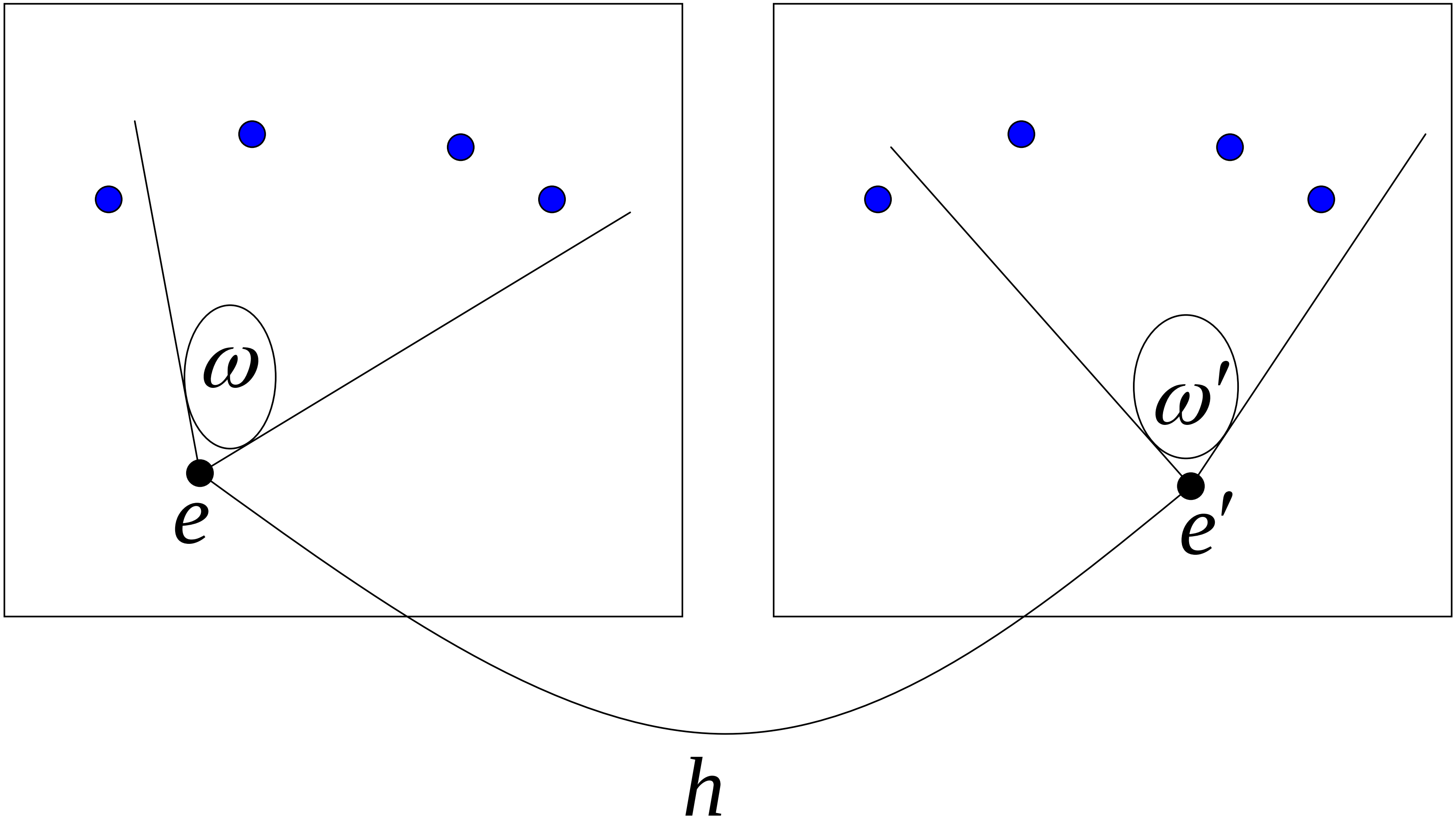


Center for Visualization & Virtual Environments

The Epipolar Constraint



The Kruppa Constraints



The Five Point Problem

Given five point correspondences,

What is R,t ?

E. Kruppa,
Zur Ermittlung eines Objektes aus zwei Perspektiven
mit Innerer Orientierung,
1913.

O. Faugeras and S. Maybank,
Motion from Point Matches: Multiplicity of Solutions,
1990.

J. Philip,
A Non-Iterative Algorithm for Determining all
Essential Matrices Corresponding to Five Point Pairs,
1996.

B. Triggs,
Routines for Relative Pose of Two Calibrated Cameras from 5 Points,
2000.

D. Nister,
An Efficient Solution to the Five-Point Relative Pose Problem,
2002.

The solution is minimal in two respects:

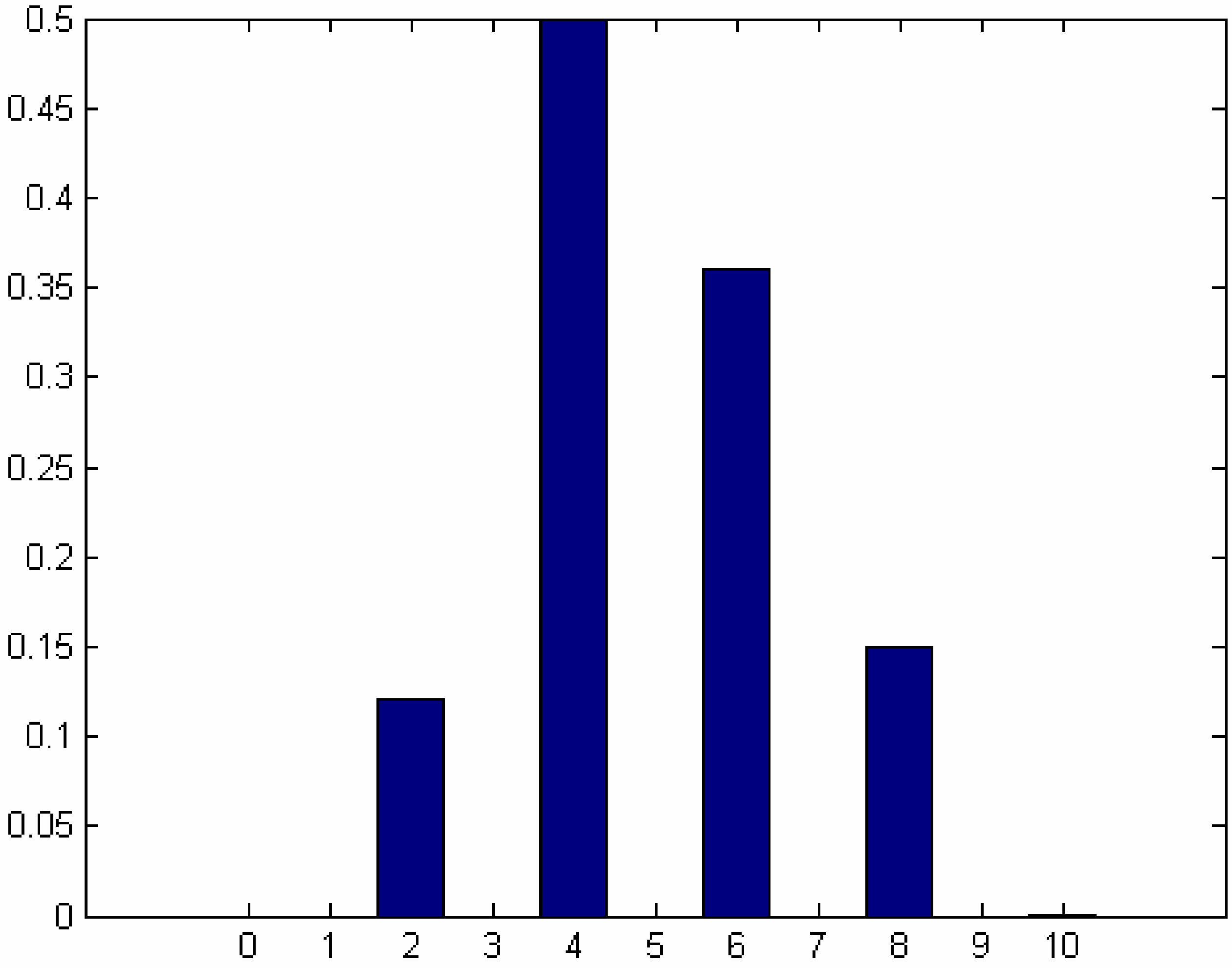
It can operate on the smallest number of points required to get a finite number of solutions.

Closed form derivation of 10th degree polynomial.

↓

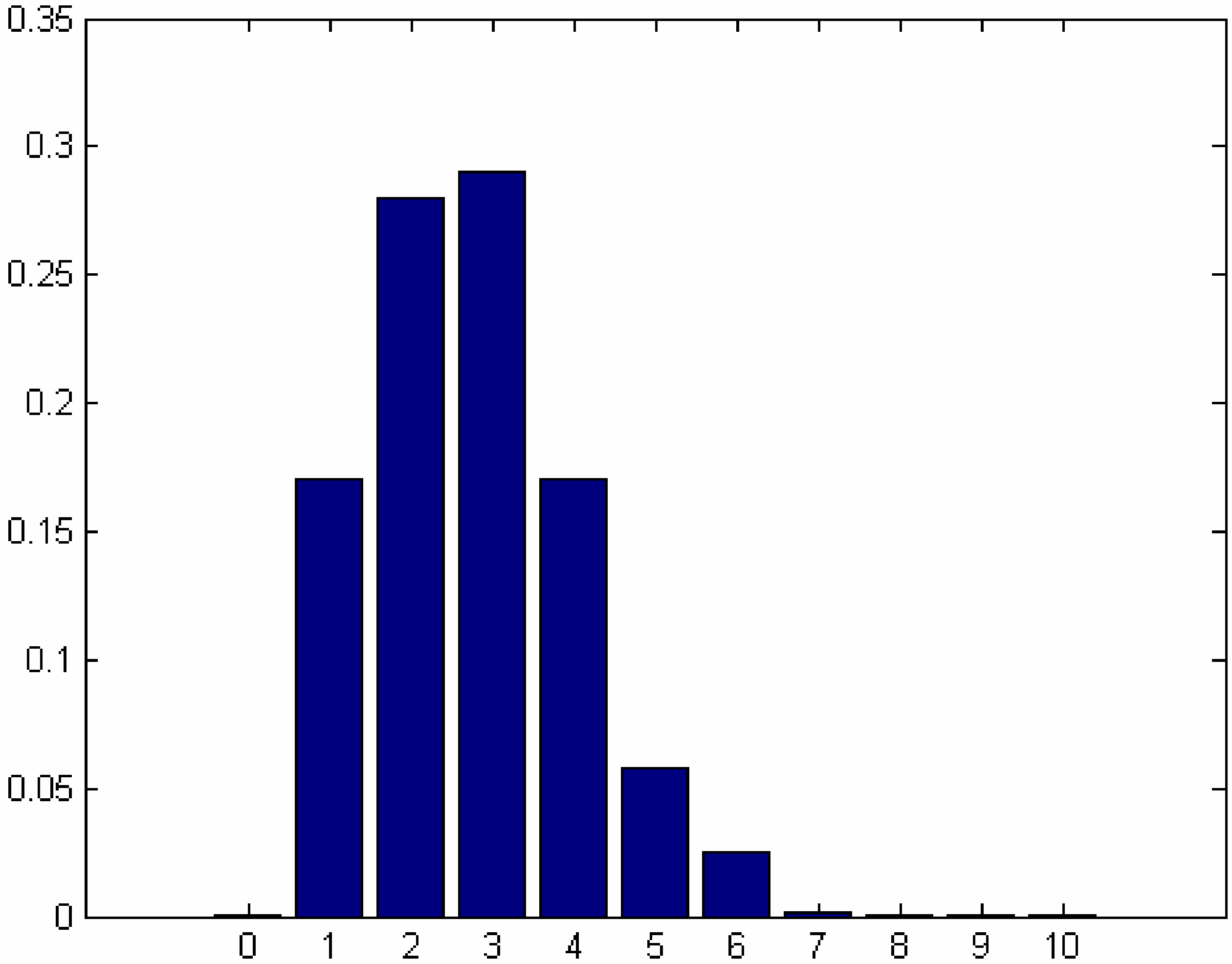
First solution suited for numerical implementation that corresponds directly to the intrinsic degree of difficulty of the problem.

Nr of Roots



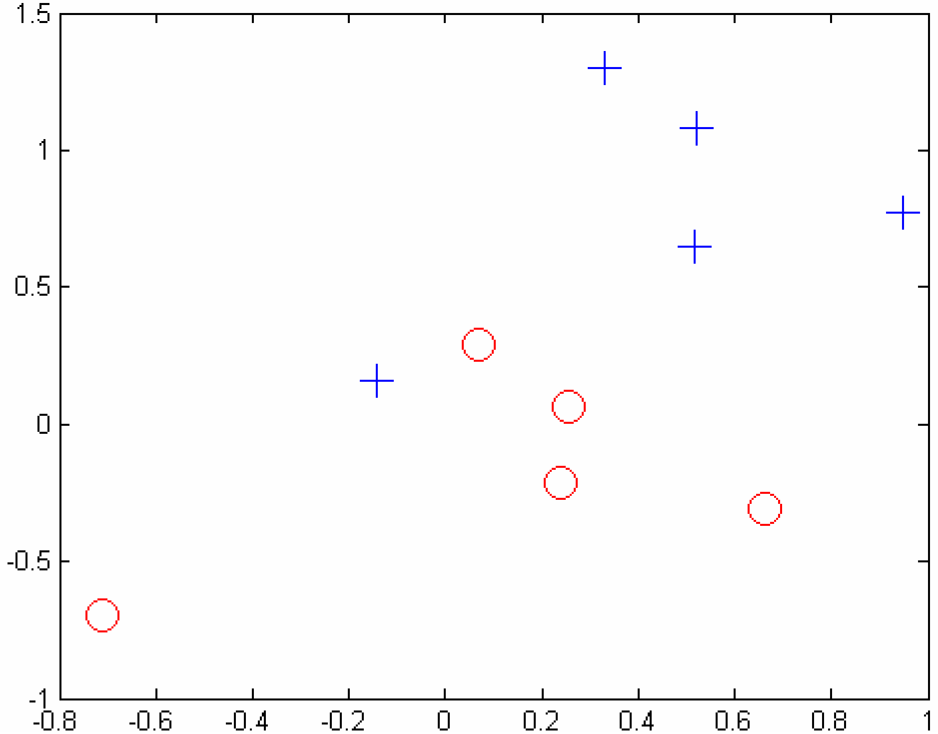
Average 4.55

Nr of Solutions



Average 2.74

10 Solutions



[0.067, 0.287] < > [0.329,1.297]

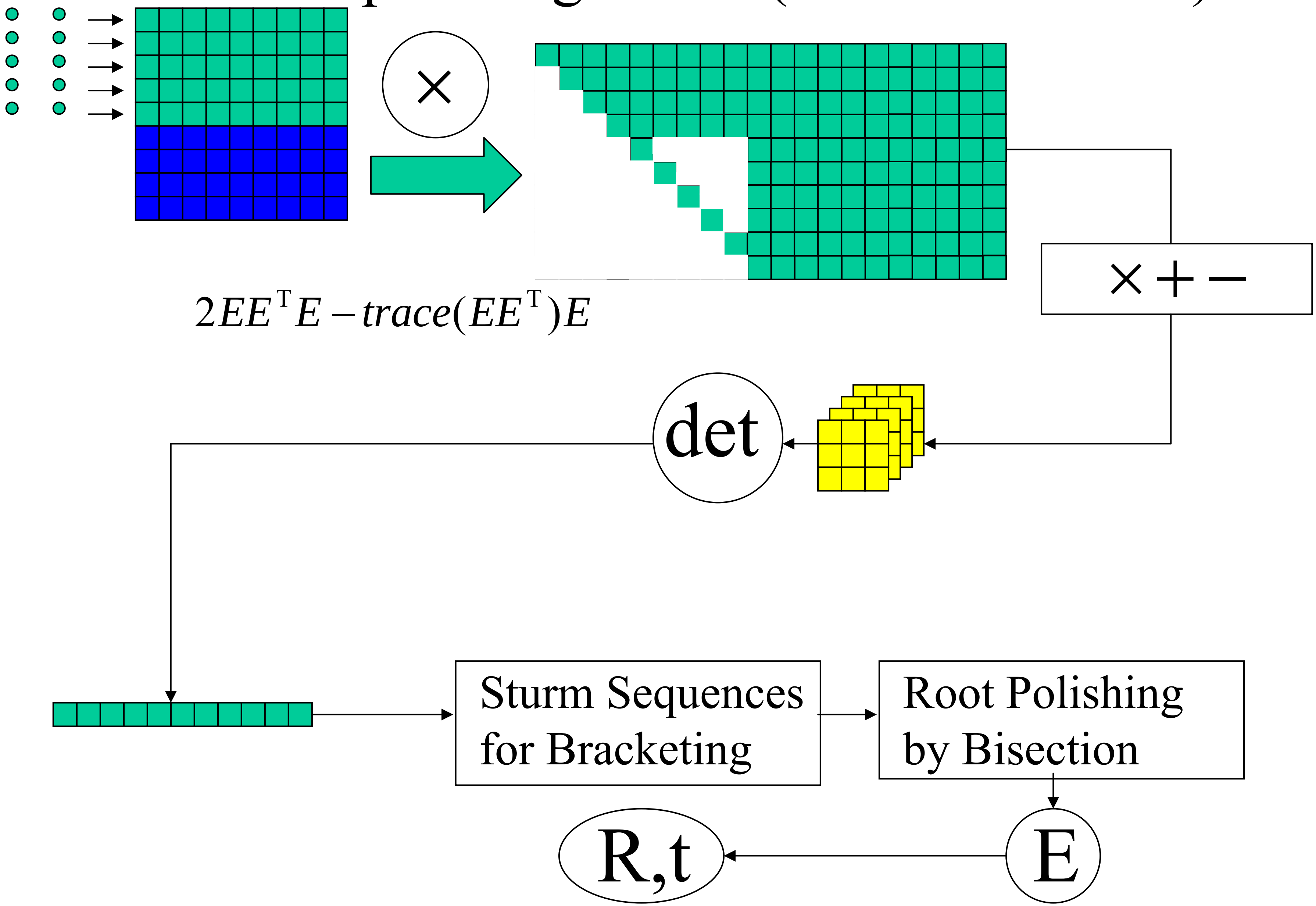
[0.254, 0.0646] < > [0.523,1.0807]

[0.239, -0.213] < > [0.517,0.645]

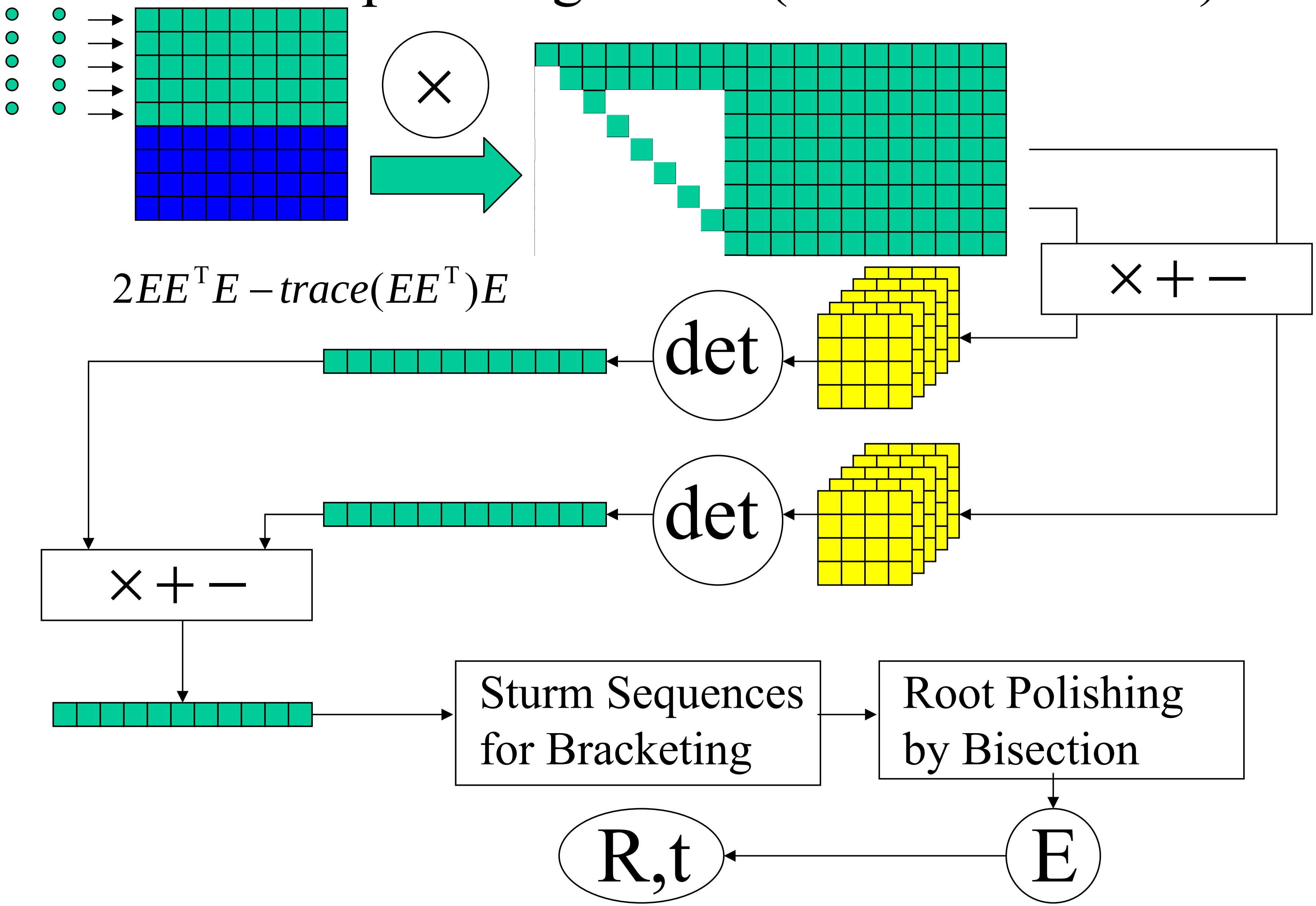
[-0.710, -0.693] < > [-0.141,0.157]

[0.661, -0.307] < > [0.950, 0.773]

The 5-point algorithm (Nistér PAMI 04)



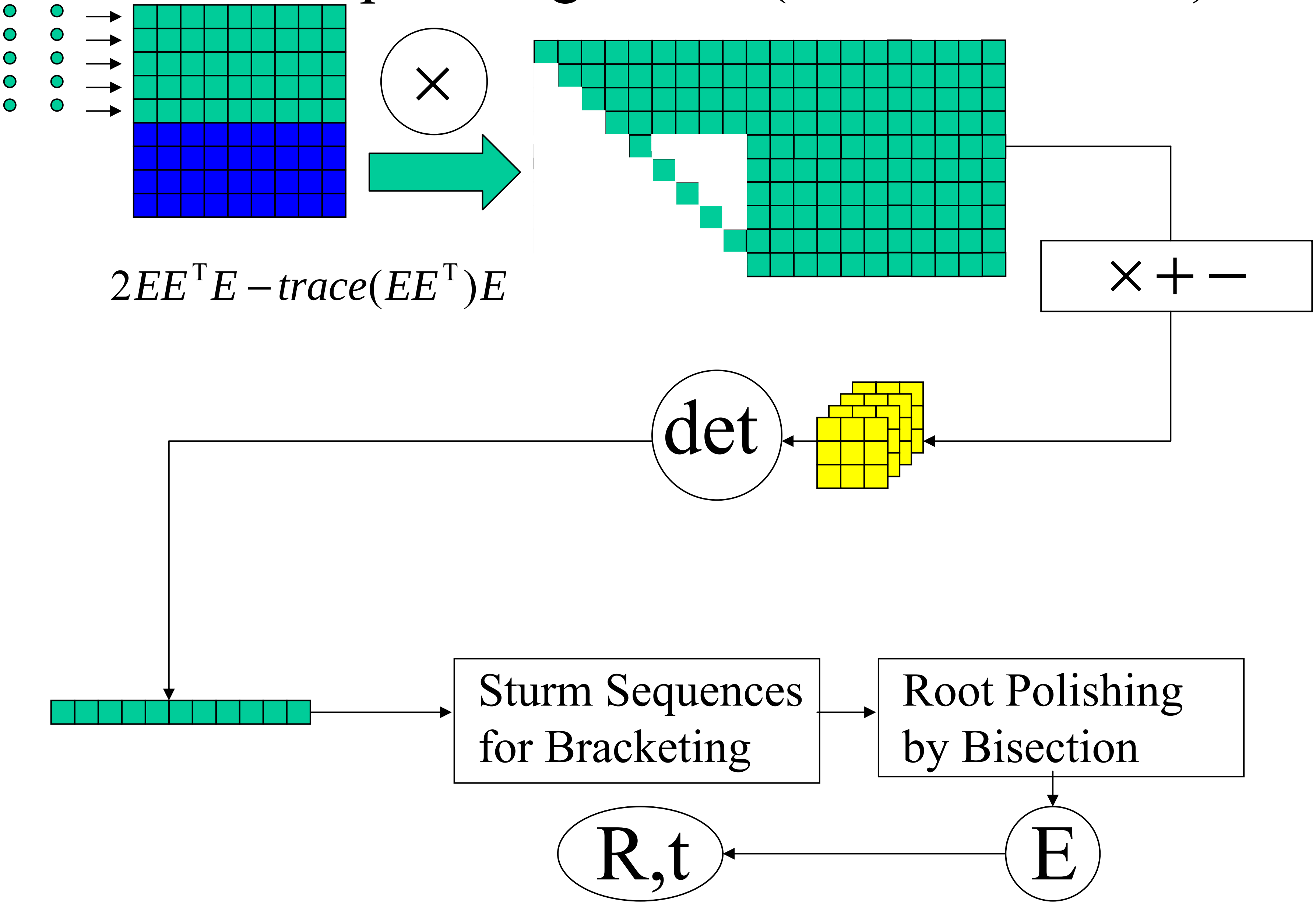
The 5-point algorithm (Nistér CVPR 03)



UK

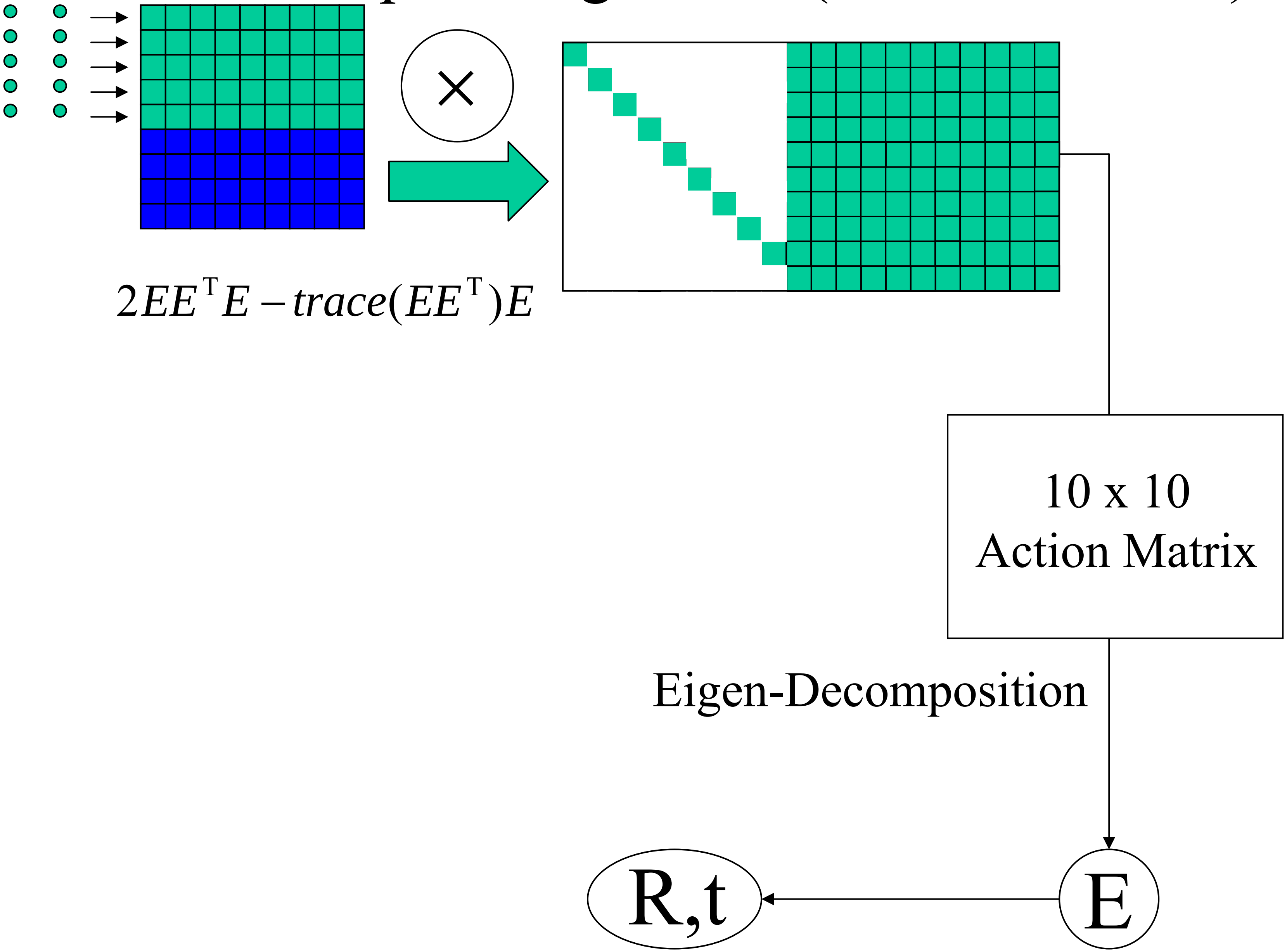
Center for Visualization & Virtual Environments

The 5-point algorithm (Nistér PAMI 04)



Center for Visualization & Virtual Environments

The 5-point algorithm (Stewénius et al)

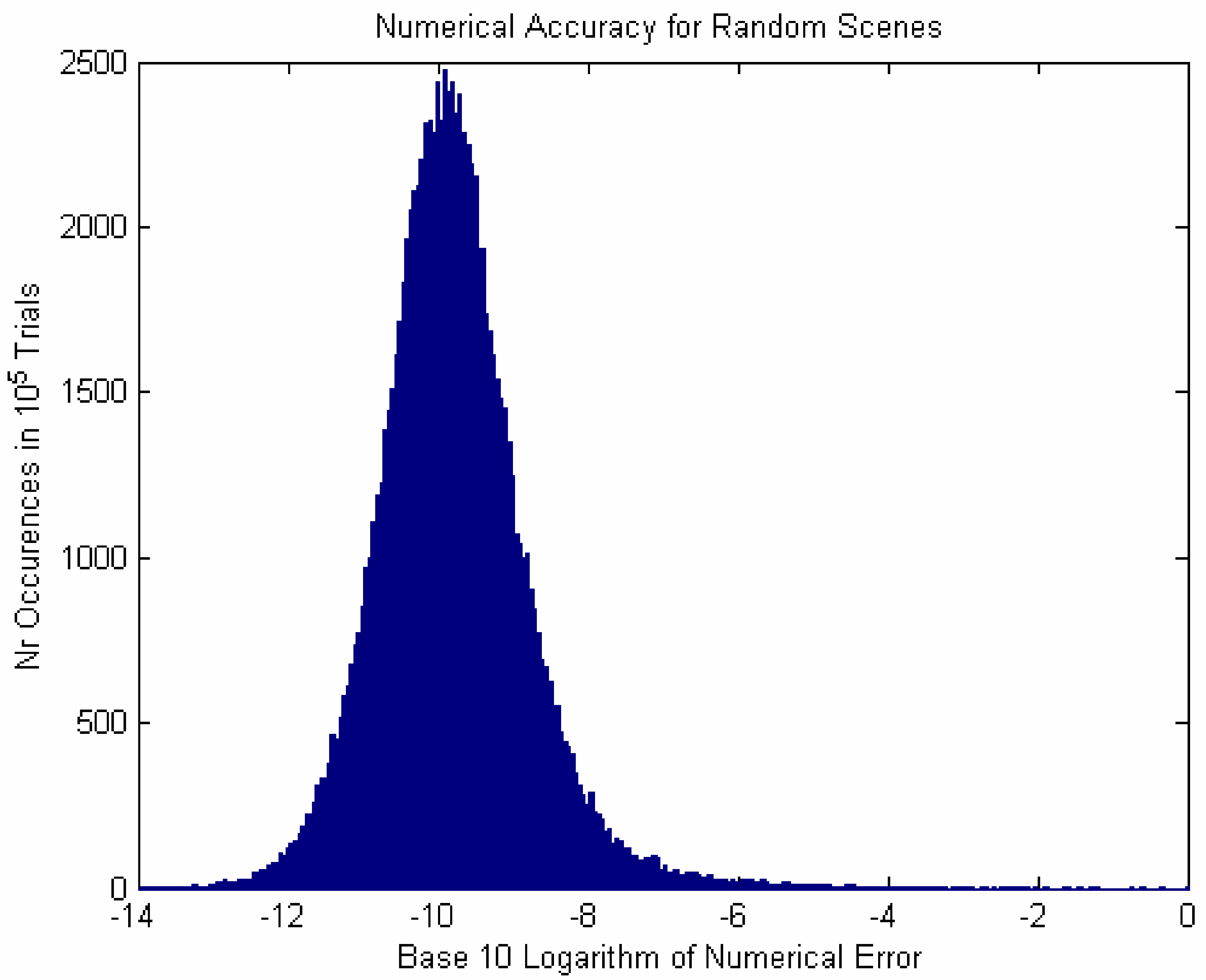


5-Point Matlab Executable

Recent Developments on Direct Relative Orientation,
Henrik Stewenius, Christopher Engels, David Nister,
ISPRS Journal of Photogrammetry and Remote Sensing

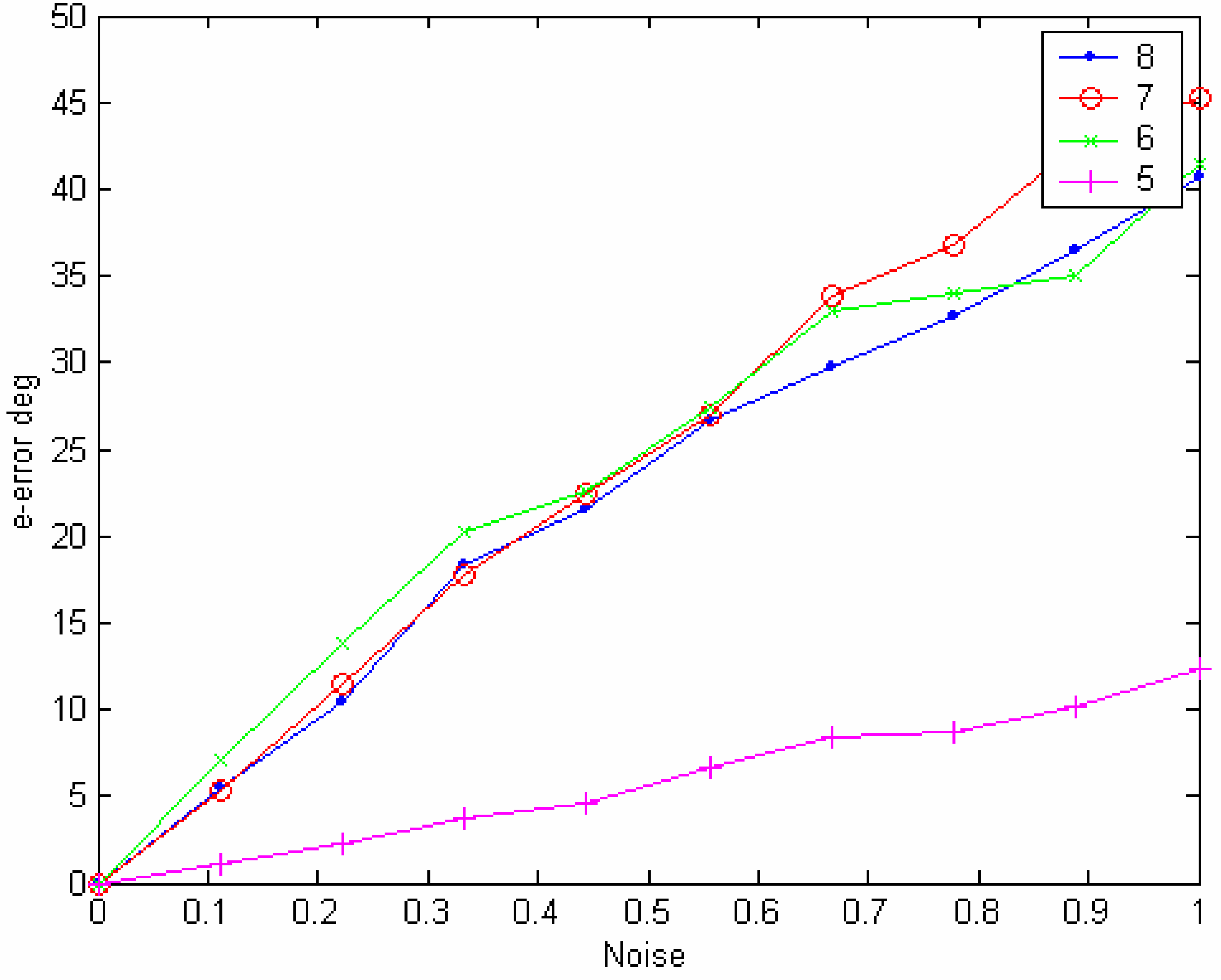
www.vis.uky.edu/~dnister

UK Center for Visualization & Virtual Environments



Center for Visualization & Virtual Environments

Noise

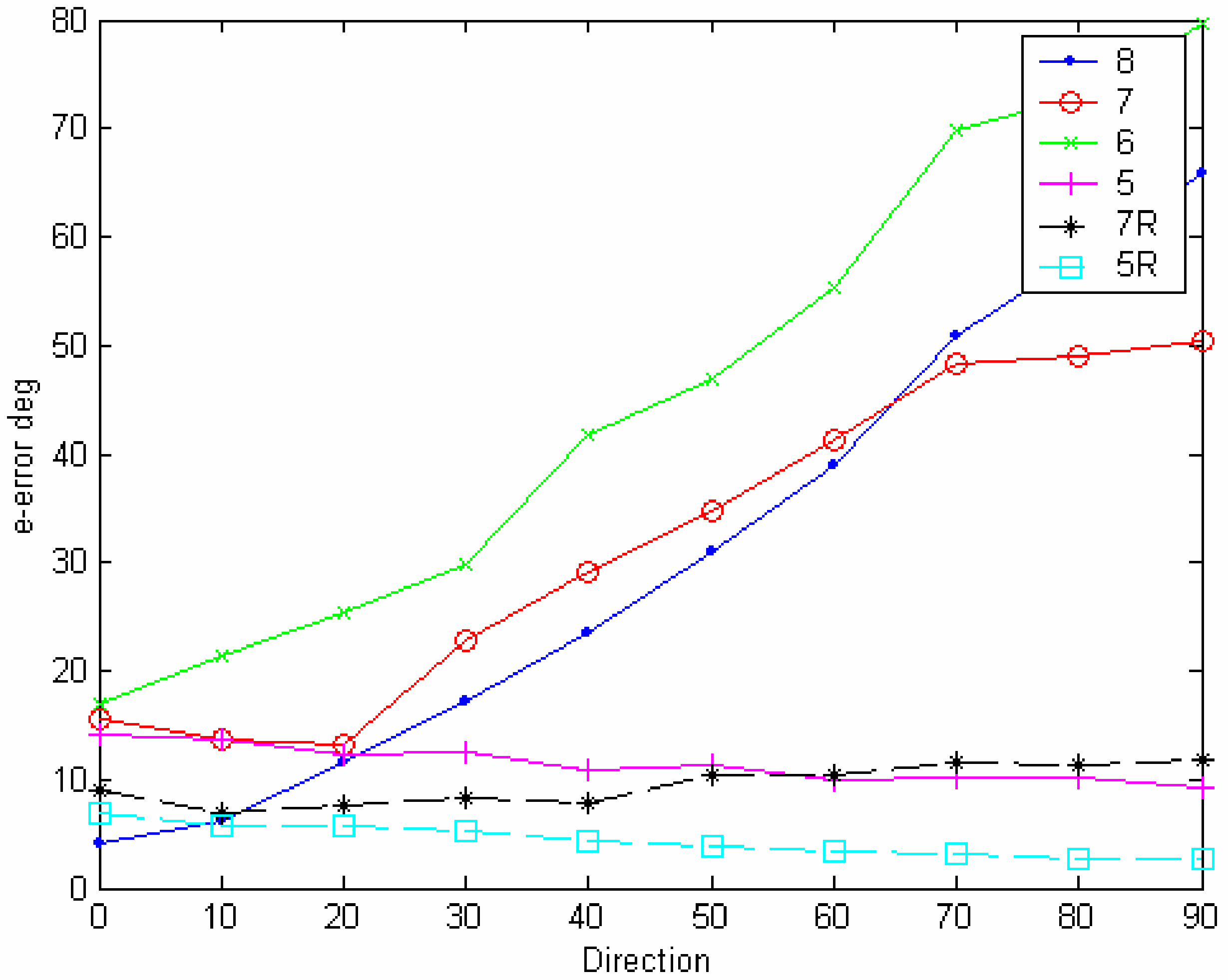


Minimal Cases, Sideways Motion

Depth 0.5
Baseline 0.1
Field of View 45 degrees

UK Center for Visualization & Virtual Environments

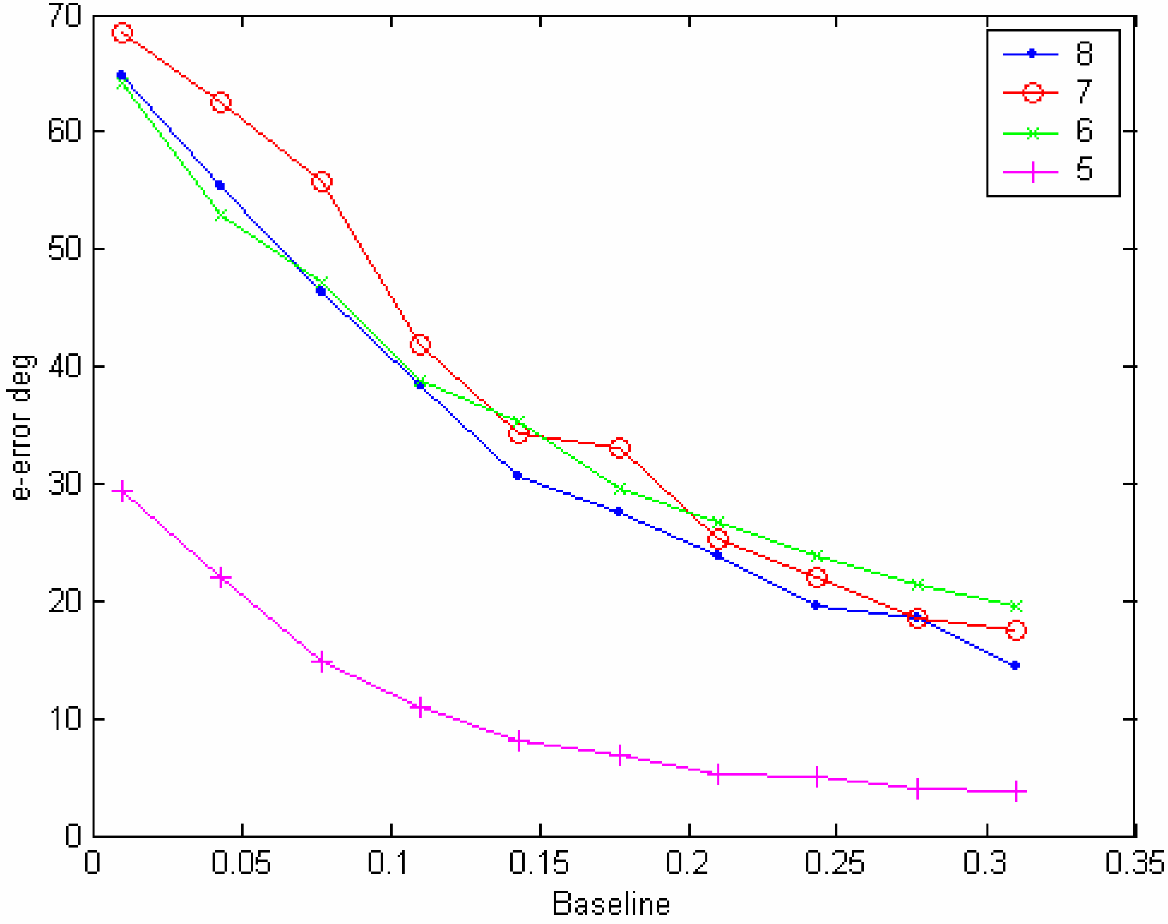
Direction



50 points

Depth 0.5
Baseline 0.1
Field of View 45 degrees

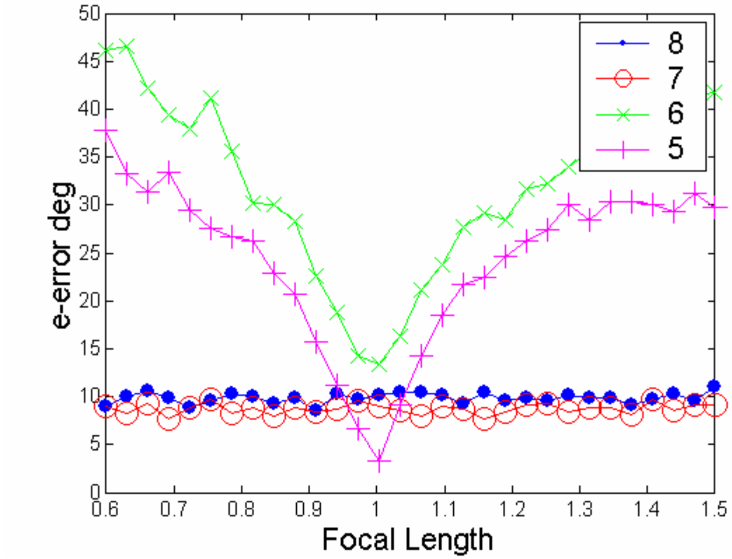
Baseline



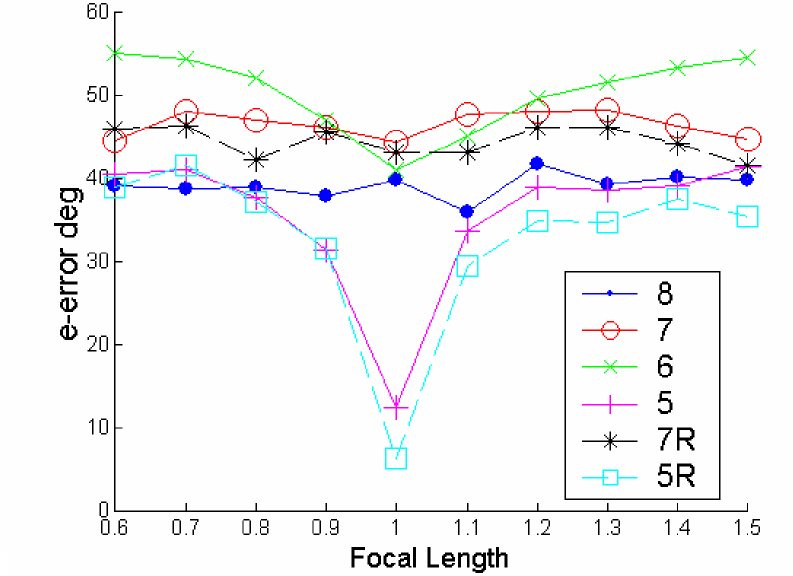
Minimal Cases, Sideways Motion

Depth 0.5
Baseline 0.1
Field of View 45 degrees

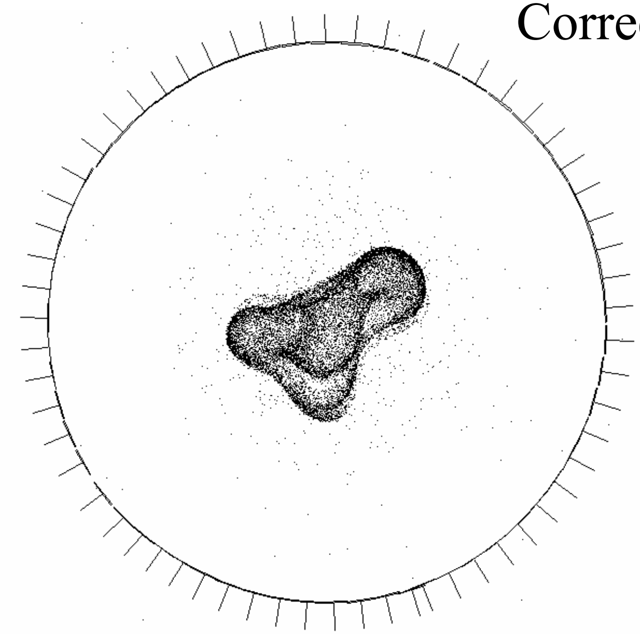
Easy Conditions

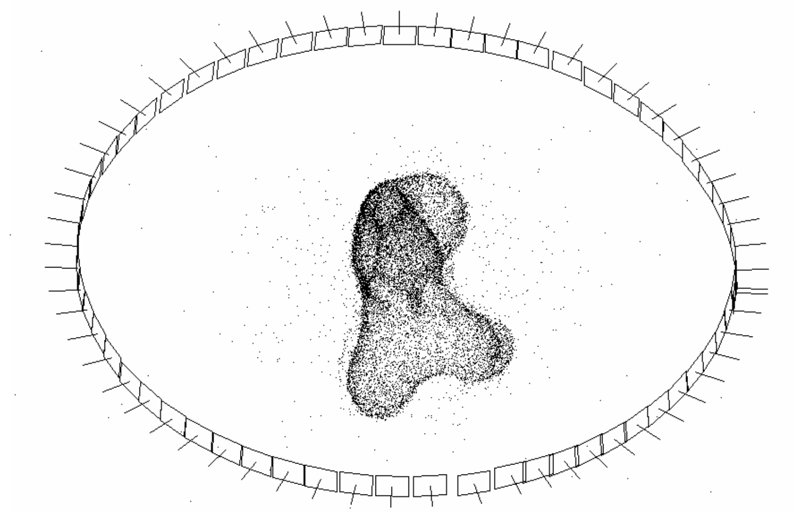


Realistic Conditions



Correct Calibration

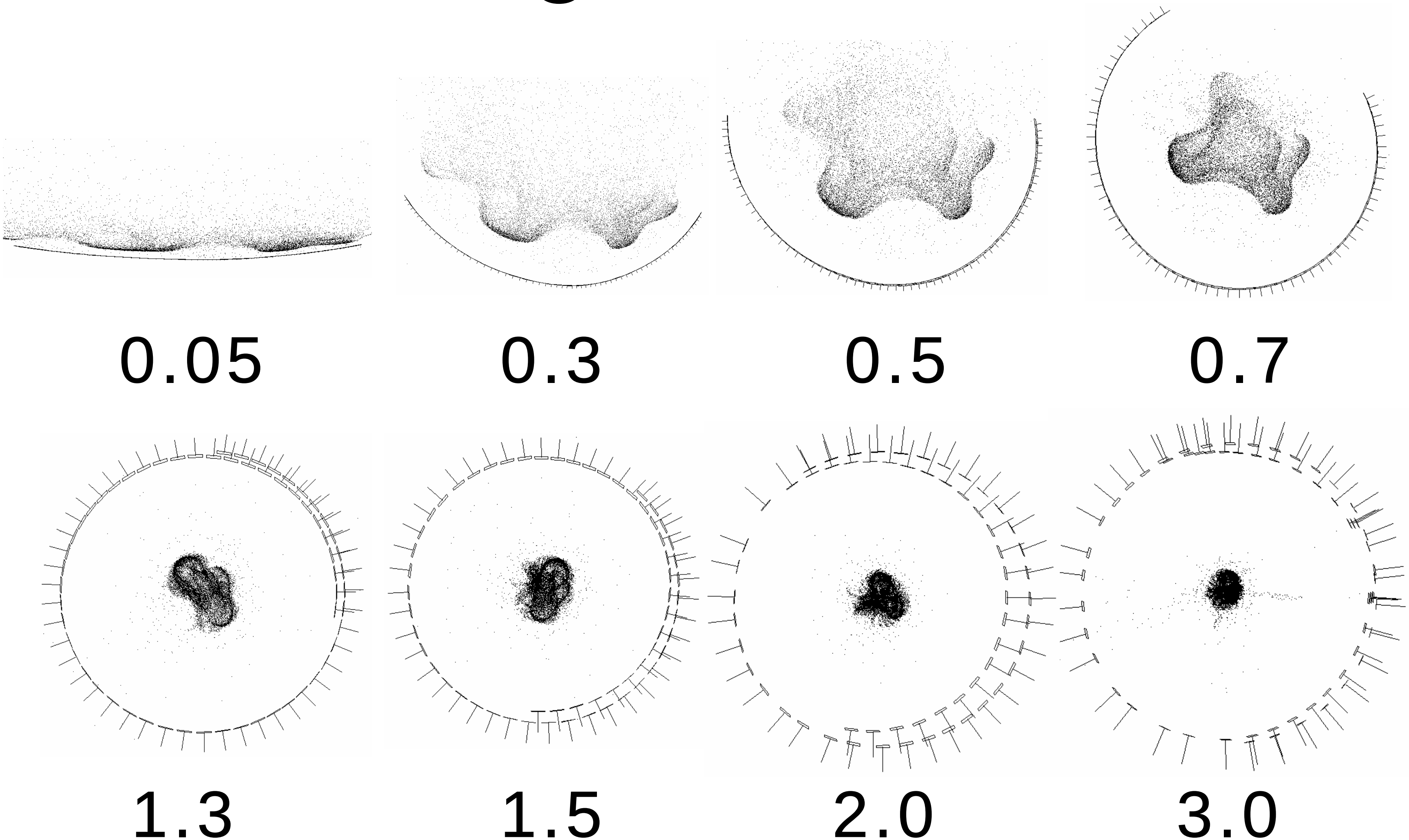




UK

Center for Visualization & Virtual Environments

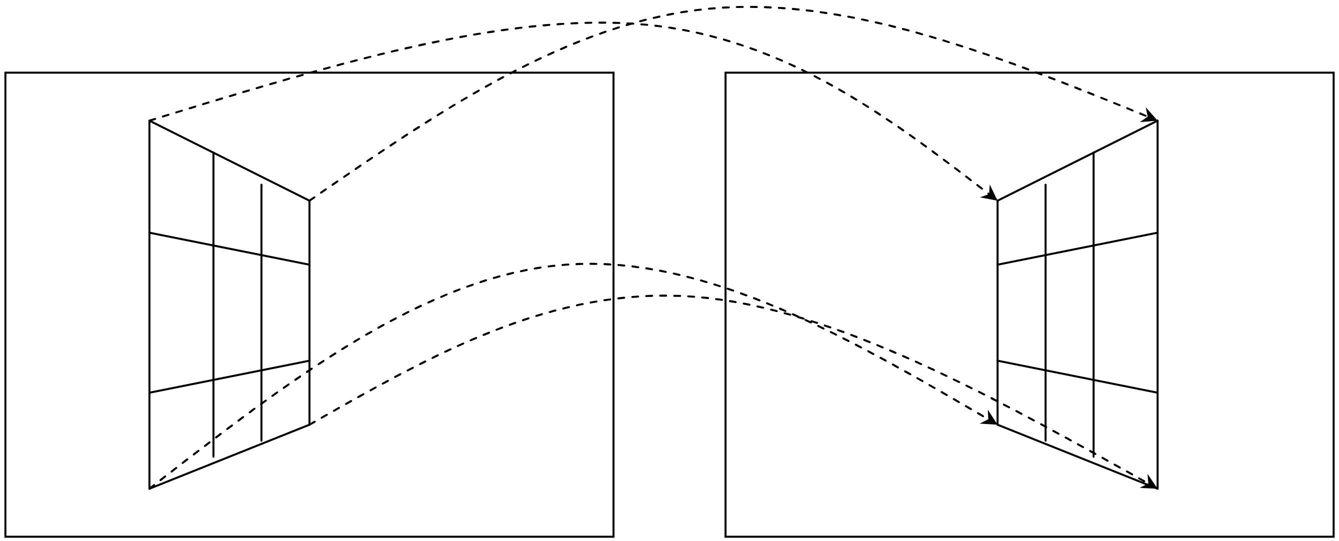
Focal Length Miscalibration



UK

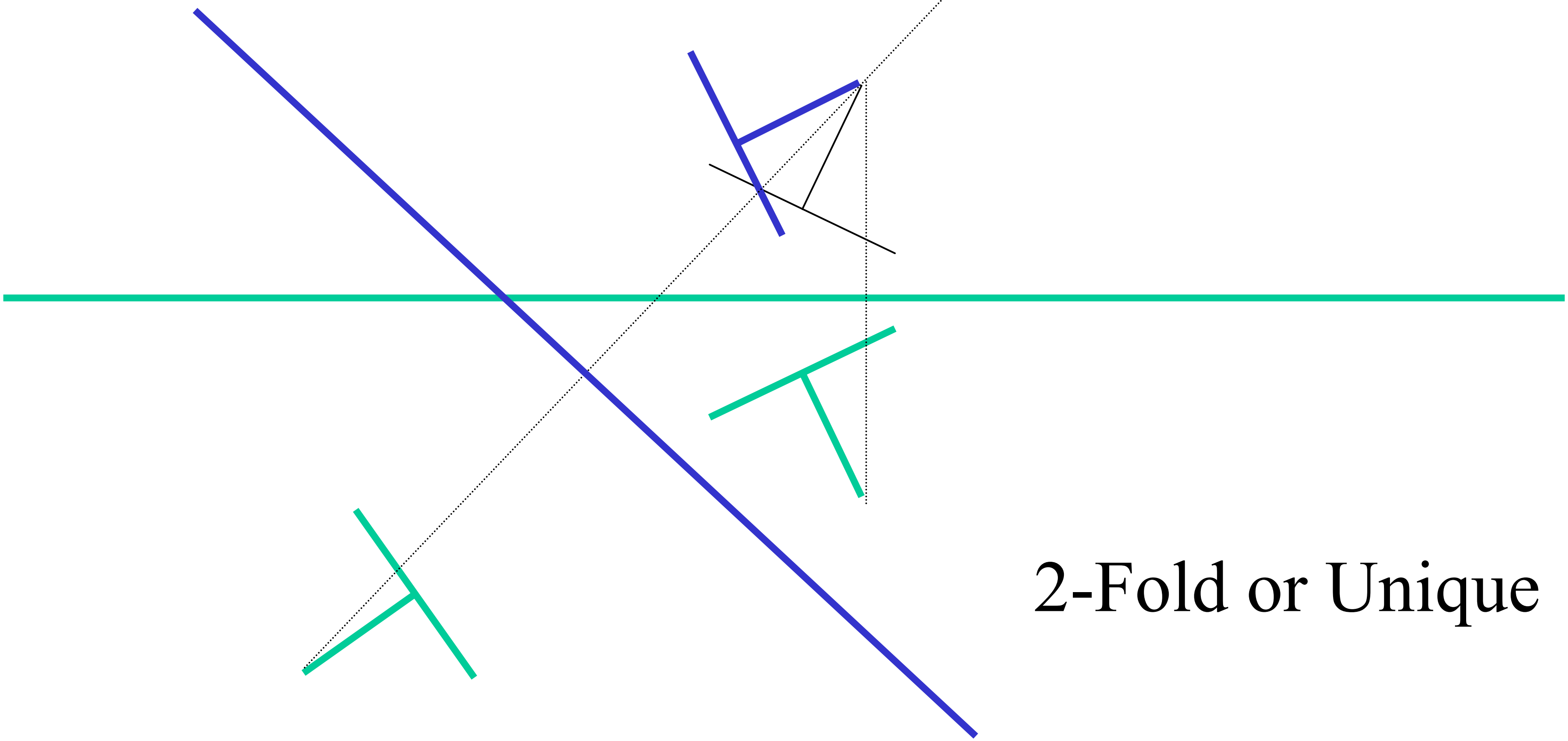
Center for Visualization & Virtual Environments

Planar Ambiguity, Uncalibrated



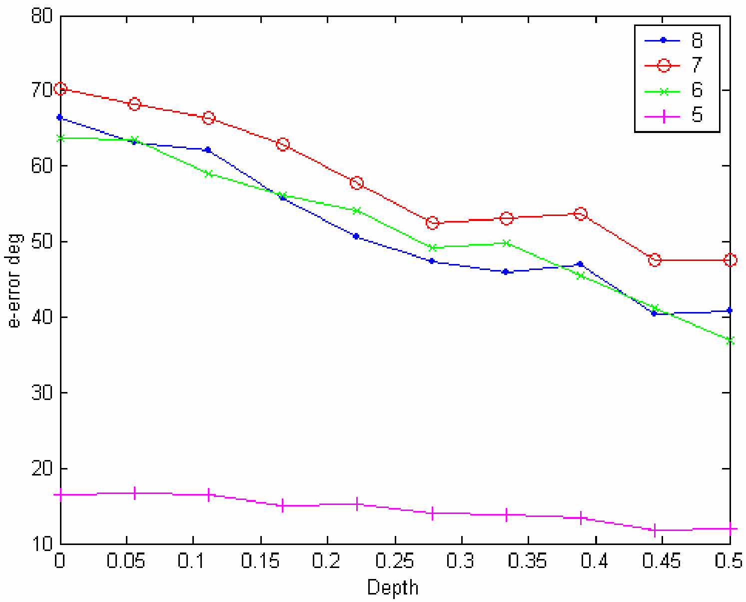
2Degrees of Freedom

Planar Ambiguity, Calibrated



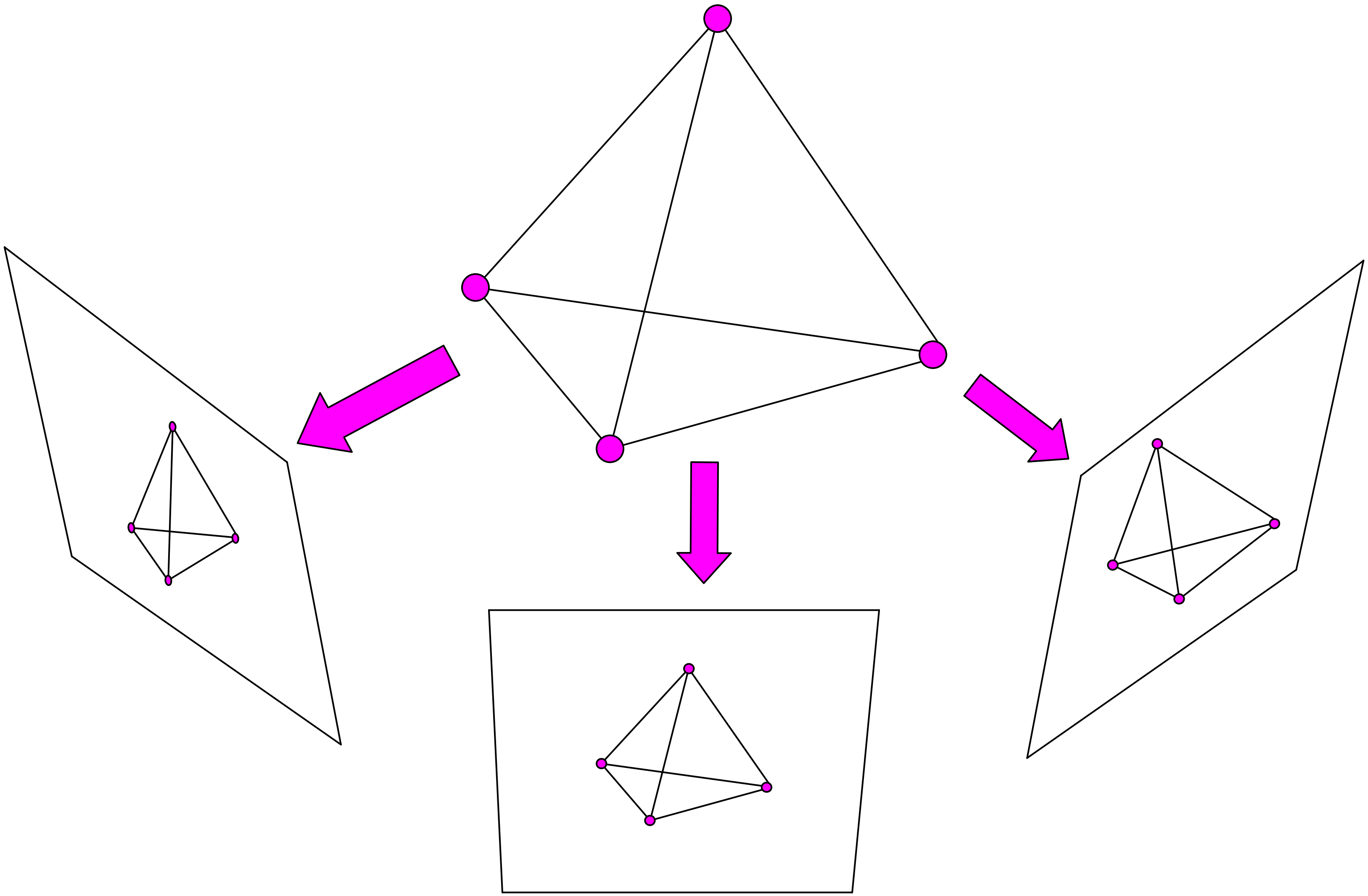
2-Fold or Unique

Depth

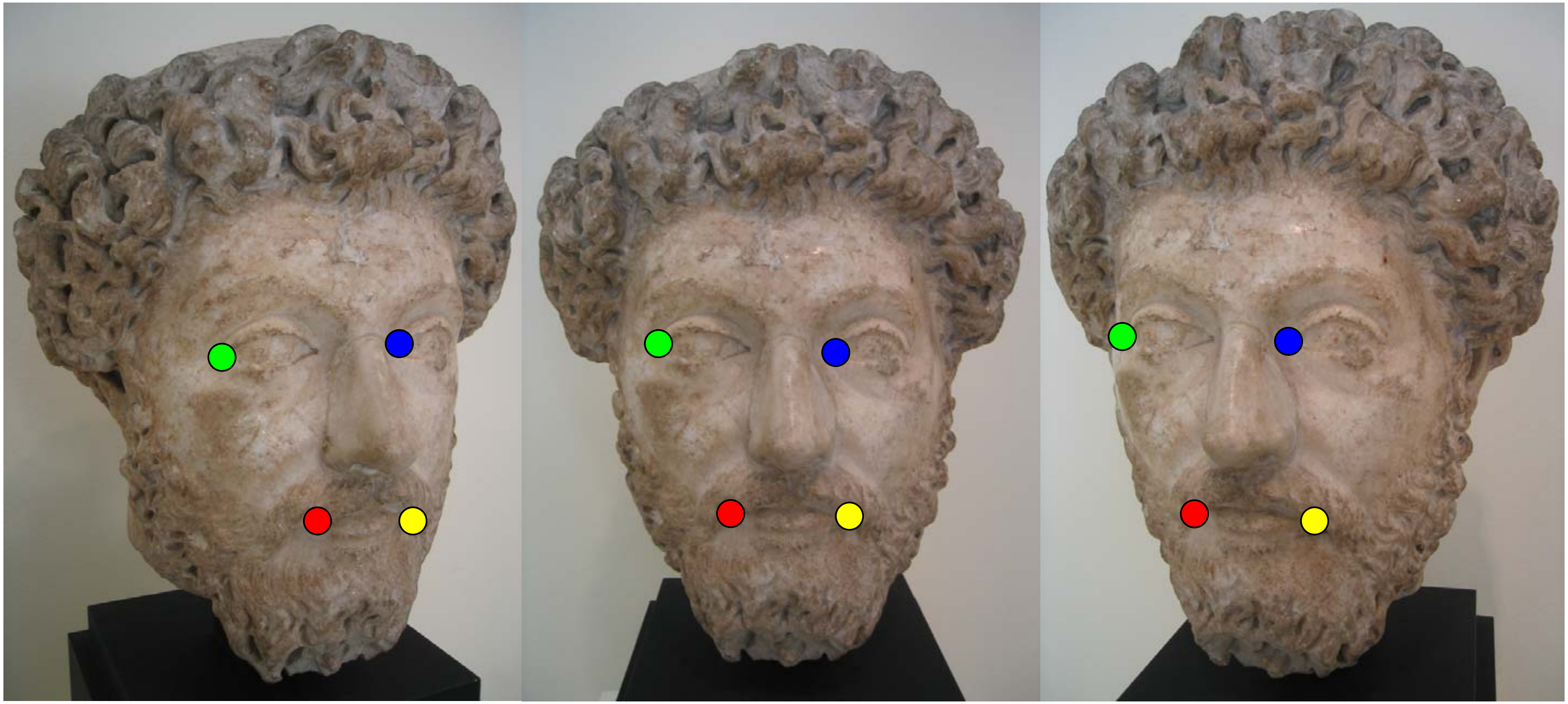


Center for Visualization & Virtual Environments

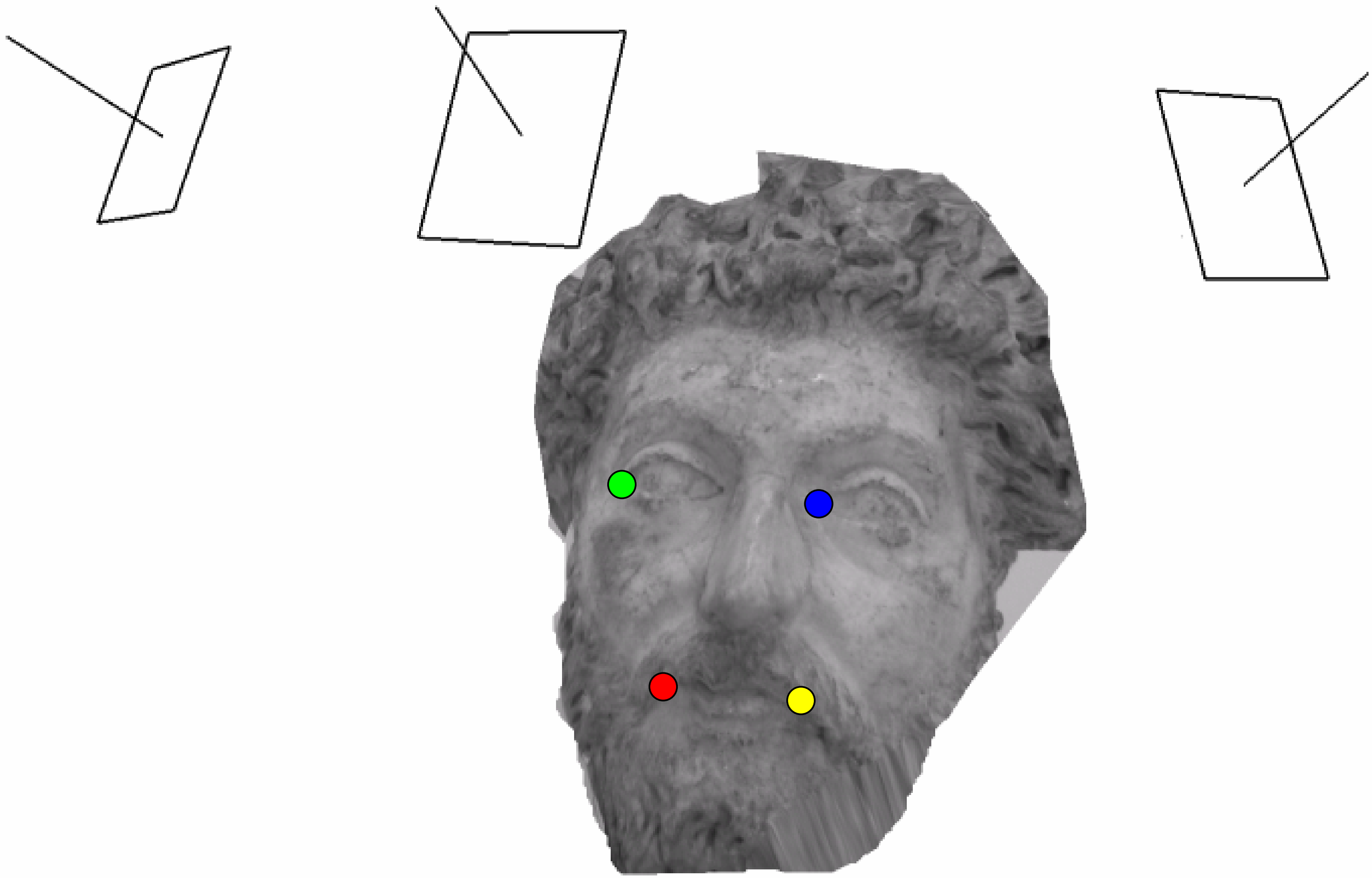
The 3 View 4 Point Problem



Center for Visualization & Virtual Environments

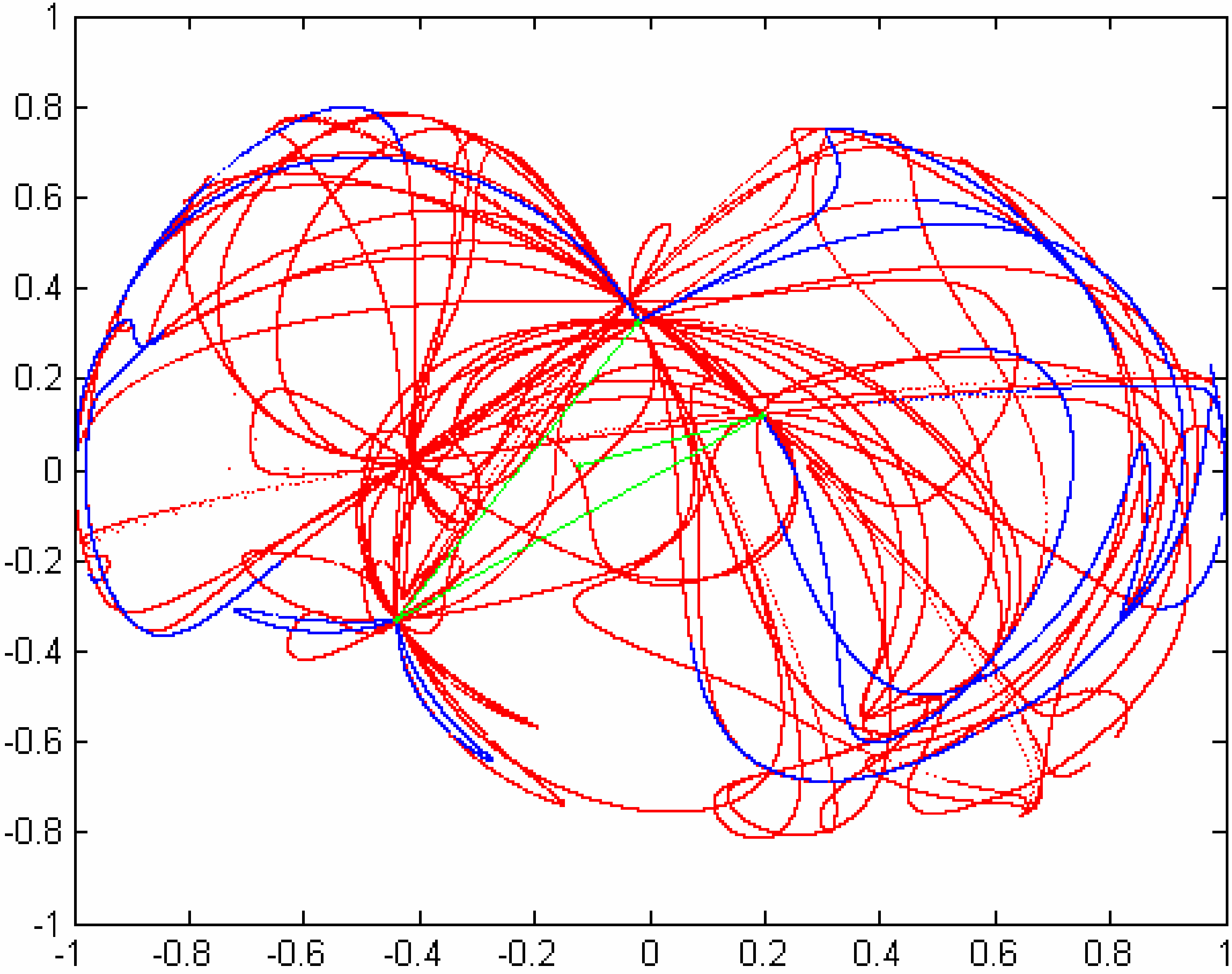


Center for Visualization & Virtual Environments



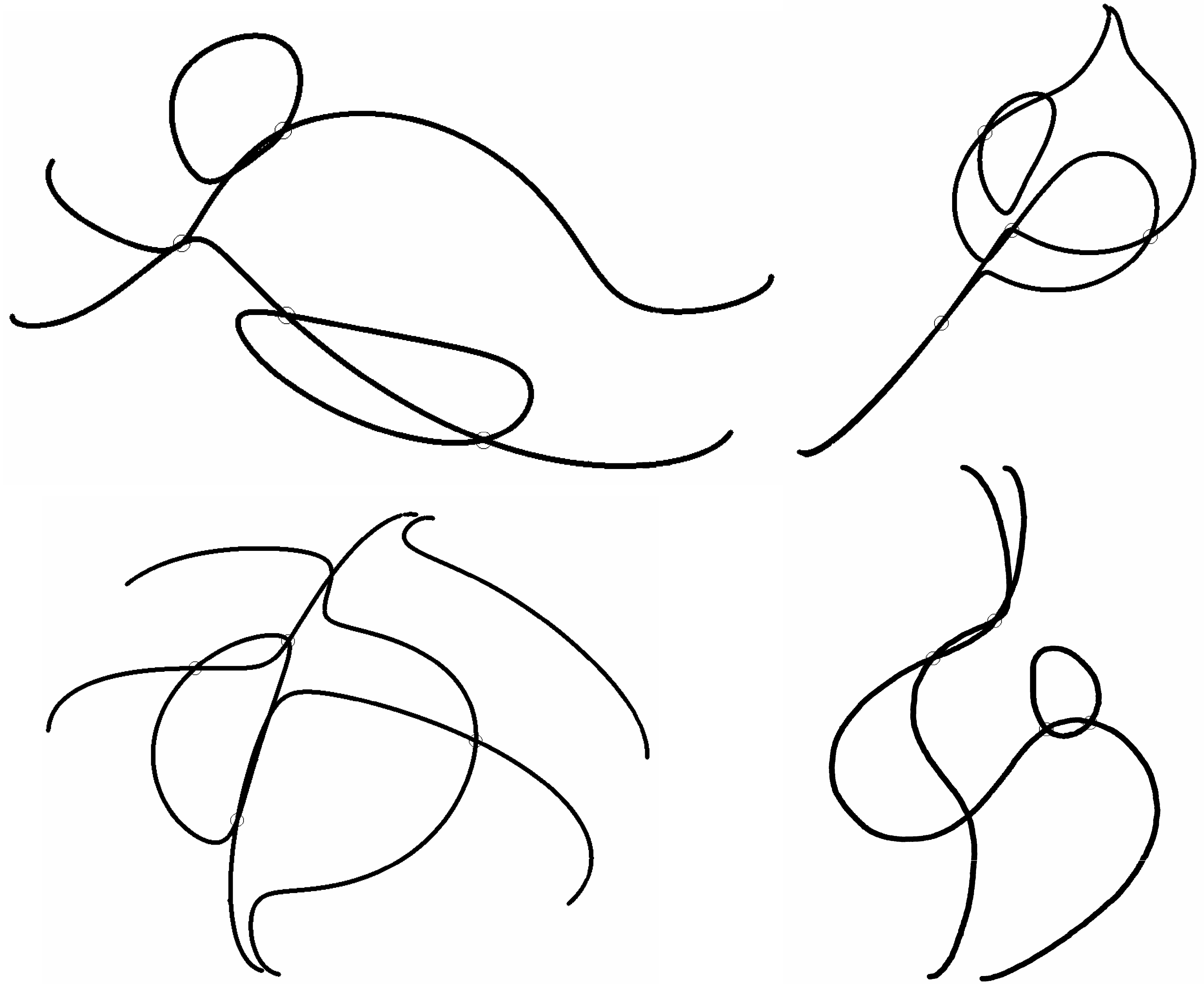
How Hard is this Problem?

Approximately This Hard



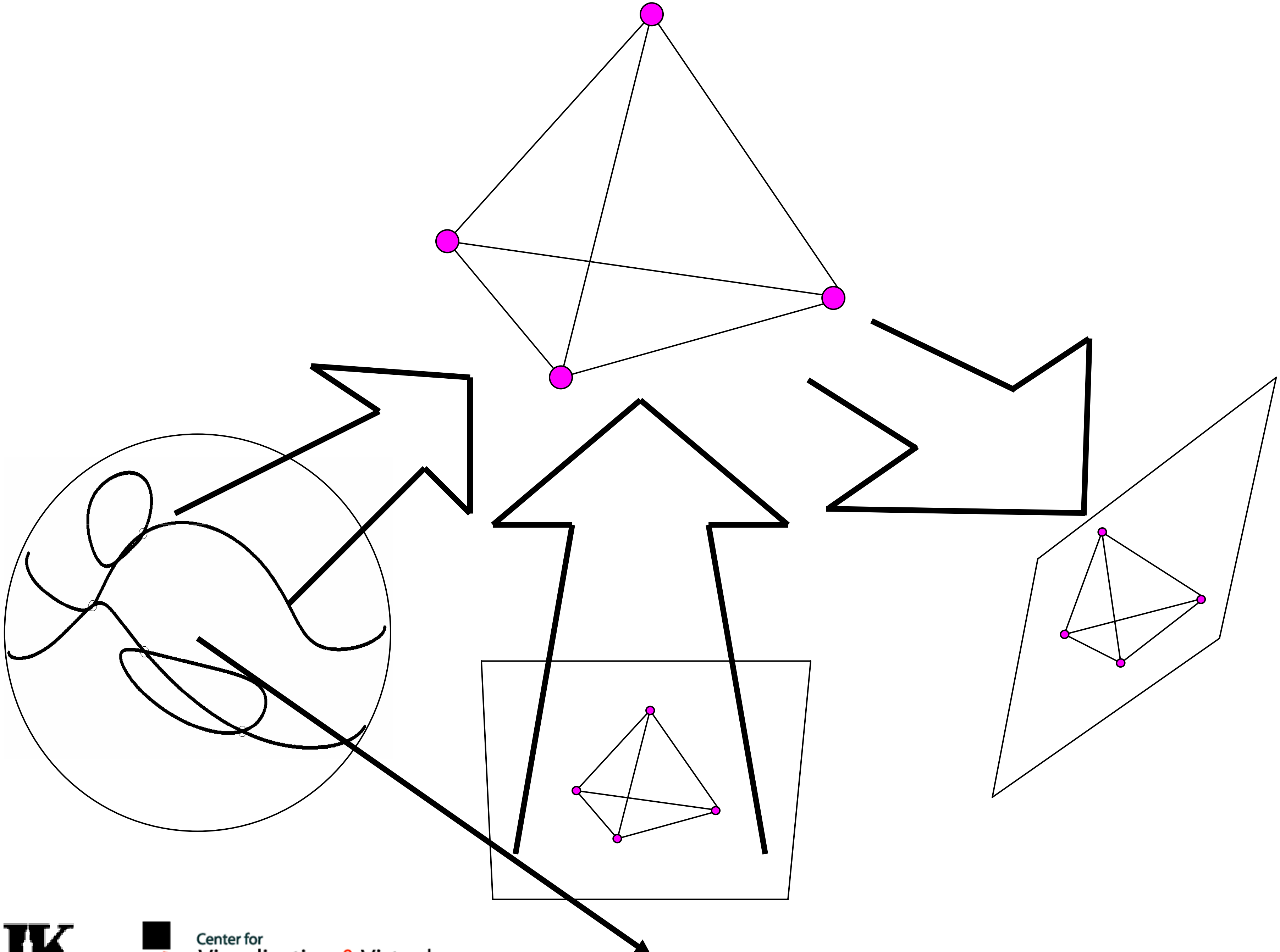
UK

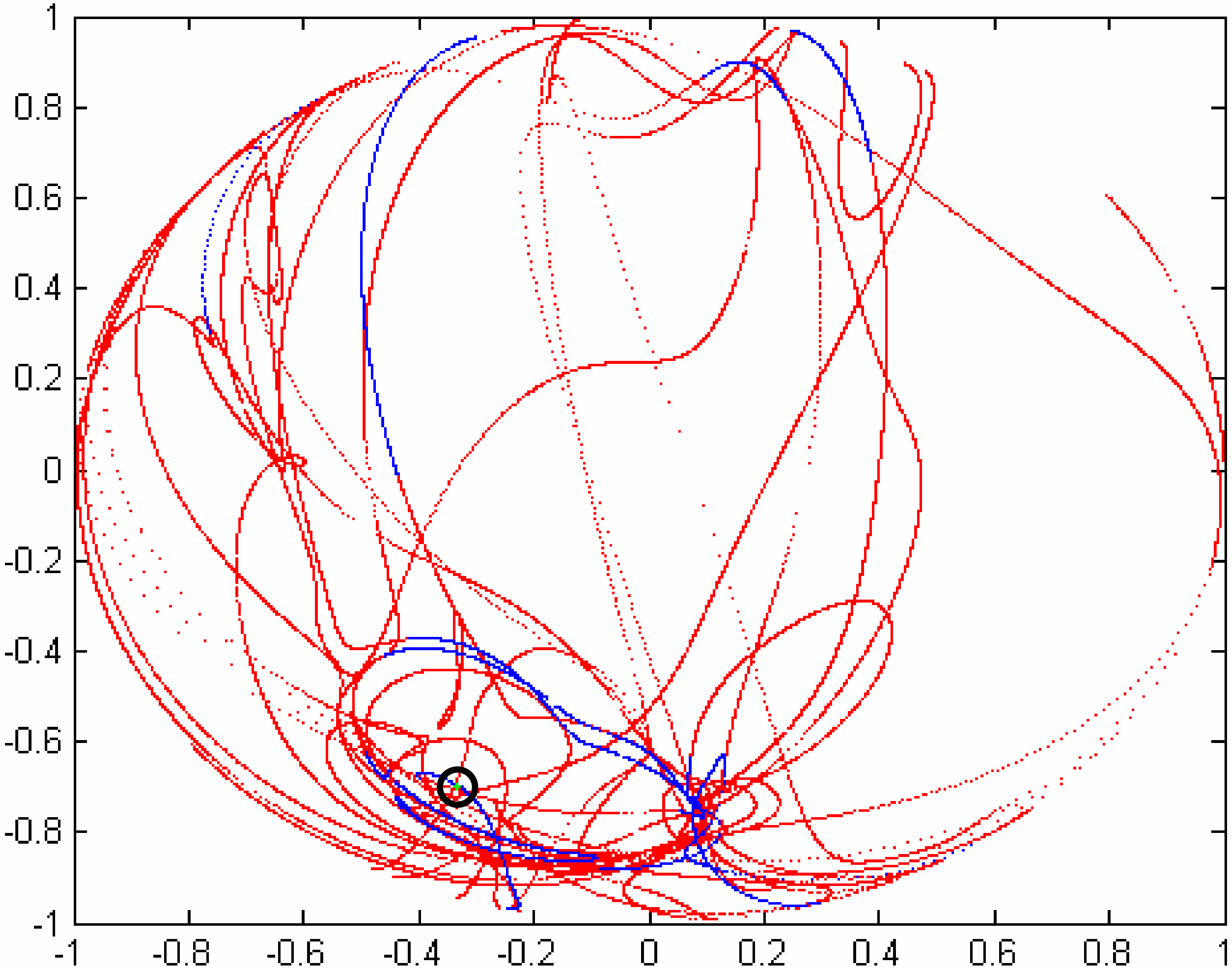
Center for Visualization & Virtual Environments



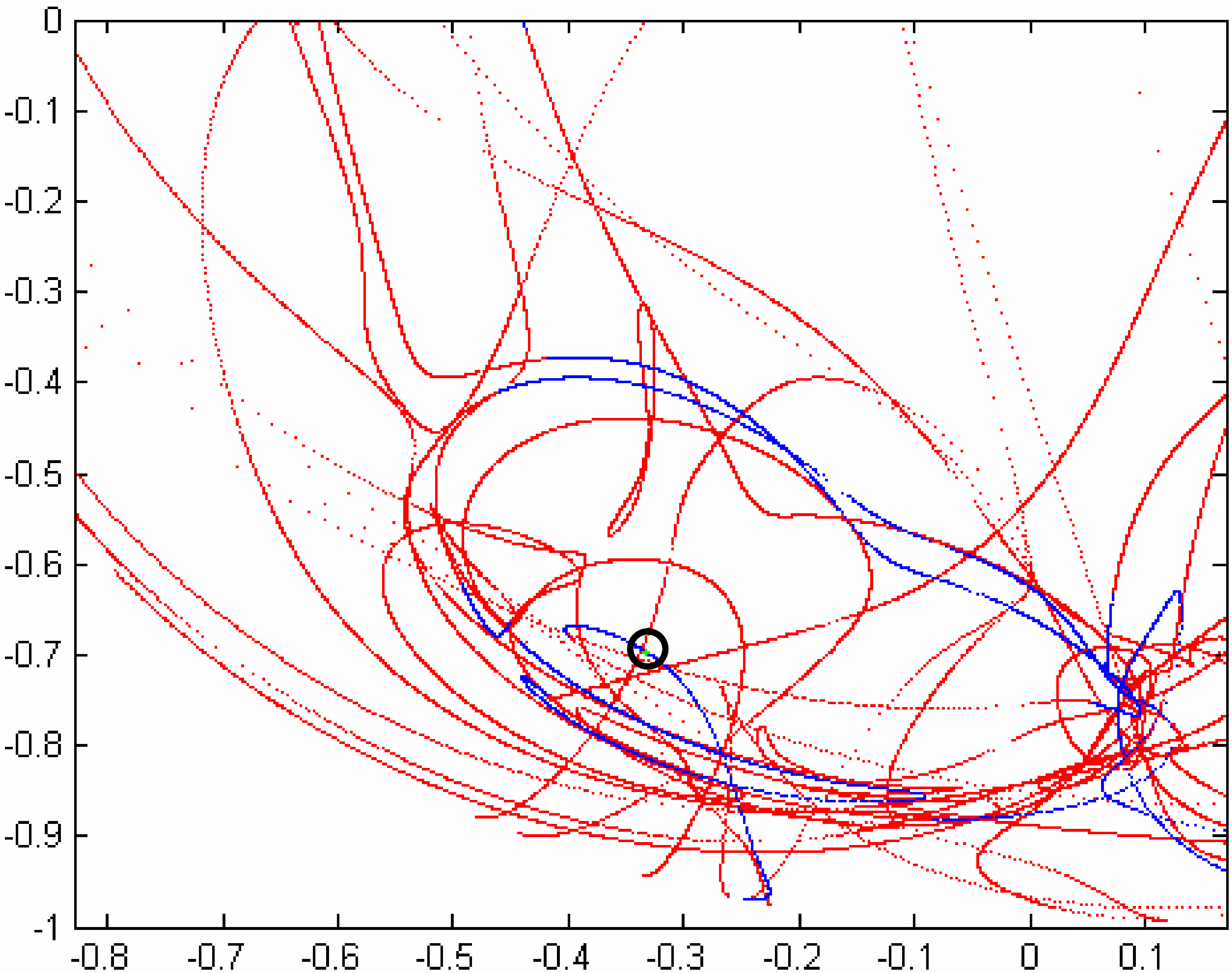
UK

Center for Visualization & Virtual Environments



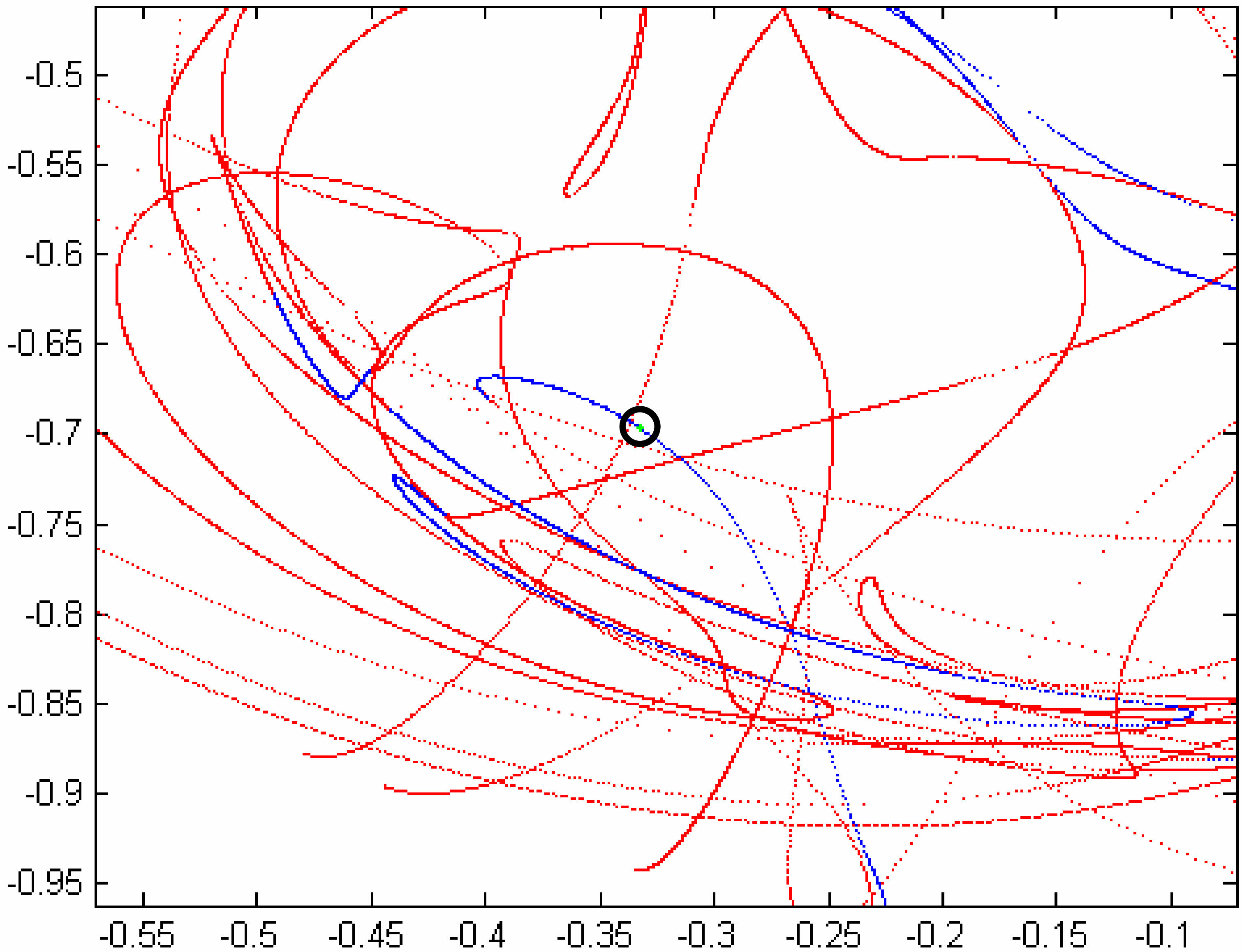


Center for Visualization & Virtual Environments

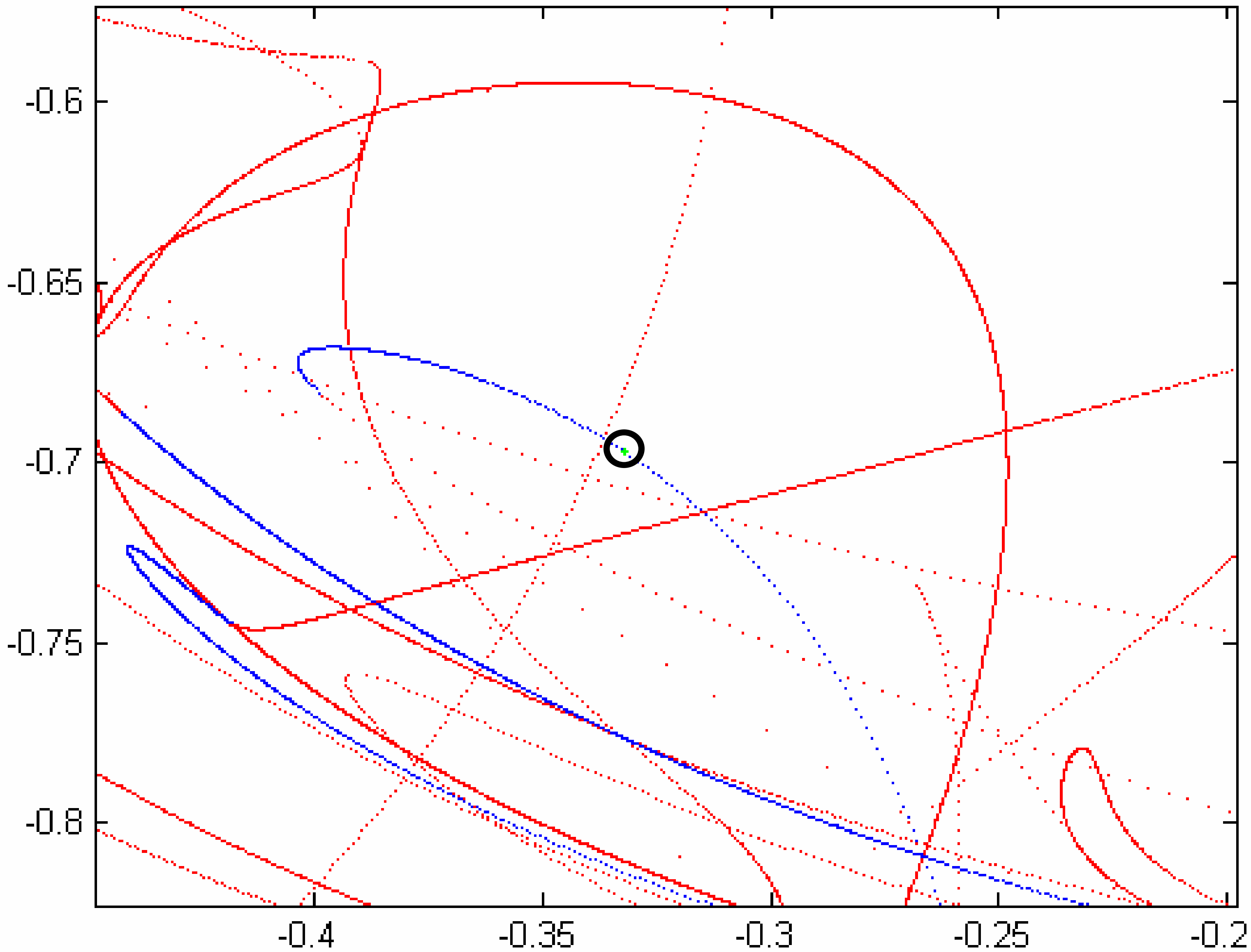


UK

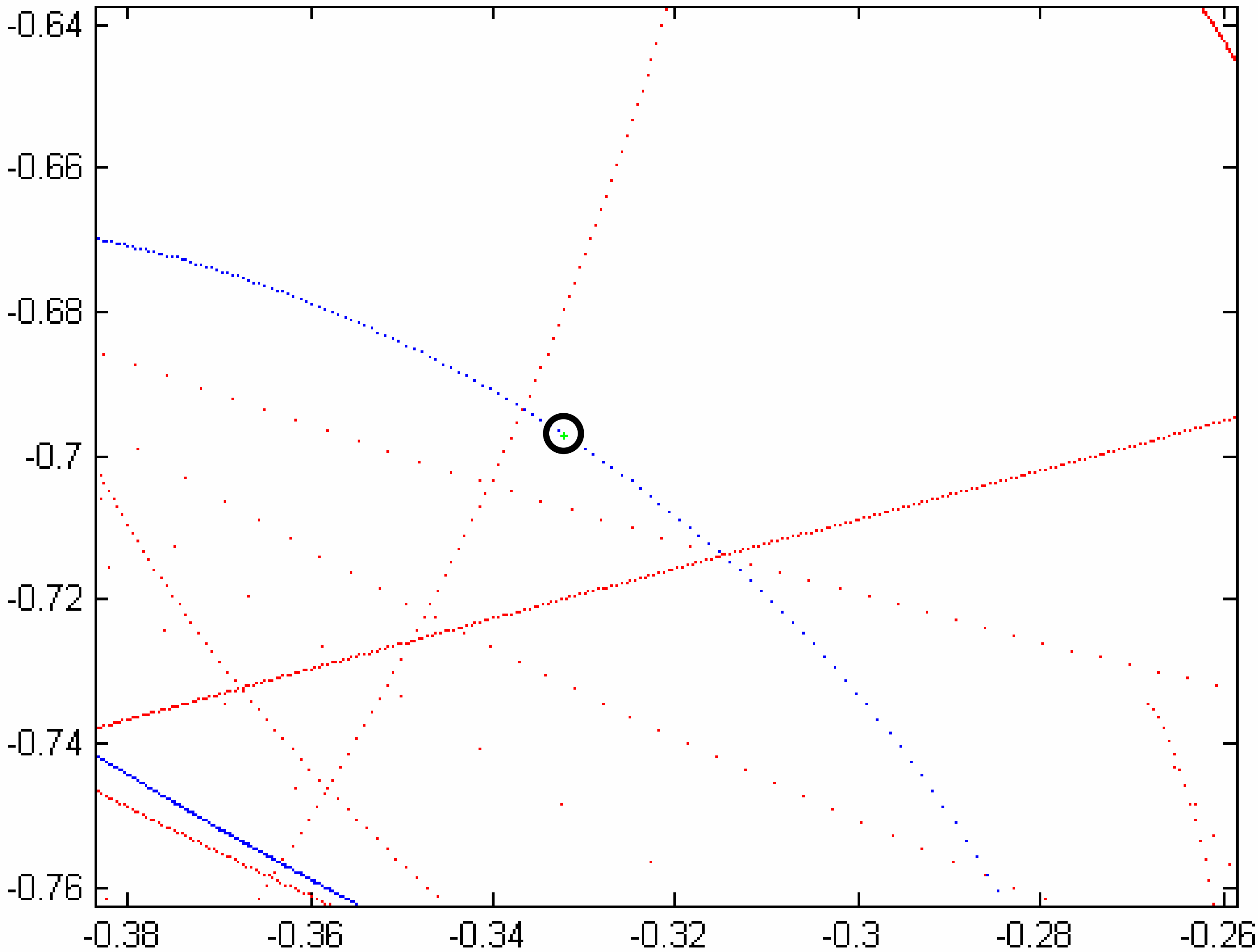
Center for Visualization & Virtual Environments



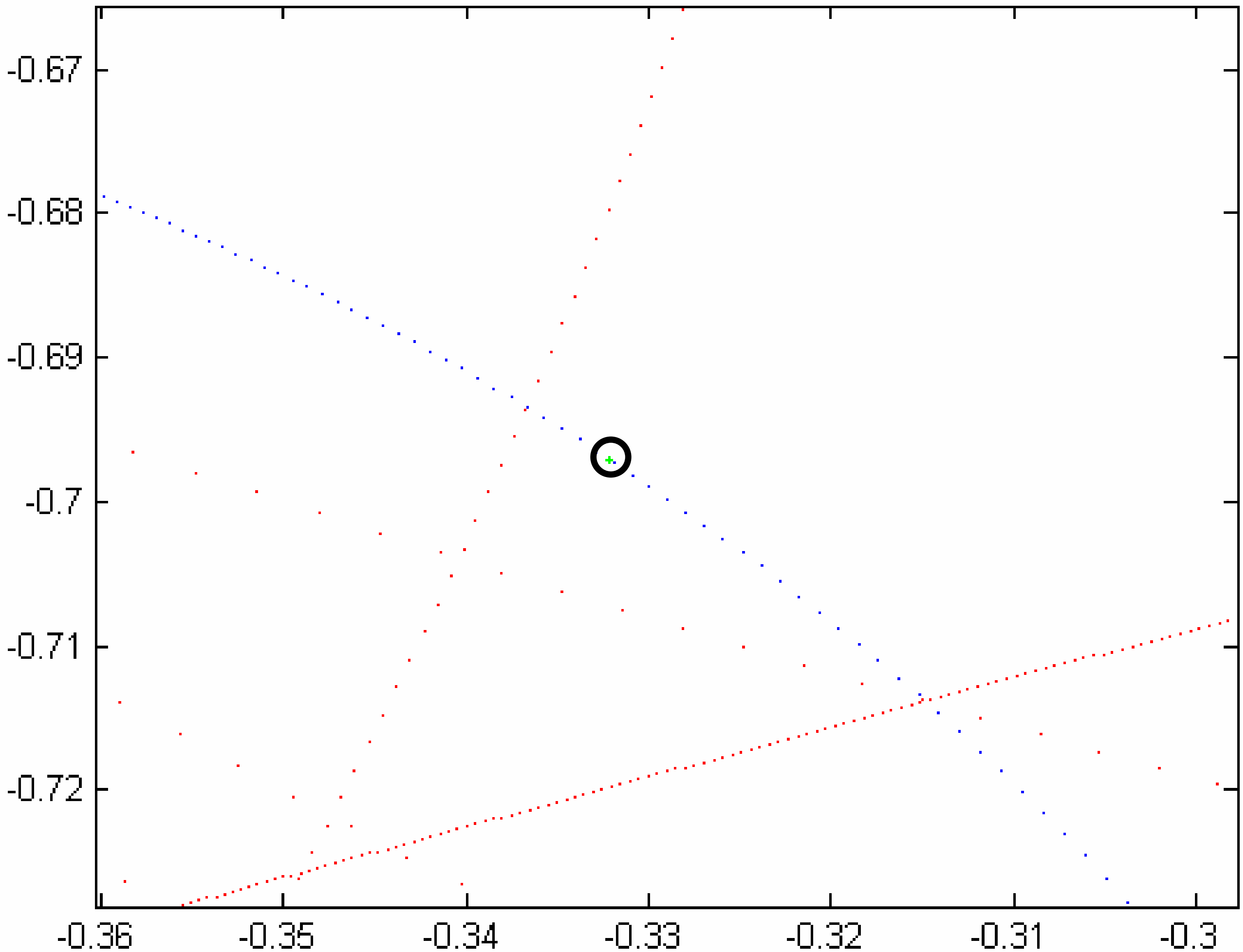
Center for Visualization & Virtual Environments



Center for Visualization & Virtual Environments



Center for Visualization & Virtual Environments



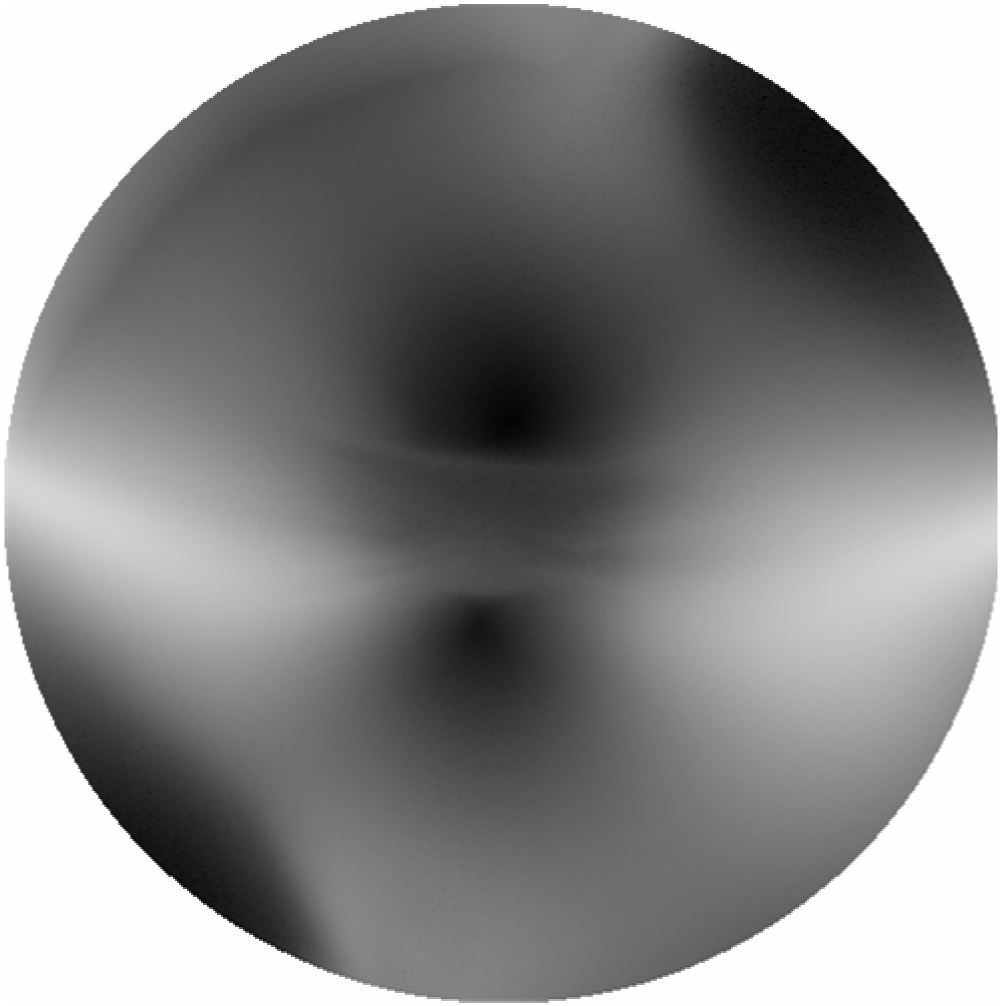
UK Center for Visualization & Virtual Environments

Uncertainty in Epipolar Geometry

work with Chris Engels

Single Estimate often ill posed

Representation of posterior likelihood well posed, but computationally challenging



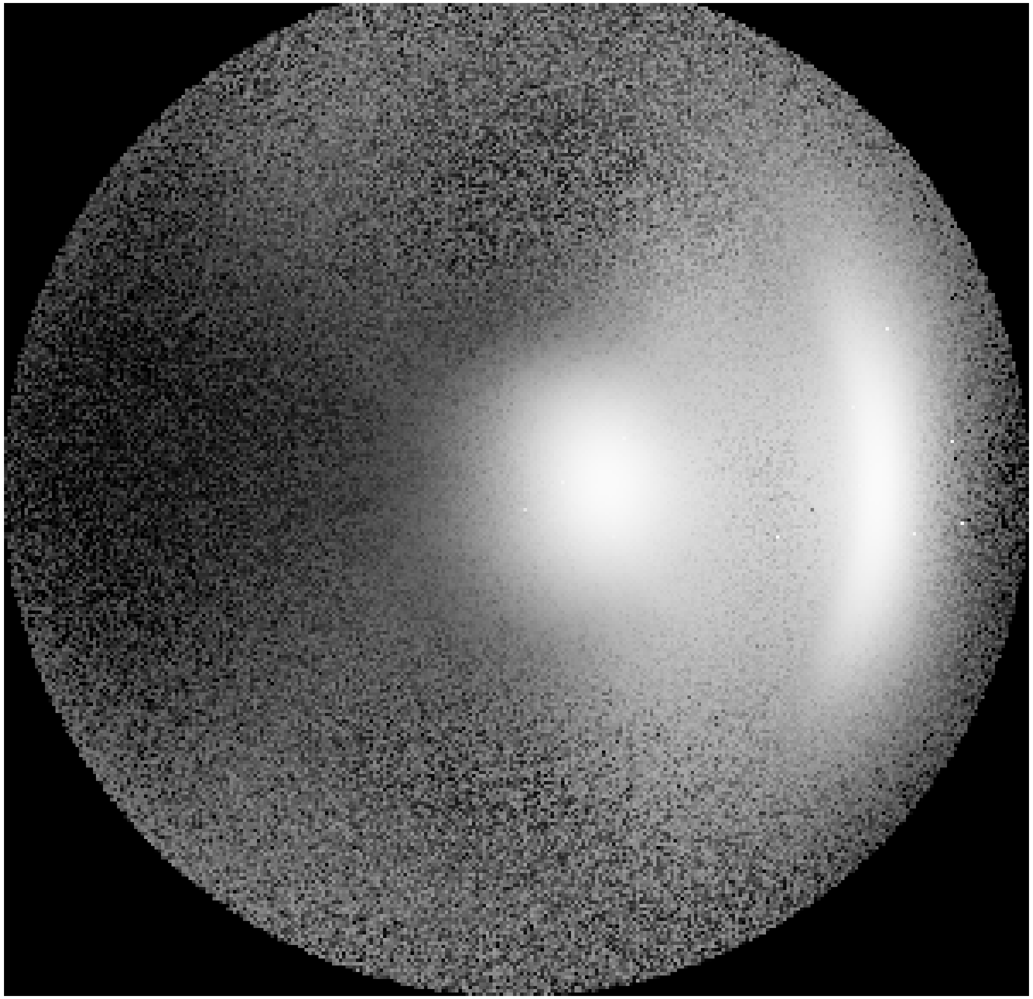
Center for Visualization & Virtual Environments

Uncertainty in Epipolar Geometry

work with Chris Engels

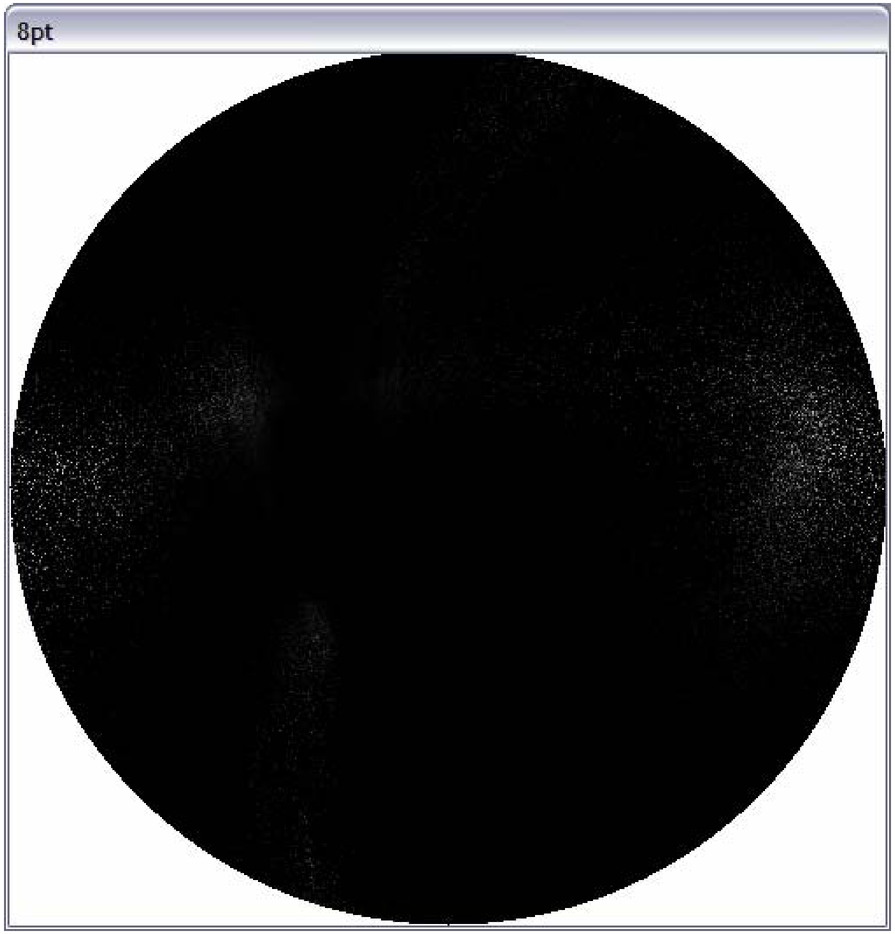
Single Estimate often ill posed

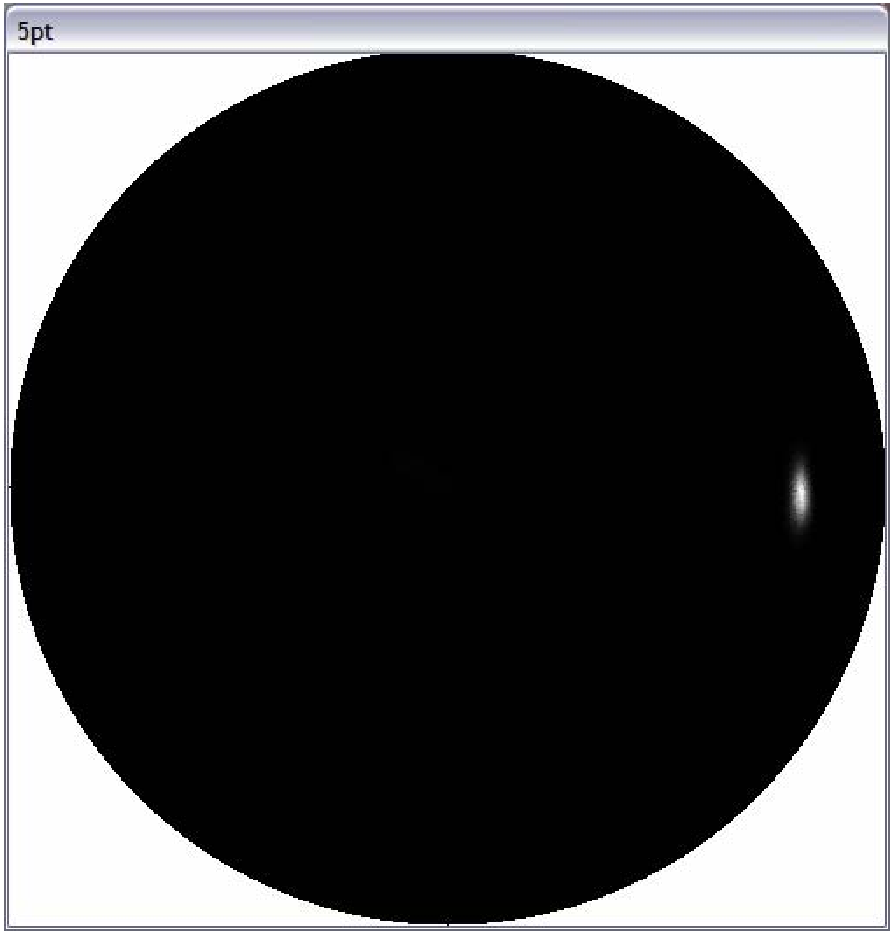
Representation of posterior likelihood well posed, but computationally challenging



Epipoloscope

work with Chris Engels

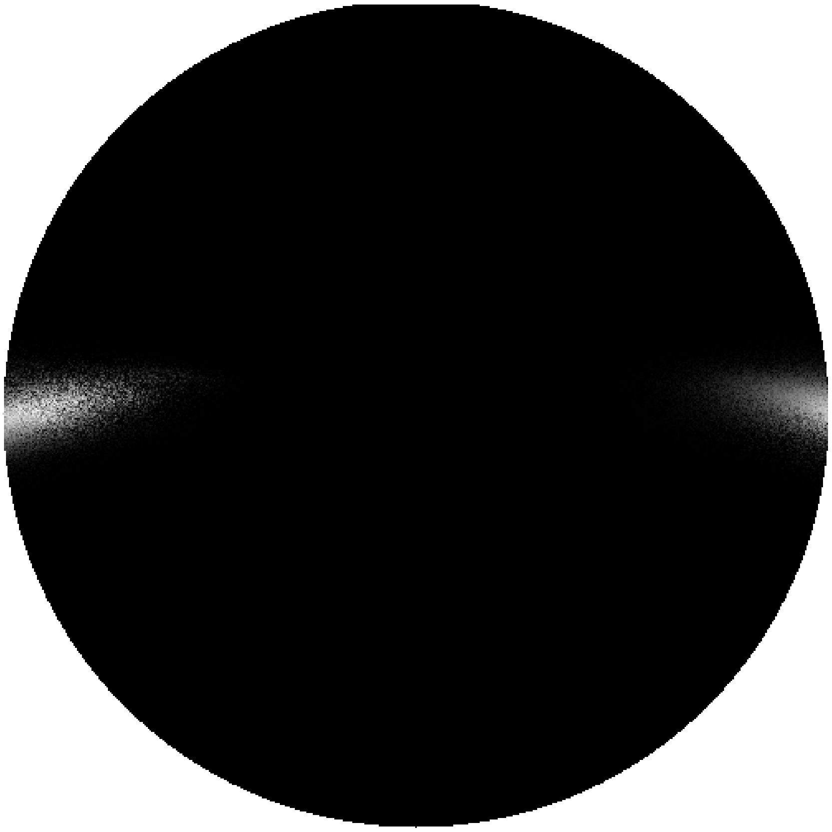




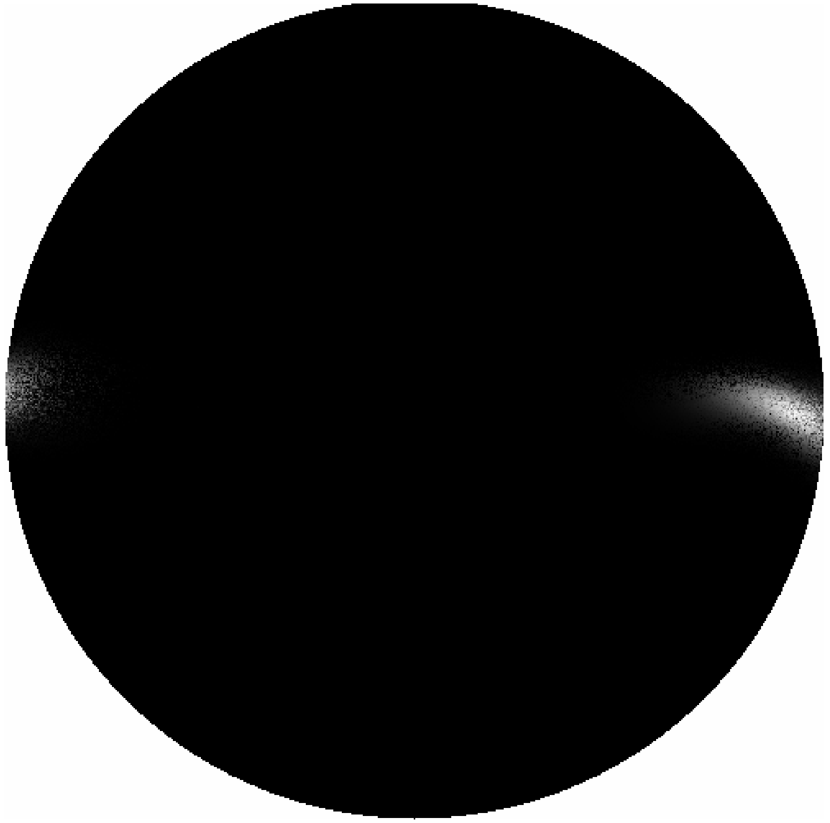
UK Center for Visualization & Virtual Environments

Epipoloscope

work with Chris Engels



8 point



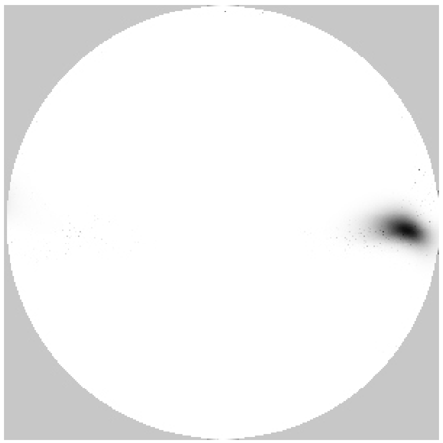
5 point

Hypothesis Generators

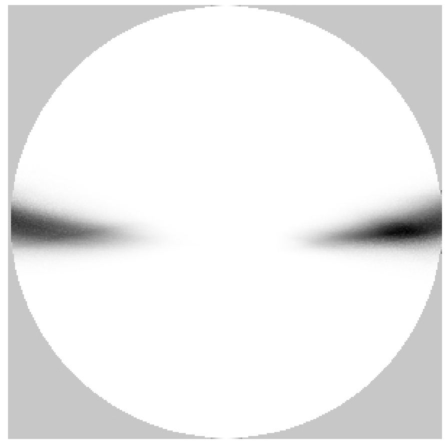
- Partially data-driven methods
  - Five-point + epipole
  - Three-point + epipole (uses intrinsic calibration)
- Fully data-driven methods:
  - Eight-point
  - Seven-point
  - Five-point (uses intrinsic calibration)

Results

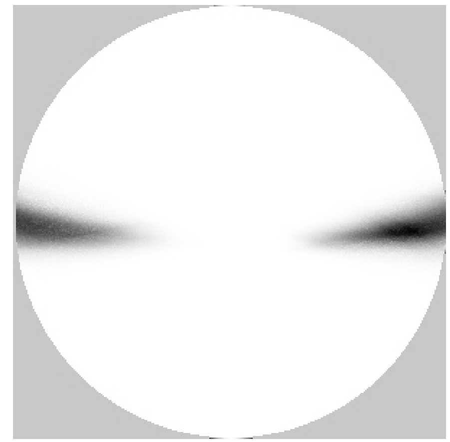
- Likelihood image using different methods



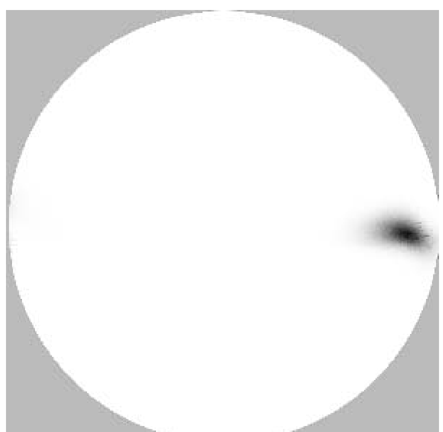
Five-Point



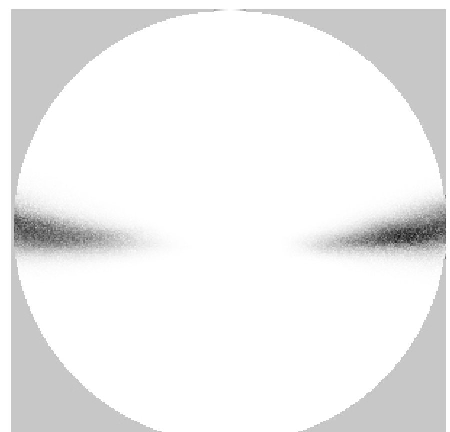
Seven-Point



Eight-Point



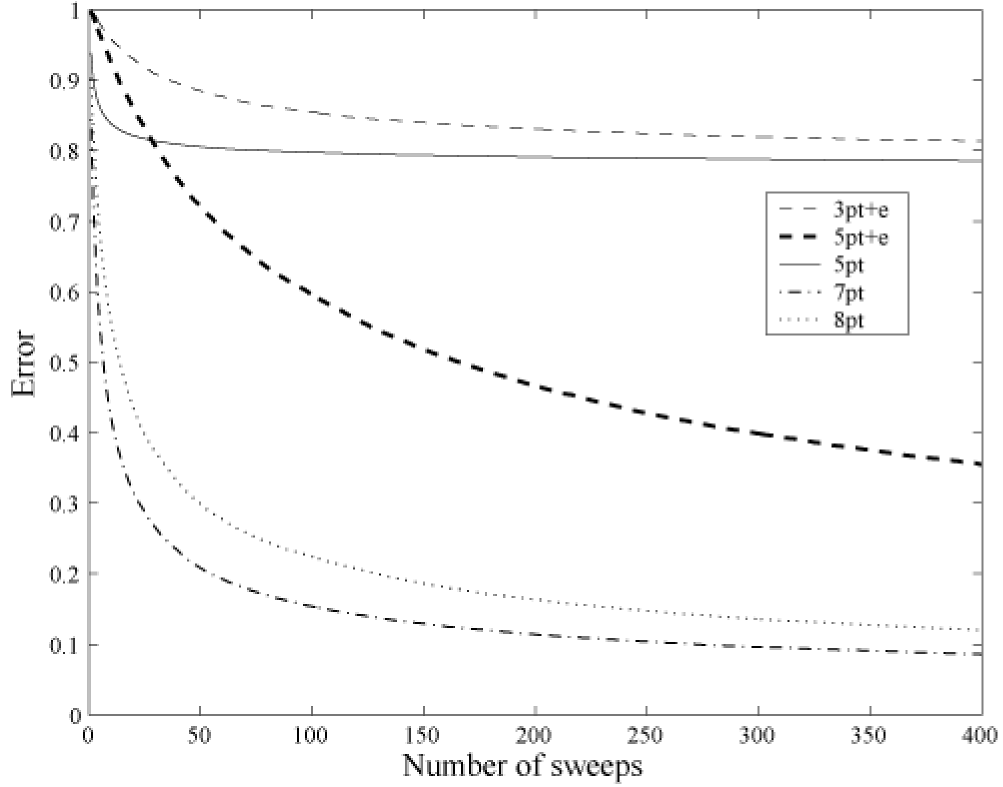
Three-Point + epipole



Five-Point + epipole

Results

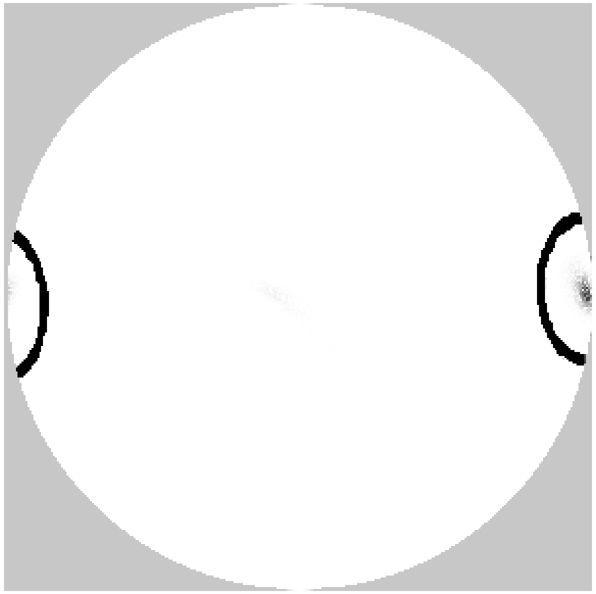
- Convergence of the posterior

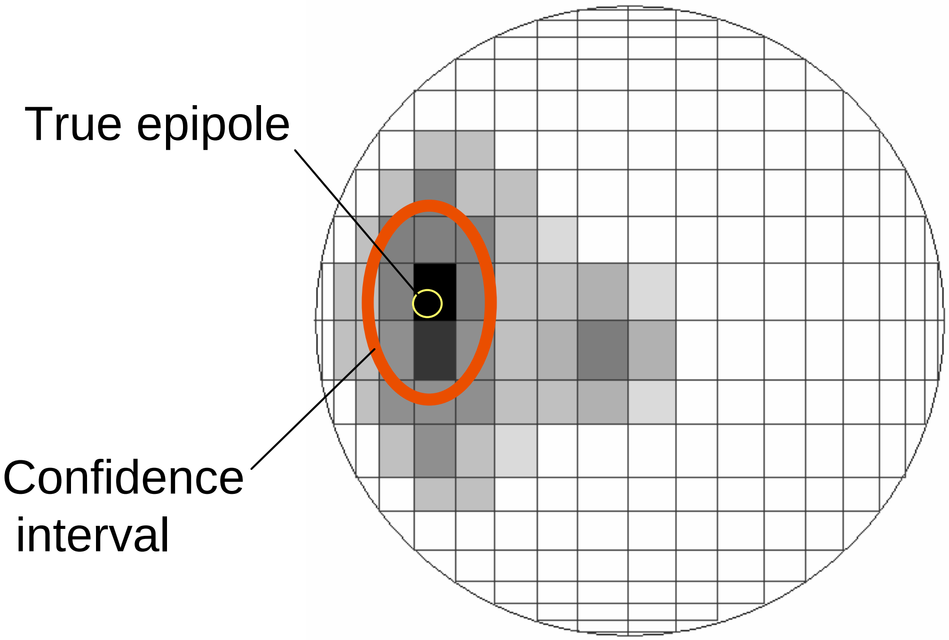


Center for Visualization & Virtual Environments

Results

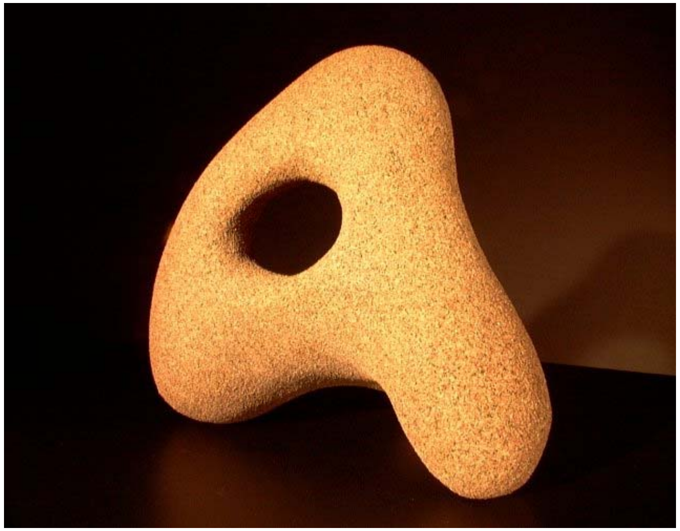
- Estimation of Confidence Interval
  - Confidence estimated by probability mass contained within certain interval





Results

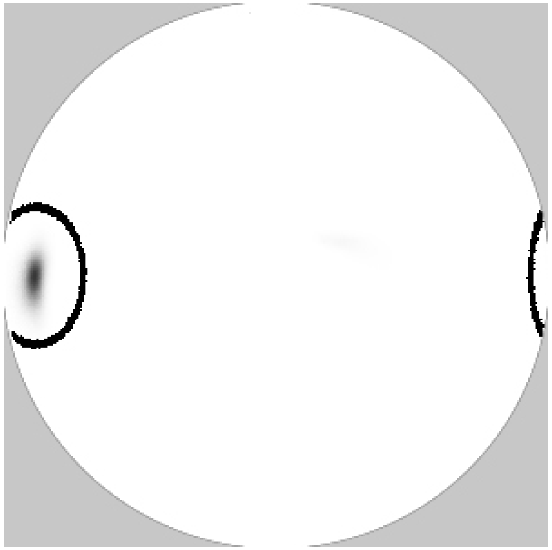
- Comparison of Confidence Intervals



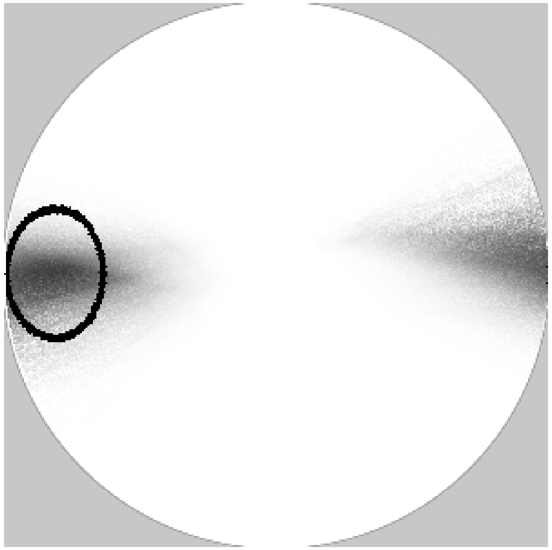
Results

- Comparison of Confidence Intervals
  - Fully Data-driven

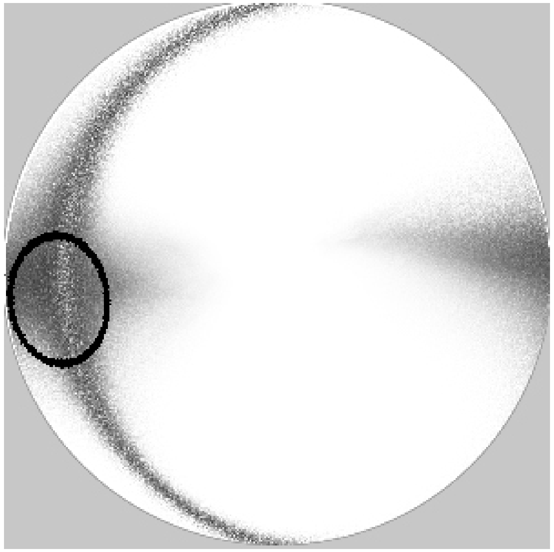
Five-Point
0.935666



Seven-Point
0.395411



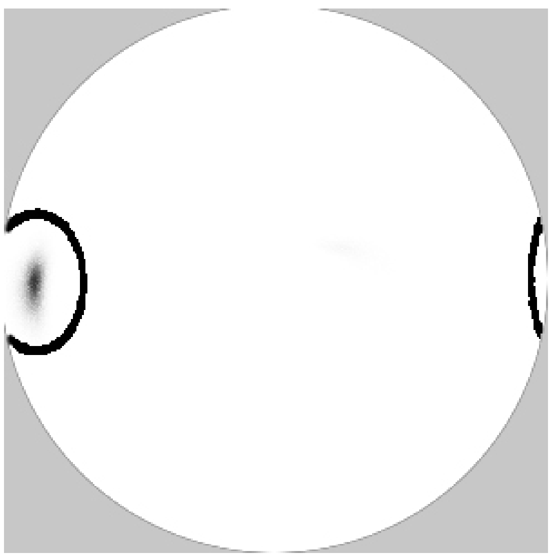
Eight-Point
0.277246



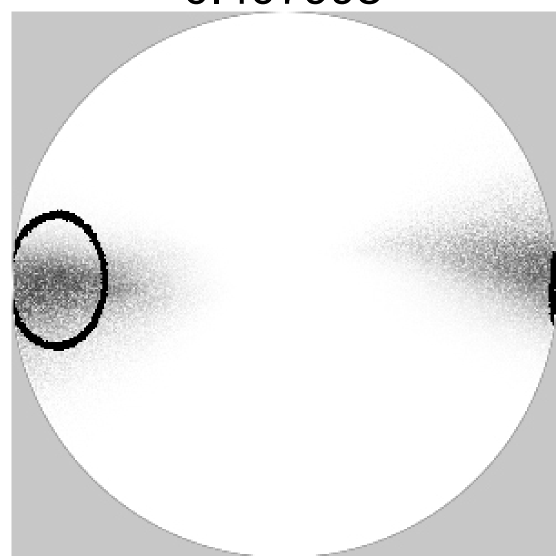
Results

- Comparison of Confidence Intervals
  - Partially Data-driven

Three-Point + epipole
0.937596

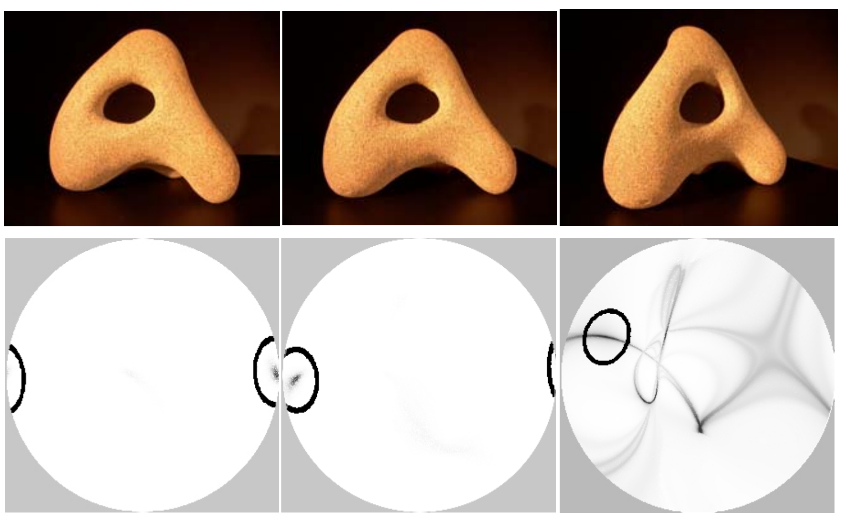


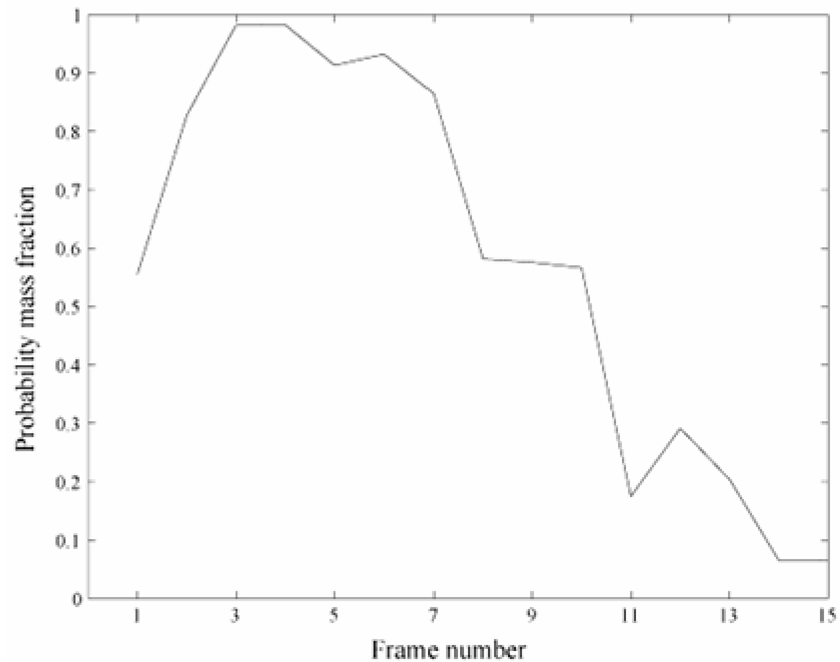
Five-Point + epipole
0.407995



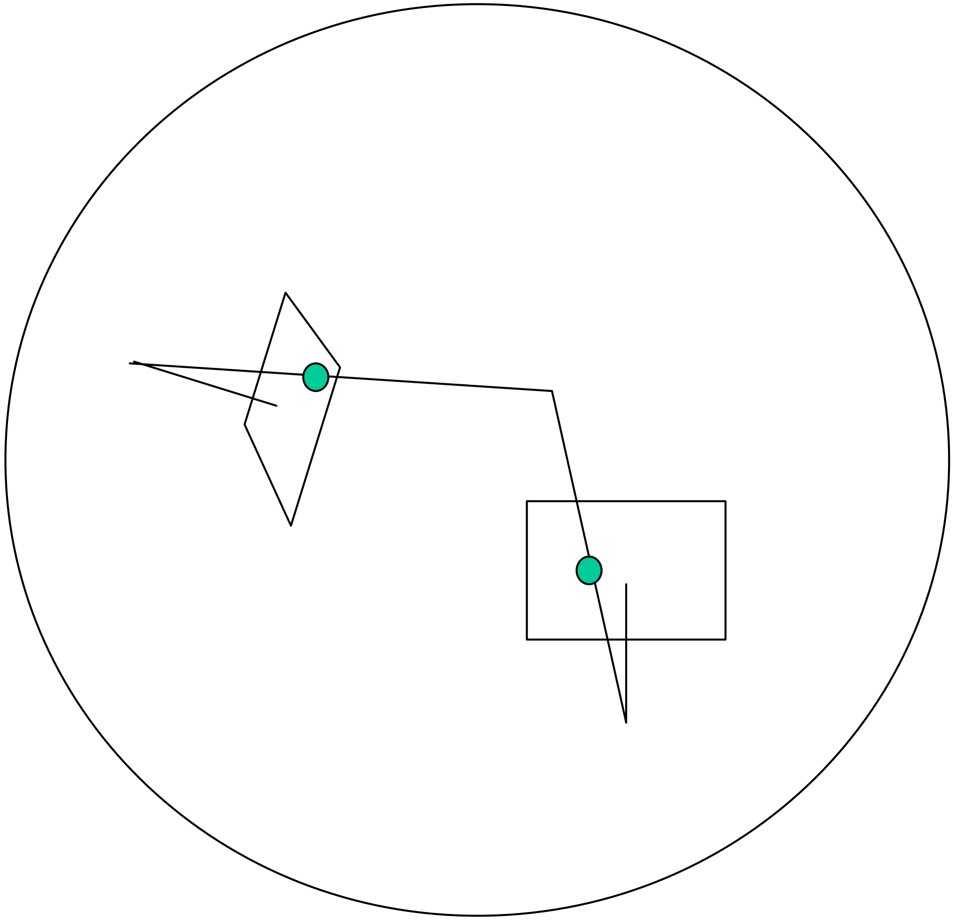
Results

- Baseline Selection
  - Choose best pair of frames for pose, stereo, etc.



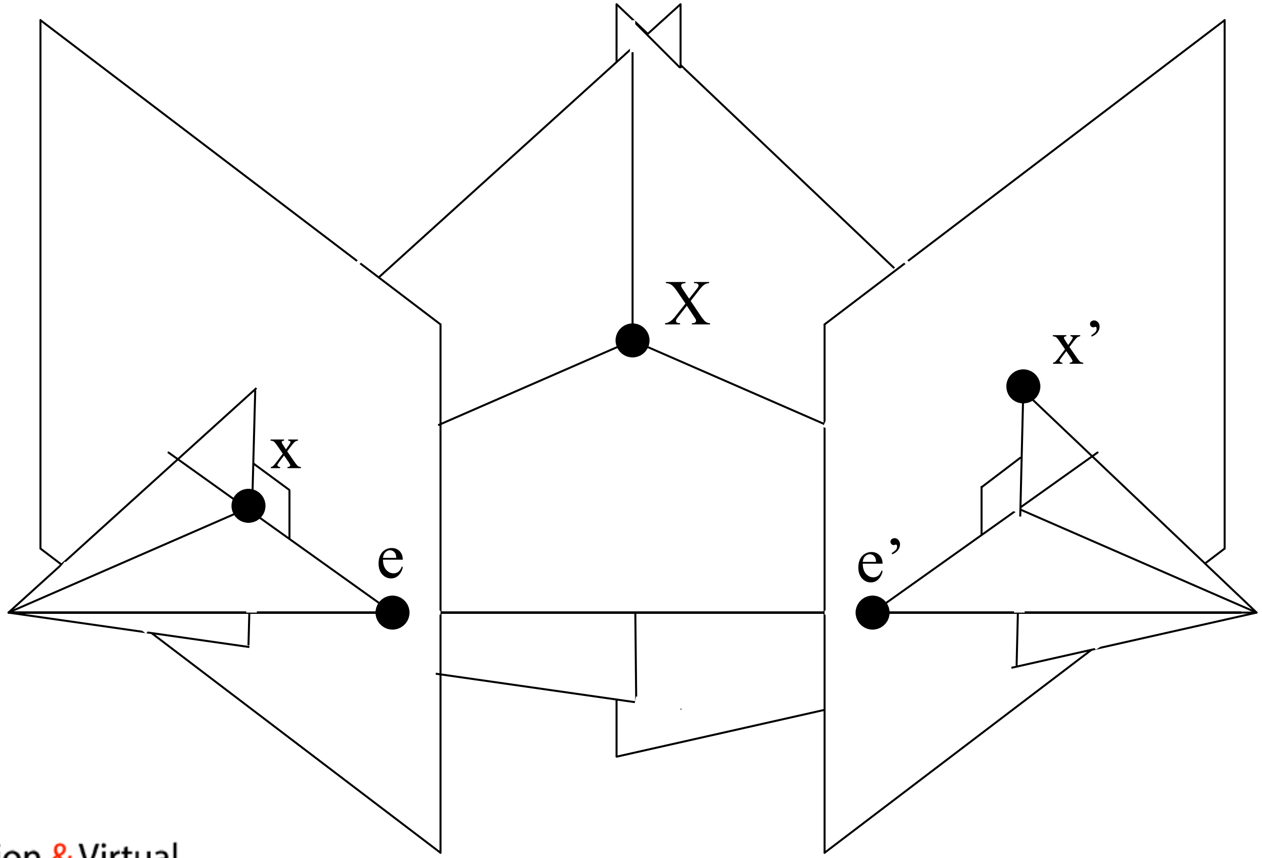


Triangulation



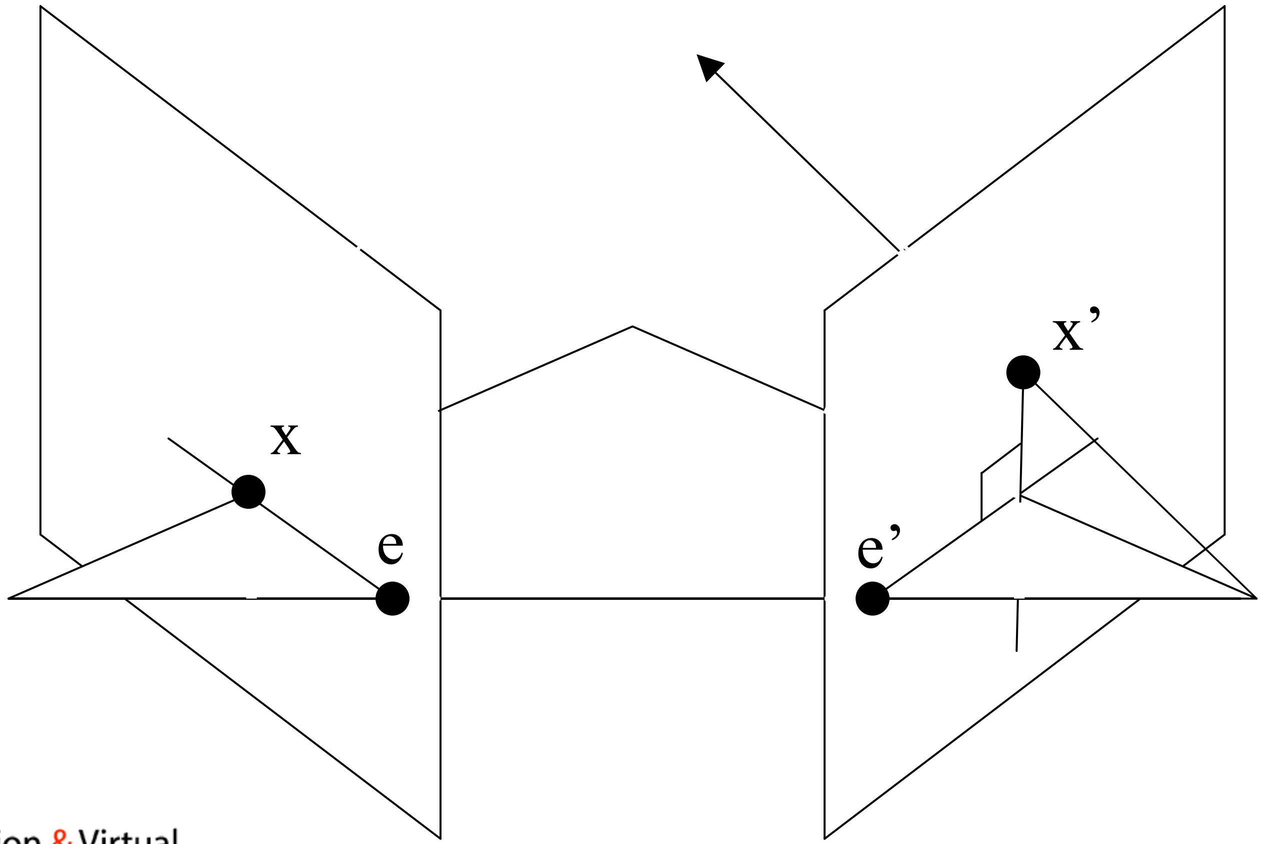
Triangulation

- 2 Stages: Correction & Ideal Triangulation



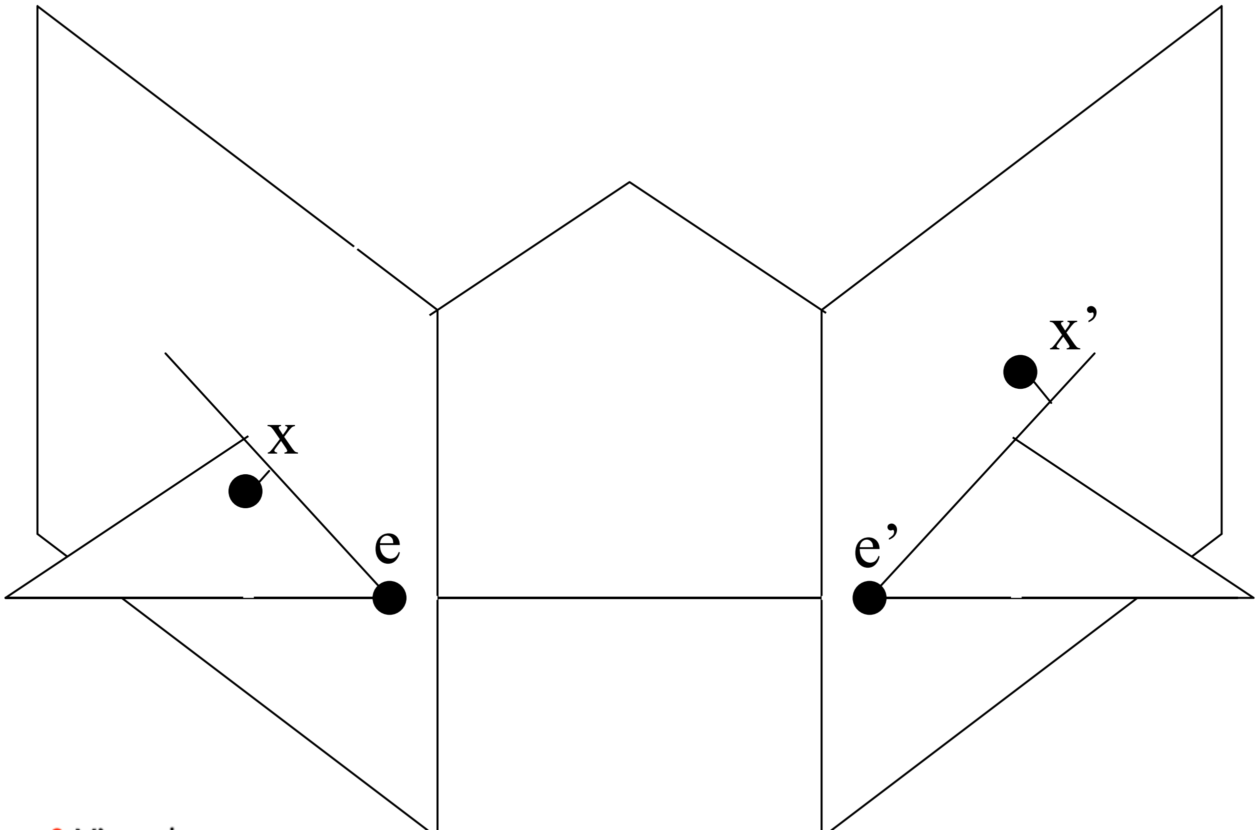
Triangulation

- Rays Intersect <-> Rays Coplanar



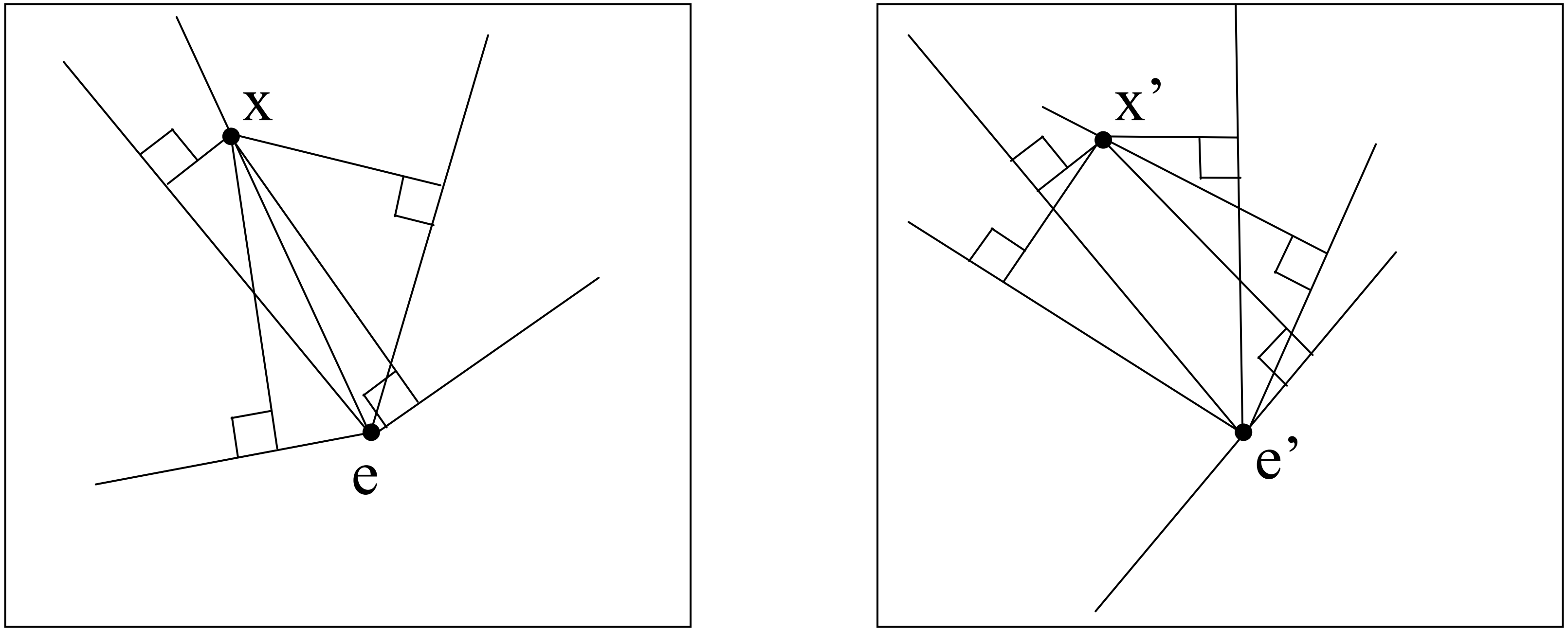
Triangulation

- One parameter family – Balance the error



Triangulation

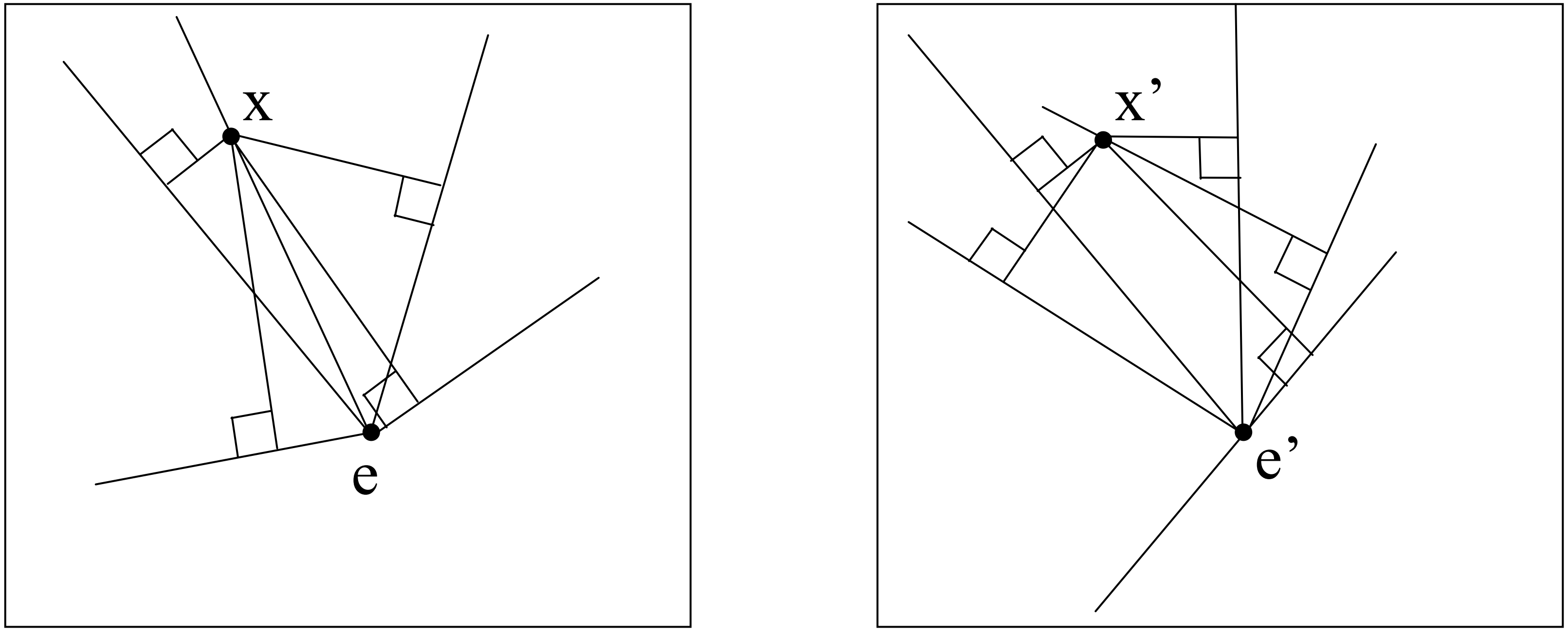
- One parameter family – Balance the error



UK Center for Visualization & Virtual Environments

Triangulation

- One parameter family – Balance the error



Triangulation

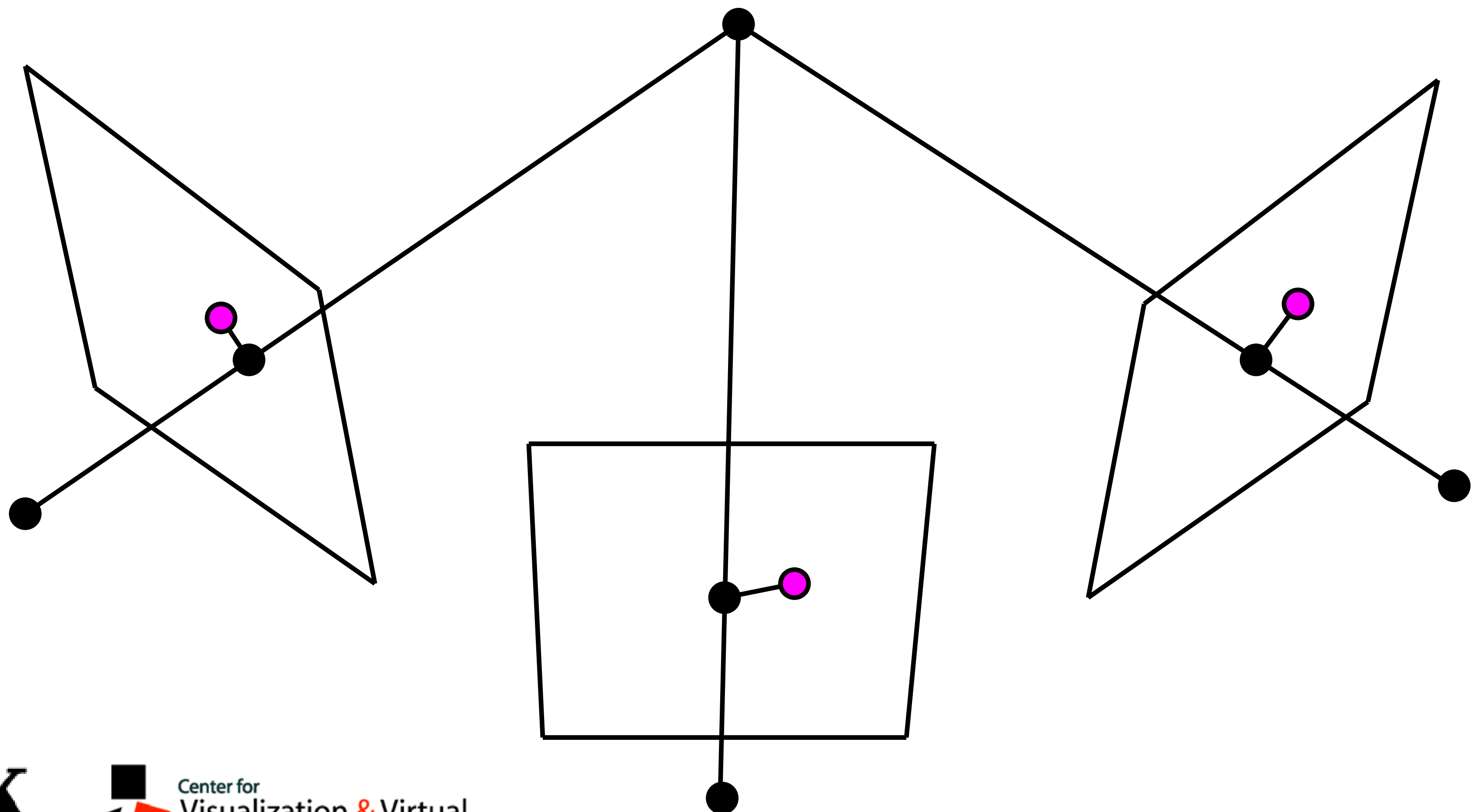
- One parameter family – Balance the error
- L2-Norm -> Sextic (Hartley & Sturm)
- Max-Norm -> Quartic (Closed form, Nistér)
- Directional Error -> Quadratic (Oliensis)

UK Center for Visualization & Virtual Environments

Optimal 3 View Triangulation

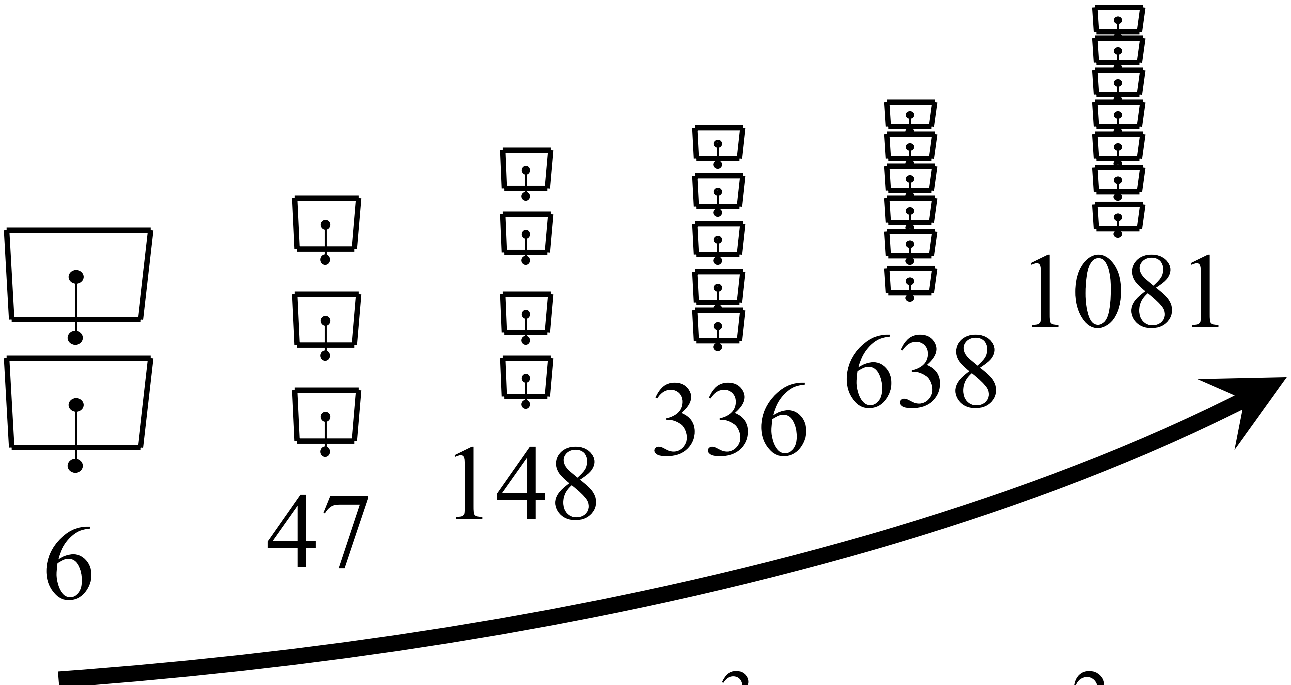
work with Henrik Stewenius and Fred Schaffalitzky

47 Stationary Points



Center for Visualization & Virtual Environments

Nr of Stationary Points for Triangulations in N Views



$$4.5N^3 + 3N^2 + 0.5N - 2$$

Sampson Approximation

Squared Mahalanobis Distance

$$M(x) = x^{\mathrm{T}} C_{xx}^{-1} x$$

Covariance Propagation

$$C_{JxJx} = J C_{xx} J^{\mathrm{T}}$$

Combine

$$M(f) \approx f^{\mathrm{T}} \left(J C_{xx} J^{\mathrm{T}} \right)^{-1} f$$

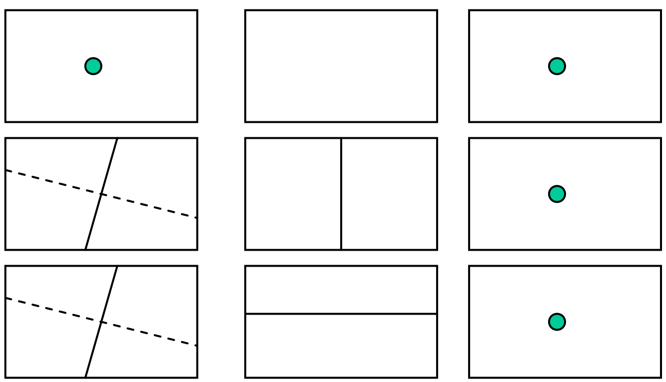
Where C_{xx} is the covariance matrix of detected image features and f and J are the incidence function and its Jacobian

Sampson Approximation

For two views this leads to

$$M\left(F,x,x'\right)=\frac{\left(x'^{\mathrm{T}}Fx\right)^2}{\left(Fx\right)_1^2+\left(Fx\right)_2^2+\left(x'^{\mathrm{T}}F\right)_1^2+\left(x'^{\mathrm{T}}F\right)_2^2}$$

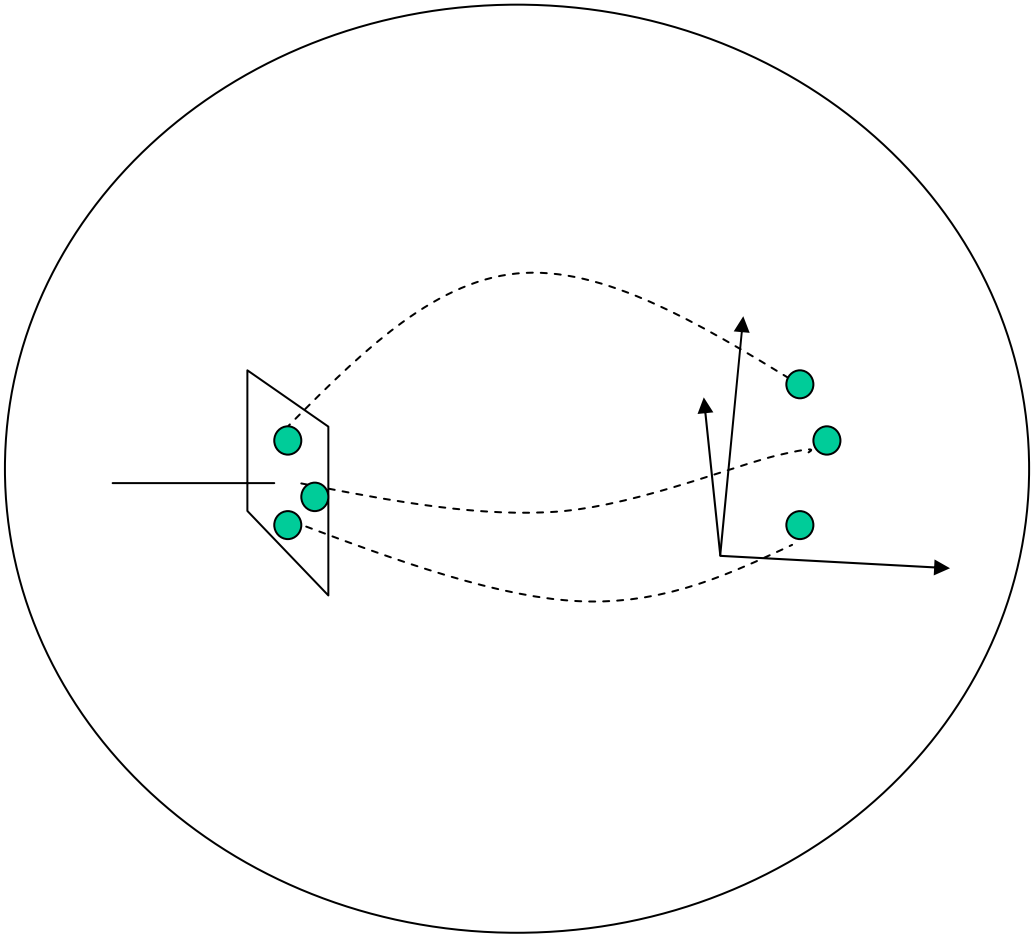
For three views, an approximation of the distance to trifocal incidence can be found by tensor contractions and Cramer's rule in <1 microsecond



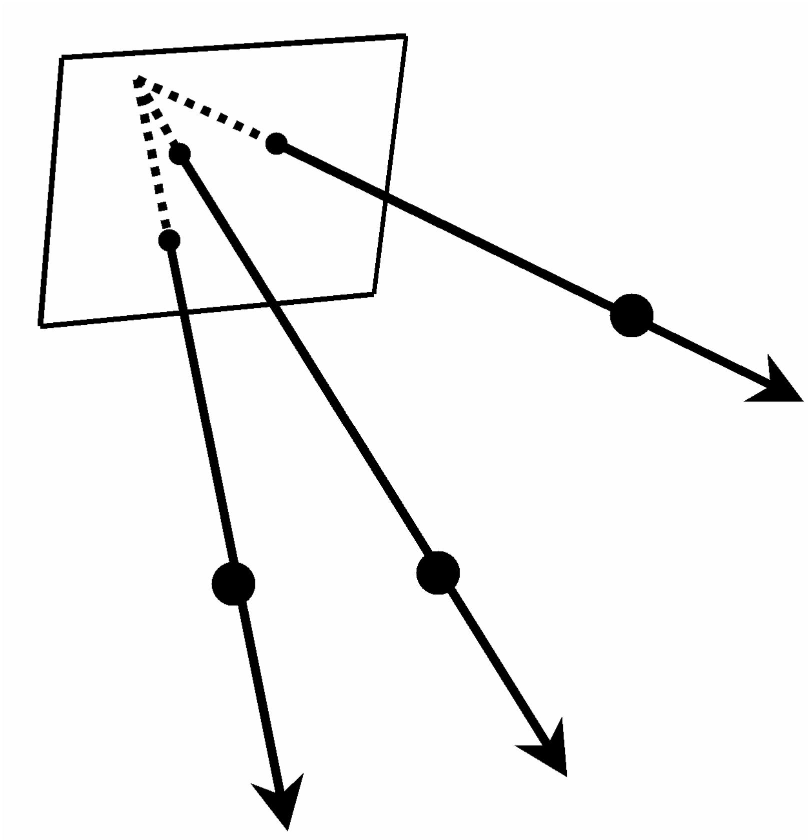
Assuming Cauchy distribution

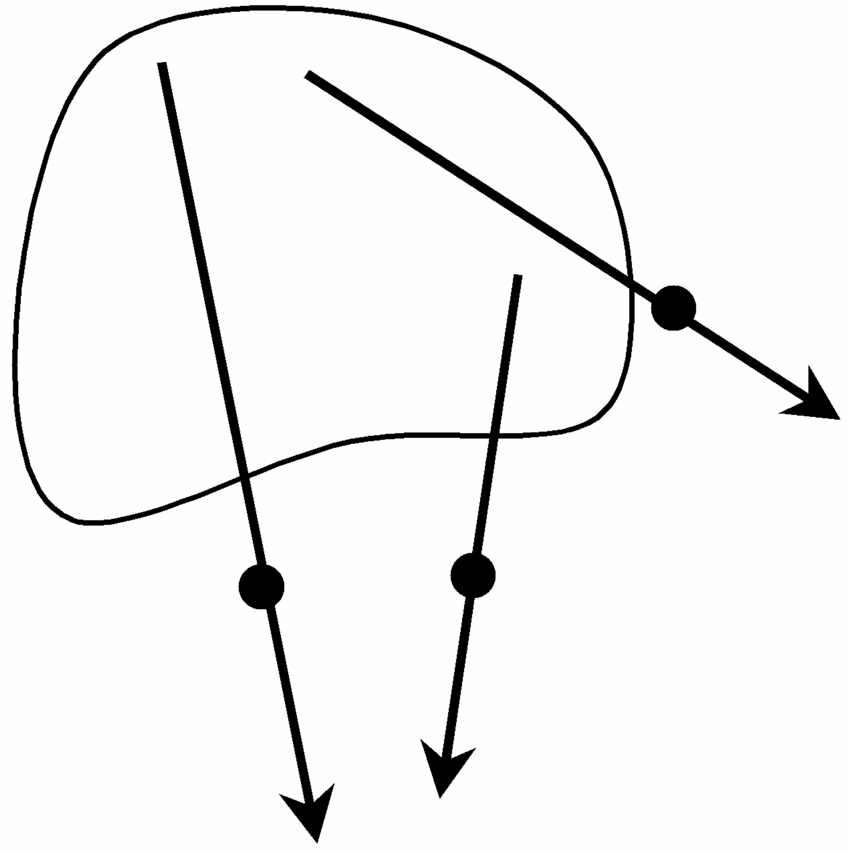
$$D=\ln\left(1+M\right)$$

2D-3D Pose

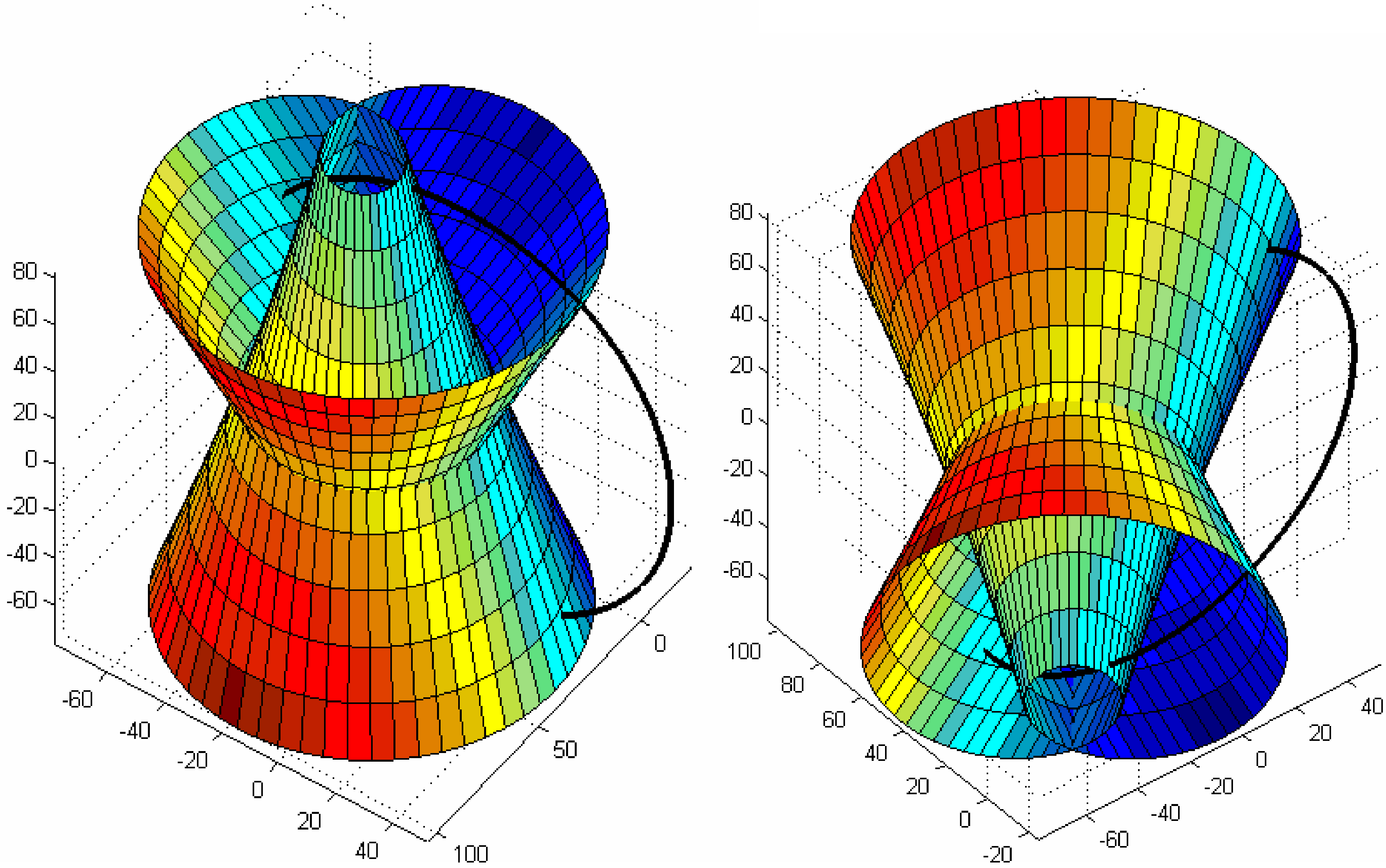


The 3-Point Problem

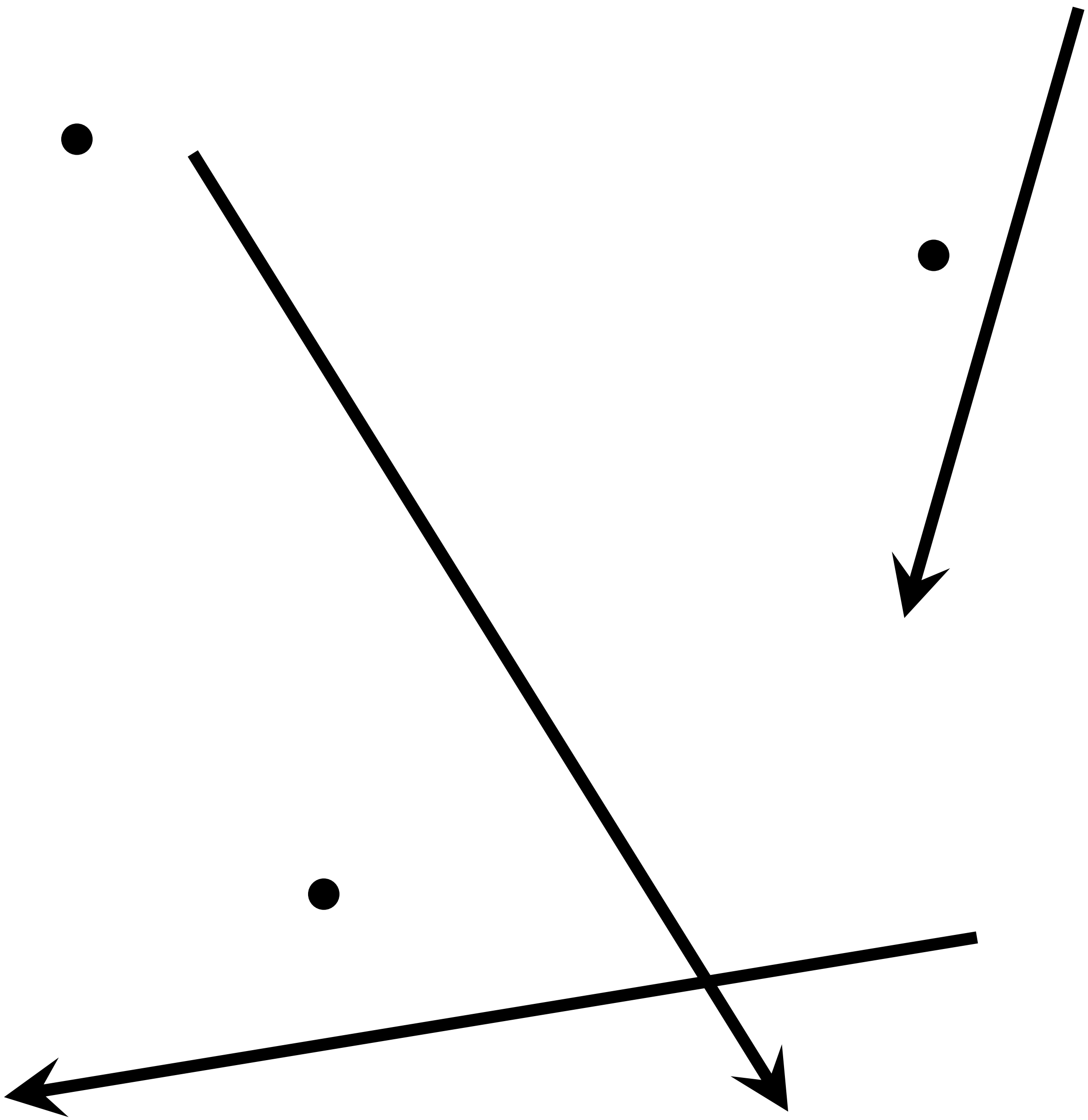




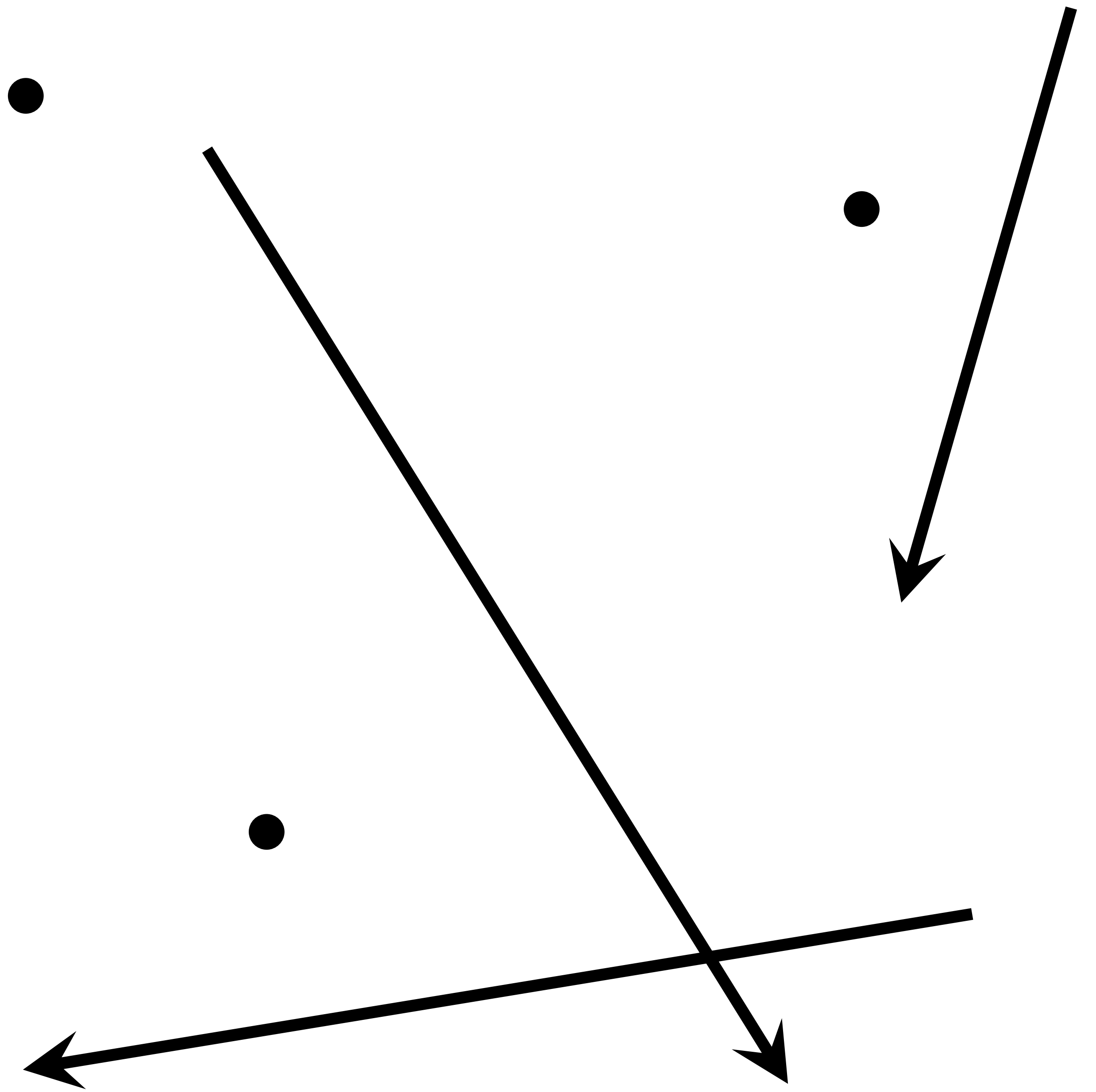
The 3-Point Problem

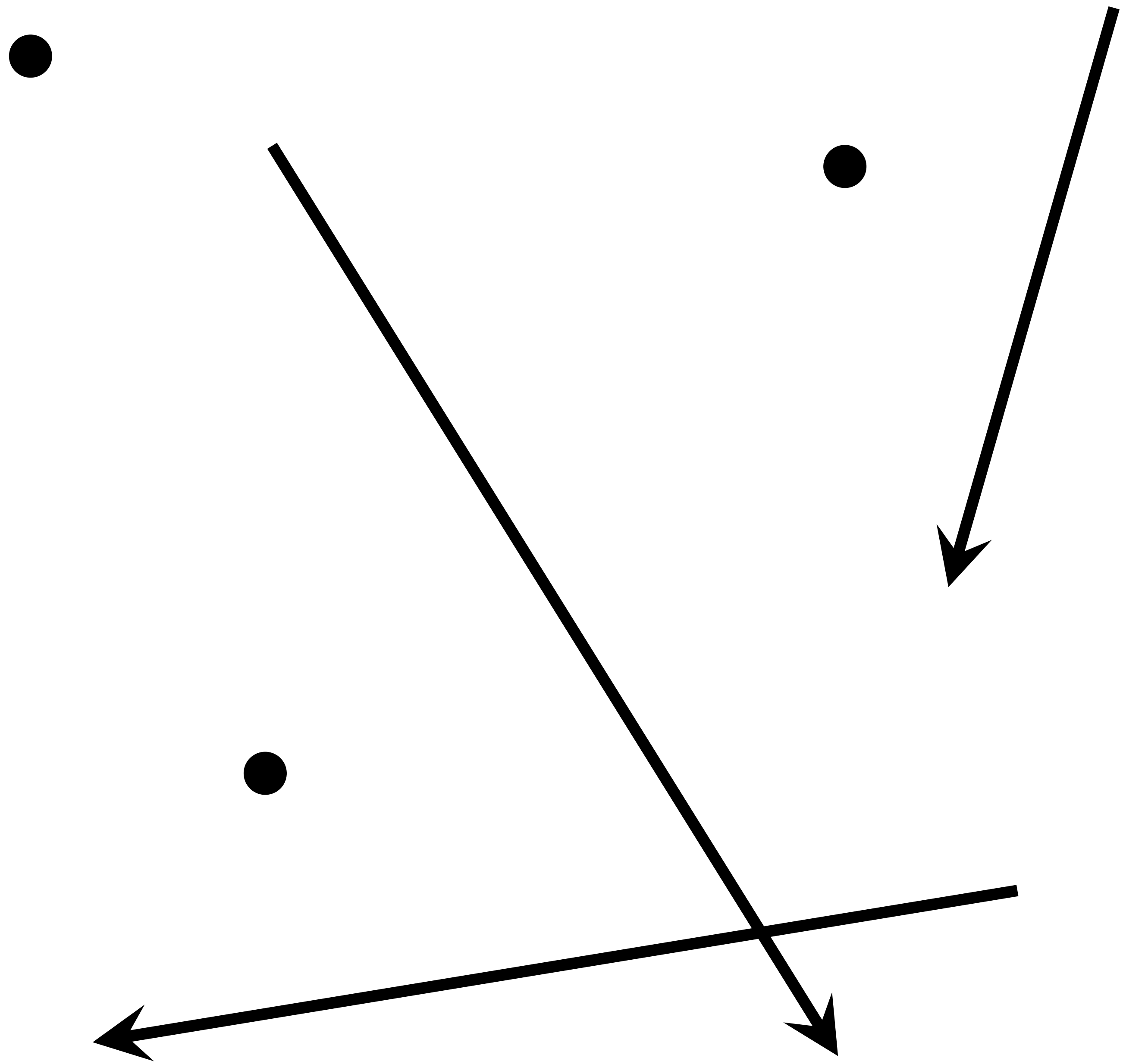


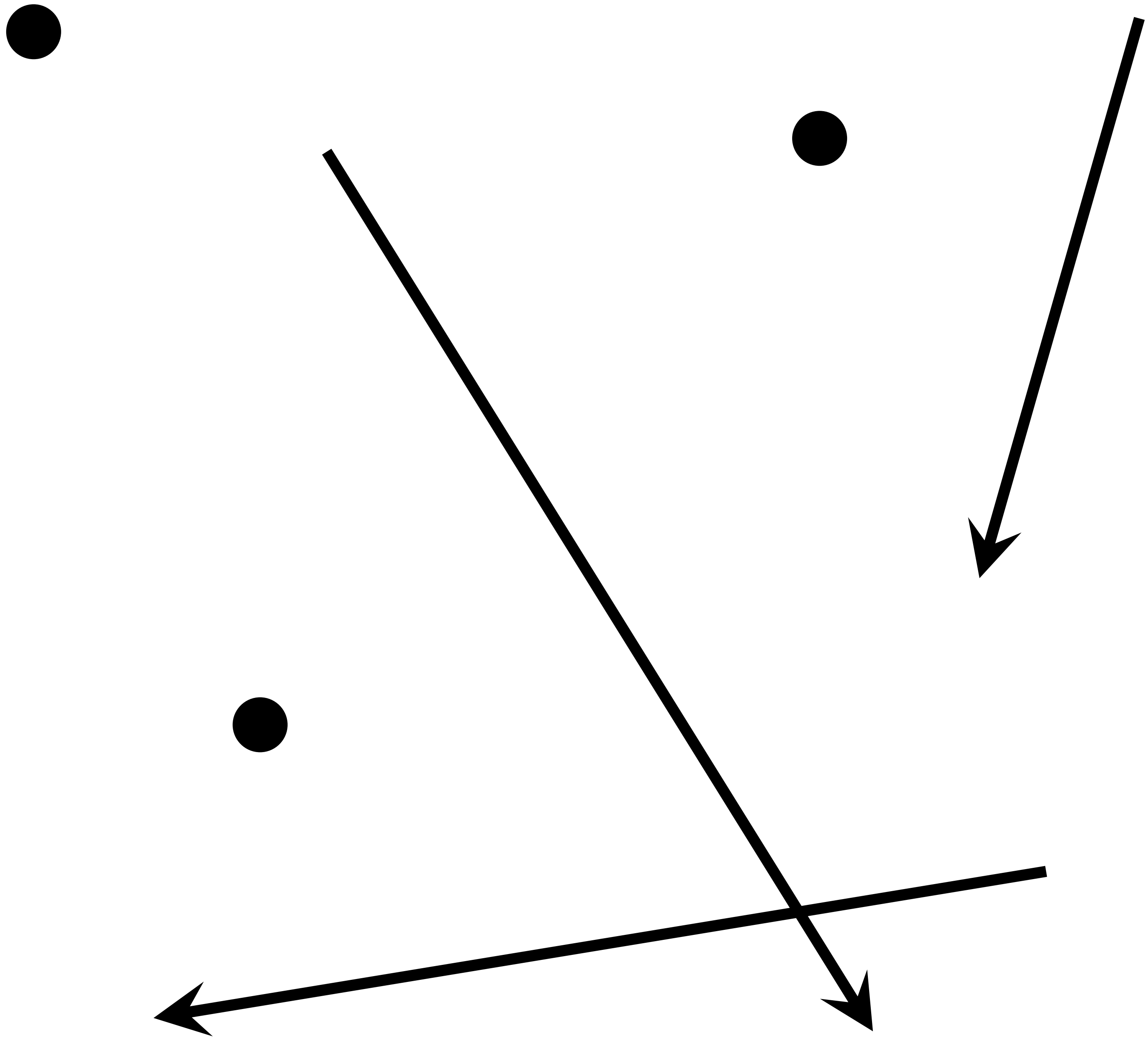
Center for Visualization & Virtual Environments

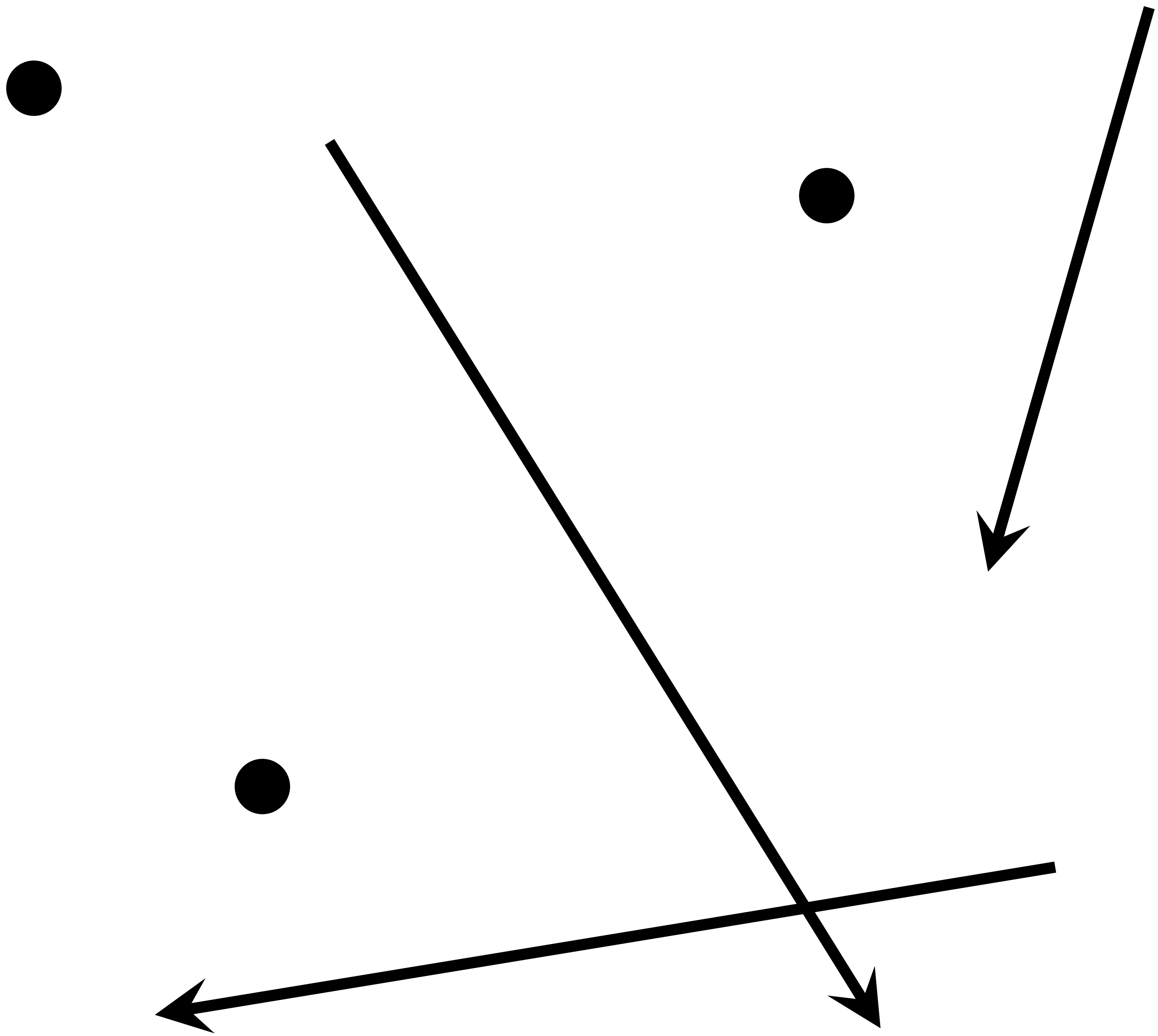


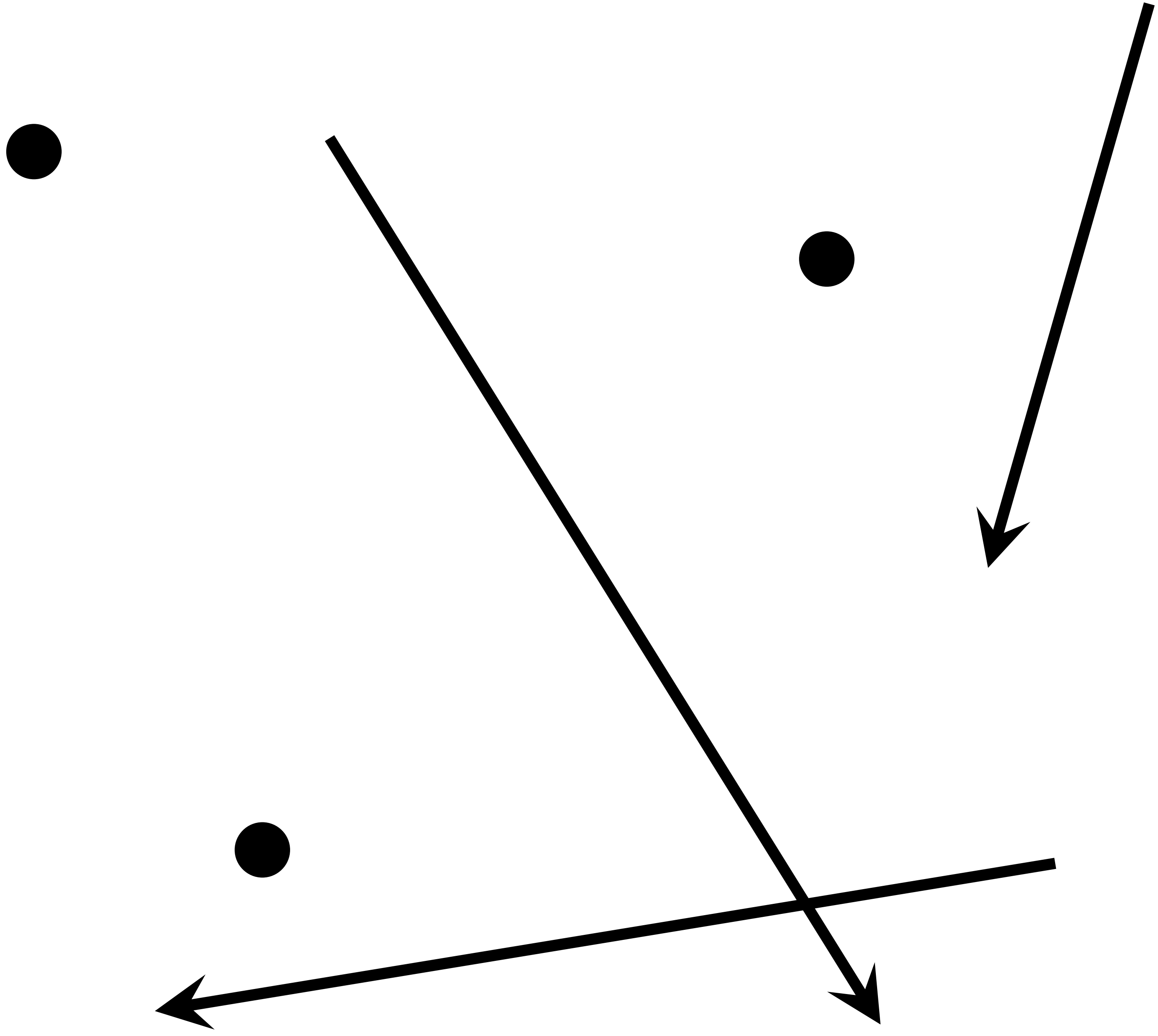
Center for Visualization & Virtual Environments

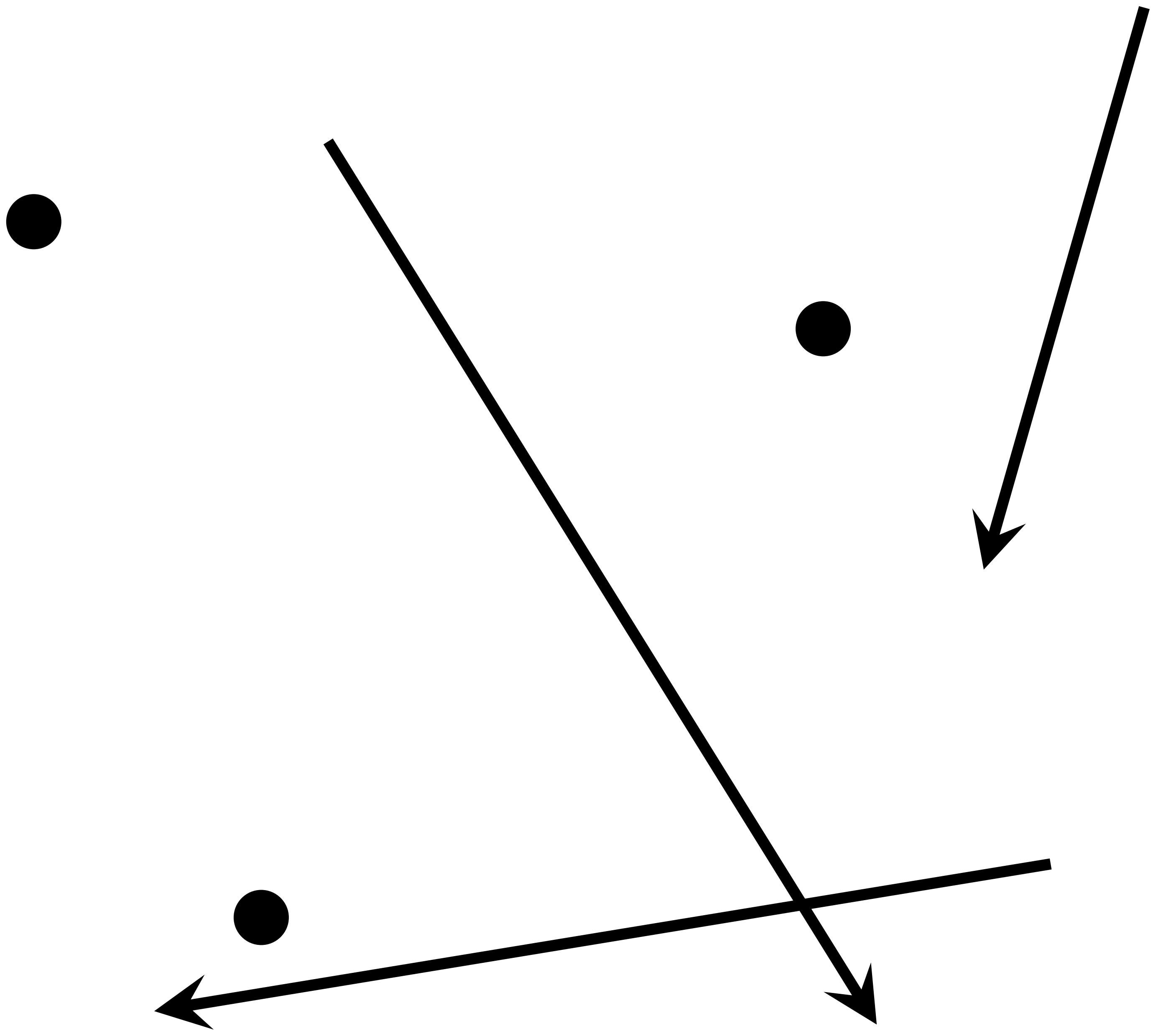




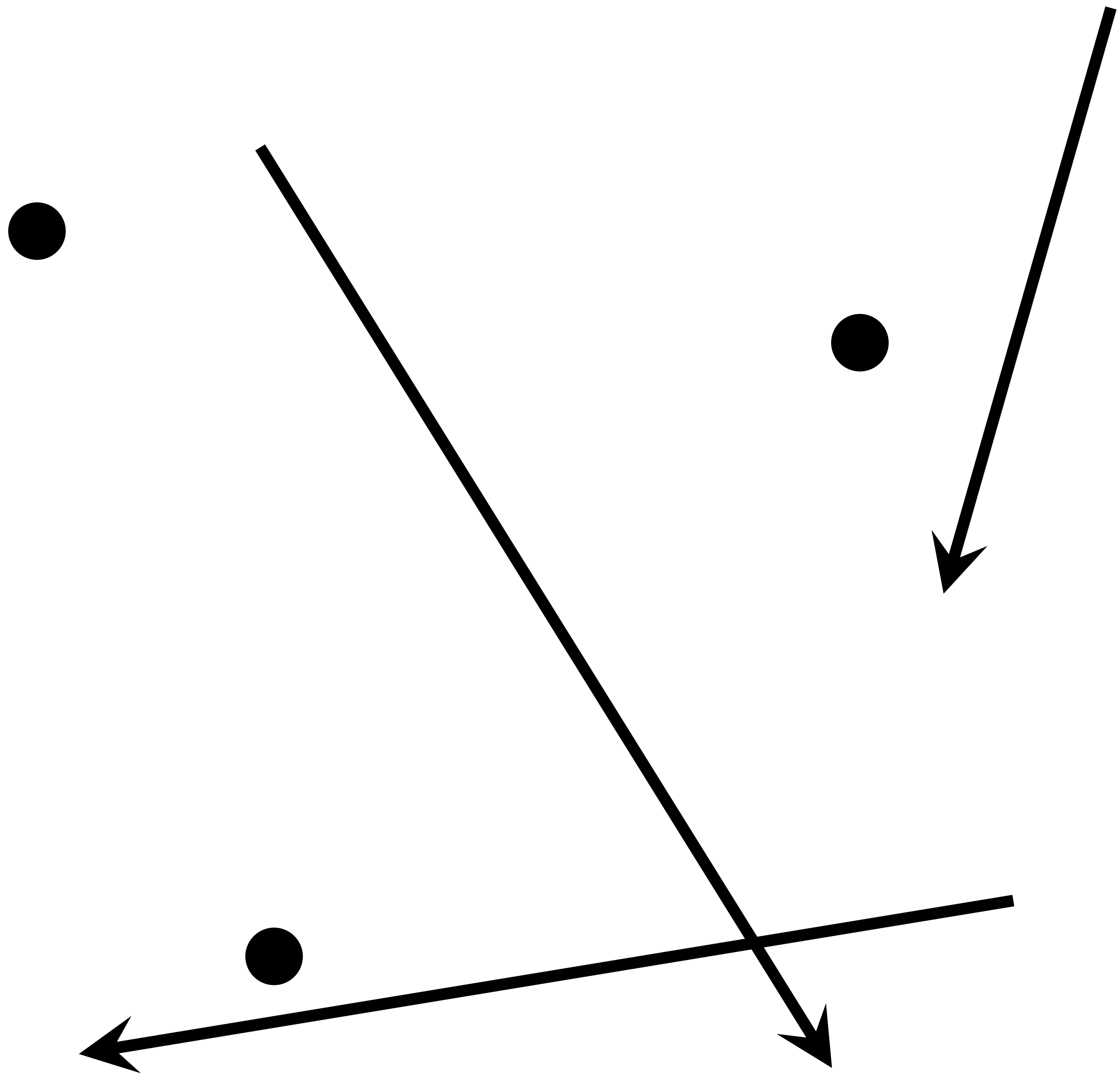


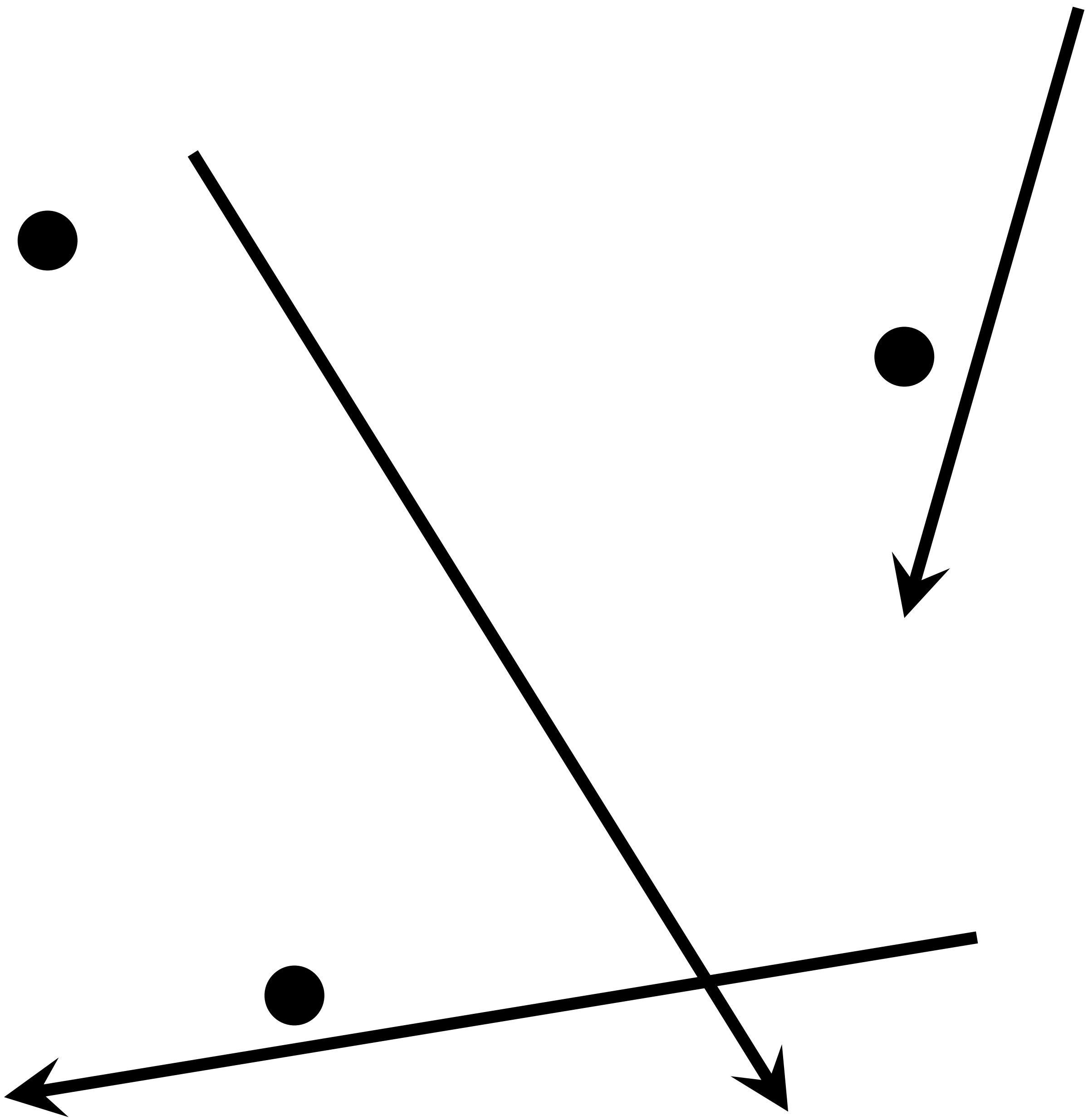


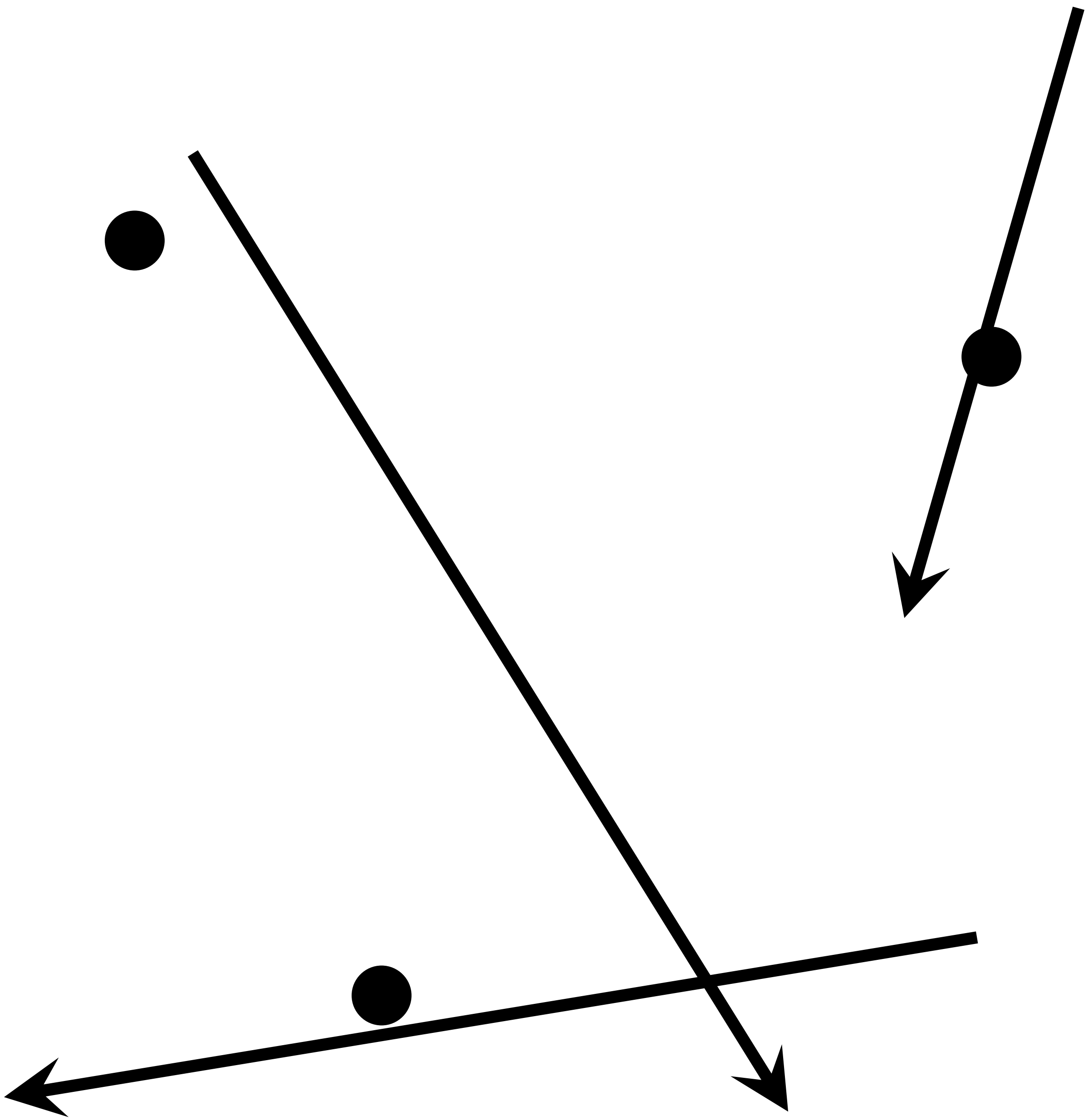


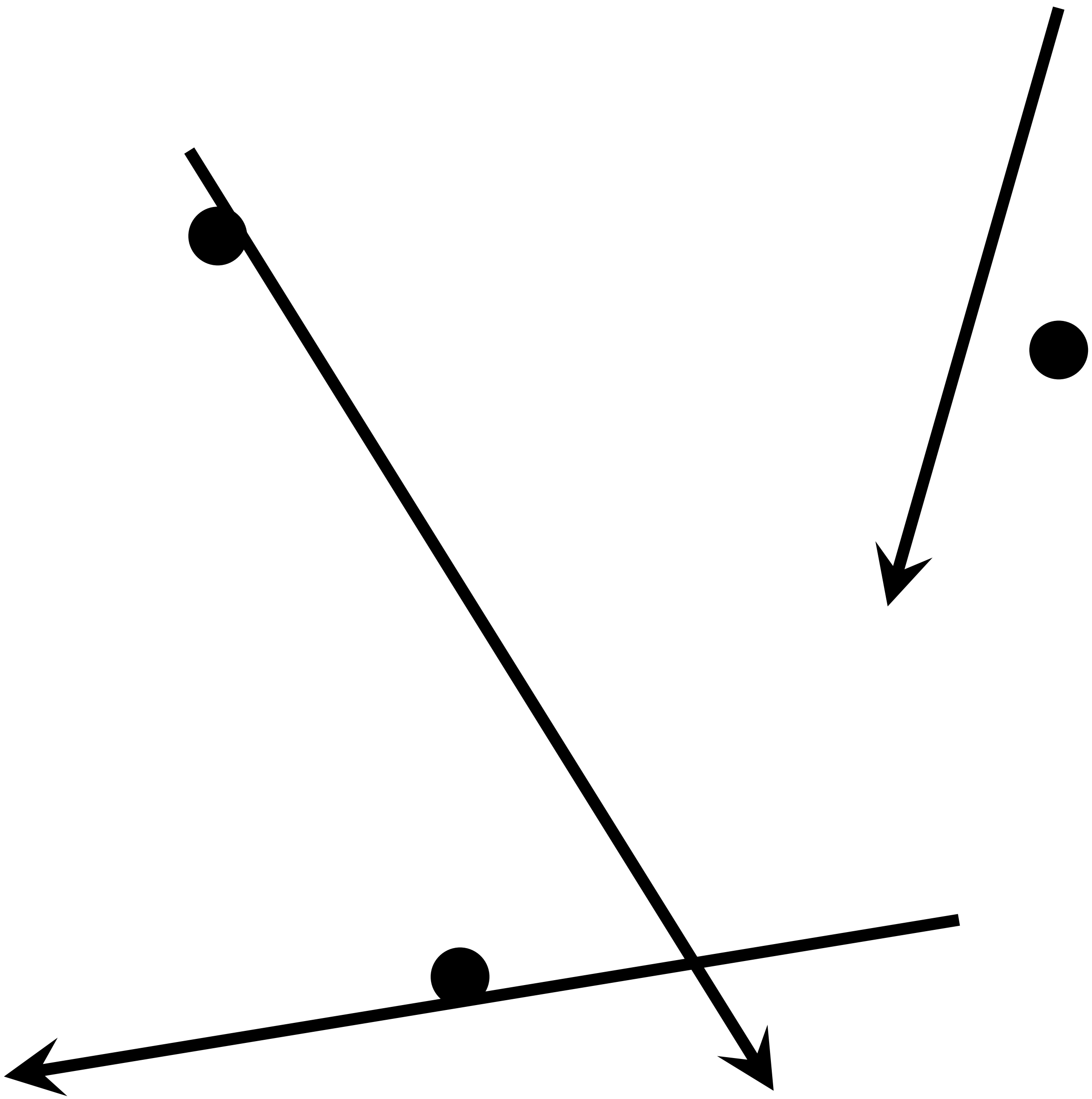


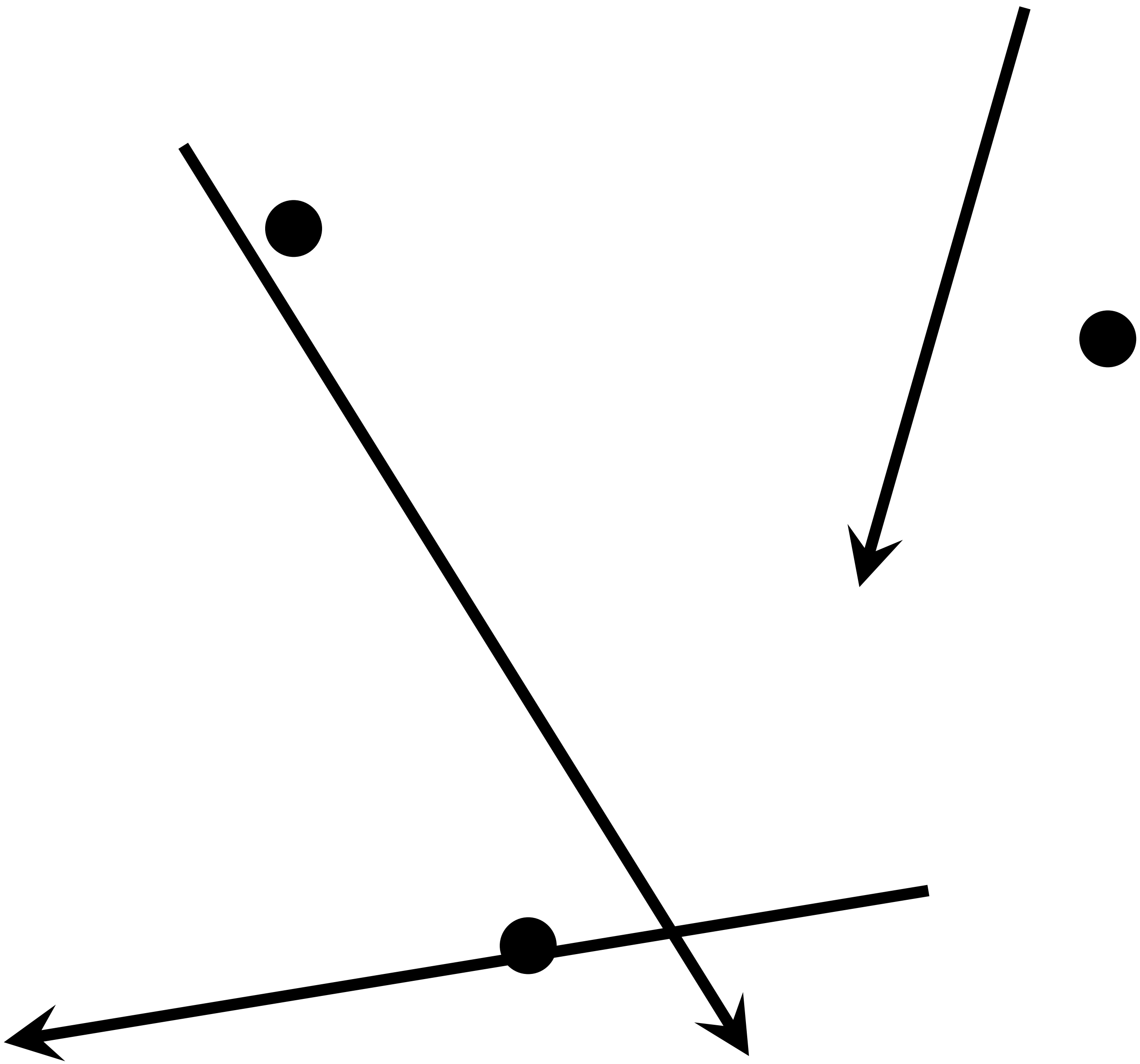
UK Center for Visualization & Virtual Environments

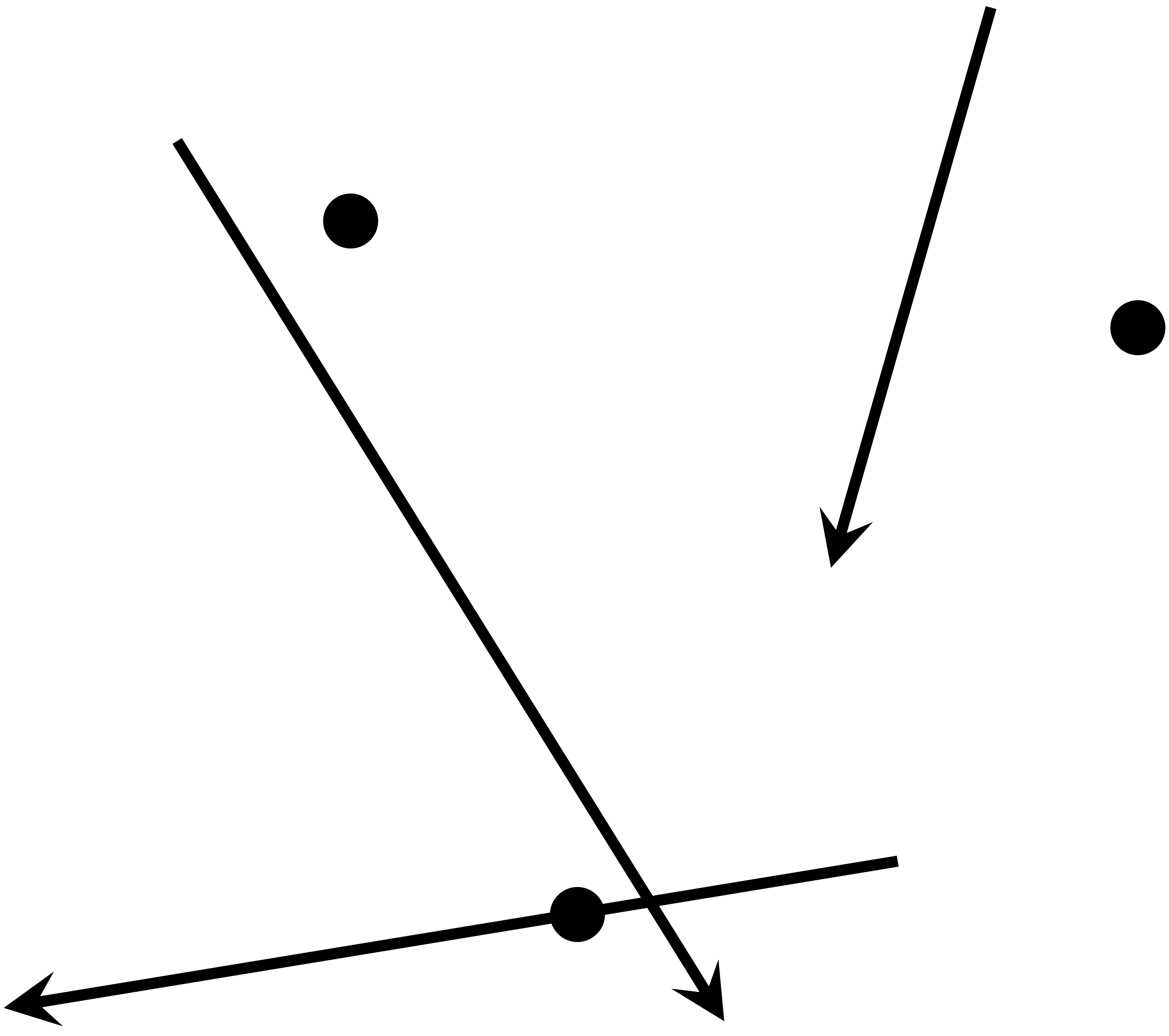


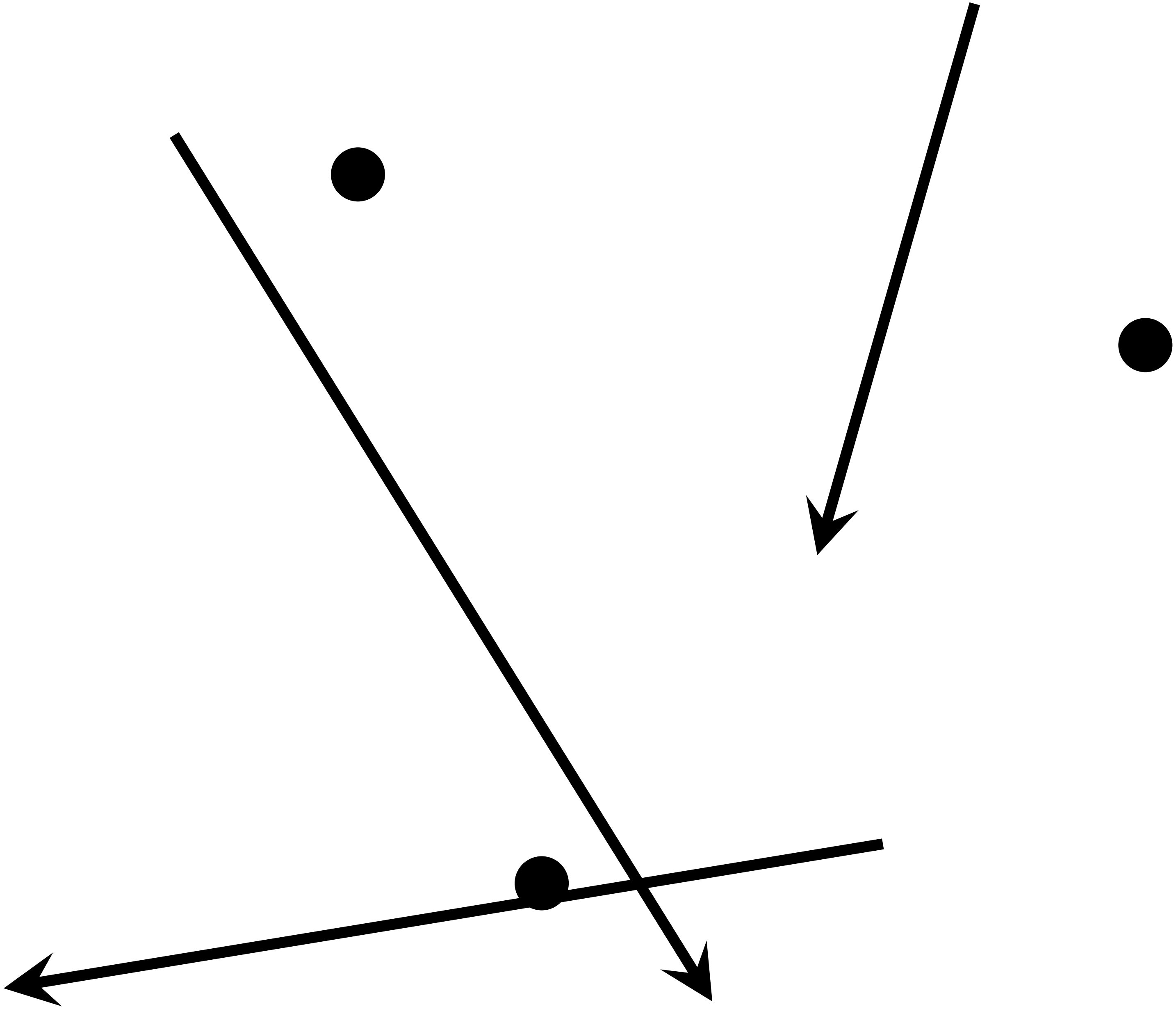


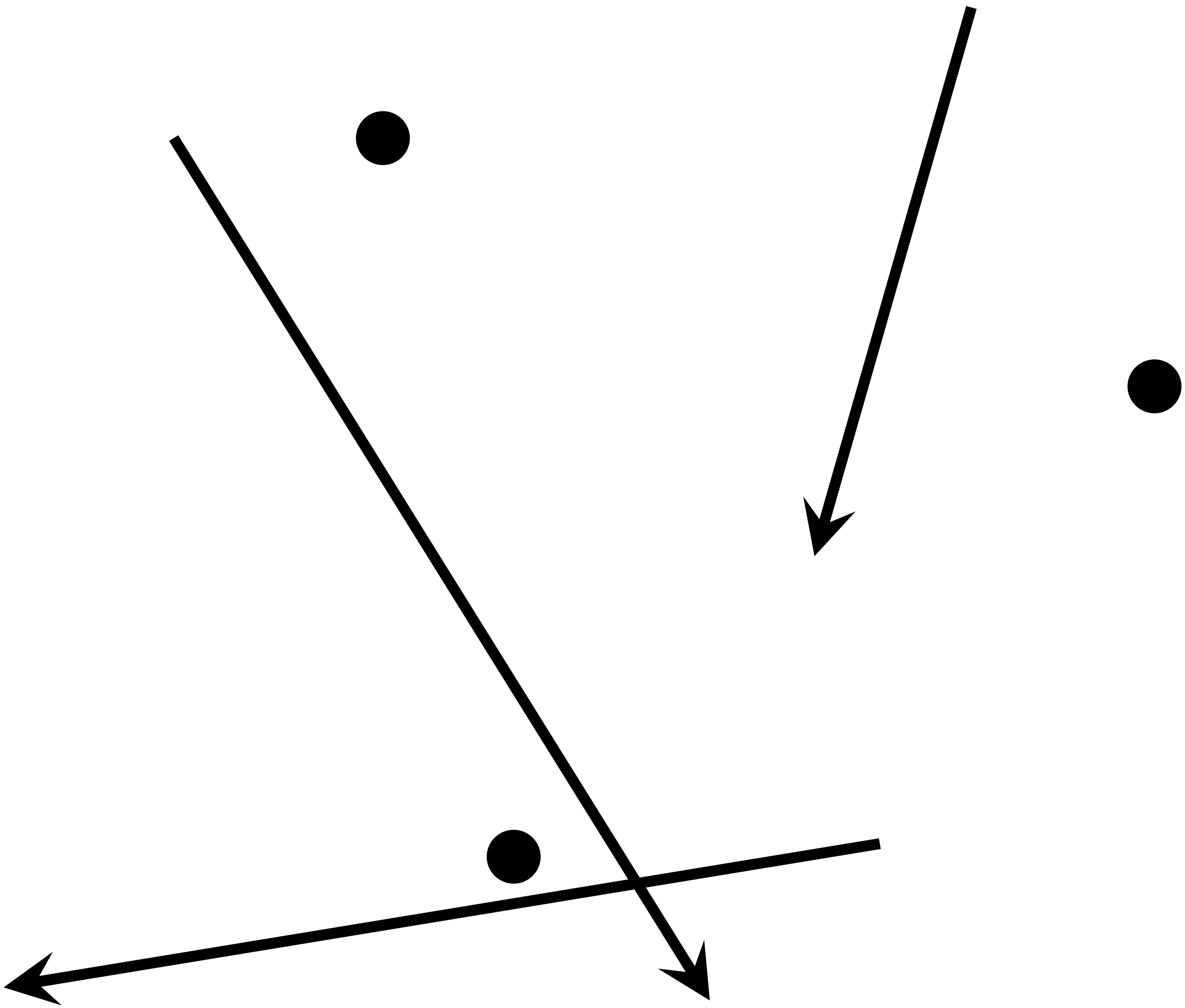


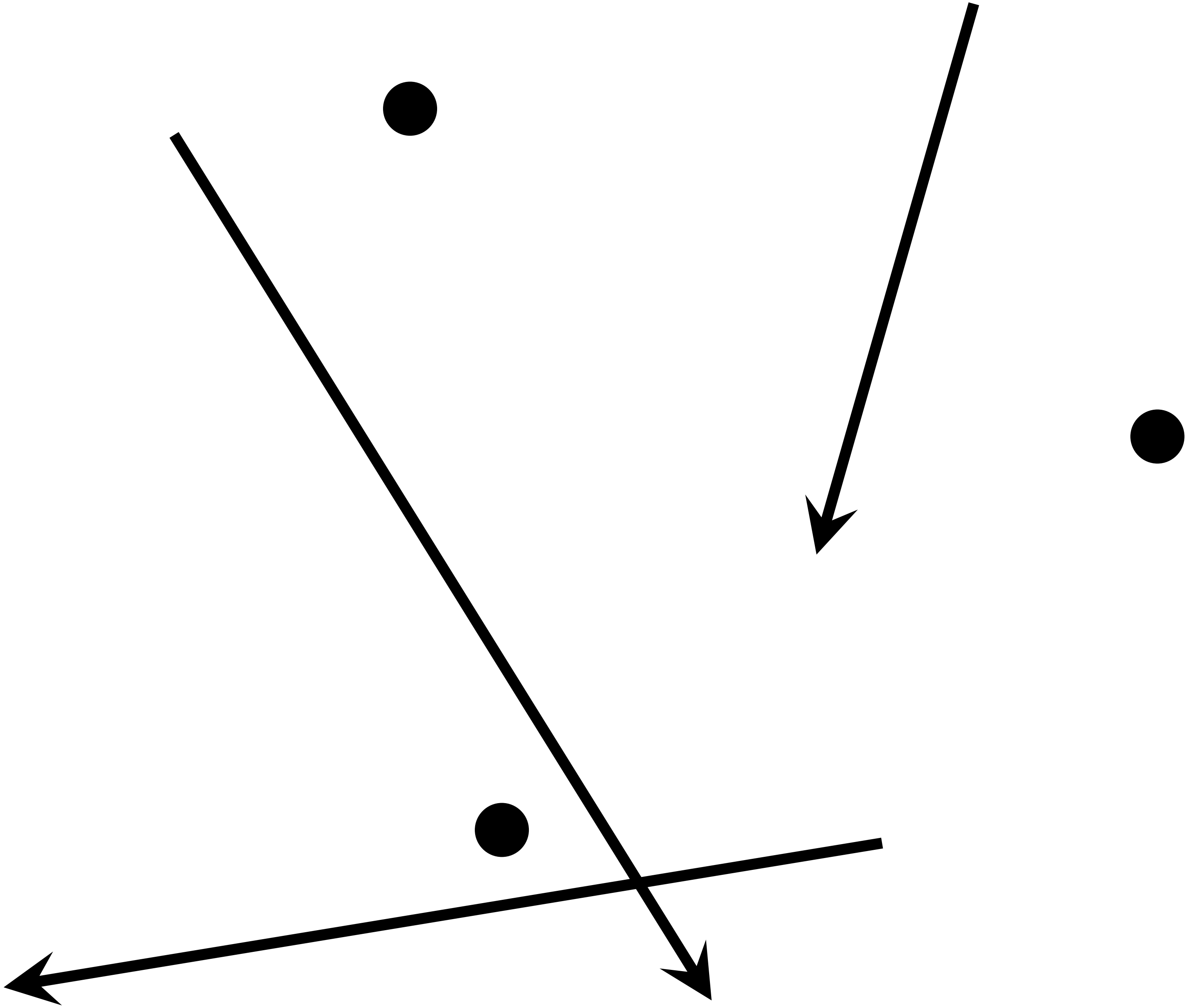




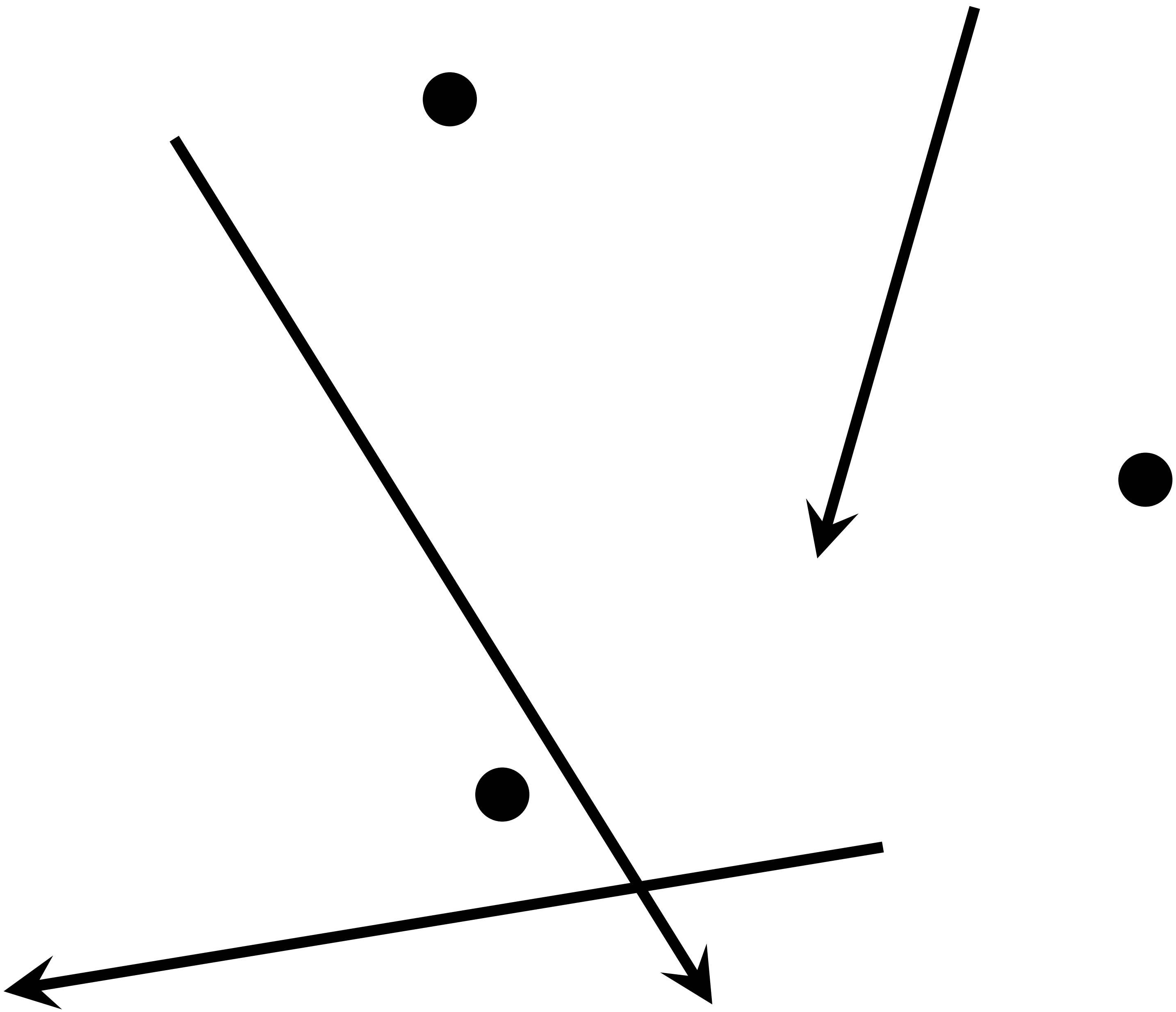


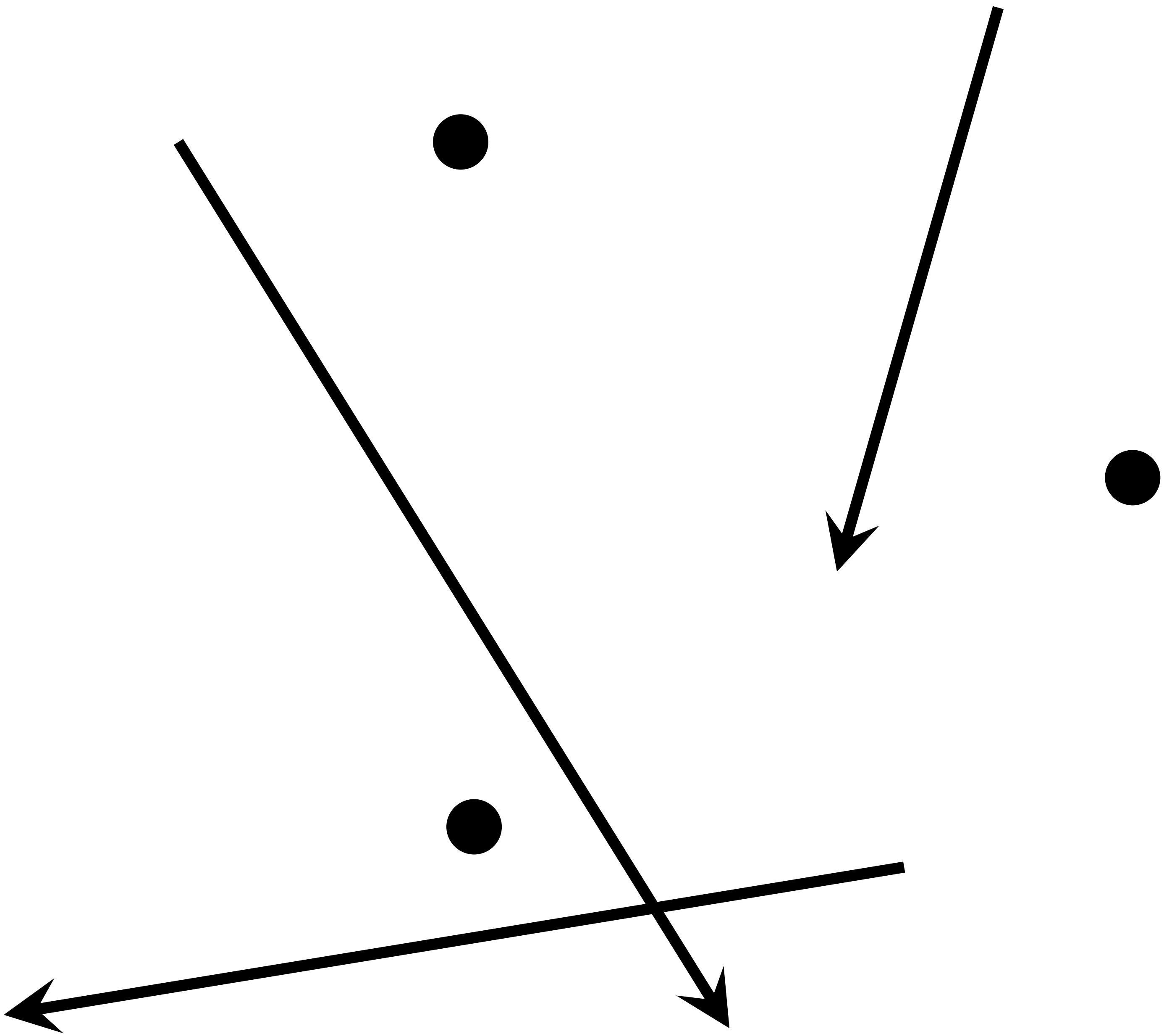




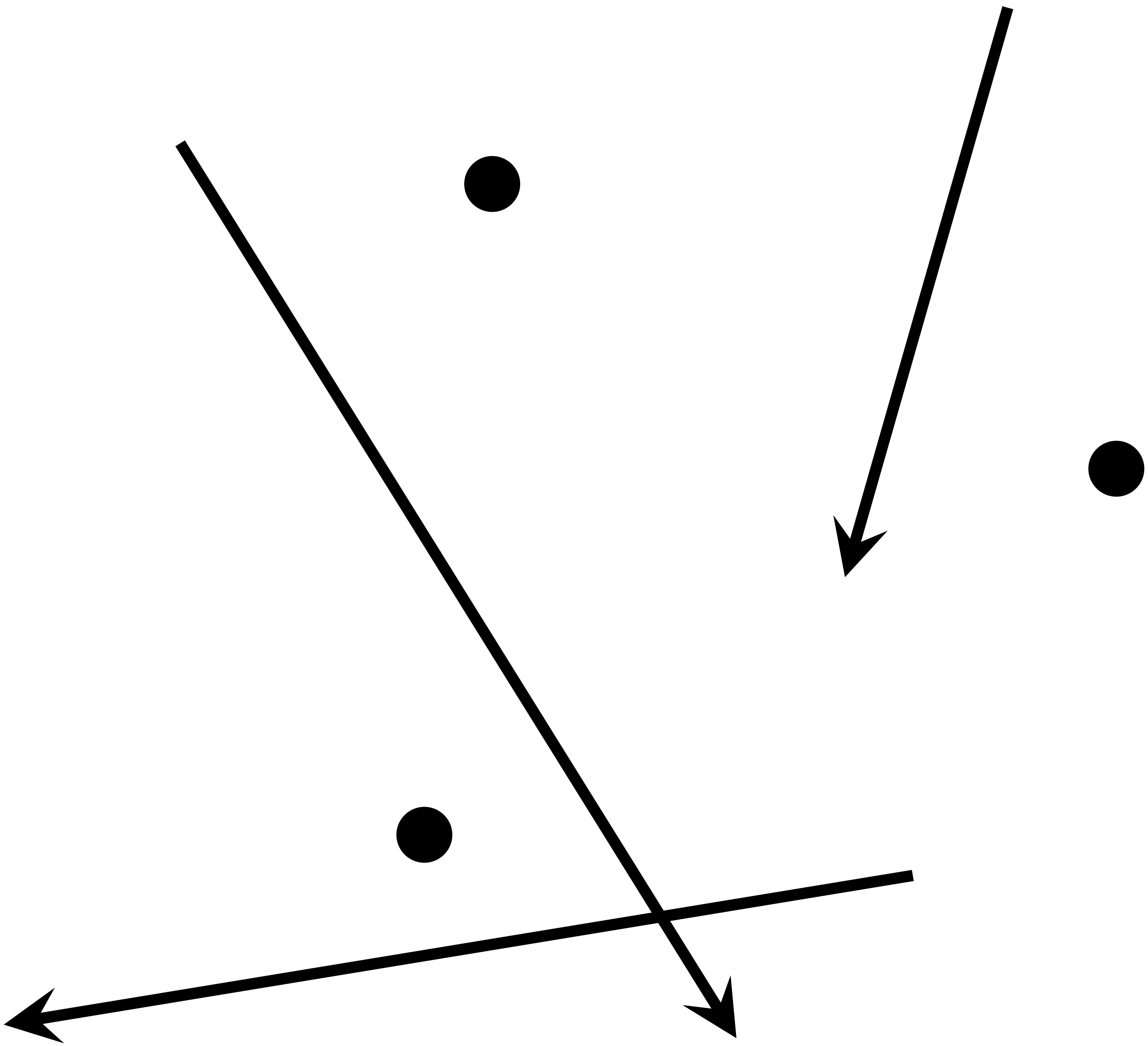


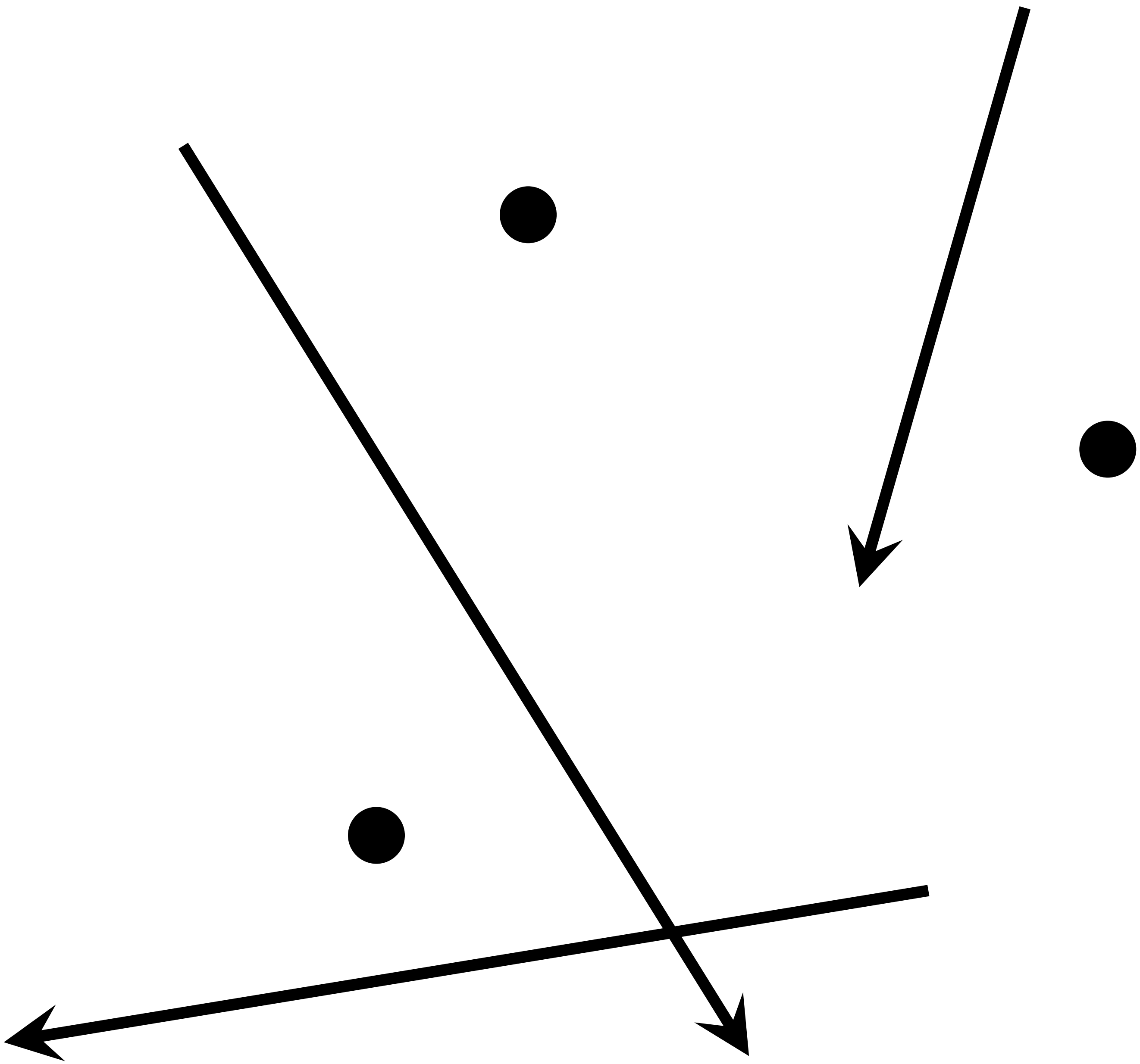
UK Center for Visualization & Virtual Environments

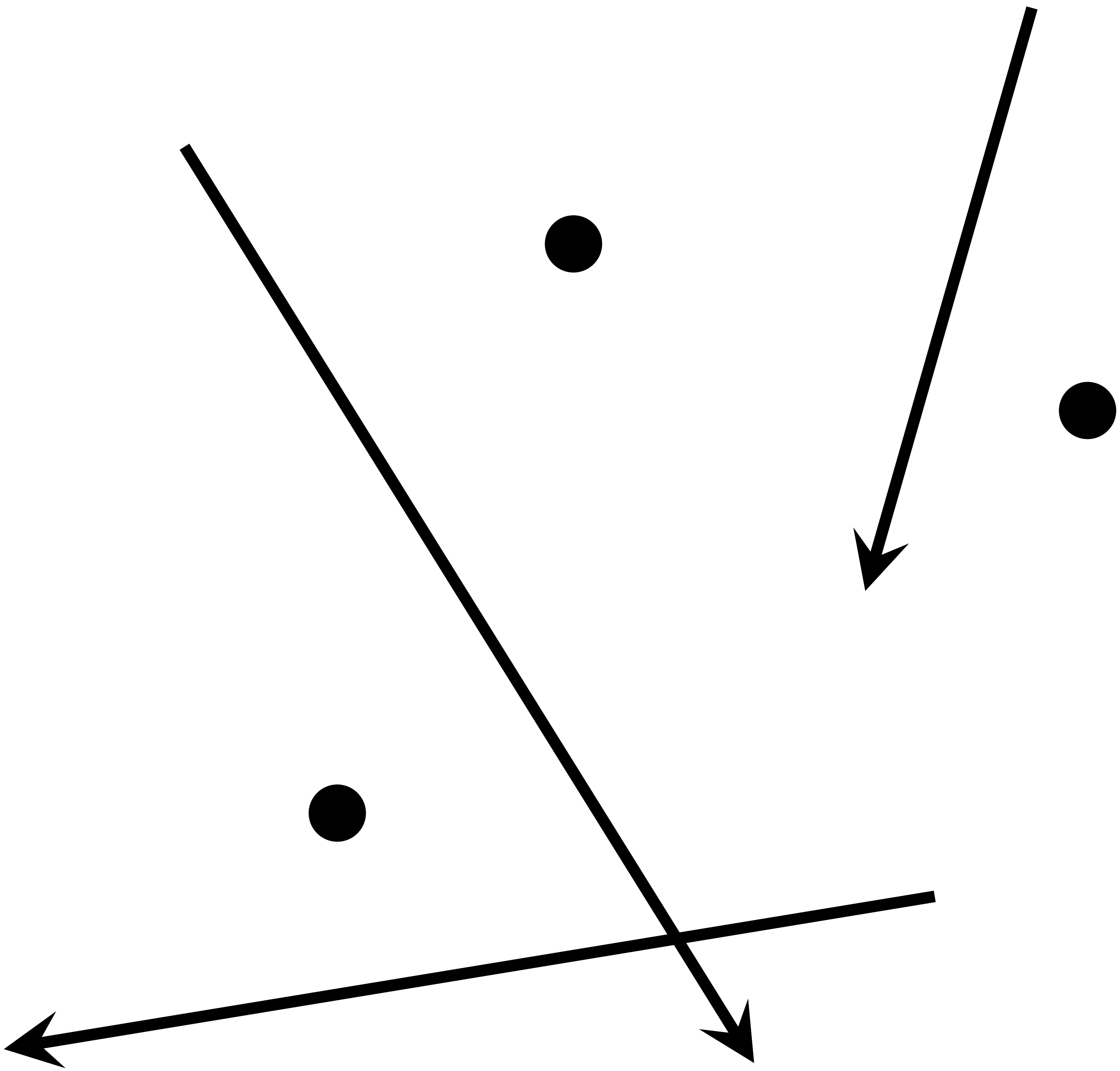


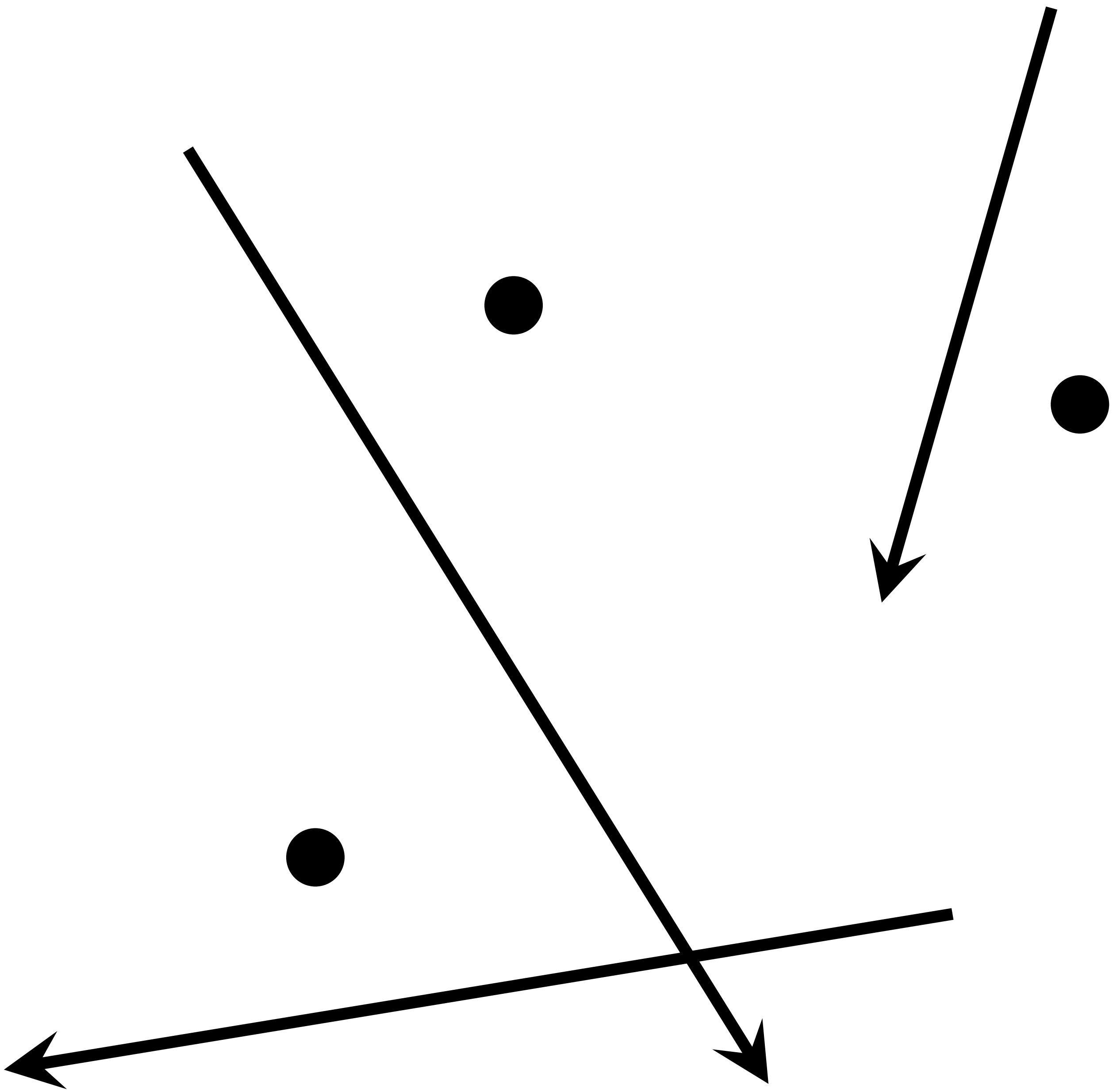


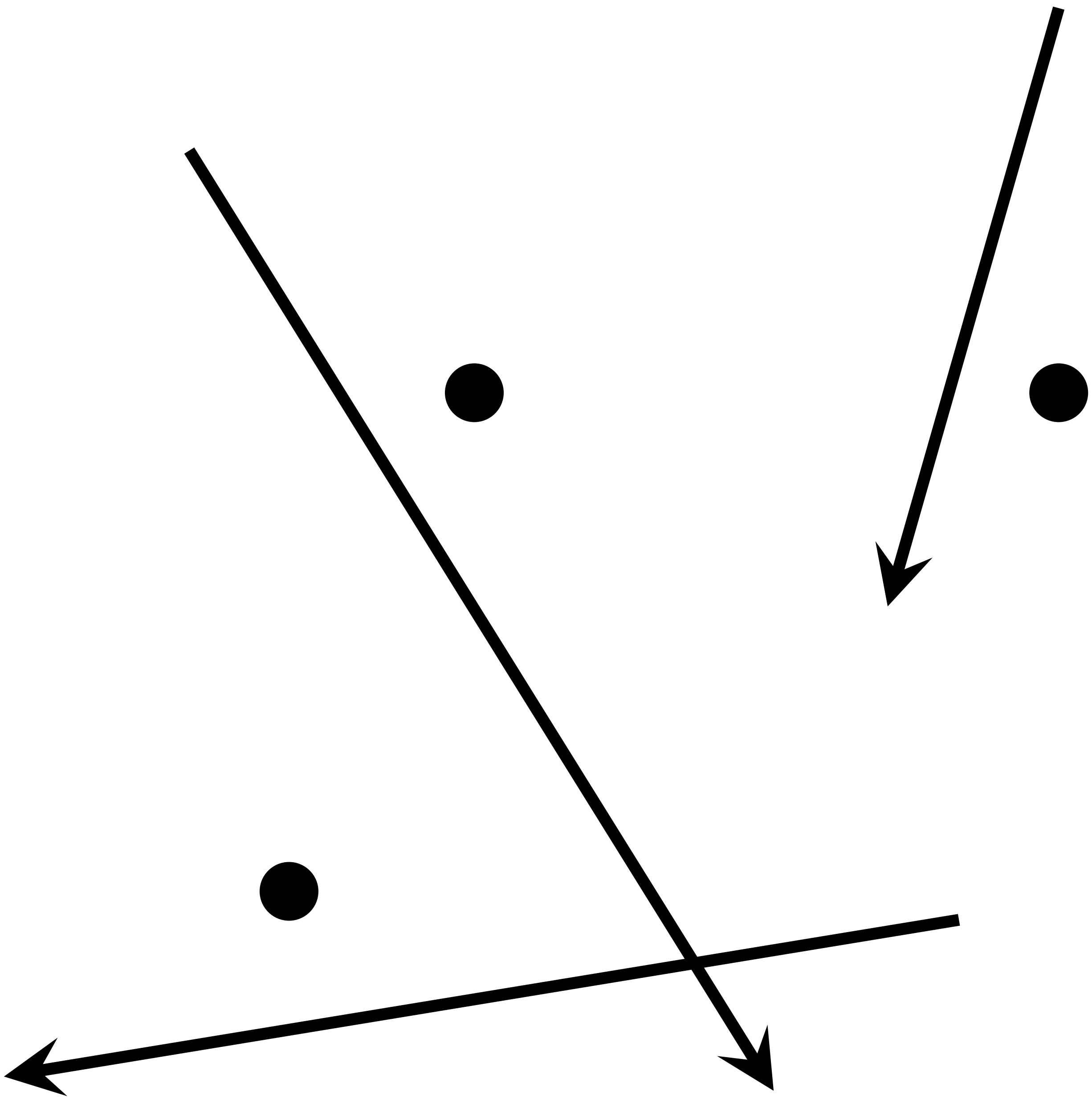
Center for Visualization & Virtual Environments

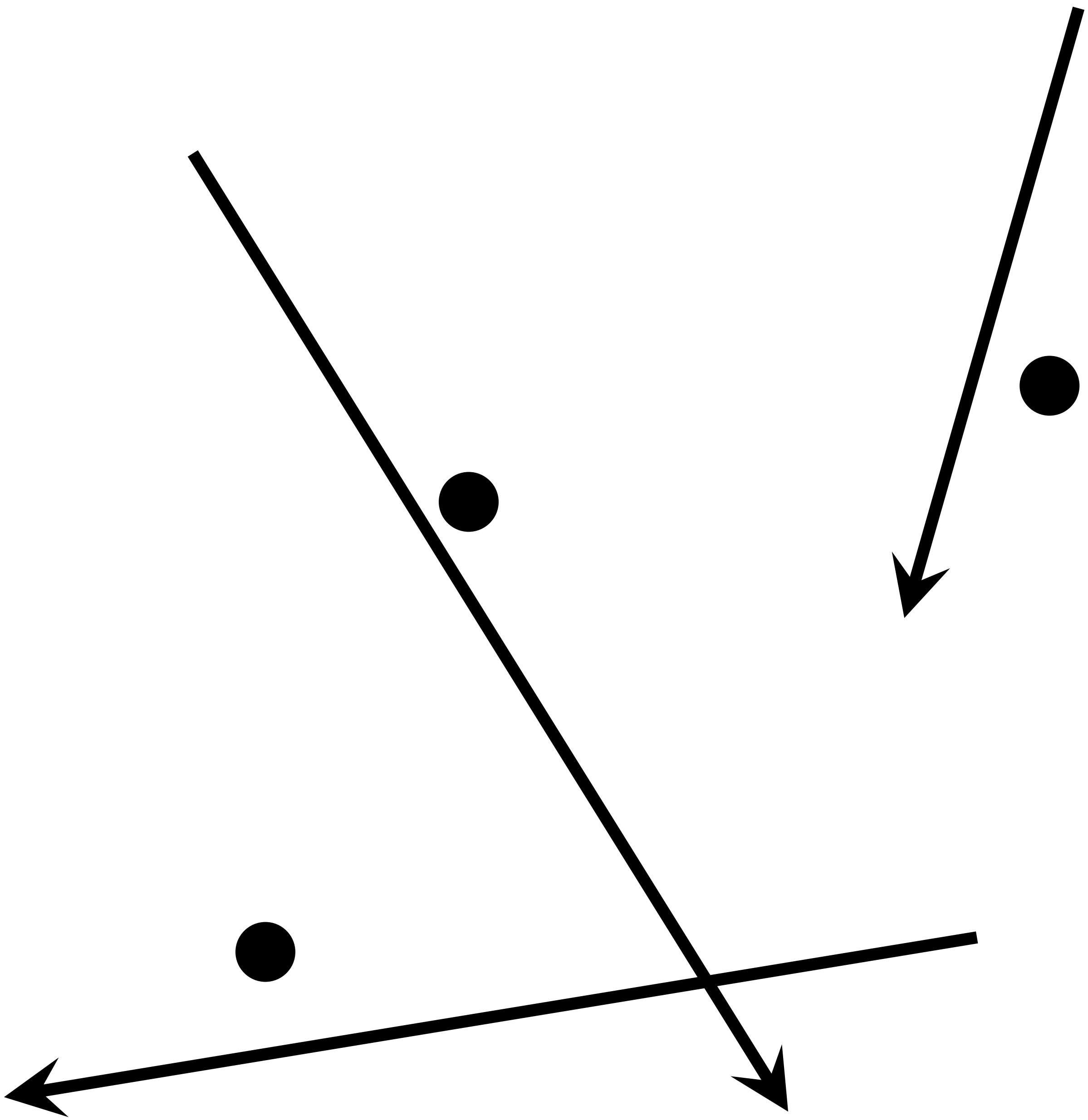




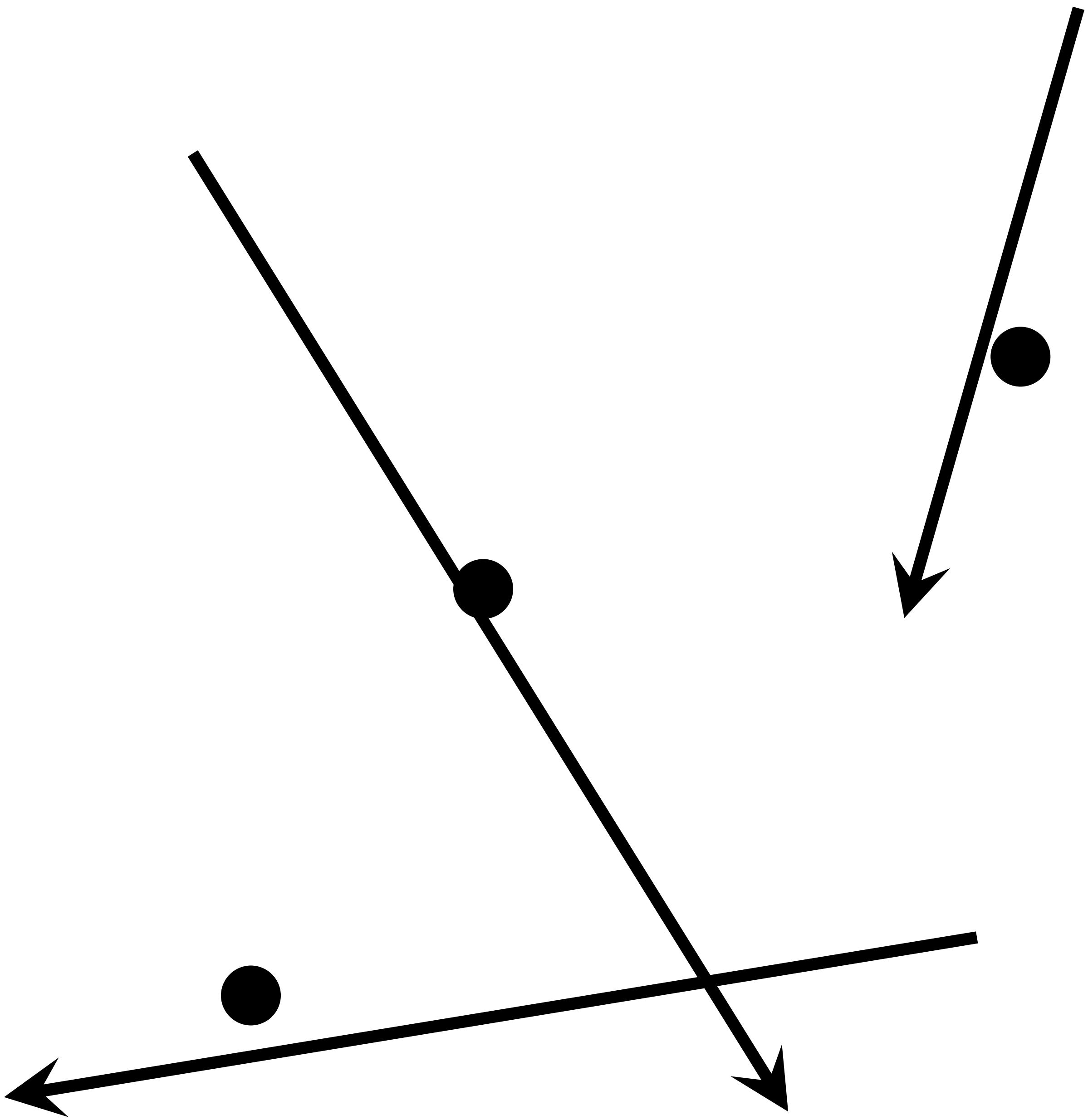




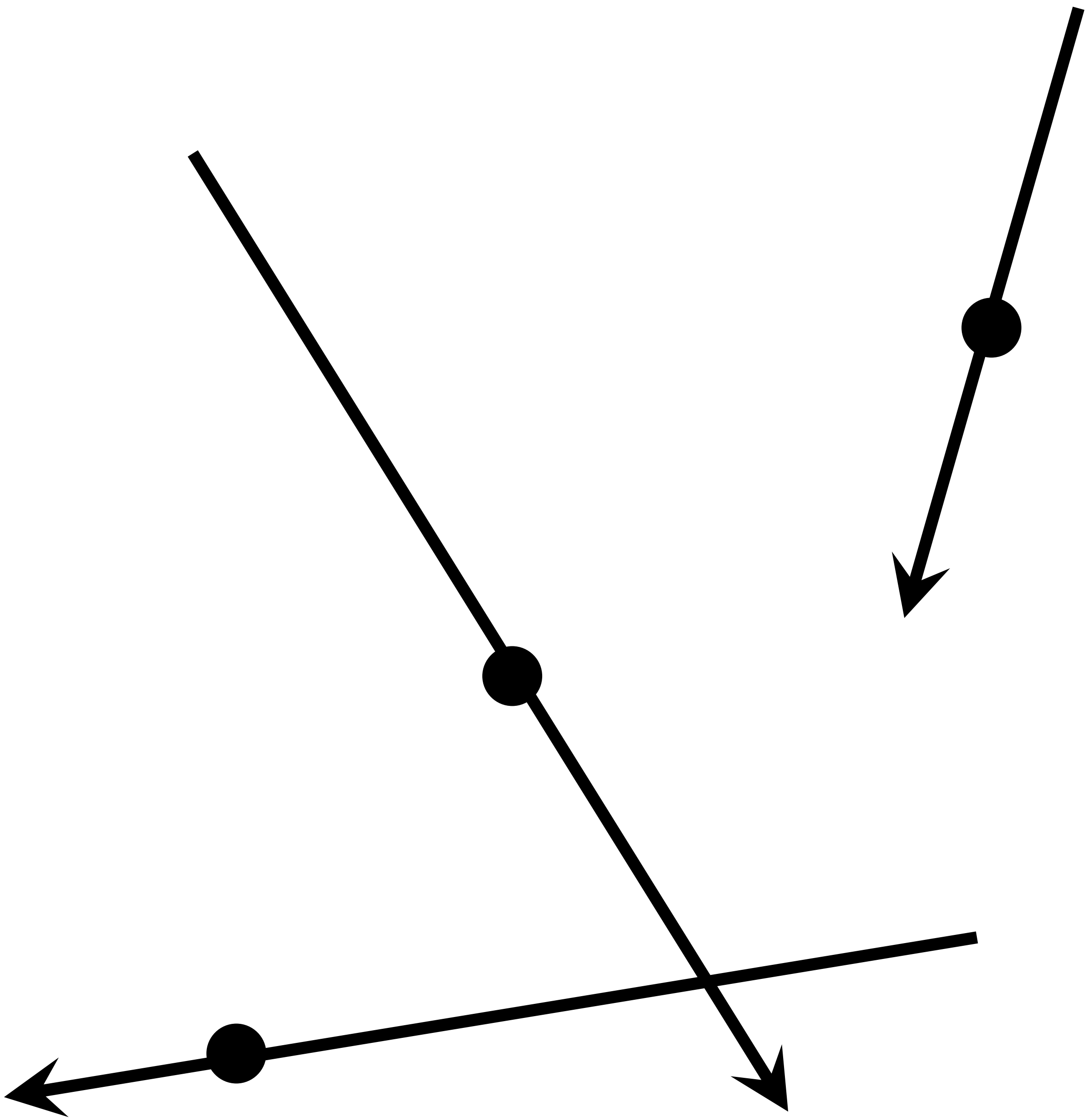


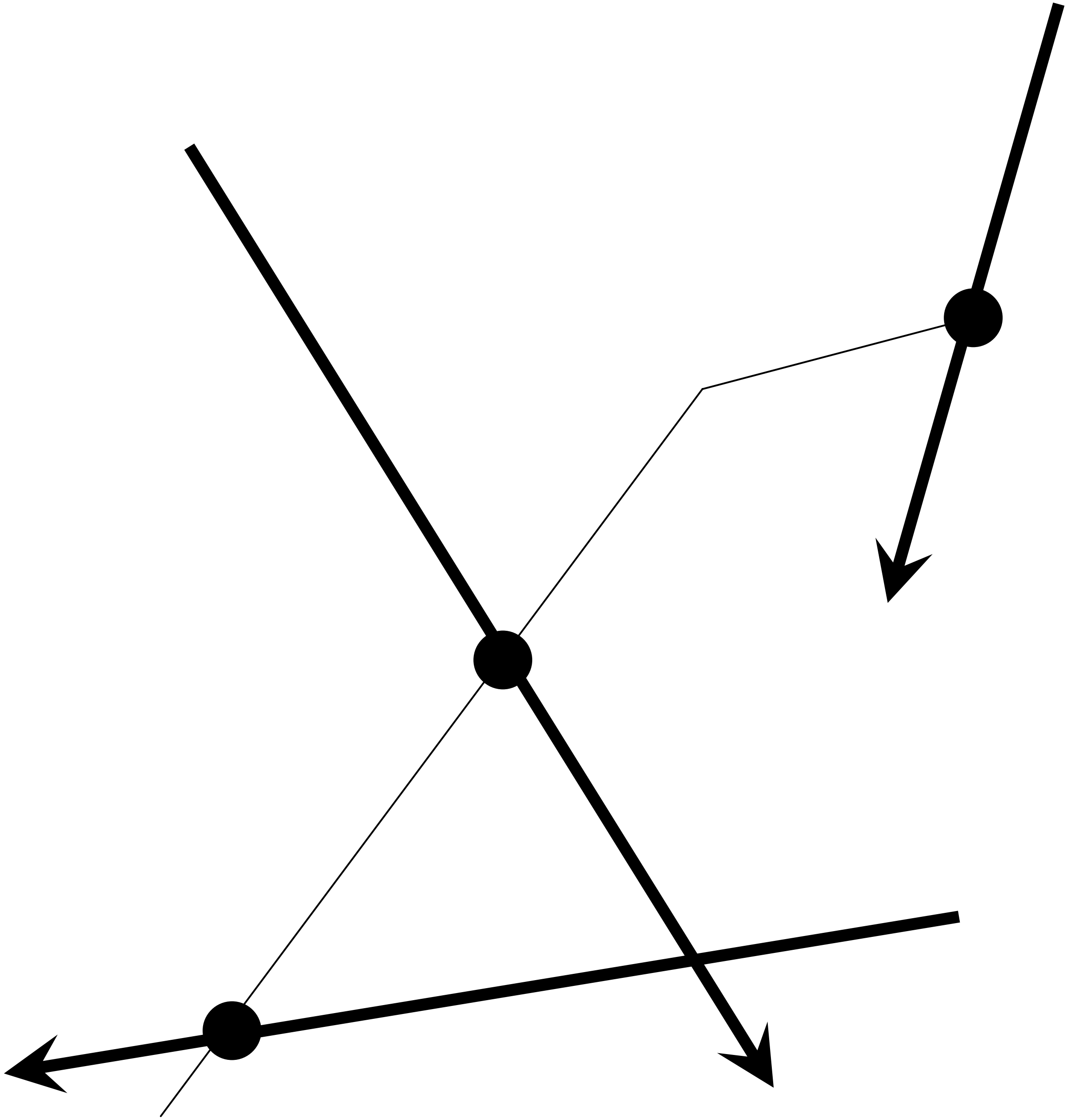


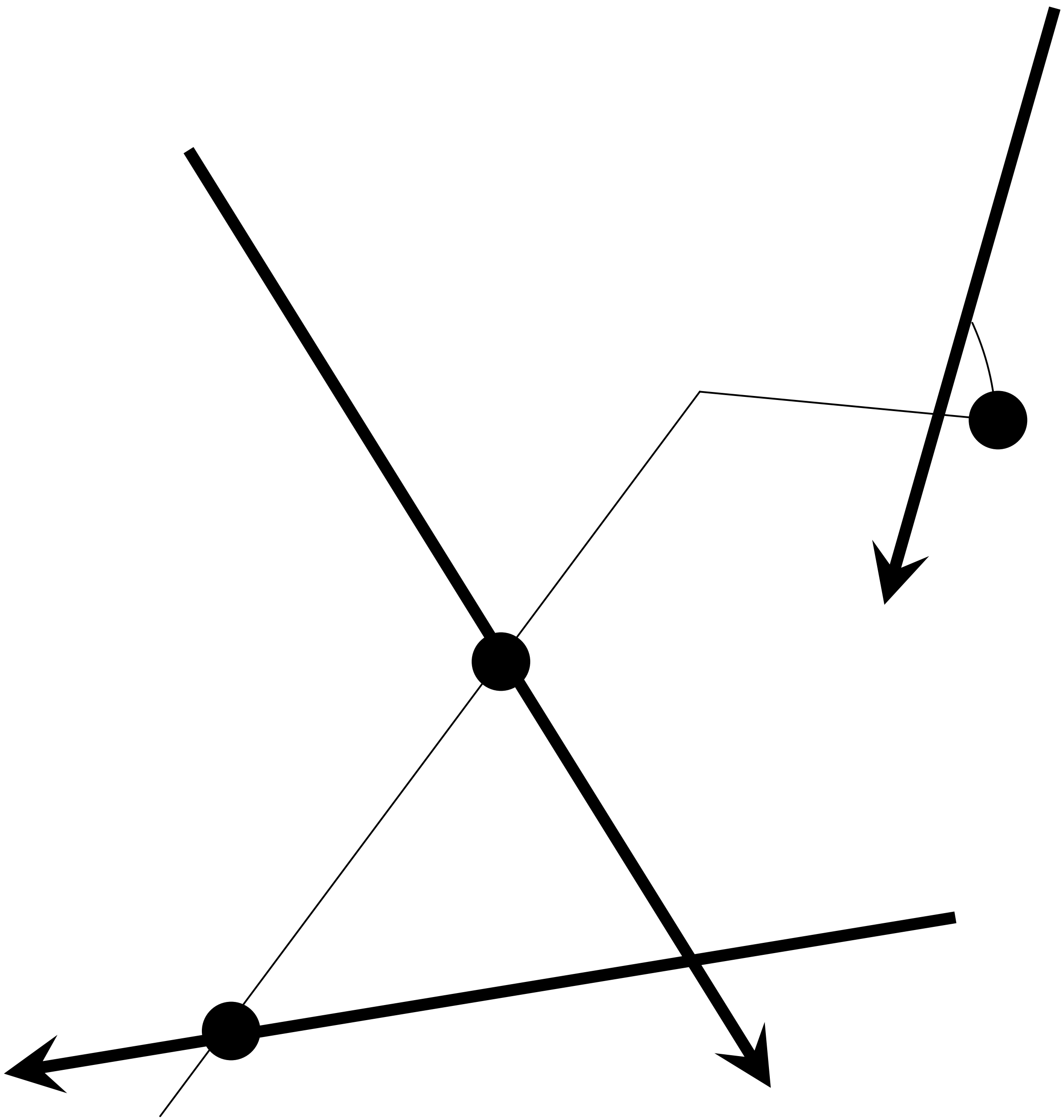
UK Center for Visualization & Virtual Environments

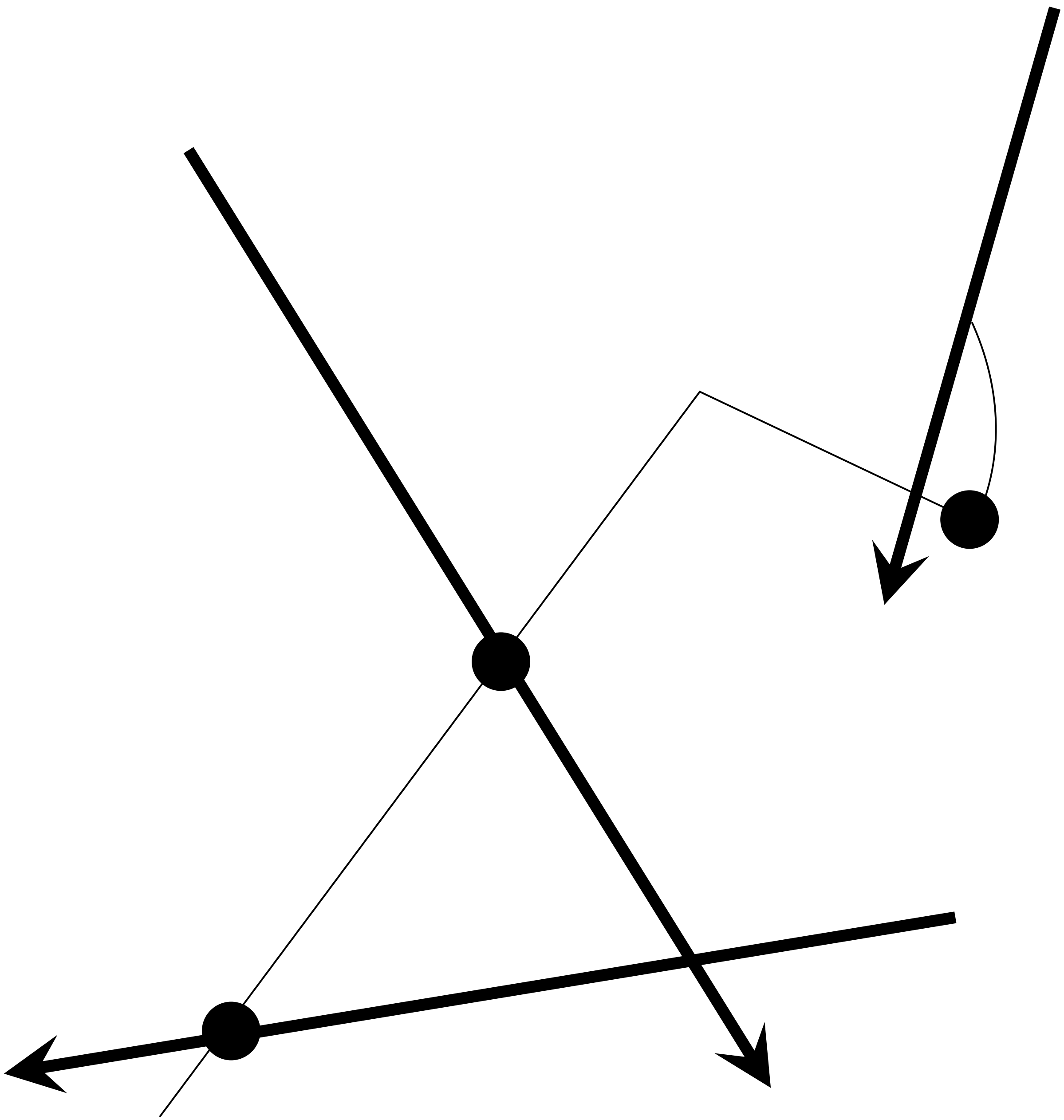


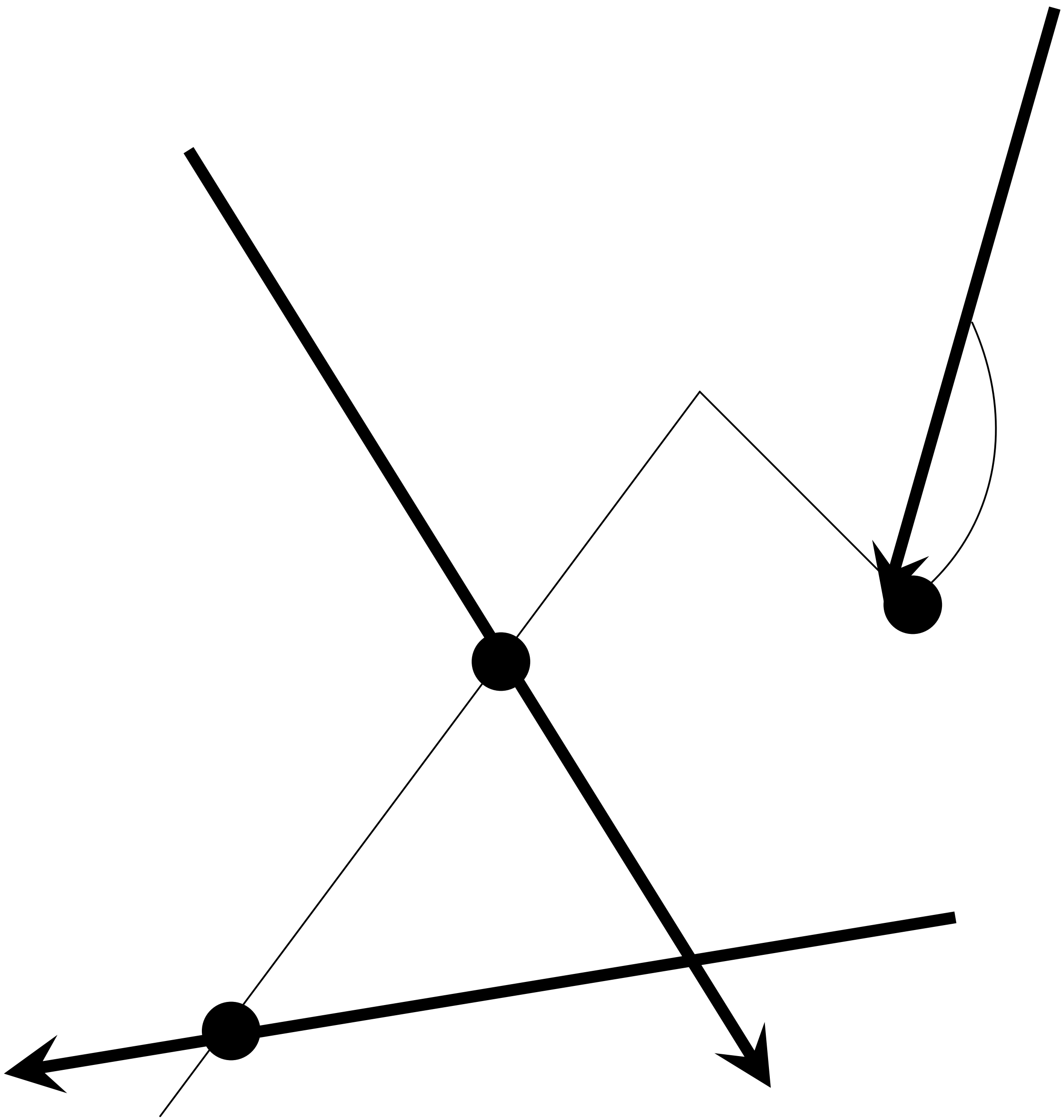
UK Center for Visualization & Virtual Environments



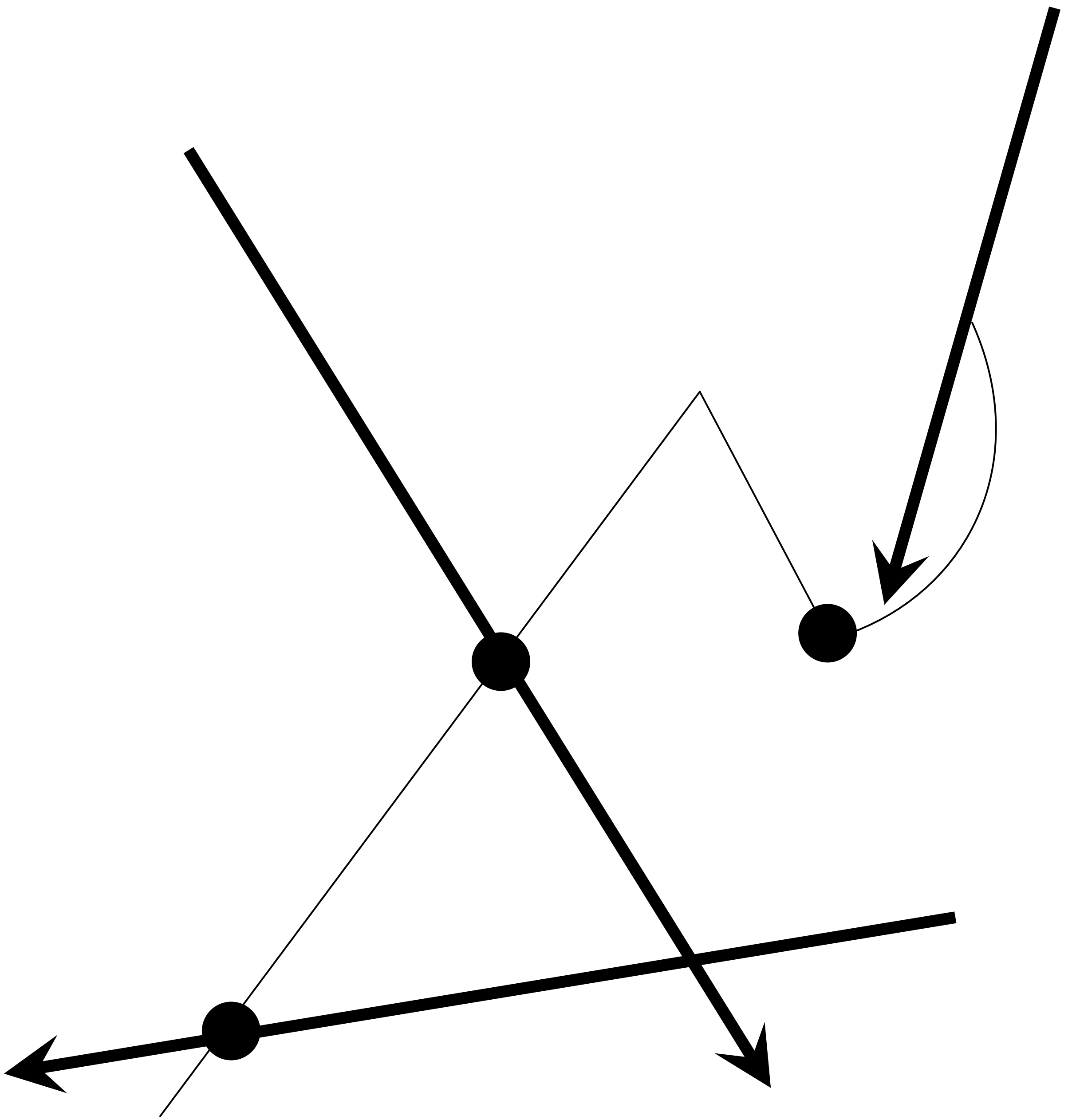


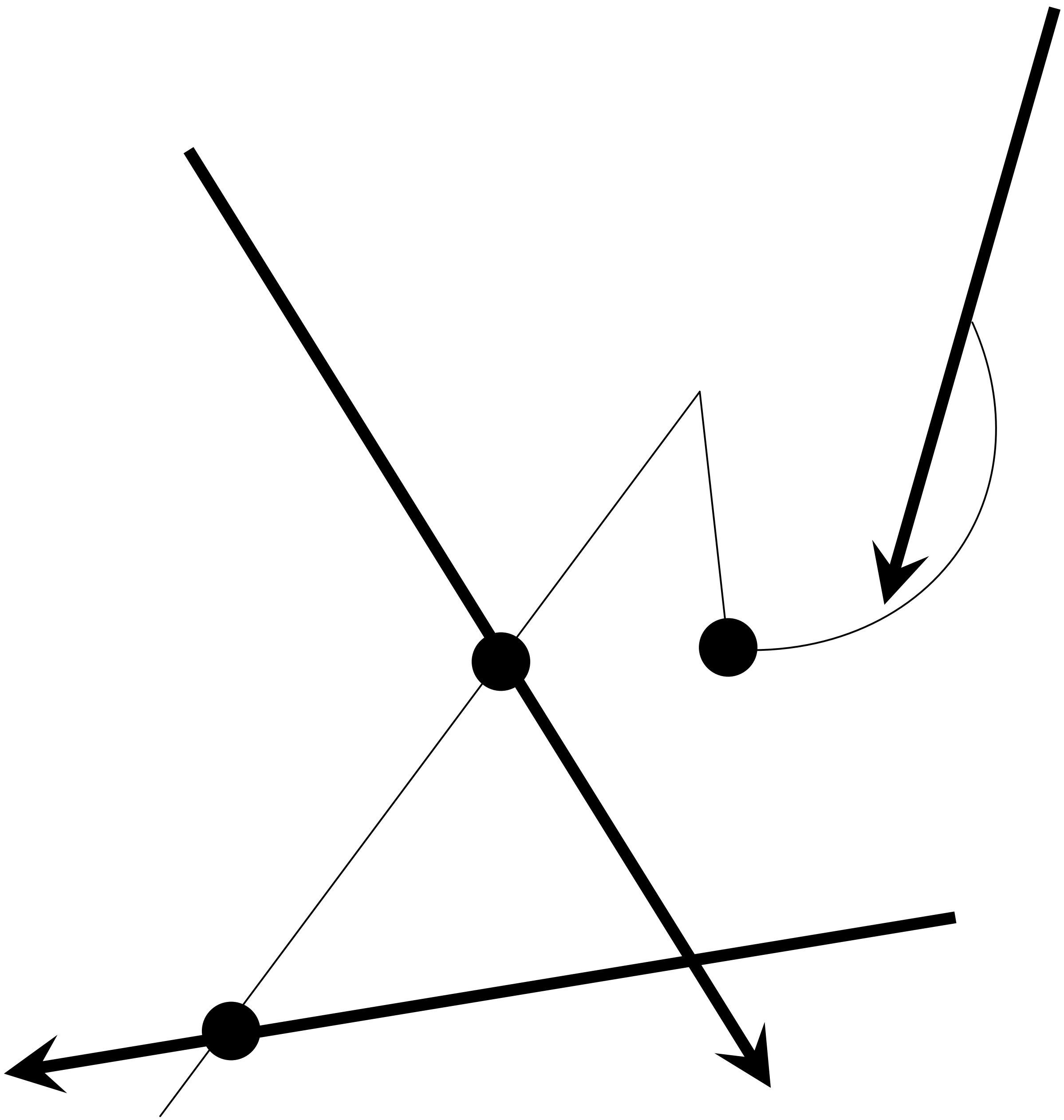


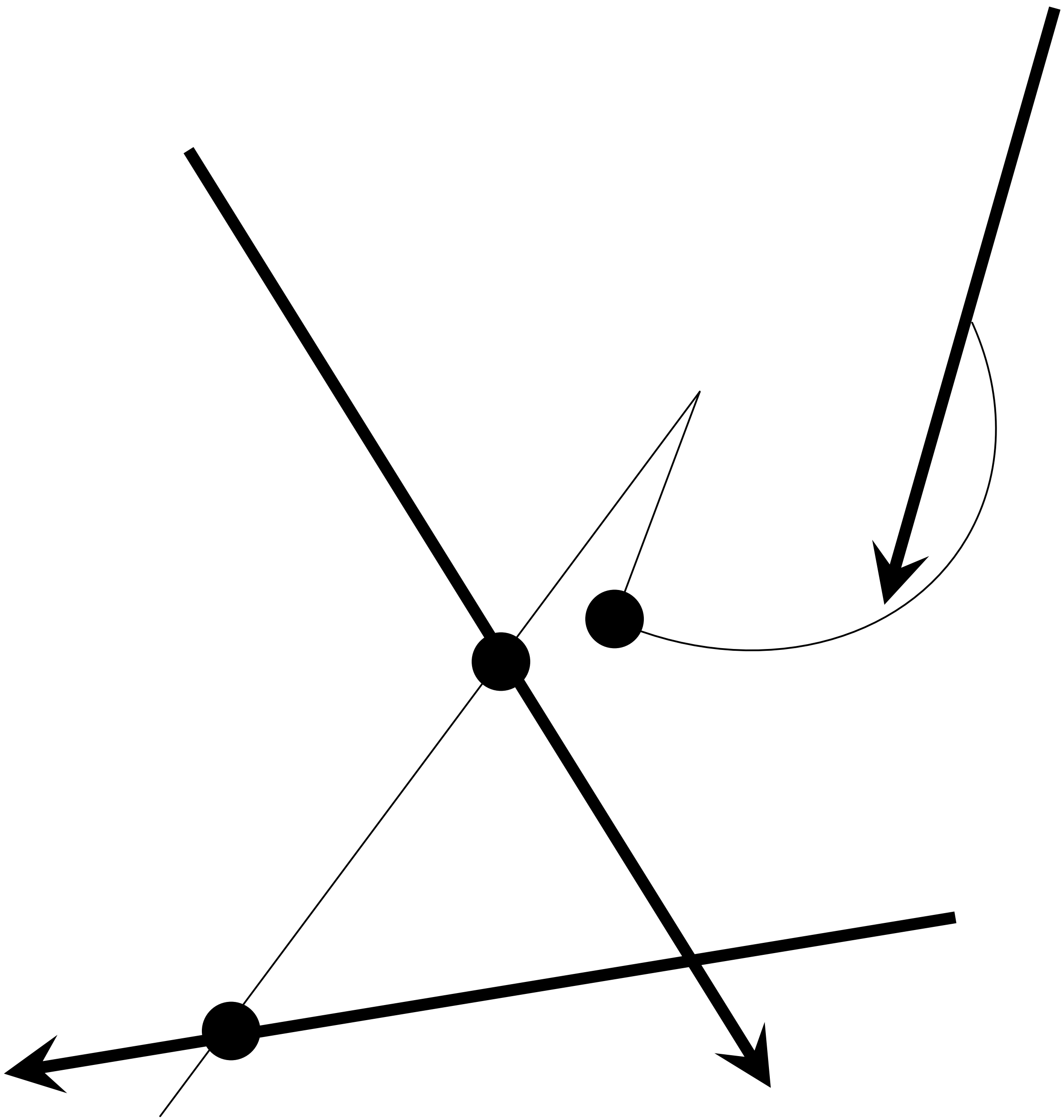


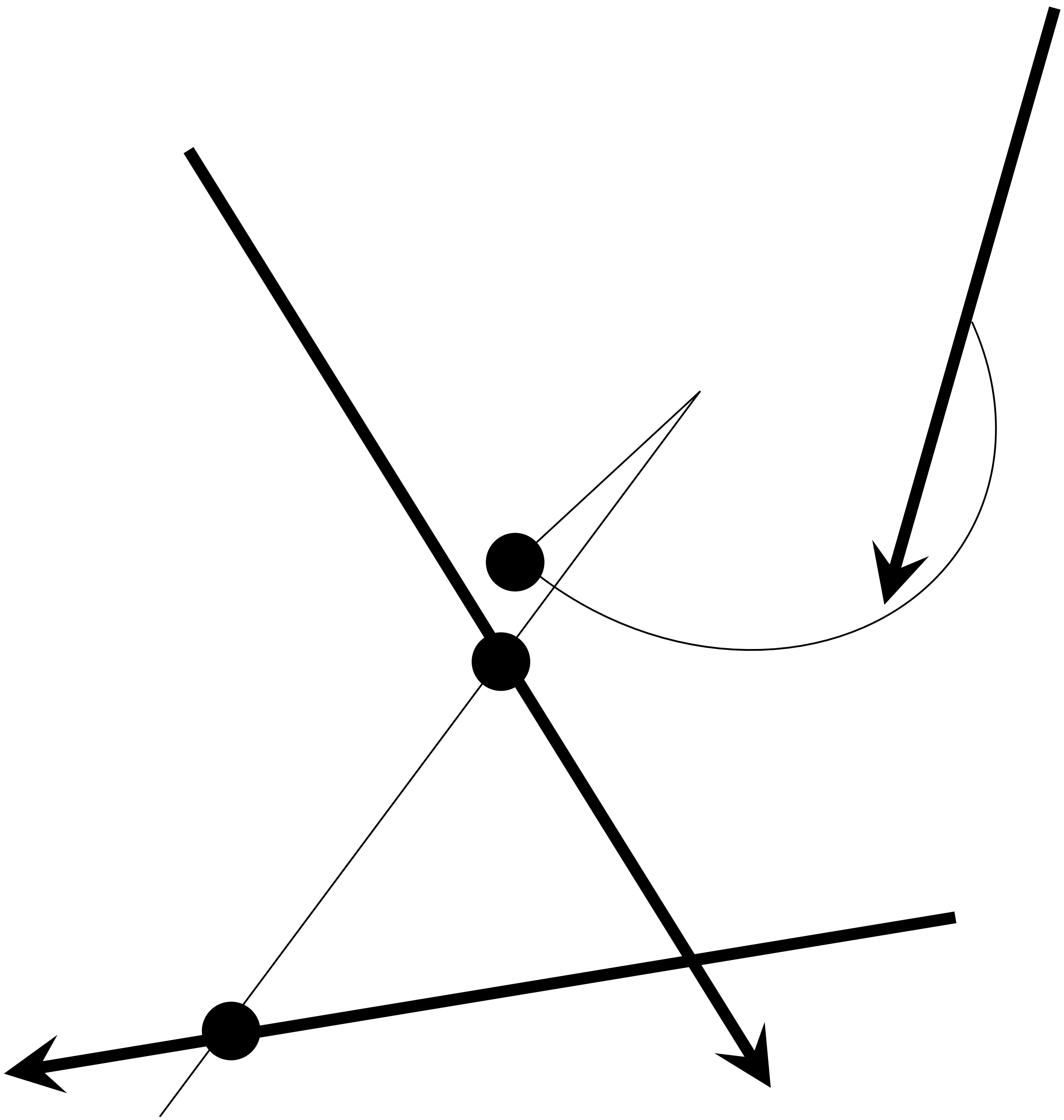


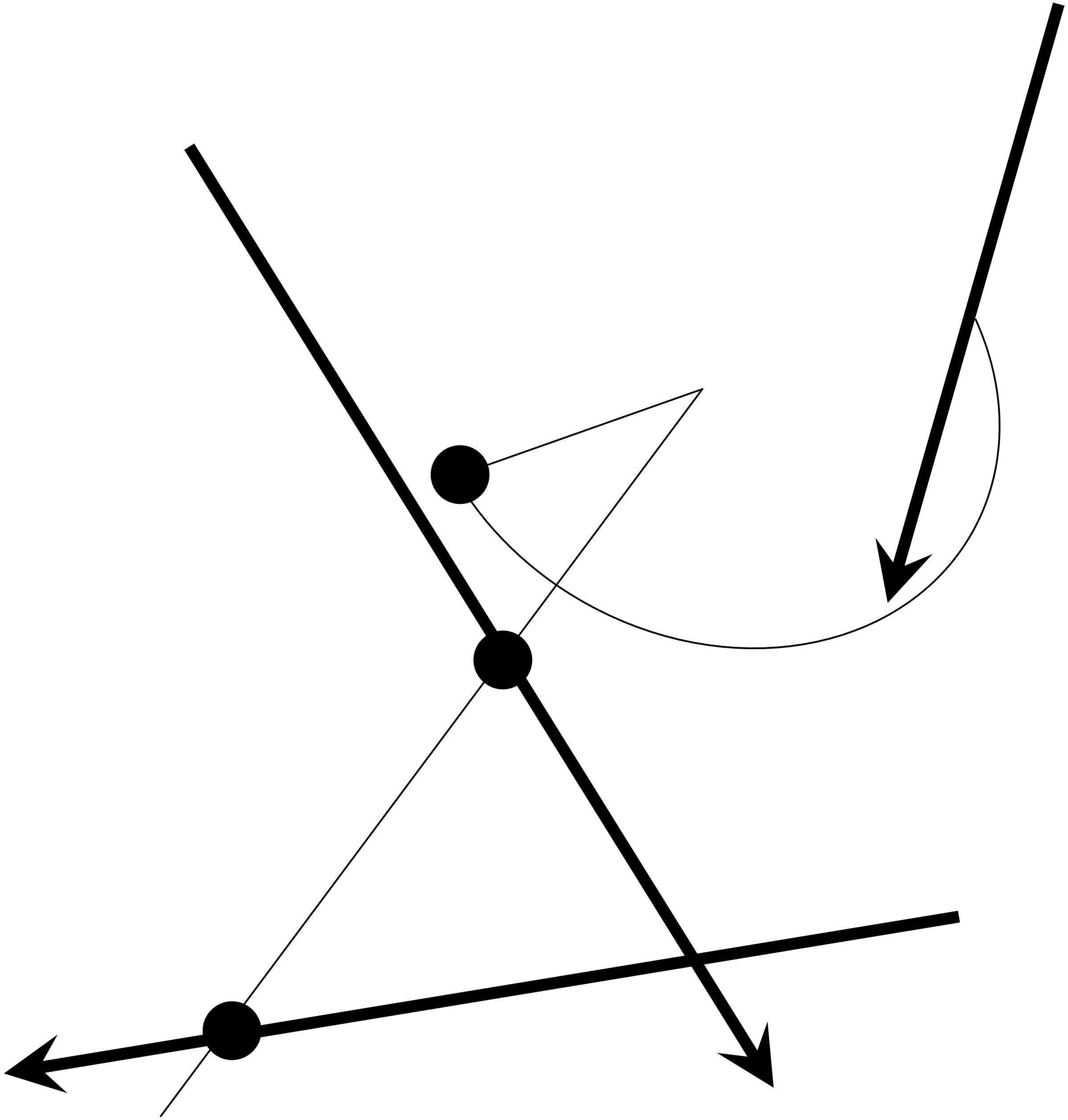
UK Center for Visualization & Virtual Environments

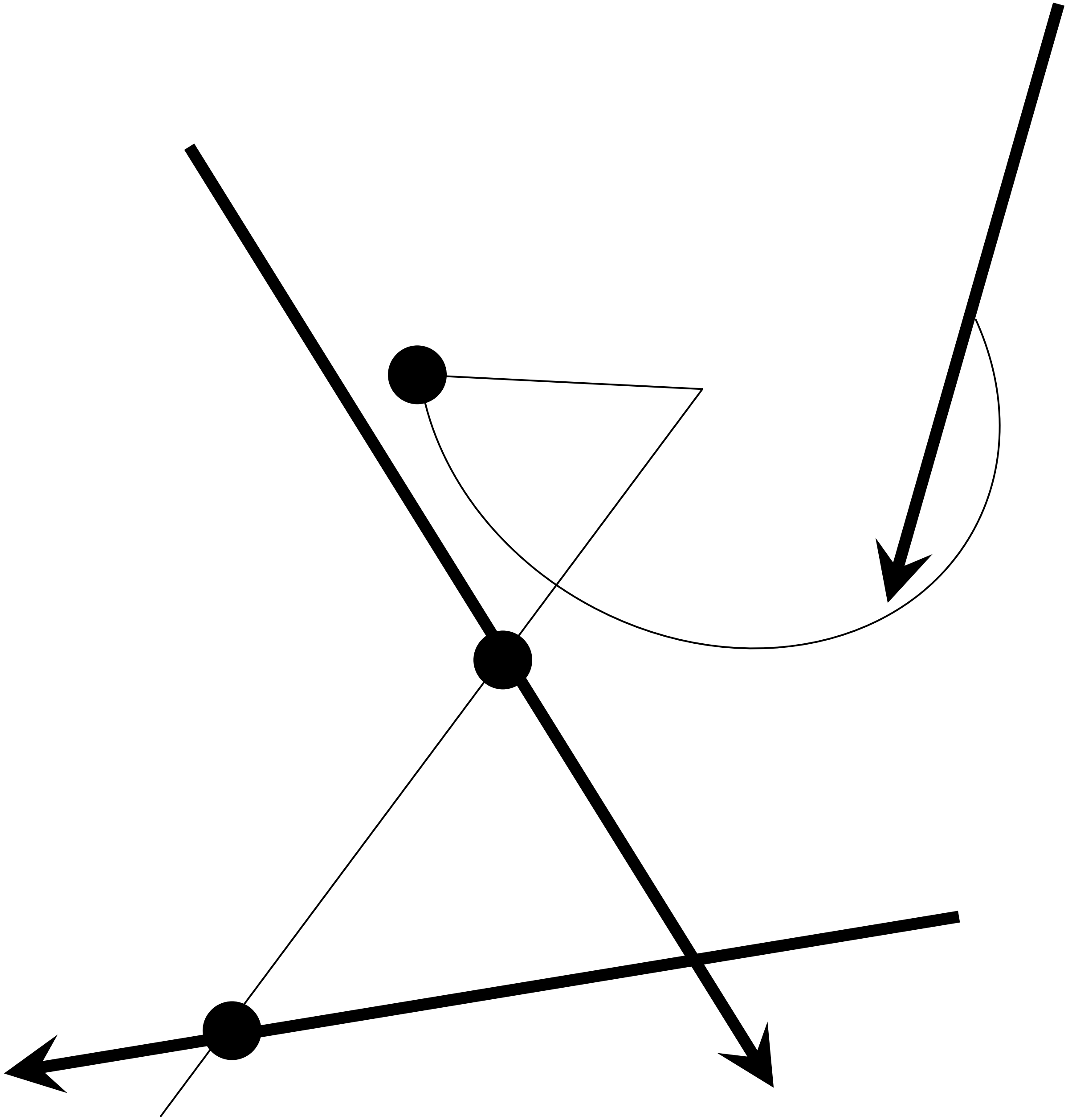


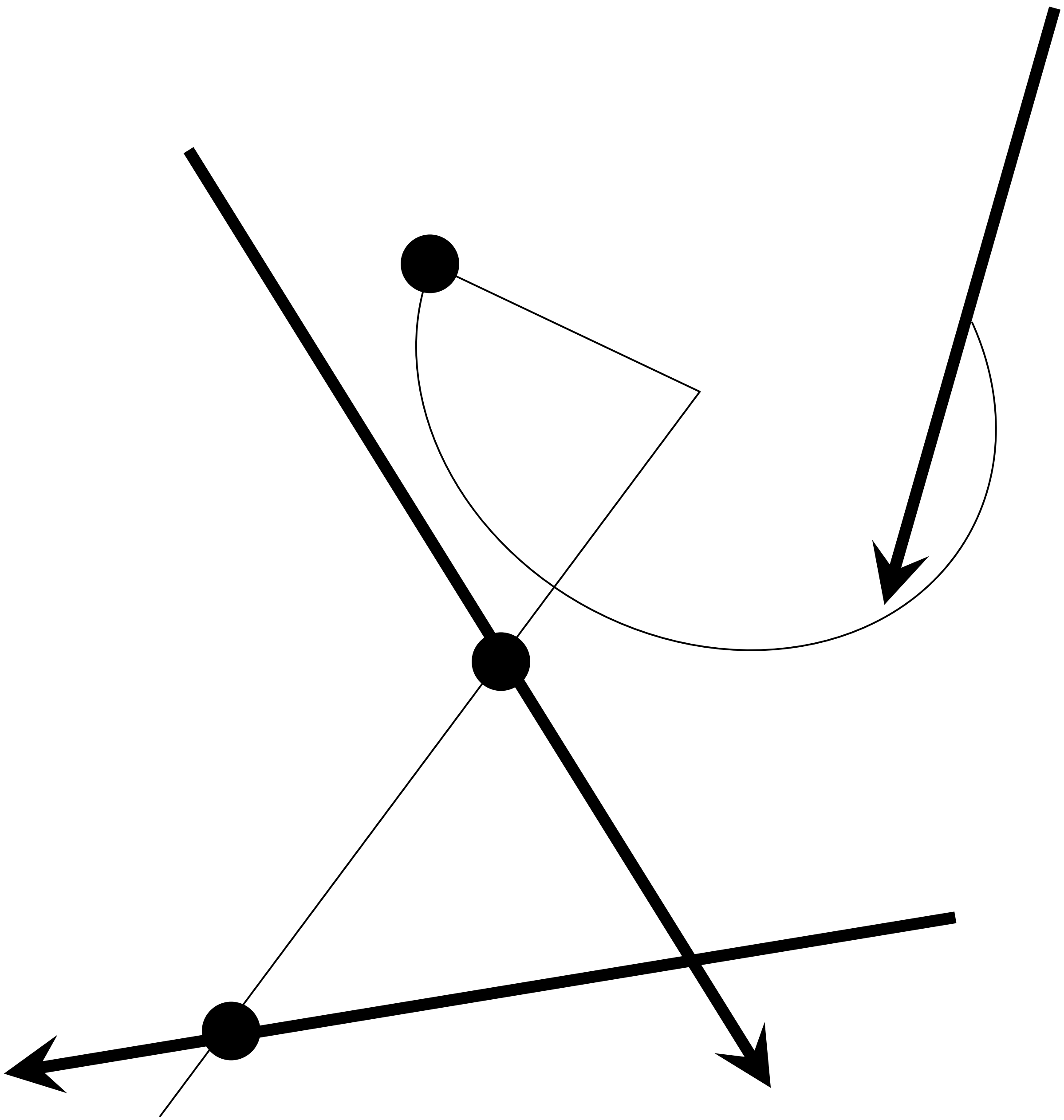


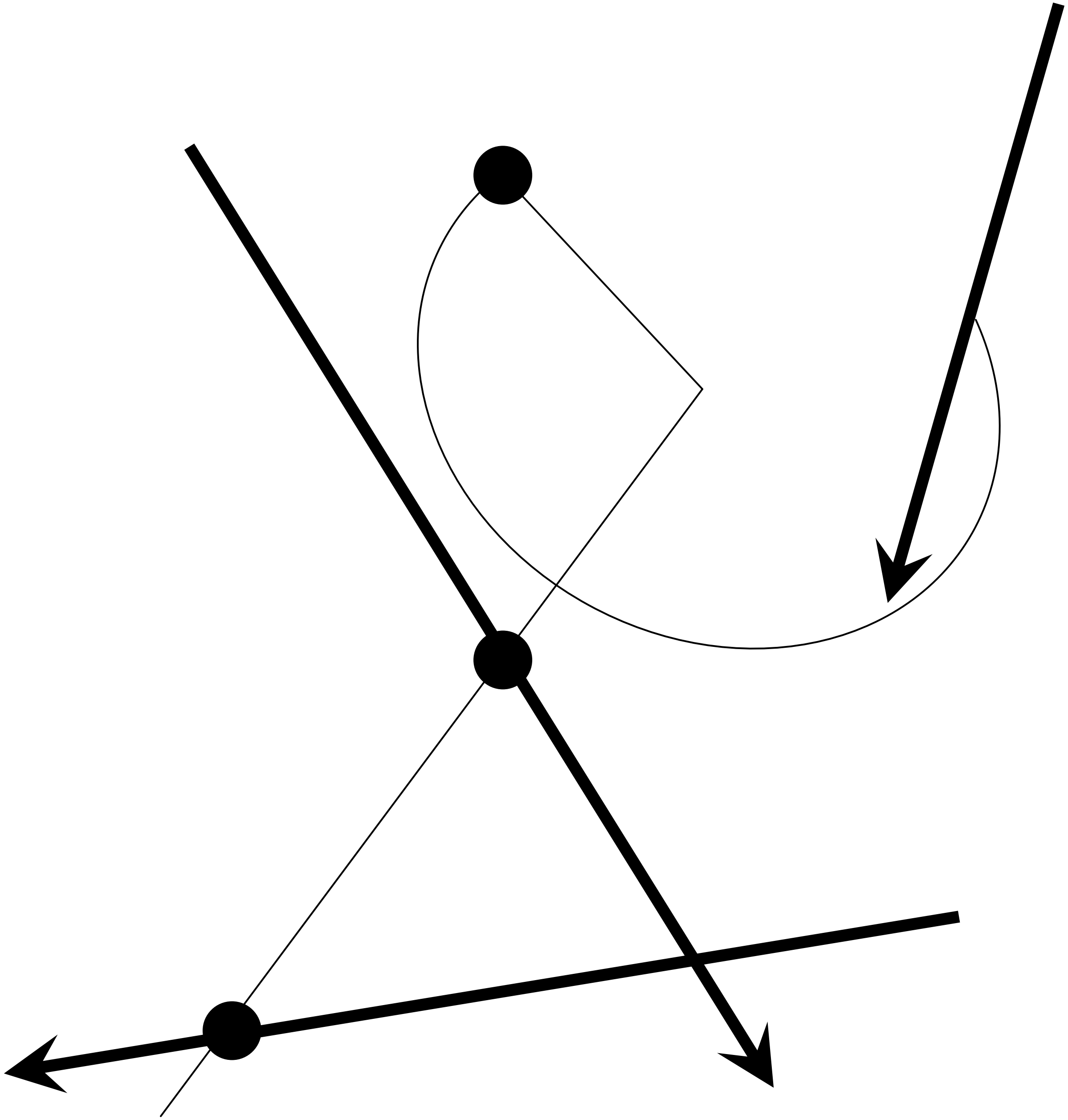




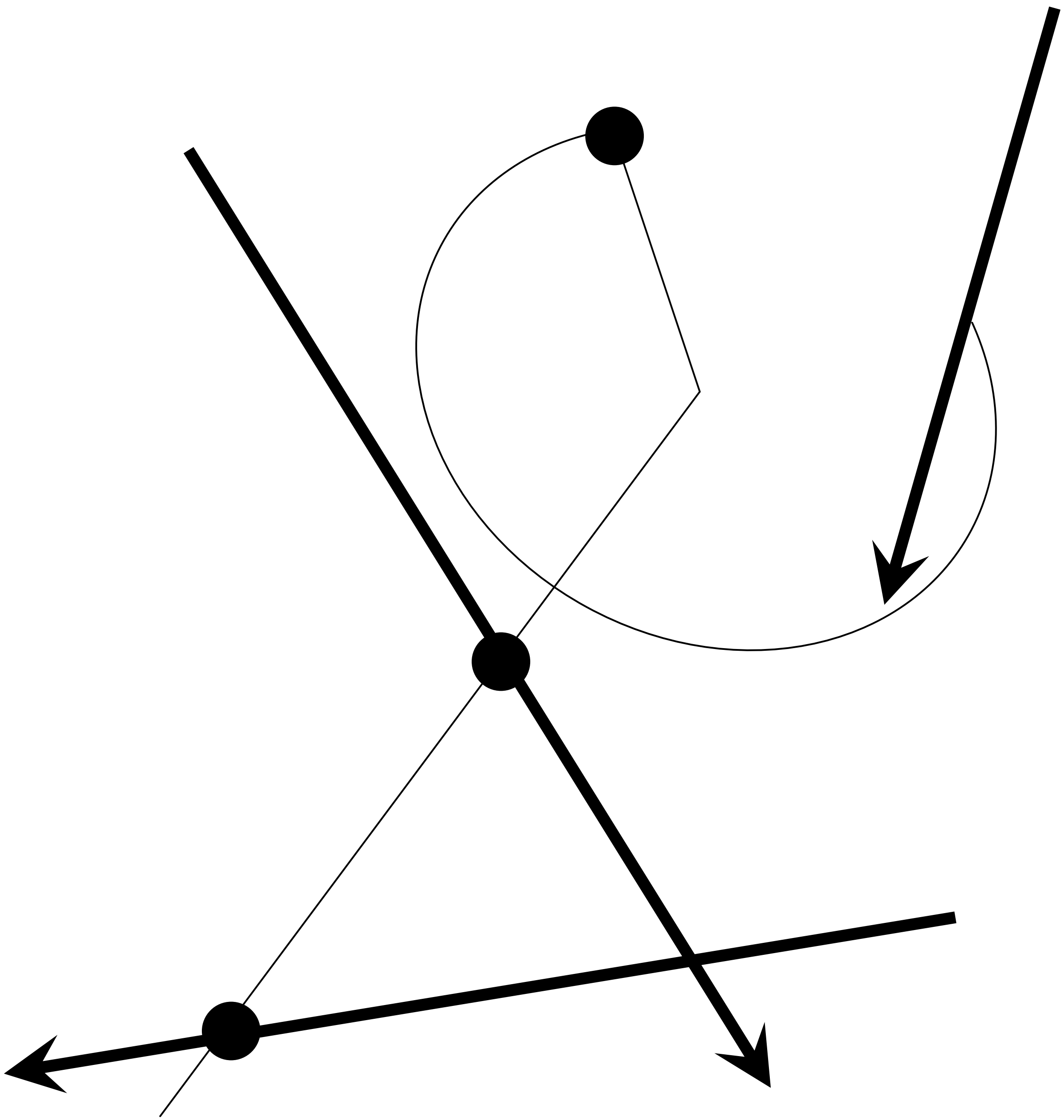


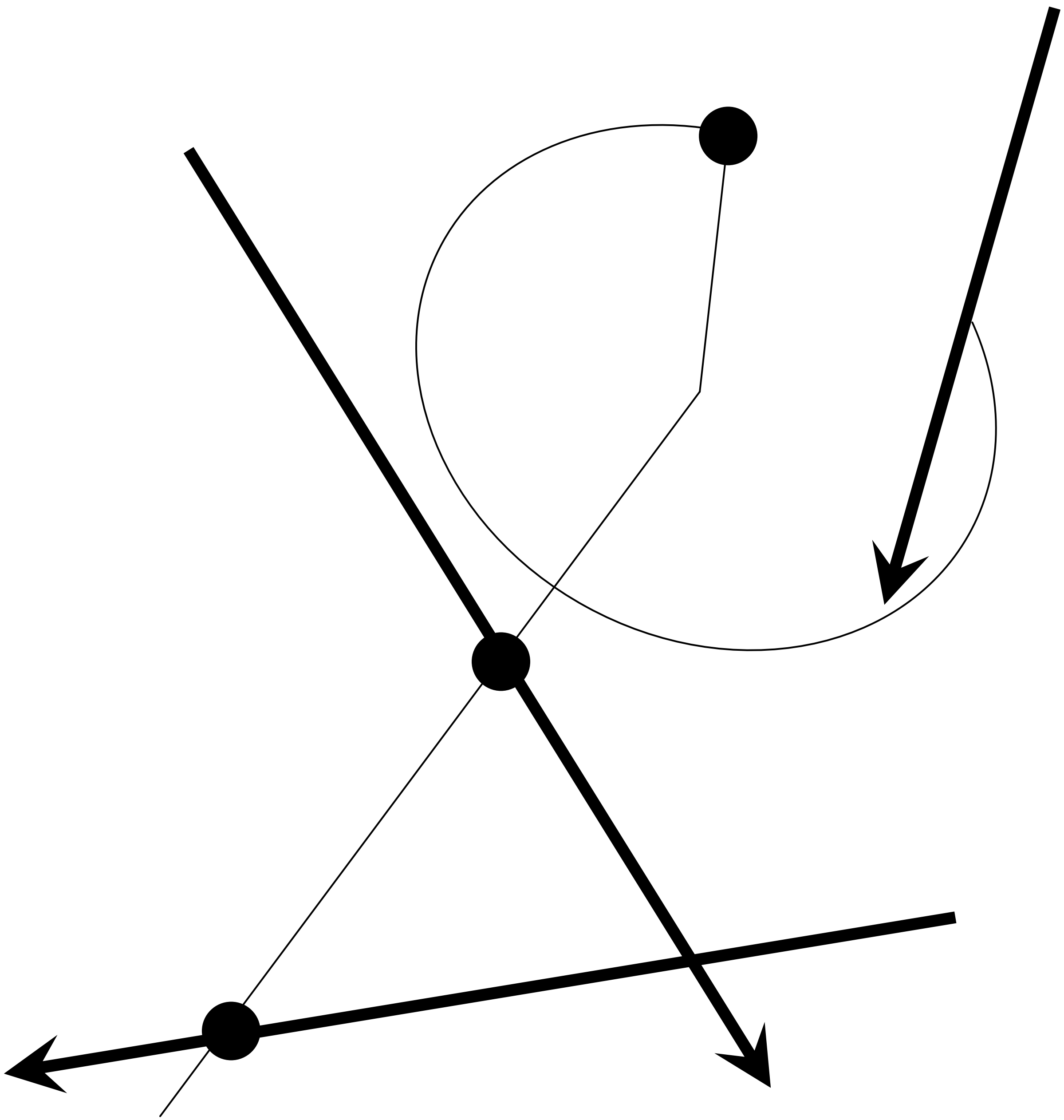


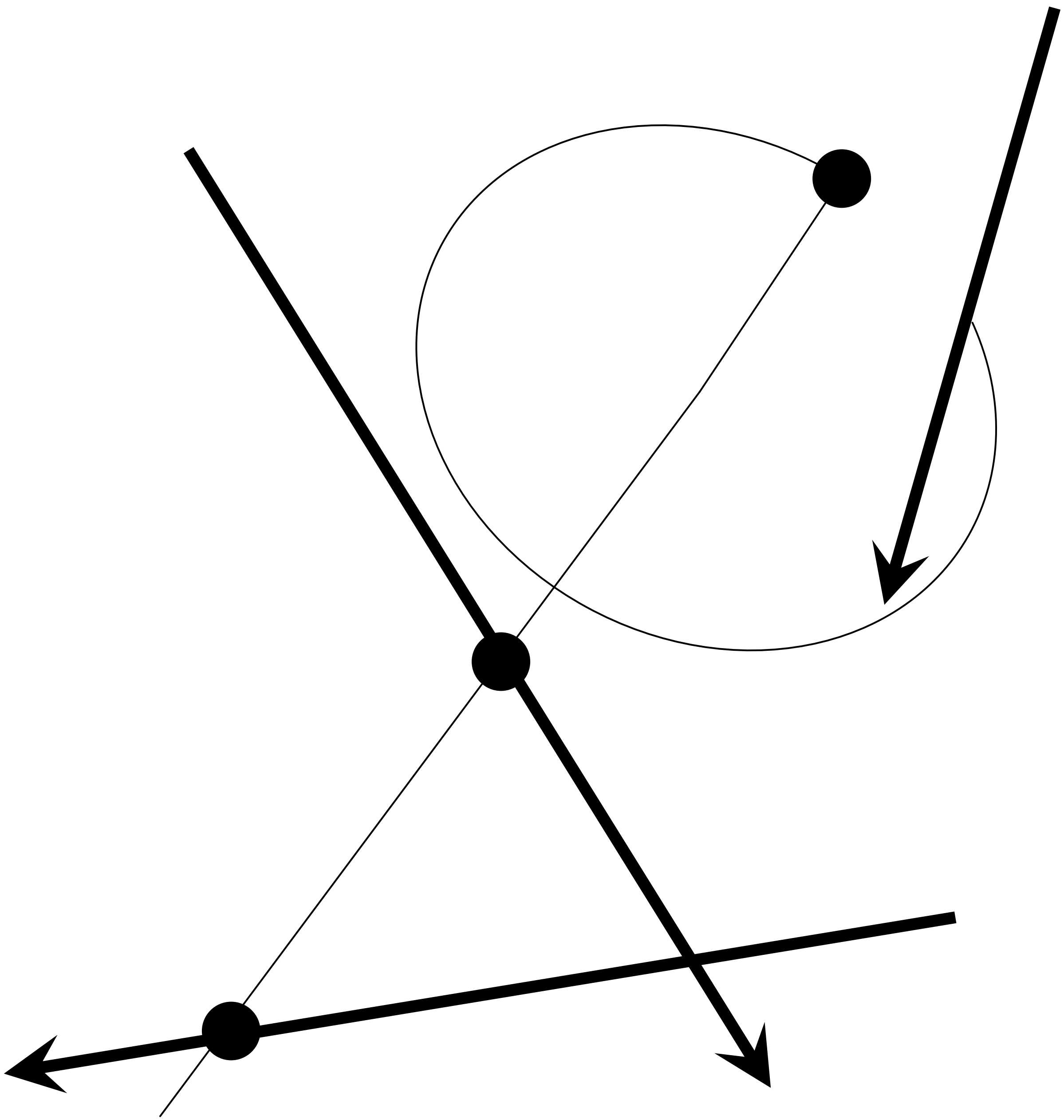


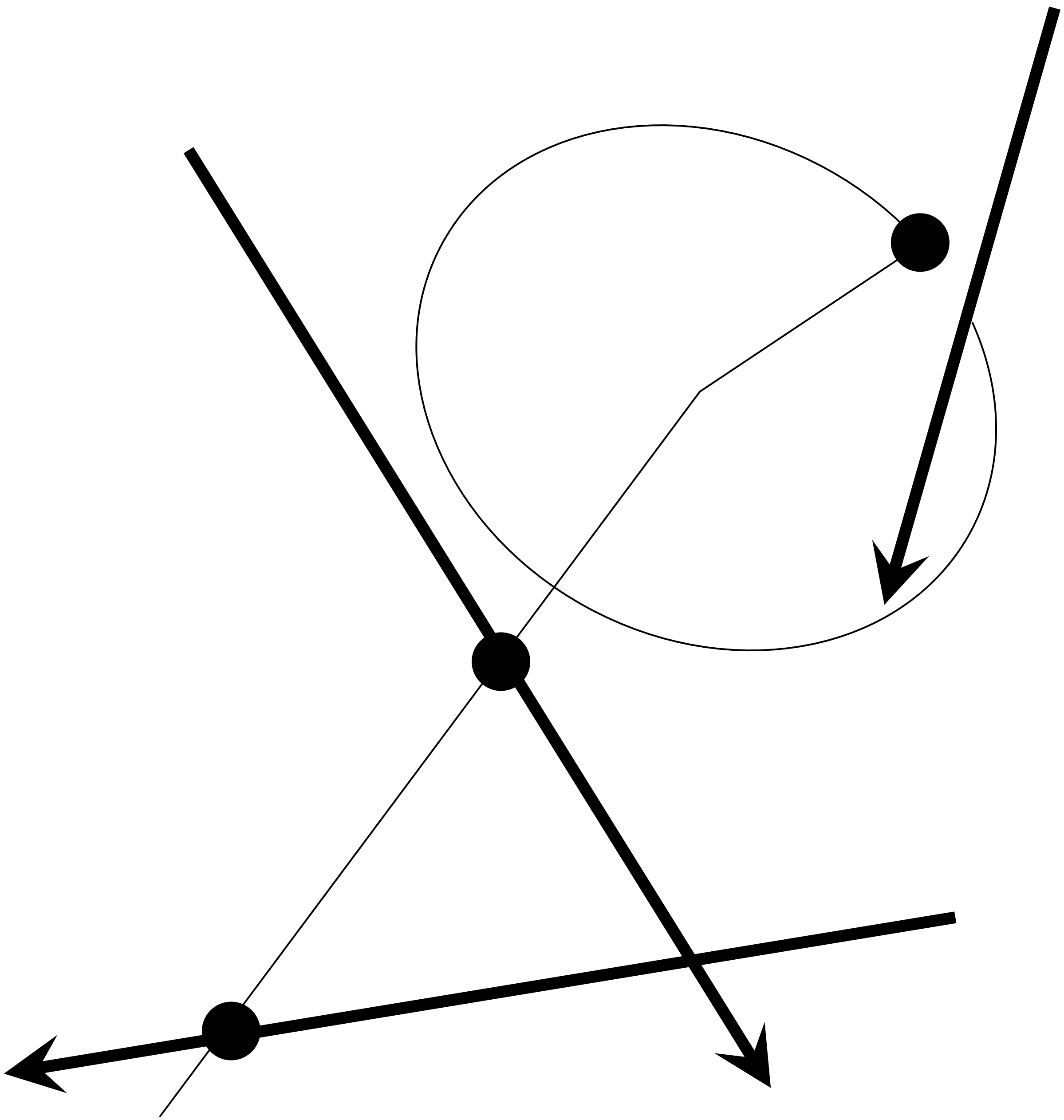


UK Center for Visualization & Virtual Environments

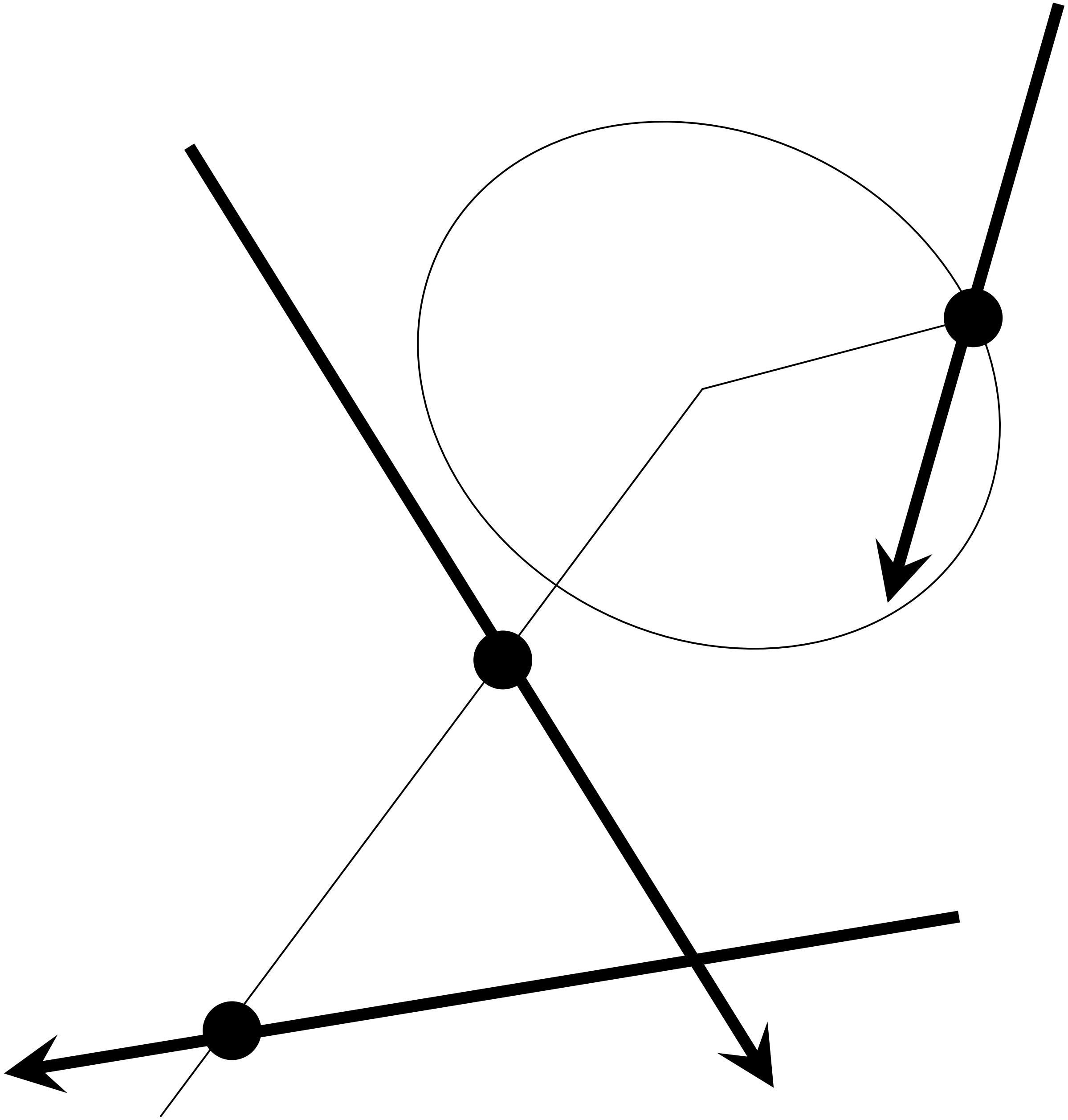




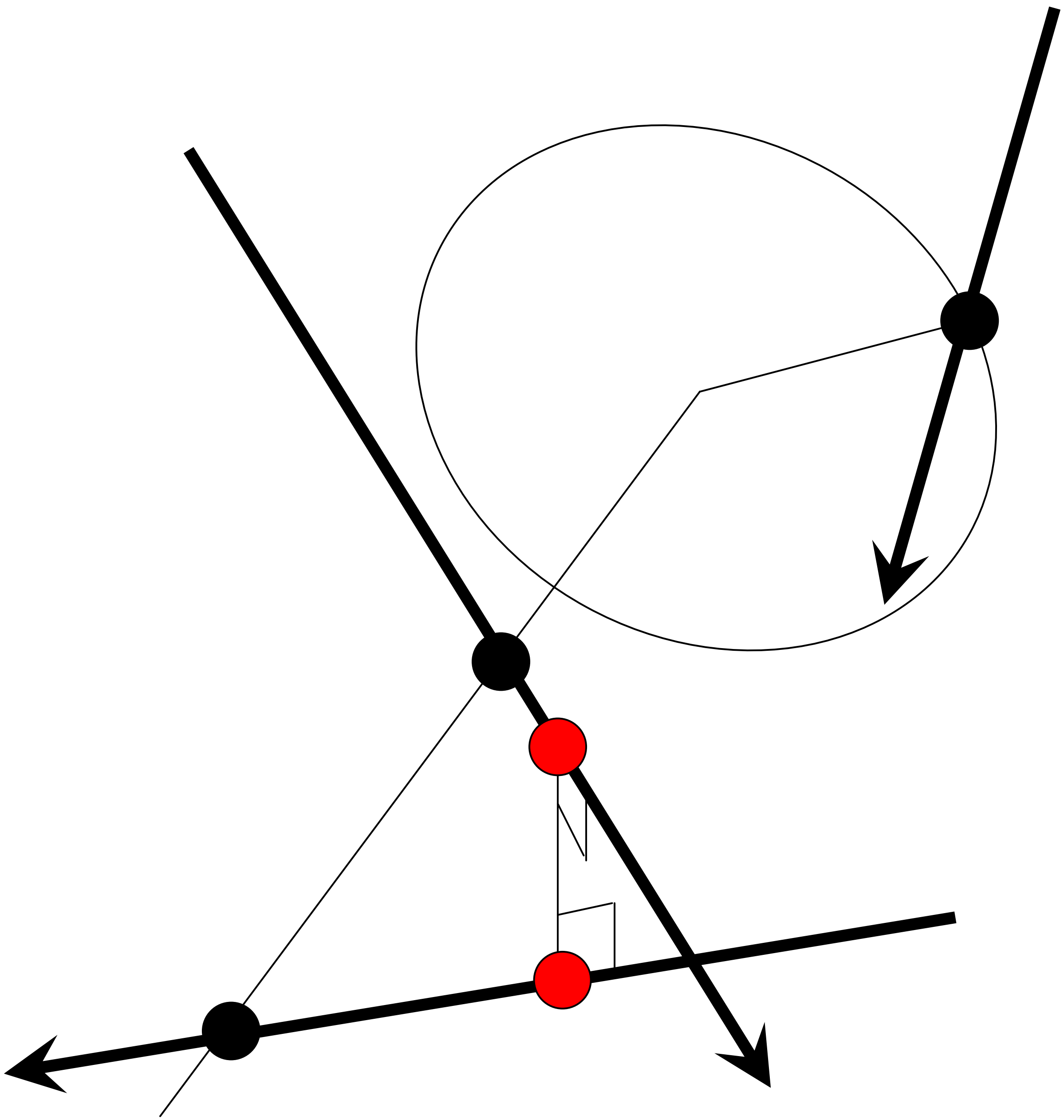


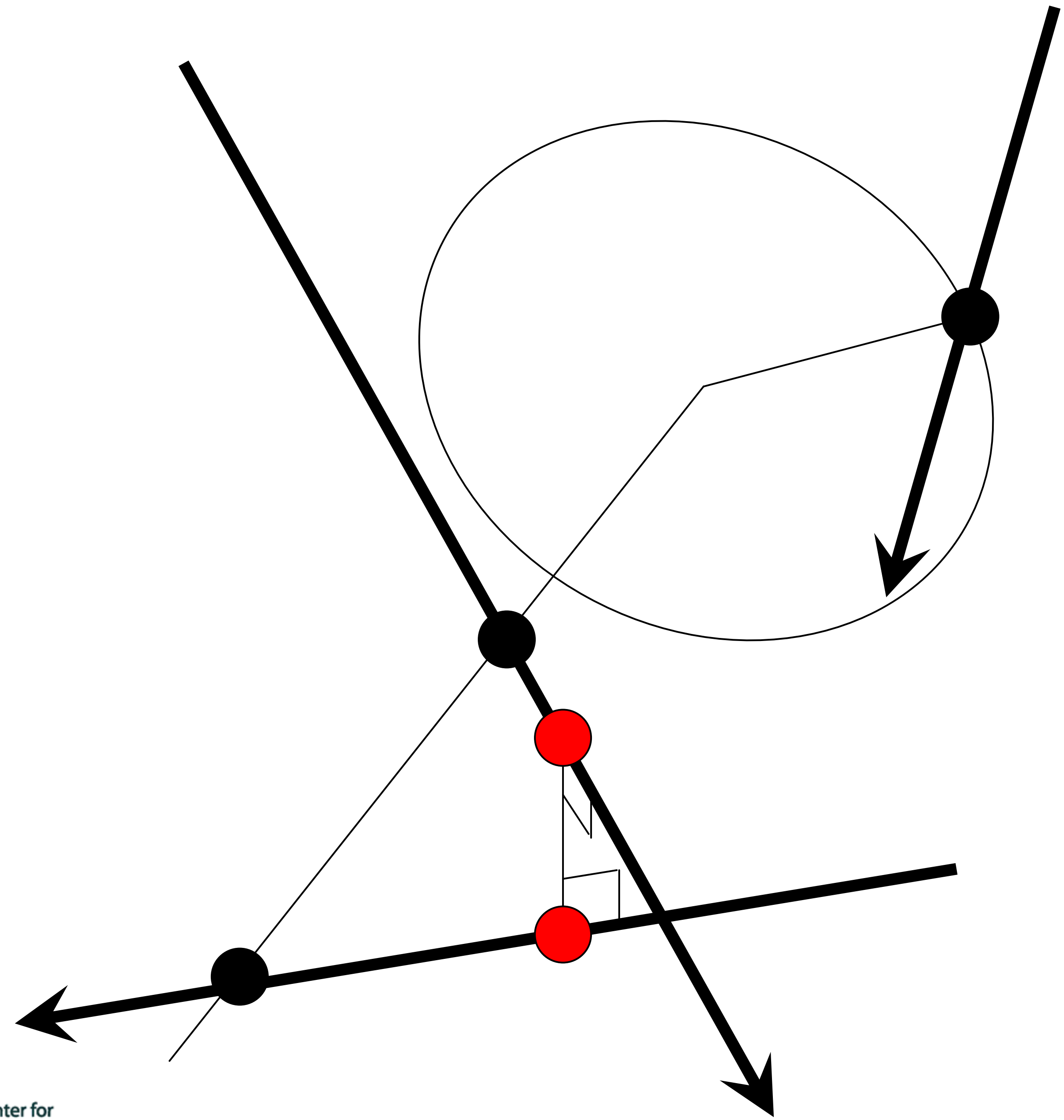


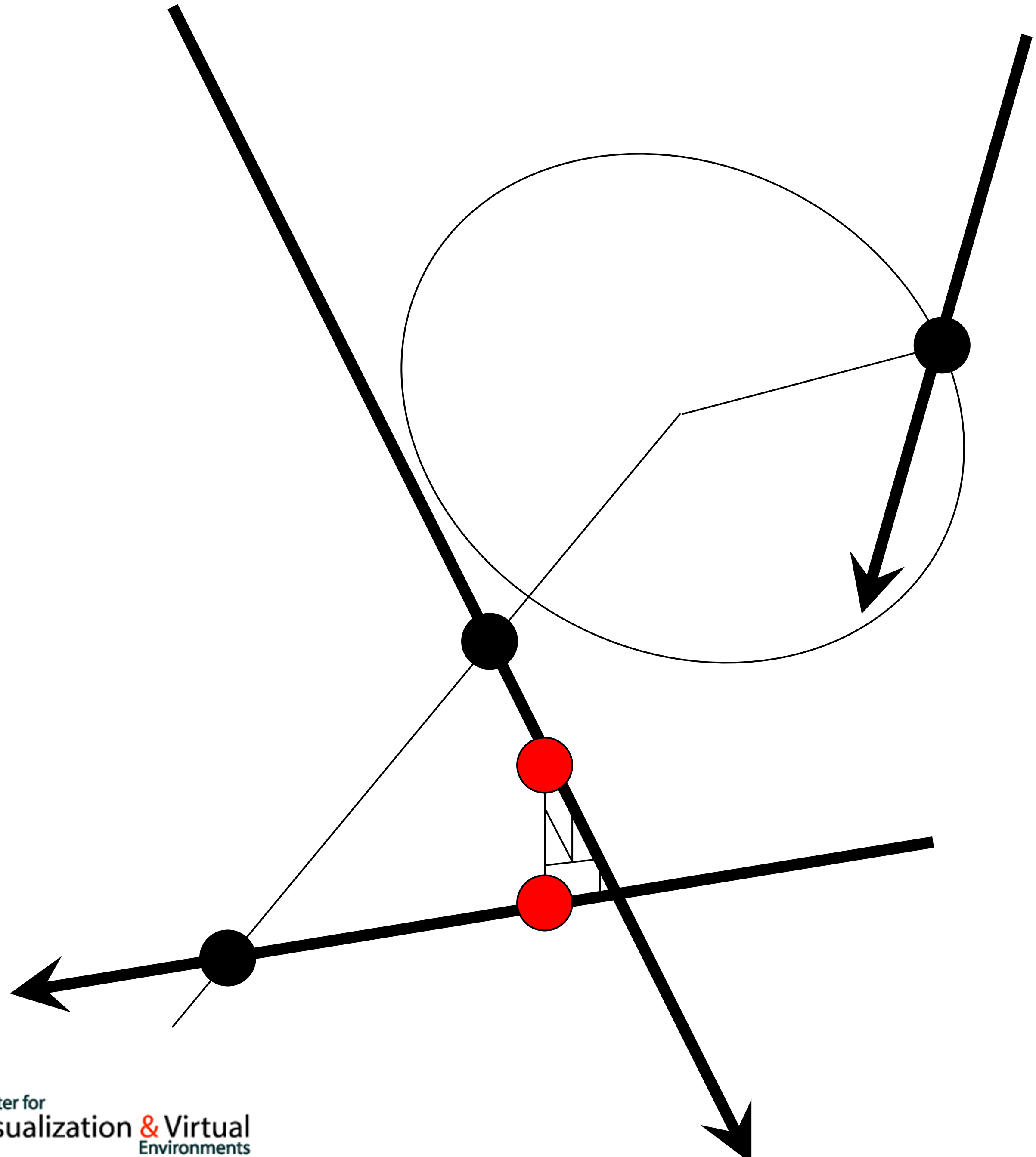
UK Center for Visualization & Virtual Environments

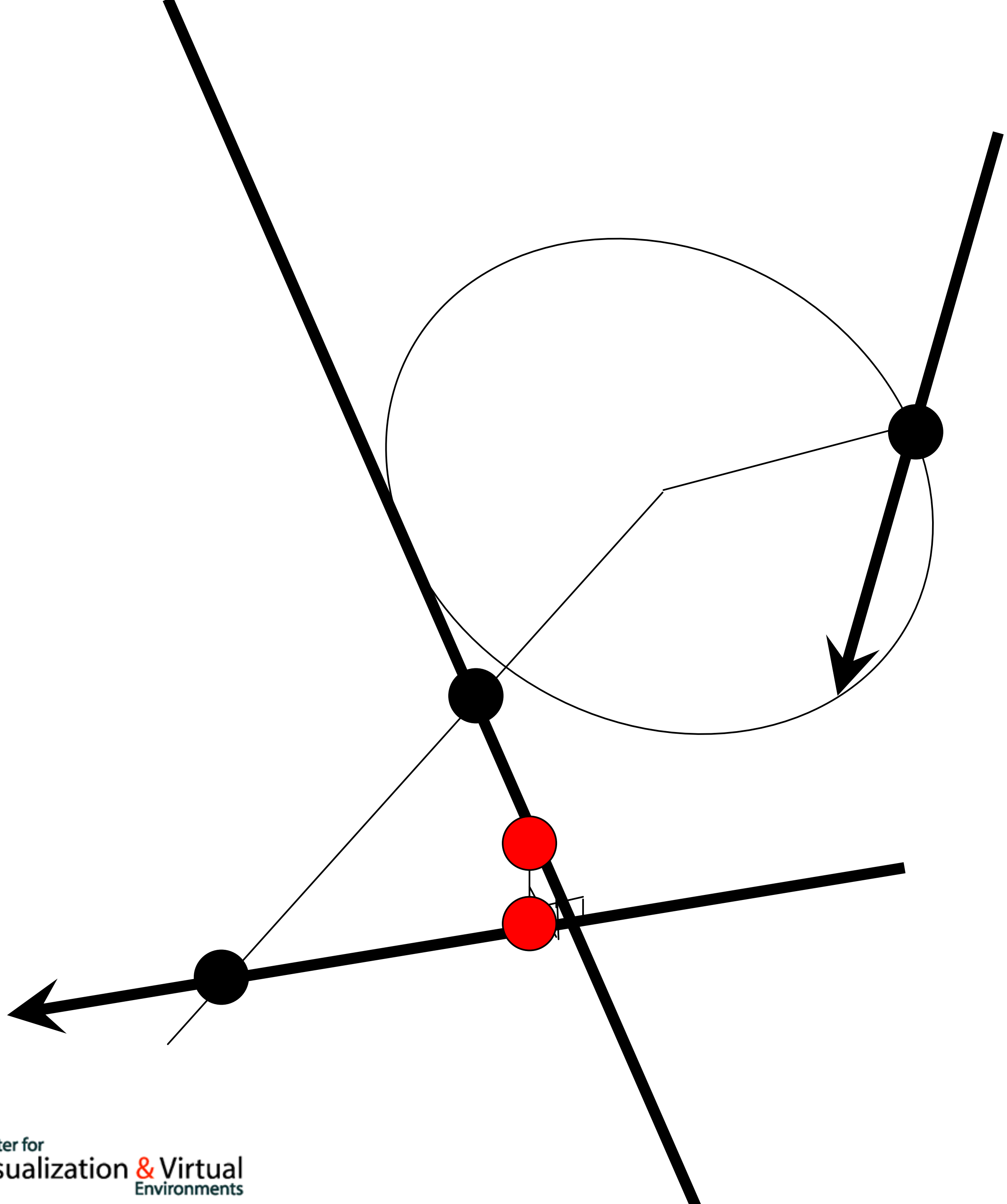


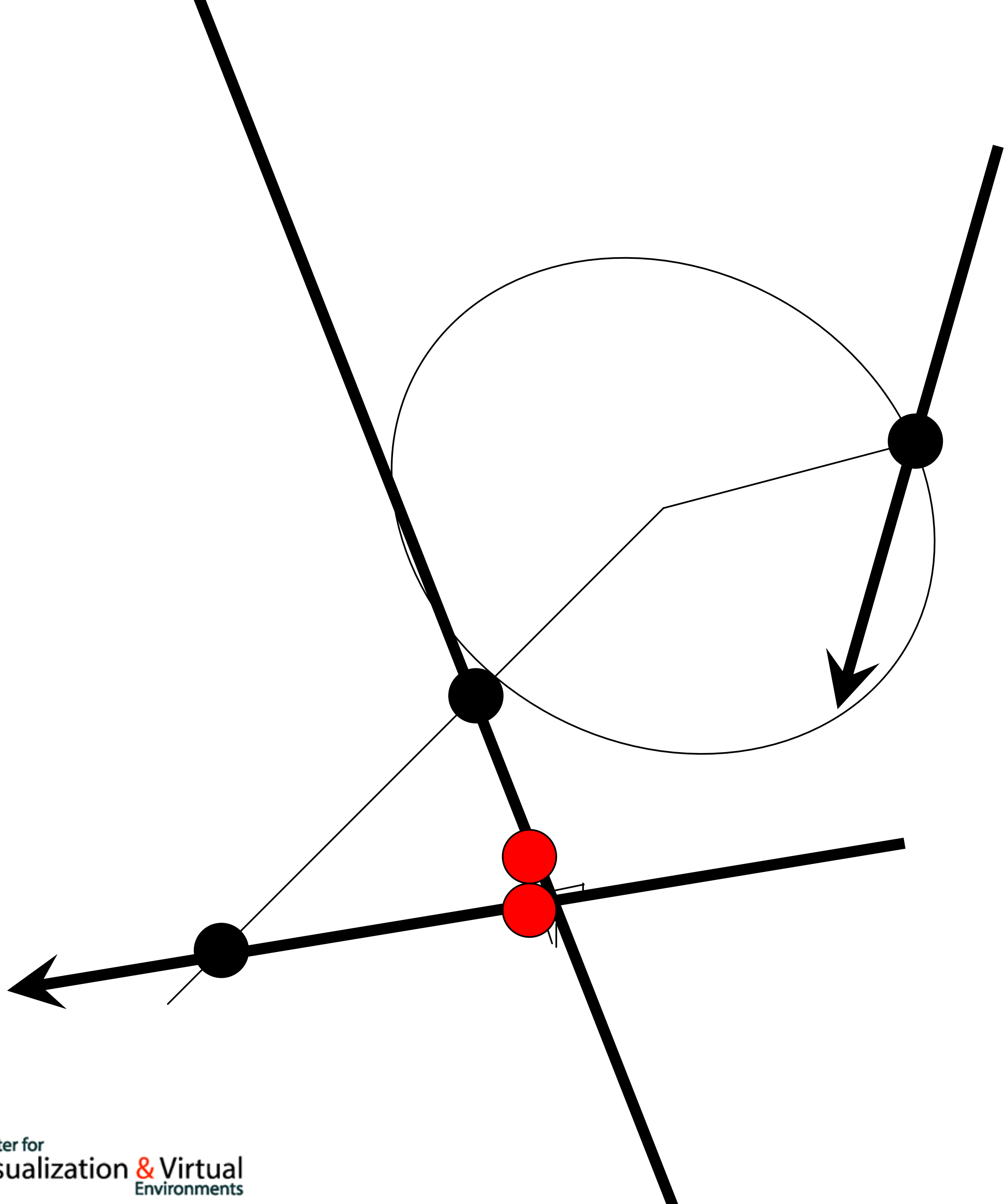
UK Center for Visualization & Virtual Environments



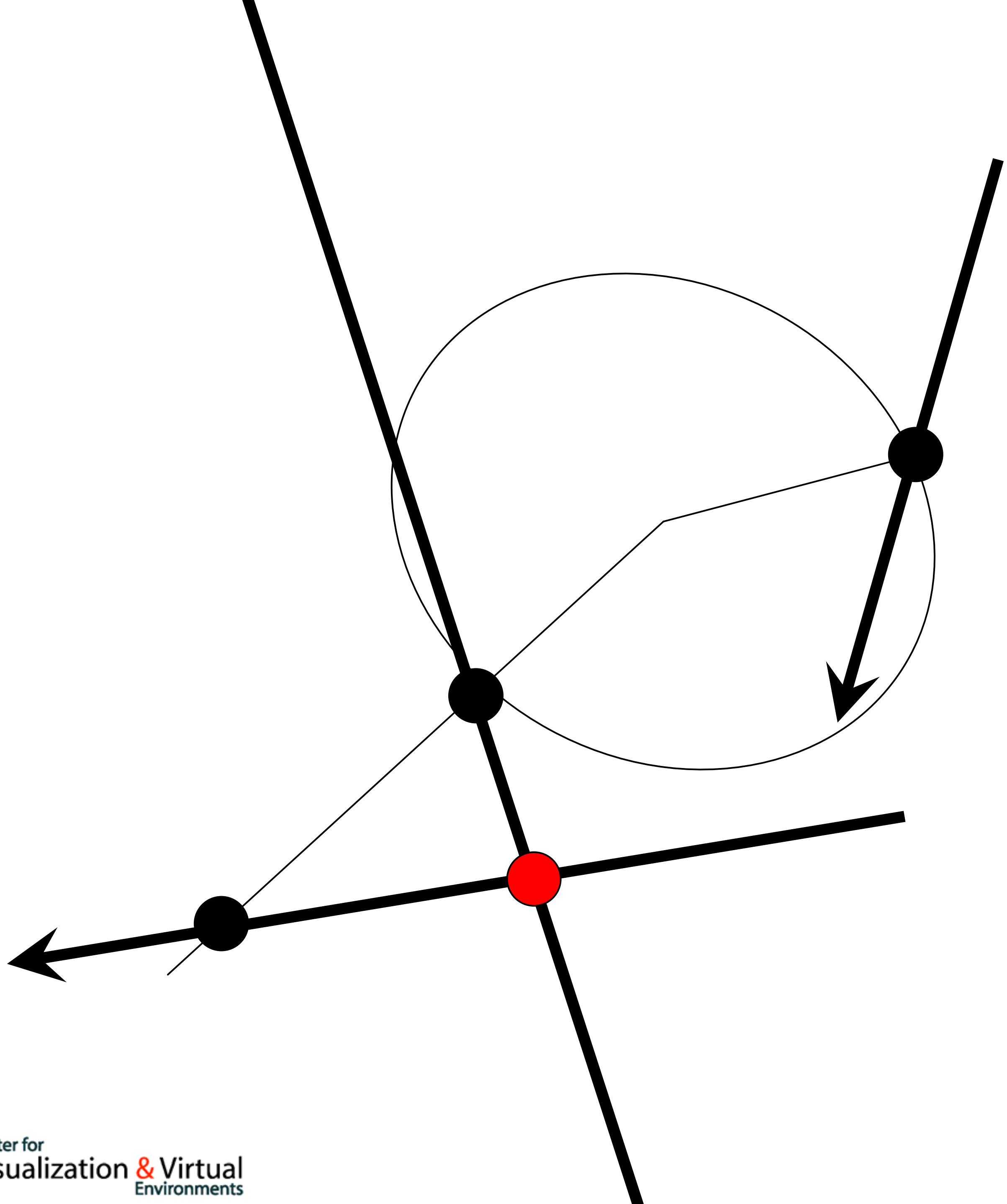


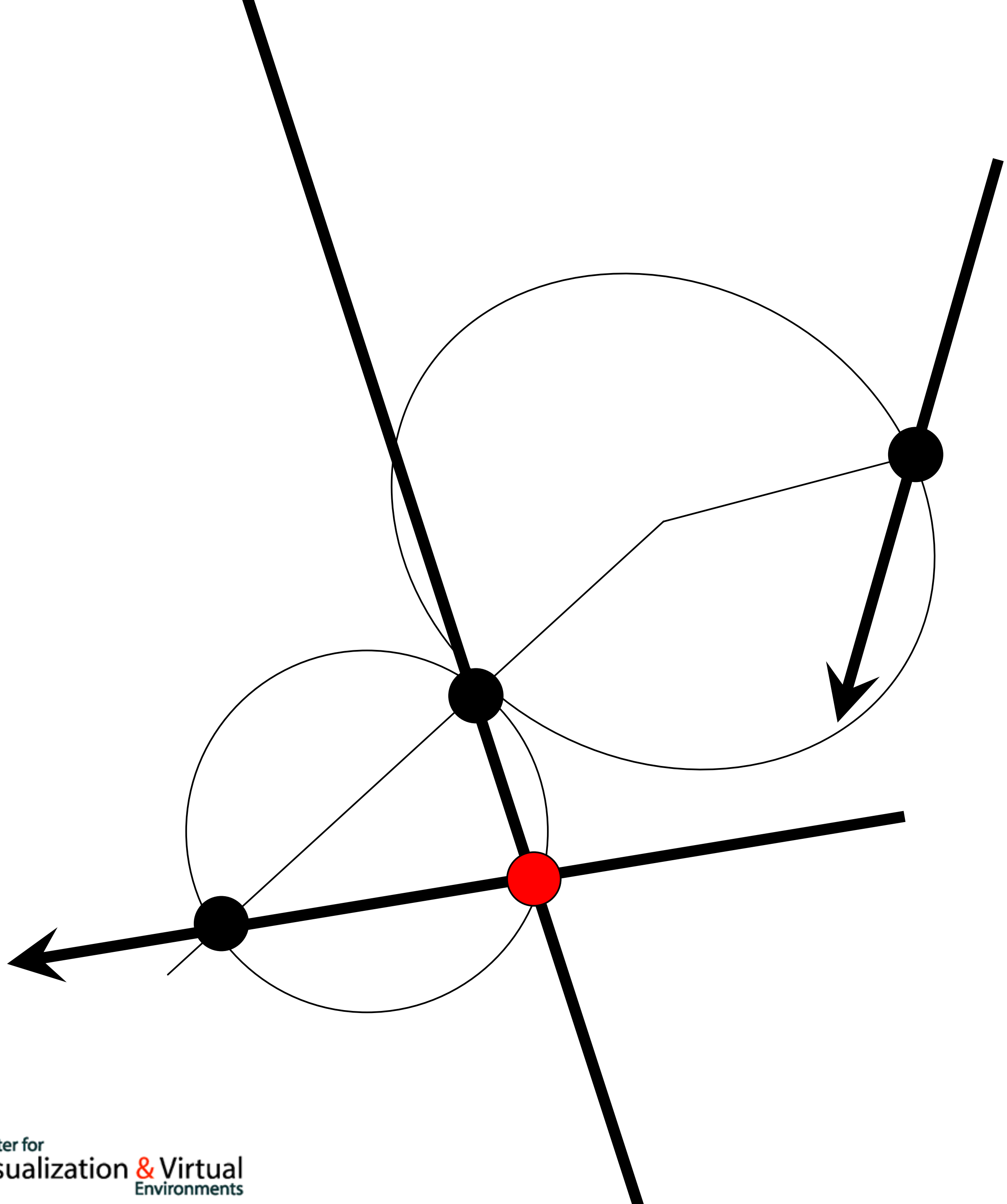


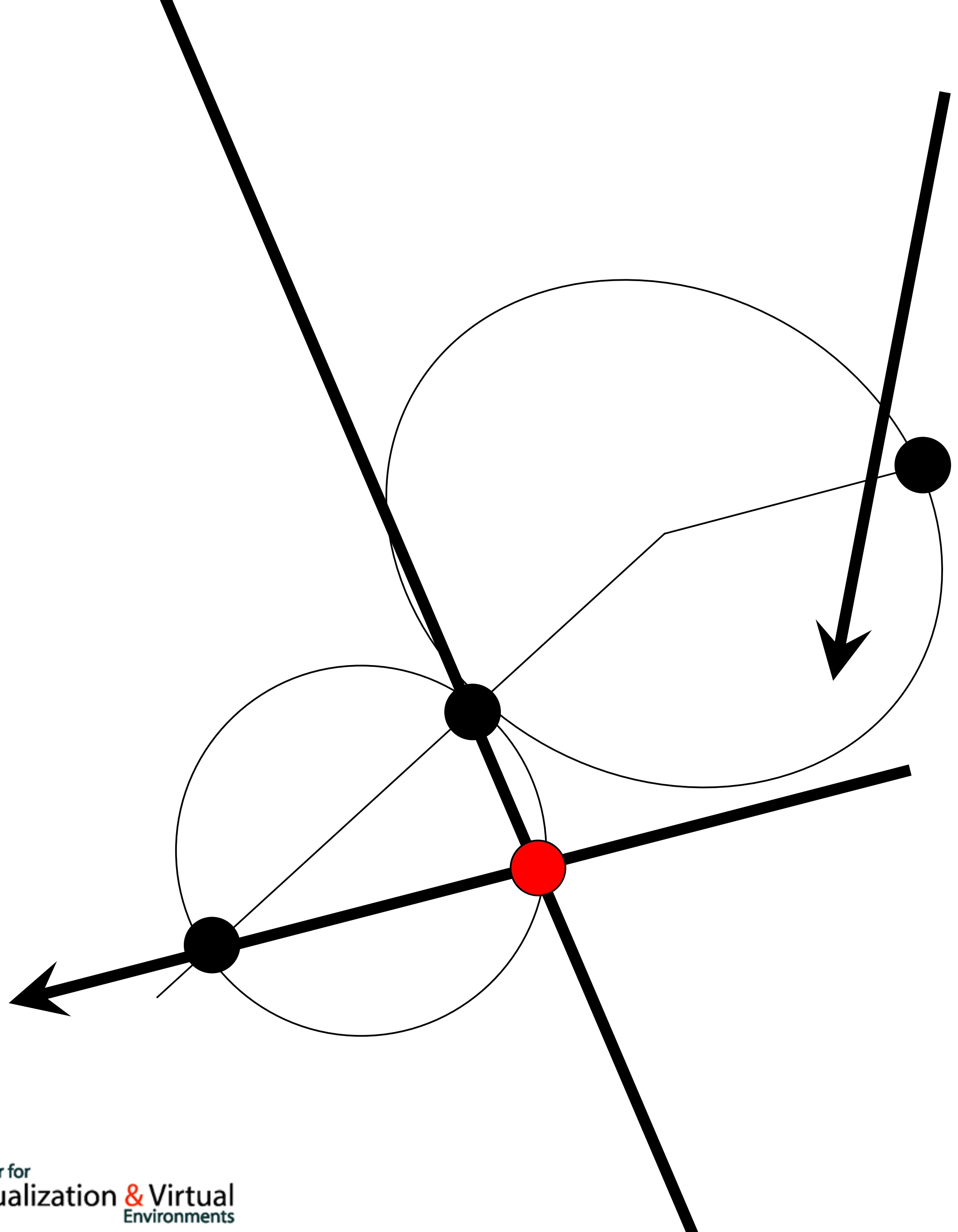


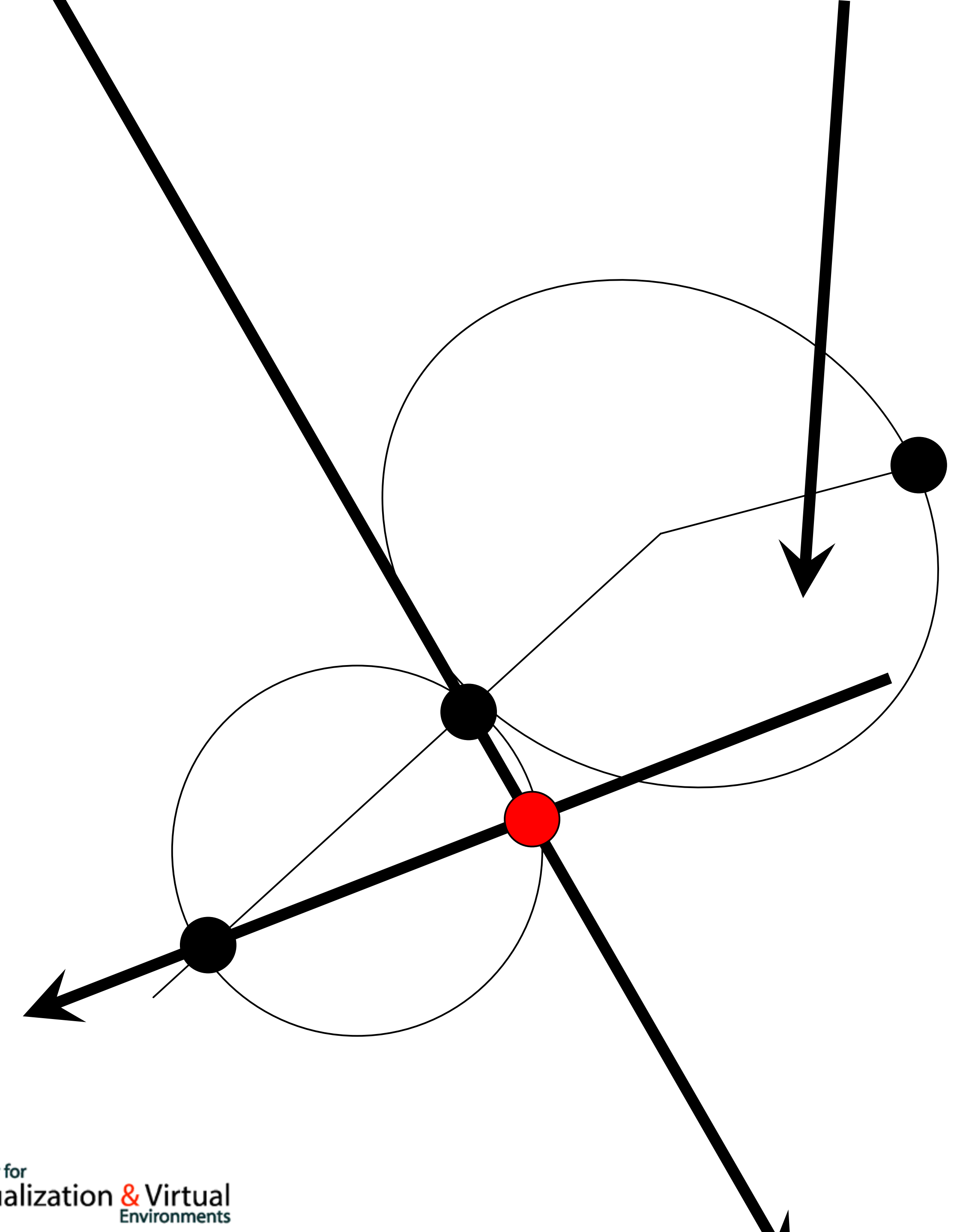


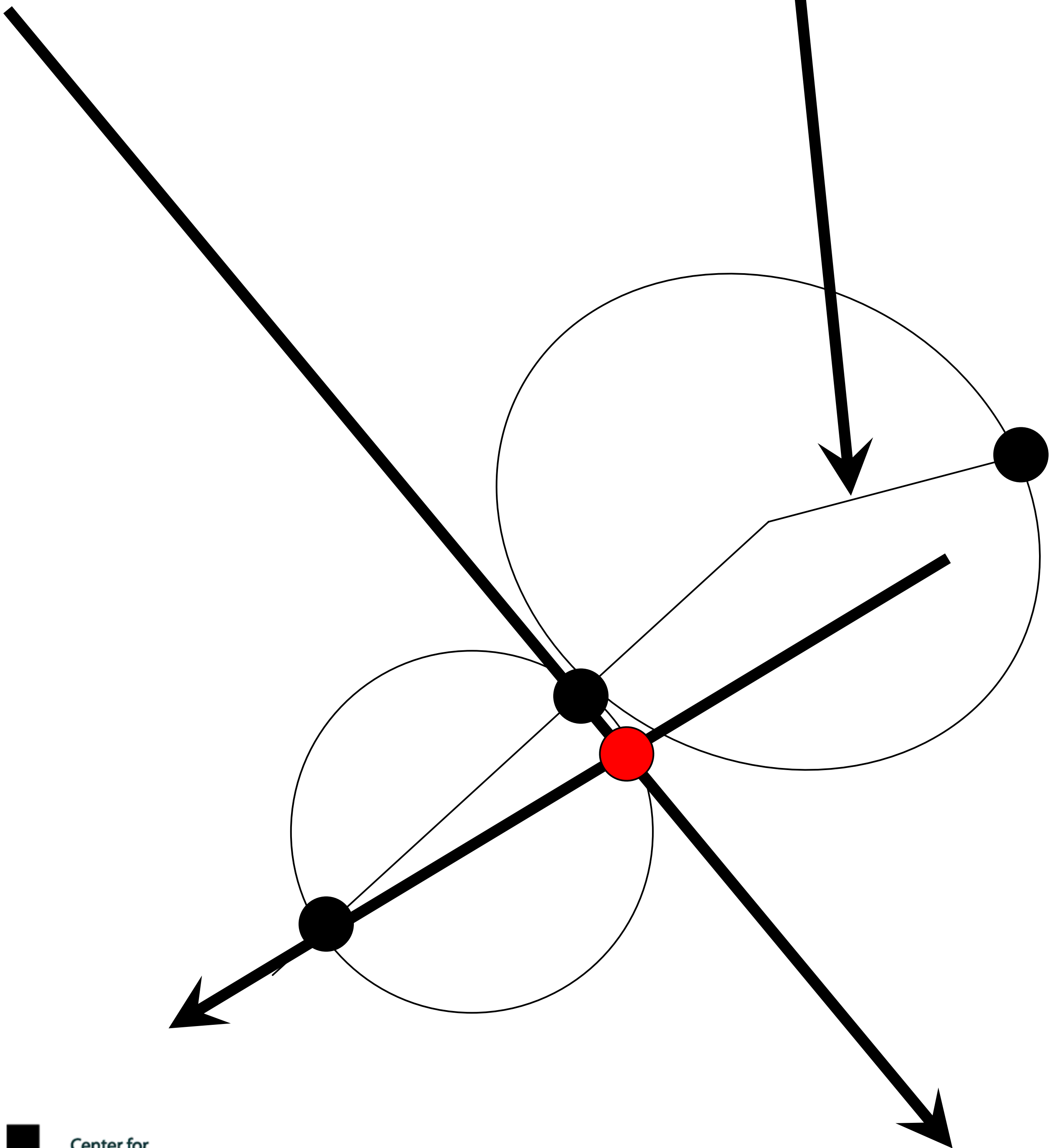
UK Center for Visualization & Virtual Environments





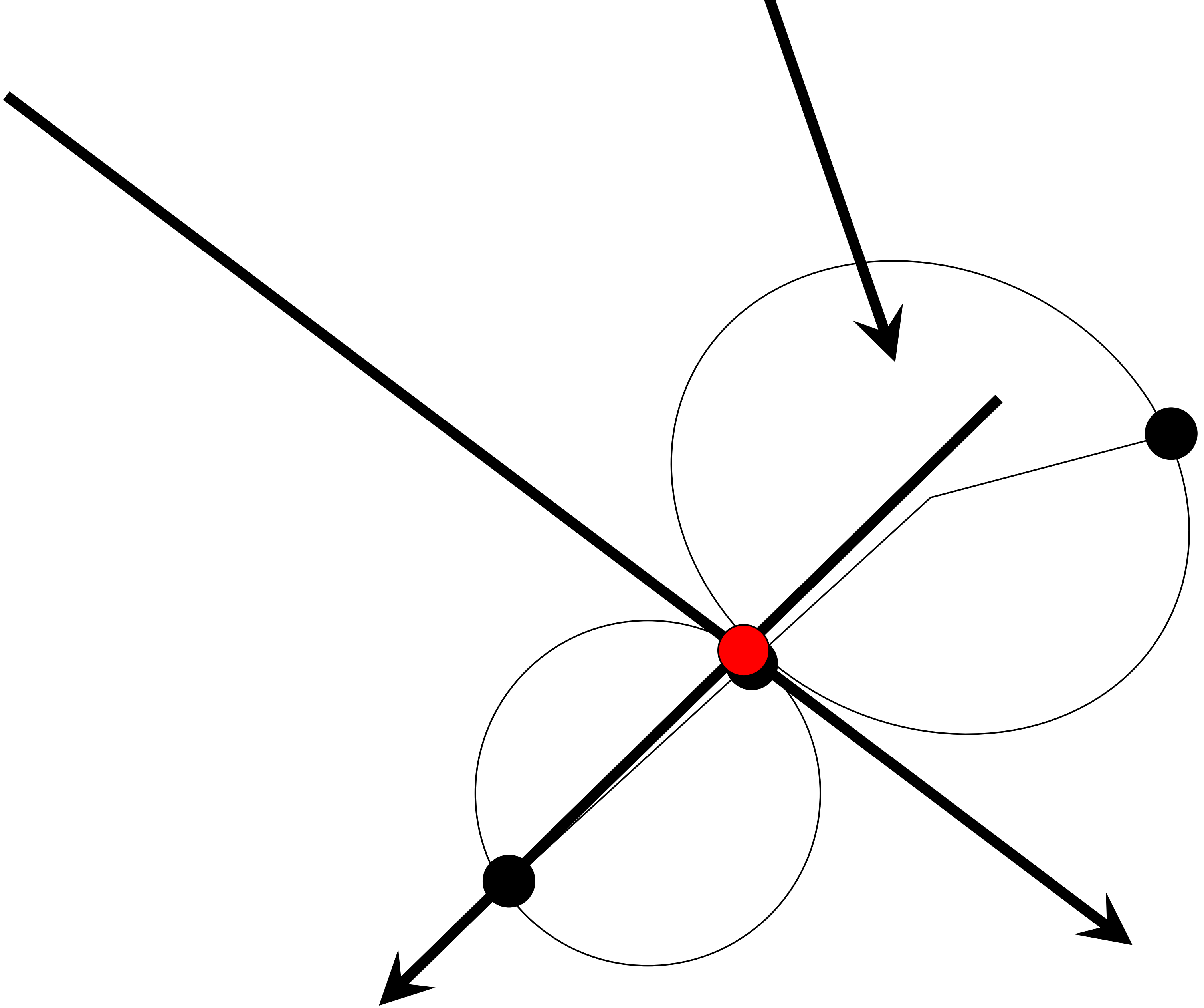


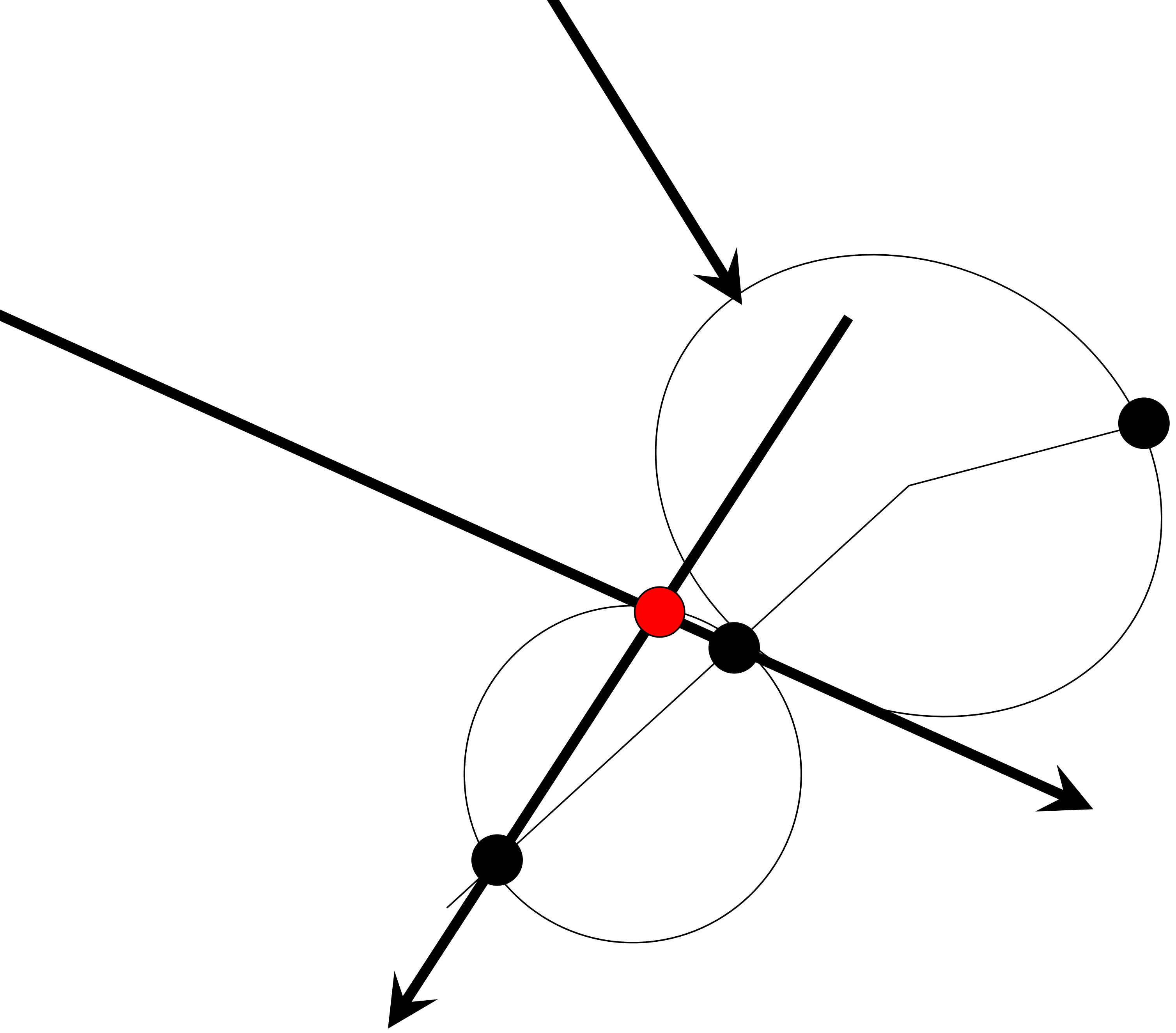


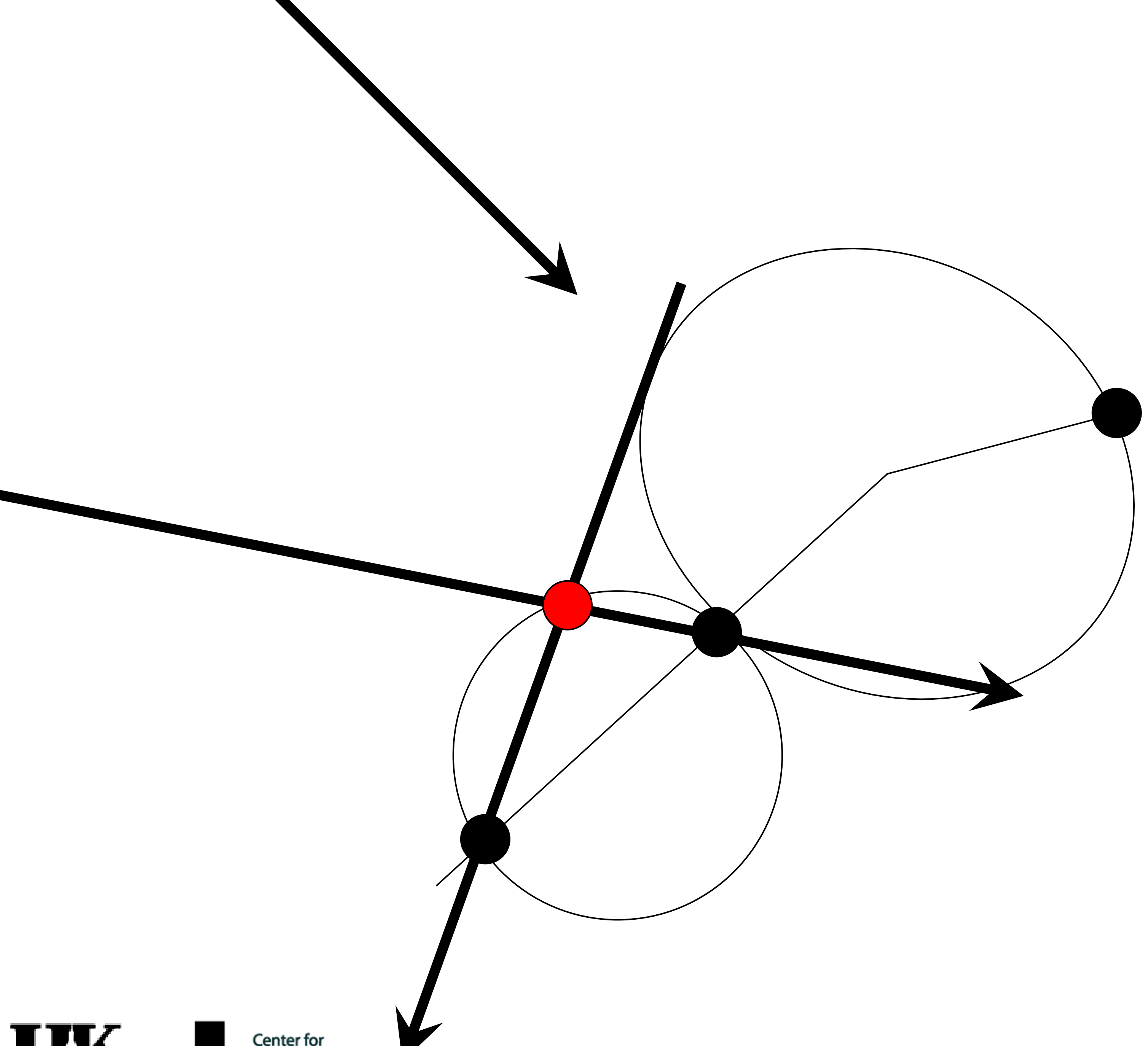


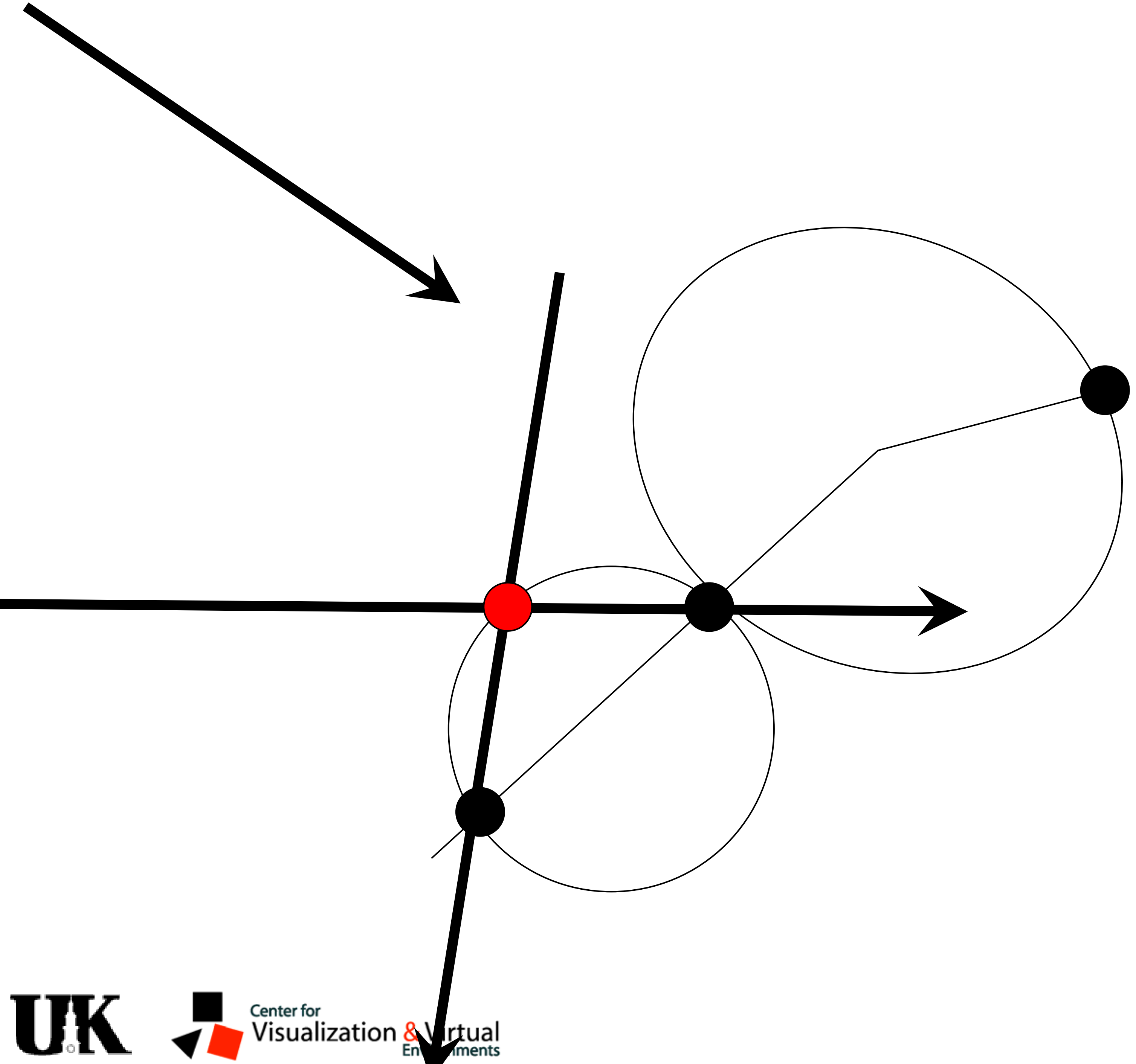
UK

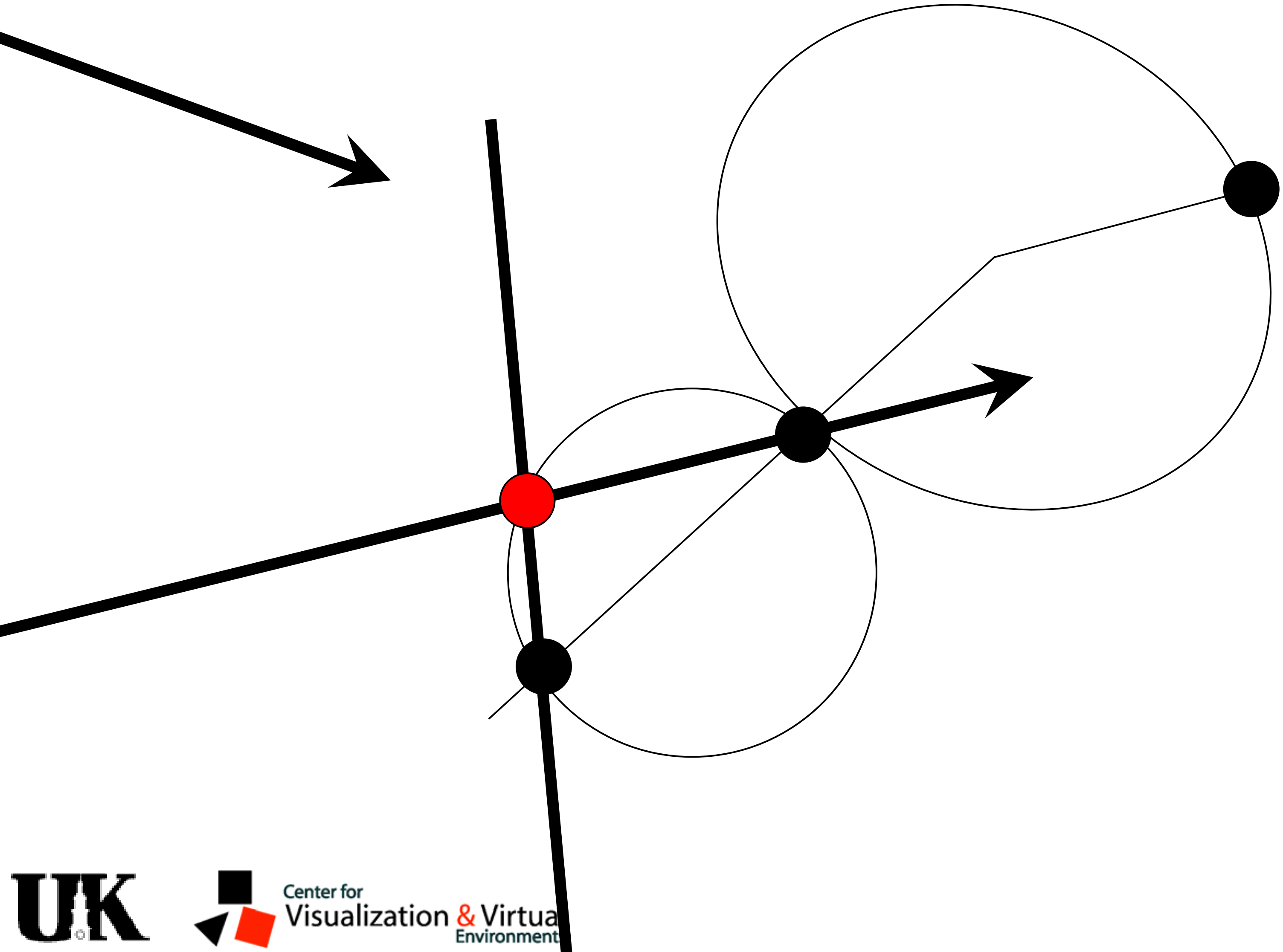
Center for Visualization & Virtual Environments

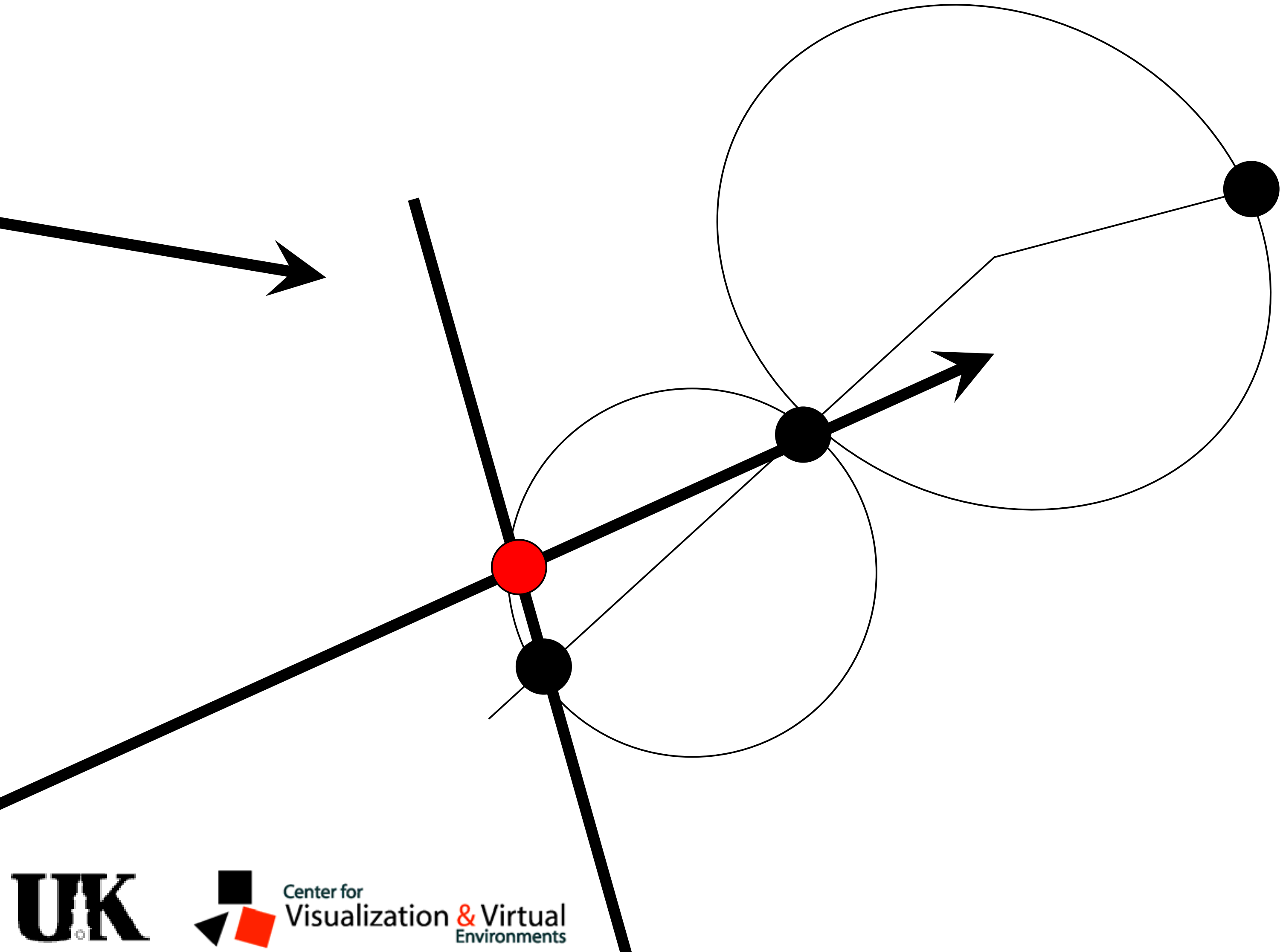


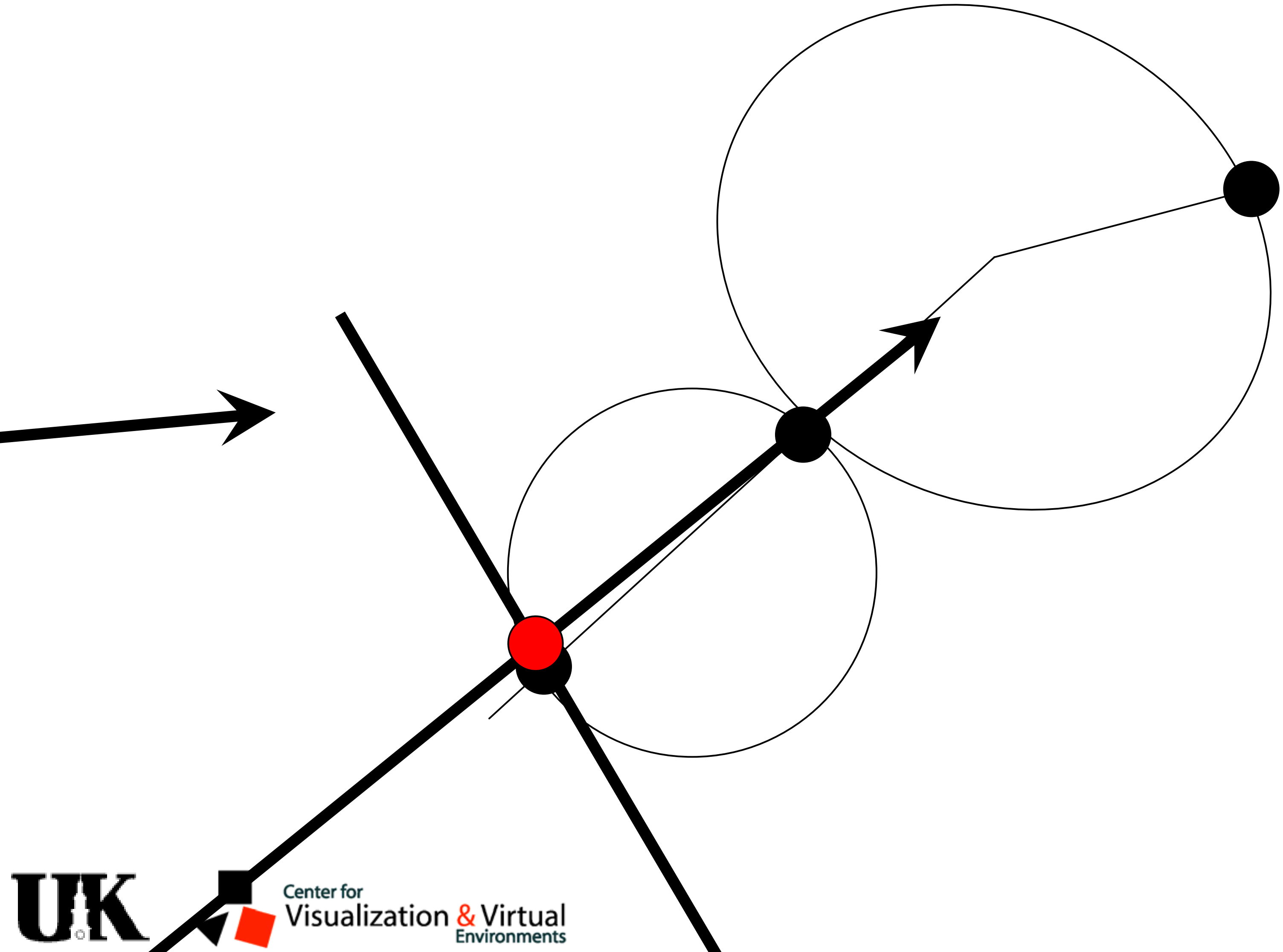




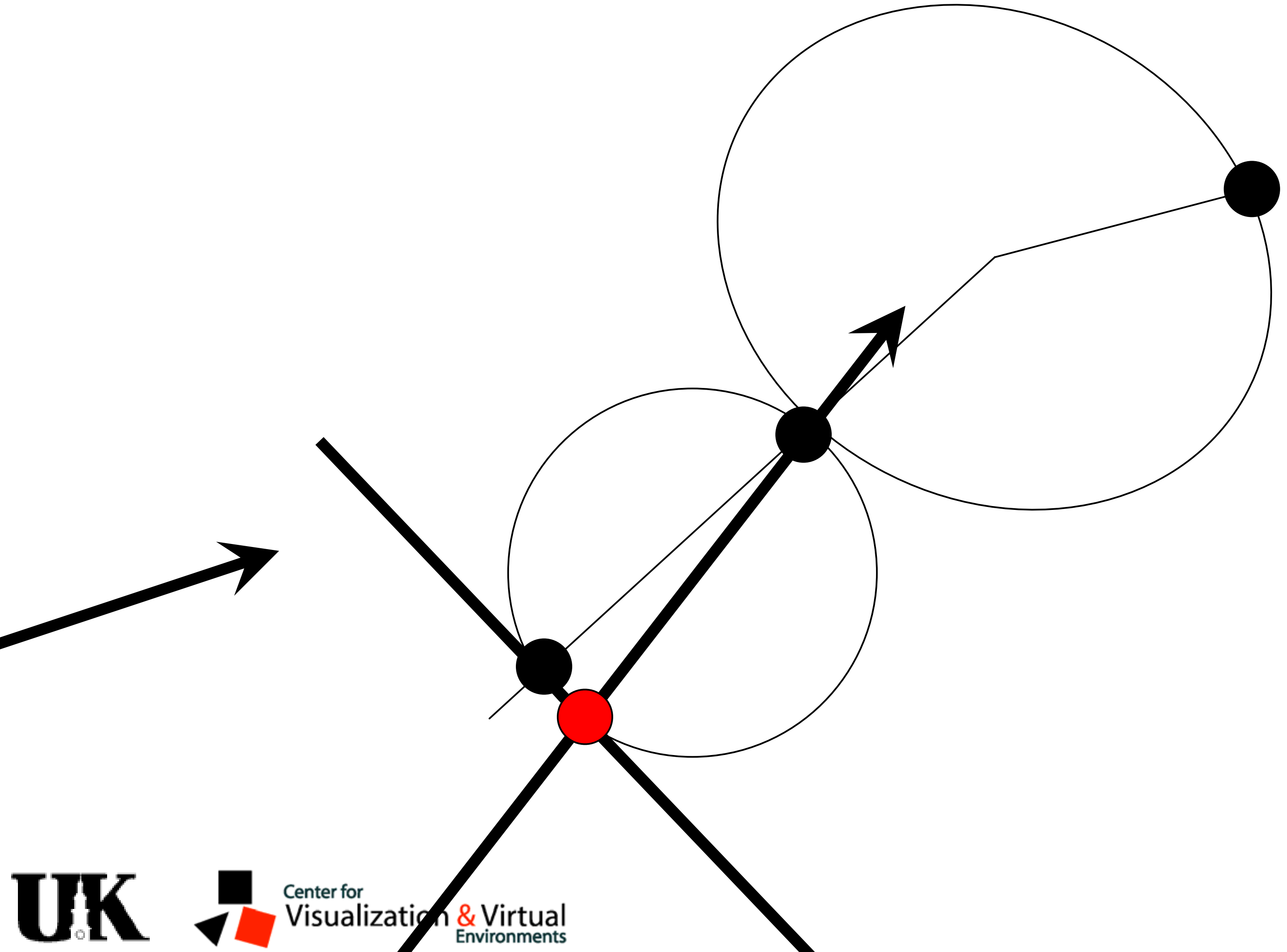


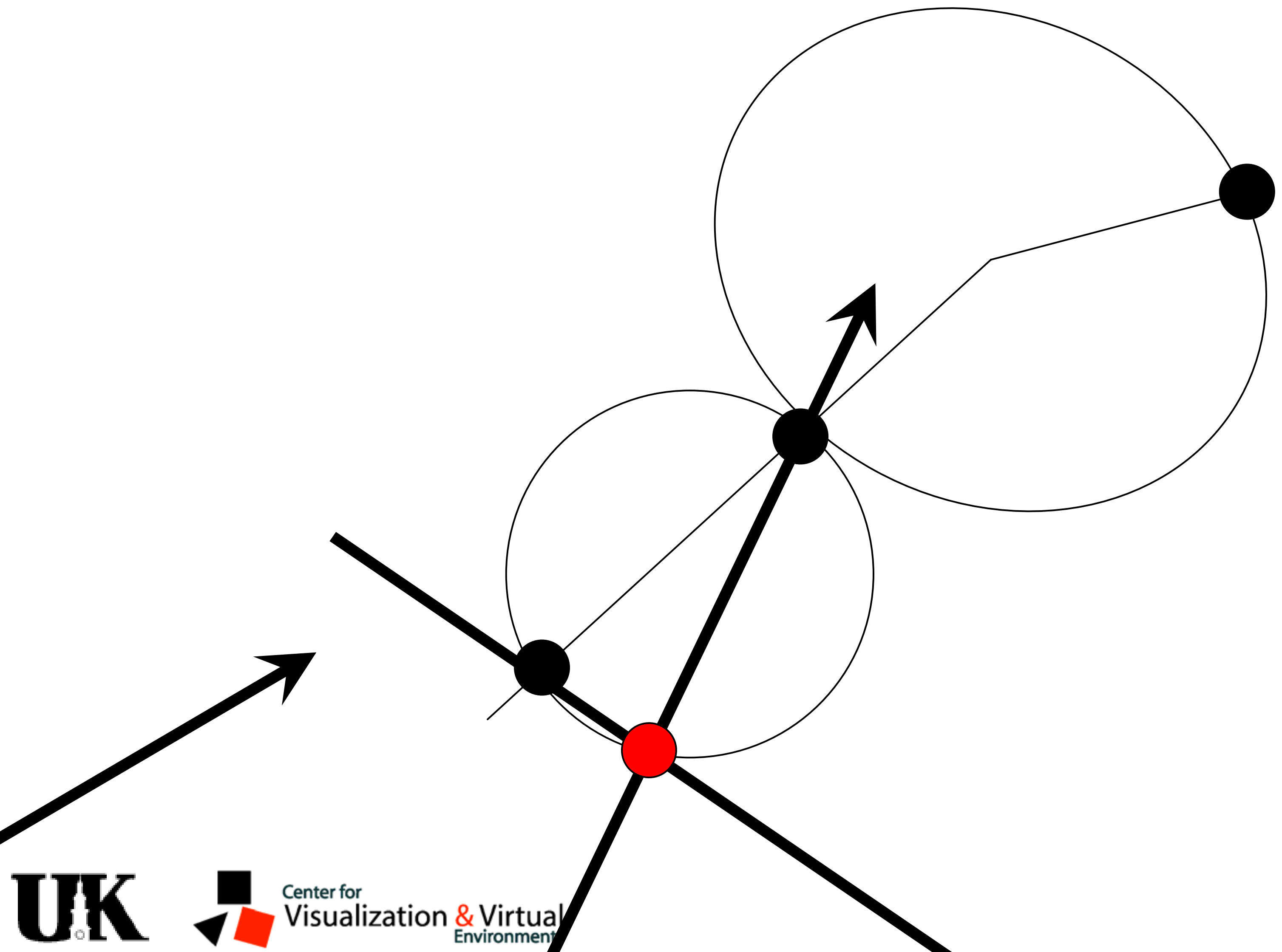






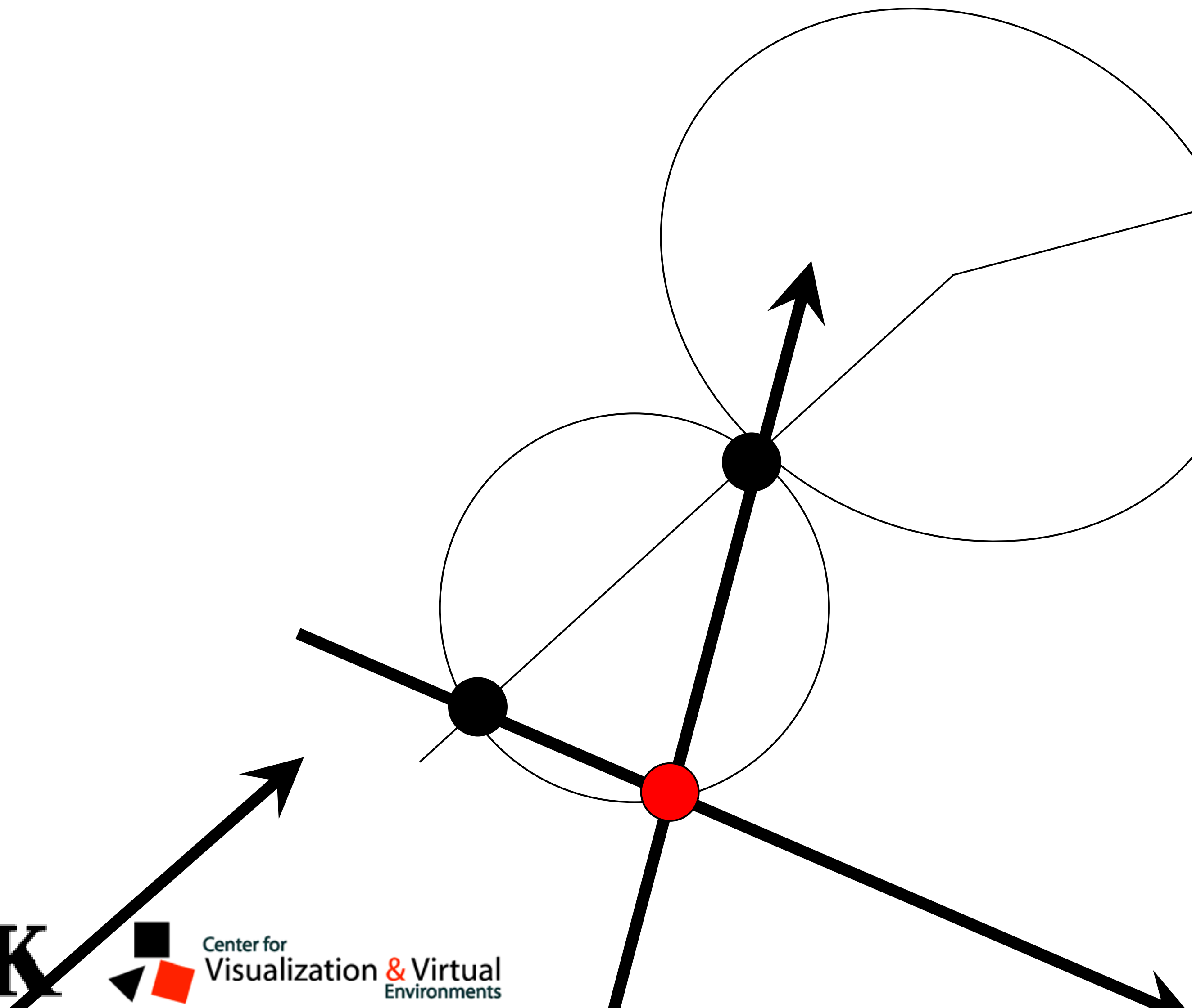
Center for Visualization & Virtual Environments

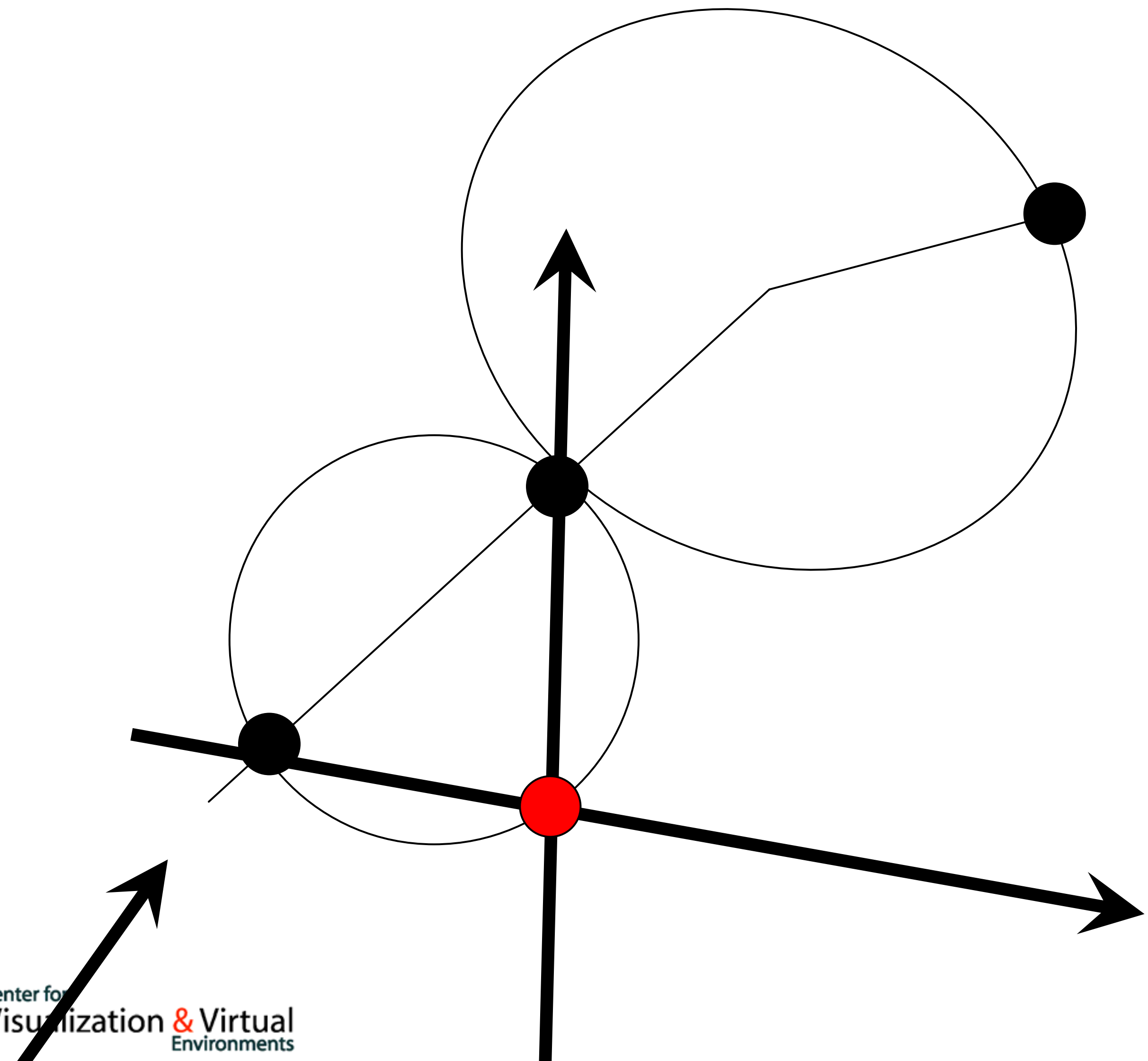




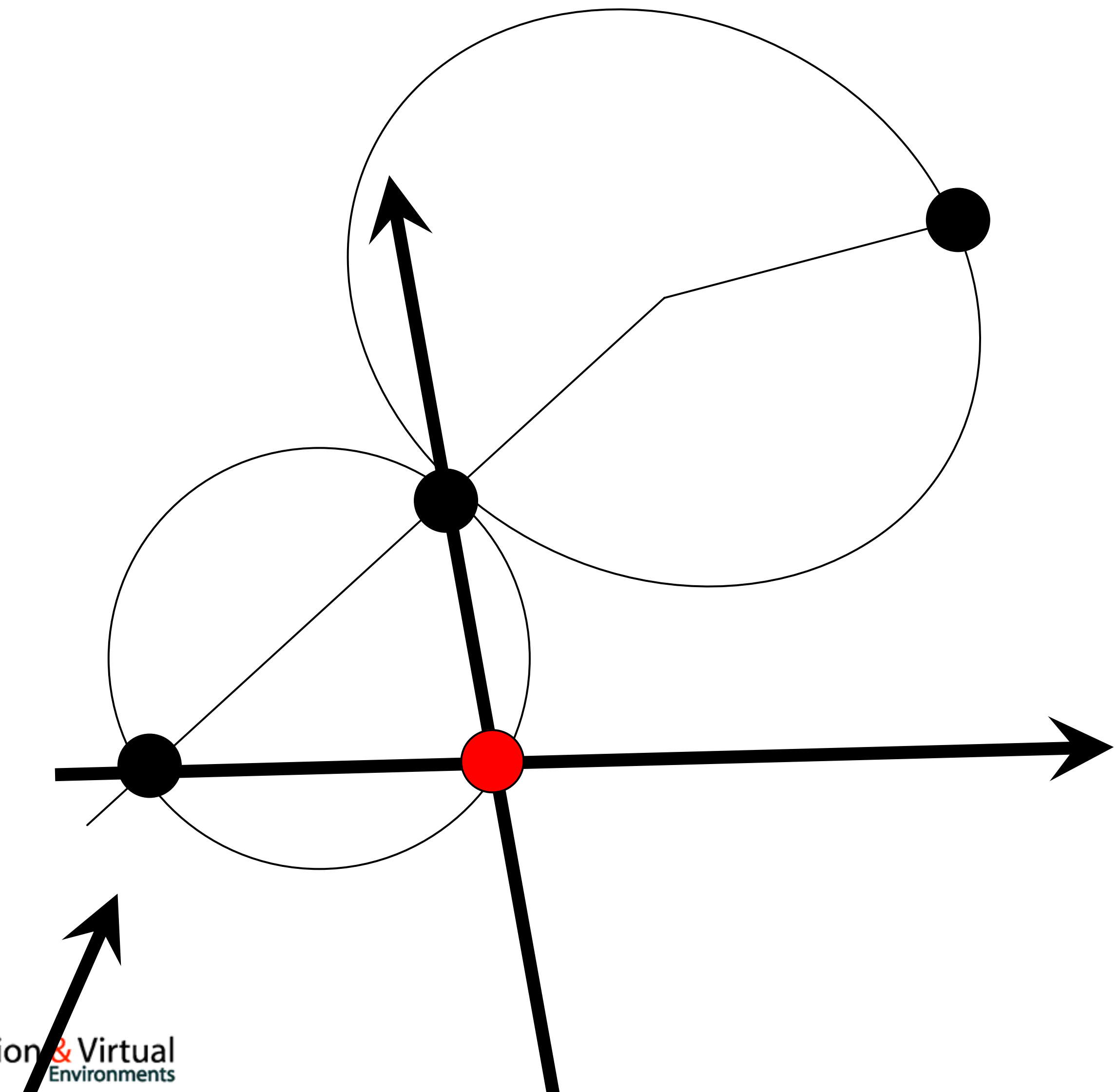
UK

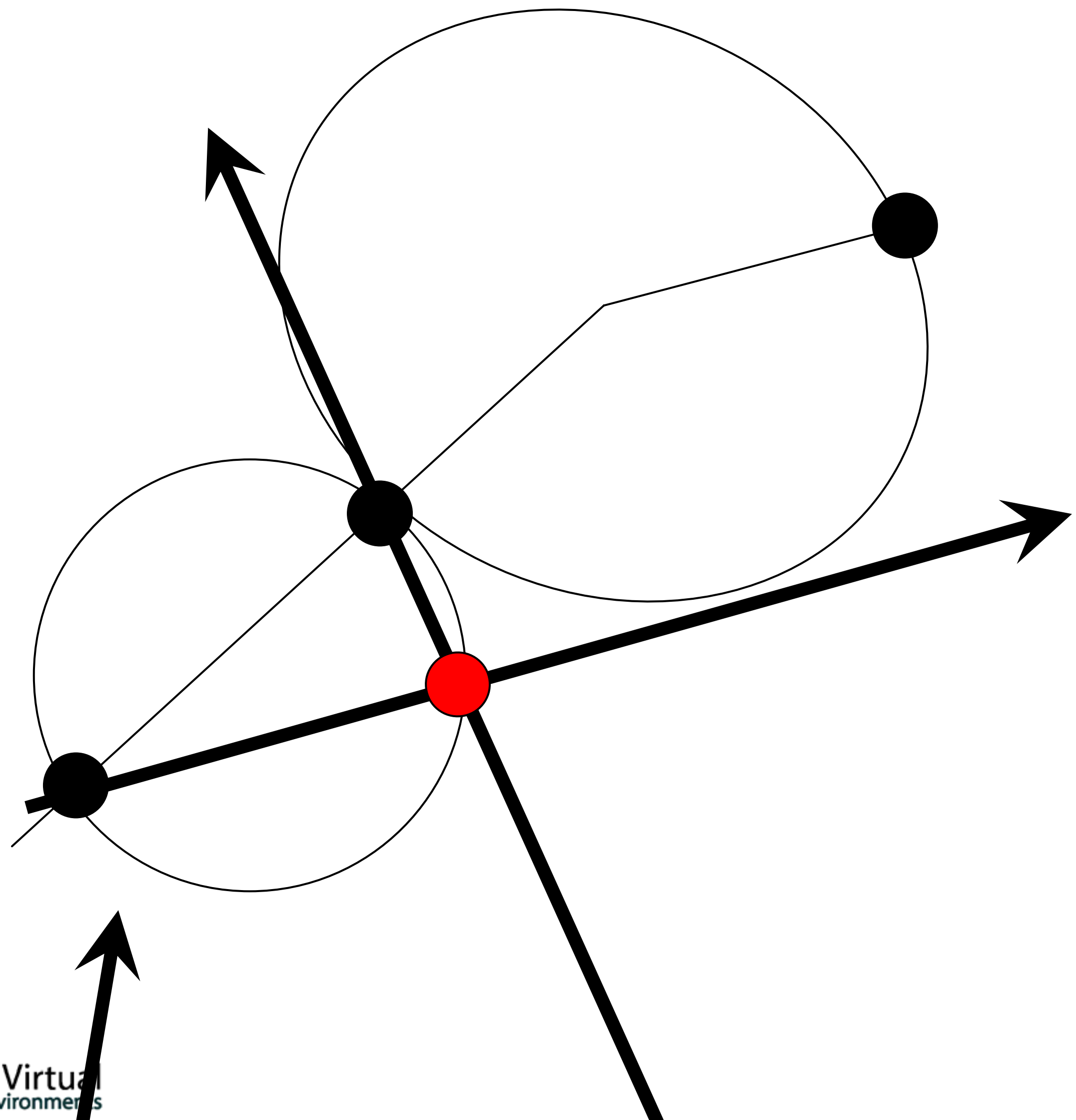
Center for Visualization & Virtual Environments

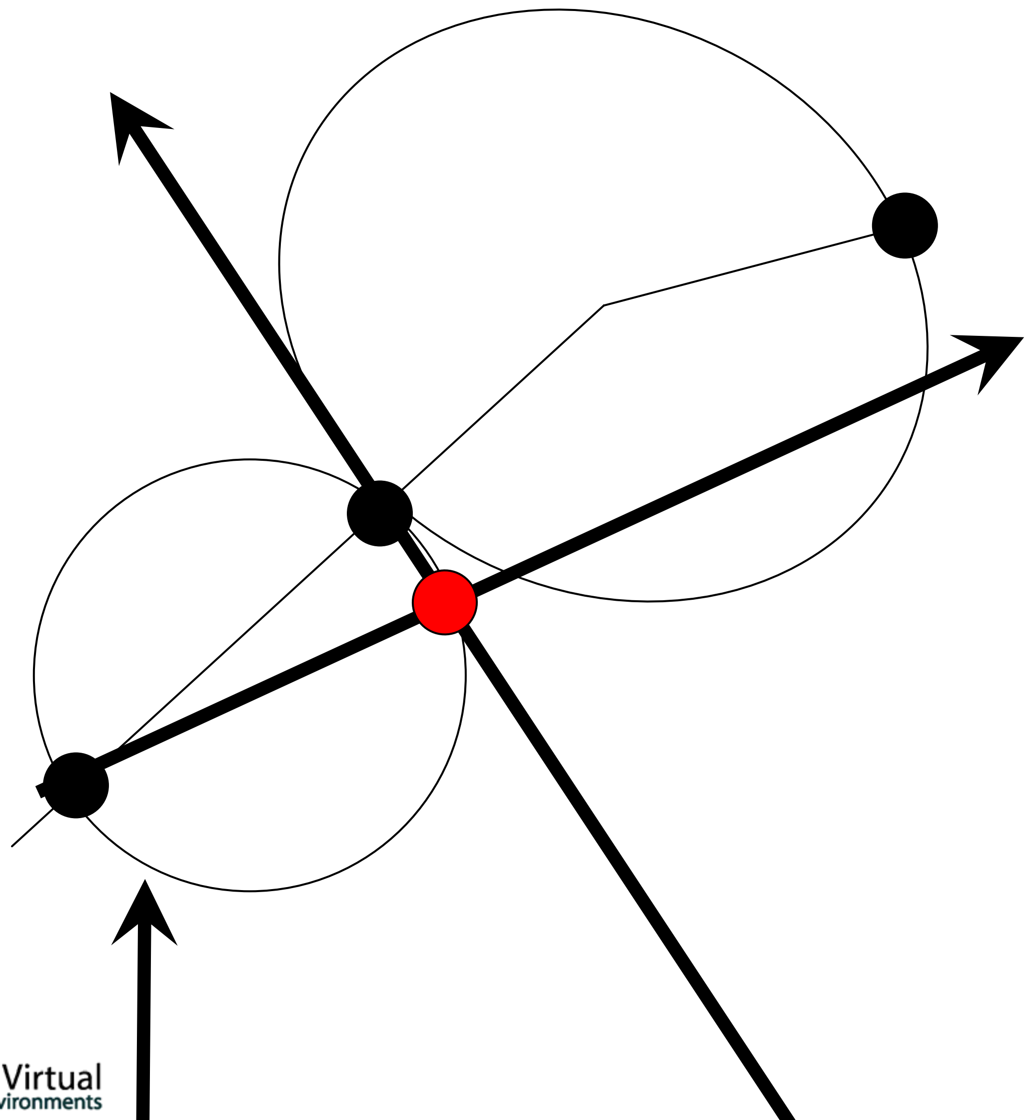




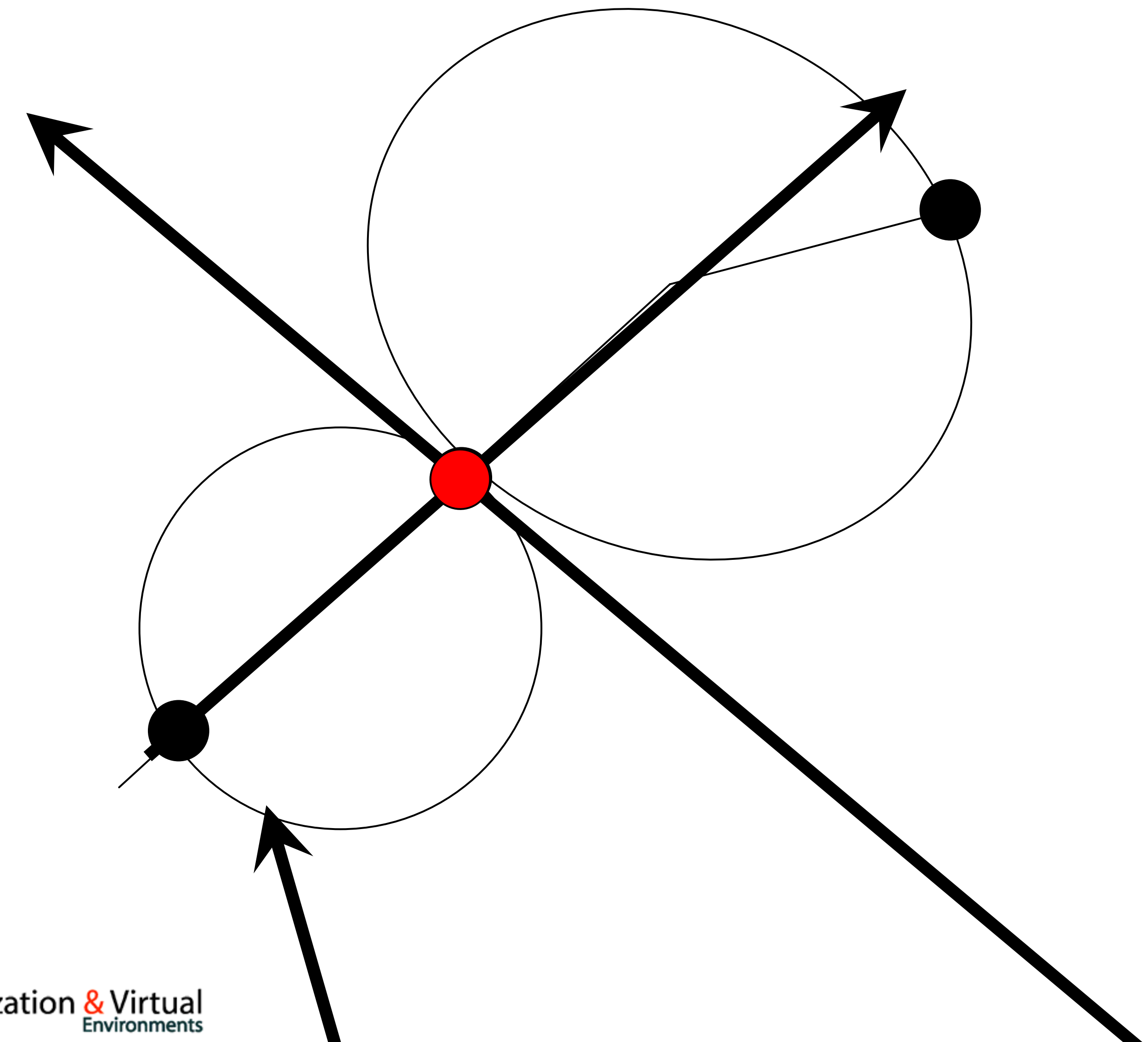
UK Center for Visualization & Virtual Environments





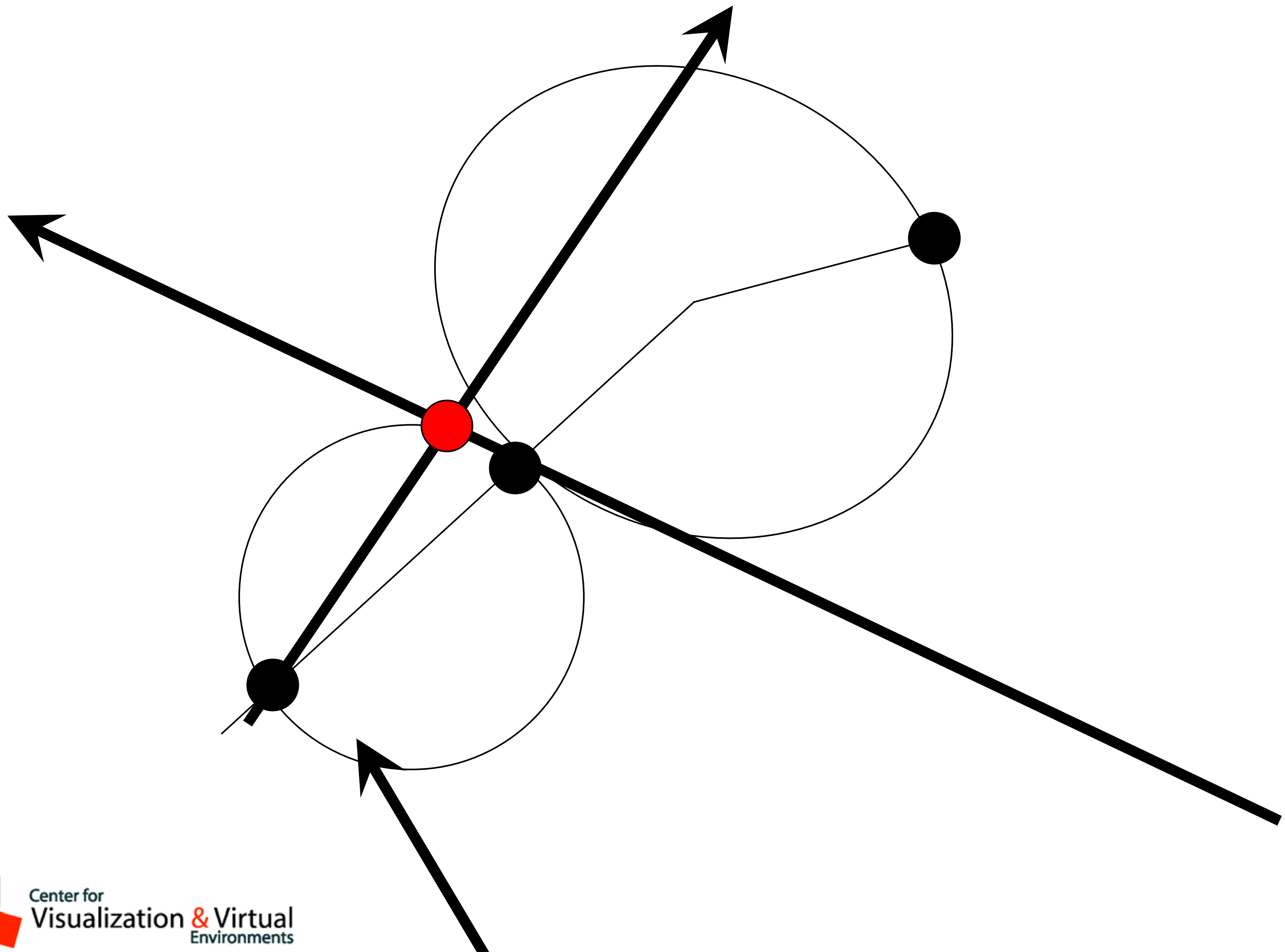


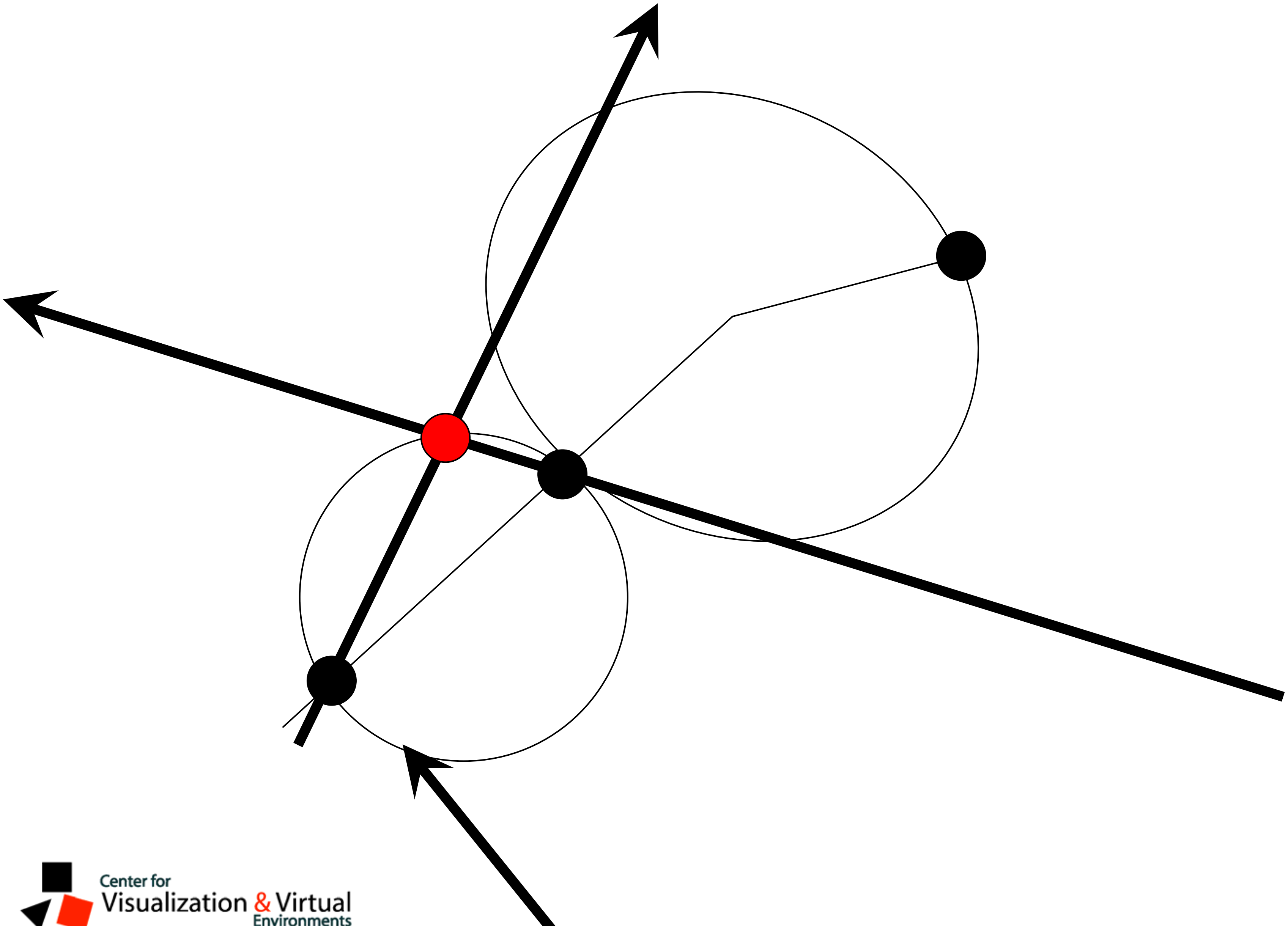
UK Center for Visualization & Virtual Environments



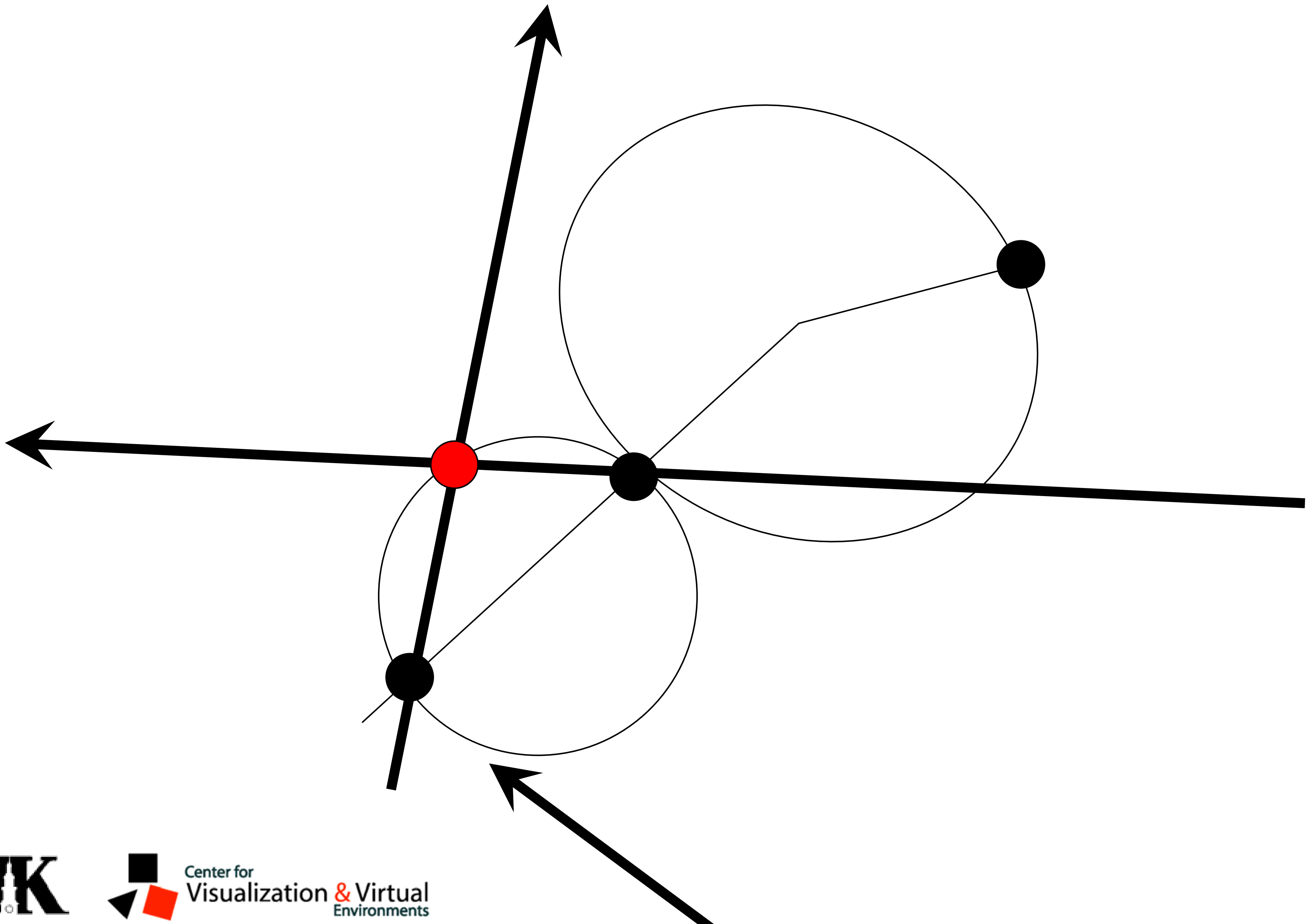
UK

Center for Visualization & Virtual Environments

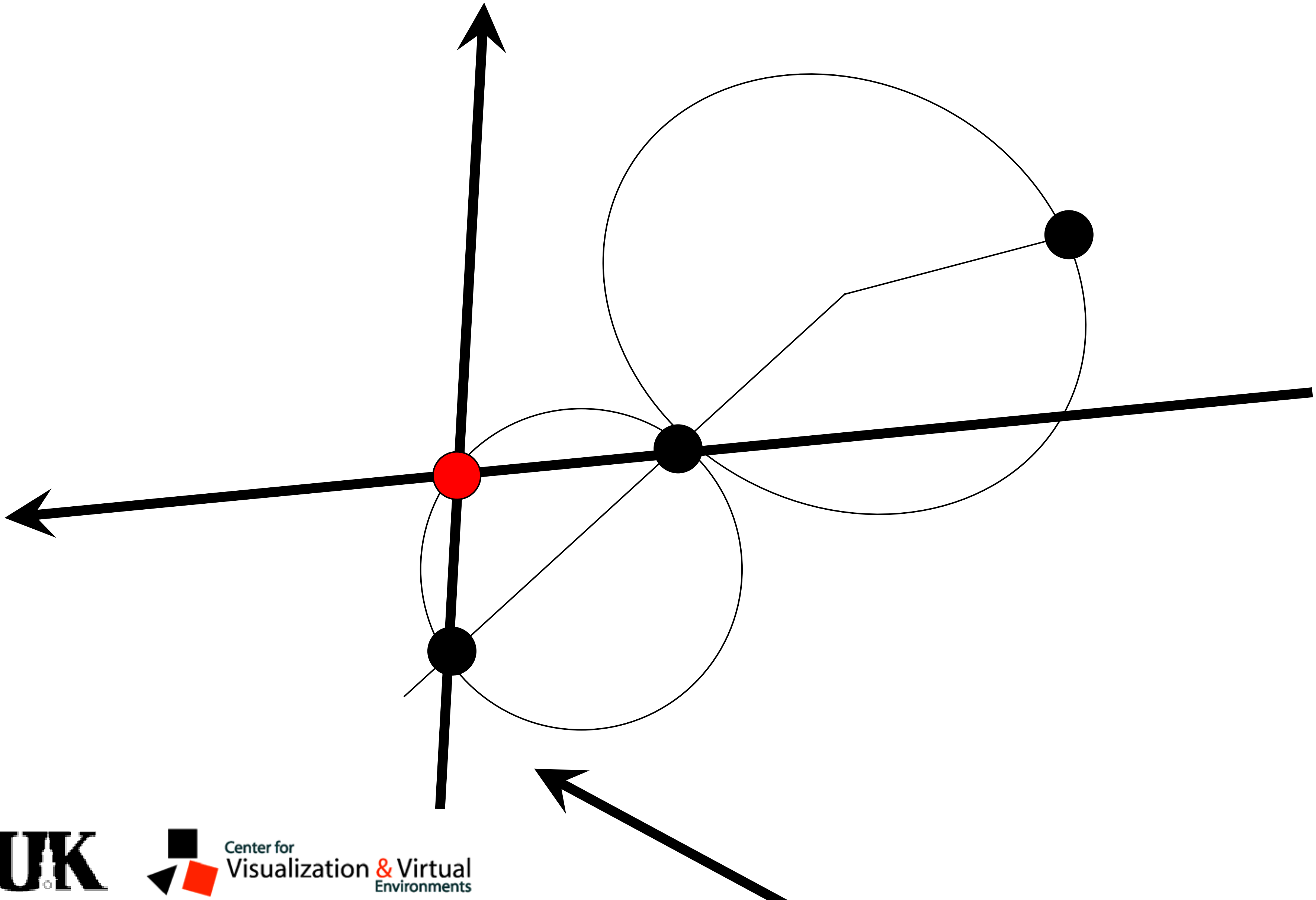


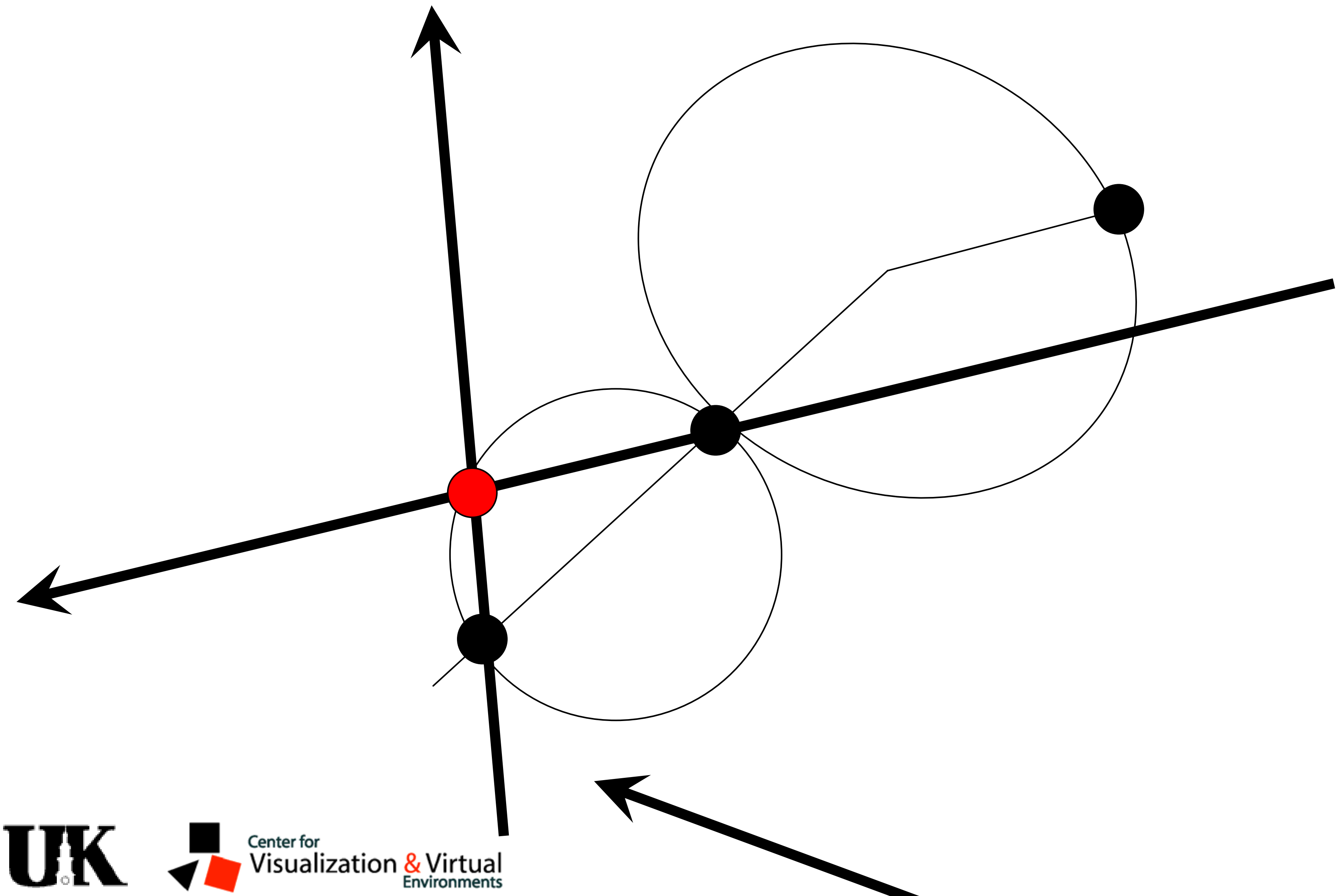


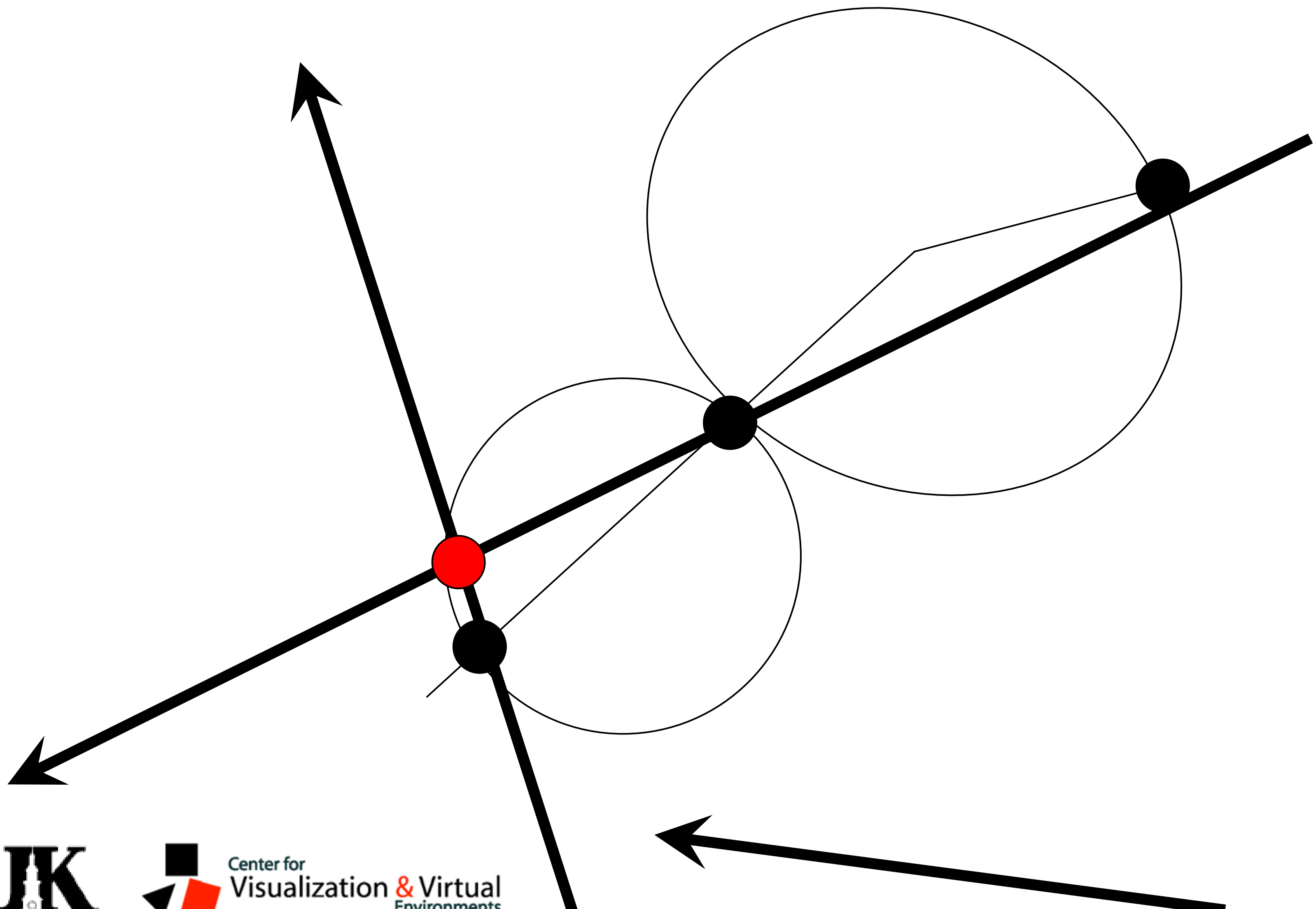
UK Center for Visualization & Virtual Environments



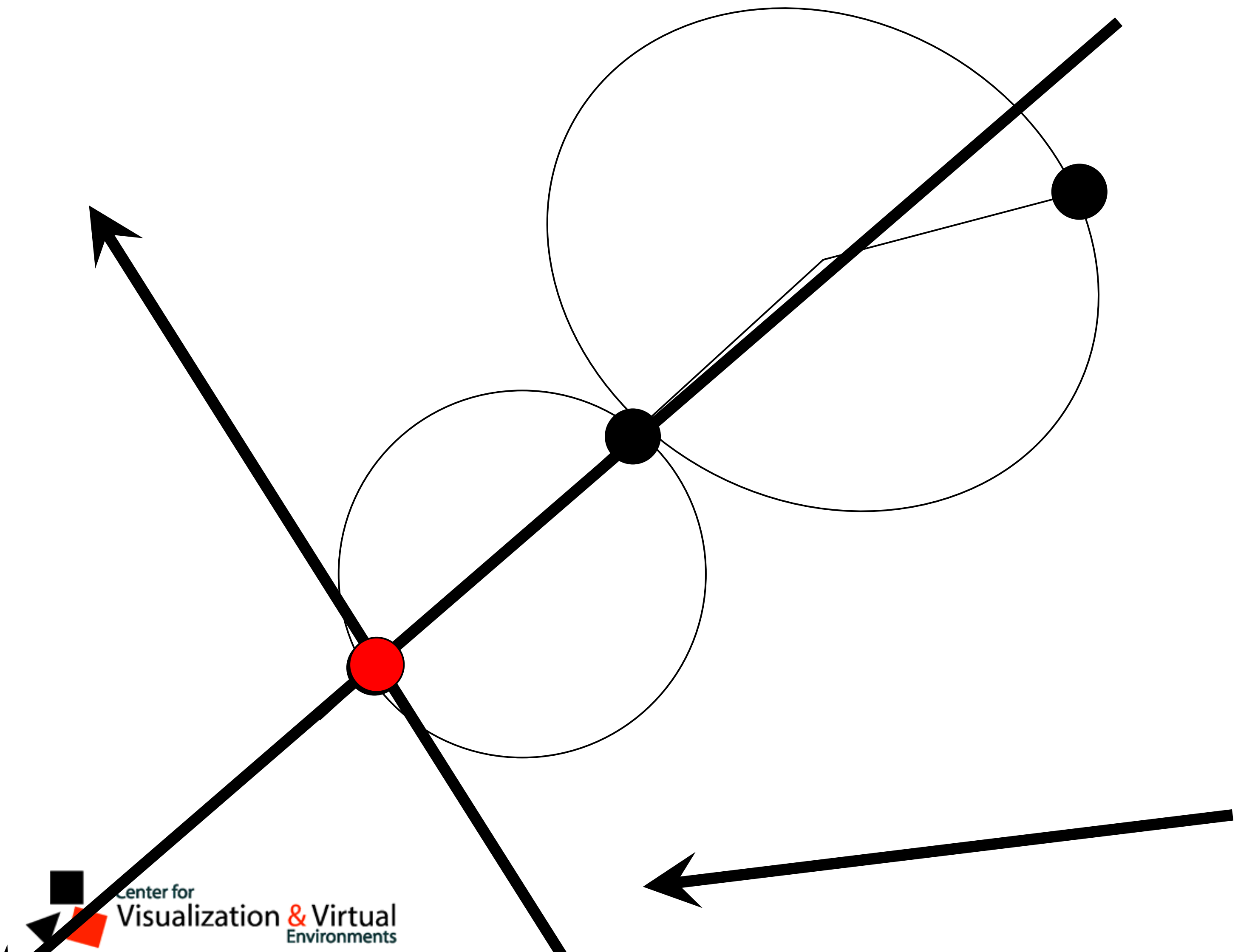
UK Center for Visualization & Virtual Environments



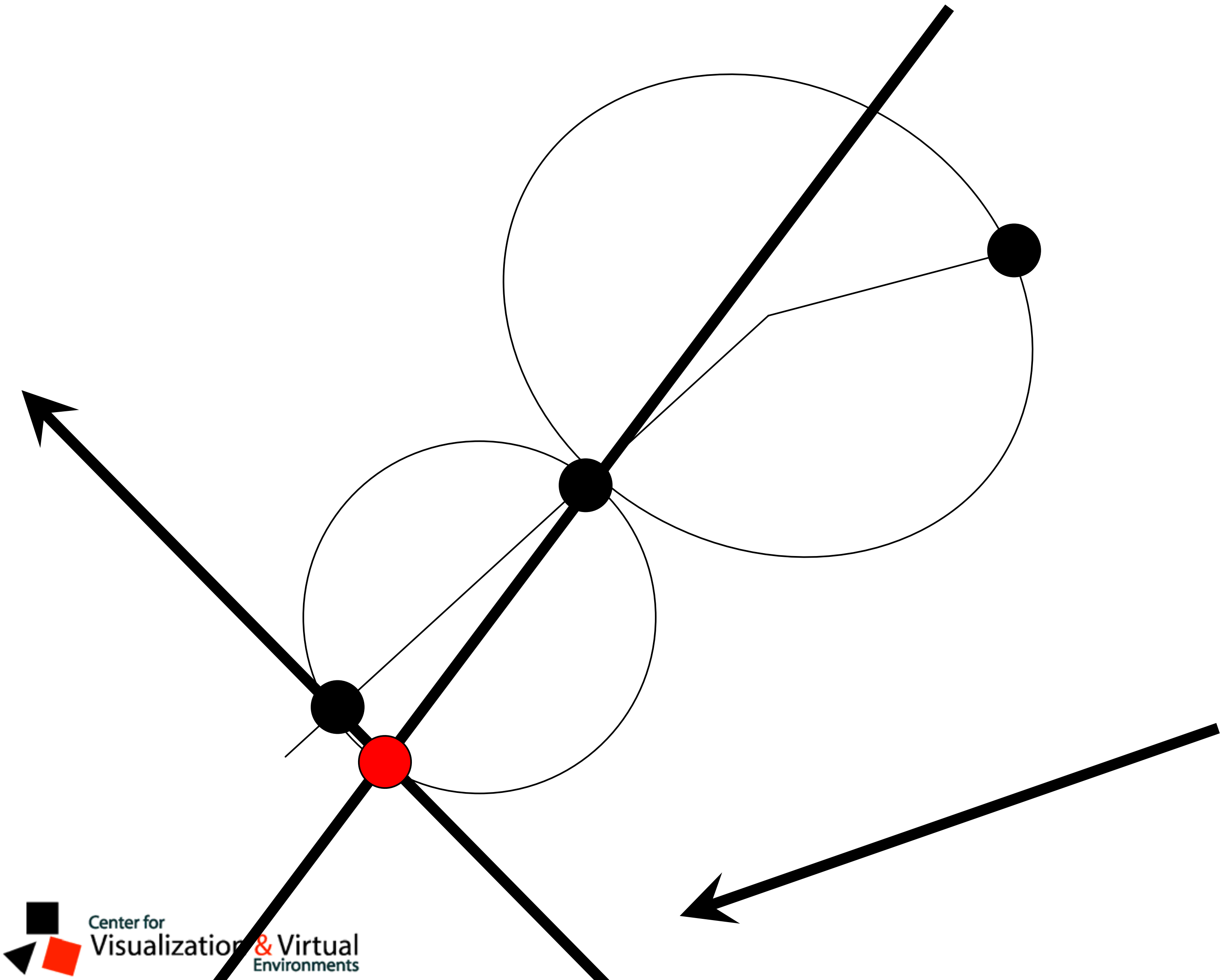


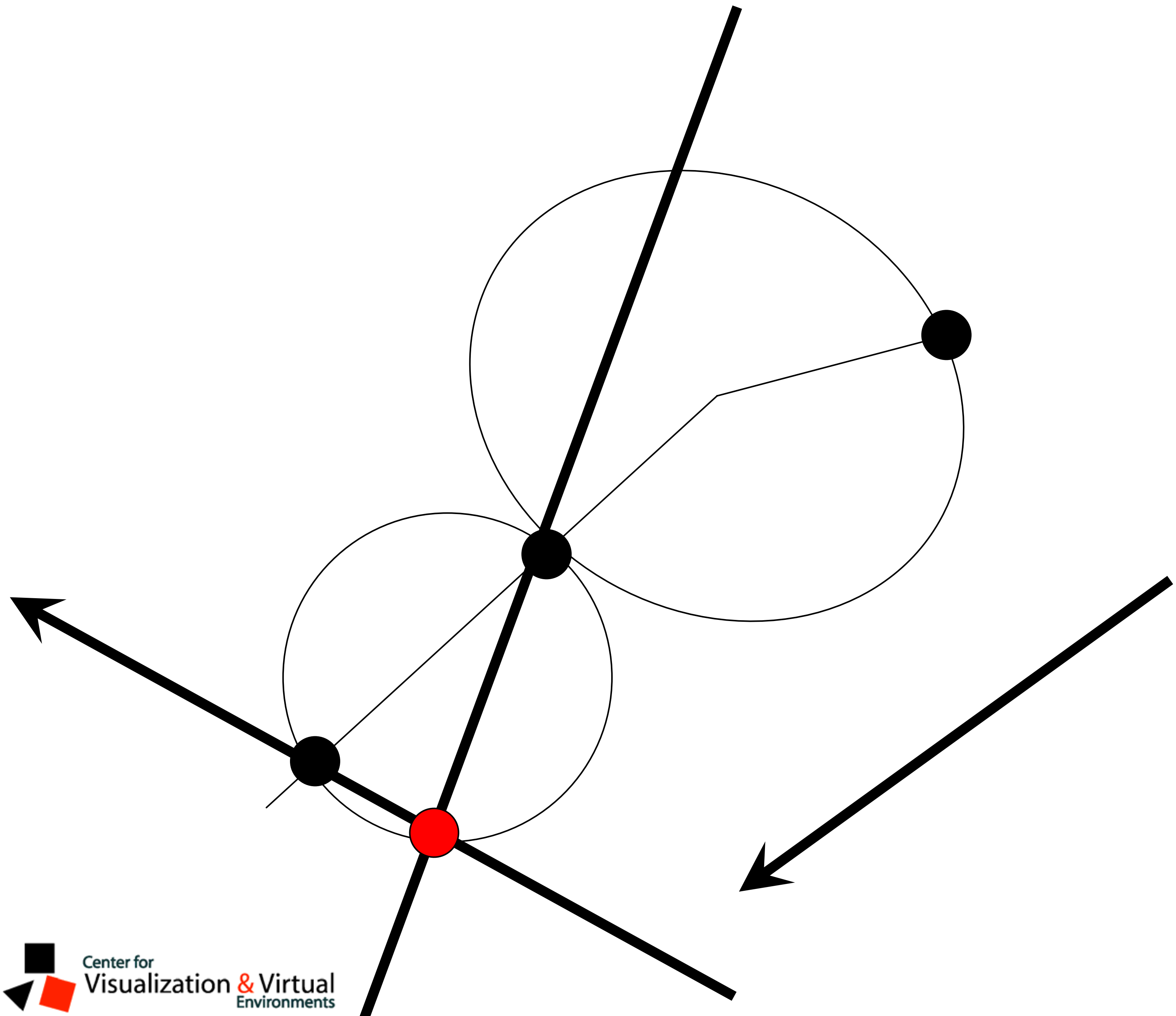


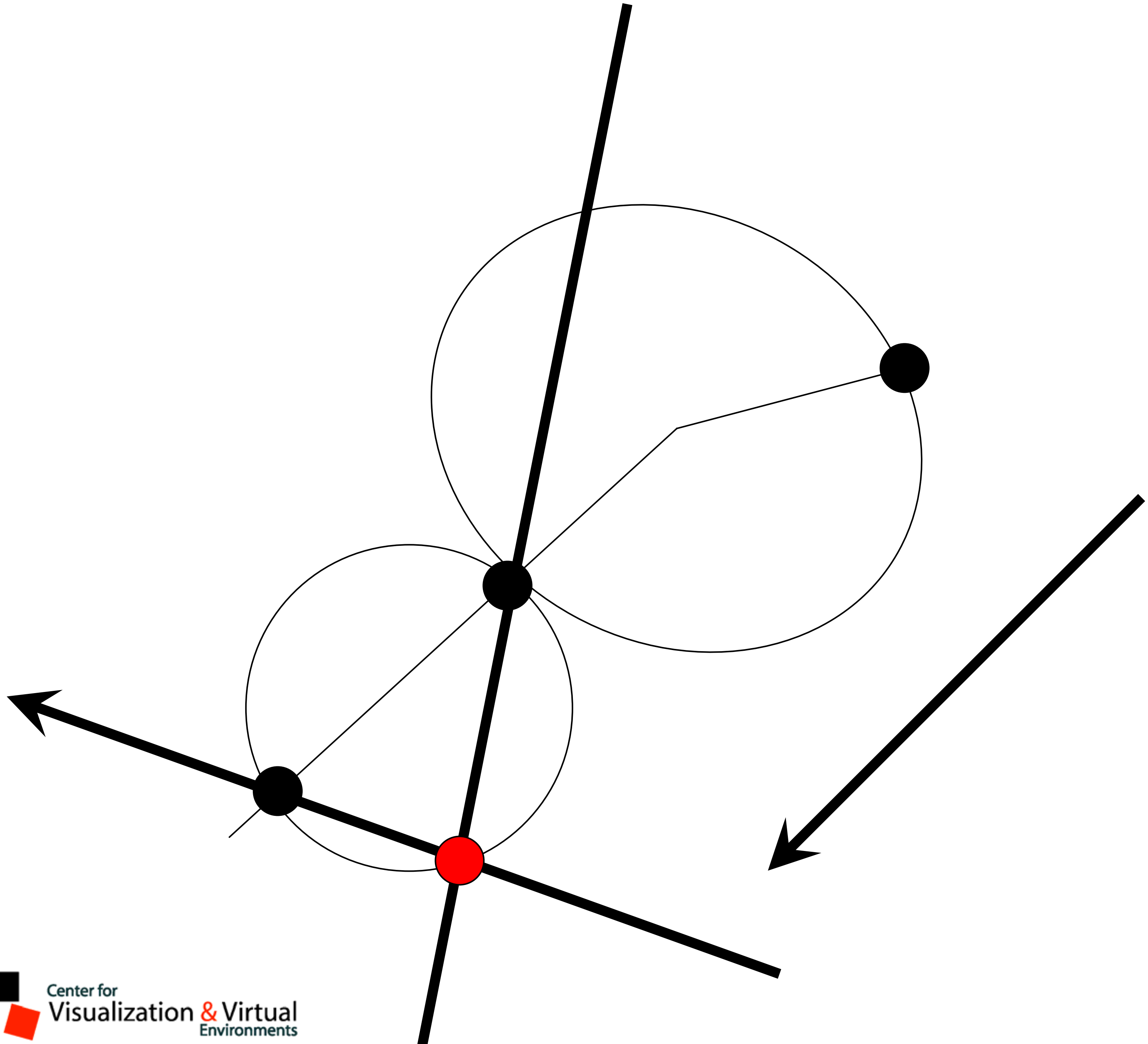
Center for Visualization & Virtual Environments

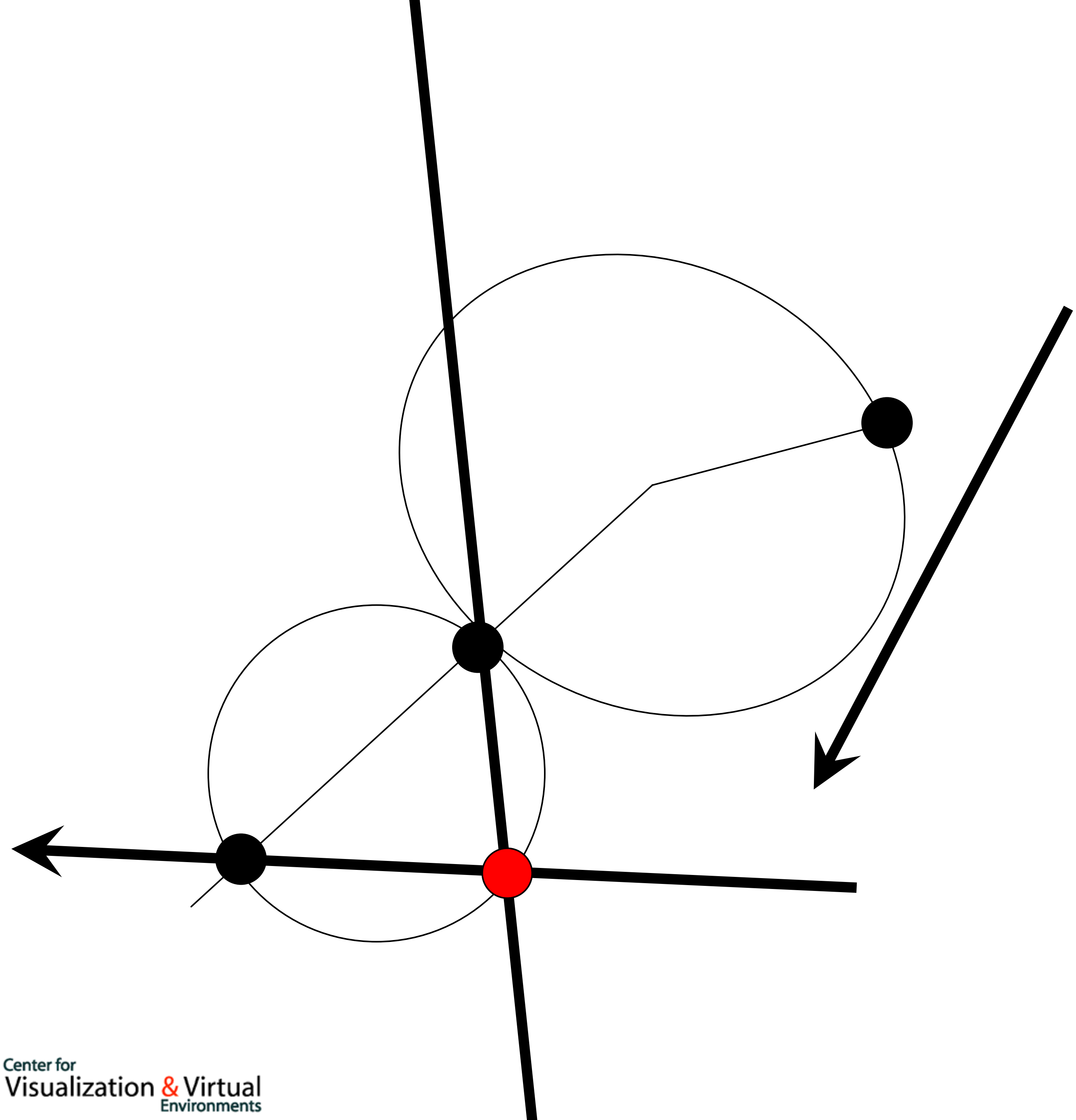


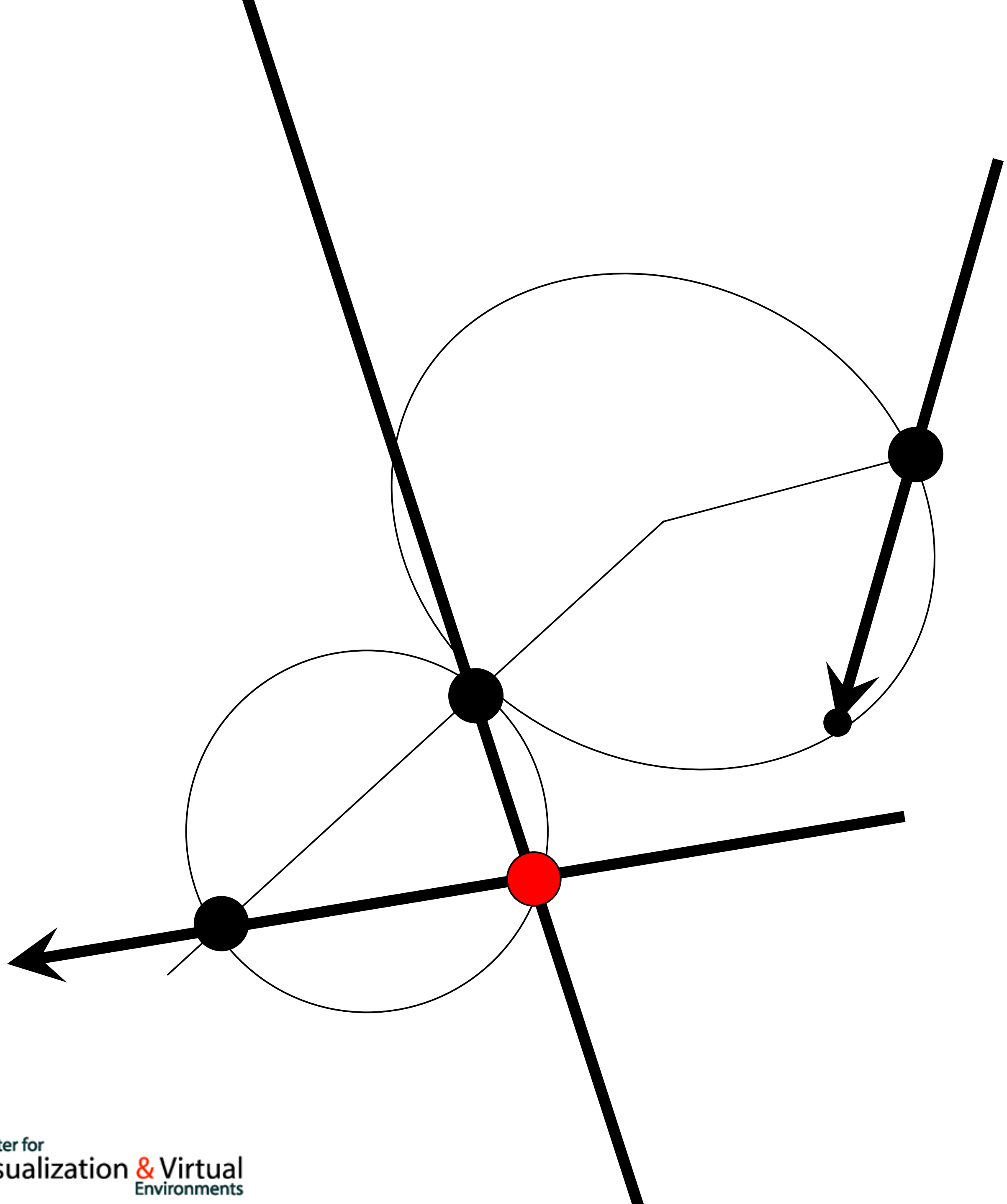
Center for Visualization & Virtual Environments



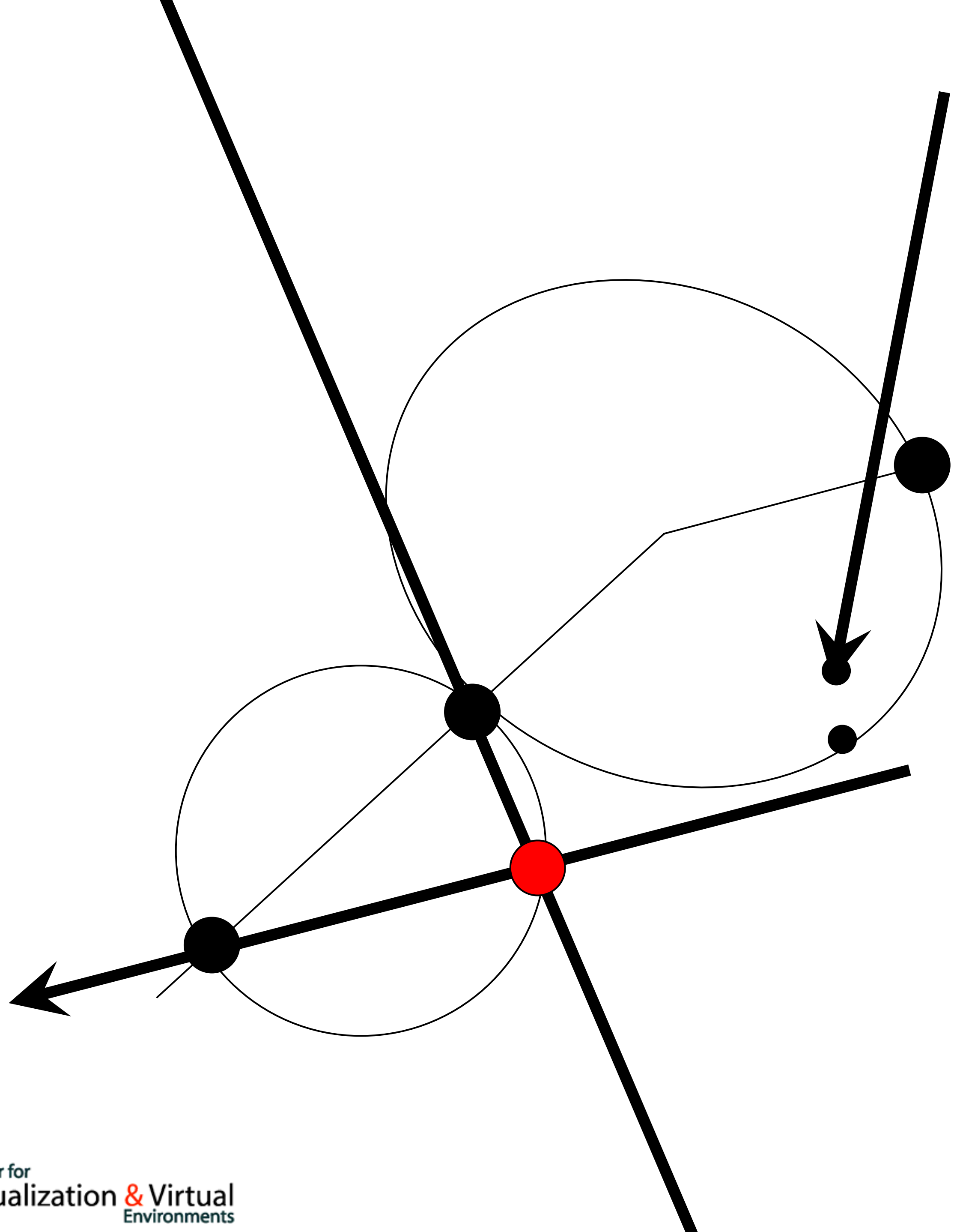


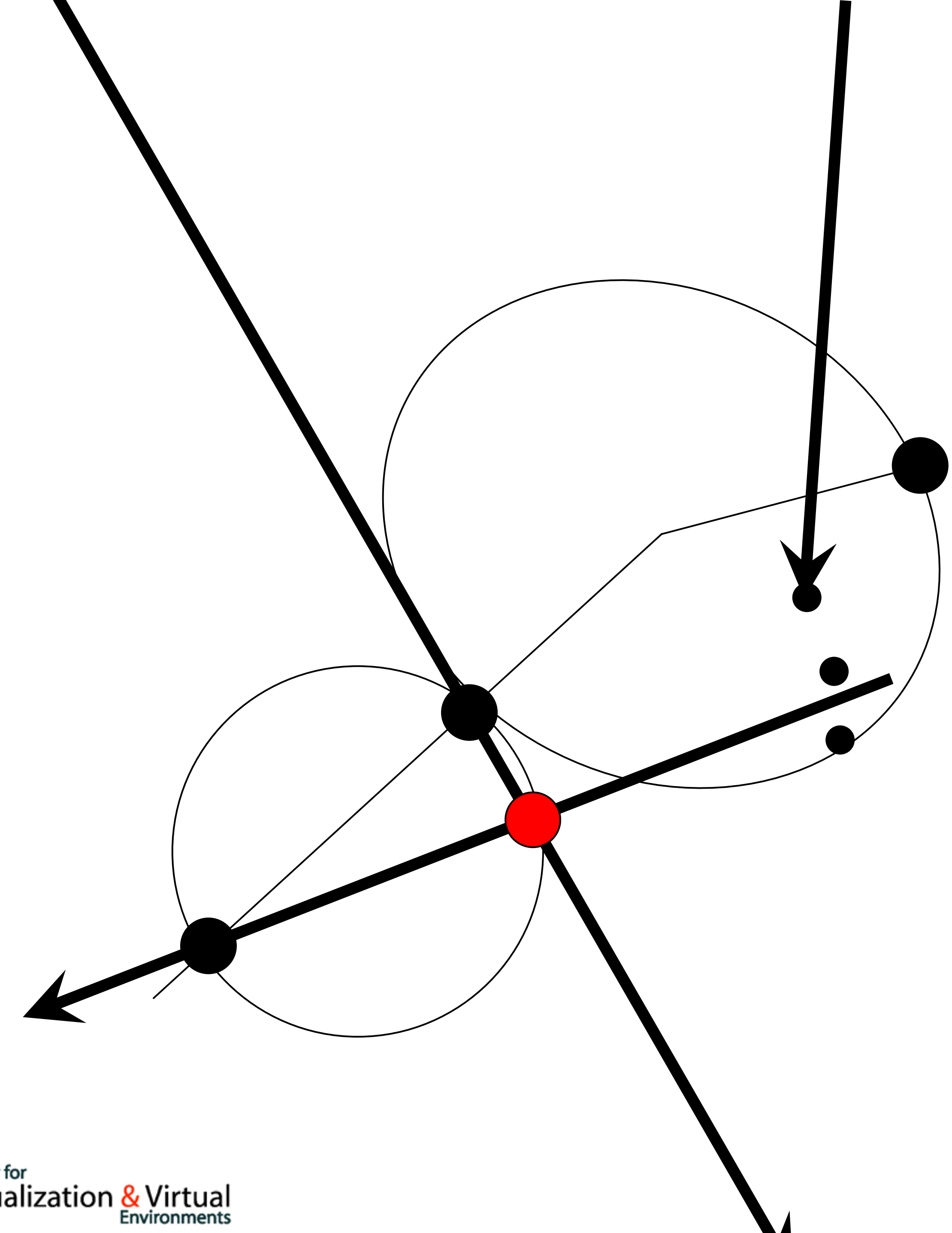


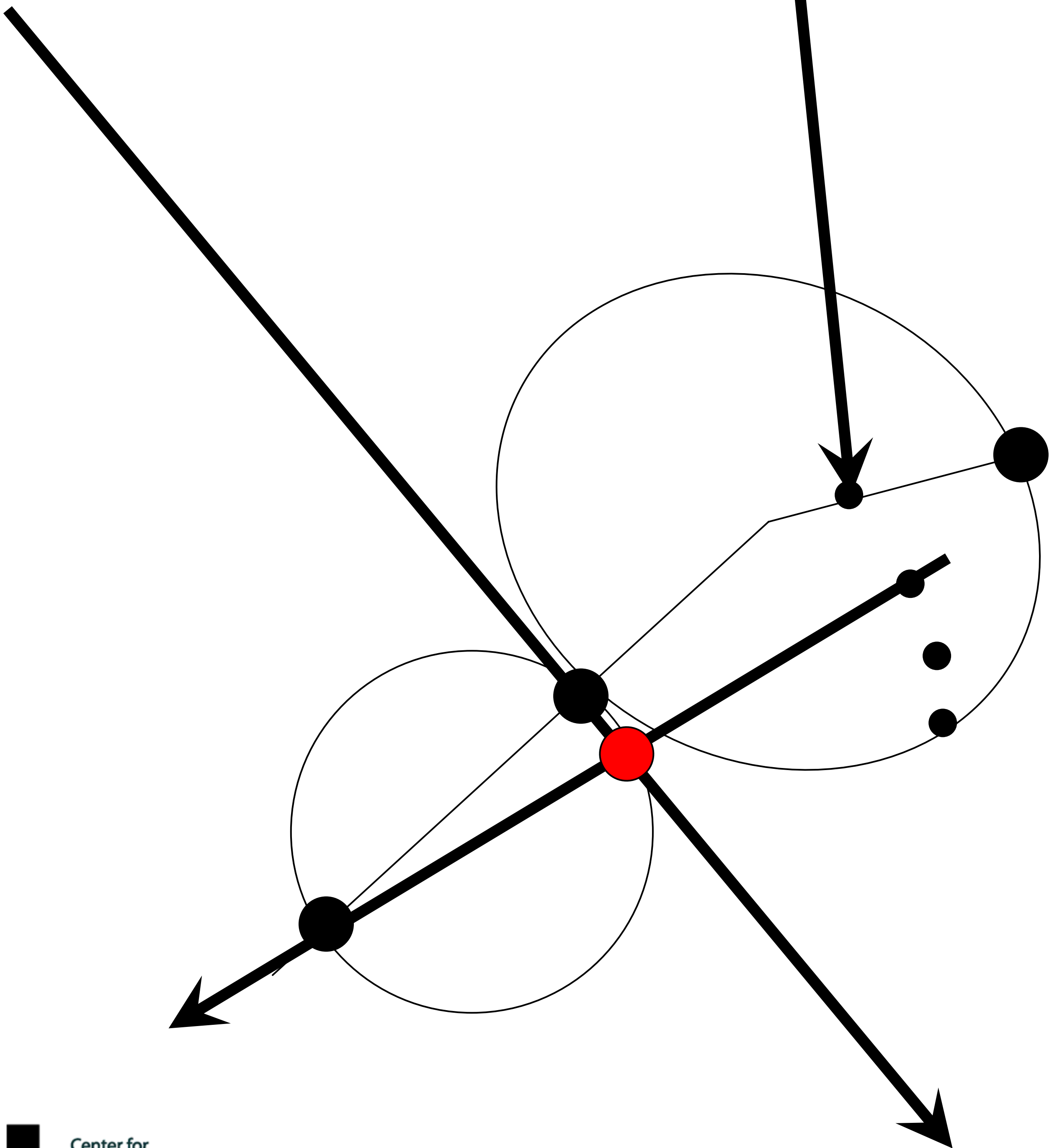


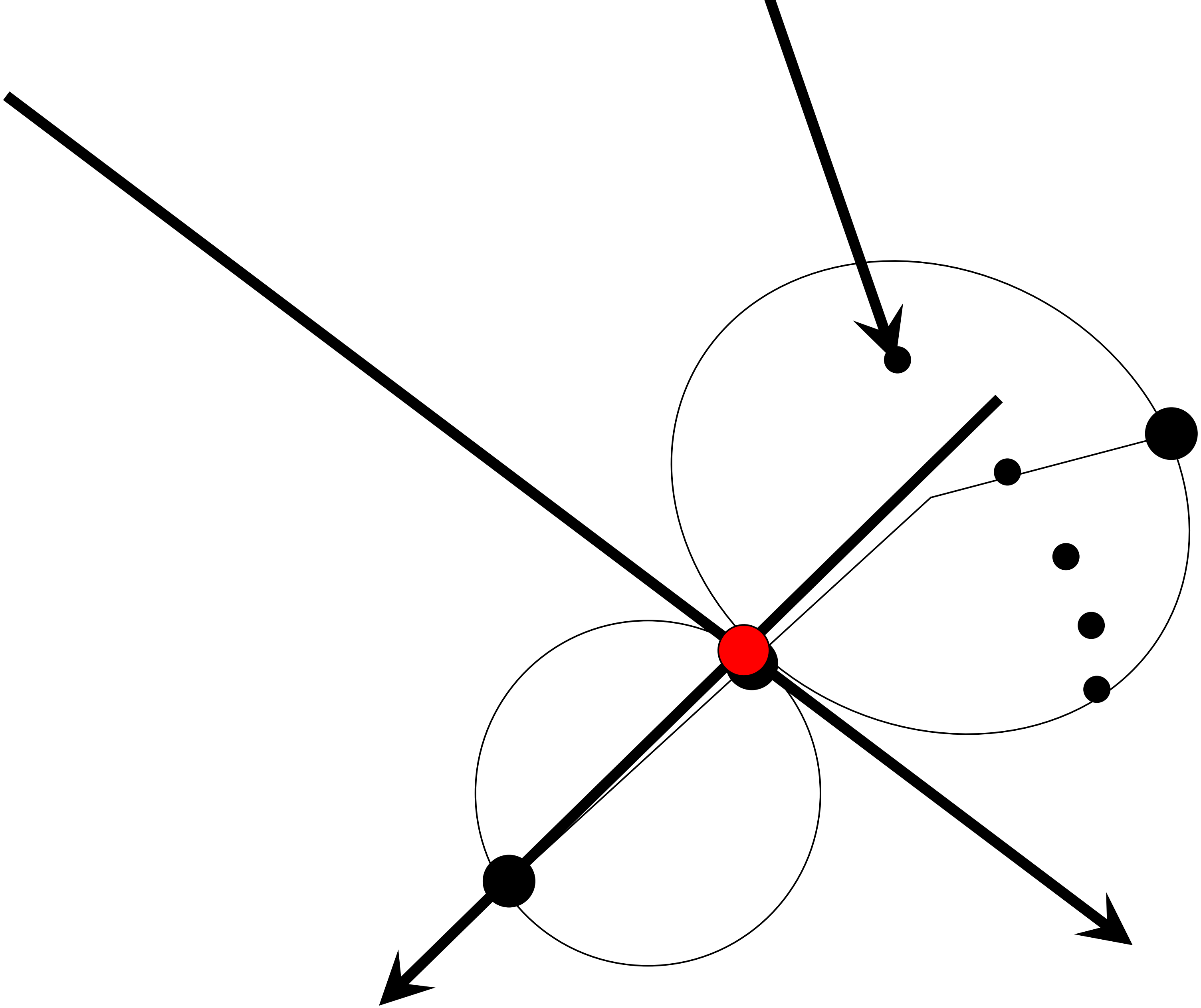


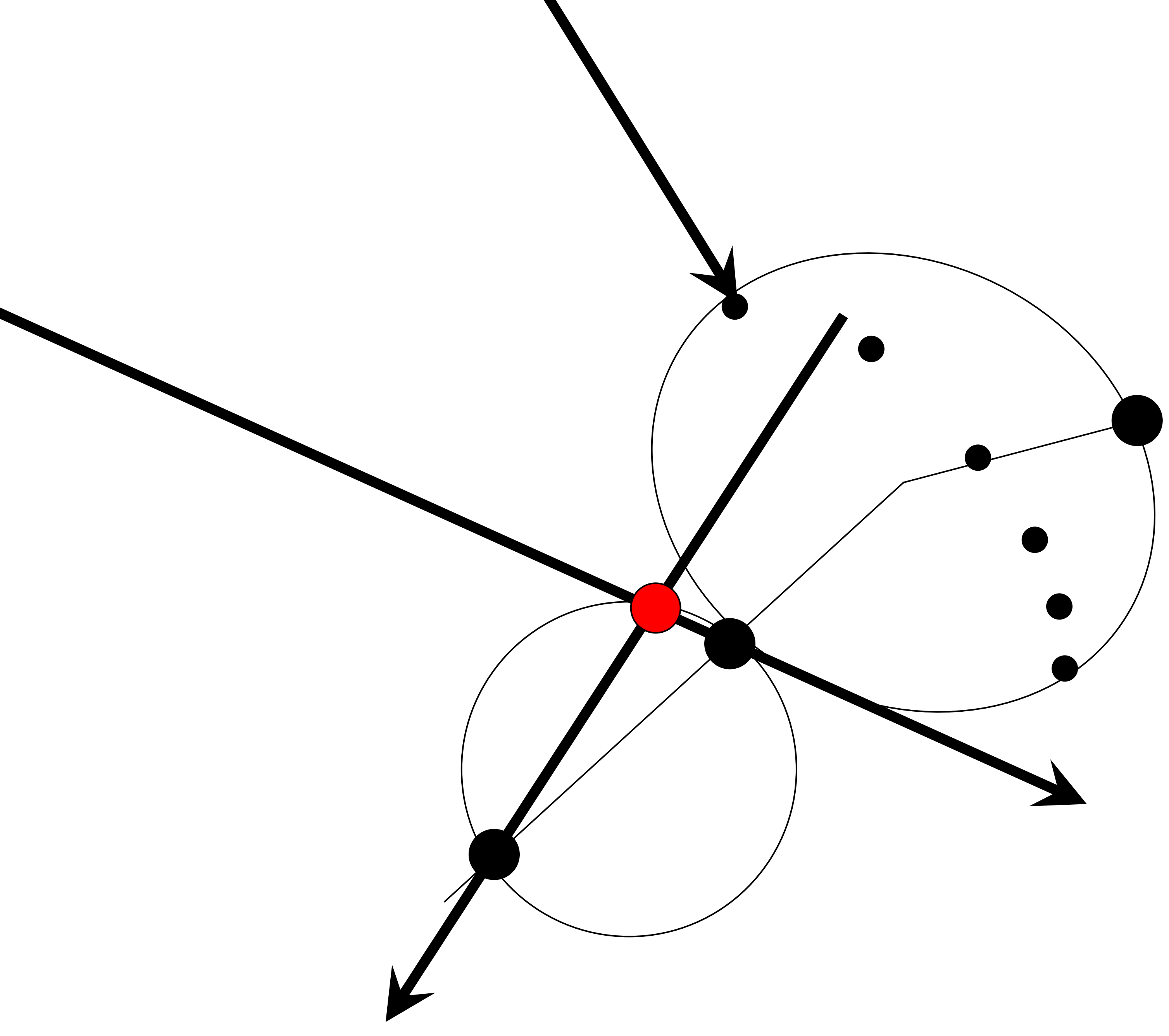
Center for Visualization & Virtual Environments



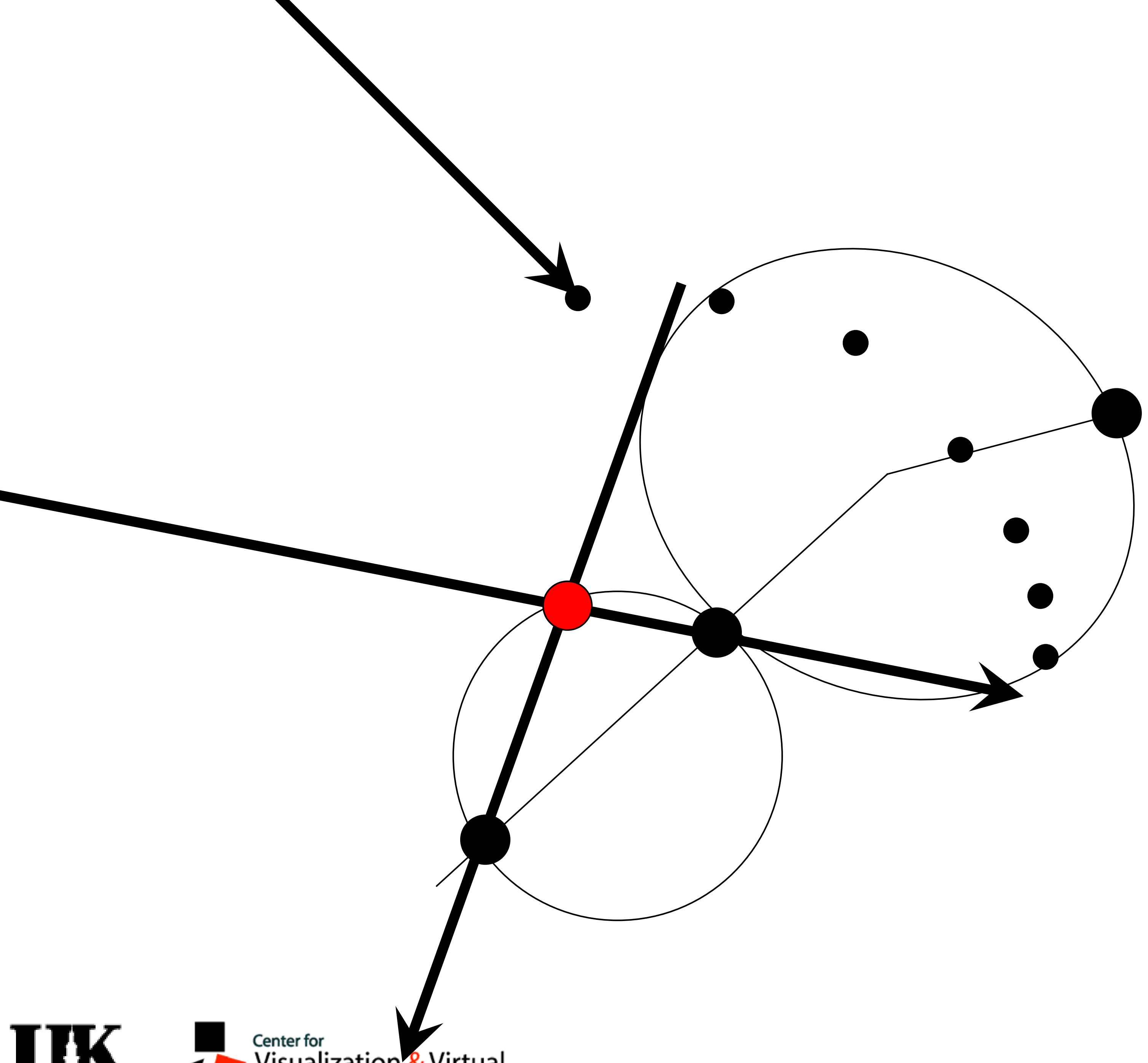


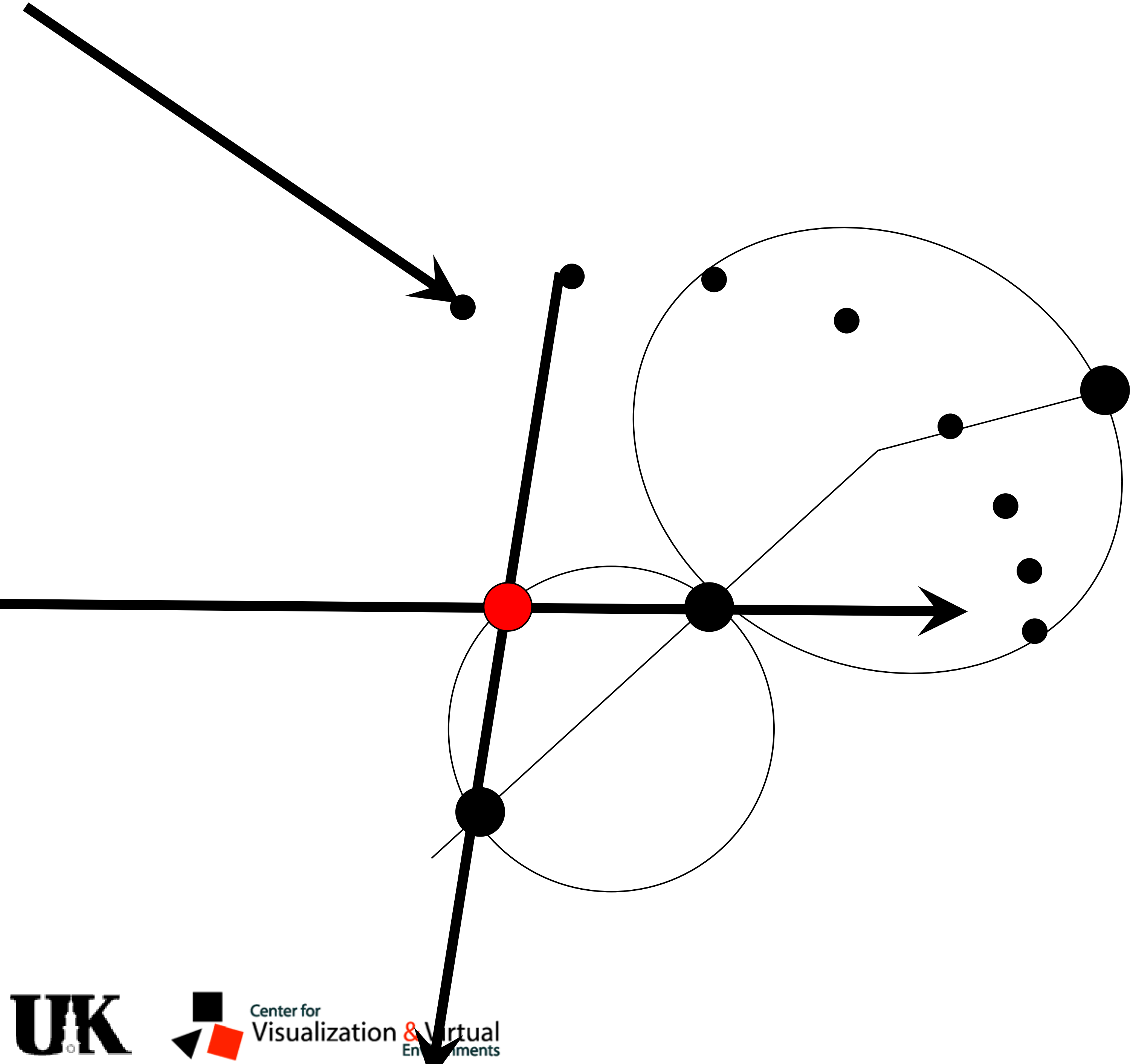


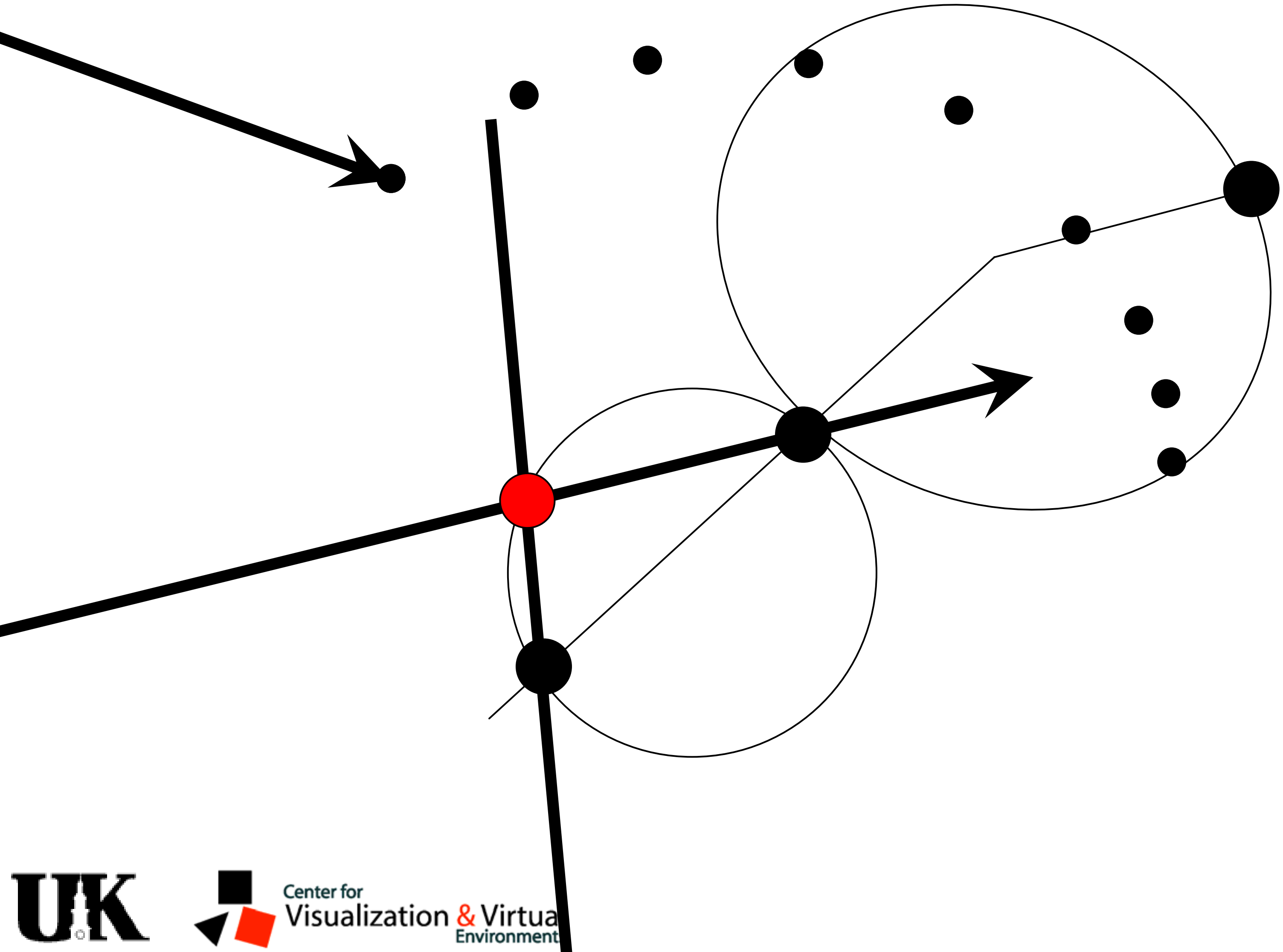




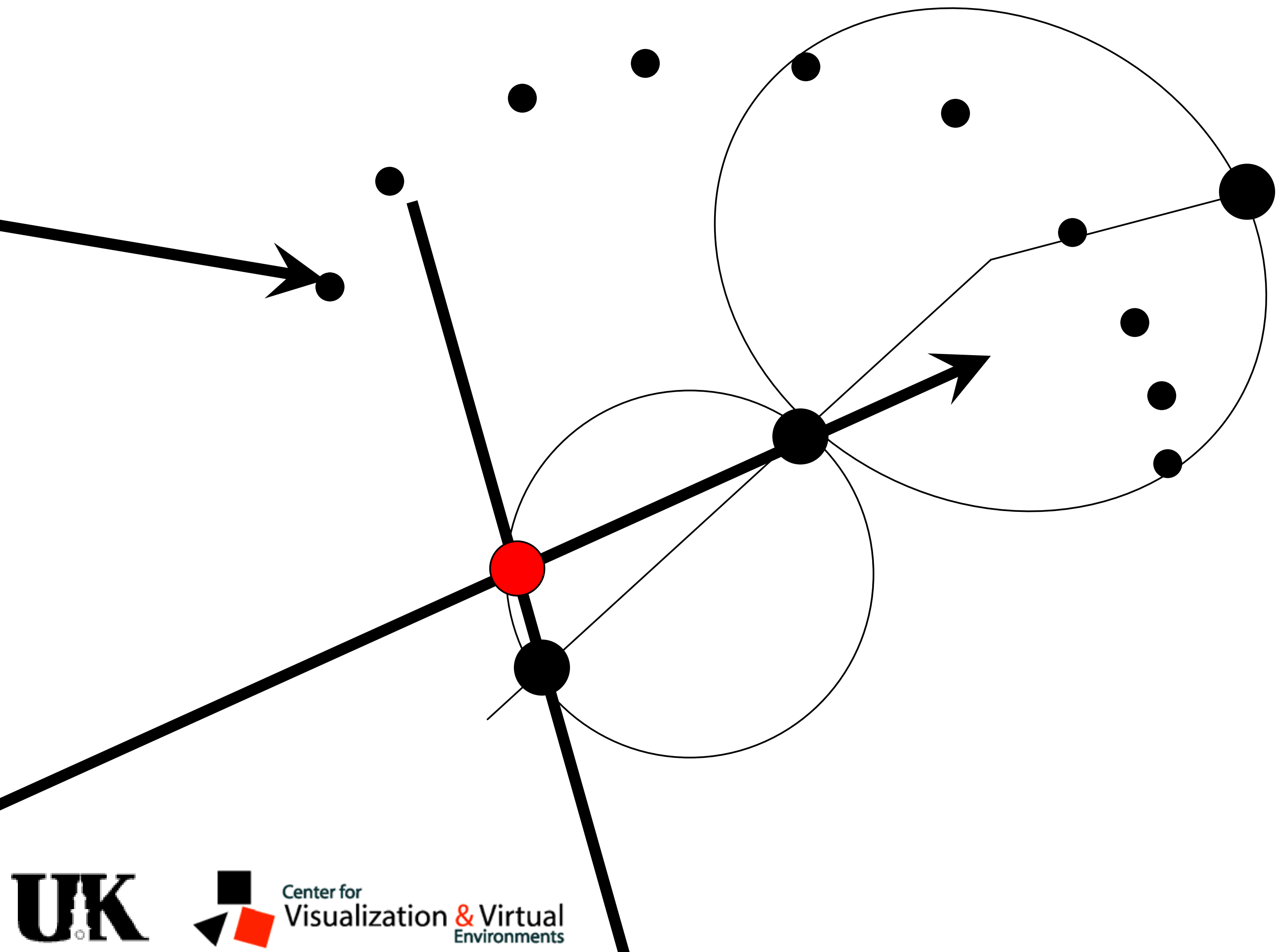
UK Center for Visualization & Virtual Environments

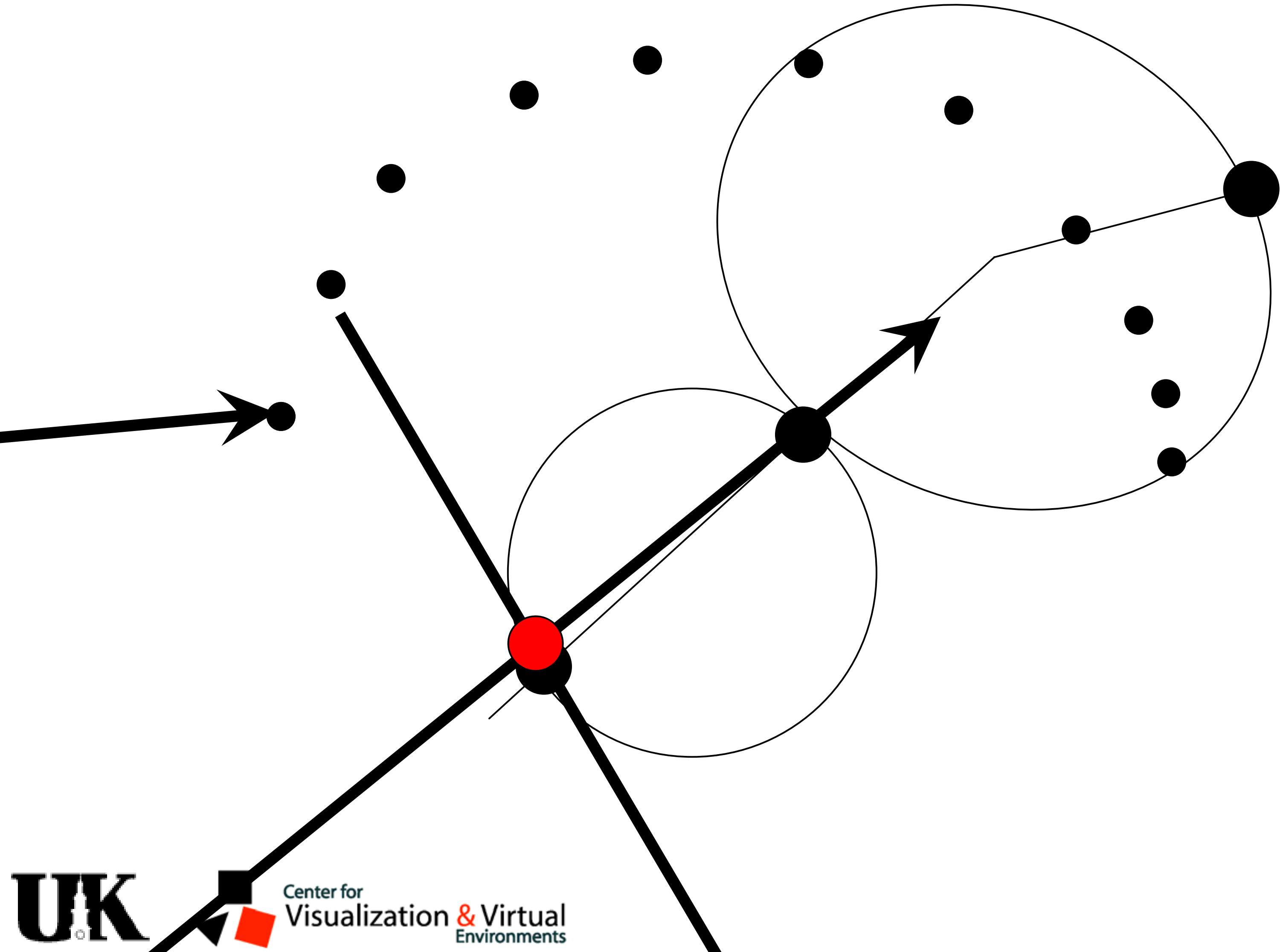




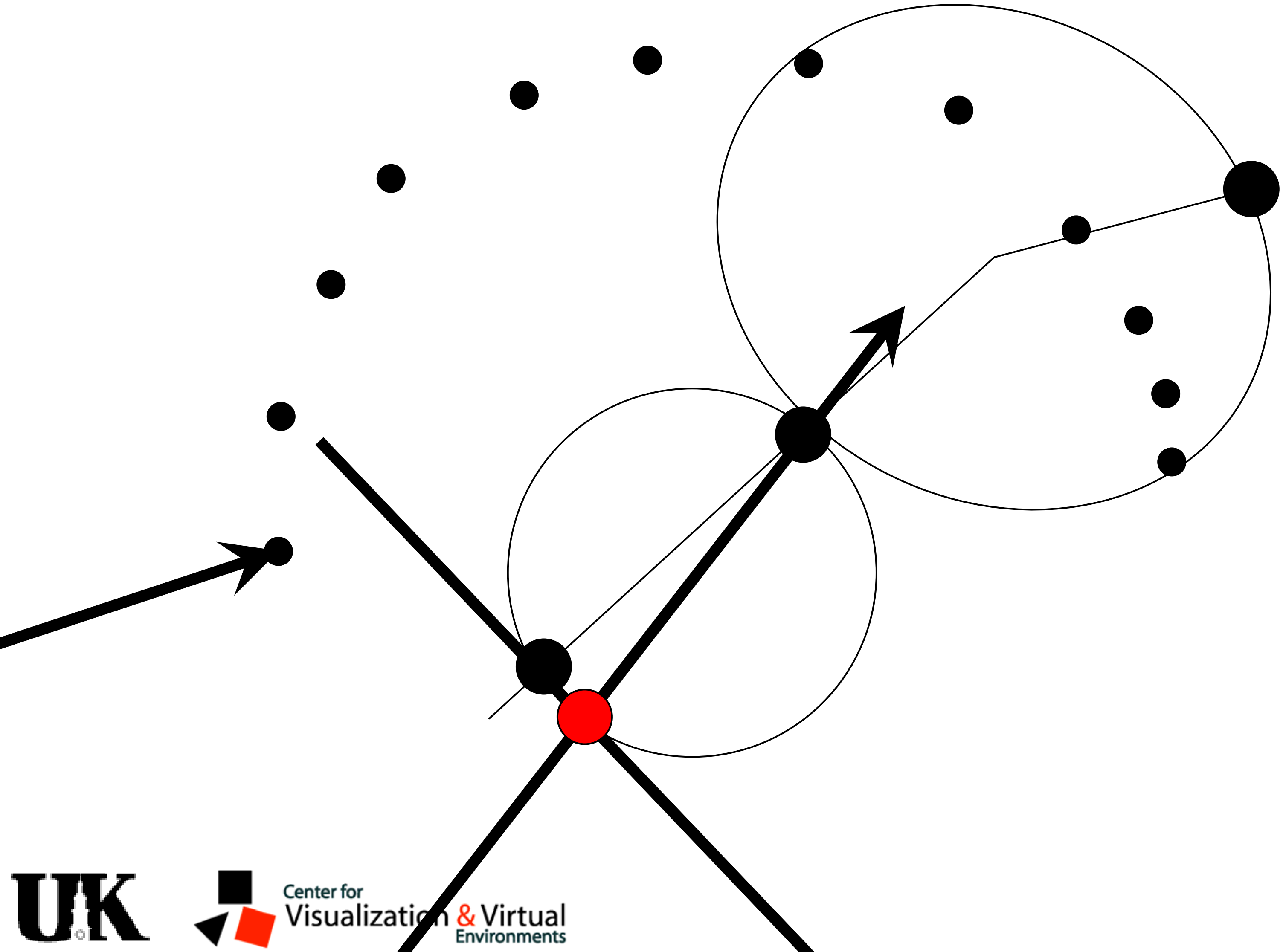


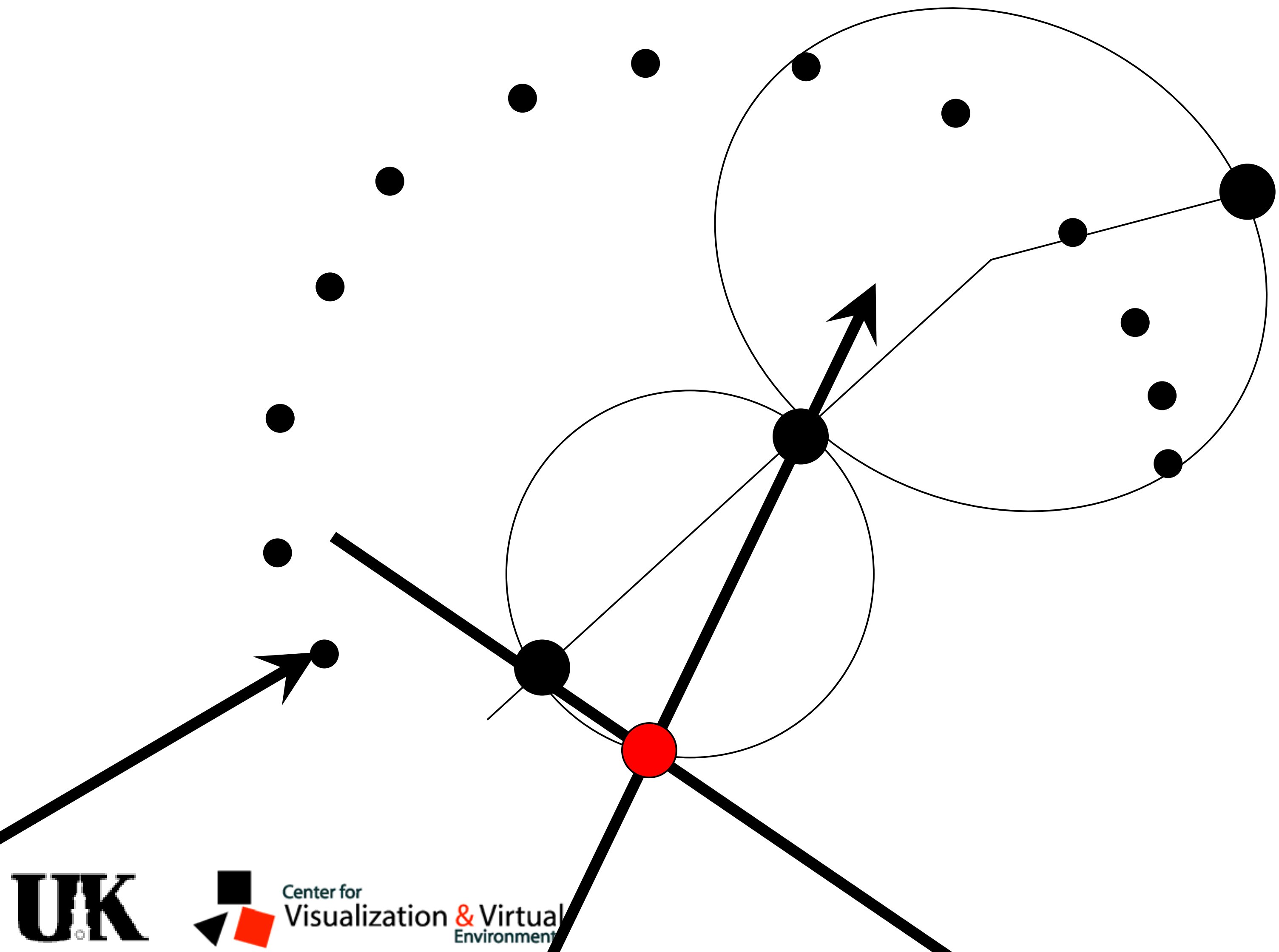
Center for Visualization & Virtual Environments

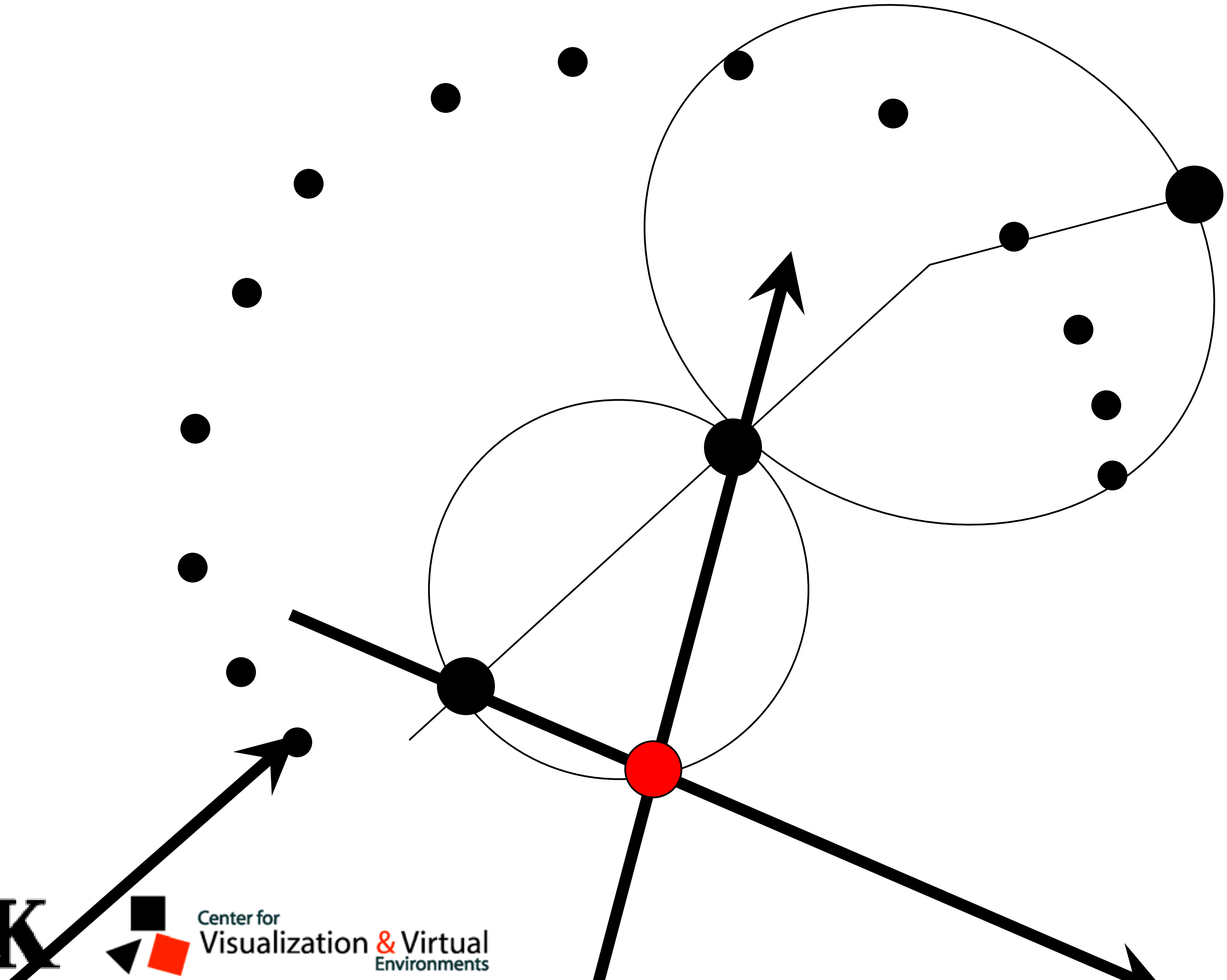


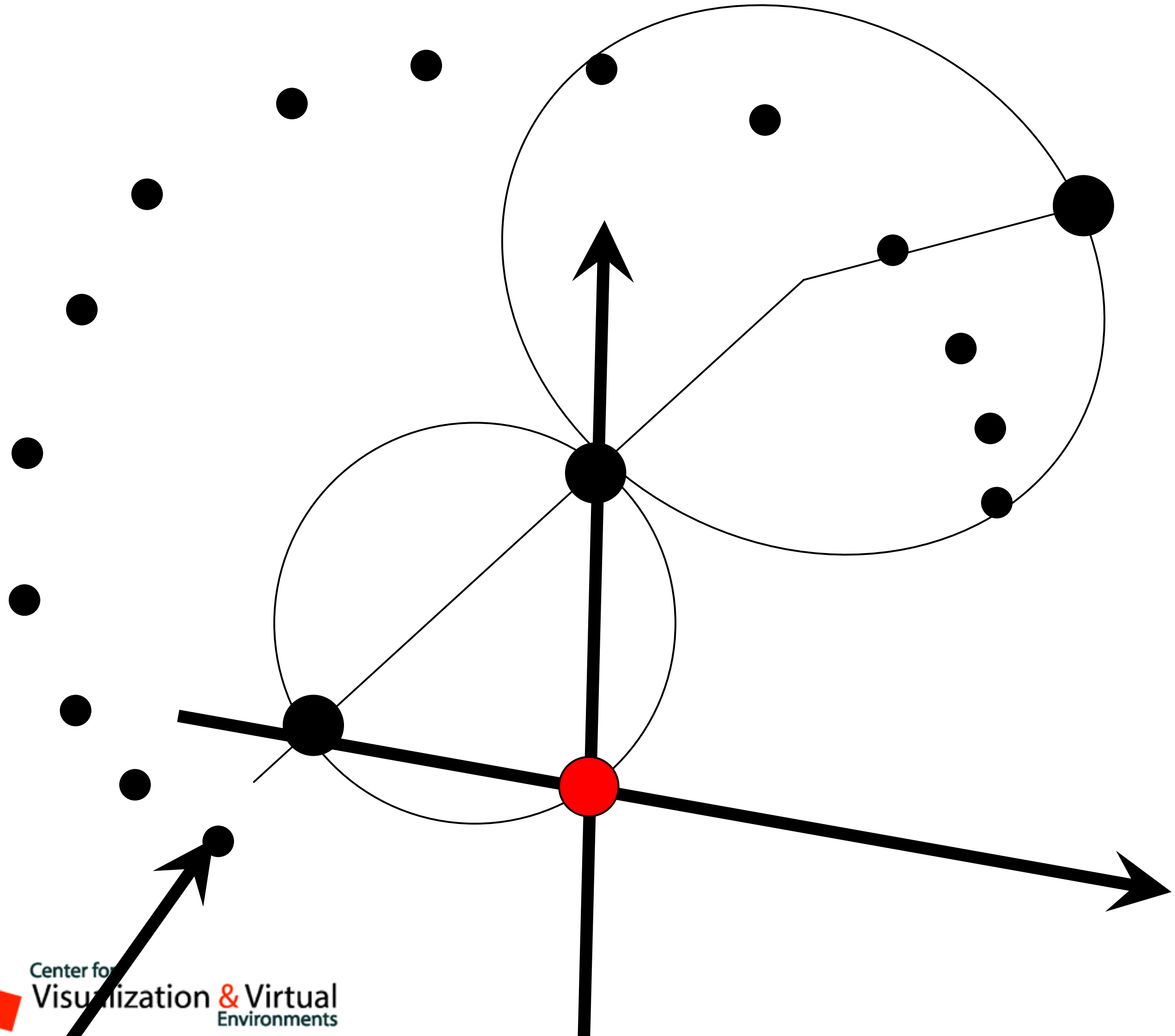


Center for Visualization & Virtual Environments

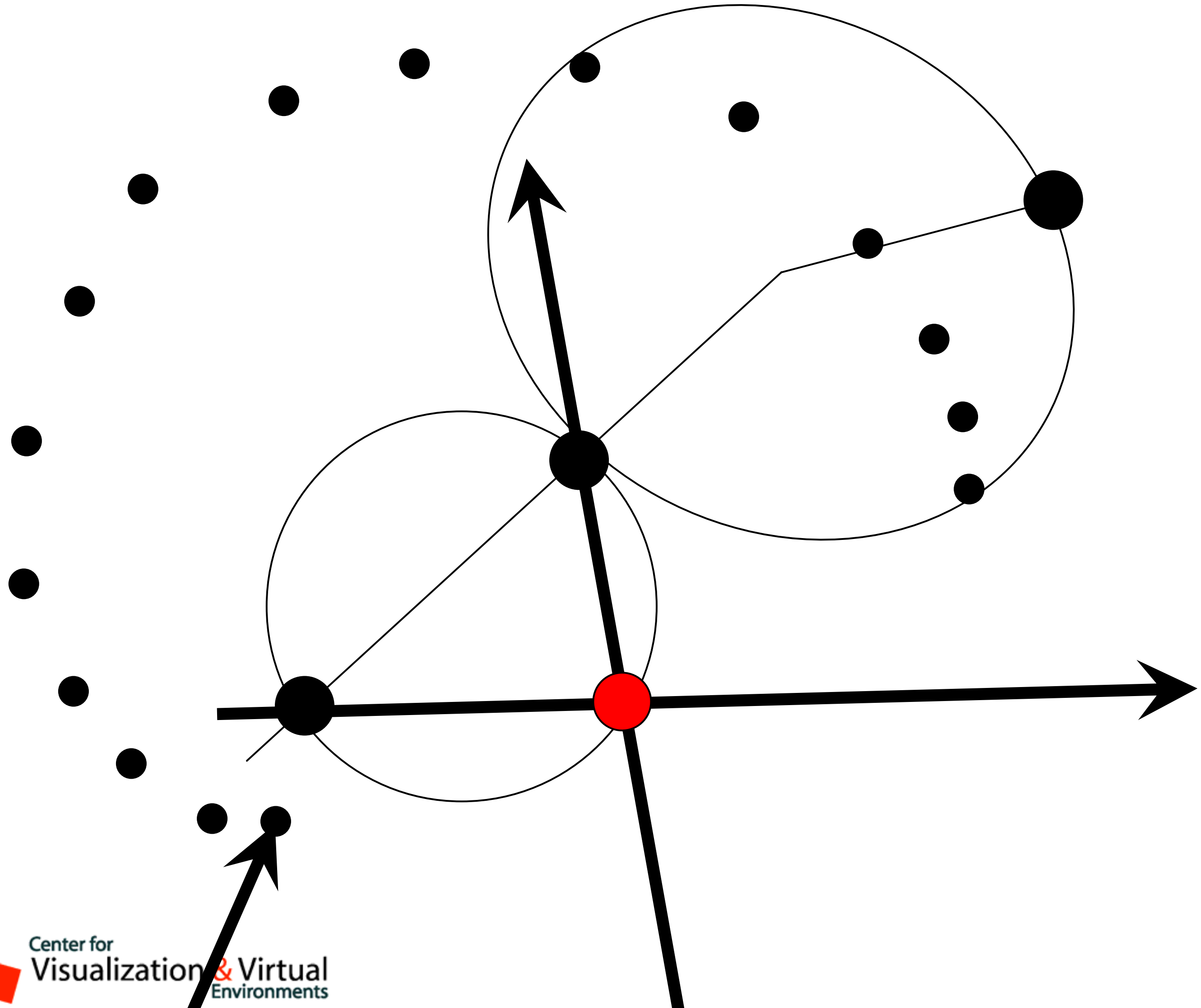




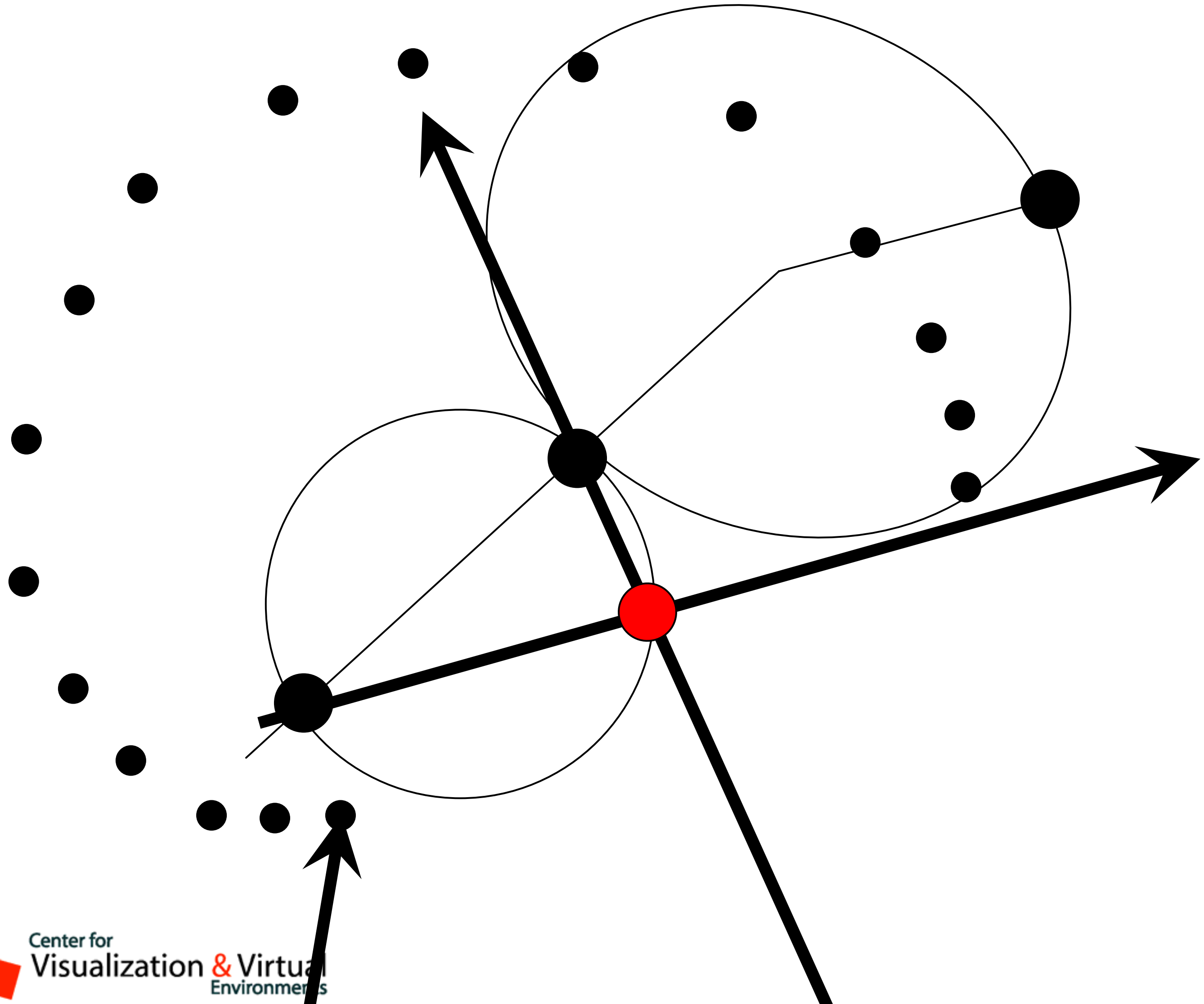


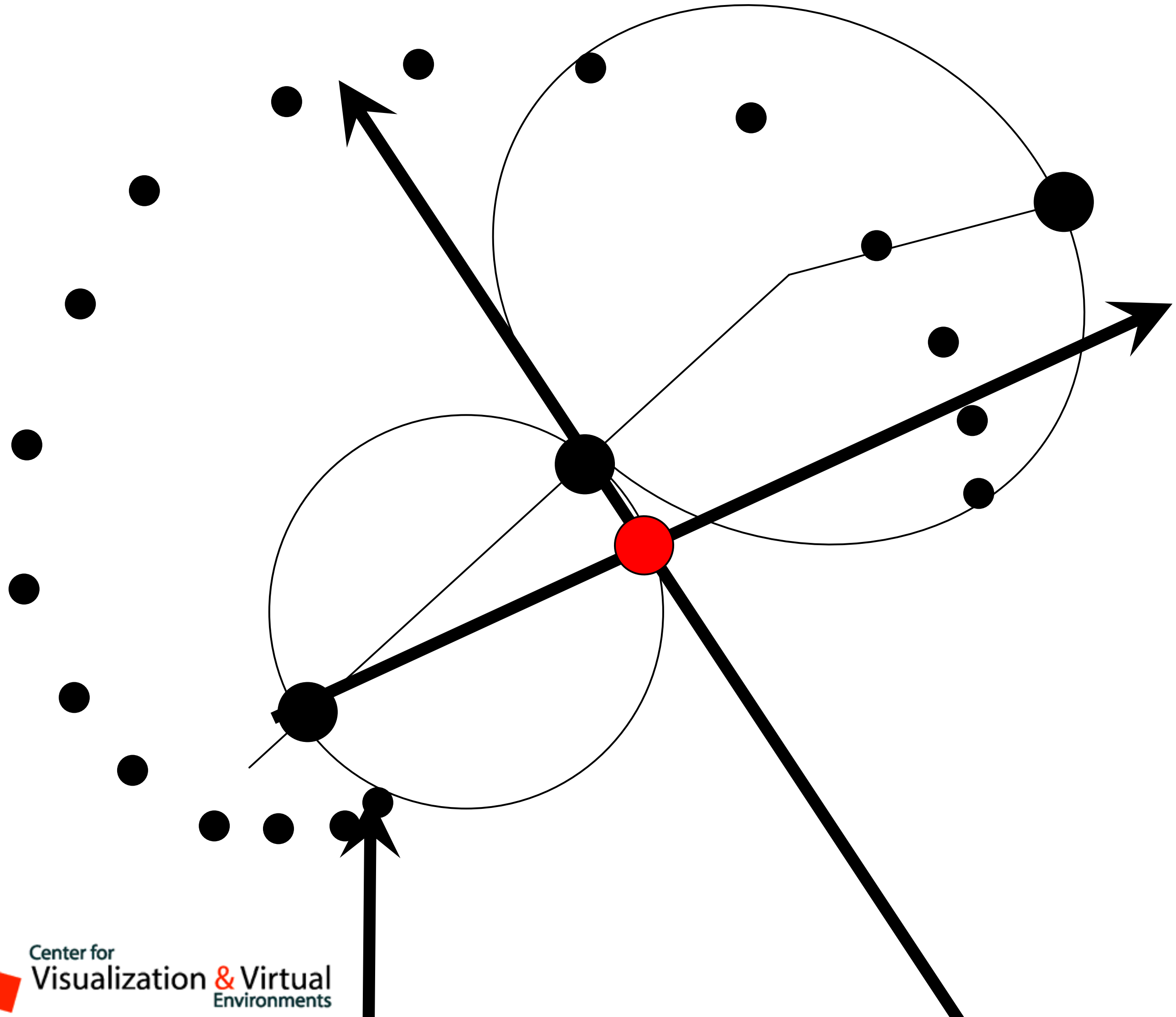


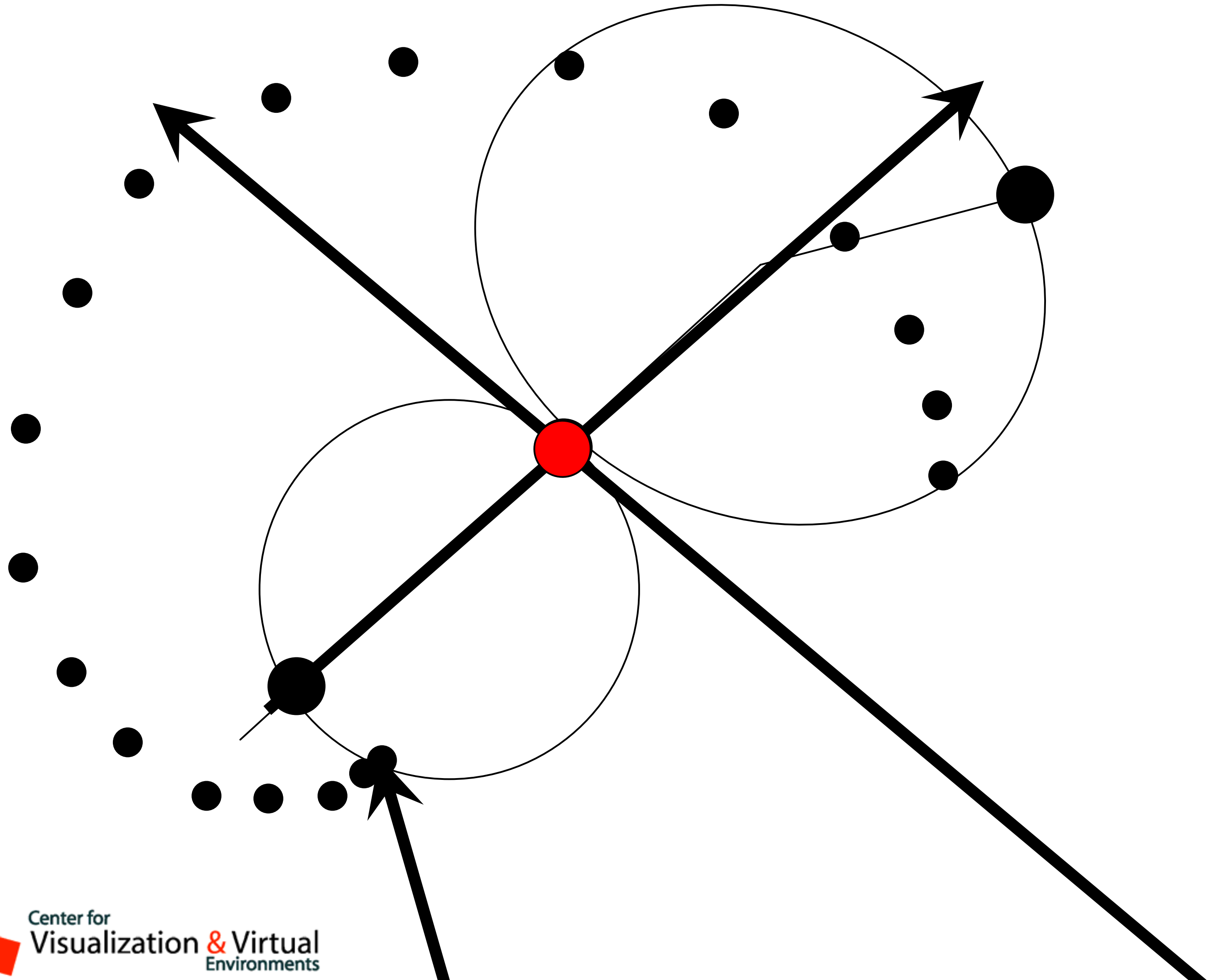
UK Center for Visualization & Virtual Environments

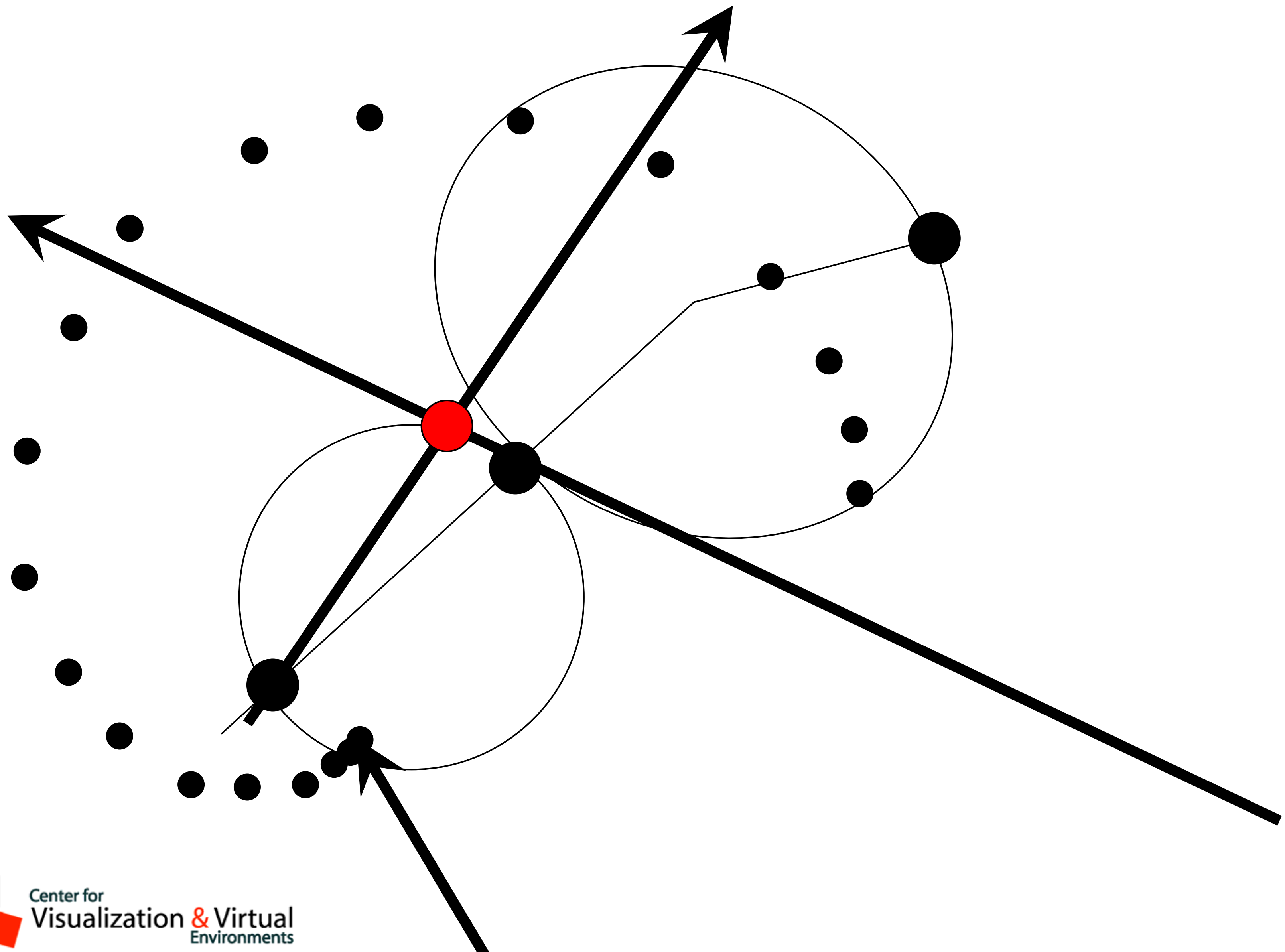


Center for Visualization & Virtual Environments

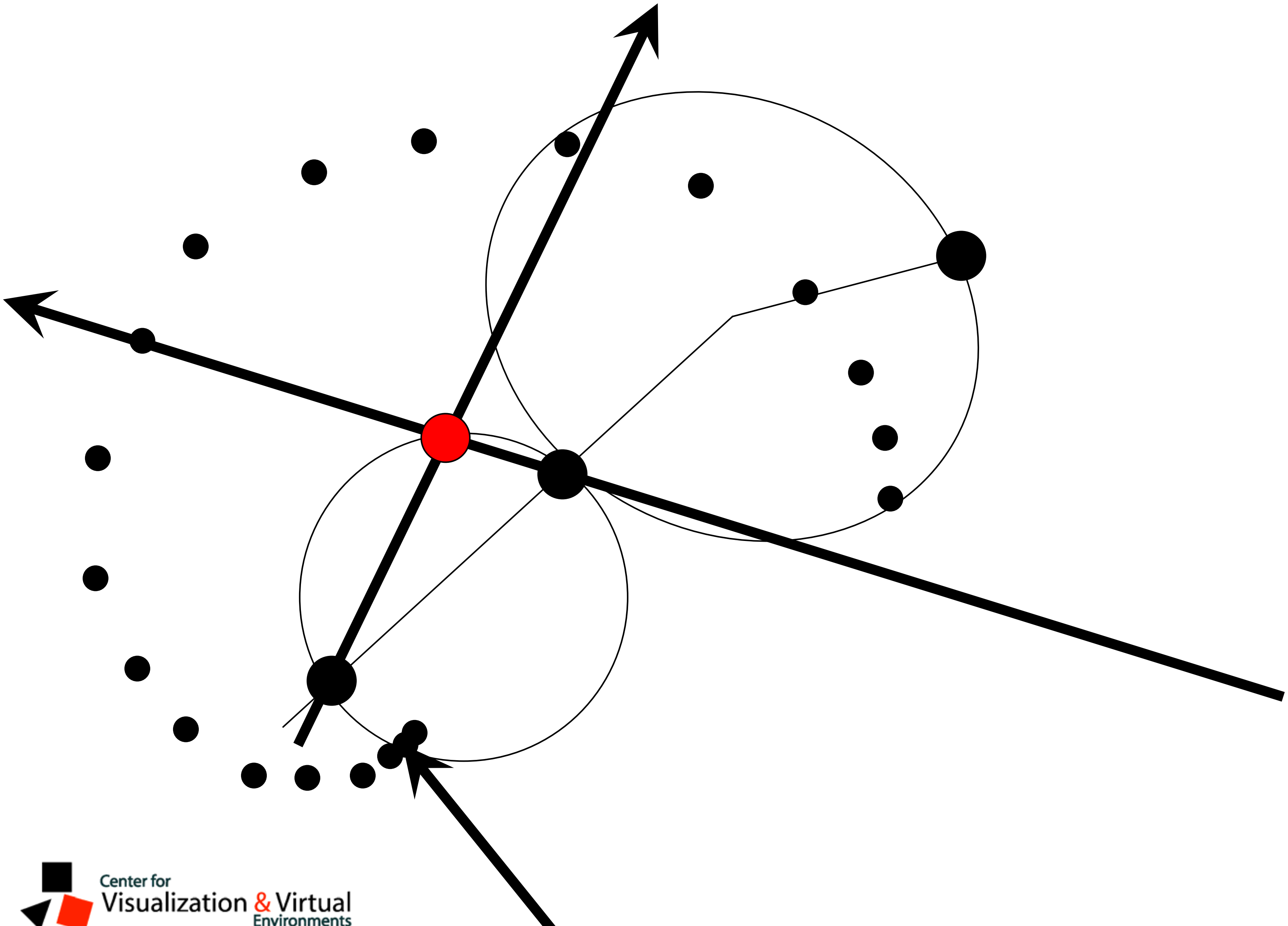


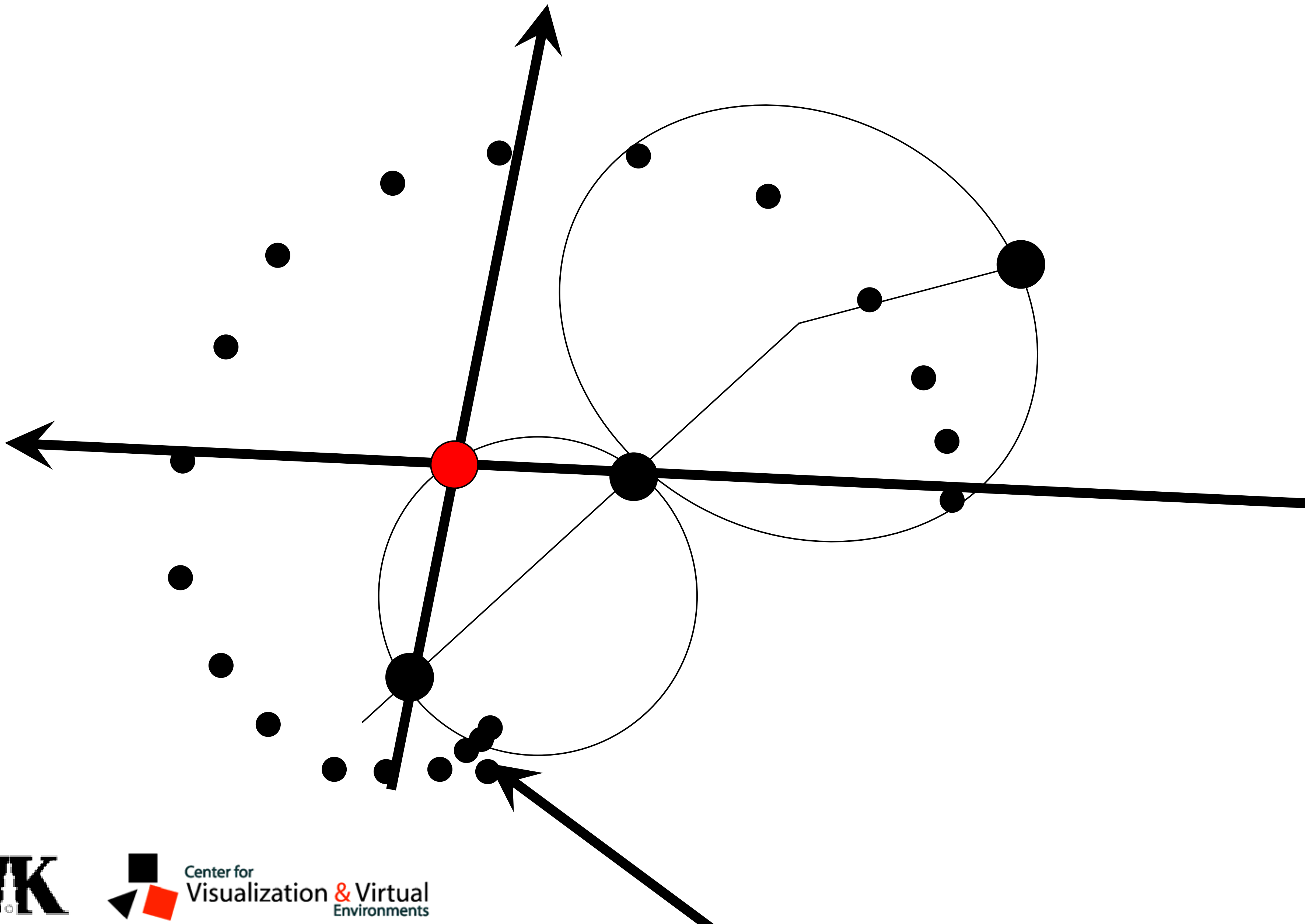


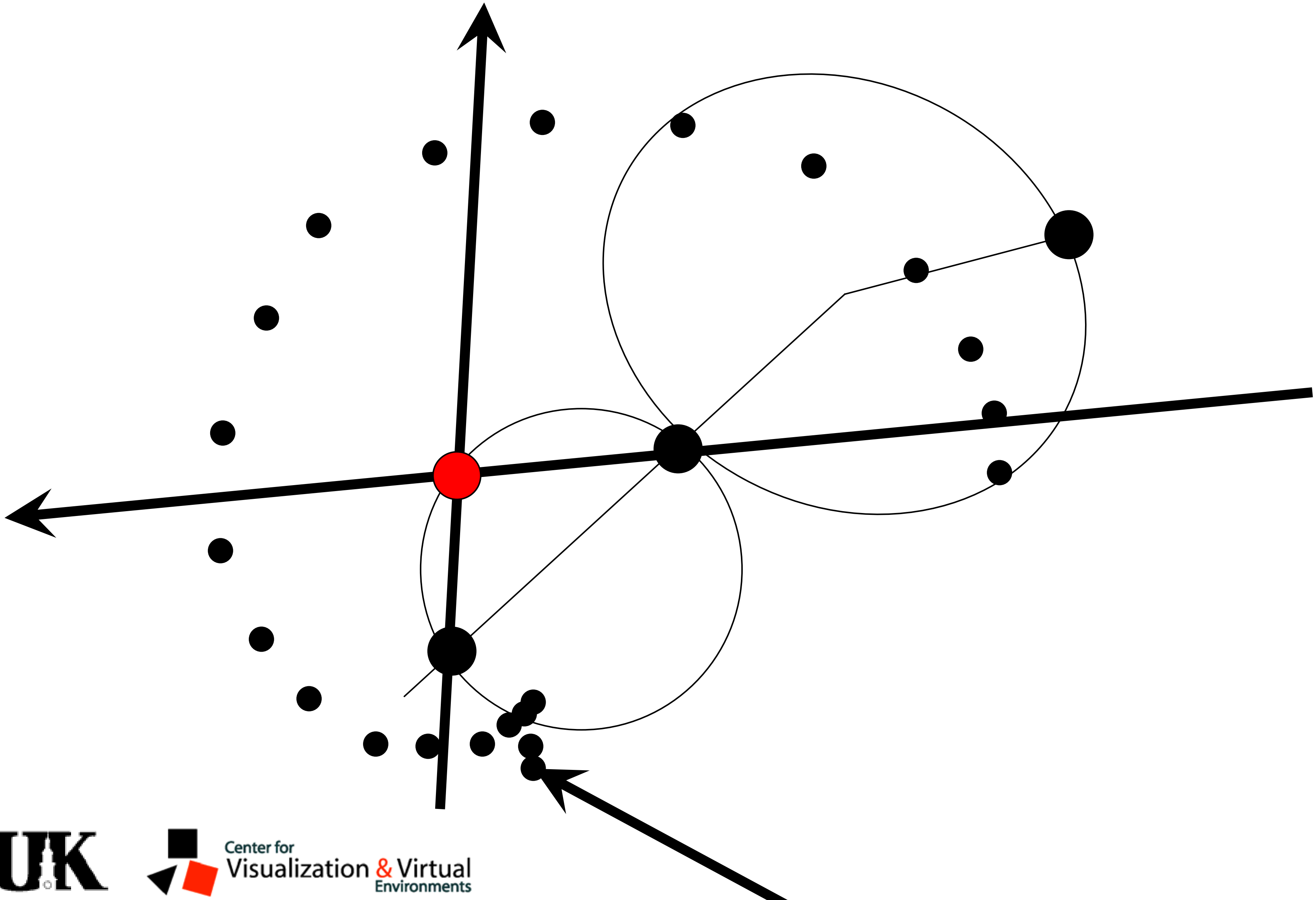


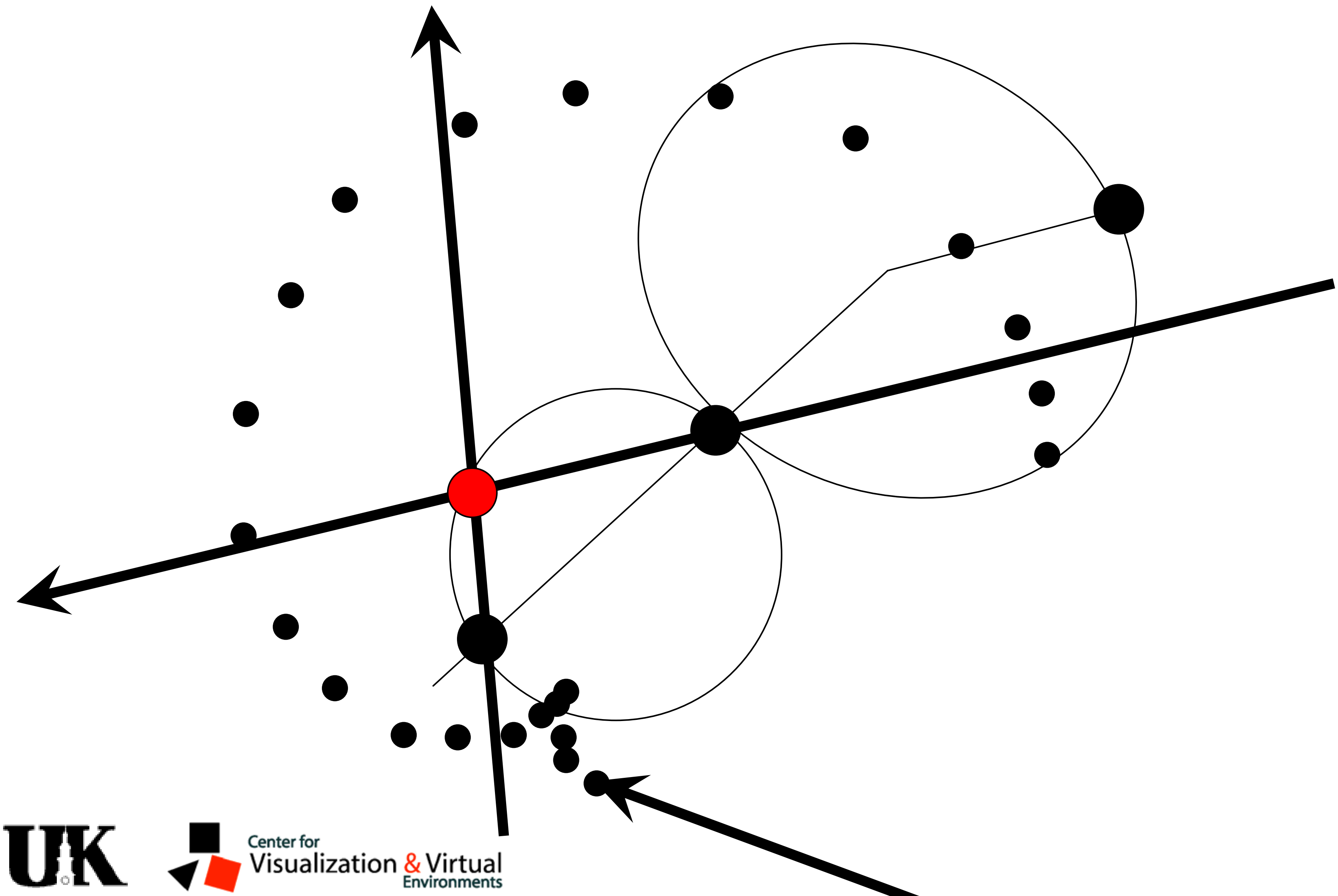


UK Center for Visualization & Virtual Environments

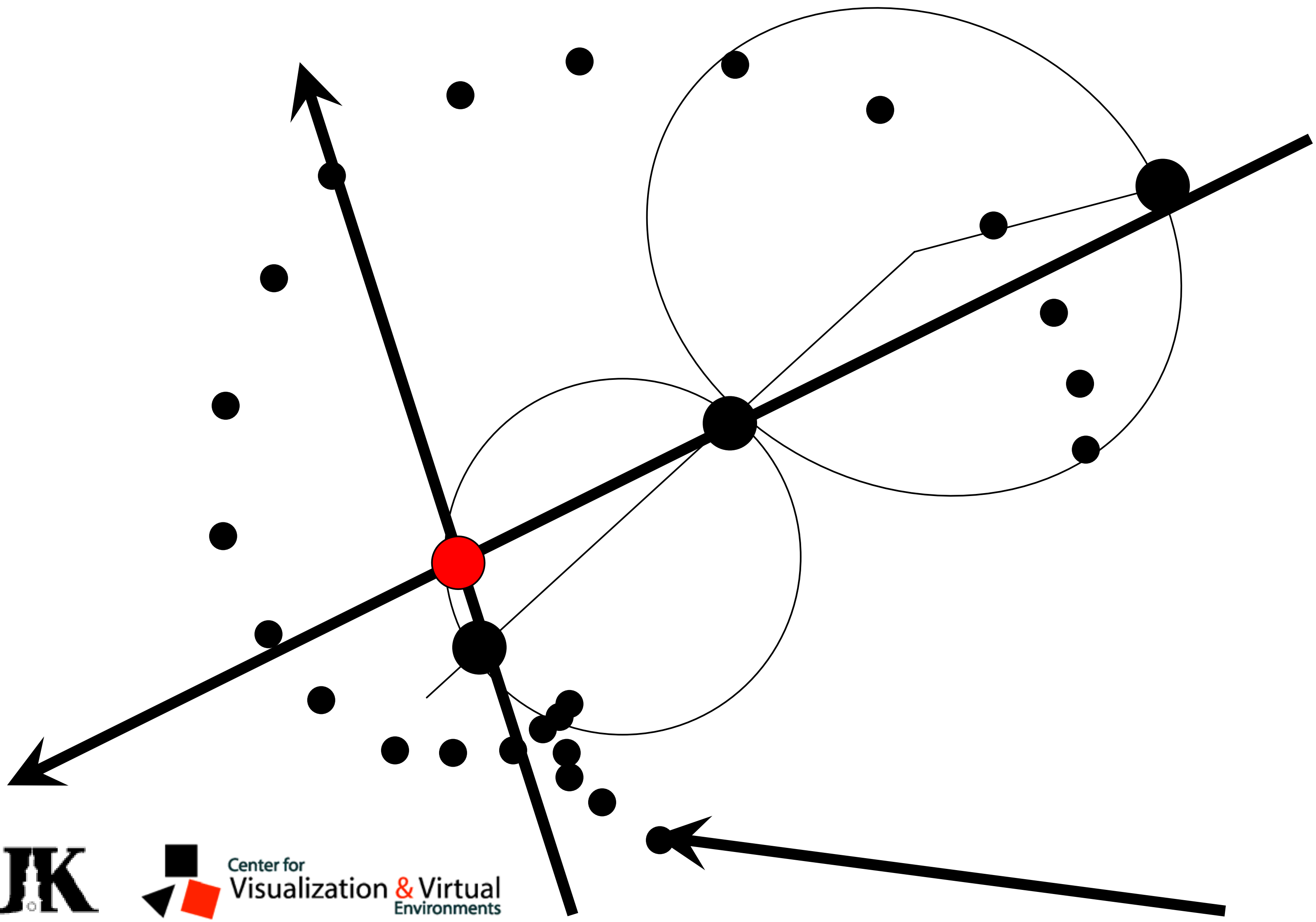


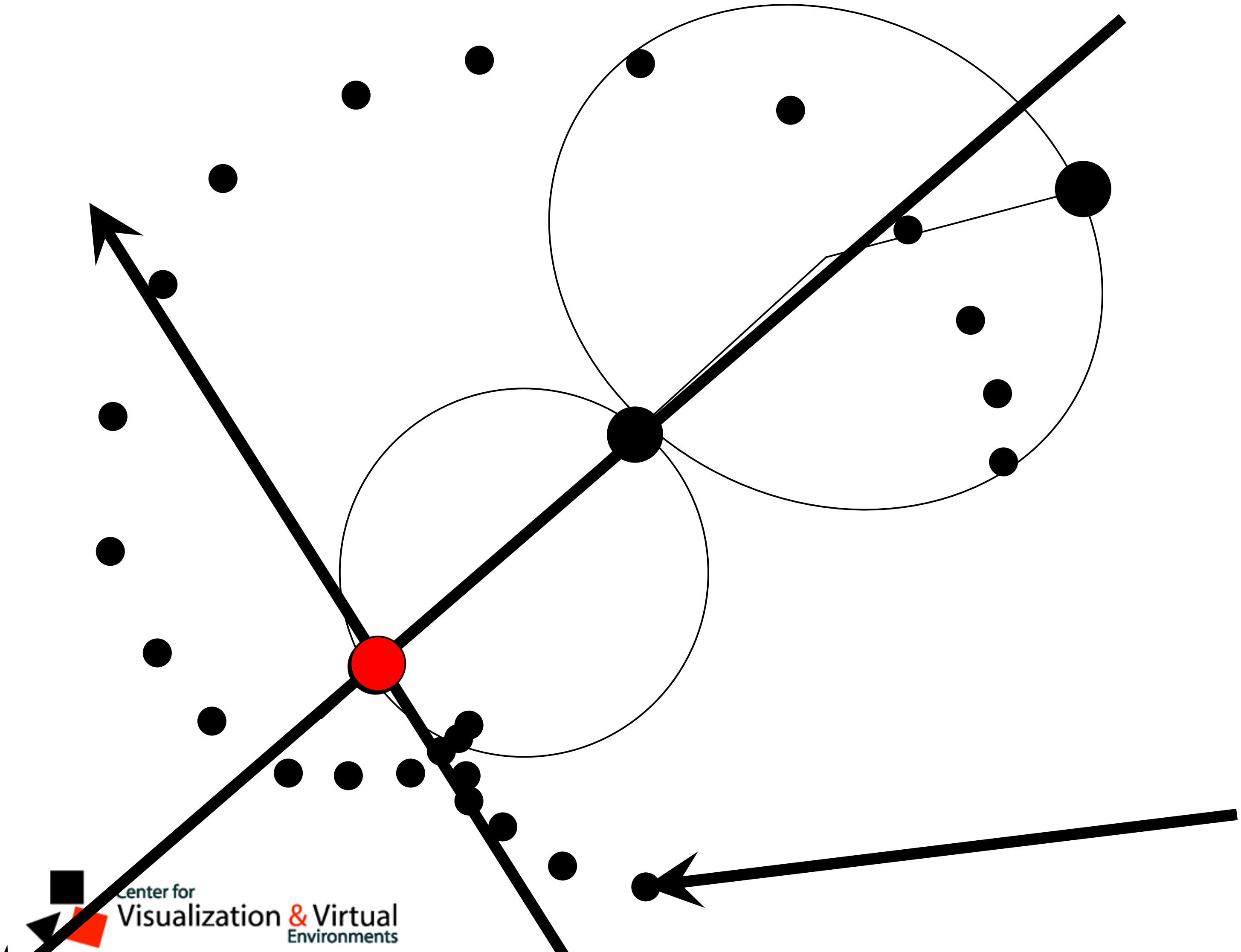


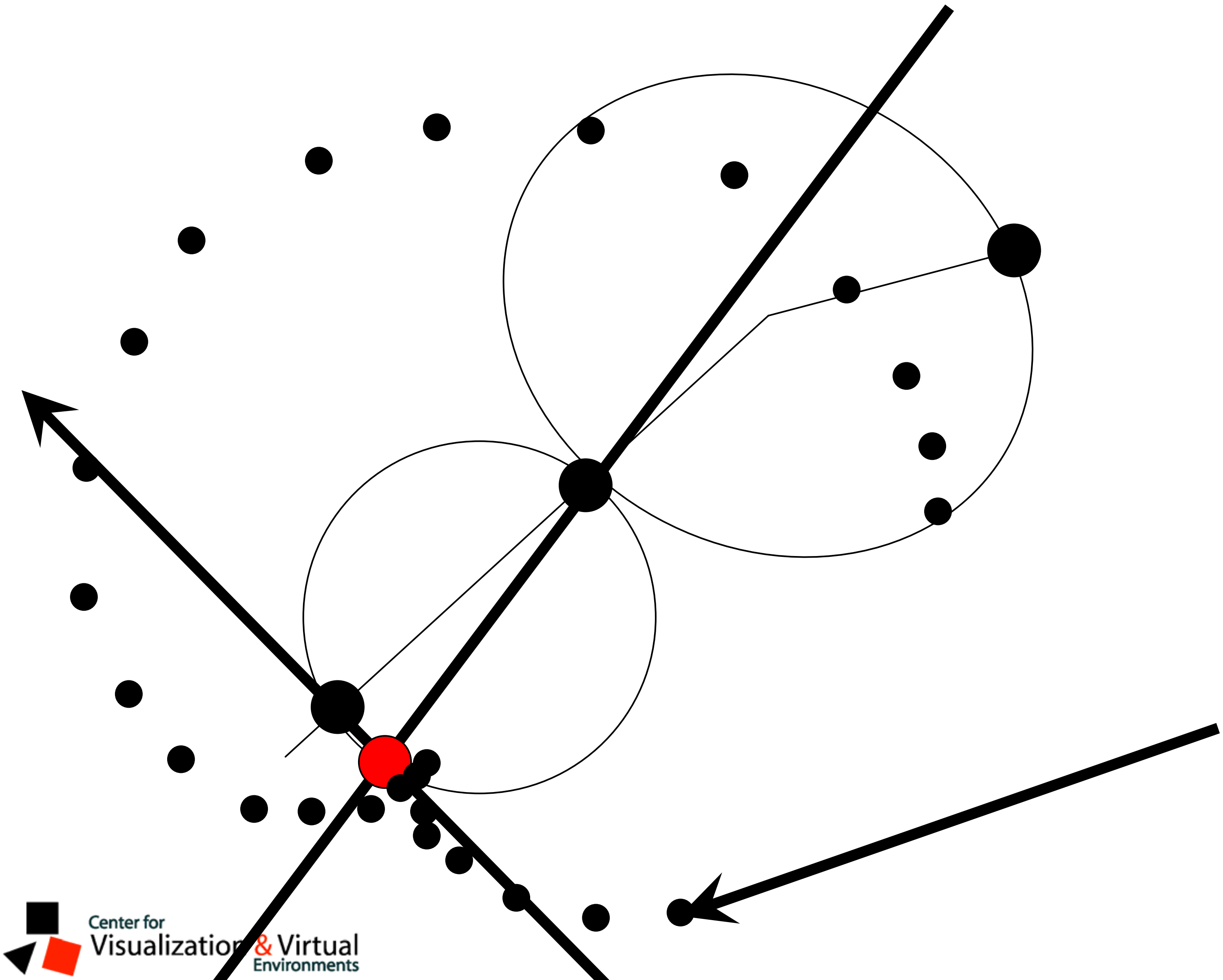


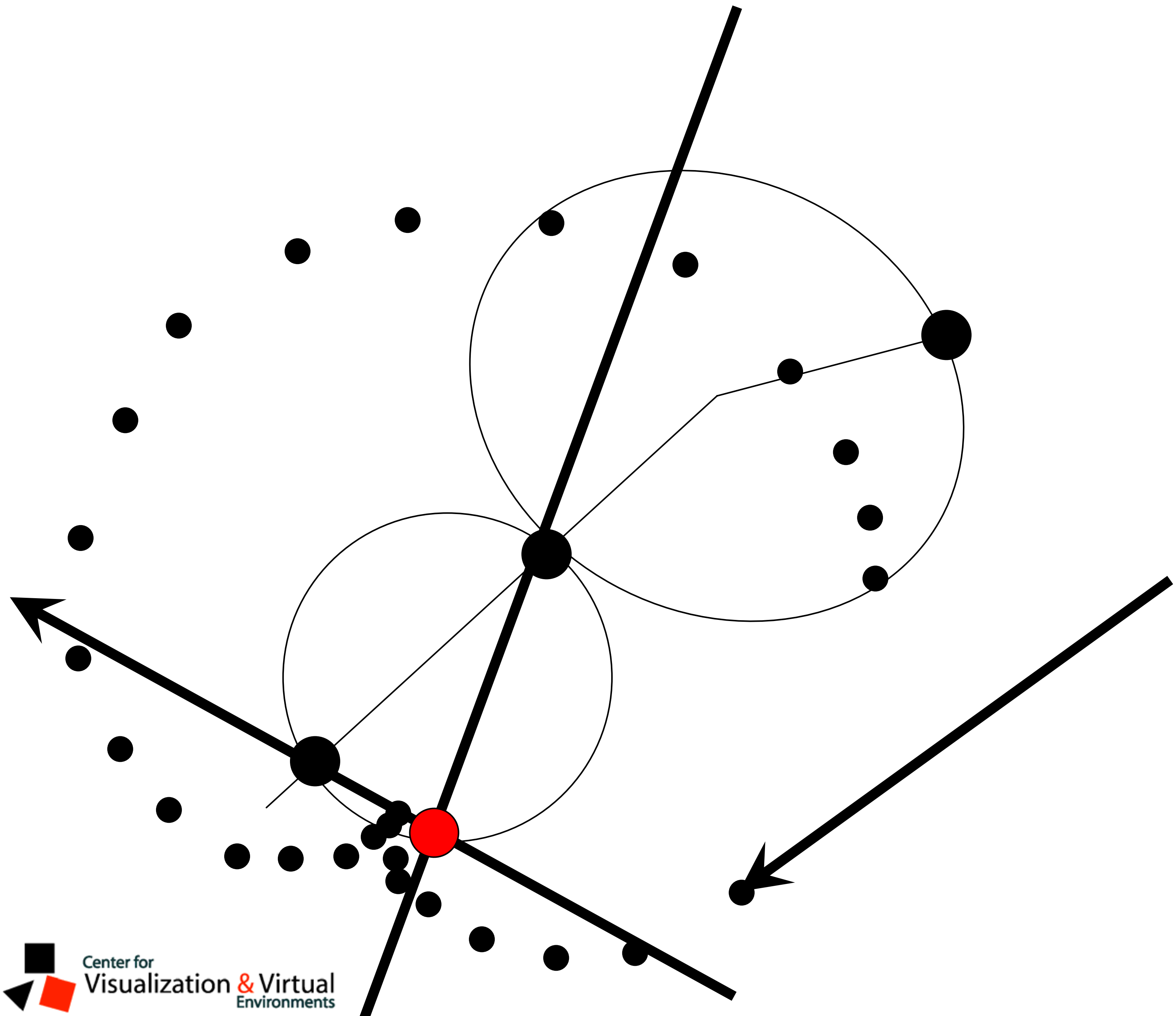


UK Center for Visualization & Virtual Environments



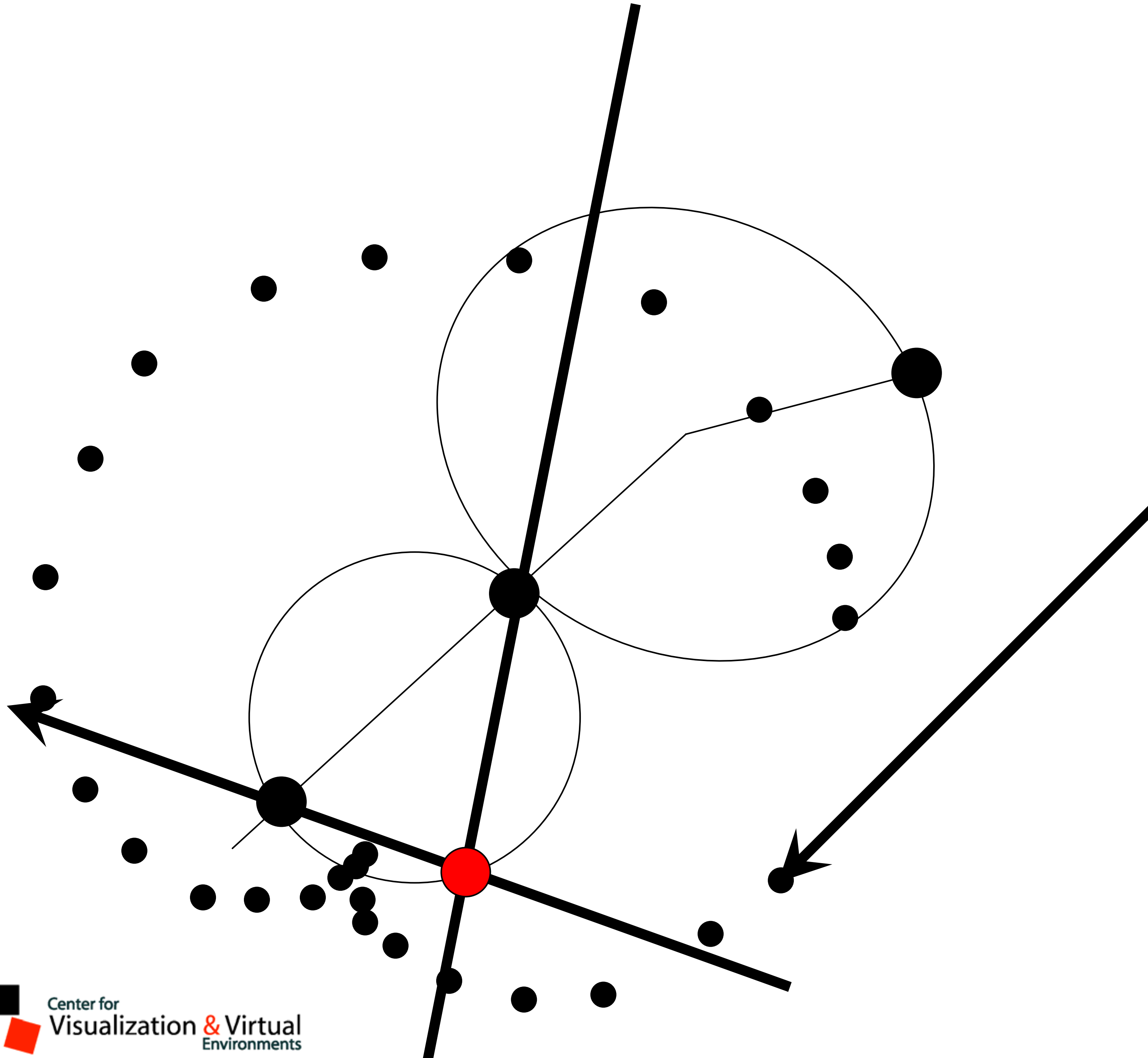


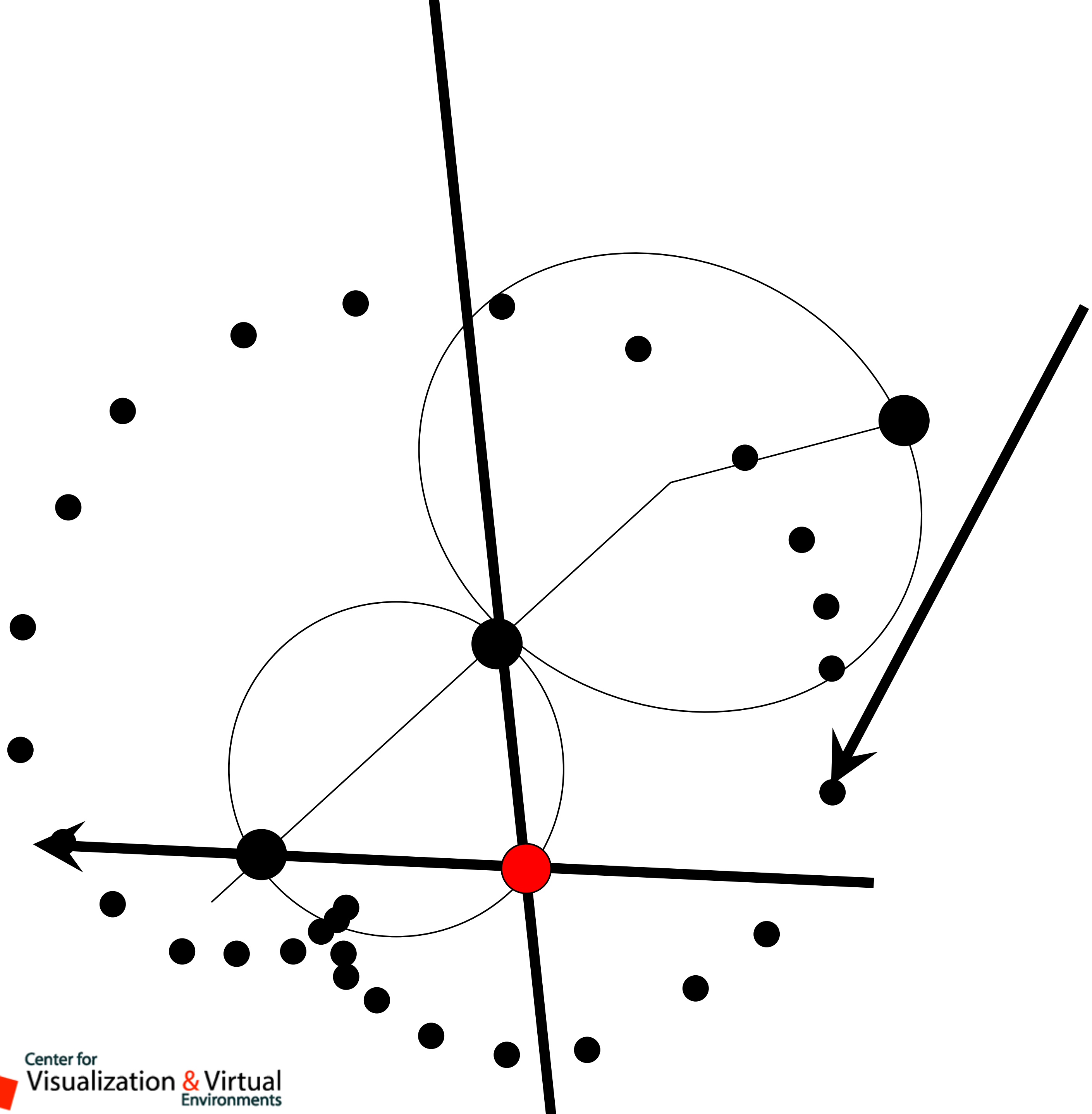




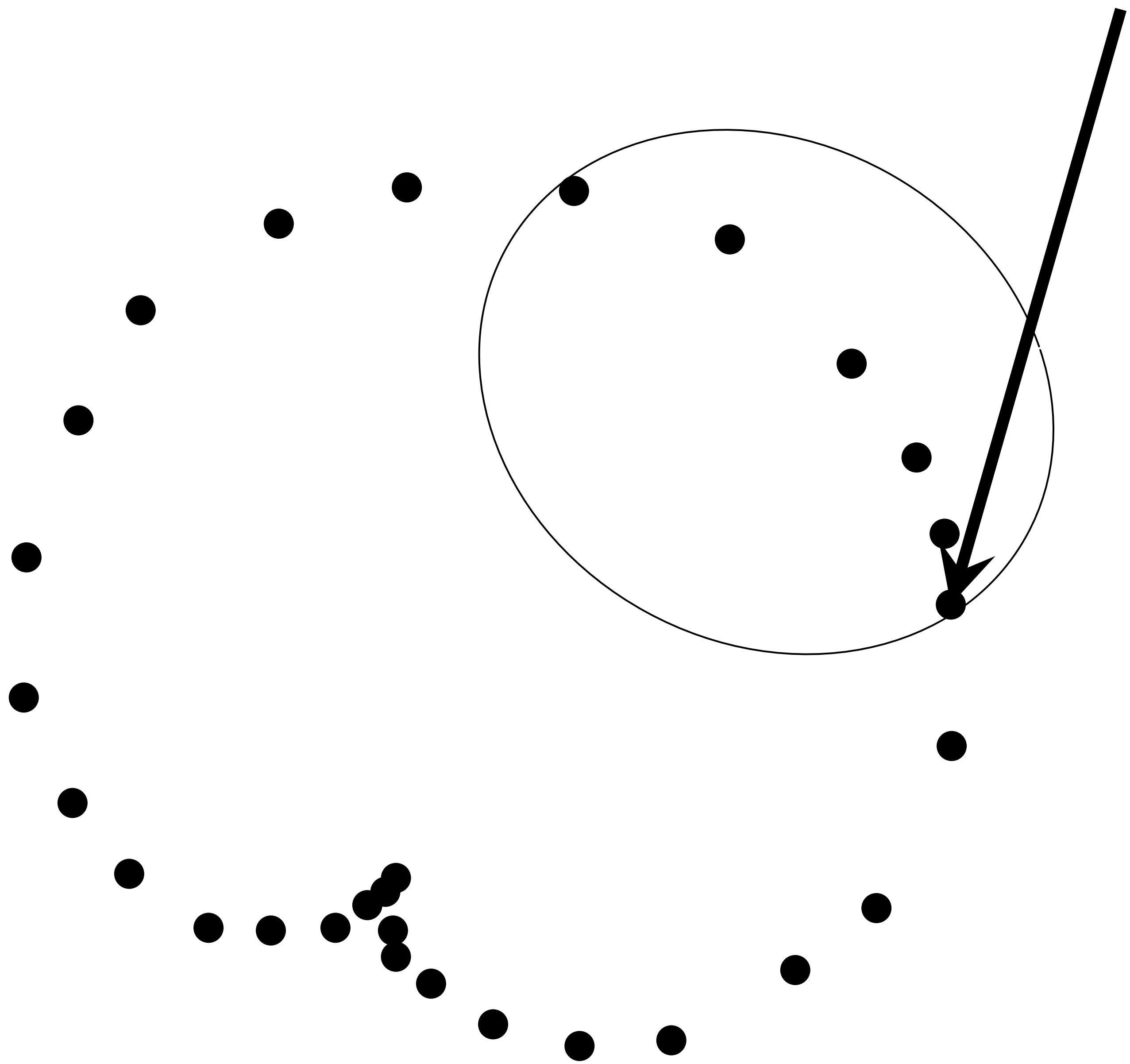
UK

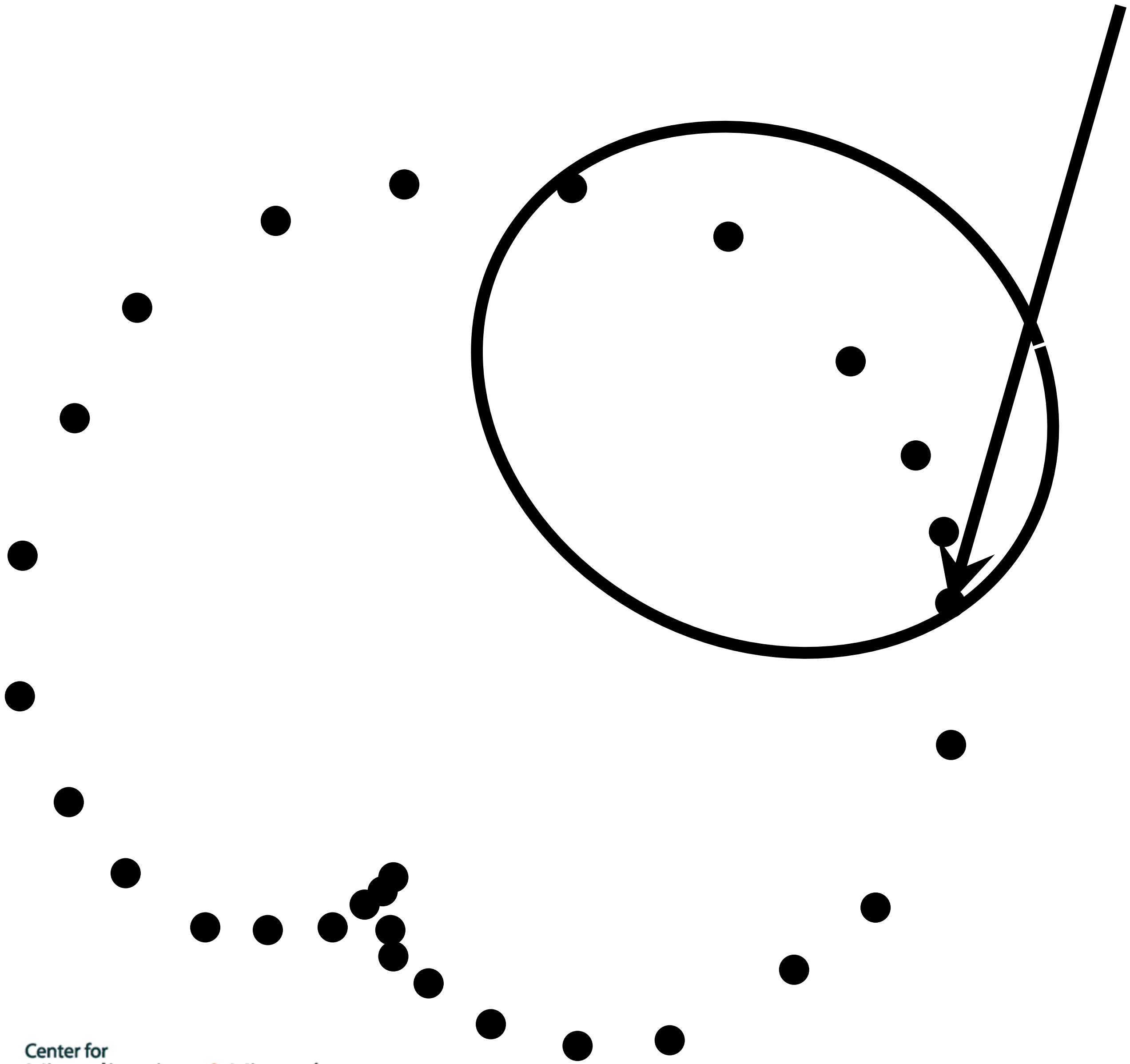
Center for Visualization & Virtual Environments

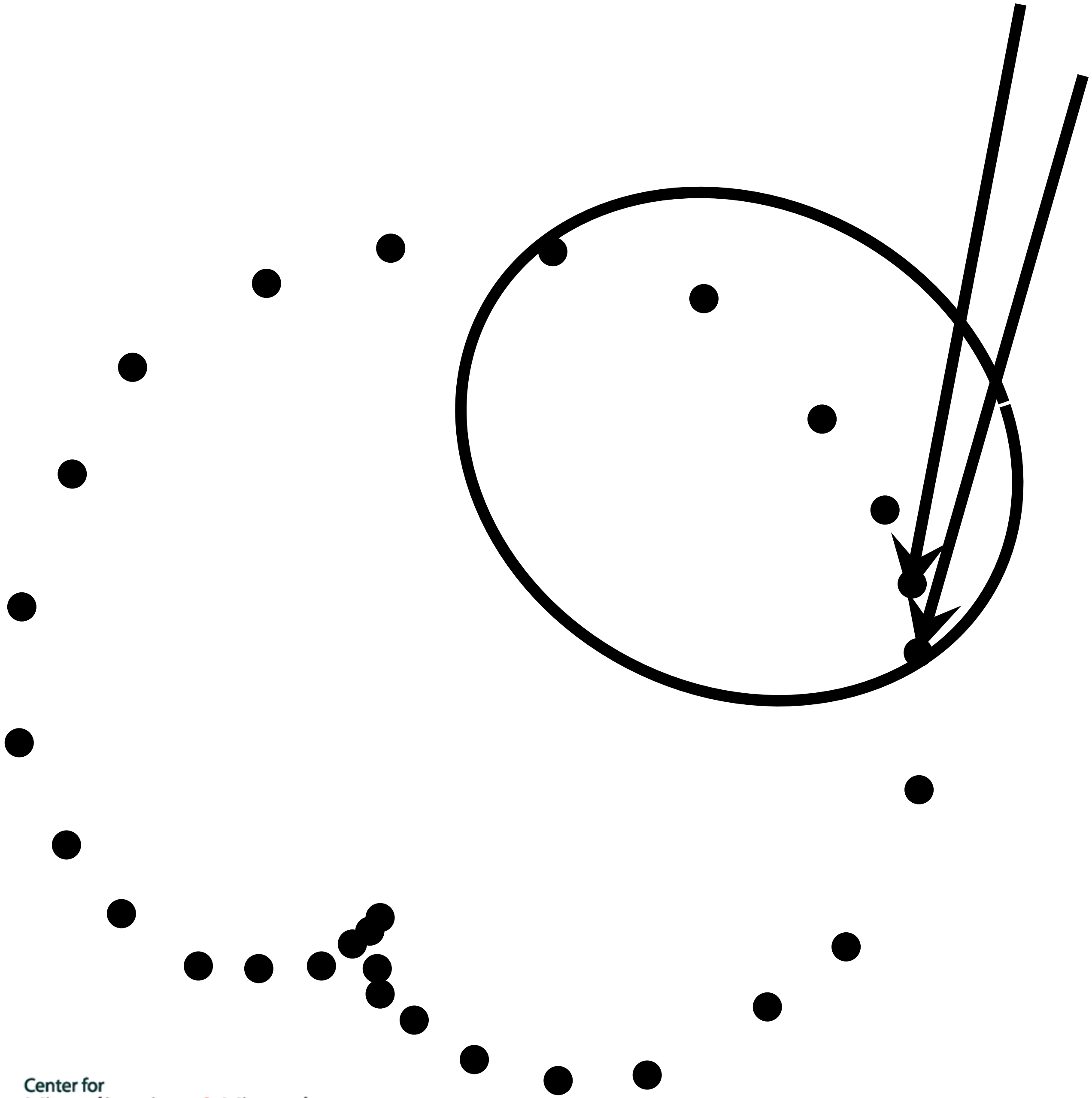


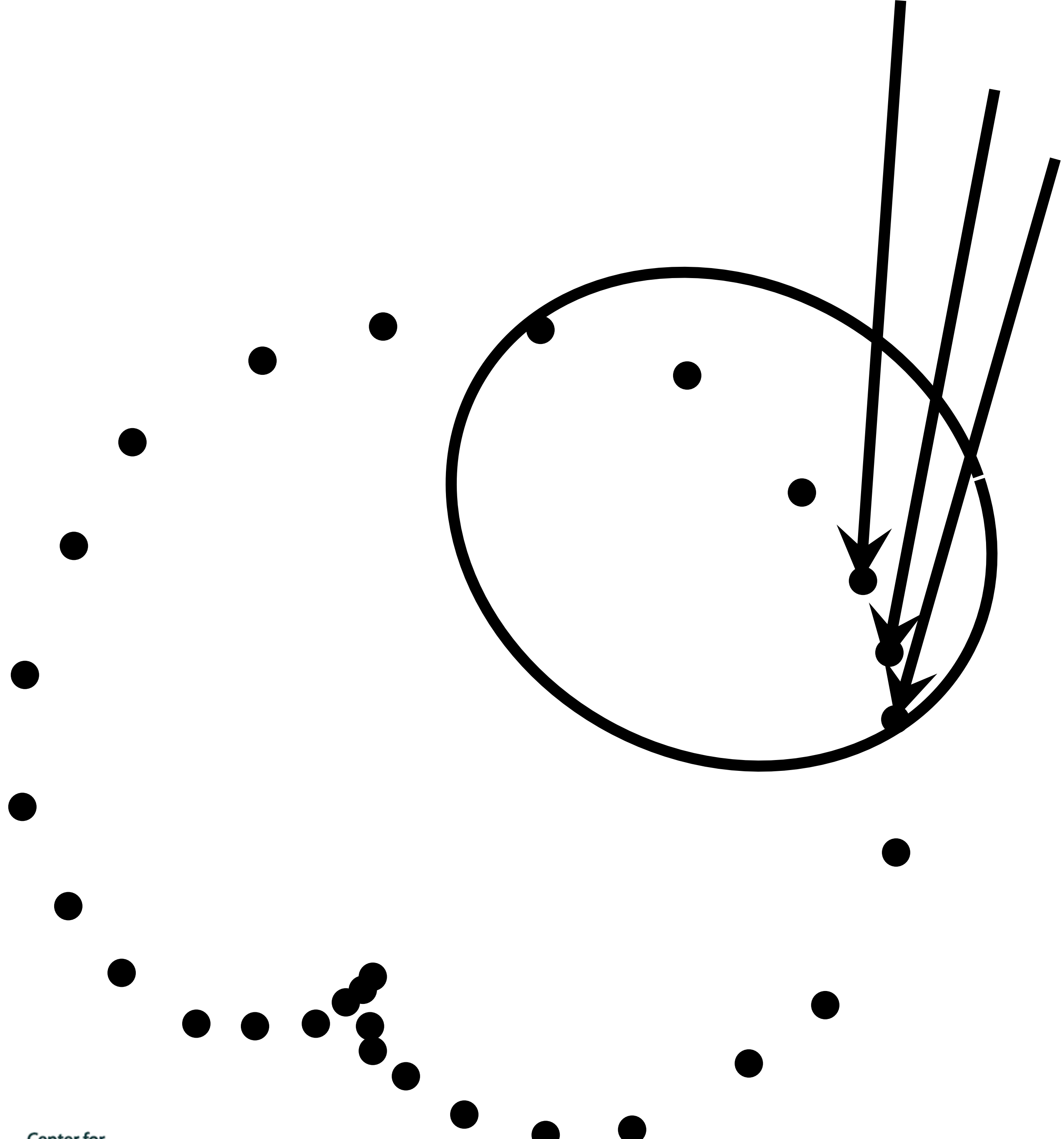


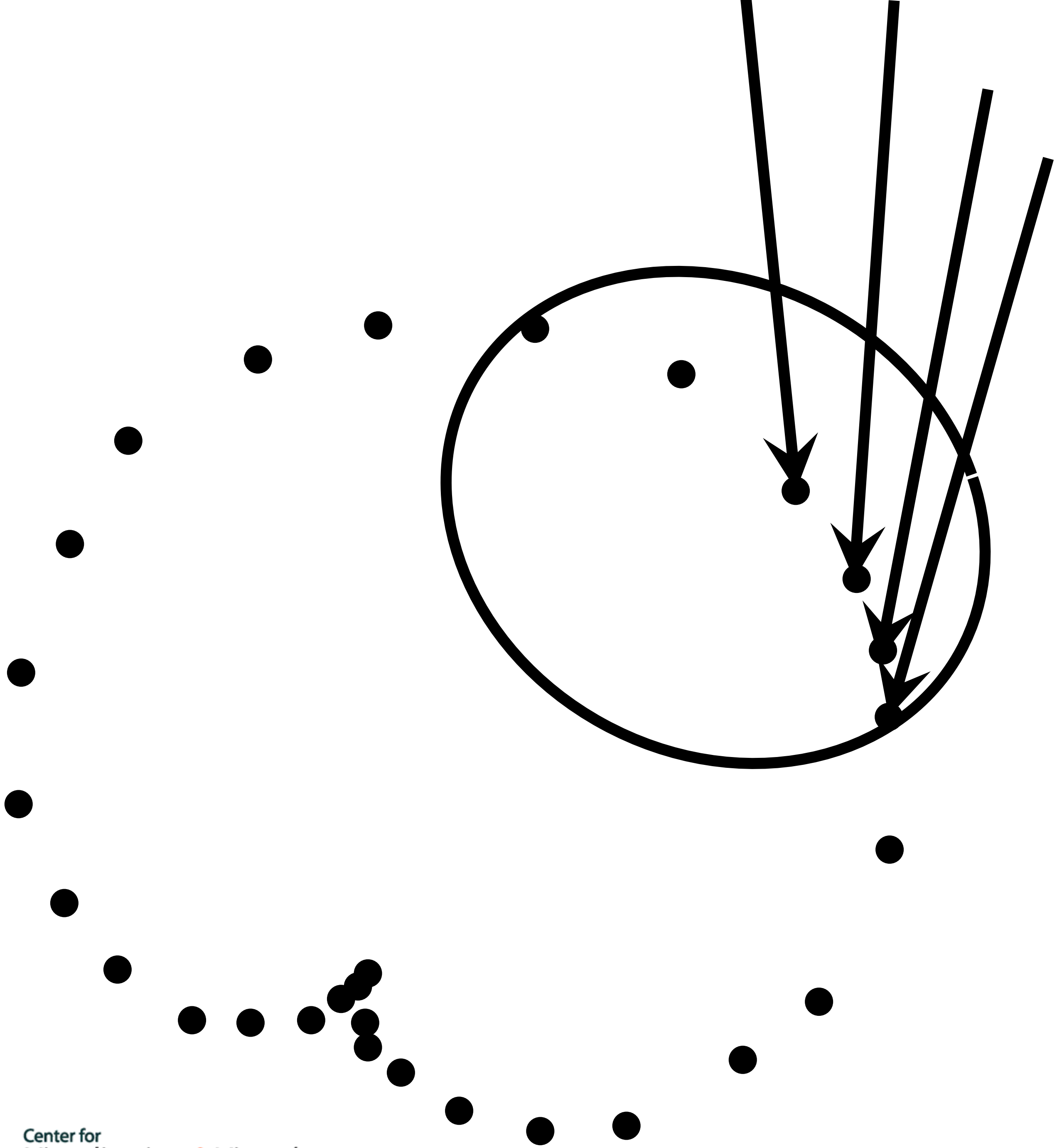
UK Center for Visualization & Virtual Environments





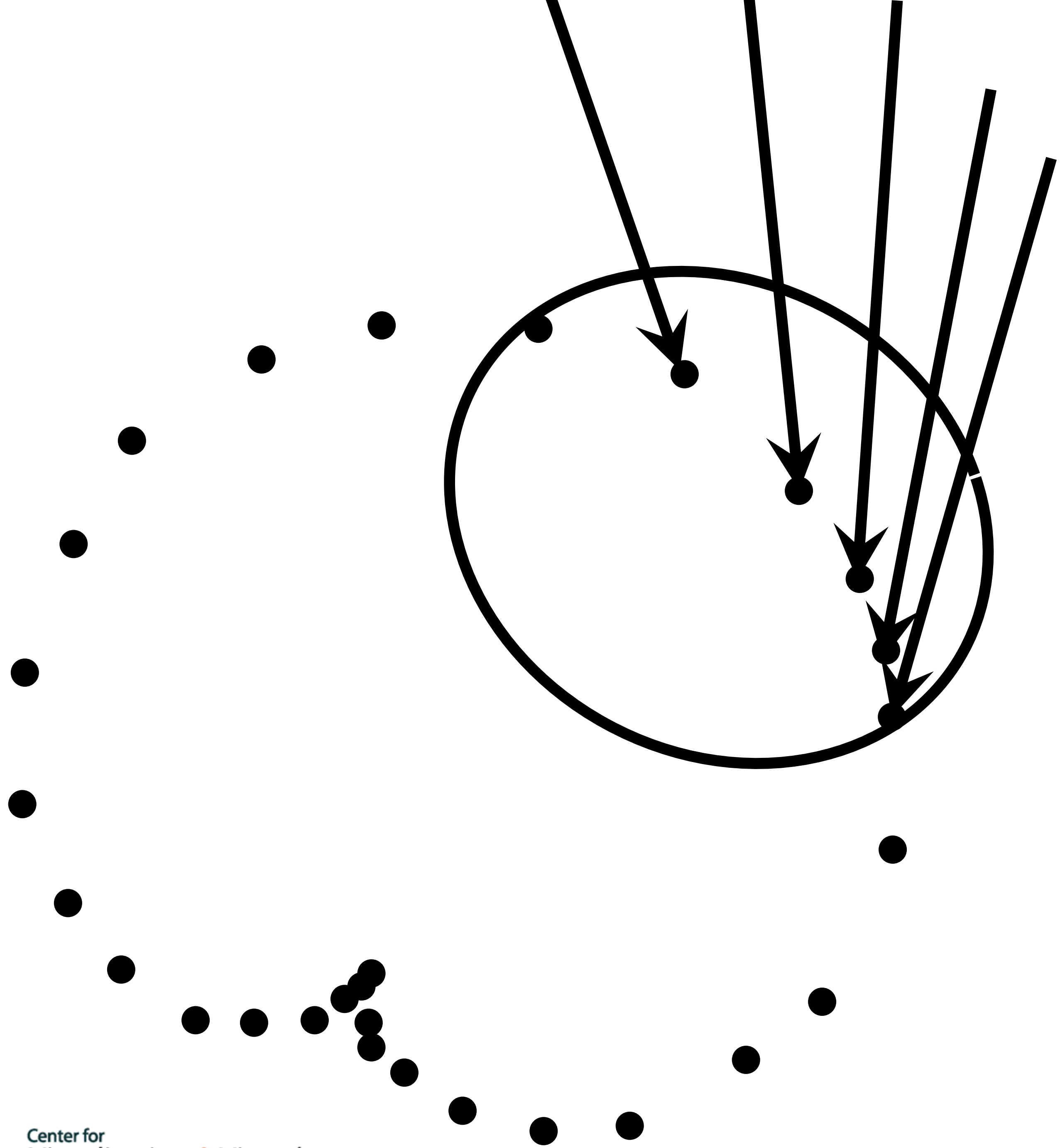




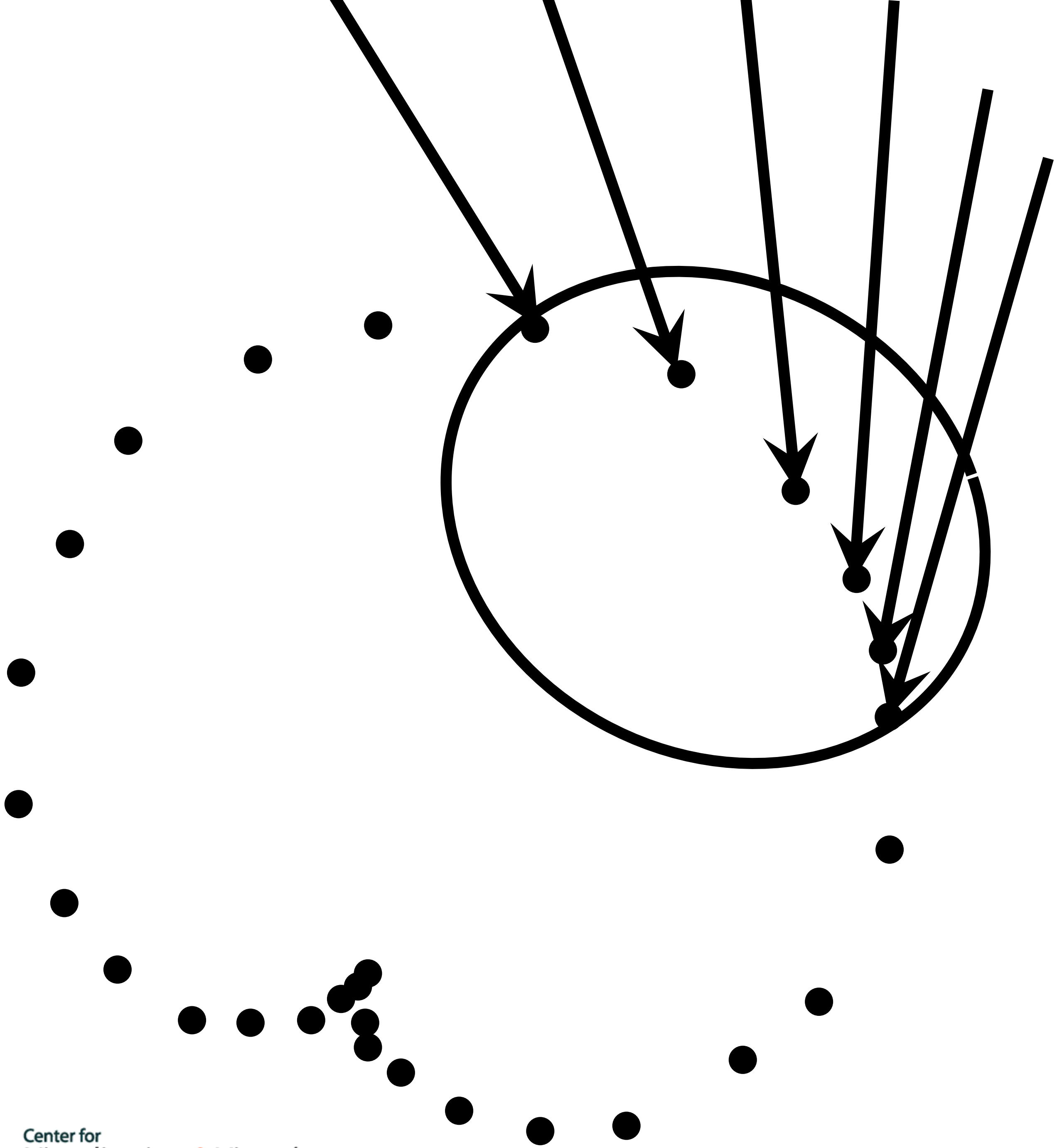


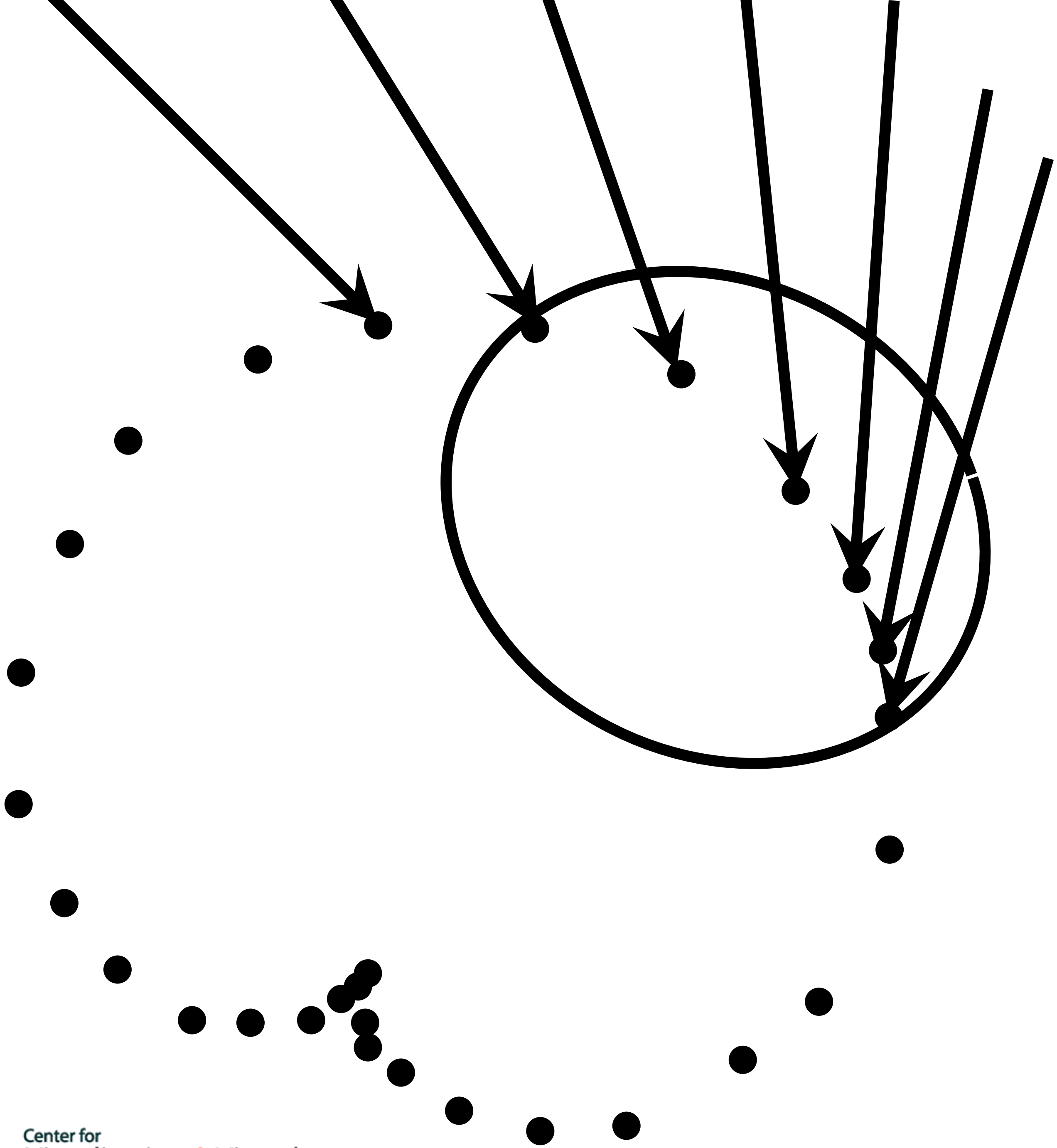
UK

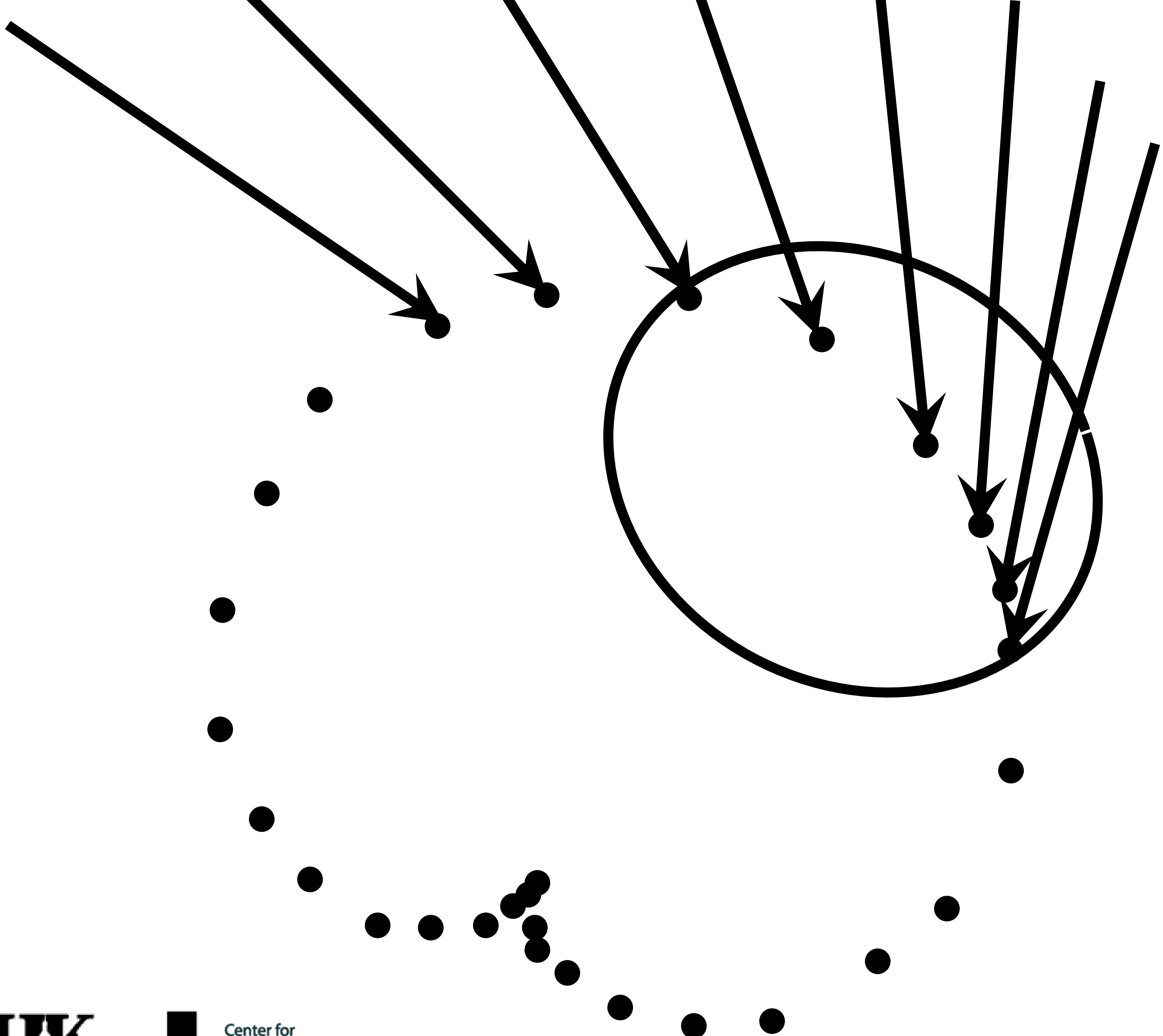
Center for Visualization & Virtual Environments

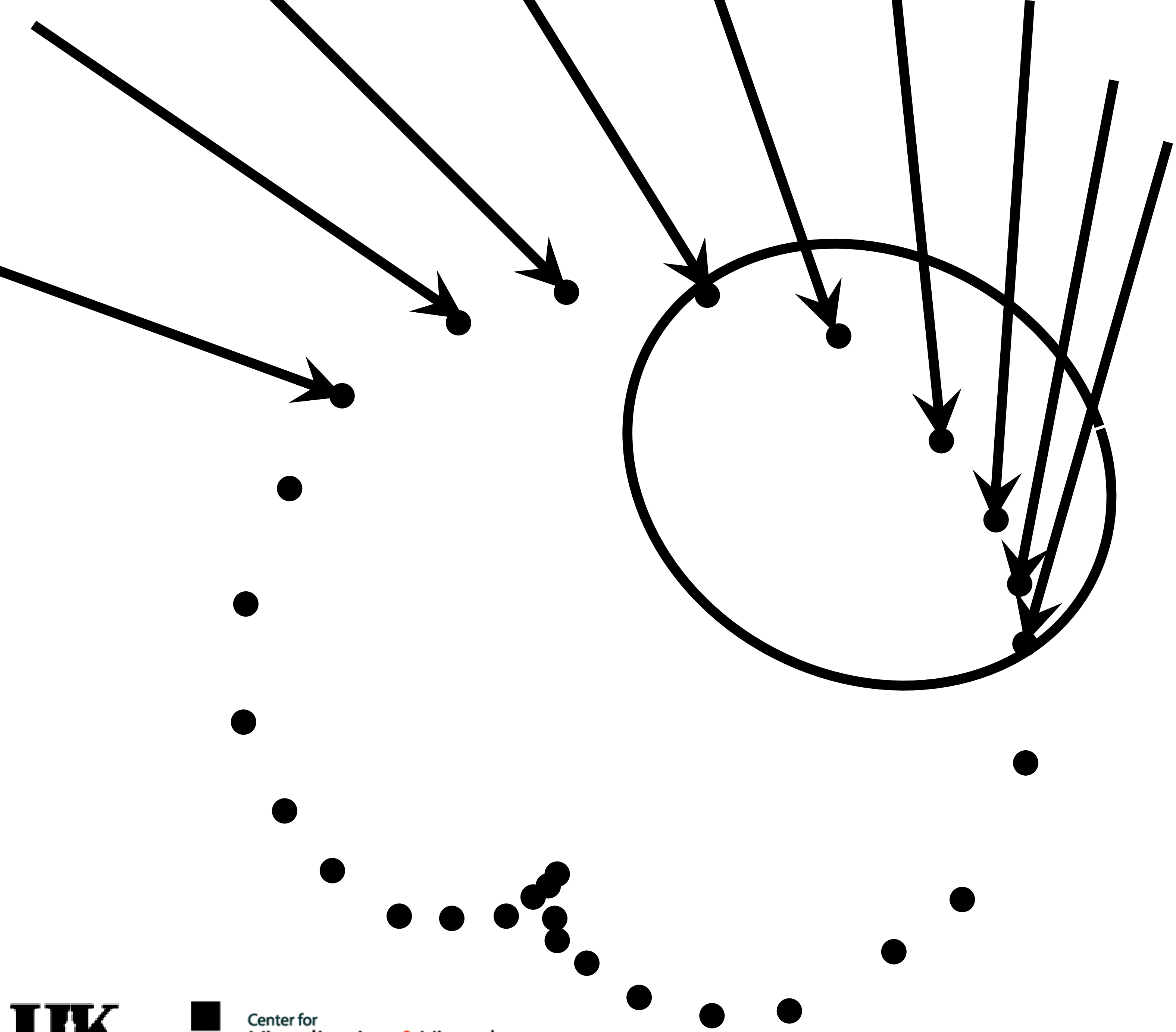


UK Center for Visualization & Virtual Environments

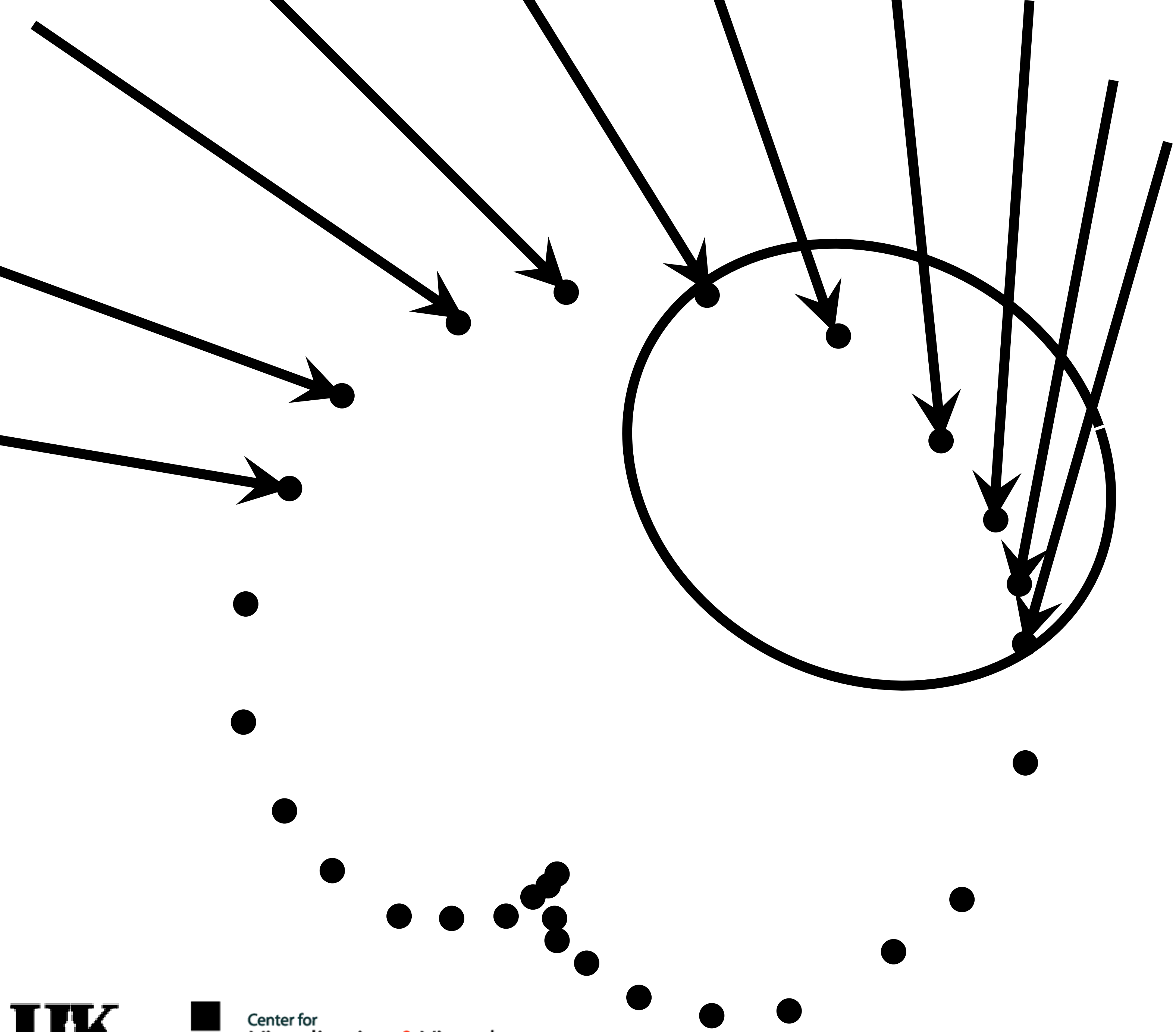


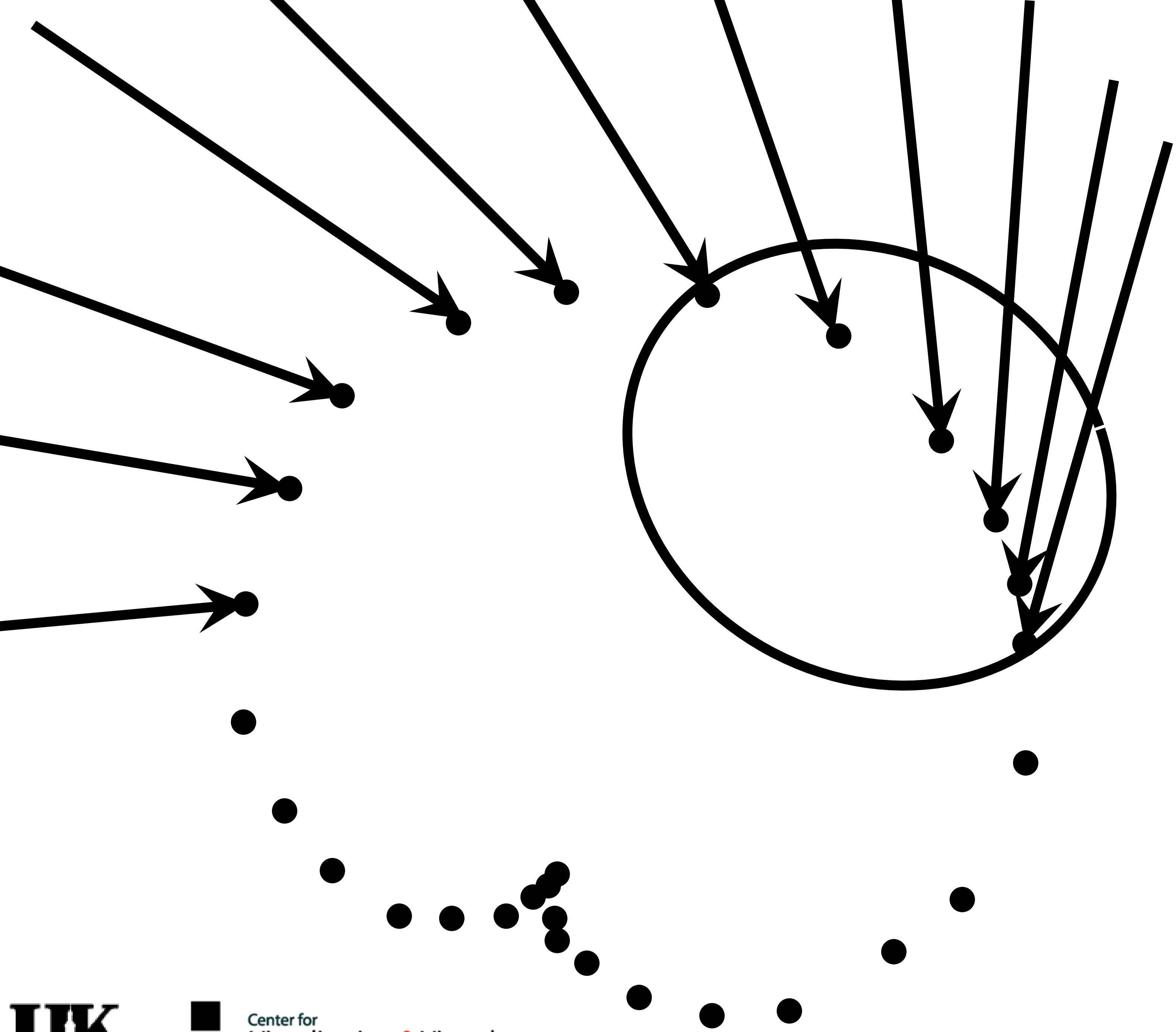


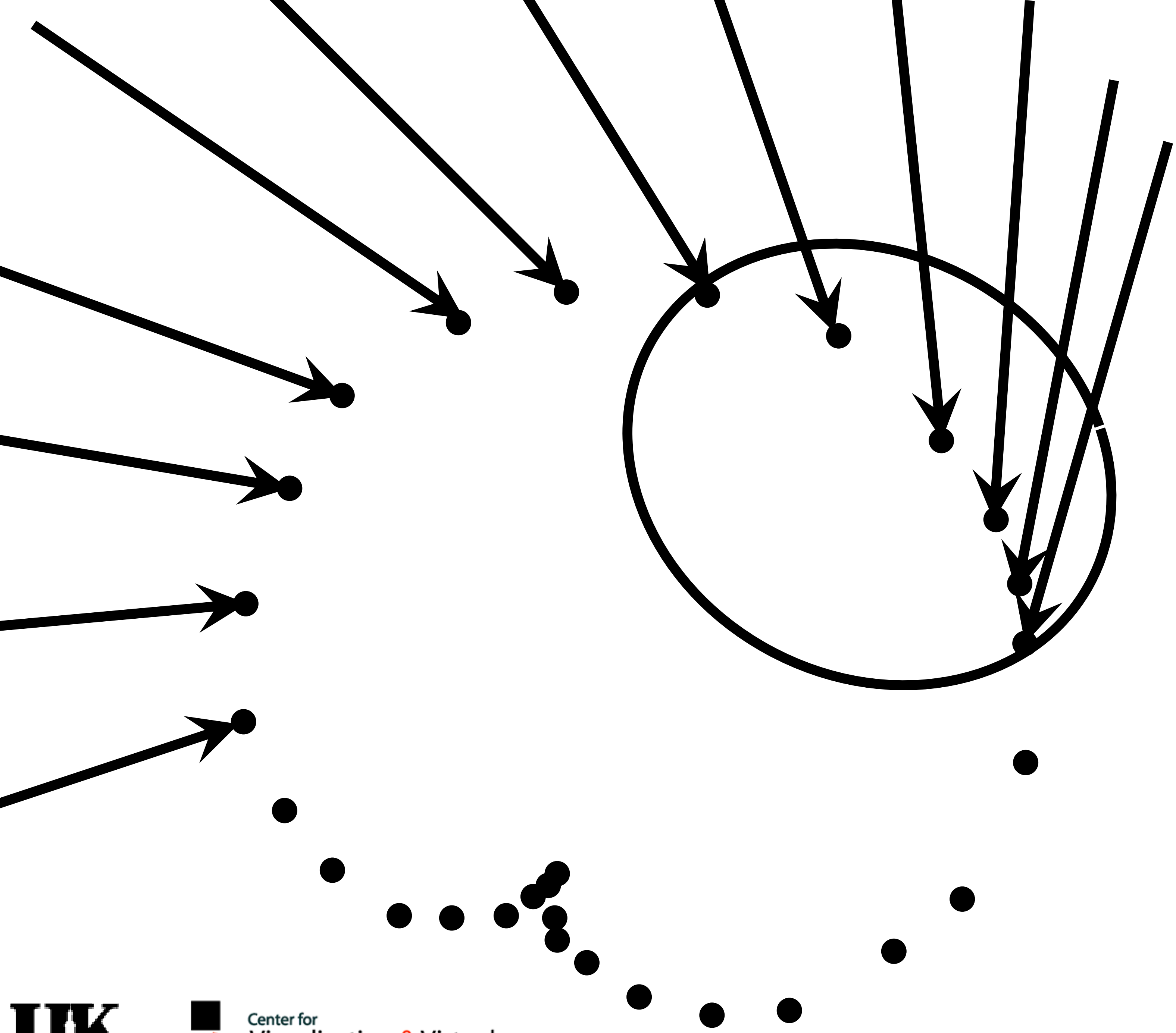




UK Center for Visualization & Virtual Environments

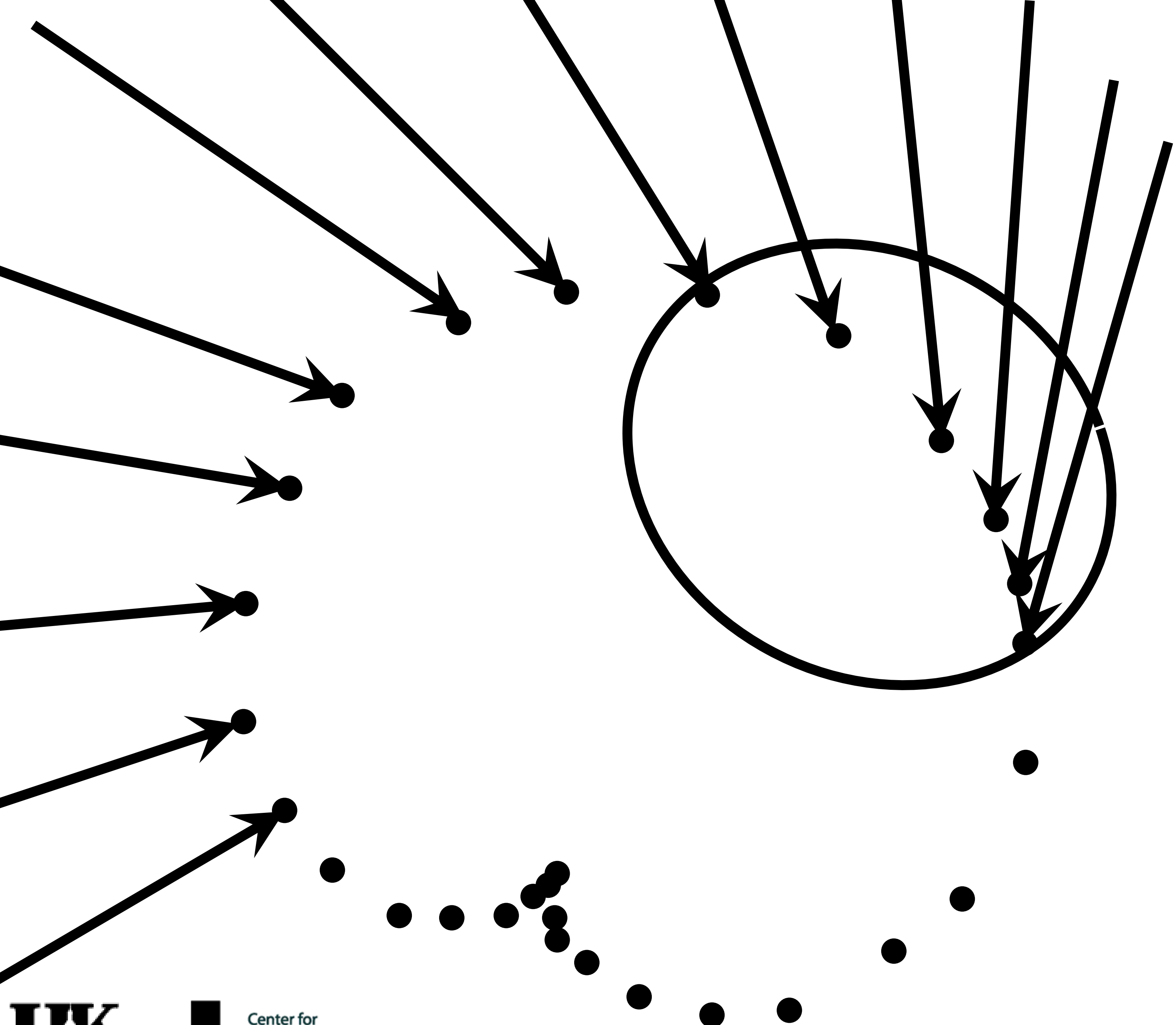






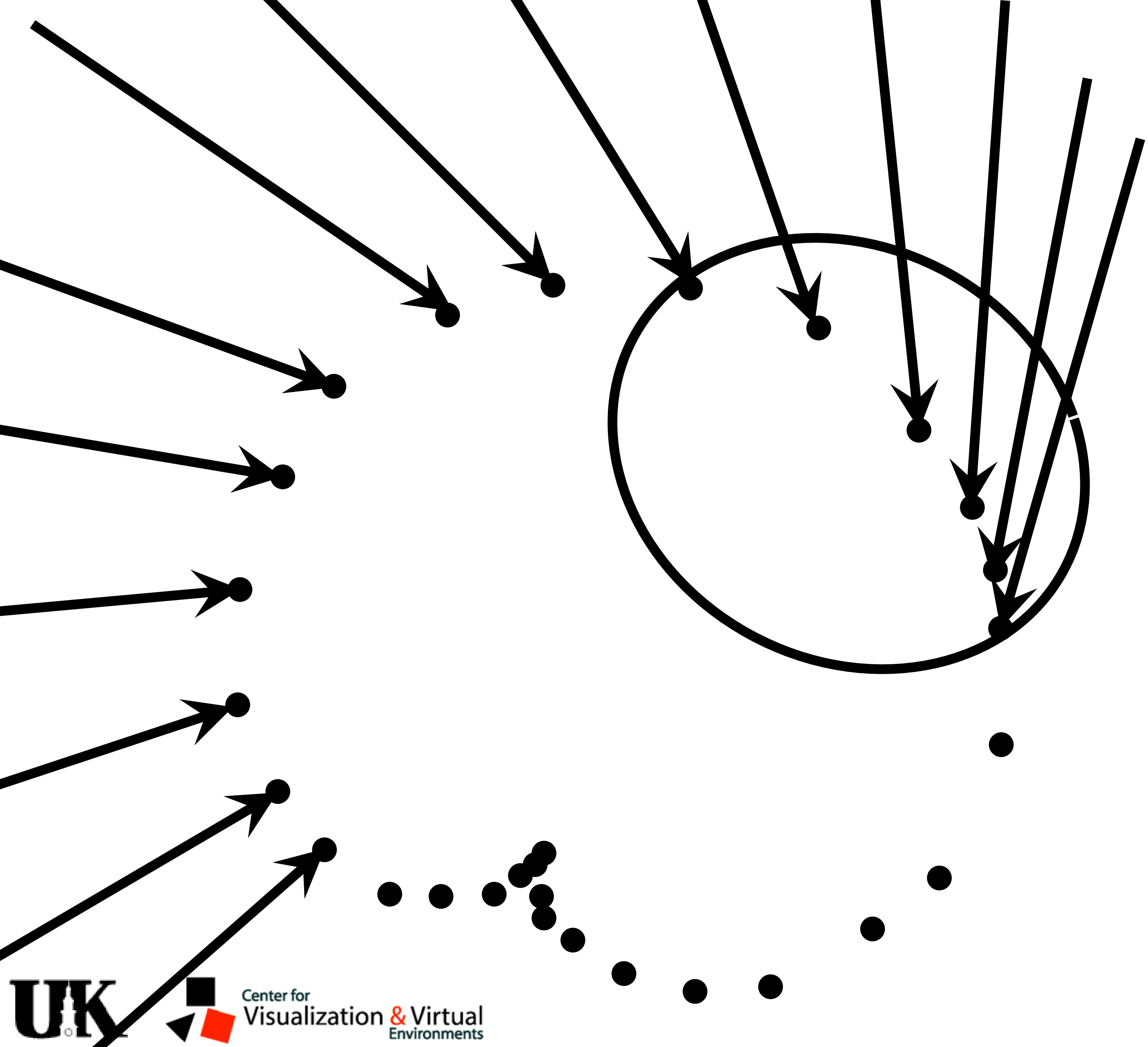
UK

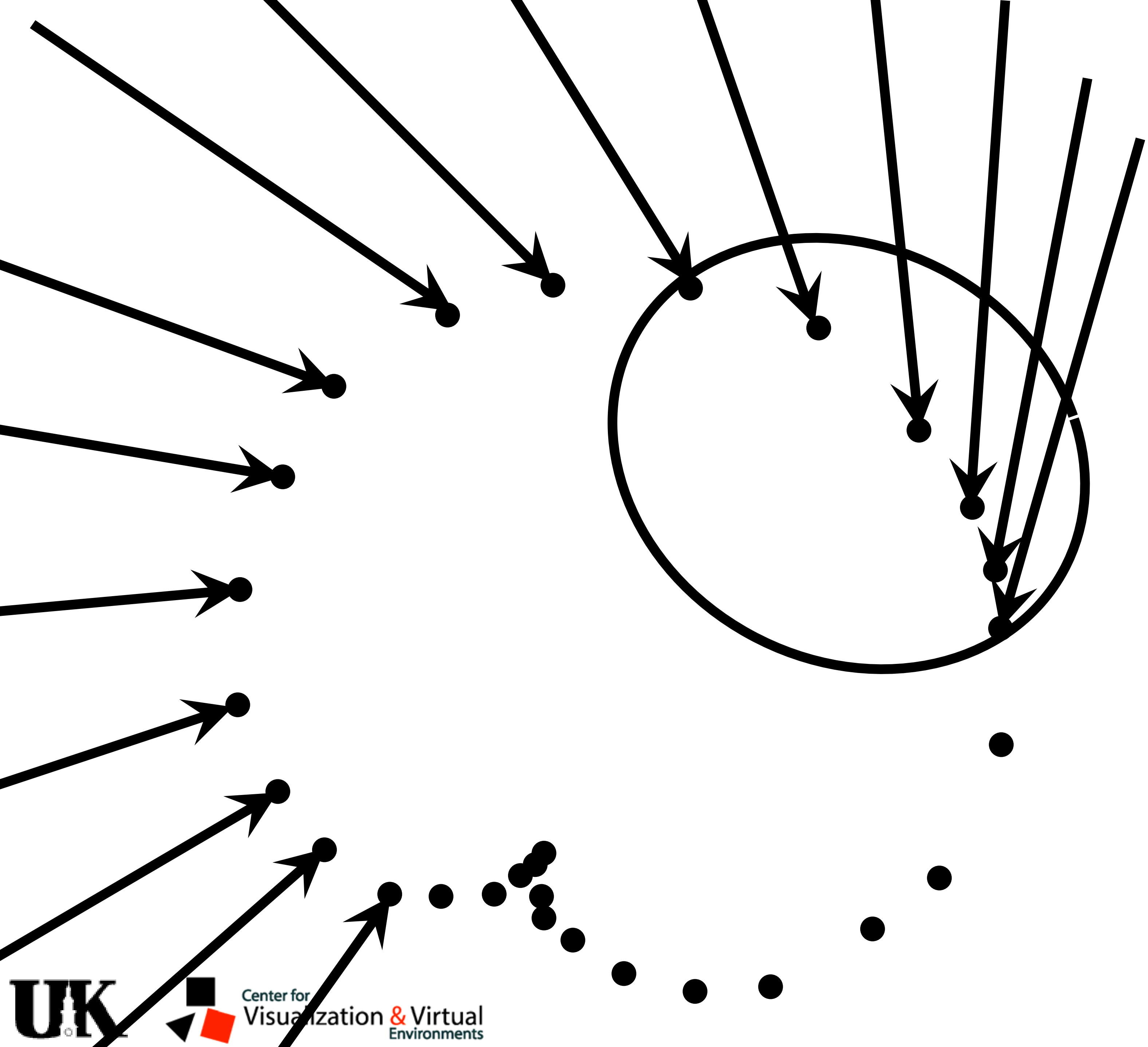
Center for Visualization & Virtual Environments



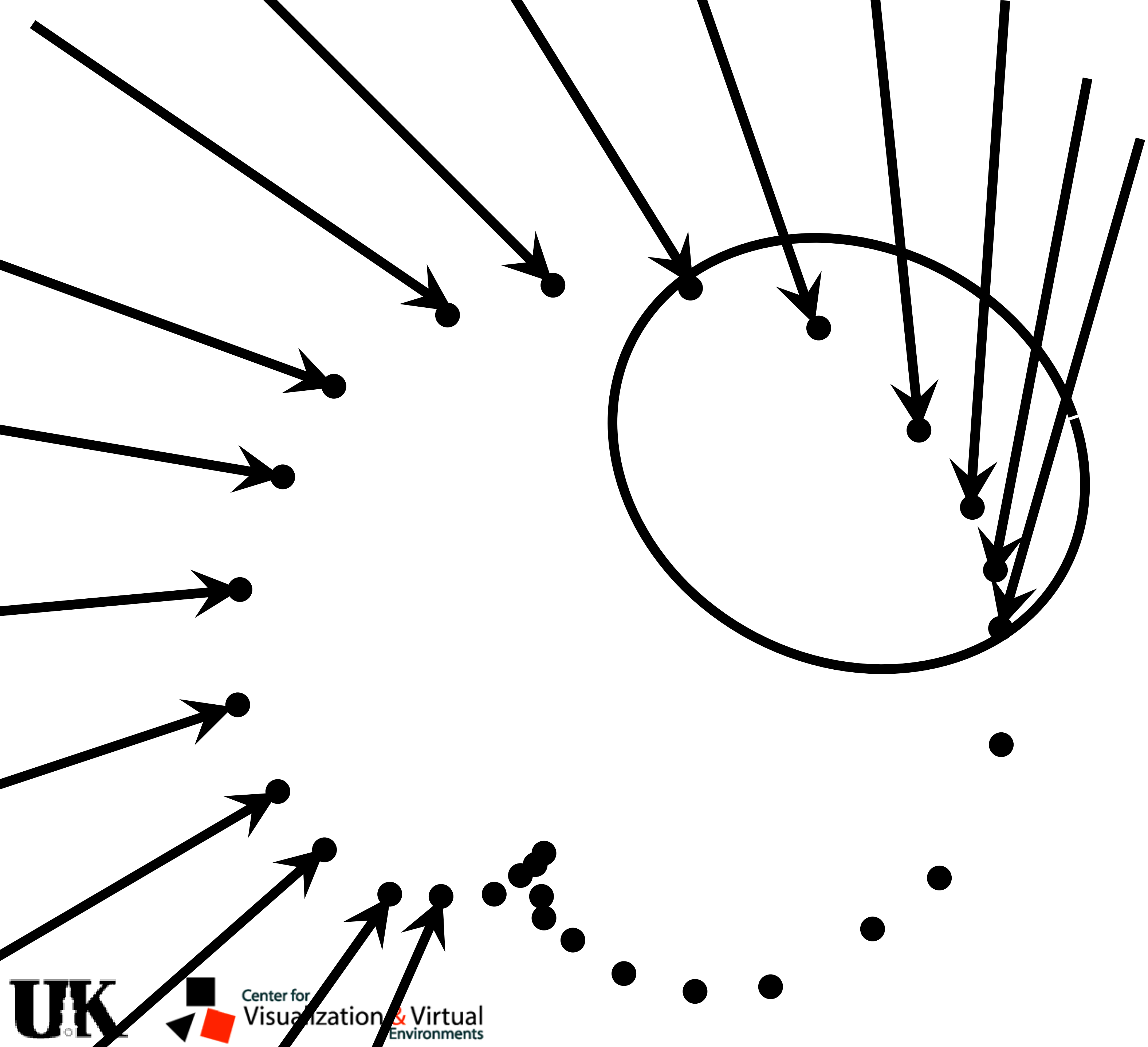
UK

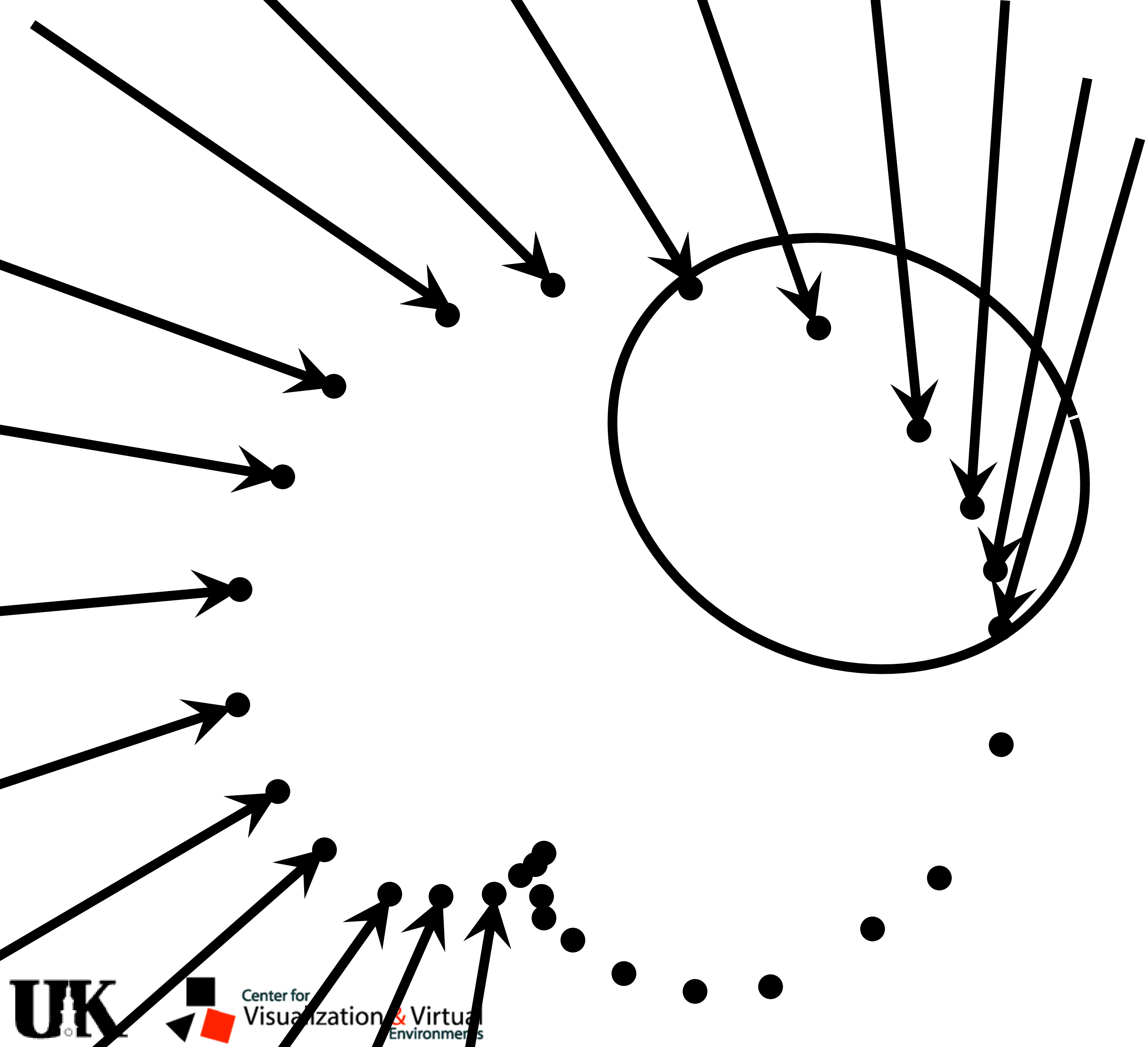
Center for Visualization & Virtual Environments



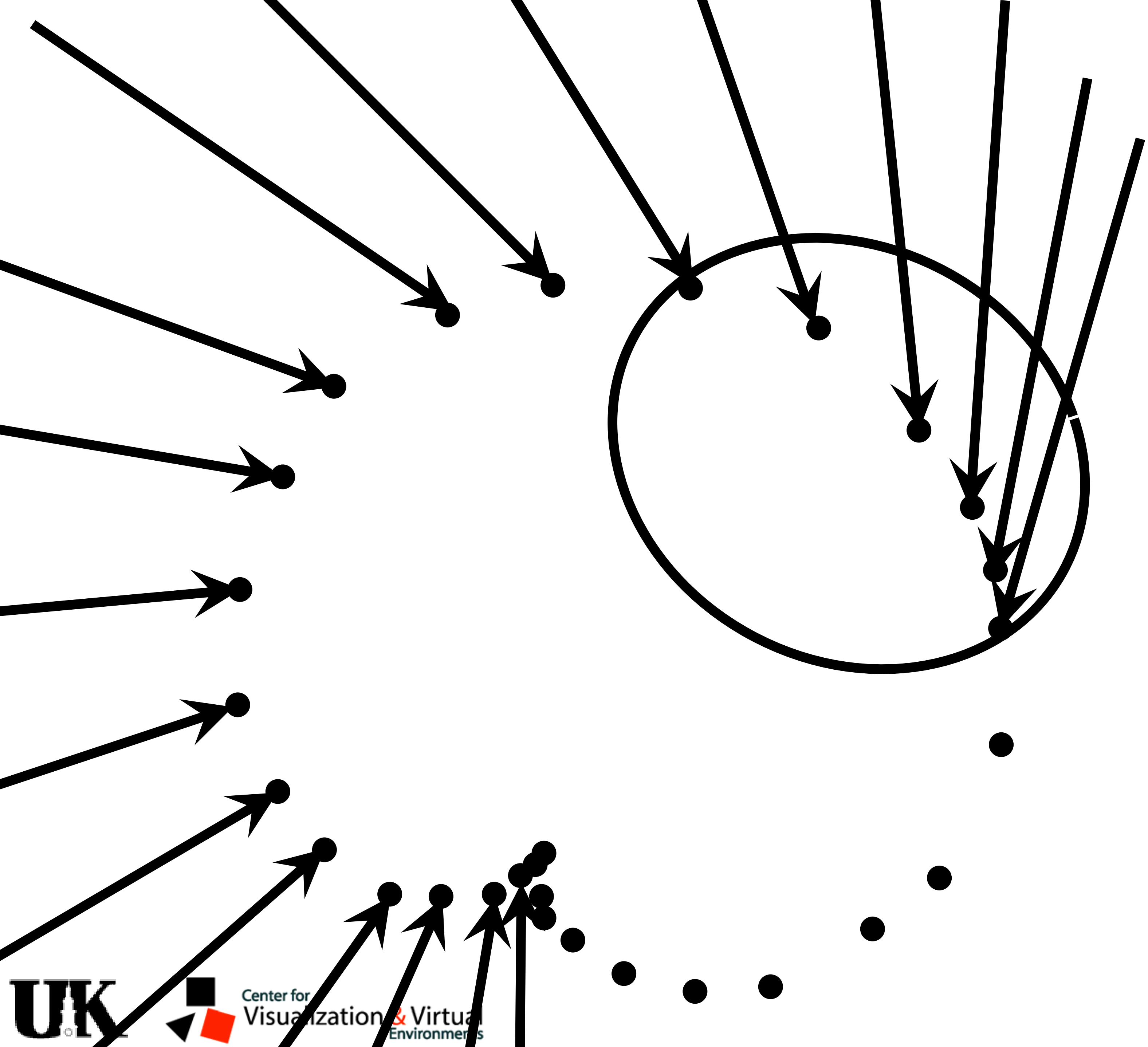


UK Center for Visualization & Virtual Environments

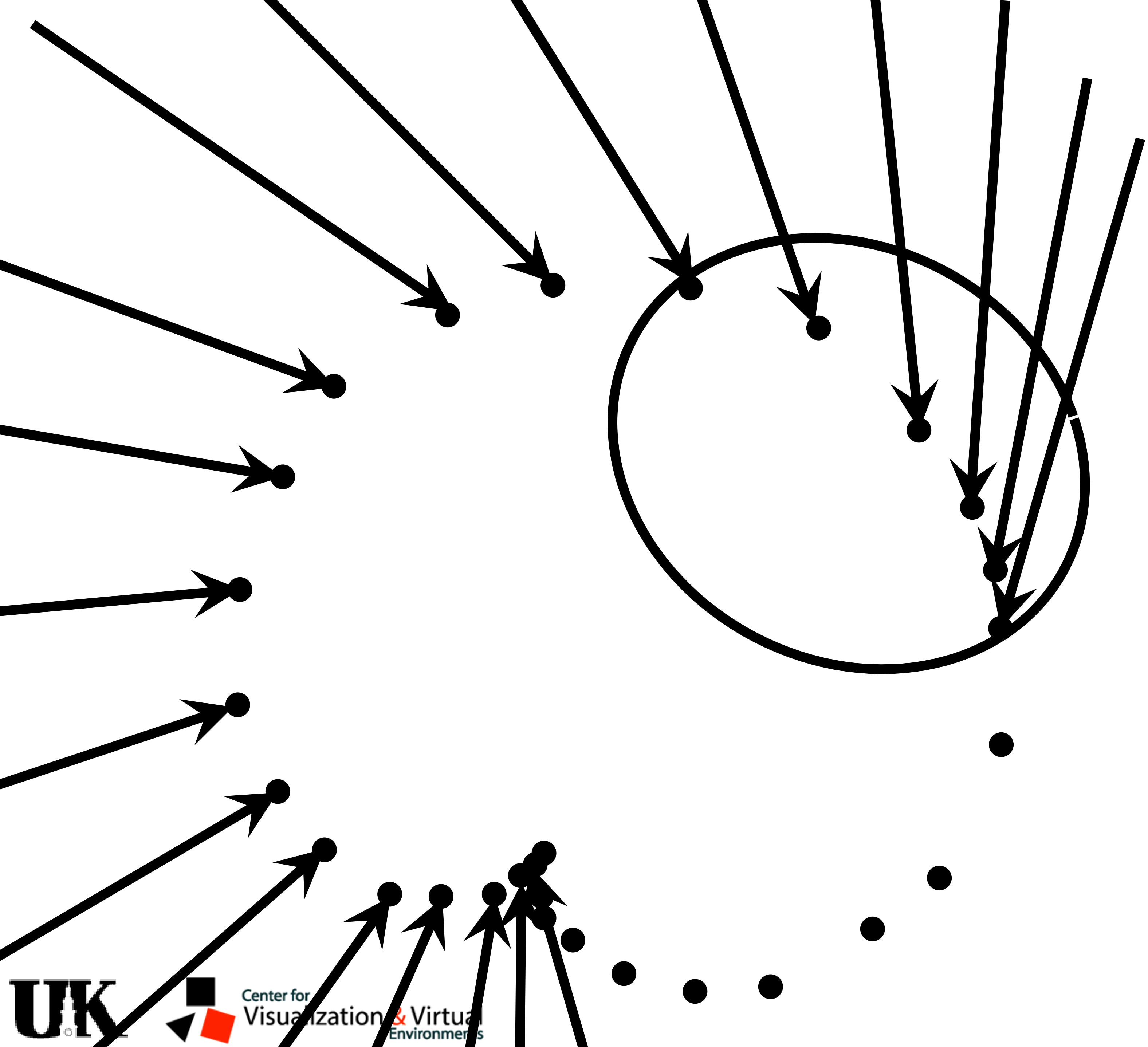


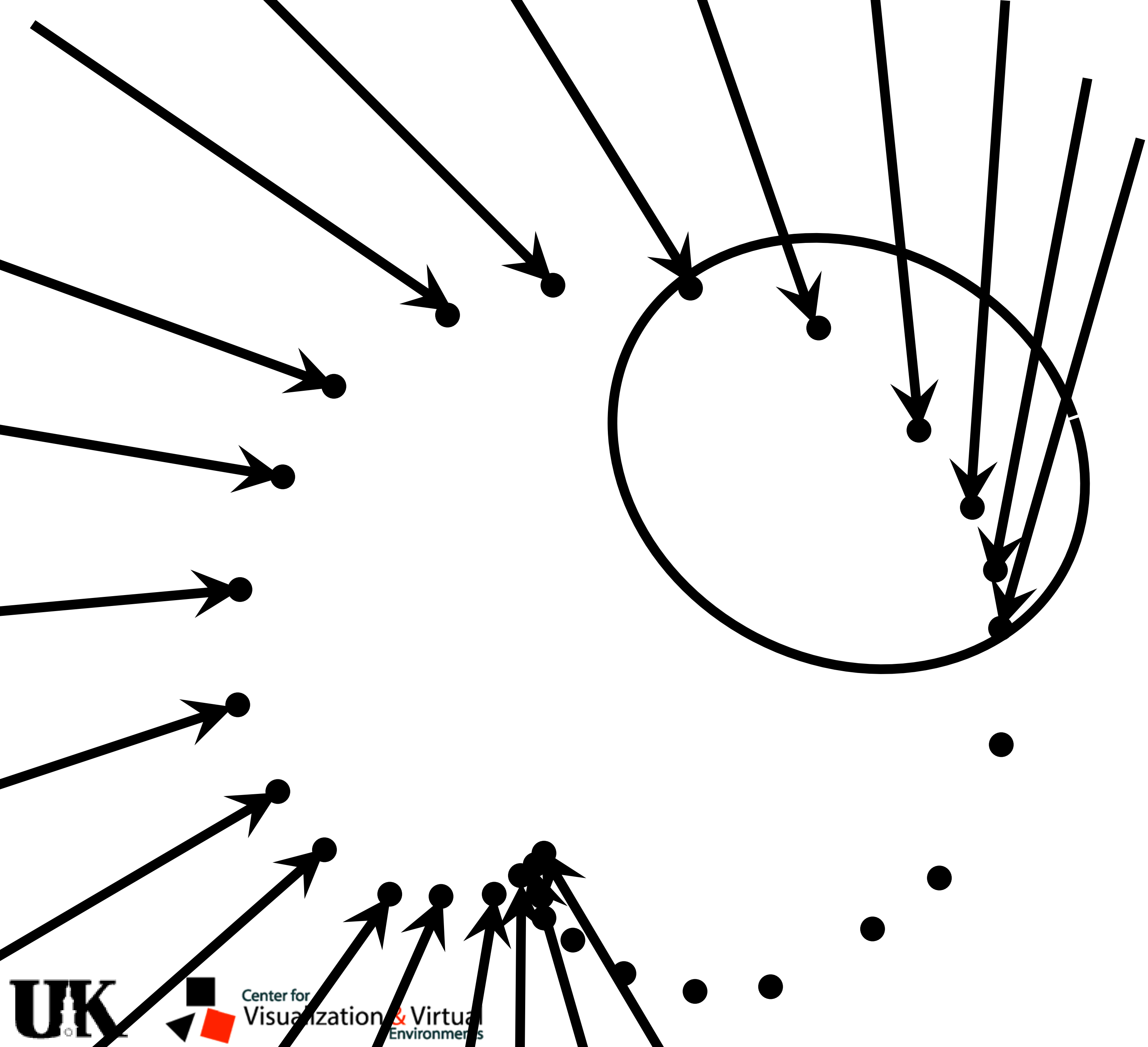


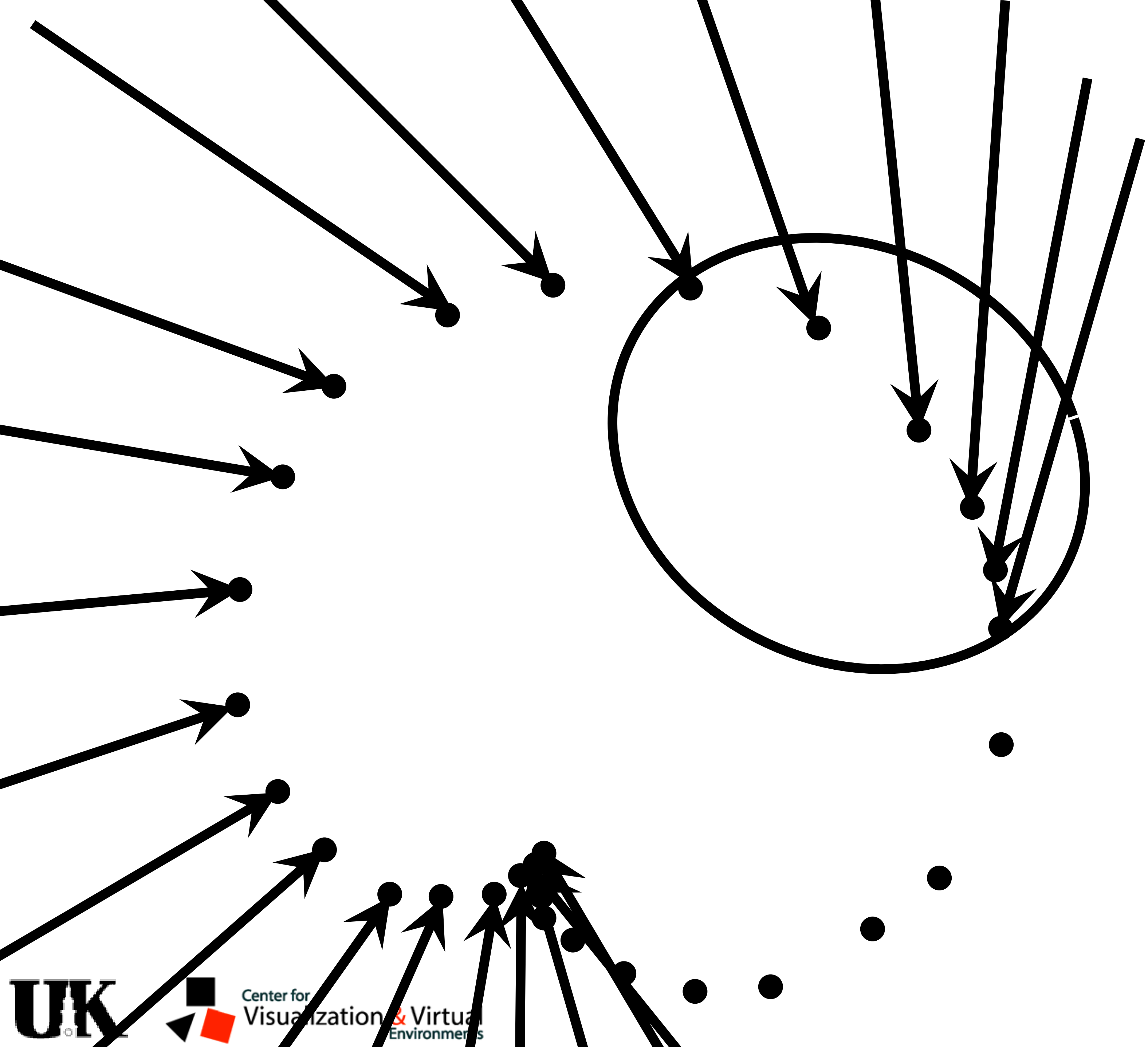
Center for Visualization & Virtual Environments

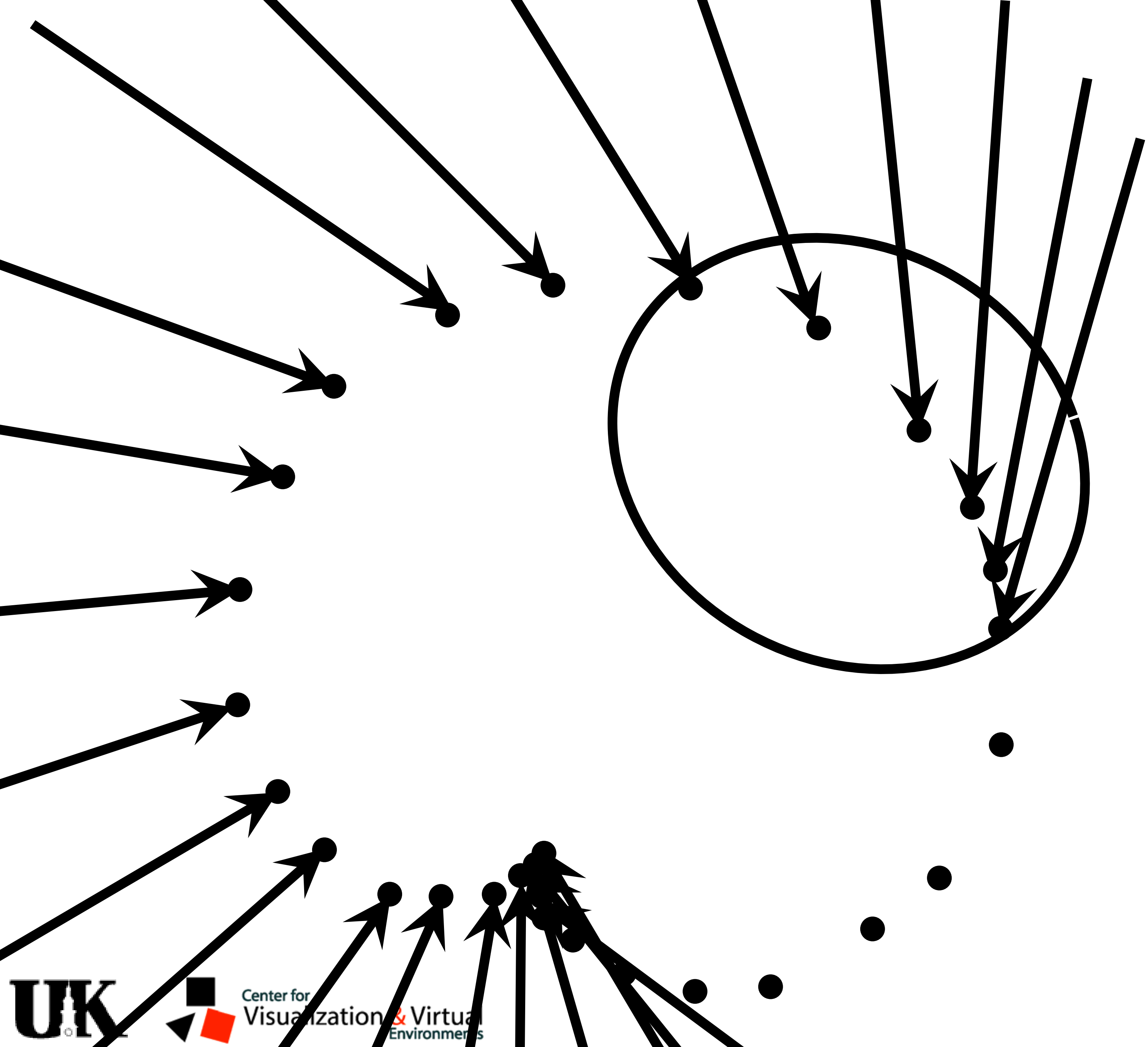


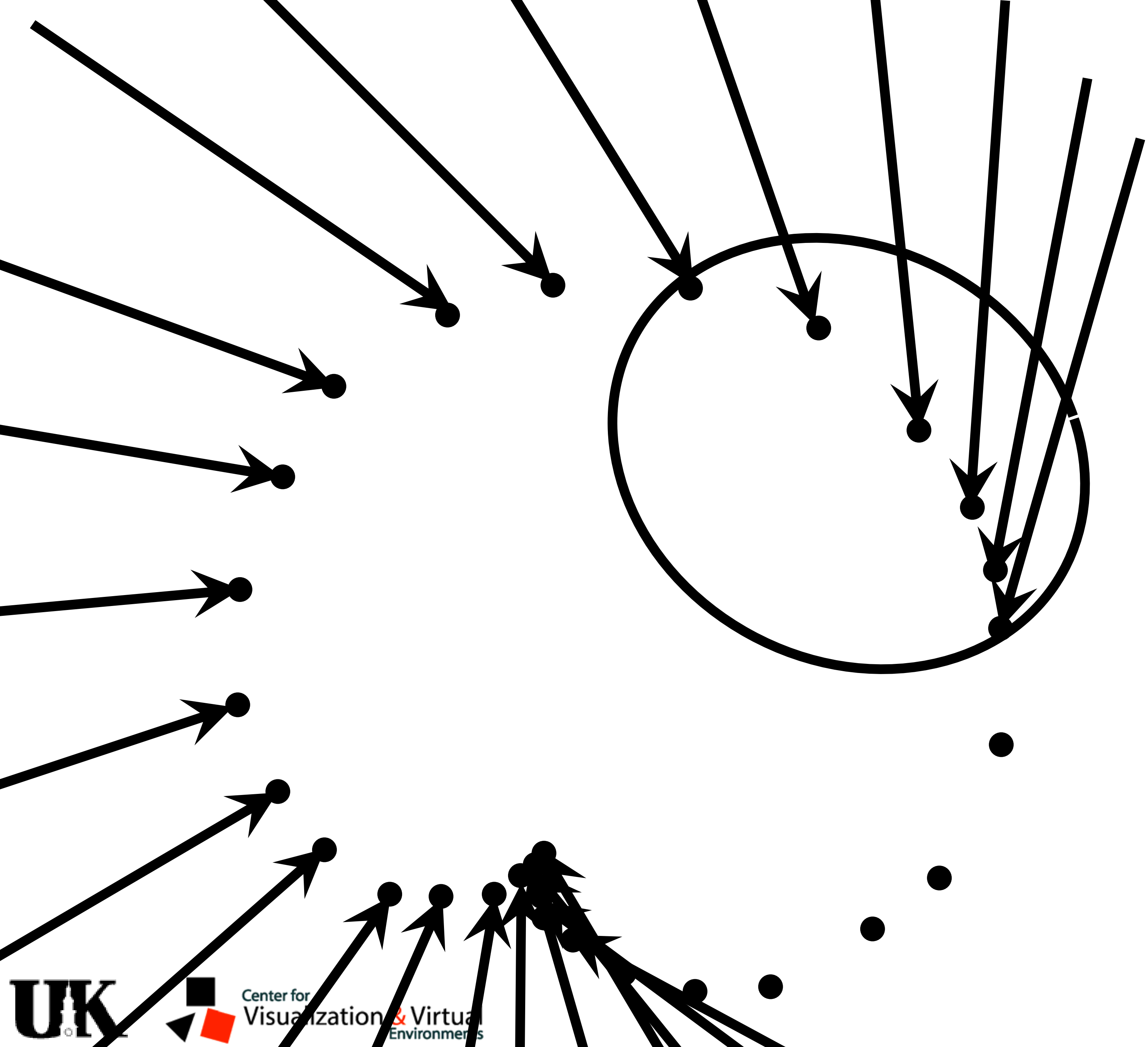
Center for Visualization & Virtual Environments

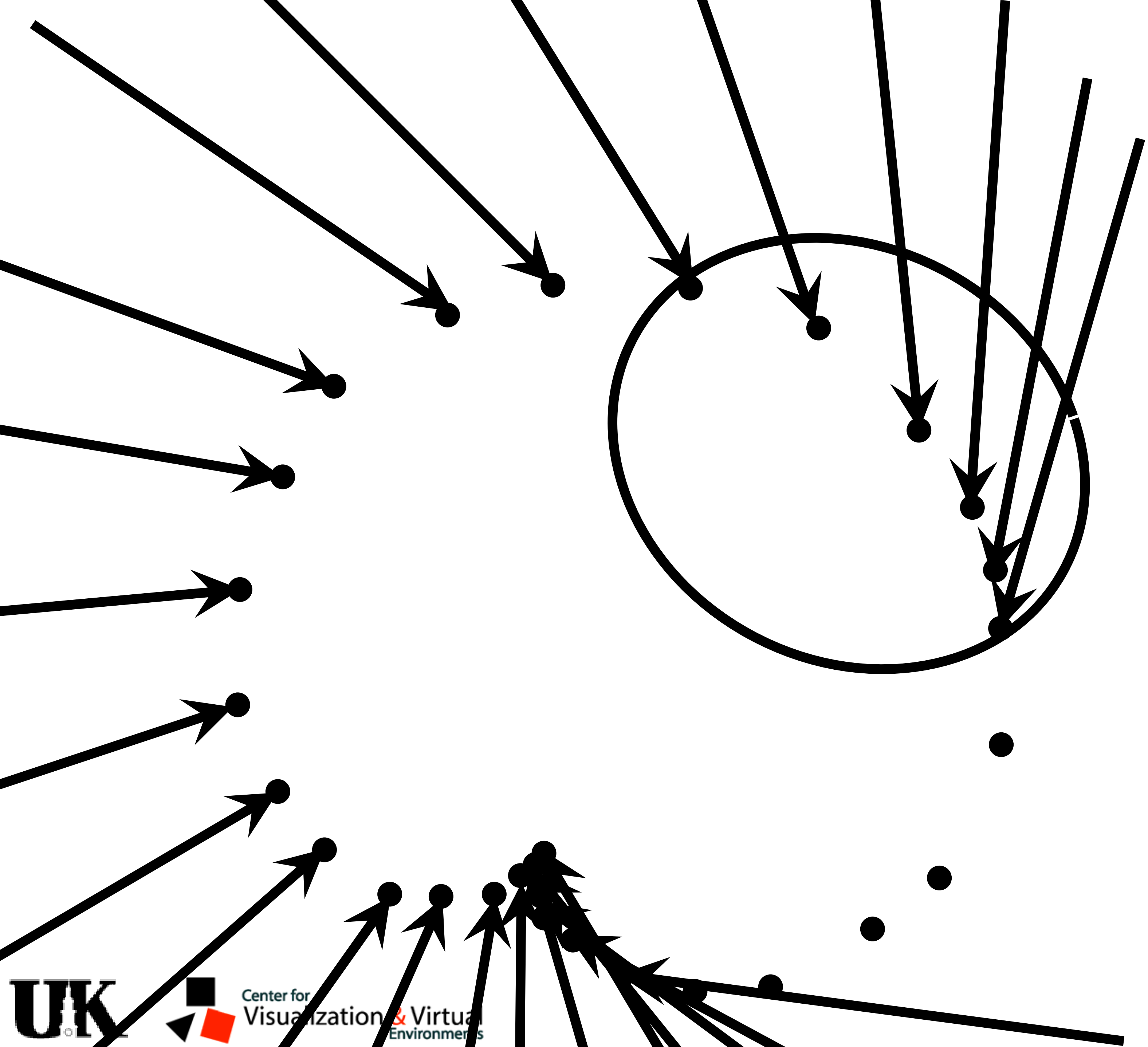


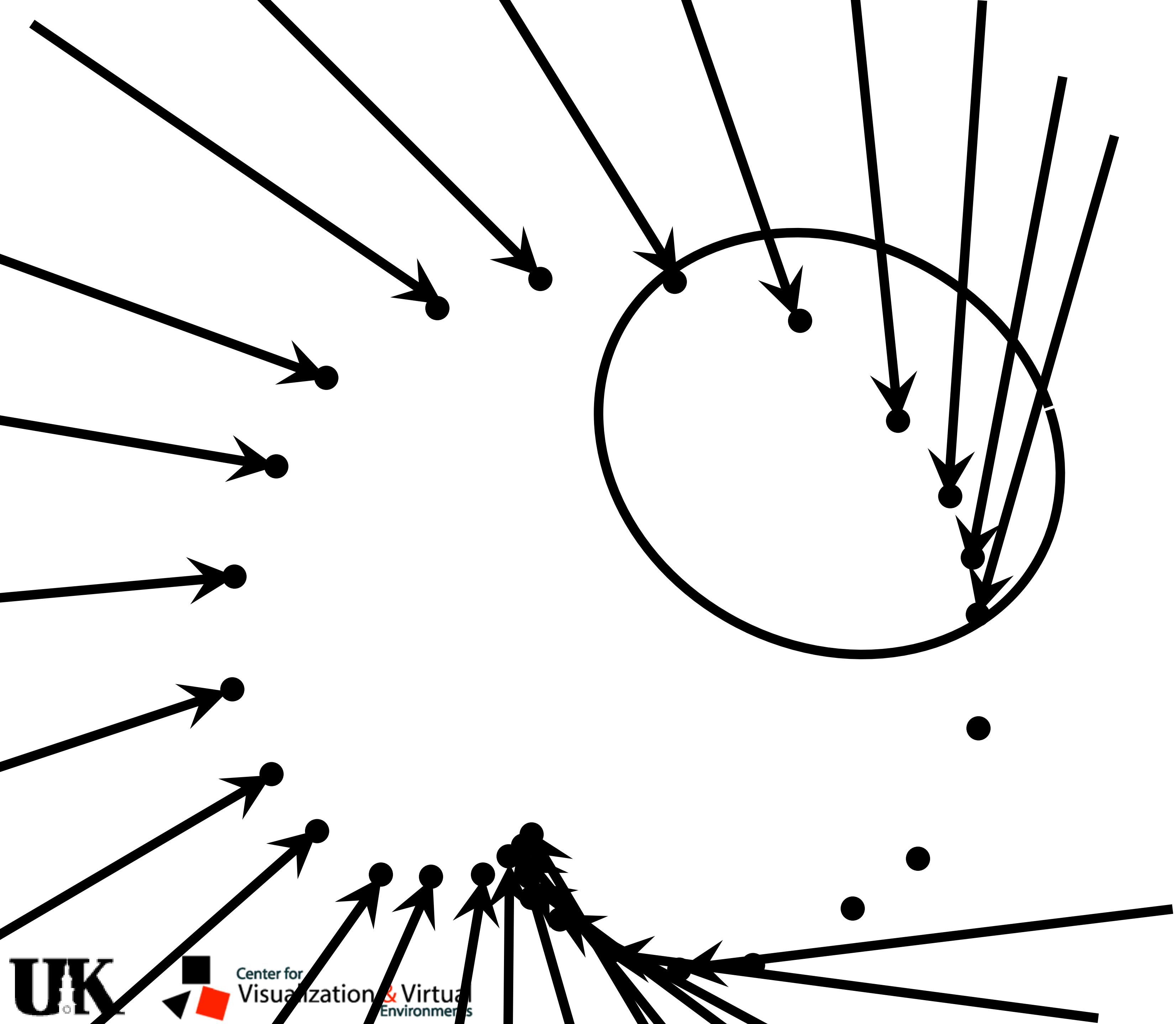

UK
Center for
Visualization & Virtual
Environments

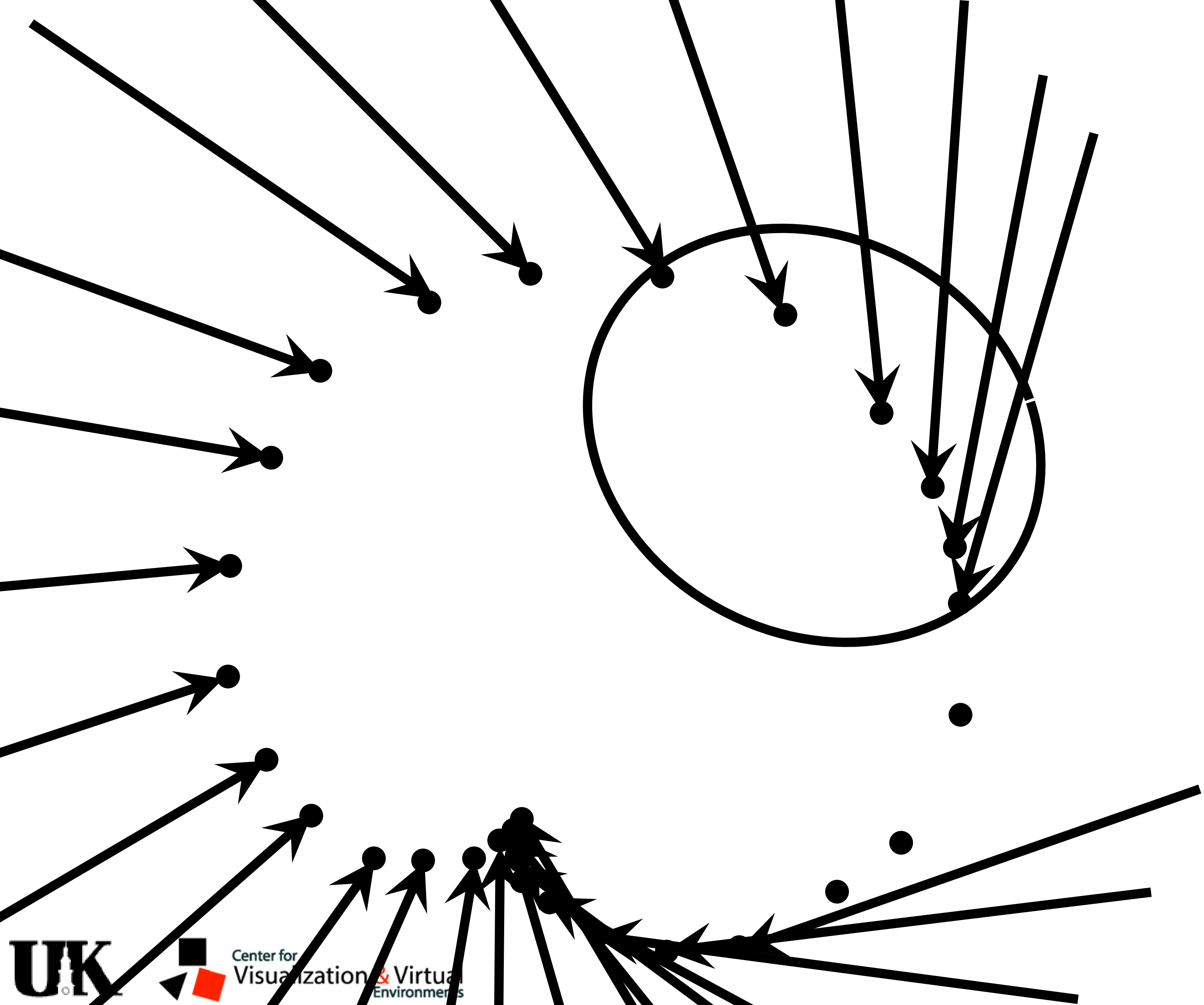


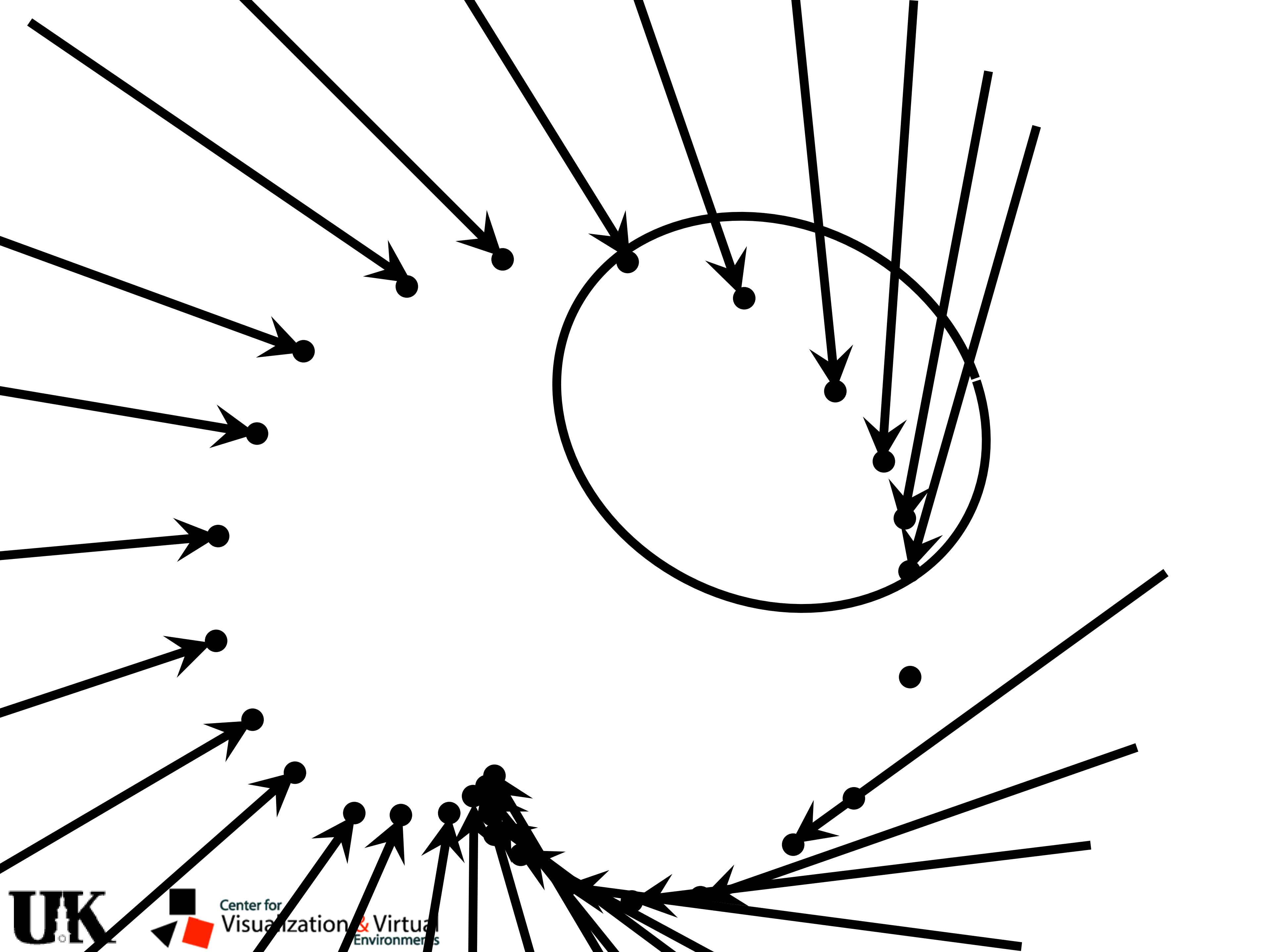


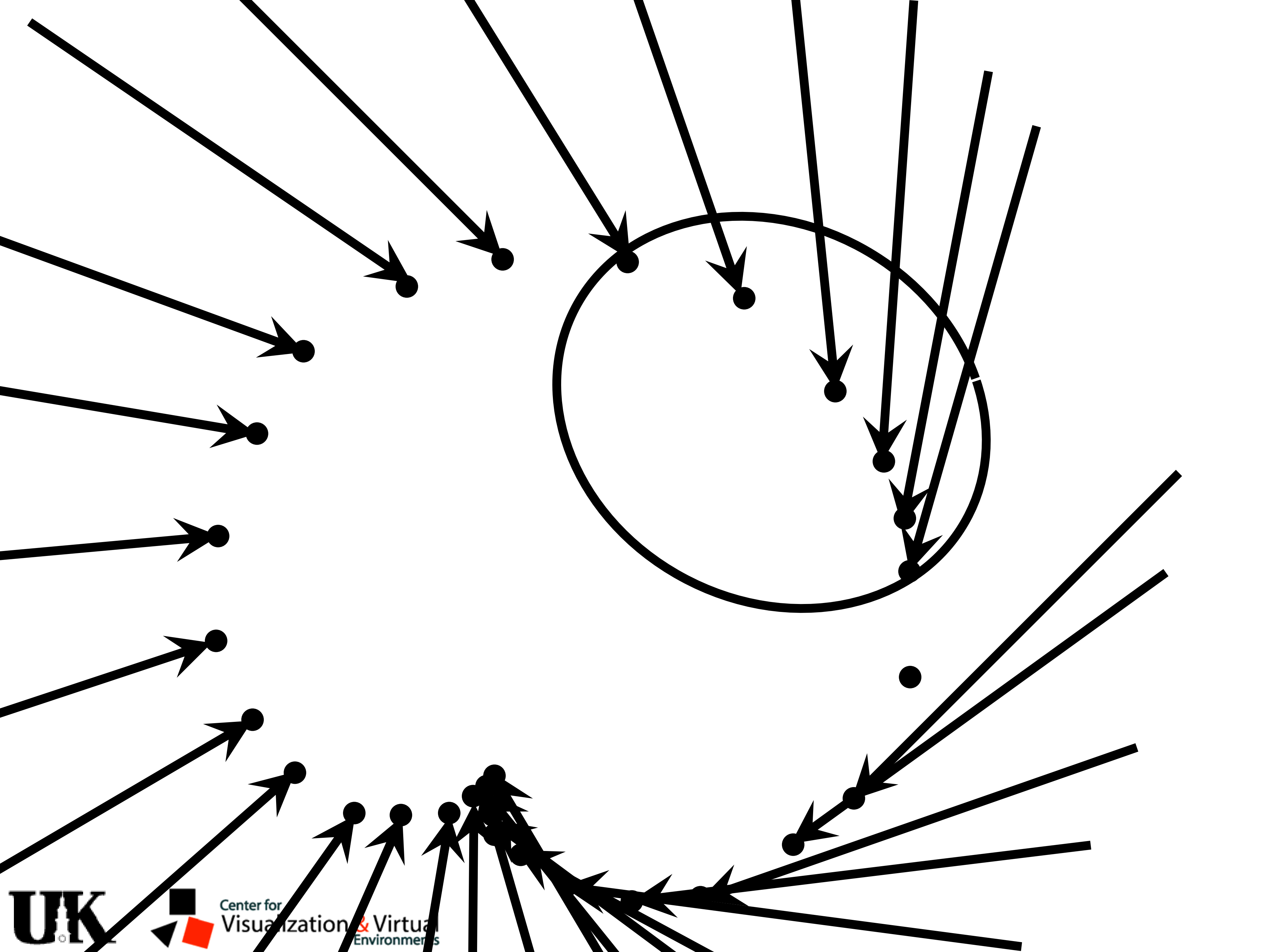




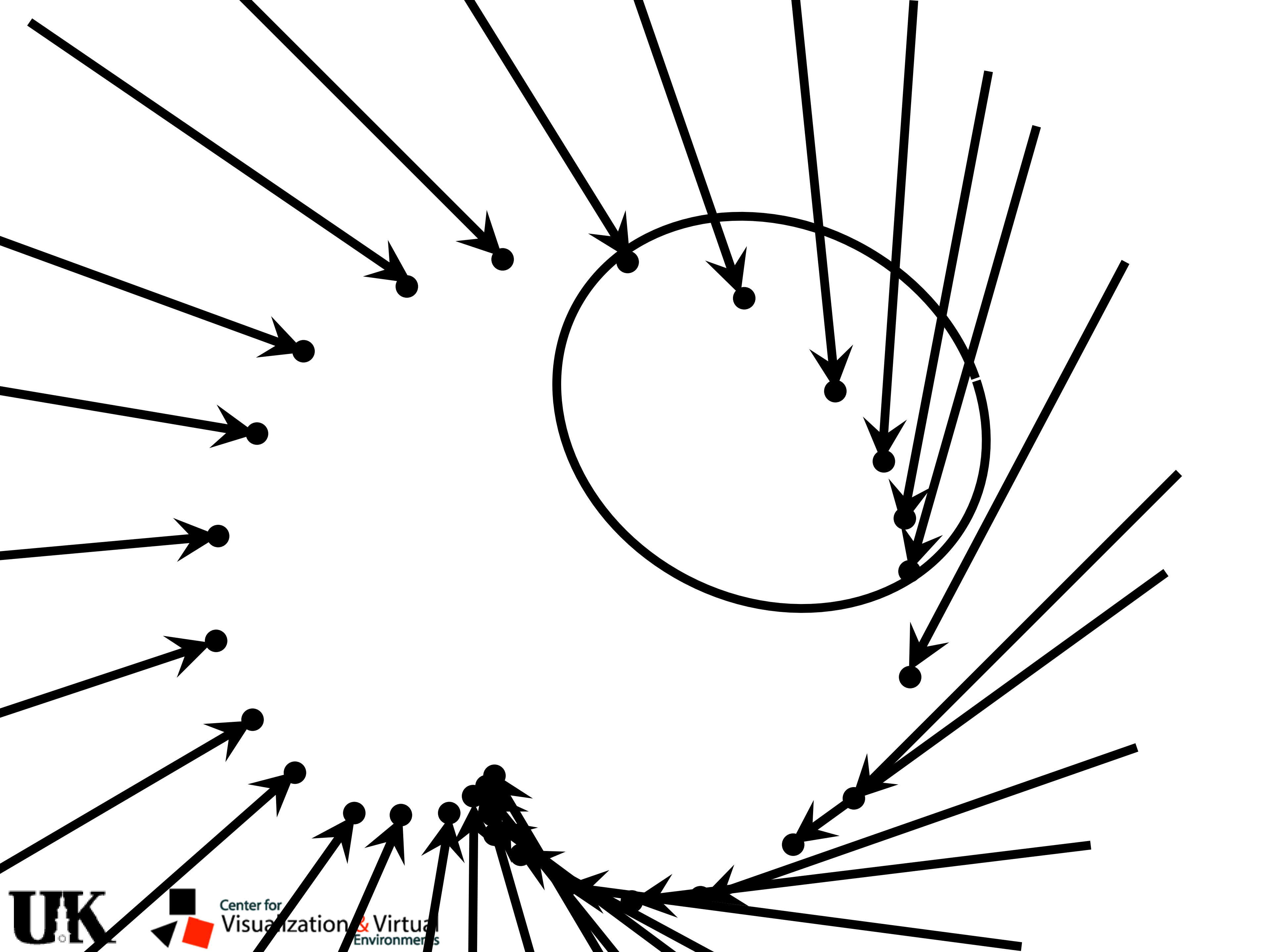

UK
Center for
Visualization & Virtual
Environments




UK
Center for Visualization & Virtual Environments

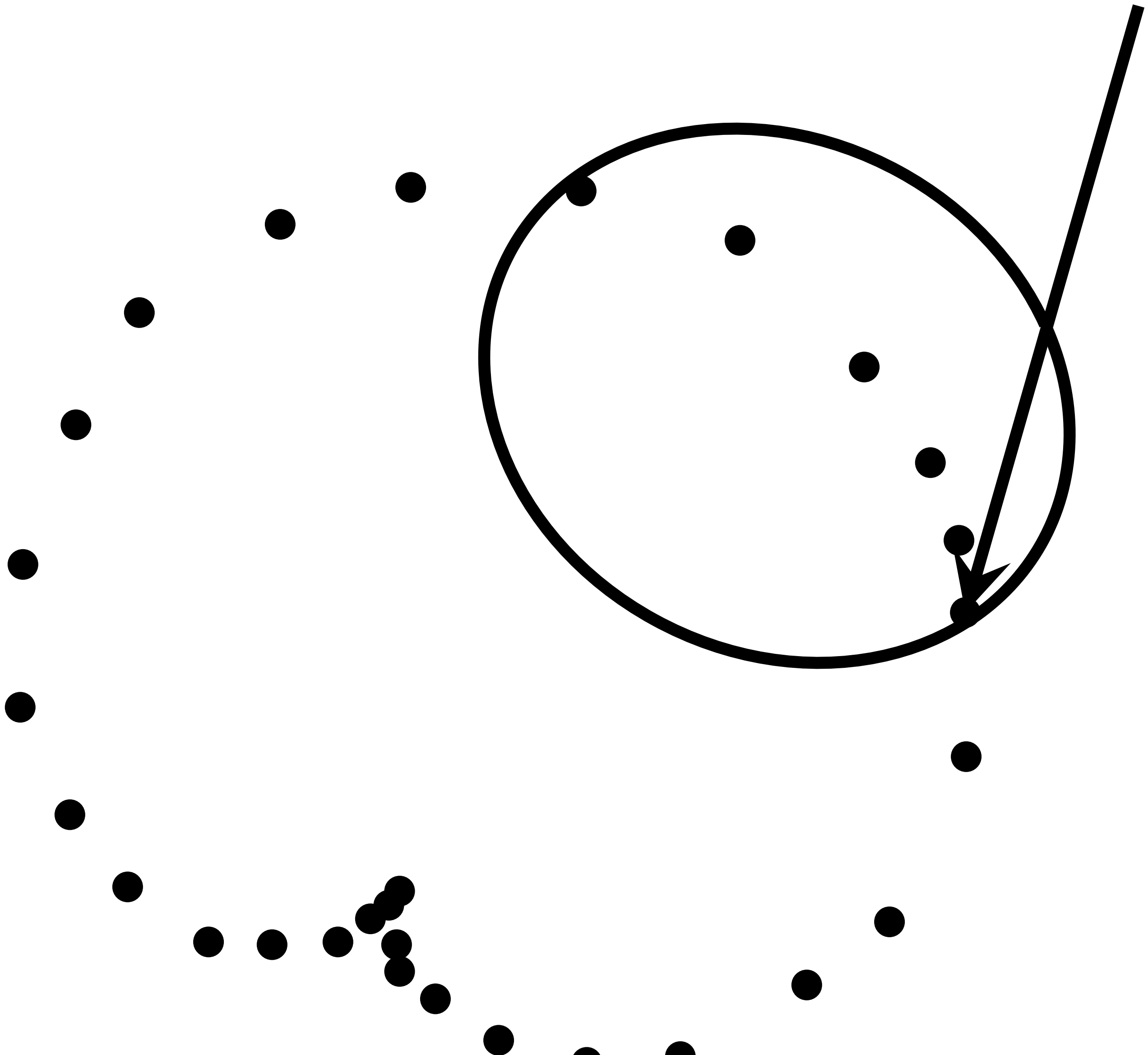


UK Center for Visualization & Virtual Environments

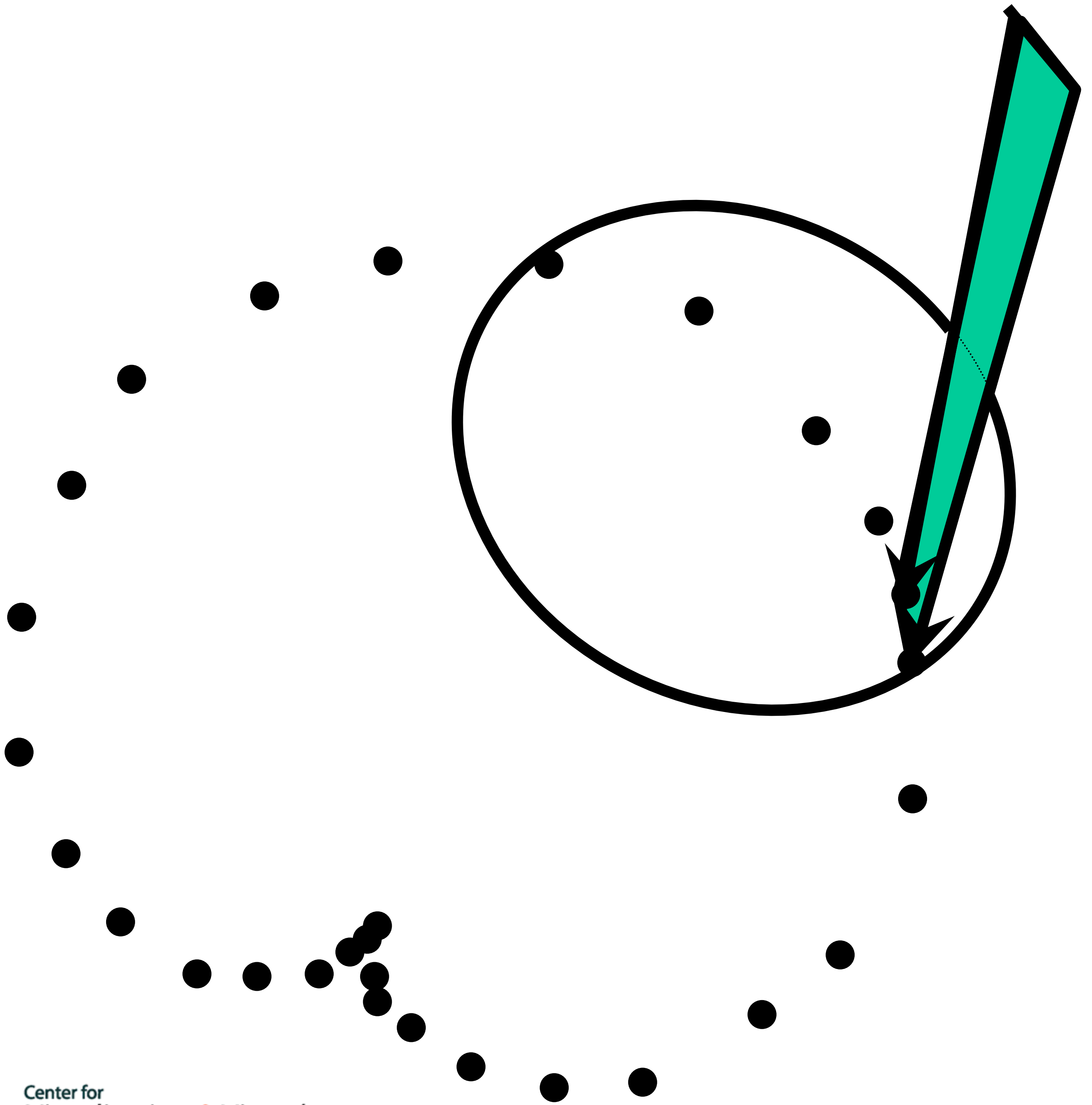


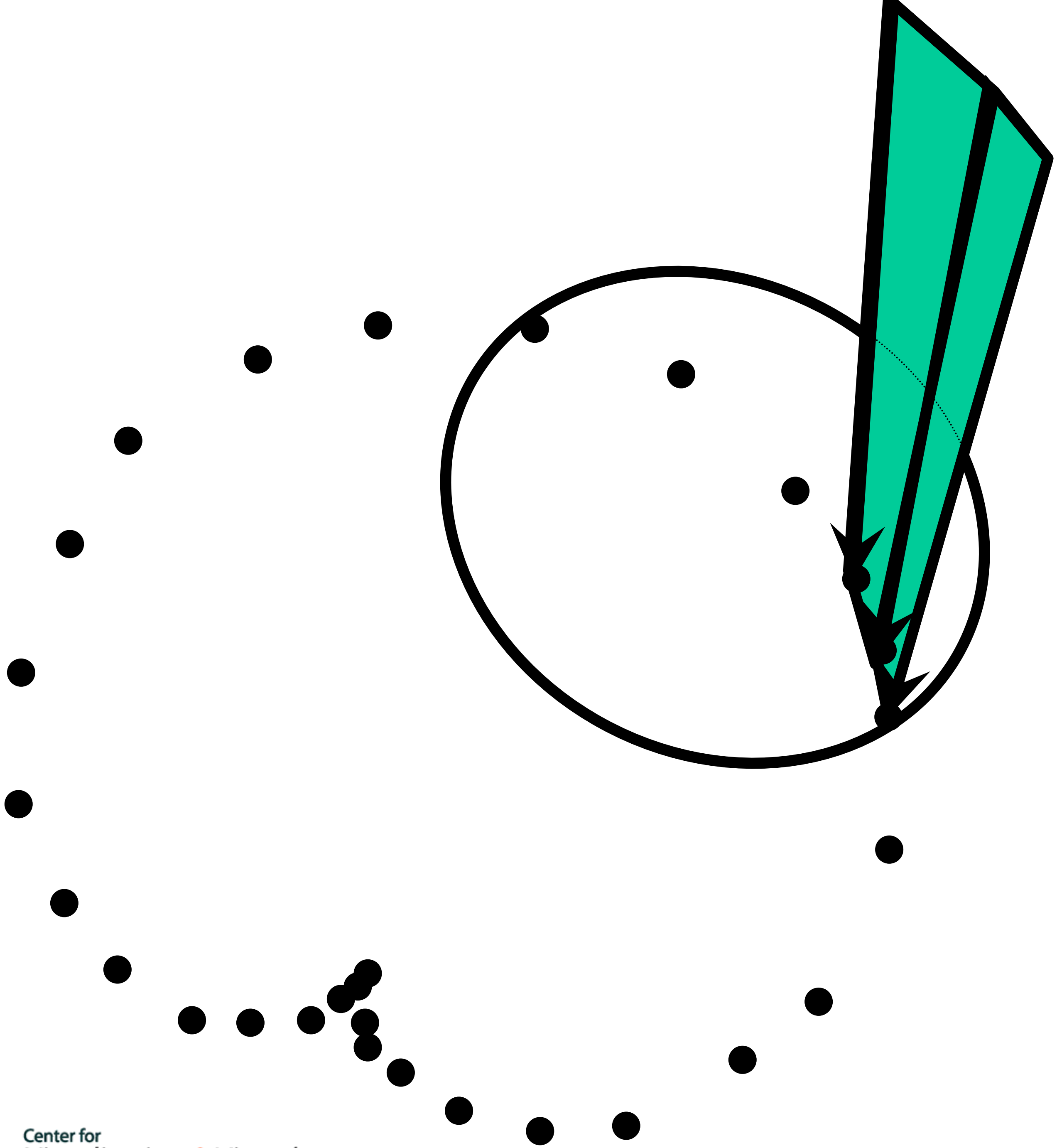
UK

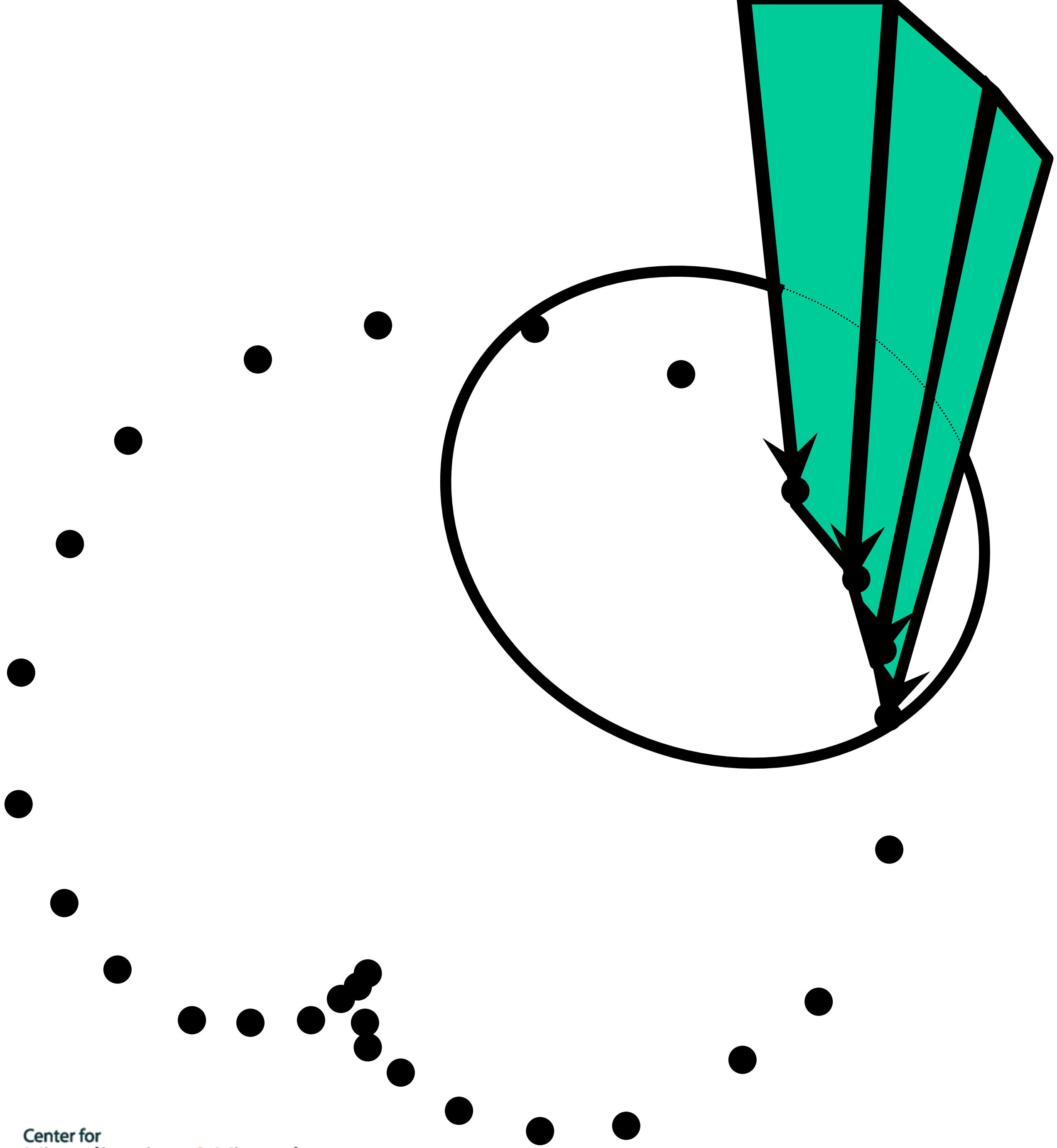
Center for Visualization & Virtual Environments



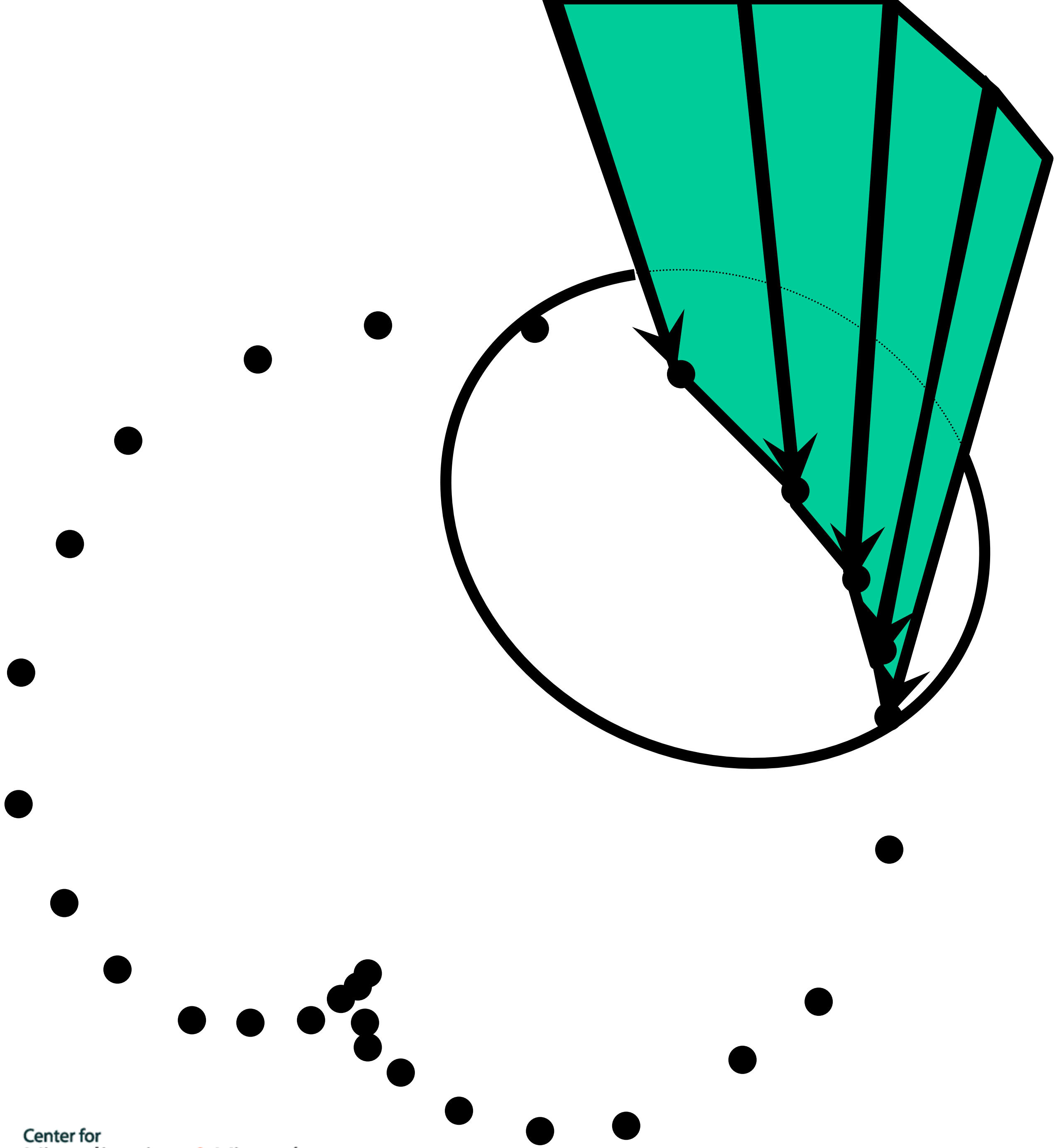
Center for Visualization & Virtual Environments

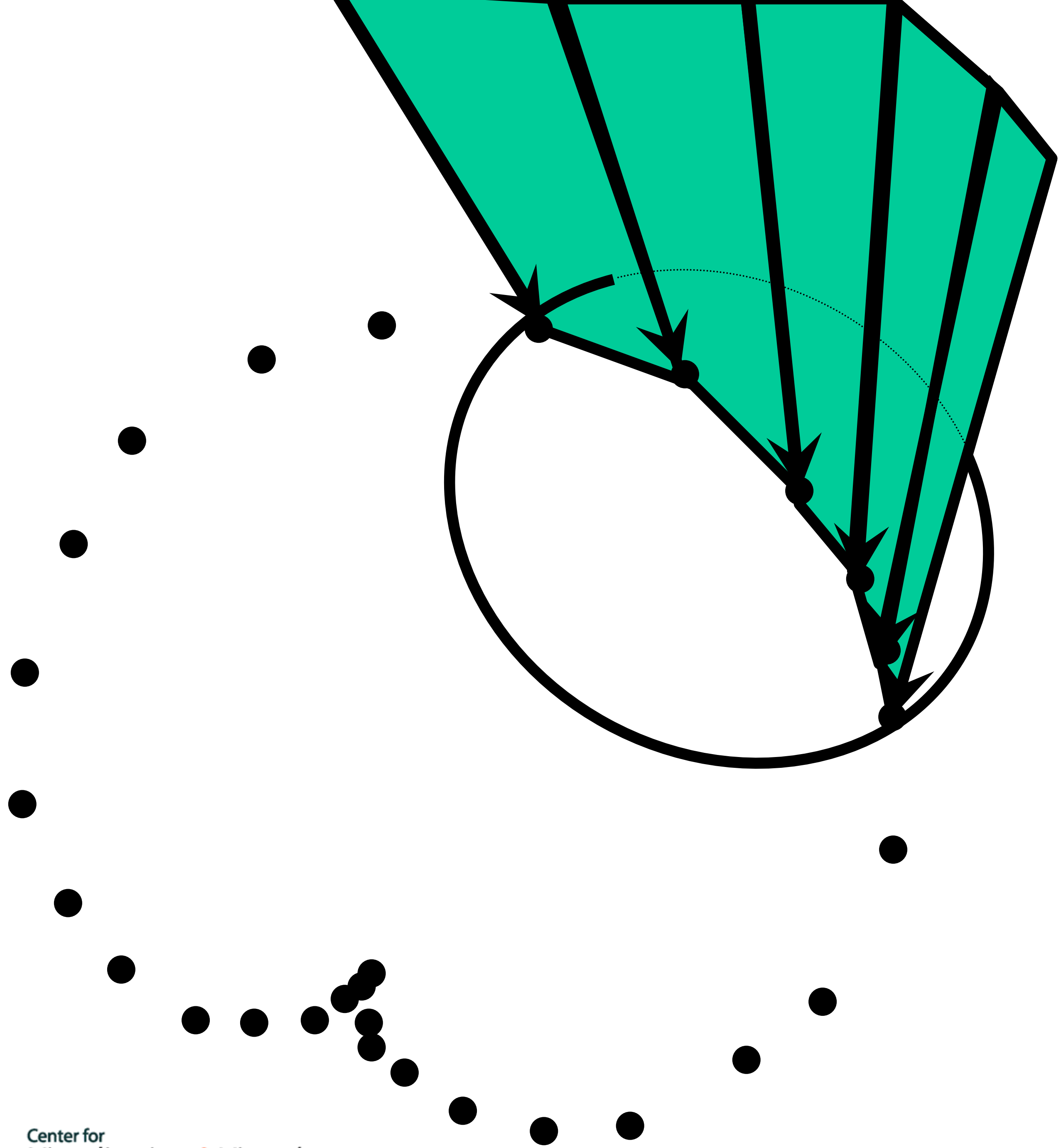


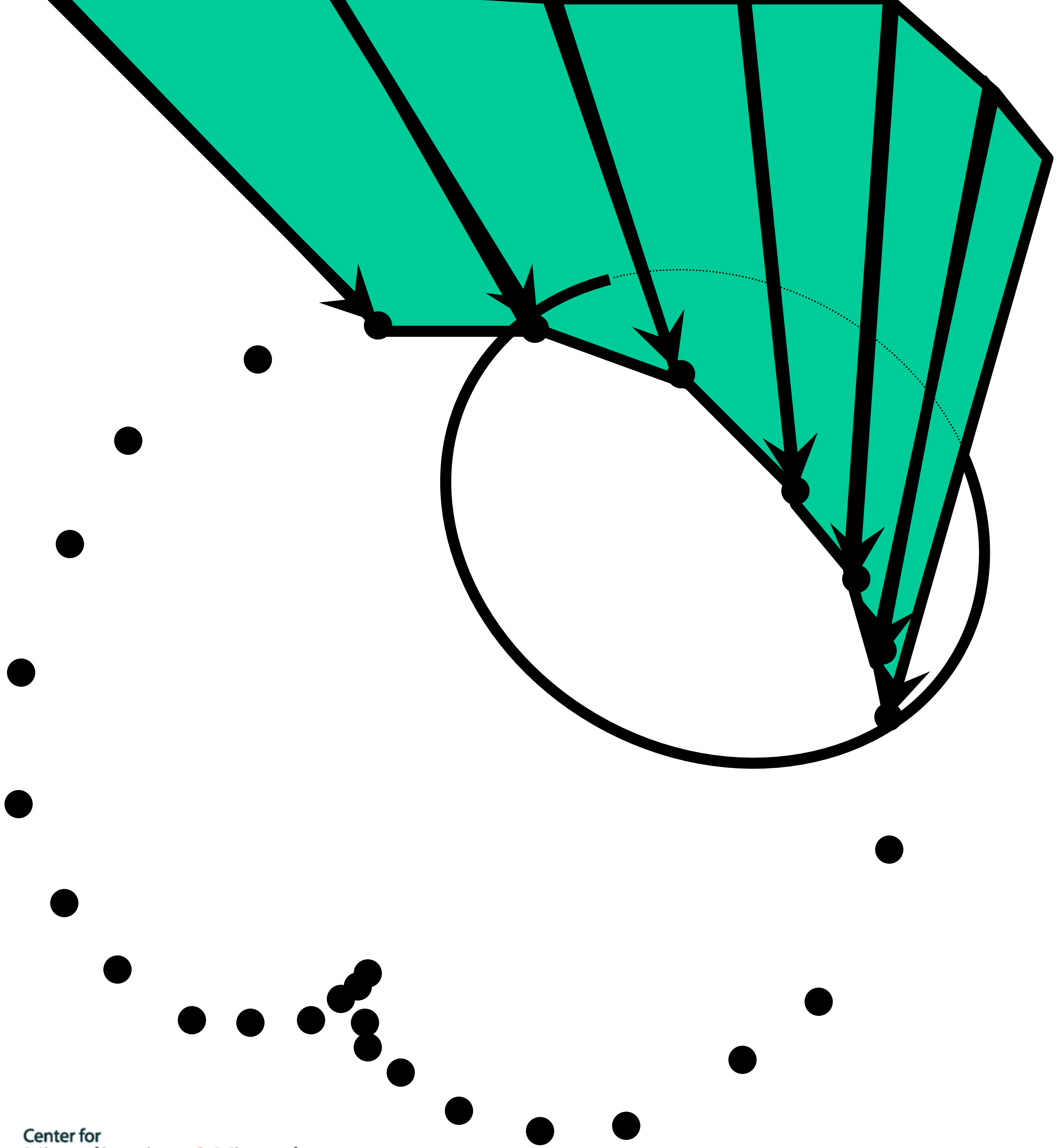




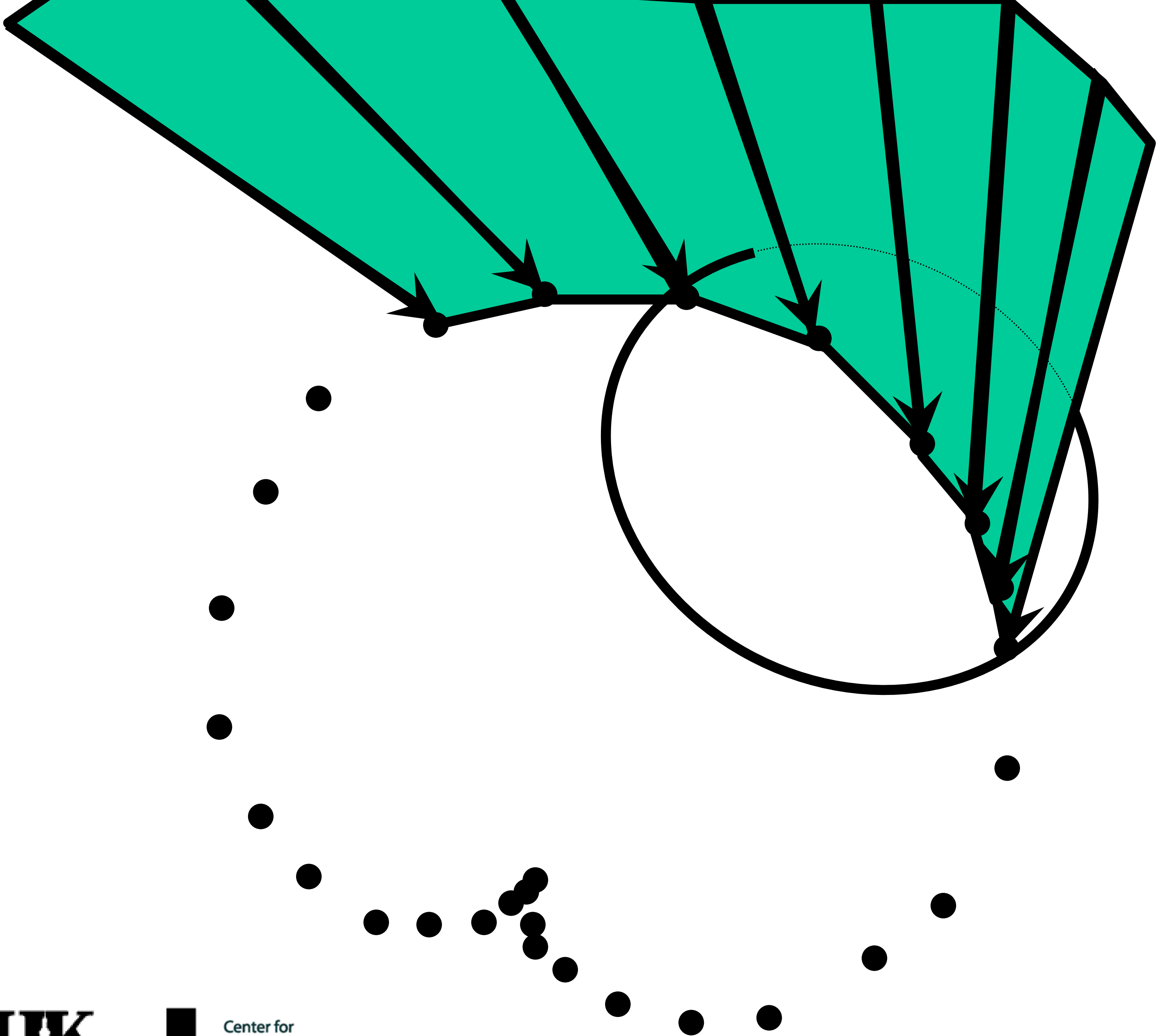
Center for Visualization & Virtual Environments

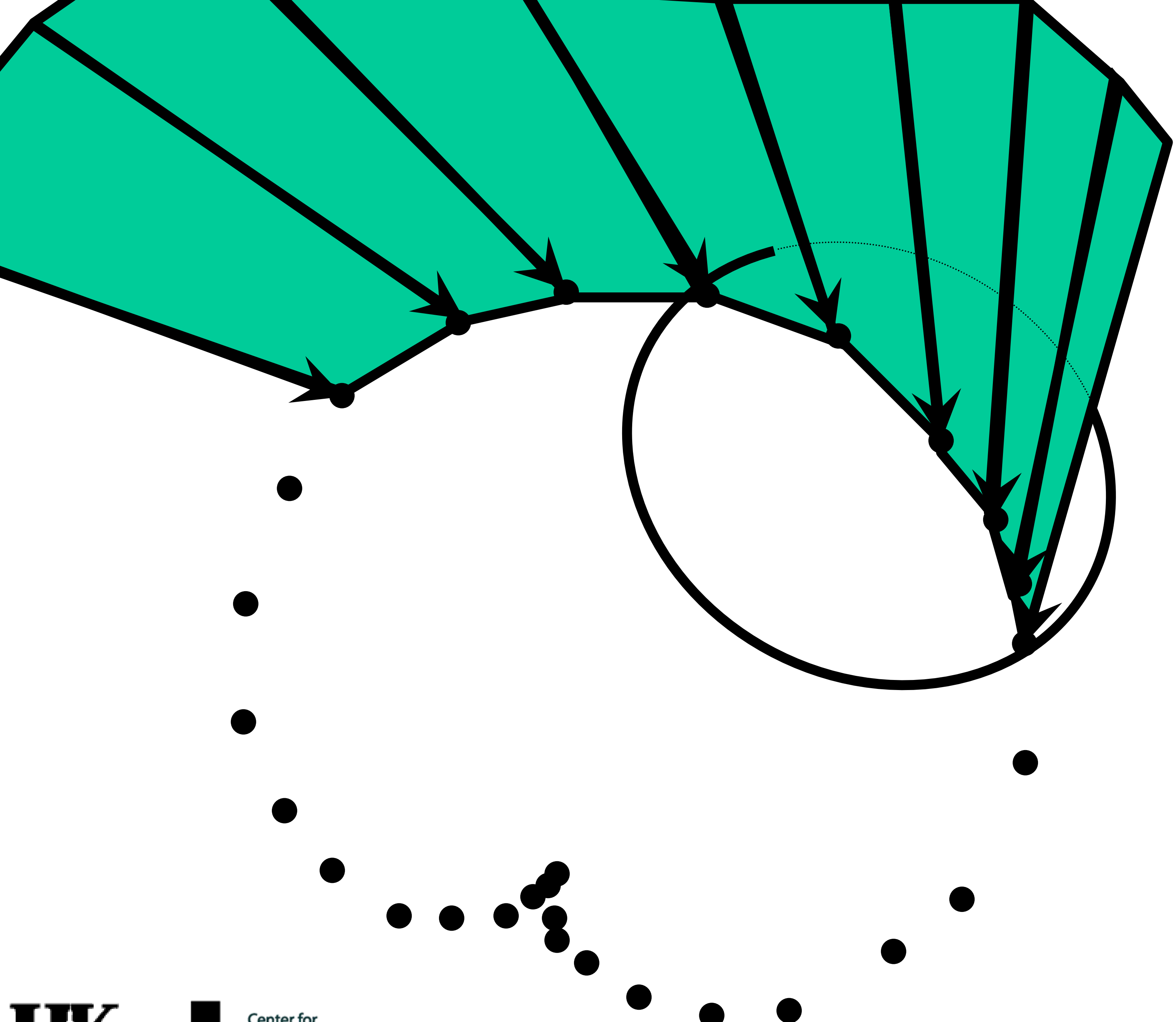


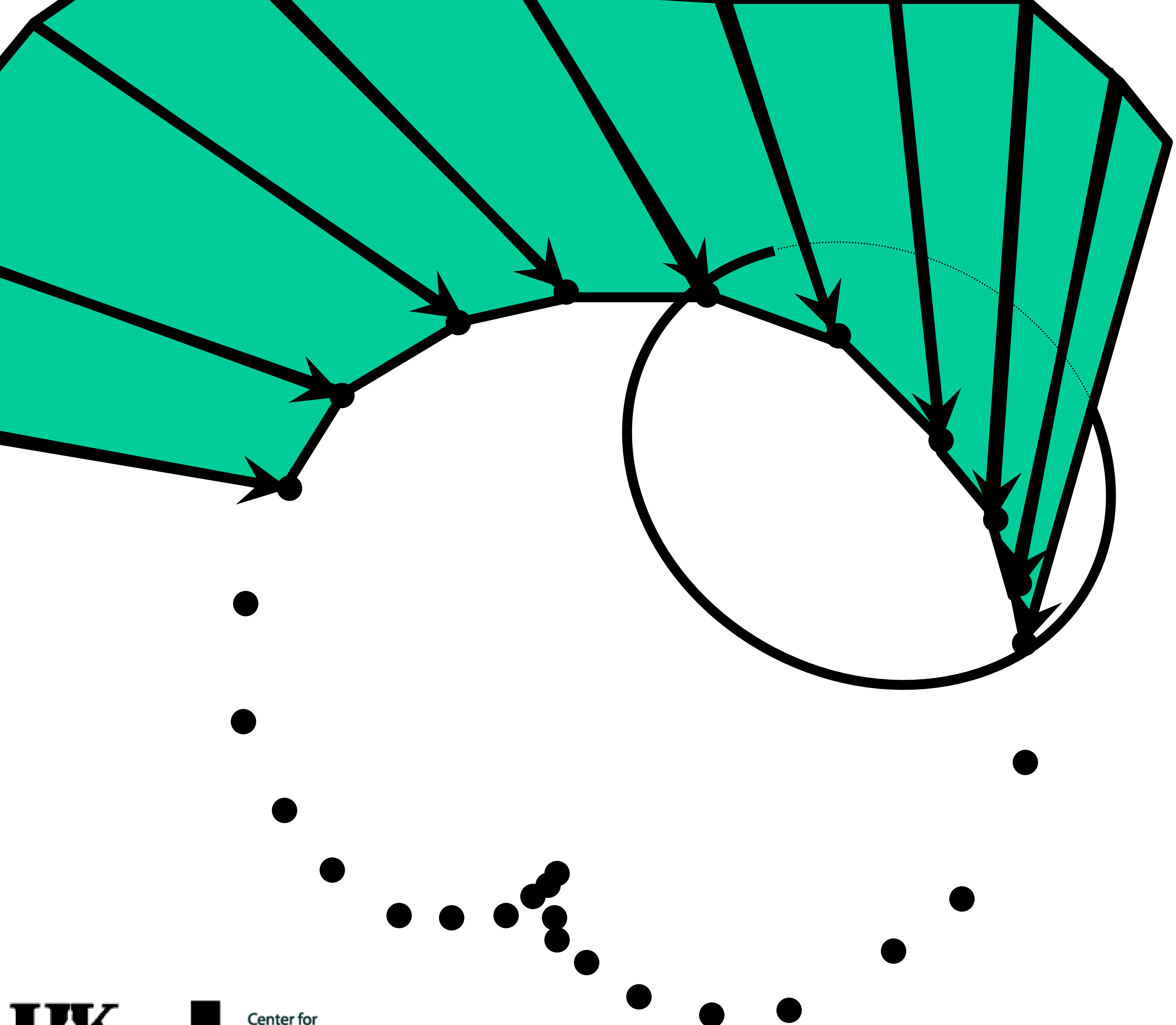


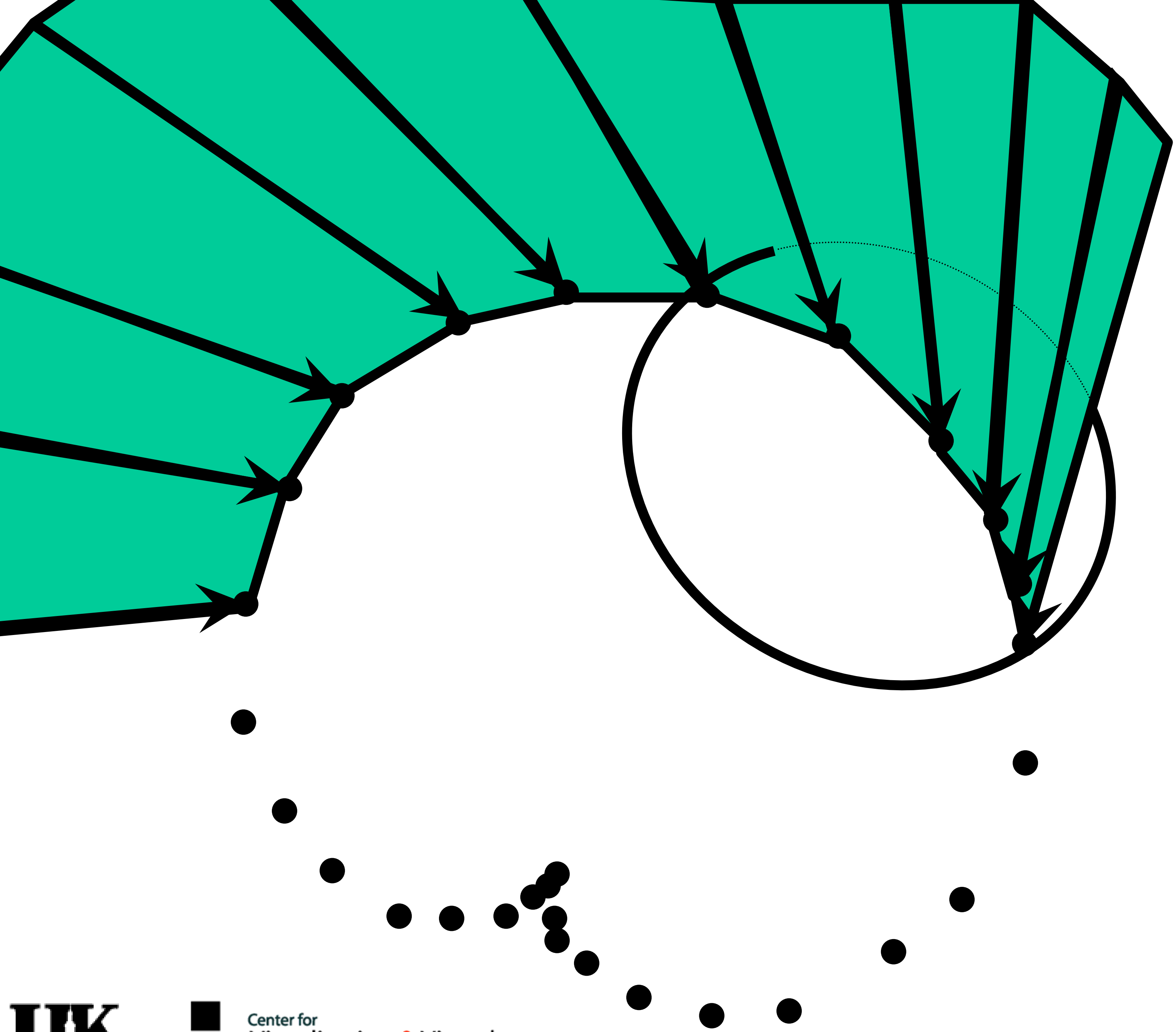


Center for Visualization & Virtual Environments

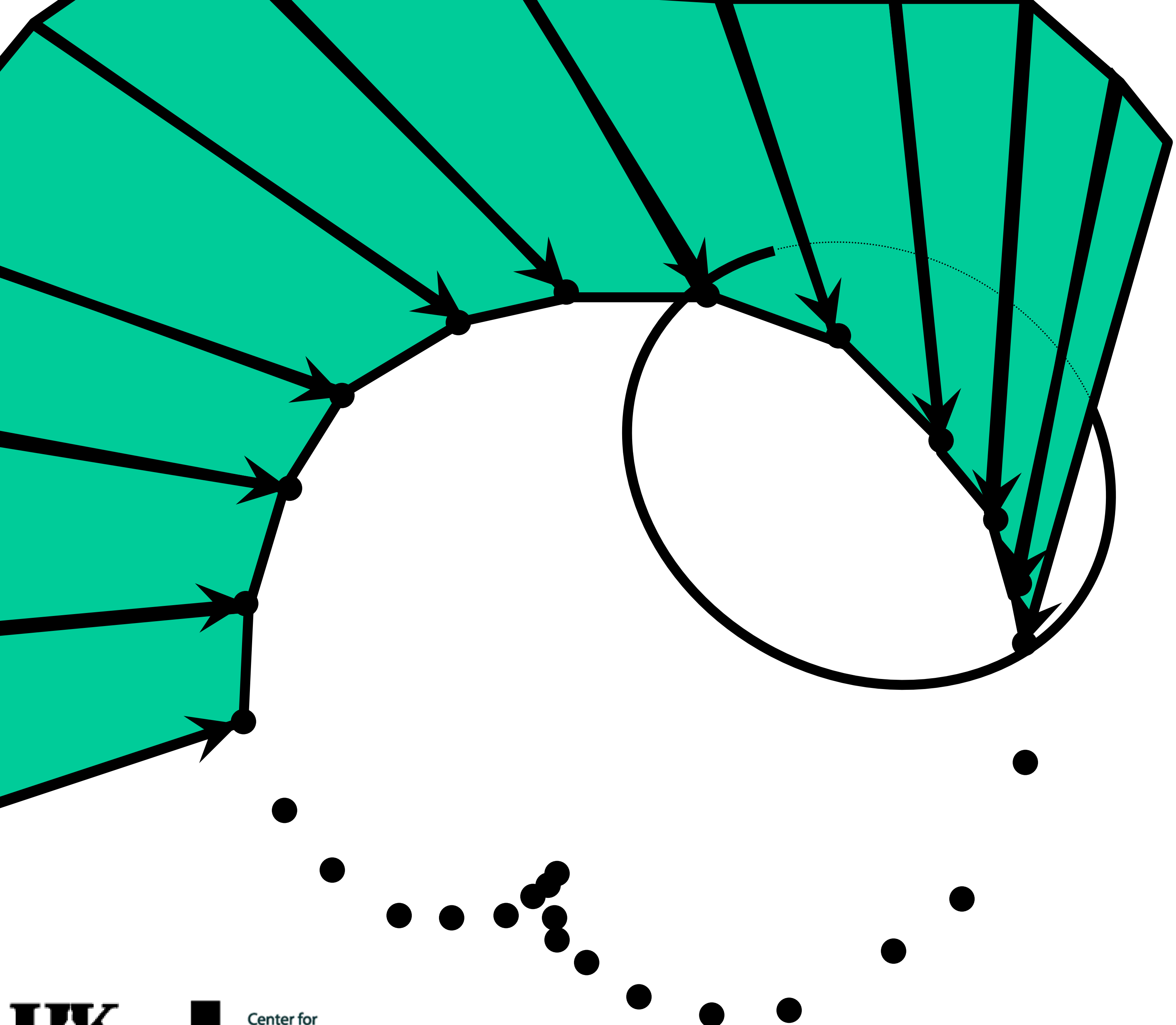




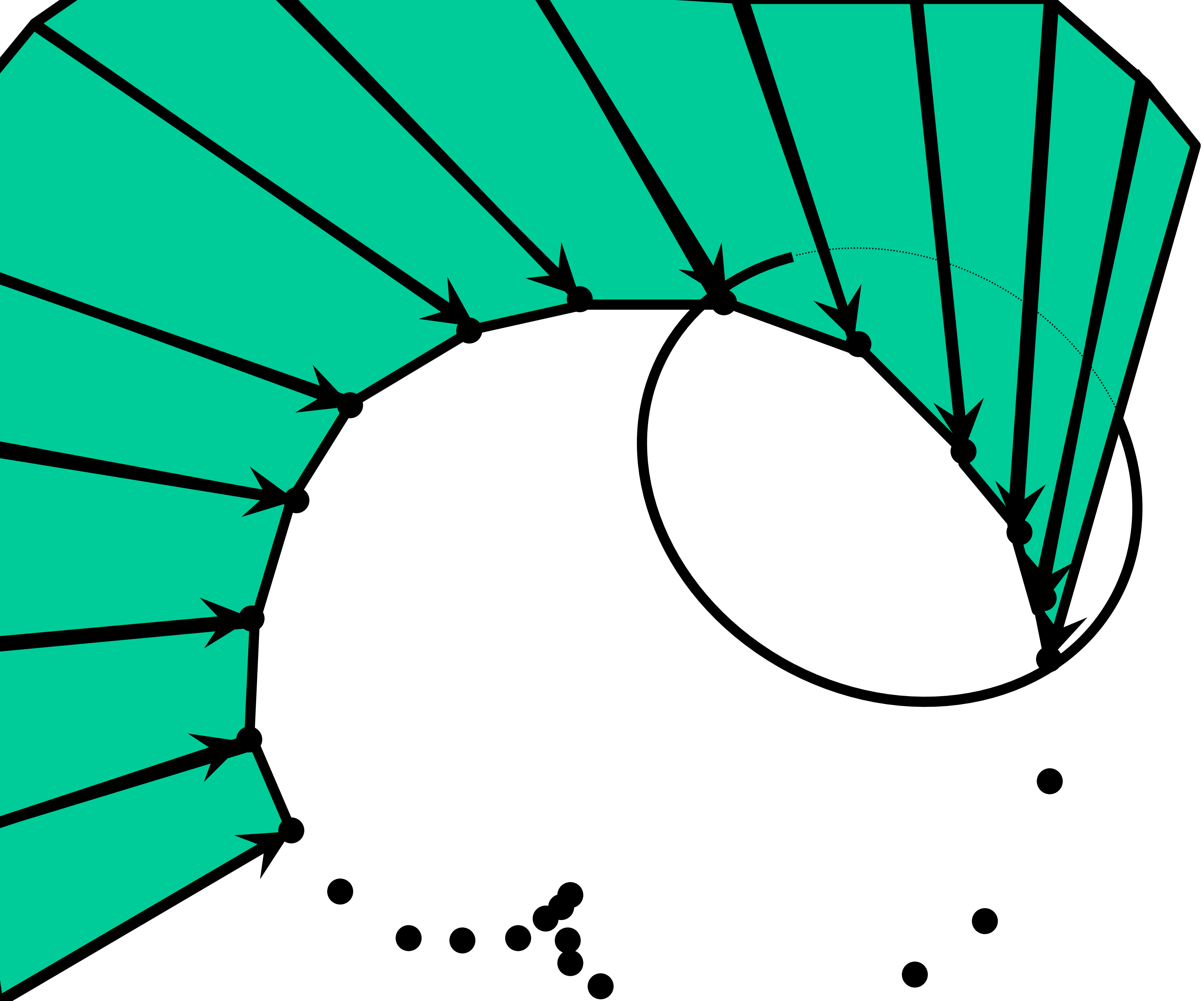




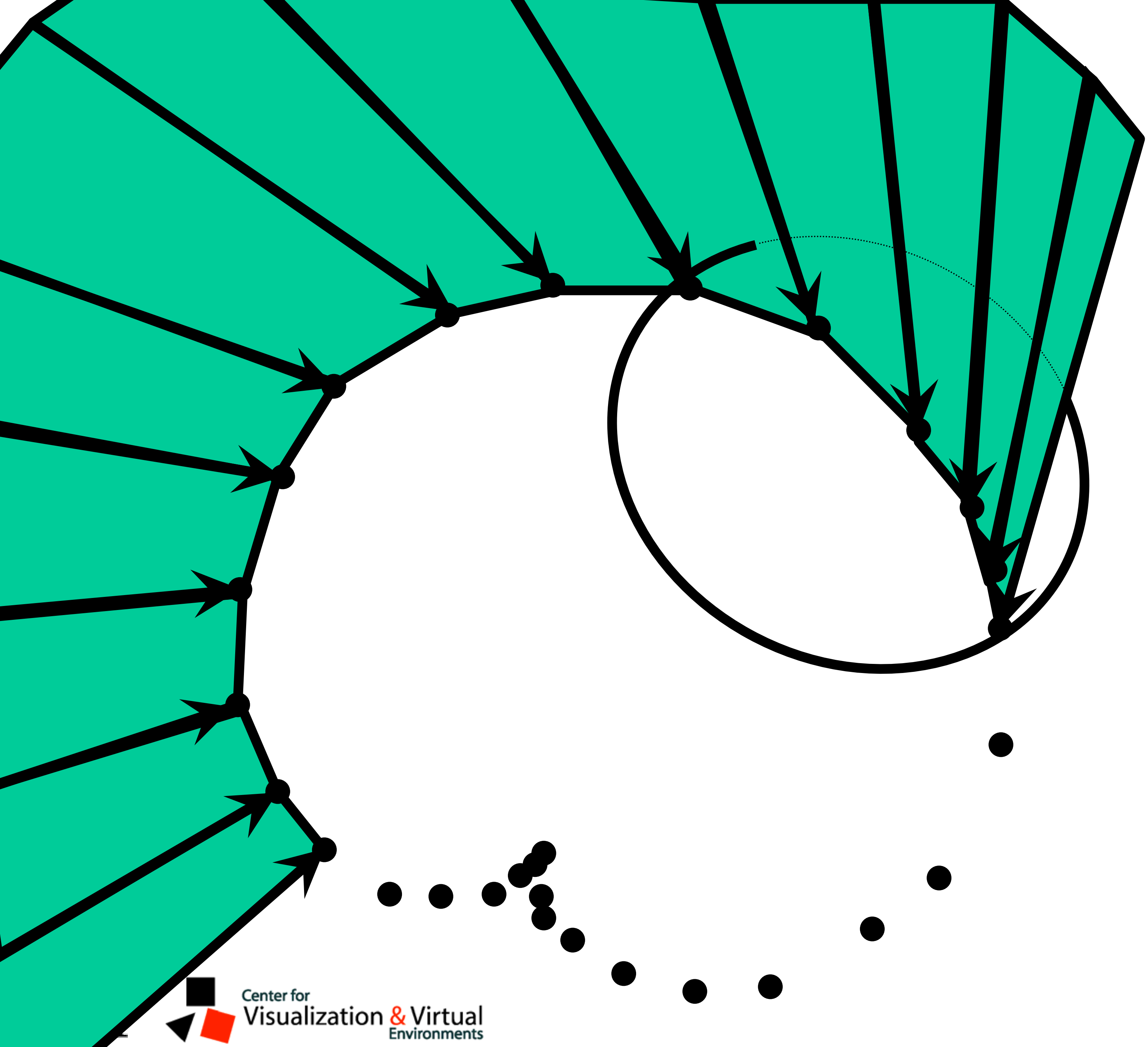
UK Center for Visualization & Virtual Environments

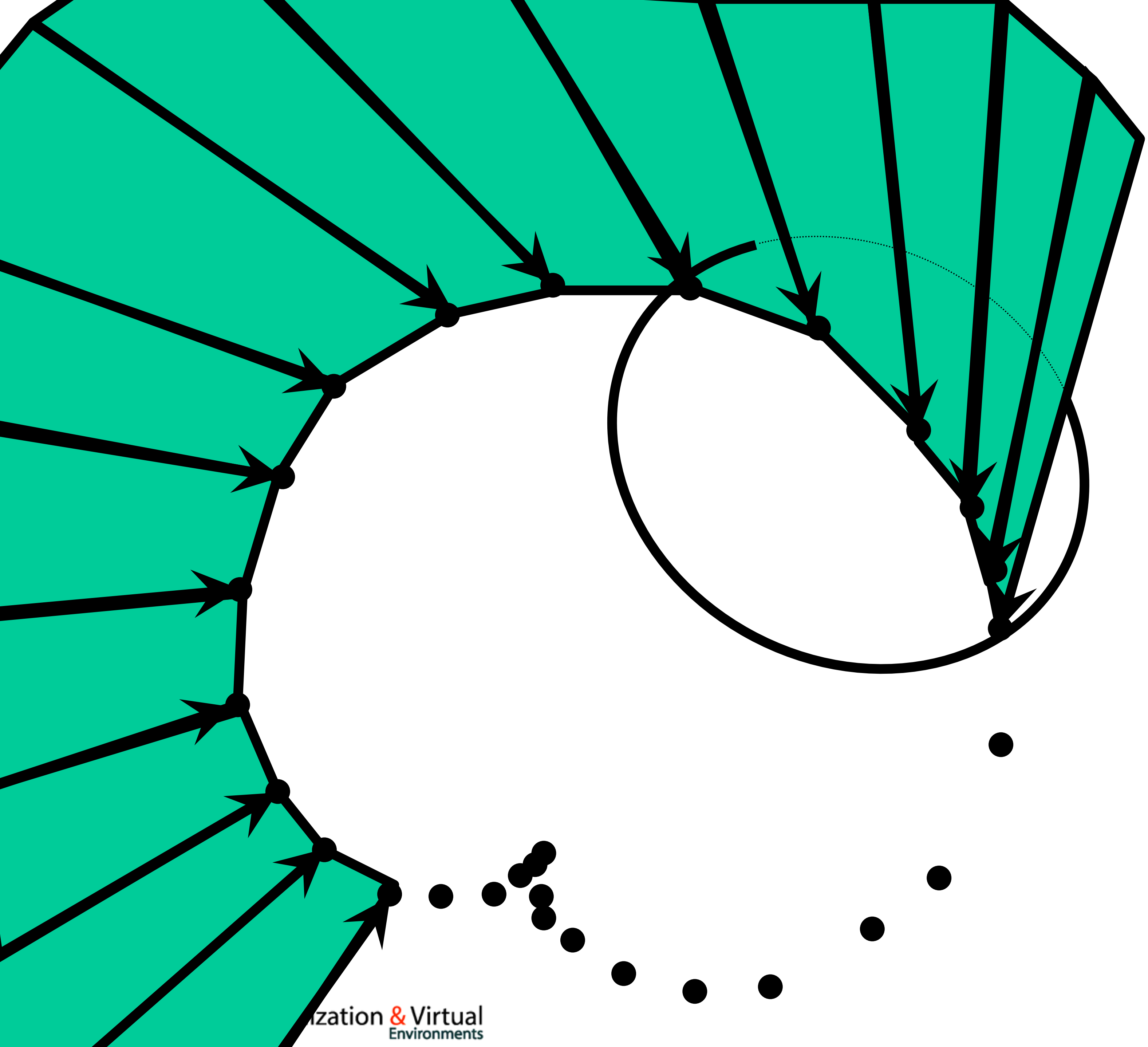


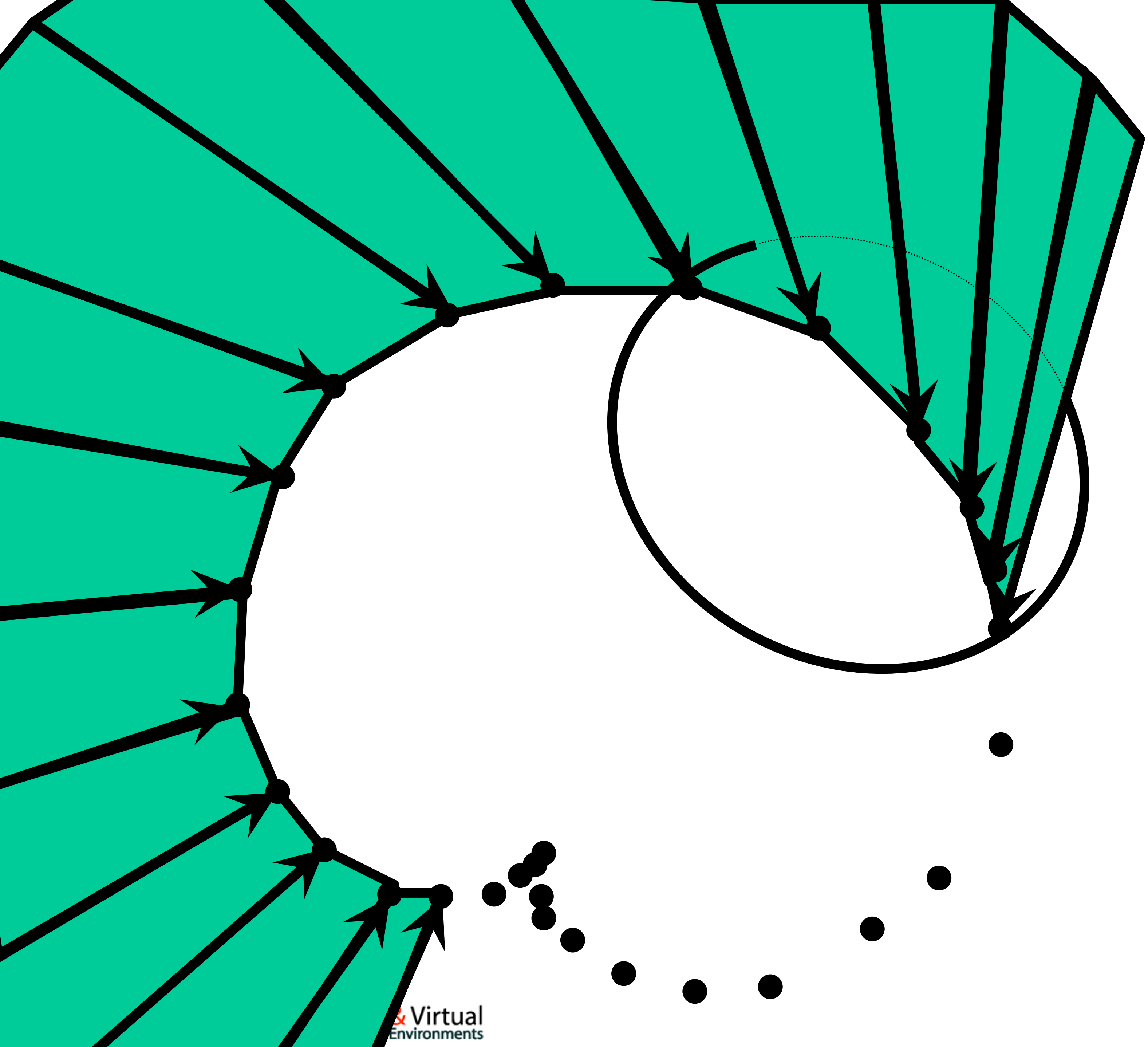
Center for Visualization & Virtual Environments

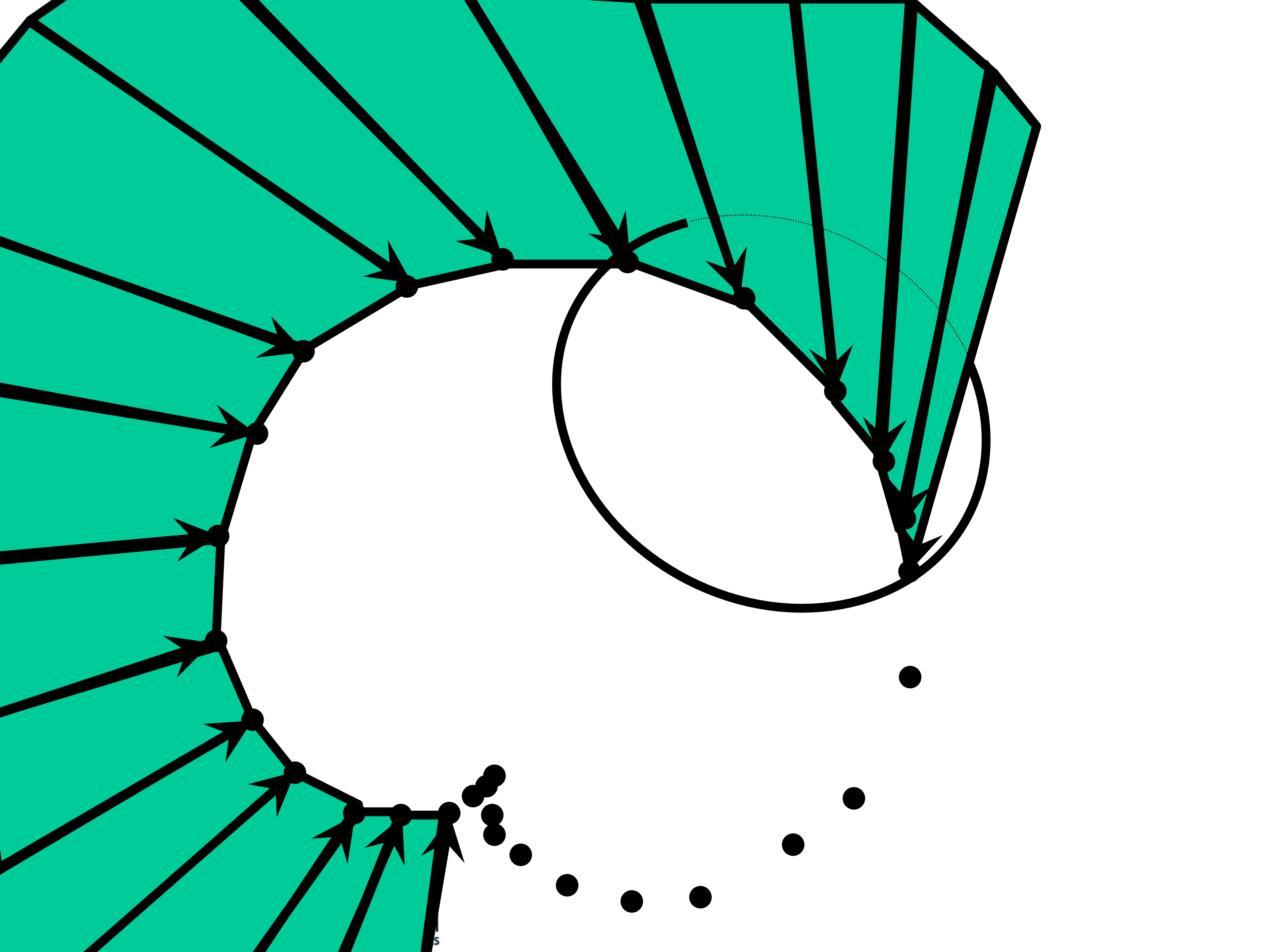


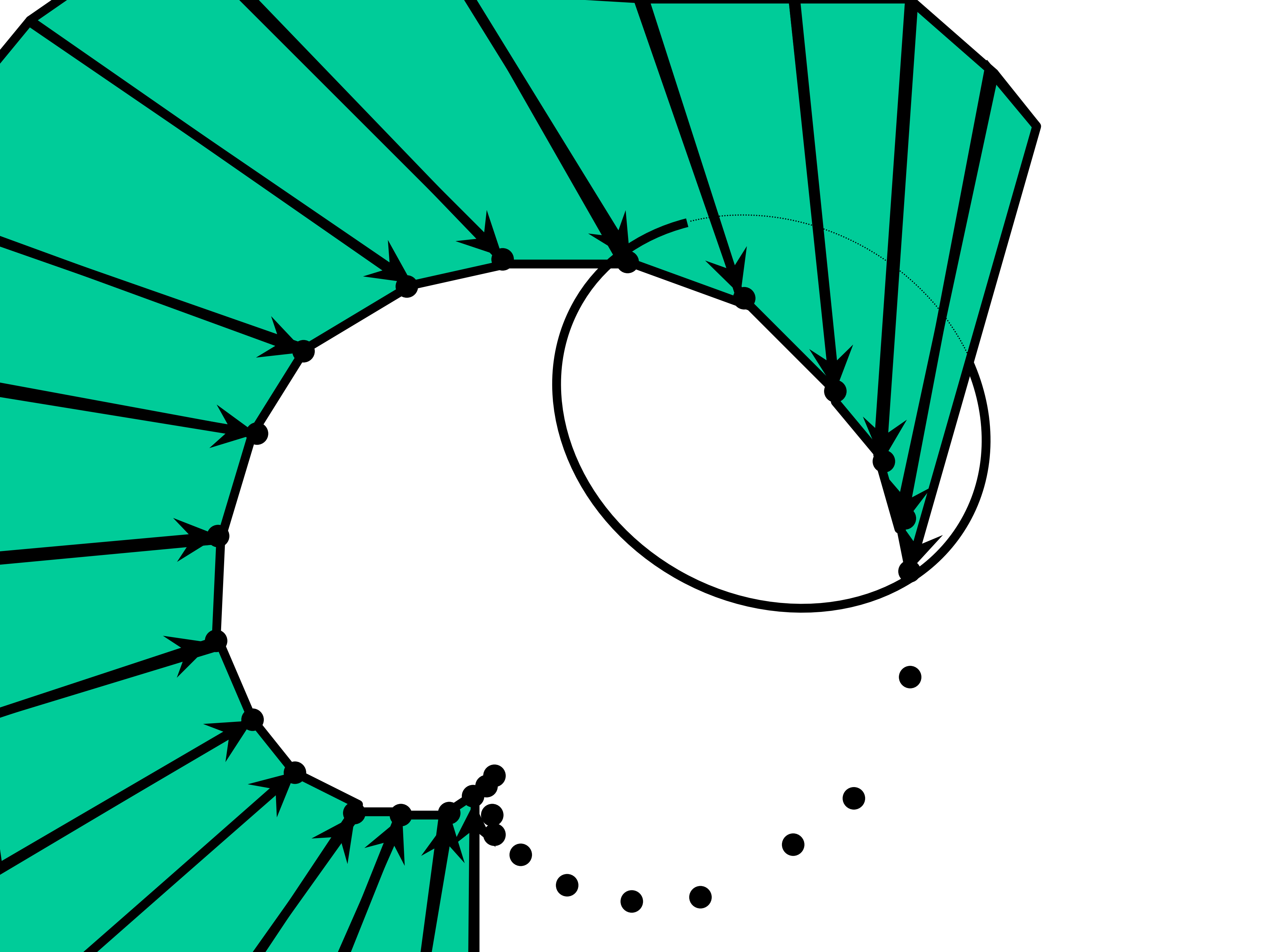
UK Center for Visualization & Virtual Environments

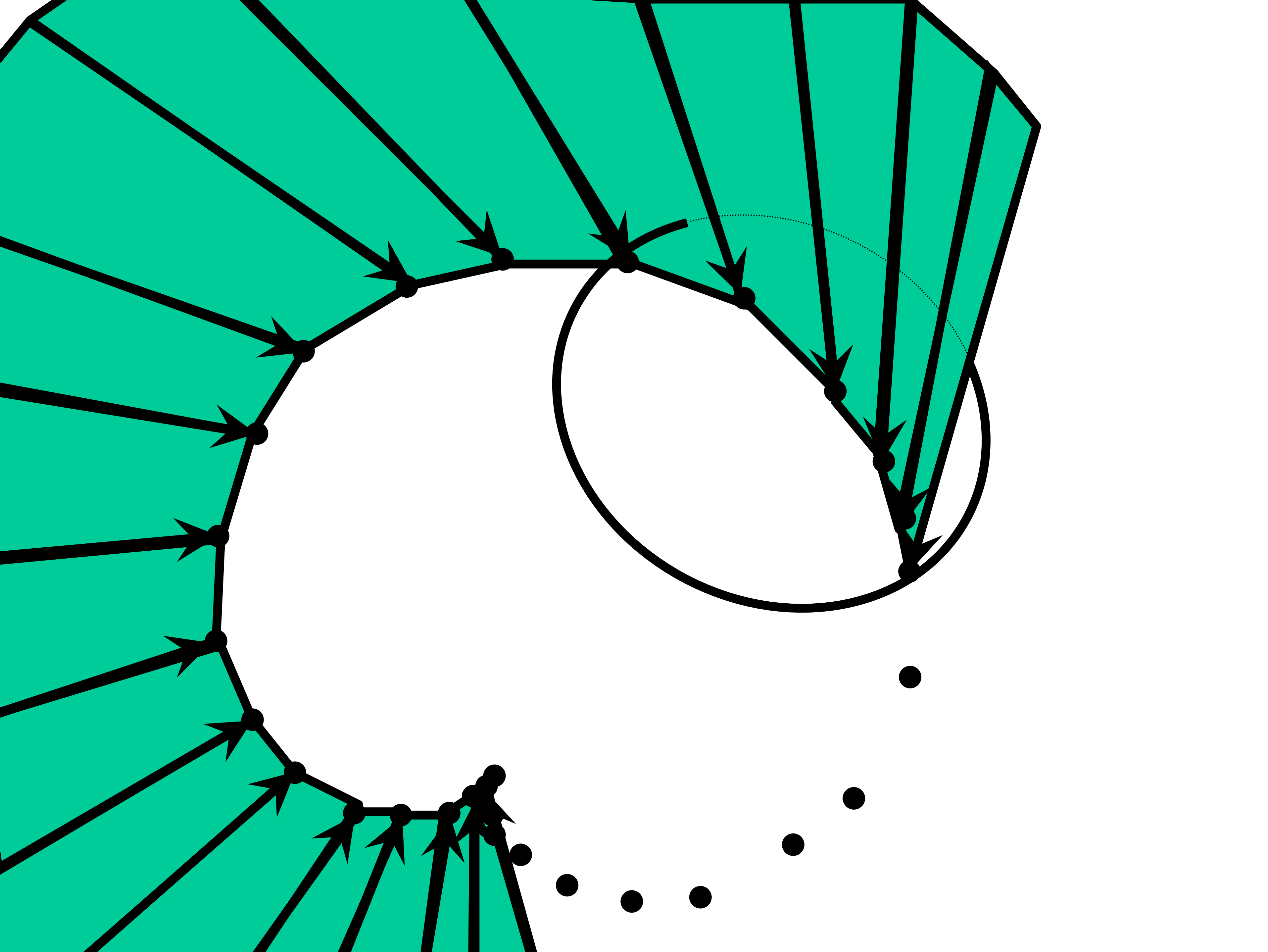


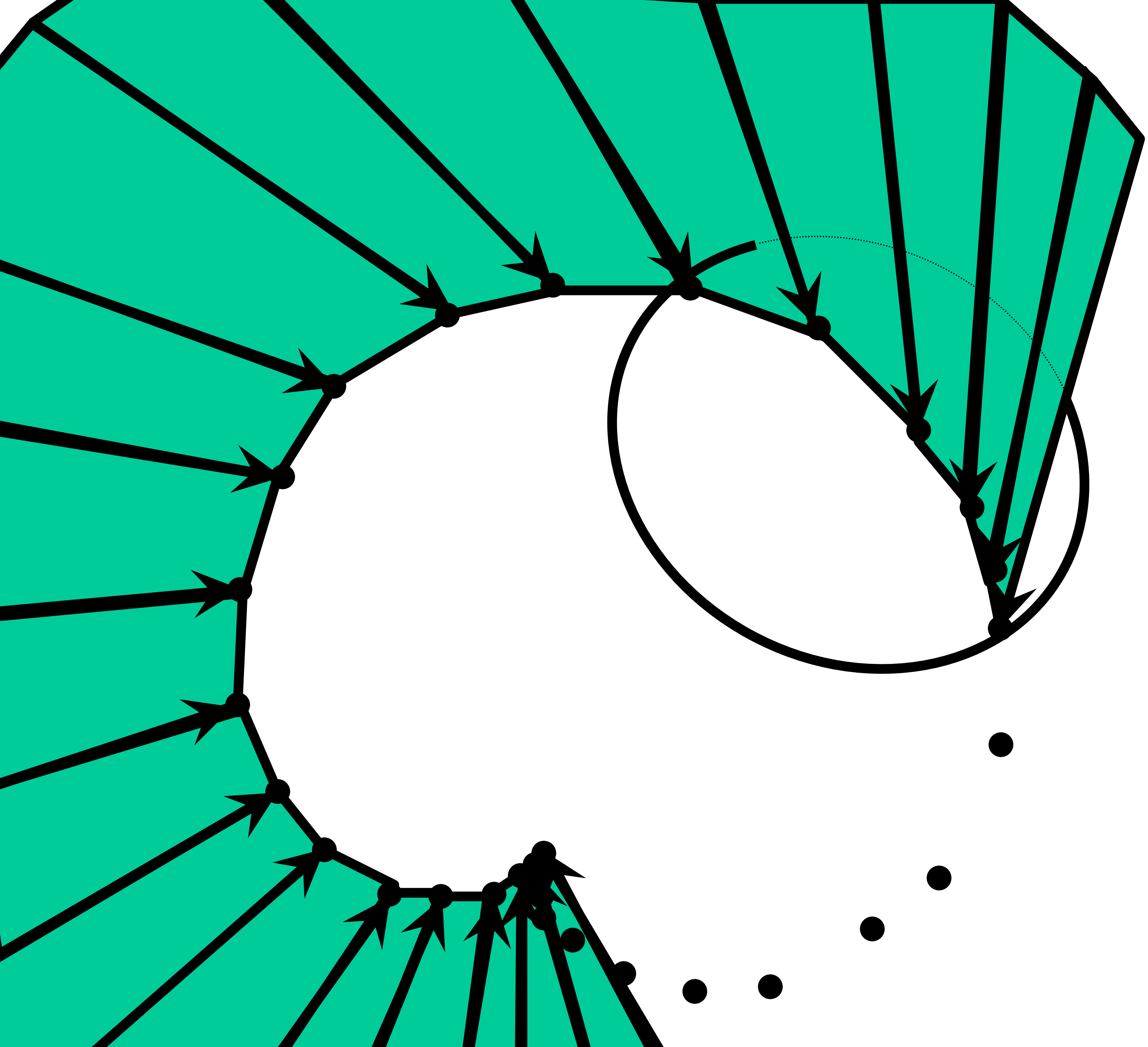


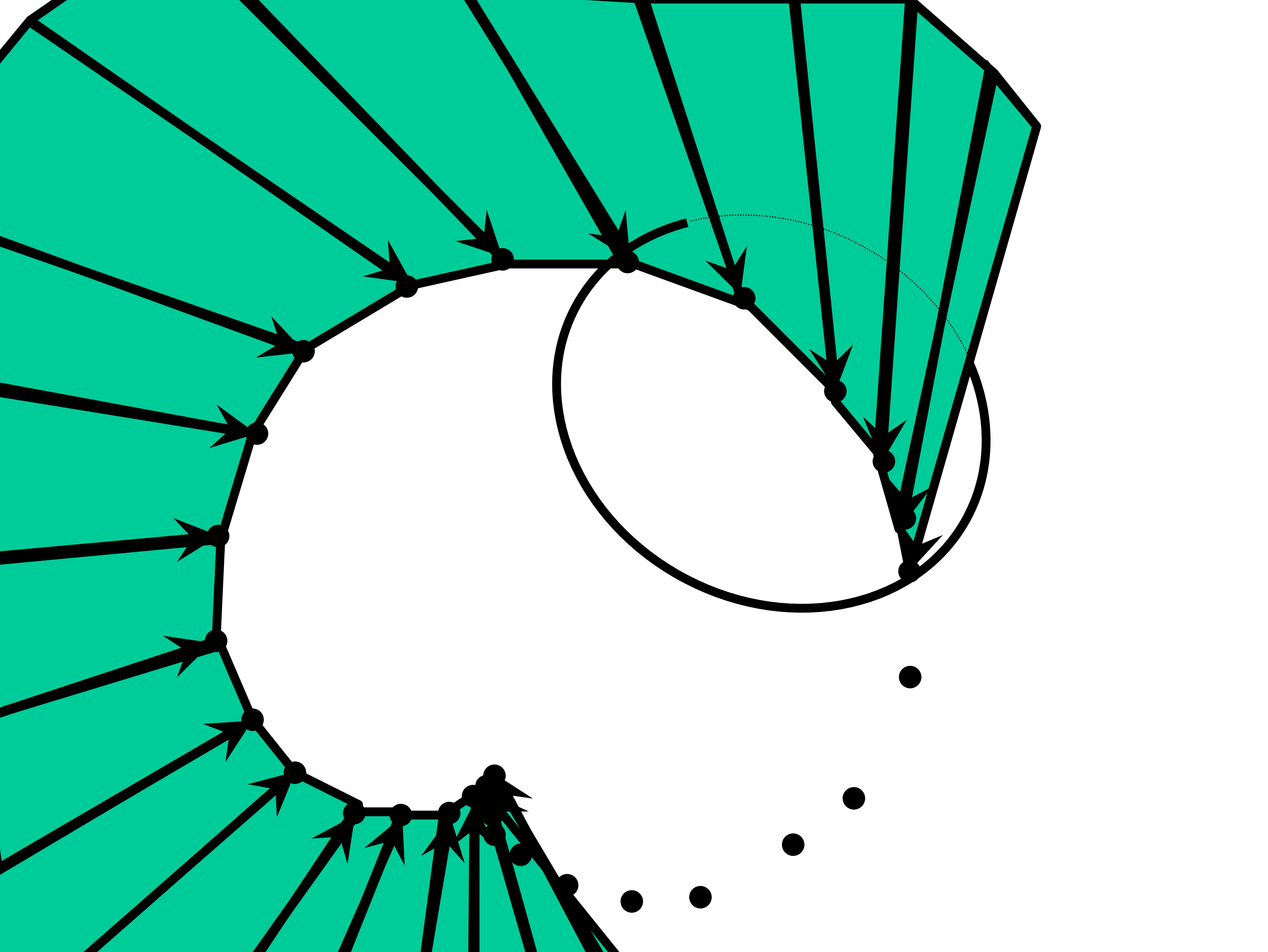

& Virtual
Environments

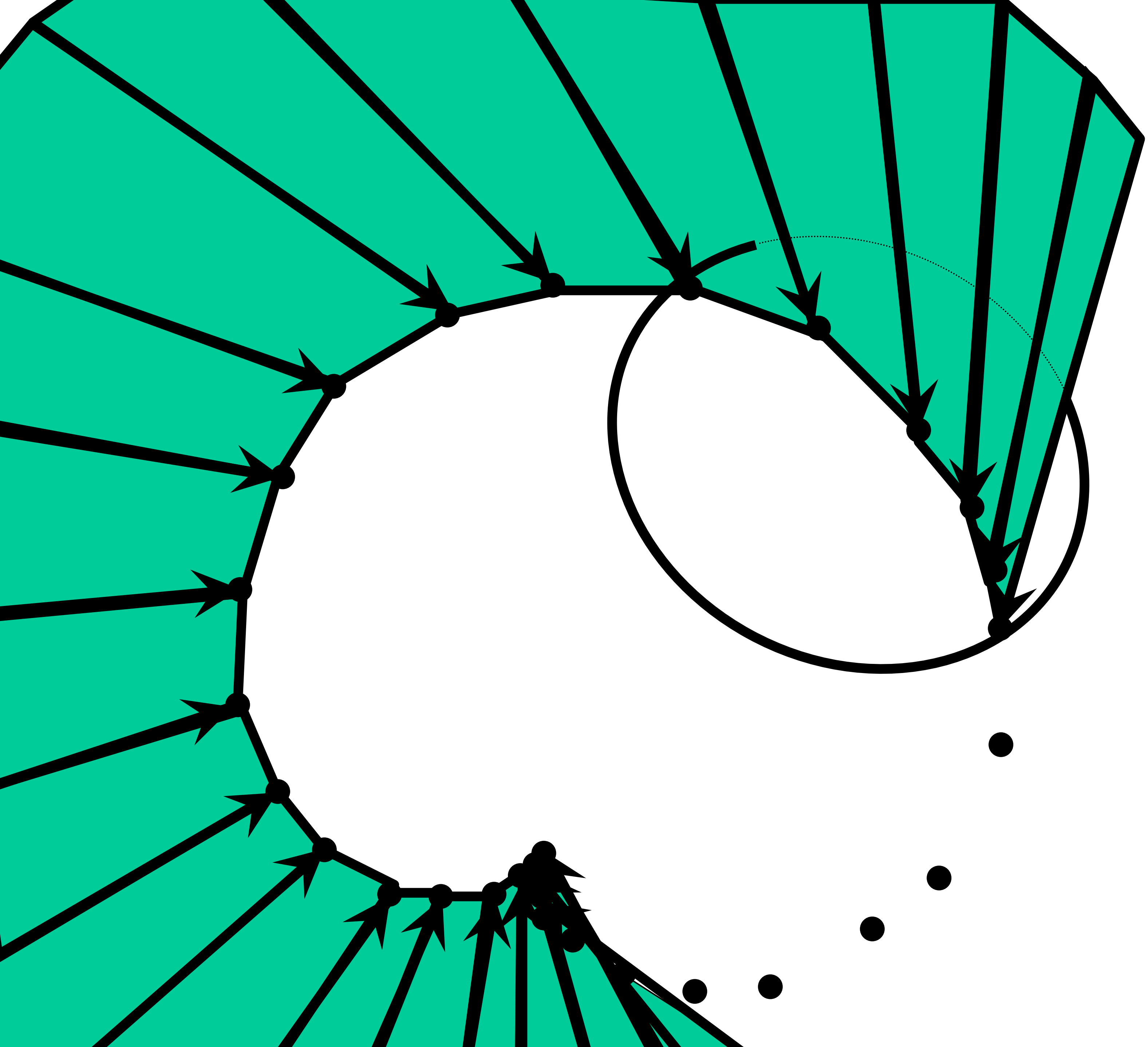


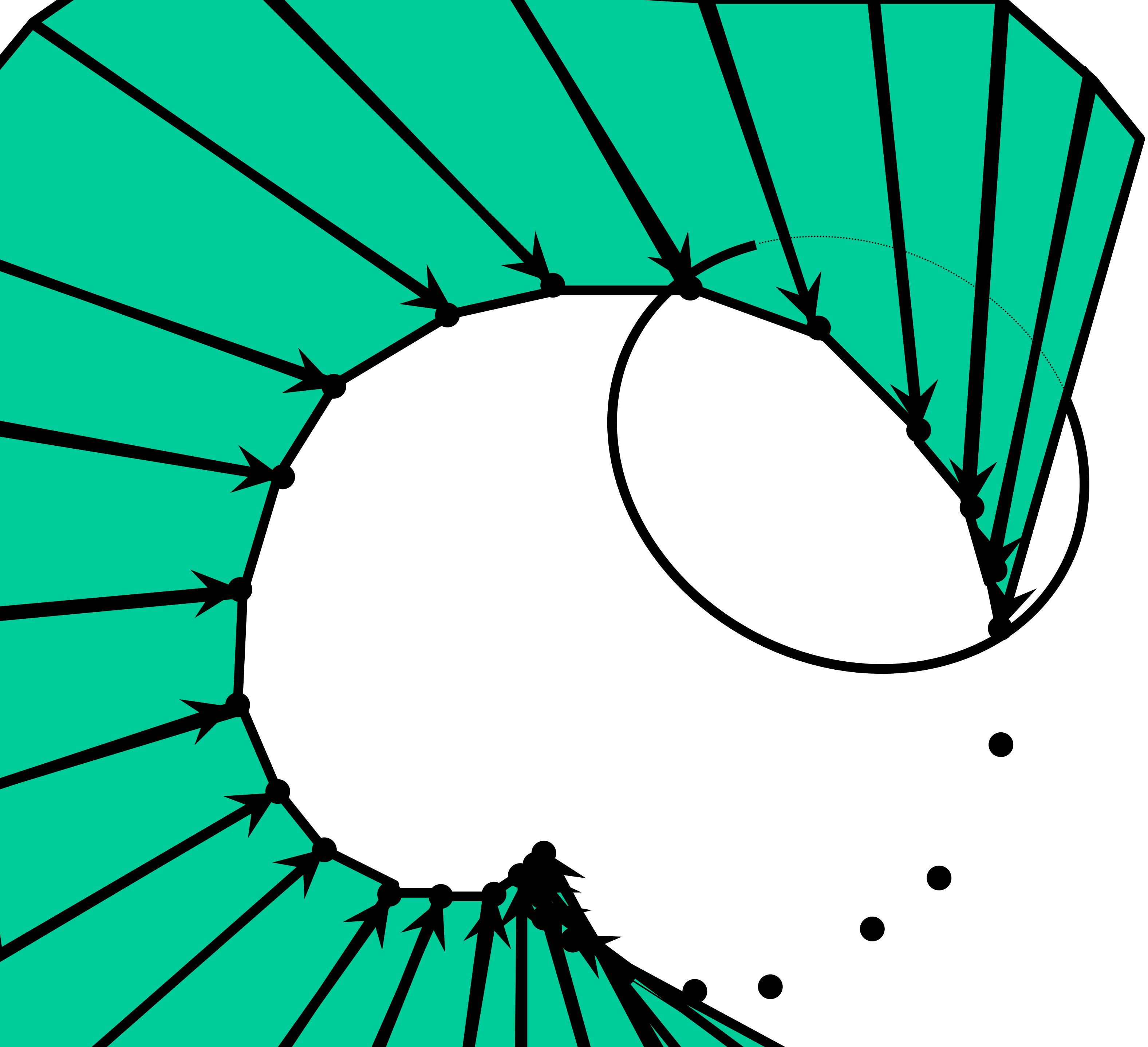


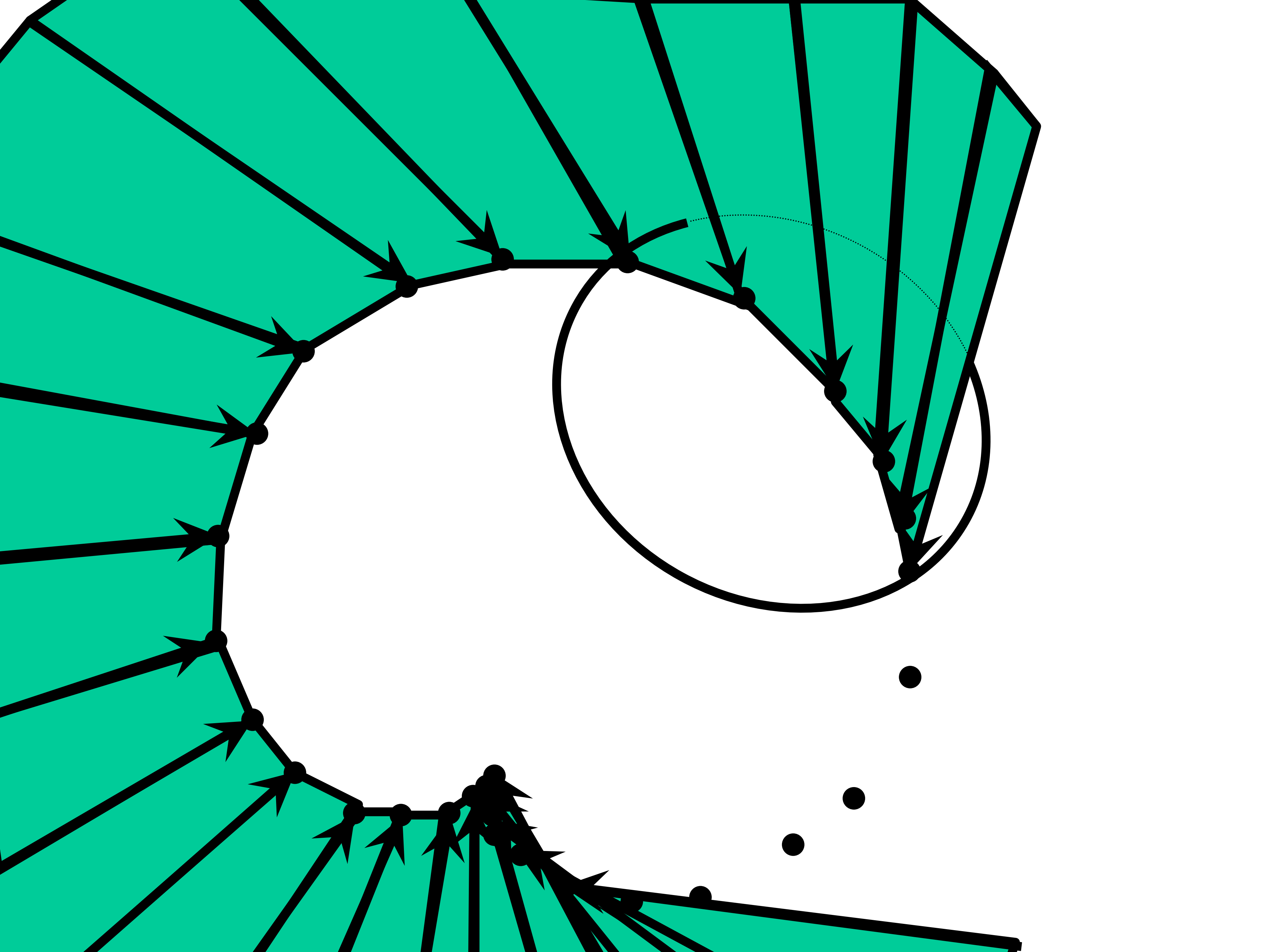


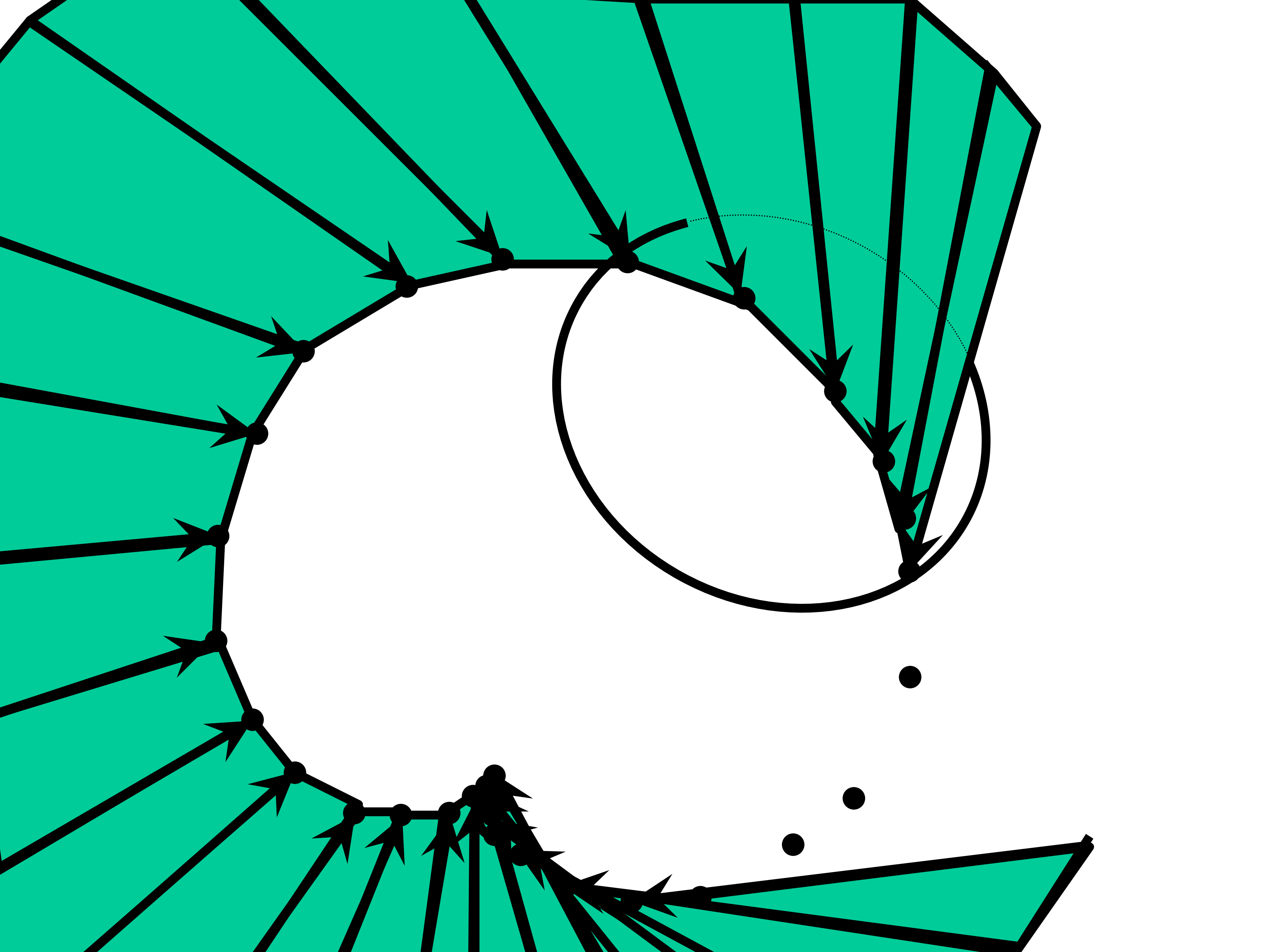


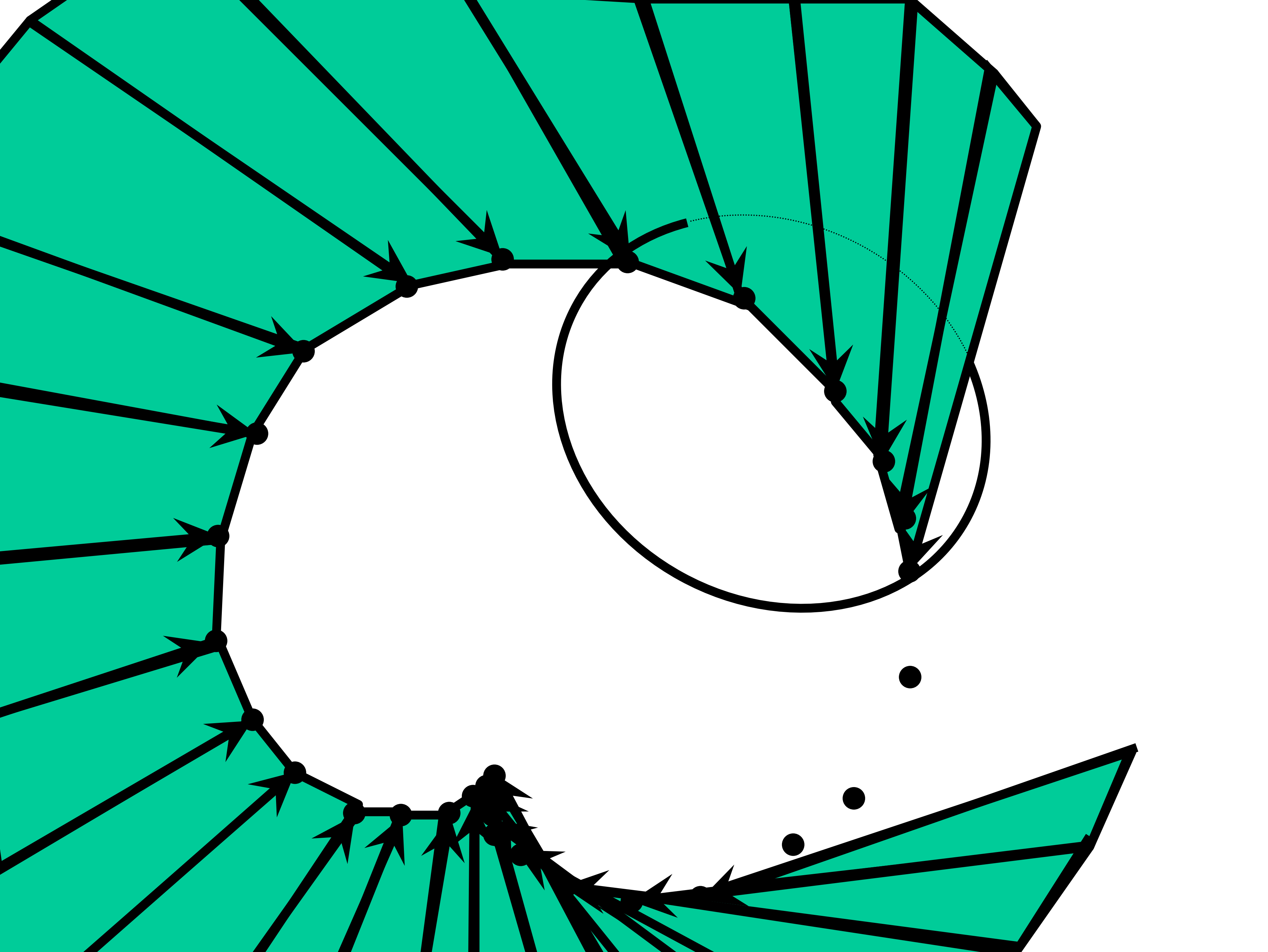


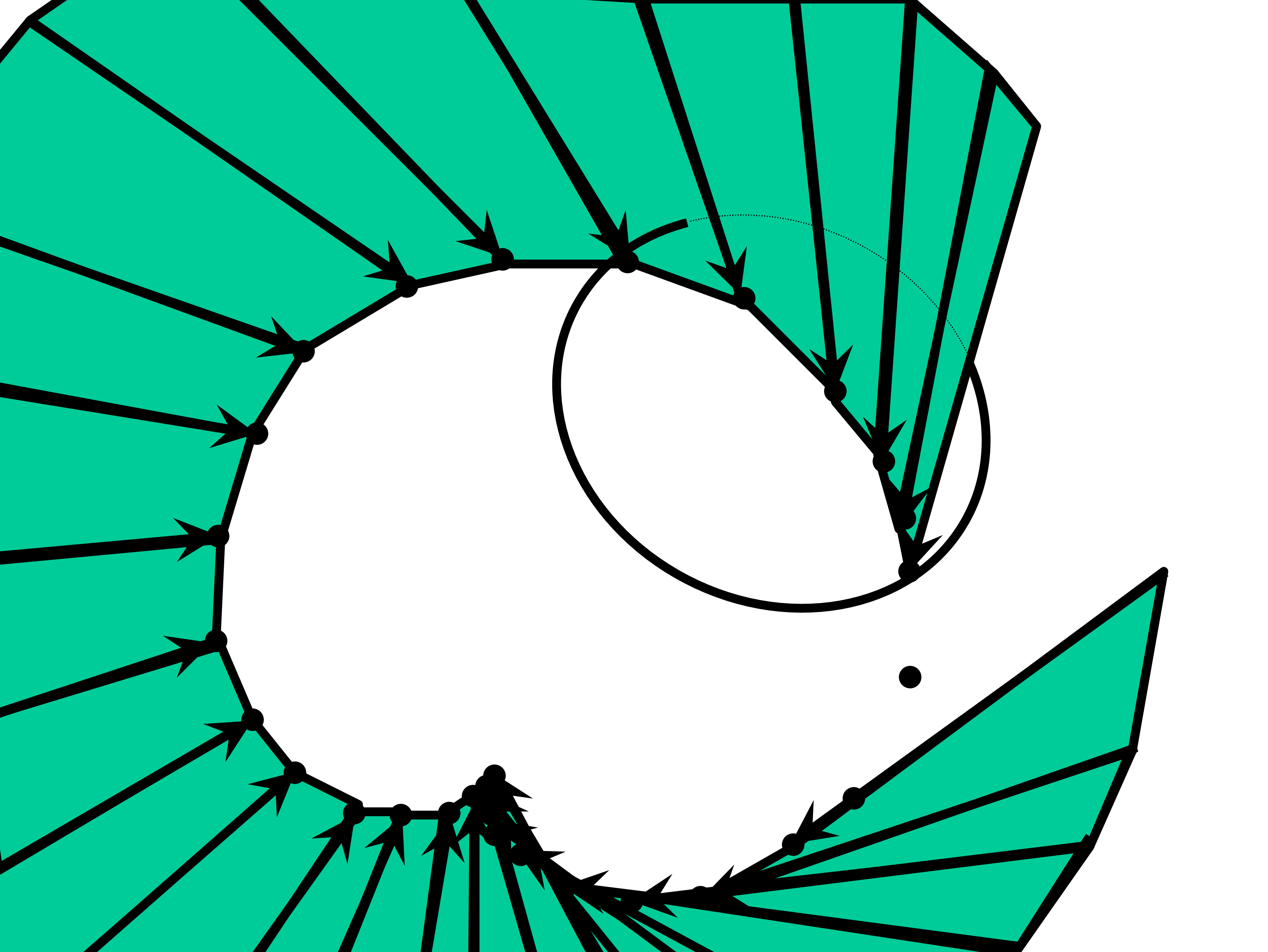


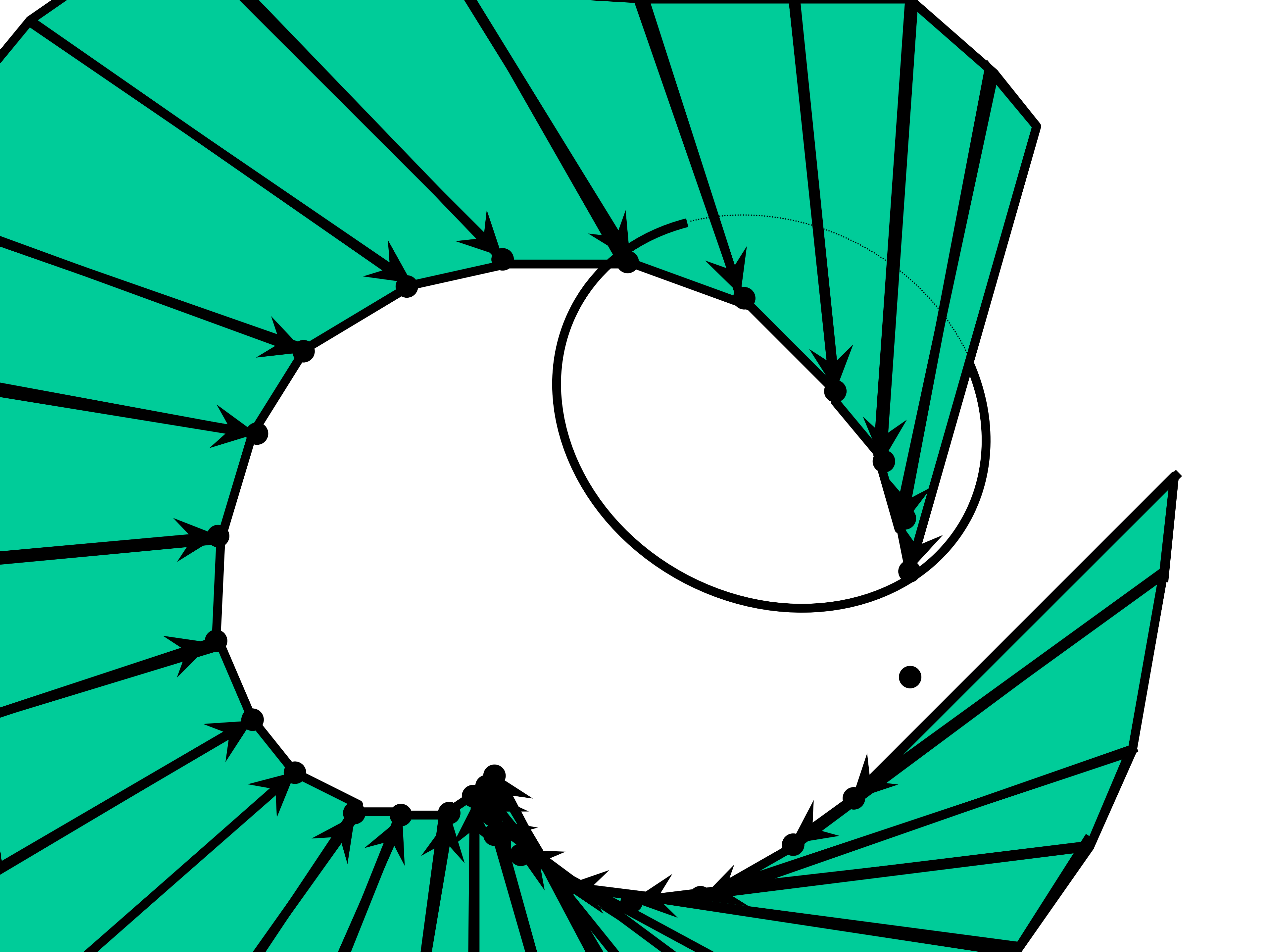


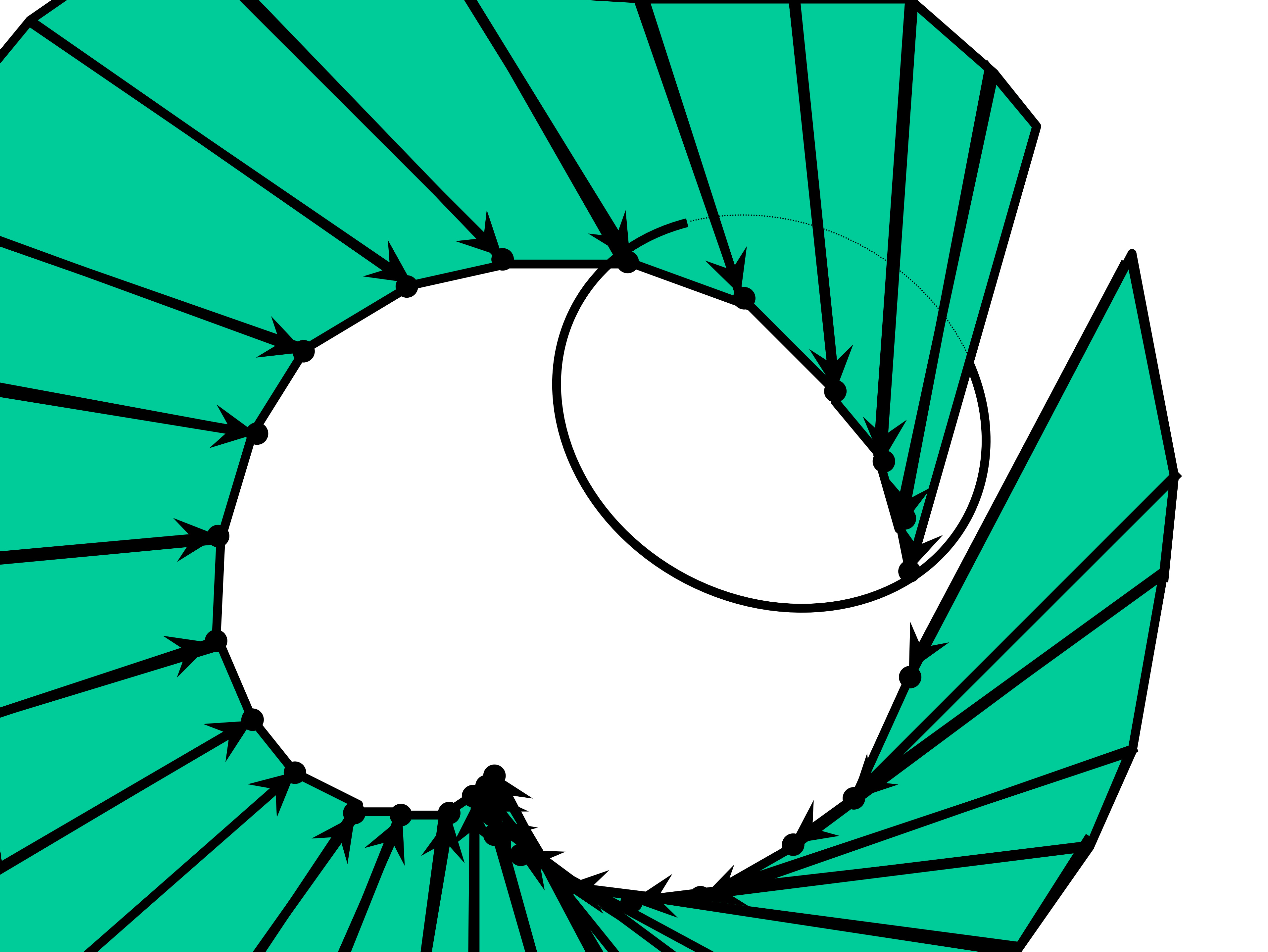


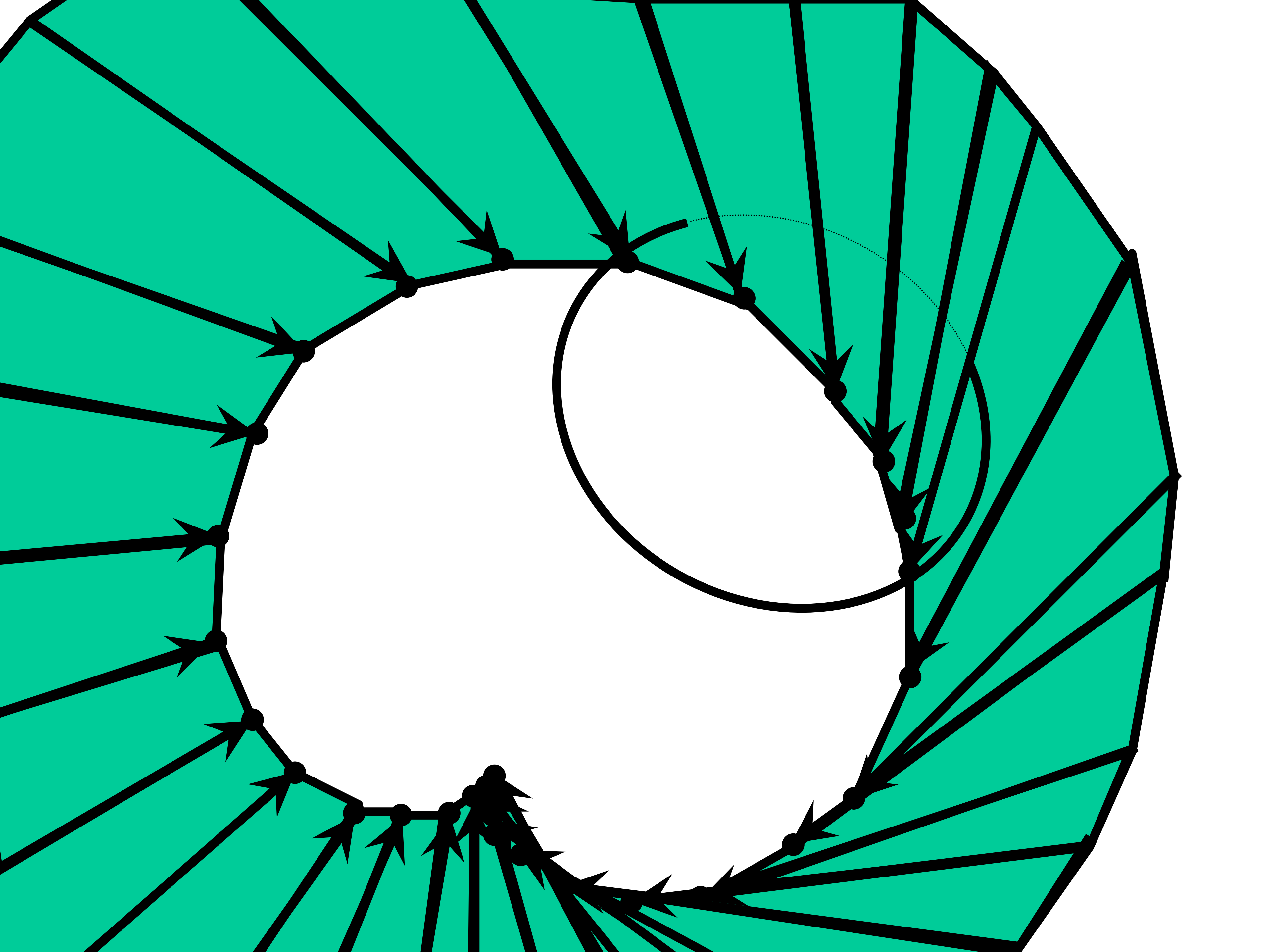


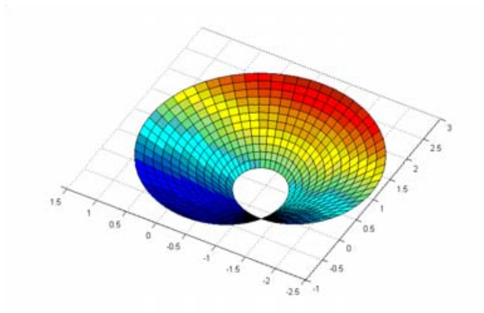
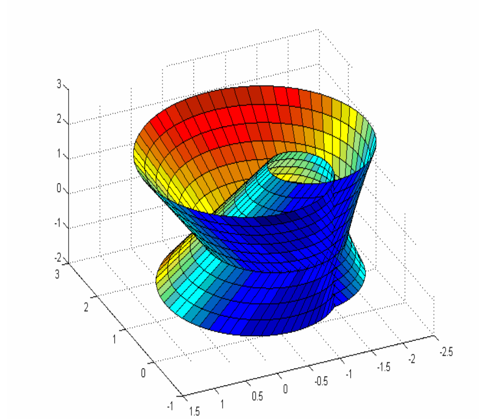
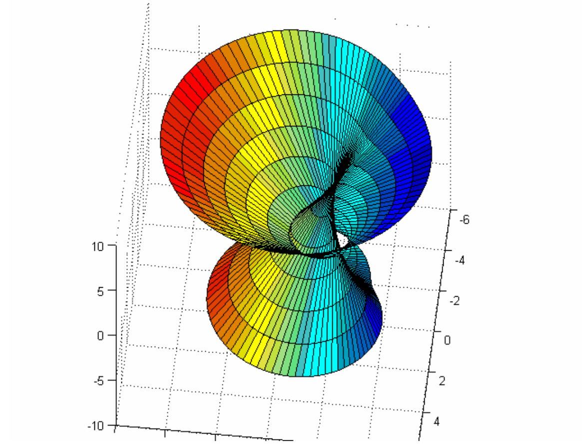
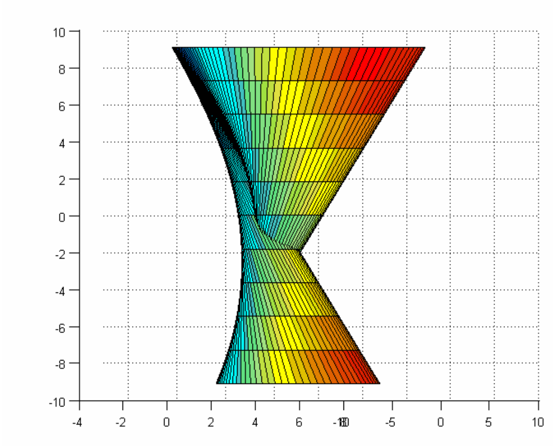
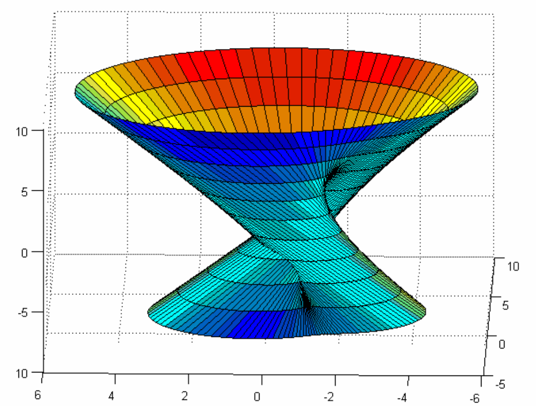
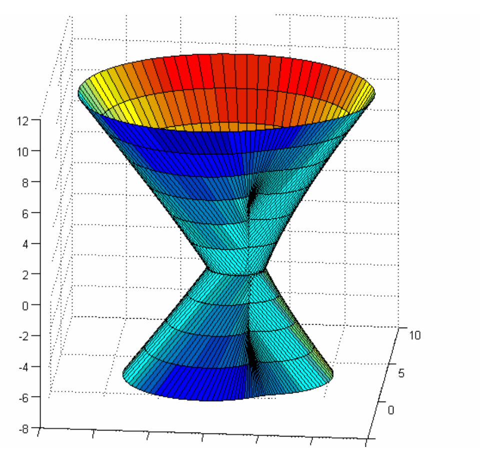
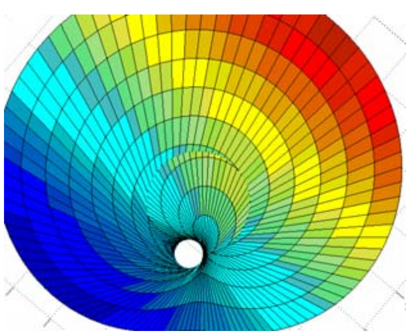
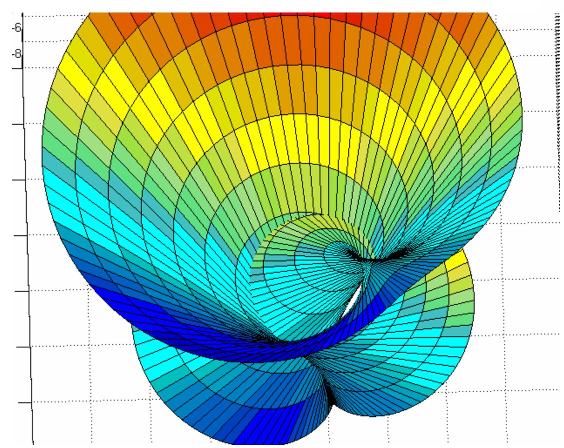
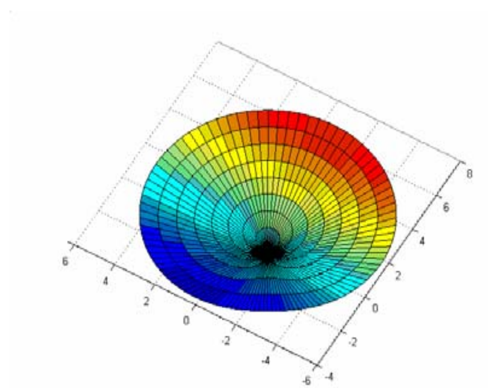






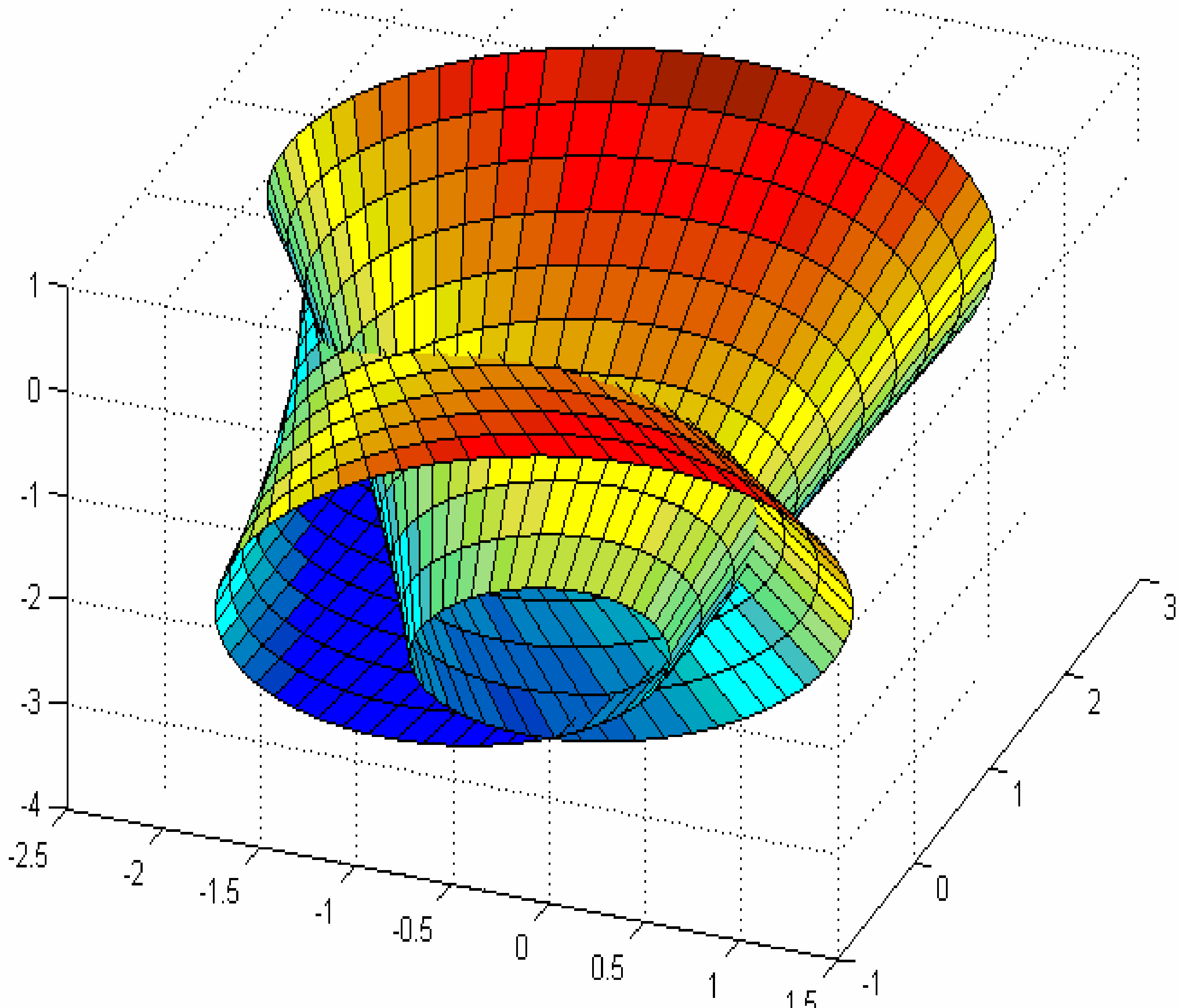




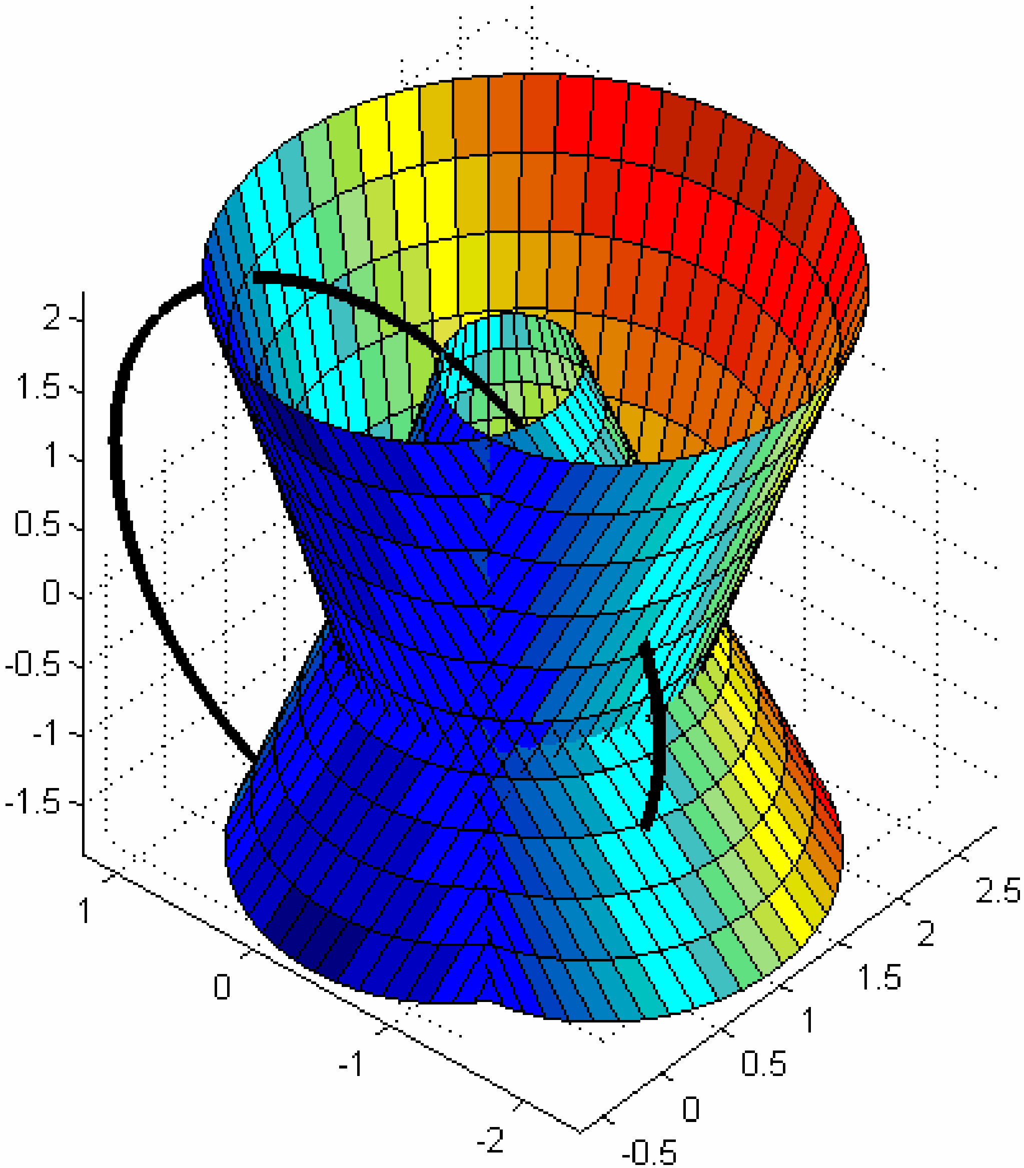










UK

Center for Visualization & Virtual Environments

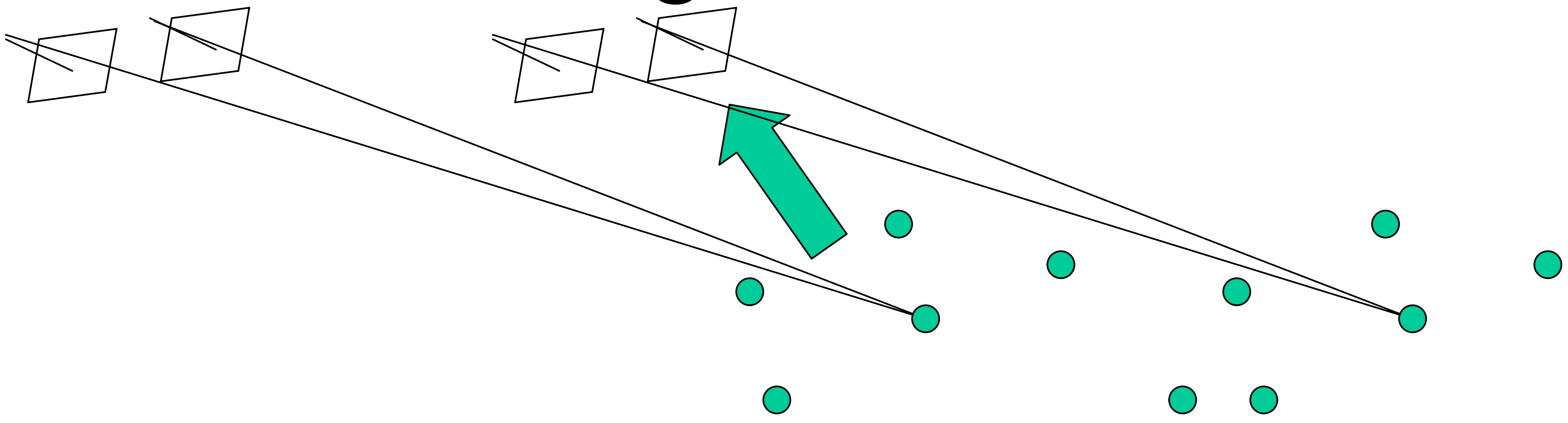

1
0
-1
-2
-3
-4
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
-1
0
1
2
3

Seamlessly into the classical case

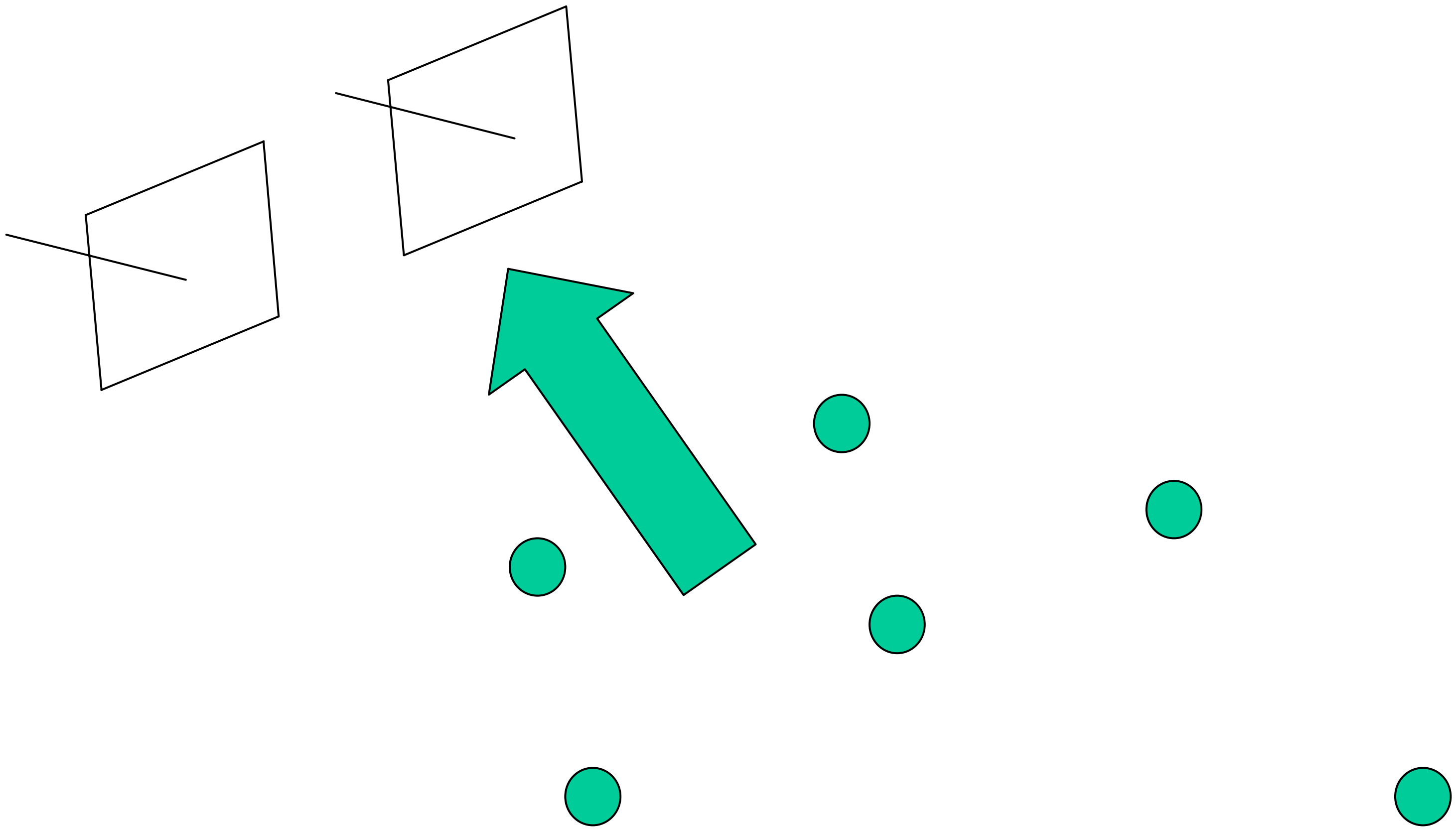


Environments

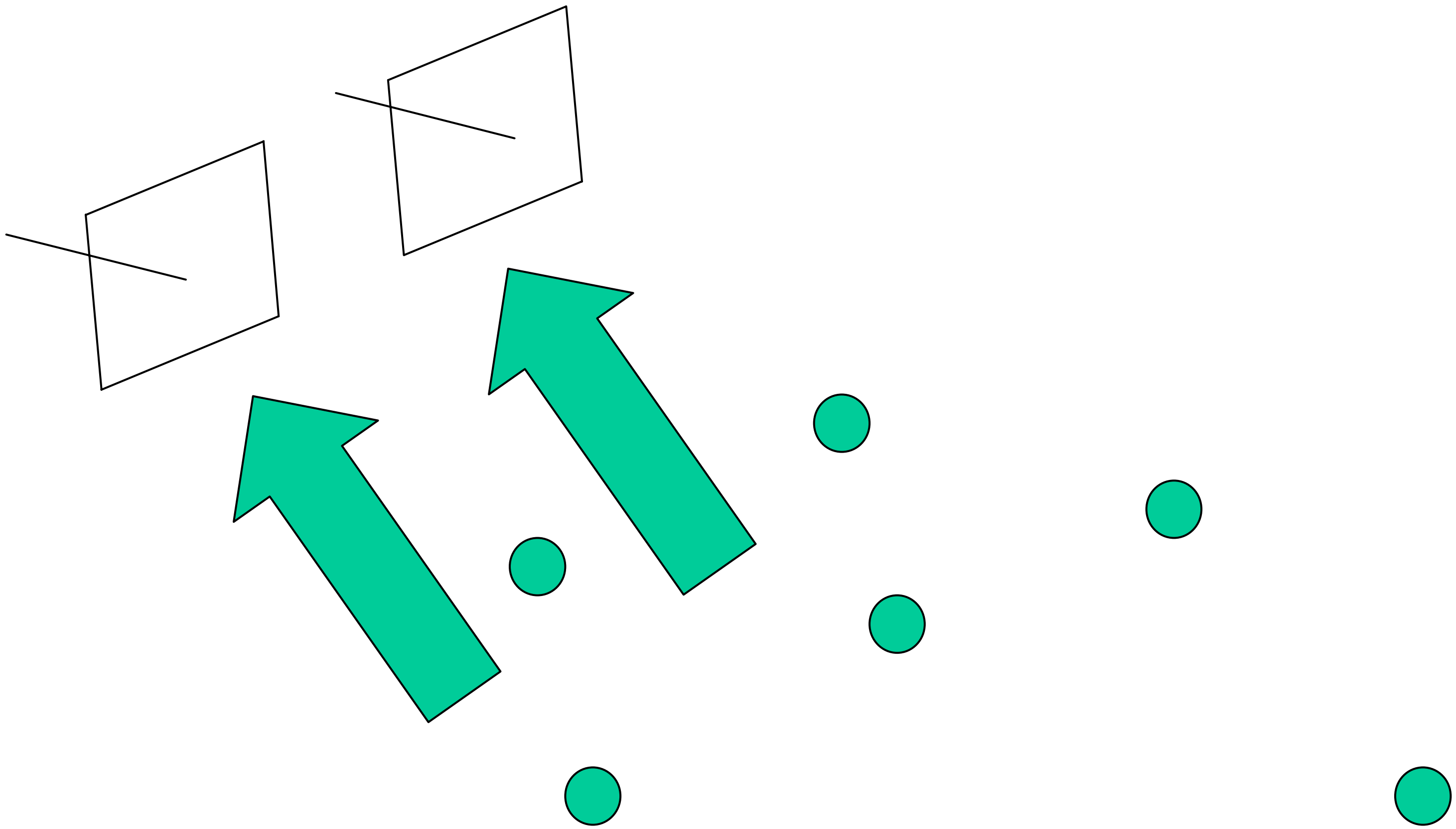
Moving Stereo Pair



Moving Stereo Pair



Moving Stereo Pair



Center for Visualization & Virtual Environments

6-point pose

$$[x]_\times PX = 0$$

Linear, stack 5 point constraints, results in pencil of cameras:

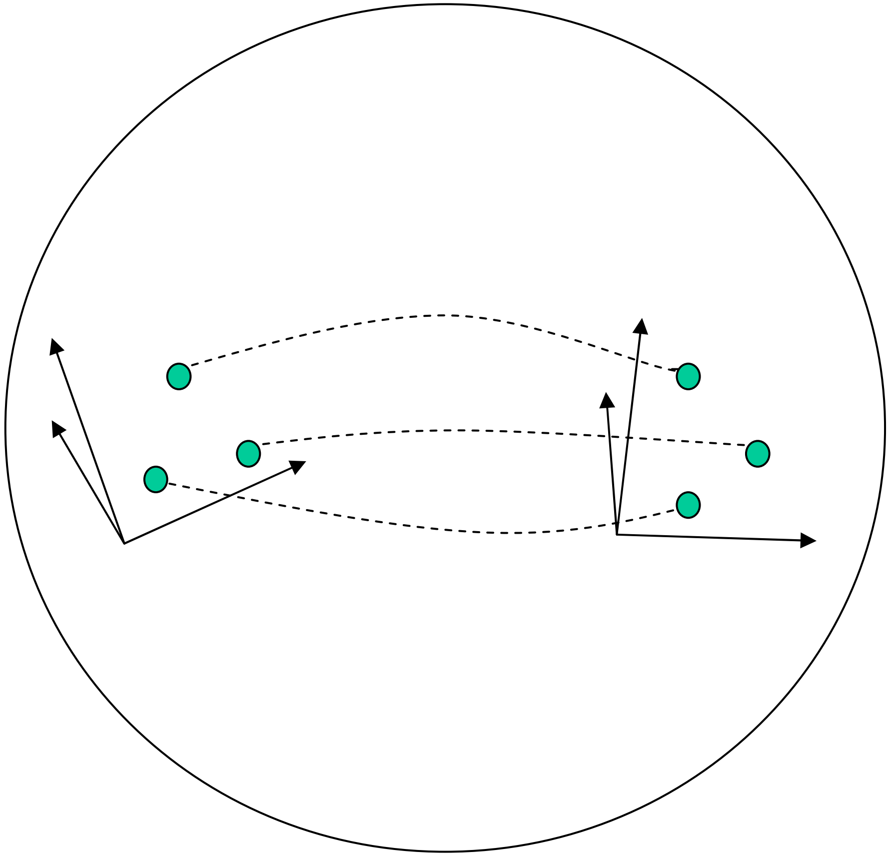
$$P = (1-a)P_1 + aP_2$$

Projects world point onto image line

$$x = (1-a)P_1 X + aP_2 X$$

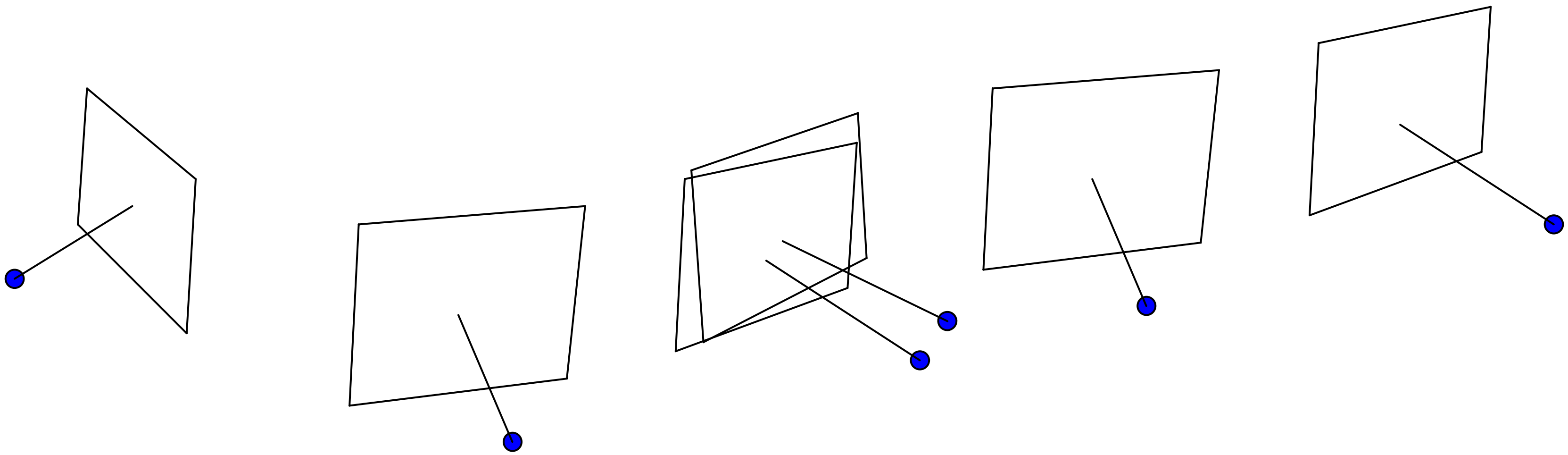
Correct point by perpendicular projection.
Add constraint and solve uniquely

Absolute Orientation
'Stitching'



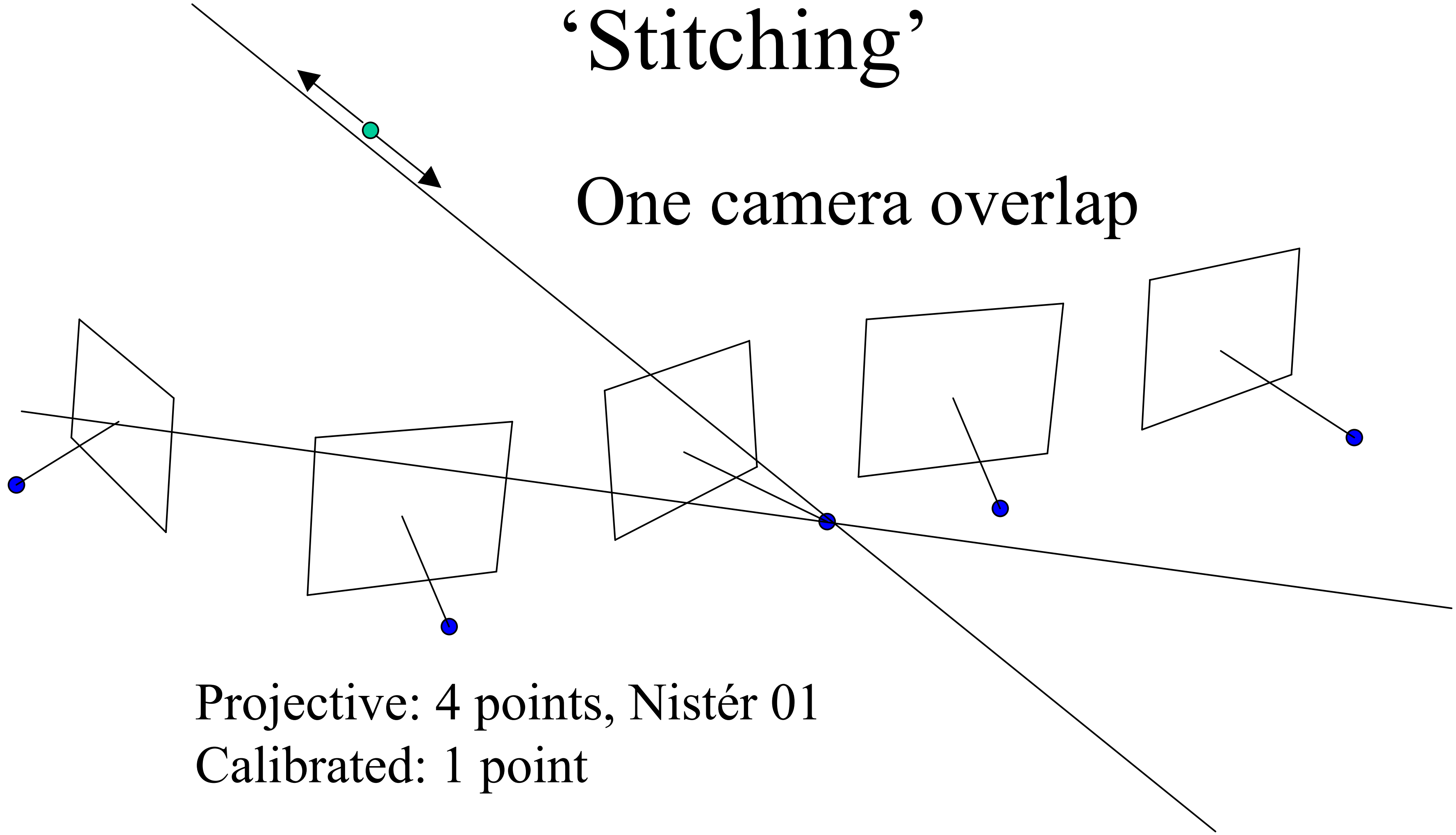
B. Horn,
Closed-Form Solution of Absolute Orientation using Unit Quaternions

Absolute Orientation
'Stitching'



Center for
Visualization & Virtual
Environments

Absolute Orientation
‘Stitching’



One camera overlap

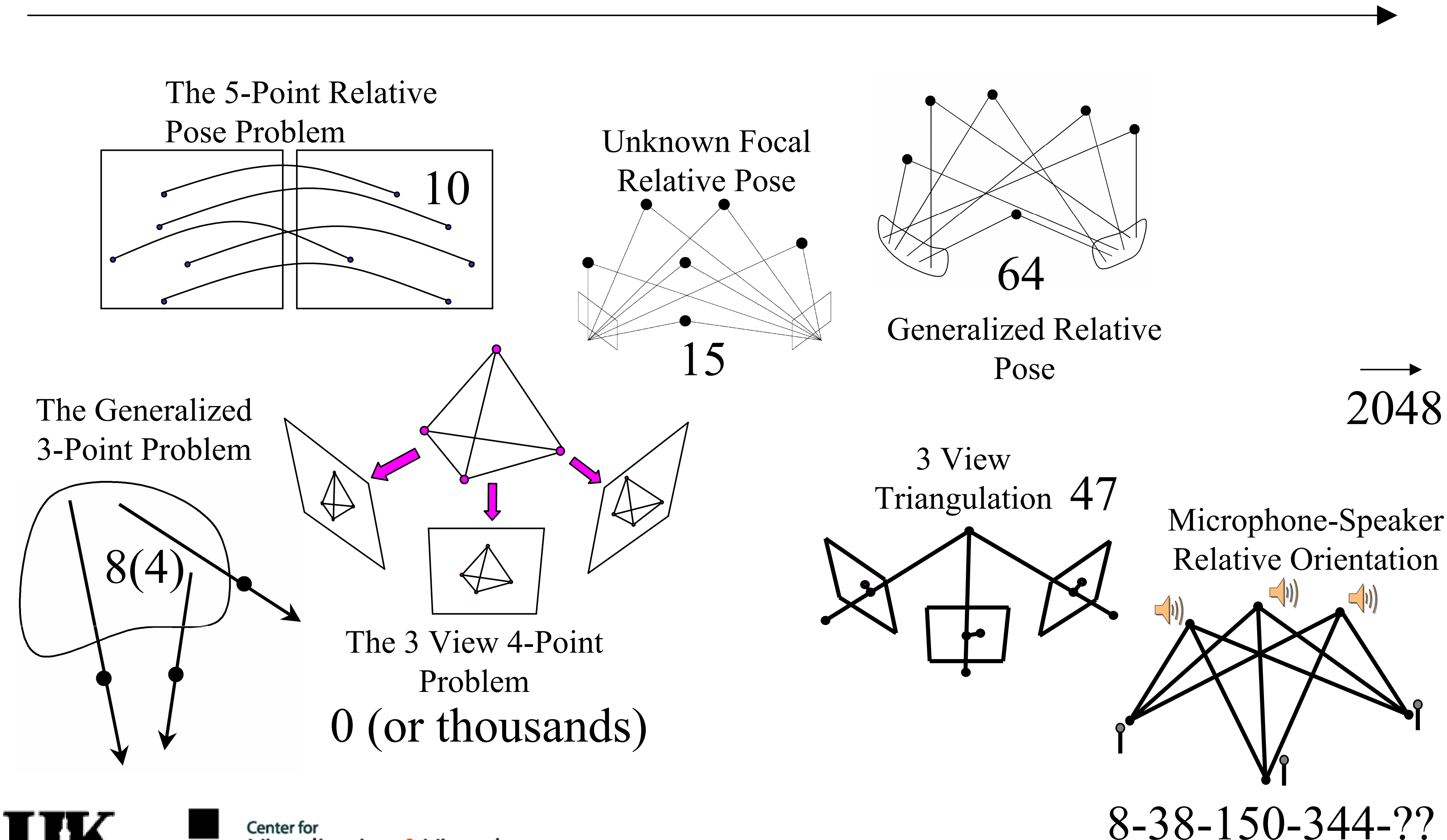
Projective: 4 points, Nistér 01
Calibrated: 1 point

Algebraic Geometry

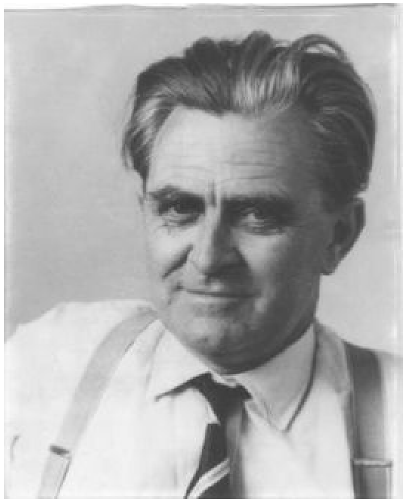
Geometry-Algebra ‘Dualism’

- Hilbert’s Nullstellensatz $\mathrm{I}(\mathrm{V}(J)) = \sqrt{J}$

Hypothesis Generation



UK Center for Visualization & Virtual Environments



Wolfgang Gröbner (1899-1980)

Bruno Buchberger

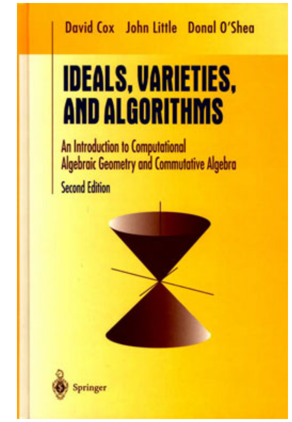
RISC

Research **I**nstitute for **S**ymbolic **C**omputation

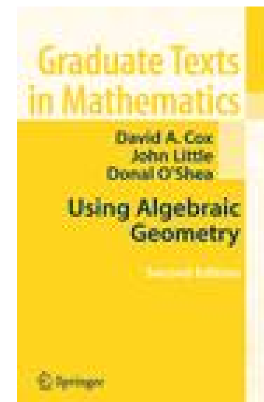
Linz, Austria

UK Center for Visualization & Virtual Environments

Suggested Literature

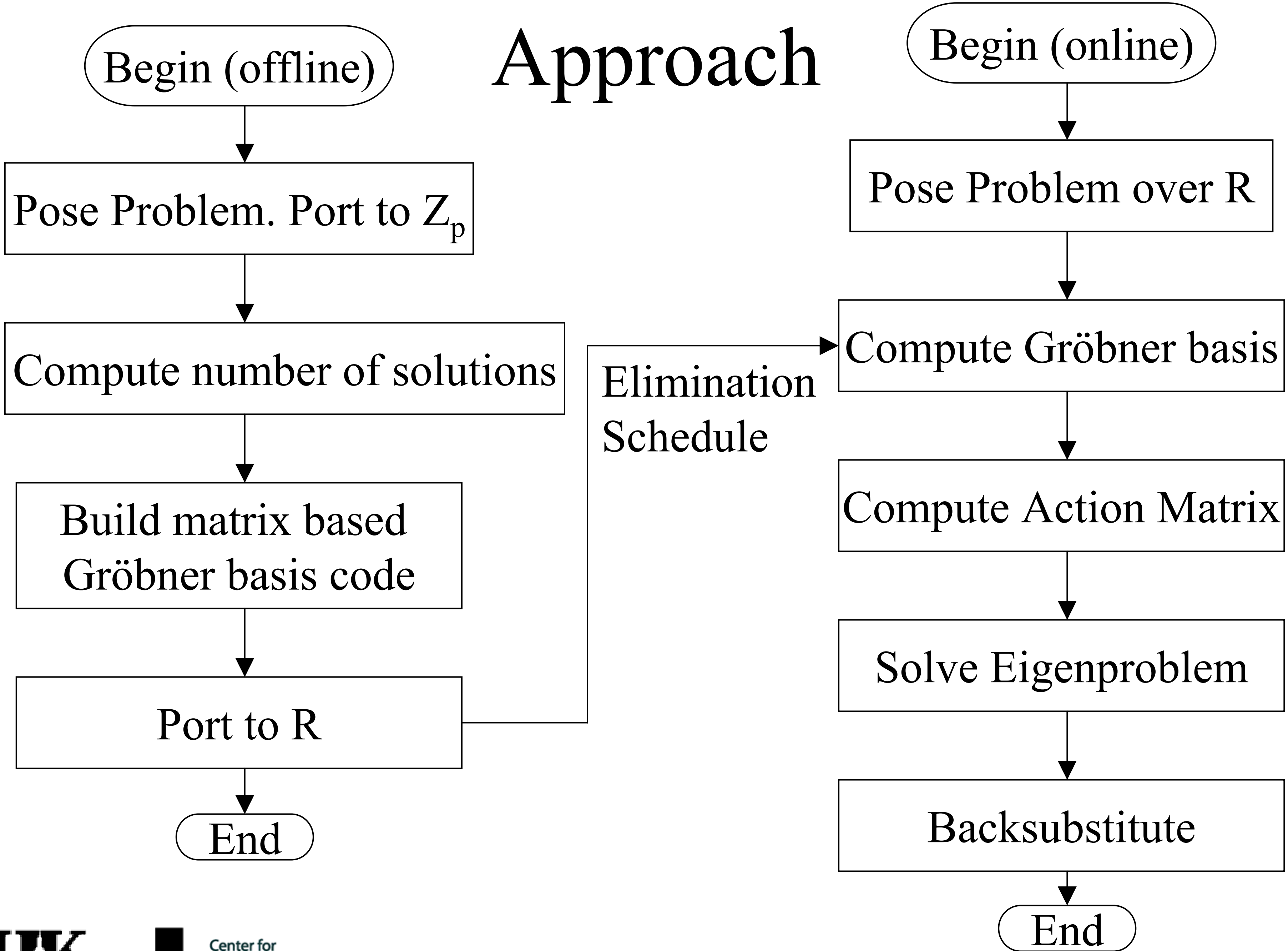


- D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms*, Second Edition, 1996.



- D. Cox, J. Little, D. O'Shea, Using Algebraic Geometry, Springer 1998.
- T. Becker and Weispfennig, Gröbner Bases, A Computational Approach to commutative Algebra, Springer 1993.

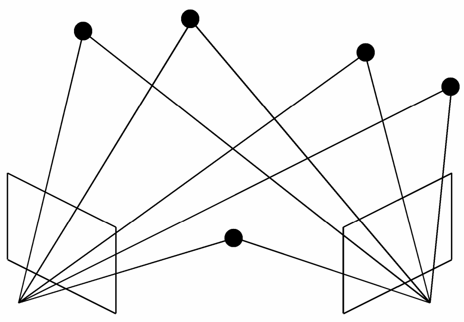
Approach

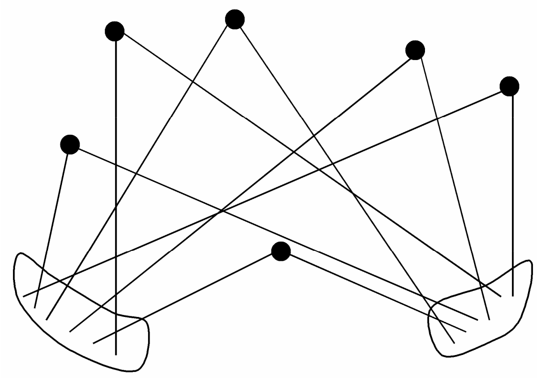


Begin (offline)
Pose Problem. Port to Z_p
Compute number of solutions
Build matrix based Gröbner basis code
Port to R
End
Elimination Schedule
Begin (online)
Pose Problem over R
Compute Gröbner basis
Compute Action Matrix
Solve Eigenproblem
Backsubstitute
End

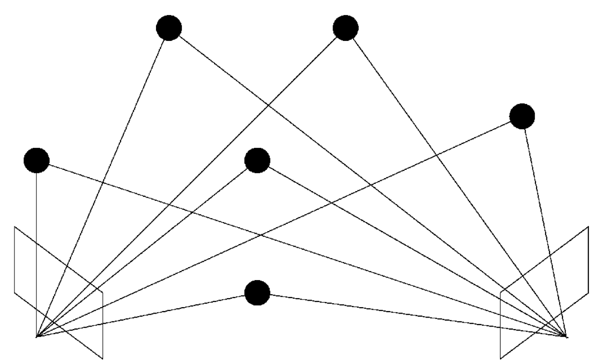
Examples of Solved Problems

6-point generalized relative orientation (64 solutions) (Stewenius, Nistér, Oskarsson and Åström, Omnivis 2005)





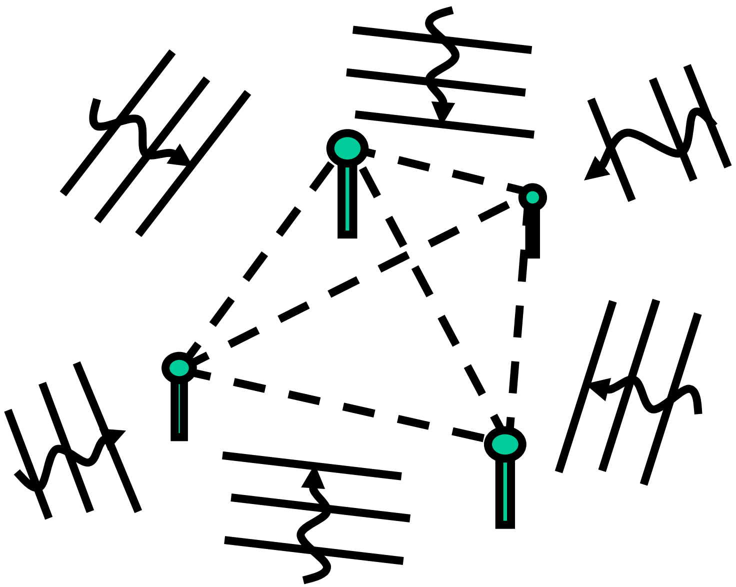
6-point relative orientation with common but unknown focal length (15 solutions) (Stewenius, Nistér, Schaffalitzky and Kahl, CVPR 2005)

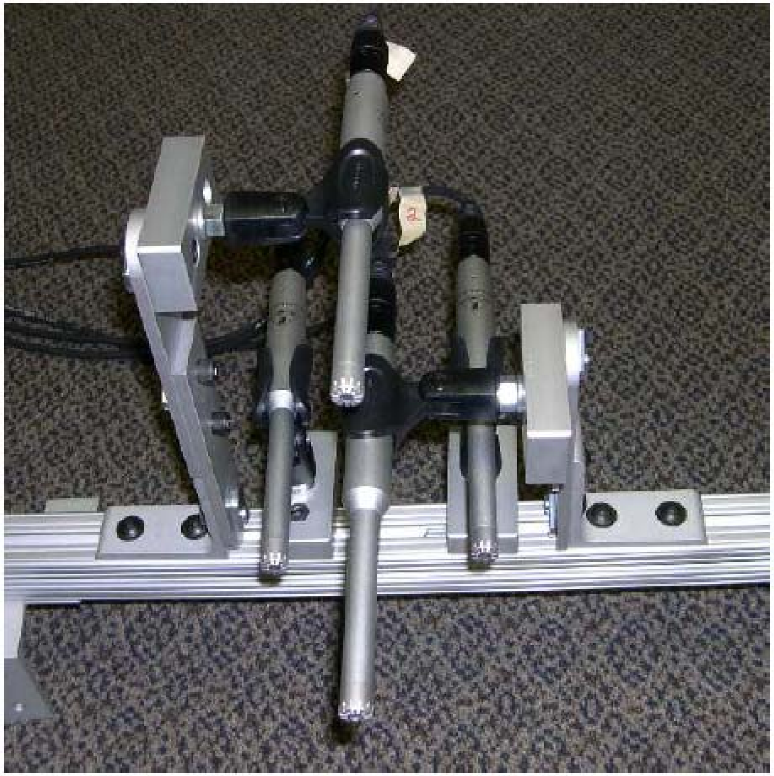


UK Center for Visualization & Virtual Environments

"Audio-Grammetry"

work with Henrik Stewenius, Jens Hannemann, Kevin Donahue



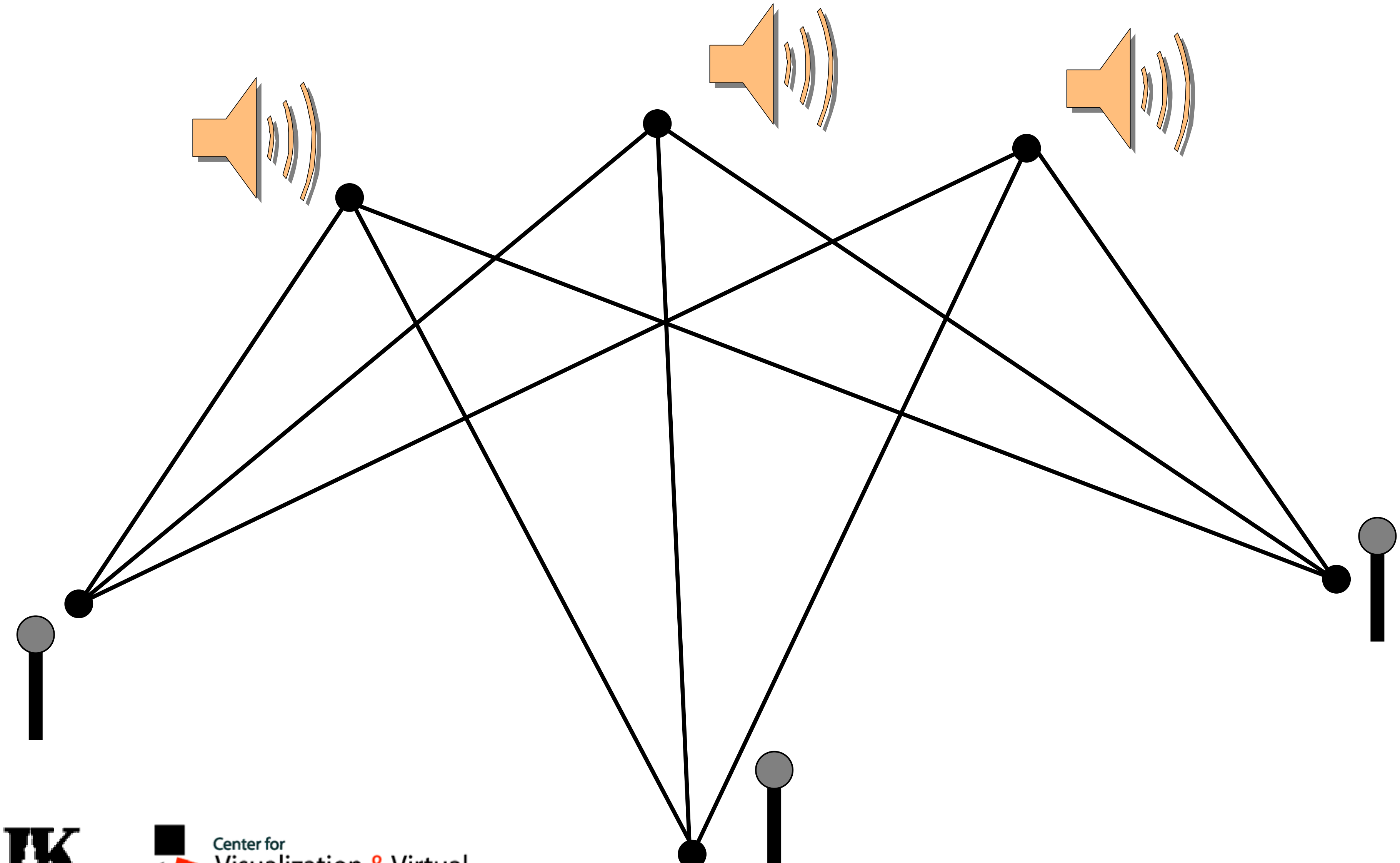


UK

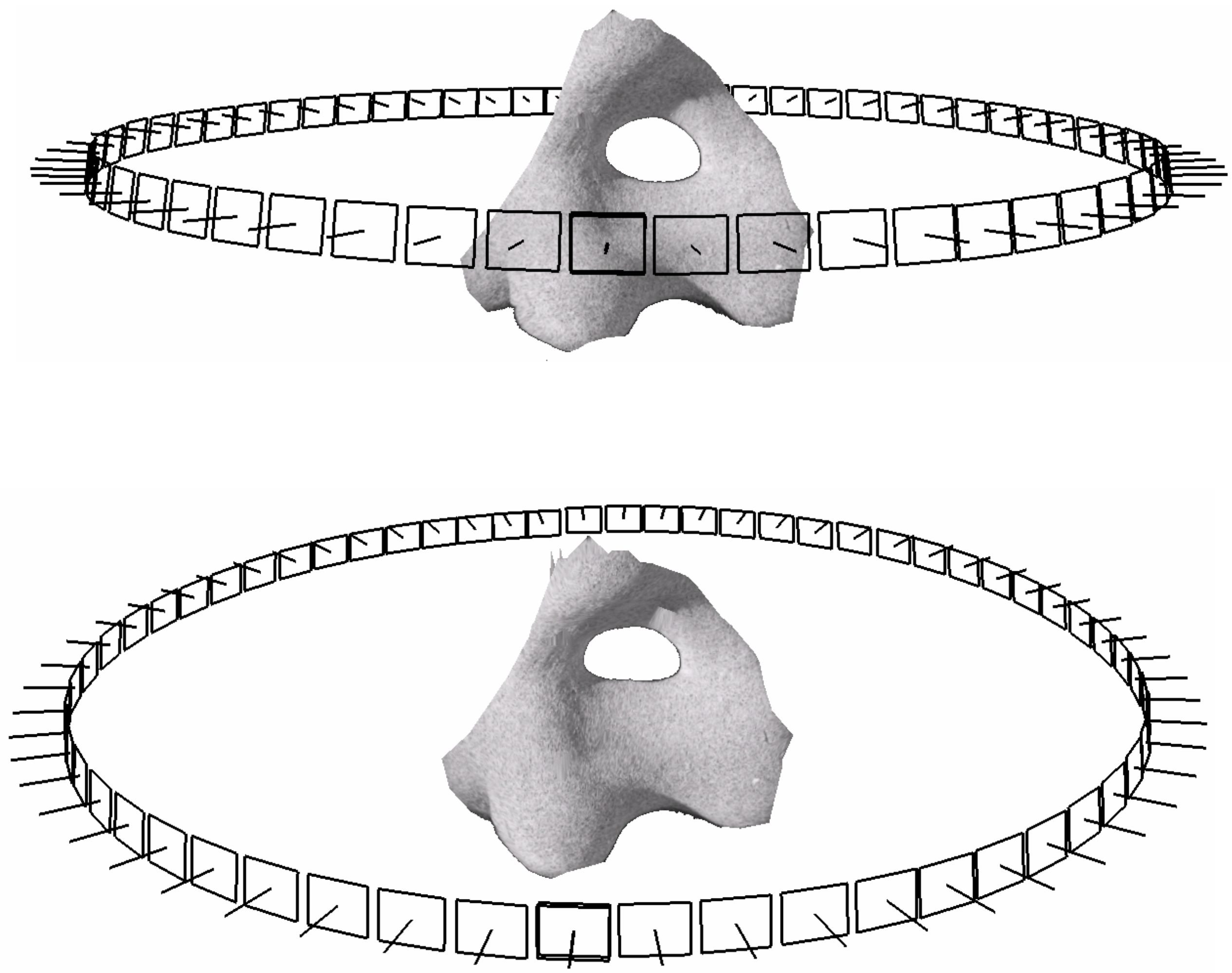
Center for Visualization & Virtual Environments

Microphone-Speaker Location

work with Henrik Stewenius, Jens Hannemann, Kevin Donahue



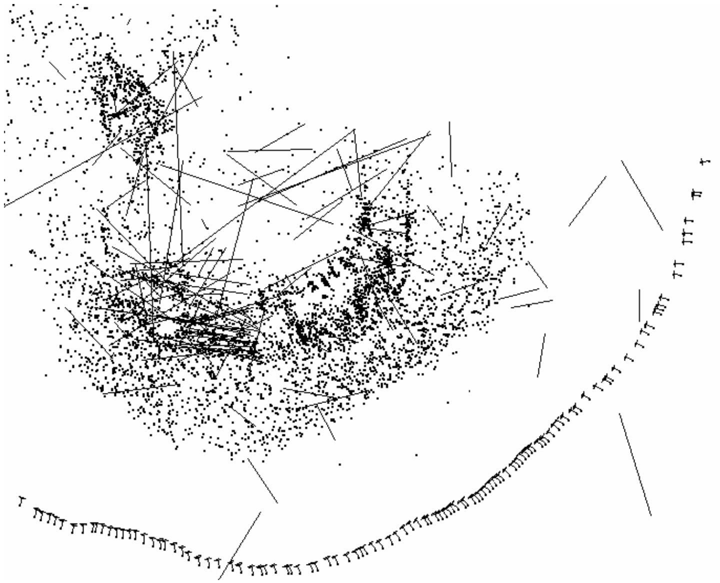
Center for Visualization & Virtual Environments



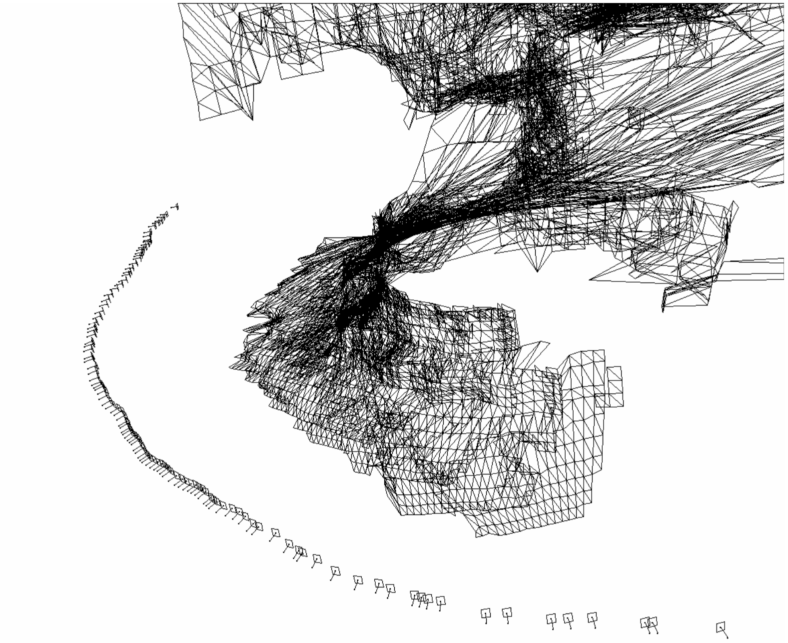
UK

Center for Visualization & Virtual Environments

Sparse



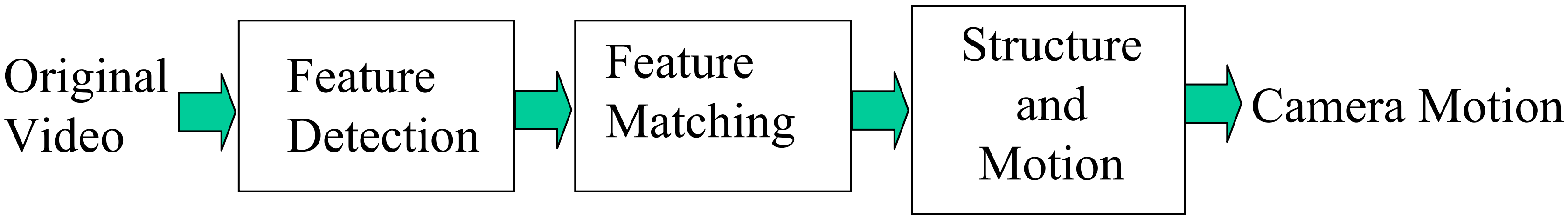
Dense



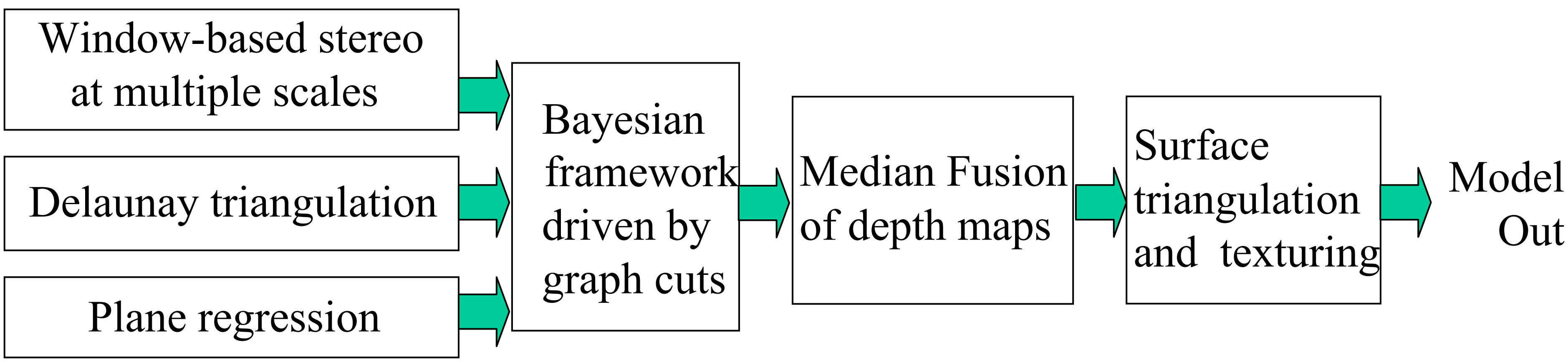
UK

Center for Visualization & Virtual Environments

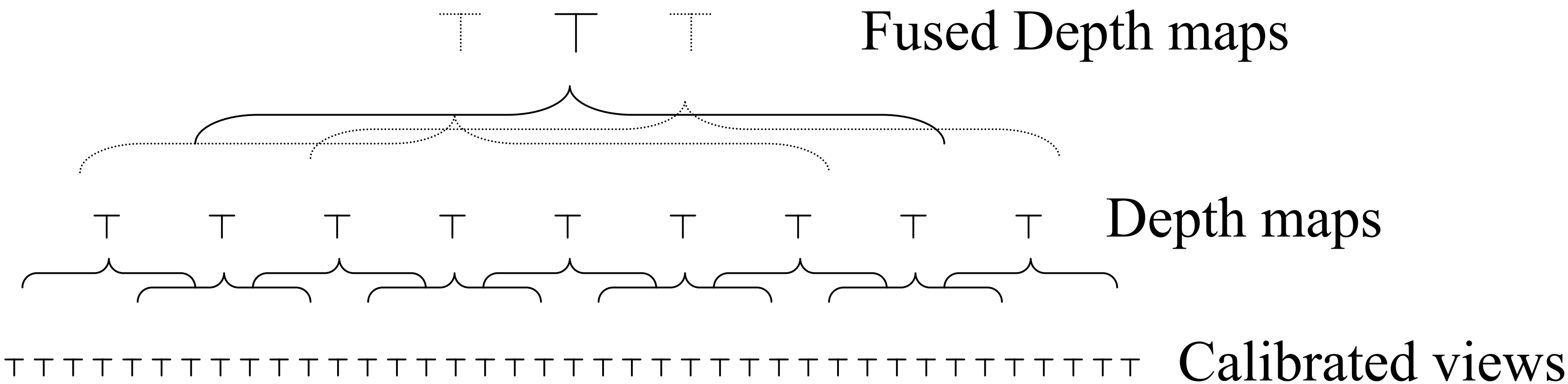
Sparse Reconstruction



Dense Reconstruction



Dense Reconstruction



Center for Visualization & Virtual Environments

Stereo

- Feature Based Stereo
- Classical Stereo
- Dynamic Programming
- Belief Propagation
- Graph Cuts
- Color Segmentation
- Plane Sweep
- Level Sets

Discontinuity Energy

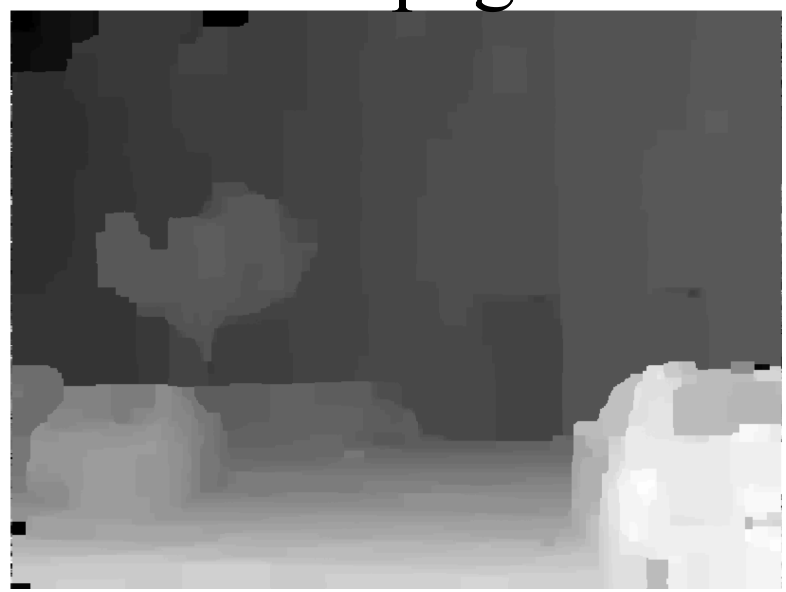
$$w_{MAP} = \arg\min_{w} \left[E_{d_2|w,d_1} + E_{w|d_1} \right]$$

Dissimilarity Energy

Multi-View Depth Reconstruction

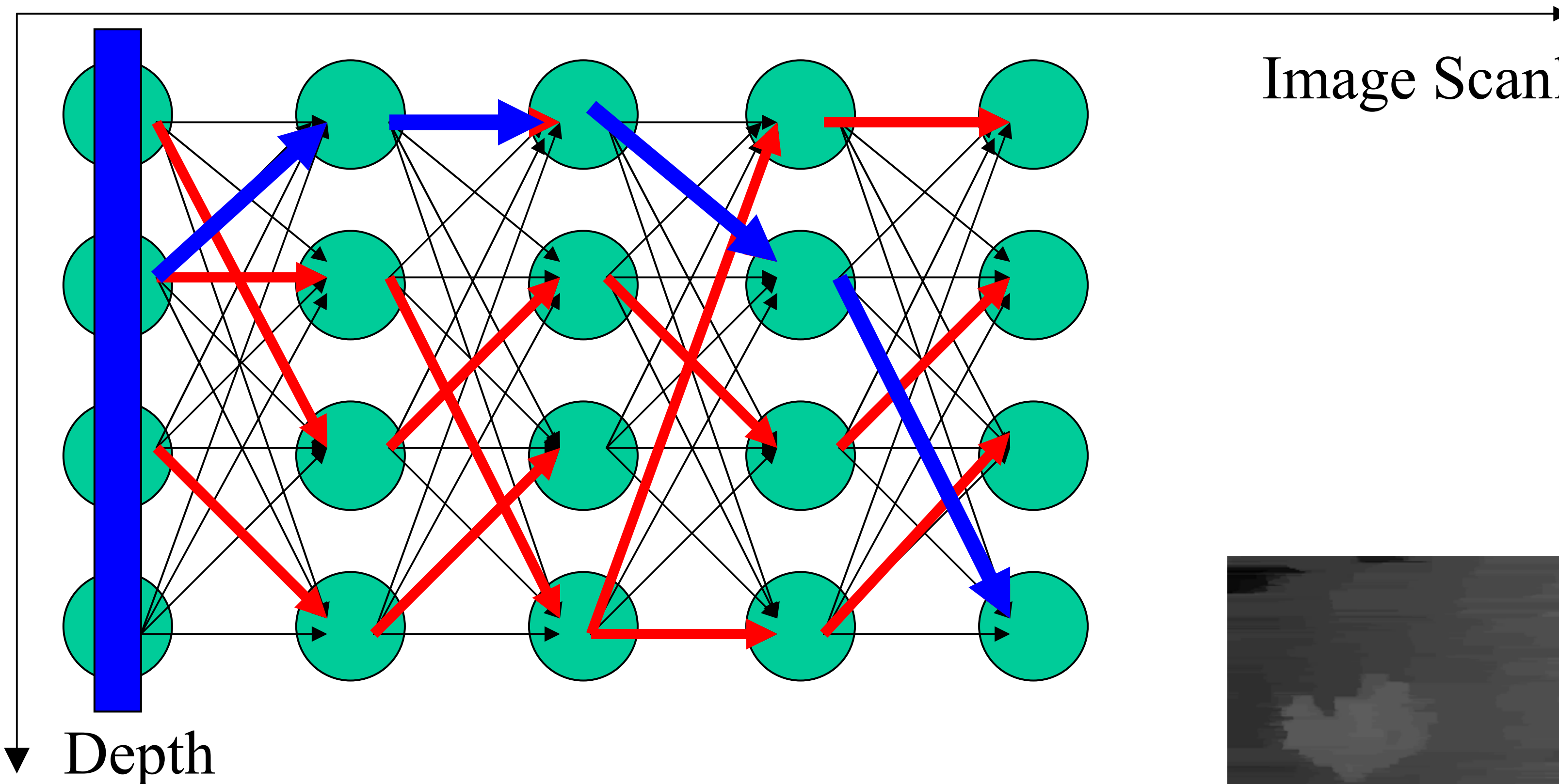
Dynamic Programming

Belief Propagation

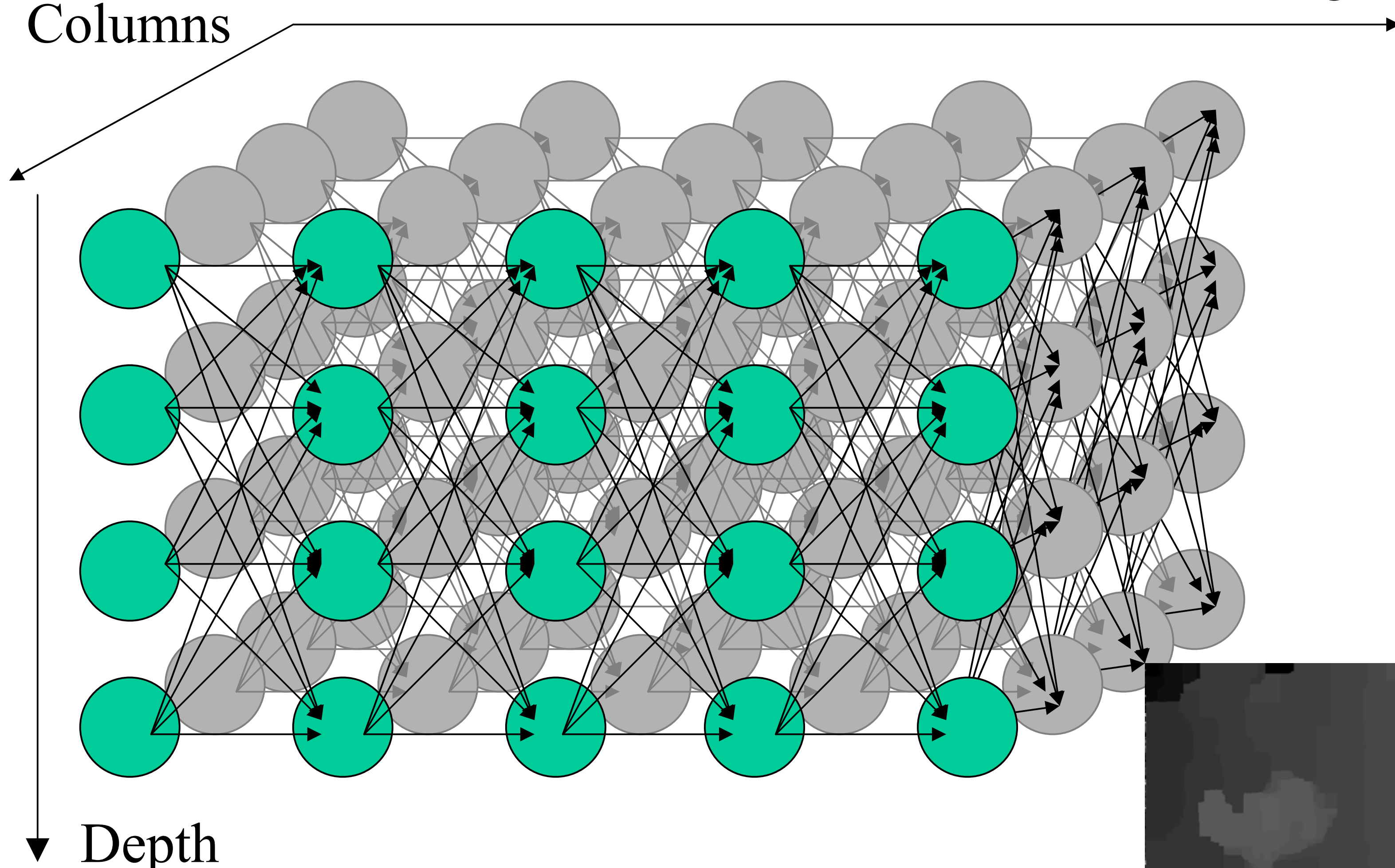


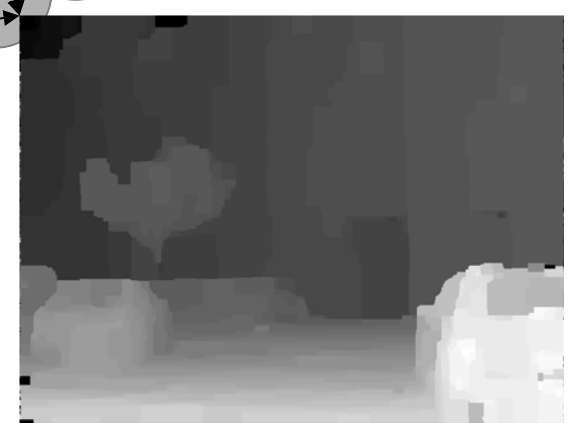
Center for Visualization & Virtual Environments

Dynamic Programming

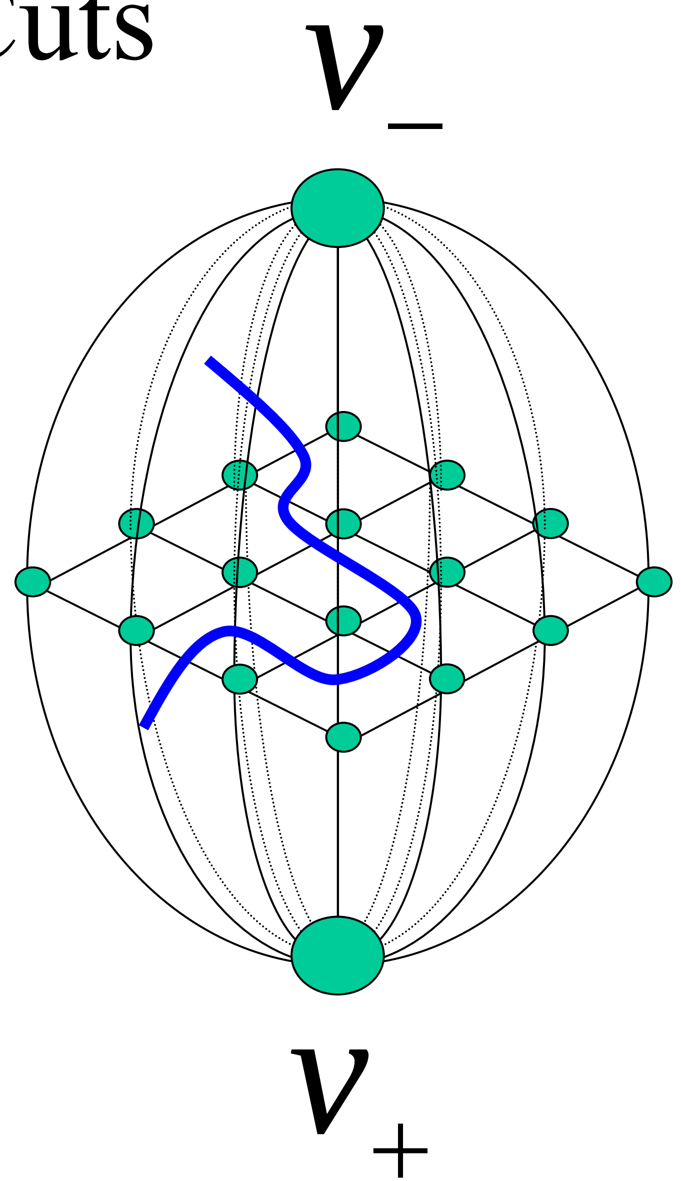


Belief Propagation

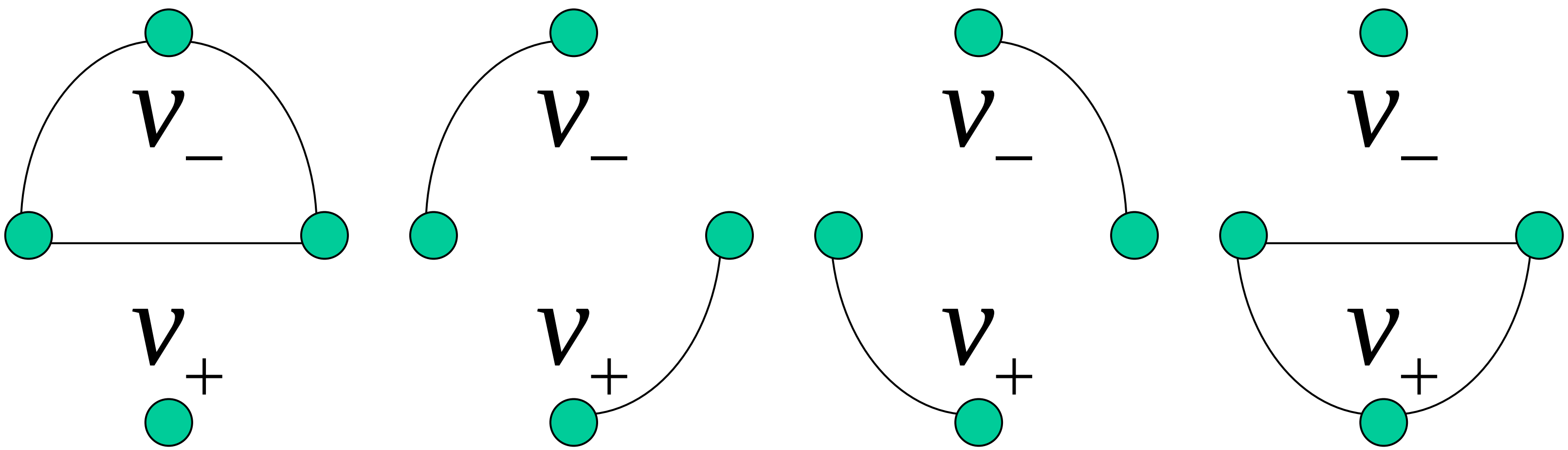




Graph Cuts



Graph Cuts

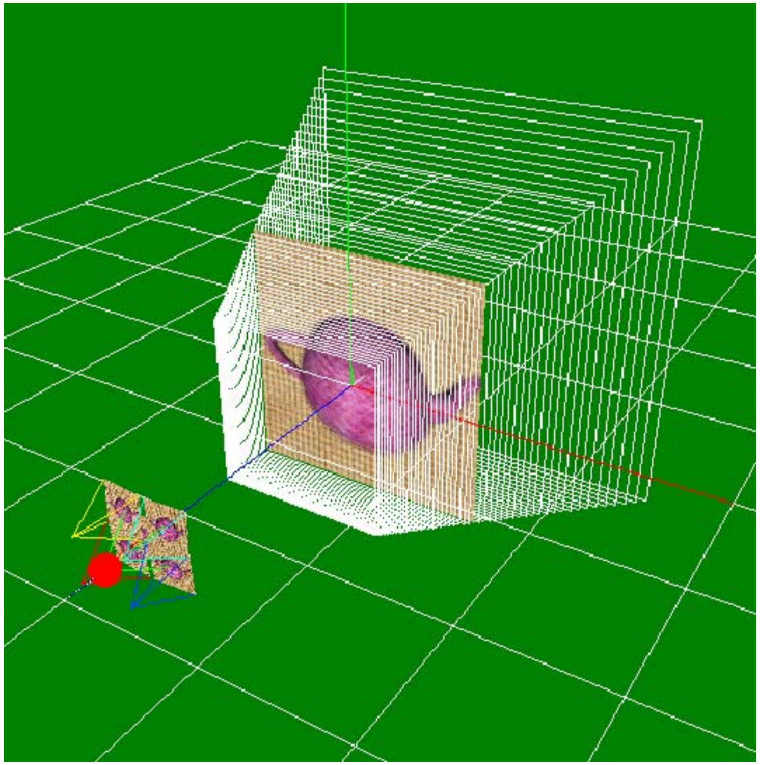


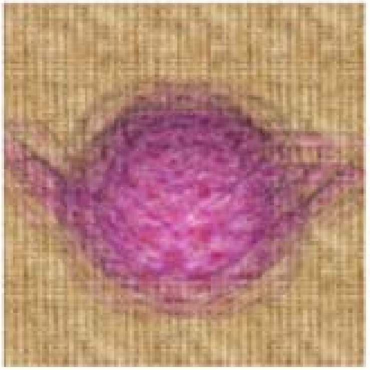
(f, f) (f, g) (g, f) (g, g)

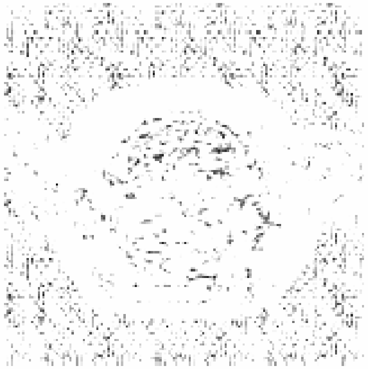
Multi-View Depth Reconstruction

work with Q. Yang, L. Wang, R. Yang

- Plane-sweep stereo on GPU







UK Center for Visualization & Virtual Environments

Middlebury Stereo Record

work with Q. Yang, L. Wang, R. Yang

Double-BP

Highly computationally demanding

even for small images

Color-weighted correlation

Real-time for small images and

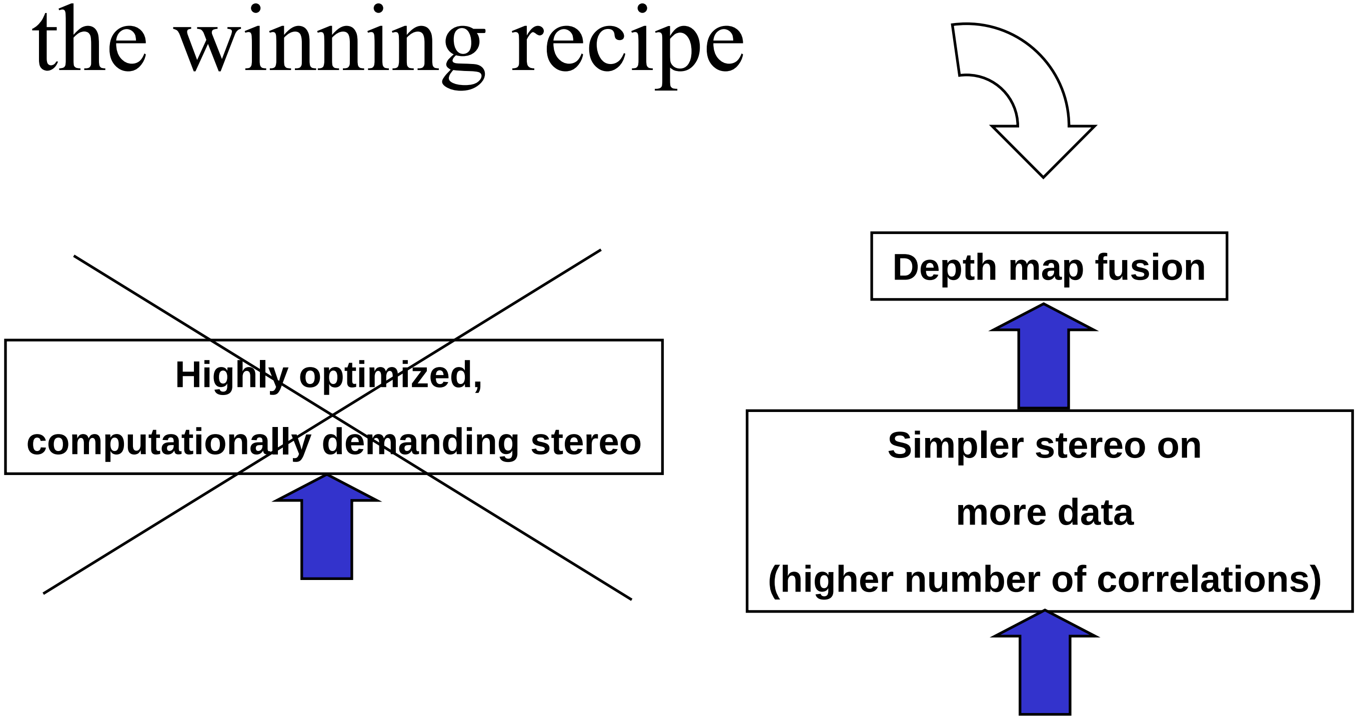
few disparity levels

Error Threshold = 1 | Error Threshold... | Sort by nonocc | Sort by all | Sort by disc

Algorithm	Avg. Rank	Tsukuba (ground truth) nonocc	Tsukuba all	Tsukuba disc	Venus (ground truth) nonocc	Venus all	Venus disc	Teddy (ground truth) nonocc	Teddy all	Teddy disc	Cones (ground truth) nonocc	Cones all	Cones disc
Double-BP [15]	1.3	**0.88** 1	**1.29** 1	**4.76** 1	**0.14** 1	0.60 2	**2.00** 1	**3.55** 1	8.71 2	**9.70** 1	**2.90** 1	9.24 2	**7.80** 1
Segm+visib [4]	3.3	1.30 5	1.57 2	6.92 6	0.79 4	1.06 3	6.76 6	5.00 2	**6.54** 1	12.3 2	3.72 3	**8.62** 1	10.2 4
SymBP+occ [7]	3.4	0.97 2	1.75 3	5.09 2	0.16 2	**0.33** 1	2.19 2	6.47 4	10.7 3	17.0 4	4.79 7	10.7 6	10.9 5
AdaptWeight [12]	4.7	1.38 7	1.85 4	6.90 5	0.71 3	1.19 4	6.13 4	7.88 6	13.3 5	18.6 7	3.97 5	9.79 4	8.26 2
SemiGlob [6]	6.3	3.26 12	3.96 10	12.8 15	1.00 5	1.57 5	11.3 10	6.02 3	12.2 4	16.3 3	3.06 2	9.75 3	8.90 3
Layered [5]	7.8	1.57 8	1.87 5	8.28 8	1.34 7	1.85 6	6.85 7	8.64 8	14.3 6	18.5 6	6.59 11	14.7 11	14.4 10
GC+occ [2]	7.9	1.19 3	2.01 7	6.24 3	1.64 10	2.19 9	6.75 5	11.2 11	17.4 11	19.8 9	5.36 9	12.4 9	13.0 9
MultiCamGC [3]	8.4	1.27 4	1.99 6	6.48 4	2.79 14	3.13 12	3.60 3	12.0 12	17.6 12	22.0 11	4.89 8	11.8 8	12.1 7
TensorVoting [9]	9.3	3.79 13	4.79 13	8.86 9	1.23 6	1.88 7	11.5 11	9.76 9	17.0 10	24.0 13	4.38 6	11.4 7	12.2 8
CostRelax [11]	10.1	4.76 15	6.08 15	20.3 16	1.41 9	2.48 10	18.5 14	8.18 7	15.9 8	23.8 12	3.91 4	10.2 5	11.8 6
RealTime-GPU [14]	10.2	2.05 11	4.22 12	10.6 12	1.92 12	2.98 11	20.3 15	7.23 5	14.4 7	17.6 5	6.41 10	13.7 10	16.5 12
Reliablty-DP [13]	11.4	1.36 6	3.39 9	7.25 7	2.35 13	3.48 14	12.2 13	9.82 10	16.9 9	19.5 8	12.9 17	19.9 17	19.7 14
TreeDP [8]	11.7	1.99 10	2.84 8	9.96 11	1.41 8	2.10 8	7.74 8	15.9 15	23.9 15	27.1 16	10.0 14	18.3 14	18.9 13

Depth Map Fusion

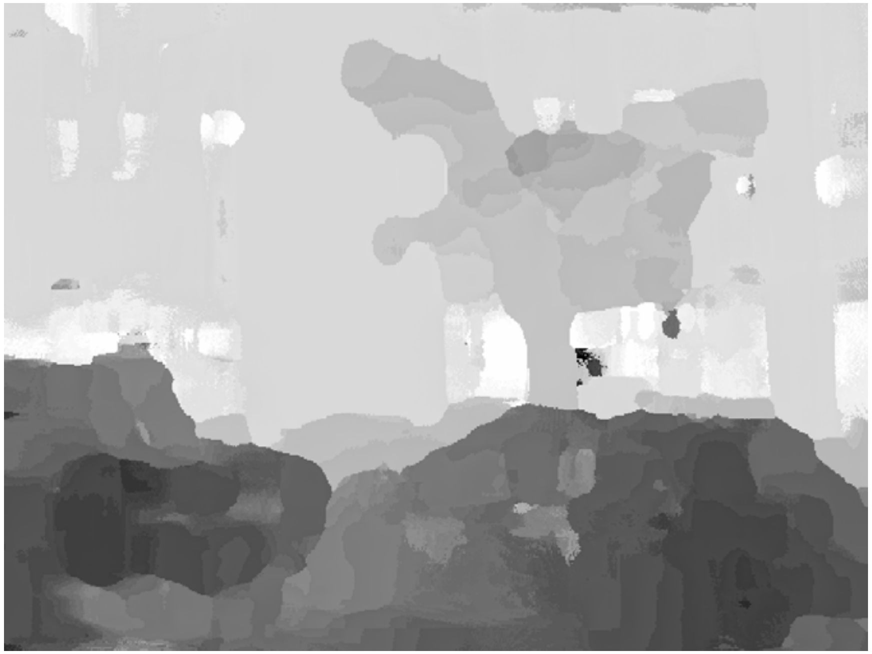
- Main lesson: simple stereo with many correlations on many images + fusion is the winning recipe



GPU Stereo

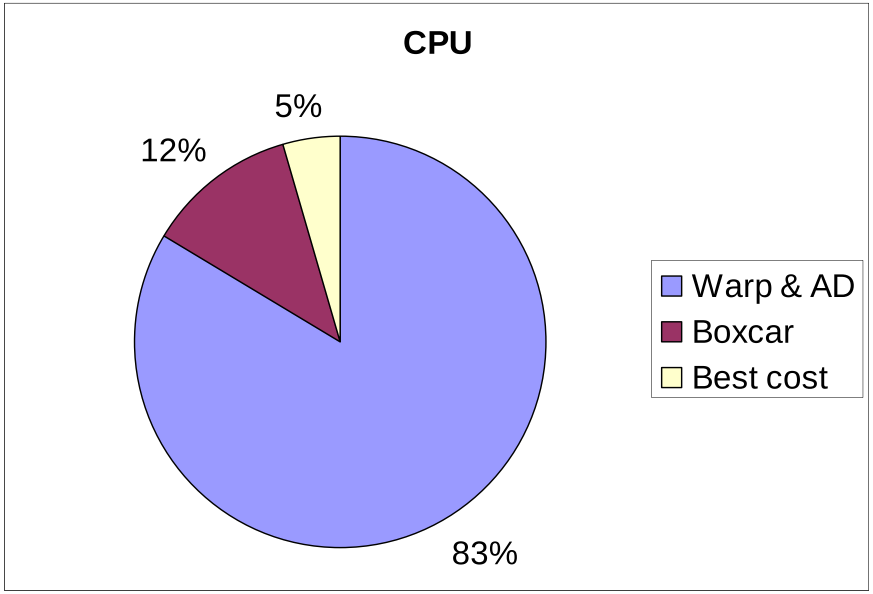
CPU

GPU

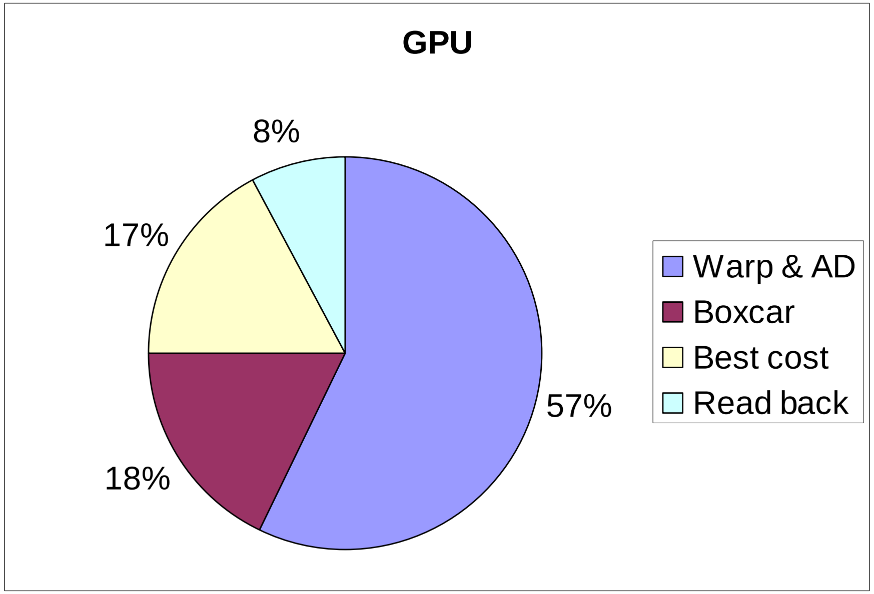


GPU Stereo

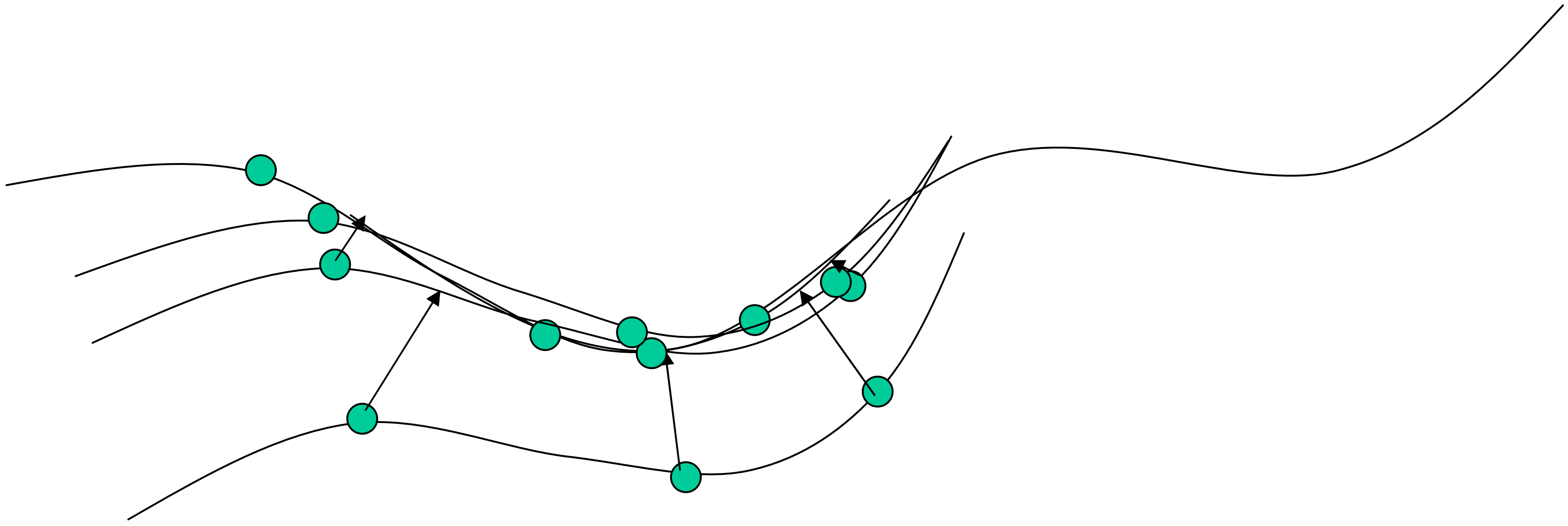
CPU (Xeon 3GHz): 3.2s



GPU (NVIDIA 7800 GTX): 70ms



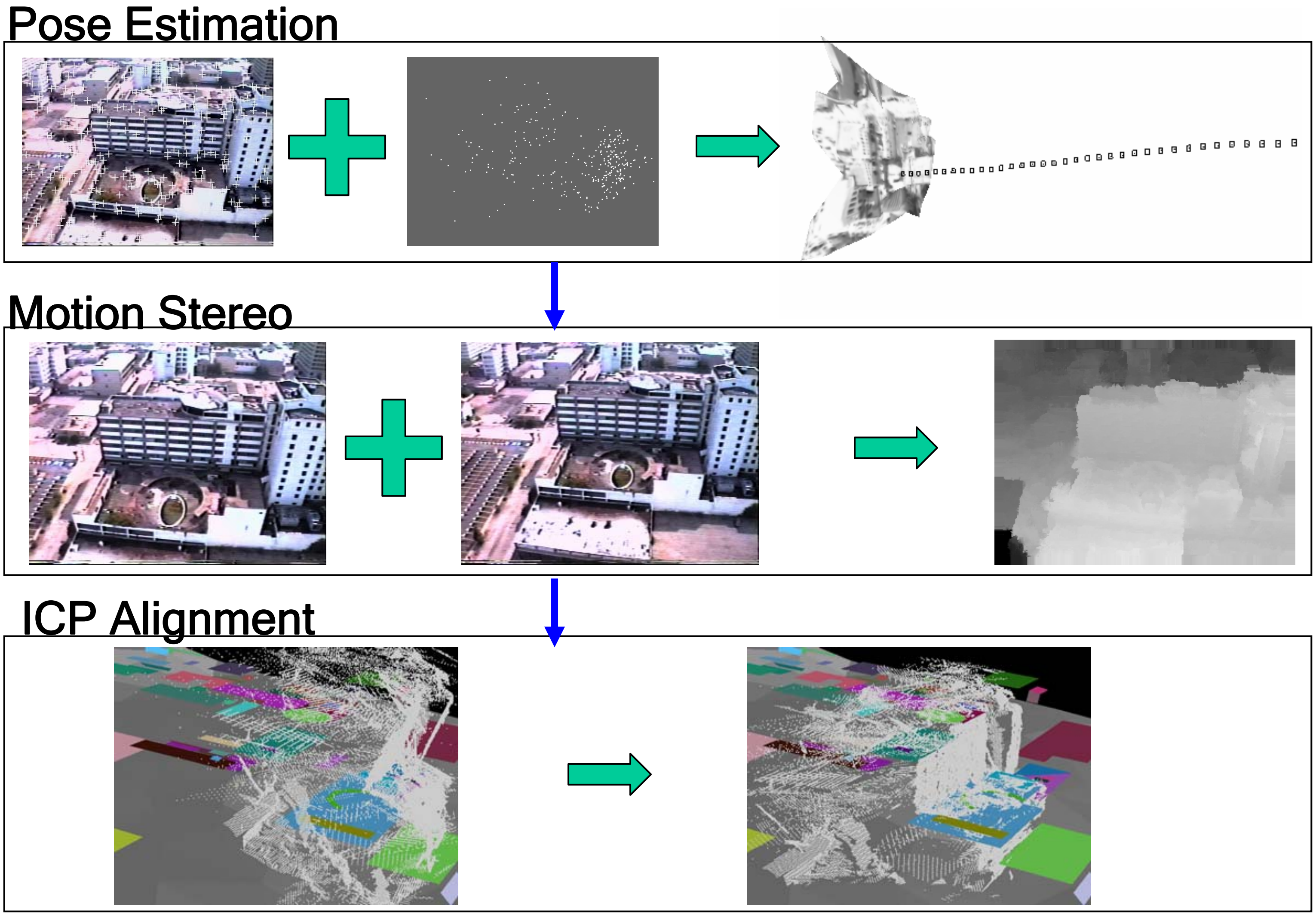
ICP



Center for Visualization & Virtual Environments

Alignment of Video onto 3D Point Clouds

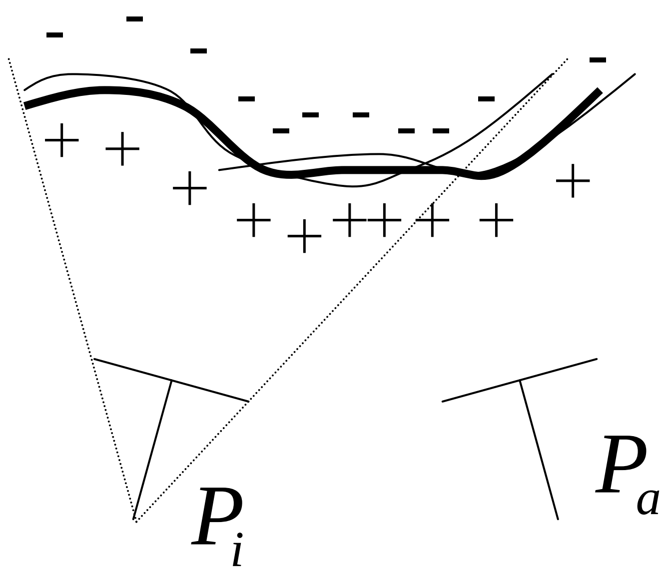
work with Wen-Yi Zhao and Steve Hsu



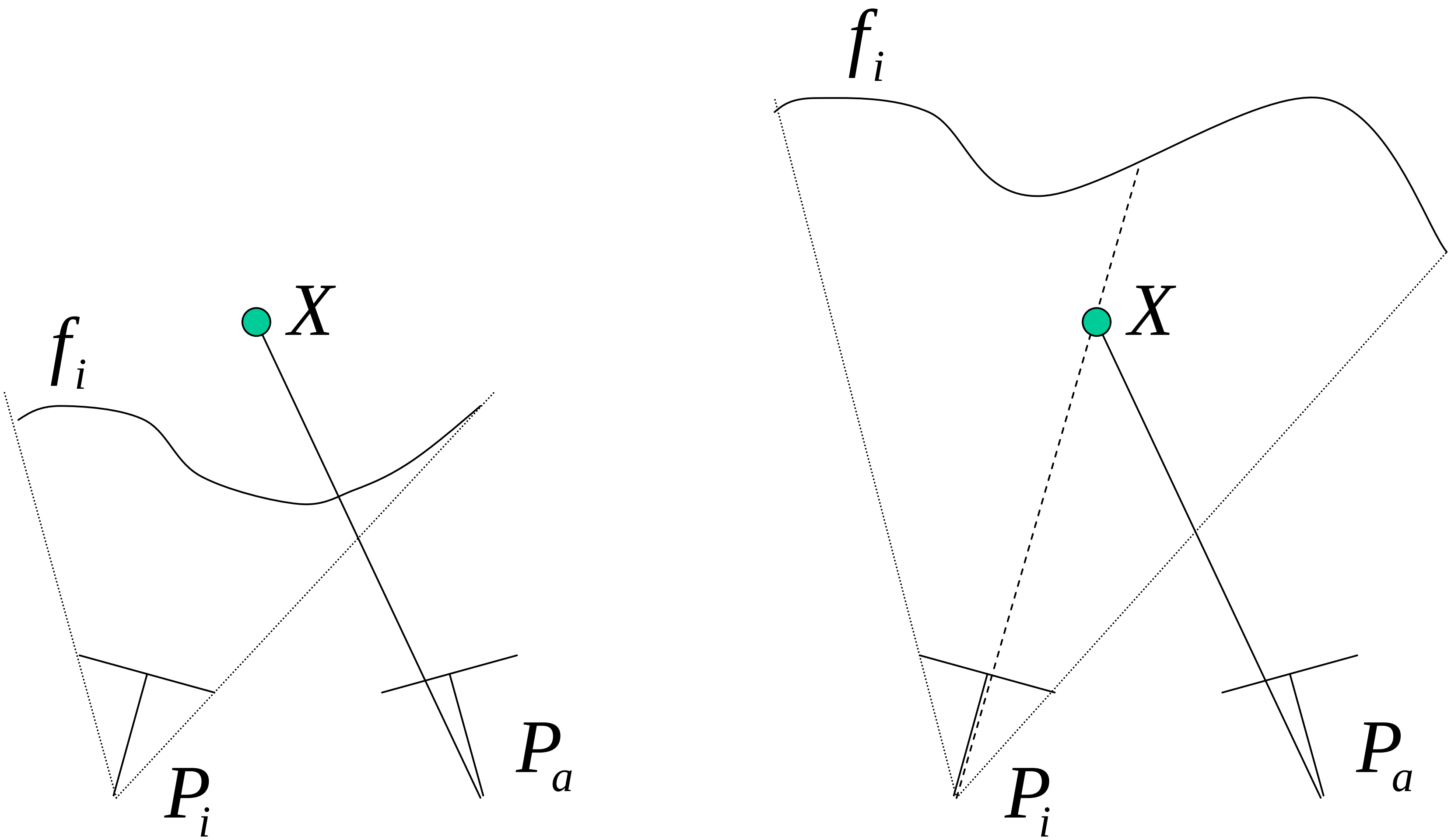
UK Center for Visualization & Virtual Environments

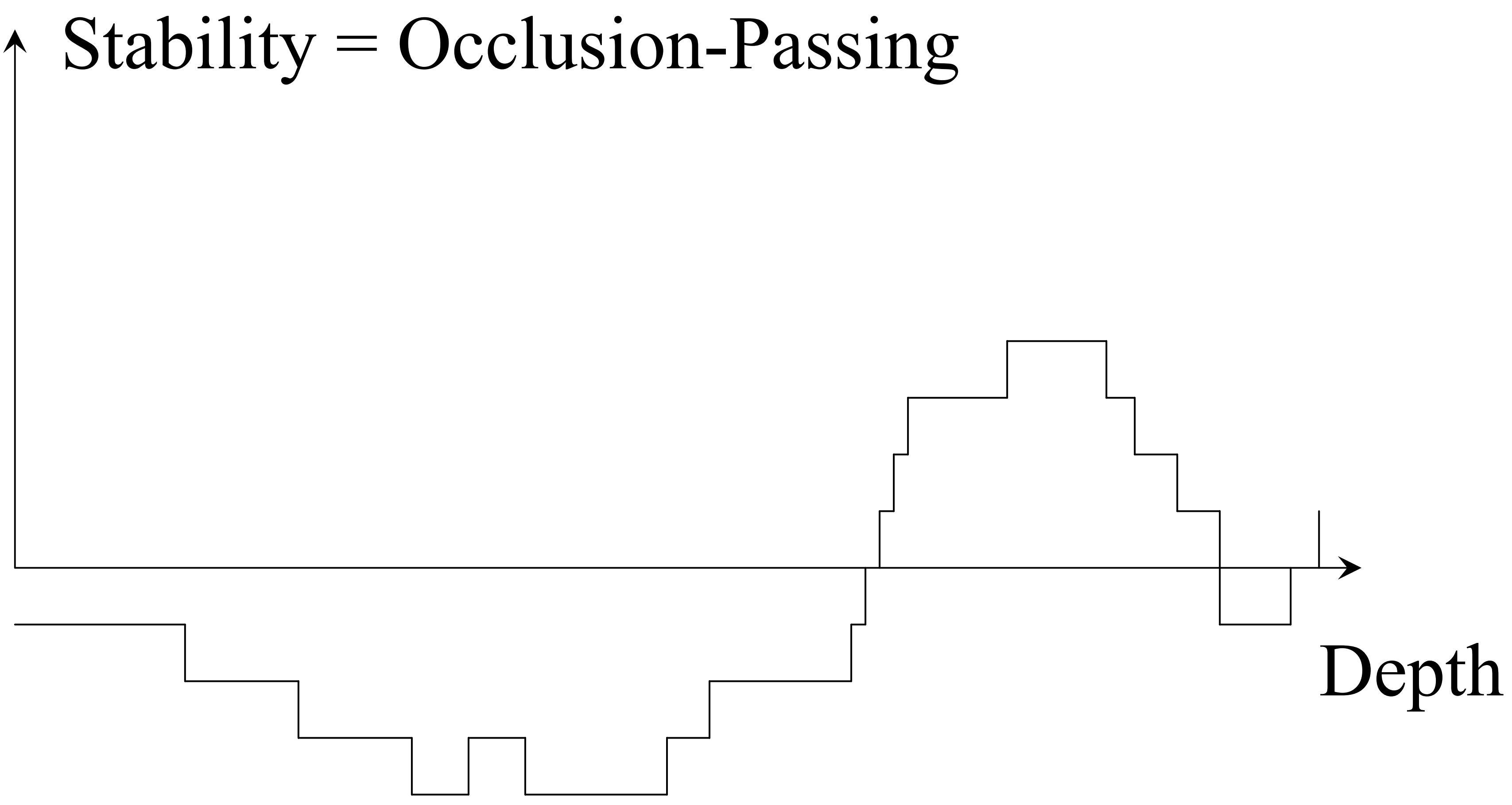
Fusion

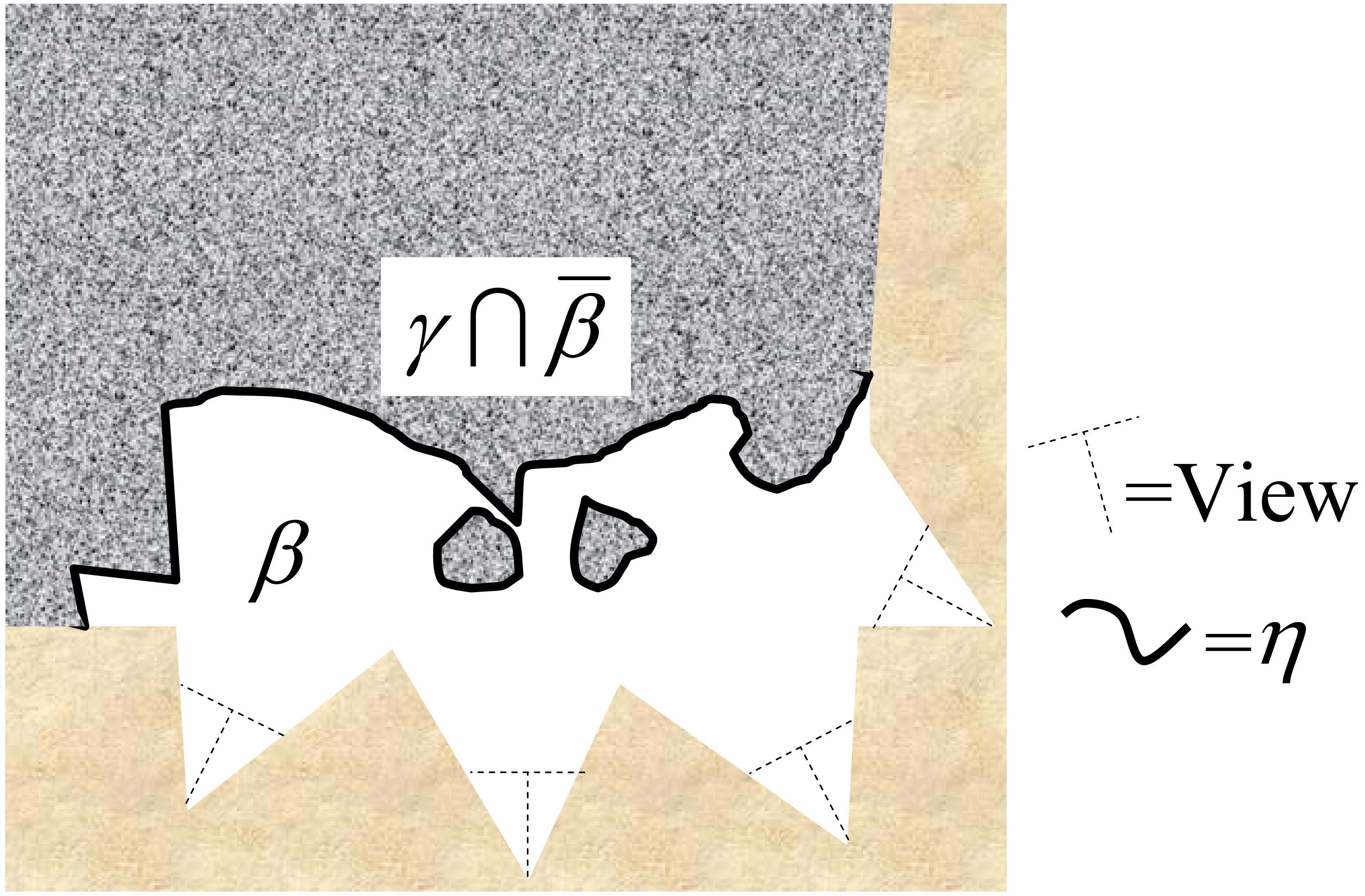
- Curless & Levoy



Median Fusion



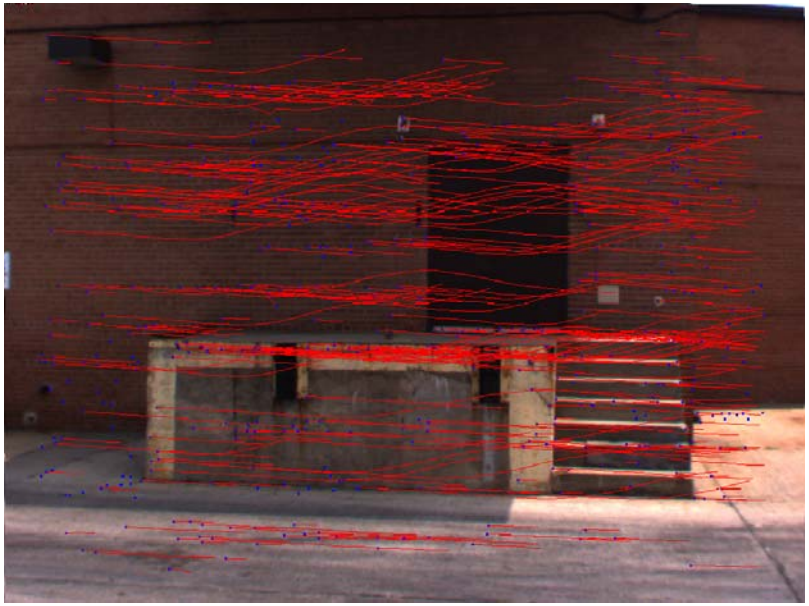

Stability = Occlusion-Passing
Depth

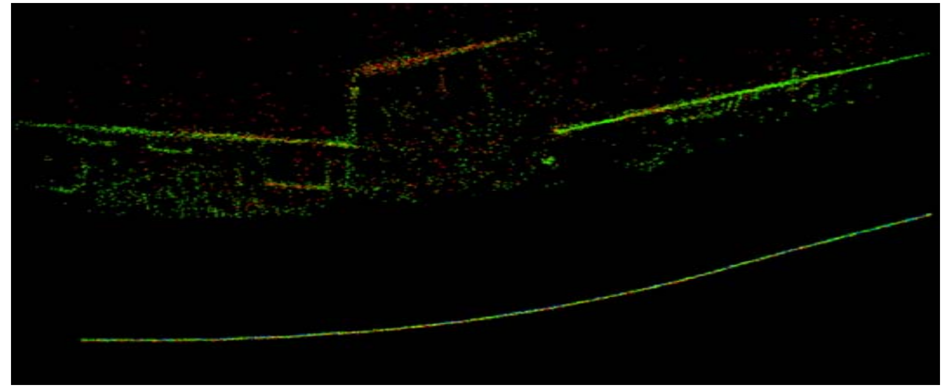


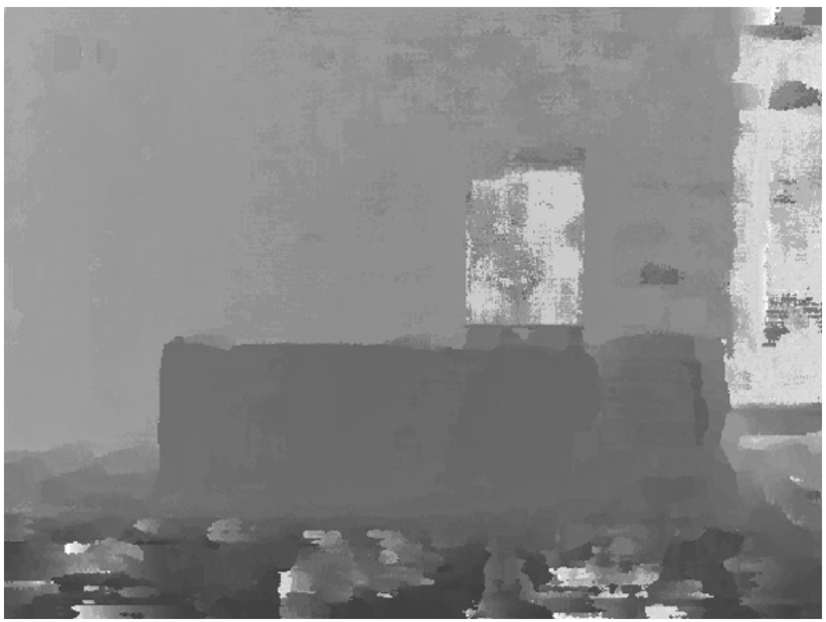
Center for Visualization & Virtual Environments

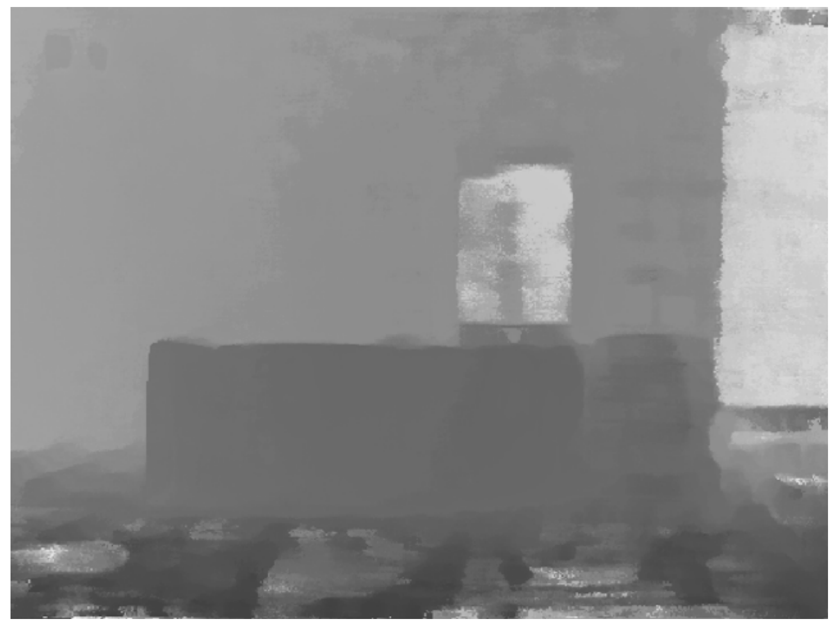
Depth Map Fusion

- Resolves inconsistencies. Cleans up results very efficiently
- Suited for GPU implementation (essentially consists of rendering back and forth many times)



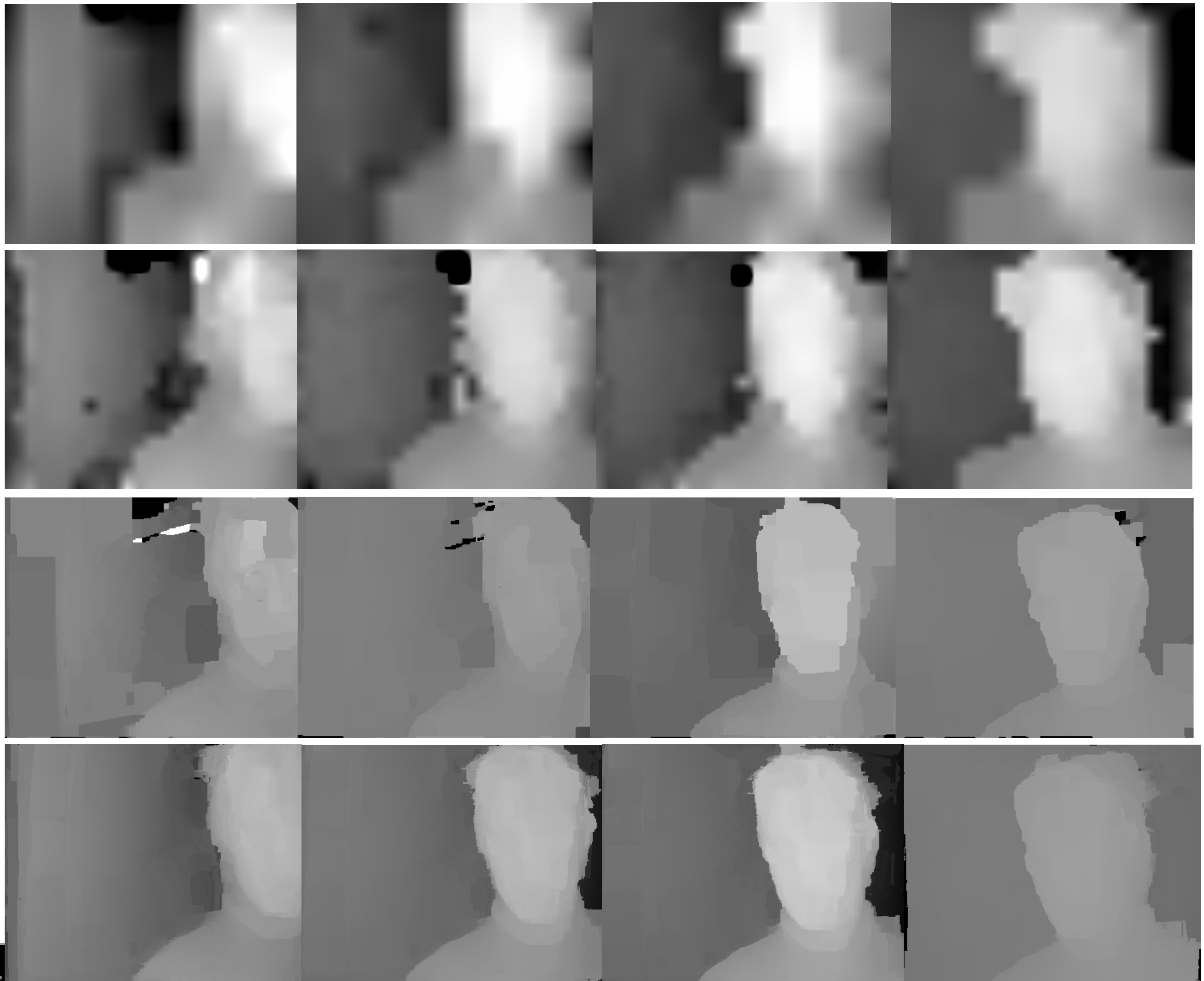


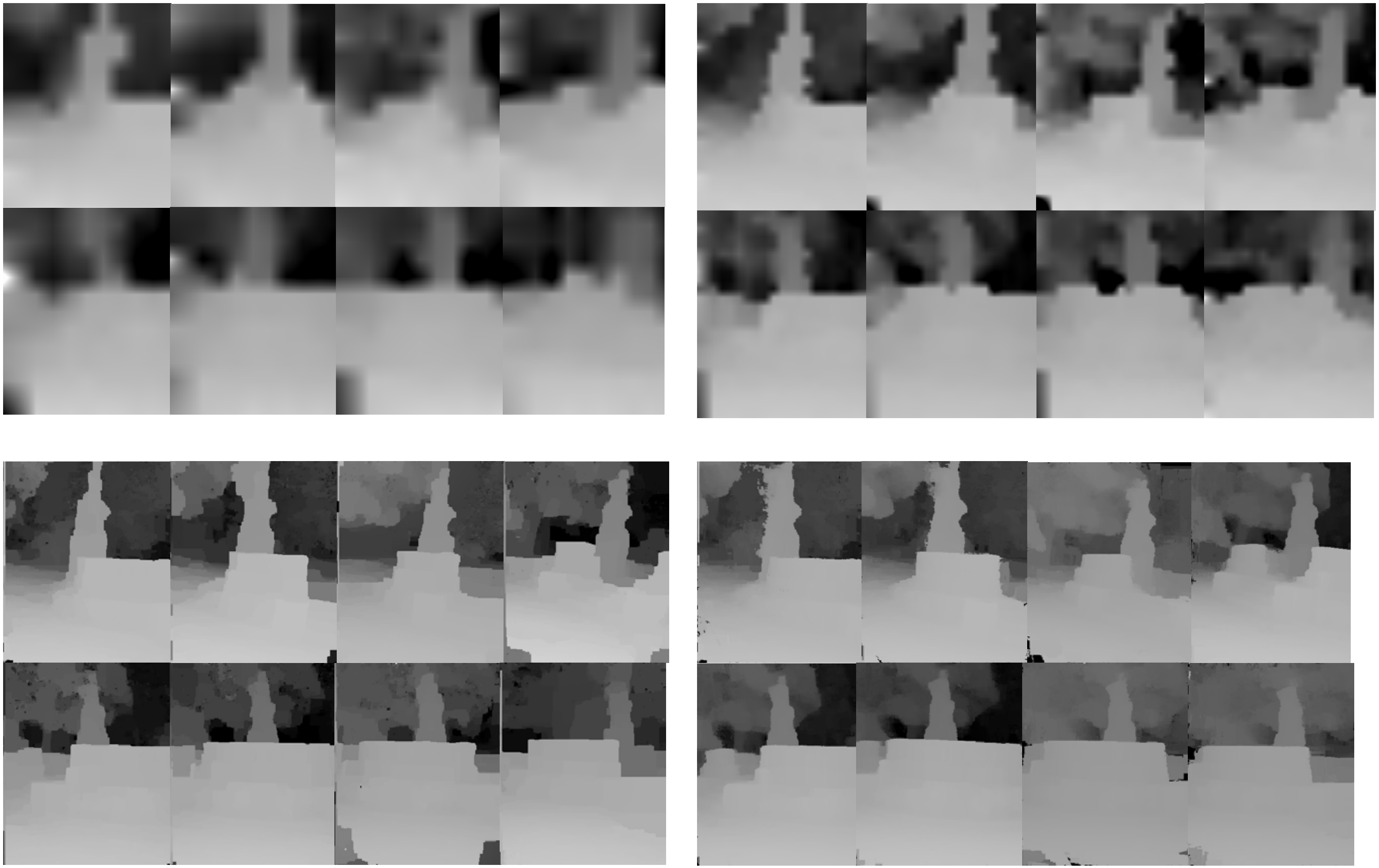




UK

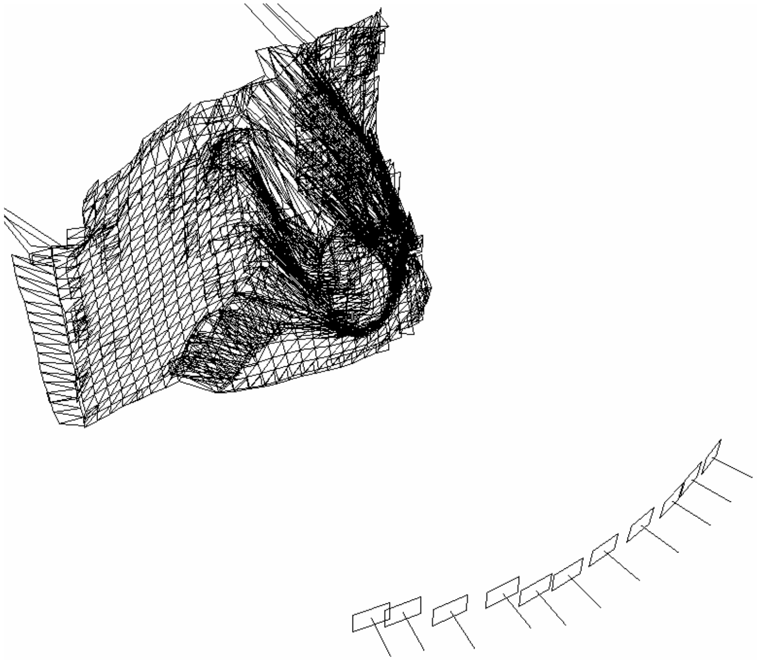
Center for
Visualization & Virtual
Environments

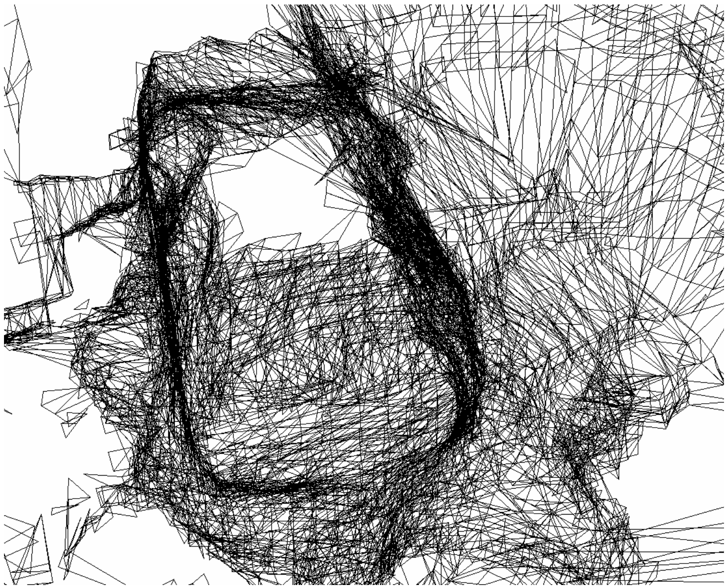


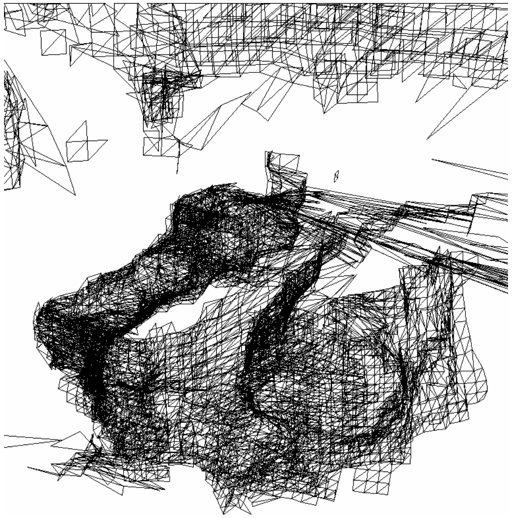


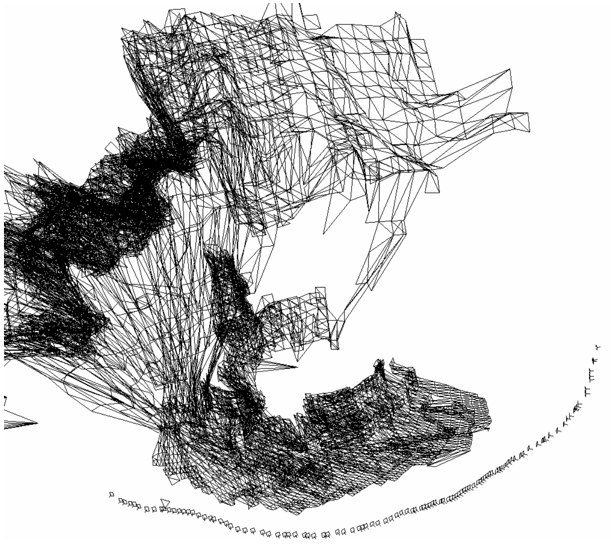
UK

Center for Visualization & Virtual Environments









UK

Center for Visualization & Virtual Environments

Depth Map Fusion

Sparse Mesh Generation

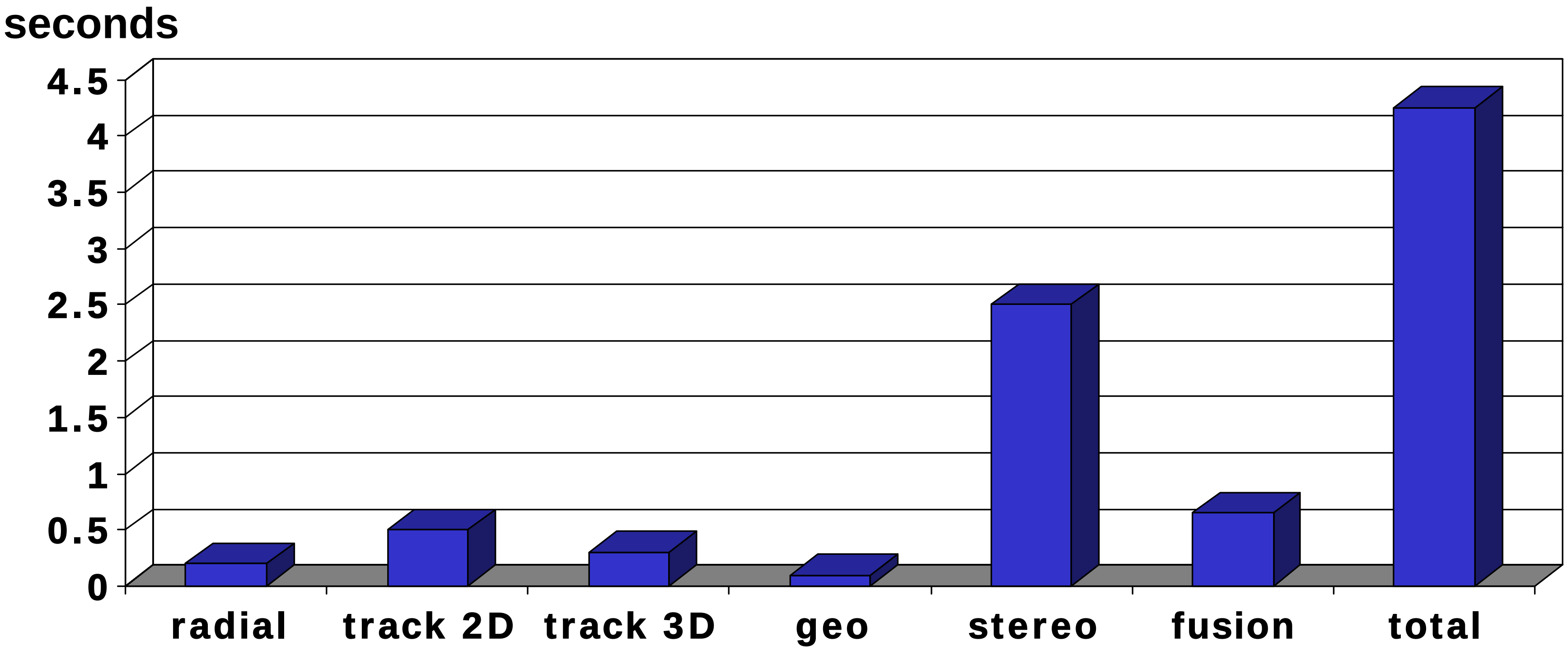
UK

Center for Visualization & Virtual Environments

Computation times CPU

Single CPU processing times for single video stream

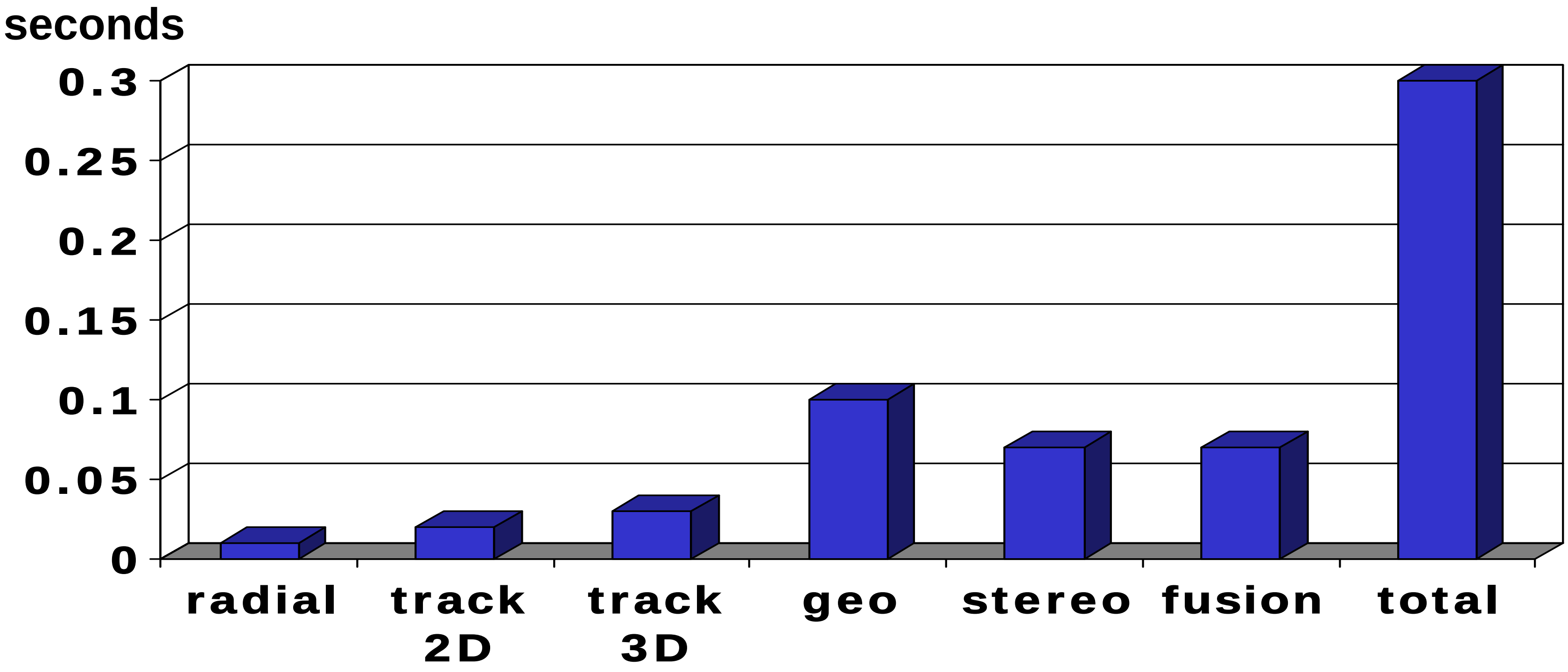
Running the whole system with:
1024x768 resolution for Radial, Tracker 2D, Tracker 3D, Geo registration
512x384 resolution for Stereo, Fusion, 3D model generation



Computation times CPU+GPU

Single CPU + GPU processing times for single video stream

Running the whole system with:
1024x768 resolution for Radial, Tracker 2D, Tracker 3D, Geo registration
512x384 resolution for Stereo, Fusion, 3D model generation



Center for Visualization & Virtual Environments

UK Center for Visualization & Virtual Environments

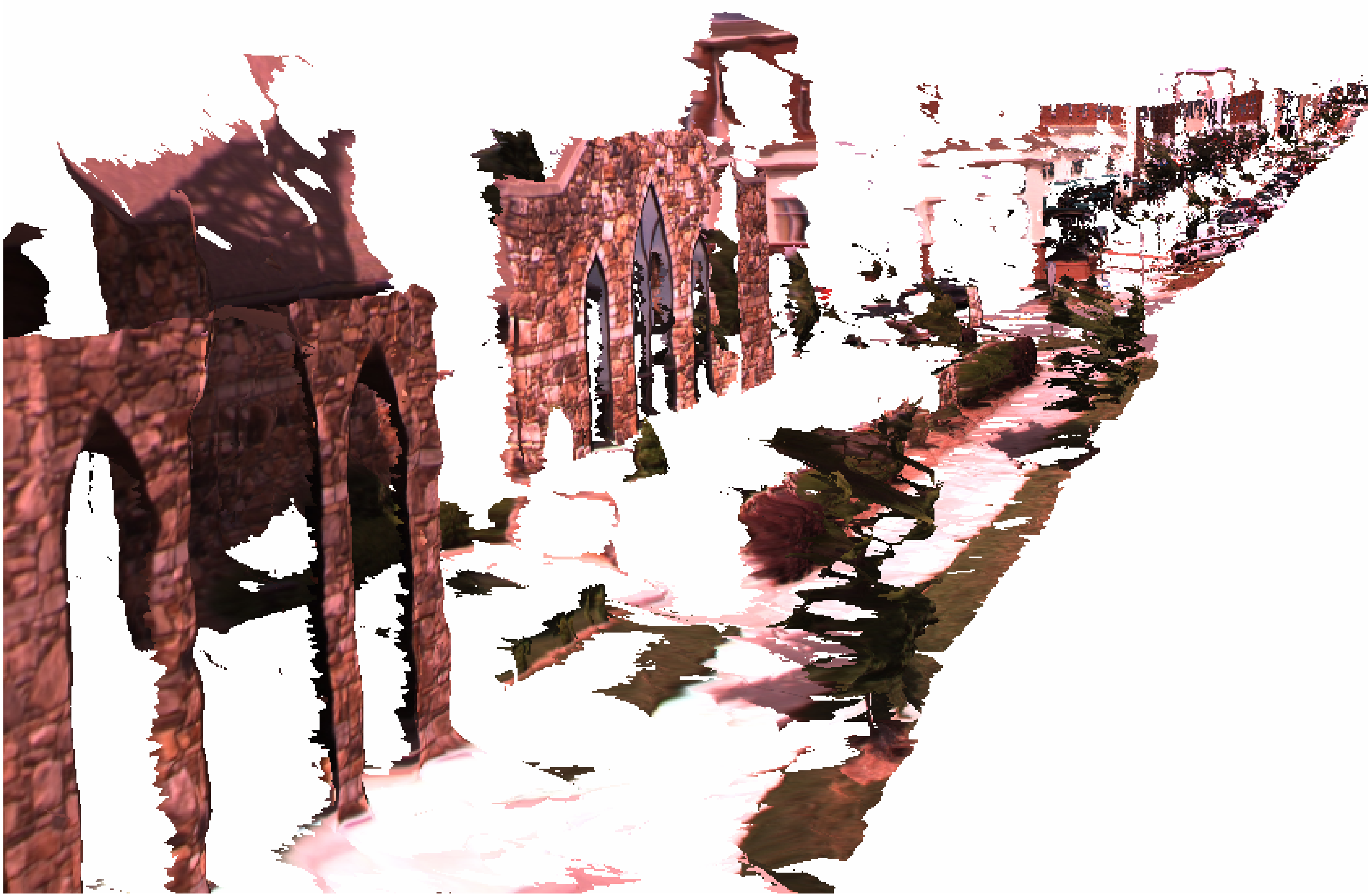
UK

Center for Visualization & Virtual Environments

UK Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments



UK

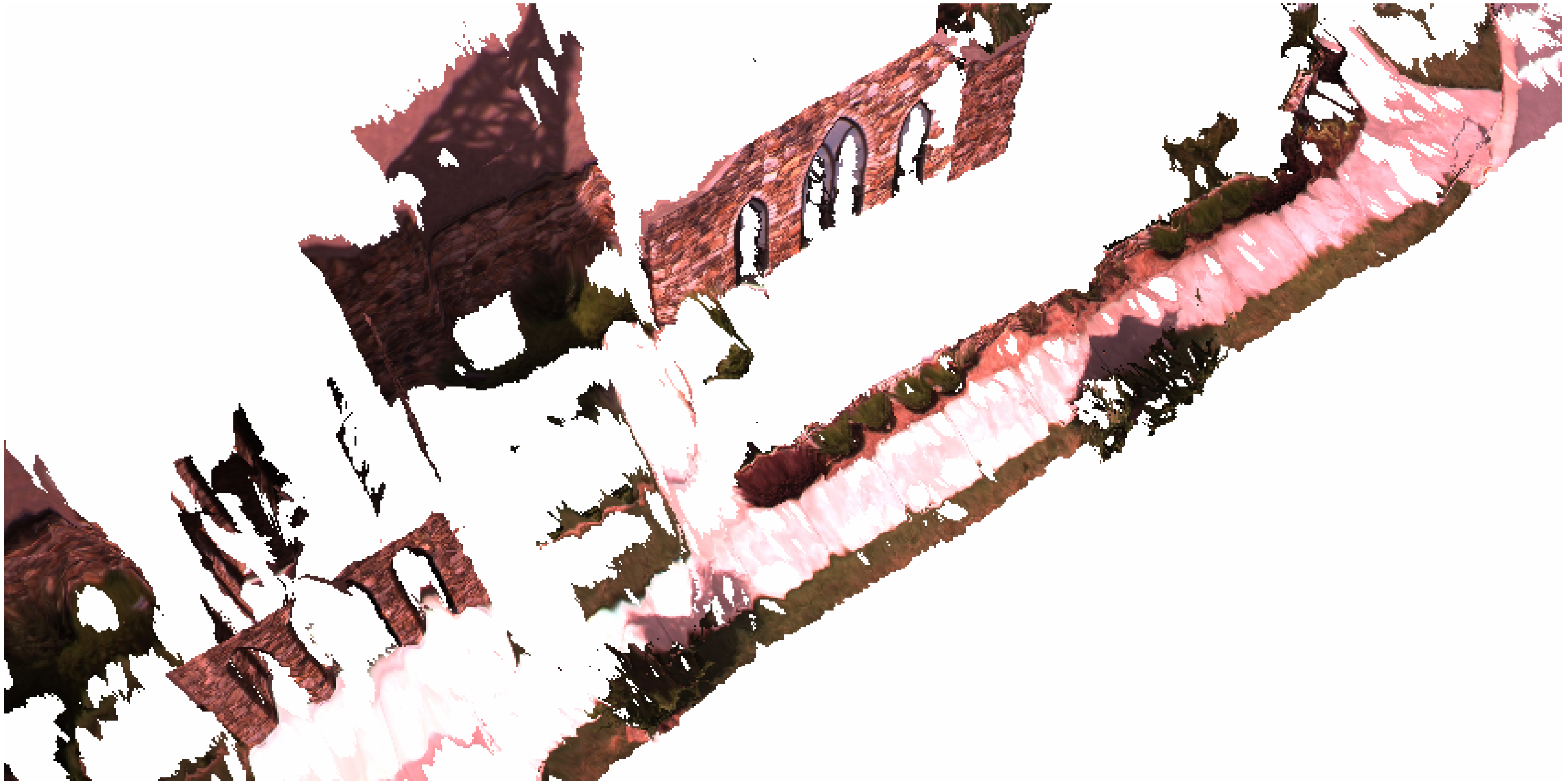
Center for Visualization & Virtual Environments

UK

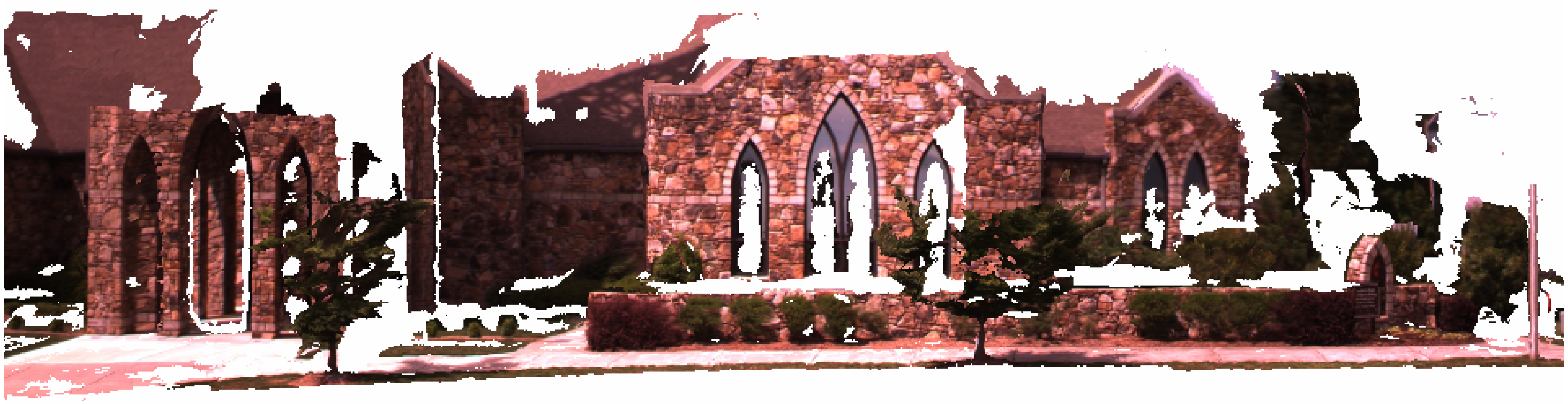
Center for Visualization & Virtual Environments

UK

Center for Visualization & Virtual Environments



Center for Visualization & Virtual Environments



UK

Center for Visualization & Virtual Environments

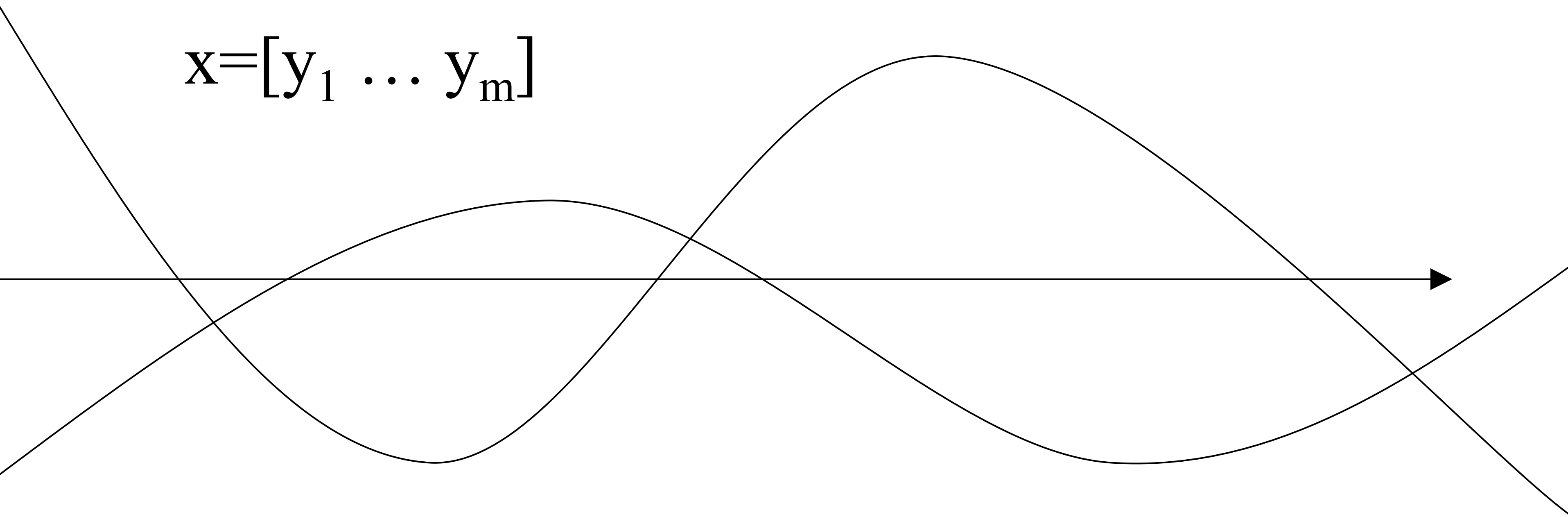
Camera Geometry

- Often leads to polynomial formulations, or can at least very often be formulated in terms of polynomial equations

Polynomial Formulation

- $p_1(x)$, … , $p_n(x)$= A set of input polynomials
  (n polynomials in m variables)

  $x=[y_1 \dots y_m]$



Center for Visualization & Virtual Environments

Algebraic Ideal

- $I(p_1 , \ldots , p_n)$= The set of polynomials generated by the input polynomials

(through additions and multiplications by a polynomial)

p and q in I => p+q in I

p in I => pq in I

The ideal I consists of ‘Almost’ all the polynomials implied by the input polynomials
(More precisely, the radical $\sqrt{I}$ of the ideal consists of all)

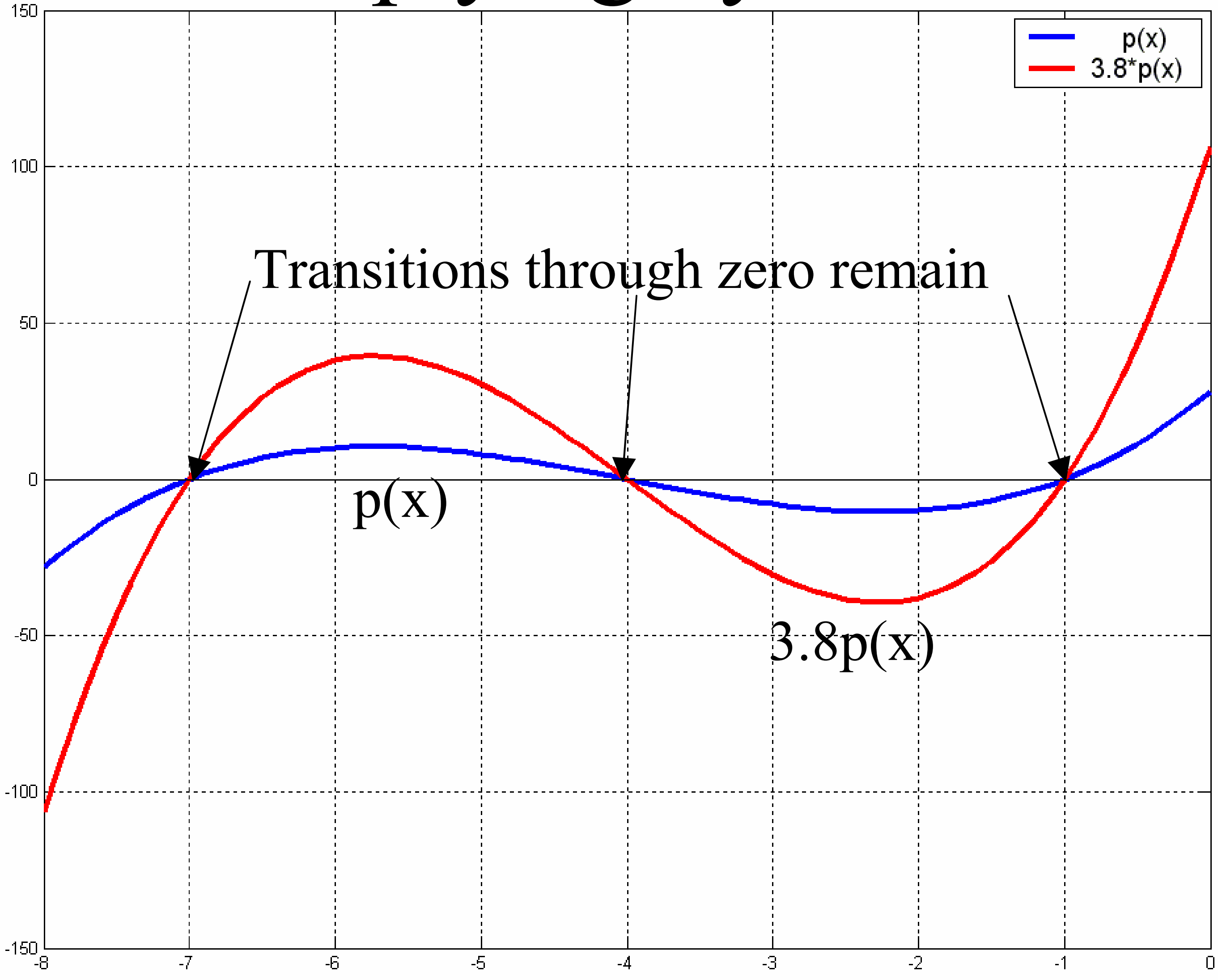
Remember Row Operations:

- Multiplying a row by a scalar
- Subtracting a row from another
- Swap rows

Add:

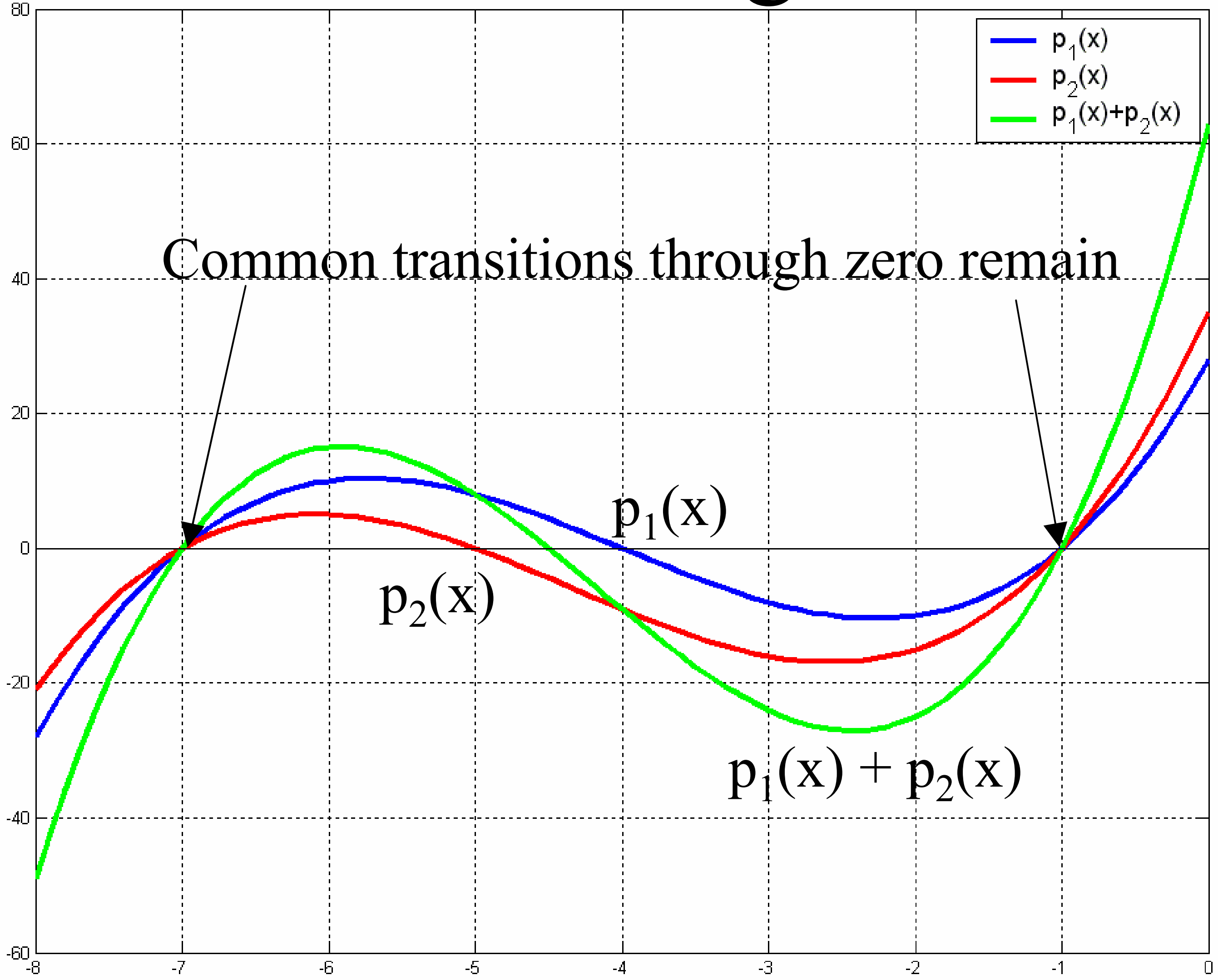
- Multiplying a row by any polynomial

Multiplying by a Scalar



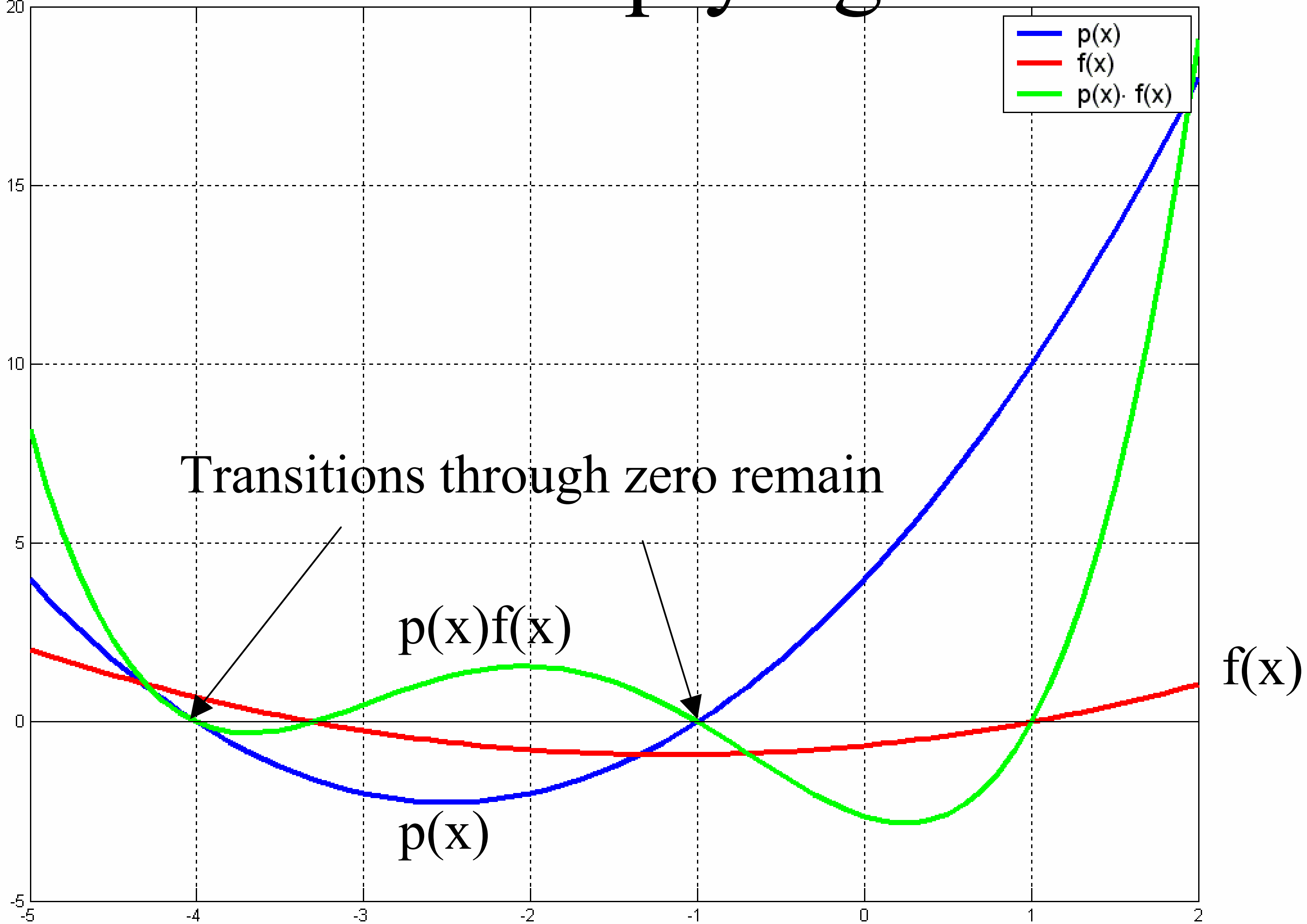
Center for Visualization & Virtual Environments

Adding



p1(x)
p2(x)
p1(x)+p2(x)
Common transitions through zero remain
p1(x)
p2(x)
p1(x) + p2(x)
80
60
40
20
0
-20
-40
-60
-8
-7
-6
-5
-4
-3
-2
-1
0

Multiplying



Center for Visualization & Virtual Environments

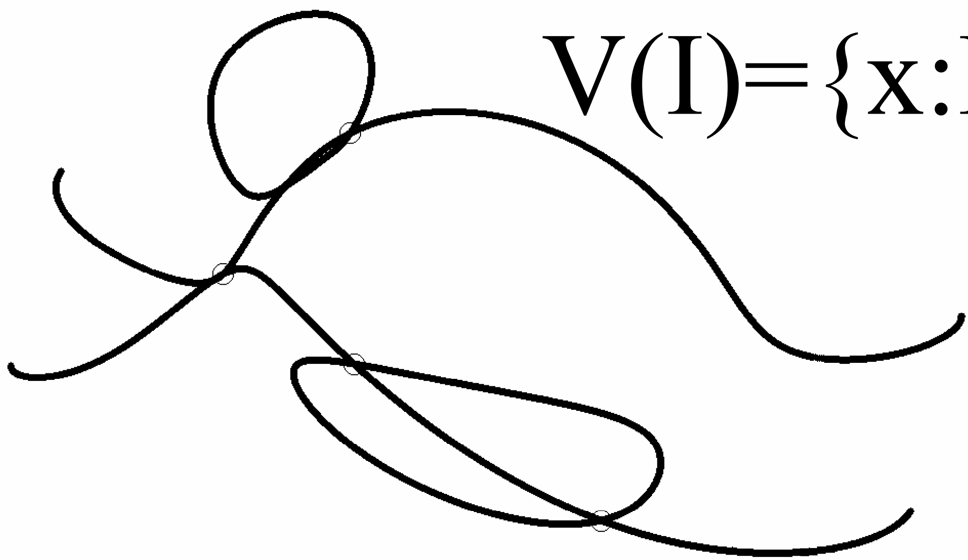
Basis (for Ideal)

- A basis for I is a set of polynomials $(p_1, \ldots, p_n)$ such that $I = I(p_1, \ldots, p_n)$

Algebraic Variety

- The solution set

(the vanishing set of the input polynomials)

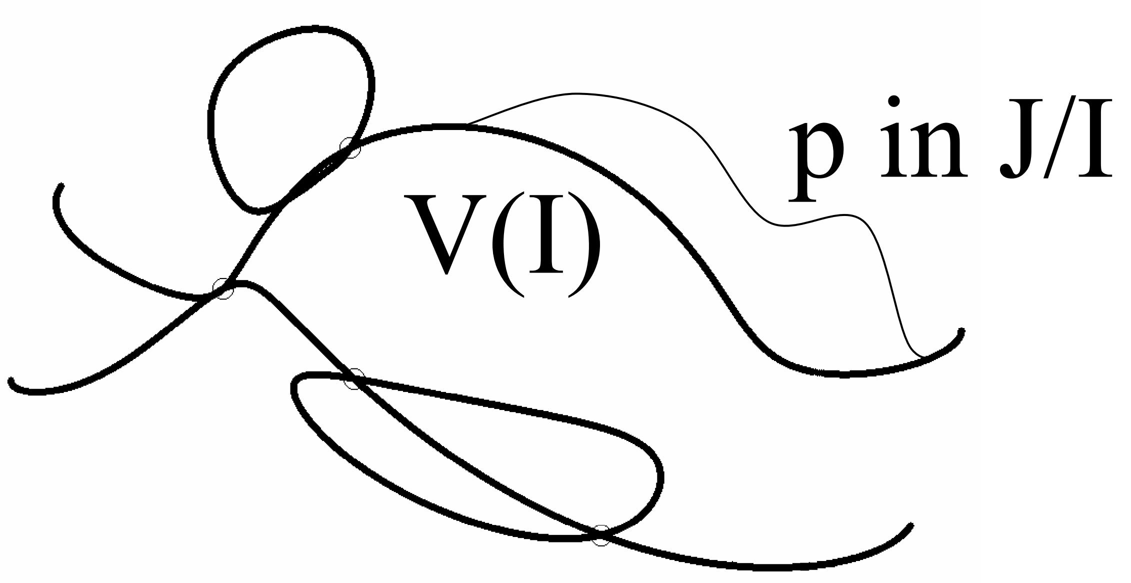


$$V(I)=\{x:I(x)=0\}$$

More precisely
p(x)=0 for all p in I

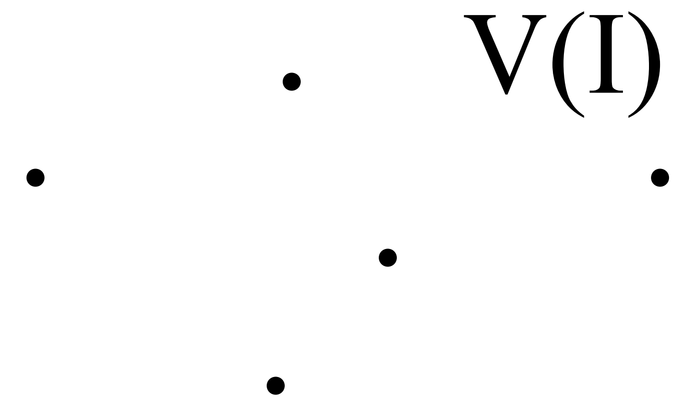
Quotient Ring J/I

- The set of equivalence classes of polynomials when only the values on V are considered (i.e. polynomials are equivalent iff $p(x)=q(x)$ for all x in V)

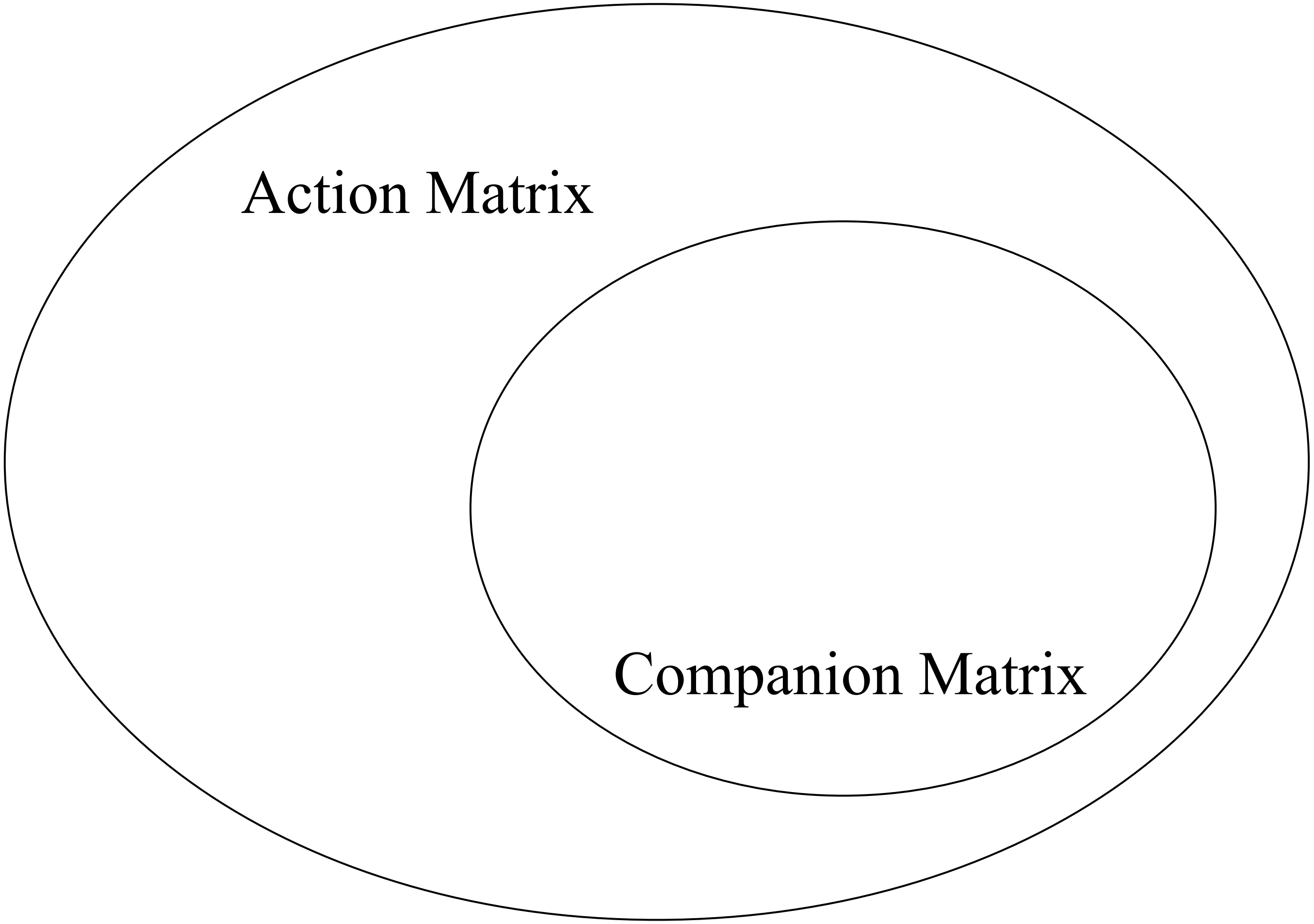


Action Matrix

- For multiplication by polynomial on finite dimensional solution space



Action Matrix



An ‘Equivalence’

Compute Companion Matrix

Finding the Roots of a Polynomial

Finding the Eigenvalues of a Matrix

Compute Characteristic Polynomial

Requires Gröbner Basis for Input Equations

Compute Action Matrix in Quotient Ring (Polynomials modulo Input Equations)

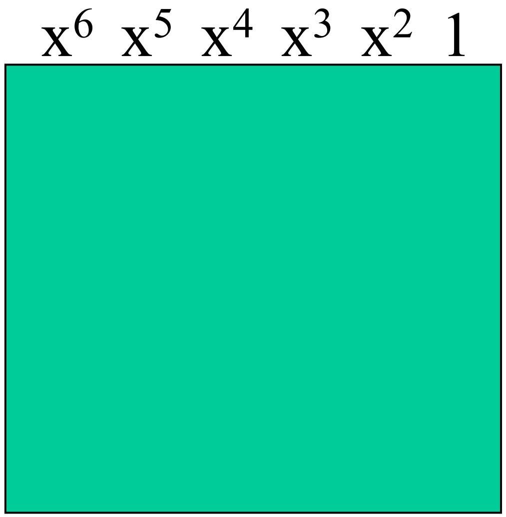
Finding the Roots of Multiple Polynomial Equations

Finding the Eigenvalues of a Matrix

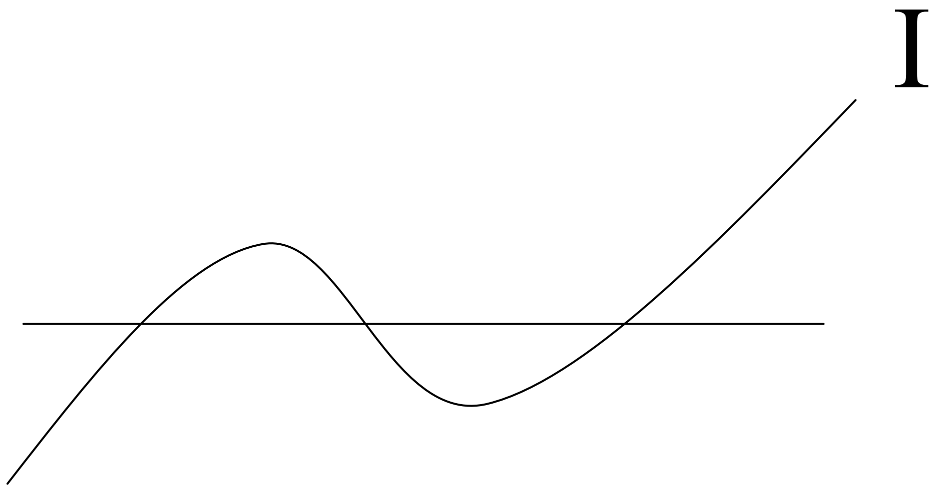
Compute Characteristic Polynomial

Companion Matrix

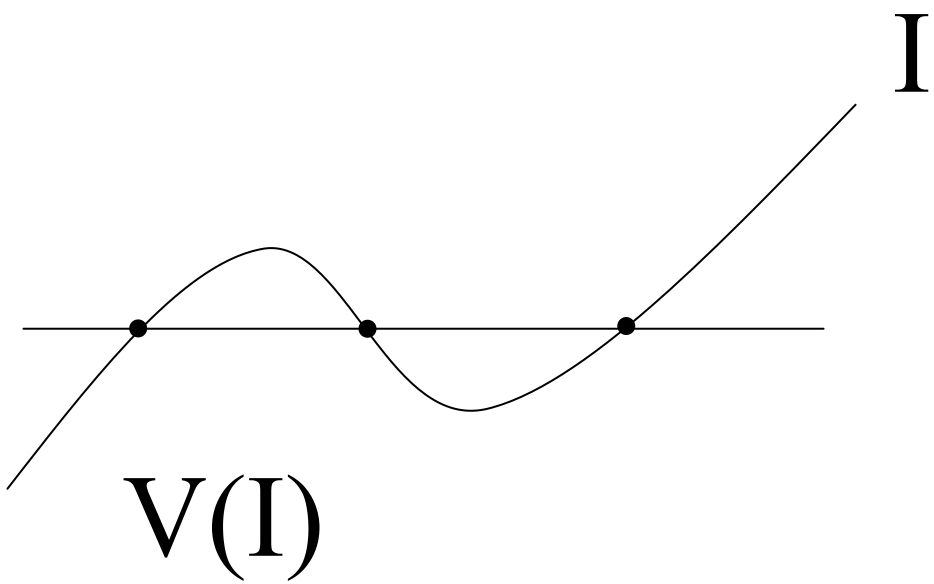
$$a_7x^7 + a_6x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$$



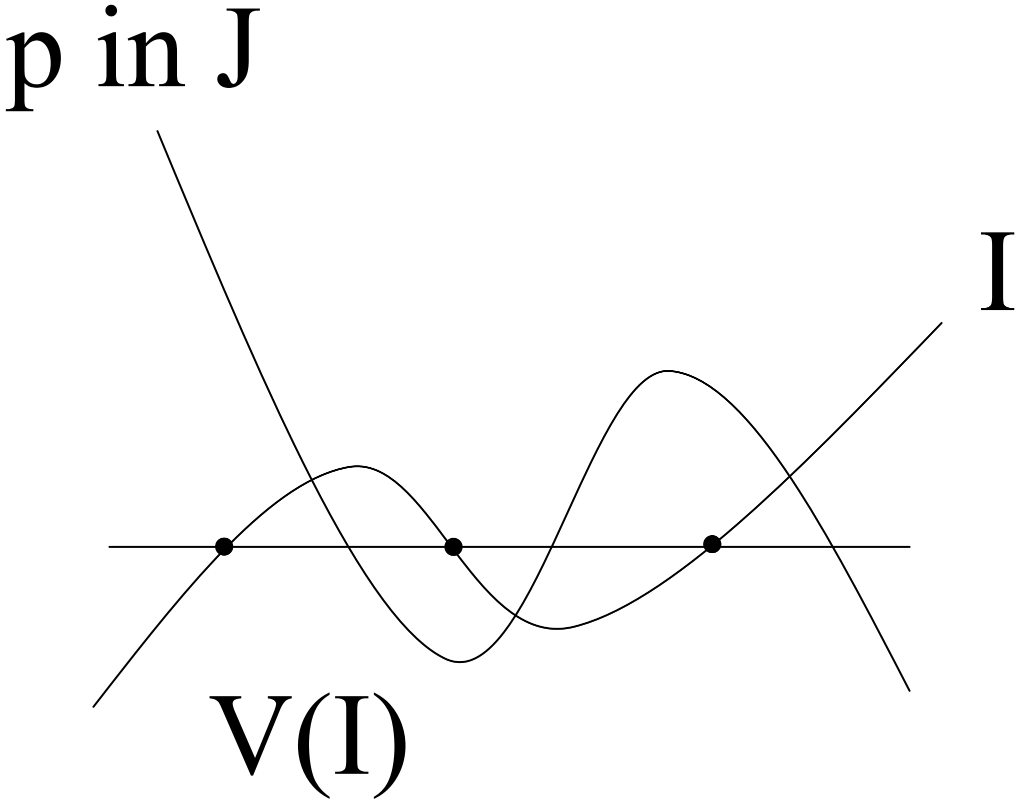
Action Matrix



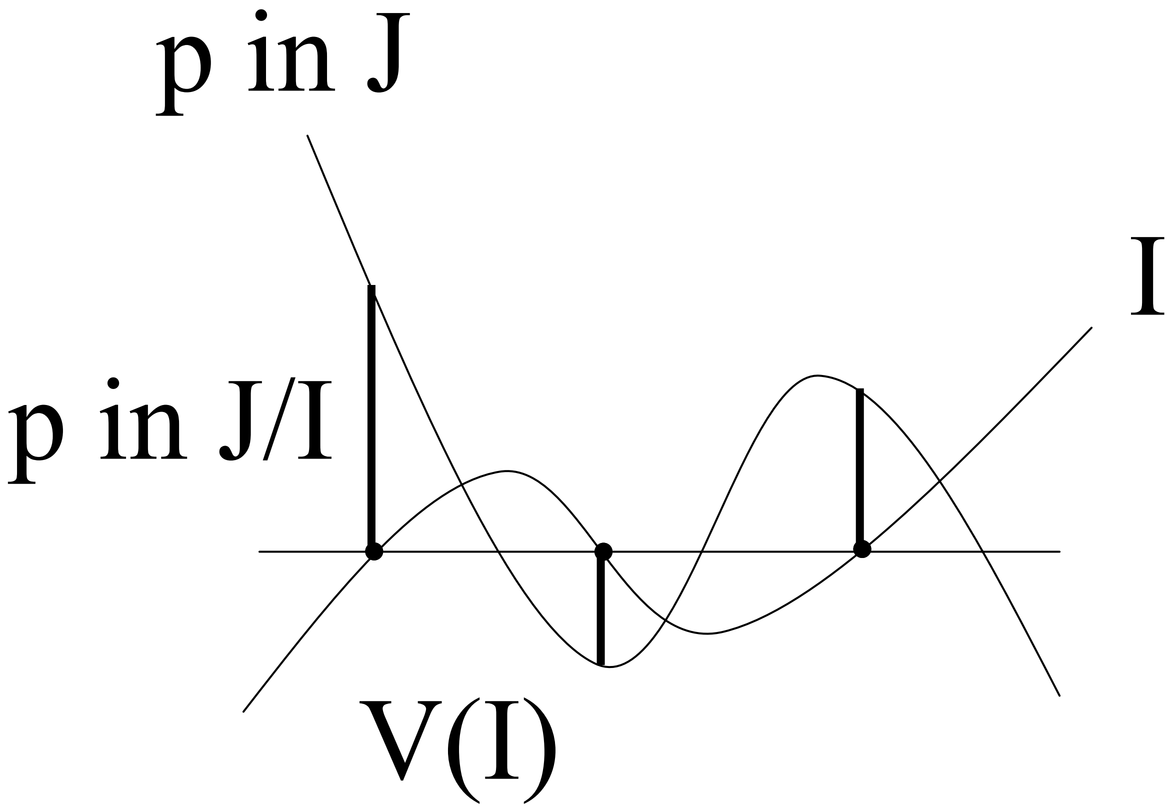
Action Matrix



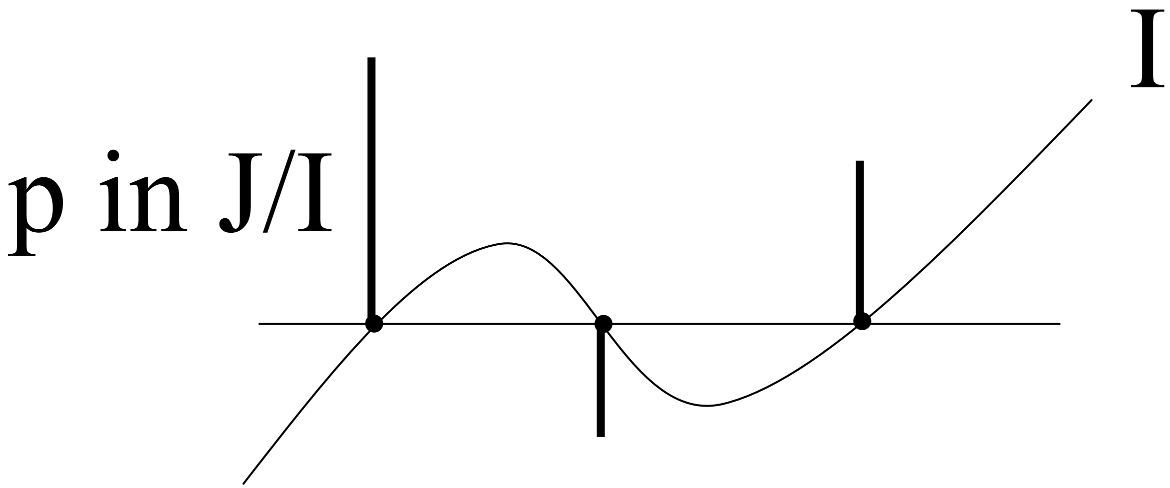
Action Matrix



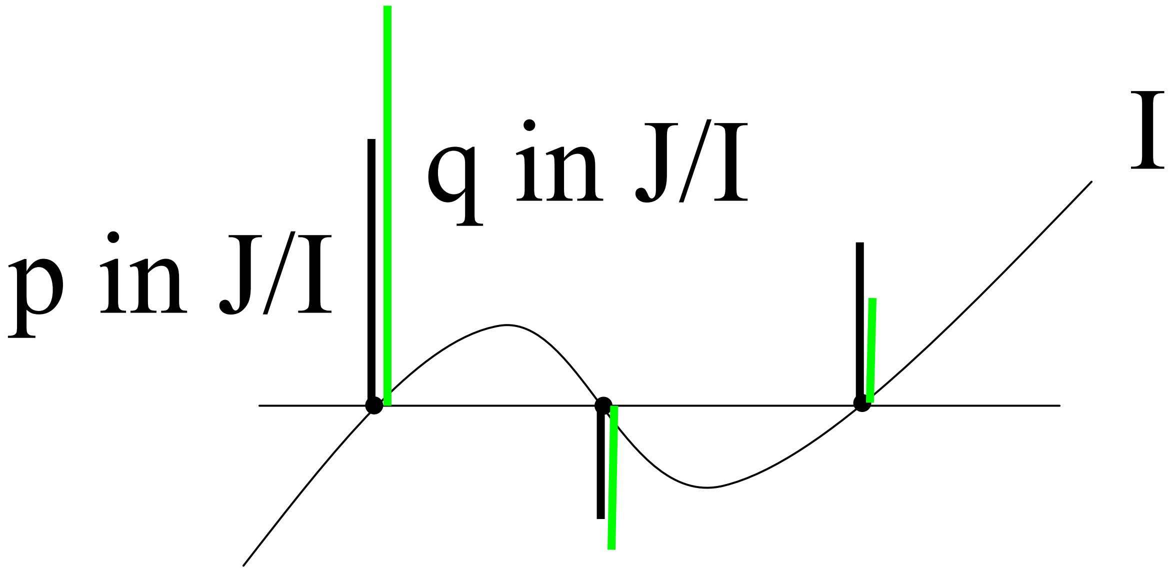
Action Matrix



Action Matrix

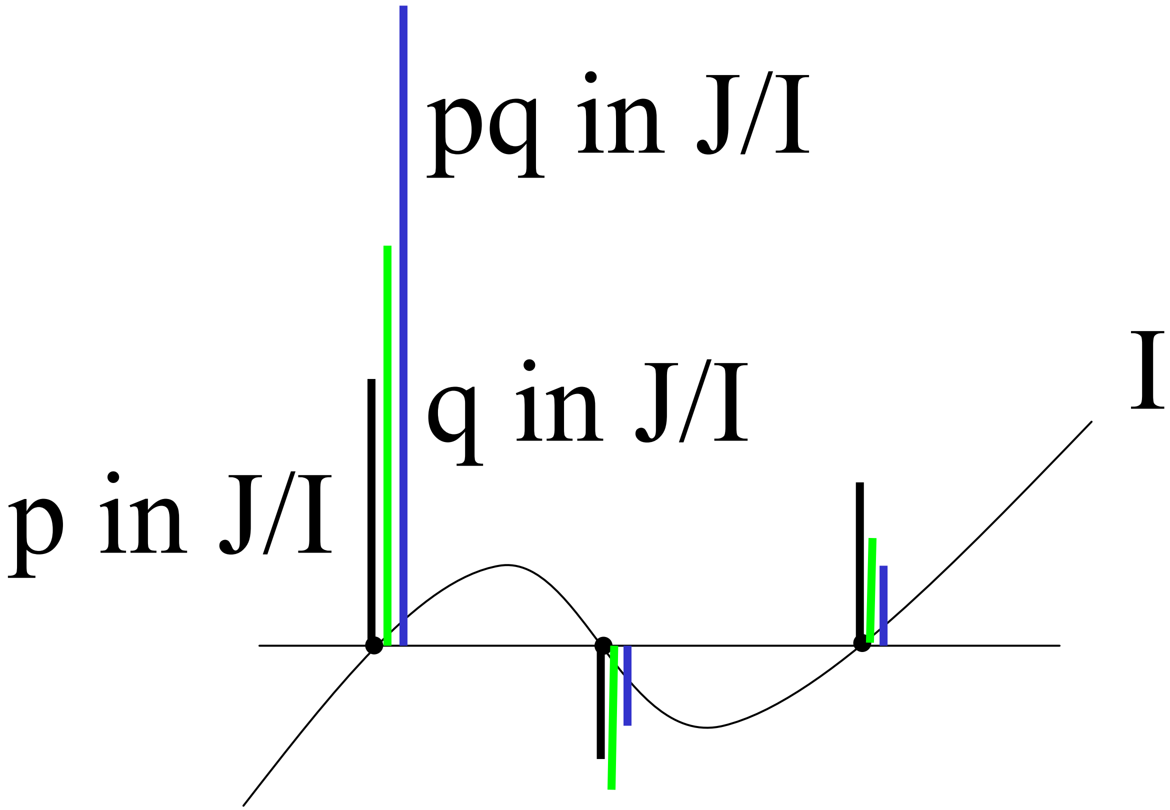


Action Matrix



Center for Visualization & Virtual Environments

Action Matrix



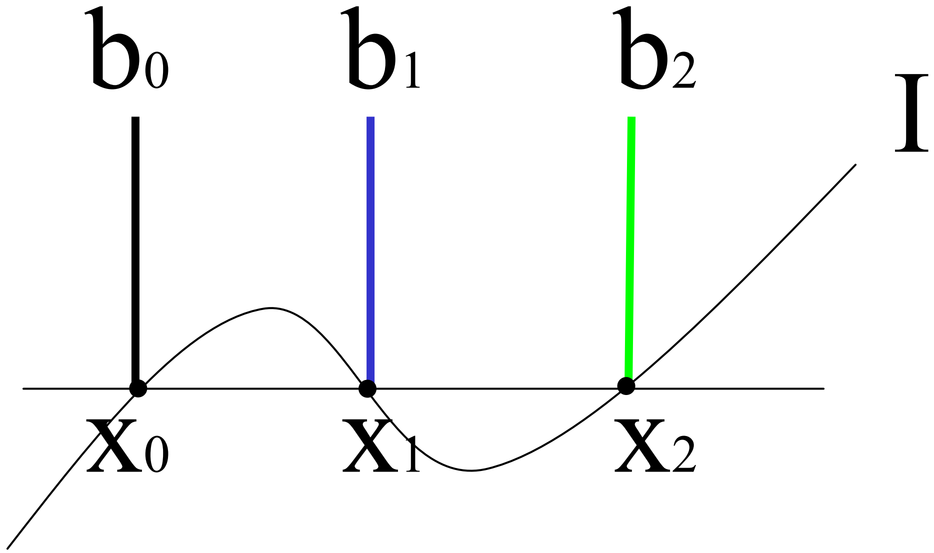
Action Matrix

Multiplication by a polynomial q is a linear operator A_q

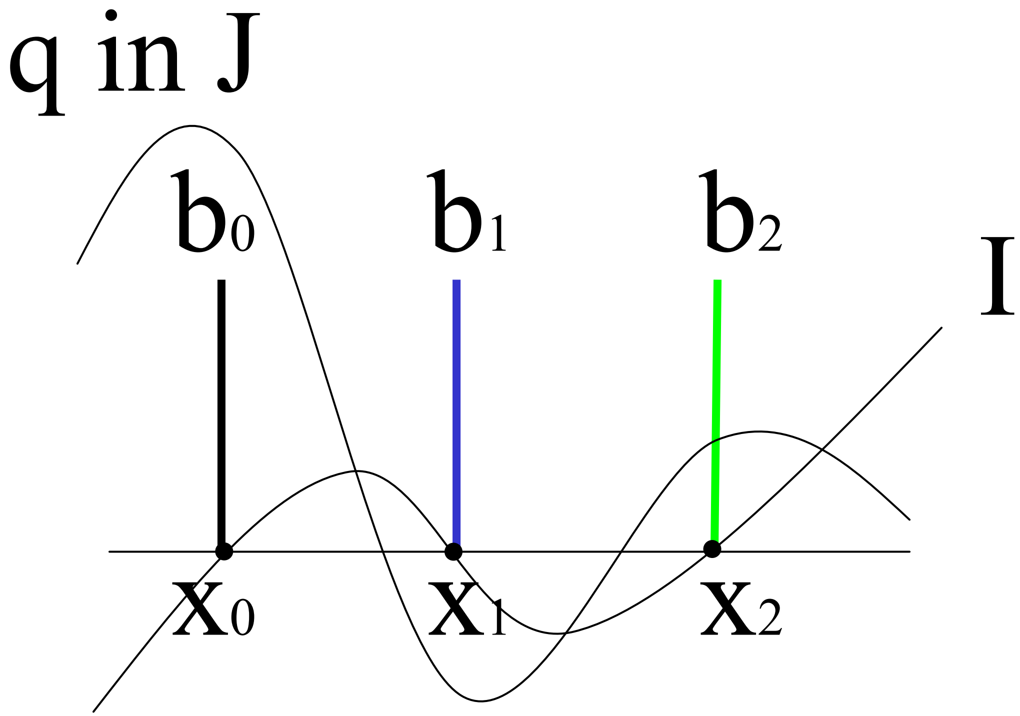
$$(\alpha p+\beta r)q=\alpha(pq)+\beta(rq)$$

The matrix A_q is called the action matrix for multiplication by q

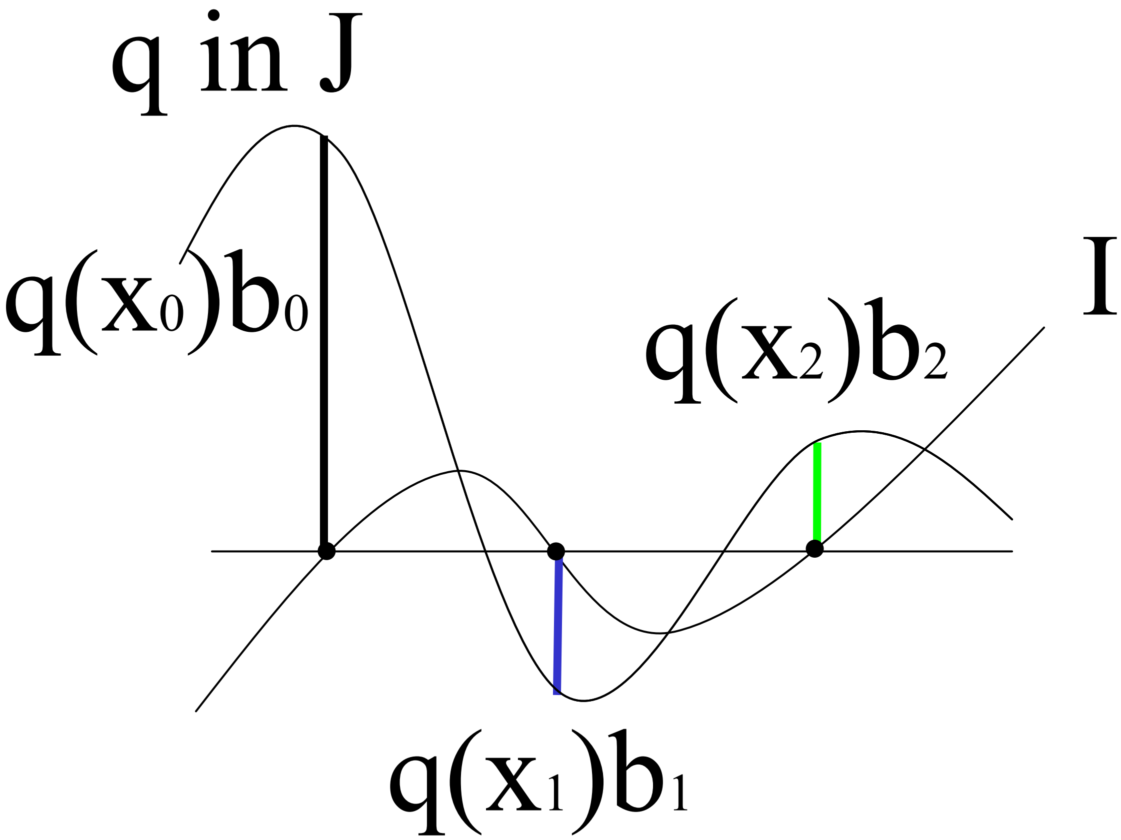
Action Matrix



Action Matrix



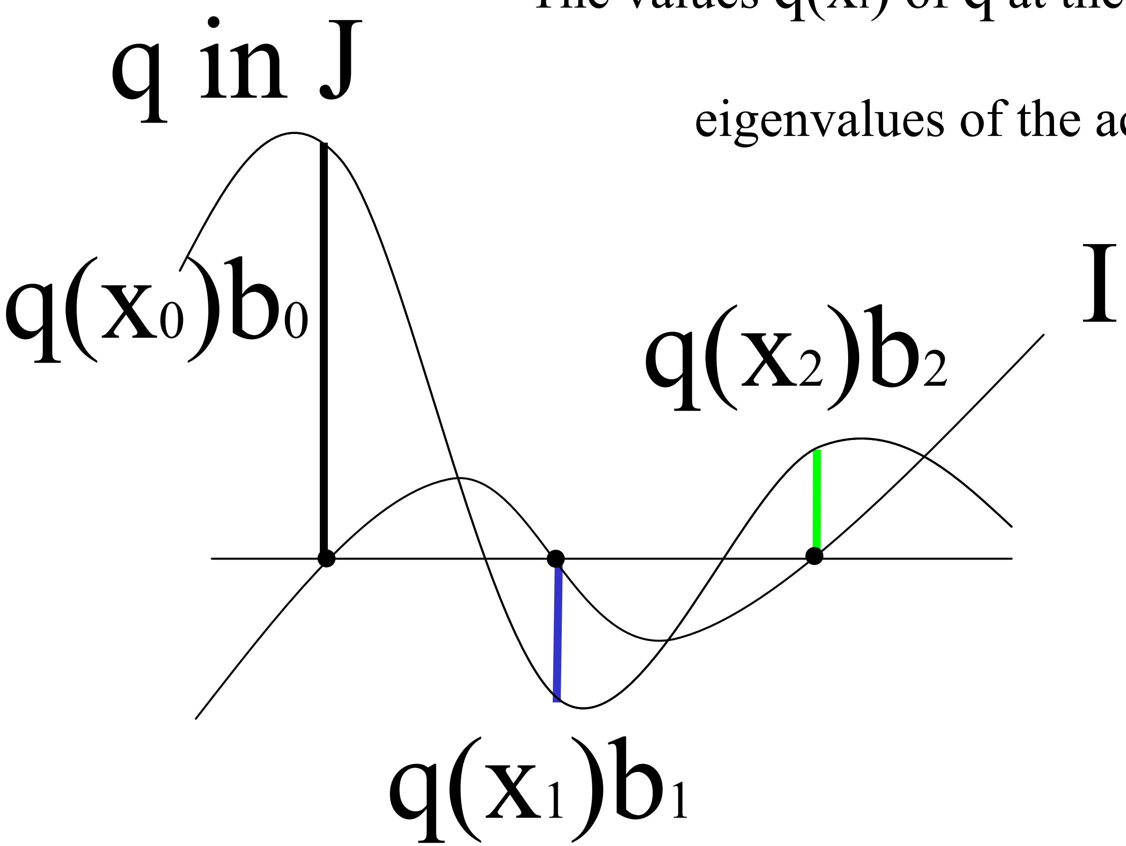
Action Matrix



Center for Visualization & Virtual Environments

Action Matrix

The values $q(x_i)$ of q at the solutions x_i are the eigenvalues of the action matrix



Action Matrix

The values $q(x_i)$ of q at the solutions x_i are the eigenvalues of the action matrix

If we choose $q=y_1$, the eigenvalues are the solutions for y_1

Action Matrix

$$b' = [r_1 \ \dots \ r_o]$$

$$b'(x) A_q \, p = q(x) b'(x) p$$

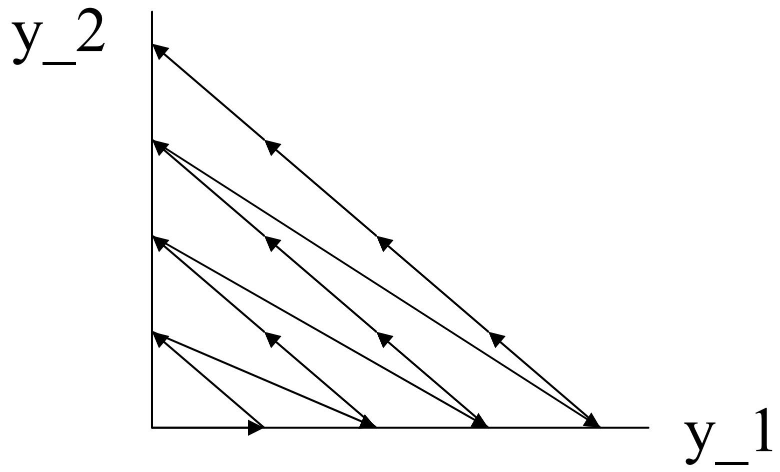
for all p in J/I and x in V(I)

$$b'(x) A_q = b'(x) q(x)$$

b(x) is a left nullvector of A_q corresponding to eigenvalue $q(x)$

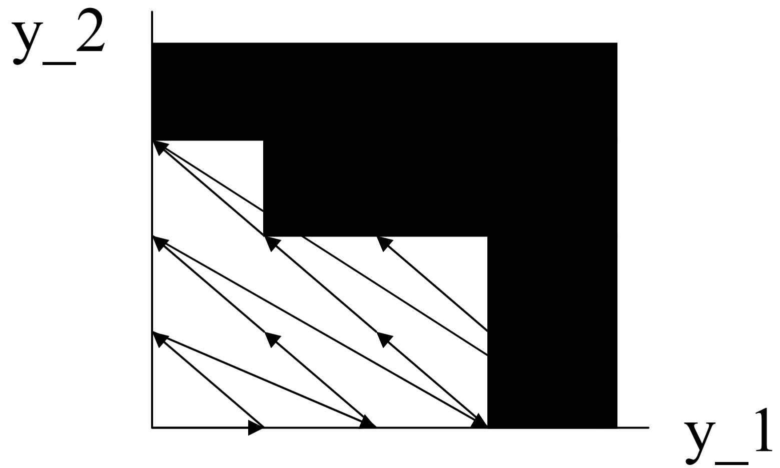
Monomial Order

- Needed to define leading term of a polynomial
- Grevlex (Graded reverse lexicographical) order usually most efficient



Gröbner Basis

- A basis for ideal I that exposes the leading terms of I (hence unique well defined remainders)
- Easily gives the action matrix for multiplication with any polynomial in the quotient ring

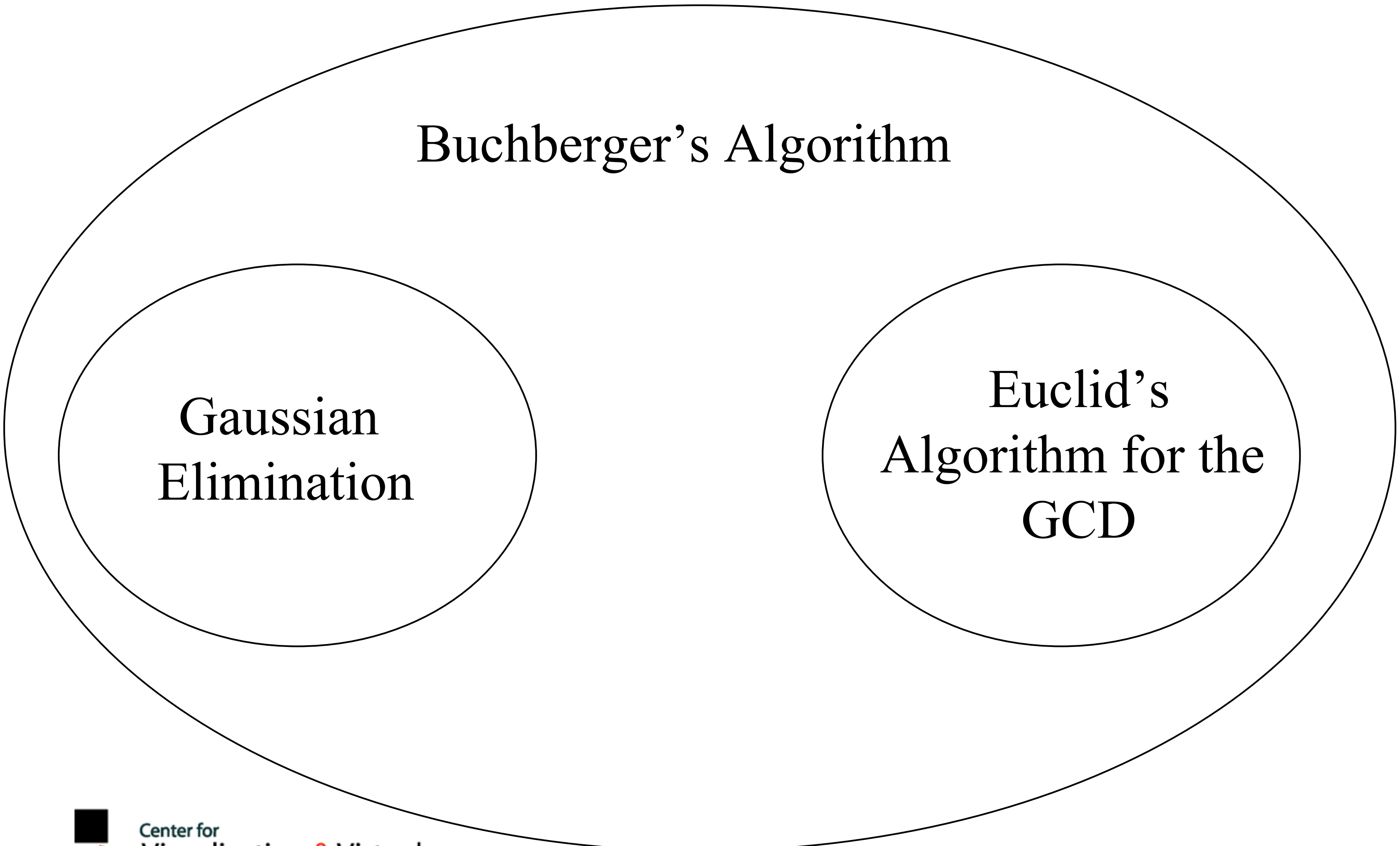


UK Center for Visualization & Virtual Environments

A Reduced Gröbner Basis is a Basis in the normal sense

- A polynomial in the ideal I can be written as a unique combination of the polynomials in a reduced Gröbner basis for I
- The monic Gröbner basis for I is unique

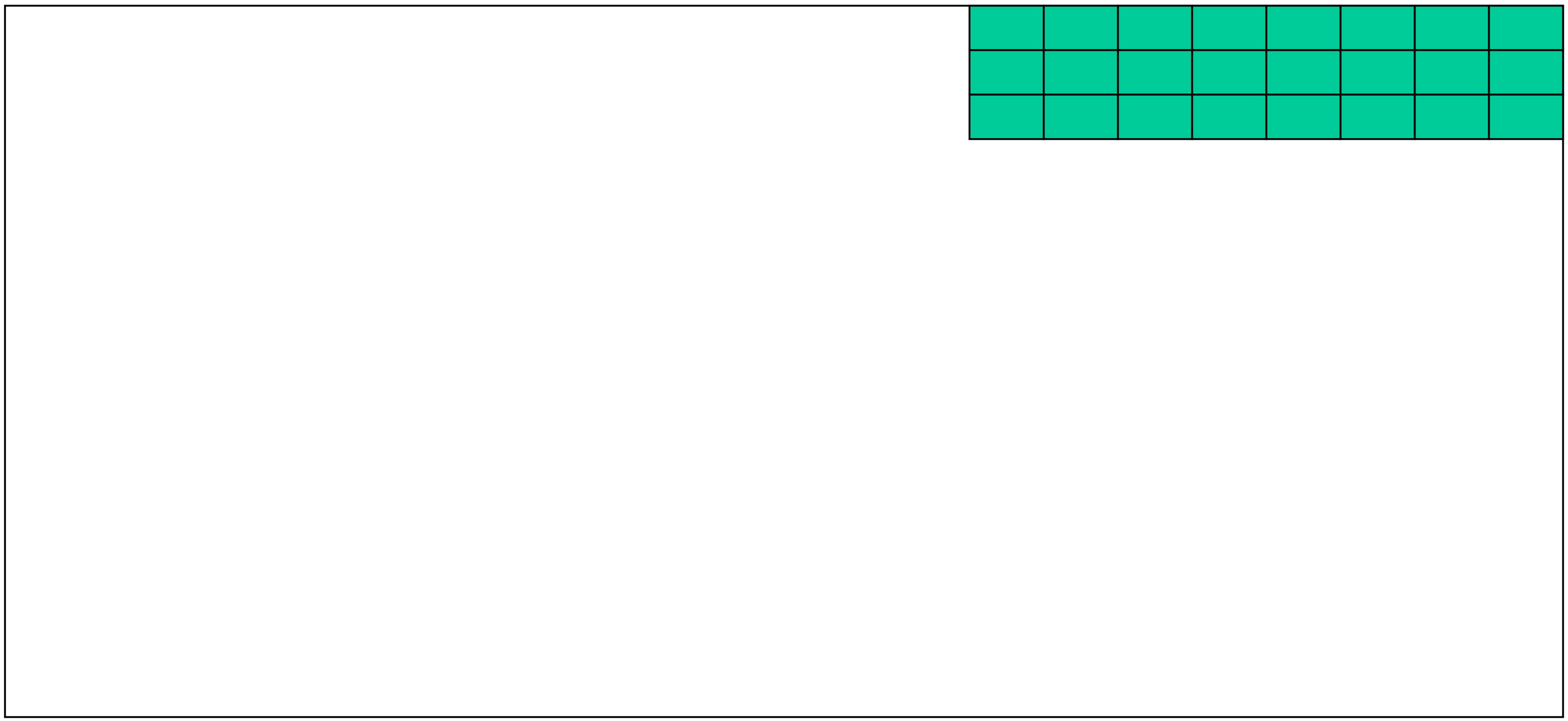
Buchberger's Algorithm



Center for Visualization & Virtual Environments

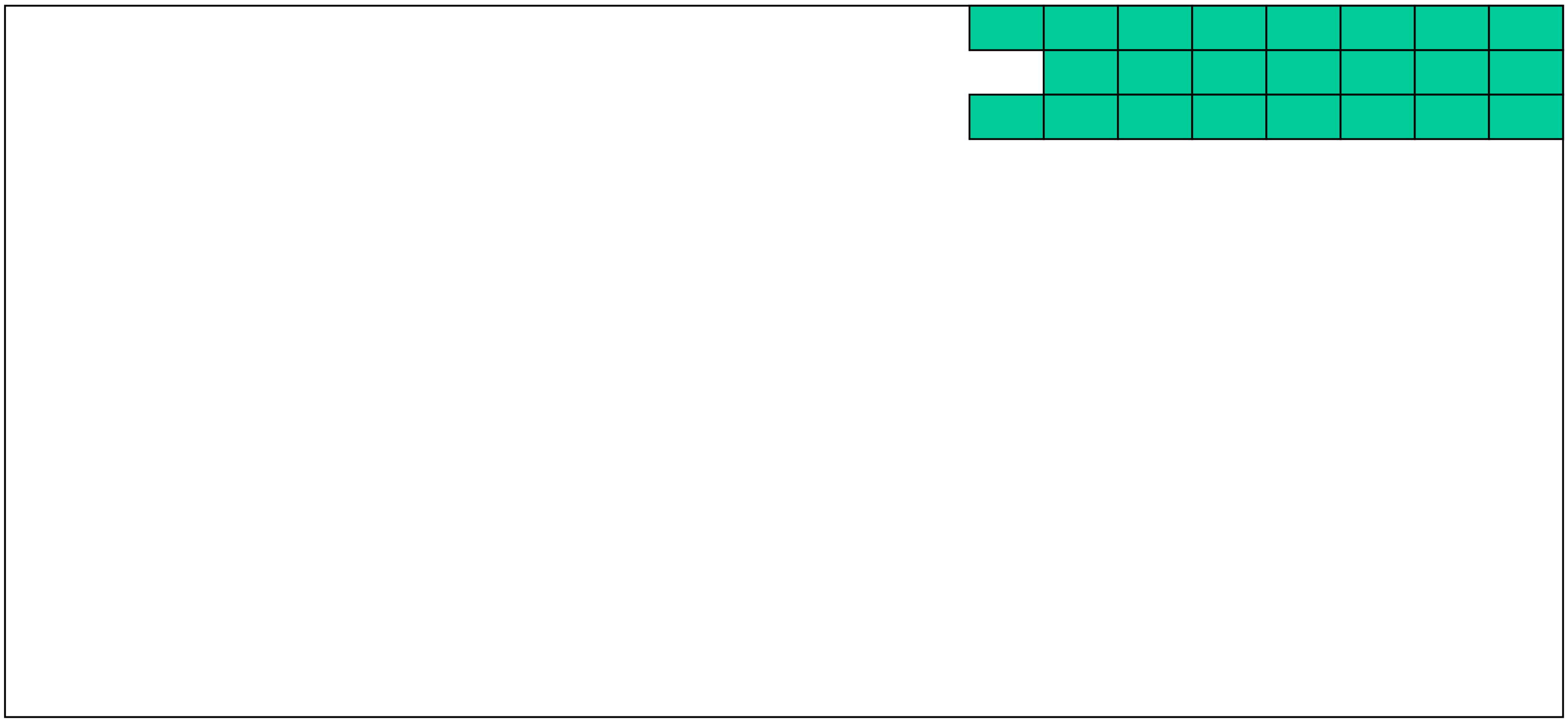
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



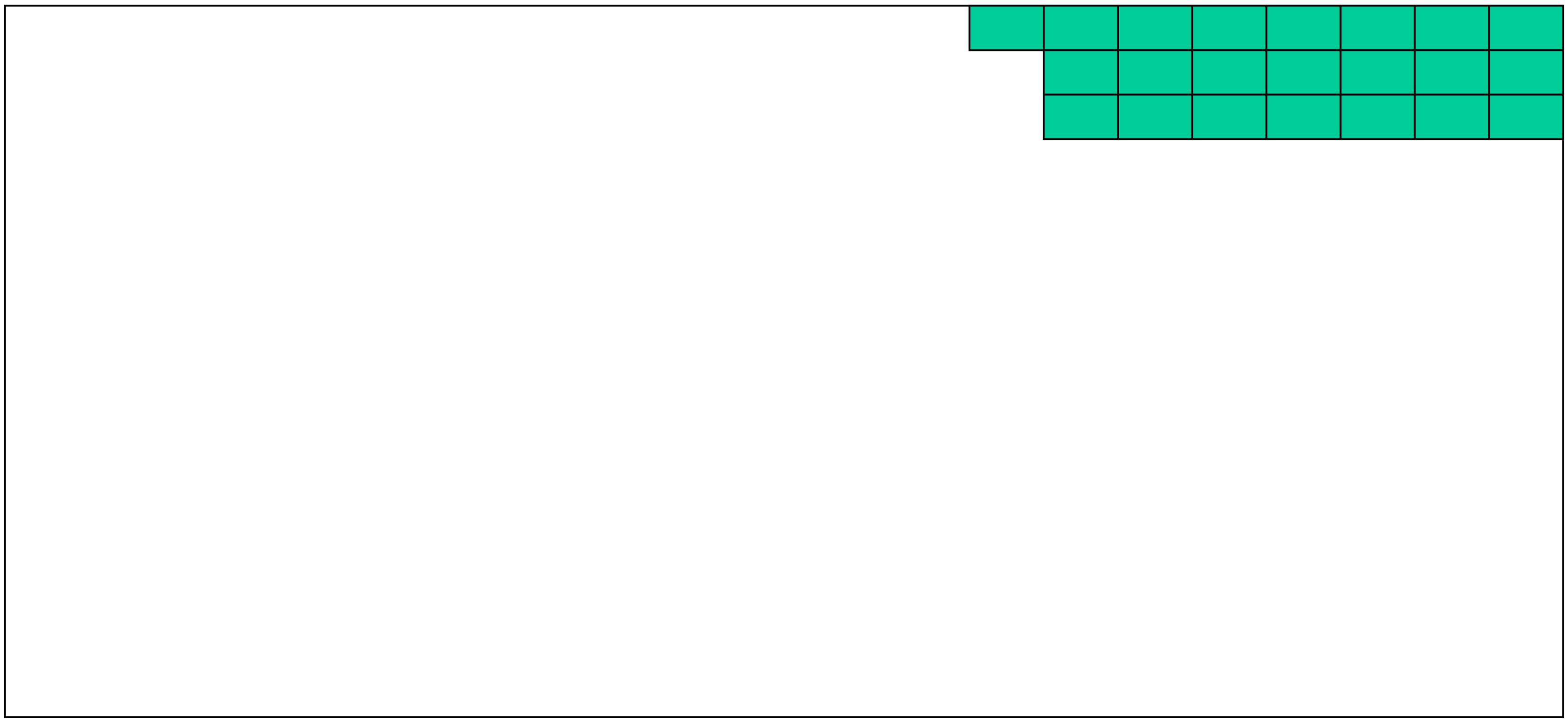
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



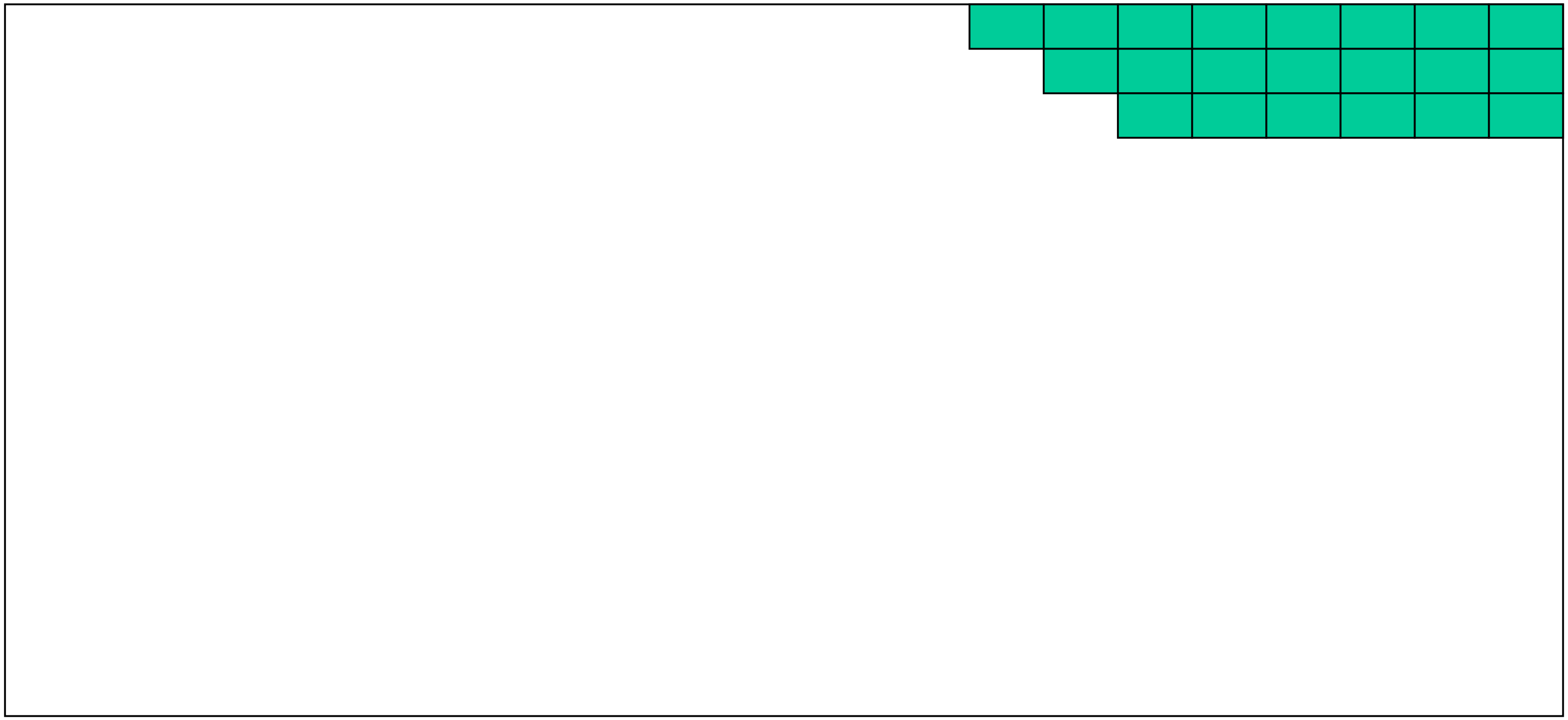
Buchberger’s Algorithm

Compute remainders of S-polynomials until all remainders are zero



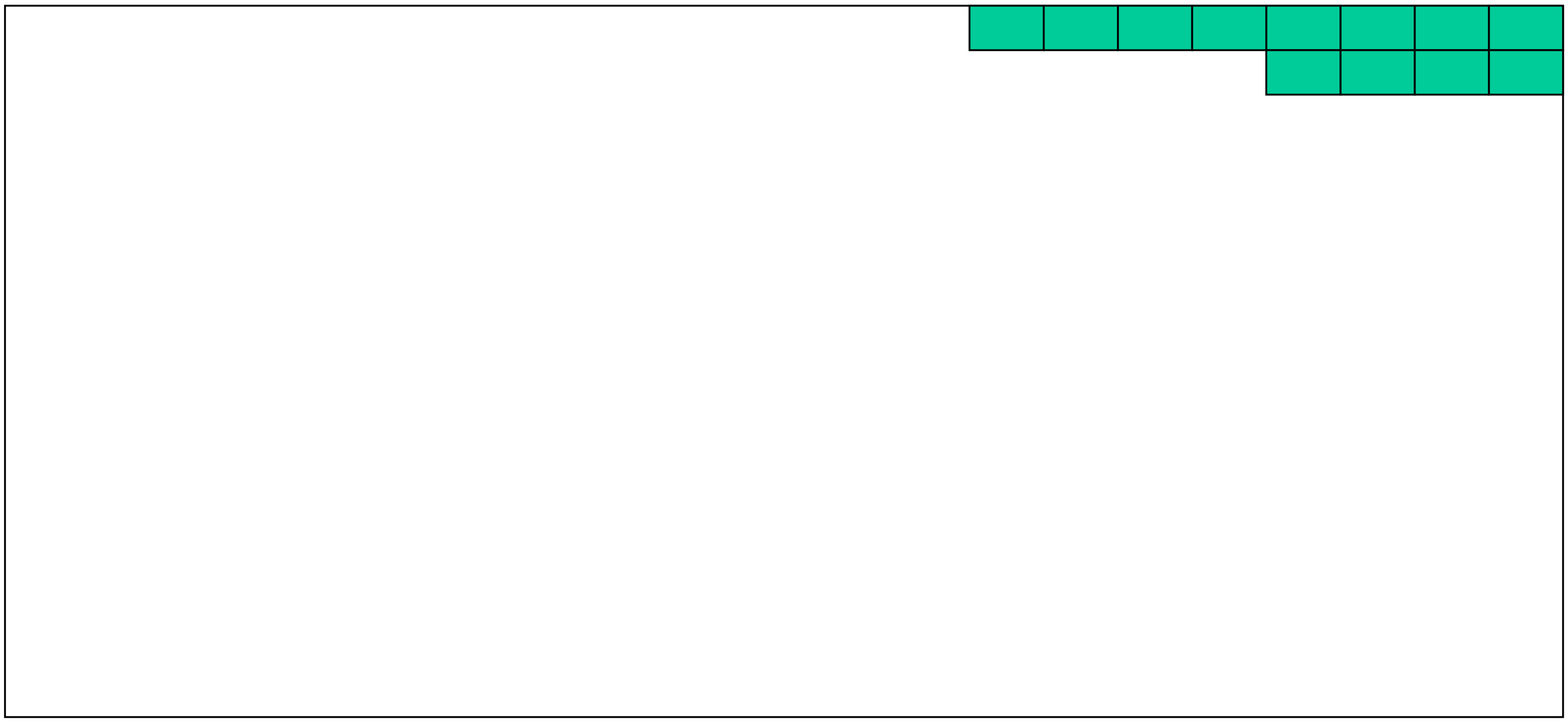
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



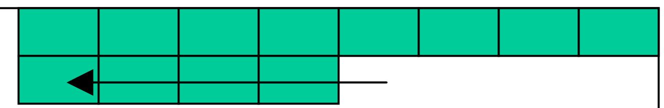
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



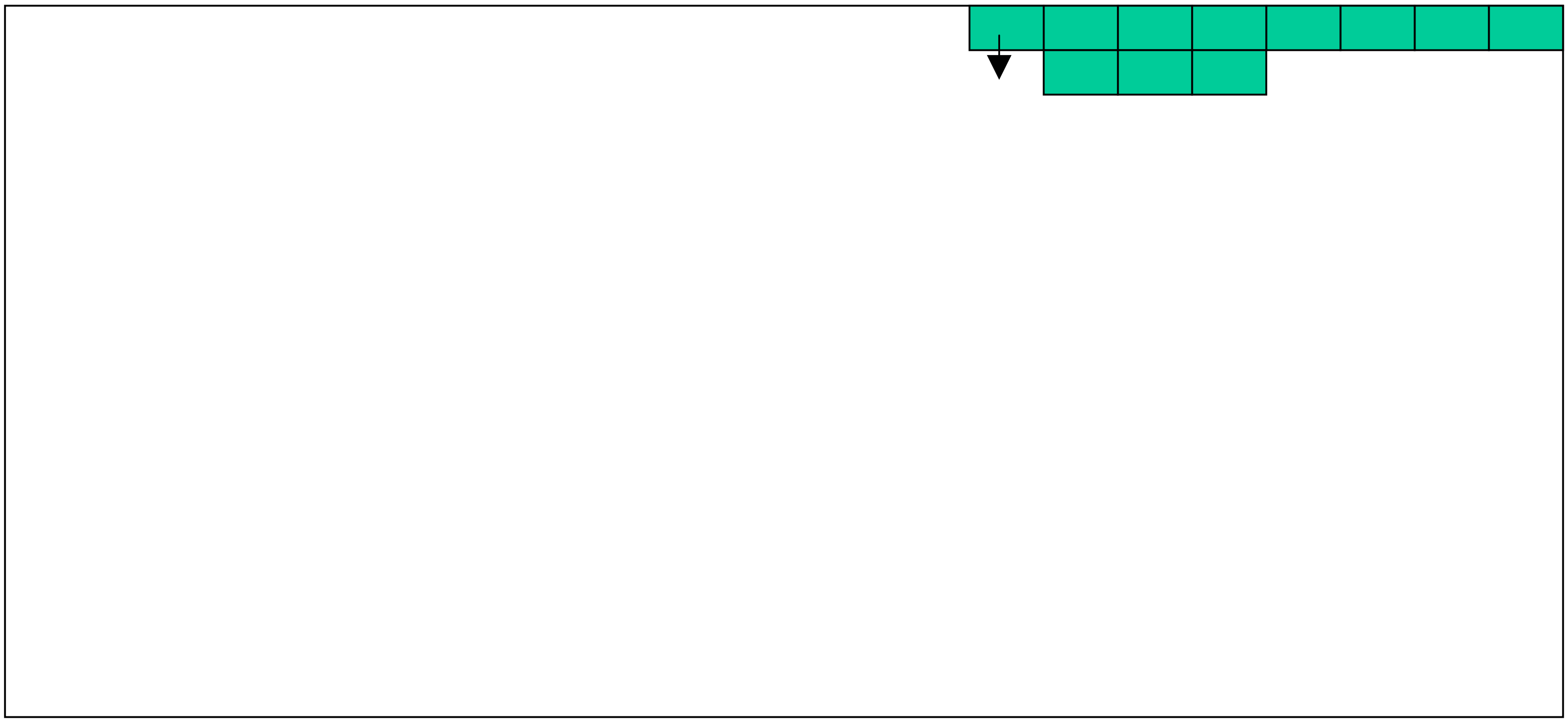
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



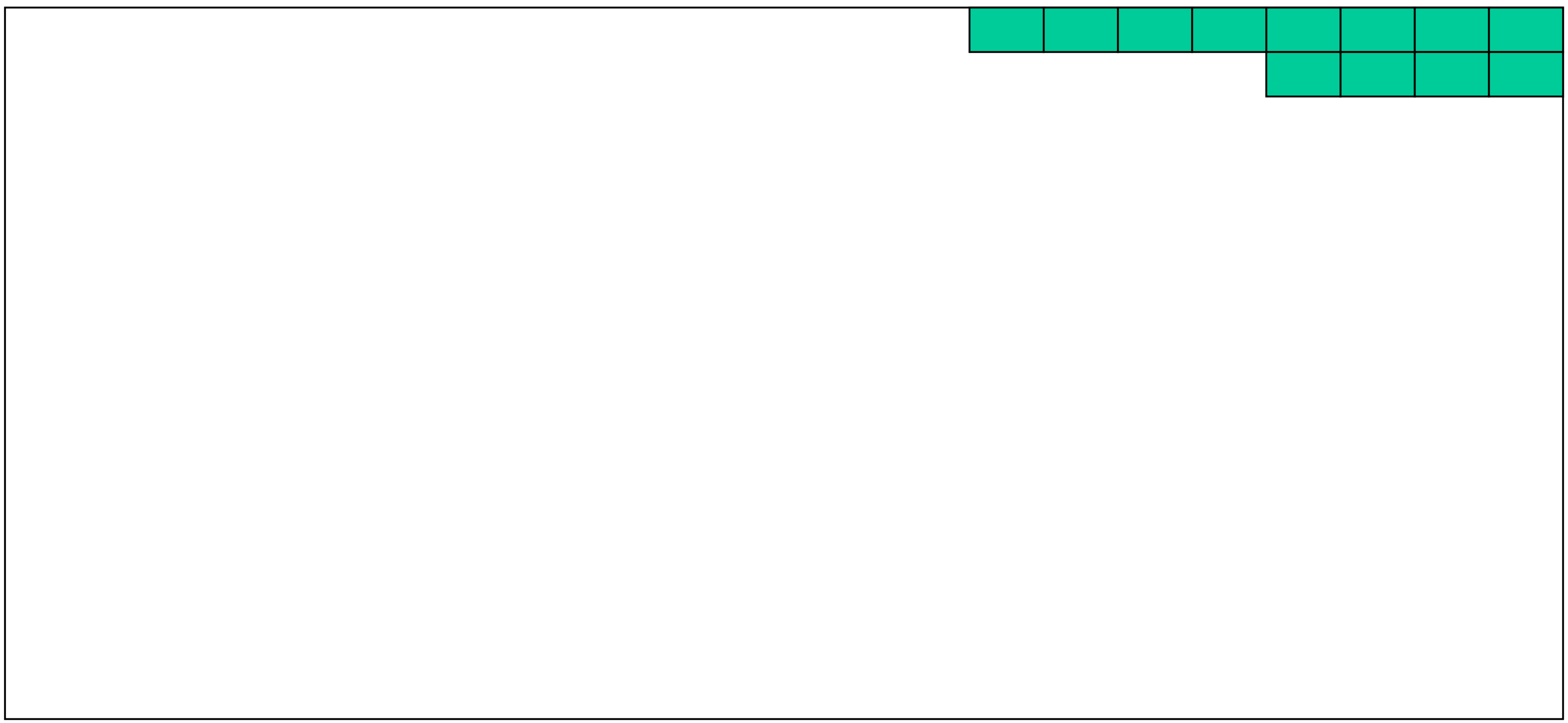
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



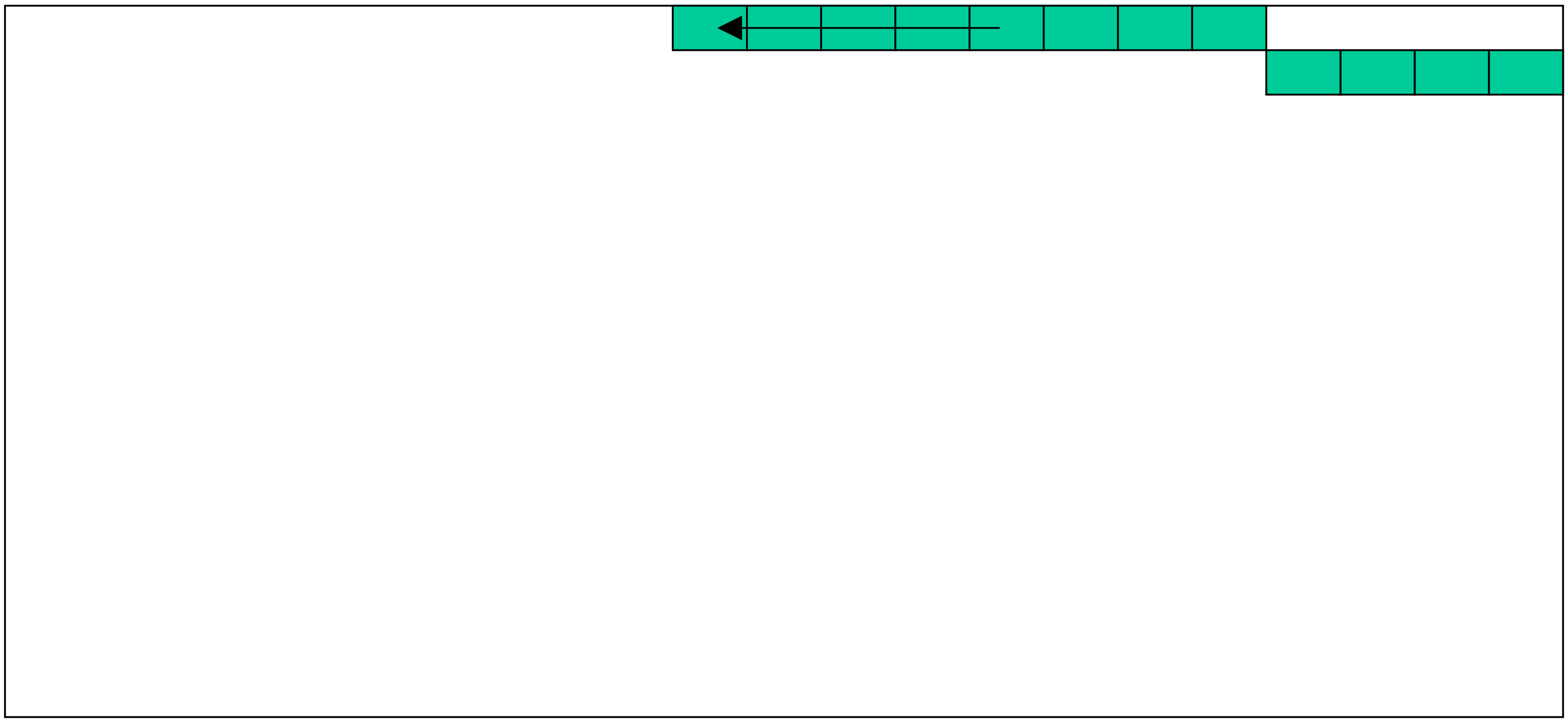
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



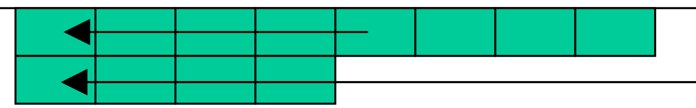
Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



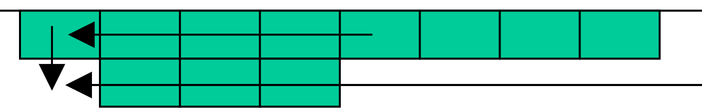
Buchberger’s Algorithm

Compute remainders of S-polynomials until all remainders are zero



Buchberger's Algorithm

Compute remainders of S-polynomials until all remainders are zero



Prime Field Formulation

- Reals => Cancellation unclear
- Rationals => Grows unwieldy
- Prime Field => Cancellation clear, size is limited, only small risk of incorrect cancellation if prime is large

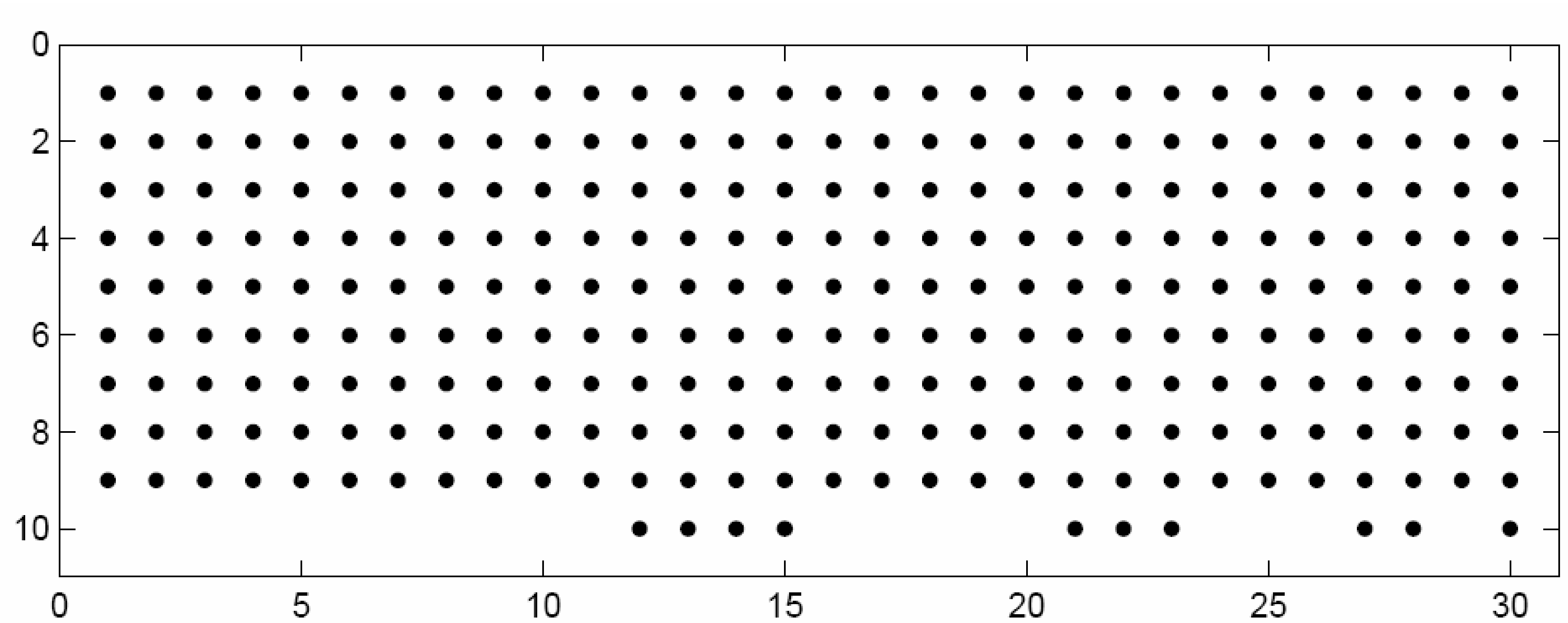
Gaussian Elimination

- Expanding all polynomials up to a certain degree followed by Gaussian elimination allows pivoting

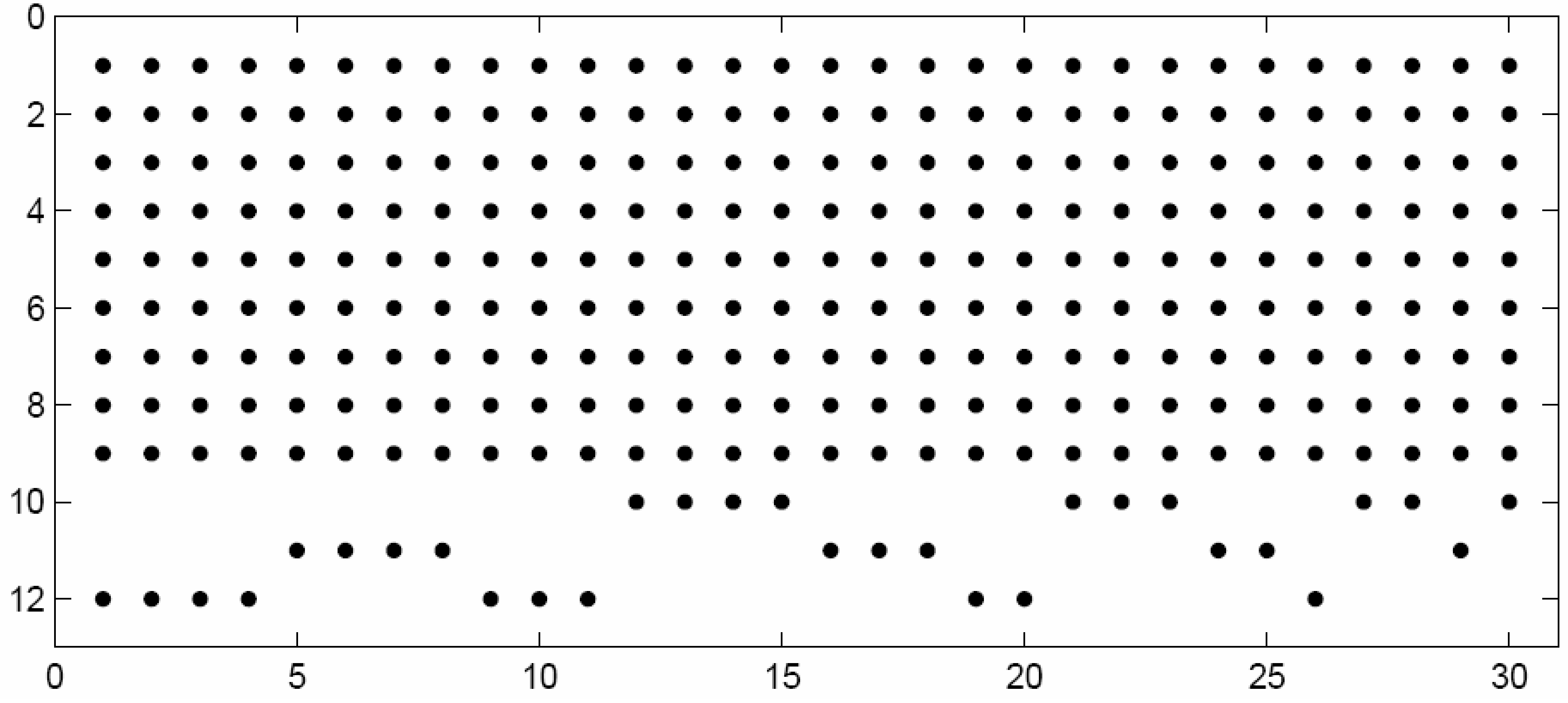
Unwanted Solutions

Can be removed by ideal quotients, or more generally saturation

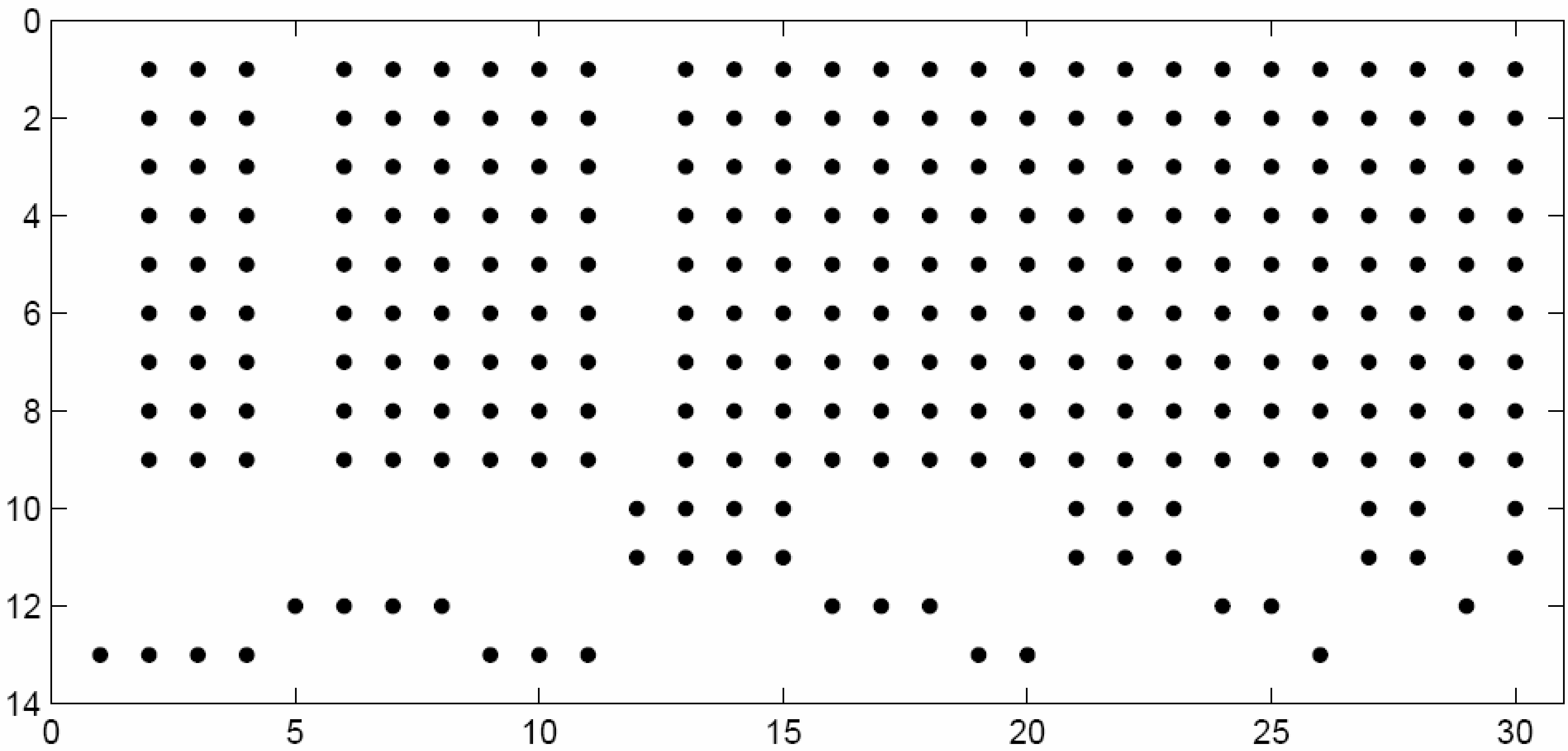
Elimination Example



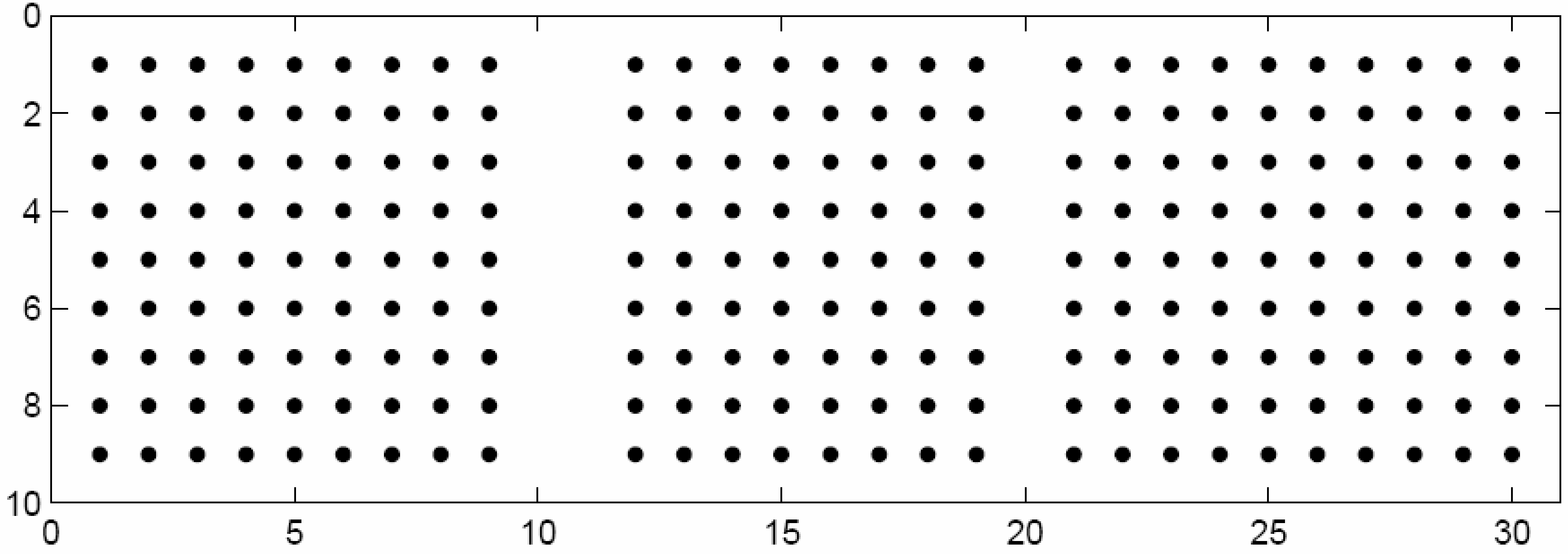
Elimination Example



Elimination Example

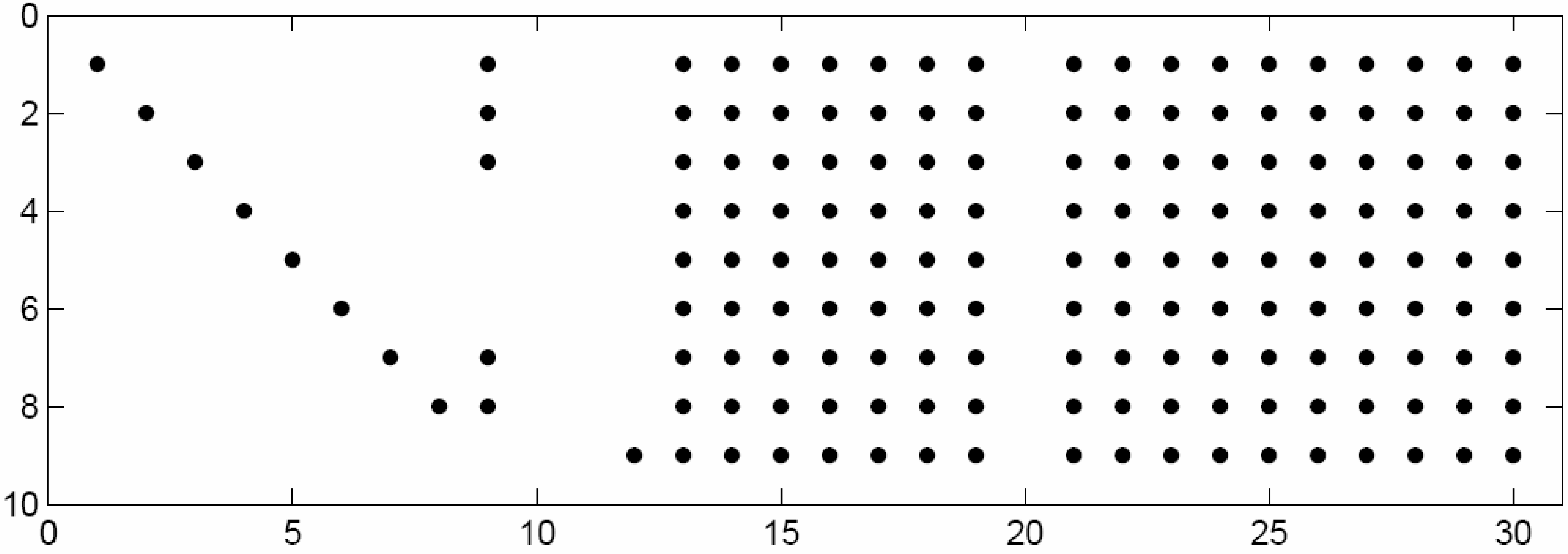


Elimination Example

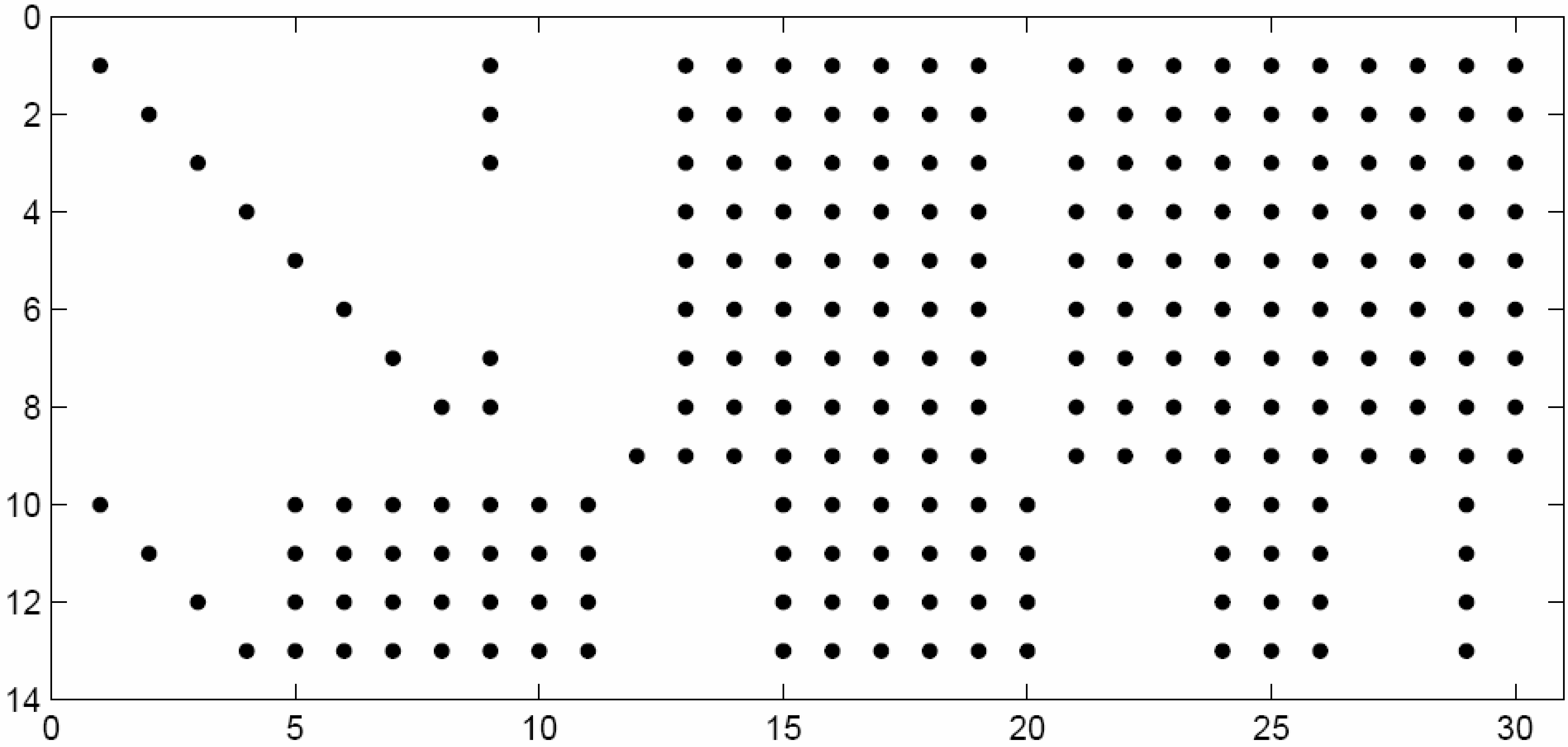


Center for Visualization & Virtual Environments

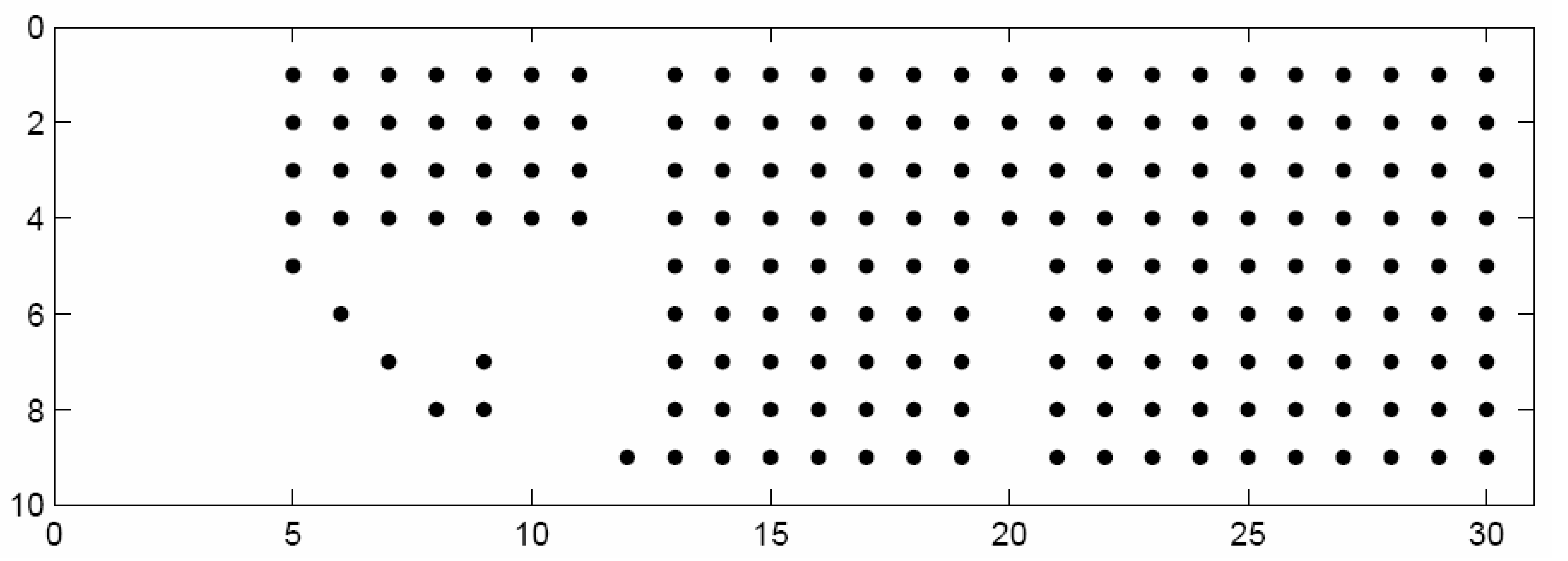
Elimination Example



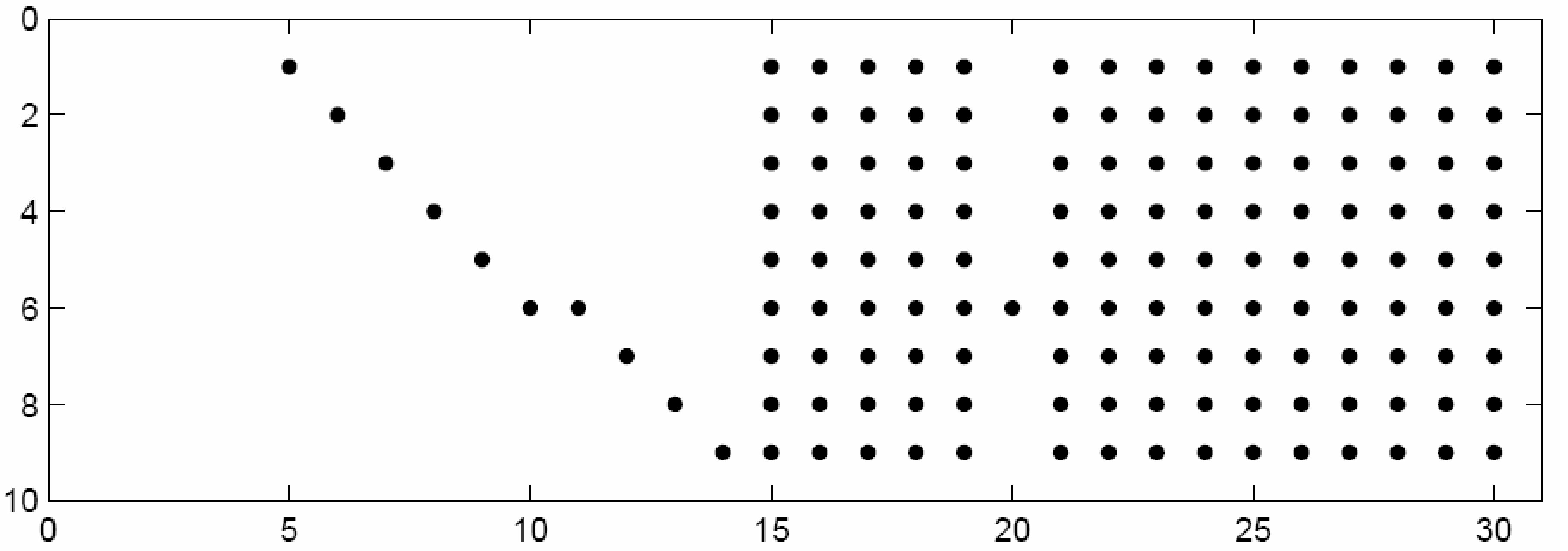
Elimination Example



Elimination Example

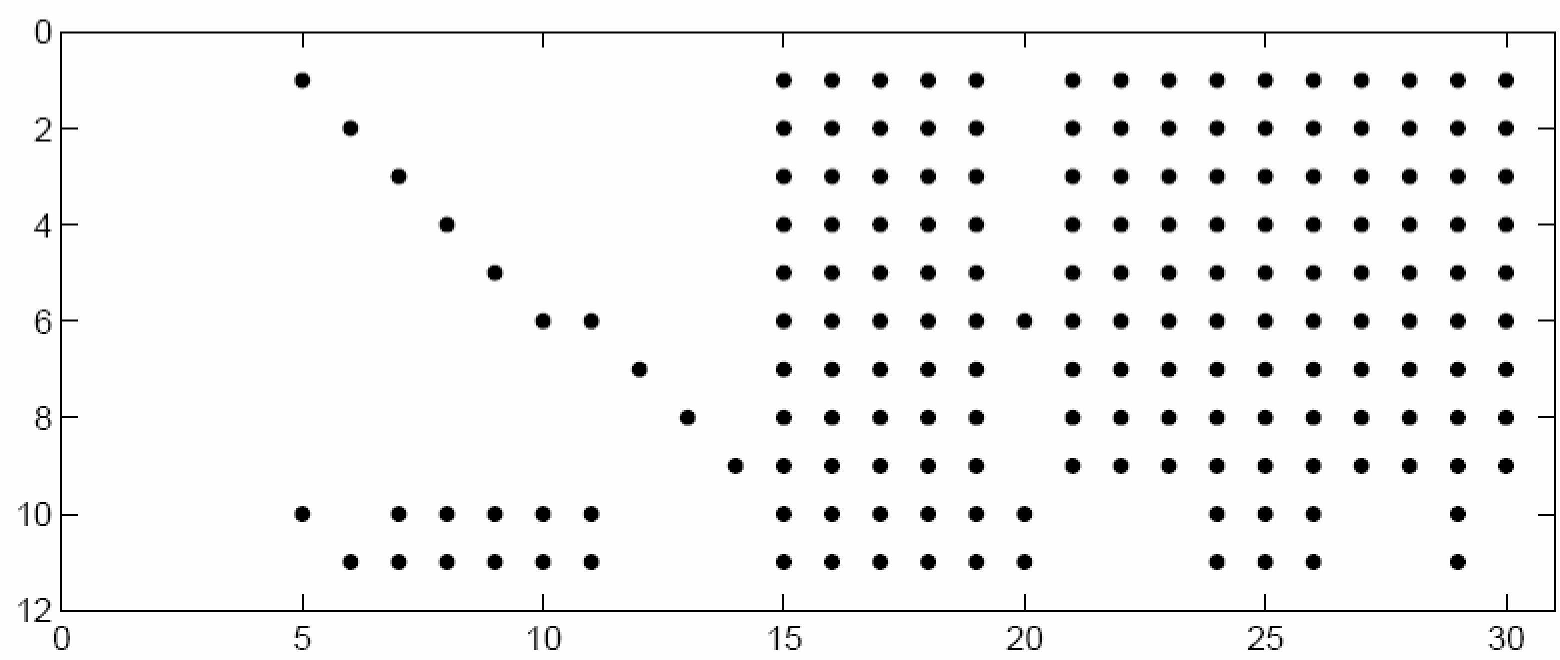


Elimination Example

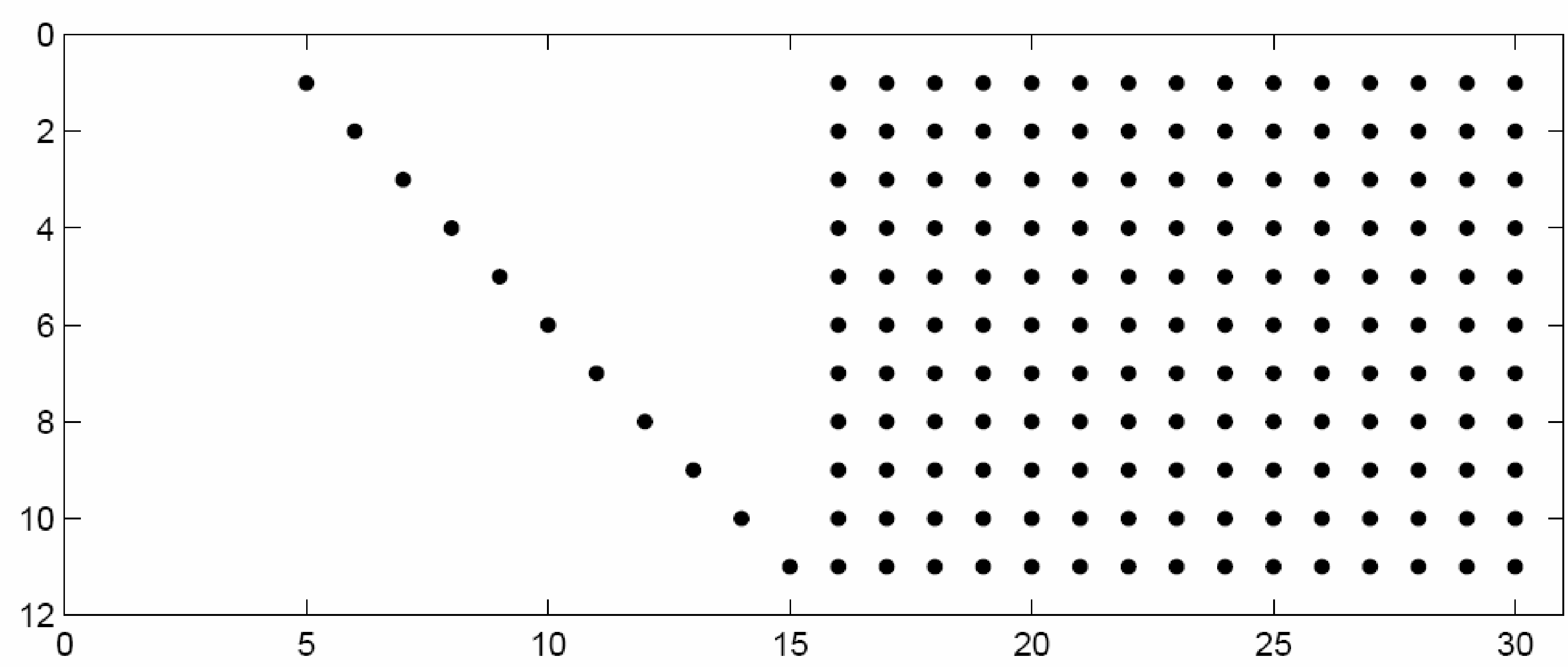


Center for Visualization & Virtual Environments

Elimination Example

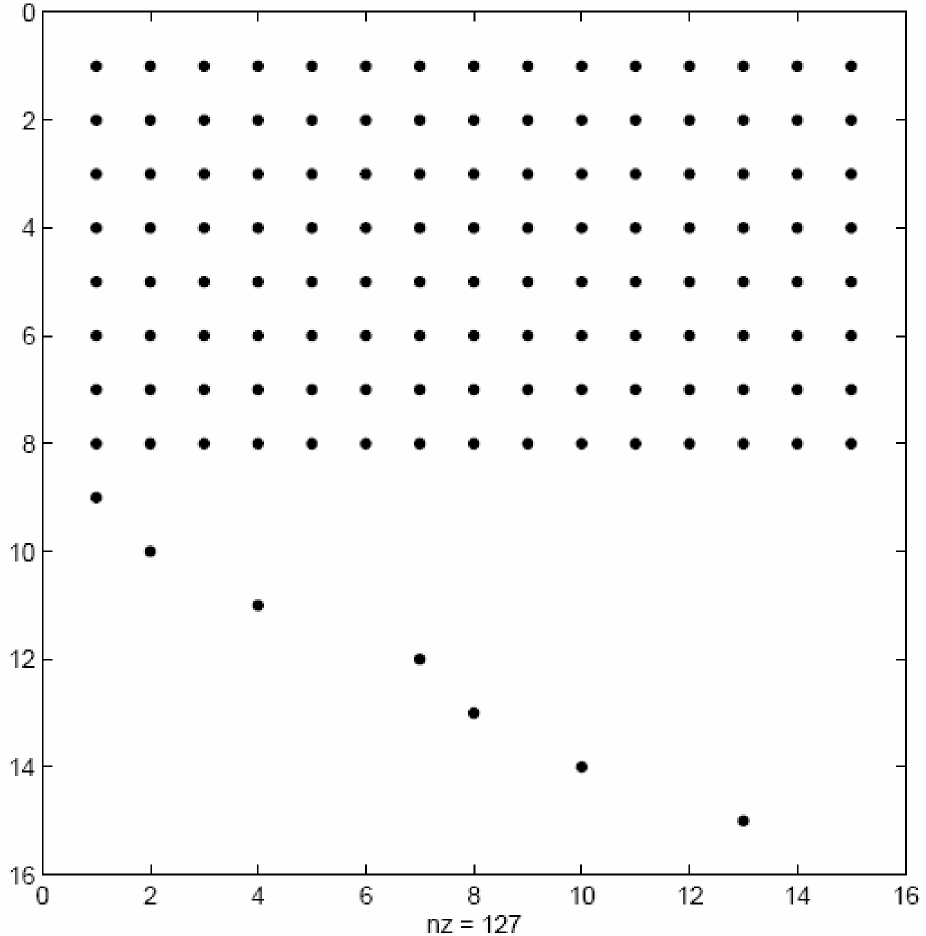


Elimination Example



Center for Visualization & Virtual Environments

Action Matrix



Center for Visualization & Virtual Environments

Generalized Camera in M2

- Generalized Camera (good problem formulation)

```
mmmmmmmmmmooomoomommooooooooommmmmmmm
mmmmmmommoommmooooooooooooooooooomooo
ooooooooooororrrororoo
```

64 solutions

- Generalized Camera (bad problem formulation)

```
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm
mmmmmmmmmmmommmommmmommmmmmmoommmmmmm
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm
            ...... (total 168 lines)
rrrrrorrrrrrrorrrrrrrrrrrrorrrrrrrrrr
```

64 solutions

UK Center for Visualization & Virtual Environments

Stratified Self-Calibration

Introduction

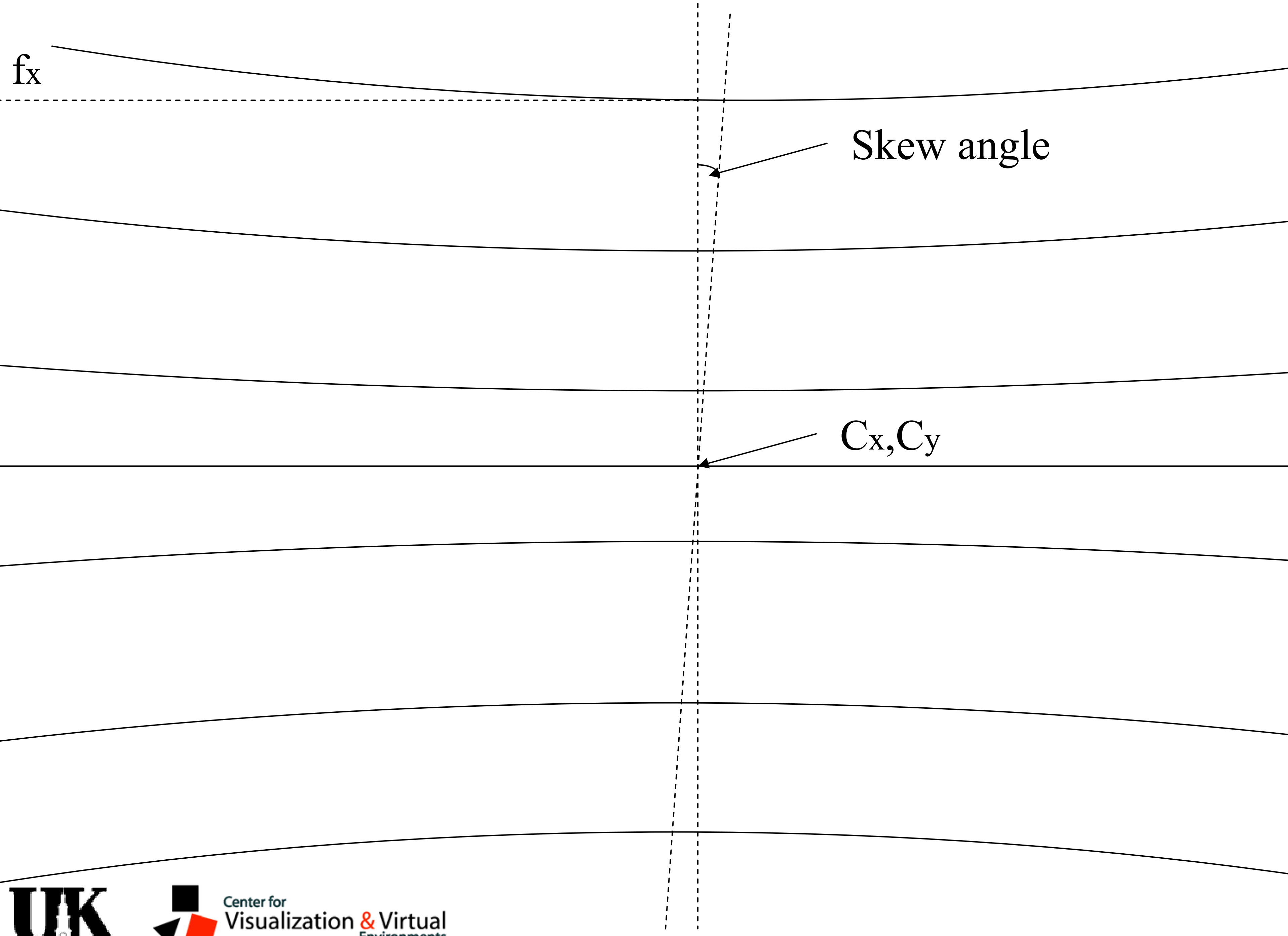
Camera calibration and the search for infinity
Hartley, Hayman, de Agapito, Reid

Calibration with robust use of cheirality by quasi-affine reconstruction of the set of camera projection centres
Nister

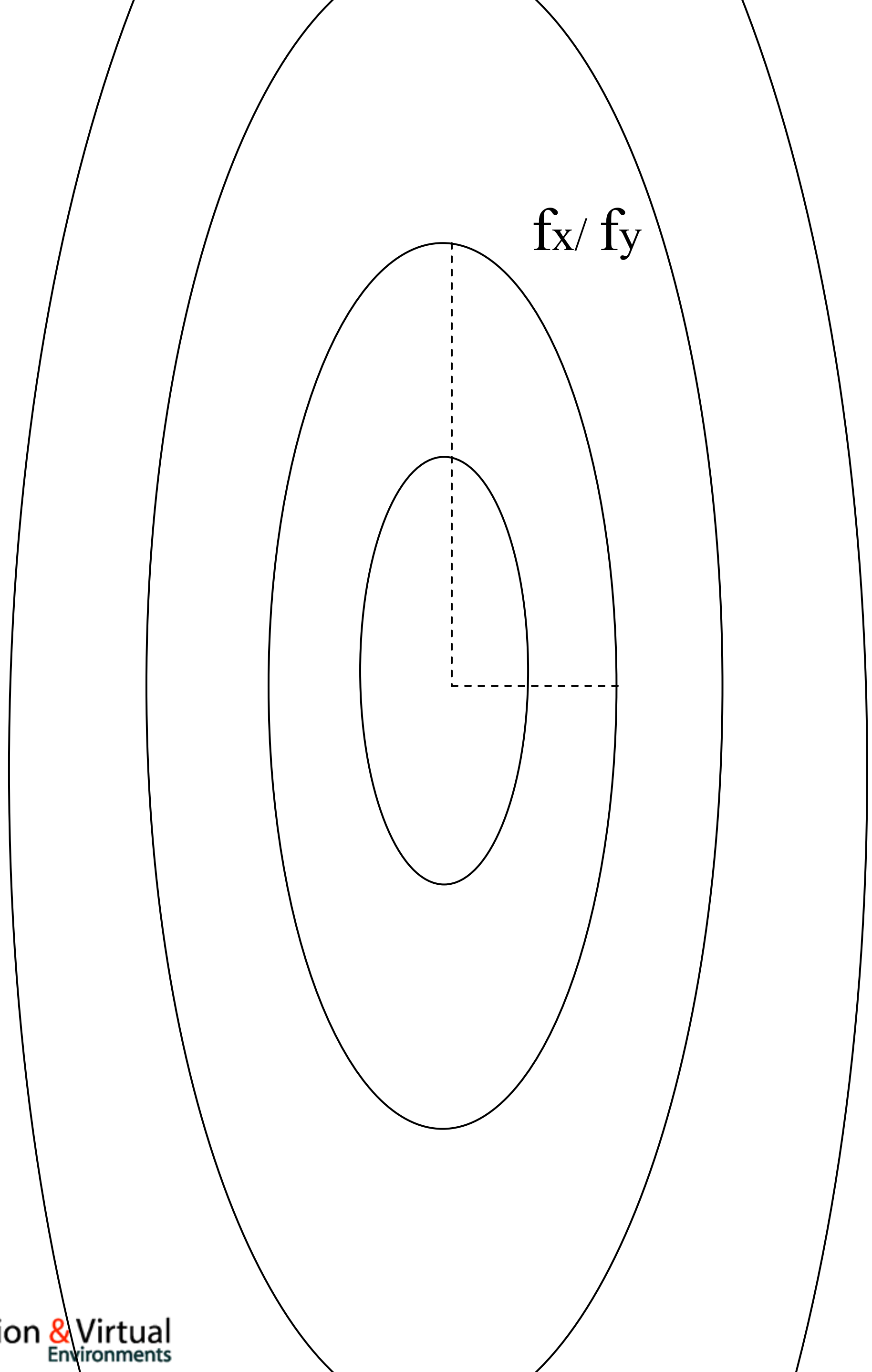
Center for Visualization & Virtual Environments

Self-calibration	Pre-calibration
Flexible	Less problems with critical surfaces (when information used correctly)

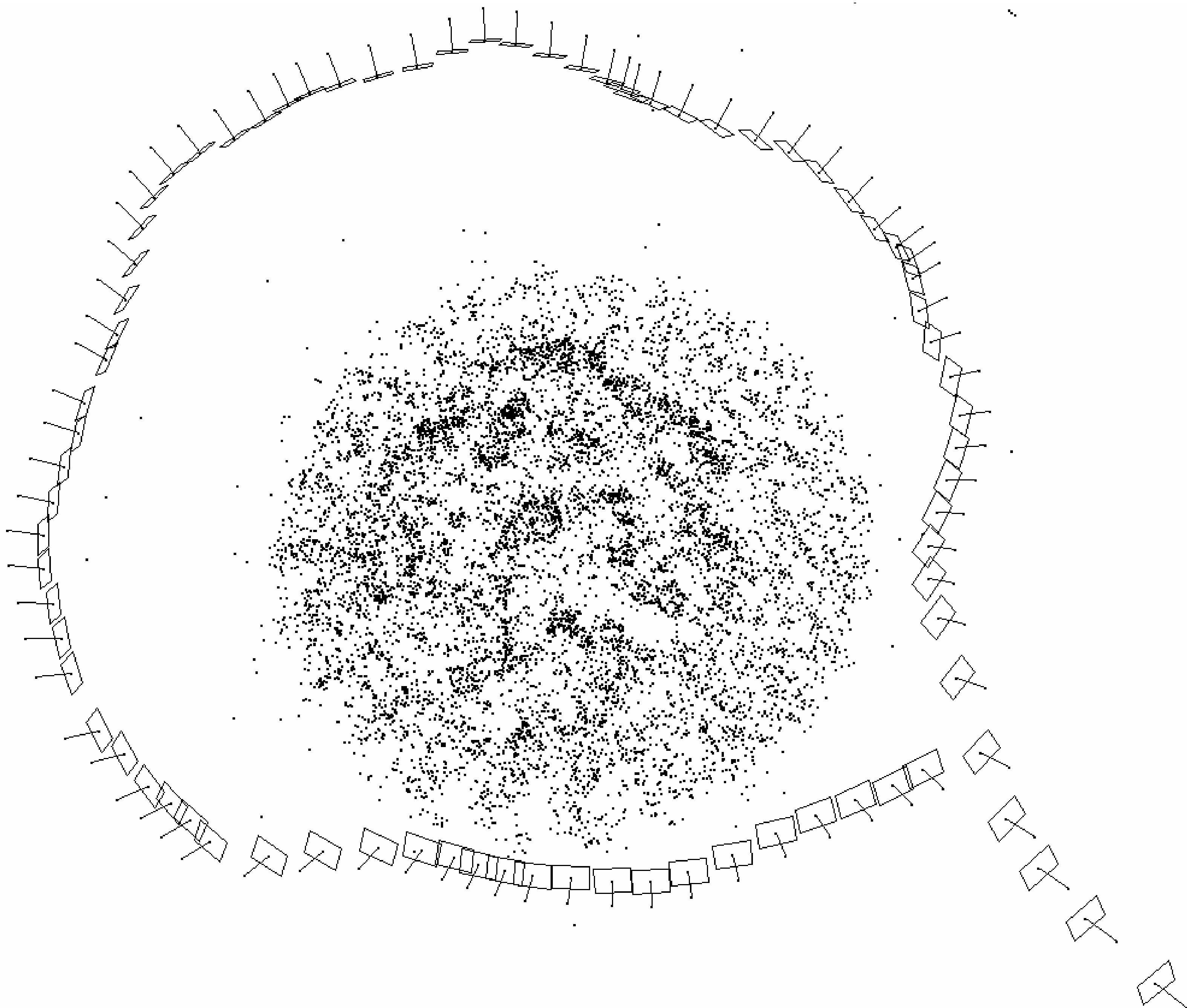
What is the cue in self-calibration?

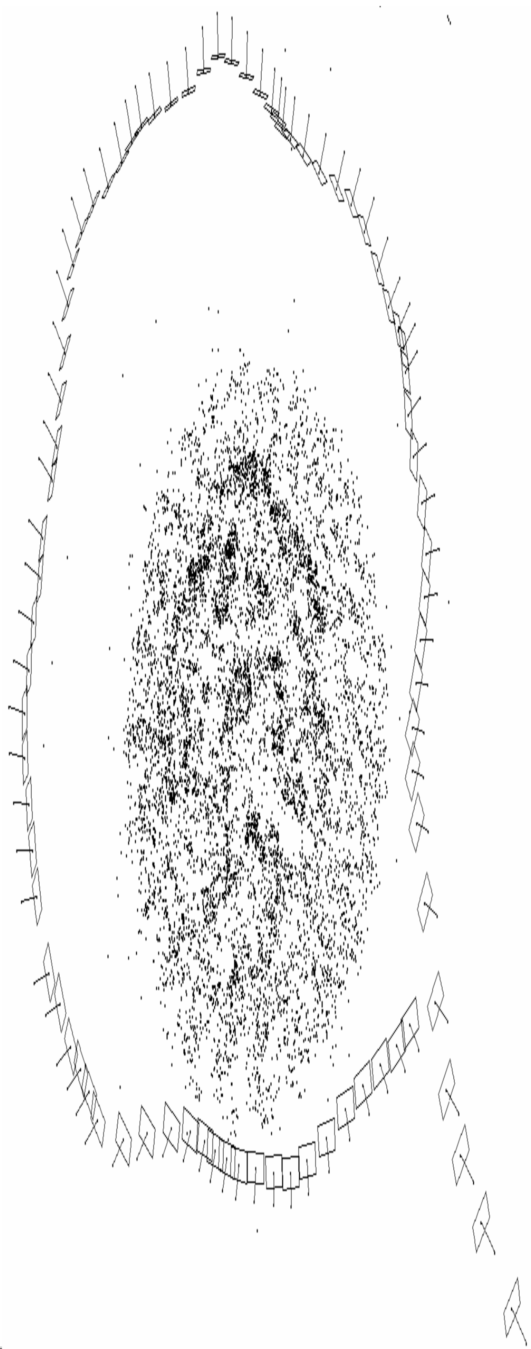


UK Center for Visualization & Virtual Environments

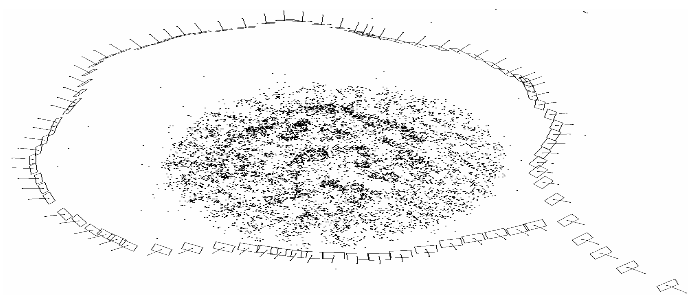


f_x / f_y

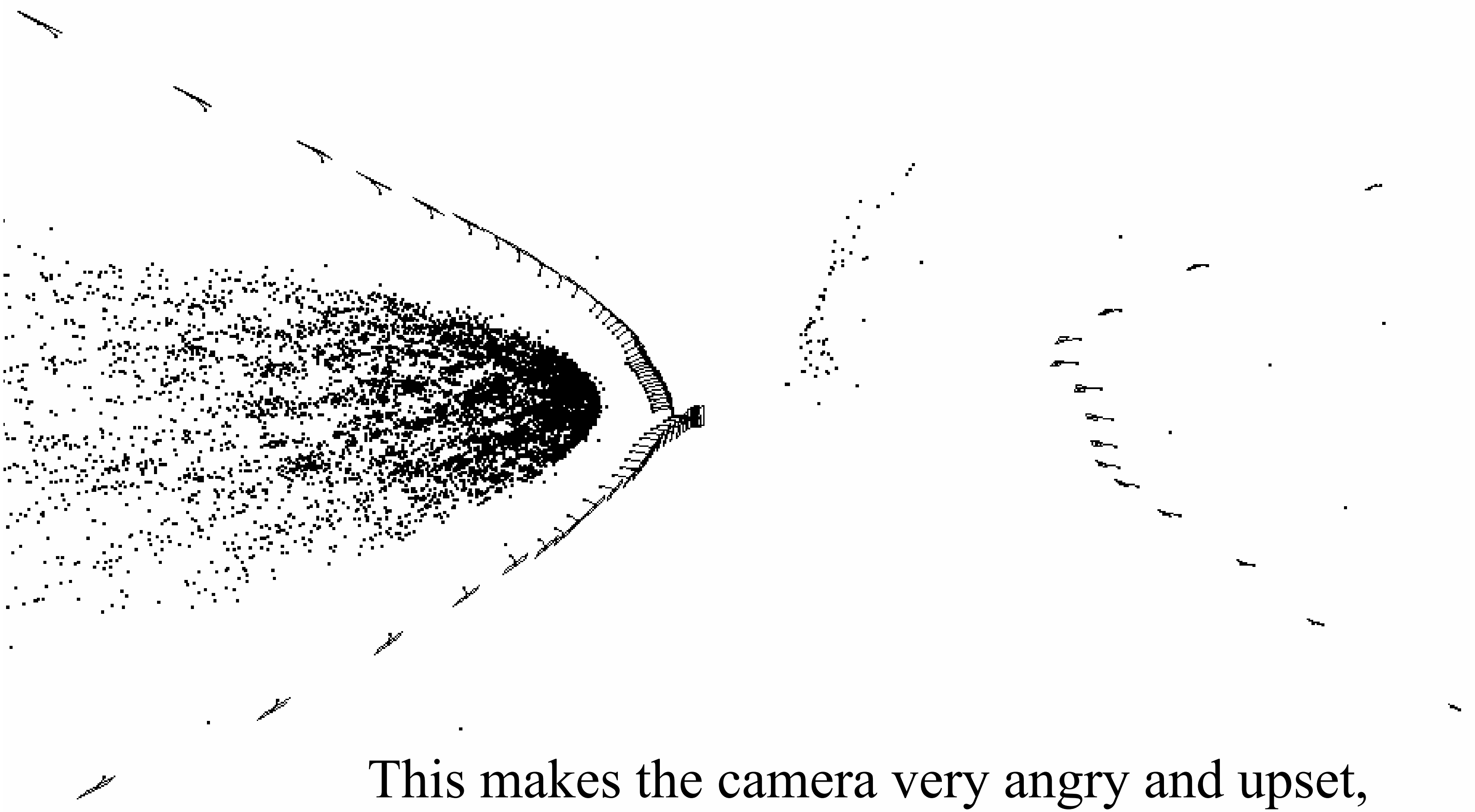




Distortion of the cameras is the cue that drives self-calibration



To move across the plane at infinity, a camera has to go through a ‘geometric wormhole’

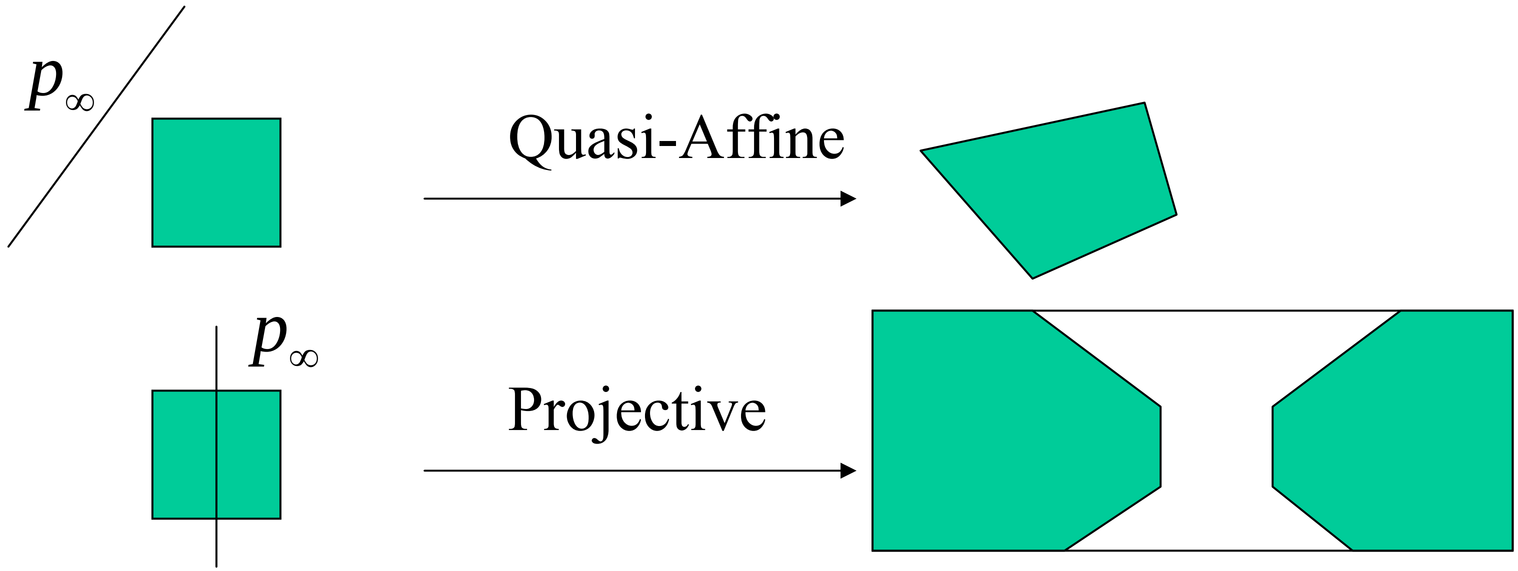


This makes the camera very angry and upset, in fact it will refuse

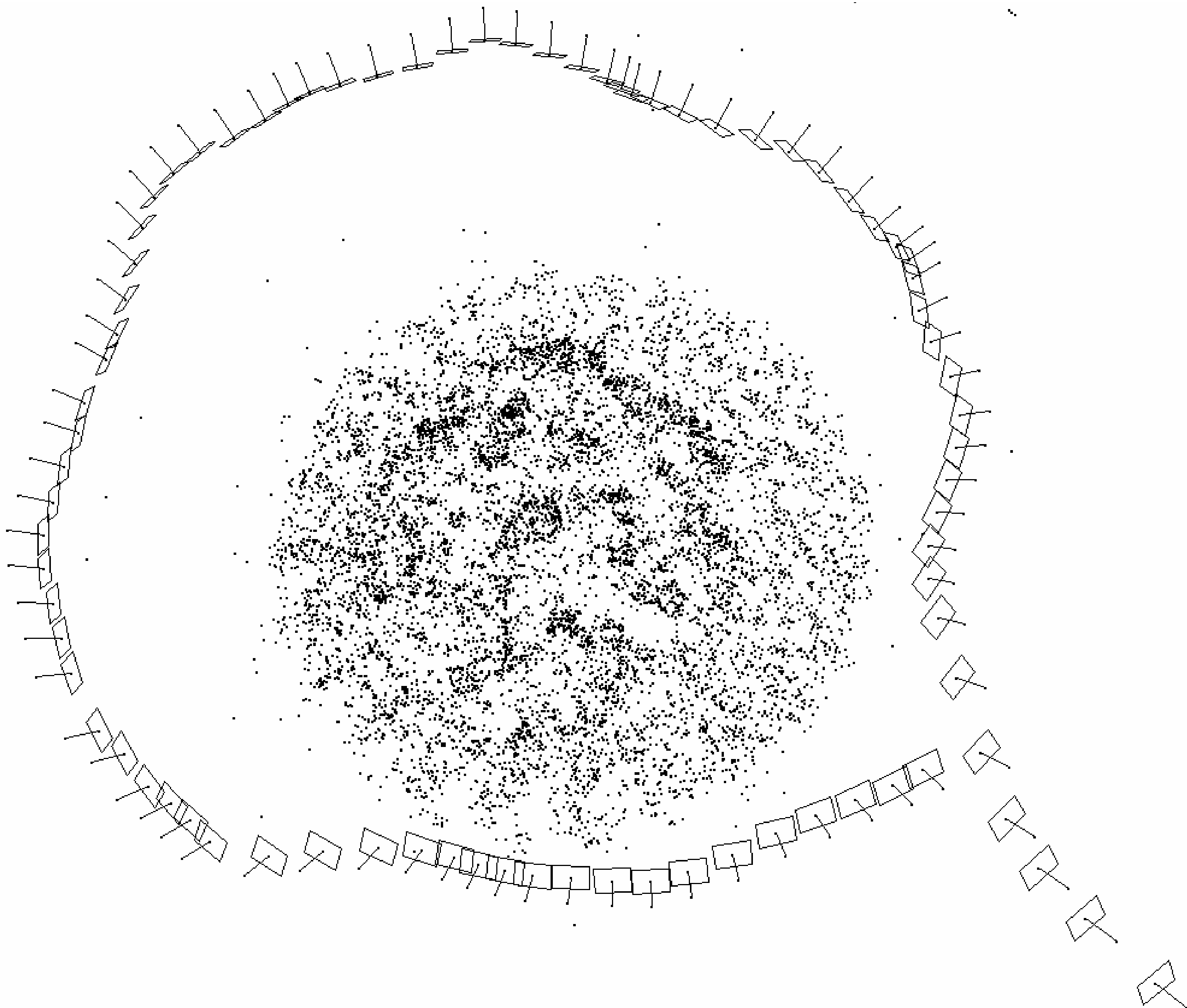
Quasi-affine transformations and cheirality

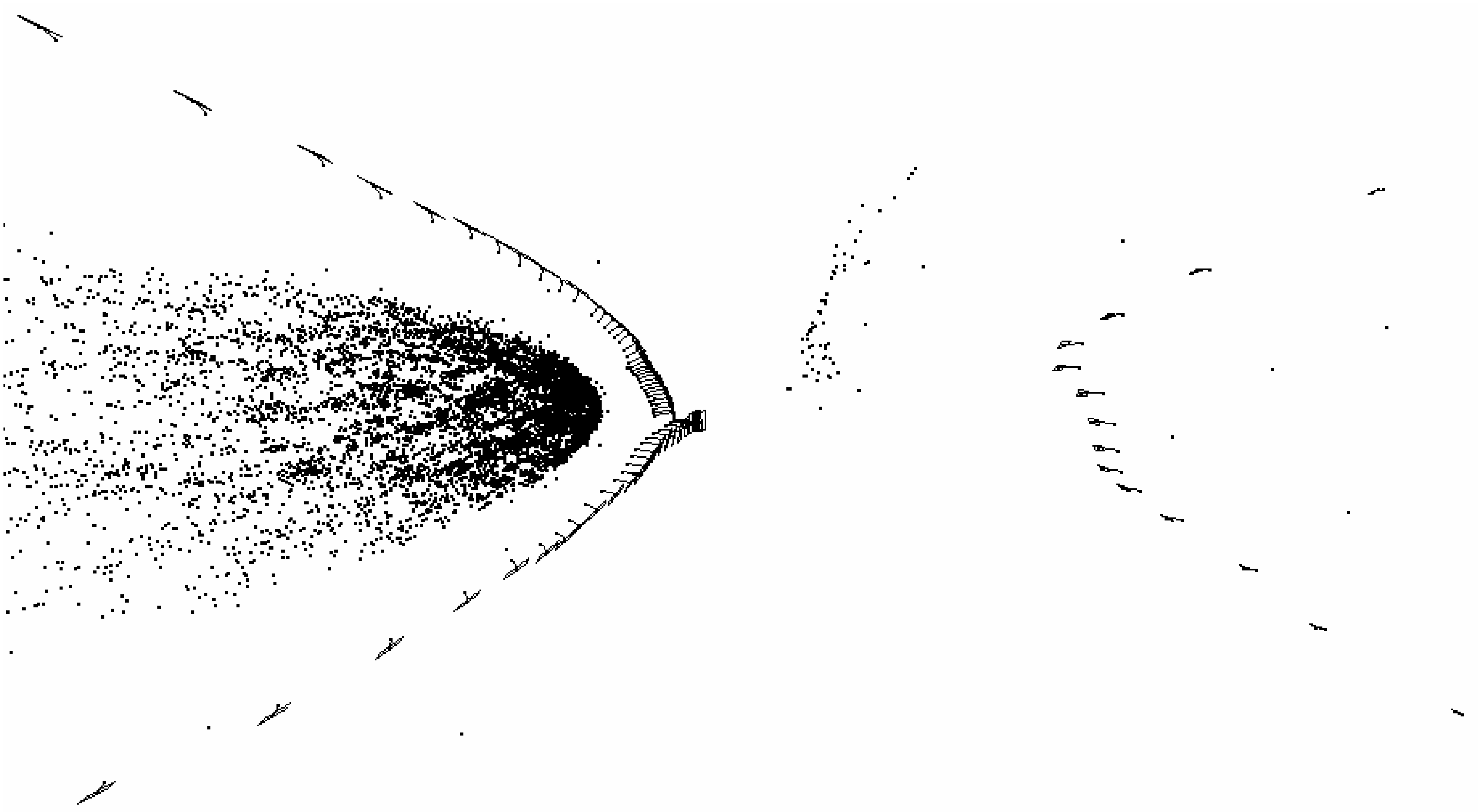
A projective transformation is quasi-affine with respect to a set iff it preserves the convex hull of the set

$$ConvexHull(H(A)) \overset{?}{=} H(ConvexHull(A))$$

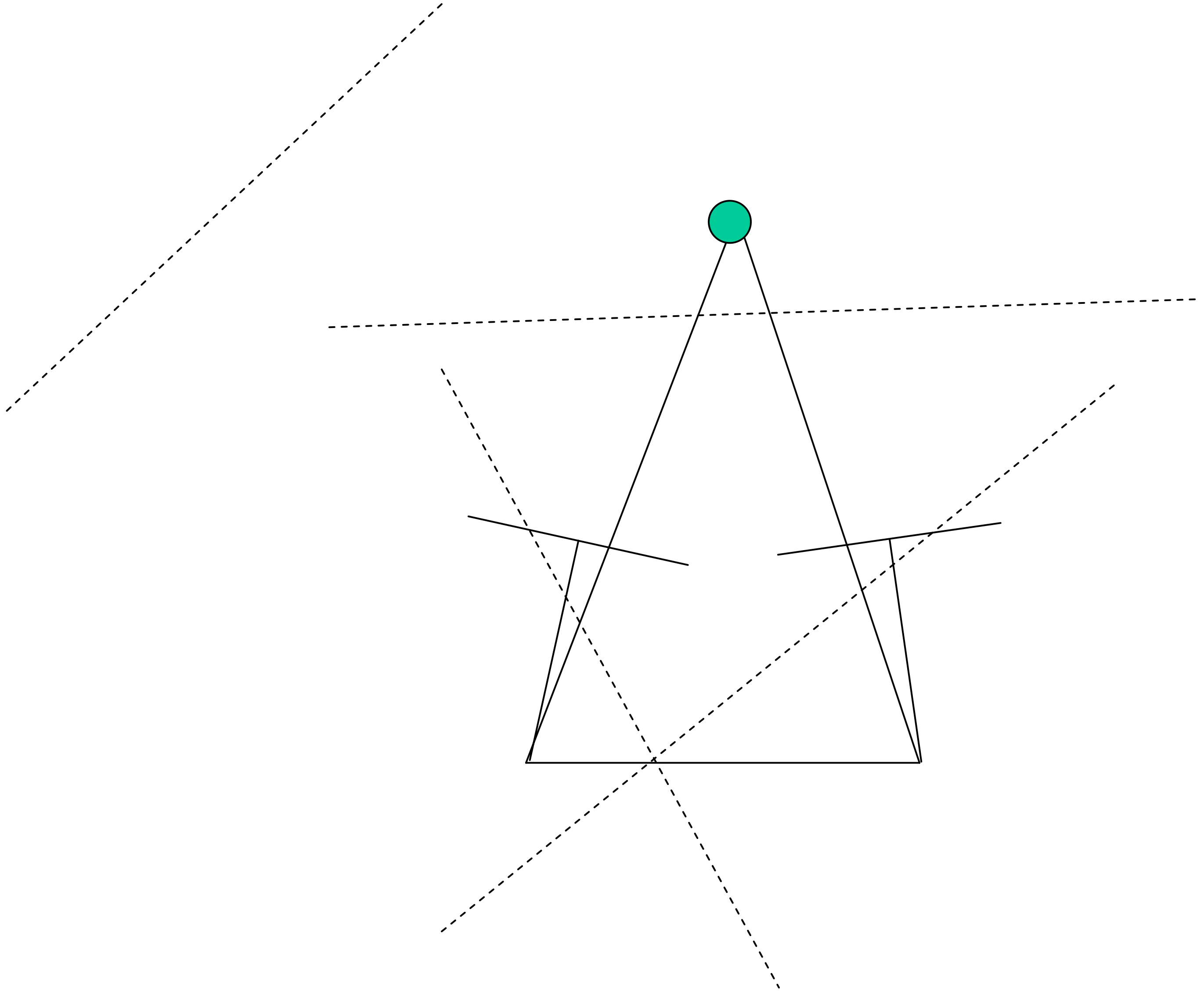


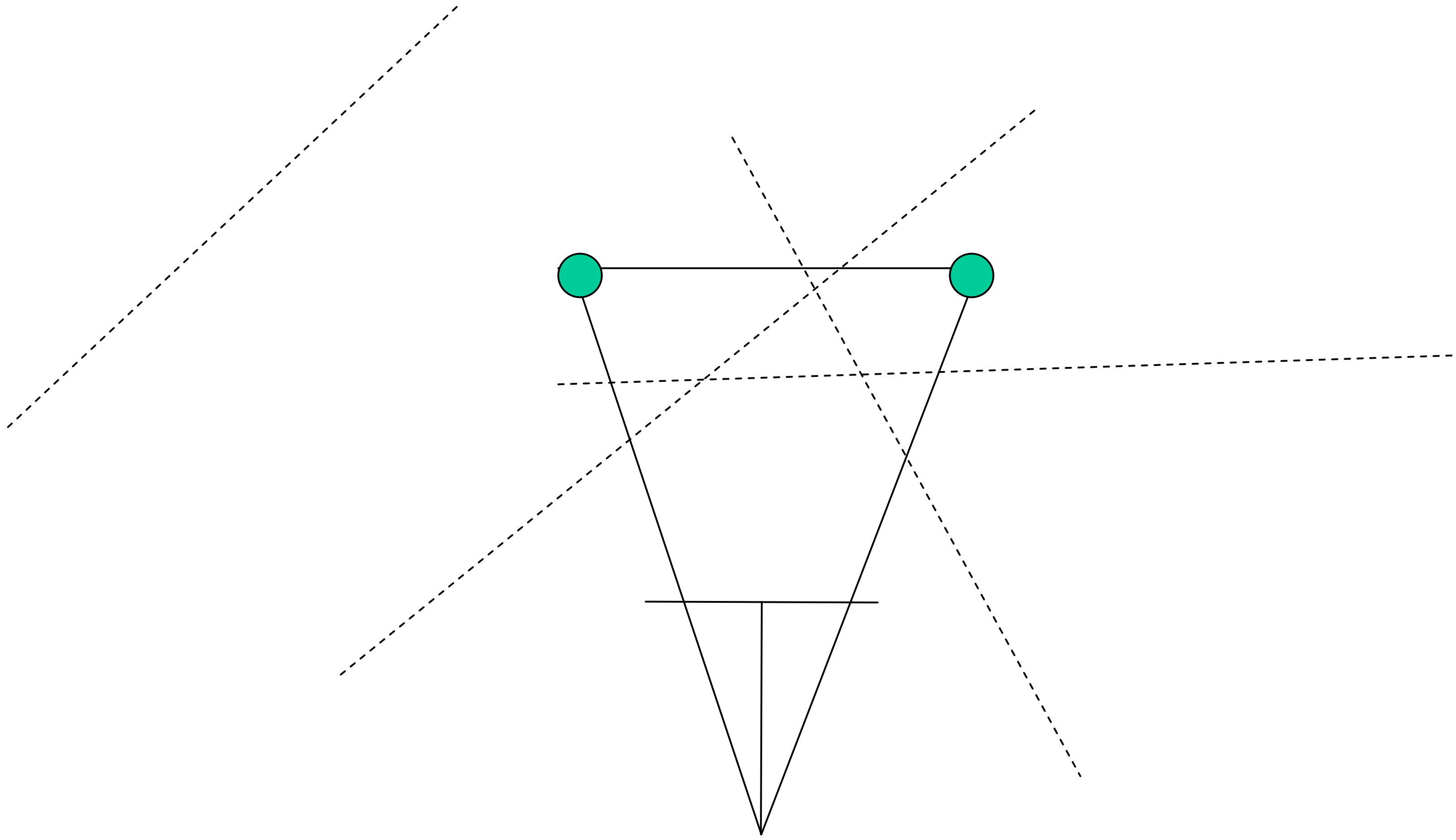
A projective transformation is affine iff
it is quasi-affine with respect to the set of all finite points



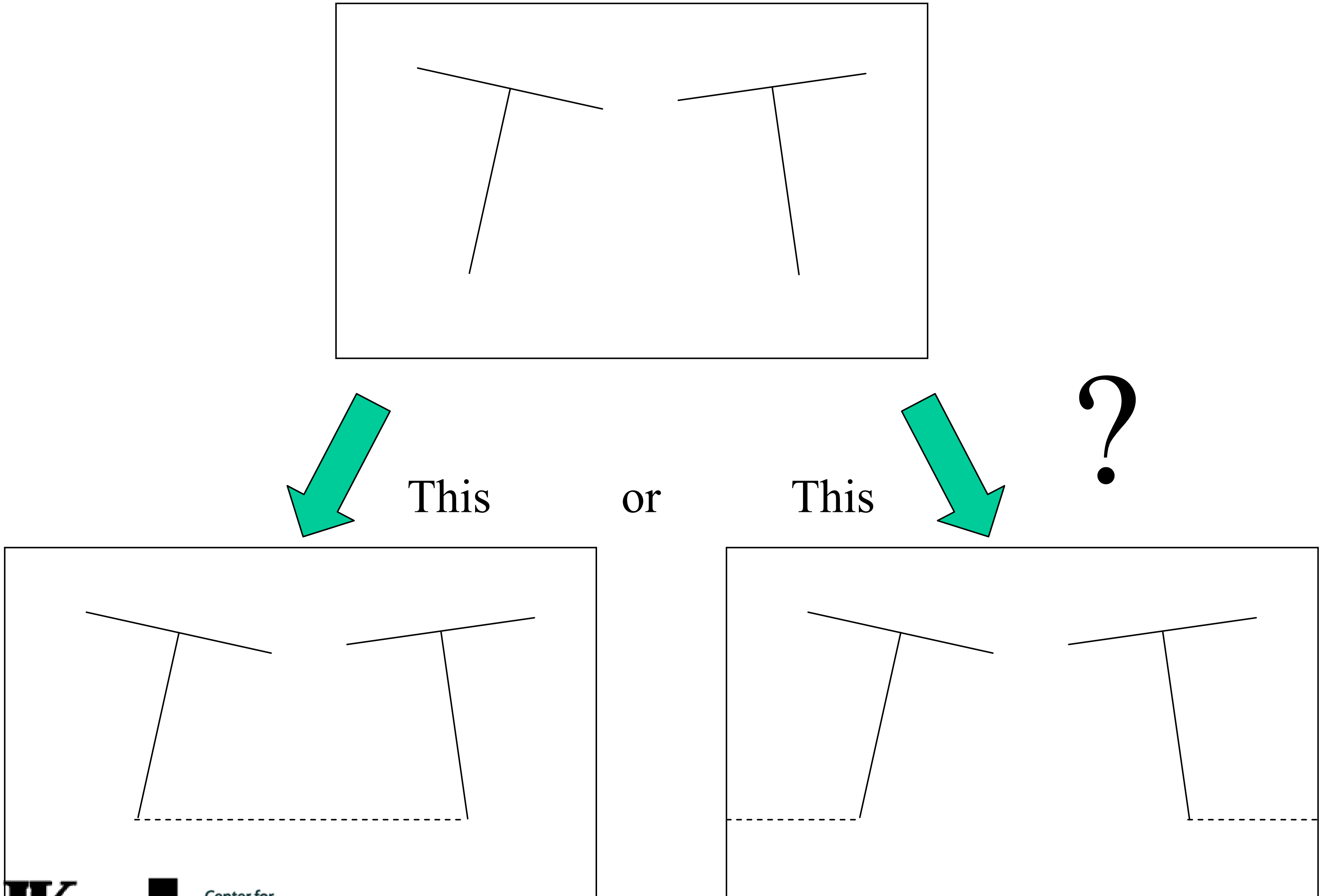


UK Center for Visualization & Virtual Environments



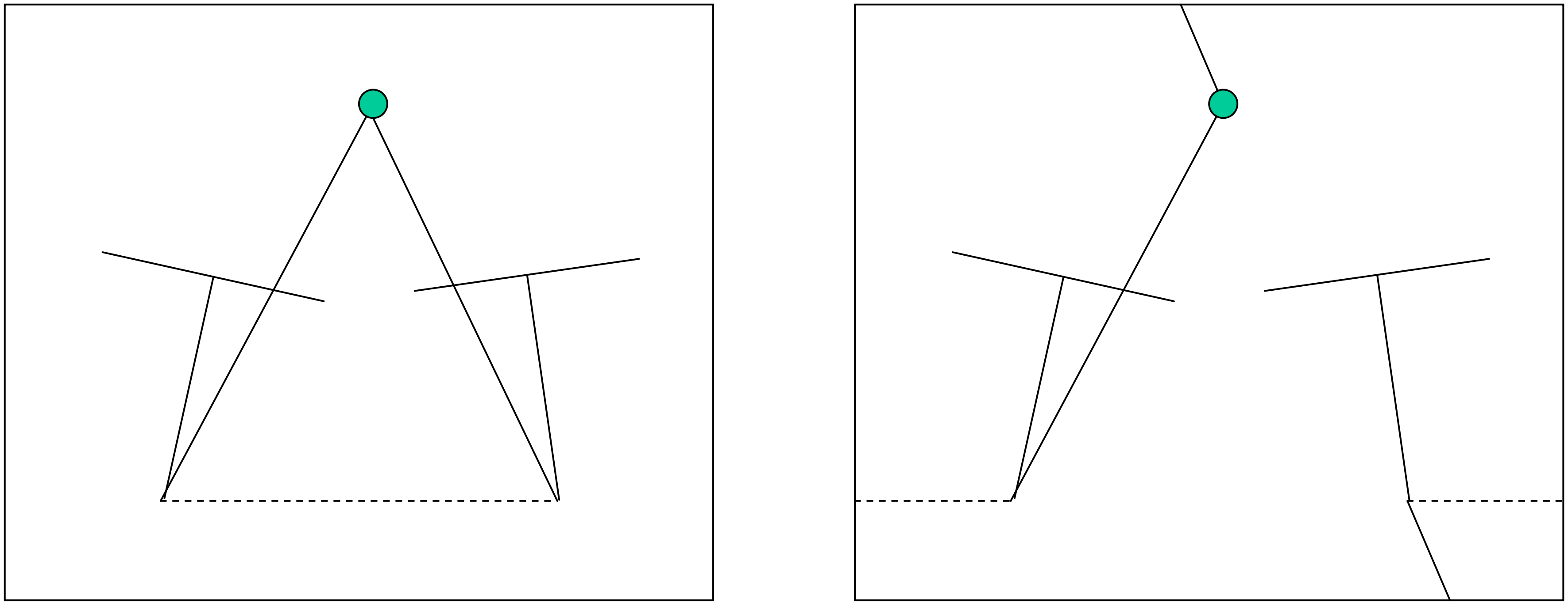


Each camera pair poses a question regarding the metric baseline



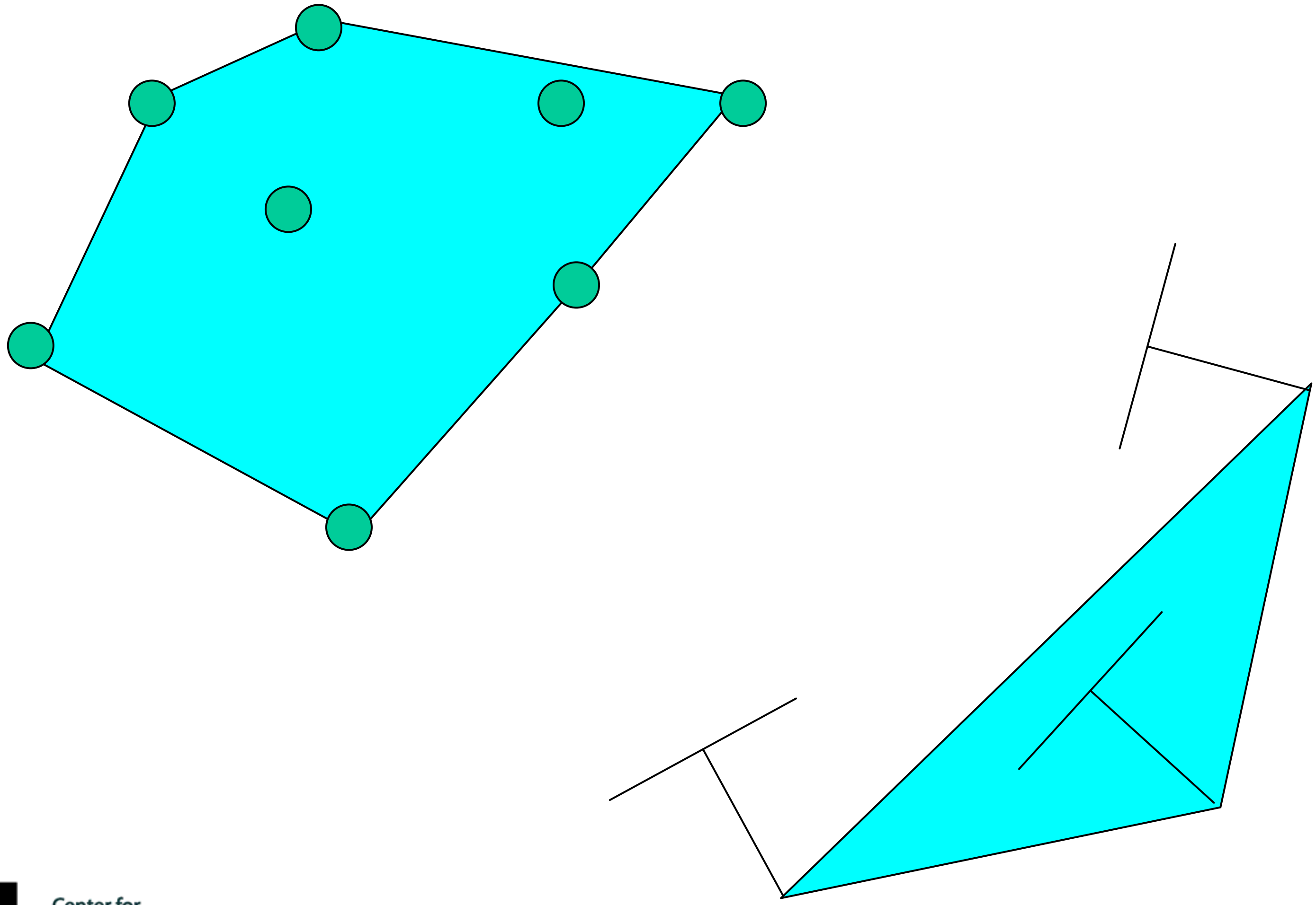
Center for Visualization & Virtual Environments

The question is easily answered by cheirality since a point in front of or behind both cameras supports the former case and a point on different sides supports the latter.



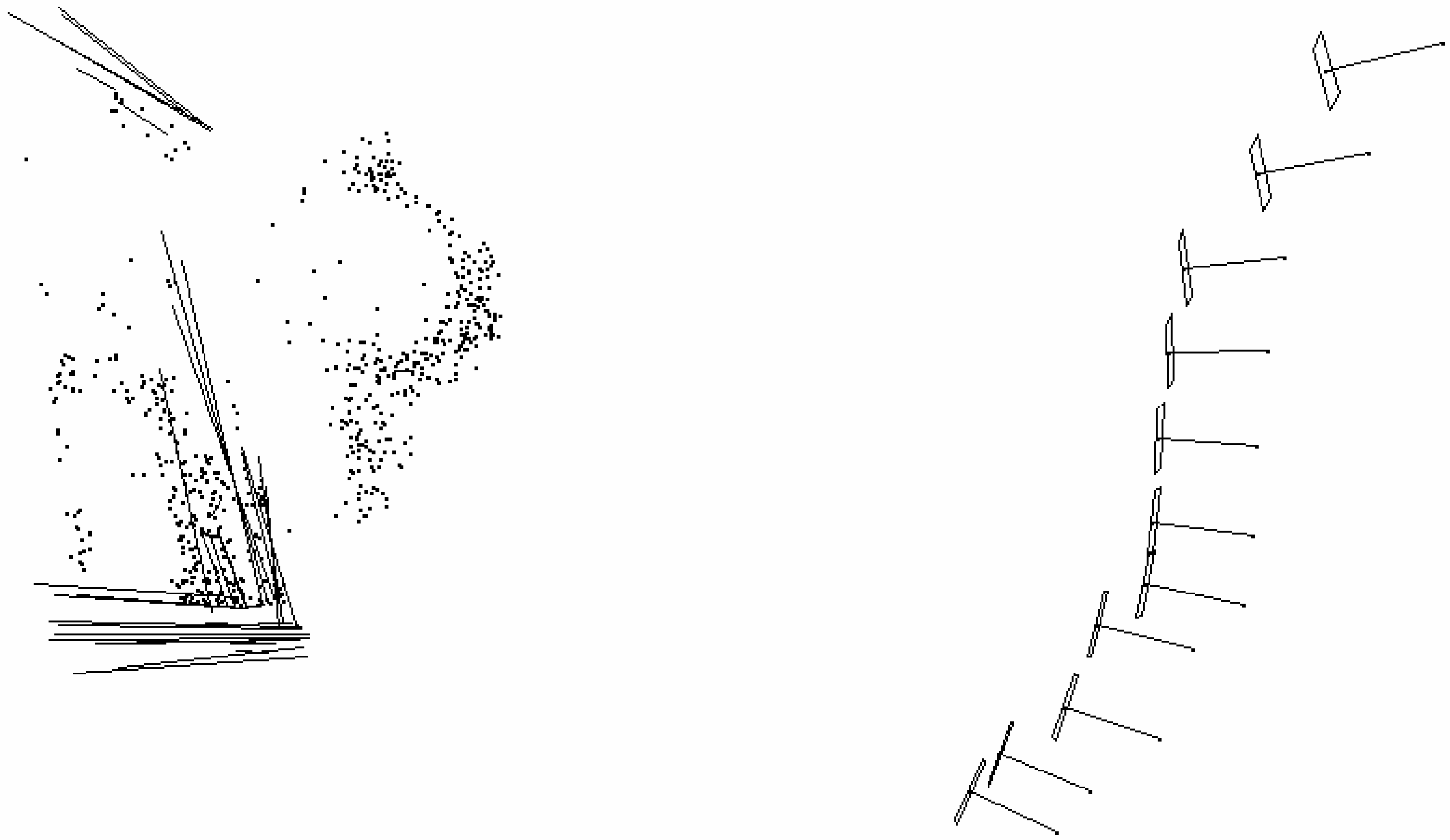
A sequence of such binary decisions then deduces the convex hull of the camera centres.

Using cheirality, the convex hull of the points and the convex hull of the cameras can be respected
(But not necessarily the convex hull of the union)

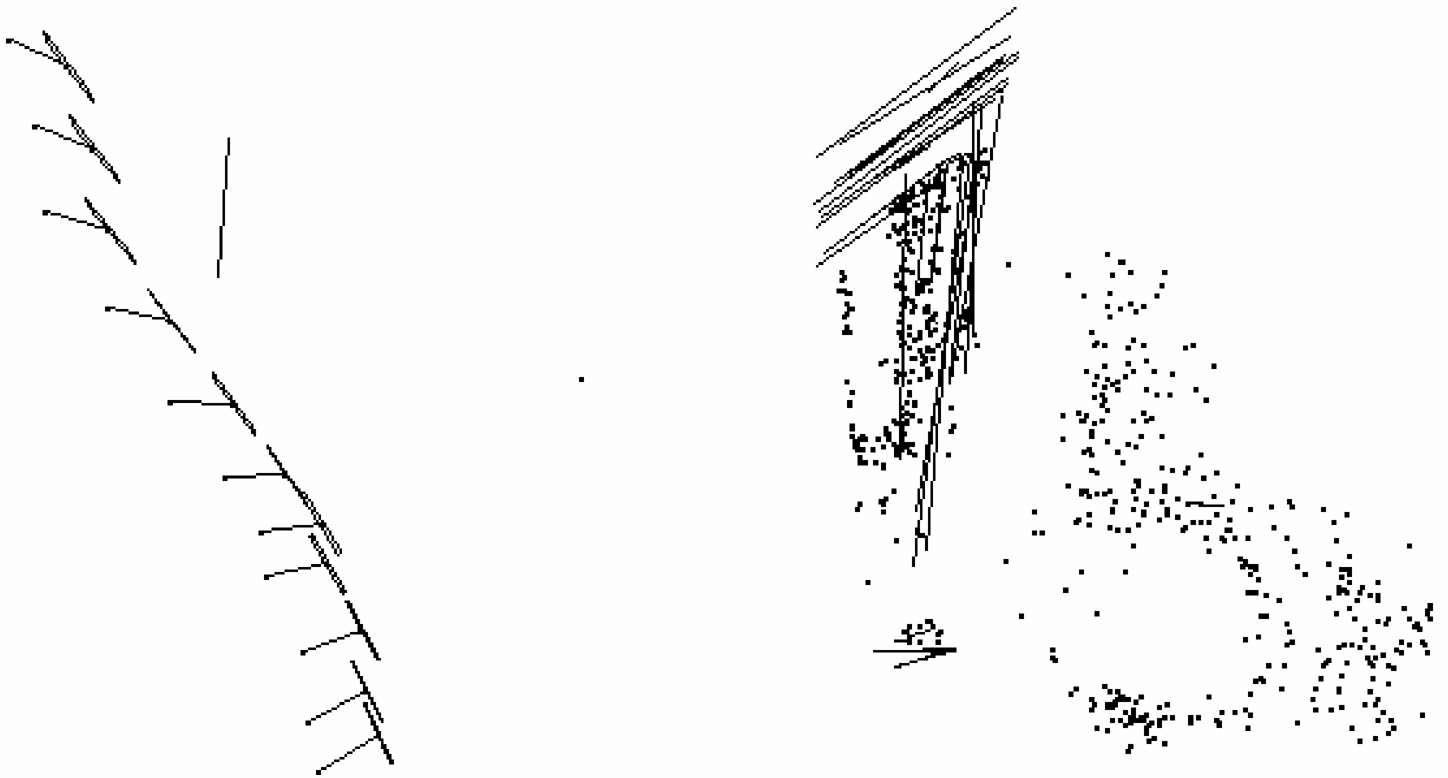


Center for Visualization & Virtual Environments

Metric configuration



Center for Visualization & Virtual Environments



Cheirality (QUARC reconstruction)

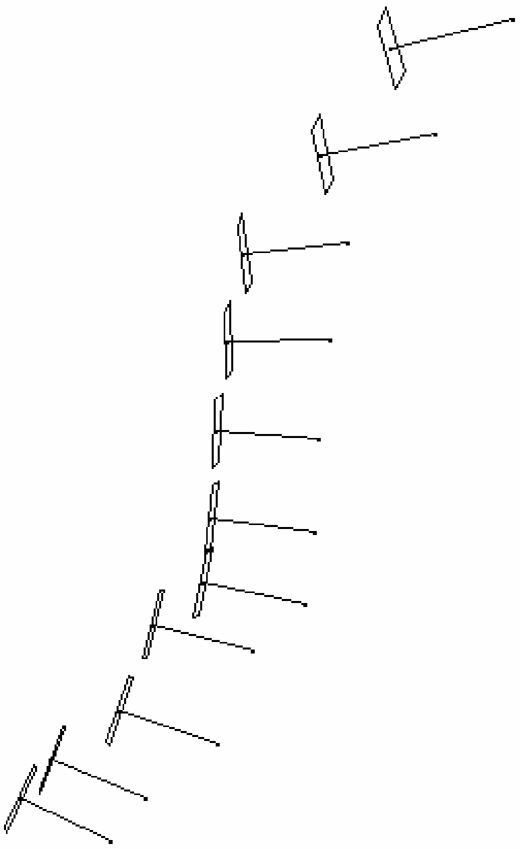
$$P_n H^{-1} \cong K_n \left[R_n \mid -R_n t_n \right]$$

$$K_n = \begin{bmatrix} k_1 & k_2 & k_3 \\ & k_4 & k_5 \\ & & 1 \end{bmatrix}$$

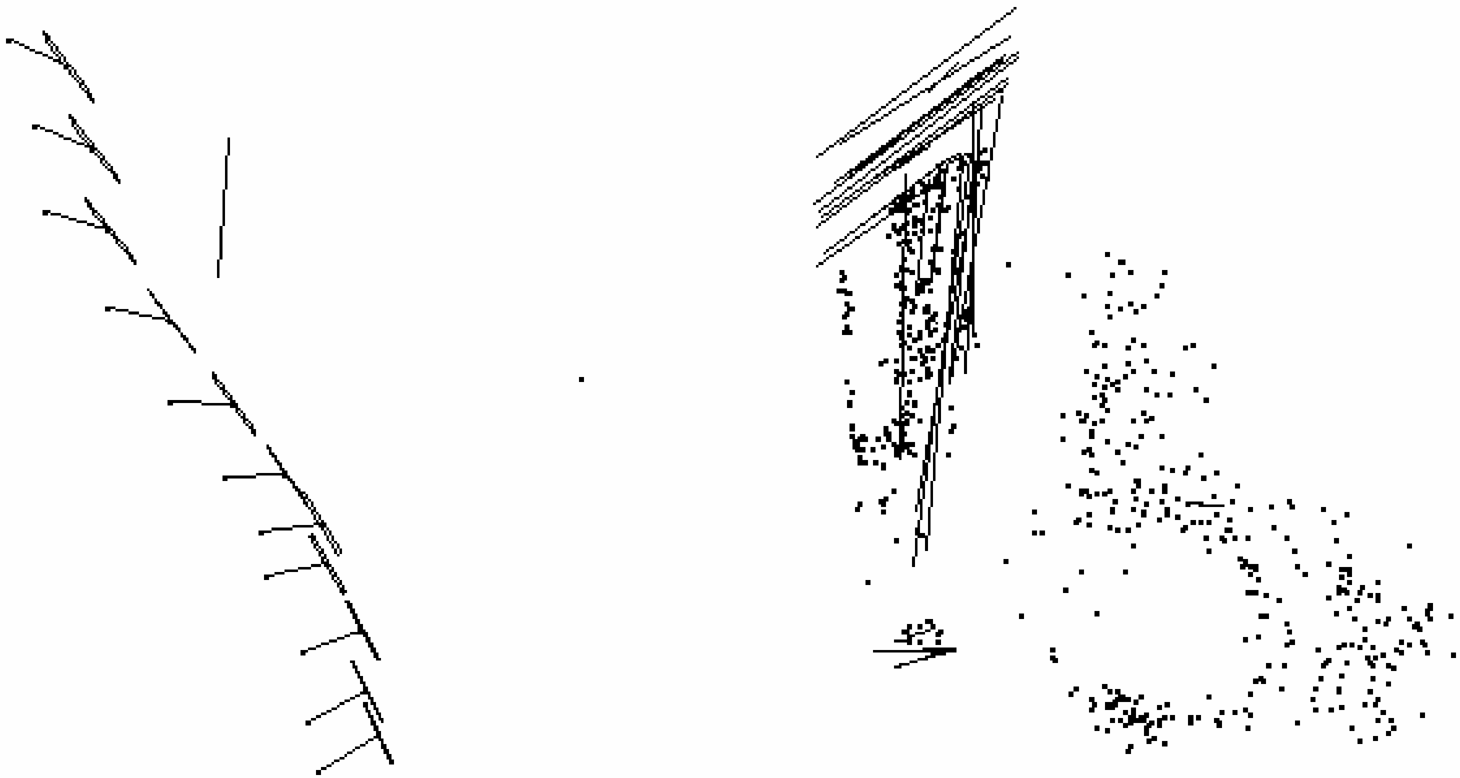
$$\left(\frac{k_2}{f}\right)^2 + \left(\frac{k_3}{f}\right)^2 + \left(\frac{k_5}{f}\right)^2 + \left(\frac{k_1 - k_4}{f}\right)^2, \; f = \frac{k_1 + k_4}{2}$$

Center for Visualization & Virtual Environments

Metric configuration



The points are not essential, convergence occurs even from this projective equivalent



Center for Visualization & Virtual Environments