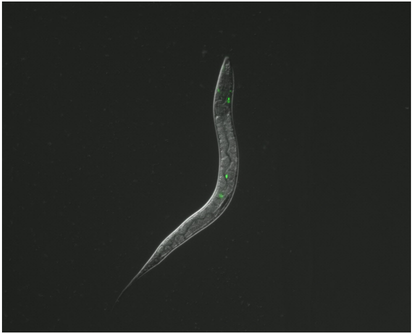
Analysis and Classification of C.Elegans in High-Throughput Experiments

Matthias Demant

Chair of Pattern Recognition and Image Processing,
Albert-Ludwigs-University Freiburg

Outline

C.Elegans



- C.elegans genome fully sequenced in December 1998
- 50-65 % of the currently known human genes have a homologue in the model organism
- Model organism for drug treatment (Alzheimer)
- Green Fluorescent Protein

Problem Statement

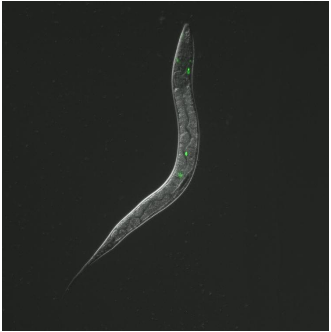


Figure: C.elegans with fluorescent CAN neurons

- CAN neurons develop in the head
- Migrate to the vulva

COPAS Sorter

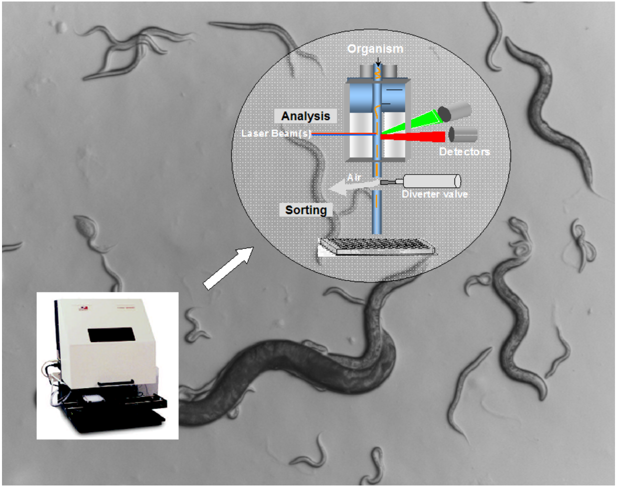


Figure: Workflow of the COPAS sorter

Problem Statement

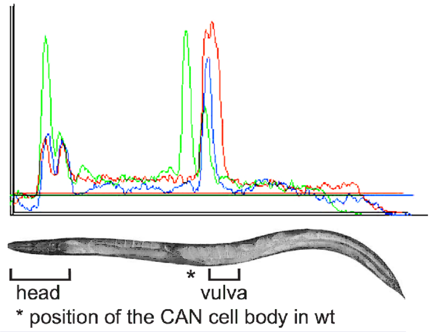


Figure: Exemplary fluorescent profiles

- Readout of COPAS sorter
- Peaks in the head and in the center

Scope of Pattern Recognition

1. Compare individual worm sequences
2. Description of a population
3. Comparison of populations
4. Classification of individual worm sequences

Scope of Pattern Recognition

1. Compare individual worm sequences
2. Description of a population
3. Comparison of populations
4. Classification of individual worm sequences

Scope of Pattern Recognition

1. Compare individual worm sequences
2. Description of a population
3. Comparison of populations
4. Classification of individual worm sequences

Scope of Pattern Recognition

1. Compare individual worm sequences
2. Description of a population
3. Comparison of populations
4. Classification of individual worm sequences

Outline

Outline

Euclidean Distance

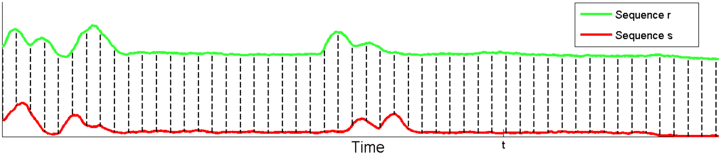


Figure: Euclidean distance measure

Euclidean distance Compare uniformly sampled elements

Disadvantage Small shift → completely different result

Dynamic Time Warping

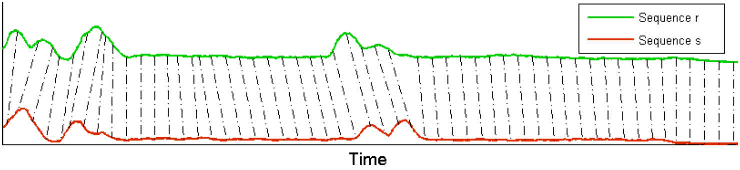


Figure: DTW approach

DTW distance Compare signals at corresponding points

Advantage Small shift → small increment of distance

Dynamic Time Warping

1. Local cost measure: Normalized cross-correlation of patches $\mathbf{s}_i$ and $\mathbf{r}_j$ centered at i, j with regularization term

$$\text{Dist}(i,j) = 1 - \frac{\left\langle \mathbf{s}_i - \mu_{\mathbf{s}_i}, \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\rangle}{\left\| \mathbf{s}_i - \mu_{\mathbf{s}_i} \right\| \cdot \left\| \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\| + \epsilon}$$

2. Search path through cost matrix with minimal costs
Ordering, boundary constraint

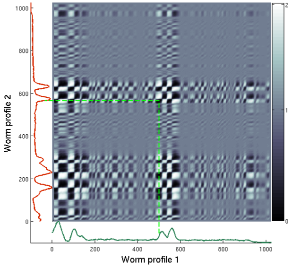


Figure: Distance matrix between the patches of the signals

Dynamic Time Warping

1. Local cost measure: Normalized cross-correlation of patches $\mathbf{s}_i$ and $\mathbf{r}_j$ centered at i, j with regularization term

$$\mathrm{Dist}(i,j) = 1 - \frac{\left\langle \mathbf{s}_i - \mu_{\mathbf{s}_i}, \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\rangle}{\left\| \mathbf{s}_i - \mu_{\mathbf{s}_i} \right\| \cdot \left\| \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\| + \epsilon}$$

2. Search path through cost matrix with minimal costs
   Ordering, boundary constraint

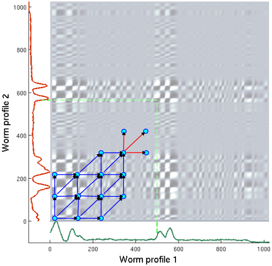


Figure: Path search within a trellis

Dynamic Time Warping

1. Local cost measure: Normalized cross-correlation of patches $\mathbf{s}_i$ and $\mathbf{r}_j$ centered at i, j with regularization term

$$\text{Dist}(i,j) = 1 - \frac{\left\langle \mathbf{s}_i - \mu_{\mathbf{s}_i}, \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\rangle}{\left\| \mathbf{s}_i - \mu_{\mathbf{s}_i} \right\| \cdot \left\| \mathbf{r}_j - \mu_{\mathbf{r}_j} \right\| + \epsilon}$$

2. Search path through cost matrix with minimal costs
   Ordering, boundary constraint

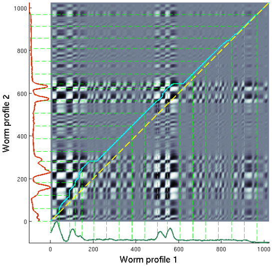


Figure: Path with minimum costs

DTW and Time-Delayed DTW

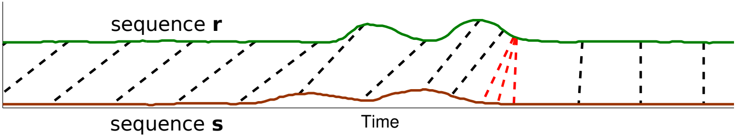


Figure: One-to-many alignment

- DTW may align an element to a segment
- Viterbi algorithm can be extended on second order terms or refined with an open snake.

Time-delayed Dynamic Time Warping

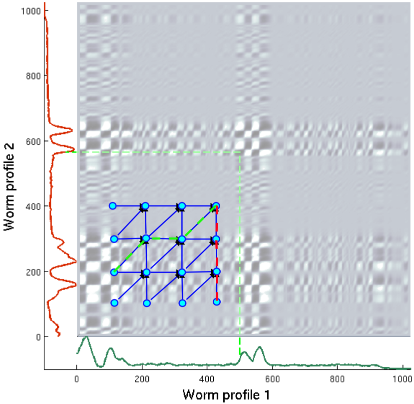


Figure: Path search within a trellis with a time-delayed decision

Time-delayed Dynamic Time Warping

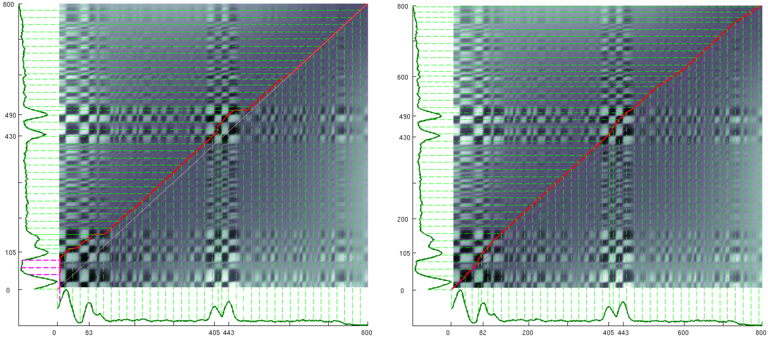
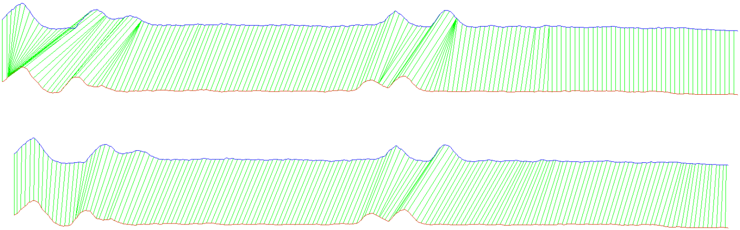


Figure: DTW and refined DTW

DTW and Time-Delayed DTW

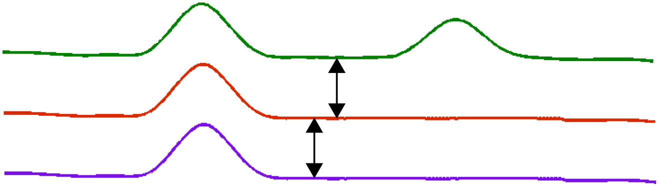


- DTW extended on second order terms

⇒ Smooth alignment

Distance measure and Noise

- Accumulated costs along warp path
- Problem:



- Deformation as similarity measure.
- Low variance as indicator for noise.

⇒ Weighting and penalizing of correlation results.

- Penalize pathes with little signal to signal matches
- Weight deformation with minimum signal level

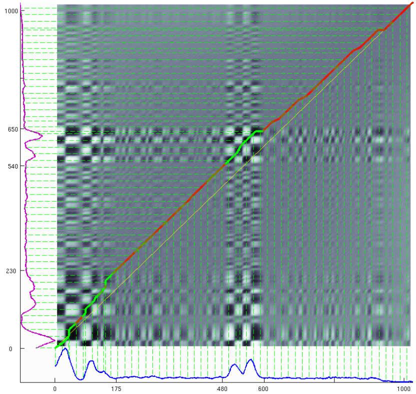
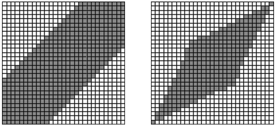


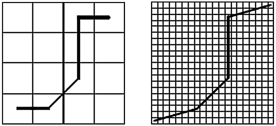
Figure: Signals and the expected noise value along the warp path.

Speeding up DTW

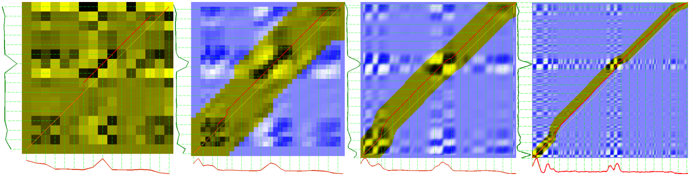
- Runtime DTW: $O(n^2)$
- Evaluate less cells



- Compute path at lower resolution and project onto finer resolution.



- Multiscale DTW



Outline

Gabor Wavelets

- Multiplication of Gaussian with a complex exponential

$$f(x) = \underbrace{\exp(-i\mu_0(x - x_0))}_{\text{complex exponential}} \underbrace{\exp(-\frac{(x - x_0)^2}{2\sigma^2})}_{\text{Gaussian}}$$

- Expand patches in frequency domain.
- Resolution in spatial and frequency domain.
- Multiresolution analysis with self-similar family of Gabor wavelets.

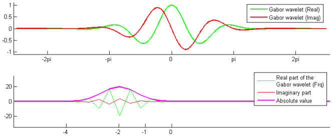


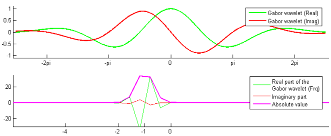
Figure: Gabor filter in spatial and frequency domain.

Gabor Wavelets

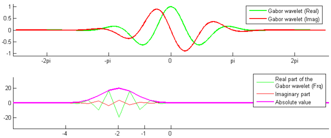
- Multiplication of Gaussian with a complex exponential

$$f(x) = \underbrace{\exp(-i\mu_0(x - x_0))}_{\text{complex exponential}} \underbrace{\exp(-\frac{(x - x_0)^2}{2\sigma^2})}_{\text{Gaussian}}$$

- Expand patches in frequency domain.
- Resolution in spatial and frequency domain.
- Multiresolution analysis with self-similar family of Gabor wavelets.



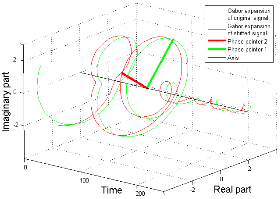
(a) Gabor filter (lower frq)



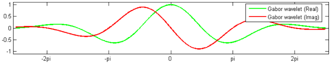
(b) Gabor filter (higher frq)

Phase Shift and Gabor Wavelets

- Displacement between signals ⇒ phase shift
- Increase displacement:
  Smooth phase shift
- Different effect on different Gabor features



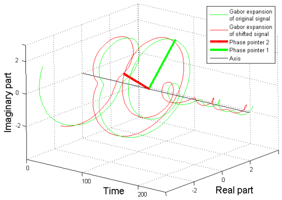
(c) Signals expanded (lower frq)



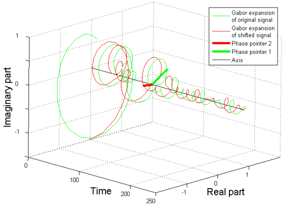
(e) Gabor filter (lower frq)

Phase Shift and Gabor Wavelets

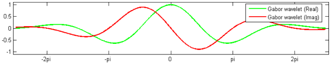
- Displacement between signals ⇒ phase shift
- Increase displacement:
  Smooth phase shift
- Different effect on different Gabor features



(e) Signals expanded (lower frq)



(f) Signals expanded (higher frq)



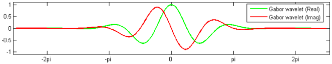


Figure: Gabor wavelets and phase shifts

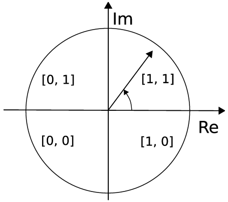
Distance Measure in Gabor Feature Space

- Encoding and demodulation of signal s at scale k

$$\mathrm{h}^k_{\{\mathrm{Re},\mathrm{Im}\}}(t) = \mathrm{sgn}_{\{\mathrm{Re},\mathrm{Im}\}} \int_x s(x-t) e^{-i(k\omega)(x-t)} e^{\frac{-(x-t)^2}{2(\sigma/k)^2}} dx$$
$$= \mathrm{sgn}_{\{\mathrm{Re},\mathrm{Im}\}} (s * f_k)(t)$$

- $\mathrm{h}^k_{\{\mathrm{Re},\mathrm{Im}\}}$ is a complex valued bit sequence.
- Bit sequences at different scale ⇒ Code to describe a worm
- Compare sequence codes using the Hamming distance:

$$\mathrm{HD}_{\mathrm{worm}} = \|(codeA \otimes codeB)\|$$



(a) Quadrant Demodulation Code

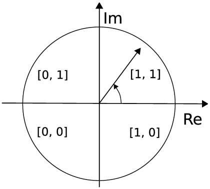
Distance Measure in Gabor Feature Space

- Encoding and demodulation of signal s at scale k

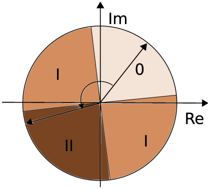
$$\mathrm{h}^k_{\{\mathrm{Re,Im}\}}(t) = \mathrm{sgn}_{\{\mathrm{Re,Im}\}} \int_x s(x-t) e^{-i(k\omega)(x-t)} e^{\frac{-(x-t)^2}{2(\sigma/k)^2}} dx$$
$$= \mathrm{sgn}_{\{\mathrm{Re,Im}\}} (s * f_k)(t)$$

- $\mathrm{h}^k_{\{\mathrm{Re,Im}\}}$ is a complex valued bit sequence.
- Bit sequences at different scale ⇒ Code to describe a worm
- Compare sequence codes using the Hamming distance:

$$\mathrm{HD}_{\mathrm{worm}} = \|(codeA \otimes codeB)\|$$



(g) Quadrant Demodulation Code



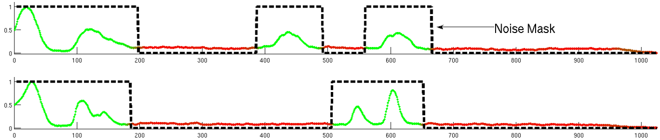
(h) Cosine difference

Comparing Bit Sequences with Noise Handling

- Exclude noise patches

⇒ Fractional Hamming distance

$$\mathrm{HD}_{\mathrm{worm}} = \frac{\|(codeA \otimes codeB) \bigcap (maskA \bigcup maskB)\|}{\|maskA \bigcup maskB\|} \qquad (1)$$



(i) Two masks created with the assumed noise model.

Outline

Hierarchical Clustering

- Group sequences in a tree structure
- Initialization: Each sequence is a cluster
- Merge sequences with the distances of the DTW and a linkage function:

  Nearest neighbor, average distance, Ward's variance criteria

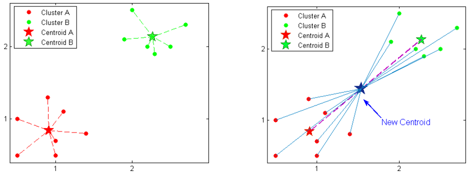


Figure: By merging two groups the centroid changes. Ward's linkage merges clusters with the lowest increment of variance.

Self-Organizing Maps - Motivation

Goal Quantitative description of population

- Population consist of different subgroups
- Continuous transitions between subgroups

⇒ Self-Organizing Maps

Self-Organizing Maps - Motivation

Goal Quantitative description of population

- Population consist of different subgroups
- Continuous transitions between subgroups

⇒ Self-Organizing Maps

Self-Organizing Maps - Motivation

Goal Quantitative description of population

- Population consist of different subgroups
- Continuous transitions between subgroups

⇒ Self-Organizing Maps

Self-Organizing Maps - Structure

- SOM consists of neurons n_k.
- Connected to model vectors $\mathbf{m}_k$ and to input vectors.
- During the matching process the BMU is detected.
- Activation of neuron depends on distance to the BMU.
- Update model vector according to activation of connected neuron.

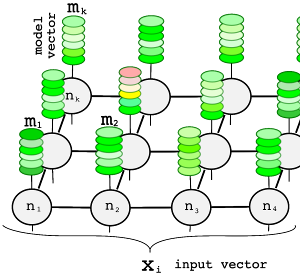


Figure: Model of a Self Organizing map.

Self-Organizing Maps - Structure

- SOM consists of neurons n_k.
- Connected to model vectors $\mathbf{m}_k$ and to input vectors.
- During the matching process the BMU is detected.
- Activation of neuron depends on distance to the BMU.
- Update model vector according to activation of connected neuron.

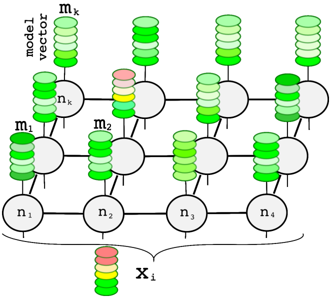


Figure: Model of a Self Organizing map.

Self-Organizing Maps - Structure

- SOM consists of neurons n_k.
- Connected to model vectors $\mathbf{m}_k$ and to input vectors.
- During the matching process the BMU is detected.
- Activation of neuron depends on distance to the BMU.
- Update model vector according to activation of connected neuron.

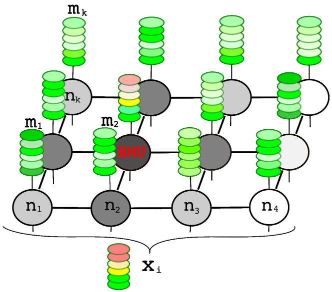


Figure: Model of a Self Organizing map.

Self-Organizing Maps - Learning

Initialization Model vectors = random sequences

Matching Compute position of BMU n_b: $\mathbf{r}_b = (x, y)^T$

$$\mathbf{r}_b = \text{argmin}_{\mathbf{r}_k} \left\{ \text{dist}\left(\mathbf{x}_i, \mathbf{m}_k\right) \right\} \tag{2}$$

Update

$$\mathbf{m}_k^{(t+1)} \leftarrow \mathbf{m}_k^{(t)} + h_{bk}(t) \left\| \mathbf{x}_i - \mathbf{m}_k^{(t)} \right\| \tag{3}$$

$h_{bk}(t)$ is the "neighborhood" function.

Activation

$$h_{bk}(t) = \alpha(t) \cdot \underbrace{\exp\left(- \frac{\|\mathbf{r}_b - \mathbf{r}_k\|}{2\sigma^2(t)} \right)}_{\text{Gaussian centered at BMU}} \tag{4}$$

$\alpha(t)$ returns a learning rate $\alpha(t) \in [0,\, 1]$ at time step t.
$\sigma(t)$ implies the width of the Gaussian kernel.

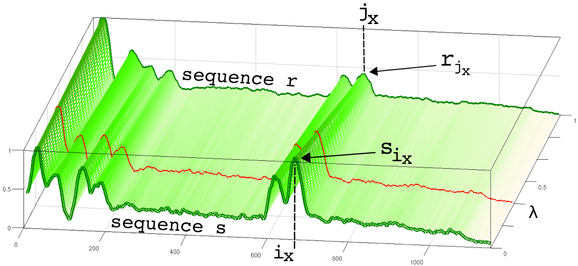
SOMs and DTW

- Update process requires weighted average.
- Registration between **s** and **r**: $s_{i_x} \leftrightarrow r_{j_x}$

⇒ Morphed model vector.

$$w_x = (1 - \lambda) \cdot s_{i_x} + \lambda \cdot r_{j_x}$$
$$t_x = (1 - \lambda) \cdot i_x + \lambda \cdot j_x$$

- $\lambda \in [0, 1]$ warping factor
- **t** are sampling instances of the weighted average **w**. $f(t_i) := w_i$ describes the morphed signal. Interpolate f at uniformly scaled sampling points.



Comparing Populations

1. Learn SOM on all worm sequences of all populations

$\Rightarrow$ Prototypes

2. Quantification of each population by histogram over SOM codebook
3. Comparison of histograms:

$$D(i,j) = \frac{\text{hist}_{\text{popA}}(i,j)}{\sum_{i,j} \text{hist}_{\text{popA}}(i,j)} - \frac{\text{hist}_{\text{popB}}(i,j)}{\sum_{i,j} \text{hist}_{\text{popB}}(i,j)}$$

$\Rightarrow$ Typical differences

Outline

Exemplary Color Assignment

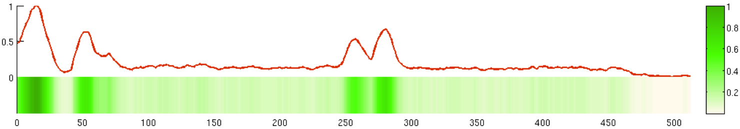


Figure: An example for the color assignment of a worm sequence to its image illustration.

COPAS Data

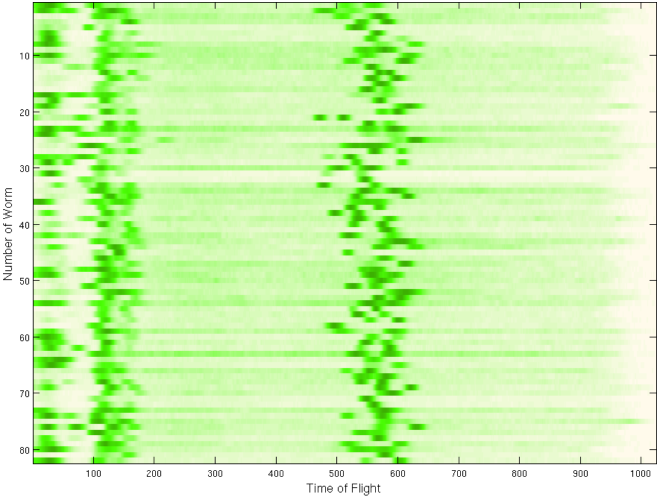


Figure: Wild type (82 worms)

COPAS Data

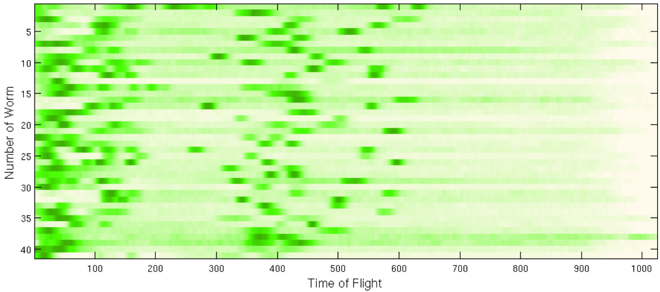
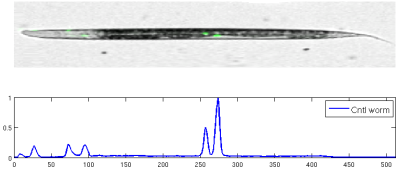
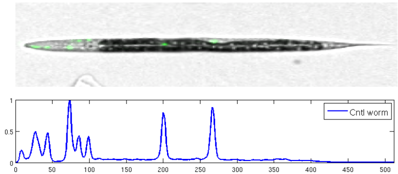


Figure: Mutants (41 worms)

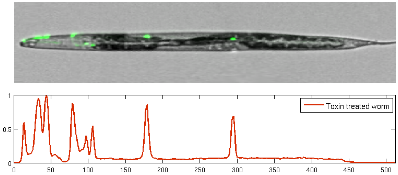
Microscopic Data



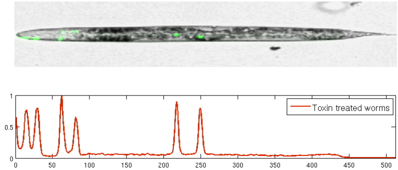
(a) Control worm



(b) Control worm



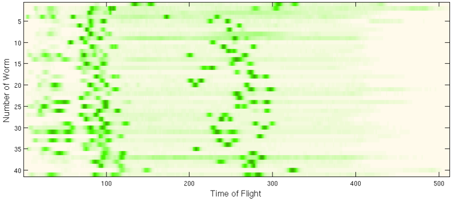
(c) Toxin treated worm



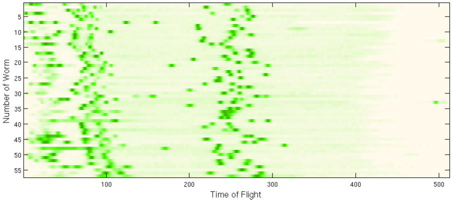
(d) Toxin treated worm

Figure: Toxin treated and control worms

Microscopic Data



(a) 41 Control worms



(b) 56 Toxin treated worms

Figure: Toxin treated worms and control worms

Microscopic and COPAS Data

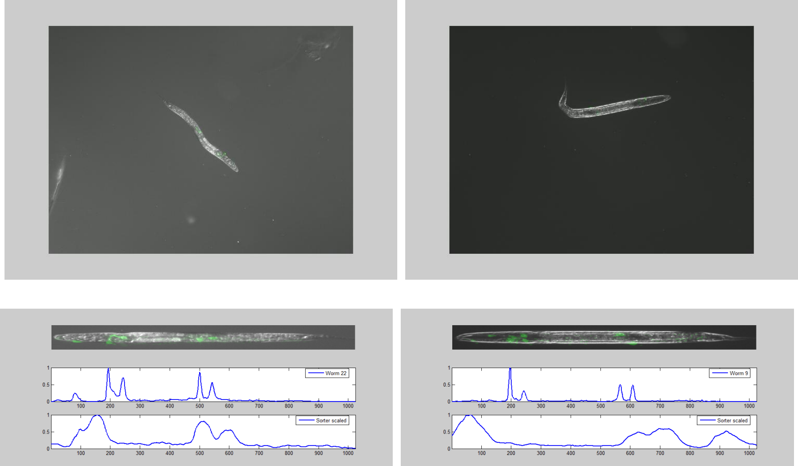


Figure: Top-down: original image, images of segmented and aligned worm, the extracted GFP sequence and the corresponding COPAS sorter result.

Microscopic and COPAS Data

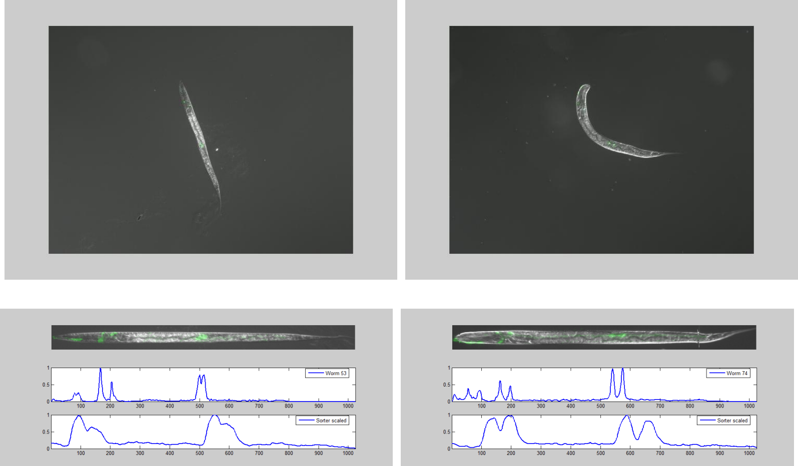
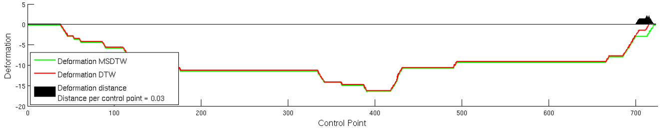


Figure: Top-down: original image, images of the segmented and aligned worm, the extracted GFP sequence and the corresponding COPAS sorter result.

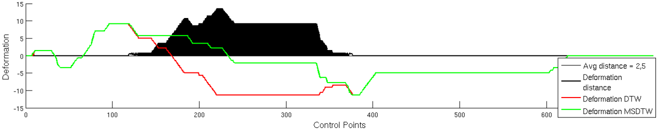
Outline

MSDTW and DTW

Quality metric: Deformation difference



(a) Most of the deformation models differ insignificantly.

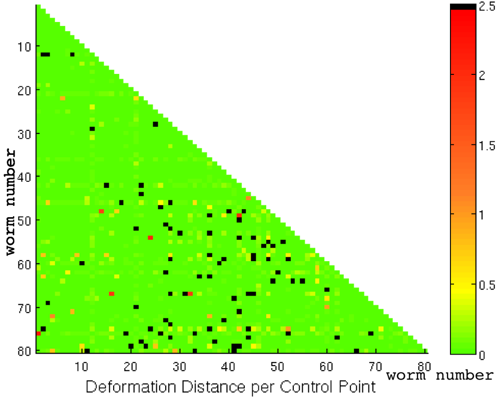


(b) An outlier with strong deformation differences.

Figure: The deformation models are plotted in red and green.

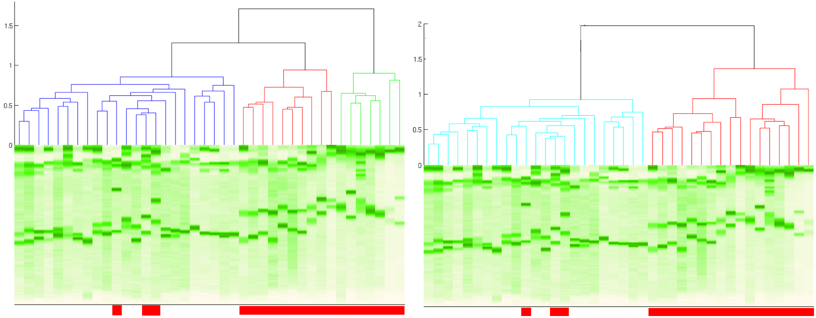
- Deformation and deformation difference between the MSDTW and the DTW.
- Black area indicates the absolute difference.

MSDTW and DTW



- Deformation distance between the MSDTW and DTW matrix of 80 worms
- Black points: 85 outliers with an average deformation $\geq$ 2.5 pixel

Clustering



(a) Correlation along the warp path. (b) Penalizing pathes with a low signal to signal relation

Left Clustering with summed up correlation along warp path

Right Penalizing pathes with little signal elements

⇒ Distance of the clusters increases

Clustering - Example

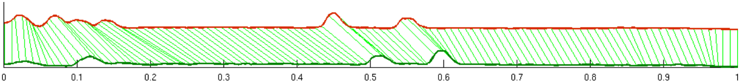
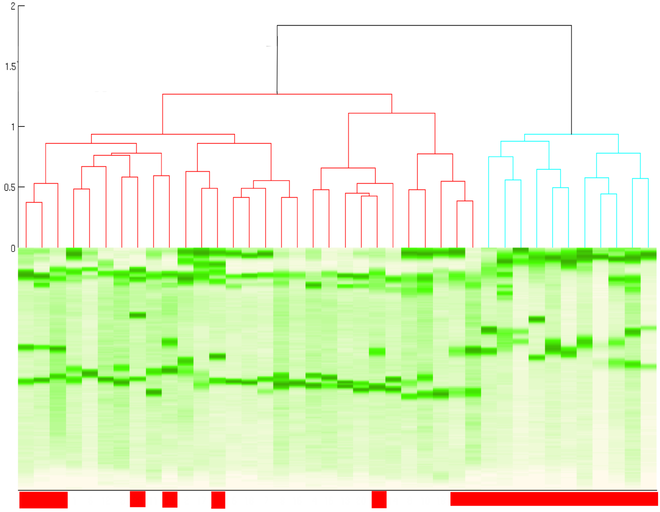


Figure: Wild type and mutant signals

Fast Comparison - Clustering



Self-Organizing Maps

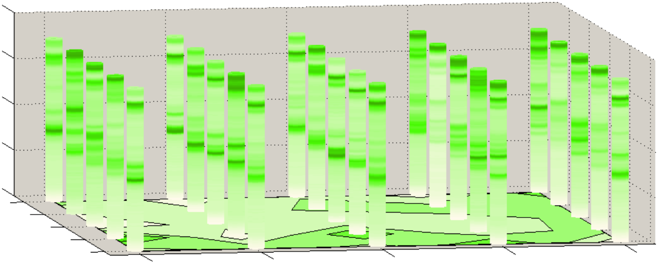
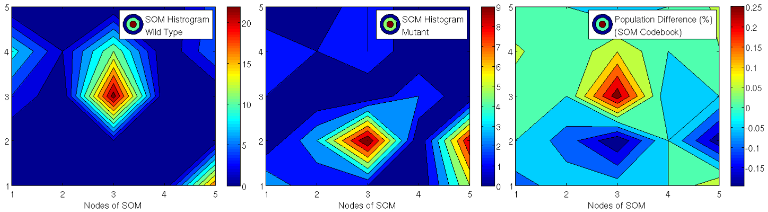


Figure: 5 × 5 SOM

- SOM after 500 iterations
- Cylindric objects represent the model vectors
- Ground plot → How often BMU.

Comparing Populations



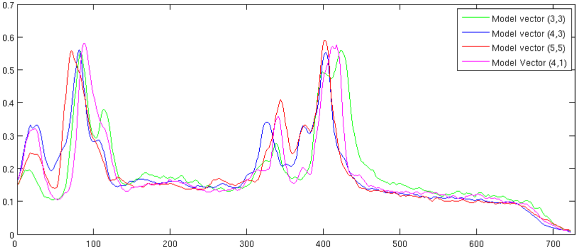
(a) Histogram of wild type population

(b) Histogram of mutant population

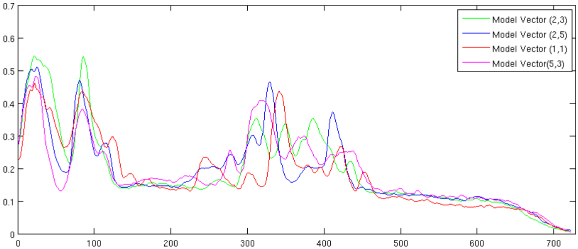
(c) Histogram differences

- Quantification of the populations according to the SOM codebook.
- Each element of a population is assigned to its best matching unit (BMU) on the SOM.
- Difference of normalized histograms (right).
- Preferred areas are visible.

Comparing Populations



(d) Prototypical sequences of the wild type population



(e) Prototypical sequences of the mutant population

Outline

Conclusion

DTW	Considers shape of the profiles and alignment Time-delayed DTW ⇒ excellent registration and similarity results
MSDTW	MSDTW yields nearly same results. Adequate: Long sequences with weak deformations.
Cluster	Grouping from coarse to fine structure differences DTW distance measure ⇒ intuitive groups
SOM	SOM combined with DTW to model a sparse representation of all populations Trying to enforce a global topological order ⇒ Quality of prototypes decreased SOM could partially model the two worm populations
DTW	Runtime $O(n^2)$, 0.8 seconds with $n = 512$
Gabor	80 Worms $n = 1024$ Quadrant: 9 sec Cosine: 22 sec Noise: 43 sec

Outlook

SOM Incorrect registration leads to artefacts

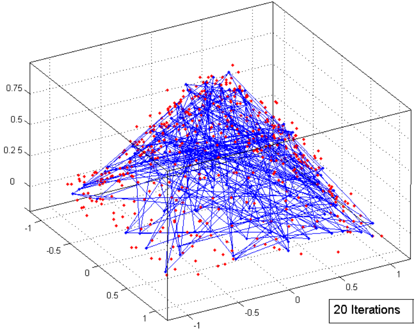
Evaluation on huge datasets

COPAS Improve quality of sorter data.

Thank you for your attention.

Self-Organizing Maps - Example

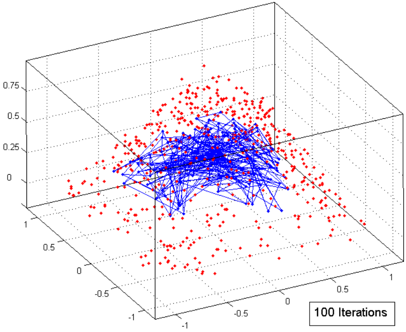
Figure: The SOM was initialized with random data values. It appears like a 'haystack'



A bell-shape was formed with 20000 data points. Some of them are illustrated in the red points. They were added with Gaussian noise. The blue lines indicate a SOM with its neighborhood relation. The SOM was created with 12×12 neurons and an Euclidean distance measure.

Self-Organizing Maps - Example

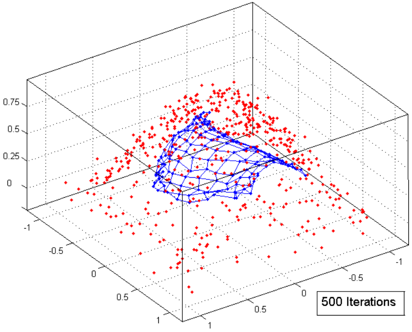
Figure: After 100 Iterations. The SOM learns fast within a huge neighborhood.



A bell-shape was formed with 20000 data points. Some of them are illustrated in the red points. They were added with Gaussian noise. The blue lines indicate a SOM with its neighborhood relation. The SOM was created with 12×12 neurons and an Euclidean distance measure.

Self-Organizing Maps - Example

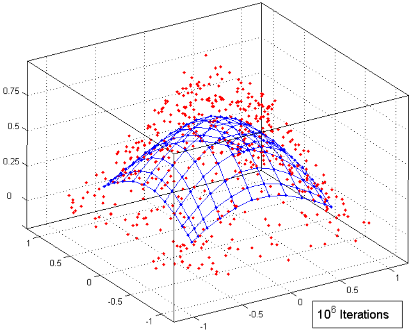
Figure: After 500 Iterations. The topology of the data gets visible.



A bell-shape was formed with 20000 data points. Some of them are illustrated in the red points. They were added with Gaussian noise. The blue lines indicate a SOM with its neighborhood relation. The SOM was created with 12×12 neurons and an Euclidean distance measure.

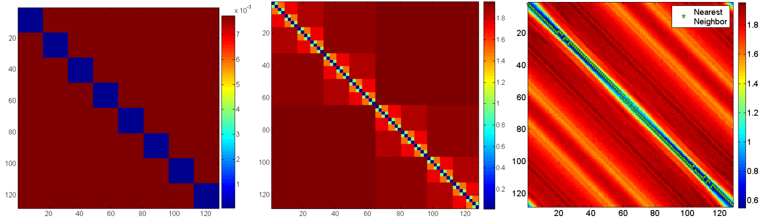
Self-Organizing Maps - Example

Figure: After 10^6 iterations the SOM is in the refinement stage. The topology of the bell was nearly reconstructed.



A bell-shape was formed with 20000 data points. Some of them are illustrated in the red points. They were added with Gaussian noise. The blue lines indicate a SOM with its neighborhood relation. The SOM was created with 12×12 neurons and an Euclidean distance measure.

Fast Comparison - Shift invariance



(a) Average Feature (b) Haar Feature (c) Gabor Feature (Cosine Difference)

Figure: Applying the feature based methods onto a shifted delta impulse. The illustrated similarity matrices show that only the results of the Gabor feature comparisons (c and d) are invariant to a shift of the signals.