# Convex Optimization of the Sammon Transformation

## Thesis Introduction

Susanne Westphal
12/16/2013
Computer Science Dept. 5 (Pattern Recognition)
Friedrich-Alexander University Erlangen-Nuremberg

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

## Convex Optimization of the Sammon Transformation

- The Sammon Transformation
- Test data sets
- Lagrange Multipliers
- Following steps

# The Sammon Transformation

## Sammon Transformation

- In 1969 John Sammon published an article about a non linear mapping for data structure analysis
- It is a mapping from a high-dimensional space to a lower-dimensional space
- The inner point distances of the points are preserved as good as possible over the transformation
- The Stress Function is an indicator for how big the difference of the inner point distances in the different spaces is
- For finding the best fitting points in the low dimensional space we have to minimize this equation.

# Sammon Stress Function:

$$E = \frac{1}{\sum_{i<j} d_{ij}} \sum_{i<j}^{N} \frac{(d_{ij} - ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2)^2}{d_{ij}}$$

$d_{ij}$ are the inner point distances in the original space.
$\boldsymbol{x}_i, \boldsymbol{x}_j$ are the projected points in the low-dimensional space.

# Properties of the Sammon Mapping

# Properties of the Sammon Mapping

- Preserves the grouping of the points

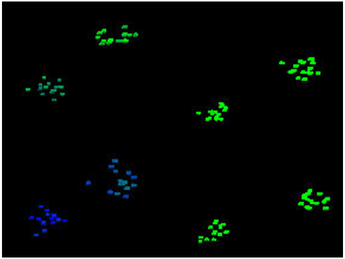
# Properties of the Sammon Mapping

- Preserves the grouping of the points
- Preserves the overall structure of the points

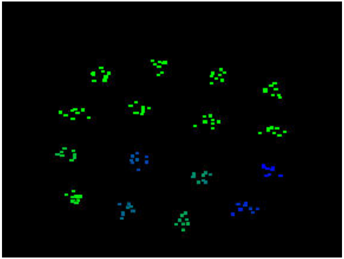
# Test data sets

# Hypercube

Original points in the 3D-space

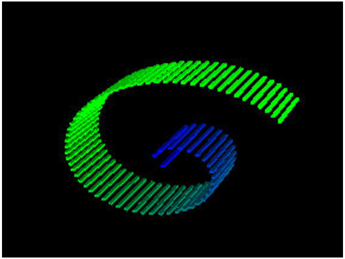


A good solution would be:

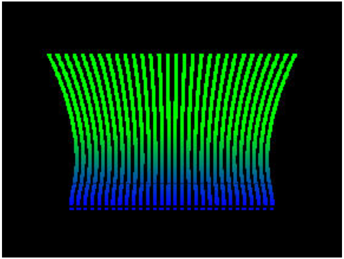


# Swiss Roll

Original points in the 3D-space



A good solution would be:



# Lagrange Multipliers

# Definition

$$\begin{array}{lll} \text{minimize} & f_0(\boldsymbol{x}) & \\ \text{subject to} & f_i(\boldsymbol{x}) \leq 0\ , & i = 1, ..., m; \\ & h_i(\boldsymbol{x}) = 0\ , & i = 1, ..., p; \end{array}$$

# Definition

$$\begin{array}{lll} \text{minimize} & f_0(\boldsymbol{x}) & \\ \text{subject to} & f_i(\boldsymbol{x}) \leq 0\ , & i = 1, ..., m; \\ & h_i(\boldsymbol{x}) = 0\ , & i = 1, ..., p; \end{array}$$

The Lagrangian is defined as:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) + \sum_{i=1}^{p} \nu_i h_i(\boldsymbol{x})$$

# Lagrangians of the Sammon Transformation

- Linear constraint: $d_{ij} = ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 \; \forall i, j$

$$L(\boldsymbol{x}, \boldsymbol{\nu}) = \sum_{i,j} \nu_{ij}(||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 - d_{ij})$$

# Lagrangians of the Sammon Transformation

- Linear constraint: $d_{ij} = ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 \ \forall i, j$

$$L(\boldsymbol{x}, \boldsymbol{\nu}) = \sum_{i,j} \nu_{ij}(||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 - d_{ij})$$

- Quadratic constraint: $d_{ij}^2 = ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2 \ \forall i, j$

$$L(\boldsymbol{x}, \boldsymbol{\nu}) = \sum_{i,j} \nu_{ij}(d_{ij}^2 - ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2)$$

# Lagrangians of the Sammon Transformation

- Linear constraint: $d_{ij} = ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 \; \forall i, j$

$$L(\boldsymbol{x}, \boldsymbol{\nu}) = \sum_{i,j} \nu_{ij}(||\boldsymbol{x}_i - \boldsymbol{x}_j||_2 - d_{ij})$$

- Quadratic constraint: $d_{ij}^2 = ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2 \; \forall i, j$

$$L(\boldsymbol{x}, \boldsymbol{\nu}) = \sum_{i,j} \nu_{ij}(d_{ij}^2 - ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2)$$

Based on these Lagrangians, we found different objective functions, in which we can chose different weighting factors. For example for a stronger weighting of small distances.

# Following steps

# Following steps of my thesis

# Following steps of my thesis

- Comparison of the different objective functions

# Following steps of my thesis

- Comparison of the different objective functions
- Finding the best possible weighting factors

# Following steps of my thesis

- Comparison of the different objective functions
- Finding the best possible weighting factors
- Looking for a faster method to find the global minimum than a gradient descent

# Following steps of my thesis

- Comparison of the different objective functions
- Finding the best possible weighting factors
- Looking for a faster method to find the global minimum than a gradient descent
- Apply the found solution on real datasets with pathological speech and movement fields of MR data

# Following steps of my thesis

- Comparison of the different objective functions
- Finding the best possible weighting factors
- Looking for a faster method to find the global minimum than a gradient descent
- Apply the found solution on real datasets with pathological speech and movement fields of MR data

## Aim of the thesis:

Finding a convex function with the same properties like the Sammon Transformation which also minimizes the Stress Function.

The End