SCALABLE INTERACTIVE DATA VISUALIZATION

Florian Chen and Thomas Gärtner — TU Wien, Vienna, Austria

MOTIVATION

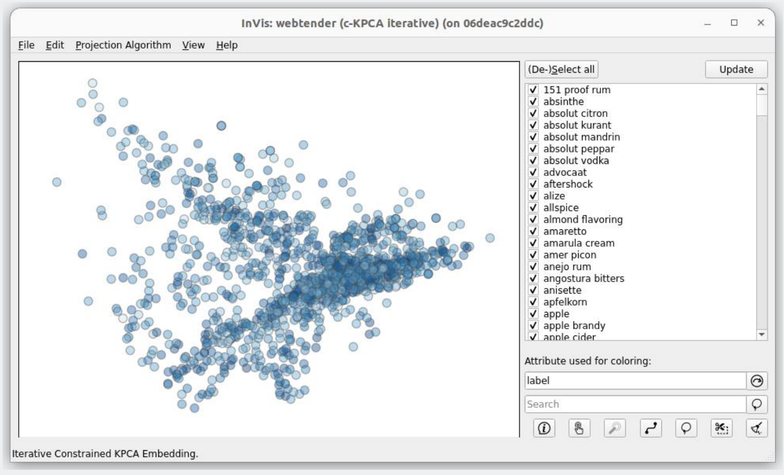
Data visualization is crucial for exploration, and interactive visualizations enable dynamic adjustments.

Control points offer direct interaction, allowing users to move points, affecting other points based on similarity.

Prior implementations are often slow on large datasets due to recalculations needed as a point is dragged around.

APPLICATION

Check it out!



*We implemented all our algorithms in an updated version of **InVis** [1]!*

KNOWLEDGE-BASED KERNEL PCA

Knowledge-based kernel PCA [2] extends kernel PCA with constraints for user interaction. (e.g. control point)

We seek functions $f_1 \dots f_d$ that maximize:

$$f_s^* = \underset{f \in \mathcal{H}}{\operatorname{argmax}} \quad \frac{1}{n}\sum_{i=1}^{n}(f(x_i) - \langle f, \mu\rangle)^2 - \nu\sum_{s'=1}^{s-1}\langle f_{s'}, f\rangle^2 - \rho\frac{1}{m}\sum_{i=1}^{m}\|f(x_i) - y_{si}\|^2,$$

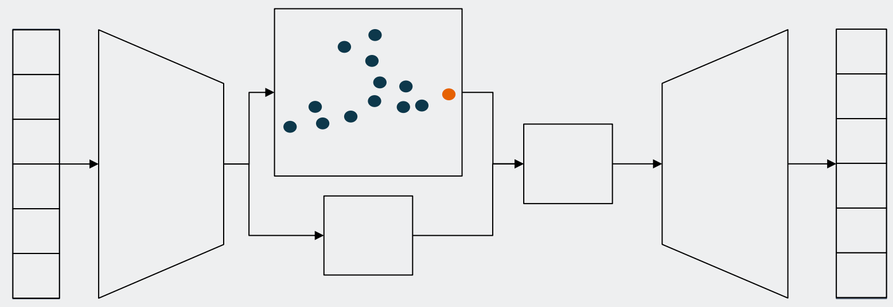
$$\text{subject to} \quad \|f\|_{\mathcal{H}_X} = 1$$

Maximize variance — Maximize orthogonality — Control point constraint

VARIATIONAL AUTOENCODER

Visualize means of a two-dimensional embedding learned by a variational autoencoder (VAE)

Fine-tune VAE with added control point location loss

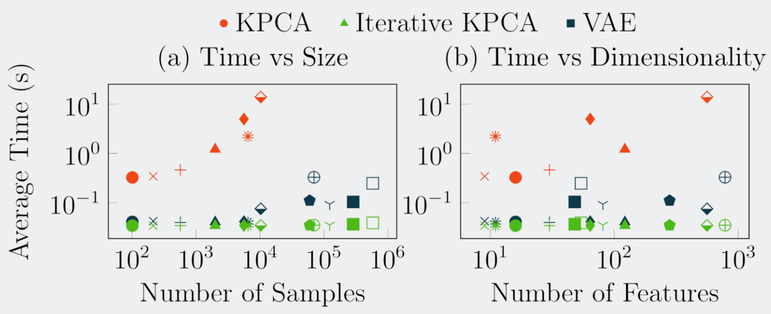


EXPERIMENTS AND RESULTS

Scalability

Test update time on various datasets

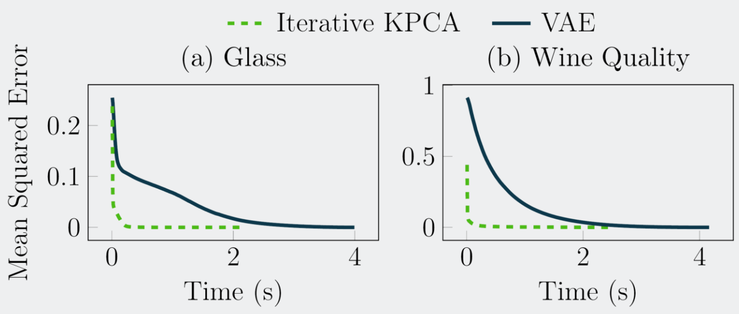
Iterative approaches scale better



Convergence

Measure MSE to the final embedding

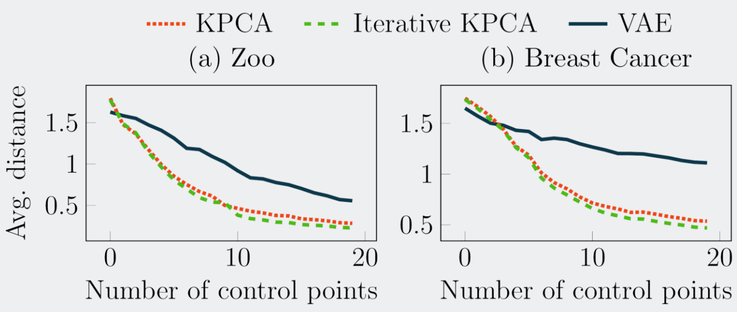
Largest changes at the beginning



Flexibility

Use control points to approx. embedding

More points improve approximation



1. Paurat, D., Gärtner, T.: *InVis: A tool for interactive visual data analysis* (2013)
2. Oglic, D., Paurat, D., Gärtner, T.: *Interactive knowledge-based kernel PCA* (2014)