HAAG Weekly Report

Nikita Angarski – 3D Modeling

Week 4

Time-Log

- What did you do this week?
  - Met with Dr. Porto and discussed next steps, which are finalized to be applying a PCA analysis on given data and find a way to integrate this information into the pycpd algorithm, which is most easily done in the smoothing kernel level.
  - Read some literature about PCA's as kernel functions, which has proven to be useful implementation-wise, while also reviewing Dr. Porto's code that he graciously shared with us for a similar process, but with gradient descent.
  - Presented my findings on the algorithms and associated math my research group, and related it to the code. This proved useful in
  - Reviewed past lectures and assigned myself to the AV role for the seminars program, and reached out to Victor who did this role in semesters' past.
  - For seminar programs, I met with Victor who handled AV last year and got access to all the requisite resources to aid in my role.
- What are you going to do next week
  - Perform PCA process on the point set data.
  - Use these PCA's on getting a replacement kernel to use that's not the Gaussian example used in the original algorithm.
- Blockers, things you want to flag, problems, etc.
  - Not currently

Abstracts:

Principal component analysis: a review and recent developments

https://pmc.ncbi.nlm.nih.gov/articles/PMC4792409/pdf/rsta20150202.pdf

Abstract: Large datasets are increasingly common and are often difficult to interpret. Principal component analysis (PCA) is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance. Finding such new variables, the principal components, reduces to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, not a priori, hence making PCA an adaptive data analysis technique. It is adaptive in another sense too, since variants of the technique have been developed that are tailored to various different data types and structures. This article will begin by introducing the basic ideas of PCA, discussing what it can and cannot do. It will then describe some variants of PCA and their application.

Summary: **Principal Component Analysis (PCA) is a statistical technique employed to reduce the dimensionality of large datasets while preserving essential information.** The provided text reviews PCA's foundations, detailing its mathematical underpinnings and adaptive nature across diverse data types. **It explores various PCA adaptations, including functional PCA for curve analysis, simplified PCA for enhanced interpretability, robust PCA to mitigate outlier influence, and symbolic data PCA for complex data structures.** Examples are also discussed, along with the challenges associated with applying PCA in specific contexts like atmospheric science and fossil analysis. **This review highlights PCA's widespread applicability and continuous evolution, showcasing its utility in numerous disciplines through ongoing adaptations and methodological advancements.** The article references numerous research papers, books and journals, highlighting the depth and importance of PCA research.

What did you do and prove it

As this week was more research, review, and going over codebases with the research team, I don't have anything to show, but the videos of meetings with team and with Dr. Porto can give a good idea of what this week was attempting to tackle. But we got all members' code working and I think that all of us have gained knowledge on the algorithm itself, so we are in a good position to make actual changes to the algorithm next week, and make a file to run the PCA and using the PCA based kernel to test and compare with the existing code.

Steve's link to our meeting with notes.
Foryoung, Steve Tse Ngwa: 3d-Modeling Researchers Meeting (Fed 3 2025)