Folding Time to Visualize Patterns of Temporal Evolution in Data

TIME CURVES

TEMPORAL EVOLVING DATA

- Audio
- Video
- The series editions of a book
- Global temperature change

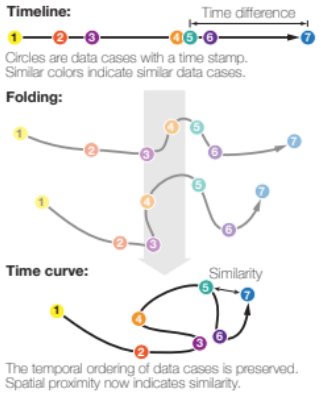
…

Bunch of examples !!

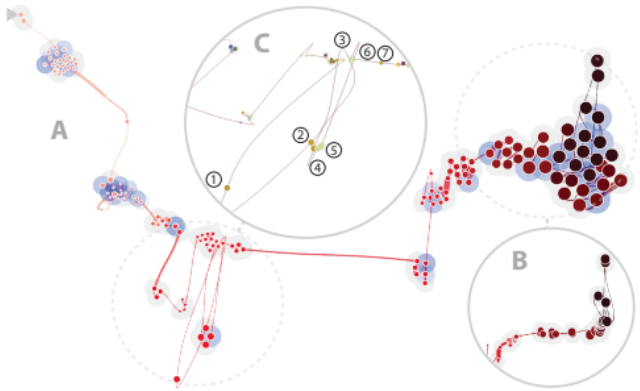
TIME CURVES

Time Curves: Folding Time to Visualize Patterns of Temporal Evolution in Data

Benjamin Bach, Conglei Shi, Nicolas Heulot, Tara Madhyastha, Tom Grabowski, Pierre Dragicevic



(a) Folding time



(b) History of the Wikipedia article on Palestine

Fig. 1. The time curve principle: a) a timeline is folded into itself in such a way that similar time points end up being close to each other; b) Example: a time curve showing the evolution of a Wikipedia article.

Abstract—We introduce *time curves* as a general approach for visualizing patterns of evolution in temporal data. Examples of such patterns include slow and regular progressions, large sudden changes, and reversals to previous states. These patterns can be of interest in a range of domains, such as collaborative document editing, dynamic network analysis, and video analysis. Time curves employ the metaphor of folding a timeline visualization into itself so as to bring similar time points close to each other. This metaphor can be applied to any dataset where a similarity metric between temporal snapshots can be defined, thus it is largely datatype-agnostic. We illustrate how time curves can visually reveal informative patterns in a range of different datasets.

Index Terms—Temporal data visualization, information visualization, multidimensional scaling

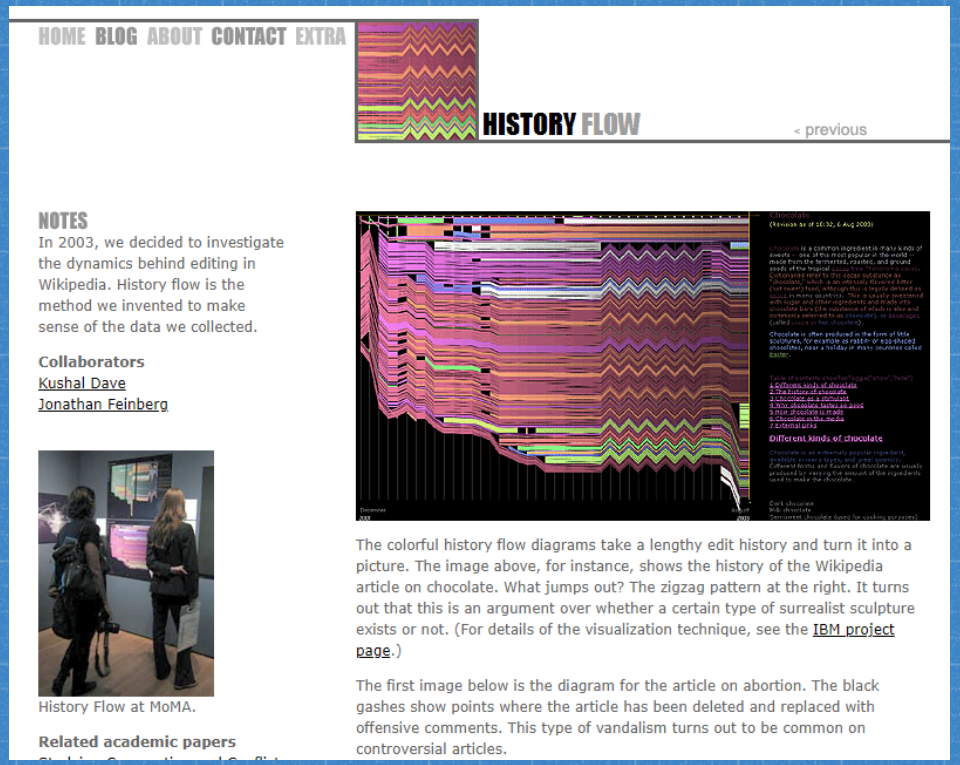
Bach, B., Shi, C., Heulot, N., Madhyastha, T., Grabowski, T., & Dragicevic, P. (2015). Time curves: Folding time to visualize patterns of temporal evolution in data. *IEEE transactions on visualization and computer graphics*, 22(1), 559-568.

1
RELATED WORK

History Flow, MDS,
Arc Diagram, MultiPiles,
& Video Trajectory

VISUALIZING TEMPORAL DATA

- History Flow
  - Visualizing the editing history of a Wiki page

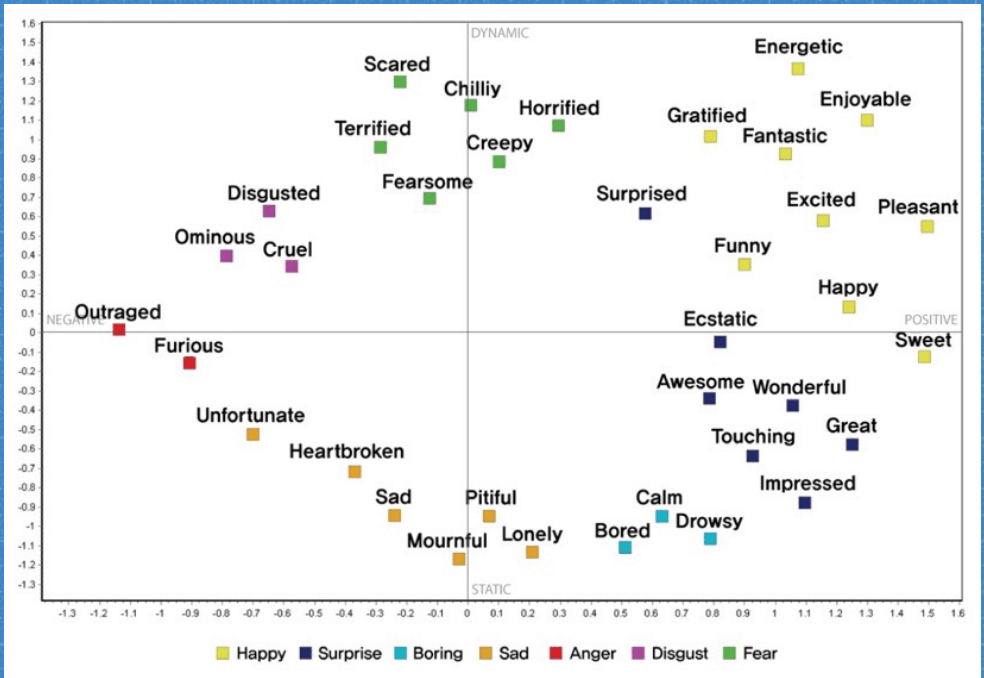


http://hint.fm/projects/historyflow/

Viégas, F. B., Wattenberg, M., & Dave, K. (2004, April). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 575-582).

VIZUALIZING SIMILARITY DATA

- Multidimensional Scaling (MDS)
  - Lay out data points on a low-dim space

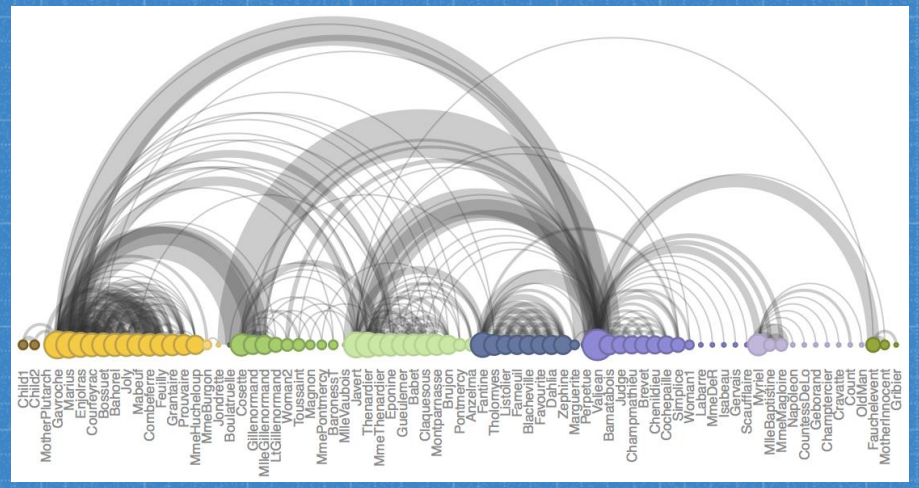


Kruskal, J. B. (1978). Multidimensional scaling (No. 11). Sage.

Borg, I., & Groenen, P. J. (2005). Modern multidimensional scaling: Theory and applications. Springer Science & Business Media.

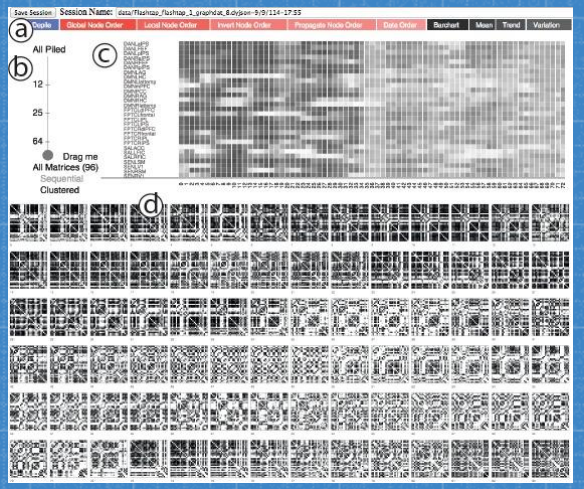
VISUALIZING BOTH TEMPORAL & SIMILARITY DATA

- Arc Diagram



Wattenberg, M. (2002, October). Arc diagrams: Visualizing structure in strings. In *IEEE Symposium on Information Visualization*, 2002. INFOVIS 2002. (pp. 110-116). IEEE.

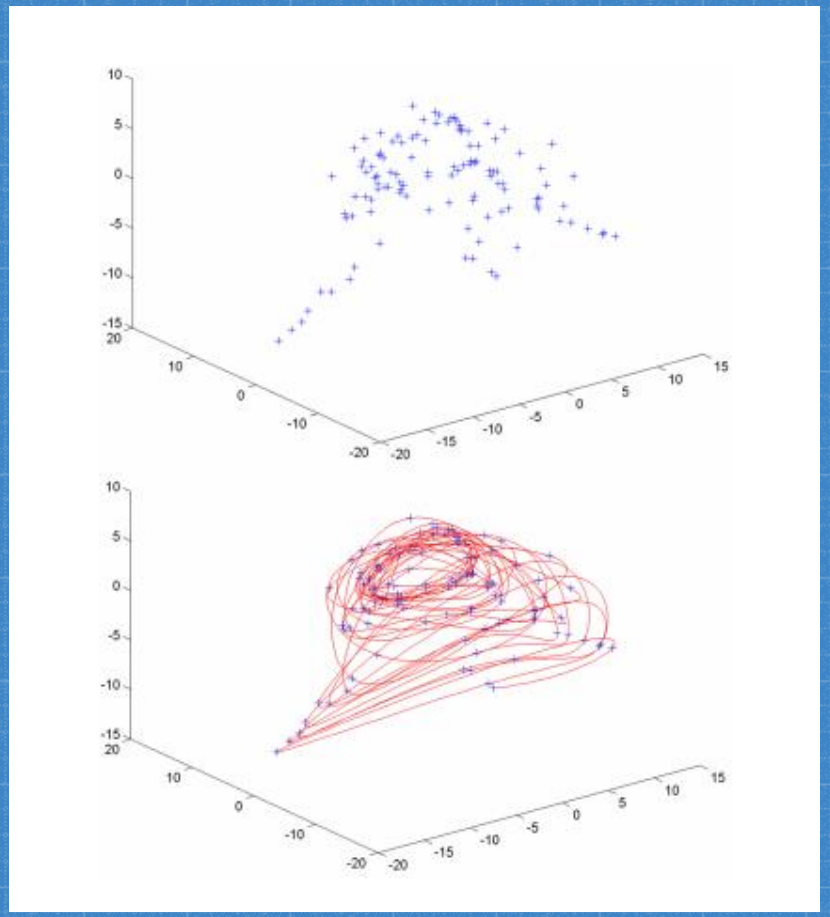
- MultiPiles



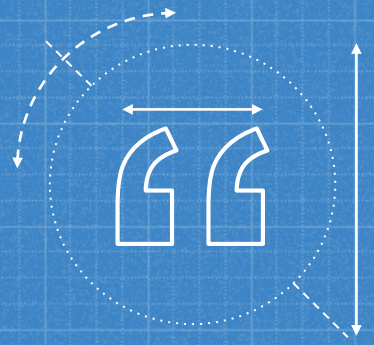
Bach, B., Henry-Riche, N., Dwyer, T., Madhyastha, T., Fekete, J. D., & Grabowski, T. (2015, June). Small MultiPiles: Piling time to explore temporal patterns in dynamic networks. In *Computer Graphics Forum* (Vol. 34, No. 3, pp. 31-40).

VISUALIZING BOTH TEMPORAL & SIMILARITY DATA

- Video Trajectory



Pless, R. (2003, October). Image Spaces and Video Trajectories: Using Isomap to Explore Video Sequences. In *ICCV* (Vol. 3, No. 2, pp. 1433-1440).



The goal of time curves is to offer a generic way of producing simple visual overviews for a range of temporal datasets.

2
BASIC IDENTITIES

Definition & Methodology

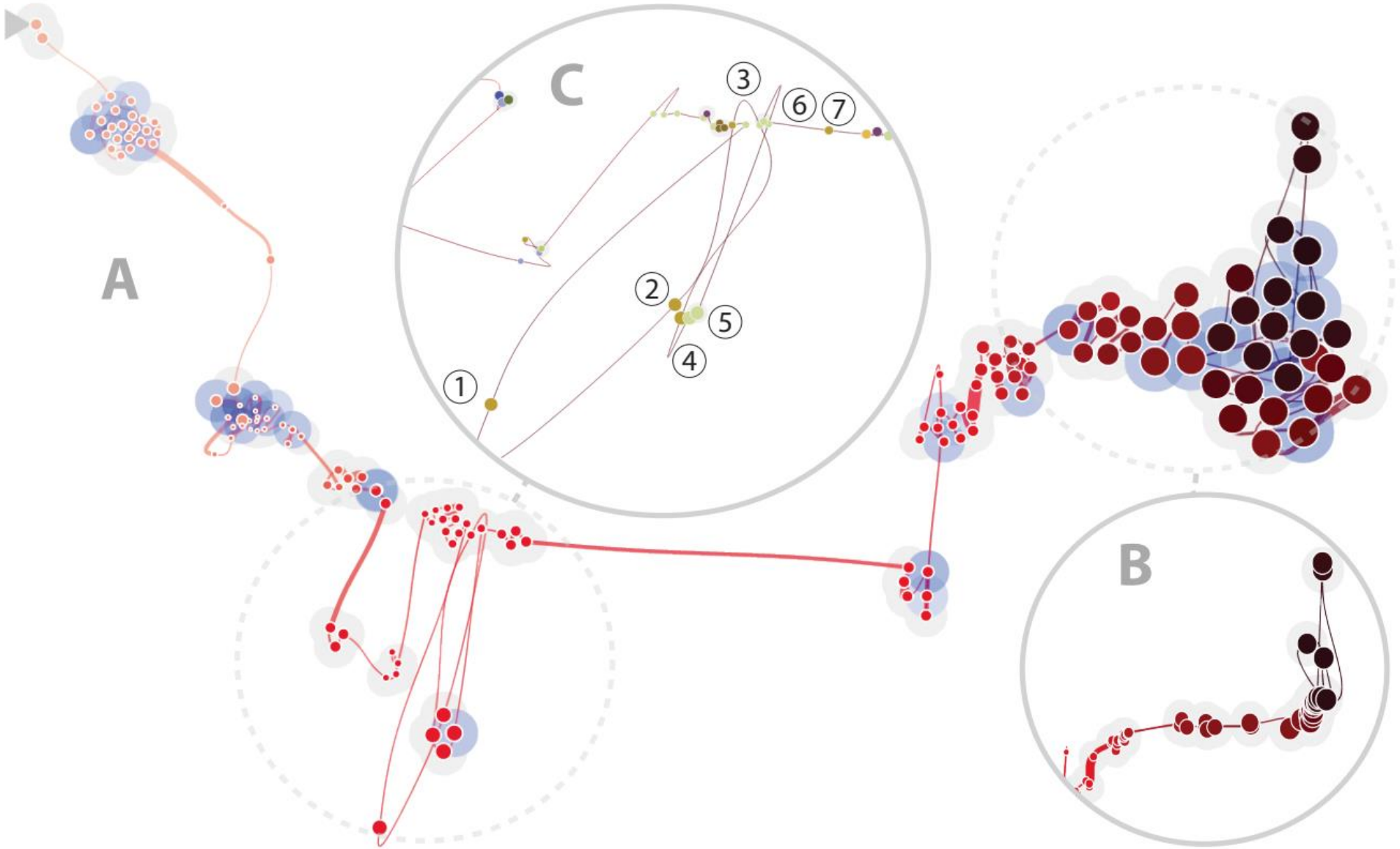
DEFINITION

- Given a temporal dataset $P = p_0, p_1, p_2, \ldots, p_n$ as a set of **time points**
- For each time point, $p_i = (t_i, s_i)$, with $t_i \in \mathbb{R}$ as the **time stamp** and $s_i \in \mathbb{S}$ as the **snapshot**.
  - Domain $\mathbb{S}$ can be any type

- With P, one can define a **distance matric** $D = \{d_{ij}\}$
  - Distance Metrice: $d: \mathbb{S}^2 \to \mathbb{R}^+$
  - Order: $\forall (i,j), i > j \Rightarrow t_i \geq t_j$
- Define P_D as the **temporal similarity dataset**

METHODOLOGY

- The **distance matric** D determines the location of data points.
  - The overlaps are removed by slightly moving the points
  - Using grey halos around the data points to indicate overlapping
  - Using blue halos around the data points to indicate a second visit to the exact same point
    - The darker the color the more the times the point has been visited



(b) History of the Wikipedia article on Palestine

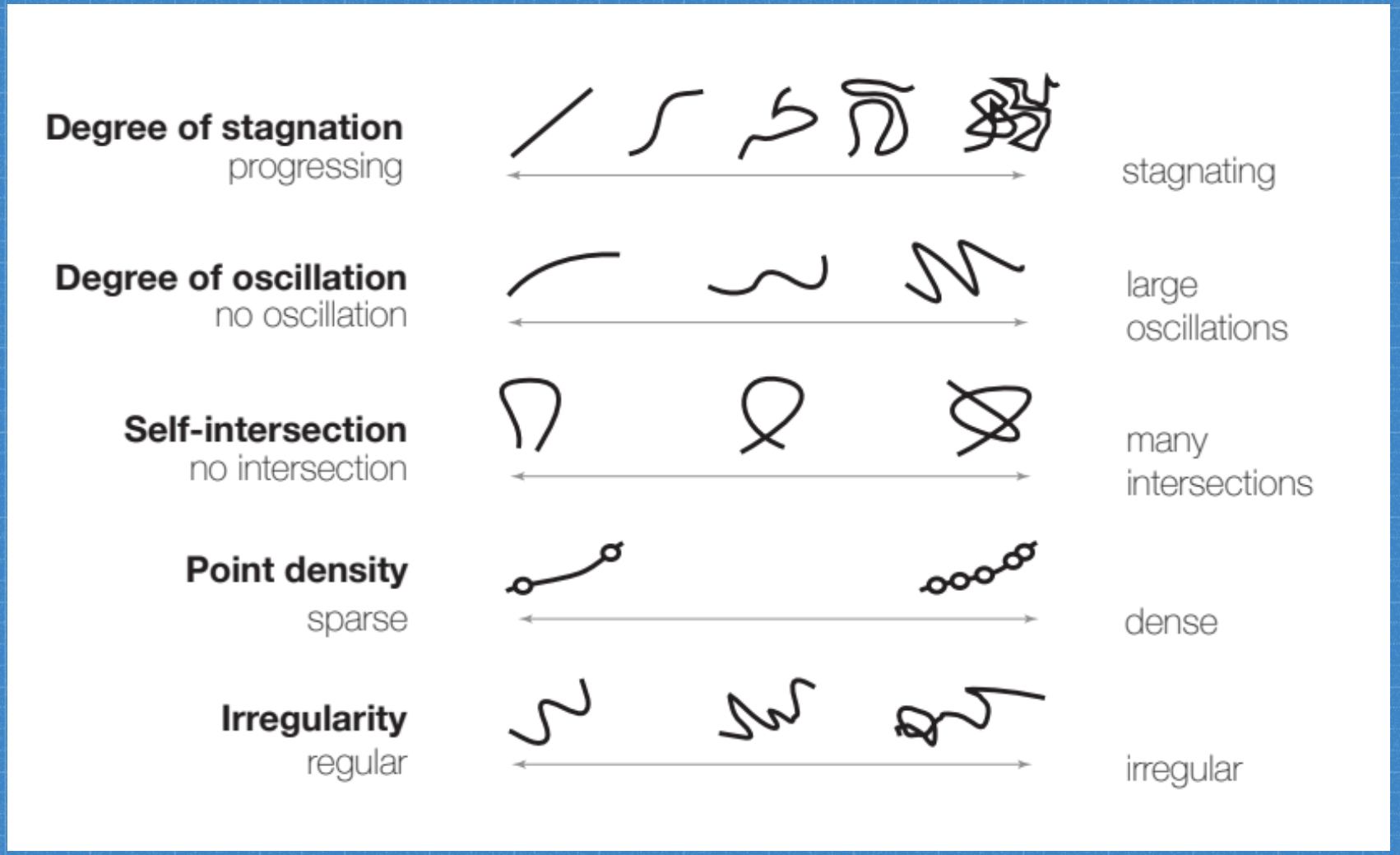
METHODOLOGY

- The color of the dots may be determined by time stamps t_i
  - The darker the color, the more recent the data point
- The color may also be determined by other aspect
  - E.g., the editors of for Wiki edition data
- The strength of a connecting curves is determined by t_i
  - The strengthen the curve, the longer the period
- For article editions specifically, the dot size represents the length of the article

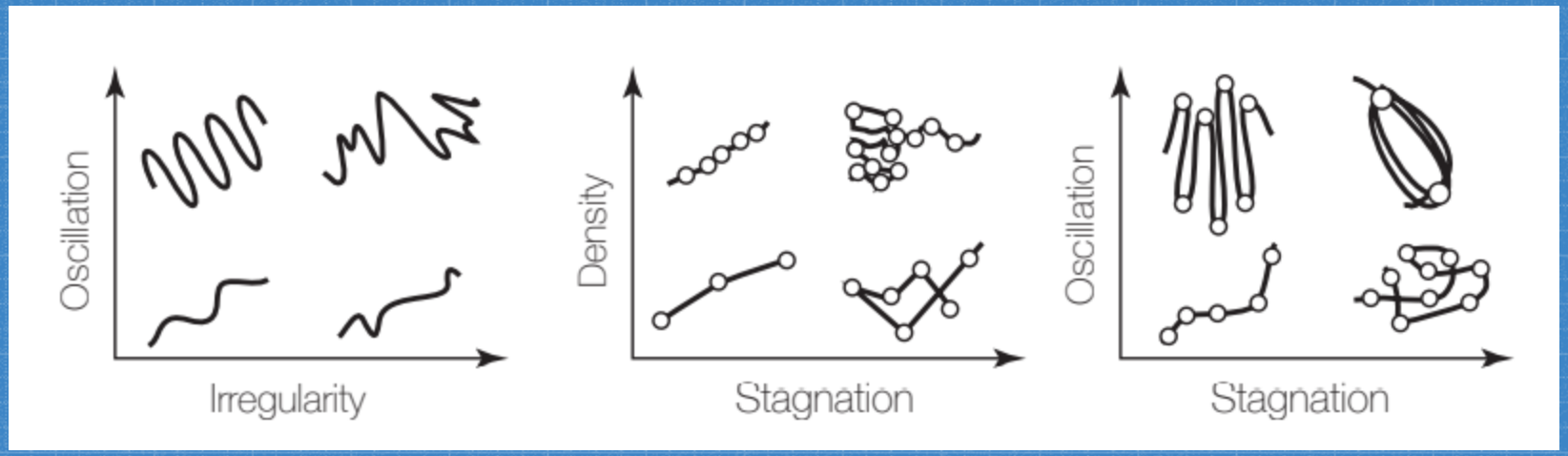
3 Characteristics

Combination of Characteristics and Patterns

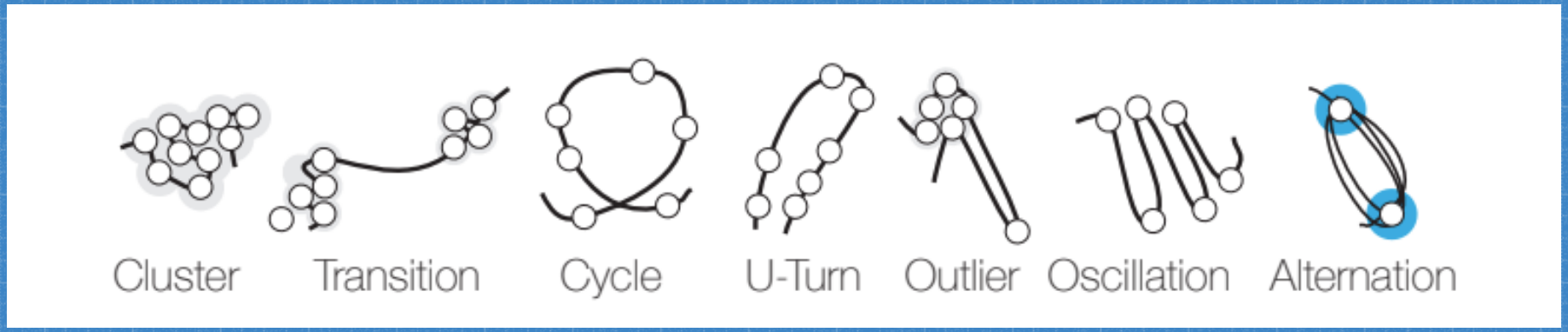
GEOMETIRC CHARACTERISTICS OF CURVES



COMBINATION OF CHARACTERISTICS



PATTERNS



4 Applications

Document Histories,
Video Recordings,
& Longitudinal Dataset

EXPLORING DOCUMENT HISTORY

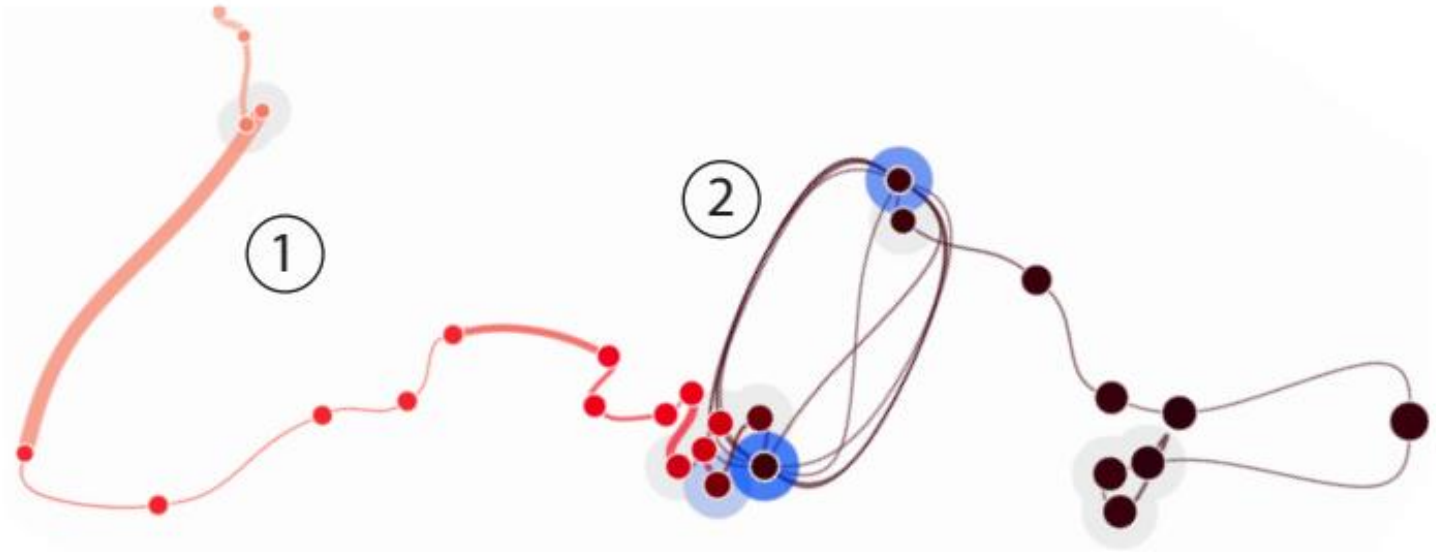


Fig. 2. Wikipedia article on *Chocolate* showing an edit war in stage (2). Blue halos indicate identical revisions.

EXPLORING DOCUMENT HISTORY

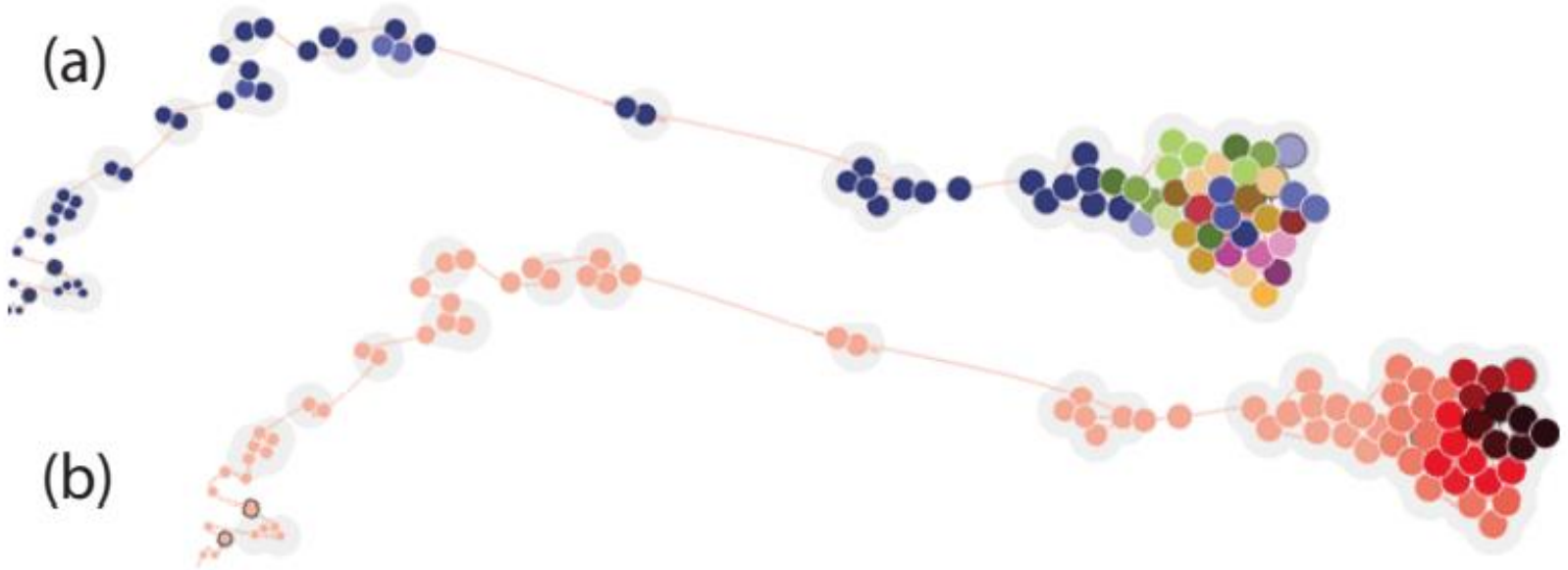


Fig. 3. Wikipedia article about former leader of German Democratic Republic *Erich Honecker*, using different color encodings for nodes.

EXPLORING DOCUMENT HISTORY

Fig. 4. Vandalism on the Wikipedia article *Crimea*. The time point on the right (very small) is a revision that contains a single link only.

EXPLORING DOCUMENT HISTORY

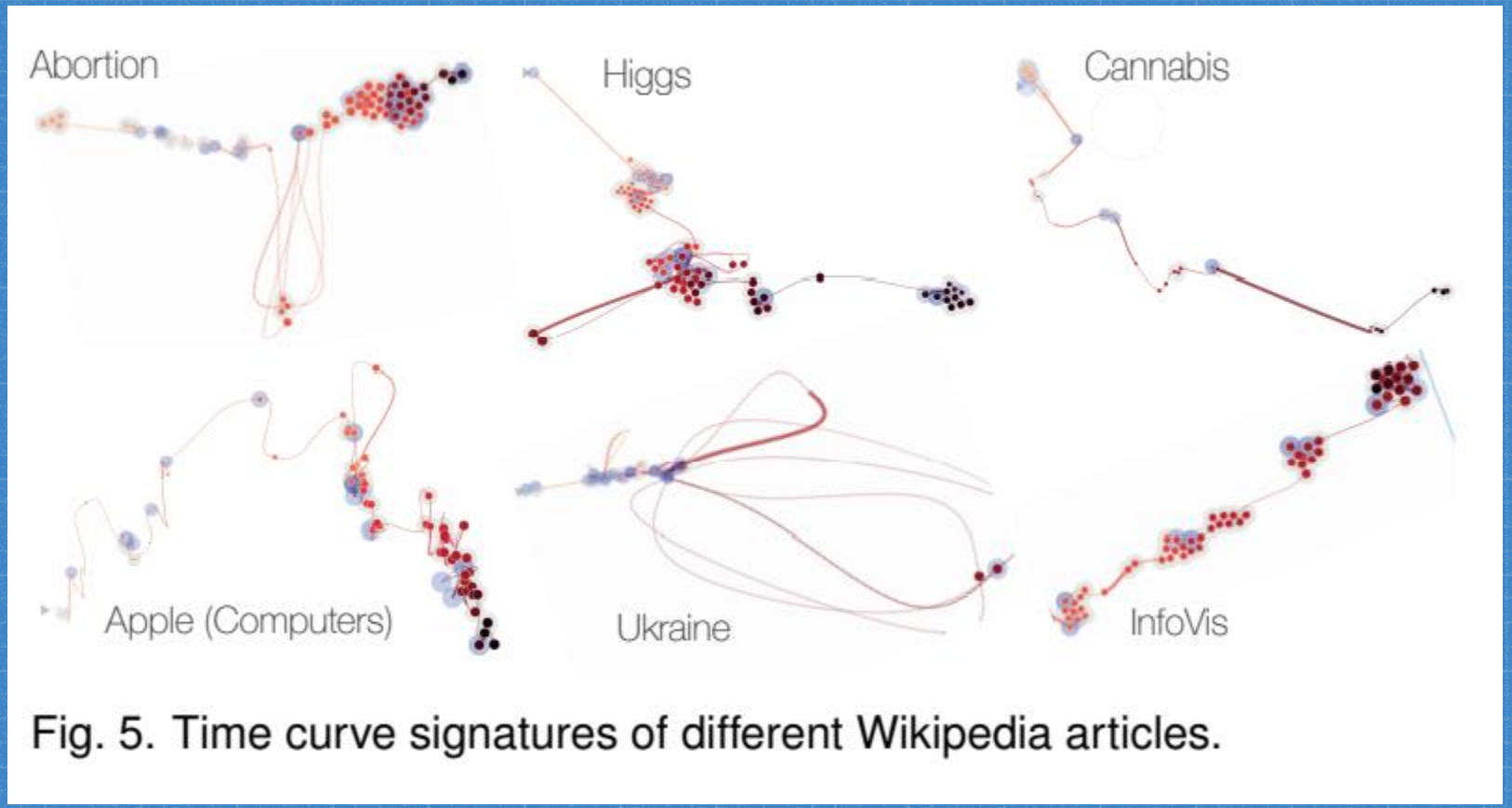


Fig. 5. Time curve signatures of different Wikipedia articles.

VIDEO RECORDINGS

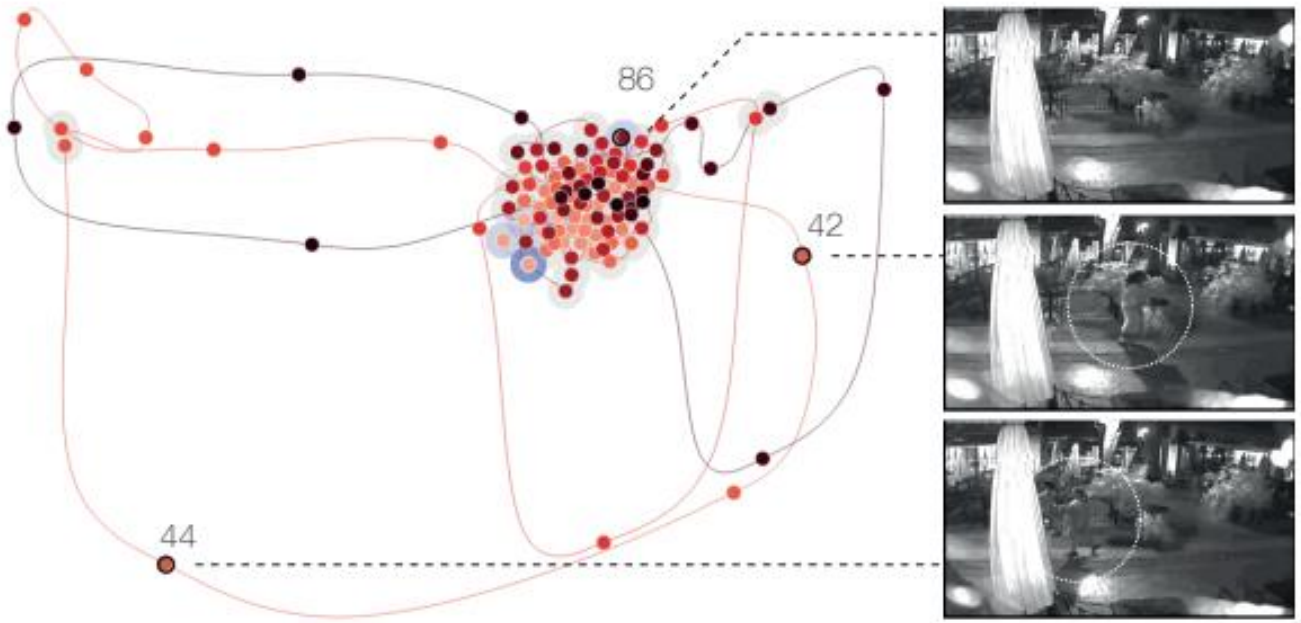


Fig. 6. A surveillance video of a street. Outliers are passing pedestrians.

VIDEO RECORDINGS

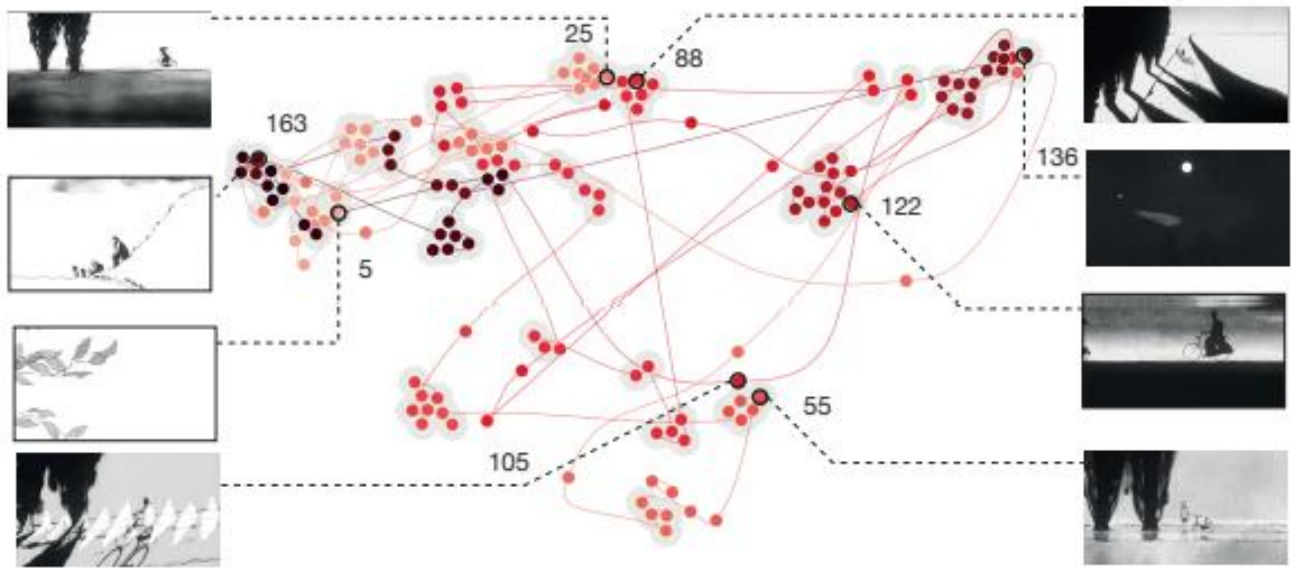


Fig. 7. Animated movie. Scenes appear as clusters. Video stills are not ordered chronologically, but placed next to the corresponding time point on the curve, showing similar scenes and scenes that get revisited.

ANALYZING DYNAMIC VIZUALIZATIONS

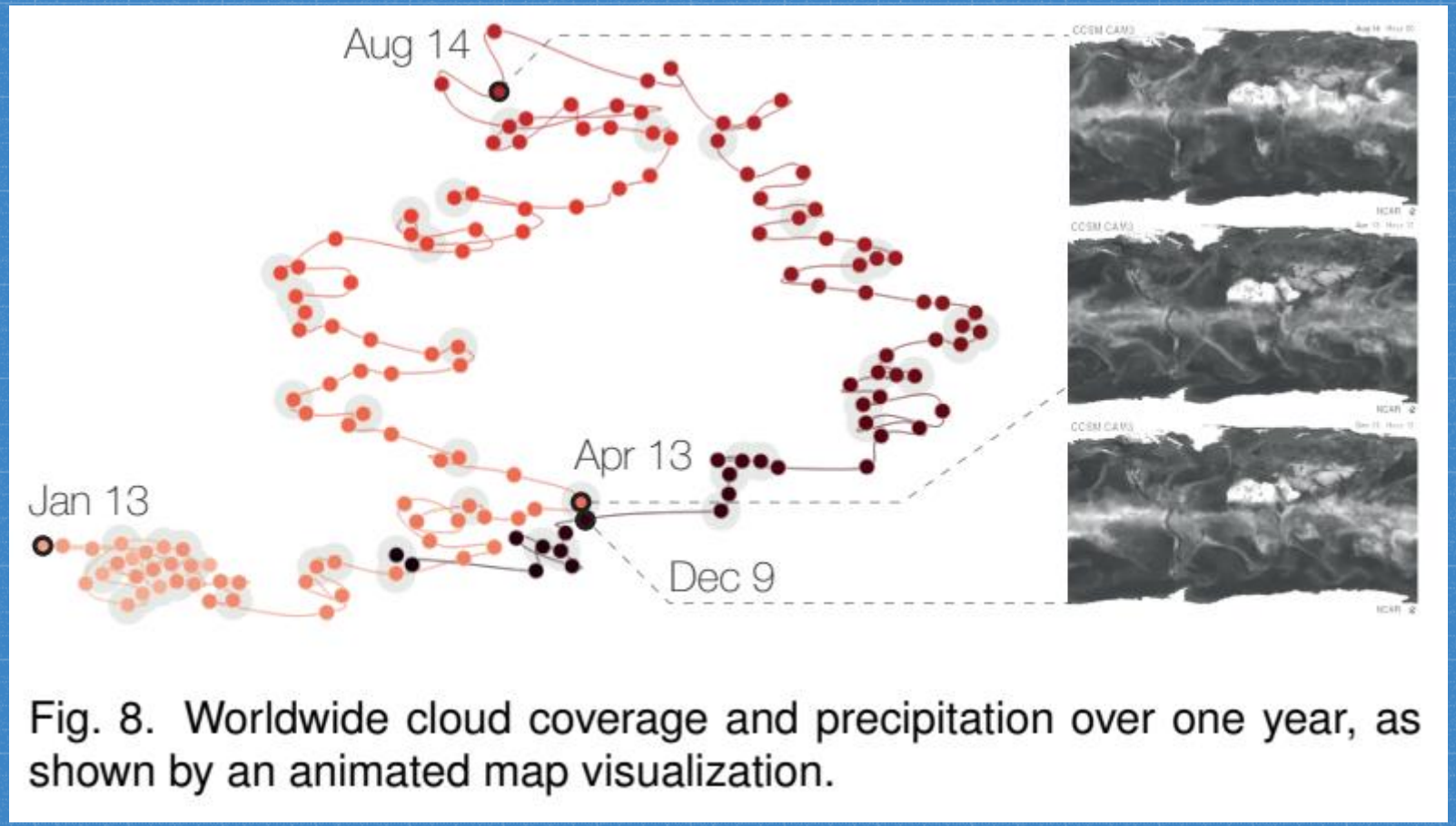


Fig. 8. Worldwide cloud coverage and precipitation over one year, as shown by an animated map visualization.

ANALYZING DYNAMIC VIZUALIZATIONS

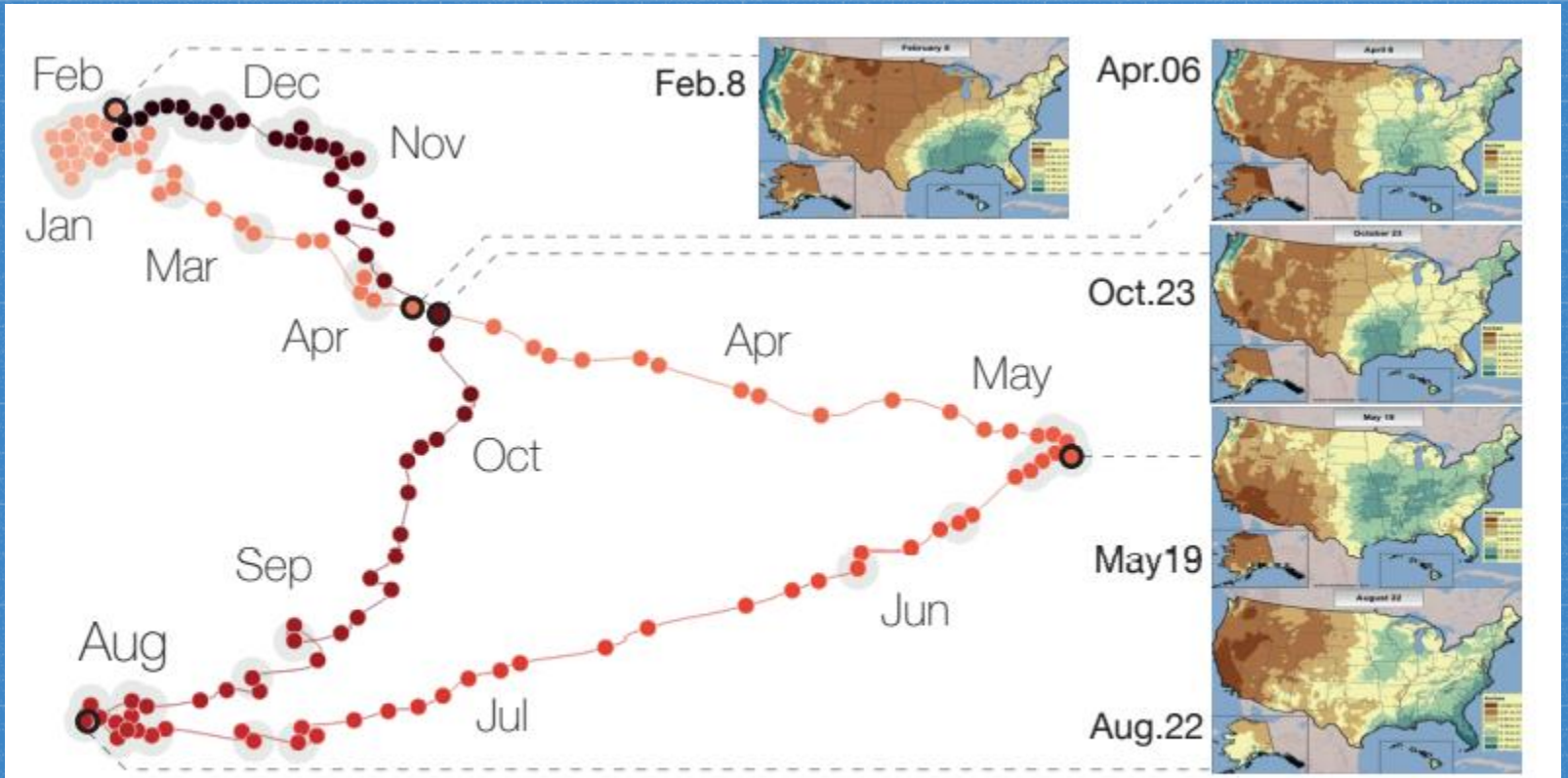


Fig. 9. Precipitation in the US across one year.

ANALYZING DYNAMIC VIZUALIZATIONS

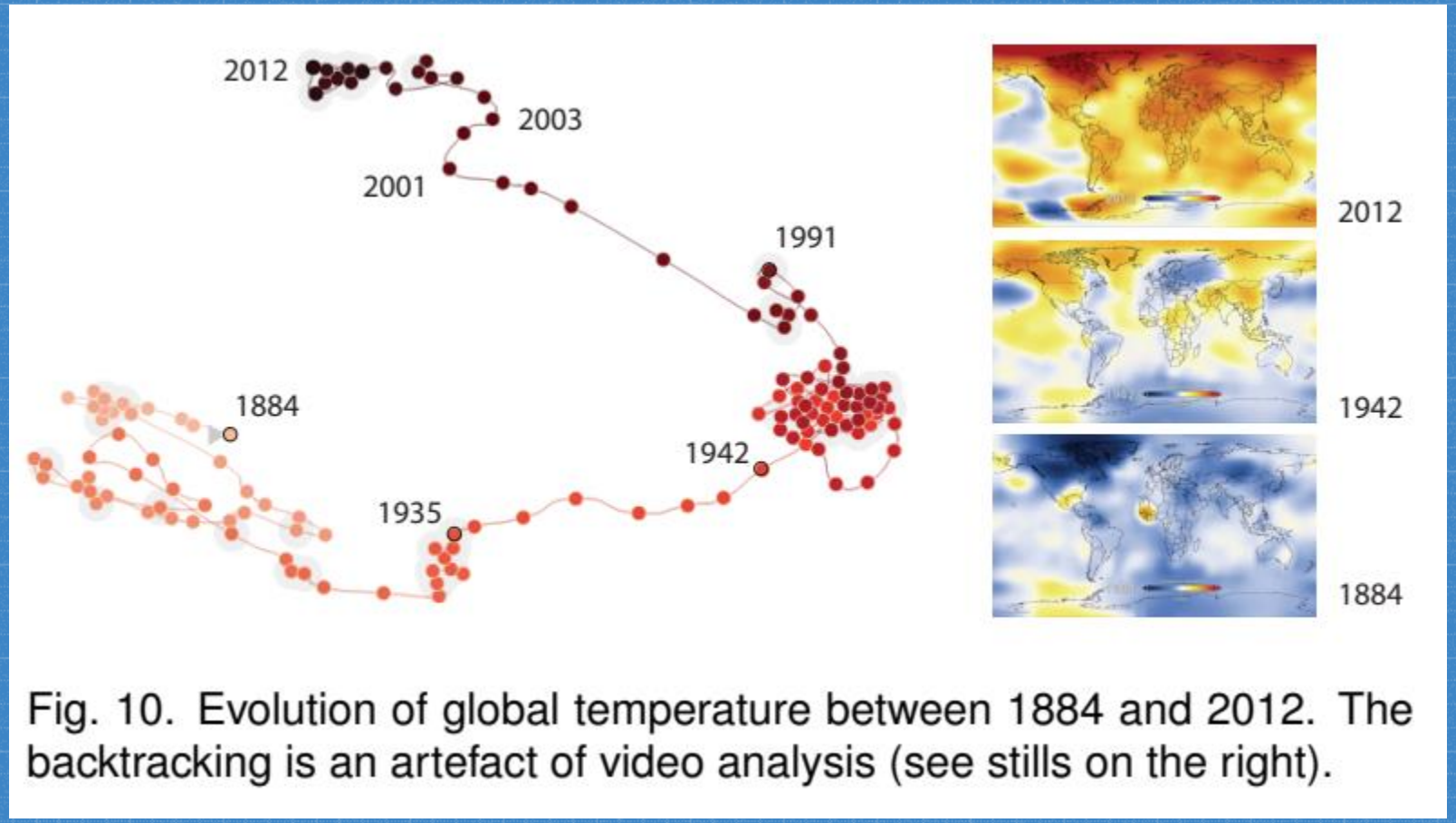


Fig. 10. Evolution of global temperature between 1884 and 2012. The backtracking is an artefact of video analysis (see stills on the right).

5 Evaluation

Time Curve in Neuroscience Research

TIME CURVE IN NEUROSCIENCE RESEARCH – with a neuroscientist T.M.

- Datasets:
  - Activity in different regions of the brain
    - Measured by blood-oxygen-level dependent signal
    - Using functional magnetic resonance imaging (fMRI)
- Tasks:
  - Identify noise in data
  - Identify specific state of brain connectivity
  - Compare patterns of connectivity
- Focus on:
  - How time curves helped domain experts formulate hypotheses

INDIVIDUAL fMRI SCANS

- Curves were embedded in their own MDS space, so cannot be accurately compared

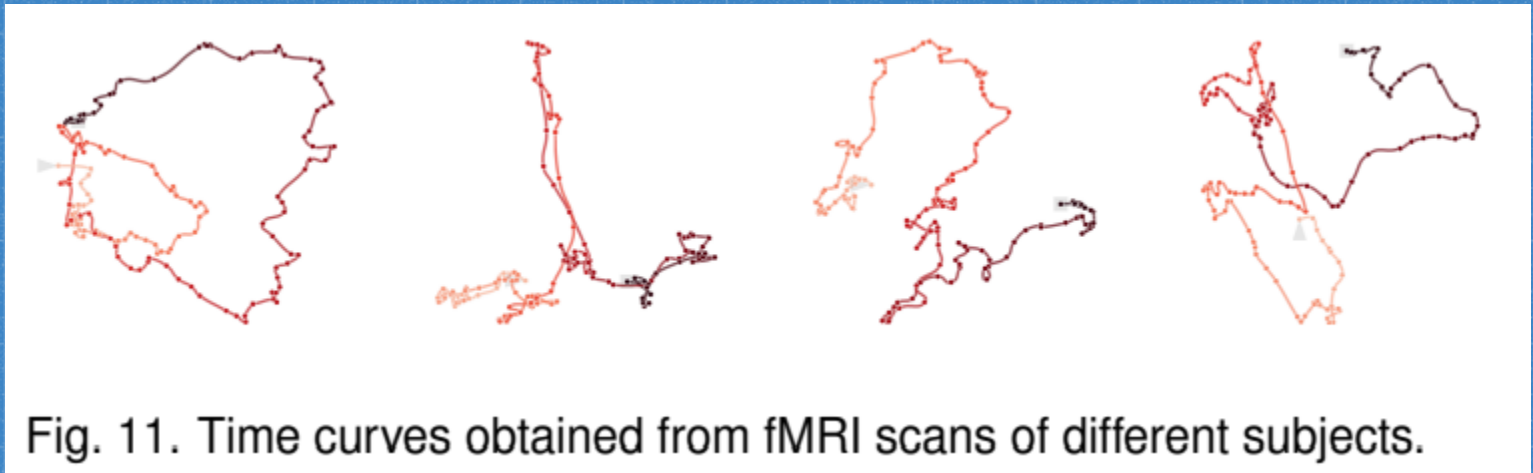


Fig. 11. Time curves obtained from fMRI scans of different subjects.

COMPARING DENOISING TECHNIQUES

- The shapes of the curves are similar

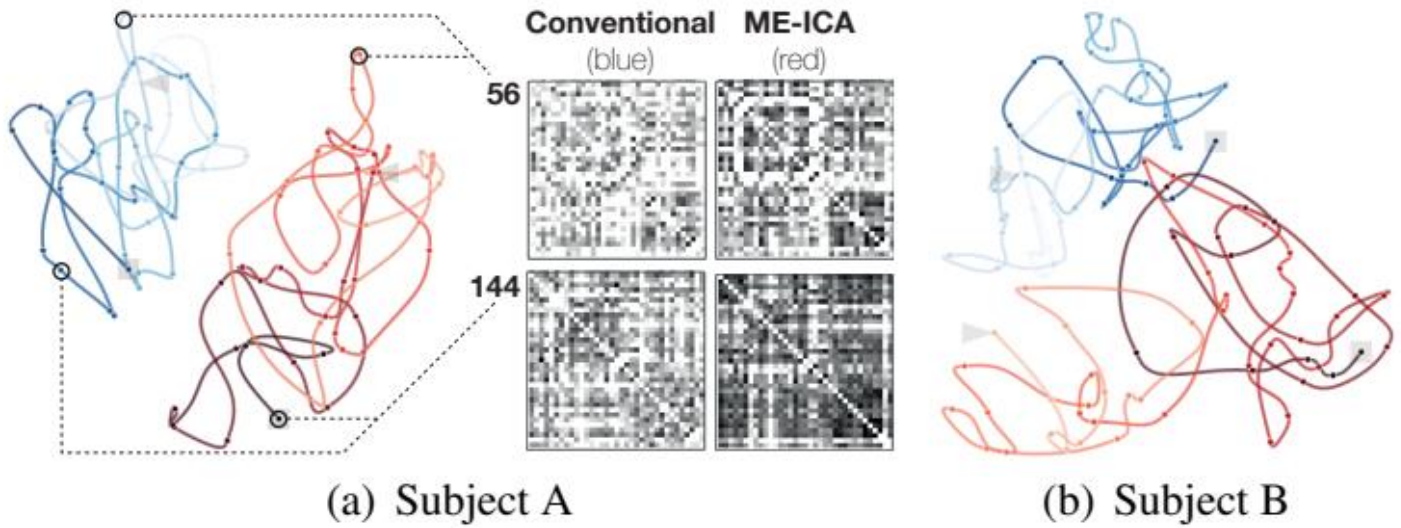


Fig. 12. Two curve pairs (subject A and subject B) showing offsets between conventionally denoised scans (blue) and scans denoised using ME-ICA (red). Matrix representations for timepoints 56 and 144 show an increase in signal strength through ME-ICA denoising.

COMPARING CONNECTIVIT ACROSS INDIVIDUALS

- The outliers and the disjointed individuals inspired future research

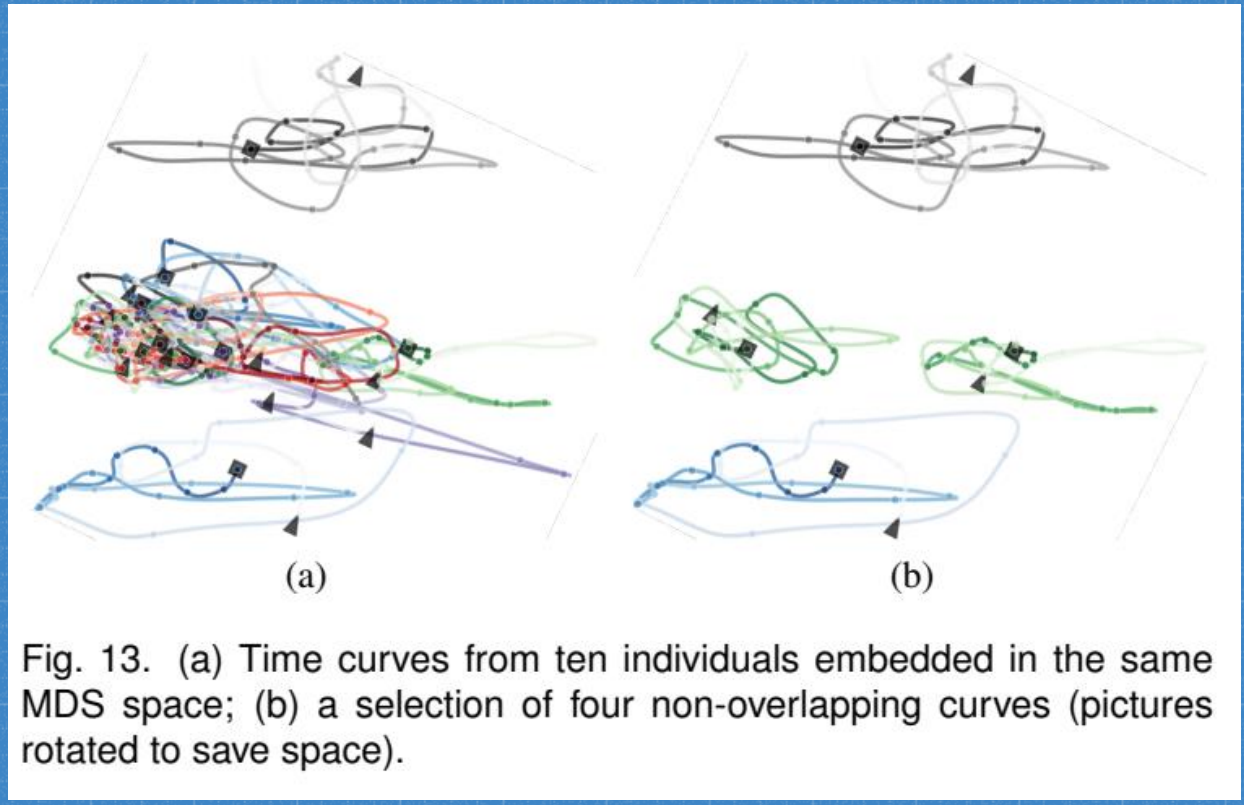


Fig. 13. (a) Time curves from ten individuals embedded in the same MDS space; (b) a selection of four non-overlapping curves (pictures rotated to save space).

USER FEEDBACKS & FUTURE DIRECTIONS

- Feedbacks:
  - Easy to understand
  - Can deliver quick insights

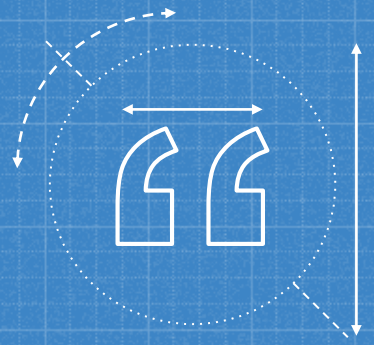
- Future Directions:
  - Apply the time curves to examine and compare different measures for brain connectivity
  - Apply the time curves to various domain of longitudinal studies

6 Limitation

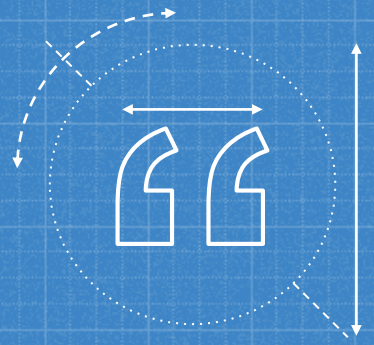
Time Curve benefits from MDS but also inherit all its weakness

LIMITATIONS

- Limited capability in dealing with qualitative information
- Do not convey all information on similarity
- Limitation on distance metrices
- Scalability. 500 seconds for 500 time points.



In what other scenario could the Time Curves be valuable for domain expertise?



How could the Time Curves be improved?

Thank You

Ruochen