

Visual Data Mining of Multimedia Data

Chen Yu, Yiwen Zhong, Thomas Smith, Ikhyun Park, and Weixia Huang

Multimedia Data



- With advances in computing and sensing techniques, multimedia data are ubiquitous.
e.g. a large amount of high-resolution high-quality multimedia data (video, audio, EEG, fMRI, etc.) in behavioral studies.
e.g. surveillance cameras.
- How to automatically and effectively discover new patterns and knowledge from rich multimedia data poses a compelling challenge.

Interaction Studies



- A pair of participants interacted with each other.
- We are interested in fined-grained measures of their behaviors, such as what they visually attend to moment by moment, what actions they generate, what they speak, the coupling of their actions, and how one person keeps track of the actions generated by another person and adjusts his own actions accordingly.
- For each pair of participants, we have collected a large amount of data in a 6-10 minutes interaction.
 - Video and audio: three video streams recorded simultaneously.
 - Motion tracking data: 200 Hz with 8 sensors. Each sensor provides 6-dimensional data (x-y-z-h-p-r).
 - Eye movement data: 120 Hz.

Multisensory Study



Multisensory Study



ASL Eye Tracker

CCD camera

Microphone



position
sensors

Data Mining of Multimedia data



multimedia data

Data Mining of Multimedia data



multimedia data



derived data

Data Mining of Multimedia data

multimedia data

derived data

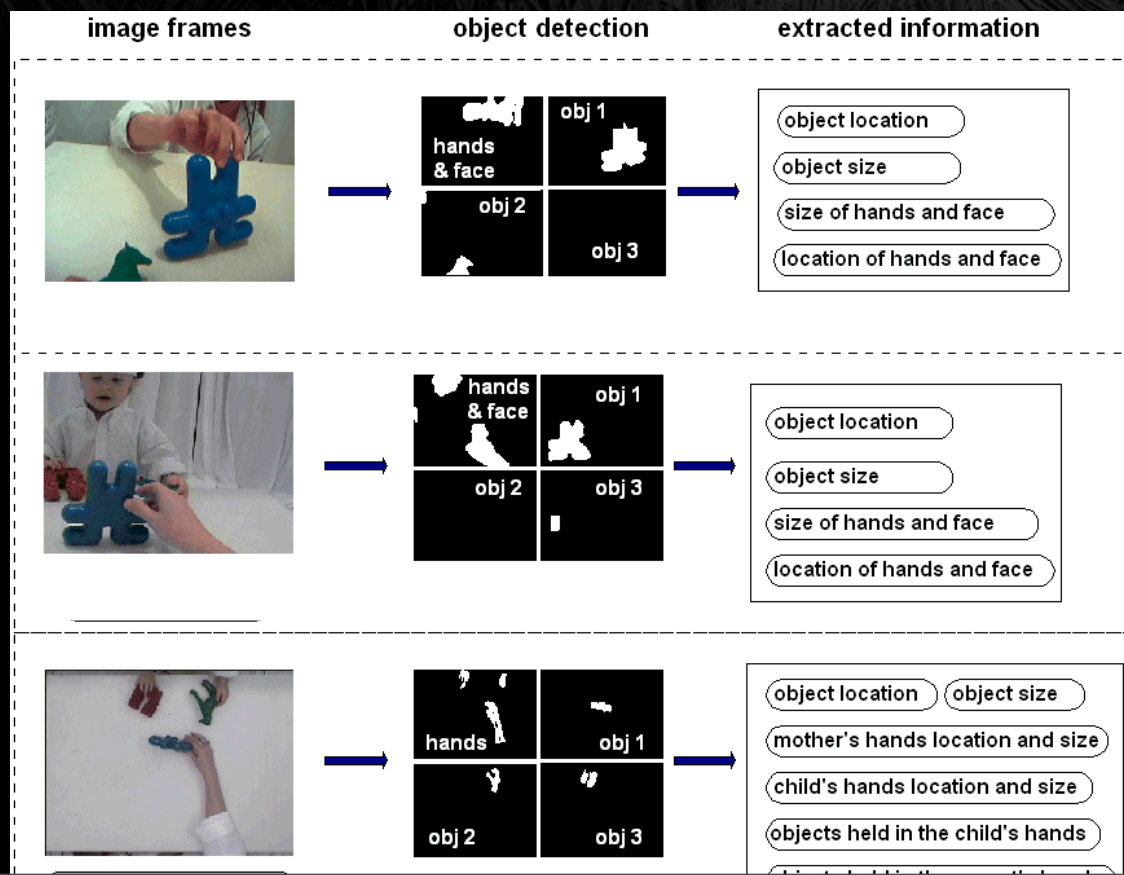
patterns



Multimedia data processing



- In the **first** step, researchers extract some derived data from raw multimedia data.



multimedia data

?

derived data

Multimedia data processing



Multimedia data processing



- Multimedia data contains various kinds of information.
 - Video: n (cameras) \times m (objects/people) \times k (properties)

Multimedia data processing



- Multimedia data contains various kinds of information.
 - Video: n (cameras) \times m (objects/people) \times k (properties)
 - Speech: In addition to linguistic content, speech data also contains other information, such as a speaker's gender and age, the speaker's emotion state, and so on.

Multimedia data processing



- Multimedia data contains various kinds of information.
 - Video: n (cameras) \times m (objects/people) \times k (properties)
 - Speech: In addition to linguistic content, speech data also contains other information, such as a speaker's gender and age, the speaker's emotion state, and so on.
- A chicken-and-egg problem: to discover new knowledge researchers may not know in advance what information is most interesting and should be extracted first. But meanwhile, without extracting some derived data first and compute some results, we may not know how to start.

Multimedia Data Mining

- In the **second** step, researchers work on derived data (time series, etc.) with the goal to find interesting **patterns**.

derived data



patterns

Multimedia Data Mining



- In the **second** step, researchers work on derived data (time series, etc.) with the goal to find interesting **patterns**.

derived data



patterns

- However, data mining algorithms can only search and discover only **pre-determined** patterns that also need to be **statistically reliable**.

Multimedia Data Mining



- In the **second** step, researchers work on derived data (time series, etc.) with the goal to find interesting **patterns**.

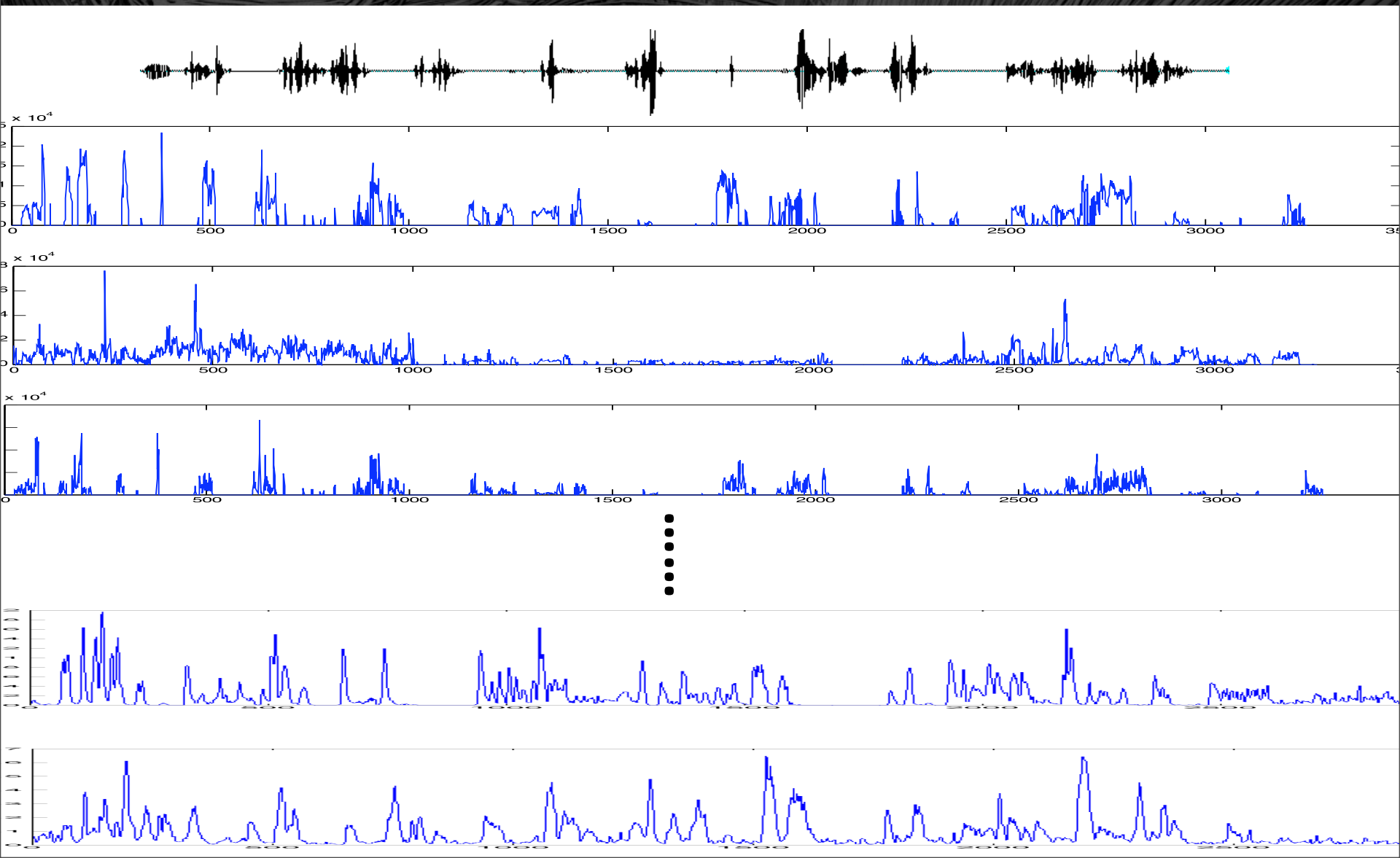
derived data



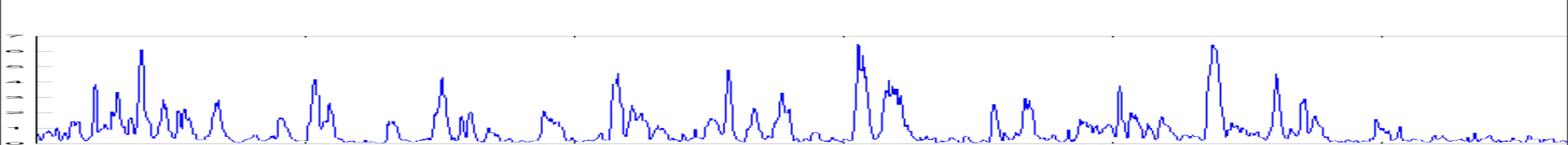
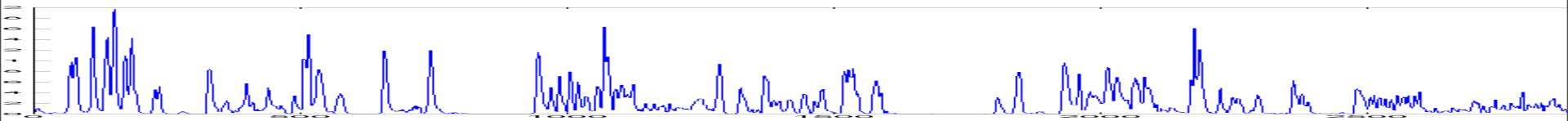
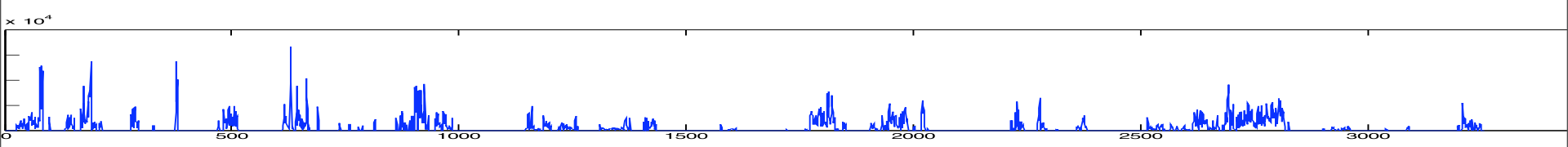
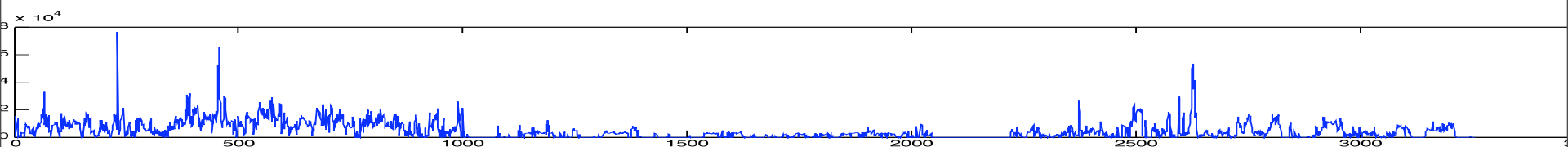
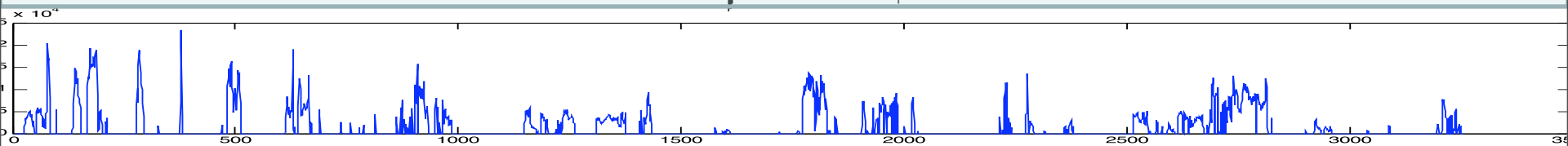
patterns

- However, data mining algorithms can only search and discover only **pre-determined** patterns that also need to be **statistically reliable**.
- The **exploratory** nature of discovering new knowledge requires the ability to detect uncommon (but interesting) patterns.

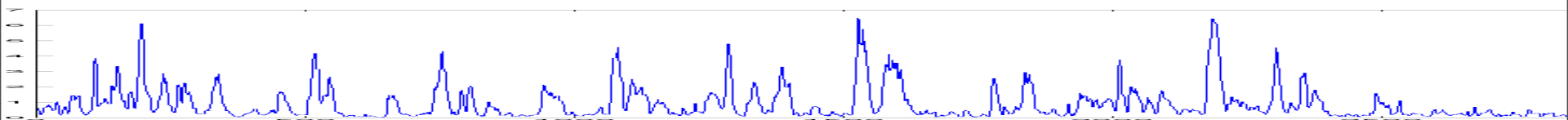
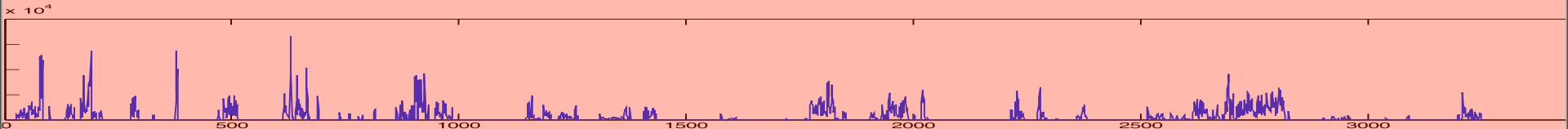
Multi-Streaming Multimodal Data



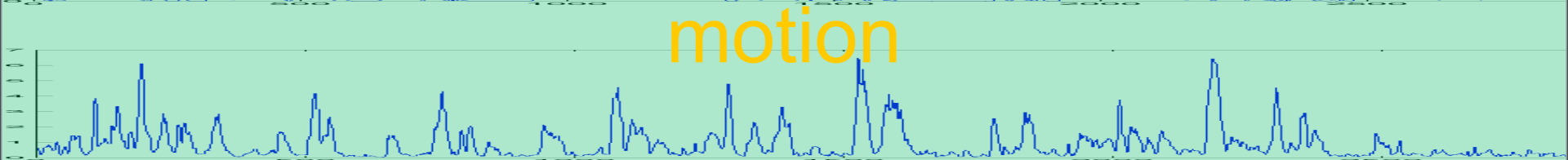
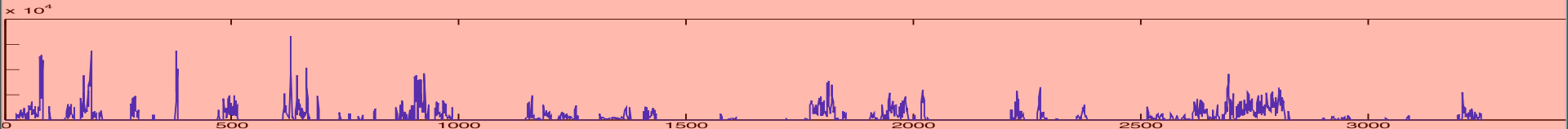
Multi-Streaming Multimodal Data



Multi-Streaming Multimodal Data



Multi-Streaming Multimodal Data



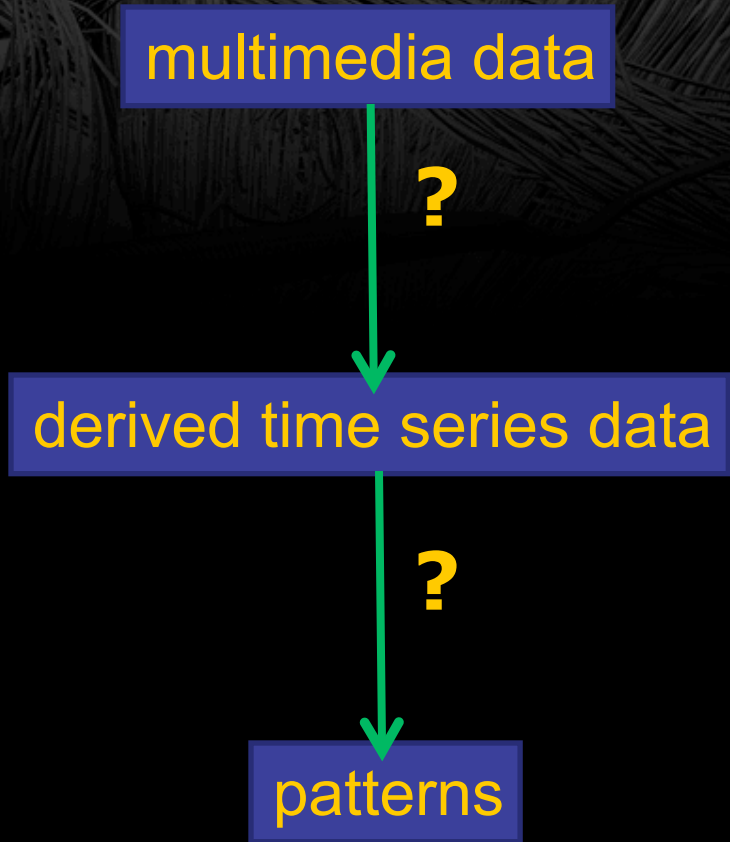
multimedia data

?

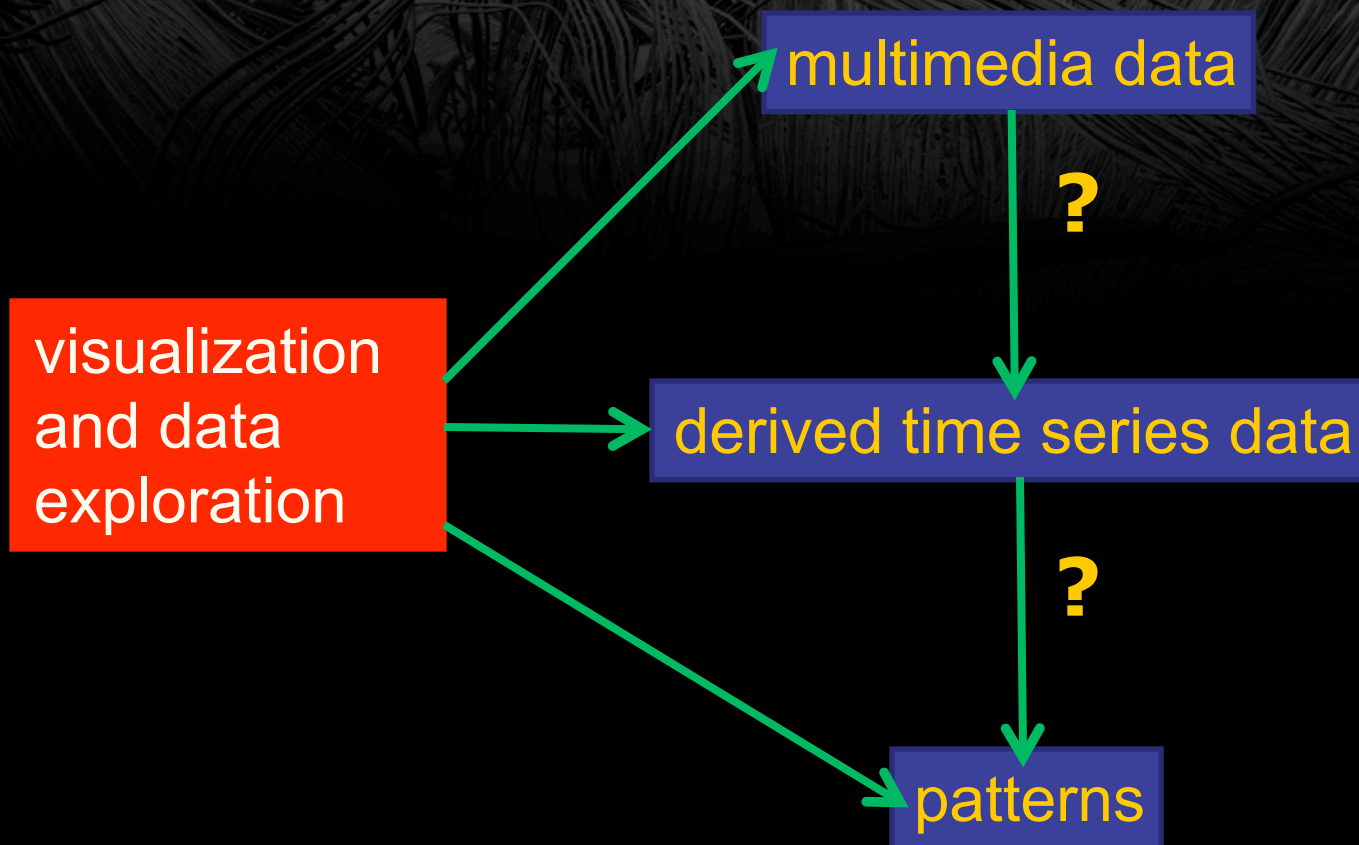
derived time series data

?

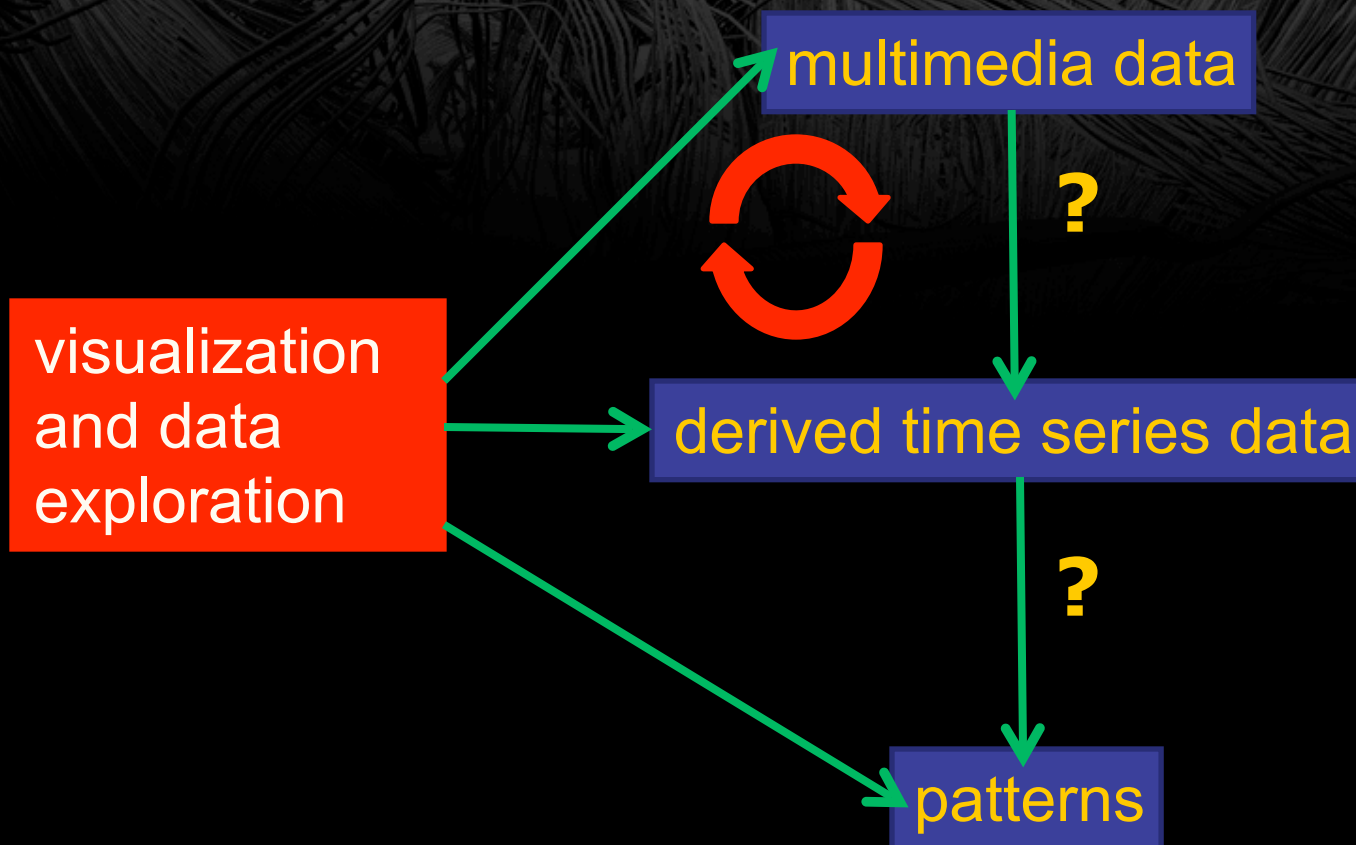
patterns



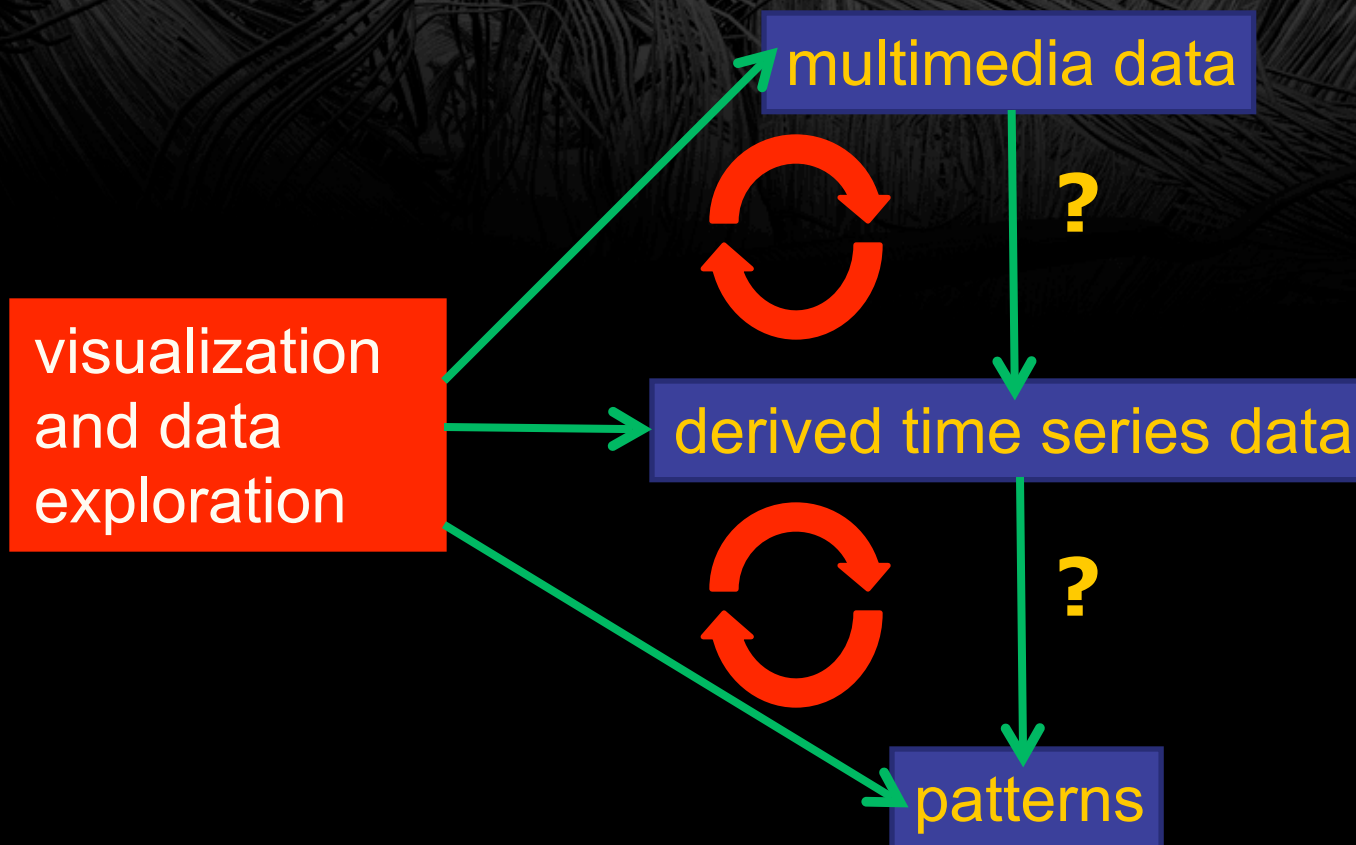
Visual Data Mining (Shneiderman, 2002)



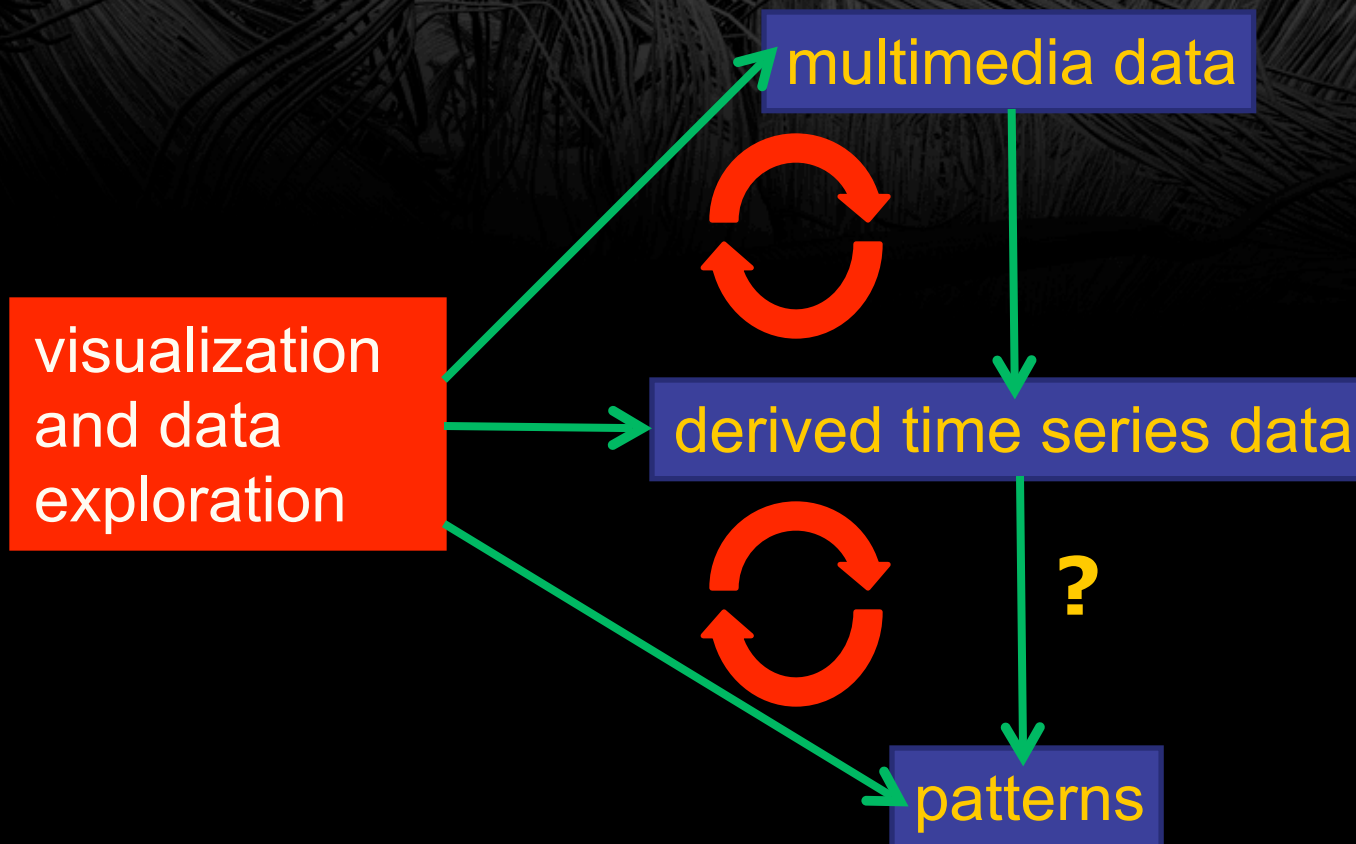
Visual Data Mining (Shneiderman, 2002)



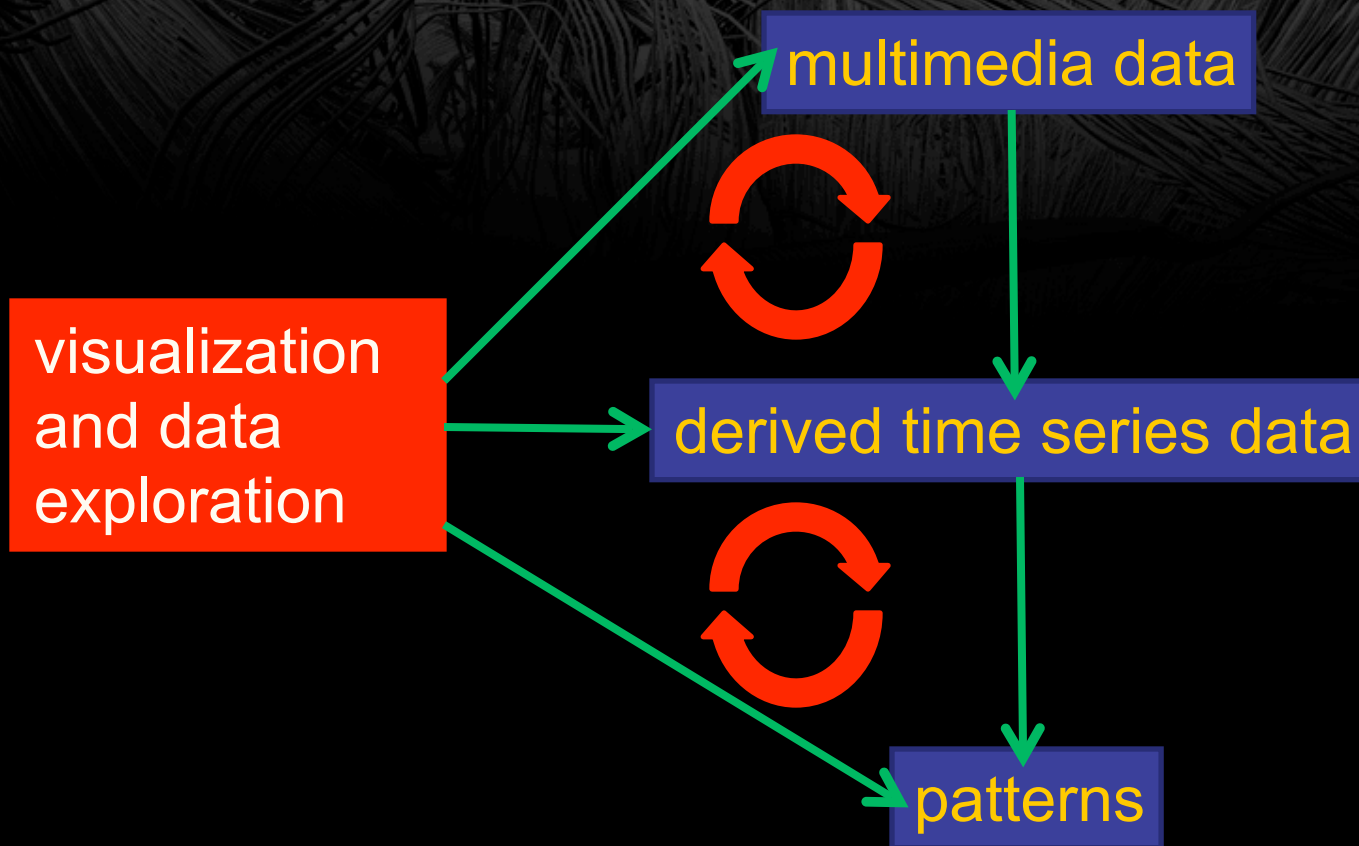
Visual Data Mining (Shneiderman, 2002)



Visual Data Mining (Shneiderman, 2002)



Visual Data Mining (Shneiderman,2002)



Visual data mining



- **Data mining:** Statistics and measures extracted from raw data may **exclude** interesting patterns embedded or even be **misleading**. We need a mechanism to capture the overall statistics and meanwhile make fine-grained data accessible.

Visual data mining



- **Data mining:** Statistics and measures extracted from raw data may **exclude** interesting patterns embedded or even be **misleading**. We need a mechanism to capture the overall statistics and meanwhile make fine-grained data accessible.
- **Visualization** techniques can present the data in various informative ways and by doing so make it easier for researchers to detect new patterns, to gain new insights, and to generate new hypotheses which will lead to new discoveries.

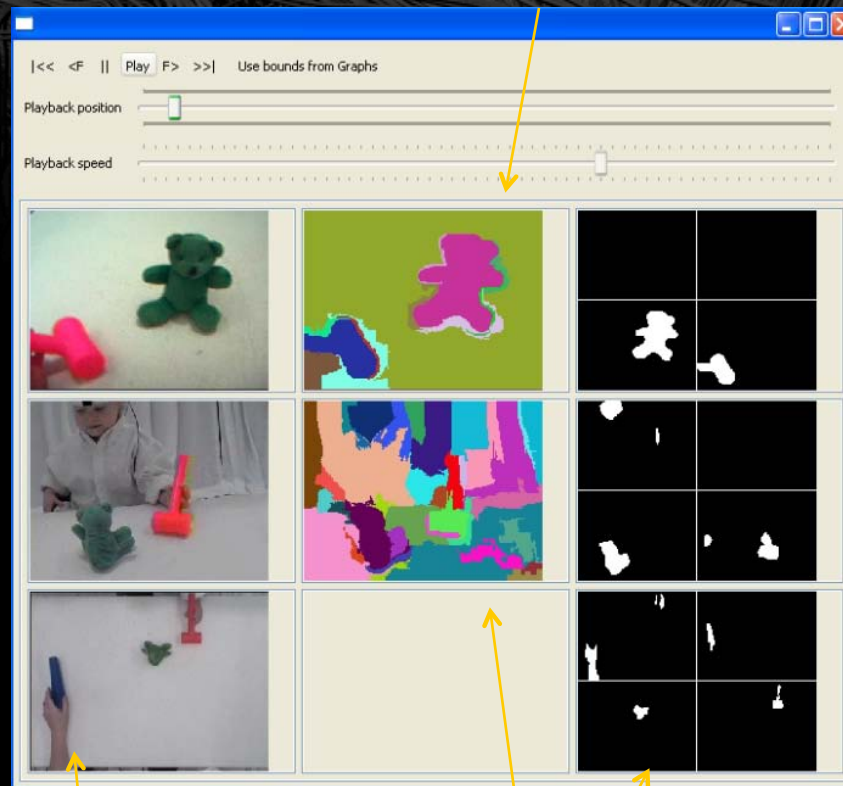
Visual data mining



- **Data mining:** Statistics and measures extracted from raw data may **exclude** interesting patterns embedded or even be **misleading**. We need a mechanism to capture the overall statistics and meanwhile make fine-grained data accessible.
- **Visualization** techniques can present the data in various informative ways and by doing so make it easier for researchers to detect new patterns, to gain new insights, and to generate new hypotheses which will lead to new discoveries.
- **Interactive mode:** data mining and visualization can bootstrap each other – the more informative visualization based on new results will lead to the discovery of more complicated patterns which in turn can be visualized again to lead to more findings.

Accessing the raw data

Multimedia playback panel



raw video

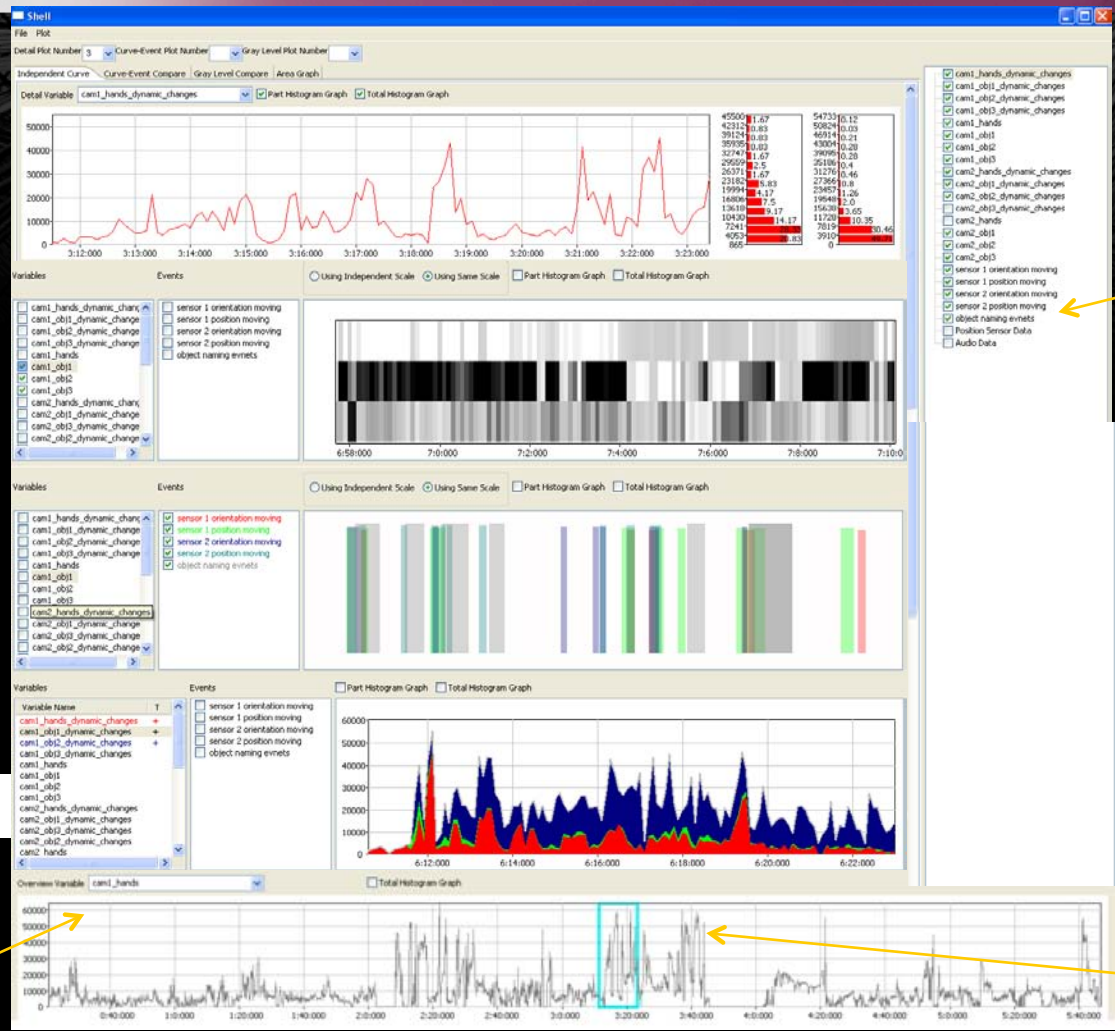
processed image frames

Derived Temporal Data



- **Continuous variables:** related to time points (a series of single measurement at particular moments in time).
E.g. the location of an object in video
- **Event variables:** related to time intervals (the onset and offset of an event).
E.g. a person is speaking or not.

Visualizing derived data based on TimeSearcher

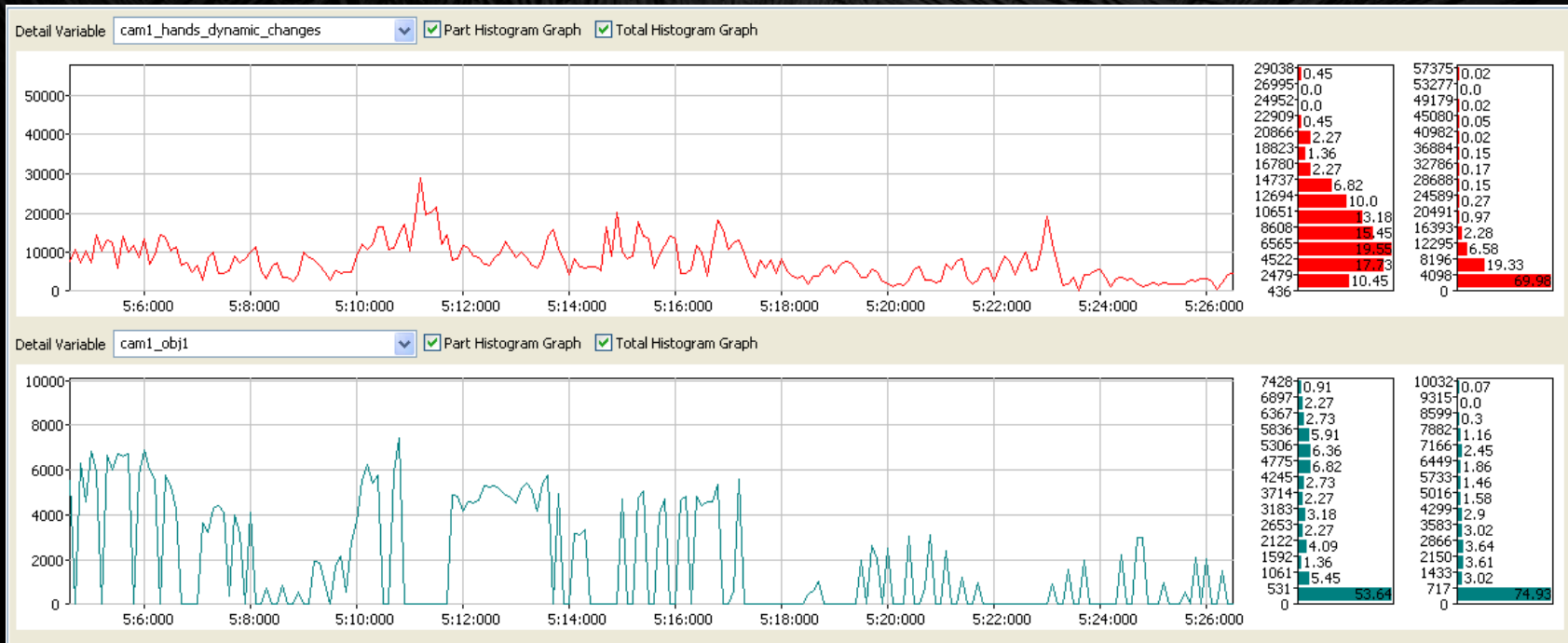


variable list

overview panel

zoom box

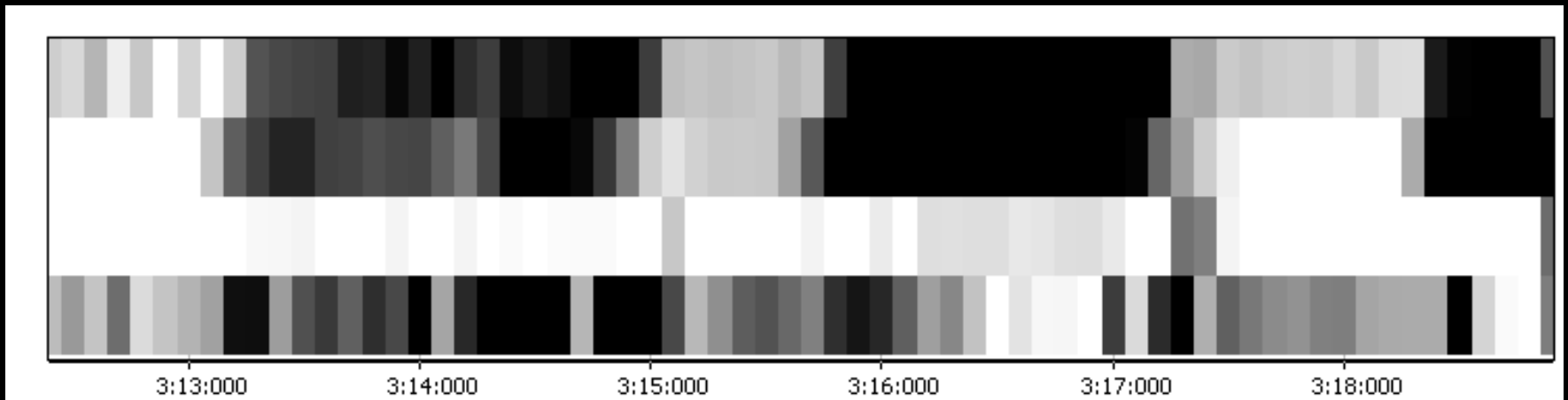
Continuous Temporal Series



Gray-level representation of multiple time series

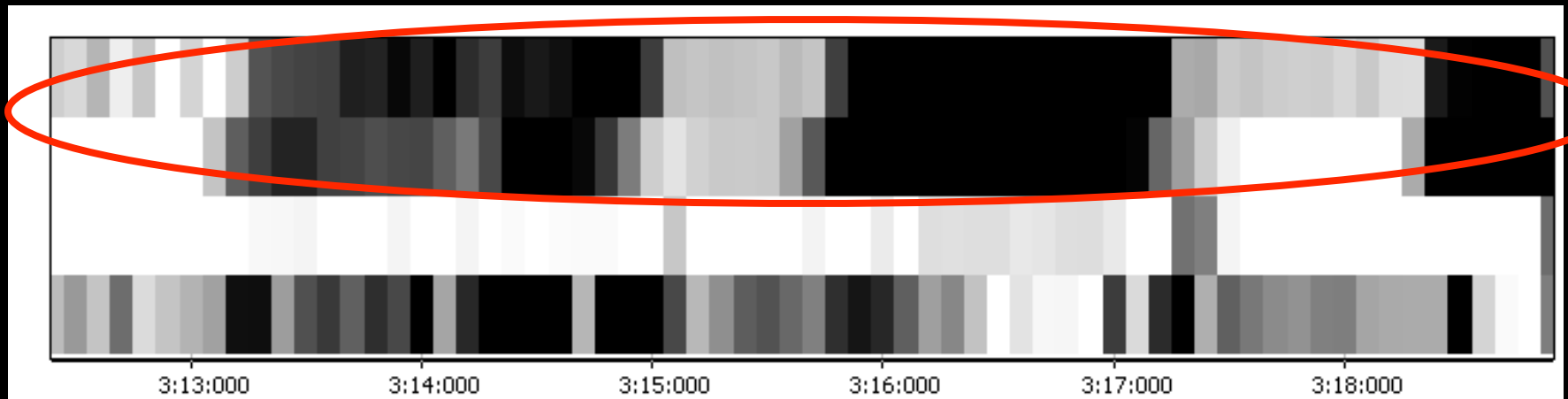


- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.



Gray-level representation of multiple time series

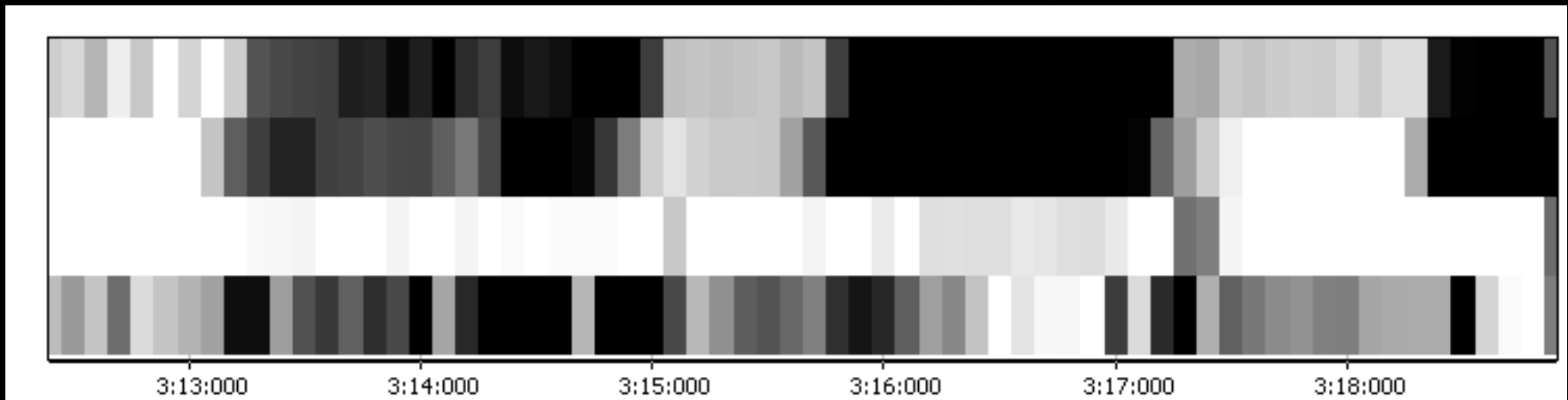
- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.



Gray-level representation of multiple time series

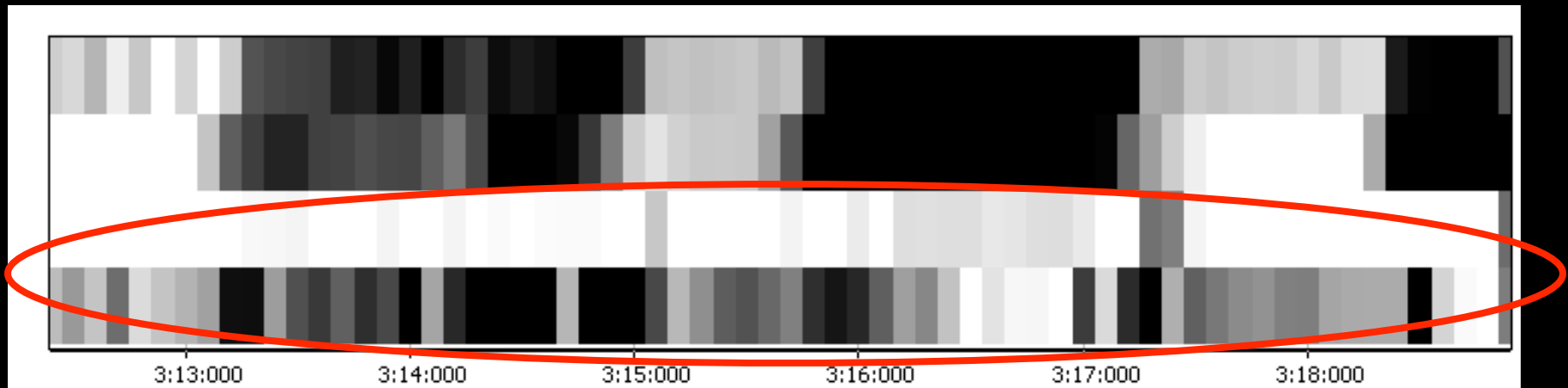


- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.



Gray-level representation of multiple time series

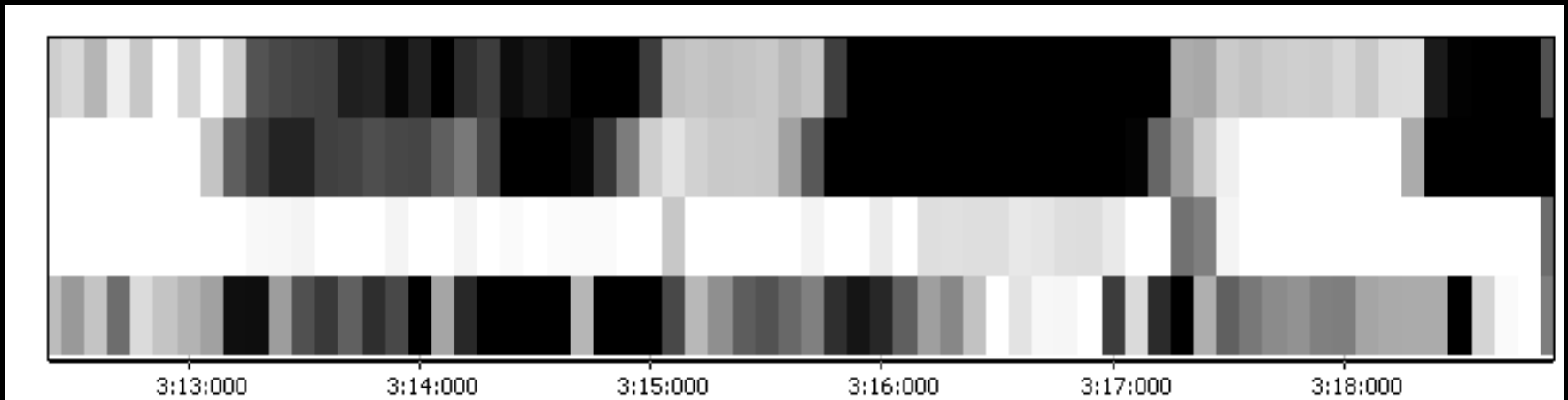
- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.



Gray-level representation of multiple time series

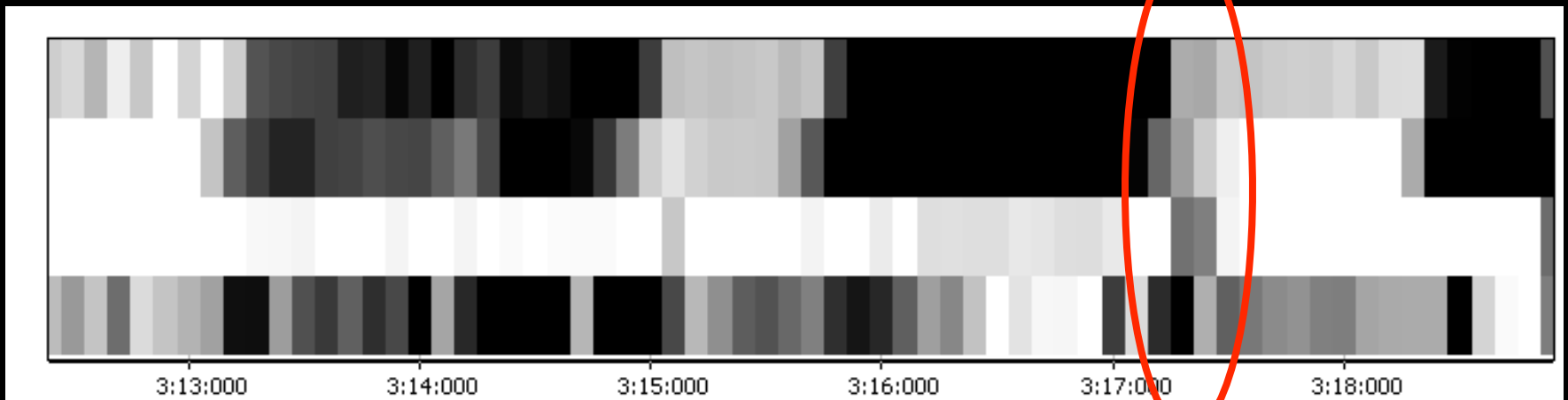


- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.

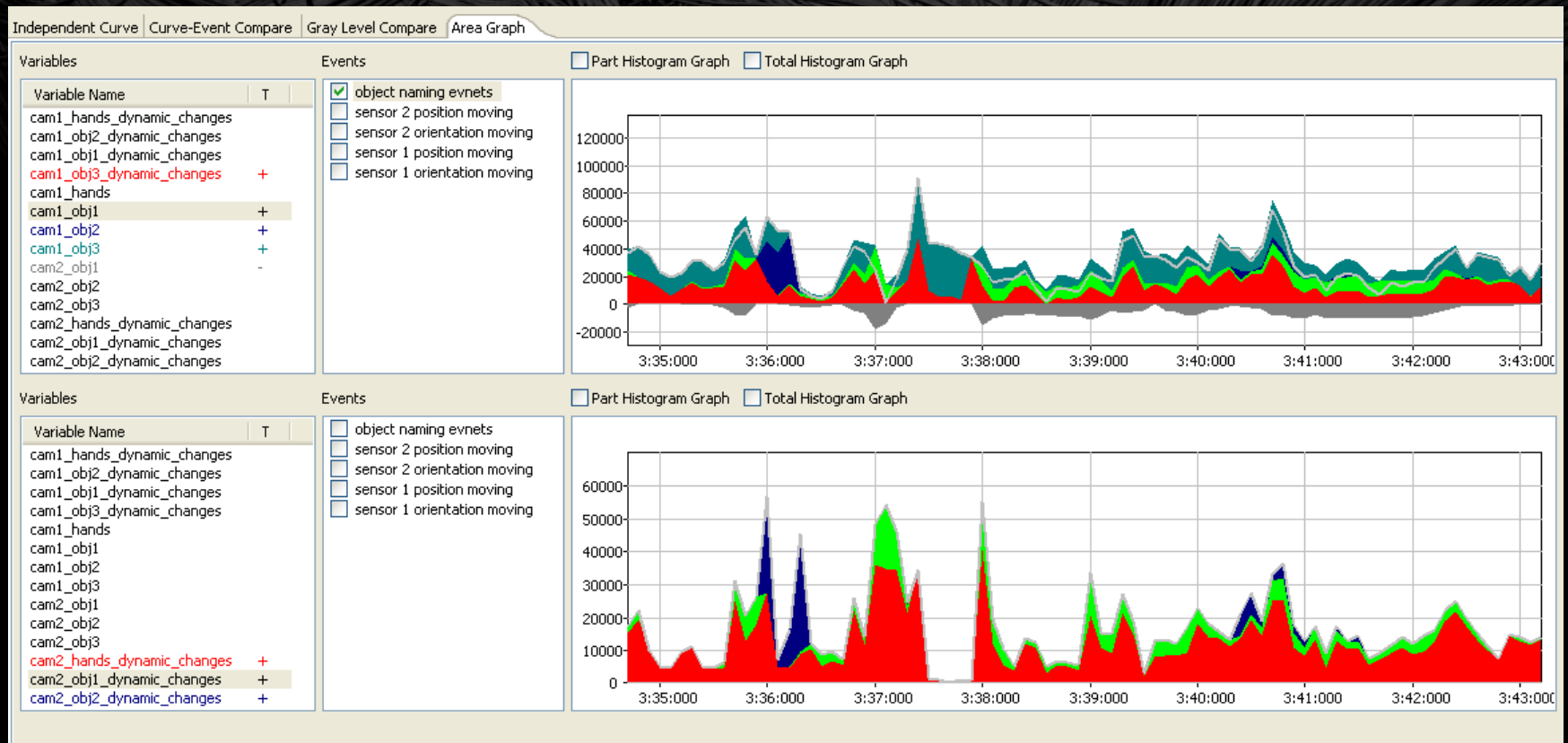


Gray-level representation of multiple time series

- The correlation between multiple data streams.
- Interesting joint patterns across multiple data streams.

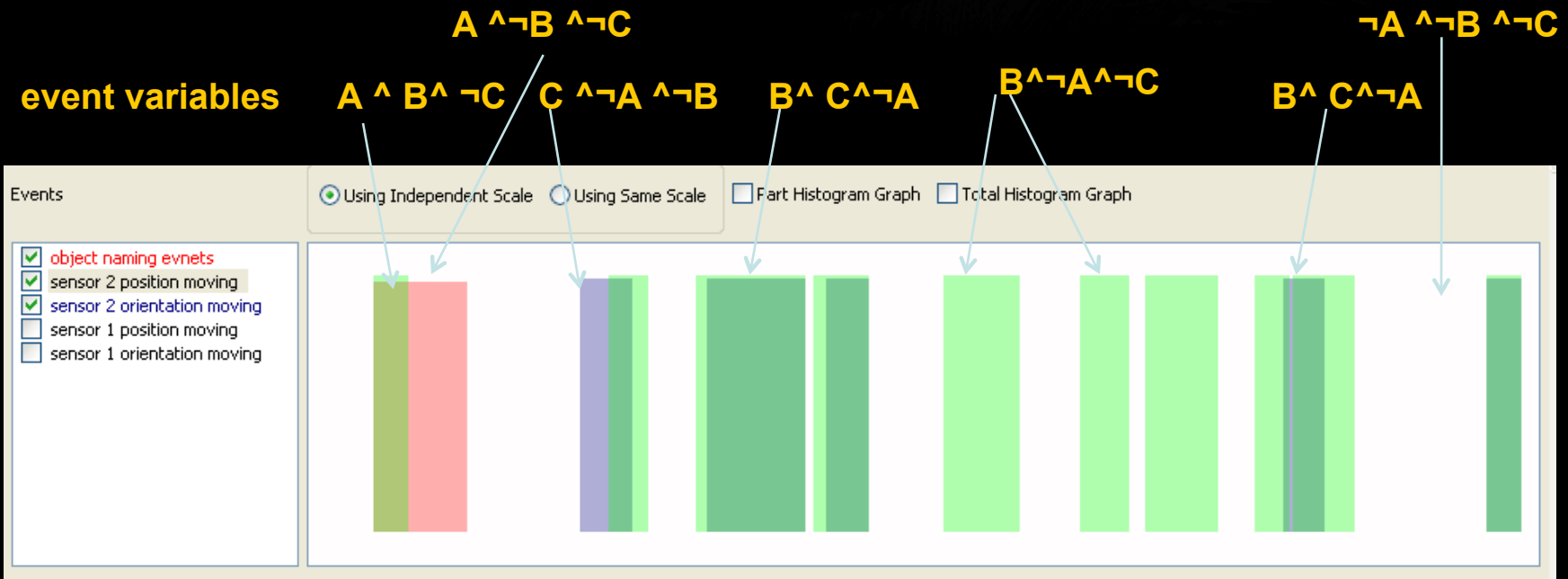


Continuous time series data



Event Data

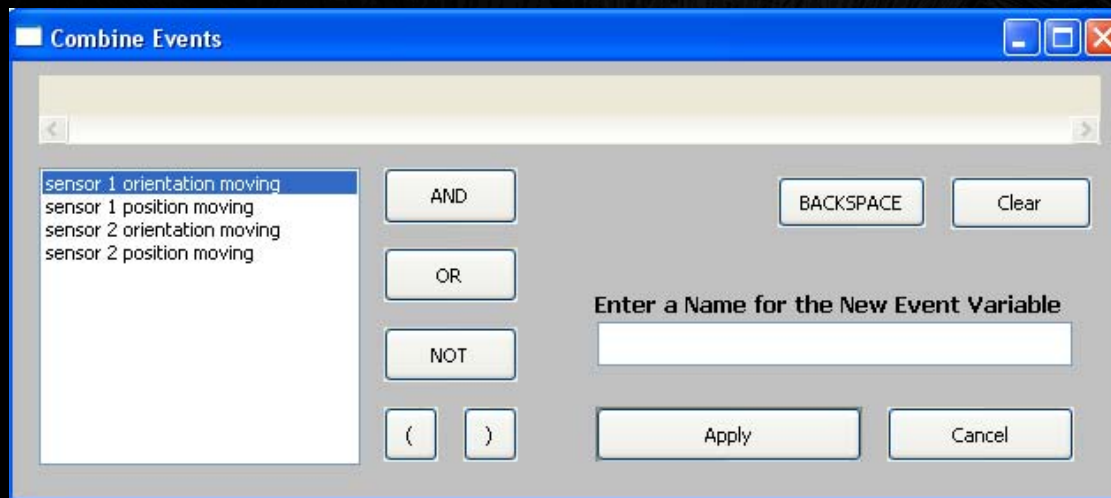
- Single events: frequency, duration, and its periodicity.
- Joint events: $P(A|B)$ and $P(A \wedge B)$.
- Causality and Prediction.



High-order Probabilities



- User-defined events to create new event variables
- AND, OR and NOT



Concurrent visualization of continuous and event variables



continuous variables event variables

event variables (color) overlapped on
continue variables (gray level)



We are interested in exploring the potential complex patterns hidden in continuous variables **conditioned on event variables** – what trends and patterns exist in the continuous variables when certain events happen.

Event-based Interactive Visual Exploration

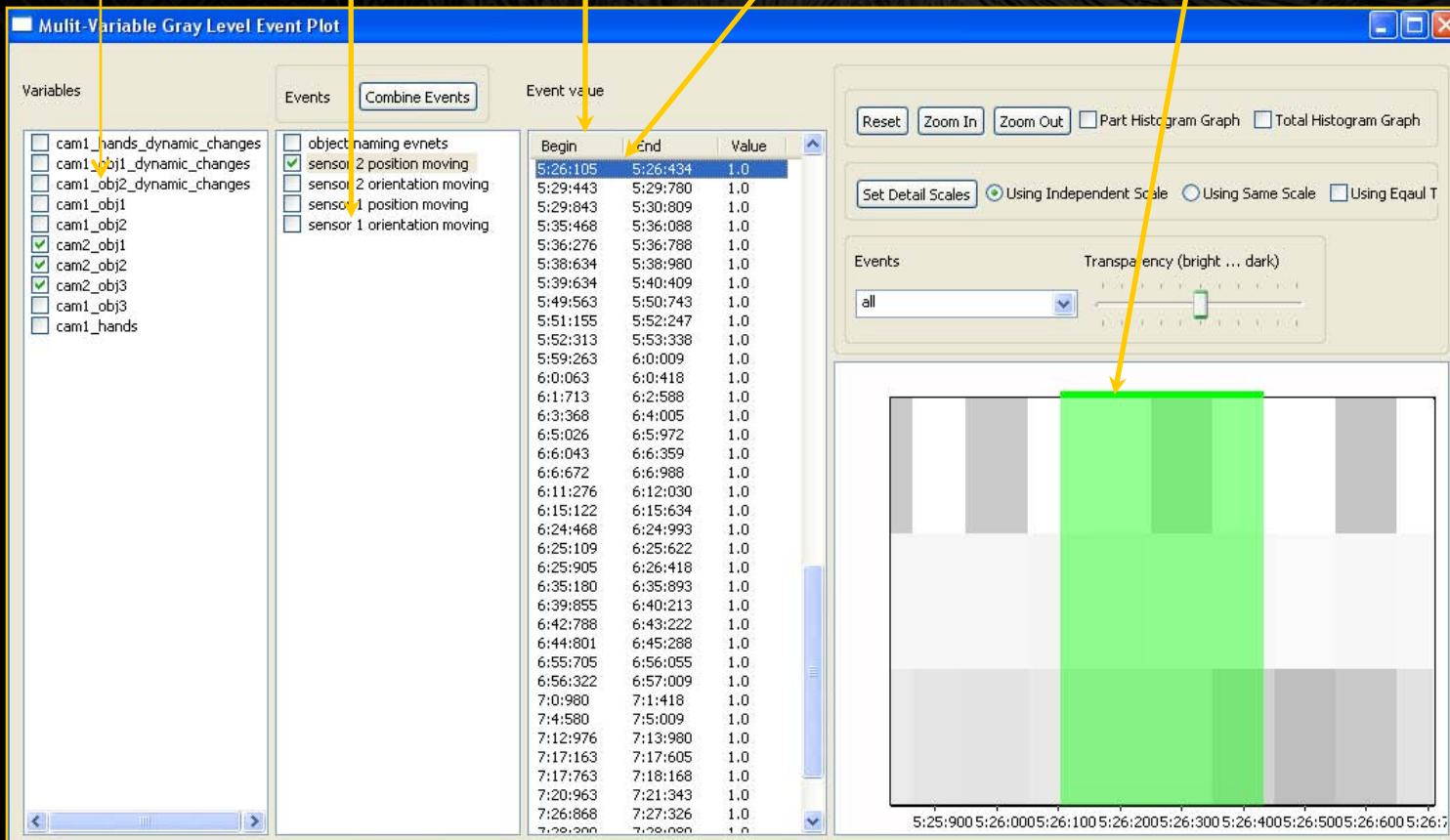
cont.
variable
list

event
variable
list

instances
of the
current
event

onset and
offset
timestamps

the
current
instance



Event Grouping



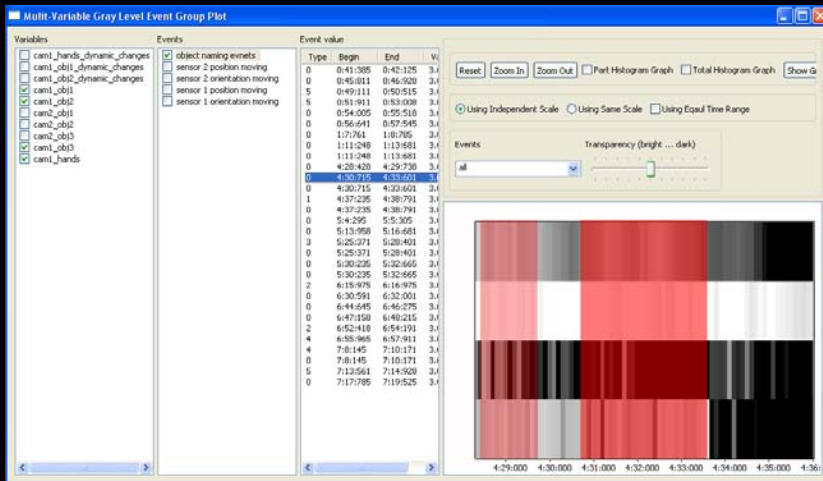
cont.
variable list
event variable list

instances of the current event
current instance

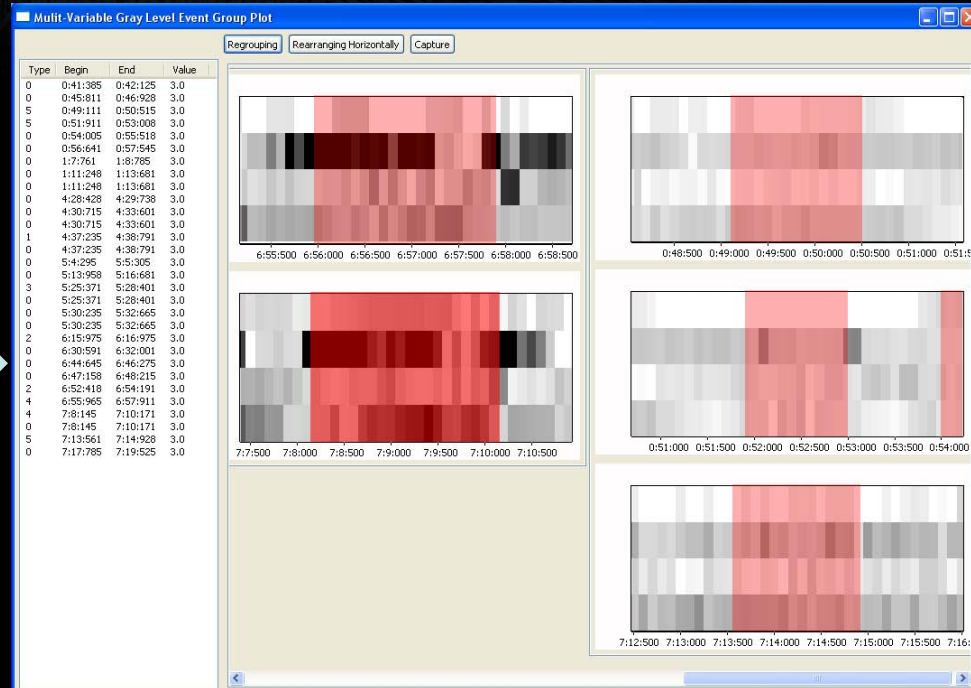
instances of the current event

two instances in Group 1

three instances in Group 2



assign a group number



event grouping results

Multimedia Data → Derived Data



- We notice that derived data are extracted from raw multimedia data and can only include some information available in the raw multimedia data. Therefore, it is important that users can refer to the raw multimedia data while exploring derived data. We design and implement one critical component to connect multimedia playback with visual data mining.
- By visually exploring the current derived data and meanwhile examining the corresponding raw data, users can gain insights about what **additional variables** are missing and should be extracted from raw data.

Visual Exploration ↔ Data Mining



- In order to **quantify** and extend observations, researchers need to develop and use data mining algorithms and statistics to extract and measure the patterns detected in visual exploration.
- To increase the flexibility to be compatible with data mining, our system allows users to use any programming language to obtain new results, as far as the output files follow pre-defined formats for either continuous or event variables.
- We propose a tight **loop** between visual exploration and data mining. The insights gleaned from visualization can be used to guide further data mining. Meanwhile, the results from the next round of data mining can be visualized which allows users to obtain new insights and develop more hypotheses with the data.

Summary



- Decompose and represent multimedia data into continuous and event variables.
- Allow users to examine both raw and derived data.
- Develop various ways to visualize and explore data. E.g. Event-driven exploration.
- Allow visualization and data mining to bootstrap each other.
- Allow users to not only easily examine and synthesize information into new ideas, but also quickly quantify and test the insights gained from visualization.

Acknowledgement



Thanks!

The research was supported by NIH RO1 HD056029 and NSF BCS0544995.

Large Display

