Data Mining

Associate Professor Dr. Raed Ibraheem Hamed

University of Human Development,
College of Science and Technology
Department of Computer Science

2016 – 2017

Introduction

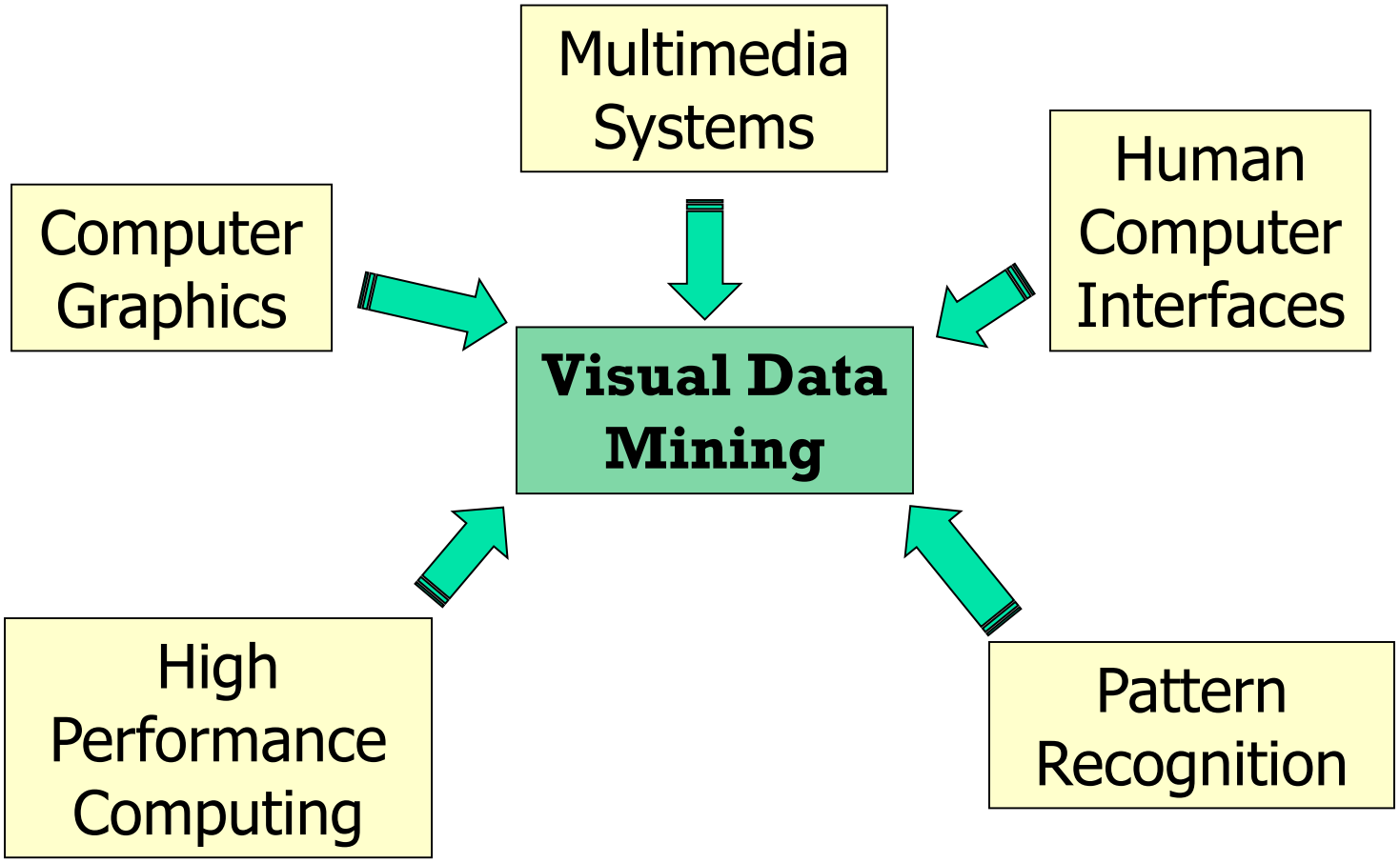
- Visual Data Mining?
- Purpose of Visualization for Data Mining
- Visual Data Mining & Data Visualization
- Data Mining Result Visualization
- Data Mining Process Visualization
- Visual Classification
- Visualization of Data Mining Processes by Clementine
- Methods of Data Visualization

What is Visual Data Mining

- Visualization: Use of computer graphics to create visual images which aid in the understanding of complex, often massive representations of data.

- Visual Data Mining: Visual Data Mining presents the data in some visual form, allowing users to mine and gain insight into the data, draw conclusions and directly interact with the data.

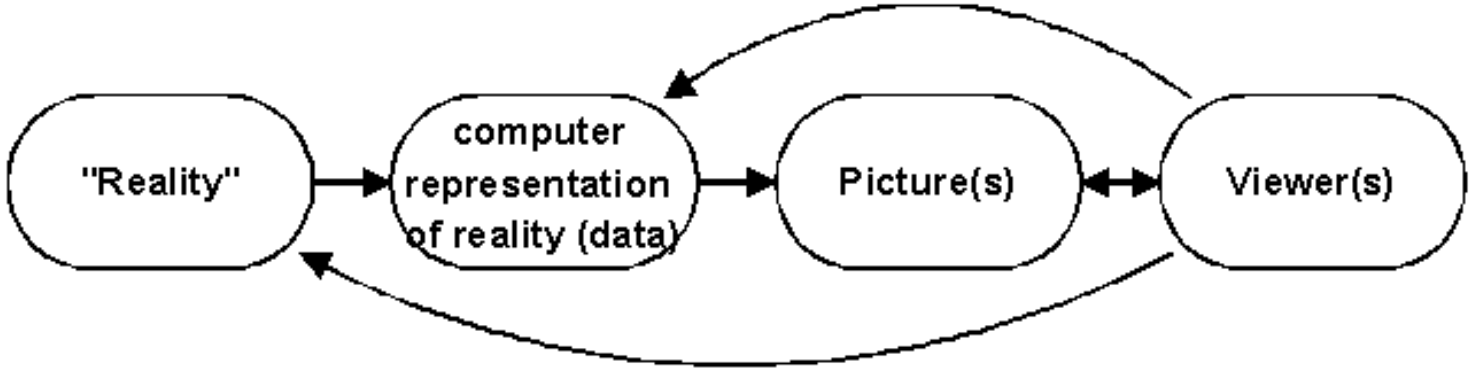
Visual Data Mining is Closely Related to :



Purpose of Visualization for Data Mining

1. Gain insight into an information space by mapping data onto graphical primitives
2. Provide qualitative overview of large data sets
3. Search for patterns, trends, structure, irregularities, relationships among data.
4. Help find interesting regions and suitable parameters for further quantitative analysis.
5. Provide a visual proof of computer representations derived

Computer Representations Of Reality



As depicted by the above figure, visualization is essentially a mapping process from computer representations to perceptual representations, choosing techniques to maximize human understanding. The goal of a viewer might be a deeper understanding of physical phenomena, but it also might be a visual proof of computer representations derived from such an initial stage.

Data Mining Algorithms Verses Visualization

	Data Mining Algorithms	Visualization
Actionable	+	–
Evaluation	+	–
Flexibility	–	+
User Interaction	–	+

Visual Data Mining & Data Visualization

- Integration of visualization and data mining
  - data visualization
  - data mining result visualization
  - data mining process visualization
- Data visualization
  - Data in a database or data warehouse can be viewed
    - at different levels of abstraction
    - at different combinations of attributes or dimensions
  - Data can be presented in various visual forms

Data Mining Process Visualization

- Presentation of the various processes of data mining in visual forms so that users can see
  - Data extraction process
  - Where the data is extracted
  - How the data is cleaned, integrated, preprocessed, and mined
  - Method selected for data mining
  - Where the results are stored
  - How they may be viewed

Data extraction is the process of retrieving data out of data sources

VDM Approach

VDM takes advantage of both,

- The power of automatic calculations, and
- The capabilities of human processing.
  - Human perception offers phenomenal abilities to extract structures from pictures.

Levels of VDM

- No or very limited integration
  - Corresponds to the application of either traditional information visualization or automated data mining methods.
- Loose integration
  - Visualization and automated data mining methods are applied sequentially.
  - The result of one step can be used as input for another step.
- Full integration
  - Automated data mining and visualization methods applied in parallel.
  - Combination of the results.

Methods of Data Visualization

Different methods are available for visualization of data based on type of data

Data can be

- Univariate
- Bivariate
- Multivariate

Univariate data

- Measurement of single quantitative variable
- Characterize distribution
- Represented using following methods
  - Histogram
  - Pie Chart

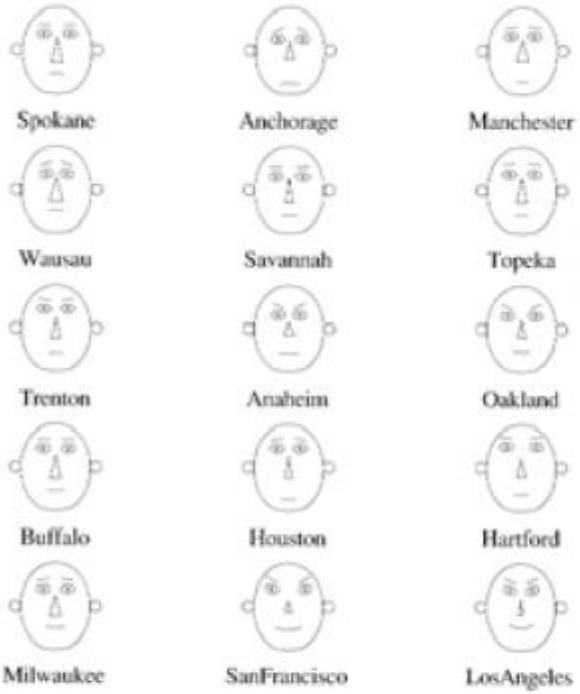
Bivariate Data

- Constitutes of paired samples of two quantitative variables
- Variables are related
- Represented using following methods
  - Scatter plots
  - Line graphs

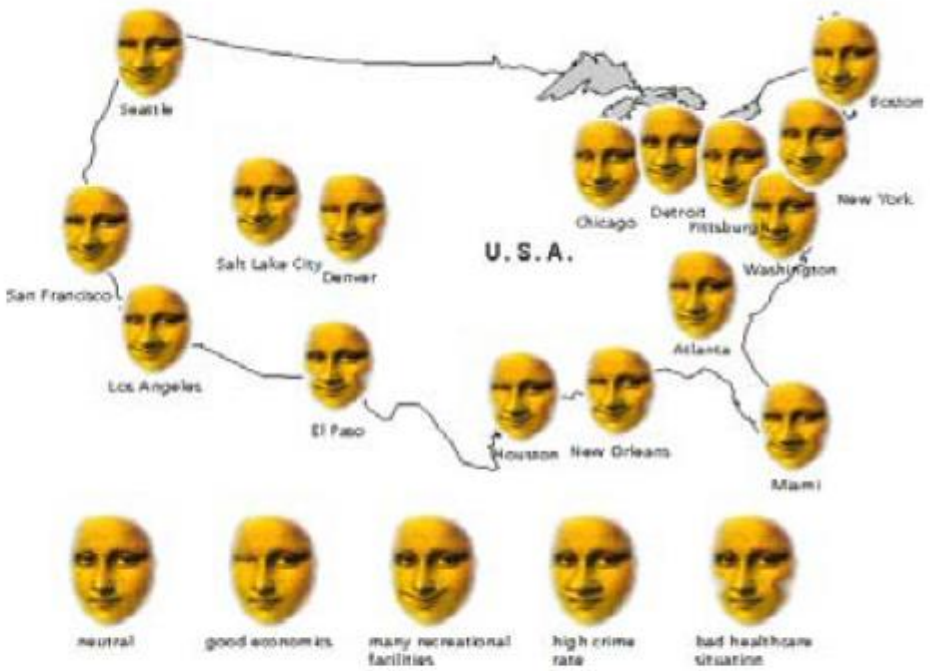
Multivariate Data

- Multi dimensional representation of multivariate data
- Represented using following methods
  - Icon based methods
  - Pixel based methods
  - Dynamic parallel coordinate system

Icon based Methods



Visualizing town data with Chernoff-faces, (Spence, 2001)

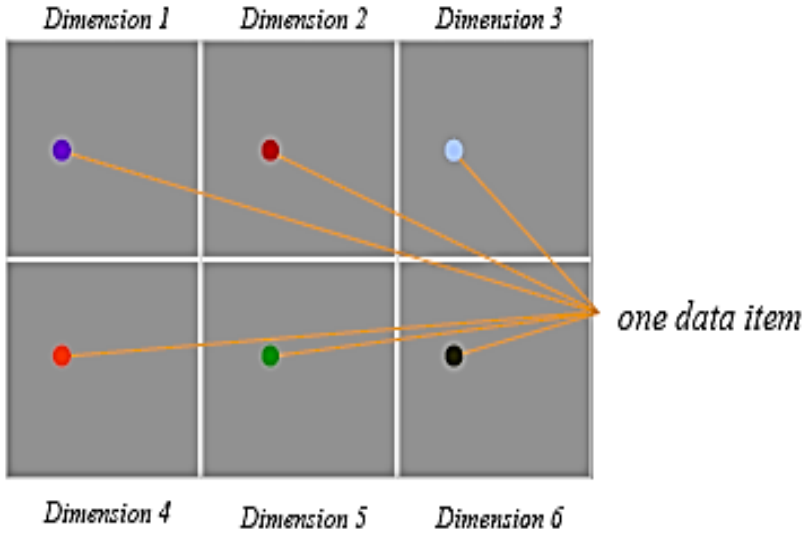


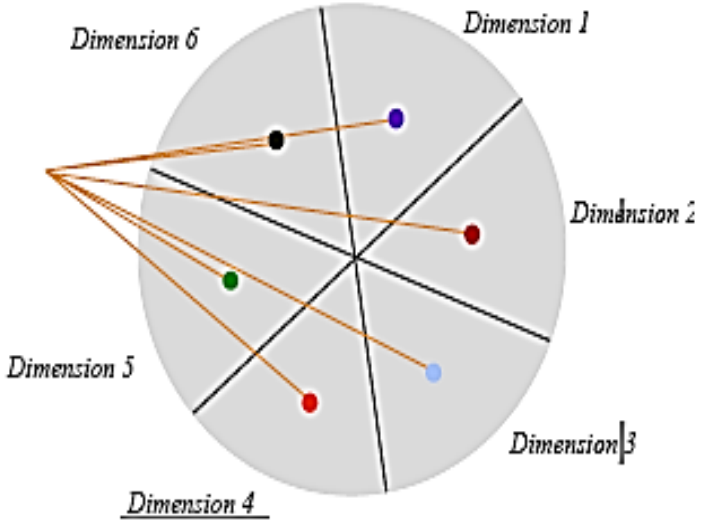
Visualization of town data using morphed faces (Alexa, Müller, 1998)

Dense Pixel Display

Approach:

- Each attribute value is represented by one colored pixel.






Thank You