Course Information

• **Location:** M 113
• **Time:** 1~2:30PM Tu/Th
• **Instructor Office Hours:** 2:30~3:30PM Tu /Th
  or by appointment via emails
• **Instructor Office:** PGH 566

🌟 **Course webpage:**
http://www2.cs.uh.edu/~chengu/Teaching/Fall2019/Visualization_fall2019.html

• **TA Info:**
  – Lieyu Shi,
  – Office: PGH 313 (TA office),
  – Office hours: TBD,
Course Information

• Prerequisites:
  – *Knowledge and experience in C/C++ programming (Javascript?)
  – *Knowledge on data structure and algorithm design (required)
  – Knowledge on linear algebra, calculus, geometry.
  – Experience on computer graphics and OpenGL is a plus but not required

• Textbook: (recommended, NOT required)

  – Reading materials and course notes given in class or on course page
Course Information

• **Grading:**
  – 7 assignments (45%)
  – 1 mid-term exam and a final open-note quiz (25%)
  – 1 final project and presentation (25%)
  – course participation (5%)

• **Late Policy:** Late assignments will be marked off 20% for *each weekday* that it is late. You can ask for a delay for *one* assignment for up to *5 weekdays*.

• **Academic Dishonesty:** Do your own work! *No code copy!*
# Course Schedule (tentative)

<table>
<thead>
<tr>
<th>TIMELINE</th>
<th>MATERIAL COVERED</th>
</tr>
</thead>
</table>
| WEEK 1 (08/20, 22) | **Class Preliminaries**  
Introduction – History of visualization, Visualization pipeline; Data type and data representation |
| WEEK 2 (08/27, 29)   | Visual perceptions; Principles of effective plots (Assignment 1 out);          |
| WEEK 3 (09/03, 05)    | Colors in visualization; OpenGL Tutorial (Assignment 2 out);                   |
| WEEK 4 (09/10, 12)    | Color coding for scalar data visualization; Iso-contouring/ Iso-surfacing; Direct (Assignment 3 out) |
| WEEK 5 (09/17, 19)    | Direct Volume rendering – Ray casting and Splatting                             |
| WEEK 6 (09/24, 26)    | Transfer function design (Assignment 4 out); Vector field/flow visualization introduction; |
| WEEK 7 (10/01, 03)    | Flow visualization techniques in 2D ‐arrow and color plots, streamlines, texture‐based (Assignment 5 out) |
| WEEK 8 (10/08, 10)    | Vector field feature‐based visualization                                         |
| WEEK 9 (10/15, 17)    | 3D vector field visualization; unsteady flow visualization (Assignment 6 out)     |
| WEEK 10 (10/22, 24)   | Mid‐term review; **Mid-term exam** (IEEE Visualization 2019)                   |
| WEEK 11 (10/29, 10/31)| Tensor data application – introduction and math; Geometric‐based and texture‐based tensor visualization; |
| WEEK 12 (11/05, 07)   | Glyph‐based technique (Assignment 7 out); Information visualization – graph and hierarchy data visualization |
| WEEK 13 (11/12, 14)   | Information visualization – high dimensional data visualization;                |
| WEEK 14 (11/19, 11/21)| **Final project presentations**                                                |
| WEEK 15 (11/27)       | Multifield visualization or other advanced topics, **Thanks giving holiday**    |
Background
Data Explosion

DATA NEVER SLEEPS 5.0

How much data is generated every minute?

With data and information now crawling the internet and social media in an endless loop, it’s important to always stay informed and informed about the data generated every minute. What would you do with this information? Would you make use of it in your business or life? How can you make use of this data to improve your decision-making in any way?
How to make use of them for knowledge discovery and decision making?
Data Mining or Visualization?

Data mining is the process of discovering patterns and knowledge in large data sets involving methods at the intersection of machine learning, statistics, and database systems.

- groups of data records (cluster analysis)
- unusual records (anomaly detection)
- dependencies (association rule mining, sequential pattern mining)

From Wikipedia
Data Mining or Visualization?

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.
Let us start with a simple example
can you tell me the trend of these numbers from left to right?
Cognitive study has shown that human visual system is the most effective channel to transport information to the brain. Visual representation is one effective way to convey information. Leading to the introduction of **visualization**.
Which gender and income group of people has different trend of triglyceride (the percentage of fat) level over years from the others?

<table>
<thead>
<tr>
<th>Income\Age</th>
<th>Male &lt;65</th>
<th>Male 65 and above</th>
<th>Female &lt;65</th>
<th>Female 65 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-$24,999</td>
<td>250</td>
<td>200</td>
<td>375</td>
<td>550</td>
</tr>
<tr>
<td>$25,000+</td>
<td>430</td>
<td>300</td>
<td>700</td>
<td>500</td>
</tr>
</tbody>
</table>
History of Visualization

• Visualization is actually rather old


• Often an intuitive step: graphical illustration

http://www.visualinformation.info/old-international-cables-map-year-1924/
Now, What is Visualization?

• In 1987
  – the National Science Foundation (of the U.S.) started “Visualization in scientific computing” as a new discipline, and a panel of the ACM coined the term “scientific visualization”
  – Scientific visualization, briefly defined: The use of computer graphics for the analysis and presentation of computed or measured scientific data.

  – to form a mental vision, image, or picture of (something not visible or present to the sight, or of an abstraction); to make visible to the mind or imagination

• Visualization transforms data into images that effectively and accurately represent information about the data.
  – Schroeder et al. The Visualization Toolkit, 2nd ed. 1998
Now, What is Visualization?

• In 1987
  – the National Science Foundation (of the U.S.) started “Visualization in scientific computing” as a new discipline, and a panel of the ACM coined the term “scientific visualization”
  – Scientific visualization, briefly defined: The use of computer graphics for the analysis and presentation of computed or measured scientific data.

  – to form a mental vision, image, or picture of (something not visible or present to the sight, or of an abstraction); to make visible to the mind or imagination

• Visualization transforms data into images that effectively and accurately represent information about the data.
  – Schroeder et al. The Visualization Toolkit, 2nd ed. 1998

Tool/Means to enable a User insights into Data via Visual Representation
(or an understanding of the story behind the data)
Why is Visualization Important?

- To effectively convey information to data stakeholders

Table 7.2: Direct global warming potentials of several well-mixed trace gases relative to CO₂. The GWPs of the various non-CO₂ species are calculated for each of five time horizons (20, 50, 100, 200 and 500 years) using, as in IPCC, the carbon cycle model of Seganthaler (1988). (Note that IPCC contained a typographical error which led to incorrect values for the direct GWP of methane.)

<table>
<thead>
<tr>
<th>Gas</th>
<th>Lifespan (years)</th>
<th>20 years</th>
<th>50 years</th>
<th>100 years</th>
<th>200 years</th>
<th>500 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CH₄</td>
<td>10.3</td>
<td>35</td>
<td>19</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>N₂O</td>
<td>132</td>
<td>260</td>
<td>270</td>
<td>270</td>
<td>240</td>
<td>170</td>
</tr>
<tr>
<td>CPC-11</td>
<td>53</td>
<td>4500</td>
<td>4100</td>
<td>3400</td>
<td>2400</td>
<td>1400</td>
</tr>
<tr>
<td>CPC-12</td>
<td>116</td>
<td>7100</td>
<td>7400</td>
<td>7100</td>
<td>6200</td>
<td>4100</td>
</tr>
<tr>
<td>HFC-122</td>
<td>15.3</td>
<td>4200</td>
<td>2800</td>
<td>1600</td>
<td>970</td>
<td>340</td>
</tr>
<tr>
<td>CPC-113</td>
<td>190</td>
<td>4000</td>
<td>4700</td>
<td>4500</td>
<td>3900</td>
<td>2500</td>
</tr>
<tr>
<td>CPC-114</td>
<td>220</td>
<td>6100</td>
<td>5700</td>
<td>7000</td>
<td>7000</td>
<td>5800</td>
</tr>
<tr>
<td>CPC-115</td>
<td>550</td>
<td>5500</td>
<td>6200</td>
<td>7000</td>
<td>7800</td>
<td>8500</td>
</tr>
<tr>
<td>HCFC-123</td>
<td>1.71</td>
<td>330</td>
<td>150</td>
<td>90</td>
<td>55</td>
<td>30</td>
</tr>
<tr>
<td>HCFC-124</td>
<td>6.9</td>
<td>1300</td>
<td>760</td>
<td>440</td>
<td>270</td>
<td>150</td>
</tr>
<tr>
<td>HFC-125</td>
<td>40.5</td>
<td>5200</td>
<td>4900</td>
<td>3400</td>
<td>2200</td>
<td>1200</td>
</tr>
<tr>
<td>HFC-134a</td>
<td>15.6</td>
<td>3100</td>
<td>1900</td>
<td>1200</td>
<td>750</td>
<td>400</td>
</tr>
<tr>
<td>HFC-141b</td>
<td>10.8</td>
<td>1800</td>
<td>980</td>
<td>580</td>
<td>250</td>
<td>200</td>
</tr>
<tr>
<td>HFC-142b</td>
<td>22.4</td>
<td>4000</td>
<td>2800</td>
<td>1800</td>
<td>1100</td>
<td>620</td>
</tr>
<tr>
<td>HPC-143a</td>
<td>24.2</td>
<td>4700</td>
<td>4500</td>
<td>3800</td>
<td>2800</td>
<td>1600</td>
</tr>
<tr>
<td>HFC-152a</td>
<td>1.8</td>
<td>530</td>
<td>250</td>
<td>150</td>
<td>89</td>
<td>49</td>
</tr>
<tr>
<td>CCCL</td>
<td>67</td>
<td>1800</td>
<td>1660</td>
<td>1380</td>
<td>860</td>
<td>480</td>
</tr>
<tr>
<td>CH₃Cl</td>
<td>6.1</td>
<td>360</td>
<td>170</td>
<td>100</td>
<td>62</td>
<td>34</td>
</tr>
<tr>
<td>CP,Br</td>
<td>77</td>
<td>5600</td>
<td>5300</td>
<td>4900</td>
<td>3800</td>
<td>2300</td>
</tr>
</tbody>
</table>

SAOD Table 7.2 (p. 7.6)
Why is Visualization Important?

• To effectively convey information to data stakeholders
• To make invisible phenomena visible

Astrophysics
source: VACET

Combustion
source: VACET

Automobile design [Chen et al. Vis11]
Why is Visualization Important?

• To effectively convey information to data stakeholders.
• To make invisible phenomena visible.
• Create intuitive representation for abstract concepts

$$\frac{dy}{dx} = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

Fusion physics
What Does Visualization Do?

- Three types of **goals** for visualization
  - ... to **present**
    - “everything” known about the data,
    - Vis. used for **Communication of Results**
What Does Visualization Do?

- **Three types of goals for visualization**
  - **... to present**
    - “everything” known about the data,
    - Vis. used for Communication of Results
  - **... to analyze**
    - There are hypotheses (expect to see...),
    - Vis. used for **Verification or Falsification**
What Does Visualization Do?

• Three types of goals for visualization
  – ... to present
    • “everything” known about the data,
    • Vis. used for Communication of Results
  – ... to analyze
    • There are hypotheses,
    • Vis. used for Verification or Falsification
  – ... to explore
    • Nothing (or little) is known,
    • Vis. used for **data exploration**/knowledge discovery
• This is a very rich and highly inter-disciplinary area that combines knowledge from various disciplines (including the application areas and many sub-fields in computer science).
Evolution of Visualization Research

• From **direct** visual representation of the data to visualize **derived** information.

• From **simple** data forms (e.g., scalar values) to **more complex** forms (tensor, multi-modal data, etc.).

• From **represent** the data to **reveal new findings** (**explore!!!**).

• From **scientific visualization** to information visualization, to bio-visualization, geographical data visualization, multi-modal (heterogeneous) data and beyond.
Types of Data and Their Visualizations
SciVis vs. InfoVis

• Scientific visualization is mostly concerned with:
  – 2, 3, 4 dimensional (i.e., physical space+time) data
  – discretized or sampled data (for continuous function)

• Information visualization focuses on:
  – high-dimensional, abstract data
  – discrete data
  – financial, statistical, etc.
  – visualization of large trees, networks, graphs
  – data mining: finding patterns, clusters, voids, outliers
SciVis vs. InfoVis

- Information visualization focuses on:
  - high-dimensional, **abstract** data (no natural coordinates!)
  - inherently **discrete** data
    - financial, statistical, etc.
    - visualization of large trees, networks, graphs
  - data mining: finding patterns, clusters, voids, outliers
Age of Big Data

DATA NEVER SLEEPS 5.0

How much data is generated every minute?

The world of data

The world's internet population has grown from 3.5 billion in 2013 to 3.7 billion today.

Google conducts
3.607 billion
searches

Buzzfeed processes
5,925.92
millions

Venmo processes
51,892
transactions

Uber riders take
547,854 trips

Spotify
13 million

Americans use
2,657,700 GB
of internet data

Will reach 350,000,000,000 by 2020

2012
2013
2014
2015
2016
2017

Global internet population growth 2010-2017

Learn more at domo.com

We are not concerned with the big data issue in this course but rather building the foundation that could be used to eventually address the big data challenges!!!
Goals and Topics
Goals of this Course

• Know the types of visualization problems! Does it belong to scientific visualization, information visualization or their combination?

• Know the representative/standard methods for some classic visualization problems.

• Be able to develop the customized visualization techniques and systems for practical and research needs.
After this course, you can be a developer of visualization systems not just a user
Topics to Be Covered in this Course

• What will be covered?
  – Colors, perception, plots
  – Scalar field visualization
  – Vector field visualization
  – Tensor field visualization
  – Information visualization (graphs, high-dimensional, etc.)
  – Hop topics in visualization

• What mathematical knowledge will be needed?
  – Calculus, trigonometry, linear algebra, discrete math, topology, dynamical systems, numerical algebra, etc.

• What knowledge from the other sub-fields of computer science will be needed?
  – Algorithms, data structures, C++ programming, computer graphics, human-computer interaction, imaging processing, etc.

• Anything else?
  – Art and design, psychophysics, scientific computing, other domain related knowledge
Preview of Topics
Why visualization can easily go wrong?

http://yifanhu.net/GALLERY/GRAPHS/

Images from the storytelling with data
Perception

Graphs and Charts

Abstract

Bad graphs vs. good graphs

Images from the storytelling with data
Graphs and Charts

FOUR QUESTIONS

1. What data is important to show?
2. What do I want to emphasize in the data?
3. What options do I have for displaying this data?
4. Which option is most effective in communicating the data?
Colors

• Color theory and visual representation

HSV Color Wheel

How to use colors effectively?

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

<table>
<thead>
<tr>
<th>Country</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUS</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>BRA</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>CAN</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>CHI</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>FRA</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>GER</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>IND</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>ITA</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>MEX</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>NLD</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>SPA</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>TUR</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>US</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Images from the storytelling with data
Scalar Field Visualization

• Applications
  – Medical
  – Physics
  – Engineering
  – Climate
  – Environment
  – ......

• 2D/2.5D

Global temperature 2008

Global temperature average
Scalar Field Visualization

- 3D

Iso-surfacing

Microvasculature

Volume Rendering
Vector Field Visualization

• Applications
  – Medical
  – Physics
  – Engineering
  – Climate
  – Environment
  – Oceanography
  – ……

• 2D

Texture-based
Geometry-based
Feature-based
Vector Field Visualization

- 3D

Feature-based

Geometry-based

Texture-based

Hurricane Isabel
Tensor Field Visualization

• Applications
  – Physics
  – Engineering
  – Medical
  – ......

• A rather complex but powerful notation

\[ T = \begin{bmatrix} T_{00}(x, y, z) & \ldots \end{bmatrix} \]

Second order tensor
Tensor field visualization

Texture-based

Geometry-based

Glyph-based
Information visualization

Graph visualization http://gephi.org

High-dimensional data visualization http://ncva.itn.liu.se/great-statistics-visualization/

Text visualization http://www.poynter.org

Tree/hierarchy data visualization http://infosthetics.com
Other hot topics in visualization

– Multi-field visualization
– Uncertainty visualization
– Big data visualization
– Visualization in practice
– New applications