Graphics & Visualization

Chapter 10

Visualization Principles
Introduction

• If you unexpectedly see a picture of a person you care about, you feel the love you have for them.
  - Information flows from your visual system through your brain.
  - Artificial stimuli can produce similar effects (visual stimuli most effective)

• Modern scientific experiments & simulations produce vast amounts of data.
  - The nature of data is symbolic: hard for humans to comprehend.
  - A picture is worth a thousand words!!

<table>
<thead>
<tr>
<th>23</th>
<th>24</th>
<th>25</th>
<th>27</th>
<th>26</th>
<th>25</th>
<th>25</th>
<th>24</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>26</td>
<td>28</td>
<td>30</td>
<td>29</td>
<td>27</td>
<td>26</td>
<td>28</td>
<td>31</td>
</tr>
<tr>
<td>26</td>
<td>28</td>
<td>29</td>
<td>31</td>
<td>32</td>
<td>29</td>
<td>30</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>30</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>32</td>
<td>34</td>
<td>35</td>
<td>37</td>
<td>41</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>31</td>
<td>33</td>
<td>36</td>
<td>38</td>
<td>40</td>
<td>42</td>
<td>43</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>32</td>
<td>32</td>
<td>35</td>
<td>37</td>
<td>41</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>30</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>36</td>
<td>38</td>
<td>41</td>
<td>42</td>
<td>44</td>
</tr>
<tr>
<td>32</td>
<td>34</td>
<td>27</td>
<td>29</td>
<td>40</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>22-25</th>
<th>26-29</th>
<th>30-33</th>
</tr>
</thead>
<tbody>
<tr>
<td>34-37</td>
<td>38-41</td>
<td>42-45</td>
</tr>
</tbody>
</table>
Introduction (2)

• Applications of visualization can be:
  ■ Exploration of large **acquired** data sets, e.g.:
    ♦ Medical data
    ♦ Oil and gas data
    ♦ Weather data
  ■ Exploration of large **simulation** data sets, e.g.:
    ♦ Engineering simulations
    ♦ Meteorological forecasts
    ♦ Computational fluid dynamics
    ♦ Finance
  ■ In some cases we merge the above categories (e.g. meteorology)
Goals of Visualization

- To increase human understanding of complex data
  - By taking advantage of high-bandwidth human visual channel
  - Using techniques from Computer Graphics
- Transform data into information
- Bring data to life!!
- To be useful, it must be the medium that enables information to be effectively communicated to the user
Definitions of Visualization

• Given by a 1987 NSF panel:
  ■ Visualization is a method of computing
  ■ It transforms the symbolic into the geometric, enabling researchers to observe their simulations & computations
  ■ Offers a method for seeing the unseen
  ■ Enriches the process of scientific discovery
  ■ Fosters profound and unexpected insights

• From a modern visualization course:
  ■ Visualization is a cognitive process using the powerful information processing & analytical functions of the human vision system
  ■ Flexibility across scales, from interstellar to subatomic
  ■ The fundamental objective is to acquire new knowledge rather than generating pictures

• From P. Hanrahan:
  ■ Conveying information using graphical techniques
Construct model of a phenomenon (description in mathematical framework)

The creation of a model is an iterative process:
- After sufficient observation of phenomenon, an initial model is proposed
- Attempt to validate it with real data \( \rightarrow \) discrepancies often arise
- Correct the parameters of the model or even the model itself

The model can then be used to make predictions
- E.g., Newton’s gravitational model
- Better hardware & more efficient algorithms allow for more complicated (& thus accurate) models, e.g. weather prediction
Methods of Scientific Exploration (2)

- Example of an evolving model: The Illumination Model in CG
  - Initially models consisted of simple depth cuing (1960’s)
  - The Phong model (1970’s):
    - diffuse & specular reflections only
    - no account of the interactions of light between objects
    - assumed a constant ambient illumination value
  - The Ray-tracing & Radiosity models (1980’s):
    - include light-interaction computations
    - produce more photorealistic images
Data Aspects & Transformations

• Sources of visualization data:
  - Experiments
    - External data: produced externally to the visualization system
  - Simulations
    - Internal data (but not always; may be acquired from other sources)

• Original (raw) data vs Derived (processed) data
  - Examples of processing: normalization, filtering
Data Aspects & Transformations (2)

• Properties of data, e.g.:
  ■ Data type
  ■ Sampling domain
  ■ Sampling pattern
  ■ Dimensionality
  ■ Format

• Visualization systems need to provide the user with powerful data-import modules

• The type of data largely determines the kind of visualization algorithm to be applied
Data Aspects & Transformations: Normalization

• The process of converting a given data range into a standard input range.
  ■ Data can assume arbitrary ranges
  ■ Visualization packages require a standard input range (e.g. \([0.0, 1.1]\)) e.g. for a standard color map
• Normalization functions are often linear
• If \(i_{\text{min}}, i_{\text{max}}\): the minimum / maximum input data values
  \(i\): an arbitrary input data value
• we can linearly normalize \(i\) into the range \([n_{\text{min}}, n_{\text{max}}]\) as:

\[
  i_{\text{norm}} = \frac{i - i_{\text{min}}}{i_{\text{max}} - i_{\text{min}}} \cdot (n_{\text{max}} - n_{\text{min}}) + n_{\text{min}}
\]
Data Aspects & Transformations: Filtering

- Filtering techniques are used to remove noise, smooth, sharpen, and generally improve the quality of the data.
  - Experimental data in particular may contain noise
- The median filter is an example noise-removal filter
  - Preserves detail
  - Replaces each data value with the median of
    - Itself and
    - Its neighbors within a certain radius
Data Aspects & Transformations: Coregistration

- The process of unifying coordinate systems (cs)
  - Different data sources may produce data in different coordinate systems
  - Apply coordinate transformations to the data to ensure compatibility between:
    - The cs of the source and the visualization system
    - The cs of different sources when co-displaying multi-source data
- Generally uses affine transformations
Coregistration Case Study

MEG signals within a Generic Model Brain

- Display in 3D, magnetoencephalographic (MEG) patient-specific signals within a transparent model of a generic brain.
- Two data sets:
  - The generic brain
  - The MEG signals (displayed using arrow glyphs)
- Must be co-displayed after coregistration
Coregistration Case Study (2)

MEG signals within a Generic Model Brain

- Coregistration:
  1. Establish 2 coordinate systems in the 2 data sets
     - Let the coordinate system of the generic brain model be \( CS_B \)
     - Let the coordinate system of the MEG signals be \( CS_M \)
     - Three non-collinear points define a coordinate system
     - For the \( CS_B \) the points are: \( a_B, p_B, f_B \)
     - For the \( CS_M \) the points are: \( a_M, p_M, f_M \)
  2. Convert one of the data sets to the coordinate system of the other
     - Suppose we transform the MEG data to the generic brain model
     - Suppose the \( a \) points mark the origin of the 2 coordinate systems
     - Suppose the \( \text{ap} \) vectors mark the \( +x \)-axis
     - Suppose the \( f \) points indicate the ‘up’ direction, from which the \( +z \)-axis is derived
Coregistration Case Study (3)

MEG signals within a Generic Model Brain

2. Convert one of the data sets to the coordinate system of the other

- The directions of the 3 axes in each coordinate system are computed as:

\[
\begin{align*}
\vec{f} &= \vec{f} - \vec{a}, \\
\vec{x} &= \vec{p} - \vec{a}, \\
\vec{y} &= \vec{x} \times \vec{f}, \\
\vec{z} &= \vec{y} \times \vec{x}
\end{align*}
\]
Coregistration Case Study (4)

MEG signals within a Generic Model Brain

2. Convert one of the data sets to the coordinate system of the other

Transformations steps:

- Translate the MEG data set so that the origins of the 2 coordinate systems (the \( a \) points) coincide:

\[
\mathbf{MEG}' = \mathbf{T}(\mathbf{a}_B - \mathbf{a}_M) \cdot \mathbf{MEG}
\]

- Align the +x-axes using 2 rotations, about the y- and z-axes:

\[
\mathbf{MEG}'' = \mathbf{R}_z(\theta_2) \cdot \mathbf{R}_y(\theta_1) \cdot \mathbf{MEG}'
\]

- Align the other 2 axes using another rotation about the x-axis:

\[
\mathbf{MEG}''' = \mathbf{R}_x(\theta_3) \cdot \mathbf{MEG}''
\]

- As the size of the model and patient brains may differ, scale the MEG vectors according to the ratios of the respective measurements:

\[
\mathbf{MEG}'''' = \mathbf{S}\left(\frac{X_{\text{SIZE}}_B}{X_{\text{SIZE}}_M}, \frac{Y_{\text{SIZE}}_B}{Y_{\text{SIZE}}_M}, \frac{Z_{\text{SIZE}}_B}{Z_{\text{SIZE}}_M}\right) \cdot \mathbf{MEG}'''
\]
Coregistration Case Study (5)

MEG signals within a Generic Model Brain

• The composite transformation that coregisters the MEG data onto the generic brain model is:

\[ S \left( \frac{X_{\text{SIZE}}_B}{X_{\text{SIZE}}_M}, \frac{Y_{\text{SIZE}}_B}{Y_{\text{SIZE}}_M}, \frac{Z_{\text{SIZE}}_B}{Z_{\text{SIZE}}_M} \right) \cdot R_y(\theta_1) \cdot R_z(\theta_2) \cdot R_x(\theta_3) \cdot T(a_B - a_M) \]

• The 2 data sets can now be correctly displayed together.
Time-Tested Principles for Good Visual Plots

- If a visualization includes coordinate axes:
  - They should be clearly marked with their quantities & units
  - Legends should never be omitted, even when obvious
  - An overloaded visualization is hard to comprehend
  - Split a visualization into multiple units, if a large number of variables must be presented (avoid overloading).
Time-Tested Principles for Good Visual Plots(2)

- Scale and coordinate-axis origins:
  - The wrong scale relative to the data values can result in large data fluctuations appearing small & vice versa
  - Setting axis origins at a non-zero value can result in an apparent reduction in data values
  - The initial value of the axis should be clearly indicated
Time-Tested Principles for Good Visual Plots

- Comparison of unlike quantities leads to misinformation
  - Last bar is current year so far

- Transition from quantitative to visual information:
  - A well-chosen color-mapping can bring out information that would otherwise be uncapturable
  - Choosing colors with sufficient disparity to display different variables avoids clutter
Tone Mapping

- Need to accentuate important value transitions in the scalar data domain & compress the rest:
  - Scalar data come from various sources → their range varies
  - The scale of the raw data is not necessarily compatible with the sensory response curve of human photoreceptors
  - A direct linear mapping of the input data to light intensity doesn’t have the desired effect
  - The domain is so large that only a poor discretization & scaling of the input range can be perceived
- Convert the raw input data domain scale to a meaningful range of intensity & color values, so that the desired information is pinpointed & extracted intuitively
- Transform the data with the help of
  - Transfer functions
  - Color mapping
Transfer Functions

- The original signal (a) is a thermal sensor capture
  - The sensor temperature-sensitivity range is mapped to a linear 8-bit grayscale gradient
  - The visualization provides only a general idea of the heat distribution
  - The original source is not easily spotted
  - The smooth shade transition makes it impossible to classify the heat into temperature zones
- The original signal was modified by different transfer functions into (b), (c), (d)
Transfer Functions (2)

- A transfer function has the form:
  \[ i_{\text{out}} = f_{\text{transfer}}(i_{\text{in}}) \]
- The function \( f_{\text{transfer}} \) is not necessarily linear or continuous
- In the above example:
  \[
  i_{\text{out}} = f_{\text{quant}}(f_{\text{contrast}}(i_{\text{in}})) ,
  \]
  \[
  f_{\text{contrast}}(x) = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \cdot v_{\text{max}} ,
  \]
  \[
  f_{\text{quant}}(x) = x_{\text{min}} + \frac{x_{\text{max}} - x_{\text{min}}}{N} \cdot \left[ N \cdot \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} + \frac{1}{2} \right]
  \]

where:
- \( x_{\text{min}}, x_{\text{max}} \): the minimum / maximum input signals values
- \( v_{\text{max}} \): the maximum allowed range value
- \( N \): the number of quantization steps
Transfer Functions (3)

- Some commonly used transfer functions:
  - Horizontal axis: input values
  - Vertical axis: output values

![Sigmoid](image1)
![Thresholding](image2)
![Linear Quantized](image3)
![Noise](image4)
Color Maps

• Intensity alone cannot always convey an intuitive idea about data

• Human beings
  ■ attribute certain colors to particular states of mind
  ■ recognize quantities and qualities through color

• Examples:
  ■ Map: land mass is brown for high altitudes – green for low flatlands
    sea is blue
  ■ The indication of critical levels: green (low/safe) – yellow – red (high/critical)

• Colors
  ■ Have a better separation than grayscale values
  ■ Can clearly highlight important value ranges
Color Maps (2)

• Color map:
  ■ Used to go from grayscale to an arbitrary color gradient
  ■ Is a look-up table of colors corresponding to specific sorted intensity values
  ■ An input intensity that matches one of the table records is directly mapped to the associated color
  ■ Other values are interpolated from the closest table entries
Color Maps (3)

- Let
  - $N_C$: number of color entries in the color map
  - $c_i, \ i=0... \ N_C-1$: the color entries in a color map, sorted in ascending order
  - $s_i$: the input values
- Calculate, via interpolation, the output color $c$ for an input intensity $s$:

```java
if (Nc < 2)
    c = colormap[0].col;
i=Nc-2;
while ( colormap[i].val > s && i > 0 )
    i--;
s1 = colormap[i].val;
s2 = colormap[i+1].val;
if ( s1 == s )
    c = colormap[i].col;
else {
    t = (s - s1) / (s2 - s1);
    c = interp(colormap[i].col, colormap[i+1].col, t);
}
```
In visualization, characteristics of the human visual system must be taken into account.

Eye consists of:
- The pupil (entry point for light)
- The lens (focuses the light onto the retina)
- The retina (“projection wall”)

The retina is made up of nerve cells called photoreceptors.

Photoreceptors capture and transmit visual information to the brain.

The center of the retina is the fovea.

There are two types of photoreceptors:
- Rods
- Cones

Matters of Perception
Matters of Perception (2)

- **Rods:**
  - Sensitive to variations in *intensity*
  - Approximately 200,000,000
  - Spread evenly over the retina

- **Cones:**
  - Sensitive to variations in *chromaticity*
  - Approximately 5,000,000
  - Located close to the fovea
  - Have the ability to separately sense 3 different parts of the spectrum
  - Sensitive to either red, green, or blue light
  - Green cones → 64% of the total number of cones
  - Red cones → 32% of the total number of cones
  - Blue cones → 4% of the total number of cones
  - Red and green cones are located closer to the fovea
  - Blue cones form a ring around them
Matters of Perception (3)

- Rods and cones over the retina:
Matters of Perception (4)

- Different color wavelengths $\rightarrow$ lens assumes different focal lengths
- Many people have a deficiency in distinguishing certain colors (color-blindness)
  - Usually between red and green $\rightarrow$ malfunction of red and green photoreceptors
- The above facts of the human visual system have important consequences in visualization:
  - Cones are located close to the fovea $\rightarrow$ better color vision near the center of the viewing direction
  - Rods significantly outnumber the cones $\rightarrow$ variations in intensity are more effective in a visualization than variations in chromaticity
    - Intensity is linked to variations in value
    - Chromaticity is useful for area segmentation
  - Area around fovea has no blue cones $\rightarrow$ blue is unsuitable for text and other detail
    - Blue is excellent for backgrounds
  - There are no red or green cones on the periphery of the retina $\rightarrow$ red and green are unsuitable for backgrounds
Matters of Perception (5)

- Colors with significantly different wavelengths should not be displayed close to each other, since they require different focusing and the eye gets tired.

- Avoid colors that differ only in their red-green ratio to cater for color-blind individuals.

---

Graphics & Visualization: Principles & Algorithms  Chapter 10
Matters of Perception (6)

- The perceived effect of intensity variations is logarithmic:
  - Apparent difference between intensity pairs (0.2, 0.4) & (0.4, 0.8) is the same
- Perception of intensity levels relates to the relative intensity of their neighborhood
  - An object of the same intensity will appear darker in a light background and lighter in a dark background
Matters of Perception (7)

• Perception of visual stimuli is divided into:
  - Conscious processing
  - Preconscious processing

• Preconscious visual processing:
  - Takes place involuntarily
  - Is extremely fast
  - Precedes conscious visual processing

• Mapping values into visuals \(\rightarrow\) take advantage of preconscious processing \(\rightarrow\) done in a number of ways which include:
  - The use of intensity rather than chromaticity as a value discriminator
  - The use of change to attract attention to detail
  - The mapping of large values to nearer (and therefore larger) objects

• Visual perception is not a mere physical but rather a psychophysical phenomenon \(\rightarrow\) one must also consider the emotional response that different colors have on humans
Matters of Perception (8)

- Significance of certain colors in different cultures:
  - Red $\rightarrow$ danger, stop, negative, excitement, hot
  - Dark blue $\rightarrow$ stable, calming, trustworthy, mature
  - Light blue $\rightarrow$ youthful, masculine, cool
  - Green $\rightarrow$ growth, positive, organic, go, comforting
  - White $\rightarrow$ pure, clean, honest
  - Black $\rightarrow$ serious, heavy, death
  - Gray $\rightarrow$ integrity, neutral, cool, mature
  - Brown $\rightarrow$ wholesome, organic, unpretentious
  - Yellow $\rightarrow$ emotional, positive, caution
  - Gold $\rightarrow$ conservative, stable, elegant
  - Orange $\rightarrow$ emotional, positive, organic
  - Purple $\rightarrow$ youthful, contemporary, royal
  - Pink $\rightarrow$ youthful, feminine, warm
  - Pastels $\rightarrow$ youthful, soft, feminine, sensitive
  - Metallic $\rightarrow$ elegant, lasting, wealthy
Visualizing Multidimensional Data

- In graph plotting, each variable is assigned a separate coordinate axis (dimension)
- In complex problems, we need to visualize many variables simultaneously → cannot be easily accommodated in the few dimensions that we can handle
- Most displays are 2D → 2 variables can be plotted simultaneously
- Virtual-reality systems are 3D → extends the visualization capabilities by just one extra simultaneously displayable variable
- A straightforward mathematical method of reducing the problem is to project onto a subset of the dimensions
  - E.g. by assigning a constant value to some of the variables (orthogonal projection)
Visualizing Multidimensional Data (2)

• For example:
  - Given a data set of \( d \) variables \( \{v_1, v_2, v_3, \ldots, v_d\} \), we project \( d \) variables onto the first 2 by assigning constant values to the rest of them, and the data set becomes \( \{v_1, v_2, v_3 = c_3, v_4 = c_4, \ldots, v_d = c_d\} \), thus achieving a 2D data set

• The common technique of \textit{slicing} is an example of a 2D projection:
Visualizing Multidimensional Data (3)

- 1 extra variable can be accommodated by exploiting the time dimension
  - Preferably map onto the time dimension a variable that is itself related to time
  - Animation techniques are very relevant
Color, grayscale or fill patterns can also map the value of 1 extra variable

Glyphs can be used to display more variables in a visualization
A glyph is a visual object onto which variable values may be mapped, each onto a different visual attribute of the glyph
The type of glyph used should be chosen so as to invoke the desired human perception of the data represented
Visualizing Multidimensional Data (5)

- For example:
  - for vector data, the obvious glyph to use is the arrow
- Spheres, disks, crosses, and cylinders are also commonly used glyphs
- Up to 3 positional variables can be mapped onto the position of the glyph
- A number of extra variables can be mapped onto other glyph attributes
  - This number can be no more than 2, otherwise the glyphs get overloaded
- A common way is to map one variable onto the scale of the glyph and the other onto its color / intensity
Mihalisin proposed a hierarchical method for visualizing functions of $N$ independent variables $f(x_1, x_2, \ldots, x_N)$ as 2D graphs.

- Independent variables take values from finite, discrete, contiguous range.
- Vertical axis displays the function value.
- The independent variables are assigned a unique priority and are hierarchically mapped onto the horizontal axis:
  - Variable with highest priority ($x_1$) varies the fastest.
  - The variable with the lowest priority ($x_N$) varies the slowest.
  - For each value of $x_N$, which maps onto a line segment on the horizontal axis, all other variables cycle through their values like a nested for-loop, with $x_1$ cycling the fastest.

- Value of $f$ is plotted for each set of values that independent variables take.
- The function values for each cycle of the variables can be hierarchically nested in bounding boxes → better visualization.
Visualizing Multidimensional Data (7)

- EXAMPLE (Mihalisin):

\[ f(x_1, x_2, x_3) \]

\[ x_1 \]
values

\[ x_2 \]
values

\[ x_3 \]
values