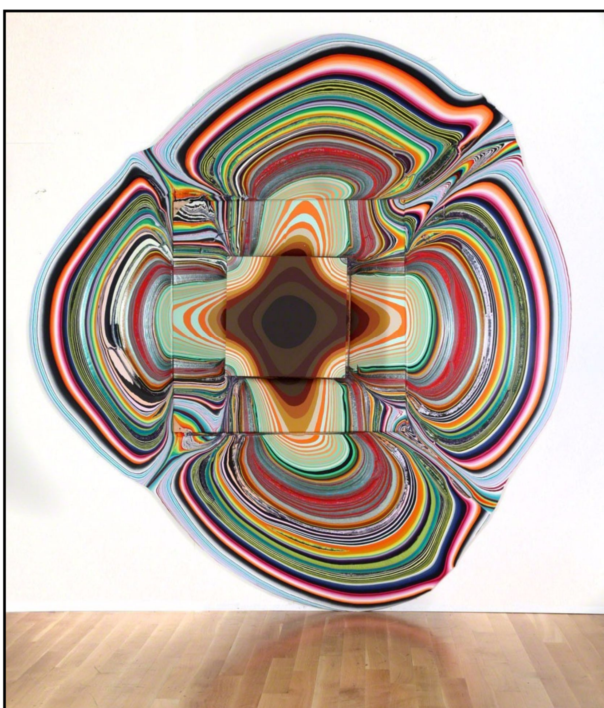
CSCI 4550/6550 Interactive Visualization

https://www.cs.rpi.edu/~cutler/classes/visualization/S24/

Lecture 9: Streamgraphs, Gestalt Psychology & Line Arrangements

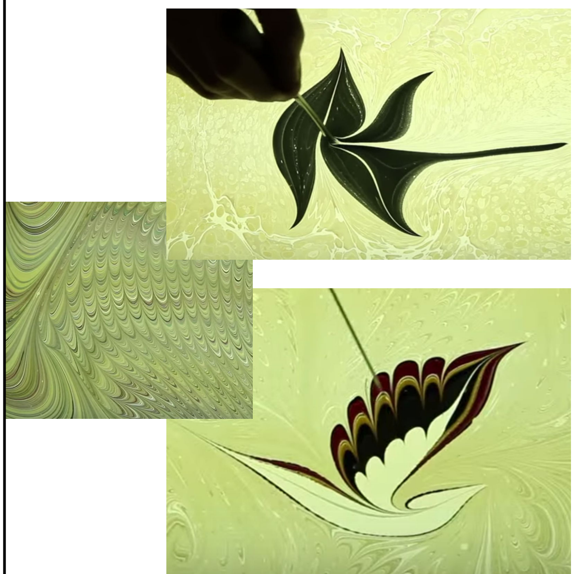


Holton Rower
“Sometimes I Have to Look in the Mirror to See if I'm Still There”
2011

https://www.youtube.com/watchz?v=Gyktr2OI4v4

http://seaburyschoolnavigators.blogspot.com

Ebru Art @ American Islamic College



https://www.youtube.com/watch?v=lILIFcq3K3U

What is the science behind these art forms? What physics/fluid would we need to accurately model to build a simulation?

What is the science behind these art forms? What physics/fluid would we need to accurately model to build a simulation?

- water/oil interaction (chemistry too)
- pen/stick/needle/stylus actions (water depth, speed through & speed in/out, angle, thickness, material, momentum)
- Momentum of water/color
- Surface tension
- Paper step looks hard
- Layered transparency blending? (not sure)
- Use this technique, paint on flat water, apply pattern to curved surface
- Fluid rigid body simulation
  - Surface tension
  - Viscosity
- Paint mixing vs displacement
- How to dye the oil? (not water based)

Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - "Stacked Graphs – Geometry & Aesthetics"
  - "Baby Names, Visualization, and Social Data Analysis"
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

Homework Assignment 5: Experimenting with Color

- Revisit an earlier assignment/data/toolkit
  - Make a non-color-related improvement to this visualization
- Prepare many versions of the same visualization experimenting with different color palettes, e.g.:

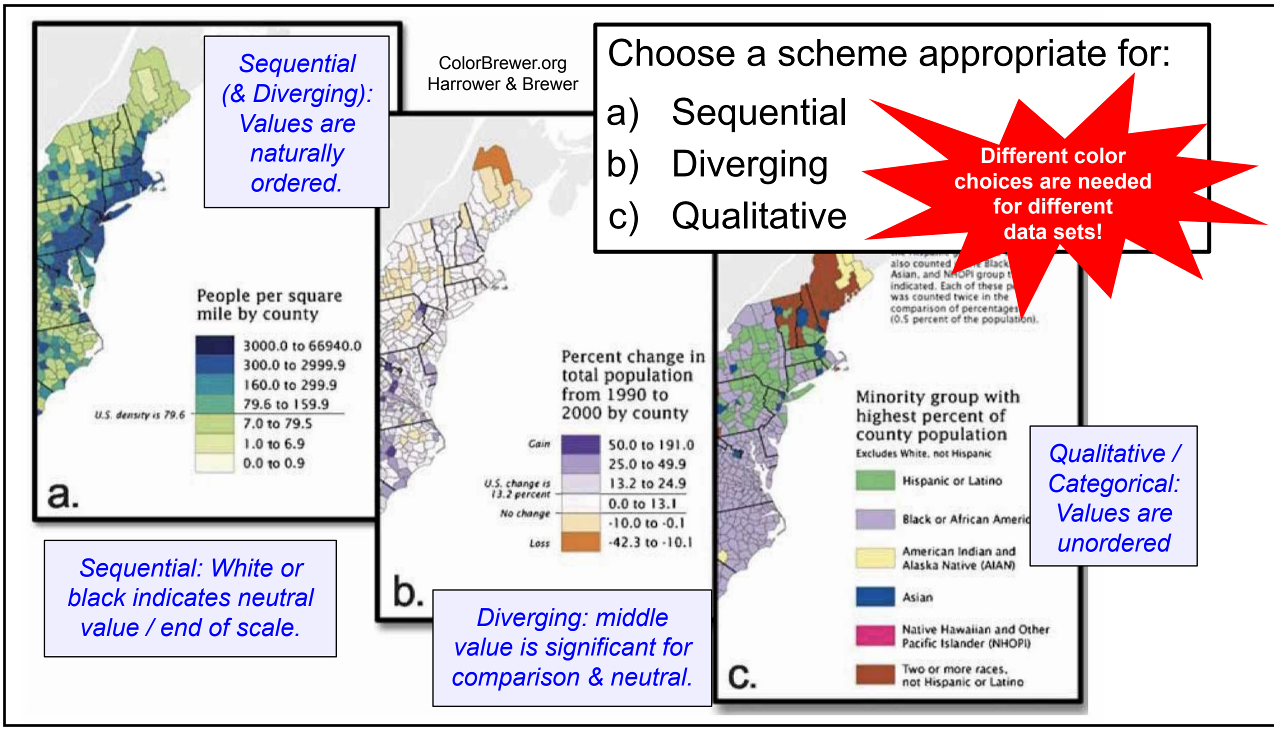
Teams of 2 encouraged!

  – Shades of grey
  – Black & white
  – Cool vs. warm tones
  – Bold/saturated vs. pastel
  – Colorblind aware
  – Light vs dark background and/or color negation
  – Etc.

- Analyze the effectiveness of the color scheme for each visualization.
- How well does it convey the message? Or mislead the viewer?
- Compare the visualizations to each other.

Matlab Color Schemes

- “when using many colors of different hues, I get a blended hot mess of color that completely washes out and detail of the underlying 3D rendered model”
- The Matlab “jet” color scheme is the default (and it looks so pretty!), but it is misleading!
- *Color Brewer is here to help!*



"Modeling Color Difference for Visualization Design"
Szafir, IEEE TVCG / IEEE VIS 2017

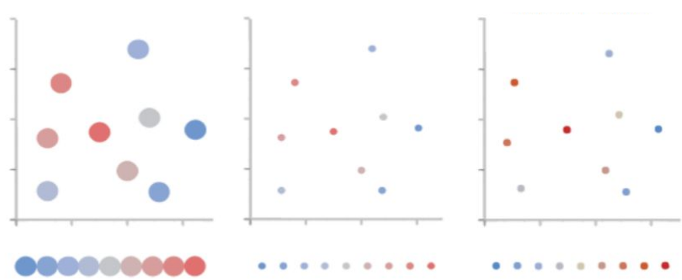


Table 1. Regression results for points, where $p = m_x * \Delta X$.

Axis	Size (s)	Size in Px	Slope	R^2	$ND(50\%)$ in ΔE
L	0.25°	6 px	0.059	0.948	8.37
L	0.5°	12 px	0.074	0.97	6.74
L	0.75°	18px	0.087	0.981	5.75
L	1°	25px	0.087	0.965	5.75
L	1.5°	37 px	0.082	0.996	6.08
L	2°	50 px	0.091	0.974	5.47
a	0.25°	6 px	0.031	0.984	16.11
a	0.5°	12 px	0.05	0.988	9.98
a	0.75°	18px	0.059	0.987	8.52
a	1°	25px	0.064	0.992	7.81
a	1.5°	37 px	0.073	0.985	6.87
a	2°	50 px	0.073	0.994	6.84
b	0.25°	6 px	0.026	0.978	19.46
b	0.5°	12 px	0.037	0.988	13.34
b	0.75°	18px	0.044	0.994	11.35
b	1°	25px	0.05	0.979	10.03
b	1.5°	37 px	0.056	0.979	8.97
b	2°	50 px	0.063	0.99	7.99

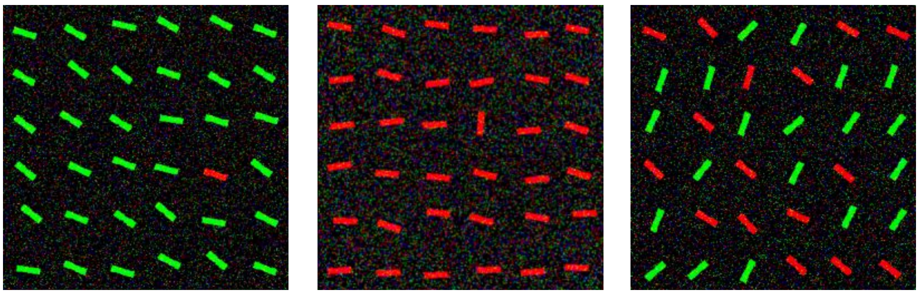
Table 2. Regression results for lines, where $p = m_x * \Delta X$.

Axis	Size (s)	Size in Pixels	Slope (m)	R^2	$ND(50\%)$ in ΔE
L	0.05°	2px	0.033	0.876	15.35
L	0.1°	3px	0.042	0.92	11.98
L	0.15°	4px	0.058	0.921	8.69
L	0.25°	6px	0.065	0.955	7.74
L	0.3°	7px	0.069	0.947	7.23
L	0.35°	9px	0.072	0.96	6.92
a	0.05°	2px	0.036	0.978	13.92
a	0.1°	3px	0.043	0.956	11.57
a	0.15°	4px	0.049	0.959	10.28
a	0.25°	6px	0.053	0.94	9.39
a	0.3°	7px	0.061	0.933	8.15
a	0.35°	9px	0.064	0.919	7.79
b	0.05°	2px	0.026	0.981	19.47
b	0.1°	3px	0.031	0.967	16.15
b	0.15°	4px	0.033	0.934	15.17
b	0.25°	6px	0.036	0.918	13.75
b	0.3°	7px	0.04	0.927	12.43
b	0.35°	9px	0.045	0.945	11.05

Visual Salience

Laurent Itti (2007), Scholarpedia, 2(9):3327.
http://www.scholarpedia.org/article/Visual_salience

- … is the perceptual quality that makes some items in the world stand out from their neighbors and grab our attention



- Designers use saliency to create objects (such as this emergency triangle) that appear highly salient in a wide range of viewing conditions

- Berlin & Kay's 11 basic color terms:
  black, white, red, green, yellow, blue, brown, purple, pink, orange, gray

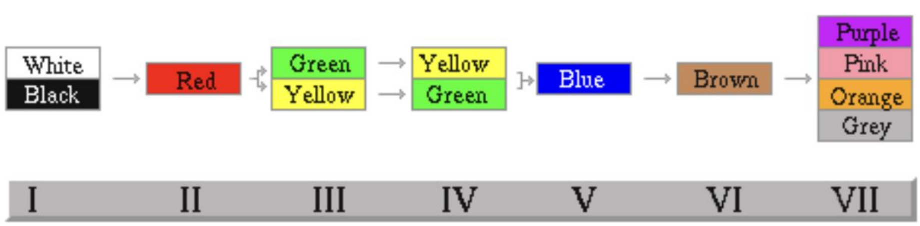


Figure 3. Berlin and Kay's hypothesis about seven evolutionary stages of colour terms (1969:4).

"The Colour and The Word", Virve Sarapik, Folklore, 1997.

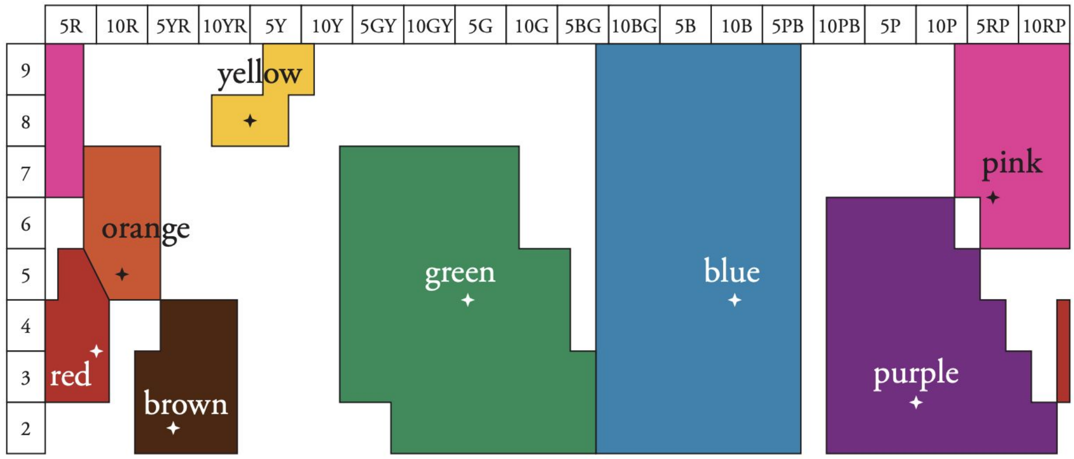


Fig. 6. Berlin and Kay's diagram of the eight basic colours in English (redrawn from [BK69, Appendix I, p.119]). As in Figure 5, the colour space is that of the Munsell colour system. See Figure 4 for an explanation of the notation. Berlin and Kay used a 320 chip Munsell array. They asked participants to determine, for each basic colour term, ***x***, (1) all those colour chips which they would, under any conditions, call ***x***, and (2) the best, most-typical examples of ***x***. The small crosses mark the locations of the "best, most-typical example" for each colour. The colour of each area matches that best most-typical example, within the limits of the available gamut. The white areas represent colour chips that were not given an unequivocal colour name.

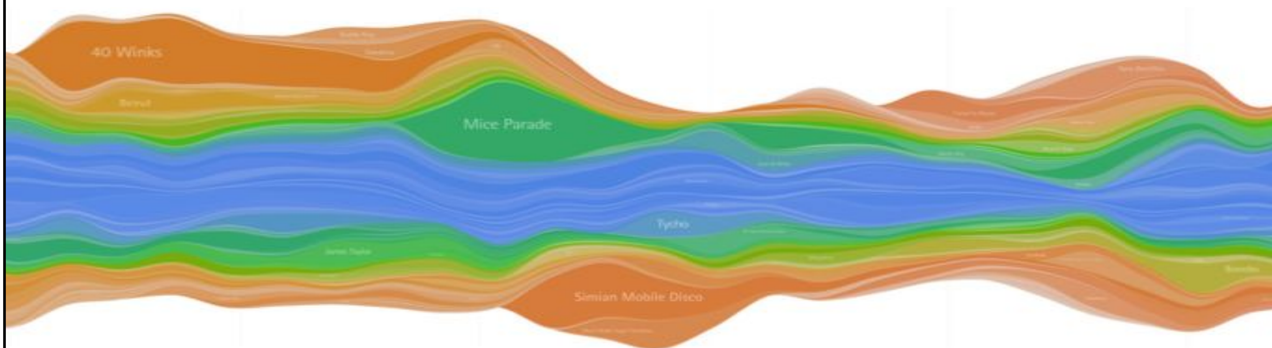
"What is the "opposite" of "blue"? The language of colour wheels."
Dodgeson, Perceptual Imaging 2019.

Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - "Stacked Graphs – Geometry & Aesthetics"
  - "Baby Names, Visualization, and Social Data Analysis"
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

Reading for Tuesday *pick one*

“Stacked Graphs – Geometry & Aesthetics”
Lee Byron & Martin Wattenberg, IEEE TVCG 2008



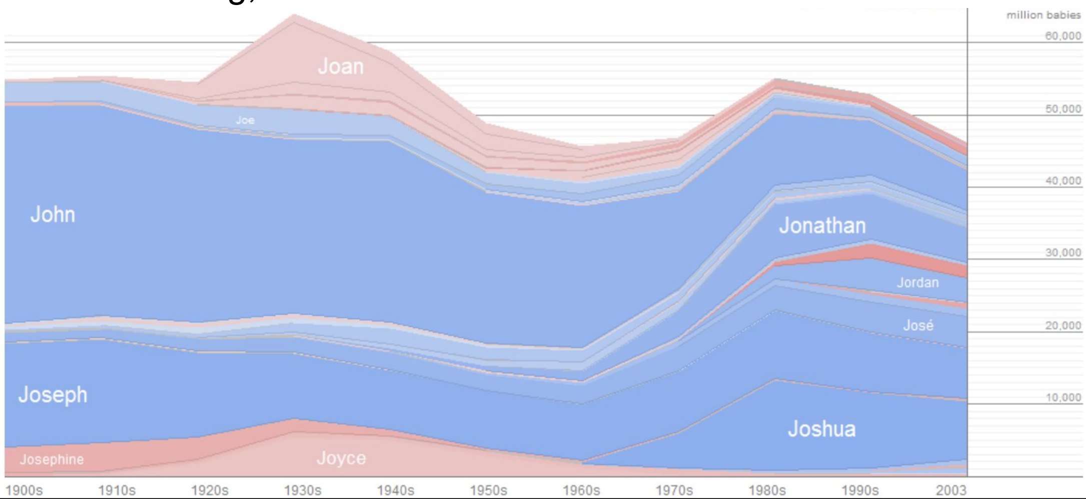
- Internet feedback was a large part of paper
  - Unnecessary?
  - started off as a story or documentary rather than an academic paper, strange
  - vulgar language :(, seemed unprofessional (?)
  - Cannot measure “organic and emotionally pleasing”, proper analysis of comments beyond ‘some people liked it / some did not’ was not attempted/difficult/impossible
- Like other art… it is controversial, and that’s ok!
  - Eye magnets
- Layout
  - Tufte’s macro/micro: show both the sum and the individual values
  - Minimize “wiggle” of extremes & middle curves, thicker layers weighted more highly (but does not display data as accurately?)
  - Keep graph centered (don’t drift up or down)
  - Border/space between layers? If required by media
  - Labels are important
  - Interaction is good
  - Looks like mountains, foreground/background, perceived overlap implies some data is bigger than just the visible area

- Colors
  - Natural & pleasing, not too loud or distracting, (boring?)
  - For a particular dataset is it necessary to match disjoint regions by color?
  - Choose color from relevant 2D axes mapped to color & saturation
  - Don't all need to be unique, aren't limited by # of distinguishable colors
- Local contrast
  - Display data, e.g., time of onset, popularity,
  - Bias color selection to keep image balanced
- Order
  - Lack of prescribed ordering is unsettling
  - Sort by time of offset *or* "measure of volatility"
  - Generally alternating top & bottom seems to be best
  - Entrance position discussion interesting
  - This is not a black box one-size-fits-all visualization technique. It requires thought to choose colors & tweak layout. It won't work for all data. Could be done badly…
  - Difficult for newbies - no clear, straightforward solution/guidance for layout or color
  - Waviness of output appropriate for music data
  - Handles massive amount of data
  - This should be a graph option in Excel!

- Legend mandatory to understand and interpret a particular streamgraph (no conventions)
  - The plots are so interesting you are motivated to figure it out
- Abstract was awkward talking about an image that wasn't right there
- Insufficient comparison images within the paper
  (copyright thing? page limit? assumed to be "common knowledge"?)
- Didn't explain improvements streamgraph vs. theme river
- Personalized visualizations are cool
- Purpose of this visual: to look cool & be engaging & draw interest or to scientifically measure & conclude things?
- Some sloppy figure callouts/captions/labels. Some captions too short. Why was this image included? Need to do more than just describe the source of the data
- Vertical vs horizontal: horizontal most appropriate for time data.
- Helpful reference:
  http://www.visualisingdata.com/index.php/2010/08/making-sense-of-streamgraphs/

Reading for Tuesday *pick one*

"Baby Names, Visualization, and Social Data Analysis"
Martin Wattenberg, IEEE InfoVis 2005

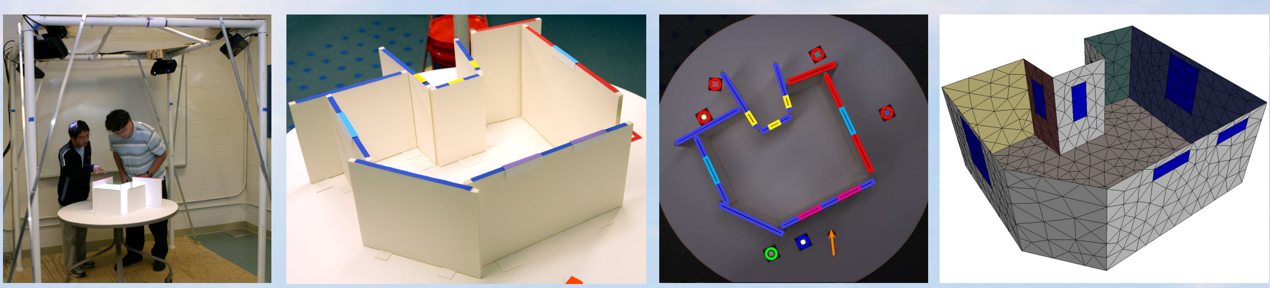


- Value in a tool that encourages exploration
- Users spending considerable time
- Data analysis is fun, becomes a game, exploration not just for an individual, also fun as a group or social activity, family connections
- Tool is a minor fad, broadly popular, no formal advertising investment
- Not scientific, anecdotal review of lots of blog posts
- Data from ~100 years of U.S. SSN data
- Time series data for approximately 6,000 distinct names
- Darker shades of blue/pink for names that are currently more popular
- "Overview first, zoom and filter, details on demand" [Shneiderman 1996]
- Visualization updates on each keystroke (don't need to press enter)
- Compares to other contemporary interactive visualization designs
- Total data drawn regulated to avoid performance problems
- Bartles taxonomy: achievers, socializers, explorers, killers
- What makes, or how to create a successful visualization / interaction app?
  Common but unique perspective, expressive spectator interface, discovery transfer

Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - “Stacked Graphs – Geometry & Aesthetics”
  - “Baby Names, Visualization, and Social Data Analysis”
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

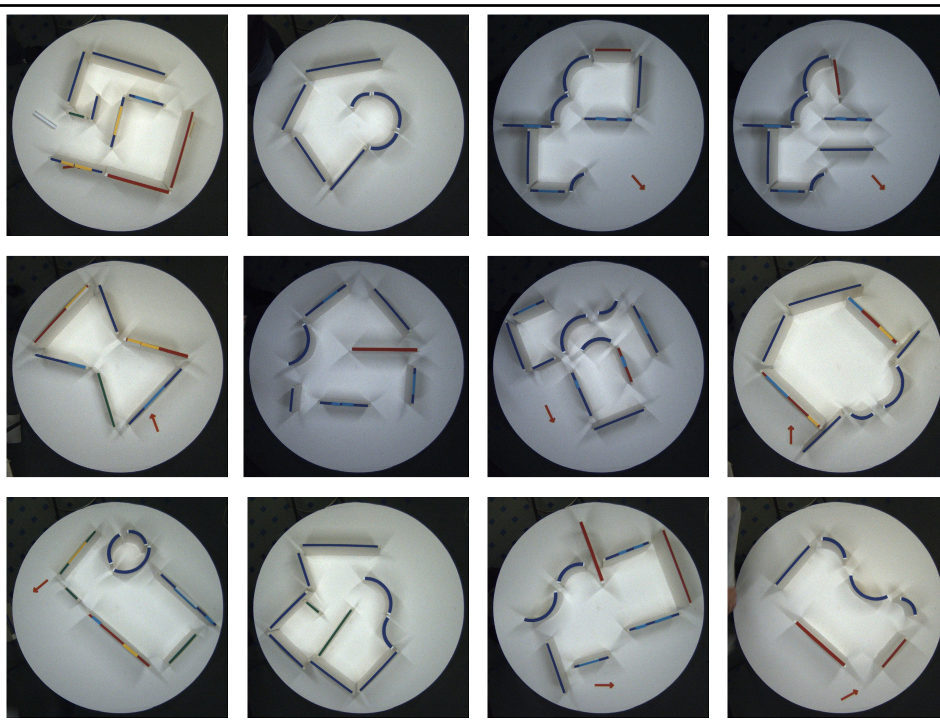
Interpreting Physical Sketches as Architectural Models



Barbara Cutler and Joshua Nasman

Advances in Architectural Geometry 2010

Spatially Augmented Reality (SAR) Projection



Our Contributions

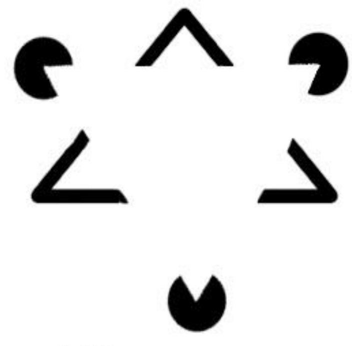
- Algorithm for automatic interpretation of interior space vs. exterior space
- Construction of a watertight 3D mesh
- User study collected >300 example designs
- Validation of algorithm
  - Compare to annotations by the original designer
- Quantify design ambiguity
  - Compare annotations of a design by other users

Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - “Stacked Graphs – Geometry & Aesthetics”
  - “Baby Names, Visualization, and Social Data Analysis”
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

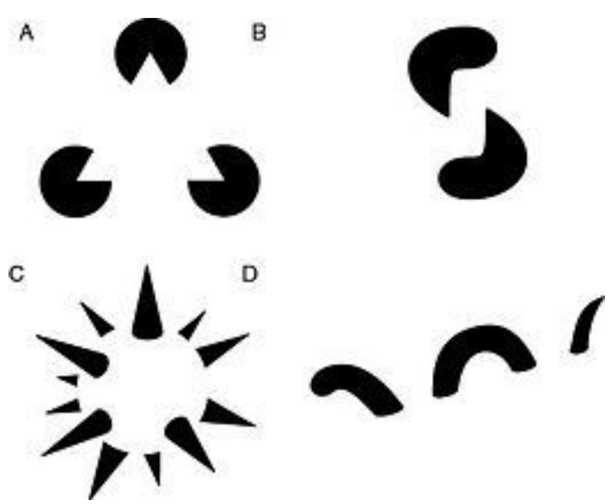
Related Work

- Tangible User Interfaces [e.g., Ben-Joseph et al. 2001]
- Sketch-based Modeling User Interfaces [Zeleznik et al. 1996; Igarashi et al. 1999; Dorsey et al. 2007]
- Pen-based Sketch Recognition [e.g., Wacom 2010]
- Automated Recognition of CAD Drawings [Aoki et al. 1996; Llados et al. 1997; Ah-Soon and Tombre 2001; Kulikov 2004; Lu et al. 2005]
- Human Vision, Perception, & Gestalt Psychology
  - Seek the simplest explanation for an incomplete diagram
  - Closure, proximity, symmetry, continuity, collinearity, & parallelism



Gestalt Principles: Reification

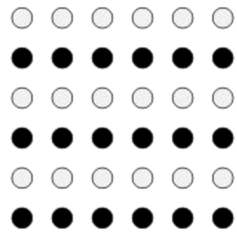
- Constructive or generative aspect of perception



http://en.wikipedia.org/wiki/File:Reification.jpg

Pragnanz: Gestalt Laws of Grouping

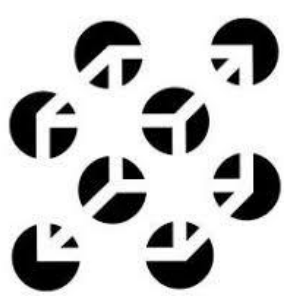
- Proximity
- Similarity
- Closure
- Symmetry
- Common fate
- Continuity
- “Good” gestalt
  (regular, simple & orderly,
  eliminate complexity
  & unfamiliarity)
- Past experience

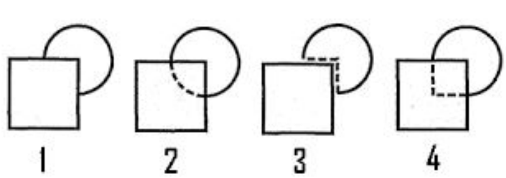


http://en.wikipedia.org/wiki/Gestalt_psychology

Gestalt Principles

- Figure/Ground





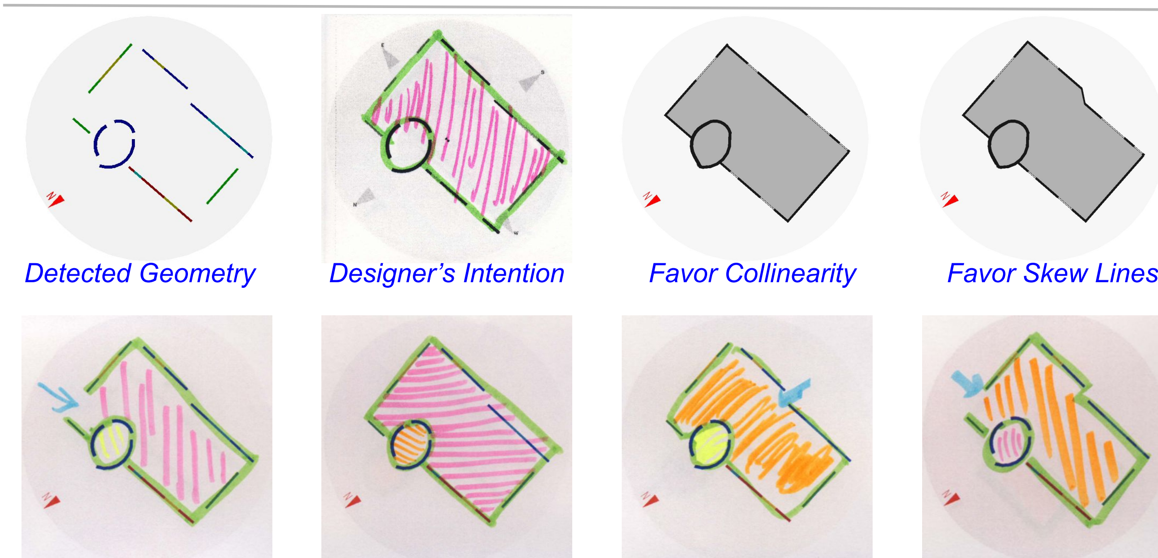
Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - “Stacked Graphs – Geometry & Aesthetics”
  - “Baby Names, Visualization, and Social Data Analysis”
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

Overview of Algorithm

- Image Processing (our earlier publications)
- Lack of precision in sketch:
  - Should elements be parallel?
  - collinear?
  - perpendicular?
  - tangent?
- Link tangent walls to form continuous chains that divide space into zones
- Determine interior vs. exterior
- Generate floor plan diagram & watertight mesh geometry

Tolerance Example: Collinearity

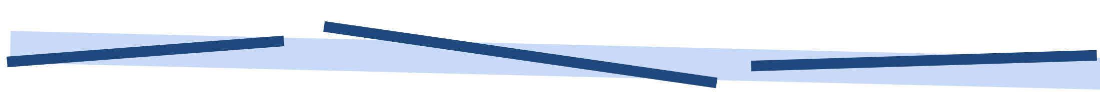


Detected Geometry *Designer's Intention* *Favor Collinearity* *Favor Skew Lines*

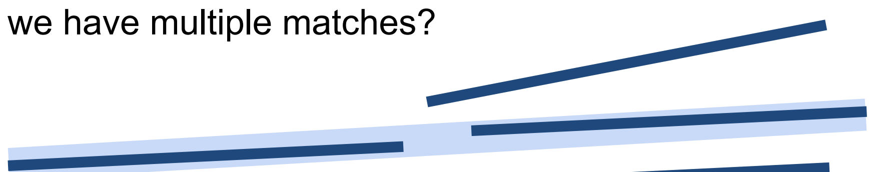
Other Users' Interpretations

Connecting Segments Together

- End points near each other (what epsilon/tolerance?)
- Approximately parallel (what epsilon/tolerance?)

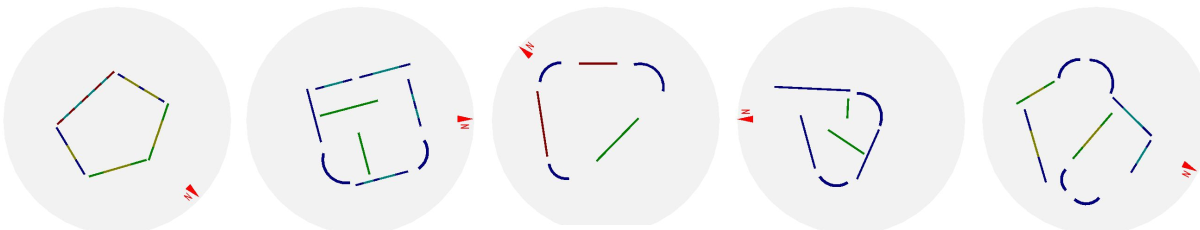


- Snap to perfect line? Or preserve original shape?
- What if we have multiple matches?

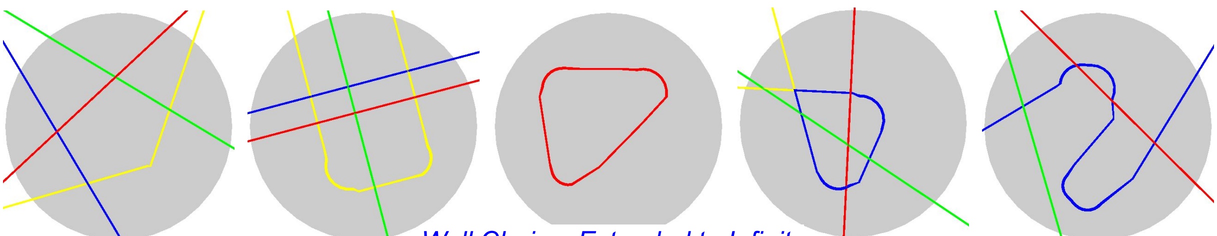


Linking Elements to Form Chains

- Nearby walls with similar tangents can be joined into a chain



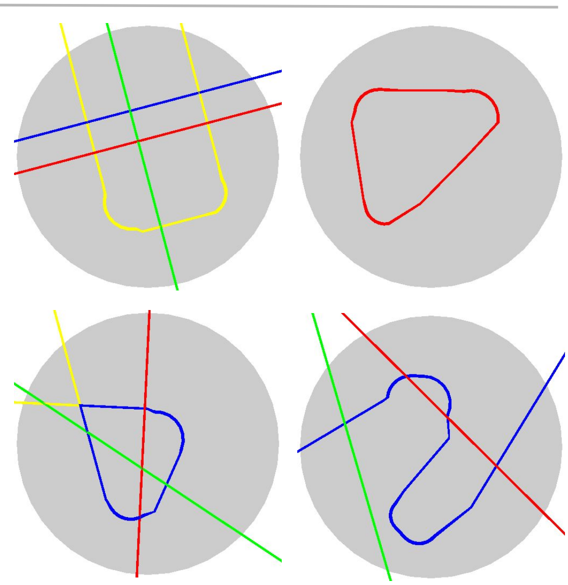
Detected Geometry



Wall Chains, Extended to Infinity

A Modified “Line Arrangement”

- In addition to infinite straight lines, a “wall chain” may:
  - Bend or be curved!
  - Be a closed loop!
  - Cross itself!
  - Cross another wall chain more than once!

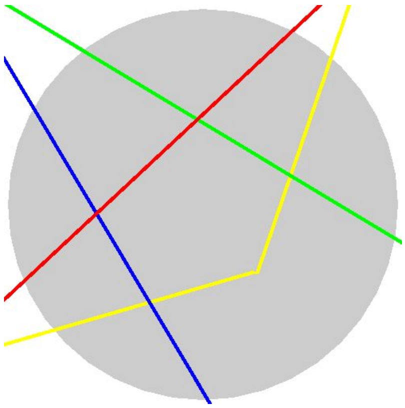


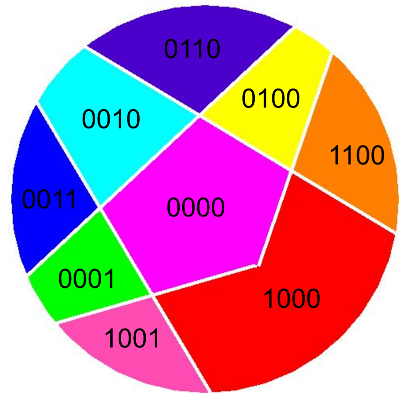
Complexity Analysis?

- Input: *n* physical walls or sketched pen strokes
- Consider connecting into chains
  - How many endpoint-endpoint connections?
  - How to reduce # of comparisons?
- After joining into *m* “wall chains” ($m \le n$)
  - How many point intersections (between wall chains)?
    What if the wall chains are perfect lines?
    What if they wall chains are general curves?
- How many zones/cells/faces (assume lines)?
- How to uniquely label faces?
- Labeling Interior vs. Exterior
- How many ways to label entire diagram interior vs. exterior?
  Assume *f* faces, and each face should be labeled interior or exterior.

Complexity Analysis?

- Input: *n* physical walls or sketched pen strokes
- Consider connecting into chains
  - How many endpoint-endpoint connections? $(2n * 2(n-1)) / 2 = O(n^2)$
  - How to reduce # of comparisons? *Spatial data structure, like a quadtree*
- After joining into *m* “wall chains” ($m \le n$)
  - How many point intersections (between wall chains)?
    What if the wall chains are perfect lines? $O(m^2)$
    What if they wall chains are general curves? $O(\infty)$
- How many zones/cells/faces (assume lines)? $O(2^m)$
- How to uniquely label faces? *Binary code, each bit represents which “side” of each wall chain it lies on.*
- Labeling Interior vs. Exterior
- How many ways to label entire diagram interior vs. exterior? $O(2^f) = O(2^{2^m})$
  Assume *f* faces, and each face should be labeled interior or exterior.





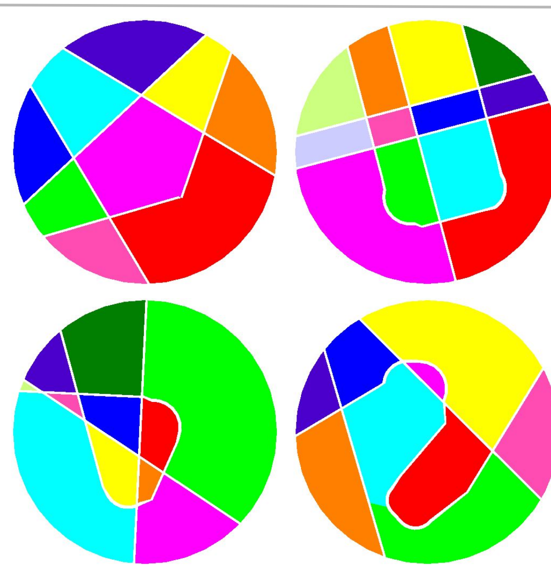
4 lines (wall chains)

with this configuration,
limited to this circle,
we have 9 faces/zones

Halfspace Zones & Enclosure

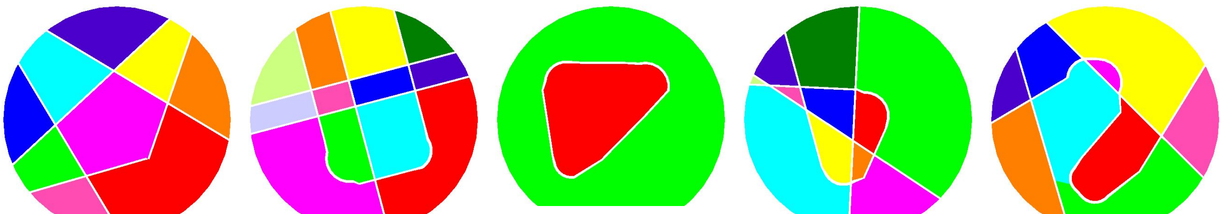
- For n wall chains
- *For simplicity, assume they are infinite straight lines*
- We will have $O(n^2)$ faces/cells in the arrangement
- Each face/cell can be "interior" or "exterior"

- → $2^{O(n^2)}$ *possible buildings*
  Not feasible to check all of them!!

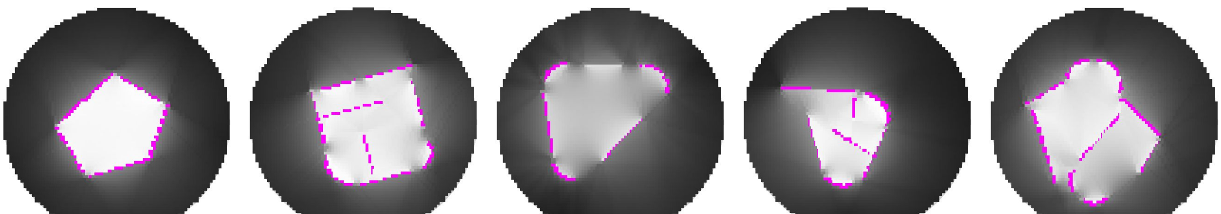


Halfspace Zones & Enclosure

- Further subdivided using GraphCuts (if needed)



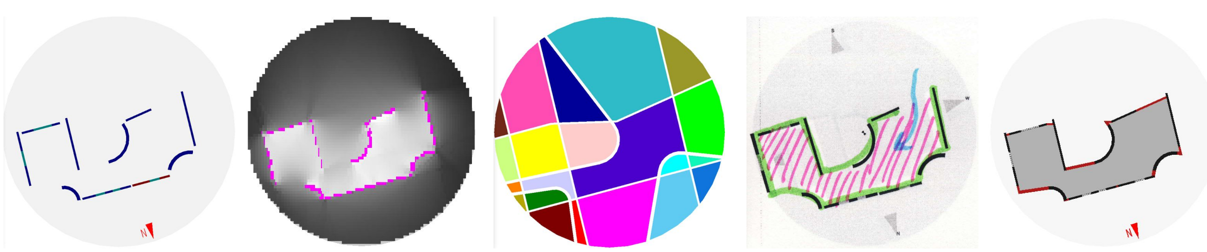
Halfspace Zones



Densely Sampled Enclosure (Visibility Test)

Interior/Exterior Optimization

- Analyze histogram of point-sampled enclosure values
- Maximize usage of lengths of real wall elements
- Minimize length of inferred (added) walls
- Minimize area assigned in opposition of simple threshold metric

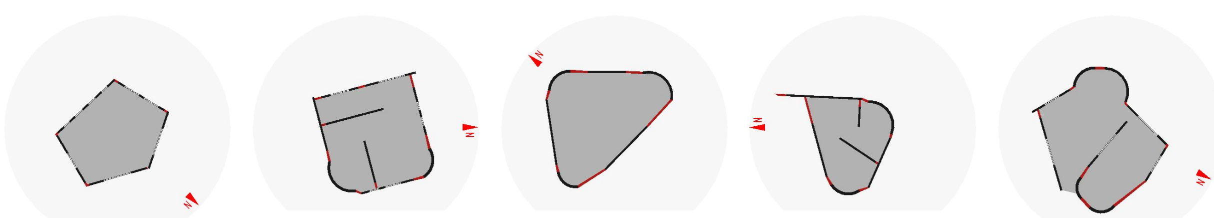


Interior/Exterior Optimization

- Analyze histogram of point-sampled enclosure values
- Maximize usage of lengths of real wall elements
- Minimize length of inferred (added) walls
- Minimize area assigned in opposition of simple threshold metric

Interior/Exterior Enclosure Threshold

There is no universal threshold – varies design-to-design, and *within-a-design*

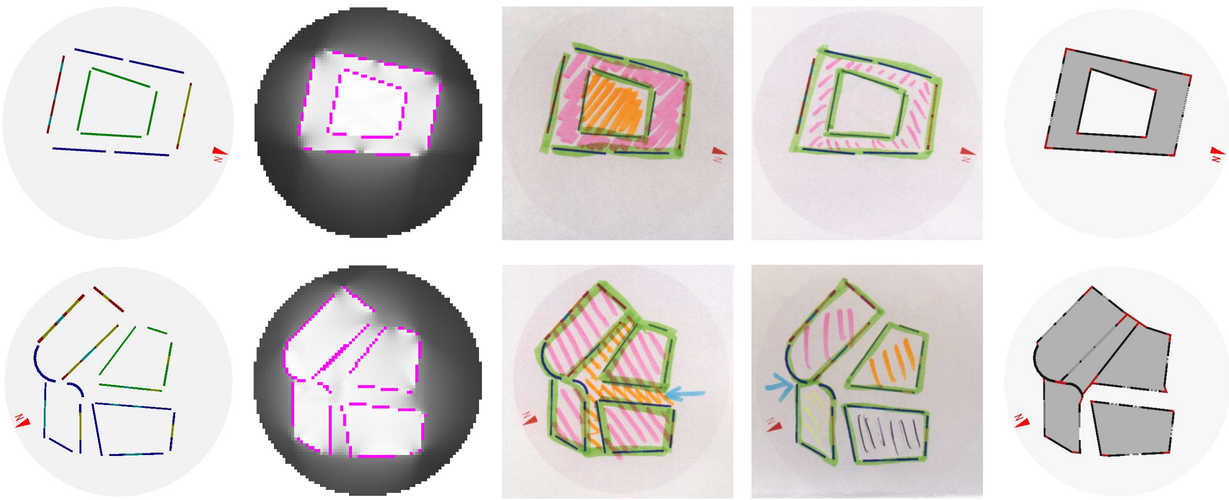


Automatic Interior/Exterior Determination & Final Floorplan

Compare to Designer's Intention

Interior/Exterior Optimization

- (Courtyard option) Minimize total enclosed area



Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - “Stacked Graphs – Geometry & Aesthetics”
  - “Baby Names, Visualization, and Social Data Analysis”
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

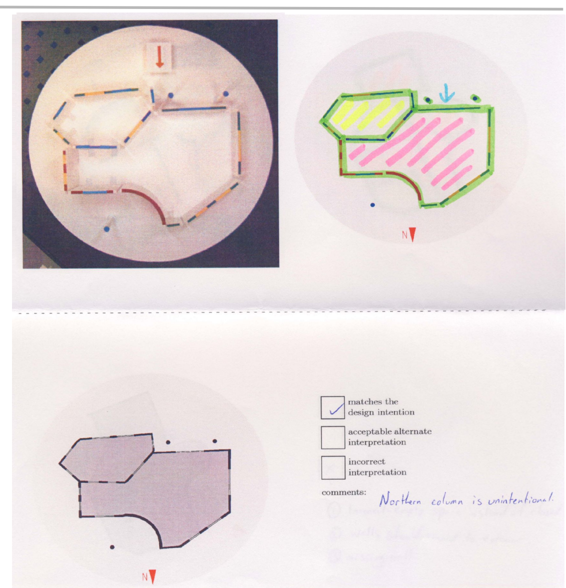
Our Goals in Conducting User Study Design

- Understand range of designs possible
- Improve physical sketching user interface
- Improve algorithm for sketch recognition of interior/exterior space
  - Learn common human interpretation “rules”
  - Quantify design ambiguity
- Measure effectiveness of *Virtual Heliodon* as an architectural daylighting design tool

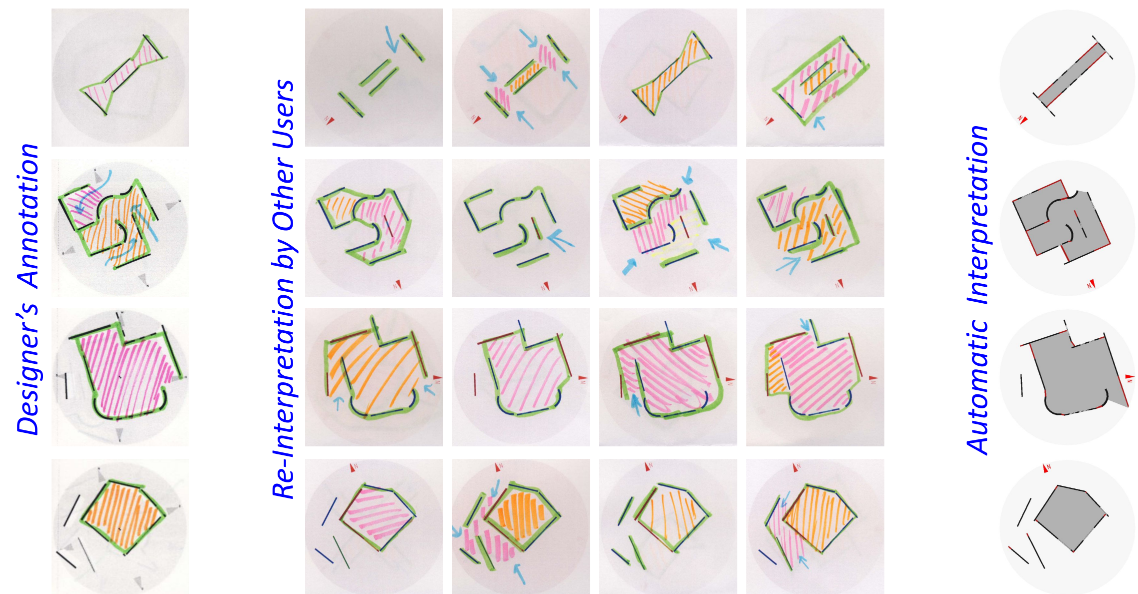
User Study 1: Open-Ended Design

- 30 participants (15 architects)
- 20 mins of sketching
- 329 unique designs (154 by architects)

- After design session:
- Designer annotates each design
- Then, we reveal our automatic interpretation

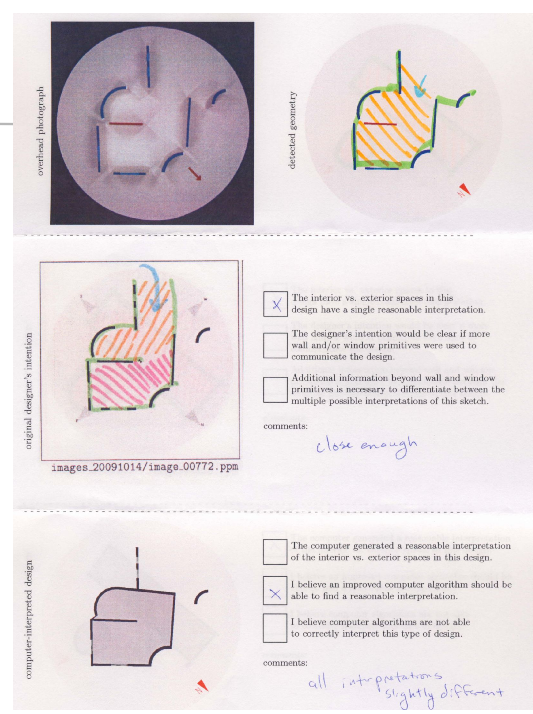


User Study: Identify/Quantify Ambiguous Designs



User Study 2: Re-Interpretation

- 114 designs from Study 1
  - All ambiguous designs included
  - Some clear designs (as controls)
- 15 participants
- Re-interpreted by another user
  - 3-6 new annotations for each
  - 346 total (124 by architects)
- Then compare to original designer's annotation
- And finally, to our automatic interpretation



Re-Interpretation Results

- No correlation found between background (architecture/arts/none) & interpretation accuracy
- We will continue to improve the robustness of our software

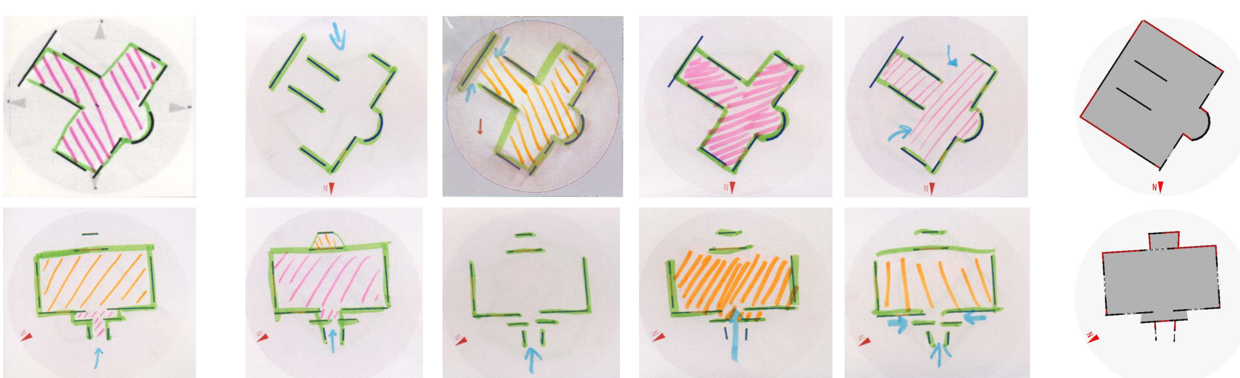
matches the original designer's intention

	correct		mostly correct		incorrect		total
clear	155	78%	17	9%	26	13%	198
ambiguous	74	56%	35	27%	22	17%	131
total	229	70%	52	15%	48	15%	329

multiple interpretations possible

Domain-Specific Knowledge Required

- Standard vocabulary of architectural forms (e.g., cruciform, portico, etc.)
- Maybe architectural data and modern AI/ML would help in these cases?



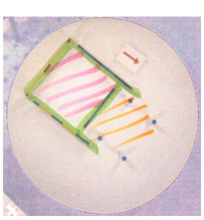
Designer's Annotation *Re-Interpretation by Other Participants* *Automatic Interpretation*

Future Work

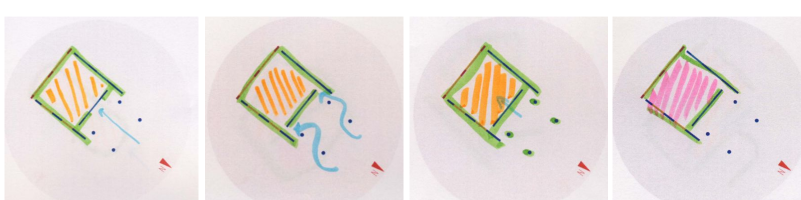
- Improve/robustify interpretation algorithm
  - Detect symmetry & repetition
  - Multi-zone interiors & circulation paths
- Incorporate domain-specific knowledge
- Enhance user interface
  - Additional tokens, more complex element shapes
  - Alternative to sketching in plan:
    sketch (double height, multi-floor) vertical sections
- Apply to pen-based sketch interpretation

Thanks!

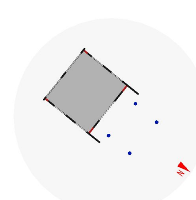
- Yu Sheng, Ted Yapo, & Andrew Dolce
- Our user study volunteer participants
- Funding from NSF & IBM



Designer's Annotation



Re-Interpretation by Other Participants



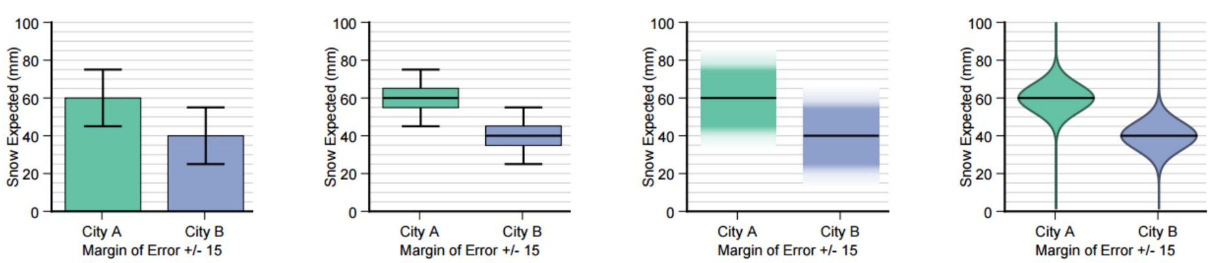
Automatic Interpretation

Today

- Homework 5: Experimenting with Color
- Continuation on Color from Lecture 8 and Friday's Worksheet
- Readings for Today
  - “Stacked Graphs – Geometry & Aesthetics”
  - “Baby Names, Visualization, and Social Data Analysis”
- Research Example: Architectural Sketching
- Related Work: Gestalt Psychology
- Computational Geometry: Line Arrangements & Big O Notation
- User Study Design
- Readings for Friday

Reading for Friday *pick one*

"Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error", Correll & Gleicher, TVCG 2014



(a) **Bar chart** with error bars: the height of the bars encodes the sample mean, and the whiskers encode a 95% t-confidence interval.

(b) **Modified box plot**: The whiskers are the 95% t-confidence interval, the box is a 50% t-confidence interval.

(c) **Gradient plot**: the transparency of the colored region corresponds to the cumulative density function of a t-distribution.

(d) **Violin plot**: the width of the colored region corresponds to the probability density function of a t-distribution.

Reading for Friday *pick one*

"Visual Encodings of Temporal Uncertainty: A Comparative User Study", Gschwandtner, Bogl, Federico, & Miksch, TVCG 2016

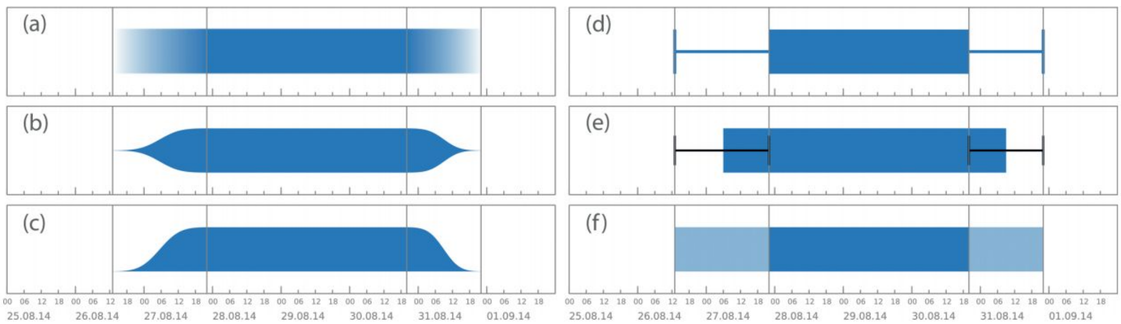


Fig. 1: Six different visual encodings of start/end uncertainty of temporal intervals used in the user study: (a) gradient plot, (b) violin plot, (c) accumulated probability plot, (d) error bars, (e) centered error bars, and (f) ambiguation. We designed encodings (a)–(c) to encode statistical uncertainty and encodings (d)–(f) to encode bounded uncertainty. All encodings were used to estimate earliest start, latest start, earliest end, and latest end, as well as minimum, maximum, and average interval duration. Moreover, encodings (a)–(c) were used to estimate the probability that the interval has already started/ended at a marked position in time.