

Lecture 5. Marks and Channels

(Munzner Ch.5)

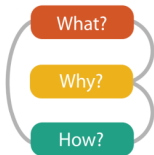
Sim, Min Kyu, Ph.D.
mksim@seoultech.ac.kr



서울과학기술대학교 데이터사이언스학과

Marks and Channels

Moving on to “How”



- 3단계의 vis 사례분석 framework
- What-why-how는 data-task-idiom과 대응된다.
 1. (Data) What data the user sees.
 2. (Task) Why the user intends to use a vis tool.
 3. (Vis Idiom) How the **visual encodings** are constructed.
 - Vis Idiom: 관용적으로 사용되는 vis의 표현들
- Visual encoding은 **Marks**와 **Channels**로 구성된다.

What is Marks?

- 이미지에 등장하는 기초적 그래픽 원소
- Geometric primitive objects classified by spatial dimension

➞ Points



➞ Lines



➞ Areas



Marks Types

- In table datasets, a mark always represents an item.
 1. Marks as Items/Nodes
 2. Marks as Links

1. Marks as Items/Nodes

Marks as Items/Nodes

➡ Points



➡ Lines



➡ Areas



- 1D, 2D, 3D에 해당하는 basic geometric primitive
- 3D mark인 volume은 거의 사용되지 않는다.

2. Marks as Links

Marks as Links

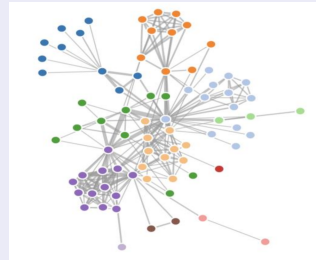
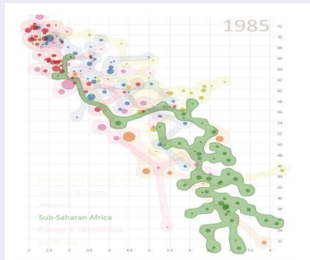
➡ Containment



➡ Connection



- A link may be represented by
 - connection mark for pairwise relationship between items using line
 - containment mark for hierarchical relationships using areas.



What is Channels?

- Marks의 외형을 결정하는 수단
- Marks의 dimensionality와 독립적이다.

➡ Position



➡ Color



➡ Shape



➡ Tilt



➡ Size



➔ Volume



Channel types

- 인간의 인지 시스템은 두 개의 지각 양식을 가지고 있다.

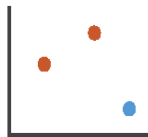
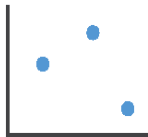
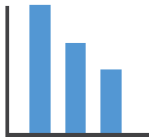
1. Identity channels

- Tell us **what something is** or **Where it is**
- Works for **categorical**
- What shape we see: a circle, a triangle, a cross
- It does not make sense to ask magnitude question for shape

2. Magnitude channels

- Tell us **How much of something there is**
- Works for **ordered**
- Regarding magnitude with line length, how much is this line than that line?
- In this case, identity is not a productive question.
- Luminance question: 얼마나 밝은가라고 물어야지 무슨 색인지 물을필요 없다.
- Angle, area, volume, saturation, tilt

Using marks and channels



- Marks & channels
 - Marks represent items or links
 - Channels change appearance of marks based on attributes



(a) Barchart는 2개의 attribute를 표현한다.

- Mark - line
- Channels
 - vertical position channel for quantitative attribute
 - horizontal position channel for categorical attribute

(b) Scatterplot은 2개의 attribute를 표현한다.

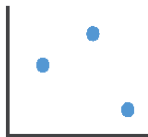
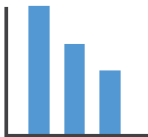
- Mark - point
- Channels
 - vertical and horizontal position channel for quantitative attributes

(c) Additional categorical attribute is encoded to the scatterplot

- Additional channels
 - color channel for categorical attribute

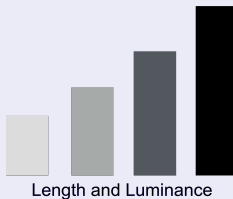
(d) Additional quantitative attribute is encoded to the scatterplot

- Additional channels
 - size channel for quantitative attribute



Multiple channels

- 위의 예시에서는 하나의 attribute가 하나의 channel로만 표현이 되어 있지만 항상 그래야 하는 것은 아니다.



- Multiple channel로 강한 message를 전달할 수 있다. 그러나 과도하게 channel을 소비할 위험이 있다.

Effective Channel

Two principles for channels

- 모든 채널의 효과가 동일한 것은 아니다. 같은 데이터가 서로 다른 채널로 표현되었을 때에 인간의 지각/인지 과정에 따라서 다른 정보를 전달할 수 있다.
- 따라서 marks와 channels은 아래 두 가지의 원칙에 입각해서 정해져야 한다.
- Two principles
 1. Expressiveness: Match channel type to data type
 2. Effectiveness: Some channels are better than others

1. Expressiveness principle

- 데이터셋 attribute의 정보를 모두 표현해야 하고, 그것만을 표현해야 한다. (all of and only the information)
- **Match channel type to data type**
- Ordered data는 magnitude channel을 통해 ordered로 인식되도록 표현되어야 한다.
- Unordered data는 identity channel을 통해 unordered로 인식되도록 표현되어야 한다.

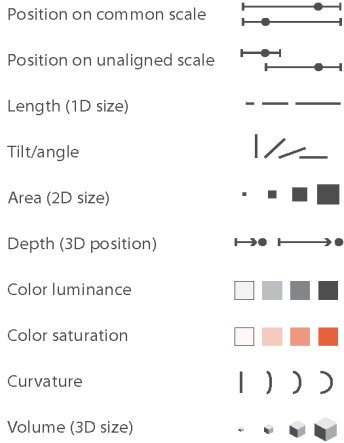
2. Effectiveness principle

- 중요한 attribute는 채널에 특징에 부합하여 눈에 잘 띄어야 (noticeable) 한다.
- **Some channels are better than others**
- 가장 중요한 attribute는 가장 효과적인 채널과 매치하여 가장 눈에 잘 띄게 해야한다.
- Attribute의 중요도가 낮아지는 순서로 덜 효과적인 채널과 매치해야 한다.

Channel rankings

Channels: Expressiveness Types and Effectiveness Ranks

➡ **Magnitude Channels: Ordered Attributes**



➡ **Identity Channels: Categorical Attributes**



Channel Effectiveness

Measures

- Motivation
 - 직전 슬라이드의 ranking은 어떻게 정당화 될 수 있는가?
 - 다양한 channel들은 각각 얼마나 많은 정보를 표현할 수 있는가?
 - 왜 어떤 channel이 다른 채널보다 우월한가?
 - 여러개의 channel들이 함께 사용할 수 있는가? 아니면 독립적으로 사용되어야 하는가? (compatibility)
- 위의 질문은 각 채널이 아래의 관점에서 얼마나 effective한지에 대한 분석을 요구한다.
 1. Accuracy
 2. Discriminability
 3. Separability
 4. Visual popout을 제공할 수 있는지의 여부

1. Accuracy

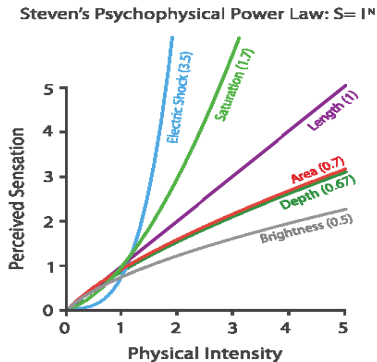


그림 1: Stevens 75, p.15

- 인간의 지각채널은 물리적 강도를 재해석하여 받아들인다.
- Super-linear
 - Electric Shock
 - Saturation
- Linear
 - Length
- Sub-linear
 - Area
 - Depth
 - Brightness

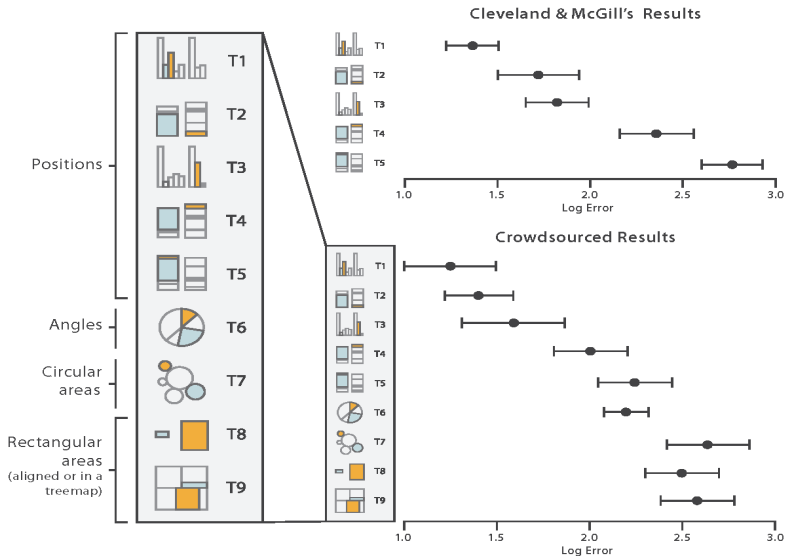


그림 2: Cleveland and McGill 84a과 Heer and Bostock 10

2. Discriminability

- Visual channel이 정해진 경우에 item 들 간의 차이가 인식되는가?
- 각각의 visual channel에 따라서 number of bins의 적정 수치가 있다.
- Some channels have a very limited number of bins.
- Ex) Line width
 - Changing line size only works for a fairly small number of steps
 - Too thick line looks like polygon.
 - Next figure provides example of linewidth.
 - Linewidth works only up to three to four different values for an attribute.

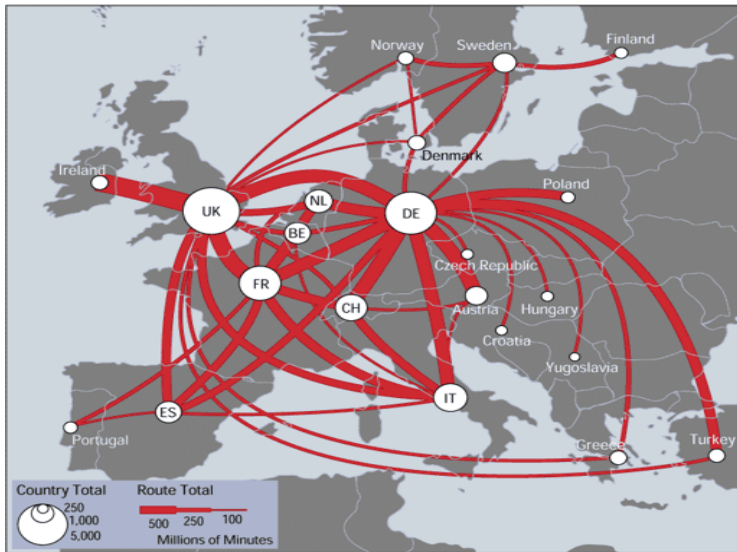
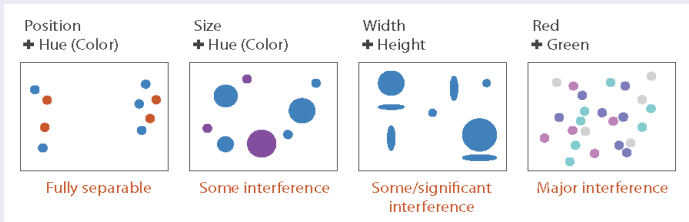


그림 3: Linewidth has a limited number of discriminable bins (Fig 5.9.)

3. Separability

- Not all visual channels are compatible.
- Some channels are independently **separable**.
- Some channels interact.

Separable and interacting



- 1) Position + Color - Fully separable
- 2) Size + Color - Some inference
 - For small circles, hard to perceive color.
 - Size interacts with many visual channels, including shape.
- 3) Width + Height - Some/significant inference
 - We perceive planar size (area) before width and height.
 - We perceive small/medium/large, instead of thick/thin and short/tall.
- 4) Red + Green - Major inference
 - We perceive green as green, instead of yellow + blue.
 - This example simply leads to four different color.
 - Mixing color is poor idea.

4. Popout

- Many visual channels provide visual **popout**, where a distinct item stands out immediately.

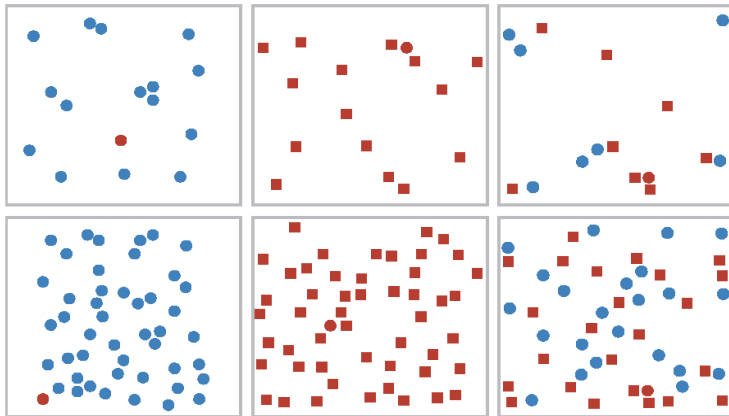
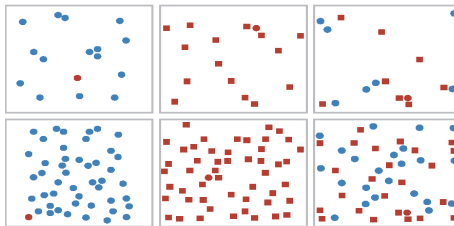
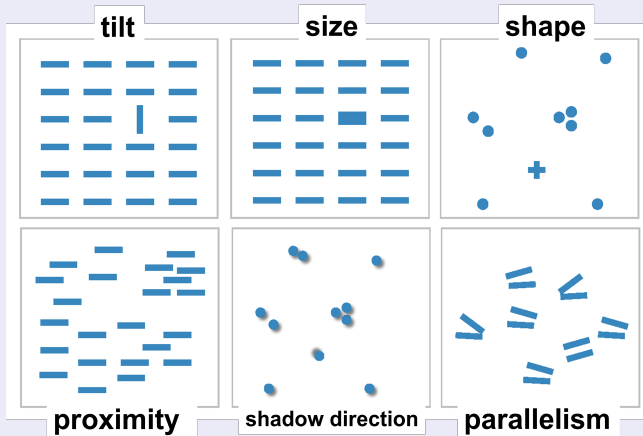


그림 4: Spotting a red circle



- (Left) Spotting a red object from a sea of blue ones.
 - Very quick in both
- (Middle) Spotting a circle from a sea of squares.
 - Not much different spotting time from small or large set.
- (Left) vs (Middle)
 - (Middle) is slower than (Left).
 - The difference between red and blue on the color channel is larger than the difference in shape between filled-in circles and filled-in squares.
- (Right)
 - Number of distracting objects matter.
 - From the large set, it takes **serial search**. The search time is linear to the number of distracting objects.

Many channels support visual popout.



Multiple channels

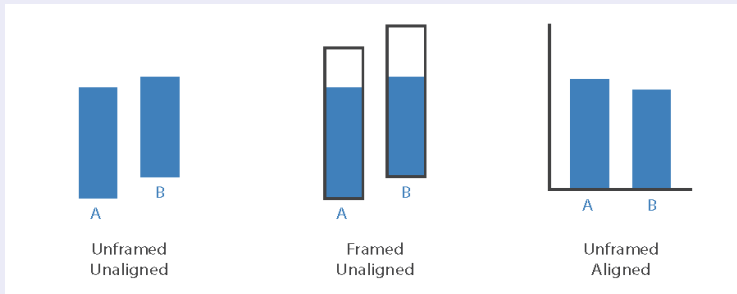
- Most pairs of channels do not support popout
- A few pairs do:
 - space and color
 - motion and shape
- Definitely not possible with three or more channels
- Vis designer should count on using popout for a single channel.

Relative vs. Absolute judgement

Relative vs. Absolute judgement

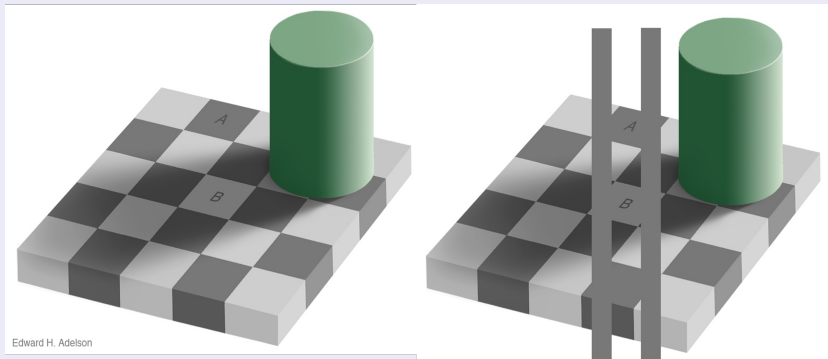
- Human perception is based on relative judgements, not absolute ones.
- This principle holds true for all sensory modalities.
- Weber's law
 - $K = \frac{\delta I}{I}$, where K is detectable difference and I is stimulus intensity.
 - For instance, the amount of length difference we can detect is a percentage of the object's length.
- When considering accuracy and discriminability, we must distinguish relative and absolute judgements.

Position along a scale can be more accurately



- (left) Difficult to make length judgement
- (middle) Easier with framing
- (right) Aligning provides a common scale

Luminance is completely contextual



- (left) A and B appear different
- (right) With the aid of gray mask, they are in fact identical.

```
"Data Visualization"
```

```
## [1] "Data Visualization"
```