Communicating Research

CS 197 | Stanford University | Michael Bernstein

Special thanks to Kayvon Fatahalian and Lauren Gillespie for slides

Administrivia

Draft talk is **due in section** the Tuesday after Thanksgiving

- Six minutes
- You'll be getting feedback from each other and from the TA

Draft talk due in week 10, final paper and talk due during finals

Today's goals

How to structure and deliver a stellar research talk

So, you’re giving a talk.

Talks are how people get to know your work

A lot of people might scroll past your social media post about your paper

Few are masochistic enough to scroll through every paper in the archives to run into yours

People might not make it to your poster

But, **we have a culture of showing up to and listening to talks**.

Clarity is highly prized

Great talks get work more attention, adoption, and conversation

“Hey, did you see X’s talk yesterday?”

“Hey, great talk—we should chat about an internship.”

It's a huge shame if you spend months on the project, but can’t go the last mile to communicate it to others

Michael’s take: one **difference between good and great researchers** is being able to explain how and why the work matters

And now, principles and tips

Tip 1

Identify your one narrative goal

What do they need to know?

Aim for the audience to understand everything you're saying, and to **feel smarter** at the end of the talk

> "Aha! I understand something that I didn't understand before!"

Even if you are targeting experts in your area, they have not been spending months thinking about this topic

Every step of the talk should be leading them up to an understanding, or helping them understand the consequences of that understanding

> A talk is not a step-by-step report—it is an **explainer**

What is the one idea they should walk away with?

You have spent a lot of time in the guts of your project, and know all the details and the caveats and the webs of relationships. Unfortunately, the audience will remember none of that

They will **remember one moment—one idea—from your talk**. What will that be?

"This claim."

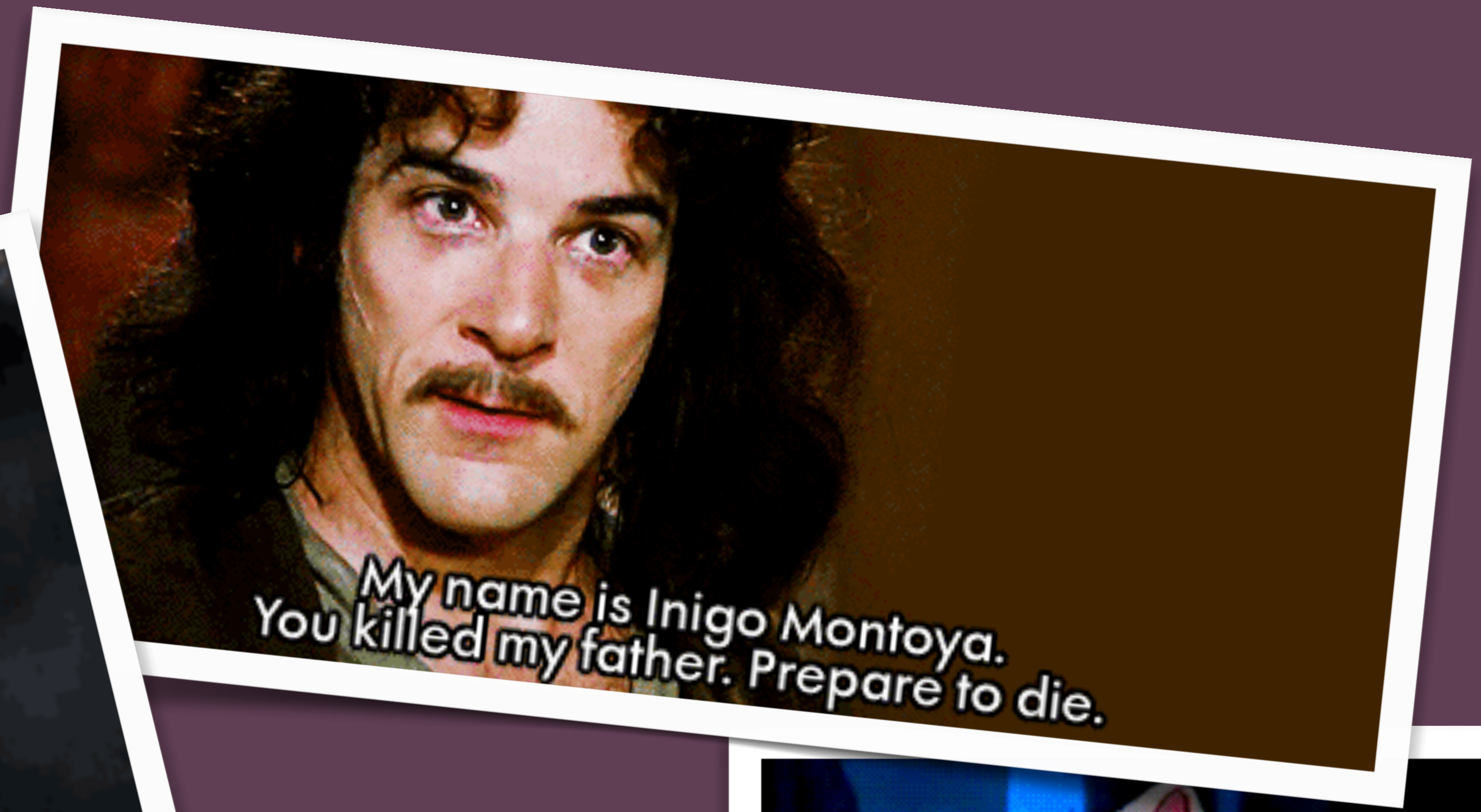
"This diagram."

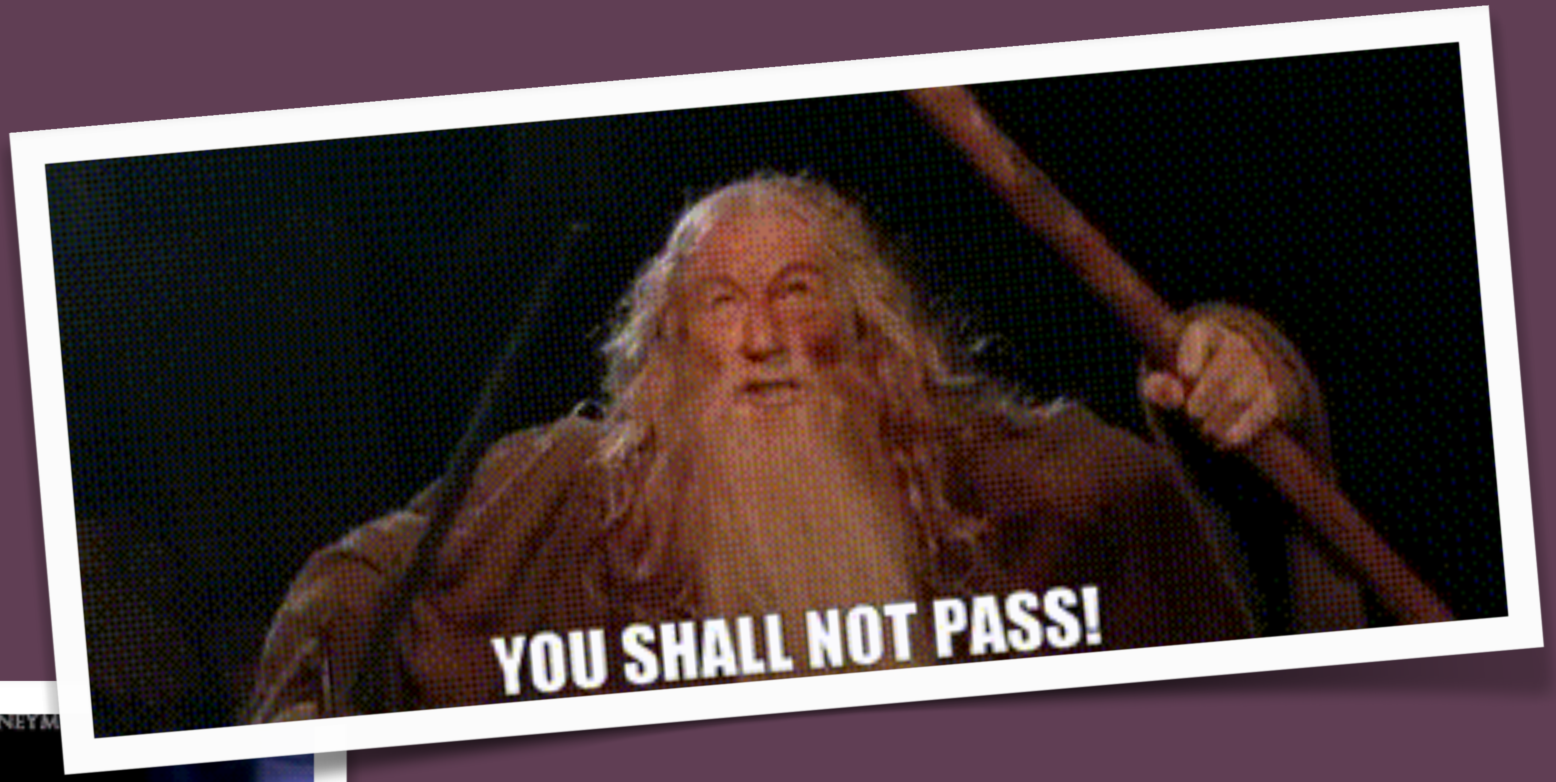
"This graph."

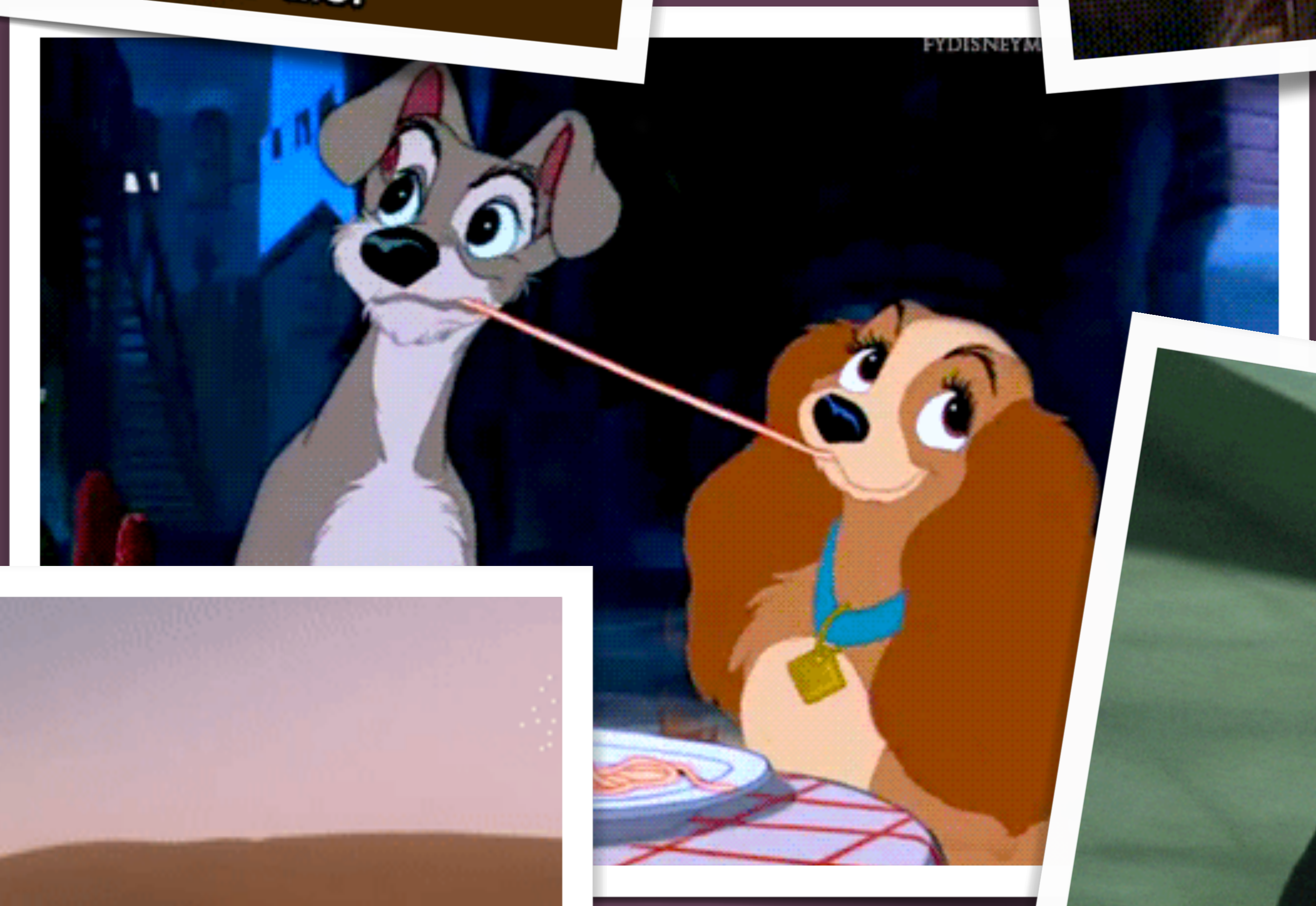
"This result."

Movies do this too

What is the one image or moment that the whole movie was built around?







Work backwards from your one idea to design the talk

Once you know what the high point of your talk will be, ask yourself:

- What is necessary (and what's not) to lead up to it?
- How do I highlight that it's the high point so that they remember it?
- How do I hit a denouement afterwards?

This means that **your talk will likely not cover everything**

Tip 11

Every sentence matters.

Intros are often very poor

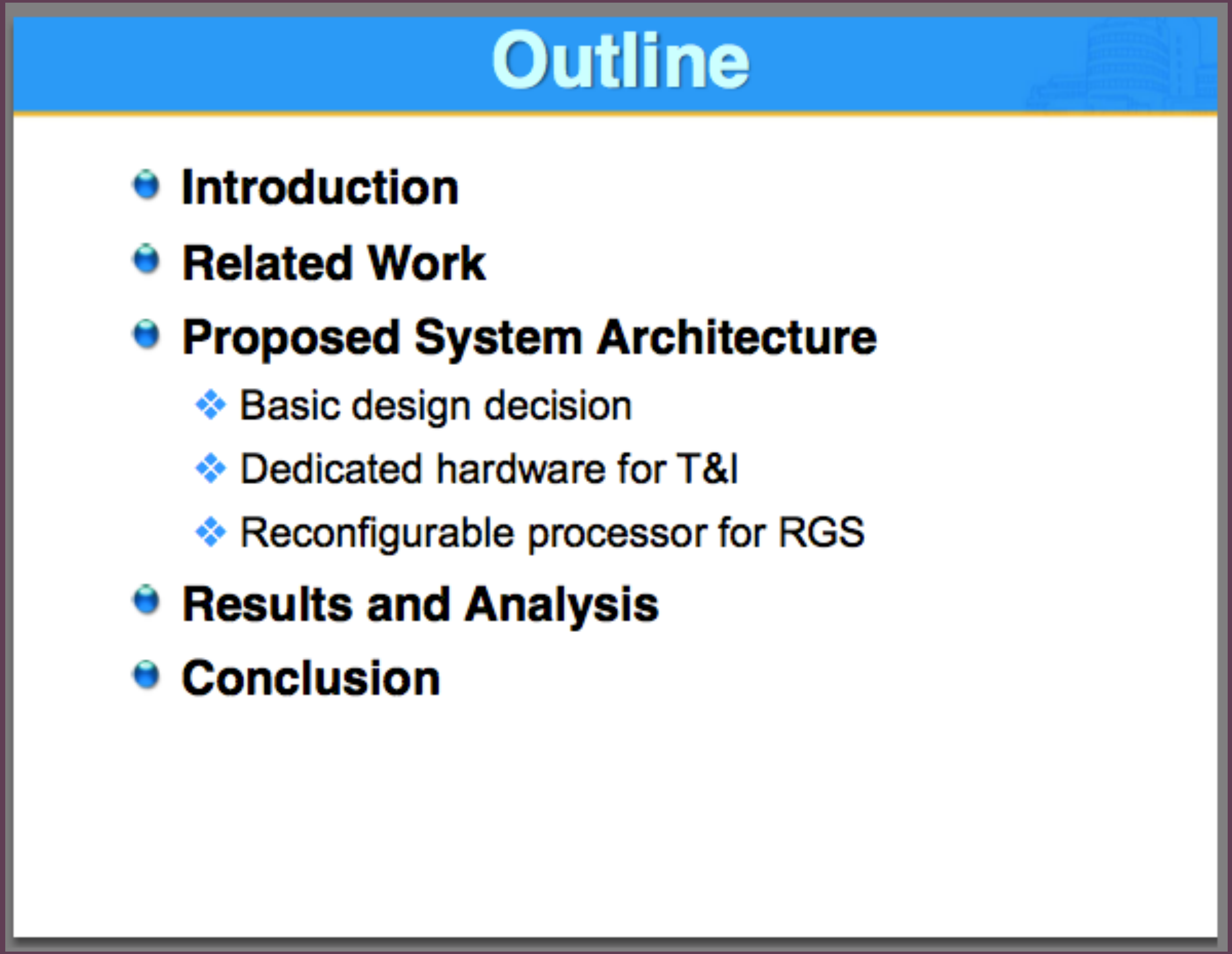
"LLMs have been shown to have good task performance on a variety of tasks…" Everyone knows this. So don't say it.

Unlike a paper, related work in a talk does not exist for academic completeness

The goal of the intro and background is to tell the listener: **"In this talk, here is the way I want you to think about the problem I am trying to solve."**

Bad example 1

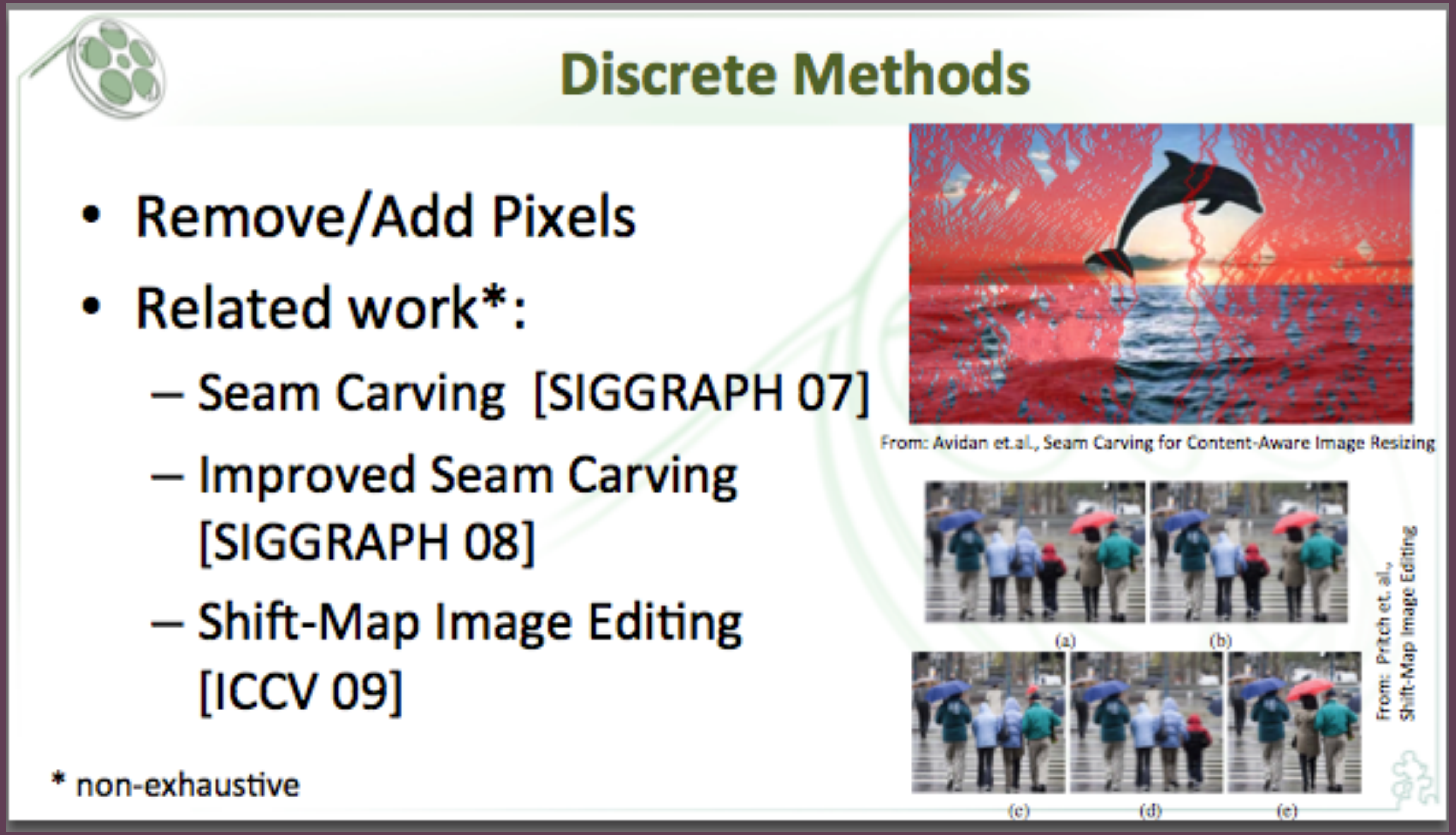
Never do this:



All I now know is that you will have the talk outline I would have already expected

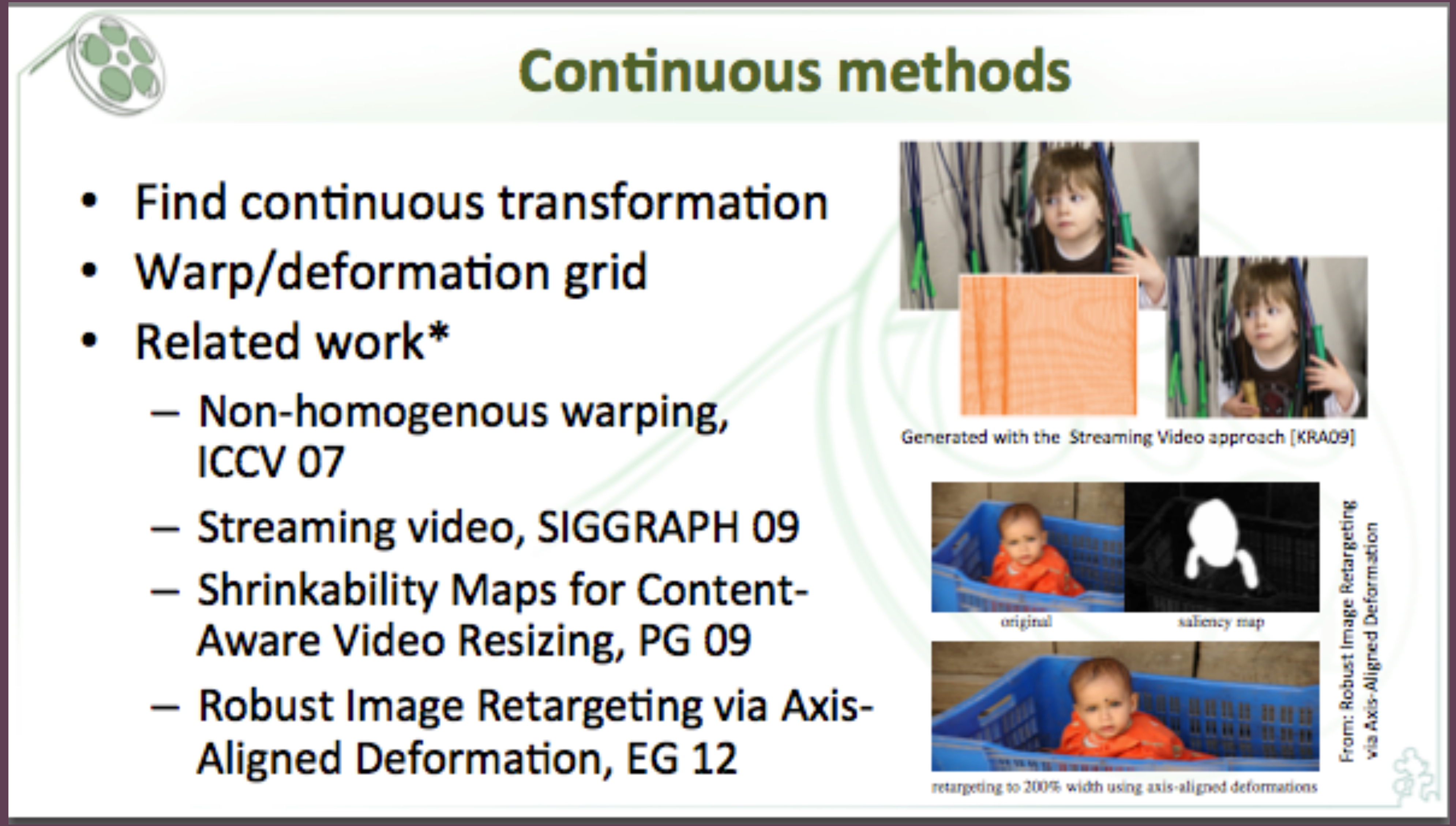
Bad example 2

What is the idea or bit flip that you're building towards?



Experts already know these papers exist, and non-experts aren't going to learn these papers from seeing their names

This type of related work section says little more than "others have worked in this area before". I suspect your audience already assumes this is the case.



"Think of it this way:"

The goal of the intro and background is to tell the listener: "During this talk, here is the way I want you to think about the problem I am trying to solve."

An excellent strategy to catch the audience's attention and frame the story is to make them aware that there is something they didn't know they didn't know.

A better example [Mullapudi 2019]

The talk intro you're about to see asked the audience to consider the following question:

> *Why are we working so hard to train general ML models that work in many different situations?*

> *Why don't we just use simple, less general ML models and retrain them constantly for the specific conditions at hand?*

Here's the sequence used to do that…

What a traffic camera sees

It seems like it should be easy to train a simple, low-cost ML model to detect cars/ pedestrians/etc in this one scene…right?

Problem: distribution shift

But any one scene can change dramatically due to weather, time-of-day, types of vehicles, etc…

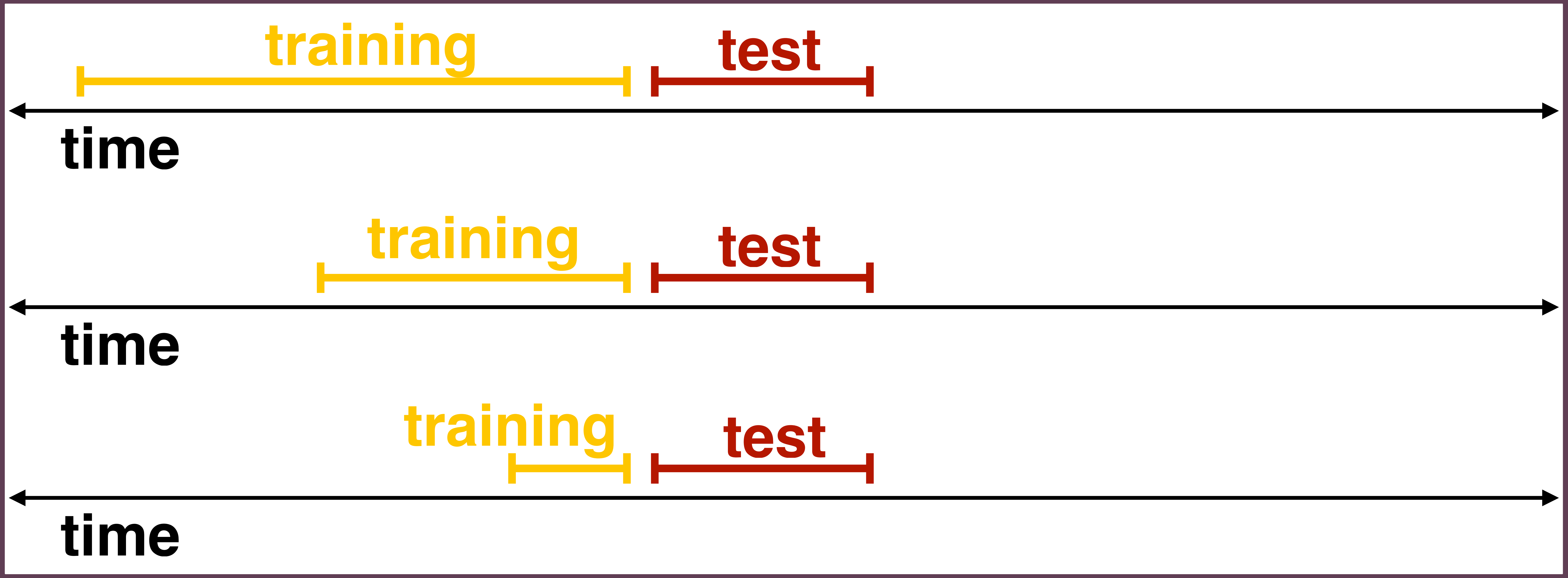
We need to acquire a diverse training set…right?

Consider this experiment

Plop camera down in a new environment

We want a specialized model for processing the stream from this camera

How much training data is needed to train an accurate model?



Now the audience is thinking, and ready for you to help them feel smart for figuring it out

Continuous adaptation!

The premise: low cost efficient models can retain high accuracy for complex tasks in challenging environments if they are continuously specialized to the contents of video streams

> a.k.a. Don't worry about carefully curating the perfect training set up front, just make sure you can adapt quickly online when you see it

Audience member: "But how do I adapt an ML model quickly on the fly with little data or supervision? **That sounds hard.**"

> "Ah, that's the point of the talk…"
>
> "This speaker has something to say, maybe I shouldn't check my phone."

Now that it's framed…

…related work can now be discussed in the context of this framing

Establishing framing is the primary value of the talk intro

In this example: how have other folks tried to quickly adapt ML models to new contexts?

- The "meta learning" field tries to train models that can be quickly adapted later using a few examples…
- There's a whole body of work on "domain adaptation": taking a model that works well in context X and modifying it to work well in Y…
- Instead, we adopt an approach based on "model distillation"

Tip III

Inputs, outputs, constraints

Establish goals early

"Given these **inputs**, we wish to generate these **outputs**…"

"We are working under the following **constraints**"

- Example: the outputs should have these properties…
- Example: the computer graphics algorithm...
  - Should run in real time
  - Should be widely parallelizable so it can run efficiently on a GPU
- Example: the system…
  - Need not compile all of Python, only this subset that we care about…
  - Should realize about 90% of the performance of hand-tuned code

Why goals and constraints?

Your contribution is typically a system or algorithm that meets the **stated goals under the stated constraints**

Understanding whether a solution is "good" requires having this problem context.

Tip IV

Show, don’t tell

Example

In a recent project, Kayvon Fatahalian and team asked the question: given enough video of tennis matches of a professional athlete, could we turn it into a controllable video game character?

Words are ambiguous. Consider possible listener questions?

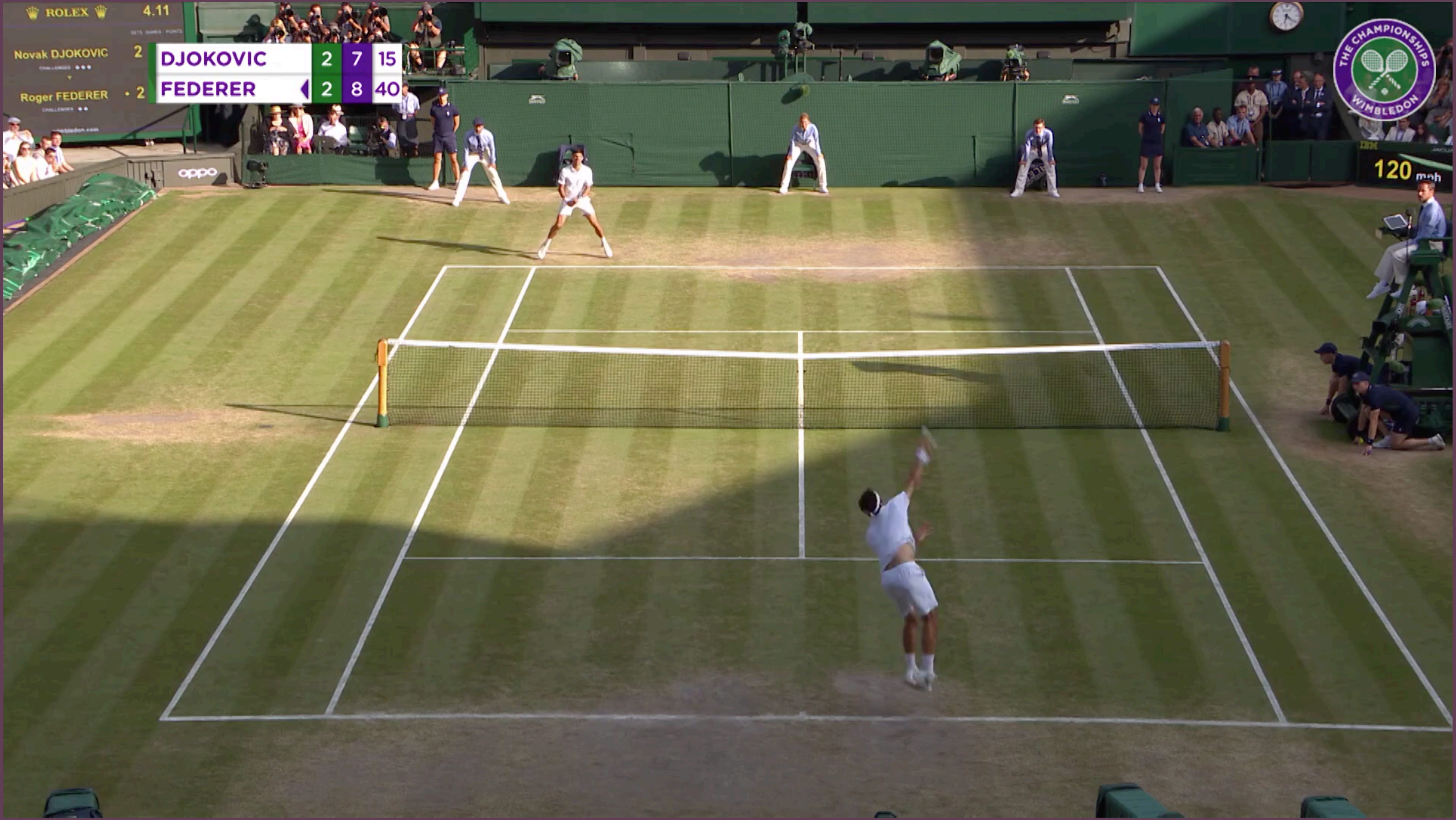
What does this video contain? Is it easy to analyze?

What do you mean by controllable characters?

What output are you trying to create? Video? 3D animation?

Compare the description above to the following sequence…

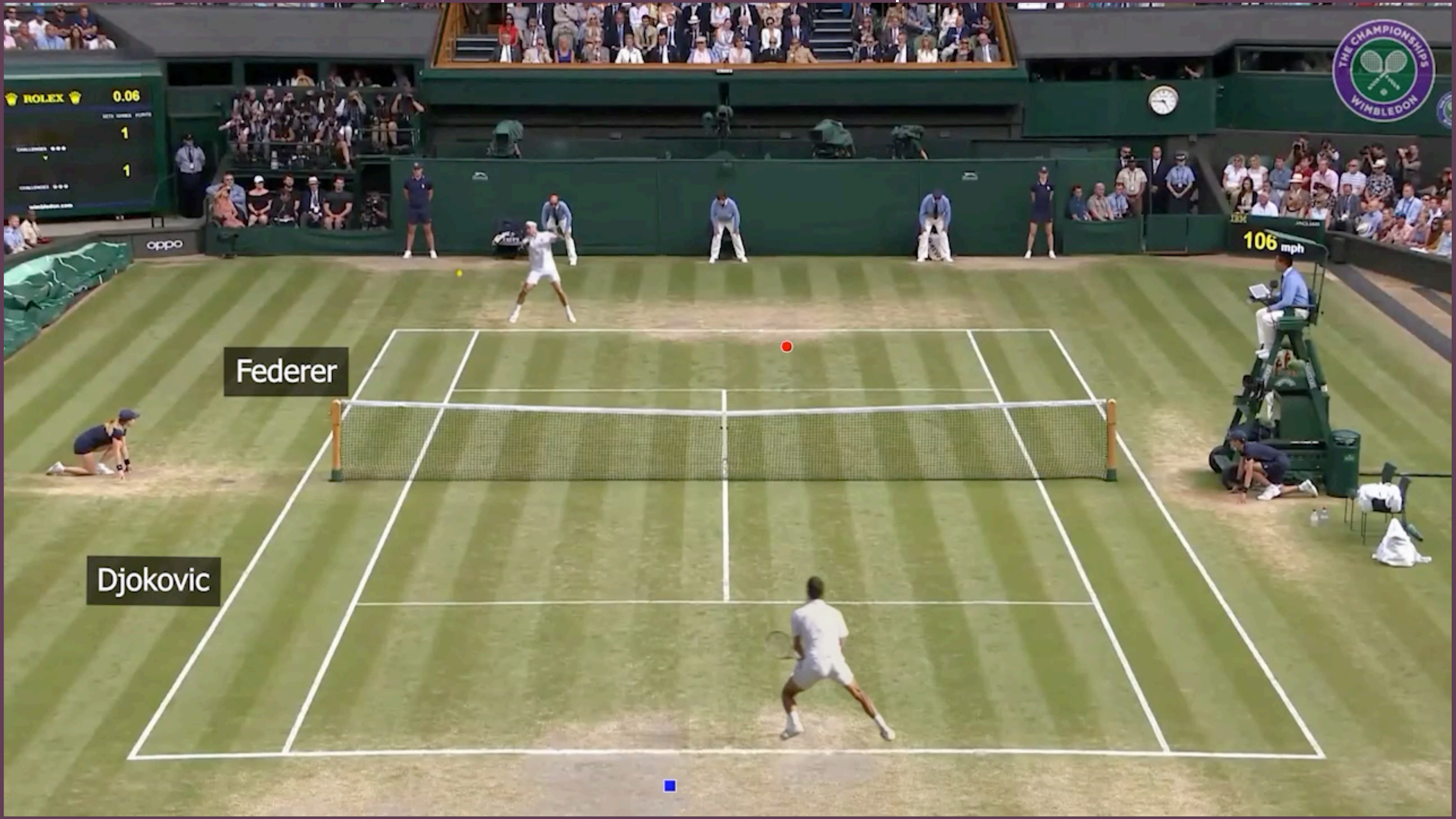
Example source video



The best way to describe the input data is just show it!

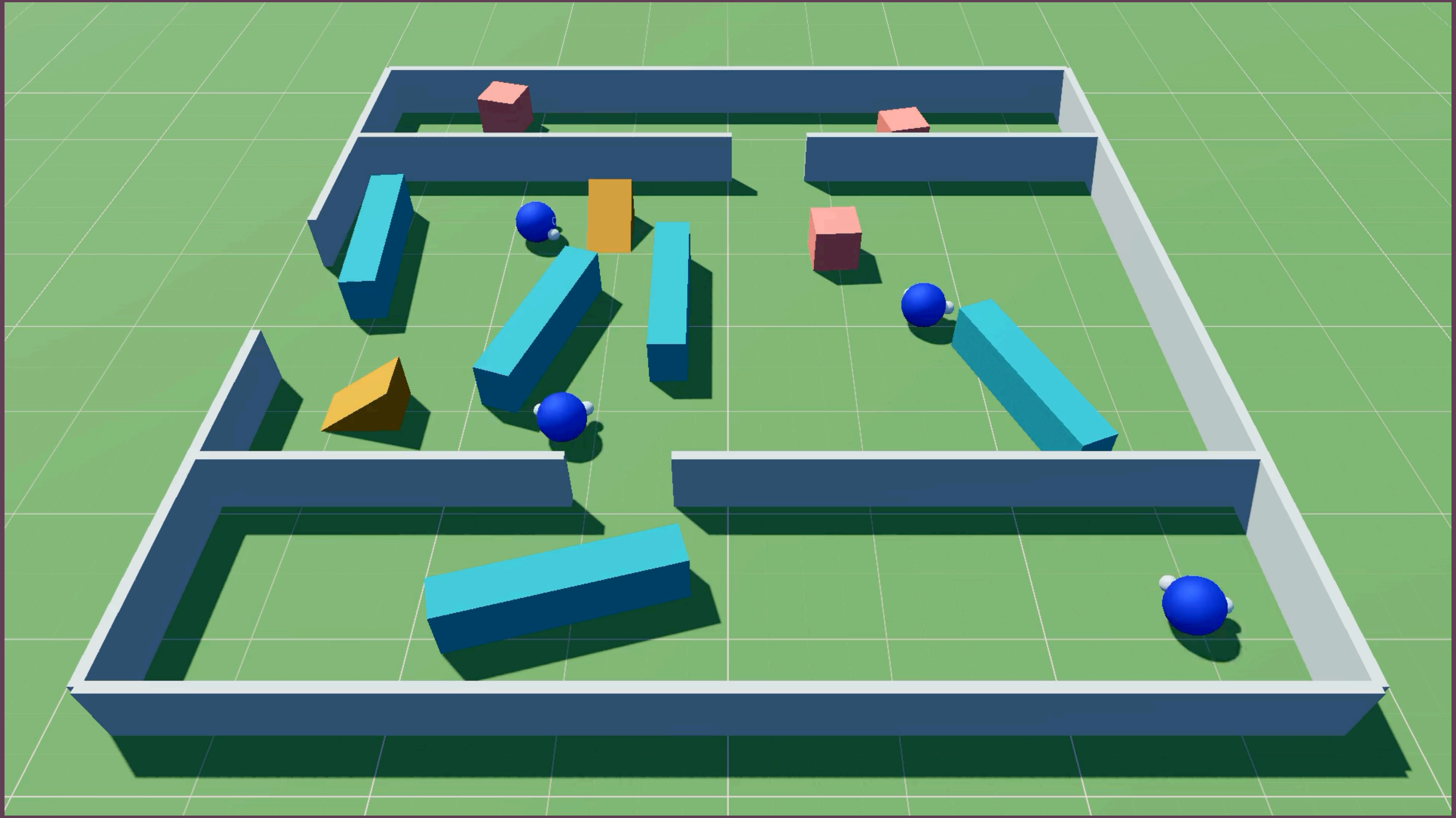
And there’s a lot of it!

Here's an example of controllable output



The best way to describe the output we seek is just show the result of the system

Another example

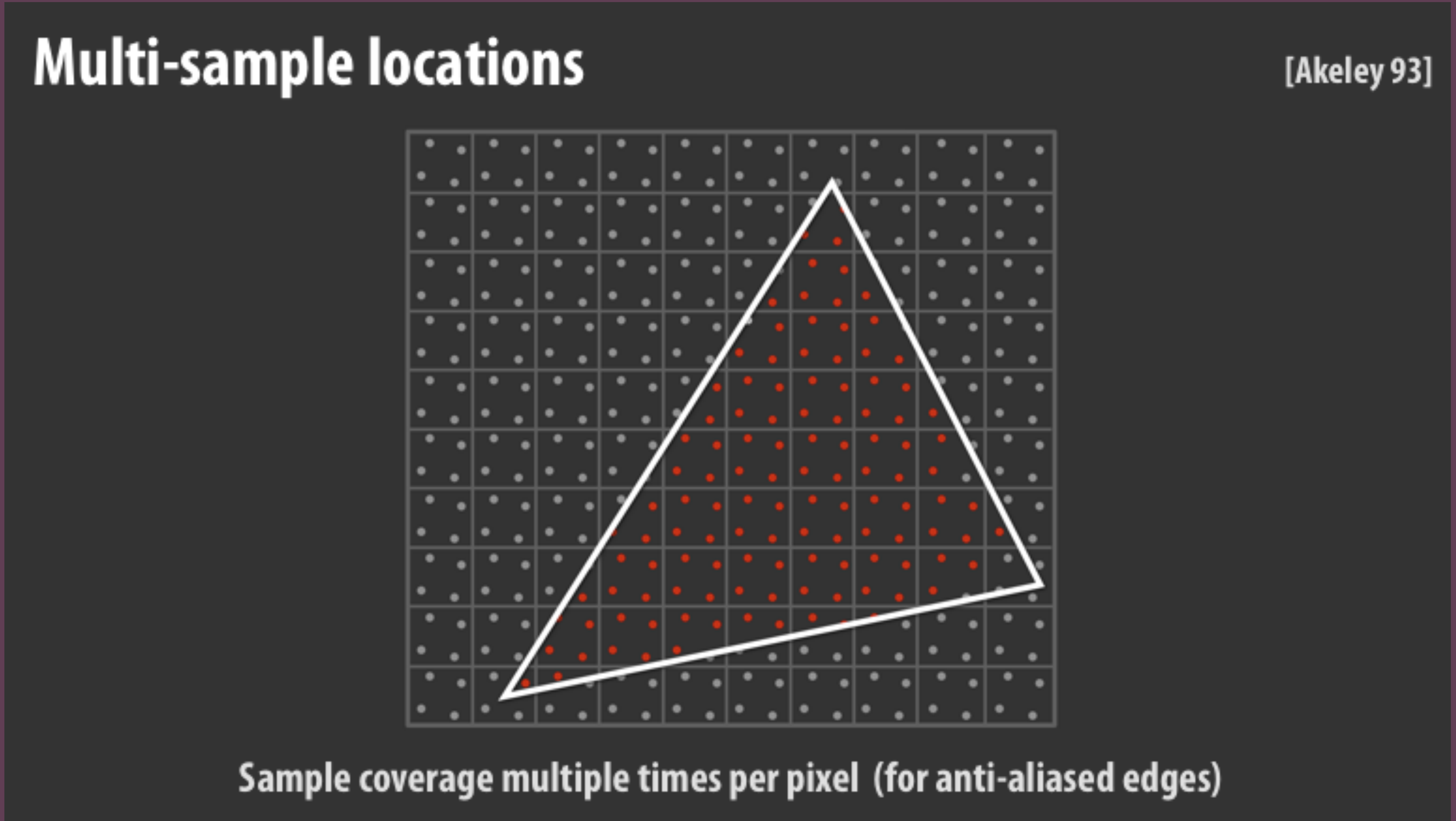


A game engine that can play many games in parallel simultaneously

Tip V

Always explain any figure or graph

Walk the audience through any figure

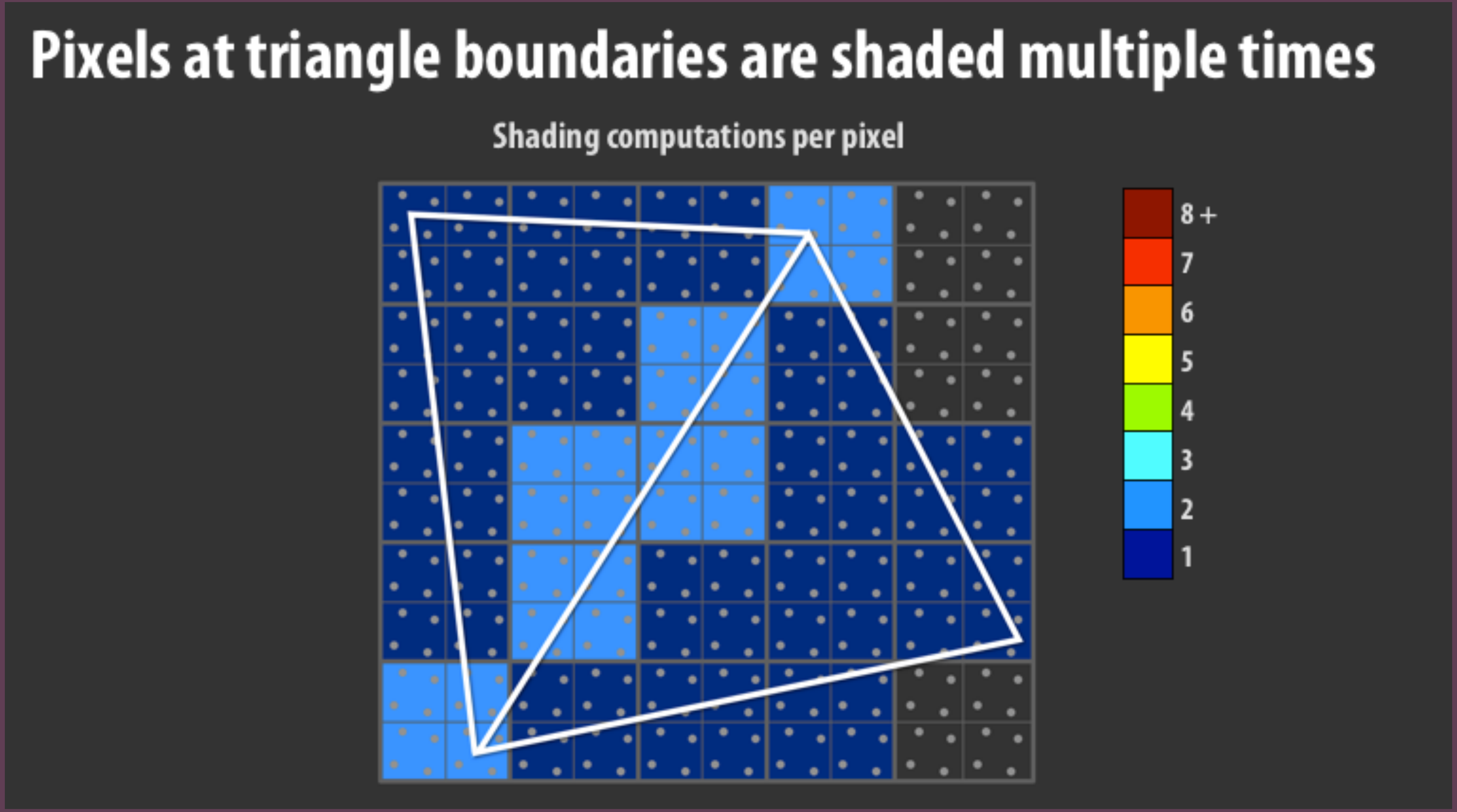


What are we looking at?

Looking at this, you don't know what it is or why I'm showing it to you

How about: "Here I'm showing you a pixel grid, a projected triangle, and the location of four sample points at each pixel. Sample points falling within the triangle are colored red."

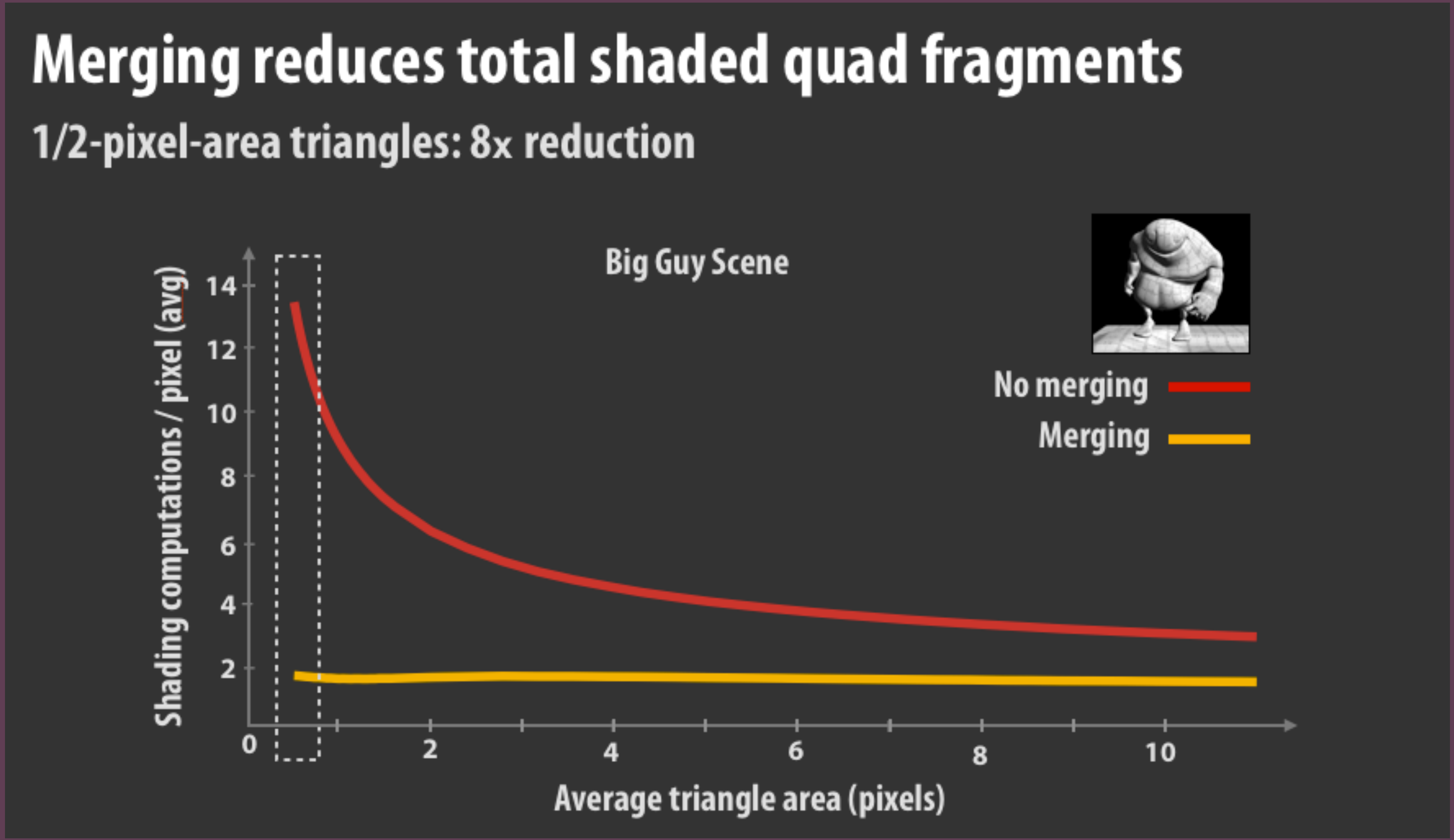
Walk the audience through any figure



What should I take away?

"Now I'm showing you two adjacent triangles, and I'm coloring pixels according to the number of shading computations that occur at each pixel as a result of rendering these two triangles. As you can see from the light blue region, pixels near the boundary of the two triangles get shaded twice."

Walk the audience through any figure



Explain the axes and result I should notice

"This figure plots the number of shading computations when rendering tessellations. X-axis gives triangle size. If you look at the left side of the graph, which corresponds to a high-resolution micropolygon mesh, you can see that merging, shown by yellow line, shades over eight times fewer per pixel."

Tip VI

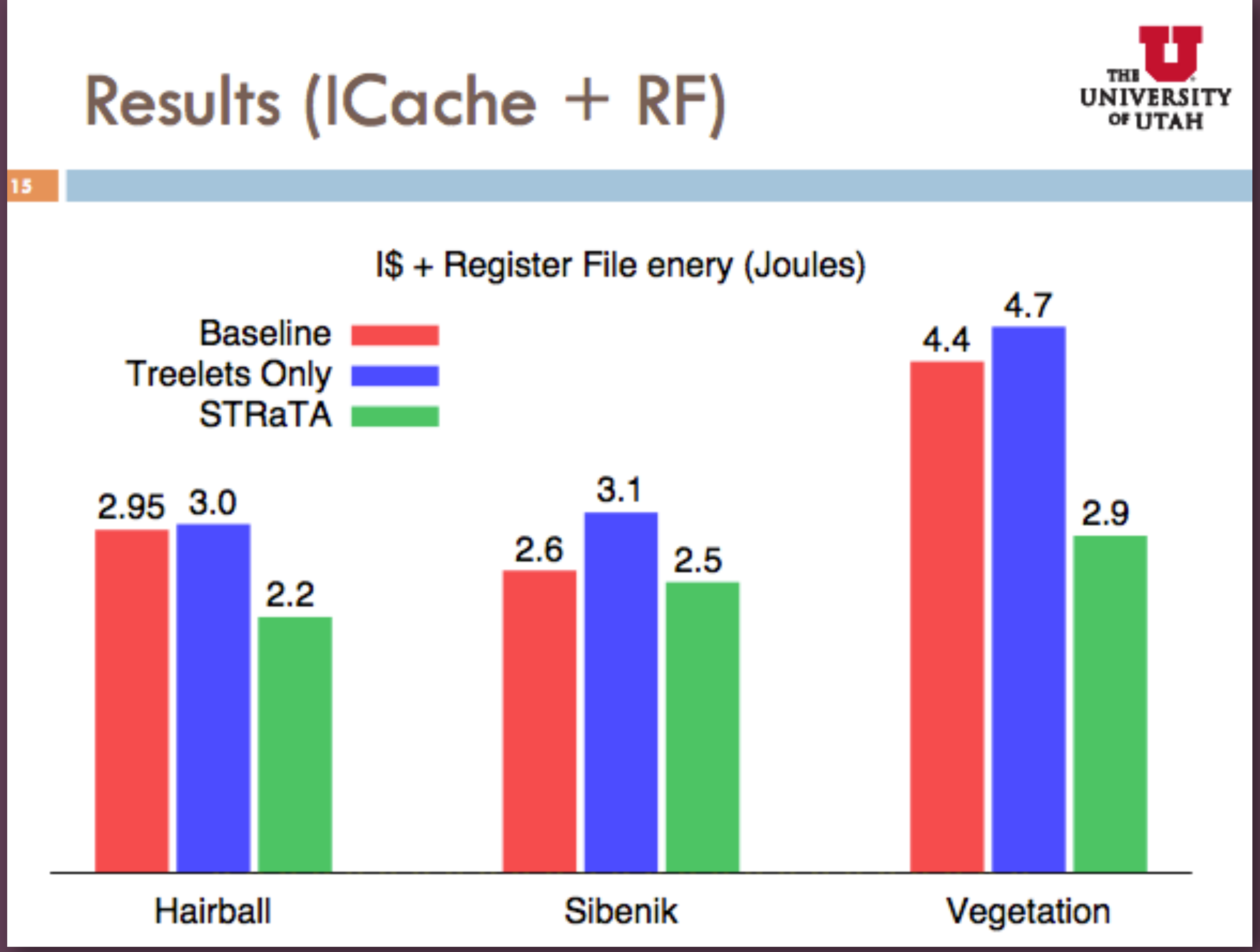
One point per slide
One point per slide
One point per slide

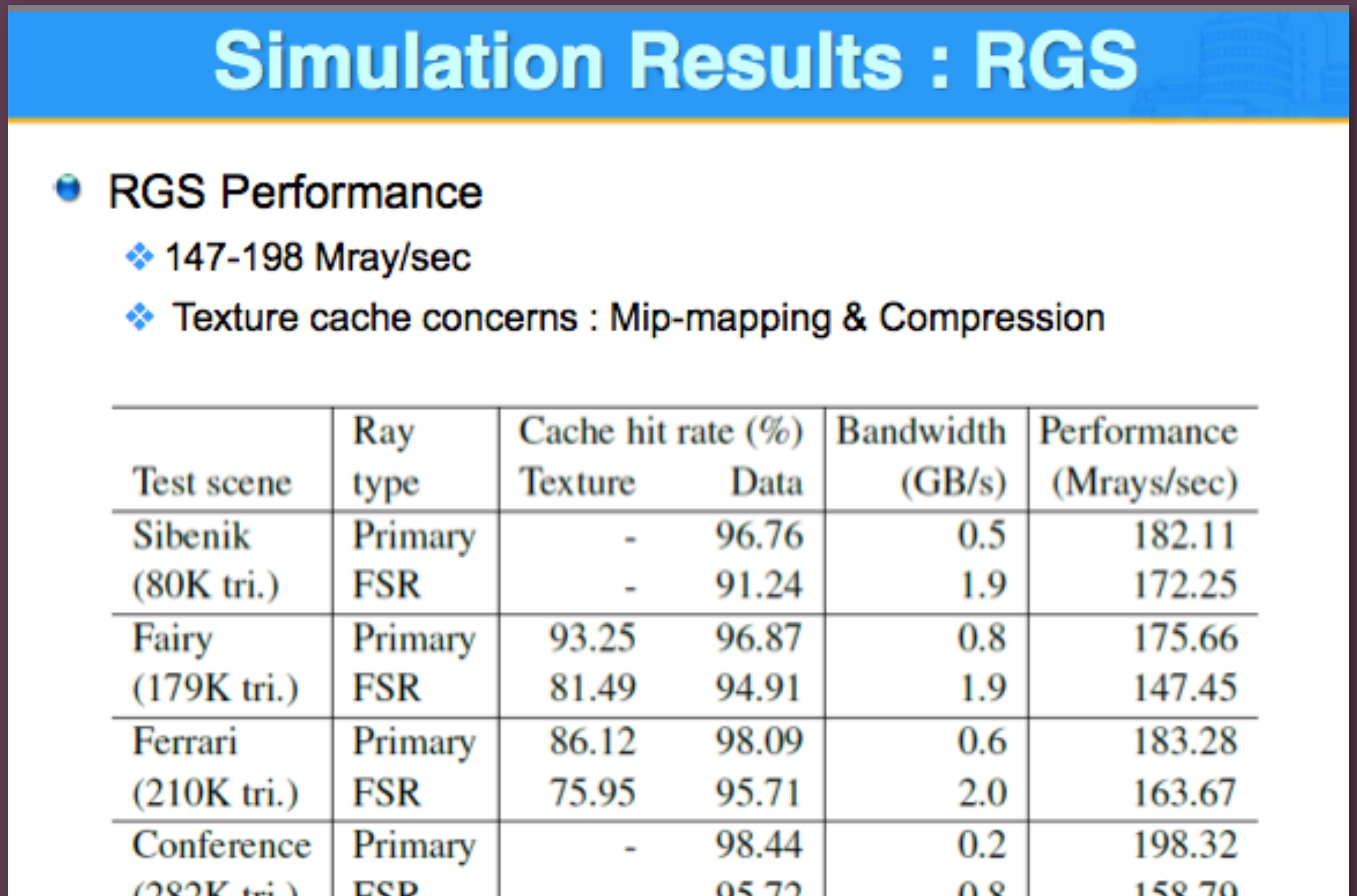
Don’t do this:

Results

1. Quantitative
    1. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec metus dolor, accumsan feugiat porta a, placerat ut risus. Curabitur eu sapien
2. Qualitative
    1. Integer sit amet justo risus. Nunc at facilisis urna. Pellentesque cursus et lorem sed fringilla. Etiam tincidunt tortor at tincidunt iaculis. Maecenas
3. Performance
    1. Nullam odio turpis, ultricies a arcu id, feugiat pretium sem. Nullam arcu erat, faucibus id laoreet ullamcorper, lacinia vitae ligula.

Bad examples

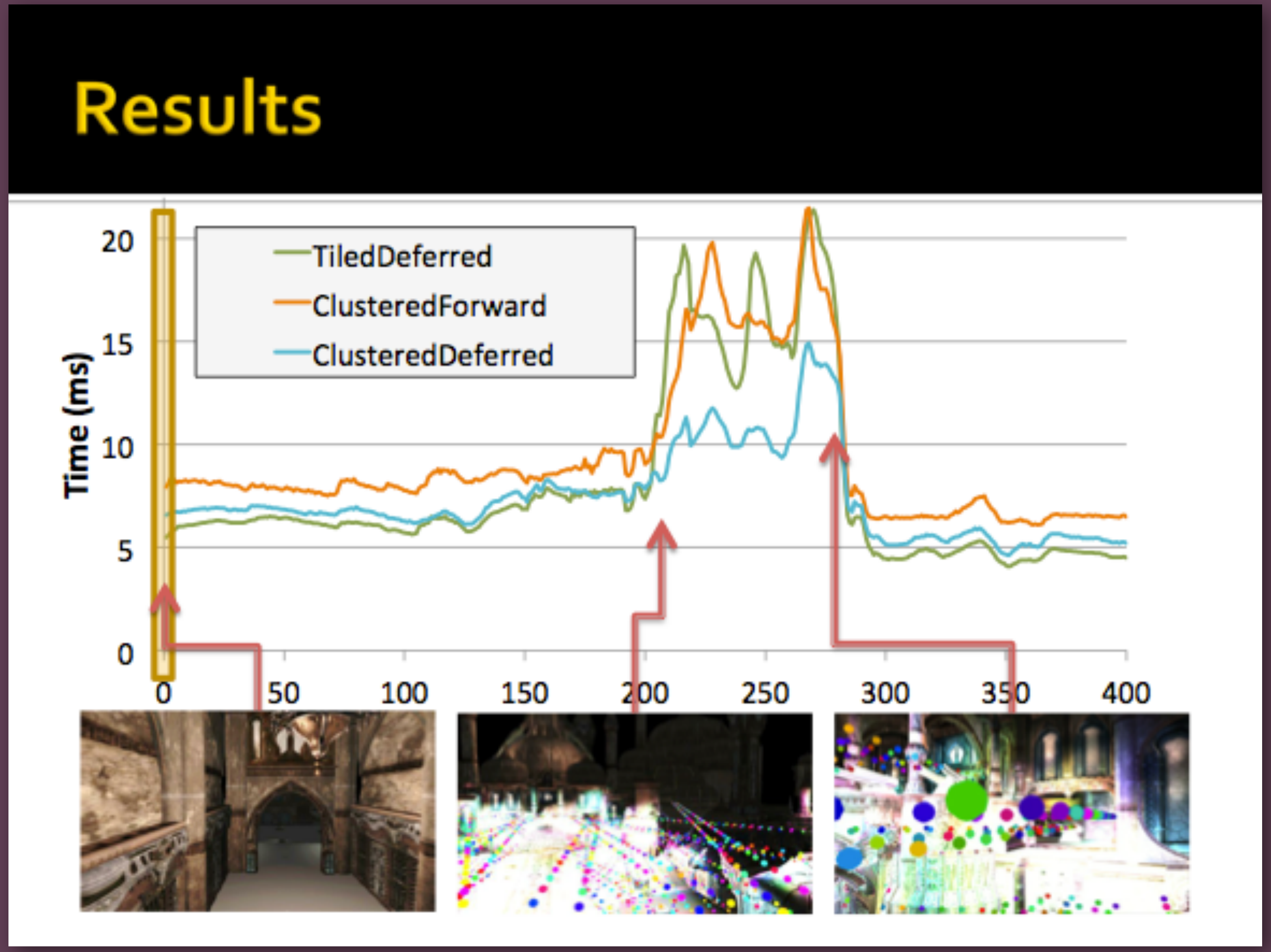




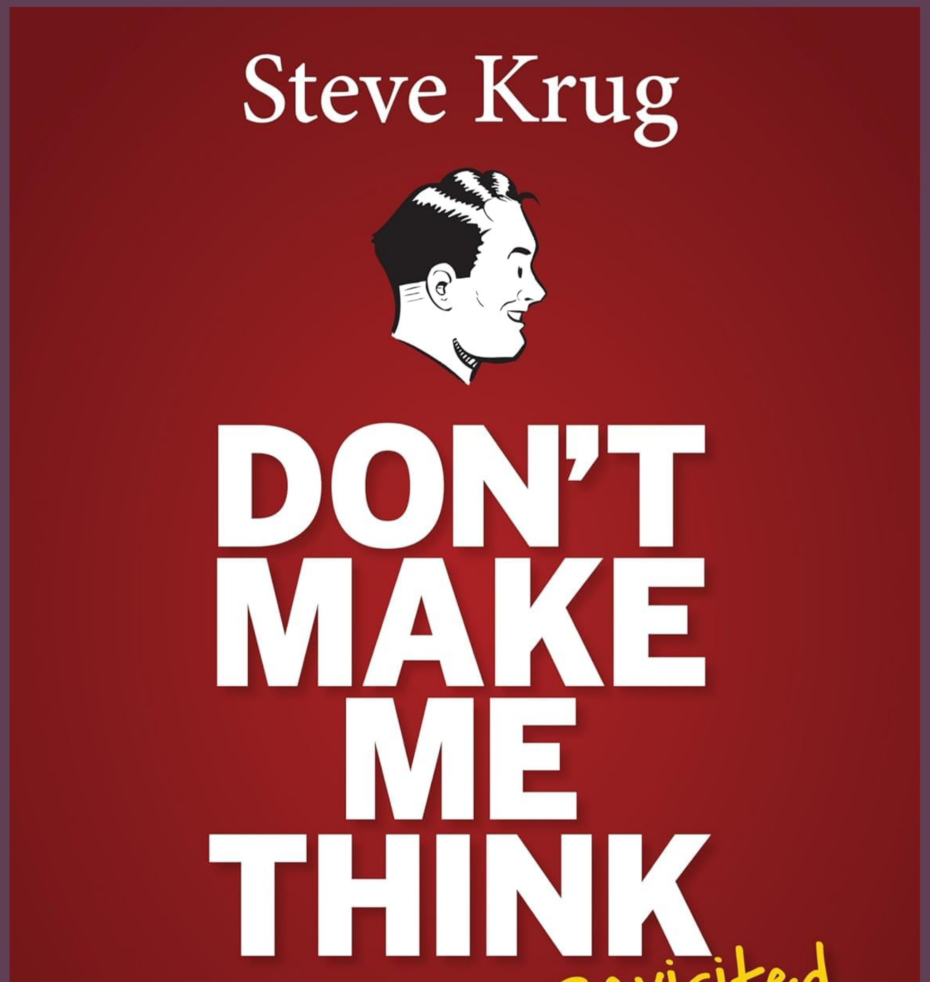
Simulation Results : RGS

- RGS Performance
  - 147-198 Mray/sec
  - Texture cache concerns : Mip-mapping & Compression

Test scene	Ray type	Cache hit rate (%) Texture	Data	Bandwidth (GB/s)	Performance (Mrays/sec)
Sibenik	Primary	-	96.76	0.5	182.11
(80K tri.)	FSR	-	91.24	1.9	172.25
Fairy	Primary	93.25	96.87	0.8	175.66
(179K tri.)	FSR	81.49	94.91	1.9	147.45
Ferrari	Primary	86.12	98.09	0.6	183.28
(210K tri.)	FSR	75.95	95.71	2.0	163.67
Conference	Primary	-	98.44	0.2	198.32
(282K tri.)	FSR		95.72	0.8	158.79

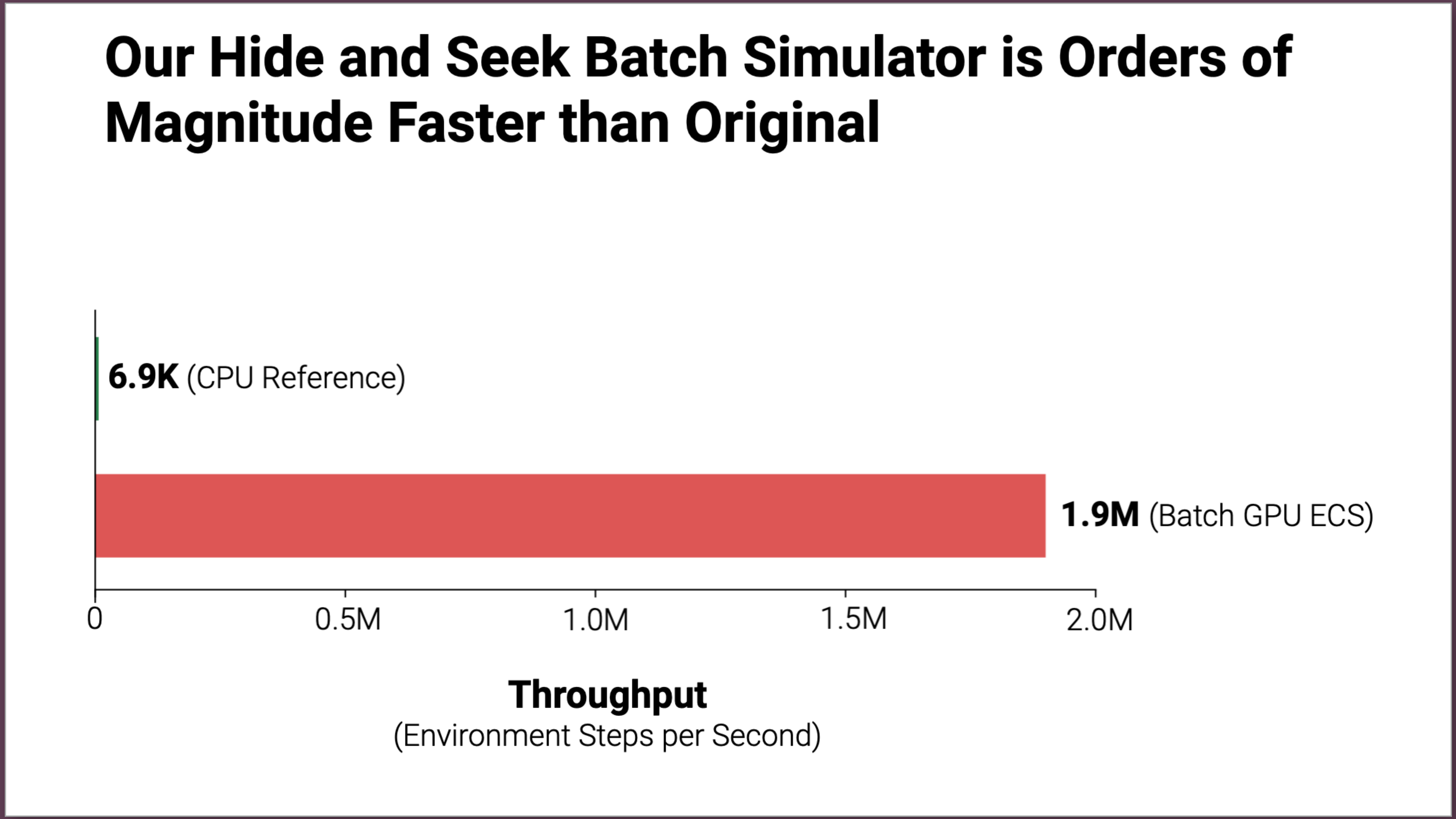


You have to work very hard to interpret these data—what are you supposed to take away?

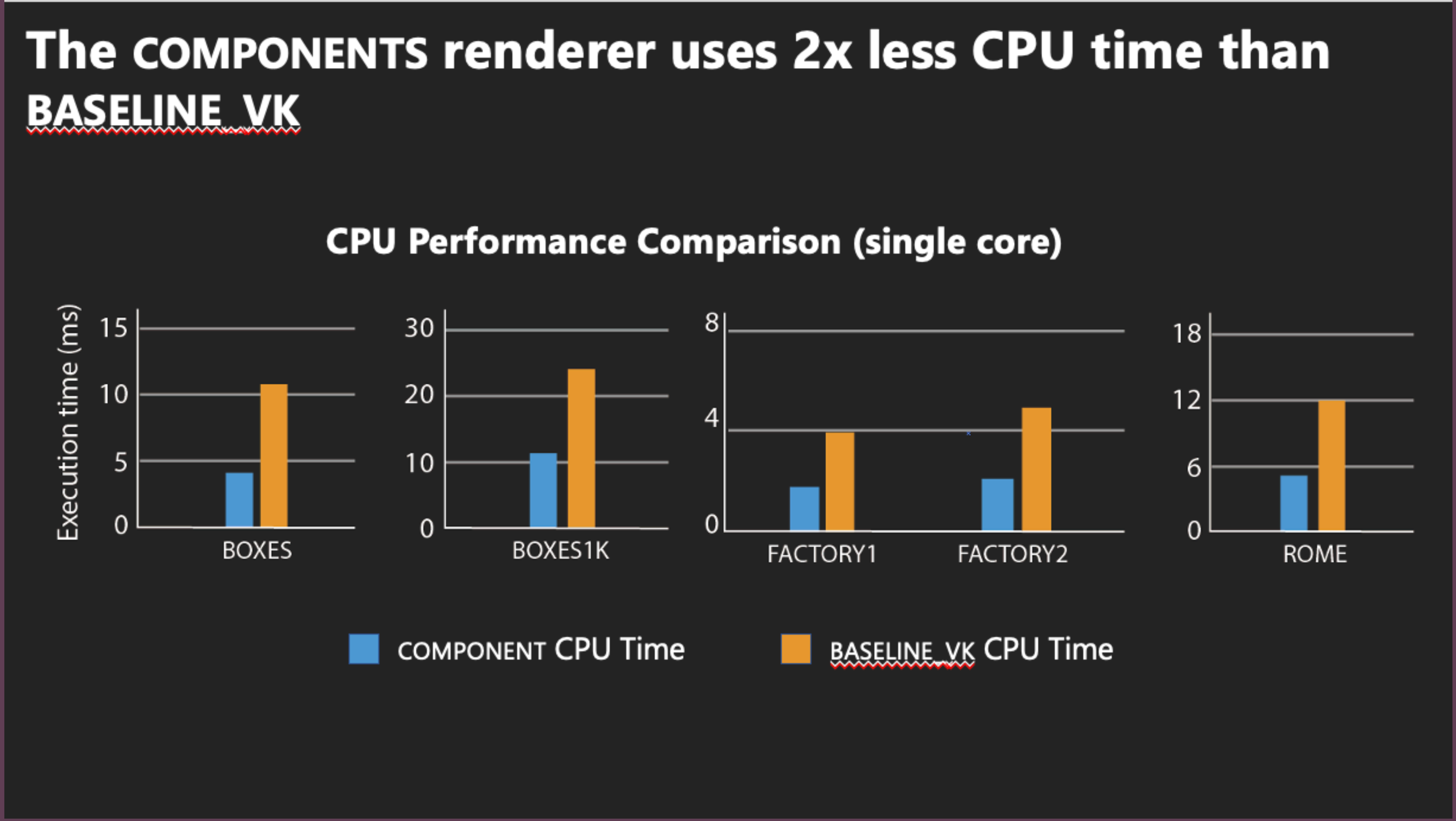


The point is the slide title.

Every slide should have a thesis you are proving on that slide, and that thesis should be the title of the slide

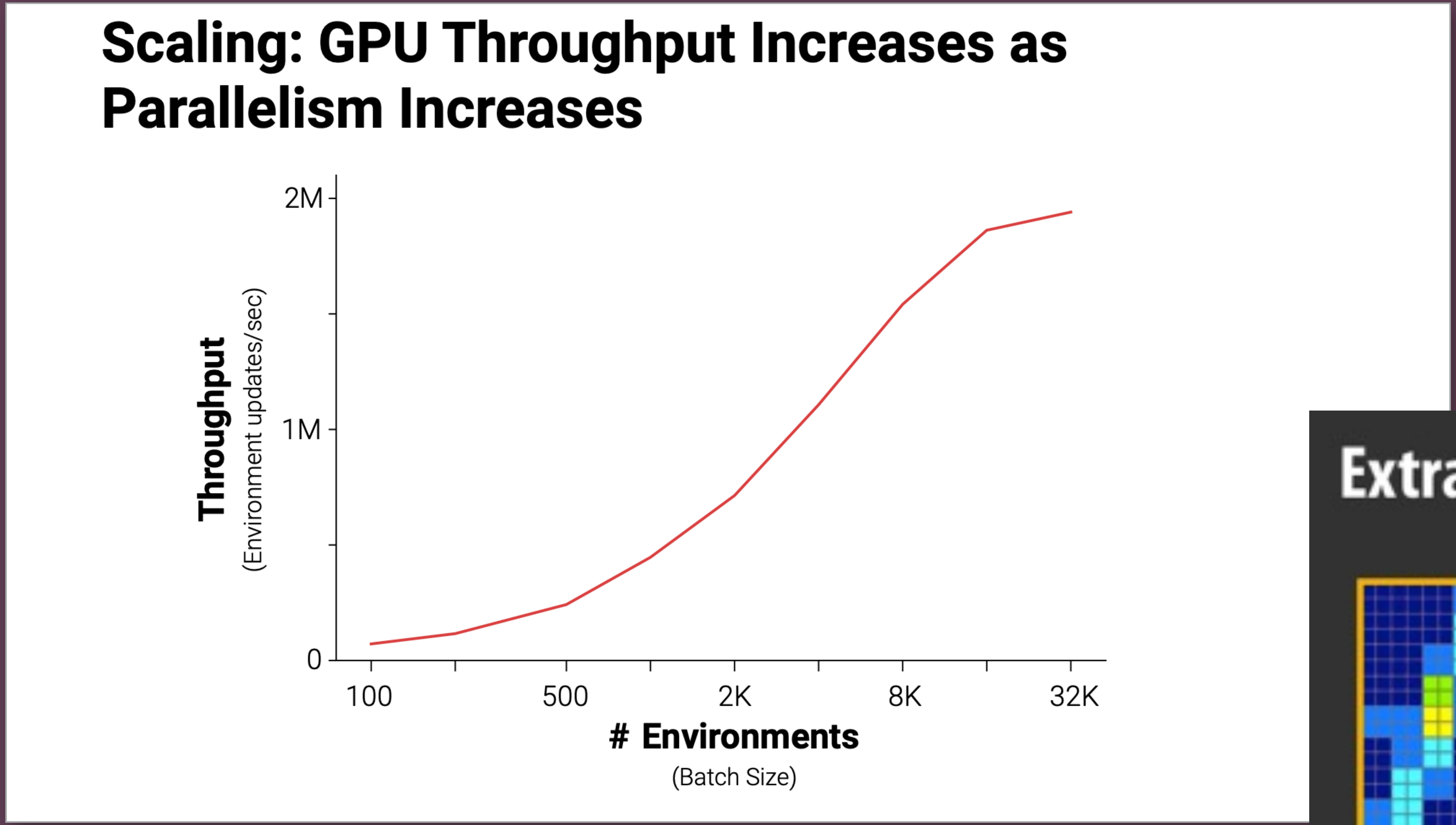


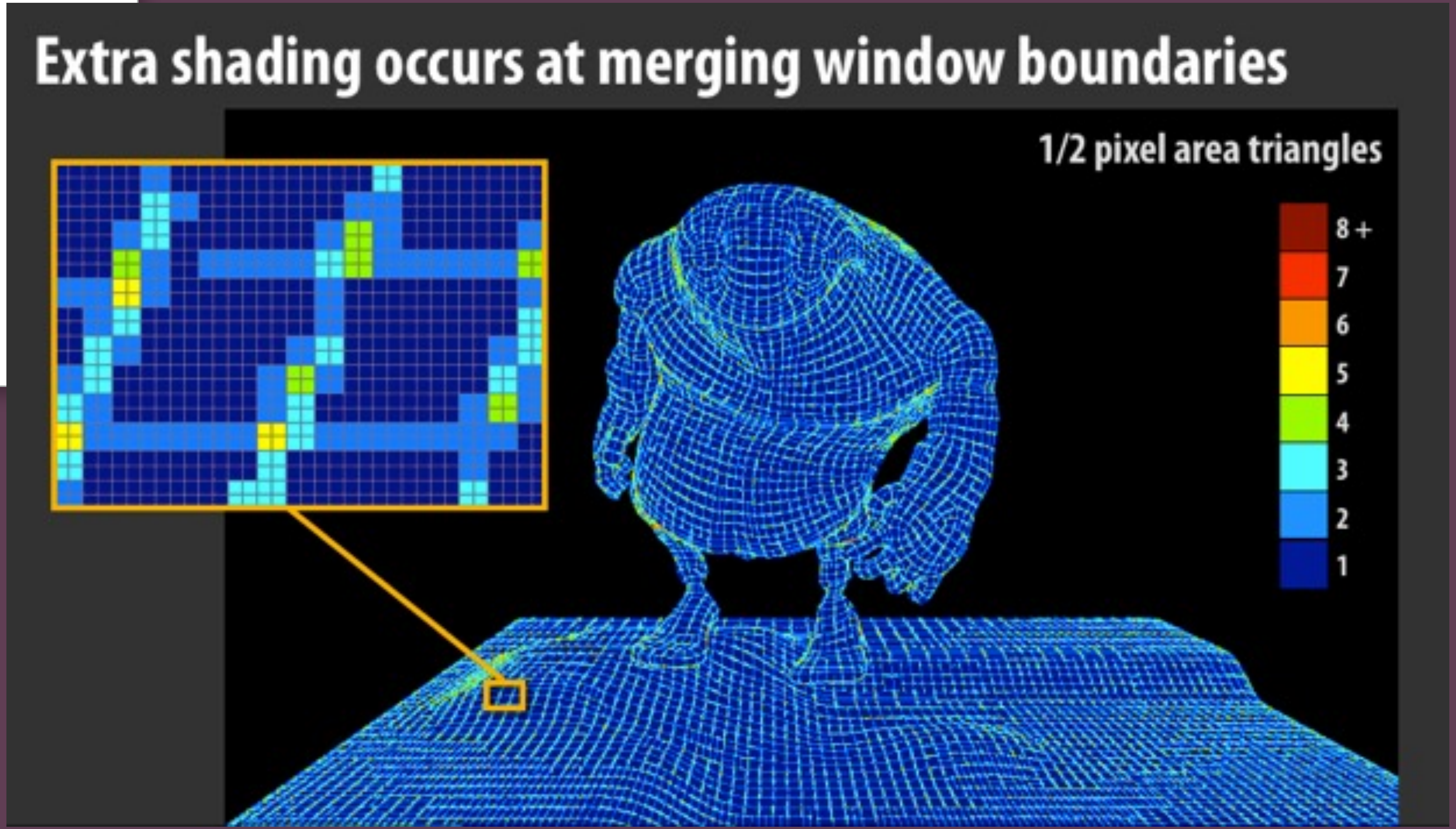
Example



The COMPONENTS renderer uses 2x less CPU time than BASELINE_VK
CPU Performance Comparison (single core)
Execution time (ms)
15
10
5
0
BOXES
30
20
10
0
BOXES1K
8
4
0
FACTORY1
FACTORY2
18
12
6
0
ROME
COMPONENT CPU Time
BASELINE_VK CPU Time

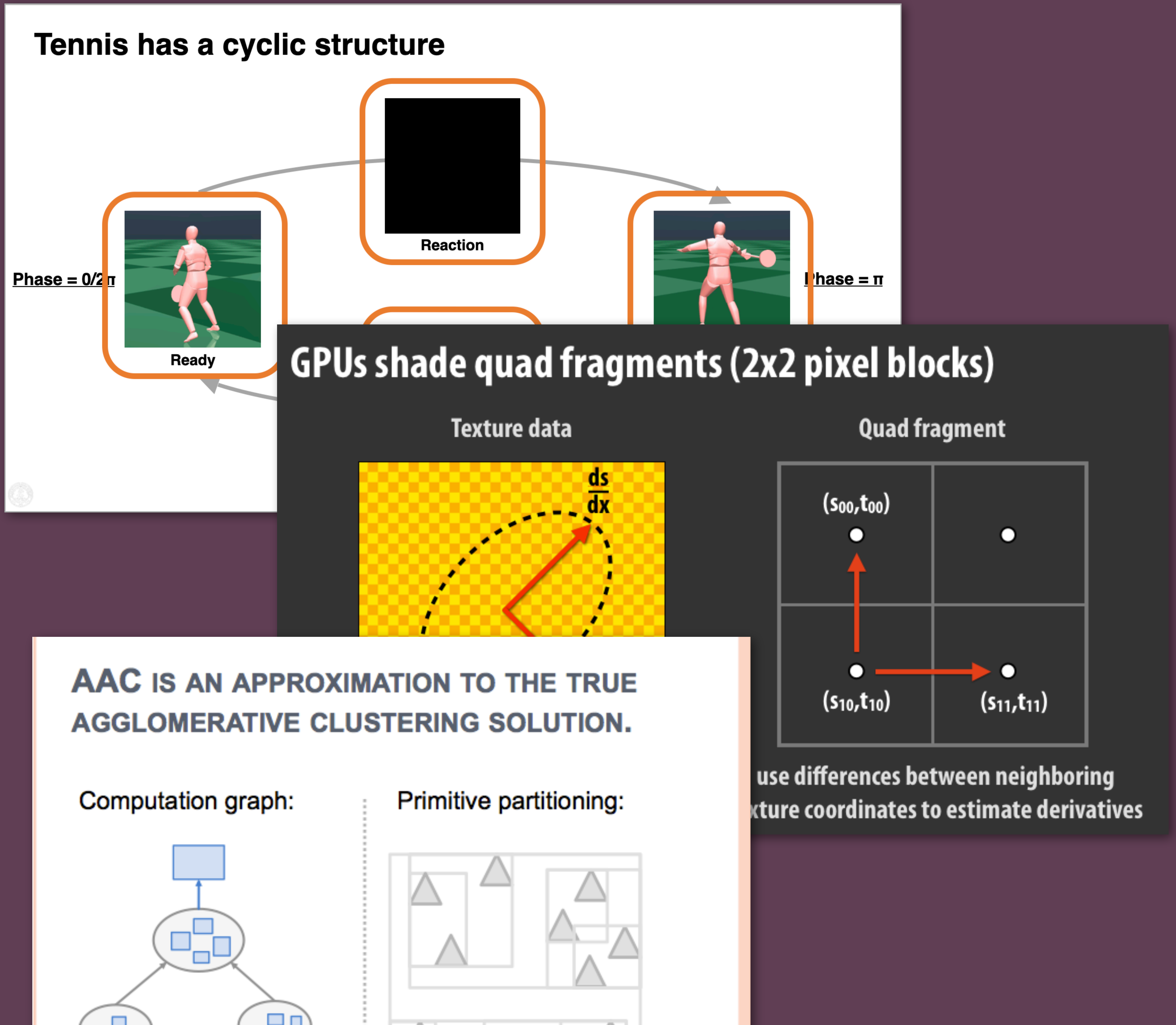
More examples



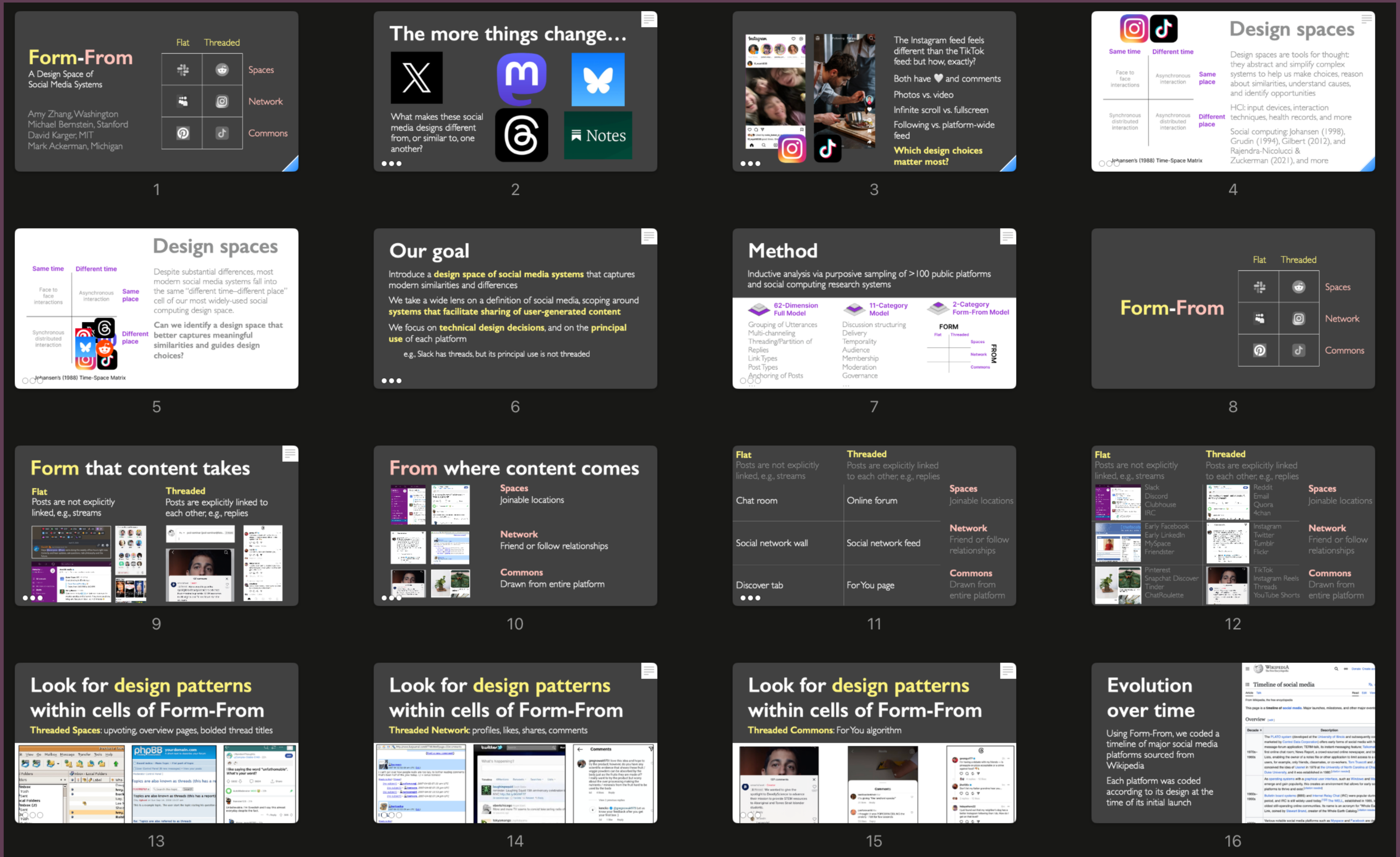


Slide titles matter!

If you read the titles of your talk all the way through, it should be a great summary of the talk.



Read your talk as thumbnails



Do they tell the story that you want to tell?

(Side benefit: assume people can't read anything too small in this view)

In sum:

1. Identify your one narrative goal
2. Every sentence matters
3. Inputs, outputs, constraints
4. Show, don't tell
5. Always explain any figure or graph
6. One point per slide

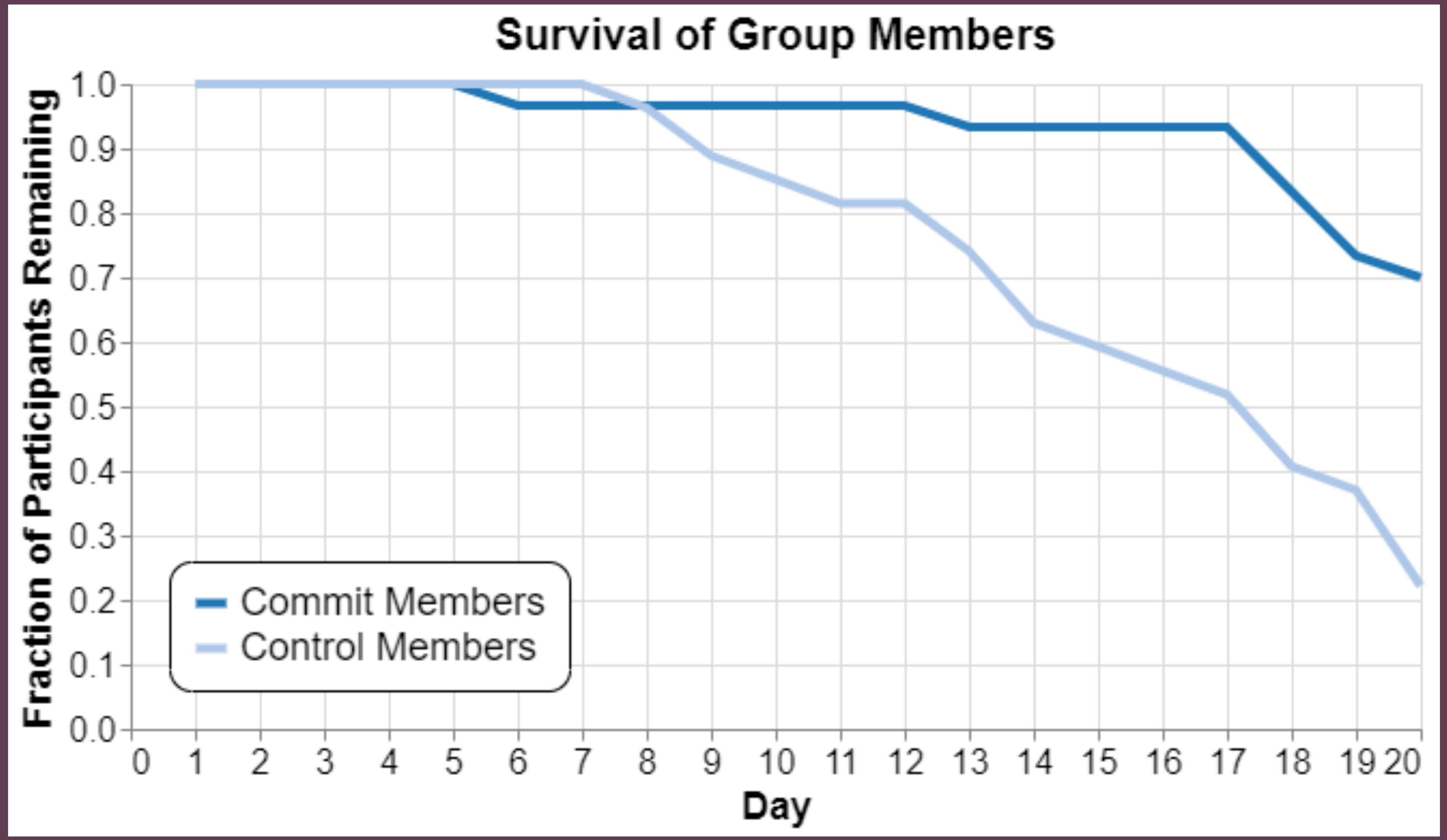
Figures

Each figure makes one point

The first sentence of the caption of the figure should describe what the person takes away from the figure, not what is in the figure

Not: "A survival plot comparing the Commit to the Control condition."

Instead: "Participants in the Commit condition were four times as likely to stay active until the end of the study."



Each figure makes one point

Ergo, ask yourself: **what is the point being made by this figure?** What justifies its existence?

"We made a simulated town"

Common figure types

Introductory / overview of what you did

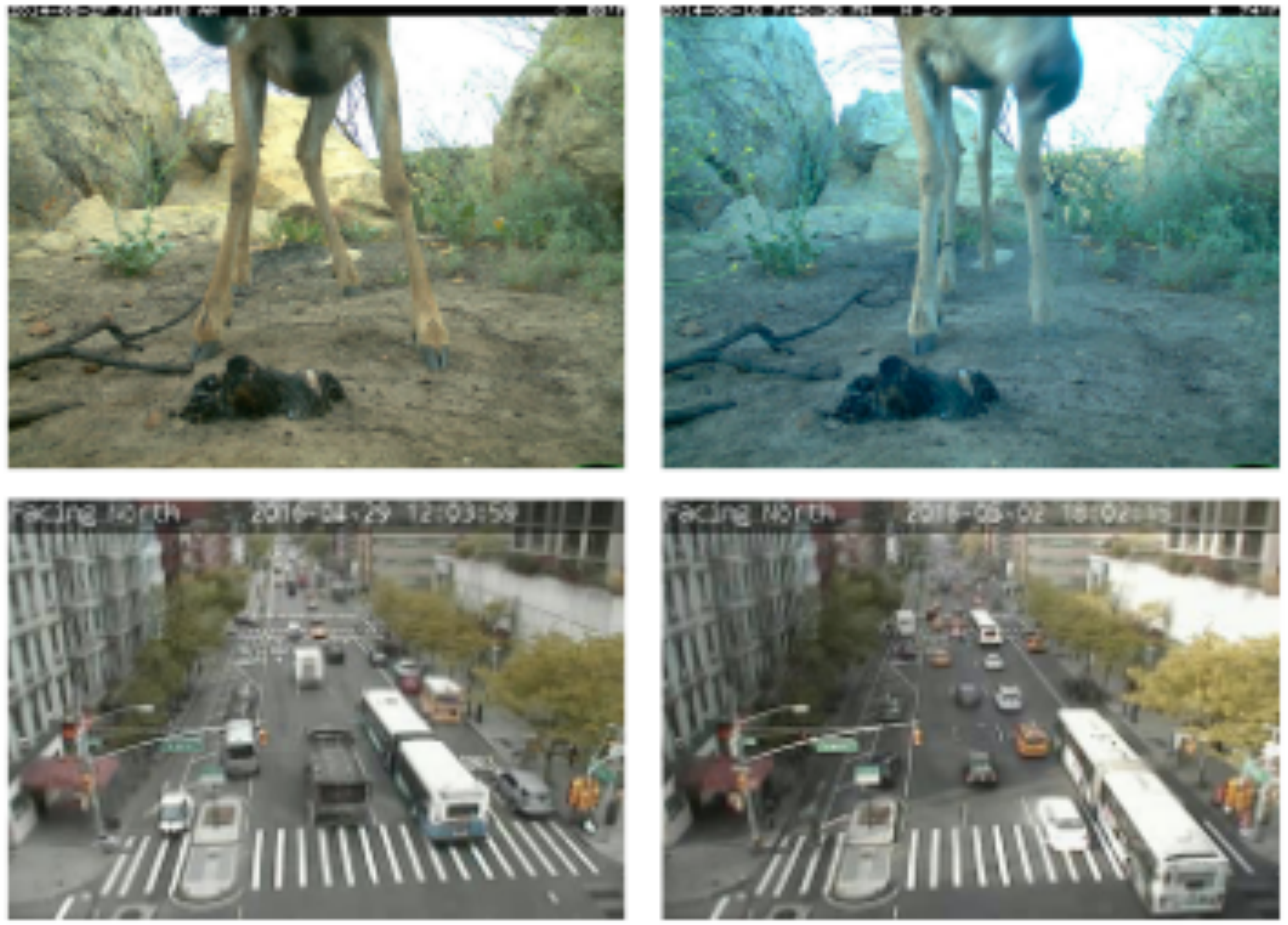
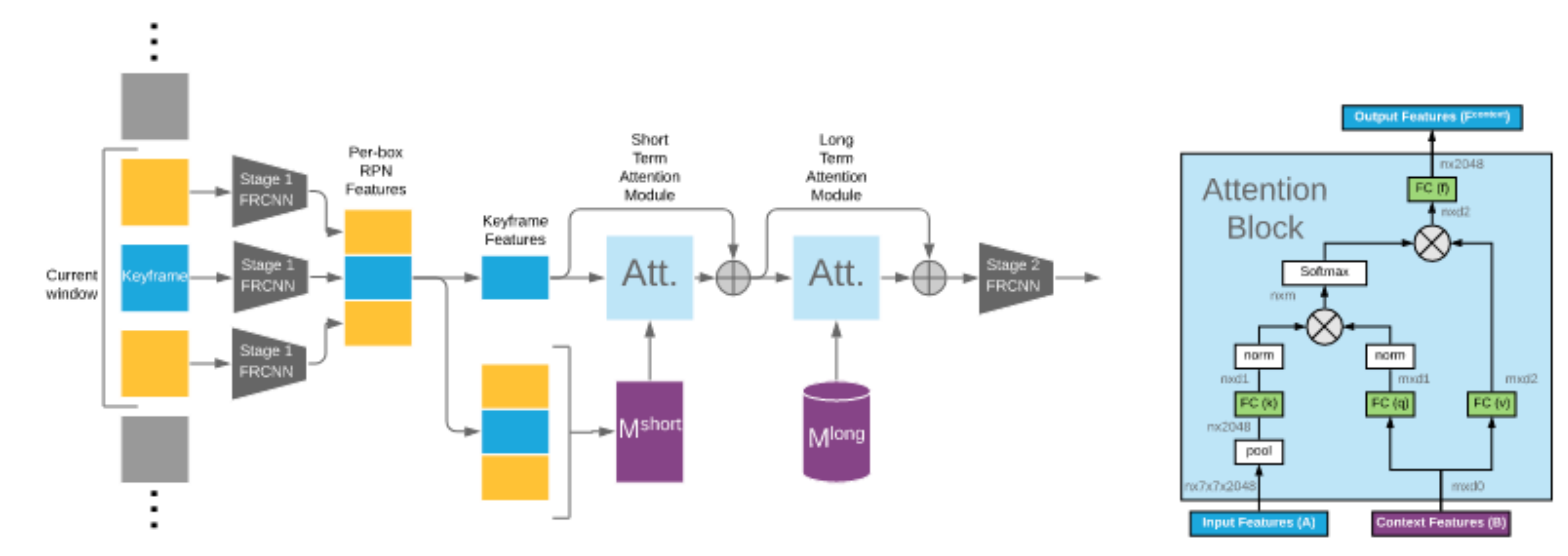


Figure 1: **Visual similarity over long time horizons.** In static cameras, there exists significantly more long term temporal consistency than in data from moving cameras. In each case above, the images were taken on separate days, yet look strikingly similar.

Beery et. al. 2020

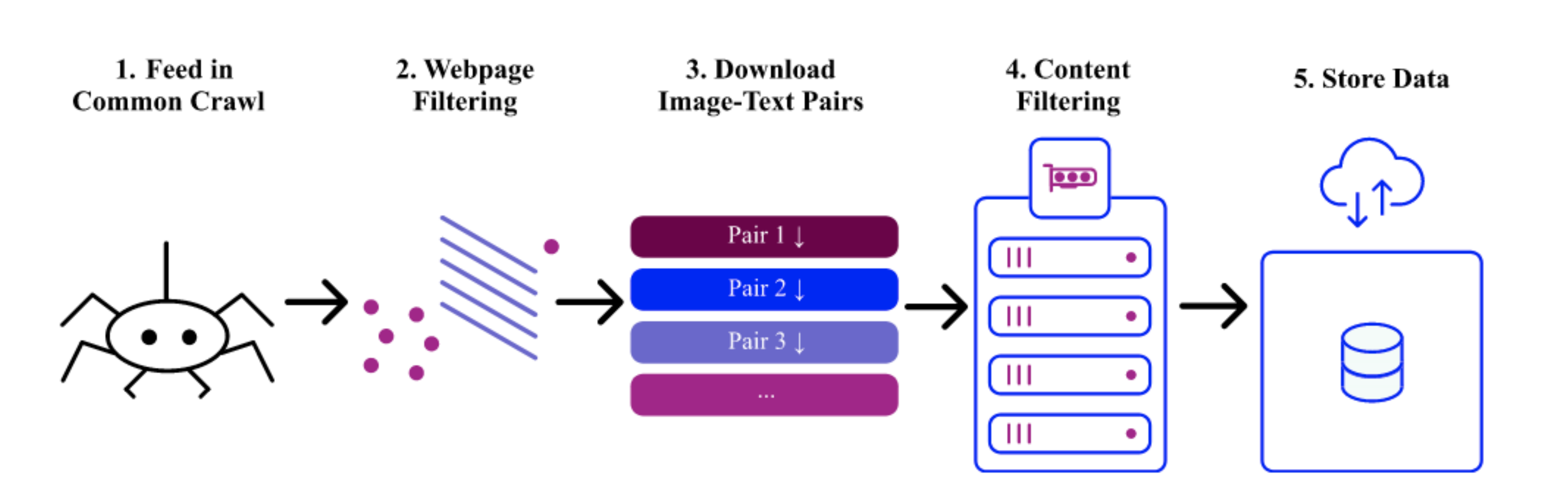
Common figure types

Framework / pipeline / model



(a) High-level architecture.

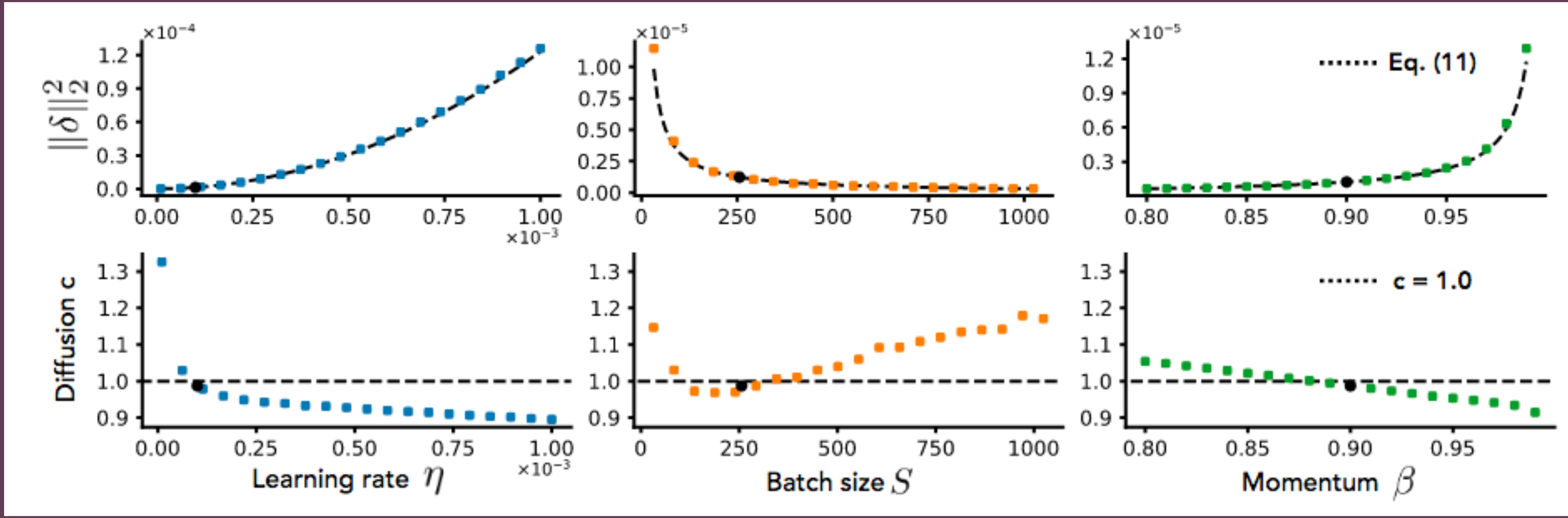
(b) Single attention block.



Schuhmann et. al. 2022

Common figure types

Experimental results



Kunin et. al. 2020

Common figure types

Applications or use in practice

Fig. 10. CoolMoves uses limited input of the user's hands and head to recover plausible legs motions as showcased in this scene of a user climbing a ladder.

Ahuja et. al. 2021

What makes a good figure?

A good figure does:

- Summarize one key takeaway at a time
- Make it immediate and easy to see your bit flip
- Remain true to the data

A good figure does **not**:

- Become a giant information dump of all your findings
- Bury your bit flip in a hard-to-see format
- Hide mediocre results with questionable display

Tuning your tables

Tables are the neutron stars of your paper: incredibly information-dense

Ideal for summarizing 3+ treatments

Centering your bit flip:

- Can **bold** best model per-treatment
- Clever alternatives: arrows, colored deltas

Use colors, lines, bolding to distinguish treatments & improve readability

Model	Pre-training	INet	INet-v2	INet-R	INet-S	ObjNet
B/32	CLIP WIT	63.3	56.0	69.4	42.3	44.2
	LAION-400M	$62.9^{-0.4}$	$55.1^{-0.9}$	$73.4^{+4.0}$	$49.4^{+7.1}$	$43.9^{-0.3}$
	LAION-2B-en	$65.7^{+2.4}$	$57.4^{+1.4}$	$75.9^{+6.5}$	$52.9^{+10.6}$	$48.7^{+4.5}$
B/16	CLIP WIT	68.3	61.9	77.7	48.2	55.3
	LAION-400M	$67.0^{-1.3}$	$59.6^{-2.3}$	$77.9^{+0.2}$	$52.4^{+4.2}$	$51.5^{-3.8}$
B/16+	LAION-400M	69.2	61.5	80.5	54.4	53.9
L/14	CLIP WIT	75.6	69.8	87.9	59.6	69.0
	LAION-400M	$72.8^{-2.8}$	$65.4^{-4.4}$	$84.7^{-3.2}$	59.6	$59.9^{-9.1}$
	LAION-2B-en	$75.2^{-0.3}$	$67.7^{-2.0}$	$87.4^{-0.5}$	$63.3^{+3.7}$	$65.5^{-3.6}$

Schuhmann et. al. 2022

Read your paper as figures

Just like the talk thumbnail view, you should be able to understand the main narrative of the paper by just skimming the figures and tables

Communicating Research