Handling Scene Detection of Lecture Videos With Annotations

Ashley Li, Charitha Nannapaneni, Enya Chen

Research Mentor: Professor Lawrence Angrave

Motivation

Currently, there are limited alternatives to video format lectures—thus many educational materials are not universally accessible.

In response, we ask this: given a video lecture, how do we extract its key frames and create a concise PDF?

ClassTranscribe

media space Illinois

YouTube

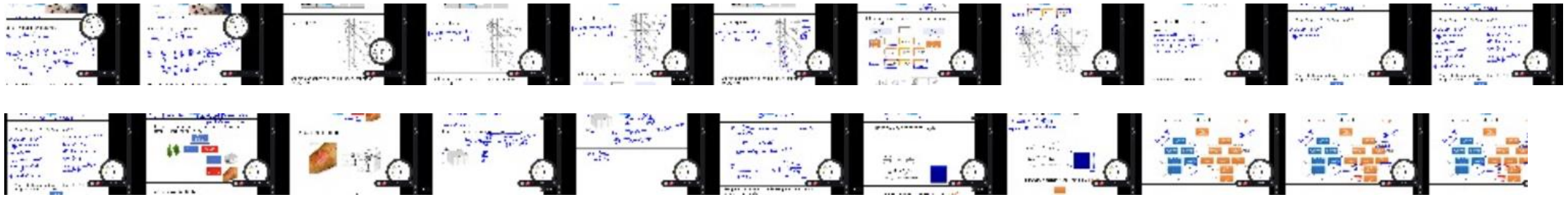
Khan Academy

Problem context

Scene detection attempts to find semantic boundaries between segments of a video

However, existing research (Li, ASEE 2022) does not apply well to engineering lecture videos

How would you segment this video into chapters?

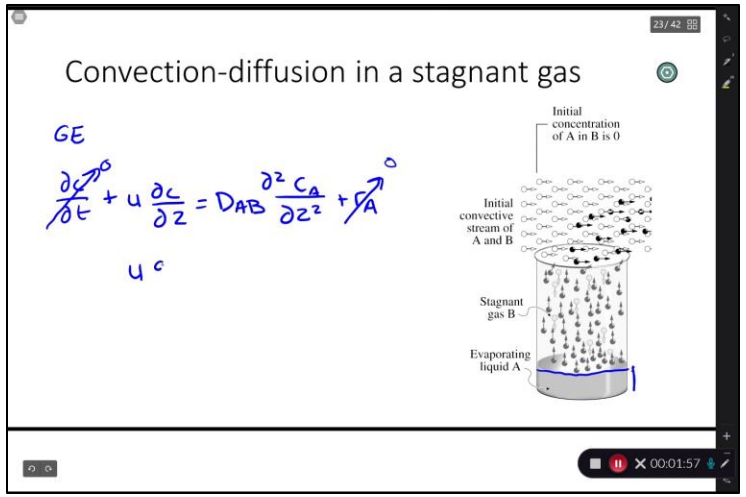


test1_bio360_2_12_2022.mp4

How would you segment this?

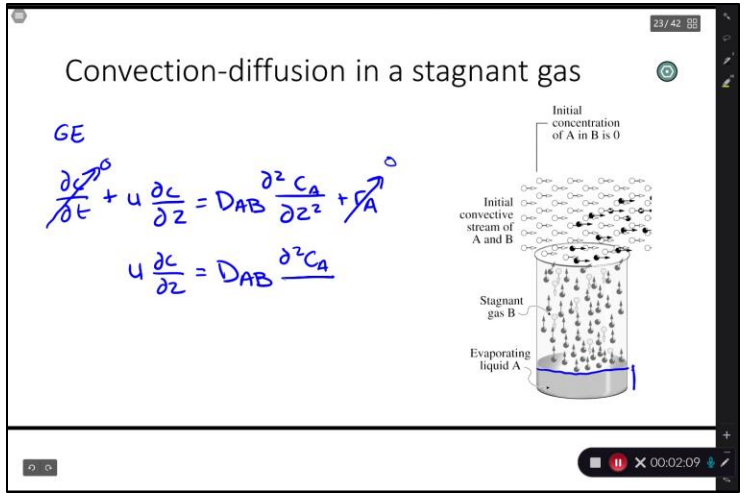
Slides from a BIOE360 on February 12th, 2022

1



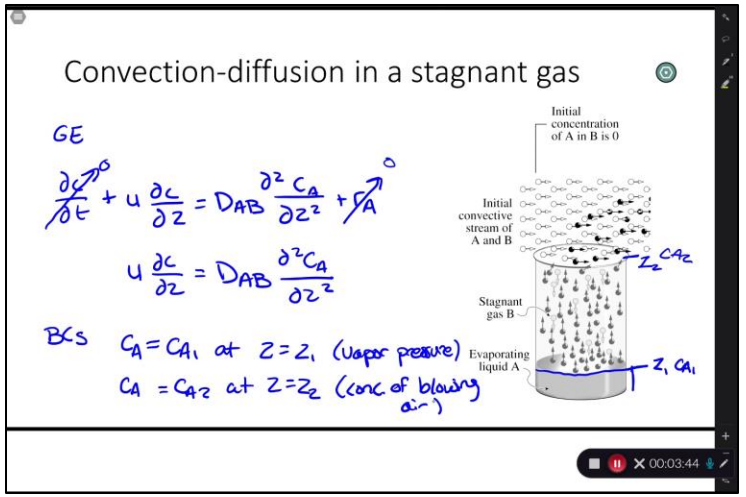
1:58

2



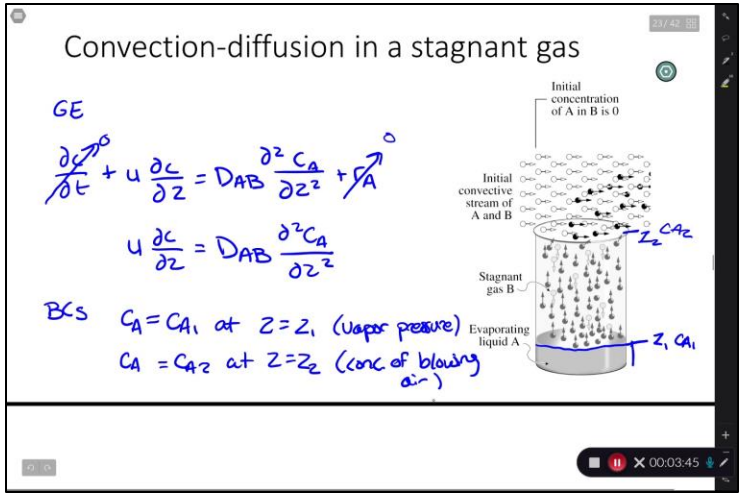
2:10

3



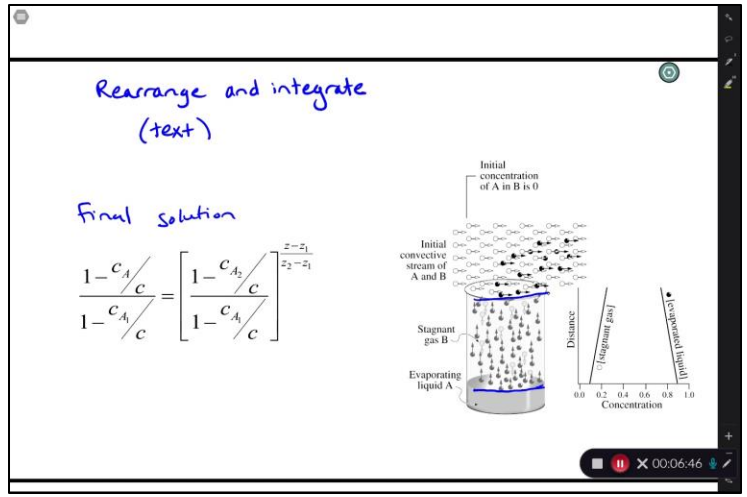
3:45

4



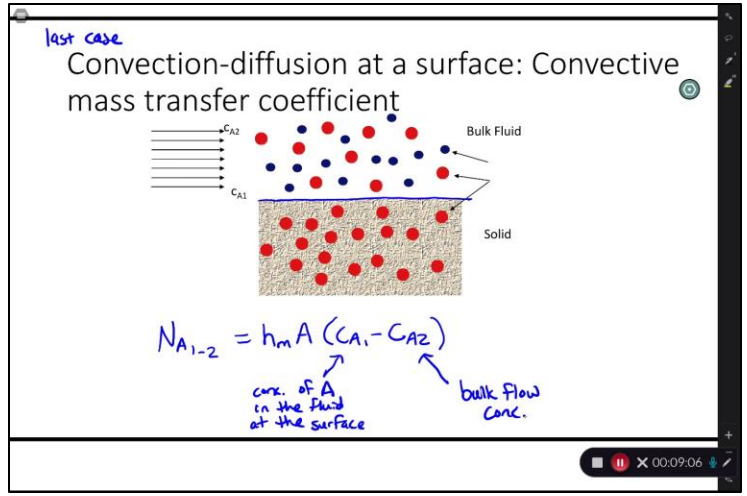
3:46

5



6:46

6



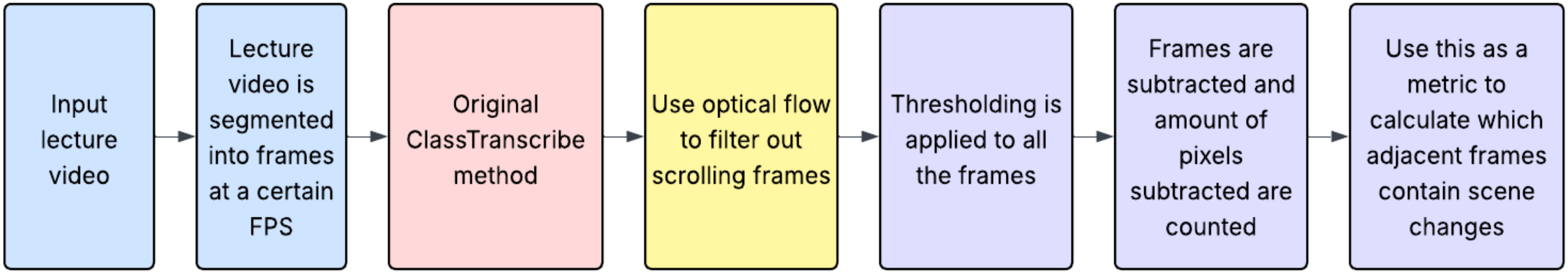
9:06

Our Approach

Segmenting video into proper format

Handling Scrolling

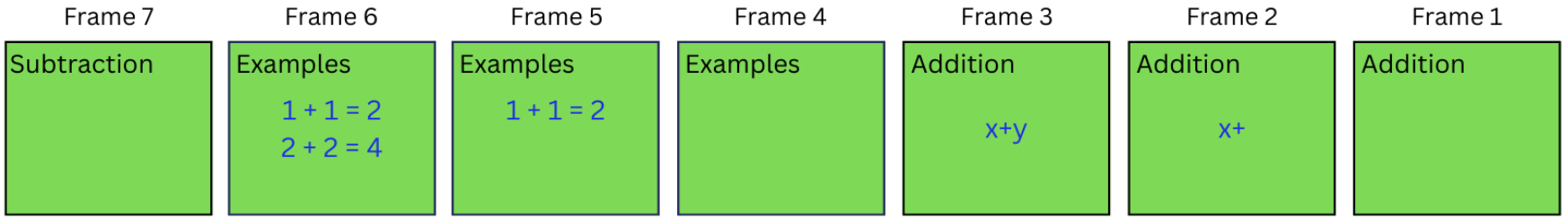
Handling Annotations



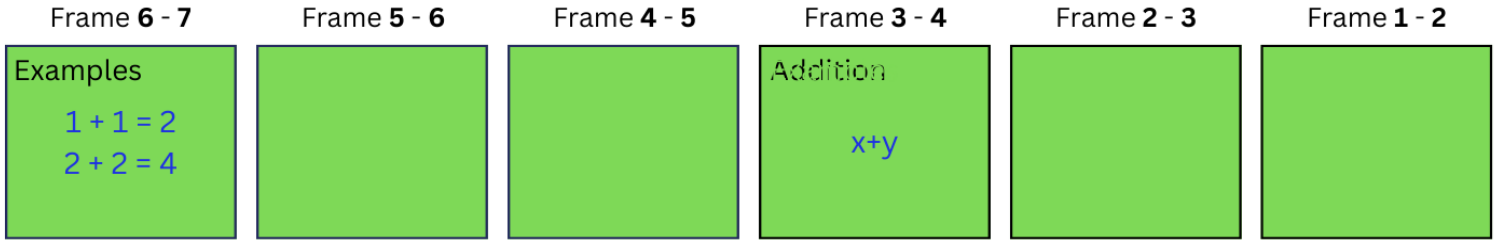
Handing annotations and scrolling

through a process of masking and subtracting the frames.

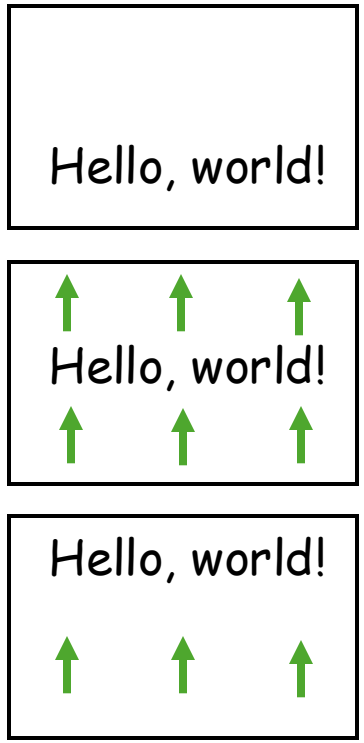
Masking + Subtraction (backwards)



Subtraction results



Scrolling



Purpose of masking

If remaining pixels (non-green pixels here) are related to the information on the screen, not masking leaves misleading pixels

Slide 38 – 39 =

38

Sequential Pattern Mining

- What kind of patterns are sequential?
- Sequential – The order really matters. You can not swap two items in a sequence and have the same sequence.
- Example: The English language is sequential : Subject -> Verb -> Object.
- Other points:
  - For Sequential Pattern Mining, the time at which the items occur is **not** considered.
  - Time Series Analysis does take into account the time at which an item occurred.

38

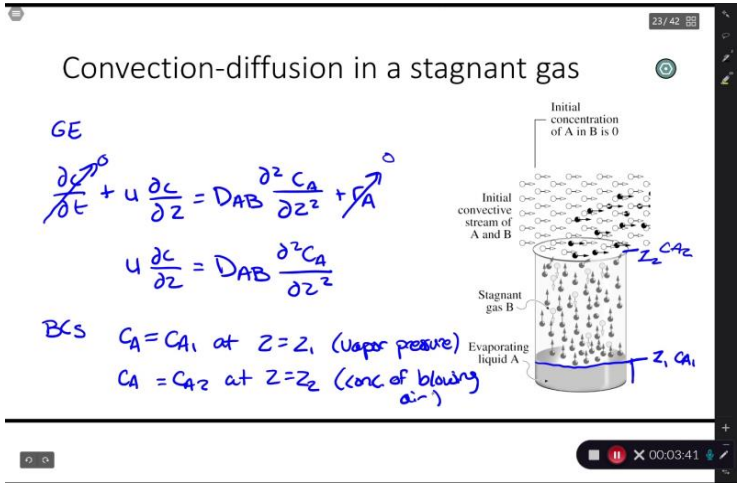
39

Sequential Pattern Examples

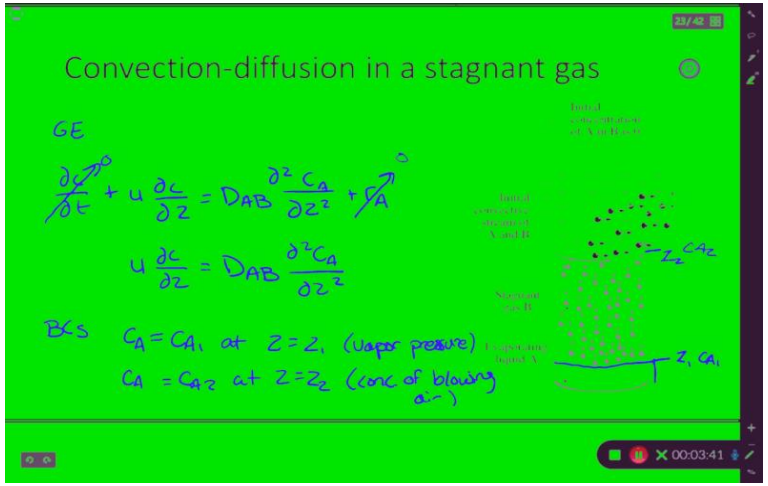
- Application of Sequential pattern Mining
  - **Customer shopping** → Purchase a laptop first, then a digital camera, and then a smartphone.
  - **Medical treatments** → Go to the doctor, get drugs, doctor monitors progress, doctor reacts accordingly -> more/less drugs
  - **Natural disasters** -> Before the disaster, during the disaster, after the disaster.
  - **Scientific Experiments** → Step 1, Step 2, Step 3.
  - **Stocks Markets** → Stocks go up and down together.
  - **Biological sequences, DNA /Protein**→ If you change the order of proteins, it is a different gene.

39

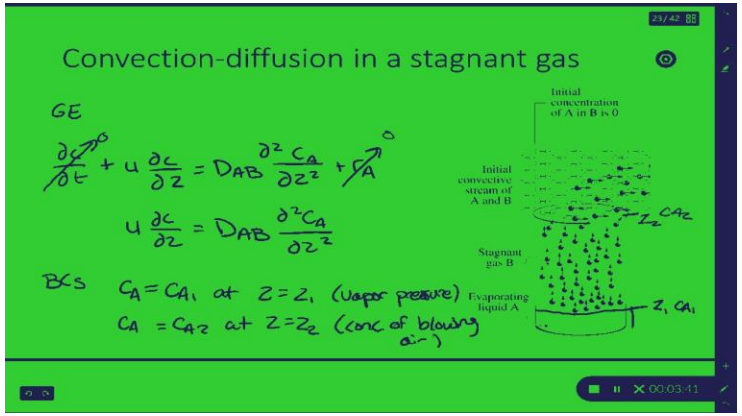
Our Masking Methods



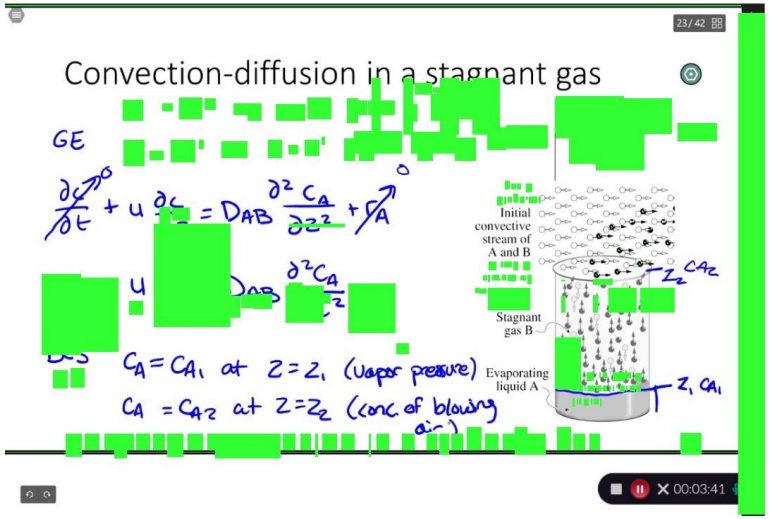
Selected Represented Frame



Thresholding method, $a_{x,y} < t$



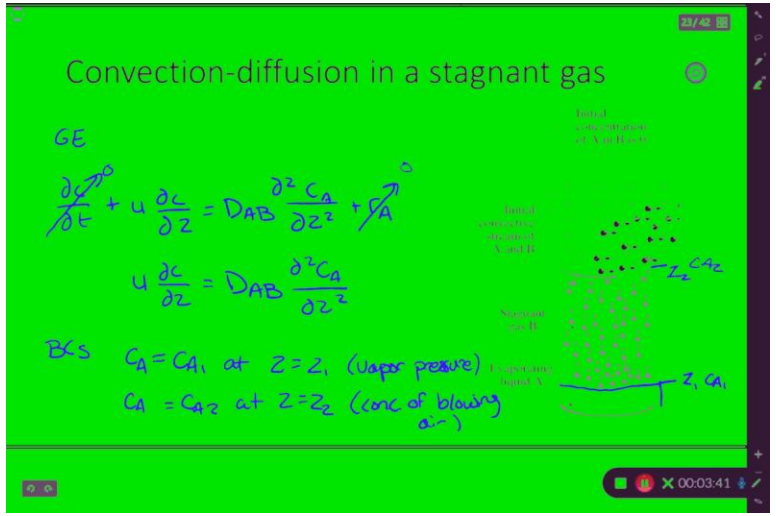
Brute force method, $p_{x,y} - b_{x,y} < t$

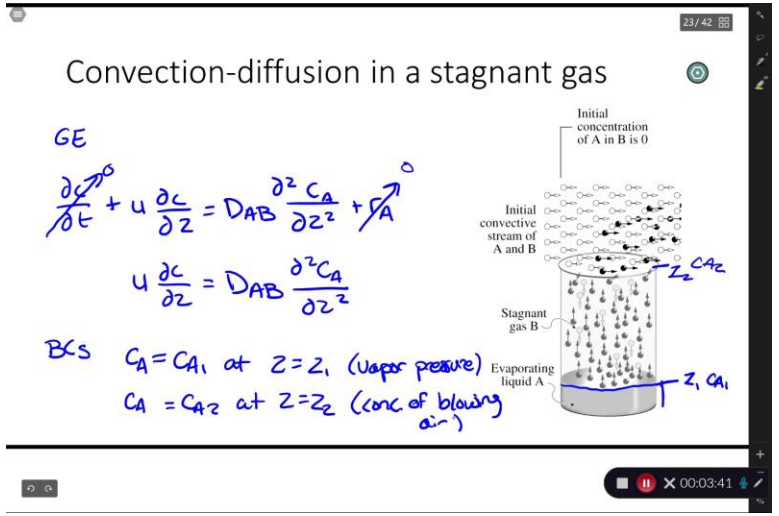


Spectral Clustering method, $O(mnk)$

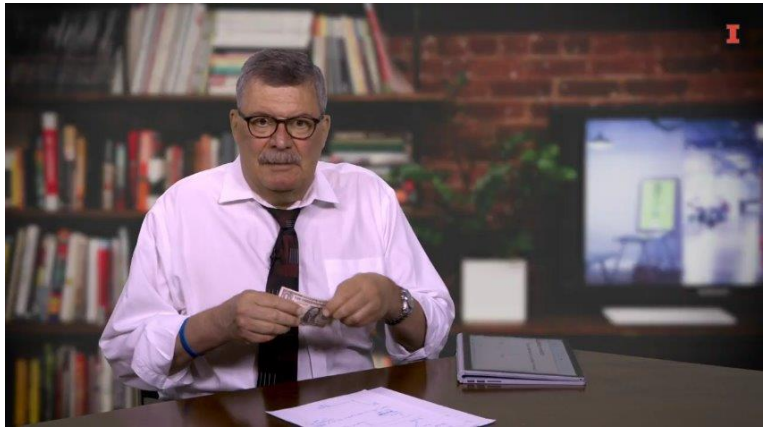
Thresholding

$$a_{x,y} < t$$





Thresholding on a lecture frame



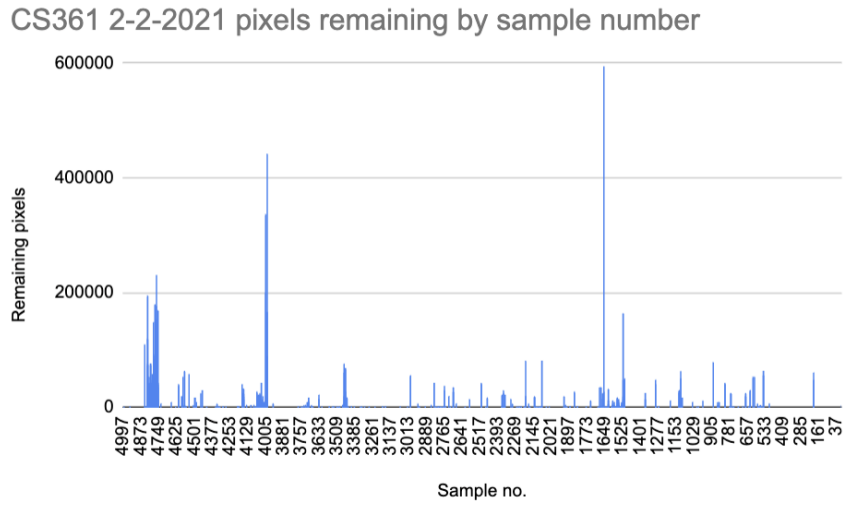
Thresholding on a non-lecture formatted frame

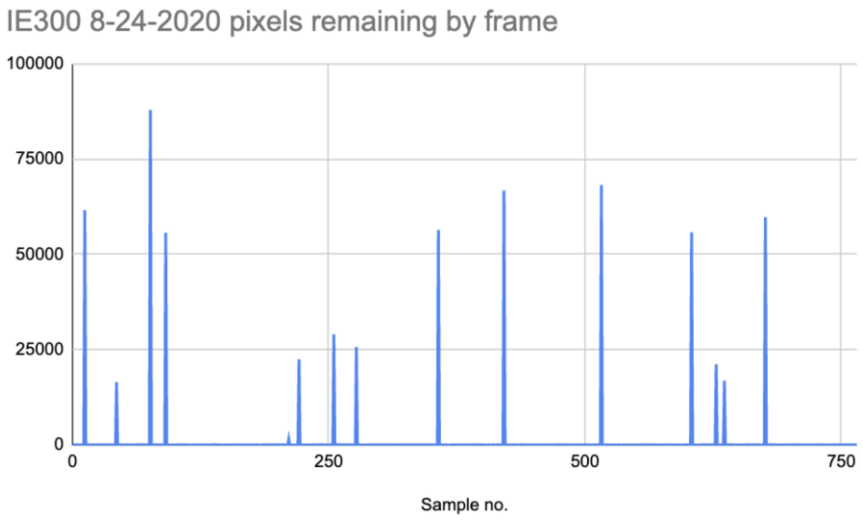
Mask comparison

Currently, we're using a brute force subtraction method.

Next, we are improving our method to pick out the significant outliers from our data.

Current output





Future Improvements

Broader background color support

Integration with scrolling methods + ClassTranscribe

Thank you for listening!

Questions?

Ashley Li
(ashleyl7@Illinois.edu)

Enya Chen
(enya2@Illinois.edu)

Charitha Nannapaneni
(cn29@Illinois.edu)

Subtraction in depth

There are more or less four cases:

1. The slide has not changed
2. The slide is being annotated
3. The slide has changed
4. (The slide has changed to blank)

start end

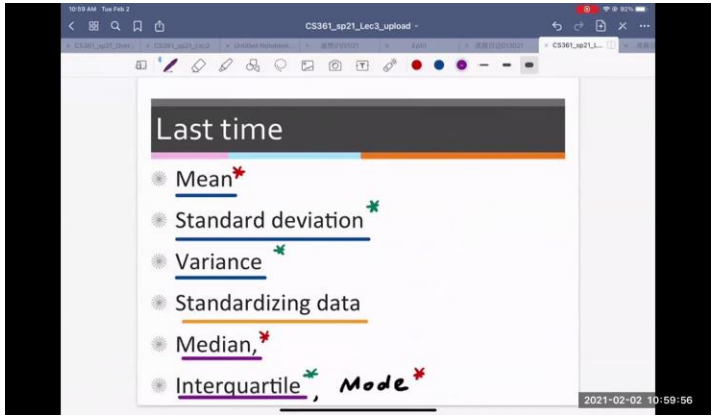
Case 1: The slide has not changed

Frames

Masks

Subtraction

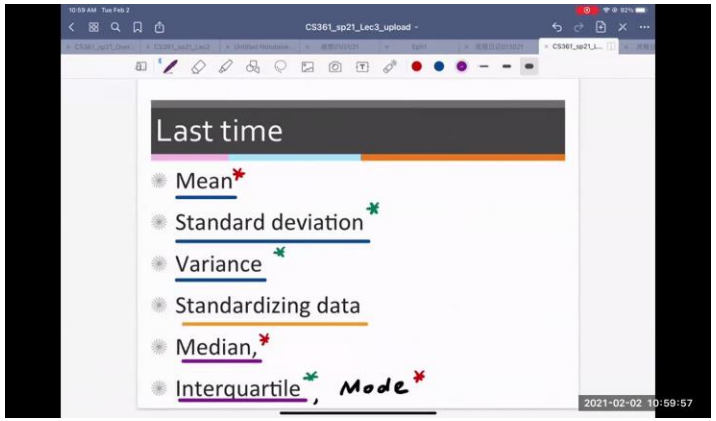
218



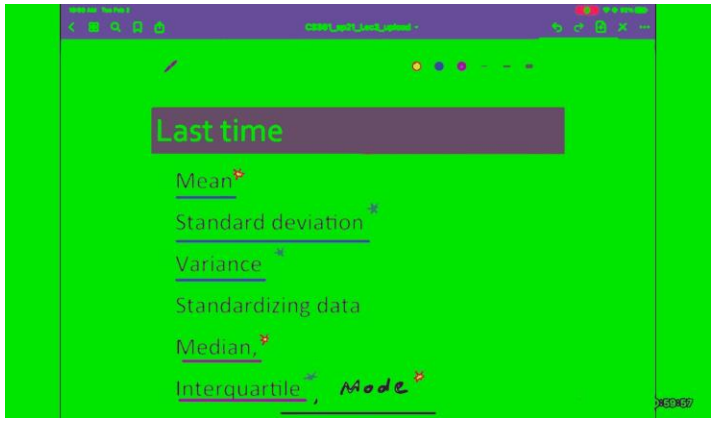
→

−

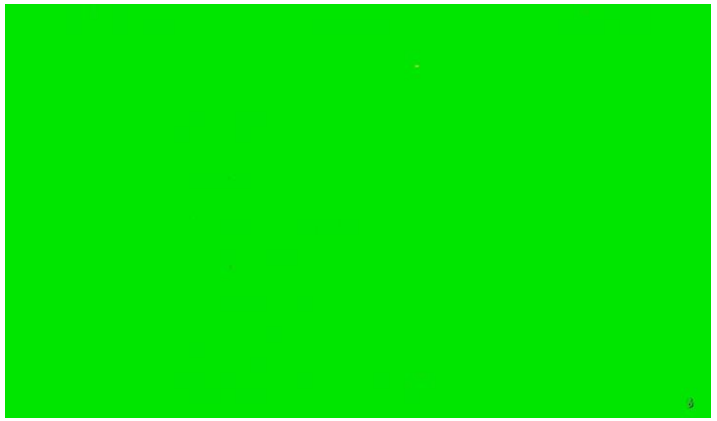
219



→



=



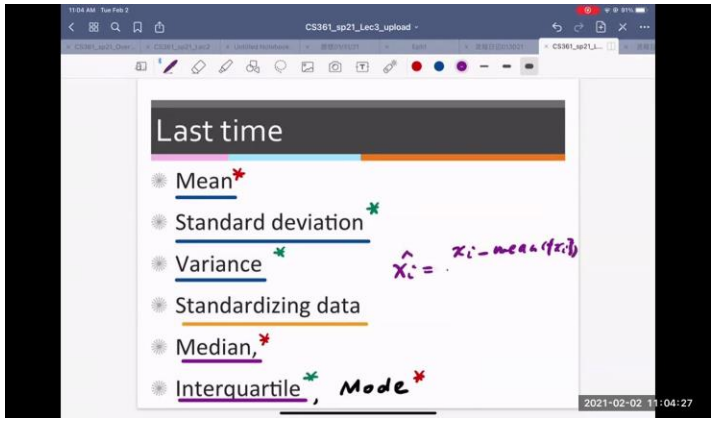
Case 2: The slide is being annotated

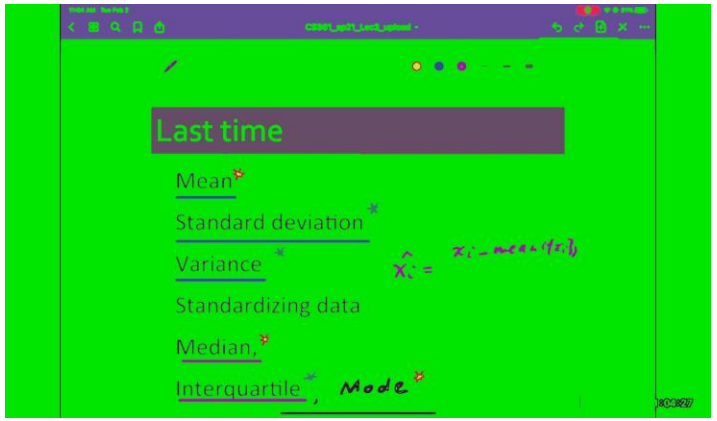
Frames

Masks

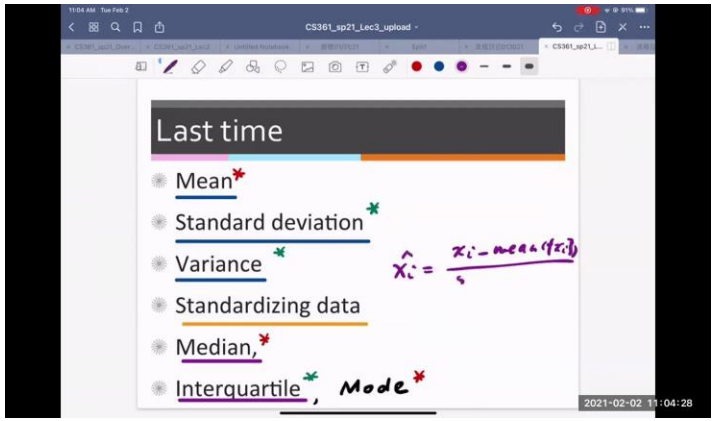
Subtraction

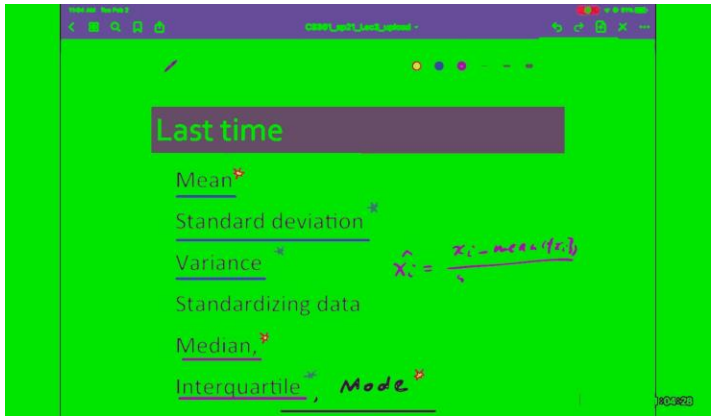
489

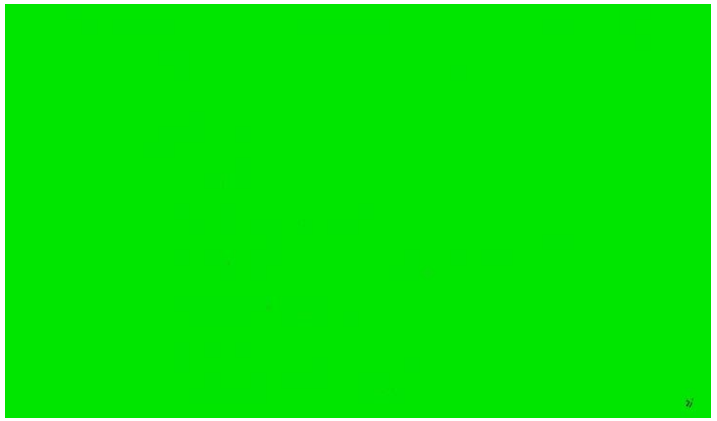




490







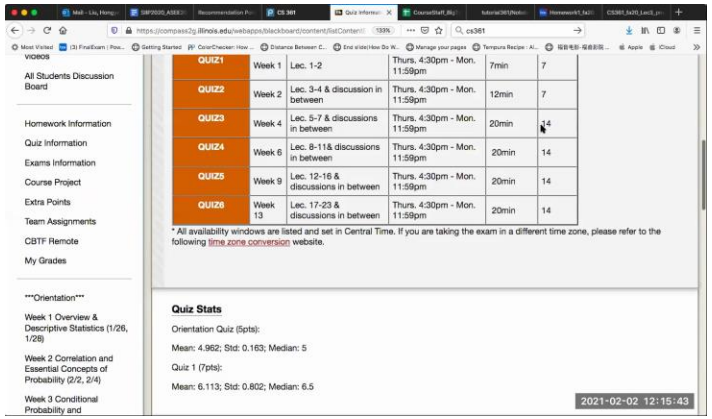
Case 3: The slide has changed

Frames

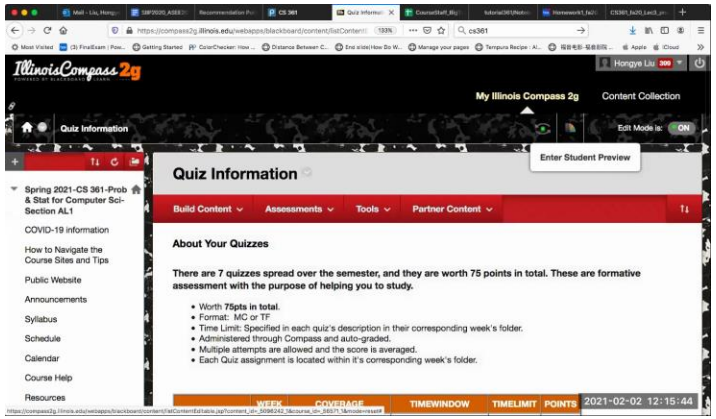
Masks

Subtraction

4765



4766



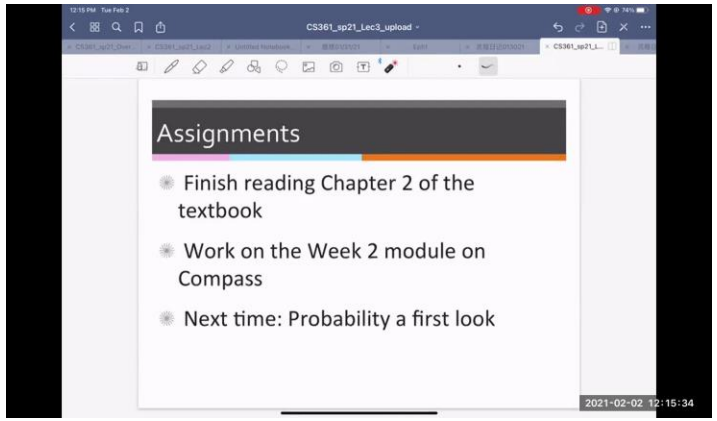
(Case 4: The slide has changed to blank)

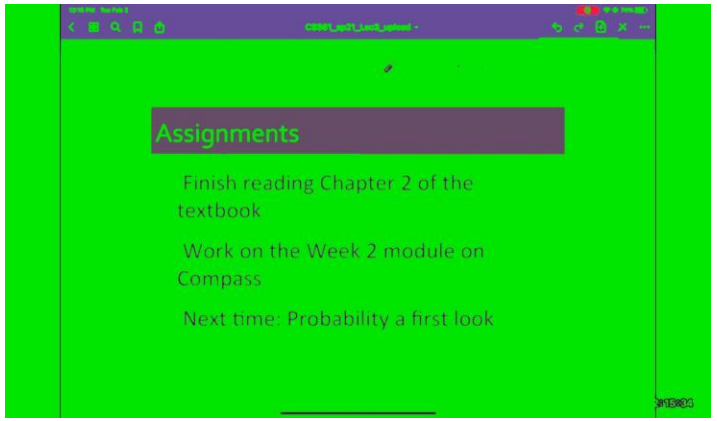
Frames

Masks

Subtraction

4756





4757