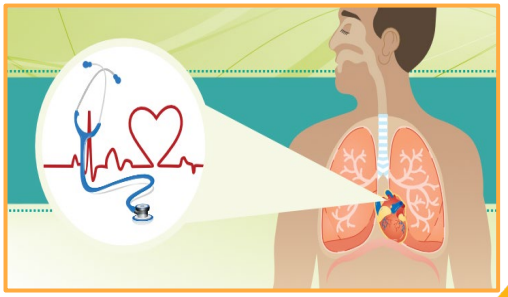
Breathing Rate from Video

Matthew Makila, Zahid Hasan, Azim Khan, Dr. Nirmalya Roy

Problem Statement

- **Respiratory Rate:** Number of breaths (respiration activities) for one minute.
  - Indicator of potential respiratory dysfunction, physiology of acutely-ill patients and early warning
  - Actions/Emotions (crying, sleeping, agitation, exercise, age) have large influence on respiratory rate.
  - Facilitate identification of changes in physiology along with other vital signs.
- ***Physics:*** Some body-parts motion due to breathing
  - *Our abdomen, shoulder, head moves with our breathing as the diaphragm moves.*
  - Video contains Human breathing rate induced motion in its spatio-temporal frames
- Idea and Task from *VIDEO* Signal
  - **Detect the corresponding body-parts**
  - **Track their movements** from video



Problem Breakdown: Complexities vs. Simplicities

- **Noise**
  - ***Talking***, ***moving***, *types of clothing, facial expressions*
- **Video issues**
  - *Complex* ***backgrounds*** *(other people or objects), misalignment of body within frames*
- **Starting simple …**
  - **Manually** define regions to measure Respiratory Rate (RR) from the chest, abdomen, and shoulder
  - **The idea**: observe how spatial signal changes over time to predict RR
  - **Assumptions:**
    - We can simplify the problem for now by ***reducing complexities and focusing on calculations***
    - We test a person in several static positions for several minutes at a time
    - We do not have them talk
    - We keep the background for testing minimal

Challenges in Related Work

- Lack of dataset and no variability in subject positions
  - There are few works that utilize large datasets with a **diverse subject pool** or **variety of testing positions**
- Video processing computation
  - Previous methods utilize computationally-heavy ML networks to find RR
    - AHN, CNN, etc (take *lots* of training and resources)

Building a Dataset (Expectations)

- We need sample breathing data
- Build a ***holistic*** and ***representative*** dataset
  - We want to include data for varying circumstances that exist within our major limitations
    - Samples taken of different individuals from different backgrounds, genders, etc
    - Samples taken before and after eating
    - Samples taken before and after exercising
    - Some samples taken with known issues to *outline boundaries & limitations*

Data Collection

- Details
  - 3-5 min length
  - Volunteers are positioned either **lying down** or **sitting (side and front profiles)**
  - **Front-facing camera** of phone
- Methods
  - (1) Setup **ground truth** (GT) measuring device
  - (2) Setup camera and tripod to fit body in frame
  - (3) Start video and then afterwards start GT device (***on camera!***)
  - (4) Record video entire video while subject remains still
  - (5) Disable GT device, then video

Fig. 1. Ground truth respiration monitoring device

Methods (1)

- Installed Anaconda + packages for Pytorch, OpenCV, and their related dependencies
- Processing raw sample data provided into a video displaying different edge detection methods
  - The best: **Canny edge detection**

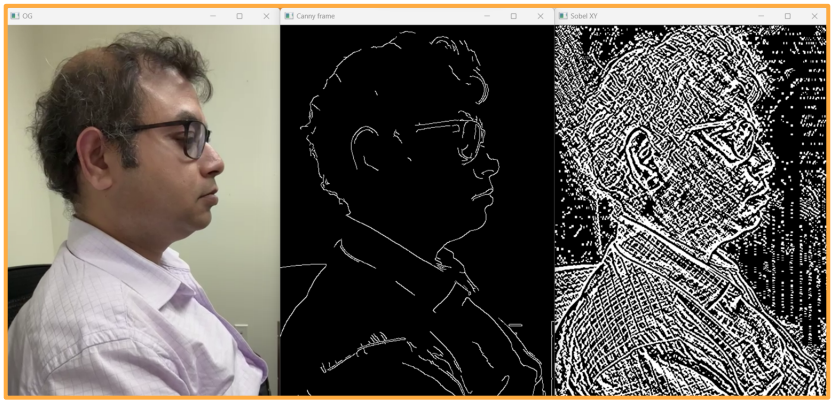


Fig. 2. Sample raw video processed with edge detection(s).
From left to right, original video, Canny filter, and Sobel filter.

Methods (2)

- Working with ***Anaconda + Spyder*** to analyze sample video data by **extracting specific regions of interest (ROIs)** from *edge-detected* frames to observe RR

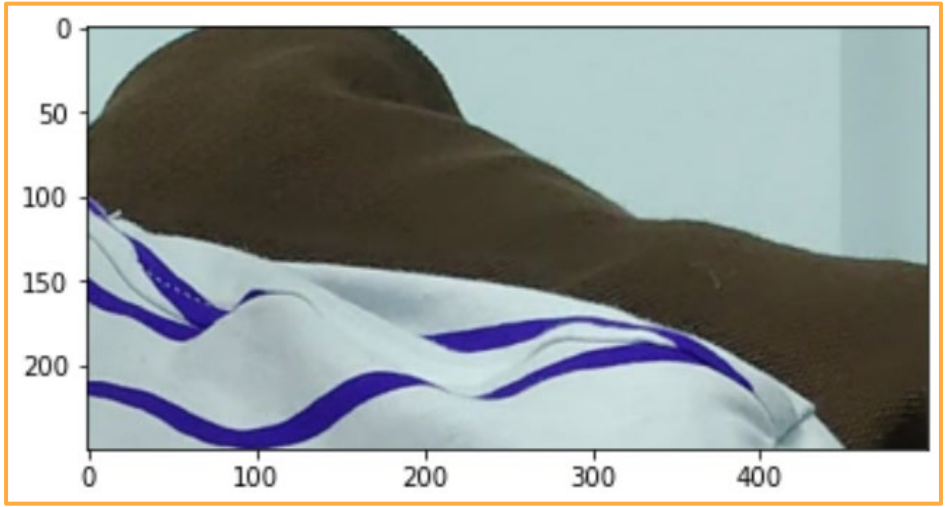


Fig. 3. Manually selected ROI from raw video.

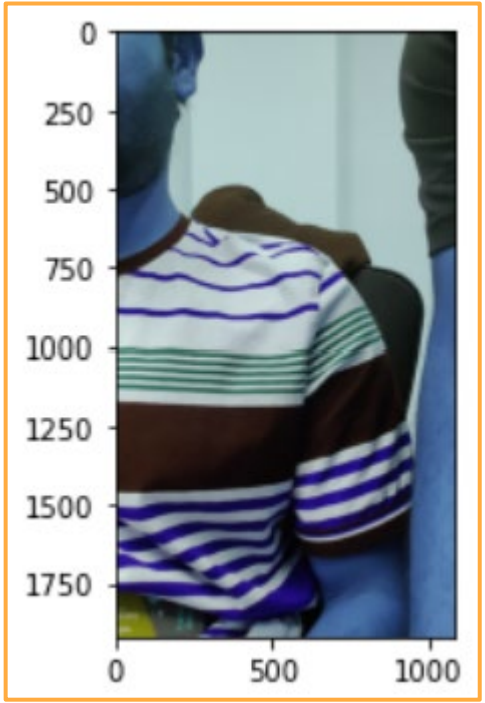


Fig. 4. Raw frame from sample video.

Methods (3): An Energy-based Approach

- Row, column, and full sum techniques
  - We want to check how **point concentration** shifts within each frame
  - And where their highest concentrations are per row/col
- Requires **filtering** in frequency domain
  - RR frequency is limited within a range (typically **low**)
- Filter out the noise range
  - Low-pass filtering

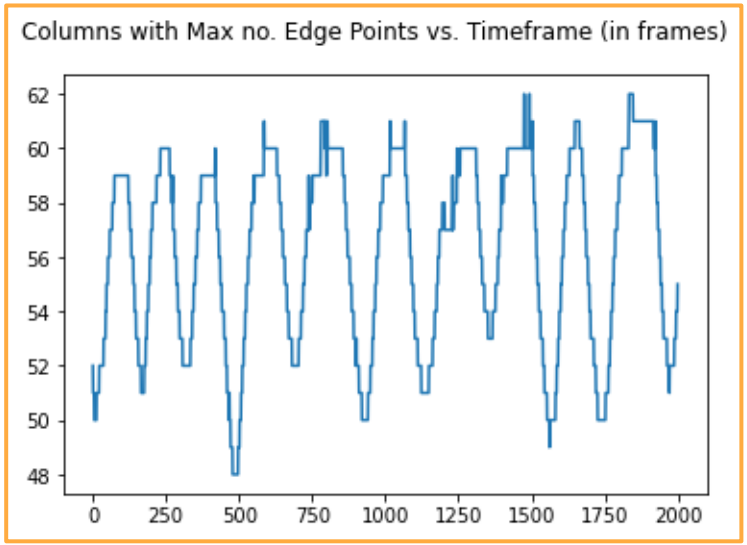


Fig. 5. A sample video's raw noisy signal plotted as the column numbers featuring the maximum number of edge points against the timeframe, specified by the number of frames

Methods (4): Cleaning Data

- To better examine our video data, we can look into balancing different ***filtering***
  - **Low-pass**
    - Extract *low-frequency* components of signal, emphasize slower and larger change but less sporadic noise
      - **Designed Butterworth filter**
  - **FFT**
    - Breaks apart **various components** of signals and helps **identify patterns in temporal changes**
    - Allows **automatic choice** between row, column, or full sum technique for samples

Data Analysis (Example)

- Filtering + better ROI extraction has made for the production of *cleaner + more meaningful data*
  - A rough sketch of the **subject movement** shown in fig. 1.
  - A **distinct pattern** emerges with movement corroborated visually with video data (can *eventually* get us to the RR)

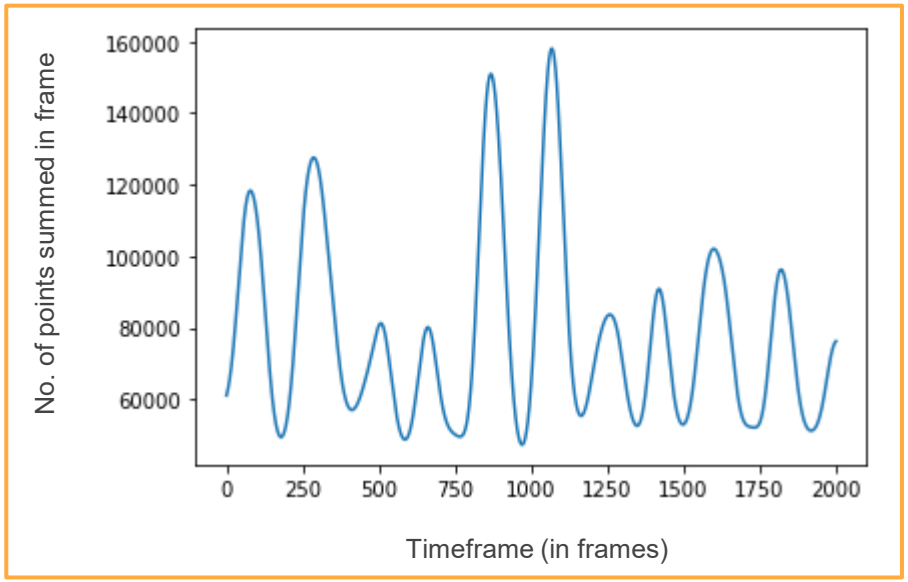


Fig. 6. Plot of sample no. 20 using the full sum technique. The Butterworth low-pass filter was used to produce a clean plot.

Future Endeavors

- Continuing to work in the lab
- Polishing robust method for extracting RR
  - Proving its **merit** against the GT device and other metrics
    - Formulating the exact RR
    - Finding accuracy, errors, etc
    - Testing effect of camera distance
  - Creating an **algorithm** and/or incorporating **ML** for anomaly removals (talking and movement)
    - Some directions may include **Viola-Jones** algorithm
- Shifting from recorded video to live testing
- Working alongside Zahid to *publish a report* for **SMART** and **CHASE** conferences

Acquired Skills

- Learning **data filtering**, **signal processing** uses/techniques, and **python libraries** to process and analyze data
- **Reading** papers and **deciphering** important information

Research Experience

- ***The Mantra:*** Measuring the BR of a living person without touching them is a *difficult task*
- I have learned tons about the research environment, the importance of mentor relationships, and the tenacity research requires
- **In summary:** *it's been falling down, getting up, falling again, … and getting back up again*
    - It takes resilience to bounce back in the face of failure and adversity; that's what discovering new ideas can feel like at times

Acknowledgements

- NSF REU Grant
- MPSC Lab
- Special thanks to Zahid Hasan for project support and Azim Khan for initial dataset