

Video File Uniqueness

Daniel Barry, Cole Flemmons, Rayhan Hossain, Jack Povlin, and Cole Schwerzler

Problem Statement

- Is there a systematic method for selecting a video's thumbnail?
 - The most distinctive frame in each video in a collection of videos
- The following conditions should be met:
 - Selected frame for a video is distinct to other selected frames from other videos
 - Selected frame is a good representation of the video

Motivation

- Make it easier to recall a scene or event that occurred in a video recording
- Challenging to find a particular scene in a collection of videos
 - Tedious to manually look through each video
- Videos often organized by metadata (i.e. timestamps, file size, file name)
 - Not organized by actual video content
- Video hosting sites, e.g. YouTube, let you choose your own thumbnail



Data Collection

- Video game capture
- What about automated collection?
 - Not necessary - search tool intended for those who *have* video data
- Simplifies development of algorithms
 - Less noise/contrast issues



Mario_1



Mario_2



StarFox



Zelda



BanjoKazooie



GoldenEye

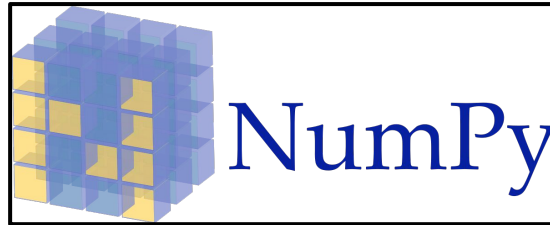


Mario_3



Mario_4

Explanation - Tools and Dependencies



Approach

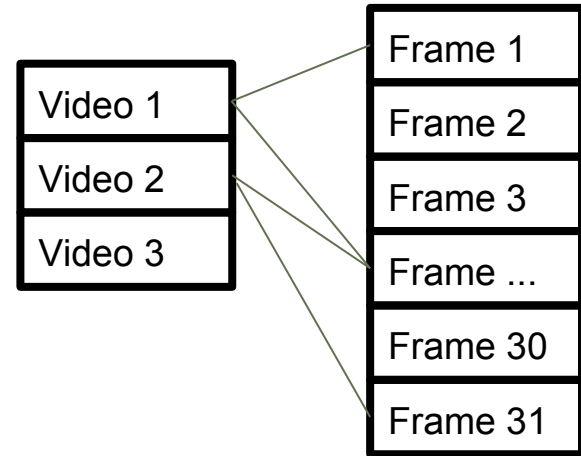
- Roles of Team Members
 - Cole F. -- Data Management
 - Cole S. -- Storage
 - Rayhan -- Image Processing
 - Jack -- Data Analytics
 - Daniel -- Workflow Management

Data Management

- Noise Reduction
 - Considered so that we'd have higher quality video data
 - Decided against it because we're just choosing a frame from the video
- Use a more manageable amount of video data
 - Our original videos were 30 seconds to 2 minutes long
 - 900 to 3600 frames each
 - Operations are expensive
 - 3 Minutes to process 15 second video
 - Split the videos up into 7 - 15 second chunks

Storage

- Need for data structure to search for frames from videos and vice versa
- SQLite
- Stores video attributes:
 - Name
 - ID no. (index)
- Stores frame attributes:
 - ID no. (index)
 - Video ID no.
 - Top 5 “most likely” RGB triples (15 elements, 5 each)
 - Top 5 most intense geometric properties - (10 elements, 5 reals, 5 imaginary)
 - 25-Dimensional data

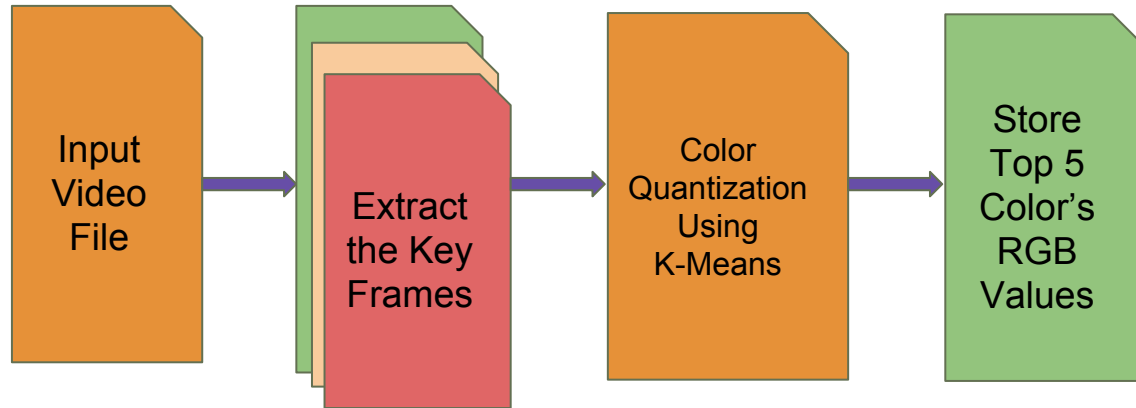


Sample Data

- Here is a sample video file. Data is collected using the N64 Console from the Zelda: Ocarina of Time video game.



Image Processing



Color Quantization

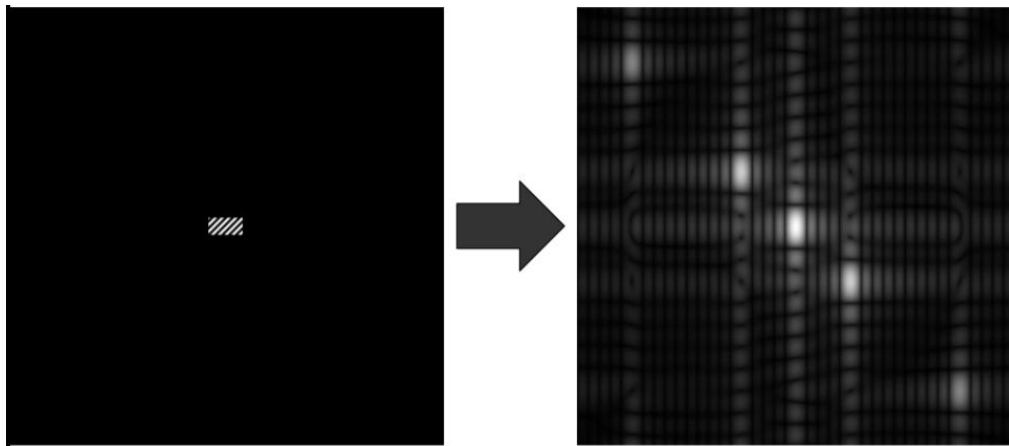
Color Quantization is the process of reducing number of colors in an image. Here is an example of quantization using K-Means for $k = 2, 4$, and 8 .



Image Source: https://docs.opencv.org/3.0-beta/_images/oc_color_quantization.jpg

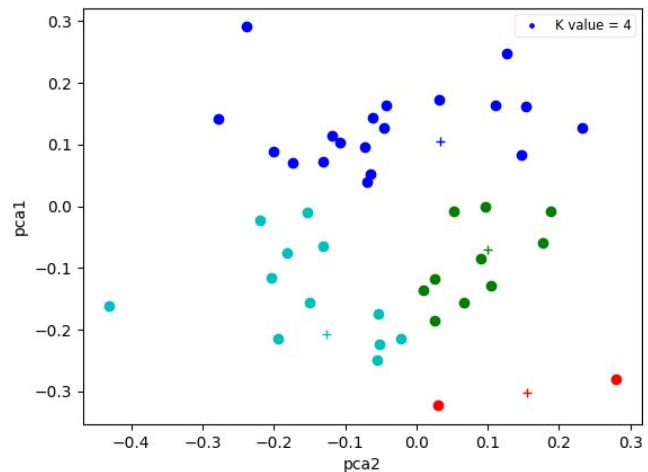
Geometric Properties Extraction

- 2D DFT → Most intense points indicate most prevalent geometric components of image
- This example is to illustrate the extraction (not actually from video)
- This operation is very computationally expensive
- Real and imaginary components



Data Analytics

- Query DB
- Perform PCA on video data retrieved from DB
- Perform K-means clustering on first 2 PCA's
 - Returns K best frames closest to cluster's center, as they best represent the video
- Once all videos have been processed, we perform PCA again to compare the frames
 - Frames with largest sum of distance to frames in other videos are chosen as distinct to other videos

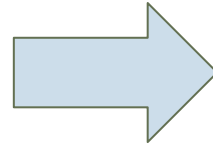


Project/Workflow Management

- Directly capture video data signals generated from hardware
- Install dependencies and libraries
- Ensure that commits integrate cleanly
 - Data structures compatible with routines
 - Routines produce expected results...
- Validating final outputs is difficult → not autonomous



Results



Results



Results - Homogeneous Case

- Three separate videos containing similar scenery...



Results - Conclusions

- It works fairly well
- Performance improves with larger samples
- It's possible that title screens are more distinguished than the rest of average gameplay
 - More unique geometry
- This tool alerts user of overlapping frames
 - Unintended but useful

Limitations

- Humans had to validate
- Peculiar cases with “fades to black” or moody frames
 - Increase # of sample videos?

Improvements

- Deinterlacing (only if absolutely necessary - thumbnails might be interlaced)
 - Loss of information for deinterlacing algorithms
- Specialized hardware (GPU)
- Parallel programming
- Dynamic hyperparameter optimization
- Optimizing database to reduce expensive video processing

Thank you!

Questions?