Video File Uniqueness Final Project Proposal

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

I. OBJECTIVE

Our team intends to develop a systematic method for selecting the most unique frame in each video in a collection of videos to serve as its thumbnail. Through developing such a methodology, we hope to provide organization-aware custodians of data a tool to more efficiently index video data by focusing on data content versus traditional metadata.

II. MOTIVATION

When trying to recall a scene or event that occurred in a video recording amongst a collection of videos, it can be a challenge to locate the video and temporal location in the video in which the event occurred. This problem is exacerbated by the tendency of files to be organized by metadata (i.e. timestamps, file size, file name), which do not explicitly indicate the events or scenes of the video. Metadata that does accomplish indicating the content, to some extent, is the video thumbnail. Manually traversing a video by jumping to various temporal locations can be tedious and often unproductive. Furthermore, certain online video-hosting websites, such as Youtube, allow users to select a ”custom” thumbnail for video content. This can be misleading because such a ”custom” thumbnail need not be an actual frame from the video it represents, further complicating a manual traversal through videos to find an event.

III. DATA USE

In this section, we emphasize the methods we plan to use to acquire, process, integrate, and validate the video data.

A. ACQUISITION

We have two primary means of obtaining data:(1) gathering videos from cell phones by manually selecting all videos of reasonable length and content and (2) scraping various legal online video-hosting websites. For source (1), we will have to manually obtain the videos due to legal policy and cell-phone user security purposes. For source (2), we can use such resources as torrenting clients to access the content from such video-hosting websites as BitChute. Neither of the sources allow for a significantly streamlined, automated means for collecting video data. However, this is not an issue because the purpose of this project is to serve as a provisional tool to those who are already in possession of a collection of video data. The collection of data we use is merely to serve as a proof of concept. (It is possible that we may find more source types for collecting the video data, and if we find them to be feasible, then we will incorporate them into the workflow.)

B. PROCESSING

In order to process the video data using a content-focused means, we have to first transform a video into a data structure that is comparable to other videos which have undergone the same transformation. We will then develop a ”video comparator” operation which will serve as a metric that can be consistently applied to all videos, regardless of time-length of the videos, resolution, file type, and otherwise distinguishing metadata, as such metadata does not indicate the true content of the video.

C. INTEGRATION

The integration of the video data in this project will serve to observe (1) the feasibility and quality of a framework to compare data structures that function as unique-enough signatures for videos while not excluding important features of a given video (2) the performance of such a framework as a video collection is scaled up. Performance refers to the scaling of execution time needed to achieve results of consistent quality.

D. VALIDATION

The validation for this project is complex in nature. Due to the notion of the thumbnails needing to be maximally unique with respect to one another, the results need to be validated by visual inspection by the development team. The difficulty in relying on manual inspection is the scale at which this is viable. A team of humans cannot readily remember hundreds of images and recall if each has a certain threshold of contrast. Therefore, we will develop a lightweight, post-processing tool that verifies the thumbnail images are reasonably unique compared to each other.

IV. MODELS AND ALGORITHMS

We will be using image processing to develop an automated method that collects metadata from a video and analyzes the video for distinct frames in it.

We plan to use existing image-processing algorithms, such as histogram equalization, colorspace conversions, and 2D Discrete Fourier Transforms (DFT). These will be used to construct signatures of the color and geometric attributes of a each frame in the video. Each of these signatures will be stored in the video's data structure. (There may be more attributes than color and geometric, and if they matter and are computationally feasible, then we will include them in the workflow.)

We will compare each of each frame's attributes in a given video to all of those in each frame of each other video.

We then will combine the results of the comparisons in a weighted manner to determine the optimal weights.

Finally, we will pass the results to the post-processing tool to improve the weights. We repeat the process until the weight coefficients cease to change beyond a certain percent.

A sketch of the algorithm is as follows:

```
Set w_1 = w_2 = 0.5
For each video i:
  For each frame j in video[i]:
    2ddft(frame[j]) → struct[i][j].geometric
    histequalzxn(frame[j]) → struct[i][j].color
While |max − prevmax| > TOL:
  For each video i:
    For each frame j in video[i]:
      Set prevmax = max = 0.0
      ig_1 = struct[i][j].geometric
      ic_1 = struct[i][j].color
      For each video k ≠ i :
        For each frame l in video[k]:
          ig_2 = struct[k][l].geometric
          ic_2 = struct[k][l].color
          buff = w_1 · gdiff(ig_1, ig_2) + w_2 · cdiff(ic_1, ic_2)
          if buff > max
            max = buff
          optimize(max, prevmax, w_1, w_2)
```

Fig. 1. The basic format of our algorithmic approach to this problem.

V. TIMELINE AND RESPONSIBILITIES

A. Responsibilities

- **Team Leader**:
  This person is responsible for assigning tasks, upholding deadlines, verifies the quality of computational results, and communicating with team members. (Daniel Barry)
- **Systems Programmer**:
  This person is responsible for ensuring that the development environment is up-to-date and all necessary software libraries and infrastructure are available. This person will also develop the main driving program for the workflow. (Jack Povlin)
- **Graphics Programmer**:
  This person is responsible for programming any specialized hardware used in the project. (Rayhan Hossain)
- **Data Officer**:
  This person is responsible for developing and maintaining the data structures needed to complete the project. (Cole Schwerzler)
- **Software Engineer**:
  This person is responsible for ensuring that the project follows good software engineering practices and tracking the compatibility of the steps in workflow. (Cole Flemmons)

B. Timeline

We have the following timeline:

- **Proposal** due on Monday, October 1, 2018.
- **Sample of data acquired** by Friday, October 12, 2018.
- **Verify readability of data** by Wednesday, October 17, 2018. (Jack Povlin and Cole Schwerzler)
- **Image processing code functional** by Wednesday, October 24, 2018. (Daniel Barry and Rayhan Hossain)
- **Data structure functional** by Wednesday, October 24, 2018. (Cole Flemmons and Cole Schwerzler)
- **Graphics-hardware ports of code completed** by Wednesday, October 31, 2018. (Daniel Barry and Rayhan Hossain)
- **Comparison functions for data structure** completed by October 31, 2018. (Cole Flemmons and Jack Povlin)
- **Quality assurance and testing completed** by Friday, November 9, 2018.
- **Final project report and presentation** due late November 2018.