Experimenting with Video Features

Submitted by: Dhananjayan Santhanakrishnan

Group Members – Group # 19

Bharat Kumar Suresh

Darshan Shetty

Dhananjayan Santhanakrishnan

Shivam Gaur

Uthara Keerthi

1. Abstract

The first phase of the project involves the study of the feature extraction from a video file in various formats that would be the key when we are analyzing the video for different operations. The first task involves creation of color histogram for the cells in the frame of a video based on the given input "R" which defines the division of the frame as R*R array. Whereas, the next task is mainly of SIFT vector extraction for a video, as in the previous case we extract SIFT vectors for each frame of the video. The final task was to extract motion vectors for the cells of the frame and this gives us the motion details of the vectors in the cell to the successive position in a video. All these features are explained in detail in this report.

Keywords: Motion Vectors, Histogram, SIFT, cells, feature extraction

2. Introduction

2.1. Terminology:

Color Histogram - A color histogram is a representation of the distribution of colors in an image. [3]

The Scale-Invariant Feature Transformation approach transforms an image into a large collection of local feature vectors, each of which is invariant to image translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. [2]

Motion Vector - a motion vector is the key element in the motion estimation process. It is used to represent a macroblock in a picture based on the position of this macroblock (Macroblock [4] is a processing unit in image and video compression formats) in another picture, called the reference picture. [1]

2.2. Goal Description

The goal of this Phase of project is to extract three key features pertaining to video that will help us analyze the video as per requirement in various junctures. These features give us insights about the videos while operating on them. All the three tasks have a commonality that they extract details of the image based on the frames of the video. We extract frames of the video individually and we extract three distinct features from them as, color histogram, SIFT vectors and motion vectors. These three features are extracted as a whole for the frame, however for convenience we separate the given frame in to cells that will logically partition the frame like matrix and divide the frames as equal or unequal chunks for convenience.

2.3. Assumptions

The following are assumed for the task completion

- The programs are executed in Windows 10 64 Bit operating system.
- All dependencies of the program are in the same path as in the Zipped folder.
- The given value of R is assumed not to exceed video height and width
- If either width or height is not divisible by R the resulting float value is floored to integer and the remaining pixels are appended to the last cell
- All the video inputs are assumed to have *.mp4 extension
- Code Name for the tasks are as follows, with reference from the folder given:
    - First Task to compute color histogram: MyCode/colorHist.m
    - Second Task to extract SIFT Vectors: sift/siftVectEx.m
    - Third task to extract motion vectors: MyCode.m
- Output files are as follows ‘out_file.chst’, ‘out_file.sift’ and ‘out_file.mvect’. They are created in current working directory.
- Path of the video files should be given with appropriate trailing slash or backslash as per the environment (Windows \ and Linux /)

3. Implementation

Task – 1 Color Histogram for frames in vector

The value of R, N and the video path is prompted from the user. After that the videos are read from the path. The videos are read as frames and divided into R * R cells for which we will calculate the histogram values and store in the output file. The code first converts the read frame to grayscale using Matlab’s rgb2gray() function whose output will be later used by imHist function to compute the color histogram values.

Task – 2 SIFT Vector extraction for Frames

The videos are divided into frames and the SIFT vectors are extracted from each frames. This task is done using the SIFT library.

Task – 3 Motion Vector extraction of frames

This task includes extraction of motion vectors for each frame which will give the values where the vectors will be in the next frame. This information is useful to manipulate videos. This is extracted through the FFMPEG library where we build a basic version of the C++ code which takes an input and outputs the motion vector parameters for each frame of the video. This C++ file is built in Visual Studio with all the dependencies (DLL’s in case of windows) and exported as an executable. After exporting it as an executable, we are executing that through matlab to get the motion vectors which is later manipulated and mapped to the appropriate cells in matlab.

4. System Requirements

- Windows 10 – 64 bit Operating System
- Matlab 2016a with Image Processing toolbox
- Visual Studio 2015 – to customize Motion Vector Output from the CPP in FFMPEG code library [7]
- Hardware Specs – Hard disk free space not less than 1 Giga bytes, RAM not less than 4 GB

5. Execution Instruction

Task # 1

a. First run 'firstTask/colorHist.m' file in the given package through matlab
b. After that you will be prompted for r value, please enter a value for r
c. The second prompt will be for bin value, please enter that
d. The final prompt will be the path of the video files, please enter the path with trailing slash or backslash according to the running environment

Task # 2

a. First run 'sift/siftVectEx.m' file in the given package through matlab
b. After that you will be prompted for r value, please enter a value for r
c. The final prompt will be the path of the video files, please enter the path with trailing slash or backslash according to the running environment

Task # 3

a. Please do not change the folder structure given, as the third task depends on the executable in the folder structure
b. The first prompt will be for the folder path of the video files, please pad the path with trailing slash or backslash according to the environment
c. The final prompt will be about the R value which will divide the frames based on this input

6. Related Work

The key points in the feature extraction above, has its motive to represent the features of the image in various ways that will help us to operate on them efficiently for different operations pertaining them. The color histogram is a basic representation of the image as histogram of color distribution in a given image. The distribution would be very helpful in comparing the image with another or classifying images under a particular category based on the inference from the color histogram distribution. For digital images, a color histogram represents the number of pixels that have colors in each of a fixed list of color ranges, that span the image's color space, the set of all possible colors. [3]

SIFT Vectors is Scale Invariant Feature Transform is the extraction of the distinctive features of the image that can be used for various matchings between different objects or scene. The representation is robust enough that it can provide matching across a substantial range of distortion and noise. [5] The SIFT vectors of the frames of the video are extracted using the SIFT Library provided from UCLA by Andrea Vedladi. [7]

Motion vectors on the other hand is a key component in motion estimation process, which is the representation of the image that transforms from one frame to another in a video sequence. [6] The motion vectors are extracted through the libraries of FFMPEG.

All these representation helps us to extract features from an image or frame after which it will be helpful to analyze and operate on the image for its manipulation, representations etc.,

7. Conclusion

This phase of the project had various tasks that encompassed on the feature extraction for an image or frame in a Video Sequence, which is a representation or perspective of the entity in different ways. This representation will be used to do various manipulations with the image or frame which extend to various application in multimedia processing. Through this experiment we were able to learn the highly basic representations to manipulate and do operations with multimedia data effectively.

8. Bibliography

[1] Motion Vector - Wikipedia

[2] David G. Lowe "Object Recognition from Local Scale-Invariant Features"

[3] Color Histogram – Wikipedia

[4] Macroblock – Wikipedia

[5] David G. Lowe - Distinctive Image features from Scale – Invariant Keypoints

[6] Motion Estimation – Wikipedia

[7] FFMPEG Library – Used for Motion Vector extraction process

[8] SIFT Library – Lowes SIFT Implementation for matlab by Andrea Vedladi