Roman Kolacz
Downing College
rk476

Part II Project Proposal

Time-Lapse Based Weather Classification

19th October 2014

Project Originator: *Alan Blackwell*

Resources Required: See attached Project Resource Form

Project Supervisor: *Advait Sarkar*

Signature:

Director of Studies: *Robert Harle*

Signature:

Overseers: *Markus Kuhn* and *Neal Lathia*

Signatures:

Introduction and Description of the Work

Time-lapse videos are a relatively unexplored aspect of data visualisation, and have a rather unique property amongst video of being able to hold information about incredibly large periods of time in a relatively small amount of time and space. This is especially true in the case of events which happen slowly and gradually, such as the weather conditions in a particular place over a long period of time.

This project will aim to use time-lapse video as an alternative way of associating time-lapse videos with external data sources by attempting to determine the weather from time-lapse footage and evaluating it against true weather information. This will be demonstrated graphically.

The time-lapse video will be made as part of the project, though any existing video will also suffice for the project.

Literature Review

Classification of Weather Situations on Single Colour Images

An approach which forms the basis of the weather-labelling per frame is described in the paper "Classification of Weather Situations on Single Colour Images"[4]. This paper describes classifying weather using image metrics to be used to assist driving-assistance systems which are heavily dependent on clear conditions. It achieves very accurate results using only two labels (clear and rain), and reasonable accuracy when adding a third, "light rain" label. This project will build on this by adding further labels to the weather classifying, as well as incorporating a cloud-cover measure.

The paper also mentions inaccuracies relating to weather situations containing heavy fog or rain, as visibility is greatly reduced. These will have to be taken into consideration and/or mentioned in the evaluation.

Weather Identifying System Based on Vehicle Video Image

"Weather Identifying System Based on Vehicle Video Image"[5], describes using image comparison to get the weather conditions on a road. A 'background' image is obtained (and updated should a more suitable image be detected), and then weather conditions are inferred by analysing the images obtained by subtracting each of 1000 frames in a video from the background image. Although this technique isn't going to be used to label weather in the approach which will be presented in this project, a similar technique may very well be employed to determine cloud cover percentage.

Classification of Road Conditions

The paper, "Classification of Road Conditions"[6], also applies weather classification to driving conditions. It uses an inverse-least-squares method of inferring the weather from not only images, but a multitude of weather data including temperature, humidity and wind speed. This project will not be using these additional variables, though the results will likely be evaluated with their aid.

Time-Lapse Photography Applied to Educational Videos

Although unrelated to weather classifying, "Time-Lapse Photography Applied to Educational Videos"[7] explores applying time-lapse videos to educational videos to aid with analysing long term trends and discarding short term 'noise'. This is comparable to what this project is aiming to achieve through the use of time-lapse video.

Two-Class Weather Classification

"Two-Class Weather Classification"[8] also uses image processing techniques to classify weather into two distinct labels: cloud and sun. It initially explores analysing pixel brightness distribution, which appears to fail in a general case. The paper then explores various techniques for looking for signs of a sunny day. One such technique involves detecting the sky in the image (something which this project may well have to incorporate) and checking for a blue hue. The paper also explores using edge detection to find shadows, which is found to be very vulnerable to false positives in a cloudy case; as well detecting bright reflections in reflective objects- two techniques which will not be employed in this project, though may be useful for discussion purposes.

Resources Required

- Obtaining the data will require a Raspberry Pi[1] with an attached camera module.

- Depending on the frequency of frame capture and resolution, some additional disk space may be required to store the data.

- OpenCV[2] will need to be installed.

- If the machine-learning part of the project is to be undertaken, Weka[3] will need to be installed.

Substance and Structure of the Project

The OpenCV library will be used to obtain the following image metrics from individual frames:

- Brightness

- Average Hue/Saturation

- Brightness curve

This data will then be analysed and compared to the images so that a set of hard-coded rules can be derived which will determine which weather label can be attached to each image.

Frames in which the sun is shining are likely to have have brightness curves with very high peaks and low troughs, in contrast to relatively flat curves of cloudy images. This would come as a result of shadow and reflections skewing the curve in sunlit images, whereas cloudy images will be generally more evenly (and dimly) lit as a result of the clouds dispersing and absorbing sunlight.

Snow will likely cause the entirety of the image to generally appear brighter than expected and with less saturation, as a result of the reflective properties of snow.

Rain will likely be the most challenging weather label to place, as the conditions will decrease the brightness of the frame in a similar way to dawn and dusk would, and otherwise be relatively similar to cloudy conditions. One potential situation to look out for would be reflections in puddles, though these aren't guaranteed to happen, and may have insignificant effects on the image even if they do. A probable situation is that the saturation of the frame decreases and the hue shifts towards a blue/grey shade, though this will need to be explored during the progress of the project, and the accuracy of which will need to be heavily analysed.

The cloud cover will be inferred by isolating the sky from the rest of the image, and then analysing the brightness and hue of the relevant part of each frame.

Evaluation

The weather classifier will need to be evaluated alongside real weather data. This will be obtained from forecast.io[9], which contains an API[10] for obtaining precise (latitude and longitude) weather information. This is free to use for up to 1000 polls a day, which will easily suffice for this project.

The data will be split into training and testing sections with a 4:1 split respectively, as (especially in the extension case involving machine learning), data which is used to train

the classifier will have a positively skewed accuracy. The testing section will be used to evaluate the classifier.

Each aspect of the classifier will then be evaluated separately.

Weather Labels

The weather labels, being discrete sets of data, will be tested for accuracy by simply calculating the percentage of correctly assigned labels. A truth matrix will be created to highlight the concentration of errors on a per-case basis.

Cloud Cover

The cloud cover will be calculated as a percentage and compared to the real value which is available from the Forecast API. The accuracy of this will be manually checked, given that a camera pointing outside will cover a relatively large section of the sky. The accuracy of this aspect of the classifier will be calculated using root mean square error.

Sunset and Sunrise

The sunset and sunrise times will be tested for accuracy by also using root mean square error, using the difference between actual and inferred time, in minutes.

Extensions

Depending on the progress of the project at specific points, the following additional features may be implemented, in order of priority:

- Multiple images per frame may be taken at varying exposure and gain settings as a large difference in luminance within a single frame (which will likely happen on a sunny day), may make other parts of the image unnaturally dark should the camera use automatic settings.

- Weather prediction and likely trends can be derived from existing data. Information predicted in such a way can be compared to inferred data and accurate data when the related time period passes.

- Machine learning can be used instead of hard-coded rules on determining weather, as both the image metrics and accurate weather information will be available as a learning data set. Weka will be used for this.

Success Criterion

For the project to be deemed a success the following items must be successfully completed.

1. The sunrise and sunset times can be inferred from a time lapse video.

2. The weather in each frame can be inferred and labelled one of the following:

   - Clear
   - Cloud
   - Rain
   - Snow

3. The cloud cover can be inferred from each frame within the video.

4. A measure of confidence can be assigned to each of the above criteria to reflect the probability of incorrectly inferred data.

5. The above data can be graphically demonstrated and shown to work to some degree of accuracy.

6. The accuracy of the data must be evaluated.

7. The dissertation must be planned and written.

Timetable and Milestones

24th October - 6th November

Research how to use the Raspberry Pi and camera module to take time lapse videos, and set up a fixed position for it to gather data from for the rest of the project.

Research the OpenCV library and how it can be used to gather the data required, as well as how the data gathered can be used to infer the conditions described in the success criteria.

Milestones

- Produce a short, not necessarily usable time lapse video using the Raspberry Pi.

- Begin gathering data for remainder of project.

- Found/made a relatively hassle free mount for the Raspberry Pi and camera.

- Produce some small demo's and examples using OpenCV.

7th November - 18th December

Use the research gathered to design and implement algorithms which can establish weather conditions from frames based on hard coded rules.

Milestones

- A working library for inferring weather conditions from images.

- Tests written for said library.

Key Dates

- NPR End: 14th December

19th December - 1st January

Review algorithms and adjust timetable if further tests and research is needed for a working, usable library.

Implement the library to the thus-gathered time lapse video, and design a graphical demonstration program to view the data through. Assign a confidence value to each frame to reflect the chance of a correct result.

Milestones

- A graphical visualisation of the data extracted from the time lapse video.

- A confidence measure for the weather at each given frame.

2nd January - 29th January

Work on evaluating the data by gathering accurate weather information for each obtained frame and automatically labelling them with said data. Compare some specific instances to vaguely gauge the accuracy of the program.

Milestones

- Accurate data gathered for evaluation.

- A general idea of how accurate the data is.

Key Dates

- NPR Start: 4th January

30th January - 19th February

Work on more formally evaluating the accuracy of the data, along with the reliability of the confidence measure assigned to each piece of data. This can also be graphically shown.

Milestones

- Complete data on accuracy of algorithm.

- Graphical visualisation of said accuracy.

Key Dates

- Progress Report Deadline: 30th January

- Progress Report Presentation: 5th, 6th, 9th or 10th February

- Dissertation Lecture: 11th February

20th February - 5th March

Review work and adjust timetable if needed. Enhance test program for a more friendly UI which is more suitable for demonstrative purposes.

Milestones

- Program should be fully complete, usable and presentable.

6th March - 2nd April

Write dissertation. Notes should have been gathered throughout the previous stages, so these should be read first for familiarisation of the previous weeks of work.

Milestones

- First draft of dissertation should be complete.

Key Dates

- NPR End: 15th March

3rd April Onwards

Take a look at what needs attention, if anything, and work on that. Look through dissertation and ensure it's up to submission standards.

Milestones

- Final dissertation should be complete.

Key Dates

- NPR Start: 12th April

- Project Deadline: 15th May

- Supervisor Report Deadline: 20th May

- Exams: 2nd-4th June

- Viva Announcement: 12th June

- Vivas: 15th June

- NPR End: 21st June

References

[1] Raspberry Pi Computer
http://www.raspberrypi.org/

[2] Open CV Library
http://opencv.org/

[3] Weka 3 Library
http://www.cs.waikato.ac.nz/ml/weka/

[4] Classification of Weather Situations on Single Colour Images
Martin Roser and Frank Moosmann, 2008
http://www.mrt.kit.edu/z/publ/download/Roser_al2008iv.pdf

[5] Weather Identifying System Based On Vehicle Video Image
Hong Jun Song, 2011
`http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5970722`

[6] Classification of Road Conditions
Patrik Jonsson, 2011
`http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6059917`

[7] Time-Lapse Photography Applied To Educational Videos
Wei Liu and Hongyun Li, 2012
`http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6202303&`
`tag=1`

[8] Two-Class Weather Classification
Cweu Lu, Di Lin, Jiayi Jia and Chi-Keung Tang, 2014
`http://www.cse.cuhk.edu.hk/leojia/papers/weatherclassify_cvpr14.pdf`

[9] Forecast.io
`https://forecast.io/`

[10] Forecast.io Developers API
`https://developer.forecast.io/docs/v2`