**Automatic Annotation of Family Photo Albums**

**By: Taylor Luke and Angela Garcia**

PROJECT REPORT

1.Introduction

-What is the general problem

-Why is it important; what will be the impact of a solution

-How is the rest of the report structured

For this assignment we chose to work on Automatic Annotation of Family photo album. Our main goal was to find an efficient way to store and retrieve a large number of digital photos that are stored in multiple locations without any annotation. To achieve this goal we wanted to create a software to accurately locate a given picture that contains a timestamp so that it makes it easier to identify members of a photo. Using the location and time the software should accurately label the names of the people in the photo.Facial recognition is necessary to help identify people. Automatic annotation is necessary because it allows the user to keep track of when and where the photo was taken. In this report we will discuss our evolution of our project. Discuss what we were able to achieve and in what we struggled on the most. Along with the future plans for this project.

2.Proposed problem formulation

-What is the specific thing you are trying to do

-Explain clearly the specific components of the problem

-Explain how the formulation fits into solving the general problem you have talked about earlier

-NOTE: This section is key, since it tells readers what to look for in your work

For this assignment our goal is to create a program that is able to locate the time and location that the photo was being taken. We want to identify the time and location of the photo so that it can be organized and create easy access for the user. In order to achieve this development of the automatic annotation of the photo we need to take into consideration in the place and time that the photo was being taken. To help facilitate organization among the photos it is essential to take note of the social context that the photo was being taken. Keeping track of the frequency that the person appears in each photo. Some analytical considerations that should be taken when creating a program for the automatic annotation of the photo are analyzing the physical features of each individual for example: body structure, eye alignment, and facial expression. Along with analyzing graphical details in each photo like color, format, digitalization , and alignment of the photo. To better organize and categorize who appears in the photos implementation of Metadata needs to be expanded along with systems like the Photo 4W system continuously analyzing who appears in the photos along with the time and location that the photo was taken. Evaluating components like alignment and graphical structure of the photo keeping focus on a specific group and event of photos. Although facial recognition is a contributing factor of automatic annotation, the format and structure of the photo needs to be recorded and analyzed before actually stepping into facial recognition.

3.High-Level description of your idea

-Write about how you want to go about solving the problem

-Present thoughts on why your approach makes sense

-SELL YOUR IDEAS! POINT OUT NOVELTY! SOUND EXCITED!

In order to approach this problem our main focus will be time and location for the development of our automatic annotation program. We will define social context as time and location and visual context and body. With social context information takes advantage of the fact that in a family setting the same group of people tend to appear in the same social events and they tend to wear the same clothing. This social context information is used to cluster photos into events, so that the visual context of the recognized faces can be used to find other presence of the same person. With this information it can be estimated the probability of the presence of the person based on the results of the face and body recognition.

4.Prior work in the area

-Very Important Section

-If you have done any surveys, read papers, articles, books etc. related to the problem

–discuss them here

-IMPORTANT: relate how the prior work is related to the general problem and toyour specific formulation

-Compare advantages and disadvantages as impartially as possible

-Show at the end, how these ideas inspire your work. Also show why they are not sufficient to solve the problem (if this is applicable)

-If you are doing something applied, show which works you are

basing your project on.

With so many pictures that can be taken at once the users want an efficient way to pick the best photo, but it becomes very time consuming to pick the right one. There are multiple methods that are being developed so that the user can efficiently select a photo from their list of photos. Allowing them to select and remove any photo that they do not want. There are methods that allow the user to flip through multiple pictures just like most of our phones contain a photo gallery that allows us to select and organize our photos. This is where we can sort and thumbnail (or pin) any photo that we like.

During certain methods of automatic annotation what researchers look for are ways to analyze how people organize their photos meaning self organization and pinning pictures that the users consider as their favorite. Taking into consideration the visual, temporal, and spatial context of the photo. Analysing the visual annotation of the photo it is taken into consideration the features of the photo analysing the shape, color, texture, and structure of the photo. There have been issues with temporary metadata in which time and date the photo was taken. Being able to extract the photo with the photo metadata file data. Many methods have had issues in retaining information about the who, where, when the photo was taken. It has been emphasized that it is easier to annotate photos when it is taken during a big event.

Brushing and dragging allows the user to carefully compare a set of photos. It allows the user to focus on certain areas of the photo and compare it to multiple similar photos. Using their finger to brush over a specific spot of the photo and compare that part with the other photos. Then once the user is content with the photos the user can drag the photos and organize them in any order they like. Giving them the option to remove or select any photo they like. With the Brush and Drag method it emphasises focus and context of the photo. Brushing supports the focus part in which the user chooses an area of the photo by just simply brushing over the area with their fingers. And the context is emphasized by how the photo was ranked and retained.

Lastly photo maps, with the photoMaps method its automatic annotation is used by spatial, temporal, and social context that a photo has. Like we have discussed previously. PhotoMaps collects metadata based on weather conditions and nearby significant objects. It provides the user with a way to organize, share, and retrieve photos. Photomap uses bluetooth in the mobile device to identify and record the spatial moment of the photo being taken. Just like world wide media exchange, photo maps also have a spatial temporal interface. Photomaps is more diverse compared to the previous methods because it is able to connect to web services to gain more context information. The issues that have been encountered is that it has not been able to automatically organize photos into events and has not been able to develop a mechanism to query the metadata of the photos.

5.Description of your work(Details of the structure will depend on what you are doing)

-Overview of system description

-Specific technical details for each part of the system on which work was done

-Relate how the technical work supports your idea

-Cover each part. Relate the design and development to the

high-level description of your idea as well as to prior work whenever

Applicable

-In the case of a team project: Each section should indicate the

names of those who worked on it.

For our system we used python code. An important first step was downloading openCV to work with image processing. From here we implemented simple function to make a time stamp on a photo, along with user input to also stamp the location of the photo. Once use input was taken we then sorted each jpg by order of latest to most recent. Our main idea was time and location and we used the time() function and sort() for technical reasons to improve our code. These functions helped us reach our goal of showing how to sort photos by time and location.

In an attempt to understand how facial recognition is developed among programs we researched demos to follow so that we get the general idea of how facial recognition implementation works. We tried implementing our own and realized that it is a lot more complicated than anticipated. Finding that one of our biggest challenges is having an environment that supports facial recognition libraries. Giving us access to facial recognition methods. For research purposes we followed a demo by Adarsh Menon who walked us through the basics of what it takes to implement a program that involves facial recognition. For this demo we used google colab which allowed us to have a workspace that supported the facial recognition libraries. The facial recognition program takes a face image as an input and generates 128 encodings that will be used to be compared to another image. In order to see if the two images are similar it will compare the encodings of each image. The more similar the encoding is, the more likely the image contains the same person. If the encodings are not similar then it means that the faces in the image do not match. In this facial recognition program we have a list of known faces and a list of unknown faces. In the program it will take the image encodings and will compare it to the known and unknown faces. If the encodings match then it will create a red box around the face along with the name of the identified person.

6.Experimental evaluations

-Quantitative data which shows how well your ideas and

implementations work

-Design experiments (essential for multi-student groups) that show

the efficacy of the design and each component that was implemented

-The experiments should provide QUANTITATIVE data to analyze

your works

-In some cases, QUALITATIVE data –such as user surveys are o.k

and even required (e.g. Analysis of the UI)

-It is not necessary that the experimental results are all good.

-Both good and poor results should be analyzed in this section. The

analysis should try to answer the question: Why are we getting these

results? What are their implications for the project?

For our project we didn’t necessarily have data to analyze like most of the other projects. We analyzed based on our code sorting accuracy and ever time the code was run, the sort by date feature was always correct. Our design was based on time and location. Our code design first started off by time stamping our photos. To get a time stamp we used a simple time() function. Next we focused on sorting these photo by time. We used a sort files function which we then specified the files to be a jpg. After this we took user input to label each photo with a specific location. Our results were always conclusive. We ran the code multiple times and every time user input was prompted, then pasted correctly on the photo followed by an accurate time stamp. Then all these jpg’s were sorted correctly in terminal output.

7.Conclusions

-Summarize the project and results. Relate what you were able to

develop to the original goals you had set in the introduction and problem formulation sections

-Point out next steps you may want to consider

Overall, we started this project a little of track. We were confused as to the whole concept of this project. We thought we were just deeply analyzing the facial recognition issue and so when midterms came around and we were notified we were off track we started to get a little panicked. We then quickly came up with a language we wanted to use and format of the type of program we wanted to make. Installing open cv was the biggest problem of this problem. It is very sensitive to the type of python version you have and the IDE version you have as well. Pip was also very sensitive and if you didn't have the right version it would not work. After finally getting this install we were able to start working on our code. We went in order of time stamp, commentary stamp, sorting, user input and then implementation of user input. We faced only a few problems when figuring the code out and eventually we got it to our goal of how we want to take on the issue of time and location annotation. Our goal in the very beginning was to far to reach, creating a facial recognition program would take way more time than a semester could allot. We are very happy with our final product because we feel it covers our whole idea of focusing on time and location. If we had next steps I would then try to sort out the photos by user input, changing the order to be by event and not time.

8.References

-References should be listed alphabetically in terms of the last name

of the first author

De Figueirêdo, Hugo Feitosa, Lacerda, Yuri Almeida, De Paiva, Anselmo Cardoso, Casanova, Marco Antonio, and De Souza Baptista, Cláudio. "PhotoGeo: A Photo Digital Library with Spatial-temporal Support and Self-annotation." Multimedia Tools and Applications 59.1 (2011): 279-305. Web.

Kim, Seon Joo, Ng, Hongwei, Winkler, Stefan, Song, Peng, and Fu, Chi-Wing. "Brush-and-drag." (2012): 59-68. Web.

Pagare, Reena & Shinde, Anita. (2012). A Study on Image Annotation Techniques. International Journal of Computer Applications. 37. 42-45. 10.5120/4616-6295.

Rowland, Duncan, Appiah-Kubi, Kwamena, Shipp, Victoria, Mortier, Richard, and Benford, Steve. "Annotation and Anonymity." (2015): 1-10. Web.

Vermeulen, Jo, MacDonald, Lindsay, Schöning, Johannes, Beale, Russell, and Carpendale, Sheelagh. "Heartefacts." (2016): 712-23. Web.

[Video: Face Recognition in Python using face\_recognition Library (in Google Colab)](https://youtu.be/987QtKPZ-P0)

Work done by each team member:

For the initial two powerpoints, we did equal research and equal slides. For the final, Taylor worked on the team’s code in python and did that demo. Angela for the final worked on the demo for the google facial recognition software.