



Hat Trick Coach 2019

Divyanshu Pandey



School of Information Technologies

Faculty of Engineering & IT

ASSIGNMENT/PROJECT COVERSHEET - GROUP ASSESSMENT

Unit of Study: Multimedia Retrieval

Team name: Hat Trick Coach

Tutorial time: 8:00 Pm Professor name: Zhiyong Wang

DECLARATION

We the undersigned declare that we have read and understood the [*University of Sydney Academic Dishonesty and Plagiarism in Coursework Policy*](#), and, except where specifically acknowledged, the work contained in this assignment/project is our own work, and has not been copied from other sources or been previously submitted for award or assessment.

We understand that failure to comply with the *Academic Dishonesty and Plagiarism in Coursework Policy* can lead to severe penalties as outlined under Chapter 8 of the *University of Sydney By-Law 1999* (as amended). These penalties may be imposed in cases where any significant portion of my submitted work has been copied without proper acknowledgement from other sources, including published works, the internet, existing programs, the work of other students, or work previously submitted for other awards or assessments.

We realise that we may be asked to identify those portions of the work contributed by each of us and required to demonstrate our individual knowledge of the relevant material by answering oral questions or by undertaking supplementary work, either written or in the laboratory, in order to arrive at the final assessment mark.

By

Divyanshu Pandey

Contents

Abstract.....	4
Project Summary.....	4
Related work	5
Value of Project.....	5
Milestones and Current Progress	5
Procedure Followed	6
Detailed Results	8
Scene Detection (PySceneDetect).....	8
Morphological Transformations	9
Finding the statistical data	10
Output.....	11
Reflection on the Group Project (University of North Dakota Standard, USA)	11
Bibliography	12

Abstract

In this report we provide our detailed implementation of feature similarity in provided videos of Soccer teams. The technique we are trying is described by Baraldi, L., C. Grana, and R. Cucchiara [1], and later we refined the techniques introduced by Datta, S. and N. Oschlag-Michael [2] along with Kapela, R., K. McGuinness, and N.E. O'Connor [3]. We implement the main techniques and applications introduced in the papers. We also compared the methods and tried to overcome the limitations in the stated techniques.(techniques used by the authors)

Project Summary

The goal of our project is to implement video feature comparison, by applying and analysing the videos footages of the sport effectively we will be finding out the feature similarity in different teams from the recorded matches [2], which will be helpful in the providing the coaches and the team management teams about the team performance and playing pattern and formation so that they can make changes in the team for a better team performance.

Broadcasting of sports is a huge market with the viewers estimated to reach 3.4 billion by the year 2023 in the industry of sports broadcast. With the most fans in the world, broadcast of football has its various applications in the commercial market. For example, specialized videos help the development of a football club and highlights packages provide a better interaction with fans. However, due to a faster speed in producing videos, higher demand for better quality of videos, more competitive market and limited manpower handling videos, automated content-based analysis and retrieval technology becomes an essential tool for a better service in football industry.

In this project, Hat Trick Coach, we aim to build a video retrieval system where coaches find the benchmarks of his football team to follow. “Hat Trick” is a term meaning a player scoring equally or more than three goals in one match. When putting this title on coaches, we have a vision to let our system help a coach produce more hat trick players in his team.

With the records of videos in matches, coaches can find out the team’s performance in the match along with the statistical results. As a result, coaches can deliver players which style of their team is and how to advance further based on the benchmark of the most similar world-class football clubs in the world.

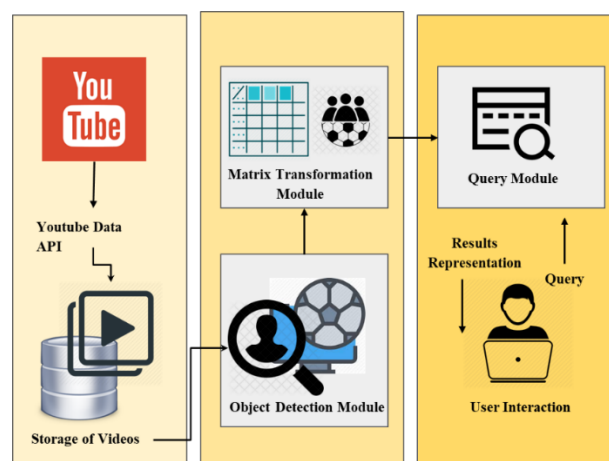


Figure 1: Activity Flowchart

Related work

There are basically two simple approaches for content-aware image retargeting: which include discrete and continuous. The discrete approaches tend to manipulate the individual pixels, on the other hand continuous methods perform a complete mapping of the source image to the target image. These both methods could also make use of extra information and constraints in order to optimize their decisions (e.g. object detection, face detection, user supplied region selection, etc.).

Known approaches include: Shift-mapping, scale-and-stretch, energy-based deformation, and nonhomogeneous warping. Each method suffers from certain limitations and perform better on some tasks over others.

The basic idea includes performing a series of detection on the scenes to find the related frames and out of those given frames in order to decide whether the features (Patterns, Textures, content) could be extracted with the help of the scene detection tools.

Value of Project

- To contribute towards improvement of efficiency and increase in effectiveness of the Football teams and further development of the team play.
- To find out **the game pattern of the teams** and different areas of the game where the team's lack or have further room for development as per player position in the team.
- Enhanced Gameplay monitoring feature will **facilitate regular and sustainable monitoring** of players performance, coordination with each other and level of skill set to the coaches for timely and appropriate advice **and training-based policy formulation and decision making**.
- **Grouping the teams with the similar playing style** and pattern through gameplay recognition.

Milestones and Current Progress

TASK NAME	START DATE	DUE DATE	DURATION (Days)	% DONE	MILESTONE
Draft Proposal	27-03-2019	03-04-2019	7	100	Phase 1
Data Acquisition	03-04-2019	12-04-2019	9	100	Phase 1
Scene Detection	12-04-2019	18-04-2019	6	100	Phase 2
Feature Extraction (Player)	18-04-2019	26-04-2019	8	100	Phase 2
Analytics Functionality	26-04-2019	09-05-2019	13	100	Phase 3
Documentation and Manual	09-05-2019	19-05-2019	10	100	Phase 4

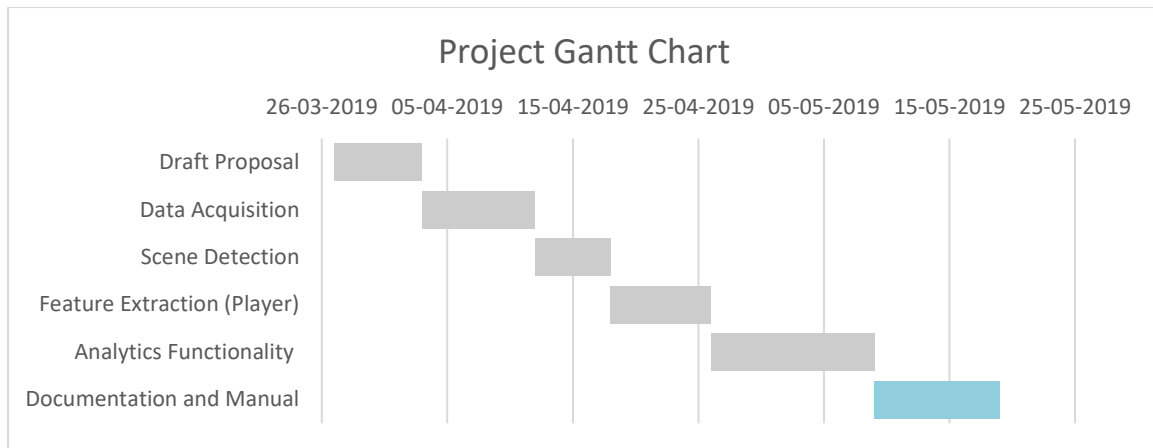


Figure 2: Milestone of the projects

Procedure Followed

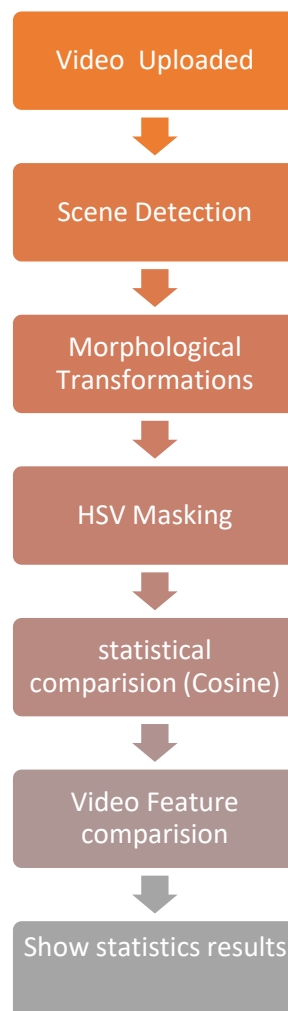


Figure 3: Procedure followed for feature similarity

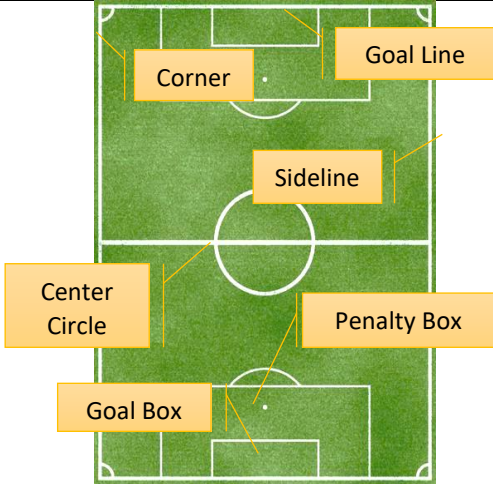
Football Field	Ontology of a Football Match
	<p>Object:</p> <ul style="list-style-type: none"> • People: Goalkeeper, Defender, Midfielder, Striker, Coach, Referee, Audience • Facility: Ball, Goal Post, Grass, Line <p>Event: Goal, Corner Kick, Foul, Injury, Running, Special Skills</p> <p>Concept: Condensed Press, Sparse Formation, Win, Lose and so on.</p>

Figure 4: Elements of Ontology in a Football Match (put in related work)

The entire project is divided into five major segments which are:-

- Data Acquisition
- Scene Detection
- Texture Extraction
- Statistical Analysis
- Similarity Analysis

In the initial step we collect the data from the data source (YouTube) through the API, then we filter out the correct frames which are supposed to be taken for the Analysis.

This is followed by the HSV masking of the background which in turn is followed by morphological transformation which helps in detecting the desired objects (players) in the given frame for which we had to specify the condition (height and width of the players pixel) and when the player is identified they are categorized into different teams based on the masking of their team's jersey colour.

Then different frames have a statistical comparison and the results show the distance between the team players and the domination of the team in the different frames.

Detailed Results

Scene Detection (PySceneDetect)

It is a command-line tool, written in Python and uses the OpenCV, which helps in analysing a given video, it seeks and finds for scene changes or cuts. The output generated timecodes can then be used with another tool for splitting the video into different individual clips. For a given video using this as a frame-by-frame analysis can also be generated, in order help with determining optimal threshold values or detecting patterns/other analysis methods for a particular video.

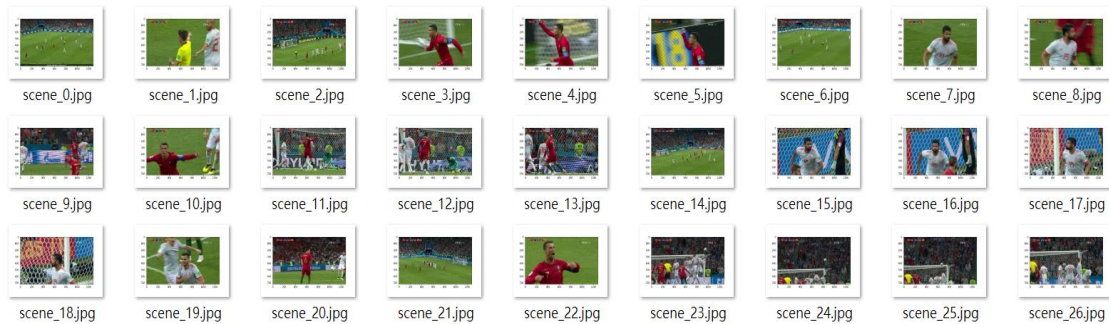


Figure 5: Detected scenes in video

So, in the given video we are given a number of scenes but all are not related to the gameplay and some are player close up shots and other are simply penalty shoot outs.

Thus, we need to select the frames in which only the entire gameplay is being shown.

For example,

The scenes selected will be, the following as they cover the entire background rather than the

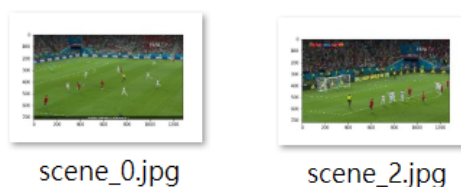


Figure 6: Selected Scenes

On the other hand the scenes rejected will be

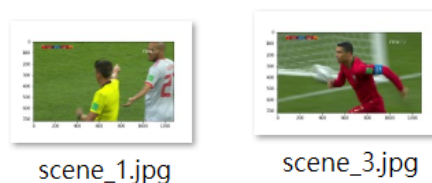


Figure 7: Rejected Scene

Once the scenes are selected, even then there are other details which remain to be extracted along with features in the given frame so that, we are able to identify the activities which are being performed in the frame and the important scene is selected

Morphological Transformations

Morphological transformation can be termed as the simple operations which are based on the image shapes and are usually performed on the binary images. It requires minimum of two inputs one is our original image, second one is called **structuring element** or **kernel** which decides the nature of operation.

Erosion and Dilation are two of basic morphological operators. Though there are variation which also come into play these include Opening, Closing, Gradient etc. These are obtained through the erosion of an image followed by a dilation.

```
dst = open(src, element) = dilate(erode(src, element))
```

It is Useful for removing small objects (it is assumed that the objects are bright on a dark foreground)

To identify players, we need to extract 2 image features:

- ▶ Shape Feature: Masking the field and conduct the morphological operation. (contour)
- ▶ Color Feature: With contours, we set a filtering condition to the pattern of shape and use masking operation to identify players' teams.
 - ▶ $h \geq (1.2) * w$ and $h > 10$ and $w > 10$
 - ▶ Number of colored pixel > 30

- Masking- It returns a binary mask (an `ndarray` of 1s and 0s) the size of the image where values of 1 indicate values within the range, and zero values indicate values outside.



Figure 8 :Masking of the background

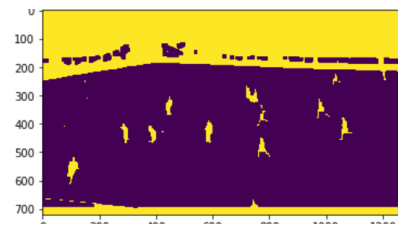


Figure 9 : Morphological transformation

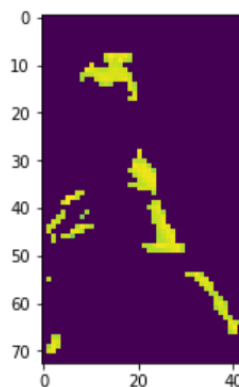


Figure 10: Masking of the player

Finding the statistical data

Number of Players

```

Number of team 1 players:  7
Number of team 2 players:  4
Domination Ratio of Team 1 in this frame:  0.6363636363636364
Domination Ratio of Team 2 in this frame:  0.36363636363636365

```

Distance of Players

- Distance Metric: *norm2*
- Normalization of Distance Matrix:

$$distance = \frac{\sqrt{(x1 - x2)^2 + (y1 - y2)^2}}{max_distance}$$

```

team_2_distance matrix
[[0.    0.31 0.51 0.7 ]
 [0.31 0.    0.2  0.38]
 [0.51 0.2   0.    0.2 ]
 [0.7   0.38 0.2   0.   ]]

```

In order to calculate the statistical data between the frames we used the norm2 to calculate the distance between the players appearing in the given frame of the same team along with finding the distance between the players of different teams as well.

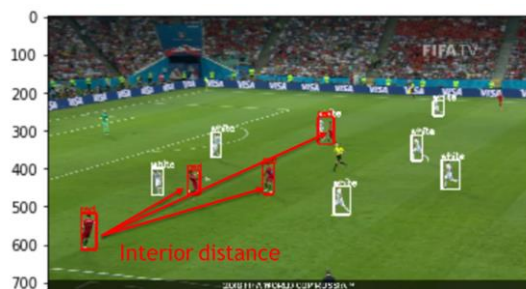


Figure 11: Interior team distance

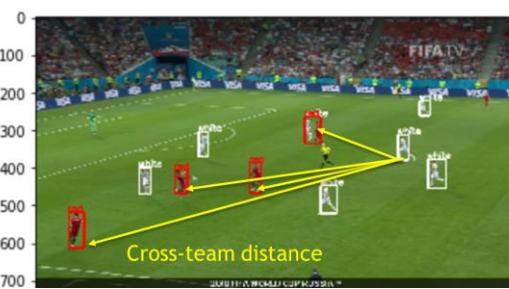
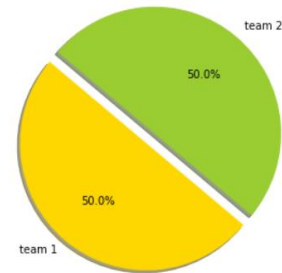
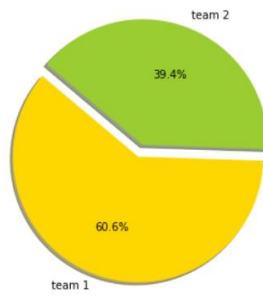


Figure 12: Cross Team Distance

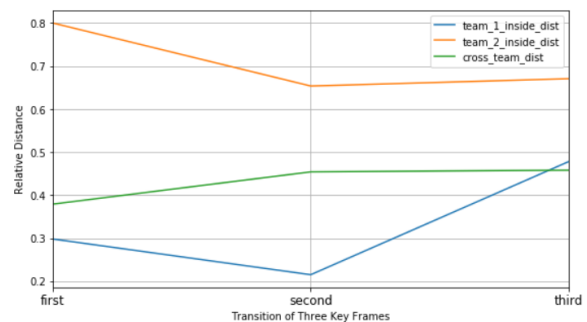
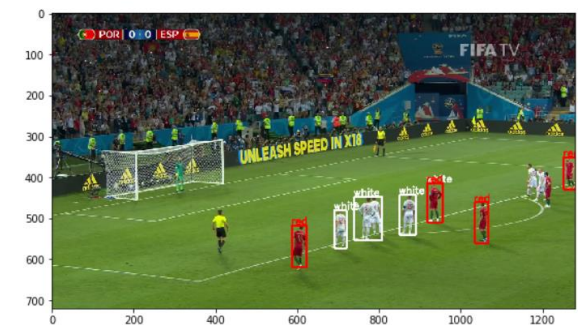
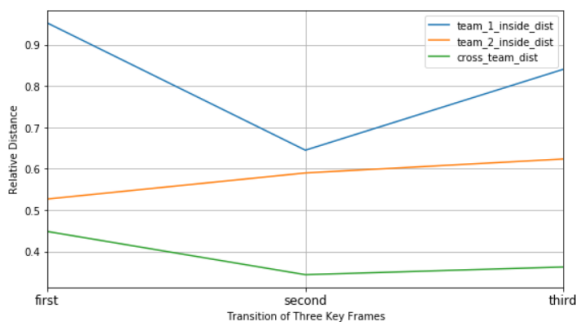
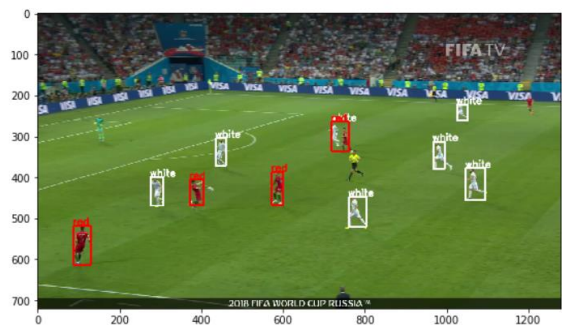
Output



Select 2 Scenes:

Scene 1 Index: 0

Scene 2 Index: 1



Reflection on the Group Project (University of North Dakota Standard, USA)

Working together is a part of life, and it only gets better when you are working with someone who thinks like you. That makes working fun as this was the case with us as we worked together on our project 'Hat trick Coach'.

As part of COMP5425, Multimedia Retrieval Project, We were assigned to come up with the system design along with its implementation using the multimedia management system such as retrieval, browsing, summarization, analytics, social media applications, and mobile media applications. We were successful in drafting the proposal and now converting that in an actual and working system. We developed a system which compares the video feature similarity in the football match.

Through making this project's presentation, we learned about several aspects of collaboration that need to be in place in order for a group to smoothly and successfully reach their goal. One of these aspects is good lines of communication. We needed to know what the goal is and we finally decided what the presentation would look like for which we had agreed-upon idea of how the goal will be accomplished timely.

For deciding the different tasks that we needed to perform we decided to have a thorough research, for which we decided to use multiple resources in order to know about the similar projects and how they were achieved. We researched multiple databases for related articles such as **Gartner** as well as **Google scholar**, we also went through multiple **GitHub** repository to evaluate the similar developments.

We also decided to have an open line of communication on **Facebook** which made communications simpler also we used **Google Drive** so they our development could be easily seen by each other which also the sharing of ideas and receiving constructive feedback. Positive interdependence is another important part of collaboration.

We decided to make the presentation precise, crisp and to the point as we had only a few minutes to cover the project and show what all things we had covered. We wanted to attain the maximum attraction from the audience along with maximum clarity for which followed the guidelines specified by the **GCFglobal.org** also we went through the guidelines specified by the University of North Dakota.

Bibliography

1. Baraldi, L., C. Grana, and R. Cucchiara, *A deep siamese network for scene detection in broadcast videos*. arXiv preprint arXiv:1510.08893, 2015.
2. Datta, S. and N. Oschlag-Michael, *Understanding and Managing it Outsourcing: A Partnership Approach*. 2015, Basingstoke: Palgrave Pivot.
3. Kapela, R., K. McGuinness, and N.E. O'Connor, *Real-time field sports scene classification using colour and frequency space decompositions*. Journal of real-time image processing, 2017. **13**(4): p. 725-737.
4. Abdullah et al. (2018), "Soccer Event Detection," Computer Science & Information Technology (CS & IT), pp. 119-129.
5. Adrian Rosebrock (2018), "Yolo Object Detection with OpenCV," Deeping Learning, Objection, Tutorials, pyimagesearch