

COMP3419 Assignment 1*

Nick Zhou 460363707

October 2018



THE UNIVERSITY OF
SYDNEY

*Option 1

Contents

1	Introduction	1
1.1	Preamble	1
1.2	Task	1
2	Implementation	2
2.1	Process	2
2.1.1	Capturing Motions	2
2.1.2	Replacing Background and Marionette	3
2.1.3	Adding Intelligent Objects	3
2.1.4	Adding a Sound Track	3
2.2	Algorithm	3
2.3	Experimental Results	4
3	Conclusion	4
3.1	Motion Capture	4
3.2	Background and Marionette Replacement	5
3.3	Intelligent Objects	6
3.4	Sound Track	6
3.5	Final Words	6

Contents

1 Introduction

1.1 Preamble

The purpose of this report is to demonstrate the pipeline used to complete the COMP3419 Assignment (Option 1). This report consists of a three main sections; introduction, implementation and conclusion. This report was created with L^AT_EX as required by the assignment specification.

1.2 Task

The task of this assignment was to program a short video involving digital video processing, compositing and 2D animation techniques. The output is an animation piece based on the provided clip. The original video was of a marionette monkey with red markers on the head, hands and feet. Our task involved tracking these red markets in order to determine the motion of the marionette monkey and recreate it with our own character. The programming language used to complete this assignment was python3 on the windows operating system.

This assignment task was broken into a set of 4 subtasks excluding this report:

- **Motion Capture:** The body of the monkey is labelled with red markers. The subtask involved the segmentation of the red markers in order to track the motions of the monkey.
- **Replace Background and Marionette:** The blue background was to be replaced with a dynamic background. The monkey marionette was to be replaced with our own character with the original motions of the old marionette being preserved.
- **Intelligent Objects:** Two randomly moving objects were to be added to the video and required to interact with the marionette in at least two different ways. Special effects were to be triggered when the interactions between the marionette and intelligent objects occur.
- **Sound Track:** At least two sound tracks were to be added to the video. The sound track should be relevant to the interaction between the intelligent objects and the marionette.

2 Implementation

2.1 Process

The process of implementing the assignment involved performing the subtasks in steps.

2.1.1 Capturing Motions

The red parts of the monkey were found by first converting the frames to HSV (Hue Saturation Value) and then applying a threshold. The use of the HSV threshold to identify the red parts of the monkey proved to be far superior to using RGB. The segmentation of the red markers was enhanced by performing two iterations of morphological erosion operations, followed by a single iteration dilation operation and finally another iteration of erosion. An elliptical structuring element was used to perform these morphological operations.

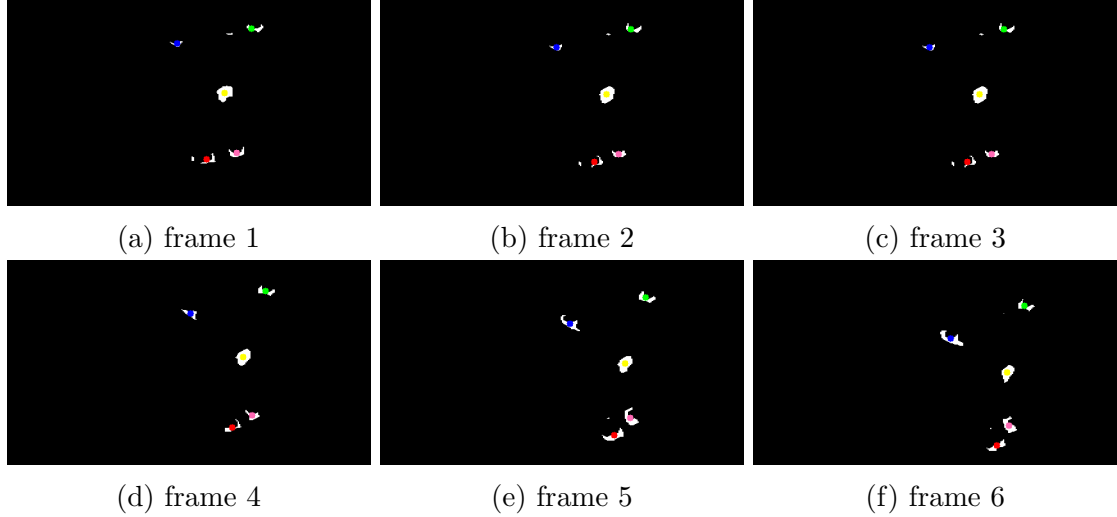


Figure 1: Example of the marked centers being accurate over a series of moving frames

2.1.2 Replacing Background and Marionette

The background is replaced by not using any part of the original video at all. Instead the only inputs in this stage from the last one are the movements of the monkey. The marionette replacement is achieved by creating a Donald Trump character with each of his 5 body parts superimposed on where there monkey's red markers would have been.

2.1.3 Adding Intelligent Objects

The intelligent objects were created by loading textures from the image folder. Their movement vector is randomly chosen between a sequence of -10 and 10 inclusive. Collisions between intelligent objects and marionette body parts are treated as circular objects when determining their resultant velocities. When a collision between an intelligent object and marionette body part occurs, a pow image is played on top of the collision area.

2.1.4 Adding a Sound Track

Whenever an intelligent object collides with the marionette, a sound track is queued to play. The sound is played using the winsound library.

2.2 Algorithm

The K-Medoids algorithm was used on the red dots to distinguish them into proper clusters and identify each part of the monkey, the discernible parts being the head, two arms and two legs. The initial centers of the clusters were manually set to be as close to the middle of each of the monkey parts as possible.

2.3 Experimental Results

In attempting to locate the red dots on the monkey, various morphological techniques were used. Performing a closing operation resulted in "clusters" that had very low dissimilarity.

An opening operation yielded a result where the 5 red markers were easier to discern and had a high dissimilarity.

Various other operations were performed, such as alternating between erosion and dilation operations with different number of iterations. Eroding for more than 2 iterations before performing a dilation resulted in too few red pixels being retained. If the dilation was performed first, the red dots joined too close to each other, which resulted in poorer cluster centers.

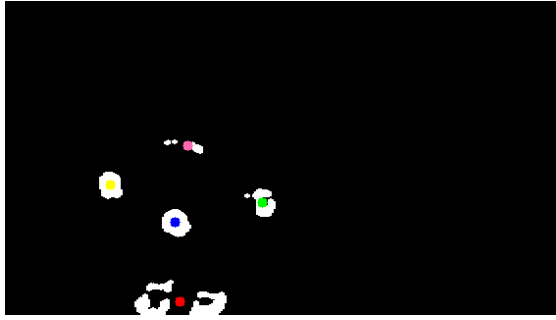
Various structuring elements were used when experimenting with the morphological operations. A square structuring element of size 3x3 tended to work better than one of size 5x5. A 5x5 square structuring element would result in too much being eroded in erosion operations, and too much noise with dilation operations. Though the 3x3 structuring element worked well, it was inferior to an elliptical 5x5 structuring element.

3 Conclusion

The implementation of K-Medoids was highly susceptible to noise. Fast movements would often result in the deformation of the clusters. The cross-over between the two legs markers did not break the clustering algorithm in matching the motions of the monkey. The heavy amount of erosion performed on the image prevented the two clusters from merging into a single one.

3.1 Motion Capture

Despite the many draw backs of the implementation, all four tasks required for the assignment were certainly met successfully. Most of the monkey's motions were accurately captured using the K-Medoids classifier.



(a) Example of the clusters being misidentified due to noise



(b) As a result of the cluster centroids being misidentified in previous frames such as the previous figure, the initial medoids result in incorrect classification of the monkey parts

Figure 2: When the clusters start being misclassified, future frames will also be misclassified

3.2 Background and Marionette Replacement

The blue background and the monkey was replaced as required by the specification. The character replacement (Donald Trump) simulated the gestures of the monkey and had all 5 connected components.



Figure 3: This still shows a frame of the original video with the blue background and monkey marionette



Figure 4: This still shows a frame of the output video in which the background and marionette is replaced

3.3 Intelligent Objects

The condition of having two intelligent objects interact with the marionette was met and their motions affected when they came into contact with each other. The requirement of a special effect being triggered during interactions was satisfied as demonstrated in following figure.



Figure 5: A still image of a special effect being triggered when an intelligent object collides with the marionette

3.4 Sound Track

Two soundtracks were programmed for the video. They are triggered to play when the intelligent objects collide with the marionette. The sound track that is played depends on the type of intelligent object that collides with the marionette. If it is an Obama object, the sound track "you're fired" is played. Alternatively, when a Hillary object collides with the marionette, the "such a nasty woman" sound track is played.

3.5 Final Words

In conclusion, though the algorithms used in the implementation of the assignments tasks were not perfect, they still allowed the completion of the tasks to a very high standard. Thus this report can note that this assignment has been completed satisfactorily.

List of Figures

1	Example of the marked centers being accurate over a series of moving frames	3
2	When the clusters start being misclassified, future frames will also be misclassified	5
3	This still shows a frame of the original video with the blue background and monkey marionette	5
4	This still shows a frame of the output video in which the background and marionette is replaced	5

5	A still image of a special effect being triggered when an intelligent object collides with the marionette	6
---	--	---