

Informatics Institution of Technology

In Collaboration with

University of Westminster, UK.

Automatic Video Descriptor for Human Action Recognition

A dissertation by

Ms M.V. Perera

Supervised by

Mr. M. Cassim Farook

Submitted in partial fulfillment of the requirements for the

BEng (Hons) in Software Engineering degree

Department of Computing

April 2016

Abstract

Assistive software such as screen readers are unable to describe images or videos for visually impaired people. Although recent research have found ways to automatically describe images, automatically describing the content of a video is still an ongoing issue. Visually impaired people find it difficult to understand video content without an indication of sound. The current solution of video description is only provided through digital television for selected programs and movies. Since descriptions are manually added extra cost, time and effort is needed. As an initiative to describe video content for visually impaired people, the solution act as video player which automatically understand the ongoing human action on screen, associate textual descriptions and narrate it to the blind user.

The human actions in the video should be recognized in real time, hence fast, reliable feature extraction and classification method must be adopted. A feature set is extracted for each frame and is obtained from the projection histograms of the foreground mask. The projection histograms contains the number of moving pixels for each row and column of the frame. These values provide sufficient information to identify the instant position of a person. Support Vector Machine is used to classify extracted features of each frame. The final classification is given by analyzing frame wise classifications in segments. The classified actions will be converted from text to speech.

Key words : Human Action Recognition, Classification, Support Vector Machine, Machine Learning, Feature Extraction, Action Detection