Abstract

Having survived the ordeal of a laryngectomy, the patient must come to terms with the resulting loss of speech. With recent advances in portable computing power, automatic lip-reading (ALR) may become a viable approach to voice restoration. This thesis addresses the image processing aspect of ALR, and focuses three contributions to colour-based lip segmentation.

The **first contribution** concerns the colour transform to enhance the contrast between the lips and skin. This thesis presents the most comprehensive study to date by measuring the overlap between lip and skin histograms for 33 different colour transforms. The hue component of HSV obtains the lowest overlap of 6.15 %, and results show that selecting the correct transform can increase the segmentation accuracy by up to three times.

The **second contribution** is the development of a new lip segmentation algorithm that utilises the best colour transforms from the comparative study. The algorithm is tested on 895 images and achieves percentage overlap (OL) of 92.23 % and segmentation error (SE) of 7.39 %.

The **third contribution** focuses on the impact of the histogram threshold on the segmentation accuracy, and introduces a novel technique called *Adaptive Threshold Optimisation (ATO)* to select a better threshold value. The first stage of ATO incorporates ϵ-SVR to train the lip shape model. ATO then uses feedback of shape information to validate and optimise the threshold. After applying ATO, the SE decreases from 7.65 % to 6.50 %, corresponding to an absolute improvement of 1.15 pp or relative improvement of 15.1 %. While this thesis concerns lip segmentation in particular, ATO is a threshold selection technique that can be used in various segmentation applications.