Abstract
Producing caption for the deaf and hearing impaired is a labor intensive task. We implemented a software tool, named SmartCaption, for assisting the caption production process using automatic visual detection techniques aimed at reducing the production workload. This paper presents the results of an eye-tracking analysis made on facial regions of interest to understand the nature of the task, not only to measure of the quantity of data but also to assess its importance to the end-user; the viewer. We also report on two interaction design approaches that were implemented and tested to cope with the inevitable outcomes of automatic detection such as false recognitions and false alarms. These approaches were compared with a Keystoke-Level Model (KLM) showing that the adopted approach allowed a gain of 43% in efficiency.

Keywords
Caption production, eye-tracking analysis, facial recognition, Keystoke-Level Model (KLM).