

UNIVERSITAT POLITÈCNICA DE VALÈNCIA
DEPARTAMENT DE SISTEMES INFORMÀTICS I COMPUTACIÓ



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Different Contributions to Cost-Effective Transcription and Translation of Video Lectures

Thesis

presented by Joan Albert Silvestre Cerdà

supervised by Dr. Alfons Juan Císcar and Dr. Jorge Civera Saiz

November 18, 2015

Durante estos últimos años, los repositorios multimedia on-line han experimentado un gran crecimiento que les ha hecho establecerse como fuentes fundamentales de conocimiento, especialmente en el área de la educación, donde se han creado grandes repositorios de vídeo charlas educativas para complementar e incluso reemplazar los métodos de enseñanza tradicionales. No obstante, la mayoría de estas charlas no están transcritas ni traducidas debido a la ausencia de soluciones de bajo coste que sean capaces de hacerlo garantizando una calidad mínima aceptable. Soluciones de este tipo son claramente necesarias para hacer que las vídeo charlas sean más accesibles para hablantes de otras lenguas o para personas con discapacidades auditivas. Además, dichas soluciones podrían facilitar la aplicación de funciones de búsqueda y de análisis tales como clasificación, recomendación o detección de plagios, así como el desarrollo de funcionalidades educativas avanzadas, como por ejemplo la generación de resúmenes automáticos de contenidos para ayudar al estudiante a tomar apuntes.

Por este motivo, el principal objetivo de esta tesis es desarrollar una solución de bajo coste capaz de transcribir y traducir vídeo charlas con un nivel de calidad razonable. Más específicamente, abordamos la integración de técnicas estado del arte de Reconocimiento del Habla Automático y Traducción Automática en grandes repositorios de vídeo charlas educativas para la generación de subtítulos multilingües de alta calidad sin requerir intervención humana y con un reducido coste computacional. Además, también exploramos los beneficios potenciales que conllevaría la explotación de la información de la que disponemos a priori sobre estos repositorios, es decir, conocimientos específicos sobre las charlas tales como el locutor, la temática o las transparencias, para crear sistemas de transcripción y traducción especializados mediante técnicas de adaptación masiva.

Las soluciones propuestas en esta tesis han sido testeadas en escenarios reales llevando a cabo numerosas evaluaciones objetivas y subjetivas, obteniendo muy buenos resultados. El principal legado de esta tesis, *The transLectures-UPV Platform*, ha sido liberado públicamente como software de código abierto, y, en el momento de escribir estas líneas, está sirviendo transcripciones y traducciones automáticas para diversos miles de vídeo charlas educativas en numerosas universidades e instituciones Españolas y Europeas.