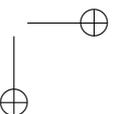
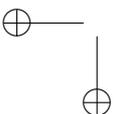


# Resumen

Gracias a avances significativos en aprendizaje profundo, la traducción del habla (ST) se ha convertido en un campo consolidado, lo que permite la utilización de la tecnología ST en soluciones para entornos de producción. Como consecuencia del aumento constante del número de horas de contenido audiovisual generado cada año, así como una mayor sensibilización sobre la importancia de la accesibilidad, la ST está preparada para convertirse en un elemento clave para la producción de contenidos audiovisuales, tanto de ocio como educativos.

A pesar de que se ha progresado significativamente en ST, la mayor parte de la investigación se ha centrado en el escenario en diferido (*offline*), en el cual todo el audio de entrada está disponible. En cambio, la ST en directo (*online*) es una temática en la que falta mucho por investigar. En concreto, existe un caso de traducción en directo, la traducción continua (*streaming*), que traduce un flujo continuo de palabras en tiempo real y bajo unas estrictas condiciones de latencia. Este es un problema mucho más realista, que es necesario resolver para que sea posible aplicar la ST a una variedad de tareas de la vida real.

Esta tesis está centrada en investigar y desarrollar las técnicas claves que son necesarias para una solución de ST continua. En primer lugar, de cara a permitir el desarrollo y la evaluación de sistemas de ST, se ha recopilado un nuevo conjunto de datos para ST multilingüe, que expande significativamente el número de horas disponibles para ST. A continuación se ha desarrollado un segmentador preparado para la condición continua, que se utiliza para segmentar las transcripciones intermedias de nuestra solución por etapas, que consiste



---

en un sistema de reconocimiento automático del habla (ASR), seguido de un sistema de traducción automática (MT) encargado de traducir las transcripciones intermedias al idioma de destino elegido. Diversas investigaciones han concluido que la calidad de la segmentación es un factor muy influyente es la calidad del sistema MT, por lo que el desarrollo de un segmentador efectivo es un paso fundamental en el proceso de ST continua. Este segmentador se ha integrado en la solución por etapas, y estas se optimizan de manera conjunta para alcanzar el equilibrio óptimo entre calidad y latencia.

La ST continua tiene unas restricciones de latencia mucho más estrictas que la ST en directo, ya que el nivel deseado de latencia tiene que mantenerse durante todo el proceso de traducción. Por tanto, es crucial ser capaz de medir de manera precisa esta latencia, pero las métricas estándar de ST en directo no se adaptan bien a esta tarea. Como consecuencia de esto, se proponen nuevos métodos para la evaluación de ST continua, que garantizan unos resultados precisos a la vez que interpretables.

Por último, se presenta un nuevo método para mejorar la calidad de la traducción continua mediante el uso de información contextual. Mientras que los sistemas tradicionales de ST en directo traducen audios de manera aislada, existe abundante información contextual que está disponible para mejorar los sistemas de ST continua. Nuestra propuesta introduce el concepto de historia continua, que consiste en el almacenamiento de la información más reciente del proceso de traducción, que se utiliza más adelante por el modelo para mejorar la calidad de la traducción.