
Resumen

El reconocimiento de texto manuscrito (HTR) en árabe y la traducción automática (MT) del árabe al inglés se han tratado habitualmente como dos áreas de estudio independientes. De hecho, la idea de crear un sistema que combine las dos áreas, que directamente genere texto en inglés a partir de imágenes que contienen texto en árabe, sigue siendo una tarea difícil. Este proceso se puede interpretar como la traducción de imágenes de texto en árabe. En esta tesis, se propone un sistema que reconoce las imágenes de texto manuscrito en árabe, y que traduce el texto reconocido al inglés. Este sistema está construido a partir de la combinación de un sistema HTR y un sistema MT.

En cuanto al sistema HTR, nuestro trabajo se enfoca en el uso de los *Bernoulli Hidden Markov Models* (BHMMs). Los modelos BHMMs ya han sido probados anteriormente en tareas con alfabeto latino obteniendo buenos resultados. De hecho, existen resultados empíricos publicados usando corpus conocidos, tales como IAM o RIMES. En esta tesis, estos resultados se han extendido al texto manuscrito en árabe, en particular, a las bases de datos IfN/ENIT y NIST OpenHaRT.

En aplicaciones reales, la transcripción del texto en árabe no se limita únicamente al texto manuscrito, sino también al texto impreso. El texto impreso se puede interpretar como una forma simplificada de texto manuscrito. Por lo tanto, para este tipo de texto, también proponemos el uso de modelos BHMMs. Además, estos modelos se han comparado con tecnología del estado del arte basada en redes neuronales.

Una idea clave que ha demostrado ser muy eficaz en la aplicación de modelos BHMMs es el uso de una ventana deslizante (*sliding window*) de anchura adecuada durante la extracción de características. Esta idea ha permitido obtener resultados muy competitivos tanto en el reconocimiento de texto manuscrito en árabe como en el de texto impreso. De hecho, un sistema basado en este tipo de extracción de características quedó en la primera posición en el concurso *ICDAR 2011 Arabic recognition competition* usando la base de datos *Arabic Printed Text Image (APT)*. Además, esta idea se ha perfeccionado mediante el uso de técnicas de reposicionamiento aplicadas a las ventanas extraídas, dando lugar a nuevas mejoras en el reconocimiento de texto árabe.

En el caso de texto manuscrito, este refinamiento ha conseguido mejorar el sistema que ocupó el primer lugar en el concurso *ICFHR 2010 Arabic handwriting recognition competition* usando IfN/ENIT. En el caso del texto impreso, este refinamiento condujo a un sistema mejor que ocupó el segundo lugar en el concurso *ICDAR 2013 Competi-*

tion on Multi-font and Multi-size Digitally Represented Arabic Text en el que se usaba APTI. Por otro lado, esta técnica se ha evaluado también en tecnología basada en redes neuronales, lo que ha llevado a resultados del estado del arte.

Respecto a la traducción automática, el sistema se ha basado en la combinación de tres tipos de modelos estadísticos del estado del arte: los modelos *standard phrase-based*, los modelos *hierarchical phrase-based* y los modelos *N-gram phrase-based*. Esta combinación se hizo utilizando el método *Recognizer Output Voting Error Reduction (ROVER)*. Por último, se han propuesto tres métodos para combinar los sistemas HTR y MT con el fin de desarrollar un sistema de traducción de imágenes de texto árabe a inglés. El sistema se ha evaluado sobre la base de datos NIST OpenHaRT, donde se han obtenido resultados competitivos.