

Generación eficiente de transcripciones y traducciones automáticas en poliMedia

Juan Daniel Valor Miró, Carlos Turró, Jorge Civera y Alfons Juan

jvalor@dsic.upv.es, turro@cc.upv.es, jcivera@dsic.upv.es, ajuan@dsic.upv.es
Universitat Politècnica de València

Abstract

Usage of educational videos in higher education has raised quickly for different educational applications, which leads to platforms and services such as poliMedia, of the Polytechnic University of Valencia (UPV), which enables the creation, publication and dissemination of this educational multimedia content.

Through various research projects, and specifically the European project transLectures, UPV implemented a system that automatically generates subtitles in various languages for all poliMedia videos. Creation of these subtitles is made by an automatic speech recognition and machine translation system that provides high accuracy in both recognition and translation to the main European languages.

Transcripts and translations are not only used to improve accessibility, but also allow search and retrieval of video contents within the video portal. So a user can locate the video, and the time within it, where a certain word is said for later viewing.

In this article we also extend previous work in the assessment of the review process, including transcription of French and Spanish to Catalan translation.

Keywords: transcription, translation, video, e-learning, docencia en red

Resumen

El uso de videos docentes en el mundo universitario se ha extendido e implantado rápidamente en diferentes aplicaciones docentes, lo que da pie a plataformas y servicios como poliMedia, de la Universitat Politècnica de València (UPV), que permite la creación, publicación y difusión de este contenido multimedia docente.

Por medio de diversos proyectos de investigación, y específicamente el proyecto europeo transLectures, la UPV implementó un sistema que genera automáticamente subtítulos en diversos idiomas para todos los videos de poliMedia. Esta generación de subtítulos está basada en sistemas de reconocimiento automático del habla y de traducción automática, que proporcionan una elevada precisión en dicho reconocimiento y traducción a los principales idiomas europeos.

Las transcripciones y traducciones generadas no se utilizan únicamente para mejorar la accesibilidad, sino que también permiten la búsqueda y localización de los contenidos de vídeo dentro del portal. Así se puede localizar el vídeo, y el instante dentro de él, donde se dijo una determinada palabra.

En este artículo también ampliamos los trabajos previos sobre la evaluación del proceso de revisión, incluyendo la transcripción de francés y la traducción de español a catalán.

Palabras clave: *transcripción, traducción, vídeo, e-learning, docencia en red*

Introducción

La Universitat Politècnica de València (UPV) lleva años apostando por el uso y la integración de vídeos docentes en sus programas formativos. Los usos más comunes de estos vídeos docentes en la propia universidad son los siguientes: integrados en cursos MOOC online en la plataforma UPVx, como material complementario a través de la plataforma poliMedia al ofrecido en la docencia presencial (Turró, 2009), como grabaciones de clases presenciales puestas a disposición del alumnado a través de la plataforma vídeoApuntes, o en la incorporación de metodologías docentes semipresenciales mediante la aproximación de clase inversa (flip-teaching).

De todos estos usos de los vídeos docentes, sin duda alguna el que posee un mayor recorrido y penetración es su uso como material complementario a la docencia presencial a través de la plataforma poliMedia. Esta plataforma permite la creación, publicación y difusión de vídeos docentes por parte del profesorado de la UPV.

Adicionalmente, poliMedia comprende desde la preparación de los vídeos docentes hasta la distribución y publicación de los mismos; ofreciendo de esta forma instrumentos, materiales y personal técnico para facilitar la grabación de los vídeos docentes. Estos vídeos son grabados mostrando una vista fija conjunta del profesor y de la proyección de la pantalla del ordenador, habitualmente diapositivas o una aplicación informática, como se puede ver en la Figura 1.



Fig. 1 vídeo Polimedia con subtítulos generados automáticamente

En el marco de diversas iniciativas de investigación de la UPV, y especialmente con el proyecto europeo transLectures (Silvestre, 2012), se generaron transcripciones y traducciones de forma automatizada en español, catalán e inglés para todos los vídeos de poliMedia (Valor Miró, 2012).

transLectures es el acrónimo de “*Transcription and Translation of video Lectures*”, en el cual se aplicaron técnicas de reconocimiento automático del habla y de traducción automática con el objetivo de proporcionar transcripciones y traducciones automáticas a grandes repositorios de vídeos docentes, entre los cuales estaba poliMedia.

El uso de transcripciones y traducciones aporta a los vídeos docentes un gran valor añadido posibilitando la accesibilidad a personas con discapacidad auditiva o por aquellos que simplemente desconocen la lengua del vídeo (Papadopoulos, 2012).

Además, los subtítulos permiten enriquecer a los repositorios de nuevas funcionalidades (Glass, 2007) como búsqueda en el contenido, recomendaciones de vídeos mucho más precisas, búsqueda de palabras clave dentro del propio vídeo, generación de apuntes de manera semi-automática, etc.

La integración de las búsquedas en las transcripciones se ha realizado mediante el uso de un buscador de texto completo, que se ha implementado en el portal media.upv.es, que se muestra en la Figura 2, donde la página de inicio del portal muestra de forma señalada un buscador, que es la forma de navegación preferida por los usuarios.

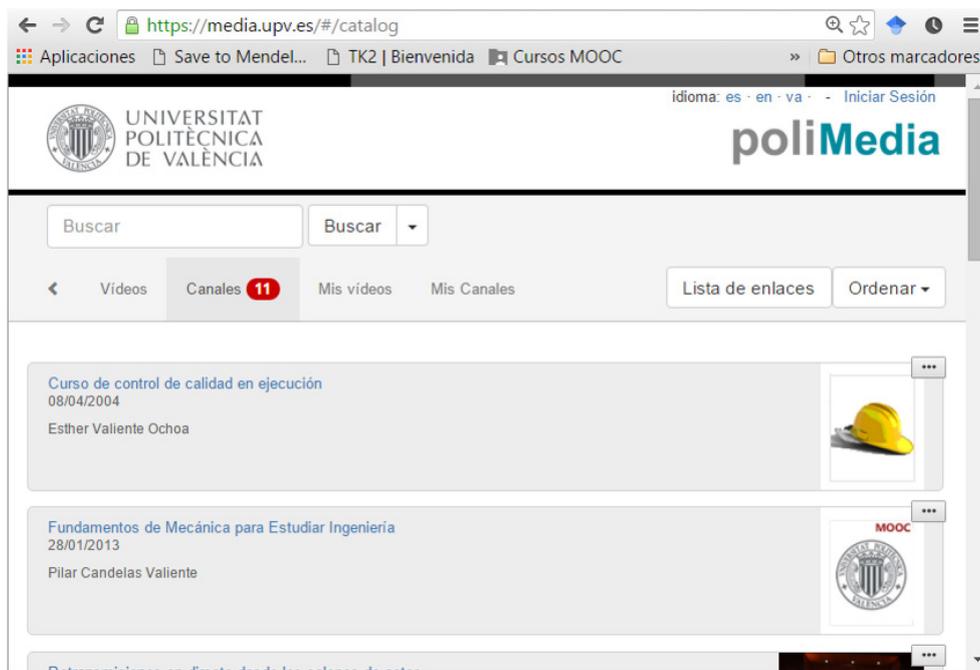


Fig. 2 Portal de videos media.upv.es

1. Objetivos

Los objetivos fundamentales del proyecto participante de la iniciativa Docencia en red durante el curso 2015-2016 eran tres:

1. Mejorar de forma continua la calidad de las transcripciones y traducciones a partir de la revisión de las mismas por los profesores de la UPV
2. Añadir nuevos idiomas de transcripción y traducción
3. Integrar de forma eficiente las transcripciones y traducciones en el proceso de búsqueda de videos del portal media.upv.es

Respecto del primer objetivo, es conocido que debido al proceso automático de obtención de transcripciones y traducciones, estas no están libres de poseer errores, por lo que requieren de la revisión por parte del profesorado de las mismas. Trabajos previos demuestran que supervisar transcripciones automáticas es siempre más eficiente que generarlas desde cero (Valor Miró, 2015).

Este trabajo se llevó a cabo durante el programa de la UPV “Docencia en red” del año 2013-2014 (Valor Miró, 2015) obteniendo mejoras sustanciales de ahorro de tiempo respecto a la transcripción totalmente manual del 73%, 44% y 38%, en la revisión de transcripciones en español, catalán, e inglés, respectivamente; y del 65% la revisión de traducciones del español al inglés.

Este año planteamos la ampliación de los idiomas de transcripción, introduciendo la revisión de vídeos en francés; y de traducción, con la revisión de traducciones desde el español al catalán. Asimismo, actualizamos los resultados obtenidos en (Valor Miró, 2015) como consecuencia de la mejora continua de los sistemas de transcripción y traducción implicados.

Adicionalmente se ha trabajado en la implementación de los mecanismos de búsqueda sobre las transcripciones y traducciones, focalizándose en dos servicios:

1. Búsqueda de portal, donde se ha realizado un tratamiento transcripciones para que se usen conjuntamente con el resto de los metadatos de los vídeos (Autor, título, etc)
2. Búsqueda dentro del vídeo, que permite localizar las palabras dentro de la transcripción una vez se está visualizando el vídeo.

2. Desarrollo de la innovación

Esta innovación se ha llevado a cabo en el marco del programa *Docencia en red*, que incentiva la adopción de las nuevas tecnologías en la elaboración de materiales docentes por parte del profesorado.

Para ello, 72 profesores realizaron la revisión de transcripciones y traducciones de vídeos docentes a través de la plataforma transLectures, actualmente integrada con éxito en el repositorio poliMedia de la UPV.

En esta edición se ha realizado un seguimiento y evaluación de la revisión por parte del profesorado de la UPV de transcripciones en francés y de traducciones del español al catalán. Concretamente se han revisado las transcripciones de 20 vídeos docentes en francés (2.5 horas) y las traducciones del español al catalán de 9 vídeos (0.7 horas).

Además, la revisión de transcripciones en español, catalán, e inglés, y la revisión de traducciones del español al inglés también se han llevado a cabo en la presente edición, obteniendo resultados muy similares a los expuestos en la edición pasada del programa *Docencia en red*.

Para realizar estas revisiones se ha empleado un innovador reproductor de vídeo web (Valor Miró, 2012) que permite la modificación de los subtítulos mediante una interfaz que puede verse en funcionamiento en la Figura 3.



Fig. 3 Reproductor de supervisión de transcripciones y traducciones

Respecto del interfaz de búsquedas, primero se realiza un preprocesado de las transcripciones en la cual se efectúa la eliminación de stopwords y la búsqueda de raíces semánticas (stemming):

- Stopwords: es una lista de palabras de uso frecuente que, tanto en la indexación como en la búsqueda no se tienen en consideración y se omiten. En general se trata de palabras comunes que no aportan información (artículos, preposiciones, conjunciones, etc.)
- Stemming: es un método para obtener la raíz semántica de una palabra. Las palabras se reducen a su raíz o stem (tema), de modo que, si buscamos por “perdidos” encontrará “perdidos” pero también “pérdidas”, “perdimos”, ... porque, en realidad, estamos buscando por “perd-”.

Evidentemente, las stopwords y los algoritmos de stemming tienen que estar específicamente diseñados para cada lengua.

A continuación se realiza una compensación entre los metadatos. En la actualidad se muestran en el portal primero las coincidencias en el título y el autor y luego los resultados ponderados de las búsquedas en las transcripciones.

Respecto de las búsquedas dentro del vídeo, se usa una funcionalidad específica del reproductor Paella Player (Turró, 2015) que permite la búsqueda en las transcripciones aplicando estos mismos algoritmos de búsqueda en texto completo y que se muestran en la Figura 4.

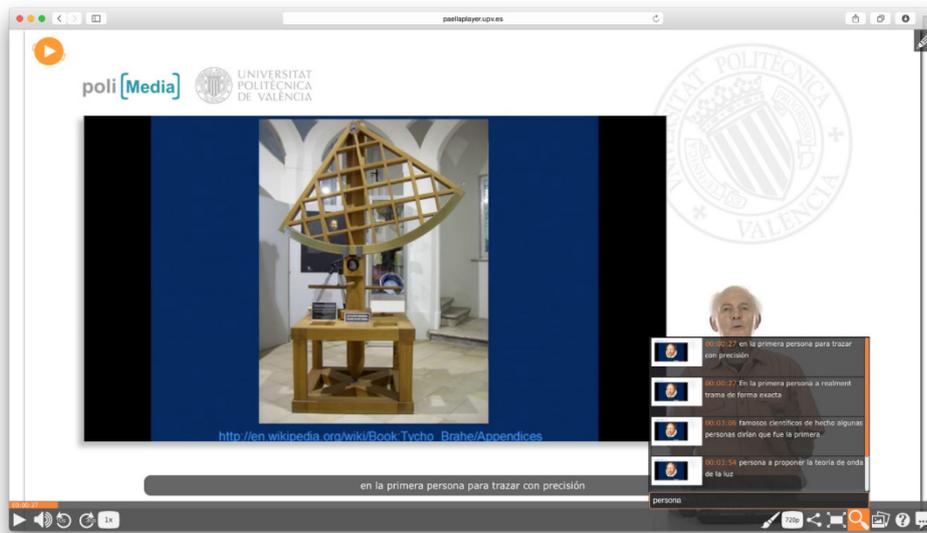


Fig. 4 Búsqueda en las transcripciones integrada en Paella Player

3. Resultados

Tras realizar el estudio descrito, consistente en el análisis de la revisión de transcripciones en francés y traducciones del español al catalán, se ha confirmado una reducción significativa del tiempo dedicado a la subtitulación de los vídeos, si lo comparamos con la generación completamente manual de subtítulos.

Así pues, obtenemos reducciones de dedicación temporal del 47% en la revisión de transcripciones en francés, en comparación a realizarlas desde cero. Lo cual supone una reducción del esfuerzo de transcripción muy significativo. Este ahorro de tiempo permite al profesor dedicarse a otro tipo de tareas docentes en lugar de emplearlo en la transcripción de los vídeos.

Por otro lado, en las revisiones de traducciones automáticas del español al catalán, la reducción es del 62%. De nuevo este ahorro de tiempo es muy significativo ya que permite realizar la traducción en casi la mitad de tiempo que si la hiciéramos desde cero.

En cualquier caso la mejora obtenida por nuestros sistemas y presentada en este artículo es significativa y muy satisfactoria. Teniendo en cuenta que en español, inglés, catalán, y traducción del español al inglés se han obtenido resultados similares a los del pasado año, el presente estudio confirma la conveniencia de realizar la revisión de transcripciones y traducciones generadas automáticamente en lugar de realizarlas desde cero, con una reducción muy significativa del tiempo dedicado a la generación de subtítulos.

4. Impacto educativo

Los subtítulos automáticos, corregidos posteriormente de forma muy eficiente, tiene un alto impacto sobre los alumnos que visitan el portal de poliMedia y los diferentes cursos asociados, que se cuantificó en más de 10.000 visualizaciones de vídeos con los subtítulos activados, lo cual supone una amplia mejora en el alcance de los vídeos docentes, en especial a personas con discapacidad auditiva o que no conocen el idioma original del vídeo.

Además, a esto hay que añadirle los beneficios de la búsqueda por palabras de vídeos en el portal, e incluso dentro del propio vídeo. Estas funciones han sido del agrado de los alumnos y usuarios de la plataforma ya que reducen de forma significativa el coste temporal de la búsqueda de contenido de interés aumentando al mismo tiempo la visibilidad de los contenidos del portal. Todo esto permite basar la mayor parte del funcionamiento del portal en la funcionalidad de búsqueda por palabras.

5. Conclusiones

En este artículo hemos evaluado la aplicación de técnicas de reconocimiento del habla y traducción automática sobre el repositorio poliMedia para la generación de subtítulos.

Más concretamente, hemos ampliado la evaluación de los sistemas automáticos de subtitulación de vídeos docentes en el marco del programa *Docencia en red* para incluir también la transcripción de francés y traducción del español al catalán.

Este sistema de subtitulación está actualmente integrado en la plataforma poliMedia de la UPV, posibilitando la revisión de la transcripción o traducción de cualquier vídeo docente disponible en ella.

La subtitulación automática de vídeos docentes ha demostrado ser una herramienta poderosa, que permite una revisión de los subtítulos de una forma mucho más eficiente respecto a generarlos desde cero.

Además, es necesario hacer hincapié en la importancia de los subtítulos en los vídeos docentes, pues aportan un valor añadido facilitando su comprensión y ampliando su alcance a personas con dificultades idiomáticas o con una discapacidad auditiva, así como permitir una localización sencilla de los contenidos del portal.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 287755



(transLectures) and ICT Policy Support Programme (ICT PSP/2007-2013) as part of the Competitiveness and Innovation Framework Programme (CIP) under grant agreement no 621030 (EMMA), and the Spanish MINECO MORE (TIN2015-68326-R) research project.

Referencias

CERDA, S., ALBERT, J., DEL AGUA TEBA, M. A., GARCÉS DIAZ-MUNIO, G. V., GASCÓ MORA, GUILLEM, GIMENEZ PASTOR, A., et al. « Translectures ». In IberSPEECH 2012-VII Jornadas en Tecnología del Habla and III Iberian SLTech Workshop (pp. 345-351). IberSPEECH 2012.

GLASS, J. R., HAZEN, T. J., CYPHERS, D. S., MALIOUTOV, I., HUYNH, D., & BARZILAY, R. (2007). « Recent progress in the MIT spoken lecture processing project ». In Interspeech (pp. 2553-2556).

VALOR MIRÓ, J. D., DE MARTOS, A. P. G., CIVERA, J., & JUAN, A. (2012). « Integrating a state-of-the-art ASR system into the opencast Matterhorn platform ». In Advances in speech and language technologies for Iberian languages (pp. 237-246). Springer Berlin Heidelberg.

VALOR MIRÓ, J. D., SILVESTRE-CERDA, J. A., CIVERA, J., TURRO, C., & JUAN, A. (2015). Efficient Generation of High-Quality Multilingual Subtitles for video Lecture Repositories. In Design for Teaching and Learning in a Networked World (pp. 485-490). Springer International Publishing.

VALOR MIRÓ, J. D., SILVESTRE-CERDA, J. A., CIVERA, J., TURRO, C., & JUAN, A. (2015). Efficiency and usability study of innovative computer-aided transcription strategies for video lecture repositories. *Speech Communication*, 74, 65-75.

PAPADOPOULOS, M., & PEARSON, E. (2012). « An intelligent system to support accurate transcription of university lectures ». In Intelligent Tutoring Systems (pp. 718-719). Springer Berlin Heidelberg.

TURRO, C., BUSQUETS, J., & DESPUJOL, I. (2015). « Grabación automatizada de clases magistrales: el proyecto vídeoapuntes de la UPV ». *Revista de Educación a Distancia* nº40 <<http://revistas.um.es/red/article/view/234411/0>> [Consulta : 30 de Abril de 2016]

TURRO, C., FERRANDO, M., BUSQUETS, J., & CAÑERO, A. (2009). « Polimedia: a system for successful video e-learning ». In Proc. of EUNIS.

VALOR MIRÓ, J. D., SPENCER, R. N., PEREZ GONZALEZ DE MARTOS, A., GARCÉS DIAZ-MUNIO, G., TURRO, C., CIVERA, J., & JUAN, A. (2014). Evaluating intelligent interfaces for post-editing automatic transcriptions of online video lectures. *Open Learning: The Journal of Open, Distance and e-Learning*, 29(1), 72-85.