

Seguimiento de ROI en videos de LSU

La traducción automática de lengua de señas es un campo reciente de la investigación en lingüística, las ciencias de la computación y el procesamiento de imágenes. Para ello, es necesario construir sistemas de aprendizaje automático que permitan realizar la tarea a partir de asociaciones y reconocimiento de patrones en los datos empleados.

La lengua de señas es propia de cada comunidad, región o país, hecho que implica que el desarrollo de soluciones locales en este campo requiere de la colecta y el curado de conjuntos de datos propios. Con este fin, Facultad de Ingeniería obtuvo desde la Dirección Nacional de Telecomunicaciones una serie de grabaciones de video de la televisión pública y de sesiones del Parlamento del Uruguay que incluyen un intérprete de lengua de señas uruguaya (LSU) en un recuadro rectangular situado próximo a la esquina inferior derecha de la pantalla, de aquí en adelante llamado ROI (*region of interest*).

Las primeras etapas de curado de los datos incluyen el recorte de la ROI del resto de la escena. En aquellos videos donde la ROI es fija a lo largo del tiempo es sencillo aislarla de forma manual a partir de cuatro coordenadas 2D, por ejemplo las cuatro esquinas del rectángulo. No obstante, una parte de los vídeos obtenidos, presentan una ROI móvil en la escena, tratándose en general de un movimiento continuo. En este proyecto se propone como objetivo tomar estos últimos casos y realizar un seguimiento de la ROI a lo largo del tiempo. Se espera que luego de realizar el recorte de la ROI, el recuadro del señante permanezca estable.

Tutores (y mail de contacto):

Ariel E. Stassi, astassi@cup.edu.uy

Gregory Randall, randall@fing.edu.uy