Automated Illustration of Text to Improve Semantic Comprehension

Mirchandani, Suvir

Urban, Konrad

Over a million Americans suffer from aphasia, a disorder that severely inhibits language comprehension. Medical professionals suggest that individuals with aphasia have a noticeably greater understanding of pictures than of the written or spoken word. Accordingly, we design a text-to-image converter that augments lingual communication, overcoming the highly constrained input strings and predefined output templates of previous work. This project offers four primary contributions. First, we develop an image processing algorithm that finds a simple graphical representation for each noun in the input text by analyzing Hu moments of contours in images from The Noun Project and Bing Images. Next, we construct a dataset of 700 human-centric action verbs annotated with corresponding body positions. We train support vector machines to match verbs outside the dataset with appropriate body positions. Our system illustrates body positions and emotions with a generic human representation created using iOS's Core Animation framework. Third, we design an algorithm that maps abstract nouns to concrete ones that can be illustrated easily. To accomplish this, we use spectral clustering to identify 175 abstract noun classes and annotate these classes with representative concrete nouns. Finally, our system parses two datasets of pre-segmented and pre-captioned real-world images (ImageClef and Microsoft COCO) to identify graphical patterns that accurately represent semantic relationships between the words in a sentence. Our tests on human subjects establish the system's effectiveness in communicating text using images. Beyond people with aphasia, our system can assist individuals with Alzheimer's or Parkinson's, travelers located in foreign countries, and children learning how to read.

Awards Won:

Fourth Award of $500

European Organization for Nuclear Research-CERN: Third Award $500