**Docket #:** S21-205

# Txt2Vid: Compressing Talking-head Videos to Text

Researchers at Stanford University, UCSB and MIT have invented a novel video compression pipeline, called Txt2Vid, which substantially reduces data transmission rates by compressing webcam videos ("talking-head videos") to a text transcript. The text transcript can be transmitted and decoded on the recipient's end into a realistic reconstruction of the original video using recent advances in deep learning based voice cloning and lip syncing models.

This generative pipeline achieves two to three orders of magnitude reduction in the bitrate as compared to the standard audio-video codecs (encoders-decoders), while maintaining equivalent quality-of-experience based on a subjective evaluation by users (n=242) in an online study. The Txt2Vid framework opens up the potential for creating novel applications such as enabling audio-video communication during poor internet connectivity, or in remote terrains with limited bandwidth. Additionally, the text transmitted has the potential to be translated into different languages or decoded into different voices and faces to create a custom end user experience for teaching and more.

## Applications

- Video compression with extremely low bit rate for good quality communication in areas of low internet connectivity
- Video conferencing with text as the compression format, enabling real-time language translation and/or hybrid communication (i.e., where either typed messages or spoken messages can be transmitted and then reconstructed into voice with the same experience for the end user)
- Transmission of pedagogical content for remote learning and online instruction, with the ability to reconstruct compressed content into more engaging formats (e.g., generate a math lesson in which a favorite movie character is the

teacher)

## Advantages

- Audio-video communication compression of two-to-three orders of magnitude reduction with similar quality-of-experience (500-1000x compression)
- Functions with bitrates as low as 100bps
- Extremely low bandwidth requirements make communication accessible in areas of poor Internet availability
- Flexibility to operate as a video player platform, a video streaming platform, or a real-time communication platform

## Publications

- Tandon, Pulkit, et al. "Txt2Vid: Ultra-Low Bitrate Compression of Talking-Head Videos via Text." arXiv preprint arXiv:2106.14014 (2021).

## Patents

- Published Application: 20220417291
- Issued: 12,567,398 (USA)

## Innovators

- Pat Pataranutaporn
- Pulkit Tandon
- Shubham Chandak
- Yimeng Liu
- Misha Sra
- Anesu Mapuranga
- Tsachy Weissman
- Pattie Maes

# Licensing Contact


## Imelda Oropeza

Senior Licensing Manager, Physical Sciences

Email