Pictures of Jap Girls in Synthesis

Eyal Gruss
eyalgruss@gmail.com

Ayelet Sapirshtein

Vered Heruti
Hamidrasha, Beit-Berl College
Kfar Saba, Israel

Abstract
Pictures of Jap Girls in Synthesis\(^1\), is a machine that creates live, visual poetry (Figure 1). The machine reads, or listens to poetry in a variety of languages. It synthesizes in real-time, visual images illustrating its understanding and interpretation of the text. As an interactive installation, poets can perform live poetry, with the machine simultaneously translating it into visual imagery, allowing a synesthetic and a more universal experience (Figure 2).

Introduction
As poets and digital artists our main forms of expression are based on text, either human words or computer code. We are interested in how to augment our words, give them a place in space, and realize them in other mediums. How can we visualize poetry? How can we define and create visual poems? How can we allow speakers of different languages to share a universal poetic experience? (Figure 3) How can we create a synesthetic experience for poetry, involving multiple senses, and combining words, sounds, space, and visual imagery into a unified experience? Can we automate this process? We call it synthetic synesthesia.

To quote Mitchell: "Word and Image" seems to be better understood as a dialectic trope... It is a dialectical trope because it resists stabilization as binary opposition, shifting and transforming itself from one conceptual level to another, and shuttles between relations of contrariety and identity, difference and sameness. Mitchell's statement about the relations between image and text can also be connected to this project. Mitchell (1996) formulates these relations through one welded word – "imagetext". Mitchel objects the perception of image and text as binary opposition concepts and claims that there are similarities as well as differences between the two terms.

On the technical side, we are using state-of-the-art deep learning algorithms for speech-to-text\(^3\), language translation\(^4\), and the AttnGAN\(^5\) for text-to-image synthesis (Xu et al. 2017). While the results are far from photorealistic, we get images of surrealistic quality, containing colors, shapes, textures, patterns and objects related to the words and their context, and creating a poetic experience. This project was created in a 30-hour hackathon, at Geekcon-Art\(^6\) Hamidrasha 2018.

References


\(^1\) "Pictures of Jap girls in synthesis" is a quote from David Bowie's Ashes to Ashes (1980).


\(^3\) https://cloud.google.com/speech-to-text

\(^4\) https://cloud.google.com/translate

\(^5\) https://github.com/taoxugit/AttnGAN

\(^6\) https://geekcon.org/art2018

Figure 1: Generated image for the phrase from David Bowie's Ashes to Ashes, for which the work is named after.
Figure 2: Savyon, an Israeli poet, performing one of her poems in Hebrew, live with the interactive installation.

Figure 3: An example of verse in Farsi. While we do not understand the text, the visual image may be universal.