Virtual Singers Empowered by Machine Learning

Siyao Li
City University of Macau, Macao

Haoyu Liu
City University of Macau, Macao

Pi-Ying Yen
Macau University of Science and Technology, Macao

INTRODUCTION

The field of machine learning has been developing over decades, and machine learning techniques are utilized in many areas, such as medicine (Obermeyer & Emanuel, 2016; Rajkomar et al., 2019), natural language processing (Hirschberg & Manning, 2015), and pattern recognition (Fang et al., 2016). With the theory and practice of machine learning becoming more mature, its application areas have also been expanded (Wang et al., 2017).

Empowering virtual singers is one of the newer applications of machine learning. The basis of virtual singers is synthesis software that can create life-like or concert-like voices with computers. Inputting tones and lyrics, producers can synthesize songs through software, and the machine learning technology is used to produce three-dimensional images of virtual singers.

The virtual singer industry was first introduced in the 2000s, when it was a big hit and created history. With the advance of technology during the past twenty years, the industry has been upgraded. Although virtual singers cannot yet overtake traditional singers, they demonstrate a promising and powerful future. According to Chia (2021), China has the biggest potential audience for virtual singers worldwide, and an estimated 390 million people are watching virtual singers. The accompanying animation industry, including TV series and comics, achieved a market value of $35 billion in 2020. Virtual singers have come into people's daily lives. Nescafe and KFC use songs by Luo Tianyi, one of the rising superstars in the virtual singer industry, in advertising campaigns, and Harper's Bazaar put Luo Tianyi's image on the cover of its Chinese edition.

BACKGROUND

One of the most successful virtual singers globally is Hatsune Miku from Japan, who represents a benchmark in the virtual singer area. She won many fans after her debut in 2007 (Liang, 2020), and the song "Tell Your Word", featuring her, ranked first on Japan's iTunes in 2012. Notably, Hatsune Miku is also the first virtual singer to use holographic projection technology to hold concerts: Using a computer-generated voice and 3D graphics, she performs like a real person on the stage. Hatsune Miku performed her first concert in 2009 at Japan's Saitama Super Arena (Corey, 2011). As her influence has increased

V

all over the world, she has started to hold concerts overseas. In 2018, she held her first live concert tour in Europe, Hatsune Miku Expo 2018 Euro.

In recent years, more and more virtual singer groups have been emerging in Korea. Aespa, introduced by SM Entertainment and debuting in 2020, is the first girl group to bring the machine learning concept to K-pop (Rowley, 2020). Their latest single song "Next Level" has caught enormous attention. The song was released on 17 May 2021 and entered at Number 33 on the Billboard K-Pop 100 chart, peaking at Number 2 and staying in the top five for weeks (Billboard, 2021). If Hatsune Miku marked the successful emergence of the virtual singer industry, then Aespa has continued this success and is pushing the industry to new heights.

Asia has been the most developed region for the virtual singer industry so far (Chia, 2021). Outside Japan and South Korea, China is also a market of tremendous potential. The latest rising virtual singer is Luo Tianyi, the first virtual singer to make a profit in China. She has attracted hordes of young fans in China, as indicated by her 5 million followers on Weibo. Her songs appear in various advertisements, and the Chinese state-controlled broadcaster has invited her to perform in the New Year's Gala. Luo Tianyi is undoubtedly a great hit in China's virtual singer industry.

Beyond Asia, virtual singers are also gaining popularity in Europe and America. Lil Miquela is a Los Angeles based character who performs rhythm and blues (R&B), electronic music, and popular music. Miquela released her first single song "Not Mine" in August 2017, and one more piece "Over You" in September 2017 (The New York Times, 2019). She released a further 11 songs from 2018 to 2020. By January 2022, Miquela has around three million Instagram followers, and over 80000 streams on Spotify each month. France also has a famous virtual singer named ALYS. ALYS not only sings in French but in Japanese as well. She started her career in March 2014 and was designed as a 21-year-old with a long blue braided ponytail. Her voice was developed by a collaboration between VoxWave and Plogue Art & Technologies. ALYS had her first show, Reve de Machine, at the Trianon in Paris in December 2016 and successfully attracted many fans coming to celebrate Christmas together.

From a technical perspective, the use of machine learning for the music industry has been investigated widely in academia. Readers interested in the technical papers mentioned below may refer to the additional readings.

Nachmani and Wolf (2019) propose an unsupervised method based on a single convolutional neural network (CNN) encoder and a single conditional WaveNet decoder. Their method can create natural voices and high-quality singing from target speakers. Zhang et al. (2020) introduce a singing voice conversion model DurIAN-SC that can produce high-quality voices recognizable as target singers by only adopting their speech data, while Zeng et al. (2021) introduce a large-scale pre-trained model called MusicBERT for music understanding.

Gu et al. (2021) develop the ByteSing system, a Chinese singing voice synthesis (SVS) system that adopts duration-allocated Tacotron-like acoustic models and WaveRNN neural vocoders. They prove by both objective and subjective tests that their proposed SVS method can improve a songs' pitch and spectrogram prediction accuracy, and produce highly natural and high-fidelity songs. In particular, subjective evaluation demonstrates that the system can achieve more than 80 percent of human singing levels, which suggests the effectiveness of their proposed SVS system.

Tae et al. (2021) develop MLP Singer, a parallel Korean singing voice synthesis system, which solves the problem of slow inference speed. MLP Singer consists of multi-layer perceptrons (MLPs). Compared to autoregressive conditional generative adversarial network (GAN) based systems, MLP Singer performs better in terms of audio quality and synthesis speed, especially inference speed.

Related Content

Challenges and Chances of Classical Cox Regression
Mengying Xiaand Leigh Wang (2023). *Encyclopedia of Data Science and Machine Learning (pp. 2438-2449).*
www.irma-international.org/chapter/challenges-and-chances-of-classical-cox-regression/317683

Development of a Charge Estimator for Piezoelectric Actuators: A Radial Basis Function Approach
Morteza Mohammadzaheri, Mohammadreza Emadi, Mojtaba Ghodsi, Issam M. Bahadur, Musaab Zarogand Ashraf Saleem (2020). *International Journal of Artificial Intelligence and Machine Learning (pp. 31-44).*
www.irma-international.org/article/development-of-a-charge-estimator-for-piezoelectric-actuators/249251

Integrated Regression Approach for Prediction of Solar Irradiance Based on Multiple Weather Factors
Megha Kambleand Sudeshna Ghosh (2021). *International Journal of Artificial Intelligence and Machine Learning (pp. 1-12).*
www.irma-international.org/article/integrated-regression-approach-for-prediction-of-solar-irradiance-based-on-multiple-weather-factors/294105

Comparison of Brainwave Sensors and Mental State Classifiers
Hironori Hiraishi (2022). *International Journal of Artificial Intelligence and Machine Learning (pp. 1-13).*
www.irma-international.org/article/comparison-of-brainwave-sensors-and-mental-state-classifiers/310933

Methodology and Application of Information Technology for Carbon-Based Nano-Composites
Vibhooti Narayan Mishra, Divya Pratap Singh, Shweta Singh, Ashish Kumar Singhand Savendra Pratap Singh (2024). *Methodologies, Frameworks, and Applications of Machine Learning (pp. 52-65).*
www.irma-international.org/chapter/methodology-and-application-of-information-technology-for-carbon-based-nano-composites/342648