

「初音ミク」はなぜ注目されているのか

後藤 真孝

産業技術総合研究所



歌声合成, 歌声情報処理, 初音ミク, CGM, ニコニコ動画

1. はじめに

2007年8月31日に、歌声合成ソフトウェア「初音ミク」⁽¹⁾がクリプトン・フューチャー・メディア(株)から発売され、活発な創作活動と連鎖反応の源泉となった。このソフトウェアには音楽のクリエイター以外にも関心を持つさまざまな工夫がなされており、「初音ミク」という名前の仮想の歌手をイメージしたアニメ風のキャラクターのイラストがソフトウェアパッケージの箱に描かれ、年齢やプロフィールなども設定されていた。その歌声合成の手軽さと特徴的な歌声、キャラクターなどが人気を呼び、多くの人々が初音ミクやその周辺の技術・文化・現象に興味を持って多彩な活動をしたことで、未来を切り拓く画期的な素晴らしい現象が起き続けている。

この「初音ミク」のソフトウェアを購入すれば、誰でも自分の楽曲を歌わせることができるため、アマチュアを中心とした多くのクリエイターによって、次々と楽曲が発表されてきた。そして、それ以上にこの現象が特徴的なのは、発売元が的確に権利を解放したことで、多様な連鎖反応が起きたことである。例えば、キャラクターや楽曲のイメージのイラストを描いたり、他のクリエイターが作った楽曲に対して新たな映像を付けて動画にしたり、キャラクターのイラストをスライドショーのように自分の曲に付けて動画にしたりする二次創作(派生した著作物の創作)が盛んに取り組まれた。こうした一般の人々が活発に創作活動を繰り返し広げて内容が生成されていくメディアはCGM(Consumer Generated Media, 消費者生成メディア)と呼ばれ、クリエイターだけでなく視聴者も増え続けている。

この初音ミクの登場が「人間の歌声でなければ聴く価値がない」という旧来の価値観を打破し、「合成された歌声がメインボーカルの楽曲を積極的に楽しむ文化」が世界で初めて日本に誕生した。これは初音ミク発売以前には考えられなかった快挙である。歌声合成技術の研究者自身ですら、その用途として、人間の歌声に合わせるコーラスや、人間が吹き替える前の制作段階の仮歌としての使用を想定していた。しかし今日では、歌声合成技術が普及することは歴史的必然であると確信できる状況となった⁽²⁾。

本稿では、こうした過去5年間の新しい現象について解説し、なぜこれほどまでに異なる分野の多くの人々が、そ

してさまざまな研究者が、歌声合成ソフトウェアでかつキャラクターでもある初音ミクを発端としたCGM現象に注目しているのかを議論する。初音ミク現象の発展では二次創作が本質的であるが、本稿も筆者自身による「初音ミク、ニコニコ動画、ピアプロが切り拓いたCGM現象」⁽³⁾の二次創作として位置づけられ、初音ミク現象や技術についての初歩的な解説などを加筆して構成されている。また、今日では初音ミク以外にもさまざまな声質や原理の歌声合成ソフトウェアが流通しており、下記の解説にはそれらにも当てはまる記述が多いが、便宜上、「初音ミク」を用いる。

2. 歌声合成独特の表現を含む音楽文化の誕生

歌声合成技術による歌声は、単なる人間の歌声の代用だけではなく新たな音楽表現を生み出し始めている。これは、楽音合成技術(シンセサイザ)による楽器音が、単なる既存の楽器の再現だけではなく新たな音楽表現を生み出してきたのと同様である。ただし楽器音とは異なり、歌声合成の場合には技術自体の進歩だけでは不十分で、キャラクターという身体性を合成歌声に持たせたことがその普及のために本質的役割を果たした。

2.1 歌声合成技術 VOCALOID

「初音ミク」はヤマハ(株)の歌声合成技術「VOCALOID 2」^{(4)~(6)}に基づいた製品である。VOCALOID 2では、メロディの音符を画面上で並べて歌詞を入力すると、人間の歌声を収録したデータベース「歌手ライブラリ」に基づいて歌声を出力することができる。歌声ライブラリを変えれば異なる個性(声質)を持つ歌声を合成でき、さまざまな企業がそうした製品を出している。従来よりも自然な歌声の合成が可能になり、ユーザインタフェースが改善され、声の高さや大きさなどをパラメータとしてユーザが細かく調整できることが、人気が出る技術的な要因となった。

歌声合成技術自体は、1950年代後半からさまざまな研究開発がなされ⁽⁷⁾、計算機や信号処理の進歩とともにその品質・機能も向上し続けてきた⁽⁶⁾。例えば、録音した歌声の断片(素片)を変形しながらなめらかに接続して合成する方式や、人間の発声機構を物理モデルでシミュレーションして合成する方式、歌声の生成過程を声帯振動(ソース、励振源)と声道共鳴(フィルタ)に分けてモデル化するソースフィルタモデルに基づく合成方式、歌声の周波数

スペクトルなどの変化を隠れマルコフモデル (HMM) でモデル化して合成する方式などが提案されている。商用の歌声合成技術は、「VOCALOID 2」以前にも、NTT AT (株) の歌声合成エンジン「ワンダーホルン」やヤマハ (株) のフォルマントシンキングプラグインボード「PLG-100SG」、古くは日本電気(株) のパソコン「PC-6601」の歌声合成コマンドなどが存在していた。

ヤマハ(株) の「VOCALOID」は、録音した歌声の素片を接続する合成方式に基づいており、2003年に発表されて2004年に最初の製品が発売された。2007年に発売された「VOCALOID 2」までは、ある音素から別の音素への変化部分のような2音素の連鎖を素片としていたが、2011年に発売された「VOCALOID 3」では3音素の連鎖も扱えるようになってきている⁽⁵⁾。歌声合成では話し声の音声合成と異なり、言葉を伝えるだけでなく、楽譜に応じてさまざまな音高で異なる長さの音符に対応した歌声を合成する必要がある。そこで、音高が異なれば声色も異なるため、複数の異なる音高の素片を用意し、伸ばした母音を高品質に合成する工夫がなされている。さらに歌声の歌詞が適切なタイミングで聞こえることも必須であり、例えば日本語の「さ」(/s a/) のような「子音+母音」を合成する場合には、音符の開始時刻に母音が鳴るように子音はその前にずらすタイミング調整機構を持っている⁽⁶⁾。

2.2 キャラクタという身体性が果たす本質的役割

いくら歌声合成技術が進歩しても、それだけでは歌声合成技術がメインボーカルとして広く利用されるのは難しかった。「初音ミク」が画期的だったのは、女性声優の声のデータベースを高品質に収録したことに加え、それとイメージが合致するキャラクタという身体性を合成歌声に持たせたことである。このクリプトン・フューチャー・メディア(株)の本質的な貢献により、初めて、「誰が歌っているのか」、「誰に歌わせたいのか」が明確になった。

これにより、初音ミクの発売直後から、初音ミクにしか歌えない曲が次々と発表され始めた。例えば、キャラクタ側の視点から、クリエイターとの関係性を感情表現した歌詞(歌声合成の立場でどういうふうに歌わせてほしいかを述べた歌詞^{*1}や、実在せず歌うことしかできないと語る歌詞^{*2}など)が登場した。人間には歌うのが困難な高速な歌詞や広い声域の楽曲なども登場した。これにより、初音ミクのために曲を作る文化が誕生した。ここで、クリエイターの側には「このキャラクタ(歌手)に歌わせたいから曲を作る」という強い動機が生まれ、視聴者の側には「このキャラクタ(歌手)の歌だから聴きたい」という積極的な楽しみ方が生まれたことが重要である。ほかにも、歌声合

成ソフトウェアのキャラクタ(声質)が、さまざまな企業の貢献により「鏡音リン」、「鏡音レン」、「がくっぽいど」、「巡音ルカ」、「Megpoid」、「氷山キヨテル」、「歌愛ユキ」、「SF-A2 開発コード miki」などと増えていくにつれて、表現も多彩になり、キャラクタ間の関係性を表現した歌詞なども登場している。

こうして、人間の代わりに歌声合成を用いるのではなく、歌声合成でなければならぬ表現を含む音楽文化が誕生した。

2.3 積極的に歌声合成を使うクリエイターの増加

クリエイターにとっては、歌声合成技術を用いることで、自分の思いどおりの表現を追求しやすくなる。人間の歌手に自分の曲を歌ってもらおうとすると、その機会を得るだけで大変であり、自分の望む声で思いどおりの表現をしてくれるとは限らない。それが、初音ミクの歌声で、納得のいく表現が得られるまで何度でも修正しながら再現性高く合成できるのは、大きな利点である。結果的に、初音ミクという同一歌手の曲が数万曲生まれた。そこから淘汰されて多くの視聴者に聴かれるようになった有名曲は、ポピュラリティの高い良質な作品である。そうした作品がさらに人気を呼んで、初音ミクに歌わせれば聴いてもらえる可能性が高まり、彼女を歌わせたいという動機をクリエイターに与えるという連鎖反応が起きている。これは、人間の歌手ではあり得ない現象である。

3. N次創作の連鎖反応

このように「初音ミク」では歌声だけでなく、キャラクタも重要であるが、それを誰でも合法的に非商用利用できるようにライセンスが整備されたことで、創作活動の連鎖反応が持続した。さらに、創作を促進し連鎖反応が盛り上がる場として「ニコニコ動画」が貢献したことで、人気は加速した。

3.1 権利を解放するためのライセンスの重要性

初音ミクの歌声は自分の作品で自由に使えても、そのキャラクタのイラストは、通常であれば権利的に自由には使えない。歌声合成ソフトウェアは楽器の位置づけだが、キャラクタの再利用は著作権の制約を受けるからである。そこでクリプトン・フューチャー・メディア(株)は、「初音ミク」発売の3か月後にいち早く、「キャラクター利用のガイドライン」^{*3}を宣言し、二次創作を応援するCGM型コンテンツ投稿サイト「ピアプロ」^{*4}を2007年12月3日に立ち上げた。そして「ピアプロ・キャラクター・ライセンス(PCL)」^{*5}を発行することで、個人の創作意欲を引

^{*3} http://piaopro.jp/license/character_guideline

^{*4} <http://piaopro.jp/>

^{*5} <http://piaopro.jp/license/pcl/summary>

^{*1} <http://www.nicovideo.jp/watch/sm1050729>

^{*2} <http://www.nicovideo.jp/watch/sm1274898>

き出し、誰でも合法的に非商用利用することを可能にした⁽¹⁾。つまり、積極的に権利を開放したのである。

これはとても重要で、このおかげでアマチュアを中心としたクリエイターは権利問題を心配せずに、初音ミクのキャラクターを使って安心して創作し、発表できるようになった。こうした権利の開放により自由な環境が意図的に整えられた意義は大きく、ほかの歌声合成ソフトウェアにも波及して、この現象の発展を決定づけた。

この権利開放の結果、初音ミク発売元は数パターンのイラストしか公表していないにもかかわらず、アマチュアを中心としたさまざまなクリエイターが描いた十数万通りの初音ミクのイラストが創作された。また、キャラクターを中心とした動画制作を容易にする三次元CG（コンピュータグラフィクス）ソフトウェア「MikuMikuDance（MMD）」^{*6}等も無償公開されて大きな貢献を果たし、音楽に連動した動画作品が多数生まれた。

しかし、それはあくまでキャラクターを非商用利用できるだけで、他のクリエイターの作品から二次創作する権利までは含まれない。その権利を解放することを意図して立ち上げられ、重要な役割を果たしたのが、CGM型コンテンツ投稿サイト「ピアプロ」である。クリエイターがピアプロに自分の音楽、イラスト、テキスト（歌詞）、3Dモデル（三次元形状モデル）をアップロードする際には、他のクリエイターが非営利に利用することに必ず同意している。つまりピアプロ上の作品は、クリエイターに感謝の意を示して使用報告さえすれば、安心して二次創作で利用できるのである。さらに派生作品の親子関係も表示されるので、多数の二次創作作品に利用された親作品であることが、人気のバロメータにもなっている⁽¹⁾。

3.2 クリエータのやる気を引き出すニコニコ動画

初音ミクの楽曲や動画が多数投稿されている動画コミュニティサービス「ニコニコ動画」^{*7}は、初音ミク現象を支え、創作を促進している重要な社会装置である。ニコニコ動画は、YouTubeやUstreamのような従来の動画共有サービスとは異なる先進的な機能を数多く持ち、それがさまざまなCGM現象の土壌となっている。動画共有サービスでは、そこに集う人々は動画の視聴を主目的としているのに対し、動画コミュニティサービスでは、そこに集う人々はコミュニケーションや共同視聴体験を主目的としている。動画の視聴を共通体験とした上での、作品内の時刻に同期したコメント（時刻同期コメント）などによるメタレベルのコミュニケーションを、視聴者同士が、あるいは動画のクリエイターと視聴者間で、楽しんでいるのである。

初音ミクなどの楽曲や動画の作品を発表する場として、

このニコニコ動画は最適であった。動画作品を投稿すると、時刻同期コメントが視聴者から得られ、クリエイターの励みになるからである⁽⁸⁾。作品全体に対する通常のコメントに比べ、時刻同期コメントはその内容が質的に異なり、その瞬間ごとの感情をより多く表現する傾向がある。そのため、「ここが良い」、「ここをこう変えては？」というような感想・提案も含めて、創作活動にさまざまなフィードバックが得られる環境になっており、クリエイターのやる気が引き出されやすい。さらに、現在の歌声合成技術を用いた楽曲では歌詞が聞き取りにくい場合があるが、動画中の字幕やコメントで歌詞が表示されると聞き取りやすくなる利点もある。

このニコニコ動画の場で、クリエイターと視聴者の創意工夫の相乗効果により多様な二次創作が生まれた。例えば、歌声合成による元楽曲を人間が歌ったり、演奏したり、踊ったりする実写動画や、その実写の踊りをさらにCGキャラクターで再現した動画なども発表されている。既存動画を切り貼りして新しい動画を作る活動も活発であり、例えば、多数の動画からランキングを作ったり、音楽に合う映像を断片的に切り貼りして新たな動画を作ったりしている。多くの人に見てもらえたり、コメントで感想がもらえたりするだけで、クリエイターはさらに嬉しくなって創作の連鎖が起きていく。

こうした創作が創作を呼び起こす連鎖反応（集団的創造現象）は「N次創作」と呼ばれる⁽⁹⁾。その本質にはコミュニケーションの要素もあり、「この作品に対して、私はこういう作品を二次創作して応えよう」といった掛け合いのような創作の連鎖が起きている。そして、ニコニコ動画と前述のピアプロでは、再利用可能なコンテンツを集め、コミュニケーションの支援もすることで、こうしたN次創作をさらに後押ししている。その過程で複数の作品の良いところが掛け合わされて良質な作品が生まれ、さらに人気が高まっていった。

4. 初音ミク現象の広がり

初音ミクの楽曲や動画を楽しむ視聴者のために、物理空間での大規模なライブコンサートも国内外で開催され、その先進性はこれまで初音ミクに興味を持っていなかった人々にも大きな衝撃を与えた。歌声合成技術はまだ発展途上で今後も性能は向上し続け、新たな未来を切り拓いていく。

4.1 初音ミクのライブコンサート

ライブコンサートの歌手が人間ではなく、初音ミクというキャラクターであっても、数千人規模の観客を動員する大型のコンサートが成功するようになったのは画期的である。初音ミク主演のライブコンサートは、国内では2009年8月31日、2010年3月9日、2011年3月9日、2012

^{*6} <http://www.geocities.jp/higuchuu4/>

^{*7} http://www.nicovideo.jp/video_top/

年3月8日・9日に東京で、2011年8月16日・17日に札幌で開催された。国外では、2011年7月2日にLos Angelesで5000人以上の観客を集め、2011年11月11日にSingaporeで2500人以上の観客を集めた。特に2012年3月の東京公演は2日間で1万人が来場し、国内外34か所（日本24か所、台湾7か所、香港2か所、上海1か所）の映画館に生中継（パブリックビューイング）されて大規模に開催された。

メインボーカルの歌手だけが歌声合成と3次元CGキャラクターで表現され、彼女以外のステージ上の楽器演奏者と観客は全員人間という特殊なコンサートであり、その様子を初めて観ると大きな衝撃を受ける。歌手の声と映像はコンサート前に静的に用意されているにもかかわらず、観客が大きな声援を送らずにはいられないのも興味深い。

しかも国外でのコンサートや生中継では、演奏された楽曲は日本語の歌詞のままであり、それでも世界に通用する音楽となったことは本当に素晴らしい。アマチュアを中心としたクリエイターが制作した楽曲ではあるが、上記の淘汰を経たポピュラリティの高い音楽であることも寄与している。

4.2 歌声合成の未来

初音ミク現象は単なるブームではなく、持続発展する文化であり、歌声合成技術が普及することは歴史的必然である⁽²⁾。楽音合成技術がポピュラー音楽制作で不可欠となったのと同様、歌声合成技術もいつの日か不可欠になる。そして楽音合成と同様に歌声合成も、独自の表現を生み出し、多様な音楽の創造に寄与していくはずである。特に、^{のど}喉という物理的な制約のない歌声がどのような表現を生み出していくのかは興味深い。また、技術の発展によって表現手段を得た一般の人々が、アマチュアクリエイターとしてCGM現象に参加する流れは、今後ますます発展していくと考えられる。

歌声合成技術が発展すると、人間の歌手はいらなくなるのだろうか。筆者はそうは思わない。なぜなら、我々人類

は歌うことをやめないからである。人は、強制されて歌うわけではなく、歌いたいから歌っている。また聴き手も、この人（自分の子供、友だち、好きな歌手）が歌うから、と思って聴く。そこに初音ミクという選択肢が増えるに過ぎないのである。

しかも、歌声合成技術が普及すれば、歌で自己表現するクリエイターが増えていく可能性がある。もし歌手が「歌による表現者」を意味するとすれば、そうした歌声合成技術の使い手も、広義の「歌手」だと呼べる時代が来るのかもしれない。

5. おわりに

初音ミクは、ニコニコ動画やピアプロとともに、新たな未来を切り拓き続けている。単に「歌やキャラクターを好む人々のための趣味的な世界」だと見なすと、この本質を見誤る。この現象では、日本の技術・社会・文化のすべての強みが結集したと思えるような明るい未来が築かれつつあり、一度、そのすごさを知ってしまうと、ほかの人に伝えずにはいられない気持ちになってしまう。それゆえ、これだけ多方面に支持者がいるのである。特に、本現象に魅了されている研究者には、これまでのキャラクターを中心とした文化には馴染みが薄く、初音ミクで初めて議論している人々も多い。

音楽情報処理分野の部分領域である「歌声情報処理」⁽¹⁰⁾では、近年も活発に歌声に関連した研究開発が続いており、またさまざまな商用利用も進んでおり、それらがさらに新たな現象を生んでいくことが期待される。技術の切り拓く未来を描く手段の一つに、SF（サイエンスフィクション）小説がある。すでに、SF小説「南極点のピアピア動画」⁽¹¹⁾では、この日本発の現象の本質が切り拓く明るい未来の一つの形が描かれている。現実世界の我々は、どのような明るい未来を切り拓いていくことになるのかが楽しみである。

文献

- (1) 伊藤博之:「初音ミク as an interface」, 情報処理 (情報処理学会誌), Vol.53, No.5, pp.477-482 (2012)
- (2) 後藤真孝:「音楽情報学」, 情報処理 (情報処理学会誌), Vol.51, No.6, pp.661-668 (2010)
- (3) 後藤真孝:「初音ミク、ニコニコ動画、ピアプロが切り拓いたCGM現象」, 情報処理 (情報処理学会誌), Vol.53, No.5, pp.466-471 (2012)
- (4) H.Kenmochi and H.Ohshita: "VOCALOID—Commercial Singing Synthesizer Based on Sample Concatenation", Proc. of Interspeech 2007 (2007)
- (5) 剣持秀紀:「歌声合成の過去・現在・未来:「使える」歌声合成のためには」, 情報処理 (情報処理学会誌), Vol.53, No.5, pp.472-476 (2012)
- (6) 剣持秀紀:「歌声合成システム VOCALOID の開発」, システム/制御/情報, Vol.56, No.5, pp.244-248 (2012)
- (7) P.R.Cook: "Singing Voice Synthesis: History, Current Work, and Future Directions", Computer Music Journal, Vol.20, No.3, pp.38-46 (1996)

- (8) 戀塚昭彦:「ニコニコ動画の創造性: 動画コミュニティサービス「ニコニコ動画」の5年間」, 情報処理 (情報処理学会誌), Vol.53, No.5, pp.483-488 (2012)
- (9) 濱野智史:「ニコニコ動画はいかなる点で特異なのか:「擬似同期」[N次創作]「Fluxonomy (フラクソノミー)」」, 情報処理 (情報処理学会誌), Vol.53, No.5, pp. 489-494 (2012)
- (10) 後藤真孝, 齋藤毅, 中野倫靖, 藤原弘将:「歌声情報処理の最近の研究」, 日本音響学会誌, Vol.64, No.10, pp.616-623 (2008)
- (11) 野尻抱介:「南極点のピアピア動画」, 早川書房 (2012)



後藤 真孝

ごとう・まさたか

1998年早稲田大学大学院博士後期課程修了。現在、産業技術総合研究所情報技術研究部門上席研究員兼メディアインタラクション研究グループ長。音楽情報処理を20年間研究し、能動的音楽鑑賞サービス <http://songle.jp> や歌声合成技術「ほかりす」等を実現。博士 (工学)。