

# 1 初音ミク，ニコニコ動画，ピアプロが切り拓いた CGM 現象

後藤真孝（産業技術総合研究所）

## 初音ミク，ニコニコ動画，ピアプロが切り拓いた CGM 現象の先進性

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でもある「初音ミク」<sup>☆1</sup>を発端とした CGM (Consumer Generated Media, 消費者生成メディア) 現象に注目し、議論しているのか。それは、そこに情報処理技術の力で切り拓かれつつある新たな未来があるからである。単に「歌やキャラクタを好む人々のための趣味的な世界」だとみならず、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。

この現象を支え、創作を促進する重要な社会装置の1つに、初音ミクの楽曲や映像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」<sup>☆2</sup>がある。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、時刻同期コメント等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。

さらに、「初音ミク」発売の3カ月後にいち早く、「キャラクター利用のガイドライン」<sup>☆3</sup>を発売元が宣言し、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」<sup>☆4</sup>を立ち上げたのも重要であった。これにより、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」等のキャラクタ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。

## 何がすごいのか

歌声合成ソフトウェア「初音ミク」が発売された2007年8月以降、実際にどんなすごいことが起きてきたのかを解説する。

### ■ 楽曲のメインボーカルが歌声合成になった

「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

☆1 <http://www.crypton.co.jp/mp/pages/prod/vocaloid/cv01.jsp>  
 ☆2 [http://www.nicovideo.jp/video\\_top/](http://www.nicovideo.jp/video_top/)  
 ☆3 [http://piapro.jp/license/character\\_guideline](http://piapro.jp/license/character_guideline)  
 ☆4 <http://piapro.jp/>

めて日本に誕生した。これは 2007 年以前には考えられなかった快挙であり、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。それを大きく変えたのが、歌声合成技術「VOCALOID2」<sup>☆5</sup>に基づく「初音ミク」である。

ヤマハ（株）の歌声合成技術「VOCALOID2」（最新版は VOCALOID3）では、メロディの音符と歌詞を入力すると、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（本特集<sup>1</sup>の剣持氏の記事参照）。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、声の高さや大きさ等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。

2007 年 8 月 31 日に発売された「初音ミク」は、この「VOCALOID2」を使用したクリプトン・フューチャー・メディア（株）の歌声合成ソフトウェアである。初音ミクでの同社の代表的な貢献は、女性声優の声のデータベースを高品質に収録したことと、そ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をパッケージに描き、年齢やプロフィール等の設定を決めたことにある。

実はこの貢献こそが本質的であり、このキャラクターという身体性を合成歌声に持たせたことで初めて、「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。たとえば、キャラクター側の視点から、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風に歌わせてほしいかを述べた歌詞<sup>☆6</sup>や、実在せず歌うことしかできないと語る歌詞<sup>☆7</sup>等）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等も登場した。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「このキャラクター（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「このキャラクター（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ー（声質）が「鏡音リン」「鏡音レン」「ぐっぽいど」「巡音ルカ」と増えていくにつれて、キャラクター間の関係性を表現した歌詞等も登場している。

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。

### ■ 活発な創作活動が起きた

誰でもソフトウェアさえ購入すれば、初音ミクに自分の曲を歌わせることができる。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通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 CGM 的に数万曲生まれた。そこから淘汰されて数十万～数百万再生されるような有名曲は、ポピュラリティの高い良質な作品となる。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。

ただし、初音ミクの歌声は自分の作品で自由に使っても、そのキャラクターのイラストは、通常は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、キャラクター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「キャラクター利用のガイドライン」を宣言し、「ピアプロ・キャラクター・ライセンス（PCL）」<sup>☆8</sup>を発行することで、個人の

☆5 <http://www.vocaloid.com/lineup/vocaloid2/>

☆6 <http://www.nicovideo.jp/watch/sm1050729> 等

☆7 <http://www.nicovideo.jp/watch/sm1274898> 等

☆8 <http://piaapro.jp/license/pcl/summary>

創作意欲を引き出し、誰でも合法的に非商用利用することを可能にした（本特集<sup>1)</sup>の伊藤氏の記事参照）。つまり、権利を開放したのである。この取り組みの成功は他社にも波及し始めており、たとえば日本マイクロソフト（株）も自社のキャラクタに対してPCLとほぼ同等のライセンス<sup>☆9</sup>を定めている。

この権利開放により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CGM的に創作された。また、キャラクタを中心とした動画制作を容易にする3次元CG（コンピュータグラフィクス）ソフトウェア「MikuMikuDance（MMD）」<sup>☆10</sup>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。

こうした楽曲や動画の作品を発表する場として、ニコニコ動画は最適であった。動画作品を投稿すると、作品内の時刻に同期したコメント（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（本特集<sup>1)</sup>の戀塚氏の記事参照）。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。

### ■ N次創作の連鎖反応が起きた

さらに他人が作った曲に合う映像を付けて動画にしたり、他人が描いた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も盛んになった。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キャラクタで再現した動画等も発表されている。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、たとえば、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。

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）を「N次創作」<sup>2)</sup>と呼ぶ（本特集<sup>1)</sup>の濱野氏の記事参照）。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、コンテンツ投稿サイト「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。

### ■ ライブコンサートの歌手が歌声合成とCGキャラクタになった

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。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5,000人以上の観客を集め<sup>☆11</sup>、2011年11月11日にSingaporeで2,500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カ所（日本24カ所、台湾7カ所、香港2カ所、上海1カ所）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

☆9 <http://msdn.microsoft.com/ja-jp/hh298798>

☆10 <http://www.geocities.jp/higuchuu4/>

☆11 <http://www.youtube.com/watch?v=WzYssX-CcSY>

送るのも興味深い。なぜ自分の声が相手（初音ミク）に届かないと分かっているにもかかわらず、人々は声援を送ったり初音ミクに呼びかけて叫んだりせずにはいられないのか。それは声を出す自己表現手段であると同時に、会場の聴衆間のコミュニケーション手段であるからである。ステージ上の受け手の実在は声援行為に本質的ではなかったことが、こうしたコンサートだと浮き彫りになる。むしろ、ほかの観客が周囲にいることが本質的なのであろう。今後、研究開発が進めば、歌手の声と映像がコンサート中に動的に生成されるようになるのは間違いなく、そうなったときの歌手と観客とのインタラクションも興味深い。

上記の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。

### ■ 人と人をつなげるハブとなった

初音ミクには、人と人をつなげるハブ (hub) の効果もある。異分野にいる人々が、初音ミクや歌声合成をキーワードに次々とつながってきた。クリエイター同士も、N次創作のように多くのコンテンツが絡み合う過程で交流が促進され、視聴者を交えたコミュニケーションも活発である。後述するような学術的な研究開発だけでなく、周辺のソフトウェア・ハードウェア開発や動画ランキング制作等も活発で、さまざまなコラボレーションが生まれる土壌となっている。

### 歌声合成の技術・社会・文化

初音ミク、ニコニコ動画、ピアプロが切り拓いた CGM 現象は、従来にない連鎖反応を引き起こした

ために、本特集<sup>1)</sup>の他の記事にもあるように、さまざまな観点から議論されてきた。以下では、歌声合成技術という側面に焦点を当てて議論する。なお、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、上記の解説や下記の議論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。

### ■ 一過性のブームなのか

初音ミク、ニコニコ動画、ピアプロが切り拓いた CGM 現象は、単なるブームではなく、持続発展する文化である。文献<sup>3)</sup>でも述べたように、歌声合成技術が普及することは歴史的必然である。これまでピアノやギター、シンセサイザのような新たな技術の登場が新たな音楽表現を生んできており、今後も必然的に、新たな音楽表現が生まれ続ける。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、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、今後もますます発展していくと考えられる。

すでに、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視聴者は多く、特に、10代の女性の比率が高いという調査結果が発表されている<sup>☆12</sup>。もし若い頃に好きだった音楽を一生好むことが多いとすれば、歌声合成による新たな音楽表現は今後も好まれ続ける可能性がある。特に、上記の初音ミク主演のライブコンサートを会場で観れば、一生忘れられないぐらい強烈な体験となり得る。



<sup>☆12</sup> <http://blog.piapro.jp/2011/12/youtubevocaloid.html> 等。

## ■ 人間の歌手はいらなくなるのか

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。

## ■ なぜ日本なのか

日本での音楽情報処理の研究開発が先進的であるという背景はあるが、本現象がこのタイミングで日本において起きたのは、技術の優位性だけではない。日本の技術、日本の社会、日本の文化のそれぞれの強みが相乗効果で発揮されたからである。

上記のように日本の社会には先進的なコメント機能を持つニコニコ動画が普及しており（登録会員数が2,500万人で日本人口の約19%以上）、誰もが作った曲をすぐに不特定多数の人たちに聴いてもらい、時刻同期コメントをもらうことができる。さらには、キャラクタを大切にす文化が日本にはあり、キャラクタを軸としたさまざまな展開が起きる土壌がある。その上、音楽やイラスト、映像等のさまざまなコンテンツで潜在能力が高く熱意あるクリエイターが日本には多く、従来は埋もれていたのが、1次創作やN次創作等で活躍するようになった。新しいものを好む人たちも多く、楽曲・動画のクリエイター・視聴者や、技術のユーザとして本現象に貢献している。魅力的な声を提供できた声優・歌手の貢献もある。

過去に遡って考えれば、たとえば、能や歌舞伎にはすでにキャラクタを大切にす文化があり、連歌はN次創作そのものである。ほかにも、人形浄瑠璃との関連性も指摘されている<sup>☆13</sup>。日本の伝統は生きている。

## 初音ミク, ニコニコ動画, ピアプロに 関連した研究

情報処理研究者の立場からの研究開発も活発に取り組まれている。音楽情報処理分野の部分領域である「歌声情報処理」<sup>4)</sup>は以前から研究開発が進んできたが、初音ミク現象以降、社会的な関心が高まり、さまざまな大学や企業、研究機関、個人が参入してきている。たとえば歌声合成関連では、上記のVOCALOID 以外に

- お手本歌唱を模倣することで初音ミク等をより自然な歌い方で歌わせることができる歌声合成技術「VocaListener (ばかりす)」<sup>☆14</sup>
  - HMM 歌声合成技術に基づいて任意の歌詞の歌声を合成する Web 上のサービス「Sinsy」<sup>☆15</sup>
  - 歌声合成ツール「UTAU」<sup>☆16</sup>
  - 音声分析合成システム「WORLD」<sup>☆17</sup>
- 等が研究者以外にも知られつつある。歌声合成以外にもさまざまな歌声情報処理技術の研究開発・実用化が進んでいる<sup>☆18</sup>。

ニコニコ動画上での創作活動も研究されており、たとえば、初音ミク関連のN次創作における協調的創造活動（引用関係や役割等）が分析された<sup>5)</sup>。その先進性から、ニコニコ動画全般に関連した研究も多くなされており、たとえば、N次創作によって音楽に合ったダンス動画を自動生成するシステム「DanceReProducer」や、ランキング動画全自動生成システム「ニコニコランキングメーカー」、ニコニコ動画の時刻同期コメントの自動生成システム「MusicCommentator」、時刻同期コメントを利用した信頼度・感情分析やサビ検出、ニコニコ動画のタグ共起の分析等が挙げられる<sup>☆19</sup>。

☆13 <http://ascii.jp/elem/000/000/634/634796/>

☆14 <http://staff.aist.go.jp/t.nakano/VocaListener/index-j.html>

☆15 <http://www.sinsy.jp/>

☆16 <http://utau2008.web.fc2.com/>

☆17 <http://www.aspl.is.ritsumeai.ac.jp/morise/world/>

☆18 <http://staff.aist.go.jp/m.goto/IPSJ/event20100728.htm> および <http://staff.aist.go.jp/m.goto/IPSJ/event20120203.htm> 等。

☆19 <http://staff.aist.go.jp/masahiro.hamasaki/niconico.html> にその多くが列挙されている。

ピアプロに関しては、ピアプロのためだけの研究ではないが、能動的音楽鑑賞サービス「Songle」<sup>☆20</sup>が2012年2月から試験的に連携を始めた。ピアプロ上の楽曲等を音楽音響信号理解技術によって自動分析して、その内容の可視化機能やサビ出し機能等を Web 上で誰でも体験できる形で実現している。

いずれも研究開発の余地はまだ大きく、研究テーマの宝庫ともいえる状況なので、今後もより多くの研究者の参入とさまざまな角度からの貢献が望まれる。

### CGM 現象の未来

アマチュアのクリエイターが生き生きと活躍する CGM 現象は、技術の進展に伴って創作の敷居がさらに下がり、今後ますます発展していくであろう。なぜなら、上記の「歌いたいから歌っている」ように、あるいは「スポーツをしたいからしている」ように、人間は「表現したいから表現する」ためである。根源的には、そこに必ずしも見返りは必要ない。だから、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、さらに嬉しくなって創作の連鎖が起きていく。それが「一億総クリエイター時代」である。

筆者は、今後は「単位エネルギー当たりの幸せ」（幸せ ÷ エネルギー）という新たな概念が重要になると提唱している。音楽等のメディアコンテンツは、豊かで人間らしい生活を送る上でもはや不可欠であり、幸せや充足感を得る手段の有無は生活の質に直結してくる。文献3)でも環境問題やエネルギー問題への寄与について述べたが、特に音楽コンテンツは、必要なエネルギーが少なく繰り返しの視聴に耐える良質な娯楽であり、制作資源・エネルギーはデジタルコンテンツ制作環境の普及により下がってきた。N次創作は、既存の複数のコンテンツの良い

点を凝縮して幸せを増すのに有効な手段として位置づけられる。結果的に、コンテンツがリサイクル(再利用)される側面もあるので、エネルギー効率の良いコンテンツ制作手段に意図しないうちになっていると解釈できる。こうした「単位エネルギー当たりの幸せ」を向上する(たとえば、消費エネルギーを減らしつつ幸せを増やす)ための研究開発は、今後ますます重要になっていくであろう。

それでもなお、こうした CGM 現象、初音ミク現象を単なるサブカルチャーだと捉える見方もあるかもしれない。しかし、現在メインストリームだと思われているほかのさまざまな文化も、その発端ではサブカルチャーであったのではないだろうか。そういう意味では、この情報処理技術が主導する CGM 現象を、世界も注目する日本発のメインストリーム文化として育て上げていけるかどうかは、一人一人の行動にかかっているといえる。技術の切り拓く未来を描く手段の1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」<sup>6)</sup>では、この日本発の CGM 現象の本質が切り拓く明るい未来の1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だろうか。10年後が楽しみである。

#### 参考文献

- 1) 特集「CGMの現在と未来:初音ミク、ニコニコ動画、ピアプロの切り拓いた世界」(解説5件)、情報処理(情報処理学会誌)、Vol.53, No.5, pp.464-494 (May 2012)。
- 2) 濱野智史:アーキテクチャの生態系—情報環境はいかに設計されてきたか、NTT出版(2008)。
- 3) 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 (June 2010)。
- 4) 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)。
- 5) 濱崎雅弘、武田英明、西村拓一:動画共有サイトにおける大規模な協調的創造活動の創発のネットワーク分析—ニコニコ動画における初音ミク動画コミュニティを対象として、人工知能学会論文誌、Vol.25, No.1, pp.157-167 (2010)。
- 6) 野尻抱介:南極点のピアピア動画、早川書房(2012)。(2012年3月19日受付)

後藤真孝(正会員) | m.goto[at]aist.go.jp

1998年早大大学院博士後期課程修了。博士(工学)。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(<http://songle.jp>等)を20年間、音声言語情報処理(<http://podcastle.jp>等)を14年間研究。

☆20 <http://songle.jp>