

# Miku Miku Face: HANAUTAUと 顔プロジェクトマッピングによるバーチャル「初音ミク」

竹本 拓真<sup>1,a)</sup> 片寄 晴弘<sup>1</sup>

**概要:** 歌唱合成ソフト“初音ミク”は、バーチャルアイドルとして人気を確立しており、初音ミク主演のライブが開催されている。しかし現在開催されている初音ミクのライブは、あらかじめ用意された映像を流しているだけであり、歌手と観客とのインタラクティブなライブを実現できていない。また平面スクリーンでは前方向以外の映像が無く、観客の視点が一方向に固定され、観客が多方向から視聴するアリーナ形式のライブを行うことができない。本稿では、リアルタイム歌唱生成システム“HANAUTAU”と立体顔形状ディスプレイの組み合わせによって歌唱と歌手の動作をリアルタイム生成し、多方向からの視聴に耐えるインタフェース“Miku Miku Face”を提案する。

## Miku Miku Face: Virtual ” Hatsune Miku” realized with HANAUTAU and Face Projection Mapping

TAKEMOTO TAKUMA<sup>1,a)</sup> KATAYOSE HARUHIRO<sup>1</sup>

**Abstract:** singing voice synthesizer “Hatsune Miku” has been developed as a virtual-idol. The live of Hatsune Miku starring is held. However, it is only by passing a static animation. So, it impossible to the interactive live between singer and audience. Addition, flat screen there is no video other than forward. Therefore, audience viewpoints are limited to one direction, and it is not possible to perform arena style live the audience view from multiple directions. In this paper, we propose a interface “Miku Miku Face” to perform generation of motion and singing of the singer in realtime by using “HANAUTAU” and “3D face shape display”

### 1. はじめに

歌唱合成ソフトウェア“VOCALOID2 初音ミク”(Crypton, 2007)は、歌唱音源の声優起用、アニメ風の身体などのキャラクタ設定の付与により、歌唱合成ソフトとしてだけでなく、バーチャルアイドルとしても人気を確立している。一般アイドルと同様に、初音ミクのライブイベント(以下ライブ)開催への要望が強まり、初音ミク主演のライブが世界各地で開催されている(図1<sup>\*1</sup>)。

二次元キャラクタである初音ミクのライブを実現するにあたっては、実世界上に等身大の初音ミクを何らかの方法

で存在させる必要がある。現在開催されている初音ミク主演のライブではステージ上に透過スクリーンを配置し、プロジェクタ数台の映像を透過スクリーン上に重ね合わせて映写することで、視認性の高い初音ミクの3DCG映像を実現している。これにより初音ミクがあたかもステージ上に立っているかのように見せている。歌唱はあらかじめ用意した音源を流しており、伴奏は人間による生演奏となっている。

本来のライブにおける重要な要素として、観客の声に対する歌手の応答(アドリブ・トーク・視線)などの、歌手と観客とのインタラクション及び、歌手がそこにいると感じる臨場感の高さの2点がある。初音ミクのライブにおいて、前者については、初音ミクの歌唱や映像はあらかじめ用意された静的なものであるため歌手と観客とのインタラクションが完全に失われている。後者については、透過ス

<sup>1</sup> 関西学院大学院  
Kwansei Gakuin University, Sanda, Hyogo, Japan  
<sup>a)</sup> takumarakan@kwansei.ac.jp  
<sup>\*1</sup> <http://oh-news.net/live/?p=97150>



図 1 初音ミクのライブ風景  
Fig. 1 Live of Hathune Miku

クリーンでは初音ミクの映像が平面であるため、観客が前方以外の場所からみた場合、モノリザ効果<sup>\*2</sup>が発生し、高い臨場感を得にくいというのが現状である。

本稿では、これらの問題点を解決することで、初音ミクと観客とのインタラクションを高臨場感で体感できるインターフェース“Miku Miku Face”を提案する。我々の先行研究である“HANAUTAU”[1]により、ユーザの鼻歌と歌詞入力によってリアルタイムに歌唱を生成できるシステムを開発した。これらの機能は歌手と観客とのインタラクションの再現に有用であるが、本稿では HANAUTAU に、顔の形をしたスクリーンを立体的に成型しそこに顔プロジェクションマッピングすることで遠隔者と臨場感のあるインタラクションが可能な立体顔形状ディスプレイを組み込み、ユーザ頭部の 3 次元の動きや表情を取得して初音ミクの視線や仕草を再現することで、臨場感のある歌手と観客とのインタラクションの向上を計る。

## 2. 本研究のアプローチ

現在催されている初音ミクのライブは一方向的に初音ミクが歌い、一方向的に観客が声援を送っている状況であり、本来のライブにおける重要な要素が失われてしまっている。初音ミクのライブをより高臨場感かつインタラクションのあるライブとするためには、まず初音ミクの歌唱・会話のリアルタイム生成する必要がある。我々の先行研究である“HANAUTAU”[1]により歌唱・会話をリアルタイムに生成可能である。

次にライブ会場のステージ上に初音ミクの身体を存在させる必要がある。現在の初音ミクのライブにおいては透過スクリーンによりステージ上に初音ミクを映写しているが、スクリーンがステージの大部分を占有するためその他の演出の妨げになってしまっている。また、透過スクリーンでは初音ミクの映像は前方向のみであり、モノリザ効果やスクリーンの視野角の制約により観客の視点が一方に限定

されてしまっている。その為現在催されている初音ミクのライブは全てシアター形式であり、多方向に観客の視点があるアリーナ形式のライブは実現できていない。スクリーンでは存在感や臨場感を十分に再現できない。様々な立体化の手法の中で、存在感や臨場感の高さに関して実世界上に実物が存在することに勝る手法は確立されておらず、ロボットによる実体化が最善であると考えられる。本稿では、まず歌手と観客のインタラクションにおいて特に重要と思われる、頭部の製作を行う。

石黒らのジェミノイド<sup>\*3</sup>など、物理的な構造によるヒューマノイドロボットが開発されているが、顔部分は稼働部が多く複雑に駆動している。顔の動きを細部まで再現するには大量のアクチュエータが必要であり、それに伴い重量・動作音・消費電力・故障率が増加してしまう。この問題を解決する手法として立倉らは“Mask-bot”[2]を提案している。

Mask-bot は立体顔形状ディスプレイの後方からプロジェクタによる顔映像を映写することで立体的な顔を実体化させる手法である(図 2<sup>\*4</sup>)。これによりヒューマノイドロボットの顔に使用されている全てのアクチュエータをプロジェクタ 1 台に集約することを実現した。また、二次元キャラクターの顔は極端にデフォルメされ、物理的な機構では表現不可能な変形が生じる(図 3)。この点においても立体顔形状ディスプレイでは容易に表現可能であり、本研究では初音ミクの頭部の実体化に立体顔形状ディスプレイを採用する。既存の立体顔形状ディスプレイでは、頭部の顔部分しか作られていない。多方向からの視点でも臨場感を維持するためには顔だけでなく後頭部まで作りこまなければならない。ユーザの頭部の動作や顔の表情をセンシングし、ディスプレイの首部分に取り付けられたサーボモータと連動させる。また、人間の表情を二次元キャラクターにマッピングする際、二次元キャラクターの表情は極端にデフォルメされており同じ部位を連動させるだけでは不十分である。本研究では、人間が笑う際の口を横に広げる動作を二次元キャラクターの目を細める動作に連動させることで、観客との掛け合い時における初音ミクの仕草や表情を再現しインタラクションの向上を計る。

この機能によって実現されるリアルタイムに歌唱と頭部動作を生成するシステムを以下、“Miku Miku Face”と呼ぶ。以下、Miku Miku Face を実装するにあたって利用した技術 HANAUTAU と立体顔形状ディスプレイについて紹介する。

<sup>\*2</sup> 平面ディスプレイに顔を表示した場合、どの方向から見ても目線があっているように感じてしまう現象

<sup>\*3</sup> 大阪大学や ATR が参画する文部科学省グローバル COE プログラム「認知脳理解に基づく 未来工学創成」(拠点リーダー: 石黒浩)において開発された。

<sup>\*4</sup> 立倉らの“Mask-bot”[2]論文内より引用

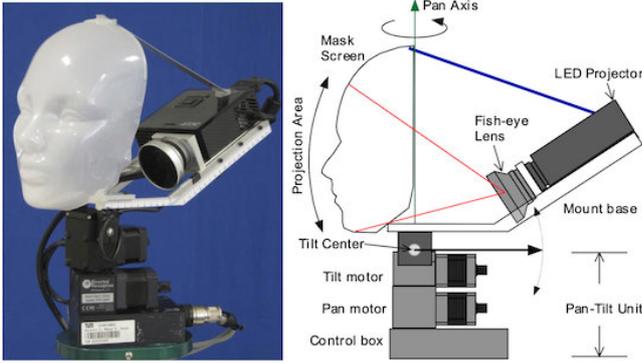


図 2 Mask-bot



図 3 二次元キャラクタにおけるデフォルメされた表情  
Fig. 3 internal of Geminoid F

## 2.1 HANAUTAU

我々の先行研究である HANAUTAU[1] は、ユーザの鼻歌によりピッチを、タイピングにより歌詞を、それぞれ与えることによって初音ミクをリアルタイムで歌わせることのできるシステムである。タイピングにより指定された歌詞の音源のピッチをマイク入力された鼻歌から推定した基本周波数にフェイズボコーダを使用してリアルタイムに変更することで実現している。鼻歌によるピッチ指定とすることで鼻歌に含まれているピッチに関する歌唱表現をそのまま初音ミクの音源に転写できるため自然な歌唱を生成することが可能となっている。鼻歌からのピッチ指定であるた細かなピッチ変化が可能で会話の抑揚も再現でき、歌唱だけでなく会話もリアルタイムに行うことが可能である。

## 2.2 立体顔形状ディスプレイ

立体顔形状ディスプレイとは、人間の顔の形状をした立体的なディスプレイに裏側から実際の人間の映像を3次元映写したものである。倉立らの“Mask-bot”[2]は、顔部分をプロジェクタによる映写とすることで物理的な構造を排除し、ロボット製作にかかるコストを大幅に削減している。Mask-botの首部分にはサーボモータが取り付けられており、3自由度で話者の仕草や表情を再現している。三澤ら

の研究[3]により立体顔形状ディスプレイは、人物認識が可能な角度が平面ディスプレイに対して増幅し、モナリザ効果が解消されることが分かっており、立体顔形状ディスプレイは観客に対して平面ディスプレイへの映写以上の高い臨場感を与えることが可能である。しかし、これらの立体顔形状ディスプレイを用いた研究において、どのシステムにおいても顔部分のみの製作だけとなっており、後頭部まで製作されておらず、側面や背後からの視線を考慮していない。

## 3. Miku Miku Faceの実装

Miku Miku Faceのシステム概要図を図4に、見取り図を図5に示す。本システムのリアルタイム歌唱合成部分には HANAUTAU[1]機能を用いてマイクから鼻歌を、PCキーボードから歌詞を入力し、初音ミクの歌唱を生成する。立体顔形状ディスプレイに背面から初音ミクの頭部3DCG映像を映写し、PCキーボードから入力された歌詞の母音によって初音ミクの口を動作させることで、リップシンクを行う。ユーザの首の向き、表情の情報を Miku Miku Face

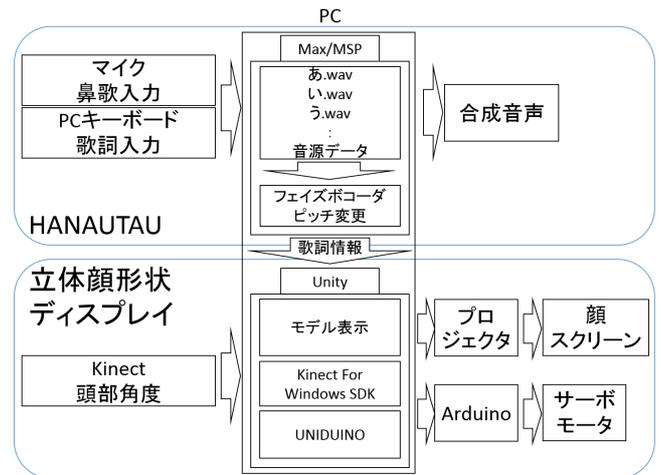


図 4 Miku Miku Face の概要図  
Fig. 4 Overview of Miku Miku Face

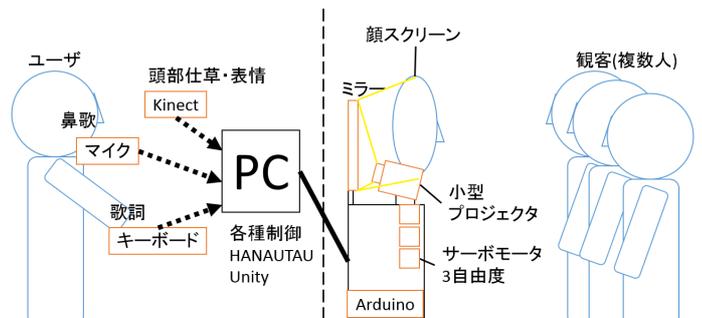


図 5 Miku Miku Face 見取り図  
Fig. 5 Sketch of Miku Miku Face

に転写するために、Kinect を使用する。Microsoft 社から提供されている Face Tracking SDK を使用し、ユーザの頭部をトラッキングし、顔が正面を向いている場合を初期位置とした 3 軸の回転角度・口の開き具合を検出し、それらを 3DCG キャラクタに反映させることで初音ミクの仕草や表情を表現する。

モデル表示や各モジュールとの通信・制御は Unity により行う。HANAUTAU と立体顔形状ディスプレイは別々の PC により制御されており、ユーザは立体顔形状ディスプレイの設置位置とは別の場所で HANAUTAU の操作が可能である。PC 間の通信には OSC(Open Sound Control) を使用した。

立体顔形状ディスプレイのモデルは“ままま式あびミク”を使用する。髪・後頭部は 3D プリンタによりナイロン樹脂で出力した後にアクリル塗料により着色する。顔スクリーン部は 3D プリンタにより型を出力し、透明な PET 板を加熱した後に型に押し当て PET 板に顔の形状を転写する。顔スクリーンの裏面に背面投影用塗料を均一に添付する。顔スクリーンの後方に光量 85 ルーメンの小型プロジェクタを搭載し、投影距離を確保するために鏡に 2 度反射させ、顔スクリーンに顔プロジェクションマッピングすることで初音ミクの顔映像を映写している。首部分にサーボモータを 3 個取り付けて 3 自由度を確保する。サーボモータの制御には arduino を使用し、PC とはシリアル通信で接続する。サーボモータの最大トルクは電圧 11V 時において 41.7kgf・cm であり、本研究では 6V で駆動させている。

#### 4. プロトタイプ展示と検討

EC2014 学会において Miku Miku Face のプロトタイプの展示を行った [4]。EC2014 の展示において気持ち悪い・不気味だなどの否定的な意見が見受けられた。原因として



図 6 プロトタイプ Miku Miku Face と Miku Miku Face  
Fig. 6 Prototype Miku Miku Face and Miku Miku Face

首の稼働位置が実際の人間とかけ離れている、顔の前部分しかいという点が挙げられた。本稿の Miku Miku Face では首の駆動位置の修正と後頭部まで製作することでこの点を改良し、本バージョンを NT 東京展示会 \*5 に展示した。NT 東京では、否定的な意見はほぼ無くなり、肯定的な意見が多かった。

Miku Miku Face はプロジェクタを用いて顔部分を投影しているため、明るい展示会場では顔が薄くなってしまう。プロジェクタの性能は頭部に搭載可能な大きさと重量を考慮する必要があり、光量を増加させることによる改良が困難であるため、できる限り光量の減衰を抑える必要がある。高性能な魚眼レンズを使用し、至近距離から顔スクリーンに直接映像を映写することで鏡の反射による光量の減少を排除する方法の検討及び、顔スクリーンの背面投影用塗料の添付量を最適にすることでより精細な映写ができるよう改良していく。また、暗い展示会場では顔がはっきり見える反面、顔だけが輝いて見えてしまう。照度センサを使用し、環境光に応じてプロジェクタの光量を変化させる必要がある。

#### 5. まとめ

本稿では、HANAUTAU と立体顔形状ディスプレイを用いて歌唱と歌手の動作の実時間生成を行うことにより、インタラクティブでより臨場感のあるバーチャル初音ミクを実現させるためのインタフェース “Miku Miku Face” を紹介した。立体顔形状ディスプレイを用いた臨場感向上に関する既存研究は実際の人間の顔を対象としているため、デフォルメされたキャラクタの顔であっても同様の臨場感が得られるかどうかについては検証の必要がある。現在 kinect から取得している表情パラメータを全て使用しきれしていない。今後二次元キャラクタのデフォルメされた表情を人間の表情とどのように関連させるべきか検討し、Miku Miku Face のさらなるインタラクティブ及び、臨場感の向上を計っていく。

#### 参考文献

- [1] 竹本 拓真：リアルタイムに初音ミクを歌わせるタイプソングシステム「HANAUTAU」とそのアジャイル型開発事例報告, インタラクシオン 2014(2014)
- [2] T. Kuratate “Mask-bot”: a life-size robot head using talking head animation for human-robot communication, Humanoids 2011, pp.99-104(2011)
- [3] 三澤 加奈：LiveMask：立体顔形状ディスプレイを用いたテレプレゼンスシステムにおけるコミュニケーションの評価, インタラクシオン 2012, pp.41-48(2012)
- [4] 竹本 拓真：Miku Miku Face：HANAUTAU と立体顔形状ディスプレイを用いたバーチャル「初音ミク」, EC2014, pp.73-74(2014)

\*5 <http://j.nicotech.jp/nt.tokyo.2014>