



録画録音の発注資料

この度は数ある依頼の中からお選びいただき、誠にありがとうございます。

近年弊社ではAI事業に取り組んでおり、現在は表情の動きから心情に合った声色で発音をさせる研究も行っています。

今回ご協力いただきたいのは、

- ・ライブやコンサートで発される叫び声の録音と、その際の表情の移り変わりをデータ化したいため
- ・顔を正目から撮影した動画です。

納品して頂いたデータは当方で編集させていただきますので、予めご了承ください。

※撮っていただいた音声、動画は、AIを作成する音声の参考にさせていただきますのでそのままの使用、顔の公開は一切ございません。①、⑥で補足を記述しております。

今日需要の高まっているVTuber界の新技术の火付け役となる為ために重要な、何色もの声を持つエキストラ声優さんを探し求めておりました。初の試みとなるので不安な点は多くあると思いますが、その都度対応していく所存です。何でもご相談くださいませ。AI技術は需要が高まっている一方ですので、今後様々な場面で使用されていく可能性を秘めています。

もくじ

- ① 使用目的
- ② 手順・納品日
- ③ 参考音声、読んでいただくセリフ一覧
- ④ 納品していただく物まとめ
- ⑤ 留意いただきたい点
- ⑥ 顔画像の肖像権、個人情報の保護について

① 使用目的

※音声データだけでなく顔の表情を撮影する理由

右図のようにソフトを用い、AIの特徴である無機質な発音と、今回撮っていただく表情の変化を撮影した動画を組み合わせ、【感情的な声色】を作成する為です。
ですから第三者に顔画像を公開することは、声優担当者様の事前の同意、承諾を得た場合を除き、一切ございません。



以前から日中共同プロジェクトの「直感 x アルゴリズム

(<https://project-algorithm.com/index.html>)」さんとリアルタイムのモーションキャプチャを利用した

VTuberシステムにも力を入れて取り組んでおり、今後ライブイベントを展開していく予定です。

今回は音声データだけでなく表情のデータを絡めていくことにより、感情的な読み上げ技術の強化を図ります。

これによりAI読み上げ機能の正確性を高め、応用させていく予定です。

② 手順・納品日

手順		納品日	形式
1	裏声、アニメ声等、声優担当者様の出せる声色の種類を、いくつか動画、音声データで送っていただきます。	依頼受注日の翌日終日	mp4 形式の動画 撮影環境が分かるもの (何個か任意)
2	頂いたデータから、録画録音をきちんとした音響の場所とマイクで撮れているかを当方で確認させていただき、問題が無いようでしたら同じ環境で以下③のセリフを録音、録画していただきます。	この時に、最終納品日のご相談をさせていただきます。(当方としては早ければ早い程ありがたく存じます。)	mp4 形式の動画ファイル 120 個 (+指笛の動画ファイル)
3	③ を納品していただけたら、再びこちらでデータを確認、編集させていただきます。お互い不備、質問等なければ一旦お取引終了となり、報酬お支払い等行っていきます。		

③ 参考音声、読んでいただくセリフ一覧

・参考音声

2つのセリフを、AI が別のイントネーションで発音した比較音声データです。
【[AI_サンプル音声 - Google ドライブ](#)】をタップ/クリックすると共有 drive ページへ飛びます。
このように、撮らせていただいたデータをもとにより高性能な AI を目指している所存です。

・セリフ一覧

コンサートやイベントでのシチュエーション別 掛け声		
1 がんばってー	11 会いたかった（よ）ー	21 キャー
2 すきー	12 ありがとー	22 ギャー
3 だいすきー	13 なに（なに）ー？	23 うおおー
4 さいこうー	14 俺もー	24 イエー
5 カッコいいー	15 ダメー	25 わー
6 いいよー	16 愛してるー	26 ああああ
7 泣かないでー	17 おめでとー	27 ええ～！？
8 かわいいー	18 いけー	28 【裏声】フゥー
9 こっち向いてー	19 アンコール！	29 【掛け声】ハイッハイッ
10 イケメン	20 せーのっ	30 【合いの手】〇〇の瞳に恋してる

上記セリフ 30 個を、以下 4 つの異なる声量で録音、録画（30 種×4 パターン）

求めている 声量	例	参考動画 URL	
叫び声 強	怒鳴り声 最大音量	https://twitter.com/124_jin_p/status/1455946155427131399?s=21	https://youtu.be/pbskmjCs9BA?t=186
叫び声 中	語尾伸ばしめ	https://youtu.be/v0yCu9oFC5I?t=131	https://youtu.be/52sF6qoXLQE?t=560

叫び声 小	バラード後 直後の掛け声	https://youtu.be/GkX4w_KJZSI?t=68	https://youtu.be/vRyeA60TZwk?t=1765
気持ち大きい話し声	普通、 ガヤガヤ程度	https://youtu.be/F32epua9VQs?t=75	https://youtu.be/McnDJsm4VrI?t=7
(指笛)	曲の合いの手	https://youtu.be/-32gw7khILg?t=10	

数千人規模の大観衆のライブを想定しております。

声質はできるだけ統一して録音していただくと幸いです。

④ 納品していただく物まとめ

1. 声優担当者様の声質、撮影環境の確認できるデータ : サンプルとして③-1「がんばれー」を4種類の音量で収録したMp4形式のファイル4つ

2. 「コンサートやイベントでのシチュエーション別 掛け声」を4パターンの声量で言葉を発している際の、正面から顔を撮影した視点固定のmp4形式の動画 : 全120個
(+指笛できる方は、指笛を鳴らす動画)

※それぞれ1つずつ区切ってファイル分けをお願いいたします。

⑤ 留意いただきたい点

・この依頼は、録画録音をきちんとした音響の場所とマイクで録れる方にのみお願いしたく存じます。

条件が満たされていない場合での収録料金はお支払い出来兼ねますのでご了承ください。そうならないように相談しながら進めていけたらと思っております。

・納品して頂いた音声データは、当方で編集を施しますので予めご了承ください。

・ボイスメモでの録音は可能ですが、内部マイクではなく**外部マイクが必要**になります。

iPhoneで撮影/録音されるかたは、movファイル(それかmp4形式)で送付、

androidで撮影/録音されるかたは mp4形式での納品をお願いいたします。

ご注意くださいのは、大きな声が、潰れずに適切に録音されていることが必要です。そのため、外部マイクで適切に録音レベルが調整されていることが必要になります。

・顔動画を撮影する際、動画は画面サイズ解像度 **1280x720i以上**、fps25以上、

(iPhoneで撮影の方は設定→カメラ→ビデオ撮影→4K/60fpsに設定。SEはpixel数が不足しているので不可)

・顔のサイズは **128x128pix以上** (顔はできるだけ真正面を向くこと)

・撮影の際、カメラ位置は正面固定

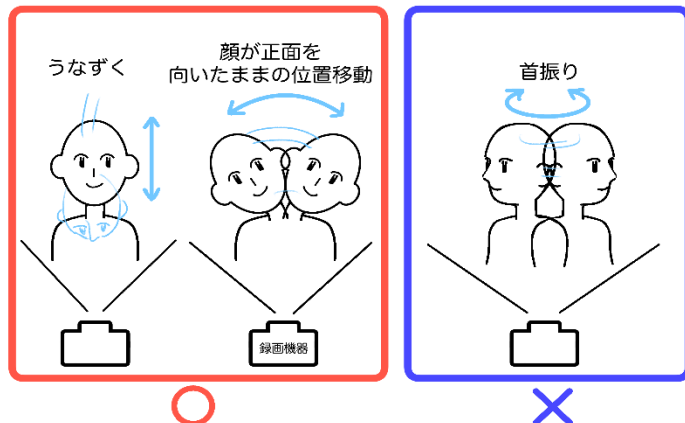
・撮影時の顔の動きについて

○：頷き、顔が正面を向いたままの傾き

×：扇風機みたいな首振り、

正面を向いていない

上記でのご協力、何卒よろしくお願いいたします。



⑥ 顔画像の肖像権、個人情報の保護について

顔画像は AI の学習用に使う用途以外に使用しません。

当社が収集させて頂いた顔画像は、以下の何れかに該当する場合を除き、第三者に提供することはありません。

お客様の事前の同意・承諾を得た場合。

その他法令等により提供が必要な場合。

さいごに

丁寧な説明を心がけてはおりますが、専門分野な上に前例が乏しく分からないことがまだまだあるかと思います。質問、不安な点があれば遠慮せずお問い合わせくださいませ。多彩な声色が拝聴できますことを楽しみにしております。

これから何卒よろしくお願い申し上げます。

長文最後までお読みいただきましてありがとうございました。