

# キャスト支援のための台本を入力とした 登場人物の顔タイプ推定方式

平野 龍一† 岡田 龍太郎† 峰松 彩子† 中西 崇文†  
武蔵野大学データサイエンス学部データサイエンス学科†

## 1. はじめに

近年、インターネットの普及によって、多くの作家が多様な物語を表現した台本を手軽に世間に配信・共有している。このことから、インターネット上に、多様な物語を表現した台本が散在しており、これらを取得できる機会が増大している。これらの物語を記述した台本に基づき、舞台作品や映画作品を創作する際に、どの役を誰に演じてもらうかというキャスト問題がある。台本に記述された登場人物について、その登場人物に合致した演者を探し出すことは、舞台作品や映画作品の制作者にとっての負担となっている。

台本のテキストデータに基づき、登場人物を表す特徴を抽出し、その特徴に合致したキャラクターを持つ演者をマッチングするシステムが実現されれば、舞台作品や映画作品の制作者の負担が軽くなるだけでなく、より品質の高い作品を創造することが可能になると考えられる。

本稿では、キャスト支援のための台本を入力とした登場人物の顔タイプ推定方式を示す。本方式は、glamfields[1]で提唱されている人種、性別を問わず世界的に用いられている7種類の顔のタイプを用いて、セリフとそのセリフの印象に合致した顔のタイプを推定する。本方式を用いることで、一つの登場人物に対して複数の演者候補がいる場合に、推論した顔タイプから、その候補を絞ることが可能となる。

本稿では、実際のシステムを構築する際に、シナリオ公募ナビ[2]で公開されているハリウッドの台本と実際に演じた俳優の顔画像を集約したデータセットを作成し、俳優の顔画像とあらかじめ準備したそれぞれの顔タイプの代表的顔との類似度比較を行うことでラベル付けを行い、セリフ-顔タイプデータベースを作成している。

## 2. 関連研究

Face Type Estimation Method of Characters from Script for Casting Support  
Ryuichi Hirano†, Ryotaro Okada†, Ayako Minematsu†, Takafumi Nakanishi†

† Musashino University, Department of Data Science

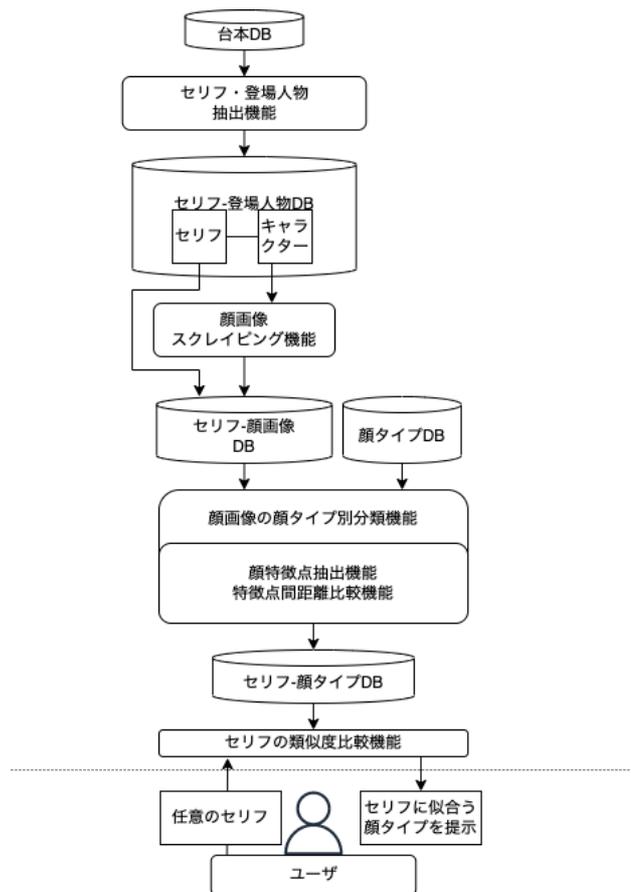


図1: キャスト支援のための台本を入力とした

## 登場人物の顔タイプ推定方式

関連研究としては、台詞に含まれる感情表現を反映したロボット動作の自動生成[3]がある。この研究ではセリフを分析し、含まれる感情表現からロボットの動作を決めると言うものだが本方式はセリフ全体を対象とする点、そしてそれらと顔タイプを結びつける点異なる。また、出力も顔タイプと言う点異なる。

## 3. 台本を入力とした登場人物顔タイプ推定方式

### 3.1 顔タイプ推定方式の全体像

図1に顔タイプ推定方式の概要を示す。本方式

は、ユーザが指定した任意のセリフ、台本データからのセリフとそのセリフに対応する登場人物の抽出、画像スクレイピングによる登場人物の顔画像の取得、台本データより抽出したセリフとスクレイピングで取得した顔画像とを対応させたセリフ-顔画像 DB の作成、顔タイプデータとセリフ-顔画像 DB の顔画像との類似度比較を用いたセリフ-顔タイプの作成、そしてユーザが入力した任意のセリフとセリフ-顔画像 DB 内のセリフとの類似度比較によるセリフに最適な顔タイプの提示によって構成される。

### 3.2 セリフ・登場人物抽出機能

シナリオ公募ナビ[2]より取得した PDF ファイルをテキスト形式に変換をする。テキストを登場人物、セリフ、その他に分類をして DB に格納する。書き出した際の頻出文が登場人物であると仮定をし、記号を含んでいる文章や括弧文のみで構成されている文章はその他として扱う。今回使用した台本の構造では登場人物とそのセリフの間に改行がなされているため、任意の登場人物の次の文章をその登場人物のセリフとして格納した。

### 3.3 顔画像スクレイピング機能

3.2 で作成した DB の”登場人物”を検索語として Google 画像検索にて検索をし、スクレイピングを行った。スクレイピングした画像を上記 DB に追加し、新たな DB を作成した。

### 3.4 顔画像の顔タイプ別分類機能

事前に glamfields[1]をもとに7つの顔タイプの標準的顔画像を入手する。これらの顔画像に対して mediapipe face mesh を使用して 468 の顔特徴点を抽出する。この抽出した特徴点のうち、鼻尖を顔の中心と仮定し、鼻尖の特徴点から他の特徴点への距離を numpy ライブラリを用いて三平方の定理を元に算出する。これによって一つの顔画像あたり 467 次元の配列ができる。同様にして 3.3 でスクレイピングした顔画像に関しても 467 次元の配列を得る。7つの顔タイプの標準的顔画像とスクレイピングした顔画像の配列を比較することで類似度を測り、最も類似度が高い顔タイプを 3.2 の DB にラベル付けをしていく。具体的には、2つの 467 次元の配列間のユークリッド距離と差の平均の二つを用いて類似性を測る。まずはユークリッド距離を算出し、正規化を行う。次に二つの配列の相互に対応する次元ごとに差を求め、その差の平均を取得し、こちらも正規化を行う。この正規化されたそれぞれの結果を足すことで類似性を求める指数とした。この指数が最も小さいものが類似性が最も高いものとなる。これによってセリフ-

顔タイプ DB が完成する。

### 3.5 セリフの類似度比較機能

ユーザによって入力された任意のセリフとセリフ-顔タイプ DB のセリフをそれぞれ比較し、最も類似度が高いものの顔タイプを出力する流れとなる。文章の類似度比較には、sentence-transformers/bert-base-nli-mean-tokens を使用した。上記モデルを使用して、入力された任意のセリフと最も類似度が高かったセリフ-顔タイプ DB の顔タイプを出力する。

## 4. 実験

実験として、映画”Good Will Hunting”の主人公である”Will”のセリフを入力し、以下に実験結果を示す。

diamond	37.35%	round	58.78%
heart	63.84%	square	51.27%
oblong	62.02%	triangle	61.12%
oval	69.79%		

多くのハリウッド俳優の顔タイプをまとめている VIAJE sunglasses[4]を参照すると”Will”を演じた Matt Damon の顔タイプは oval であると示されている。実験結果と合致していることが確認できる。

## 5. おわりに

本稿では、ユーザによって入力された任意のセリフをもとに、セリフ-顔タイプ DB と比較をして、最もセリフの印象に合致する顔タイプを出力する手法を示した。これによって、演技を披露する人々がキャスティングをする際の効率化とよりセリフの印象に近い役の割り当てが容易になることが期待される。

今後の課題としては、多言語への対応、より多くのデータ収集をし、より細かな顔タイプへの分類、アンケートなどを用いた演技を見る人々のセリフと役に対する印象の把握などが挙げられる。

### 参考文献

- [1] glamfields, <https://www.glamfields.com>  
 [2] シナリオ公募ナビ,  
<http://scenarionavi.blog81.fc2.com>  
 [3] 台詞に含まれる感情表現を反映したロボット動作の自動生成,  
[https://www.jstage.jst.go.jp/article/itetr/39.8/0/39.8\\_41/\\_pdf](https://www.jstage.jst.go.jp/article/itetr/39.8/0/39.8_41/_pdf)  
 [4] VIAJE sunglasses,  
<https://viajesunglasses.com/pages/face-shape-guide>