

オーディオブック自動生成のための 2次元キャラクタ特徴に基づく音声生成の検討



○大道昇^{†1} 大井翔^{†2} 佐野睦夫^{†1}
†1:大阪工業大学 †2:立命館大学

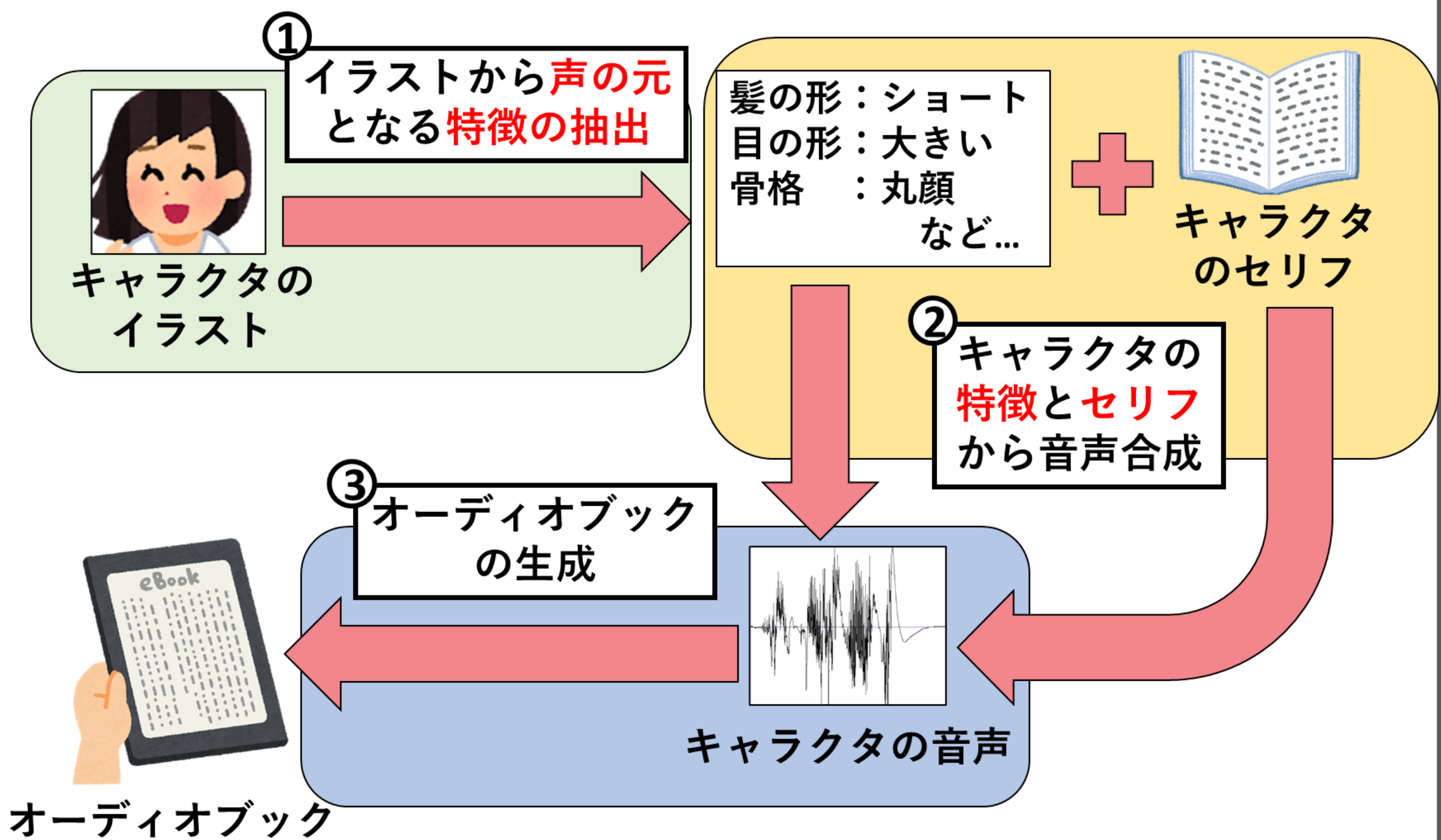


研究背景

近年のコミックや文庫は多数**電子書籍化**されており、**オーディオブック**も増えている

➡ 電子書籍にキャラクタにあった声を**ユーザが任意**で再生できる右図のようなシステムを検討する

本研究では右図の①である「イラストから声の元となる**特徴の抽出**」を試みる



提案手法

既存研究では、キャラクタの**顔全体のイラスト**画像のみから**声優を推定**できるか実験を行ったが、分類精度は0.4と悪かった

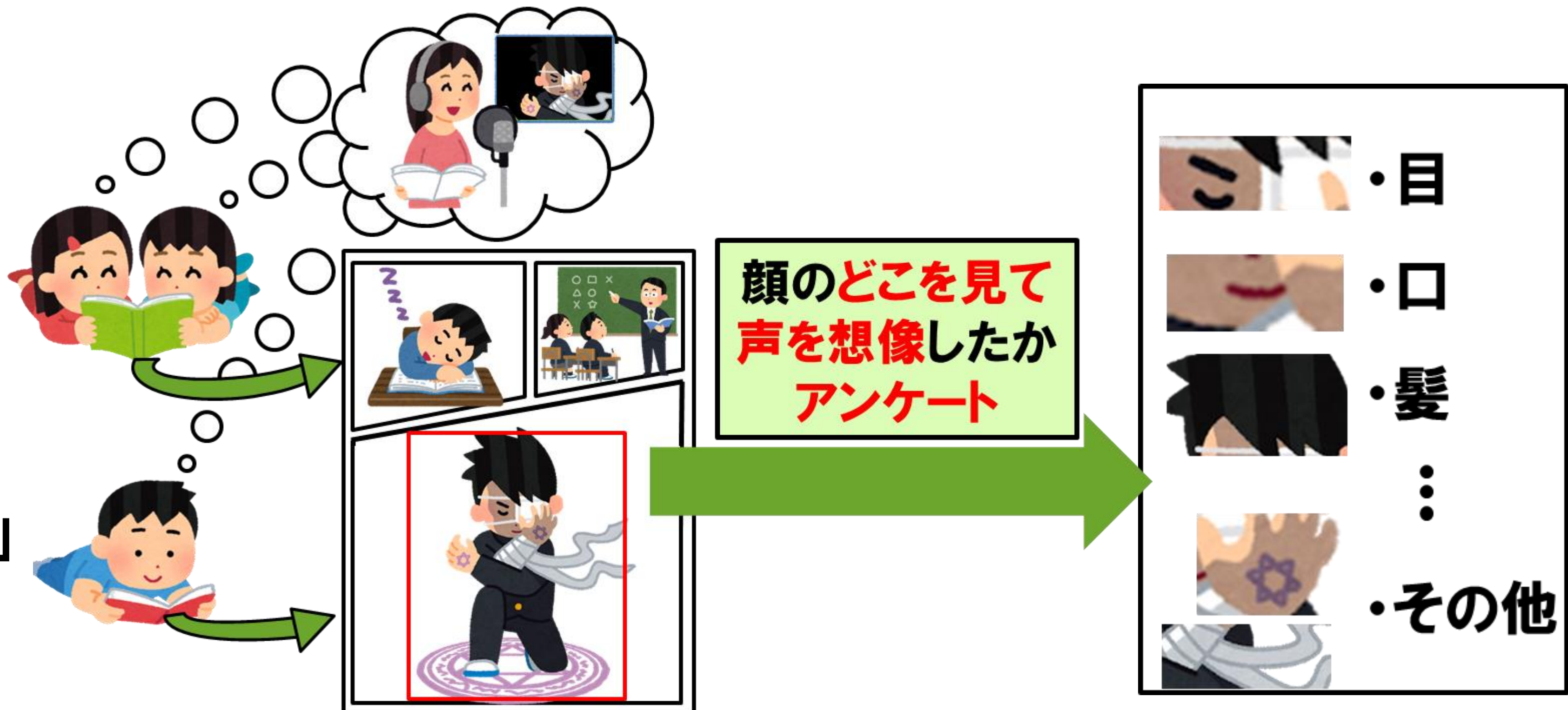
そこで本研究ではキャラクタの**顔イラストのパーツ**に着目し、どのパーツが**声を推定**するために必要なかを厳選し、各々で学習する必要があると考えた

実験

キャラクタの顔のどのパーツから、どのような声を想像したか**アンケート**を行い**必要なパーツ**を検証する

アンケート項目

「髪の色」「髪の色」「目の色」「目の形」「肌の色」「骨格」「服装」「装飾品」「その他（項目外の自由記述）」

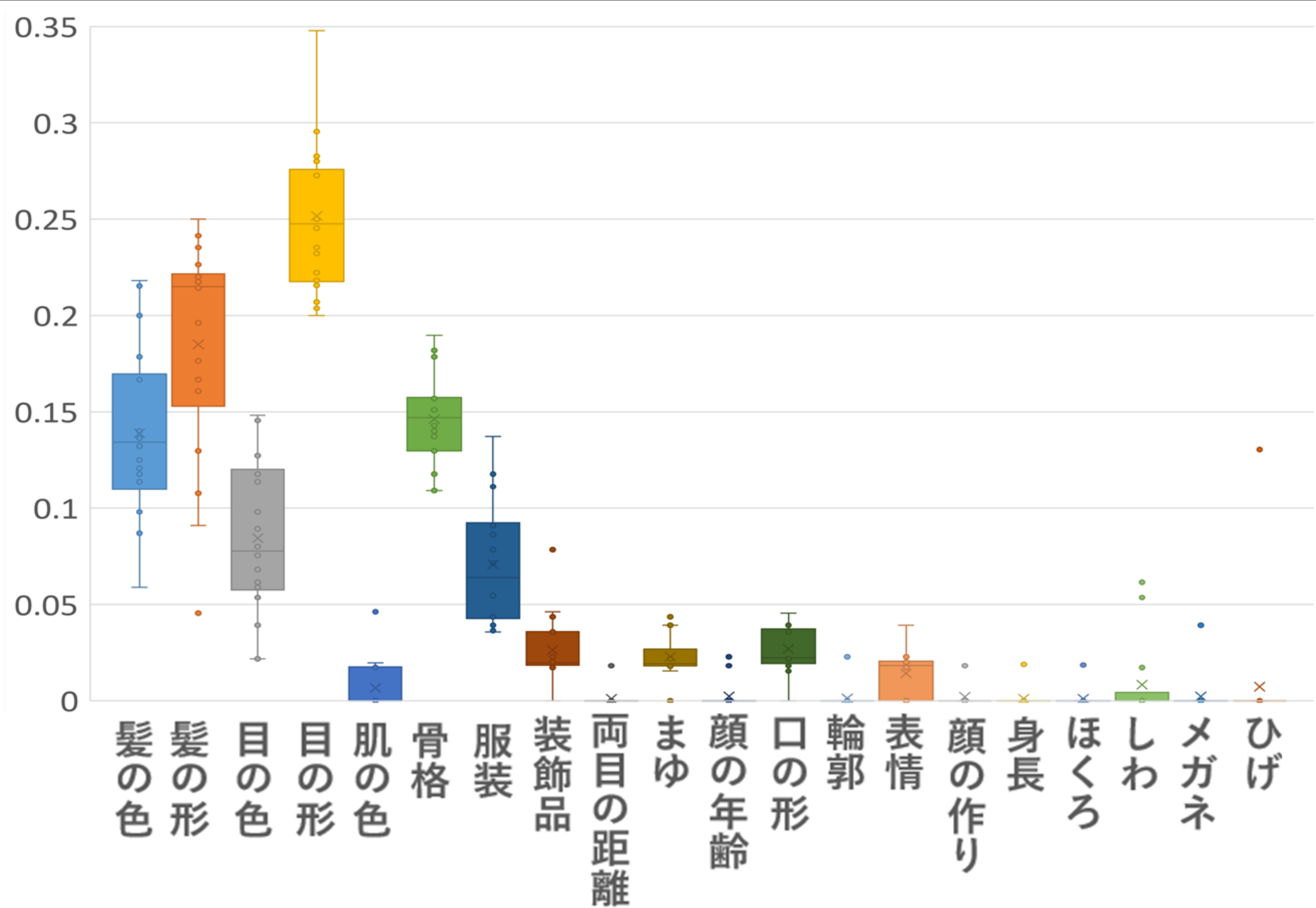


結果と考察

1つのキャラクタのイラストの中でそれぞれのパーツの割合を求め、18キャラクタ分をパーツ別に箱ひげ図で作成した

結果として、「**目の形**」「**髪の色**」「**骨格**」の割合が高くなっていた

また、キャラクタによっては「**ひげ**」や「**しわ**」などの**独自のパーツ**が一定の割合を占めていることが分かった



今後の展望

今後の展望として、キャラクタの声を想像するために必要なパーツのうち、**高い割合**を占めていた「**目の形**」「**髪の色**」「**骨格**」を用いてキャラクタの声の推定精度が向上するか検証し、キャラクタに合った声を生成できるか確かめる