

# 類似画像検索を用いた イラスト中キャラクター名のタグ付け支援

上野研究室 中川 雄貴

本研究では、アニメやマンガ等の有名キャラクターが描かれたイラストに対するタグ付け支援を目的とする。有名キャラクターは特定の髪や服の色、形を特徴として持っている事が多い。そこで、これらの特徴が類似するキャラクターは同一のキャラクターである可能性が高い事に注目し、類似画像検索を用いてイラスト中のキャラクターを判別する手法を提案する。

提案手法の評価の結果、同キャラクター間では類似画像検索による類似度が高く、異なるキャラクター間では類似度が低くなる事が確認できた。この結果から、提案手法がキャラクターの判別に有効であると考えることができ、実際にシステムを構築した。

システムは、入力されたイラスト画像内のキャラクターと髪や服の色、形が類似するデータベース内のキャラクター名を類似度の順位にしたがって出力する。システムの評価実験として、Webから収集した多量のキャラクター画像と名前をデータベースにした実験を行った。

実験の結果、いくつかの入力イラストに関して類似度順位が高い結果を得る事ができ、本研究の提案手法の有用性が確認できた。

はじめに.....	1
関連研究.....	2
Folksonomy	2
Auto-annotation	2
顔検出と顔認識	3
アニメキャラクターのイラスト画像分析	3
研究背景と提案手法.....	5
実写人物の特徴量とキャラクターの特徴量との相違	5
色の分布を用いたキャラクターの特定	5
提案手法	7
類似画像検索システムImgSeek	8
アニメ顔検出モジュールImager::AnimeFace	9
システム構成.....	10
支援モジュール	10
実験 1：認識手法の評価.....	12
概要	12
各画像データの準備	12
結果	13
考察 1.....	15
実験 2：システムを用いた実験.....	16
概要	16
各画像データの準備	16
キャラクターの選定	16
実験結果	17
考察 2.....	20
おわりに.....	22
謝辞.....	23

## 1. はじめに

近年、ブロードバンド回線の一般家庭への普及により、インターネット上では画像や動画、音声データなど、テキスト情報以外のコンテンツの利用が急速に増えている。また、“ニコニコ動画”<sup>[1]</sup>、“Youtube”<sup>[2]</sup>や“Flicker”<sup>[3]</sup>など、そのサービス内で画像や動画等のコンテンツを投稿・共有し、それを媒介として他のユーザーとの交流を行うWebサービスも多数存在しており、広く利用されている。

これらのサービス上では、サービス利用者の増加に伴って、扱われるコンテンツの種類や量も膨大となり、利用者が欲しい情報を見つけ出すのが困難となる。特に、画像や動画、音声などのコンテンツは、キーワードによって検索することができない。このような状況において、膨大かつ多様なコンテンツを効率よく管理し、必要な情報を入手する手法が必要とされている。このような手法の一つとして、タグを用いたコンテンツの分類、検索が行われている。

タグとはコンテンツに付加され、そのコンテンツの特徴を表す語の集合である。利用者は興味のあるキーワードを入力してタグ検索を行い、検索条件にかかるタグが付与されているコンテンツを得られる。これらのタグ情報は、コンテンツの投稿者や他のサービス利用者等の人手によって付けられている事が多い。

タグによる検索を用いる場合、各コンテンツごとに多様なタグが付与されていると多くの検索キーワードに対して検索にかかりやすく、また複数の語によって検索条件を絞り込む事ができる。しかしながら、膨大な量のコンテンツに対して、多くのタグを人手によって入力するのは困難である。

そこで、これまでもシステムによってタグを自動的に付与する、Auto-annotationと呼ばれる技術に関して研究が行われてきた。Auto-annotationとは、コンテンツの内容を解析することにより、コンテンツに対し自動的にタグを付加する技術である。

一方で、“pixiv”<sup>[4]</sup>や“PiXA”<sup>[5]</sup>に代表されるような、タグを用いたコンテンツ管理が行われているイラスト画像投稿・共有Webサービスに関しても、同様に多量のタグ入力のわずらわしさという問題が存在している。しかしながら、イラスト画像に特化したAuto-annotationに関する研究は存在しておらず、システムによるタグ付け支援が行われていない。

そこで本研究では、イラストに対応したAuto-annotationの研究として、様々な人々が描いたイラスト画像中のキャラクターを認識し、キャラクター名の表示を行うことを目的とする。特に国内のイラスト画像投稿・共有サービスにおいて、多くの利用者が著名な漫画やアニメなどの有名キャラクターを描いて投稿する点に着目する。実写画像に対して被写体となっている人物を認識し、その人物名のタグを付与する技術と同様に、キャラクターが描かれたイラスト画像に対してキャラクター認識を行い、データベースからキャラクター名等の情報を取得する手法を提案する。これにより、イラスト画像投稿・共有Webサービスにおいて、イラスト画像に対してタグ情報を付与する際の支援が可能となる。また、提案した手法を実装したシステムの構築及び評価を行い、提案手法の有効性や精度の向上について考察、検討を行う。

## 2. 関連研究

### 2.1. Folksonomy

Folksonomyとは、タグの用いられているWebサービス上において、コンテンツの作成者や投稿者、サービスの管理者等の特別な権限を持つ人物だけでなく、他の一般利用者が自由にタグを追加、削除、編集する事で多数の利用者の意見を反映したタグがコンテンツに付けられる分類法である<sup>6)</sup>。

タグによるコンテンツ管理、及びFolksonomyが導入されているWebサービスの代表的な例として、写真、イラストなどの画像投稿・共有サイトでは“Flicker”や“pixiv”，動画投稿サイトでは“ニコニコ動画”が挙げられる。

例えば、イラスト投稿・共有サービスである“pixiv”に投稿されている図1の画像には“オリジナル、入道雲、子ども、川、水、背景、親子、風景”の8個のタグが付いている。この画像の作成者は投稿時に“オリジナル、子ども、入道雲”の3つのタグしか付けずに投稿した。それ以外のタグはこのイラストを閲覧した別の利用者によって付けられた物である。

### 2.2. Auto-annotation

Auto-annotationとは、コンテンツに対して、“メタデータ”をシステムにより自動的に付ける技術である。メタデータとはコンテンツに付与され、コンテンツの内容を表すデータである。メタデータの例としてタグが挙げられ、一般的に広く利用されている例では、アーティスト名やアルバム名等の情報を音楽コンテンツに付けられるID3タグなどがある。

画像に対してメタデータを付与する場合はタイトルや制作者等の情報の他に、その画像が表現している事柄や、どんな物体が映っているか、どのような表現の画像か（実写か、絵画か、または記号か）等、画像の内容についての情報を持たせる事もある。このことから、画像に対するAuto-annotationは画像の自動分類や類似画像検索、画像認識といった画像内容に関する研究分野との関連性が高い。



図1 ”pixiv”におけるタグのついたイラスト画像例

画像に対して自動的にメタデータを付与する場合、各画像ごとに適切なメタデータの生成が必要となる。このメタデータの生成には、各画像の周辺に存在しているテキスト情報（例えば、ブログにアップロードされた写真とブログ記事の内容）を用いる方法や、画像内容を解析して得た画像の特徴をもとにメタデータを生成する方法がある。

画像の周辺に存在しているテキスト情報を利用する例としては、Smithらの研究がある。SmithらはWeb上を走査して画像を収集し、画像URLをもとに画像内容を推定し、データベースに各画像を分類する”Webseek“と呼ばれるシステムを構築し、その有効性を検証した<sup>[7]</sup>。しかし、この研究のシステムでは画像付近のテキスト情報に無い内容に対応して分類する事は難しく、このシステムで自動的にキャラクター名のAuto-annotationを行うならば、イラスト画像付近のどこかにキャラクター名情報が必要となる。

画像内容を解析してタグ付けを行うシステムとしては、画像間の類似を用いる方法がある。瀬崎らの研究<sup>[8]</sup>では、色の分布が類似している画像間では類似したタグが付いていると仮定した。既にタグの付いた画像データベースを利用し、入力画像と類似した画像に付けられたタグによってタグ入力の支援を行った。この手法は色の分布が画像の内容をよく反映している画像に対し、精度よく画像に関するタグを追加する事ができた。しかし、実験は実写画像に対して行われ、イラスト画像への有効性は示されていない。

また、顔認識などの画像認識・物体認識<sup>[9][10]</sup>の手法を利用する事によって画像から自動的にメタデータを生成する手法が提案されているが、これらの手法も主に実写画像が研究対象とされており、本研究で扱うようなイラスト画像への適用は行われていない。

### 2.3. 顔検出と顔認識

現在、広く一般的に顔認識と呼ばれるものは、1) 画像内でどの位置に顔が存在するか認識する、2) 画像内に存在する顔が誰であるかを認識（特定）する、の2つの意味で混同されている事が多い。本論文では、1) を“顔検出”、2) を“顔認識”とする。

顔認識は、静止画を用いた人物認識としては最も多く利用され、研究されている分野の一つである。生体認証として本人確認に用いられる他、写真管理ソフトウェアなどで画像内に映っている人物により写真を自動的に分類するなどの事例で用いられている。写真画像に対するAuto-annotationの分野でも用いられており、例えば、写真投稿・共有サイト”facebook”<sup>[11]</sup>や”picasa”<sup>[12]</sup>では、専用のアプリケーションと連携することによって投稿された写真に顔認識を適用し、写真内の人物名をタグとして付与することで、写真を自動的に管理・分類できる。

画像からの顔認識手法は、大きく分けて顔の部分的な構造の特徴（配置、距離）を用いる幾何学的特徴ベースの手法と、顔画像全体の濃淡を用いる見え方ベースの手法の2つに分けられる<sup>[13]</sup>。

### 2.4. アニメキャラクターのイラスト画像分析

アニメ等キャラクターのイラスト画像に対する画像分析の事例は少なく、未だ未発達分野である。河谷らの研究ではシステムによるキャラクターの顔の輪郭線や目の形状、顔に対する目の占める割合等に注目し、アニメらしさを定量的に評価した<sup>[14]</sup>。しかし、同研究では異なるイラスト制作者が描いた同じキャラクターの特徴量がどのような関係になるかは示されていない。

また、アニメ顔検出モジュールImager::AnimeFaceの制作者は類似キャラクター検索のシステムを作成し、Webサービスとして公開している<sup>[15]</sup>。このシステムでは、キャラクターの顔領域の検出を行い、髪の色などの顔の情報が類似したキャラクター画像の検索を行っている。しかし、キャラクターの名前などの情報は保持していない。

### 3. 研究背景と提案手法

#### 3.1. 実写人物の特徴量とキャラクターの特徴量との相違

イラスト画像に対して自動的にイラスト中のキャラクター名のタグを付与する手法の1つとして、イラスト画像中のキャラクターを認識し、その結果をタグとして付ける方法が考えられる。

イラスト画像に描かれる漫画やアニメなどの有名キャラクターは、特定の服装、髪の色・髪型、体格や顔のパーツの形・位置等の特徴によって表現されている。

実写人物の認識においても、主に目の位置や大きさ、口の位置などの顔のパーツの特徴を用いる方法がある。しかし、イラスト画像の場合、同じキャラクターであってもイラスト制作者の癖や好みによって顔の描かれ方が変わる。図2は、同一のキャラクターについて、それぞれ異なるイラスト制作者が描いた8枚のキャラクター顔部分である。目の大きさや口の形状などは各イラスト間で大きく変化し、統一性が無く、共通する点は髪の色と髪型であると言える。

また、1人の制作者が描いたキャラクターであれば、たとえ異なるキャラクターであっても、癖や好みにより顔が似る傾向にある。図3は、それぞれ異なるキャラクターについて、1人のイラスト制作者が描いた8枚のキャラクター顔画像である。目の大きさや形状、口の形状などが非常に似ており、大きな相違点は髪の色と髪型のみであると言える。

このように、キャラクターの目や口などの顔の部分的特徴はイラストの制作者による変動が大きく、2.2節で述べた幾何学的特徴ベースの人物認識手法をそのまま用いてキャラクター認識を行う事は難しい。

同様に2.2節で述べた、顔全体の濃淡を用いた見え方ベースの手法に関しても、イラストのキャラクターの顔には陰影が付けられる事が少なく、陰影が付けられていたとしても制作者によって陰影の付け方が異なり、キャラクターごとに決まった陰影の付け方は存在しない（キャラクターごとに陰影が描き分けられる事は無い）。よって、顔画像の濃淡によるキャラクター認識も難しい。

#### 3.2. 色の分布を用いたキャラクターの特定

一方で、異なる作者によるイラストであっても同じキャラクターである場合、服装や髪の色や形などの特徴は変化しにくい。例えば、服が赤いキャラクターは異なる作者によるイラストにおいても、多少の違いはあれど視覚的に赤と認識できる色で描かれる事が多



図2 異なる制作者による同一キャラクターの顔画像群



図3 同じ制作者による異なるキャラクターの顔画像群

い. このことから、構図が同じである場合、同じキャラクターが描かれた複数のイラスト画像はそれぞれの色の分布が類似すると考えられる。例えば、図4に示す2枚の画像は、同じキャラクターについて類似した構図で異なる作者が描いたイラストであり、色の明るさ等の微妙な違いはあるが、長い緑の髪や灰色の服装などの特徴により色の分布は類似している。構図が同様な別のキャラクターの画像を追加し、類似画像検索システムImgSeek<sup>[16]</sup>によって画像の類似度を測定した結果を表1に示す。ImgSeekは色の分布が類似している画像間に対して高い類似度を示す、類似画像検索のシステムとしては一般的なシステムである。表1から見て取れるように、入力画像と同キャラクターの画像間では類似度が高くなっている。このように、構図が類似する同キャラクターの画像間では、色の分布を用いた類似度は高くなる。

しかしながら、実際にはイラスト画像中のキャラクターは様々な構図や背景色などで描かれており、色の分布は構図などの影響を受ける。例えば、表2のNo.3に示す画像はNo.0と同じキャラクターが描かれ構図の異なるイラスト画像であり、No.2に示す画像はNo.0と異なるキャラクターが描かれ構図の類似しているイラスト画像である。これらの画像をシステムに入力すると、異なるキャラクターが描かれているにも関わらず、後者は前者の類似度を上回る。すなわち、同じキャラクターよりも同じ構図のイラストを類似画像として識別してしまう。



図4 色分布の類似した同キャラクターの画像

表1 構図の類似した画像間の類似度

No.	0	1	2	3
備考	入力画像	入力画像と同じキャラクター	入力画像と異なるキャラクターA	入力画像と異なるキャラクターB
画像				
入力画像との類似度	-	類似度 51.11%	類似度 13.33%	類似度 0.00%

表2 構図の異なる画像間の類似度

No.	0	1	2	3
キャラクター	入力画像	入力画像と同じキャラクター	入力画像と異なるキャラクター	入力画像と同じキャラクター
構図		入力画像と同じ構図	入力画像と同じ構図	入力画像と異なる構図
画像				
入力画像との類似度	-	類似度 50.76%	類似度 31.28%	類似度 21.13%

このように、色の分布によって画像の類似度を測定する類似画像検索システムは、画像の構図の影響により、同じキャラクターの描かれている画像を類似画像として出力できない場合がある。

### 3.3. 提案手法

本研究では、同一のキャラクターであれば、異なる制作者が描いても服装や髪型が統一される事に着目する。実写画像の人物認識に用いられている目・口の位置や濃淡の代わりに、キャラクターの髪・服装の色・形を用いるキャラクター認識手法を提案し、入力画像へのタグ付け支援システムに採用する。

キャラクターの認識の方法として、まず、多くのキャラクターの画像と名前を網羅したデータベースをあらかじめ用意し、次に、入力画像に描かれたキャラクターがどのキャラクターであるかデータベース内から特定を行う。これにより入力画像とキャラクター名の対応が付けられ、キャラクター名のタグ付け支援が可能であると考えられる。

キャラクターの特定には、3.2節で述べたように類似画像検索システムImgSeekを用いる。同じキャラクターであれば同様の髪や服の色・形で表現されるため、構図が統一されている場合、色の分布が類似する。よって、類似画像検索によって入力画像とデータベース内の画像全てとの色の分布の類似度を測定し、類似度の最も高いデータベース内の画像が入力画像に描かれているキャラクターである可能性が高い。

しかし、3.2節で述べているように、実際に存在するイラストは様々な構図で描かれており、そのままでは類似画像検索によってキャラクターの特定を行う事は難しい。

この影響を緩和するため、キャラクター顔検出を用いてキャラクターの顔位置を推定し、キャラクターの周りを切り出して新たな画像とする事で、キャラクター位置を揃えて構図の統一化を図る。本研究では、PerlモジュールImager::AnimeFace<sup>[17]</sup>を用いる。

まず、Imager::AnimeFaceはイラストなどの画像を入力すると、図5 a) のようにキャラクターの顔と推定される正方形領域を返す。

次に、推定された領域を切り出して類似画像検索に入力するが、切り出し前の処理として、検出された顔領域の拡大を行う。この処理により、キャラクターの顔だけでなく、服装や髪型が切り出し後の画像に含まれるようにする。図5 b) の場合、出力された正方形領域をそれぞれ各横方向に0.5倍、上方向に0.5倍、下方向に2.0倍に拡大した範囲である。以降、本論文ではこの拡大した範囲を、“キャラクター領域”と定義して説明を行う。ここで下方向への拡大率が高いのは、イラスト中キャラクターは画面下端から直立しているとし

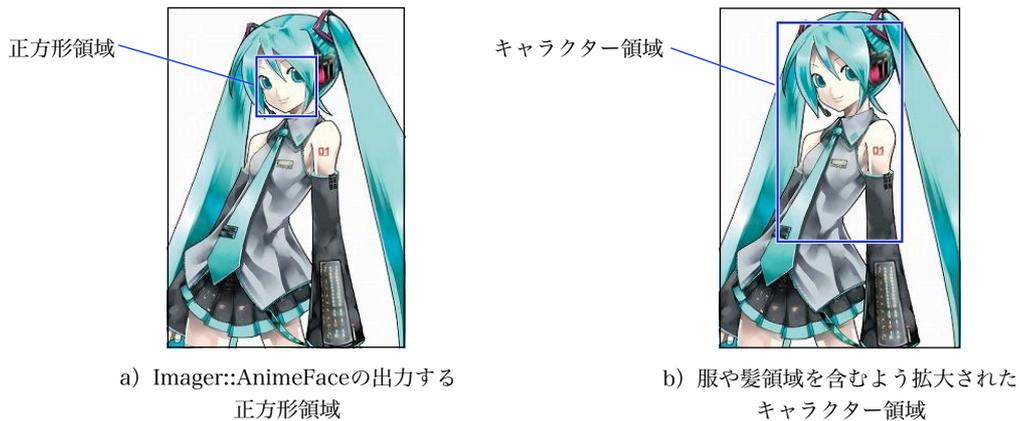


図5 Imager::AnimeFaceの検出する顔領域と拡大したキャラクター領域



図6 縦方向に各顔位置のそろったキャラクター領域

て、領域内にキャラクターの服装を含める目的がある。また、顔領域を相対的に拡大する事で、図6のように切り出した後の各画像は顔位置が揃う。この操作により類似画像検索において、データベース内の画像の構図を統一し、各画像の構図の違いによる検索精度の減少を防ぐ事が可能である。

Imager::AmineFaceの出力する正方形領域の大きさは入力画像によって変わるため、相対的に拡大を行っているキャラクター領域にも大きさの変動が生じる。しかし、事前の調査から入力画像とデータベース画像の大きさはImgSeekの出力する類似度に大きな影響を及ぼさないと判明しているため、画像の大きさの統一は行っていない。

### 3.4. 類似画像検索システムImgSeek

ImgSeekは、Haarウェーブレットの多重解像度解析により算出される画像特徴量を用いたコンテンツベース類似画像検索 (Content-based image retrieval : CBIR) システムである。コンテンツベース画像検索とは、画像自体を解析し、入力画像と内容が類似した画像を検索する技術である。ImgSeekのデータベースに画像を登録し、データベース内からクエリ画像 (類似した画像を検索したい問い合わせ画像) を選択すると、クエリ画像と色の分布や形が類似している画像をデータベース中から検索し、類似度順にソートされて出力される。図7の類似画像検索例では図の上部に表示された写真に対して、青い色領域が画像内の上半分を占める青空などの画像が類似画像として出力されている。本研究では、類似画像の抽出に本システムを用いる。

### 3.5. アニメ顔検出モジュールImager::AnimeFace

Imager::AnimeFaceは、Web上をクローリングして集められた多量のアニメや漫画などのキャラクターの顔画像をもとに、教師あり学習によって求められた識別関数を用いた画像認識用Perlモジュールである。入力画像を様々な位置と大きさの正方領域に分割し、その領域それぞれについて識別関数を用いてキャラクターの顔であるか否かを判別する事によって、イラスト中のキャラクターの顔位置と大きさを推定できる。本研究では、キャラクターの顔検出に本モジュールを用いる。

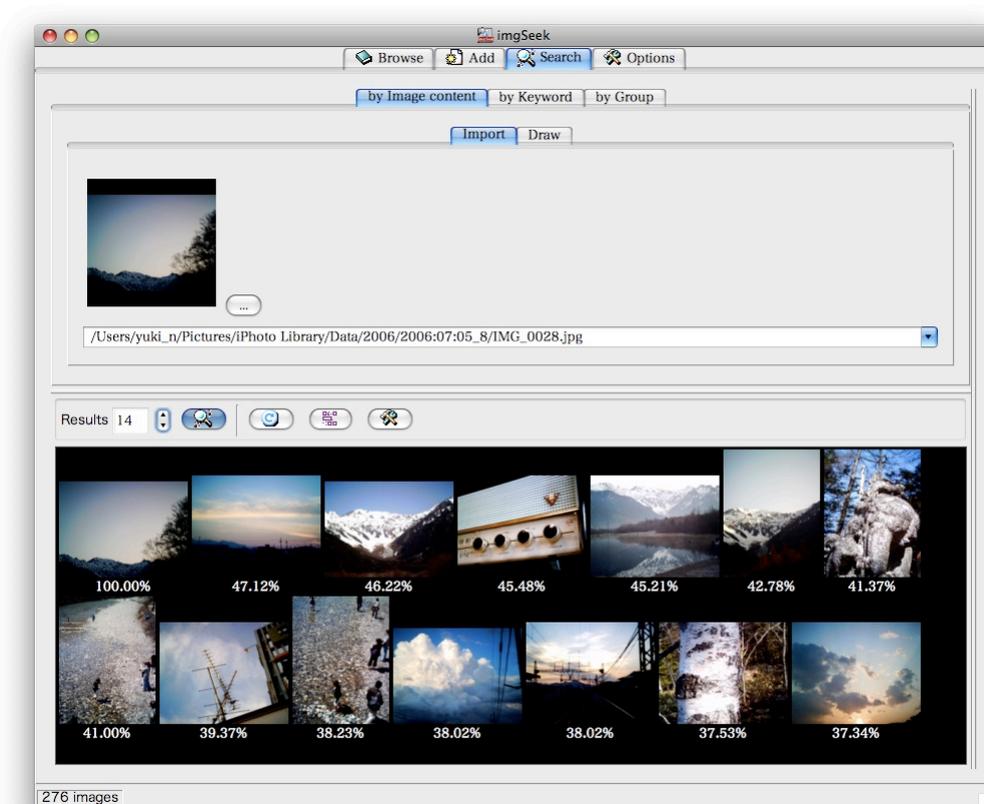


図7 ImgSeekの類似画像検索結果例

## 4. システム構成

入力されたイラスト画像に対して自動的にキャラクター位置の検出、及びデータベース内から入力画像に描画されたキャラクターと一致するキャラクターの特定を行い、キャラクター名を出力するシステムを作成した。図8に本システムの概要を示す。

本システムは、Perlによって実装された350行のプログラムである。まず、画像からキャラクターの顔位置を検出し、キャラクター領域を切り出す。次に、類似画像検索エンジンを用いてキャラクター領域と類似性の高い画像をデータベースから取得する。このデータベースは、本章で後述する支援モジュールによって構築された物を用いる。最後に、ImgSeekの出力結果と類似性の高い画像に付与されている名前と作品名を出力する。出力は類似度順にソートされたhtmlファイルまたはCSVファイルとして出力される。図9に本システムのhtmlファイルによる出力例を示す。図9の最上部に示された画像で検索を行った結果が、キャラクター名、作品名と共に示されている。

次節以降では本システムを構成する各モジュールについて説明する。

### 4.1. 支援モジュール

提案システムのデータベース構築を支援するため、Webページからキャラクター情報を取得するデータ取得モジュールと、取得した画像と各データをもとにしたデータベースを構築するデータベース構築モジュールを作成した。

データ取得モジュールはPerlによって実装された240行のプログラムである。本モジュールはAnime Characters Database<sup>[18]</sup>からキャラクターの顔画像や名前などの情報を取得するモジュールである。Anime Characters Databaseは多くの利用者の投稿によって構築された膨大なキャラクター画像や名前などのデータベースサイトであり、20,700キャラクター分のキャラクター名や作品名、外見的特徴などがキャラクターごとに記載されている。本モジュールはAnime Characters Databaseのhtmlタグ解析を行い、Webページ上から画像と名前などのデータを抜き出す。その際にCSVファイルに画像ファイル名とキャラクター名の対

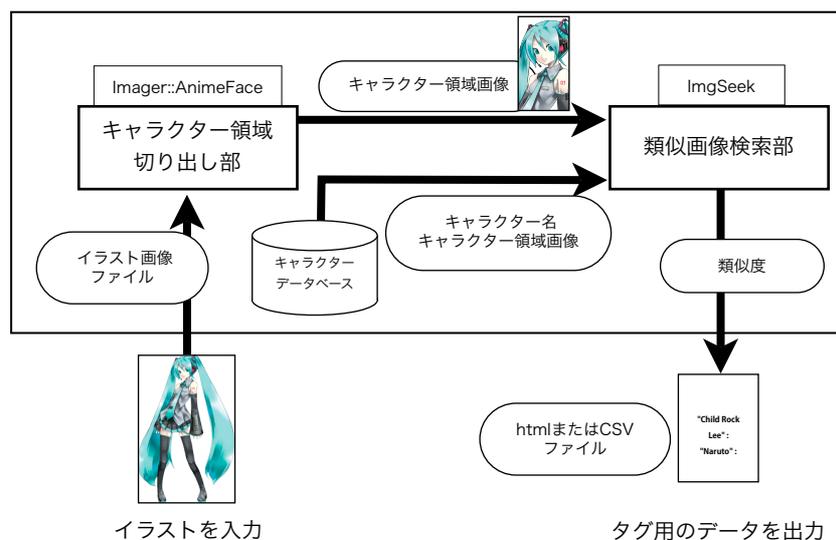


図8 システム全体の構成図

応を保持する。また、取得動作を有名アニメ、漫画、ゲームのキャラクターに限定する機能もそなえている。

データベース構築モジュールはPerlによって実装された120行のプログラムである。本モジュールはデータ取得モジュールで取得した画像に対してアニメキャラクター顔検出を適用し、各キャラクターごとにデータベース用のキャラクター領域画像を切り出すモジュールである。



図9 本システムのhtmlファイルによる出力結果例

## 5. 実験 1：認識手法の評価

実際に提案システムを構築する前に、本研究で提案したキャラクター認識手法の評価実験を行った。

### 5.1. 概要

3.3節で提案した、キャラクター認識手法の有効性を検証するため、精度評価の実験を行った。まず、同じキャラクターのキャラクター領域画像を集めた正解データと、それぞれ違うキャラクターのキャラクター領域画像を集めた不正解データを用意した。次に、各正解データ *Correct\_1~7* と他の正解・不正解画像データとの類似度を *ImgSeek* を用いて測定した。登録する各画像は全てイラスト制作者が違う物を用意し、制作者が違っていても同じキャラクター間であれば、類似画像検索の類似度が高く、異なるキャラクター間であれば類似度が低い結果が現れるか実験をおこなった。

### 5.2. 各画像データの準備

正解データの選定方法として、まず、”pixiv”の2009年12月16日付の“人気のタグ”のページから、キャラクター名として使われるタグのうち利用数1位のキャラクターを選択した。次に、制作者の違う同キャラクターの画像をWeb上から7枚収集した。この際には、全画像中の同キャラクターの服装が同じ画像を選択した。最後に、*Imager::AnimeFace*により自動的にキャラクター領域を切り出すツールを用いて、各画像に関してキャラクター領域画像を作成した。本実験では、多くの画像に対して頭全体から腰までがキャラクター領域に含まれるように、*Imager::AnimeFace*の出力する正方形領域の大きさから上方向に0.9倍、下方向に2.5倍、左右方向にそれぞれ0.35倍に拡大した領域を採用した。以降、この7枚の画像を正解データ画像 *Correct\_1~7* として扱う。図10に *Correct\_1~7* を示す。

不正解データの選定の際には、“人気のタグ”ページの上位キャラクターに同じ作品シリーズのキャラクターが多く含まれる事を考慮した。上位から単純にキャラクターを選ぶと不正解データのキャラクターに画風やデザインの偏りが現れる可能性があるため、“人気のタグ”ページのキャラクター名として利用されるタグのうち、利用数上位80個分のタグについて、そのタグ名のキャラクターの登場作品を調べた。これにより上位80キャラクターは17作品のキャラクターの集合である事がわかった。この結果を受け、正解データと違う7作品からタグ利用数の多いキャラクターを1人ずつ選択した。次に、7作品の各キャラク



図10 正解画像データ *Correct\_1~7*



a) *InCorrect\_1* b) *InCorrect\_2* c) *InCorrect\_3* d) *InCorrect\_4* e) *InCorrect\_5* f) *InCorrect\_6* g) *InCorrect\_7*

図11 不正解画像データ*InCorrect\_1~7*

ターについてWeb上から制作者の違うイラスト画像を収集した。最後に、正解データと同様に自動的にキャラクター領域を切り出すツールを用いて、各画像に関してキャラクター領域画像を作成した。以降、この7枚の画像を不正解データ画像*InCorrect\_1~7*として扱う。図11に*InCorrect\_1~7*を示す。

また、画像背景が類似度への影響を与える可能性を考慮し、影響を低減するために本実験で用いる全ての画像に対して画像編集ソフトウェア“GIMP”にてキャラクター以外の背景領域を白色に変更した。

### 5.3. 結果

各正解データ*Correct\_1~7*間の類似度測定結果を表3に、各正解データ*Correct\_1~7*と各不正解データ*InCorrect\_1~7*間の類似度測定結果を表4に示す。各類似度の算出方法としては、図12に示すように、各正解画像についてその正解画像以外の全ての正解画像及び不正解画像との類似度をSeekを用いて測定した。

表3 各正解データ*Correct\_1 ~ 7*と他の正解データとの類似度

	<i>Correct_1</i>	<i>Correct_2</i>	<i>Correct_3</i>	<i>Correct_4</i>	<i>Correct_5</i>	<i>Correct_6</i>	<i>Correct_7</i>	平均類似度
<i>Correct_1</i>	-	19.91	30.41	23.26	15.86	23.68	17.37	21.75
<i>Correct_2</i>	19.91	-	22.09	8.97	9.03	12.40	12.06	14.08
<i>Correct_3</i>	30.41	22.09	-	24.64	10.98	22.72	11.81	20.44
<i>Correct_4</i>	23.26	8.97	24.64	-	11.86	18.02	12.27	16.50
<i>Correct_5</i>	15.86	9.03	10.98	11.86	-	11.22	7.55	11.08
<i>Correct_6</i>	23.68	12.40	22.72	18.02	11.22	-	16.52	17.43
<i>Correct_7</i>	17.37	12.06	11.81	12.27	7.55	16.52	-	12.93

表4 各正解データ*Correct\_1 ~ 7*と不正解データとの類似度

	<i>InCorrect_1</i>	<i>InCorrect_2</i>	<i>InCorrect_3</i>	<i>InCorrect_4</i>	<i>InCorrect_5</i>	<i>InCorrect_6</i>	<i>InCorrect_7</i>	平均類似度
<i>Correct_1</i>	11.74	16.10	7.19	12.98	15.73	13.48	15.49	13.24
<i>Correct_2</i>	11.36	16.99	9.05	13.83	17.35	8.39	11.97	12.71
<i>Correct_3</i>	7.86	4.07	4.03	7.75	7.73	5.12	5.35	5.99
<i>Correct_4</i>	7.70	14.74	16.99	14.60	6.38	14.04	7.25	11.67
<i>Correct_5</i>	2.63	6.83	5.11	5.83	0.00	3.47	1.08	3.56
<i>Correct_6</i>	8.97	15.33	13.44	13.30	6.20	12.79	10.47	11.50
<i>Correct_7</i>	9.61	9.21	6.47	4.43	9.67	16.35	3.44	8.45

次章では本実験結果をより詳細に分析し、検定によるキャラクター特定アルゴリズムの有効性を検証する。

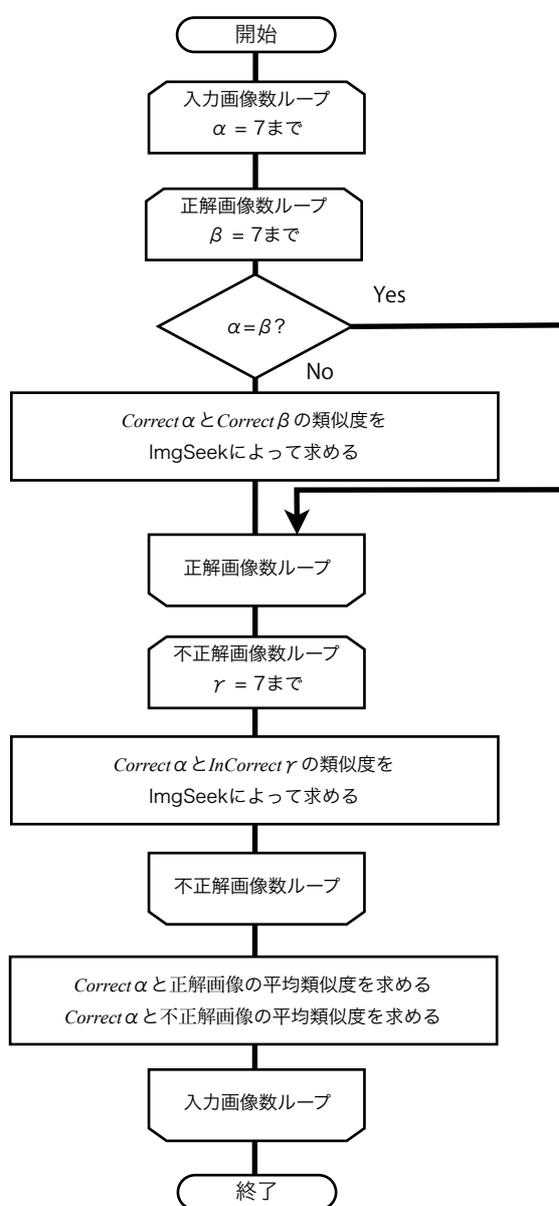


図12 実験1の手順流れ図

## 6. 考察 1

各正解データに対する正解データ群と不正解データ群の平均類似度とその標準偏差、及びWelchのt検定により求めたp値を表5に示す。

t検定は2つの標本群の平均値に有意な差があるか検定を行い、特にWelchのt検定は2つの標本群の母集団分散が異なる場合に用いられる。今回の実験の検定においては、各正解データ群・不正解データ群が正規分布に従うと仮定し、また分散はそれぞれ異なるものとして検定した。

結果として、7つの正解データのうち、*Correct\_1, Correct\_3, Correct\_5, Correct\_6*の4つの正解データの平均値について有意水準95%で有意差がみられた。この結果から、正解データによるが半数程度の入力画像に対して、提案手法によってキャラクターの区別が可能であると言える。

結果の内、*Correct\_2, Correct\_4*と*Correct\_7*に関しては類似度が低かった。これは色の違いが類似度低下の原因の一つと考えられる。例えば、これらの画像の髪の色は視覚的に見て全て緑と認識できる色だが、*Correct\_2*に関しては少し赤みがかっており、*Correct\_4*は暗く、*Correct\_7*は色が薄いなどの画像による色の差に影響を受けていると考えられる。

画像による若干の色味の違いの対処については、人間が同色であると知覚できる範囲の色を同色として扱い、画像に使われる色を減色する(青緑、深緑、黄緑を緑にまとめる等)等の方法が考えられる。この際には、微妙な色の違いによって区別されているキャラクターを区別しづらくなる事を考慮に入れる必要がある。

また、検出したキャラクターの顔の輪郭や、目の大きさ、位置などからキャラクターの頭身を推定できる可能性があり、これによって切り出し範囲を調整する事により、キャラクターの頭身の違いによる影響を低減できると考えられる。例えば、6~8頭身程度の頭身の高いキャラクターは目は小さく、位置は高く、顔の輪郭は縦に長く描かれる傾向にあり、3~5頭身程度の頭身の低いキャラクターは目の大きく、位置は低く、顔の輪郭は丸く描かれる傾向にある。Imager::AnimeFaceは目の大きさや位置を検出する機能も備えているため、検証の手間も少ない。しかし、3.1節で述べている通り、目や口などの顔のパーツ特徴は描き手による変動が大きい事を考慮に入れる必要がある。

表5 各正解データに対する2群の類似度平均とt検定のp値

	正解データ群		不正解データ群		Welchのt検定 p値
	平均類似度	標準偏差	平均類似度	標準偏差	
<i>Correct_1</i>	21.75	5.26	13.24	3.12	0.0086
<i>Correct_2</i>	14.08	5.60	12.71	3.55	0.6185
<i>Correct_3</i>	20.44	7.60	5.99	1.75	0.0049
<i>Correct_4</i>	16.50	6.49	11.67	4.38	0.1579
<i>Correct_5</i>	11.08	2.83	3.56	2.51	0.0005
<i>Correct_6</i>	17.43	5.14	11.50	3.13	0.0393
<i>Correct_7</i>	12.93	3.58	8.45	4.30	0.0652

## 7. 実験 2：システムを用いた実験

5章の実験および6章の考察において、提案システムのキャラクター判別に有効性が見られたため、実際にシステムを構築し、多くのキャラクターが登録されたデータベースを用意して評価実験を行った。本章ではその手順と詳細、および結果を示す。

### 7.1. 概要

まず、多くのキャラクターのデータセットを取得するため、データ取得モジュールを用いて3,645キャラクター分の画像データと、キャラクター名等のデータをAnime Characters Databaseから取得した。次に、データベース構築モジュールを用いて、取得した各画像から2,437データのキャラクター領域を切り出し、提案システムのデータベースに登録した。その上でデータベースに登録されたキャラクターから4キャラクターを選び、各キャラクターが描かれたイラスト画像6枚ずつを用意した。最後に、用意した24枚のイラスト画像を提案システムに入力し、類似度が上位200位までのキャラクター画像とキャラクター名が記載されたhtmlファイルとして出力した。

### 7.2. 各画像データの準備

本実験では、提案システムに登録されキャラクターの特定に用いられるデータベース画像と、提案システムに入力される入力イラスト画像を用いる。

データベース画像とは、Anime Characters Databaseから取得した3,645データ分のイラスト画像に対し、データベース構築ツールを用いてキャラクター領域を切り出した画像である。本実験ではImager::AnimeFaceの出力する正方形領域の大きさから、上方向に0.3倍、下方向に2.0倍、左右方向にそれぞれ0.2倍に拡大したキャラクター領域を採用した。実験1のキャラクター領域とサイズが異なる理由としては、背景色の除去を行わないため、頭の上や首周りの背景色の影響を低減させる目的がある。キャラクター顔検出の際に、1,208データについては顔検出に失敗した。また、3データについては取得時に画像データが壊れていたため、除外した。結果としてキャラクター領域を切り出す事のできたデータ数は2,437データである。これらのデータベース画像を提案システムに登録した状態で以降の実験を行った。

### 7.3. キャラクターの選定

実験は提案システムのデータベースに登録した2,437キャラクターから、4キャラクターを選択し、実験対象とした。図13に4キャラクターの画像を示す。以降、この画像を正解データ *CorrectA~D* として扱う。各キャラクターの選択基準として、実際にイラストに描かれる事が多いキャラクターに対し、正しいタグを精度良く付けられる事が望ましい。そこで5章の実験と同様に、“pixiv”において多く利用されているタグが並べられた“人気タグ一覧”のページ上から、キャラクター名として利用されているタグを4キャラクター分選択した。この際に、1) 同一の作品からは一人しか選ばない、2) 作品名内に含まれるキャラクター名である場合、作品名の略称としてキャラクター名と同じ物が使われる頻度が高い (“ルパン三世”は作品名略称として“ルパン”が用いられる、等) ため、そのキャラクター



図13 正解画像データ *CorrectA~D*

は使用しない，この2つの条件を満たすようにキャラクターを選択した．画像はAnime Characters Databaseから取得し，提案システム内に登録されている画像を用いた．

その後，選択した各キャラクターに対して，制作者の異なるイラスト画像を6枚ずつWeb上から用意した．

#### 7.4. 実験結果

用意した4キャラクター各6枚の画像をシステムに入力し，画像ごとの類似度や類似度順位，キャラクターごとの平均類似度・平均類似度順位を算出した．図14に示す通り，算出の際には1キャラクターについて用意した6枚の画像を1枚ずつシステムに入力し，各画像ごとに得られた類似度と類似度順位の値を平均する処理を4キャラクターについて行った．

入力した24枚のイラスト画像全てにおいて顔検出漏れは発生せず，すべてのキャラクターについてキャラクター領域が切り出された．以降，切り出されたキャラクター領域画像の内，*CorrectA*と同一キャラクターのキャラクター領域画像6枚を*QueryA\_1~6*として扱う．*QueryB\_1~6*や*QueryC\_1~6*，*QueryD\_1~6*についても同様である．

また，キャラクター顔検出の結果，*QueryB~D*の各画像6枚のうち1枚ずつ，計3枚についてキャラクター顔検出にて誤検出が発生し，全体で24出力されるべきhtmlデータが27となった．このように一枚の画像から誤検出により2つ以上のキャラクター領域画像が切り出された場合には，提案システムの出力順に*QueryB\_3a*，*QueryB\_3b*，...のようにギリシャ文字を末尾につけて扱う．

表6に，*CorrectA*と*QueryA\_1~6*との類似度，及びデータベース内の他のデータと比べた際の類似度の順位，類似度が1位の画像を示す．同様に，表7~9に*QueryB~D*について示す．

今回は全体で2,437データある中で，結果を扱いやすいデータサイズにするため類似度が上位200位までのhtmlデータの生成を行った．この処理により順位及び類似度が取得できなかった入力画像は27画像中8であった．誤検出によって本来の顔領域以外の箇所が顔領域として切り出された3枚の画像 (*QueryB\_3b*, *QueryC\_6b*, *QueryD\_3b*) に関しては全て順位が取得できなかった．

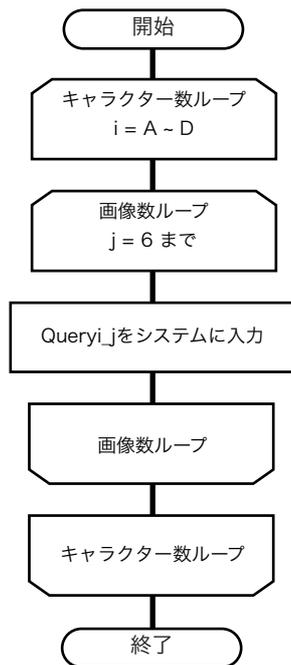


図14 各画像の類似度及び順位算出流れ図

表6 *CorrectA*と*QueryA\_1~6*の類似度および順位

<i>Query</i>							
	<i>QueryA_1</i>	<i>QueryA_2</i>	<i>QueryA_3</i>	<i>QueryA_4</i>	<i>QueryA_5</i>	<i>QueryA_6</i>	
類似度 [%]	-	23.64	22.30	31.60	20.47	-	
順位	201以下	164	127	26	41	201以下	
一位画像							
類似度 [%]	31.34	36.60	34.19	38.40	26.94	35.74	

表7 *CorrectB*と*QueryB\_1~6*の類似度および順位

<i>Query</i>							
	<i>QueryB_1</i>	<i>QueryB_2</i>	<i>QueryB_3</i> $\alpha$	<i>QueryB_3</i> $\beta$	<i>QueryB_4</i>	<i>QueryB_5</i>	<i>QueryB_6</i>
類似度 [%]	34.63	44.09	23.62	-	34.89	35.35	40.49
順位	85	1	60	201以下	18	7	7
一位画像							
類似度 [%]	48.52	44.09	31.50	38.42	44.12	39.86	46.66

表8 *CorrectC*と*QueryC\_1~6*の類似度および順位

<i>Query</i>							
	<i>QueryC_1</i>	<i>QueryC_2</i>	<i>QueryC_3</i>	<i>QueryC_4</i>	<i>QueryC_5</i>	<i>QueryC_6</i> $\alpha$	<i>QueryC_6</i> $\beta$
類似度 [%]	19.44	19.40	20.10	25.30	-	-	-
順位	104	163	149	19	201以下	201以下	201以下
一位画像							
類似度 [%]	31.89	32.43	35.44	29.85	33.73	38.18	37.63

表9 *CorrectD*と*QueryD\_1~6*の類似度および順位

<i>Query</i>							
	<i>QueryD_1</i>	<i>QueryD_2</i>	<i>QueryD_3</i> $\alpha$	<i>QueryD_3</i> $\beta$	<i>QueryD_4</i>	<i>QueryD_5</i>	<i>QueryD_6</i>
類似度 [%]	28.55	31.58	28.31	-	29.13	-	20.34
順位	20	7	2	201以下	15	201以下	26
一位画像							
類似度 [%]	38.13	40.67	38.40	42.07	35.80	36.69	27.33

## 8. 考察 2

本章では、7.4節の結果に対する考察を行う。表10に各キャラクターごとの平均順位、平均類似度、最高順位、標準偏差、類似度順位201位以下の結果数と誤検出数を示す。平均順位に関しては順位201位以下のデータを除いた値であり、順位201位以下の結果数に関しては、キャラクター領域の誤検出によって順位201位以下となっている結果数を除いている。

24画像を入力した結果、順位201位以下となった画像は全体の2割程度であり、それ以外の画像の平均順位は54.8位であった。画像ごとに個別に見ると、QueryB\_2は1位であり、また QueryB\_5, QueryC\_6, QueryD\_2, QueryB\_3aの4つのデータに関しても10位以上となる高い順位の結果が得られた。この結果は、キャラクターの髪や服の形や色を用いる提案手法によって、キャラクターの特定ができる可能性がある事を示唆している。しかし、入力イラストのキャラクター名のタグを正しく付与するためには、正解となる画像との類似度順位が1位となる必要があり、現状の精度では入力イラストに対して自動的にタグ付けを行う事は難しい。

各キャラクターについての結果は、全体としての平均順位より各キャラクターについての平均順位の方が高いQueryBとQueryDのグループと、各キャラクターについての平均順位の方が低いQueryAとQueryCのグループとの2つのグループに分けられる。グループ全体の平均順位は、QueryBとQueryDのグループが22.6位、QueryAとQueryCのグループが99.1位であった。

その他のデータについても同様の事が言え、QueryBとQueryDのグループは全体の標準偏差より各キャラクターについての標準偏差の方が小さく、最高順位も平均より高い。QueryAとQueryCのグループは全体の標準偏差より各キャラクターについての標準偏差の方が大きく、最高順位も平均より低い。このことから、QueryBとQueryDのキャラクターに関しては安定して上位に結果が現れる傾向にあり、精度が高く認識できるといえる。一方、QueryAとQueryCのキャラクターに関しては順位が安定せず、認識精度が悪いといえる。QueryAとQueryCのキャラクターの順位が低くなる要因の1つとして、データベース画像CorrectAやCorrectCの質の問題が考えられる。

異なるイラスト制作者がキャラクターのイラストを描く際には、何らかのイラストを参考にすると考えられる。この際には、同じイラストを参考にする事もあれば、それぞれ異

表10 各キャラクターごとの結果

	平均順位	平均類似度	最高順位	標準偏差	順位201以下	誤検出
QueryA_1~6	89.5	24.50	26	66.69	2	0
QueryB_1~6	29.7	35.51	1	34.55	0	1
QueryC_1~6	108.8	21.06	19	64.91	2	1
QueryD_1~6	14.0	27.58	2	9.67	1	1
全体平均	54.8	28.06	12	58.16	合計 5	合計 3

なるイラストを参考にする事もあると考えられる。その参考にしたイラストにより、同じキャラクターであっても色や髪型等の細部に影響を受ける可能性が高い。特に、本実験の *Correct\_A*はこのキャラクターの元作品のイラストであるが、この絵以外に多くの人々がこのキャラクターのイラストを描いており、*Query\_A1~6*はこの絵とは異なるイラストを参考にして描かれた可能性がある。また、元作品のイラストの画風等が時代とともに移り変わる事もあり、こちらも参考元の絵が異なるため、同様に対応が必要である。しかし、現状のシステムでは類似画像検索に用いられるデータベースの画像が1枚しか登録されていない。この問題を解決するためには、1キャラクターについてキャラクター領域画像を複数登録することや、複数の同じキャラクター領域画像を合成して構図や色が平均的な画像を生成する事で、そのキャラクターに対する表現の移り変わりや、描き手による色の微妙な違いへの対応が考えられる。

また、類似した服装や髪型のキャラクターが多く存在するキャラクターに関しては、正解データと不正解データの区別が難しくなる。例えば、学生服のキャラクターが存在する作品は非常に数が多く、また1作品あたりの学生服のキャラクターの登場数も多い。この様なキャラクター群に対して、正確な認識を行うように対策を行う必要がある。

本実験を行う際にはキャラクター領域の大きさが実験1とは異なり、小さくなっている。しかし、その影響により単純に実験1との比較ができなくなったため、データベースに登録されている画像数の増大による結果への影響についての情報が得られない事は実験上の不備であった。

## 9. おわりに

本研究では、イラストに対応したAuto-annotationの基礎研究として、様々な人々が描いたイラスト画像中に描画されているキャラクターを認識し、キャラクター名を表示することを目的とした。特に国内のイラスト画像投稿・共有サービスにおいて、多くの利用者が著名な漫画などの有名キャラクターを描いて投稿する点に注目して研究を行った。実写画像に対して被写体となっている人物を認識し、その人物名のタグを付与する技術と同様に、キャラクターが描かれたイラスト画像に対してキャラクター認識を行い、データベースからキャラクター名等の情報を取得する手法を提案し、評価した。これによりイラスト画像投稿・共有Webサービスにおいて、イラスト画像に対してタグ情報を付与する際の支援をする事が可能となる。また、提案手法をシステムとして実装、評価を行い、提案手法の有効性や精度の向上について考察、検討を行った。

提案手法の評価実験として、同じキャラクターのイラストを集めた正解データ群と、異なるキャラクターの画像を集めた不正解データ群に対して、正解データ内の1枚と、残りのデータとの類似度を測定した。その結果、正解データとの類似度平均値は不正解データとの類似度平均値より高く、またこの2組の類似度平均値に有意差がみられた。これにより、認識手法がキャラクターの判別に有効であることがわかった。

システムの評価実験として、提案システムのデータベースとしてWeb上から収集した2,437キャラクター分のデータセット中を用意し、4キャラクターの24枚の入力イラストについて精度の評価を行った。いくらかのデータに関しては精度よく結果があらわれたが、類似度1位となった結果は1件のみであり、有用なタグを付与するためには、更なる改善が必要と考えられる。

今後、より精度を上げるため、背景色などの影響を大きく受ける問題に関し対応していく必要がある。今回作成したシステムのアルゴリズムは、全ての色に対して類似度の測定を行うため、背景色と髪の色などは同等に扱っている。キャラクターの類似度を測定する際に髪や服の色に対して重み付けを行って類似度の測定を行うことにより、認識精度が向上すると考えられる。

また本研究では、1つのイラストに対し1人のキャラクターが描かれているとして認識を行ったが、実際のイラストには複数のキャラクターが描かれている事も多い。この場合、キャラクター間の関係を考慮する事によって精度の向上が望めると考えられる。今回作成したシステムにおいても、キャラクター名の情報以外に作品名の情報を取得している。そこで、1枚のイラスト中に描かれた複数のキャラクターは同じ作品のキャラクターであると仮定して、同作品に含まれるキャラクターの中で最も類似度の高いキャラクターの組み合わせになるように認識を行う事により、認識精度の向上が見込めると考えられる。

## 謝辞

本論文を執筆するにあたり，忙しい中にも関わらず推敲に力添えをしてくださった上野秀剛先生，及び，査読して下さった松尾先生に感謝します。

また，共に議論を交わして下さった研究室の方々，発表の場にて本研究に対し質疑して下さった山口智宏先生方，本論で使用させていただいたImager::AnimeFace開発者のultraist(twitterユーザー名)氏，ImgSeek開発者の方々，イラスト制作者の方々，研究テーマをひらめく元となった初音ミクさんに感謝します。

## 参考文献

- [1] <http://www.nicovideo.jp/>
- [2] <http://www.youtube.com/>
- [3] <http://www.flickr.com/>
- [4] <http://www.pixiv.net/>
- [5] <http://www.pixa.cc/>
- [6] A. Mathes : “Folksonomies - cooperative classification and communication through shared metadata, ”Computer Mediated Communication, UIC Technical Report (2004).
- [7 ] John R. Smith, Shih-Fu Chang : “Visually Searching the Web for Content, ” IEEE MultiMedia, Vol.4, pp.12-20 (1997).
- [8] 瀬崎直人, 黄瀬浩一 : “タグの共起と類似画像を利用したタグ付け支援システム, ” IEEJ transactions on electronics, information and systems, Vol.129, No.3, pp.475-480 (2009).
- [9] 柳井啓司 : “一般物体認識の現状と今後, ” 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol.48, No.SIG 16 (CVIM 19) (2007).
- [10] 中山 英樹, 原田 達也, 國吉 康夫, 大津 展之 : “画像・単語間概念対応の確率構造学習を利用した超高速画像認識・検索方法, ”Technical report of IEICE. PRMU, Vol.107, No.384, pp. 65-70 (2007).
- [11] <http://www.facebook.com/>
- [12] <http://picasa.google.co.jp/>
- [13] 岩井 儀雄, 勞 世竝, 山口 修, 平山 高嗣 : “画像処理による顔検出と顔認識, ” 情報処理学会研究報告. CVIM, No.38, pp.343-368, (2005).
- [14] 河谷 大和, 柏崎 礼生, 高井 昌彰, 高井 那美 : “アニメキャラクターの特徴抽出に基づくアニメ度の評価, ”情報処理学会研究報告. グラフィクスとCAD研究会報告, IPSJ SIG Notes, Vol. 2008, No.80, pp.35-38, (2008).
- [15] <http://anime.udp.jp/face-search-v1/>
- [16] <http://www.imgseek.net/>
- [17] <http://anime.udp.jp/imager-animeface.html>
- [18] <http://www.animecharactersdatabase.com/>