


Title	イラスト推薦と可変グリッドを用いた模写練習支援システムの提案
Author(s)	金山, 春香; 謝, 浩然; 宮田, 一乗
Citation	情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション, 2023-HCI-202(43): 1-8
Issue Date	2023-03-06
Type	Journal Article
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/18202">http://hdl.handle.net/10119/18202</a>
Rights	<p>社団法人 情報処理学会, 金山春香, 謝浩然, 宮田一乗, 情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション, 2023-HCI-202(43), 2023, 1-8. ここに掲載した著作物の利用に関する注意 本著作物の著作権は情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	 <p>The logo for JAIST (Japan Advanced Institute of Science and Technology) is displayed. It features the acronym 'JAIST' in a large, stylized blue font. Below it, the full name 'JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY' is written in a smaller, blue, sans-serif font. A green circular element is positioned at the bottom left of the logo.</p>

# イラスト推薦と可変グリッドを用いた 模写練習支援システムの提案

金山 春香<sup>1,a)</sup> 謝 浩然<sup>1</sup> 宮田 一乗<sup>1</sup>

**概要:**近年、タブレット PC などのデジタルデバイスの普及により、独学でデジタルイラスト制作を行う人が増えている。しかし、初心者に対してデジタルイラスト練習を支援するシステムは少ない。イラスト練習でよく用いられる模写は観察が重要であるが、既存の模写支援システムは、観察による結果をガイドとして示すことで模写を支援しているものが多く、観察自体を支援しているものは少ない。本研究では、模写で用いるグリッドを可変にすることで、観察に焦点を当てた効果的な模写支援を可能にするという仮説に基づいて、ラフスケッチを入力とした参考イラストの推薦と、模写における可変グリッドシステムを組み合わせた模写練習支援のフレームワークを提案する。本システムでは、ユーザの描いたラフスケッチを入力とし、データセットから特徴量の近い類似イラストを候補として提示する。ユーザはその中から好ましいイラストを選び、グリッドの細かさを変化できる模写インターフェースにより模写を行う。評価実験とアンケート調査の結果、本システムの模写練習支援における有用性を確認できた。

## 1. はじめに

近年、SNS の発展により、イラスト制作についての情報が得やすくなっている。また、タブレット型 PC やペンタブレットなど、ペン型デバイスで操作することのできる製品の普及や、ペイントソフトなど、デジタルイラスト制作のためのツールの普及が進んでいる。このように、デジタル環境下でのイラスト制作に必要な機材が身近になっていることから、デジタルイラスト制作を行う人が増えている。ペンタブレットメーカーである wacom 社が 2020 年に行った調査 [1] では、イラスト制作において、ユーザの 85 % がデジタル画材 (PC, タブレット, イラスト制作ソフト等) を取り入れており、66 % が制作の半分以上においてデジタル画材を使用していることがわかった。

イラストを上達するための練習として、模写がある。プロのイラストレータの作品を観察し再現することで、作風や意図を理解することが目的である。単なるトレースでは観察が伴わないので練習としては適さない。模写とトレースの違いを図 1 に示す。グリッドとよばれる画面を水平線と垂直線で分断した格子状の線を手本と自らのキャンバスに追加することで補助を行うこともある。このグリッドは、自画像を描くときなどにも使用される。また、他の練習法としてデッサンやクロッキーがあるが、それらは輪郭

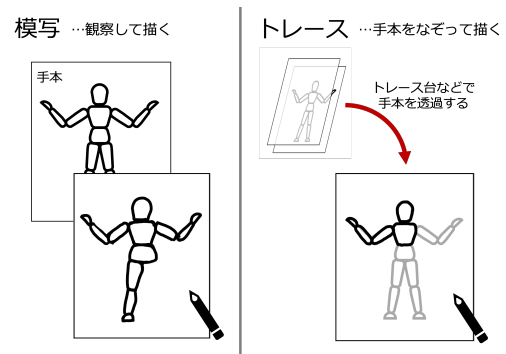


図 1: 模写とトレースの違い

線のない立体を 2 次元のキャンバスに起こすため、線を選ぶのが初心者にとって難しい。そのため、初心者が初めて行うのに最適なイラストの練習法は線がはっきりと描かれているイラストの模写を行うことであると考えられる。

模写を行うためには、参考となるイラストが必要である。キャラクタイラスト制作において参考に用いられるのは、ポーズ集や 3D モデルなどが挙げられる。しかし、それらを活用するにはある程度の知識や経験が必要になるため、初心者にとってはハードルが高いと考えられる。また、キャラクタイラストにおいては、顔が人物の印象の多くを決定してしまう [2] ため、特に顔の作画のクオリティやバランスが重要となる。そのため、初心者はまずキャラクタ顔を練習することで、イラストの上達を実感できやすく、モチベーションの維持につながると考えられる。

<sup>1</sup> 北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology  
<sup>a)</sup> haru-kana@jaist.ac.jp

キャラクターのイラスト制作に関する研究がいくつか報告されている。例えば、元となるキャラクターの顔イラストと生成したい表情のランドマーク(顔のパーツや輪郭を決める特徴点)からその表情をしたキャラクタイラストを生成する手法 [3] や StyleGAN を用いて全身イラストを生成する手法 [4] がある。しかし、これらの手法はインタフェースが用意されておらず、初心者にとって使いやすさとはいえない。インタラクティブに画像の生成を行う手法も報告されている [5]。この手法では、スケッチベース画像編集ネットワークを用いて、手描きスケッチから実世界に基づいた画像の生成や編集をすることができる。このように、イラスト生成を行う研究は行われているが、イラスト練習を支援する研究は少ない。

また、デジタル環境で模写を行う場合であっても、参考イラストを見つけることとそのイラストと模写で用いるキャンバスにグリッドを描画する必要がある。そのためには、参考イラストをイラストサイト等で探したり、ペイントソフトの理解が必要であることから、初心者にとってはハードルが高いと考えられる。

本研究では、これらの問題を解決するとともに効果的な模写支援を行うための、デジタル環境における初心者向け模写練習システムを提案する。模写で必要な参考イラスト探しから完成イラストの比較まで支援することで、初心者にとって使いやすいシステム構築を目指す。

## 2. 関連研究

### 2.1 イラスト制作支援

イラスト制作支援の研究事例として、ShadowDraw [6] がある。この研究では、ユーザの描いている対象の関連画像を大規模データセットから検索することで複数見つけ出す。それらの画像を合成したものを、ユーザの描いているキャンバスに影としてガイドを表示する。ユーザの描画にリアルタイムで適応し、影を更新することができる。画像検索はヒストグラムを用いているが、ハッシュ化を行うことで局所的と大域的な類似性を高速に求めることが可能となった。

デジタルイラスト制作における自動レイヤ分けシステム [7] がある。デジタルイラストでは、レイヤとよばれる層に分けてイラストを描くことができる。線画と着色、背景とキャラクターなどを別のレイヤに分けて描くことで、描画や修正をより簡単に行うことができる。レイヤ分けは、ペイントソフトの塗りつぶしや選択範囲ツールを用いて自動で行うこともあるが、描写が細かい箇所は失敗することも多く、手動で修正する必要がある。この研究では、pix2pix [8] を用いて、キャラクターのパーツ毎に学習を行うことで、パーツを認識し、レイヤ分け作業の自動化を行っている。

また、アーティストの作品制作過程をモデル化した画像生

成・編集フレームワーク [9] が提案されている。CGAN [10] をベースにして制作過程を段階ごとに学習することで、作品の制作支援を行っている。また、既存の作品の制作過程を推測することができ、制作の途中過程を編集することで、自然かつ元とは異なる作品を生成することができる。

このように、既存のイラスト制作支援はユーザのイラスト練習を目的とはしておらず、手本をなぞることでクオリティの高いイラストを制作することやイラスト制作における作業を自動化することで支援を行っている。本研究では、模写を用いることで手本イラストを直接ガイドとして与えることなく、観察を伴ったイラスト練習支援を行う。

### 2.2 模写支援

実人間の模写を通して、模写を行う上で重要な観察を支援する研究がある [11]。この研究では、イラスト描画の初心者の問題として挙げられる絵画対象の観察をおろそかにするという点を解決するために、自身の絵の修正点への気づきを促すシステムを提案した。提案システムでは、観察したことを言語化することによって意識的に観察を促すことや、部分的に遮蔽することでより深い観察をさせるシステムを実装している。

また、模写の支援として、様々なガイドを用いて描画を支援するシステム [12] がある。このシステムで用いられているインタフェースでは、既存研究とは異なり、ガイドとして模写対象自体を用いることはせず、模写対象から抽出した幾何学的特徴からガイドを作成する。

本研究はキャラクタイラストを対象にして、模写で用いるグリッドを可変にすることが、観察に焦点を当てた効果的な模写支援を可能にするという仮説に基づく。そして、ラフスケッチを入力とした参考イラストの推薦と、模写における可変グリッドシステムを組み合わせた模写練習支援のフレームワークを提案し、仮説の検証を行う。

## 3. 提案システム

提案システムでは、ユーザによるスケッチを入力として、それを元としたキャラクタイラストを推薦する。また、スケッチの作成・編集と生成したイラストを表示するインタフェースを作成する。イラスト画像の推薦においてはラフスケッチから線画の洗練化 [5] を行い、入力とする。入力画像の特徴量を抽出し、データセット内の画像と比較し、類似度が高い上位5枚のイラストを推薦する。その中からユーザが手本にしたいイラストを選択し、模写過程に移る。模写過程では、グリッドを変化させ、手本イラストと比較することができる。

文献 [5] では、洗練化においてデータセットとして実世界の人の顔画像である CelebA-HQ データセット [13] を用いているが、本手法ではキャラクタイラストから作成したデータセットでネットワークを学習させる。システム概要

図を図 2 に示す。

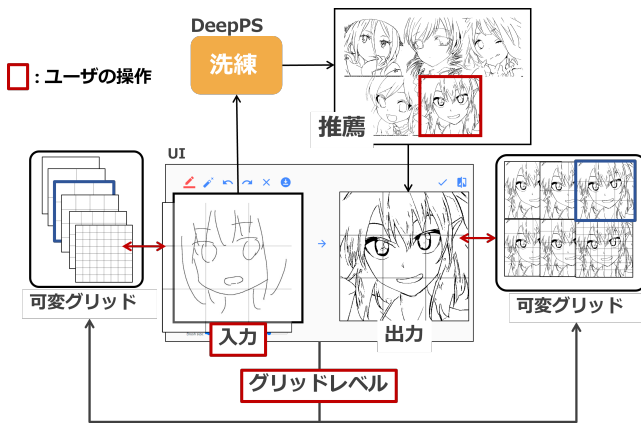


図 2: システム概要図

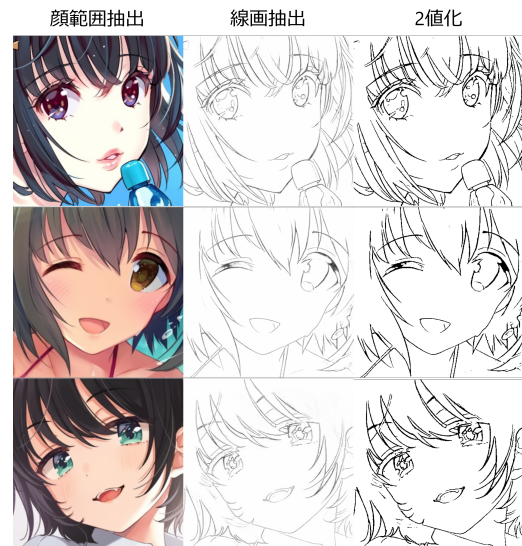


図 3: 線画抽出・2 値化例 [16] [17][18]

### 3.1 データセット

本研究において、データセット作成のために、イラスト画像データセットとして Face of Pixiv Top Daily Illustration 2018 ~ 2020<sup>\*1</sup>\*2<sup>\*3</sup>と Tagged Anime Illustrations<sup>\*4</sup>を使用している。

データセット作成のために、イラスト画像内のキャラクターの顔範囲画像の抽出、線画抽出、2 値化処理を行う。この処理は、Python と OpenCV を用いて行う。

イラスト画像内のキャラクターの顔範囲画像の抽出は、アニメ顔検出用のライブラリ lbpcascade\_animeface [14] を使用する。本実験ではサイズが 512 × 512 px の顔イラスト画像を使用するが、データセットから顔範囲画像の抽出を行うと、512 × 512 px 以上のサイズの顔範囲画像はごく少数となる。そのため、サイズが 300 × 300 px 以上の顔範囲画像を対象とし、OpenCV の関数 cv2.resize を用いて 512 × 512 px に拡大する。そして、waifu2x-ncnn-vulkan [15] を用いて画像拡大において発生したノイズ除去を行う。これらの処理の結果、データセットとして使用する顔イラスト画像は 4,847 枚である。

顔範囲の抽出後、Anime2Sketch [16] [17] を用いて線画抽出を行う。最後に 2 値化処理を行う。この処理では、画像ごとに最適なしきい値を求めるため、判断分析法 (Discriminant Analysis Method; 大津の手法 [18]) を用いる。例を図 3 に示す。

\*1 <https://www.kaggle.com/datasets/stevenevan99/face-of-pixiv-top-daily-illustration-2018>

\*2 <https://www.kaggle.com/datasets/stevenevan99/face-of-pixiv-top-daily-illustration-2019>

\*3 <https://www.kaggle.com/datasets/stevenevan99/face-of-pixiv-top-daily-illustration-2020>

\*4 <https://www.kaggle.com/datasets/mylesoneill/tagged-anime-illustrations>

### 3.2 画像推薦システム

本研究では、模写における手本イラストの推薦を行う。推薦については、ユーザのラフスケッチを入力として、データセットから類似画像を検索し、類似度の高い上位 5 枚を推薦する。本節では、類似画像検索手法の検討と、ユーザのラフスケッチの処理について示す。

#### 3.2.1 特徴量を用いた類似画像検索

現在、類似画像検索で用いられる代表的な手法には、CNN を使った特徴量抽出や、Siamese Network [19] などがあげられる。これらの手法は色情報を用いて学習を行っていることが多い。しかし、モノクロ画像である手本イラストの推薦には、色情報を必要としない手法のほうが適していると考えられる。色情報を用いずに画像の類似度を求める手法として、ヒストグラムベースの方法や、特徴量抽出、テンプレートマッチングなどがあげられる。ヒストグラムベースの方法は画像内の輝度値をヒストグラムで表す手法であり、輝度の分布を使用する。特徴量抽出は画素中に存在する重要な点である特徴量のうち、周辺の濃淡の変化が多い特徴点の周辺領域の輝度や画素の微分値から得られる特徴ベクトルを抽出する方法である。テンプレートマッチングはテンプレート画像に一致する部分を見つける手法である。本研究ではラフスケッチからデータセット内の画像検索を行うため、ラフスケッチの大まかな特徴を捉えられるかが重要である。そのため、ヒストグラムベースとテンプレートマッチングは適さないと考えられる。したがって、本研究では類似画像検索の手法として特徴量抽出を用いる。

画像の特徴量には様々なものがあるが、代表的なものとして SIFT と HOG がある。本研究では、この 2 つ特徴量のどちらを使用するかを決定するための予備実験を行う。

#### 3.2.2 特徴量の比較 (予備実験)

予備実験として、HOG 特徴量 [20] と SIFT 特徴量 [21] を利用した類似画像検索を行い、両者の比較を行う。実験

の内容は、データベース内の画像の特徴量を保存するインデックスの作成時間の比較、インデックスを利用したラフスケッチを入力とした類似画像検索の質と検索時間の比較である。類似画像検索の質は5段階評価を行う。どちらの特徴量に対しても、特徴量のマッチングはOpenCVのBFMatcher.knnMatch関数を用いて行った。本予備実験はBFMatcher.knnMatchの引数を $k = 2$ とし、類似画像検索は類似度が高い5枚のイラストを示すことで評価を行った。本予備実験の被験者は5人で、全員20代男性であった。

### 3.2.3 結果

データベース内の画像の特徴量を保存するインデックスの作成時間は、HOG特徴量が3分48秒、SIFT特徴量が2分56秒となった。また、平均検索時間に関しては、HOG特徴量が0.36秒、SIFT特徴量が17.26秒となった。類似画像検索の質を5段階評価で行った。評価は、1:とても悪い~5:とても良いの5段階とした。平均値はHOG特徴量が4、SIFT特徴量が2.2となった。検索結果例を図4に表す。

平均検索時間はHOG特徴量のほうが非常に早く、主観的な評価ではあるが画像検索の質もHOG特徴量がSIFT特徴量に対し評価が高かったため、提案システムではHOG特徴量を使用することにする。

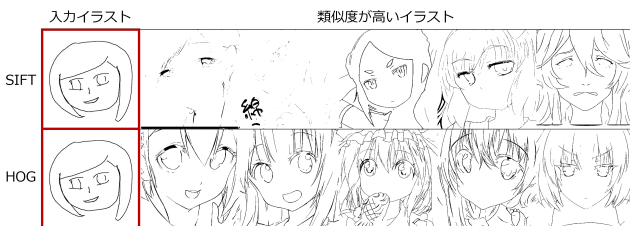


図4: 予備実験の結果例

### 3.2.4 ラフスケッチの洗練化

類似画像検索においてユーザが描いたラフスケッチを入力とするが、ラフスケッチはデータベース内のイラストとは違い詳細度は低く、未完成である。そのため、類似画像検索に使用するためには詳細度を上げることでより良い精度が得られると考えられる。本研究では、ラフスケッチの詳細度を上げるためにDeepPS [5]を参考としたスケッチ洗練ネットワークを利用する。

DeepPSでは、線画の膨張(Dilation)を行い、膨張レベル $l$ ごとに膨張した線画から元線画を出力するような学習をpix2pix [8]を用いて行っている。この過程では、洗練レベル $l$ に応じて畳み込み層の出力を調整している。またpix2pixで線画をカラー画像に出力するように学習し、GANを用いて精度を向上させている。DeepPSでは実際の写真であるCelebA-HQ [13]を使って学習を行っているが、本研究はイラストを対象にしているため3.1節で作成したデータセットで学習を行った。本研究におけるデータ

セットを適用させた洗練例を図5に示す。図5では、上部



図5: スケッチ洗練例 ( $l = 1$ )

の粗いイラストの詳細度が上がっているが、下部の比較的細かなイラストでは線が統合することで減っていることが確認できる。本研究におけるデータセットはDeepPSで用いているCelebA-HQデータセット [13](30,000枚、サイズ $1024 \times 1024$ )に比べて小規模なため、精度が落ちてしまうと考えられる。また、実際の顔写真ではなくキャラクターイラストを扱っているため、それを考慮した処理を行う必要があると考えられる。

### 3.3 インターフェース



図6: 提案インタフェース

図6に示すインタフェースの開発を行った。本インタフェースはJavaScriptの2Dキャンバスライブラリであるkonva.jsとJavaScriptのフレームワークであるVue.js、Pythonを用いて実装した。本インタフェースでは、スケッチの作成・グリッド操作を行うために、ペンと消しゴム、操作の取り消しとやり直しといった基本の機能から、本システムの特徴である参考イラスト生成、グリッド変化、手本としたイラストとの比較などの機能を有している。

グリッドレベルはグリッドなし、 $5 \times 5$ 、 $4 \times 4$ 、 $3 \times 3$ 、 $2 \times 2$ を設定していたが、もう一段階細かなグリッドが必要と考えたため、 $5 \times 5$ よりグリッド1つあたりの面積が半分以下の $8 \times 8$ を加えて6段階の設定を行った。例を図7に示す。

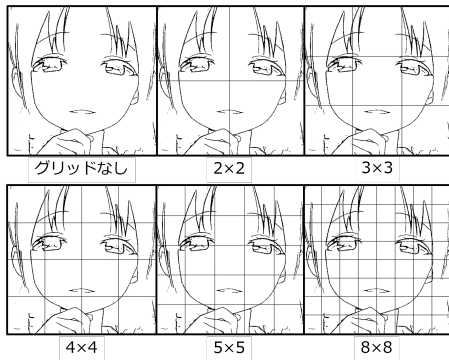


図 7: グリッド追加例

#### 4. ユーザスタディ

提案システムを評価するための実験を行った。本実験では、20代から30代の男女計10人(男8,女2)に対して本システムを利用してもらった。提案システムによる支援ありと支援無しでアンケート調査を行い、最後にSUS(System Usability Scale) [22]と日本語版NASA-TLX [23]による評価を行った。

被験者を半数ずつランダムにパターン①とパターン②に分け、それぞれの流れに沿って実験を行った。実験の流れを図8に示す。

実験手順としては、まず本研究の趣旨や提案システムの概要説明を行い、被験者のイラスト経験などを調査する事前アンケートを行った。次に、提案インターフェースの説明を行い、5分程度の練習を行った。提案インターフェースの説明では、実際に著者がインターフェースを利用した際の動画を流しながら説明を行った。練習では、提案インターフェースに慣れるために、実験の一連の流れを練習した。

事前アンケートでは、表1に示す内容の調査を行った。

表 1: 事前アンケート内容

- |                          |
|--------------------------|
| 1 イラストをどのくらいの頻度で描きますか    |
| 2 模写をしたことはありますか          |
| 3 (前設問であると答えた場合) 何回ありますか |
| 4 グリッドを用いて模写をしたことがありますか  |

その後、インターフェースを利用してシステムによる支援あり・支援無しで模写の制作を行った。その際、制作過程を動画として収録し、制作時間と用いたグリッドを記録した。使用するグリッドは被験者が選択したものを基準とした。支援ありの場合はグリッドを変化させることができるが、必ずしも変化させる必要はないことを説明した。

模写の制作を行った後、表2に示すアンケート調査を行った。加えて、支援ありで模写の製作を行う場合は、表3の内容も調査した。支援あり・無しで模写制作を行いアンケートを回答した後、SUSとNASA-TLXによる評価を行った。

表 2: 模写後のアンケート内容

- |                                 |
|---------------------------------|
| 1 絵の満足度 (5段階評価)                 |
| 2 絵の完成度 (5段階評価)                 |
| 3 上手く描けたと思うか (5段階評価)            |
| 4 気をつけたこと                       |
| 5 難しかったこと                       |
| 6 作成した模写に違和感があるか (ある・無し)        |
| 7 (前設問で違和感があると答えた場合) どこに違和感があるか |
| 8 修正するとしたらどこをどのようにするか           |
| 9 感想                            |

表 3: 模写後のアンケート内容 (支援あり)

- |                                   |
|-----------------------------------|
| 10 グリッドが可変なのは模写で有用であったか (5段階評価)   |
| 11 前設問の理由                         |
| 12 画像推薦システムは模写で有用であったか (5段階評価)    |
| 13 前設問の理由                         |
| 14 またこのシステムを使って模写をしたいと思うか (5段階評価) |
| 15 前設問の理由                         |

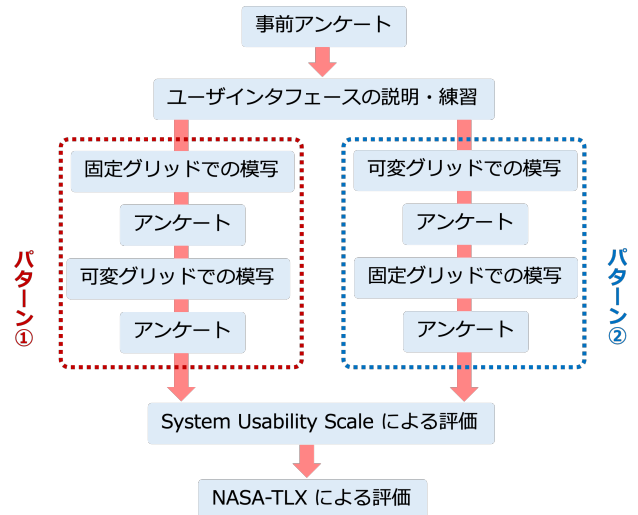


図 8: 実験の流れ

#### 5. 結果

本章では、提案システムを利用したユーザスタディの結果より、本研究の有効性を示す。

##### 5.1 事前アンケート調査の結果

被験者10人のうち、イラストを描いたことがあるのは8人で、そのうち2人が月に数回以上描いていた。また、模写の経験については8人が経験があり、そのうち2人は10回程度、1人は頻繁に行うと答えていた。グリッドを使った模写の経験がある被験者は1人であった。

##### 5.2 アンケート調査の結果

制作時間の平均は支援なしで14分42秒、支援ありで14分34秒となった。比較を図10(a)に示す。

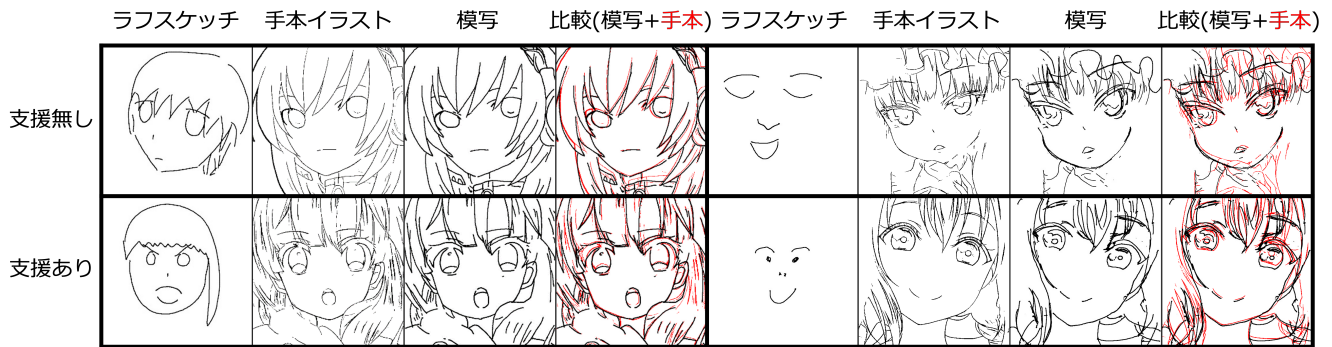
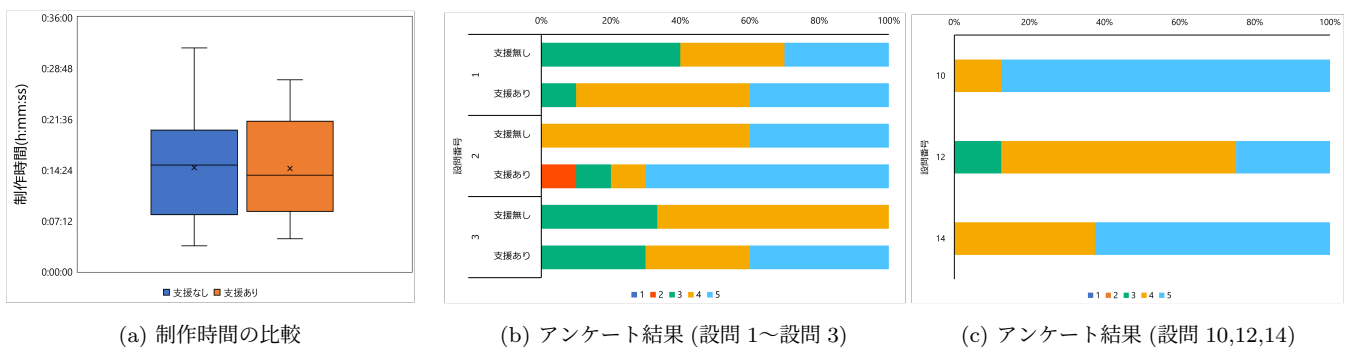


図 9: 制作例



(a) 制作時間の比較

(b) アンケート結果 (設問 1~設問 3)

(c) アンケート結果 (設問 10,12,14)

図 10: 実験結果

図 10(a) から読み取れるように、制作時間の最大値は支援なしのほうが大きかったものの、支援ありと支援なしの制作時間に違いが見られるか検定を行った。サンプル数が少ないため、ノンパラメトリック検定の 1 つであるマン・ホイットニーの U 検定を用いた。その結果、制作時間に有意差は確認されなかった ( $U = 47, p > 0.10$ )。支援なしの制作時間の最大値が大きかった理由は、被験者が指定した参考イラストが他で選ばれた参考イラストより線の描写が多く、時間がかかってしまったためだと考えられる。

次に、表 2 で示したアンケート項目の比較を行う。5 段階評価である設問 1~設問 3 の結果を図 10(b) で示す。

また、設問 1, 2, 3 において支援なしの評価の平均値はそれぞれ 3.9, 4.4, 3.9、支援ありの評価の平均値はそれぞれ 4.3, 4.4, 4.1 であった。

設問 1: 絵の満足度と設問 3: 上手く描けたと思うかは、支援ありの方が評価が高かった。一方、設問 2: 絵の完成度は同じ値となっている。つまり、平均ではどの被験者も支援ありと支援なしで同程度の完成度で制作したということを示している。制作時間の比較と同じく、マン・ホイットニーの U 検定を用いた結果、設問 1~3 に有意差は確認されなかった (設問 1:  $U = 36.5, p > 0.05$ , 設問 2:  $U = 41, p > 0.05$ , 設問 3:  $U = 42.5, p > 0.05$ )。

支援ありの場合のみに行った、表 3 で示したアンケート結果を図 10(c) で示す。また、設問 10,12,14 において評価の平均値はそれぞれ 4.9, 4.1, 4.6 であった。

### 5.3 実験結果のまとめ

アンケート評価の結果、模写過程においてグリッドを可変にすることは設問 10 の平均値が 4.8 であることや、設問 11 のコメントから高評価であることが確かめられた。また、顔イラストの模写において特に目や顎、髪が難しいことがわかった。目は細かく描写されていることが多く、加えてイラストの印象を大きく左右するため、特に時間をかける被験者が多いことが実験を観察していると分かった。髪や顎は逆に描写が粗いため、周囲の線の位置などから空間を把握して描写する必要がある。描写が粗い箇所に関しては、特に細かなグリッドを使用しないと初心者にとっては模写が難しいと考えられる。また、初めから細かなグリッドを利用するとグリッド同士の対応関係を間違えろといった意見もあった。グリッドが粗いと大域的な意識ができ、グリッドが細かいと局所的に意識を向けることができるという意見を踏まえると、グリッドを粗いものから細かいものにするによって、イラスト描画の基本の流れである大域的に描くこと (ラフ) から局所的な箇所を仕上げる (線画) を自然に身に着けられるのではないかと考えられる。また、模写過程の初めの方でグリッドを細かくして、そのまま最後まで描き切る被験者も多かったため、グリッドの細かさはユーザの描画具合を読み取ることで、システム内で自動的に行う必要を感じた。

グリッドを用いた模写自体に関しては、支援あり・無しどちらのどちらについても設問 3 の平均値が 4 であることや、設

表 4: SUS の質問内容と結果

番号	質問内容	平均	標準偏差
1	デジタル環境で模写を行う場合、このシステムをしばしば使用したいと思う。	4.3	0.64
2	このシステムを使用するには説明が必要となるほど複雑であると感じた。	2.1	0.70
3	このシステムは使いやすいと感じた。	4.0	0.63
4	このシステムを利用するには、専門家・技術者のサポートが必要だと思う。	2.0	0.89
5	このシステムの様々な機能には統一感があると感じた。	3.8	0.60
6	このシステムには一貫性が無いところが多いと感じた。	1.5	0.50
7	ほとんどの人がこのシステムはすぐ使いこなせるようになると思う。	4.4	0.49
8	このシステムは操作しづらいと感じた。	2.0	0.45
9	このシステムを自信を持って使える。	4.1	0.83
10	このシステムを使いこなすには事前にたくさんのことを学ぶ必要があると思う。	1.7	0.64

問 9 のコメントに上手なイラストが描けたと意見があることから、上手く描けたと評価されている。一方で、グリッドとイラストが重なっている箇所が書きづらいとの意見もあった。これに対しては、グリッドの色を可変にすることで対処できると考えられる。

また、インタフェースの問題として細い線が書きづらい(ペンのサイズが太い)ことや線の太さの扱いに困るといった意見もあった。これに関しては、ペンのサイズを調整するスライダとペンのサイズの対応関係がよくわからなかったのではないかと考えられる。スライダを調整する際に、調整したペンのサイズを視覚的に把握できるようにすることで解決できると考えられる。SUS に関する質問内容と結果を表 4 に示す。評価は 5 段階評価で行い、1: とてもそうは思わない、5: とてもそう思うとした。各被験者の SUS スコアとして、0 から 100 のスケールに変換するための処理を行った結果、最小値が 67.5、最大値は 90、平均は 78.25 であった。[24] では、SUS スコアと形容詞評価 (Adjective ratings; Best Imaginable, Excellent, Good, OK, Poor, Worst Imaginable の 6 段階の評価) の間に有意な正の相関が確認されている。[24] によると、本システムの SUS スコアである 78.25 は  $71.4 \pm 11.6$  の区分に含まれており、形容詞評価で「Good」に該当し、提案インタフェースが良好な評価を持つことが確認できた。

SUS の各設問結果は、全体の平均と比較して、設問 1, 6, 7, 10 が高評価、設問 2 と 5 が低評価となった。設問 1, 6, 7, 10 はそれぞれシステムの使用意欲、一貫性、理解の簡単さ、事前知識の不必要さを質問している。これらが高評価であることは、提案システムが使用しやすく、かつ使用したいという評価を得ていると考えられる。設問 2 はシステムの複雑さ、設問 5 はシステムの統一感を表している。これらが低評価であることは、ユーザの操作する機能が多いことが考えられる。特に、グリッドの操作はスライダに加え決定ボタンを押す必要があるため、煩わしかったのではないかと考えられる。また、決定ボタンは推薦された参考イラストの決定とグリッド操作の決定の両方に使

NASA-TLX Score

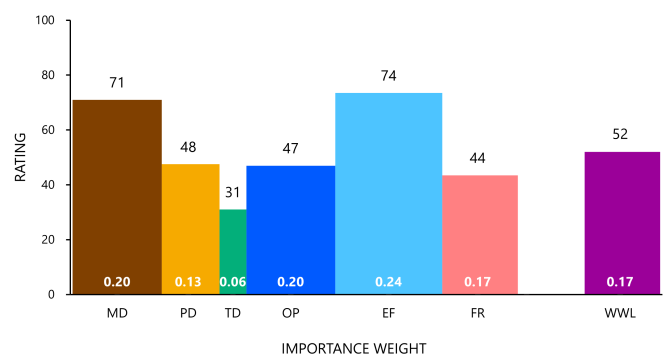


図 11: NASA-TLX 結果

用するため、複数の機能を持つことでまとまりがなく、統一感が無いと判断されたのではないかと考える。

NASA-TLX に関する結果を図 11 に示す。図 11 における MD, PD, TD, OP, EF, FR, WWL はそれぞれ知的・知覚的要求、身体的要求、タイムプレッシャー、作業成績、努力、フラストレーション、全体的な評価である。アナログスケールと重みの積に対して、特に評価の大きいものは知的・知覚的要求と努力、特に評価の小さいものはタイムプレッシャーとなった。

ユーザスタディは、提案システムを用いて模写を制作することであり、参考イラストの観察と完成までの努力が必要不可欠である。そのため、知的・知覚的要求と努力の評価が大きくなることは自明である。タイムプレッシャーについては、特に制限時間を設けなかったため評価が小さくなったと考えられる。

また、重みにおいては作業成績についても大きく評価されていた。つまり、提案システムを使用する際には達成度や満足度が重要であるが、それによる負荷はあまり発生しないことを示唆していると考えられる。



## 6. 結論と今後の展望

本研究では、模写で用いるグリッドを可変にすることで、観察に焦点を当てた効果的な模写支援を可能にするという仮説に基づいて、ラフスケッチを入力とした参考イラストの推薦と、模写における可変グリッドシステムを組み合わせた模写練習支援のフレームワークを提案した。本システムでは、ユーザの描いたラフスケッチを入力とし、データセットから特徴量の近い類似イラストを候補として提示する。ユーザはそこから好ましいイラストを選び、グリッドの細かさを変化できる模写インタフェースにより模写を行う。ユーザスタディを行った結果、グリッドを可変にすることで効果的な模写を可能にすることが確かめられた。また、インタフェースの評価を行ない、提案インタフェースの有用性を確認できた。

本研究で提案したインタフェースは、ユーザが全て手動で操作する必要がある。ユーザスタディからも、ユーザの負担が低くはないことが示されている。そのため、ユーザが行う動作をある程度自動化することを考えている。例えば、グリッドをユーザ側のキャンバスと参考イラストで同じタイミングで変化するように実装を行うことや、ユーザの描画具合を読み取ることで、適切なグリッドを設定することなどが考えられる。また、カラーイラストからデータセットを作成したことを活かし、モノクロだけでなくカラーイラストの模写にも適応させていきたい。

### 参考文献

- [1] wacom. 画材・文具に対する意識調査. <https://tablet.wacom.co.jp/article/regarding-art-supplies-and-stationery-awareness-survey>, 2021.
- [2] 中島 楓華, 山西 良典, 巽 優人, 他. アニメキャラの顔パーツの位置バランスとキャラクタ属性の関係性に関する基礎検討. 第6回コミック工学研究会予稿集, pp. 24–28, 2021.
- [3] 大芝 潤気, 岩田 基, 黄瀬 浩一. C2GAN を用いたランドマークに基づくアニメキャラクタの顔画像自動生成手法. 第3回コミック工学研究会予稿集, pp. 57–62, 2020.
- [4] 小笠原 渚, 清 雄一, 田原 康之, 大須賀 昭彦. アニメキャラクター風イラストの全身画像生成手法の提案. 電子情報通信学会技術研究報告, Vol. 120, No. 379, pp. 30–35, 2021.
- [5] Shuai Yang, Zhangyang Wang, Jiaying Liu, and Zongming Guo. Deep plastic surgery: Robust and controllable image editing with Human-Drawn sketches. January 2020.
- [6] Yong Jae Lee, C Lawrence Zitnick, and Michael F Cohen. ShadowDraw: Real-time user guidance for free-hand drawing. *ACM Transactions on Graphics (TOG)*, Vol. 30, No. 4, pp. 1–10, 2011.
- [7] 渡邊優, 阿倍博信. pix2pix を用いたデジタルイラスト制作における自動レイヤ分けシステム. 研究報告コンシューマ・デバイス&システム (CDS), Vol. 2020-CDS-27, No. 27, p. 1–8, Jan 2020.
- [8] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *CVPR*, 2017.
- [9] Hung-Yu Tseng, Matthew Fisher, Jingwan Lu, Yijun Li, Vladimir Kim, and Ming-Hsuan Yang. Modeling artistic workflows for image generation and editing. In *Computer Vision – ECCV 2020*, pp. 158–174. Springer International Publishing, 2020.
- [10] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *CoRR*, Vol. abs/1411.1784, , 2014.
- [11] 菅野一平, 中村聡史. 個人のイラスト制作における観察に対する支援手法の検討. 第2022-DCC-30 巻, pp. 1–8, 2022.
- [12] Emmanuel Iarussi, Adrien Bousseau, and Theophanis Tsandilas. The drawing assistant: Automated drawing guidance and feedback from photographs. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13*, p. 183–192. Association for Computing Machinery, 2013.
- [13] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In *International Conference on Learning Representations*, 2018.
- [14] lbpcascade\_animeface: A face detector for anime/manga using OpenCV. [https://github.com/nagadomi/lbpcascade\\_animeface](https://github.com/nagadomi/lbpcascade_animeface).
- [15] waifu2x-ncnn-vulkan: waifu2x converter ncnn version, runs fast on intel / amd / nvidia / apple-silicon GPU with vulkan. <https://github.com/nihui/waifu2x-ncnn-vulkan>.
- [16] X. Xiang, D. Liu, X. Yang, et al. Anime2sketch: A sketch extractor for anime arts with deep networks. <https://github.com/Mukosame/Anime2Sketch>, 2021.
- [17] X. Xiang, D. Liu, X. Yang, et al. Adversarial open domain adaptation for sketch-to-photo synthesis. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 944–954. IEEE Computer Society, jan 2022.
- [18] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, pp. 62–66, 1979.
- [19] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a "siamese" time delay neural network. NIPS'93. Morgan Kaufmann Publishers Inc., 1993.
- [20] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 886–893 vol. 1, 2005.
- [21] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, Vol. 60, No. 2, pp. 91–110, 2004.
- [22] John Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, Vol. 189, pp. 189–194, 11 1995.
- [23] 芳賀繁, 水上直樹. 日本語版 nasa-tlx によるメンタルワークロード測定 各種室内実験課題の困難度に対するワークロード得点の感度. 人間工学, Vol. 32, No. 2, pp. 71–79, 1996.
- [24] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *J. Usability Studies*, Vol. 4, No. 3, p. 114–123, may 2009.