

Title	How Experts Draw: 描画過程における熟練者と初心者の描画運動特性の比較分析
Author(s)	天野, 一平; 福里, 司; 謝, 浩然
Citation	情報処理学会第217回ヒューマンコンピュータインタラクション研究発表会, 2026-HCI-217(28): 1-8
Issue Date	2026-03-02
Type	Conference Paper
Text version	publisher
URL	<a href="https://hdl.handle.net/10119/20350">https://hdl.handle.net/10119/20350</a>
Rights	<p>社団法人 情報処理学会, 天野一平, 福里司, 謝浩然, 情報処理学会第217回ヒューマンコンピュータインタラクション研究発表会, Vol.2026-HCI-217, No.28, 2026, 1-8.</p> <p>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(社)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	情報処理学会第217回ヒューマンコンピュータインタラクション研究発表会, 芝浦工業大学豊洲キャンパス, 東京, 2026年3月9日-11日

# How Experts Draw: 描画過程における熟練者と初心者の描画運動特性の比較分析

天野 一平<sup>1a)</sup> 福里 司<sup>2b)</sup> 謝 浩然<sup>1,2c)</sup>

**概要:** 描画技術は高度な認知・運動スキルの統合体である。しかし、描画プロセスにおいて視線移動や身体制御が必要となり、経験者と初心者の技能差がどう反映されるかは未だに解明されていない。そこで本研究は、描画過程における経験者と初心者の運筆速度や筆圧の変化に注目し、定量的な分析を試みる。分析の結果、経験者と初心者の間には、運動制御および筆圧調整の戦略において有意な差異が得られた。まず、ストロークの構造において、初心者は経験者より長いストロークを生成しているものの、各ストロークあたりの長さが有意に短く、断片的な運筆が目立った。その一方、経験者は筆圧の平均値、最大値、および標準偏差のすべてにおいて、極めて大きな効果量で初心者を有意に上回った。これは経験者が描画表現に応じて筆圧を広範かつ動的に制御していることを示している。また、ストロークの平均速度に大きな差はなかったものの、最大速度および速度の標準偏差において、経験者は高い値を示している。本知見は、筆圧や速度といった運動特徴量に基づき、学習者の「非効率な状態（迷い）」をリアルタイムに検知・是正する、次世代の描画支援システムへの応用が期待される。

## 1. はじめに

描画は、視覚情報を運動指令に変換する高度な認知プロセスであり [1], [2], その習得には多大な時間を要する。経験者は、迷いのないストロークで効率的に形状や質感を表現できる一方、初心者は試行錯誤を繰り返し、過剰な力みや時間の浪費に陥りやすい [3], [4]。効果的な描画支援を行うためには、このような経験者と初心者の間に存在するスキルのギャップが、具体的にどのような運動特性として現れるのかを解明する必要がある。

従来の描画研究においては、完成された作品（静止画）の視覚的評価や、アンケートによる主観評価が主であった [5]。しかし、描画とは本来、時間の経過とともに生成される動的なプロセスである [6], [7]。「なぜその線が引かれたのか」「どこで時間を費やしたのか」といったプロセス情報と、最終的な「絵の質」との関係性を定量的に結びつける試みは、十分になされてこなかった。

そこで本研究では、描画中の運動データ（筆圧・速度など）と、深層学習モデルによる画像評価（CLIP・DINOv2[8], [9]）を統合した新たな分析フレームワークを提案する。独自に開発したロギングシステムを用いて、熟練者と初心者の描画プロセスを詳細に記録し、比較分析を行った。具体的には、熟練者 12 名と初心者 11 名の計 23 名を対象に、iPad を用いたアニメの顔画像の模写実験を行った。課題には、マンダラート法に基づき選定された、微細な筆圧制御を要する「老人」と、ストロークの安定性を要する「若者」の 2 種類の線画を用いた。収集した筆圧・速度などの運動データと、作成された画像の AI 評価スコアを統合し、スキルの差異を多角的に検証した。

本研究の主な貢献は以下の通りである。

- (1) 運動制御において、初心者が「筆圧と速度のトレードオフ（力むと描けない）」という制約を抱えていることを定量的に示した。
- (2) 最新の評価指標を用いた分析により、初心者は熟練者と統計的に同等の「形」（DINO スコア：初心者 0.42 vs 熟練者 0.38,  $p > 0.05$ ）を生成できているものの、そのため有意に長い描画時間（初心者 388s vs 熟練者 231s,  $p < 0.01$ ）を費やしている「非効率性」を明らかにした。

<sup>a)</sup>s2510002@jaist.ac.jp

<sup>b)</sup>tsukasafukusato@waseda.jp

<sup>c)</sup>xie@jaist.ac.jp

<sup>1</sup> 北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology

<sup>2</sup> 早稲田大学  
Waseda University

(3) 描画対象(老人・若者の顔画像)の複雑さに応じて, 経験者が柔軟に筆圧制御を行っている実態を可視化した. 本研究の見解は, 単なるスキルの比較にとどまらず, 学習者の「脱力」や「効率化」を促す次世代の描画支援システムの基盤となるものである.

## 2. 関連研究

### 2.1 描画プロセスの認知・運動的分析

描画における経験者と初心者の差異は, 長年にわたり認知科学およびデザイン工学の観点から調査されてきた. Chenらは, 概念設計段階におけるデザイナーのスケッチ活動を分析し, 経験者が思考の外化としてスケッチを戦略的に利用していることを指摘している [3]. また, 描画行為は単なる手の運動ではなく, 視覚情報と運動制御が密接に連携したプロセスである. Tchalenko は, 画家が対象を観察してから紙面にペンを走らせるまでの視線と手の協調 (Eye-Hand Coordination) メカニズムを詳細に調査し, 経験者特有の視覚運動ループの存在を示唆した [6]. Lohmeyer らも同様に, 視線計測を用いて熟練デザイナーの視覚的注意を分析し, 彼らが対象の機能的特徴に効率的に注意を向けていることを明らかにしている [4]. Suwa らはプロトコル分析により, 建築家がスケッチを通して新たな知覚的発見を行っていることを示した [10].

近年では, ウェアラブルセンサを用いてこれらの動作を定量化する試みも進んでいる. Akamatsu らは, 視線データと身体モーションデータを組み合わせたマルチモーダル解析により, 作業者の熟練度 (Expert/Novice) を高精度に分類する手法を提案した [11]. しかし, これらの先行研究は主に「動作パターンの分類」や「視線戦略の記述」に焦点を当てており, 実際に生成された「絵の質」と「運動コスト(筆圧や時間の浪費)」との定量的な関係については, 十分に議論されていない.

### 2.2 深層学習を用いた描画支援と計算論的解析

コンピュータビジョン分野においては, 手書きスケッチの認識や生成に関する研究が急速に進展している [5]. 初期の代表的な研究として, Ha らはスケッチを画素集合ではなくペンの移動軌跡 (ベクタ形式の時系列データ) としてモデル化した SketchRNN を提案し, 描画プロセスの逐次的な生成を可能にした [7]. 近年では, 大規模な画像-テキストペアで学習された基盤モデルが登場し, スケッチの意味論的 (Semantic) な評価が可能となっている. Radford らが提案した CLIP [8] や, Oquab らによる DINOv2 [9] は, 抽象的なスケッチ画像からも高次元特徴量を抽出できるため, CLIPasso [12] のように, スケッチの抽象度を制御しながら意味を保つ生成モデルへの応用も進んでいる.

これらの生成技術を応用したインタラクティブな描画支援システムも提案されている. 特に, Huang らは, ユーザー

の描画ストロークに基づいてリアルタイムにアニメ顔画像を生成・補完する AniFaceDrawing [13] を提案した. このシステムは, ユーザーの未熟なスケッチと高品質な生成結果との間のギャップを埋めることで, 創作のハードルを大きく下げること成功している. また, 初心者に対する直接的な指導方法として, DualFace [14] が提案されている. 同研究では, 描画プロセスを「大域的な構造」と「局所的な詳細」の2段階に分け, それぞれの段階で適切なガイドを提示することで, 初心者が整った顔スケッチを描けるよう支援した.

しかし, これらの計算論的アプローチや支援システムは, 主に「描画された結果 (画像)」の解析や品質向上に留まっている. 経験者がどのように筆圧を制御してその線を描いたのか, あるいは初心者がどこで無駄な力を込めてしまったのかといった「身体的プロセス」の情報は, 最終的な画像からは欠落している. システムがユーザーのスキルレベルに合わせてより適応的な支援を行うためには, まず人間側の運動学的・視覚的な振る舞いの差異を定量的に理解する必要があり, 学習支援への直接的な応用には課題が残る.

## 3. 実験設定

本研究では, 描画動作の定量的データを取得するため, 独自に開発したデータロギングシステムを用いた. 以下にその詳細を述べる.

### 3.1 描画対象の選定 (マンダラート分析)

実験タスクの選定にあたり, 描画に必要なスキル要素を網羅的に抽出するため, マンダラート法 (Mandalart) を用いた分析を行った. 「描画スキル」を中心核とし, 周辺に関連要素 (筆圧制御, 形状把握, 線の安定性など) を展開・構造化した結果, 対照的な運動特性を要する**複数種類のアニメーションキャラクターの線画**を課題として選出した.

本実験では計6種類の画像を使用したが, 本稿ではその中から, 特に特性の差異が著しい2つの代表例を取り上げて分析する. 一つは, 複雑な皺や曲線を多く含み, 微細な筆圧制御 (脱力) を要する「老人男性 (Old Man)」の例である (図3(a)). もう一つは, シンプルな輪郭線で構成され, 一定の速度とストロークの安定性が求められる「若年男性 (Young Man)」の例である (図3(b)).

### 3.2 データ収集システム

計測には, 11 インチ iPad Pro (第3世代) および Apple Pencil (第2世代) を使用した. これらのデバイスを用いることで, 高精度な座標データに加え, 筆圧やペンの傾きといった詳細な描画情報の取得が可能である. 描画データの記録には, Xcode を用いて独自に開発した描画アプリケーションを使用した (図4). 本システムでは, 描画領域の左上隅を原点 (0, 0) とする座標系を採用し, 以下の項目をサンプリ



図 1: 老人男性モデルの例 [15]



図 2: 若年男性モデルの例 [16]

図 3: 実験で用いた描画対象の代表例



(a) データ収集アプリの画面



(b) 実験中の描画風景

図 4: 実験システムの概要.(a) 独自開発したロギングアプリケーション.(b) 120Hz でのデータ収集を行っている様子.

ングレート 120Hz で記録する.

- 座標 ( $x, y$ ): ペン先の位置情報 [px]
- 筆圧 (force): ペン先にかかる圧力 (正規化値 0.0 ~ 1.0)
- 傾斜角 (alt)・方位角 (azi): ペンの物理的な傾きと向き [rad]
- 時刻 (time): システム上の絶対時刻 [sec]
- LSL タイムスタンプ (LSL\_T): 他の計測デバイスとの同期用タイムスタンプ [sec]

### 3.3 データ収集実験

#### 3.3.1 被験者

本実験には、計 23 名の大学生が参加した。参加者は、描画学習の経験有無に基づき以下の 2 群に分類された。

**経験者 (12 名)** 芸術系大学で美術・デザインを専攻する学生 11 名 (全員女性), および建築分野において描画を学んだ男子学生 1 名 (20 代) である。芸術専攻学生の年齢構成は 10 代が 5 名, 20 代が 6 名であり、建築専攻学生を含めると経験者群全体では 10 代が 5 名, 20 代が 7 名となる。なお、本研究では「半年以上の継続的な描画学習経験」を有する者を経験者と定義した。

**初心者 (11 名)** 過去に専門的な描画教育を受けた経験のない者である。内訳は男性 5 名 (10 代 3 名, 20 代 2

名), 女性 6 名 (10 代 3 名, 20 代 3 名) であった。

#### 3.3.2 描画タスクと手順

被験者に対し、アニメ調のキャラクターの顔を模写するタスクを課した。描画対象として、以下の計 6 種類のキャラクター画像を提示した。描画対象として、若年層・中年層・老年層のそれぞれについて男女を組み合わせた、計 6 種類のキャラクター画像を提示した。実験では、手本となる画像を提示し、それを iPad 上で模写する形式をとった。描画に関する時間制限や、描画順序などの制約は一切設けず、被験者自身の自由なスタイルで描画を行わせた。

#### 3.4 分析手法

本研究では、経験者と初心者の描画プロセスの差異を多角的に明らかにするため、ストロークデータに基づく「運動学的特徴」と、深層学習モデルを用いた「視覚的類似性」の両面から分析を行った。

#### 3.5 ストロークデータの処理と運動特徴量の算出

収集された生ログデータ (log.txt) から、タイムスタンプ  $t$ 、ペン先座標 ( $x, y$ ), および筆圧  $f$  の時系列データを抽出した。筆圧が 0 より大きい区間を有効なストロークとして定義し、以下の特徴量を算出した。

- (1) **総描画時間 (Total Duration):** 最初のストローク開始から最後のストローク終了までの経過時間
- (2) **総筆記距離 (Total Distance):** 各サンプリング点間のユークリッド距離の総和
- (3) **平均筆記速度 (Average Speed):** 各ストローク内の移動距離を経過時間で除した値の平均
- (4) **平均筆圧 (Average Force):** 全サンプリング点における筆圧値の平均

#### 3.6 深層学習モデルによる画像特徴抽出

描かれたスケッチの「質」を客観的に評価するため、異なる特性を持つ 2 つの事前学習済みモデル, CLIP および DINOv2 を用いた。評価には、被験者が描画したスケッチ画像 (Student) と、模写対象である手本画像 (Teacher) の比較を用いた。

##### 3.6.1 画像の前処理 (Image Preprocessing)

画像入力における空白領域の影響を排除し、描画内容に焦点を当てるため、ログデータから得られた描画軌跡の最大・最小座標に基づき、関心領域 (Region of Interest: ROI) を定義した。

$$\text{ROI} = [x_{\min} - m, y_{\min} - m, x_{\max} + m, y_{\max} + m] \quad (1)$$

ここで、 $x_{\min}, x_{\max}, y_{\min}, y_{\max}$  は収集された全ストロークデータの座標における最小値および最大値 (バウンディングボックス) を表す。また、 $m$  は画像切り出しの際に設ける余白 (マージン) であり、本研究では  $m = 20$  [px] と設定

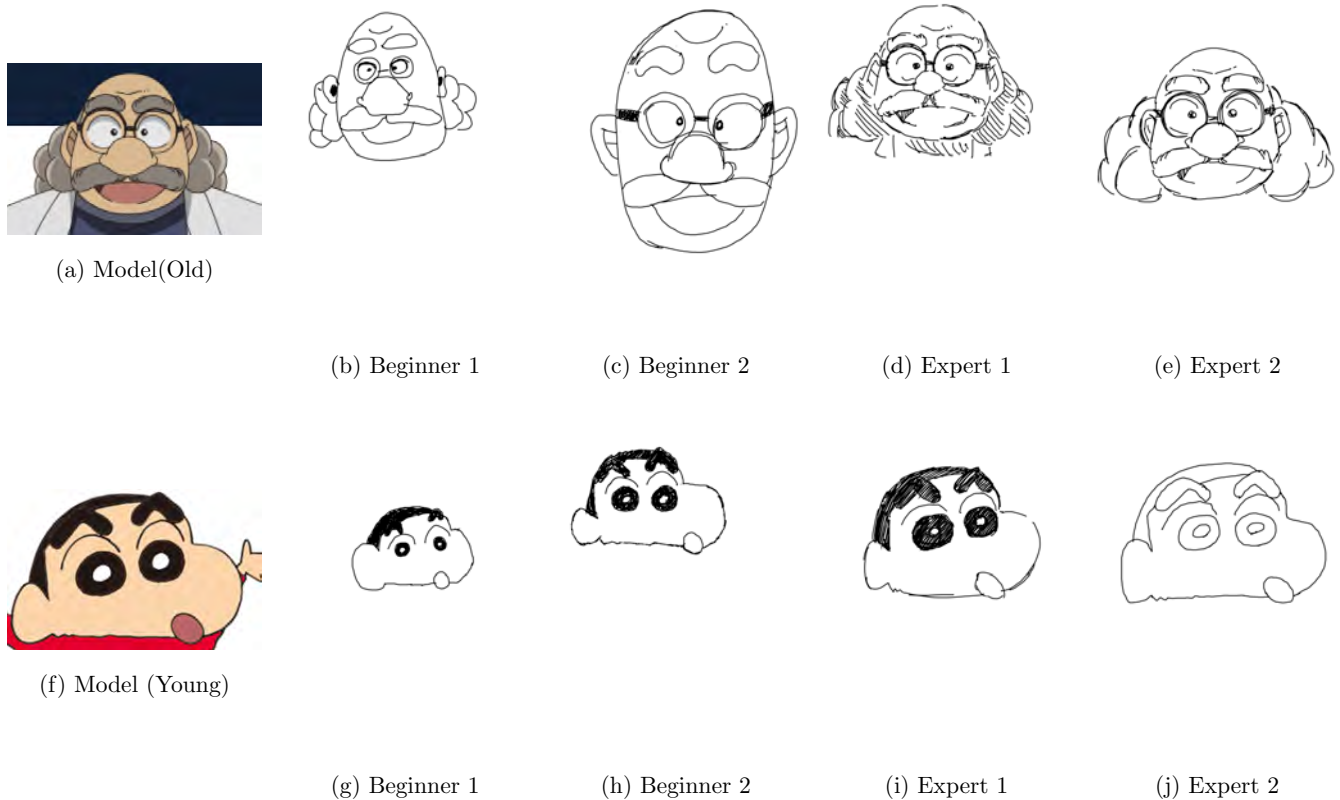


図 5: 描画結果の比較. 上段は「老人男性 (阿笠博士 [15])」, 下段は「若年男性 (野原しんのすけ [16])」の結果を示す. 左列の手本画像に対し, 初心者 (左から 2,3 番目) は線の震えや形状の著しい歪みが見られる. 一方, 経験者 (右側の 2 名) は安定したストロークにより対象の構造を正確に捉えている.

した. 切り出された画像は正方形にパディングされた後, 各モデルの入力解像度に合わせてリサイズおよび正規化処理 (ImageNet 平均・標準偏差) を適用した.

### 3.6.2 CLIP による意味的整合性の評価 (Semantic Similarity)

描画対象の概念的な特徴 (「何が描かれているか」) を評価するため, CLIP (ViT-B/32) [8] を用いた. CLIP は画像とテキストの大規模ペアデータセットで学習されており, 大域的な意味情報の抽出に優れている. 入力画像  $I$  をエンコーダ  $E_{clip}$  に入力し, 得られた埋め込みベクトル  $v_{clip}$  を用いて, 手本画像とスケッチの意味的な類似度を算出した.

### 3.6.3 DINOv2 による構造的整合性の評価 (Structural Similarity)

描画の幾何学的な正確さや空間的な配置 (「線が正しい位置にあるか」) を評価するため, 自己教師あり学習モデルである DINOv2 (ViT-Small/14, 518px) [9] を用いた. DINOv2 の特徴量は, 物体のパーツレベルでの空間的な対応関係 (Dense Correspondence) を捉えることに特化している. これにより, 線画のようなテクスチャ情報の少ないドメインにおいても, 手本の構造をどの程度正確に模倣できているかを定量化した.

### 3.7 類似度スコアの定義

各モデルから抽出された手本画像の特徴ベクトル  $v_{teacher}$  と, 被験者のスケッチの特徴ベクトル  $v_{user}$  の間のコサイン類似度 (Cosine Similarity) を計算し, これを評価スコアとした.

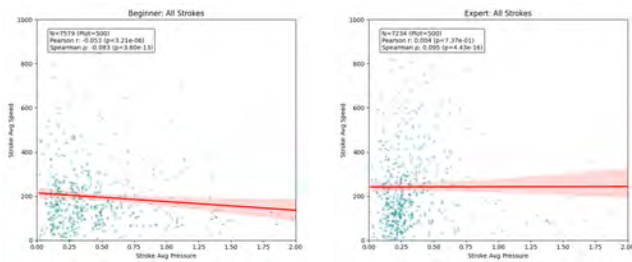
$$\text{Similarity} = \frac{v_{teacher} \cdot v_{user}}{\|v_{teacher}\| \|v_{user}\|} \quad (2)$$

ここで,  $\cdot$  はベクトルの内積,  $\|v\|$  はベクトルのユークリッドノルム (大きさ) を表す. このスコアは  $-1$  から  $1$  の値をとり, 値が  $1$  に近いほど手本画像との特徴空間上の距離が近く, 高い再現度であることを示す.

本研究では, CLIP によるスコアを「意味的類似度 (Semantic Similarity)」, DINOv2 によるスコアを「構造的類似度 (Structural Similarity)」と定義し, 熟練度による差異を比較検証した.

## 4. 分析結果

本章では, 収集したログデータと評価スコアに基づき, 経験者と初心者の差異を定量的に分析する. なお, 後述する被験者間の比較散布図 (図 7 以降) において, プロットの色は被験者属性を表しており, 青色は初心者 (Beginner), 赤色は経験者 (Expert) に対応する. また, 各群の全体的な



(a) 初心者 (Beginner) (b) 熟練者 (Expert)

図 6: ストロークごとの筆圧と速度の関係。(a) 初心者は筆圧増加に伴い速度が低下する傾向(負の相関)が見られるが、(b) 熟練者は無相関であり、筆圧と速度を独立して制御している。(赤線: 回帰直線, 帯域: 95%信頼区間)

傾向を可視化するため、データ点に加え、線形回帰直線(実線)および95%信頼区間(帯状の半透明領域)を併記した。

#### 4.1 運動制御の特性

個別の部位別分析に先立ち、全ストロークを対象とした平均筆圧と平均速度の相関分析を行った。結果を図6に示す。分析の結果、初心者と経験者の間で対照的な運筆特性が確認された。図6に示す通り、初心者においては、平均筆圧と平均速度の間に統計的に有意な負の相関(Pearson  $r = -0.053, p < 0.001$ )が認められた。これは、描画に際して筆圧を強めるほど運筆速度が低下するという、運動制御上の制約を受けている実態を反映している。一方で、図6に示す経験者においては、筆圧と速度の間に相関は見られなかった( $r = -0.004, p > 0.05$ )。この結果から、経験者は筆圧の強弱を自在にコントロールしながらも、一定の描画リズムや速度を維持できる高度な運動制御能力を有していることが推察される。

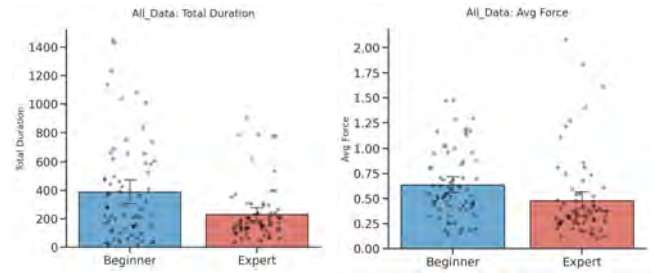
#### 4.2 効率性と描画クオリティ

次に、運動制御の特性が最終的な「絵の質」と「効率」にどう影響したかを、AI評価指標(CLIP, DINOv2)を用いて定量的に分析した。

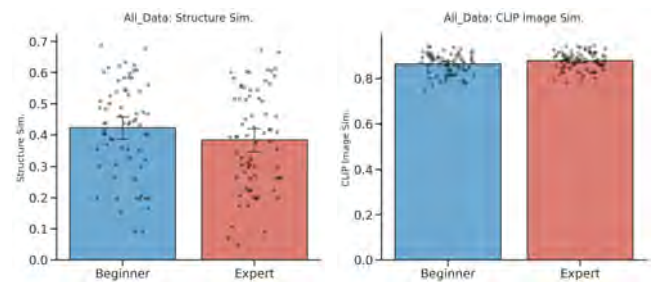
##### 4.2.1 コストと成果の不均衡

図7に、運動コスト(時間・筆圧)とAI評価スコアの比較を示す。まず、成果物(下段)に注目すると、幾何学的な「形の正確さ(Structure Sim.)」においては、初心者と経験者に統計的に有意差が見られなかった( $p > 0.05$ )。これは、初心者が時間をかけて丁寧になぞることで、座標的な正確さは達成できていることを示唆している。

しかし、その同等のスコアを出すために支払ったコスト(上段)には劇的な差がある。初心者は経験者に比べて、有意に高い筆圧( $p < 0.01$ )と、約1.7倍もの描画時間(Total Duration)を費やしている。一方で、「画風や雰囲気(CLIP Image Sim.)」のスコアにおいては経験者が有意に高く、初



(a) 描画時間(コスト) (b) 平均筆圧(コスト)



(c) 形の正確さ(DINO) (d) 画風・雰囲気(CLIP)

図7: コストと成果の比較。初心者は(a)(b)の高コストを支払っているが、(c)形の正確さに差はなく、(d)画風では劣る「非効率」な状態にある。

心者は時間をかけても経験者の持つ「ニュアンス」までは再現できていないことが示された。

##### 4.2.2 非効率性の可視化

この「努力が報われていない状態」をより詳細に分析するため、個人の全描画時間とDINOスコア(形の正確さ)の関係を図8に示す。初心者のデータ分布(青)を見ると、横軸(時間)が右に伸びても、縦軸(スコア)は上昇せず、むしろ横ばいになっている。これは、初心者が費やしている膨大な時間が、スコア向上に寄与しない「浪費」に近い状態であることを示唆している。対照的に、熟練者(赤)は左上に集中しており、短時間で高いスコアを安定して出していることがわかる。

さらに、この数値的な傾向が実際の描画プロセスにおいてどのように現れているかを図9に示す。図の上段に示す経験者は、開始から160秒の時点で描画を完全に終了しており、その時点で既に高い完成度(CLIPスコア)を達成している。対照的に、下段の初心者は同時点(160秒)において、ようやく顔の輪郭を描き終えた段階に留まっている。初心者はここから細部の描き込みや修正に時間を費やし、最終的な完成までには熟練者の約3.3倍となる530秒を要した。

この時系列比較から、初心者は「画力がない(描けない)」のではなく、「正解の線を探るプロセス(細部の修正)」に膨大なコストを支払っているという非効率性が視覚的にも確認できる。表1に、各指標における経験者と初心者の比較結果(t検定)を示す。まず、運動コストの観点を確認する

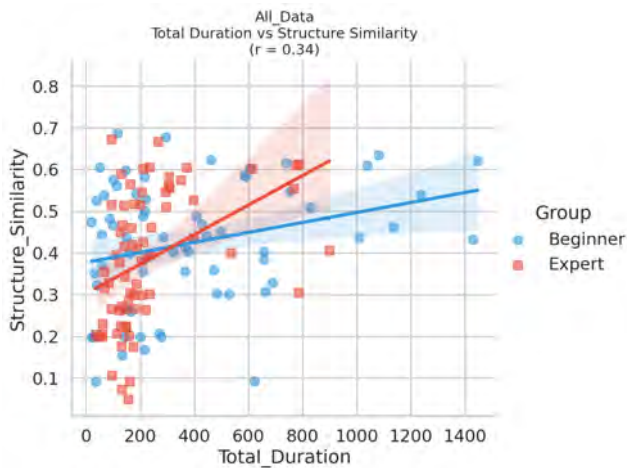


図 8: 描画時間と DINO スコアの散布図. 初心者は時間をかけてもスコアが頭打ちになっている (非効率性) .

と、総描画時間 (Total Duration) において有意な差が見られ ( $p < 0.01$ ), 初心者は熟練者よりも長時間作業を行っていることが示された。また、平均筆圧 (Avg Force) においても有意差 ( $p < 0.05$ ) が確認され、初心者は常に高い筆圧負荷をかけている実態が明らかになった。一方で、平均速度 (Avg Speed) には有意な差が見られなかった ( $p > 0.05$ )。これは、初心者が単に手を動かすのが遅いのではなく、描画と停止を繰り返すなどの制御面で時間を費やしていることを示唆している。

次に、深層学習モデルを用いた客観的指標に基づき、成果物の品質を検証する。興味深いことに、幾何学的な構造評価である DINO スコアにおいては、両群に有意な差は見られなかった ( $p = 0.13$ )。これは、「初心者は時間をかければ、経験者と同等の『形の正確さ』を達成できる」ことを統計的に裏付けている。しかし、画風や意味的整合性を評価する CLIP スコアにおいては有意な差 ( $p < 0.05$ ) が認められ、経験者は効率的に高い表現力を発揮しているのに対し、初心者は時間をかけても画風の再現には至っていないことが定量的に示された。

#### 4.3 描画対象による特性の違い

全体的な傾向に加え、描画対象 (キャラクターの属性) による詳細な違いについても分析を行った。図 10(a) は、複雑な線画を要する「老人男性 (Old Man)」における、ストローク数と平均筆圧の関係である。全体的に負の相関 ( $r = -0.52$ ) が見られ、特に経験者 (赤) はストローク数が増えても筆圧を低く保っている傾向がある。これは、シワや陰影などの細かいディテールを描き込む際、経験者が筆圧を抜いて繊細なタッチを用いていることを示唆している。

一方、図 10(b) は、比較的シンプルな線で構成される「若年男性 (Young Man)」における、描画速度と画風スコア (CLIP) の関係である。ここでは負の相関 ( $r = -0.51$ ) が

見られ、速度が速すぎる (雑になる) と、AI による画風評価が低下する傾向が確認された。さらに、図 10(c) は、同対象 (若年男性) における描画時間と構造的類似度の関係を示している。全体として正の相関 ( $r = 0.63$ ) があり、時間をかけるほど手本の構造に近づく傾向が見られる。特筆すべきは、同一の描画時間帯において、経験者 (赤) の回帰直線が初心者 (青) よりも常に高い位置にある点である。これは、経験者が初心者と同じ時間を費やした場合により正確な形状を描けること、すなわち「構造把握の効率性」が高いことを示している。初心者は時間をかけても経験者の短時間のスコアに及ばないケースがあり、無駄な修正や迷いが生じている可能性が推察される。

以上の結果から、効果的な描画支援においては以下の 3 点が重要であると結論付けられる。第一に、複雑な対象 (老人) を描く際の「脱力 (筆圧制御)」の習得。第二に、単純な対象 (若者) における「速度と質のトレードオフ」の理解。そして第三に、時間をかけるだけでなく、効率的に正しい形状を捉えるための「構造的観察眼」の育成である。

## 5. 考察

### 5.1 「非効率な努力」の正体

本研究の結果、初心者は経験者と同等の「形の正確さ (DINOv2 類似度)」を達成しているにもかかわらず、そのために膨大な時間と筆圧を費やしていることが明らかになった。これは、初心者が創造的な「描画 (Drawing)」ではなく、既存の線をなぞるような「位置合わせ (Tracing)」に近い作業に認知資源を割きすぎていることを示唆している。すなわち、初心者は局所的な「位置の正解」を探索することに固執するあまり、運動の自由度を凍結させ、「迷い続ける状態 (高筆圧かつ低速な運筆)」から脱却できていない可能性が高い。

### 5.2 支援システムへの示唆

この知見は、今後の描画支援システムの設計に重要な指針を与える。従来のように「形がズレていること」を指摘するだけでは、初心者はさらに慎重になり、筆圧が上がり、非効率性が加速する恐れがある。むしろ、形が多少ズレていても、「筆圧を抜くこと」「一定の速度でストロークを描き切ること」をフィードバックする方が、結果として経験者のような「画風 (CLIP スコア)」の向上につながる可能性がある。効率的なストロークへの変容を促すことで、初心者は「作業」から解放され、より本質的な「表現」に注力できるようになると考えられる。

### 5.3 今後の課題

本研究にはいくつかの課題が残っている。第一に、被験者数が計 23 名と小規模である点、およびタスクが「アニメ調の顔画像の模写」に限定されていた点である。第二に、

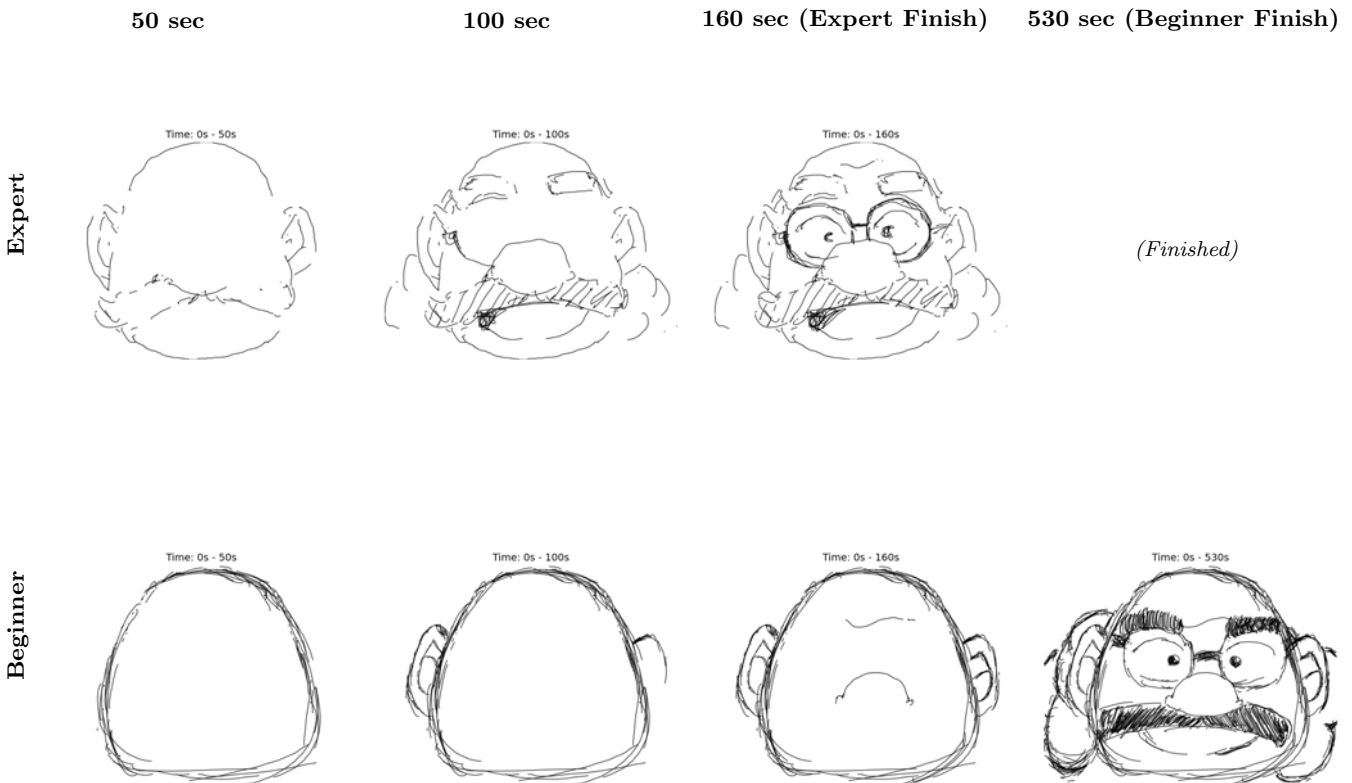


図 9: 描画プロセスの時系列比較. 上段の熟練者は 160 秒で完成しているのに対し, 下段の初心者は同時点では輪郭のみであり, 完成までに約 3 倍の時間を要している (非効率性).

表 1: Comparison of metrics between Beginners and Experts. (Mean  $\pm$  SD)

Metric	Beginner	Expert	<i>t</i> -value	<i>p</i> -value
Total Duration	388.15 $\pm$ 353.78	230.96 $\pm$ 189.23	3.25	0.002**
Avg Force	0.64 $\pm$ 0.33	0.48 $\pm$ 0.39	2.59	0.011*
Avg Speed	247.46 $\pm$ 153.45	226.58 $\pm$ 104.26	0.94	0.351
Structure (DINO)	0.42 $\pm$ 0.15	0.38 $\pm$ 0.16	1.52	0.130
Style (CLIP)	0.87 $\pm$ 0.05	0.88 $\pm$ 0.04	-2.14	0.034*

\**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001

実験設定における「描画スタイルの自由度」に関する課題が挙げられる。本実験では被験者に自由なスタイルでの描画を求めたため、「着色（塗り）まで行う被験者」と「線画のみで終了する被験者」が混在することとなった。着色動作（塗りつぶし）と線画動作（輪郭線）では、筆圧や速度といった運動特性が大きく異なる可能性があるが、本分析ではこれらを区別せずに解析を行っている。今後は、線画と着色を明確に分けたタスク設定や、ストロークの種類（線か塗りか）を自動判別する手法の導入により、より精緻な分析を行う必要がある。

今後の展望として、本研究で得られた知見（初心者の高筆圧と非効率性）に基づき、リアルタイム描画支援システムの開発を行う。具体的には、筆圧が過剰に高まった瞬間

にフィードバックを送ることで、初心者の運動制御を直接的に矯正し、効率的な描画スタイルへ誘導する効果を検証していく予定である。

## 6. 結論

本研究では、描画プロセスにおける熟練者と初心者の差異を定量的に解明するため、ストロークに基づく運動解析と、深層学習モデル（CLIP・DINOv2）による画像評価を統合した分析フレームワークを提案した。検証実験として、独自開発したロギングシステムを用い、熟練者 12 名と初心者 11 名を対象とした顔画像の模写タスクを実施した。課題には、マンダラート分析に基づき選定された、異なる運動特性（筆圧制御・ストローク安定性）を要する 2 種類のモ

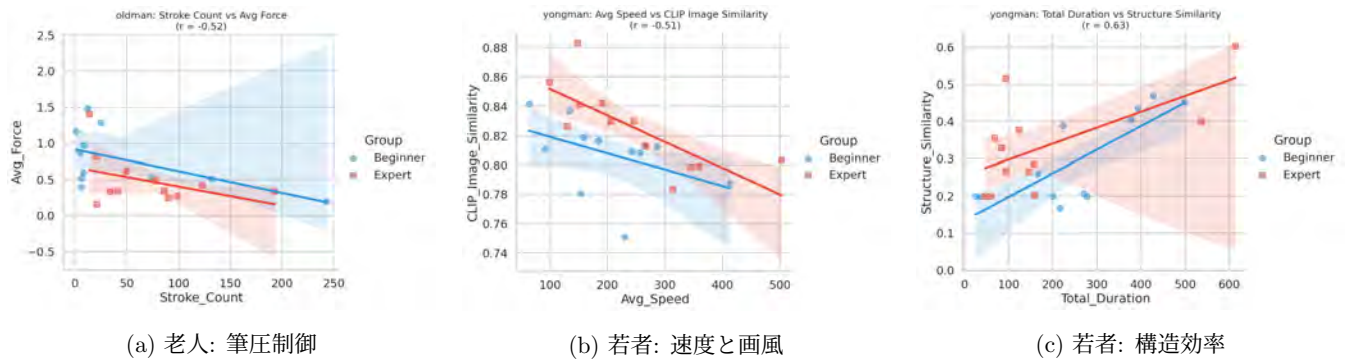


図 10: 描画対象別のスキル分析.(a) 複雑な対象 (老人) では低筆圧化が重要.(b) 単純な対象 (若者) では速度超過による画風低下が見られる.(c) 若者画における時間と構造再現性の関係.

chiefs を用い、収集した多次元データ (筆圧・速度・時間・画像スコア) を分析した。

その結果、以下の知見が得られた。第一に、運動制御の観点において、初心者は筆圧と速度の間に強いトレードオフ (不自由さ) を抱えており、円滑な運筆が阻害されていることが確認された。第二に、最終的な成果物において、初心者は時間を費やすことで「形 (DINO スコア)」を合わせることはできているものの、熟練者のような「画風 (CLIP スコア)」の再現には至っていないことが明らかになった。これらの知見から、初心者が直面している本質的な課題は、単なる「画力の低さ」ではなく、投入したコスト (時間・筆圧) を成果に効率よく変換できない「非効率性」にあると結論付けられる。

## 謝辞

本研究は、JST,BOOST,JPMJBY24D6 の支援を受けたものである。本実験の遂行にあたり、多大な時間と労力を割いて描画タスクにご協力いただいた被験者の皆様に、心より感謝申し上げます。

## 参考文献

- [1] Goel, V.: *Sketches of thought*, MIT press (1995).
- [2] Schön, D. A.: *The reflective practitioner: How professionals think in action*, Routledge (2017).
- [3] Chen, H.-H. and You, M.: The comparison between novice and expert designers' sketching in conceptual design, *Futureground-DRS International Conference*, pp. 17–21 (2004).
- [4] Lohmeyer, Q., Meboldt, M., Matthiesen, S. et al.: Analysing visual strategies of novice and experienced designers by eye tracking application, *DS 76: Proceedings of E&PDE 2013, the 15th International Conference on Engineering and Product Design Education, Dublin, Ireland, 05-06.09. 2013*, pp. 202–207 (2013).
- [5] Xu, P., Hospedales, T. M., Yin, Q., Song, Y.-Z., Xiang, T. and Wang, L.: Deep learning for free-hand sketch: A survey, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 45, No. 1, pp. 285–312 (2022).
- [6] Tchalenko, J.: Drawing from life: The inner eye and external support, *BMC Neuroscience*, Vol. 10, No. 1, p. 74

- (2009).
- [7] Ha, D. and Eck, D.: A Neural Representation of Sketch Drawings (2017).
- [8] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J. et al.: Learning transferable visual models from natural language supervision, *International conference on machine learning*, PMLR, pp. 8748–8763 (2021).
- [9] Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A. et al.: DINOv2: Learning robust visual features without supervision, *arXiv preprint arXiv:2304.07193* (2023).
- [10] Suwa, M. and Tversky, B.: What do architects and students perceive in their design sketches? A protocol analysis, *Design studies*, Vol. 18, No. 4, pp. 385–403 (1997).
- [11] Akamatsu, Y., Maeda, K., Ogawa, T. and Haseyama, M.: Classification of expert-novice level using eye tracking and motion data via conditional multimodal variational autoencoder, *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 1360–1364 (2021).
- [12] Vinker, Y., Pajouheshgar, E., Bo, J. Y., Bachmann, R. C., Bermano, A. H., Cohen-Or, D., Zamir, A. and Shamir, A.: Clipasso: Semantically-aware object sketching, *ACM Transactions on Graphics (TOG)*, Vol. 41, No. 4, pp. 1–11 (2022).
- [13] Huang, Z., Xie, H., Fukusato, T. and Miyata, K.: Anifacedrawing: Anime portrait exploration during your sketching, *ACM SIGGRAPH 2023 conference proceedings*, pp. 1–11 (2023).
- [14] Huang, Z., Peng, Y., Hibino, T., Zhao, C., Xie, H., Fukusato, T. and Miyata, K.: dualface: Two-stage drawing guidance for freehand portrait sketching, *Computational Visual Media*, Vol. 8, No. 1, pp. 63–77 (2022).
- [15] 読売テレビ: 名探偵コナン 「キャラクター紹介 阿笠博士」, <https://www.ytv.co.jp/conan/character/agasa/>. (参照 2026-01-06).
- [16] 双葉社・シンエイ・テレビ朝日・ADK: クレヨンしんちゃん 公式ポータルサイト 「作品紹介」, <https://www.shinchan-app.jp/about/index2.php>. (参照 2026-01-06).