

# AIによるモーションピクトグラム自動生成と ユーザビリティへの影響に関する研究

Research on AI-Based Automatic Generation of Motion Pictograms and Effects on Usability

岡谷 夏実<sup>1)</sup>, 塩谷 隆二<sup>2)</sup>, 中林 靖<sup>3)</sup>

Natsumi Okatani, Ryuji Shioya and Yasushi Nakabayashi

- 1) 東洋大学大学院 (〒350-0815 埼玉県川越市鯨井2100, E-mail: s4b102210019@toyo.jp)  
2) 工博 東洋大学 総合情報学研究科 教授 (〒350-0815 埼玉県川越市鯨井2100, E-mail: shioya@toyo.jp)  
3) 工博 東洋大学 総合情報学研究科 教授 (〒350-0815 埼玉県川越市鯨井2100, E-mail: nakabayashi@toyo.jp)

This research aims to improve usability by using AI to automatically create motion pictograms. Traditional still pictograms have limitations in conveying complex information and dynamic concepts, but motion pictograms can overcome these limitations through dynamic visual expression. However, the creation of motion pictograms requires specialized skills and high costs, which has prevented their widespread use. The automatic generation of motion pictograms based on AI will solve these problems.

**Key Words :** Motion, Pictograms, AI, User Experience (UX)

## 1. はじめに

ピクトグラムは文字情報を必要とせず情報を伝達するグラフィックシンボルであり、特定の意味を視覚的に認知させる(1)。文字情報がほとんど不要である特性から、特に公共空間のように多様な人々が利用する場面において、言語よりも迅速かつ直感的に伝達する手段として有効である。ピクトグラムは危険の警告や、災害時の適切な行動提示など、安全の確保に使用される場合もあり、公共の安全性においてもその有効性が活用されている。

しかし、ピクトグラムの有効性は、普遍性と文化や文脈による解釈の多様さという二面性を持つ。広く認知されたものは効果的に機能する一方、新たなピクトグラムの浸透には時間がかかり、複雑な情報を表現する際には誤解が生じる可能性が指摘されている。静止画による表現の限界は、動的な状況や複雑な概念の伝達において特に問題となる。既知のピクトグラムであれば迅速な理解が可能だが、未知のもの、加えて文化や表現の違いを含む場合、理解の難易度は高まり、静止画ではその内容を十分に表現できないこともある。その結果、誤解や混乱を生じ、かえって危険を引き起こす可能性もある(2)。これに対し近年では、モーショングラフィックスを用いて情報を拡張した「モーションピクトグラム」が注目されている。静止画ピクトグラムに比べて動的な表現が可能であり、複雑な情報を直感的に伝達できる。

AIによる画像から動画への自動生成技術は、教育、エンターテイメント、広告など、幅広い分野で導入されつつある。特に、静止画の生成コストは映像に比べて低い傾向にある。そのためAIによる自動生成は、低コストで高品質な動画コンテンツ制作を可能にし、情報伝達の幅

を広げる新たな可能性を開拓している。しかしAIによる画像認識や判定、自動生成には、生成結果の不確実性が課題として残る。エンターテインメント分野であれば、最終的なリリースに至るまでに人の手による検証と修正が行われ、出力時点での不確実性は許容される場合が多い。しかし、交通や医療といった安全性が求められる分野におけるAI導入では、その不確実性を極力低減しなければならない。特に自動運転など、AIの出力に対して人間による確認を通してユーザーへと提供するシステムにおいては、誤動作が人命に関わる可能性もあるため、高い信頼性と安全性の実現が求められる。

## 2. 目的

本研究の目的は、AIを用いて静止画ピクトグラムを認識し、生成したモーションピクトグラムを活用することで、より効果的かつ効率的な情報伝達を実現し、その普及を促進することである。静止画では表現が困難であった動的情報や複雑な意味を、モーションピクトグラムによって直感的に伝えることで、理解の促進と誤解の防止を目指す。また、従来は専門知識や技術を要していたモーションピクトグラムの作成を自動化することで、低コストかつ迅速な制作が可能となり、利用の拡大が期待される。またその一方で、AIによる生成には品質と信頼性の担保が求められる。とくに緊急時や災害時など、ピクトグラムが安全確保に直結する場面では、生成結果の正確さが極めて重要となる。意図しない意味解釈が誤った行動を誘発する可能性があるため、生成されたピクトグラムが正しく情報を伝えるかを評価・検証することが不可欠である。

### 3. 方法

本研究では、静止画ピクトグラムからその表す意味に即したモーションピクトグラムを生成するシステムの開発検討を考える。生成は大きく以下の3つのプロセスからなる。第一にAIおよび画像認識による静止画ピクトグラムの認識および解析、第二にAIによる映像生成、第三に生成結果の確認である。図-1に提案システムの構成を示す。

#### (1) 提案システムの構成

**静止画ピクトグラムの解析** 入力として静止画ピクトグラムを与える、AIによってピクトグラム内の物体や形状

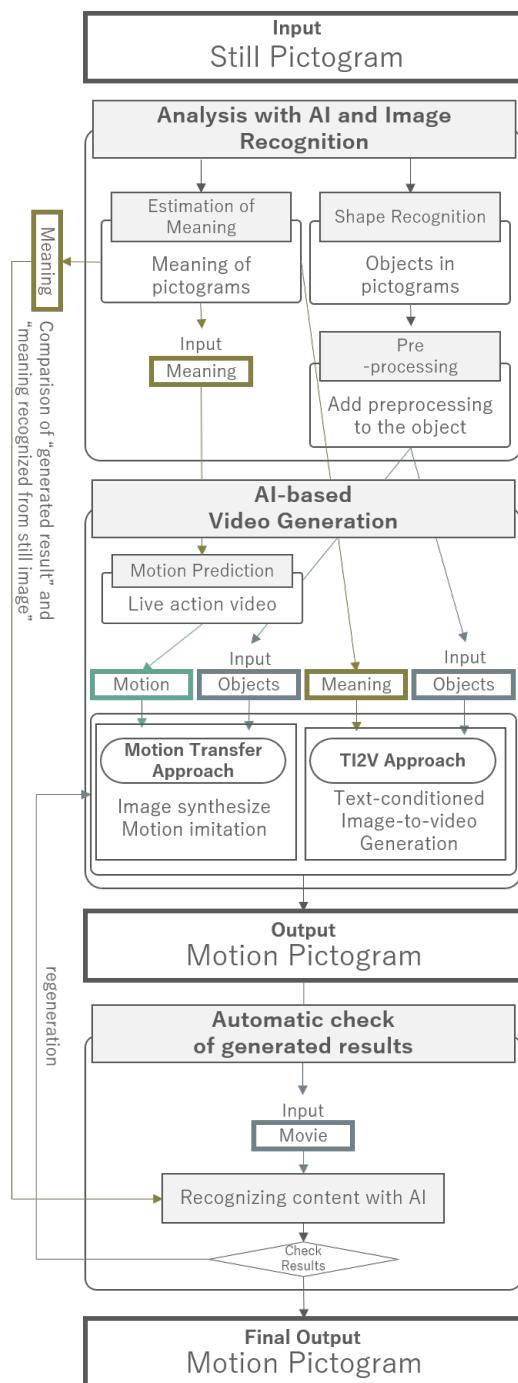


図-1 提案システム構成

の認識を行う。必要に応じて認識された物体に対する前処理を行い、映像生成時に変換しやすいオブジェクトとして次ステップへと渡す。一方で、ピクトグラムの表す意味の推定を行い、次項の映像生成プロセスにおける条件情報として入力プロンプトの一部に利用する。

#### (2) 映像生成

静止画像から得られたオブジェクトおよび意味情報、条件情報をもとに、AIを用いてモーションピクトグラムの映像生成を行う。AIによる映像生成については2通りのアプローチを考える。第一に、モーション転移によるアプローチである。判定した対象に応じた実写映像やアニメーションのモーションを対象の静止画ピクトグラムへと転移させることで、未知のデータであってもモーショングラフィックス化へと対応させることを目指す。第二に、TI2V (Text-guided Image-to-Video) 型の生成AIを用いたアプローチである。このアプローチでは、静止画像とともに、前ステップにおいて推定した意味情報および条件情報をテキストとして条件付けに用い、画像から映像への変換を行う。これにより、テキストによる情報の補完と条件づけを通じて、静止画ピクトグラムに基づいたモーション生成が可能となる。

#### (3) 生成結果の確認と再生

生成されたモーションピクトグラムについては、その内容の品質と信頼性の確認が必要となる。生成映像中に不適切なオブジェクトや表現が含まれていないことや、ピクトグラムが本来意図する意味が適切に表現されているかをチェックすることが求められるため、JIS規格ピクトグラム(5)のような公式ピクトグラムを参照し、人間による主観的評価と一致するように判定を行う。たとえば、「転倒注意」という意味を持つピクトグラムであれば、人物の転倒動作や、注意を喚起するシンボルが明示的に含まれている必要がある。検証の結果、必要な構成要素が不足している場合や意味と映像が一致しない場合には、再度、前段の映像生成プロセスに戻り再生成を行う。これにより、意味と動作の整合性を備えた高品質なモーションピクトグラムの生成を目指す。

以上のようなプロセスにより、静止画ピクトグラムを入力として意味に即したモーションを付与し、安全表示や注意喚起における視覚的伝達力の強化を図る。

### 4. モーション転移による生成

モーションピクトグラム生成の手法として、既存の映像やアニメーションから動作情報を抽出し、それを静止画ピクトグラムへと転移させる、モーション転移による生成を行った。この手法では、動作の自然さや一貫性を保持したまま、静的な図像に対してモーション表現を付与する。以下の2つの事前学習済みモデルを使用しモーションピクトグラム生成を試行した。

## (1) Thin-Plate Spline Motion Model



図-2 静止画ピクトグラムとモーションピクトグラム「非常口」の生成結果

モーションピクトグラム生成のため、Thin-Plate Spline Motion Model(6)を用い、事前学習済みモデル「TaiChiHD」による動作転移を試行した。実写映像を参照とし、静止画ピクトグラムへ動作を転移したが、前処理なしの転移ではノイズが発生し、右足先端の動作も反映されなかつた。背景を除去する前処理を加えることでノイズは軽減されたが、右足の動作は固定されたままであった。これは、実写とピクトグラムの初期姿勢の差や、学習モデルが実写向けであることが原因と考えられる。図-2上側に前処理を加えず行った結果、下側に背景を除去して行った結果を示す。また、人体検出モデルがデフォルメ表現に適しておらず、2~3頭身のイラストでは一部の動作に不完全さが見られた。初期姿勢の調整や認識したオブジェクトに対する前処理の必要性が示された。さらに背景除去や首の追加といった前処理の自動化、OpenPose等を用いた関節検出によるピクトグラム対応型の動作生成システムの開発が求められる。

現状のシステムでは、実写画像の特徴に基づいた検出や生成を行っているため、ピクトグラムの表現に適した改善が必要である。例えば、今回手動で行った認識における背景部分の一時削除や首部分の追加といった前処理を自動で行い、一度モーションを生成したのち、再び背景と合成、モーションピクトグラム化を行うといったプロセスの開発や、ピクトグラムのデフォルメされた形状にOpen Poseなどの人体姿勢推定モデルを適用し、関節の位置座標を取得といった方法が考えられる。こうして認識した特定の関節の位置座標を、今回使用したiPERCoreの人体認識部分に渡すこと、モーションピクトグラム化を自動的に動作させる。

## (2) Liquid Warping GAN with Attention

Liquid Warping GAN with Attention(7)は、3次元人体メッシュの復元により回転や関節位置に対応できる。本研究ではその再現手法であるiPERCoreを用いて、モーションピクトグラム生成を試行したが、「非常口」などのピクトグラムでは人体が認識されず、誤認識により動作転移が不完全となった。そこで、比較的シンプルな「男性」

「女性」のピクトグラムに対して、動作転移を再試行した。これは、前処理を行う事で動作した。このため、人体認識を促進するための前処理として、カラー化や首の補完、実写写真との合成を行い、動作転移の精度を検証した。「男性」ピクトグラムに対する結果を図-3に示す。さらに、実写・シルエット・写実性の異なるイラストを

Input (Still-pictogram)	Output (Motion-pictogram)
Original	The human body was not recognized
Colored	The human body was not recognized
	Added neck part
	clipped from a photo

図-3 ピクトグラム「男性」における実行結果

Input (Still-pictogram)	Output (Motion-pictogram)
Original	

図-4 シルエットにおける実行結果

用いて、人体認識に必要な条件を調査した。検証の結果、首の追加により一部動作の転移、実写画像のクロッピングでは安定した転移が動作した。写実性の高いイラストやシルエットでも良好な結果が得られたが、元ピクトグラムのカラー化のみでは人体を認識しなかった。

## 5. モーションピクトグラムの効果検証

上述の通りモーションピクトグラムの自動生成には課題が残るが、ピクトグラムのモーション化を行う意義と効果を明確化および、自動生成時の手法の改善や最適化に役立てるため、静止画ピクトグラムから手動で作成したモーションピクトグラムを用いて、ユーザ評価実験を通じた静止画ピクトグラムとモーションピクトグラムの効果比較検証を行った。この検証により、ユーザビリティ視点および情報の伝達効果といった方向性から、モーションピクトグラムが情報伝達に与える影響を確認し、意味理解の正確性や主観的評価の変化を測定した。対象としたピクトグラムはJIS規格に基づく5種（A: 転落注意、B: 乗り継ぎ、C: 靴を脱いでください、D: スロープ、E: 洪水／内水氾濫）である。

## (1) 検証実験構成

使用したモーションピクトグラムは、各静止画をもとに動画作成ソフトで手動作成したものであり、3~7秒程度のモーションを経て元の静止画と一致する状態に戻る構成とした。評価実験の対象は10代~20代の大学生48名とし、次の3ステップで実施した。(1) 静止画ピクトグラムを提示し、その意味を自由記述で回答させる。(2) モーションピクトグラムを提示し、その意味を尋ねるととも

に、静止画との印象の違いを5段階で評価。(3)全体を通しての感想や印象を記述式で収集。

各ピクトグラムについて、正解とみなす条件は以下のように設定した。

- A. 「転落」「注意」の語句やこれを表す語句が解答に含まれる場合
- B. 「飛行機を乗り継ぐ」ことが言及されている場合
- C. 「靴を脱ぐ」ことが言及されている場合
- D. 「スロープ」「車いすで登れる」のいずれかが言及されている場合
- E. 「洪水」「氾濫」について言及している場合

## (2) 結果

実験の結果、最も大きな効果が確認されたのはBの「乗り継ぎ」であり、静止画での正解率2.9%に対し、モーション化後には76.3%と大幅に上昇した。5段階評価の平均も4.29と高く、多くの回答者がモーションによって意味の理解が促進されたと回答している。他のピクトグラムにおいても概ね正解率は向上し、特に動作や状況の変化が

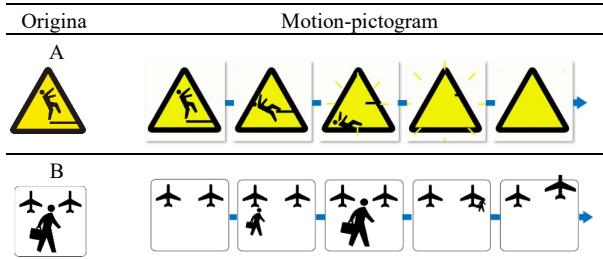


図-5 実験で使用したモーションピクトグラム

(A: Caution, drop; B: Connecting flights)

表-1 正解率平均の比較

	A	B	C	D	E
Pictograms	Caution, drop	Connecting flights	Take off your shoes	Slope	Flood from rivers/Flood from inland waters
$x_1 : Q1$ Average	85.3%	2.9%	63.4%	80.0%	5.9%
$x_2 : Q2$ Average	93.2%	76.3%	88.1%	81.0%	12.8%
$\Delta x = x_2 - x_1$	+ 7.9%	+ 73.3%	+ 24.7%	+ 1.0%	+ 6.9%
Sensory comparison between still and moving pictograms					
Average	3.92	4.29	3.95	3.32	3.50
Max 5.00					

\*今回は予め設定した条件により採点を行ったうえで算出したが、解答は記述式としているため、採点者、採点条件によっては、正解率に若干の誤差が生じる可能性がある。

視覚的に明確であったものほど効果が高い傾向が見られた。一方で、すべてのピクトグラムで一様に効果が現れたわけではなかった。たとえば、意味が抽象的または文化的前提知識を必要とするものについては、モーションによっても理解が困難な例が存在した。また、「動作はわかったが、それが禁止なのか指示なのか判断できなかった」といった指摘も複数あった。これは、日本の災害標識において色や形により「指示」「注意」「禁止」などの分類が視覚的に設計されていることと関係する。ユーザがそうした前提知識を持たない場合、モーションのみでは意図を完全に伝えることは困難であることが示唆された。このことから、モーションピクトグラム生成システムには、単に静止画に動きを付けるだけでなく、ピクトグラムの意図を正確に表現できるよう、動作の選定や構成の工夫が必要である。たとえば、意図する「指示の分類」（誘導・警告・禁止など）をシステム側で認識し、それに応じたモーションを用いた出力や表現を考えられる。また、前処理段階でオブジェクト抽出に加えて意味認識を行い、分類に応じて最適な動作を合成し再生成する処理の実装が求められる。

さらに、自由記述による感想では、「動画にすることでイメージしやすくなった」「危険度が伝わりやすい」といった肯定的な意見が全体の72.5%を占めた一方、「アニメーションの意図が曖昧」といった改善を示唆する意見もあった。これらのフィードバックは、モーションによる伝達効果や一貫性を確保するための指針となり得る。

以上の検証結果によりモーションピクトグラムの意義と効果が示され、また、今後のシステム開発における伝達内容の指示部分の明確化や、モーション表現の改良および再生産の重要性が示された。

## 6. 結論

本研究では、静止画ピクトグラムの情報伝達機能を拡張する手法として、AIによる自動生成によるモーションピクトグラムの可能性を検討した。AIの画像認識および生成技術を活用し、静止画から意味情報を抽出、対応するモーションを付与することで、複雑な情報の直感的な伝達や誤解の低減を行うとともに、従来は専門性を要していたモーションピクトグラムの作成を自動化することで、その普及の促進を目指す。

事前学習済みモデルを用いたモーション転移による生成を通じて、現段階では静止画ピクトグラムの特性に適した前処理の必要性や、人体認識精度の向上、動作の一貫性確保など複数の課題が明らかとなった。一方、ユーザ評価実験では、特に動作を示す場合や元のピクトグラムの認知度が低いピクトグラムにおいて、モーション化による理解促進効果が顕著に現れた。特に「乗り継ぎ」など静止画では理解が困難であった内容が、モーションの効果によって改善することが示された。

また、モーションピクトグラムの評価においては、意

味理解の正確性のみならず、「指示」「注意」「禁止」といった意図の区別や、文化的背景に基づく認識の違いにも配慮する必要があることが示唆された。これにより、単なる動作の付与にとどまらず、伝達すべき内容の明示的な構成の必要性がわかった。

今後はAIによる自動生成の信頼性を確保するとともに、ピクトグラムの用途に応じた分類情報の導入や、ARなどを活用した多様な提示方法の検討を進める必要がある。本研究の成果は、安全性が重視される場面における情報提示の質の向上に貢献し、ピクトグラムの新たな展開可能性を示すものである。

結論として、AI生成モーションピクトグラムの活用は、ピクトグラムの効率化と高度化に大きく貢献する可能性を持っている。コミュニケーションの障壁を低減し、より包容的な社会の実現に貢献することが期待される。その実現のために、自動チェックシステムの開発、生成モデルの高度化、そして人間による評価の継続といった多角的なアプローチが必要不可欠である。将来的な展望として、ARグラスをはじめとした関連デバイスの発展および一般への普及とともに、本研究における認識や自動生成といった利点の活用にも期待したい。

### 参考文献

- [1] 太田幸夫，“国際安全標識のピクトグラムデザインの研究”，平成15年度 共同研究費研究成果報告書, 2004.
- [2] Tijus, C., et al. “The design, understanding and usage of pictograms”. In G. Rijlaarsdam (Series Ed.) and D. Alamargot, P. Terrier, & J.-M. Cellier (Vol. Eds.), Studies in Writing, Vol. 21, Written Documents in the Workplace (pp. 17–31). 2007.
- [3] 岡谷夏実, 塩谷隆二, 中林靖, 多田光利, “AIを用いたピクトグラムの モーション化システムの開発”, 計算工学講演会論文集 Vol.28, C-10-01, 2023.
- [4] 岡谷夏実, 塩谷隆二, 中林靖, 多田光利, “AI と AR を用いた平面図の 3 次元モデル化システムの開発”, 日本計算工学会論文集, 2022巻, p.20220009, 2022
- [5] 国土交通省: 案内用図記号 (JIS Z8210) , 案内用図記号 (JIS Z8210) (令和元年 7 月 20 日) , 2019.
- [6] Jian Zhao Hui Zhang: Thin-Plate Spline Motion Model for Image Animation, School of Software, BNRIst, Tsinghua University, Beijing, China, 2022.
- [7] Wen Liu, Zhixin Piao, Zhi Tu, Wenhan Luo, Lin Ma, Shenghua Gao: Liquid Warping GAN with Attention: A Unified Framework for Human Image Synthesis, School of Software, BNRIst, Tsinghua University, Beijing, China, 2022.