

ピクトグラムの 自動モーション化システムの開発

Instructions for Preparation of Manuscripts
for Proceedings of the Conference on Computational Engineering and Science

岡谷夏実¹⁾, 塩谷隆二²⁾, 中林靖²⁾

Natsumi Okatani, Ryuji Shioya and Yasushi Nakabayashi

1) 東洋大学 総合情報学研究科 (〒350-8585 埼玉県川越市鯨井2100, E-mail: s4b102210019@toyo.jp)

2) 工博 東洋大学 総合情報学部 教授 (〒350-8585 埼玉県川越市鯨井2100, E-mail: shioya@toyo.jp)

2) 工博 東洋大学 総合情報学部 教授 (〒350-8585 埼玉県川越市鯨井2100, E-mail: nakabayashi@toyo.jp)

Pictograms convey meaning through graphic symbols; however, there are limitations when using static pictograms, which has motivated research into using animated pictograms (referred to as motion pictograms). Motion pictograms offer intuitive comprehension; however, their limited adoption stems from higher production costs compared to static pictograms. A potential solution to realize improved adoption of motion pictograms lies in automating the process of generating pictograms from still images using artificial intelligence technologies. Thus, in this study, a pretrained model was used to identify relevant key constraints and related challenges. Experimental results demonstrate the effectiveness of motion pictograms, highlighting their many advantages and paving the way for broad adoption and improved accessibility.

Key Words : Artificial Intelligence; Image Recognition; Motion Graphics; Pictograms.

1. はじめに

(1) 背景

ピクトグラムとは意味するものの形を使ってその意味概念を表すグラフィックシンボル(図記号)であり, 言語に頼ることなく, 特定の意味を視覚的に認知させる[1]. 公共・一般施設, 交通施設, 商業施設, 観光・文化・スポーツ施設のほか, 街の中の道路や空間, あらゆる施設で目にすることができる. 指示や規制, 強制, 警告, 禁止情報を伝える機能を持っており, 情報が迅速に処理される必要がある場合や, ユーザーが非母国語話者である場合, 識字能力の差異や教育的格差により言語能力が制限されている場合, 高齢者などで視覚の問題がある場合などに役立てられる. 言葉よりも直感的かつすばやく意味を伝達する目的で制作され, できる限り文字を使わないという原則のもと, シンプルかつ明確な表現が求められる[2].

ピクトグラムは, 情報や注意を示す汎用性が利点であるが, 欠点として, 認知度の高いピクトグラムは少数であること, 新たなピクトグラムが浸透してから, 効果を最大限発揮できるようになるまでには多くの時間を要することが挙げられる. 意味の理解に時間を要したり, 違う意味に誤解されたりしてしまうものも存在している. 文化や表現が異なる場合には, さらに理解が難しくなる. 静止画のシンプルな表現では, 複雑な内容を表しきれない状況もあり, 誤解や混乱により, かえって危険を生む可能性がある[3].

これに対し, アニメーションにより情報を拡張する, モーショングラフィックス化されたピクトグラムが登場した. アニメーションであれば, 利用者はより直感的に, 素早く意味を理解することができ, 提供者はより正確に指示や注意を伝えることが可能になる. 「言語を必要とせず意味内容を伝える」というピクトグラムの意義を強化でき, 危険の回避や注意を効果的に促すことが可能となる. 本研究では, これを「モーションピクトグラム」と称する.

(2) 目的

ピクトグラムはJIS規格ピクトグラム[4]として策定されている104項目をはじめ, 公用, 民間, 国内外で多種多様な表現が存在し, 社会のニーズに応じて日々新たに作成, 更新が行われている. 一方で, 映像作成は静止画に比べ高度な知識や技術が必要になり, コストも多くかかる. 例えば, 世界最大級のクラウドソーシングサイトのひとつ, upwork[5]では, モーショングラフィックスデザイナーの雇用は1時間あたり28ドル~325ドルで行われている. こうした高額な費用は, モーションピクトグラムが浸透しにくい一因でもある.

そこで, 本研究では, AIを用いて静止画のピクトグラムを認識し, モーショングラフィックス化を行うシステムの開発を目指す. 従来の制作コストを大幅に削減でき, 設置場所の資金力に依存せずに導入, 設置できる可能性が高まる. モーションピクトグラムによる情報伝達を, より現実的な手段として, 広く普及させることが可能になると考えられる. 本研究では主に, 以下の2つを行った.

1. 静止画ピクトグラムからモーションピクトグラムを自動生成する手法の研究
2. 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

これにより、静止画ピクトグラムの課題を克服し、より効果的な情報伝達を実現するための方法を明らかにする。

(3) 関連研究 :AIと画像認識を用いたコンテンツ生成

AIと画像認識による認識結果を用いたコンテンツの生成は、以前に、教育分野で活用を行った。紙面に描かれた空間図形を認識し、結果をもとにした3Dモデルを生成することにより、空間図形の学習支援に適用してきた[6]。本研究では、これを公共空間デザインの分野に適用する。なかでも、静止画ピクトグラムの判定、画像認識や特徴量抽出といった部分に応用・活用することができる。以前の研究と比べて、認識内容をもとにしたコンテンツ生成という部分では共通するが、3Dモデルからアニメーションへと生成内容が変化する点で、これに応じた新たなシステムの構築が必要となる。

2. ピクトグラムモーション化の自動化の検討

静止画のピクトグラムからモーションピクトグラムを生成する過程の自動化を検討するにあたり、事前学習済みモデルを使用したモーションピクトグラム生成を試行した。

(1) 事前学習済みモデルを用いたモーション化

2種類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行った。ピクトグラムには、人型のモチーフを使用したピクトグラムが数多く存在する。そこで、はじめに、人型ピクトグラムのなかでもメジャーな、「非常口」を示す静止画ピクトグラムに対し、実写映像から「走る」動作を転移し、新たなモーションピクトグラム生成を行った。

a) Thin-Plate Spline Motion Model

モーションピクトグラム生成のため、Thin-Plate Spline Motion Model[7]を用い、事前学習済みモデル「TaiChiHD」を使用した動作転移を行った。各フレームの左側にモーション遷移に使用する実写映像、右側が対象となる静止画ピクトグラムを示す。はじめに元のピクトグラムに何も前処理を加えず行った結果では、右足先端部分が動作せずノイズの発生もみられたため、背景を除去しての転移も実施した。Fig.1上側に前処理を加えず行った結果、下側が背景を除去して行った結果を示す。

背景除去により、ノイズは低減できたが、右足先端部分の動作は転移できず固定されたままであった。これは、使



Fig.1 Result of generating the still pictogram and the motion pictogram "Emergency exit".

用した実写映像と対象の人型ピクトグラムの初期姿勢の違いと、使用した学習モデルが原因と考えられる。また、動作転移において、人体検出に使用した学習済みモデルは実写映像に対応したものであり、ピクトグラムのようなデフォルメされた表現には適していなかったものと考えられる。これらの問題を解決するためには、初期姿勢の差分を修正する処理を加え、ピクトグラムの表現に対応した学習モデルを作成、対応させることが必要である。

b) Liquid Warping GAN with Attention

Liquid Warping GAN with Attention[8]は、3次元人体メッシュ復元モジュールによりモデルの体型を再現し、回転や関節位置に対応可能である。Liquid Warping GAN with Attention の再現手法であるiPERCoreを用いた動作転移を試行した。しかしながら、a)で使用した「非常口」の静止画ピクトグラムを含め、多くのピクトグラムで人体が認識されず、誤認識により不完全な動作となった。

そこで、ピクトグラムの中でもシンプルで比較的人体を認識しやすい「男性」「女性」で動作転移を試行した。どちらのピクトグラムも、前処理なく実行した場合には動作しなかった。そこで、認識に必要とされる前処理を検討するため、手動でのカラー化、首部分の補完、実写写真によるクロッピング後の動作転移を試行した。「男性」のピクトグラムにおける結果をTable 1に示す。また、システムが人体構造を認識する要素として必要な条件を調査するため、実写写真、シルエット、写実性・デフォルメ度合いの異なるいくつかのイラストにおいて、動作転移を試行した。

結果として、カラー化のみでは人体は認識されなかった。首部分を追加する前処理を行うと、認識位置は正確でなかったものの、動作はある程度転移された。実写写真のクロッピングでは、どちらも比較的良好な結果となった。一方、実写および頭身がある程度正確なイラスト・シルエットでは動作転移に成功した。

Table 1 Result of ‘Male’ Pictogram

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	

Table 2 Result of Illustration with Silhouette

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

シルエットにおける結果の一部をTable 2に示す。2～3頭身のイラストでは、動作の転移はしたもの、不完全な箇所がいくつかみられた。

このように、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、認識時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システムを開発や、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て、関節の位置座標を取得できるように改良する方法が考えられる。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で、モーションピクトグラム化を自動的に動作させる。このように、現状のシステムでは解決すべき点が多くあり、モーションピクトグラムの生成の自動化には至っていないため、今後も開発を進めていく必要がある。

3. ピクトグラムモーション化の効果検証

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するため、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、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化の意義を明らかにし、システムの開発に役立てる。

(1) 検証実験の実施

実験の構成は以下のとおりである。

- 1) 静止画ピクトグラムの表す意味を尋ねる
 - 2) モーションピクトグラムの表す意味を尋ねる
 - i. 1)のモーション化ピクトグラムの表す意味を尋ねる
 - ii. 静止画での回答時との感覚の違いを5段階で尋ねる
 - 3) 全体を通しての感想、体験の結果を尋ねる
- 対象は大学生を主とする10代～20代48名である。実験で意味を尋ねたピクトグラムは以下の5つで、いずれもJIS規格ピクトグラム[4]である。
- A: 転落注意 Caution, drop
 - B: 乗り継ぎ Connecting flights
 - C: 靴を脱いでください Take off your shoes
 - D: スロープ Slope
 - E: 洪水／内水氾濫 Flood from rivers/Flood from inland waters

Table 3 Motion Pictogram
(A: Caution, drop; B: Connecting flights)

Original	Motion-pictogram
A 	
B 	

Table 4 Comparison of correct percentage averages

	A	B	C	D	E
	Caution, drop	Connecting flights	Take off your shoes	Slope	Flood from rivers/Flood from inland waters
Pictograms					
x1 :Q1	85.3%	2.9%	63.4%	80.0%	5.9%
Average					
x2 :Q2	93.2%	76.3%	88.1%	81.0%	12.8%
Average					
$\Delta x = x_2 - x_1$	+ 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。

使用するモーションピクトグラムは対象のピクトグラムをもとに動画作成ソフトを使用して作成した。モーションは各3秒～7秒程度で、静止画のピクトグラムと一致する状態から動き出し、モーション後、再び元の静止画と一致する状態に戻る。Table 3に例として、A、Bの静止画をもとに作成したモーションピクトグラムをフレームごとに並べた様子を示す。

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」について言及されている場合

(2) 結果と考察

Table 4に正解率と五段階評価の平均をまとめた結果を示す。5つの問題の中で、モーションピクトグラムにした際に最も正解率が向上したのは、Bの「Connecting flights」だった。静止画では正解率が2.9%だったのに対し、動画では正解率が73.3%上昇し、76.3%になった。変化を感じたかについて5段階評価で集計した結果も、平均で4.29と高評価で、実感を伴う効果があったことが分かる。その他

の問題についても、5段階評価およびフィードバックと正解率を照らし合わせながら分析を行った結果、モーショングラムは、静止画と比べて、理解度の向上に一定の効果があることが分かった。ただし、その効果には静止画ピクトグラム自体の特徴、モーシヨンの特徴によって差異がある。加えて、高い効果を発揮するものには、いくつかの傾向が推察された。これらの結果から、モーショングラムを自動生成するシステムには、ピクトグラムの特徴や意味に応じて、分かりやすいオブジェクトを取り出すこと、そのオブジェクトにとってシンプルで自然な動きを推定し、生成することが求められる。

また、フィードバックの内容をみると、静止画から動画になることで「わかりやすくなった」「イメージしやすくなった」といった意見が72.5%の回答者にみられた。一方、表現されている行動が読み取れても、「依頼しているのか禁止しているのかわからなかった」といった意見もみられた。日本のピクトグラムでは、例えば災害標識であれば、その色と形で「指示」「誘導」「注意」「禁止」の分類がなされており、その対応を知っていれば理解できるように工夫がされている[9][10]。しかし、より多様な人が利用することを想定すると、前提知識はできるだけ求めずに理解を促すことが重要視される。モーシヨンの設計の際には、その指示の「分類の明確化」が重要であることが分かった。指示しようとする内容が、色や形状などにより体系的に設計されているものであるならば、オブジェクトや動きの推定とは別に、システム側で指示の内容を認識したのち、指示内容を示すモーシヨンに変換し、再びオブジェクトや推定した動きと合成する方法も考えられる。

4. まとめ

本研究では、静止画ピクトグラムからモーショングラムを自動生成することにより、現状の静止画ピクトグラムの抱える課題を解決し、より直感的で効果的な情報供給の普及へと貢献することを目指す。

AIを用いたモーショングラムの自動生成手法の研究では、事前学習済みモデルを使用したモーショングラム生成を試行した。これにより、ピクトグラムの表現に適した改善が必要であることが明確になった。求められた効果的な前処理や認識上の課題をもとに、前処理の自動化や関節の位置座標推定など、ピクトグラムに対する精度改善を検討・開発することが必要である。モーショングラムの効果を検証するため、静止画ピクトグラムとモーショングラムの効果と比較検証する実験を行った。その結果、効果の傾向と要件が示された。傾向として、行動を表すものや比較的複雑な意味を表すピクトグラムに対して効果が高く、特に、元のピクトグラムの認知度が低いものや、動作を表すものにおいて効果が顕著であった。要件としては、オブジェクトや動作の分かりやすさに加え、明示的な指示が伝わるモーショングラムの設計が挙げられる。なかでも、ピクトグラム

の指示の「分類の明確化」が重要な観点となる。

今後の計画としては、Text-to-Video (T2V)やText-and-Image-to-Video (T2I2V): など、テキスト駆動型の生成AIを使用した自動化を検討している。また、AIによる生成時と人間による手作業での作成時のモーショングラムの効果比較を行い、自動生成されたモーショングラムが発揮できる効果を調査予定である。

将来的な展望としては、認識したピクトグラムをマーカー、生成結果を表示コンテンツとしたAR化への活用が考えられる。ARによるモーショングラム化が実現できれば、現実に設置されている静止画のピクトグラムをそのままに、拡張的にモーショングラムを取り入れることが可能となる。新たなピクトグラムに対してもリアルタイムで情報の拡張、更新を行うことができる。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] 田中克己, 黒橋禎夫: 情報デザイン, 共立出版株式会社, 2018.
- [3] 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, 17–31, 2007.
- [4] 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.
- [5] upwork: How much does a Motion Graphics Designer cost?. upwork, 2023.
- [6] 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.
- [7] Jian Zhao Hui Zhang: Thin-Plate Spline Motion Model for Image Animation, School of Software, BNRist, Tsinghua University, Beijing, China, 2022.
- [8] Wen, L., et al.: Liquid Warping GAN with Attention: A Unified Framework for Human Image Synthesis. School of Software, BNRist, Tsinghua University, Beijing, China, 2022.
- [9] 国土交通省 水管理・国土保全局 防災課: 防災標識でそれぞれの自然災害が起きたときに逃げる場所を知ろう, 小4以上対対象 社会科自然災害から人々を守る活動 児童用ワークシート, 2023.
- [10] 国土交通省 水管理・国土保全局 防災課: 防災標識でそれぞれの自然災害が起きたときに逃げる場所を知ろう 教員向け手引き, 小4以上対対象 社会科自然災害から人々を守る活動, 2023.