

# 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

Development of a system to convert pictograms into motion-pictograms using AI

岡谷 夏実<sup>1)</sup>, 塩谷 隆二<sup>2)</sup>, 中林 靖<sup>3)</sup>, 多田 光利<sup>4)</sup>

Natsumi Okatani, Ryuji Shioya, Yasushi Nakabayashi and Terutoshi Tada

- 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)
- 4) 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: t\_tada@toyo.jp)

Pictograms are graphic symbols that expresses the meaning or concept of an idea by using the form of the target item. They are used for emergency exit signs, public facilities, etc., in order to inform the information in a non-verbal way. However, due to the limitations of expression through still images and differences in expression by nationality and culture, not all pictograms can be understood clearly by everyone. Therefore, a system is developed that recognizes still pictograms using AI and converts them into motion graphics.

**Key Words :** AI, Image Recognition, Pictograms, Motion Graphics

## 1. はじめに

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]. 従来は静止画により, 非常口サインや公共施設の案内表示などに使用されてきたが, 近年では, 東京五輪において「動くピクトグラム」[2]が登場したように, アニメーションにより情報を拡張, モーショングラフィックス化されたピクトグラムが用いられ始めている. 本研究では, これを「モーションピクトグラム」と称する.

ピクトグラムは, 意味伝達の目的と, 設置場所の景観を損なわない程度の存在感が求められることから, シンプルかつ安易な表現で, また, 利用者が素早く理解できるデザインが求められる. しかしながら, 表そうとする事象によっては伝達内容の複雑さや国籍・文化による表現の違いを乗り越えなければならないといった課題もあり, 静止画による表現では限界がある. 誰もが正確な意味を一見して理解できるものばかりではないのが現状である.

そこで, アニメーションによって情報補完することで, 複雑な事象や静止画では伝わりづらい動きも表現することができ, より直感的で正確な理解が可能となる. 「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することとなり, 例えば, 表示が母国語でない場合や, 利用者が子どもである場合のように, 文字での説明が理解できなくても情報を伝えられることで, 危険の回避や注意に役立つ. 国際化, 多様化の進む現代社会において, 誰もが使いやすい街づくりに貢献することができる.

## 2. 目的

ピクトグラムはJIS規格ピクトグラム (案内用図記号 JIS Z8210) [3]として策定されている104項目をはじめ, 公用, 民間合わせ数多くの種類が存在する. 国内外を合わせればさらに多種多様なピクトグラムがみられる. また, 新型コロナウイルスによる社会的な影響を受けて, ソーシャルディスタンスやマスク着用, 黙食といった数多くのピクトグラムが考案されたように, 社会の状況や必要に応じて日々新たなピクトグラムが作成, 更新されている.

これに対し映像作成の時間は静止画の作成に比べ多大かつ専門性も求められ, 時間や人的資源も要する. こうしたアニメーション作成のコストは, モーションピクトグラムが浸透しづらい理由のひとつでもある.

そこで, 静止画のピクトグラムからモーションピクトグラムを自動的に生成することができれば, モーションピクトグラムによる情報伝達を, 実現可能かつ一般に取り入れやすい手段として浸透させることができる. 言語を要しないコミュニケーション手段として, 現状の静止画ピクトグラムの抱える課題を克服し, 安定かつ広範な情報供給が可能になると考えた. 本研究では, AIを用いて静止画のピクトグラムを認識することで, モーショングラフィックス化を行うシステムを開発する.

## 3. AIと画像認識を用いたコンテンツ生成

AIと画像認識を用いて認識結果からコンテンツを生成することは, これまで, 教育分野に適用し, 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. [4]

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が活用できる。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、これに応じたシステムの構築が必要となる。

#### 4. モーションピクトグラム

最終的な成果のイメージとして、静止画ピクトグラムをもとにしたモーションピクトグラムを手動で作成した。例として、図1に静止画のピクトグラムを示す。

図1は単体かつ静止画の状態では何を示したもののかわかりにくいですが、JIS Z8210「転落注意（Caution, drop）」を示したピクトグラムである。そこで、モーションピクトグラム化を行えば、落下の動作はアニメーションによって明確に示され、転落に対する注意喚起の意義を強めることになる。作成したモーションピクトグラムについて、フレーム分割した様子を図2に示す。

#### 5. システム構成

対象のピクトグラムが表現しようとしているモチーフ、意味するものの形を、AIおよび画像認識を用いて判定し、判定結果から対象の動きを推定する。例えば、対象が人型ピクトグラムであれば「歩く」「走る」といった動作、対象が水や煙であれば物理法則に基づいた「流れる」「漂う」「落下する」といった動作、矢印による図示があればその方向に対象が移動、といったように対象の動作を推測し決定する。



図1 静止画ピクトグラム「転落注意」



図2 モーションピクトグラム「転落注意」  
(フレーム分割)

こうした推定結果をもとにモーショングラフィックスを生成し、モーションピクトグラム化を行う。特に、人型や動物型のピクトグラムでは、判定した対象に応じた実写映像や既存のアニメーションの動きを対象の静止画ピクトグラムへと転移させることで、新たなもの、未知のものであってもモーショングラフィックス化へ対応することを考える。

#### 6. 人型ピクトグラムのモーション化

図-3は「非常口」を示す静止画ピクトグラムである。この人型ピクトグラムに対し、実写映像から「走る」動作の転移を行い、新たなモーションピクトグラム生成を行った。

##### (1) Thin-Plate Spline Motion Modelの使用

モーションピクトグラム生成のため、Thin-Plate Spline Motion Model[5]を用いた動作転移を行った。動作転移には事前学習済みモデル「TaiChiHD」を使用した。生成したモーションピクトグラムについて、動作転移の様子を図-4に示す。各フレーム左下部分に動作転移に使用した実写映像、右上に対象とする静止画ピクトグラムを配置している。

図-4に示した生成結果では、人型ピクトグラムの周囲にグレーのノイズが生じた。また、右足先端部分の動作が転移せずに一点に固定され、右膝のみが動作している状態となった。これは、人型ピクトグラムと非常口の周囲の壁面を示す枠が重なっており、それらが同色で表現されていることが原因と考えられる。



図3 静止画ピクトグラム「非常口」

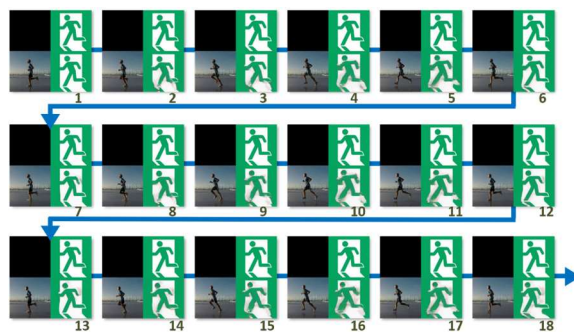


図4 モーションピクトグラム「非常口」  
(フレーム分割)

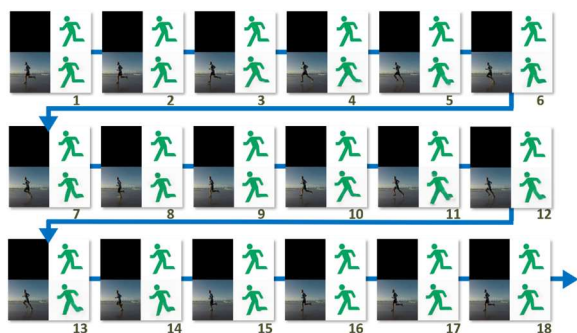


図5 モーションピクトグラム  
「非常口（背景なし）」（フレーム分割）

そこで、非常口周囲の壁面部分を除去し、人型ピクトグラムにフォーカスした状態での動作転移を試行した。動作転移の様子を図-5に示す。

図-5に示した生成結果では、図-4の生成結果で見られたようなノイズは低減することができた。しかし、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。人型ピクトグラムでは右足が地面に平行な状態まで曲げられているが、実写映像ではそこまで高く曲げられておらず、右足先端部分がやや下の位置にある。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が、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。

## (2) Liquid Warping GAN with Attentionの使用

Liquid Warping GAN with Attention[6]では、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ことが示されている。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、図-3「非常口」の静止画ピクトグラムでは、画像内に人体が認識されず、動作転移を行うことができなかった。また、背景部分を手動で削除した画像を用いての実行でも、同様に画像内に人体が認識されなかったため、使用した人体検出モデルが、デフォルメされた人型ピクトグラムに対しては適応されないことが示唆された。更に、静止画ピクトグラムの種類を変え、図-6に示される「転落注意」「更衣室」といった姿勢や色の異なる対象でも試行したものの、人体は認識されない結果となった。一方、「陸上競技場」



図6 ピクトグラム  
「転落注意」「更衣室マーク」「陸上競技場」

場」のピクトグラムでは人体の認識に成功した。しかしながら、動画との対応クロッピングおよび正面画像の生成部分が動作せず、モーションピクトグラムの生成はできなかった。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。対応する認識モデルを作成、適用させることで、人型ピクトグラム専用の3次元人体メッシュ復元の方法を検討、動作転移の実行を目指す。

## 7. まとめ

本研究では、静止画ピクトグラムからモーションピクトグラムの自動生成を行うことにより、モーションピクトグラムの利活用をより一般的で現実的な手段とすることを目指す。静止画ピクトグラムの課題である表現の限界や複雑な事象の伝わりにくさをモーションピクトグラムの活用によって解決することで、言語によらない情報伝達をより安定的で効果的なものとし、公共の場での注意喚起や情報伝達といった役割を高める。静止画のピクトグラムを入力として用い、ピクトグラムが表現する意味概念を考慮した適切な動作を自動的に生成システム開発に取り組んだ。

将来的な展望としては、駅や博物館などの公共施設に設置されている電子ディスプレイでの利用および、認識したピクトグラムをマーカ、生成結果を表示コンテンツとしたAR化への活用も期待できる。ARによるモーションピクト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ピクト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] 野波健祐: 入場行進にドラクエ 世界に魅力伝わった? 五輪開会式, 朝日新聞デジタル, 2021-7-27.
- [3] 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.
- [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] Jian Zhao Hui Zhang: Thin-Plate Spline Motion Model for Image Animation, School of Software, BNRist, Tsinghua University, Beijing, China, 2022.
- [6] 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.