

# ピクトグラムとアニメーションを用いて文章概要を表現するシステム

宇佐美 寛 小川 純矢 杉浦 俊弥 吉村 直美 濱川 礼

中京大学 情報理工学部 情報システム工学科

## 1. 概要

本論文では、ピクトグラム(図 1)とアニメーションを用いて、文章を直感的に理解するシステムについて述べる。本研究では、ピクトグラムにアニメーションを付加することで、文章が表す概要をより直感的に把握できるのではないかと考え、システムとして実装し、検証を行った。



図 1 ピクトグラム例(自作)

## 2. 背景・目的

近年、街中や身の回り等あらゆるところにピクトグラムを用いた標識やマークが普及している [1]。ピクトグラムを用いて、普段読んでいるニュースなど記事の内容を把握できないかと考え開発を行った。

## 3. 事前調査

時系列順に並んでいることが多くアニメーションにし易いことから、扱う文章をニュースの事件・事故に絞り、過去 1 年から 530 件のニュース記事 [2] について調査を行った。多くの記事に対応するため、10 回以上出現した動詞から 45 種類の動作を定義し、210 種類の動詞に関連付けた。また、10 回以上出現した名詞から 354 種類の名詞を JIS 規格ピクトグラムの一部と 117 個の自作ピクトグラムへ関連付けた。これは名詞の出現回数 40430 回中の 29159 回(72.1%)、動詞の出現回数 4469 回中の 3008 回(72.4%)にあたる(表 1)。

表 1 事前調査による頻出単語ベスト 5

順位	出現回数	動詞	順位	出現回数	名詞
1	158	調べる	1	633	人
2	78	出す	2	287	署
3	75	話す	3	262	男性
4	69	見つかる	4	252	者
5	61	負う	5	228	事故

## 4. 提案手法

図 2 に実装したシステムでの処理手順を示す。文章をテキストファイルから読み込み、形態素解析を行う。形態素解析には形態素解析エンジン「MeCab」 [3] を用いている。以降の説明について以下に説明する。

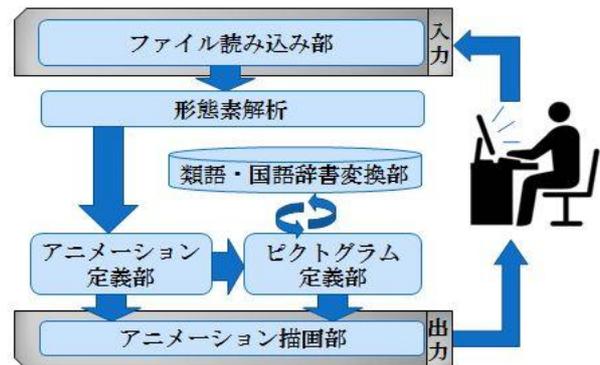


図 2 処理の手順

### ▶ アニメーション定義部

アニメーション定義を行うために形態素解析データからオブジェクト群を作成する。オブジェクトとは形態素に品詞などを付加した独自のデータ形式で、[オブジェクト名, 品詞情報, オブジェクト情報, 使用フラグ]で構成される(表 2)。

表 2 オブジェクト例

変数	値
オブジェクト名	神奈川県厚木市
品詞情報	名詞
オブジェクト情報	地域
使用フラグ	false

形態素が名詞でない場合はオブジェクト名と品詞情報に、形態素解析データの原型と品詞を当てはめ、オブジェクト情報と使用フラグは利用されない。

形態素が名詞である場合は連続した名詞を結合しオブジェクト名とし、品詞情報には名詞を当てはめる。オブジェクト情報には形態素解析データの品詞詳細を当てはめる。品詞詳細が地域、日付、時刻ではない場合にはオブジェクト名をピクトグラム定義部へ渡し、分類を取得する。

使用フラグはアニメーション定義された時に false から true に変わる。

オブジェクト群が作成されたらアニメーション定義のため rad ファイルを作成する. rad ファイルはアニメーション描画部で利用されるデータの一覧であり, [オブジェクト名, オブジェクト名の原型]と, 45 種類の予め定義された動作パターン[描画開始タイミング, 描画時間, 描画位置, 透過度, 回転角度, 反転フラグ, 画像分割フラグ]からオブジェクトに対応するものが記述される.

オブジェクト群の中でオブジェクト情報を持っているものに対応する動作定義を rad ファイルへ書き出す. また, 動詞からさかのぼって名詞が出てきた場合に動詞と名詞から動作定義を rad ファイルへ書き出す. さらに, ここまで何もされていない名詞でピクトグラム定義部から分類が得られたものの動作定義を rad ファイルへ書き出す.

#### ▶ ピクトグラム定義部

アニメーション定義部から受け取った名詞(以下, 元単語)に対応するピクトグラムの画像ファイル名を独自データベースから探し関連付けを行う. このデータベースは[単語 画像ファイル名分類]で構成され, JIS 規格ピクトグラム 313 種と自作ピクトグラム 117 種の画像ファイル名に対して 667 語の単語情報, それに対する単語分類が記述されている.

関連付けには全文一致検索, 部分一致検索, 類語検索, 国語辞書検索を利用する. 全文一致検索では元単語を基にデータベースの検索を行う. 元単語が見つかった場合, アニメーション定義部へ分類を返し, 画像ファイル名を外部ファイル(以下, orz ファイル)へ書き出す. orz ファイルは[元単語, 画像ファイル名]の組み合わせで構成され, アニメーション描画部で利用される.

#### ▶ 類語・国語辞書変換部

類語辞書検索では元単語に対応する類語を検索し, 複数候補が該当した場合, 複数単語を返す.

国語辞書検索では元単語の意味を検索し, 解説文の最後の単語を返す. 単語 1000 件について調査した結果, 本研究で利用している辞書は, 解説文の最後の単語が本質を示しているものが 606 件あったため, この手法を用いた.

本研究の類語辞書・国語辞書には「日本語 WordNet」 [4] を利用している.

#### ▶ アニメーション描画部

orz, rad 形式の 2 つのファイルを利用し描画を行う. orz ファイルから元単語を取得し, その関連データとして画像を読み込む. 次に, rad ファイルからデータを読み込み時間軸上に登録する. 加えて, rad ファイルの開始タイミングに実行時間を加えて最も長いものを全体時間として登録し,

クバーとの対応を行っている. アニメーションの描画にはダブルバッファリングを用いた. 描画は(図 3)のように時間軸に従って進んでいく.

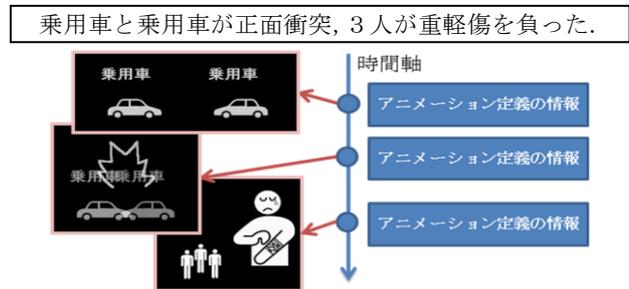


図 3 アニメーション遷移

時間軸上にアニメーション定義が登録されていたとき, その登録情報の元単語と一致する画像の描画フラグを true にし, アニメーション定義を情報として, 追加登録する. この追加登録を行った情報は設定されている実行時間の間だけ実行され, 実行時間が 0 となった際に描画フラグは false にされ描画されなくなる. 時間軸は 1 コマあたり 30 分の 1 秒に設定されている.

### 5. 評価

濱川研究室学生 28 名に対し, 事件・事故を扱った記事 5 件, 評価者が自由に選んだ記事 5 件について, 評価を行ってもらった. ストレスなくアニメーション変換処理が行われたとする意見が 8 割を超え, 対応する単語やアニメーションが充実し, 表現の幅が広がれば本システムを利用したいといった意見が 7 割あった.

### 6. 考察

上記の評価より, より分かりやすく文章概要を表現するために, 対応している単語やアニメーションを増やす事が重要であるとわかる. また, 事前調査で頻出単語に対応したにもかかわらずうまくアニメーションへ変換できなかった理由として, 時事的な単語が多かったことがあげられる.

### 7. 参考文献

- [1] 財団法人 共用品推進機構  
<http://www.kyoyohin.org/index.php>
- [2] MSN 産経ニュース  
<http://sankei.jp.msn.com/>
- [3] 形態素解析エンジン「MeCab」  
<http://mecab.sourceforge.net/>
- [4] 日本語 WordNet  
<http://nlpwww.nict.go.jp/wn-ja/>