

マルチモーダルノベライズシステムの提案

吉田尚平^{†1}, 村田嘉利^{†1}, 鈴木彰真^{†1}

スマートフォンやタブレット端末、電子書籍リーダの普及により、電子書籍が一般生活に浸透してきた。特に、小説や漫画など、手軽に楽しめるジャンルは電子書籍化されやすく、数多くリリースされている。小説の電子書籍のほとんどは、紙媒体の本と同様のテキストから、端末に合わせたフォーマットに変換されて利用されている。これらは、動画や音声を埋め込むことができるため、電子書籍ならではの表現が可能であるが、実際は既存の紙媒体の本を再現しているだけで、スマートフォンやタブレット端末が持つマルチモーダルなインタフェースを活用できていない。本論文では、顔の表情をもとにした喜怒哀楽の感情入力と、背景色の変化による感情表現といったマルチモーダル処理により、感情が正確に伝わり、より多様な表現が可能なるノベライズシステムを提案した。また、実際にシステムを構築し、アンケートによって提案手法の有用性を評価した。

Proposition of Multi-modal Novelize System

SHOHEI YOSHIDA^{†1}, YOSHITOSHI MURATA^{†1}, AKIMASA SUZUKI^{†1}

1. はじめに

スマートフォンやタブレット端末、電子書籍リーダの普及により、電子書籍が一般生活に浸透してきた。特に、小説や漫画など、手軽に楽しめるジャンルは電子書籍化されやすく、数多くリリースされている。

小説の電子書籍のほとんどは、紙媒体の本と同様のテキストから、端末に合わせてEPUB (Electronic PUblication)[1] やXMDF (ever-eXtending Mobile Document Format)[2] などのフォーマットに変換されて利用されている。これらは、動画や音声を埋め込むことができるため、電子書籍ならではの表現が可能であるが、実際は既存の紙媒体の本を再現しているだけで、スマートフォンやタブレット端末が持つマルチモーダルなインタフェースを活用できていない。

その一方、プレーンなテキストの文章を読者が読む場合、ある文字列が持つ意味として、喜怒哀楽がはっきり表現しきれない場合があり、作者の意図したものとは異なる感情が誤って伝わってしまうことがある。

そこで本研究では、マルチモーダル処理により、感情がより正確に伝わり、多様な表現が可能なるノベライズシステムを提案する。さらに、構築したアプリケーションを用いたアンケートを行うことで、その有用性の評価する。

2. 関連研究

メールやSMSなどの電子的なコミュニケーションシステムでは、プレーンなテキストで文章を送る他に、様々な表現を用いる。例えば、絵文字や顔文字やエモーティコンなどの顔を用いた表現や、文字のサイズや色を変えるなどの

文字を利用した表現、またキャラクターが動いたり装飾が輝いたりするようなアニメーションを用いた表現がある。

しかし、これらの表現は、紙媒体の小説においてはあまり利用されておらず、それを変換した電子書籍でも同様である。また、このようなシステムを用いたコミュニケーションでは、必ずしも書き手の感情が正確に伝わっているとは限らない。

藤原ら[3]は、BBS (Bulletin Board System) の利用において、書き込みを行っているときのタイピングから、書き込みスピード・入力誤りの頻度・顔文字・擬態語の挿入状況を検出することにより、書き手の感情を自動的に判断した上で、その感情に合わせて掲示板に書き込まれる内容に装飾を行うシステムを提案している。書き込んだ内容の文字について、背景の色・文字の色・文字の大きさを変更し、感情を表現している。また、掲示板のトップにある風景の写真が、掲示板全体の雰囲気によって、天候が変わる表現をしている。

また、箕田ら[4]は、メールでのコミュニケーションについて、色のみを送ることによる感情表現の試みを提案している。この中で、人間の基本的な感情は、色と感情の一致度が高いということから、喜怒哀楽の感情を、色のみでメールで伝えることができることを示している。

しかし、藤原らのシステムは必ずしも書き手の感情が正しく入力されるとは限らない。箕田らのシステムは、色以外の表現方法についての検証を行っていない。さらに、電子書籍の小説において、装飾による感情表現が行われている例は見当たらない。

^{†1} 岩手県立大学
Iwate Prefectural University, 152-52, Sugo, Iwate, 020-0693, Japan

3. マルチモーダルノベライズシステムの提案

3.1 システムの概要

以上の現状を踏まえて、マルチモーダル処理により、感情がより正確に伝わり、多様な表現が可能なノベライズシステムを提案する。提案手法では、文章に感情をマルチモーダルに付加する方法として、XML(Extensible Markup Language)の独自タグを用意し、パラメータに感情の値を入力したXML文書を電子書籍のフォーマットとして用いる。また、感情をマルチモーダルに表現するシステムとして、リーダとエディタによるアプリケーションを提案する。作者に対しては、文章に対して感情を付加できるエディタアプリケーションを提案し、表情を用いた入力を可能とする。また、読者に対しては、感情が付加された文章を色の変化やアニメーションなどによって、その文字に付加された感情を分かりやすく表現するリーダアプリケーションを提供する。システム構成図を図1に示す。

以下、それぞれのアプリケーションにおけるシステム概要について説明する。

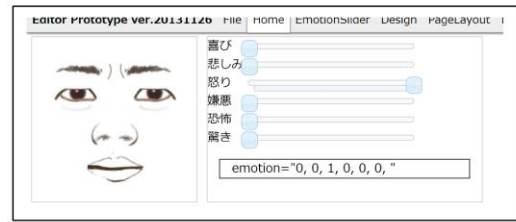
3.2 感情を付加できるエディタアプリケーションの概要

小説等の作者が、文章に対して感情を入力し表現するためには、作者が意図した感情を数値化して入力する必要がある。基準がなければ感情値を決めることは難しい。顔の表情は感情を最も効率的に発する[5]ことから、Paul EkmanらによるFacial Action Coding System[6] (以下、FACSと記す)を用いて作成された表情を見ながら感情値を入力することにより、感情を文章に与えるための補助手段にする。

Ekmanらは、「喜び、驚き、恐怖、嫌悪、怒り、悲しみ」という6つの表情が汎文化的であると考えている[7]。そこで、その6つの表情を6つの感情に変換し、独自タグに与えるパラメータとする。

エディタを利用する作者は、FACSを用いて作成された表情を基準としながら、感情の値を文章に与えることができる。表情の生成には、FACSで定義されたAU (Action Unit)の動きを用いる。AUとは、顔の領域の動作部分や度合いを定義したものであり、エディタではその組み合わせによって表情を作成する。Ekmanらによる6つの表情は、表1のようなAUの組み合わせで表現することができる。そこで、各表情について、最小(AUが動作していない)を0、最大(AUが動作した)を1としてパラメータ設定をできるようにし、その値から表1で示す各顔の部位のAUを同時に変更する。

また、このエディタアプリケーションでは、感情の値を得ることが重要であるため、表情をわかりやすくするために、AUの動作を強調する。どのAUを強調するかについては、各感情の表情を表示したときに感情を認識できるかを確かめるアンケートを行い、わかりにくかった感情のAUについて強調して調整をする。また、2つ以上の感情



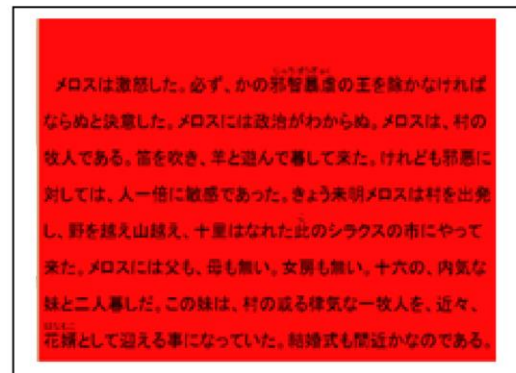
エディタアプリケーション

①文章に対して感情を入力

```
<emotion anger="1">メロスは激怒した</emotion>
```

感情タグ

②入力された感情を表現



リーダアプリケーション

図1 ノベライズシステムの構成

の合成も可能にする。たとえば、恐怖を0.7、嫌悪を0.5のようにすると、何か不気味なものが近づいてきている、というような表情を生成することができる。

一方、表1によれば動作させるAUが競合する場合がある。たとえば、驚きと恐怖の表情では、口を開けるというAUが両方に存在している。そこで、AUの競合が起こった場合は、AUの動作におけるパラメータの最大値が1となるよう、競合する表情の各動作割合から計算しなおす必要がある。これを解決するために、競合する2つ以上の数値を合計したものと1を比べて、小さい方のパラメータをAUの程度として出力することとする。

3.3 付加された感情を分かりやすく表現するリーダアプリケーションの概要

入力された感情をマルチモーダルに表現する方法として、背景や文字の色の変化、文字の大きさの変化、エフェクトやアニメーション、バイブレーション、サウンド再生、表情の表示などが考えられる。

また、表現が適用されるタイミングは、画面にその文字が表示されたとき、目でその文字を捉えたときが考えられる。また、読書のスピードを計測し、現在読んでいる位置を推測することも考えられる。これらは、画面に対するタッチ操作の位置取得[8]や、インカメラからの情報[9]を活用することで実現可能である。また、読んでいる間に徐々に

表 1 表情と AU の組み合わせ

表情	AU	
喜び	頬	上げる
	口角	釣り上げる
驚き	両眉	上げる
	まぶた	見開く
	口	開ける
恐怖	両眉	上げて寄せる
	目	見開く
	口角	横に引く
	口	開ける
悲しみ	眉内側	上げて寄せる
	鼻唇溝	深める
	口角	下げる
	頭部	前に傾ける
	視線	下に向ける
怒り	両眉	中央寄せる
	まぶた	力をいれる
	目	見開く
	唇	硬くする
嫌悪	鼻	皺寄せする
	下唇	押し下げる
	唇	開く
	口	開ける

変化を起こすこともできる。本システムでは、これらの技術が利用可能な状況を前提とした上で、本を読んでいる位置における文字と背景の色を変化させ、感情が付加された文字が画面に表示されたときに起こる表現の付与する試行を行う。

Schaie は、色彩の情緒的意味についてまとめている[10]。その中から、本システムで用いる 6 つの感情に対応する色の選択を行った。リーダアプリケーションでは、表 2 に示す表情との関係を用いて、リーダアプリケーションにおける背景の色を変化させる。複数の感情値が同じ文字列に対して付与されていた場合は、それぞれの感情が与える色の RGB 値を各要素で平均した結果を最終的に出力する色とする。

また、文字色については、背景色として最終的に出力する 8bit の RGB 値の各要素の平均が 255 / 2 よりも大きい場合は黒色、それ以下の場合は白色を出力する。

表 2 リーダアプリケーションで用いる感情と色の関係

感情	色	RGB
驚き	橙色	(255, 106, 0)
恐怖	黒色	(0, 0, 0)
嫌悪	紫色	(170, 0, 204)
怒り	赤色	(255, 0, 0)
悲しみ	青色	(0, 12, 153)
喜び	黄色	(238, 255, 51)

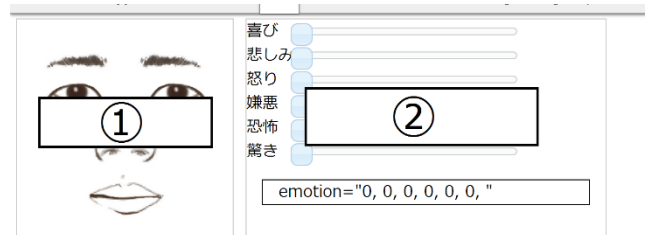


図 2 感情の値を入力するインターフェース

4. システムの実装

以上の検討内容をもとにして、アプリケーションの実装を行った。以下、エディタアプリケーションとリーダアプリケーションについて、順に説明する。

4.1 エディタアプリケーション

4.1.1 感情を表現するスライダを用いた XML タグの作成

作者が文章/感情の入力を行うエディタアプリケーションを、HTML5 と JavaScript を利用して作成した。表情を表示するために、HTML5 の機能のひとつである canvas 要素を用いる。canvas 要素は、動的に 2 次元ビットマップ画像を描画するための HTML 要素である[11]。アプリケーションの感情入力ウィンドウのインターフェースを図 2 で示す。図 2 右側② の領域には、6 つの感情「喜び、驚き、恐怖、嫌悪、怒り、悲しみ」の値を設定するスライダを用意する。スライダが変化した時点左側①の canvas 領域に顔の表情が生成される。顔を領域ごとに分け、画像として canvas 領域に配置し、拡大/縮小、歪み、回転によって表情を作成する。それぞれの感情の値は 0 以上 1 以下の数値で出力され、XML タグを作成する。

4.1.2 顔の表情を表す CG の実装

最初に、図 3 のような顔の CG を、ペイントソフトを用いて作成した。



図 3 作成した CG

この CG は、以下の複数の要素が階層構造によって表現されている。

- ① 眉
- ② まぶた
- ③ 目の縁
- ④ 眼球
- ⑤ 上唇
- ⑥ 下唇
- ⑦ 口の中

図 2 の②の領域にあるようなスライダを操作し、その値を用いて①に描画されている要素を動かすことで、3.2 で説明したような AU を実現する。

このスライダは、左に表示されているそれぞれの感情について、0 以上 1 以下の数値を入力するために用いる。スライダが変化するたびに、canvas 領域に配置されているレイヤを動作させ、AU を再現する。図 3 は、実際に図 2 の②領域にあるスライダを実際に変化させ、それぞれの感情を 1 としたときに生成された表情である。図 4 で示したような表情は、一つの感情のみが反映されている。今回は、まず基本的な 6 つの感情について、その感情一つひとつが表情に反映されているかを確認するために、8 人に対してアンケートを行った。アンケートの方式は、ランダムに並び替えて表示された 6 つの表情の CG を見て、6 つ感情の選択肢中から選択するという方式で行った。

アンケートの結果を図 5 に示す。図 5 は、図 4 の各表情の CG に対してアンケートによって選択された表情が意図した表情と合致するかどうかを認識率として表している。図 5 を見ると、恐怖、嫌悪、怒りについては、表情を認識できた割合が 50% を下回っており、改良を行う必要があるということが確認できた。また、その他の意見として、人間の写真から表情を生成したほうが認識しやすくなるのではないか、アニメーション作品やマンガのような表現を加えると良い、AU の動きを強調することで分かりやすくなる、というものが挙げられた。

4.1.3 写真から作成した CG を用いた表情の実装

以上の結果に基づいて、写真をもとに、顔 CG を、ペイントソフトを用いて作成した。より表情を分かりやすくするために、眉間の皺、鼻、ほうれい線を表情で変化させることにした。そのため、要素の追加を行った。追加後の要素の構成は以下の通りである。

- ① 眉
- ② 眉間の皺
- ③ まぶた
- ④ 目の縁
- ⑤ 眼球
- ⑥ 鼻
- ⑦ ほうれい線
- ⑧ 上唇

- ⑨ 下唇
- ⑩ 口の中

これらの要素を canvas 要素に表示した結果を図 6 で示す。左は CG を作成するために撮影した顔の写真である。これをペイントソフトでなぞることによって顔の要素を作成した。

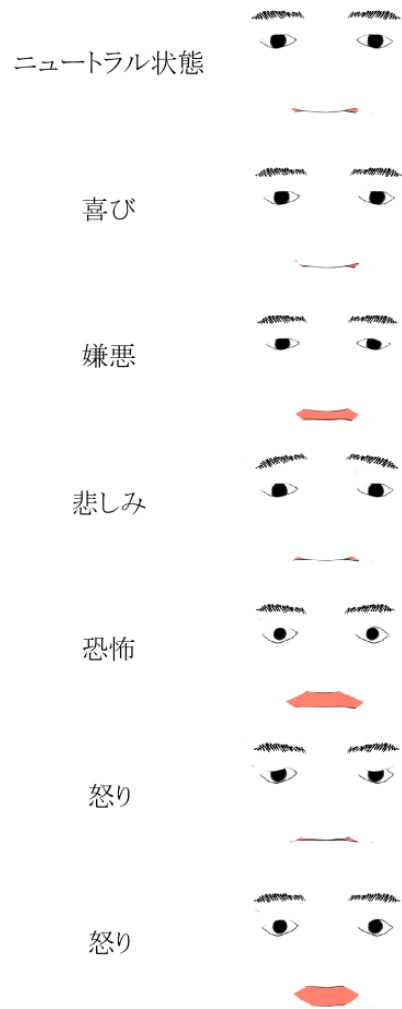


図 4 それぞれの感情を 1 にしたときに生成された顔の CG

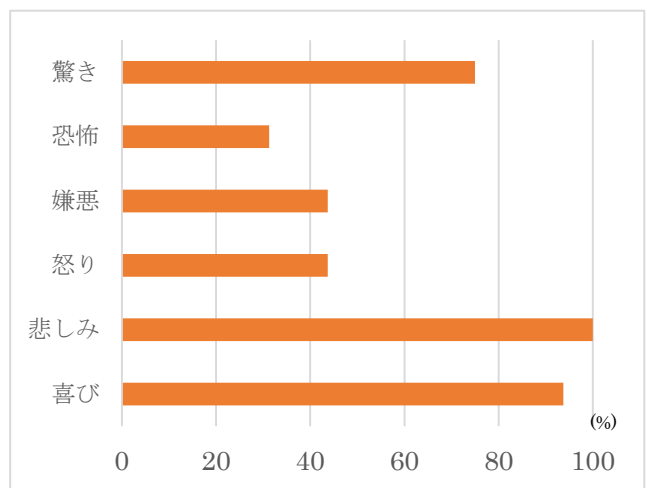


図 5 各感情における表情認識率

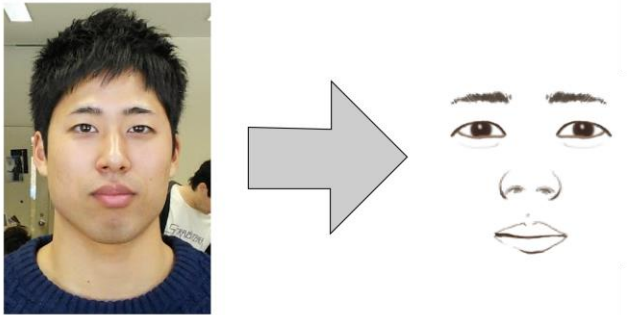


図 6 写真を基に作成した顔 CG

AU の動作についても変更を行った。FACS にもとづいて動作させる AU の他に、各感情の表情の特徴を表すことのできる AU については以下に示す 6 つの強調を行った。

- ① 喜びについては、口角を上げるという動作だけでなく、口を少し開き、それに伴ってほうれい線が出現するようにした。また、目を細める調整を行い、頬が上がる表現を追加した。
- ② 嫌悪については、とくに変更した部分はなく、4.1.1 で示された AU と同じ動作をするようにした。
- ③ 悲しみについては、眉内側を大きく上げるように変更した。
- ④ 恐怖については、アニメや漫画で用いられるような表現を使い、感情表現を強調した。額から鼻にかけて、血の気が引いたような影を表示した。また、右頬に汗を表示させることによって、冷や汗をかいているような表現を行った。
- ⑤ 怒りについては、両眉を寄せた時に発生する眉間の皺を表示させた。
- ⑥ 驚きについては、口の開け方と目の開きを大きくした。

図 7 は、以上の修正を行い、実際にスライダを変化させ、それぞれの感情を 1 としたときに生成された表情を示している。

4.2 リーダアプリケーションの実装

感情表現が入力された XML 文書を読むためのリーダーアプリケーションを、HTML と JavaScript を利用して作成した。本アプリケーションは、文章を表示する他、XML 文書に入力された感情タグを解析し、感情に合わせて画面の装飾を行う。画面の装飾を行うために、今回は読者が読んでいる文章位置の推定として画面のスクロール位置を用いた。感情タグが画面の中央に近づくにつれて、感情に対応する色が背景に強く出力される。図 8 は、走れメロスの冒頭の”メロスは激怒した”という文字列に対して、怒りの感情を 1 としたときの画面の出力結果である。

ニュートラル状態

喜び

嫌悪

悲しみ

恐怖

怒り

驚き



図 7 それぞれの感情を 1 にしたとき生成された顔の CG

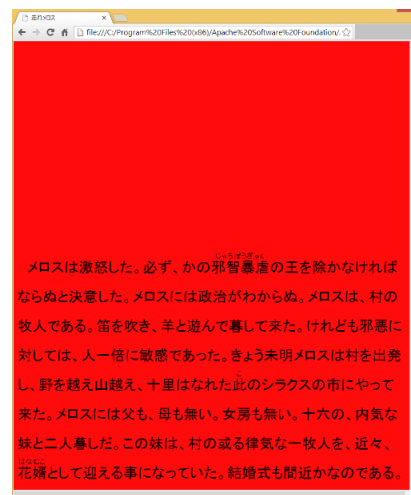


図 8 リーダアプリケーションの画面

5. アプリケーションの有用性考察

4章では、エディタアプリケーションとリーダアプリケーションの実装し、6つの各感情についてのアンケートをとった。その結果から、それぞれのアプリケーションについての有用性について考察する。

5.1 エディタアプリケーションの有用性考察

改良した表情のCGを用いて、4.1.2と同様のアンケートを、20人に対して行った。その結果を図9に示す。

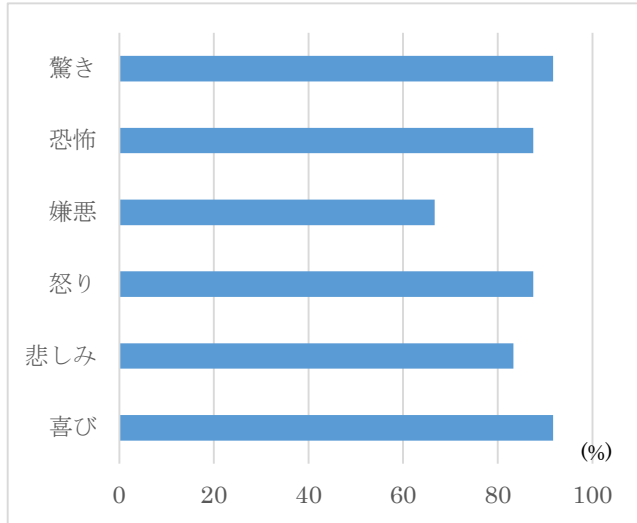


図9 各感情における表情認識率

図9は図5と同様、横軸に表情認識率を示している。図9を見ると、図5に示した認識率よりは全体的に高くなっていることがわかる。また、他の似ている表情の認識率の向上に伴い、悲しみや喜びの認識率が図5と比較して低下しているのは、他の似ている表情の認識率の向上に伴うものと考えられる。6つの感情(驚き、恐怖、嫌悪、怒り、悲しみ、喜び)の表情の平均をとると、8割を超える。そのため、6つの基本的な表情は、文章に感情値を与えるための補助としては、十分使えるシステムと考えられる。

嫌悪の表情については、感情認識率が8割を下回っており、改善する必要がある。4.1.3で行ったアンケートにおけるその他の意見として、嫌悪の表情は目を細めたほうがよい、悲しみを強調させたほうがよいというものがあった。今後はこの意見をもとに、さらに修正を行い、認識率の向上を目指す。

5.2 リーダアプリケーションの有用性考察

色を変更することで文章に与えられた感情がよりよく伝わるかの確認を行うアンケートを15人に対して行った。各感情が最大だった場合について、白い背景に黒い文字で表示されたものと、背景と文字の色が変化したものを比べて、より感情が伝わる表現であると感じた場合はチェックをする方式を採用した。図10にその結果を示す。図10は各感情における背景色の出力結果に対して、意図する感情が得られたかどうかの割合を表情認識率として表している。

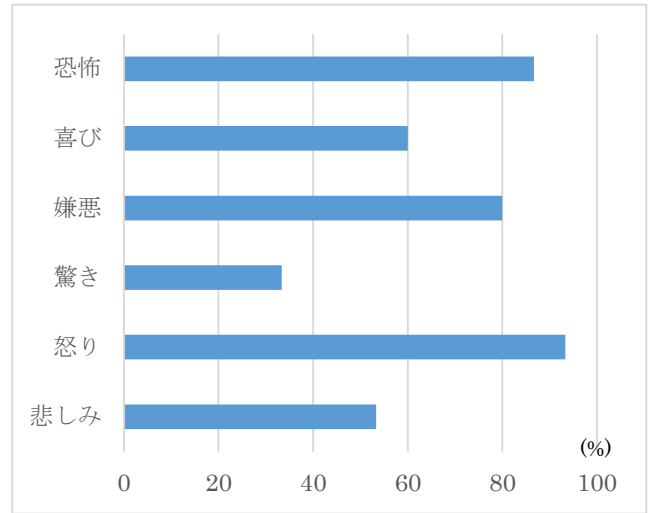


図10 各感情における背景色変化時の感情認知度

図10を見ると、恐怖、嫌悪、怒りの感情については、感情がより伝わりやすくなったと評価できる。しかし、喜び、驚き、悲しみについては、感情がうまく伝わっていない。アンケートの結果をもとに、色についての修正を行う必要があると考える。

喜び、驚き、悲しみの感情について、より正確に感情を伝えるために、色の変更を行う必要があると考えている。4.2のアンケートにおけるその他の意見として、それぞれの感情に対応する色の意見を得ることができた。例えば、喜びについては桃色のほうがよい、怒りについては、さらに濃い(暗い)色にしたほうがよいなどである。今後は、これらの意見をもとに色の変更を行い、感情の認知度の向上を目指す。

また、本論文ではスクロール位置によって感情が与えられた文章の感情を背景色と文字色の変化によって表現した。しかし、他にもタッチ操作やインカメラからの情報を用いた入力によって感情表現ができると考えられる。そこで、他の入力方法についても実装を行い、有用性を検証していく。

さらに、表現方法について、色の変化だけでなく、アニメーションや音、振動などの表現も考えられる。例えば、悲しみの感情が与えられた場合、背景が暗くなり、雨が降り始めるなどといった表現も可能である。今後はそれらについても実装を行い、有用性の検証を行う。

6. おわりに

本論文では、マルチモーダル処理によって感情の入力と出力を行うことができるノベライズシステムを提案した。

提案手法では、Ekmanによる人類普遍の6つの基本表情「喜び、驚き、恐怖、嫌悪、怒り、悲しみ」を感情として利用した。

エディタアプリケーションは、作者がマルチモーダル処理によって文章に対して感情を与えることができる。感情

を与えるために、入力する感情に応じて表情が変化する表情プレビューを作成した。このプレビューが正確に入力したい感情を表しているかを確認するために、評価実験を行い、6つの感情それぞれについて、表情プレビューが有用であることが分かった。

リーダアプリケーションは、感情が入力された文章を読み込み、その感情に応じてマルチモーダルに感情表現を行う。本論文では、背景色と文字色の変化を用いて感情表現を行った。画面のスクロール位置を用いて、感情が入力された文章が画面中央に近づくにつれて、感情の表現を強く出力する実装を行った。色の変化によって入力された感情が正確に伝わるかを確認する評価実験を行った結果、恐怖、嫌悪、怒りの感情については、感情がより伝わりやすくなったと評価できた。しかし、喜び、驚き、嫌悪の感情について、色の変更を行う必要があることが分かった。

以上の検討から、構築したアプリケーションに対して、システムの有用性を評価することができた。今後は複数の感情が合成された場合のアプリケーションの改良と、システムの有用性を評価する。

参考文献

- 1) EPUB | International Digital Publishing Forum, 入手先<<http://idpf.org/epub/>> (2014.05.14)
- 2) XMDF 情報スクエア, 入手先<<http://www.xmdf.jp/>> (2014.05.14)
- 3) 藤原光照, 山根信二, 村山優子: 書き手の感情をグラフィカルに表現する BBS の構築, インタラクション 2004 予稿集, pp.239-240 (2004)
- 4) 箕田妃希, 長幾朗: 「空メール」 - 色彩メールによる感情表現の試み - インタラクション 2003 予稿集, pp.65-66 (2003)
- 5) 中村真: コミュニケーションにおける表情と感情判断, 人文社会科学研究所年報, No.5, pp.85-91 (2007)
- 6) Ekman, P., & Friesen, W. V.: Facial Action coding. Consulting Psychologists Press. (1978)
- 7) Ekman, P., Friesen, W.V.: Constants across cultures in the face and emotion, Journal of Personality and Social Psychology 17, pp.124-129 (1971).
- 8) jQuery API Documentation, 入手先<<http://api.jquery.com/>> (2014.05.15)
- 9) 今淵貴志, 菊池輝, Prima Oky Dicky Ardiansyah, 伊藤久祥: 非接触型視線追尾システムに向けた虹彩検出手法の検証, 情報処理学会第 76 回全国大会, 6ZA-8, pp.167-168(2016)
- 10) Schaie, K. W.,: On the relation of color and personality, Journal of projective techniques & Personality assessment, Vol.4, No.6, pp.512-524
- 11) HTML Canvas 2D Context, 入手先<<http://www.w3.org/TR/2010/WD-2dcontext-20100624/>> (2014.05.15)