

LD\_003

## 記事の印象を伝えるニュースリーダー wEE の開発と評価

The wEE: News Reader Expressing Emotional Aspects of News Articles

熊本 忠彦†  
Tadahiko Kumamoto灘本 明代†  
Akiyo Nadamoto田中 克己††  
Katsumi Tanaka

## 1. まえがき

近年、日常生活の一部として、Web コンテンツを閲覧する機会が増えつつあるが、一人のユーザがインターネットにアクセスできる時間と場所は限られているため、実際に閲覧できる Web コンテンツの量はさほど多くない。この問題を解決するための一手段として、Web コンテンツをテレビ番組風の動画コンテンツに変換することにより、家事、入浴、食事といった日常的な作業を行いながら、あるいは車の中、公園、トイレといった様々な場所での閲覧を可能にする受動的視聴方式 [1, 2, 3] が注目を浴びている。受動的視聴という概念は、FRIEND21 プロジェクトの Nonogaki ら [1] により TV 番組メタファとして提唱され、その後、ネットニュースの議論から脚本を作成し、テレビ番組風に提示する方式 [2] や Web ページ中のテキストと画像の同期化領域を決定しながらテレビ番組風の情報番組に変換する方式 [3] などが提案されている。しかしながら、いずれの方式も静的データから時系列データへのメディア変換を行っているにすぎず、演出は常に同じであった。その結果、元の Web コンテンツの有する印象が視聴者にうまく伝わらないという問題が生じていた。

本論文では、Web コンテンツの有する印象を演出により伝えるニュースリーダー wEE (News Reader with Emotional Expressions) を提案する。wEE は、Web ニュース記事を CG 形式の動画ニュースに変換する際に、記事から抽出される印象に応じて BGM や CG キャラクター (ニュースキャスター役) の声色を選択するという点に特徴があり、(1) Web 上のニュースサイトから記事を収集する、(2) 記事の見出しと本文から印象 (明るい⇔暗い) を抽出する、(3) 抽出された印象に応じて BGM と声色を選択する、(4) 書き言葉を話し言葉に変換する、(5) NHK 放送技術研究所の番組記述言語 TVML (TV program Making Language) [4] を用いて動画ニュースを生成する、という処理を行う。

以下、2. で wEE のシステム構成と処理概要、演出方式について述べ、3. で被験者実験 (600 人) を通して、演出 (BGM や声色の選択) が視聴者の受ける印象に与える影響を分析し、wEE の有効性を検証する。最後に 4. で本論文のまとめと今後の課題について述べる。

## 2. ニュースリーダー wEE

wEE により自動生成されたニュース番組のスナップショット\*を図1に示し、wEE のシステム構成を図2に示す。wEE は、(1) ユーザの指定したニュースサイトから記事を収集し、(2) 収集された記事の各々から見出しと本文、あれば画像を抽出する。(3) 画像はサイズを調



図1: wEE が生成したニュース番組の例

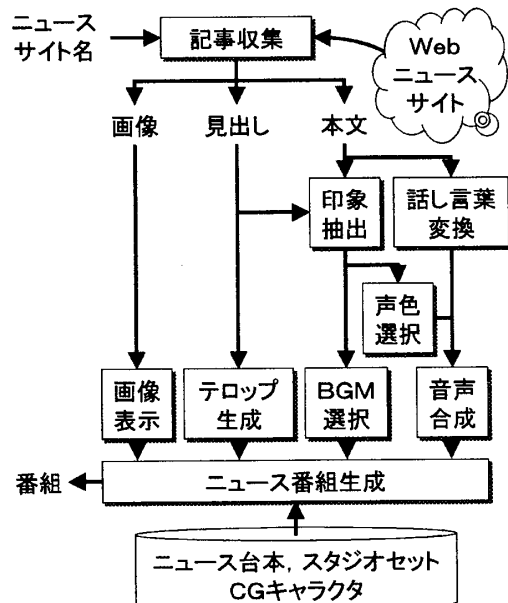


図2: wEE のシステム構成

整の上、定位置 (図1参照) に表示し、(4) 見出しは文字数に応じて、1段もしくは2段のテロップとして表示する。(5) 見出しと本文 (第1段落のみ) から、その記事の印象値 (印象尺度「明るい⇔暗い」と対応付けられた 0~1 の実数値) を算出し、その値に応じて BGM と声色を選択する。(6) 本文は書き言葉で書かれているので、自然な音声読み上げのために話し言葉に変換する。具体的には、文末および節末 (接続助詞「が」が用いられている場合のみ) のデアル調をデスマス調に変換するとともに、記事特有の略記表現 (例えば「適時打」や「W杯」) を読み上げ表現 (先の例では「タイ

† 独立行政法人情報通信研究機構・知識創成コミュニケーション研究センター・自然言語グループ

†† 京都大学大学院・情報学専攻・社会情報学専攻

\* 著作権上の問題から、wEE が自動生成したニュース番組の中の画像とテロップを著者らが別途用意したデータと置き換えた。

ムリーヒット」や「ワールドカップ」)に変換する。(7)市販の感情音声合成ソフトウェア[5]を用いて、読み上げ音声を合成する。(8)あらかじめ用意されているスタジオセットやCGキャラクターに関する情報と処理(3)から(7)において決定された画像、テロップ、BGM、読み上げ音声に関する情報をTVML形式のニュース台本に埋め込み、1本のニュース番組を生成する。(9)このニュース番組をユーザのTVMLプレーヤー[4]<sup>†</sup>上で再生する。

以下では、wEEの主要部である印象抽出に基づく演出(BGM・声色の選択)と話し言葉変換に焦点を当てる。

2.1 印象抽出に基づく演出(BGM・声色の選択)

まず、記事のテキスト部分<sup>‡</sup>から印象を抽出するための手法を示す。記事の印象は、印象尺度「明るい⇔暗い」における印象値(0~1の実数値)として算出され、テキスト部分から抽出される内容語(名詞(形式名詞、副詞的名詞、数詞を除く)、動詞、形容詞、カタカナ、副詞)の印象値と重みを印象辞書(単語と印象尺度の対応関係を示す辞書)から取得し、式(1)に代入することによって求められる。

印象辞書は、文献[6]の手法を用いて、日経新聞全文記事データベース(1990~2001年版の200万強の記事)[7]から自動構築された。文献[6]では、印象尺度を構成する印象語は1語に限られていたが、これを複数語に拡張し、ある単語 $w$ が2つの印象語群のどちらとより共起しやすいかを定式化した。この共起のしやすさを印象の強さあるいは程度と捉え、ある単語 $w$ が印象尺度 $e_1 \leftrightarrow e_2$ の左側の印象語群と共起しやすいとき、印象値 $S_{e_1 \leftrightarrow e_2}(w)$ は1に近い値をとり、右側の印象語群と共起しやすいとき、0に近い値をとるように設計された。一方、内容語の中には、出現年度数や出現回数が多いものもあれば、少ないものもある。そこで、印象値 $S_{e_1 \leftrightarrow e_2}(w)$ に対する重み $M_{e_1 \leftrightarrow e_2}(w)$ を内容語 $w$ と印象語群 $e_1, e_2$ とが共起した年度数と回数の総和(12年間分)に応じて、増減するように設計した。ここで、印象辞書の一部を表1に示し、印象尺度「明るい⇔暗い」を構成する印象語群を表2に示す。

次に、テキストの印象値を算出する手法について述べる。テキスト $TEXT$ が入力されたら、まず、日本語形態素解析システムJuman[8]を用いて、形態素解析し、 $TEXT$ に含まれる内容語を抽出する。そして、印象辞書から各単語の印象値 $S_{e_1 \leftrightarrow e_2}(w)$ と重み $M_{e_1 \leftrightarrow e_2}(w)$ を取得し、以下の式を用いて $TEXT$ の印象値 $O_{e_1 \leftrightarrow e_2}(TEXT)$ を算出する。

$$O = \frac{\sum_{TEXT} (S \times |2S - 1| \times M)}{\sum_{TEXT} (|2S - 1| \times M)}$$

この式は、 $|2S - 1| \times M$ を重みとする印象値 $S$ の重みつき平均であり、 $|2S - 1|$ 項は $S$ に依存する傾斜配分となっている。この傾斜配分は、印象語群との関係が乏しい一般的な単語(印象値は0.5に近い値をとる)

<sup>†</sup>NHK放送技術研究所の <http://www.nhk.or.jp/str1/tvml/>において無料で公開されている。

<sup>‡</sup>wEEは、記事の見出しと本文(第1段落)を解析し、印象を抽出する。本文の第1段落には記事の印象を決定する事象・事象が現れやすく、その一方で、第2段落以降には第1段落の内容を補足したり説明したりする事実や事例などが現れやすいため、現在の仕様では第2段落以降を用いないことにしている。

表1: 印象辞書のエントリー例

見出し語	品詞	印象値 $S$	重み $M$
シェフ	名詞	0.894	1.032
満喫する	動詞	0.842	1.106
待ち遠しい	形容詞	0.833	1.022
バラエティー	カタカナ	0.857	1.128
ゆったり	副詞	0.820	1.148
負債	名詞	0.075	1.183
圧迫する	動詞	0.080	1.139
低調だ	形容詞	0.125	1.048
エコノミスト	カタカナ	0.251	1.056
依然	副詞	0.176	1.314

表2: 印象尺度「明るい⇔暗い」を構成する印象語群

印象尺度	印象語群
明るい	明るい, うれしい, 嬉しい, 楽しい, 楽しみだ, 好きだ
暗い	暗い, 苦しい, 悲しい, 哀しい, 嫌いだ

表3: wEEの演出に関する仕様

話題タイプ	BGM	声色
明るい	明るいBGM	明るい声
普通	明るいBGM	普通の声
暗い	BGMなし	普通の声

が $O$ 式の平均操作に及ぼす悪影響を削減するために導入されている。

以上のようにして算出された記事の印象値 $O$ に基づいて、話題タイプ(明るい, 普通, 暗い)を判定し、表3に従い、BGMと声色を選択する。具体的には、閾値 $H_1$ と $H_2$ ( $H_1 \geq H_2$ )を設定し、 $O > H_1$ のとき「明るい話題」、 $O < H_2$ のとき「暗い話題」、それ以外のとき「普通的话题」と判定する。なお、表3は、被験者1200人(男女600人ずつ)に対する予備実験の結果に基づいて、それぞれの話題タイプに対し、ニュース番組としての分かりやすさ、親しみやすさ、好感度という点から最も有効であると判断されたBGM・声色が設定されている[9]。

BGMの選択には、印象に基づく楽曲検索システム[10]を応用した。このシステムは、ユーザが入力した印象に対し、印象の近い楽曲を提示する。このとき、ユーザの印象入力、10本の印象尺度(表4参照)の中から1本以上を選択し、それぞれを7段階評価するというGUIベースの方法で行われる。すなわち、各楽曲の印象も各印象尺度における印象値(0~8の実数値)として表現されている。wEEの印象尺度は、表2に示された印象語から構成されているので、「明るい話題」あるいは「普通的话题」と判定される $O$ 値( $H_2 \leq O \leq 1$ )を印象尺度4・印象尺度8上の「明るい」・「楽しい」に対応する値(5~8)に写像するために、

$$v = \frac{3(O - H_2)}{1 - H_2} + 5 \quad (O \geq H_2 \text{のとき})$$

という式を導入し、 $v$ 値を両印象尺度への入力とした。

表4: 印象に基づく楽曲検索のための印象尺度

番号	印象尺度
1	静かな ⇔ 激しい
2	落ち着いた ⇔ 忙しい
3	爽やかな ⇔ 重苦しい
4	明るい ⇔ 暗い
5	荘厳な ⇔ 軽々しい
6	ゆったりとした ⇔ 窮屈な
7	綺麗な ⇔ 綺麗でない
8	楽しい ⇔ 悲しい
9	気持ちが落ち着く ⇔ 気持ちが高揚する
10	心が癒される ⇔ 心が傷つく

表5: デアル調・デスマス調変換表の一部

デアル調表現	デスマス調表現
くる	きます
きた	きました
こない	きません
こなかった	きませんでした
...	...
る	ます
た	ました
ない	ません
なかった	ませんでした

表6: 略記表現・読み上げ表現変換表の例

変換元となる表現	読み上げ表現
適時打	タイムリーヒット
W杯	ワールドカップ
J1	ジェイ・ワン
左前打	レフトまえヒット
原油高	原油ダカ

## 2.2 書き言葉の話し言葉への変換

ニュース記事は、デアル調で書かれているため、そのまま読み上げてしまうと、違和感が生じる。そこで、記事を構成する各文の文末と節末（接続助詞「が」が用いられている場合のみ）をデスマス調に変換する。変換には、Juman 辞書を解析することにより自動構築した変換表（表5参照）を用いる。文末もしくは節末に表中のデアル調表現が現れたとき、対応するデスマス調表現と置き換えられる。なお、デアル調表現のマッチングは、変換表の第1行から順に行われ、マッチングした時点で終了する。

以上で述べたデアル調からデスマス調への変換と同時に、記事特有の略記表現や読みが難しい表現を音声読み上げのための表現に変換する。この変換は、手作業で構築した変換表（表6参照）を用いて行われる。

## 3. wEEの有効性検証

本章では、被験者600人（男女300人ずつ）を対象に行った被験者実験の結果から、BGMや声色がユーザの受ける印象に与える影響を分析し、wEEの有効性を検証する。なお、被験者の年齢構成は、20代が90名（15.0%）、30代が273名（45.5%）、40代が168名

（28.0%）、50代が56名（9.3%）、60代以上が13名（2.2%）であった。

まず、被験者600人を2群（各々男女150人ずつ）に分け、A群の被験者には、wEEが生成したニュース番組のBGM<sup>§</sup>と音声を録音したものを聴取してもらい、B群の被験者には、演出を固定したものの、すなわち読み上げる文章はそのまま、BGMを「普通のBGM」に、声色を「普通の声」に変更したものを聴取してもらった<sup>¶</sup>。各被験者は、アンケート画面の指示に従い、ニュース番組を聴取し、記事の印象（明るい⇔暗い、のんびり⇔緊迫、腹立つ⇔腹立たない、怖い⇔怖くない、好き⇔嫌い）をそれぞれ10点満点（例えば、10点：とても明るい⇔0点：とても暗い）で評価した。ニュース記事は全部で3つあり、普通の話題のニュースは「国庫に眠っていた幻の金貨がオークションに出された」という話、明るい話題のニュースは「ノーベル賞受賞者が秋葉原好きの親日家」という話、暗い話題のニュースは「小学校や保育園でウサギが殺された」という話であった。実験の結果を表7にまとめる。A群とB群の間で行った平均値の差の検定（大標本法）の結果も表7に示す。表中、有意水準1%もしくは5%で統計的に有意な差があるときは、その有意水準を示し、有意水準5%で有意な差がないときは、「—」と記した。

表7に示した検定の結果から、演出を変えることによって、(1) ニュース番組から受ける印象を変えられること、(2) 明るい話題のニュースをより明るく伝えられること、(3) 暗い話題のニュースを怖さや緊迫感を和らげながら、明るめに伝えられることが確認され、wEEの有効性が示されたと言える。

今回の被験者実験では、以上の実験に加え、暗い話題のニュースに対し「暗いBGM・暗い声」という演出を加えた場合の実験も行った。その結果を平均値の差の検定の結果とともに、表8に示す。なお、暗い話題のニュースとして「インフルエンザの特効薬・タミフルを服用した日本人の死者が24人に及んだ」という話を用いた。表8に示された検定の結果から、暗い話題のニュースを暗い演出で表現することにより、ニュースから受ける印象をより暗く、より怖くできることが確認されたが、その一方で印象の悪化を招くことから、今回は表3のままの演出仕様を維持することにした。

なお、現在のwEEでは、演出仕様の設計や評価を単純化するために、「明るい⇔暗い」という印象のみを利用しているが、2.1で述べた印象辞書構築手法を用いれば、怒りや不安感、緊迫感といった様々な印象を抽出することが可能であり、記事のどのような印象に対し、どのような演出を加えるべきかという点を検討していく必要がある。今後の課題とする。

## 4. まとめ

本論文では、記事の印象を演出（BGM・声色の選択）により伝えるニュースリーダーwEEを提案した。wEEは、Webニュース記事をCG形式の動画ニュースに変換する際に、記事の印象に応じてBGMと声色を選択するという点に特徴があり、被験者実験（600人）を通して、その有効性が確認された。また、今回の被験者実験

<sup>§</sup>不確定要素を減らすために、今回の被験者実験では、印象に基づく楽曲検索は行わず、明るいBGMが選択されたときは、ベートーヴェンのバイオリンソナタ「四季」より「春」第1楽章を用いた。

<sup>¶</sup>画像やCGアニメーションが被験者に与える影響を排除するために、BGMと音声を録音したものを聞いた。

表7: 演出が印象の伝わり方に及ぼす影響

普通の話題	演出選択あり (wEE)			平均値の 差の検定	演出固定		
	平均値	標準偏差	データ数		普通の話題	平均値	標準偏差
明るい⇔暗い	4.35	2.32	300	—	4.21	2.25	300
のんびり⇔緊迫	5.17	2.02	300	—	5.07	2.07	300
腹立つ⇔腹立たない	4.92	1.99	300	—	4.75	2.18	300
怖い⇔怖くない	4.43	2.30	300	5%	4.83	2.39	300
好き⇔嫌い	4.16	2.04	300	—	4.02	2.24	300
明るい話題	明るいBGM・明るい声			平均値の 差の検定	普通の話題		
	平均値	標準偏差	データ数		平均値	標準偏差	データ数
明るい⇔暗い	5.51	2.00	300	1%	4.68	2.14	300
のんびり⇔緊迫	5.25	1.72	300	—	5.10	2.02	300
腹立つ⇔腹立たない	4.74	1.84	300	—	4.63	2.24	300
怖い⇔怖くない	4.16	2.04	300	—	4.48	2.32	300
好き⇔嫌い	4.38	1.94	300	—	4.34	2.11	300
暗い話題	BGMなし・普通の声			平均値の 差の検定	普通の話題		
	平均値	標準偏差	データ数		平均値	標準偏差	データ数
明るい⇔暗い	4.23	2.00	300	1%	3.50	2.29	300
のんびり⇔緊迫	4.66	1.72	300	5%	4.31	2.03	300
腹立つ⇔腹立たない	5.34	1.91	300	—	5.48	2.46	300
怖い⇔怖くない	5.11	2.06	300	5%	5.51	2.33	300
好き⇔嫌い	4.26	2.02	300	1%	3.68	2.16	300

表8: 暗い話題のニュースのときの暗い演出が印象の伝わり方に及ぼす影響

暗い話題	暗いBGM・暗い声			平均値の 差の検定	普通の話題		
	平均値	標準偏差	データ数		平均値	標準偏差	データ数
明るい⇔暗い	2.98	2.34	300	1%	3.98	2.12	300
のんびり⇔緊迫	4.62	2.13	300	—	4.49	2.14	300
腹立つ⇔腹立たない	5.14	1.70	300	—	4.99	2.12	300
怖い⇔怖くない	5.83	2.45	300	5%	5.43	2.24	300
好き⇔嫌い	3.67	2.21	300	5%	4.07	2.04	300

では、記事の印象に応じた演出（明るい話題のニュースのときは明るいBGMと明るい声、暗い話題のニュースのときは暗いBGMと暗い声）を加えることにより、その印象を強調できることも確認された。

今後の課題としては、自然かつ多様な感情音声合成の実現、好感の持てるCGキャラクターのデザイン、文／段落単位の印象抽出手法の開発（印象抽出の高精度化）、ニュース番組に適した印象尺度の設計、記事文脈を考慮した演出の実現、略記表現・読み上げ表現変換表の自動構築などが挙げられる。

## 参考文献

- [1] H. Nonogaki, and H. Ueda, FRIEND21 project: A Construction of 21st Century Human Interface, Proc. of International Conference on Human Factors in Computing Systems (CHI'91), pp. 407-414 (1991).
- [2] 矢部純, 高橋伸, 柴山悦哉, ニューススレッドからの番組自動生成, 情処研報, Vol.1999, No.087 (99-HI-85), pp. 13-18 (1999).
- [3] 灘本明代, 服部多栄子, 近藤宏行, 沢中郁夫, 田中克己, Web コンテンツの受動的視聴のための自動変換とスクリプト作成マークアップ言語, 情報処理学会論文誌: データベース, Vol.42, No.SIG1 (TOD8), pp. 103-116 (2001).
- [4] 林正樹, めざせ! テレビ番組のクリエイター パソコンと番組記述言語 TVML で実現!!, 技術評論社 (2005).
- [5] 株式会社アニモ, FineSpeech Ver.2 Emotional Option, <http://www.animo.co.jp/products/tts/fs/>
- [6] 熊本忠彦, 田中克己, Web ニュース記事を対象とする喜怒哀楽抽出システム, インタラクシオン 2005, Vol.2005, No.4 (A-103), pp. 25-26 (2005).
- [7] 日経全文記事データベース DVD-ROM 版, 1990-1995年版, 1996-2000年版, 2001年版, 日本経済新聞社.
- [8] 黒橋禎夫, 河原大輔, 日本語形態素解析システム JUMAN version 4.0 (2003).
- [9] 熊本忠彦, 灘本明代, 田中克己, 感情を陽に伝えるニュース番組の自動生成, 第17回データ工学ワークショップ DEWS2006, No.5C-oi3 (2006).
- [10] 熊本忠彦, 太田公子, 印象に基づく楽曲検索システムの設計・構築・公開, 人工知能学会論文誌, Vol.21, No.3, pp. 310-318 (2006).