

## 感情を陽に伝えるニュース番組の自動生成

熊本 忠彦<sup>†</sup> 灘本 明代<sup>†</sup> 田中 克己<sup>†,††</sup>

<sup>†</sup> 独立行政法人情報通信研究機構 〒619-0289 京都府「けいはんな学研都市」光台 3-5

<sup>††</sup> 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町 36-1

E-mail: <sup>†</sup>{kuma,nadamoto}@nict.go.jp, <sup>††</sup>ktanaka@i.kyoto-u.ac.jp

あらまし 本論文では、感情を陽に伝えるニュース番組自動生成システム wEE (Web2TV with Emotional Expressions) を提案する。本システムは、Web ニュース記事を CG (Computer Graphics) 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 (背景音楽) やニュースキャスター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 (見出し, 本文, 画像) を収集する (2) 見出しと本文から印象 (明るい, 暗い) を抽出する (3) 印象に応じて BGM (明るい BGM, 普通の BGM, 暗い BGM, BGM なし) と声色 (明るい声, 普通の声, 暗い声) 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 TVML (TV program Making Language) を用いて動画ニュースを生成する, という処理を行う。また, 2 回の被験者実験 (1200 人と 600 人) を通じて, BGM や声色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。  
キーワード 受動的視聴, コンテンツ変換, 印象マイニング, ニュース番組演出, アニメーション生成

## Automatic Generation of TV News Program Explicitly Expressing Emotions

Tadahiko KUMAMOTO<sup>†</sup>, Akiyo NADAMOTO<sup>†</sup>, and Katsumi TANAKA<sup>†,††</sup>

<sup>†</sup> National Institute of Information and Communications Technology Hikaridai 3-5, Kansai Science City, Kyoto, 619-0289 Japan

<sup>††</sup> Graduate School of Informatics, Kyoto University Yoshida-Honmachi 36-1, Sakyo, Kyoto, 606-8501 Japan

E-mail: <sup>†</sup>{kuma,nadamoto}@nict.go.jp, <sup>††</sup>ktanaka@i.kyoto-u.ac.jp

**Abstract** We propose a system, named Web2TV with Emotional Expressions (wEE), which automatically generates a TV-like news program with animation, synthesized speech, and background music. A distinctive point of our system is that it explicitly conveys the emotional tone of news articles to the users by determining appropriate background music and matching the voice of a newscaster to the tone of a news article when generating an animated news program. First, the wEE collects news articles, consisting of a headline, text, and zero or more images, from a user-specified news site on the Internet. Next, the wEE determines whether the emotional tone of each article is bright, dark, or neutral by analyzing the headline and text. Then, the system selects a bright, dark, or neutral musical piece and matching tone of voice for the newscaster. If no appropriate music can be determined, then the system can choose to play no music. The wEE then transforms written language into spoken language. Finally, the wEE generates an animated news program with all the above elements using the TV program Making Language (TVML), which was developed by NHK Science and Technical Research Laboratories. We conducted experiments with 1,200 participants to determine the specifications of the wEE and experiments with 600 participants to verify the effectiveness of the wEE.

**Key words** passive viewing, content transformation, sentimental aspect mining, TV program production, animation generation

## 1. ま え が き

インターネットの一般社会への普及・浸透に伴い、様々な Web コンテンツを閲覧する機会が増えている。現在の Web コンテンツ閲覧環境は、能動的閲覧環境であり、ユーザ自身が閲覧したい Web コンテンツを選択し、主体性を持って閲覧する必要がある。しかしながら、ユーザがインターネットにアクセスできる時間・場所は限られているため、閲覧できる Web コンテンツの量はさほど多くない。そのため、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、Web コンテンツを視聴するための受動的視聴環境[1]~[4]が望まれている。

様々な情報をテレビ番組風の動画コンテンツに変換して提示するという基本概念は、FRIEND21 プロジェクトの Nonogaki らにより TV 番組メタファとして提唱された[1]。その後、ネットニュースの議論から脚本を作成し、台詞を CG キャラクタに割り当てることにより、議論をテレビ番組風に提示する方式[2]や Web ページをセグメント化し、セグメントごとに音声と CG キャラクタにより表現・提示する方式[3]、テキストと画像からなる Web ページを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[4]が提案されている。しかしながら、いずれの方式も単に Web コンテンツを動画コンテンツに変換するというメディア変換を行っているにすぎず、コンテンツの内容に応じて演出を変えることはできない。そのため、BGM や CG キャラクタの声色など一旦設定されたものは、コンテンツの内容に関係なく固定され、常に同じ演出(BGM や声色等)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番組自動生成システム wEE (Web2TV with Emotional Expressions) を提案する。本システムは、テキストと画像からなる 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 (背景音楽) や CG キャラクタ (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 (見出し、本文、画像) を収集する (2) 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 (明るい BGM, 普通の BGM, 暗い BGM, BGM なし) と声色 (明るい声, 普通の声, 暗い声) 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)[5] を用いて動画ニュースを生成する、という処理を行う。また、2 回の被験者実験 (1200 人と 600 人) を通して、BGM や声色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。

以下、2. で 1200 人による予備実験を行い、コンテンツの内容 (明るい話題、普通的话题、暗い話題) に応じて BGM や声色をどう変えるべきかという点を分析し、wEE の演出に関する

仕様を設計する。3. で wEE のシステム構成と処理概要、演出方式について述べ、4. でコンテンツの内容に応じて演出を変えることによってコンテンツの有する印象が伝わりやすくなることを被験者実験 (600 人) により示し、wEE の有効性を検証する。最後に 5. で本論文のまとめと今後の課題について述べる。

## 2. 演出仕様を設計するための予備実験

本章では、BGM やニュースキャスターの声色の違いがニュース番組の分かりやすさ、好感度、親しみやすさにどのような影響を与えるのかを調べるための被験者実験を行い、その結果に基づいて提案システム wEE の演出に関する仕様を設計する。被験者は、男女 600 人ずつの計 1200 人であり、20 代が 268 名 (22.3%)、30 代が 571 名 (47.6%)、40 代が 247 名 (20.6%)、50 代が 88 名 (7.3%)、その他の代が 26 名 (2.2%) であった。各被験者は、アンケート画面の指示に従い、ニュース番組 (画像が与える影響を排除するために、BGM と合成音声を録音したものを用いた) を聴取し、ニュース番組を構成する各々のニュースの分かりやすさ、好感度、親しみやすさを 10 点満点 (10 点: とても-分かりやすい / 好感が持てる / 親しみが持てる 0 点: とても-わかりにくい / 不快である / 親しみにくい) で評価した。

### 2.1 BGM の効果を調べるための予備実験

まずはじめに、BGM の有無や BGM から受ける印象の違いがニュースの分かりやすさ、好感度、親しみやすさに与える影響を調べるために、普通的话题のニュース (国の特別天然記念物であるオオサンショウウオが路上にいたという話)、明るい話題のニュース (ノーベル賞を受賞した博士が秋葉原好きの親日家という話)、暗い話題のニュース (小学校や保育園で飼育されていたウサギが殺されたという話) のそれぞれに対し、BGM なし、普通の BGM (草薙考司・Next Season)、明るい BGM (ベートーヴェン・ヴァイオリンソナタ「春」第 1 楽章)、暗い BGM (サティ・グノシエンヌ第 1 番) の 4 種類の BGM を用意した。すなわち、記事内容が同じで BGM が異なる 4 つのニュースをつなぎ合わせ、1 つのニュース番組とした。被験者らの評価結果を表 1、表 2、表 3 にまとめる。なお、ニュースキャスターの声色は、いずれのニュースでも「普通の声」であった。

表 1 は、話題のタイプによらず、「BGM なし」が高評価であったことを示している。そこで、「BGM なし」の場合と他の BGM との間で平均値の差の検定を行ってみたところ、明るい話題のニュースのときの明るい BGM に対しては、有意水準 5% でも有意な差はなかったが、明るい話題のニュースのときの普通の BGM に対しては有意水準 5% で、他の場合には有意水準 1% で有意な差があった。したがって、分かりやすさという点では、BGM はない方が良いが、明るい話題のときは明るい BGM でも良いことがわかる。但し、今回行った被験者実験では、BGM と合成音声を録音したものを用いたため、音質に劣化が見られ、その結果「BGM なし」が高評価につながった可能性もある。なお、「BGM なし」の場合を除いて、BGM から受ける印象と話題のタイプとの関係を調べてみたところ、明

表 1 BGM がニュースの「分かりやすさ」に及ぼす影響

話題	BGM	平均値	標準偏差	データ数
普通	なし	5.12	2.52	1183
	普通	4.75	2.27	1183
	明るい	4.76	2.31	1183
	暗い	4.58	2.35	1179
明るい	なし	4.84	2.44	1183
	普通	4.62	2.24	1182
	明るい	4.70	2.24	1183
	暗い	4.48	2.28	1182
暗い	なし	5.01	2.35	1181
	普通	4.67	2.19	1185
	明るい	4.41	2.18	1182
	暗い	4.63	2.27	1183

表 2 BGM がニュースの「好感度」に及ぼす影響

話題	BGM	平均値	標準偏差	データ数
普通	なし	4.53	2.28	1184
	普通	4.76	2.20	1180
	明るい	4.86	2.30	1182
	暗い	4.10	2.35	1183
明るい	なし	4.43	2.24	1187
	普通	4.69	2.20	1186
	明るい	4.86	2.25	1186
	暗い	4.02	2.31	1185
暗い	なし	4.43	2.26	1188
	普通	4.40	2.20	1185
	明るい	4.05	2.31	1185
	暗い	4.11	2.31	1183

表 3 BGM がニュースの「親しみやすさ」に及ぼす影響

話題	BGM	平均値	標準偏差	データ数
普通	なし	4.25	2.19	1183
	普通	4.61	2.16	1185
	明るい	4.83	2.28	1184
	暗い	4.03	2.26	1181
明るい	なし	4.25	2.16	1183
	普通	4.59	2.14	1184
	明るい	4.80	2.25	1183
	暗い	3.92	2.23	1183
暗い	なし	4.33	2.20	1185
	普通	4.40	2.16	1185
	明るい	4.13	2.27	1186
	暗い	4.00	2.26	1183

るい話題のニュースに対し暗い BGM を用いたり、暗い話題のニュースに対し明るい BGM を用いたりするのは、好ましくないが、他の組み合わせに対しては、分かりやすさに違いがないことが示された（有意水準 5%）。

表 2 と表 3 は、明るい話題や普通の話のニュースに関して、好感度も親しみやすさも「明るい BGM」が高評価であったことを示している。そこで、「明るい BGM」の場合と他の BGM との間で平均値の差の検定を行ってみたところ、好感度という点では、「明るい BGM」と「普通の BGM」が同程度に高評価（有意水準 5%で有意差なし）であり、親しみやすさという点で

表 4 声色がニュースの「分かりやすさ」に及ぼす影響

話題	声色	平均値	標準偏差	データ数
普通	普通	4.80	2.27	1183
	明るい	4.61	2.29	1185
	暗い	4.16	2.46	1186
明るい	普通	4.81	2.23	1182
	明るい	4.50	2.25	1186
	暗い	4.31	2.39	1185
暗い	普通	4.91	2.17	1183
	明るい	4.39	2.19	1182
	暗い	4.55	2.42	1181

表 5 声色がニュースの「好感度」に及ぼす影響

話題	声色	平均値	標準偏差	データ数
普通	普通	4.69	2.22	1186
	明るい	4.52	2.24	1186
	暗い	3.41	2.37	1184
明るい	普通	4.69	2.16	1187
	明るい	4.63	2.23	1183
	暗い	3.50	2.33	1183
暗い	普通	4.59	2.16	1187
	明るい	4.12	2.24	1185
	暗い	3.83	2.40	1183

表 6 声色がニュースの「親しみやすさ」に及ぼす影響

話題	声色	平均値	標準偏差	データ数
普通	普通	4.55	2.21	1184
	明るい	4.46	2.24	1183
	暗い	3.34	2.35	1184
明るい	普通	4.59	2.13	1185
	明るい	4.54	2.19	1185
	暗い	3.46	2.28	1186
暗い	普通	4.57	2.16	1186
	明るい	4.18	2.21	1186
	暗い	3.77	2.35	1186

は、「明るい BGM」のみが高評価（「普通の BGM」に対しては有意水準 5%で、「BGM なし」と「暗い BGM」に対しては有意水準 1%で有意差あり）であった。一方、暗い話題のニュースに関しては、好感度、親しみやすさのいずれにおいても、「BGM なし」と「普通の BGM」が高評価であり、これらの間には有意水準 5%でも有意な差はなかった。

## 2.2 声色の効果を調べるための予備実験

次に、声色から受ける印象の違いがニュースの分かりやすさ、好感度、親しみやすさに与える影響を調べるために、普通の話のニュース、明るい話題のニュース、暗い話題のニュースのそれぞれに対し、普通の声、明るい声、暗い声の 3 種類の声色<sup>(注1)</sup>を用意した。すなわち、記事内容が同じで声色が異なる 3 つのニュースをつなぎ合わせ、1 つのニュース番組とした。被験者らの評価結果を表 4、表 5、表 6 にまとめる。なお、BGM は「BGM なし」とした。

(注1): 音声合成には市販の感情音声合成ソフトウェア FineSpeech Ver.2 [6] を用いた。

表 7 演出 (BGM と声色の組み合わせ方) がニュースの「分かりやすさ・好感度・親しみやすさ」に及ぼす影響

評価項目	演出	平均値	標準偏差	データ数
分かりやすさ	A	4.80	2.12	394
	B	4.54	2.24	398
	C	4.73	2.29	394
好感度	A	4.86	2.06	395
	B	4.58	2.22	398
	C	4.69	2.19	394
親しみやすさ	A	4.66	2.05	392
	B	4.40	2.17	397
	C	4.62	2.21	395

表 4 は、話題のタイプによらず「普通の声」が高評価であったことを示している。そこで、「普通の声」の場合と他の声色との間で平均値の差の検定を行ってみたところ、「普通の話・明るい声」の組み合わせに対しては有意水準 5% で、他の組み合わせに対しては有意水準 1% で有意な差があり、分かりやすさという点では「普通の声」が有効であることが示された。

一方、表 5 と表 6 は、好感度や親しみやすさという点でも、「普通の声」が高評価であったことを示しているが、「明るい声」との差は大きくない。そこで、平均値の差の検定を行ってみたところ、普通の話のニュースと明るい話題のニュースに関しては、「普通の声」と「明るい声」は同程度に高評価 (有意水準 5% で有意差なし) であるが、暗い話題のニュースに関しては、「普通の声」の方が高評価 (有意水準 1%) であることが示された。

### 2.3 BGM と声色の組み合わせ効果を調べるための予備実験

最後に、演出 (BGM と声色の組み合わせ方) の違いがニュース番組全体の分かりやすさ、好感度、親しみやすさに与える影響を調べるために、ニュース構成が同じで BGM と声色の組み合わせ方が異なるニュース番組を 3 本作成した。すなわち、ニュース構成は皆「普通の話 (国庫に保管されていた幻の金貨がオークションで競りにかけられたという話)、普通の話 (郵便局が投函の販売を始めたという話)、明るい話題 (ノーベル賞受賞者が秋葉原好きの親日家という話)、普通の話 (路上にオオサンショウウオがいたという話)、暗い話題 (小学校や保育園でウサギが殺されたという話)」であり、演出には、パターン A 「BGM なし・普通の声」、パターン B 「普通の BGM ・普通の声」、パターン C 「話題と同じ印象の BGM ・話題と同じ印象の声」の 3 パターンを用意した。この実験では、被験者 1200 人を 3 群 (各 400 人) に分け、各群の被験者 (男女 200 人ずつ) が 3 本のうちの 1 本だけを聴取し、ニュース番組全体の分かりやすさ、好感度、親しみやすさを 10 点満点で評価した。結果を表 7 にまとめる。

表 7 から、分かりやすさ、好感度、親しみやすさのいずれにおいても、タイプ A 「BGM なし・普通の声」が高評価を得ているが、平均値の差の検定を行ってみたところ、有意水準 5% でも有意な差はなく、2.1 と 2.2 の実験において低評価であった「暗い BGM」と「暗い声」の組み合わせを部分的に用いても

表 8 ニュース番組に対する嗜好 (a)BGM に対する嗜好

選択項目	頻度	割合
話題によらず明るい BGM	59	4.9%
話題によらず暗い BGM	26	2.2%
話題によらず平静な BGM	281	23.4%
話題の印象に合った BGM	604	50.3%
BGM なし	165	13.8%
その他	65	5.4%
合計	1200	100%

(b) 声色に対する嗜好

選択項目	頻度	割合
話題によらず明るい声	66	5.5%
話題によらず暗い声	20	1.7%
話題によらず平静な声	400	33.3%
話題の印象に合った声	621	51.8%
その他	93	7.7%
合計	1200	100%

表 9 wEE の演出に関する仕様

話題	BGM	声色
普通	明るい	普通
明るい	明るい	明るい
暗い	なし	普通

(パターン C の場合)、ニュース番組全体の分かりやすさや好感度、親しみやすさにはあまり影響しないことが示された。なお、ニュース番組の嗜好に関するアンケート調査の結果 (表 8 参照) から、BGM と声色の両方において、話題の印象に合った BGM ・声色が望まれていることがわかるので、演出仕様の設計においては、この点も考慮する必要がある。

以上の分析結果から、wEE の演出に関する仕様を表 9 のとおりとする。BGM に関しては、明るい話題・普通の話のニュースのとき、好感度と親しみやすさが高評価であった「明るい BGM」を採用することにし、暗い話題のニュースのとき、分かりやすさ、好感度、親しみやすさのいずれにおいても高評価であった「BGM なし」とする。声色に関しては、分かりやすさ、好感度、親しみやすさのいずれにおいても高評価であった「普通の声」を採用することにする。但し、明るい話題のニュースに対しては、ニュース番組に対する嗜好調査の結果を踏まえ、「普通の声」と同程度の評価を得ている「明るい声」を採用することにする。

### 3. ニュース番組自動生成システム wEE

本章では、wEE のシステム構成、処理の流れ、演出方式について述べる。

まず、wEE のシステム構成を図 1 に示し、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<sup>(注2)</sup>を図 2 に示す。

次に、wEE の処理の流れを示す。(1) wEE は、ユーザが指定したニュースサイトから記事を収集する (2) 収集された記

(注2): 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。

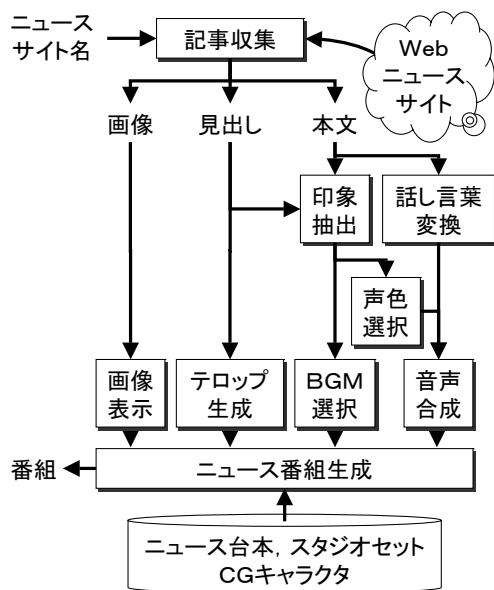


図1 wEEのシステム構成



図2 wEEが生成したニュース番組の例

事の各々から見出しと本文、あれば画像も抽出する(3) 画像はサイズを調整し、定位置(図2参照)に表示する(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」と対応付けられた実数値)を算出し、その値に応じてBGMとニュースキャスターの声色を選択する(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[6]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・CGキャラクタ(ニュースキャスター)に関する情報と処理(3)~(7)で決定された画像、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[5](注3)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)選択と話し言葉変換に焦点を当て、詳しく述べる。

### 3.1 印象抽出に基づく演出選択

まず、記事のテキスト部分(注4)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、計算式に当てはめることによって求められる。以下、印象辞書を自動構築する手法ならびにテキストから印象値を算出する手法について述べる。

印象辞書は、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]を以下の手順で解析することにより構築された。

$y$ 年版に掲載された記事のうち、印象語群 $e$ に含まれる印象語のいずれかを含む記事の数を $df(y, e)$ 、印象語群 $e$ に含まれる印象語と印象辞書において見出し語となる内容語 $w$ の両方を含む記事の数を $df(y, e&w)$ とすると、印象語群 $e$ のいずれかが現れたときに内容語 $w$ も現れる確率 $P(y, e, w)$ は、

$$P = \frac{df(y, e&w)}{df(y, e)}$$

と表される。そこで、印象尺度「 $e_1$   $e_2$ 」を構成する印象語群 $e_1, e_2$ に対し、内容語 $w$ の印象語群 $e_1$ に対する出現確率 $P(y, e_1, w)$ と印象語群 $e_2$ に対する出現確率 $P(y, e_2, w)$ の内分比 $R(y, e_1, e_2, w)$ を

$$R = \frac{P(y, e_1, w)}{P(y, e_1, w) + P(y, e_2, w)}$$

という式で求める。但し、分母=0のときは、便宜的に $R=0$ として処理する。この $R$ 値を年版ごとに求め、以下の式に代入することにより、内容語 $w$ の印象尺度「 $e_1$   $e_2$ 」における印象値 $S_{e_1, e_2}(w)$ が求められる。

$$S = \sum_{y=1990}^{2001} R(y, e_1, e_2, w) \Bigg/ \sum_{y=1990}^{2001} T(y, e_1, e_2, w)$$

但し、 $T(y, e_1, e_2, w)$ は、 $df(y, e_1&w) + df(y, e_2&w) > 0$ のとき、1、そうでないとき、0となる関数であり、 $R$ 式の分母が0となるケースを計算式から除外するために導入されている。

一方、内容語の中には、出現する年や出現頻度が多いものもあれば、少ないものもある。そこで、印象値 $S_{e_1, e_2}(w)$ に対

(注3): NHK 放送技術研究所の <http://www.nhk.or.jp/str1/tvml/> において無償で公開されている。

(注4): 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 10 印象辞書のエントリー例

品詞	見出し語	印象値 $S$	重み $M$
名詞	シェフ	0.894	1.032
	旅先	0.888	1.066
	ハイキング	0.908	1.025
	鑑賞	0.848	1.267
動詞	満喫する	0.842	1.106
	学べる	0.855	1.137
形容詞	カラフルだ	0.847	1.124
	待ち遠しい	0.833	1.022
カタカナ	アーティスト	0.876	1.083
	バラエティー	0.857	1.128
副詞	ゆったり	0.820	1.148
	ひとしお	0.796	1.071
名詞	負債	0.075	1.183
	借入金	0.073	1.142
	減産	0.026	1.064
	値上げ	0.072	1.252
動詞	取り崩す	0.063	1.080
	圧迫する	0.080	1.139
形容詞	必至だ	0.095	1.256
	低調だ	0.125	1.048
カタカナ	エコノミスト	0.251	1.056
	セクター	0.303	1.198
副詞	依然	0.176	1.314
	軒並み	0.192	1.203

する重み  $M_{e_1, e_2}(w)$  を以下のように定義し、内容語  $w$  と印象語群  $e_1, e_2$  とが共起した年数と頻度の総和 (12 年間分) に応じて、増減するように設計した。

$$M = \log_{12} \sum_{y=1990}^{2001} T(y, e_1, e_2, w) \times \log_{144} \sum_{y=1990}^{2001} (df(y, e_1 \& w) + df(y, e_2 \& w))$$

以上の方法で構築された印象辞書の一部を表 10 に示し、印象尺度「明るい 暗い」を構成する印象語群を表 11 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、 $TEXT$  に含まれる名詞、動詞、形容詞、カタカナ、副詞を抽出する。そして、印象辞書から各単語の印象値  $S_{e_1, e_2}(w)$  と重み  $M_{e_1, e_2}(w)$  を取得し、以下の式を用いて  $TEXT$  の印象値  $O_{e_1, e_2}(TEXT)$  を算出する。

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

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) 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。

表 11 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 12 印象に基づく楽曲検索システムにおける印象入力インタフェース

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 13 楽曲の印象表現の例

アイネ・クライネ・ナハトムジーク 第 1 楽章 (モーツァルト) (2.5 3.8 4.5 6.1 4.9 4.2 5.2 5.8 3.8 4.5)
ジムノペディ 第 1 番 (サティ) (6.3 6.4 4.3 2.7 5.1 6.3 5.2 3.5 5.7 5.5)
ジュ・トゥ・ヴ (サティ) (4.9 5.3 4.1 4.3 4.3 5.1 5.0 4.7 4.7 5.0)
春 第 1 楽章 (ビバルディ「四季」より) (2.5 3.6 5.5 5.2 5.0 4.2 5.7 5.2 3.8 4.6)
行進曲 (チャイコフスキー「くるみ割り人形」より) (2.8 2.7 4.5 5.7 4.2 3.7 4.9 5.6 2.9 4.5)

(小数点第 2 位で四捨五入して表示)

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、話題のタイプ (明るい, 暗い, 普通) を判定し、表 9 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」、それ以外るとき「普通的话题」と判定する。

BGM の選択には、筆者らが開発した「印象に基づく楽曲検索システム」[9] を応用した。このシステムは、ユーザが入力した印象に基づいて、印象の近い楽曲を提示する。このとき、ユーザの印象入力は、GUI (グラフィカルユーザインタフェース) を用いて 10 本の印象尺度 (表 12 参照) の中から 1 本以上を選択し、それぞれを 7 段階評価することにより、行われる。一方、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) として表現される。例を表 13 に示す。wEE の印象尺度は、表 11 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、

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

という式を導入し、 $v$  値を両印象尺度への入力とした。

### 3.2 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

表 14 デアル調・デスマス調変換表の例

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
...	...
まない	みません
まなかった	みませんでした
だ	です
る	ます
た	ました
ない	ません
なかった	ませんでした

表 15 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
稼頭央	カズオ
GW	ゴールデンウィーク
HP	ホームページ
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
SMAP	スマップ
日朝	ニッショウ
原油高	原油ダカ

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 14 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 15 参照）を用いて行われる。

#### 4. wEE の有効性検証

本章では、2. で述べた被験者実験の被験者とは異なる 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<sup>(注5)</sup>の BGM と合成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を、すなわち読み上げる文章はそのまま

(注5): 但し、今回の被験者実験では、3.1 で述べた BGM 選択を行わず、話題タイプに応じて 2.1 で用いた BGM を選択するようにした。

で、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。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、番組の好感度、内容の理解度、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（10 点：とても好感が持てる / 理解できる / 明るい / のんびりしている / 腹の立つ / 怖い / 好き 0 点：とても不快である / 全然理解できない / とても暗い / とても緊迫している / 全然腹立たない / 全然怖くない / とても嫌い）で評価した。ニュース番組は全部で 3 本あり、それぞれ 1 つのニュースから構成された。順に、普通の話のニュース（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（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（小学校や保育園でウサギが殺されたという話）であった。結果を表 16 にまとめる。A 群と B 群の間で行った平均値の差の検定の結果も表 16 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 16 に示した検定の結果から、演出を加えることによって、(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 17 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 17 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、その一方で番組好感度の低下や印象の悪化を招くことから、今回は表 9 のままの演出仕様を維持することにした。

#### 5. ま と め

本論文では、Web ニュース記事の印象を演出（BGM と声色の選択）により伝えるニュース番組自動生成システム wEE（Web2TV with Emotional Expressions）を提案した。本システムは、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 と声色を選択するという点に特徴があり、2 回の被験者実験（1200 人と 600 人）を通して、その有効性が確認された。また、今回の被験者実験では、ニュース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、同一話者を想定した自然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、

表 16 演出が印象 (sentiment) の伝わり方に及ぼす影響

評価項目	演出選択あり (wEE)			平均値の 差の検定	演出固定		
	普通の話題				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
番組好感度	4.65	2.22	300	—	4.67	2.31	300
内容理解度	4.76	2.18	300	—	4.96	2.23	300
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				普通の話題		
	平均値	標準偏差	データ数	差の検定	平均値	標準偏差	データ数
番組好感度	4.86	2.01	300	—	4.80	2.27	300
内容理解度	4.57	2.08	300	—	4.87	2.31	300
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				普通の話題		
	平均値	標準偏差	データ数	差の検定	平均値	標準偏差	データ数
番組好感度	4.54	2.04	300	—	4.35	2.23	300
内容理解度	5.47	2.06	300	—	5.33	2.14	300
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 17 暗い話題のニュースのとき暗い演出が印象 (sentiment) の伝わり方に及ぼす影響

評価項目	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
番組好感度	4.10	2.20	300	—	4.69	2.29	300
内容理解度	5.05	2.17	300	—	5.27	2.22	300
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

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。

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] T. Yamaguchi, I. Hosomi, and T. Miyashita, WebStage: An Active Media Enhanced World Wide Web Browser, Proc. of International Conference on Human Factors in Computing Systems (CHI'97), pp. 391-398 (1997).  
 [4] 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).  
 [5] 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).  
 [6] 株式会社アニモ, FineSpeech Ver.2, <http://www.animo.co.jp/products/tts/fs/>  
 [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, 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 (2006).