Information Engineering Express
International Institute of Applied Informatics
2017, Vol.3, No.2, P.1 - 10

Speech-balloon Shapes Estimation for Emotional Text Communication

Ryosuke Yamanishi * , Hideki Tanaka,
Yoko Nishihara and Junichi Fukumoto

Abstract

The purpose of this study is to visualize speech nuance using speech-balloon shapes in text-based communication. Though speech nuance disappears in a text communication, comic readers can recognize speech nuance from the speech-balloon shape, which expresses nuances, emotions, and intentions. Through the subjective evaluation experiments, we confirmed that the shapes of speech-balloon changed the emotion towards the phrases. In this paper, we focus on the relationships between the linguistic features of speech and speech-balloon shapes, and developed an estimation model for speech-balloon shapes with the linguistic features using Naive Bayes classifier. The effectiveness was confirmed especially in the estimation of "Explosion" and "Wave" shapes. The proposed method is expected to be used to visually represent a speech nuance that contains emotion in a text-based communication such as chat system and social networking service.

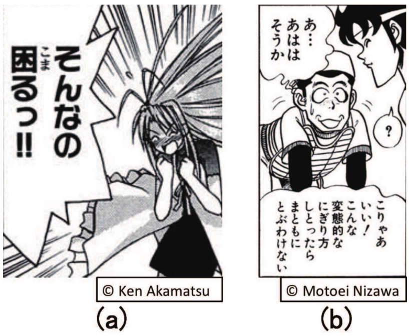
Keywords: Comic computing, Speech-balloon, Visualization, Affective computing

1 Introduction

Humans have text-based communications in our every aspect of daily lives. Text-based communication has been the main medium to communicate with each other without difference of time period: letter, E-mail, and social networking services. Text-based communication becomes more popular along with the development of communication devices such as mobile phones and tablets. In a face-to-face communication, humans sense the emotion/affection and intention of the communication partner, that is speech nuance, with not only semantics but also acoustic features such as "volume" and "quality" of voice. However, such acoustic features can not be expressed in a text-based communication. The limitation might cause the problem that the receiver can not see the sender's intention, and unexpected disagreements would happen in a text-based communication.

Comics express speech nuance by only images and texts without acoustic information [1]. Speech of a character in comics is enclosed by speech-balloon. The speech-balloon shape depends on the speech nuance. Thus, comic readers can roughly sense the speech nuance from the speech-balloon shapes. For example, in Figure 1 (a), the speech of a girl is

* Ritsumeikan University, Kusatsu, Japan



Thanks to Y. Matsui et al.: Sketch-based Manga Retrieval using Manga109 Dataset, Multimedia Tools and Applications, Springer, 2016

Figure 1: The examples of speech-balloon. The readers associate shouting voice with (a), and weak voice with (b). In (a), a girl shouts "そんなの困るっ!! (I don't like it!!) ." In (b), a boy mutters "あ…あはは　そうか' (Ha...hahaha, I see)."

included by a balloon which structure likes explosion. The reader associates shouting voice with the balloon. For Figure 1 (b), the reader associates weak voice with the wave-lined balloon. The speech nuance is visualized by using a speech-balloon shape [2].

The goal of this study is to develop a system to visualize speech nuance in a text-based communication service. Speech nuance might have a concern relation with the semantics or emotions of the utterance content. So, we believe that a sender's intended speech nuance can be indirectly estimated by using linguistic features in the text message. We analyzed the relationships between a sentence in a speech-balloon and the speech-balloon shapes in Japanese comics; the linguistic features and the estimation models are prepared for Japanese. In this paper, we develop a relation model between linguistic features and speech-balloon shapes using Naive Bayes classifier to estimate balloon-speech for input text message.

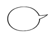
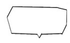
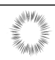
2 Related work

Comic is a mixed medium of images and texts. In comics, the situation including serial information (e.g., acoustics and movement) is spatially represented. Researchers related with comic computing have been reported in varied fields of media processing. The purposes of such researches can be roughly categorized into three types: support to enjoy comics, support to create comics, and application of comic styles.

Researches to classify and explore comics have been proposed to support to enjoy comics. Comics are classified based on drawing style [3]. The preview of a comic episode is automatically generated as thumbnail-like summaries [4]. The support systems to create comics have been reported in the field of image processing. Manga Colorization support to color comics [5]. The background of comics is generated from a real photograph with a screening method, especially for comics [6]. Also, the screentones are automatically detected from comics with a combined method of LoG and FDoG filters [7]. The comic style is applied into communication support and visualization. Comic engine support to interactively browse comic books [8]. A user interface to creation and communication environment is developed with comic styles such as flames [9].

This study can be assumed as a research applying comic styles in a communication

Table 1: Examples of speech-balloon and the generic situation towards the shape.

Speech-balloon shape	Label	Emotion
	Ordinal	None
	Explosion	Anger
	Wave	Anxiety
	Polygon	Polite
	Cloud	Joy
	Square	Explanation
	Cornered-cloud	Emphasize
	Flushing	Surprising

support and visualization. The goal of this study is to visualize speech nuance in a text-based communication by using speech-balloons.

3 Speech-balloon shapes for emotional text communication

According to the analysis of speech-balloon database, the speech-balloon types are classified into eight kinds of shapes. Table 1 shows the speech-balloon shapes according to the analysis balloon in comics. In this paper, the eight kinds of speech-balloon shape are used as the general types of speech-balloon shape. The emotion/affection that the reader generally associates with the shape is also shown in the table. The ordinal balloon and the others are dramatically different each other.

The ordinal balloon is the most common shape in comics, and does not express any kinds of emotion/affection. The other balloons show each corresponding emotion. Comic readers recognize the presence or absence of speech nuance based on whether the shape is ordinal or not. With the other seven types of balloon-shape, the comic reader recognizes the speech nuance, e.g., shouting and weak. That is to say, the recognition of the speech nuance from speech-balloon is processed in two stages: (1) The presence or absence of speech nuance, and (2) What kind of emotion/affection is expressed.

We conducted the preliminary experiment in order to verify the effectiveness of speech-balloon shapes for emotional text communications. Five phrases in a daily communication selected from a Japanese language text for foreigners were prepared as the examples in the experiment. Each of the phrases does not have any kinds of emotion by itself and the expression is normal without any decorative expressions such as an exclamation mark. The phrases were surrounded by each of the eight kinds of speech-ballon shape. Then, 19 participants in 20s–30s who are native for Japanese were asked to answer the emotion towards the phrase and the speech-balloon shape: joy, anger, sad, happy, which are the basic four types of the emotion, and the other.

Table 2 shows the results of the preliminary experiment. Even though the phrase was the

Table 2: The results of the preliminary experiment: the percentages of each emotion for each combination of phrase and speech-balloon shape. The emotions towards the phrase differ depending on the shape of the speech-balloon even though the phrase is the same.

Phrase	Shape	Joy	Anger	Sad	Happy	Other
おかえりなさい (Welcome back)	Ordinal	33.3	4.17	12.50	29.17	20.83
	Explosion	**43.48**	**52.17**	0	0	4.35
	Wave	**45.83**	0	0	**54.17**	0
	Polygon	26.32	8.33	12.50	20.83	16.67
	Cloud	11.54	3.85	38.46	11.54	34.62
	Square	19.05	14.29	23.81	9.52	33.33
	Cornered-cloud	15.79	10.53	15.79	21.05	36.84
	Flushing	**45.83**	16.67	8.33	16.67	12.50
そうなんですか (I see)	Ordinal	10.53	0	26.32	26.32	36.84
	Explosion	19.23	30.77	0	11.54	38.46
	Wave	17.39	4.35	21.74	**52.17**	4.35
	Polygon	4.55	4.55	27.27	22.73	**40.91**
	Cloud	0	0	**65.00**	10.00	25.00
	Square	4.55	9.09	**40.91**	0	**45.45**
	Cornered-cloud	10.53	5.26	36.84	10.53	36.84
	Flushing	18.52	11.11	7.41	14.81	**48.15**
ありがとうございます (Thank you)	Ordinal	**52.17**	4.35	0	34.78	8.70
	Explosion	**52.00**	24.00	0	12.00	12.00
	Wave	**51.85**	0	0	**48.15**	0
	Polygon	**45.45**	18.18	4.55	18.18	13.64
	Cloud	11.54	0	34.62	23.08	30.77
	Square	**41.67**	0	25.00	16.67	16.67
	Cornered-cloud	25.00	0	25.00	10	**40.00**
	Flushing	**68.18**	0	4.55	22.73	4.55
どうしたの？ (What's wrong?)	Ordinal	0	4.17	25.00	20.83	**50.00**
	Explosion	7.14	39.29	0	3.57	**50.00**
	Wave	16.67	0	12.50	37.50	33.33
	Polygon	9.52	9.52	19.05	23.81	38.10
	Cloud	0	3.85	**50.00**	3.85	**42.31**
	Square	4.55	4.55	36.36	0	**54.55**
	Cornered-cloud	0	5.26	42.11	15.79	36.84
	Flushing	0	5.26	36.84	5.26	**52.63**
ひさしぶり (Long time no see)	Ordinal	**48.00**	0	0	28.00	24.00
	Explosion	**41.94**	22.58	0	25.81	9.68
	Wave	37.50	0	0	**62.50**	0
	Polygon	34.78	4.35	13.04	26.09	21.74
	Cloud	11.11	0	**48.15**	11.11	29.63
	Square	21.74	4.35	26.09	21.74	26.09
	Cornered-cloud	10.00	0	25.00	30.00	35.00
	Flushing	**43.48**	8.70	8.70	26.09	13.04

same, the emotion towards the phrase and the speech-balloon was different with each other. While "Welcome back" in an explosion speech-ballon was evaluated as anger or joy, the one in a wave speech-ballon was evaluated as happy or joy which both are positive feelings. "I see" and "Whats' wrong?" in a cloud speech-ballon were greatly evaluated as sad though each phrase showed just an agreement and a question. "Thank you" was evaluated as joy and happy without difference of the shape of the speech-balloon. "Long time no see" in a wave speech-balloon was evaluated as happy though the one in a cloud speech-balloon was evaluated as sad; the polarity of the emotion differed depending on the shapes of speech-balloon. From these results, it is reasonable to say that the emotion towards the phrases is depended on the shapes of the speech-balloon in a comic.

4 Speech-balloon shapes estimation based on linguistic features

Comic creators add some expressions enhancing the emotion to the serif, for example, adding "っ (the expression for an assimilated sound)" and some symbols to the end of the serif, however we prepared the phrases without any decorative expressions in the experiment described in the section 3. We consider that such linguistic features in a serif have some relationships with the shapes of speech-balloon and can be a key to estimating the shapes of speech-balloon.

In the proposed method, the relationships between speech-balloon shape and linguistic features are modeled based on Naive Bayes classification. For our main goal —to visualize speech nuances by using speech-balloon shape, this paper proposes a method to estimate speech-balloon shape based on linguistic features of the input text message. Ordinal speech-balloon is sufficient for general text communication, however, other types of speech-balloon might be required for emotional text communication. Thus, the primal task of this paper is to estimate the seven types of speech-balloon shapes, which is related to the recognition of the speech nuance. The presence or absence of speech nuance would be effective in order to develop more efficient estimation model; the speech-balloon shape would be estimated in two stages. In this paper, each estimation model for seven classes (what kinds of emotion) and two classes (presence or absence of speech nuance) are modeled. In our previous study [10], we collected the database of speech-balloon shapes shown in Table 3. We selected four genres of comic magazines: for boys, for girls, for young men, and for ladies. From each magazine, five titles are selected. One of the authors transcribed the all speeches in each episode and the type of speech-balloon shape.

Through the pre-analysis and discussion for the database, this paper uses four types of linguistic feature to estimate speech-balloon shapes. Our previous study confirmed a certain correlation between speech-balloon shape and the ending part of speech in Japanese comics, which feature is defined as "ending features" in this paper. Also, the following three types of linguistic feature are used; "formal features" such as the number of words and word class, "semantic features" such as emotional terms, and "phonetic features" that are the pronunciation of the speech.

In Japanese, the ending part of the sentence shows the modality and politeness such as "です" and "ます." Table 4 shows the ending features. To extract ending features from the text, the text is firstly morphologically analyzed and divided into part-of-speech. As setting the last term of the text as the source, the terms are checked back while a term excepting terms shown in Table 4 is observed. The number of the observed terms shown in Table 4 is assumed as the features. Note, for f_8^{end}, "!" and "?" are assumed as a single term when these terms continuously occur in this order. For example, the ending features of "つけた…っ!?" can be expressed as $\{f_1^{ending}$=1, f_8^{ending}=1, f_{10}^{ending}=1$\}$.

Table 5 shows the formal features. The features concerning the number of characters are prepared as f_1^{form} for numbers and f_2^{form} for words. The number of words greatly depends on speech. Hence, the number of words is discretized into five levels using k-means with all speeches in the database. The five levels are assumed as the possible variable for the feature concerning the number of words. If a sentence includes a term which word class is person name, the feature concerning person name would be active: $f_3^{form} = 1$. The kinds of character are represented as f_4^{form}, f_5^{form}, and f_6^{form}. In Japanese comics, *Hiragana* is the most-used character to represent pronunciation. *Kanji* is used to show semantics by itself and has its own pronunciation. And, *Katakana* is often used for imported words and coined terms. If the sentence consists of only *Hiragana* characters, then $f_4^{form} = 1$. The sentence includes each *Katakana* and *Kanji* character, then $f_5^{form} = 1$, $f_6^{form} = 1$, respectively. The features for word class are prepared as f_7^{form} for nouns, f_8^{form} for verbs, and f_9^{form} for adjectives, respectively. The features for personal pronoun are prepared as f_{10}^{form} and f_{11}^{form}.

Table 3: Statistics of the speech-balloon database: title, episode, magazine and the number of speech-balloon for each type of speech-balloon.

Title	Episode	Magazine	Ordinal	Explosion	Wave	Polygon	Cloud	Square	Cornered-cloud	Flushing
ノゾ×キミ2年生編/ Nozo × Kimi	1	WEEKLY SHONEN SUNDAY	184	33	23	0	4	31	0	6
食戟のソーマ/ Food Wars!: Shokugeki no Soma	71	WEEKLY JUMP	38	28	5	0	0	0	0	3
エリアの騎士/ The Knight in the Area	362	Weekly Shonen Magazine	67	0	0	43	0	12	31	0
生徒会役員共/ Seitokai Yakuindomo	279	Weekly Shonen Magazine	42	7	0	0	0	2	0	0
弱虫ペダル/ Yowamushi Pedal	305	Weekly Champion	65	31	14	0	4	17	0	9
麻雀飛龍伝説 天牌/ Tenpai	749	Weekly Goraku	24	5	0	0	0	2	0	4
BILLYBAT	118	Morining	100	8	4	4	0	0	0	0
アイアムアヒーロー/ I am hero	180	Big Comic Spirits	74	9	9	0	0	0	0	9
なにわ友あれ/ Naniwa tomo are	337	Weekly Young Magazine	92	10	32	0	0	0	0	0
テラフォーマーズ/ Terra Formars	97	Weekly Young Jump	36	25	0	0	0	11	15	0
ちび☆デビ/Chibi Debi	74	Ciao	42	5	8	0	46	2	0	5
うそつきリリィ/ Liar Lily	104	Margaret	52	6	9	0	0	0	12	0
さばげぶっ！/ Sabagebu!	40	Nakayoshi	58	13	7	0	12	0	2	6
忘却の首と姫/ Bokyaku no kubi to hime	31	Hana Yume	107	3	0	3	0	0	17	0
つばさとホタル / Tsubasa to Hotaru	10	Ribon	104	8	3	6	3	30	22	3
マリアの城/ Maria's Castle	2	YOU	82	10	19	0	5	0	2	1
噂をとめられない！/ Uso wo tomerarenai	1	Harlequin	201	23	38	0	19	12	4	12
たそがれたかこ/ Tasogare Takako	19	BE LOVE	71	8	21	17	5	13	0	0
SUPER G	28	BE LOVE	62	2	16	0	0	0	0	6
バラ色の聖戦/ Barairo no Seisen	69	Kiss	56	7	8	0	0	0	0	8
Total			1557	241	216	73	98	132	105	72

The terms of the first and second person are collected from Wikipedia. If the sentence includes the first person and second person, then each f_{10}^{form} and f_{11}^{form} would respectively be active: $f_{10}^{form} = 1$ and $f_{11}^{form} = 1$.

Table 6 shows semantic features. The polarity of a sentence is represented as the features f_1^{sem}, f_{12}^{sem}, and f_{13}^{sem}. The value of these features is each number of Bad words, positive word, and negative word, respectively. Here, the Bad words are prepared from "Niconico Live: NG words," and the positive and negative words are prepared from the existing study [11]. The features f_2^{sem}–f_{11}^{sem} are prepared for emotion. The value of these features is the number of emotional terms stored in the dictionary [12] in a sentence.

Table 7 shows the phonetic features. f_1^{ph} has the number of moras as its value. Which vowel is used at the end of speech is represented as f_2^{ph}–f_6^{ph}, which features are exclusive.

4.1 Estimation Model

In this paper, the speech-balloon estimation model is constructed as the following two steps: 1) Features selection, and 2) The relationships between speech-balloon shape and the selected features are studied with a Naive Bayes method. The features selection is processed in two stages because a combinatorial explosion might occur if the features are directly

Table 4: Ending features. All of the features have the number of the symbol as the value.

ID	Feature	ID	Feature
f_1^{end}	“っ”	f_{11}^{end}	“—”
f_2^{end}	“ッ”	f_{12}^{end}	“ー”
f_3^{end}	“♡”	f_{13}^{end}	“〜”
f_4^{end}	“☆”	f_{14}^{end}	“です”
f_5^{end}	“♪”	f_{15}^{end}	“ます”
f_6^{end}	“！”	f_{16}^{end}	“ない”
f_7^{end}	“？”	f_{17}^{end}	“かな”
f_8^{end}	“ !?”	f_{18}^{end}	“けど”
f_9^{end}	“。”	f_{19}^{end}	“なんて”
f_{10}^{end}	“…”		

Table 5: Formal features.

ID	Feature	ID	Feature
f_1^{form}	# of numbers	f_7^{form}	# of nouns
f_2^{form}	# of words	f_8^{form}	# of verbs
f_3^{form}	Including person name	f_9^{form}	# of adjectives
f_4^{form}	Consisting of only *Hiragana*	f_{10}^{form}	# of the first person
f_5^{form}	Including *Katakana*	f_{11}^{form}	# of the second person
f_6^{form}	Including *Kanji*		

Table 6: Semantic features. All of the features have the number of words as the value.

ID	Feature	ID	Feature
f_1^{sem}	“Bad-words”	f_8^{sem}	“Love”
f_2^{sem}	“Joy”	f_9^{sem}	“Hate”
f_3^{sem}	“Anger”	f_{10}^{sem}	“Excite”
f_4^{sem}	“Sad”	f_{11}^{sem}	“Relax”
f_5^{sem}	“Surprising”	f_{12}^{sem}	“Positive”
f_6^{sem}	“Shame”	f_{13}^{sem}	“Negative”
f_7^{sem}	“Fear”		

Table 7: Phonetic features.

ID	Feature
f_1^{ph}	# of moras
f_2^{ph}	The speech ends vowel “a”
f_3^{ph}	The speech ends vowel “i”
f_4^{ph}	The speech ends vowel “u”
f_5^{ph}	The speech ends vowel “e”
f_6^{ph}	The speech ends vowel “o”

selected from 49 features. At first, the top 20 features for the mutual information between each feature and speech-balloon shape are selected from the 49 features shown in Table 4, Table 5, Table 6 and Table 7. Since then, the combination of features that shows the highest effectiveness for the estimation of speech-balloon shape is secondly selected from the selected 20 features based on cross-validation.

4.2 Features selection based on mutual information

The speech-balloon shape is assumed as the category c. Let each discrete random variables of f and c be U and C, respectively. The mutual information $I(U;C)$ can be calculated as the follows;

$$I(U;C) = \sum_{e_f \in U} \sum_{e_c \in C} P(U,C) log \frac{P(U,C)}{P(U)P(C)}, \tag{1}$$

where, e_f and e_c each shows presence of the feature and the category that is the speech-balloon shape, respectively.

Mutual information is a measure of the mutual dependence between the two variables. Therefore, the higher mutual information the higher the feature is related to speech-balloon shape. The top 20 features for the mutual information are selected as the effective feature candidates to estimate speech-balloon shape.

4.3 Selection of feature-combination with cross validation

The best combination of linguistic features to estimate speech-balloon shape is selected from the 20 features selected in section 4.2 based on 10-fold cross-validation while chang-

Table 8: The selected features for each two and seven classes classification.

Seven classes				Two classes			
f_1^{end}	f_{13}^{end}	f_4^{end}	f_{14}^{end}	f_1^{end}	f_{10}^{end}	f_2^{end}	f_{13}^{end}
f_6^{end}	f_1^{form}	f_7^{end}	f_4^{form}	f_4^{end}	f_{14}^{end}	f_6^{end}	f_4^{form}
f_8^{end}	f_6^{form}	f_9^{end}	f_4^{sem}	f_7^{end}	f_6^{form}	f_8^{end}	f_1^{sem}
f_{10}^{end}	f_9^{sem}			f_9^{end}	f_{13}^{sem}		

ing the combination of the features. That is, the best combination is selected considering ${}_{20}C_1$–${}_{20}C_{20}$.

Table 8 shows the selected linguistic features to estimate speech-balloon shape. The features for seven classes classification are used to estimate the nuance types of speech-balloon. The features for two classes classification are used to estimate the presence or absence of speech nuance. The cross-validation is conducted with the database collected in our previous study [10]. The speech-balloon shape is estimated by using Naive Bayes classifier in the 10-fold cross-validation.

5 Experiments

The experiments to estimate speech-balloon shapes were conducted with the selected linguistic features described in section 4.1. The speech-balloon shape was estimated by using Naive Bayes classifier. The database of speech and speech-balloon was the database collected in our previous study as same as section 4.3. The effectiveness of the estimation model was evaluated with precision, recall, and F-value for each speech-balloon shape.

5.1 Estimation of seven types of speech-balloon shapes

Table 9 shows the results estimation for each speech-balloon shape. The values in the table are the average of the 10-fold cross-validation. Relatively high F-values were confirmed for “Explosion,” “Wave” and “Square” types. It was suggested that the selected features shown in Table 8 were effective to estimate especially these types of speech-balloon shape.

For “Cornered-cloud” and “Flushing,” the estimation did not work well: both precision and recall were 0%. Through the error analysis, it was confirmed that 40% of each “Cornered-cloud” and “Flushing” was wrongly estimated as the explosion. The matter would be investigated through the discussions of scenes where these two types of speech-balloon shape were used. In the database, the “Cornered-cloud” shape was likely to be used in a scene where a lady strongly spoke in girl comics, and the “Flushing” shape was likely to be used for representing shouting in one’s heart. The “Explosion” type was often used to represent shouting and anger. The speech with these three types of speech-balloon often include “!” and “っ.” Such speeches were similar to each other in the linguistic features. From these facts, it was suggested that the selected features could not cover “shouting in one’s heart” or “a lady’s strong speech,” which were the discrimination from meta view. As the result, such speeches were wrongly detected as the “Explosion” type of balloon-speech which was also used for a speech with an emphatic.

Table 9: The results of the estimation for seven types of speech-balloon: precisions, recalls and F-values for the estimation of each speech-balloon.

Speech-balloon shape	Precision	Recall	F-value
Explosion	44.41%	**67.36%**	**53.35%**
Wave	47.97%	45.45%	**55.47%**
Polygon	50.00%	5.50%	9.88%
Cloud	50.00%	14.52%	22.50%
Square	34.59%	**57.50%**	43.19%
Cornered-cloud	0.00%	0.00%	0.00%
Flushing	0.00%	0.00%	0.00%

Table 10: The results of the estimation for two classes i.e. presence or absence of speech nuance: precisions, recalls and F-values.

Speech-balloon shape	Precision	Recall	F-value
Ordinal	68.64%	**87.63%**	**76.98%**
The others	67.61%	39.20%	49.64%

5.2 Estimation of two types of speech-balloon shapes

Table 10 shows the estimation result of presence or absence of speech nuance. The values in the table are the average of the 10-fold cross-validation. Relatively high F-value was confirmed for "Ordinal" speech-balloon that is an absence of speech nuance, though the low F-value was confirmed in absence of speech nuance. According to the results, it was suggested that the linguistic features shown in Table 8 were insufficient to estimate whether the presence or absence of speech nuance. It was considered that some other features(e.g. context of the conversation) would be necessary for such classification.

6 Conclusions

This paper proposed a method to estimate speech-balloon shape based on linguistic features of the input text to represent speech nuance in a text-based communication. The results of the features selection showed that the ending features were more effective to estimate speech-balloon shape. Through the experiment, relatively high F-values were confirmed for estimation of "Explosion", "Wave" and "Cornered-cloud" shapes. The result of the estimation whether the speech was "Ordinal" shape or not showed insufficient performance. The "Ordinal" type of speech-balloon is the most common shape in comics. Therefore, the "Ordinal" type of speech-balloon would be the default in a communication tool as same as the existing tools. Though the estimation model for "Ordinal" shape or not was insufficient, it is expected that the proposed method for the estimation of the seven types of speech-balloon shape is available when the user intends to express the speech nuance in a text-based communication.

In our future, the feature selection and machine learning method would be compared and improved. Additionally, the context of the conversation such would be taken in the consideration to estimate the speech-balloon shape; which might contribute to the effectiveness of the estimation. The discussions for each comic genres and the speakers' characteristics also will be conducted in our future.

Acknowledgments

This work is supported in part by JSPS Grant-in-Aid for Challenging Exploratory Research #15H02780.

References

[1] K. Nagatani, 漫画の構造学！ *(Structural study of comics)*. INDEX Press, 2000. (in Japanese).

[2] O. Takeuchi, マンガ表現学入門 *(A guide of expression studies of comics)*. Chikumashobo, 2005. (in Japanese).

[3] W.-T. Chu and Y.-C. Chao, "Line-based drawing style description for manga classification," in *Proc. of the 22nd ACM international conference on Multimedia*, pp. 781–784, 2014.

[4] K. Hoashi, C. Ono, D. Ishii, and H. Watanabe, "Automatic preview generation of comic episodes for digitized comic search," in *Proc. of the 19th ACM international conference on Multimedia*, pp. 1489–1492, 2011.

[5] Y. Qu, T.-T. Wong, and P.-A. Heng, "Manga colorization," *ACM Transactions on Graphics (SIGGRAPH 2006 issue)*, vol. 25, no. 3, pp. 1214–1220, 2006.

[6] Y. Qu, W.-M. Pang, T.-T. Wong, and P.-A. Heng, "Richness-preserving manga screening," *ACM Transactions on Graphics (SIGGRAPH Asia 2008 issue)*, vol. 27, no. 5, pp. 155:1–155:8, 2008.

[7] K. Ito, Y. Matsui, T. Yamasaki, and K. Aizawa, "Separation of manga line drawings and screentones," in *EG 2015 - Short Papers*, 2015.

[8] H. Tobita, "Comic engine: interactive system for creating and browsing comic books with attention cuing," in *Proc. of the International Conference on Advanced Visual Interfaces*, pp. 281–288, 2011.

[9] H. Tobita, "Comic computing: creation and communication with comic," in *Proc. of the 29th ACM international conference on Design of communication*, pp. 91–98, 2011.

[10] H. Tanaka, R. Yamanishi, and J. Fukumoto, "Relation analysis between speech balloon shapes and their serif descriptions in comic," in *Proc. of 4th International Congress on Advanced Applied Informatics*, pp. 229–233, 2015.

[11] N. Kobayashi, K. Inui, Y. Matsumoto, and K. Tateishi, "Collecting evaluative expressions for opinion extraction," *Journal of Natural Language Processing*, vol. 12, no. 3, pp. 203–222, 2005. (in Japanese).

[12] A. Nakamura, 感情表現辞典 *(Dictionary for expressions of emotion)*. Tokyodo, 1993. (in Japanese).