Emotional video ranking based on user comments

Murakami, Naomichi
Research institute for IT, Kyushu University | Graduate school of ISEE, Kyushu University

Ito, Eisuke
Graduate school of ISEE, Kyushu University | Research institute for IT, Kyushu University

https://hdl.handle.net/2324/21048

出版情報：Proc. of iiWAS2011 (The 13th International Conference on Information Integration and Web-based Applications, pp.499-502, 2011-12. ACM
バージョン：

Emotional video ranking based on user comments

Naomichi Murakami
Dept. of Advanced Information Technology,
Kyushu University
Fukuoka, Japan
mrkm978@gmail.com

Eisuke Ito
Research Institute for Information Science,
Kyushu University
Fukuoka, Japan.
ito.eisuke.523@m.kyushu-u.ac.jp

ABSTRACT

Nowadays, a lot of people post various contents to the video sharing services. To find good contents, it is necessary something high-quality content search and contents classification method, and the method should understand user's taste and user's context. In this paper, we propose a ranking method based on viewer's comments, especially amount of "funny" feelings comments given by consumer. We also evaluate the questionnaire for our method. Our proposed method is assumed to be applicable to all types of content, if it given a lot of comments from people.

Keywords

Folksonomy, Content analysis, Video content.

1. INTRODUCTION

Recently, a lot of contents are uploading and sharing on the Internet video sharing services such as *Myspace*[1], *YouTube*[2], and *deviantArt*[3], and those services provide commenting system for users. Users are not only enjoying contents on those services all over the world, but also enjoying commenting. The number of online contents is increasing rapidly, and contents are very diverse. This situation is called as the information explosion. It is difficult for users to find good contents with traditional search or ranking system. Something high-quality content search or contents classification method are necessary, and the method should be understand user's taste, user's context.

We believe that folksonomy based method may be effective for good contents search, because folksonomy is suitable for classification of massive web contents. Various techniques based on folksonomy (collective intelligence) such as tag cloud, social bookmarking and collaborative filtering are developed, and those techniques are compared in [4].

We have been focusing on the *Nico Nico Douga* (*NND* for short) movie sharing service, and *NND* is one of most famous video sharing service in Japan. We researched massive contents search methods [5, 6, 7] for NND. To make searching multimedia contents fast and efficient, using annotation, a description standard MPEG-7 has lately attracted attention [8], [9]. An attempt to use comments by viewer on NND for detect video highlights [10].

iiWAS2011, 5-7 December, 2011, Ho Chi Minh City, Vietnam..

In this paper, we propose a ranking method for video contents based on viewers emotion, which are hiding in their comments. We also report on the questionnaire evaluation for it. Our proposed method is assumed to be applicable to all types of content, if it given a lot of comments by people.

2. NICO NICO DOUGA (NND)

In this section, we briefly describe about NND service. NND is a YouTube like user-post type video sharing service, and it started service at 12th of December 2006. The total number of subscribed members of NND is over 567 million people at 31 October 2010. This member includes 18.9 million mobile *members* and 1.01 million premium members. Total number of video in NND is more than 6 million at March 2011. 2.5 million videos are uploaded in recent one year. At present, about one thousand videos are uploaded in one day.

The greatest feature of NND may be the unique comment system. In NND system, viewer is able to give comments to viewing video, and the NND system stores not only posted comment text but also playback time. In case of video replay, viewer's comments are overlaid onto the video, according to the post time. For example, if you post a comment 'Goood!!' to a video at 01:31 passed from the video start, the NND system stores the text string 'Goood!!', and the post time '01:31'. In case of replay of the video, all viewers show your comment 'Goood!!' at 01:31 onto the video movie. Figure 1 is a screenshot of NND's flash movie player, ant it shows viewer's comment texts flow on the video. Actually, figure 1 shows the English version NND flash player. Japanese version is a little bit different, but both players have the same comment system.

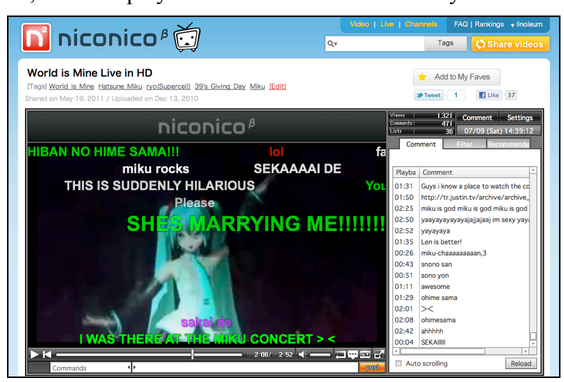

Figure 1. NND flash player. *(English site.)*

Figure 2 is the model of annotations in NND. The video publisher gives the title, description and tags for his/her video. The system determines video length. Viewer can give tags and comments, and the viewer can add the video into his/her favorite list. Number of views, comments, favorite lists can be regarded as annotations given by viewers too, and those may be an indicator of the popularity.

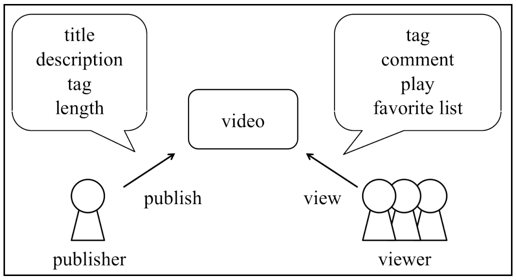

Figure 2. The annotation model of NND.

Some of these annotations given by many viewers are folksonomy resource and useful for determining video weights. We pay attention to tags and comments, because they may express viewer's emotion or feeling that help to detect some video's property.

NND provides two video filtering (searching) systems. One is tag search system, and the other one is tag search system. In NND, video category is specified by category tag(s). NND system decides some major tags as category tags. Keyword filtering is applied for the video title or the description given by the publisher. After filtering, filtered results can be sorted by six keys: "play", "comment" and "favorite list", "length of video", and recently comment. However tag space and video resource are too huge to search video in existing system.

3. TAGS AND COMMENTS RESOURCES

In this section, we describe basic data using our analysis. We collected metadata and comments of all "music" category vides in NND, since October 28 until November 05, 2009. The total number of video is 373,265. Attributes of metadata are title, description, length, and number of play, comments, and favorite lists. We could only get recent 500 viewer's comments for one video at most, even if more than 500 comments are posted by viewers.

3.1. Tags analysis

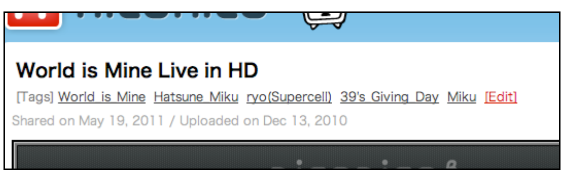

Figure 3. Example of tags.

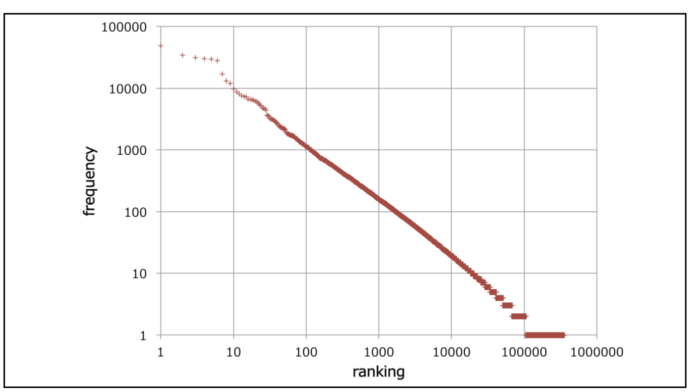

Figure 4. Tag's Frequency-Ranking. (log scale.)

Figure 3 shows an example of video tags in NND. At most 10 tags are given. Anybody can edit (delete and add) tags of a movie, so the publisher can lock tags (set disable editing), and most of locked tags are category or genre tags.

Figure 4 shows rank frequency plot, and it follows the power-law distribution. There were 355,872 unique tags in the music video set. 251,229 tags (equivalent to 70.6% of all) are occurred only once, and most of these tags were joke tags or description words of the video. These tags are not useful for search, but it may be useful for emotional ranking because these tags may have viewer's excitement.

3.2. Comments analysis

We gathered comments for each 373,265 videos. Total size of comments is over 6GB. The number of comments per video; average is 355, median is 29. Only 27,152 (about 7%) videos were given over 500 comments. The number of characters per comment; average is 11, median is 9. (Japanese 2-byte character regarded as one character). This results shows that, viewers gave comments to top 10% popular videos, and most of comments are very short.

We also checked comment texts, and we found that there are a lot of typos, and ungrammatical texts. We also fount many new words, and most of them are abbreviation of some words. Then, it is difficult to apply usual natural language processing techniques for comment analysis. These results are similar with text of twitter or text of bulletin board system. So, we used n-gram (actually, 1-gram) analysis.

TABLE I. CHARACTER'S FREQUENCY (TOP-10 CHARACTERS)

Rank	Character	Frequency	Rank	Character	Frequency
1[-]	w	31,398,314	6[16]	ー	7518,485
2[10]	い	14,963,261	7[1]	の	6,556,529
3[57]	あ	9,014,433	8[21]	っ	6,324,308
4[12]	な	8,228,804	9[17]	か	6,260,110
5[-]		7,854,930	10[37]	う	5,783,100

Table 1 shows frequently of characters in comments. For comparison, the rank of the character in "The Nikkei" newspaper [11] is also described in rank column with "[ ]". (We checked 520,685 characters used in January 2009 articles.)

The 1st character in table 1 is "w". This is very interesting, and very characteristic. The alphabet "w" doesn't occur in usual Japanese text. It never appears "w" in "The Nikkei". The word "w" is a Japanese Internet slang as same as "lol" in English. "w" comes from the first letter of "warai", where it means "laugh". When a viewer laughed a video, then (s)he enter "w" in comment.

The 3rd character is an interjection or an expression of shout, is like "Ah". It may be used for sing, because our target is music videos. The 7th character is often used in a case particle, ant it is similar with "of" in English. It was the most frequent letter in newspaper resource. Comments of NND are simpler than newspaper, but this word frequently occuer.

4. EMOTIONAL RANKING

NND has an interesting culture in which viewer gives comments include viewer's emotion such as the word "w" which means laughing. Under this situation, we propose a new emotional ranking of video contents. The definition of ranking method is shown in bellow equations.

$$W_v = \frac{\sum w_i}{n(C)} * \log(n(C)+1)$$

C is set of comment given for a video v, $n(C)$ is size of C. W_v is the weight of the video, and it average of emotional word w_i and logsclae of $n(C)$.

And we defined w_i as follows at this time.

$$w_i = \begin{cases} 1 & \text{if } c_i \text{ include "w"} \\ 0 & \text{if } c_i \text{ doesn't include "w"} \end{cases}$$

So, the weight W_v is accumulation of the ratio of comments which include the letter "w" and not. The reliability is a logarithm of comment set size, which used for weighted.

5. EVALUATION

In order to evaluate proposed ranking method, we analyzed ranking results with poplar videos. We checked "List of *Vocaloid* producers". Vocaloid is a vocal synthesizer, and it is very popular in NND. There were 1,153 publisher (vocaloid music producer) are listed in the *Nico Nico Pedia* [12] (Wikipedia like DB in NND). Only a few of them are professional music creator, that is almost publishers are amateur. 12,824 videos were given music producer's nametag, and 2,247 videos were given over 500 comments. We picked the top 10 producers who publish vocaloid music videos given over 500 comments. From 313 videos produced by the top 10 producers, we chosen 100 videos that each producer's top 5 of play, and top 5 of W_v. 6 videos were overlap, 2 videos were deleted 3 videos were difficult to access, then we got remained 89 videos as the result.

5.1. Popularity v.s. W_v

Figure 5 is scatter plot of 89 videos. The horizontal axis is the number of play (popularity), the vertical axis is W_v value. Figure 5 shows interesting results, that is, high W_v value videos are not poplar.

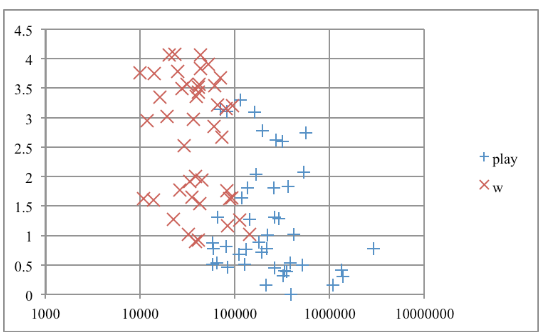

Figure 5. Picked up 89 videos. *(semi-logarithmic scale.)*

Figure 6. Result of questionnaires.

For more detail, we checked 89 videos by 8 persons, and they answered following three questions. (a) Do you think the video is aiming for laugh? (b) Is it funny for you? (c) Do you want to see again? Each person answered to each question by "yes" or "no". So each video has scores of 3 items, and given 0 … 8 points to each item. The result of questions is shown in Figure 6. It shows the number of videos given each point for each item. Whether the video is aimed for laugh was easy to match the subject's view. The scores of "*Aiming for laugh*" were divided into high and low results. On the other hand, "*Funny*" and "*Want to see again*" were concentrate in lower scores.

5.2. Spearman's Rank Correlation Coefficient

At first, we calculated Spearman's rank correlation coefficient between 5 items. Ranking videos in descending order of three questionnaire's items, value of "*Play*" and "*W*".

TABLE II. SPEARMAN'S RANK CORRELATION COEFFICIENT

(a) laugh	(a) laugh				
(b) funny	0.890	(b) funny			
(c) again	- 0.207	- 0.240	(c) again		
Play	- 0.200	- 0.172	0.426	Play	
Wv	0.730	0.742	- 0.254	- 0.468	*Wv*

" (a) *Aiming for laugh*" correlated with "(b) *Funny*" and W_v, on the other hand, it didn't other two. This result doesn't represent that W_v indicates worth watching, but W_v related with "(b) *Funny*" feelings. Then our proposed method may extract funny videos. "(b) *Funny*" had a weak negative correlation with "(c) *Want to see again*" (according to t-test; p = 0.01). This may cause our method couldn't derive worth videos and it didn't show correlation with *Play*. Correlation between "(b) *Funny*" and W_v is stronger than "(a) *Aiming for laugh*" and W_v. "(c) *Want to see again*" had a weak correlation with *Play*. *Play* had a negative correlation with W_v. It doesn't mean a video has high W_v value is worthless to watch; W_v had a positive correlation with "*Funny*" feelings. W_v includes some foolish elements often, and it might cause this.

5.3. F-measure

W_v had a tendency to take a higher value for video related to feelings of laughter. But we have not defined specific values for judging whether a video is interesting.

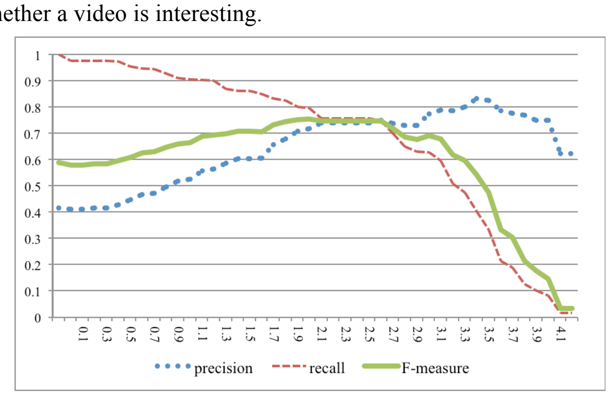

Figure 7. Transition of F-measure.

Figure 7 examined whether how much value give the better accuracy when *divided* videos into two classes by the value of W_v. The answer is using the data of the questionnaire survey results. We estimated by W_v whether each person answered videos as "*Aiming for laugh*". Using threshold in increments of 0.1, we observed the change of precision, recall, F-measure. X-axis is the threshold for W_v.

F-measure takes the maximum value of 0.76 at W_v=2.0, and keeps a value *around* 75% in the range of about 1.8-2.7. If you want to discover some funny video from large quantity videos, the precision may be more important than recall. Precision takes the maximum value of 0.84 at W_v =3.4. If the W_v is grater than 4, precision is lowered against our expectations. Those videos, which have high value of "*W*", are very few (only 3 videos have more than 4), and it cause decline of precision value.

5.4. ROC curve

Figure 8 and 9 are ROC curve. Figure 8 used threshold for W_v changing in increments of 0.1, similar to the above analysis about F-measure. We observed specificity and sensitivity for 3 points of view - "*Aiming for laugh*", "*Funny*", "*Want to see again*". W_v shows good curve for "*Aiming for laugh*" and "*Funny*". The AUC of "*Aiming for laugh*" was 0.82.

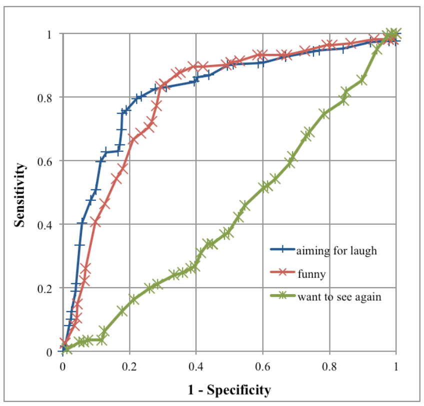

Figure 8. ROC curve by W.

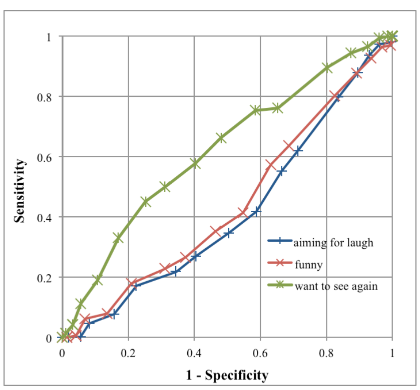

Figure 9. ROC curve by logarithmically the Number of Play.

On the other hand, Figure 9 is drawn in the same way, but it took the number of play, instead of W_v. In general, the number of play follows a power-low in NND. So we used logarithmic of play number to draw a detailed diagram, to be precise. It was dotted for each one-fiftieth of the maximum. The play number is important perspective for finding video in NND, and to support it, ROC for "*Want to see again*" shows a little good curve. Its AUC was only 0.60.

Figure 8 can be regarded as an existing search system's curve. If you want to watch videos with out concretely plan, in existing system, you should use video weighting depend on play number, whose AUC is 0.60. However, if only you have a vaguely wish to discover funny video somehow, our proposed weighting having over 0.8 AUC can help you.

6. IMPLEMENTATION

We built a video search system [13] using proposed ranking method. This system searches 373,265 music category videos using W_v ranking. Figure 10 shows the system. The background color will change tint according to dependency of the value of W_v for each video. Overall, searching by different key word will bring more interesting videos.

Figure 10. Demo system of emotional ranking.

7. CONCLUSION

We have researched the contents search method, focused on NND. In this paper, we proposed a emotional weighting method for video contents ranking using viewers comments, and report on the questionnaire evaluation for it, using correlation coefficient, F-measure, ROC curve. We showed comments given for multimedia contents contain some pieces of viewer's sentiment, and our proposal was effective video, considering its simplicity. However there are also some problems. Our method is heuristic and it is weak for SEO. It can't decide other feelings for video except "fun". We should pay attention to the other characters or words to solve these problems.

REFERENCES

[1] Myspace, http://www.myspace.com, (accessed 2011.7.10)

[2] YouTube, http://www.youtube.com, (accessed 2011.7.10)

[3] deviantART, http://www.deviantart.com, (accessed 2011.7.10)

[4] I. Guy, N. Zwerdling, I. Ronen, D. Carmel, E. Uziel: Social media recommendation based on people and tags, ACM SIGIR2010, pp.194-201, 2010.

[5] Nico Nico Douga, http://www.nicovideo.jp, (accessed 2011.7.10)

[6] Nico Nico Douga: From Wikipedia, the free encyclopedia., http://en.wikipedia.org/wiki/Nico_Nico_Douga, (accessed 2011.7.10)

[7] niconico, http://www.niconico.com, (accessed 2011.7.10)

[8] H. Sack, J. Waitelonis: Integrating Social Tagging and Document Annotation for Content-Based in Multimedia Data, 1st Semantic Authoring and Annotation WorkshopAthens, (2006)

[9] D. Yamamoto, K. Nagao: iVAS: Web-based Video Annotation System and its Applications, International Semantic Web Conference 2004, (2004)

[10] T. Nakamura, H. Aoki, H. Miyashita: NICOVIDEO Navigation System using comic-like Visualization, Information Processing Society of Japan, (2008) (in Japanese)

[11] Nihon Keizai Shimbun: From Wikipedia, the free encyclopedia., http://en.wikipedia.org/wiki/Nihon_keizai_Shimbun, (accessed 2011.7.10)

[12] Nico Nico Pedia, http://dic.nicovideo.jp, (accessed 2011.7.10)

[13] SeeMore, http://joe.cc.kyushu-u.ac.jp/seemore.