NEWS VIDEO STORY SENTIMENT CLASSIFICATION AND RANKING

Chunxi Liu[1], Li Su[1], Qingming Huang[1,2], Shuqiang Jiang[2]
[1]Graduate University of Chinese Academy of Sciences, Beijing, 100190, China
[2]Key Lab of Intell. Info. Process., Inst. of Comput. Tech., CAS, Beijing, 100190, China
{cxliu, lsu, qmhuang, sqjiang}@jdl.ac.cn

ABSTRACT

In this paper, we present a novel approach for news video story sentiment analysis. Two research challenges are addressed: news video story sentiment classification and ranking. For classification, a graph based semi-supervised learning approach is utilized to classify the news stories into sentiment classes. Graph based semi-supervised learning is able to tackle the problem of lacking labeled data. After classification, two sentiment classes are obtained: positive and negative. In order to project the news videos into sentiment space, a multimodal approach by fusing the text sentiment and visual representation scores is adopted to rank the videos in each class. For sentiment representation, inter and intra sentiment class analysis is conducted based on affinity propagation clustering and *PageRank* algorithm. A user study is conducted to evaluate the video ranking performance. The experimental results on the selected topics are promising and demonstrate the proposed approach is effective.

Keywords— News video analysis, sentiment classification, personalized news video, semi-supervised learning

1. INTRODUCTION

In order to facilitate people's access to news and search video content, news video analysis has been a hot research topic for a long time. The state-of-the-art news video analysis covers structure analysis, semantic concept detection, annotation and search. Through structure analysis, the individual news stories are segmented and news video can be accessed by story level indexing. After concept detection and annotation, the semantics in the stories and videos can be understood. News video search can retrieve the user interested stories. All above research topics address the objective aspect of the news video. However, news video is produced by collecting news from different sources and interviewing different people. For some critical events, different people and government may give different perspectives. For example, for the controversial topic "*Kosovo independence*", there are obviously two sentiment perspectives from different countries, which lead to opposite sentiment reports. Two examples collected from web are listed below.

(1) Title: *Bush backs Kosovo independence, assures help*

"*President George W Bush has said that the US will work with its allies to try to prevent violent clashes after the anticipated declaration of independence by Kosovo. …… "The United States will continue to work with our allies, to the very best we can, to make sure there's no violence,"……*"

(2) Title: *Serbia rejects Kosovo independence*

"*Serbia's government proclaimed on Thursday that any unilateral act by Kosovo's ethnic Albanian leadership to declare independence would be invalid and illegal. The government adopted the resolution with Kosovo expected to declare independence within days……"*"

The first report reflects the *United States* supports *Kosovo* independence, while the second reveals *Serbia* rejects *Kosovo* independence. This example indicates that there are different perspectives even in the same news topic. Once an important and critical political event happens, most of the video channels will conduct continuous report. The attitudes in those news videos can generally be classified into two classes [1]: 1) positive, which holds supportive and praiseful attitude towards the happened event; 2) negative, which possesses opposed and critical attitude towards the happened event. If the sentiment orientations of the video stories can be obtained, it will help news video review and personalization. For example, after news search, a lot of news video stories are presented to the user and in the order according to their similarities to the text query. It would be useful to know which story gives positive report and which presents negative. This is especially important for news commenter, whose job is to collect response from different perspectives of the same topic in news video and give comment on the event by comparing these different perspectives. It is time consuming and labor intensive for the commenter to collect the sentiment report manually.

Existing work on news video analysis focuses on structure analysis, semantic concept detection, annotation and search. The state-of-the-art approaches on sentiment analysis mainly focus on reviews [2][3], political debate [4], blog [5], *etc*. To the best of our knowledge, there is no related work on news video sentiment analysis. In this paper, we present a novel framework to detect the sentiment orientation of news video topics and rank the stories according to their sentiment representation. We focus on analyzing those important and critical political news topics, which have obviously different perspectives. There are two research

978-1-61284-350-6/11/$26.00 ©2011 IEEE

issues to be addressed: sentiment classification and ranking. The proposed framework is shown in Figure 1. For classification, firstly, the sentiment words are selected. Then, by using the selected words as features, news stories are classified into sentiment classes based on graph based semi-supervised learning. For sentiment ranking, the multimodal fusion approach is adopted to build the final ranking list in each sentiment class. The contributions of the paper include: 1) we present a novel framework to mine news video sentiment based on text and visual analysis; 2) we propose a semi-supervised learning approach for news story sentiment classification; 3) we propose a novel approach for news story visual sentiment representation by combining affinity propagation clustering and *PageRank* algorithm.

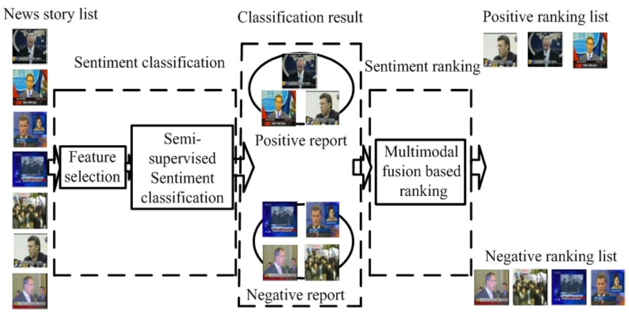

Fig. 1 Sentiment classification and ranking framework

The rest of the paper is organized as follows. Section 2 presents the news story sentiment classification approach. In section 3, a multimodal fusion approach for sentiment ranking is proposed. In section 4, the experimental results are reported and analyzed and two applications are proposed. Finally, we conclude the paper in section 5.

2. SENTIMENT CLASSIFICATION BASED ON SEMI-SUPERVISED LEARNING

In news video, the reports with different perspectives are reflected by the words of the anchorpersons and reporters. If these words can be understood, the sentiment classification will be relatively easy. Fortunately, the automatic speech recognition (*ASR*) results of the anchorpersons and reporters are available in news video. Although the *ASR* data for some speeches are not robust, it can generally help to detect the sentiment orientation of the news video stories.

2.1. Graph based Semi-supervised Learning

In order to classify news stories into sentiment classes, we encounter the insufficient training data problem. By now, there is no method on news video sentiment analysis and no benchmark for evaluation either. Manually labeling training data is time consuming and labor intensive. On the other hand, unlabeled news stories can be obtained through search engines and news websites easily. In order to tackle the problem of insufficient training data and make use of the large unlabeled data, a graph based semi-supervised learning approach [6] is adopted.

2.2. Sentiment Word Selection and News Story Sentiment Classification

In order to analyze the sentiment of news stories, we have to select proper features. The *ASR* data consist of many words. However, not all the words are suitable for sentiment representation, as many words are topic dependent and cannot reflect sentiment. Another difficulty for sentiment analysis lies in that although the overall sentiment orientation of a comment is definitely positive (negative), some words or sentences in the comment may also reveal negative (positive) sentiment. If we can distinguish the positive words from the negative words, the sentiment classification will become relatively easy. Sentiment words used in previous work (e.g. [1][3][7] *etc*) are manually generated or semi-automatically generated with seed words. We consider two methods for sentiment word generation: manual selection and semi-supervised word selection. For manual selection we choose the words with strong sentiment tendency (e.g. *support*, *oppose*, *etc*) in the news data. For semi-supervised word selection, we employ the graph based semi-supervised learning method. When constructing the affinity matrix, each word is considered as a node. Assume w_i and w_j represent two words, their occurrence in the document is represented as $p_i = \{d_1^i, d_2^i, ..., d_n^i\}$ and $p_j = \{d_1^j, d_2^j, ..., d_n^j\}$, where n is the number of document and d_k^j represents the frequency of word j in document k. The similarity between word w_i and w_j is calculated as:

$$similarity(w_i, w_j) = \begin{cases} \exp\left(-\frac{\sum_{k=1}^{n}(d_k^i - d_k^j)^2}{\sigma^2}\right) & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The assumption behind this similarity is that concurrent words tend to have same sentiment tendency. After the sentiment words are selected, the next step is to use them as features for news story sentiment classification. Due to the insufficient training samples, we also adopt the graph based semi-supervised learning approach. The key problem here is how to measure the sentiment similarity between those news stories. There are many approaches focusing on the reviews, where lot of sentiment features have been tested including unigram, bigram, trigram feature, *etc*. The experimental results show that although the unigram feature is simple, its performance is the best one [3]. There are other methods analyzing the sentiment by comparing the number of positive and negative words [7] in the text or exploring the percentage of positive and negative sentence feature [8]. Motivated by the above work, we propose to use two types of features to represent each document and compare their performance by experiment. The first is the bag-of-word feature. The second is the number of positive words and negative words in a document. After feature extraction, each document d_j is represented by a vector $\{f_1, ..., f_m\}$ and m is the number of sentiment features. For graph construction, the similarity between d_i and d_j is calculated as follows:

$$similarity(d_i, d_j) = \begin{cases} \exp\left(-\dfrac{\sum_{k=1}^{m}(f_k^i - f_k^j)^2}{\sigma^2}\right) & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

3. NEWS STORY SENTIMENT RANKING

It is straightforward to use the classification confident scores to rank the news video stories in each class. The confident score of story *i* belongs to class *j* can be calculated as follows:

$$Con\_socre_i^j = |F_{ij}^*| / \sum_{j=1}^{c} |F_{ij}^*| \quad (3)$$

where F^* is obtained from the classification result of the algorithm [6]. However, we argue that using this score is not suitable. Although the news story maybe determined to have strong sentiment tendency by text analysis, the video may only contain the anchorpersons. At the same time, the visual impact is stronger than the information conveyed by the text. Therefore, it is necessary to find a way to rank videos by combining the text and visual information. Before ranking, we propose to check the video part to see if there are any useful clues for sentiment ranking. Some typical story images collected for the topic "*Ukraine election*" are shown in Figure 2, where each image represents a story.

Fig. 2 Typical images for sentiment classes

From Figure 2, we can see that near duplicate images exist in the same and across different sentiment classes. Near duplicate image is very useful for news video analysis and has been used for topic tracking [9] and story clustering [10]. Actually not only near duplicate images but also similar scenes (*e.g.* the protesting scenes in Figure 2) in the video reveal similar sentiment. We also find that the news stories with different sentiment (e.g. positive or negative) usually contain the video scenes corresponding to the sentiment classes. For sentiment ranking, the multimodal fusion approach is adopted, which is shown in Figure 3.

For the visual part, the shots are first clustered. For clustering, the Affinity propagation (*AP*) [11] is adopted, which can automatically choose the right cluster number. In our approach, each data point is a shot and is represented by a keyframe with the *HSV* 256 bins (16 hue by 4 saturation by 4 brightness levels) global histogram. Then each visual story is scored through *PageRank* algorithm. Finally, the ranking result is obtained through multimodal fusion.

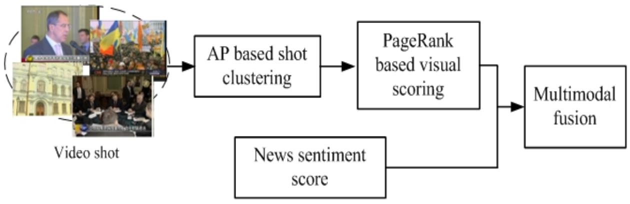

Fig. 3 Sentiment ranking approach

After classification, each story is classified into its sentiment class and with confident score Con_socre. Assume the obtained visual representation score is V_{score}. This score represents the importance of the visual story. Then, the final ranking score R_{score} can be obtained as follows:

$$R_{score} = Con\_socre * V_{score} \quad (4)$$

This fusion scheme is commonly used and has been proved to be effective [12]. By observation, the news stories with different sentiment classes usually contain the scenes corresponding to the sentiment classes. Based on this observation, two assumptions are made:

1) The visual similar scenes shared by the stories in the same class will boost its discrimination for sentiment representation.

2) The visual similar scenes shared by the stories in different classes will decrease its discrimination for sentiment representation.

Based on the above assumptions, it is natural to consider the visual representation as a page ranking problem [13]. The share of similar scenes can be considered as a link between two news stories. The problem here is how to construct the linking matrix *F* and use the information shared between the same and different classes. Assume there are totally *N* stories for a topic, and the first *m* stories are positive while the left ones are negative. We have proposed to use the *AP* algorithm to cluster the video shots. For *AP* clustering, each cluster has a true exemplar and the other data are clustered into the cluster according to it. Thus, the linking matrix *F* can be set as:

$$F(i,j) = \begin{cases} 1 & \text{if story i and j share similar scenes and in their cluster the scene in story j is the exemplar} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

This setting means that the exemplar images are important ones in the clusters. If there is only one class for ranking, this setting is feasible. However, we have two sentiment classes. Thus, we have to revise *F* in order to rank each sentiment class properly. After matrix is revised, we will obtain two matrixes: *PF* for positive class and *NF* for negative. In the following, we describe how to revise the *F* matrix for positive sentiment visual ranking. Assume there is a link from story *i* to story *j*. There are four cases:

1) i belongs to positive class and j belongs to positive class too. In this case, we set $PF(i,j)=F(i,j)$.

2) i belongs to positive class and j belongs to negative class. Because we are ranking the positive videos, the linking outside the positive class is useless. Therefore, we set $PF(i,j)=0$.

3) i belongs to negative class and j belongs to positive class. This means that the positive story j shares similar visual scenes with negative story i. If we keep this link, it will increase the ranking score of story j in the positive list. However, based on our assumptions, this link should decrease the ranking score of j in positive list. Therefore, we set $PF(i,j)=0$ and $PF(k,j)=1 \quad k \in postive$ class and $k \neq i$.

4) i belongs to negative class and j belongs to negative class too. Because we are dealing with positive class, the link in the negative class will be set to 0, that is $PF(i,j)=0$.

After the linking matrix is obtained, we use *PageRank* algorithm [41] to obtain the positive visual representation scores from *PF*. The positive scores are obtained by the following function.

$$PR(i)=(1-d)\sum_{j:j\rightarrow i}\frac{PR(j)}{O(j)}+d\frac{1}{N} \qquad (6)$$

where $PR(i)$ represents the visual representation score for story i, $O(j)$ represents the out-links of story j(that is $o(j)=\sum_{i=1}^{M}PF(j,i)$), d is a dampening factor and represents the probability jumping to a random story and N is the number of stories. The equation is recursive until the score convergence. After that, we fuse the text sentiment score with the visual important score to obtain the final positive sentiment ranking list. The *NF* is revised in the same way as *PF* and the negative scores are obtained from *NF*.

4. EXPERIMENTAL RESULT

4.1. Data Collection

In the experiment, ten representative topics are selected:(1) "*Ukrainian presidential election*, 2004", 59 stories; (2) "*Iran nuclear problem*, *2004*", 36 stories; (3) "*Lebanon-Israel conflict*, 2006", 33 stories; (4) "*North Korean nuclear six-party talks*", 16 stories; (5) "*Fiji coup, 2006*", 12 stories; (6) "*Bhutto assassination, 2007*", 15 stories; (7) "*Iran detained British sailors, 2007*", 9 stories; (8) "*The independence of Kosovo problem，2008*", 12 stories; (9) "*American and Iran naval confrontation，2008*", 11 stories; (10) "*US-Russian Anti-Ballistic missile issue, 2008*", 15 stories. These topics are typically critical international political events. The experimental data for topics (1) and (2) were collected from TRECVID 2005 data set. We use only the Chinese and English news, because we are not familiar to Arabic. The *ASR* result and *MT* transcripts of these data are provided by *NIST*. For other topics we collected the video and *ASR* data from the *CCTV* website [14].

4.2. Experimental Result for Sentiment Classification

Before classification, we have to label the ground truth. In order to do the labeling more objectively, three subjects, one female and two males and aging from 23 to 28, were asked to give their labels, positive or negative, on each story. The final labels were decided by majority voting. In the experiment, the agreements are 88% for all and 12% for two subjects. The label result is shown in Table 1, where T represents topic label, P represents the number of positive reports and N denotes the number of negative reports.

Table 1. The component of each topic

T	1	2	3	4	5	6	7	8	9	10
P	19	13	21	6	3	7	4	7	4	6
N	40	23	12	10	8	8	5	5	7	9

We downloaded some news reports from the news websites such as *CCTV* [14], *CNN* [15], *etc* as the training set. Totally, 400 news reports are collected, 200 positive reports and 200 negative reports. Three experiments were conducted to justify our assumptions. The first experiment used the bag-of-word feature. Those words were selected by pruning out the words whose frequencies are less than four times in the data set, which is the same way as in [3]. As a result, 2669 words were selected out of the 6822 words. The second experiment employed a semi-supervised learning method for word selection. 20 words for each class were labeled. After semi-supervised learning, those words were ranked by their sentiment scores and the top 25% words in each class were selected. Finally, 1334 sentiment words were obtained. The third experiment was based on the manually selected words, 148 positive words and 123 negative words. We counted the number of sentiment words in each document and used the number of positive and negative words as the feature for classification. Before testing on the news video reports, we tested the method on the downloaded 400 news stories. The classification results are shown in Figure 4.

In Figure 4, x axis represents the labeled data in each sentiment class and y axis represents the classification error rate. From Figure 4 we can see that, in experiment 1 and 2 with the number of the labeled data increasing the error rate first decreases and then increases. The best performance is achieved when the labeled number is 10 and the error rate is 37%. From these two experimental results, we conclude that the bag-of-word and the bag-of-sentiment-word feature are not as effective as expected. Experiment 3 shows that with the labeled number increasing the error rate first increases a little and then decreases sharply. The best classification performance is 77% and is achieved when the labeled number is 100. One might argue that the simple positive and negative number comparison method may also generate good results. We count the number of positive and negative word in each story and classify the document into the class with more words. The result shows the classification precision is only 36%.

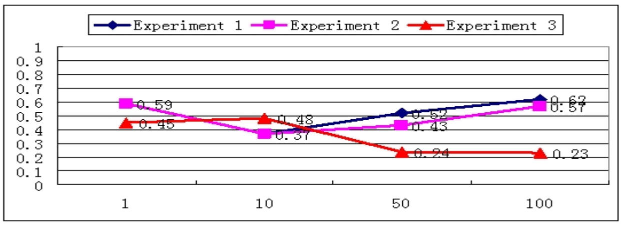

Fig. 4 News sentiment classification evaluation

Because of our limited data set, we only made a comparison experiment with the *PMI* method [2] and not with other machine learning methods, which need a lot of training samples. In the experiment of *PMI* method, the selected seed words are "*positive*" and "*negative*" and the experiment used *AND* search in *Google*. The result showed that the classification accuracy is only 56%. And most of the positive stories are classified into negative class, which is complied with the result reported in [16].

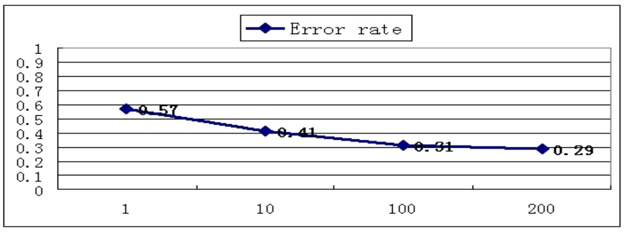

Fig. 5 The final sentiment classification result

For the final classification, we only use the number of positive and negative words as feature. The experimental result is shown in Figure 5. The labeled data are selected randomly from the collected 400 news stories. The best performance is achieved when the labeled number is 200 and the precision is 71%. Although the proposed method is effective, the performance is not as good as the result reported in [3] for movie review. This may be due to the following reasons: Firstly, compared with the movie review, the news data are more challenging and the sentiment word mixed in the news is more complicate; Secondly, the collected data set is small and cannot cover the full sentiment space; Thirdly, the similarity used in constructing the graph is not as effective as expected; Finally, the experiment in our approach contains many topics with diverse ways for sentiment representation. We need to find a way to track the sentiment representation shift from topic to topic.

4.3. Experimental Result for Sentiment Ranking

After multimodal fusion, we obtained two ranking lists for each class: text based ranking list and text visual fusion based list. In order to show the effectiveness of our approach, we need to evaluate the experiment result. In [17], the ranking results were evaluated by the mean average precision criteria (*MAP*). However, using the *MAP* criteria for our sentiment ranking evaluation is not suitable, because the *MAP* is used to evaluate how many retrieved images are the right ones in the top *N* rank. In our approach, the positive and negative stories are pre-collected and they are all related to the given topic. By now, there are no existing criteria to evaluate the sentiment ranking result. Therefore, we adopted a subjective user study approach [18]. Five users were invited to participate in the study, 4 males and 1 female aging from 23 to 30. On one hand, the news video stories in the top ranks will be the most representative ones. On the other hand, evaluate all the videos are hard and time consuming. Like the *MAP* criteria we only evaluate the top two stories in each sentiment class. The subjects are asked to vote each video based on the following two criteria:

1) Sentiment matching: whether the users are satisfied with the sentiment intensity matching between the selected video clips and the sentiment class.

2) Visual matching: whether the users are satisfied with the matching between the sentiment class and the visual information in the selected video clips.

Five scores are defined for each of the criteria: 5: better, 4: good, 3: common, 2: bad, 1: worse. For each video, the average score of the five subjects is adopted for each criterion. Finally, the overall performance of each video is calculated as the average of the above two values and is shown in Figure 6-7.

The selected videos by the two methods for negative topic *T4*, *T7*, *T9.1* and positive topic *T4*, *T5*, *T6*, *T7* are same. We can see that the evaluation performance for these videos is acceptable and the evaluation scores are all bigger than 3. From the result we can also find that for negative topic *T6.2*, *T8.2* and for positive topic *T3.2*, *T8.2*, *T9.1*, *T9.2* the performance of the fused approach is much better than the text approach. This is because the videos selected by the text approach only contain the anchorpersons. Although their sentiment scores are high, their visual matching score is low. This result justifies our initial assumptions and shows that the fused ranking approach is reasonable. For other topics in the negative class, the performance of the fused approach is generally better than the text approach. For positive topic *T1.1*, *T2.1*, *T3.1*, *T10.2*, the performance of the text based approach is a little better than the fused approach. However, we can see that both of the two approaches have better performance on these videos. By our observation, the videos selected by the fused approach are the ones with near duplicate images or similar scenes, which turn out many times in different videos. For the negative topic *T2.1*, both performances are low. This is because the selected videos by the two approaches are same and have strong sentiment tendency but only contain anchorpersons. Even by adding the visual representation score for ranking cannot balance the sentiment intensity and the visual representation.

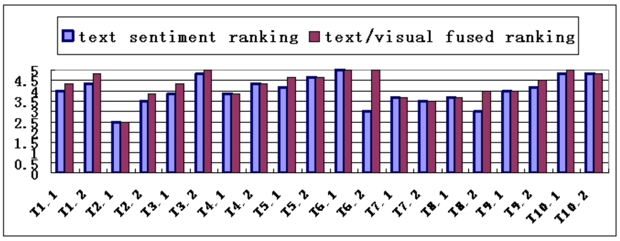

Fig. 6 The overall performance for negative videos

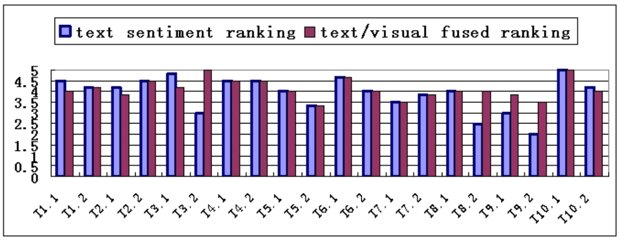

Fig. 7 The overall performance for positive videos

4.4. Application Example

An application example of our method for the topic "*Kosovo independence*" is shown in Figure 8. The retrieved news video reports are classified into different sentiment classes (positive or negative) and ranked according to their sentiment representation. By this way, we can see the news event more clearly.

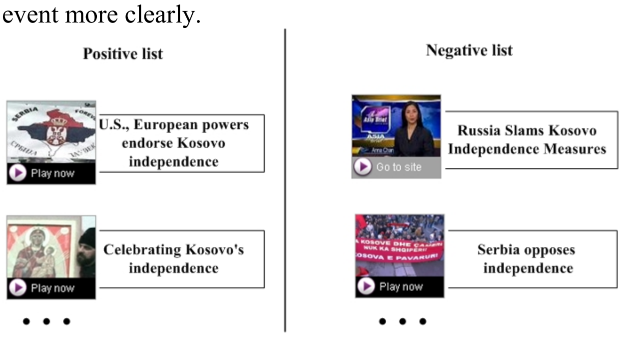

Fig. 8 Retrieved results ranked by sentiment

5. CONCLUSION AND FUTURE WORK

Most of the current news video research work focuses on structure analysis, semantic concept detection, annotation and search. The state-of-the-art approaches for sentiment analysis mainly focus on review, political debate, blog, *etc*. In this paper, we have proposed a novel approach on news video sentiment analysis. To the best of our knowledge, this is the first solution towards news video sentiment analysis. Two research issues are addressed, which are news video topic sentiment classification and news video story sentiment ranking. For the first issue, we propose to employ the graph based semi-supervised learning approach to overcome the insufficient training data problem. For the second issue, we apply the multimodal fusing approach to obtain the final sentiment ranking list, where the sentiment score is obtained by classification and the visual representation score is obtained based on the *AP* clustering and *PageRank* algorithm. The experimental results demonstrate the effectiveness of the proposed method.

6. ACKNOWLEDGEMENT

This work was supported in part by National Natural Scence Foundation of China: 61025011, 60833006 and 61070108, and in part by Beijing Natural Science Foundation: 4092042.

7. REFERENCES

[1] M. Godbole, N. Srinivasaiah, M. Skiena, "Large scale sentiment analysis for news and blogs," in Proc. Int. Conf. Weblogs and Social Media, 2007.

[2] P. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," in Proc. ACL, Philadelphia, PA, USA, 2002.

[3] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in Proc. EMNLP, pp79-86, Philadelphia, USA, 2002.

[4] M. Thomas, B. Pang, L. Lillian, "Get out the vote: determining support or opposition from Congressional floor-debate transcripts," in Proc. EMNLP 2006.

[5] T. Fan, and C. Chang, "Sentiment-oriented contextual advertising", Knowledge and Information Systems, vol.23, no.3, pp321-344, 2010.

[6] D.Y. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Scholkopf, "Learning with local and global consistency," in Proc. Neural Information Processing Systems, 2003.

[7] L. Ku, L. Lee, T. Wu, and H. Chen, "Major topic detection and its application to opinion summarization," in Proc. ACM SIGIR, pp627-628, Salvador, Brazil, 2005.

[8] B. Pang and L. Lee, "Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales," In Proc. ACL, 2005.

[9] W.H. Hsu and S.-F. Chang, "Topic tracking across broadcast news videos with visual duplicates and semantic concepts," in Proc. Int. Conf. Image Processing, 2006.

[10] X. Wu, C.W. Ngo, and Q. Li, "Threading and autodocumenting news videos: a promising solution to rapidly browse news topics," IEEE Signal Processing Magazine, vol.32, no.2, 2006.

[11] B. J. Frey, D. Dueck, "Clustering by passing messages between data points," Science, vol.315, pp972-976, 2007.

[12] Haveliwala, H. Taher, "Topic sensitive PageRank," In Proc. Int. Conf. WWW, 2002

[13] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: bringing order to the web," Technical report, Stanford University, Stanford, CA, 1998.

[14] http://www.cctv.com

[15] http://edition.cnn.com

[16] M. Taboada, C. Anthony, and K. Voll, "Methods for creating semantic orientation dictionaries," in Proc. Int. Conf. Language Resource and Evaluation, 2006.

[17] W. H. Hsu, L. S. Kennedy, S. F. Chang, "Video search reranking via information bottleneck principle," in Proc. ACM Multimedia, 2006.

[18] J. Chin, V. Diehl, and K. Norman, "Development of an instrument measuring user satisfaction of the human-computer interface," in Proc. SIGCHI on Human Factors in CS, pp213-218, Washington, D.C., United States, 1998.