

Received December 6, 2019, accepted December 16, 2019, date of publication December 23, 2019,
date of current version January 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2961392

An Intelligent Video Tag Recommendation Method for Improving Video Popularity in Mobile Computing Environment

RENJIE ZHOU^{ID}^{1,2,3}, DONGCHEN XIA^{ID}^{1,2}, JIAN WAN^{ID}^{1,2,4}, AND SANYUAN ZHANG^{ID}³

¹School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China

²Key Laboratory of Complex Systems Modeling and Simulation, Ministry of Education, Hangzhou Dianzi University, Hangzhou 310018, China

³College of Computer Science and Technology, Zhejiang University, Hangzhou 310058, China

⁴School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China

Corresponding author: Jian Wan (wanjian@zust.edu.cn)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation under Grant LY17F020030, in part by the Zhejiang Provincial Key Technology Research and Development Program under Grant 2019C03134, in part by National Natural Science Foundation of China under Grant 61972358, and in part by the National Key Technology Research and Development Program of China under Grant 2019YFB2102101.

ABSTRACT Big data generated from social media and smart mobile devices has been regarded as a key to obtain insights into human behavior and been extensively utilized for launching marketing activities. A successful marketing activity requires attracting high social popularity to their contents, since higher popularity usually indicates stronger influence, more fame and higher revenue. In this paper, we focus on the question of how to improve popularity of videos sharing on websites like YouTube in mobile computing environment. Obviously, composing high quality titles and tags is beneficial for viewers to discover videos of their interests and increase their tendency to watch more videos. However, it is not an easy task for uploaders, which is especially true since the screen is tight for most mobile devices. To this end, this paper proposes a novel hybrid method based on multi-modal content analysis that recommends keywords for video uploaders to compose titles and tags of their videos and then to gain higher popularity. The method generates candidate keywords by integrating techniques of textual semantic analysis of original tags and recognition of video content. On one hand, taking the original keywords of a video as input, the method obtains most relevant words from WordNet and related video titles gathered from the three top video sharing sites (YouTube, Yahoo Video, Bing Video). On the other hand, through recognizing video content with deep learning technology, the method extracts the entity name of video content as candidate keywords. Finally, a TF-SIM algorithm is proposed to rank the candidate keywords and the most relevant keywords are recommended to uploaders for optimizing the titles and tags of their videos. The experimental results show that the proposed method can effectively improve the social popularity of the videos as well as extend the length of video viewing time per playback.

INDEX TERMS Mobile computing, social media, big data, video tagging, video popularity, artificial intelligence, deep learning.

I. INTRODUCTION

In recent years, mobile computing becomes more and more popular, as reported by Perficient, 58% of web visits were from mobile devices. With the widespread usage of social media platforms and mobile devices, more and more people are interested in interacting, sharing, and collaborating

The associate editor coordinating the review of this manuscript and approving it for publication was Anfeng Liu^{ID}.

through online social media, causing the amount of data generated by the social media grows exponentially. Undoubtedly, the social media and mobile computing have considerably changed people's daily lifestyle and the advertisement model as well. Understanding of interests of users and how they behave are valuable for a variety of communities such as researchers, marketers, politicians, and so on. Big data analytics and cognitive computing are developed to satisfy this need by combining technologies from multiple disciplines such as

artificial intelligence, network analysis, statistics, time series analysis, and natural language processing.

Because of the great influence of the social media and the capability of carrying out precise advertisement under the help of big data and artificial intelligence techniques, an increasing number of commercial content, commodities and activities are promoted through social media and mobile devices, which yields better performance than other forms of media as reported [38]. As for ordinary users, they may become popular if their content goes viral. No matter what kind of activities, commercial or individual, the most critical point is how to attract more attention to their content.

In general, there are two directions to obtain more social media attention. The first direction is to leverage on search engines and recommender system technologies [1]. It is well known that search engine can search almost all kinds of information; however, items retrieved by search engine are usually noisy. On the contrary, the information suggested by a recommendation system has a high degree of relatedness in general, but it is not as good as the search engine in terms of information volume. Therefore, it is more effective to improve the popularity of video content by exploiting the potential of both search engines and recommendation systems. The second direction is about the metadata of social media. For videos on YouTube, the most important metadata is the title and tags of a video uploaded by users. High quality title and tags of a video can be beneficial to attract attention from viewers and get recommended by recommender systems, so how to make high quality title and tags for a video is the focus of our work to improve video popularity [21].

The method proposed in this paper adheres to two principles. Firstly, the recommended keywords must be relevant to the video content but not to attract social popularity by deceiving the viewers. Secondly, on the basis of satisfying the first principle, the method should suggest keywords which are expected to attract high social popularity. In order to adhere to the first principle, the method obtains candidate words relevant to the original title and tags of a video by leveraging on WordNet. In addition, the method uses these keywords to search videos with similar topics on the mainstream video sharing sites like YouTube, Yahoo Video, and Bing Video and then obtain relevant keywords. Furthermore, the proposed method applies deep learning driven image recognition technology to extract the entity names from video content as candidate keywords as well. According to the second principle, the paper proposes a novel algorithm named Term Frequency-Similarity (TF-SIM) to sort the collection of candidate keywords related to the video. The algorithm combines the main ideas of Term Frequency-Inverse Document Frequency (TF-IDF) algorithm and Normalized Google Distance (NGD) algorithm to recommend the most relevant keywords to users. With the recommended keywords, users could optimize the titles and tags of their videos. By comparing the original video title, tags and the optimized title, tags, of a video, it is found that the optimized titles and tags can attract more social popularity and extend the viewing time per playback.

The contributions of the paper can be summarized as the following three points:

(1) The paper proposes a method for suggesting keywords that are relevant to video content. Firstly, based on the original title and tags of a video, the collection of candidate keywords relevant to a video is suggested through querying WordNet and extracting relevant keywords from titles and tags of related videos searched from the mainstream video sharing websites. Since, search engines usually retrieved the most popular videos that match with the input keywords. Thus, in this way, we can not only guarantee the semantic relevance but also the popularity of the candidate keywords. Secondly, the method applies deep learning driven video content analysis technology to identify the entity names from video content and combine them into the set of candidate keywords.

(2) A novel algorithm TF-SIM is proposed for ranking the collection of candidate keywords relevant to a video. On one hand, the algorithm considers the frequency of a keyword by calculating the number of occurrences of the keyword in the candidate keywords; On the other hand, the algorithm uses the NGD formula to calculate the semantic similarity between the keyword and the original video keywords.

(3) The proposed method is evaluated in a real scenario. A set of 50 videos and their original titles and tags were collected from the Internet. Then, the titles and tags were optimized using the keywords recommended by our method and other two baseline methods, and then were assigned to a copy set of the 50 videos, respectively. The four sets of videos with their corresponding titles and tags were uploaded into four new YouTube accounts, respectively. The comparative experimental results show that the titles and tags optimized using our method can effectively boost the social popularity and extend viewing time of a video.

II. RELATED WORKS

With the rapid development of the Web 2.0 and mobile computing technology, the volume of images, videos and other multimedia content on the Internet is growing rapidly, various recommendation systems for helping users quickly find contents of their interests have been proposed [1], [37], [39]. In order to evaluate and improve browsing experience in mobile computing environment, quality of mobile service have been heavily studied [2]–[8]. Prediction of the popularity of web content has been attracting abundant attention from academic and commercial communities [41], [42]. Furthermore, in the face of the large-scale social media big data, how to discover the potential rules and valuable information to boost the social popularity of the video content is becoming another hot topic in the research community [16].

There have been some related works focusing on image recognition, image tag assignment, refinement and retrieval [9]–[12]. Recently, Yu et.al. proposed a several novel methods to improve performance of image content recognition [43], [44]. According to the idea that similar images share similar tags, Makadia et al. proposed the k-Nearest Neighbor (KNN) algorithm, which first finds the adjacent

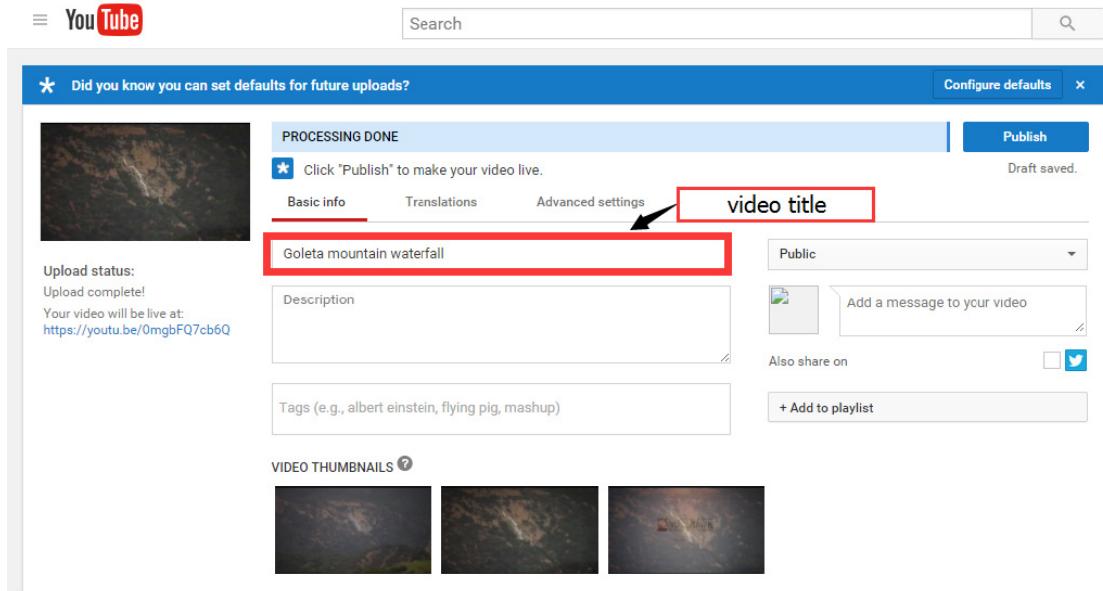


FIGURE 1. Video uploading page.

images within a range of similarity, and then calculates the frequency of adjacent image tags, and finally sorts the tags. The idea of this method can also be applied to the video tag recommendation. Video tag refinement and retrieval is also a hot research field [13]–[15]. Ballan et al. use the original tags of a video to search a set of corresponding images from the Internet, then compare the video’s key-frames with the image set, and finally extract the tags of these most similar images as the expansion of the video’s tags. This method generates candidate tags by collecting a number of large image sets on the Internet and thus improves the diversity of tags.

The study presented by Santos-Neto et al. is, to the best of our knowledge, the closest to this work [16]. In their study, they use a variety of resources associated with a video, such as news reports, reviews, blogs and so on, to recommend new tags to the video. Their study showed that the resources on the network could improve the potentiality of YouTube video tags to attract social popularity.

This research combines the ideas of the above research, but differs from the above research. Our research proposes a hybrid video tagging method based on multi-modal content analysis, which combines the textual semantic relevance analysis and deep learning driven video content recognition to generate candidate video keyword set, and then ranks the candidate keywords using a novel keyword-sorting algorithm proposed in the paper to recommend the best keywords to the uploaders. The results of experiments conducted on YouTube platform show that our method is able to optimize the original title, tags and improve the social popularity and the length of viewing time of a video.

III. PROBLEM STATEMENT

The scenario of uploading a video onto YouTube is shown in Figure 1. As can be seen from the figure, the uploader

should provide a title, a piece of description and several tags for the video s/he wants to upload. In the paper, the title and tags assigned by uploaders without assistance of our method is referred to as original title and tags.

However, it is usually difficult for the uploaders to provide complete and attractive textual information for a video. As reported in paper [17], the textual information of a video is usually incomplete and inappropriate. To this end, our work aims to recommend keywords for users to compose and optimize title and tags for a video based on the video’s content and original title and tags, which in turn can help viewers discover videos of their interests and attract more social attention to the video. In order to achieve the goal, the following three questions should be addressed:

(1) How to generate relevant keywords for a video based on the original textual information provided by the uploader. Normally, the textual information of a video semantically represents the content of the video to some extent. Hence, it is of high value to recommend relevant keywords for a video based on the original textual information. The key problem here is how to recommend keywords that are not only semantically relevant to the original text but also of high potential to be utilized by search engine and recommendation system to help viewers discover videos.

(2) How to obtain relevant keywords based on the video content. To recommend relevant keywords by recognizing and understanding the content of a video is advantageous, since it is a good way to ensure that the recommended keywords are semantically relevant to the video content. Although, it is still impossible to thoroughly understand the semantic of video content at this moment, new technologies such as deep learning have output high accuracy in recognizing the entities from images. So how to exploit the deep learning technology to identify the important entities of the video is the key to this question.

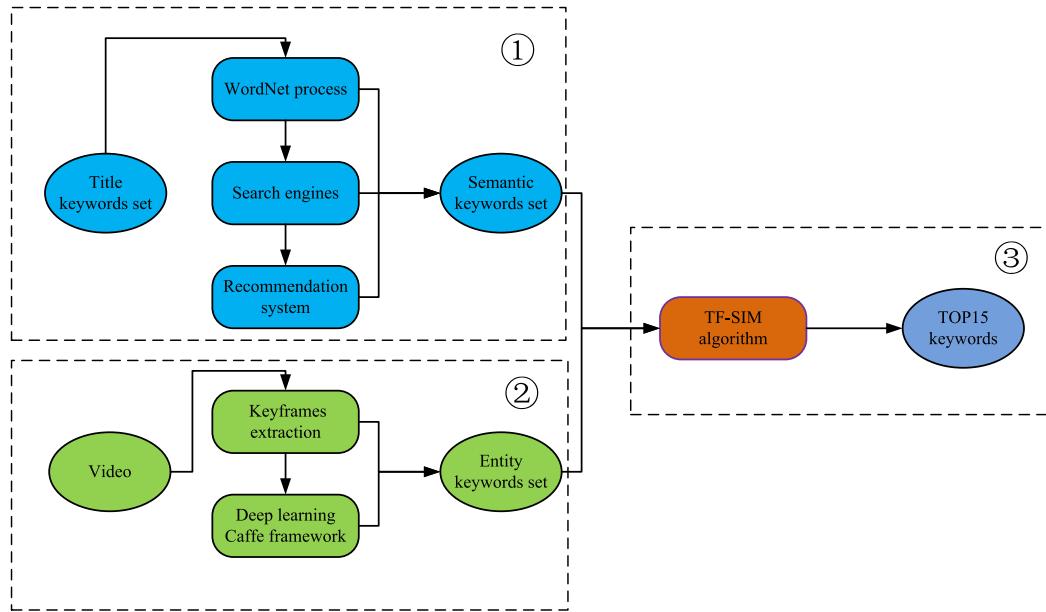


FIGURE 2. The framework of the method.

(3) After obtaining the relevant candidate keywords, the last question is how to rank them and finally recommend the best keywords to the uploader. As for ranking the candidate keywords, two principles should be adhered. Firstly, the keywords need to be relevant to the content of a video. Secondly, the keywords need to have great potential to help the video obtain high social popularity. The problem is how to rank the keywords appropriately by balancing the relevance and potentiality of attracting popularity of the keywords.

IV. METHODOLOGY

The framework of the proposed method consists of three parts, as shown in Figure 2. The first part is marked as 1 in the diagram, which addresses the problem of suggesting keywords based on the original textual information given by the uploader. In this part, the process of suggesting keywords starts from querying WordNet [18] for obtaining keywords semantically relevant to the original title and tags. Then, the method uses these keywords to search videos with similar topics from the mainstream video sharing sites like YouTube, Yahoo Video, and Bing Video and obtain the related videos of these videos recommended by recommender systems. Finally, for the top searched videos, collect the related videos of each video and gain relevant keywords.

The second part is marked as 2 in the figure, which focuses on recommending candidate keywords by analyzing the video content. It begins with extract key-frames for a given video, and then identifies the entity names from video key-frames with deep learning framework Caffe [19]. The entity names of video content are considered as the candidate keywords for the video.

The third part is the TF-SIM algorithm (marked as 3), whose function is to order the candidate keywords and

ultimately suggest the top 15 keywords excluding the initial title keywords given by the uploader.

Under the assistance of our video keyword recommendation system, the process of composing the title and tags of a video is like follows. Firstly, the uploader composes the title and tags of a video as usual. Then, our method recommends to the uploader with top 15 keywords (except the initial title keywords) which have high relevance to video content and original textual information and high potentiality to improve the video popularity. Finally, the uploader optimizes the title and tags of the video using the 15 recommended keywords and then uploads the video onto the video sharing website.

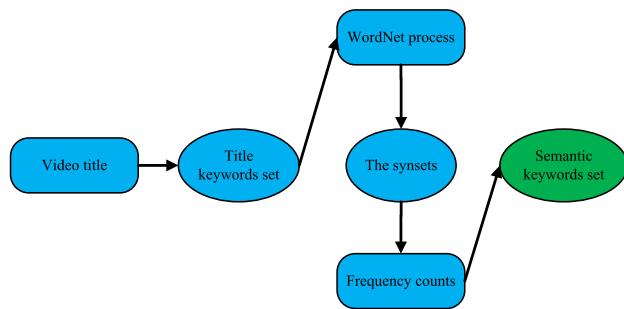
A. GENERATING CANDIDATE KEYWORDS VIA WORDNET

WordNet is an English lexical database, which divides English words into groups of synonyms (synsets), each of them expressing a concept [18], [20]. WordNet provides a brief definition and usage examples for each synset, and records the semantic relations between different synsets. Through the API of WordNet, it is able to conduct automatic search of the words that are semantically relevant to a given word.

The process of expanding candidate keywords using WordNet can be divided into the following three steps, as shown in Figure 3:

(1) Given the title and tags of a video initialized by an uploader, the first step is to extract the original keyword set. We consider simple filters: remove the stop-words and punctuation, detect named entities.

(2) For each original keyword, query the WordNet to obtain the synsets that are associated with the keyword.

**FIGURE 3.** Based on WordNet extension.

The noun family has 8 senses (first 7 from tagged texts)

1. (388) **family**, household, house, home, menage -- (a social unit living together; "he moved his family to Virginia"; "It was a good Christian household"; "I waited until the whole house was asleep"; "the teacher asked how many people made up his home")
2. (381) **family**, family unit -- (primary social group; parents and children; "he wanted to have a good job before starting a family")
3. (110) **family**, family line, folk, kinfolk, kinsfolk, sept, phratry -- (people descended from a common ancestor; "his family has lived in Massachusetts since the Mayflower")
4. (22) class, category, **family** -- (a collection of things sharing a common attribute; "there are two classes of detergents")
5. (18) **family**, fellowship -- (an association of people who share common beliefs or activities; "the message was addressed not just to employees but to every member of the company family"; "the church welcomed new members into its fellowship")
6. (7) **family** -- ((biology) a taxonomic group containing one or more genera; "sharks belong to the fish family")
7. (4) kin, kinsperson, **family** -- (a person having kinship with another or others; "he's kin"; "he's family")
8. syndicate, crime syndicate, mob, **family** -- (a loose affiliation of gangsters in charge of organized criminal activities)

FIGURE 4. Query the keyword of “family” in WordNet.

(3) The frequency count indicates the relevance strength between synsets and the given word. In order to obtain the most relevant keywords for the given word, we select the synset which has the highest frequency count and extract 2-3 keywords from it as the candidate keywords.

For example, Figure 4 shows the result of querying the WordNet with the word “family”. As can be seen from the figure, the keyword of “family” has 8 synsets. The number in parentheses is the frequency count. We select the keywords according to the frequency counts displayed in the front. Since the first synset has the highest frequency counts, selecting the first synset as the expansion keywords of “family”, three keywords (“household”, “house”, “home”) are selected here.

Since WordNet has already designed the semantic relations between the various synsets, the study can use WordNet to expand 2 or 3 keywords for each keyword in the title and tags of the original video. In this way, while extending the video keywords, the semantic relevance of the candidate keywords to the original video title and tags is ensured, and the diversity of the keywords is improved as well.

B. GENERATING CANDIDATE KEYWORDS VIA MAINSTREAM VIDEO SHARING WEBSITES

With the rapid development of the network technology and multimedia technology, more and more video sharing sites

appear on the Internet. Billions of videos are hosted in various video sharing platforms and the videos cover an extremely wide range of topics. Usually, there are repetitive or similar videos on the video sharing sites that are relevant to the video goes to be uploaded. Thus, it is promising to collect relevant and popular textual information from videos that are of the same or similar topics on video sharing sites.

In this section, we mainly discuss the problem of extending the semantic candidate keywords through the mainstream video sharing websites. As reported in previous studies [21], [36], search engines and recommendation systems are the two most important tools that help viewers discover videos of their interests. Furthermore, it has been shown that the title of a video is the most important data used by search engine to retrieve the video for viewers. The title of a video is also the main factor that impacts the initialization of a recommendation link induced by a recommendation system. Thus, we use the search engines and recommendation systems of the mainstream video websites to retrieve the videos that are relevant to the video that will be uploaded and then extract keywords from their titles as the new expansion of the semantic candidate keywords. In this way, we can not only guarantee the diversity but also the popularity of the semantic candidate keywords.

According to the above analysis, the process of expanding candidate keywords based on the mainstream video sharing website can be divided into the following two steps:

(1) We create 2 or 3 groups of keywords using the original video title keywords and the candidate keywords obtained via WordNet. One group consists of the original video title keywords. The other one or two groups are formed by replacing the keywords in the original title with the corresponding candidate keywords. Then, each group of keywords is fed to the search engines of the three mainstream video sharing sites and the top 10 result videos are gathered from each site. Finally, we extract the keywords from titles of these videos and add them to the semantic candidate keywords collection. For example, search the group of keywords “Deep Dream VR”, the result pages of the three major websites are shown in Figure 5.

As can be seen from the figure, the mainstream video sharing sites used in the paper are YouTube, Yahoo Video, and Bing Video. YouTube is the world’s largest video sharing site, including a variety of types and categories of video; Yahoo Video contains videos from a number of professional video sharing sites, including Hulu, Cable News Network, Fox, Dailymotion and so on. Bing Video contains nearly 10 large video sharing sites in the country, including Youku, Ku6, Tudou and so on. All of these video sharing sites are equipped with powerful video search engines and recommendation systems, which make the discovery of interesting videos much more convenient.

- 1) For each top video obtained through search engine in the first step, we gather the related videos recommended by recommendation system. Then, we collect

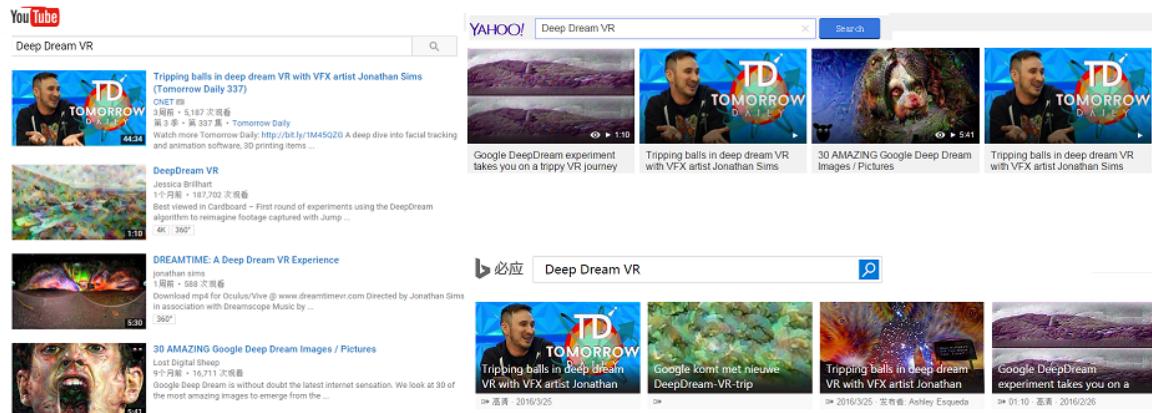


FIGURE 5. Search “Deep Dream VR” in three video sharing sites.

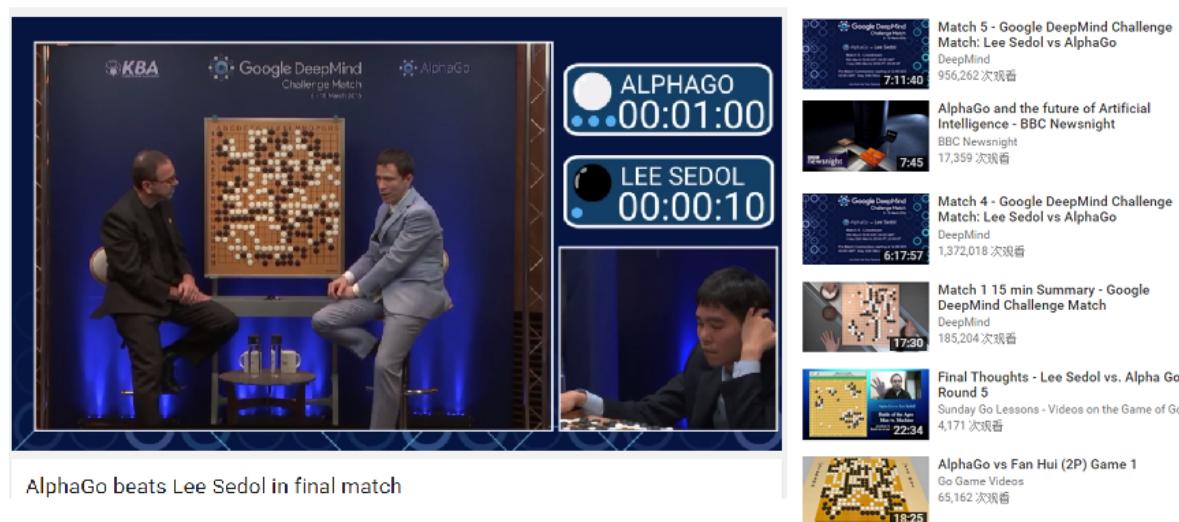


FIGURE 6. Related video list.

the title keywords of these related videos and add into the semantic candidate keywords collection. Figure 6 shows an example of related video list.

Through the above two steps, we make full use of the capabilities of search engine and recommendation system to generate the semantic candidate keywords that are not only relevant to the video topic, but also are of a high degree of social popularity.

Based on all of the above analysis, it is reasonable to argue that the semantic candidate keywords gathered using the proposed method have a high relevance with the video that will be uploaded, and have a high popularity and diversity as well.

C. GENERATING CANDIDATE KEYWORDS BASED ON VIDEO CONTENT RECOGNITION

This section mainly discusses the problem of expanding the entity candidate keywords through deep learning technology. In the process of uploading a video to a video sharing website, the uploader not only provides the textual information, but

also provides the video itself. Regarding to the fact that video content itself is the most important information on which the viewers pay their most attention, we need to obtain relevant keywords by directly recognizing the video content. To the best of our knowledge, the deep learning is the state of the art technology in video content recognition. Thus, in our paper, we use the most widely used image recognition framework, Caffe deep leaning framework, for recognizing video content. The Caffe framework takes images as input, and then outputs the recognized entity names of the images. The entity names of a video are named as the entity candidate keywords in our paper. It is reasonable to believe that the entity candidate keywords obtained as proposed have a high relevance with the video content.

Machine learning technologies have been applied in some studies concerned on image label distribution and refinement [22]–[24]. Deep learning is a relatively new machine learning technology, which always yields much better results than traditional methods in recognizing image content as reported. A video consists of a series of images, the entity

names of the video can be extracted by recognizing the images of the video. The image recognition technology used in our study is Caffe deep learning framework [19], [25], [26], which currently is a mainstream framework in the academic community. The training dataset for the Caffe framework used in our work is the ImageNet ILSVRC-2012 dataset [27], [28]. The training dataset consists of 1,200,000 images from more than 1000 categories. The large number of categories and images ensures a good comprehensiveness of the dataset. Here, we need to point out that our work is not specialized to study deep learning and Caffe framework, but just utilize them to extract the entity information of a video.

The process of video content recognition is divided into two steps, as shown in Figure 7:

(1) As shown in the left part of the figure, we intercept 15 video keyframes equably according to the length of the video. For example, for a one minute video, we intercept a keyframe every 4 seconds.

(2) We input the video's keyframes into the trained Caffe framework and it outputs the entity names of the keyframes. We then add them to the collection of video entity candidate keywords.

Through recognizing the video content with deep learning technology, the most appropriate entity information of video content are obtained. To a considerable extent, the work proposed in this section improves the relevance between the keywords and the video content, and makes the candidate keywords more diverse and comprehensive.

D. RANKING KEYWORDS WITH TF-SIM ALGORITHM

Besides generating a good collection of candidate keywords, another key point of our work is how to rank the candidate keywords and recommend the final keyword set. As to this problem, a novel method that combines the frequency count and semantic similarity of keywords is proposed in this paper.

Research results concerned on the relevance between images show that the more times an image label appears in the adjacent images [9], the higher degree of relevance between the label and the image. Similarly, it is reasonable to believe that the more times a keyword occurred in the keywords set indicate a higher relevance of the keyword to the video content.

Recently, some achievements have been made in the measurement of the semantic similarity of keywords [29]–[32]. Our study adopts the method of NGD based on the two following reasons: firstly, NGD relies on the Google search engine which is able to search for comprehensive information; secondly, NGD is suitable to calculate the correlation between the candidate keywords and the original title keywords. Specifically, the Normalized Google Distance between two search keywords x and y is shown in the Equation 1

$$NGD(x, y) = \frac{\max\{\log h(x), \log h(y)\} - \log h(x, y)}{\log N - \min\{\log h(x), \log h(y)\}} \quad (1)$$

where, $h(x)$ and $h(y)$ are the number of hits that uses the Google engine to search the two keywords x and y , respectively; $h(x, y)$ is the number of web pages on which both x and y occur; N is the total number of web pages could be searched by Google. If the distance value is close to 0, it indicates that the two keywords are highly related; on the other hand, if the distance value is close to infinity, it means the two keywords are totally unrelated.

According to the above analysis, combined with TF-IDF rationale that if a word or phrase frequently appears in an article, but rarely appears in other articles, it is regarded that the word or phrase has a good representativeness to distinguish between categories [33]–[35]. TF-IDF is defined as:

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (2)$$

$$TF(t, d) = \frac{f_{t,d}}{\sum_{w \in d} f_{w,d}} \quad (3)$$

$$IDF(t, D) = \log \frac{N}{1 + |\{d \in D : t \in d\}|} \quad (4)$$

where $f_{t,d}$ represents the number of times that term t occurs in document d , $\sum_{w \in d} f_{w,d}$ indicates the total times of all of terms in document d , D is the set of documents in the corpus, N indicates the total number of documents in the corpus $N = |D|$, $|\{d \in D : t \in d\}|$ is the number of documents where term t appears. The denominator is adjusted to $1 + |\{d \in D : t \in d\}|$ in order to avoid a division-by-zero.

As to our problem, we also values the frequency count of a keyword shown up in the collection of candidate keywords. Nevertheless, measuring the degree of correlation solely based on frequency count is not reasonable enough. Moreover, it is not appropriate to use TF-IDF directly, since we are not measuring the representativeness of a keyword in the collection of candidate keywords but the appropriateness of a keyword to be a part of the title and tags of a video. Therefore, we have to propose a novel ranking method that values both the frequency count of a candidate keyword and its semantic similarity with original keywords.

This paper proposes a novel algorithm by the combination of the frequency of keyword and the similarity between the keywords, named the TF-SIM algorithm, as shown in the Equation 5.

$$TF - SIM(t, x) = T_t \times e^{-\frac{1}{n} \sum_{i=1}^n NGD(t, x_i)} \quad (5)$$

where, T_t represents the frequency of the t keyword, x represents the initial title keyword set, and n represents the number of keywords in the original title keyword set.

As mentioned in the previous sections, there are two collections of candidate keywords: video semantic keywords collection and video entity keywords collection. The method is designed to recommend to the uploader a number of top keywords except for the initial title keywords. One more issue needs to be solved is to determine the proportion of keywords

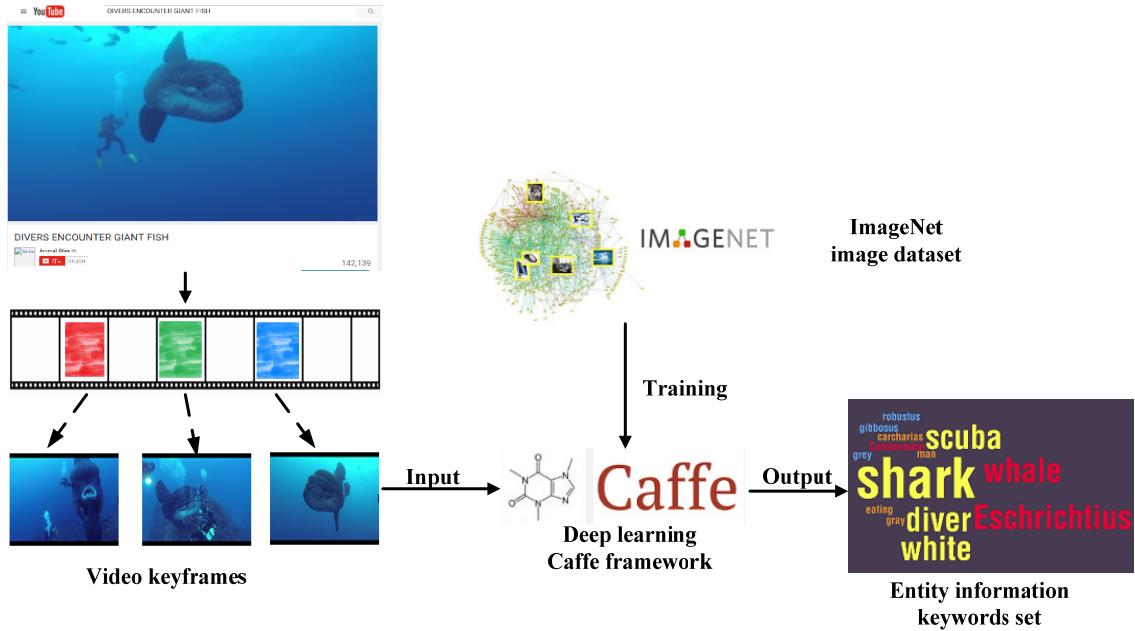


FIGURE 7. Video entity information extraction.

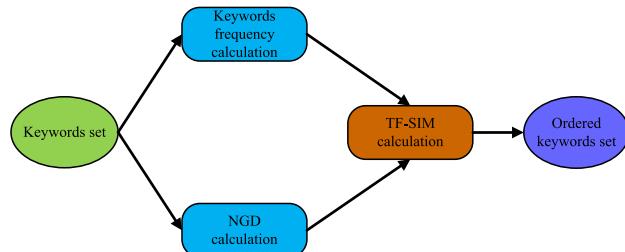


FIGURE 8. Keywords set correlation ranking.

from each collection. The proposed solution is shown in the Equation 6.

$$T_n = T_s + \delta T_s \quad (6)$$

where, T_n represents the number of keywords that needs to recommend to the users, T_s represents keyword count recommended from the semantic keywords collection, δT_s represents the number of keywords extracted from the entity keywords collection, δ could be set to 0.5 based on experience.

As shown in Figure 8, the process of ranking the keywords collection is divided into the following four steps:

(1) For each candidate keyword, we obtain its number of occurrences in the candidate keyword collection.

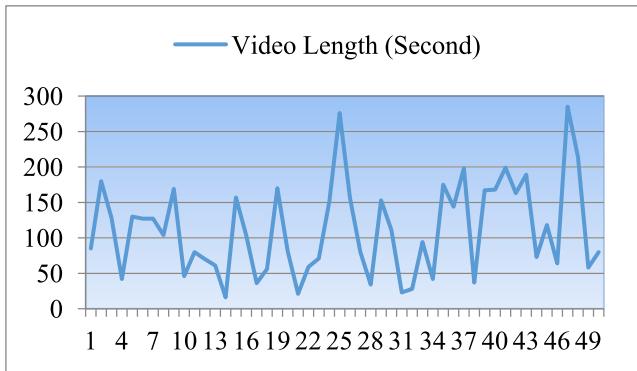
(2) Calculate the average similarity between each candidate keyword and the original title keywords using NGD formula.

(3) Calculate the TF-SIM score for each candidate keyword using its frequency count and average similarity in TF-SIM formula.

(4) Rank each collection of candidate keywords based on TF-SIM score in a descending order.

After ranking the candidate keywords, the set of recommended keywords can be finalized once the total number of keywords and the proportion of keywords from each collection are determined. For example, if the method is assigned to recommend top 15 keywords to users and the proportion parameter δ is assigned to 0.5. That is, the proposed method recommends the top10 keywords from the semantic keywords set and suggests top 5 keywords from the entity keywords set. Eventually, the method recommends top15 keywords for an uploader to optimize the title and tags of his/her video.

From the introduction of our method, it is not difficult to find that our method recommends keywords with comprehensive consideration of relevance, diversity, and popularity. The relevance of recommended keywords with a video is the most seriously followed principle. In each step of the method, including generating candidate keywords via WordNet, mainstream websites and video content recognition, and ranking the keywords with TM-SIM, the relevance is the first priority to be fulfilled. The diversity of recommended keywords is ensured by the fact that our candidate keywords are generated from three diverse and independent ways. The popularity of recommended keywords is mainly addressed in the following two aspects. Firstly, candidate keywords obtained from mainstream video sites are usually of a relatively high degree of popularity. Secondly, TM-SIM algorithm considers the frequent count of the keyword appears in the candidate collection. The more times the keywords appear in the collection, the greater the potential that attract the social attention, so as to ensure the final keywords to the uploader has very high popularity. In summary, our method is able to

**FIGURE 9.** Video length.

recommend relevant, diverse, popular keywords for uploaders to optimize titles and tags of their videos. The optimization of textual information of a video is expected to help viewers discover videos of their interests on one hand, and meanwhile to improve the popularity of the video on the other hand.

V. EXPERIMENTS

The proposed method of the paper can be used for optimizing the title and tags of a video, which can then improve the video's popularity and viewing time. In order to better evaluate the performance of our method, the following experiments are conducted: professional website evaluation, user judgement and YouTube experiment.

A. DATA SET

The videos in the experimental dataset used in our paper were collected from YouTube. The dataset consists of 50 videos, which were randomly picked from different video categories including film and animation, pets and animals, sports, travel and events, gaming, comedy, entertainment, news and politics, science and technology. Furthermore, the original title and the original tags of 50 videos were also collected. As can be seen from Figure 9, the experimental videos are mostly short, that is, in the range of five minutes. This matches the statistics reported in previous work that most of the user generated videos sharing on YouTube are less than five minutes. In contrast to a long video, a short video usually has higher popularity and has higher similarity between the adjacent video keyframes given the number of intercepted keyframes is fixed.

From the figure, we can see that most videos have a length shorter than 200 seconds, only 2 videos are longer than 200 seconds. The average length of the 50 videos is 112 seconds, which belong to the category of short video.

B. PROFESSIONAL WEBSITE EVALUATION

The performance of our method is firstly evaluated on a professional survey website Amazon Mechanical Turk (AMT). The turkers (the person who accept to participate in the survey) of AMT are required to fill out our questionnaire. The detailed description of the questionnaire is: You are given

a video, the initial title, description, tags of the video, and 15 keywords recommended by our algorithm for the video as well. You are invited to evaluate the relevance of the 15 keywords to the video. Among them, which keyword(s) will you choose to improve and optimize the quality of initial video title, description and tags for attracting more audience? Please choose them. If you feel all keywords are appropriate, do not hesitate to choose all of them.

In order to reflect the diversity of the survey, we set up 4 copies of questionnaires for each video. That is, 200 copies of questionnaires for 50 videos in total. The questionnaires are finished by 200 different turkers. The cost of each questionnaire is 0.2\$ and a total of 50\$ (plus tax) for 50 videos. Based on the results of questionnaires, the average number of selected keywords per video is calculated and shown in Figure 10.

As shown in Figure 10, the number of selected keywords per video is mostly between 6 and 9. There are 14 videos that have the number of selected keywords larger than 9. Through calculation, the average number of selected keywords per video is 8.18, accounting for 54.53% of the 15 recommended keywords. In other words, given a set of recommended keywords, more than half of them will be selected by the user.

In order to evaluate the performance of the proposed ranking algorithm, we analyze the percentage of selected keywords that are from the top 10 keywords. The results are shown in Figure 11.

As can be seen from Figure 11, the percentage of selected keywords that are from the top 10 keywords is relatively high. For almost all of the cases, the number is higher than 70%, and the number even reaches 100% for three of the videos. Through calculation, it is found that 80% of the selected keywords are from the top 10 keywords on average. It indicates that our ranking algorithm has a good ability to rank the most appropriate keywords in the front of the list of top 10 keywords.

Recognizing the high percentage of selected keywords that are from the top 10 keywords, we analyze the performance of our method if only the top 10 keywords are recommended to the user. Figure 12 shows the number of selected keywords by users from the top 10 keywords.

The average number of selected keywords is 6.42, in other words, 64.2% of the top ten keywords will be chosen by users. The percentage is increased by 10 points in the scenario of top 10 keywords, comparing to that of top 15 keywords. The high probability that the majority of top10 keywords will be chosen by users to optimize titles and tags of their videos strongly indicates the good performance of the proposed method.

C. USER SUBJECTIVE EVALUATION

In user subjective evaluation scenario, volunteers were requested to subjectively evaluate whether the optimized title and tags are better than the original title and tags of a video. There are two criteria for evaluating title and tags, which are relevance and diversity. Here, we use the method of rating the

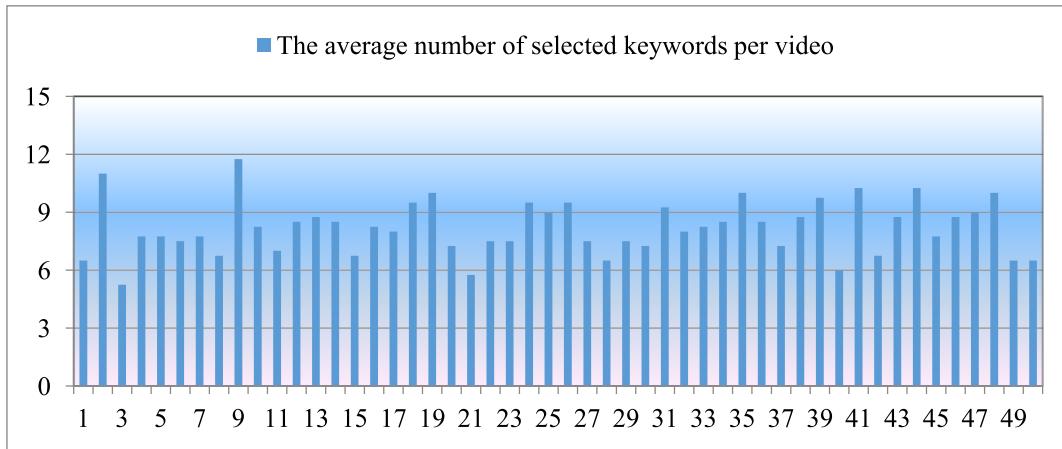


FIGURE 10. The average number of selected keywords per video.

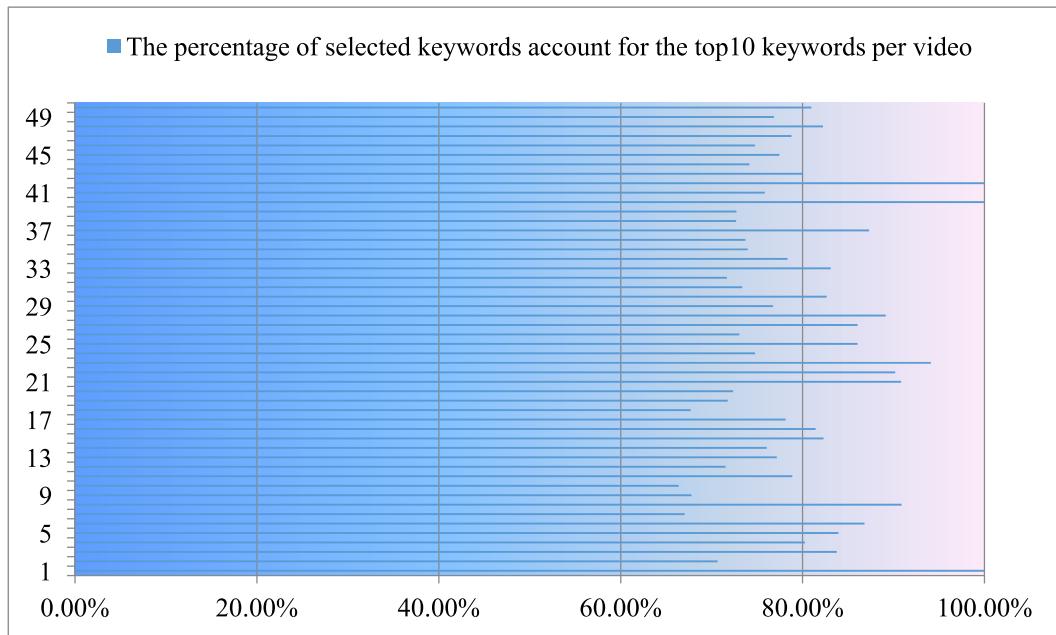


FIGURE 11. The percentage of selected keywords account for the top 10 keywords per video.

score from volunteer to assess the extent of the optimization of the video title and tags.

In order to prevent some extreme cases about rating the score from volunteers, we set up 4 copies of the same questionnaire for each video, a total of 200 questionnaires were assigned to 28 volunteers. Video content of the questionnaire assigned to each volunteer is different. The scoring system is based on 10 points.

In order to reduce the differences in subjective scales of volunteers, a detailed description of our questionnaire is like this: if you score the optimized title and tags with 6 points or above, it means that you feel that the optimized title and tags are better than the original title and tags of a video. The higher the score, the more surpasses of optimized title and tags over the original title and tags. User's subjective points for the 50 videos are shown in Figure 13.

As can be seen from Figure 13, the points of all videos are more than 6 points, and most of them are around 8 points. Through further calculation, we find that the average point of the 50 videos is 8.1. The number of videos that have a score higher than 8 is 32, among them, 3 videos have a score larger than 9 points, with the highest score of 9.3 points. The results indicate that the keywords recommended by our method are better than the original title and tags in terms of relevance and diversity.

D. YouTube EXPERIMENT

YouTube experiment is designed to investigate whether the video associated with optimized title and tags can really attract more social popularity and extend the viewing time in real environment. Therefore, the criteria for evaluating the

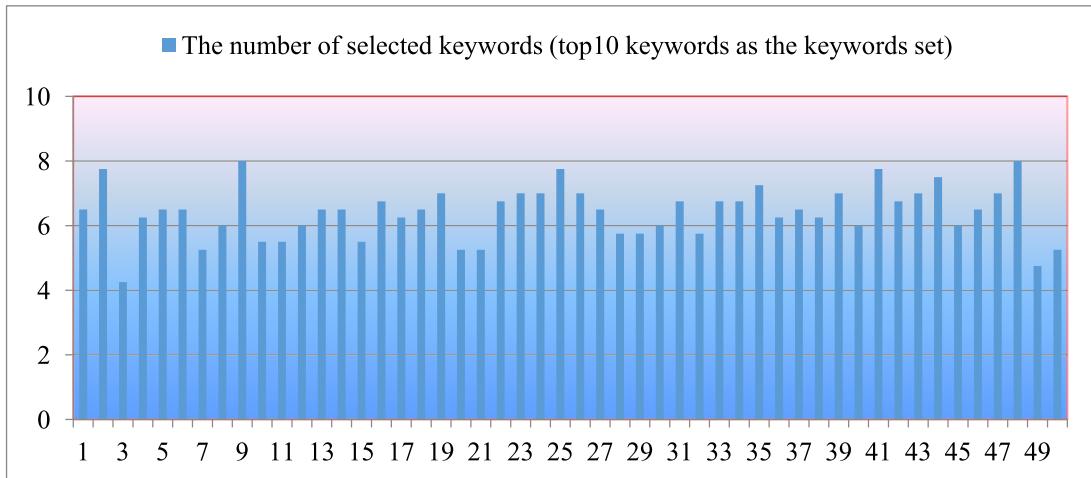


FIGURE 12. The number of selected keywords from top10 keywords.

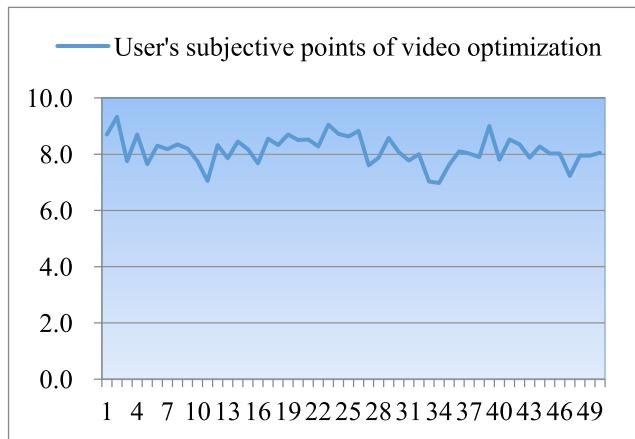


FIGURE 13. User's subjective points of video optimization.

performance of our method here is the view count and the viewing time of a video.

In order to carry out the experiment, we registered four new accounts on YouTube without any history of browsing, uploading video and so on. The first account hosts the 50 videos with original title and tags, which is called the original scenario. The second account accommodates the 50 videos with the title and tags optimized using the keywords recommended by TF-SIM algorithm, which is named TF-SIM optimized scenario. Furthermore, we also compare the performance of the proposed methods (TF-SIM) with other two ranking algorithms NGD and KNN. For each of the two algorithms NGD and KNN, we registered a new account for hosting 50 videos with keywords recommended by the corresponding algorithm. For these four accounts, the uploaded 50 videos, the description, the classification and other information of the videos are exactly the same. The only difference for the four accounts is the title and tags of the video.

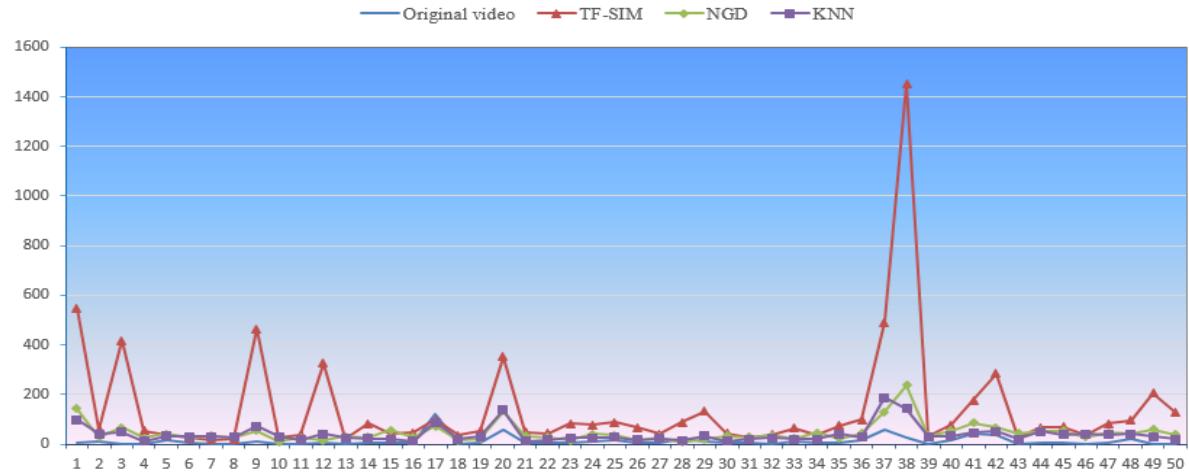
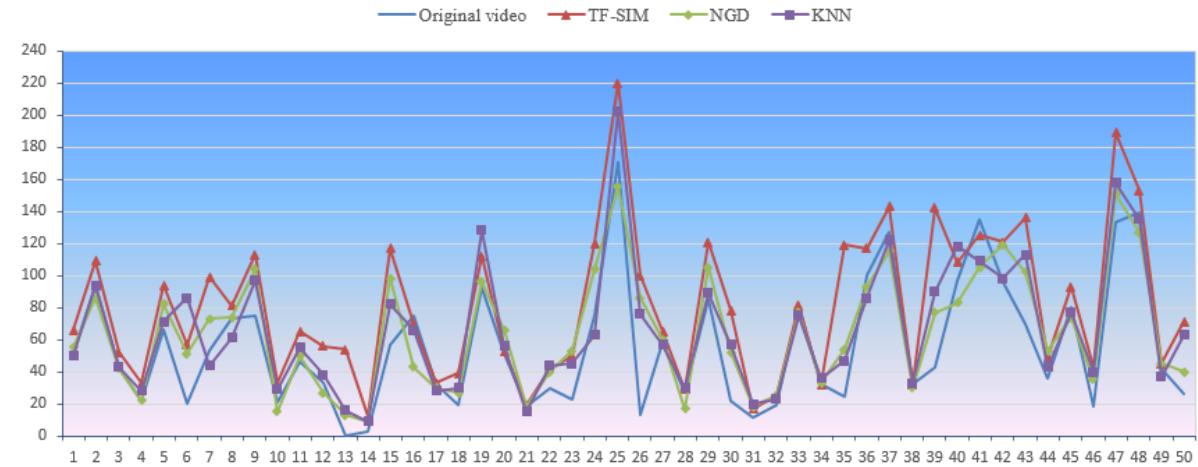
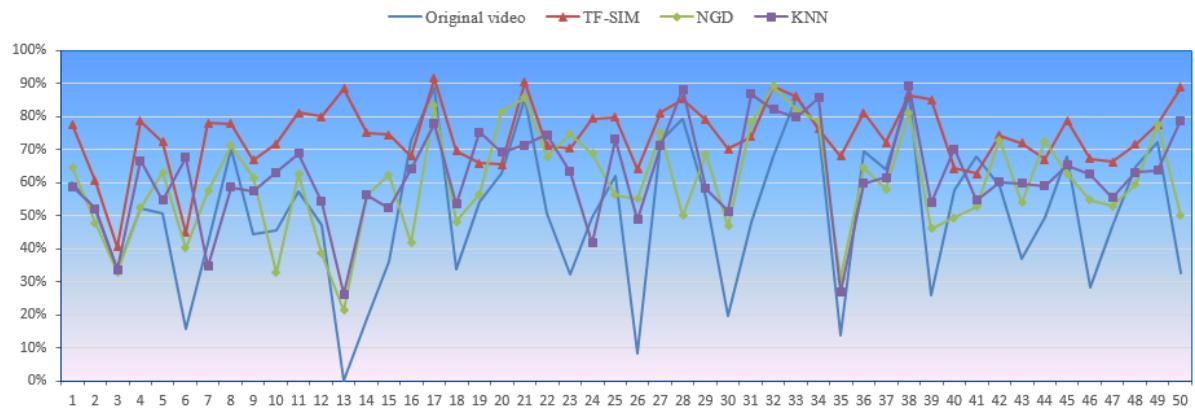
The experimental videos were uploaded to the above four different accounts. Three months later, we collected the view count and viewing time of each video hosted in the four

accounts. Figure 14 shows the view count of videos in the four accounts.

As shown in Figure 14, the view counts of videos optimized by the three algorithms are significantly higher than that of the original scenario in general. After analyzing the experiment results, we find that the view count of original scenario ranges from 0 to 122. The view count of TF-SIM optimized scenario is from 18 to 1453. In the figure, it is easy to find that most of the videos in original scenario gain less than 10 views, some even as low as 1 to 2 views. The videos optimized by TF-SIM algorithm mostly receive about 60 views. We also calculated the average view count for each scenario, and the results are 12, 141, 46, 40 for original scenario, the TF-SIM algorithm, the NGD algorithm and KNN algorithm optimized scenarios, respectively. That is to say, the view count of videos optimized with TF-SIM algorithm is 3.07 times, 3.52 times, 11.75 times as many as that of NGD algorithm, KNN algorithm and original video, respectively. So in terms of the view count, the videos optimized using TF-SIM algorithm are able to attract more social attention than the original videos and the NGD and KNN algorithm optimized videos.

In order to evaluate the performance of our method in recommending appropriate keywords more comprehensively, we also analyze another two metrics: the video's average viewing time and the percentage that video viewing time accounted for the total length of a video. These two metrics are appropriate to evaluate whether our method improves the popularity of a video by helping the video expose to viewers who are of interests. The average viewing time of the video is shown in Figure 15 and the percentage of video viewing time is shown in Figure 16.

As shown in Figure 15, the average viewing time of each optimized scenarios is longer than that of the original scenario. The average viewing time of the 50 videos in the original, TF-SIM, NGD and KNN optimized scenarios are 55.76, 80.66, 64.18 and 66.24 seconds, respectively. That is, the average viewing time of

**FIGURE 14.** The view count of videos.**FIGURE 15.** The average viewing time of the video.**FIGURE 16.** The percentage of video viewing time.

videos optimized with TF-SIM algorithm is 1.26 times, 1.21 times, 1.45 times of that of NGD, KNN optimized and original videos, respectively. Therefore, the videos optimized with TF-SIM algorithm can get longer viewing time than the original, NGD and KNN optimized videos.

As can be seen from Figure 16, the fluctuations of the percentage of length of the original videos to watch is relatively large, while that of the TF-SIM optimized videos is relatively stable, basically in about 75%. Overall, the percentage of the length of the viewing time of each optimized scenarios is usually higher than that of the original scenario. The average

percentage of video viewing time accounting for the length of videos in the original, TF-SIM, NGD, KNN optimized scenarios are 49.79%, 74.19%, 59.85%, and 62.11%, respectively. That is to say, the percentage of viewing time of videos optimized with TF-SIM algorithm surpasses NGD, KNN and the original scenario by 14.34, 12.08, and 24.4 points, respectively. By analyzing the length and the percentage of video viewing time, it is further confirmed that the videos optimized with TF-SIM algorithm can get more attention than the videos in the original, NGD and KNN algorithm optimized scenarios.

Through investigating the three types of experimental data gathered in YouTube platform, which are view count, viewing time, and the percentage of viewing time, it is demonstrated that the videos optimized with TF-SIM algorithm can attract higher social popularity and longer viewing time than videos in the original, NGD and KNN algorithm optimized scenarios. All of the results further demonstrate the effectiveness of the method proposed in our paper.

VI. CONCLUSION

Social media platforms such as YouTube, Facebook, offer great opportunities for people to entertain, interact, and advertise. Big data analytics is being developed to help researchers and marketers understand interests of users for gaining higher popularity and generate more revenue through social media. In this paper, we propose a novel hybrid tagging method based on multi-modal content analysis, that is, textual semantic analysis and deep learning driven video content recognition for boosting video popularity. The method is divided into three parts. Firstly, generating semantic candidate keywords based on textual semantic analysis. More superficially, we use WordNet and three mainstream video sharing sites (YouTube, Yahoo Video, Bing Video) to generate candidate keywords that are semantically relevant to the keywords obtained from the initial video title. Secondly, we explore the Caffe deep learning framework to recognize the key-frames of the video and then identify the entities in the video image as entity keywords set. Thirdly, we propose the TF-SIM algorithm to rank the candidate keywords. The algorithm comprehensively considers the frequency count of a candidate keyword and the semantic similarity between a candidate keyword and original keywords as well.

The proposed method is evaluated in three scenarios: professional survey website Amazon Mechanical Turk (AMT), campus volunteers and YouTube website, respectively. It is demonstrated that the video title and tags optimized by our method can attract more popularity and extend longer viewing time per playback. This result indicates that the proposed method based on textual semantic analysis and video content recognition is effective to recommend appropriate keywords for video uploaders to optimize the title and tags of their videos.

Further research will be carried out to generate complete title and description of a video by integrating natural language processing and video semantic analysis technologies, which

will provide more convenience for generators to upload their videos and also help viewers discover videos of their interests.

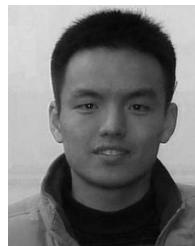
ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments, and also to the editors for their efforts that improved this paper.

REFERENCES

- [1] J. Davidson, "The YouTube video recommendation system," in *Proc. 4th ACM Conf. Rec. Syst.*, Barcelona, Spain, 2010, pp. 293–296.
- [2] Y. Yin, L. Chen, Y. Xu, J. Wan, H. Zhang, and Z. Mai, "QoS prediction for service recommendation with deep feature learning in edge computing environment," *Mobile Netw. Appl.*, to be published, doi: [10.1007/s11036-019-01241-7](https://doi.org/10.1007/s11036-019-01241-7).
- [3] G. Feng, Y. Li, Q. Zhao, H. Wang, H. Lv, and J. Lin, "Optimizing broadcast duration for layered video streams in cellular networks," *Peer-Peer Netw. Appl.*, vol. 10, no. 3, pp. 765–779, May 2017.
- [4] W. Zhang, W. Liu, T. Wang, A. Liu, Z. Zeng, H. Song, and S. Zhang, "Adaption resizing communication buffer to maximize lifetime and reduce delay for WVSNs," *IEEE Access*, vol. 7, pp. 48266–48287, 2019.
- [5] Y. Yin, W. Zhang, Y. Xu, H. Zhang, Z. Mai, and L. Yu, "QoS prediction for mobile edge service recommendation with auto-encoder," *IEEE Access*, vol. 7, pp. 62312–62324, 2019.
- [6] J. Tan, W. Liu, M. Xie, H. Song, A. Liu, M. Zhao, and G. Zhang, "A low redundancy data collection scheme to maximize lifetime using matrix completion technique," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, Jan. 2019, Art. no. 5.
- [7] X. Liu, A. Liu, T. Wang, K. Ota, M. Dong, Y. Liu, and Z. Cai, "Adaptive data and verified message disjoint security routing for gathering big data in energy harvesting networks," *J. Parallel Distrib. Comput.*, vol. 135, pp. 140–155, Jan. 2020.
- [8] M. Huang, W. Liu, T. Wang, A. Liu, and S. Zhang, "A cloud-MEC collaborative task offloading scheme with service orchestration," *IEEE Internet Things J.*, to be published, doi: [10.1109/JIOT.2019.2952767](https://doi.org/10.1109/JIOT.2019.2952767).
- [9] X. Li, T. Uricchio, L. Ballan, M. Bertini, C. G. Snoek, and A. D. Bimbo, "Image tag assignment, refinement and retrieval," in *Proc. 23rd ACM Int. Conf. Multimedia*, Brisbane, Australia, 2015, pp. 1325–1326.
- [10] M. Tan, J. Yu, Z. Yu, F. Gao, Y. Rui, and D. Tao, "User-click-data-based fine-grained image recognition via weakly supervised metric learning," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 14, no. 70, pp. 1–23, 2019.
- [11] J. Yu, X. Yang, F. Gao, and D. Tao, "Deep multimodal distance metric learning using click constraints for image ranking," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4014–4024, Dec. 2017.
- [12] J. Yu, J. Li, Z. Yu, and Q. M. Huang, "Multimodal transformer with multi-view visual representation for image captioning," *IEEE Trans. Circuits Syst. Video Technol.*, to be published, doi: [10.1109/TCSVT.2019.2947482](https://doi.org/10.1109/TCSVT.2019.2947482).
- [13] L. Ballan, M. Bertini, G. Serra, and A. D. Bimbo, "A data-driven approach for tag refinement and localization in Web videos," *Comput. Vis. Image Understand.*, vol. 140, pp. 58–67, Nov. 2015.
- [14] L. Huang and B. Luo, "Tag refinement of micro-videos by learning from multiple data sources," *Multimedia Tools Appl.*, vol. 76, no. 19, pp. 20341–20358, Oct. 2017.
- [15] R. Shetty, H. R. Tavakoli, and J. Laaksonen, "Image and video captioning with augmented neural architectures," *IEEE Multimedia Mag.*, vol. 25, no. 2, pp. 34–46, Apr./Jun. 2018.
- [16] E. Santos-Neto, T. Pontes, J. M. Almeida, and M. Ripeanu, "Towards boosting video popularity via tag selection," in *Proc. SoMuS ICMR*, 2014, pp. 1–6.
- [17] E. Santos-Neto, T. Pontes, J. Almeida, and M. Ripeanu, "On the choice of data sources to improve content discoverability via textual feature optimization," in *Proc. 25th ACM Conf. Hypertext Social Media*, Santiago, Chile, 2014, pp. 273–278.
- [18] N. K. Jha, A. Jethva, N. Parmar, and A. Patil, "A review paper on deep Web data extraction using WordNet," *Int. Res. J. Eng. Technol.*, vol. 3, no. 3, pp. 1003–1006, Mar. 2016.
- [19] Y. Jia, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, Orlando, FL, USA, 2014, pp. 675–678.

- [20] H. Redkar, S. Bhingardive, D. Kanojia, and P. Bhattacharyya, "World WordNet database structure: An efficient schema for storing information of WordNets of the world," in *Proc. AAAI*, Mar. 2015, pp. 4290–4291.
- [21] R. Zhou, S. Khemmarat, L. Gao, J. Wan, J. Zhang, Y. Yin, and J. Yu, "Boosting video popularity through keyword suggestion and recommendation systems," *Neurocomputing*, vol. 205, pp. 529–541, Set. 2016.
- [22] Y. Yin, J. Xia, Y. Li, Y. Xu, W. Xu, and L. Yu, "Group-wise itinerary planning in temporary mobile social network," *IEEE Access*, vol. 7, pp. 83682–83693, 2019.
- [23] L. Chen, D. Xu, I. W. Tsang, and J. Luo, "Tag-based image retrieval improved by augmented features and group-based refinement," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 1057–1067, Aug. 2012.
- [24] X. Li and C. G. Snoek, "Classifying tag relevance with relevant positive and negative examples," in *Proc. 21st ACM Int. Conf. Multimedia*, Barcelona, Spain, 2013, pp. 485–488.
- [25] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, "cuDNN: Efficient primitives for deep learning," 2014, *arXiv:1410.0759*. [Online]. Available: <https://arxiv.org/abs/1410.0759>
- [26] J. Yu, B. Zhang, Z. Kuang, D. Lin, and J. Fan, "iPrivacy: Image privacy protection by identifying sensitive objects via deep multi-task learning," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 5, pp. 1005–1016, May 2017.
- [27] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [28] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, "XNOR-Net: ImageNet classification using binary convolutional neural networks," in *Comput. Vis.—ECCV*. Amsterdam, The Netherlands: Springer, 2016, pp. 525–542.
- [29] R. L. Cilibrasi and P. M. B. Vitanyi, "The Google similarity distance," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 3, pp. 370–383, Mar. 2007.
- [30] Y. Yin, L. Chen, Y. Xu, and J. Wan, "Location-aware service recommendation with enhanced probabilistic matrix factorization," *IEEE Access*, vol. 6, pp. 62815–62825, 2018.
- [31] E. Agirre, C. Banea, D. Cer, M. Diab, A. Gonzalez-Agirre, R. Mihalcea, G. Rigau, and J. Wiebe, "SemEval-2016 task 1: Semantic textual similarity, monolingual and cross-lingual evaluation," in *Proc. 10th Int. Workshop Semantic Eval. (SemEval)*, San Diego, CA, USA, 2016, pp. 497–511.
- [32] S. Kim, W. J. Wilbur, and Z. Lu, "Bridging the gap: A semantic similarity measure between queries and documents," 2016, *arXiv:1608.01972*. [Online]. Available: <https://arxiv.org/abs/1608.01972>
- [33] J. H. Paik, "A novel TF-IDF weighting scheme for effective ranking," in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Dublin, Ireland, 2013, pp. 343–352.
- [34] S. Plangsangket and J. Q. Gan, "A query suggestion method combining TF-IDF and Jaccard coefficient for interactive Web search," *Artif. Intell. Res.*, vol. 4, no. 2, pp. 119–125, 2015.
- [35] S. M. H. Dadgar, M. S. Araghi, and M. M. Farahani, "A novel text mining approach based on TF-IDF and support vector machine for news classification," in *Proc. IEEE Int. Conf. Eng. Technol. (ICETECH)*, Coimbatore, India, Mar. 2016, pp. 112–116.
- [36] R. Zhou, S. Khemmarat, L. Gao, J. Wan, and J. Zhang, "How YouTube videos are discovered and its impact on video views," *Multimedia Tools Appl.*, vol. 75, no. 10, pp. 6035–6058, May 2016.
- [37] X. Yu, Y. Chu, F. Jiang, Y. Guo, and D. Gong, "SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features," *Knowl.-Based Syst.*, vol. 141, pp. 80–91, Feb. 2018.
- [38] Y. Liu, A. Liu, X. Liu, and X. Huang, "A statistical approach to participant selection in location-based social networks for offline event marketing," *Inf. Sci.*, vol. 480, pp. 90–108, Apr. 2019.
- [39] X. Yu, F. Jiang, J. Du, and D. Gong, "A cross-domain collaborative filtering algorithm with expanding user and item features via the latent factor space of auxiliary domains," *Pattern Recognit.*, vol. 94, pp. 96–109, Oct. 2019.
- [40] J. B. Schmitt, D. Rieger, O. Rutkowski, and J. Ernst, "Counter-messages as prevention or promotion of extremism?! The potential role of YouTube recommendation algorithms," *J. Commun.*, vol. 68, no. 4, pp. 780–808, 2018.
- [41] Y. L. Chen and C. L. Chang, "Early prediction of the future popularity of uploaded videos," *Expert Syst. Appl.*, vol. 133, pp. 59–74, Nov. 2019.
- [42] T. Trzcinski and P. Rokita, "Predicting popularity of online videos using support vector regression," *IEEE Trans. Multimedia*, vol. 19, no. 11, pp. 2561–2570, Nov. 2017.
- [43] J. Yu, M. Tan, H. Zhang, D. Tao, and Y. Rui, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, to be published, doi: [10.1109/TPAMI.2019.2932058](https://doi.org/10.1109/TPAMI.2019.2932058).
- [44] J. Yu, C. Zhu, J. Zhang, Q. Huang, and D. Tao, "Spatial pyramid-enhanced NetVLAD with weighted triplet loss for place recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: [10.1109/TNNLS.2019.2908982](https://doi.org/10.1109/TNNLS.2019.2908982).



RENJIE ZHOU was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Massachusetts at Amherst. He holds a postdoctoral position at the College of Computer Science and Technology, Zhejiang University, Hangzhou, China. He is an Associate Professor with the School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou. His research interests include measurement and analysis of online social networks, and data intelligence.



DONGCHEN XIA received the degree from Zhejiang Shuren University, in 2014. He is currently pursuing the master's degree with the School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China. He is currently with the Cloud Computing Research Center, Hangzhou Dianzi University. His current research interest includes measurement and analysis of online social networks.



JIAN WAN received the Ph.D. degree from Zhejiang University, Hangzhou, China, in 1996. He is a Professor and the Vice President of the Zhejiang University of Science and Technology, Hangzhou. His research interests include parallel and distributed computing systems, virtualization, and grid computing.



SANYUAN ZHANG received the Ph.D. degree in science from Zhejiang University, Hangzhou, China, in 1992. He was a Senior Visiting Scholar with the University of Hannover, Germany. He is a Professor with the College of Computer Science and Technology, Zhejiang University. His research interests include image and video processing, computer graphics, computer aided geometric design, computer mathematics, and digital design.