2013

Leave a Comment! An In-Depth Analysis of User Comments on YouTube

Peter Schultes
University of Passau, Passau, Germany, schult16@stud.uni-passau.de

Verena Dorner
University of Passau, Passau, Germany, verena.dorner@uni-passau.de

Franz Lehner
University of Passau, Passau, Germany, franz.lehner@uni-passau.de

Leave a Comment! An In-Depth Analysis of User Comments on YouTube

Peter Schultes, Verena Dorner, and Franz Lehner

University of Passau, Passau, Germany
schult16@stud.uni-passau.de,
{verena.dorner,franz.lehner}@uni-passau.de

Abstract. User comments are the most popular but also extremely controversial form of communication on YouTube. Their public image is very poor; users generally expect that most comments will be of little value or even in thoroughly bad taste. Nevertheless, heaps of comments continue to be posted every day. We propose an explanation for this contradiction in user attitudes and behaviour based on a new comment classification approach which captures salient aspects of YouTube comments. We show that, based on our new classification, we are able to perform very fast lightweight semantic video analysis. In addition, our results indicate that users' video perceptions (Likes and Dislikes) are indeed influenced by the dispersion of valuable and inferior comments.

Keywords: YouTube, comments, online video, video, social media

1 Introduction

More than 800 million people visit YouTube every month [16] and watch more than three billion hours of video material. In 2011, YouTube registered more than one trillion visits – this translates to approximately 140 visits by each member of the entire world population. Online videos fascinate and inspire internet users like no other medium. For a long time now, online videos evolved from an object of passive consumption into an object of social exchange. According to YouTube, more than 100 million people interact every week by rating, sharing and commenting videos [16]. 500 years of YouTube video are watched on Facebook each day, and over 700 YouTube videos are shared on Twitter each minute. YouTube is one of the largest platforms for user-generated content on the internet. Our own study provides evidence to support this fact: popular videos accumulate more than 500 comments each day and obtain some 100.000 ratings during their lifetime on YouTube.

The focus of this article is the most widely used communication feature on YouTube: user comments. To obtain a first impression of users' attitude towards YouTube comments, we conducted an online survey as part of another study (convenience sample of 95 participants). The participants had a rather negative view on YouTube comments: 64% perceived comments as "irrelevant", 42% as "aggressive" and 51% as stupid. Only 6% regarded comments to be "of essential importance for online videos".

11th International Conference on Wirtschaftsinformatik,
27th February – 01st March 2013, Leipzig, Germany

These findings are not generalizable to the general (online) public, but our participants' views are notably similar to the negative opinions voiced in many blogs and articles addressing this topic. In a Guardian article in 2009 [13], YouTube comments were described as follows:

"juvenile, aggressive, misspelled, sexist, […], YouTube comments are a hotbed of infantile debate and unashamed ignorance – with the occasional burst of wit shining through"

And it is not only the comments' content that is perceived to be inadequate. Even the way comments are presented to us is substandard: YouTube has implemented a sequential comment list sorted by creation date in descending order. Usually, eight to ten posts are displayed per page and the remaining comments can be viewed via paging. Unfortunately, only the first two or three comments fit in the originally visible space of the video page. The remaining posts disappear in the scroll area (at a typical vertical display resolution of 1080 pixels). Recent research on users' reading behaviour on web pages [11] has shown that we devote only 20% or less of our attention to the scroll area ("below the fold"). Consequently, places one and two in the comment list are privileged. They are likely the only ones that many users read. Now consider that we found that popular videos yield 200 to 800 comments each day (details in section 3). During peak periods, the publishing rate is even higher. This means that comments are pushed down into the scroll area within a few minutes of their publication, barely noticed – let alone read – by the users.

All in all it would appear that YouTube comments have a rather bad image among users, the press pokes fun at them, and the presentation form is so poor that only a tiny fraction manages to draw an audience. This raises the question: why do people persist in contributing comments? In our survey, 12% stated that they post comments regularly. A broader survey among 3.000 YouTube users [6] yielded approximately the same fraction of commenting users. The size of the YouTube community is estimated at 800 million users [16] – that means at least 96 million active comment authors. To illustrate, a country with a population this size would be the 13th largest country in the world. Apparently, and despite all negative publicity, comments are an essential feature of YouTube and by far the most widely used way to communicate on online videos.

Our user survey brought another interesting aspect to light: 34% (19% in some cases) stated that they read comments "often" and 53% agreed with the statement that they usually read the first two or three comments after watching a YouTube video. Considering that at any one time, a number of the over 800 million users are likely to be reading a comment, each comment is probably still read by thousands of people – even if publishing rates are very high.

In the remainder of the article, we first address the question whether YouTube comments offer added value and how we can measure it. We propose that users perceive an added value in additional information, entertainment and social exchange.

But, as we noted before, this would not explain the comments' poor image and users' contradictory attitudes and behaviour. We therefore suggest that the fraction of inferior comments is huge and mainly responsible for the negative image. Unfortunately, evidence to support our suppositions is difficult to find. YouTube comments are mainly unstructured text fragments, which makes an analysis very difficult. We therefore develop a two-tiered classification approach to help us gather structured information about comment type and quality from more than 100.000 unstructured user comments (section 3). We conduct an explorative study to test the validity of our research model and our suppositions (section 4). During model evaluation, we noticed that the distribution of our comment classes is a precise and characteristic feature of a YouTube video. In our next step, we demonstrate that comments are not just an end-user gimmick but also eminently suitable for semantic analysis of the actual video content (section 5). Our comment classification approach is, to a certain extent, a lightweight alternative to complex image processing. This is of particularly great interest to online video providers who must organize and classify enormous amounts of data. Finally, we investigate the influence of our identified comment types on users' rating behaviour (section 6). We show that there is a positive relationship between certain comment types and the dispersion of Likes and Dislikes, respectively. Obtaining many Likes, thus boosting a video's popularity, is important to the owner of a YouTube video because he can earn good money with the YouTube affiliate program. In fact, the revenues of YouTube affiliate program have more than doubled over four consecutive years [16]. YouTube, as an online video provider, also has a stake in ensuring high viewing numbers and quality (or at least popular) content: ultimately, the entire ad-based revenue model is based on viewing figures and user ratings.

2 Related work

YouTube comments are not a very common research area but there are some articles that gave direction to our work: Ammari et al. [1] investigated the question of how to identify "noisy" YouTube comments. The authors present a roadmap for filtering comments that are not relevant to a certain domain. Their "noisy" comments represent a subset of our T2 comments (details in chapter 3), and their pre-processing step resembles ours. In contrast to their approach, however, we deliberately desisted from focussing on any specific domain. We wanted a broader view on how people communicate via comments on YouTube. Mishne and Glace focused on comments in webblogs [10]. They found that comments constitute a substantial part of the blogosphere, accounting for up to 30% of the volume of weblog posts themselves. Their work underlines the importance of user comments as a way how people interact and extend primary content. Even though web blogs are a quite different domain, the authors identified comment types which are very similar to those we found in our study (e.g. discussion). Thelwall et al. [14] examined the characteristics of authors and comments that can be found on YouTube. They limited their work to descriptive statistics on YouTube comments (with a focus on discussions) without semantic statements. We compared our preliminary results with this work to ensure consistency and found a considerably higher amount of offensive posts in our data set.

In section 5, we show how our comment-based classification approach can be used for video categorization. Leung et al. present an approach based on a clustering of similar comment content [8] which leads to better search results and appropriate video categories "of interest". In comparison, our approach is less powerful but more general. Leung et al.'s procedure requires domain specific knowledge on the video category whereas our comment classes do not. A very interesting study on how user ratings for YouTube comments can be predicted was conducted by Siersdorfer et al. [12]. They found that positive and negative ratings are closely related to the video category. Their main objective is different to our study but our results could have benefitted from taking ratings into account, since it is one possible measurement of comment quality. Unfortunately, we could not include user rating in our data set since we were not able to collect them via our data collection API. Hsu et al. [5] developed an approach to rank user comments regardless of the context in which they appear. They present a regression-based procedure for automatic quality assessment of user comments depending on the preferences of the particular community. As an extension, their results could be mixed with our measures of the comment quality, which may lead to better classification.

3 Research Model and Data Collection

First of all we had to find out how users communicate via comments. We were not primarily interested in the motivations for posting messages (see e.g. [7] for detailed information on the motivations for posting comments in online communities) but rather in the observable behaviour of YouTube users. We therefore collected quite a large amount of comment data and developed a formal model for the object YouTube comment, and finally applied data mining methods. We conducted an in-depth analysis of 136.854 comments on 304 YouTube videos between 03/15/2012 and 03/21/2012. The comments were collected using the YouTube Data API [17], and stored temporarily for subsequent analysis. To keep the amount of data manageable, videos were selected automatically from standard feeds like "most popular", "most viewed" for the regions North America, Germany and Great Britain. Since classifying small amounts of data leads to unstable results [14], we only selected video categories containing at least five videos and only videos which attracted at least 50 comments during evaluation period. Our sample represents a snapshot of the most popular videos in the selected regions at the time. We decided against manually selecting additional videos for the smaller categories like Sports because there exists no obvious selection rule and we did not want to introduce selection bias into our results. Table 1 shows an overview of our data set.

Table 1. Overview of analysed video categories

	News & Politics	Comedy	Shows	Sports	Science & Technology	Gaming	People & Blogs	Music	Film & Animation	Enter-tainment	Pets & Animals
#videos	10	42	9	7	7	9	19	59	8	47	9
#com-ments	3498	35105	1825	804	1902	4156	14193	46630	1928	18481	1163

Basically, these are many unstructured text fragments. Our next challenge was to extract information from the data which would reveal the reasons for the contradictory usage behaviour of YouTube comments. We focussed on three basic comment types:

- **Discussion posts (T1)**: contains comments which are part of discussions among users. Since YouTube comments can be published as a reply to another comment, discussion threads evolve.
- **"Inferior" comments (T2)**: contains offensive statements and/or insults, comments without any relevant content or short emotional shout-outs.
- **"Substantial" comments (T3)**: contains comments without offensive statements that carry certain content information and are, ideally, directly related to the actual video content.

We propose that T1 and T3 comments provide added value for the users (information, entertainment, social exchange), whereas the “inferior” comments (T2) annoy most users. We decided to apply the definitions for the three comments types broadly and make no claims of their being complete. We operationalized our comment types on several relevant low level features (Table 2). The feature OFFENSIVE HINT was extracted by sentiment analysis with “SentiStrength” [15]. We also used manually built word lists for EMOTIONAL HINT for topics like "offensive", "amused" or "amazed" from samples of our data set. We did not rely on sentiment analysis alone because YouTube comments contain a high number of vernacular and cryptic shortcuts. These could not be processed satisfactorily by sentiment analysis. The features KEYWORD MATCH and TITLE MATCH are considered as indicators for a relationship between comment content and actual video content. A semantic analysis of video material turned out to be very difficult at this point. Automatic audio and image processing is still an extremely complex task, and the results must be treated with caution. Hence, we analysed metadata published along with a video instead [4], in particular the keywords and video title. Each comment was then searched for relevant keywords and parts of the title.

Table 2. Description of all Features

Feature	Description
SPAM HINT	Comment was reported as spam by a user
OFFENSIVE HINT	Comment has a negative sentiment strength greater than 2, a high proportion of capital letters and/or contains words for topics "aggressive" or "angry"
EMOTINAL HINT	Comment contains words for topics "amused", "amazed", "devoted" or "disgusting"
EMOTICON HINT	Comment contains at least one emoticon
PART OF THREAD	Comment is part of a discussion among users
#WORDS	Number of words
TIMESTAMP HINT	Comment contains a video timestamp (e.g. "at 1:30")
KEYWORD MATCH	Comment contains at least one relevant keyword
TITLE MATCH	Comment contains a relevant component from the title

Our comment types reflect a very general view on YouTube comments in their entirety and combine several specific communication patterns. To get a more accurate view on the characteristics of YouTube comments as a form of communication we defined several subclasses for each comment type. The formal description of each comment class was derived directly from the definition of the consisting comment type. In section 4 we will show that the comment classes revealed further insights in the way how people use comments and express their minds. In addition we were also able to consider inferior discussion posts through T1-subclasses C1 and C2. Table 3 lists the comment classes and their formal specifications that we used for classification.

Table 3. Predefined comment classes with formal definitions

Class		Title	Definition
T1	C1	*offensive discussion post*	PART OF THREAD & (OFFENSIVE | SPAM)
	C2	*insubstantial discussion post*	PART OF THREAD & (#WORDS < 4)
	C3	*normal discussion post*	PART OF THREAD
T2	C4	*spam or offensive post*	SPAM | OFFENSIVE
	C5	*short emotional shout out*	#WORDS < 9 & (EMOTIONAL | EMOTICON) & !(TIMESTAMP | KEYWORD | TITLE)
	C6	*insubstantial post*	#WORDS < 4
T3	C7	*reference to video content*	TIMESTAMP
	C8	*contribution with respect to video content*	#WORDS > 8 & (KEYWORD | TITLE)
	C9	*normal statement*	#WORDS > 10
	C10	*short statement*	#WORDS > 3 & #WORDS < 11

We designed the pre-processing step similar to [8], eliminating stop words and single characters form the text bodies. Each comment was then checked for class membership. Class assignment was carried out in a disjunctive fashion; the OFFENSIVE HINT feature, for example, does not match classes C5 - C10.

Note that we used both syntactic (e.g. #WORDS) and semantic (e.g. OFFENSIVE HINT) features for classification, which is the reason why some categories overlap. C1-comments, for example, would belong to type T2 if we ignored the syntactic feature PART OF THREAD. However, we decided to consider as many easily extractable features as possible to ensure optimal operationalization.

4 Consistency Check: Is our Operationalization Valid?

After classification, we tested the validity of our operationalization. Our goal was to verify the results of our classification by a descriptive evaluation of our test data set. This step was necessary because we cannot guarantee that the formal description of our comment classes represents their semantic meaning. We examined the distribution of comment classes in each video category to determine whether the variance can be explained based on video content only. This would provide strong evidence that our operationalization is highly reliable. The classification results are shown in table 4. The upper cell value marks the relative proportion of each subclass in each category and the lower value is the standard deviation.

C1 - C3: Users conduct comment-based discussions in a variety of ways. In the categories Sports, Science & Technology and Gaming, approximately 20% of all comments are part of discussion threads. Relatively few comments (less than 20%) belong to C1; the majority belongs to C3. Videos in these categories mainly address events and new products (for example *"Inspired Bicycles - Danny MacAskill April 2009"* or *"New iPad 3 Concept Features"*), for which there exist special-interest groups. People in this groups use the video as an opportunity for exchange, but mostly in a way which suggests that the videos do not polarize or address social issues. The category News & Politics presents a completely different picture: 57% of all comments are part of discussions. Nearly 40% of discussion posts contain offensive characteristics, suggesting that strongly polarizing topics are addressed. Considering that videos like *"9/11: Total Proof That Bombs Were Planted In The Buildings!"* or "*Appeasing Islam*" form part of this category, this phenomenon is perhaps little surprising. The distribution of the feature #WORDS is very interesting. In most categories, average word count lies between 9 and 16, but in News & Politics, it is 37, indicating longer and more substantive disputes. The ratio of #AUTHORS to #COMMENTS supports this view. In News & Politics, it is under 0.40, but higher than 0.70 in all other categories. In summary, comparatively few authors are responsible for most comments, and threads are much longer than in other categories. The category Pets & Animals in particular shows a completely different picture: Only 11% of all comments belong to classes C1 - C3 and the ratio of authors to comments is much higher at 0.94 than in News & Politics. Videos of this category apparently cause only little occasion for discussion.

Table 4. Distribution of comment classes among video categories.

		News & Politics	Comedy	Shows	Sports	Science & Technology	Gaming	People & Blogs	Music	Film & Animation	Enter-tainment	Pets & Animals
T1	C1	**0.22**	**0.05**	**0.03**	**0.03**	**0.04**	**0.04**	**0.07**	**0.03**	**0.02**	**0.04**	**0.02**
		0.10	0.04	0.02	0.03	0.03	0.02	0.07	0.02	0.01	0.05	0.02
	C2	**0.01**	**0.02**	**0.03**	**0.03**	**0.02**	**0.04**	**0.02**	**0.04**	**0.02**	**0.03**	**0.02**
		0.03	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02
	C3	**0.34**	**0.08**	**0.08**	**0.15**	**0.18**	**0.17**	**0.08**	**0.11**	**0.12**	**0.13**	**0.07**
		0.08	0.06	0.05	0.09	0.24	0.08	0.08	0.07	0.06	0.11	0.04
T2	C4	**0.18**	**0.22**	**0.16**	**0.20**	**0.19**	**0.17**	**0.28**	**0.24**	**0.19**	**0.18**	**0.16**
		0.06	0.06	0.05	0.07	0.08	0.03	0.10	0.11	0.06	0.05	0.04
	C5	**0.01**	**0.11**	**0.07**	**0.06**	**0.06**	**0.09**	**0.07**	**0.11**	**0.12**	**0.12**	**0.14**
		0.01	0.04	0.05	0.05	0.04	0.04	0.05	0.06	0.03	0.06	0.04
	C6	**0.02**	**0.15**	**0.20**	**0.12**	**0.14**	**0.09**	**0.13**	**0.11**	**0.13**	**0.13**	**0.20**
		0.12	0.05	0.11	0.02	0.07	0.04	0.06	0.03	0.06	0.05	0.07
T3	C7	**0.01**	**0.03**	**0.04**	**0.13**	**0.01**	**0.04**	**0.01**	**0.03**	**0.04**	**0.04**	**0.02**
		0.01	0.03	0.04	0.08	0.01	0.03	0.01	0.04	0.03	0.03	0.01
	C8	**0.09**	**0.11**	**0.25**	**0.11**	**0.09**	**0.12**	**0.11**	**0.13**	**0.10**	**0.11**	**0.13**
		0.03	0.05	0.17	0.04	0.06	0.05	0.09	0.06	0.07	0.09	0.08
	C9	**0.11**	**0.14**	**0.08**	**0.11**	**0.19**	**0.16**	**0.16**	**0.12**	**0.17**	**0.13**	**0.16**
		0.09	0.05	0.06	0.08	0.10	0.04	0.06	0.04	0.06	0.04	0.08
C10		**0.01**	**0.09**	**0.06**	**0.06**	**0.08**	**0.08**	**0.07**	**0.08**	**0.09**	**0.09**	**0.08**
		0.02	0.03	0.04	0.02	0.05	0.02	0.04	0.02	0.03	0.04	0.03

C4 - C6: The proportion of spam-like or offensive comments is relatively high across all categories. Only Pets & Animals exhibits low levels below 20% (16% and an additional 2% of offensive discussion posts). The category People & Blogs leads C4-comments with 28%. Videos such as *"I AM A MUSLIM!"* or *"Abortion? I love you. Life is beautiful. pregnant. My speech to President Obama"* address socially relevant topics on which there are strongly conflicting opinions in the community. YouTube users apparently tend to express their opinions in an offensive rather than a content-centred way. We also identified many C4-comments for user-generated videos (in People & Blogs) such as *"GINGERS DO HAVE SOULS!!"*. User-generated videos with controversial content seem to inspire users particularly to post offensive or derogatory comments. One out of four comments in Music category contains spam or offensive posts. We found a large proportion of comments which were reported as spam by users. In particular, we encountered an increased number of references to personal channels and hidden advertisements. Videos from Comedy feed likewise evoke a lot C4-comments (22%), which is probably due to the fact that videos such as *"Americans are NOT stupid - WITH SUBTITLES"* are not considered as funny by everyone. The largest number of offensive and spam-like comments was found in News & Politics, if adding C4- and C1-posts. Nearly 40% could be assigned to these classes. C5- and C6-comments dominate in the categories Film & Animation, Entertainment and Pets & Animals, which mainly contain entertaining or amusing videos. It is not surprising that videos such as *"World's Largest Rope Swing"* or funny animal videos instigate short emotional shout outs. In general, we found that the proportion of very short posts with no significant content (C5 and C6) is rather high at 20% to 30%.

C7 - C9: C7-comments represent only a very small proportion of comments (2% to 4%). However, these posts certainly offer added value to the viewer: With the help of C7-comments, users can skip to the most interesting scenes in the video instead of being forced to view the entire video. The only exception is the Sports category where 13% of all comments fall into class C7. In many sports events, there are few decisive or spectacular scenes (e.g. football goals being scored) which are of particular interest to users and most likely to be commented on; hence the frequency of C-7 comments C8-comments are a very important comment type as well. In our opinion, comments that contain neither offensive nor spam like statements, that are of a certain length and are related to video content are most likely to be perceived as valuable by users. In most categories, 10% to 13% of comments are C8 comments. The comparatively low proportion of 9% in News & Politics is due to disjunctive classifying. Almost half the comments had already been assigned to classes C1 - C3. If we had ignored the PART OF THREAD-feature, the total amount of C8-comments would have been much higher for this category. C9-comments are similar to C8-comments. They are content carriers without offensive or spam-like tokens but without a direct relationship to video content. In Science & Technology, C9-comments occur frequently because many posts discuss alternative or competing products (especially for videos such as *"New iPad 3 Concept Features"* or *"BAE Electromagnetic Railgun"*). Since in these cases no keywords matches are found, many posts are classified as C9 instead of C8-comments.

Implications. Our results indicate that no comment type is dominant on YouTube. Rather the opposite is true: users communicate on different video topics in different ways. This resembles verbal communication patterns: we talk about political topics in a different way than about sport events. Indicators are, for example, emotional tokens, stronger content relationship or extensive discussions. Overall, we found a relatively high proportion of offensive posts. But basically, YouTube comments appear to reflect real-life communication behaviour. Consequently, the distribution of the ten comment classes is highly different among the video categories and the variance can be explained mostly by the video content itself. This is a strong indicator for the validity of our operationalization and confirms our supposition: we found a substantial amount of comments that do not contain offensive statements and can be perceived as content carriers. Likewise, social interactions take place in 20% of all posts via discussion threads. These two comment types offer added value to the users, which we suggested is the main reason why so many users post so many comments. On the other hand, our investigation revealed that about 30% of all comments belong to comment type T2. These comments are likely perceived as disturbing and annoying by the majority of users. This fact explains why users have such a negative impression of YouTube comments and also explains the contradictory usage behaviour.

5 Comment Classes as a Basis for Automatic Video Analysis

The results of our comment classification offer interesting insights into user commenting behaviour on YouTube. One particular interesting variable is the variance of the distribution of different comment classes across different video categories. If YouTube users commenting behaviour depends on the content of the video in question, then we might be able to use the variance of the distribution of different comment classes as a means to identify which topic a video is likely to deal with.

Automatic extraction of semantic information from raw video data is still extremely difficult and error prone. But for online video providers like YouTube, these evaluations are essential because they allow better search results and indexing of video data. We examine how our results on comment class distributions can be used to overcome the "semantic gap" in some areas. We picked the video category as the subject of analysis, which, due to its very general nature, is very likely one of the first steps in a video retrieval process. Our aim is to estimate reliably which video category a video belongs to, based on available comments and the distribution of the comment classes.

For this purpose we conducted a support vector machine (SVM) based classification. We performed n-force cross validation by splitting the comments on our 226 videos (Table 3) into ten parts. One part is used as test data, the other parts as training data, and the test data switches after each run. Every test video is described by a vector composed of the actual category as well as the relative distribution of the comment classes. After the training run, the SVM estimated the category for each test vector. We used the LIBSVM implementation [2] for our investigations and achieved the best results by configuring the SMV as a nu-classifier working on a radial kernel (nu = 0.1). In our test run the classifier had an error rate of merely 2.8%. Table 5 shows the detailed results for each category. The outstanding quality of the measured values indicates that comment class distribution is a reliable decision base for providing first assessments of the actual video content and the video category, respectively. Of course we cannot ensure the reliability of this method in general yet, but our results are very encouraging. More experiments with different training and test data are needed to confirm our claim.

Table 5. Results of the SVM-classifier for our data set

	News & Politics	Comedy	Shows	Sports	Science & Technology	Gaming	People & Blogs	Music	Film & Animation	Enter-tainment	Pets & Animals
Precision	1.00	0.88	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Recall	1.00	1.00	1.00	1.00	0.67	0.67	0.89	1.00	1.00	1.00	1.00
F-Measure	1.00	0.93	1.00	1.00	0.80	0.80	0.94	1.00	1.00	1.00	1.00

The automatic determination of the video category showed that comment classes are very well suited as a lightweight approach to gather semantic information for online videos. Apart from category membership, we have identified further opportunities for using YouTube comments in semantic analysis and implemented them in prototypes.

C7-comments, for example, can be used to identify the most relevant scenes in the video. Assuming a relatively small number of 5.000 comments for a particular video, a classification run results in approximately 200 C7-posts from which it is very likely we will be able to extract the most important scenes.

Beyond that, C8-comments are suitable for summarizing video content. Since keywords are not very user-friendly as a means for providing a content summary, we employ C8-comments to fill this gap. Identifying appropriate C8-comments for the sequences that were extracted from C7-comments turned out to be manageable by performing textual similarity checks. The enrichment of C7-comments by C8-comments is recommendable because samples showed us that C7-posts tend to be relatively short. The user attitude towards a video can be measured by the amount of C4- and C5-comments as well as parts of our feature indicators for offensive and emotional posts, respectively. Even with relatively simple features, we were able to determine the degree of emotionality in a comment and form a basic attitude via accumulation. Analysing C3-comments is also a very convenient way of gaining further knowledge. We assume that, since the post raised a discussion, the content of the "root" comment is particularly interesting for users. The longer the thread length, the more relevant the post seems to be.

There are many further interesting approaches for a comment based content retrieval system for YouTube videos, which we are working on now. We would like to refer to our future research work since further details would go beyond the scope of this article.

6 Comment Types and Rating - is there a Relationship?

In the previous section, we showed that comments have the potential for some kinds of semantic analysis, which is very important for platform providers like YouTube. In the next step, we will demonstrate how comments relate to the dispersion of Likes and Dislikes. The direct effect of comments on the number of views is difficult to measure: obviously, a viewing takes place before comments appear. It is therefore doubtful whether an effect of comments on the number of views exists. Even if there were a statistic correlation, it would be difficult to arrive at a meaningful interpretation. But an increased number of Likes affects the number of views since the video indicates higher popularity to other users. Positively rated videos are ranked higher than unpopular videos and generate more views. And, of course, more views have a strong effect on platform revenue because more people watch commercials. The video owner is also interested in increasing the number of views and positive ratings. This interest is partly due to the participation in the YouTube affiliate program and partly to a higher social standing in the community.

To obtain a detailed view on the correlations between comments and the dispersion of Likes and Dislikes, we estimated two negative binomial regression models. We took the total number of T1-, T2- and T3-comments as independent variables and the relative dispersion of Likes and Dislikes as dependent variables. Negative binomial regression is an appropriate evaluation method for our data because we modelled the comment types as overdispersed count variables. The results are shown in Table 6.

Table 6. Effects of comment types on dispersion of Likes and Dislikes

Variable	%Likes			%Dislikes		
	Est.	SE	p	Est.	SE	p
(Intercept)	6.1719	0.0892	< 2e-16***	3.4850	0.0892	< 2e-16***
#T1-comments	-0.0019	0.0006	0.0010***	-0.0009	0.0006	0.139
#T2-comments	0.0014	0.0005	0.0023**	0.0018	0.0005	5.92e-05***
#T3-comments	0.0027	0.0009	0.0034**	0.0014	0.0009	0.117
Log Likelihood:	-1710.80***			-1124.89***		
Nagelkerke's R^2:	0.4236			0. 6398		

Note: '.' p < 0.1, '' p < 0.05, '**' p < 0.01, '***' p < 0.001, ' ' p < 1*

As we can see from Nagelkerke's R^2-values, our models explain a respectable amount of variance. Video content itself certainly has a major influence on user ratings, but we could not integrate it into our model for lack of a measure for the perception of the video content. We believe that this variable explains much of the remaining variance.

The estimated values indicate two key findings. First, the number of T3-comments has the strongest effect on the relative amount of Likes ($p < 0.01$, $\beta = 0.0027$). The other two comment types have a significant ($p < 0.01$) but substantially lower effect on the dispersion of Likes. T1-comments cause the opposite effect. In other words, larger numbers of T3-comments correspond to a higher number of Likes. The second key finding concerns the dispersion of Dislikes. We measured a significant ($p < 0.001$, $\beta = 0.018$) influence of the total number of T2-comments on the amount of Dislikes. The other two comment types do not have a significant effect. Put differently again, an increased number of T2-comments evokes a higher number of Dislikes. This is in accordance with the findings from our descriptive analyses (section 4):

T3-comments seem to provide added value to the users whereas T2-comments have the reverse effect.

In conclusion, the regression analyses showed that different distributions of comment types influence the users' rating behaviour. The recommendations for a video owner are clear: motivate users to post fair, substantial and content specific comments (T3-comments) to get a higher amount of Likes. To counteract Dislikes, it should be made quite clear that offensive, spam like or insubstantial posts (T2-comments) are not welcome on the video page. However, a video owner has little influence on the comments published for his video. Only YouTube as the platform provider is capable of implementing and enacting appropriate policies.

7 Discussion

In this article we dealt with a relatively new form of communication, YouTube comments. The main object of our study was the distribution of three different comment types among all YouTube comments: discussion (T1), inferior post (T2) and substantial comment (T3). Our results support our suggestions that the bad image of YouTube comments is due to the high amount of T2 comments, but that users still make frequent use of comments due to added value (information, entertainment, social exchange, etc.) derived from, T1- and T3-comments. Furthermore, with the help of our classification approach, we were able to provide further insights into the communication form "YouTube comment" and to highlight its main characteristics.

Our analyses yield several recommendations for improving user acceptance of YouTube comments. High numbers of T2-comments displace substantial comments. To counteract the negative image of YouTube comments, the added value of T1- and T3-comments ought to be emphasised. Appropriate visualizations which take context dependencies for particular video sequences into account, for example, could highlight valuable posts. Dynamic and media time based annotations, which were established in the context of interactive videos years ago [3] [9], are likely to be very well suited for visualizing substantial or context-specific user comments. Furthermore, we noticed that users express their emotional attitude towards a video via short shout outs (C5- and C6-comments), which also displace substantial contributions. It might be helpful at this point to launch a secondary rating system, which user could use to express their emotional attitude. Short emotional posts would retain their significance if a more suitable visualization form was available. Our results at least confirm the claim that users frequently communicate their emotions in comments.

Beyond that, our classification approach showed that comments are very well suited for semantic analyses of video content. First results of our prototypes showed that comments can be used for automatic video retrieval, which leads to higher-quality indexing and better search results. Since it is safe to assume that YouTube will present us with new record upload and view statistics in near future, lightweight alternatives to audio and video processing will become more and more important. Our regression analysis shows that YouTube comments affect the way we perceive online videos. We are now able to describe the main characteristics of a comment that creates added value for the users: fair, substantial and relevant for the underlying video are some of them. This knowledge forms the basis for recommendations to YouTube regarding the way comments should be presented and published, respectively.

In our opinion, the potential of video comments goes far beyond improving bad publicity. We could translate existing research results on social TV and interactive video to the context of online videos and provide a completely new experience of social exchange to millions of users. Interactive and context-sensitive comments, which are either directly embedded in the video context or in the surrounding page, may lead to a higher level of entertainment and information. We definitely think that online videos will evolve into a social medium, and communicative features such as "comments 2.0" will help get there. In future research we will focus on this trend and investigate further ways how users can participate in this global social exchange.

There are several limitations of our work that provide avenues for future research. Our data set, for instance, is based on certain YouTube feeds and contains popular online videos only. Our findings are limited to popular videos for which a reasonably large number of comments is available, because the analysis methods we used rely on these assumption. Furthermore our features do not cover the user perception of a video, since our data set did not cover this information. The regression analyses on the relationship of comment types and user ratings would certainly reveal further interesting results if we included the video content perception as another independent variable. Further research will be necessary to analyse these correlations.

References

1. Ammari, A., Dimitrova, V., Despotakis, D.: Semantically Enriched Machine Learning Approach to Filter YouTube Comments for Socially Augmented User Models. In Proceedings of the Int. Workshop on Augmented User Models, UMAP Conference, pp. 6 - 17 (2011)
2. Chang C.-C., Lin C.-J.: LIBSVM: A library for support vector machines. In ACM Transactions on Intelligent Systems and Technology, Volume 2, Issue 3, pp. 1 - 27 (2011)
3. Fagá, R.Jr., Motti V.G., Cattelan R.G., Teixeira, C.A.C. Pimentel M.d.G.C.: A social approach to authoring media annotations. In Proceedings of the 10th ACM symposium on Document engineering, pp. 17 -26 (2010)
4. Hsu C.-F., Caverlee J., Khabiri E.: Hierarchical Comments-Based Clustering. In Proceedings of the 2011 ACM Symposium on Applied Computing, pp. 1130 - 1137 (2011)
5. Hsu C.-F., Khabiri E., Caverlee J.: Ranking Comments on the Social Web. In Proceedings of the International Conference on Computational Science and Engineering 2009, pp. 90 - 97 (2009)
6. iab research: Social: A closer look at behaviour on YouTube & Facebook, Executive summary. http://www.iabuk.net/research/library/social-a-closer-look-at-behaviour-on-youtube-and-facebook
7. Lampe C., Wash R., Velasquez A., Ozkaya E.: Motivations to participate in online communities. In Proceedings of the 28th international conference on Human factors in computing systems, pp. 1927-1936 (2010)

8. Leung J. K.-W., Li C.H., Ip T. K.: Commentary-based Video Categorization and Concept Discovery. In Proceedings of the 2nd ACM workshop on Social web search and mining, pp. 49 - 56 (2009)
9. Meixner B., Siegel B., Hölbling G., Lehner F., Kosch H.: SIVA suite: authoring system and player for interactive non-linear videos. In Proceedings of the international conference on Multimedia, pp. 1563-1566 (2010)
10. Mishne, G., Glance, N.: Leave a Reply: An Analysis of Weblog Comments. In Proceedings of the Third annual workshop on the Weblogging ecosystem (2006)
11. Nielsen J.: Scrolling and attention. In Alertbox: Current Issues in Web Usability, http://www.useit.com/alertbox/scrolling-attention.html (ISSN: 1548-5552) (2010)
12. Siersdorfer S., Chelaru S., Nejdl W., San Pedro J.: How useful are your comments? Analyzing and predicting youtube comments and comment rating. In Proceedings of the 19th international conference on World wide web, pp. 891 - 900 (2010)
13. The Guardian: Our top 10 funniest YouTube comments – what are yours? From http://www.guardian.co.uk/technology/blog/2009/nov/03/youtube-funniest-comments
14. Thelwall M., Sud P., Vis F.: Commenting on YouTube videos: From guatemalan rock to El Big Bang. In Journal of the American Society for Information Science & Technology, Vol. 63 Issue3, pp. 616 - 629 (2012)
15. Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., Kappas, A.: Sentiment strength detection in short informal text. In Journal of the American Society for Information Science and Technology, 61(12), pp. 2544–2558. (2010)
16. YouTube.com: YouTube facts and figures, http://www.youtube.com/t/press_statistics
17. YouTube Data API: YouTube API Version 2.0 Developers Guide, https://developers.google.com/youtube/2.0/developers_guide_protocol_audience