THEME ARTICLE: Jan-March 2019-5G Multimedia Communications: Theory, Technology, and Application

Emotion-aware Video QoE Assessment via Transfer learning

Yixue hao
Huazhong University of Science and Technology

Jun Yang
Huazhong University of Science and Technology

Min Chen (corresponding author)
Huazhong University of Science and Technology

M. Shamim Hossain
King Saud University

Mohammed F. Alhamid
King Saud University

In this article, we proposed an emotion-aware video quality of experience (QoE) assessment. Specifically, we first detect the user's emotion via transfer learning and match the user's emotion with video content. Second, we design a personalized video QoE prediction model by consider the diversity of the users. Third, a decision tree model is utilized to illustrate the nonlinear relationship among average bitrate and buffer ratio and QoE. Experiments show that compared to the traditional video QoE assessment, emotion-aware video QoE assessment has higher score.

With the current popularity of mobile devices and the prosperity of network video services, mobile video has become the main source of mobile data. It is estimated that mobile video data accounts for more than half of the total mobile data worldwide, and, by the end of 2018, this proportion will increase to 2/3. At the same time, the industry market of network video service is expanding rapidly. For example, current commercial radio and television companies (such as the FOX Broadcasting Company and NBC) and the suppliers providing video stream service according to customer demand serve millions of users every day. These factors not only illustrate the huge demand of mobile video, but also point to a possible new service paradigm: mobile big data-driven video service computing. Thus, video quality of experience (QoE), assessment becomes more and more important[1-2].

Nowadays, there are many factors impacting video QoE[3-5]. Internet service providers, media player designers, video content providers, and content delivery networks (CDN) have different metric standards. Suppliers of video content may improve the users' experience quality by choosing a better coding rate, and network operators of content distribution may enhance the user experience through scheduling better edge servers. The most common factors affecting video QoE are as follows:

- Video attributes: the users viewing experience and satisfaction depend to a large extent on video attributes, such as the quality, content, and popularity of the video. These are the primary factors considered by video content providers. Some general factors in this category include, but are not limited to, the following: encoding bit rate, video length, video type (live broadcast or video on demand), and video popularity.
- Network quality: a network of poor quality will greatly reduce the user's viewing experience, so a lot of quality experience models include network quality during video watching as a major factor. Network quality parameters include, but are not limited to, the client, core network, and wireless access network, with impact factors being start-up delay, buffer time, and buffer ratio, etc.
- Additional factors: previous studies have shown that viewing environment has a significant effect on the user's quality experience. Factors falling within this definition include device type (smart phone, tablet or laptop), time and location.

Video QoE assessment model is the function that captures the user's diversified QoE patterns with the input of various influencing factors. For the video QoE, the most direct way to obtain it is subjective test (e.g., mean opinion score (MOS). But the subjective tests require more human and material resources. With the emergence of big data analytics for mobile video applications, data-driven video QoE assessment[6] has attracted extensive attention from both academic and industrial researchers, which is focused on the following factors (i.e., quality of service (QoS) metrics): start-up delay, buffer ratio, and average bitrate. There are three prediction models based on data-driven QoE, which are based on linear regression[7], the decision tree[8], and the quasi- experimental designs (QED) model[9]. However, such data-driven video QoE assessment do not consider the diversity of users and accuracy is not high.

Thus, in this article, we consider a user's emotional reaction to be a key factor in the user's watching experience, since the mood of the user reflects the specific reaction expected by the maker of the corresponding short video clip. We believe the newly introduced mood factor is critical to building a user video QoE model for each user who has diversified video-viewing patterns, influenced by interest in specific content, personality, instant emotion, and suitable mood related to the user's environment. Therefore, we first consider that the user's mood can be more subjectively related to video QoE, so we can use the change of the users' mood as an important metric of video QoE. Thus, we propose an emotion-aware video QoE assessment (EQA) model. Second, we establish a personalized video QoE model for every user. Third, a decision tree adjustment model is utilized to illustrate the nonlinear relationship among average bitrate, buffer ratio and QoE. In summary, the contributions of this article include:

- Based on the transfer learning, we propose a new emotion-aware video QoE evaluation metric.
- By taking into account the diversity of the user, we build an emotion-aware personalized video QoE prediction.
- Compared to the traditional video QoE assessment, emotion-aware video QoE assessment has better performance.

EMOTION-AWARE VIDEO QOE ASSESSMENT FRAMEWORK

In this section, we describe the emotion-aware video QoE assessment framework, which include sensing the emotion, emotion-aware video QoE assessment metric and emotion-aware personalized video QoE prediction, as shown in Figure 1. The designed primarily considering the following factors:

- Appropriate QoE assessment: consider good indicator of user's experience or satisfaction.
- Identifiable influencing factors: consider the commonly-used factors.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

- Consider user diversity: consider that different people have different video QoE.

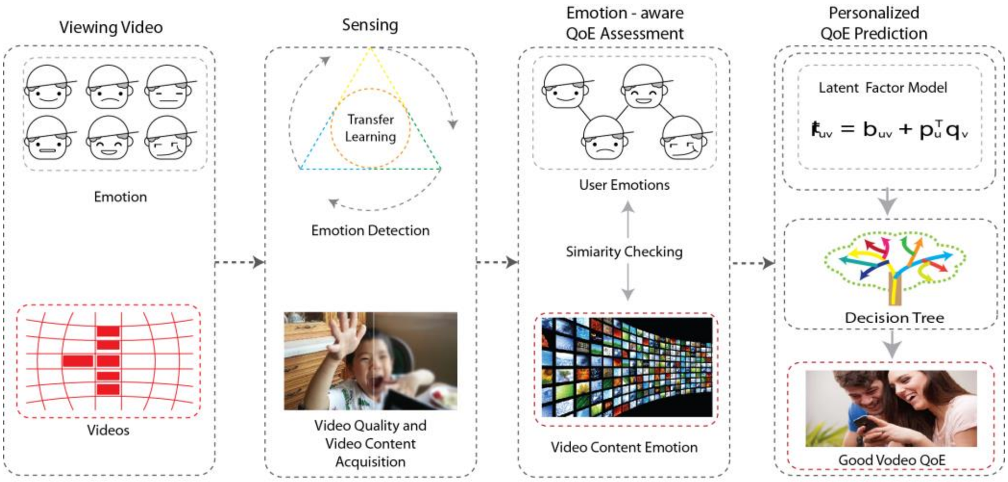

Figure 1. The emotion-aware video QoE assessment framework.

Sensing the Emotion

For the emotional sensing, we use the Circumplex mood model as shown in Figure 2. From the Figure, we can see that the model includes two dimensions, i.e., pleasure dimension and activation dimension. The dimension of pleasure is from pleasure to displeasure, while the dimension of activation is from activation to deactivation. Thus, this model can describe many emotions. When using the Circumplex mood model, we choose a set of standard and representative emotional states as shown in Figure 2.

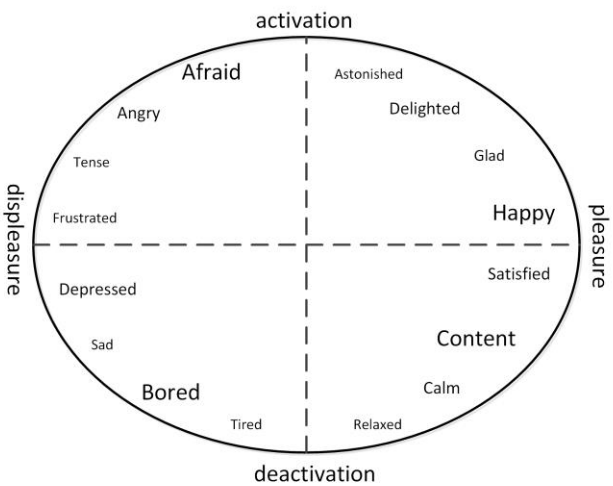

Figure 2. The Circumplex model of mood.

When users are watching the video, we establish the model by using the collected data, build classifier to obtain their emotion by using the transfer learning and hidden Markov model (HMM)[10-11], and also detect the quality and content of the video watched by users, by obtaining video content emotions through statistical methods.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

Emotion-aware Video QoE Assessment Metric

The user's emotion is the most direct manifestation of video QoE, so we put forward an EQA metric. We compare the user's mood measured by real-time detection with the emotion of the video content: when the similarity is high, the video QoE is higher; otherwise, the video QoE is relatively low.

Emotion-aware Personalized Video QoE Prediction

While considering the diversity of users, we establish a personalized model of video QoE for each user, rather than a common model for all users, which can be more accurate, given the user's QoE preference. Furthermore, using the information in a decision tree may help us to know which factor accounts for the largest proportion, and to adjust the different factors influencing the experiences of different users, so as to achieve a satisfactory users QoE.

EMOTION-AWARE VIDEO QOE ASSESSMENT METRIC

Table 1. Collection user's emotional data

	Category	**Components**
Physical data	Electrocardiogram	ECG
	Facial expression	Facial expression
	Activity	Static, walk , run
Cyber data	Video viewing log	Category and duration
	Web usage log	Type and number

The emotion-aware video QoE assessment metric is proposed in this section. Specifically, we first use transfer learning and hidden Markov model to identify user emotions. Then, we use similarity measure to give the correspondence among the user's emotion and the emotion of video content. Specifically, the emotion detection are as follows:

- Collection and feature extraction of Emotional data. The emotional data collected consists of physical and cyber data, as shown in Table 1. Specifically, the physical data includes electrocardiogram (ECG) data, facial expression data, activity levels data (i.e., static, walk and run). The cyber data includes video viewing data (i.e., category and duration) and web usage data (i.e., type and number). Based on the collected data, we clean up the data and extract the features. For emotional data labels, users can use smartphone to label.
- Automatic labeling using transfer learning. It is a time-consuming and labor-intensive task for the user to label their personal emotions. Therefore, we use the transfer learning to label emotions automatically. That is, we only have some labelled emotional data, and unlabeled data are labelled using transfer learning. To be specific, the user's labelled data is used as source domain input data, and user unlabeled data is used as target domain data. We use transfer learning to estimate the similarity of the users, and use the low-k similar distributions to calculate the maximum probability label.
- Label validation. We use transfer learning to evaluate the validity of the labels. Specifically, we detect the input of users' emotional label to applications such as Moodagent, and regard it as ground-truth labels. We observe that the emotional space we detected from the application may not match the emotional label space. Therefore, we utilize transfer learning to construct the similarity measure between emotional label space and user input emotions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

- Emotion detection**.** Based on Hidden Markov models, we can detect the users' emotions. Moreover, as the amount of emotional data increases, the accuracy of our model is higher and higher. Therefore, we get the emotion detection model.

The emotion-aware video QoE assessment are as follows: Through the above methods, we can detect the emotion of users. Then, we give the emotion of video content. In this article, we first use questionnaire to give the video' inherent emotion label, as shown in Figure 3. Second, based on the SentiWordNet[12] dictionary, we transform the emotional words into scores [−1, 1]. Finally, we use the distance similarity to compute the detection mood matching for video QoE (MMVQ). We set two thresholds and discretize the MMVQ into **{**low, medium, high**}**. That is, if the MMVQ is low, it indicates that the mood of user does not match with the emotional attributes of content video. If the MMVQ is medium, it displays that the mood of user's mood match with the video content emotion to some extends. If the MMVQ is high, it is considered that the mood of user match the video content emotion match well.

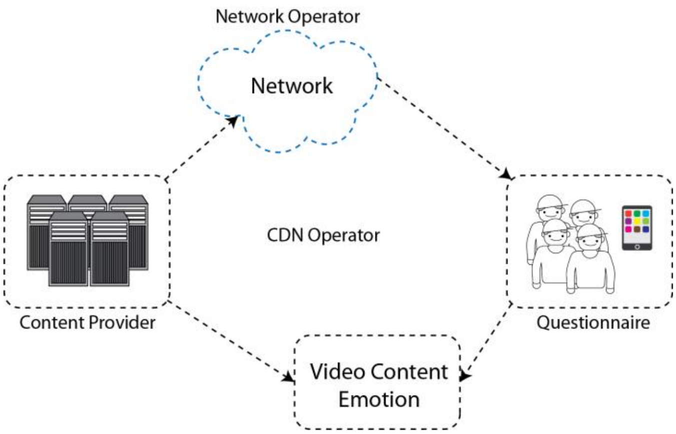

Figure 3. Detection of video content emotion.

EMOTION-AWARE PERSONALIZED VIDEO QOE PREDICTION

In this section, considering that each user has a different video QoE, we propose an emotion-aware personalized video QoE prediction model. When discussing the relationship between the QoE of the users and the influential factors, we find that individual users obviously have different sensibilities to different these factors. To handle this challenge, we need to make a special prediction model for each user, instead of constructing a universal model and considering all users as a unified whole. In other words, we need to capture the varieties of users in the model.

Supposed there are m users $\{u_i|i = 1,2,\cdots,m\}$, each with l features. We then take the user's feature matrix as U. With s videos $\{v_i|i = 1,2,\cdots,s\}$, each with k features, the video features matrix is regarded as V. We can obtain the assessment of the users for each video's QoE, which is represented as a matrix R, with elements denoted as r_{uv}. However, each user only watches a small part of the videos, so the experience quality matrix R contains sparse data. We can make use of the personal information and similarity of the videos to predict the great number of unknown values in the experience quality matrix.

To handle the varieties of users and the problem of data sparsity, we put forward a personalized video QoE model based on the latent factor model (LFM)[13-14]. It provides a method that not only enables information sharing among similar users and videos, but also allows sharing of knowledge that comes from different domains. In our model, the predicted QoE $\hat{r}_{uv}$ is presented by the following formula:

$$\hat{\mathrm{r}}_{uv} = b_{uv} + p_u^T q_v,\ \mathrm{u} = 1,2\cdots,\mathrm{m};\ \mathrm{v} = 1,2,\cdots,\mathrm{s}.$$

where b_{uv} is the baseline projection, p_u is the vector of user-factor (namely, the preference of the users to these latent factors), and q_v is the vector of video-factor (namely, the scores that the video obtains for these factors (such as V)). Thus, $p_u^T q_v$ indicate sharing the information among similar users and videos. b_{uv} defined as $b_{uv} = \mu + b_u + b_v$, where μ is the average QoE of all users, b_u stands for the bias of users, and b_v is the bias of videos. So b_{uv} indicate sharing the information among different domains of definition. That is, due to their individual personalities, some users have higher demands for the QoE of the videos. For example, some videos are so good that some people will watch them even when the network is in a bad condition.

In the above formula, the optimal values b_u, b_v, p_u and q_v are all unknown. To solve for them, we can transform this problem into solving a regularized least squares problem as follows.

$$\min_{\{b_u, b_v, p_u, q_v\}} \sum_{\{u,v\}} (r_{uv} - \mu - b_u - b_v - p_u^T q_v)^2 + \lambda_1 b_u^2 + \lambda_2 b_v^2 + \lambda_3 \| p_u \|^2 + \lambda_4 \| q_v \|^2$$

where $\lambda_i, i = 1,2,3,4$, are regularization constants to avoid over-fitting during calculation.

We use stochastic gradient descent to estimate the model parameters by minimizing the regularized squared error function. Define $e_{uv} = r_{uv} - \hat{r}_{uv}$. It follows that

$$\begin{aligned}
b_u &\leftarrow b_u + \gamma_1 (e_{uv} - \lambda_1 b_u) \\
b_v &\leftarrow b_v + \gamma_2 (e_{uv} - \lambda_2 b_v) \\
p_u &\leftarrow p_u + \gamma_3 (e_{uv}\, q_v - \lambda_3 p_u) \\
q_v &\leftarrow q_v + \gamma_4 (e_{uv}\, p_u - \lambda_4 q_v)
\end{aligned}$$

where $\gamma_i, i = 1,2,3,4$ are the learning rate. We can iterate through the above formula and obtain the unknown QoE of the users. The baseline projection b_{uv} captures the basic experience quality model and the bias brought by each user and video. Meanwhile, $p_u^T q_v$, i.e., the similar part of the low order, shows the experience quality fluctuation caused by varieties of users and videos. The intuition is that we can access the preference of the users based on these latent factors and the scores of videos for these latent factors through converting the users and videos to the same latent factor space. Each user u_i and each video v_j correspond to a preference vector p_u and a score vector q_v respectively.

According to the above discussion, we can obtain the personalized video QoE. Then we use decision tree to adjust the user's video QoE. The basic idea of the decision tree technique is that an object is classified by minimizing the impurity in the data. Based on the decision tree model, we study the relationship among buffer ratio, average bitrate and video QoE. Specifically, we discretize the buffer ratio and average bitrate into {low, high}. With the decision tree algorithm, we can determine the relationship among video QoE, buffer ratio and average bitrate. Furthermore, we can get better video QoE by adjusting those influencing factors.

PERFORMANCE EVALUATION

In this section, we demonstrate the proposed EQA metric and the personalized video QoE prediction. Specifically, we establish datasets and compare the estimation performance using our model include the MMVQ metric, video QoE predictive model and decision tree-based adjustment model.

Datasets

Video dataset: We have collected streaming video datasets, including 100 comedies and 50 other categories of video from YouTube. We use 200 participants, including undergraduate and graduate students, to label the emotion of video content. Their age ranges from 18 to 45. Each participant watched 30 videos, labels the emotion of video and gives the opinion score. Therefore, we

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

can get the emotion and corresponding MOS of video content through these participants. Moreover, the buff ratio and average bitrate when users watching videos are also collected.

Emotion dataset: To capture users' emotions, we collect users' basic information (e.g. their age, gender, location and healthcare), behaviors information (e.g., activity level, facial expression) and cyber data (e.g., call and SMS logs, application usage). Furthermore, in order to obtain users interested in video category, we also collect users' video browsing history. In the emotion recognition experiments, we choose 20 users and 40 videos to do the experiments.

In the experiment, each of the above 20users watches 40videos, so we can obtain the emotions of the users. Thus, we can get the MMVQ when the user watches video. In addition, we can use the personalized QoE model to predict the video QoE for other users. In order to evaluating the personalized QoE prediction model, each user further watches 20 videos selected among the remainder 110 videos and obtain their QoEs which can be regarded as ground truth.

Evaluation Methodology

Since the personalized QoE model can predict user video QoE, in order to evaluate the user video QoE model, we compare four measurements: accuracy, precision, recall and F1-Measure[15]. Accuracy is denoted as the number of correctly inferred video QoEs divided by the number of total samples. Precision refers to the percentage of correct QoE predictions made by the personalized QoE model. Recall is the percentage of video QoEs detected. F1-Measure is the weighted harmonic mean of precision and recall, and represents the overall performance.

Performance Evaluation

In this subsection, we give the system performance evaluation, including EQA metric MMVQ evaluation, user personalized video QoE prediction and decision tree-based adjustment model evaluation.

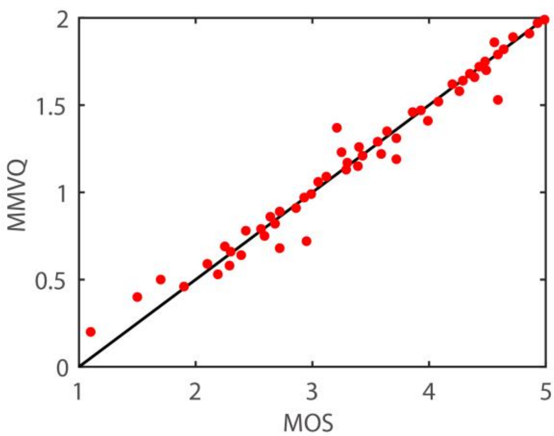

Figure 4. Scatter plot of MOS versus MMVQ.

EQA metric MMVQ evaluation:

In order to evaluate the EQA metric, we compared the MMVQ with MOS. This is because the MOS is the user's subjective test. In Figure 4, the relationship between MOS and MMVQ are drawn. From the figure, we can see that the MMVQ is highly correlated with subjective test, which shows that MMVQ metric is valid.

Furthermore, in order to verify the validity of MMVQ, we play a 100 minutes of comedy, which conform the users' interest, under the condition of good network environment. Then we compare

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

the user's mood score to the video content emotion, where MMVQ means the difference between the two scores. As shown in Figure 5, there is little fluctuation in MMVQ, which indicates that the evaluation based on MMVQ is accurate.

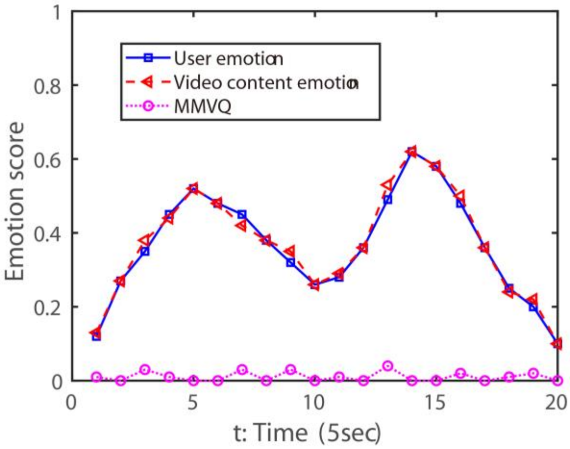

Figure 5. MMVQ metric evaluation.

User personalized video QoE prediction

In order to evaluate user personalized video QoE prediction model, we compared with two widely-used models with high performance: (1) Predictive model (PM) [8], (2) Improved predictive model (IPM). PM[8] use of decision tree to predict user's QoE. Decision tree is one of the most widely-used machine learning models for data-driven video QoE analysis, and it divides the data set into several subsets according to the characteristics selected through purity detection. IPM uses random forest method, which employs the integrated learning models based on the decision tree, and it obtains better performance by using the integrated and stimulation methods.

Table 2. Comparison of accuracy, precision, recall and F1-measure among predictive model, improved predictive model and personalized model.

Model	Accuracy (in %)	Precision (in %)	Recall (in %)	F1-Measure (in %)
Predictive model	61.2	60.5	62.8	61.6
Improved predictive model	63.5	65.2	60.2	62.6
Personalized QoE model	72.1	73.2	69.2	71.1

However, the above two models do not take into account the diversity of user and user's mood when predicting users' QoE. While our model obtains the individual model for each user through shared user-video, interactive information, and a great amount of mood information, which enables us to make significant improvement. Table 2 compares accuracy, precision, recall and F1-measure among PM, IPM and our EQA model. It can be seen that the accuracy rate of predicting video QoE of user by IPM is 2.3% higher than that by PM. This is because the IPM is used to construct a more ideal model using all the power of a set of weak learning machines, and their performance is just marginally better than the PM. Under all metric including accuracy precision, recall and F1-measure, the proposed EQA model outperforms other comparative models. This is because in PM and IPM, all users are represented as a unified whole, thus neglecting the varieties of users' behaviors and emotions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

Decision tree-based adjustment model evaluation

This model introduce the nonlinear relationship among video QoE, average bitrate and buffer ratio. As shown in the Figure 6, when MMVQ is low, i.e., the user of QoE is low, while buff ratio and average bitrate of video are high. This indicates that user is not interested in the video. Thus, it should be change video content, and recommend videos that users are interested in. The specific decision process is shown in Figure 6.

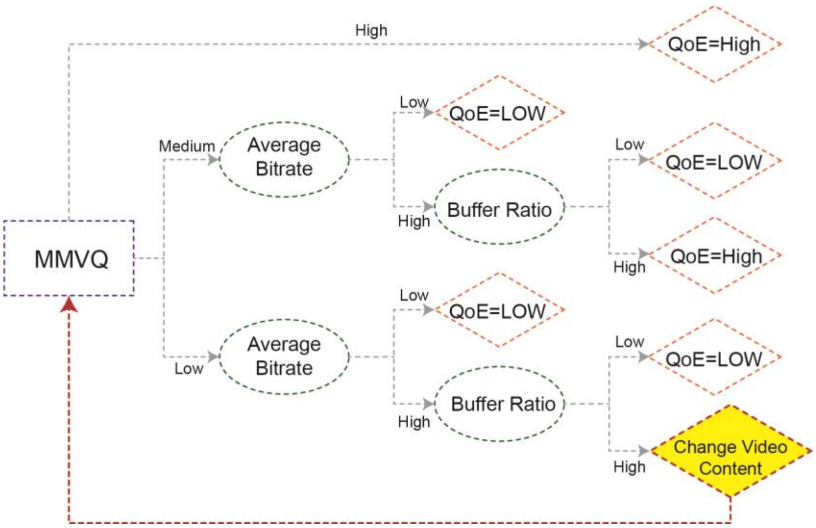

Figure 6. The video QoE adjustment model.

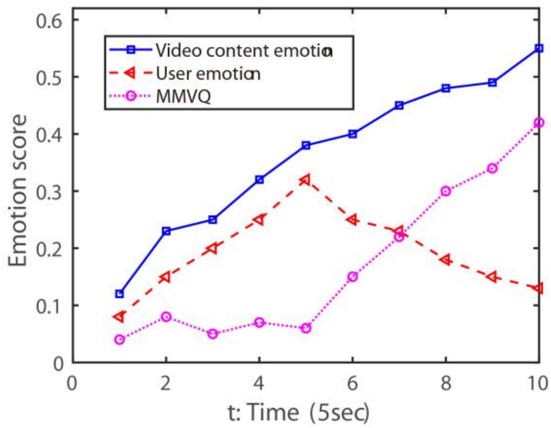

Figure 7. MMVQ before decision tree-based adjustment model.

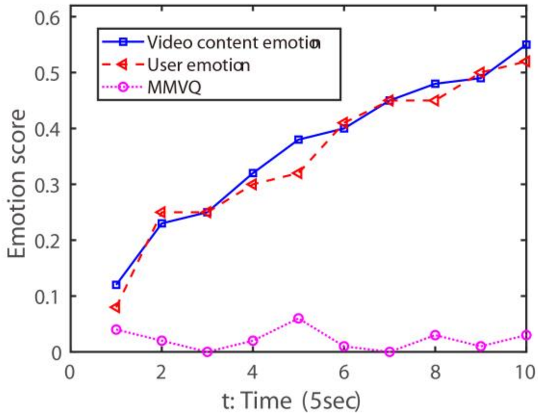

Figure 8. MMVQ after decision tree-based adjustment model.

To evaluate the decision tree model, we experiment with a comedy with a duration of 50 minutes. The effect of comedy is from weak to strong. We divide the users into two groups randomly, one group does not use the decision tree-based adjustment model, the other group uses the decision tree-based adjustment. Figure 7 and Figure 8 depicts the average values of each group. From the Figure 7 and Figure 8, we can see that the emotion score of the video content increase, this is because the effect of video comedy is from weak to strong. Furthermore, from the Figure 7, there is one sharp changes in the MMVQ, this is because the users are not interested in the video content. However, from the Figure 8, the user's emotions are very consistent with video content emotions, i.e., there is no sharp changes in the MMVQ. This is because when the user's QoE is not good, the adjustment model will be adapted to achieve better QoE.

CONCLUSION

In this article, we regard emotion as an important indicator of video QoE evaluation, and propose emotion-aware video QoE assessment model. Specifically, first, based on transfer learning, we can get users' emotions and use similarity model to match moods for video QoE. Then, we introduce video QoE predictive model by consider the diversity of the users and use adjustment model to illustrate the nonlinear relationship among average bitrate, buffer ratio and QoE. Finally, in contrast to conventional video QoE assessment methods, i.e., decision tree and random forest, EQA model has better accuracy. With the development of deep learning, the deep convolution neural networks have achieved excellent performance in extracting the features of image and speech. Thus, in the future work, we will consider using deep convolution neural network for user emotion recognition and predict personalized QoE.

ACKNOWLEDGMENTS

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia for funding this research group No. (RG-1437-042).

REFERENCES

1. L. Zhou, “QoE-Driven Delay Announcement for Cloud Mobile Media”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no.1, 2017, pp. 84-94.

2. L. Zhou, "Mobile Device-to-Device Video Distribution: Theory and Application", *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 12, no.3, 2015, pp. 1253-1271.
3. Y. Chen, K. Wu, and Q. Zhang, "From Qos to QoE: A Tutorial on video Quality Assessment," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, 2015, pp. 1126–1165.
4. M. Hemmati, B. McCormick, S. Shirmohammadi, "QoE-Aware Bandwidth Allocation for Video Traffic Using Sigmoidal Programming", *IEEE MultiMedia*, vol. 24, no. 4, 2017, pp. 80–90.
5. D. Ghadiyaram, J. Pan, A. Bovik, "Learning a Continuous-Time Streaming Video QoE Model," *IEEE Transactions on Image Processing*, vol. 27, no. 5, 2018, pp. 2257–2271.
6. Wang et al., "A Data-Driven Architecture for Personalized QoE Management in 5G Wireless Networks", *IEEE Wireless Communications*, vol. 24, no. 1, 2017, pp. 102–110.
7. F. Dobrian et al., "Understanding the Impact of Video Quality on User Engagement," *Communications of the ACM*, vol. 56, no. 3, 2013, pp. 91–99.
8. Balachandran, et al., "Developing a predictive model of quality of experience for internet video," Proc. *ACM SIGCOMM 2013*, 2013, pp. 339–350.
9. S. S. Krishnan and R. K. Sitaraman, "Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, 2013, pp. 2001–2014.
10. Zheng et al., "Cross-Domain Color Facial Expression Recognition Using Transductive Transfer Subspace Learning", *IEEE Transactions on Affective Computing*, vol. 9, no. 1, 2018, pp. 21–37.
11. Chen et al., "User Intent-oriented Video QoE with Emotion Detection Networking", Proc. *IEEE GLOBECOM 2016*, 2016, pp. 1–6.
12. S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.
13. Yang et al., "Social Collaborative Filtering by Trust", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, Aug. 2017, pp. 1633-1647.
14. Chen et al., "On Sampling Strategies for Neural Network-based Collaborative Filtering", in Proc. *KDD 2017*, 2017, pp. 767–776.
15. K. Hwang, M. Chen, Big Data Analytics for Cloud/IoT and Cognitive Computing, Wiley, U.K., ISBN: 9781119247029, 2017.

ABOUT THE AUTHORS

Yixue Hao (yixuehao@hust.edu.cn) received the B.E. degree in Henan University, China, and his Ph.D degree in computer science from Huazhong University of Science and Technology (HUST), China, 2017. He is currently working as a postdoctoral scholar in School of Computer Science and Technology at Huazhong University of Science and Technology. His research includes 5G network, Internet of Things, mobile cloud computing.

Jun Yang (junyang cs@hust.edu.cn) received Banchelor and Master degree in Software Engineering from HUST, China in 2008 and 2011, respectively. Currently, he is a Ph.D candidate at Embedded and Pervasive Computing (EPIC) Lab in School of Computer Science and Technology, HUST. His research interests include cognitive computing, software intelligence, Internet of Things, cloud computing and big data analytics, etc.

Min Chen [SM'09] (minchen2012@hust.edu.cn) has been a full professor in the School of Computer Science and Technology at HUST since February 2012. He is Chair of the IEEE Computer Society STC on big data. His Google Scholars Citations reached 10,800+ with an h-index of 52. He received the IEEE Communications Society Fred W. Ellersick Prize in 2017. His research focuses on cyber physical systems, IoT sensing, 5G networks, SDN, healthcare big data, etc.

M. Shamim Hossain [SM'09] (mshossain@ksu.edu.sa) is a Professor at the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He is the recipient of the 2016 ACM Transactions on Multimedia Computing, Communications and Applications (TOMM) Nicolas D. Georganas Best

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MMUL.2018.2879590, IEEE MultiMedia

Paper Award. He is on the editorial board of IEEE Multimedia, IEEE Network, IEEE Access, and Journal of Multimedia Tools and Application (Springer). His research interests include cloud networking, social media, IoT, cloud and multimedia for healthcare.

Mohammed F. Alhamid (mohalhamid@ksu.edu.sa) is an Assistant Professor at the Software Engineering Department, King Saud University, Riyadh, KSA. Alhamid received his Ph.D. in Computer Science from the University of Ottawa, Canada. His research interests include recommender systems, social media mining, big data, and ambient intelligent environment.