User Intent-oriented Video QoE with Emotion Detection Networking

Min Chen*, Yixue Hao*, Shiwen Mao†, Di Wu‡ and Yuan Zhou§
*School of Computer Science and Technology, Huazhong University of Science and Technology, China (minchen@ieee.org)
†Department of Computer Science, Auburn University, USA (smao@ieee.org)
‡Department of Computer Science, Sun Yat-sen University, China (wudi27@mail.sysu.edu.cn)
§School of Electrical and Information Engineering, Tianjin University, China (zhouyuan@tju.edu.cn)

***Abstract*—With the ever-growing number of users enjoying online video service in mobile environments, video streaming services have been dominating the mobile traffic. It can be predicted that a small improvement in the user's watching experience will cause a substantial leap in profitability in terms of content providers and distributors, network operators and service providers for mobile videos. Though recent years have witnessed effective efforts to improve a user's video quality of experience (QoE) by the use of big data for analyzing users' viewing behaviors based on large-scale, video-viewing history datasets, it is very challenging to precisely analyze users' hidden intents and feelings when they are watching online videos. In addition to obtain a better video QoE, we propose to introduce user's emotional reactions into QoE assessment. In this scheme, first, the user's mood is detected in a real time fashion via emotion detection networking. Then, a mood matching process is performed to gain the similarity of the user's intent and the video content property in terms of emotion design. Finally, a novel, decision tree-based adjustment model is proposed to characterize the relationship between QoE and various factors, including buffer ratio, average bitrate, and the user's emotions. Our study opens a road for improving video QoE based on emotion detection networking.**

***Index Terms*—Affective computing; Machine learning; Video Quality of Experience (QoE).**

I. Introduction

According to the report by Cisco [1], global mobile data traffic grew 69 percent in 2014, and mobile video traffic exceeded 50 percent of total mobile data traffic for the first time in 2012. Mobile video traffic exceeded 50 percent of total mobile data traffic by the end of 2012 and grew to 55 percent by the end of 2014, and it is predicted that nearly three-fourths of the world's mobile data traffic will be video related by 2019. The large number of mobile users implies that a substantial leap in profitability could be achieved with a small improvement in the user's watching experience. In order to provide a better experience of video service, we first need to answer a basic question: how can one evaluate video quality? In the early stages, video quality is quantified by service quality. However, since the video service is highly user-centered, it is difficult to achieve an objective standard for service quality which directly represents the user's perceptual experiences.

The experience quality of video should be measured by a subjective test, which can directly solicit the users' evaluation scores in a controlled environment; for example, mean opinion score (MOS) [2] is a common metric. But the cost of subjective tests is quite high. With the emergence of the big data analysis [3] and the availability of massive video concept shadow history datasets, data-driven video quality of experience (QoE) assessment (DDVQA) becomes popular [4] [5], the users' experience quality is usually obtained by the user's participation in the video service, and can be quantified by various measurement standards, such as the watching time ratio, the number of videos have been watched, and the video download rate [5]. However, such existing DDVQA metrics either have a limited accuracy, or do not take into account the diversity of users. Different from the above QoE metrics, we introduce user's emotional reactions into QoE assessment and we propose the emotion-assisted QoE assessment (EQA) model in this paper to address these shortcomings of the existing models.

In recent years, with the development of emotional computing and big data, the recognition and monitoring of people's emotions have been advanced [6]. When the users are watching a video, their mood comprises the most direct subjective test, which can be used to assess the QoE. We judge the experienced quality of the video by comparing the user's emotional state when watching videos with the video content. The similarity between the users' emotions of video watching and the video content is closely related to users' video QoE. Therefore, the metric of emotion-assisted QoE can be established. Next, we establish the user intent-oriented video based on the watching history of user, especially for the video topic. The user intent-oriented video refers to the video content as being recommended which users are interested in when users' QoE becoming lower. Finally, we use a decision tree to improve the users' QoE according to the video quality.

In summary, the main contributions of this paper include:

- A new metric of video QoE is proposed, i.e., the metric of emotion-assisted QoE assessment.
- An user intent-oriented video can then be created, which, according to the users' real-time QoE variations, will allow users to choose whether or not to switch to another video in which they are interested, before finish viewing the current video.
- Based on the metric of emotion-assisted, we propose a decision tree-based adjustment model which considers video QoE and various impact factors.

TABLE I
PHYSICAL AND CYBER DATA

Class	Type	Components
Physical Data	Physiological Data	ECG
	Activity Level	Static, walking and running
	Location	Latitude, Longitude
Cyber Data	Call and SMS Logs	Number, Duration
	Applications Usage	Type, Number

The remainder of this article is organized as follows. The system overview is described in Section II. We present the emotion-assisted video QoE metric in Section III and propose the EQA model in Section IV. Our experimental results and discussions are given in Section V. Finally, Section VI concludes this paper.

II. System Framework

In this section, we present the proposed framework, as shown in Fig. 1. The traditional QoE is based on the ability of the network to adjust the video coding and decoding strategy and network transmission strategy to improve the user's QoE. This paper asserts that even using the traditional QoE optimization method, it is still difficult to accurately reflect the users' true intention. Therefore, in this paper, we propose a QoE assessment system for streaming video, which is based on emotion detection. The general idea is to assess the mood of the user and the video and then compute the similarity between the two. If mood of video and user do not match, but no video streaming problems can be detected, a change in video content is proposed which based on users' real-time mood and the viewing history. We further propose a decision tree-based adjustment model which considers video QoE and various impact factors.

Specifically, our framework is divided into the following aspects: Firstly, we can sense user emotion and video content emotion by transfer learning and questionnaire respectively. Then we give the emotion-assisted QoE assessment metric. Secondly, we utilize singular value decomposition (SVD) model to recommend a video to users based on viewing historical records. Finally, we select important factors of impacting QoE by a decision tree, and adjust each factor to enable the user to achieve better QoE.

III. Emotion-assisted QoE Assessment (EQA) Metric

In this section, we present the EQA metric. In order to establish this metric, firstly, we employ the user's mood and content emotion video's judgment. The users' mood is derived with transfer learning, using a hidden Markov model (HMM) for classification, while the video content mood is obtained with a questionnaire. Secondly, we show how to determine their similarity.

A. User Emotion Detection

1) Emotional Data Collection and Feature Extraction: Because the metric of EQA requires real-time personal emotion, a real-time model is needed. The real-time emotional data includes users' electrocardiogram (ECG) signal collected by wearable devices, such as smart clothing, and users' life and behavior habits collected by mobile phone. We divide the data into physical and cyber data. Physical data include physiological data, activity level, and location. Cyber data consist of calls, SMSs, application usage. Table I shows the data collection in detail. We define an emotional space label as $L = \{happy, relaxed, afraid, angry, sad, bored\}$. The symbol c is defined to represent the corresponding labels. As for emotional data labels, based on our previous work [6], we design a phone application of labelling emotion which can be utilized for users to label their emotions.

According to the collected emotional data, we first introduce the preprocessing of the data, which include data cleaning, data integration and eliminate redundancy. Then we extract the data features. For EGG signal, discrete wavelet transform (DWT) based scheme is used for feature extraction. For location data, we obtain the position information of user by density-based spatial clustering of applications with noise (DBSCAN) cluster. For the cyber data, considering it is statistical data, we count the number as its feature.

2) Automatically Label based on Transfer Learning: Since users need to label their emotion by themselveswhich is the time-consuming and labor-intensive, we utilize transfer learning to label users' emotion autonomously, i.e., we only need to label some users' collected emotional data, other users' data can be labelled by transfer learning. The transfer learning [7] can be described as follows, Let denoted χ_s to be the source instance space, i.e., the data collected which have emotion label, and χ_t to be the target instance space, i.e., the data collected which don't have emotion label. F_s and F_t is the feature space correspond with χ_s and χ_t, respectively. L is the label space. The transfer learning model can represented as $x_t \to f_t \to f_s \to c$, where $x_t \in \chi_t$, $f_t \in F_t$, $f_s \in F_s$ and $c \in L$.

Now we introduce emotional autonomously labels in detail, the source domain input data x_s comes from the users' emotional data in terms of physical data, cyber data which have emotion label. The target domain input data x_t comes from other personal emotional data which do not have emotion label. Our goal is to estimate the mood c probability $p(c|x_t), c \in L$. Since x_s and x_t may be in a different feature space, we first need to find a transformation $\phi(f_t, f_s) \propto p(f_t|f_s)$ to link the two feature spaces [7]. Since f_s and f_t are features that are independent to x_s, where $x_s \in \chi_s$, we can calculate $p(f_s, f_t)$ as follows.

$$p(f_t, f_s) = \int_{\chi_s} p(f_t, x_s) p(f_s|x_s) dx_s. \quad (1)$$

Since we know $p(f_s|x_s)$, so we need to calculate the $p(f_t, x_s)$. Now we can link the unlabeled feature f_t and x_s through Jensen-Shannon divergence [8] (the symmetrized

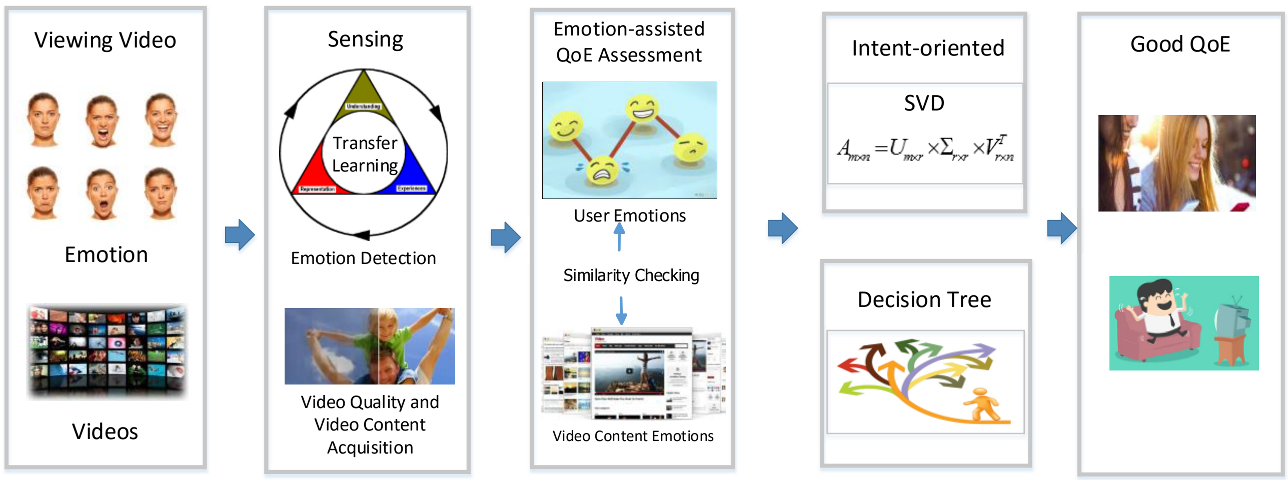

Fig. 1. Functional components in the proposed architecture for improving user's video QoE via EQA (Emotion-assisted QoE Assessment).

and smoothed version of Kullback-Leibler (KL) divergence), which can be described as follows.

$$\mathrm{JSD}(P||Q) = \frac{1}{2}(D_{KL}(P||M) + D_{KL}(Q||M)). \quad (2)$$

where $M = \frac{1}{2}(P+Q)$ and $D_{KL}(\cdot)$ is the KL-divergence which can be as $D_{KL}(P||Q) = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}$. P and Q are the probability distribution of x_s and f_t. Since the Jensen-Shannon divergence is widely used for measuring the similarity between two probability distributions, $\mathrm{JSD}(P||Q)$ equals to zero if and only if the two distributions P and Q are identical. so we take the low-k similar distributions out. Now we link the f_t and f_s, so we can calculate the most probability label of x_s.

3) Label Validation: We can validate the labels are right or not based on transfer learning. In the validation phase, we detect users' input of their own emotional reactions when they use some applications such as Moodagent and Facebook. One major problem is that, the mood space $M = \{m_1, \cdots, m_n\}$ we detect from applications is not the same as the mood label space L. Thus, we need to compare the similarity of the label space L with the user's input mood m, which is collected in each time slot as ground-truth label. Using a framework similar to transfer learning [9], we can identify the similarity between mood c and m as follows.

$$\mathrm{sim}(c,m) = \mathrm{MMD}^2[D_c, D_m], \quad (3)$$

where $D_c = \{y_i | i = 1, \cdots, n_c\}$ and $D_m = \{z_i | i = 1, \cdots, n_m\}$ are a set of documents that represent a mood in C and M, respectively, and y_i and z_i is tf-idef vector [10]. The term $\mathrm{MMD}^2[D_c, D_m]$ denotes the maximum mean discrepancy [11]. When calculating the similarity between L and M, we validate the label when $\mathrm{sim}(c,m) > thr$, where thr is pre-defined threshold.

4) Emotion Detection: Through transfer learning, the time-consuming and labor-intensive labelling can be simplified extensively. After a certain time of labeling and validation through transfer learning, the training sets are established. Employing a Hidden Markov model (HMM), the emotional data are classified into six moods. The accuracy rate of our model becomes increasingly high with the increased amount of data. After a period of time, we can obtain a more accurate detection model. Eventually we obtain a real-time emotional model.

B. Mood Matching for Video QoE (MMVQ) based Similarity Checking

Through above methods, we can recognize in real-time users' emotions. We then need to identify the emotional attributes of the video content. Due to the difficulty of identifying the video content emotion, we use statistical rules, primarily on the basis of two approaches.

- Video's inherent label: also known as the type of video, such as comedy or adventure action.
- Using questionnaire: many people make emotion labels on the same video every 5 minutes. The label space is still L, so one can obtain the emotion of the video content.

From the above, we know both the user's emotion and that of video content. We then determine their similarity as follows. During the time when users watch video, derive the users' mood and the emotional attributes of video content, and estimate their similarity every 5 minutes. The similarity criterion is given by the following formula, namely, L, for the above conditions. We use the SentiWordNet [12] dictionary to score the emotional words in L about the video content and those of mood. So the emotional words is transformed into scores in $[-1, 1]$. We then compute the matching degree of the users' mood and the emotional attributes of the video content as follows.

$$d_i = |S_u^i - S_v^i|, \quad i = 1, 2, \cdots, m. \quad (4)$$

where S_u^i is the score showing the users mood and S_v^i is the score representing video content in time slot i, which can cause

d_i to change during video viewing; the greater the d_i, the poorer the matching degree. We set two thresholds, $threshold1$ and $threshold2$, where $threshold1 < threshold2$. We define video matching degree as high, medium, and low. If $d_i \leq threshold1$, it is considered that there is a good match between the user's mood and the video's emotional attributes. So the user has a good video QoE. If $threshold1 \leq d_i < threshold2$, it is considered that the user's mood and content video's emotional attributes almost match. If $d_i \geq threshold2$, it is considered that the user's mood and the emotional attributes of content video do not match. Based on this, we establish an emotion-assisted video QoE assessment metric.

IV. Emotion-assisted QoE Assessment (EQA) Model

Because the experience quality is quite subjective, it is determined by the individual preference of the user, including video content and video quality. Even if the user is given the same network status, video attributes and viewing environment, the experience quality of different users will be greatly different. That is, for each user, according to the above matching degree of similarity between the user's mood and the emotional attributes of video content, there are below two results. If the user's mood and content video's emotional attributes match, the user's QoE is then indicated to be high; otherwise the user's QoE is low.

According to the above criterion, we first show how to enable user intent-oriented videos; namely, when a user's QoE becomes low, we can recommend the user the video in which they may be interested, through the viewing history with the video content catering to the user's interest. Second, we adjust the video QoE based on the decision tree.

A. User Intent-oriented Video QoE

During the viewing process, the video content may not conform to the user's interest, because the content of the video is measured subjectivity, while the user's interest is measured by the user's viewing history. Therefore, we select a recommended video based on the historical data (that the user watched before). Because the theme of each video has a corresponding label, for example, comedy, action, youth, etc., we can use singular value decomposition (SVD) and latent semantic indexing (LSI) to determine which themes the user is interested in, and thus recommend the themes that the user prefers in real-time according to his/her viewing history. Specifically, the proposed scheme works as follows.

Assume that the system serves m users, denoted by $\{u_i | i = 1, 2, \cdots, m\}$, with s videos in n topics $\{t_i | i = 1, 2, \cdots, n\}$. We also define the index variables a_{ij} as follows. If a user u_i watches videos with topic t_j, then we have $a_{ij} = 1$; when the user does not watch the video with topic t_j, we have $a_{ij} = 0$. Thus we obtain the following $n \times m$ matrix, denoted by $\mathbf{A} = (a_{ij})_{m \times n}$.

Assume that the order of matrix $\mathbf{A}$ is $\text{rank}(\mathbf{A}) = r$, so that the matrix can be decomposed with SVD as

$$\mathbf{A} \approx \mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^T. \tag{5}$$

where $\mathbf{U}$ is an $m \times r$ matrix with orthogonal vectors (i.e., left singular vectors), $\boldsymbol{\Sigma}$ is an $r \times r$ diagonal matrix. The diagonal elements in $\boldsymbol{\Sigma}$ are called singular values satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. In (5), $\mathbf{W}$ is an $r \times n$ matrix with orthogonal vectors, which are called right singular vectors, and $\{\cdot\}^T$ is the matrix transpose operation. This way, matrix $\mathbf{A}$ can be decomposed into the product of three matrices, i.e., $\mathbf{U} = (u_{i,j})_{m \times r}$, $\boldsymbol{\Sigma} = \mathbf{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$, and $\mathbf{W} = (w_{i,j})_{r \times n}$, which are given below.

Next, we use a LSI to determine the topic t_j in which user u_i is interested. Each row vector in $\mathbf{U}$ stands for certain characteristics of a user, while each column vector in $\mathbf{W}$ stands for certain characteristics of a topic. The singular value matrix $\boldsymbol{\Sigma}$ in the middle stands for the relationship between each row in $\mathbf{U}$ with each column in $\mathbf{W}$. We can select two dimensions and let the first two columns in $\mathbf{U}$ be $\mathbf{u}$ and the first two rows in $\mathbf{W}$ be $\mathbf{w}$. That is,

$$\mathbf{u} = \begin{bmatrix} u_{11} & u_{12} \\ \vdots & \vdots \\ u_{m1} & u_{m2} \end{bmatrix} \tag{6}$$

$$\mathbf{w} = \begin{bmatrix} w_{11} & \dots & w_{1n} \\ w_{21} & \dots & w_{2n} \end{bmatrix}. \tag{7}$$

We will project $\{(u_{11}, u_{12}), \cdots, (u_{m1}, u_{m2}), (w_{11}, w_{21}), \cdots, (w_{1n}, w_{2n})\}$ onto a two-dimensional surface to obtain the relationship between user and topic, and determine which users are interested in which topics with the clustering method. Therefore, we can recommend the topics of interest to the users, thus enhancing the QoE of users.

B. Decision Tree-based Adjustment Model

If the user has a low QoE and once the content of the video is consistent with the topic in which the user is interested, we may infer that the video quality has problems. As for the influence of video quality, the linear correlation of Pearson, the rank correlation of Spearman, and the rank correlation of Kendall show that no single factor has an obvious linear relation or monotonic relationship with the experienced quality of the users, so we can conclude that the relationship between the experience quality and influential factors is not linear or monotonic.

We therefore cannot assume that the influential factors are independent of each other. For example, high coding bit rate may result in a long buffer time. The decision tree is a non-parametric model. It does not assume the linear or monotonic relation between the experience quality and other factors, nor does it suppose the influential factors are independent of each other. Thus, we use decision tree to predict the QoE and to propose the adjustment model. The following influential factors of the experience quality: buffer time and average bitrate are considered. Other extra factors, including location, time of day or day of week, and so on, will not be considered for now. Adopting the method mentioned above, we use factors which influent the video QoE assessment as the classification attribute and denote those as S. We discretize the attributes into {low, high}. The cluster label of videos as C_i, $i = 1, 2, 3$, and set $C_{i,S}$ as a cluster of C_i tuples in S. $|S|$ and $|C_{i,S}|$ as the number of tuples in S and $C_{i,S}$, respectively. We determine which factor takes a greater proportion by the information gain. The information gain is defined based on the concept of entropy as follows $H(S) = -\sum_i P_i \log P_i$. where $P_i = |C_{i,S}|/|S|$ is the non-zero probability of C_i. The expected information required for tuple classification of S according to attribute $\mathbf{A}$ is $H_A(S)$. Then we have

$$H_A(S) = \sum_{v \in V(A)} \frac{|S_v|}{|S|} H(S_v), \tag{8}$$

where v stands for the $v-th$ subset divided from S according to attribute $\mathbf{A}$. We thus obtain the information gain as follows: $Gain(S, A) = H(S) - H_A(S)$. With the information gain, we can determine which factor has the greatest influence and mediate it to enable the users to obtain higher QoE.

V. Performance Evaluation

A. Datasets

We collect a flow video dataset which contains 100 comedy and 50 other category videos from Youtube. We also label video content emotion through 200 participants, including undergraduate and graduate students. Their ages range form 18 to 45. Each participant watches 30 videos and label the videos' content emotion. Furthermore, they give their subjective scores. The data set has been anonymized with all identifiers related to the clients made anonymous according to strict guidelines.

For user's emotions, the information and use behavior of 20 users, such as population statistical information including age, gender, and address, and other physical and cyber information and so on is collected. In order to get the video categories which users are interested in, we also collect 20 users' viewing history. Since we can get 20 users mood and 150 videos content emotion, when a user watch the video, we can get the MMVQ. For the video, we collect average bitrate and buff ratio of video. Based on the varieties of users contained in the above dataset, we design the experience to verify that different users have significantly different experience quality models, and believe our method can obtain better experience quality assessment.

B. Performance Evaluation

1) **EQA metric evaluation:** Firstly, we give the accuracy of the mood detection. As shown in Fig. 2, we can see the accuracy of our mood recognition model in more than 70% and happy is the highest one.

Secondly, we evaluation the QoE based on similarity of mood of user and video. Because the basic criteria of video QoE are users' subjective test, we describe the relevances of MMVQ and subjective test. Since the values of MMVQ and subjective test belong to [0,2] and [0,5], respectively, we give the different values normalization processing which can make

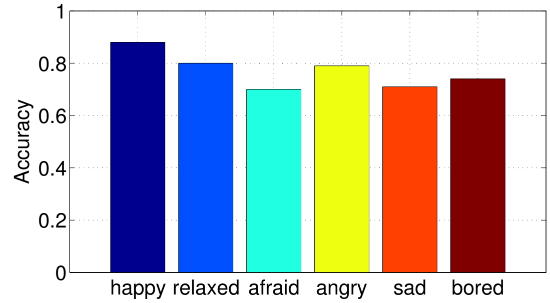

Fig. 2. Accuracy of emotion detection.

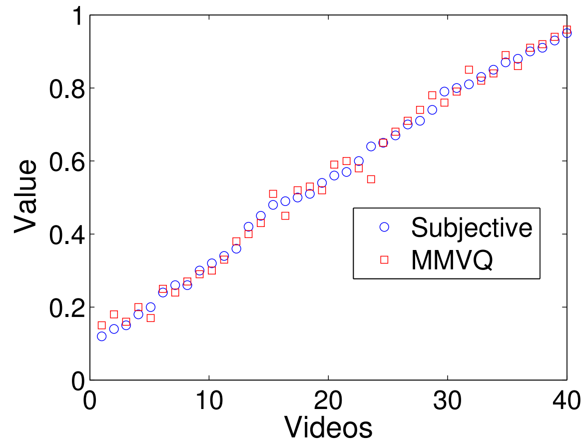

Fig. 3. Performance comparison among subjective and MMVQ assessment method.

them belong to [0,1]. As shown in Fig. 3, we can conclude the differences between values of MMVQ and subjective test are small based on normalization processing, which can illustrate our MMVQ model is effective.

Thirdly, in order to evaluate the effectiveness of measurement, we will play comedy for 100min based on ensuring the users' interest on network and video content, and compare the score of users' mood and the score of video content mood,

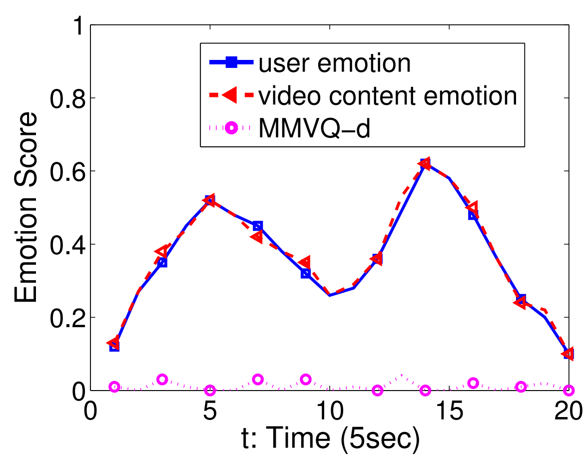

Fig. 4. MMVQ metric.

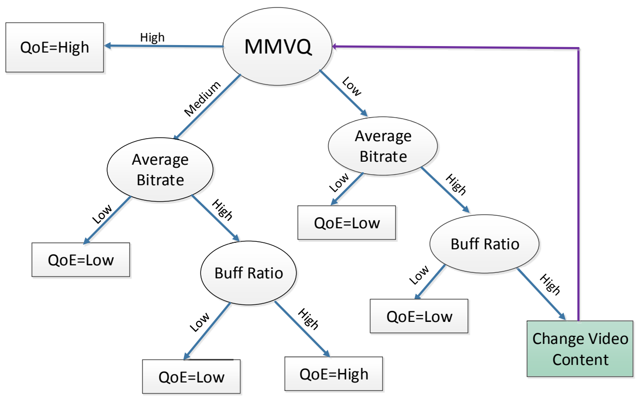

Fig. 5. The decision tree-based adjustment model for video QoE.

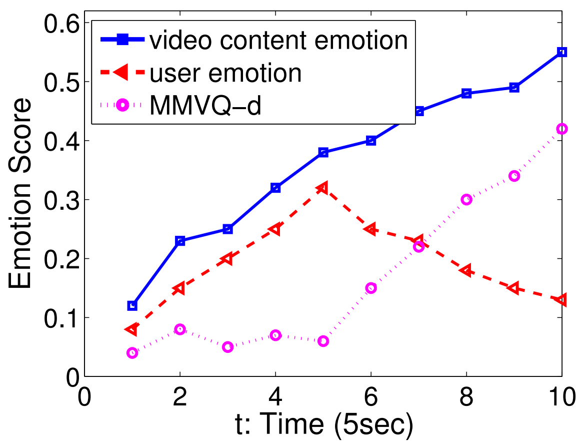

Fig. 6. Mood matching for video QoE before adjustment.

we can see d is slightly fluctuated as shown in Fig. 4, which means that the evaluation is more accurate based on mood.

2) **EQA model evaluation:** We proposed a adjustment model based on the decision tree as shown in Fig. 5. From that, we can figure out the relationship between QoE and various factors. As shown in the Fig. 5, when MMVQ of user is very low, namely QoE of user is very low, meanwhile average bitrate and buff ratio are very high, that's because the user is not interested in the content of the video, then videos the user interested in shall be recommended. For further demonstration, we play one 50-min comedy short film for test and randomly select ten users who watch the video without adjustment, and another 10 users who are served with adjustment in real-time according to the situation of the user. As shown in Fig. 6 and Fig. 7, the value in the figure is the mean value of 10 users. We can draw a conclusion that the video content emotion scores become large over time, because the user with bad video QoE will achieve a recommended video to obtain a good video QoE through adjustment.

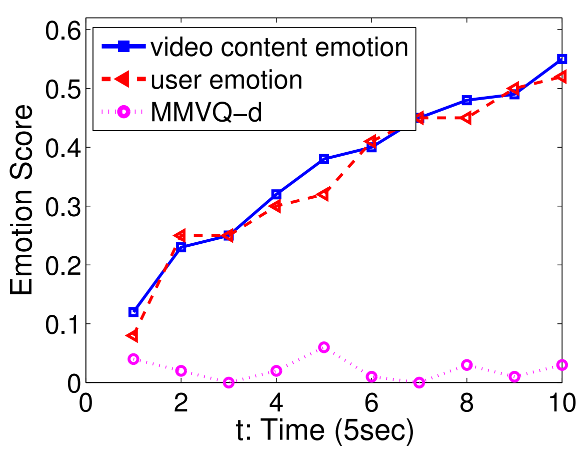

Fig. 7. Mood matching for video QoE after adjustment.

VI. Conclusion

In this paper, we propose emotion-assisted QoE assessment (EQA) model which can analyze users' hidden intents. In this model, we innovatively introduce user's emotion for the QoE assessment. First, user's real time emotion is detected via emotion detection networking. Then, mood matching for video QoE (MMVQ) based similarity checking is performed and we give emotion-assisted QoE assessment metric. Second, we introduce user intent-oriented video content model based on SVD. Third, we analyze the various factors including buffer ratio, average bitrate impact to the user's video QoE. In future work, we will consider more factors (e.g., various types of devices) which have impact on user's video QoE.

References

[1] *Cisco Visual Networking Index: Global mobile data traffic forecast update, 2014-2019, Feb , 2015.*

[2] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, and L. Hanzo, "Quality-of-experience assessment and its application to video services in lte networks," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 70–78, 2015.

[3] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.

[4] H. Yin, Y. Jiang, C. Lin, Y. Luo, and Y. Liu, "Big data: transforming the design philosophy of future internet," *IEEE Network*, vol. 28, no. 4, pp. 14–19, 2014.

[5] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "Developing a predictive model of quality of experience for internet video," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4. ACM, 2013, pp. 339–350.

[6] M. Chen, Y. Hao, Y. Li, D. Wu, and D. Huang, "Demo: Lives: Learning through interactive video and emotion-aware system," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 399–400.

[7] W. Dai, Y. Chen, G.-R. Xue, Q. Yang, and Y. Yu, "Translated learning: Transfer learning across different feature spaces," in *Advances in neural information processing systems*, 2008, pp. 353–360.

[8] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991.

[9] D. H. Hu, V. W. Zheng, and Q. Yang, "Cross-domain activity recognition via transfer learning," *Pervasive and Mobile Computing*, vol. 7, no. 3, pp. 344–358, 2011.

[10] C. D. Manning, P. Raghavan, H. Schütze *et al.*, *Introduction to information retrieval*. Cambridge university press Cambridge, 2008, vol. 1.

[11] K. M. Borgwardt, A. Gretton, M. J. Rasch, H.-P. Kriegel, B. Schölkopf, and A. J. Smola, "Integrating structured biological data by kernel maximum mean discrepancy," *Bioinformatics*, vol. 22, no. 14, pp. e49–e57, 2006.

[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.