ISSN 2278-3091
Volume 10, No.1, January - February 2021
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse481012021.pdf
https://doi.org/10.30534/ijatcse/2021/481012021

A machine learning based approach for frame work of objective video quality assessment system

Ganesh K.[1], C. M. Patil[2]
[1]Assistant Professor, Dept of Electronics & Communication Engg, BIET, Research Scholar, VVCE, Mysuru, Visvesvaraya Technological University, India, digitalgk2020@gmail.com
[2]Professor & Head, Dept of Electronics & Communication Engg, Mysuru, India, patilcm@gmail.com

ABSTRACT

Video quality assessment aims to predict viewer's opinion through objective means. A single video quality metric is not sufficient to predict and quantify the test video. Hence, more video quality metrics have to be used for the quantification of video quality. So, hybrid quality metrics are required for quantification of video quality. In this paper, we propose a quality assessment method using machine learning algorithm. The proposed method performs better than other classification techniques.

Key words : About four key words or phrases in alphabetical order, separated by commas.

1. INTRODUCTION

With the entry of Covid 19 virus in the world, the lockdowns have increased in many countries. These lockdowns have increased the demand for multimedia consumption. Multimedia consumption over the internet have substantially increased in recent year. There is increasing trend towards Over The Top(OTT) services like Amazon prime. Viewers of these multimedia platforms or services are becoming more and more video quality conscious. This has created a challenge and opportunity for engineers and scientists to design and develop algorithms which can predict the viewer's opinion about the quality of video. Objective video quality assessment can address this issue.

Video Quality Assessment (VQA) is a major research area which aims to design algorithms and to evaluate objective scores well correlated with subjective scores given by the human. Assessment of image quality metrics are also applied to evaluate the quality of the video by using temporal pooling. Image and video quality analysis plays a key point to assess the algorithms in process like enhancement, compression, reconstruction etc. LIVE (Laboratory for Image and Video Engineering) video dataset is utilised for the analysis of quality metrics for the proposed technique.

There are two types of video quality analysis i.e. subjective and objective. Subjective quality analysis is the one in which human observers are shown the reference and test videos and are asked to grade them. In objective video quality analysis method, mathematical models are built to automatically predict the quality of video which is correlated with human observer quality score. Based on the availability of reference video, video quality metrics are categorised into Full Reference (FR), Reduced Reference (FR) and No Reference (NR) video quality metrics. In our implementation, we have used FR video quality analysis using LIVE video database.

LIVE Dataset

The LIVE Video Quality database(Seshadrinathan, Soundararajan, Member, Bovik, & Cormack, 2010) (Seshadrinathan, Soundararajan, Bovik, & Cormack, 2010) is created by University of Texas at Austin. It comprises 10 uncompressed high quality videos. These videos are different contents as reference videos. Fifteen distorted videos per reference video with different distortions are created from these reference videos. So, it has set of 150 distorted videos. Four types of distortions are used. MPEG-compression, H.264, H.264 video through error prone IP network and error prone wireless network for H.264 compressed video bit stream.

To create distortion video, distortion strength is manually adjusted so that the various distorted videos are separated by distortion perception levels. For the creation of LIVE video database, 38 human subjects participated in the experimentation. A stimulus study is conducted to process each videos. In continuous quality scale is used to subjective score of the videos and removed the hidden references. LIVE database contains features like Differential Mean Opinion Score (DMOS)'s mean, variance of subjective evaluations and reference. For the analysis of video quality scores and to predict the video quality score, we have used data mining based approach.

2. BACKGROUND

2.1 Machine Learning

Machine learning is a popular technique in this era. It has various learning strategies like supervised, unsupervised, semi-supervised and reinforcement are few of them.

2.2 Supervised learning

The input dataset elements are related to predict the output. These are based on dependent and dependent variable. Linear regression is the popular example in this category.

2.3 Unsupervised learning

It does not have information about the expected output. It finds patterns from the input data.

2.4 Reinforcement learning

Based on the environment it either provides reward and punishment. An algorithm that can learn to play a game by playing and receiving feedback on its performance (victories, good and bad moves, etc.).

In this paper, we have used supervised learning mode for the prediction of quality of a video.

In addition, there are different types of machine learning algorithms available based on the system output and purpose:

- Regression: The continuous magnitude value is estimated by the system.
- Classification: It assigns a category (from a finite set) to the inputs.
- Clustering: It divides the inputs into groups. Here, groups are not known in advance.
- Dimensionality reduction: Reduce the dataset size based on either feature extraction or feature selection.

2.5 Regression

Regression analysis is a statistical framework that is used to estimate the strength and direction of the relationship between two or more variables. Simple regression analysis is used to estimate the relationship between a dependent variable (Y) and an independent variable (X). Multiple regression analysis is used to estimate the relationship between a dependent variable and two or more independent variables. We typically think of the independent variable as something we are
trying to predict and the dependent variables as quantities we can measure.

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Where

$\hat{Y}_i$ is estimated value of Y_i

$\hat{\beta}_0$ is estimated value of β_0

$\hat{\beta}_1$ is estimated value of β_1

We have used dimensionality reduction based on feature selection. In our case, features are the video quality metrics.

Generally, machine learning models are referred as predictive models especially for the particular applications. The word prediction is the nothing but estimation. In the field of VQA, the model can able to predict quality in the sense that is able to estimate how human users would score the video quality. There are different models which are used in this work such as Decision tree, naïve bays, SVM, KNN and Ensemble classifier for VQA

3. METHODOLOGY

In this section, we explain the methodology of the proposed methodology as shown in the figure 1. In our method, raw YUV420 videos of LIVE database are utilised for the experimental analysis. Here, 'Y' denotes the brightness, or 'luma' value, and 'UV' denotes the color, or 'chroma' values. Since we have not considered color in our quality analysis, we are accessing only luminance component of the video. Y component of reference frame of reference video of dataset and Y component of distorted video(test video) frame are considered for objective video quality scores like SSIM, CWSSIM etc.

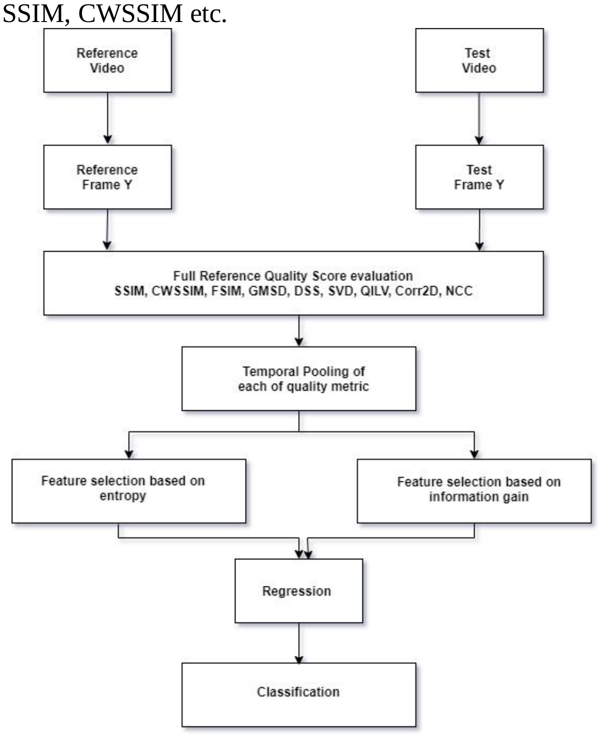

Figure 1: Proposed methodology

We have considered first 100 frames of these datasets for our analysis. Hence, 100 quality scores for each quality metric are obtained. For instance, 100 frames of bs2_25fps.yuv are considered for pooling. BS data set has 16 types of distorted test videos. For the pair of each distorted video with reference video of the dataset, quality score evaluation is done. Scatter plot of quality scores has been depicted in Figure 2. Temporal pooling is applied for 100 temporal values of quality scores of each of 9 quality metric to obtain the quality metric value. This process is done for all the calculated quality scores of distorted videos of the dataset. Averaging is used for pooling the quality scores.

Feature selection is done based on the entropy of the video quality metric in each category. Among the 9 quality metrics, 4 metrics with entropy more than value 4 are being selected for the classification. In information theory, entropy of a random variable is the average level of information in variable's possible outcomes. Entropy of a quality score whose information content is more is considered in our case. Here, we have selected the features (quality metrics) based on entropy of quality scores of the quality metric.

These selected features are used to train the machine learning algorithms. Finally the accuracy of different implementations are taken into consideration.

4. RESULTS AND DISCUSSIONS

In this section, we present the results by using the proposed method. Experimental analysis is performed with a system with 8GB RAM and 1TB hard disk. Experiments are conducted by using MATLAB 2020. Y axis is MOS of the quality score and X axis is the objective score of individual quality metric as indicated in Fig.2. For experimental analysis, we used LIVE database to measure the video quality.

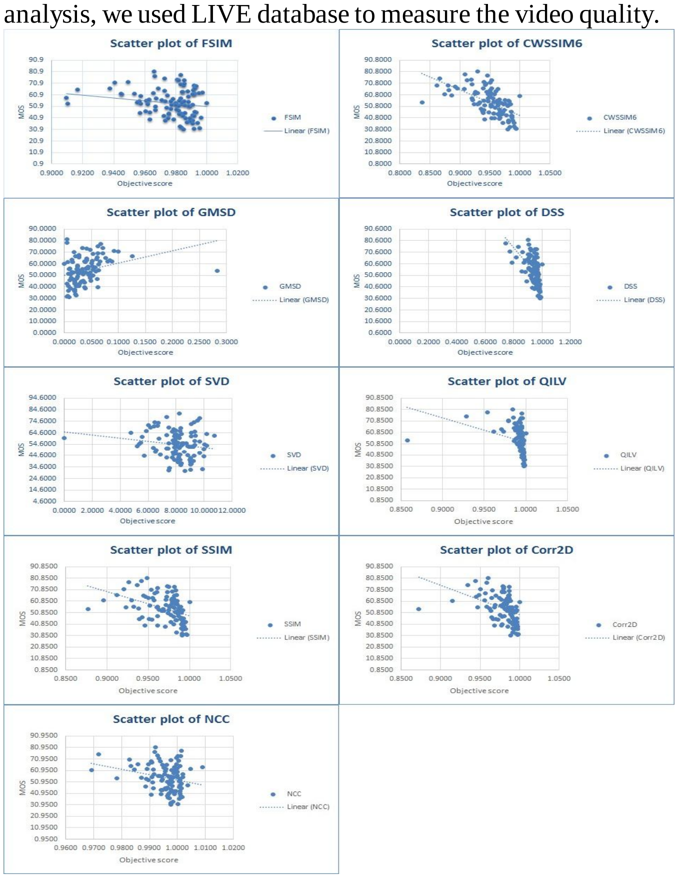

Figure 2: Scatter plots of different individual video quality metrics - FSIM, CWSSIM6, GMSD, DSS, SVD, QILV, SSIM, Corr2D, NCC

Table 1: Quality scores used and their coefficients in regression

	Coefficients
Intercept	-52.19043227
FSIM	-291.0064031
CWSSIM6	-157.8282231
GMSD	51.7846144
DSS	-133.3804311
SVD	-1.58149308
QILV	268.2605103
SSIM	315.4036913
Corr2D	-160.5034452
NCC	262.6636495

We have done analysis the predicted and actual values using Pearson's correlation. Pearson correlation is also called Pearson Linear Correlation Coefficient(PLCC). It measures the statistical relationship, association between variables of interest. PLCC between actual and predicted values of opinion scores is found to be 0.74815. Since the correlation value lies between 0.5 to 1.0, we can infer that selected quality metrics used to build model is having strong correlation.

One of the major things considered in model building is of feature selection. In our case, features means quality metrics used to predict viewers opinion. To minimize the number of variables used to define the model is related to optimization. Here, in this paper, we have used entropy based feature selection. Quality metrics used to build model are based on their entropy values. Entropy is the measure of information. It is measured in bits. More the entropy, more the information contained in the variable i.e. quality metric considered.

Table 2: Quality metrics and entropy values for LIVE database

Quality Metric	**Entropy value for LIVE database**
FSIM	3.6781
CWSSIM6	**4.6525**
GMSD	**4.2564**
DSS	**4.8002**
SVD	5.8539
QILV	2.1651
SSIM	**4.0170**
Corr2D	3.5735
NCC.	2.0807

In our approach, we have calculated the entropy of each of quality metric for LIVE database which has been listed in Table 2. From table, metrics having entropy more than 4 have been selected for training. The quality metrics in Table 2 shown in bold letters – CWSSIM6, GMSD, DSS and SSIM- are used for training. We have used regression based machine learning method and used 10 fold cross validation. Classification accuracy obtained with different classifiers and is presented in Table 3.

We have checked with decision tree, Kernel Naïve Bayes, Quadratic SVM, Cubic KNN, Subspace discriminant ensemble classifiers for the experimentation. Hence, entropy as a feature selection measure to create a hybrid quality metric is feasible. So, we can select features based on entropy to optimize the video quality prediction system.

Table 3: Classification accuracy of different classifiers for entropy based feature selection process.

Classifier	Training with 9 features	
	Classification Accuracy %	Classification Accuracy (with PCA) %
Coarse decision tree	31.9	19.0
Kernel Naïve Bayes	25.9	20.0
Quadratic SVM	33.8	14.3
Medium KNN	32.9	17.1
Ensemble – Subspace discriminant	28.7	28.6

Information gain is the measure of reduction in entropy from transforming a dataset. In this approach shown in subsequent section, Information gain used for feature selection. Information gain is evaluated by comparing the entropy of the dataset before and after a transformation. Information gain may be referred as mutual information and calculate the statistical dependence between variables.

In our approach, we have calculated the information gain of each of quality metric for LIVE database which has been listed in Table 3. From table, metrics having information gain more than 0 have been selected for training. The quality metrics in Table 3 shown in bold letters – FSIM, CWSSIM6, DSS, QILV, SSIM and Corr2D are used for model building. We have used regression based machine learning method and used 10 fold cross validation. Classification accuracy obtained with different classifiers and is presented in Table 4. From Table 5, classification accuracy improved with the feature selection based on information gain.

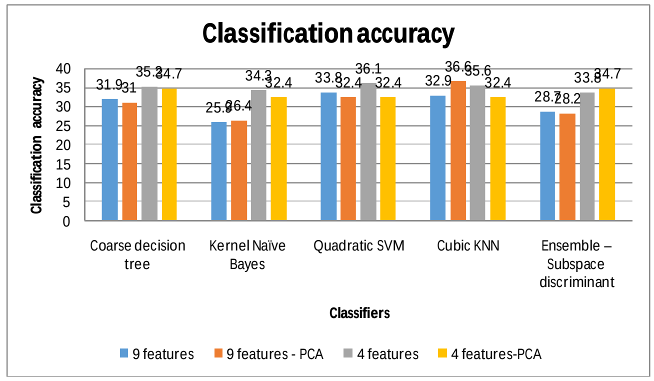

Figure 3: Bar chart of classification accuracy for different classifiers for entropy based feature selection

Classifier	Training with 4 features	
	Classification Accuracy %	Classification Accuracy (with PCA) %
Coarse decision tree	33.3	21.9
Kernel Naïve Bayes	36.2	18.1
Quadratic SVM	39.0	18.1
Medium KNN	43.8	21.0
Ensemble – Subspace discriminant	39.0	28.6

Table 4: Quality metrics and information gain values for LIVE database

Quality Metric	Information gain LIVE database
FSIM	**0.2086579**
CWSSIM6	**0.3395563**
GMSD	0.0000000
DSS	**0.2105293**
SVD	0.0000000
QILV	**0.2877562**
SSIM	**0.1975380**
Corr2D	**0.2392221**
NCC	0.0000000

5. CONCLUSION

In this paper, objective video quality assessment is performed using entropy based feature selection and information gain based feature selection. The proposed method demonstrates higher classification rate (accuracy) as compared to without feature selection technique. For entropy based feature selection, when we train with 4 selected features, Quadratic SVM shows higher classification accuracy as compared to other techniques. Feature selection by using information gain also shows improved performance as compared without feature selection

Table 5: Classification Accuracy of different classifiers for Information gain based feature selection process.

Classifier	Training with 9 features		Training with 4 features	
	Classification Accuracy %	Classification Accuracy (with PCA) %	Classification Accuracy %	Classification Accuracy (with PCA) %
Coarse decision tree	31.9	19.0	33.3	21.9
Kernel Naïve Bayes	25.9	20.0	36.2	18.1
Quadratic SVM	33.8	14.3	39.0	18.1
Medium KNN	32.9	17.1	43.8	21.0
Ensemble – Subspace discriminant	28.7	28.6	39.0	28.6

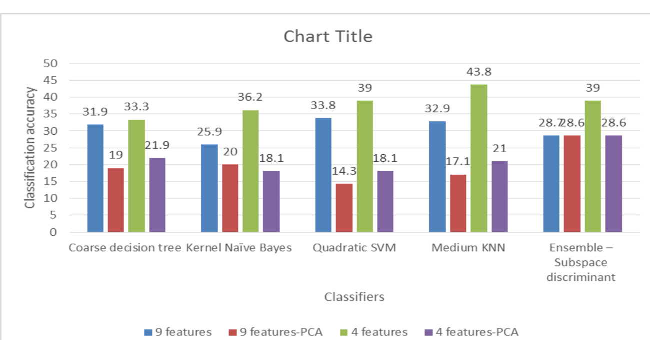

Figure 4: Bar chart of classification accuracy for different classifiers for information gain based feature selection

ACKNOWLEDGMENT

This work is supported by Research Center, Department of Electronics and Communication Engineering, VVCE, Mysuru

REFERENCES

Aja-Fernández, S., Estépar, R. S. J., Alberola-López, C., & Westin, C. F. (2006). Image quality assessment based on local variance. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. https://doi.org/10.1109/IEMBS.2006.259516

Balanov, A., Schwartz, A., Moshe, Y., & Peleg, N. (2015). Image quality assessment based on DCT subband similarity. *Proceedings - International Conference on Image Processing, ICIP*. https://doi.org/10.1109/ICIP.2015.7351172

Blaber, J., Adair, B., & Antoniou, A. (2015). Ncorr: Open-Source 2D Digital Image Correlation Matlab Software. *Experimental Mechanics*. https://doi.org/10.1007/s11340-015-0009-1

Ismael, O. F., & Hussain, Z. M. (2019). SVD-structural similarity in the wavelet-gabor domain: Improved confidence for face recognition under noise, blur and haze. *Journal of Computer Science*. https://doi.org/10.3844/jcssp.2019.1209.1224

Sampat, M. P., Wang, Z., Gupta, S., Bovik, A. C., & Markey, M. K. (2009). Complex wavelet structural similarity: A new image similarity index. *IEEE Transactions on Image Processing*. https://doi.org/10.1109/TIP.2009.2025923

Seshadrinathan, K., Soundararajan, R., Bovik, A. C., & Cormack, L. K. (2010). A subjective study to evaluate video quality assessment algorithms. *Human Vision and Electronic Imaging XV*. https://doi.org/10.1117/12.845382

Seshadrinathan, K., Soundararajan, R., Member, S., Bovik, A. C., & Cormack, L. K. (2010). *Study of Subjective and Objective Quality Assessment of Video*. *19*(6), 1427–1441.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Xue, W., Zhang, L., Mou, X., & Bovik, A. C. (2014). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, *23*(2), 668–695. https://doi.org/10.1109/TIP.2013.2293423

Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*. https://doi.org/10.1109/TIP.2011.2109730