International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 8, October - 2012

Video Quality Assessment Using Motion Features

A.Sphoorthyi[1], Ch.Madhuri Devi[2]

[1]M.tech.Student, Sri Indu College of Engineering & Technology, Hyderabad

[2]Associate Professor, Sri Indu College of Engineering & Technology, Hyderabad

Abstract

The quality of videos as estimated by human observers is of interest for a number of applications. The video quality depends on the video codec, bit-rates required and the content of video material. In this paper, we propose a new scheme for quality assessment of coded video streams has been proposed. The proposed method proposes features describing the intensity of salient motion in the frames, as well as the intensity of coding artifacts in the salient motion regions. In this paper, feature selection is used to selecting the features most correlated to video quality. The experimental results show that the intensity of the blurring and blocking effects in the salient regions has most bearing on the perceived video quality.

***Index Terms*—M5 , motion, no-reference, perceptual quality, regression trees, saliency, video quality assessment.**

I.INTRODUCTION

One of the key technologies required for efficient access and management of video library is video summarization, that is, to effectively extract important information from video data while removing redundant data. A large number of published papers exist, proposing different measures of prominent artifacts appearing in coded images and video sequences . The goal of each no-reference approach is to create an estimator based on the proposed features that would predict the Mean Opinion Score (MOS) of human observers, without using the original (not-degraded) image or sequence data. In the past, Background object extraction usually contains nonliving objects that remain passive in the scene. The background objects can be stationary objects, such as walls, doors and room furniture, or non-stationary objects such as wavering bushes or moving escalators .

The traditional video quality metrics1, such as signal-to-noise ratio (SNR), peak-signal-to-noise ratio (PSNR), and mean squared error (MSE), though computationally simple, are known to disregard the viewing conditions and the characteristics of human visual perception . The additional measures were introduced to account for the temporal dynamics of the sequence. Two motion intensity measures were used: (i) global motion intensity, calculated from the global motion field, and (ii) object motion intensity, calculated by subtracting the global motion from the MPEG motion vectors . Subjective video quality assessment methods are able to reliably measure the video quality and are crucial for evaluating the performance of objective visual quality assessment metrics. The subjective video quality methods are based on groups of trained/untrained users viewing the video content, and providing ratings for quality .

The paper is organized as follows. Feature selection based on correlation is presented in Section II. Section III describes the videoquality measurements. Results are shown in section IV. Section V describes the conclusion.

II. FEATURE SELECTION BASED ON CORRELATION

In this paper, 35 feature values have been calculated for sequences the Video Quality Experts Group (VQEG) provided as a benchmark for codec evaluation. The feature selection based on correlation is used to train an M5 decision tree, as an estimator for the MOS of new sequences. The flow chart of salient motion detection follows as shown in fig.1.

ISSN: 2278-0181
Vol. 1 Issue 8, October - 2012

1) Salient motion detection

The algorithm employs a multi-scale model of the background in the form of frames which form a Gaussian pyramid, akin to the model employed in the attention model. It produces better segmentation of dynamic objects at a small number of scales like 3-5. Moreover, it is able to do so consistently over a wide range of the amount of coding artifacts present. The background frames at each level are obtained by infinite impulse response (running average) filtering commonly used in background subtraction. This allows the approach to take into account temporal consistency in the frames. Finally, outlier detection is used to detect salient changes in the frame. The assumption is that the salient changes are those that differ significantly from the changes undergone by most of the pixels in the frame.

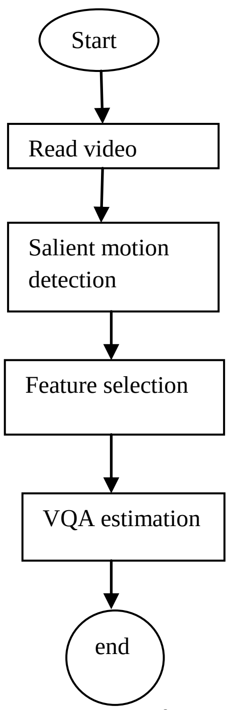

Fig.1. Detection of salient motion algorithm

The procedure for salient motion detection as shown below

a) Each frame from video is passed to a Gaussian filter and then obtains a pyramid of frames.

b) Updated the two background frames

$$b(i)=(1-\alpha)b(i)+\alpha p(i) \qquad (1)$$

where α is the learning rate used to filter the ith background frame, p(i) is the value of pixel at location in the current frame i,b(i) is the value of pixel at location in the ith background frame

c) Calculate temporal filter by inserting the current frame between the two background frames

$$f(x) = -\frac{2}{\sqrt{3}}\pi^{-\frac{1}{4}} \cdot (1-x^2) \cdot \exp\frac{-x^2}{2} \qquad (2)$$

Where x represents the Euclidean distance of the point from the center of the filter.

d) To calculate mean obsulate distance (MAD) to detect the outliers from frame

$$MAD = \frac{\sum_{i=1}^{N} |fp_i - \mu|}{N} \qquad (3)$$

e) Calculate z-score value

$$Z_i^{score} = \frac{|fp_i - \mu|}{MAD} \qquad (4)$$

z-score value is better the moving objects of interest in the scene when compared those obtained by the static saliency model.

2) Feature selection

In this paper, two features are selected. They are zero crossing rates and z-score. The feature selection was based on the results of prediction of an M5 algorithm trained using a specific feature subset. Again, a genetic algorithm was used to search solution space remain in the peripheral vision, but they have a definite effect on our perception of the video quality. They are in fact, for the lack of a better expression, very annoying and observable.

2.1.SSIM Feature

The **structural similarity** (SSIM) index is a method for measuring the similarity between two images. The SSIM index is a full reference metric, in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proved to be inconsistent with human eye perception

The SSIM metric is calculated on various windows of an image. The measure between two windows x and y of common size $N \times N$ is:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

3) VQA estimation of salient motion

The video quality assessment measure based on the selected features for half of the frames of the sequence, uniformly distributed (i.e., the frame rate was halved to make the approach more efficient). The features obtained for each evaluated frame were fed into the estimator and the measure of the quality for that frame obtained. Since the standard deviation of the estimator's prediction error over the frames of a single sequence is relatively high, robust statistics should be used to arrive at the final single measure of sequence video quality.

III. VIDEO QUALITY MEASUREMENTS

We used two measures Root Mean Square Error (RMSE) and mean absolute error (MAE).

The RMSE of original frame R and recovery frame F is given by

$$RMSE = \frac{1}{M*N}\sum_{i=1}^{m}\ \sum_{j=1}^{n}[R(i,j) - F(i,j)]^2 \qquad (5)$$

The MAE measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables. The equation is given in the library references. Expressed in words, the MAE is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. The MAE is a linear score which means that all the individual differences are weighted equally in the average.

$$\mathbf{MAE=1/n \sum |F\text{-}Y|} \qquad (6)$$

Where f is the original frame and Y is the recovery frame

IV. EXPERIMENTAL RESULTS

In table I shows the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Root Relative mean Squared Error (RRMSE) from 99th frame from train video in proposed method and Multi Layer Perceptron (MLP).

method	MAE	RMSE	RRMSE
proposed	0.25	0.41	0.78
MLP	0.37	0.49	0.87

Table.1. show the measurements values in fig.1.

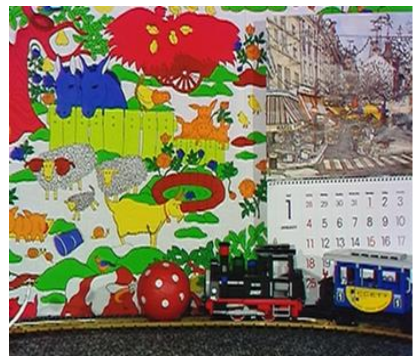

Fig.2. 99th frame from train video at 4MB/s

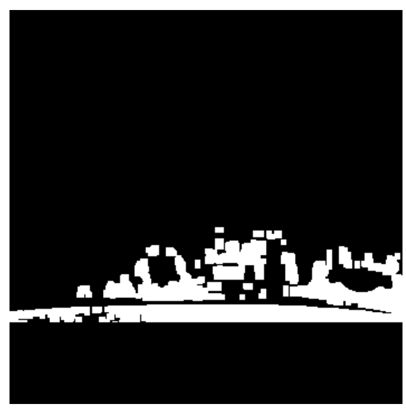

Fig.3. Salient motion detected at 4MB/s

V.CONCLUSION

In this paper, the proposed method can be used to enhance the performance of video quality assessment approach. The improvement can be achieved even when computationally inexpensive approaches, such as that proposed, are used to determine salient regions in the frame. The proposed methods, the frame intensity of the blurring and blocking effects in the salient regions have most bearing on the perceived video quality.

V1.REFERENCES

[1] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," in Proc. IEEE Int.Conf. Image Process., 2002, pp. 477–480.

[2] D. Culibrk, D. Kukolj, P. Vasiljevic,M. Pokric, and V. Zlokolica, "Feature selection for neural-network based no-reference video quality assessment," in Proc. ICANN (2), 2009, pp. 633–642.

[3] Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU-R BT.500, Video Quality Experts Group, 2002.

[4] T. Liu, X. Feng, A. Reibman, and Y. Wang, "Saliency inspired modeling of packet-loss visibility in decoded videos," in Proc. Int. Work- shop VPQM, 2009, pp. 1–4.

[4] D. Gavrila, "The visual analysis of human movement: A survey," Comput. Vis. Image Understanding, vol. 73, no. 1, pp. 82–98, 1999.

[5] L. Li and M. Leung, "Integrating intensity and texture differences for robust change detection," IEEE Trans. Image Processing, vol. 11, pp. 105–112, Feb. 2002.

[6] L. Guo and Y. Meng, —What is wrong and right with MSE?,‖ in Proc.8th Int. Conf. Signal Image Process., 2006, pp. 212–215.

[7] Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU-R Recommendation BT.500-11.

[8] G.Warwick and N. Thong, Signal Processing for Telecommunications and Multimedia,. New York: Springer, 2004, ch. 6.

Miss.**A.SPHOORTHY** graduate from SPHOORTHY Engg College in **Electronics& Communications**. Now pursuing Masters in **Digital Electronics and Communication Systems** (DECS) from Sri Indu College of Engineering & Technology.

I express my gratitude to Mrs.**CH.MADHURI DEVI** Associate Professor Department of (ECE) and for her constant co-operation, support and for providing necessary facilities throughout the M.tech program. She has 6 Years of Experience at B.Tech and 2 years of Experience at M.tech Level and working as a Associate Professor in Sri Indu College of Engg. & Technology.