Enhanced optical flow-based full reference video quality assessment algorithm

Sagar Gujjunoori[1] · Madhu Oruganti[2] · Alwyn Roshan Pais[3]

Abstract

Full reference video quality assessment based on optical flow is emerging. Human Visual System (HVS) based video quality assessment algorithms are playing an important role in effectively assessing the distortions in video sequences. There exist very few video quality assessment algorithms which consider spatio-temporal distortions effectively. To address the above issues, we present an enhanced optical flow based full reference video quality algorithm which considers the orientation feature of the optical flow while computing the temporal distortions as opposed to the use of feature, minimum eigenvalue as in the state of the art. Further, it presents an interquartile range based comparative weighted closeness (INT-CWC) measure which aimed to measure the comparative dispersion of video quality scores of any two video quality assessment algorithms with DMOS scores. Here INT-CWC measure is a novel attempt. The performance of proposed scheme is evaluated using the LIVE dataset and scheme is shown to be competitive with, and even out-perform, existing video quality assessment algorithms.

Keywords Video Qualti assessment · DMOS · Optcial flow · Human visual system (HVS)

1 Introduction and literature

1. Nowadays, video technology and its applications has huge demand in communication systems. The importance for reliable video quality assessment is also increasing. In recent past many video quality assessment methods and metrics are came to light with

✉ Sagar Gujjunoori
  sagar4u.nitk@gmail.com

  Madhu Oruganti
  oruganti.madhu@gmail.com

[1] Department of Information Technology, Vardhaman College of Engineering, Shamshabad, 501218, Telangana, India

[2] Junior Research Fellow (JRF), Department of Information Technology, Vardhaman College of Engineering, Shamshabad, 501218, Telangana, India

[3] Department of Computer Science and Engineering, NITK-Surathkal, Mangalore, India

variable computational complexity and accuracy. As per recent survey, e.g., [20], 66% of data traffic is due to video transmission from mobile devices.

2. The traditional video quality metrics, 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 [11]. Although subjective video quality assessment methods may serve the purpose, based on the groups of trained/untrained human evaluators. Further, to meet the standards ITU-T, these evaluation methods must follow straightened conditions such of viewing distance, test duration, room illumination, etc [16, 22] and these methods are time consuming expensive and laborious.

3. The validation tests, for objective video quality metrics, developed by VQEG is remarkable in this direction [23].

4. More recently, a full reference video quality metric called MOtion-based Video Integrity Evaluation (MOVIE) index was proposed by Seshadrinathan and Bovik [8], FLOSIM [13] and Ortiz-Jaramillo et al. [1]. The MOVIE model, strives to capture the characteristics of the middle temporal (MT) visual area of the visual cortex in the human brain for video quality analysis. Neuroscience studies indicate that the visual area MT is critical for the perception of video quality [2].

5. The response characteristics of the visual area MT are modelled using separable Gabor filter banks. The model described two indices, namely a Spatial MOVIE index that primarily captures spatial distortions and a Temporal MOVIE index that captures temporal distortions. However, the MOVIE index is not suitable which is considering the optical flow baged visual qualities assessment that emphasize on the HVS characteristics. In [4], proposed a method by considering direction of optical flow in video is used for video quality assessment scores. In [18] proposed a method for computing video quality assessment score by considering HVS and optical flow concepts.

Motion detection plays an important role in video analysis. Optical flow based motion estimation algorithms became more popular in estimation of motion trajectories [5, 6, 12]. There are significant contributions in the state of the art to assess the distorted video sequences [15, 19, 24, 25]. However, understanding the Human visual system yet an open area of research. Recently optical flow based video quality assessment algorithms are emerging which attempt to understand the human visual system characteristics while measuring the distortions in video sequences [13, 20]. Manasa et al. algorithm (FLOSIM) [13] have proposed an optical flow based full referenced video quality assessment algorithm by considering the statistical features of optical flow. Manasa et al. have presented the Full Reference (FR) technique to measure the perceptual annoyance that results from temporal distortions. The statistical features: mean, standard deviation of the flow magnitudes and minimum eigenvalue of a flow are considered to quantify the temporal distortions.

In the paper we claim that the consideration of the statistical feature, minimum eigenvalue, for measuring the randomness in the optical flow be improved. We can observe for certain videos, as in Fig. 1, that the video quality scores are far away from the DMOS scores. In Fig. 1, the x- axis represents various test video sequences and y- axis represents video quality scores in terms of DMOS and FLOSIM values. This motivated us to propose another model to capture the temporal distortions that results from random flow. We proposed to use the orientation feature of the optical flow which effectively quantifies the randomness in flow, as opposed to the use of minimum eigenvalue as in [13].

From the Fig. 2, we can observe that though the video quality scores we obtained using our proposed algorithm are more closer to the DMOS score when compared to the
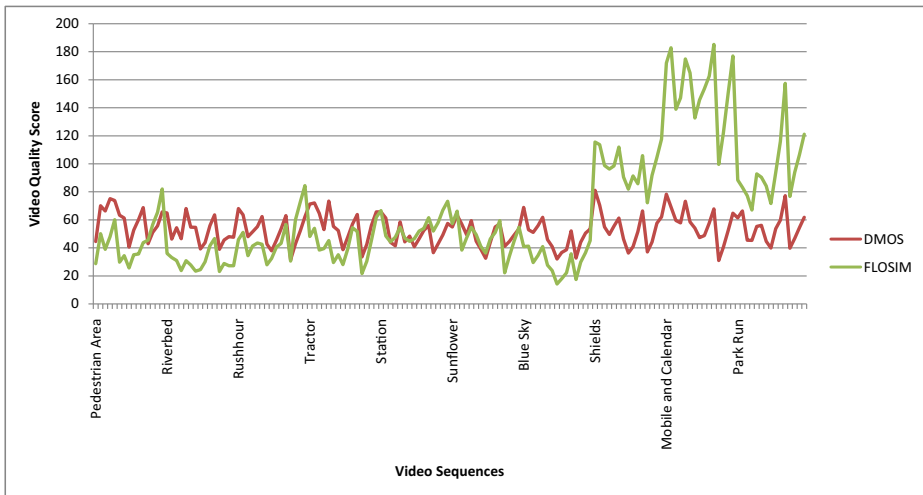
Fig. 1 Comparison of FLOSIM video quality scores with DMOS scores

FLOSIM, the correlation coefficient with DMOS are lower. This motivated us to propose yet another measure, referred as INT-CWC measure, aimed to provide a comparative closeness score with DMOS for any two video quality assessment algorithms. We hypothesize that if majority of the video quality scores of an algorithm are closer to the profile of DMOS scores, its performance should be appreciated even few extreme deviations from DMOS exist.

The summary of our contributions are as follows. The article aims at addressing the issue of video quality assessment, emphasizing on human visual system. It first propose a full reference optical flow based video quality assessment algorithm by considering orientation features and then propose interquartile based comparative weighted closeness measure(INT-CWC). Our experimental results demonstrate that the first video quality
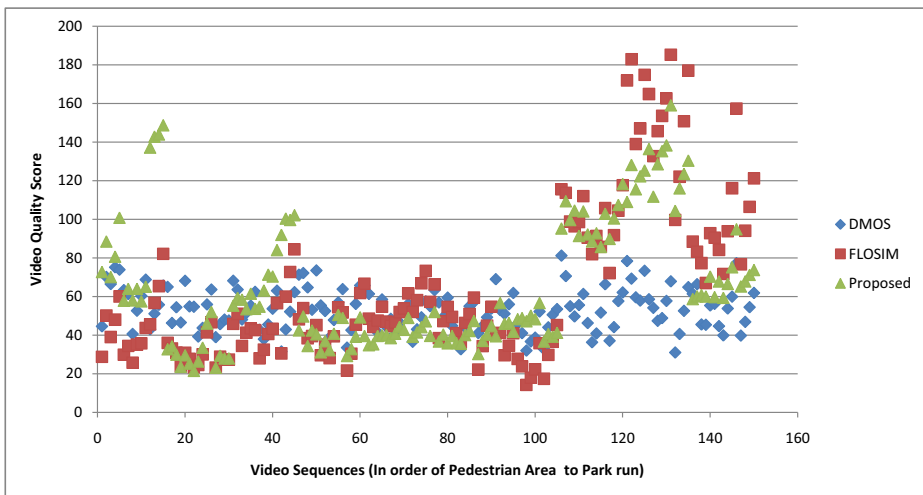

Fig. 2 Scatter plot for all the video sequences

assessment algorithm achieves much closer video video quality scores to DMOS scores and the INT-CWC measure is highly correlated with the video quality score profiles.

The rest of the article is organized as follows. In Section 2, we present our two proposed contributions. Section 3 presents the experimental results. In Section 4, we conclude the article.

2 Proposed work and methodology

The proposed work is presented in the following two sections. In Section 2.1, we present the enhanced optical flow based full reference video quality assessment algorithm. In Section 2.2, we present the INT-CWC: Interquartile based comparative weighted closeness measure.

2.1 Enhanced video quality assessment algorithm

We propose an HVS based full reference video quality assessment model which aimed to enhance that of Manasa et al. algorithm (FLOSIM) [13]. Manasa et al. algorithm at first computes temporal distortions and spatial distortions separately and then computes the combined final video quality score by their proposed pooling strategy. We follow the same approach, however the proposed model differs in computation of temporal features. Please refer the figure captioned *overview of the proposed approach* in Manasa et al. algorithm [13] and its overview can be seen in Fig. 3.

To obtain the spatial distortion features the MS-SSIM index [26] is used. To extract the temporal features any optical flow estimation algorithm can be used. The idea behind temporal distortion features computation is that there exist a deviation in local statistical properties of optical flow due to distortions when compared to the undistorted optical flow statistics [13]. Further, Manasa et al. [13] algorithm claim that local mean $\mu_{|.|}$ and local standard deviations $\sigma_{|.|}$ are well capable of capturing the local flow inconsistencies and randomness of optical flow can be represented by the minimum eigenvalue $\lambda_{|.|}$. We claim that the use of feature minimum eigenvalue, to represent the randomness is not well capable. Due to the use of this feature for some video sequences such as Mobile and Calendar and Park run [21, 22] as in Fig. 1, the distortion scores of FLOSIM are higher than the DMOS values.

2.1.1 Notations and models

Let $\mathbf{V}_x^i$ and $\mathbf{V}_y^i$, be the x and y components of ith velocity matrices of size $M \times N$ respectively, can be computed using any optical flow estimation algorithm, [5, 6, 12], between two consecutive video frames $\mathbf{F}^j$ and $\mathbf{F}^{j+1}$, each of size $M \times N$, respectively. Let $\mathbf{V}^i = (\mathbf{V}_x^i, \mathbf{V}_y^i)$. In the proposed model, we used Gunnar Farneback optical flow algorithm [5], and its the implementation is available in Matlab R2018a. From the velocity matrices obtained using Gunnar Farneback optical flow estimation algorithm, the magnitudes and scaled phase angles (in radians) can be defined as in (1) and (2) respectively.

$$m_{xy}^i = \sqrt{v_x^i + v_y^i} \tag{1}$$

where v_x^i and v_y^i are the elements of $\mathbf{V}_x^i$ and $\mathbf{V}_y^i$ respectively. Let $\mathbf{M}^i$ be the ith magnitude matrix of size $M \times N$, which contain all the magnitude values, m_{xy}^i as in (1), computed
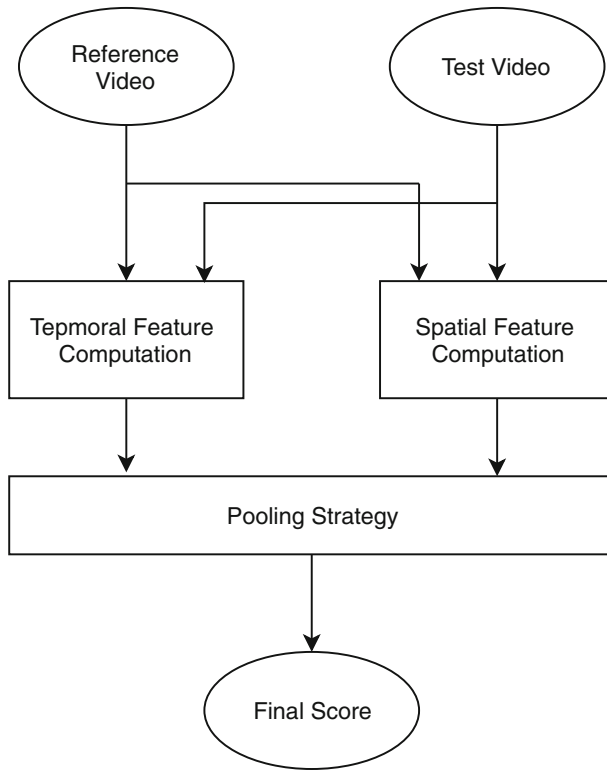
Fig. 3 Overview of Manasa et al. Algorithm

between the $\mathbf{V}_x^i$ and $\mathbf{V}_y^i$. Similarly the phase angle (orientation) between v_x^i and v_y^i is computed as in (2)

$$\theta_{yx}^i = atan2(v_y^i, v_x^i) \times (\frac{360}{\pi}) \tag{2}$$

Let $\mathbf{A}^i$ be the ith phase angle matrix of size $M \times N$, which contain the elements, θ_{yx}^i, computed between the matrices $\mathbf{V}_x^i$ and $\mathbf{V}_y^i$. And $\mathbf{A}^i(x, y) = \theta_{yx}^i$ denotes an element at coordinate (x, y).

2.1.2 Description

In Manasa et al. [13] algorithm, the temporal local statistical features are computed using the optical flow components. These local statistical features are computed frame-by-frame basis and each feature is computed on 7×7 optical flow patches. The three features computed on optical flow components are given in (3), (4), (5).

$$\mathbf{f}_1^i = [\mu_1^i, \mu_2^i, \mu_3^i, \ldots, \mu_{K \times L}^i]^T, \tag{3}$$

$$\mathbf{f}_2^i = [\sigma_1^i, \sigma_2^i, \sigma_3^i, \ldots, \sigma_{K \times L}^i]^T, \tag{4}$$

$$\mathbf{f}_3^i = [\lambda_1^i, \lambda_2^i, \lambda_3^i, \ldots, \lambda_{K \times L}^i]^T, \tag{5}$$

Where $\mathbf{f}_1^i$ denotes the per frame feature vector of $\mathbf{V}^i$, and $\mu_{|.|}$ denotes the mean of the flow magnitudes, m_{xy}^i as in (1), that are present in a local patch of size 7×7. Similarly, $\mathbf{f}_2^i$ is the

second per frame feature vector of $\mathbf{V}^i$, and $\sigma_{|.|}$ is the standard deviation of flow magnitudes, m_{xy}^i as in (1), that are present in the local patch of size 7×7. Whereas $\mathbf{f}_3^i$ is the third per frame feature vector of $\mathbf{V}^i$, and $\lambda_{|.|}$ denotes that minimum eigenvalue of flow patches of covariance matrix. There are $K \times L$ number of non-overlapping patches in a frame.

In the proposed model, the computation of first feature $\mathbf{f}_1^i$ vector and second feature vector $\mathbf{f}_2^i$ is same as Manasa et al. algorithm [13], but the computation of third feature $\mathbf{f}_3^i$ differs. In Manasa et al. algorithm, it is stated that the perceivable distortion can be effectively characterized by the measure of dispersion of the features. The measure of dispersion of data is given in (5), where $CV(.)$ denotes the coefficient of variation, $\mathbf{z}$ denotes the data vector, $\mu_{\mathbf{z}}$ is the mean of $\mathbf{z}$, $\sigma_{\mathbf{z}}$ denotes the standard deviation of $\mathbf{z}$.

$$CV(\mathbf{z}) = \frac{\sigma_{\mathbf{z}}}{\mu_{\mathbf{z}}} \tag{6}$$

Once the $CV(\mathbf{f}_1^i)$ and $CV(\mathbf{f}_2^i)$ are computed and pooled, the difference in dispersion is computed as follows in (7).

$$D(\mathbf{z}_r, \mathbf{z}_t) = CV(\mathbf{z}_r) - CV(\mathbf{z}_t) \tag{7}$$

Where $\mathbf{z}_r$ is the reference data vector and $\mathbf{z}_t$ is the test data vector. The definition in (7) is applied to $\mathbf{f}_1^i$, $\mathbf{f}_2^i$, i.e, we can be obtain $D(\mathbf{f}_{1r}^i, \mathbf{f}_{1t}^i)$ and $D(\mathbf{f}_{2r}^i, \mathbf{f}_{2t}^i)$ from (7).

We propose a different way of computing the third feature $\mathbf{f}_3^i$, which effectively computes the randomness of the optical flow that effectively quantifies the temporal distortions. The proposed method of computing the $\mathbf{f}_3^i$, is presented from (8)-(14).

$$\mathbf{A}_1^i(x, y) = \begin{cases} -1, & \text{if } \theta_{yx}^i < 0 \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

$$\mathbf{A}_2^i(x, y) = |\theta_{yx}^i| \tag{9}$$

$$\mathbf{a}_1^i = [\sigma_1^{\mathbf{A}_1^i}, \sigma_2^{\mathbf{A}_1^i}, \sigma_3^{\mathbf{A}_1^i}, \dots, \sigma_{K \times L}^{\mathbf{A}_1^i}]^T, \tag{10}$$

$$\mathbf{a}_2^i = [\sigma_1^{\mathbf{A}_2^i}, \sigma_2^{\mathbf{A}_2^i}, \sigma_3^{\mathbf{A}_2^i}, \dots, \sigma_{K \times L}^{\mathbf{A}_2^i}]^T, \tag{11}$$

$$\eta_l^i = \begin{cases} \sigma_l^{\mathbf{A}_1^i} \times \sigma_l^{\mathbf{A}_2^i}, & \text{if } \sigma_l^{\mathbf{A}_1^i} \neq 0 \wedge \sigma_l^{\mathbf{A}_2^i} \neq 0 \\ \psi_l^i, & \text{otherwise} \end{cases} \tag{12}$$

where,

$$\psi_l^i = \begin{cases} \sigma_l^{\mathbf{A}_1^i}, & \text{if } \sigma_l^{\mathbf{A}_2^i} = 0 \\ \sigma_l^{\mathbf{A}_2^i}, & \text{otherwise} \end{cases} \tag{13}$$

for $l = 1, 2, 3, \dots, K \times L$.

$$\mathbf{f}_3^i = [\eta_1^i, \eta_2^i, \eta_3^i, \dots, \eta_{K \times L}^i]^T, \tag{14}$$

The deviation in overall values of $\mathbf{f}_3^i$ in (14) of the distorted patch from the reference patch is quantified as below in (15).

$$C(\mathbf{f}_{3r}^i, \mathbf{f}_{3t}^i) = 1 - \text{corr}(\mathbf{f}_{3r}^i, \mathbf{f}_{3t}^i) \tag{15}$$

where the subscript r denotes reference and t denotes test sets, and $\text{corr}(\mathbf{x}, \mathbf{y})$ is the correlation coefficient computed between two data vectors $\mathbf{x}$ and $\mathbf{y}$. The pooling strategy and final video score computation is same as that of Manasa et al. algorithm [13]. Readers have been directed to refer [13] for more details on pooling strategy and final video score computation.

2.2 INT-CWC: Interquartile based comparative weighted closeness measure

We propose INT-CWC measure which aimed to measure the quantitative closeness of any two video quality assessment algorithms to be compared with DMOS. When majority of the video quality scores in the first algorithm are closer to DMOS when compare to second algorithm, the first algorithm performance is superior and similarly when majority of video quality scores in the second algorithm are closed to DMOS when compared to first algorithm, the second algorithm performance is superior.

$$\text{Let, } \mathbf{s^d} = [s_1^d, s_2^d, s_3^d, \dots, s_N^d]^T,$$
$$\mathbf{s^{m_1}} = [s_1^{m_1}, s_2^{m_1}, s_3^{m_1}, \dots, s_N^{m_1}]^T \text{ and}$$
$$\mathbf{s^{m_2}} = [s_1^{m_2}, s_2^{m_2}, s_3^{m_2}, \dots, s_N^{m_2}]^T$$

denotes the DMOS score, video quality scores of first algorithm and second algorithms respectively and N is the total number of video scores. The absolute difference between the DMOS scores and video quality scores of first algorithm are given in (16).

$$\Delta^{\mathbf{m_1}} = [\Delta_1^{m_1}, \Delta_2^{m_1}, \Delta_3^{m_1}, \dots, \Delta_N^{m_1}]^T \tag{16}$$

where $\Delta_i^{m_1} = |s_i^d - s_i^{m_1}|$ and for $i = 1, 2, 3, \dots, N$. Similarly, the absolute difference between the DMOS scores and video quality scores of second algorithm are given in (17).

$$\Delta^{\mathbf{m_2}} = [\Delta_1^{m_2}, \Delta_2^{m_2}, \Delta_3^{m_2}, \dots, \Delta_N^{m_2}]^T \tag{17}$$

where $\Delta_i^{m_2} = |s_i^d - s_i^{m_2}|$ and for $i = 1, 2, 3, \dots, N$. We propose to use upper whisker of interquartile range (IQR) statistical dispersion measure [27] in designing the INT-CWC measure. To compute the upper whisker, Q_1 is computed as follows in (18).

$$Q_1 = \text{median}(\Delta^{\mathbf{z}}) \tag{18}$$

Let $\Delta^{\hat{\mathbf{z}}} = [\Delta_1^{\hat{z}}, \Delta_2^{\hat{z}}, \Delta_3^{\hat{z}}, \dots, \Delta_n^{\hat{z}}]^T$ denotes all the remaining terms in $\Delta^{\mathbf{z}}$ which are greater than Q_1. Then compute Q_3 as in (19).

$$Q_3 = \text{median}(\Delta^{\hat{\mathbf{z}}}) \tag{19}$$

The upper whisker is given below in (20).

$$I = Q_3 + 1.5 \times IQR \tag{20}$$

where $IQR = Q_3 - Q_1$.

Let $I_{\Delta^{\mathbf{m_1}}}$ be the upper whisker threshold for $\Delta^{\mathbf{m_1}}$ which can be computed using (20) and $I_{\Delta^{\mathbf{m_2}}}$ be the upper whisker threshold for $\Delta^{\mathbf{m_2}}$ which can be computed using (20). We compute the updated $\Delta^{\mathbf{z}}$ into $\Delta^{\mathbf{z_1}}$ and compute $N_{\mathbf{z}}$ as below in (21), where $N_{\mathbf{z}}$ denotes the number of times $\mathbf{z}$ is updated and initially $N_{\mathbf{z}} = 0$.

$$(\Delta^{\mathbf{z_1}}(i), N_{\mathbf{z}}) = \begin{cases} (I, N_{\mathbf{z}} = N_{\mathbf{z}} + 1), & \text{if } \Delta^{\mathbf{z}}(i) > I, \\ & \quad \text{for } i = 1, 2, 3, \dots, N \\ (\Delta^{\mathbf{z}}(i), N_{\mathbf{z}}), & \text{otherwise} \end{cases} \tag{21}$$

Let $N_{\Delta^{\mathbf{m_1}}}$ denotes the number of times $\Delta^{\mathbf{m_1}}$ is updated using (21) and $T_{\Delta^{\mathbf{m_1}}}$ be the summation of updated values of $\Delta^{\mathbf{m_1}}$ computed using (21) and (22). Similarly, $N_{\Delta^{\mathbf{m_2}}}$ be the number of times $\Delta^{\mathbf{m_2}}$ is updated using (21) and $T_{\Delta^{\mathbf{m_2}}}$ be the summation of updated values of $\Delta^{\mathbf{m_2}}$ computed using (21) and (22).

$$T_{\mathbf{z_1}} = \sum_{i=1}^{N} \Delta^{\mathbf{z_1}}(i) \tag{22}$$

Table 1 Correlation between FLOSIM and proposed algorithm

Video Sequence	Correlation Coefficient
Riverbed	0.9678
Tractor	0.9430
Sunflower	0.8482
Pedestrian Area	0.8368
Shields	0.7556
Rushhour	0.6917

Later compute the $(n_{\Delta m_1}, n_{\Delta m_2})$ as in (23).

$$(n_{\Delta m_1}, n_{\Delta m_2}) = \begin{cases} n_{\Delta m_1} = n_{\Delta m_1} + 1, & \text{if } \Delta_i^{m_1} < \Delta_i^{m_2}, \\ & \quad \text{for } i = 1, 2, 3, \ldots, N \\ n_{\Delta m_2} = n_{\Delta m_2} + 1 & \text{if } \Delta_i^{m_1} > \Delta_i^{m_2}, \\ & \quad \text{for } i = 1, 2, 3, \ldots, N \\ (n_{\Delta m_1} = n_{\Delta m_1} + 1, \\ n_{\Delta m_2} = n_{\Delta m_2} + 1), & \text{otherwise} \end{cases} \quad (23)$$

where $n_{\Delta m_1}$ and $n_{\Delta m_2}$ are initially assigned with 0.

Using $n_{\Delta m_1}$, $N_{\Delta m_1}$ and $T_{\Delta m_1}$, and $n_{\Delta m_2}$, $N_{\Delta m_2}$ and $T_{\Delta m_2}$, we compute the S_1 and S_2 as follows in (24) and (26) respectively.

$$S_1 = \begin{cases} X, & \text{if } n_{\Delta m_1} \neq 0 \\ T_{\Delta m_1}, & \text{otherwise} \end{cases} \quad (24)$$

$$X = \begin{cases} n_{\Delta m_1} \times T_{\Delta m_1} \times \dfrac{1}{N_{\Delta m_1}}, & \text{if } N_{\Delta m_1} \neq 0 \\ n_{\Delta m_1} \times T_{\Delta m_1}, & \text{otherwise} \end{cases} \quad (25)$$

$$S_2 = \begin{cases} Y, & \text{if } n_{\Delta m_2} \neq 0 \\ T_{\Delta m_2}, & \text{otherwise} \end{cases} \quad (26)$$

$$Y = \begin{cases} n_{\Delta m_2} \times T_{\Delta m_2} \times \dfrac{1}{N_{\Delta m_2}}, & \text{if } N_{\Delta m_2} \neq 0 \\ n_{\Delta m_2} \times T_{\Delta m_2}, & \text{otherwise} \end{cases} \quad (27)$$

Using (24) and (26) we compute the final comparative scores as below in (28) and (29) for the first algorithm and second algorithm respectively.
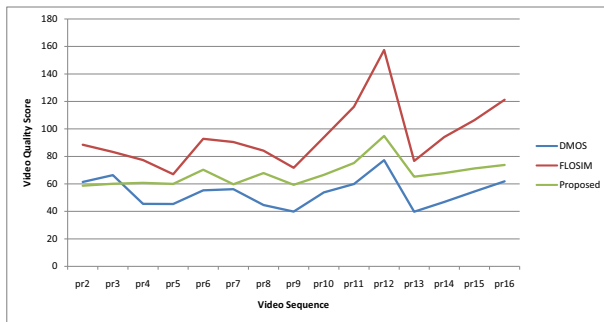
$$S_{m_1} = \frac{S_1}{S_2} \quad (28)$$
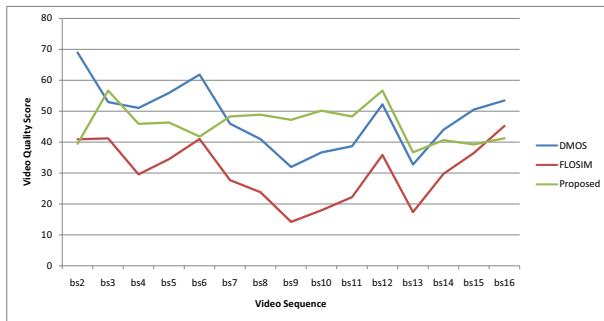
$$S_{m_2} = \frac{S_2}{S_1} \quad (29)$$

We refer S_{m_1} and S_{m_2} as INT-CWC measures for first algorithm and second algorithms to be compared with DMOS.
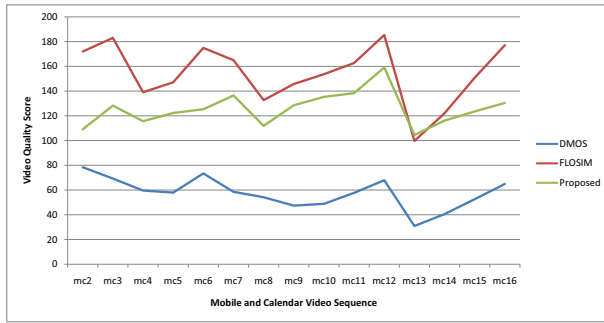
(a) Mobile and Calender

(b) Park Run

(c) Blue Sky

Fig. 4 Performance improvement

3 Experimental results and discussion

In the proposed full reference video quality assessment algorithm, we used the LIVE video dataset [21, 22]. This database consists of ten original and 150 distorted videos. The distorted videos are generated from a set of original videos by using standard distortions like H.264 compression, MPEG-2 compression, transmission of
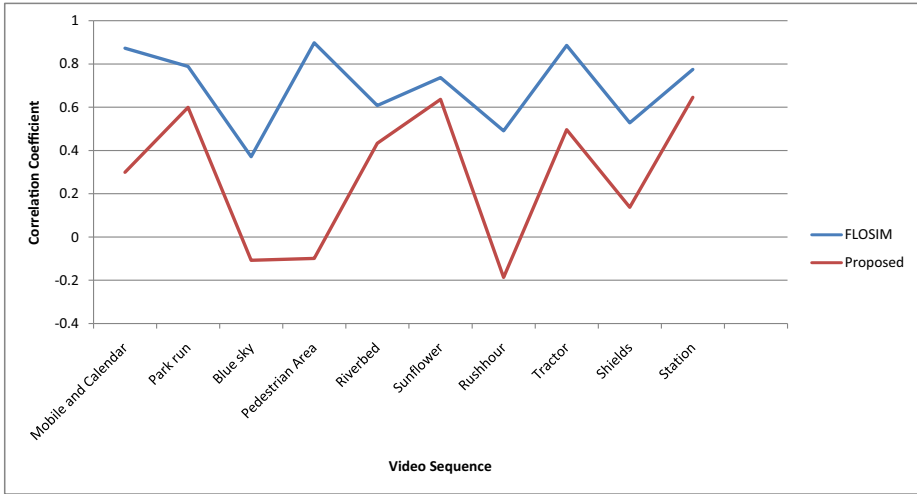
Fig. 5 Comparison of Correlation Coefficient

H.264 compressed bit streams through error-prone IP networks and through error-prone wireless networks. These distortions are applied at different levels such that by using original video, 15 corresponding distorted videos are produced. Experimental results are shown in Table 1. We compare the experimental results with Manasa et al. algorithm FLOSIM [13]. We used FLOSIM code available at [3]. The proposed optical flow based full reference video quality assessment algorithm is implemented using the Gunnar Farneback optical flow estimation algorithm available in Matlab R2018a. The proposed video quality scores for Riverbed (rb), Tractor (tr), Sunflower (sf), Pedestrian Area (pa), Shields (sh), Rushhour (rh) are highly correlated with that of the FLOSIM scores.
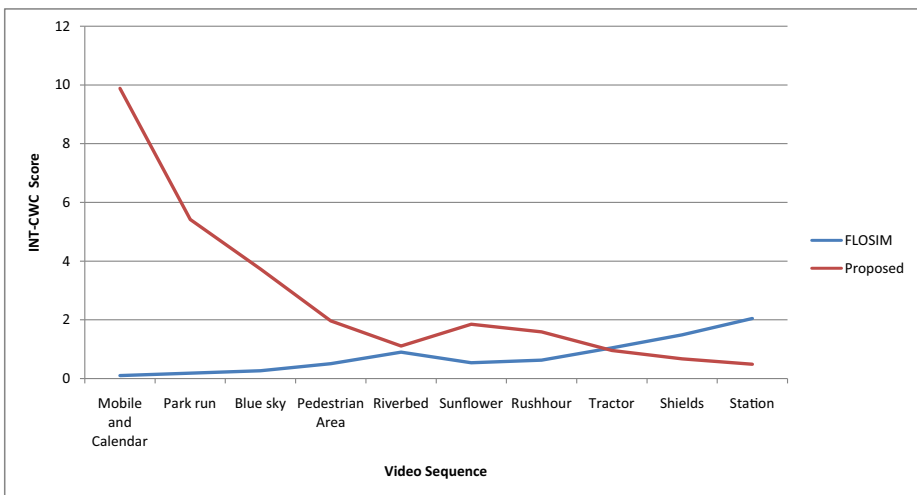
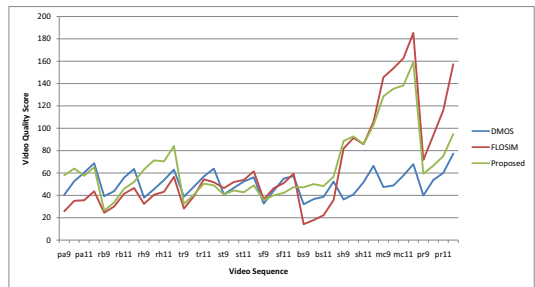**Fig. 6** Comparison of INT-CWC Score

Fig. 7 FLOSIM and Proposed
video quality scores profile
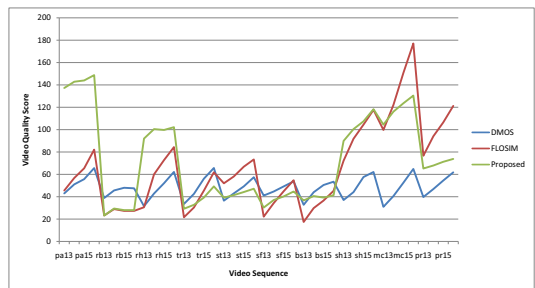compared to DMOS
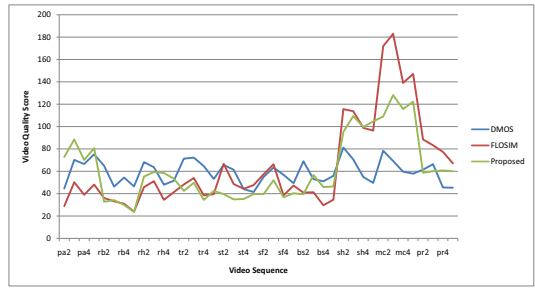
(a) Wireless distortion
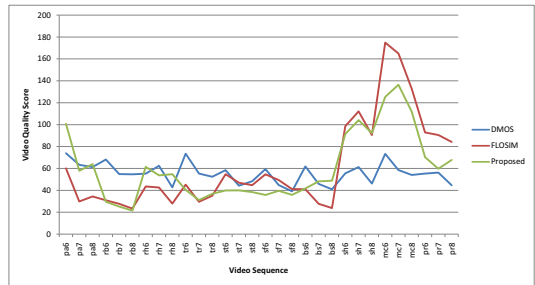
(b) IP distortions


(c) H.264 compression

(d) MPEG-2 compression

Further, from Fig. 4, we can observe that the video quality scores obtained for the video sequences Mobile and Calender(mc), Park run(pr) and Blue sky(bs) using the proposed algorithm are more nearer to the DMOS scores. Note that in Figs. 2, 4, 5, 6, 7, 8 and 9 the x-axis represents the test video sequences and y-axis represents video quality scores in terms of DMOS, FLOSIM, and proposed algorithms. The scatter plot of video quality scores for all the 150 video sequences can be observed in Fig. 2 and We can note that the video quality scores obtained using the proposed algorithm are much nearer to the DMOS
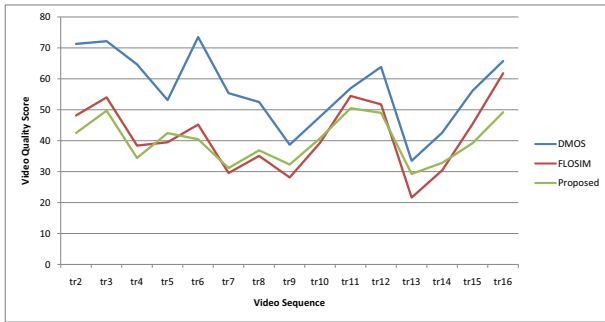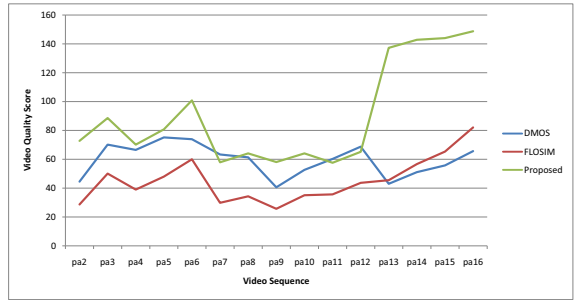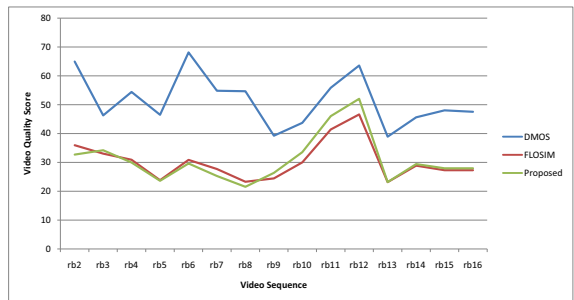

(a) Tractor

(b) Shields
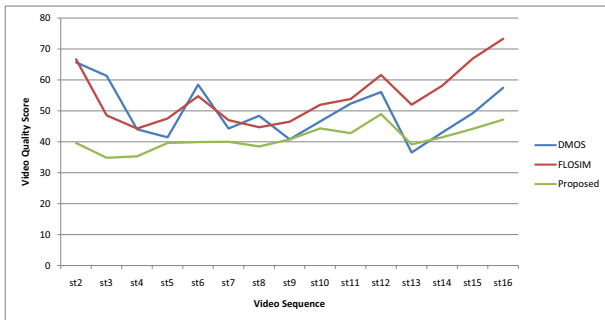
(c) Station

Fig. 8 FLOSIM and Proposed video quality scores profile compared to DMOS

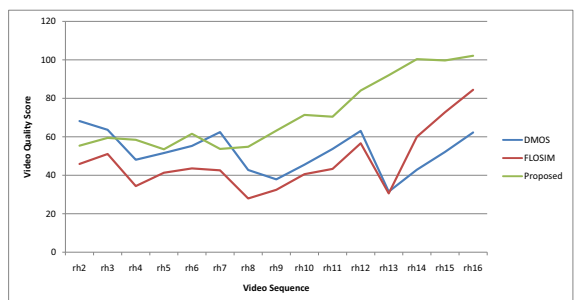**Fig. 9** FLOSIM and Proposed video quality scores profile compared to DMOS

(a) Pedestrian Area

(b) Riverbed

(c) Sunflower

(d) Rushhour

Table 2 Closeness score to DMOS

Video Sequence	FLOSIM		Proposed	
	INT-CWC (S_{m_1}) Score	Correlation with DMOS	INT-CWC (S_{m_2}) Score	Correlation with DMOS
Mobile and Calendar	0.1012	0.8728	9.8833	0.2985
Park run	0.1846	0.7879	5.4172	0.5990
Blue sky	0.2678	0.3713	3.7344	-0.1073
Pedestrian Area	0.5085	0.8974	1.9667	-0.0992
Riverbed	0.8991	0.6073	1.1123	0.4333
Sunflower	0.5404	0.7368	1.8506	0.6358
Rushhour	0.6286	0.4909	1.5908	-0.1864
Tractor	1.0411	0.8850	0.9605	0.4958
Shields	1.4863	0.5276	0.6728	0.1371
Station	2.0428	0.7746	0.4895	0.6451

scores. The video quality score profiles for video sequences with wireless distortions, IP distortions, H.264 compression, MPEG-2 compression can be seen in Fig. 7. And also all the individual video sequence video quality score profiles can be seen in Figs. 8 and 9.

From Fig. 4 and Table 2, we can observe that though the correlation coefficient is lower, for the video sequences Mobile and calendar, Park run and Blue sky(in Table 2), when compared to FLOSIM, the video quality scores obtained using the proposed algorithm are much closer to the DMOS scores. This motivates us to design the INT-CWC measure. From Table 2, Figs. 5 and 6 , we can observe that though the correlation coefficient is lower, the proposed algorithm out performs in terms of INT-CWC score when compared to the FLOSIM algorithm by Manasa et al.

4 Conclusion

Instead of using minimum eigenvalue to capture the randomness features, we used orientation feature of optical flow to improve the distortion computation. The proposed INT-CWC measure, which aimed to measure the quantitative closeness of any two video quality assessement algorithms to be compared with DMOS, is novel attempt. However, there is further scope to improve the orientation feature computation for better understanding the randomness.

References

1. Benhur O-J et al (2014) A full reference video quality measure based on motion differences and saliency maps evaluation. In: 2014 International Conference on Computer Vision Theory and Applications (VISAPP), vol 2. IEEE

2. Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157–189
3. FLOSIM (2020). [Online] http://www.iith.ac.in/lfovia/Codes/FLOSIM.zip
4. Gujjunoori S, Oruganti M (2020) An optical flow direction-based full reference video quality assessment algorithm. Int J High Performance Computing and Networking 16(Nos.2/3):148–159
5. Gunnar F (2003) Two-frame motion estimation based on polynomial expansion. Scandinavian conference on Image analysis. Springer, Berlin
6. Horn BKP, Schunck BG (1981) Determining optical flow. Artif Intell 17:185–203
7. Index (2010) Cisco Visual Networking. Global Mobile Data Traffic Forecast Update, 2009-2014, White Paper, CISCO Systems Inc 9
8. Kalpana S, Bovik AC (2009) Motion tuned spatio-temporal quality assessment of natural videos. IEEE Trans Image Process 19.2:335–350
9. Kjell B et al (2009) VQEG Validation and ITU standardization of objective perceptual video quality metrics [standards in a nutshell]. IEEE Signal Processing Magazine 26.3:96–101
10. Kui L et al (2010) Optical flow and principal component analysis-based motion detection in outdoor videos. EURASIP Journal on Advances in Signal Processing 2010.1:680623
11. Linfeng G, Meng Y (2006) What is wrong and right with MSE? Eighth IASTED. International Conference on Signal and Image Processing
12. Lucas BD, Kanade T (2020) An iterative image registration technique with an application to stereo vision. In: Proceedings of the 7th international joint conference on Artificial intelligence - (IJCAI'81), vol 2. Morgan Kaufmann Publishers Inc, San Francisco, pp 674–679
13. Manasa K, Channappayya SS (2016) An optical flow-based full reference video quality assessment algorithm. IEEE Trans Image Process 25.6:2480–2492
14. Nicolas S et al (2010) Assessing quality of experience of IPTV and video on demand services in real-life environments. IEEE Transactions on broadcasting 56.4:458–466
15. Ninassi A, Le Meur O, Le Callet P, Barba D (2009) Considering temporal variations of spatial visual distortions in video quality assess-ment. IEEE J Sel Topics Signal Process 3(2):253–265
16. Recommendation ITU-T P (1999) Subjective video quality assessment methods for multimedia applications. International Telecommunication Union
17. Recommendationitu-R BT (2002) Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union
18. Sagar G, Oruganti M (2018) HVS Based full reference video quality assessment based on optical flow. In: Proceedings of the international conference on pattern recognition and artificial intelligence
19. Seshadrinathan K, Bovik AC (2010) Motion tuned spatio-temporal quality assessment of natural videos. IEEE Trans. Image Process 19(2):335–350
20. Seshadrinathan K, Soundararajan R, Bovik AC, Cormack LK (2010) Study of subjective and objective quality assessment of video. IEEE Trans Image Process 19(6):1427–1441. https://doi.org/10.1109/TIP.2010.2042111
21. Seshadrinathan K, Soundararajan R, Bovik AC, Cormack LK (2010) A subjective study to evaluate video quality assessment algorithms. In: Rogowitz BE., Pappas TN. (eds) Human Vision and Electronic Imaging XV, vol 7527, pp 128–137. https://doi.org/10.1117/12.845382
22. Shyamprasad C et al (2011) Objective video quality assessment methods: a classification, review, and performance comparison. IEEE transactions on broadcasting 57.2:165–182
23. Video Quality Experts Group (2010) Report on the validation of video quality models for high definition video content. [Online] http://www.its.bldrdoc.gov/media/4212/vqeg_hdtv_final_report_version_2.0.zip
24. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
25. Wang Z, Lu L, Bovik AC (2004) Video quality assessment based on structural distortion measurement. Signal Process Image Commun 19(2):121–132
26. Wang Z, Simoncelli EP, Bovik AC (2003) Multiscale structural similarity for image quality assessment. In: Proc 37th Asilomar Conf Signals, Syst Comput, vol 2, pp 1398–1402
27. Wikipedia contributors (2022) Interquartile range. https://en.wikipedia.org/wiki/Interquartile_range. Accessed 22 Feb 2022