International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-09, Issue-11, November 2022

Overview of the Video Quality Measurement Techniques

Chena Ram, Subhash Panwar

Abstract— **Various network or channel errors in the network environment will result in damage or loss of video information during transmission or storage. Since very high compression has usually been applied to the compressed bit stream, any damage in the reconstructed signal at the decoder would likely lead to unpleasant visual distortion. To control the inception of potentially visible artifacts, a proper mechanized video quality measurement method is needed. This paper gives a systematic overview of available subjective and objective assessment techniques for video quality, together with performance comparisons of objective assessment methods for quality. The previous results demonstrated that the high-performance video quality predictors of perceived video quality are SSIM, MS-SSIM, RR-VQM, and VIF.**

Index Terms— **Correlation, image quality, PSNR, quality management, video coding, video compression.**

I. INTRODUCTION

Media content, such as photos, audio or videos, is extensively recycled in human daily life with the hasty development of Internet technology [1]. Globally, internet video traffic account for 82 percent of business internet traffic in 2022, up from 70 percent in 2017, according to recent projections, e.g., [2]. Internet video traffic grow four-fold worldwide from 2017 to 2022, with an annual growth rate of 33 percent. In a post COVID-19 world, understudies are utilizing video conferencing choices to go to online classes. Further, it is anticipated that the video content will undoubtedly increment by 8-10 times among now and 2023. Also, the odds are, post Coronavirus, video traffic is probably going to zoom considerably farther than the factor of 10. This has increased responsibilities on the video service providers to match the video quality expectations of the end user.

Video preparation systems can lead to certain artifact measures or deviations in the video signal, so an important problem is the estimation of video quality. Due to the growing interest in video-based applications, the reliable assessment of video quality has expanded. Over the past decade, numerous video quality evaluation techniques have been developed with shifting computational complexity and precision [3]. During collection, processing, compression, transmission, and reproduction, videos are subject to a wide range of distortions, all of which can lead to video quality degradation [4]. For example, in the quantization process, lossy video compression techniques that are almost often used to reduce the bandwidth needed for video storage or transmission, can decrease video quality. The approach of advanced video compression, storage, and transmission frameworks uncovered essential impediments of procedures and strategies that have customarily been utilized to quantify video performance [5].

There are various factors that affect the quality of video. Quality factors associated with source video include camera performance and shooting conditions like focus, contrast, and brightness. When this source video is coded, the quality of the coded video depends on the coding parameters (bit rate, frame rate, resolution, etc.) and type of video codec used. Next, the video data that is coded and transmitted over an IP network may result in some IP packet loss. Finally, the ability of user's terminal used to decode and display video data may affect the quality of video.

Since video signals are transmitted in a wide range of applications to human end users, it is extremely important that automated video quality assessment strategies are available that can result in the monitoring of the quality of video being transmitted to this basic audience [6]. As of late, the usage of video-based apps has grown due to the far-reaching use of the Internet as well as the enhancement of video technology. Consequently, video quality assessment has become critical, and numerous video quality measurement metrics have been established over the past decade. [7].

This article is arranged as follows. Section II describes the various subjective quality assessment techniques. The overview of objective quality assessments is given in section III. Section IV discusses the specifics of conventional point-based metrics. Previous results of performance comparisons of objective quality assessment techniques are given in Section V. Finally, Section VI concludes the paper.

II. SUBJECTIVE QUALITY ASSESSMENT

Subjective video quality is related to how video is perceived by the observer and gives opinion on a specific test sequence under consideration. Many aspects of viewing conditions and human psychology complicate subjective measuring methods, such as observer preference for material, vision capacity, adaptation, conversion of perception of quality into ranking score, level of ambient light, display devices, etc. Subjective video quality assessment methods are quite expensive with regard to human resources and time.

The most reliable way of video quality measurement is the subjective assessment for the reason that in different applications, human beings are the absolute receivers. For many years, the mean opinion score has been considered the trustiest method of quality assessment. For many applications, however the MOS approach is costly and slow.

Subjective quality tests are the most significant way of assessing the quality of experience of video communication

Manuscript received November 18, 2022

Chena Ram, Department of Computer Science and Engineering, Engineering College, Bikaner, India

Subhash Panwar, Department of Computer Science and Engineering, Engineering College, Bikaner, India

services. The assessment of subjective quality requires visual psychological tests where human observers go through a video stimulus and determine its quality depends on their intrinsic subjective judgment. Evaluation equipment's and specialist knowledge are required to plan and implement the tests, along with a suitable way of assessments, the selection of a video display method and adjusting the viewing conditions.

A. *Double Stimulus Continuous Quality Scale (DSCQS)*

In the DSCQS system, pair of videos consists of the source video and a damaged video is shown twice. The source video is not known to the subjects, and the videos are shown in random order. The subjects continue records the quality of videos on an assessment scale normalized to the range 0 – 100. The differential value is calculated by subtracting the score for the impaired video from the score for the original reference video and averaged against all the subjects to make a final DSCQS value. Since the DSCQS value is measured from differences in picture quality, a lower value indicates higher quality and a higher value indicates lower quality.

B. *Double Stimulus Impairment Scale (DSIS)*

ITU-R BT.500-11 defines the DSIS process. This method is often referred to as the DCR (Degradation Category Rating), in which videos are displayed in pairs. The first reference video is shown and the evaluator is well aware, and then the affected video is shown. The original and the impaired videos are shown one after the other in small segments of a few seconds each, in the same session, and subjects' rate both video sequences with the help of sliders. The disparity among the quality ratings of the original and the impaired test sequences results in a verdict of subjective impairment.

C. *Single Stimulus Continuous Quality Evaluation (SSCQE)*

The SSCQE approach is a continuous subjective video quality assessment that does not use any reference video. Instead, subjects watch a program that is usually 30 minutes in length and continuously give scores on a slide in accordance with the instantaneously perceived quality. Subjects continuously rate quality of video on a five segments linear scale marked with adjectives to act as guides.

D. *Absolute Category Rating (ACR)*

Absolute Category Rating is a single-stimulus process, where single test video is revealed to the subjects. Subjects give score for each video under test individually without using an explicit reference for comparison. The subjects give only one score for the overall quality of video with the help of a scale divided into five discrete levels ranging from bad to excellent. From the scores obtained between the mean opinion scores of each test object and its corresponding concealed reference, a differential score (DMOS) is computed. Accuracy increases with the number of participants [8].

E. *Pair Comparison (PC)*

It's a double-stimulus process. In this process, the test video clips taken under different conditions within the same scene are paired in a number of possible variations, and for each pair, the subjects give preference. A matrix tracks the outcomes of the paired comparison experiment, each vector corresponding to the frequencies preferred by the stimulus over another stimulus. The Thurstone-Mosteller's or Bradley-Terry's model then transforms these data into scale values. Although the technique has been seen to be very effective, it is time consuming [8].

III. Objective Quality Assessment

A growing interest has been seen in the evolution of quantitative video quality assessment methods that will immediately be able to determine the perceptual consistency of video sequences. Such techniques are desired for a wide array of applications and are also useful tools for image and video database systems. Objective measurements are planned to be determined by human judgement and its reliability is dependent on its relationship to the subjective test conclusions [8].

Essentially, these methods are employed in three different ways. To begin with, they may be used to track the content of video for quality management processes. Secondly, they can also be used for benchmarking algorithms and video system planning. Third, it is also possible to update parameter settings and algorithms in video handling applications. [9]. With the introduction of highly effective video coding technologies, there is a significant requirement for measurements to be able to calculate and evaluate the delivery and efficiency of video coding as required by the end-user [3].

In order to provide cost-effective, user-friendly services over networks, video communication services must be designed and managed with the help of quality assessment technology having the right stuff of quantifying quality of experience (QoE). Psychological assessment of observers is fundamentally important for evaluating the subjective video quality, but this needs special evaluation equipment and lots of time and human effort. This makes it very difficult to improve the efficiency of video quality assessment and design of services.

A. *Full Reference Methods*

The full reference model is a method of objectively predicting a processed video stream by comparing the information of a raw video sequence before distortion has occurred and a distorted video sequence. They need complete reference footage to be usable, normally in uncompressed and lossless form, and typically need exact spatial-temporal synchronization with the calibration of color and luminance between the two videos, such that each pixel in each frame can be compared to its equivalent in the other video frame for comparative analysis. Since the full reference model makes comparison between the source video and processed video, it gives a highly reliable objective video quality assessment. The main drawback is that it needs a huge amount of data from the source video footage to make an accurate comparison, and thus inculpates environments that can tackle the significant costs of acquiring the original source video and cannot be readily implemented in user's homes.

B. Reduced Reference Methods

The goal of the Reduced Reference Image Quality Assessment (IQA) is to make use of less reference image data and to achieve better evaluation accuracy [10]. Reference videos need a significant amount of disk space and are in most cases, difficult to have for certain applications. Reduced-reference content evaluation includes only partial reference information from the source video that is accessible from an ancillary data channel. The reduced reference model is a method of objectively measuring video quality of a degraded video by comparing impaired content and a limited volume of attribute data from the original source video. As the reduced reference model uses the features of the source video and the manipulated video, it is reasonably reliable, but not as reliable as the full reference method. Since the reduced reference model involves a limited amount of information from the original video for objective measurement, it is important to calculate the transfer of the feature data.

C. No-Reference Methods

The no-reference method objectively assesses quality of video by only using processed video. Since in no-reference method, information is not required from the original video, it can be used in various environments. However, since the method does not use feature data from the original video, it is less authentic in measuring the quality of video than the FR and RR methods. The development of no-reference algorithms for objective quality evaluation is very difficult. This is largely attributed to the lack of knowledge of the human sensory system as well as the related functional dimensions of the human brain. In the literature, only a few metrics have been suggested for quantitative non-reference quality measurement, but this issue has recently received a huge amount of publicity [11]. We would expect to see far-flung application of this method for regular quality monitoring in users' homes.

IV. Traditional Point-Based Metrics

The two popular methods of determining the efficiency of advanced video processing systems are signal-to-noise ratio (SNR) and peak signal-to-noise ratios (PSNR). The most common objective video content assessment method is PSNR. However, owing to the non-linear nature of the human vision, the PSNR values are not fully correlated with the perceived visual output [18].

A. Peak Signal-to-Noise Ratio (PSNR)

PSNR is most widely used for calculating the efficiency of the restored image of the lossy encoding codecs. PSNR is the ratio of signal to noise between the reference signal and the distortion signal in the picture given in dB. The signal for this is the calculation of the original input, and the noise is the inaccuracy of the compression. When correlating compression codecs, the PSNR is similar to the human understanding of the consistency of restoration. The higher the Peak Signal-to-Noise Ratio, the closer the reconstructed image is to the original image. Broadly speaking, a higher PSNR value should correlate to a higher quality image. Even though a high PSNR generally point to the reconstruction is of superior quality, in a number of cases it may not. It needs to be exceedingly cautious about the extent of usefulness of this metric; it is only decisively applicable when it is used to judge the effects of identical data of the same codec or codec category. However, PSNR is a popular quality measurement technique because it's simple and fast to work out while still giving acceptable results.

PSNR is based on the mean squared error (MSE) associated with the maximum possible luminance value (with a distinctive 8-bit value of 2^8-1 = 255) as follows:

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j)-F(i,j)]^2}{M \cdot N} \quad (1)$$

$$PSNR = 20 \cdot \log_{10}\left(\frac{255}{\sqrt{MSE}}\right) \quad (2)$$

Where $f(i,j)$ is the original signal at the pixel (i,j), $F(i,j)$ is the reproduced signal and, $M \cdot N$ is the size of the image. The outcome is a single number in dB, varying from 30 to 50 for medium to high-definition footage.

While multiple objective models of video quality have been introduced over the last two decades, Peak Signal-to-Noise Ratio remains to be the most common approximation of the quality difference between videos.

B. Structural Similarity Index (SSIM)

The signals of natural images are strongly structured: their pixels display signs of strong dependency, particularly when they are spatially near, and they provide valuable information on the structure of objects within the visual scene [4]. The structural similarity of the image quality model relies on the premise that the human visual system is very well suited to the separation of structural data from the image. A good estimate of the perceived image quality can therefore be given by a structural similarity measure [12].

The SSIM Index is a full reference quality assessment methodology. SSIM seeks to advance conventional approaches similar to PSNR and MSE, which have been found to be incompatible with the vision of the human eye. SSIM considers image deterioration as perceived alteration in structural details. Structural knowledge is a scheme that has well-built interdependencies between pixels, especially when they are spatially nearby. Such dependencies provide essential details about the structure of the objects in the image.

Structural Similarity Index [4] is specific to still image quality assessment. The SSIM Index is derived from the calculation of three components (similarity of luminance, similarity of contrast and structural similarity) and from the integration of the three components into the final result value. Luminance is modelled as average pixel intensity, contrast by means of variance among the distorted as well as reference image, and structure by means of cross-correlation among the two images. The follow-on values are combined (by means of exponents specified as alpha, beta, and gamma) and averaged to generate the final value of the SSIM index. SSIM describes a comparison of the luminance

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (3)$$

Where μ_x and μ_y denotes the mean luminance intensity of the image signals X and Y relative to each other. In the case of an image with a dynamic range L, the constant of stabilization is set to $C_1 = (K_1L)^2$ where K_1 is a small constant, so that C_1 will take place only if ($\mu_x^2 + \mu_y^2$) is small. In the same way, SSIM determines a contrast correlation function

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{4}$$

With σ_x and σ_y representing the standard luminance sample deviations of the two images and C_2, the stabilizing constant is identical to C_1. In addition, the structure comparison function can be expressed as the covariance of the luminance samples σ_{xy} as

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{5}$$

The SSIM index shall be specified as

$$SSIM(x,y) = [l(x,y)]^\alpha \cdot [c(x,y)]^\beta \cdot [s(x,y)]^\gamma \tag{6}$$

In order to adjust the relative importance of the three different functions, the positive parameters α, β, and γ adjust. Setting $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$ gives the specific form

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{7}$$

examined in [4]. The overall image quality, i.e., the mean SSIM (MSSIM) index, is calculated as the average quality map.

C. Multi Scale Structural Similarity Index (MS-SSIM)

Traditional image quality evaluation approaches are focused on a bottom-up methodology that aims to replicate the characteristics of the appropriate components of the early human visual system [12]. An alternate and complementary approach to the issue of image quality measurement is given by structural similarity. It depends on a top-down premise that the human visual system is very well adapted to taking structural details out of the scene and as a result, an elevated estimate of the observed picture quality can be a calculation of structural resemblance.

Multi-Scale Structural Similarity index is layered on SSIM. It is based on the SSIM metric of many downgraded stages of the original images. The outcome is the weighted average of these metrics. At each point, a low pass filter for the reference and impaired images and the filtered images down sampled by a factor of two are added to the MS-SSIM process. Contrast and structure comparisons are measured on the m^{th} scale and marked as $c_m(x,y)$ and $s_m(x,y)$, respectively. The luminance relation shall be measured at scale M (i.e., the largest scale gained after M − 1 iterations) and shall be marked as $l_M(x,y)$. Combining the scales would give

$$MS-SSIM(x,y) = [l_M(x,y)]^\alpha \cdot \prod_{m=1}^{M}[c_m(x,y)]^\beta \cdot [s_m(x,y)]^\gamma \tag{8}$$

Revealed to outperform the SSIM index and numerous other algorithms for assessment of still image consistency.

MS-SSIM places less emphasis on the luminance factor than on the contrast and structural elements. Overall, MS-SSIM has been established to improve the correlation between the MS-SSIM index and the subjective quality evaluation. Conversely, the trade-off is that MS-SSIM needs more time to run than the simple SSIM.

D. Video Structural Similarity Measure (V-SSIM)

A variety of changes to the SSIM have recently been suggested and adaptations to the video evaluation have been added. V-SSIM is a full reference video quality measure used to quantify structural distortion as an approximation of perceived visual distortion. The weighted feature of SSIM frame ratings is the V-SSIM rating. The selection of weights is based on the assumption that they are less powerful in the dark areas. As the darkness will rely on the screen used, this is uncertain. [8].

E. Visual Information Fidelity (VIF)

Latest literature on brain theory indicates that the HVS dynamically predicts key visual data and attempt to ignore the confusion that persists in vision interpretation and comprehension. Perceptual consistency thus depends on the precision of the information intended for primary visual information as well as the residual uncertainty [10]. Visual Information Fidelity [13] is determined by visual statistics combined with human visual system modelling.

Only gray-scale images with a luminance scale [0, 255] can be treated by VIF. Therefore, before operating with VIF for color pictures, we need to convert it to a [0, 255] gray-scale format. This can usually be achieved via the rgb2gray MATLAB routine. Visual information fidelity for fusion (VIFF) has recently been suggested in [17]. VIFF is a multi-resolution image fusion metric used to systematically test fusion efficiency using visual information fidelity. In relation to several current fusion methods, the proposed fusion calculation approach is contrasted with the subjective study database that Petrovic presented. With respect to both, computational complexity and corresponding human expectations, it is observed that VIFF performs well.

F. Video Quality Metric (VQM)

The VQM is a structured system of objective assessment of video that closely anticipates the subjective scores that a jury of human viewers will gain. VQM [5] has been established by the Institute for Telecommunication Science (ITS) to make an objective assessment of the overall quality of videos available. An initial video clip and a processed video clip are contrasted with the VQM program and a VQM which correlates with perception is recorded. The VQM scores range from zero to one with no distortion at zero and a nominal maximum distortion at one [14]. The test results indicate that VQM also has strong correlation towards subjective video quality judgment and it has been accepted by ANSI as an objective video quality standard.

G. MOtion-based Video Integrity Evaluation (MOVIE) index

In moving image scenes, motion plays a key role in human vision. Motion gives significant clues about the shape of three-dimensional objects and helps distinguish objects. Many existing algorithms for video quality estimation are capable of capturing spatial distortions that take place in video sequences, but do not do enough work to capture temporal distortions. More recently, Seshadrinathan and Bovik have introduced a FR video quality method called the MOtion-based Video Integrity Evaluation (MOVIE) Index [6]. By monitoring distortions around motion trajectories that are perceptually important, MOVIE incorporates explicit motion data into the framework of video content evaluation, thus increasing the measurement of spatial artifacts in videos.

Two indexes, the Spatial MOVIE index in particular, holds spatial artifacts and a Temporal MOVIE index which holds temporal artifacts, were defined in the model. First, using a Gabor filter family, the source and test video are subdivided into spatio-temporal bandpass channels. A system motivated loosely by the SSIM index and the theoretical methods of information for image quality assessment achieves spatial quality assessment. Using motion data from the series of source videos, temporal accuracy is measured. Finally, to achieve an aggregate video integrity ranking referred to as the MOVIE index, the spatial and temporal consistency ratings are pooled.

H. Visual Signal-to-Noise Ratio (VSNR)

The Visual Signal-to-Noise Ratio metric submitted by Chandler and Hemami [15] is basically a FR still-image quality model, but when applied on a frame-by-frame level and then averaged, it has also shown impressive results in measuring video quality.

The VSNR works through a two-stage technique. Firstly, contrast thresholds for distortion detection in the case of natural images are determined by means of wavelet-based visual masking schemes as well as visual summation to decide if distortions are noticeable in the impaired image. It is thought that the distorted image is of perfect visual fidelity (VSNR = ∞) if the distortions are smaller than the detection threshold, and no further analysis is required. A second stage is used if the distortions are above threshold, which functions on the basis of the visual low-level property of apparent contrast and the visual mid-level property of global precedence. In the distortion-contrast space of multi-scale wavelet decomposition, these two properties are modelled by Euclidean distances, and VSNR is determined on the basis of a straightforward linear sum of these distances.

V. PERFORMANCE COMPARISON

A. Figures and Tables

In recent years, the estimation of perceptual video quality has become particularly important and different video quality measuring methods have been introduced. The question of their relative merits and demerits inevitably emerges with too many quality evaluation algorithms proposed.

The authors in [3], have conducted independent performance comparisons for six objective quality metrics and results from common objective video quality estimation approaches have been seen with sequences from the LIVE video database. It is observed that the natural visual statistics related MS-SSIM, the natural visual feature related VQM, as well as the spatio-temporal frequency-domain related MOVIE index provide high performance in favour of LIVE Video Quality Database.

In [7], a comparison of different methods was made with respect to precision, stability, monotonicity, in addition to the criterion for complexity vs. quality. The findings suggest that the content of the video, the resolution and the type of distortion have a substantial effect on the accuracy of metrics for video quality evaluation. FMSE and MOVIE are measures that typically achieve a good correlation by means of subjective effects for any of the datasets along with all forms of distortion; however, MOVIE has complexity greater than FMSE. An example is the distortion induced by IP transmission where none of the methods evaluated revealed reasonable precision and stability. In [8], the consistency among objective and subjective measurements is determined by the estimation of the coefficients of correlation for the total calculated points. None of the metrics evaluated in the findings cross 50 percent of individual judgement.

Table 1: COMPARISON OF OBJECTIVE VIDEO QUALITY METRICS

Method	Test-Details	Performance			
		PCC	SROCC	OR	RMSE
SSIM [4]	VQEG Phase I, LIVE Image database	0.967	0.963	0.041	5.06
V-SSIM [9]	VQEG Phase I	0.864	0.812	0.578	-
MS-SSIM [12]	VQEG Phase I, LIVE Image database	0.969	0.966	1.160	4.91
RR-VQM [5]	VQEG FRTV Phase II	0.938	-	0.460	-
MOVIE [6]	VQEG FRTV Phase I	0.821	0.833	0.644	8.09
VIF [13]	29 test images, 5 distortion types	0.949	0.949	0.013	5.08
RR-VIF [10]	LIVE Image database	0.725	0.732	-	17.15
VSNR [15]	LIVE Image database	0.889	0.889	-	7.39

We compare the approaches mentioned in Table 1 to review the output of a representative set of video performance measures surveyed. The illustrated results are obtained for given metrics on specific image or video quality database available online. From Table I, we find that, relative to other metrics, the SSIM, MS-SSIM, RR-VQM, and VIF metrics contributes to the maximum PCC and SROCC values, suggesting a strong correlation with subjective ratings. Furthermore, relative to the other metrics, the SSIM, MS-SSIM, and VIF methods have the smallest values for OR and RMSE suggesting lowest difference between the objective and subjective ratings. Thus, of these eight quantitative quality metrics, SSIM, MS-SSIM, and VIF are the top scoring approaches.

VI. CONCLUSION

This paper offers a detailed review of available subjective and objective video quality evaluation approaches, along with performance comparisons with objective quality evaluation methods. It is assumed that the best video quality metrics are SSIM, MS-SSIM, RR-VQM, and VIF. There is a wider potential for enhancing accurate video quality measures which, using a variety of different databases and media content, obtain maximum efficiency. A more sequenced affirmation process should be sought after as indicated in [16] to show important outcomes just as to have a typical reason for the examination of different methods.

REFERENCES

[1] S. Lian, Z. Liu, Z. Ren, and H. Wang, “Commutative Encryption and Watermarking in Video Compression,” *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 17, no. 6, pp. 774-778, Jun. 2007.

[2] Cisco, Inc., “Visual networking index: Global Networking Trends Report, ” 2020.

[3] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, “Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison,” *IEEE Trans. broadcasting*, vol. 57, no. 2, pp.165-182, Jun. 2011.

[4] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[5] M. Pinson and S. Wolf, “A New Standardized Method for Objectively Measuring Video Quality,” *IEEE Trans. Broadcasting*, vol. 50, no.3, pp. 312-322, Sep. 2004.

[6] K. Seshadrinathan and A. C. Bovik, “Motion tuned spatio-temporal quality assessment of natural videos,” *IEEE Trans. Image Processing*, vol. 19, no. 2, pp. 335–350, Feb. 2010.

[7] M. Vranjes, S. Rimac-Drlje, and K. Grgic, “Review of objective video quality metrics and performance comparison using different databases,” *ELSEVIER Journal Signal Processing: Image Comm.*, vol. 28, issue 1, pp. 1-19, Jan. 2013.

[8] E. Bosc, P. Callet, L. Morin, and M. Pressigout, “Visual quality assessment of synthesized views in the context of 3DTV,” *3D-TV System with Depth-Image-Based Rendering Architectures, Techniques and Challenges*, hal-00748518, ver. 1, pp. 439-474, Nov. 2012.

[9] Z. Wang, L. Lu, and A. Bovik, “Video quality assessment based on structural distortion measurement,” *Signal Processing Image Comm.*, vol. 19, no. 2, pp. 121–132, Feb. 2004.

[10] J. Wu, W. Lin, G. Shi, and A. Liu, “Reduced-Reference Image Quality Assessment with Visual Information Fidelity,” *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1700-1705, Mar. 2013.

[11] Z. Wang, H. R. Sheikh, and A. C. Bovik, “Objective Video Quality Assessment,” Chapter 41 in the Handbook of Video Databases: *Design and Applications*, B. Furht and O. Marqure, ed., CRC Press, pp. 1041-1078, Sep. 2003.

[12] Z. Wang, E. Simoncelli, and A. Bovik, “Multiscale structural similarity for image quality assessment,” *Conf. Rec. 37th IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1398–1402, Nov. 2003.

[13] H. Sheikh and A. Bovik, “Image information and visual quality,” *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[14] M. Noorkami and R. M. Mersereau, “Digital Video Watermarking in P-Frames with Controlled Video Bit-Rate Increase,” *IEEE Trans. Info. Forensics and Security*, vol. 3, no. 3, pp. 441-455, Sep. 2008.

[15] D. Chandler and S. Hemami, “VSNR: A wavelet-based visual signal-to-noise ratio for natural images,” *IEEE Trans. Image Processing*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[16] C. Keimel, T. Oelbaum, and K. Diepold, “Improving the verification process of video quality metrics,” *Proc. of IEEE International Workshop on Quality Multimedia Experience (QoMEX)*, pp. 121-126, Jul. 2009.

[17] Y. Han, Y. Cai, Y. Cao, and X. Xu, “A New Image Performance Metric Based on Visual Information fidelity,” *Info. Fusion Journal*, vol. 14, issue 2, pp. 127-135, Apr. 2013.

[18] C. Ram and M. Dutta, " Performance analysis of High Efficiency Video Coding (HEVC) relative to H.264 AVC," *PISER Journal*, vol. 2, issue 5/6, pp. 79-93, Sep. 2014.

Authors

Chena Ram received the Bachelor of Engineering degree in Electronics and Communication Engineering from University of Rajasthan, Jaipur, India, in 2004 and M. E. degree in Computer Science and Engineering from Punjab University, Chandigarh, India, in 2014. He is currently pursuing PhD degree in the Department of Computer Engineering, Rajasthan Technical University, Kota, India. Since 2004, he is working in Department of Electronics and Communication Engineering, Government Engineering College, Bikaner, India. His research interests are in image and video processing, deep learning, computer network, and information theory and coding.

Subhash Panwar received his BE degree in Computer Science and Engineering from University of Rajasthan, Jaipur, India in 2004, MTech degree in Computer Engineering from Motilal Nehru National Institute of Technology, Allahabad, India, in 2010, and PhD degree in Computer Engineering from the Malaviya National Institute of Technology, Jaipur, India in 2015. He is working in Department of Computer Science and Engineering, Government Engineering College Bikaner, Rajasthan, India from September 2004. His current research interests include image fusion, computer vision, computational intelligence, machine learning and pattern recognition.