RESEARCH ARTICLE OPEN ACCESS

Performance of Full Reference Video Frames quality evaluation metrics by using First order and Second order Edge detection operators

Dr. K.Sai Prasad Reddy*, Prof. K. Nagabhushan Raju**
* Department of Instrumentation, Sri Krishnadevaraya University, Anantapur-03
** Department of Instrumentation, Sri Krishnadevaraya University, Anantapur-03

ABSTRACT
Assessment of visual quality is critical for many image and video processing applications. A quality assessment (QA) algorithm predicts visual quality by comparing a distorted video frame against a reference video frame. Various procedures are proposed for measuring the quality of the video frame. Video frame quality assessment plays an important role in the field of video processing. Structural Similarity Index (SSIM) and Visual Information Fidelity (VIF) has become a standard among image quality metrics. SSIM is a framework for quality assessment based on the degradation of structural information of video frame. Structural Similarity Index (SSIM) is used to assess the resemblance between the reference video frame and the distorted video frame. Measure of structural similarity index is able to provide a good approximation to perceived video frame quality. This procedure evaluates the visual impact of changes in luminance, contrast and structure in an image. In this paper Structural Similarity Index and Visual Information Fidelity values are calculated and compared for live video frames by applying different approaches using different color models to assess the quality of the frames. Experimental results comparisons demonstrate the effectiveness of the proposed method.
Keywords – Color Models, Edge Detection, Edge Detection Operators, Structural Similarity (SSIM) Index, Visual Information Fidelity (VIF)

Date of Submission: 10-07-2021 Date of Acceptance: 26-07-2021

I. INTRODUCTION

Edges are major local intensity changes in video frames as well as in images. These edges are key feature for analyzing an image and Video frame. To separate regions in an object or to make out changes in illumination or color edges are important clues. Edges are main features in early vision stages in human eye. Detection of edges recognizes abrupt changes or discontinuities in image. The major objective is to extract information of two-dimensional projection of three-dimensional scene. Image segmentation, region separation, are the secondary goal which also includes objects description and recognition. Coordinates [i, j] are edge point coordinates in image at the location where major local intensity changes occur in image or video frame. Edge fragment is a tiny line segment with reference to the size of a pixel, or a point with orientation attributes. Edge is commonly used for edge points and edge fragments. Edge detection algorithm generates a set of edges from image or video frame. Some edge detection algorithms generate tangent direction of the arc which passes through the pixel. Contour is a list of edges of mathematical curve which models the edges list[1]. Edge linking is a process which forms structured list of edges from unordered list. Edge following is procedure of searching the edge image to decide contours. Source of edges are Surface normal discontinuity, Depth discontinuity, Surface color discontinuity and Illumination discontinuity. Important features like corners, lines and curves of an image or video frame can be extracted by using edge detection algorithms. These features are used by computer vision algorithms. The optimal edge detector is supposed to trace all real edges by ignoring noise and other artifacts. The detected edges must be same to the true edges. The edge detector has to return only one point for every true edge point. Cues of edge detection are differentiation in color, intensity, or texture across the boundary and continuity and closure.

II. DIFFERENT COLOR MODELS

A color model is conceptual mathematical model which describes how the colors can be characterized as database of numbers, typically represents as three or four values or color components. A color space provides normal process to indicate order, manipulate and effectively displays the object colors taken into consideration. The selected color model should be well matched in order to deal with the problem's statement and solution.

2.1 RGB Model

In order to reproduce a wide array of colors, Red (R), Green (G) and Blue (B) colors are put together added in different ways. This RGB color model is additive color model. The name of this RGB color model is derived from the initials of the Red, Green and Blue colors. These three colors are additive primary colors. In this model an image is represented by these image planes Red, Green and Blue. Stating a particular color is by specifying the amount of each of the primary components present. Geometry of the RGB color model represents the identifying colors using Cartesian coordinate system. In order to obtain any color in the visual spectrum these three colors Red, green, and blue are combined in a variety of proportions [1]. The range of levels of Red, Green and Blue is from 0 to 100 percent of full intensity. Numerical value of each intensity value will be in between 0 to 255 and in hexadecimal range will be from 00 to FF.

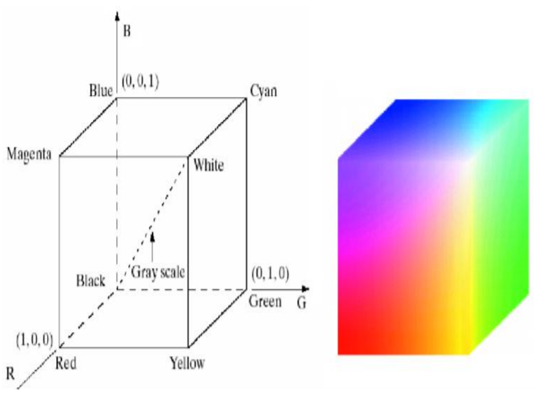


Fig.1: The RGB color cube. The grey scale spectrum lies on the line joining the black and white vertices.

2.2 YUV Color Model

YUV color model is extracted from RGB color model. It comprises of the luminance(Y) and two color difference (U, V) components. The luminance value will be calculated as the weighted sum of Red, Green and Blue colors [2]. The color difference which is chrominance component is formed by subtracting luminance from the colors Blue and Red .YUV is a color space typically used as part of a color image pipeline. The YUV model characterizes a color space in terms of luma (Y′) and chrominance (UV) components. Luma component which is the brightness is represented by Y′ and two chrominance which are color components are represented by U and V. Representation of luminance is by Y where as luma by Y′. The gamma compression is denoted by prime symbol (') with luminance as perceptual brightness, where as luma is electronic brightness [3]. The value of Y ranges from 0 to 1, in digital format ranges from 0 to 255, where as the values of U and V ranges from -0.5 to 0.5, in signed digital form ranges from -128 to 127, in unsigned form ranges from 0 to 255.

Conversion between RGB – YUV

$$R = Y + 1.4075*(V-128) \quad (1)$$
$$G = Y - 0.3455*(U-128) - (0.7169*(V-128)) \quad (2)$$
$$B = Y + 1.7790*(U-128) \quad (3)$$
$$Y = R*0.2990 + G*0.5870 + B*0.1140 \quad (4)$$
$$U = -0.14713*R - 0.28886*G + 0.436*B \quad (5)$$
$$V = 0.615*R - 0.51499*G - 0.10001*B \quad (6)$$

2.3 YCbCr Color Model

For digital video the most extensively used color space is YCbCr color space. In YCbCr color space, luminance information is stored as single component (Y), and chrominance or color information is stored in form of two color difference components (Cb and Cr)[4]. The difference between blue value and a reference value represents Cb and Cr is the difference between red value and a reference value [4]. YCbCr data is able to be double precision. The range of Y data is [16, 235], and for Cb and Cr the range is [16, 240].

$$Y = 0.299*R + 0.587*G + 0.114*B \quad (7)$$
$$Cb = 128 - 0.168736*R - 0.331264*G + 0.5*B \quad (8)$$
$$Cr = 128 + 0.5*R - 0.418688*G - 0.081312*B \quad (9)$$

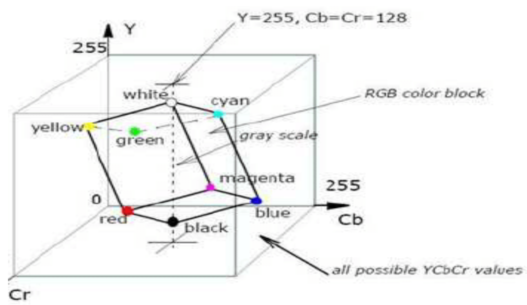


Fig.2: YCbCr and YUV color models[5],[6]

2.4 CIE XYZ Color Model

The XYZ color space is an international standard developed by the Commission Internationale de I Eclairage (CIE). This model is based on three hypothetical primaries, XYZ and all visible colors can be represented by using only positive values of X, Y, Z. The CIE XYZ primaries are hypothetical because they do not correspond to any real light wave lengths. The Y primary is calculatedly defined in order match to luminance, where as X and Z primaries provides color information. One of the major advantages of the CIE XYZ model is it is completely device independent. The block location of RGB representable colors in the XYZ space is shown in Fig.3

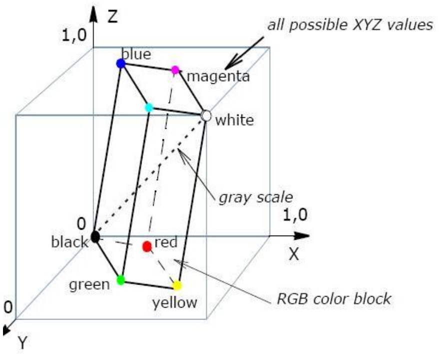


Fig.3 RGB colors Cube in the XYZ color space [5]

Conversion between RGB and XYZ color space:

$X=0.412453*R+0.35758*G+0.180423*B$ (10)
$Y=0.212671*R+0.71516*G+0.072169*B$ (11)
$Z=0.019334*R+0.119193*G+0.950227*B$ (12)
$R=3.240479*X-1.53715*Y-0.498535*Z$ (13)
$G=-0.969256*X+1.875991*Y+0.041556*Z$ (14)
$B=0.055648*X-0.204043*Y+1.057311*Z$ (15)

III. EDGE DETECTION OPERATORS

3.1 Laplacian Operator

To find edges in an image or video frame derivative mask called Laplacian Operator is used. The key distinction between Laplacian and other operators like Prewitt, Sobel, Robert is all these are first order derivative masks where as Laplacian is second order derivative operator[7]. There are two types of Laplacian masks. The first one is Positive Laplacian mask and other is Negative Laplacian mask. One more difference is Laplacian operator takes out edges as Inward edges and Outward edges but other operators take out edges in particular direction.

3.1.1 Positive Laplacian Operator

In Positive Laplacian mask center element should be negative and all corner elements should be zero shown in Fig.4.1. Positive Laplacian Operator takes out outward edges in an image or in video frame.

3.1.2 Negative Laplacian Operator

In negative Laplacian operator center element should be positive and all corner elements are zero and rest of the elements in mask should be -1 as shown in Fig.4.2.

0	1	0
1	-4	1
0	1	0

Fig.3.1 Positive Laplacian

0	-1	0
-1	4	-1
0	-1	0

Fig.3.2 Negative Laplacian

3.2 Sobel operator

The Sobel operator will have a pair of 3 X 3 convolution kernels. The Fig.5.1 represents 3 X 3 mask along x-axis and Fig.5.2 represents 3 X 3 mask along y-axis. These two kernels are designed in such way that they respond utmost to edges running vertically as well as horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels will be applied independently to input image in order to produce separate measurements of the gradient component in each orientation. These kernels can be combined together to find absolute magnitude of the gradient at each point and the orientation of that gradient [8].

-1	0	1
-2	0	2
-1	0	1

Fig.3.3: 3 X 3 mask along X-axis

-1	0	1
-2	0	2
-1	0	1

Fig.3.3: 3 X 3 mask along Y-axis

3.3 Robert operator

Robert operator is a first-order edge detection operator. This operator exercises partial differential operator to find the edge. Robert operator uses approximation between the two adjacent pixels of the diagonal direction of the gradient amplitude to detect edge [8]. In 2x2 diagonal derivative two convolution kernels Gx and Gy are shown in Fig.6 respectively.

Roberts operator is defined as:

$$G(x, y) = \{[\sqrt{f(x, y)} - \sqrt{f(x+1, y+1)}]2 + [\sqrt{f(x+1, y)} - \sqrt{f(x, y+1)}]2\}1/2 \quad (16)$$

Gradient size of Robert operator represents the edge strength of the edge and direction of the gradient and the edge are vertical. The operator edge has higher positioning accuracy, but it is easy to lose a part of the edge. The operator with a steep low-noise image corresponds best.

1	0
0	-1

1	0
0	-1

Fig. 6 Robert Mask

3.4 Prewitt operator

The Prewitt operator is one type of an edge model operator. This model operator is prepared from ideal edge sub-image composition. Detect the image using edge model one by one, and take the maximum value of the model operator that is most similar to the detected region as the output of the operator [9]. Prewitt operator and Sobel operator uses the similar differential and filtering operations and the difference is template do not use the similar image. This gradient based edge detector is estimated in the 3x3 neighborhood for 8 directions [8]. All the eight convolution masks are calculated. The convolution mask with the largest module is then selected. The convolution masks of the Prewitt detector are given below:

1	1	1
0	0	0
-1	-1	-1

-1	0	1
-1	0	1
-1	0	1

Fig.7 Prewitt Mask

IV. STRUCTURAL SIMILARITY INDEX (SSIM)

Structural similarity (SSIM) index is quality assessment method to calculate the perceived quality of images and video frames. At first it was designed and developed in Laboratory for Image and Video Engineering (LIVE) in university of Texas at Austin. The Structural similarity (SSIM) index algorithm was designed and developed jointly with Laboratory for Computational Vision (LCV) in New York University. Similarity between two images or video frames can be evaluated by using Structural similarity (SSIM) index [10]. The Structural similarity (SSIM) index is a full reference metric where distortion-free image as reference in measuring or predicting the quality of image or video frame. The traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE) are inconsistent with human visual perception. To overcome this problem Structural similarity (SSIM) index is designed and developed. "Mean squared error(MSE) and Peak signal-to-noise ratio (PSNR) estimates absolute errors where as SSIM is a perception-based model that considers image degradation as perceived change in structural information by incorporating important perceptual phenomena, including both luminance masking and contrast masking terms. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene. Luminance masking is a phenomenon whereby image distortions tend to be less visible in bright regions, while contrast masking is a phenomenon whereby distortions become less visible where there is significant activity or "texture" in the image"[10] . The SSIM index is calculated on various windows of an frame. The measure between two windows and of common size N×N is

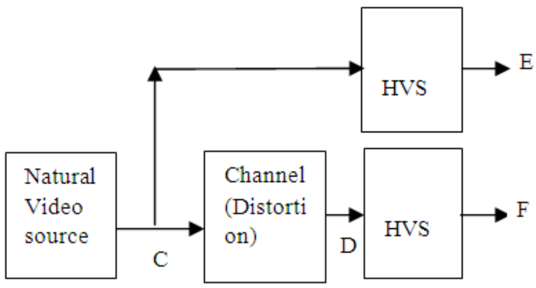
$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (17)$$

μx the average of x
μy the average of y
$\sigma x2$ the variance of x
$\sigma y2$ the variance of y
σxy the covariance of x and y
c1=(k1L)2 , c2=(k2L)2 two variables to stabilize the division with weak denominator
L the dynamic range of the pixel-values
K1=0.01 and k2 = 0.03 by default

"The SSIM index satisfies the condition of symmetry: SSIM(x,y)=SSIM(y,x). In order to evaluate the image quality, this formula is usually applied on luma, color values or chromatic values. The resultant SSIM index is a decimal value between 0 and 1" [1].

V. VISUAL INFORMATION FIDELITY (VIF)

Visual Information Fidelity (VIF) is based on the amount of mutual information shared by the reference and distorted video frames. The visual quality of the distorted video is strongly related to relative information present in the distorted video. The distortion is considered as the loss of video information and this is used to calculate the Video Quality Assessment metrics.



The VIF metrics have shown improved performance over many of the existing Video Quality Assessment algorithms. The main disadvantage of VIF is its computational complexity [11]. Visual information fidelity measurement for Video Quality Assessment quantifies the loss of video information to the distortion process and explores the relationship between video information and visual quality [12].

Distortion Model:

$$D = GC + V = \left\{ g_i \vec{C_i} + \vec{V_i} : i \in I \right\} \quad (18)$$

D denotes the random field (RF) from a sub band in the reference signal. G denotes deterministic scalar gain field. C stands for the RF from a sub band in the reference signal. V denotes additive white Gaussian noise RF. Distortions like blur, additive noise, and global or local contrast changes will be captured by this model.

Properties of VIF:

- VIF is bounded below by zero
- VIF is exactly unity if calculated between the reference image and its copy
- A linear contrast enhancement of the reference image will result in a VIF value larger than unity, signifying a superior visual quality

VI. PROPOSED ALGORITHM

Edge detection & calculation of SSIM and VIF values based on different color models & edge detection operators for live video frames can be explained by the following steps:

Step 1: Constructing a video input object
Step 2: Select the source to use for video frames acquisition
Step 3: View the properties for the selected video source object.
Step 4: Preview the stream of video frames
Step 5: Acquire and display a single video frame
Step 6: Video frame is sub divided into R, G and B layers.
Step 7: A Laplacian mask of size 3 X 3 is convolved with R layer of Video Frame to detect the edges to obtain edge detected R layer.
Step 8: Edge detected R layer and G, B layers of video frame are concatenated to obtain edge detected frame.
Step 9: The SSIM(R) and VIF(R) parameters are measured. The same process is repeated to obtain the edge detected G layer and edge detected B layer along with SSIM (G & B) and VIF(G & B) parameters
Step 10:The average of the three edge detected layers SSIM(R,G&B) gives SSIM of the RGB frame. Similarly the average of the three edge detected layers VIF(R,G&B) gives VIF of the RGB frame.
Step 11:Repeat steps 6 to 10 for Sobel, Prewitt's and Robert's Operators
Step 12:Repeat steps 6 to 11 for XYZ, YCbCr and YUV color models.

VII. FLOW CHART

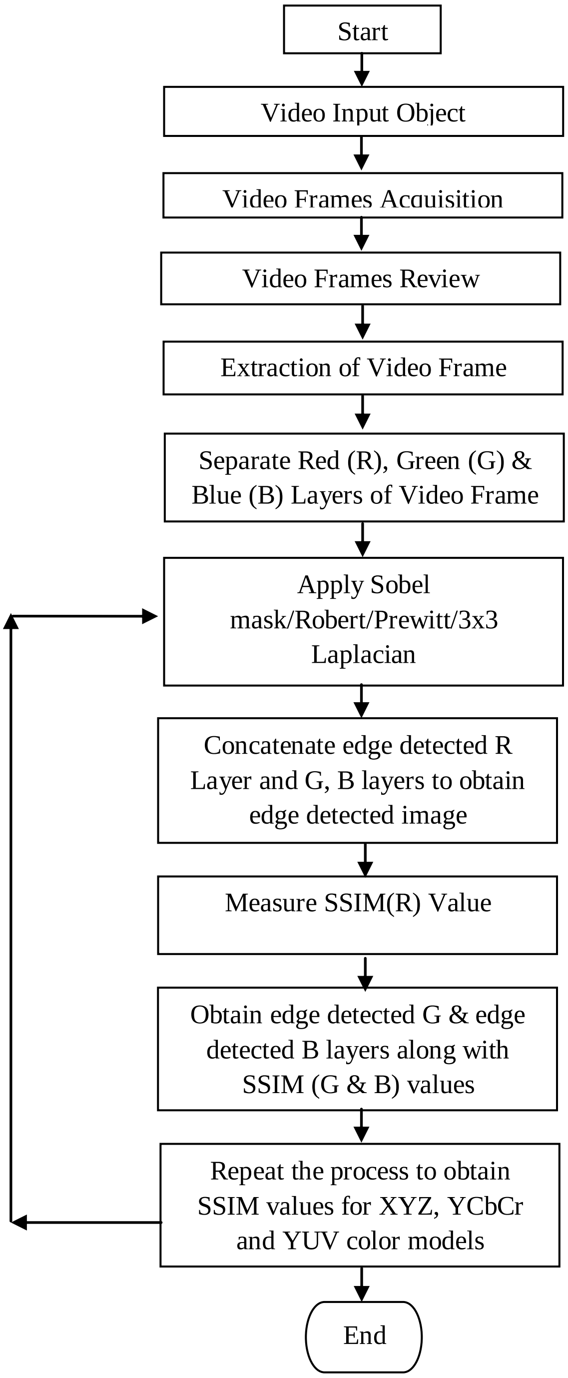


Fig. 7: Flow Chart of the proposed Algorithm

VIII. EXPERIMENTAL RESULTS

Fig.8.1: Original Video Frame

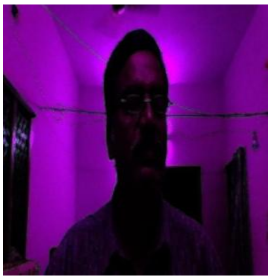


Fig.8.2: Edge Detected Y Concatenated with X & Z by Laplacian Operator

Fig.8.3: Edge Detected X Concatenated with Y & Z by Robert Operator

Fig.8.4: Edge Detected Green Concatenated with Red & Blue by Sobel Operator

Fig.8.5: SSIM value of Green with Red & Blue by Laplacian Operator is: 0.2854

Fig.8.6: SSIM value of Cb with Y & Cr by Sobel Operator is: 0.0223

Fig.8.7: SSIM value of Y With U & V by Laplacian Operator is: 0.0019

Fig.8.8: SSIM value of Cb with Y & Cr by Laplacian Operator is: 0.3117

Table: SSIM & VIF Values of 10 Live Video Frames

Video Frame No	Color Model	SOBEL		LAPLACIAN		ROBERT		PREWITT	
		SSIM	VIF	SSIM	VIF	SSIM	VIF	SSIM	VIF
FRAME 1	RGB	0.0318	0.6243	0.0954	0.6687	0.0975	0.6677	0.1073	0.6738
	XYZ	0.9945	0.6142	0.9953	0.6663	0.9926	0.3904	0.9951	0.6678
	YCbCr	0.0409	0.4557	0.1322	0.6750	0.1357	0.6743	0.1198	0.6790
	YUV	0.0010	0.2804	0.0549	0.2688	0.0457	0.2684	0.0540	0.2707
FRAME 2	RGB	0.0362	0.6240	0.1130	0.6671	0.1147	0.6662	0.1272	0.6717
	XYZ	0.9954	0.6150	0.9960	0.6662	0.9937	0.3394	0.9959	0.6678
	YCbCr	0.0414	0.4563	0.1347	0.6742	0.1382	0.6735	0.1224	0.6783
	YUV	0.0230	0.2674	0.0848	0.2565	0.0719	0.2568	0.0811	0.2593
FRAME 3	RGB	0.0365	0.6234	0.1169	0.6668	0.1187	0.6660	0.1318	0.6716
	XYZ	0.9959	0.6213	0.9964	0.6662	0.9940	0.3419	0.9962	0.6677
	YCbCr	0.0382	0.4560	0.1287	0.6745	0.1320	0.6740	0.1166	0.6786
	YUV	0.0080	0.2684	0.0579	0.2580	0.0496	0.2577	0.0579	0.2610
FRAME 4	RGB	0.0364	0.6217	0.1219	0.6679	0.1236	0.6664	0.1366	0.6719
	XYZ	0.9940	0.6121	0.9948	0.6662	0.9919	0.3515	0.9947	0.6677
	YCbCr	0.0479	0.4586	0.1280	0.6751	0.1310	0.6744	0.1175	0.6797
	YUV	0.0219	0.2715	0.0409	0.2615	0.0351	0.2606	0.0418	0.2642
FRAME 5	RGB	0.0390	0.6243	0.1051	0.6683	0.1072	0.6671	0.1187	0.6741
	XYZ	0.9953	0.6164	0.9959	0.6662	0.9934	0.3643	0.9958	0.6676
	YCbCr	0.0526	0.4566	0.1373	0.6754	0.1401	0.6745	0.1273	0.6798
	YUV	0.0047	0.2760	0.0533	0.2656	0.0454	0.2655	0.0544	0.2685
FRAME 6	RGB	0.0422	0.6244	0.1059	0.6684	0.1116	0.6673	0.1233	0.6744
	XYZ	0.9949	0.6167	0.9956	0.6662	0.9931	0.3382	0.9955	0.6679
	YCbCr	0.0503	0.4547	0.1518	0.6752	0.1550	0.6747	0.1414	0.6802
	YUV	0.0036	0.2688	0.0550	0.2578	0.0458	0.2578	0.0521	0.2610
FRAME 7	RGB	0.0406	0.6242	0.0949	0.6687	0.0971	0.6675	0.1051	0.6736
	XYZ	0.9957	0.6180	0.9963	0.6662	0.9940	0.3418	0.9962	0.6678
	YCbCr	0.0362	0.4550	0.1395	0.6748	0.1432	0.6742	0.1277	0.6793
	YUV	0.0140	0.2665	0.0625	0.2540	0.0506	0.2537	0.0585	0.2569
FRAME 8	RGB	0.0390	0.6248	0.0956	0.6687	0.0976	0.6676	0.1057	0.6739
	XYZ	0.9952	0.6178	0.9958	0.6663	0.9933	0.3515	0.9957	0.6680
	YCbCr	0.0196	0.4557	0.1345	0.6759	0.1375	0.6753	0.1259	0.6808
	YUV	0.0122	0.2670	0.0600	0.2547	0.0493	0.2543	0.0567	0.2577
FRAME 9	RGB	0.0400	0.6246	0.0955	0.6688	0.0979	0.6676	0.1075	0.6752
	XYZ	0.9941	0.6155	0.9950	0.6662	0.9924	0.3630	0.9949	0.6673
	YCbCr	0.0273	0.4562	0.1413	0.6753	0.1442	0.6745	0.1329	0.6799
	YUV	0.0125	0.2665	0.0598	0.2547	0.0490	0.2542	0.0566	0.2576
FRAME 10	RGB	0.0317	0.6261	0.0875	0.6684	0.0900	0.6677	0.0983	0.6746
	XYZ	0.9963	0.6187	0.9968	0.6661	0.9948	0.3458	0.9966	0.6687
	YCbCr	0.0446	0.4571	0.1429	0.6752	0.1455	0.6743	0.1349	0.6797
	YUV	0.0138	0.2721	0.0433	0.2619	0.0358	0.2613	0.0440	0.2647

IX. CONCLUSIONS

The edge detected video frames are obtained for RGB, XYZ, YCbCr and YUV color models by applying Sobel, Robert, Prewitt's and Laplacian operators.. The detected edges are more exact based on the XYZ color model and are detected effectively when compared with RGB, YCbCr and YUV color models. SSIM values are obtained and shown in Table. Based on the performance metrics SSIM, XYZ color model is having high similarity index which indicates low data loss during transformation. The images obtained through YUV color model are losing structural information and highly distorted when compared with the RGB, XYZ, YCbCr color models

REFERENCES

[1]. A. Koschan, "A comparative study on color edge detection" in Proceedings of the 2nd Asian Conference on Computer Vision ACCV'95, pp. 574-578, Singapore, 1995.

[2]. Image Color Conversion, https://software.intel.com/en-us/node/503870

[3]. X. W. Li and X. R. Zhang, "A perceptual color edge detection algorithm", International Conference on Computer Science and Software Engineering, vol. 1, pp.297-300, 2008.

[4]. https://www.researchgate.net/figure/284923134_fig1_Fig-1-RGB-color-cube-in-the-YCbCr- space-13

[5]. Abhishek Gudipalli and Ramashri Tirumala, "Comprehensive Infrared Image Edge Detection Algorithm," International Journal of Image Processing (IJIP), Volume (6) : Issue (5) : 2012

[6]. Gnanatheja Rakesh V, T Sreenivasulu Reddy / "YCoCg color Image Edge detection", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 2,Mar- Apr 2012, pp.152-156

[7]. R.S.Ji, Comparison of color image edge detection methods," University of Science and Technology of China,2007.

[8]. R C Gonzalez, R.E Woods. Digital Image Processing (Second Edition)[M]. New York: Prentice Hall,2003

[9]. Sonam Saluja, Aradhana Kumari Singh, Sonu Agrawal "A Study of Edge-Detection Methods" International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 1, January 2013

[10]. Wikipedia, Structural similarity. http://en.wikipedia.org/wiki/Structural-similarity.

[11]. K.R. Joy, E. Gopalakrishna Sarma, "Recent Developments in Image Quality Assessment Algorithms: A Review", Journal of Theoretical and Applied Information Technology, Vol. 65 No.1, 10 th July 2014.

[12]. Hamid R. Sheikh, "A Visual Information Fidelity approach to Video Quality Assessment", Texas Instruments Inc., USA, Alan C. Bovik, The University of Texas at Austin, USA