Jurnal
Teknologi

Full Paper

OBJECT QUALITY ENHANCEMENT OF MULTI-FRAME LOW-RESOLUTION VIDEO

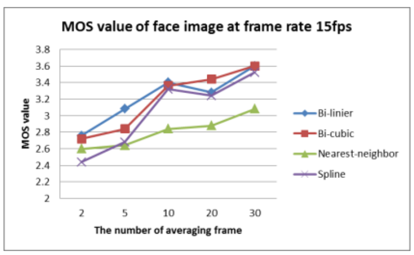
Siti Aisyah[a*], Fitri Arnia[b]

[a] Electrical Engineering Department, Politeknik Negeri Batam, Batam, Indonesia
[b] Electrical Engineering Department, Syiah Kuala University, Aceh, Indonesia

Article history
Received
27 April 2015
Received in revised form
15 June 2015
Accepted
25 November 2015

*Corresponding author
siti_aisyah@polibatam.ac.id

Graphical abstract



Abstract

A good quality image is required in various applications such as object identification and authentication. This research presents the performance of image resolution enhancement method, in which the low-resolution image originated from low-resolution CCTV video. The enhancement method is initialized by averaging video frames and continued by interpolating the resulted images using the existing interpolation techniques, namely bilinear, bi-cubic, nearest neighbor and spleen. Frame rate of 15 and 25 frames per second (fps) has been applied to the testing video. The result shows that the differences of frame rate and number of the averaged frame would affect image quality. Subjective assessment of respondents of MOS above 3 has been obtained by increasing the frame rate and the number of averaging frames.

Keywords: Digital video; frame rate; interpolation; super-resolution; MOS

1.0 INTRODUCTION

Video technology development is really helpful to record human activities at every moment in their life. For the safety reason, many industries, business departments, public areas and other places use CCTV video to monitor every event at any time. But, in many cases, most of CCTV video quality used does not meet any requirement to be an evidence to proof any criminal cases. It happened because of the low specification, such as frame rate, resolution, of that device compared to a video camera, handy-cam and any other devices. Therefore, it is relevant to introduce any data process techniques which would enhance video quality to help identifying any objects.

Digital video is a temporal-continuous image frames [1]. There are many image/object enhancement techniques applied today, from the basic to the advanced method. Generally, enhancement technique of image quality is divided into two categories namely statistic technique of image grey-level, such as image histogram manipulation in grey-level, and spatial technique or image frequency [2].

Interpolation technique is classified as an enhancement technique of image quality in spatial domain [3]. This interpolation technique has been proven to be used in various purposes, including analysis of forensic evidence [4], scientific analysis either space condition or space objects [5], analysis of low resolution video object [3], and the most popular lately is object identification in medical [6] and in geography mapping [7,8,9].

Nailul et al. focused on the use of some interpolation techniques for image quality enhancement [3]. This research extends the past researches by enhancing image quality based on the differences in the number of averaging frame and frame rate, in which the method will be implemented in tested image at different interpolation techniques.

This research has been initialized by the idea that digital video is sampled in horizontal, vertical and temporal domain and a 2-dimensional image is formed from horizontal and vertical operation of digital video in domain spatial [1].

In digitization process, digital data will be obtained by conducting a sampling and quantizing process. Digitization of analog video does not perform sampling process horizontally, but execute vertical sampling, i.e. scanning process, and temporal sampling at a certain frame rate [1].

Digital video is obtained from temporal shifting of video frame at a certain rate. This rate would affect viewer's subjectivity of video visual. If the rate is high enough, then video frame shifting will look like a moving animation. The number of frames per second (fps), i.e. frame rate, of 10fps is enough to produce a smooth-shifting video, while less than 10 fps will deliver a time-lapse video [1].

A video frame represented by an image is depicted by the number of pixels obtained from image width and height. This image width and height is usually called as image matrix and have a certain value in each vector [1].

Averaging method can be implemented in the video frame to obtain an image as an averaging result [3]. This is conducted by averaging the pixel value in the relative matrix dimension to the number of averaging frame [3]. This relation is represented in mathematical formula as follows:

$$Y[k,l] = \frac{1}{N}\sum_{i=1}^{N} Xi[k,l] \dots\dots\dots\dots\dots (1)$$

$Y[k,l]$ represents a 2-dimension matrix of output image as averaging results of N input image. $Xi[k,l]$ is a 2-dimension matrix of input image, while i represents the image/frame index and N is the number of image/frame.

To enlarge a small image, super-resolution techniques is usually applied. The existing techniques have their own benefits relative to the implementation. Some literature said that nearest-neighbor interpolation technique is the simplest among other interpolation techniques [10, 11]. Ranjeet [11] believed that the sequence of interpolation technique, seen from the high quality of the fair one, is spline, bi-linier and nearest-neighbor. This kind of interpolation techniques is applied to this research.

To assess the resulted image, there are some methods available such as PSNR, MSE, MOS, etc. This research would apply Mean Opinion Score (MOS). MOS is usually applied in voice assessment [12]. Some used this to assess video quality [13]. The subject of this method is human perception, while the object is the tested image/object. It means that people would make an assessment based on their opinion, subjectively to the tested object. Some research performed that for any type of voice or video assessment, MOS is a more suitable assessment method than other methods. This assessment method introduces a way to assess object quality based on human perception level, which can be described as very good, good, enough, fair, bad, and so on [12,13].

2.0 EXPERIMENTAL

The format of recorded video data used in this research is an audio video interleave (.avi) file format. Algorithm of multi-resolution image derived from multi frame rate video CCTV at different numbers of averaging frame is represented in figure 1.

The sequence method used to produce multi-resolution image derived from the multi frame rate CCTV video is initialized by capturing the video frame, cropping the frame, averaging the video frame and interpolating the averaging frame. The criteria of the process data are described as follows:

1. Frame rate used at CCTV video recorded are 15 and 25 frames per second. In algorithm, variable x represents frame rate.
2. The number of averaging frame is different for each object. In algorithm, variable y represents the number of averaging frame.
3. Super-resolution technique used is interpolation technique such as bi-linier, bi-cubic, nearest-neighbor and spline. Enlargement of 23 times is applied to the original image.

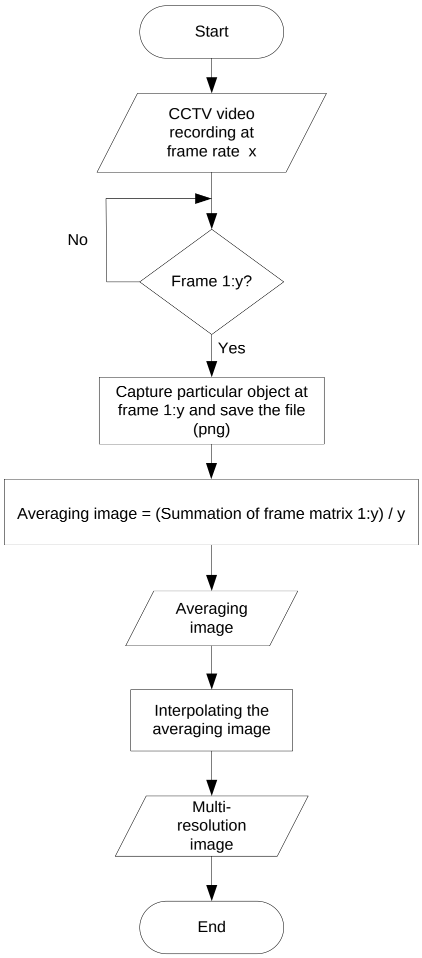


Figure 1 Algorithm of multi-resolution image derived from multi frame rate video applied to CCTV video recording

The complete process of multi-resolution image method is described in Pseudocode 1 applied for 15fps frame rate video using bi-cubic interpolation. The same algorithm is also applied for video data with different frame rate and other interpolation methods.

Pseudocode 1 Algorithm of multi-resolution image derived from multi frame rate video applied to recorded CCTV video at frame rate of 15fps for 5 frames using bi-cubic interpolation technique

```
For recording data of CCTV video at frame rate
15 fps do
For frame 1:5 do
Change RGB video to grayscale
Crop the frame at matrix dimension[,]
Save frame file in file format 'png'
End
For image file in file format 'png' 1:5 do
Image Mean = (Summation of image matrix
1:5)/5;
End
Save image mean as averaging image
Multi-resolution      image=interp2(averaging
image, 'cubic');
End
```

Assessment of multi-resolution image has been carried out by 25 (twenty five) respondents, in which the assessment scale is shown in Table I.

Table I MOS Scale

Value Scale	Quality	Image perception
5	Very good	Image is well visualized.
4	Good	Image is clearly visualized. The image line is rather smooth and clear.
3	Enough	Image is still recognizable and slightly clear. The image line is less smooth and unclear.
2	Fair	Image is almost unrecognizable. The image line is blurry.
1	Bad	Image is not recognizable.

The tested videos are focused to object 'face' and 'clock'. Both objects were used as samples to represent moving and unmoving object, respectively. Figure 2 shown face image at size of 81x111 and size of 121x121 was applied for the clock image.

(a). **(b).**

Figure 2 (a). Face image at size of 81x111; (b). Clock image at size of 121x121

Visualization of comparison was performed for both objects with the differences of frame rate, the number of averaging frames and applied interpolation technique.

3.0 RESULTS AND DISCUSSION

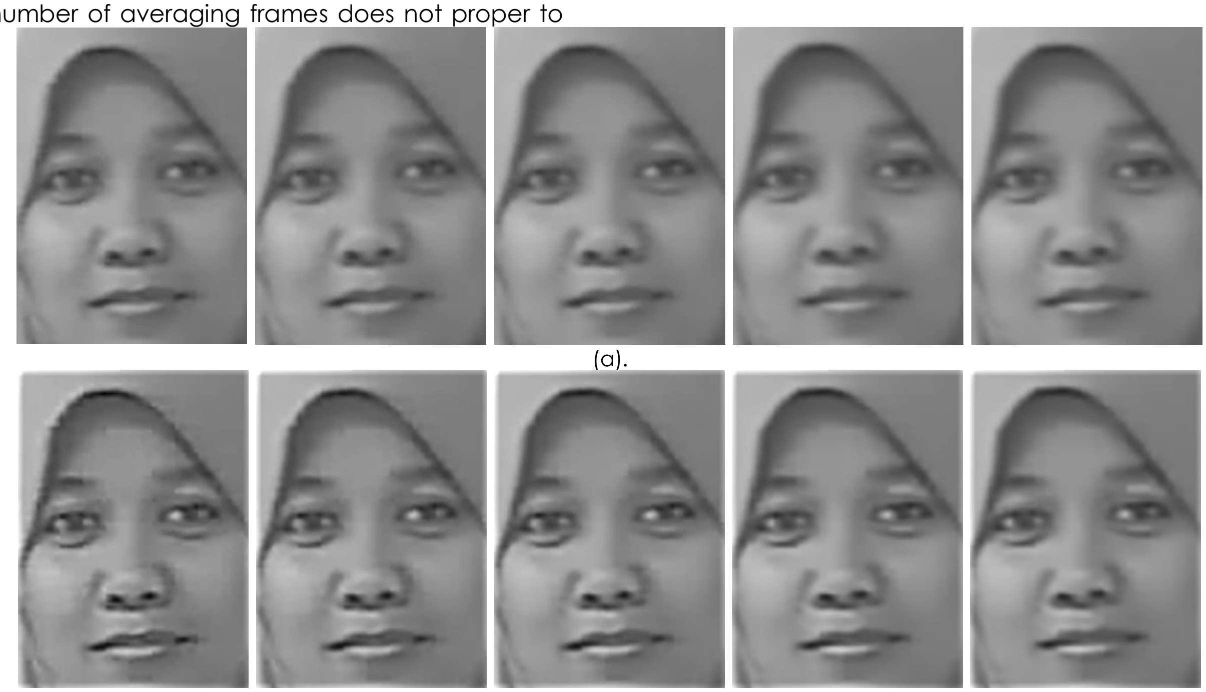
3.1 Multi-Resolution Image At Differences Number Of Averaging Frames

The first test was conducted by comparing the multi-resolution image for all four interpolation methods with different number of averaging frames namely 2, 5, 10, 20 and 30 at face image at frame rate 15fps. The face image has been enlarged for 23 times. Figure 3 shows that the multi-resolution image of face image at frame rate 15 fps which was applied with bi-cubic interpolation technique and differences number of averaging frame.

MOS assessment applied to face and clock image at some combinations of number of averaging frame and interpolation techniques has been shown in

Figure 4. From the MOS value of object 'face', it can be seen that the higher number of averaging frame, the better image would be visualized, which is represented by the MOS value higher than 3. This result is different with image clock, in which the MOS value tends to be deffered for each interpolation type. For bi-linear, bi-cubic and spline interpolation, MOS value tends to decrease as the number of averaging frame increases. In contrary, the MOS value of nearest-neighbor interpolation is getting higher for the higher number of averaging frames. Figure 4 of the clock image shows that most of higher MOS values occur in the number of averaging frame of 5. It indicates that the number of averaging frames does not proper to determine image quality of unmoving objects because it tends to result in uncertain MOS value.

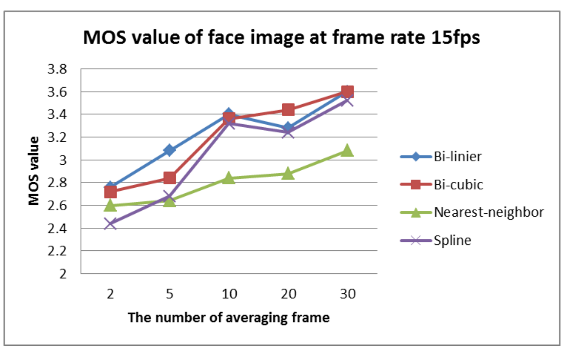
Even the MOS value is getting higher as the increasing of the number of averaging frames, it can be clearly seen in Figure 3.a that the image lost its detail in the higher number of averaging frames, while the lower number of the averaging frames would obtain an unsmooth edge (saw tooth) image. Applying smoothing techniques to the proposed method would produce a better image in the higher number of averaging frames as represented in Figure 3.b.



(a).

(b).

Figure 3 Multi-resolution image with enlargement of 23 times at frame rate 15fps with the number of averaging frame of 2; 5; 10; 20; 30 (left-to-right) (a). using bi-cubic interpolation technique; (b). using bi-cubic interpolation technique and applying the smoothing technique (Gaussian Filter)



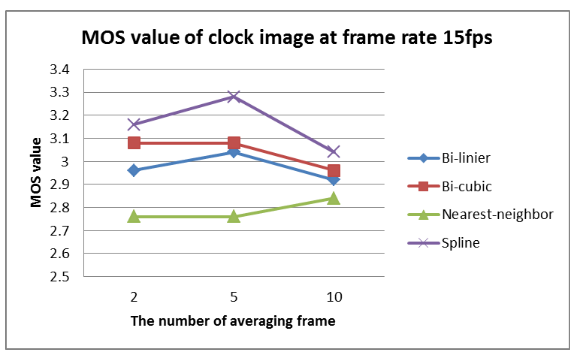


Figure 4 The result of MOS value for face image (left) and clock image (right) at frame rate 15 fps which has been applied to different number of averaging frame

3.2 Multi-Resolution Image At Different Number Of Frame Rate

The second experiment was carried out onto the two objects, i.e. face and clock, by comparing object visualization at different frame rates of15fps and 25fps at any different number of averaging frames. The MOS value of face image at different number of frame rate applied to different interpolation type is represented in Figure 5.

The MOS value of face image shown that the higher frame rate, the higher MOS value would be obtained. It means that the quality of image is getting better as the MOS value increases. Data represented in Figure 5 is also shown that the best MOS value of each interpolation sampling derived from the highest number of averaging frame and frame rate. This result can applied to most of interpolation techniques which been used. But in case of nearest-neighbor interpolation, MOS value tends to decline as the increasing of frame rate.

Figure 6 representing MOS data of the clock image shown that interval of MOS value is 2.5 to 3.5 except for the one which applying Nearest-neighbor technique only get MOS value less than 2.5. In this figure, it depicts that MOS value tends to decrease while frame rate rise up except for that which applying bi-linear technique. It has also shown that the number of averaging frame and interpolation type would not affect the MOS value. This result indicates that the method used is unsuitable for unmoving videos/objects.

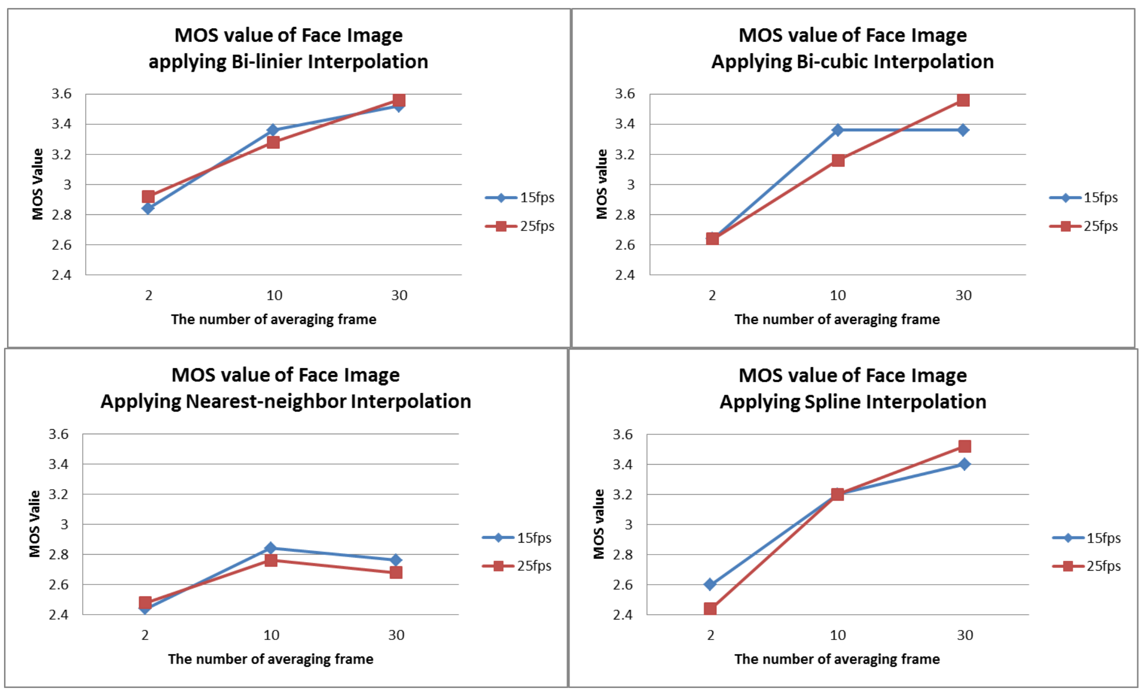


Figure 5 MOS value of face image at frame rate 15fps and 25fps with number of averaging frame of 2, 10 and 30 frames

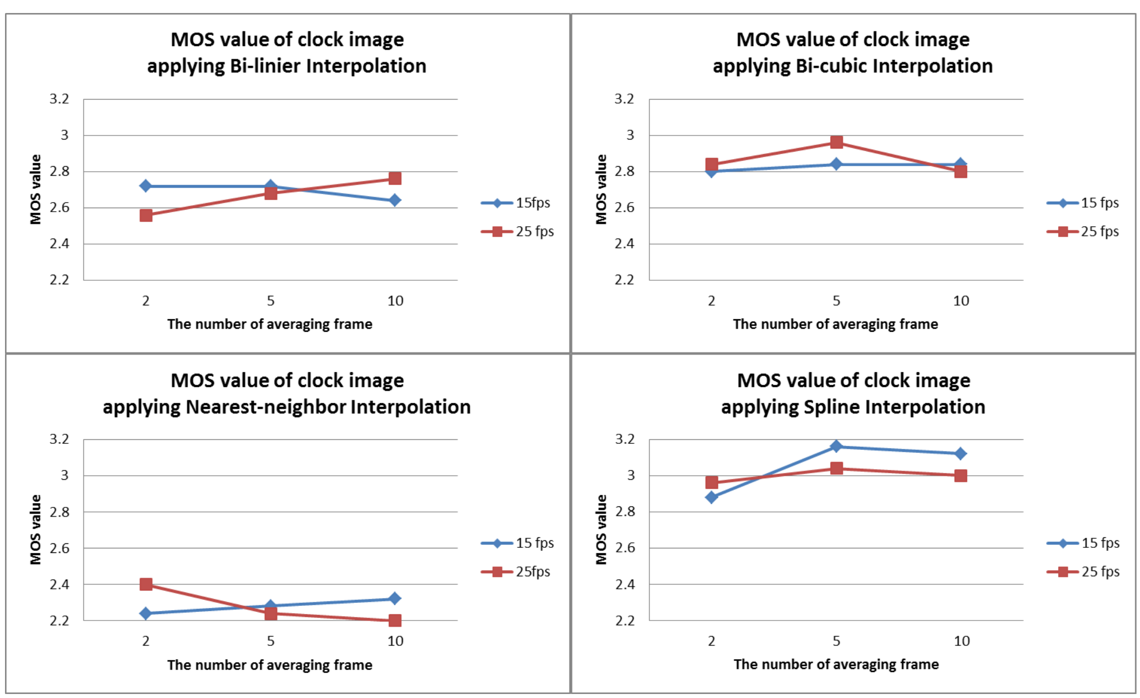


Figure 6 MOS value of the clock image at frame rate 15fps and 25fps with number of averaging frame of 2, 10 and 30 frames

3.3 Multi-Resolution Image At Different Interpolation Techniques

The third experiment was accomplished by comparing any kind of interpolation techniques applied to video with different number of averaging frame and frame rate. The MOS value of tested object applied to all interpolation type is represented in Figure 7. The MOS value of face image shows that the best interpolation type for this applied method is bi-linear and bi-cubic interpolation with MOS value higher than 3.

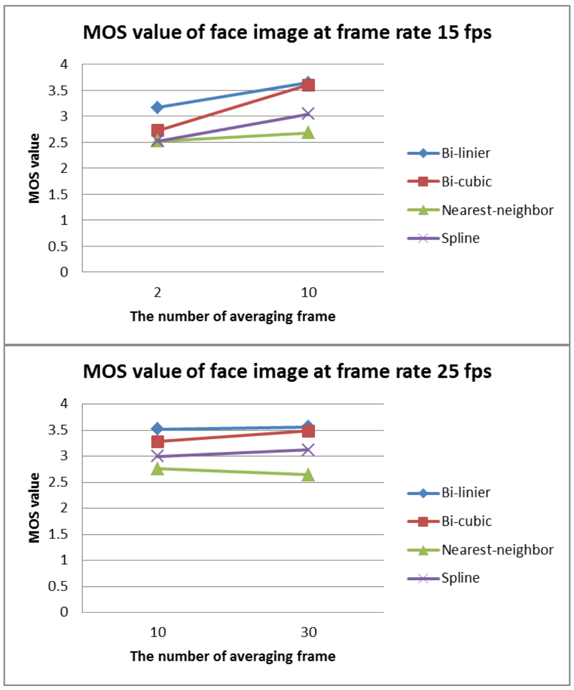


Figure 7 MOS value of face image (top) and clock image (bottom) at any interpolation techniques

4.0 CONCLUSION

From the experiments in this study, the researchers can conclude that:

1. The number of averaging frames would affect image quality. The higher number of averaging frames would impact image details, while the lower number of averaging frames would produce saw-tooth edge images. A better image would be obtained in the higher number of the averaging frame by applying smoothing techniques orGaussian filter to the proposed method.
2. Object quality of averaging image would be improved by enlarging frame rate. It has been shown by the increasing MOS value of face image at the higher frame rate that is 25fps. But, the frame rate does not affect image quality for unmoving video, object 'clock' - as the sample in this study, in which it results in uncertain MOS value.
3. The best interpolating technique applied to this method is bi-linear interpolation technique compared to other interpolation techniques. It is shown by the average MOS value equals to 3.61 for face image. While nearest-neighbor interpolation technique results in the lowest MOS value of 2.67 when applied to face image.

Acknowledgement

We are really grateful to Politeknik Negeri Batam for its financial support for this research.

References

[1] Poynton, C. 1996. *A Technical Introduction to Digital Video*. England: John Wiley & Sons Ltd.
[2] Petrou, M. and Bosdogianni, P. 1999. *Image Processing The Fundamentals*, West Sussex, England: John Wiley & Sons Ltd.
[3] Nailul Mustaqim Abdi, et al. 2011. Peningkatan Kualitas Citra Digital Menggunakan Metode Super Resolusi Pada Domain Spasial. *Jurnal Rekayasa Elektrika*. 9(3).
[4] Shen Wang, et al., 2010. A Novel and Secure Image Interpolation Methods for Image Disguise. *Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing.* IEEE Computer Society. 482–485.
[5] R.Swaminathan, and Manoj Wadhwa. 2014. Satellite Image enhancement using Combination of Transform Techniques and Interpolation Methods. *International Journal of Engineering and Computer Science*. 3: 5529-5532.
[6] George J. Grevera, and Jayaram K. Udupa. 1998. An Objective Comparison of 3-D Image Interpolation Method. *IEEE Transactions on Medical Imaging*. 17(4).
[7] Fei He1, et al. 2011. An Effective Method for Interpolation. *19th International Conference on Geoinformatics*, Shanghai. 1-6.
[8] Jing M. Chen, et. al. 2006. Locally Adjusted Cubic-Spline Capping for Reconstructing Seasonal Trajectories of a Satellite-Derived Surface Parameter. *IEEE Transactions on Geoscience and Remote Sensing*. 44(8).
[9] Paolo Villa, and Marco Gianinetto. 2006. Multispectral Transform and Spline Interpolation for Mapping Flood Damages. *IEEE International Conference on Geoscience and Remote Sensing Symposium*, Denver CO. 275-278.
[10] Philippe Thévenaz, Thierry Blu and Michael Unser. *Image Interpolation and Resampling*. Swiss Federal Institute of Technology—Lausanne.
[11] Ranjeet Roy, Maninder Pal and Tarun Gulati. 2013. Zooming Digital Images using Interpolation. *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*. 2(4).
[12] ITU-T P.800.1. 2006. *Methods For Objective And Subjective Assessment Of Quality*. Mean Opinion Score (MOS) terminology.
[13] ITU-R BT.500-11. 2002. *Methodology for The Subjective Assessment of The Quality of Television Pictures*