MOTION DETECTION IN COMPRESSED VIDEO USING MACROBLOCK CLASSIFICATION

M.Usha

Department of ECE, MKCE, Karur, Tamil Nadu

ABSTRACT

In this paper, to detect the moving objects between frames in compressed video and to obtain the best compression video and the noiseless video. We describe a video in which frames by classifying macroblocks (MB), and describe motion estimation (ME), motion vector field (MV) and motion compensation (MC). we propose to classify Macroblocks of each video frame into different classes and use this class information to describe the frame content based on the motion vector. MB class information video applications such as shot change detection, motion discontinuity detection, Outlier rejection for global motion estimation. To reduce the noise and to improve the clarity of the compressed video by using contrast limited adaptive histogram equalization (CLAHE) Algorithm.

KEYWORDS

Motion estimation, Macro block Classification, Motion vector, Motion Compensation, CLAHE.

I. INTRODUCTION

Video is a prominent multimedia data form in today‘s communication systems. Hence it‘s processing and analysis is of vital consequence. Video processing techniques such as video compression, video content analysis, compensation, extraction, etc are important in many applications. Motion based features play an important role in video signal processing since they manipulate the real-time, ―dynamic‖ parameters of video signal.

Most professional projects have an offline phase that uses compressed video and then an online, finishing phase that uses uncompressed video recaptured at full resolution. Uncompressed video requires expensive VTRs and large, high-speed hard disks.

Noise is a very important factor for image quality. Noise is a random variation of image density, visible as grain in film and pixel level variations in digital images. It arises from the effects of basic physics the photon nature of light and the thermal energy of heat inside image sensors. Typical noise reduction (NR) software reduces the visibility of noise.

The video compression is achieved by identifying high intensity moving object in compressed video frames. To estimate the motion by using macroblock classification method and based on motion vector field by using three classes. Three applications such as shot change detection, motion discontinuity detection and by using contrast limited adaptive histogram equalization (CLAHE) Algorithm to reduce the noise and to improve the clarity of the compressed video.

The rest of the paper organized as follows: video compression overview described, Macroblock(MB) classification method described, Compression method and proposes two applications such as shot change detection, motion discontinuity detection.Last Section describes CLAHE algorithm.

2. OVERVIEW

To estimate the motion by using macroblock classification method and based on motion vector field by using three classes. There are many applications such as shot change detection, motion discontinuity detection and outlier rejection using global estimation. Thus shot change detection, one of the applications was chosen and proved using macroblock classification method and by comparing threshold values and PSNR values.

By using contrast limited adaptive histogram equalization (CLAHE) Algorithm to reduce the noise and to improve the clarity of the compressed video. Thus the compressed video was further compressed, it will less in memory space and good in clarity while displaying.

3. VIDEO COMPRESSION

Film frame or video frame is one of the many still (or nearly so) images which compose the complete moving picture. Since the Video data may occupy more bandwidth than the other media data during transmission, it should be given more emphasis in the wireless multimedia communication. Once a video signal is digital, it requires a large amount of storage space and transmission bandwidth.

Sources of redundancy:

- Temporal – Adjacent frames highly correlated.
- Spatial – Nearby pixels are often correlated with each other.

To reduce the amount of data, several strategies are employed that compress the information without negatively affecting the quality of the image.FIG.1 means input video is converted to number of frames and thus frames undergoes macroblock classification and motion estimation motion vector field for the prediction and the intensity of motion pixels. Thus compensation down for compressed video frames and using of CLAHE algorithm enhanced clarity compressed video can be obtained without any error.

Basically video process having four important processes follows:

- Frame conversion
- Motion Estimation
- Motion vector field calculation
- Motion compensation

FRAMES CONVERSION: First of all, compressed video is going to convert number of frames. Because process depends on frames only

MOTION ESTIMATION: Motion estimation explores the temporal redundancy, which is inherent in video sequences, and it represents a basis for lossy video compression. Motion estimation uses the comparison of the adjacent frames. It is an important process in which comparison between frames.

MOTION VECTOR FIELD: The displacement of the reference macroblock to the target macroblock is called a motion vector MV.

MOTION COMPENSATION: It is an algorithm technique employed in the encoding of video data for video compression.

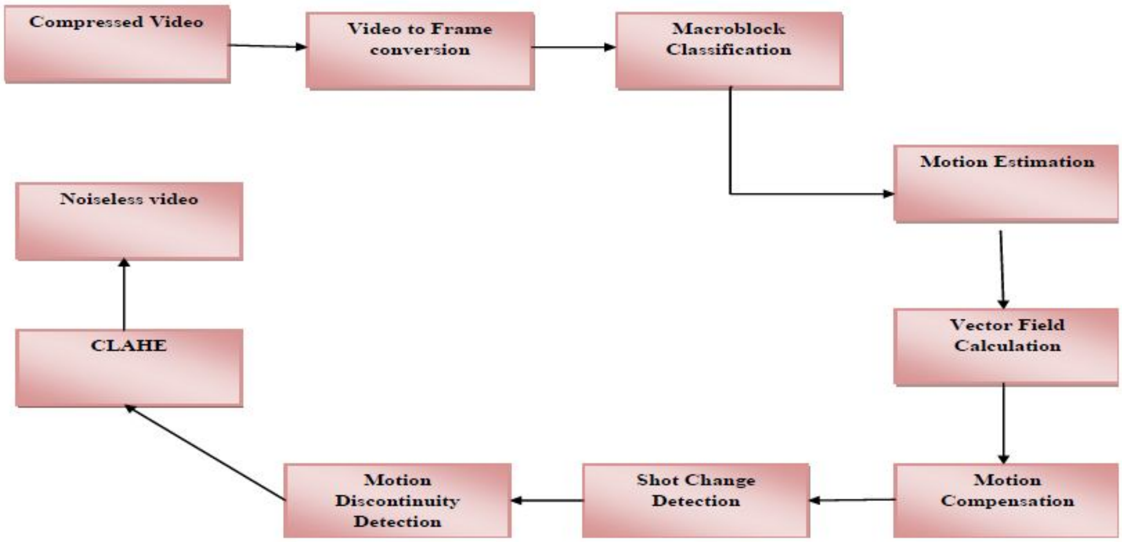

Figure 1. Block Diagram

4. MOTION ESTIMATION

Motion estimation used as the basis for powerful video analysis and video processing. Motion estimation explores the temporal redundancy, which is inherent in video sequences, and it represents a basis for lossy video compression. Motion estimation is often performed in the macroblock domain. It identifies same pixel position in the reference frame by comparing the current frame. Estimation in video compression efficiency of the system is mainly reflected in image quality, compression rate and search speed. The basic principle is the use of adjacent frames in video sequences, the temporal correlation and spatial correlation. Establish the relationship between the sequence adjacent to the inter-frame expression, thereby reducing the temporal redundancy and spatial redundancy; improve the efficiency of video coding.

In a video solution, the motion estimation computation is generally 60-80% of the total computation; the results directly affect the quality of the video image coding efficiency and recovery. Therefore, efficient motion estimation algorithm has a very important significance to improve the video data compression coding efficiency. Improve image quality, speed up the estimated speed and reduce the bit rate is the goal of motion estimation algorithm.

Motion estimation is that one block b of a current frame C is sought for in a previous frame If a block of pixels which is similar enough to block b is found in R , then instead of transmitting the whole block just a ―motion vector‖ is calculated. After finding out motion estimation and motion vector field motion compensation is achieved.

The name motion picture comes from the fact that a video, once encoded, is nothing but a sequence of still pictures that are shown at a reasonably high frequency. That gives the viewer the illusion that it is in fact a continuous animation. Each frame is shown for one small fraction of a second, more precisely 1/ k seconds, where k is the number of frames per second.

Coming back to the definition of a scene, where the frames are captured without interruption, one can expect consecutive frames to be quite similar to one another, as very little time is allowed until the next frame is to be captured.

5. MOTION VECTOR FIELD

In video compression, a motion vector is the key element in the motion estimation process. It is used to represent a macroblock in a picture based on the position of this macroblock (or a similar one) in another picture, called the reference picture. A two-dimensional vector used for inter prediction that provides an offset from the coordinates in the decoded picture to the coordinates in a reference picture.

All macroblocks in a video frame are processed in raster scan order in the space domain, so the adjacent macro blocks in the upper left, upper right and the left up can be well used as a reference macroblock. Use this algorithm to support regional prediction MV of the target macroblock D is decided by A, B, C, three macro blocks of the MV. A, B, C, macroblock MV in one to predict the MV of the target macroblock D and to get the D macro block MV predictive value using the three classes.

Full advantage of the spatial and temporal correlation of video sequences, the use of the adjacent macroblock motion vector to block movement by type starting point for prediction using different search strategies on the macro block. The results show that, in the case of the image quality is slightly improved, compared with original MVFAST algorithm, the improved algorithm can effectively improve the encoding speed.

6. MB CLASSIFICATION METHOD

Image compression component and technique based on wavelet transform used on still image and video frames. MB is usually composed of two or more blocks of pixels. Size of the block is usually a multiple of 4. Typically, pictures (frames) are segmented into macroblocks and individual prediction types can be selected on a macroblock basis rather than being the same for the entire picture.

Each image frame is divided into a fixed number of usually square blocks. For each block in the frame, a search is made in the reference frame over an area of the image that allows for the maximum translation that the coder can use. The search is for the best matching block, to give the least prediction error, usually minimizing either mean square difference, or mean absolute difference which is easier to compute. Typical block sizes are of the order of 16x16 pixels, and the maximum displacement might be +-64 pixels from a block's original position. Several search strategies are possible, usually using some kind of sampling mechanism, but the most straightforward approach is exhaustive search.

Figure 2. Macroblocks in a frame

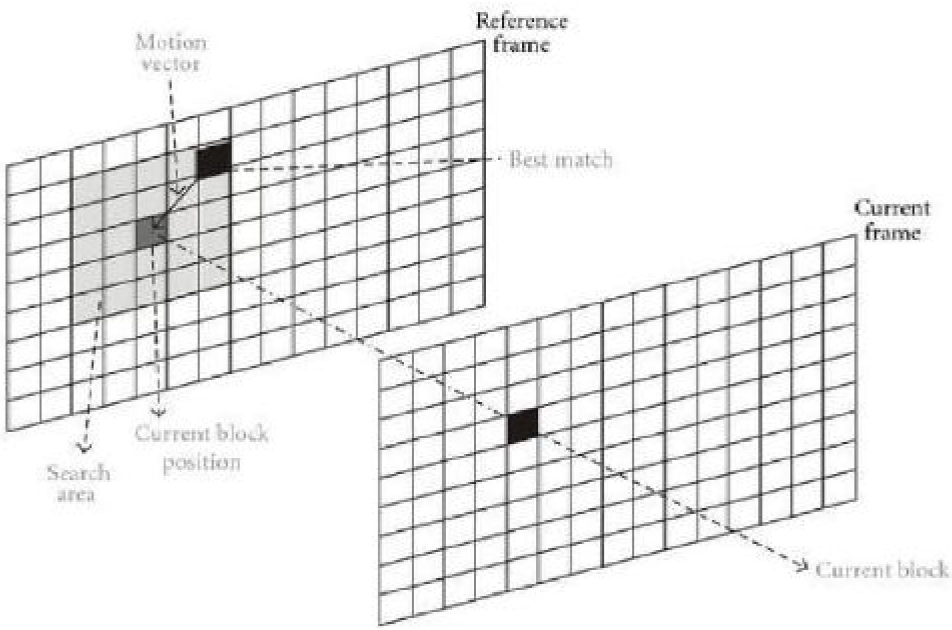

Figure 3.Comparison between frames

Motion estimation used as the basis for powerful video analysis and video processing. The macroblock motion decision was proposed to reduce the computational complexity of the motion estimation process in compressed video [3].Low-complexity encoding can be realized by limiting the amount of motion estimation performed at the encoder [4]. It identifies same pixel position in the reference frame by comparing the current frame.

For each reference frame, the motion estimation algorithm is executed for all possible modes of an MB[3]. Low-complexity encoding can be realized by limiting the amount of motion estimation performed at the encoder[4]. FIG.3 Motion estimation is that one block b of a current frame C is sought for in a previous frame R. If a block of pixels which is similar enough to block b is found in R, then instead of transmitting the whole block just a ―motion vector‖ is calculated. After finding out motion estimation and motion vector field motion compensation is achieved.

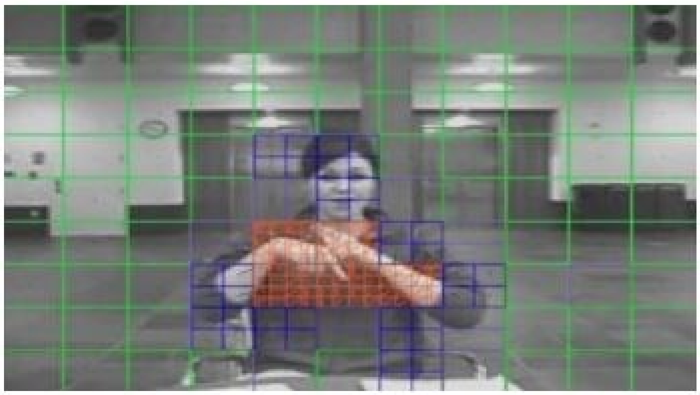

Figure 4.Macroblock Classification

Sequence statistics are used to predict macroblock type prior to coding, enabling selective computation of functions such as motion estimation (ME), motion vector field (MV) and motion compensation (MC) [2]. Each slice consists of macroblocks, which are blocks of 16x16.

However, each macroblock is also divided into sub macroblock partitions for motion-compensated prediction[1].Without loss of generality, the MB classification method can be described as,

$$
\text{Classcur\_MB}
\begin{cases}
1. & \text{if } inti\_COST < Th1 \\
2. & \text{if } inti\_COST \geq Th1 = \text{and } |PMVcur\_MB - MVpre\_final| > Th2. \\
3. & \text{if } inti\_COST \geq Th1 \text{ and } |PMVcur\_MB - MVpre\_final| \leq Th2
\end{cases}
$$

The motivations of classifying MBs according to three classes can be summarized as follows:

1)According to three classes, MBs in Class 1 have two features: (a) their MVs can be predicted accurately (i.e., is calculated based on the motion information of spatial or temporal neighboring

MBs). This means that the motion patterns of these MBs are regular (i.e., can be predicted) and smooth (i.e., coherent with the previous-frame motions).

(b) They have small matching cost values. This means that these MBs can find good matches from the previous frames. Therefore, the Class 1 information can be viewed as an indicator of the content correlation between frames

2) Class 2 includes MBs whose motion cannot be accurately predicted by their neighboring information and their previous motion information. This means that the motion patterns of these MBs are irregular and unsmooth from those of the previous frames. Therefore, the Class 2 information can be viewed as an indicator of the motion un smoothness between frames.

3) Class 3 includes MBs whose are close to the and whose matching cost values are large. Therefore, Class 3 MBs will include areas with complex textures but similar motion patterns to the previous frames.

Since is only available in the ME process, (1) is more suitable for applications where video coding and other video processing are performed at the same time, such as global motion estimation, rate control, computation control coding, as well as labeling shot changes in the process of compressing videos. However, it should be noted that (1) is only an implementation example of the proposed classification method. The idea of the proposed MB classification is general and it can be easily extended to other forms for different applications like shot change detection.

CLASS 1: differentiate the pixels in grey color
CLASS 2: differentiate the pixels in black color
CLASS 3: differentiate the pixels in white color

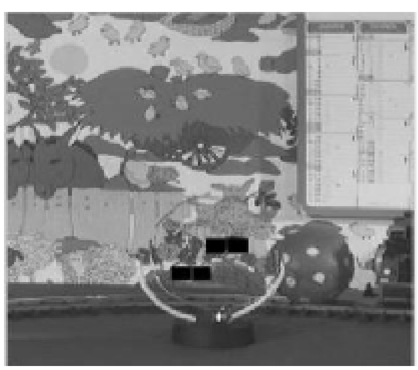

Figure 5.Class 2

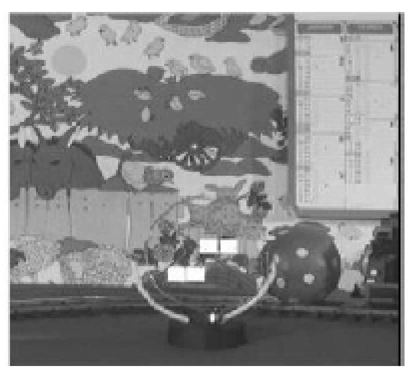

Figure 6.Class 3

6. MOTION COMPENSATION

It is an algorithm technique employed in the encoding of video data for video compression. When images can be accurately synthesized from previously transmitted/stored images, the compression efficiency can be improved. The blocks are not transformed in any way apart from being shifted to the position of the predicted block. This shift is represented by a motion vector. By using the motion estimation and motion vector field , motion compensation can be achieved.

The choice of block-size to use for motion compensation is always a compromise, smaller and more numerous blocks can better represent complex motion than fewer large ones. This reduces the work and transmission costs of subsequent correction stages but with greater cost for the motion information itself. And they conclude that the choice of block-size can be affected not only by motion vector accuracy but also by other scene characteristics such as texture and inter-frame noise.

7. RESULTS

Figure 7. Original Video

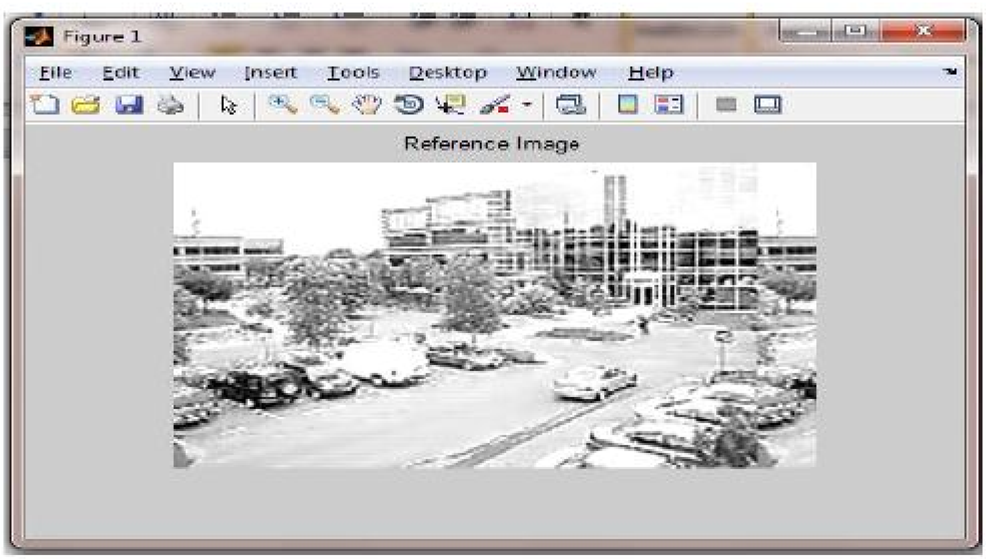

Figure7.Reference Frame

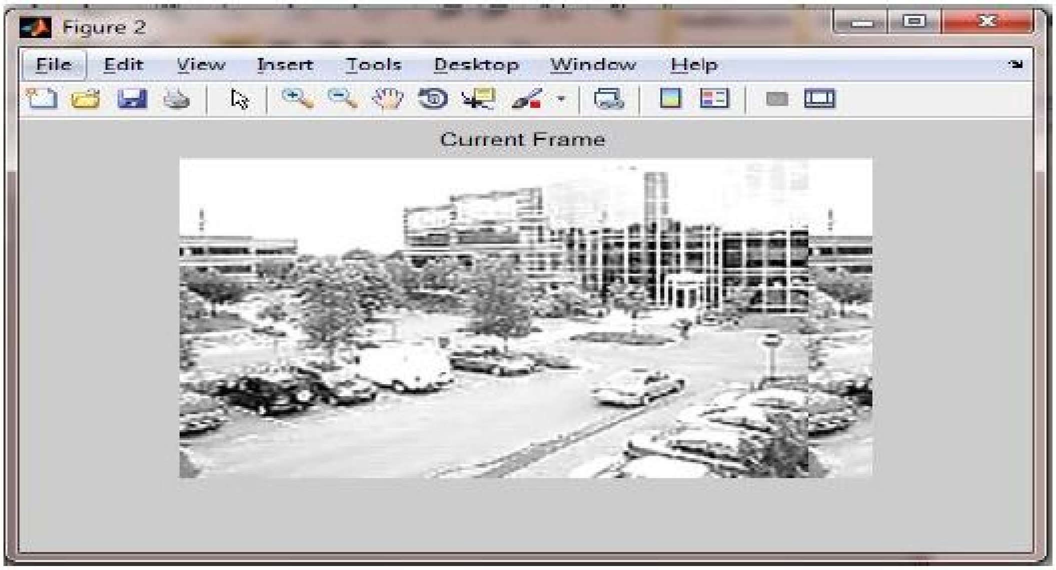

Figure 8.Current Frame

Figure 9.Motion Compensation

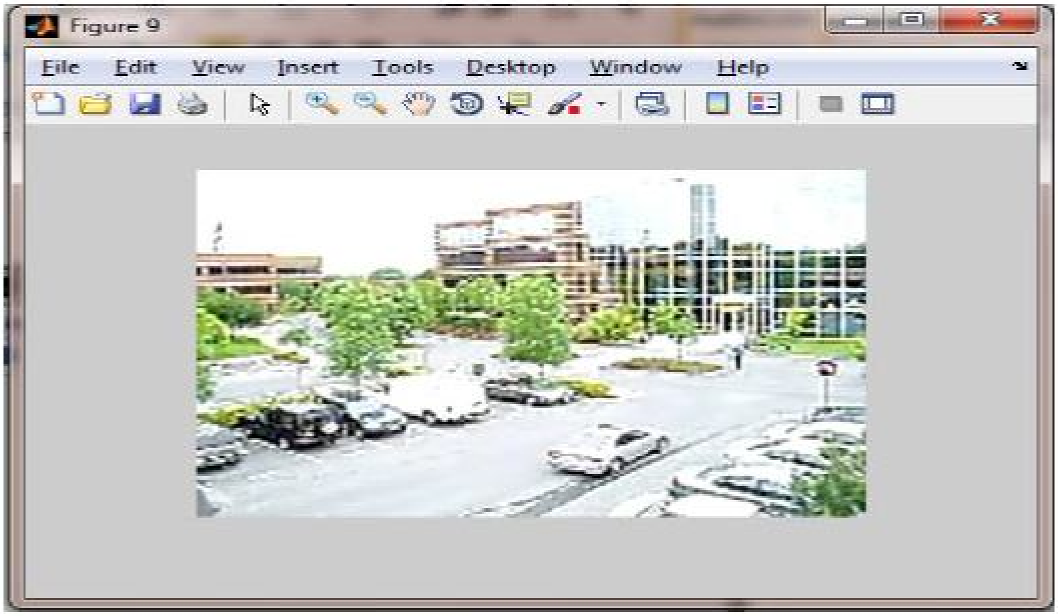

Figure 10.CLAHE Output

8. APPLICATIONS

8.1. SHOT CHANGE DETECTION

A crucial step in multimedia processing is that of reliable video segmentation into visually coherent video shots (i.e., scene change detection). FIG. 9 Frames captured by one camera action (i.e., a continuous operation of one camera), and a „shot change" as the boundary of two shots. Video shots,as tracking of selected video information [5].

Therefore, we can use the information of Class 1 as the primary feature to detect shot changes. Since the motion pattern will also change at shot changes, the information of Class 2 and Class 3 can be used as additional features for shot change detection. However, they can be easily extended for use with different block-based transform video compression methods[5]. FIG.10(a)&(b) transitions from one shot to another, either abrupt or gradual, may take place. An abrupt transition, or hard cut, occurs between twoconsecutive frames and is the most common type.[6]

There are two types of shot changes .They are:

- Hard cut(abrupt change)
- Fades(gradual change)

8.1.1 Hard cut(Abrupt change):

A cut is an instantaneous transition from one scene to the next. There are no transitional frames between two shots.

Figure 11. Example for Hard Cut

8.1.2 Fades (gradual change):

A fade is a gradual transition between a scene and a constant image (fade-out) or between a constant image and a scene (fade-in).

Figure 12. Example for Fades Out

Since, shot changes (including abrupt, gradual, fade-in or fade-out) always happen between two uncorrelated video shots, the content correlation between frames at shot changes will be low. Therefore we can use the information of class 1 as the primary feature to detect shot changes.
Furthermore , since the motion pattern will also changes, the information of class 2 and class 3 can be used as additional shot change detection. Thus according to the three classes the class based algorithm (CB-Shot) for shot change detection is proposed as below equation (a)

$$
Fgshot(t)=\begin{cases} 1 & \text{if } Nclass\_1(t) \le T1 \text{ and } NIntra\_MB(t)\text{-}NIR(t) \ge T4 \\ & \text{or if } \begin{cases} Nclass\_1(t) \le T2 \text{ and } NIntra\_MB(t)\text{-}NIR(t) \ge T4 \text{ and} \\ |Nclass\_2(t)\text{-}Nclass\_2(t\text{-}1)| + |Nclass\_3(t)\text{-}Nclass\_3(t\text{-}1)| \ge T3 \end{cases} \\ 0 & \text{else} \end{cases}
$$

From the equation(a) the class based algorithm ,where 't' is the frame number and Fgshot(t) is a flag indicating wheather a shot change happens at the current frame t or not. Fgshot(t) will equal to 1 if there is a shot change and will equal to 0 else.

- CLASS 1 → Indicate low content correlation
- CLASS 2 }
- CLASS 3 } → Detect rapid motion changes in frame.

NIntra_MB(t) is the number of intra coded macroblocks at the frame 't' . NIR(t) is the number of intra-refresh macroblocks in the current frame.

Nclass_1(t) , Nclass_2(t) and Nclass_3(t) are the total number of class 1 ,class 2and class 3 MBs in the current frame t respectively.

$$
\begin{cases} Fgshot(t) = 1 & \text{shot change occurs} \\ Fgshot(t) = 0 & \text{no shot change} \end{cases}
$$

T1,T2,T3 and T4 are the thresholds for deciding the shot change. T1-T4 are calculated by equation (b)

$$
\begin{cases} T1=(NMB\ (t)) - (NIR(t))/40 \\ T2=(NMB\ (t)) - (NIR(t))/30 \\ T3=(NMB\ (t)) - (\ NIR(t))/4, T4=T1 \end{cases}
$$

Where NMB (t) is the total number of MBs of all classes in the current frame. It should be noted that in Fgshot(t) equation the Class 1 information is the main feature for detecting shot changes (i.e. Nclass_1(t), Nclass_1(t) ≤ T2 in equation(a)). The intuitive of using the Class 1 information as the major feature is that it is a good indicator of the content correlation between frames. The Class 2 and Class 3 information is used to help detect frames at the beginning of some gradual shot changes where a large change in motion pattern has been detected but the number of Class 1 MBs has not yet decreased to a small number.

The intra-coded MB infor-mation can help discard the possible false alarm shot changes due to the MB misclassification. From, (a) and (b), we can also see that when intra-refresh functionality is enabled (i.e., when NIR(t) > 0), our algorithm can be extended by simply ex-cluding these intra-refreshed MBs and only performing shot change detection based on the remaining MBs.

Furthermore, note that (a) is only one implementation of using our class information for shot change detection. We can easily extend (a) by using more sophisticated methods suchas cross-validation to decide the threshold values in an automatic way. Besides, other machine learning models can also be used to decide the shot detection rules and to take the place of the manually-set rules in (a)

8.2. MOTION DISCONTINUITY DETECTION

Automatic detection of independently moving targets in a potentially large collection of surveillance video data obtained from an unmanned sensor(such as video camera in an unmanned aerial vehicle).Consequently this work is about independent motion estimation[7]. Since our class information, especially Class 2 information, can efficiently reflect the irregular motion patterns, it can be easily used for motion discontinuity detection.

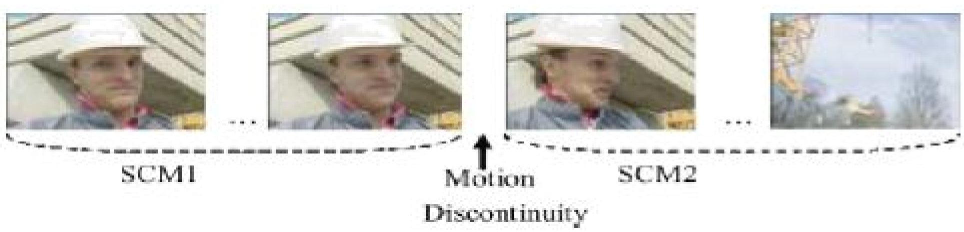

Figure 13. Example for Motion Discontinuity

$$FgMD(t)=\begin{cases} 1 & \text{if } Nclass\_1(t) \geq th1\text{^}MD \text{ and } \\ & \Sigma \;\; I(Nclass\_2(t\text{-}i) \geq \\ & th3\text{^}MD)=k+1 \\ 0 & \text{else} \end{cases}$$

Basically, motion discontinuity can be viewed as motion unsmoothness or the change of motion patterns. Where I(f) is an indicator I will equal to 1 if f is true, and 0 if f is false. FIG.11 means that an MD will be detected only if the number of Class 2 MBs is larger than a threshold for k+1 consecutive frames. This is based on the assumption that an obvious camera motion change will affect several frames rather than one.

9. CLAHE ALGORITHM

Contrast limiting procedure has to be applied for each neighbourhood from which a transformation function is derived. Contrast limited adaptive histogram equalization (CLAHE) was developed to prevent the over amplification of noise. FIG.12 A „Claire‟ video sequence was used to study the effect of the Quality Scaling Factor (QSF) of a standard Codec on the compression ratio of the proposed encoder.

QSF is the value of the quantizing parameter of a standard encoder .[8] CLAHE was developed to prevent the over amplification of noise that adaptive histogram equalization a block in a standard encoder are divided by this QSF for achieving higher compression. This higher the value of QSF, higher will be the compression achieved. The quality however goes down (because increased quantization error). Thus QSF provides a mean for trade off between quality and compression. [8]

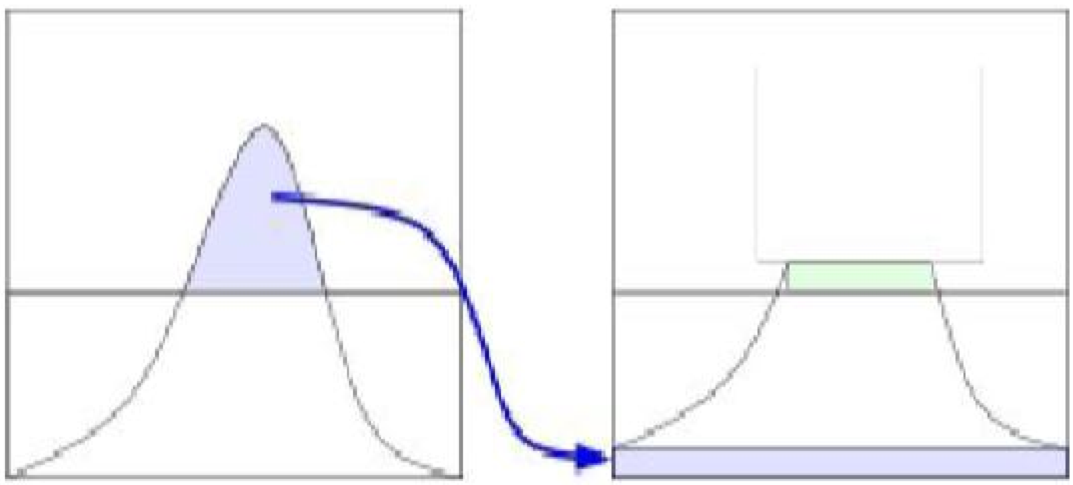

Figure 14. CLAHE Algorithm

Use the displayed image to estimate the intensity transfer that is then applied to all channels individually. This is the desired mode for getting local contrast compression for color photographs with a higher bit-depth and higher dynamic range than appropriate for a digital display.

Figure 15.Original image &CLAHE processed

10. CONCLUSION

In this paper, a new Macro Block class information is proposed for various video processing applications. We have to classify Macro blocks of each frame into different classes and use this class information to describe the frame content. Thus by using class information many algorithms can be achieved for many video applications. The CLAHE algorithm is the promising candidate for the noise reduction and enhancement of video quality of compressed video. However, CLAHE is still a young–born research field and several issues need to be addressed to fully understand its potential and limitations in practical.

REFERENCES

[1] Weiyao Lin, Ming-Ting Sun, Fellow, IEEE, Hongxiang Li, Senior Member, IEEE, Zhenzhong Chen, ‗‘Macroblock Classification Method for Video Applications Involving Motions‘‘ Ieee Transactions on Broadcasting, Vol. 58, no. 1, march 2012

[2] W. Lin, M.-T. Sun, H. Li, and H. Hu, ―A new shot change detection method using information from motion estimation,‖ in Pac.-Rim Conf. Multimedia, 2011, vol 15.

[3] Ranjana, Kedar, M.R.Ingle,P.M.Jadhav ―Analysis of Fast Computing Algorithms Using Macro block Motion Decision in H.264‗‗‗‗ 1st International Conference on Recent Trends in Engineering & Technology, Vol 42 Mar-2010 .

[4] Yafan Zhao and Iain Richardson ―Macroblock Classification for Video Encoder Complexity Management Image Communication Technology Group, The Robert Gordon University, Schoolhill, Aberdeen, UK.,Vol 12 April 2009.

[5] Ranjana, Kedar, M.R.Ingle,P.M.Jadhav ―Analysis of Fast Computing Algorithms Using Macro block Motion Decision in H.264‗‗‗‗ 1st International Conference on Recent Trends in Engineering & Technology, Mar-2008

[6] D. Swanberg, C. F. Shu, and R. Jain, ―Knowledge guided parsing in video database,‖ in Proc. SPIE—Storage and Retrieval for Image and Video Database, San Jose, CA, 1993, vol. 1908, pp. 13–24 , 2008

[7] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, ―Overview of the H.264/AVC video coding standard,‖ IEEE Trans. Circuit Syst. , Member,IEEE Video Technol., vol. 13, pp. 560–576, 2005.

[8] Dan Lelescu and Dan Schonfeld "Statistical Sequential Analysis for Real-Time Video Scene Change Detection On Compressed Multimedia Bitstream'' Signal and Image Research Laboratory , University of Illinois , Chicago,Vol 19, 2005.

[9] S. Shu and L. P. Chau, ―A new scene change feature for video transcoding,‖ IEEE Int. Symp. Circuits Syst., vol. 5, pp. 4582–4585, 2004.

[10] K. Yoon, D. DeMenthon, and D. Doermann, ―Event detection from MPEG video in the compressed domain,‖ in Int. Conf. Pattern Recognition., 2002, vol 21, pp. 819–822.

Author

Usha.M received the B.E. degree from Anna University Trichy, Tamil Nadu, in 2011, the M.E. degree from Anna University Chennai, Tamil Nadu, in 2013,Since 2013, she has been an Assistant Professor with Department of Electronic and Communication Engineering. Her research interests include video processing, video coding & compression.