Journal of Information Hiding and Multimedia Signal Processing ©2012 ISSN 2073-4212
Ubiquitous International Volume **3**, Number **1**, January **2012**

A Survey of Video Enhancement Techniques

Yunbo Rao, Leiting Chen

School of Computer Science and Engineering,
University of Electronic Science and Technology of China, Chengdu, China, 611731
uestc2008@126.com

Received August 2011; revised December 2011

Abstract. *Video enhancement is one of the most important and difficult components in video research. The aim of video enhancement is to improve the visual appearance of the video, or to provide a "better"transform representation for future automated video processing, such as analysis, detection, segmentation, recognition, surveillance, traffic, criminal justice systems. In this paper, we present an overview of video enhancement processing and analysis algorithms used in these applications. The existing techniques of video enhancement can be classified into two categories: Self-enhancement and Context-based fusion enhancement. More specifically, we categorize processing methods based representative techniques of video enhancement. Thus, the contribution of the paper has fourfold: (1) to classify and review video enhancement processing algorithms, (2) to discuss the advantages and disadvantages of these algorithms, (3) according to this integrated consideration, attempt an evaluation of shortcomings and general needs in this field of active research, and (4) we will point out promising directions on research for video enhancement for future research.*
Keywords: Video enhancement, self-enhancement, frame-based fusion enhancement, spatial-based domain enhancement, transform-based domain enhancement

---

1. **Introduction.** Video enhancement problem can be formulated as follows: given an input low quality video and the output high quality video for specific applications. How can we make video more clearer or subjectively better?

Digital video has become an integral part of everyday life. It is well-known that video enhancement as an active topic in computer vision has received much attention in recent years. The aim is to improve the visual appearance of the video, or to provide a "better"transform representation for future automated video processing, such as analysis, detection, segmentation, and recognition [1-5]. Moreover, it helps analyses background information that is essential to understand object behavior without requiring expensive human visual inspection [6]. There are numerous applications where digital video is acquired, processed and used, such as surveillance, general identity verification, traffic, criminal justice systems, civilian or military video processing et al.

Carrying out video enhancement understanding under low quality video is a challenging problem because of the following reasons [2-3, 6-9]. (i)Due to low contrast, we cannot clearly extract moving objects from the dark background. Most color-based methods will fail on this matter if the color of the moving objects and that of the background are similar. (ii) The signal to noise ratio is usually very low due to high ISO (ISO is the number indicating camera sensors sensitivity to light). Using a high ISO number can produce visible noise in digital photos. Low ISO number means less sensitivity to light. (iii)The information carrying video signal is a degraded version of a source or original video signal which represents the three dimensional continuous world. These degradations can be a result of the acquisition process, or the rate and format conversion processes. (iv)Environmental information affects the way people perceive and understand what has happened. Hence, dealing with moving tree, fog, rain, behavior of people in nighttime video are the difficult because they lack background context due to poor illumination. (v)Inter-frame coherence must also be maintained i.e. the moving objects region as weights in successive images should change smoothly. (vi)One pixel from a low quality image may be important even if the local variance is small, such as the area between the headlights and the taillights of a moving car. (vii)The poor quality of the used video device and lack of expertise of the operator.

There are two main methods to process an image as defined by the domain in which the image is processed, namely spatial-based domain and frequency-based domain. Spatial-based domain refers to the image plane itself, and approaches in this category are based on direct manipulation of pixels in an image. Frequency-based domain processing techniques are based on modifying the spatial frequency spectrum of the image as obtained by transform. Enhanced techniques based on various combinations of methods from these two categories are not unusual and the same enhancement technique can also be implemented in both domains, yielding identical results. With the same image processing, a lot of video enhancement methods have been proposed. However, in all of these methods, there still are no general standards, which could be used as a design criterion of video enhancement algorithms. There is also no general unifying theory of video enhancement. The survey of available techniques is based on the existing techniques of video enhancement, which can be classified into two broad categories: spatial-based domain video enhancement and transform-based domain video enhancement [1,9-14]. Spatial-based domain video enhancement operates directly on pixels. The main advantage of spatial-based domain technique is that they are conceptually simple to understand, and the time complexity of these techniques is low which favors real time implementations. But these techniques generally lacks in providing adequate robustness and imperceptibility requirements. A survey of spatial-based domain enhancement techniques can be found in [4,15-18]. Transform-based domain video enhancement is a term used to describe the analysis of mathematical functions or signals with respect to frequency, and operate directly on the transform coefficients of the image, such as Fourier transform, discrete wavelet transform(DWT), and discrete cosine transform(DCT) [1,13-14,18-20]. The basic idea in using this technique is to enhance the video by manipulating the transform coefficients. The advantage of transform-based video enhancement include (i) Low complexity of computations, (ii) Ease of viewing and manipulating the frequency composition of the image, and (iii) the easy applicability of special transformed domain properties. The basic limitations including (i) it cannot simultaneously enhance all parts of the image very well, and (ii) it is difficult to automate the image enhancement procedure.

In this paper, according to if enhanced video embed high quality background information, the existing techniques of video enhancement can be classified into two broad categories: Self-enhancement and frame-based fusion enhancement. Traditional methods of video enhancement are to enhance the low quality video itself. It doesn't embed any high quality background information. Such as contrast enhancement method, HDR-based video enhancement, compressed-based video enhancement, and wavelet-based transform video enhancement. These approaches are uniformly called self-enhancement of low quality video. It don't enough luminous of low quality video. The reason is that in the dark video, some areas are so dark that all the information is already lost in those regions. No matter how much illumination enhancement you apply, it will not be able to bring back lost information. Frame-based fusion enhancement refers to low quality video, which fuse illumination information in different time video. The approach is that it is by extracting high quality background information to embed low quality video. How would one combine information from two (or more) background images in a meaningful way? How would one pick high-quality background parts while keeping all the low-quality important information? To these problems, the previous researchers have abundant research. Fig.1 shows the more detail categories of video enhancement.

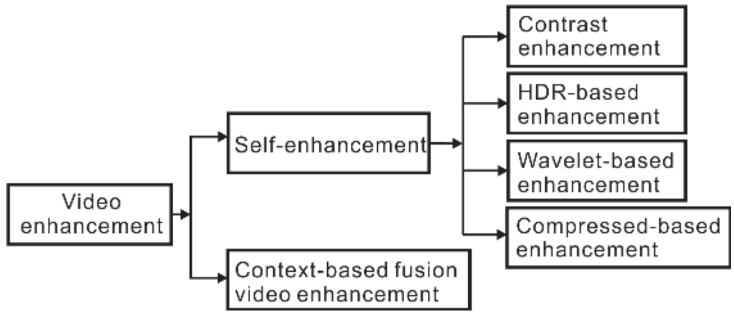


FIGURE 1. the block diagram of video enhancement categories

In this paper, we focus on video enhancement considering both areas of self-enhancement and frame-based fusion enhancement. Research in the field started as early as in the 70s with the advent of computers and the development of efficient video processing techniques. We also discuss related image enhancement techniques, since most video enhancement techniques are based on frame enhancement. We don't aim at covering the whole field of video enhancement and its applications. It is a broad subject that is still evolving. E.g. we don't discuss contributions, which are made by ITU and ISO standard in this area.

The remainder of the paper is organized as follows. Sections 2 review self-enhancement of low quality video, Sections 3 survey context-based fusion video enhancement, Sections 4 give some discussions and the proposed future directions, and Section 5 conclude the paper.

2. **Self-enhancement of low quality video.**

In this section, we focus on self-enhancement of low quality video. The approaches can be classified into four categories: contrast enhancement, HDR-based video enhancement, compressed-based video enhancement, and wavelet-based video enhancement. An overview of some of the well-known methods in these categories is given below.

2.1. **Contrast enhancement.**

Video enhancement techniques involve processing an image/frame to make it look better to human viewers. It is usually used for post processing by modifying contrast or dynamic range or both in an image. The aim of contrast enhancement process is to adjust the local contrast in different regions of the image so that the details in dark or bright regions are brought out and revealed to the human viewers. Contract enhancement is usually applied to input images to obtain a superior visual representation of the image by transforming original pixel values using a transform function of the form.

$$g(x, y) = T[r(x, y)] \tag{1}$$

where $g(x, y)$ and $r(x, y)$ are the output and input pixel values at image position. Usually for correct enhancement it is desirable to impose certain restrictions on the transformation function T [20].

The existing techniques of contrast enhancement techniques can be broadly categorized into two groups: direct methods [22,23] and indirect methods [10, 24-28]. Direct methods define a contrast measure and try to improve it. Indirect methods, on the other hand, improve the contrast through exploiting the under-utilized regions of the dynamic range without defining a specific contrast term. In fact, there are other type of algorithm for contrast enhancement, such as gamma enhancement, power-low rule, logarithmic approach, automatic gain/offset, and transform enhancement. In this paper, contrast enhancement techniques can be broadly categorized into two groups: histogram equalization(HE), tone mapping.

A) Histogram equalization

Histogram equalization is one of the most commonly used methods for contrast enhancement. It attempts to alter the spatial histogram of an image to closely match a uniform distribution. The main objective of this method is to achieve a uniform distributed histogram by using the cumulative density function of the input image[24-28]. The advantages of the HE include (i) it suffers from the problem of being poorly suited for retaining local detail due to its global treatment of the image. (ii) small-scale details that are often associated with the small bins of the histogram are eliminated. The disadvantage is that it is not a suitable property in some applications such as consumer electronic products, where brightness preservation is necessary to avoid annoying artifacts. The equalization result is usually an undesired loss of visual data, of quality, and of intensity scale. Fig.2 shows the experimental result of histogram equalization [21].

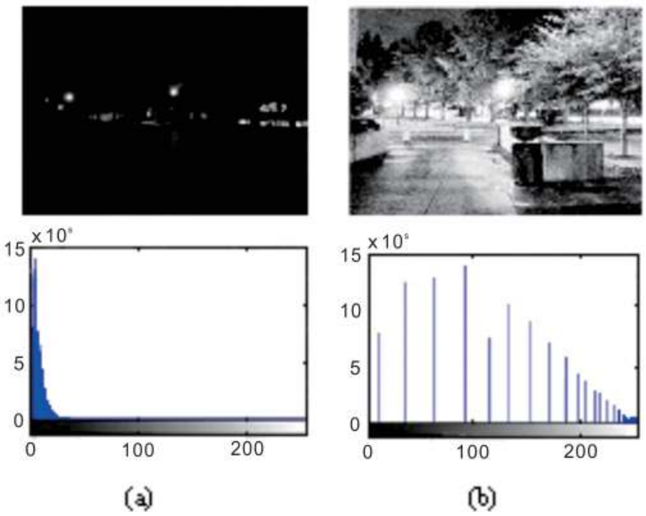


Figure 2. Histogram equalization. (a) original frame of low quality and histogram image, and (b) the result of histogram equalization and histogram image.

HE methods are divided into two major categories: global and local methods. Global histogram equalization (GHE) attempts to alter the spatial histogram of an image to closely match a uniform distribution [21]. In this approach, the contrast stretching is limited in gray levels with high frequencies. This causes significant contrast loss for gray levels having lower frequencies. To overcome this problem, different local histogram equalization (LHE) methods have been proposed. Typical histogram specification, histogram equalization, and gamma correction to improve global contrast appearance only stretch the global distribution of the intensity. More adaptive criterions are needed to overcome such drawback. [26] uses two adaptive histogram equalization techniques to modify intensity's distribution inside small regions. To unsharp masking for contrast enhancement of images/videos,[27]employes an adaptive filter that control the contribution of the sharpening path in such a way that contrast enhancement occurs in high detail areas and little or no image sharpening occurs in smooth areas. Some methods to transform domain by way of a transform coefficient histogram have been fully explored. In another local method which is called shape preserving histogram modification, instead of a rectangular block, connected components and level-sets are used for contrast enhancement. Partially overlapped sub block HE is another local method in [28]. Some local methods that don't use histogram have been proposed in literature also. Statistic-based properties of the image method determine a transformation function for each pixel by considering the local minimum/maximum and local average in a window centered at that pixel [29]. Another local method is based on using 2D teager-kaiser energy operator to compute the value of local contrast of each pixel [30]. The computed value is transformed by a predefined function to emphasize the pixel's contrast. Then, a reverse process is performed to obtain the new value of the pixel according to the new value of the contrast. The local histogram equalization and adaptive histogram equalization can provide better results but are computationally intensive. Recently, a novel and effective video enhancement algorithm for low lighting video is proposed. The algorithm works by first inverting the input low-lighting video and then applying an image de-haze algorithm on the inverted input. To facilitate faster computation and improve temporal consistency, correlations between temporally neighboring frames are utilized [23].

For preserving the input brightness of the image, which is required to avoid the generation of non-existing artifacts in the output image, different methods based on histogram equalization have been proposed. Mean preserving bi-histogram equalization (BBHE), equal area dualistic sub-image histogram equalization (DSIHE), minimum mean brightness error bi-histogram equalization (MMBEBHE), recursive mean-spread histogram equalization (RMSHE), and multi-HE are HE based methods which tend to preserve the image brightness with a significant contrast enhancement [31]. In BBHE, histogram of the input image is separated into two parts according to the mean of gray levels and each part is equalized independently. DSIHE is similar to BBHE except that it separates the histogram at the median of gray levels instead of the mean. MMBEBHE is an extension of BBHE and provides maximal brightness preservation. In RMSHE, scalable brightness preservation is achieved by partitioning the histogram recursively more than once. Multi-HE consists of decomposing the input image into several sub-images, and then applying the classical HE process to each one. This methodology performs a less intensive image contrast enhancement [32]. This technique is a generation of BBHE. Although these methods preserve the input image brightness on output, they may fail to produce images with natural looks [31]. In order to overcome this drawback, two multi histogram equalization methods, i.e. Minimum within-class level squared error MHE (MMLSEMHE), have been proposed. In these methods, number of sub-images is determined by a cost function. They usually perform a less intensive image contrast enhancement [31]. This is the cost that is paid for achieving contrast enhancement, brightness preservation and natural looking images at the same time. [33]uses the histogram of each frame, along with upper and lower bounds computed per shot in order to enhance the current frame. This ensures that the artifact introduced during the enhancement is reduced to a minimum. Traditional methods don't compute per-shot estimates tend to over-enhance parts of the video such as fades and transitions.

Histogram specification technique is another approach for contrast enhancement [21]. In this method, the shape of the histogram is specified manually, and then a transformation function is constructed based on this histogram to transform input image at gray levels. Dynamic histogram equalization (DHS) method tends to preserve the details of the input image[25]. Image histogram is partitioned based on local minima and specific gray level ranges that are assigned to each partition. After partitioning, HE is applied on each partition. Another modified HE approach is presented in [34]. The histogram is divided into three regions as dark, mid and bright. In order to keep original histogram features, the differential information is extracted from the input histogram, and then desired histogram is specified based on this information and some extra parameters such as direct current and gain value of the input image. [35]propose a modified version of histogram specification, in which a block around each pixel is defined and the desired histogram for that block is specified automatically. Histogram specification is done based on an optimization problem, whose main constraint is preserving the mean brightness of the block [32]. In histogram specification techniques, to reduce noise in enhancement produce, an efficient technique for real-time enhancement of video containing inconsistent and complex conditions like non-uniform and insufficient lighting is proposed. The method enhances video in low lighting conditions without any loss of color information and makes real-time enhancement for homeland security application successfully realized [17].

Different genetic approaches have been applied for images/videos contrast enhancement[36]. [37]uses a local enhancement technique. Genetic algorithms are meta-heuristic optimization techniques based on natural theory and survival of the fittest. [32]uses a simple chromosome structure and genetic operators to increase the visible details and contrast of low illumination images especially with high dynamic range. In particular the enhancement of very dark and blurred images has been of particular interest as many aspects of these images are ambiguous and uncertain [38]. The method based on transform function that stretches the occupied gray scale range for the image secondly the transformation function is optimized using genetic algorithms with respect to the test image.

To more clearly show contrast enhancement of HE-based, we attempt a brief review of existing systems of contrast enhancement methods. It should be mentioned that this review does not by any means cover all existing systems, but it rather considers representative algorithm that highlights the major trends in the area. Table 1 show a brief survey of HE-based.

B) Tone mapping

Tone mapping is another approach contrast enhancement technique. In this method, if we want to output high dynamic range (HDR) image on paper or on a display, we must somehow convert the wide intensity range in the image to the lower range supported by the display [39]. However, most LCD or CRT displays and print-outs have low dynamic range. Tone mapping technique used in image processing and computer graphics to map a set of colors to another, often approximate the appearance of high dynamic range images in media with a more limited dynamic range. Tone mapping is done in the luminance channel only and in logarithmic scale. It is used to convert floating point radiance map into 8-bit representation for rendering applications. The two main aims of tone mapping algorithm: preserving image details and providing enough absolute brightness information in low dynamic range tone mapped image. The existing techniques of tone mapping can be classified into categories: global tone mapping and local tone mapping. Global tone mapping function is based on logarithmic compression of luminance. It is a need to map the large range into a range that can be displayed given a HDR image with a dynamic range spanning many orders of magnitude. Simple mapping methods, such as the function $y = x/(x, y)$ will map a range of $x = [0, \infty)$ to the range $x = [0, 1)$. This method is considered a local operator since the operation only affects pixel values in an image individually on a pixel-by-pixel basis and each pixel is mapped in the same way. The global are independent of local spatial context. It performs the same operation on each pixel and don't work well when illumination varies locally. The simplest tone

Table 1. A brief survey of HE enhancement

Author	Year	Operating domain	Model	processing techniques	application
Agaian SS[1]	2007	Spatial domain	HE-based Logarithmic transform LTHS	log-reduction zonal magnitude technique; Logarithmic transform histogram shifting;	Traffic monitoring;Security Surveillance;
Hao Hu[4]	2010	Spatial domain Transform	Content adaptive video processing model	Content classification and adaptive processing	Computer vision
Tarik Arici[10]	2009	Spatial domain	HE-based modification	Histogram modification framework, content-adaptive algorithm	LCD display device;Low quality video
Sangkeun Lee[13]	2007	Spatial domain Transform domain	Dynamic range compression	Discrete cosine transform(DCT);Retinex theory;	Image/video compressing;
Viet Anh Nguyen[19]	2009	Transform domain	Cauchy distribution model; AC transform coefficient	Video reconstructed from multiple compressed copies of video content	Compression video
R.C. Gonzalez [21]	2008	Spatial domain Transform domain	HE	Global Histogram Equalization; Histogram specification technique	Image/video; Security surveillance;
Xuan Dong[23]	2010	Spatial domain	Image inverting model	Inverting the input low-lighting video; de-haze algorithm	Traffic monitoring;Medical image;
Shan Du[23]	2010	Spatial domain	ARHE model	Adaptive Region-based Method	Face Recognition
A.-A-Wadud, M[25]	2007	Spatial domain	Dynamic histogram equalization(DHE)	Dynamic Histogram Equalization technique	Medical image; Low quality video
Boudraa, A.O[30]	2008	Spatial domain	2DTKEO model	2D Teager-Kaiser Energy Operator technique	Medical image; Satellite image;
David Menotti[31]	2007	Spatial domain	MHE model	Multi-histogram equalization methods	Image processing; Computer vision;
Sara Hashemi[32]	2010	Spatial domain	Improve HE	Genetic algorithms	High dynamic range image processing
George D[33]	2009	Spatial domain	Improve HS and HE	Histogram-based video enhancement technique	Image processing;
Jafar,I [35]	2007	Spatial domain	ALHS Model	Automatic Specification of Local Histograms	Image processing; low quality video
Abhijit Mustafi [38]	2009	Spatial domain	Optimally enhance model	Genetic algorithms	Image processing; low quality video

reproduction is a linear mapping which scales the radiances to the range between 0 and 255. The logarithm of the radiances is taken and linearly scaled to [0, 255]. Tone mapping algorithm designed for high contrast images is widely accepted. Multipass-based technique first estimates local adaptation level, and applies simple tone mapping function to it, and then puts back image details. The approach follows functionality of human visual system (HVS) without attempting to construct its sophisticated model. The definition local contrast C at a pixel is as follows.

$$C(x, y) = L(x, y)/La(x, y) - 1 \tag{2}$$

where L is pixel luminance and La is local adaptation level, which we take to be just an average luminance over some neighborhood around pixel position (x, y).

Gradient-based domain tone mapping algorithm is used to display high dynamic range video sequences in low dynamic range devices. [40] obtain a pixel wise motion vector field and incorporates the motion information into the Poisson equation. Then, by attenuating large spatial gradients, the algorithm can yield a high-quality tone mapping result without flickering artifacts. Tone mapping technique can adjust image or video content for optimum contrast visibility taking into account ambient illumination and display characteristics.

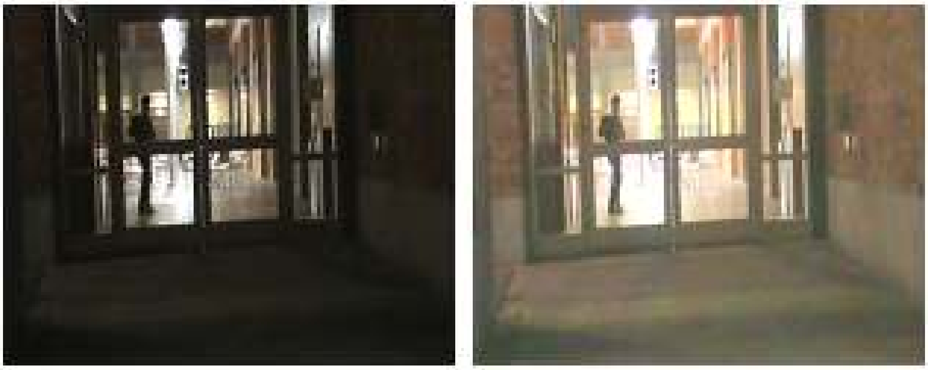
The operator weights contrast distortions to minimize given a display model that enforces constraints [41]. The approximation of an inverse tone mapping function can reduce the high dynamic range to displayable range. The most significant difference from the conventional methods is the use of an inverse tone mapping function. Similar technique with [41], for producing high-quality brightness enhancement functions for real-time reverse tone mapping of images and videos. [42] uses bilateral filter to obtain smooth results while preserving sharp luminance discontinuities, and can be efficiently implemented on GPUs. [43] makes use of two layer coding algorithm for high dynamic range images. First layer, a low dynamic range image is encoded by a conventional codec. Second layer, the residual information is represented the difference between an original and the decoded images using inverse tone mapping.

Tone mapping operator can minimize visible contrast distortions for a range of output devices. To enhance underexposed and low dynamic range videos, [15]uses adaptively and independently varying exposure at each photoreceptor in a post-process. The non-linear exposure variation and denoising filters smoothly transition from temporal to spatial for moving scene elements. System outputs restored video sequences with significantly reduced noise, increased exposure time of dark pixels, intact motion, and improved details. A tone mapping specialized for underexposed video should therefore associate a confidence level for details based on their luminous intensity. The mapping is given as follows [44].

$$m(x, \psi) = \frac{log(\frac{x}{xmax}(\psi - 1) + 1)}{log(\psi)} \tag{3}$$

In [7] uses similar the mapping functions with Ref.[15,44] to enhance the context of nighttime surveillance. Meanwhile, it also improves the contrast and signal to noise ratio. The method demonstrates a successful application of this tone mapping function to the nighttime video enhancement. Fig.3 shows the experimental result of Ref.[44] method.

2.2. **HDR-based video enhancement.** High dynamic range imaging (HDRI or just HDR) is a set of techniques that allow a greater dynamic range of luminances between the lightest and darkest areas of an image than standard digital imaging techniques or photographic methods. This wider dynamic range allows HDR images to more accurately represent the wide range of intensity levels found in real scenes, ranging from direct sunlight to faint starlight [45]. However, the dynamic range in real-world environments



(a) (b)

FIGURE 3. (a) Original nighttime video frame, (b) Enhanced nighttime video use tone mapping method in [44].

thus far exceeds the range represented in 8-bit per-channel texture maps. Extended dynamic range can be reached by combining multiple images of the same scene taken with different, known exposure times. The result is a floating point radiance map with radiance values being proportional to those observed in the real scene. The two main sources of HDR imagery are computer renderings and merging of multiple photographs, which in turn are known as low dynamic range (LDR) [39].

Video can encode higher contrast or dynamic range. HDR image and video formats are designed to encode all luminance levels that may vary from cd/m2(Illumination of the moonless sky) to cd/m2(Illumination of the sun). Similar to HDR imaging, there are several different ways to obtain HDR video, including capturing HDR images using video cameras, fixed mask, adaptive light modulator, and varying exposures for alternate frames. HDR video uses tone mapping across frames to convert the floating points to the 8-bits representative for rendering in the standard monitors. According to model of the human visual system, [46]propose a new definition of visible distortion based on the detection and classification of visible changes in the image structure. The metric is carefully calibrated and its performance is validated through perceptual experiments.

HDR video can be captured with HDR video capable hardware and merging of varied exposure frames. The hardware is limited and expensive. Therefore, the merging of varied exposure frames method is desirable. There are a lot of methods to obtain the sets of frames for use in generating radiance maps, and result in a tradeoff of either spatial resolution or temporal resolution. In the process of formatting HDR video for display, frame by frame tone mapping is inappropriate because variation in luminance values across frames can cause artifacts, thus temporal tone mapping is used. There are likewise many different temporal tone mappings methods which take into account global or neighboring frame luminance values, each with associated advantages and disadvantages. How to enhance video enhancement? An overview of some of the methods in these categories follows.

A) Radiance maps

To generate HDR image from a set of LDR images, an algorithm is required to combine and correctly map the relative luminance values for the images. The resulting map of luminance values is called a radiance map, which can contain values that can range over many orders of magnitude, depending upon the set of exposures. The major issue in implementation of the radiance map generation is the non-linearity of the capture devices.

During capture of a scene, the true relative luminance values recorded and saved by the device do not have a linear correspondence, where if one point of the scene has twice the luminance value of another point, the luminance value of the brighter pixel may not necessarily be twice that of the darker pixel. This non-linearity makes the process of merging photographs more difficult. To resolve this issue, [47] uses the constraint of sensor reciprocity to derive the response function and relative radiance values directly from a set of images taken with different exposures. The algorithm can fuse the multiple photographs into a single photograph. HDR radiance map of pixel values are proportional to the true radiance values in the scene. To correct lighting for new adding objects, [48] uses measured scene radiance and global illumination. To compute the illumination, the scene is considered as three components: the distant scene, the local scene, and the synthetic objects. Global illumination is used to simulate the interplay of light amongst all three components, except that light reflected back at the distant scene is ignored. Estimates of the geometry and material properties of the local scene are used to simulate the interaction of light between it and the synthetic objects.

B) HDR-based context enhancement

Video image quality improving technologies produce remarkable achievements as display devices make rapid progress. However, there exist some limitations on intensity representation of display and acquisition devices to reproduce the real world video images. Various context-based contrast enhancement algorithms have been developed and applied to overcome those limitations on intensity representation of display and acquisition devices to reproduce the real world video images. However, there are still some limits including color information change, loss of the lighting information, and excessive enhancement. To resolve these problems, [49]propose HDR imaging context enhancement algorithm. An input image is determined whether or not it requires context enhancement in a pre-processing step using the proposed auto exposure algorithm. Multiple images are generated by applying the intensity mapping function to an input image. An HDR image is constructed with multiple images, in which registration of multiple images is not required. The algorithm can increase the dynamic range and thus increase the contrast of an input image. A novel histograms approach using histograms to estimate the camera response function and the radiance mapping is proposed in [50]. The method applies histogram correspondences to register neighbor frames and computes pixel radiances performing ghost elimination.[51]establishes criteria for the display of both LDR and HDR images, given a specific HDR display device. The approach of the problem from the viewpoint of a user of HDR displays devices, rather than a designer of such devices.

C) HDR-based illumination enhancement

Video cameras with HDR output are particularly suitable for driving assistance applications, where lighting conditions can strongly vary, going from direct sunlight to dark areas in tunnels. However, common visualization devices can only handle a low dynamic range, and thus a dynamic range reduction is needed. Many algorithms have been proposed in the literature to reduce the dynamic range of still pictures. Some of the available methods to video are not straightforward to reduce the dynamic range of still pictures, due to the peculiar nature of video data [52]. For reducing the dynamic range of video sequences and enhancing its appearance, thus improving visual quality and reducing temporal artifacts.[53]provides an optimized version of the algorithm for a viable hardware implementation on an FPGA. A possible concern for the extended use of HDR displays is the potential to cause visual fatigue. Furthermore, ambient illumination has a significant effect on the perception of the imagery displayed, and its impact on user preferences for brightness and contrast must be understood. To examine these issues, [54]uses two user studies. In each study, subjects watched video content on an HDR display in several different ambient illumination environments, and are asked to adjust the brightness and black level of the display to their preference.

D) HDR-based temporal properties enhancement

The temporal properties of glare are a strong means to increase perceived brightness and to produce realistic and attractive renderings of bright light sources. Based on the anatomy of the human eye, [55] makes use of GPU to enable real-time simulation of dynamic glare. This allows an improved depiction of HDR images on LDR media for interactive applications like games, feature films, or even by adding movement to initially static HDR images. By conducting psychophysical studies, the method improves perceived brightness and that dynamic glare-renderings are often perceived as more attractive depending on the chosen scene. Furthermore, based on an adaptive spatio-temporal connective (ASTC) noise filter and an adaptive piecewise mapping function (APMF), for ill-exposed videos or those with much noise,[56]introduces novel local image statistic to identify impulse noise pixels, and incorporate it into the classical bilateral filter to form ASTC, aiming to reduce the mixture of the most two common types of noises-Gaussian and impulse noises in spatial and temporal directions. After noise removal, the methods enhance the video contrast with APMF based on the statistical information of frame segmentation results.

2.3. **Compressed-based video enhancement.** Compressed-based domain methods operate directly on the transform coefficients of the images that are compressed. The efficient method to handle and edit the compressed data has been introduced. And lots of efficient video coding algorithms laso has been introduced in [57,58]. There are three ways to enhance the compressed images/video [59, 60].

(i) Enhance the image/video before compression. However, there are two disadvantages of this approach. One is that enhancement will reduce the compressibility of the original image, and the other is that it will affect all the receivers.

(ii) Enhance the image/video after decompression. Because the post compression approach does not affect the compressibility of the original image, it is often adopted.

(iii) Enhance the image/video in the compressed domain. The basic idea of this method is to enhance the image by manipulating the DCT coefficients. Compared with the image enhancement in the spatial domain, this method can reduce storage requirements and computational expense as the majority of the coefficients in the DCT domain are zeros after quantization.

A) Discrete cosine transform

The enhancement algorithm, which enhances the images in the discrete cosine transform (DCT) domain by weighting the quantization table in the decoder, has seven advantages[13,61]: (i) the algorithm is fast because it operates directly in the compressed domain, (ii) suitability for real-time application, (iii) ease of adjustment by end-users (for example, adjusting a single parameter), (iv) less severe block artifacts as compared with conventional (post compression) enhancements, (v) the algorithm doesn't affect the compressibility of the original image because it enhances the images in the decompression stage, (vi) the approach is characterized by low computational complexity, and (vii) the algorithm is applicable to any DCT-based image compression standard, such as JPEG, MPEG , and H.26X, without any significant modification. A simple JPEG system is composed of an encoder and a decoder [60]. In the encoder, the image is first divided into nonoverlapping $8X8$ blocks. Then, the 2-D DCT is computed for each $8X8$ block via a forward DCT, defined for a $8X8$ block $I(i,j)$ as:

$$d(u,v)=\frac{1}{4}c(u)(v)\sum_{0}^{7}\sum_{0}^{7}cos\frac{(2i+1)u\pi)}{16}\times cos\frac{(2i+1)v\pi)}{16}I(ij) \tag{4}$$

For $u,v=0.1.2...7$
Where

$$c(\eta)=\begin{cases}\frac{1}{\sqrt{2}},\eta & =0\\ 1, otherwise\end{cases} \tag{5}$$

Once the DCT coefficients are obtained, they are quantized using a specified quantization table. Quantization of the DCT coefficients is a lossy process, and in this step, many small coefficients (usually high frequency) are quantized to zeros. The zig-zag scan of the DCT matrix followed by entropy coding makes use of this property to lower the bit rate required to encode the coefficients.

In the decoder, the compressed image is decoded and then dequantized by point-to-point multiplication with the quantization table and inverse DCT transformed. The DCT inverse transformation can be expressed as

$$I(x,y)=\frac{1}{4}\sum_{0}^{7}\sum_{0}^{7}c(u)(v)cos\frac{(2i+1)u\pi)}{16}\times cos\frac{(2i+1)v\pi)}{16}d(u,v) \tag{6}$$

For $u,v=0.1.2...7$, Each block of an image is reconstructed from the weighted sum of the DCT coefficients that correspond to the specific spatial frequency contributions. Thus, the distribution of the DCT coefficients provides a natural way to define a spectral content measure of the image in the DCT domain.

The enhancement algorithm is operating on the macroblock level of the decompressed video sequence. During the encoding process, the encoder generates several pieces of information, including macro-block type, quantization step size, and forward motion vectors. The enhancement algorithm captures both spatial and temporal correlation properties in an image sequence.

B) Compressed-based domain video enhancement

To compress image dynamic range and to enhance image contrast without boosting block artifacts and noise in the compressed domain, generally algorithm separate the DCT coefficients into illumination (dc coefficients) and reflectance (ac coefficients) components. The dc coefficients are adjusted based on the Retinex theory to compress the image dynamic range. To enhance the contrast, the coefficients are modified according to a newly defined measure of spectral content of the image[13, 61, 63]. To reduce block artifacts boosting in the target area of the image, [13] makes use of several DCT coefficients in low frequency bands to receive a special treatment during the enhancement process. In addition, a simple scheme to estimate and reduce the noise information directly in the DCT domain is employed for handling the image corrupted by noise. Furthermore, an image contrast enhancement algorithms for block discrete cosine transform (BDCT) based on compressed images is achieved by modifying the quantized DCT coefficients based on a contrast measure defined within the DCT domain in [63]. For compressing image dynamics and enhancing image details in the DCT domain, [61]introduces a simple multi-scale image enhancement algorithm. The algorithm can first divide an image into illumination and reflectance components. Then illumination component is manipulated adaptively for image dynamics by using a content measure, and reflectance component is altered by a multi-scale alpha-rooting method for enhancing image details based on human visual perception. The main advantage of the algorithm is that it enhances the details in the dark and the bright areas with low computations without boosting noise information and affecting the compressibility of the original image since it performs on the images in the compressed domain. Block-based DCT is widely used in many video compression standards to exploit the spatial redundancy of visual signal. Due to quantization errors, the decoded video may suffer from undesirable coding artifacts at moderate to low bit rates, such as blocking artifact and ringing noise, which may result in severe loss in visual quality and fidelity of the reconstructed video[19].

JPEG is a commonly used method of lossy compression for photographic images. The degree of compression can be adjusted, allowing a selectable tradeoff between storage size and image quality. Recently, some image/video based JPEG enhancement algorithms is presented [59,64-65]. For improving low-vision patients and enhancement image/video, enhancement algorithm of the JPEG standard is based on the contrast measure defined within the DCT domain [59]. To resolve the problem of the frequency content of each coefficient block in the DCT encoded JPEG image and the problem of implementing a nonlinear operator, some researchers use fuzzy theory compressed domain processing. [64]uses the compressed domain processing (no decompression/compression) and the pixel level processing (enhance the image after decompression) to enhance the images compressed with JPEG. However, for improving the visual quality of the image before it is decompressed, some researchers integrated sharpening the magnitude of the DCT coefficients into the JPEG compression processing in [65].

C) Compressed-based domain video enhancement

Resolution enhancement methods are becoming widely studied, but only a few procedures have been developed to work with compressed video, despite the fact that compression is a standard component of most image- and video-processing applications.

1) Based on motion vector resolution enhancement

Recent worldwide growth of digital video applications such as home AV network or mobile multimedia has been demanding higher video compression ratio in order to meet system requirements due to limited bandwidth or storage capacity. For improving video resolution enhancement from compressed video data, some researchers use interpolating motion vectors to resolve enhancement in [66,67]. For interpolating the transform domain coefficients and the motion vectors to zoom a compressed video stream, enhancement of temporal resolution using frame-rate up conversion (FRC) at decoder side is attractive in a sense that the R-D performance can be maintained even at low bitrates by dropping frames during encoding process and recovering them at the decoder. To provide accurate estimates of motion vectors for recovered frames based on compressed domain information, a new low-cost video FRC technique in compressed domain information is present in[67]. The aim of the study is to realize a decoder design with robust FRC over a variety of implementation platforms from mobile to consumer set-top.

The frame rates of video sequences differ because of the source of their video signals. To reproduce these video sequences on various display devices, it has to convert the frame rate of the input video sequence to the frame rate of the target display devices [68]. Temporal frame interpolation (TFI) algorithms are used to generate the intermediate frames at the frame rate required for the output display devices [69]. TFI algorithms are composed of motion estimation and motion compensation. In order to estimate true motion vectors (MVs) instead of using transmitted MVs for high-resolution displays, even though this increases the complexity of the hardware system, a bidirectional motion estimation method based on tracking feature trajectories and compensating for occlusion to enhance the temporal resolution of an input video sequence is used in [68].

2) Super-resolution enhancement

Super-resolution is a process of image enhancement by which low quality, low resolution (LR) images are used to generate a high quality, high resolution (HR) image. There are numerous applications of super-resolution in the areas of image processing and computer vision such as target detection, recognition, tracking etc. It has many applications in the consumer products such as cell phone, webcam, high-definition television, closed circuit television[70]. There have been several techniques for super-resolution which can be classified in two categories: Single image super-resolution [71,72] and super-resolution from several frames [70, 73]. Single image super-resolution is no additional information available to enhance the resolution. The algorithms are based on smoothing and interpolation techniques. Super-resolution from several frames is iterative back projection similar to the projections in computer aided tomography. The method starts with an initial guess of the HR image and simulated to generate a set of LR images which are compared with the observed image to update the HR image.

Bayesian framework can incorporate variable noise information and fuse the super-resolution and post-processing problems. The method establishes relationships between algorithm parameters and information in the compressed bit stream [72]. To increase the frame rate of video compressed, the algorithm inserts images between received frames of the sequence. However, the result still is blurred and subsamples to create a low-resolution image with half the number of original pixels in each dimension. An initial analytic interpolation, such as bicubic interpolation, is applied to the low-resolution image to generate an image of the desired size that lacks high-resolution detail[71].

Due to low cost sensors or physical limitation of the hardware, the resolution enhancement technique could be used as an inexpensive software alternative. [70] proposes kernel regression approach to reconstruct a high resolution image from several low resolution video frames. [73]expands kernel regression ideas for using in image denoising, up scaling, interpolation, fusion, and more. The method contacts with the field of nonparametric statistics and present a development and generalization of tools and results for use in image processing and reconstruction. In general, the classical super-resolution reconstruction (SRR) algorithms are usually based on translational observation model hence these SRR algorithms can be applied only on the sequences that have simple translation motion[74]. In order to cope with real video sequences and complex motion sequences,[75]propose SRR algorithm is based on maximum posteriori estimation technique by minimizing cost function. The classical L1 and L2 norm are used for measuring the difference between the projected estimate of the high-resolution image and each low resolution image, removing outliers in the data and errors due to possibly inaccurate motion estimation.

3) Bit-rate based resolution enhancement

For improving a low bit-rate compressed video sequence, general algorithms use the information provided by the encoder, which is spatio-temporally adaptive and enforces different degrees of between-block, within-block, and temporal smoothness of the decompressed frames based on macroblock types. Based on the rate-control mechanism and pre- and post-processing procedures within the context of MPEG is completely controlled by the system designer. Enhanced video is with low computations and no noise information [76]. Super-resolution algorithms recover high-frequency information from a sequence of low-resolution observations. For improving MPEG video enhancement, an MPEG-based image contrast enhancement algorithm for people with low vision is used as well. Contrast enhancement is achieved by modifying the inter- and intra-quantization matrices in the MPEG decoder during the decompression stage [77].

2.4. **Wavelet-based transform video enhancement.** Wavelet transform is the most exciting development in the last decade. The method focuses on wavelet-based image resolution enhancement and suitable for processing the image/video resolution enhancement. In mathematics, a wavelet series is a representation of a square-integrable (real- or complex-valued) function by a certain orthonormal series generated by a wavelet. The wavelet can be classified into four categories: continuous wavelet transform, discrete wavelet transform, complex wavelet transform, and dual wavelet. DWT is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution. It captures both frequency and location information. The most video enhancement in wavelet transforms domain use the DWT [78]. The integral wavelet transform is the integral transform defined as follows.

$$[W\psi, f](a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \psi(\frac{x-a}{b}) f(x) dx \tag{7}$$

The wavelet coefficients c_{jk} are then given by.

$$c_{jk} = [W\psi, f](2^{-j}, k2^{-j}) \tag{8}$$

Here $a = 2^{-j}$ is called the binary dilation or dyadic dilation, and $b = k2^{-j}$ is the binary or dyadic position.

A) Video enhancement based on wavelet shrinkage denosing

The method of wavelet shrinkage denosing is developed principally in [79]. The noise is reduced using threshold the empirical wavelet coefficients. The threshold is adaptive and a threshold level is assigned to each dyadic resolution level by the principle of minimizing the stein unbiased estimate of risk for threshold estimates. The enhanced video will reduce noise using wavelet shrinkage denosing method. Suppose a one-dimensional signal f from a noisy observation g. i.e.

$$g(n) = f(n) + q(n) \tag{9}$$

For $n = 0, 1, 2, ...N - 1$ where q is additive noise. The method attempts to reject noise by damping or thresholding in the wavelet domain. The estimate of the signal f is given by.

$$\hat{f} = w^{-1} T_\lambda w_g \tag{10}$$

where the operators w_g and w^{-1} stand for the forward and inverse discrete wavelet transforms, respectively, and T_λ is a wavelet-domain point-wise thresholding operator with a threshold λ.

The removal of noise in video signals has not been studied seriously. Since the success of the wavelet transform over other mathematical tools in denoising images, some researchers believe that wavelets may be successful in the removal of noise in video signals as well. The combination of wavelet image denoising and temporal filtering outperforms both wavelet based image denoising techniques [80]. In the case of video denoising, a robust, high-quality video denoising algorithm is required to not only be scalable to differing levels of noise corruption, but also scalable to differing amounts of motion in the original signal. Unfortunately, this principle has not been seriously considered in video denoising. [81]uses selective wavelet shrinkage in all three dimensions of the image sequence and proves to outperform the few video denoising algorithms given in the relevant literature. First, the individual frames of the sequence are denoised by using method in[80],which had developed earlier. Then a new selective wavelet shrinkage method is used for temporal-domain processing.

B) Video enhancement based on wavelet coefficients

To better preserve significant image features, which are identified via the spatial correlation of the wavelet coefficients at different scales, some researchers present a wavelet domain adaptive threshold scheme. Threshold scheme was performed only on the wavelet coefficients that do not correspond to any image features. The significant wavelet coefficients were determined via recursive hypothesis tests [20]. The wavelet-domain image resolution enhancement algorithm based on the estimation of detail wavelet coefficients at high resolution scales exploits shape function according to wavelet coefficient correlation in a local neighborhood and employs undedicated discrete wavelet transform to estimate the unknown detail coefficients [82]. For resolution enhancement of Omni-directional images based on wavelet transform, the degradation model of Omni-directional image is given [83]. The resolution enhancement of an image is achieved by using local extreme extrapolation of wavelet coefficients. Fusion operation is applied to the coefficients of registered pixels in the enhanced images of an image sequence. A fine resolution enhancement image is reconstructed via inverse wavelet transform. Recently, to obtain estimates of local edge orientation from a wavelet decomposition of the available low-resolution image, in [84] introduces directional variant of the cycle spinning methodology. This information to influence the choice of cycle spinning parameters is employed for resolution up scaling. The advantages include (i) lower computational complexity compared to the conventional cycle spinning, (ii) the outperforms competing methods for a wide range of images offering modest but consistent improvements both in objective as well as subjective terms.

For improving the clarity and continuity of ridge structures based on the multi-resolution analysis of global texture and local orientation by the wavelet transform, some the effect algorithm of image enhancement is proposed in[20,82-84].

C) Video enhancement based on shift invariant wavelet

The wavelet transform has been shown to be an invaluable tool in signal processing applications such as data compression and fast computations. However, the most commonly used implementation WT. The critically sampled DWT is shift variant and so is unsuitable for many signal analysis applications shift invariant. Based on the optimal shift invariant wavelet (SIWP) representation at the encoder, the optimal SIWP basis is searched using a fast optimization algorithm and the location of the best basis in the entire SIWP library is transmitted as overhead information to the decoder. The selected basis is jointly optimal in terms of both the time-frequency tiling and the relative time-domain offset (or shifts) between a signal and its wavelet packet representation. After the decoder reconstructs the compressed image, the postprocessor performs wavelet shrinkage using the optimal basis. However,[8,9]uses image fusion technique of shift-invariant discrete wavelet to integral all those context information in the final result of video enhancement. To overcome the shift dependency of the wavelet fusion method, the input images are decomposed into shift invariant wavelet representation and a composite shift invariant wavelet representation is built by the incorporation of an appropriate selection scheme. Using final SIWP fusion, the ghost problems of video enhancement is the better resolved for nighttime surveillance video.

D) Video enhancement based on dual-tree complex wavelet transform

Using the dual-tree complex wavelet transform (DT-CWT) with the sub-band coefficients modeled as Cauchy random variables can decompose image/video frame into multiresolution representations. [6]uses convolution of Cauchy distributions as a probabilistic prior to model the fused coefficients, and the weights used to automatically combine images of the same scene captured at different times or seasons are optimized via maximum likelihood estimation. The important map is produced to construct the composite approximation image. A unique characteristic of the algorithm is its ability to extract and maintain the meaningful information in the enhanced image while recovering the surrounding scene information by fusing the background image[85]. The algorithm has a lot of advantages. (i) Using convolution of Cauchy models, it is able to develop a generative model where the distribution of the fused sub band is determined by the distributions of the input sub-bands. (ii) The new model leads to a more accurate and reliable optimization process and doesn't take into account any assumption about the input images. And (iii) the applied DT-CWT provides near shift invariance and good directional selectivity while preserving the usual properties of perfect reconstruction and computational efficiency.

3. **Context-based fusion video enhancement.** Context-based fusion refers to insert high quality information from the same scene. E.g. to overcome bright regions and blurred details to enhance low visibility video, [86]proposes enhancing image features by using the information gathered from multiple images. Context-based enhancement is used in numerous applications such as surveillance and civilian or military image/video processing. There is a wealth of methods of context-based video enhancement. Some of them share common characteristics and guidance. It aims to detect, recognize and track objects such as people and cars from the image while being aware of the existing surroundings.

In this section, we firstly summarize a general algorithm of context-based fusion video enhancement. Then, we summarize techniques of video enhancement. At last, we analyze the typical representative systems of video enhancement.

3.1. **General algorithm of context-based fusion video enhancement.** The basic idea of context-based video enhancement is to extract and fuse the meaningful information of video sequence captured from a fixed camera under different illuminations[3]. The algorithm of video enhancement is automatically combining images of a scene at different time intervals by image fusion. All the important information of the original low quality video is combined with the context from a high quality background image at the same viewpoint. The fused image contains a comprehensive description of the scene which is more useful for human visual and machine perception. Fig.4 shows general algorithm framework of context-based video enhancement.

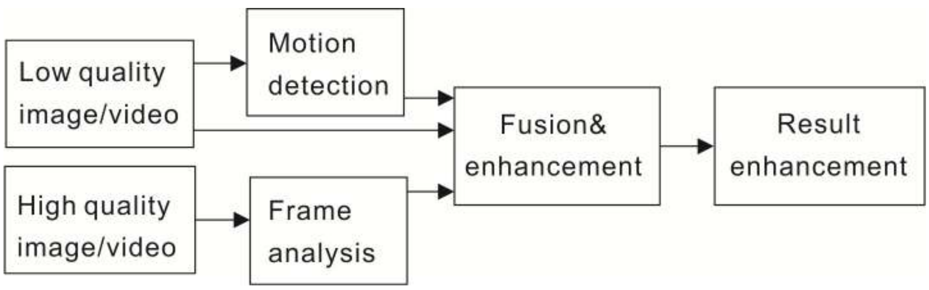


FIGURE 4. the block of general algorithm of context-based video enhancement

In the following, we analyze the techniques of the algorithm framework in Fig. 4.

A)Retinex theory

Retinex theory has been known for more than 30 years as a simple and effective model of the human vision. The name retinex [87-88], which comes from the contraction of two words "retina"and "cortex"indicate the intention to take into account the biological elements that influence our visual perception. The basic concept in the retinex theory is to separate the illumination and reflectance components of an image. It is assumed that the available illuminance data in the image is the product between illumination and reflectance. The reflectance component can be estimated as the ratio between the illuminance and an estimate of illumination. To estimate illumination information, it can use a low-pass filter.The methods include single scale Gaussian estimation [89], multi-scale Gaussian estimation[90], and bilateral filter estimation [91].

Using retinex theory to separate reflectance image and illumination image has several advantages[3,87-88].(i)the reflectance image and illumination image can be obtained from a single image instead of a sequence of images,(ii)this method does not require any learning and no training images are needed, and(iii) there is no assumption about lighting sources and shadow.

An input color image is decomposed into intensity image $I(x,y)$ and a color layer $C(x,y)$. The color layer $c \doteq (r,g,b)$ is given by dividing the input pixel values by the intensity $I(x,y)$ .Color space including RGB, HSV, Ycbcr, La*b*, HIS et al. Then, the intensity image $I(x,y)$ is decomposed into the illuminance layer $L(x,y)$ and the reflectance layer $R(x,y)$ by retinex theory. Intensity image $I(x,y)$ is represented by the product of the illuminace $L(x,y)$ and the reflectance image $R(x,y)$.

$$I(x,y) = L(x,y)R(x,y) \tag{11}$$

Illuminance $L(x,y)$ is assumed to be the low frequency component of an image $I(x,y)$. The reflectance image $R(x,y)$ is estimated as the ratio of the image $I(x,y)$ and the illuminace $L(x,y)$.

The bilateral filter is an edge-preserving smoothing filter that is developed in [92]. The output of the bilateral filter of an input image $I(x,y)$ at pixels, with the Gaussian function $g(x,\sigma) = exp(-x^2/\sigma^2)$, is defined by:

$$BF(I(s)) = \frac{1}{k(s)} \sum_p g(||p-s||;\sigma_s) g(|I(p)-I(s)|;\sigma_r) I(p) \tag{12}$$

$$k(s) = \sum_p g(||p-s||;\sigma_s) g(|I(p)-I(s)|;\sigma_r) \tag{13}$$

Where $g(||p-s||;\sigma_s)$ and $g(|I(p)-I(s)|;\sigma_r)$ are parameters of the Gaussian function in the spatial domain and the range of intensity difference, respectively. The output of the bilateral filter from the intensity image $I(x,y)$ provides the illumination layer $L(x,y)$.

$$I(x,y) = BF(I(x,y)) \tag{14}$$

The reflectance layer $R(x,y)$ is calculated from the ratio of the intensity image and the illumination layer from equation (14). Fig.5 shows the experimental result of using retinex theory.

B) Gaussian mixed model

Gaussian mixed model (GMM) is adapted to motion detection. The advantages of GMM include (i) it is an effective solution to real time motion detection due to its self learning capacity, and (ii) it is robust to variations in lighting, moving scene clutter, multiple moving objects compared with other methods. GMM for real time motion detection can be briefly summarized as follows[93].

Each pixel in the scene is modeled by a mixture of K Gaussian distributions. The probability that a certain pixel has a value of X_N at time N can be written as

FIGURE 5. Illumination image use retinex theory. (a) Daytime image illumination, (b) Nighttime image illumination

$$p(X_N) = \sum_{j=1}^{k} w_j \eta(X_n; \theta_j) \tag{15}$$

Where w_j is the weight parameter of the K^{th} Gaussian component. $\eta(X_n; \theta_j)$ is normal distribution of K^{th} component represented by

$$\eta(X; \theta_k) = \eta(X; u_k; \sum_k) = \frac{1}{(2\pi)^{D/2} |\sum_k|^{1/2}} e^{-\frac{1}{2}(X-u_k)^T \sum_k^{-1}(X-u_k)} \tag{16}$$

Where u_k is the mean and $\sum_k = \sigma_k^2 I$ is the convariance of the K^{th} component. K distributions are ordered based on the fitness value w_k/σ_k and the first B distributions are used as a model of the background of the scene where B is estimated as

$$B = argmin(\sum_{j=1}^{b} w_j > T) \tag{17}$$

The threshold is the minimum fraction of the background model. Background subtraction is performed by marking a foreground pixel any pixel that is more than 2.5 standard deviations away from any of the distributions. Fig.6 shows the experimental result of GMM.

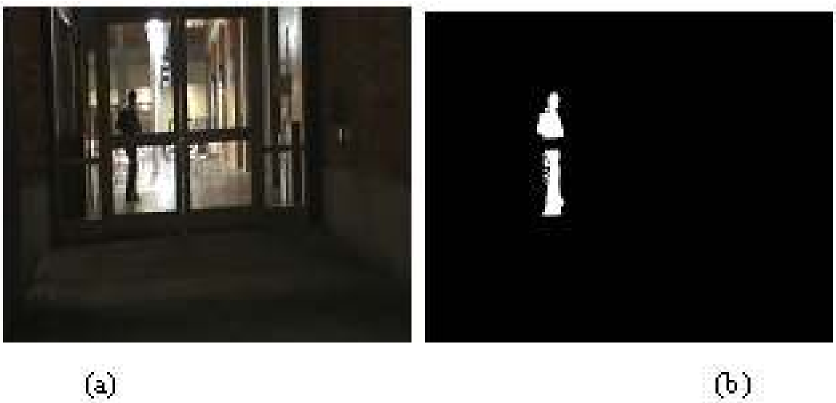


FIGURE 6. (a) Original nighttime video, and (b) the experimental result using GMM method.

C) Fusion methods

Image fusion is the process of combining information from two or more images of a scene into a single composite image that is more informative and is more suitable for visual perception or computer processing. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite that contains a "better"description of the scene than any of the individual source images. An example review of the state of the art fusion algorithms can be found in [94,95].

Fusion algorithms can be categorized into low, mid, and high levels. In some literature, this is referred to as pixel, feature, and symbolic levels. Pixel-level algorithms work either in the spatial domain or in the transform domain. Feature-based algorithms typically segmentation the images into regions and fuse the regions using their various properties. High-level fusion algorithms require correspondence between image descriptions for comparison and fusion. The actual fusion process can take place at different levels of information representation. Recently surveys of image fusion theory and applications can be found in[96, 97].

In video enhancement filed, some commonly used fusion techniques focus on pixel-level including gradient pyramid[3], shift invariant discrete wavelet transform[8,9], weighted combination[17,98], optimization approach and biologically based approaches such as neural networks [99] and contourlet transform[100,101], bio-inspired weight average image[102]. To analyze the existing algorithms of context-based video enhancement[3-9,94], we find a common fusion equation of video enhancement, which can be obtained by computing the weight average image at each pixel:

$$F(x, y) = N(x, y) * w(x, y) + D(x, y)(1 - w(x, y)) \tag{18}$$

Where $F(x, y)$ is the final fusion image, $N(x, y)$ is the low quality video, $D(x, y)$ is the high quality background. $w(x, y)$ is weight, value is set in the range[0, 1].The process of determining importance weights $w(x, y)$ depends on the specific application.

3.2. **Context-based fusion video enhancement.** Context-based video enhancement algorithms extract and maintain the meaningful information in the enhanced image the background image. Image fusion is and will be an integral part of many existing and future surveillance systems. Enhance low quality video has some common problems:(i)the obtained low quality video appear much noise, due to sensor noises or very low luminance. (ii) high light or dark areas in which the scene information can't be seen clearly by the observers. We can classify the methods to combine information from multiple images into one by noting which parameter of the scene or the camera is changing between successive images. Context-based enhancement method in most real surveillance scenes is based on the following two assumptions[2-3,7-9,103-106]. (i) the camera is fixed and can observe the same scene all day long.(ii)the scene model is coincident from day to night.

Combining regions of interest together by a multi-resolution based fusion method such as shift-invariant discrete wavelet transform, weighted combination, and optimization approach. The limitations include: (i) Define high illumination are in the night video as meaningful area not reasonable, (ii) Motion-based background model estimation and moving object segmentation heuristic and requires various thresholds which need adjustment when the scene is changed, and (iii) If an object of one source image is partly clear and partly blurry, the blurry part may be selected as part of the fused image when considering the integrality of the segmented part [8, 9]. Context-based enhancement generally adopts retinex theory to separate reflectance image and illumination image [3, 8, 9, 102-107]. It also adopts Gaussian mixed model to motion detection of video enhancement. However, using Gaussian mixed model will produce segmentation problems [7,8]. In [104] descript adaptive and integrated neighborhood dependent approach for nonlinear enhancement for improving the visual quality of digital images captured under extremely low or non-uniform lighting conditions. The method can decrease artifacts such as aliasing and ghosting by computing adaptive dynamic range compression of illuminance and adaptive enhancement for mid-tone frequency components.

[103]propose a denighting method to enhance the nighttime videos. The algorithm uses the illumination ratios of the daytime background and nighttime background videos to enhance the nighttime videos. The illumination component of the enhanced nighttime video is obtained by:

$$L_{eng} = \frac{L_{DB}(x,y)}{L_{NB}(x,y)} L_N(x,y) \tag{19}$$

where $L_{DB}(x,y)$ and $L_{NB}(x,y)$ represent the illumination components of the day-time and night-time background images, respectively. $L_N(x,y)$ represents the illumination component of the input night-time video. The enhanced results lose the static illumination. Fig.7 (a) shows experimental result of in [103]. The reason is that in those regions, the illumination ratios of the daytime background images and nighttime background images can be much smaller than 1, which tends to transform the static illumination of the night-time video back to the illumination of the day-time background video. To improve the contrast and signal to noise ratio in the fused image. [7]exploits more information from the fusion image than a single nighttime image. The fused image increases the information density and provides good input to high-level behavior analysis and understanding. Fig.7(b) shows experimental result. The limitation is that if the segmentation is correct, enhancement method using image segmentation and object extraction are powerful. But if errors occur in their process, unnatural mixture images may be generated in the result image.

Figure 7. (a) The experimental result use method in [103], (b) The experimental result use method in [7].

One of the changeling problems in video enhancement is enhance underexposed visible-spectrum video in non-visible spectra, such as short wave IR or near IR. Some researchers[15,16,105-106] introduce related technique to enhance image. IR sensors can capture video in low-light for night-vision applications. However it lacks the color and the relative luminances of visible spectrum sensors. RGB sensors capture color and correct relative luminances, but are underexposed, noisy, and lack fine features due to the short exposure times necessary for video. To remove noise from the RGB source, such as (i) removes shot noise and increase the color accuracy of the RGB footage, and (ii) normalized IR video to ensure cross-spectral compatibility with the visible-spectrum video using ratio images. Some fusion techniques utilizing both temporal and spatial filtering are used to ensure coherency from frame-to-frame. [15]introduces enhanced fusion output and reconstruction of RGB input assisted by the IR data, not an incorporation of elements imaged only in IR. To aid fusion, it decomposes the video sources with edge-preserving filters, which utilize the less-noisy IR for edge detection but also preserve strong visible-spectrum edges not in the IR. [15]uses related adaptive filtering approach that simulates virtual exposure control and transitions from temporal to spatial filter depending on the motion in the scene. For merging CCD and thermal images, [105,106]uses improved dual-tree complex wavelet transform fusion method. It can operate at either long wavelength infrared rang or short wavelength infrared range. The system applies nonlinear neighborhood dependent image enhancement to improve the visibility of images captured.

Gradient-based enhancement is another method of video enhancement. In image processing, gradient is defined as a gradual blend of color which can be considered as even gradation from low to high values, as used from white to black in the images. Some recently methods are that work in the gradient space rather than intensity space.

Gradient-based method doesn't improve the quality of the pixels themselves. It simply gives sufficient context to improve human interpretation. Consequently, operations such as contrast enhancement, histogram equalization, Gaussian mixed models for background estimation are orthogonal to approach and can be easily used alongside to improve the final result. Generally, two heuristics to decide what information to carry from daytime images into the desired result: (i) the gradients from the nighttime snapshot that appear to be locally important, and (ii) the gradients from the daytime snapshot to provide context to locally important areas while maintaining intra-image coherence.[3]report a different class of image and video enhancement techniques to enhance nighttime traffic or surveillance videos using daytime images taken from the same viewpoint. Fig.8 shows the experimental result using method in[3].

For rendering HDR images on conventional displays, it uses the gradient field of the luminance image of attenuating the magnitudes of large gradients. Then, a new low dynamic range image is obtained by solving a Poisson equation on the modified gradient field. The method can also be used to enhance ordinary images. By attenuating strong gradients and rescaling the reconstructed image back to the original 0...255 ranges, small contrasts in dark regions become easier to see. However, if low quality image have shadows in the same scene, the method doesn't able to significantly enhance ordinary images. To remove shadows in an image, first compute its gradient. For distinguishing shadow edges, setting the gradient values at the shadow edges to zero and finally reintegrating the image. The major disadvantages in [2, 3] are that extra processes are needed to deal with observable color shift, which is a common problem if gradient-based approaches. It can decrease artifacts such as aliasing and ghosting by computing in a gradient domain. However, these methods still produces aliasing and ghosting.

3.3. **Representative System and analysis.** Based on the previous categorization of context-based fusion video enhancement methods, we attempt a brief review of existing systems. We attempt a categorization of these methods in terms of their fusion method, the basic processing techniques used. This categorization is summarized in Table 2.

4. **Discussions.** The video enhancement is still an active area of research by many experts. There are still many problems of video enhancement, such as false background

FIGURE 8. Experimental result using the method in [3], (a) Gradient filed at axis (b) Gradient filed at axis, (c) Gradient field using and axis fusion, (d) Gradient field using and axis fusion with color.

problem, camera shift problem, color shift problem et al. So one of key problems is image/frame fusion problem to ensure better image reconstruction and color assignment [2, 3, 8, 9, 16, 94, 95, 102].

4.1. **Eliminate false background problem.** When using a combination of illumination fusion different time video or motion-based fusion to enhance low quality video, the most keys step is how to get the high-quality and clearly background image. General method uses medians of several images [2, 3, 7, 103], or select the Gaussian mixture model to exact moving objects [8, 9]. However, if we use medians of several images, there are many static objects in background image. E.g. a scene of daytime parking lot may contain many cars which parked at the parking lot for the whole day, while a scene of nighttime parking lot may be almost empty. So, the daytime background and night-time background could be quite different. If we use Gaussian mixture model, the background image is still false. E.g., in some cases some objects may remain in the frame for a long time. Previous researchers maintain running average or database of high quality images may alleviate this problem. However, there isn't the more appropriate method to resolve this problem.

4.2. **Camera shift problem.** To combine the high-quality image information to enhance low quality video, one of keys problem is to get the same surveillance scenes between the high-quality image and low quality video. For previous researchers, the assumption is that the camera is fixed and can observe the same scene all day long, and the scene model is coincident from day to night [2, 3, 7-9, 103-104]. However, most of the camera will shift usually, such as highway camera and parking camera. How to resolve camera shift problem is the focusing problem in surveillance. Recently, some researchers used the global motion compensation from coarsely sampled motion vector fields to camera

TABLE 2. User study results

Author	Year	Fusion techniques	Retinex theory	Motion detection	Application
YB RAO[2]	2010	Image-based fusion	yes	yes	Highway monitoring;
Raskar.R [3]	2005	Gradient domain methods	no	yes	Traffic monitoring; Low video quality;
H.Hu[4]	2010	Content classification and adaptive process	yes	yes	Computer vision; Video processing
Wan T [6,85]	2008 2007	Dual-tree complex wavelet transform; convolution of cauchy distributions	no	yes	Traffic monitoring; Campus monitoring;
Cai Y[7]	2006	Weighted combination	yes	yes	Traffic monitoring; Campus monitoring;
Li J[8,9]	2006 2009	Shift-invariant wavelet-transform	yes	yes	Night Surveillance
Bennett EP[15,16]	2005, 2007	Multispectral bilateral video fusion;	yes	no	underexposed visible-spectrum video
Melkamu H[86,100]	2010, 2009	contourlet transform	no	no	Image enhancement; Multi-sensor image enhancement
Wei Huang[101]	2007	Pulse coupled neural network source	no	no	Multi-focus image fusion using
Akito.Y [103]	2008	Denighting method;	yes	no	Surveillance camera
Li T[104, 105,106]	2006, 2005, 2004	Integrated neighborhood dependent approach; multi-sensor image fusion	yes	yes	Nonlinear enhancement of color images; Drivers in poor lighting conditions,

shift problem. However, when the scene/background is very dark, the method is still not appreciated.

4.3. **Color shift problem.** A common problem in gradient-based approaches is that how to deal with observable color shift problem. This phenomenon unfortunately has been a common problem of gradient-based approaches and can be observed in most previous works [2, 3]. There are two major reasons that cause the color shifts. (i) A valid vector field is not guaranteed to be maintained when modifying it with nonlinear operators. The gradient field of the result is only an approximation of the desirable one. (ii) In some cases, it is difficult to maintain the perception of high contrast in the result because the high quality and low quality snapshots are taken at significantly different exposure times [2, 3, 8].

DCT and wavelet transforms in video enhancement application is an interesting problem. It focuses on combine optimization method with spatial-domain, such as genetic algorithm. The other possible extension is to enforce the validity of the vector field when computing the gradients result. This requires using analytical operators to approximate nonlinear mask and blending function.

5. **Conclusions.** Video enhancement is one of the most important and difficult component of video security surveillance system. The increasing use of night operations requires more details and integrated information from the enhanced image. However, low quality video of most surveillance cameras is not satisfied and difficult to understand because they lack surrounding scene context due to poor illumination. A large number of techniques have been proposed to address this problem.

In this survey, we focus on survey the existing techniques of video enhancement, which can be classified into two broad categories: (i) Self-enhancement and (ii) Frame-based fusion enhancement. We show the existing technique of image/video enhancement and discuss the advantages and disadvantages of these algorithms. We also have described recent developments methods of video enhancement and point out promising directions on research for video enhancement for future research.

6. **Acknowledgment.** The authors would like to thank the anonymous reviewers for their helpful comments. This work is partly supported by National High-Tech Program 863 of China (Grant No. 2007AA010407 and 2009GZ0017) and China Scholarships Council.

REFERENCES

[1] S. S. Agaian, S. Blair and K. A. Panetta, Transform coefficient histogram-based image enhancement algorithms using contrast entropy, *IEEE Trans. Image Processing*, vol. 16, no. 3, pp. 741-758, 2007.
[2] YunBo Rao, W. Lin and L. T. Chen, Image-based fusion for video enhancement of nighttime surveillance, *Optical Engineering Letters*, vol. 49, no. 2, pp. 120501-1- 120501-3, 2010.
[3] A. Ilie, R. Raskar and J. Yu, Gradient domain context enhancement for fixed cameras, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 4, pp. 533-549, 2005.
[4] H. Hu, Video enhancement: content classification and model selection, Ph. D. Thesis, Technique Universities Eindhoven, Eindhoven, Netherlands, 2010.
[5] P.Y Liu, K.B Jia, Research and Optimization of Low-Complexity Motion Estimation Method Based on Visual Perception, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 3, pp. 217-226, 2011.
[6] T. Wan , T. George, T. Panagiotis,C. Nishan and A. Alin, Context enhancement through image fusion: a multi-resolution approach based on convolution of Cauchy distributions, *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing* , pp. 1309-1312, 2008.
[7] Y. Cai, K. Huang, T. Tan and Y. Wang, Context enhancement of nighttime surveillance by image fusion, *Proc. of the IEEE 8th International Conference on Pattern Recognition*, pp. 980-983, 2006.
[8] J. Li, Z.li Stan, P. Quan and Y. Tao, Illumination and motion-based video enhancement for night surveillance, *Proc. of the IEEE 2nd International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pp. 169-175, 2005.
[9] J. Li, Y. Tao, P. Quan and Y. Cheng, Combining scene model and fusion for night video enhancement, *Journal of Electronics*, vol. 26, no. 1, pp. 88-93, 2009.
[10] YunBo Rao and Leiting Chen, An efficient contourlet transform-based algorithm for video enhancement, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 3, pp. 282-293, 2011.
[11] V. Kastrinaki, M. Zervakis and K. Kalaitzakis, A survey of video processing techniques for traffic applications, *Image and Vision Computing*, vol. 21, pp. 359-381, 2003.
[12] R. Jafri and H. R. Arabnia, A survey of face recognition techniques, *Journal of Information Processing Systems*, vol. 5, no. 2, pp. 41-68, 2009.
[13] S. Lee, An efficient content-based image enhancement in the compressed domain using Retinex theory, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, no. 2, pp. 199-213, 2007.
[14] K. Junga, K. Kimb and A. K. Jainc, Text information extraction in images and video: a survey, *Pattern Recognition*, vol. 37, pp. 977-997, 2004.
[15] E. P. Bennett and L. McMillan, Video enhancement using per-pixel virtual exposures, *ACM Trans. Graphics*, vol. 24, no. 3, pp. 845-852, 2005.
[16] E. P. Bennett, J. L. Mason and L. McMillan, Multispectral bilateral video fusion, *IEEE Trans. Image Processing*, vol. 16, no. 5, pp. 1185-1194, 2007.

[17] G. Mittal, S. Locharam, S. Sasi, G. R. Shaffer and A. K. Kumar, An efficient video enhancement method using LA*B* analysis, *Proc. of IEEE International Conference on Video and Signal based Surveillance*, pp. 66-71, 2006.

[18] Yunbo Rao, Zhongho Chen, Ming-Ting Sun, Yu-Feng Hsu and Zhengyou Zhang, An effective night video enhancement algorithm, *Visual Communications and Image Processing*, pp. 1-4, 2011.

[19] V. A. Nguyen, Y. P. Tan and Z. H. Chen, On the method of multicopy video enhancement in transform domain, *Proc. of the 16th IEEE International Conference on Image Processing*, pp. 2777-2780, 2009.

[20] S. Walid and A. Ibrahim, Real time video sharpness enhancement by wavelet-based luminance transient improvement, *Proc. of the 9th International Symposium on Signal Processing and Its Applications*, pp. 1- 4, 2007.

[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, New Jersey, 2008.

[22] S. D. Chen and A. R. Ramli, Minimum mean brightness error bihistogram equalization in contrast enhancement, *IEEE Trans. Consumer Electronic*, vol. 49, no. 4, pp. 1310-1319, 2003.

[23] X. Dong, Y. Pang and J. Wen, Fast efficient algorithm for enhancement of low lighting video, *Proc. of the 37th International Conference and Exhibition on Computer Graphics and Interactive Techniques*, 2010.

[24] S. Du and R. K. Ward, Adaptive region-based image enhancement method for robust face recognition under variable illumination conditions, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 99, pp. 1-12, 2010

[25] A. A. Wadud, M. Kabir, M. H. Dewan and M. C. Oksam, A dynamic histogram equalization for image contrast enhancement, *IEEE Trans. Consumer Electronic*, vol. 53, no. 2, pp. 593-600, 2007.

[26] J. A. Stark, Adaptive image contrast enhancement using generalizations of histogram equalization, *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 889-896, 2000.

[27] A. Polesel, G. Ramponi and V. J. Mathews, Image enhancement via adaptive unsharp masking, *IEEE Trans. Image Processing*, vol. 9, no. 3, pp. 505-510, 2000.

[28] J. Y. Kim, L. S. Kim and S. H. Hwang, An advanced contrast enhancement using partially overlapped sub block histogram equalization, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 475-484, 2001.

[29] Z. Yu and C. Bajaj, A fast and adaptive method for image contrast enhancement, *Proc. of the IEEE International Conference on Image Processing* , vol. 2, pp. 1001-1004, 2004.

[30] A. O. Boudraa and E. H. S. Diop, Image contrast enhancement based on 2D teager-kaiser operator, *Proc. of the IEEE International Conference on Image Processing* , pp. 3180-3183, 2008.

[31] D. Menotti, L. Najman, J. Facon and Arnaldo de A. Araujo, Multi-histogram equalization methods for contrast enhancement and brightness preserving, *IEEE Trans. Consumer Electronics*, vol. 53, no. 3, pp. 1186-1194, 2007.

[32] S. Hashemi, S. Kiani, N. Noroozi and M. E. Moghaddam, An image contrast enhancement method based on genetic algorithm, *Pattern Recognition Letters*, vol. 31, no. 13, pp. 1816-1824, 2010.

[33] D. G. Toderici and J. Yagnik, Automatic, efficient, temporally-coherent video enhancement for large scale applications, *Proc. of the 17th ACM International Conference on Multimedia*, pp. 609-612, 2009.

[34] S. Srinivasan and N. Balram, Adaptive contrast enhancement using local region stretching, *Proc. of the 9th Asian Symposium on Information Display* , pp. 152-155, 2006.

[35] I. Jafar and H. Ying, A new method for image contrast enhancement based on automatic specification of local histograms, *International Journal of Computer Science and Network Security* , vol. 7, no. 7, pp. 1-10, 2007.

[36] Z. J. Chang. and X. D Wang, Global and local contrast enhancement for image by genetic algorithm and wavelet neural network, *LNCS Book Series*, Springer, pp. 910-919, 2006.

[37] M. Paulinas and A. Usinskas, A survey of genetic algorithms applications for image enhancement and segmentation, *Information Technology and Control*, vol. 36, no. 3, pp. 278-284, 2007.

[38] A. Mustafi and P. K. Mahanti, An optimal algorithm for contrast enhancement of dark images using genetic algorithms, *Computer and Information Science, Studies in computational intelligence*, vol. 208, pp. 1-8, 2009.

[39] E. Reinhard, G. Ward, S. Pattanaik and P. Debevec, High dynamic range imaging: acquisition, display, and image-based lighting, *Morgan Kaufmann Publishers Incorporated*, San Francisco, CA, 2005.

[40] C. Lee, C. S. Kim, Gradient domain tone mapping of high dynamic range videos, *Proc. of IEEE International Conference on Image Processing*, no. 3, pp. 461-464, 2007.

[41] R. Mantiuk, S. Daly and L. Kerofsky, Display adaptive tone mapping, *ACM Trans. Graphics*, vol. 27, no. 3, pp. 681-690, 2008.

[42] R. P. Kovaleski and M. M. Oliveira, High-quality brightness enhancement functions for real-time reverse tone mapping, *Visual Computer*, vol. 25, no. 5, pp. 539-547, 2009.

[43] T. Jinno, M. Okuda and N. Adami, Acquisition and encoding of high dynamic range images using inverse tone mapping, *Proc. of IEEE International Conference on Image Processing*, vol. 4, pp. 181-184, 2007.

[44] F. Durand and J. Dorsey, Fast bilateral filtering for the display of high dynamic range images, *Proc. of the 29th International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 257-266, 2002.

[45] Z. G. Li, S. Rahardja, Z. J. Zhu, S. L. Xie and S. Q. Wu, Movement detection for the synthesis of high dynamic range images, *Proc. of the IEEE International Conference on Image Processing*, pp. 3133-3136, 2010.

[46] T. O. Aydin, R. Mantiuk, K. Myszkowski and H. P. Seidel, Dynamic range independent image quality assessment, *Proc. of the 35th International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 1-10, 2008.

[47] Y. S. Heo, K. M. Lee, S. U. Lee, Y. SMoon and J. Y. Cha, Ghost-free high dynamic range imaging, *Proc. of Asian Conference on Computer Visio*, Lecture Notes in Computer Science, pp. 486-500, 2011.

[48] P. E. Debevec, Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography, *Proc. of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 189-198, 2008.

[49] W. Zhang and W. K. Cham, Gradient-direct composition of multi-exposure images, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 530-536, 2010.

[50] L. Velho, Histogram-based HDR video, *Proc. of the 34th International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 62, 2007.

[51] A. O. Akyuez, B. E. Riecke, E. Reinhard and H. H. Buethoff, Do HDR displays support LDR content A psychophysical evaluation, *ACM Trans. Graph*, vol. 26, no. 3, pp. 38, 2007.

[52] P. Didyk, R. Mantiuk, M. Hein and H. P. Seidel, Enhancement of bright video features for HDR displays, *Computer Graphics Forum*, vol. 27, no. 4, pp. 1265-1274, 2008.

[53] S. Marsi, G. Impoco, A. Ukovich and G. Ramponi, Video enhancement and dynamic range control of HDR sequences for automotive applications, *EURASIP Journal on Advances in Signal Processing*, pp. 1265-1274, 2007.

[54] G. R. Allanel, W. Heidrich1, H. Li and R. Mantiuk, Video viewing preferences for HDR displays under varying ambient illumination, *Proc. of the 6th Symposium on Applied Perception in Graphics and Visualization*, pp. 45-52, 2009.

[55] T. Ritschel, M. Ihrke, JR Frisvad, J. Coppens, K. Myszkowski and H. P. Seidel, Temporal Glare: real-time dynamic simulation of the scattering in the human eye, *Compute Graph Forum*, vol. 28, no. 2, pp. 183-192, 2009.

[56] C. Wang, L. F. Sun, B. Yang, Y. M. Liu and S. Q. Yang, Video enhancement using adaptive spatio-temporal connective filter and piecewise mapping, *EURASIP Journal on Advances in Signal Processing*, pp. 1-13, 2008.

[57] H. Wang, Joyce Liang and C. C. Jay Kuo, Overview of Robust Video Streaming with Network Coding, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 1, pp. 36-50, 2010.

[58] J. Lou, Shan Liu, Anthony Vetro and Ming-Ting Sun, Trick-Play Optimization for H.264 Video Decoding, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 2, pp. 132-144, 2010.

[59] J. Tang, E. Peli and S.Acton, Image enhancement using a contrast measure in the compressed domain, *IEEE Signal Processing Letters*, vol. 10, no. 10, pp. 289-292, 2003.

[60] D. Liu, X. Y Sun, F. Wu, S. P Li and Y. Q. Zhang, Image compression with edge-based in painting, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1273-1286, 2007.

[61] S. Lee, Content-based image enhancement in the compressed domain based on multi-scale alpha-rooting algorithm, *Pattern Recognition Letters*, vol. 27, pp. 1054-1066, 2006.

[62] R. Ansari and N. Memon, The JPEG lossy image compression standard, *Handbook of Image and Video Processing*, Academic Press, pp. 513-526, 2000.

[63] A. Kebin, Q. Ni, J. Sun, A contrast enhancement method for compressed images, *IEICE Electronics Express*, vol. 1, no. 18, pp. 582-587, 2004.

[64] P. Camelia, A. Vlaicu, M. Gordan and B. Orza, Fuzzy contrast enhancement for images in the compressed domain, *Proc. of the International Multi-Conference on Computer Science and Information Technology*, pp. 161-170, 2007.

[65] M. Amore, S. K. Mitra, M. Carli and J. S. McElvain, Image enhancement in the compressed domain, *Proc. of the International Symposium on Circuits and Systems*, pp. 1-4, 2007.

[66] N. Kaulgud, J. Karlekar and U. B. Desai, Compressed domain video resolution enhancement: use of motion vectors, *Proc. of the 3rd Indian Conference on Computer Vision, Graphics and Image Processing*, 2002.

[67] S. Sekiguchi, Y. Idehara, K. Sugimoto and K. Asai, A low-cost video frame-rate up conversion using compressed-domain information, *Proc. of IEEE International Conference on Image Processing*, vol. 2, pp. 974-977 ,2005.

[68] Y. H. Cho, H. Y. Lee, D. S. Park and C. Y. Kim, Enhancement for temporal resolution of video based on multi-frame feature trajectory and occlusion compensation, *Proc. of the IEEE International Conference on Image Processing*, pp. 389-392, 2009.

[69] Q. D. Le, Next generation of frame-rate conversion algorithm, *SID Symposium Digest of Technical Papers*, vol. 39, no. 1, pp. 998-1001, 2008.

[70] M. M. Islam, V. K. Asari and M. A Karim, Super-resolution enhancement technique for low resolution video, *IEEE Trans. Consumer Electronics*, vol. 56, no. 2, pp. 919-924, 2010.

[71] D. Li, S. Simske and R. M. Mersereau, Single image super-resolution based on support vector regression, *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2898-2901, 2007.

[72] A. C. Segall, A. K. Katsaggelos, R. Molina and J. Mateos, Bayesian resolution enhancement of compressed video, *IEEE Trans. Image Processing*, vol. 13, no. 7, pp. 898-911, 2004.

[73] H. Takeda, S. Farisu and P. Milanfar, Kernel regression for image processing and reconstruction, *IEEE Trans. Image Processing*, vol. 16, no. 2, pp. 349-366, 2007.

[74] V. Patanavijit and S. Jitapunkul, General observation model for an iterative multiframe regularized super-resolution reconstruction for video enhancement, *Proc. of International Symposium on Intelligent Signal Processing and Communication Systems* , pp. 1-4, 2008.

[75] V. Patanavijit and S. Jitapunkul, A Modified three-step search algorithm for fast affine block base motion estimation, *Proc. of International Workshop on Advanced Image Technology*, pp. 99-104, 2006.

[76] C. Andrew C. Segall and Aggelos K. Katsaggelos, Enhancement of compressed video using visual quality measurements, *Proc. of IEEE International Conference on Image Processing*, vol. 2, pp. 645-648, 2000.

[77] J. Kim, A. Vora and E. Peli, MPEG-based image enhancement for the visually impaired, *Optical Engineering*, vol. 43, no. 6, pp. 1318-1328, 2004.

[78] B. Zhu, A. Y. T. Leung, C. K. Wong and W. Z. Lu, On-line health monitoring and damage detection of structures based on the wavelet transform, *International Journal of Structural Stability and Dynamics*, vol. 8, no. 3, pp. 367-387, 2008.

[79] D. L. Donoho and I. M. Johnstone, Adaptation to unknown smoothness via wavelet shrinkage, *Journal of the American Statistical Association* , pp. 1200-1244, 1995.

[80] E. J. Balster, Y. F. Zheng and R. L. Ewing, Feature-based wavelet shrinkage algorithm for image denoising, *IEEE Trans. Image Processing*, vol. 14, no. 12, pp. 2024-2039, 2005.

[81] E. J. Balster, Y. F. Zheng and R. L. Ewing, Combined spatial and temporal domain wavelet shrinkage algorithm for video denoising, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 220-231, 2006.

[82] A. Temizel, Image resolution enhancement using wavelet domain hidden Markov tree and coefficient sign estimation, *Proc. of International Conference on Image Processing*, pp. 381-384, 2007.

[83] Peng Qimin and Jia Yunde, Wavelat-based resolution enhancement of omnidirectional images, *Acta Electronica Sinica*, no. 11, pp. 125-132, 2004.

[84] A. Temizel and T. Vlachos, Wavelet domain image resolution enhancement using cycle-spinning, *Journal of Electronic Imaging, Electronics letters*, vol. 41, no. 3, pp. 119-121, 2005.

[85] Tao Wan , N. Canagarajah and A. Achim, Multiscale color-texture image segmentation with adaptive region merging, *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. I-1213 - I-1216, 2007.

[86] M. H. Asmare, V. S. Asirvadam, L. Iznita and A. F. M. Hani, Image enhancement by fusion in contourlet transform, *International Journal on Electrical Engineering and Informatics*, vol. 2, no 1, pp. 29-42, 2010.

[87] E. H. Land and J. J. McCann, Lightness and retinex theory, *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1-11, 1971.

[88] B. Funt, F. Ciurea and J. MacCann, Retinex in matlab, *Proc. of the 8th IS&T/SID Eighth Color Imaging Conference*, pp. 112-121, 2000.

[89] D. Jobson, Z. Rahman and G. Woodell, Properties and performance of a center/surround Retinex, *IEEE Trans. Image Processing*, vol. 6, no. 3, pp. 451-462, 1997.

[90] D. Jobson, Z. Rahman and G. Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 965-976, 1997.

[91] C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, *Proc. of IEEE International Conference on Computer Vision*, pp. 836-846, 1998.

[92] F. Durand and J. Dorsey, Fast bilateral filtering for the display of high dynamic range images, *Proc. of the 29th International Conference Exhibition on Computer Graphics and Interactive Techniques*, pp. 257-266, 2002.

[93] C. Stauffer and G. Wel, Learning patterns of activity using real-time tracking, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747-757, 2000.

[94] F. Sadjadi, Comparative image fusion analysis, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 8, pp. 1029-1051, 2005.

[95] A. Ardeshir Goshtasby and S. Nikolov, Image fusion: advances in the state of the art, *Information Fusion*, vol. 8, no. 2, pp. 114-118, 2007.

[96] M. I. Smith and J. P. Heather, Review of image fusion technology in 2005, *Proc. of the Symposiumon on Defense and Security* , pp. 29-45, 2005.

[97] RS. Blum and Z. Liu, Multi-sensor image fusion and its applications, *CRC Press, Taylor and Francis Group*, Florida, 2006.

[98] E. Lallier and M. Farooq, A real time pixel-level based image fusion via adaptive weight averaging, *Proc. of the 3rd International Conference on Information Fusion*, pp. WeC3/3 - WeC3/13, 2000.

[99] Yunbo Rao, Weiyao Lin and Leiting Chen, A global-motion-estimation-based method for nighttime video enhancement, *Optical Engineering*, vol. 50, no. 5, pp. 057203-1 - 057203-5, 2011

[100] H. A. Melkamu, V. S Asirvadam and L. Iznita, Multi-sensor image enhancement and fusion for vision clarity using contourlet transform, *Proc. of International Conference on Information Management and Engineering*, pp. 352-356, 2009.

[101] W. Huang and Z. L Jing, Multi-focus image fusion using pulse coupled neural network source, *Pattern Recognition letters*, vol. 28, no. 9, pp. 1123-1132, 2007.

[102] F. Sadjadi, Comparative image fusion analysis, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 8-15, 2005.

[103] A. Yamasaki, H. Takauji, S. Kaneko, T. Kanade and H. Ohki1, Denighting: enhancement of night-time image for a surrveillance camera, *Proc. of ICPR 19th International Conference on Pattern Recognition* , pp. 1-4, 2008.

[104] T. Li and V. K. Asari, An efficient illuminance-reflectance nonlinear video stream enhancement model, *Proc. of SPIE Conference on Real-Time Image Processing* ,vol. 6063, pp. 60630I. 1-60630I. 12, 2006.

[105] T. Li, H. Ngo, M. Zhang, A. Livingston and V. Asari, A multi-sensor image fusion and enhancement system for assisting drivers in poor lighting conditions, *Proc. of the 34th Applied Imagery and Pattern Recognition workshop* , pp. 106-113, 2005.

[106] T. Li and K. V. Asari, An integrated neighborhood dependent approach for nonlinear enhancement of color images, *Proc. of International Conference on Information Technology: Coding and Computing*, pp. 138-139, 2004.