LOW LIGHT VIDEO PROCESSING FOR ENGINEERS: AN OVERVIEW

Abstract

"Low Light Video Processing for Engineers: An Overview", offers a comprehensive exploration of the challenges and advancements in the domain of low-light video processing. As the demand for high-quality video in low-light conditions continues to rise across various applications, this paper presents a valuable overview of image quality in challenging lighting scenarios. From understanding the underlying principles of image sensing in low light to exploring state-of-the-art denoising, illumination correction, and image fusion techniques, this abstract aims to equip engineers with a foundational understanding of low-light video processing. Moreover, it highlights the key considerations and trade-offs involved in implementing such processing methods, emphasizing real-world applications in surveillance, automotive vision, and night-time photography. By grasping the fundamental concepts presented in this overview, engineers can embark on a journey to create cutting-edge solutions that significantly improve video quality and reliability under adverse lighting conditions.

Keywords: Low Light, Surveillance, Illumination, Video Processing, Noise.

Author

C. Anitha
Professor
CSE Department
Saveetha School of Engineering,
SIMATS, Chennai, TamilNadu, India.
canithathomas@gmail.com

I. INTRODUCTION

Digital images and video have become integrated in our daily lives. It is commonly known that video enhancement has achieved a significant amount of popularity attention recently, as a big issue in computer vision.The aim is to make the video seem better visually or in the future, to provide a "better" transform representation for automated video processing including analysis, detection, segmentation, and recognition.Digital video may be captured, analyzed, and utilized for a variety of purposes, including surveillance, automated driving, long range monitoring, image tracking, identity verification in general, traffic congestion, criminal justice systems, and civilian or military video processing are all areas where video processing is used.

Low light video is not necessarily just night video. Different levels of light can come from many sources, and anything that is lower than midday outside light is termed lowlight.Indoor photography of low ambient light (such as in most of our homes), as well as night photography with light that is hardly visible to our eyes and likewise videos affected by environmental condition such as fog, haze, shadow, rain, cloud etc., is also considered to be low-light.

Usually, under typical lighting circumstances, computer vision may produce considerable results. In lowlight situations, however, an image becomes noisy and dark, making future computer vision tasks challenging. A low-light image improvement work is required to make information that has been hidden more visible moreoverreduce the amount of blur and noise in a low-light image.Many distinct approaches have been subjected to extensive investigation.However, performing low-light image improvement with most of these methods requires a lot of effort or expensive equipment. For example, in order to be processed, the image must be shot as a raw camera format, in addition, the addressing approach is ineffective in low-light situations. Low light videos are affected by high level of noise and low dynamic range.Lumens per Square Meter (Lux) is a measure of a camcorder's low-light capabilities. Extremely low light videos are the videos with relative luminance lower than 0.1Lux.

Figure 1.1: A video with moderately low light (on the left) and extremely low light (on the right)

Furthermore, it aids in the analysis of background data i.e. necessary toward comprehend object behaviors without the need for intensive and visual inspection is costly. Because of the following reasons, doing video enhancement understanding under low-quality video is a difficult task.We can't see moving things against a dark background because of low contrast. If the colour of the moving items and the background are similar, most color based techniques will fail in this case. An example of a video with moderately low-light plus extremely low-light is shown Figure 1.1.Due to high International Standards Organization (ISO) noise levels, the Signal to Noise Ratio (SNR) is generally quite low. ISO stands for the sensitivity of a camera's sensor to light.In digital pictures, using a high ISO level might result in apparent noise. The lower the ISO value, the less sensitive the camera is to light. The video signal is a distorted version of the actual or source video signal, which shows a three-dimensional continuous environment. The acquisition procedure, as well as the rate and format conversion operations, might cause these performance degradation.Environmental knowledge has an impact on how individuals observe and comprehend the situation.As a result, dealing with moving trees, fog, rain, and people's behavior in nighttime videos is challenging owing to a lack of background context due to inadequate illumination. Inter-frame coherence should be maintained, which means that the weights in subsequent images should vary smoothly in the region of moving objects.Even though the local fluctuation is modest, one pixel as of a low quality image, such as the region between a moving car's headlights and taillights, can be significant. The operator's lack of competence and the low quality of the employed video equipment.As a result, distinct enhancing processes must be used for this type of video in order to extract comprehensive information from it.

Low-light visionimprovement has a long history. To improve the subjective and objective quality of low-light images, many approaches have been developed.There are various traditional techniques, such as Histogram Equalization (HE) (Bassiou, 2007), Dynamic Histogram Equalization (DHE) (Abdullah-Al-Wadud, 2007), contrast enhancement of low-light images using histogram equalization and illumination adjustment (Banik, 2018), the Retinex-based theory (Land, 1977), and the Multi Scale Retinex model (MSR-net) (Jobson, 1997). HE presents as a most popular algorithm for image enhancement because it increases the intensity of the image for better quality. Adaptive Histogram Equalization (AHE) does the task of converting pixel intensity into a proportionate display range while simultaneously improving the local intensity histogram, which reduces the shadowing edges. These techniques often produce unnatural and unrealistic results, because some priors or assumptions are not good enough to hold for different illumination conditions. Because the image can be identified from two factors: reflectance and illumination, Retinex was the first effort at the Retinex theory–based approach. Reflectance is the ultimate enhancing result of Multiscale Retinex with colour restoration, and it is utilized to improve the estimation accuracy of the illumination component by utilizing a variety of complex filtering algorithms. However, the typical performance outcomes are artificial, and the image is often over-enhanced.

Aside from the aforementioned techniques, some research have gone farther and suggested deep learning-based algorithms to address these difficulties. The powerful capability of a deep neural network (Gonzalez, 2007) has led to robust improvements over object recognition. The LLNet proposed in (Lore, 2017)is a deep learning-based method to capture images in low-light environments, which enhances and denoises image captures simultaneously. LLNet applied the relationships between a lightless image and a ground truth

augmented and denoised image using an existing deep neural network. As a consequence, this deep learning-based solution outperformed other low-light image improvement methods.

II. PRINCIPLES OF VIDEO PROCESSING

1. **Digital -Video Formats :** A video sequence is made up of a series of images. An image is a 2 Dimensional signal taken at a particular time. The sequence of images in a video changes with respect to time. Number of images displayed in one unit of time is called video display rate which is normally 25 or 30 Frames Per Second (FPS). In other words, a video sequence with one second duration will consist of an around of 30 frames(one minute-1800 frames) and the contents in a video clip of one second duration is very less. The structure of a general video is shown in Figure 1.2. A video sequence may be seen as a well-organized document that can be broken down into logical parts at four distinct granularity levels: (a) Frame level (b)Shot level (c) Scene level (d) Sequence/story level. Each video is a collection of distinct scenes that are connected by a meaningful link. i.e., Scene 2 should take place after Scene 1.
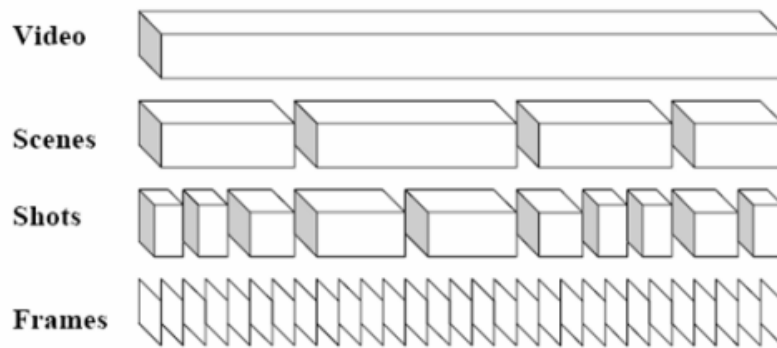
Figure 1.2: The Structure of a video

A shot is a collection of frames captured in a single camera operation, whereas a scene is a collection of several shots. Different shots, each of varying lengths, may be used to create a semantically significant scene. The video signal's properties are as follows:

- Frame size: It represents the total amount of pixels in a frame ie., row * column.
- The Aspect Ratio: It depicts the ratio of width to height
- Frame rate: It indicates the number of frames required to display the video in a unit of time.
- Bitrate: This is the number of bits used to indicate a video or audio data. In general, the higher the bitrate, the better the quality.
- The audio sample rate: It represents the frequency at which the audio is sampled when converted from an analog source to a digital file.

2. **Types of Video Formats:** Video and audio data are normally stored in compressed format, which takes less space than raw video. Compression schemes of a video are represented by codec; lossless codec preserves the original data with perfect accuracy. But lossy codec systems may degrade a signal. Least identifiable data signals are thrown away to do lossy compression. Some of most common video formats are written below

- **Motion Pictures Experts Group (MPEG):** It integrates audio and video compression. MPEG-1 formats are used in Video Compact Disc (VCD), MPEG-2 formats are used in Digital Video Disc(DVD) and MPEG-4(MP4)formats are used for broadcasting.
- **Audio Video Interleave (AVI): It** offers less compression than MPEG. It fixes the aspect ratio which cannot be changed manually. Now AVI was replaced by Microsofts Windows Media Video (WMV)
- **QUICKTIME:** It supports wide range of codecs. It started its journey in 1991. Now it is available for Mac and Windows. QuickTime was a sophisticated programme that could handle any number of streams and multi segmented files.
- **Digital Video Express(DivX):** DivX format compress video with less quality loss. It is recognized by DVD players and not by VCD players.
- **FlashVideo: It** is used for sending video through internet, but it is not supported in iPhones or iPads.

III. LOW LIGHT VIDEO

1. **Low Light Video Characteristics:** The main obstacle for capturing the scenery in a low light condition is the lack of visibility of the subject in the video. To increase visibility the amount of light captured by a camera must be increased. The most important parameters that need to be considered to increase the lighting in the video are ISO sensitivity, shutter speed and aperture settings(Peterson, 2008). These three essentials determine the final exposure of the input video which in turn decides the visibility of it. The ISO sensitivity of an image sensor is an algorithmic value that indicates how sensitive it is to light. On increasing the lighting in the video with the help of these elements there will also be an increase in amount of light that the camera captures. As a result, noise levels rise. Low dynamic range and excessive noise are two of the most important characteristics of a low-light video. It is necessary to examine the different noise sources in order to reduce the quantity of noise in a low-light video.Fig 1.3 shows low light and high quality image fromLOw-Light (LOL) dataset Wei et al. [2018].
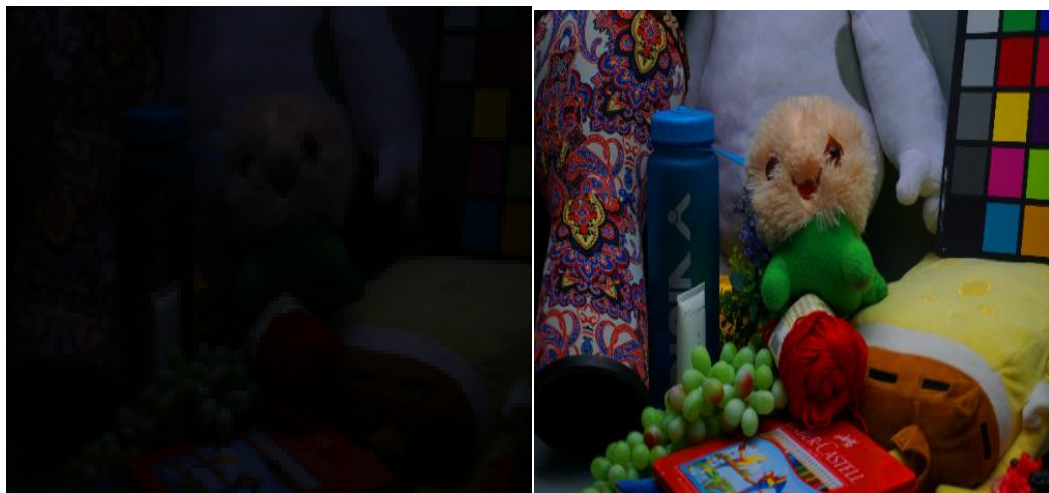

Figure 1.3: Low light image and High quality image

2. **Understanding ISO:** ISO stands for the sensitivity of your camera's sensor to light. As a result, it has an impact on the image's exposure, or what percentage of light is captured via the sensor. Depending on the scenario, change the ISO level to make the sensor more or less sensitive to light.

ISO is usually measured on a scale ranging from 100 to 1600, or even higher, on most cameras. The preceding value on the scale is doubled by two for each value on the scale. This is just an example of a typical range: 100, 200, 400, 800, 1600, 3200, and 6400.

The camera's sensor is twice as sensitive to light at ISO 200 as it is at ISO 100. ISO 800 is half as light-sensitive as ISO 1600.
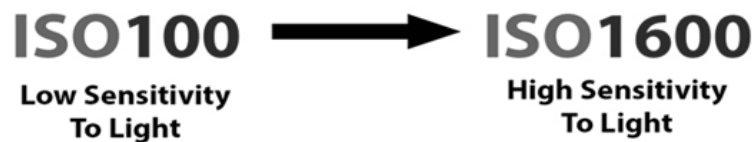
Figure 1.5: ISO Sensitivity

Low ISO (for example, ISO100 or ISO200) indicates poor sensitivity to light. This is just what is required in bright settings to avoid overexposed images.

A high ISO rating (800, 1600, or more) indicates that the camera is sensitive to light. This comes particularly handy in low-light situations where the camera needs to capture higher light to generate a better exposed image.

This is especially true when alternative choices for capturing more light aren't available, such as a wider aperture or slower shutter speed.

ISO is among the three primaryfeatures that affect exposure on a digital camera and are part of the exposure triangle. Aperture and Shutter Speed are the other two.

3. **Signal to Noise Ratio:** The Signal to Noise Ratio (SNR) is a critical measure for describing the overall performance of an imager. The SNR of a Complementary Metal Oxide Semiconductor (CMOS) imager is defined as the signal-to-noise ratio at a certain input light intensity. It may be stated in the following way:

$$SNR = 20log\left(\frac{QE \cdot P}{\sqrt{n_{SN}^2 + n_{DN}^2 + n_{RN}^2 + n_{FPN}^2}}\right)[DB] \tag{1.1}$$

The Quantum Efficiency of the pixels is denoted by QE, the number of input photons on the pixel region is denoted by P,photon shot noise is represented by n_{SN}^2, n_{DN}^2 signify the dark current shot noise, n_{RN}^2 represent the read noise, and n_{FPN}^2 indicatesFixed Pattern Noise (FPN). The first three noise terms indicate temporal

(random) noise sources, whereas the final one is the spatial noise source of CMOS imagers. Every imaging system and imager aspires to have the greatest SNR possible. If the SNR is more than 20dB, the image is considered to be of excellent quality.

4. **Dynamic Range:** A system's Dynamic Range (DR) is the ratio of the strongest signal to the lowest signal it can detect. In the dark, a CMOS imager's dynamic range is defined as the ratio of signal saturation to noise floor. It assesses an imager's ability to capture both bright highlights and dark shadows in a single image.In common, it may be stated as follows:

$$DR = 20 log \left(\frac{S_{sat}}{n_{DK}} \right) [DB] \qquad (1.2)$$

Where,S_{sat} denotes the signal level at saturation, and n_{DK} is the dark noise floor, which comprises pixel dark current shot noise and read noise. Equation (1.2), shows that there are two methods to enhance the DR: raising the greatest amount of signal an imager can process or lowering the dark noise level. It may therefore be deduced that the read noise level is an essential element for both the SNR and the DR of a CMOS imager in low-light situations.

5. **Low Light Issues:** Low-light difficulties can cause a variety of issues.

- **Video noise:**When the amplifier's gain is increased, the noise level rises.
- **Streaking:** Under low light conditions, a bright spot created by a candle, reflections, panning, or zooming causes streaks because the bright spot's after image stays on the image.
- **Blurred motion:**Reduced shutter speed to compensate for low illumination will result in a blurry image in the video. If there is any motion, such as camera movement, the same thing happens.
- **Poor focus:**In low-light circumstances, focusing can be difficult, and the camera's auto focus system may fail to accurately focus on the target.
- **Spotlight look:**When a video grapher uses an on-board spotlight to combat low light conditions, if no other light is available, it may be beneficial. If any other light is present, however, it will project circles of light with sharp edges, which may be unpleasant.

IV. MOTIVATION

There has been a significant increase in the sensitivity and resolution of digital cameras in recent years. Despite the fact that image sensor sensitivity has increased, current digital cameras still have limitations when it comes to capturing High Dynamic Range (HDR) image in low-light.Digital cameras, like the human eye, struggle to record extremely low light settings. Because more light is received by the images sensors, digital cameras perform better when shot at higher ISO settings and slower shutter speeds. In low-light situations, these are the settings to use. But, higher ISO levels usually result in higher noise and slower shutter speeds result in motion blur. As the lighting gets worse, the noise becomes more intense than the signal, making it difficult to reconstruct the image. As a result, improving the visual look of films shot in low-light settings has been an active study topic that might be useful in a variety of video processing applications. Video telephony has quickly established

itself as a significant technology, eclipsing the idea of a traditional phone conversation. With the mobile market expanding with respect to computers and laptops in the near future video telephony will be the main source of communication. With the introduction of high-definition video codecs and high-speed internet, video telephony became a practical technology with regular use. H.264 is one of the most widespread codec standard used for video telephony. As low-quality video devices such as webcams and mobile phones have become increasingly common for video conferencing, the demand for dependable video enhancement solutions to increase their output has grown. Noise picture construction and transmission frequently cause video sequences to become disorganized, especially in low-light environments. Low light videos have extremely low dynamic range, as a result the quality of videos in low light is limited. Image sequences captured in low light conditions have very low Signal to Noise ratio. As a result, it is necessary to enhance the low-light video quality. Similarly, enhancing low-light video is meant for several specified applications such as automated vehicles, video conference, security and surveillance industries, satellite videography, traffic management, digital photography.

V. SURVEILLANCE VIDEO

Surveillance videos contribute significantly to unstructured big data. Closed Circuit Television (CCTV) cameras are installed in any location where security is a priority.Manual surveillance seems tedious and time consuming. Security possiblycharacterized in a variety of ways depending on the situation, such as detecting theft, detecting violence, predicting the likelihood of an explosion, and so on. In crowded public areas, the phrase "security" refers to nearly any unusual occurrence. Because it includes group activities, detecting violence among them is tough. Due to many real-world restrictions, detecting anomalous or aberrant behaviour in a crowd video scene is extremely challenging. Security surveillance systems are used to provide secure and safe transportation, entertainment, and supported living in our daily lives.

Static Camera: In a visual security system, a fixed digital camera collects all dynamic scene information within a static surveillance area background scene. As a result, the surveillance video database for a whole day is made up of the surveillance area's daytime and nighttime frames.The dark and light variations of the monitored region are affected by scene information collected from a surveillance system during the night. While analyzing and tracking a security danger, the surveillance area's non-uniform lighting creates a problem.
Dynamic Camera: Moving camera of variable focal length.

VI. PROBLEM STATEMENT

The loss of visibility is the most significant challenge with videos taken in extremely low-light. This obstacles might perhaps be characterized into low dynamic range in addition to high amount of noise. There are various types of noise sources in a low-light video which include quantization noise, thermal noise, read out noise and photon shot noise. Straight away stretching the dynamic range of a low light video exhibits various undesirable artifacts such as noise amplification, intensity saturation and loss of resolution.So, suitable denoising technique is used as well as deeplearning method is used for enhancing videos in extremely low-light.

VII. CONCLUSION

A main context about "Low Light Video Processing for Engineers: An Overview", equips engineers with the knowledge and tools to meet the challenges of low-light conditions head-on, inspiring innovation and progress in the dynamic world of video processing.

REFERENCES

[1] Bassiou N. and Kotropoulos C., (2007), Color Image Histogram Equalization by Absolute DiscountingBack-off,*ComputerVisionandImageUnderstanding*,107(1-2),108–122.

[2] Abdullah-Al-Wadud M., Kabir M., Dewan M., and Chae O., (2007), A Dynamic Histogram Equalization for Image Contrast Enhancement, *IEEE Transactions on Consumer Electronics*, 53(2),593–600.

[3] BanikP.,SahaR.,andKimK.,ContrastEnhancementofLow-lightImageusingHistogram Equalization and Illumination Adjustment, *In International Conference on Electronics, Information,andCommunication(ICEIC),24-27Jan,Honolulu,HI,USA*.2018,1–4.

[4] Land E., (1977), The Retinex Theory of Color Constancy, *Scientific American*, 237(6), 108–129.

[5] Jobson D., Rahman Z., and Woodell G., (1997*a*), A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Cenes, *IEEE Transactions on Image Processing*, 6(7),965–976.

[6] Gonzalez T., ImageNet Classification with Deep Convolutional Neural Networks, *In HandbookofApproximationAlgorithmsandMetaheuristics*,2007,1–1432.

[7] Lore K., Akintayo A., and Sarkar S., (2017*a*), LLNet: A Deep Autoencoder Approach to NaturalLow-lightImageEnhancement,*PatternRecognition*,61(6),650–662.

[8] Peterson B., *Understanding Shutter Speed: Creative Action and Low-Light Photography Beyond 1/125 Second*. Amphoto,2008.

[9] WeiC.,WangW., Yang W.,andLiuJ.(2018),DeepRetinexDecompositionforLow-light Enhancement. ArXiv preprintarXiv:1808.04560.

[10] Anitha, C., & Kumar, R. M. S. (2019). Extremely low light video enhancement along with no-reference video quality measurements. International Journal of Advanced Trends in Computer Science and Engineering, Trends in Computer Science and Engineering, 8(5).