Dark Feature Enhancement [provisional]

Student Name: George Musker

Supervisor Name: Dr. George-Alex Koulieris

November 9, 2019

I INTRODUCTION

This paper will outline and review similar work in the field of video and image enhancement emphasizing each component of the pipeline.

The displays that we view content are always improving, currently there is a large interest in the use of high dynamic range (HDR) displays. These displays can express much more detail and contrast in images and videos. As well as this, there is a large focus on improving capture technology, whether it is to improve capture explicitly - namely capturing HDR content from home - or capturing superior footage in difficult conditions. Our project aims to restore contrast and detail to low light, underexposed video.

II LITERATURE REVIEW

A *Tone Mapping*

This section outlines work using tone mapping and its inverse for increasing the dynamic range of image and video. For a more in-depth analysis, (Banterle et al. 2009) created a survey of all work before 2010.

The problem of expanding LDR content to HDR content is well established by (Banterle et al. 2006). Their project introduces the problem of inverse tone-mapping for increasing the dynamic range of LDR content. The goal here is to tackle the large number of legacy videos and images that have a LDR whether this is from the nature of capture, or the equipment used to acquire it. Although the work outlined here is separate to our project it is still of note, as low light video is effectively video with low dynamic range. In the paper, they seek to increase the dynamic range of images by locating the over saturated areas of an image and expanding the colour depth at these locations. This is done by using a median cut algorithm to locate areas of high luminance and then applying an expand map. The paper does apply this algorithm to video, but given the nature of the expand map, the algorithm is not temporally coherent resulting in flickering artefacts. However, in (Rempel et al. 2007), a temporally coherent algorithm is designed. Based off of the design of (Banterle et al. 2006), they use a spatial bilateral filter to reduce noise then apply a brightness enhancement function. The brightness enhancement function increases depth in saturated areas but maintains clear edge contrast by using flood fill edge stopping function. In this paper the focus is to create a real-time, automatic converter.

(Banterle et al. 2010) address the issue of temporal coherence on LDR video with their work in this paper. Where the previous design suffered from the expand map being a frame by frame algorithm, here a temporal expand map. The median cut algorithm previously mentioned is expanded to use time as a third dimension. However, temporal density approximation is costly and therefore the solution cannot run in real-time a problem which we hope to address.

A semi-automatic system is proposed by (Didyk et al. 2008). Again, the paper focuses on enhancing bright regions within the image. The videos are pre-processed to detect motion flow

and clipped regions. A classifier is then used on these regions to identify scene objects, which results in a system robust to pronounced noise and contouring problems. However, this system requires the manual entry of variables for tuning the system, something that an automatic system cannot have. Machine learning is also used in (Endo et al. 2017) and (Eilertsen et al. 2017) to create deep reverse tone mappers. Although our project does not wish to use a completely ML system, the use of machine learning components may be a useful way to improve computation times.

B Image Filters

Low-light video is inherently a noisy source, due to the low exposure and high camera gain. Most frameworks will use some form of filter to smooth noisy areas, however it is paramount to not lose any of the fine detail that is hidden within the video. This section will review methods for improving existing image filters.

A large number of papers have made use of some form of the bilateral filter. The bilateral filter uses a Guassian kernel to smooth the source image while preserving edges by accounting for difference in colour intensity and range differences. A number of papers have implemented more computationally efficient bilateral filters, which are especially relevant if the a low-light video is to be adjusted in real-time. (Adams et al. 2010) propose a method which "which tessellates high-dimensional space with uniform simplices". This reduces both the space and time complexities of such algorithms. Similarly, (Adams et al. 2009) propose a method of accelerating filters like the bilateral. Makes use of Monte-Carlo kd tree sampling algorithm. The algorithm takes a significant amount of time to build the tree however, a problem for real-time solutions. (Frosio et al. 2015) uses machine learning to create a bilateral filter, this filter outperforms other filters that require ruled-based tuning.

C Low Light video

This problem was tackled in (Bennett & McMillan 2005), where each pixel's exposure is adaptively varied, using a spatiotemporal filter. This filter is integrated temporally, expanding the dynamic range of the frame while reducing noise. The result is a temporally coherent video. Our project hopes to address real time low-light enhancement, a problem that isn't addressed in this paper. Similarly, the problem of enhancing low light video low light video is also done by (Malm et al. 2007). The algorithm shares a similar design with (Bennett & McMillan 2005), but rather than tone mapping with a logarithmic function they instead use a contrast limited histogram equalisation function. Their results showed a better noise reduction than (Bennett) but there was a clear loss of contrast, a shortcoming resultant of an automatic algorithm. The paper does not produce a real-time solution.

Work was also done on deblurring noisy/blurred pairs of images by (Yuan et al. 2007). Given long exposure leads to shaky captures and low exposure leads to dark and noisy captures, they seeked to combine each image into high qualtiy image. The model required a spatially invariant kernel and two images, a method inappropriate for automatic video enhancement with motion, however should be of note given the overall effort is to regain detail my integrating over similar frames - much like videos.

Real time work continued by (Kovaleski & Oliveira 2009). The algorithm uses bilateral filtering, improved by the use of a bilateral grid for faster computation. The result is fully automatic

and is implemented effectively onto GPUs.

(Zhang et al. 2012) propose an algorithm which uses luminance maps to acquire detail from low-light frames. The images are then de-hazed by inverting the luminance map, producing a higher contrast image albeit with noise. This noise is addressed by applying a joint-bilateral filter using the green channel to supply edge detail. This method would not work on videos as temporal coherence is not adddressed.

(Li et al. 2015) use a superpixel method to split image into patches. The paper found that the change in gradients at superpixels was large for noisy regions, assuming the noise is additive white Guassian noise, and the change was low in textured regions. Using this, the de-noising filter can be adaptively applied. Path based denoising is also done by (Kim et al. 2015). Here, they reduce termporal noise - the knaive solution would be to average along temporal dimension but this would leave ghosting artefacts on regions containing motion. Again, a spatio-temporal filter is used. The result is then tone-mapped with a clipping function using gamma correction. Uses non local means to handle remaining spatial noise.

References

Adams, A., Baek, J. & Davis, M. A. (2010), Fast high-dimensional filtering using the permuto-hedral lattice, *in* 'Computer Graphics Forum', Vol. 29, Wiley Online Library, pp. 753–762.

Adams, A., Gelfand, N., Dolson, J. & Levoy, M. (2009), 'Gaussian kd-trees for fast high-dimensional filtering', *ACM Trans. Graph.* **28**(3), 21:1–21:12.
URL: *http://doi.acm.org/10.1145/1531326.1531327*

Banterle, F., Debattista, K., Artusi, A., Pattanaik, S., Myszkowski, K., Ledda, P. & Chalmers, A. (2009), 'High dynamic range imaging and low dynamic range expansion for generating hdr content', *Computer Graphics Forum* **28**(8), 2343–2367.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2009.01541.x*

Banterle, F., Ledda, P., Debattista, K. & Chalmers, A. (2006), Inverse tone mapping, *in* 'Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia', GRAPHITE '06, ACM, New York, NY, USA, pp. 349–356.
URL: *http://doi.acm.org/10.1145/1174429.1174489*

Banterle, F., Ledda, P., Debattista, K. & Chalmers, A. (2010), Expanding low dynamic range videos for high dynamic range applications, *in* 'Proceedings of the 24th Spring Conference on Computer Graphics', SCCG '08, ACM, New York, NY, USA, pp. 33–41.
URL: *http://doi.acm.org/10.1145/1921264.1921275*

Bennett, E. P. & McMillan, L. (2005), 'Video enhancement using per-pixel virtual exposures', *ACM Trans. Graph.* **24**(3), 845–852.
URL: *http://doi.acm.org/10.1145/1073204.1073272*

Didyk, P., Mantiuk, R., Hein, M. & Seidel, H. (2008), 'Enhancement of bright video features for hdr displays', *Computer Graphics Forum* **27**(4), 1265–1274.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2008.01265.x*

Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R. K. & Unger, J. (2017), 'Hdr image reconstruction from a single exposure using deep cnns', *ACM Trans. Graph.* **36**(6), 178:1–178:15.
URL: *http://doi.acm.org/10.1145/3130800.3130816*

Endo, Y., Kanamori, Y. & Mitani, J. (2017), 'Deep reverse tone mapping.', *ACM Trans. Graph.* **36**(6), 177–1.

Frosio, I., Egiazarian, K. & Pulli, K. (2015), Machine learning for adaptive bilateral filtering, *in* 'Image Processing: Algorithms and Systems XIII', Vol. 9399, International Society for Optics and Photonics, p. 939908.

Kim, M., Park, D., Han, D. K. & Ko, H. (2015), 'A novel approach for denoising and enhancement of extremely low-light video', *IEEE Transactions on Consumer Electronics* **61**(1), 72–80.

Kovaleski, R. P. & Oliveira, M. M. (2009), 'High-quality brightness enhancement functions for real-time reverse tone mapping', *The Visual Computer* **25**(5), 539–547.
URL: *https://doi.org/10.1007/s00371-009-0327-3*

Li, L., Wang, R., Wang, W. & Gao, W. (2015), A low-light image enhancement method for both denoising and contrast enlarging, *in* '2015 IEEE International Conference on Image Processing (ICIP)', IEEE, pp. 3730–3734.

Malm, H., Oskarsson, M., Warrant, E., Clarberg, P., Hasselgren, J. & Lejdfors, C. (2007), Adaptive enhancement and noise reduction in very low light-level video, *in* '2007 IEEE 11th International Conference on Computer Vision', IEEE, pp. 1–8.

Rempel, A. G., Trentacoste, M., Seetzen, H., Young, H. D., Heidrich, W., Whitehead, L. & Ward, G. (2007), 'Ldr2hdr: On-the-fly reverse tone mapping of legacy video and photographs', *ACM Trans. Graph.* **26**(3).
URL: *http://doi.acm.org/10.1145/1276377.1276426*

Yuan, L., Sun, J., Quan, L. & Shum, H.-Y. (2007), 'Image deblurring with blurred/noisy image pairs', *ACM Trans. Graph.* **26**(3).
URL: *http://doi.acm.org/10.1145/1276377.1276379*

Zhang, X., Shen, P., Luo, L., Zhang, L. & Song, J. (2012), Enhancement and noise reduction of very low light level images, *in* 'Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)'.