JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Critical Analysis of Image Representation and Compression by using Segmentation Methods

Dr. Meenakshi Choudhary
University Computer Centre
Dr. Bhimrao Ambedkar University, Agra

Abstract

In contrast with existent approaches, which compute a single image with reduced range, close in a given sense to the original data, we propose to look for a representative set of images. The goal is then to produce a minimal set of images capturing the information all over the high dynamic range data, while at the same time preserving a natural appearance for each one of the images in the set. A specific algorithm that achieves this goal is presented and tested on natural and synthetic data.

1. Introduction

Numerous applications exist for high dynamic range images. One is computer graphics and the production of synthetic images with realistic or hyper-realistic appearance. Another application covers high dynamic range photographs, which are able to capture much more detailed scene information than standard photographs.

Recently, methods to acquire such photographs have been developed [7, 20, 16]. These systems make it possible to capture a highly detailed range representation of the scene and later process the data in order to select the image/s that better fulfils the given requirements.
These images could also improve computer vision and image analysis algorithms that usually rely on limited range data. This is particularly relevant in scenarios where we do not have complete control over the illumination, like medical applications for instance. Therefore, there is a need to develop algorithms to perform the translation from scenes to images, algorithms as the ones discussed and presented in this research paper.

We can classify the existing approaches for the translation from scene to image in two main groups. The first group consists of algorithms that map the original range to the output range while attempting to preserve the subjective perception of the scene, e.g., [5, 8]. Although this idea of "tone mapping" works quite well, it has some caveats. First, it is not able to reproduce all the details present in the scene. Second, the method breaks down when the input range is too wide (and uniform) compared with the available output range.

In the second group, we have algorithms that favor the visualization of details instead of the subjective perception of the scene. As examples of this approach, we have the works of Tumblin and Turk [34] and DiCarlo and Wandell[24]. Both apply a multi-scale decomposition to discriminate between illumination and details.

The main problem with this idea is that, although correct in theory, it usually introduces halos in the output image. The local mappings produced by these techniques violate the basic monotonicity principle, the pixel

value order is not necessarily preserved and darker (brighter) regions in the scene might become brighter (darker) in the image. Moreover, these approaches tend to have a large number of parameters, generally hard to control in an automatic fashion.

Our proposed paradigm attempts to have the best of both groups mentioned above. We propose a method that captures the details while preserving the natural appearance of the scene. As we will explain below, there is an intrinsic limitation in representing a high dynamic range image with only one standard output image.

Sometimes it is practically impossible to find an output image containing all the relevant information in the high dynamic range image that represents the scene. For this reason, we argue for a method to obtain a set of images containing all the relevant information of the original scene and displayed in a suitable way.

This set also includes a single satisfactory image that can be used if the specific application is limited to a single output. In general, different applications can select different sub-sets from the output of our algorithm. Our proposed approach is particularly useful when the high dynamic range image contains detailed information in regions with very different light levels, thereby requiring the use of different images to accurately visualize each region. This will be exemplified with our experiments.

Our paradigm is also of particular use when the display device effortless permits the use of multiple images. This is the case for example of electronic displays, and less the case of printed media.

2. Related Works

In [28] the authors introduce two methods to display high dynamic range images. The first one is intended to display synthetically generated images. The image is decomposed into layers of lighting and surface properties. The light layer, which contains most of the high contrast, is compressed and added back to the surface layers containing details and texture.

In this way, high contrast is reduced while preserving the details and texture from the original image. This method is designed to work only for synthetic images, for which the different layers are easily computed. (We will come back to this point later when reviewing the work of Tumblin and Turk [35].) For natural scenes, they proposed a locally adaptive method, denoted as the foveal display, which is inspired by eye movements.

The user selects a point of attention and the algorithm computes an output image with preserved contrast in the foveal region (a region around the selected point).

It is important to remark that this approach is dynamic in the sense that a set of images is generated with the aid of user interaction. The drawback with this approach is that the user needs to select a point of attention. Moreover, the user should be able to "see" everywhere in the image to choose the points of attention. Problems could then arise if the image presented to the user is not adequate for inspection. On the other hand, we should note that this approach is connected to ours in the sense that we also suggest to compute a set of images instead of a single one.
Similar to the work of Pattanaik et. al. we have the multi-scale generalization [9, 23] of the Retinex algorithm [7]. In this approach, to avoid the appearance of halos close to strong edges [43], is necessarily to finely tune the weighting of the different scales [13].

In terms of dynamic range compression, it performs well for moderate dynamic range compression but not for high dynamic range compression [13]. Continuing with the idea of segregating the image into layers of lighting and details [12],Tumblin and Turk [30] proposed a multi-scale approach to extract a hierarchy of details and boundaries.

Their idea is to mimic the way artists work from coarse to fine when recreating highly contrasted scenes in low contrasted mediums. Artists start with a sketch of strong features and progressively add small details. They propose to decompose the image into strong and weak features using a multi-scale operator, then, only strong features are compressed. Although the idea is very attractive, their algorithm cannot avoid halos completely and, as pointed out by the authors, needs the difficult tuning of some crucial parameters.

In our view of the problem, the two groups of works just described are not completely equivalent; they basically address slightly different problems.

Furthermore, although similar, their solutions cannot be easily compared, since one representation tries to capture the global subjective appearance of the scene under the limitation of the output/display device, while the second group attempts to visualize the scene details (note that details may be difficult to see even in front of the scene).

While enhancing details, we may be adding information not perceptually present in the original scene. On the other hand, while preserving visual appearance, details might be omitted.

The approaches described above produce a single image per scene or per focused scene region. It is quite an optimistic assumption that we can accurately represent and visualize the information of an image with tens of thousands of values, and with details in regions with very different luminance, with just one image and with limited dynamic range.

This is what motivates the paradigm here proposed, meaning the use of a set of images to represent such highly detailed information. We could say that while the algorithms described above deal with the reproduction of the scene, the technique here proposed deals with its visualization.

Moreover, we argue that not only the set of images has to accurately visualize the relevant information present in the scene, but also has to do it in a visually pleasant form. We present a particular algorithm to exemplify this new paradigm. In the proposed algorithm, each image is produced by a different monotonic global map, thereby avoiding gradient reversals typical of the local schemes.

3. Analysis of Method

Let us assume we have a scene with dark and bright areas, and details all over it. If we want to visualize all the details, we need first to have a minimum dynamic range available, and second to be able to “see” in every region. In order to understand what is meant by “see,” we present a couple of simple examples.

Consider a dark room, in order to “see” the details we would usually turn on the lights. Now, suppose we are in the beach and everything is too bright, in this case to “see” we would probably wear sunglasses to reducethe amount of light. In both cases, the information is out there but it cannot be seen. Inphotographic terms, this is due to under-exposure or over-exposure. All the proposed operations are simply contrast changes.In this way, artefacts such as halos are not introduced.

The first observation is that if we modify a region’s luminance, expanding its range ofvariation, it will be easier to perceive details in it. On one hand we display this region with“good illumination,” while on the other hand we might be compressing and missing detailsin other regions.

In Figure 1.1 we show a piecewise linear mapping which illustrates ouridea. Hence, there is clearly a competition between the output fidelity of different regions,and unfortunately, it is difficult or impossible to

find a satisfactory solution with a singleoutput. To overcome this, we propose to generate a sequence of images with differentresolutions in each scene region (space varying resolution).

This sequence could be eitherobserved as a movie or as a set of still images could be extracted from it. The key pointhere is that for many applications more than one output image is a reasonable solution.

The basic idea is then to distribute the existing resources among different output images.In the case of only two regions, we first display dark regions expanding its luminancerange, and then we slowly swap resources to bright areas.

4. Proposed Algorithm

We are now ready to describe the different steps of the algorithm.

1. Compute scene luminance: From the (r, g, b) primaries compute the luminance L (in cd/m2) and the color information (r/L, g/L, b/L). We process the luminance while preserving the color information.

2. Segment the scene: Divide the scene into two or more regions of interest. This is achieved splitting the histogram into sub-intervals [Lwmin, L1], [L2, L3], ..., [Ln, Lwmax].

If we segment the scene arbitrary, we could lose the monotonic restriction of the tone mapping. We then choose a simple procedure that segments the scene into bright and dark areas to illustrate the idea, see section 1.3.1. From now on then, in this section, we assume two regions, [Lwmin, Lw] and [Lw, Lwmax].

3. Modify the luminance: Apply Larson's et al. histogram adjustment algorithm, [8], to each interval. Map [Lwmin, Lw] to [Ldmin, Ld] and [Lw, Lwmax] to [Ld, Ldmax].

This algorithm computes the output brightness Bd with the formula

$$Bd = \log(Ldmin) + [\log(Ldmax) - \log(Ldmin)] \hat{} \; P(Bw)$$

This is done constraining the slope of the operator using an estimation of the just noticeable differences (JND) for a given adaptation luminance level. That is, the slope of the operator is constrained to the ratio of the just noticeable differences at world and display levels ($dLd/dLw \leq JND(Ld)/JND(Lw)$).

5. Experimental Results

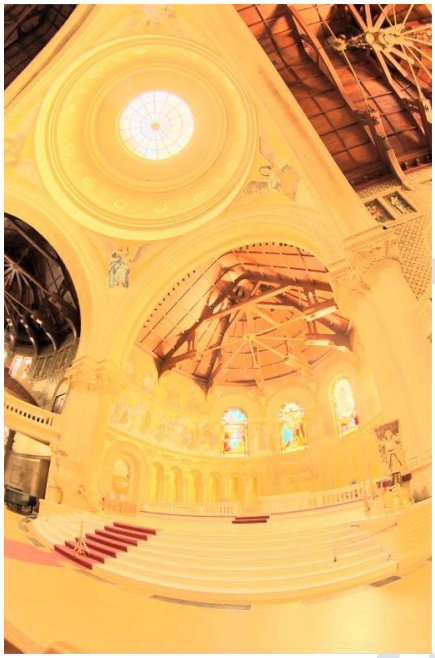
To begin with, the first regions have very small luminance values. This means that is very difficult, even being in front of the real scene, to perceive details there. The dynamic range for this region is extremely small, representing only the $5.32 \times 10{-}3$ % of the total input range.

In other words, although region 0 covers a significant part of the scene, it expands an insignificant range from the total. Once again, it is clear that high dynamic range data it is not just a problem of output range, it is in fact a problem of resolution within regions in the scene.

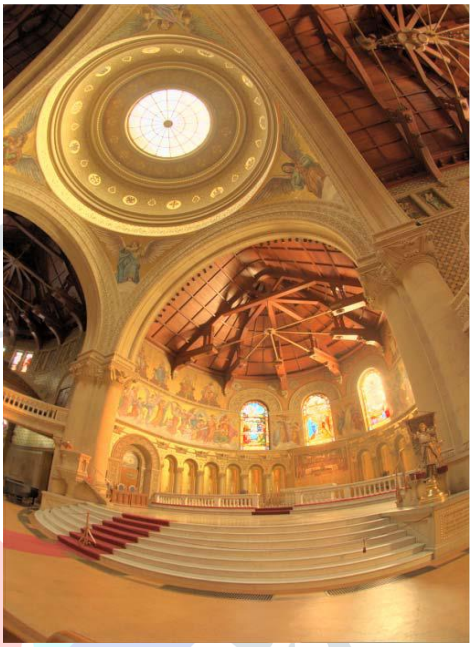
In Figure 1.3 we show the set of images obtained with our algorithm. Each image represents the image with maximal entropy per region. With these images we can see more details than in the single image processed with Larson's algorithm (see Figure 1.3d and also Figure 1.4).

The resulting images in Figure 1.3 are fairly natural, specially for the last two images. The three images selected visualize more information, especially in very dark and very bright regions. Hence, we managed to display the original scene in aset of images, making visible some information that was obscured in the original scene.

Note for example the details in the dark corners and in the bright windows. In the first image observe the upper left and upper right corners; the details in the ceiling are clearly visible now. In the second image observe the walls. Finally, the third image represents the windows.

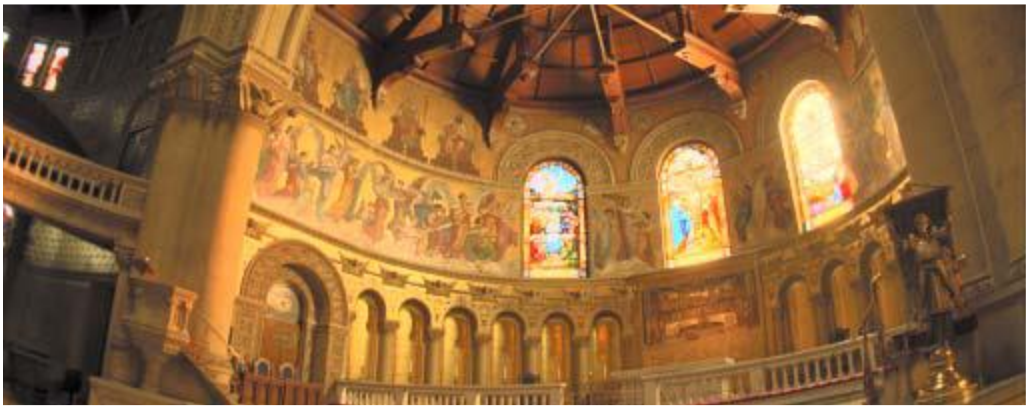


(Fig 1.1 : Image with entropy level 0)



(Fig 1.2: image with entropy level 1)

(Fig 1.3 : Zoomed Image)



(Fig 1.4 : Zoomed Image with low illumination)

(Fig 1.5 : Zoomed Image with high illumination)

6. Conclusion

We have presented a new paradigm for the reproduction and visualization of information in high dynamic range images. We argued for the use of a set of images instead of a single one as in traditional approaches.

This is particularly important when the image contains details in regions with very different illumination. The set of produced images also includes a single image that is satisfactory for applications limited to a single output.More than being the last word about the problem of visualizing high dynamic range data, with this work we attempted to illustrates the intrinsic limitation of working with only one image.

To select a single image from the set, we consider the image that maximizes theglobal entropy. If we compare it with the image computed with Larson's algorithm, seeFigure 1.4, we can see that it contains enough details everywhere. This is due to themodification of the luminance distribution; we expanded regions that were originally verycompressed. However, it is clear from the zoomed images, Figure 1.5 and Figure 1.6, thatit does not capture dark and bright details as well as the images with maximum resourcesper region.

7. References

1. A. Basso, I. Dalgic, F. Tobagi, and C. van den Branden Lambrecht. Study of MPEG-Coding Performance bsed on a Perceptual Quality Metric. In *Proc. Picture CodingSymposium*, pages 263–268, 1996.
2. T. S. Bauer and B. Makai. Efficiency of a Shape Adaptive 2d Transform for Codingof Arbitrarily Shape Image Segments. *IEEE Transactions on Circuits and Systemsfor Video Technology*, 5(3):254–258, June 1995.
3. M. Bi, S.-H. Ong, and Y.-H. Ang. A Hybrid Shape-Adaptive Orthogonal Transformfor Coding of Image Segments. *IEEE Transactions on Circuits and Systems forVideo Technology*, 10(8):1471–1477, December 2000.
4. J. Canny. A computational Approach to Edge Detection. *IEEE Transactions onPattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
5. J. Casas and L. Torres. Coding of Details in Very Low Bit-Rate Video Systems.*IEEE Transactions on Circuits and Systems for Video Technology*, 4(3):317–327,June 1994.
6. J. Casas and L. Torres. A Feature-based Subband Coding Scheme. In *IEEE InternationalConference on Acoustics, Speech and Signal Processing*, volume 4, pages2355–2358, May 1996.
7. V. Caselles, B. Coll, and J.-M. Morel. Scale Space Versus Topographic Map forNatural Images. In *Scale-Space Theory in Computer Vision. First Conference onScale Space 1997*, volume 1252 of *Lecture Notes in Computer Science*, pages 29–49.Springer, 1997.
8. V. Caselles, B. Coll, and J.-M. Morel. Topographic Maps and Local ContrastChanges in Natural Images. *International Journal of Computer Vision*, 33(1):5–27, September 1999.
9. T. Chan and J. Shen. Variational restoration of non-flat image features: Models andalgorithms. *SIAM Journal on Apllied Mathematics*, 61(4):1338–1361, 2000.
10. S.-Y. Chien, S.-Y. Ma, and L.-G. Chen. Efficient Moving Object Segmentation AlgorithmUsing Background Registration Technique. *IEEE Transactions on Circuitsand Systems for Video Technology*, 12(7):577–586, July 2002.
11. S. H. Chon, R. C. Kim, S. S. Ho, and S. U. Lee. A Coding Technique for the contoursis Smoothly Perfect Eight-Connectivity Based on Two-Stage Motion Compensation.*IEEE Transactions on Circuits and Systems for Video Technology*, 9(5):59–69,February 1999.
12. M. Crouse and K. Ramchandran. Joint thresholding and quantizer selection fortransform image coding: Entropy-constrained analysis and applications to jpeg. *IEEETransactions on Image Processing*, 6(2):285–297, February 1997.
13. P. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps fromPhotographs. In *SIGGRAPH 97 Conference Proceedings*, pages 369–378, 1997.
14. R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edgedetector. *International Journal on Computer Vision*, pages 167–187, 1987.
15. A. Kaup. Object-Based Texture Coding of Moving Video in MPEG-4. *IEEE Transactionson Circuits and Systems for Video Technology*, 9(1):5–15, February 1999.
16. A. Kaup and T. Aach. Coding of Segmented Images Using Shape Independent BasisFunctions. *IEEE Transactions on Image Processing*, 7(7):937–947, July 1998.
17. S. Khan and M. Shah. Object Based Segmentation of Video using Color, Motionand Spatial Information. In *CVPR2001 - Int. Conf. Computer Vision and PatternRecognition*, volume 2, pages 746–751, 2001.
18. C. Kim and J.-N. Hwang. Fast and Automatic Video Object Segmentation andTracking for Content-Based Applications. *IEEE Transactions on Circuits and Systemsfor Video Technology*, 12(2):122–129, February 2002.
19. M. Kim, J. G. Choi, D. Kim, H. Lee, M. H. Lee, C. Ahn, and Y.-S. Ho. A VOPGeneration Tool: Automatic Segmentation of Moving Objects in Image SequencesBAsed on Spatio-Temporal Information. *IEEE Transactions on Circuits and Systemsfor Video Technology*, 9(8):1216–1226, December 1999.
20. W. Kohler. *Gestalt Psychology*. Liveright Publishing Corporation, 1947.

21. M. Kunt, A. Ikonomopoulos, and M. Kocher. Second Generation Image CodingTechniques. *Proceedings of the IEEE*, 73(4):549–574, April 1985.
22. E. Land and J. McCann. Lightness and retinex theory. *Journal of the Optical Societyof America*, 63(1):1–11, 1971.
23. G. W. Larson, H. Rushmeier, and C. Piatko. A Visibility Matching Tone ReproductionOperator for High Dynamic Range Scenes. *IEEE Transactions on Visualizationand Computer Graphics*, 3(4):291–306, October-December 1997.
24. A. Pardo and G. Sapiro. Vector Probability Diffusion. *IEEE Signal ProcessingLetters*, 8(4):106–109, April 2001.
25. A. Pardo and G. Sapiro. Visualization of High Dynamic Range Images. In *InternationalConference on Image Processing ICIP2002*, volume 1, pages 633–636,2002.
26. A. Pardo and G. Sapiro. Visualization of High Dynamic Range Images. *IEEETransactions on Image Processing*, 12(6):639–647, June 2003.
27. H. S. Park and J. B. Ra. Homogenous Region Merging Approach for Image SegmentationPreseving Semantic Object Contours. In *Proceedings of VLBV'98*, pages149–152, 1998.
28. S. Pattanaik, J. Ferwerda, M. Fairchild, and D. Greenberg. A Multiscale Modelof Adaptation and Spatial Vision for Realistic Image Display. In *SIGGRAPH 98Conference Proceedings*, pages 287–298, 1998.
29. R. Stasinski and J. Konrad. A New Class of Fast Shape-Adaptive Orthogonal Transforms and Their Application to Region-Based Image Compression. *IEEE Transactionson Circuits and Systems for Video Technology*, 9(1):16–34, February 1999.
30. B. Tang, G. Sapiro, and V. Caselles. Diffusion of general data on nonflat manifolds via harmonic maps theory: The direction diffusion case. *International Journal ofComputer Vision*, 36:149–161, February 2000.
31. P. Teo, G. Sapiro, and B. Wandell. Creating Connected Representations of Cortical Gray Matter for functional MRI Visualization. *IEEE Transactions on MedicalImaging*, 16(6):852–863, December 1997.
32. T. Tran and R. Safranek. A locally adaptive perceptual masking threshold model for image coding. In *International Conference on Acoustic Speech and Signal Processing*, volume 4, pages 1882–1886, 1996.
33. Y. Tsaig and A. Averbuch. Automatic Segmentation of Moving Objects in Video Sequences: A Region Labelling Approach. *IEEE Transactions on Circuits and Systemsfor Video Technology*, 12(5):597–612, 2002.
34. J. Tumblin, J. Hodgins, and B. Guenter. Two Methods for Display of High Contrasted Images. *ACM Transactions on Graphics*, 18(1):56–94, January 1999.
35. J. Tumblin and H. Rushmeier. Tone Reproduction for Realistic Images. *IEEEComputer Graphics and Applications*, 13(6):42–48, November 1993.
36. J. Tumblin and G. Turk. LCIS: A Boundary Hierarchy for Detail-Preserving Contrast Reduction. In *SIGGRAPH 99 Conference Proceedings*, 1999.
37. B. Usevitch. A Tutorial on Modern Lossy Wavelet Image Compression: Foundations of JPEG2000. *IEEE Signal Processing Magazine*, 18(5):22–35, September 2001.
38. R. L. D. Valois and K. K. D. Valois. *Spatial Vision*. Oxford University Press, 1988.
39. C. J. van den Branden Lambrecht. *Perceptual Models and Architectures for Video Coding Applications*. PhD thesis, ´EcolePolytechniqueF´de´erale de Lausanne, 1996.
40. O. Verscheure, A. BAsso, M. El-Maliki, and J. Hubaux. Perceptual bit allocation for mpeg-2 cbr video coding. In *International Conference on Image Processing*, pages 117–120, 1996.
41. G. Wallace. The JPEG still picture compression standard. *Communications of the ACM*, 34(4):30–44, April 1991.