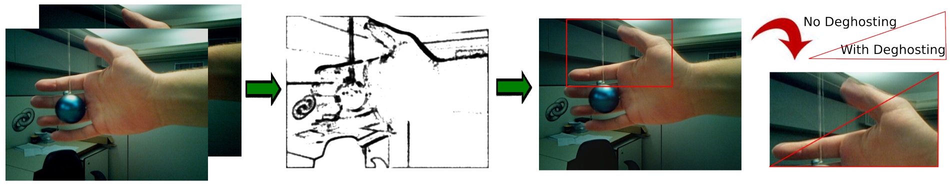
Filter Based Deghosting for Exposure Fusion Video

Alexandre Chapiro   Marcelo Cicconet   Luiz Velho
Visgraf laboratory, Institute of Pure and Applied Mathematics, IMPA, Brazil.



(a) (b) (c) (d)

Figure 1: *Video Creation Pipeline: (a) Inputs, (b) Ghosting Coefficients (after High-Pass filtering) (c) Deghosted result (d) Comparison*

Introduction and Related Work This work deals with a well known problem - the fact that consumer cameras are unable to capture the whole range of color luminance variations the human eye can perceive. Many techniques deal with this, the most widespread probably being High Dynamic Range Imaging. These works focus mainly on still images. Our focus in this work, however, is to improve videos taken from mobile devices by extending one such technique, Exposure Fusion, introduced in [Mertens et al.].

Exposure Fusion is a very efficient procedure that produces results similar to those of tone-mapped HDR images through a clever per-pixel averaging process. This process, however, was not designed for video processing. Attempts to apply it in such a way result in undesirable artifacts, or "ghosts" during the combination of multiple differently exposed images containing moving objects. Although several deghosting methods exist, most are inefficient when applied to this situation. For instance, Optical Flow does not work well with changing exposures in the input frames.

In this work we present a novel method that deals with the elimination of such artifacts by using several carefully selected filters and performing a local analysis. In order to tackle the variations brought by the changing exposure times between frames, we used features that are known to stay reasonably constant in this case - edges, saliences and textures as detected by a High-Pass filter. This method was inspired by our past work dealing with HDR video on mobile devices [Castro et al.] and a deghosting method for HDR that involved the use of pixel regions as an estimator [Gallo et al.].

Deghosting Method Exposure Fusion relies on three numerical parameters assigned to each pixel in the inbound images - Well-Exposedness, Detail and Saturation. We propose a fourth parameter - Ghosting. This parameter will define the likelihood of presence of object movement on the scene at the current pixel. Given a set of images $I = I_1, ..., In$, we wish to construct another set of n improved images (in this work, $n = 2$). We begin by analyzing I_1 and I_2. In order to obtain a better reference of possible object movement, we analyze pixel regions instead of single pixels.

The original images' pixel color variations were found too steep due to the implicit exposure variation to generate reliable results. Because of this, the process outlined below is applied to the result of a regular High-Pass Laplacian filter applied to each image, described as a function $HP(I_k)$. Thus to find the Ghosting parameter of pixel (i, j), $G(i, j)$ we analyze the regions A and B as given by $(i - l : i + l, i - l : i + l)$ in each image.
The process is then repeated for each pixel using a different input. The initial images I_1, I_2 are now subject to a Low-Pass Gaussian filter, given by $LP(I_k)$. Following, we take $HP(LP(I_k))$ as our inputs and proceed to analyze their pixel regions. This method is then repeated with additional Low-Pass filter steps. The pixel areas are evaluated according to the following formula: $G(i, j)_i = 1 - \sum_{n=1}^{2l} \sum_{n=1}^{2l} |A - B|_{n,m}$.
The resulting obtained Ghosting coefficients are multiplied to obtain the final pixel Ghosting parameter value. This process attenuates the contribution of capture noise and irrelevant weaker high-frequencies, which disappear after consecutive Low-Pass applications, resulting in less erroneous detections of non-movement high-frequency variations and a strengthened Ghosting parameter for pixels that involve true object movement.

Finally, the ghosting parameter is used in the same way as the three initially proposed coefficients and allows us to reduce the contribution of pixels from I_2 that show moving objects in relation to I_1.
This process is then repeated for consecutive pairs of images I_k, I_{k+1}, $k = 1, ..., n - 1$ to form the final video. The whole process is outlined in Fig.1. More information is available at w3.impa.br/ achapiro/deghost, where a video is also presented.

Obtention and Registration of Input Images In order to obtain the video frames used in this work, a Nokia N900 running Maemo 5 and the FCam API was used. We proceed to perform a multi-resolution alignment based on image pyramids. This step is necessary, as background pixel correspondence is crucial for this work. Both are explained in detail in our previous work [Castro et al.].

Results and Future Work Testing of the algorithm showed good results when applied to regular situations, removing most of the ghosting artifacts and resulting in improved video quality. Some issues arise with quick movement relative to the camera's capture rate, where sometimes ghosting is not properly treated.

Future work may include improvements to the algorithm's robustness to quick object and camera movement as well as improvements of the filter-based deghosting technique through the use of multi-resolution to aid in the location of shifting objects.

References

CASTRO, T., CHAPIRO, A., CICCONET, M., AND VELHO, L. Towards mobile hdr video. *Eurographics 2011*.

GALLO, O., CHEN, W., GELFAND, N., TICO, M., AND PULLI, K. Artifact-free high dynamic range imaging. *ICIP 2009*.

MERTENS, T., KAUTZ, J., AND VAN REETH, F. Exposure fusion. *Pacific Graphics 2007*.

SIGGRAPH 2011, Vancouver, British Columbia, Canada, August 7 – 11, 2011.
ISBN 978-1-4503-0921-9/11/0008