Journal of Real-Time Image Processing
https://doi.org/10.1007/s11554-020-01001-x

ORIGINAL RESEARCH PAPER

De-ghosted HDR video acquisition for embedded systems

Ghost-free HDR video of motion objects from stationary cameras

Martin Musil[1] · **Svetozar Nosko**[1] · **Pavel Zemcik**[1]

Received: 19 February 2020 / Accepted: 7 July 2020

Abstract
This paper proposes a novel ghost-free High Dynamic Range (HDR) multi-exposure video acquisition suitable for real-time implementation in embedded systems. While the method is limited to stationary cameras, it achieves, with low requirements on resources, results comparable to state-of-the-art de-ghosting methods that are often very computationally expensive and almost impossible to implement in smart cameras and embedded systems. The paper describes the method itself and includes an evaluation of the performance on selected embedded platforms and a comparison of the results to the state of the art using HDR datasets.

Keywords HDR · HDR Acquisition · Embedded systems · Real-time HDR processing · HDR De-ghosting

1 Introduction to HDR imaging

High Dynamic Range video acquisition is a very popular topic in photography and the film industry. Contemporarily, HDR is also becoming increasingly more established in surveillance, traffic monitoring, quality control in the industry, etc. in environments, where lighting conditions are challenging.

HDR videos are typically acquired through multi-exposure by sensors with a limited (standard) dynamic range [2, 14, 18], since this is both technologically and economically feasible as compared to other possible alternatives. Unfortunately, such videos typically suffer from "ghosts" caused by individual exposures of the motion objects taken at different times (and thus also capturing the objects in different positions) or by camera motion. Alternatives include theoretically ghost-free approaches, such as systems using beam-splitters with several CCD/CMOS sensors [25] or expensive and technologically demanding specific HDR sensors [19, 31].

In this paper, we propose a novel ghost-free HDR acquisition method for stationary cameras that is powerful yet well implementable even in embedded systems in real-time with low resource requirements. While de-ghosting has been researched for a long time, the state-of-the-art methods with good performance are computationally very demanding and so they are impossible to implement in smart cameras and/or embedded systems attached to cameras. The novel HDR de-ghosting method proposed in this paper was designed with respect to real-time processing in embedded hardware.

2 Related work

In this paper, we focused on embedded systems and real-time processing; therefore, we reviewed only simple, computationally unpretentious methods. We excluded optical flow and patch based algorithms due to their high computational demands. We concentrated on the methods categorised by Tursun [26], Srikantha [23] and other authors as "moving

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11554-020-01001-x) contains supplementary material, which is available to authorized users.

✉ Martin Musil
imusil@fit.vutbr.cz

Svetozar Nosko
inosko@fit.vutbr.cz

Pavel Zemcik
zemcik@fit.vutbr.cz

[1] Faculty of Information Technology, Centre of Excellence IT4Innovations, Brno University of Technology, Brno, Czech Republic

Published online: 30 July 2020

object selection" methods, and similar approaches. The proposed algorithm does not address global image registration and works with images from stationary cameras.

Debevec and Malik [2] proposed an algorithm serving as a baseline for HDR acquisition; however, this approach only addresses static scenes, without either camera or object motion. The algorithm fuses multiple images into a HDR radiance map, whose pixel values are proportional to the true radiance values in the scene. HDR image H is calculated as a weighted average of the exposures L:

$$H = \frac{\sum_{i=1}^{N} w(L_i) \cdot L_i \cdot \frac{L_i}{t_i}}{\sum_{i=1}^{N} w(L_i)} \tag{1}$$

where N is the number of exposures, t_i exposure times and function w the triangle weight function.

The algorithm devised by Gallo et al. [3] assumes a linear dependency between couples of pixels when they "see" the same radiance levels, based on knowledge of exposure times. Any image spot violating this linear relation is considered as containing a motion. Patches with large numbers of not corresponding pixels are omitted from merging, causing visible artifacts that occur at their boundaries. Raman et al. [17] extended the work of Gallo et al. [3] so that it does not require knowledge of exposure times or the CRF (Camera Response Function [2, 14, 18]), the response function of the camera sensor to the incident light.

Grosch [4] proposed a simple method based on pixel value estimation from the known exposure time and the CRF estimated by the histogram-based method proposed by Grossberg and Nayar [5]. The algorithm put forth by Wu et al. [29] estimates the CRF from regions where the RGB vectors remain fixed with respect to the exposure changes. The algorithm refines motion detection by combination of pixel order relation [21] and pixel value estimation [4]. The algorithm devised by Wang et al. [28] normalizes all images L_i according to the reference image L_{ref}. A ghost mask is obtained by adaptive thresholding of absolute differences of pixel intensities from each input image L_i; however, the threshold is expected to be tuned manually. Ghost mask is further refined by morphological operations.

The algorithm of Jacobs et al. [7] calculates pixel variance over exposures to detect motion. The variance is compared to a fixed threshold which results in a *variance image* (a *ghostmap*). The *variance image* is supplied by an *uncertainty image*, which is calculated using the local variance obtained from a histogram of a small window.

Min et al. [12] improved the method of Pece et al. [16] and introduced a multi-level threshold map, where thresholds are selected to divide the image into multiple regions according to the pixel ntensity, each region having the same number of pixels. Any difference between the threshold maps of input images and the reference image, presented typically by the mid-exposure, is marked as a motion-region. The algorithm depends on the scene composition and image histogram layout. Min et al. [13] further improved the algorithm by employing a noise reduction phase incorporating an additional set of rules for spatially neighbouring pixels. Due to using a large spatio-temporal filter, the algorithm is memory and performance demanding. The methods 12, 13, 16] are using coarse morphological operators, such as erosion and dilatation, to suppress false detection occuring on edges or by noise.

Bouderbane et al. [1] implemented a simple ghost removing algorithm on an FPGA based platform. They were inspired by the work of Sidibe et al. [21]. The algorithm is based on modification of the weighting functions used for the HDR merging.

3 Novel ghost-free HDR merging

In this paper, we propose a novel HDR merging method that produces ghost-free results. Our approach is based on pixel value matching, the idea being similar to the solutions proposed by Grosch [4], Wu [29], and Wang [28] but with quite different and improved processing. The exposure time of each image is known; therefore, it is possible to estimate and match pixel values in the adjacent images, except for the over or under-exposed patches where the pixel values will obviously not match. Such estimation is not very precise, the captured image data is affected by factors such as noise, sensor quantization errors, CRF, etc. The reviewed methods generally use fixed or user-guided thresholds which must be employed to introduce user-defined tolerance to these factors. These fixed or user-defined thresholds often cause adverse effects in the final HDR images, such as visible transitions between static and motion areas etc. We propose a method to overcome such problems.

3.1 Obtaining a certainty map

In our approach, every image L_i is assigned a *Certainty map* C_i related to the reference image L_{ref}, which is generally considered to be the middle (exposure) image in the sequence. The *Certainty map* C contains values representing the estimated level of certainty that the individual pixels contain the same patch of the scene as the reference pixel, but obtained under a different exposure. Unlike ghostmaps, *Certainty maps* hold not only the patches containing motion, but rather all patches inappropriate for merging—such as under and over-exposed pixels (Fig. 1).

The probability distribution of low level value pixels is Poisson [10] due to the discrete nature of the incoming photons. With higher intensities, the distribution transforms into

Fig. 1 Example of ghost artefacts and result of the proposed ghost free merging. Top left—stripes of original images with a significant car motion. Top middle and top right - Images representing coefficients used for the HDR merging (*certainty maps*, see Sect. 3). Bottom left—ghosted HDR image. Bottom right—HDR image merged using the proposed method

Fig. 2 Two *Certainty maps* (bottom) obtained from the sequence on the top. The *Certainty map* on the left was obtained from top left and top middle (reference) image, the *Certainty map* on the right was obtained from top middle (reference) and top right image

Normal (Gaussian). Therefore, we use the Gaussian function to derive the certainty (estimated probability) that the two luminance levels, estimated and measured, match. The *Certainty map* C_i (see Fig. 2) replaces the binary *ghostmap* with soft assigned values, obtained using the information from the reference image L_{ref}, the estimated image $\overline{L_i}$, the exposure times t_i and t_{ref}, as well as the CRF. Note, please, that in this paper the inverse CRF was implicitly applied to all images L_i. Image $\overline{L_i}$ is estimated by the following equation:

$$\overline{L_i} = L_{ref} \cdot \left(\frac{t_i}{t_{ref}} \right) \tag{2}$$

Consequently, the estimated value for image i is processed along with the actual value of L_i to get the probability based *Certainty map* C_i as:

$$C_i = e^{-\frac{\left(L_i - \overline{L_i}\right)^2}{2\sigma^2}} \tag{3}$$

where σ reflects the standard deviation of the pixel measurement (affecting the "softness" weight). The lower σ is, the sharper or more strict the *Certainty* maps are, which results mainly in the dynamic range reduction. On the other hand, a high σ causes "softer" *Certainty* maps, which may start to be ghosted. Ghost detection generally, and indeed inherently, cannot work well for the over and under-exposed spots of an image; thus the *Certainty* map algorithm contains a boundary condition: If the estimated value lies beyond the point of saturation, the Certainty is assigned at maximum value.

3.2 Multi-exposure merging algorithm

Our modification of Debevec's [2] merging algorithm incorporates the weights from the *Certainty map*, obtained through Eq. 3. The HDR image H is calculated as the weighted sum of pixels from n images using the following equation:

$$H = \frac{C_i \cdot w(L_i) \cdot L_i \cdot \frac{t_i}{t_{min}}}{\sum_{i=1}^{n} (C_i \cdot w(L_i))} \tag{4}$$

The C_i for reference image certainty is considered to be 1. The $w(L_{ref})$ is considered to be 1, as the reference image is a "pattern" with the desired object layout; it is not desirable to weight out the pixels, even if poorly exposed. A scheme illustrating the Eq. 3 is shown in Fig. 3.

4 Implementation

The proposed algorithm was designed with respect to real-time processing using embedded hardware—we are considering mainly FPGA based platforms and SoC equipped with GPU (e.g. NVIDIA Tegra). The standard desktop CPU implementation is included mainly for comparison.

The ghost-free merging unit consists of two components, *Certainty map* creation (Sect. 3.1) and HDR merging (Sect. 3.2). The *Certainty map* is obtained by predicting and matching the luminance levels, it is thus necessary to provide luminance images. The RAW data from sensors, e.g. in embedded devices/cameras, should, therefore, be converted to luminance because if the individual RGB channels are processed separately, their saturation, which is independent for each channel, may lead to certainty results different for individual colour channels and thus to adverse colour shifts during HDR merging.

The value of σ should be adjusted based on the image sensor noise, including the quantisation noise, and the Exposure Value (EV) step (exposure time ratio) between the individual

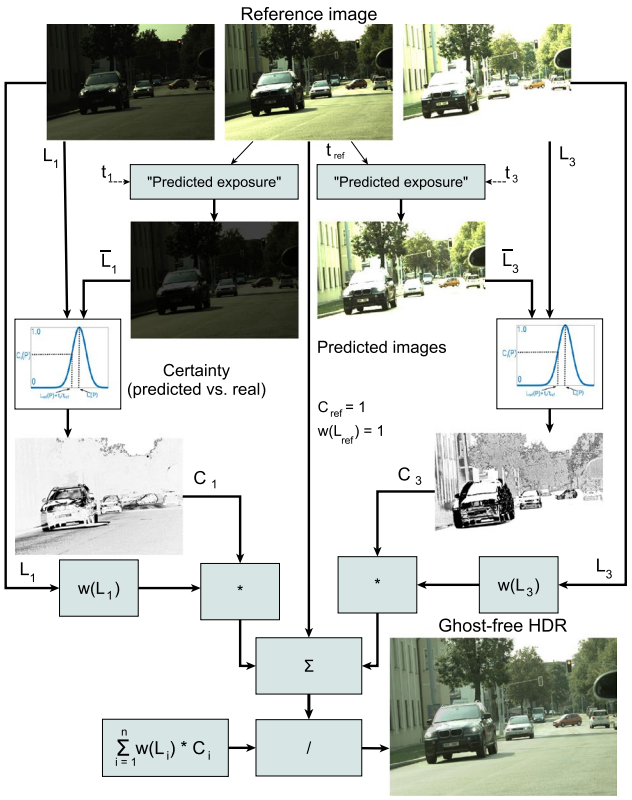

Fig. 3 A scheme illustrating the proposed ghost-free merging of according to Eq. 4 on a sequence of three images

images in the HDR sequence. In our reference implementation, we achieved the best results with $\sigma = 5$ for the exposure step of 1EV and $\sigma = 11$ for 2EV.

The proposed merging algorithm (Sect. 3.2) is applicable on grayscale, RGB and also RAW image data. The RAW data can be merged in the original form before debayering and debayered afterwards, as proposed by Tamburrino et al. [24]; thus, this implementation can save approximately 2/3 of the operations comparing to merging in RGB space.

The proposed ghost-free merging is applicable to an arbitrary number of images in the sequence of exposures. However, the following implementation and performance comparisons are related to merging of three images, unless stated otherwise.

4.1 Computing optimisations

The proposed algorithm performs per-pixel processing and requires a relatively small number of per-pixel operations. Some of its functionality is computationally demanding (e.g. division and Gauss function calculation), however, it can be optimised and/or tabulated. Inverse CRF and triangle weight functions can be tabulated thanks to the limited number of possible LDR (Low Dynamic Range) pixel values. The ratio between exposures t_i and t_{ref} in Eq. 2 can be calculated once for each setting of LDR exposure times. The Gaussian function (Eq. 3) can be convenient because the pixel values are discrete and only a finite combination of pixel values is possible, especially when considering only the differences between the captured and predicted values. The number of the Gaussian function results with relevant certainty, e.g. $> 1\%$ is limited, especially for a higher σ. The evaluation in Sect. 5 is performed with $\sigma = 11.0$, which leads only to 35 various results.

The functions represented in the tables are pre-calculated using the processors present in the embedded acceleration platforms. If needed, they can be updated while the accelerator executes the main algorithm.

4.2 Precision evaluation

The CPU and GPU reference implementations are written in C++ and CUDA, using standard 32-bit floating point data type. The whole FPGA design is implemented using only fixed point data representation and arithmetic, which is natural and also efficient for FPGA hardware. The ranges of numerical values in the individual pipeline stages are known; therefore, it is feasible to adapt the bit width of the individual parts of the pipeline to achieve a sufficient range (and precision) without using the floating point representation, whose resource requirements are generally much higher. The FPGA implementation is fixed for merging three LDR images with up to 10 bit depth. The input of the ghost detection block consists in three corresponding pixels in 10.8 fixed point representation (10 bits for integer and 8 bits for decimal part). The fractional part can be used for the data after the linearization process (application of CRF—Camera Response Function [2, 14, 18]). The resulting *Certainty maps* are in the 1.10 format and all further mathematical operations during the HDR merging are performed using 10.12 precision. The accuracy of fixed point arithmetic comparing to the software float implementation is evaluated using PSNR and MSSIM metrics. The ghost detection and merging achieved PSNR of 51.1 and 58, MSSIM is over 99% for both algorithms, using the above mentioned 12bit fractional bits.

4.3 Performance evaluation

The performance of the algorithm on the relevant platforms is summarised in Table 1. We measured only the core algorithms, without any data preprocessing—we assume that at

least in the FPGA and GPU implementations, the images are transferred into the memory using DMA in the background, without any performance losses. With the proposed optimisations, the algorithm is single-pass only. Table 1 compares the performance of the proposed Ghost-free merging of three LDR images on FPGA, SoC GPU and CPU platforms. In case of FPGA, the design achieves target frequency of 200MHz and is fully pipelined; therefore, it allows production of result pixels every clock cycle. Unlike in the sequential CPU and GPU processing, increasing the amount of work that the FPGA pipeline performs leads to consumption of more resources and prolonging the processing pipeline, which has a negative influence on latency; however, the data throughput remains the same (see Table 1).

The Table 2 presents an FPGA resource consumption of proposed design. The abbreviations in the table describes the FPGA primitives: LUT—Look-up Table; FF—registers; BRAM – Block RAM (36kbit block of distributed memory); DSP—Digital Signal Processing block (used as a multiplier); LUTRAM—LUT-based small distributed memory.

A line "Total (HLS)" indicates the amount of resources estimated by Xilinx High Level Synthesis (HLS) design tool[1]. Such resources are quite often overrated and the Place and Route process optimises out an unnecessary logic (see line "Total (Routing)") for the target FPGA. The Table 2 shows e.g. most of BRAM resources were conveniently converted into LUTRAM, probably due to only a few Gauss coefficients needed to store, as explained in Sect. 4.1.

We implemented the proposed algorithm into FPGA based HDR video acquisition pipeline proposed by Nosko et al. [15]. The proposed algorithm was designed to replace the original and very simple "Deghosting & merging" block (please refer to Nosko et al. [15]). The Table 3 compares the resources consumed by such pipeline with pipeline from Bouderbane et al. [1]; unfortunately, they do not provide more detailed statistics. For detailed description regarding pipeline, please refer to the article by Nosko et al. [15]. Please note that proposed design is built on Xilinx Zynq and Bouderbane camera on Virtex-6 and also that in Nosko's pipeline, more than 1/3 of LUT and Register resources and most of BRAM and DSPs are occupied by local tone-mapping operator [15].

The implementation of proposed ghost-free merging algorithm increases the power consumption of the complete HDR pipeline about 12 mW @ 96.4FPS; we consider relevant only the difference to "standard" HDR merging as it has to be present in the pipeline anyway. The overall HDR camera power consumption [15] is approx. 8 W, out of which the FPGA pipeline consumes only 0.94 W (estimated by Xilinx Vivado); therefore, the measurement of 12 mW difference would be under the level of measurement error, so we kept Vivado estimation.

Comparing to a reference CPU implementation, which consumes 25W @ 25.25FPS (see Table 2), the FPGA power consumption is marginal. The CPU requires 990mJ per frame, out of which the ghost-free part consumes 500 mJ, while the FPGA requires 0.2 mJ per frame, out of which the ghost-free part consumes 0.12 mJ.

5 Results

We performed evaluation and comparison of the proposed method to various state-of-the-art method on several HDR datasets [9, 26, 27] and image sets from many de-ghosting related articles [3, 8, 20]. We also performed additional evaluation on our data. In general, the visual results are comparable to the state-of-the-art; however, the proposed algorithm is capable of running in real-time, while he state-of-the-art algorithms require long offline processing in terms of seconds or even minutes per image.

The results of the proposed ghost-free merging are presented in Figs. [1, 4, 6, 7, 8, 9, 10 and 11] . Our method is suitable for almost any application with stationary cameras. Besides the evaluation of various generic datasets, the ghost removing capability was evaluated on a traffic monitoring task, where the main goal was to preserve the greatest possible level of detail so that the images can serve as evidence, with the readability of the licence plates of the vehicles in motion playing the most important part. Figure 1 contains a car approaching the camera at approximately 50km/h. Still, six exposures (∼ 66ms at 90FPS) were intentionally omitted between the images to show the capability of the ghost removing for e.g. faster moving objects. According to our experiments and/or based on the authors' claims, most of the de-ghosting methods related to our approach are very dependent on scene composition, luminance distribution, or other assumptions (see Fig. 5). Our approach does not have such limitations, it is more robust, and does not require user-guided tuning of parameters, unlike algorithms with similar complexity.

In general, the existing methods are more or less using fixed or user-adjusted thresholds and binary ghost maps, which either includes the pixel into the merging process or omits it completely. Such approach negatively affects the merging process and appearance of the resulting HDR image, causes higher noise on the affected patches around the moving objects, and also on wrongly detected patches.

5.1 Dataset evaluation and comparison

We performed the evaluation on datasets [9, 26, 27], containing sequences of images of various scenes and different

[1] http://www.xilinx.com.

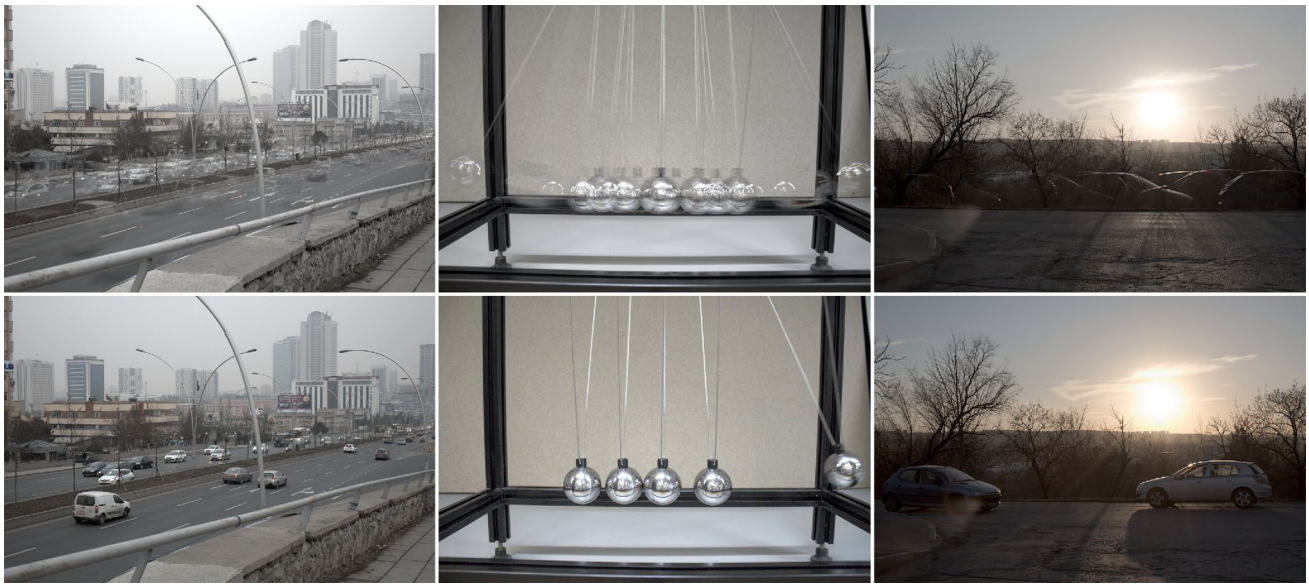

Fig. 4 Ghosted HDRs (top line) and HDRs merged using proposed ghost-free method (bottom line) on sequences "Fast cars" [26] (left), "105" [27] (middle) and "117" [27] (right). Datasets contains 9 LDR (Low Dynamic Range) images

Fig. 5 Sample outputs of related deghosting algorithms [16] (left) and [12] (right) on the scene from Fig. 1. Our experiments showed that the algorithms listed should be successful only on images with convenient histogram distribution

types of motion. The results provide a comparison of the proposed method with generally more precise and computationally demanding methods, commonly based on optical flow, which were not even included into the related work due to their complexity and high computational demands.

One of the datasets [9] contains multiple scenes with artificial objects movements. Its advantage consists in the existence of the ground truth image, which allows a comparison to the results as well as to many results of various published methods [6, 20, 22]. Figures 6 and 7 show the capabilities of the proposed method, showing that it provides results visually comparable to optical flow based methods.

Tursun et al. [26, 27] published two datasets and proposed metrics for evaluation of HDR de-ghosting quality. The evaluated samples from the datasets are shown in Fig. 4 and the HDR quality metric [26] is evaluated in Table 4. The metric evaluates the dynamic range achieved inside the motion regions, considering also the correctness of the de-ghosting. The image sets, in which we got worse results than other algorithms, were successfully de-ghosted anyway; however, the worse results were probably caused by losses in the dynamic range. Evaluation of the proposed method on these datasets also proves that the proposed method is generally usable for sequences larger than two/three images, commonly used in cameras. In all the referenced datasets [9, 26, 27], the proposed algorithms achieved results visually comparable or even better than more complex algorithms (see Fig. 9). However, the proposed method and also many HDR de-ghosting methods may yield artifacts in regions where the moving objects in the reference image are poorly-exposed, as Tursun et al. concluded [26].

Another metric we found useful is HDR-VDP2 by Mantiuk et. al. [11]. The metric evaluates the visibility and quality differences in image pairs and represents a probability that an average observer will notice a difference in the images in the pair (see Fig. 10). The essential problem for the metric evaluation is the absence of ground truth images. Applying this metric on image sets without ground truth reference seems useless, as even the state-of-the-art algorithms may fail in ghost detection and/or changes in the image quality e.g. by bluring of motion regions (see top of Fig. 9). As a result, the metric output obtained on such data does not have any meaningful value.

Karaduzovic's [9] dataset contains ground truth images, because it contains scenes with artificial object motion. We evaluated the metric on "complex" scenes and used the HDR merged from the ground truth sequence as a reference. The ground truth sequence is processed also by our algorithm (with de-ghosting disabled) to eliminate the effect

Fig. 6 Output of proposed ghost-free merging method on the sequence of Gallo [3] (top). Previews of the various algorithm results are shown at the bottom: Gallo et al. [3] (**a**), Jacobs et al. [7] (**b**), Pece et al. [16] (**c**), Zhang et al. [30] (**d**) and proposed algorithm (**e**). The previews A to D are published online at http://www.vsislab.com/projects/IPM/HDR/project.html

Table 1 The table compares the performance of the proposed ghost-free merging of 3 LDR images (Fig. 2) with a resolution of 1920×1080 on following platforms: FPGA Xilinx Zynq, embedded CPU and GPU Nvidia Tegra TX2 and CPU Intel Core i7-3770 (single core)

	FPGA	TX2 GPU	TX2 CPU	CPU
Certainty map [ms]	10.3	1.59	45.9	16.6
Merging [ms]	10.3	4.58	112.3	23.0
Total [ms]	10.3	6.17	158.2	39.6
Overall FPS	96.45	162.07	6.32	25.25

Table 2 FPGA Resource utilisation for merging 3 LDR images of 1920×1080 pixels. Design is routed for Xilinx Zynq Z-7020

	LUT	LUTRAM	FF	BRAM	DSP
Certainty maps	3532	–	3339	4	4
HDR merging	893	–	2570	10	16
Total (HLS)	4425	–	5909	14	20
Total (Routing)	**1057**	**252**	**2052**	**2**	**16**
available	53200	17400	106400	280	220
utilisation [%]	1.99	1.45	1.93	0.72	7.27

The total resource consumption for this FPGA is marked as bold

Table 3 Resource utilization of complete camera solution of Nosko et al. [15] enhanced by the proposed ghost-free merging block, comparing to Bouderbane [1]

	LUT	LUTRAM	FF	BRAM	DSP
Prop. pipeline	39145	3137	53592	51	58
Bouderbane [1]	49193	–	50399	35	20

of unrelated image enhancements. Table 5 contains an overall "quality" metric of the produced ghost-free HDR output according to HDR-VDP2 [11] metric. Figure 10 shows "scene 1" with highlighted differences between ground truth HDR and ghost-free HDR.

All of the datasets were evaluated with either $\sigma = 11$ or $\sigma = 5$ based on the EV step in the data. The source codes and evaluated datasets are available online[2].

[2] https://github.com/ghostfreehdr/HDR.

Fig. 7 Output of the proposed HDR ghost-free merging method for *Complex Scene 1* of dataset [9] (left). Ghosted HDR image is shown on the right. Previews of various algorithm results are shown at the bottom. No de-ghosting (**a**), Silk et al. [22] (**b**), Sen et al. [20] (**c**), Photoshop (**d**), Photomatix (**e**) and proposed algorithm (**f**). The previews A to E are published as a part of a Karaduzovic dataset [9]

Table 4 Results of the "Dynamic Region Dynamic Range" metric proposed by Tursun [26] and evaluated on their dataset

Metric "DR"	[4]	[20]	[22]	None	This work
Cafe	**2.63**	2.61	2.60	2.47	2.42
FastCars	1.12	1.18	1.10	1.10	**1.38**
Flag	1.40	1.50	1.49	1.45	**1.59**
Gallery1	1.59	1.59	1.56	1.55	**1.70**
Gallery2	2.41	**2.56**	2.14	2.29	2.05
LibrarySide	1.78	1.93	1.60	1.76	**3.20**
Shop1	2.20	2.39	2.00	2.10	**2.42**
Shop2	2.68	2.72	**2.89**	2.55	2.42
WalkingP.	1.94	**2.07**	1.83	2.05	1.58

The metric evaluates the resulting dynamic range within regions containing movement; the higher the value, the better. The best achieved values are in bold

6 Conclusion

In this paper, we proposed a novel ghost-free HDR merging algorithm suitable for real-time implementation in embedded devices. The algorithm produces results comparable to the state-of-the-art, as shown in the evaluation. While the algorithm is constrained to optically aligned images, it has a very low computational complexity as compared to the state of the art. Furthermore, it is suitable for implementation in embedded systems, as well as in programmable hardware (FPGA) or GPU, as experimentally demonstrated, or even custom chips. Thanks to

Fig. 8 The source sequence (top left) is merged with (bottom) and without (top right) the proposed ghost-free merging algorithm. Images retrieved from Sing Bing Kang [8]

Fig. 9 Figure shows scene "Cafe" from Tursun's [26] dataset processed with Sen [20] (top) and the proposed ghost-free algorithm (bottom). Sen [20] produce a heavily blurred image, which precludes the HDR-VDP metric [11]; moreover, the de-ghosting method fails (see marked areas, where objects are shadowed and blured)

Table 5 Evaluation of the HDR-VDP2[11] metric on a "complex" scene from Karaduzovic's [9] dataset

	Scene1	Scene2	Scene3	Scene 4
Q	73.24	76.20	82.83	71.58

the unique features of algorithm, the systems exploiting our solution can produce high quality HDR and/or tone-mapped video while maintaining low cost and achieving low power consumption and the small footprint suitable e.g. for smart cameras. The source codes of the proposed algorithm and evaluated datasets are available online[2].

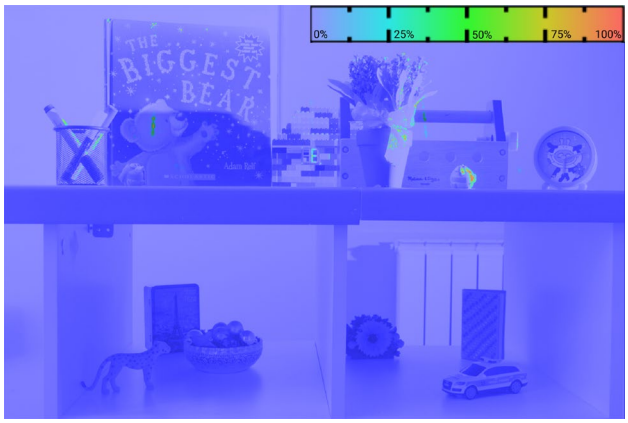

Fig. 10 Figure shows "scene 1" from Karaduzovic's [9] dataset processed by the HDR-VDP2 [11] metric. The colour bar reflects the probability that an average observer will notice a difference between ghost-free HDR and ground truth HDR. De-ghosted HDR visual quality, according to HDR-VDP2 metric [11] is 73.24 (see Table 5)

Fig. 11 A car passing by the camera—ghosted HDR (left) and result of the proposed ghost-free merging algorithm (right)

Acknowledgements This work was supported by VRASSEO (VI20172020068) and by the Ministry of Education, Youth and Sports of the Czech Republic from the National Programme of Sustainability (NPU II); project IT4Innovations excellence in science - LQ1602.

References

1. Bouderbane, M., Dubois, J., Heyrman, B., Lapray, P.J., Ginhac, D.: Ghost removing for hdr real-time video stream generation. In: Real-time image and video processing (2016)
2. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: ACM Trans. Graph., SIGGRAPH '97 (1997)
3. Gallo, O., Gelfandz, N., Chen, W.C., Tico, M., Pulli, K.: Artifact-free high dynamic range imaging. In: 2009 IEEE International conference on computational photography (ICCP), pp 1–7 (2009). https://doi.org/10.1109/ICCPHOT.2009.5559003
4. Grosch, T.: Fast and robust high dynamic range image generation with camera and object movement (2006)
5. Grossberg, M.D., Nayar, S.K.: Determining the camera response from images: what is knowable? IEEE Trans. Pattern Anal. Mach. Intell. **25**(11), 1455–1467 (2003). https://doi.org/10.1109/TPAMI.2003.1240119

6. Hu, J., Gallo, O., Pulli, K., Sun, X.: Hdr deghosting: How to deal with saturation? In: 2013 IEEE Conference on computer vision and pattern recognition (2013)
7. Jacobs, K., Loscos, C., Ward, G.: Automatic high-dynamic range image generation for dynamic scenes. IEEE Comput. Graph. Appl. **28**(2), 84–93 (2008). https://doi.org/10.1109/MCG.2008.23
8. Kang, S.B., Uyttendaele, M., Winder, S., Szeliski, R.: High dynamic range video. ACM Trans. Graph. **22**(3), 319–325 (2003)
9. Karaduzovic-Hadziabdic, K., Hasic, T.J., Mantiuk, R.K.: Multi-exposure image stacks for testing hdr deghosting methods (2017)
10. Mandel, L.: Fluctuations of photon beams: the distribution of the photo-electrons. Proc Phys Soc **74**(3), 233 (1959)
11. Mantiuk, R., Kim, K.J., Rempel, A.G., Heidrich, W.: Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. ACM Trans. Graph. **30**(4), 40:1–40:14 (2011). https://doi.org/10.1145/2010324.1964935
12. Min, T.H., Park, R.H., Chang, S.: Histogram based ghost removal in high dynamic range images. In: Multimedia and Expo, 2009. ICME 2009. IEEE international conference on. IEEE (2009)
13. Min, T.H., Park, R.H., Chang, S.: Noise reduction in high dynamic range images. Signal, Image and Video Processing **5**(3) (2011). https://doi.org/10.1007/s11760-010-0203-7
14. Mitsunaga, T., Nayar, S.K.: Radiometric self calibration. In: Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149), vol. 1, p. 380 Vol. 1 (1999)
15. Nosko, S., Musil, M., Zemcik, P., Juranek, R.: Color hdr video processing architecture for smart camera. J. Real-Time Image Process. (2018). https://doi.org/10.1007/s11554-018-0810-z
16. Pece, F., Kautz, J.: Bitmap movement detection: Hdr for dynamic scenes. In: Visual media production, 2010 conference on, pp. 1–8. IEEE (2010)
17. Raman, S., Kumar, V., Chaudhuri, S.: Blind de-ghosting for automatic multi-exposure compositing. In: ACM SIGGRAPH ASIA 2009 Posters, SIGGRAPH ASIA '09. ACM, USA (2009)
18. Robertson, M.A., Borman, S., Stevenson, R.L.: Estimation-theoretic approach to dynamic range enhancement using multiple exposures. J. Electron. Imaging **12**(2), 219–228 (2003)
19. Sakakibara, M., Kawahito, S., Handoko, D., Nakamura, N., Satoh, H., Higashi, M., Mabuchi, K., Sumi, H.: A high-sensitivity cmos image sensor with gain-adaptive column amplifiers. IEEE J. Solid-State Circuits **40**(5), 1147–1156 (2005)
20. Sen, P., Kalantari, N.K., Yaesoubi, M., Darabi, S., Goldman, D.B., Shechtman, E.: Robust Patch-Based HDR Reconstruction of Dynamic Scenes. ACM SIGGRAPH Asia (2012)
21. Sidibe, D., Puech, W., Strauss, O.: Ghost detection and removal in high dynamic range images. In: 2009 17th European Signal Processing Conference, pp. 2240–2244 (2009)
22. Silk, S., Lang, J.: Fast high dynamic range image deghosting for arbitrary scene motion. In: Proceedings of Graphics Interface 2012, GI '12, pp. 85–92. Canadian Information Processing Society, Toronto, Ont., Canada, Canada (2012). http://dl.acm.org/citation.cfm?id=2305276.2305291
23. Srikantha, A., Sidibé, D.: Ghost detection and removal for high dynamic range images: recent advances. Sig. Process. **27**(6), 650–662 (2012). https://doi.org/10.1016/j.image.2012.02.001https://www.sciencedirect.com/science/article/pii/S0923596512000306
24. Tamburrino, D., Alleysson, D., Meylan, L., Süsstrunk, S.: Digital camera workflow for high dynamic range images using a model of retinal processing. In: IST/SPIE electronic imaging: digital photography IV, vol. 6817 (2008)
25. Tocci, M.D., Kiser, C., Tocci, N., Sen, P.: A versatile hdr video production system. In: ACM SIGGRAPH 2011 Papers, SIGGRAPH '11. USA (2011)
26. Tursun, O.T., Akyüz, A.O., Erdem, A., Erdem, E.: The state of the art in hdr deghosting: a survey and evaluation. Comput. Graphics Forum **34**(2) (2015). https://doi.org/10.1111/cgf.12593
27. Tursun, O.T., Akyüz, A.O., Erdem, A., Erdem, E.: An objective deghosting quality metric for hdr images. Comput. Graph. Forum **35**(2), 139–152 (2016). https://doi.org/10.1111/cgf.12818
28. Wang, C., Tu, C.: An exposure fusion approach without ghost for dynamic scenes. In: 2013 6th International congress on image and signal processing (CISP), vol. 2, pp. 904–909 (2013). https://doi.org/10.1109/CISP.2013.6745293
29. Wu, S., Xie, S., Rahardja, S., Li, Z.: A robust and fast anti-ghosting algorithm for high dynamic range imaging. In: 2010 IEEE International Conference on Image Processing, pp. 397–400 (2010). https://doi.org/10.1109/ICIP.2010.5654196
30. Zhang, W., Cham, W.K.: Gradient-directed composition of multi-exposure images. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 530–536. IEEE (2010)
31. Zhao, H., Shi, B., Fernandez-Cull, C., Yeung, S.K., Raskar, R.: Unbounded high dynamic range photography using a modulo camera. In: ICCP (2015)

Martin Musil received the MS degree at the Faculty of Information Technology, Brno University of Technology, Czech Republic. He is currently a PhD student and member of the Graph@FIT group at the Department of Computer Graphics and Multimedia at FIT, Brno University of Technology.

Svetozar Nosko received the MS degree at the Faculty of Information Technology, Brno University of Technology, Czech Republic. He is currently a PhD student and member of the Graph@FIT group at the Department of Computer Graphics and Multimedia at FIT, Brno University of Technology.

Pavel Zemcik received his PhD degree from the Faculty of Electrical Engineering and Computer Science, Brno University of Technology, Czech Republic. He works as a full professor, dean and member of the Graph@FIT group at the Department of Computer Graphics and Multimedia at FIT, Brno University of Technology. His interests include acceleration of computer vision and graphics algorithms, programmable hardware and also applications.