2014 22nd International Conference on Pattern Recognition

Simultaneous HDR and Optic Flow Computation

David Hafner, Oliver Demetz, and Joachim Weickert
Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science,
Saarland University, Saarbrücken, Germany
Email: {hafner, demetz, weickert}@mia.uni-saarland.de

***Abstract*—Camera shakes and moving objects pose a severe problem in the high dynamic range (HDR) reconstruction from differently exposed images. We present the first approach that simultaneously computes the aligned HDR composite as well as accurate displacement maps. In this way, we can not only cope with dynamic scenes but even precisely represent the underlying motion. We design our fully coupled model transparently in a well-founded variational framework. The proposed joint optimisation has beneficial effects, such as intrinsic ghost removal or HDR-coupled smoothing. Both the HDR images and the optic flows benefit substantially from these features and the induced mutual feedback. We demonstrate this with synthetic and real-world experiments.**

I. Introduction

Many modern cameras have an auto exposure bracketing mode that creates a sequence of images with varying exposure times. This option of capturing differently exposed *low dynamic range* (LDR) images can be exploited to recover a larger dynamic range of the scene that goes beyond the contrast range of the sensor. However, standard methods require perfectly aligned images to compute the *high dynamic range* (HDR) composite. In a practical scenario this requirement is often not fulfilled: Even small camera shakes or moving objects lead to severe artefacts in the final result.

Most previous research on HDR imaging of dynamic scenes tries to compensate for this motion by preprocessing the LDR images *before* fusing them. However, since the alignment and the HDR reconstruction highly influence each other, it would be natural to solve them in a *joint* ansatz that benefits from mutual interactions. In contrast to existing methods, we are not only interested in a nice-looking image with high contrast, but also in an accurate representation of the scene motion. On the one hand, this opens the possibility for further computer vision tasks that are based on precise motion estimations such as inter-frame computation, scene understanding and many more. On the other hand, in this way we want to ensure that only pixels are merged in the composite that belong to the same object. This is particularly desirable if the physical correctness of the computed HDR irradiance values is of great importance.

Contributions: The goal of our paper is to pursuit this joint ansatz by proposing a fully coupled approach that simultaneously computes

(i) the aligned HDR irradiance map,
(ii) the dense and accurate optic flow field for each image.

Our model is based on an energy functional that is simultaneously minimised w.r.t. these two quantities. Although our functional is composed of relatively simple and intuitive assumptions, the resulting overall model turns out to be very powerful. We show that the inherent feedback among the HDR irradiance map and the optic flow is extraordinarily beneficial for the quality of *both* estimates.

Fig. 1. Given a set of unaligned images taken with different exposure times (*first row*), we simultaneously compute the optic flow fields (*second row, left and right*) and the aligned HDR composite (*second row, middle*). In this way, we can cope with camera as well as object motion. Here, the camera is shaking while the rabbit moves from frame to frame.

Outline: We start with a discussion of related work in Sec. II. Subsequently we present our observation model (Sec. III) that builds the basis of the energy functional in Sec. IV. After its algorithmic realisation in Sec. V, we demonstrate the performance of our approach (Sec. VI). Finally, we conclude the paper with a summary and an outlook in Sec. VII.

II. Related Work

The pure HDR reconstruction problem with perfectly aligned LDR images has been widely researched in the last two decades; see e.g. [1], [2] or the book of Reinhard et al. [3] for a broad overview. More recently, there has also been a lot of research tackling the problem of motion in the HDR acquisition pipeline. In general one can distinguish two classes of algorithms: ghost rejection and image alignment.

Assuming the images to be globally aligned, the task of the so-called *deghosting* methods is to identify and eliminate remaining motion artefacts that may show up as ghosts in the final HDR image. An extensive and detailed review of these deghosting methods is beyond the scope of our paper. For this purpose, we would like to refer to the recent survey of Srikantha and Sidibé [4] and references therein.

Our approach is conceptually different from those deghosting techniques. Rather, it can be attributed to the image alignment methods that try to register all parts of the images. This generally allows to fuse all available information in the HDR composite. Contrary to early approaches (e.g. Ward [5]) that assume one global transformation for each image, the

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.360

IEEE computer society

following techniques perform a local registration of the images under generally arbitrary motion patterns. Kang et al. [6] align the LDR images by first computing a global transform for the whole image and afterwards refining this motion field with a local optic flow approach. Similarly, Menzel and Guthe [7] as well as Mangiat and Gibson [8] perform first a block matching approach that is locally refined afterwards. Both methods are equipped with an additional post-deghosting step to account for remaining misalignments. Hu et al. [9] make use of an advanced patch matching algorithm to find corresponding image regions. Also this method contains a deghosting step after the registration to handle wrong correspondences. In their follow-up work [10], they explicitly improve the handling of saturated regions. Zimmer et al. [11] pre-register the images with a dense energy-based optic flow method. Hereby, they assume that the image gradients stay almost constant under varying exposure times. In this context Gupta et al. [12] propose an exposure bracketing technique to enhance the accuracy of existing optic flow methods that are applied for the pre-alignment step. Most related to our idea is the paper of Sen et al. [13]. In contrast to a decoupled pre-alignment, they are the first to *jointly* estimate the aligned input images and the HDR composite. However, in contrast to their method we do not only output aligned images. In fact, we also compute accurate representations of the scene motion (optic flow) which may serve as input for further tasks. Moreover, our model incorporates the camera-specific response function explicitly. First, this allows us to refrain from a preliminary transformation of the LDR images to the linear domain as required in the method of Sen et al. Second, and even more important, it allows our optimisation to adapt to the shape of this function in an accurate way.

III. Observation Model

We assume that the unknown HDR irradiances $r(\boldsymbol{x})$ are related to the observed LDR images $f_i(\boldsymbol{x})$, $i = 1 \dots N$, via the *camera response function* (CRF) Φ as follows:

$$f_i(\boldsymbol{x}) \;=\; \Phi(t_i \cdot r(\boldsymbol{x}))\,, \tag{1}$$

where $\boldsymbol{x} = (x, y)^\top$ denotes the position on the 2D rectangular image domain $\Omega \in \mathbb{R}^2$, and t_i the exposure time corresponding to f_i. As mentioned in the introduction, varying the exposure times and merging the differently exposed images allows to reconstruct the HDR content of the scene. W.l.o.g. we assume that the CRF Φ maps the incident light energy $t_i \cdot r(\boldsymbol{x})$ to the interval of 0 and 1. After this mapping, the values are quantised and stored with a certain amount of bits per brightness value.

Equation 1 only holds for a stack of perfectly registered input images. However, even small camera shakes or moving objects break this assumption. Hence, to account for such motion, we modify our model in the following way:

$$f_i(\boldsymbol{x} + \boldsymbol{w}_i(\boldsymbol{x})) \;=\; \Phi(t_i \cdot r(\boldsymbol{x}))\,, \tag{2}$$

where $\boldsymbol{w}_i(\boldsymbol{x}) \;=\; (u_i(\boldsymbol{x}), v_i(\boldsymbol{x}))^\top$ denotes the displacement field, i.e. the optic flow. More precisely, an optic flow vector field $\boldsymbol{w}_i : \Omega \to \mathbb{R}^2$ allows for each position $\boldsymbol{x}$ in the irradiance map r to specify the corresponding position $\boldsymbol{x} + \boldsymbol{w}_i(\boldsymbol{x})$ in the unaligned input image f_i. Thus, once the optic flow has been computed, the registration of the input images is straightforward. For the sake of readability, we will from now on omit the argument $\boldsymbol{x}$ of $\boldsymbol{w}_i$ and r.

IV. Variational Approach

A. Energy Functional

Let us now embed the discussed observation model as data term into a variational framework. To this end, we develop an energy functional whose minimiser yields the irradiance map r as well as the optic flow $\boldsymbol{W} := (\boldsymbol{w}_0, \dots, \boldsymbol{w}_N)$. Basically our energy consists of three main components:

$$E(r, \boldsymbol{W}) \;=\; \int_\Omega \left(\sum_{i=1}^{N} M_i + \alpha \cdot \sum_{i=1}^{N} S_{\boldsymbol{w}_i} + \beta \cdot S_r \right) \mathrm{d}\boldsymbol{x}\,. \tag{3}$$

Every LDR image yields one data term

$$M_i \;=\; \Psi\Big(\big(f_i(\boldsymbol{x} + \boldsymbol{w}_i) \;-\; \Phi(t_i\, r) \big)^2 \Big) \tag{4}$$

that represents the proposed observation model: For each position $\boldsymbol{x} + \boldsymbol{w}_i$ in the image f_i, the measured brightness value should be similar to the "predicted" brightness value that is computed by $\Phi(t_i\, r)$. We penalise the difference between those two quantities with the sub-quadratic function

$$\Psi(z^2) \;=\; \sqrt{z^2 + \epsilon^2}\,, \qquad \epsilon > 0\,, \tag{5}$$

to achieve a robustness w.r.t. outliers, caused e.g. by occlusions or noise.

The regularisation terms

$$S_{\boldsymbol{w}_i} \;=\; |\boldsymbol{\nabla} u_i|^2 + |\boldsymbol{\nabla} v_i|^2 \qquad \text{and} \qquad S_r \;=\; |\boldsymbol{\nabla} r|^2 \tag{6}$$

reward smooth flow fields and a smooth irradiance map, respectively. Here, $\boldsymbol{\nabla} := (\partial_x, \partial_y)^\top$ denotes the standard 2D gradient operator. Last but not least, the positive parameters α and β allow to steer the respective amount of smoothing.

We select the image with the fewest saturated pixels to be the reference, to which the final irradiance map should be aligned. Consequently, the flow field $\boldsymbol{w}_{\text{ref}}$ corresponding to the reference frame f_{ref} is obviously identical to $\mathbf{0}$ everywhere.

B. Minimisation

According to the calculus of variations, the minimiser of the energy (3) must necessarily fulfil the so-called Euler-Lagrange equations. With the abbreviation

$$\Psi_i' \;:=\; \Psi'\Big(\big(f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \Phi(t_i\, r) \big)^2 \Big) \tag{7}$$

they read for the flow variables u_i and v_i as follows:

$$\Psi_i' \cdot \big(f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \Phi(t_i r) \big) \cdot \partial_x f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \alpha \Delta u_i = 0\,, \tag{8}$$

$$\Psi_i' \cdot \big(f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \Phi(t_i r) \big) \cdot \partial_y f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \alpha \Delta v_i = 0\,, \tag{9}$$

and for the irradiance part

$$\sum_{i=1}^{N} \Big(\Phi'(t_i r) \cdot t_i \cdot \Psi_i' \cdot \big(f_i(\boldsymbol{x} + \boldsymbol{w}_i) - \Phi(t_i r) \big) \Big) \;+\; \beta \Delta r \;=\; 0\,. \tag{10}$$

The corresponding homogeneous Neumann boundary conditions read

$$\boldsymbol{n}^\top \boldsymbol{\nabla} u_i = 0\,, \quad \boldsymbol{n}^\top \boldsymbol{\nabla} v_i = 0\,, \quad \text{and} \quad \boldsymbol{n}^\top \boldsymbol{\nabla} r = 0\,, \tag{11}$$

where $\boldsymbol{n}$ is the outer normal vector to the boundary of Ω.

C. Discussion

In contrast to previous HDR reconstruction methods, we define our data terms (4) in the *intensity domain* and not in the irradiance domain, i.e. we do not apply the inverse camera response function. This seems at first glance like a small change. However, it turns out to be very pivotal as we will see in the following paragraphs.

1) Selective Weighting: To compute the irradiances r in (10), the product $\Phi'(t_i\, r) \cdot t_i$ by construction provides an intuitive weighting of the different LDR images. First, the derivative of the camera response function Φ' is a natural confidence measure of the intensity values that accounts for quantisation noise and saturated pixels [1]: The smaller the derivative of the CRF, the larger the quantisation noise and the less the resulting weight of the considered brightness value. Second, the additional weighting by the exposure time t_i accounts for the fact that images taken with longer exposure times contain less sensor noise and thus, should get a higher weight [14]. In other words, larger exposure times stretch the irradiance range in such a way that small values are quantised more finely, which consequently yields less under-saturated pixels.

Contrary to the irradiance computation, the discussed weighting term $\Phi'(t_i\, r) \cdot t_i$ does *not* show up in the Euler-Lagrange equations (8) and (9) of the optic flow. This is an advantageous property that naturally arises from the carefully designed energy, too: If the estimated irradiance or rather the predicted brightness value $\Phi(t_i\, r)$ suggests that a pixel should be saturated in the i-th input image, we do not discard this information. Instead, we make optimal use of this knowledge by implicitly enforcing the matching of over- and under-saturated pixels.

2) Inherent Deghosting: Besides the described discriminative weighting of the intensity values, the term $\Phi'(t_i\, r) \cdot t_i$ plays an additional useful role in the irradiance computation. It intrinsically performs ghost removal by an intuitive *plausibility check*: The estimated irradiance r yields the predicted light energy $t_i\, r$. If this prediction is larger than 1, Φ' vanishes and we do not consider the corresponding brightness value f_i in the irradiance computation since it is physically not meaningful.

In addition, this ghost removal behaviour is substantially supported by the robustified data terms (4). From an optic flow point of view, the penaliser Ψ increases the robustness w.r.t. outliers. For instance, such outliers can be caused by noise or occluded regions, i.e. parts of the reference image that are not visible in the other frames or vice versa. However, in our setting, the nonlinear terms Ψ_i' also appear in the computation of the irradiance map in (10). Here they play a different but nevertheless important role: If the motion estimation is incorrect in some parts of the image, the brightness value $f_i(\boldsymbol{x} + \boldsymbol{w}_i)$ differs significantly from its prediction $\Phi(t_i\, r)$. Fortunately, the proposed penalisation function handles these cases and weights those terms down that would cause artefacts in the HDR image. This leads to ghost-free images which further help to improve the motion estimation, and the other way around.

D. Anisotropic Modification

The optic flow smoothness terms $S_{\boldsymbol{w}_i}$ in (6) lead to the Laplacian terms Δu_i and Δv_i in the Euler-Lagrange equations (8) and (9). They provide an *isotropic* diffusion of the flow fields that is equal in all directions. However, such a smoothing is not always desirable: We clearly do not want to smooth across object structures if the motion differs between the objects. To this end, we adapt the amount of smoothing to the local structure in an *anisotropic* way. Inspired by [15], we propose a novel joint irradiance- and flow-driven smoothing that is steered by $2{\times}2$ diffusion tensors $\boldsymbol{D}_i$. These tensors are constructed from the normalised vectors $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ that point along and across edges of the evolving HDR image; i.e. they are given by the orthonormal eigenvectors of the irradiance structure tensor

$$G_\rho * \big(\boldsymbol{\nabla}(G_\sigma * r)\; \boldsymbol{\nabla}^\top (G_\sigma * r)\big)\;, \tag{12}$$

where $G_\rho *$ and $G_\sigma *$ denote convolutions with Gaussian kernels of standard deviation ρ and σ, respectively. The corresponding eigenvalues of $\boldsymbol{D}_i$ determine the amount of smoothing and are computed by

$$g\big((\boldsymbol{s}_k^\top \boldsymbol{\nabla} u_i)^2 + (\boldsymbol{s}_k^\top \boldsymbol{\nabla} v_i)^2\big) \quad \text{with} \quad k = 1, 2 \tag{13}$$

and with the Charbonnier diffusivity [16]

$$g(z^2) \;=\; \frac{\epsilon}{\sqrt{z^2 + \epsilon^2}}\;, \qquad \epsilon > 0\,. \tag{14}$$

Finally, we exchange the homogeneous diffusion terms Δu_i and Δv_i by their anisotropic counterparts

$$\operatorname{div}(\boldsymbol{D}_i\, \boldsymbol{\nabla} u_i) \qquad \text{and} \qquad \operatorname{div}(\boldsymbol{D}_i\, \boldsymbol{\nabla} v_i)\,, \tag{15}$$

where div is the 2D divergence operator. The reliance on the evolving HDR image instead of the LDR input images (irradiance- vs. image-driven) is very advantageous: It even allows to adapt to structures that are hardly visible (e.g. saturated) in the single LDR images, but clearly distinctive in the HDR composite which combines all information from the whole exposure sequence.

In a similar way, we also apply such an anisotropic modification to the smoothness term of the irradiances:

$$\operatorname{div}(\boldsymbol{D}_r\, \boldsymbol{\nabla} r)\,. \tag{16}$$

Here, the eigenvalues of the irradiance diffusion tensor $\boldsymbol{D}_r$ with eigenvectors $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are determined by

$$g\big((\boldsymbol{s}_1^\top \boldsymbol{\nabla} r)^2\big) \quad \text{and} \quad g\big((\boldsymbol{s}_2^\top \boldsymbol{\nabla} r)^2\big)\,. \tag{17}$$

Thanks to its edge-preserving property, the anisotropic diffusion of the irradiances is particularly beneficial in the presence of noise or under high ISO settings. This extends the work of Rameshan et al. [17]. They showed (using isotropic smoothness terms) that an incorporation of the denoising task in the HDR fusion is preferable to a decoupled denoising of the input images before composing the HDR map.

E. Colour Images

For didactic reasons, we restricted ourselves to greyscale images so far. However, an extension of our method to colour images is straightforward: First, we perform a joint robustification of all colour channels, i.e. we sum up the channels

in the argument of Ψ in (4). Second, we use the combined structure tensor [18] to determine the eigenvectors $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ of the diffusion tensors. To compute the eigenvalues of $\boldsymbol{D}_r$, we add up the projections of all colour channels in the argument of the diffusivity function g in (17).

V. Algorithmic Realisation

The Euler-Lagrange equations in Sec. IV-B form a non-convex and nonlinear system of equations. Similar to [19], we transform this nonconvex problem to a series of convex subproblems. To avoid being trapped in local minima and to handle large motions, we embed this series in a coarse-to-fine pyramid approach. In each step, we solely compute small incremental values of u, v and r. These increments are then successively used to update the solutions from the previous iterations. Hereby, all optic flow vectors are initialised with $\mathbf{0}$ while we determine the initialisation of the irradiance map by means of the reference frame f_{ref}.

Following Brox et al. [19], we apply in each subproblem a *lagged nonlinearity* method to handle the occurring nonlinear terms. With this procedure, we have to solve a sparse linear system of equations in each intermediate step, while the nonlinearities are updated in an outer loop. We solve these linear systems iteratively with a positivity constraint on the irradiances. To this end, we apply a variant of the the Gauß-Seidel method, namely projected successive overrelaxation [20].

VI. Experiments

Our experiments consist of two parts. First, we consider a synthetic HDR scene (Sec. VI-A). The main reason for the use of such a synthetic data set is the availability of ground truth data that is particularly required for a quantitative comparison of the optic flow fields. Second, the experiments in Sec. VI-B show the performance of our method on a real-world exposure series. To visualise the computed HDR irradiance maps we apply the popular tone mapping operator of Fattal et al. [21] (implementation: http://pfstools.sf.net/). Furthermore, we applied a known (nonlinear) CRF for the synthetic scene. For the real-world scene, we calibrated the CRF in advance with a set of perfectly aligned images; following the method of Grossberg and Nayar [22].

A. Synthetic Data Set

The tree input images of our synthetic data set are depicted in Fig. 2. The exposure series shows camera as well as object motion, and every LDR image contains large over- or under-saturated regions.

Let us now compare our approach to state-of-the-art HDR alignment methods, in particular to the optic flow-based method of Zimmer et al. [11] and the patch-based methods of Sen et al. [13] and Hu et al. [10]. Since Zimmer et al. and Hu et al. do not output an HDR image directly, we used their aligned images as input for the HDR algorithm of Robertson et al. [14] (implemented in the *pfstools*). As mentioned, a comparison with pure deghosting approaches is out of the scope of our work (cf. Sec. II). Rather, we want to focus on methods that do not "simply" reject moving objects but try to align all parts of the input images. Otherwise, it would not be possible to reconstruct the HDR content of moving objects. Additionally,

Fig. 2. Synthetic data set modeled with *Blender*. The exposure times of the three input images vary by ± 3 exposure values (EV). During the acquisition, the camera is rotating clock-wise while the balloon is moving towards the top right corner.

we want to stress that we are also especially interested in the underlying scene motion: We consider the motion estimation not only as a side-product, but rather as a valuable part of the available information about the scene.

First we compare our approach with the state-of-the-art method for optic flow-based HDR registration by Zimmer et al. Figure 3 shows that while the method of Zimmer et al. is not able to detect the motion of the balloon, our approach yields very reliable flow fields. This observation is underlined by Table I that rates the quality of the computed flow fields in terms of the average angular error (AAE) [23]. Here, we also see that our joint method clearly outperforms the recent optic flow method by Demetz et al. [24] that is specifically tailored to such illumination changes. It is worth mentioning that also further pre-alignment tests with top-ranked optic flow methods, which are not especially designed for such severe brightness variations, produced unsatisfactory results. To summarise, the computed flow fields demonstrate that our proposed *coupled computation of the irradiances and the optic flow is clearly preferable to a decoupled optic flow-based pre-alignment*. Our joint optic flow computation marks the state-of-the-art in this HDR setting.

Concerning the irradiance maps (cf. Fig. 4), it is obvious that an additional post-deghosting step in the approach of Zimmer et al. is unavoidable. In contrast, our approach contains such deghosting features inherently. This is very advantageous: Obviously, it removes ghosts in the evolving HDR images. On top of that, ghost-free HDR irradiance maps yield better motion estimations; see e.g. irradiance-driven smoothing or explicit enforcing of the mapping of saturated pixels. In this way, our joint optimisation and the induced mutual interactions successively improve both estimates and produce very accurate final results.

As mentioned in Sec. II, Sen et al. also follow the idea of a joint computation of the HDR image and the aligned LDR images. However, due to the patch-based nature of their approach they do not compute a meaningful displacement map with sub-pixel precision. Unfortunately, they do not output a displacement map at all. Thus, it is not possible to accurately specify the scene motion which may serve as input for further computer vision tasks. Regarding the HDR image, we observe artefacts especially in regions that are saturated in one of the LDR images like the sky region. Here, their algorithm produces unpleasant results with low contrast.

Hu et al. do not explicitly exploit the benefits of a joint alignment and irradiance computation since their goal is "only" the alignment of the image stack — not the HDR composite.

TABLE I. QUANTITATIVE COMPARISON OF THE COMPUTED OPTIC FLOW FOR THE SYNTHETIC DATA SET IN FIG. 2.

method	AAE (in degree)
proposed	3.47
Demetz et al. [24]	5.17
Zimmer et al. [11]	12.60
Hu et al. [10]	44.20
Sen et al. [13]	not provided

The final image looks visually similar to our result. Nevertheless, differences can again be found especially in the sky region (Fig. 4, *fourth row*). Further, the unreliable displacement maps in Fig. 3 (*left*) illustrate that the method does not always fuse the correct pixels. This may result in visually attractive HDR images but not in accurate, physically meaningful irradiances.

Last but not least, it is very important to note that Sen et al. as well as Hu et al. extremely weight down the influence of all other input images wherever the reference image is not saturated. Hence, they do not make optimal use of all available information which results in less accurate HDR maps.

B. Real-World Data Set

Let us now experimentally verify our observations on real-world data. To this end, we consider a sequence of five images each separated by one exposure value. Figure 1 (*first row*) depicts three of these five images where again camera and object motions are present. While all other methods have problems particularly with the specularities on the window sill, our final HDR image shows no such artefacts (cf. Fig. 5). Due to the patch-based approach, the algorithm of Sen et al. even produces new content in this region. This clearly illustrates that optic flow is required to produce a reliable representation of scene motion and thereby fuse the correct corresponding pixels. In this regard, our flow fields in Fig. 1 (*second row, left and right*) again show highly accurate motion estimations.

Limitation: There is one general limitation of all warping-based optic flow methods that unavoidably carries over to our approach: If the sequence is undersampled in time such that small objects undergo a too large displacement, the motion estimation tends to fail. Clearly, it is the task of future research to overcome this drawback. To tackle it, an interesting way may be the design of appropriate hardware with sufficiently small time intervals between the differently exposed images.

VII. CONCLUSIONS

We have presented the first fully coupled approach that simultaneously computes the HDR irradiance map as well as accurate displacement fields. Especially the introduced feedback in this joint computation of the irradiances and the optic flow turns out to be very beneficial. We have shown that the proposed strategy is clearly preferable to a decoupled optic flow-based pre-alignment, presented in previous work. In contrast to existing patch-based approaches, our dense displacement maps represent the underlying motion with sub-pixel precision. This is particularly of great importance for tasks that require precise motion estimations, such as object removal, artificial motion blur generation or bullet time effects in an HDR context. To conclude, we do not only produce a nice-looking HDR image, we additionally compute accurate motion estimations of the dynamic scene. In summary, whenever optic flow is required in an exposure series, we recommend to consider a *simultaneous* computation of the HDR image and the flow fields as demonstrated by this work.

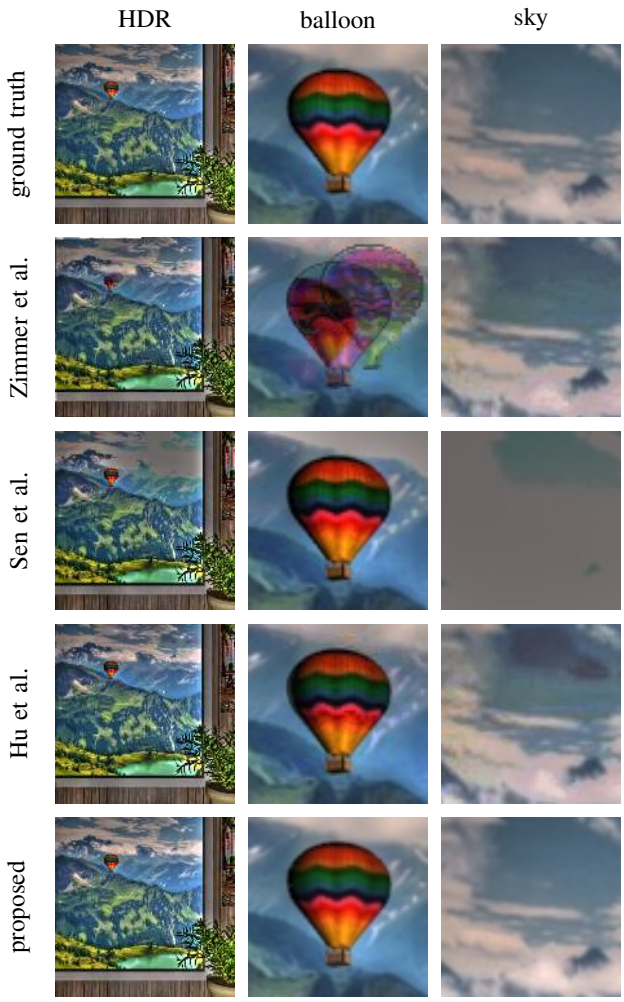
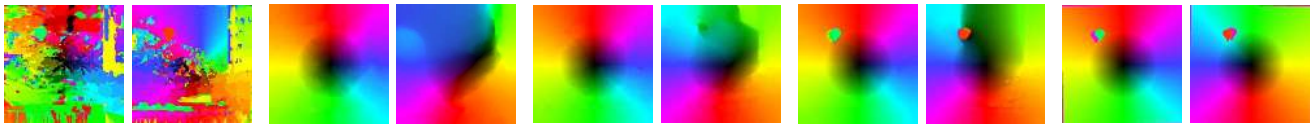


Fig. 4. Visual comparison of the irradiance maps. The full HDR composites (*first column*) with their zoom-ins (*second and third column*) illustrate clearly that our method produces the most accurate irradiances.

Moreover, thanks to the reliance on optic flow we believe that our method is especially suited for continuous image sequences. Thus, we intend to test and adapt it to the production of HDR videos [6], [25]. Last but not least, the explicit integration of the response function in our model generally allows to additionally compute the CRF on-the-fly if it is unknown. First experiments in this direction show very promising results.

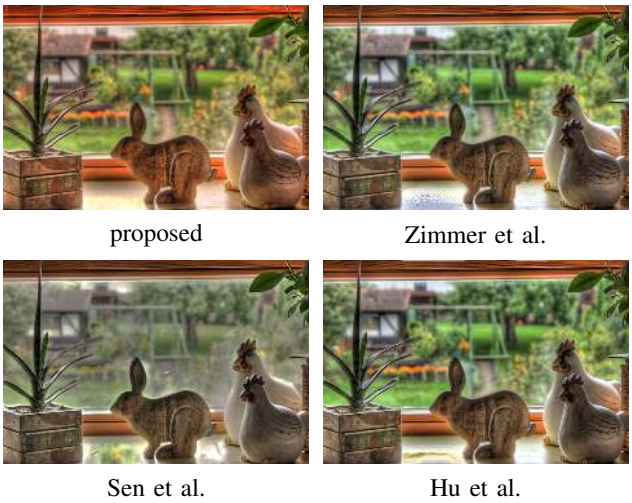
ACKNOWLEDGEMENTS

We thank Zimmer et al., Sen et al. and Hu et al. for providing access to their algorithms. Furthermore, we gratefully



Hu et al. Zimmer et al. Demetz et al. proposed ground truth

Fig. 3. Compared to all other methods, our approach produces the most accurate displacement maps. Thanks to our joint optimisation, we are able to estimate the motion of the balloon (besides the camera motion) with high precision even though it is hardly visible in some of the LDR images.



proposed Zimmer et al.

Sen et al. Hu et al.

Fig. 5. Visual comparison of the HDR results for the real-world example from Fig. 1. While the competing methods have especially problems with the highlights on the window sill, our method yields a convincing irradiance map.

acknowledge the partial funding by the Deutsche Forschungsgemeinschaft (DFG) through a Gottfried Wilhelm Leibniz Prize for Joachim Weickert.

References

[1] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *Proc. 48th IS&T Annual Conference*, Springfield, VA, May 1995, pp. 442–448.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH*, Los Angeles, CA, Aug. 1997, pp. 369–378.

[3] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd ed. Morgan Kaufmann, 2010.

[4] A. Srikantha and D. Sidibé, "Ghost detection and removal for high dynamic range images: Recent advances," *Signal Processing: Image Communication*, vol. 27, no. 6, pp. 650–662, Jul. 2012.

[5] G. Ward, "Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures," *Journal of Graphics, GPU, and Game Tools*, vol. 8, no. 2, pp. 17–30, 2003.

[6] S. B. Kang, M. Uyttendaeleand, S. Winder, and R. Szeliski, "High dynamic range video," in *Proc. ACM SIGGRAPH*, San Diego, CA, Jul. 2003, pp. 319–325.

[7] N. Menzel and M. Guthe, "Freehand HDR photography with motion compensation," in *Proc. Vision, Modeling, and Visualization Conference*, Saarbrücken, Germany, Nov. 2007, pp. 127–134.

[8] S. Mangiat and J. Gibson, "High dynamic range video with ghost removal," in *Proc. Applications of Digital Image Processing XXXIII*, vol. 7798, San Diego, CA, Aug. 2010.

[9] J. Hu, O. Gallo, and K. Pulli, "Exposure stacks of live scene with hand-held cameras," in *Computer Vision – ECCV 2012*, ser. Lecture Notes in Computer Science, K. M. Lee, Y. Matsushita, J. M. Rehg, and Z. Hu, Eds. Berlin: Springer, 2012, vol. 3024, pp. 499–512.

[10] J. Hu, O. Gallo, K. Pulli, and X. Sun, "HDR deghosting: How to deal with saturation?" in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, Jun. 2013, pp. 1063–1170.

[11] H. Zimmer, A. Bruhn, and J. Weickert, "Freehand HDR imaging of moving scenes with simultaneous resolution enhancement," *Computer Graphics Forum*, vol. 30, no. 2, pp. 405–414, Apr. 2011.

[12] M. Gupta, D. Iso, and S. Nayar, "Fibonacci exposure bracketing for HDR imaging," in *Proc. IEEE International Conference on Computer Vision*, Sydney, Australia, Dec. 2013, pp. 1473–1480.

[13] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtma, "Robust patch-based HDR reconstruction of dynamic scenes," in *Proc. ACM SIGGRAPH*, vol. 31, Los Angeles, CA, Aug. 2012, pp. 203:1–203:11.

[14] M. Robertson, S. Borman, and R. Stevenson, "Estimation-theoretic approach to dynamic range improvement using multiple exposures," *Journal of Electronic Imaging*, vol. 12, no. 2, pp. 219–228, Apr. 2003.

[15] H. Zimmer, A. Bruhn, and J. Weickert, "Optic flow in harmony," *International Journal of Computer Vision*, vol. 93, no. 3, pp. 368–388, Jul. 2011.

[16] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Two deterministic half-quadratic regularization algorithms for computed imaging," in *Proc. 1994 IEEE International Conference on Image Processing*, vol. 2. Austin, TX: IEEE Computer Society Press, Nov. 1994, pp. 168–172.

[17] R. M. Rameshand, S. Chaudhuri, and R. Velmurugan, "High dynamic range imaging under noisy observations," in *Proc. 18th IEEE International Conference on Image Processing*, Brussels, Belgium, Sep. 2011, pp. 1333–1336.

[18] J. Weickert, "Coherence-enhancing diffusion of colour images," *Image and Vision Computing*, vol. 17, no. 3–4, pp. 199–210, Mar. 1999.

[19] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Computer Vision – ECCV 2004, Part IV*, T. Pajdla and J. Matas, Eds. Berlin: Springer, 2004, vol. 3024, pp. 25–36.

[20] C. W. Cryer, "The solution of a quadratic programming problem using systematic overrelaxation," *SIAM Journal on Control*, vol. 9, no. 3, pp. 385–392, 1971.

[21] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," in *Proc. ACM SIGGRAPH*, San Antonio, TX, Jul. 2002, pp. 249–256.

[22] M. D. Grossberg and S. K. Nayar, "Modeling the space of camera response functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1272–1282, Oct. 2004.

[23] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[24] O. Demetz, D. Hafner, and J. Weickert, "The complete rank transform: A tool for accurate and morphologically invariant matching of structures," in *Proc. 2013 British Machine Vision Conference*, T. Burghardt, D. Damen, W. Mayol-Cuevas, and M. Mirmehdi, Eds. Bristol, UK: BMVA Press, Sep. 2013.

[25] N. K. Kalantari, E. Shechtman, C. Barnes, S. Darabi, D. B. Goldman, and P. Sen, "Patch-based high dynamic range video," in *Proc. ACM SIGGRAPH ASIA*, Hong Kong, Nov. 2013, pp. 202:1–202:8.