



Ghost Detection and Removal for High Dynamic Range Images: Recent Advances

Abhilash Srikantha, Désiré Sidibé

► To cite this version:

Abhilash Srikantha, Désiré Sidibé. Ghost Detection and Removal for High Dynamic Range Images: Recent Advances. *Signal Processing: Image Communication*, Elsevier, 2012, pp.10.1016/j.image.2012.02.001. 10.1016/j.image.2012.02.001 . hal-00671579

HAL Id: hal-00671579

<https://hal-univ-bourgogne.archives-ouvertes.fr/hal-00671579>

Submitted on 9 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ghost Detection and Removal for High Dynamic Range Images: Recent Advances

Abhilash Srikantha, Désiré Sidibé*

Université de Bourgogne - LE2I, CNRS, UMR 6306, 12 rue de la fonderie, 71200 Le Creusot, France

Abstract

High dynamic range (HDR) image generation and display technologies are becoming increasingly popular in various applications. A standard and commonly used approach to obtain an HDR image is the multiple exposures fusion technique which consists of combining multiple images of the same scene with varying exposure times. However, if the scene is not static during the sequence acquisition, moving objects manifest themselves as ghosting artefacts in the final HDR image. Detecting and removing ghosting artefacts is an important issue for automatically generating HDR images of dynamic scenes. The aim of this paper is to provide an up-to-date review of the recently proposed methods for ghost-free HDR image generation. Moreover, a classification and comparison of the reviewed methods is reported to serve as a useful guide for future research on this topic.

Key words: High dynamic range images, Exposures fusion, Ghost detection, Ghost removal

1. Introduction

Conventional digital cameras can only capture a limited luminance dynamic range and most monitors and displaying media also have limited dynamic range due to the limited capacity of digital sensors, to about two orders of magnitude. On the other hand, the dynamic range of real world scenes varies over several orders of magnitude, up to ten. As a consequence, when taking a photograph of a scene bright areas tend to be overexposed while dark regions tend to be underexposed. These bright and dark regions appear saturated in the image. An auto-exposure mechanism can be used to minimize the number of saturated pixels or to correctly expose a region of interest such as a face, but fails to correctly expose the entire image and recover the whole dynamic range of the captured scene.

To enlarge the dynamic range spanned by conventional cameras a very interesting and powerful technique has been developed in the last few years: high dynamic range imaging. The obtained images are called high dynamic range (HDR) images and represent the scene more faithfully than conventional low dynamic range (LDR) images [1].

*Corresponding author

Email addresses: aalibash@gmail.com (Abhilash Srikantha), dro-desire.sidibe@u-bourgogne.fr (Désiré Sidibé)

HDR images can be obtained using either hardware or software methods. Hardware methods to capture HDR images include the use of multiple imaging devices, or devices with special sensors [1]. For example, Mitsunaga and Nayar describe the process of spatially varying pixel exposures [2]. They place an optical mask adjacent to a conventional image detector array. The mask has a pattern with spatially varying transmittance, thus adjacent pixels on the detector are given different exposures to the scene. Other methods use a different CCD design to achieve HDR imaging. For example, Wen [3] and Street [4] use a CCD camera where each detector cell includes two sensing element of different size. This way, two measurements are made within each cell and are combined on-chip to produce an HDR image. Recently, Tublin *et al.* [5] develop a new camera design that first measures the difference between adjacent pixels pairs and then quantizes the differences appropriately to capture an HDR image. Unfortunately, these devices are just beginning to enter the market and are so far expensive for mainstream consumers. Moreover, due to the limitations of digital image sensors, it is not generally possible to capture the full dynamic range of a scene with a single exposure [1].

The most common method for HDR image generation is based on the combination of multiple distinct exposures. The motivation behind this technique is that different exposures capture different dynamic range characteristics of the scene. For instance, bright regions are captured in the shorter exposures while dark regions are captured in the longer ones. Using pixel values, shutter times, and the camera response function, it is possible to estimate a scene-referred, high dynamic range radiance map that captures all details of the scene. However, this simple and easy to implement technique suffers from two main problems: i) *Misalignment*: global camera motion, from hand-held camera for instance, results in misaligned images that cause the combined HDR image to look blurry. ii) *Ghosting*: moving objects in the scene while capturing the images, will appear in different locations in the combined HDR image, creating what are called *ghost* or *ghosting* artefacts.

The first problem can be solved by placing the camera on a tripod or by using an image registration method. In particular, the median threshold bitmap (MTB) technique proposed by Ward [6] is an efficient solution. The method is fast and can accurately recover the small displacements between images. Other registration methods based on keypoints extraction and matching can be used as well. The most used keypoints detectors are Harris corners [7] and SIFT features [8]. The second problem is a more severe limitation of the multiple exposures technique since motion is hardly avoidable in outdoor environments. This drawback limits the application of HDR imaging in practice and a lot of work have been carried out to detect and remove ghosting artefacts in dynamic environments.

The aim of this paper is to provide a comprehensive survey of the most recent methods that have been developed to deal with the ghost problem in HDR image generation. Moreover, a classification and comparison of different methods is described to highlight the performances and advantages of each technique. For a general overview of all aspects of HDR image acquisition and reproduction, the reader is referred to the excellent book by Reinhard *et al.* [1] which entirely covers the topic.

The remainder of this paper is organized as follows. First, the multiple exposures fusion technique and the ghost problem in HDR image generation are described in Section 2. Next, different ghost detection techniques are presented in Section 3 and ghost removal methods are discussed in Section 4. Then, a classification and comparison of the surveyed methods is proposed in Section 5. The paper ends with the conclusions in Section 6.

Table 1: Notations used in this paper

Notation	Explanation
N	Number of exposures (images)
$U \times V$	Resolution of each LDR image
$\{L_k\}_{k=1\dots N}$	Set of low dynamic range (LDR) images
$\{\Delta t_k\}_{k=1\dots N}$	Exposure time associated with L_k
Z_{uv}^k	Pixel value at position (u, v) in exposure L_k
E_{uv}^k	Estimated radiance value at position (u, v) in exposure L_k
$w(Z_{uv}^k)$	Weight of pixel at position (u, v) in exposure L_k
G	Ghost map
$f()$	Camera response function

2. The Multiple Exposures Combination Technique and The Ghost Problem

High dynamic range images may be captured from real scenes or rendered by computer graphics techniques. The most common approach to obtain an HDR image is to take multiple images of the same scene with different exposure times, and combine them into a single HDR image [9, 1]. The multiple exposures technique is based on the observation that taking multiple images with different exposures, each pixel will be properly exposed in at least one image. Therefore, an HDR image is obtained by appropriately combining the LDR images. In the following subsections, we start by providing explanations of the important terminologies and notations used in this paper, followed by a description of the HDR image generation methods and the ghost problem.

2.1. Definitions and Notations

We first present important terminologies that will be referred to in the rest of the manuscript.

Dynamic range of an image can be defined as the ratio between the lightest and darkest pixels. For a camera, the dynamic range is the ratio of the luminance that just saturates the sensor and the luminance that lifts the camera response to one standard deviation above the noise level [10].

Radiance is a radiometric quantity that measures the amount of light passing through or emitted from a particular point in a given direction. For a digital camera, the radiance values correspond to the physical quantity of light incident on each element of the sensor array.

Camera response function of a digital camera, $f()$, is a function that maps the radiance values of a scene to the pixel values in the captured image. This function models the effect of non-linearities introduced in the image acquisition process such as non-linear dynamic range compression and quantization [11].

The various notations used henceforth are summarized in Table 1.

2.2. Multiple Exposures Combination

The fusion of a set of LDR images into an HDR image can be achieved in different methods which can be classified into two main approaches: fusion in the radiance domain and fusion in the image domain.

2.2.1. Fusion in the radiance domain

This HDR image generation method introduced in [12, 9, 13] consists of three steps. First, the camera response function is recovered to bring the pixel brightness values into the radiance domain. This function models the effect of non-linearities introduced in the image acquisition process. Since the camera response function is not always provided by manufacturers, different methods are proposed for its estimation from a sequence of differently exposed images [9, 13, 14]. Secondly, all radiance maps are combined into an HDR image encoded specially to store the pixel values that span the entire tonal range of the scene. Finally, a tone mapping operator is used to make the HDR image displayable on common low dynamic range monitors [15, 16, 17].

More precisely, let $\{L_k\}_{k=1 \dots N}$ be a set of N images with exposure times $\{\Delta t_k\}_{k=1 \dots N}$. Given the camera response function $f()$, the HDR image is computed as the weighted average of pixels values across exposures using the following equation:

$$R_{uv} = \frac{\sum_{k=1}^N w(Z_{uv}^k) f^{-1}(Z_{uv}^k) / \Delta t_k}{\sum_{k=1}^N w(Z_{uv}^k)}, \quad (1)$$

where R is the combined radiance map, Z_{uv}^k is the pixel value at location (u, v) in exposure L_k and $w(Z_{uv}^k)$ is the weight of that pixel. The weighting function $w()$ is designed to reduce the influence of unreliable pixels such as saturated ones.

Several methods have been proposed to select a good weight function. A good overview can be found in [1, 18]. Mann and Picard [12] propose to use the derivative of the camera response function using the argument that the reliability of pixel values is correlated with the camera sensitivity to light changes. Debevec and Malik [9] use a simple hat shaped function based on the assumption that the pixels that are in the middle of the range are more reliable. Mitsunaga and Nayar [13] multiply Mann and Picard weighting function by the linearized camera output since signal-to-noise ratio increases with signal intensity. Finally, Ward suggests to multiply Mitsunaga and Nayar's weighting function with a broad hat filter to exclude unreliable pixels near extremes [1].

In order to display the obtained HDR image on a low dynamic range monitor, a tone mapping operator is applied. Tone mapping techniques can be classified into global and local methods. Global methods specify one mapping curve that applies equally to all pixels, while local methods provide a space-varying mapping curve that takes into account the local content of the image [19]. For more details about tone mapping techniques, the reader is referred to [1, 15, 16, 17].

2.2.2. Fusion in the image domain

Alternative methods combine multiple exposures directly without the knowledge of the camera response function [20, 21, 22, 23]. These methods combine LDR images by preserving only the best parts of each exposure. The final HDR image is obtained as a weighted average of pixel values across exposures:

$$I_{uv}^C = \sum_{k=1}^N w(Z_{uv}^k) Z_{uv}^k, \quad (2)$$

where I^C is the composite image.

The choice of the weighting function is crucial to get good and accurate results. Mertens *et al.* [20] combine multiple exposures using contrast, saturation and well-exposedness as parameters for weighting functions. They also use a Laplacian pyramid blending framework to avoid artefacts in the composite image. Zhang and Cham [23] use gradient information to compute the

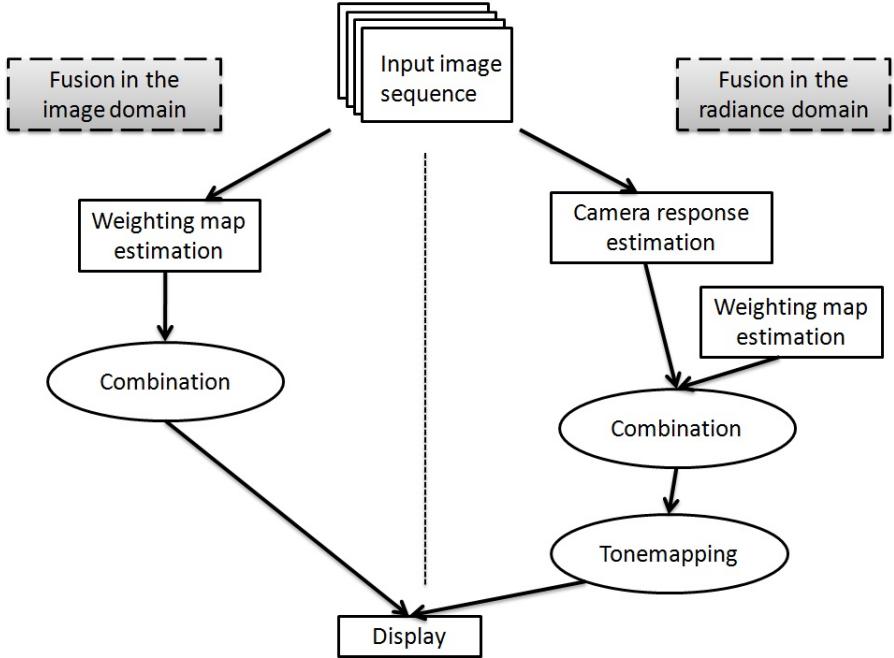


Figure 1: HDR image generation process.

weights. Raman and Chaudhuri [22] use a bilateral filter to define the weighting function and Goshtasby [21] uses an entropy measure defined on image blocks to combine multiple exposures.

The two different HDR image generation processes are depicted in Fig. 1. The performance of the methods that combine images in the radiance domain highly relies on an accurate estimation of the camera response function, which is sensitive to image noise and misalignment. Moreover, these methods require tone mapping operators for HDR images reproduction. Methods that combine exposures in the image domain are more efficient since they avoid the estimation of the camera response function and do not require tone mapping. They directly produce a tonemapped-like HDR image. However, with methods of the first approach, a true HDR radiance map is obtained in the combination step which contains the whole dynamic range of the captured scene. This radiance map can later be used for different processing or display applications.

2.3. The Ghost Problem

The main limitation of the multiple exposures combination technique is the requirement of a complete static scene when capturing the images. Indeed, any object movement in the scene can cause ghosting artefacts in the resulting HDR image. The ghosting problem is a severe limitation of the multiple exposures technique since motion can hardly be avoided in outdoor environment which contain moving entities such as automobiles, people and motion caused naturally; due to wind for example. Even a very small or limited movement will produce a very noticeable

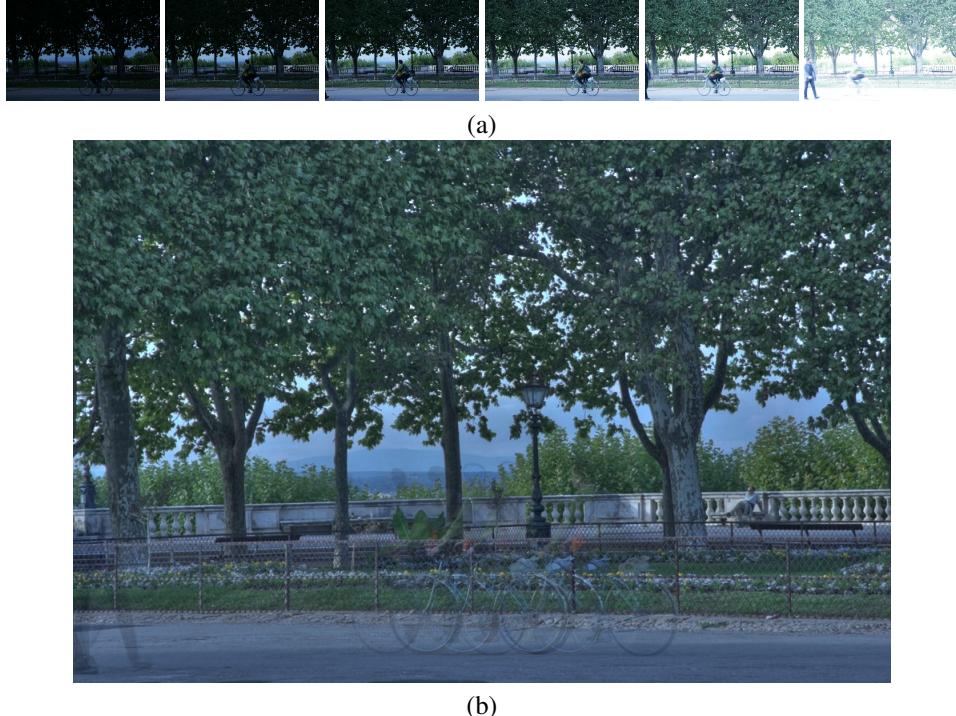


Figure 2: The ghost problem. (a) Six exposures of a dynamic scene; (b) HDR image generated by the multiple exposure combination technique showing ghosting artefacts.

artefact in the combined HDR image. Therefore, detecting and removing ghosting artefacts is an important issue for the automatic generation of HDR images of dynamic scenes.

An example of HDR image generated with moving object and ghosting is shown in Fig. 2. The ghosting artefacts created by the moving cyclist are visible in Fig. 2(b).

3. Ghost Detection Methods

Several methods have been developed in literature to solve the ghost problem in dynamic scenes. Most of the methods employ a two-step strategy: first, regions affected by ghost are detected, then ghost artefacts are removed. Therefore, we will first describe various techniques that have been proposed for ghost detection in this section, and ghost removal techniques will be discussed in Section 4.

Ghost detection methods are based on motion detection in the exposures sequence. Basically, we can identify two type of motions in a dynamic scene: (i) a moving object on a static background, e.g. moving people or cars; (ii) a moving background with static or dynamic objects, e.g. windblown leaves or waves. Some of the following methods can detect only the first type of motion while others can detect both.

3.1. Variance based ghost detection

This method detects ghost regions based on a weighted variance measure [1, 24]. First, the camera response function is estimated and the radiance maps are computed. Then, a *Variance Image* (VI) is generated by evaluating the variance of radiance values at each spatial location (u, v) :

$$\text{VI}_{uv} = \frac{\sum_{k=1}^N w(Z_{uv}^k)(E_{uv}^k)^2 / \sum_{k=1}^N w(Z_{uv}^k)}{(\sum_{k=1}^N w(Z_{uv}^k)E_{uv}^k)^2 / (\sum_{k=1}^N w(Z_{uv}^k))^2} - 1, \quad (3)$$

where the weighting function is defined by:

$$w(Z_{uv}^k) = \begin{cases} Z_{uv}^k & \text{if } Z_{uv}^k \leq 127 \\ 255 - Z_{uv}^k & \text{if } Z_{uv}^k > 127 \end{cases}. \quad (4)$$

As regions affected by movement exhibit high variance, the VI can be used as a likelihood measure for intra-image movements. Regions where this local variance measure is above a defined threshold are detected as ghost regions:

$$G_{uv} = \begin{cases} 1 & \text{if } \text{VI}_{uv} \geq \text{threshold} \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

In [24], the threshold is set to 0.18 for the normalized VI. For color images, the VI is calculated as the maximum over the three color channels and morphological operations (erosion and dilation) are applied to remove outliers, false detections and to obtain closed and well defined structures.

3.2. Entropy based ghost detection

In their work, Jacobs *et al.* [24] define two types of motions: high contrast movement and low contrast movement. The former type of motion occurs when the moving object is different from the background and can be detected using the variance measure above. The latter type of motion occurs when the dynamic object and the background are similar in color and cannot be detected by the variance measure. Hence, they introduce another measure derived from entropy. First, a local neighbourhood based entropy map is computed for each LDR image. For each pixel (u, v) in L_k , the entropy is calculated from a local histogram computed in the window of size $(2r + 1) \times (2r + 1)$ around (u, v) :

$$H_{uv}^k = - \sum_{x=0}^{B-1} P(X = x) \log(P(X = x)), \quad (6)$$

where B is the total number of bins of the histogram and the probability $P(X = x)$ is obtained from the normalized histogram. It is to be noted that the product term in Eq. (6) is set to zero if $P(X = x) = 0$.

An *Uncertainty Image* (UI) is then derived from the weighted difference of the precomputed entropy images as follows:

$$\text{UI}_{uv} = \sum_{k=1}^N \sum_{l=1}^{l < k} \frac{v^{kl}}{\sum_{k=1}^N \sum_{l=1}^{l < k} v^{kl}} h_{uv}^{kl}, \quad (7)$$

with $h_{uv}^{kl} = |H_{uv}^k - H_{uv}^l|$ and $v^{kl} = \min(w(Z_{uv}^k), w(Z_{uv}^l))$. The weighting function is defined by:

$$w(Z_{uv}^k) = \begin{cases} (Z_{uv}^k \times 0.9/127) + 0.05 & \text{if } Z_{uv}^k \leq 127 \\ ((255 - Z_{uv}^k) \times 0.9/127) + 0.05 & \text{if } Z_{uv}^k > 127 \end{cases}. \quad (8)$$

This uncertainty image is used to find ghost regions based on thresholding:

$$G_{uv} = \begin{cases} 1 & \text{if } UI_{uv} \geq \text{threshold} \\ 0 & \text{otherwise} \end{cases}. \quad (9)$$

The threshold value is set to 0.7 for a normalized UI computed from the entropy images obtained with $r = 40$ and $B = 200$ [24]. Similarly to the variance measure, for color images the UI is calculated as the maximum over the three color channels.

3.3. Prediction based ghost detection

In this method [25], the deviation between the predicted intensity value of a pixel and the actual intensity is used as a measure to decide between ghost and non-ghost pixels. More precisely, given two images L_k and L_l , one tests if the value of a pixel in L_l is well approximated by the predicted value from L_k using the estimated camera response function. The prediction is based on the following equation:

$$\check{Z}_{uv}^l = f\left(\frac{\Delta t_l}{\Delta t_k} f^{-1}(Z_{uv}^k)\right), \quad (10)$$

where $f()$ is the camera response function and, Δt_k and Δt_l are the exposure times of L_k and L_l , respectively.

For each pair of consecutive input LDR images, pixels that show a significant difference between the predicted value and the actual one, are marked as ghost pixels in the corresponding ghost map:

$$G_{uv} = \begin{cases} 1 & \text{if } |\check{Z}_{uv}^l - Z_{uv}^l| \geq \text{threshold} \\ 0 & \text{otherwise} \end{cases}. \quad (11)$$

The default value for the threshold is not given by the author in [25]. For the results shown in Section 5, we set the threshold to 5.

3.4. Pixel order relation

This method relies on the order relation between pixels values in differently exposed images to find ghost areas [26]. More precisely, it is possible to relate pixel values to radiance values using the camera response function:

$$Z_{uv}^k = f(E_{uv}^k \Delta t_k).$$

Then, assuming that $f()$ is monotonic, which is a reasonable assumption since an increase in radiance values always produces an increased or equal recorded pixel values [18], it can be shown that for each pixel location (u, v) the intensity values in different exposures must satisfy:

$$Z_{uv}^k \leq Z_{uv}^l, \text{ if } \Delta t_k < \Delta t_l. \quad (12)$$

Therefore, if the input LDR images are arranged in increasing order of exposure times, the ghost map is generated by the following equation:

$$G_{uv} = \begin{cases} 0 & \text{if } Z_{uv}^1 \leq Z_{uv}^2 \leq \dots \leq Z_{uv}^N \\ 1 & \text{otherwise} \end{cases}. \quad (13)$$

As the above order relation works only if the pixel is not under- or over-exposed, saturated pixels are excluded from the ghost map computation.

3.5. Multi-level thresholding based ghost detection

This method detects moving areas in the scene based on multi-level threshold maps [27]. Roughly speaking, it imposes the condition that the grey levels at a particular pixel location must exhibit a non-decreasing property when the images are scanned from lowest to highest exposure values.

First, for each image L_k , a set of P threshold values is found such that the image is divided into P levels, each having the same number of pixels. The multi-level threshold maps $TM^k, k = 1 \dots N$ are then computed by classifying the intensity values of L_k into P levels using these thresholds. The ghost map estimate is generated using the multi-threshold maps as follows:

$$G_{uv} = \begin{cases} 1 & \text{if } |TM_{uv}^{ref} - TM_{uv}^k| \geq 1, k \neq \text{ref} \\ 0 & \text{otherwise} \end{cases}, \quad (14)$$

where the mid-exposure is taken as reference image L_{ref} . In the experiments, we divide the images into 8 levels.

3.6. Bitmap based ghost detection

In this method, ghost regions are detected based on median bitmaps which impose relations between pixels in each single exposure [28]. The algorithm relies on the fact that if a pixel is not affected by ghost, then its relation to the median intensity of the image must be the same in all LDR images. For each exposure L_k , a binary median bitmap M_k is obtained by thresholding L_k based on its median pixel value. Dark regions of M_k indicate pixels whose values are lesser than or equal to the median intensity value of L_k . Bright regions of M_k indicate the pixels whose values are greater than the median intensity value.

The ghost map is recovered from the median bitmaps as follows:

$$G_{uv} = \begin{cases} 0 & \text{if } S_{uv} = 0 \text{ or } S_{uv} = N \\ 1 & \text{otherwise} \end{cases}, \quad (15)$$

where S_{uv} is the sum of the bitmaps values at location (u, v) : $S_{uv} = \sum_{k=1}^N M_{uv}^k$.

3.7. RANSAC based ghost detection

In this method [29], patches of ghost regions are detected using RANSAC procedure [30]. The method is based on the fact that the intensity values at any location (u, v) in any two input images L_k and L_l are related by:

$$\frac{Z_{uv}^k}{\Delta t_k} = \frac{Z_{uv}^l}{\Delta t_l}. \quad (16)$$

Apart from saturated pixels, the above equation deviates only at locations affected by ghost. However, in order to be robust to noise, the processing is performed on a patch level. First, saturated regions of each exposure are computed and the least saturated image is selected as reference. Then, in order to determine if an $r \times r$ patch in L_k is affected by ghost, log intensities of the patch in L_k are plotted against the log intensities of the corresponding patch in the reference image. A best fit line through the plot is obtained by the RANSAC procedure and the percentage number of outliers is calculated using a distance threshold. If this percentage is greater than a predefined threshold, the patch is decided to be affected by ghost. In [29], the distance threshold is set to 0.75 and a 40×40 patch is affected by ghost if its percentage of outlier is greater than 0.5% of the patch's size.

3.8. Graph-Cuts based ghost detection

In this method [31], joint probability densities are employed to roughly detect ghost regions and these regions are further refined using energy minimization based on graph-cuts methods [32]. First, joint intensity histograms are constructed in order to study the intensity correspondence within various exposures. Joint histogram $P^{ref,k|c}$ for the color channel $c \in \{R, G, B\}$ between the reference image L_{ref} and another image L_k is constructed as:

$$P_{ij}^{ref,k|c} = \sum_{u=1}^U \sum_{v=1}^V G_{uv}^k \times T[(i, j) == (Z_{uv}^k, Z_{uv}^{ref})], \quad (17)$$

where $T[.]$ is one if the argument is true, and zero otherwise. The ghost map G_{uv}^k is initialized to one for the first iteration. Next, each joint histogram $P^{ref,k|c}$ is convolved with a 5×5 Gaussian filter and normalized to represent a pdf.

For each exposure a ghost map is defined by:

$$G_{uv}^k = \begin{cases} 1 & \text{if } P^{ref,k|c}(Z_{uv}^{ref|c}, Z_{uv}^{k|c}) \leq \text{threshold} \\ 0 & \text{otherwise} \end{cases}. \quad (18)$$

The default value for the threshold is set to 10^{-5} . However, because the ghost regions estimated by the above equation are noisy, they are refined by an energy minimization approach using graph-cuts [32]. The energy to minimize is defined as:

$$E(f_n) = \sum_p D_p(f_n(p)) + \sum_p \sum_{q \in N(p)} V_{pq}(f_n(p), f_n(q)), \quad (19)$$

where the boolean label $f_n(p) \in \{0, 1\}$ represents whether a pixel $p = (u, v)$ in exposure L_n is affected by ghost or not. $f_n(p) = 0$ if the pixel is a ghost and $f_n(p) = 1$ otherwise. $N(u, v)$ represents the neighborhood of p .

The energy $E(f_n)$ is composed of two terms: a data cost function D and a smoothness term V_{pq} . The data cost term is equal to zero if the label $f_n(p)$ assigned to a pixel matches with its binary value in the ghost map G :

$$D_p(f_n(p)) = \begin{cases} 0 & \text{if } f_n(p) = 1 - G_{uv}^n \\ \beta & \text{otherwise} \end{cases}, \quad (20)$$

where β is a constant value set to 2.5 in [31].

The smoothness function V_{pq} is based on the intensity difference between neighboring pixels and is defined as:

$$V_{pq}(f_n(p), f_n(q)) = \lambda_{pq} \cdot \min(|f_n(p) - f_n(q)|, V_{max}), \quad (21)$$

$$\lambda_{pq} = \begin{cases} \lambda_L & \text{if } (|Z_{uv}^{ref} - Z_{u'v'}^{ref}| \leq \eta) \text{ or } (|Z_{uv}^n - Z_{u'v'}^n| \leq \eta) \\ \lambda_S & \text{otherwise} \end{cases}. \quad (22)$$

The values λ_L and λ_S are chosen such that $\lambda_L > \lambda_S$ and in order to emphasize more smoothness (larger λ_L) if the difference of intensity values between neighboring pixels is smaller than the threshold η . The default values for the different parameters are given as $V_{max} = 1$, $\eta = 5$, $\lambda_L = 3$ and $\lambda_S = 1$ in [31]. The total energy $E(f_n)$ is optimized using the graph-cuts method [32]. The optimized label map $f_n()$ is used to update the ghost map in Eq. (17) and the process is repeated iteratively until convergence. In [31], the authors found that two or three iterations are sufficient for convergence.

3.9. Motion compensation based ghost detection

Since ghosting artefacts are mainly due to moving objects in the scene, it is possible to estimate ghost affected areas if the motion between exposures is known. First, a global motion between two exposures L_k and L_{k+1} is found by estimating an affine transform that maps L_k into L_{k+1} . Then, a gradient-based optical flow technique such as Lucas and Kanade algorithm [33] is used to compute a dense local motion field. The estimated motion parameters are used to warp pixels in the exposures so that all scene features are correctly aligned [34, 35]. Differences between a warped image of L_k and L_{k+1} indicates ghost regions. More specifically, one needs to compensate the differences in exposure values before motion estimation because motion estimation algorithms are based on the so-called *brightness constancy assumption* which clearly do not hold for images with different exposure times. Therefore, differences in exposure values are compensated using the camera response function to bring pixel values into the radiance domain [34]. After this radiometric alignment step, the motion parameters are estimated based on the following equation:

$$E^k(\mathbf{x}) = E^{k+1}(\mathbf{x} + \mathbf{u}(\mathbf{x}; \mathbf{p}_m)), \quad (23)$$

where $\mathbf{x} = (u, v)$ denotes the spatial image position, $\mathbf{u}(\mathbf{x}; \mathbf{p}_m)$ the displacement vector at that point, and \mathbf{p}_m is a vector representing the parameters of the motion model. For example, if we assume an affine motion model defined by $\mathbf{a} = [a_1, a_2, a_3, a_4, a_5, a_6]^T$, the motion vector can be written as follows:

$$\mathbf{u}(\mathbf{x}; \mathbf{p}_m) = \begin{bmatrix} 1 & u & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & u & v \end{bmatrix} \mathbf{a}. \quad (24)$$

The parameter vector \mathbf{a} is estimated using, for instance, Lucas and Kanade algorithm [33, 36] and the result is used to warp L_k into L_{k+1} .

The ghost problem is not restricted to HDR image generation but also appears in the process of image stitching or mosaic construction from multiple images. For those applications, one first need to estimatite the homography between two images (this can be done by extracting keypoints and matching them in the overlapping area), and then, align the images to a reference frame [37]. If the scene is not static, moving regions of the composite image will contain combinations of pixels values from different parts of the scene, hence, creating ghosting artefacts. Shum and Szeliski [38] use optical flow estimation to remove ghosting due to small misregistration while constructing panoramas. Uyttendaele *et al.* [39] propose a method to detect and remove ghost in image mosaics which is based on the detection of moving regions in the overlap areas of the images to combine. The detected regions are then treated as nodes in a graph and a vertex cover algorithm is used to selectively remove all but one instance of each object.

It is important to mention that in the case of HDR image generation, the motion between images is usually limited since the images are taken continuously using, for instance, the auto-bracketing function embedded in many cameras today. Futhermore, if a tripod is used, the only motion to estimate is that of moving objects in the scene. Therefore, optical flow methods can be applied after radiometric alignment of the exposures.

4. Ghost Removal Techniques

Removing ghosting artefacts in the combined HDR image is the ultimate aim of any method that address the ghost problem. Different methods produce different results and can be classified into two main categories. We can first distinguish methods which remove ghosting artefacts while keeping a single occurrence of the moving object. For example, in the case presented in

Fig. 2(b), those methods will keep the moving cyclist at a fixed location in the final HDR image. Other methods, on the contrary, will completely remove the moving object in the composite image.

4.1. Keeping a single occurrence of moving object

If the moving object is of interest for the photographer, then it is desirable to keep it at a fixed location in the final HDR image, avoiding ghosting artefact due to multiple appearances at different locations, rather than completely removing it.

Many ghost removal techniques are based on the detected ghost map and the simplest approach is to apply the standard multiple exposure fusion method in ghost-free regions while selecting a single reference exposure in ghost affected areas. This approach is based on the observation that each exposure is self-consistent [1]. The reference exposure is typically the image that is least saturated [1, 24] or the image whose ghost regions are best kept in range [25]. Another approach developed by Gallo *et al.* [29] is to determine the correct number of exposures to use in different ghost affected areas. This number is obtained, for each $r \times r$ patch ($r = 40$), as the number of images in which the patch does not deviate from the patch in the reference image. The algorithm then builds the HDR image using different number of exposures on each detected ghost region.

However, using a single reference exposure introduces new artefacts in the combined HDR image. Indeed, it creates seams at ghost regions boundaries and these boundary effects have to be removed. For a seamless composition of exposures, Pece and Kautz [28] and Mertens *et al.* [20] use a Laplacian pyramid blending framework. This blending technique works at multiple resolutions using a pyramidal image decomposition for seamlessly blending two images [40]. The input images are decomposed into a Laplacian pyramid, which basically contains band-pass filtered versions at different scales and blending is performed for each level separately. Gallo *et al.* [29] use a gradient domain approach to avoid boundary effects in the final HDR image. The method is based on estimating an image whose gradient is closest, in the mean squared error sense, to the gradient of the estimated radiance map. This results in solving a partial differential equation subject to some conditions which are solutions of Poisson equation [16].

Both Laplacian pyramid blending and Poisson editing frameworks are used to avoid boundary effects introduced by using a single reference exposure in ghost affected regions. A simpler method which produces good results is based on weights adaptation. The idea is to adjust pixel weights based on the deviation from the reference image. More precisely, a pixel whose value differs significantly from the reference value will be assigned a lower weight according to the following equation:

$$w(Z_{uv}^k) = \frac{[a(Z_{uv}^{ref})]^2}{[a(Z_{uv}^{ref})]^2 + [(f^{-1}(Z_{uv}^k) - f^{-1}(Z_{uv}^{ref})) / f^{-1}(Z_{uv}^{ref})]^2}, \quad (25)$$

where $a()$ is a function of the pixel value in the reference LDR image normalized to the range $[0, 1]$:

$$a(x) = \begin{cases} 0.058 + 0.68(x - 0.85) & \text{if } x \geq 0.85 \\ 0.04 + 0.12(1 - x) & \text{if } x < 0.85 \end{cases}. \quad (26)$$

Using these formulas, regions that are consistent with the reference image are averaged over, whereas regions affected by ghost are downgraded [1]. The methods described in [27, 28, 31] use a similar weight adaptation approach.

Another ghost-free HDR image generation method using gradient information is proposed by Zhang and Cham [23]. The method is based on the observation that the gradient direction in stationary regions remains stable in different exposures, provided that these regions are neither under-exposed nor over-exposed. On the contrary, if the content changes due to object movement, the gradient direction varies accordingly. Therefore, a consistency measure based on gradient direction changes between different exposures and the reference one is computed, and used as weighting function in the HDR image generation equation. More precisely, for each input LDR image L_k , the gradient information is extracted by convolution with the first derivative of a 2D Gaussian kernel. The gradient magnitude and the gradient direction of the pixel located at (u, v) in L_k are denoted by M_{uv}^k and D_{uv}^k , respectively. First, a visibility measure that indicates the relative visibility of a pixel (u, v) in exposure L_k is defined by:

$$V_{uv}^k = \frac{M_{uv}^k}{\sum_{i=1}^N M_{uv}^i + \epsilon}, \quad (27)$$

where ϵ is a small value used to avoid singularities.

To deal with ghost artefacts, the consistency measure S_{uv}^k is computed for each pixel (u, v) in L_k as:

$$S_{uv}^k = \sum_{l=1}^N \exp\left(\frac{-(d_{uv}^{kl})^2}{2\sigma_s^2}\right), \quad (28)$$

where $\sigma_s = 0.2$ and d_{uv}^{kl} is the gradient direction change between L_k and L_l at position (u, v) and is calculated in the window of size $(2r+1) \times (2r+1)$, $r = 9$, as follows:

$$d_{uv}^{kl} = \frac{\sum_{x=-r}^r |D_{(u+x)(v+x)}^k - D_{(u+x)(v+x)}^l|}{(2r+1)^2}. \quad (29)$$

In order to eliminate the effect of saturated regions, a refined score C is derived from S , as follows:

$$C_{uv}^k = \frac{S_{uv}^k \times w_{uv}^k}{\sum_{i=1}^N (S_{uv}^i \times w_{uv}^i) + \epsilon}, \quad (30)$$

where the weighting function is given by:

$$w_{uv}^k = \begin{cases} 1 & \text{if } 25 \leq Z_{uv}^k \leq 225 \\ 0 & \text{otherwise} \end{cases}. \quad (31)$$

The final HDR image is obtained, without employing tonemapping techniques, as a weighted sum of pixel values across exposures. The weights are derived from the previously generated visibility and consistency measures. The final HDR image is obtained using the following equation:

$$I_{uv} = \sum_{k=1}^N W_{uv}^k Z_{uv}^k, \quad (32)$$

where the weights are given by:

$$W_{uv}^k = \frac{V_{uv}^k \times C_{uv}^k}{\sum_{i=1}^N (V_{uv}^i \times C_{uv}^i) + \epsilon}. \quad (33)$$

4.2. Removing all moving objects

In some cases, it could be desirable to completely remove all moving objects in the final HDR image. For example, considering the case of a building, the object of interest for the photographer could be the building itself and not moving persons in the scene. To achieve this goal, a simple approach is to discard in the combination step, exposures that are affected by ghosting at each pixel location. This idea is used by Sidibe *et al.* [26] who identify, for each pixel location (u, v) , two sets of exposures: A_{uv} and B_{uv} . The former is the set of exposures containing ghosting at location (u, v) , while the latter represents exposures that do not contain ghosting. Therefore, combining only exposures in B_{uv} lead to a ghost-free HDR image. Gallo *et al.* [29] use a similar approach to generate ghost-free HDR images. However, their algorithm is based on image patch processing rather than working with pixels individually. They start by determining the number of exposures to use in different ghost affected areas and use these exposures to generate an HDR image.

Other methods [41, 42] directly remove ghosting by adjusting the weighting function used in the HDR image generation equation (Eq. (1)). Such methods do not need explicit ghost detection as they directly and iteratively change pixels weights to minimise the number of visible artefacts. Khan *et al.* [41] propose a kernel density estimation method that iteratively estimates the probability that a pixel belongs to the static part of the scene. Pedone and Heikkilä [42] suggest a similar iterative approach. They estimate bandwidth matrices for computing the accurate probability that a pixel belongs to the background, and propagate the influence of the low probabilities to the surrounding regions using an energy minimization technique. The final probabilities are used as weights in the HDR image generation equation. The main assumption in these work is that the exposures sequence predominantly captures the static parts of the scene or, equivalently, that moving objects appear in a small number of images at each pixel location. Moreover, these methods require a sufficiently large number of images to produce good results and can be computationally expensive since they require a certain number of iterations.

5. Comparison and Classification of Ghost Detection and Removal Algorithms

In this section, we compare and classify the different ghost detection and removal methods that have been described in Sections 3 and 4. The comparison is based on a quantitative evaluation of the generated ghost maps, i.e. we evaluate the accuracy of the different methods in detecting moving objects in the scene. The classification is based on several criteria: the fusion domain, the need for a ghost map computation, the number of exposures required, the setting of parameters and the final generated HDR image.

5.1. Comparison of Ghost Detection and Removal Algorithms

For a fair comparison of the different ghost methods, we use a sequence of seven exposures with moving objects. Five of the seven LDR images of the sequence are shown in Fig. 3(a) and the resulting HDR image with ghosting artefacts is shown in Fig. 3(b). The sequence is taken with a Canon EOS 50D camera, and the exposure times are set to $[\frac{1}{3}, \frac{1}{5}, \frac{1}{8}, \frac{1}{13}, \frac{1}{20}, \frac{1}{30}, \frac{1}{50}]$ seconds respectively.

The sequence is designed to test the algorithms on various aspects. Ghosts formed by the pen on the table serve as *small ghost* that tests the ghost detecting resolution of the algorithms. The similarity in colors between the container and the table tests the ability of detecting low contrast ghosts. The shadows of the container and the high variance of the background tests the



(a)



(b)

Figure 3: The sequence used for comparison. (a) Five exposures containing two moving objects; (b) HDR image generated showing ghosting artefacts.

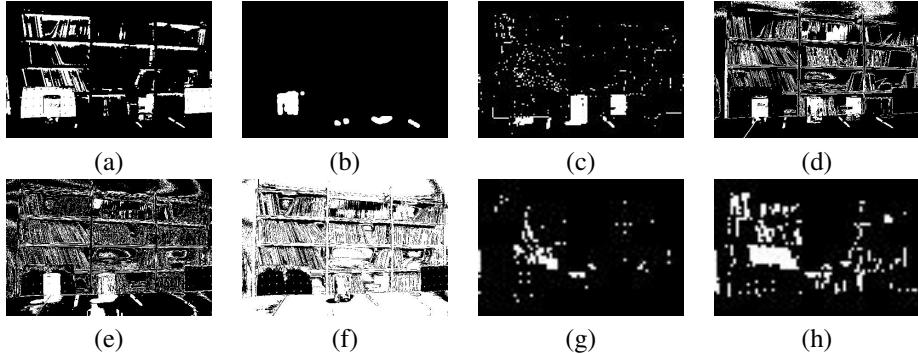


Figure 4: Some ghost detection results. (a) Detection with the variance based method [1]; (b) Detection with the entropy based method [24]; (c) Detection with the pixel order method [26]; (d) Detection with the bitmap based method [28]; (e) and (f) Detection with the multi-thresholding based method [27]; (g) and (h) Detection with the RANSAC based method [29].

sensitivity to ghosts. Finally, ghosts in each image either overlap completely or do not overlap at all, which simplifies obtaining the ground truth data. For each image of the exposures sequence, the exact positions of the moving objects are manually segmented and give the ground truth for ghost pixels in the scene. The ghost detection methods described in Section 3 are applied to generate ghost maps indicating the areas of the scene that are affected by ghost.

Some examples of detected ghost maps are shown in Fig. 4. It is important to mention that some methods such as the variance method [1], the entropy method [24], the bitmap method [28] or the pixel order based method [26], generate a single ghost map using the entire sequence of exposures. On the contrary, methods such as the multi-thresholding technique [27] and the Ransac based method [29] generate a different ghost map for each pair of images formed by the reference exposure and another exposure. The prediction based method [25] and the graph-cuts based approach [31], generate a ghost map for each pair of consecutive exposures. Furthermore, many ghost detection methods are based of thresholding and the detection results depend on the value of the threshold. We have tried different thresholds and have selected the values producing the best results. The results shown in Fig. 4 are obtained with a threshold value of 0.45 for the variance method, 0.45 for the entropy method and for the ransac based method, an outlier probability and a threshold value of 0.05% and 0.5, respectively.

Ghost detection can be viewed as a classification problem in which each pixel is classified as been either a ghost pixel or a nonghost pixel. We can therefore compare the detected ghost maps with the ground truth ghost maps in terms of sensitivity and specificity. The sensitivity is the percentage of ghost pixels correctly classified as ghost and the specificity indicates the percentage of nonghost pixels correctly classified as nonghost. Therefore, a good ghost detection method should have high sensitivity, i.e. correctly detect all ghost pixels, and high specificity, i.e. not misclassify nonghost pixels as ghost.

The detection results for different ghost detection algorithms are summarized in Table 2. Note that in the cases where a particular method generates multiple ghost maps, the presented sensitivity and specificity values are the average of the individual values for each ghost map. As can be seen, most of the detection methods show a high specificity value, more than 80%, meaning that they do not misclassify many nonghost pixels as ghost. However, the multi-thresholding

Table 2: Quantitative evaluation of ghost detection methods

Method	Reference	Sensitivity	Specificity
Variance	[1]	0.467	0.833
Entropy	[24]	0.320	0.992
Prediction	[25]	0.679	0.861
Pixel order	[26]	0.617	0.930
Multi-thresholding	[27]	0.870	0.540
Bitmap	[28]	0.664	0.807
Ransac	[29]	0.111	0.928
Graph-Cuts	[31]	0.913	0.845

method [27] shows a specificity of only 54%, which means that most pixels of the scene are incorrectly detected as ghost pixels. This can be observed in the images of Fig. 4(e) and (f). The methods that achieve the best sensitivity values are the graph-cuts [31] based method and the multi-thresholding based method. They achieve a specificity of 91.3% and 87%, respectively. Based on the obtained results, we can conclude that the best ghost detection methods, i.e. methods showing both high sensitivity and specificity values, are the graph-cuts based method [31], the prediction based method [25] and the pixel order method [26].

Some examples of ghost removal for the sequence in Fig. 3 are shown in Fig. 5. For this sequence containing slow motion and low contrast ghost, all ghost removal methods fail to completely remove all artefacts. As can be seen in Fig. 5(a) and (b), the variance based method and the entropy based method fail to remove artefacts created by the pen and the container when trying to keep both objects at fixed locations. A satisfactory result is obtained by the multi-thresholding method shown in Fig. 5(c). The method almost correctly remove all visible artefacts and keep one occurrence of the moving objects. A method such as the density estimation method of Khan [41] which tries to completely remove all moving objects fails in this case because of slow motion as shown in Fig. 5(e). Note that this iterative method would eventually remove all ghost artefacts as many iterations are performed, but as we can see in Fig. 5(e), after three iterations the final HDR image starts to be blurred (e.g. the books on the bottom left of Fig. 5(e)).

5.2. Classification of Ghost Detection and Removal Methods

We classify the different ghost detection and removal techniques using several criteria: the fusion domain, the need for a ghost map computation, the number of exposures required, the setting of parameters and the final generated HDR image. The classification of the reviewed methods is shown in Fig. 7.

5.2.1. Fusion domain

Most of the proposed ghost removal methods generate the ghost-free HDR image in the radiance domain [1, 24, 34, 25, 26, 27, 29, 31, 41, 42]. In order to display the final HDR image in common LDR monitors, a tone mapping operator is used [15, 16, 17]. As previously mentioned, the performance of these methods highly depends on an accurate estimation of the camera response function which is used to convert pixel brightness values into the radiance domain. But the estimation of the camera response function is sensitive to noise and misalignment. The results shown in Fig. 5(a), (b), (c) and (e) are obtained in the radiance domain and are displayed

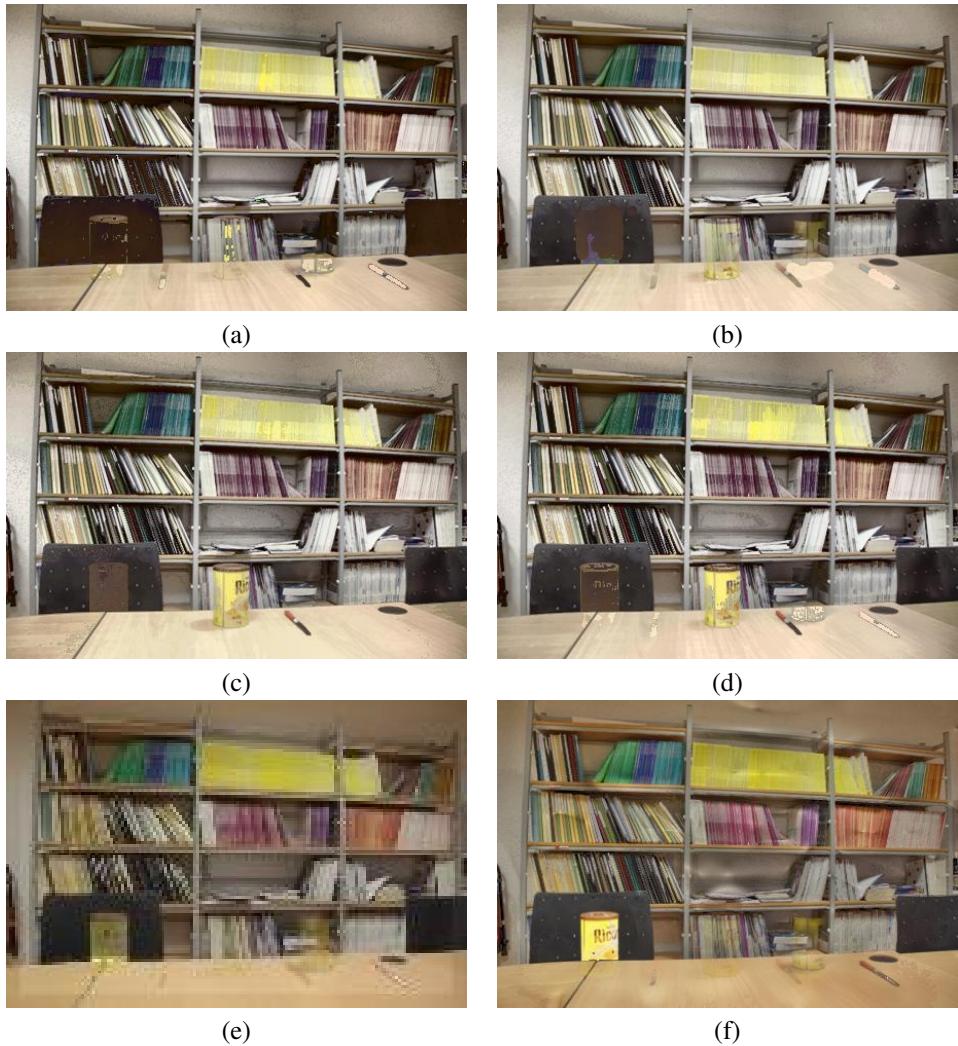


Figure 5: Some ghost removal results. (a) Ghost removal with the variance method [1]; (b) Ghost removal with the entropy based method [24]; (c) Ghost removal with the multi-thresholding based method [27]; (d) Ghost removal with the bitmap based method [28]; (e) Ghost removal with kernel density estimation method [41]. (f) Ghost removal with the gradient based method [23].

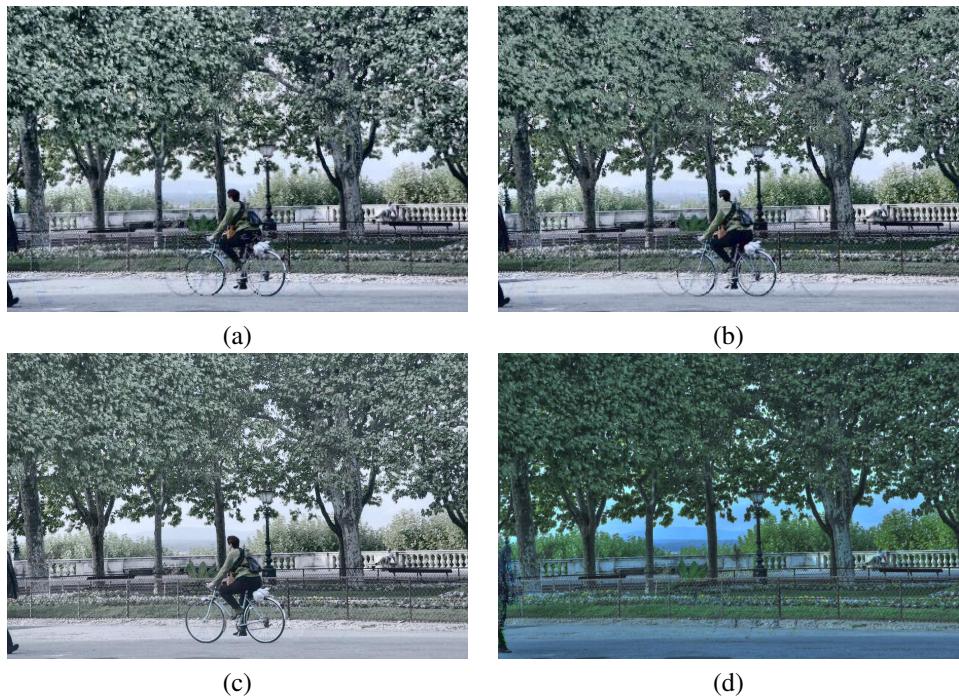


Figure 6: Ghost removal results with the sequence in Fig. 2. (a) Ghost removal with the variance method [1]; (b) Ghost removal with the bitmap based method [28]; (c) Ghost removal with the multi-thresholding based method [27]; (d) Ghost removal with the pixel order based method [26].

after a tone mapping operation. A few methods, on the other hand, directly combine the different exposures in the image domain by weighting pixels intensities [28, 23]. They are more time-efficient since they avoid the camera response function estimation and tone mapping. Example results are shown in Fig. 5(d) and (f). However, methods that combine exposures in the radiance domain have the advantage of producing a true HDR radiance map which can later be used for different processing or display applications.

5.2.2. Ghost map detection

Many methods are based on a first step of ghost detection [1, 24, 25, 26, 27, 28, 29, 31]. Ghosting artefacts are then removed based on the detected ghost map as explained in Section 4. On the other hand, methods such as [23, 41, 42] directly remove ghosting artefacts without the need for a ghost map. They are based on pixels weights adaptation. Zhang and Cham [23] use a reference image while Khan *et al.* [41] and Pedone and Heikkilä [42] use an iterative process. Methods that are based on an explicit ghost map computation are most appropriate for removing ghosting artefacts while keeping the moving object at fixed location, as shown in Fig. 6(a), (b) and (c). Iterative methods [41, 42] can accurately remove all moving objects in the combined HDR image, but they can be very expensive in computation as many iterations are required to achieve good results. For instance, the kernel density estimation technique [41] requires about one hour while the gradient based method [23] requires about eleven seconds for the sequence in Fig. 3. Note that we use our own Matlab implementations for the experiments. Furthermore, applying many iterations can cause blur if the final HDR image as shown in the example of Fig. 5(e).

5.2.3. Number of input LDR images

The minimum number of LDR images required to create an HDR image is two. But, most photographers use a set of three LDR images because the auto-bracketing function embedded in many cameras allows to capture three different exposures in one shot. Thus, ghost detection and removal methods must be able to produce good results with two or three exposures. Methods such as gradient based [23], bitmap based [28] or multi-level thresholds based [27] give accurate results with few images (three exposures). On the contrary, methods that depend on estimating the statistical distribution of pixel values [26, 41, 42] require more than three images to produce good visible results. For instance, a minimum number of five exposures is recommended in [26] and Khan *et al.* [41] use at least seven exposures. Using few images leads to fast computation but the methods might fail to detect and remove ghosting artefacts in case of slow motion of the object in the scene. In Fig. 5, we can see that none of the methods can successfully remove all ghosting artefacts in the final HDR image. On the other hand, using a large number of images requires a higher computation time. With our Matlab implementations, the bitmap based method [28] requires about 0.5 seconds for ghost detection with three exposures while the gradient based method [23] and the pixels order based method [26] require, respectively, 11 seconds and 53 seconds for seven exposures.

5.2.4. Parameters setting

In all reviewed methods, some parameters are used and tuned to get good final results. For instance, most of the ghost detection methods discussed in Section 3 need a threshold value to classify a pixel as ghost or not. Methods such as [1, 24, 25, 29, 41] employ manually set parameters in their algorithms to ensure best reconstruction of HDR images. This provides a complete control of the algorithm to the user and these methods can be employed when the input

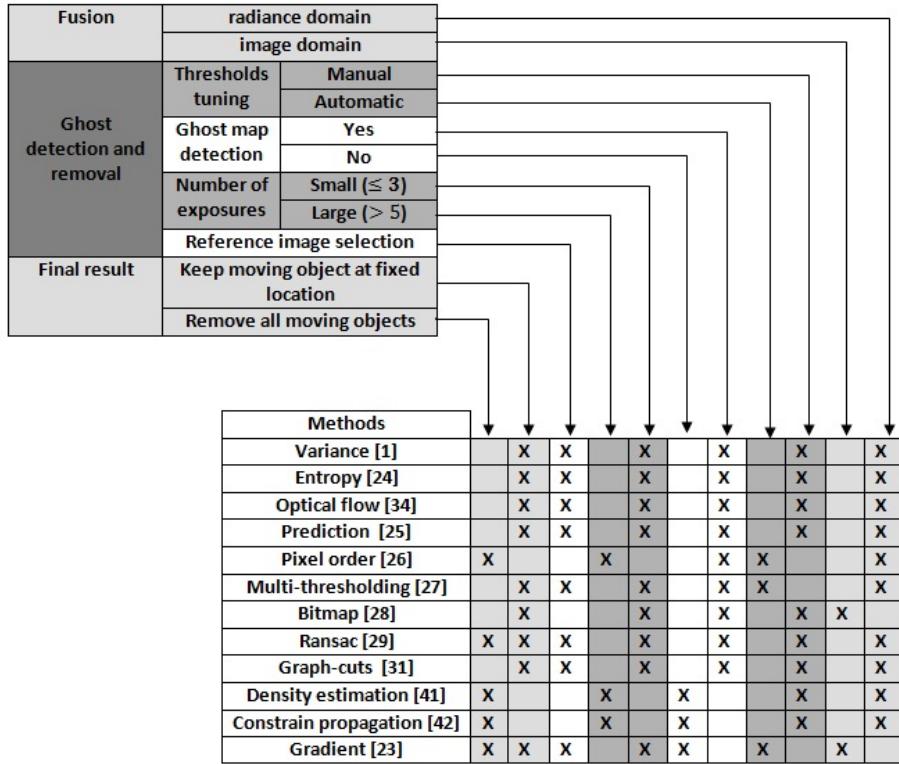


Figure 7: Classification of ghost detection methods.

LDR images do not change drastically, like in the construction of an HDR family portrait in a photo-studio, for instance. On the contrary, in algorithms such as [26, 27], the internal thresholds are set automatically by the algorithm, thus eliminating human intervention.

5.2.5. Final HDR image

Depending on the user’s interest, the moving objects can be either completely removed or kept at a fixed position in the final HDR image. Methods such as [1, 24, 25, 27, 28, 29, 31] remove all but one occurrence of the moving object by selecting a reference exposure in ghost regions. In order to avoid seams at ghost regions boundaries, a Laplacian pyramid blending framework or a Poisson editing technique can be used. Some examples are shown in Fig. 6 where we can observe that the multi-thresholding based method [27] produces a very good result (Fig. 6(c)). It keeps the cyclist at a fixed location while eliminating other artefacts. On the contrary, the variance [1] and the bitmap method [28] fail to correctly remove the ghosting artefacts. Other methods such as [26, 41, 23] completely remove all moving object in the final HDR image. They can achieve satisfactory results as shown in Fig. 6(d), but fail to remove all ghosting artefacts in case of slow motion as can be seen in Fig. 5(e) where Khan’s method [41] fail to eliminate the ghosting artefacts.

6. Conclusion

In this paper, the ghost problem in high dynamic range (HDR) imaging is presented and recently proposed methods to solve this problem are reviewed. Each method is described in detail and a comparison and classification of the reviewed methods is proposed. The comparison is based on a quantitative evaluation of the accuracy of the different methods in detecting ghost regions in a given sequence of exposures. The results show that high contrast movement, i.e. a moving object different from the background, can be correctly detected while small and low contrast movements, i.e. similarity in colors between the object and the background, are more difficult to detect. We classify the methods based on the fusion domain, the need for a ghost map computation, the number of exposures required, the setting of parameters and the final generated HDR image.

Methods that combine exposures in the image domain are time-efficient as they avoid the camera response function estimation and tone mapping. On the other hand, methods that combine exposures in the radiance domain give a true HDR radiance map which might be useful for later processing or display applications.

Generally speaking, there is no single best method and the selection of an approach depends on the user's goal. For removing all moving objects in the final HDR image, iterative methods achieve good results but are computationally expensive. Methods that discard exposures affected by ghost in the combination, thus using a subset of the input exposures sequence, provide good results with low complexity. For keeping the moving object at a fixed location in the combined HDR image, a good ghost map is required at the ghost detection stage. In addition, it is better to use a weight adaptation approach rather than a single exposure to avoid boundary effects in the combined image.

References

- [1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, Morgan Kauffman, 2005.
- [2] T. Mitsunaga, S. K. Nayar, High Dynamic Range Imaging: Spatially Varying Pixel Exposures, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2000, pp. 472–479.
- [3] D. D. Wen, High Dynamic Range Charge-Coupled Device, US Patent 4873561 (1989).
- [4] R. A. Street, High Dynamic Range Segmented Pixel Sensor Array, US Patent 5789737 (1998).
- [5] J. Tumblin, A. Agrawal, R. Raskar, Why I Want A Gradient Camera, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2005, pp. 103–110.
- [6] G. Ward, Robust Image Registration for Compositing High Dynamic Range Photographs from Hand-Held Exposures, *Journal of Graphics Tools* 8 (2003) 17–33.
- [7] C. Harris, M. Stephens, A Combined Corner and Edge Detector, in: Proceedings of the 4th Alvey Vision Conference, 1988, pp. 147–151.
- [8] G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision* 60 (2) (2004) 91–110.
- [9] P. Debevec, J. Malik, Recovering High Dynamic Range Radiance Maps from Photographs, in: Proceedings of the SIGGRAPH 97 Conference, 1997, pp. 369–378.
- [10] J. DiCarlo, B. Wandell, Rendering High Dynamic Range Images, in: Proceedings of the SPIE Electronic Imaging Conference, 2000, pp. 392–401.
- [11] C. Manders, C. Aimone, S. Mann, Camera Response Function Recovery From Different Illuminations of Identical Subject Matter, in: Proceedings of the IEEE International Conference on Image Processing, 2004, pp. 2965–2968.
- [12] S. Mann, R. Picard, Being 'Undigital' With Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures, in: Proceedings of the IS&T 46th Annual Conference, 1995, pp. 422–428.
- [13] T. Mitsunaga, S. K. Nayar, Radiometric Self Calibration, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 1999, pp. 374–380.

- [14] M. Grossberg, S. K. Nayar, What Can Be Known About the Radiometric Response from Images, in: Proceedings of the European Conference on Computer Vision, 2002, pp. 189–205.
- [15] F. Durand, J. Dorsey, Fast Bilateral Filtering For The Display of High Dynamic Range Images, Computer Graphics Forum 21 (2002) 257–266.
- [16] R. Fattal, D. Lischinski, M. Werman, Gradient Domain High Dynamic Range Compression, in: Proceedings of the SIGGRAPH Conference, 2002, pp. 249–256.
- [17] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic Tone Reproduction For Digital Images, in: Proceedings of the SIGGRAPH Conference, 2002, pp. 267–276.
- [18] S. Battiatto, A. Castorina, M. Mancuso, High Dynamic Range Imaging for Digital Still Camera: An Overview, Journal of Electronic Imaging 12 (3) (2003) 459–469.
- [19] V. Brajovic, Brightness Perception, Dynamic Range and Noise: A Unifying Model for Adaptive Image Sensors , in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition - Volume 2, 2004, pp. 189–196.
- [20] T. Mertens, J. Kautz, , F. V. Reeth, Exposure Fusion: A Simple and Practical Alternative to High Dynamic Range Photography, Computer Graphics Forum 28 (2009) 161–171.
- [21] A. A. Goshtasby, Fusion of Multi-Exposure Images, Image and Vision Computing 23 (2005) 611–618.
- [22] S. Raman, S. Chaudhuri, Bilateral Filter Based Compositing for Variable Exposure Photographs, in: Proceedings of the Eurographics Conference, 2009, pp. 1–4.
- [23] W. Zhang, W.-K. Cham, Gradient-Directed Composition of Multi-Exposure Images, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2010, pp. 530–536.
- [24] K. Jacobs, C. Loscos, W. G., Automatic High Dynamic Range Image Generation of Dynamic Environments, IEEE Computer Graphics and Applications 28 (2) (2008) 84–93.
- [25] T. Grosch, Fast and Robust High Dynamic Range Image Generation with Camera and Object Movement, in: Proceedings of Vision, Modeling and Visualization Conference, 2006, pp. 277–284.
- [26] D. Sidibe, W. Puech, O. Strauss, Ghost Detection and Removal in High Dynamic Range Images, in: Proceedings of the 17th European Signal Processing Conference - EUSIPCO, 2009, pp. 2240–2244.
- [27] T.-H. Min, R.-H. Park, S.-K. Chang, Histogram Based Ghost Removal in High Dynamic Range Images, in: Proceedings of the International Conference on Multimedia and Expo - ICME, 2009, pp. 530–533.
- [28] F. Pece, J. Kautz, Bitmap Movement Detection: HDR for Dynamic Scenes, in: Proceedings of Visual Media Production (CVMP), 2010, pp. 1–8.
- [29] O. Gallo, N. Gelfand, W.-C. Chen, M. Tico, P. K., Artifact-free High Dynamic Range Imagin, in: Proceedings of the IEEE International Conference on Computational Photography (ICCP), 2009, pp. 1–7.
- [30] M. A. Fischler, R. C. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Comm. of the ACM 24 (1981) 381–395.
- [31] Y.-S. Heo, K.-M. Lee, S.-U. Lee, Y. Moon, J. Cha, Ghost-free High Dynamic Range Imaging, in: Proceedings of the 10th Asian conference on Computer vision - ACCV, 2010, pp. 486–500.
- [32] Y. Boykov, O. Veksler, R. Zabih, Fast Approximate Energy Minimization via Graph Cuts, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (2001) 1222–1239.
- [33] B. Lucas, T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, in: Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981, pp. 674–679.
- [34] S. Kang, M. Uyttendaele, S. Winder, R. Szeliski, High Dynamic Range Video, ACM Transactions on Graphics 22 (2003) 319–325.
- [35] L. Bogoni, Extending Dynamic Range of Monochrome and Color Images through Fusion, in: Proceedings of the International Conference on Pattern Recognition, 2000, pp. 7 – 12.
- [36] J. R. Bergen, P. Anandan, K. J. Hanna, R. Hingorani, Hierarchical model-based motion estimation, in: Proceedings of European Conference on Computer Vision, 1992, pp. 237–252.
- [37] M. Brown, D. G. Lowe, Recognising panoramas, in: Proceedings of the IEEE International Conference on Computer Vision - Volume 2, 2003, pp. 1218–1225.
- [38] H.-Y. Shum, R. Szeliski, Construction of Panoramic Image Mosaics with Global and Local Alignment, International Journal of Computer Vision 36 (2) (2000) 101–130.
- [39] M. Uyttendaele, A. Eden, R. Szeliski, Eliminating ghosting and exposure artifacts in image mosaics, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern recognition, 2001, pp. 509–516.
- [40] P. Burt, T. Adelson, The Laplacian Pyramid as a Compact Image Code, IEEE Transactions on Communication 31 (1983) 532–540.
- [41] E. A. Khan, A. O. Akyüz, E. Reinhard, Ghost Removal in High Dynamic Range Images, in: Proceedings of the IEEE International Conference on Image Processing, 2006, pp. 2005–2008.
- [42] M. Pedone, J. Heikkilä, Constrain Propagation for Ghost Removal in High Dynamic Range Images, in: Proceedings of the International Conference on Computer Vision Theory and Applications - VISAPP, 2008, pp. 36–41.