ARTICLE IN PRESS

Optics Communications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Underwater video enhancement using multi-camera super-resolution

E. Quevedo [a,b,*], E. Delory [a], G.M. Callicó [b], F. Tobajas [b], R. Sarmiento [b]

[a] Oceanic Platform of the Canary Islands, Telde, Las Palmas 35214, Spain
[b] Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria, Las Palmas 35017, Spain

ARTICLE INFO

Keywords:
Underwater applications
Image enhancement
Multi-camera
Super-resolution

ABSTRACT

Image spatial resolution is critical in several fields such as medicine, communications or satellite, and underwater applications. While a large variety of techniques for image restoration and enhancement has been proposed in the literature, this paper focuses on a novel Super-Resolution fusion algorithm based on a Multi-Camera environment that permits to enhance the quality of underwater video sequences without significantly increasing computation. In order to compare the quality enhancement, two objective quality metrics have been used: PSNR (Peak Signal-to-Noise Ratio) and the SSIM (Structural SIMilarity) index. Results have shown that the proposed method enhances the objective quality of several underwater sequences, avoiding the appearance of undesirable artifacts, with respect to basic fusion Super-Resolution algorithms.

1. Introduction

High-resolution images can be an important performance requirement for automated pattern recognition and scene analysis. High resolution may also improve pictorial information for human interpretation. Image resolution can be defined as a measure of the sharpness of the details of an image: the higher the resolution, the more image details can be observed. The resolution of a digital image can be categorized in many different ways: pixel resolution, spatial resolution, spectral resolution, temporal resolution (when used in videos), and radiometric resolution. In this paper, the main interest is devoted to spatial resolution. A digital image is made of small picture elements called pixels and the spatial resolution refers to the significant pixel density, measured in pixels per unit area [1].

Image spatial resolution is primarily constrained by the imaging sensors or the acquisition system. A modern image sensor is typically a Charge-Coupled Device (CCD) or a Complementary Metal–Oxide–Semiconductor (CMOS) active-pixel sensor. These sensors are usually arranged in a two-dimensional array to capture image signals. The sensor size, or equivalently the number of sensor elements per unit area, determines in the first place the spatial resolution of the image to capture. While the sensors limit the spatial resolution, image details (high-frequency bands) are also constrained by the optics, due to lens blurs (associated with the sensor Point Spread Function, PSF), lens aberration effects, aperture diffractions, and optical blurring due to motion. Constructing imaging integrated circuits and optical components to capture very high-resolution images is prohibitively expensive and not practical in most real applications. Besides the cost, the resolution of a surveillance camera is also limited by the camera speed and storage capacity. In other scenarios such as satellite or underwater imaging, it is difficult to use high resolution sensors due to size constraints. Another way to address this problem is to accept the image degradations and to use signal post-processing techniques; one of these techniques is addressed in this paper and is referred to as Super-Resolution (SR) reconstruction [2].

In this paper, the approach for SR is to construct High-Resolution (HR) video sequences from several observed Low-Resolution (LR) images, thereby increasing the high-frequency components and removing the degradations inherent to LR cameras. In our approach, every LR frame is an undersampled, aliased observation of the true scene. This idea is possible only if there is subpixel motion between these low-resolution frames, and thus the ill-posed upsampling problem can be better conditioned. In the imaging process, the camera captures several LR frames, which can be considered as a downsampled version from an HR scene with subpixel shifts between each other. SR reconstruction reverses this process by aligning the LR observations to subpixel accuracy and combining them into an HR image grid (interpolation), thereby overcoming the limitation of the camera. In terms of applications, SR arises in many areas such as surveillance video, remote sensing, medical imaging, underwater environment, or lenses design [3]. In this scope, state of the art underwater image restoration and enhancement

* Corresponding author at: Oceanic Platform of the Canary Islands, Telde, Las Palmas 35214, Spain.
E-mail addresses: eduardo.quevedo@plocan.eu, equevedo@iuma.ulpgc.es (E. Quevedo), eric.delory@plocan.eu (E. Delory), gustavo@iuma.ulpgc.es (G.M. Callicó), tobajas@iuma.ulpgc.es (F. Tobajas), roberto@iuma.ulpgc.es (R. Sarmiento).

http://dx.doi.org/10.1016/j.optcom.2017.06.054
Received 16 February 2017; Received in revised form 11 May 2017; Accepted 15 June 2017
Available online xxxx
0030-4018/

techniques [4,5] could be applied as a pre-processing step for a subsequent SR stage.

The proposed system is based on capturing various LR images from several sensors, attached one to one another by an $R \times S$ array. This framework is known as a Multi-Camera (MC) system [6]. Due to the cost reduction in camera sensors, using multiple sensors along with appropriate processing techniques will be a common strategy for image enhancement in future cameras [7]. For instance, in [8] a lens attachment of 8 LR low-quality side cameras arranged around a high quality lens turns a standard DSLR (Digital Single-Lens Reflex) camera and lens into a light field camera.

The paper is structured as follows: Section 2 provides a brief review of the state of the art in underwater image enhancement. Section 3 introduces the algorithm used to enhance underwater video sequences, Section 4 addresses the proposed experimental setup to evaluate the system, Section 5 analyzes the most significant results, and Section 6 provides the most relevant conclusions.

2. Underwater imaging enhancement

Model-based SR reconstruction techniques for underwater imaging have been previously studied, as in [9], where in order to improve the resolution an imaging model based on beam propagation is established and applied to image super-resolution reconstruction techniques for an underwater range-gated pulsed laser imaging system. However, in general terms there is no commonly accepted method for image quality evaluation in the literature, or in most cases the assessment is just visual. This limitation therefore also applies to underwater imaging. In [10,11], an acoustic sequence is taken as a reference, and then denoising is applied by means of simply averaging different acoustic images. In a similar way, in [11], Chen et al. carried out a quantitative evaluation through GMG (Gray Mean Grads) and LS (Laplacian Sum) proposed by Sheikh and Bovik [12] to compare evaluation metrics of single-frame reconstruction in underwater imaging. These evaluation metrics are based on interpolations: nearest, bilinear, cubic-spline and wavelet-based, or the Papoulis–Gerchberg (PG) method, proposed and studied by Papoulis [13] and Gerchberg [14]; and evaluation metrics of single-frame reconstruction such as Projection Onto Convex Set (POCS), which is also used in [15] to enhance low contrast, strong noise and image blur of underwater video images as an outset to create a new SR reconstruction and enhancement algorithm.

Some studies use the PSNR (Peak Signal to Noise Ratio) in order to compare the bilinear interpolated image with the super-resolved image using a Basic Super-Resolution (BSR) approach, based on fusion SR techniques [16]. PSNR is calculated according to the expression in (1) based on the Mean Square Error (MSE) presented in (2):

$$PSNR = 20 * \log\left(\frac{255}{\sqrt{MSE}}\right) \tag{1}$$

$$MSE = \frac{\sum_{i=1}^{P}\sum_{j=1}^{Q}[f(i,j) - F(i,j)]^2}{P * Q} \tag{2}$$

where $f(i,j)$ is the original signal at pixel (i,j), $F(i,j)$ is the reconstructed signal (interpolated or super-resolved), and $P \times Q$ is the image size. The result of PSNR is a value in decibels (dB).

On the contrary, in [17], Pezham Firoozfam uses the RMSE (Root Mean Squared Error) to compare different turbidity levels in terms of Nephelometric Turbidity Units (NTU), and the SSIM (Structural SIMilarity) index is introduced in [18]. The SSIM expression is presented in (3):

$$SSIM = \frac{(2xy + C_1)\cdot(2\sigma_{xy} + C_2)}{(\bar{x}^2 + \bar{y}^2 + C_1)\cdot(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

where $\bar{x}$ is the mean of x, σ_x is the variance of x, σ_{xy} is the covariance between x and y, and $C_1 = (k_1 L)^2$, $C_2 = (k_2 L)^2$ are two variables to stabilize the division with weak denominator where L is the dynamic range of the pixel values, being $k_1 = 0.01$ and $k_2 = 0.03$ by default.

As these metrics complement each other with respect to the information they convey, it was decided to use both PSNR and SSIM as a basis for the proposed test bench. Additionally, other works to improve the quality of underwater images not related to SR have been based on image restoration and image enhancement techniques, as will be described in the following sections.

2.1. Image restoration techniques

The objective of these techniques is to recover the original image from the observed image, using (if available) explicit knowledge about the degradation function (the so-called Point Spread Function, PSF) and noise characteristics. This section presents the most relevant state of the art related to this approach.

In [4] a self-adjustable filter based on a simplification of the Jaffe–McGlamery model is introduced. This simplified model is appropriate for surface waters with limited light flickering, although experimental results show that this approach can be followed in a wide variety of circumstances. As an alternative, the approach proposed in [5] is based on physical characteristics, recovering information related to scene structure (distances). In this case, it is commented that the lack of image definition is not the main cause of contrast degradation and it is concluded that polarization is the main perturbation of underwater visibility.

Another underwater image restoration algorithm is presented in [19], based on an atmospheric turbulence model. This automatic algorithm does not require previous knowledge of the acquisition conditions. This method clearly improves the contrast and image definition of underwater images. In a similar way, the study introduced in [20] shows that turbidity or marine snow is an important perturbation which can strongly diminish the performance of the color restoration algorithms, so a specific Histogram Stretching Method (HSM) is proposed to increment the robustness in environments where marine snow is present.

2.2. Image enhancement techniques

These methods consider neither the image composition process nor the previous knowledge of the environment. Therefore, these are simpler and faster methods than other image restoration techniques.

In [21] a contrast adjustment of the RGB color space is proposed together with intensity saturation and expansion of the HIS (Hue, Saturation and Intensity) color space. The advantage of applying two different expansion models is that it helps to equilibrate color contrast, considering at the same time the problem of illumination (light flickering). Other techniques such as a median filter, together with a dark channel are proposed in [22], where color correction is also integrated before producing the final output. Contrast and color correction are also considered in [23] using quaternionic image correction.

2.3. Underwater environment characteristics to be considered in the SR process

After reviewing the state of the art on image restoration and image enhancement techniques, it was concluded that the main underwater environment characteristics to be considered in the SR process are the following:

- **Color loss**: the absorption of radiations composing light is selective depending on the wavelength. Considering the visible spectrum, absorption is maximal for red color and minimum for green, and especially blue. Color loss is not only quantitative, but also qualitative, as intensity changes also depend on the specific color.

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

- **Brightness and contrast**: these characteristics are mainly affected by refraction and diffraction effects. Refraction is produced when the light wave passes from one medium to another and changes direction. Diffraction is produced when the light beam changes direction when it finds an obstacle.
- **Light flickering**: this characteristic is due to the lens effect of surface water motion, which focuses and blurs solar light when depth is reduced.
- **Turbidity**: it depends on the quantity of underwater small particles. This effect is known as "marine snow". Particle concentration depends on the zone, season, water temperature, or oceanic currents, among others.
- **Artificial light**: the lack of light, as depth increases, implies the use of artificial light devices in order to illuminate the environment. Artificial light provokes a non-uniform illumination, generating a very illuminated center zone, surrounded by a region with limited lighting.

Next section details the designed Multi-Camera Super-Resolution algorithm in order to take into account the underwater environment characteristics presented above. Specific filters (coarse filter for working window and fine filter for macro-blocks) have been implemented and tuned according to these characteristics. Based on the results, the effects of the underwater environment in the proposed SR process are presented in the Conclusions section.

3. Multi-Camera Super-Resolution algorithm

The SR algorithm used as a basis in this paper belongs to the fusion category [7]. This algorithm provides both static and dynamic SR, depending on the number of frames, i.e. whether the input is a single frame or a frame sequence, respectively. For each frame (*current frame*) a *working window* is generated, that contains the adjacent frames (*reference frames*), from which the new sub-pixel information will be obtained. Afterwards, the current frame is divided into uniform square regions, called Macro-Blocks (MB), and the motion estimation is performed between the current frame and each reference frame, following a MB basis. The algorithm uses a block-matching method to obtain the sub-pixel motion vectors associated to each MB.

In this paper, this algorithm has been tested in an underwater environment in order to validate it in a real application and to check how the main underwater environment characteristics affect its behavior. The algorithm is called MCABMSR for Multi-Camera Adaptive Block Matching Super Resolution, and combines two selective filters for the frames and Macro-Blocks (MBs, frames divisions) to be included in the SR process, and an MC framework, as shown in Fig. 1. For simplicity, the figures including MC systems in this paper show a 2×2 array framework, composed by 4 cameras. The larger the number of cameras is used, the higher the quality of the method can be. This is because the algorithm has the opportunity to find more information to improve the output sequence.

The algorithm includes a pre-processing step integrated in the frames reordering process algorithms. The target is to discern whether some parts of the captured frames should be discarded in order to gain higher quality by avoiding artifacts and/or reducing computational cost. Some physical restrictions to the MC array are considered:

- The geometry of the MC array must be rectangular or quadrangular.
- The cameras must be set in the same plane.
- The cameras must record parts of the same global scene from different locations.

Considering an array which complies with these requirements, as the one shown in Fig. 2(a), it can be seen that there will be a common part (or *overlap*) of the recorded information by the cameras of the MC array (or at least for a subset of cameras). Surrounding the *overlap* there will be a *border*, as shown in Fig. 2(b). The geometry and separation between cameras of the MC array will affect the disposition of the borders and the way to obtain them, and they will influence on the *overlap* between the cameras, as shown in Fig. 2(b). The overlap region may be combined (or not) with the borders, giving rise to several ways to process the information of the frames captured by the MC array:

- *Full Frame* mode: It considers the whole frame information. This is the basic mode, in which the full frame is captured by every camera.
- *Overlap* mode: It only considers the *overlap* between cameras. In order to obtain the *overlap,* the *offset* in pixels between cameras must be obtained. Once the *offset* is known, all the captured frames are cropped saving only the *overlap*. After cropping the frames a new sequence is obtained, which will be the input of SR process.
- *Overlap + Borders* mode: It considers dividing the frame between the *borders* and the *overlap.* This mode not only considers the *overlap* to apply SR, but also the *borders*. In this way it is possible to analyze if it is convenient to discard these *borders*, to apply SR separately, or to apply SR with the full frame without applying any pre-processing. As shown in Fig. 2(b), this mode provides 9 different parts to be super-resolved: the *overlap*, 4 *corners* and 4 *sides*.

Previous research has shown that the information contained in the overlap region maximizes the spatial resolution, confirming that the Full-Frame mode optimizes the results [7].

The first stage is called *Window Selective Filter* (WSF). As shown in Fig. 1, the combination of information from several cameras in different timestamps provides different *Working Windows* (WW_{t-1}, WW_t, WW_{t+1}, …). Each *Working Window* contains spatially and temporally frames adjacent to the current frame of one specific camera (Cam 2 in Fig. 1), from which the information is obtained. WSF uses a selection of the best candidate frames based on SSIM, depending on the characteristics of the image, therefore improving quality and reducing computational cost.

The second stage, known as *Motion Estimation*, determines the motion between two or more views of the same parts of the scene with sub-pixel accuracy, which depends on the selected *Scale Factor*, which is the factor applied to the input sequence to generate the super-resolved image/sequence (i.e. obtaining an output sequence, the size of which is twice the horizontal and vertical size of the input sequence, would mean *a Scale Factor* of 2). The MAD (*Mean Absolute Difference*) between frames is also calculated and a second dynamic decision maker is used to determine the size of MBs using a *Variable Block Size* (VBS). This proposed contribution is a capability based on the characteristics of the frame. The algorithm uses a block-matching method to obtain the motion vectors of each variable size MB. Once all the motion vectors have been calculated, the third stage, called *Shift & Add*, is executed. A Very High Resolution (VHR) grid is filled with the information (contributions) pointed by the estimated motion vectors of each frame from the ones considered by the *Window Selective Filter* (WSF). Besides, the MBs to be considered can also be selected in this stage using a *Block Selective Filter* (BSF), which is the last contribution of this paper. The selection of frames and Macro-Blocks considered in the *Motion Estimation* process depends on the comparison with thresholds which will be further detailed.

The final stage, called *Fill Holes*, considers the fact that it is possible that the candidate frames do not contain sufficient information to fill all the locations on the grid for the current frame, so that there could be some empty positions on it, denoted as *holes* in the context of this paper. In such case, each empty pixel is filled by using a bilinear surface interpolator. The whole process is repeated for each frame. Finally a HR super-resolved image is obtained with the weighted summation of all the motion compensated pixels from the frames contained in the *Working Window*, and the super-resolved sequence is stored.

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

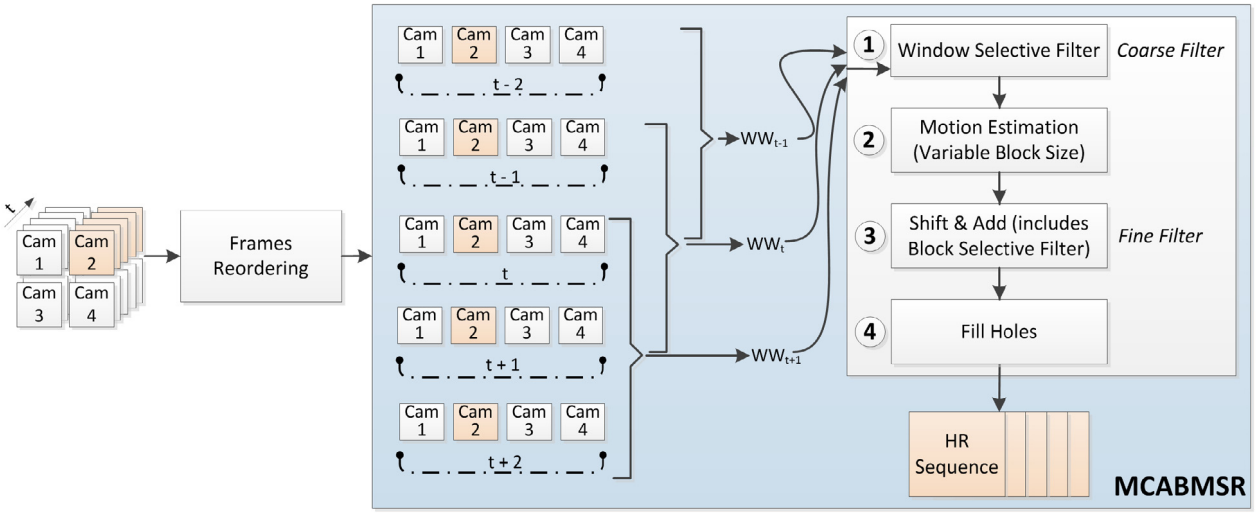


Fig. 1. Stages of the MCABMSR algorithm.

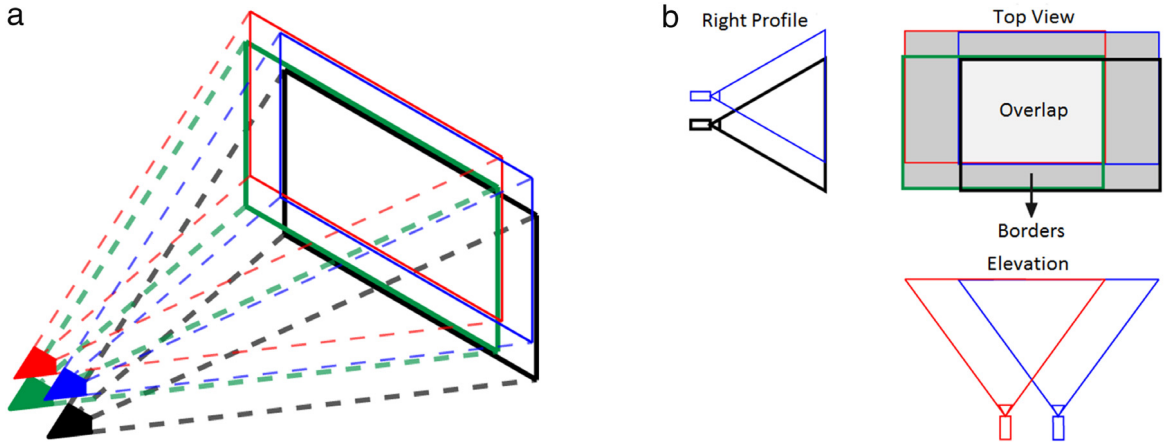


Fig. 2. Borders and Overlap in an MC array. (a) Frames recorded by an MC array in perspective. (b) Side and front elevations, and plan views of the recorded frames.

It is expected that the combination of a high amount of temporal and spatial information could provoke the appearance of artifacts in an underwater environment. To prevent this, a filter with two levels is included: a coarse level (*Window Selective Filt* er, WSF), identified with the *Working Window* WW (the selection of frames to be considered in the SR process), and a fine level (*Block Selective Filter*, BSF), associated with each MB (the MBs are categorized in different sizes of 16×16, 8×8 and 4×4 pixels following an adaptive MB topology). Low values of MB size allow having more independent motion vectors, but this largely increases the computation time. High values of MB size increase the coherence of the motion vectors, but this degrades the quality when several objects with different motions are considered together. Each filter integrates smart decision makers to select which frame or MB should be included in the SR process:

- **Coarse level**: *Window Selective Filter Threshold* (WSF_THR). This threshold is compared with the SSIM (*Structural Similarity Index*) [17] of each frame of the *Working Window* for the current frame in order to determine if a frame is considered in the SR process. Then, an updated *Working Window* (*Window Selective Filter Working Window, WSFWW*) is generated, as shown in (4):

$$-WSFWW = CF + \Sigma(FWW/SSIM(WW, CF)) > WSF\_THR, \quad (4)$$

where CF is the *Current Frame*, and FWW is a *Frame of the Working Window*.

- **Fine level**: *Block Selective Filter Threshold* (BSF_THR). This threshold is compared with the MAD of each block of a frame (MAD as it is defined in Eq. (5)) in order to determine if an MB is considered in the SR process. An MB is considered in the SR process if the expression in (6) is complied:

$$MAD = \frac{1}{M * N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left| cntBlock_{ij} - refBlock_{ij} \right| \quad (5)$$

$$MAD(cntBlock, refBlock) < BSF\_THR, \quad (6)$$

where $M * N$ is the size of the blocks, i and j are the indexes of each pixel, *cntBlock* is the MB to be enhanced, and *refBlock* is each one of the MBs from the *Working Window*.

- **Adaptive MB topology**: The selection of the appropriate size of MBs is based on a combination of the two thresholds defined for the *Window Selective Filter* and the *Block Selective Filter*. This combination defines thresholds VBS_THR1 and VBS_THR2 (*Variable Block Size Threshold 1 & 2*), which are respectively correlated with WSF_THR and BSF_THR.

The proposed thresholds are dynamically adapted using internal parameters of the algorithm based on the motion vectors. Another important parameter for the SR algorithm is the *Search Area (SA)*, which

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

is the number of pixels around every MB considered when searching the motion-vectors. For scenes containing high amounts of motion, high values of SA are required. For low-motion scenes, high values of SA are not self-defeating, to the extent that they will not decrease the output quality, but a great amount of inefficient computation would be performed. From the performed experiments it is deduced that values from 8 to 16 are usually enough for the majority of the used underwater sequences.

4. Test bench

As far as the authors are concerned, there is no accepted specific underwater images database for comparison, so in order to set up a validation environment for the proposed SR method, a new test bench has been developed. The general diagram of the proposed *testbench* is shown in Fig. 3 considering a 2×2 MC array.

Firstly, the HR sequence (any underwater sequence taken as an input) is decimated (↓D) considering the *Scale Factor (SF)*, and reordered so as to compose the LR input sequence. At the same time, all the frames in the video sequence of one of the cameras are taken as a reference (Cam 2 in Fig. 3). Secondly, the LR frames of the camera taken as a reference are interpolated (↑I) in order to get the interpolated sequence and the super-resolved sequence (↑ SR) together with the frames of the rest of the cameras in order to compose the MCABMSR sequence. Finally, both the interpolated sequence and the super-resolved sequence are compared with the HR reference in order to get the objective metrics PSNR and SSIM.

5. Results

In this section, the most significant results based on underwater test sequences adapted for comparison are presented. Table 1 summarizes the level of the main characteristics for the studied sequences. Values between "1" and "5" have been subjectively assigned, identifying the "1" as the minimum (not present characteristic in the sequence) and "5" as the maximum (dominant characteristic in the sequence), covering a wide range of combinations. The first frame for each one of the selected sequences (30 frames have been considered per sequence) for comparison is shown in Fig. 4.

Figs. 5, 6 and Table 2 show the results for the considered sequences. Figs. 5 and 6 evaluate the objective quality of the proposed method (MCABMSR), including images composed with a 3×3 camera array, versus the proposed method without using filters (MCABMSR_wof), a Basic Super-Resolution (BSR) algorithm no adapted for underwater sequences [7], and a bilinear interpolation (INT) in order to demonstrate the suitability of the followed approach. Two available SR algorithms have been considered in this paper for comparison purposes: Iterative Back Projection (IBP) and Robust Super-Resolution (RSR). They have been included in Figs. 5, 6 and Table 2, showing that the provided results clearly improve the state of the art. Iterative Back Projection [24] presents a novel real-time SR system based on a weighted mean SR algorithm combined with previous approaches, which use fast and robust multi-frame SR algorithms. Robust SR [25] is especially focused on the presence of artifacts. Such artifacts may be due to motion errors, inaccurate blur models, noise, moving objects, motion blur, etc. Robustness is needed since SR methods are very sensitive to such errors. As it is shown in Figs. 5, 6 and Table 2, the provided results clearly enhance the performance of the considered algorithms; the main reason is the resolution enhancement due to the usage of a Multi-Camera framework combining spatial and temporal Super-Resolution adapted to the underwater environment, which is the main innovation of this paper. At the same time, the introduction of two specific filters, at working-window level and macro-block level, reduces the computation time with respect to state of the art for all the considered sequences.

Table 2 shows the computation times per MB in milliseconds (ms) for these methods, showing that it is especially reduced for sequences with high local motion, such as *Coral*. The MCABMSR method not only improves the objective quality with respect to the BSR method, but also reduces substantially computational cost with respect to MCABMSR_wof, as it is shown in Table 2. This is highly desirable in sequences with high local motion such as *Coral and Reef*. In these sequences, the peak improvements of the MCABMSR method with respect to the BSR method are the following:

- *Coral*: up to 1.84 dB in PSNR and 12.19% in SSIM.
- *Reef*: up to 2.28 dB in PSNR and 5.41% in SSIM.

In sequences such as *Atlantis*, *Cavern*, *Fish_coral*, or *Titanic*, mainly characterized by global motion, *Window Selective Filter* (WSF) and *Block Selective Filter* (BSF) consider that most of the information should be included in the SR process in order to improve the objective quality. This is coherent due to the fact that global motion works very well with fusion SR techniques. In this case there are noticeable peak improvements:

- *Atlantis*: up to 2.54 dB in PSNR and 9.85% in SSIM.
- *Cavern*: up to 3.77 dB in PSNR and 3.69% in SSIM.
- *Fish_coral*: up to 2.72 dB in PSNR and 9.75% in SSIM.
- *Titanic*: up to 3.25 dB in PSNR and 5.85% in SSIM.

In order to further study the quality of the analyzed sequences, charts for the first 30 frames of each sequence were analyzed frame by frame. As a reference, Fig. 7 shows the results for the *Fish_coral* sequence and Fig. 8 shows the results for the *Reef* sequence. As shown in both figures, the proposed MCABMSR method provides not only the best results, but also stability in objective quality:

- *Fish_coral* sequence (Fig. 7): 93.33% of the frames (28 out of 30) present values of PSNR between 28.5 and 29.4 dB (less than 1 dB of variability) and SSIM values between 0.85 and 0.87 (2% of variability).
- *Reef* sequence (Fig. 8): 70% of the frames (21 out of 30) present values of PSNR between 26 and 27 dB (1 dB of variability) and SSIM values between 0.935 and 0.945 (1% of variability).

Finally, if the subjective quality of the sequences is analyzed, it can be seen that the annoying artifacts are substantially reduced. As a qualitative reference, Fig. 9 shows frame 19 of the *Fish_coral* sequence and Fig. 10 shows frame 30 of the *Reef* sequence. A subjective inspection of Fig. 9 presents how the annoying artifacts which appear in INT and BSR versions of the frame have been reduced in the coral profile. Similarly, it can be seen in Fig. 10 how the resolution of the fish has been clearly improved.

6. Conclusions

After studying a wide set of representative sequences, the effects of the underwater environment in the SR process have been analyzed as follows:

- The turbidity phenomenon, mainly characterized by marine snow and the light flickering phenomenon, produces a local motion effect which makes more difficult to reconstruct the image using fusion SR techniques. Therefore, the number of frames to be used in the SR process will be limited in order to avoid the appearance of artifacts. This has been achieved limiting the *Window Selective Filter* threshold in order to minimize the number of frames, and reducing the MB size, to consider the local motion. This situation has been detected in sequences as *Coral*, *Cavern* and *Titanic*. Therefore, it is concluded that a pre-processing step to avoid or at least reduce light flickering and turbidity [18,19] can help to improve the SR process.

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

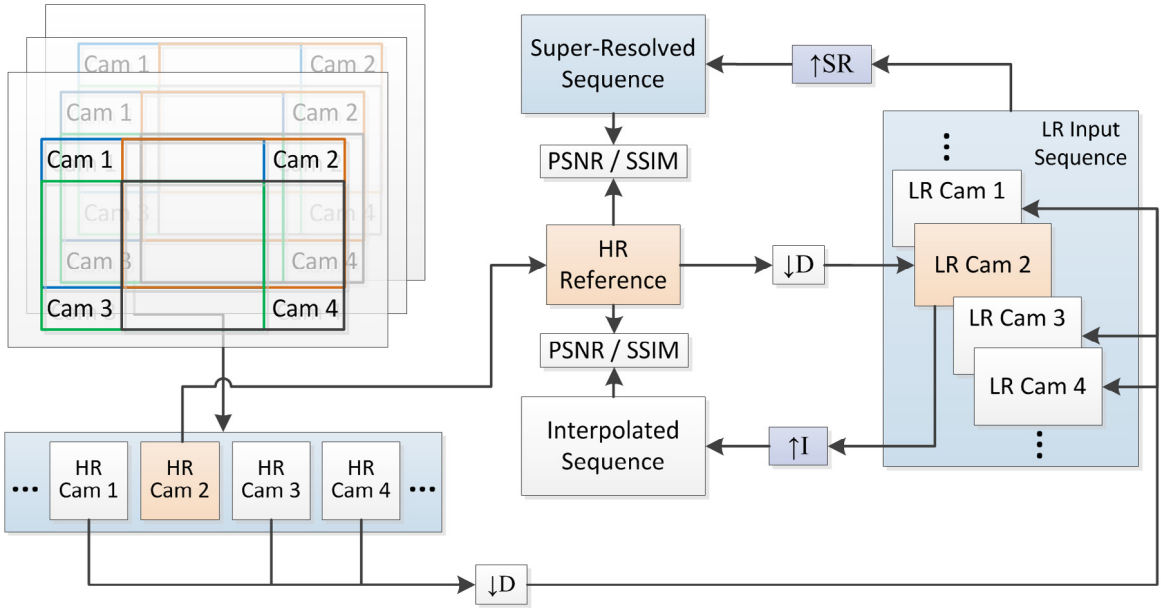


Fig. 3. Proposed testbench.

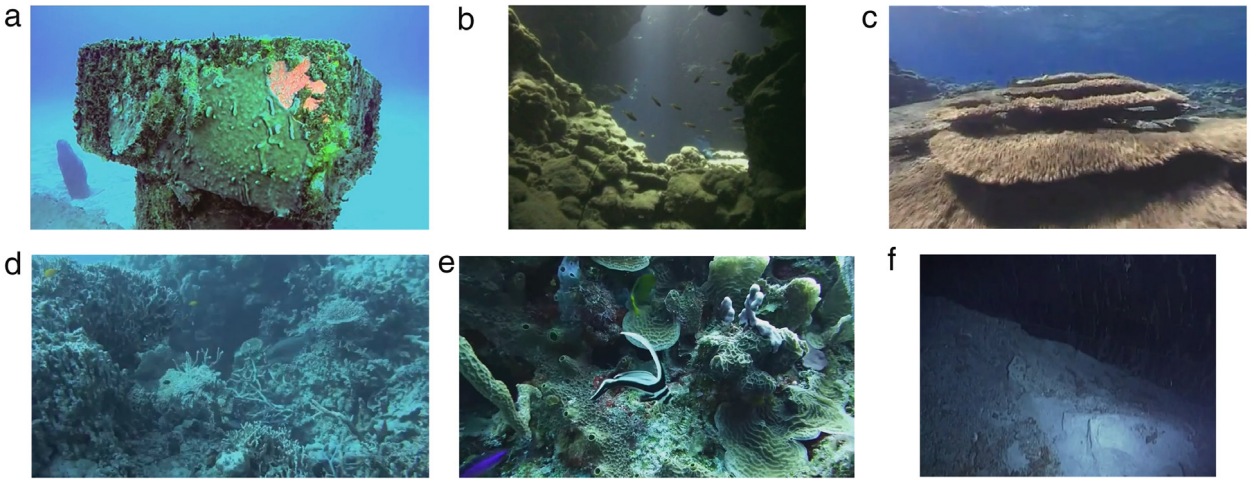


Fig. 4. Representative frames of the underwater sequences used for comparison: *Atlantis* (a), *Cavern* (b), *Coral* (c), *Fish_coral* (d), *Reef* (e) and *Titanic* (f).

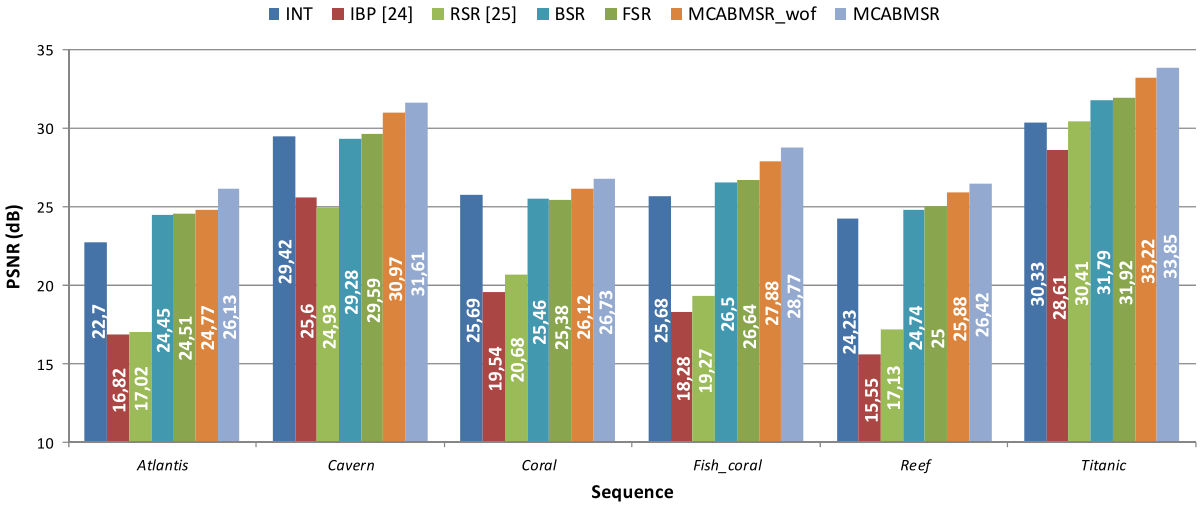


Fig. 5. Average PSNR results in dB.

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

Table 1
Main characteristics of the selected sequences.

Characteristic	Sequence					
	Atlantis	*Cavern*	*Coral*	*Fish_coral*	*Reef*	*Titanic*
Frame size	512 × 256	384 × 256	512 × 256	512 × 256	768 × 640	384 × 256
Global motion	3	2	4	3	3	4
Local motion	1	2	4	3	4	2
Color loss	3	4	2	3	2	5
Brightness and contrast	3	3	2	4	2	3
Light flickering	1	4	5	1	4	1
Turbidity	1	3	1	2	2	4
Artificial light	5	1	1	1	1	5

Table 2
Execution time per Macro-Block in ms.

Sequence	Time					
	t_{INT} (ms)	t_{IBP} (ms) [24]	t_{RSR} (ms) [25]	t_{BSR} (ms)	$t_{MCABMSR\_wof}$ (ms)	$t_{MCABMSR}$ (ms)
Atlantis	34.23	1378.83	1459.38	258.56	1293.74	700.55
Cavern	35.55	1450.93	2064.06	297.59	1263.43	668.99
Coral	17.05	848.67	898.35	258.53	1364.07	84.64
Fish_coral	31.41	1541.95	1665.47	511.52	1296.34	713.58
Reef	12.94	1677.68	1821.37	194.18	1324.48	505.58
Titanic	37.41	1526.67	1821.37	410.03	1231.96	594.18

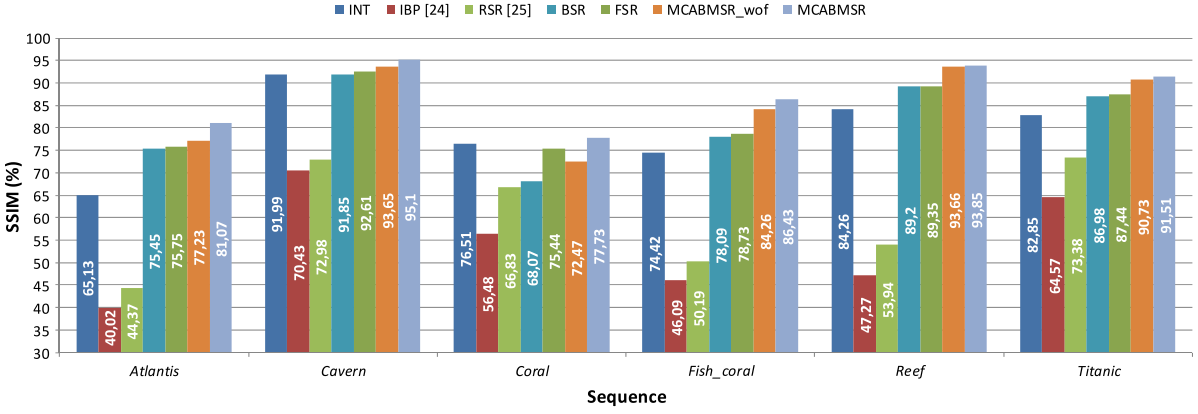


Fig. 6. Average SSIM results in %.

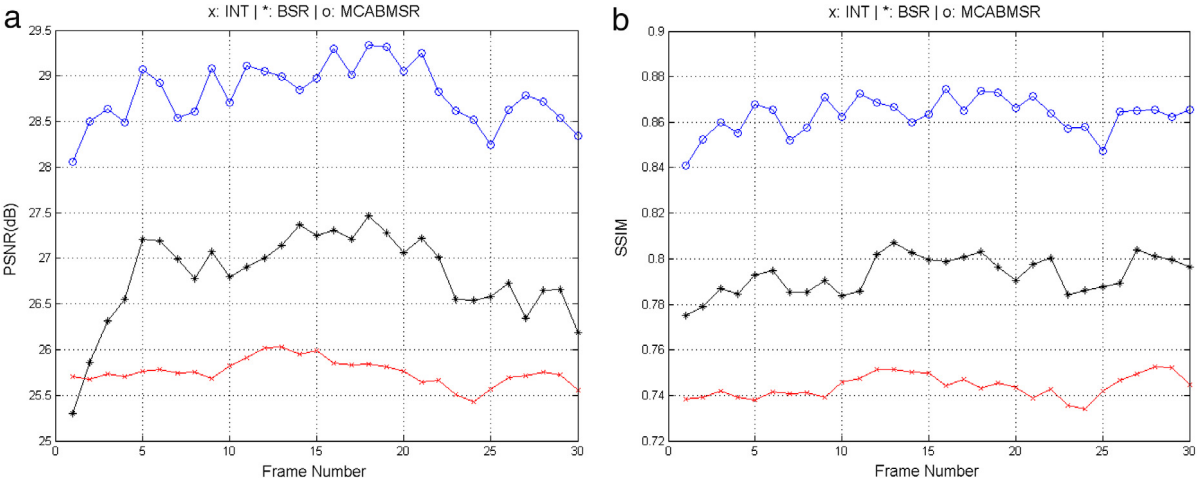


Fig. 7. INT, BSR and MCABMSR PSNR (a) and SSIM (b) results for the *Fish_coral* sequence.

- Artificial light in underwater sequences is camera-motion dependent. Artificial light can therefore be identified as a type of global motion. Sequences including artificial light but with low turbidity and light flickering, for as it is in the case of *Atlantis*, work appropriately when SR is applied, improving clearly the image details. In this case, the more the number of frames, the higher the obtained objective quality. At the same time a large MB size can be used, as global motion is assumed.
- Color loss effect is neutral when SR is applied since the considered algorithms only analyze the luminance component, so the

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

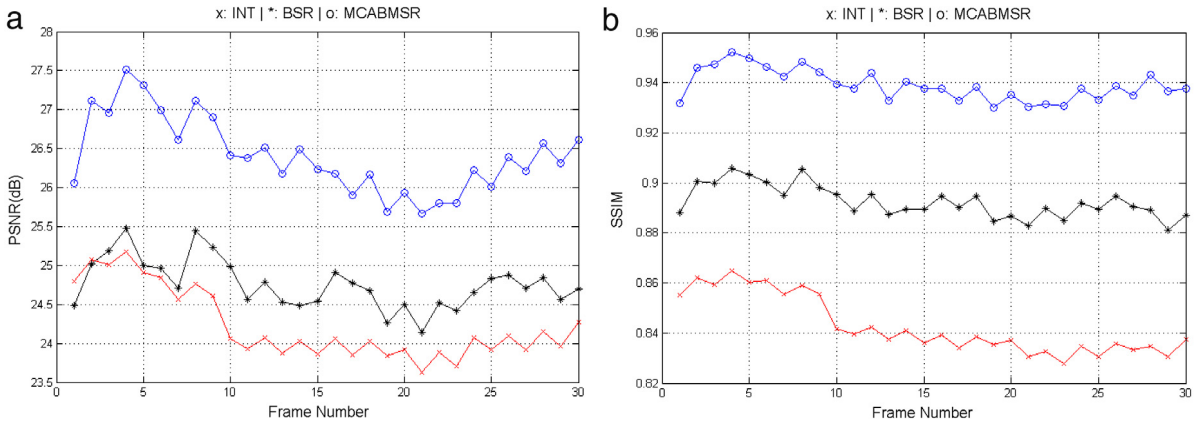
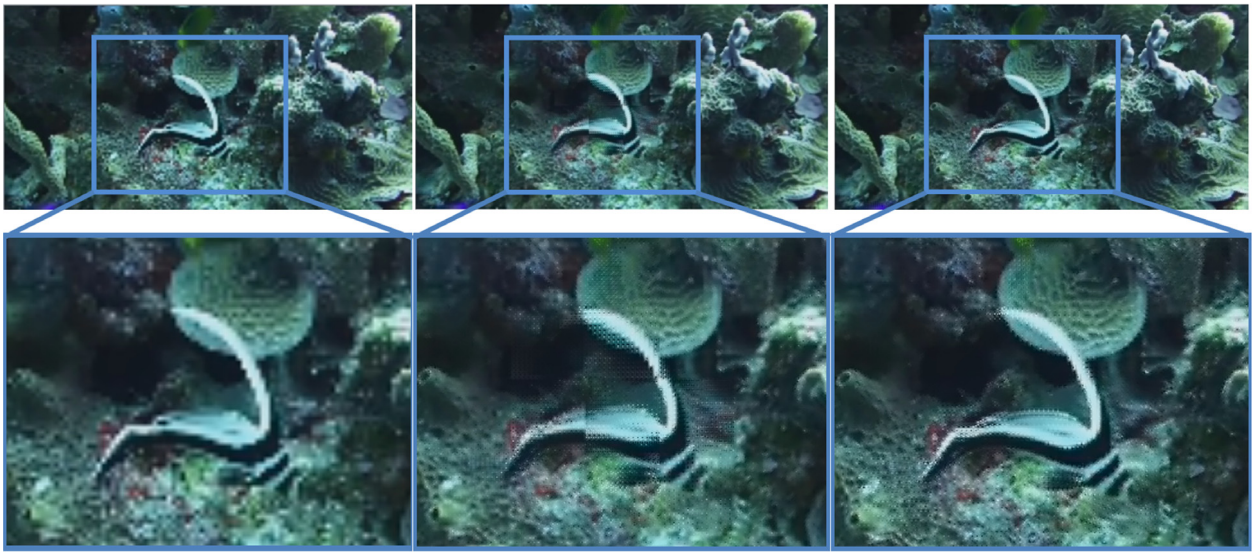


Fig. 8. INT, BSR and MCABMSR PSNR (a) and SSIM (b) results for the *Reef* sequence.

(a) INT. (b) BSR. (c) MCABMSR.

Fig. 9. Complete and detailed view of frame 19 of the *Fish_coral* sequence for INT (a), BSR (b), and MCABMSR (c).



(a) INT. (b) BSR. (c) MCABMSR.

Fig. 10. Complete and detailed view of frame 30 of the *Reef* sequence for INT (a), BSR (b), and MCABMSR (c).

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.

absence of some colors does not directly affect the SR process. At the same time, brightness and contrast do not have a significant influence in the SR process.

- For sequences recorded with a camera installed on a moving platform, usually operated by a diver or a submarine vehicle, the camera displacement is an essential aspect to be considered when applying SR. In case of ROVs (Remotely Operated Vehicles) it is possible to determine vehicle speed using georeferencing information and the number of frames to be considered in the SR process, as proposed in [26].

In this paper, a fusion Super-Resolution algorithm integrated with a Multi-Camera system has been tested in underwater environment. The combination between spatial and temporal correlations has shown appropriate results, avoiding the appearance of annoying artifacts with respect to other basic fusion Super-Resolution techniques. A wide set of underwater sequences have been considered, concluding that turbidity and light flickering are the main underwater characteristics to be avoided when SR is applied.

The results show that the proposed method enhances the objective quality of several sequences reducing computational cost using frame and Macro-Block filters. It is also remarkable that in complex situations for fusion Super-Resolution techniques such as high global and local motion, the proposed method has shown good results. Future research will be based on reducing the computation time of the algorithm in order to reach real-time execution.

References

[1] J. Yang, T. Huang, Image super-resolution: Historical overview and future challenges, in: Super-Resolution Imaging, CRC Press, Boca Ratón, 2010.
[2] Jiang Tiang, Kai-Kuang Ma, A survey on Super-Resolution imaging, J. Signal Image Video Process. 5 (2) (2011) 329–342.
[3] Ninfa del C. Lozano-Rincón, Noé Alcalá Ochoa, Design of super-resolving lenses in far-field imaging, Opt. Commun. 383 (2017) 316–322.
[4] E. Trucco, A.T. Olmos-Antillon, Self-tuning underwater image restoration, IEEE J. Ocean. Eng. 31 (2) (2006) 511–519.
[5] Y.Y. Schechner, N. Karpel, Recovery of underwater visibility and structure by polarization analysis, IEEE J. Ocean. Eng. 30 (3) (2005).
[6] Bing Ouyang, Frase Dalgeish, Anni Vuorenkoski, Walter Britton, Brian Ramos, Benjamin Metzger, Visualization and image enhancement for multistatic underwater laser line scan system using image-based rendering, IEEE J. Ocean. Eng. 38 (3) (2013) 566–580.
[7] E. Quevedo, J. de la Cruz, G.M. Callicó, F. Tobajas, R. Sarmiento, Video enhancement using spatial and temporal Super-Resolution from a Multi-Camera system, IEEE Trans. Consum. Electron. 60 (3) (2014) 420–428.
[8] Yuwang Wang, Yebin Liu, Wolfgang Heidrich, Qionghai Dai, The light field attachment: turning a DSLR into a light field camera using a low budget camera ring, IEEE Trans. Vis. Comput. Graphics (2016).
[9] Yuzhang Chen, Min Xia, Wei Li, Kecheng Yang, Xiaohui Zhang, Model-based super-resolution reconstruction techniques for underwater imaging, in: Photonics and Optoelectronics Meetings, Wuhan, China, Nov 2011.
[10] K. Kim, N. Neretti, N. Intrator, Acoustic camera image mosaicing and super-resolution, in: IEEE MTTS OCEANS International Conference, Kobe, Japan, Nov 2004.
[11] Yuzhang Chen, Wei Li, Min Xia, Qing Li, Kecheng Yang, Super-resolution reconstruction for underwater imaging, Opt. Appl. 41 (4) (2011).
[12] H.R. Sheikh, A.C. Bovik, Image information and visual quality, IEEE Trans. Image Process. 15 (2) (2006) 430–444.
[13] A. Papoulis, A new algorithm in spectral analysis and band-limited extrapolation, IEEE Trans. Circuits Syst. 22 (9) (1975) 735–742.
[14] R.W. Gerchberg, Super-resolution through error energy reduction, Int. J. Opt. 21 (9) (1974).
[15] Li Jia-Jia, Li Qing-wu, Wang Dan, Super-resolution reconstruction and enhancement for underwater video image, China Nat. Knowl. Infrastruct. (2011).
[16] G.M. Callicó, S. López, O. Sosa, J.F. López, R. Sarmiento, Analysis of fast block matching motion estimation algorithms for video Super-Resolution systems, IEEE Trans. Consum. Electron. 54 (3) (2008) 1430–1438.
[17] Pezham Firoozfam, Multi-Camera imaging for 3-D mapping and positioning; stereo and panoramic conical views (Ph.D. thesis), 2004.
[18] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Eero P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Imaging Process. 13 (4) (2004) 600–612.
[19] Yang Miao, Underwater image adaptive restoration and analysis by turbulence model, in: World Congress on Information and Communication Technologies, Trivandrum, India, 2012.
[20] M. Boffety, F. Galland, Phenomenological marine snow model for optical underwater image simulation: Applications to color restoration, in: OCEANS Conference, Yeosu, South Korea, May 2012.
[21] Kashif Iqbal, Rosalina Abdul Salam, Azam Osman, Abdullah Zawawi Talib, Underwater image enhancement using an integrated colour model, Int. J. Comput. Sci. 34 (2) (2007) 239–244.
[22] Hung-Yu Yang, Pei-Yin Chen, Chien-Chuan Huang, Ya-Zhu Zhuang, Yeu-Horng Shiau, Low complexity underwater image enhancement based on dark channel prior, in: Second International Conference on Innovations in Bio-inspired Computing and Applications, Shenzhan, Chine, 17–20, Dec 2011.
[23] F. Petit, A. Capelle-Laize, P. Carré, Underwater image enhancement by attenuation inversion with quaternions, in: IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, Apr. 2009, pp. 1177–1180.
[24] O. Bowen, C.-S. Bouganis, Real-time image super resolution using an FPGA, in: International Conference on Field Programmable Logic and Applications, Heidelberg, Germany, September 2008, pp. 89–94.
[25] A. Zomet, A. Rav-Acha, S. Peleg, Robust super-resolution, in: Proceedings on computer Vision and Pattern Recognition, Vol. 1, December 2001, pp. I-645–I-650.
[26] E. Quevedo, J. Rodríguez, D. Horat. A. Quesada-Arencibia, F. Tobajas, G.M. Callicó, R. Sarmiento, Improving underwater video navigation systems using Georeferencing and Super-Resolution techniques, in: Institute of Electrical and Electronic Engineers (IEEE) Marine Technology Society (MTS) OCEANS Conference, Taipei, Taiwan, Apr. 2014, pp. 1–7.

Please cite this article in press as: E. Quevedo, et al., Underwater video enhancement using multi-camera super-resolution, Optics Communications (2017), http://dx.doi.org/10.1016/j.optcom.2017.06.054.