We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Approach to Super-Resolution Through the Concept of Multicamera Imaging

Eduardo Quevedo, Gustavo Marrero and
Félix Tobajas

Additional information is available at the end of the chapter

Abstract

Super-resolution consists of processing an image or a set of images in order to enhance the resolution of a video sequence or a single frame. There are several methods to apply super-resolution, from which fusion super-resolution techniques are considered to be the most adequate for real-time implementations. In fusion, super-resolution and high-resolution images are constructed from several observed low-resolution images, thereby increasing the high-frequency components and removing the degradations caused by the recording process of low-resolution imaging acquisition devices. Moreover, the proposed imaging system considered in this work is based on capturing various frames from several sensors, which are attached to one another by a $P \times Q$ array. This framework is known as a multicamera system. This chapter summarizes the research conducted to apply fusion super-resolution techniques to select the most adequate frames and macroblocks together with a multicamera array. This approach optimizes the temporal and spatial correlations in the frames and reduces as a consequence the appearance of annoying artifacts, enhancing the quality of the processed high-resolution sequence and minimizing the execution time.

Keywords: super-resolution, multicamera, camera array, video enhancement, fusion

1. Introduction

The limitations of imaging devices directly affect the spatial resolution of images and video. The super-resolution (SR) reconstruction concept is considered in the literature as the process of combining information from multiple low-resolution images with subpixel displacements

to obtain a higher resolution image. Even though numerous methods have been developed to this end, there are still multiple future research challenges [1]. SR arises in several fields, such as remote sensing, surveillance, and an extensive set of consumer electronics applications [2–4].

This chapter proposes an imaging system in which high-resolution (HR) images are generated from low-resolution (LR) sensors through a SR image reconstruction process. In order to get several LR images minimizing the local motion, several digital cameras are attached to each other by a $P \times Q$ array frame. This framework is known as a multicamera (MC) system. The image reconstruction problem using an MC system applying an SR process could be stated as follows: *Given a set of multiview low-resolution frames of size $M \times N$ pixels taken with a multicamera system, and a scale factor s, reconstruct a higher resolution frame of size $sM \times sN$ pixels that accomplishes the definition of resolution enhancement.*

After a comprehensive review of the state of the art [5–21], it has been concluded that the application of SR to an MC system involves some preceding and subsequent steps. These steps are summarized in the proposed block diagram shown in **Figure 1**. However, in some cases, many steps could be skipped. For instance, the previous steps: MC system prototyping and construction and sometimes the MC system adjustment are not applicable if a commercial camera array is used [5], image capture is almost always considered, and pre- and postprocessing are sometimes omitted.
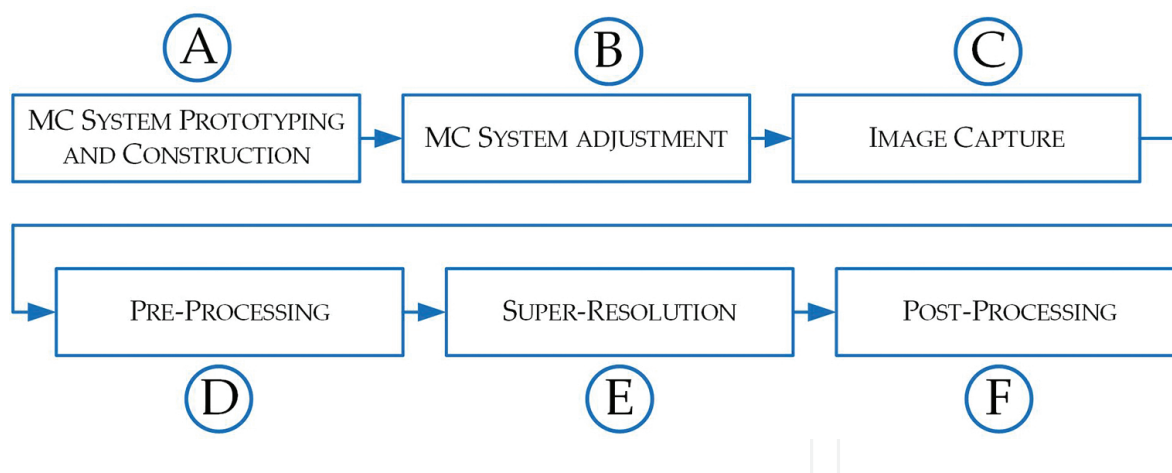
Figure 1. Block diagram of SR applied to an MC system.

This chapter is organized as follows: the state of the art of the steps described in **Figure 1** is discussed in Section 2, whereas Section 3 shows dedicated preprocessing schemes implemented together with three different methods which maximize the combination between SR and MC, proposed by the authors. These methods exploit the temporal correlation of the recorded videos and the spatial correlation among cameras. Finally, the conclusions are highlighted in Section 4.

2. Applying SR to an MC system

2.1. Multicamera system prototyping and construction

The prototyping of the MC system mainly depends on the application. One of the most used approaches is image-based rendering [4]. Using a MC array for this purpose can allow a device to process a higher quality frame in real time by means of multiple observations recorded simultaneously.

Some advisable factors to consider in any prototype of MC system are the following [4]:

- *User-friendliness*: The system should be designed to be easily created, i.e., cameras should require minimal setup and time-consuming calibration procedures should be avoided.

- *Flexibility*: Addition or removal of cameras according to their participation in the network and to a certain extent the flexibility in placing the cameras physically.

- *Off-the-shelf components*: It is desirable to keep reduced costs.

SR through the concept of MC imaging has been considered in the literature:

- Fanaswala [5] introduced in his thesis a commercial camera array of 25 cameras arranged on a 5 × 5 grid (ProFUSION25) of the commercial brand Pointgrey.

- Park et al. [6] used a prototype of a MC system based on a 3 × 3 array composed of nine digital cameras, CCD (charged coupled device).

- In Agrawal et al. [7], an implementation using four PointGrey Dragonfly2 cameras, each one equipped with 12 mm lens and triggering the cameras with a microcontroller (PIC) is presented.

- Finally, Firoozfam [8] presented a stereo MC conical system with 6 and 12 cameras, showing that increasing the number of cameras makes it possible to take advantage of several scenes observations at each time instant.

All these camera array systems are illustrated in **Figure 2**. In real-time applications, there is a compromise between the number of cameras and the computational cost, so it is useful to have a flexible architecture in order to select the cameras to be used, often in a N^2 configuration: 4, 9, 16, 25, … [5].

### 2.2. Multicamera system adjustment

The success of SR recovery from multiple views in real applications mainly depends on two factors [9]:

- The accuracy of multiple view registration results.

- The accuracy of the camera and data acquisition model.

Hence, in order to have a good level of SR, it is very important to perform a detailed adjustment of the MC system. The approach of using software located in a central server for the system

adjustment is usually taken into account. For instance, Park et al. [6] developed a software that shows previews and the status of the images, grabbing them simultaneously. The intensity and focusing indexes are also included in the implementation in order to adjust the lenses for the purpose of intensity and blur uniformity.
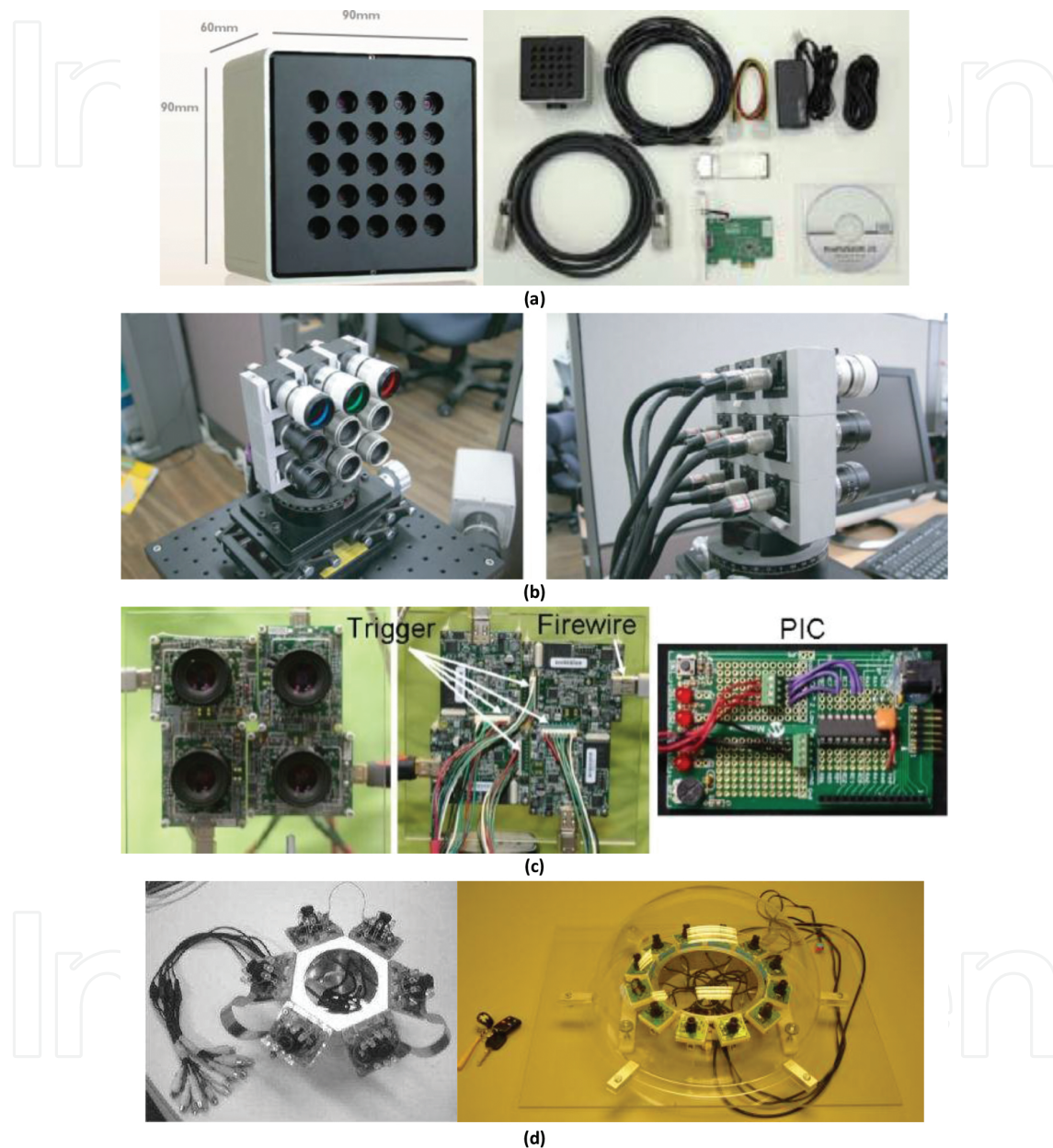
Figure 2. MC systems used for SR: (a) Fanaswala [5], (b) Park [6], (c) Agrawal et al. [7], and (d) Firoozfam [8].

There are also calibrations based on other elements, for example, Agrawal et al. [7] assumed that the scene is planar and perform geometric calibration using a checkerboard. Meanwhile color calibration is done using a Macbeth chart by computing a 3 × 3 color transformation for every camera. Finally, in this system, all cameras are triggered using microcontrollers, which avoid temporal synchronization issues. In the same way, both determining the camera

parameters and rectifying the images for lens distortion are achieved by intrinsic calibration with a checkerboard by Firoozfam [8].

The adjustment process of the MC system is a basic step to settle a solid basis for the rest of the steps in the multicamera-super-resolution (MC-SR) approach. If a commercial system is used (as for instance Fanaswala in [5]), this process is simplified, but the calibration step is limited by the system performance; however, this makes the comparison between SR algorithms easier.

2.3. Image capture

Low-resolution images are captured by cameras generally using software implemented in a central server. The variation between researches depends on where the software is included: in an external computer, in the MC system, or sharing both systems.
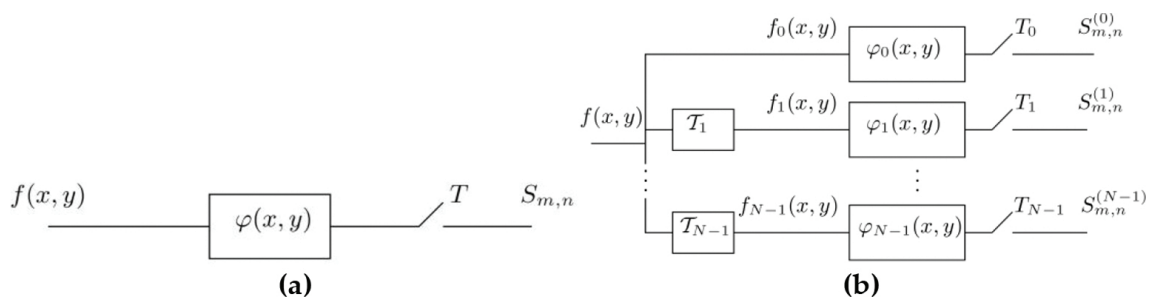
Figure 3. Classic VS distributed acquisition device [11]. a) Classic device, b) Distributed device.

- *Using an external computer*: Directo et al. [10] uses a vision server as the core of the SR system. It organizes the image capture from camera nodes, and processes the images using a high-resolution image reconstruction algorithm. In order to capture low-resolution images, three image transmission protocols, are used. In a similar way, Baboulaz et al. [11] introduced a distributed acquisition system. **Figure 3** shows a classic versus a distributed acquisition device. On the one hand, in the classic case of a single acquisition device, see **Figure 3(a)**, the incoming 2D projection $f(x, y)$ of the 3D scene is first filtered with a smoothing kernel modeling the point spread function of the lens of the camera and returning the set of samples $S_{m,n}$. On the other hand, in a distributed acquisition system, N cameras P_i, $i = 0,…, N−1$, are observing the same 3D scene from different unknown locations. Therefore, the incoming 2D projections $f_i(x, y)$ at each sensor will differ. Every projection $f_i(x, y)$ is the result of a transformation with T_i of the projection of reference $f(x, y)$. By choosing a camera as a reference (e.g., $i = 0$) the distributed acquisition system can be modeled as depicted in **Figure 3(b)**. Examples of transformations T are translation, rotation, or affine transformation according to the observed scene and to the locations of the cameras. Similarly to the single camera case, each sensor outputs a set of samples $S^{(i)}_{m,n}$. Finally, Park et al. [6] proposed a system in which low-resolution images of the same scene with different subpixel displacements from each other, are taken simultaneously by frame grabbers and a controlling software within the computer of the MC system.

- *Using the MC system*: Agrawal et al. [7] used a dedicated device, a microcontroller, to trigger all the cameras of the system, since they found that this was more stable than using a PC's parallel port [12], which could have triggered variations of 1 ms.

- *Using a shared approach with the MC system and a computer*: This approach utilizes the MC system to perform part of the multiimaging capture process. In Ref. [5], the ProFUSION25 camera array outputs raw 8-bit gray-scale images of pixel resolution 640 × 480 using one-shot mode to restrict the possibility of temporal motion of objects in the scene. In such a way, the multiview images captured by this system are well fitted for SR applications. The small baseline between each camera in the array allows the multiple views to sample the high-resolution image appropriately. Then, the images are ready to be sent to a PC using a PCI Express external cable. This connection provides more than 200 MB/s effective bandwidth and transfers 25 images at 25 FPS to the PC. In Ref. [12], a mapping between a conical view and the MC realization is performed by forming overlapping areas on the images of neighboring cameras. Unlike a rotational conical camera, a MC configuration can have multiple observations for some points of the scene. This allows recovering 3D information for these points from multiple view cues using the computational power of external software in a PC. **Figure 4** shows a sample image taken by the six-camera prototype system.
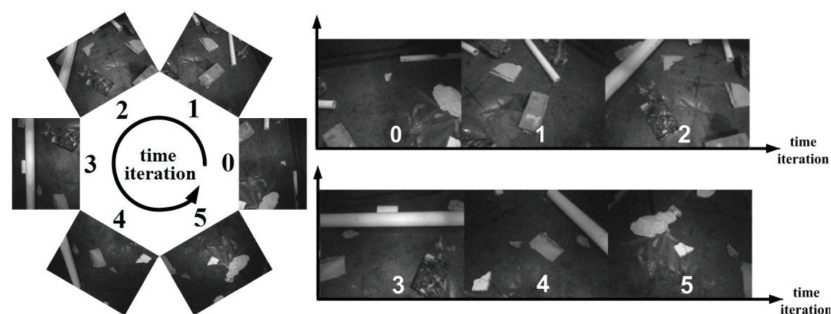
Figure 4. A sample-image taken from a MC conical system [8].

2.4. Preprocessing steps

In an MC set-up, low-resolution images are acquired by different cameras, which have different positions in space and are possibly not synchronized [9]. This causes some spatial and temporal misalignments among the sequences. On the one hand, the temporal misalignment results as a result of the possible frame rate and time offset differences among the cameras and can be modeled by a 1D affine transformation. On the other hand, the spatial misalignment between the two sequences results from the fact that the two cameras could have different internal and external parameters and has been mainly described by one of next two different models:

- *Homography*: Describes the exact image motion of an arbitrary planar surface between two discrete uncalibrated perspective views. The spatial transformation among the low-

resolution sequences can be approximated by a homography when a planar scene assumption can be made [9, 10].

- *Fundamental matrix*: Homography assumption is no longer valid when there are a significant amount of camera translations and nonplanar depth variations. Such scenarios require 3D motion models, which consist a set of local parameters (per pixel) for the representation of the 3D structure and global parameters for the camera motion.

In Ref. [11], the use of the continuous moments, instead of the discrete moments, together with the approach described in Ref. [13], allowed to perform an affine registration of very low-resolution sampled images with the accuracy of the original image.

In Ref. [7], coded sampling is used, demonstrating that it is optimal by considering a linear invertible combination of time samples.
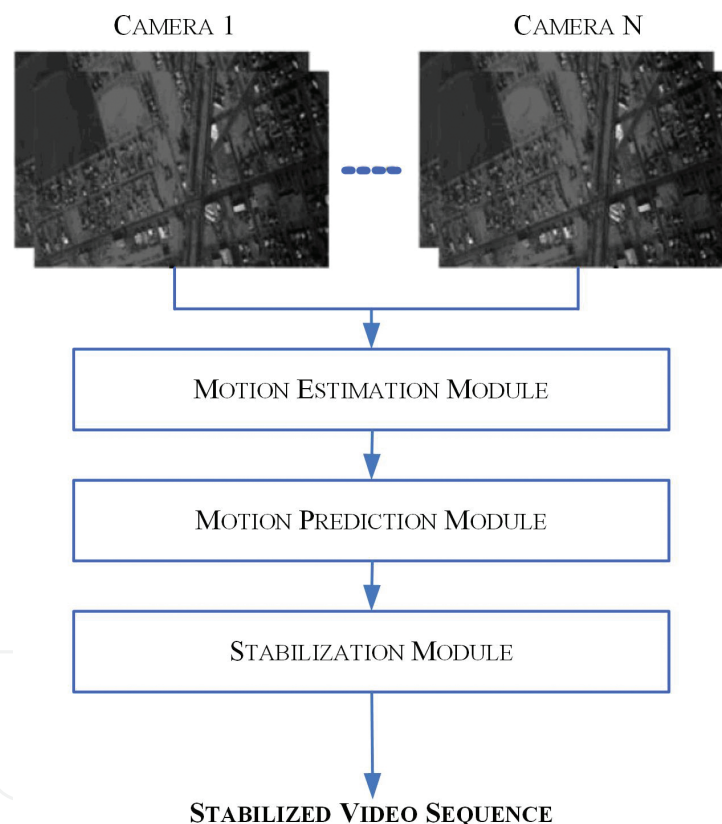
Figure 5. Video stabilization algorithm flow chart [14].

In Ref. [6], preprocessing steps consist of selecting one of the frames to be the referenced, meanwhile the contrast and lightness of the other frames are accordingly adjusted with the reference image by histogram specification. As a result, the relative global motions of the other images are calculated in accordance with the reference frame.

Images rendered by remote sensing MC platforms are basically considered to contain jitter caused by decoding timing delays, target movement and platform motion. The problem of

stabilizing large-frame and low-frame rate imagery acquired from a multicamera array system for persistent surveillance and monitoring is dealt with in Ref. [14] with an algorithm based on temporal coherence properties between the cameras, thus eliminating the need to perform motion estimation on every individual camera sequence. The video stabilization algorithm is shown in **Figure 5** and consists of the following three steps:

1. *Motion estimation module*: It computes the interframe or frame-to-frame (F2F) transformation between adjacent frames for the primary camera(s).

2. *Motion prediction module*: It predicts the F2F transformations for the video frames from the secondary set of cameras, which relies on a prior knowledge of the camera-to-camera transformations and the outputs from the motion estimation module.

3. *Stabilization module*: It temporally aligns the final image sequence from each camera so that the jitter is reduced.

There are also techniques to reduce motion blur as presented in Ref. [15], where the main idea is to capture the same static scene with a hand-held MC array by keeping different exposure settings for different cameras and subsequently reconstruct high-resolution space time volume to get less motion blur image frames.

2.5. Super-resolution

In order to apply SR, many different approaches have been adopted in the literature. A good classification for all of them could be established depending on the approach to the real conditions of the MC system: first, in some cases, a popular technique is used without introducing any modification; second, one of these techniques could be slightly adapted to the camera array; finally, the algorithm may be prepared to be adjusted to the MC system, using an observation model.

• *Using a well-known technique*: There are several approaches to the SR reconstruction of a reference image from multiple still images or video sequences. Among the popular SR recovery, techniques are the projection onto convex sets (POCS) approach [10, 16] and the Bayesian approach [17]. These techniques assume global motion between successive frames of video, as in the case of camera motion with static scenes.

• *Modifying a well-known technique*: In order to consider the MC system, Eren et al. [18] extended POCS method to object-based super-resolution from video by proposing segmentation and validity maps. In the same way, the formulation used in Ref. [6] is an extended form of the SR algorithms in Refs. [19, 20] to estimate the local accuracy of the motion estimation results and incorporate it in the minimization functional as a local regularization parameter. According to the research in Ref. [19], the result of inaccurate motion estimation is proportional to the partial derivatives of the image, which can be interpreted as the amount of high-frequency data.

• *Adjusting the algorithm to the MC system*: In this approach, real conditions of the camera array are assumed. The exposure time, for example, is critical in order to achieve a good quality super-resolved image. In Ref. [17], a Bayesian SR algorithm based on an imaging model is

shown. It includes camera response function, exposure time, sensor noise, and quantization error in addition to spatial blurring and sampling. SR reconstruction is then presented as an inverse problem, where the input q is estimated from a set of observations z_i, as shown in **Figure 6**.
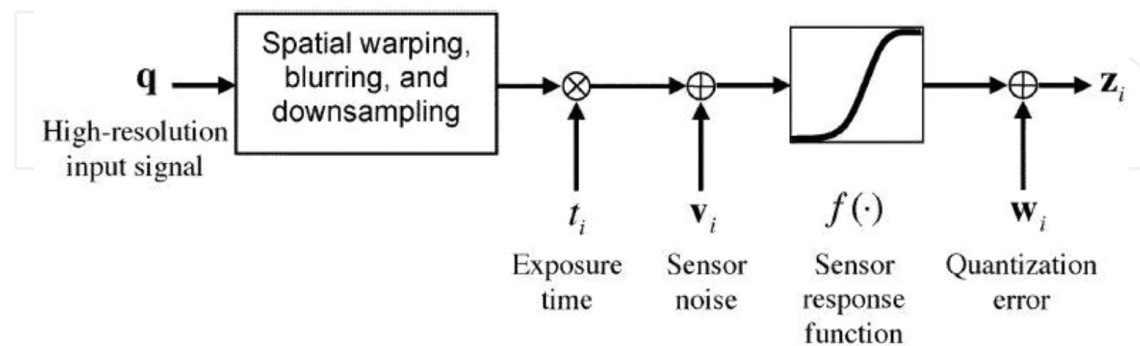

Figure 6. SR algorithm proposed in Ref. [20].

Including characteristics of the camera array in the SR process is also considered in Ref. [5] (see **Figure 7**), where a detailed observation model is integrated in the SR restoration method (see **Figure 8**). This model is sometimes referred to as the forward model to emphasize the fact that SR is an inverse problem (as shown in **Figure 6**). The accurate description of the observation model is vital for the success of the SR process. This involves characterizing the imaging sensor as fully as possible and making appropriate assumptions about the type of scene being imaged.
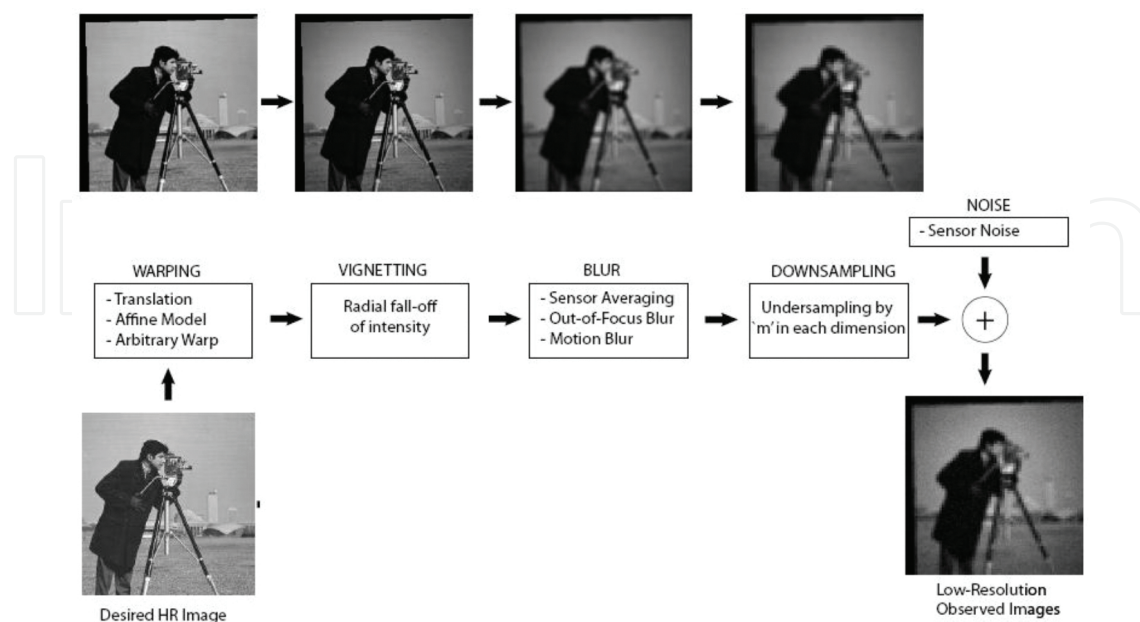
Figure 7. General observation model proposed in Ref. [5].
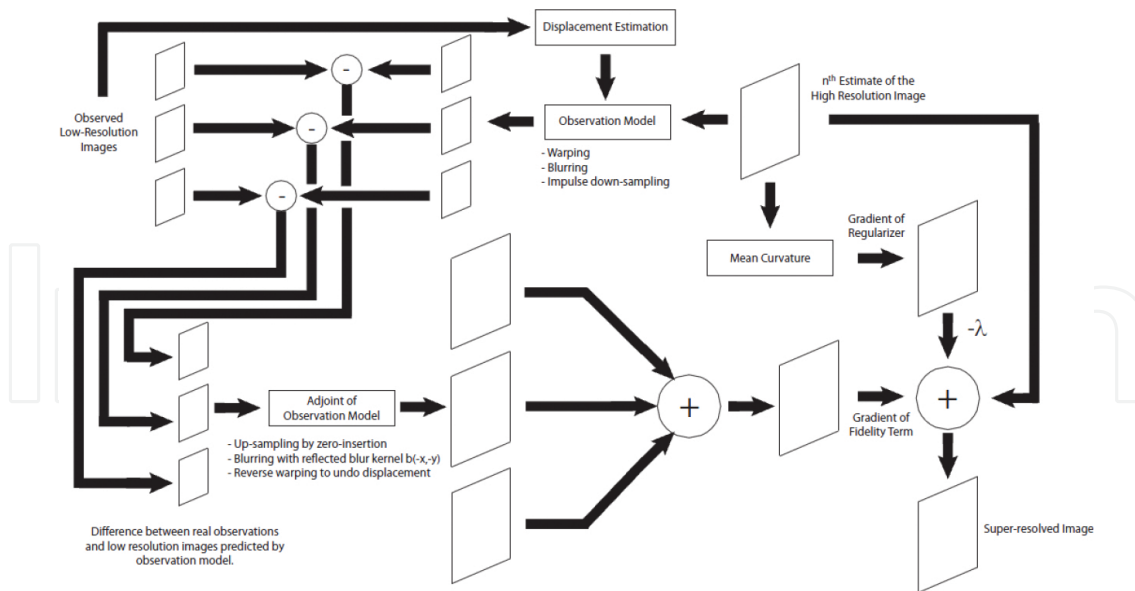
Figure 8. Single iteration in regularized super-resolution restoration [5].

The idea of the observation model is well presented in **Figures 6** and **7**. Including an observation model in the SR restoration minimizes the preprocessing steps. The fundamental components comprise the warp operator, the blur operator, and the downsampling operator:

- *Warp operator*: It describes the existing displacement between two images in a sequence, which could arise from camera motion, object motion in the scene, or a combination of both.

- *Blurring operator*: It defines the cumulative blurring effects from sensor averaging, motion blur, and out-of-focus blur.

- *Downsampling operator*: It applies a magnification factor "*m*" in each dimension (undersampling).

It is also important to consider the noise that is directly added by the system once an image is captured. This is the reason why the sensor noise is added directly in a typical observation model. Besides these fundamental components, there are some specific parameters such as the sensor response function in **Figure 6** or the vignetting operator in **Figure 7**, which are introduced due to the characteristics specific to every system. The observation model is flexible enough to include many different applications. For example, in Ref. [8], a similar observation model, which is represented in **Figure 9**, is presented for a 3D-SR application, only including geometric projection, which is based on the 3D model of the scene and position of every camera. The geometric transformation of X (3D SR scene) to the coordinates of each image (Y_n^L, low-resolution image) is computed using the camera projection model. In this situation, the accuracy of the 3D model and the camera positions are critical to the performance of the 3D-SR algorithm.

The method to measure the quality of the reconstructed image should also be considered. It is demonstrated [21] that although the image quality is usually measured by the expected and the actual mean squared error (MSE), an alternative performance measure might be based on

edge errors, since edges are often the first step in more complex image analysis for both image processing systems and biological systems. Finally, the use of MC commercial applications, as in Ref. [5], is interesting since it can be exploited by different researchers in order to compare the suitability of the SR algorithm.
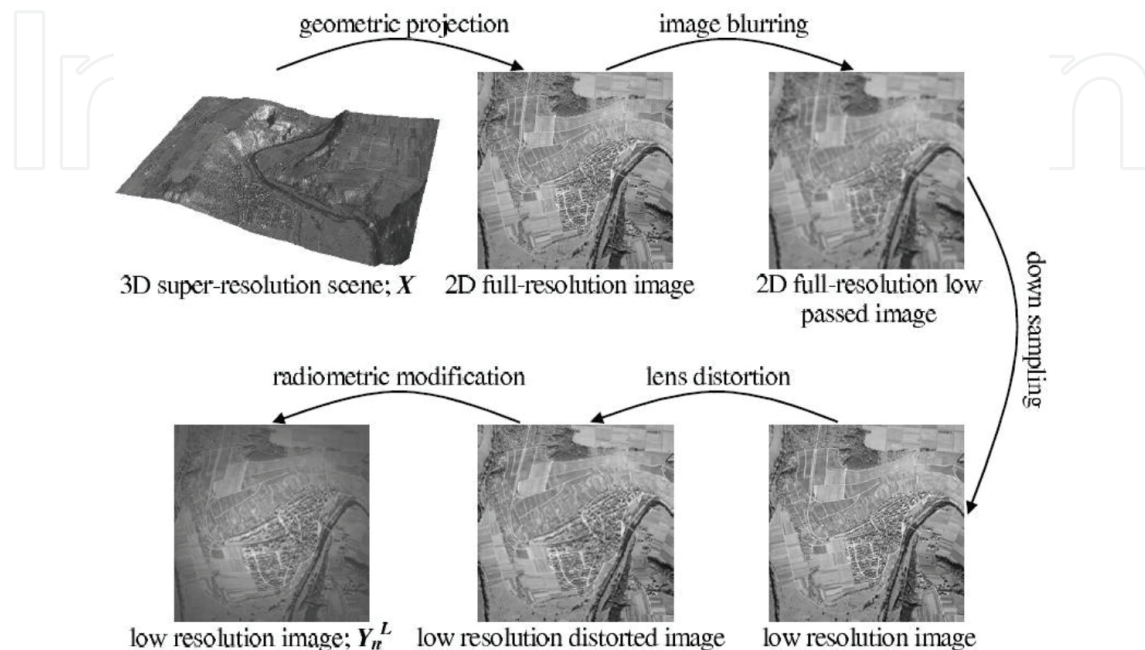
Figure 9. 3D-SR observation model [9].

2.6. Postprocessing

The implementation of postprocessing steps is not very common in the literature. In fact, SR is usually the last link in the chain of SR applied to MC systems. However, there are some researchers that continue working with the images once the SR step has concluded.

- In Ref. [6], the quality of the SR solution can be enhanced with the application of spatially adaptive regularization parameters. Also an image fusion algorithm is applied for merging the high-resolution image reconstructed by the SR algorithm and color channel resolution images. By combining image fusion with the color difference domain, which is widely used in color interpolation, the proposed image fusion algorithm can produce clearer multispectral images, even when the spectral low-resolution channels are not perfectly registered one to another.

- In Ref. [8], as it was explained in Subsection 2.5, it is demonstrated that when an accurate 3D model of the scene is available, or can be estimated, perspective projection of the scene can be exploited in place of the image alignment/warping step in the 2D super-resolution technique, so the 3D adaptation could be considered as part of the observation model or a postprocessing step.

3. Spatial and temporal SR through a MC system

It has been shown in the previous section that, there are several methods where either temporal information between frames captured by a single camera or spatial information between cameras when using a camera array is used to obtain an HR sequence using SR. Accordingly, these method are named in the literature as spatial SR and temporal SR. In this section, the implementation of algorithms to enhance video sequences combining spatial and temporal SR with an MC approach is presented, together with some associated preprocessing steps [22, 23]. According to the block diagram of **Figure 1** this corresponds to Subsections 2.4 and 2.5. The video SR algorithm used as a basis in this work belongs to the "fusion" category. The *baseline super-resolution* (BSR) algorithm execution can be divided into three independent stages: *Motion Estimation, Shift & Add* and *Fill Holes*:

- The *Motion Estimation* stage determines the motion between two or more frames with subpixel accuracy; depending on the selected *scale factor* (i.e., obtaining an output frame whose size is twice the vertical and horizontal size of the input frame would mean *a scale factor* of 2). In order to obtain the motion vectors of each MB, a block-matching method is considered.

- The second stage, known as *Shift & Add*, is executed once the motion vectors have been calculated. A grid is filled with the contributions given by the estimated motion vectors.

- Finally, the *Fill Holes* stage considers that it is possible that there could be some empty positions in the grid for the current frame, as the candidate frames do not contain information enough to fill all the locations. These empty positions are denoted as *holes* in the scope of this work. In this case, a bilinear surface interpolator is used to fill each empty pixel. For each one of the frames, the whole process is repeated. As a result, a HR super-resolved image is obtained and the SR sequence is stored.

Figure 10 summarizes a general scheme using LR frames captured by multiple cameras to generate an HR sequence from a user selected camera. It includes the following steps:
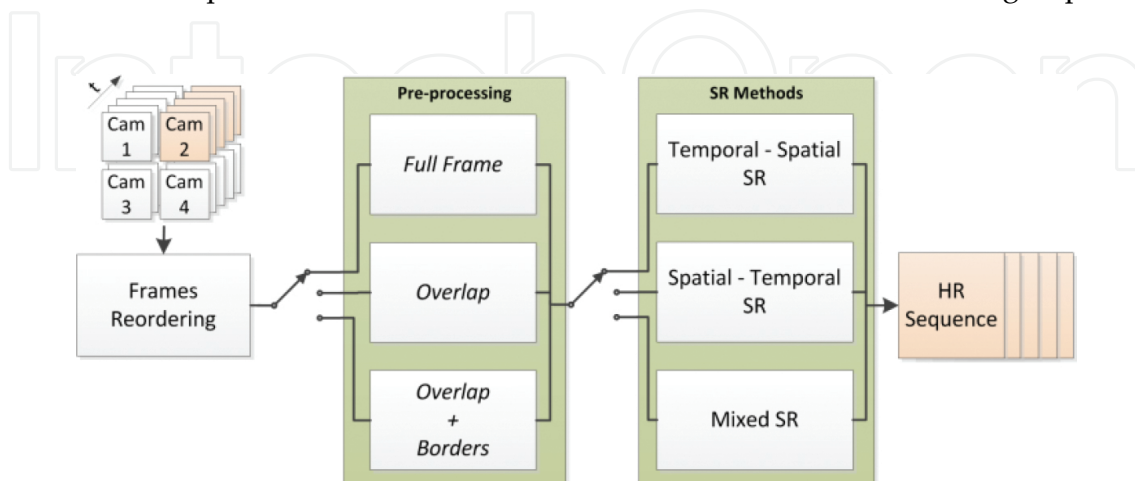
Figure 10. General scheme showing preprocessing modes and SR methods [22].

- First, the frames captured by the cameras are lexicographically reordered.

- Second, a preprocessing step is considered, including three possible options: *Full Frame, Overlap + Borders* and *Overlap*.

- Finally, the SR process is applied considering three methods: *temporal-spatial SR* method, *spatial-temporal SR* method, and *mixed SR* method. These methods combine spatial and temporal information.

3.1. Preprocessing

After reordering the frames recorded from the different cameras of the MC array, the first stage of the algorithm implementation is based on preprocessing algorithms. The target consists on deciding whether some regions of the captured frames should be discarded in order to enhance quality by avoiding artifacts and/or reducing the execution time. Some constraints to the MC array configuration are considered:

- A rectangular geometry.

- Location of the cameras in the same plane.

- Common parts from the same global scene are recorded by the cameras from different locations.

Considering a MC system which complies with these constraints, as the one shown in **Figure 11(a)**, a common region (or *overlap*) of the recorded information by the cameras of the MC system could be available (or a subset of cameras). Surrounding the *overlap*, there will be a *border*, as presented in **Figure 11(b)**. The separation and geometry between cameras of the MC system affects the way to obtain the borders and the *overlap* as shown in **Figure 11(b)**.
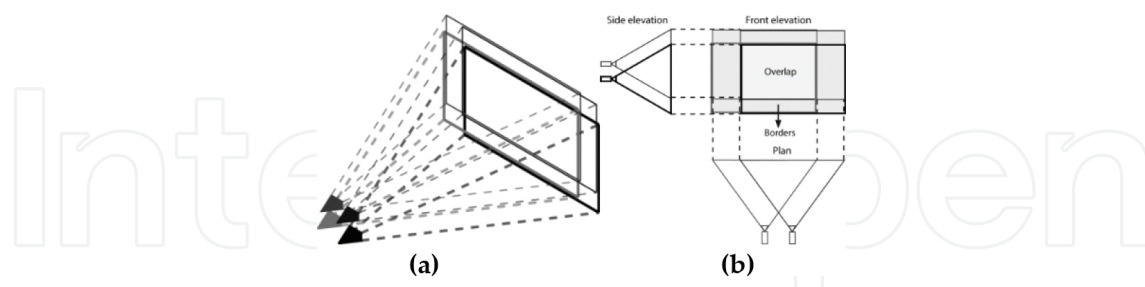

Figure 11. Borders and overlap in an MC array. (a) Frames recorded by an MC array in perspective. (b) Side and front elevations, and plan views of the recorded frames [22].

From this analysis, several ways to process the frames captured by the MC array arise:

- Considering the whole frame information (*Full Frame* mode). This is the basic mode, in which the full frame is captured by every camera.

- Considering the *overlap* between cameras (*Overlap* mode). In order to obtain the *overlap*, it is necessary to know the *offset* in pixels between cameras.

- Considering dividing the frame between the *overlap* and the *borders* (*Overlap + Borders* mode). As shown in **Figure 12**, this method provides 9 different parts to be super-resolved: 4 *sides*, 4 *corners*, and the *overlap*.
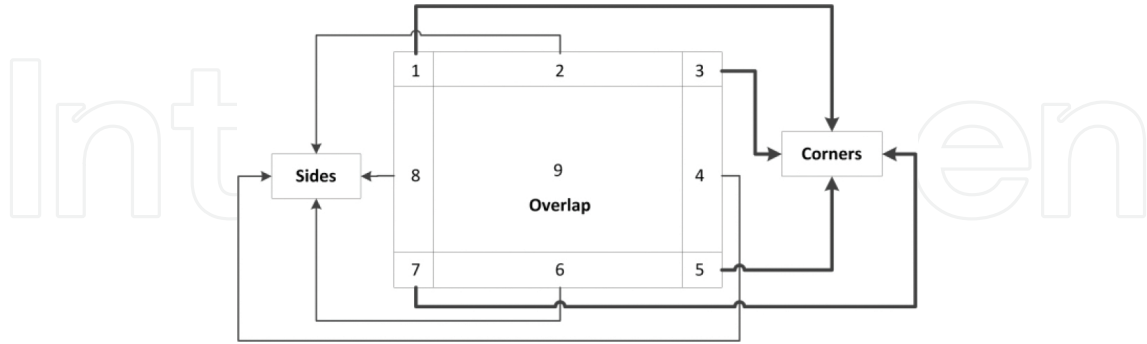
**Figure 12.** Frames division of the MC array in Overlap + Borders mode [22].

3.2. Temporal-spatial SR method

This method considers information provided by the MC array in order to obtain an HR sequence in two phases. In the first phase, only temporal information is considered, while in the second phase spatial information is processed. **Figure 13** presents this method. First, temporal SR is applied to the LR frames of each camera of the MC system. This SR process considers a *temporal working window* (TWW), which determines the number of frames used in the SR process. The output of this phase consists on a sequence of a resolution named *Medium Resolution$_{temporal}$* (MR_t) determined by the MC array dimensions: $P \times Q$.
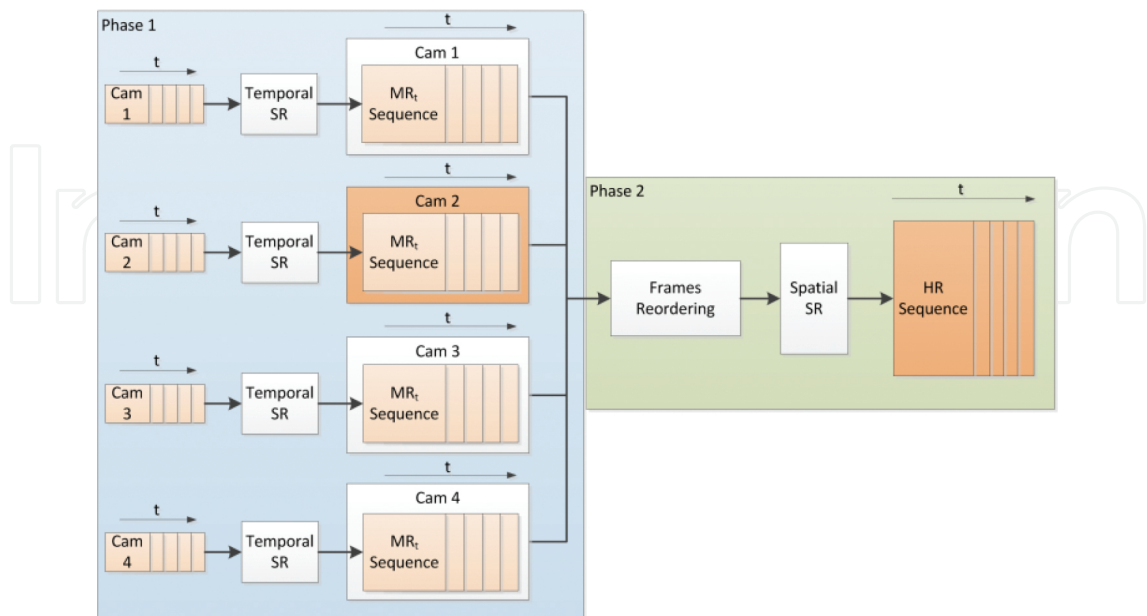
Figure 13. Temporal-spatial SR method [22].

The second phase begins with a *frame* reordering process, storing the frames in lexicographical order, from left to right and top to bottom, as shown in **Figure 14**. Then, the spatial informa-tion of the MR_t sequences is used to obtain a super-resolved HR output sequence. In order to perform the *spatial SR*, the working window considers the frames p_f the *spatial SR* process (*spatial working window*, SWW).
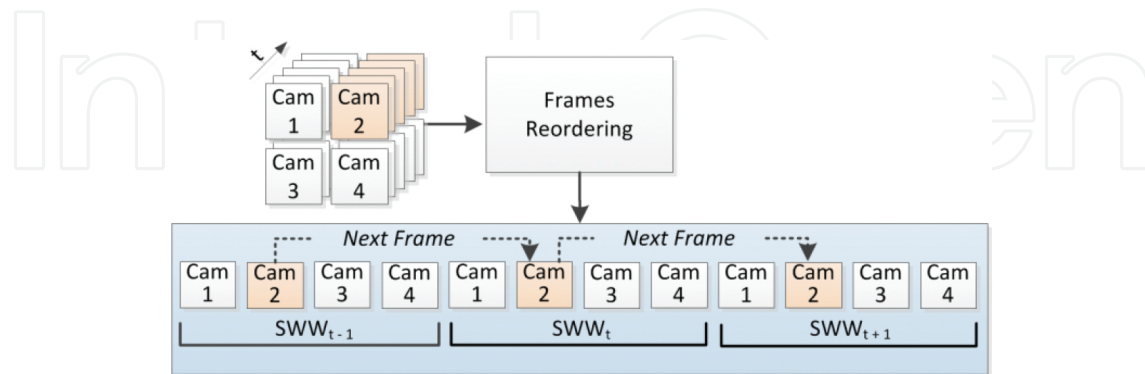

Figure 14. Frames reordering process [22].

The computational cost of this method is high, since the dimensions of the MC array directly affect the number of SR processes to be applied in the first phase, plus a *Spatial SR*, determining a total of $(P \times Q) + 1$ SR processes.

3.3. Spatial-temporal SR method

The *spatial-temporal SR* method, presented in **Figure 15**, is similar to the *temporal-spatial SR* method but reversing the order of the SR processes. In this case *Spatial SR* is applied first. As a result, the output sequence resolution is named *Medium Resolution$_{spatial}$* (MR_s). After applying *Spatial SR* a *Temporal SR* process is applied considering a *temporal working window* as in BSR, and obtaining as output an HR sequence.
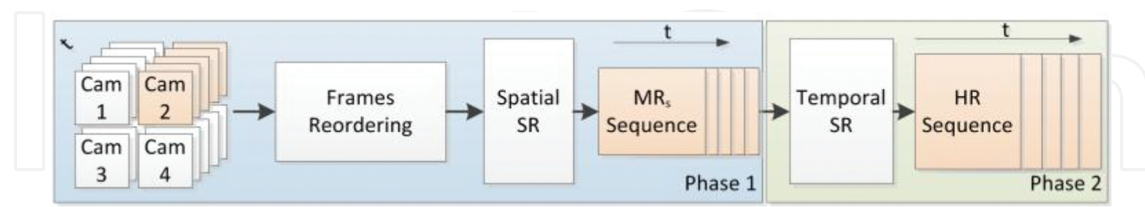
Figure 15. Spatial-temporal SR method [22].

In this case, only two SR processes are applied to obtain the HR output, which reduces considerably the computational cost.
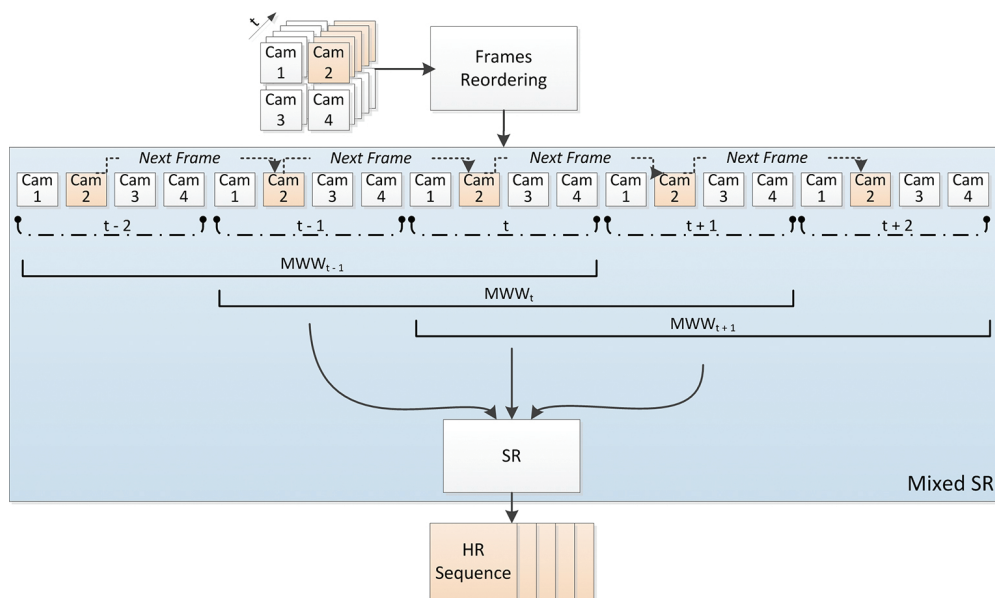
3.4. Mixed SR method

After analyzing the characteristics of the previous methods, the *mixed SR* method was considered. The advantage of this method consists on integrating the spatial and the temporal

information in a combined SR process way to generate the HR output. In this scope, a new WW is defined (*mixed working window*, MWW). The frames selection process is performed in a smart way in the *block matching* process of the *motion estimation* stage, widening the possibilities to find more appropriate information in the SR process.

An example of how MWW is defined is shown in **Figure 16**. A 2 × 2 MC array is selected and SR is applied to camera #2. As it is shown, MWW considers the information of a backward time slot and a forward time slot from the frame to be processed. For instance, the time slot "*t*" considers a WW including both the frames of the MC system captured in the time slots "*t* − 1," "*t*," and "*t* + 1." After processing the frame of the camera #2 in the time slot "*t*," SR is applied to the same frame in the time slot "*t* + 1." In this case, a similar MWW is generated, but considering the time slots "*t*," "*t* + 1," and "*t* + 2" and proceeding in the same way for the subsequent frames.

Figure 16. Mixed SR method [22].

This method reduces the computational cost of the previous presented methods, as it only consists on one SR, but the memory requirements of the algorithm are higher.

3.5. Results

In this section, significant results based both on test sequences adapted for comparison and on real MC acquisition systems are shown. The *Water Cooler* sequence [24] was recorded by using a 5 × 5 MC system. The selected cameras for this sequence are the *leftmost center, topmost center, bottommost center*, and *rightmost center*, forming a rhomboid. Additionally, several tests were completed by using a rectangular 3 × 3 camera array, to demonstrate the versatility of the proposed methods. As the rhomboid configuration of *Water Cooler* has no information of the corners, no results for the preprocessing mode *Overlap + Borders* can be shown.

Sequence/method	Water cooler		Mobcal		Stockholm		Shields		Parkrun	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
TS SR	*28.33*	*0.879*	26.55	0.879	27.05	0.901	26.92	0.927	19.82	0.825
ST SR	27.85	0.868	26.51	0.879	26.64	0.895	26.61	0.923	19.57	0.820
Mixed SR	27.67	0.869	*27.26*	*0.920*	*28.32*	*0.939*	*27.73*	*0.952*	*20.47*	*0.900*
BSR	27.66	0.868	26.47	0.891	26.31	0.891	25.91	0.903	19.30	0.808
INT	27.20	0.858	26.32	0.858	25.47	0.834	25.44	0.870	18.67	0.736

Table 1. PSNR and SSIM results for full frame preprocessing mode (best values are represented in italics).

| Sequence/method	Water cooler		Mobcal		Stockholm		Shields		Parkrun	
	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
TS SR	*27.73*	*0.871*	26.28	0.848	26.70	0.844	27.24	0.885	*19.52*	0.717
ST SR	27.03	0.855	26.07	0.847	26.26	0.830	26.83	0.872	19.20	0.690
Mixed SR	27.31	0.867	*27.34*	*0.893*	*28.26*	*0.912*	*27.51*	*0.905*	19.31	*0.746*
BSR	26.94	0.858	25.82	0.842	25.82	0.826	26.03	0.850	18.65	0.656
INT	26.06	0.838	26.03	0.829	25.06	0.761	25.57	0.809	18.32	0.613

Table 2. AVG PSNR and SSIM results for Overlap preprocessing mode (best values are represented in italics).
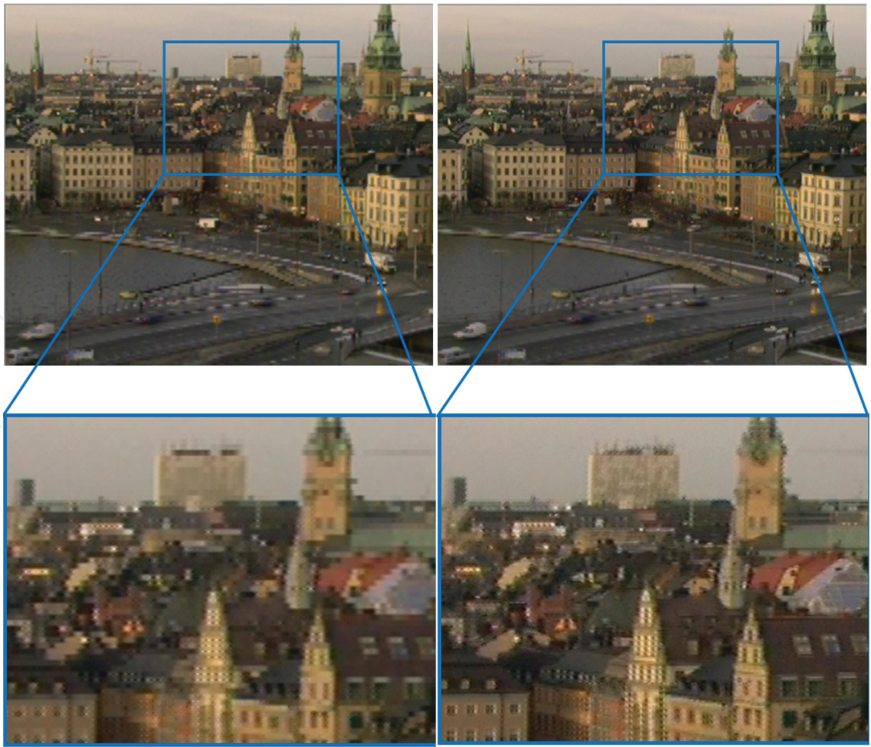
Figure 17. Complete and detailed view of a frame of *Stockholm* sequence [22]. BSR; (left); *Mixed* SR method (right).

Tables 1 and 2 show the results for the sequences under test: *Water Cooler, Mobcal, Stockholm, Shields*, and *Parkrun*. The three presented methods are evaluated using the *peak signal-to-noise ratio* (PSNR) and *structural similarity index* (SSIM) metrics: *temporal-spatial* SR (TS SR), *spatial-temporal* SR (ST SR), and *mixed SR*, together with BSR and interpolation (INT) as a reference. Each table represents the results for the preprocessing modes: *Full Frame* mode (Table 1) and *Overlap* mode (Table 2). As it can be seen from these tables, the preprocessing modes *Full Frame* and *Overlap* follows a similar behavior. It can be concluded from the results in Table 2, that the *mixed SR* method outperforms the other methods in the majority of the cases. In terms of subjective comparison, two sets of frames are shown for the *Stockholm* sequence and the *Water Cooler* sequence in Figures 17 and 18, respectively. In Figure 17, there is a relevant enhancement in the roofs and the facades of the buildings in the *mixed* SR method. Figure 18 shows that a higher definition is provided by the TS SR method versus the BSR, minimizing the number of artifacts in the items on the table [22].
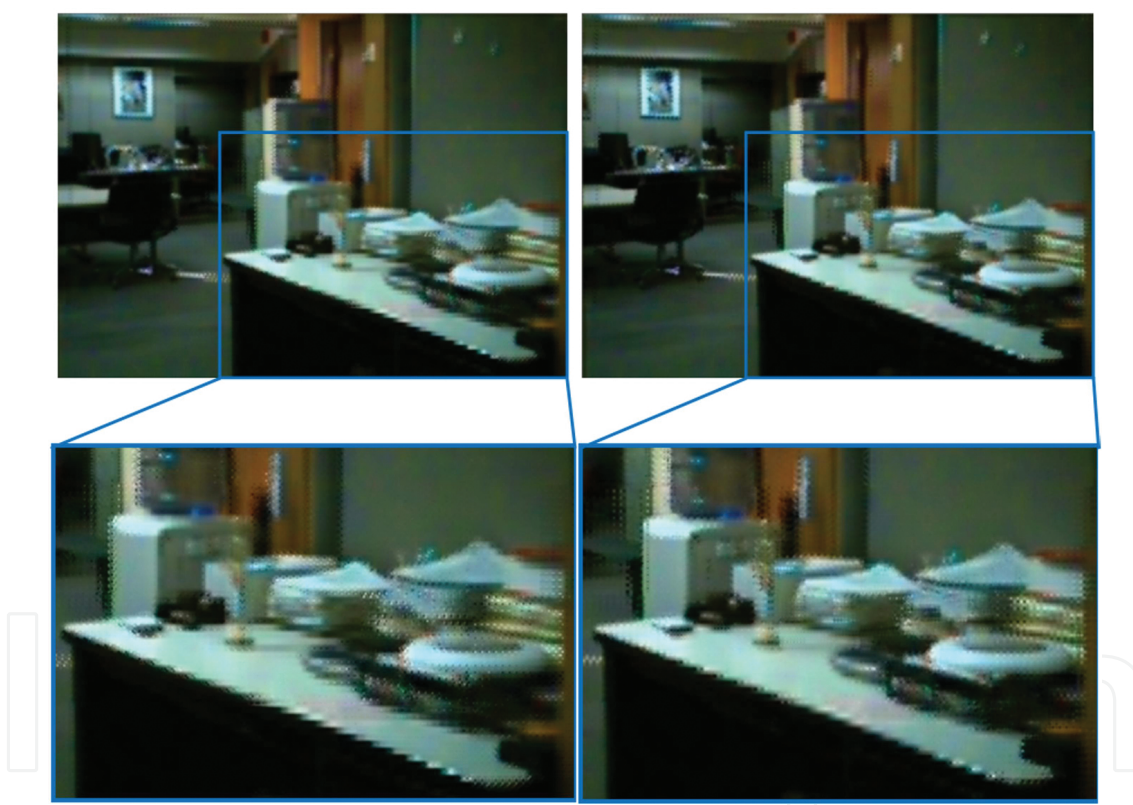
Figure 18. Complete and detailed view of a frame of *Water Cooler* sequence [22]. BSR (left); TS SR (right).

4. Conclusions

In conclusion, the main analyzed characteristics of the studied documents which combine SR and MC are shown in Tables 3–5:

Ref.	Multi-camera system		
	N°	Type	Arrangement
[5]	25	ProFUSION 25	5 × 5 array
[6]	9	CCD VCC-G20E20	3 × 3 array (3 C +6 I)
[7]	4	CCD Dragonfly	2 × 2 array (in PIC micro)
[8]	6–12	CCD cameras	Conical array (6 or 12 cam.)
[9]	2	Pulnix	Closely and overlapped
[10]	2	Logitech Quickcam	Flexible and overlapped
[11]	100	Cheap LR camera	Circularly shifted

Table 3. Multi-camera system.

Ref.	Image characteristics	
	Correspondence	Registration
[5]	One-shot mode	Two-parameter shift
[6]	Lenses adjustment and Blur uniformity	Preprocessing steps
[7]	Accurate trigger using PIC16F690	3D fundamental matrix
[8]	Planar checkerboard	Conical acquisition
[9]	Sequence to sequence alignment	2D homography
[10]	Harris feature points	2D homography
[11]	Continuous moments	Distributed acquisition

Table 4. Image characteristics.

Ref.	SR algorithm	Application	Limitations
[5]	Dense displacement estimation	Rendering	Grayscale cameras and vignetting
[6]	Bayesian multichannel	Rendering	Nonextensible MC system
[7]	Temporal: Point sampling	Rendering	No spatial resolution, only temporal resolution
[8]	3D mosaicing	Underwater	Non-real-time
[9]	POCS	Surveillance	Planar scene assumption and no tracking of the interest region
[10]	POCS + PSF	Rendering	Planar scenes and placement of cameras at a constant distance
[11]	-	Rendering	Restoration step to be improved and don't use "real-world" data

Table 5. SR Algorithm, applications, and limitations.

- *MC system:* Number, type, and position of cameras.

- *Image characteristics:* Correspondence and registration of images.

- *SR algorithm:* Algorithm used to perform the SR.

- *Application*: System use.

- *Limitations:* Not considered issues in the system.

As it can be noticed in the tables, there are limited approaches combining MC system and SR and what is more important, there are some issues which are almost always disregarded, such as:

- The implementation in MC systems of SR algorithms for real-time performance, using approaches as dedicated hardware or distributed systems. This is an important limitation for applications which require real-time solutions such as surveillance.

- The real-time self-reconfiguration of cameras in a camera array for SR applications, which has been used in the past for planning and control as a form of nonuniform sampling (or adaptive capturing) of image-based rendering scenes [4].

- The issue of SR of color images is another important research field. Although some color correction methods for multiview images have been introduced [25], monochrome processing by means of independently applying SR to every color channel is not optimal because it does not take into account the spectral correlation between the channels [26]. If the channels can be decorrelated using a transform like the *Karhunen Loeve Transform* (KLT) [27] or in a suitable color space, then the SR algorithm can be applied to every decorrelated channel separately and transformed back to the original domain or color space. The only reference found which analyzes this issue is the one proposed by Park et al. [6].

- The concept of learning-based SR [28, 29] has not been developed for MC applications. It exploits the prior knowledge between the HR examples and the corresponding LR examples through the so-called learning process. Most example-based SR algorithms usually employ a dictionary composed of a large number of HR patches and their corresponding LR patches, which may be useful for MC applications.

The study of the issues mentioned above is a clear field to research. The potential application of combining SR with a MC system is focused on research areas related to 3D mosaicing, surveillance applied to extreme conditions (underwater or blurred environments), and the improvement of more researched techniques such as medical imaging, video enhancement, remote sensing, or sporting events.

According to this review, it is shown that the exploitation of the spatial and temporal super-resolution is something novel which has been implemented by the authors. Section 3 presented a novel image enhancement SR technique integrated with an MC system to take advantage from the spatial and temporal correlations between the recorded sequences. Three different methods have been proposed: *temporal-spatial SR*, *spatial-temporal SR*, and *mixed SR*. Besides, three different preprocessing steps were introduced: *Full-Frame, Overlap* and *Overlap + Borders*. According to the results [22, 23], it has been shown that the *mixed SR* method with *Full-Frame* preprocessing outperforms the other methods.

Author details

Eduardo Quevedo[1,2*], Gustavo Marrero[1] and Félix Tobajas[1]

*Address all correspondence to: equevedo@iuma.ulpgc.es; eduardo.quevedo@plocan.eu

1 Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria, Las Palmas, Spain

2 Oceanic Platform of the Canary Islands, Carretera de Taliarte, Las Palmas, Spain

References

[1] J. Tiang and K.-K. Ma. A survey on Super-Resolution imaging. Journal of Signal, Image and Video Processing. 2011;5(2):329–342.

[2] T. Goto, Y. Kawamoto, Y. Sakuta, A. Tsutsui, and M.Sakurai. Learning-based super-resolution image reconstruction on multi-core processor. IEEE Transactions on Consumer Electronics. 2012;(3):941–946.

[3] M. M. Islam, V. K. Asari, M. N. Islam, and M. A. Karim. Super-resolution enhancement technique for low resolution video. IEEE Transactions on Consumer Electronics. 2010;56(2):919–924.

[4] B. A. Stancil, C. Zhang, and T. Chen. Active multicamera networks: From rendering to surveillance. IEEE Journal of Selected Topics in Signal Processing. 2008;2(4):597–605.

[5] M. H. Fanaswala. Regularized Super-Resolution of Multi-View Images [dissertation]. Carleton University; 2009.

[6] J. Hyun Park, H. Mook Oh, and M. Gi Kang. Multi-Camera imaging system using Super-Resolution. In: The 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), Shimonoseki; 2008.

[7] A. Agrawal, M. Gupta, A. Veeraraghavan, and S. G. Narasimhan. Optimal coded sampling for temporal super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2010.

[8] P. Firoozfam. Multi-Camera Imaging for 3-D Mapping and Positioning; Stereo and Panoramic Conical Views [dissertation]; 2004.

[9] G. Caner, A. Murat Tekalp, and W. Heinzelman. Super-Resolution recovery for Multi-Camera surveillance imaging. In: International Conference on Multimedia and Expo, Baltimore, MD; 2003.

[10] M. Directo, S. Shirani, and D. Capson. Wireless camera network for image Super-Resolution. In: Canadian Conference on Electrical and Computer Engineering, Ontario; 2004. DOI: 10.1109/CCECE.2004.1345204

[11] L. Baboulaz and P.Dragotti. Distributed acquisition and image Super-Resolution based on continuous moments from samples. In: International Conference on Image Processing; Atlanta, GA. 2006. pp. 3309–3312. DOI: 10.1109/ICIP.2006.312880

[12] A. Agrawal and Y. Xu. Coded exposure deblurring: Optimized codes for PSF estimation and invertibility. In: IEEE Conference on Computer Vision and Pattern Recognition; 20–25 June 2009; Miami, FL. IEEE. pp. 2066–2073. DOI: 10.1109/CVPR.2009.5206685

[13] J. Heikkilä. Pattern matching with affine moment descriptors. Pattern Recognition. 2004;37(9):1825–1834. DOI: 10.1016/j.patcog.2004.03.005

[14] W. D. Reynolds and D. S. Campbell. A scalable video stabilization algorithm for multi-camera systems. In: 6th IEEE International Conference on Advanced Video and Signal Based Surveillance; 2–4 Sept. 2009; Genova. pp. 250–255. DOI: 10.1109/AVSS.2009.91

[15] E. Shechtman, Y. Caspi, and M. Irani. Space-time super resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005;27(4):531–545. DOI: 10.1109/TPAMI.2005.85

[16] A. J. Patti, M. I. Sezan, and A. M. Tekalp. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. IEEE Transactions on Image Processing. 1997;6(8):1064–1076. DOI: 10.1109/83.605404

[17] B. K. Gunturk. High-resolution image reconstruction from multiple differently exposed images. IEEE Signal Processing Letters. 2006;13(4):197–200. DOI: 10.1109/LSP.2005.863693

[18] E. Eren, M. I. Sezan, and A. M. Tekalp. Robust, object-based high-resolution image reconstruction from low-resolution video. IEEE Transactions on Image Processing. 1997;6(10):1446–1451. DOI: 10.1109/83.624970

[19] E. S. Lee and M. G. Kang. Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration. IEEE Transactions on Image Processing. 2003;12(7):826–837. DOI: 10.1109/TIP.2003.811488

[20] M. K. Park, M. G. Kang, and A. K. Katsaggelos. Regularized Super-Resolution image reconstruction considering innacurate motion information. SPIE Optical Engineering. 2007;46(11):1–12.

[21] S. L. Wood, H.-B. Lan, M. P. Christstensen, and D. Rajan. Edge detection performance in super-resolution image reconstruction from camera arrays. In: IEEE Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop; 24–27 Sept. 2006; Teton National Park, WY. pp. 38–43. DOI: 10.1109/DSPWS.2006.265427

[22] E. Quevedo, J. de la Cruz, G. M. Callicó, F. Tobajas, and R. Sarmiento. Video enhancement using spatial and temporal super-resolution from a multi-camera system. IEEE

Transactions on Consumer Electronics. 2014.;60(3):420–428. DOI: 10.1109/TCE. 2014.6937326

[23] E. Quevedo, J. de la Cruz, L. Sánchez, G. M. Callicó, and F. Tobajas. Super resolution with adaptive macro-block topology applied to a multi camera system. IEEE Transactions on Consumer Electronics. 2015;61(2):230–235. DOI: 10.1109/TCE.2015.7150598

[24] B. M. Smith, L. Zhang, H. Jin, and A. Agarwala. Light field video stabilization. In: IEEE International Conference on Computer Vision; Sept. 29 2009–Oct. 2 2009; Kyoto, Japan. pp. 341–348. DOI: 10.1109/ICCV.2009.5459270

[25] F. Shao, G. Jiang, M. Yu, and Y.-S. Ho. Highlight-detection-based color correction method for multiview images. ETRI Journal. 2009;31(4):448–450.

[26] M. K. Ng and N. K. Bose. Mathematical analysis of super-resolution methodology. IEEE Signal Processing Magazine. 2003;20(3):62–74. DOI: 10.1109/MSP.2003.1203210

[27] B. Hunt and O. Kubler. Karhunen-Loeve multispectral image restoration, part I: Theory. IEEE Transactions on Acoustics, Speech and Signal Processing. 1984;32(3):592–600. DOI: 10.1109/TASSP.1984.1164363

[28] S. C. Jeong and B. C. Song. Fast super-resolution algorithm based on dictionary size reduction using k-means clustering. ETRI Journal. 2010;32(4):596–602.

[29] G. Anbarjafari and H. Demirel. Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image. ETRI Journal. 2010;32(3):390–394.