

High Resolution Images from a Sequence of Low Resolution Observations

L. D. Alvarez, R. Molina

Department of Computer Science and A.I.
University of Granada, 18071 Granada, Spain.

A. K. Katsaggelos

Department of Electrical and Computer Engineering
Northwestern University, Evanston, Illinois 60208-3118, U.S.A.

1 Introduction

This chapter discusses the problem of obtaining a high resolution (HR) image or sequences of HR images from a set of low resolution (LR) observations. This problem has also been referred to in the literature by the names of super-resolution (SR) and resolution enhancement (we will be using all terms interchangeably). These LR images are under-sampled and they are acquired by either multiple sensors imaging a single scene or by a single sensor imaging the scene over a period of time. For static scenes the LR observations are related by global subpixel shifts, while for dynamic scenes they are related by local subpixel shifts due to object motion (camera motion, such as panning and zooming can also be included in this model). In this chapter we will be using the terms image(s), image frame(s), and image sequence frame(s) interchangeably.

This is a problem encountered in a plethora of applications. Images and video of higher and higher resolution are required, for example, in scientific (e.g., medical, space exploration, surveillance) and commercial (e.g., entertainment, high-definition television) applications. One of the early applications of high resolution imaging was with Landsat imagery. The orbiting satellite would go over the same area every 18 days, acquiring misregistered images. Appropriately combining these LR images produced HR images of the scene.

Increasing the resolution of the imaging sensor is clearly one way to increase the resolution of the acquired images. This solution however may not be feasible due to the increased associated cost and the fact that the shot noise increases during acquisition as the pixel size becomes smaller. On the other hand, increasing the chip size to accommodate the larger number of

pixels increases the capacitance, which in turn reduces the data transfer rate. Therefore, signal processing techniques, like the ones reviewed in this chapter, provide a clear alternative.

A wealth of research considers modeling the acquisition of the LR images and providing solutions to the HR problem. Literature reviews are provided in [1] and [2]. Work traditionally addresses the resolution enhancement of image frames that are filtered (blurred) and down-sampled during acquisition and corrupted by additive noise during transmission and storage. More recent work, however, addresses the HR problem when the available LR images are compressed using any of the numerous image and video compression standards or a proprietary technique. In this chapter we address both cases of uncompressed and compressed data.

To illustrate the HR problem consider Figs. 1 and 2. In Fig. 1(a) four LR images are shown of the same scene. The LR images are under-sampled and they are related by global sub-pixel shifts which are assumed to be known. In Fig. 1(b) the HR image obtained by bilinearly interpolating one of the LR images is shown (the upper left image of Fig. 1(a) was used). In Fig. 1(c) the HR image obtained by one of the algorithms described later in the chapter by combining the four LR images is shown. As can be seen considerable improvement can be obtained by combining the information contained in all four images. In Figs. 2(a) and 2(b) three consecutive HR frames and the corresponding LR frames of a video sequence are shown, respectively. The LR frames resulted by blurring, downsampling and compressing the HR frames using the MPEG-4 standard at 128 Kbps. In Fig. 2(c) the bilinearly interpolated frames corresponding to the LR frames in Fig. 2(b) are shown, while in Fig. 2(d) the HR frames obtained by one of the algorithms described later in the chapter are shown. Again, a considerable benefit results by appropriately utilizing the available information.

An important question is what makes HR image reconstruction possible. We address this question with the aid of Fig. 3. In it the grid locations of four LR images are shown. Each of these images represents a sub-sampled (i.e. aliased) version of the original scene. The four images however are related by subpixel shifts, as indicated in Fig. 3. Each LR image therefore contains complimentary information. With exact knowledge of shifts the four LR images can be combined to generate the HR image shown in the right-hand side of Fig. 3. If we assume that the resolution of this HR image is such that the Nyquist sampling criterion is satisfied, this HR image then represents an accurate representation of the original (continuous) scene. Aliased information can therefore be combined to obtain an alias-free image.

This example shows the simplest scenario and the ideal situation. However, there are several problems that should be taken into account, namely, the fact that the displacements between images may not be known in a realistic application and the motion may not be global but local instead. Furthermore, there are the blur and downsampling processes that have to be taken into account together with the noise involved in the observation process. This noise will even be more complicated when the images (or sequences) are compressed.

In this chapter we overview the methods proposed in the literature to obtain HR still images and image sequences from LR sequences. In conducting this overview we develop and present all techniques within the Bayesian framework. This adds consistency to the presentation and facilitates comparison between the different methods.

The concept of an LR sequence will be understood in the broad sense to include the cases of global motion between images in the sequence and different motions of the objects in the image. We will also examine the case when the LR images have additionally been compressed. It is interesting to note that not much work has been reported on LR compressed observations with global motion between the observed images.

The HR techniques we discuss in this chapter do not directly consider the problem of HR video sequence reconstruction but instead concentrate on obtaining an HR still image from a short LR image sequence segment. All of these techniques may, however, be applied to video sequences by using a sliding window for processing frames. For a given high resolution frame, a "sliding window" determines the set of low resolution frames to be processed to produce the output.

The rest of this chapter is organized as follows. In section 2 the process to obtain the low resolution observations from the high resolution image is described. In section 3 the regularization terms used in high resolution problems are presented; these regularization terms will include in some cases only the high resolution image while in others they will include information about the motion vectors as well. In section 4 the methods proposed to estimate the high resolution images together with the high resolution motion vectors will be presented. Both regularization and observation process depend on unknown parameters; their estimation will be discussed in section 5. Finally, in section 6 we will discuss new approaches to the high resolution problem.

2 Obtaining low resolution observations from high resolution images

In this section we describe the model for obtaining the observed LR images from the HR targeted image. We include both cases of a static image and an image frame from a sequence of images capturing a dynamic scene. As we have discussed in the previous section, the LR sequence may or may not be compressed. We first introduce the case of no compression and then extend the model to include compression. In this section notation is also introduced.

2.1 Uncompressed Observations

The pictorial depiction of the acquisition system is shown in Fig. 4, in which the LR time-varying scene is captured by the camera (sensor). Let us denote the underlying HR image in

the image plane coordinate system by $f(x, y, t)$ where t denotes temporal location and x and y represent the spatial coordinates. The size of these HR images is $PM \times PN$, where $P > 1$ is the magnification factor.

Using matrix-vector notation, the $PM \times PN$ images can be transformed into a $PM \times PN$ column vector, obtained by lexicographically ordering the image by rows. The $(PM \times PN) \times 1$ vector that represents the l -th image in the HR sequence will be denoted by \mathbf{f}_l with $l = 1, \dots, L$. The HR image we seek to estimate will be denoted by \mathbf{f}_k .

Frames within the HR sequence are related through time. Here we assume that the camera captures the images in a fast succession and so we write

$$f_l(a, b) = f_k(a + d_{l,k}^x(a, b), b + d_{l,k}^y(a, b)) + n_{l,k}(a, b), \quad (1)$$

where $d_{l,k}^x(a, b)$ and $d_{l,k}^y(a, b)$ denote the horizontal and vertical components of the displacement, that is, $d_{l,k}(a, b) = (d_{l,k}^x(a, b), d_{l,k}^y(a, b))$, and $n_{l,k}(a, b)$ is the noise introduced by the motion compensation process. The above equation relates an HR gray level pixel value at location (a, b) at time l to the gray level pixel value of its corresponding pixel in the HR image we want to estimate \mathbf{f}_k . For an image model that takes into account occlusion problems see [3].

We can rewrite Eq. (1) in matrix-vector notation as

$$\mathbf{f}_l = \mathbf{C}(\mathbf{d}_{l,k})\mathbf{f}_k + \mathbf{n}_{l,k}, \quad (2)$$

where $\mathbf{C}(\mathbf{d}_{l,k})$ is the $(PM \times PN) \times (PM \times PN)$ matrix that maps frame \mathbf{f}_l to frame \mathbf{f}_k , and $\mathbf{n}_{l,k}$ is the registration noise.

In Eq. (2) we have related the HR images, now we have to relate the LR images to their corresponding HR ones, since we will only have LR observations. Note that the LR sequence results from the HR sequence through filtering and sampling. Camera filtering and down-sampling the HR sequence produces at the output the LR sequence. This LR discrete sequence will be denoted by $g(i, j, t)$, where i and j are integer numbers that indicate spatial location and t is an integer time index. The size of the LR images is $M \times N$. Using matrix-vector notation, each LR image will be denoted by \mathbf{g}_l , where l is the time index of the image vector with size $(M \times N) \times 1$. The LR image \mathbf{g}_l is related to the HR image \mathbf{f}_l by

$$\mathbf{g}_l = \mathbf{A}\mathbf{H}\mathbf{f}_l \quad l = 1, 2, 3, \dots, \quad (3)$$

where matrix \mathbf{H} of size $(PM \times PN) \times (PM \times PN)$ describes the filtering of the HR image, and \mathbf{A} is the down-sampling matrix with size $MN \times (PM \times PN)$. We assume here for simplicity that all the blurring matrices \mathbf{H} are the same, although they can be time dependent. The matrices \mathbf{A} and \mathbf{H} are assumed to be known. We also assume that there is no noise in Eq. (3). Either we can add a new term which represents this noise or we can model it later when combining this equation with Eq. (2).

Equation (3) expresses the relationship between the low and HR frames \mathbf{g}_l and \mathbf{f}_l , while Eq. (2) expresses the relationship between frames l and k in the HR sequence. Using these equations we can now obtain the following equation that describes the acquisition system for an LR image \mathbf{g}_l from the HR image \mathbf{f}_k that we want to estimate

$$\mathbf{g}_l = \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k + \mathbf{e}_{l,k} , \quad (4)$$

where $\mathbf{e}_{l,k}$ is the acquisition noise. This process is pictorially depicted in Fig. (5).

If we assume that the noise during the acquisition process is Gaussian with zero mean and variance σ^2 , denoted by $N(0, \sigma^2 I)$, the above equation gives rise to

$$P(\mathbf{g}_l | \mathbf{f}_k, \mathbf{d}_{l,k}) \propto \exp \left[-\frac{1}{2\sigma^2} \|\mathbf{g}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k\|^2 \right] . \quad (5)$$

Note that the above equation shows the explicit dependency of \mathbf{g}_l on both the HR image \mathbf{f}_k and the motion vectors $\mathbf{d}_{l,k}$. Both image and motion vector are unknown on most HR problems. This observation model was used by Hardie *et al.* [4], Elad and Feuer [5], Nguyen *et al.* [6] and Irani and Peleg [7], among others. The acquisition model of the HR frequency methods initially proposed by [8] can be also understood by using this model (see [9] for an excellent review of frequency based super resolution methods).

A slightly different model is proposed by Stark and Oskoui [10] and Tekalp *et al.* [11], see also [12] and [13]. The observation model used by these authors is oriented toward the use of the Projections Onto Convex Sets (POCS) method in HR problems. This model results in one case as the limit of Eq. (5) when $\sigma = 0$ and in others imposes constraints on the maximum value of the difference between each component of \mathbf{g}_l and $\mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k$. Note that this corresponds to uniform noise modelling. We will encounter these noise models again in the compressed case.

2.2 Compressed Observations

When the LR images have also been compressed we have to relate them to the HR image we want to estimate, \mathbf{f}_k . The new scenario is shown in Fig. 6.

When we have the compressed LR sequence, the uncompressed LR sequence is no longer available, so we cannot use \mathbf{g}_l , $l = 1, \dots, L$. These images are however used as an intermediate step since the compressed LR sequence results from the compression process applied to the LR sequence.

Let us now briefly describe a hybrid motion-compensated video compression process. The LR frames are compressed with a video compression system resulting in $y(i, j, t)$, where i and j are integer numbers that indicate spatial location and t is a time index. The size of the LR compressed images is $M \times N$. Using matrix-vector notation, each LR compressed image will be denoted by \mathbf{y}_l , where l is the time index and the image size is $(M \times N) \times 1$. The

compression system also provides the motion vectors $v(i, j, l, m)$ that predict pixel $y(i, j, l)$ from some previously coded \mathbf{y}_m . These motion vectors that predicts \mathbf{y}_l from \mathbf{y}_m are represented by the $(2 \times M \times N) \times 1$ vector $\mathbf{v}_{l,m}$ that is formed by stacking the transmitted horizontal and vertical offsets.

During compression, frames are divided into blocks that are encoded with one of two available methods, intracoding or intercoding. For the first one, a linear transform such as the DCT (Discrete Cosine Transform) is applied to the block (usually of size 8×8). The operator decorrelates the intensity data and the resulting transform coefficients are independently quantized and transmitted to the decoder. For the second method, predictions for the blocks are first generated by motion compensating previously transmitted image frames. The compensation is controlled by motion vectors that define the spatial and temporal offset between the current block and its prediction. Computing the prediction error, transforming it with a linear transform, quantizing the transform coefficients, and transmitting the quantized information refine the prediction.

Using all this information, the relationship between the acquired LR frame and its compressed observation becomes

$$\mathbf{y}_l = T^{-1}Q \left[T \left(\mathbf{g}_l - MC_l(\mathbf{y}_l^P, \mathbf{v}_l) \right) \right] + MC_l(\mathbf{y}_l^P, \mathbf{v}_l) \quad l = 1, \dots, L, \quad (6)$$

where $Q[\cdot]$ represents the quantization procedure, T and T^{-1} are the forward and inverse-transform operations, respectively, and $MC_l(\mathbf{y}_l^P, \mathbf{v}_l)$ is the motion compensated prediction of \mathbf{g}_l formed by motion compensating the appropriate previously decoded frame/frames depending on whether the current frame at l is an I , P or B frame, see [14]. This process is pictorially depicted in Fig. 7.

Note that, to be precise, we should make clear that MC_l depends on \mathbf{v}_l and only a subset of $\mathbf{y}_1, \dots, \mathbf{y}_L$; however we will keep the above notation for simplicity and generality.

We now have to rewrite Eq. (6) using the HR image we want to estimate. The simplest way is to write

$$\begin{aligned} \mathbf{y}_l &= T^{-1}Q \left[T \left(\mathbf{g}_l - MC_l(\mathbf{y}_l^P, \mathbf{v}_l) \right) \right] + MC_l(\mathbf{y}_l^P, \mathbf{v}_l) \\ &\approx \mathbf{g}_l \quad \text{and using Eq. (4)} \\ &= \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k + \mathbf{e}_{l,k}. \end{aligned}$$

If we assume that the noise term $\mathbf{e}_{l,k}$ is $N(0, \sigma_l^2 I)$ (see [15] and references therein) we have

$$P(\mathbf{y}_l | \mathbf{f}_k, \mathbf{d}_{l,k}) \propto \exp \left[-\frac{1}{2\sigma_l^2} \|\mathbf{y}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k\|^2 \right]. \quad (7)$$

Two additional models for the quantization noise have appeared in the literature. The first one, called the *quantization constraint*, (see Altunbasak *et al.* [16], [17], and also Segall *et al.*

[18]) has associated probability distribution given by

$$P_{QC}(\mathbf{y}_l | \mathbf{f}_k, \mathbf{d}_{l,k}) = \begin{cases} \text{constant if } -\frac{q(i)}{2} \leq [\mathbf{T}(\mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k - MC_l(\mathbf{y}_l^P, \mathbf{v}_l))] (i) \leq \frac{q(i)}{2}, \forall i \\ 0 \text{ elsewhere} \end{cases}, \quad (8)$$

where $q(i)$ is the quantization factor for the coefficient i .

The second model is a Gaussian distribution and it models the quantization noise as a linear sum of independent noise processes, see Park *et al.* [19], [20] and Segall *et al.* [18], [21], [15], and [22]. This second distribution is written as

$$P_K(\mathbf{y}_l | \mathbf{f}_k, \mathbf{d}_{l,k}) \propto \exp \left[-\frac{1}{2} (\mathbf{y}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k)^T \mathbf{K}_Q^{-1} (\mathbf{y}_l - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k) \right], \quad (9)$$

where \mathbf{K}_Q is the covariance matrix that describes the noise (see Segall *et al.* [15] for details).

The LR motion vectors \mathbf{v}_l used to predict the LR images are available after compression and so we can also include them in the observation process. We are interested in using the LR motion vectors $\mathbf{v}_{l,k}$ provided by the decoder; note, however, that not all \mathbf{y}_l are, during the coding process, predicted from frame k .

Although it has been less studied in the literature than the quantization noise distribution, the LR motion vectors are also important for estimating \mathbf{f}_k and \mathbf{d} . Various authors have provided different approaches to model $P(\mathbf{v}_{l,k} | \mathbf{f}_k, \mathbf{d}_{l,k}, \mathbf{y}_l)$. Chen and Shultz [23] propose the following distribution

$$P_{CS}(\mathbf{v}_{l,k} | \mathbf{f}_k, \mathbf{d}_{l,k}, \mathbf{y}_l) = \begin{cases} \text{constant if } |v_{l,k}(j) - [\mathbf{A}_D \mathbf{d}_{l,k}](j)| \leq \Delta, \forall j \\ 0 \text{ elsewhere} \end{cases}, \quad (10)$$

where \mathbf{A}_D is a matrix that maps the displacements to the LR grid, Δ denotes the maximum difference between the transmitted motion vectors and estimated displacements, and $[\mathbf{A}_D \mathbf{d}_{l,k}](j)$ is the j^{th} element of the vector $\mathbf{A}_D \mathbf{d}_{l,k}$. This distribution enforces the motion vectors to be close to the actual sub-pixel displacements and represents a reasonable condition.

A similar idea, is proposed in Mateos *et al.* [24] where the following distribution is used

$$P_M(\mathbf{v}_{l,k} | \mathbf{f}_k, \mathbf{d}_{l,k}, \mathbf{y}_l) \propto \left[-\frac{\gamma_l}{2} \|\bar{\mathbf{v}}_{l,k} - \mathbf{d}_{l,k}\|^2 \right], \quad (11)$$

where γ_l specifies the similarity between the transmitted and estimated information and $\bar{\mathbf{v}}_{l,k}$ denote an upsampled version of $\mathbf{v}_{l,k}$ to the size of the HR motion vectors.

Segall *et al.* in [21] and [22] model the displaced frame difference within the encoder using the distribution

$$P_K(\mathbf{v}_{l,k} | \mathbf{f}_k, \mathbf{d}_{l,k}, \mathbf{y}_l) \propto \left[-\frac{1}{2} (MC_l(\mathbf{y}_l^P, \mathbf{v}_l) - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k)^T \mathbf{K}_{MV}^{-1} (MC_l(\mathbf{y}_l^P, \mathbf{v}_l) - \mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k) \right], \quad (12)$$

where \mathbf{K}_{MV} is the covariance matrix for the prediction error between the original frame once in LR ($\mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k$) and its motion compensation estimate ($MC_l(\mathbf{y}_l^P, \mathbf{v}_l)$).

Note that from the observation models of the LR compressed images and motion vectors we have described, we can write the joint observational model of the LR compressed images and LR motion vectors given the HR image and motion vectors as

$$P(\mathbf{y}, \mathbf{v}|\mathbf{f}_k, \mathbf{d}) = \prod_l P(\mathbf{y}_{l,k}|\mathbf{f}_k, \mathbf{d}_{l,k})P(\mathbf{v}_{l,k}|\mathbf{f}_k, \mathbf{d}_{l,k}, \mathbf{y}_l). \quad (13)$$

3 Regularization in HR

The super-resolution problem is an ill-posed problem. Given a (compressed or uncompressed) LR sequence, the estimation of the HR image and motion vectors maximizing any of the conditional distributions that describe the acquisition models shown in the previous section, is a typical example of an ill-posed problem. Therefore, we have to regularize the solution or, using statistical language, introduce *a priori* models on the HR image and/or motion vectors.

3.1 Uncompressed Observations

Maximum likelihood (ML), maximum *a posteriori* (MAP), and the set theoretic approach using projection onto convex sets (POCS) can be used to provide solutions of the SR problem (see Elad and Feuer [5]).

In all published work on HR in the literature it is assumed that the variables \mathbf{f}_k and \mathbf{d} are independent, that is,

$$P(\mathbf{f}_k, \mathbf{d}) = P(\mathbf{f}_k)P(\mathbf{d}). \quad (14)$$

Some HR reconstruction methods give the same probability to all possible HR images \mathbf{f}_k and motion vectors \mathbf{d} (see for instance Stark and Oskoui [10] and Tekalp *et al.* [11]). This would also be the case of the work by Irani and Peleg [7] (see also references in [9] for the so called *simulate and correct methods*).

Giving the same probability to all possible HR images \mathbf{f}_k and motion vectors \mathbf{d} is equivalent to using the non-informative prior distributions

$$P(\mathbf{f}_k) \propto \text{constant}, \quad (15)$$

and

$$P(\mathbf{d}) \propto \text{constant}. \quad (16)$$

Note that although POCS is the method used in [10], [11] to find the HR image, no prior information is included on the image we try to estimate.

Most of the work on POCS for HR estimation, use the acquisition model that imposes constraints on the maximum difference between each component of \mathbf{g}_l and $\mathbf{AHC}(\mathbf{d}_{l,k})\mathbf{f}_k$. This

model corresponds to uniform noise modelling, with no regularization on the HR image. See however [9] for the introduction of convex sets as *a priori* constraints on the image using the POCS formulation, and also Tom and Katsaggelos [25].

What do we know *a priori* about the HR images?. We expect the images to be smooth within homogeneous regions. A typical choice to model this idea is to use the following prior distribution for \mathbf{f}_k

$$P(\mathbf{f}_k) \propto \exp \left[- \left(\frac{\lambda_1}{2} \|\mathbf{Q}_1 \mathbf{f}_k\|^2 \right) \right], \quad (17)$$

where \mathbf{Q}_1 represents a linear high-pass operation that penalizes the estimation that is not smooth and λ_1 controls the variance of the prior distribution (the higher the value of λ_1 the smaller the variance of the distribution).

There are many possible choices for the prior model on the HR original image. Two well known and used models are the Huber's type proposed by Schultz and Stevenson [3] and the model proposed by Hardie *et al.* [4] (see Sung *et al.* [26] for additional references on prior models). More recently, total variation methods [27], anisotropic diffusion [28] and compound models [29] have been applied to super resolution problems.

3.2 Compressed Observations

Together with smoothness, the additional *a priori* information we can include when dealing with compressed observations is that the HR image should not be affected by coding artifacts.

The distribution for $P(\mathbf{f}_k)$ is very similar to the one shown in Eq. (17), but now we have an additional term in it that enforces smoothness in the LR image obtained from the HR image using our model. The equation becomes

$$P(\mathbf{f}_k) \propto \exp \left[- \left(\frac{\lambda_3}{2} \|\mathbf{Q}_3 \mathbf{f}_k\|^2 + \frac{\lambda_4}{2} \|\mathbf{Q}_4 \mathbf{A} \mathbf{H} \mathbf{f}_k\|^2 \right) \right], \quad (18)$$

where \mathbf{Q}_3 represents a linear high-pass operation that penalizes the estimation that is not smooth, \mathbf{Q}_4 represents a linear high-pass operator that penalizes estimates with block boundaries, λ_3 and λ_4 control the weight of the norms. For a complete study of all the prior models used in this problems see Segall *et al.* [15].

In the compressed domain, constraints have also been used on the HR motion vectors. Assuming that the displacements are independent between frames (an assumption that maybe should be reevaluated in the future), we can write

$$P(\mathbf{d}) = \prod_l P(\mathbf{d}_{l,k}). \quad (19)$$

We can enforce \mathbf{d}_k to be smooth within each frame and then write

$$P(\mathbf{d}_{l,k}) \propto \exp \left[- \frac{\lambda_5}{2} \|\mathbf{Q}_5 \mathbf{d}_{l,k}\|^2 \right], \quad (20)$$

where \mathbf{Q}_5 represents a linear high-pass operation that, once again, penalizes the displacement estimation that is not smooth and λ_5 controls the variance of the distribution (see again Segall *et al.* [15] for details).

4 Estimating HR images

Having described in the previous sections the resolution image and motion vector priors and the acquisition models used in the literature, we turn now our attention to computing the HR frame and motion vectors.

Since we have introduced both priors and conditional distributions, we can apply the Bayesian paradigm in order to find the maximum of the posterior distribution of HR image and motion vectors given the observations.

4.1 Uncompressed sequences

For uncompressed sequences our goal becomes finding $\hat{\mathbf{f}}_k$ and $\hat{\mathbf{d}}$ that satisfy

$$\hat{\mathbf{f}}_k, \hat{\mathbf{d}} = \arg \max_{\mathbf{f}_k, \mathbf{d}} P(\mathbf{f}_k, \mathbf{d}) P(\mathbf{g}|\mathbf{f}_k, \mathbf{d}), \quad (21)$$

where the prior distributions, $P(\mathbf{f}_k, \mathbf{d})$, used for uncompressed sequences have been defined in section 3.1 and the acquisition models for this problem, $P(\mathbf{g}|\mathbf{f}_k, \mathbf{d})$ are described in section 2.1.

Most of the reported works in the literature on SR from uncompressed sequences first estimate the HR motion vectors either by first interpolating the LR observations and then finding the HR motion vectors or by first finding the motion vectors in the LR domain and then interpolating them. So, classical motion estimation or image registration techniques can be applied to the process of finding the HR motion vectors. See, for instance, Brown [30], Szeliski [31] and Stiller and Konrad [32].

Interesting models for motion estimation have also been developed within the HR context. These works also perform segmentation within the HR image (see for instance, Irani and Peleg [7] and Eren *et al.* [13]).

Once the HR motion vectors, denoted by $\bar{\mathbf{d}}$, have been estimated all the methods proposed in the literature proceed to find $\bar{\mathbf{f}}_k$ satisfying

$$\bar{\mathbf{f}}_k = \arg \max_{\mathbf{f}_k} P(\mathbf{f}_k) P(\mathbf{g}|\mathbf{f}_k, \bar{\mathbf{d}}). \quad (22)$$

Several approaches have been used in the literature to find $\bar{\mathbf{f}}_k$, for example, gradient descent, conjugate gradient, preconditioning and POCS among others (see Borman and Stevenson[9] and Park *et al.* [26]).

Before leaving this section, it is important to note that some work has been developed in the uncompressed domain to carry out the estimation of the HR image and motion vectors simultaneously (see Tom and Katsaggelos [33],[34],[35] and Hardie *et al.* [36]).

4.2 Compressed sequences

For compressed sequences our goal becomes finding $\hat{\mathbf{f}}_k$ and $\hat{\mathbf{d}}$ that satisfy

$$\hat{\mathbf{f}}_k, \hat{\mathbf{d}} = \arg \max_{\mathbf{f}_k, \mathbf{d}} P(\mathbf{f}_k, \mathbf{d}) P(\mathbf{y}, \mathbf{v} | \mathbf{f}_k, \mathbf{d}) , \quad (23)$$

where the distributions of HR intensities and displacements used in the literature have been described in section 3.2 and the acquisition models have already been studied in section 2.2. The solution is found with a combination of gradient descent, non-linear projection, and full-search methods. Scenarios where \mathbf{d} is already known or separately estimated are a special case of the resulting procedure.

One way to find the solution of Eq. (23) is by using the cyclic coordinate descent procedure [37]. An estimate for the displacements is first found by assuming that the HR image is known, so that

$$\hat{\mathbf{d}}^{q+1} = \arg \max_{\mathbf{d}} P(\mathbf{d}) P(\mathbf{y}, \mathbf{v} | \hat{\mathbf{f}}_k^q, \mathbf{d}) , \quad (24)$$

where q is the iteration index for the joint estimate. (For the case where \mathbf{d} is known, Eq. (24) becomes $\hat{\mathbf{d}}^{q+1} = \bar{\mathbf{d}} \forall q$.) The intensity information is then estimated by assuming that the displacement estimates are exact, that is

$$\hat{\mathbf{f}}_k^{q+1} = \arg \max_{\mathbf{f}_k} P(\mathbf{f}_k) P(\mathbf{y}, \mathbf{v} | \mathbf{f}_k, \hat{\mathbf{d}}^{q+1}) . \quad (25)$$

The displacement information is re-estimated with the result from Eq. (25), and the process iterates until convergence. The remaining question is how to solve equations (24) and (25) for the distributions presented in the previous sections.

The non-informative prior in Eq. (16) is a common choice for $P(\mathbf{d})$. Block-matching algorithms are well suited for solving Eq. (24) for this particular case. The construction of $P(\mathbf{y}, \mathbf{v} | \mathbf{f}_k, \mathbf{d})$ controls the performance of the block-matching procedure (see Mateos *et al.* [24], Segall *et al.* [21, 22] and Chen and Schultz [23]).

When $P(\mathbf{d})$ is not uniform, differential methods become common methods for the estimation of the displacements. These methods are based on the optical flow equation and are explored in Segall *et al.* [21, 22]. An alternative differential approach is utilized by Park *et al.* [19, 38]. In these works, the motion between LR frames is estimated with the block based optical flow method suggested by Lucas and Kanade [39]. Displacements are estimated for the LR frames in this case.

Methods for estimating the HR intensities from Eq. (25) are largely determined by the acquisition model used. For example, consider the least complicated combination of the quantization constraint in Eq. (8) with the non-informative distributions for $P(\mathbf{f}_k)$ and $P(\mathbf{v}|\mathbf{f}_k, \mathbf{d}, \mathbf{y})$. Note that the solution to this problem is not unique, as the set-theoretic method only limits the magnitude of the quantization error in the system model. A frame that satisfies the constraint is therefore found with the projection onto convex sets (POCS) algorithm [40], where sources for the projection equations include [17] and [18].

A different approach must be followed when incorporating the spatial domain noise model in Eq. (9). If we still assume a non-informative distribution for $P(\mathbf{f}_k)$ and $P(\mathbf{v}|\mathbf{f}_k, \mathbf{d}, \mathbf{y})$, the estimate can be found with a gradient descent algorithm [18].

Figure 8 shows one example of the use of the techniques just described to estimate a HR frame.

5 Parameter estimation in HR

Since the early work by Tsai and Huang [8], researchers, primarily within the engineering community, have focused on formulating the HR problem as a reconstruction or a recognition one. However, as reported in [9], not much work has been devoted to the efficient calculation of the reconstruction or to the estimation of the associated parameters. In this section we will briefly review these two very interesting research areas.

Bose and Boo [41] use a block semi-circulant (BSC) matrix decomposition in order to calculate the *maximum a posteriori* (MAP) reconstruction, Chan *et al.* [42], [43] and [44], and Nguyen [45], [46] and [6] use preconditioning, wavelets, as well as BSC matrix decomposition. The efficient calculation of the MAP reconstruction is also addressed by Ng *et al.* ([47] and [48]) and Elad and Hel-Or [49].

To our knowledge only the works by Bose *et al.* [50], Nguyen [51], [46], [6], [52] and to some extent [44] and [53], [54] and [34] address the problem of parameter estimation. Furthermore, in those works the same parameter is assumed for all the LR images, although in the case of [50] the proposed method can be extended to different parameter for LR images (see [55]).

Recently, Molina *et al.* [56] have used the general framework for frequency domain multi-channel signal processing developed by Katsaggelos *et al.* in [57] and Banham *et al.* in [58] (a formulation that was also obtained later by Bose and Boo [41] for the HR problem) to tackle the parameter estimation in HR problems. With the use of BSC matrices the authors show that all the matrix calculations involved in the parameter maximum likelihood estimation can be performed in the Fourier domain. The proposed approach can be used to assign the same parameter to all LR images or make them image dependent. The role played by the number of available LR images in both the estimation procedure as well as the quality of the reconstruction

is examined in [59] and [60].

Figure 9(a) shows the upsampled version, 128×64 , of one 32×16 LR observed image. We ran the reconstruction algorithm in [60] using 1, 2, 4, 8 and 16 LR subpixel shifted images.

Before leaving this section we would like to mention that the above reviewed works address the problem of parameter estimation for the case of multiple undersampled, shifted, degraded frames with subpixel displacement errors and that very interesting areas to be explored are the cases of general compressed or uncompressed sequences.

6 New approaches towards HR

To conclude this chapter, we want to identify several research areas that we believe will benefit the field of super-resolution from video sequences.

A key area is the simultaneous estimate of multiple HR frames. These sequence estimates can incorporate additional spatio-temporal descriptions for the sequence and provide increased flexibility in modelling the scene. For example, the temporal evolution of the displacements can be modelled. Note that there is already some work in both compressed and uncompressed domains (see Hong *et al.* [61], [62], Choi *et al.* [63] and Dekeyser *et al.* [64]).

Alvarez *et al.* [65] have also addressed the problem of simultaneous reconstruction of compressed video sequences. To encapsulate the statement that the HR images are correlated and absent of blocking artifacts, the prior distribution

$$P(\mathbf{f}_1, \dots, \mathbf{f}_L) \propto \left(\prod_{l=1}^L \exp \left[-\frac{\lambda_1}{2} \|\mathbf{f}_l - \mathbf{C}(\mathbf{d}_{l,k})\mathbf{f}_k\|^2 \right] \right) \exp \left[-\frac{\lambda_2}{2} \|\mathbf{Q}_2 \mathbf{f}_k\|^2 - \frac{\lambda_3}{2} \|\mathbf{Q}_3 \mathbf{A} \mathbf{H} \mathbf{f}_k\|^2 \right] \quad (26)$$

is utilized. Here, \mathbf{Q}_2 represents a linear high-pass operation that penalizes super-resolution estimates that are not smooth, \mathbf{Q}_3 represents a linear high-pass operator that penalizes estimates with block boundaries, and λ_1, λ_2 and λ_3 control the weight given to each term. A common choice for \mathbf{Q}_2 is the discrete 2D Laplacian; a common choice for \mathbf{Q}_3 is the simple difference operation applied at the boundary locations. The HR motion vectors are previously estimated. Figure 10 shows just one image obtained by the method proposed in [65] together with the bilinear interpolation of the compressed LR observation.

Accurate estimates of the HR displacements are critical for the super-resolution problem. There is work to be done in designing methods for the blurred, sub-sampled, aliased, and in some cases blocky observations. Towards this goal, the use of probability distribution of optical flow as developed by Simoncelli *et al.* [66] and [67], as well as the coarse-to-fine estimation seem areas worth exploring. Note that we could incorporate also the coherence of the distribution of the optical flow when constructing the super resolution image (or sequence).

The use of band-pass directional filters on super resolution problems seems also an interesting area of research (see Nestares *et al.* [68] and Chamorro-Martínez [69]).

Let us consider the digits in a car plate shown in Fig. 11(a). The digits after blurring, subsampling and noise are almost undistinguishable, as shown in Fig. 11(b). Following Baker and Kanade [70] there are limits for the approach discussed in Molina *et al.* [56] (see Fig. 9). However, any classification/recognition method will benefit from resolution improvement on the observed images. Would it be possible to learn the prior image model from a training set of HR images that have undergone a blurring, subsampling and in some cases compression process to produce images *similar* to the observed LR ones?. This approach has been pursued by several authors and it will be described now in some depth.

Baker and Kanade [70] approach the super resolution problem in restricted domains (see the face recognition problem in [70]) by trying to learn the prior model from a set of training images. Let us assume that we have a set of HR training images, the size of each image being $2^N \times 2^N$. For all the training images we can form their feature pyramids. These pyramids of features are created by calculating at different resolutions, the Laplacian, L , the horizontal, H , and the vertical, V , first derivatives and the second horizontal, H^2 , and vertical, V^2 derivatives. So, for a given high training HR image, I we have at resolution $j, j = 0, \dots, N$, (the higher the value of j the smaller the resolution), the data

$$F_j(I) = (L_j(I), H_j(I), V_j(I), H_j^2(I), V_j^2(I)) . \quad (27)$$

We are now given a LR $2^K \times 2^K$ image Lo , with $K < N$, which is assumed to have been registered (motion compensated) with regard to the position of the HR images in the training set. Obviously we can also calculate for $j = 0, \dots, K$

$$F_j(Lo) = (L_j(Lo), H_j(Lo), V_j(Lo), H_j^2(Lo), V_j^2(Lo)) . \quad (28)$$

Let us assume that we have now a pixel (x, y) on the LR observed image we want to improve its resolution. We can now find the HR image in the training set whose LR image of size $2^K \times 2^K$ has the pixel on its position (m_x, m_y) with the most similar pyramid structure to the one built for pixel (x, y) . Obviously there are different ways to define the concept of more similar pyramid of features (see [70]).

Now we note that the LR pixel (m_x, m_y) comes from $2^{N-K} \times 2^{N-K}$ HR pixels. On each of these HR pixels we calculate the horizontal and vertical derivatives and as HR image prior we force the similarity between these derivatives and the one calculated in the corresponding position of the HR image we want to find.

As acquisition model the authors use the one given by Eq. (5). So, several LR observations can be used to calculate the HR images, although, so far, only one of them is used to build the prior. Then for this combination of learned HR image prior and acquisition model, the authors estimate the HR image

Instead of using the training set to define the derivative based prior model, Capel and Zisserman [71] use it to create a *Principal Component* basis and then express the HR image

we are looking for as a linear combination of the elements of a sub-space of this basis. So, our problem has now become the calculation of the coefficient of the vectors in the sub-space.

Priors can then be defined either on those coefficients or in terms of the real HR underlying image (see [71] for details). Similar ideas have later been developed by Gunturk *et al.* [72].

It is also important to note that again for this kind of priors, several LR images, registered or motion compensated, can be used in the degradation model given by Eq. (5).

The third alternative approach to super resolution from video sequences we will consider here is the one developed by Bishop *et al.* [73]. This approach builds on the one proposed by Freeman *et al.* [74] for the case of static images.

The approach proposed in [74] involves the assembly of a large database of patch pairs. Each pair is made of a HR patch usually of size 5×5 and a corresponding patch of size 7×7 . This corresponding patch is built as follows: the HR image, I , is first blurred and downsampled (following the process to obtain a LR image from a HR one) then this LR image is upsampled to have the same size as the HR one, let us call it I_{ud} . Then to each 5×5 patch in I , we associate the 7×7 patch in I_{ud} centered in the same pixel position as the 5×5 HR patch. To both, the HR and the upsampled LR images their low frequencies are removed and so it could be assumed that we have the high frequencies in the HR image and the mid frequencies in the upsampled LR one.

From those 5×5 and 7×7 patches we can build a dictionary, the patches are also contrast normalized to avoid having a huge dictionary. We now want to create the HR image corresponding to a LR one. The LR one is upsampled to the size of the HR image. Given a location of a 5×5 patch we want to create, we use its corresponding 7×7 upsampled LR one to search the dictionary and find its corresponding HR one. In order to enforce consistency of the 5×5 patches, they are overlapped, see Freeman *et al.* [74]. Two alternative methods for finding the best HR image estimate are proposed in [75].

Bishop *et al.* [73] modify the cost function to be minimized when finding the most similar 5×5 patch by including terms that penalize flickering and that enforce temporal coherence.

Note that from the above discussion on the work of Freeman *et al.* [74], work on vector quantization can also be extended and applied to the super resolution problem (see Nakagaki and Katsaggelos [76]).

Acknowledgements

The work of L. D. Alvarez and R. Molina was partially supported by the “Comisión Nacional de Ciencia y Tecnología” under contract TIC2000-1275, and the work of A. K. Katsaggelos by the Motorola Center for Communications, Northwestern University.

References

- [1] S. Chaudhuri (Ed.), *Super-resolution from compressed video*, Kluwer Academic Publishers, 2001.
- [2] M. G. Kang and S. Chaudhuri (Eds.), “Super-resolution image reconstruction,” *IEEE Signal Processing Magazine*, vol. 20, no. 3, 2003.
- [3] R.R. Schultz and R.L. Stevenson, “Extraction of high resolution frames from video sequences,” *IEEE Transactions on Image Processing*, vol. 5, pp. 996–1011, 1996.
- [4] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, “High resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system,” *Optical Engineering*, vol. 73, pp. 247–260, 1998.
- [5] M. Elad and A. Feuer, “Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images,” *IEEE Transactions on Image Processing*, vol. 6, pp. 1646–1658, 1997.
- [6] N. Nguyen, P. Milanfar, and G. Golub, “A computationally efficient superresolution image reconstruction algorithm,” *IEEE Transactions on Image Processing*, vol. 10, pp. 573–583, 2001.
- [7] M. Irani and S. Peleg, “Motion analysis for image enhancement: Resolution, occlusion, and transparency,” *JVCIP*, vol. 4, pp. 324–335, 1993.
- [8] R. Y. Tsai and T. S. Huang, “Multiframe image restoration and registration,” *Advances in Computer Vision and Image Processing*, vol. 1, pp. 317–339, 1984.
- [9] S. Borman and R. Stevenson, “Spatial resolution enhancement of low-resolution image sequences a comprehensive review with directions for future research,” Tech. Rep., Laboratory for Image and Signal Analysis (LISA), University of Notre Dame, Notre Dame, IN 46556, USA, July 1998.
- [10] H. Stark and P. Oskoui, “High resolution image recovery from image-plane arrays, using convex projections,” *J. Opt. Soc. Am. A*, vol. 6, pp. 1715–1726, 1989.
- [11] A.M. Tekalp, M.K. Ozkan, and M.I. Sezan, “High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1992, pp. 169–172.

- [12] 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*, vol. 6, pp. 1064–1076, 1997.
- [13] P.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*, pp. 1446–1451, 1997.
- [14] A.M. Tekalp, *Digital Video Processing*, Prentice Hall, Signal Processing Series, 1995.
- [15] C. A. Segall, R. Molina, and A. K. Katsaggelos, “High-resolution images from low-resolution compressed video,” *IEEE Signal Processing Magazine*, vol. 20, pp. 37–48, 2003.
- [16] Y. Altunbasak, A.J. Patti, and R.M. Mersereau, “Super-resolution still and video reconstruction from MPEG-coded video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, pp. 217–226, 2002.
- [17] A.J. Patti and Y. Altunbasak, “Super-resolution image estimation for transform coded video with application to MPEG,” in *Proceedings of the IEEE International Conference on Image Processing*, 1999, vol. 3, pp. 179–183.
- [18] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, *Super-resolution from compressed video*, chapter 9 in *Super-Resolution Imaging*, S. Chaudhuri, Ed., pp. 211–242, Kluwer Academic Publishers, 2001.
- [19] S.C. Park, M.G. Kang, C.A. Segall, and A.K. Katsaggelos, “High-resolution image reconstruction of low-resolution DCT-based compressed images,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 2, pp. 1665–1668.
- [20] S.C. Park, M.G. Kang, C.A. Segall, and A.K. Katsaggelos, “Spatially adaptive high-resolution image reconstruction of low-resolution DCT-based compressed images,” *IEEE Transactions on Image Processing*, to appear, 2003.
- [21] C.A. Segall, R. Molina, A.K. Katsaggelos, and J. Mateos, “Reconstruction of high-resolution image frames from a sequence of low-resolution and compressed observations,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 2, pp. 1701–1704.
- [22] C. A. Segall, R. Molina, A. K. Katsaggelos, and J. Mateos, “Bayesian resolution enhancement of compressed video,” *IEEE Transactions on Image Processing*, to appear, 2004.

- [23] D. Chen and R.R. Schultz, “Extraction of high-resolution video stills from MPEG image sequences,” in *Proceedings of the IEEE International Conference on Image Processing*, 1998, vol. 2, pp. 465–469.
- [24] J. Mateos, A.K. Katsaggelos, and R. Molina, “Simultaneous motion estimation and resolution enhancement of compressed low resolution video,” in *IEEE International Conference on Image Processing*, 2000, vol. 2, pp. 653–656.
- [25] B.C. Tom and A.K. Katsaggelos, “An iterative algorithm for improving the resolution of video sequences,” in *Proceedings of SPIE Conference on Visual Communications and Image Processing*, 1996, pp. 1430–1438.
- [26] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical overview,” *IEEE Signal Processing Magazine*, vol. 20, pp. 21–36, 2003.
- [27] D. P. Capel and A. Zisserman, “Super-resolution enhancement of text image sequence,” in *International Conference on Pattern Recognition*, p. 600.
- [28] H. Kim, J-H Jang, and K-S Hong, “Edge-enhancing super-resolution using anisotropic diffusion,” in *Proceedings IEEE Conference on Image Processing*, 2001, vol. 3, pp. 130–133.
- [29] D. Rajan and S. Chaudhuri, “Generation of super-resolution images from blurred observations using an MRF model,” *J. Math. Imaging Vision*, vol. 16, pp. 5–15, 2002.
- [30] L.G. Brown, “A survey of image registration techniques,” *ACM Computing surveys*, vol. 24, 1992.
- [31] R. Szeliski, “Spline-based image registration,” *International Journal of Computer Vision*, vol. 22, pp. 199–218, 1997.
- [32] C. Stiller and J. Konrad, “Estimating motion in image sequences,” *IEEE Signal Processing Magazine*, vol. 16, pp. 70–91, 1999.
- [33] B. C. Tom and A. K. Katsaggelos, “Reconstruction of a high resolution image from multiple-degraded and misregistered low-resolution images,” in *Proceedings of SPIE Conference on Visual Communications and Image Processing*, 1994, vol. 2308, pp. 971–981.
- [34] B. C. Tom and A. K. Katsaggelos, “Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images,” in *Proceedings of the IEEE International Conference on Image Processing*, 1995, vol. 2, pp. 539–542.

- [35] B. C. Tom and A. K. Katsaggelos, “Resolution enhancement of monochrome and color video using motion compensation,” *IEEE Transactions on Image Processing*, vol. 10, pp. 278–287, 2001.
- [36] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, “Joint map registration and high-resolution image estimation using a sequence of undersampled images,” *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, 1997.
- [37] D.G. Luenberger, *Linear and Nonlinear Programming*, Reading, MA: Addison-Wesley Publishing Company, Inc., 1984.
- [38] S.C. Park, M.G. Kang, C.A. Segall, and A.K. Katsaggelos, “Spatially adaptive high-resolution image reconstruction of low-resolution DCT-based compressed images,” in *Proceedings of the IEEE International Conference on Image Processing*, 2002, vol. 2, pp. 861–864.
- [39] B.D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proceedings of Imaging Understanding Workshop*, 1981, pp. 121–130.
- [40] D. C. Youla and H. Webb, “Image restoration by the method of convex projections: Part 1-theory,” *IEEE Transactions on Medical Imaging*, vol. MI-1, no. 2, pp. 81–94, 1982.
- [41] N. K. Bose and K. J. Boo, “High-resolution image reconstruction with multisensors,” *International Journal of Imaging Systems and Technology*, vol. 9, pp. 141–163, 1998.
- [42] R. H. Chan, T. F. Chan, M. K. Ng, W. C. Tang, and C. K. Wong, “Preconditioned iterative methods for high-resolution image reconstruction with multisensors,” in *Proceedings to the SPIE Symposium on Advanced Signal Processing: Algorithms, Architectures, and Implementations*, 1998, vol. 3461, pp. 348–357.
- [43] R. H. Chan, T. F. Chan, L. Shen, and S. Zuowei, “A wavelet method for high-resolution image reconstruction with displacement errors,” in *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2001, pp. 24–27.
- [44] R. H. Chan, T. F. Chan, L. X. Shen, and Z. W. Shen, “Wavelet algorithms for high-resolution image reconstruction,” Tech. Rep., Department of Mathematics, Chinese University of Hong Kong, 2001.
- [45] N. Nguyen and P. Milanfar, “A wavelet-based interpolation-restoration method for super-resolution,” *Circuits, Systems, and Signal Processing*, vol. 19, pp. 321–338, 2000.
- [46] N. Nguyen, *Numerical Algorithms for superresolution*, Ph.D. thesis, Stanford University, 2001.

- [47] M. K. Ng, R. H. Chan, and T. F. Chan, “Cosine transform preconditioners for high resolution image reconstruction,” *Linear Algebra and its Applications*, vol. 316, pp. 89–104, 2000.
- [48] M. K. Ng and A. M. Yip, “A fast MAP algorithm for high-resolution image reconstruction with multisensors,” *Multidimensional Systems and Signal Processing*, vol. 12, pp. 143–164, 2001.
- [49] M. Elad and Y. Hel-Or, “A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur,” *IEEE Transactions on Image Processing*, vol. 10, pp. 1187–93, 2001.
- [50] N. K. Bose, S. Lertrattanapanich, and J. Koo, “Advances in superresolution using L-curve,” *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 433–436, 2001.
- [51] N. Nguyen, P. Milanfar, and G. Golub, “Blind superresolution with generalized cross-validation using Gauss-type quadrature rules,” in *33rd Asilomar Conference on Signals, Systems, and Computers*, 1999, vol. 2, pp. 1257–1261.
- [52] N. Nguyen, P. Milanfar, and G.H. Golub, “Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement,” *IEEE Transactions on Image Processing*, vol. 10, pp. 1299–1308, 2001.
- [53] B. C. Tom, N. P. Galatsanos, and A. K. Katsaggelos, “Reconstruction of a high resolution image from multiple low resolution images,” in *Super-Resolution Imaging*, S. Chaudhuri, Ed., chapter 4, pp. 73–105. Kluwer Academic Publishers, 2001.
- [54] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos, “Reconstruction of a high-resolution image from registration and restoration of low-resolution images,” in *Proceedings of the IEEE International Conference on Image Processing*, 1994, vol. 3, pp. 553–557.
- [55] M. Belge, M. E. Kilmer, and E. L. Miller, “Simultaneous multiple regularization parameter selection by means of the L-hypersurface with applications to linear inverse problems posed in the wavelet domain,” in *Proc. SPIE’98: Bayesian Inference Inverse Problems*, 1998.
- [56] R. Molina, M. Vega, J. Abad, and A.K. Katsaggelos, “Parameter estimation in bayesian high-resolution image reconstruction with multisensors,” *IEEE Transactions on Image Processing*, to appear, 2003.
- [57] A. K. Katsaggelos, K. T. Lay, and N. P. Galatsanos, “A general framework for frequency domain multi-channel signal processing,” *IEEE Transactions on Image Processing*, vol. 2, pp. 417–420, 1993.

- [58] M. R. Banham, N. P. Galatsanos, H. L. Gonzalez, and A. K. Katsaggelos, “Multichannel restoration of single channel images using a wavelet-based subband decomposition,” *IEEE Transactions on Image Processing*, vol. 3, pp. 821–833, 1994.
- [59] M. Vega, J. Mateos, R. Molina, and A.K. Katsaggelos, “Bayesian parameter estimation in image reconstruction from subsampled blurred observations,” in *Proceedings of the IEEE International Conference on Image Processing*, to appear, 2003.
- [60] F.J. Cortijo, S. Villena, R. Molina, and A.K. Katsaggelos, “Bayesian superresolution of text image sequences from low-resolution observations,” in *IEEE Seventh International Symposium on Signal Processing and Its Applications (ISSPA 2003)*, 2003, vol. I, pp. 421–424.
- [61] M-C. Hong, M. G. Kang, and A. K. Katsaggelos, “An iterative weighted regularized algorithm for improving the resolution of video sequences,” in *Proceedings IEEE International Conference on Image Processing*, 1997, vol. II, pp. 474–477.
- [62] M.-C. Hong, M. G. Kang, and A. K. Katsaggelos, “Regularized multichannel restoration approach for globally optimal high resolution video,” in *Proceedings SPIE Conference on Visual Communications and Image Processing*, 1997, pp. 1306–1316.
- [63] M. C. Choi, Y. Yang, and N. P. Galatsanos, “Multichannel regularized recovery of compressed video sequences,” *IEEE Trans. on Circuits and Systems II*, vol. 48, pp. 376–387, 2001.
- [64] F. Dekeyser, P. Bouthemy, and P. Pérez, “A new algorithm for super-resolution from image sequences,” in *9th International Conference on Computer Analysis of Images and Patterns, CAIP 2001, Springer Verlag, Lecture Notes in Computer Science 2124*, 2001, pp. 473–481.
- [65] L.D. Alvarez, R. Molina, and A.K. Katsaggelos, “Multi-channel reconstruction of video sequences from low-resolution and compressed observations,” in *8th Iberoamerican Congress on Pattern Recognition (CIARP’2003)*, to appear, 2003.
- [66] E.P. Simoncelli, E.H. Adelson, and D.J. Heeger, “Probability distributions of optical flow,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp. 310–315.
- [67] E.P. Simoncelli, “Bayesian multi-scale differential optical flow,” in *Handbook of Computer Vision and Applications*. 1999, Academic Press.

- [68] O. Nestares and R. Navarro, “Probabilistic estimation of optical flow in multiple band-pass directional channels,” *Image and Vision Computing*, vol. 19, pp. 339–351, 2001.
- [69] J. Chamorro-Martínez, *Desarrollo de modelos computacionales de representación de secuencias de imágenes y su aplicación a la estimación de movimiento (in Spanish)*, Ph.D. thesis, University of Granada, 2001.
- [70] S. Baker and T. Kanade, “Limits on super-resolution and how to break them,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1167–1183, 2002.
- [71] D.P. Capel and A. Zisserman, “Super-resolution from multiple views using learnt image models,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 2, pp. 627–634.
- [72] B.K. Gunturk, A.U. Batur, Y. Altunbasak, M.H. Hayes, and R.M. Mersereau, “Eigenface-domain super-resolution for face recognition,” *IEEE Trans. on Image Processing*, vol. 12, pp. 597–606, 2003.
- [73] C. M. Bishop, A. Blake, and B. Marthi, “Super-resolution enhancement of video,” in *C. M. Bishop and B. Frey (Eds.), Proceedings Artificial Intelligence and Statistics*, 2003.
- [74] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example based super-resolution,” *IEEE Computer Graphics and Applications*, vol. 22, pp. 56–65, 2002.
- [75] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning low-level vision,” *International Journal of Computer Vision*, vol. 40, pp. 25–47, 2000.
- [76] R. Nakagaki and A. K. Katsaggelos, “A VQ-based blind image restoration algorithm,” *IEEE Transactions on Image Processing*, September 2003.
- [77] J. Mateos, A.K. Katsaggelos, and R. Molina, “Resolution enhancement of compressed low resolution video,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2000, vol. 4, pp. 1919–1922.

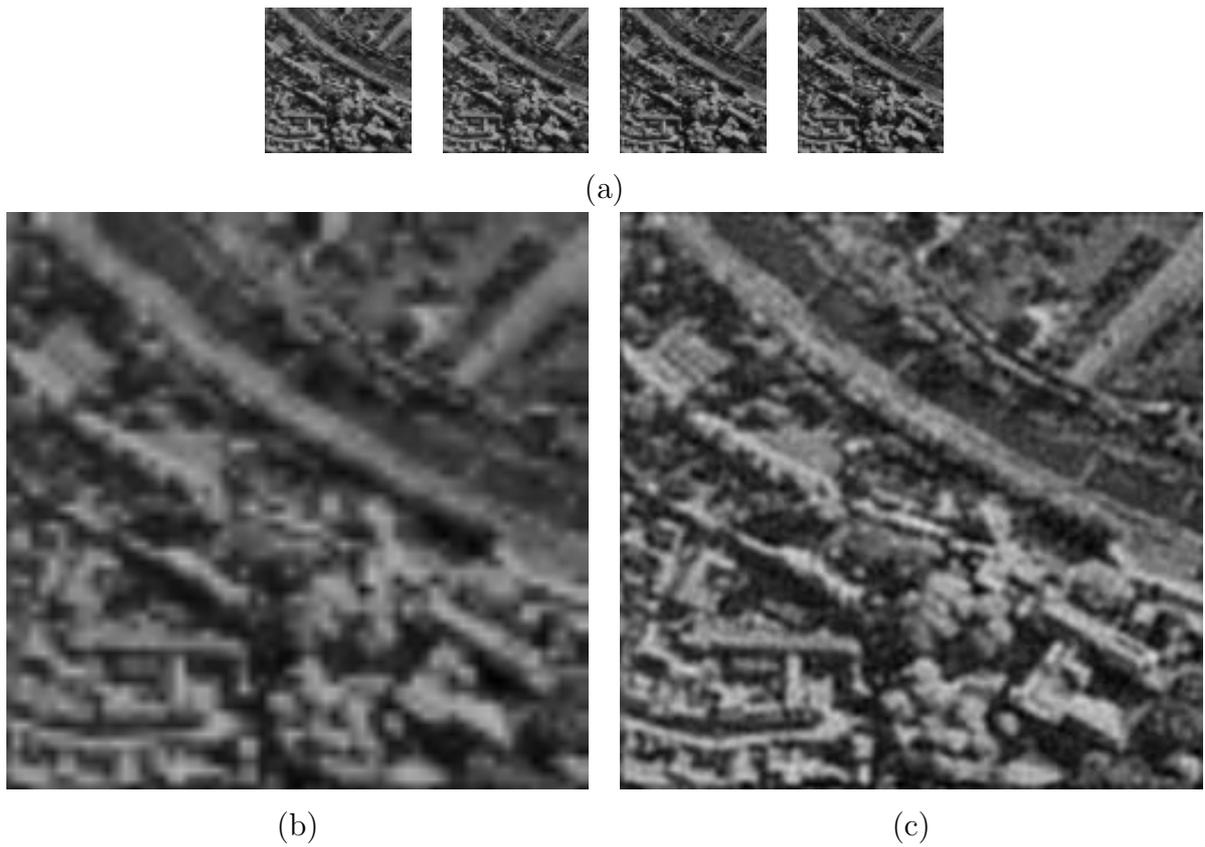


Figure 1: (a) Observed LR images with global sub-pixel displacement among them; (b) Bilinearly interpolated image using the upper left LR image; (c) HR image obtained by combining the information in the LR observations using the algorithm in [59].

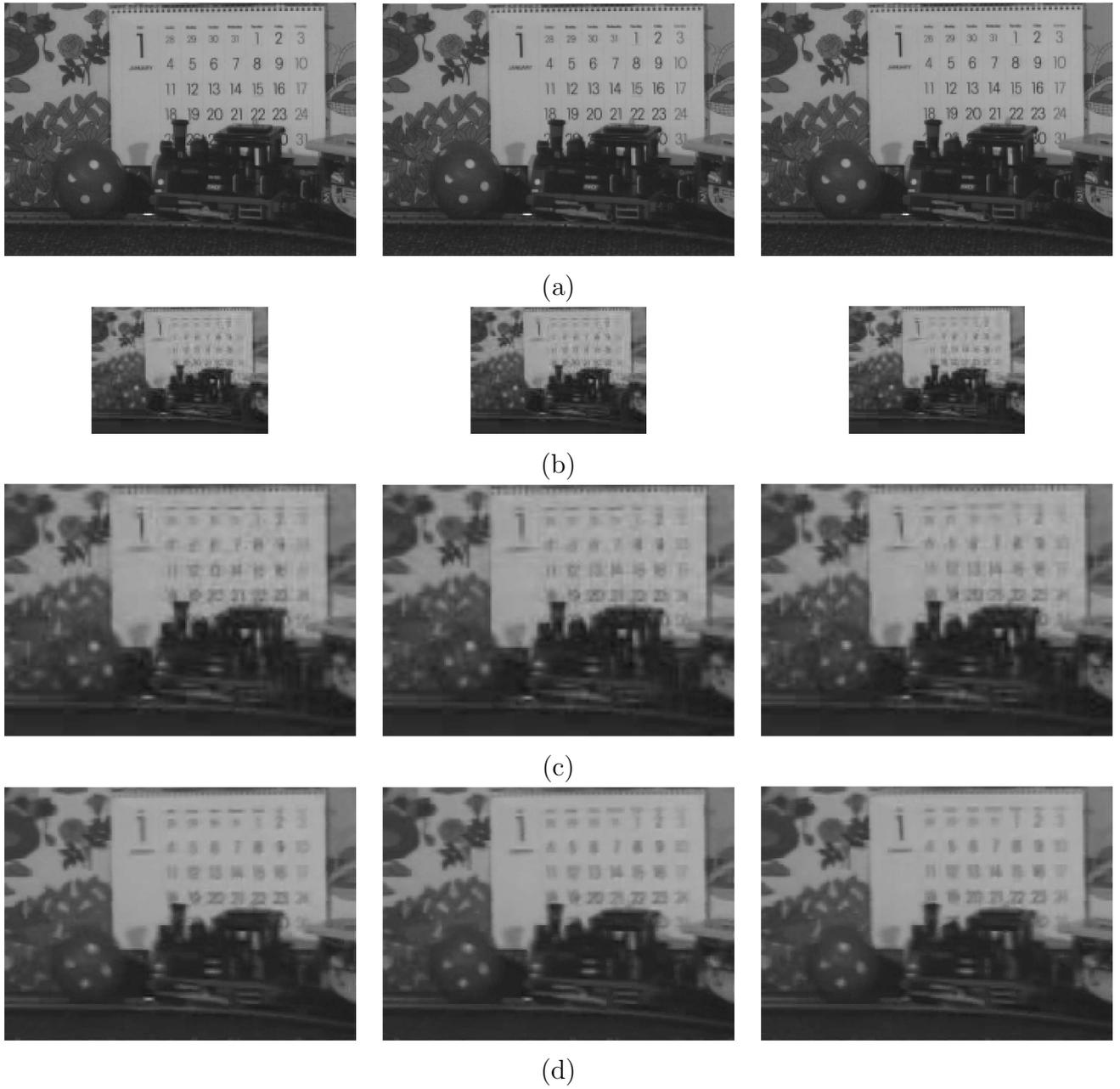


Figure 2: (a) Three HR consecutive image frames from a video sequence; (b) Corresponding LR observations after blurring, subsampling and compression using *MPEG-4* at $128Kbps$.; (c) Bilinearly interpolated frames; (d) HR frames using the algorithm in [77].

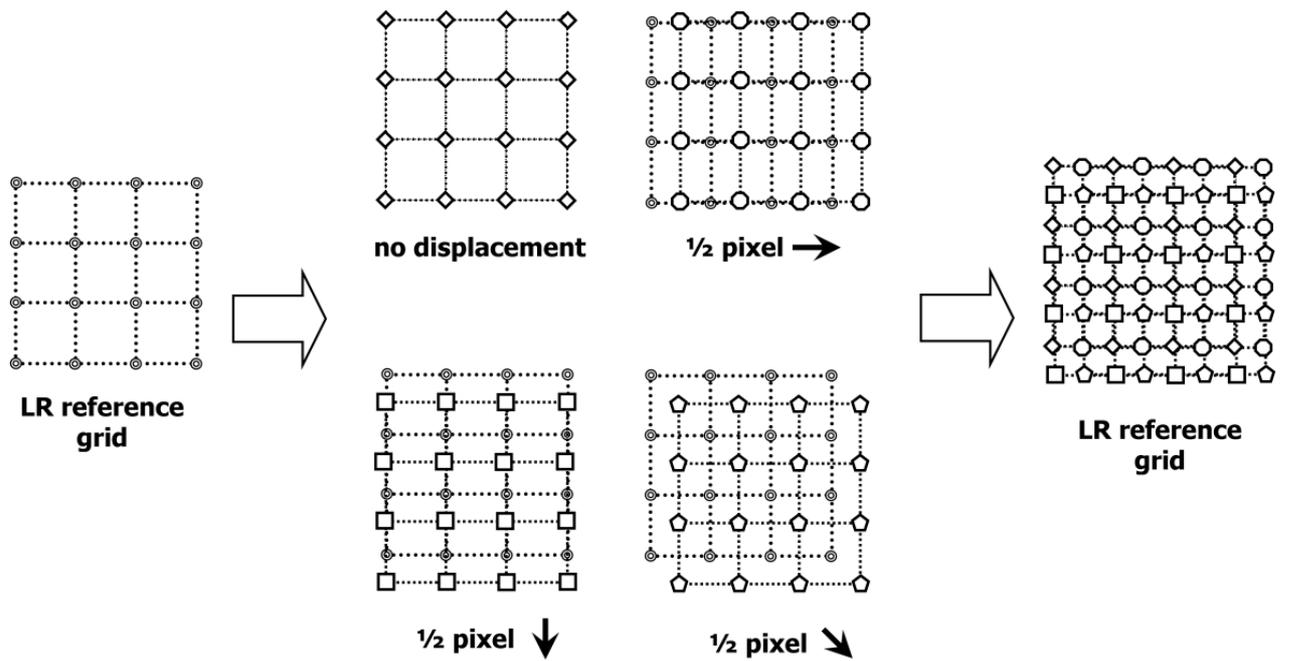


Figure 3: After taking several images of the same scene with subpixel displacement among them (several images from the same camera or one image from several cameras) or after recording an scene with moving objects in it, we can combine the observations to improve resolution.

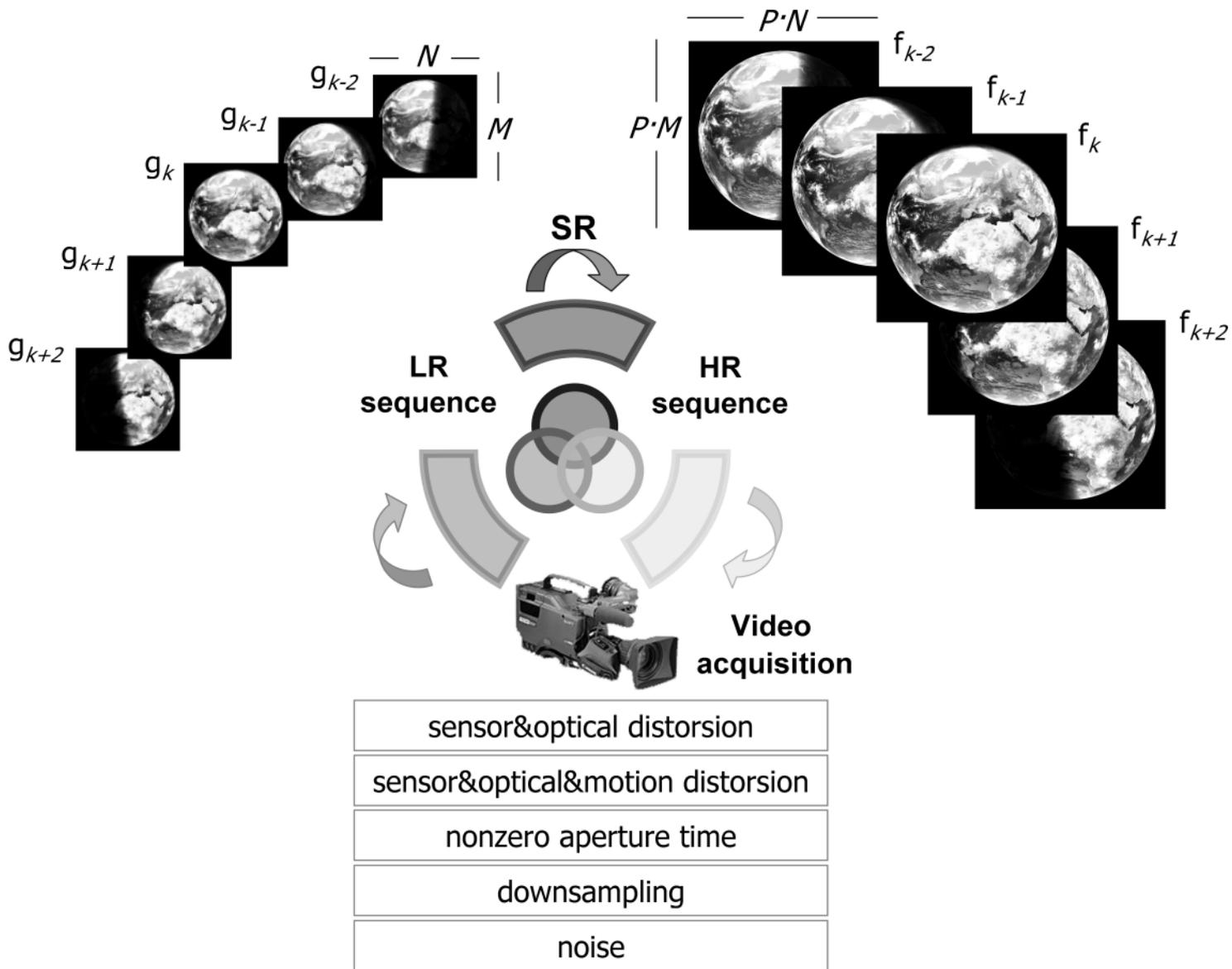


Figure 4: Acquisition model for the uncompressed case.

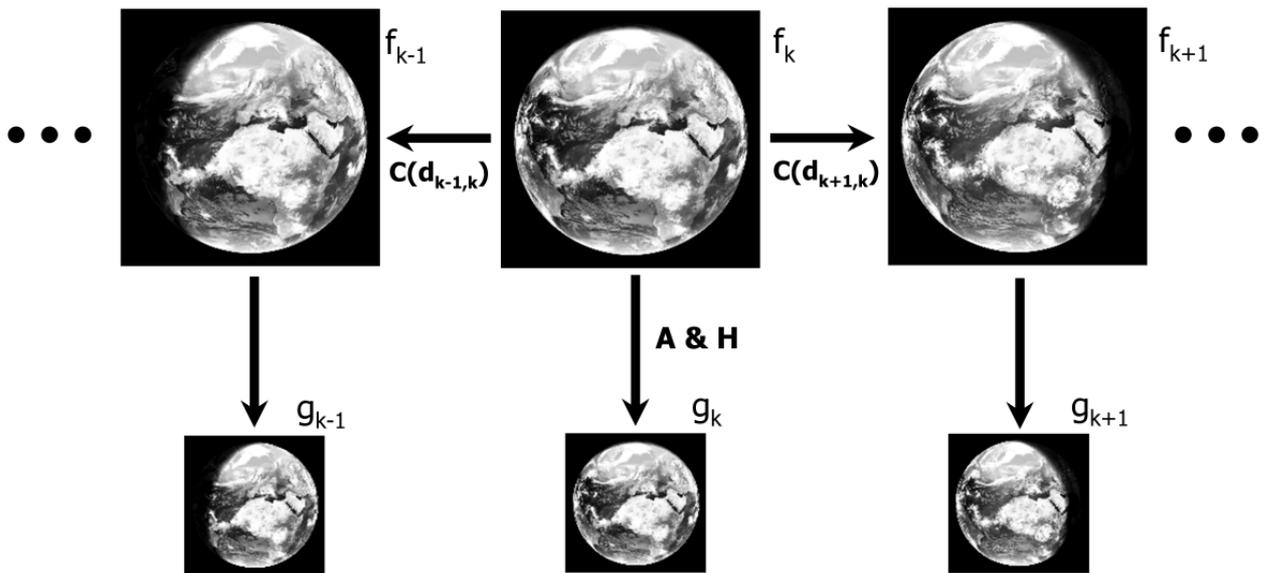


Figure 5: Relationship between LR and HR images. Note the use of motion compensation together with blurring and downsampling.

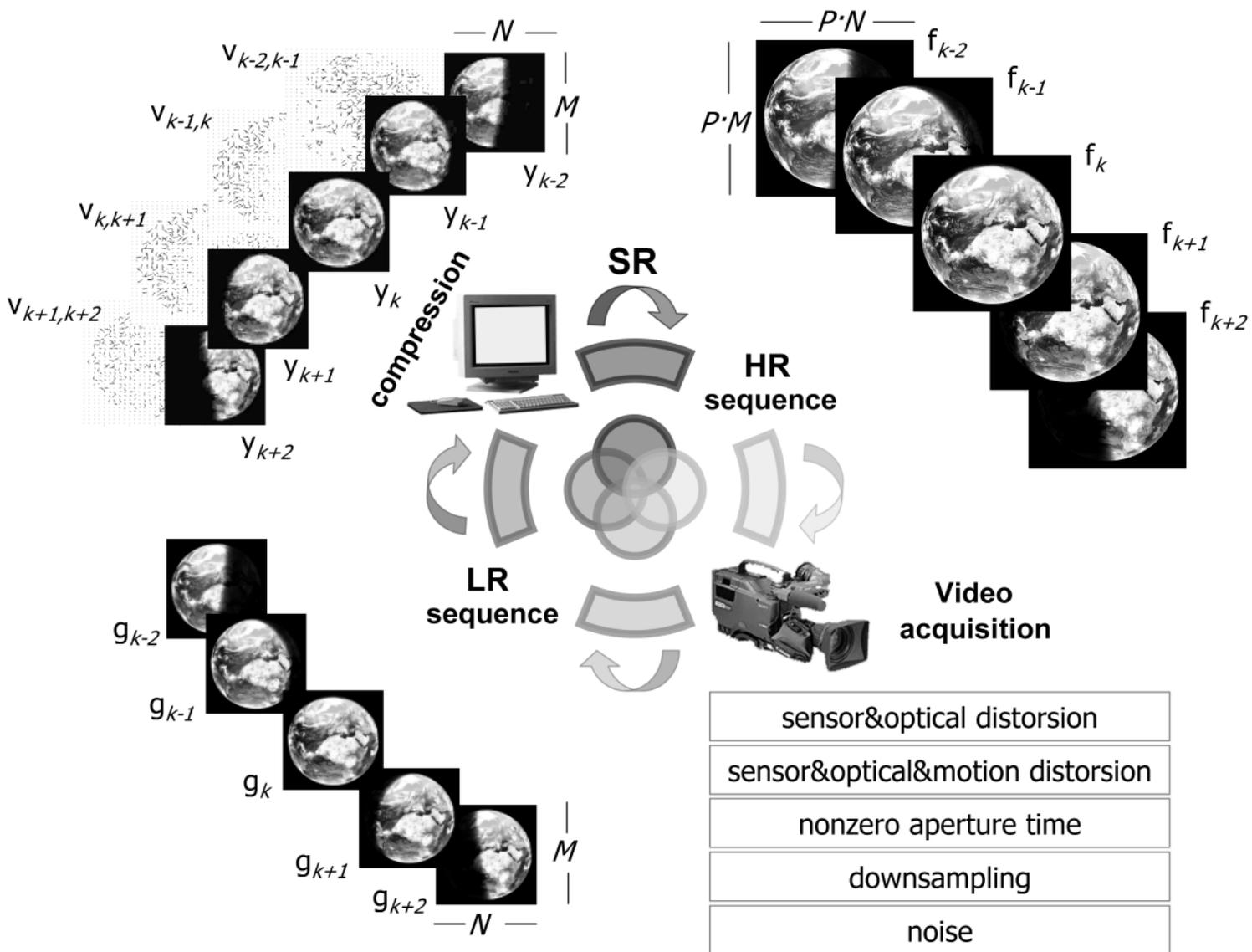


Figure 6: Acquisition model for the compressed problem.

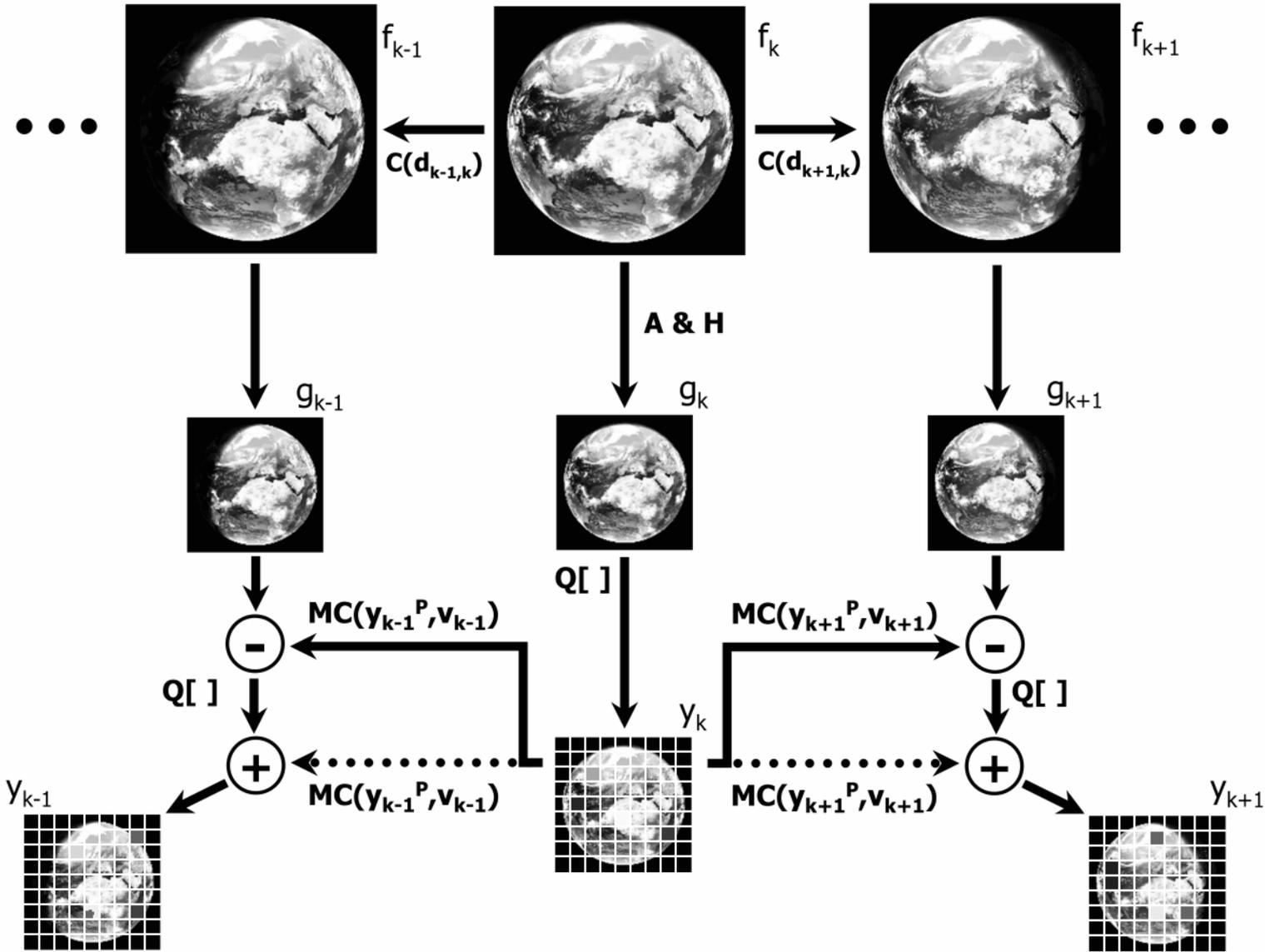


Figure 7: Relationship between compressed LR and HR images. Note the use of motion compensation together with blurring, downsampling and quantization.

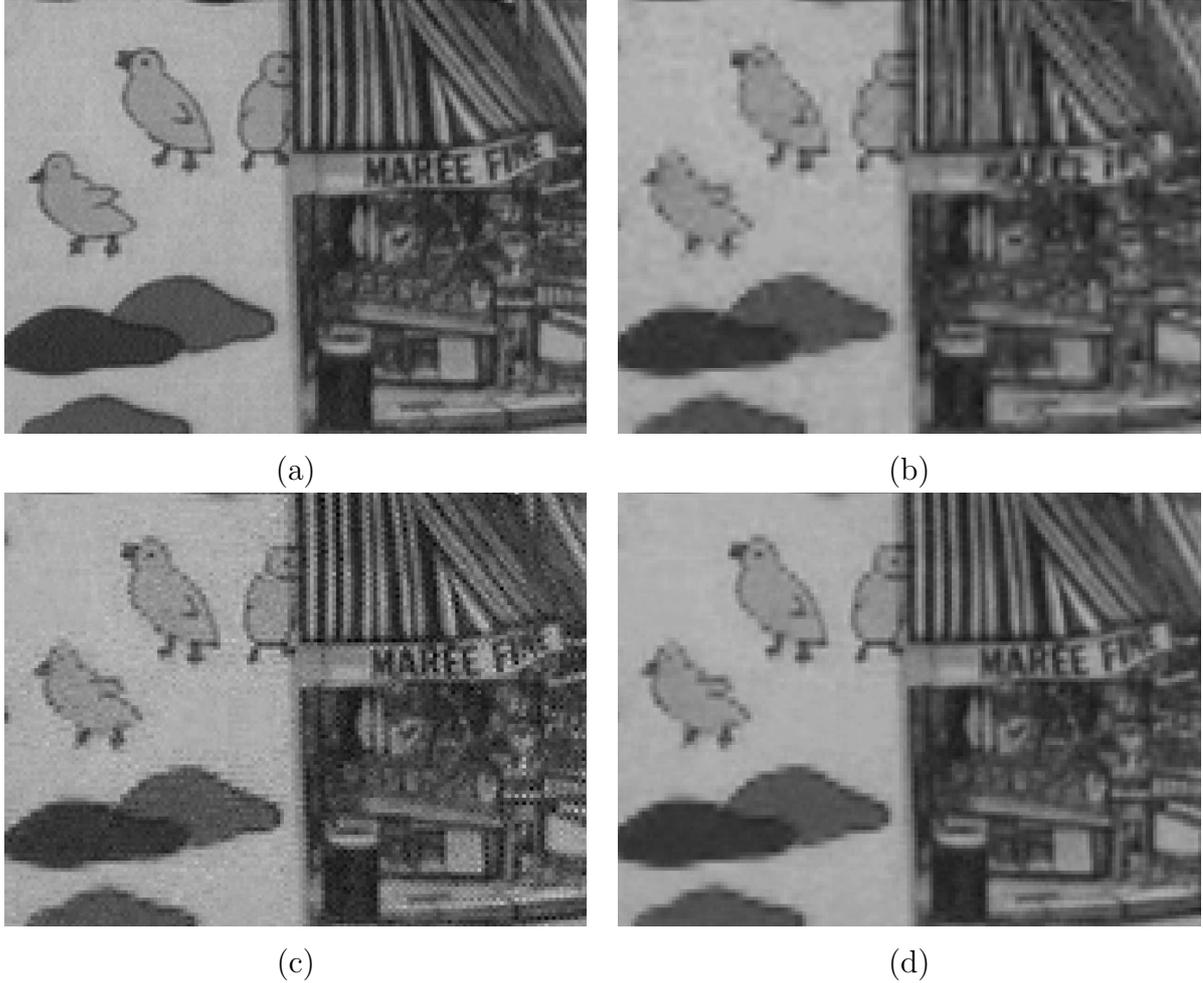


Figure 8: From a video sequence: (a) original image, (b) decoded result after bilinear interpolation. The original image is down-sampled by a factor of two in both the horizontal and vertical directions and then compressed with an MPEG-4 encoder operating at 1 Mb/s, (c) super resolution image employing only the quantization constraint in Eq. (8) and (d) super resolution image employing the normal approximation for $P(\mathbf{y}|\mathbf{f}_k, \mathbf{d})$ in Eq. (9), the distribution of the LR motion vectors given by Eq. (12) and the HR image prior in Eq. (18).

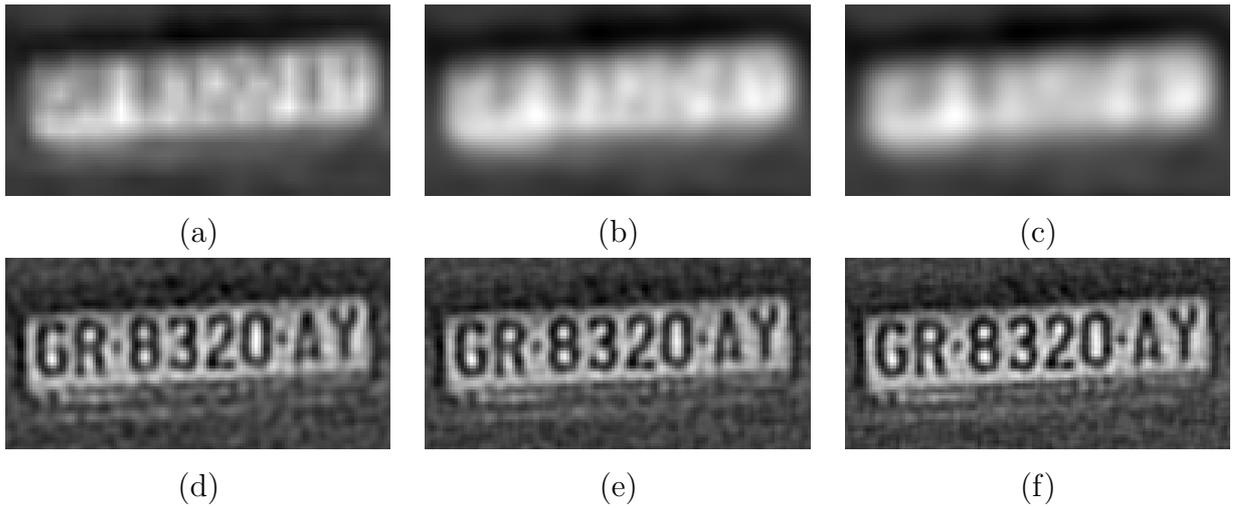


Figure 9: Spanish car license plate example: (a) upsampled observed LR image (b)-(f) reconstruction using 1, 2, 4, 8 and 16 LR images.

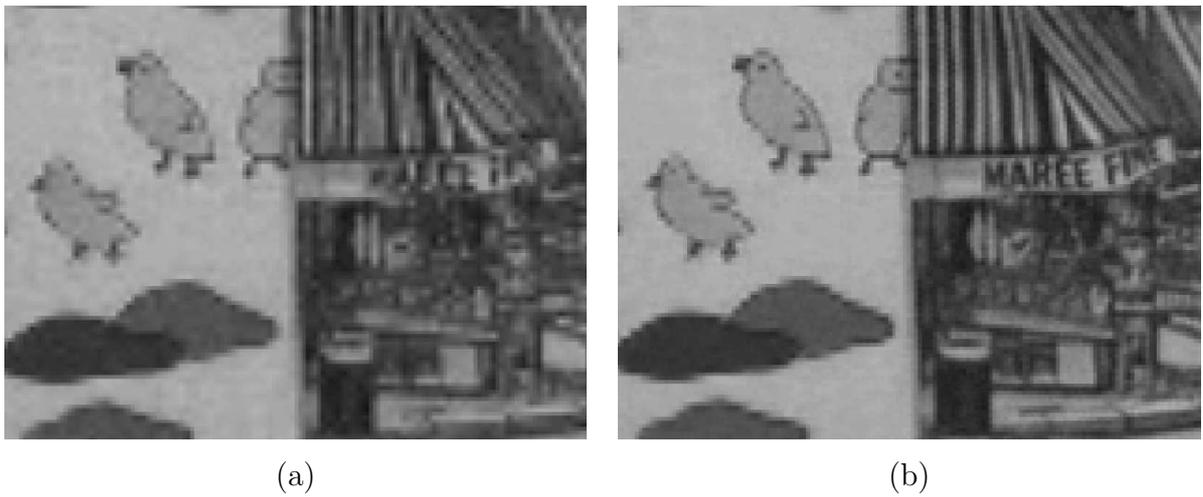


Figure 10: (a) bilinear interpolation of a LR compressed image, (b) image obtained when the whole HR sequence is processed simultaneously.

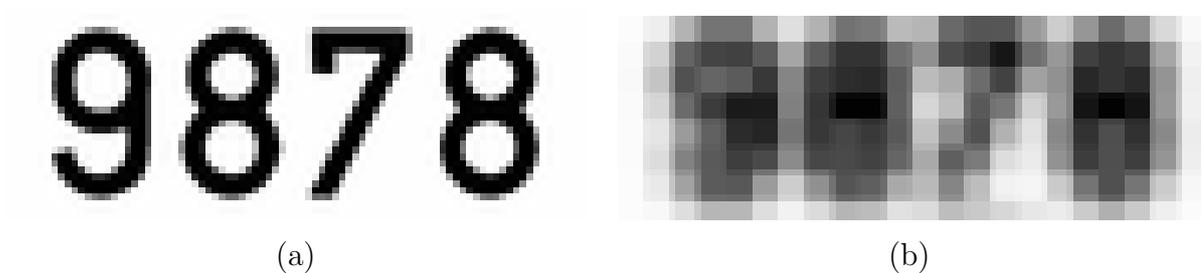


Figure 11: (a) HR image, (b) its corresponding LR observation.