J. Vis. Commun. Image R. 23 (2012) 94–99

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

Total least square kernel regression

Hiêp Luong *, Bart Goossens, Aleksandra Pižurica, Wilfried Philips

Department of Telecommunications and Information Systems, IBBT, Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

A R T I C L E I N F O

Article history:
Received 17 February 2011
Accepted 6 September 2011
Available online 10 September 2011

Keywords:
Kernel regression
Total least square
Super-resolution
Orthogonal distance regression
Image fusion
Gauss-Newton
Non-uniform resampling
Registration error

A B S T R A C T

In this paper, we study the problem of robust image fusion in the context of multi-frame super-resolution. Given multiple aligned noisy low-resolution images, image fusion produces a new image on a high-resolution grid. Recently, kernel regression is presented as a powerful image fusion technique. However, in the presence of registration errors, the performance of kernel regression is quite poor. Therefore, we present a new kernel regression method that takes these registration errors into account. Instead of the ordinary least square metric, we employ the total least square metric, which allows for spatial perturbations of the image samples. We show in our experiments that our method is more robust to noise and/or registration errors compared to the traditional kernel regression algorithm.

1. Introduction

In the last decades, the use of multiple images in the restoration process has gained a lot of popularity among various researchers. One of the image restoration problems being studied is the creation of a clean high-resolution (HR) image from multiple noisy low-resolution (LR) images, i.e., multi-frame super-resolution (SR) restoration problem. Multi-frame SR restoration becomes most successful if there is a non-integer displacement between the frequency-aliased LR images [1]. A typical multi-frame SR framework consists of image registration, image fusion and image deblurring [2]. After (proper) alignment, the LR images provide samples at non-uniform or irregular positions on the HR grid. Image fusion then converts these LR samples into samples that are placed on a regular Cartesian HR grid. Finally, the HR image is deconvolved to obtain a clean and sharp image. In this paper, we focus on the image fusion process in the presence of registration errors and image noise.

From the interpolation point of view, there are two main strategies to process non-uniformly distributed samples: we can use the same interpolation kernel everywhere and fit these kernels to the measurement data in a way that the reconstructed signal fits the measurements or we can define tailored basis functions (such as radial basis functions) that are better suited to the underlying non-uniform structure. Note that in higher dimensions the B-spline formalism is no longer applicable unless the grid is separable [3]. A more general approach is to use radial basis functions, which are closely related to splines as well, such as the membrane and thin-plate splines [4]. In [5], each triangle patch in the spatial Delaunay tessellation is approximated by a bivariate polynomial in order to reconstruct the HR image. In [6], the reconstruction of non-uniformly sampled signals is based on wavelets in a multiresolution setting.

The main drawback of these interpolation techniques is the sensitivity to image noise and in addition, a conflict could arise if there are multiple noisy samples at the same position or very close to each other.

Iterative simulate-and-correct approaches to non-uniform interpolation are intuitively very simple. A well-known method is the Papoulis–Gerchberg algorithm [7,8] in which alternately, the known set of irregularly placed samples are projected onto the HR grid and an ideal low-pass filter is applied on the HR image to enforce band-limitation. In the more general POCS algorithms, the ideal low-pass filter is substituted by other convex set operations (e.g. Gaussian blur). Iterative back-projection methods update the current estimated HR image by projecting the residual errors between the observed and the simulated LR images [9]. The simulated LR images are simply obtained by resampling the current HR image.

Alternatively, a very fast and memory efficient way to aggregate multiple LR images into one HR image is the *shift-and-add* method. This method assigns each pixel of the LR image to the nearest HR

* Corresponding author.
E-mail address: hiep.luong@telin.ugent.be (H. Luong).

1047-3203/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2011.09.002

grid point after proper registration and upsampling. If several samples are located on the same HR grid point, the HR pixel is estimated as the mean or median value of these samples [2]. Because the samples are snapped to the nearest grid points, the shift-and-add algorithm additionally generates positional errors on top of the typical registration errors. This effect adds another kind of correlated noise and artifacts to the reconstructed images such as undesirable and false zipper artifacts around edges.

Another way to solve the problems of missing HR pixels is to enlarge the footprint of each sample of the LR images. The *variable-pixel linear reconstruction* algorithm, or informally known as *drizzling*, computes each HR pixel as the weighted average from all contributing surrounding samples [10]. A sample contributes to a HR pixel if the HR grid position is lying inside a square window around the sample, while the weight is determined by the degree of overlap between this square window and the area of the HR pixel lattice. An alternative to square windows is the use of adaptive ellipses, which results in *elliptical weighted area* (EWA) filtering techniques, where the ellipses are oriented according to the transformation [11]. Both concepts interpret samples as tiny *waterdrops* (hence the term drizzling) raining on the HR grid.

In the drizzling and EWA fusion techniques, all HR pixels within the coverage of a sample receive the same weight no matter how far the HR pixel is lying from the sample position. Assigning weights in function of the spatial distance between the HR pixel position and the sample position, results in the *Nadaraya–Watson* estimator [12]. In [13], *structure adaptive normalized convolution* approximates the local signal by a set of polynomial basis functions. The values on the HR grid is then computed from the combination of these basis functions. In [14], *kernel regression* is presented as a unified framework that combines the concepts of drizzling, EWA, Nadaraya–Watson estimator and normalized convolution methods resulting in a powerful image fusion technique.

The main drawback of the mentioned fusion methods is that these techniques do not explicitly take positional or registration errors into account. However, such errors are very common for registration algorithms used in practical SR applications, especially in the presence of severe image noise. Some existing robust image fusion and SR methods tackle image noise and outliers in general. In [15], the authors proposed a robust higher-order normalized convolution, which is an extension of the work in [16], that adds extra tonal weighting according to the confidence or certainty values. In [13], the structure adaptive normalized convolution is iteratively updated with a robust Gaussian-weighted error norm, in which outlier pixels are automatically neglected. This works quite well in SR applications in case there are only a very few misalignments or images with heavy-tailed distributed noise, but it is not designed in case a lot of (noisy) images are misaligned. In this paper, we propose a new and improved data measurement model that can cope with positional errors. From this formulation, we derive a novel kernel regression method in the TLS sense.

In the following sections, we briefly discuss the kernel regression technique. We then unify the steering kernel regression with the total least square formalism. We report numerical simulations on image reconstruction problems and finally, we end this paper with a conclusion.

2. Standard kernel regression

We briefly describe the kernel regression method for solving the resampling problem in the ordinary least square sense (KROLS) as proposed by Takeda et al. [14]. Suppose that we have to estimate the pixel value $f(\mathbf{x})$ at position $\mathbf{x}$ on the HR grid. In the surrounding neighbourhood, we have a set of p noisy measurements g_i at irregularly sampled positions $\mathbf{x}_i$, the data measurement model is then given by:

$$g_i = f(\mathbf{x}_i) + n_i, \quad i = 1, \ldots, p, \tag{1}$$

where $f(.)$ is the unknown HR image, which is also referred to as the *regression function* and n_i are independently and identically distributed zero-mean noise values. In a local neighbourhood, we can approximate the regression function by its local expansion of degree N. For example, we use the second order Taylor's series expansion ($N = 2$) of $f(\cdot)$, which is denoted by:

$$\begin{aligned} f(\mathbf{x}_i) &\approx f(\mathbf{x}) + \{\nabla f(\mathbf{x})\}^T \tilde{\mathbf{x}}_i + \frac{1}{2}\tilde{\mathbf{x}}_i^T \{\mathcal{H} f(\mathbf{x})\} \tilde{\mathbf{x}}_i \\ &\approx \beta_0 + \boldsymbol{\beta}_1^T \tilde{\mathbf{x}}_i + \tilde{\mathbf{x}}_i^T \boldsymbol{\beta}_2 \tilde{\mathbf{x}}_i, \end{aligned} \tag{2}$$

where $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \mathbf{x}$; ∇ and $\mathcal{H}$ are respectively the gradient and Hessian operators. The coefficients of this polynomial are estimated by the following weighted least-squares optimization problem $\left(\hat{\boldsymbol{\beta}} = \left\{\hat{\beta}_0, \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2\right\}\right)$:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{p} \left[g_i - \beta_0 - \boldsymbol{\beta}_1^T \tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_i^T \boldsymbol{\beta}_2 \tilde{\mathbf{x}}_i\right]^2 k_{\mathbf{H}}(\tilde{\mathbf{x}}_i), \tag{3}$$

which can easily be solved and where $\hat{f}(\mathbf{x}) = \hat{\beta}_0$ is the estimated pixel value at the position $\mathbf{x}$ on the HR grid, which we are looking for. The kernel function $k_{\mathbf{H}}(\cdot)$ (which has typically a Gaussian or exponential form) penalizes positions that are located further away from the grid position and its strength is controlled by the smoothing matrix $\mathbf{H}$:

$$k_{\mathbf{H}}(\tilde{\mathbf{x}}_i) = |\mathbf{H}|^{-1} k(\mathbf{H}^{-1}\tilde{\mathbf{x}}_i). \tag{4}$$

In the special case of $N = 0$, the solution of the kernel regression algorithm corresponds to the Nadaraya–Watson estimator:

$$\hat{f}(\mathbf{x}) = \hat{\beta}_0 = \frac{\sum_{i=1}^{p} g_i k_{\mathbf{H}}(\tilde{\mathbf{x}}_i)}{\sum_{i=1}^{p} k_{\mathbf{H}}(\tilde{\mathbf{x}}_i)}. \tag{5}$$

This estimator only models locally flat signals, but does not model edges, ridges and blobs very well. On the other hand, the estimator given by Eq. (3) also takes these edges, ridges and blobs into account.

In most applications, the 2×2 smoothing matrix $\mathbf{H}$ is equal to $h\mathbf{I}$ with h being the bandwidth parameter such that the kernel's footprint is isotropic. This is referred to as *classic* kernel regression. Iteratively adapting the kernel's footprint locally and anisotropically according to the samples prevent oversmoothing across edges. Therefore, the use of anisotropic kernel functions is referred to as *steering* kernel regression [14].

3. Proposed method

Multiframe SR algorithms require a very accurate estimation of the positions of the LR samples on the fine HR grid. However, in practice, small errors on the registration parameters or the use of limited motion models cause relatively large positional errors of the LR samples with the result that the quality of the SR generated image degrades dramatically. Therefore, it is important that image fusion also takes these spatial inaccuracies into account. The following improved data measurement model specifies that the relative positions $\mathbf{x}_i - \mathbf{x}$ can be subject to perturbations:

$$g_i = f(\mathbf{x}_i + \mathbf{u}_i) + n_i, \quad i = 1, \ldots, p, \tag{6}$$

where $\mathbf{u}_i$ is the relative positional error of $\mathbf{x}_i = (x_i, y_i)$ compared to the position $\mathbf{x} = (x, y)$ on the HR grid; $\mathbf{u}_i$ and n_i are assumed to be zero-mean distributed.

In case f is modeled by a linear regression function, we can find the parameters via the basic TLS algorithm using the singular value