Recent Advances in Signal Processing

Recent Advances in Signal Processing

Edited by
Ashraf A. Zaher

In-Tech
intechweb.org

Published by In-Teh

In-Teh
Olajnica 19/2, 32000 Vukovar, Croatia

www.intechweb.org
Additional copies can be obtained from:
publication@intechweb.org

First published November 2009
Printed in India

Technical Editor: Maja Jakobovic

Recent Advances in Signal Processing,
Edited by Ashraf A. Zaher

p. cm.
ISBN 978-953-307-002-5

Preface

The signal processing task is a very critical issue in the majority of new technological inventions and challenges in a variety of applications in both science and engineering fields. Classical signal processing techniques have largely worked with mathematical models that are linear, local, stationary, and Gaussian. They have always favored closed-form tractability over real-world accuracy. These constraints were imposed by the lack of powerful computing tools. During the last few decades, signal processing theories, developments, and applications have matured rapidly and now include tools from many areas of mathematics, computer science, physics, and engineering. This was mainly due to the revolutionary advances in the digital technology and the ability to effectively use digital signal processing (DSP) that rely on the use of very large scale integrated technologies and efficient computational methods such as the fast Fourier transform (FFT). This trend is expected to grow exponentially in the future, as more and more emerging technologies are revealed in the fields of digital computing and software development.

It is still an extremely skilled work to properly design, build and implement an effective signal processing tool able to meet the requirements of the increasingly demanding and sophisticated modern applications. This is especially true when it is necessary to deal with real-time applications of huge data rates and computational loads. These applications include image compression and encoding, speech analysis, wireless communication systems, biomedical real-time data analysis, cryptography, steganography, and biometrics, just to name a few. Moreover, the choice between whether to adopt a software or hardware approach, for implementing the application at hand, is considered a bottleneck. Programmable logic devices, e.g. FPGAs provide an optimal compromise, as the hardware configuration can be easily tailored using specific hardware descriptive languages (HDLs).

This book is targeted primarily toward both students and researchers who want to be exposed to a wide variety of signal processing techniques and algorithms. It includes 27 chapters that can be categorized into five different areas depending on the application at hand. These five categories are ordered to address image processing, speech processing, communication systems, time-series analysis, and educational packages respectively. The book has the advantage of providing a collection of applications that are completely independent and self-contained; thus, the interested reader can choose any chapter and skip to another without losing continuity. Each chapter provides a comprehensive survey of the subject area and terminates with a rich list of references to provide an in-depth coverage of the application at hand. Understanding the fundamentals of representing signals and systems in both time, spatial, and frequency domains is a prerequisite to read this book, as it is assumed that the reader is familiar with them. Knowledge of other transform methods, such as the Laplace transform and the Z-transform, along with knowledge of some computational intelligence techniques is an assist. In addition, experience with MATLAB programming (or a similar tool) is useful, but not essential. This book is application-oriented and it mainly addresses the design, implementation, and/or the improvements of existing or new technologies, and also provides some novel algorithms either in software, hardware, or both forms. The reported techniques are based on time-domain analysis, frequency-domain analysis, or a hybrid combination of both.

This book is organized as follows. The first 14 chapters investigate applications in the field of image processing, the next six chapters address applications in speech and audio processing, and the last seven chapters deal with applications in communication systems, real-time data handling, and interactive educational packages, respectively. There is a great deal of overlap between some of the chapters, as they might be sharing the same theory, application, or approach; yet, we chose to organize the chapter into the following five sections:

I. Image Processing:

This section contains 14 chapters that explore different applications in the field of image processing. These applications cover a variety of topics related to segmentation, encoding, restoration, steganography, and denoising. Chapters (1) to (14) are arranged into groups based on the application of interest as explained in the following table:

Chapter(s)	Main topic (application)
1 – 3	Image segmentation and encoding
4 – 6	Medical applications
7 & 8	Data hiding
9 & 10	Image classification
11& 12	Biometric applications
13 & 14	Noise suppression

Chapter (1) proposes a software approach to image stabilization that depends on two consequent steps of global image registration and image fusion. The improved reliability and the reduced size and cost of this approach make it ideal for small mobile devices. Chapter (2) investigates contour retrieval in images via estimating the parameters of rectilinear or circular contours as a source localization problem in high-resolution array processing. It presents a subspace-based line detection algorithm for the estimation of rectilinear contours based on signal generation upon a linear antenna. Chapter (3) proposes a locally adaptive resolution (LAR) codec as a contribution to the field of image compression and encoding. It focuses on a few representative features of the LAR technology and its preliminary associated performances, while discussing their potential applications in different image-related services.

Chapter (4) uses nonlinear locally adaptive transformations to perform image registration with application to MRI brains scan. Both parametric and nonparametric transformations, along with the use of multi-model similarity measures, are used to robustify the results to tissue intensity variations. Chapter (5) describes a semi-automated segmentation method for dynamic contrast-enhanced MRI sequences for renal function assessment. The superiority of the proposed method is demonstrated via testing and comparing it with manual segmentation by radiologists. Chapter (6) uses a hybrid technique of motion estimation and segmentation that are based on variational techniques to improve the performance of cardiac motion application in indicating heart diseases.

Chapter (7) investigates the problem of restricting color information for images to only authorized users. It surveys some of the reported solutions in the literature and proposes an improved technique to hide a 512-color palette in an 8-bit gray level image. Chapter (8) introduces a novel application of the JPEG2000-based information hiding for synchronized and scalable 3D visualization. It also provides a compact, yet detailed, survey of the state of the art techniques in the field of using DWT in image compression and encoding.

Chapter (9) uses a content-based image-retrieval technique to validate the results obtained from defects-detection algorithms, in Ad-hoc features, to find similar images suffering from the same defects in order to classify the questioned image as defected or not. Chapter (10) explores a novel approach for automatic crack detection and classification for the purpose of roads maintenance and estimating pavement surface conditions. This approach relies on image processing and pattern recognition techniques using a framework based on local statistics, computed over non-overlapping image regions.

Chapter (11) proposes a robust image segmentation method to construct a contact-free hand identification system via using infrared illumination and templates that guide the user in order to minimize the projective distortions. This biometric identification system is tested on a real-world database, composed by 102 users and more than 4000 images, resulting in an EER of 3.2%. Chapter (12) analyzes eye movements of subjects when looking freely at dynamic stimuli such as videos. This study uses face detection techniques to prove that faces are very salient in both static and dynamic stimuli.

Chapter (13) reports the use of specialized denoising algorithms that deal with correlated noise in images. Several useful noise estimation techniques are presented that can be used when creating or adapting a white noise denoising algorithm for use with correlated noise.

Chapter (14) presents a novel technique that estimates and eliminates additive noise inherent in images acquired under incoherent illumination. This technique combines the two methods of scatter plot and data masking to preserve the physical content of polarization-encoded images.

II. Speech/Audio Processing:

This section contains six chapters that explore different applications in the field of speech and audio processing. These applications cover a variety of topics related to speech analysis, enhancement of audio quality, and classification of both audio and speech. Chapters (15) to (20) are arranged into groups based on the application of interest as explained in the following table:

Chapter(s)	Main topic (application)
15 & 16	Speech/audio enhancement
17 & 18	Biometric applications
19 & 20	Speech/audio analysis

Chapter (15) proposes an improved iterative Wiener filter (IWF) algorithm based on the time-varying complex auto regression (TV-CAR) speech analysis for enhancing the quality of speech. The performance of the proposed system is compared against the famous linear predictive coding (LPC) method and is shown to be superior. Chapter (16) introduces a robust echo detection algorithm in mobile phones for improving the calls quality. The structure for the echo detector is based on comparison of uplink and downlink pitch periods. This algorithm has the advantage of processing adaptive multi-rate (AMR) coded speech signals without decoding them first and its performance is demonstrated to be satisfactory.

Chapter (17) investigates the problem of voice/speaker recognition. It compares the effectiveness of using a combination of vector quantization (VQ) and different forms for the Mel frequency cepstral coefficients (MFCCs) when using the Gaussian mixture model for modeling the speaker characteristics. Chapter (18) deals with issues, related to processing and mining of specific speech information, which are commonly ignored by the mainstream research in this field. These issues focus on speech with emotional content, effects of drugs and Alcohol, speakers with disabilities, and various kinds of pathological speech.

Chapter (19) uses narrow-band filtering to construct an estimation technique of instantaneous parameters used in sinusoidal modeling. The proposed method utilizes pitch detection and estimation for achieving good analysis of speech signals. Chapter (20) conducts an experimental study on 420 songs from four different languages to perform statistical analysis of the music information that can be used as prior knowledge in formulating constrains for music information extraction systems.

III. Communication Systems:

This section contains three chapters that deal with the transmission of signals through public communication channels. Chapters (21) to (23) discuss the problems of modeling and simulation of multi-input multi-output wireless channels, multi-antenna receivers, and chaos-based cryptography, respectively. Chapter (21) discusses how to construct channel simulators for multi-input multi-output (MIMO) communication systems for testing physical layer algorithms such as channel estimation. It also presents the framework, techniques, and theories in this research area. Chapter (22) presents a new approach to the broadcast channel problem that is based on combining dirty-paper coding (DPC) with zero-forcing (ZF) precoder and optimal beamforming design. This approach can be applied to the case when several antennas coexist at the receiver. It also introduces an application that deals with the cooperation design in wireless sensor networks with intra and intercluster interference. Chapter (23) investigates three important steps when establishing a secure communication system using chaotic signals. Performing fast synchronization, identifying unknown parameters, and generating robust cryptography are analyzed. Different categories of systems are introduced and real-time implementation issues are discussed.

IV. Time-series Processing:

This section contains three chapters that deal with real-time data handling and processing. These data can be expressed as functions of time, sequence of images, or readings from sensors. It provides three different applications. Chapter (24) introduces an application, which is based on the fusion of electronecephalography (EEG) and functional magnetic resonance imaging (fMRI), for the detection of seizure. It proposes a novel constrained spatial independent component analysis (ICA) algorithm that outperforms the existing unconstrained algorithm in terms of estimation error and closeness between the component time course and the seizure EEG signals. Chapter (25) introduces the design and implementation of a real-time measurement system for estimating the air parameters that are vital for effective and reliable flights. The proposed system is installed in the cockpit of the aircraft and uses two embedded PCs and four FPGA signal processing boards. It utilizes laser beams for estimating the air parameters necessary for the safety of the flight. Chapter (26) discusses the performance of the target signal port-starboard discrimination for underwater towed multi-line arrays that have typical applications in military underwater surveillance and seismic exploring

V. Educational Systems:

Chapter (27) introduces an open source software package that can be used as an educational tool for teaching signal processing in a variety of fields including image and audio processing. It provides an interactive environment that is easy to use with GUI and web interface that is XML-based. This package can be used as an alternative to other existing packages including J-DSP, Simulink and SciLab.

November 2009

Ashraf A. Zaher

Contents

1

Digital Image Stabilization

Marius Tico
Nokia Research Center
Palo Alto, CA, USA

1. Introduction

The problem of image stabilization dates since the beginning of photography, and it is basically caused by the fact that any known image sensor needs to have the image projected on it during a period of time called integration time. Any motion of the camera during this time causes a shift of the image projected on the sensor resulting in a degradation of the final image, called motion blur.

The ongoing development and miniaturization of consumer devices that have image acquisition capabilities increases the need for robust and efficient image stabilization solutions. The need is driven by two main factors: (i) the difficulty to avoid unwanted camera motion when using a small hand-held device (like a camera phone), and (ii) the need for longer integration times due to the small pixel area resulted from the miniaturization of the image sensors in conjunction with the increase in image resolution. The smaller the pixel area the less photons/second could be captured by the pixel such that a longer integration time is needed for good results.

It is of importance to emphasize that we make a distinction between the terms "digital image stabilization" and "digital video stabilization". The latter is referring to the process of eliminating the effects of unwanted camera motion from video data, see for instance Erturk & Dennis (2000); Tico & Vehviläinen (2005), whereas digital image stabilization is concerned with correcting the effects of unwanted motions that are taking place during the integration time of a single image or video frame.

The existent image stabilization solutions can be divided in two categories based on whether they are aiming to correct or to prevent the motion blur degradation. In the first category are those image stabilization solutions that are aiming for restoring a single image shot captured during the exposure time. This is actually the classical case of image capturing, when the acquired image may be corrupted by motion blur, caused by the motion that have taken place during the exposure time. If the point spread function (PSF) of the motion blur is known then the original image can be restored, up to some level of accuracy (determined by the lost spatial frequencies), by applying an image restoration approach Gonzalez & Woods (1992); Jansson (1997). However, the main difficulty is that in most practical situations the motion blur PSF is not known. Moreover, since the PSF depends of the arbitrary camera motion during the exposure time, its shape is different in any degraded image as exemplified in Fig. 1. Another difficulty comes from the fact that the blur degradation is not spatially invariant over the image area. Thus, moving objects in the scene may result in very different blur models in certain image areas. On the other hand, even less dynamic scenes may contain different blur models in different regions in accordance to the distance between the objects and the camera,

Fig. 1. Different camera motions cause different blur degradations.

i.e., during a camera translation close objects have larger relative motions than distant objects, phenomenon known as "parallax".

In order to cope with the insufficient knowledge about the blur PSF one could adopt a blind de-convolution approach, e.g., Chan & Wong (1998); You & Kaveh (1996). Most of these methods are computationally expensive and they have reliability problems even when dealing with spatially invariant blur. Until now, published research results have been mainly demonstrated on artificial simulations and rarely on real world images, such that their potential use in consumer products seems rather limited for the moment.

Measurements of the camera motion during the exposure time could help in estimating the motion blur PSF and eventually to restore the original image of the scene. Such an approach have been introduced by Ben-Ezra & Nayar (2004), where the authors proposed the use of an extra camera in order to acquire motion information during the exposure time of the principal camera. A different method, based on specially designed high-speed CMOS sensors has been proposed by Liu & Gamal (2003). The method exploits the possibility to independently control the exposure time of each image pixel in a CMOS sensor. Thus, in order to prevent motion blur the integration is stopped selectively in those pixels where motion is detected.

Another way to estimate the PSF has been proposed in Tico et al. (2006); Tico & Vehviläinen (2007a); Yuan et al. (2007), where a second image of the scene is taken with a short exposure. Although noisy, the secondary image is much less affected by motion blur and it can be used as a reference for estimating the motion blur PSF which degraded the principal image.

In order to cope with the unknown motion blur process, designers have adopted solutions able to prevent such blur for happening in the first place. In this category are included all optical image stabilization (OIS) solutions adopted nowadays by many camera manufactures. These solutions are utilizing inertial senors (gyroscopes) in order to measure the camera motion, following then to cancel the effect of this motion by moving either the image sensor Konika Minolta Inc. (2003), or some optical element Canon Inc. (2006) in the opposite direction. The miniaturization of OIS systems did not reach yet the level required for implementation in a small device like a camera phone. In addition, most current OIS solutions cannot cope well with longer exposure times. In part this is because the inertial motion sensors, used to measure the camera motion, are less sensitive to low frequency motions than to medium and high frequency vibrations. Also, as the exposure time increases the mechanism may drift due to accumulated errors, producing motion blurred images (Fig. 2).

An image acquisition solution that can prevent motion blur consists of dividing long exposure times in shorter intervals, following to capture multiple short exposed image frames of

(a) (b) (c)

Fig. 2. Optical image stabilization examples at different shutter speeds. The images have been captured with a hand-held camera using Canon EF-S 17-85mm image stabilized lens. The exposure times used in taking the pictures have been: (a) 1/25sec, (b) 1/8sec, and (c) 1/4sec. The images get increasingly blurred as the shutter speed slows down.

the same scene. Due to their short exposure, the individual frames are corrupted by sensor noises (e.g., photon-shot noise, readout noise) Nakamura (2006) but, on the other hand, they are less affected by motion blur. Consequently, a long exposed and motion blur free picture can be synthesized by registering and fusing the available short exposed image frames (see Tico (2008a;b); Tico & Vehviläinen (2007b)). Using this technique the effect of camera motion is transformed from a motion blur degradation into a misalignment between several image frames. The advantage is that the correction of the misalignment between multiple frames is more robust and computationally less intensive than the correction of a motion blur degraded image.

In this chapter we present the design of such a multi-frame image stabilization solution, addressing the image registration and fusion operations. A global registration approach, described in Section 2, assists the identification of corresponding pixels between images. However the global registration cannot solve for motion within the scene as well as for parallax. Consequently one can expect local misalignments even after the registration step. These will be solved in the fusion process described in Section 3.

2. Image registration

Image registration is essential for ensuring an accurate information fusion between the available images. The existent approaches to image registration could be classified in two categories: feature based, and image based methods, Zitova & Flusser (2003). The feature based methods rely on determining the correct correspondences between different types of visual features extracted from the images. In some applications, the feature based methods are the most effective ones, as long as the images are always containing specific salient features (e.g., minutiae in fingerprint images Tico & Kuosmanen (2003)). On the other hand when the number of detectable feature points is small, or the features are not reliable due to various image degradations, a more robust alternative is to adopt an image based registration approach, that utilizes directly the intensity information in the image pixels, without searching for specific visual features.

In general a parametric model for the two-dimensional mapping function that overlaps an "input" image over a "reference" image is assumed. Let us denote such mapping function by $\mathbf{t}(\mathbf{x};\mathbf{p}) = [t_x(\mathbf{x};\mathbf{p})\ t_y(\mathbf{x};\mathbf{p})]^t$, where $\mathbf{x} = [x\ y]^t$ stands for the coordinates of an image pixel, and $\mathbf{p}$ denotes the parameter vector of the transformation. Denoting the "input" and "reference" images by h and g respectively, the objective of an image based registration approach is to estimate the parameter vector $\mathbf{p}$ that minimizes a cost function (e.g., the sum of square differences) between the transformed input image $h(\mathbf{t}(\mathbf{x};\mathbf{p}))$ and the reference image $g(\mathbf{x})$.

The minimization of the cost function, can be achieved in various ways. A trivial approach would be to adopt an exhaustive search among all feasible solutions by calculating the cost function at all possible values of the parameter vector. Although this method ensures the discovery of the global optimum, it is usually avoided due to its tremendous complexity. To improve the efficiency several alternatives to the exhaustive search technique have been developed by reducing the searching space at the risk of losing the global optimum, e.g., logarithmic search, three-step search, etc, (see Wang et al. (2002)). Another category of image based registration approaches, starting with the work of Lucas & Kanade (1981), and known also as gradient-based approaches, assumes that an approximation to image derivatives can be consistently estimated, such that the minimization of the cost function can be achieved by applying a gradient-descent technique (see also Baker & Matthews (2004); Thevenaz & Unser (1998)). An important efficiency improvement, for Lucas-Kanade algorithm, has been proposed in Baker & Matthews (2004), under the name of "Inverse Compositional Algorithm" (ICA). The improvement results from the fact that the Hessian matrix of the cost function, needed in the optimization process, is not calculated in each iteration, but only once in a pre-computation phase.

In this work we propose an additional improvement to gradient-based methods, that consists of simplifying the repetitive image warping and interpolation operations that are required during the iterative minimization of the cost function. Our presentation starts by introducing an image descriptor in Section 2.1, that is less illumination dependent than the intensity component. Next, we present our registration algorithm in Section 2.2, that is based on matching the proposed image descriptors of the two images instead their intensity components.

2.1 Preprocessing

Most of the registration methods proposed in the literature are based on matching the intensity components of the given images. However, there are also situations when the intensity components do not match. The most common such cases are those in which the two images have been captured under different illumination conditions, or with different exposures.

In order to cope with such cases we propose a simple preprocessing step aiming to extract an illumination invariant descriptor from the intensity component of each image. Denoting by $H(\mathbf{x})$ the intensity value in the pixel $\mathbf{x}$, and with $\mathrm{avg}(H)$ the average of all intensity values in the image, we first calculate $\bar{H}(\mathbf{x}) = H(\mathbf{x})/\mathrm{avg}(H)$, in order to gain more independence from the global scene illumination. Next, based on the gradient of $\bar{H}$ we calculate $H_g(\mathbf{x}) = |\bar{H}_x(\mathbf{x})| + |\bar{H}_y(\mathbf{x})|$ in each pixel, and $\mathrm{med}(H_g)$ as the median value of $H_g(\mathbf{x})$ over the entire image.

Finally, the actual descriptor that we are using in the registration operation is given by the following binary image

$$h(\mathbf{x}) = \begin{cases} 1 & \text{if } H_g(\mathbf{x}) > \text{med}(H_g) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

2.2 Registration algorithm

In the following we describe an image based registration method that is using a multi-resolution coarse to fine strategy. Typically in such algorithm, at each iteration step one of the images should be warped in accordance to the parameters estimated so far. In our method this warping operation is highly simplified on the expense of increase memory usage.

The levels of the multi-resolution representation are over-sampled, and they are obtained by iteratively smoothing the original image descriptor h, such that to obtain smoother and smoother versions of it. Let $\tilde{h}_\ell$ denotes the smoothed image resulted after ℓ-th low-pass filtering iterations ($\tilde{h}_0 = h$). The smoothed image at next iteration can be calculated by applying one-dimensional filtering along the image rows and columns as follows:

$$\tilde{h}_{\ell+1}(x, y) = \sum_r w_r \sum_c w_c \tilde{h}_\ell \left(x - 2^\ell c, y - 2^\ell r\right), \quad (2)$$

where w_k are the taps of a low-pass filter.

The registration approach takes advantage of the fact that each decomposition level ($\tilde{h}_\ell$) is over-sampled, and hence it can be reconstructed by a subset of its pixels. This property allows to enhance the efficiency of the registration process by using only a subset of the pixels in the registration algorithm. The advantage offered by the availability of over-sampled decomposition level, is that the set of pixels that can be used in the registration is not unique. A broad range of geometrical transformations can be approximated by simply choosing a different set of pixels to describe the sub-sampled image level. In this way, the over-sampled image level is regarded as a "reservoir of pixels" for different warped sub-sampled versions of the image, which are needed at different stages in the registration algorithm.

Let $\mathbf{x}_{n,k} = [x_{n,k} \; y_{n,k}]^t$, for n, k integers, denote the coordinates of the selected pixels into the smoothed image ($\tilde{h}_\ell$). A low-resolution version of the image ($\hat{h}_\ell$) can be obtained by collecting the values of the selected pixels: $\hat{h}_\ell(n,k) = \tilde{h}_\ell(\mathbf{x}_{n,k})$. Moreover, given an invertible geometrical transformation function $\mathbf{t}(\mathbf{x}; \mathbf{p})$, the warping version of the low resolution image can be obtained more efficiently by simply selecting another set of pixels from the area of the smoothed image, rather than warping and interpolating the low-resolution image $\hat{h}_\ell$. This is: $\hat{h}'_\ell(n,k) = \tilde{h}_\ell(\mathbf{x}'_{n,k})$, where $\mathbf{x}'_{n,k} = \text{round}\left(\mathbf{t}^{-1}(\mathbf{x}_{n,k}; \mathbf{p})\right)$.

The process described above is illustrated in Fig.3, where the images shown on the bottom row represent two low-resolutions warped versions of the original image (shown in the top-left corner). The two low-resolution images are obtained by sampling different pixels from the smoothed image (top-right corner) without interpolation.

The registration method used in our approach is presented in Algorithm 1. The algorithm follows a coarse to fine strategy, starting from a coarse resolution level and improving the parameter estimate with each finer level, as details in the Algorithm 2. The proposed algorithm relies on matching image descriptors (1) derived from each image rather than image intensity components.

Algorithm 2 presents the registration parameter estimation at one resolution level. In this algorithm, the constant N_0, specifies the number of iterations the algorithm is performing

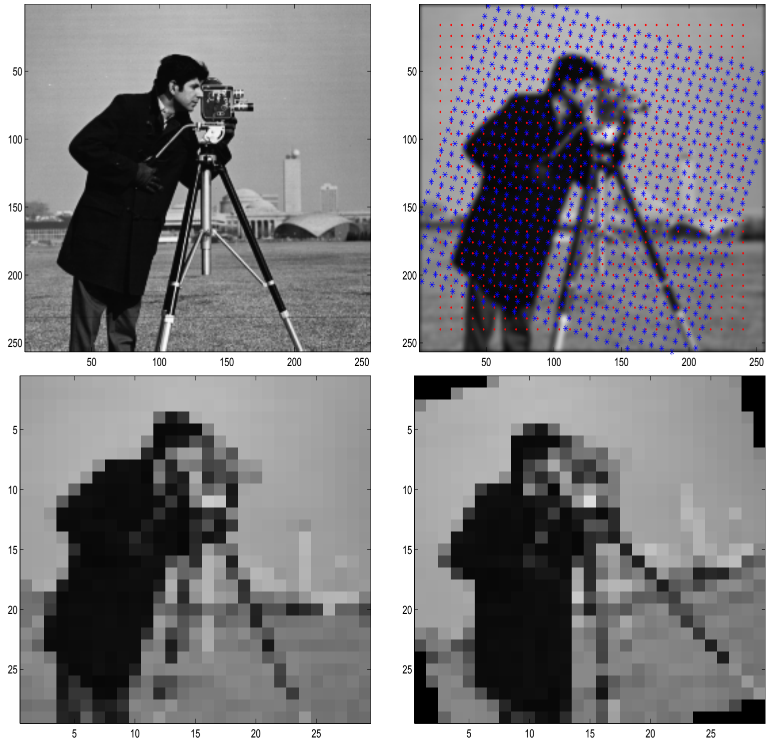

Fig. 3. Low-resolution image warping by re-sampling an over-sampled image decomposition level.

Algorithm 1 Global image registration

Input: the input and reference images plus, if available, an initial guess of the parameter vector $\mathbf{p} = [p_1\ p_2\ \cdots\ p_K]^t$.
Output: the parameter vector that overlaps the input image over the reference image.

1- Calculate the descriptors (1) for input and reference images, denoted here by h and g, respectively.

2- Calculate the decomposition levels of the two image descriptors $\{\tilde{h}_\ell, \tilde{g}_\ell \mid \ell_{min} \leq \ell \leq \ell_{max}\}$.

3- For each level ℓ between ℓ_{max} and ℓ_{min}, do Algorithm 2.

---

after finding a minima of the error function. This is set in order to reduce the chance of ending in a local minima. As shown in the algorithm the number of iterations is reset to N_0, every time a new minima of the error function is found. The algorithm stops only if no other minima is found in N_0 iterations. In our experiments a value $N_0 = 10$ has been used.