Image restoration using truncated SVD filter bank based on an energy criterion

X. Zhang and S. Wang

Abstract: Image restoration is formulated using a truncated singular-value-decomposition (SVD) filter bank. A pair of known data patterns is used for identifying a small convolution operator. This is achieved by matrix pseudo-inversion based on SVD. Unlike conventional approaches, however, here SVD is performed upon a data-pattern matrix that is much smaller than the image size, leading to an enormous saving in computation. Regularisation is realised by first decomposing the operator into a bank of sub-filters, and then discarding some high-order ones to avoid noise amplification. By estimating the noise spectrum, sub-filters that produce noise energy more than that of useful information are abandoned. Therefore high-order components in the spectrum responsible for noise amplification are rejected. With the obtained small kernel, image restoration is implemented by convolution in the space domain. Numerical results are given to show the effectiveness of the proposed technique.

Notation

$\boldsymbol{b}$	1D restoration PSF in separable case, or column-stacked vector of $\boldsymbol{B}$ in non-separable case
$\boldsymbol{b}_r$	vector corresponding to the rth SVD filter
$\boldsymbol{B}$	large matrix equivalent to $\boldsymbol{b}$
$\boldsymbol{B}$	restoration PSF of the blurring process
$\boldsymbol{B}^{(0)}$	portion of $\boldsymbol{B}$ with $C_r = 0$
$\boldsymbol{B}^{(1)}$	portion of $\boldsymbol{B}$ with $C_r = 1$
$\boldsymbol{B}_r$	large matrix derived from $\boldsymbol{b}_r$
C_r	switches corresponding to $\boldsymbol{b}_r$
$\boldsymbol{E}_r^{(y)}$	total output energy from the rth SVD filter
$\hat{\boldsymbol{E}}_r^{(n)}$	estimation of output noise energy from the rth SVD filter
$\boldsymbol{f}$	standard test pattern
$[\boldsymbol{F}]$	circulant matrix representing $\boldsymbol{f}$
$\boldsymbol{g}$	degraded test pattern
$[\boldsymbol{G}]$	circulant matrix representing $\boldsymbol{g}$
$\boldsymbol{h}$	1D degradation PSF in separable case, or column-stacked vector of $\boldsymbol{H}$ in non-separable case
$\boldsymbol{H}$	PSF of the blurring process
$\boldsymbol{n}$	additive noise
$\boldsymbol{x}$	ideal image
$\hat{\boldsymbol{x}}$	restored image
$\boldsymbol{y}$	degraded image

IEE Proceedings online no. 20045200

doi:10.1049/ip-vis:20045200

Paper first received 1st October 2004 and in final revised form 17th June 2006

The authors are with the School of Communication and Information Engineering, Shanghai University, 149 Yanchang Road, Shanghai 200072, People's Republic of China

E-mail: xzhang@staff.shu.edu.cn

1 Introduction

Using a vector representation, image degradation is commonly described by the following model [1]

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{1}$$

In (1), $\boldsymbol{x}$ and $\boldsymbol{y}$ are the ideal and observed (degraded) images, sized $N_1N_2 \times 1$ and $M_1M_2 \times 1$, respectively; and $\boldsymbol{n}$, sized $M_1M_2 \times 1$, is independent additive zero mean noise. Entries in these vectors are arranged in a lexicographical order. $\boldsymbol{H}$ is an $M_1M_2 \times N_1N_2$ block Toeplitz matrix representing a linear and shift-invariant degradation filter. The objective of image restoration is to find an estimate of the ideal image, $\hat{\boldsymbol{x}}$, given the observation $\boldsymbol{y}$ and the degradation operator $\boldsymbol{H}$. The estimate should resemble the original image as closely as possible. The Moore–Penrose pseudo-inverse of $\boldsymbol{H}$ may be used to produce a least-square minimum-norm solution to (1)

$$\hat{\boldsymbol{x}} = \boldsymbol{H}^{+}\boldsymbol{y} \tag{2}$$

which minimises the following quantity

$$Q(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 \tag{3}$$

In this solution, $\boldsymbol{H}^{+}$ is a simple inverse filter suffering from noise amplification and instability due to the ill-posed nature of image restoration. Methods for dealing with this ill-posed problem have been reported extensively, among which constrained regularisation approaches introduce an operator $\boldsymbol{C}$ and a parameter λ into the least square criterion [2–7]

$$Q_\lambda(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \lambda\|\boldsymbol{C}\boldsymbol{x}\|^2 \tag{4}$$

$\boldsymbol{C}$ is usually chosen to reflect some *a priori* knowledge of $\boldsymbol{x}$; and the regularisation parameter λ should be optimised to provide a good tradeoff between data fidelity and smoothness of the restored image. A detailed study on ill-posed problems and the regularisation methods can be found in the work of Tikhonov and Arsenin [8].

Karayiannis and Venetsanopoulos [2] formulated image restoration problems as constrained minimisation of a stabilising functional. Their choice of a quadratic functional is related to the *a priori* knowledge of the original image through formulation of a maximum *a posteriori* estimation problem, based on certain stochastic partial differential equation image models. Thompson *et al.* [3] discussed ways for choosing a scalar regularisation parameter in order to achieve satisfactory performance. Galatsanos and Katsaggelos [4] used an error analysis based on an objective mean-square-error (MSE) criterion to motivate regularisation, and proposed two regularisation approaches. The first method is a direct minimisation of MSE, which requires knowledge of noise variance. The second method does not require noise variance, but assumes Gaussian posteriors and priors, and estimates a hyper-parameter for regularisation by maximising a marginal likelihood function.

Kang and Katsaggelos [5] used a regularisation functional instead of a constant regularisation parameter. In their iterative restoration, the required prior information was extracted from the available data at the previous iteration step, that is, the partially restored image. The parameter and restoration of the degraded image are obtained simultaneously. The functional to be minimised is convex leading to a global minimisation, and optimisation was independent of the initial conditions. Taratorin and Sideman [6] presented a consistent approach to the solution of a constrained regularised restoration problem. It was based on minimising the Tikhonov's regularising functional [8] subject to a set of constraints. The authors demonstrated that the regularised solution set resulting from minimisation of Tikhonov's functional is closed and convex, and a projection operator on this set is then derived.

Rajagopalan and Chaudhuri [7] described a regularised solution to a depth-from-defocus problem. They introduced a smoothness constraint to the estimates of the blur parameters, and developed a variational approach leading to results better than those obtained without the constraint. In general, regularisation heavily depends on noise variance. Estimation of noise variance may be accomplished by using various methods [3, 4, 9]. Heuristic or empirical methods are also employed [1, 10].

Reeves and Higdon [11] conducted a perceptual study to deal with the question as to how a good choice of the regularisation parameter such as minimum MSE (MMSE) can relate to human perceptual judgement of the restoration quality. In their investigation, they found that MMSE did not agree with, but was close enough to, the perceptually optimal choice of the parameter. Nguyen *et al.* [12] derived a generalised cross-validation formula for underdetermined linear system in image restoration and resolution enhancement problems, and obtained the regularisation parameters for stabilisation of their solution by using the generalised cross-validation technique. In an order statistics based image restoration approach [13], a least mean square algorithm was used to update the weighting coefficients of L-filter recursively.

The pseudo-inverse $\boldsymbol{H}^{+}$ in (2) can be computed using singular-value-decomposition (SVD). In this case, regularisation may be achieved by discarding some components corresponding to small singular values (SVs). This is termed as truncated SVD (TSVD) [10, 14–17]. Another method is to introduce a regularising parameter or a set of parameters to properly modify the SVD so that noise effect is significantly suppressed to give a sufficiently smooth restoration [18–20]. Both analytical and *ad hoc* methods exist for optimisation of the regularisation parameters. In all these approaches, regularisation is closely related to the signal-to-noise ratio of the observation.

Image restoration usually requires huge computation capacity. The conventional SVD restoration, for instance, involves operations on matrices as large as $M_1M_2 \times N_1N_2$. Efforts have been made to find efficient algorithms. Fish *et al.* [21] proposed a scanning SVD approach in which SVD is performed over a scanning window much smaller than the image. In order to alleviate blocking effects, only the centre portion of the window is retained. This means overlapping of the scanning windows. To design a small-sized FIR filter to solve a real-time restoration problem, Schutten and Vermeij [22] first formed their filter in the frequency domain. The coefficients are then inverse-transformed, and truncated heuristically. Reichenbach and Park [23] proposed a method using a small convolution kernel that is a constrained Wiener filter based on a compact support. Wiener filtering resolves the ill-posed difficulty by incorporating power spectra of both image and additive noise, and is most conveniently implemented in the frequency domain. After a small kernel is obtained, image restoration is performed in the space domain [23]. Under the constraint of a spatially limited support, the performance of Reichenbach's filter is optimal in the MMSE sense under the finite-support constraint, and quite close to the unconstrained Wiener filter. The method requires knowledge of power spectra of both the ideal image and noise. If this knowledge is unavailable, the power spectra have to be estimated from the observed image.

In this paper, we introduce a small-sized image restoration operator using a TSVD filter bank. An algebraic solution is sought to give an algorithm that is easy to implement in the space domain. Unlike conventional approaches, we look for a pseudo-inverse of a characteristic data pattern that is a blurred version of a special test pattern. In other words, the SVD operation is performed not on a huge matrix $\boldsymbol{H}$ representing the blurring filter, but rather, on a fabricated data matrix that is only a fraction of $\boldsymbol{H}$ in size. The restoration operator is decomposed into a bank of SVD sub-filters and some high-order filters responsible for noise amplification are truncated. The estimated noise power spectrum is obtained from the observed degraded image, and an appropriate truncating approach is based on analysis of energy contained in estimation errors of the restored image.

2 Formulation of image restoration in terms of test patterns

2.1 Separable case

In this section, a deconvolution kernel is first derived for a simple case in which the degradation operator is separable. Thus the kernel can be expressed as an outer product of two vectors, $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$. (In this paper, bold sans serif fonts, for example, $\boldsymbol{f}$ and $\boldsymbol{H}$, are used to represent small-sized convolution kernels and data patterns with the same order of magnitude in size, to be distinguished from vectors and matrices in much larger sizes such as $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{H}$ that are in bold serif fonts.)

$$\boldsymbol{H} = \boldsymbol{h}_1\boldsymbol{h}_2^{\mathrm{T}} \tag{5}$$

The subscripts 1 and 2 correspond to the column and row operations, respectively. In general, the point spread function (PSF) of the blurring process may be considered as 'compact', that is, $\boldsymbol{H}$ is a small matrix, sized $L_1 \times L_2$ where L_i is considerably smaller than the dimension of the image. It is reasonable to assume that there exists a

restoration PSF, or rather, a 'point converging operator', sized $K_1 \times K_2$, expressed in a similar way

$$\boldsymbol{B} = \boldsymbol{b}_1 \boldsymbol{b}_2^{\mathrm{T}} \tag{6}$$

Thus, the separable restoration reduces to a 1D problem, in which a vector-operator, $\boldsymbol{b}$, is sought from the given degradation operator $\boldsymbol{h}$. The restored image may be obtained by simply convolving the degraded image with $\boldsymbol{B}$ calculated from (6). Strictly speaking, the support of $\boldsymbol{B}$ may well cover a large spatial area. In practice, however, only the central elements of the PSF are significant. Therefore we assume a limited support for $\boldsymbol{B}$, which should usually be equal to, or greater than, that of $\boldsymbol{H}$ ($K \geq L$). The choice of K will be accounted for later.

Our aim is to construct a small-sized operator for restoration that takes the degraded observation as its input, and generates an output being an estimate of the original image. In identifying the restoration system, one needs a pair of known data sequences to be used as the input and output, respectively.

Let us first pass a binary step data sequence (length N) through the degradation filter $\boldsymbol{h}$ (length L)

$$\begin{aligned} \boldsymbol{f} &= [f_1 \quad f_2 \quad \cdots \quad f_N]^{\mathrm{T}} \\ &= [0 \quad 0 \quad \cdots \quad 0 \quad 1 \quad 1 \quad \cdots \quad 1]^{\mathrm{T}} \end{aligned} \tag{7}$$

The choice of a test pattern is, to some extent, arbitrary. For example, the transition from zero to one may be shifted. A square pulse with two opposite transitions may also be used. To fully explore the characteristics of the blurring process, however, it is desirable to let the test sequence contain sharp changes implying rich frequency contents, and flat periods of sufficient length so that the blurred output can reach a steady state before the next transition. The step sequence in (7) is used in the present work because it is simple and can lead to a short restoration operator. Convolving $\boldsymbol{f}$ with $\boldsymbol{h}$ yields a blurred output sequence, $\boldsymbol{g}$

$$\boldsymbol{g} = [g_1 \quad g_2 \quad \cdots \quad g_M]^{\mathrm{T}} \tag{8}$$

where

$$\begin{aligned} g_m &= \sum_{l=1}^{L} f_{m+l-1} h_{L-l+1}, \\ & m = 1, 2, \ldots, M; \quad M = N - L + 1 \end{aligned} \tag{9}$$

Here, the degradation operator is normalised

$$\sum_{l=1}^{L} h_l = 1 \tag{10}$$

We will use $\boldsymbol{f}$ and $\boldsymbol{g}$ as the input and output signals, respectively, in identifying the restoration system.

As the objective is to find a linear operator, given both input and output signals, the convolution is expressed in the following manner

$$\begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_M \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & \cdots & f_L \\ f_2 & f_3 & \cdots & f_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ f_M & f_{M+1} & \cdots & f_{M+L-1} \end{bmatrix} \begin{bmatrix} h_L \\ h_{L-1} \\ \vdots \\ h_1 \end{bmatrix} \tag{11}$$

or in a compact matrix form

$$\boldsymbol{g} = [\boldsymbol{F}]\boldsymbol{h} \tag{12}$$

Note that the signal components here are deemed as coefficients of a system of linear equations, while the elements in the linear operator, $\boldsymbol{h}$, are in the position of unknowns. The $M \times L$ matrix $[\boldsymbol{F}]$, consisting of 0s and 1s, has a shift property and may be termed the standard test pattern (STP). An upper-case italic letter in square brackets represents a test data pattern.

The degraded data sequence can be written in a straightforward way

$$\begin{aligned} \boldsymbol{g} = [0 \quad \cdots \quad & h_1 \quad h_2 + h_1 \quad \cdots \quad h_L + \cdots \\ & + h_2 + h_1 \quad \cdots \quad 1]^{\mathrm{T}} \end{aligned} \tag{13}$$

Next, pass $\boldsymbol{g}$ through a restoration filter, $\boldsymbol{b}$ (length K), to yield an estimate of the original test data sequence

$$\begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \vdots \\ \hat{f}_J \end{bmatrix} = \begin{bmatrix} g_1 & g_2 & \cdots & g_K \\ g_2 & g_3 & \cdots & g_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ g_J & g_{J+1} & \cdots & g_{J+K-1} \end{bmatrix} \begin{bmatrix} b_K \\ b_{K-1} \\ \vdots \\ b_1 \end{bmatrix}, \quad J = M - K + 1 \tag{14}$$

Denote the estimate as a vector $\boldsymbol{f}_0$

$$\boldsymbol{f}_0 = [\boldsymbol{G}]\boldsymbol{b} \tag{15}$$

The $J \times K$ matrix $[\boldsymbol{G}]$ may be termed as the degraded test pattern (DTP). Given $\boldsymbol{f}$ and $\boldsymbol{h}$, $[\boldsymbol{G}]$ is available for the subsequent treatment. If the degradation filter is unknown, $[\boldsymbol{G}]$ may be measured from the observed image [1]. The length of $\boldsymbol{f}_0$ is smaller than that of the original sequence $\boldsymbol{f}$ ($J < N$), and we hope that $\boldsymbol{f}_0$ is very similar to the centre part of $\boldsymbol{f}$. According to (13) and (14), the first few elements of $\boldsymbol{f}_0$ must be zeros, and the last few elements of $\boldsymbol{f}_0$ must be $\sum_{k=1}^{K} b_k$. As redundant elements on both ends of $\boldsymbol{f}_0$ do not provide any useful information, minimum lengths of the data sequences should be used in order to save computation loads.

1. Restored data sequence $\boldsymbol{f}_0$: $J_{\min} = K + L - 1$
2. Degraded data sequence $\boldsymbol{g}$: $M_{\min} = 2K + L - 2$
3. Original data sequence $\boldsymbol{f}$: $N_{\min} = 2K + 2L - 3$

By making $\boldsymbol{f}_0$ identical to the centre part of $\boldsymbol{f}$, a least-square minimum-norm solution of $\boldsymbol{b}$ is obtained

$$\hat{\boldsymbol{b}} = [\boldsymbol{G}]^{+}\boldsymbol{f}_0 \tag{16}$$

In this solution, $[\boldsymbol{G}]^{+}$ is a Moore–Penrose pseudo-inverse of the overdetermined DTP matrix ($[\boldsymbol{G}][\boldsymbol{G}]^{+}[\boldsymbol{G}] = [\boldsymbol{G}]$, $[\boldsymbol{G}]^{+}[\boldsymbol{G}][\boldsymbol{G}]^{+} = [\boldsymbol{G}]^{+}$, $([\boldsymbol{G}]^{+}[\boldsymbol{G}])^{\mathrm{T}} = [\boldsymbol{G}]^{+}[\boldsymbol{G}]$ and $([\boldsymbol{G}][\boldsymbol{G}]^{+})^{\mathrm{T}} = [\boldsymbol{G}][\boldsymbol{G}]^{+}$). Note that pseudo-inversion of a matrix in a size comparable with the degradation operator is computed here. This is in contrast with the conventional approaches in which manipulations of a huge $M_1M_2 \times N_1N_2$ matrix $\boldsymbol{H}$ are required. By intuition, a longer $\boldsymbol{b}$ may lead to a better restoration. However, experiments show that an excessively long filter does not further improve the restoration performance in the MSE sense, but generates images with a worse visual quality.

2.2 Non-separable case

Now we generalise the above discussion to cases in which both degradation and restoration operators are non-separable. Assume that the original image size is $N_1 \times N_2$, and the size of the degradation operator is

$L_1 \times L_2$. Thus the degradation is expressed as

$$\begin{bmatrix} \boldsymbol{g}_1 \\ \boldsymbol{g}_2 \\ \vdots \\ \boldsymbol{g}_{M_1} \end{bmatrix} = \begin{bmatrix} [\boldsymbol{F}]_{11} & [\boldsymbol{F}]_{12} & \cdots & [\boldsymbol{F}]_{1,L_1} \\ [\boldsymbol{F}]_{21} & [\boldsymbol{F}]_{22} & \cdots & [\boldsymbol{F}]_{2,L_1} \\ \vdots & \vdots & \ddots & \vdots \\ [\boldsymbol{F}]_{M_1,1} & [\boldsymbol{F}]_{M_1,2} & \cdots & [\boldsymbol{F}]_{M_1,L_1} \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{L_1} \\ \boldsymbol{h}_{L_1-1} \\ \vdots \\ \boldsymbol{h}_1 \end{bmatrix} \tag{17}$$

or more concisely

$$\boldsymbol{g} = [\boldsymbol{F}]\boldsymbol{h} \tag{18}$$

In these equations, the $M_1M_2 \times 1$ vector $\boldsymbol{g}$ is the degraded image stacked in a lexicographic order. $\boldsymbol{g}_{m_1}$ is an $M_2 \times 1$ vector representing the m_1th row

$$\boldsymbol{g}_{m_1} = [g_{m_1,1} \quad g_{m_1,2} \quad \cdots \quad g_{m_1,m_2} \quad \cdots \quad g_{m_1,M_2}]^{\mathrm{T}}, \quad m_1 = 1, 2, \ldots, M_1 \tag{19}$$

Similarly, $\boldsymbol{h}$, sized $L_1L_2 \times 1$, is the degradation vector, and $\boldsymbol{h}_{l_1}$ is an $L_2 \times 1$ vector

$$\boldsymbol{h}_{l_1} = [h_{l_1,L_2} \quad h_{l_1,L_2-1} \quad \cdots \quad h_{l_1,l_2} \quad \cdots \quad h_{l_1,1}]^{\mathrm{T}}, \quad l_1 = 1, 2, \ldots, L_1 \tag{20}$$

The entries in $\boldsymbol{g}$ are

$$\boldsymbol{g}_{m_1,m_2} = \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} f_{m_1+l_1-1,m_2+l_2-1} h_{L_1-l_1+1,L_2-l_2+1}, \quad m_1 = 1, 2, \ldots, M_1; \quad m_2 = 1, 2, \ldots, M_2 \tag{21}$$

where $M_i = N_i - L_i + 1$, $i = 1, 2$.

The 2D STP $[\boldsymbol{F}]$ is a block matrix composed of $M_1 \times L_1$ sub-matrices, denoted $[\boldsymbol{F}]_{m_1,l_1}$. It has the following shift property

$$\begin{aligned} [\boldsymbol{F}]_{12} &= [\boldsymbol{F}]_{21} \\ [\boldsymbol{F}]_{13} &= [\boldsymbol{F}]_{22} = [\boldsymbol{F}]_{31} \\ [\boldsymbol{F}]_{14} &= [\boldsymbol{F}]_{23} = [\boldsymbol{F}]_{32} = [\boldsymbol{F}]_{41} \end{aligned} \tag{22}$$

In (17), $[\boldsymbol{F}]_{m_1,l_1}$ is an $M_2 \times L_2$ matrix with the same shift property, in which the elements are pixels of the test image

$$[\boldsymbol{F}]_{m_1,l_1} = \begin{bmatrix} f_{m_1+l_1-1,1} & f_{m_1+l_1-1,2} & \cdots \\ f_{m_1+l_1-1,2} & f_{m_1+l_1-1,3} & \cdots \\ \vdots & \vdots & \ddots \\ f_{m_1+l_1-1,m_2} & f_{m_1+l_1-1,m_2+1} & \cdots \\ \vdots & \vdots & \ddots \\ f_{m_1+l_1-1,M_2} & f_{m_1+l_1-1,M_2+1} & \cdots \end{bmatrix} \begin{bmatrix} f_{m_1+l_1-1,l_2} & \cdots & f_{m_1+l_1-1,L_2} \\ f_{m_1+l_1-1,l_2+1} & \cdots & f_{m_1+l_1-1,L_2+1} \\ \vdots & \ddots & \vdots \\ f_{m_1+l_1-1,m_2+l_2-1} & \cdots & f_{m_1+l_1-1,m_2+L_2-1} \\ \vdots & \ddots & \vdots \\ f_{m_1+l_1-1,M_2+l_2-1} & \cdots & f_{m_1+l_1-1,M_2+L_2-1} \end{bmatrix} \tag{23}$$

For the same reason as stated in the previous sub-section, we choose a rectangle with four alternate black and white quadrants, denoted by a $J_1 \times J_2$ matrix $\boldsymbol{F}$, as the test image

$$\boldsymbol{F} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 & 1 & \ldots & 1 \\ 0 & 0 & \ldots & 0 & 1 & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 \end{bmatrix} \tag{24}$$

Convolving $\boldsymbol{F}$ with the degradation mask $\boldsymbol{H}$ ($L_1 \times L_2$) yields a 2D degraded image, $\boldsymbol{G}$. Thus, the DTP, $[\boldsymbol{G}]$, is obtained with the entries

$$G_{j,k} = g_{j_1+k_1-1,j_2+k_2-1}, \quad \begin{aligned} j &= (j_1 - 1)J_2 + j_2, \\ k &= (k_1 - 1)K_2 + k_2, \\ j_i &= 1, 2, \ldots, J_i, \quad j = 1, 2, \ldots, J_1J_2 \\ k_i &= 1, 2, \ldots, K_i, \quad k = 1, 2, \ldots, K_1K_2 \end{aligned} \tag{25}$$

A least-square minimum-norm restoration operator is expressed in the same form as in the separable case

$$\boldsymbol{b} = [\boldsymbol{G}]^{+}\boldsymbol{f}_0 \tag{26}$$

This is then rearranged to give a $K_1 \times K_2$ convolution kernel $\boldsymbol{B}$. For the sake of clarity in presentation, a caret on top of the letter $\boldsymbol{b}$ has been omitted. In this case, $\boldsymbol{f}_0$ is a vector version of the 'chess board' test image, having a length of J_1J_2, with the border pixels deleted. $\boldsymbol{b}$ is a $K_1K_2 \times 1$ vector

$$\boldsymbol{b} = [\boldsymbol{b}_{K_1} \quad \boldsymbol{b}_{K_1-1} \quad \cdots \quad \boldsymbol{b}_{k_1} \quad \cdots \quad \boldsymbol{b}_1]^{\mathrm{T}} \tag{27}$$

where the sub-vectors

$$\boldsymbol{b}_{k_1} = [b_{k_1,K_2} \quad b_{k_1,K_2-1} \quad \cdots \quad b_{k_1,k_2} \quad \cdots \quad b_{k_1,1}]^{\mathrm{T}} \tag{28}$$

Therefore the restoration problem is formally resolved.

3 SVD filter bank interpretation

It is clear that the operator obtained in the previous section is essentially an inverse filter. Due to the ill-posed nature of restoration problems, the estimate given in (26) is sensitive to noise. For a well-behaved solution, regularisation of the pseudo-inverse operation must be introduced. In this section, SVD is performed on a small-sized DTP, rather than on a much larger degradation matrix $\boldsymbol{H}$, leading to an SVD expansion of the restoration operator or an SVD filter bank. Some high-order sub-filters will be eliminated to achieve a smoothly restored image.

First, express the restoration filter by expanding the operator $\boldsymbol{b}$ into a spectrum of its SV components [24]

$$\boldsymbol{b} = \sum_{r=1}^{R} \frac{1}{\lambda_r} (\boldsymbol{v}_r \boldsymbol{u}_r^{\mathrm{T}}) \boldsymbol{f}_0 = \sum_{r=1}^{R} \frac{1}{\lambda_r} \boldsymbol{b}_r \tag{29}$$

Here $\boldsymbol{v}_r$ and $\boldsymbol{u}_r$ are the rth eigenvectors of $[\boldsymbol{G}]^{\mathrm{T}}[\boldsymbol{G}]$ and $[\boldsymbol{G}][\boldsymbol{G}]^{\mathrm{T}}$, respectively. R is the rank of $[\boldsymbol{G}]$. The λ_rs are the SVs of $[\boldsymbol{G}]$, arranged in a descending order with respect to r. We call $\boldsymbol{b}_r$s the SV components of $\boldsymbol{b}$. It can be shown that matrix $\boldsymbol{V}^{\mathrm{T}}([\boldsymbol{G}]^{\mathrm{T}}[\boldsymbol{G}])\boldsymbol{V}$ is diagonal where columns of $\boldsymbol{V}$ are $\boldsymbol{v}_r$s. Therefore $[\boldsymbol{G}]\boldsymbol{v}_r$ and $[\boldsymbol{G}]\boldsymbol{v}_{r'}$ ($r \neq r'$) are mutually orthogonal. From (29), the SV components

and corresponding $\boldsymbol{v}_r$s are linearly correlated, $[\boldsymbol{G}]\boldsymbol{b}_r$ are also mutually orthogonal.

The observed image may be viewed as a sum of a signal component $[\boldsymbol{S}]$ and a noise term $[\boldsymbol{N}]$. Both $[\boldsymbol{S}]$ and $[\boldsymbol{N}]$ are in the same form of the test data pattern. Therefore the estimate obtained from the restoration operation is expressed as

$$\hat{\boldsymbol{x}} = [\boldsymbol{S}]\boldsymbol{b} + [\boldsymbol{N}]\boldsymbol{b} = \hat{\boldsymbol{s}}_0 + \hat{\boldsymbol{n}}_0 \tag{30}$$

Both the signal and noise parts of the restored output can be expanded into SV spectra

$$\hat{\boldsymbol{s}}_0 = \sum_{r=1}^{R} \frac{1}{\lambda_r}[\boldsymbol{S}]\boldsymbol{b}_r = \sum_{r=1}^{R} \hat{\boldsymbol{s}}_{0r} \tag{31}$$

$$\hat{\boldsymbol{n}}_0 = \sum_{r=1}^{R} \frac{1}{\lambda_r}[\boldsymbol{N}]\boldsymbol{b}_r = \sum_{r=1}^{R} \hat{\boldsymbol{n}}_r \tag{32}$$

Equation (29) implies that the restoration operator is equivalent to a bank of filters, each associated with a weight equal to the reciprocal of the corresponding SV.

An example for the decomposition of a 1D restoration operator is given in Table 1. In this case, the degradation filter is assumed to have a Gaussian shape with $L = 5$

$$\boldsymbol{h} = [0.0856 \quad 0.2427 \quad 0.344 \quad 0.2427 \quad 0.0856]^{\mathrm{T}} \tag{33}$$

The length of $\boldsymbol{b}$ is chosen as $K = 7$, and the used test data sequence $\boldsymbol{f}$ is made up of ten zeros and 11 ones, which has a minimum length. Entries of the seven SV components $\boldsymbol{b}_r$, $r = 1, 2, \ldots, 7$, are listed in the table. Also listed in the bottom row are the numbers of zero-crossings in $\boldsymbol{b}_r$s. Fig. 1 gives the frequency domain representation of these sub-filter characteristics. In calculating the frequency response, a sampling frequency of 23.62 pixels/mm (600 dpi) was assumed. It is observed from both the table and the figure that $\boldsymbol{b}_1$ is a low-pass operator with all seven entries positive, whereas the others are band-pass/high-pass filters consisting of both positive and negative entries. The number of zero-crossings is a crucial feature for each sub-filter. It generally increases with respect to the index r. Consequently, main lobes of the frequency-responses move progressively towards high frequencies with increasing r (with a few exceptions in some high-order components when K is fairly large, as our numerical experiments indicated).

The frequency response of a 2D non-separable restoration filter bank with $K = 5$ is visualised in Fig. 2 in which dark areas indicate spectral humps. The degradation kernel is $\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}$ where $\boldsymbol{h}$ is given by (33). Again, the first sub-filter (top-left) is low-pass, and the rest are band-pass/high-pass with the passband generally moving towards the border and spreading over the frequency domain as r increases.

When acting on a degraded image, the individual sub-filters extract different frequency contents from the data. The outputs are then multiplied by $1/\lambda_r$, and added together to form the restored image. It is clear that, when passing the blurred data with additive noise through a filter bank as expressed in (30)–(32), noise is amplified by the large values of $1/\lambda_r$ in the high-frequency region. In this region, the data contain relatively less energy of the image. This is shown in Fig. 3 where the SVD spectra of the pure image data (solid), noise (dotted) and the noisy image (dashed) are plotted. A 64×64 block in the image 'Lena' sized 256×256 was used in the computation. The additive noise was Gaussian with a zero mean, and independent of the image. The signal-to-noise ratio (SNR) was 30 dB. A non-separable 9×9 restoration kernel was assumed so that $[G]$ has 81 SVs. As some SV components are quite small, making deep pits in the spectra, they have been removed in order to reduce the ruggedness of the curves. From the plot, it is observed that the signal power is basically concentrated in the region of small SV indices corresponding to large SVs and low frequencies. It quickly decays towards the high end (small SVs), while the decay of noise power is much slower. Considerable amount of noise power is distributed in the high SV index region, significantly exceeding the signal power.

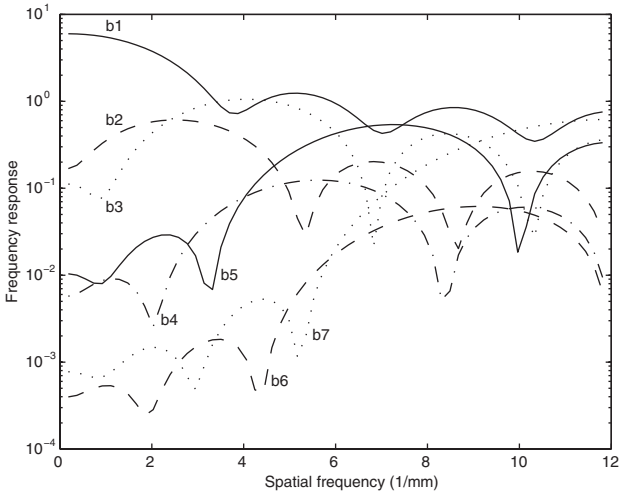

Fig. 1 *Frequency responses of the SV components of* $\boldsymbol{b}$

First SV sub-filter is low-pass, whereas the rest are band-pass/high-pass

Pass-bands move towards the high-frequency region as r increases

Table 1: Entries in SV decomposed sub-filters and numbers of zero-crossings

	$\boldsymbol{b}_1$	$\boldsymbol{b}_2$	$\boldsymbol{b}_3$	$\boldsymbol{b}_4$	$\boldsymbol{b}_5$	$\boldsymbol{b}_6$	$\boldsymbol{b}_7$
$\boldsymbol{b}_i(1)$	0.5044	0.1447	−0.2949	−0.0321	0.1049	0.0067	−0.0217
$\boldsymbol{b}_i(2)$	0.6558	0.1457	−0.1194	0.0135	−0.1455	−0.0164	0.0784
$\boldsymbol{b}_i(3)$	0.7903	0.1110	0.1156	0.0312	−0.0259	0.0144	−0.1427
$\boldsymbol{b}_i(4)$	0.9030	0.0473	0.2572	0.0051	0.1366	0.0018	0.1667
$\boldsymbol{b}_i(5)$	0.9908	−0.0294	0.2051	−0.0283	0.0110	−0.0162	−0.1309
$\boldsymbol{b}_i(6)$	1.0524	−0.1000	−0.0101	−0.0223	−0.1551	0.0152	0.0648
$\boldsymbol{b}_i(7)$	1.0896	−0.1510	−0.2654	0.0271	0.0845	−0.0051	−0.0154
Zero-crossings	0	1	2	3	4	5	6

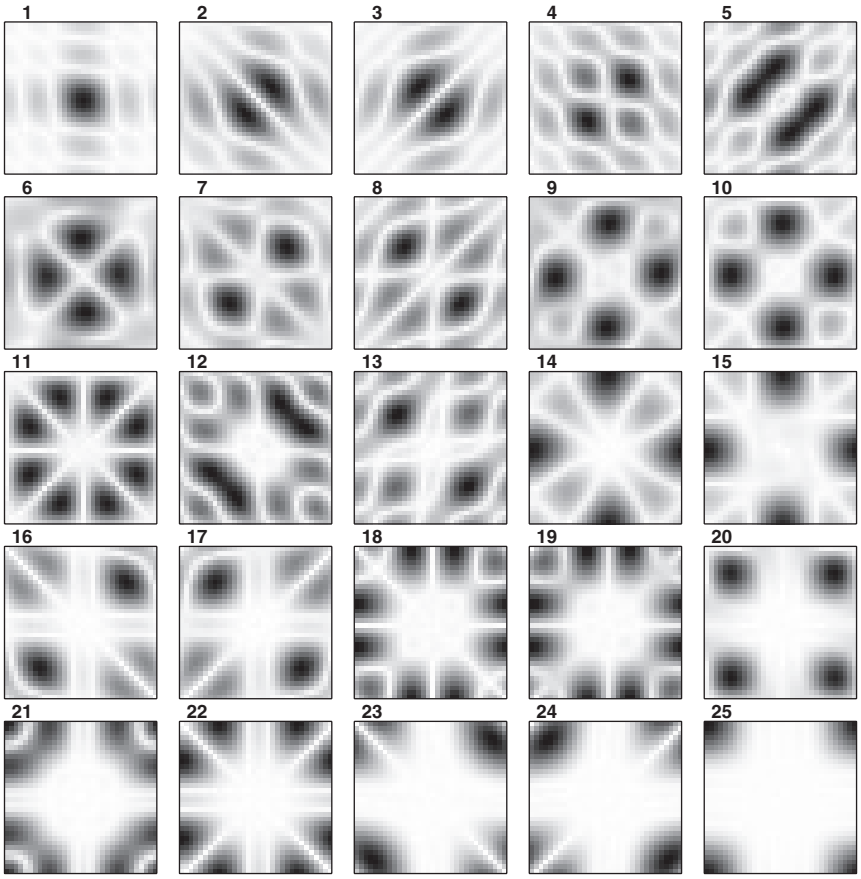

Fig. 2 *Frequency responses of non-separable SVD 2D sub-filters ($K = 5$)*

SV index r goes from 1 for the first image at the top-left, through 25 for the last at the bottom-right
First sub-filter is low-pass, whereas the rest are band-pass/high-pass, with spectral humps (dark areas) moving progressively from the centre (low frequency) towards the border (high frequency) with respect to r

4 Truncation of filter bank based on an energy criterion

In order to prevent the restored image from being corrupted by noise amplification, a TSVD technique can be used to drop some of the sub-filters corresponding to the smallest SVs. Analytical or empirical methods for choosing a truncating threshold in conventional SVD restoration approaches have been discussed in the literature [14, 21, 25]. But some was found to be inconsistent in different applications [26]. As most of these methods are not directly applicable to the present filter-bank formulation, we propose an SVD filter truncation technique that is based on an energy consideration. A sub-filter is switched off if energy of the output noise exceeds that of the output image component.

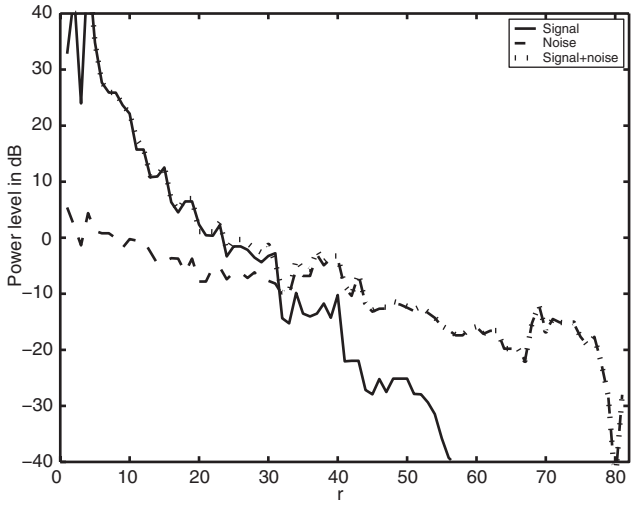

Fig. 3 *SV spectra of the signal and noise components in the observed image data (Lena, SNR = 30 dB)*

A 9×9 non-separable restoration kernel is used
Signal energy concentrates in the lower SV index region, and decays steeply towards the high end
In the high index region, noise dominates

4.1 Energy in the estimation error

Assume that only part of the sub-filters are used in the restoration, resulting in a partially restored image

$$\hat{\boldsymbol{x}}^{(\mathrm{T})} = \sum_{r=1}^{R} \frac{C_r}{\lambda_r} [\boldsymbol{Y}] \boldsymbol{b}_r = \sum_{r=1}^{R} \frac{C_r}{\lambda_r} \boldsymbol{B}_r \boldsymbol{y}, \quad C_r = 0 \text{ or } 1 \tag{34}$$

where $[\boldsymbol{Y}]$ and $\boldsymbol{B}_r$ are large matrices obtained from $\boldsymbol{y}$ and $\boldsymbol{b}_r$, respectively, and C_r switches the corresponding sub-filter $\boldsymbol{b}_r$ on or off.

Note that, in (34), the operator in a matrix form, $\boldsymbol{B}_r$, is only used for analysis. In practical computations, convolution of the image with a compact $L \times L$ kernel is conducted instead of the formidable matrix multiplication. Substituting (1) into (34)

$$\hat{\boldsymbol{x}}^{(\mathrm{T})} = \sum_{r=1}^{R} \frac{C_r}{\lambda_r} \boldsymbol{B}_r \boldsymbol{H} \boldsymbol{x} + \sum_{r=1}^{R} \frac{C_r}{\lambda_r} \boldsymbol{B}_r \boldsymbol{n} \tag{35}$$

If a good truncation, that is an appropriate choice of C_r values, is used in the TSVD, the restored image given by (34) or (35) should be close to that in the original image $\boldsymbol{x}$. In other words, a criterion for a good SVD truncation is obtainable by minimising the following parameter representing energy contained in the error

$$d = \boldsymbol{E}\left[\left(\hat{\boldsymbol{x}}^{(\mathrm{T})} - \boldsymbol{x}\right)^{\mathrm{T}}\left(\hat{\boldsymbol{x}}^{(\mathrm{T})} - \boldsymbol{x}\right)\right] \tag{36}$$

After some manipulation, (36) can be expressed as

$$d = \boldsymbol{E}\left\{\left[(\boldsymbol{BHx} - \boldsymbol{x}) - \boldsymbol{B}^{(0)}\boldsymbol{Hx} + \boldsymbol{B}^{(1)}\boldsymbol{n}\right]^{\mathrm{T}} \times \left[(\boldsymbol{BHx} - \boldsymbol{x}) - \boldsymbol{B}^{(0)}\boldsymbol{Hx} + \boldsymbol{B}^{(1)}\boldsymbol{n}\right]\right\} \tag{37}$$

where

$$\boldsymbol{B} = \sum_{r=1}^{R}\frac{1}{\lambda_r}\boldsymbol{B}_r, \quad \boldsymbol{B}^{(0)} = \sum_{C_r=0}\frac{1}{\lambda_r}\boldsymbol{B}_r, \quad \boldsymbol{B}^{(1)} = \sum_{C_r=1}\frac{1}{\lambda_r}\boldsymbol{B}_r \tag{38}$$

If noise is negligible, every sub-filter should be on, namely all C_r values should be equal to 1, and an ideal inverse filter $\boldsymbol{B}$ is used resulting in a nearly perfect restoration

$$\boldsymbol{E}\left[(\boldsymbol{BHx} - \boldsymbol{x})^{\mathrm{T}}(\boldsymbol{BHx} - \boldsymbol{x})\right] \simeq 0 \tag{39}$$

If, on the other hand, the image is contaminated by zero mean noise, assumimg noise is uncorrelated with $\boldsymbol{Hx}$ and $\boldsymbol{B}_r$

$$\boldsymbol{E}\left[(\boldsymbol{B}_r\boldsymbol{Hx})^{\mathrm{T}}(\boldsymbol{B}_{r'}\boldsymbol{n})\right] = 0 \tag{40}$$

This is true for both $r = r'$ and $r \neq r'$. Equation (37) then becomes

$$d = \boldsymbol{E}\left[\left(\boldsymbol{B}^{(0)}\boldsymbol{Hx}\right)^{\mathrm{T}}\left(\boldsymbol{B}^{(0)}\boldsymbol{Hx}\right)\right] + \boldsymbol{E}\left[\left(\boldsymbol{B}^{(1)}\boldsymbol{n}\right)^{\mathrm{T}}\left(\boldsymbol{B}^{(1)}\boldsymbol{n}\right)\right] \tag{41}$$

From (29), the restoration sub-filters $\boldsymbol{b}_r$s are low-pass, band-pass and high-pass, respectively. They remove correlation between columns of $[\boldsymbol{G}]$. Similarly, SVD filters can substantially reduce correlation in $\boldsymbol{Hx}$ and $\boldsymbol{n}$. Thus, $[(\boldsymbol{B}_r\boldsymbol{Hx})^{\mathrm{T}}(\boldsymbol{B}_{r'}\boldsymbol{Hx})]$ and $[(\boldsymbol{B}_r\boldsymbol{n})^{\mathrm{T}}(\boldsymbol{B}_{r'}\boldsymbol{n})]$ are much smaller than $[(\boldsymbol{B}_r\boldsymbol{Hx})^{\mathrm{T}}(\boldsymbol{B}_r\boldsymbol{Hx})] + [(\boldsymbol{B}_{r'}\boldsymbol{Hx})^{\mathrm{T}}(\boldsymbol{B}_{r'}\boldsymbol{Hx})]$ and $[(\boldsymbol{B}_r\boldsymbol{n})^{\mathrm{T}}(\boldsymbol{B}_r\boldsymbol{n})] + [(\boldsymbol{B}_{r'}\boldsymbol{n})^{\mathrm{T}}(\boldsymbol{B}_{r'}\boldsymbol{n})]$, respectively. Therefore

$$d \simeq \sum_{C_r=0}\left\|\frac{1}{\lambda_r}\boldsymbol{B}_r\boldsymbol{Hx}\right\|^2 + \sum_{C_r=0}\left\|\frac{1}{\lambda_r}\boldsymbol{B}_r\boldsymbol{n}\right\|^2 \tag{42}$$

Having obtained the expression of error energy as a sum of signal and noise contributions, a method of SVD filter bank truncation that minimises the noise energy can be developed as follows.

4.2 Estimation of noise energy

When present, noise usually dominates in high frequency bands as shown in Fig. 3. In this case, output from the highest SVD filters may be considered as purely contributed by noise. Consider the following three cases:

1. Noise with low energy: output energy from several sub-filters of the highest orders is less than a predefined threshold T_1.
2. Noise with a roughly flat spectrum: output energy from the highest order sub-filters is greater than T_1, but its variation across these filters is less than a second threshold, T_2.
3. Noise with uneven spectrum: variation of the output noise energy is greater than T_2.

Note that the range of highest-order filters used for noise classification should be well within the noise dominant region of the SV spectrum.

According to the above discussion, the output noise energy from the highest-order SVD filters can be estimated as

$$\hat{\boldsymbol{E}}_r^{(n)} = \|[\boldsymbol{N}]\boldsymbol{v}_r\|^2 \simeq \|[\boldsymbol{Y}]\boldsymbol{v}_r\|^2 = \frac{\|[\boldsymbol{Y}]\boldsymbol{b}_r\|^2}{\|\boldsymbol{b}_r\|^2} = \frac{\|\boldsymbol{B}_r\boldsymbol{y}\|^2}{\|\boldsymbol{b}_r\|^2} \tag{43}$$

When noise energy is low, $\hat{\boldsymbol{E}}_r^{(n)}$ from each SVD filter is equal to $\hat{\boldsymbol{E}}_R^{(n)}$, where R refers to the highest order. If noise is relatively high, an effective range of r in which noise dominates the output from $\boldsymbol{b}_r$ can be found empirically based on $\hat{\boldsymbol{E}}_R^{(n)}$. This range is denoted R_A. In general, the larger the value of $\hat{\boldsymbol{E}}_R^{(n)}$, the wider the range R_A.

If the noise spectrum is uneven, the power spectrum is modelled as Gaussian. In a separable case, the noise energy in the sub-filters can be arranged as a $K \times K$ matrix, whereas in a non-separable case they are arranged as an $R \times 1$ vector

$$\hat{\boldsymbol{E}}^{(n)} = \left[\hat{E}_1^{(n)} \quad \hat{E}_2^{(n)} \quad \cdots \quad \hat{E}_r^{(n)} \quad \cdots \quad \hat{E}_R^{(n)}\right]^{\mathrm{T}} \tag{44}$$

where $R = K^2$, and

$$\hat{\boldsymbol{E}}_r^{(n)} = \alpha \cdot \exp\left[\frac{-(r-1)^2}{2\sigma^2}\right], \quad r = 1, 2, \ldots, R \tag{45}$$

α and σ are chosen to make the Gaussian template most closely approximate the measured output noise power spectrum within the range R_A.

In case the noise spectrum is even, all $\hat{\boldsymbol{E}}_r^{(n)}$ values are equal

$$\hat{\boldsymbol{E}}_r^{(n)} = \underset{\rho \in R_A}{\mathrm{mean}}\left(\|[\boldsymbol{Y}]\boldsymbol{v}_\rho\|^2\right) = \underset{\rho \in R_A}{\mathrm{mean}}\left(\frac{\|\boldsymbol{B}_\rho\boldsymbol{y}\|^2}{\|\boldsymbol{b}_\rho\|^2}\right), \quad r = 1, 2, \ldots, R \tag{46}$$

The last equation is true because $\boldsymbol{v}_r = \boldsymbol{b}_r/\|\boldsymbol{b}_r\|$ as can been seen from (29).

4.3 Truncation of SVD filters

By using (40), the total energy output from the rth sub-filter can be expressed as

$$\boldsymbol{E}_r^{(y)} = \|[\boldsymbol{Y}]\boldsymbol{v}_r\|^2 = \frac{\|\boldsymbol{B}_r\boldsymbol{y}\|^2}{\|\boldsymbol{b}_r\|^2} = \frac{\|\boldsymbol{B}_r\boldsymbol{Hx}\|^2}{\|\boldsymbol{b}_r\|^2} + \boldsymbol{E}_r^{(n)} \tag{47}$$

Thus, a criterion for truncation of SVD filter bank that minimises the energy contained in restoration error in the form of a C_r pattern as defined in (34) is obtained

$$C_r = \begin{cases} 1 & \boldsymbol{E}_r^{(y)} > 2\hat{\boldsymbol{E}}_r^{(n)} \\ 0 & \boldsymbol{E}_r^{(y)} \leq 2\hat{\boldsymbol{E}}_r^{(n)} \end{cases} \tag{48}$$

5 Numerical results

In the numerical experiments, restoration results from blurred and noise-contaminated images were obtained

using the proposed approach. The test images 'Lena' and 'Cameraman' were blurred with a separable 5×5 kernel $\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}$ where $\boldsymbol{h}$ is given in (33), and additive noise of different strength was introduced. Peak SNR (PSNR) is used in this work to evaluate the effectiveness of image restoration operations, in which energy contained in the estimate error is measured referenced to the peak image energy that is 256^2 multiplied by the total number of pixels

$$\mathrm{PSNR} = 10 \log_{10} \frac{256^2 N_1 N_2}{\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|^2} \tag{49}$$

where $\boldsymbol{x}$ is an ideal image, $\tilde{\boldsymbol{x}}$ a degraded or restored image and N_1 and N_2 are the numbers of rows and columns of the ideal image, respectively.

In the first test, a degraded 'Lena' with PSNR = 27.9 dB was slightly contaminated by additive noise. The peak signal-to-additive-noise-ratio was 68 dB. After restoration using the low-energy noise model, PSNR became 35.5 dB. As the noise level was very low, only 11 SV sub-filters among 81 were turned off in this case representing some highest order sub-filters.

Figs. 4 and 5 present restoration results from images degraded with the same separable blurring filter, and then contaminated by additive white Gaussian noise (AWGN) at a peak signal-to-additive-noise-ratio of 44 dB. When the even spectrum noise is identified, patterns of C_r values indicating on–off status of the SVD sub-filters used in the restoration are given in Table 2. It can be seen from these patterns that more sub-filters were switched on for 'Cameraman' as it had more high-frequency content than 'Lena'. PSNR in the degraded and restored 'Lena' were 27.90 and 29.93 dB, respectively, whereas that in 'Cameraman' were 25.13 and 27.30 dB, respectively. Improvements of more than 2 dB were obtained in both cases.

As the patterns of C_r status given in Table 2 are optimised in the sense of error energy minimisation, any deviation from the desirable pattern would provide no significant improvements to the restored images. This is verified as shown in Tables 3 and 4 for the two tested images. Each listed value represents a PSNR of the restored image obtained by using the optimised TSVD filter bank except for a flipped C_r status at the corresponding position. It is observed that the flips did not produce a higher PSNR, with only one exception in 'Lena' as marked with an asterisk in the table. The discrepancy, which is only 0.14 dB, should be attributed to the fact that $\boldsymbol{B}_r\boldsymbol{H}\boldsymbol{x}$ and $\boldsymbol{B}_r\boldsymbol{n}$ are not rigorously orthogonal.

If the degraded 'Lena' and 'Cameraman' are restored using a Wiener filter, the resulting PSNR values are, respectively, 30.81 and 28.11 dB, higher than the results obtained using the TSVD method. With the Wiener filter, however, power spectra of both the ideal image and noise must be available. If, as is usually the case, power spectrum density of the ideal image is unknown and therefore can

Fig. 4 *Restoration of a blurred and noise-contaminated image (Lena 256 × 256)*

Blurring filter is $\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}$ where $\boldsymbol{h}$ is defined by (33)
AWGN is present with PSNR = 44 dB
PSNR values in the degraded and restored images are 27.90 and 29.93 dB, respectively
a Original
b Degraded
c Restored

Fig. 5 *Restoration of a blurred and noise-contaminated image (Cameraman)*

Blurring filter and noise are the same as in Fig. 4
PSNR values in the degraded and restored images are 25.13 and 27.30 dB, respectively
a Original
b Degraded
c Restored

Table 2: C_r patterns: values of C_r used in restoration of degraded 'Lena' (top) and 'Cameraman' (bottom)

	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	0	0
2	1	1	1	1	1	1	0	0	0
3	1	1	1	1	1	1	0	0	0
4	1	1	1	1	1	0	0	0	0
5	1	1	1	1	0	0	0	0	0
6	1	1	1	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0

	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	0	0
2	1	1	1	1	1	1	1	0	0
3	1	1	1	1	1	1	0	0	0
4	1	1	1	1	1	0	0	0	0
5	1	1	1	1	1	0	0	0	0
6	1	1	1	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0

Numbers 1–9 indicate sub-filter orders in the horizontal and vertical directions, respectively

only be estimated from the degraded image, the performance of the Wiener filter becomes inferior to the TSVD method, which does not require the power spectrum density of the ideal image. Experiments have shown that PSNR of the Wiener filtered images becomes 29.62 and 26.72 dB, respectively, when using approximate power spectrum estimated from the observed image, which are lower than the TSVD results.

When the blurred 'Lena' was contaminated by noise having an uneven spectrum with PSNR = 48 dB, the optimised C_r pattern as listed in Table 5 was used. This led to an improvement of PSNR from 27.95 dB in the degraded image to 30.59 dB in the restored image. PSNR values resulting from C_r flips are listed in Table 6, in which one can see no further improvement in the restoration quality in terms of PSNR. Using a Wiener filter with a power spectrum of both the ideal image and noise available, PSNR of the restored image was 31.17 dB. By estimating the image power spectrum from the degraded observation, however, PSNR of the restored image obtained with the Wiener filter became 29.83 dB, again lower than the result of the proposed TSVD method.

Different sizes of the restoration operator have been tested. Generally speaking, the resulting PSNR does not change substantially over a fairly wide range of the K value, for example $K = 5–13$ when $L = 5$. A further increase of K does not significantly reduce the error and, a larger K also requires more computations. Experiments show that an operator equal to, or moderately larger than, the degradation filter, for example $K = 5–9$ for $L = 5$, usually gives a satisfactory result. If too small an operator is used, sharpness of the restored

Table 3: PSNR of restored 'Lena' with C_r values flipped at critical positions in the pattern

	1	2	3	4	5	6	7	8	9
1							29.83	29.93	
2						29.93	29.93		
3						29.93	29.30		
4					29.93	29.93			
5				29.93	30.07*				
6	29.93	29.93	29.93	29.93					
7	29.65	29.93	29.12						
8									
9									

Table 4: PSNR of restored 'Cameraman' with C_r values flipped at critical positions in the pattern

	1	2	3	4	5	6	7	8	9
1							26.99	27.30	
2							27.30	27.30	
3						27.30	27.07		
4					27.30	27.30			
5				27.30	27.21	27.30			
6		27.30	27.30	27.30	27.30				
7	26.99	27.30	26.84						
8	27.30								
9									

image will be sacrificed, whereas a large operator tends to require higher computation complexity.

The present method has also been compared experimentally with a previous published work in which a more complicated criterion was used [24]. The image 'Lena' was blurred with the same separable 5×5 kernel $\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}$, and AWGN with an SNR of 50 dB was added, resulting in PSNR = 27.98 dB of the degraded image. With the present approach, PSNR of the restored images were 31.87 dB, 0.98 dB higher than the previous method. When the SNR of the AWGN was 30 dB, PSNR in the observed and restored images were 27.51 and 28.85 dB respectively. The latter is 0.51 dB better than that of the work of Wang and Zhang [24].

Table 5: C_r patterns used for restoration of blurred 'Lena' with additive colour noise

	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	0	0
2	1	1	1	1	1	1	1	0	0
3	1	1	1	1	1	1	0	0	0
4	1	1	1	1	1	1	0	0	0
5	1	1	1	1	1	0	0	0	0
6	1	1	1	1	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0

The proposed method is also applicable to the reconstruction of images damaged by different types of process other than blurring, provided that the degradation can be modelled as a convolution kernel. Experiments have been conducted in which some high-passed images with additive noise are satisfactorily restored. Fig. 6 shows an example in which 'Lena' is sharpened with a separable 5×5 kernel $\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}$ where $\boldsymbol{h} = [0.1 \quad -0.3 \quad 1.4 \quad -0.3 \quad 0.1]^{\mathrm{T}}$ and AWGN of the same level as used in the experiment of Fig. 4 is added. PSNR of the observed and restored images are 30.83 dB and 33.90 dB, respectively.

6 Concluding remarks

By using a pair of known test data patterns as input and output, respectively, a small-sized image restoration operator is identified through SVD. In conventional SVD restoration approaches, matrices to be manipulated have sizes of $M^2 \times M^2$ (assuming a square matrix). Performing SVD on such a large matrix requires formidable computation: the number of operations is in the order of M^6. With the proposed approach, on the other hand, SVD is performed on a matrix with a size substantially smaller than the image. The size of matrices under consideration is only related to L and K, the sizes of the degradation and restoration operators, respectively, and is independent of the image size. For a non-separable operator, $[\boldsymbol{G}]$ has a size $(K+L-1)^2 \times K^2$. Assuming $L = 5$, $K = 9$ and the image is 256×256, for example, $[\boldsymbol{G}]$ is only $1/10^6$ of the huge matrix $\boldsymbol{H}$ as in (1). Therefore the ratio between the numbers of operations required by conventional SVD restoration methods and

Table 6: PSNR of restored 'Lena', which was blurred and contaminated by additive colour noise, with C_r values flipped at critical positions in the pattern

	1	2	3	4	5	6	7	8	9
1							30.26	30.59	
2							30.59	30.59	
3						30.59	30.47		
4						30.59	30.59		
5					30.21	30.59			
6	30.59	30.59	30.59	30.59	30.59				
7	30.48	30.59	30.35	30.59					
8									
9									

a *b*

Fig. 6 *Restoration of a sharpened and noise-contaminated image (Lena, PSNR = 30.83 dB)*

PSNR of the restored image is 33.90 dB
a Observed
b Restored

the proposed approach is 10^9, much higher than is obtained by a scanning SVD method [22]. The advantage increases by almost one order of magnitude if the image area is doubled.

The proposed SVD expansion of the restoration operator makes possible to partially separate noise contributions from useful information buried in the observed data. The operator is viewed as a bank of sub-filters with pass-bands ranging from DC to high frequencies, extracting different contents over the entire frequency range. Outputs from the sub-filters form a spectrum of the restored image components. After estimating the noise spectrum of SVD filter outputs, an optimised truncation is obtained.

An operator much larger than the degradation filter is not recommended, because it requires excessive computation resources but does not offer meaningful improvement of the restoration quality. Numerical experiments show that an operator equal to or moderately larger than the blurring kernel usually provides favourable results.

Having obtained the operator, image restoration is implemented in the space domain by convolution. Image convolution with a small operator can be made very efficient, whether by software or hardware, therefore is desirable in many applications.

7 Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 60372090 and 60502039), Key Project of Shanghai Municipality for Basic Research (No. 04JC14037) and Shanghai Leading Academic Discipline Project (No. T0102).

8 References

1 Andrews, H., and Hunt, B.: 'Digital image restoration' (Prentice-Hall, 1977)
2 Karayiannis, N.B., and Venetsanopoulos, A.N.: 'Regularization theory in image restoration – the stabilizing functional approach', *IEEE Trans. Signal Process.*, 1990, **38**, pp. 1155–1179
3 Thompson, A.M., Brown, J.C., Kay, J.W., and Titterington, M.: 'A study of methods of choosing the smoothing parameter in image restoration by regularization', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1991, **13**, pp. 326–338
4 Galatsanos, N.P., and Katsaggelos, A.K.: 'Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation', *IEEE Trans. Image Process.*, 1992, **1**, pp. 322–336
5 Kang, M.G., and Katsaggelos, A.K.: 'General choice of the regularization functional in regularized image restoration', *IEEE Trans. Image Process.*, 1995, **4**, pp. 594–602

6 Taratorin, A.M., and Sideman, S.: 'Constrained regularized image restoration using projection on a set of regularized solutions', *IEEE Trans. Signal Process.*, 1996, **44**, pp. 1547–1548
7 Rajagopalan, A.N., and Chaudhuri, S.: 'A variational approach to recovering depth from defocused images', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1997, **19**, pp. 1158–1164
8 Tikhonov, A., and Arsenin, V.: 'Solutions of ill-posed problems' (Winston, Washington, DC, 1977)
9 Meer, P., Jolion, J., and Rosenfeld, A.: 'A fast algorithm for blind estimation of noise variance', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990, **12**, pp. 216–223
10 Moayeri, N., and Konstantinides, K.: 'Blind restoration of blurred and noisy images'. Proc. ICASSP-97, 1997, vol. 4, pp. 2573–2576
11 Reeves, S.J., and Higdon, A.C.: 'Perceptual evaluation of the mean-square error choice of regularization parameter', *IEEE Trans. Image Process.*, 1995, **4**, pp. 107–110
12 Nguyen, N., Milanfar, P., and Golub, G.: 'Efficient generalized cross-validation with application to parametric image restoration and resolution enhancement', *IEEE Trans. Image Process.*, 2001, **10**, pp. 1299–1308
13 Chen, T., and Wu, H.R.: 'Recursive implementation of constrained LMS L-filters for image restoration', *Signal Process.*, 2001, **81**, pp. 1101–1107
14 Shen, J., Erdogan, H., and Ebbini, E.S.: 'An optimal operator design technique for coded excitation ultrasound imaging system'. Proc. IEEE Ultrasonics Symp., 1994, vol. 3, pp. 1777–1781
15 Sullivan, B.J., and Liu, B.: 'On the use of singular value decomposition and decimation in discrete time band-limited signal extrapolation', *IEEE Trans. Acoust. Speech Signal Process.*, 1984, **ASSP-32**, pp. 1201–1212
16 Scharf, L.L., and Tuft, D.W.: 'Rank reduction for modeling stationary signals', *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **ASSP-35**, pp. 350–355
17 Hansen, P.C.: 'Truncated singular value decomposition solution to discrete ill-posed problems with ill-determined numerical rank', *SIAM J. Sci. Stat. Comput*, 1990, **11**, pp. 503–518
18 Rushforth, K., Crawford, A.E., and Zhou, Y.: 'Least-square reconstruction of objects with missing high-frequency components', *J. Opt. Soc. Am.*, 1982, **72**, pp. 24–35
19 Maeda, J., and Murata, K.: 'Restoration of band-limited images by an iterative regularized pseudo-inverse method', *J. Opt. Soc. Am. A*, 1984, **1**, pp. 28–34
20 Hui, S.K., and Er, M.H.: 'Adaptive recursive algorithm for image restoration and array processing'. Proc. ICASSP-90, 1990, vol. 5, pp. 2859–2862
21 Fish, D.A., Grochmalicki, J., and Pike, E.P.: 'Scanning singular-value-decomposition method for restoration of images with space-variant blur', *J. Opt. Soc. Am. A*, 1996, **13**, pp. 464–469
22 Schutten, R.W., and Vermeij, G.F.: 'The approximation of image blur restoration filter by finite impulse response', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1980, **PAMI-2**, pp. 176–180
23 Reichenbach, S.E., and Park, S.K.: 'Small convolution kernels for high-fidelity image restoration', *IEEE Trans. Signal Process.*, 1991, **39**, pp. 2263–2274
24 Wang, S., and Zhang, X.: 'Least square convolution mask for image restoration'. Proc. Int. Congress of Image Science 2002, (ICIS'02), Tokyo, Japan, 2002, pp. 688–689 (*J. Imaging Soc. Japan*, 2002, **41**, (4), pp. 392–397)
25 Sono, A.: 'Optimally regularized inverse of singular value decomposition and application to signal extrapolation', *Signal Process.*, 1993, **30**, pp. 163–176
26 Gorodnitsky, I.F., and Rao, B.D.: 'Analysis of error produced by truncated SVD and Tikhnov regularization methods'. Conf. Record of the 28th Asilomar Conf. on Signal, Systems and Computers, 1994, vol. 1, pp. 25–29