Subspace Techniques for Image Understanding

Hamid K. Aghajan Thomas Kailath

Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305

Abstract

A unifying framework for solving several problems in image understanding is proposed based on the exploitation of subspace fitting techniques originally developed for antenna array processing applications. The SLIDE (Subspace-based LIne DEtection) algorithm is a framework for introducing the concept of subspace fitting into problems in image understanding and computer vision. The basic formulation of SLIDE provides a solution to the problem of multiple line fitting in binary or gray-scale images. However, as is shown in this paper, its modified versions can be applied to several other image analysis problems.

1 Introduction

An interesting relationship can be established between image analysis and the notion of subspace fitting by thinking of the image under consideration in a wave propagation context in which the digitized image is regarded as a snapshot of a spatial wave field at a fixed instant in time. This formulation leads to an efficient solution to several image understanding problems by exploiting ideas and tools from sensor array processing, communications theory, and projection-based transformations.

SLIDE is a model-based algorithm based on the partitioning of the induced measurement space into a signal subspace that is defined by the few desired parameters, and a noise subspace that includes all the undesired contributions. The signal subspace is determined by exploiting the spatial (and/or temporal) coherency that exists between contributions of the desired components, *e.g.* straight lines, in the image.

After establishing the framework for multiple straight line fitting applications, it is possible to extend the coherency accumulating notions of *SLIDE* to other vision applications by exploiting such tools of communications theory as chirping/dechirping and modulation. Circle and ellipse fitting, uniform motion estimation, linewidth measurement and alignment in microlithography for manufacturing of integrated circuits, skew detection in text analysis, estimation of the location of the axis of symmetry, and determining the focus of expansion in motion sequences are examples of such extensions. In this paper, the fundamental form of the *SLIDE* algorithm will be briefly introduced first, followed by its application to several of the mentioned application areas.

2 The *SLIDE* Algorithm

Starting with a simplified case of fitting a straight line to a binary image with a set of colinear pixels, let us imagine that there is an array of sensors in front of the vertical axis of the image. A simple sketch of such arrangement is shown in Figure 1. If we now consider the straight line to be the wavefront of a propagating wave, the measurements received at the sensors will have the form:

$$z(y) = e^{-j\mu x} = e^{-j\mu x_0} e^{j\mu y \tan\theta} \tag{1}$$

where μ is a constant parameter, and can be interpreted as the speed of propagation.

In our model, μ is free to choose, and its choice gives us a handle to develop different applications. The measurements $z(y)$ have the form of a complex sinusoid with a frequency related to the line angle θ; the line offset has been separated and encoded in a constant complex number. The varying part of the measurements, *i.e.* the part related to θ, can be lumped into a term $a_y(\theta) = e^{j\mu y \tan\theta}$, and be called the array response vector. The above formulation readily generalizes to the multiple line case,

$$\begin{aligned} z(y) &= \sum_{k=1}^{d} e^{j\mu y \tan\theta_k} e^{-j\mu x_{0_k}} + n(y) && (2) \\ &= \sum_{k=1}^{d} a_y(\theta_k) s_k + n(y). && (3) \end{aligned}$$

1058-6393/95 $4.00 © 1995 IEEE

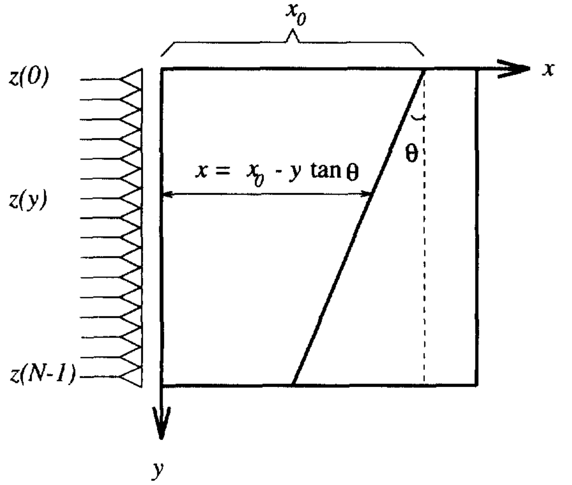

Figure 1: Image matrix and *hypothetical* sensors.

This equation is the starting point of extensive research in the last decade on the so-called subspace-based high resolution direction finding (and signal copy) algorithms (known as MUSIC, ESPRIT, WSF, etc., see *e.g.* [3]). In these methods, a sample covariance matrix is computed in a certain way from the measurements, and its eigendecomposition is examined. The basic concept of subspace fitting is that the d dominant eigenvectors of this covariance matrix span the same subspace that is spanned by the array response vectors for the desired angles. *SLIDE* uses efficient numerical methods for solving this problem, and yields high resolution estimates for the line angles. There is no search procedure involved in the implementation of *SLIDE* , and its computational complexity is an order of magnitude less than that of the traditional and search-based techniques such as the Hough transform method.

The next step would be to estimate the line offsets. It can be shown that by modifying the propagation scenario, new measurements are obtained that contain chirp (quadratic frequency) contributions from the angles, and linear contributions from the offsets. In other words, the offsets will be encoded as frequencies of complex sinusoids. Dividing the *chirped* measurements $z(y)$ by the array response $a_y(\theta)$ results in new *dechirped* measurements $w(y)$

$$w(y) = \frac{z(y)}{a_y(\theta)} = \sum_{k=1}^{d} e^{-j\alpha y x_{0_k}} + n'(y) \tag{4}$$

on which fast high resolution spectral estimation methods ('dual' to the previously mentioned array processing methods) can be applied to obtain estimates of the line offsets. The above formulation also generalizes to gray-scale images by assigning the value of the gradient at each pixel to the amplitude of the wave emanating from it. Details of the implementation of the *SLIDE* algorithm can be found in [1, 2].

3 Uniform 2D Motion Estimation

An interesting application of the *SLIDE* algorithm is in finding the velocities and positions of multiple moving objects in a frame sequence. Motion estimation is an important problem in computer vision, and over the years, several classes of techniques have been developed by researchers for handling this problem [4]. Specification of *SLIDE* to 2D motion estimation results from making the following basic observation.

If the projections of the frames on one of the main (horizontal or vertical) axes are stacked up to produce a synthetic image in t and x, say, then the moving patterns in the original sequence will be represented by skewed bands in this image. Figure 2 shows a simple case, where there are two moving objects and a stationary object in the sequence. Propagating the image parallel to the x-axis produces measurements $z(y)$ that encode the x components of the velocities of the moving patterns as frequencies of complex sinusoids:

$$z(t) = e^{-j\mu v_x t} e^{-j\mu x_0} Mx(\mu). \tag{5}$$

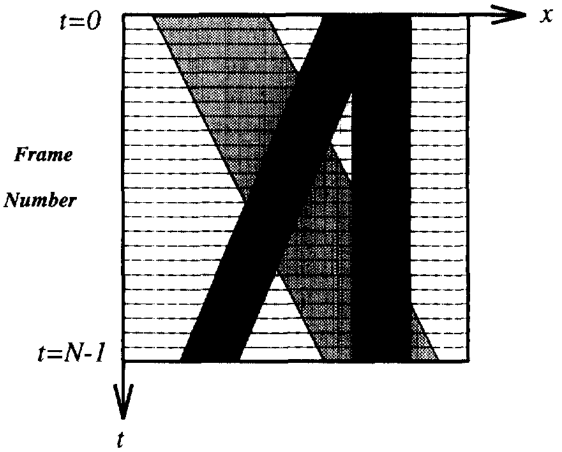

Figure 2: Stack of the frame projections.

The difference in this case would be the addition of a constant number as the amplitude of each sinusoid. This constant is the value of the Fourier transform of the projection of the pattern evaluated at a fixed frequency point. A similar approach would result in the estimation of the positions of the moving

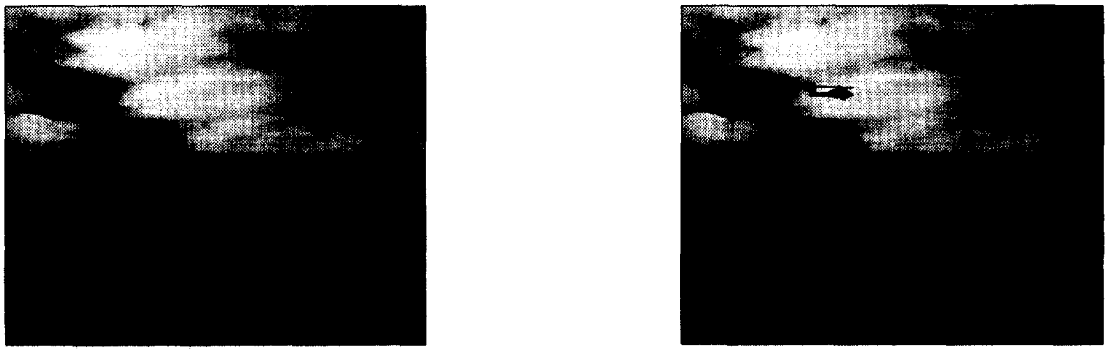

Figure 3: First and last frames of the test sequence.

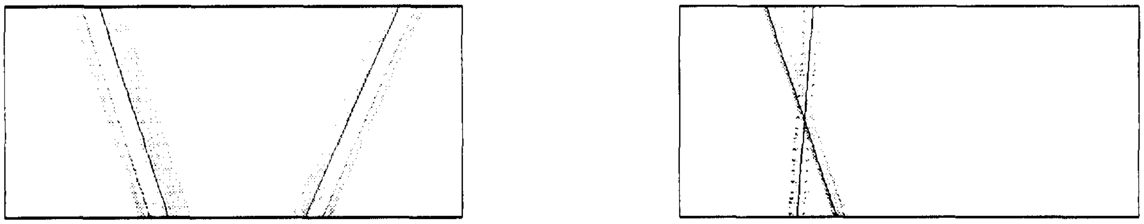

Figure 4: Stack of projections on the x and y axes, and detected motions.

patterns. By decomposing the two-dimensional problem of finding the velocities into two independent one-dimensional problems, *SLIDE* achieves superiority in the computational cost over the techniques that use 2D search. Another advantage of *SLIDE* is that it can handle multiple moving objects in the same frame sequence, whereas most other methods have to narrow down their working window such that only one moving pattern appears in it. *SLIDE* exploits the coherency that exists between the location of pixels corresponding to the moving patterns in the temporal and spatial directions, to enhance and distinguish a signal subspace that contains the desired parameters.

In Figure 3, the first and last frames of a sequence are shown, in which two helicopters move and the background is stationary. The goal is to find the velocities and the positions of the helicopters.

Figure 4 shows the synthetic images obtained by stacking the projections of the frames on the x and y axes. The derivative of the projections in time is then taken by simply subtracting each frame projection from its neighboring projection. As we can observe, each helicopter is represented by a skewed band in these images. Then, the algorithm is applied to these synthetic images independently, and estimates of the velocities and locations of the center of gravity of the moving objects are obtained. These estimates have been used to plot the trajectories of the objects.

4 Text Skew Detection

Another application of the *SLIDE* algorithm is the estimation of the skew angle in scanned text images. Skew occurs when pages of text are printed, copied, or scanned, and its detection is an important step in optical character reading tasks or in computerized filing systems [5, 6]. Although the lines of text contribute to more than one pixel per image row, and the character pixels introduce fuzziness on the lines, the estimate obtained by *SLIDE* is accurate enough for pursuing the other steps of processing the text. An example of applying *SLIDE* to a scanned text image is shown in Figure 5. The original image has been scanned with some amount of skew. The estimated value of the skew has been used to rotate the text image back to the proper orientation.

The robustness of the algorithm to model violations is also evident from this example. The broken lines, formulas, and the diagram give rise to contributions in the measurements that do not satisfy the assumptions made on the structure of the lines in an image. However, *SLIDE* provides reasonable and accurate estimates of the skew angle. This robustness stems from the fact that *SLIDE* exploits the great amount of coherency that exists in the text images between the locations of the pixels on parallel lines. This coherency contributes to enhancing the signal subspace component of the space spanned by the sample covariance

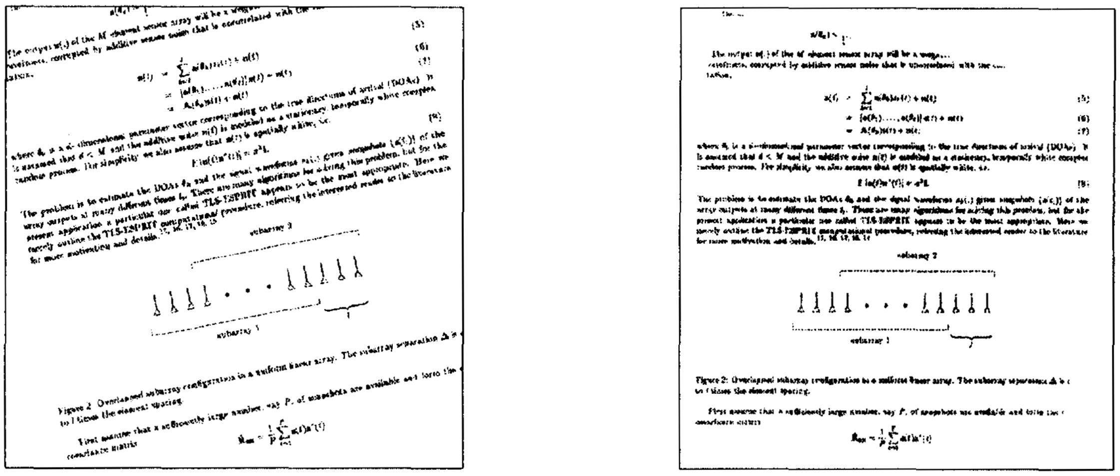

Figure 5: Scanned text image with a skew and the skew-compensated image.

matrix.

5 Axis of Symmetry

By exploiting the tools of communication theory, other application areas can be added to the framework of the *SLIDE* algorithm. One such tool is signal modulation, in which a passband signal can be represented by the location of a central carrier frequency, surrounded by two sidebands. A particular case of interest in this application category is the estimation of the axis of symmetry of patterns. This problem is an important image understanding problem in its own right; however, an interesting practical case of the problem exists in microlithography. Symmetric alignment patterns are etched on the wafer in consecutive steps, and the alignment task is to bring together the axes of symmetry of such patterns on the wafers and masks. *SLIDE* is able to efficiently detect any possible rotational misalignment angle, and then estimate the position of the axis of symmetry.

Consider the simple case of Figure 6. Assuming that the pattern profile in each row of the image can be represented by

$$f(x - x_0 + y \tan\theta), \tag{6}$$

the measurements after dechirping would be

$$w(y) = e^{-j\alpha x_0 y} F(\alpha y) \tag{7}$$

which have the structure of a modulated signal with a central carrier frequency αx_0. The shape of the sidebands around the carrier frequency in the frequency domain is a scaled version of $f(.)$, and hence is symmetric.

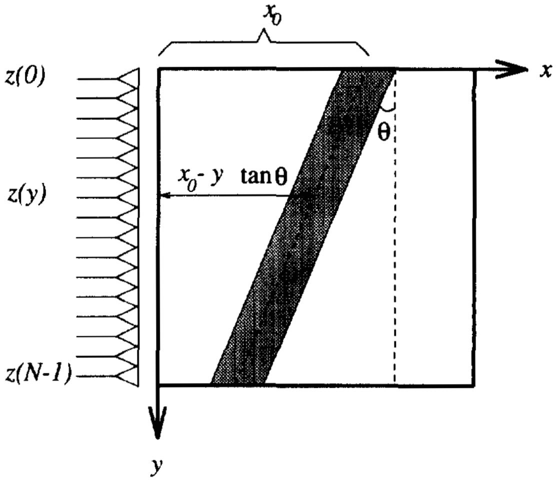

Figure 6: Image with a symmetric pattern.

It can be shown that applying subspace fitting techniques to such modulated signals results in an estimate of the mean frequency. In this application, the mean frequency is proportional to the position of the axis of symmetry, x_0. Figure 7 presents an example of applying the *SLIDE* algorithm to finding the axis of symmetry of an alignment mark on a wafer. The detected axis of symmetry has been superimposed on the original gray scale image taken from the wafer.

Detection of horizontal and vertical axes of symmetry can be extended to arbitrary symmetric shapes. The basic observation in applying subspace fitting methods is that due to the symmetry in the original shape, the measurements will have symmetric sidebands around the carrier frequency in the frequency domain. Subspace fitting results in an estimate of the

Figure 7: Semiconductor image used for alignment.

carrier frequency. Figure 8 shows an image with a pattern symmetric in the horizontal direction. The location of the detected axis of symmetry has been plotted on the image.

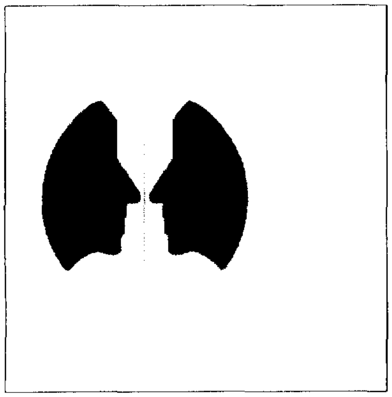

Figure 8: Image with a vertical axis of symmetry.

6 Circle and Ellipse Fitting

The circle is the most symmetric pattern. Therefore, it is expected that the algorithm should be able to find the center of a circle by finding the two horizontal and vertical axes of symmetry in the image. In fact, in addition to the estimation of the location of the center, it can be shown that *SLIDE* can also find an estimate for the value of the circle radius. This value appears as the carrier frequency of a modulated signal, which in this case is itself variable with the measurement index y. The form of the measurements after shifting the origin to the circle center will be

$$z(y) = \begin{cases} e^{j\alpha y h(y)} & , y = 0, \ldots, r \\ 0 & , \text{elsewhere} \end{cases} \tag{8}$$

where

$$h(y) = \begin{cases} \sqrt{r^2 - y^2} & , \quad y = 0, \ldots, r \\ 0 & , \quad \text{elsewhere} \end{cases} \tag{9}$$

It can be shown that the mean frequency is proportional to the value of the radius. This technique is directly applicable to both hollow circles and disks, and can readily be applied to finding the parameters of ellipses as well.

Acknowledgments

The authors wish to thank the helpful discussions with and comments by Babak Khalaj and Professor A. Paulraj of the Information Systems Laboratory, Stanford University.

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Office of Scientific Research under contract F49620-93-1-0085.

References

[1] H. K. Aghajan and T. Kailath, Sensor array processing techniques for super resolution multi-line fitting and straight edge detection, *IEEE Trans. on Image Processing*, **2**(4):454–465, Oct. 1993.

[2] H. K. Aghajan and T. Kailath, *SLIDE:* subspace-based line detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, to appear, 1994.

[3] A. Paulraj, R. Roy, and T. Kailath, Estimation of signal parameters by rotational invariance techniques (ESPRIT), In *Proc. of 19th Asilomar Conference on Circuits, Systems and Comp.*, pages 564–567, 1985.

[4] P. J. Burt, R. Hingorani, and R. J. Kolczynski, Mechanisms for isolating component patterns in the sequential analysis of multiple motion, In *Proc. of IEEE Workshop on Visual Motion*, pages 187–193, Princeton, NJ, 1991.

[5] H. S. Baird, "Anatomy of a versatile page reader", *Proceedings of the IEEE*, **80**(7):1059–1065, July 1992.

[6] S. N. Srihari, "High-performance reading machines", *Proceedings of the IEEE*, **80**(7):1120–1132, July 1992.