BMVC 2002

Orientation Correlation

A.J. Fitch, A. Kadyrov, W.J. Christmas and J. Kittler
Centre for Vision, Speech and Signal Processing
University of Surrey, Guildford, GU2 7XH, UK
a.fitch@eim.surrey.ac.uk

Abstract

A new method of translational image registration is presented: *orientation correlation*. The method is fast, exhaustive, statistically robust, and illumination invariant. No existing method has all of these properties. A modification that is particularly well suited to matching images of differing modalities, *squared orientation correlation*, is also given.

Orientation correlation works by correlating *orientation images*. Each pixel in a orientation image is a complex number that represents the orientation of intensity gradient. This representation is invariant to illumination change. Angles of gradient orientation are matched. Andrews robust kernel function is applied to angle differences. Through the use of correlation the method is exhaustive. The method is fast as the correlation can be computed using Fast Fourier Transforms.

1 Introduction

Image registration, or matching, is the process of aligning two or more images [7]. The topic has a wide range of applications, including: super-resolution, face detection, video coding, medical imaging, database classification, mosaicking, and post-production video effects. In selecting a suitable image registration method one must consider, *i*) the nature of the transformation aligning the images, *ii*) how to evaluate a given transformation.

Transformations can be parametric, e.g. translational, isometric, similarity, affine, projective, or polynomial; or non-parametric, e.g. elastic deformations or thin-plate splines [11]. Orientation correlation finds translational transformations. This is a prominent transformation in many of the above applications. Applications which require a more complex transformation may still use a translational model as the first stage of the estimation process [1, 8].

For evaluation of a transformation two matching methodologies are prevalent in the literature; area-based methods (also known as direct methods) and feature-based methods. Area-based methods match measurable image quantities, e.g. brightness [16] or phase [17], [12]. Feature-based methods match features extracted from the images, e.g. corners [14], lines [9], or junctions [8]. Orientation correlation matches the feature of gradient orientation for each pixel. Orientation correlation is a feature based method with many of the advantageous properties of area-based methods.

With the transformation selected and method of evaluating a transformation defined, image registration becomes an optimisation task to find the best transformation parameters. In the case of a translational transformation, correlation (more exactly cross-

BMVC 2002 doi:10.5244/C.16.11

correlation) has a unique ability to accomplish an exhaustive search quickly [2]. Correlation avoids the fundamental weakness of optimisation techniques; no initial "best guess" of the correct registration is required. Matching with correlation does not require a-priori knowledge. Since the search is exhaustive, local minima are not a problem for correlation-based methods.

The Fast Fourier Transform (FFT) [24] allows fast correlation of digital signals. Correlation using the FFT was first applied to image registration by Anuta [2]. A variety of improvements to correlation have been proposed over the years [17, 19, 22]. These have improved correlation results and addressed issues such as illumination invariance. Comparative studies have been made [3, 20]. In the fields of computer vision and image processing robustness is important. While some of these methods claim to be robust all are based on the square error kernel. Unlike orientation correlation they are not statistically robust.

Huber defines robust statistics as " ... insensitivity to small deviations from the assumptions." [15]. Note that "small" may imply small deviations for all data (e.g. Gaussian noise) or large deviations for a small quantity of data (outliers). In relation to image registration, robustness implies correct registration in the presence of effects such as: noise, occlusion, revealed regions, new objects, and highlights; in general any effect that may cause deviation from a perfect match. A squared error kernel excessively weights outliers. Standard correlation methods do not have a mechanism to handle outliers caused by mismatches at the correct registration.

Robust statistics have been applied to image registration [18], and the related fields of motion estimation [6, 21] and optic flow [4, 5]. Median, used in [4], is robust but computationally expensive. The most prevalent robust method in the literature is the solution of M-estimators with Iteratively Reweighted Least Squares [5, 18, 21]. None of the existing robust methods are fast and exhaustive.

The purpose of this paper is to present orientation correlation. Orientation correlation is a method of translatory image matching which is: statistically robust, illumination invariant, exhaustive, and fast. The method derives statistical robustness by using Andrews' robust M-estimator [15]. The method is illumination invariant as the matching is performed on orientation of image intensity gradient. The method derives its exhaustive property from the use of correlation. Speed is derived from computing correlation with FFTs.

The remainder of the paper is arranged as follows. The method of orientation correlation is presented in Section 2. In Section 3 the method is analysed and its aforementioned properties proven. Section 4 experimentally demonstrates the advantages of orientation correlation in the applications of video coding and multimodal microscopy image registration. Conclusions are drawn in Section 5.

2 Method

This section details our orientation correlation method. Sufficient detail is given for the results presented in Section 4 to be reproduced. Analysis highlighting the properties of the method is in Section 3.

We wish to match discrete images f and g. Images are indexed using (x, y), where x and y are integers. From f and g we construct *orientation images* (orientation of intensity gradient) f_d and g_d. Taking i as the complex imaginary unit and sgn(x) to represent the signum function, f_d is:

$$f_d(x,y) \quad = \quad \text{sgn}\left(\frac{\partial f(x,y)}{\partial x} + i\frac{\partial f(x,y)}{\partial y}\right) \tag{1}$$

$$\text{where} \quad \text{sgn}(x) = \begin{cases} 0 & \text{if } \ |x| = 0 \\ \frac{x}{|x|} & \text{otherwise} \end{cases}$$

Partial differentials are calculated using central differences [23]. To avoid a reduction in image size forward / backward differences are used for border pixels. g_d is constructed in the same fashion as f_d.

Orientation images are matched using correlation. Correlation is computed quickly with Fast Fourier Transforms (FFTs). Given $F_D(k,l)$, the FFT of $f_d(x,y)$, $G_D(k,l)$, the FFT of $g_d(x,y)$, and IFFT() the Inverse Fast Fourier Transform function, the orientation correlation matching surface is:

$$\Re\Big\{\text{IFFT}\big(F_D(k,l)G_D^*(k,l)\big)\Big\} \tag{2}$$

The registration of f and g is measured from the position of the maximum in (2).

Note that frequency domain correlation has the effect of correlating f_d with g_d^*, which is a necessary part of the algorithm. Note also that the correlation is cyclic. This derives from the cyclic nature of the FFT. If shifts greater than half the image size are expected the four possible interpretations of the maximum must be considered, and the best selected. Finally note that (2) assumes f_d and g_d are the same size. Images of differing sizes are correlated with FFTs by zero padding the smaller size to the size of the larger prior to taking the forward FFTs.

Squared orientation correlation If the images to be matched are the inverse of one another, e.g. for a no shift perfect match $f(x,y) = -g(x,y)$, then squared orientation correlation should be applied. In squared orientation correlation the right hand side of (1) is squared. The method is able to match images where one, both, or neither have been inverted. Images with inverted regions will also be matched.

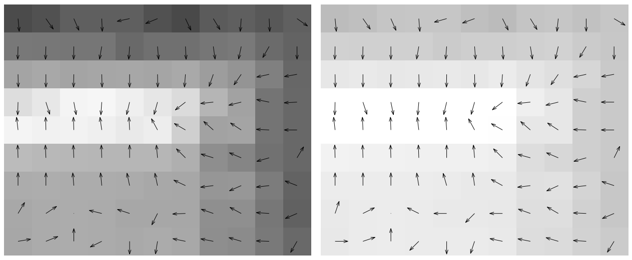
3 Analysis

In this section we analyse orientation correlation. The methods properties of illumination invariance, statistical robustness, and speed are explained.

3.1 Illumination invariant representation

Orientation correlation matches orientation images f_d and g_d, (1). An orientation image is invariant to both scale and offset illumination changes.

Each pixel in a orientation image is a complex number. Each complex number represents the orientation of intensity gradient at that pixel. The magnitude of a pixel is either one or, in the case of a uniform region of the image with no gradient, zero. Since correlation is used for matching, a $0 + 0i$ pixel will have no effect. This is a desirable property -



(a) Original (b) 0.5 scale, 150 offset.

Figure 1: Sub-image and overlaid arrows representing the orientation image.

a uniform area of an image provides no information for a match invariant to illumination change.

Orientation correlation is well suited to matching images of different modalities, e.g. matching infrared images with intensity images. This is because all orientation images have the same units; each pixel is a measure of an angle.

Figure 1(a) shows a region of a video frame with arrows overlaid to show the complex pixel values of the orientation image. Figure 1(b) shows the region scaled and offset. Note the orientation images of Figure 1(a) and 1(b) are the same.

3.2 Statistically robust matching

Orientation correlation applies a kernel based on the M-estimator proposed by Andrews [15] to differences in gradient orientation. Andrews proposed an influence function (derivative of kernel) as:

$$h'(d) = \begin{cases} \sin(d) & \text{for } -\pi \leq d \leq \pi \\ 0 & \text{elsewhere} \end{cases}$$

where d is the quantity to be minimised. Thus Andrews kernel function is of the form:

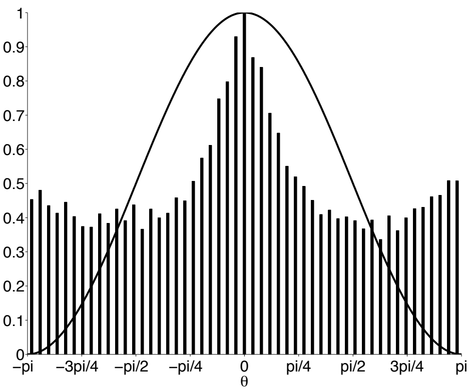
$$h(d) = \begin{cases} -\cos(d) & \text{for } -\pi \leq d \leq \pi \\ 1 & \text{elsewhere} \end{cases} \tag{3}$$

To show how orientation correlation uses the Andrews kernel we consider two orientation image pixels, p and q. As p and q are orientation image pixels their magnitude is either one or the pixel is $0 + 0i$. Since correlation is multiplicative a $0 + 0i$ pixel has no effect. Therefore we consider: $|p| = |q| = 1$. Taking the complex conjugate of q, and expressing the pixels in a polar form gives:

$$p = e^{i\phi}, \quad q^* = e^{-i\psi}$$

Correlation shifts one image with respect to the other and measures the sum of the products. The product of pixel p and the complex conjugate of pixel q is:

$$p.q^* = e^{i(\phi-\psi)}$$



(a) Robust kernel with normalised histogram of θ at the OC match.

(b) Mosaic of the OC match.

Figure 2: Orientation correlation (OC) matching the first and 80th frames of the coastguard sequence.

Returning to Cartesian coordinates, taking the real part only, and substituting $\theta = \phi - \psi$ we have:

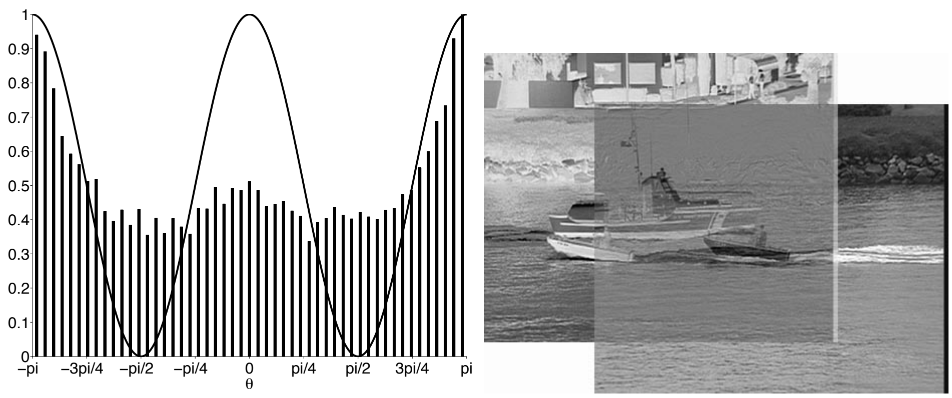
$$\Re\{p.q^*\} \quad = \quad \cos(\theta) \tag{4}$$

From this last equation we see that the real part of the multiple of p and q^* is the same as differencing angles (orientations) of each pixel and applying a cosine kernel function.

Comparing (4) with (3) we see that over the range $-\pi \leq d \leq \pi$, the functions differ only in a coefficient of -1. Selecting the range of θ to be $[-\pi, \pi]$, (4) is equivalent to the Andrews kernel function. The significance of the differing coefficient of -1 is that the best orientation correlation match will be indicated by a maximum.

Orientation correlation's robust kernel is shown in Figure 2(a). The kernel has been scaled and shifted for display with a normalised histogram. The normalised histogram is of θ at the correct cyclic match of the first and 80th frames of the coastguard sequence. A mosaic of the coastguard frames as found by orientation correlation is shown in Figure 2(b).

The histogram in Figure 2(a) contains two distributions. Correctly matching orientations are spread about $\theta = 0$. Effects such as object motion and cyclic match wrapping cause locally incorrect orientation matches. These are uniformly spread across the range of θ. Figure 2(a) shows that the robust kernel maximises the correct match orientations without overly weighting the contribution of incorrect orientation matches. The kernel is wider than the distribution of correct match orientations of the coastguard frames. From this observation we can expect orientation correlation to work equally well with images containing a higher level of noise.



(a) Robust kernel with normalised histogram of θ at the SOC match.

(b) Mosaic of the SOC match.

Figure 3: Squared orientation correlation (SOC) matching the first frame and inverted 80th frame of the coastguard sequence.

Squared orientation correlation: In squared orientation correlation orientation image pixels are squared. Considering squared orientation images pixels p and q^*, setting their magnitudes to one, and expressing them in polar form we have: $p = e^{i2\phi}$ and $q^* = e^{-i2\psi}$. Using $\theta = \phi - \psi$ the real part of pq^* is: $\Re\{p.q^*\} = \cos(2\theta)$. Thus we see that correlating one squared orientation image with the complex conjugate of another applies a $\cos(2\theta)$ error kernel to the difference of pixel orientations.

The robust kernel of squared orientation correlation is shown in Figure 3(a). The kernel has been scaled and shifted for display with a normalised histogram. The normalised histogram is of θ at the correct cyclic match of the first frame and inverted 80th frame of the coastguard sequence. A mosaic of the frames as found by squared orientation correlation is shown in Figure 3(b).

As with Figure 2(a) the histogram in Figure 3(a) contains two distributions. Again the distribution of incorrectly matching orientations is approximately uniform. However, with one frame inverted the distribution of correctly matching orientations is centered about $\theta = \pi$ (which is the same as $\theta = -\pi$). Figure 3(a) shows that the robust kernel of squared orientation correlation maximises distributions centred about both $\theta = 0$ and $\theta = \pi$. Thus squared orientation correlation can match images where neither, one, or both are inverted.

3.3 Computational cost

The majority of the computational cost of orientation correlation is with the FFTs. Orientation correlation requires two forward and one backward FFTs of complex value images. Three complex FFTs of size X by Y require $6XY\log_2(XY)$ real multiplications [24]. Equivalent data domain image registration requires X^2Y^2 operations.

Method	MAϵ					
	bus		coastguard		foreman	
		as % of FS		as % of FS		as % of FS
NC	10.3	130%	6.57	115%	4.44	157%
PC	11.2	141%	7.40	130%	4.75	167%
OC	9.13	115%	6.06	106%	3.57	126%
FS	7.91	100%	5.69	100%	2.84	100%

Table 1: Mean absolute prediction error (MAϵ) for normalised correlation (NC), phase correlation (PC), orientation correlation (OC), and full search (FS) methods.

4 Experiments

This section experimentally demonstrates the advantages of orientation correlation over the standard cyclic correlation techniques of normalised correlation and phase correlation [17].

4.1 Video coding

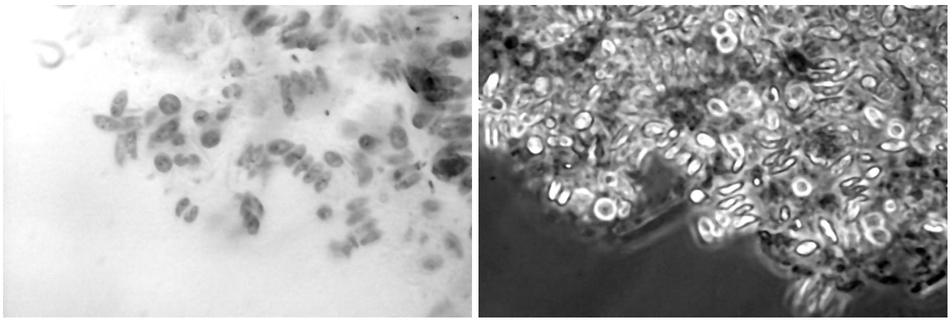
Here orientation correlation is compared to normalised correlation and phase correlation in the application of block motion estimation for video coding. Experiments are undertaken on the first 150 frames of three Cif size (288×352) video sequences, using standard MPEG size (16×16) blocks. A well defined goal, minimising prediction error, makes for a good comparative test. Motion vectors are limited to $[-8 : 7]$ with integer pixel accuracy. Comparison is made in relation to baseline full search (FS) method. The performance is measured by the ability of the method to minimise mean absolute prediction error (MAϵ). The MAϵ of frame f is defined as:

$$\text{MA}\epsilon \quad = \quad \text{mean}_{x,y}|f(x,y) - f'(x,y)|$$

where f is an original frame and f' is its reconstruction from motion compensation of the previous frame.

Temporally adjacent blocks are correlated to generate motion vectors. Correlating sequential blocks is efficient, each block need only be transformed once; compared to non-sequential correlation the overall number of forward FFTs can be halved. A relatively small block size does not suit correlation techniques; benefits of the FFT are greater for larger images. Even so a significant gain in performance over the full search method is achieved. Full search requires $16^4 = 65,536$ subtractions and counter increments per block. Orientation correlation requires $2 \times 2 \times 16 \times 16 \times \log_2(16 \times 16) = 8,192$ real multiplications per block. Modern CPU perform add or multiply operations in one clock cycle, these numbers are comparable operation counts.

MAϵ for the first 150 frames of bus, coastguard, and foreman sequences is shown in Table 1. On all three sequences orientation correlation (OC) outperforms both normalised correlation (NC) and phase correlation (PC).



(a) brightfield (b) phase contrast

Figure 4: Multimodal microscope images from [10]

4.2 Registration of multimodal microscopy images

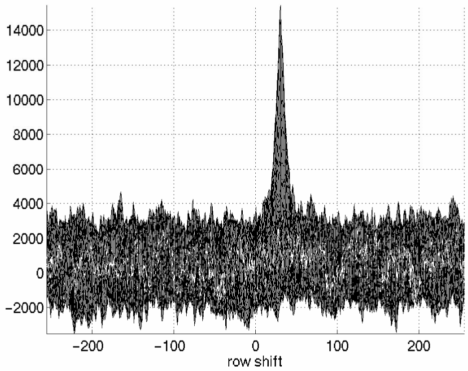
Here we demonstrate squared orientation correlation successfully registering a particularly challenging pair of multimodal microscopy images. The images, shown in Figure 4, are of algal and bacterial cells [13]. During the change in optics a shift in position is introduced. Automatic registration of the images is required for analysis.

The images are challenging to register due to their size (512 by 768), size of shift (30,169), widely differing modalities of the images (light regions in Figure 4(a) are dark in Figure 4(b)), and the differing features of each image modality. [13] reports that normalised correlation and phase correlation are unable to register the images. Squared orientation correlation successfully registers the images. The squared orientation correlation matching surface is shown in Figure 5. Figure 5 shows a clear peak at the correct registration position.

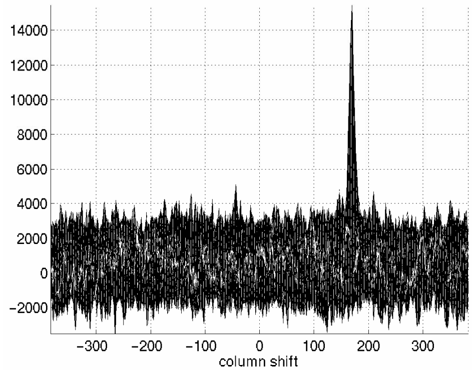
[13] registers the images by correlating absolute gradient. This is not robust. To demonstrate the advantage of squared orientation correlation over correlating absolute gradient images, we crop the right and bottom of the brightfield image and the left and top of the phase contrast image. The more the images are cropped the harder the images are to register. We measure the percentage overlap of the correct registration. For example, with no cropping 73% of each image overlaps at the correct registration. Correlating absolute gradient fails even when the images have a 59% overlap. Squared orientation correlation is able to correctly register the images with only a 12% overlap.

5 Conclusion

Orientation correlation, a fast, exhaustive, illumination invariant, statistically robust, translational image matching technique has been presented. Analysis has been made of the methods illumination invariance, statistical robustness, and computational cost. Advantages of the method has been experimentally shown in the applications of video coding and registration of multimodal microscopy images.



(a) side view showing row shift



(b) side view showing column shift

Figure 5: Squared orientation correlation matching surface of the images in Figure 4.

References

[1] R.J. Althof, M.G.J. Wind, and J.T. Dobbins III. A rapid and automatic image registration algorithm with subpixel accuracy. *IEEE Trans. Med. Imag.*, 16(3):308–316, June 1997.

[2] P.E. Anuta. Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques. *IEEE Trans. on Geoscience Electronics*, GE-8(4):353–368, October 1970.

[3] P. Aschwanden and W. Guggenbühl. Experimental results from a comparative study on correlation-type registration algorithms. In W. Förstner and St. Ruwiedel, editors, *Robust computer vision : quality of vision algorithms*, pages 268–289. Wichmann, 1992.

[4] A. Bab-Hadiashar and D. Suter. Robust optic flow computation. *International Journal of Computer Vision*, 29(1):59–77, 1998.

[5] M.J. Black and P. Anandan. The robust estimation of mulitple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, 1996.

[6] M. Bober and J. Kittler. Estimation of complex multimodal motion: An approach based on robust statistics and hough transform. *Image and Vision Computing*, 12(10):661–668, December 1994.

[7] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(2):325–376, December 1992.

[8] A. Can, C.V. Stewart, and B. Roysam. Robust hierarchical algorithm for constructing a mosaic from images of the curved human retina. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 286–292, June 1999.

[9] W.J. Christmas, J. Kittler, and M. Petrou. Structural matching in computer vision using probabilistic relaxation. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(8):749–764, August 1995.

[10] C.A. Glasbey. Microscope images. http://www.blackwellpublishers.co.uk/rss/Readmefiles/mardia.htm.

[11] C.A. Glasbey and K.V. Mardia. A review of image warping methods. *Journal of Applied Statistics*, 25:155–171, 1998.

[12] C.A. Glasbey and K.V. Mardia. A penalized likelihood approach to image warping. *Journal of the Royal Statistical Society, Series B*, 63:465–514, 2001.

[13] C.A. Glasbey and N.J. Martin. Multimodality microscopy by digital image processing. *Journal of Microscopy*, 181:225–237, 1996.

[14] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2001.

[15] Peter J. Huber. *Robust Statistics*. John Wiley & Sons, 1981.

[16] M. Irani and P. Anandan. All about direct methods. In W. Triggs, A. Zisserman, and R. Szeliski, editors, *Proc. of the Int. Workshop on Vision Algorthims*, pages 267–277, 1999.

[17] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. In *Proc. of the IEEE Int. Conf. on Cybernetics and Society*, pages 163–165, 1975.

[18] Shang-Hong Lai. Robust image matching under partial occlusion and spatially varying illumination change. *Computer Vision and Image Understanding*, 78:84–98, 2000.

[19] J.P. Lewis. Fast normalized cross-correlation. In *Vison Interface*, pages 120–123, 1995. http://www.idiom.com/∼zilla/Work/nvisionInterface/nip.html.

[20] J. Martin and J.L. Crowley. Experimental comparison of correlation techniques. In *Proc. International Conf. on Intelligent Autonomous Systems*, 1995.

[21] J.M. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models applied to complex scenes. Technical Report 788, Institut de Recherche en Informatique et Systèmes Aléatoires, France, 1994.

[22] W.K. Pratt. Correlation techniques of image registration. *IEEE Trans. Aerosp. Electron. Syst.*, AES-10(3):353–358, 1974.

[23] A. Rosefeld and A.C. Kak. *Digital Picture Processing*, volume 2. Academic Press, 2nd edition, 1982.

[24] James S. Walker. *Fast Fourier Transforms*. CRC Press, 2nd edition, 1996.