Resolving Handwriting from Background Printing Using Photometric Stereo

Abstract

We propose a scheme to resolve handwriting from background printing. The scheme detects the indentations made by the pen in the paper. Photometric stereo is used to recover the surface; a matched filter and classifier are used to detect the stroke indentation. We assess the effect of uniform and textured backgrounds on the recovery of the stroke and test the scheme on practical examples. The technique was found to work well with script written with a ballpoint pen and could effectively suppress even dark and strongly textured backgrounds. We conclude that this is a useful complement to existing techniques for background removal and is especially useful when there is no template available.

Keywords

Background removal, document image analysis, preprocessing, off-line handwriting analysis,photometric stereo, thresholding.

I. Introduction

We propose a technique to resolve handwriting from background printing. Our approach relies on the fact that handwriting leaves indentations on the paper surface whereas most forms of printing do not. Photometric stereo (PS) is used to estimate the topography of the surface; classical signal processing techniques are used to detect the indentations. The technique is useful for forensic document analysis and automatic processing of forms and cheques.

In this paper we assess how well the technique can recover handwriting from background. The ability of PS to recover topography and suppress printing is tested. We assess the consistency of the indentations with different pens, writers and inks. The topography of the paper surface is modelled and combined with the earlier findings to generate a detector. The technique is applied to six practical examples.

We found the PS algorithm to be accurate over the range of slope angles required for handwriting recovery. The indentations of ballpoints could be easily detected, though those made by propelling pencil were much less obvious and those made by felt tip could not be identified at all. The indentations made by ballpoints were found to be consistent for the two types of pen and the 4 different colours of ink tested. Although there was some variation in amplitude this seems to be largely due to the writer. We found we could recover the pen indentations from very dark or textured surfaces as well as the easier cases.

We proposed a detector based on a matched filter, thresholding and non-linear processing. This was found to be effective at resolving handwriting from background printing in real examples.

We conclude that the technique works well at separating ballpoint handwriting from printed background. The technique is more versatile than the techniques reported in the literature: it makes no requirements about the nature of the background or its alignment and does not require a template. Set against this, it does require a photometric set of images—though these can be captured quickly with the appropriate hardware. We conclude that that the technique is complementary to existing algorithms, and is especially useful where no template is available.

II. Literature Review

The recognition of handwriting is an established and extensive field of pattern recognition, see [12] for survey. It encompasses techniques from low level image processing operations to high level language modelling. The field may be split into two broad categories: *on-line*—dealing with input from a graphics tablet in real-time; and *off-line*—the examination of imaged, paper documents. Off-line recognition is most relevant to this work. Plasmodon and Srihari identify three stages in off-line recognition: preprocessing, character recognition and word recognition[12]. This paper is concerned with preprocessing.

Wang and Pavlidis present an algorithm that operates on the grey level image and recognises parts of characters as characteristic landmarks on a three dimensional surface [19]. However, most recognition algorithms operate on binary images and thresholding has been extensively studied: Sankur groups forty-two thresholding algorithms in his classification [13]. Solihin and Leedham divide thresholding algorithms into global and adaptive techniques [16]. Because most handwriting is on a uniform background and is evenly lit, global algorithms have been most thoroughly researched. Trier and Jain present a comparison of four global algorithms and report Otsu's [9] to be the most effective [18].

A related, though more complex problem, is identifying handwriting on a non-uniform background. Xiangyu presents an algorithm that uses a 'double-edge' model of the handwriting stroke to detect strokes while avoiding other edges [22]. Franke and Köppen [3] developed a forensic document examination system that applied filters, designed with a genetic algorithm, to remove background from handwritten text. In many applications, for example reading handwriting from cheques or forms, a template is available and many schemes have been proposed to take advantage of this, e.g. [8] [1] [5]. However, even with this simplification, handwriting often overlaps printed matter and several algorithms have been developed for line removal, e.g. [23] [24] [6]. Shetty et al. [15] isolated handwritten and machine printed text from a cheque by generating a 'pseudo-template' from the *completed* cheque. Morphological operators were used to suppress the (high contrast) handwritten text, and the result was then used to remove background printing from the completed cheque.

This paper proposes a novel preprocessing technique for off-line handwriting recognition. The aim of the technique, like [3], is to isolate handwriting from background printing. The novelty of this paper lies in the use of photometric stereo to recover the topography of the handwriting. Like [3] we use linear filters to recover the desired signal—though we describe the task in terms of a classical detection problem—and the approach is analogous to the geometric techniques used by [19] and [22]. As with most techniques we use a global thresholding algorithm. We found Otsu's algorithm gave noisy results and that a modified form of the *K-means* algorithm was more effective. Finally we applied morphological operators to the binarised image to reduce spurious detections.

III. Accuracy of Surface Recovery

Our approach is based on photometric stereo (PS). PS infers the slope of a surface patch by measuring how the intensity of the patch changes as the direction of the light source is varied [20]. The slope of the facet is described by the partial derivatives of the surface p and q. Cho et al. [14] use a two-light shape from shading algorithm to recover the global shape of a book surface. Hansson and Johansson applied photometric stereo to paper surfaces for an inspection task [4].

Classical photometric stereo assumes the surface obeys Lambert's law and requires three images of the surface lit from different directions. We used a modified version of photometric stereo which allows the use of many images to increase the robustness of the estimation. Throughout this paper, we use 12 images, with illumination at 30° increments of azimuth.

For each surface facet we calculate the Fourier Series of the intensity as a function of lighting azimuth. The Lambertian component of reflection contributes to the mean and the fundamental of the Fourier Series; specularities and shadowing contribute to these components, but also to the higher harmonics. By suppressing the higher harmonics, we reduce the effect of the non-Lambertian components of reflection. This will increase the validity of the assumption that the surface is Lambertian, allowing classical photometric stereo to be applied more accurately, see [7] for a fuller description.

It is important to assess the accuracy and limitations of PS when applied to paper under raking light. We do this by wrapping several sheets of paper round a cylinder (i.e. a known shape) and estimating the surface derivatives. We keep the same imaging setup for all the experiments in this paper. A Vösskühler, 12 bit, monochrome camera with constant gain is used. The sample is lit from a zenith of 70°.

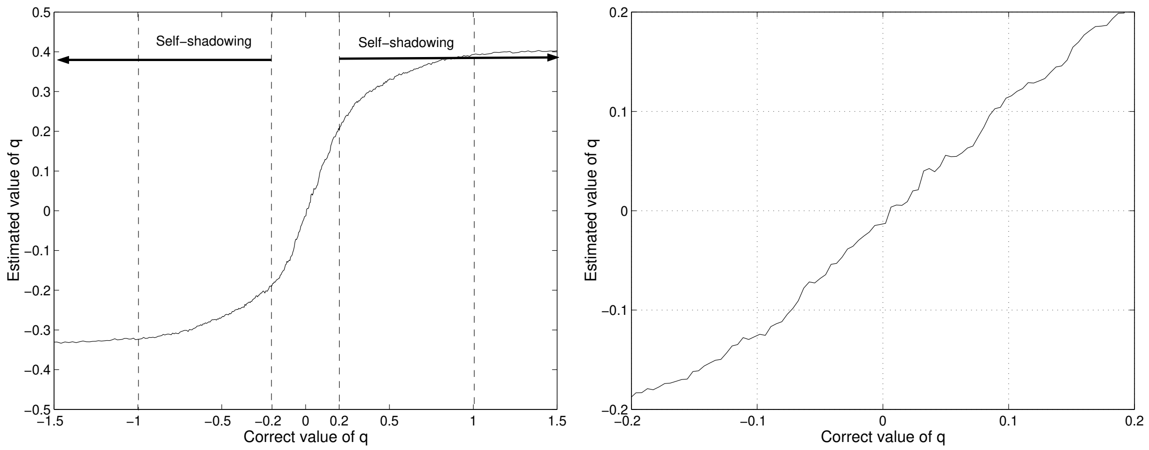

Fig. 1. Estimation accuracy.

Raking light is required to accentuate subtle aspects of the surface topography. However, it introduces artefacts: most significantly the spatial variation in the amount of incident light. This effect is present under more moderate lighting and is generally corrected using flat-fielding. Under raking light the effect is severe, and we have found flatfielding alone is not sufficient. In addition, we apply an isotropic, bandpass filter with a mean wavenumber of $1136m^{-1}$ and standard deviation of $568m^{-1}$. This attenuates low frequency components due to variations in incident light and high frequency noise components giving an obvious improvement in the image.

In Figure 1 we plot the estimated derivatives q for the cylinder against the known derivatives. The relationship between the estimate and the correct value falls into two categories: for shallow gradients there is a linear relationship; however, when the facet is steeper than the illuminant vector, i.e. $| q |> 0.2$, self-shadowing occurs and the estimate begins to saturate. The estimator is accurate for shallow angles and underestimates the slope of steeper facets.

IV. Consistency of the Recovered Signal

We recover the indentations made by the pen on the paper surface. To identify the indentations we must know their form—and assess the variabilty of that form. We would expect the writer, the writing implement and natural variation in the handwriting of a single writer to affect the shape of the indentation. We might also expect the type of ink to affect the photometric recovery of the indention. In this section we assess the effect of the type of pen and ink used, and the variation within script from the same author.

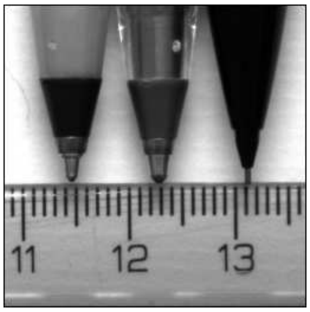

Fig. 2. BIC Fine, BIC Medium and Propelling pencil.

A naive writer was asked to write a series of strokes with different pens: a felt tip, a propelling pencil, and several ballpoints, Figure 2. Two common types of ballpoint were used: BIC Medium and BIC Fine, different colours of ink were used. PS was applied and a series of profiles, shown as white lines on Figure 3, were measured along the p derivative field and are plotted in Figure 4.

We found that the recovery algorithm was unable to detect any indentations made by the felt tip. The pencil did make detectable indentations—though these are not apparent in the profile they are visible in the derivative field. The ballpoints all made obvious indentations: these look like localised sinusoids. Note that the range of slopes measured are

Fig. 3. Test Samples

well within the linear region of photometric estimation discussed in the previous section—we would not expect the estimate to saturate. There is some variation between different pens, however, the greatest source of variation appears to be the natural variation of the writer. Furthermore, this is a variation in amplitude rather than shape.

V. Recovery from Uniform Backgrounds

Our aim is to identify handwriting (topography) and suppress background printing (albedo). We estimate topography by measuring how the direction of lighting affects the variation the amount of light reflected from surface—a non-reflecting surface would confound the algorithm. In this section we assess the effect of dark backgrounds on the recovered data.

We printed two, vertically aligned, blocks on a single page with a modern laser printer (HP 2200 DTN) using a two day old toner cartridge. The first block was set to black, the second block was set to mid-grey, Figure 5 (left). The printer has a specified resolution of 1200 dpi, corresponding to a sampling frequency of approximately 47000 m^{-1}. This is well above the nominal sampling frequency of the camera (approx. 9000m^{-1}), and the dithering associated with the mid-grey block will be averaged out. A second naive writer was asked to draw a single vertical, ruled line through the blocks. The surface derivatives were estimated, and profiles of the ruled line were taken in the black, and mid-grey regions. For comparison, a third profile was taken from the unprinted region midway between the blocks.

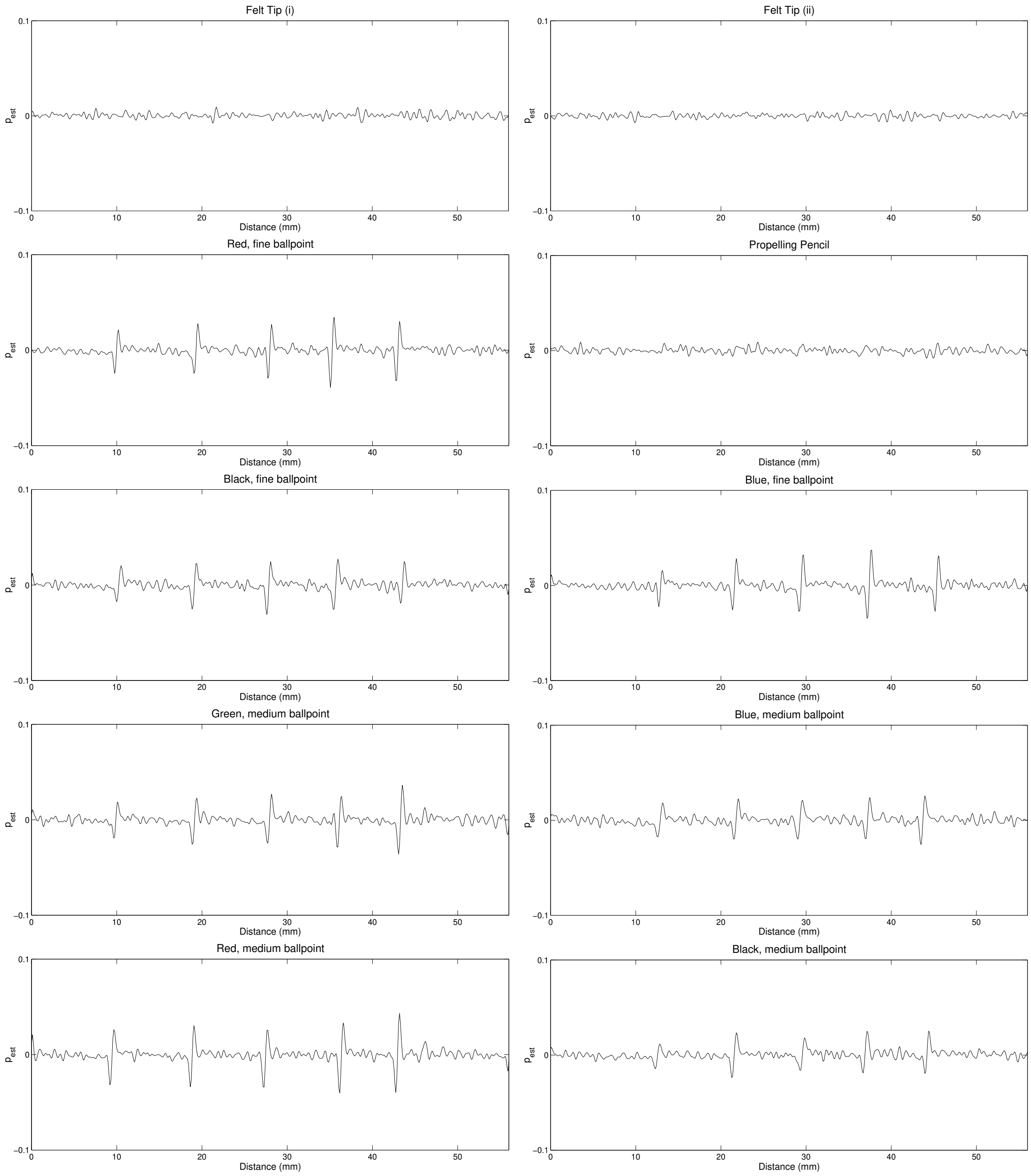

Fig. 4. Profiles of the derivative field.

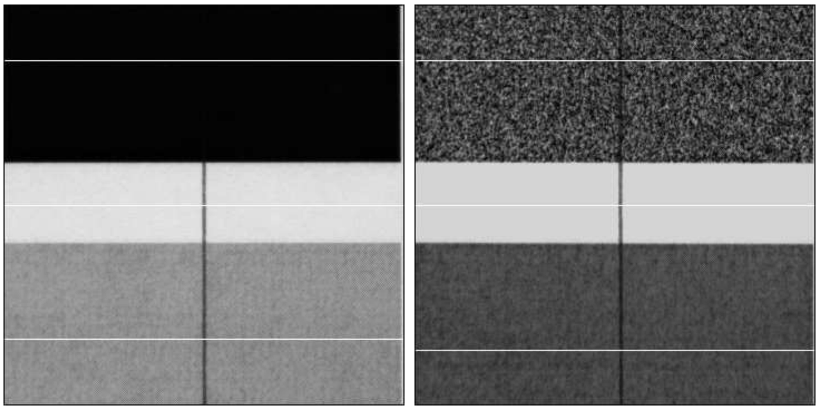

Fig. 5. Test blocks: uniform background (left) and textured background (right)

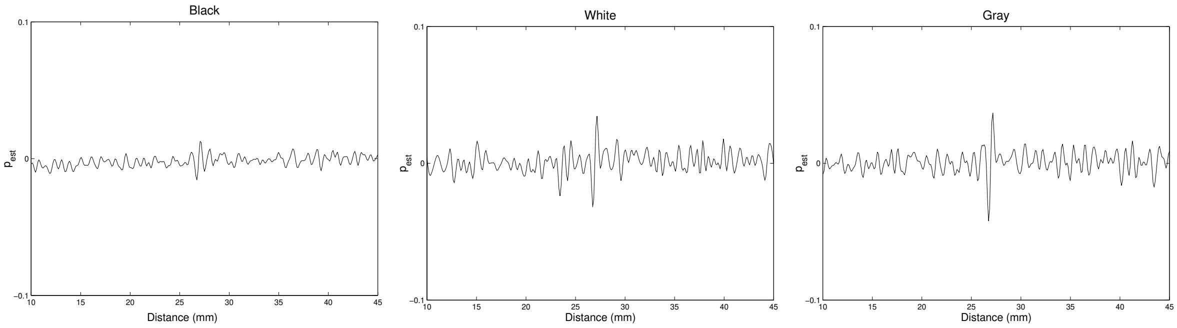

Fig. 6. Recovered profiles from uniform background.

Comparison between the profiles taken in the mid-grey and the unprinted regions shows that the mid-grey print has had little effect and has been effectively resolved, Figure 6. The profile from the black region is attenuated compared to the other two profiles. However, both the signal and the background noise have been attenuated, and the indentation is still obvious.

VI. Recovery from Textured Backgrounds

In many applications handwriting occurs on rapidly varying backgrounds. These are particularly difficult for image-based algorithms. In this section we assess the ability of the algorithm to deal with textured backgrounds. We printed two vertically aligned blocks on a single page, Figure 5 (right). Both are white noise random processes, with means equal to the mid-grey block of the previous experiment. The blocks differ in their standard deviation. As with the previous experiment we estimate the p derivatives along three profiles. All profiles show well developed peaks corresponding to the ruled line, Figure 7. The peaks are similar in shape and amplitude, though there is some variation in the background noise level.

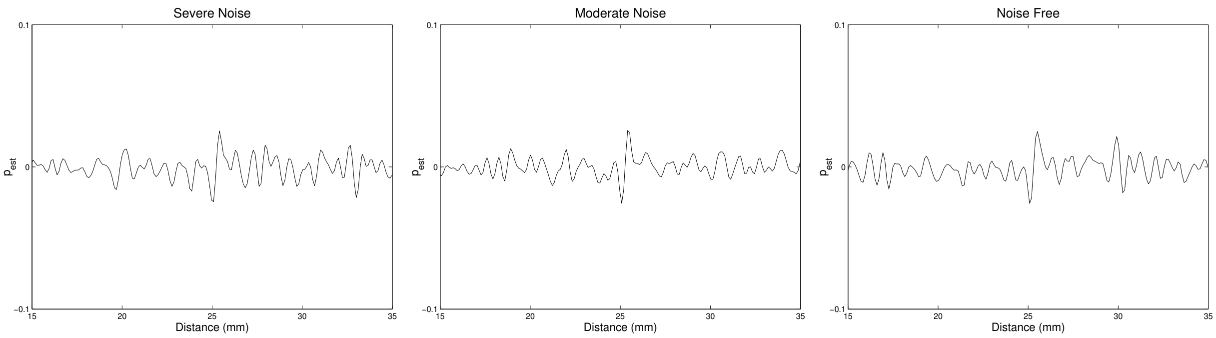

Fig. 7. Profiles recovered from textured background.

As an additional experiment we measured the standard deviation of three areas of the image. The measured standard deviations are all similar and what variation there is does not correspond to the variance of the noise block, Table I. This suggests the major source of background noise is the topography of the paper or errors in its estimation.

Noise	p_{rms}
Severe	0.0061
Moderate	0.0056
None	0.0061

TABLE I

Estimated roughness of different regions.

VII. Papers

It is clear from the previous experiments that the majority of background noise is due to the topography of the paper. In this section we model the topography with the aim of reducing its effect. We recover the topography of three different types of paper: $100gm^{-2}$, $75gm^{-2}$ and $70gm^{-2}$, Figure 8. The root mean square (rms) values of the surface derivatives are shown in Table II. The $100gm^{-2}$ surface is clearly the roughest, the other papers are approximately equal. The combined power spectra of the surface derivatives are shown in Figure 9. All the spectra have a broadband signal component, whose shape is dominated by the detrending filter. However both the $100gm^{-2}$ and the $75gm^{-2}$ papers have obvious additional evanescent components—vertical for the $100gm^{-2}$ case, horizontal and diagonal for the $75gm^{-2}$ case. The radial spectra are shown in Figure 10. All the spectra are low pass, and approximate a f^{-2} process.

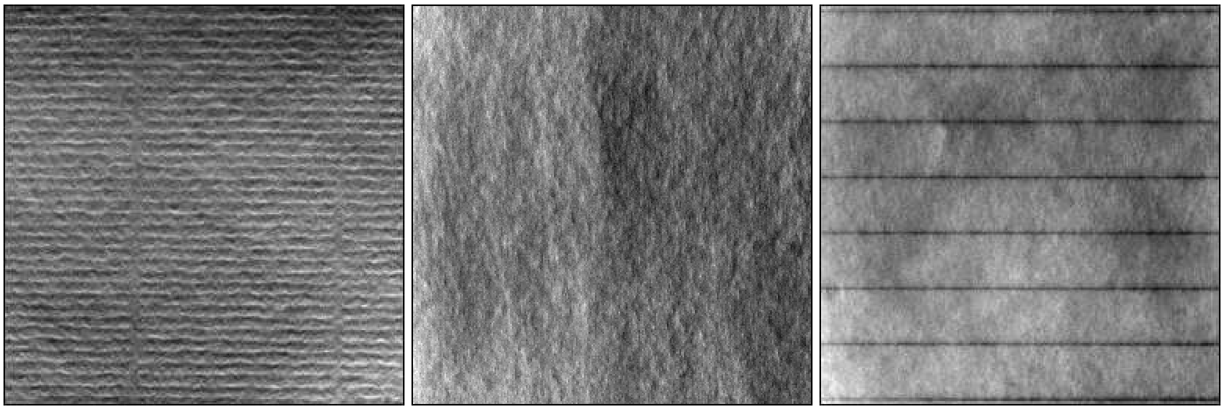

Fig. 8. Test papers (lit to maximise visible texture)

Paper	p_{rms}	q_{rms}
100 gm^{-2}	0.012	0.013
80 gm^{-2}	0.006	0.004
60 gm^{-2}	0.005	0.006

TABLE II

Estimated roughness of different paper types.

VIII. A System

In the previous sections we analysed the shape and variability of the indentations, the effect of background printing and the topography of the paper itself. We now use this knowledge to develop a system to detect the indentations.

We can treat the recovery of indentations either as an exercise in enhancement or detection. That is we can optimise either mean square error (MSE) or signal to noise ratio (SNR). The MSE criterion is suitable where we must determine the uncorrupted form of a signal that we know is present. The SNR criterion is suitable where the form of the signal

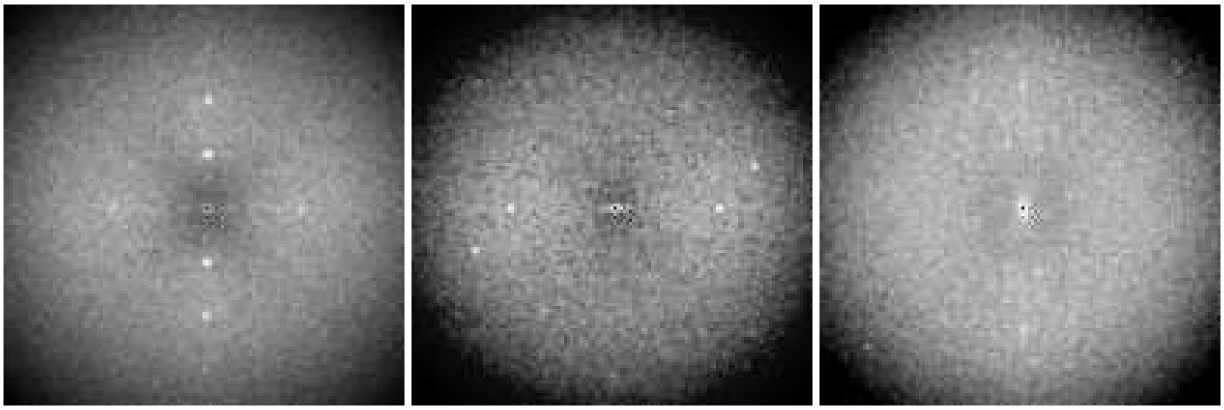

Fig. 9. Log power spectra of paper samples.

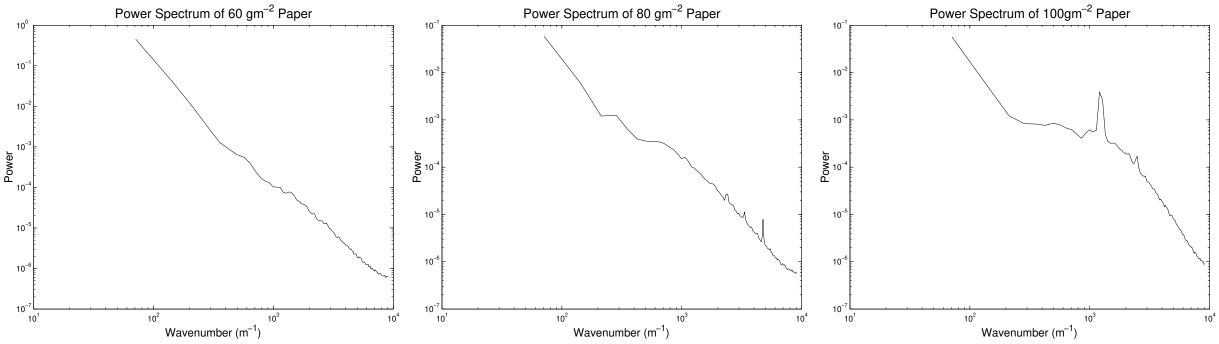

Fig. 10. Radial (log/log) power spectra of test papers.

is known, but its amplitude and time of occurrence are not [2]. This is more suited to our application.

The SNR criterion leads to the *matched filter*, Equation 1.

$$H(\omega) = \frac{S^*(\omega)}{N(\omega)} \tag{1}$$

If the noise is white, the filter has the form of the time-reversed signal. Figure 10 shows that the noise is coloured: in this case the matched filter may not be realisable [11]. The usual approach is to pre-filter the signal to whiten the noise spectrum. Because we do not know what type of paper we are dealing with we cannot guarantee the result of pre-filtering will be white noise, i.e. the matched filter will be sub-optimal. However, by using a generic high pass filter (a wide band Gabor filter) we can make the background noise *whiter* and can improve the performance of the matched filter.

The matched filter requires a prototype of the desired signal. We used the recovered indentation of the horizontal line shown in Figure 5. The pre-filter must also be applied to the training signal before the matched filter can be designed. Until now we have considered the matched filter as being one dimensional. The desired signal is made up of a series of randomly oriented strokes. If the filter is made more directional (and therefore wider) then it will improve the recovery of some of the strokes. However, it will attenuate strokes that are not aligned with the filter. The partial derivative is a directional operator; by applying the filter to a derivative field we are filtering a particular range of directions. Because the partial derivative is a steerable basis we can gain selectivity in any direction. In practice, the level of selectivity is poor and improvements can be gained by making the filters more directional.

We adopt the following approach. The matched filter is derived from a (reversed) horizontal profile of the indentation of the vertical line from Figure 5. The filter is widened and weighted with a Gaussian. This form of the filter is applied to the p-derivative, a rotated version is applied to the q-derivative.

We treat the application as a problem of detection: implemented as a matched filter followed by a threshold. We use the K-means algorithm to set the threshold. K-means does not assume an *a priori* probability of the classes. In fact, most of the written page is white space, and we have found that applying K-means alone result in many spurious detections. To reduce the number of false alarms we decided to introduce some domain knowledge into the algorithm.

We estimated the approximate ratio of white space to black space using 12 images from the CEDAR database [17]. This database contains examples of on-line handwriting—the samples are binary, and do not need to be thinned. It therefore gives an upper bound on the amount of white space. The results are shown in Table III.

We used this result to modify the K-means algorithm: if the hypothetical threshold results in less than 85% of the page being classified as white space, then the means of the clusters are revised to increase separation and the algorithm continues to iterate. Despite the arbitary nature of the modification, it does give a significant improvement in results.

After thresholding we apply an empirical, non-linear filter. For each pixel that is classified as being black, we examine the surrounding 5×5 region. If none of the boundary

Sample	White space (%)	Sample	White space (%)	Sample	White space (%)
Shop	97	Doctor	97	Informal	95
Legal	97	Teacher	98	Meetings	96
Press	98	Movie	97	Insurance	91
Driving	97	Coach	94	Personal	97

TABLE III

PERCENTAGE OF WHITE SPACE IN DATABASE SAMPLES.

pixels are also black then the pixel is re-classified as being white space, otherwise the classification is allowed to stand.

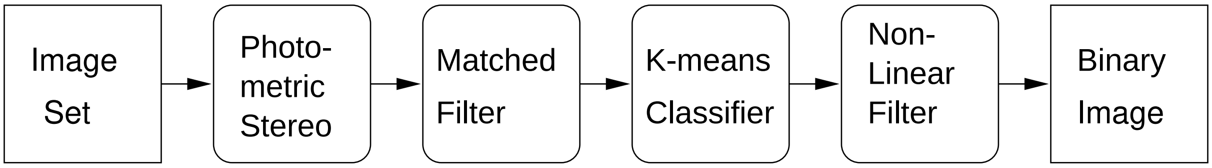

Fig. 11. General approach of detector.

IX. ILLUSTRATIVE EXAMPLES

The previous experiments have all been analytical. We now assess the ability of the proposed system to resolve handwriting from background in practical examples. We used 6 samples: *Map, Form, Pay-in*, Figure 12 (lefthand column) and *Equation, Card, Checklist*, Figure 13 (lefthand column). Each contains script written by a different writer. The central columns show the output of the matched filter, and the right-hand columns show the results of thresholding and non-linear operations.

In all cases PS and the matched filter effectively accentuate the desired signal and attenuate the background. In the *Map* example there are still traces of the printed background: the most obvious are the black lines used to highlight streets. The *Form* example has an obvious artefact in the top, left-hand corner: this is caused by a crease in the paper which is detected as an indentation. In the *Equation* example the background has been effectively removed, though the exponents of the equation, especially the diagonal strokes, have not been properly detected. The *Card* example gives the poorest results due to the resistance and topography of the card. The *Checklist* sample is written on smooth, glossy paper. The specular component of reflection does not seem to affect the surface recovery, and the smoothness of the writing surface causes fewer false detections—this is the most successful classification.

X. Discussion

We have found the technique is effective at resolving ballpoint handwriting from background printing. The shape of the indentations is consistent—though amplitude does vary. The greatest source of variation appears to be the natural variability of a given writer and variations between different writers. Interestingly, Yamazaki et al. [21] use writing pressure as a feature for writer identification in their on-line system. The most attractive feature of the scheme is that it requires few assumptions about the handwriting or the background and does not require a template. Because the main method of reducing background is photometric stereo—a local process—there are no assumptions about the nature or alignment of the background.

Set against these advantages, the technique has several limitations. Firstly, photomteric stereo requires several images of the same scene: this requires extra resources either of time or hardware. Secondly, the use of raking light causes the average intensity of the surface to vary. Flatfielding can reduce this effect, but in practice it is still a significant problem. We reduce its effects by restricting ourselves to small areas of the page approx 56mm $\times$ 56mm. Thirdly the technique measures topography: it will therefore detect structures such as creases in the paper and indentations from writing on previous pages. Both of these problems can be reduced if the recovered text is combined with the test sample under ambient lighting.

XI. Conclusions

We conclude that the combination of photometric stereo and a detector based on a matched filter can resolve handwriting from background printing. This technique has three main disadvantages: it requires several images of each sample; it is sensitive to creases in the paper and indentations made by writing on the previous page; *and* the variation of incident light over the surface restricts us to relatively small regions. Set against these disadvantages, the technique requires fewer assumptions and domain knowledge of the printed sample than the techniques reported in the literature. The technique has a different set of advantages and disadvantages from conventional separation techniques. We conclude that the proposed approach is complementary to existing techniques and is best suited to applications where domain knowledge of the sample is not available.

References

[1] E.S. Angelotti, E.E. Scalabrin, B.C. Ávila, and F. Bortolozzi. A paraconsistent system of autonomous agents for brazilian bank check treatment. In *Computer Science Society, 2000. SCCC '00. Proceedings. XX International Conference of the Chilean*, pages 89–98, 2000.

[2] G. Cooper and C. McGillen. Probabilistic methods of signal and system analysis. In *Second Edition, CBS Publishing, Japan*, 1986.

[3] K. Franke and M. Köppen. A framework for document pre-processing in forensic handwriting analysis. In *IWFHR00*, pages 73–81, 2000.

[4] Peter Hansson and Per-Åke Johansson. Topography and reflectance analysis of paper surfaces using a photometric stereo method. *Opt. Eng.*, 39(9):2555–2561, September 2000.

[5] Kye Kyung Kim, Jin Ho Kim, Yun Koo Chung, and C.Y Suen. Legal amount recognition based on the segmentation hypotheses for bank check processing. In *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, pages 964–967, 2001.

[6] S.H. Kim, S.H. Jeong, and H.K. Kwag. Line removal and character restoration using bag representation of form images. In *IWFHR00*, pages 43–52, 2000.

[7] G. McGunnigle and M.J. Chantler. Segmentation of rough surfaces using reflectance. In *British Machine Vision Conference*, pages 323–332, 2001.

[8] Lin Woan Ning, Yap Keem Siahand, and M. Yusof. Khalid. Design of an automated data entry system for hand-filled forms. In *TENCON 2000. Proceedings , Volume: 1*, pages 162–162, 2000.

[9] N. Otsu. A threshold selection method from grey-level histograms. *SMC*, 9(1):62–66, January 1979.

[10] J.R. Parker. Gray level thresholding in badly illuminated images. *PAMI*, 13(8):813–819, August 1991.

[11] P.Z. Peebles. Probability, random variables, and random signal principles. In *Second Edition, McGraw-Hill*, 1987.

[12] R. Plamondon and S.N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *PAMI*, 22(1):63–84, January 2000.

[13] B. Sankur, A.T. Abak, and U. Baris. Assessment of thresholding algorithms for document processing. In *ICIP99*, page 25PP4, 1999.

[14] I.C. Seong, H. Saito, and S. Ozawa. Shape recovery of book surface using two shade images under perspective condition. *Transactions of the Institute of Electrical Engineers of Japan, Part C*, 117-C(10):1384–1390, 1997.

[15] S. Shetty, M. Shridhar, and G. Houle. Background elimination in bank checks using grayscale morphology. In *IWFHR00*, pages 83–91, 2000.

[16] Y. Solihin and G.C.. Leedham. Integral ratio: A new class of global thresholding techniques for handwriting images. *PAMI*, 21(8):761–768, August 1999.

[17] R.K. Srihari. Cedar online handwritten database. In *http://www.cedar.buffalo.edu/Linguistics/database.html*.

[18] O.D. Trier and T. Taxt. Evaluation of binarization methods for document images. *PAMI*, 17(3):312–315, March 1995.

[19] L. Wang and T. Pavlidis. Direct gray-scale extraction of features for character recognition. *PAMI*, 15(10):1053–1067, October 1993.

[20] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *OptEng*, 19(1):139–144, January 1980.

[21] Y. Yamazaki, Y. Mizutani, and N. Komatsu. Extraction of personal features from stroke shape, writing

pressure and pen inclination in ordinary characters. In *Document Analysis and Recognition, 1999. ICDAR '99. Proceedings of the Fifth International Conference on*, pages 426–429, 1999.

[22] X. Ye, M. Cheriet, and C.Y. Suen. Stroke-model-based character extraction from gray-level document images. *IP*, 10(8):1152–1161, August 2001.

[23] Jin-Yong Yoo, Min-Ki Kim, Sang Yong Han, and Young-Bin Kwon. Line removal and restoration of handwritten characters on the form documents. In *Document Analysis and Recognition. Proceedings of the Fourth International Conference o*, volume 1, pages 128–131, 1997.

[24] B. Yu and A.K. Jain. A generic system for form dropout. *PAMI*, 18(11):1127–1134, November 1996.

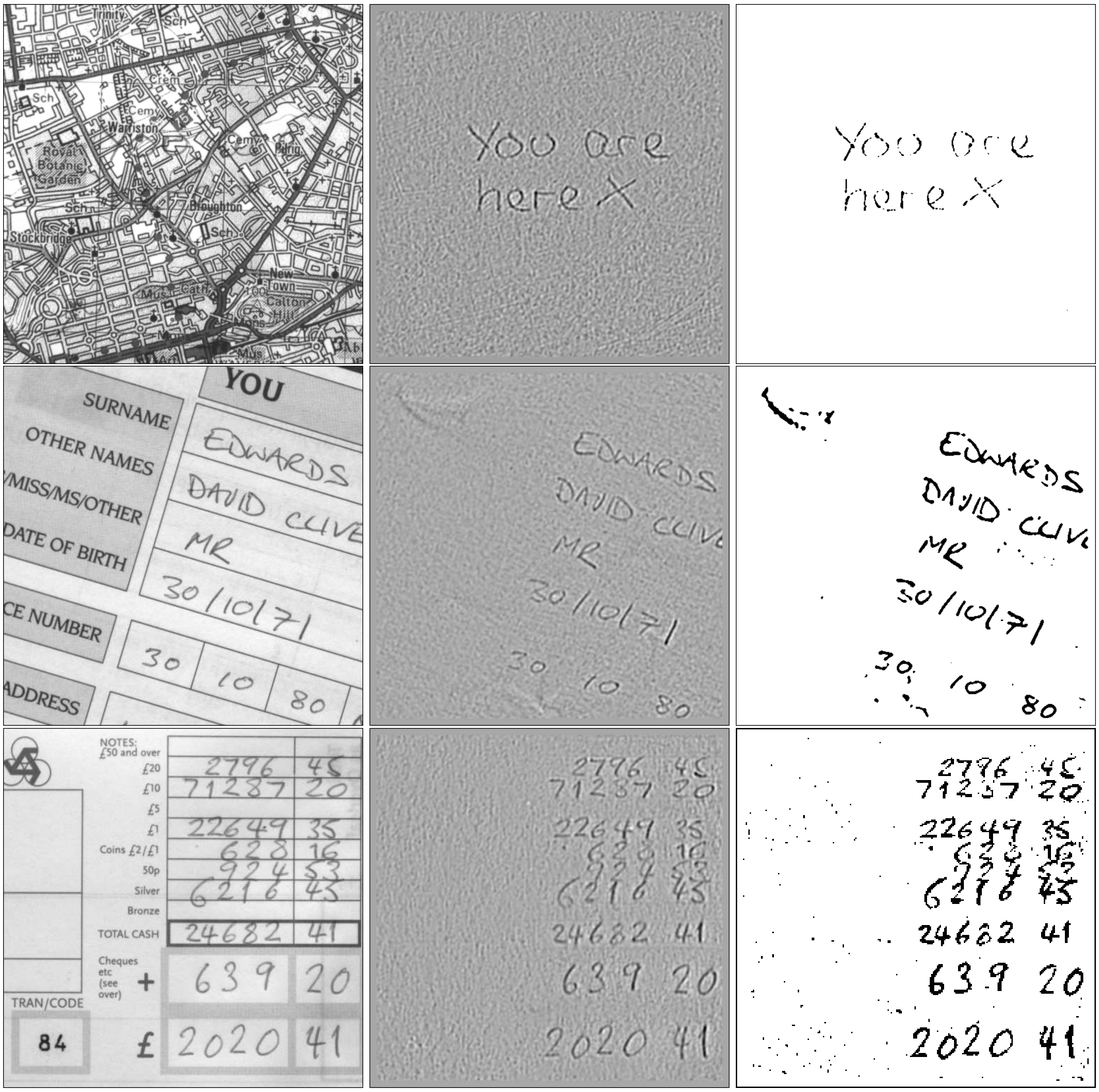

Fig. 12. Examples: ambient images (left column), output of matched filter (centre column) and segmentation (right column)

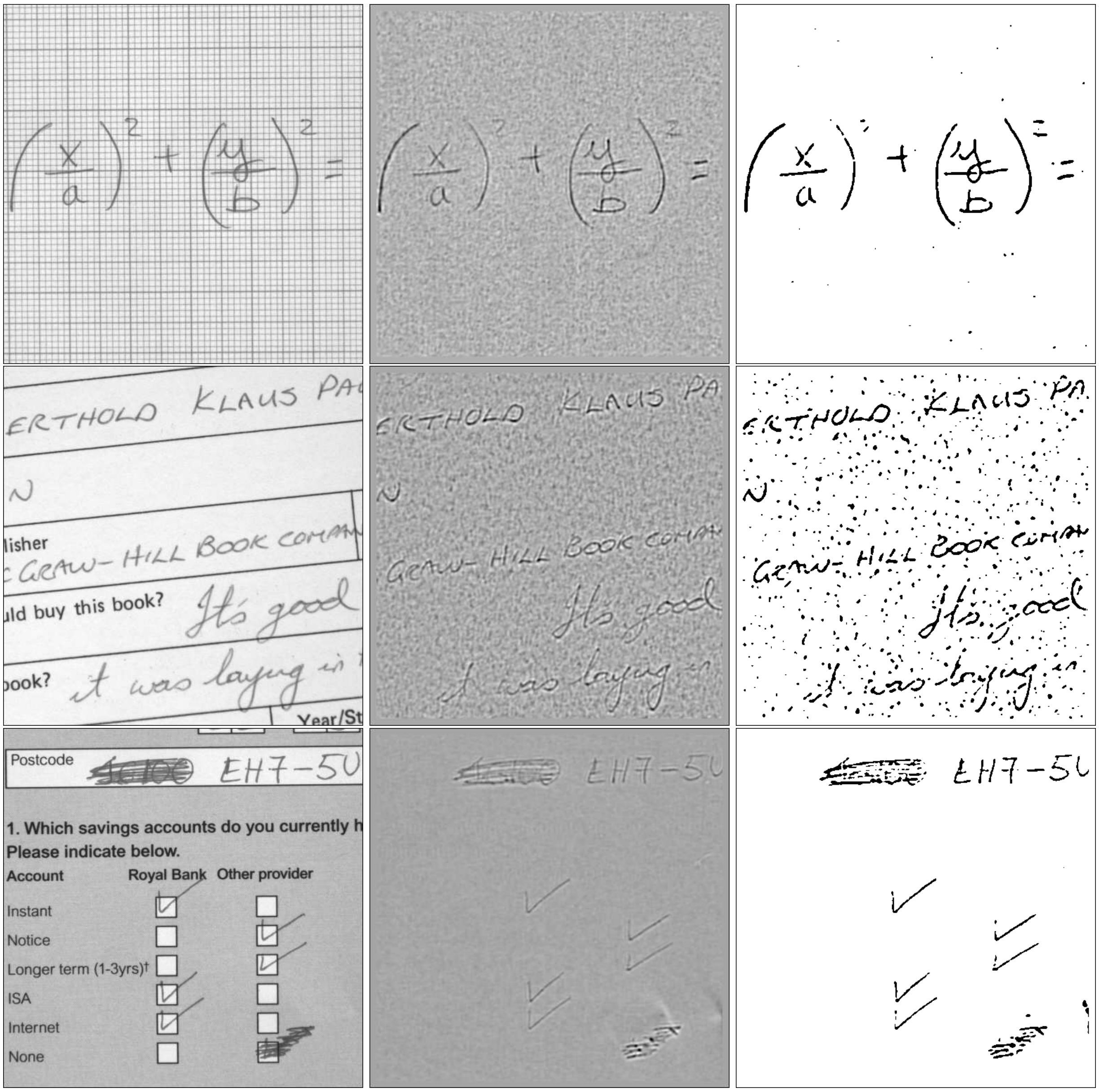

Fig. 13. Examples 2: ambient images (left column), output of matched filter (centre column) and segmentation (right column)