

Original citation:

Bhalerao, Abhir and Reynolds, Gregory. (2014) Ruler detection for autoscaling forensic images. International Journal of Digital Crime and Forensics, Volume 6 (Number 1). pp. 9-27

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/63292>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A note on versions:

The version presented in WRAP is the published version or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Ruler Detection for Autoscaling Forensic Images

Abhir Bhalerao, Department of Computer Science, University of Warwick, Warwick, Coventry, UK

Gregory Reynolds, Pattern Analytics Research Ltd, Solihull, UK

ABSTRACT

The assessment of forensic photographs often requires the calibration of the resolution of the image so that accurate measurements can be taken of crime-scene exhibits or latent marks. In the case of latent marks, such as fingerprints, image calibration to a given dots-per-inch is a necessary step for image segmentation, preprocessing, extraction of feature minutiae and subsequent fingerprint matching. To enable scaling, such photographs are taken with forensic rulers in the frame so that image pixel distances can be converted to standard measurement units (metric or imperial). In forensic bureaus, this is commonly achieved by manual selection of two or more points on the ruler within the image, and entering the units of the measure distance. The process can be laborious and inaccurate, especially when the ruler graduations are indistinct because of poor contrast, noise or insufficient resolution. Here the authors present a fully automated method for detecting and estimating the direction and graduation spacing of rulers in forensic photographs. The method detects the location of the ruler in the image and then uses spectral analysis to estimate the direction and wavelength of the ruler graduations. The authors detail the steps of the algorithm and demonstrate the accuracy of the estimation on both a calibrated set of test images and a wide collection of good and poor quality crime-scene images. The method is shown to be fast and accurate and has wider application in other imaging disciplines, such as radiography, archaeology and surveying.

Keywords: Autoscaling, Forensic Image Analysis, Fundamental Frequency/Pitch Detection, Method, Ruler Detection

INTRODUCTION

In forensic science, many high resolution digital images of crime scenes are often taken with reference to only rulers or scales placed within the frame of the image. The graduation spacing on the calibrated rulers (in centimetre and millimetre spacing, or in inches) are then subsequently used by forensic officers to determine the size

of the crime-scene marks. Fingerprint marks and ballistic marks require precise measurement so that the image features can be used to match marks to criminal records, or otherwise be quantified. This is achieved by manually picking locations of one or more graduation marks to determine the scale. Although it may seem to be easy, it can be time-consuming and is error prone, especially when the image

DOI: 10.4018/ijdcf.2014010102

resolution is insufficient to accurately identify the graduation spacing. Consider for example a photograph with a field of view of 0.5m taken with 10M pixel camera, it will only result in about 5 pixels per mm of spatial resolution. Even when pixel resolution is not a problem, such as when a flat-bed scanner is used, the precise calibration of the scanner still remains, see for example Poliakow et al. (2007).

The general problem of determining size of objects from rulers or scales within the image is common in other disciplines, such as museum archiving, archaeology and medical imaging. Typically, the ruler graduation marks are sought by combining line filtering and line segment grouping. Commonly used approaches are by Hough transform grouping, Illingworth and Kittler (1988), eigen value analysis Guru et al. (2004), or direct image gradient analysis as proposed in Nelson (1994). Herrmann et al. presented a system for measuring the sizes of ancient coins by detecting the graduations on a ruler in the image, Herrmann et al. (2009); Zambanini and Kämpel (2009). Their method used a Fourier transform of the entire image to filter the input image, thus suppressing the image of the coin, to leave the ruler graduation marks. They used a simple method to track along the ridges. Their method is relatively primitive and will only work on simple plain backgrounds when the ruler is presented either horizontally or vertically. They report an accuracy of about 1% in the detection of the graduation marks and a corresponding average error of about 1.19% in the diameter of the measured coins.

Poliakow et al. (2007) reported on a detailed analysis of the problem of calibration of commodity flat-bed scanners for the purpose of digitising large numbers of astronomical plates. Their solution uses a pair of graduated, photolithographically etched glass rulers which are placed along with the item to be scanned (the photographic plate). The rulers appear in the scanned image and can then be used to calibrate each image. They do not elaborate on any automated method to detect the ruler graduations however.

Ruler detection was used by Gooßen et al. (2008) and Supakul et al. (2012) to automatically stitch together overlapping radiographs. They present a four stage ruler recognition method that uses a Radon transform to find the orientation of the ruler. The rulers present in the images act as specialised synthetic landmarks and the method compliments techniques which use anatomical landmarks for registration, e.g. Johnson and Christensen (2002). They then proceed by projecting the ruler graduations onto the ruler line and then autocorrelating this with a template ruler. Their method works by matching a template ruler (and its graduations) to the given image and using optical character recognition to find the graduation numbers. This process allows them to register a number of radiographs with an accuracy of 3mm or less (testing over 2000 image pairs). Lahfi et al. (2003) required the precise location and measurement of a specially designed plastic ruler which was placed in the field of a digital angiographic image during intervention, see also Raji et al. (2002). They used a correlation approach to match a template ruler to the image to enable the precise augmentation of a pre-operative segmented image.

The estimation is performed on a square image block around a chosen point \mathbf{p} . The Ruler Angle is first estimated from the peak of the Power Spectral Density Function of Y_B . Next, the spatial profile through \mathbf{p} is used to estimate the period of the ruler, the Graduation Spacing, λ . The Graduation Phase can be estimated by correlating a sine wave grating at the correct wavelength and orientation with a spatial ruler block. This is achieved by a 2D sine wave model and a line search to refine λ .

Ueda et al. (2005); Baba et al. (2005) presented a method for the detection of scale intervals in images for the general problem of object sizing. Their method is notable because their approach tried to not only detect the graduation interval of the ruler, and hence the image scaling, but tried to locate automatically the ruler in the image. Their approach used a block-based Discrete Fourier Transform (DFT)

to: (1) locate candidate ruler regions; (2) estimate the ruler graduation spacing (frequency) from the candidate regions. Their results work on both bamboo and steel rulers and show errors between 1% and 17% when compared with manually estimates. Their method was shown to work for rulers which are horizontal or vertical only and they made no attempt to locate graduation marks on the ruler. Their estimation scale estimates are not resolved to sub-pixel accuracy which limits the accuracy – an error of 1% or more is oftentimes insufficient for forensic use, such as for comparing fingerprints marks. The paper also does not present any performance analysis for images containing noise.

We present a novel method for ruler estimation that uses a two stage approach that does not make assumptions on the orientation of the ruler in the image or require a template ruler to be known: we first detect the ruler orientation using a 2D DFT taken from a candidate ruler point, and then use an image model to estimate the ruler graduation spacing. The ruler is modelled locally as a 2D sine wave at a given orientation, wavelength and phase. The frequency is robustly estimated using a 1D approach designed to determine the fundamental frequency of music signals, see de Cheveigné and Kawahara (2002); McLeod and Wyvill (2005). The ruler model enables us to refine an initial wavelength estimate to sub-pixel accuracy, which is essential to obtain a precise estimate of the calibrated image resolution in

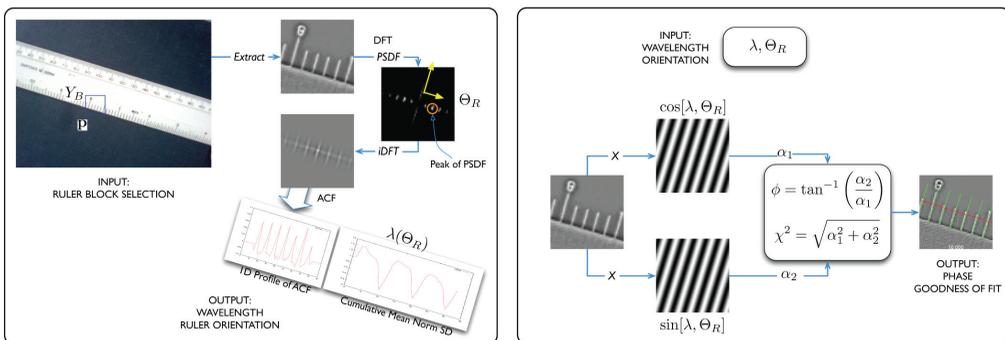
dots-per-inch (DPI). After detailing the ruler estimation method, we show how the position of the ruler can be estimated using a Hough transform approach, where candidate ruler regions are highlighted using the statistics of their local frequency using a block based DFT. We then show results and evaluate their accuracy on a set of calibration images captured using a flat-bed scanner. We illustrate results on a selection of crime scene marks and discuss possible problems and improvements of the technique.

SCALE ESTIMATION

If we assume that we are given a position in the image, \mathbf{p} , near which the ruler is present, the ruler estimation problem is broken down into three steps (illustrated in Figure 1):

1. Estimate the Ruler Angle, Θ_R , which defines the direction of the edge of the ruler, and the Graduation Angle, Θ_G , which define the angles of the graduation. We assume that the two angles are at 90° to each other, i.e. $\Theta_G = \Theta_R + \pi / 2$. An estimate of the ruler directions can be obtained from the local 2D autocorrelation function;
2. Estimate the Graduation Spacing (or wavelength, λ , in pixels per distance unit), which defines the regular spacing between

Figure 1. An overview of the ruler estimation algorithm



successive graduation marks. The algorithm uses the spatial profile of the graduations in the ruler direction through point \mathbf{p} and a pitch estimation technique to find the period. Since the pitch estimate is given to the nearest integer period, i.e. to the nearest pixel, to attain sub-pixel accuracy, we iteratively refine the wavelength estimate using a line search;

3. Estimate the Graduation Phase, ϕ . This calculates the *offset* in pixels, of the graduation marks relative to the given point \mathbf{p} . The ruler graduations are modelled as a linear combination of sine and cosine functions at the estimated wavelength. A linear least squares fit is used to estimate the phase offset of the model. The goodness of fit of this model is also the cost in the wavelength refinement step 2.

For the purposes of autoscaling an image, the essential parameter is λ , since knowing this value to a sub-pixel precision allows the image DPI to be estimated. For example, if $\lambda = 15$ and the graduations are 1/10 inch apart, then the DPI of the image is 150. The ruler/graduation angles allow the accurate estimate of wavelength to occur. The graduation phase is in itself less important, but it is essential for the refinement of λ by the sine wave model fitting (step 3), and valuable as a visual check that the ruler spacing is correct and enables a synthetic ruler to be overlaid on to the input image.

The above steps are detailed below and assume that the input is an image of a given block size, B , taken from a high-pass version of the image: A greyscale image, Y , is high-pass filtered using a un-sharp mask:

$$Y_H = Y - (Y * G(\sigma)) \quad (1)$$

where $G(\sigma)$ is a small Gaussian convolution kernel with standard deviation of σ pixels. This high-pass filtering emphasises the ruler edge and ruler graduation marks in the image and is

essential to enable local graduation angle and phase to be extracted from the analysis of the local image spectra. An image block, Y_B , is extracted from this pre-filtered image centred on user defined ruler point, \mathbf{p}

$$Y_B(\mathbf{x}) = W_B(\mathbf{x})Y_H(\mathbf{x} - \mathbf{p}), \\ B/2 \leq \|\mathbf{x}\| < B/2 \quad (2)$$

where $W_B(\mathbf{x})$ is a suitable window such as a Gaussian or a raised cosine. The windowing is essential to prevent spectral leakage and is the basis of all short-time Fourier analysis Harris (1978).

Ruler/Graduation Angle

The ruler/graduation angle is estimated from the Power Spectral Density Function, PSDF, of the image block around the chosen point, \mathbf{p} . The PSDF is simply the spectrum of the spatial autocorrelation function of a signal and can be calculated by a local Fourier transform, Gonzalez and Richard (2002); Oppenheim et al. (1989):

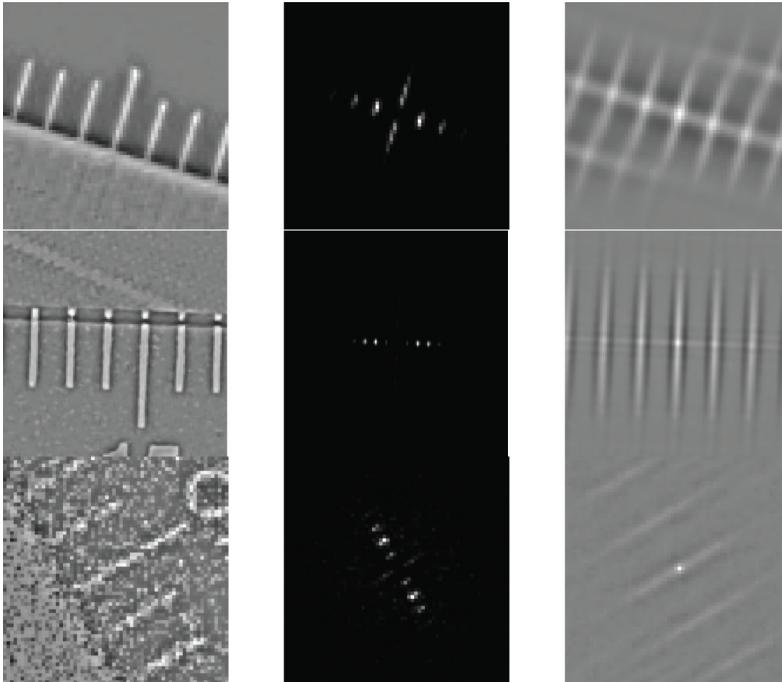
$$PSDF(\omega) = F\{Y_B\}F^*\{Y_B\} \quad (3)$$

where F is the Fourier transform operator and $*$ denotes the complex conjugate. Regions along the ruler graduations will have a distinctive energy spectra with a peak magnitude at the Ruler Angle, see Figure 2. The auto correlation function is the inverse Fourier transform of PSDF, and it can be seen that the ruler edge is suppressed and the ruler graduations dominate. The Ruler Angle is then given by:

$$\Theta_R = \arg\left[\max_{\omega} PSDF(\omega)\right] \quad (4)$$

where \arg takes the angle peak energy coefficient $\omega = [u, v]$ of the centred PSDF, namely:

Figure 2. Ruler blocks (left), their power spectral density functions (centre) and the auto correlation functions. The ruler angle is estimated from the position of the peak magnitude of the PSDF. Bottom row: Input image of imperial ruler at 150 DPI with Gaussian white noise added such that SNR of image is 2.



$$\Theta_R = \arctan\left(\frac{v}{u}\right) \quad (5)$$

Graduation Spacing

The Graduation Spacing is estimated as the period of the image amplitude of the pixels on the line which pass \mathbf{p} through at an angle Θ_R : i.e. the line which bisects the ruler block in a direction parallel to the ruler edge. Ideally, this 1D profile has a box-car profile and is periodic with the desired wavelength, namely the spacing of the ruler graduation marks. However, because of imaging noise, poor spatial resolution, poor quality ruler markings, the ideal box-car shape is not seen. The Autocorrelation Function (ACF) of the ruler block is less affected by noise (see Figure 2), but at a loss of the ruler Graduation Phase – the offset

of the graduations with respect to the selected block centre, \mathbf{p} . A robust estimate of the period of the signal is now required and we apply the YIN algorithm, which was developed for speech and music signals, de Cheveigné and Kawahara (2002).

The classical method to determine the fundamental frequency of a periodic signal is to use the autocorrelation function of the signal, ACF:

$$\text{ACF}_t(T) = \sum_{j=t+1}^{t+W} y(j)y(j+T) \quad (6)$$

Defined as the product of the signal with a delayed version of itself over a window, it has the property of being periodic also. The ACF will have maxima other than at zero lag at multiples of the signal period. The choice of the window size and search method for the

ACF maxima becomes critical in determining the correct fundamental period: if the window is too small, then the period will be incorrect; too big and the period of a harmonic will be wrongly estimated. The YIN algorithm de Cheveigné and Kawahara (2002) uses instead the squared difference (SD) function as its basis:

$$SD_t(T) = \sum_{j=1}^W (y(j) - y(j + T))^2$$

$$= ACF_t(0) + ACF_{t+T}(0) - 2ACF_t(T)$$

which can be written as a sum of ACF terms and two energy terms, one constant and a second that varies with the delay, T . If this term was constant, it would make $SD_t(T)$ an inverted form of $ACF_t(T)$. However, since $ACF_t(0)$ is not constant, the YIN algorithm proposes a normalised form: the Cumulative Mean Normalised Square Difference Function (CSDF), defined as:

$$CSDF_t(T) = \begin{cases} 1, & \text{if } T = 0, \\ SD_t(T) / \frac{1}{T} \sum_{j=1}^T SD_t(j), & \text{else} \end{cases} \quad (7)$$

This normalised form of the squared difference has a number of useful features: it starts at 1 remaining above 1 at low lag periods until the squared difference falls below the average; its minima indicate the periods of the fundamental and associated harmonics; there is no need to set an upper-bound on the expected maximum period in the search, McLeod and Wyvill (2005). In our application, we have found the first true minimum is the period of the fundamental. This is the case, even in the poorest signal to noise ratio images.

Figure 3 presents three examples of estimating the wavelength using the YIN algorithm. The plots in the left column are the 1D spatial intensity profile of rulers taken across the ruler

block at the estimated Ruler Angle for the example ruler images shown in Figure 2. The corresponding plots in the right hand column show the Cumulative Mean Normalised Difference Functions, using $CSDF(T)$ in Equation 7. The periods are approximately, 10, 21 and 15 pixels. The results for the noisy profile, lower plots in Figure 3, are particularly impressive given that the signal to noise ratio of the source data is 2 (see Figure 2).

Graduation Phase

To estimate the Graduation Phase, which is the shift of the ruler graduation relative to the coordinates of the ruler block, we correlate the spatial domain ruler block with a pure sine wave grating. The grating is orientated such that the principal variation is in the direction of the ruler edge, Θ_R . We can construct the grating by:

$$M(x, y; \Theta_R, \lambda) = \alpha_1 \cos[F(x, y; \Theta_R, \lambda)] + \alpha_2 \sin[F(x, y; \Theta_R, \lambda)] \quad (8)$$

$$F(x, y; \Theta_R, \lambda) = \frac{2\pi}{\lambda} (x \cos(\Theta_R) + y \sin(\Theta_R)) \quad (9)$$

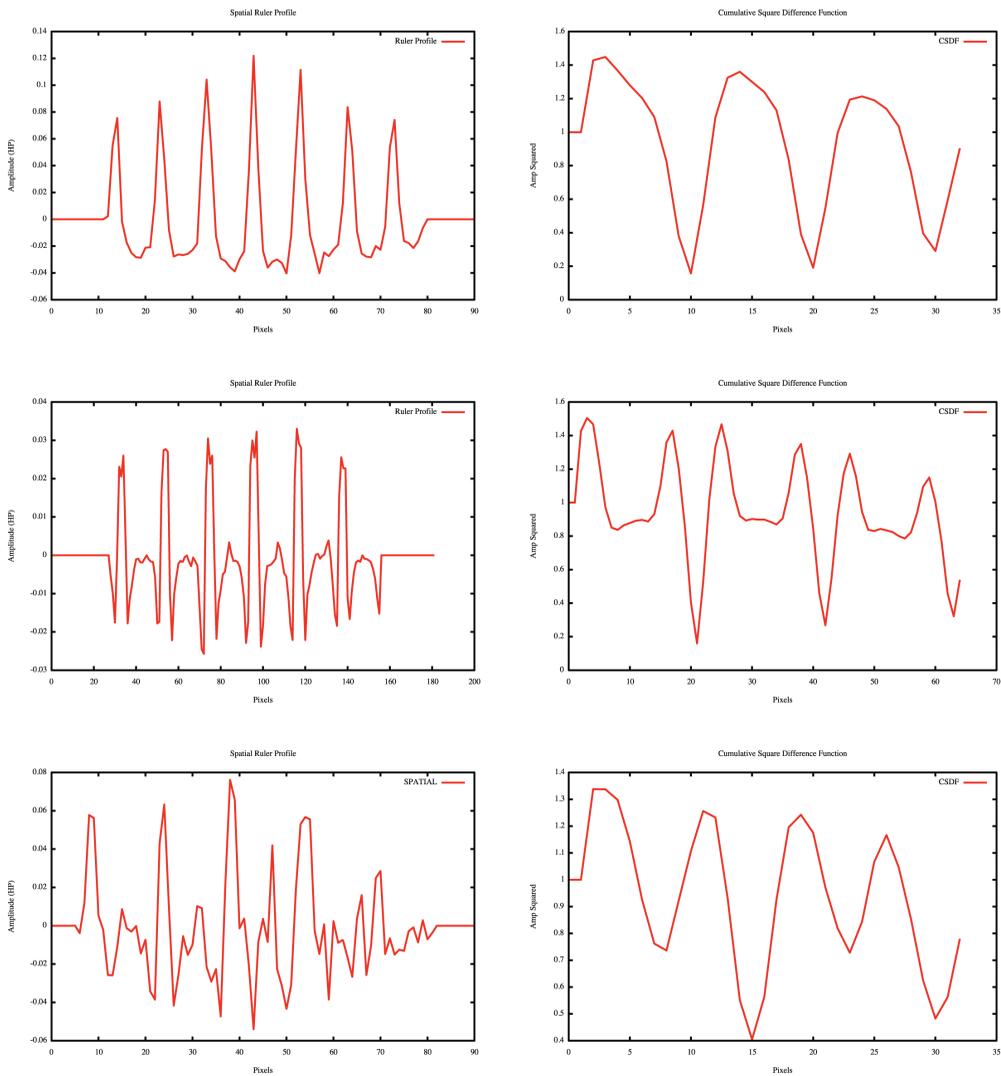
where F is the spatial frequency $2\pi / \lambda$ cycles per pixel in the ruler direction, Θ_R .

There are two amplitudes, $\alpha = \{\alpha_1, \alpha_2\}$ to be determined, for the sine and cosine functions. The grating in the block will then have a phase (in pixels) of:

$$\phi = \frac{\lambda}{2\pi} \tan^{-1} \left(\frac{\alpha_2}{\alpha_1} \right) \quad (10)$$

and the resultant wave an amplitude, $\sqrt{\alpha_1^2 + \alpha_2^2}$. Equation 8 is linear in the unknown amplitudes and therefore these can be estimated using LinearLeastSquaresto minimize $(Y_B(\mathbf{x}) - \mathbf{M}(\mathbf{x}; \alpha))^2$

Figure 3. Wavelength estimation using cumulative mean normalised difference functions. Plots of ruler spatial profiles taken at estimated ruler angles from images in Figure 2 and their corresponding CSDF functions. The period is when $CSDF_t(T)$ is a minimum. The bottom figure is taken of a 150DPI image of an imperial ruler with SNR 2, produced by adding Gaussian white noise. The period is estimated to be 15 pixels for 1/10 inch.

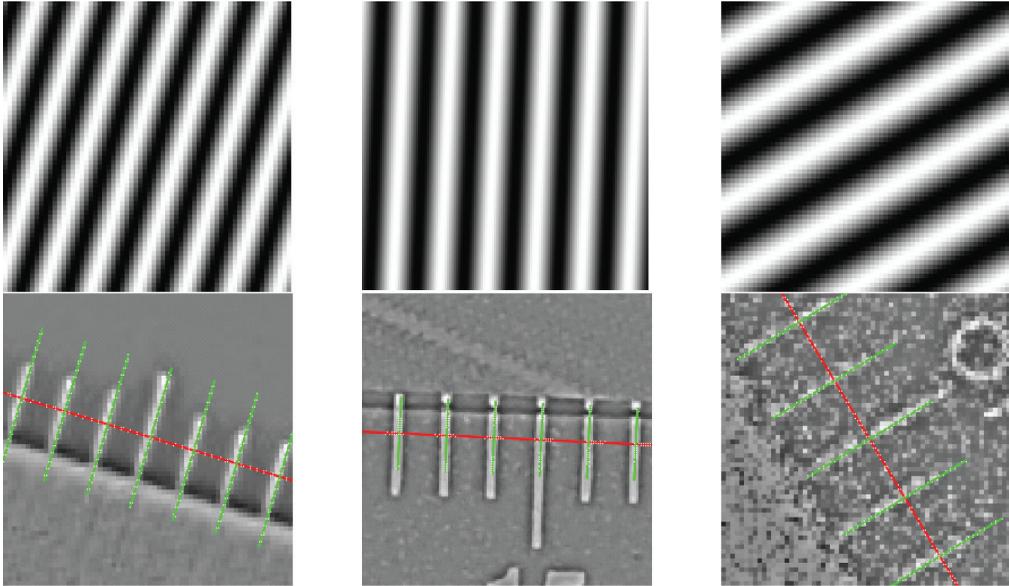


w.r.t α . Figure 4 illustrates the sine wave model for $\phi = 0$ for ruler blocks from Figure 2. The ruler estimate is overlaid on to the last example, showing its accuracy in both wavelength and phase, despite the noise.

Sub-Pixel Wavelength Estimation

The wavelength estimation using the CSDF is to the nearest integer period, T . For practical purposes, images containing a ruler normally

Figure 4. A sine wave grating is used to estimate the phase of the ruler (its shift relative to the centre of the block). The top row illustrates the grating with zero phase, with the estimated wavelength and orientation. Bottom row shows the original data which is correlated with the model to estimate phase. The ruler estimate is overlaid on the examples showing its accuracy in both wavelength and phase.



have ruler graduations at fractional spacings, so it is important to refine the initial estimate given by the YIN algorithms. A convenient and robust way is to use the sine wave model fitted above to estimate the Graduation Phase, ϕ .

The goodness of fit or the amplitude, $\sqrt{\alpha_1^2 + \alpha_2^2}$, can be used as a cost for a line search around the integer estimate, λ . In our implementation, we use the Nelder-Mead Simplex search. The algorithm proceeds as shown in Algorithm 1. Here, *PHASE_AMP()* implements the sine wave model fitting using linear least-squares, and *SIMPLEX_PROPOSAL()* uses the Nelder-Mead method to make searches around a given starting Simplex. The cost is taken as negated amplitude of the model fit, and a tolerance on this cost is used as the stopping criterion. Convergence is fairly rapid, requiring between 10 and 30 evaluations, which can be almost in-

stantaneous on small image regions using a non-optimized implementation in C++ on a standard PC.

AutoRuler Detection

Being able to identify the scale of an image without the user picking a likely ruler region is a desirable feature. For example, for the purposes of rapidly rescaling many thousands of images for batch processing, such as for archiving, see for example Poliakow et al. (2007). The commercial forensic software CSipix (iSYS Corp., St. John's, NL, Canada) has an 'AutoScaling' feature. In their demonstration video, a forensic image ruler is detected automatically and it is claimed that the measurement units (imperial or metric) do not have to be indicated. The user is prompted to confirm that the graduation marks have been correctly found. The method is

Algorithm 1. Refine(λ), which improves on an estimate of the wavelength λ

```

function Refine( $\lambda$ )
   $t=0, \lambda(t)=\lambda, \alpha(t)=0$ 
  repeat
     $\lambda(t') = \text{SIMPLEX\_PROPOSAL}(\lambda(t))$ 
     $\{\varphi(t+1), \alpha(t+1)\} = \text{PHASE\_AMP}(\lambda(t'), \Theta_R)$ 
    if  $-\alpha(t+1) < -\alpha(t)$ 
       $\lambda(t+1) = \lambda(t')$ 
       $t=t+1$ 
    endif
  until  $\alpha(t+1) - \alpha(t) > \varepsilon$ 
return  $\{\lambda(t+1), \varphi(t+1)\}$ 

```

proprietary and it is not possible to judge how robust or accurate it over a range of different input types.

To our knowledge, only Ueda et al. (2005) have proposed how this could be achieved: by looking at the frequency content of blocks over the whole image. Their approach operates over the whole image and provided a local search method to group together potential horizontal or vertical blocks which satisfied their ruler block test. This method relies greatly on a robust starting point: a distinctive ruler block, and the local continuity of neighbouring image blocks which pass the ruler detection threshold. If the threshold is set too low, non ruler-blocks enter the ruler estimation and an over segmentation will result. If the threshold is set too high, no connected set of blocks are discovered to be part of any ruler in the image. Furthermore, they did not provide methods for arbitrary oriented rulers or show how more than one ruler edge could be detected. They also did not characterise the quality of the method against noisy images, typical of many applications. Our method addresses both problems in a novel way: by estimating a robust Ruler Likelihood Map based on statistics of the local PSDFs and by using a Hough Transform approach for grouping potential ruler regions in a given parametric space. We use an entropy based feature to weight the accumulation of 'line segments' into the Hough accumulator space.

Hough Transform Grouping of Likely Ruler Blocks

The standard Hough Transform (HT), Duda and Hart (1972); Illingworth and Kittler (1988), for estimating the positions of line in an image can be parameterised by considering lines (ρ, θ) that pass through the set of image points (x, y) :

$$\rho = x \cos \theta + y \sin \theta \quad (11)$$

One way to build the HT accumulator space is by discretising the parameter space (ρ, θ) to a set of offsets $[\rho_{min}, \rho_{max}]$ and a set of angles $[0, \pi]$ in steps of $\Delta\Theta$. All candidate line points (x_i, y_i) are then binned by solving the line equation for all possible Θ_i , namely:

$$\begin{aligned} r_i &= P(x_i \cos(t\Delta\theta) \\ &+ y_i \sin(t\Delta\theta)), t \\ &= 0.. N_\theta, N_\theta = \frac{\pi}{\Delta\theta} \end{aligned} \quad (12)$$

where:

$$P(\rho) = \frac{1}{N_\rho} \left[\frac{\rho - \rho_{min}}{\rho_{max} - \rho_{min}} \right] \quad (13)$$

for N_ρ bins. We can choose to do a weighted accumulation into the Hough space:

$$H[r, t] = \sum_{j \in I[r, t]} w_j \quad (14)$$

where the index set $I[r, t]$ is all indices i which fall into bin $[r, t]$. For images, w is typically chosen to be some certainty feature of the line, such as its gradient magnitude. In this way, the HT accumulation concentrates only on strong line segments and is somewhat robust to noise. Of course, after the accumulation is complete, the maxima in the accumulator space indicate groupings of line segments at the corresponding parameter values.

For the purposes of our AutoRuler detection, we adapt this HT parameterisation to estimate the likely position of rulers in the image and use the output of the first stage of the ruler estimation (Scale Estimation section) on a set of windowed, overlapping DFT blocks across the input image.

The image is divided into a set of square blocks of size B which overlap by 50%. If an image is of size $W \times H$, there will be, $2W / B + 1$ and $2H / B + 1$ blocks in each of the two image dimensions. We can denote the centres of block coordinates by (x_i, y_i) .

For each block, we estimate its orientation Θ_i , by using the direction of the peak magnitude of its PSDF. Note that the image has to be high-pass filtered before being divided into blocks and each block is windowed by a cosine squared function centred on the block. The number of angular steps is fixed to N_θ , and Equation 12 used to determine the range of offsets $[\rho_{min}, \rho_{max}]$ and an empty accumulator space, $H[r, t]$, with $N_\theta \times N_\rho$ bins allocated.

Entropy Weighted Hough Accumulation

The accumulator space is filled by accumulating the entropy of the PSDF of each DFT block, i with domain Ω :

$$w_i = -\sum_{\Omega} p(\omega) \log_2 p(\omega) \quad (15)$$

$$p(\omega) = PSDF(\omega) / \sum_{\Omega} PSDF(\omega)$$

where $\omega = [u, v] \in \Omega$ are the frequency coordinates of the magnitude spectrum $[DFT]$. The entropy of the spectrum captures the ‘peakyness’ of the spectrum: ruler blocks are more peaky and have a smaller entropy than background or other regions.

Note that we do not need to sweep over θ as we have an estimate, Θ_i , for each input block. Also, the resulting accumulation will have maxima where the ruler blocks group together. By keeping track of which blocks accumulate to which bins, it is possible to estimate the end-points of the ruler on the image. Remembering that $I[r, t]$ is the set of input block indices which fall into any Hough accumulator bin $[r, t]$, we can calculate an expected value for the ruler angle $\hat{\theta}$ and its offset, $\hat{\rho}$, by taking the weighted averages (summing over $i \in I[r, t]$):

$$\hat{\theta} = \frac{\sum_i w_i t_i \Delta\theta}{\sum_i w_i} \quad (16)$$

and similarly for the offset:

$$\hat{\rho} = \frac{\sum_i w_i r_i}{\sum_i w_i} \quad (17)$$

We do this for the peaks of the accumulator space and therefore obtain a sub-block estimate of the auto-ruler parameters. From this, we can choose one of more points along this line as inputs to our Ruler Estimation of Graduation Spacing.

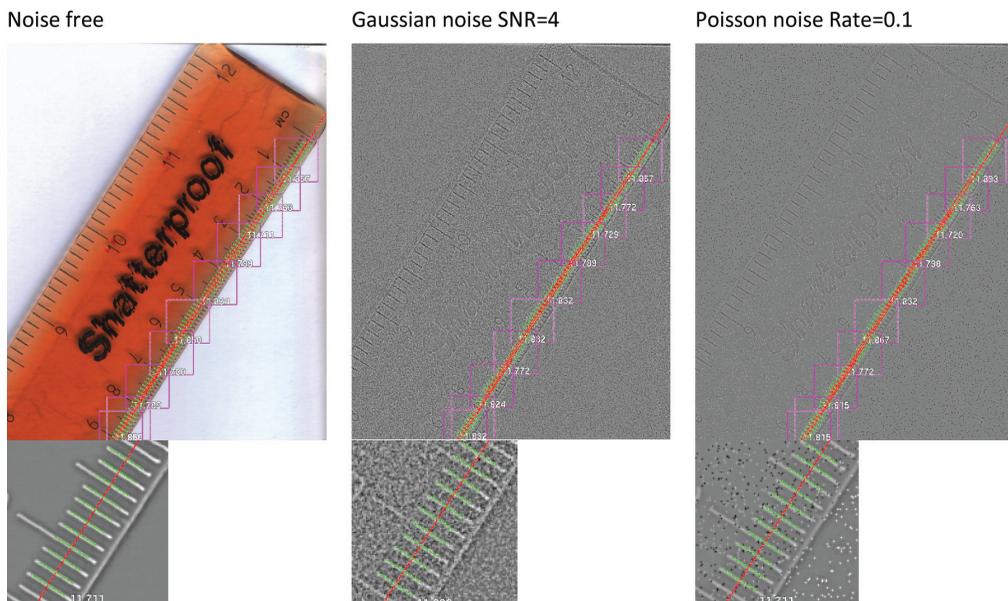
RESULTS AND DISCUSSION

The accuracy of the presented scale estimation and AutoRuler detection algorithms was tested by a series of experiments on ruler images scanned using a flat-bed scanner, at DPI resolutions varying from 150, 300, 600 and 1200, to which was added either Gaussian white noise or Poisson noise. To assess the ruler scale detection accuracy on real scenes-of-crime images, we compared the automatic scale estimation with a hand-marked ruler line, which was defined by carefully selecting two end-points. The pixel difference error, expressed as a percentage of the total marked ruler length, was then tabulated (see Table 3). Finally, we used the Hough based AutoRuler detector on the scenes-of-crime images to detect the location of the ruler graduations, these results were only qualitatively compared (see Figure 9).

Scale Estimation Accuracy on Scanned Ruler Images

We scanned the same set of plastic rulers using a desktop flat-bed scanner (Mustek, BearPaw, 1200 CU Plus) at a set of DPI values (150, 300, 600, 1200) to produce images illustrated in Figure 5. We then manually selected a set of points where the scale estimates were to be made (approximately 10 per image) and then ran a batch program to tabulate the average ruler estimates at differing levels of two types of noise: additive Gaussian white noise, and Poisson rate noise. The Gaussian white noise was added to the pre-processed, high-pass filtered input image, with zero-mean, and a noise variance such that the output image signal-to-noise ratio, measured by:

Figure 5. Illustration of results on scanned ruler at 300 DPI (11.811 Pixels/mm) used to produce accuracy tables. The ruler blocks are manually marked and the estimation repeated for different types of noise (Gaussian white noise and Poisson noise), at varying rates. Example illustrates results on noise free (left); noise value of SNR = 4 (middle); and Rate = 0.1 (right).



$$\text{SNR} = 10 \log_{10} \left[\frac{\text{var}(I)}{\sigma^2} \right] \quad (18)$$

where $N(0, \sigma^2)$ represents the noise process. Table 1 shows the results for SNR values in the range 32 to 1. We repeated the tests, this time adding Poisson noise (see Table 2), to produce speckle typically seen in forensic images of fingerprints using Ninhydrin chemical markers. A random Poisson variate is used to decide whether a given pixel should be subjected to noise by setting its amplitude to either the minimum or maximum of the range of the image, thus producing a ‘spike’ at that position. The Poisson variate’s mean (or rate) is set as a proportion of pixels: so a rate of 0.1 would result in roughly every 10th pixel incurring noise. Examples of the effect of the two noise processes used are shown in Figure 5.

The results show that scale estimation at lower DPIs is more affected by noise than higher DPIs but the performance is still acceptable (<0.5%) even at an SNR of 8 (at 150 DPI). For DPIs greater than 150, the noise has little or no effect on the accuracy of the scale estimation. Poisson noise has a greater detrimental effect on the scale estimation. Noise rates of 0.25 or higher affect the estimator at greater

DPIs, although this implies approximately $\frac{1}{4}$ of the data being speckle, which is unlikely.

Scale Estimation Accuracy on Scene-of-Crime Rulers

To compare the automatic scale estimation on scene-of-crime photographs, we measured the error between the automatic scale estimation and manually delineated rulers. Note that the manual marking was done by carefully placing two end points at selected positions on the ruler to obtain a pixel distance. The automatic scale estimation was run by choosing a suitable block size, B , for the estimator (such that the region covered a portion of the graduation marks along the edge of a given ruler). Note that the automatic process did not rely on the user being particularly careful about the position chosen, whereas for the manual process, the end point had to be picked precisely at the start and end of certain graduation marks. It was also critical to ensure that the manually drawn line was closely parallel to the ruler edge. We illustrate the results on four examples in Figure 6. The automatic estimates are shown in the boxes, below which is the estimated pixel spacing in white letters. The right-hand image shows the

Table 1. Results on metric scale estimation of scanned rulers at different DPI and varying amounts of Gaussian white-noise. The four resolution columns equate to DPI resolutions of 150, 300, 600 and 1200 respectively. Percentage error in estimate given in brackets beside each Pixels/mm entry. The worst errors are seen for the poorest resolution (150DPI). All other entries are less than |0.1|%.

SNR	Resolution – Pixels/mm (Error%)			
	5.906	11.811	23.622	47.244
∞	5.912 (0.100)	11.807 (-0.040)	23.618 (-0.020)	47.234 (-0.023)
32	5.913 (0.113)	11.812 (0.004)	23.613 (-0.042)	47.241 (-0.007)
16	5.933 (0.459)	11.807 (-0.040)	23.621 (-0.005)	47.237 (-0.015)
8	5.922 (0.273)	11.808 (-0.026)	23.613 (-0.042)	47.238 (-0.014)
4	5.799 (-1.817)	11.822 (0.092)	23.602 (-0.086)	47.241 (-0.007)
1	3.823 (-35.270)	11.821 (0.077)	23.618 (-0.020)	47.238 (-0.014)

Table 2. Results on metric scale estimation of scanned rulers at different DPI and varying amounts of Poisson noise. The four resolution columns equate to DPI resolutions of 150, 300, 600 and 1200 respectively. Percentage error in estimate given in brackets beside each Pixels/mm entry. Poisson noise at these rates has a more adverse affect on the estimator with 150DPI images performing worst as the rate parameter is a per-area value. Acceptable estimates of the ruler spacing is obtained at higher resolutions when the Poisson noise has rates lower than 0.25.

Rate	Resolution – Pixels/mm (Error%)			
	5.906	11.811	23.622	47.244
0.063	5.912 (0.100)	11.807 (-0.040)	23.618 (-0.020)	47.234 (-0.023)
0.125	5.525 (-6.458)	11.808 (-0.026)	23.639 (0.068)	47.238 (-0.014)
0.250	5.170 (-12.457)	11.817 (0.048)	23.613 (-0.042)	47.230 (-0.031)
0.500	3.690 (-37.531)	11.839 (0.235)	23.625 (0.010)	47.227 (-0.036)
1.000	4.150 (-29.727)	15.114 (27.966)	23.611 (-0.049)	47.276 (0.066)
2.000	3.092 (-47.649)	11.582 (-1.947)	32.611 (38.053)	47.224 (-0.044)

Table 3. Table comparing error between ruler scale estimates and manually marked ruler line (see Figure 6). Note that ruler in CS 003 is imperial and has 16 divisions per inch.

Ruler ID	Scale Estimate (Pixels)	Manual (Pixels)	Error (%)	DPI
CS 001	$40 \times 10.149 = 405.96$	405.173	-0.19	257.8
CS 002	$50 \times 29.547 = 1477.35$	1467.490	-0.67	521.9
CS 003	$130 \times 14.352 = 1865.76$	1869.56	0.20	351.6
CS 004	$450 \times 7.10 = 3195.00$	3152.32	1.26	173.0

processed block with the graduation marks overlaid. The manually marked ruler line is shown in yellow, with the total pixel distance annotated below in white letters. For the metric rulers (CS 001, CS 002 and CS 004), the graduation spacing is 1mm. Note that for image CS 004, the lower (Imperial) graduations are in $\frac{1}{16}$ of an inch.

To assess the accuracy of the results, we calculated the percentage of the pixel error ratio (see Box 1) the assumption being that the manual distance to be the 'true' distance. We can see that the results are favourable with an average error of about 0.5% (CS 002 has an error of 10 pixels in about 1500). Note also that the effective DPI of these photography images are between 175 and 500 DPI, with 175 DPI being

Box 1.

$\frac{\text{Number of Graduation Marks} \times \text{Scale} - \text{Manual distance}}{\text{Manual distance}} \%$	(19)
--	------

Figure 6. Crime-scene rulers. The AutoRuler results are compared with manually marked ruler line (in yellow). Pixel error estimates given in Table 3.



at the lower end of what is possible with the higher-resolution digital cameras and resulting in the worst error of about 1.26% (CS 004).

Figure 7 illustrates two issues with scale estimation, one which is related to use of flexible ruler markers on non-planar surfaces (ruler tapes, adhesive markers), and the other due to the fact that ruler images can be distorted by projective transformation (such as elation): where the ruler does not lie on a plane which is parallel to the camera’s viewplane. In Figure 7 we show both cases. For the non-planar ruler (tape), the direction of the graduation marks become skewed with a commensurate error in the mark spacing. A projective transformation

results in linear variation in the spacing along the ruler. The latter is a common issue as often the camera is hand-held and it is hard to hold it perfectly parallel to surface on which the crime scene mark appears and it would require some other technique, such as using right-angled rulers to find the projective transformation (i.e. the planar homography) to a proportional scale.

AutoRuler Detection on Scenes-of-Crime Images

To demonstrate the effectiveness of the Hough-transform based auto-ruler detection we show the results on the same set of crime-scene

Figure 7. Distorted rulers. CS 005 is distorted because of projective geometry effects as the camera viewplane is not parallel to the ruler plane resulting in a linear change in wavelength along the edge of the ruler. In CS 006, the ruler is a flexible tape and causes the ruler directions to be angled.

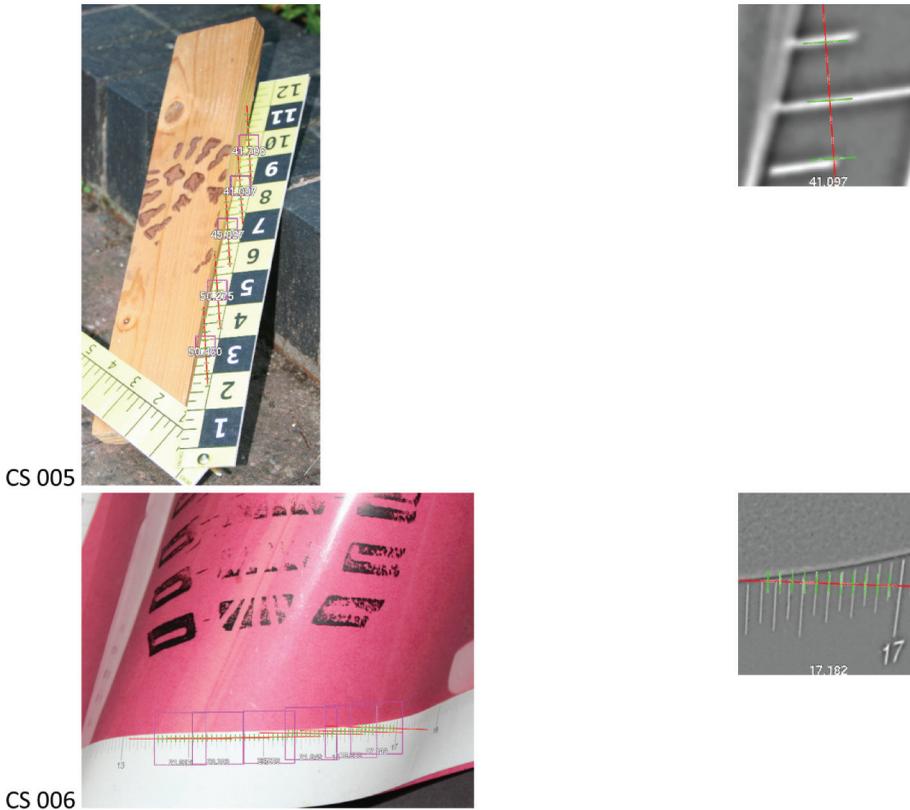


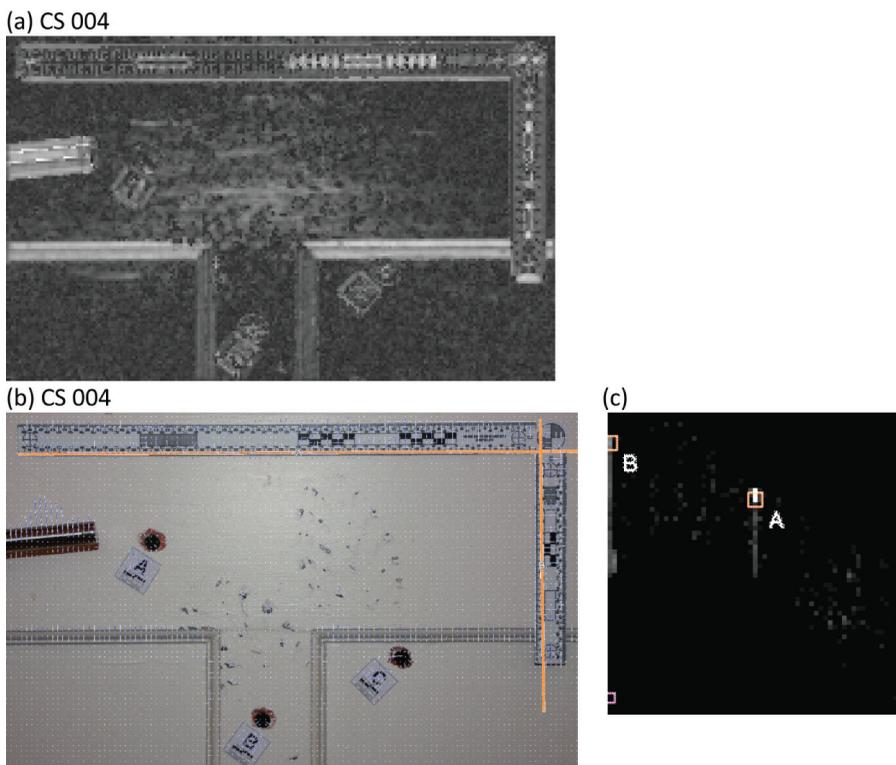
image (CS 001 to CS 004). In each case, we only specified the ruler block size to generate the results. Figure 8(a) shows the entropy map (Equation 15) on CS 004. The values of this image act as weights to the blocks accumulated in the Hough space (Equation 17). In (b), the blue vectors show the estimated graduation angle in each block, where for non-ruler blocks, these will be invalid. The Hough accumulator space is shown in (c), where the first two peaks have been highlighted as A and B, and the two lines they represent are shown in orange on (b). Note that in such a visualisation, it is easy to explore the Hough space by selecting other accumulator bins to see to which image lines they belong.

More illustrations are given in Figure 9, and again the principal Hough accumulator peaks correspond well to the ruler edges.

CONCLUSION

We have detailed algorithms for automated detection of rulers in images and accurately performing scale detection: the scaling of an image or determining its resolution in DPI from the presence of a ruler or scale in the image. Scale detection in images is an important problem in forensic imaging where scenes-of-crime marks such as ballistic holes, footprints

Figure 8. AutoRuler features using entropy and Hough accumulation on CS 004: (a) the PSDF Entropy map which is used to weight the block based Hough accumulation; (b) blue vectors mark the estimated graduation angles in each block; (c) Hough accumulation space (weighted by (a)) where bright regions are significant ruler lines. A and B mark the two largest peaks which correspond to the principal edges of the right-angle ruler.



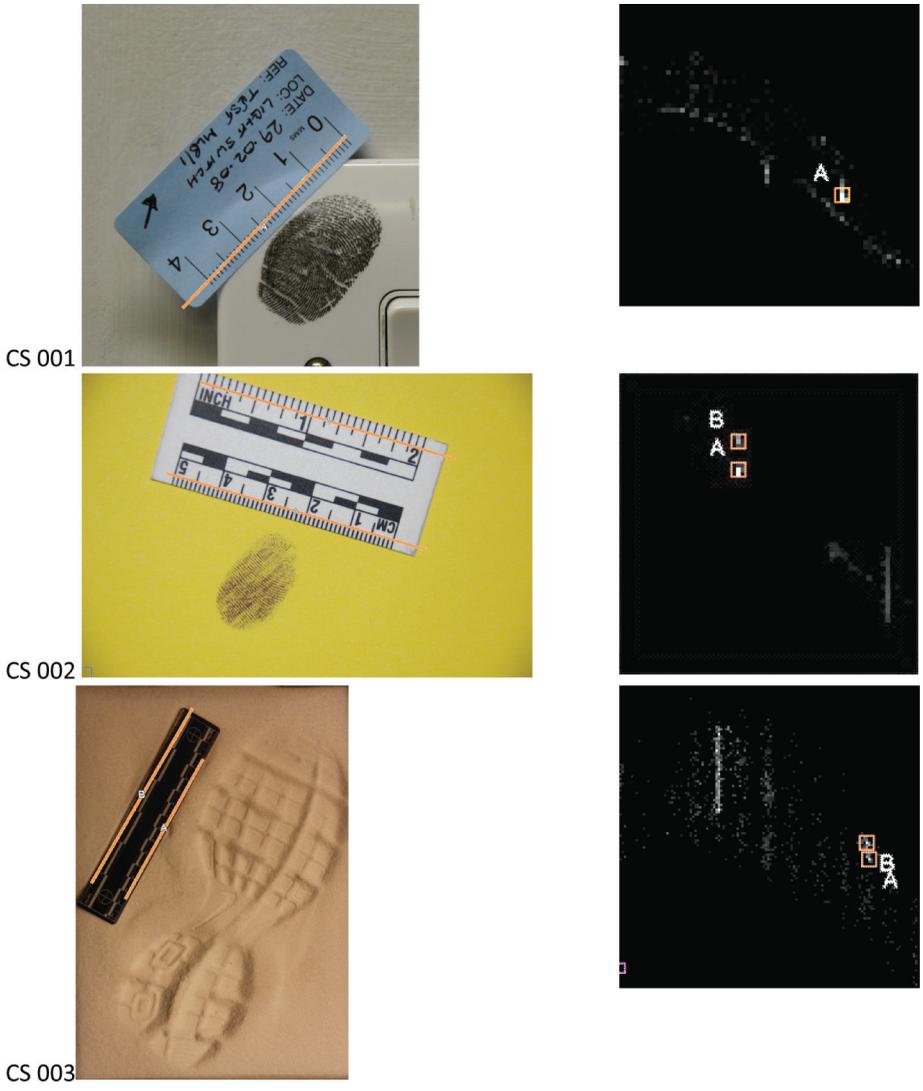
or fingerprints, have to be correctly scaled or rescaled so they can be identified. Other applications of scale estimation is in medical imaging (radiography), archaeology and archiving of artefacts. Although the ruler scale can be manually estimated, accuracy and repeatability can be hard to achieve, and if the hundreds or thousands of images have to be processed by hand, the task can be laborious.

In this work, we have defined the steps of the problem and proposed solutions to them: locating the orientation of the ruler graduation marks, estimating the spacing between them and their offset, and also automatically finding the location of rulers in an image. Our method uniquely combines 2D local Fourier analysis, using the Power Spectral Density Function

(PSDF), and the YIN pitch estimator for 1D signals (de Cheveigné and Kawahara (2002)) that works on the normalised autocorrelation of gradation mark profile. We presented a novel method to refine the discrete wavelength estimate to provide sub-pixel accurate resolution and went on to demonstrate its robustness in the presence of Gaussian and Poisson noise.

A second principal contribution of this work is an extension of the block-based PSDF analysis to produce a Ruler Likelihood Map. When used as a weighting in a Hough type accumulation, it is able to extract out the parameters of one or more ruler edges within an image. This AutoRuler method could enable digitisation and archival projects to autoscale images in large numbers without manual intervention.

Figure 9. Examples of AutoRuler estimation on crime-scene rulers. The right hand image beside each ruler image is the Hough accumulator space. A and B mark the principal peaks which correspond to the edges of the rulers in the image.



Our methods is sensitive to the size of the block used for local analysis, and this has to be chosen such that the ruler graduation marks are well represented within the block region of interest. This is an issue for both the scale estimation but less so for the AutoRuler detection based on Hough accumulation. If the relative size of the ruler to the image is known and remains constant, the block size need only

be specified for the first image of a given set. To alleviate the need to specify the block size, one approach may be to successively test multiple block sizes, e.g. by forming a pyramid of the input image or by building a scale-space, and taking the wavelength estimate which is stable over scale changes. A second issues with our method is the problem of projective geometry warping of the ruler relative to the imaging

plane. This was illustrated in one example, in Figure 7, and shown to create linear variations in the resolution estimates. A way to deal with this sort of image distortion is to use right-angled rulers and un-warp the imaging geometry by an estimated planar homography.

The noise analysis and the comparisons of the scale estimation with manual estimation shows that our ruler scale estimation is accurate in most cases to below 0.5%. This is adequate for forensic use, such as fingerprint mark digitisation where 0.5% of 500DPI image will create a 2.5 pixel offset error; for minutiae matching in fingerprint identification, this error can be tolerated by fingerprint examiners and latent matching systems.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Pattern Analytics Ltd., UK.

REFERENCES

- Baba, T., Ueda, K., & Nakagawa, Y. (2005). Detection of scale interval on a ruler in digital image. In *Proc. Data Engineering Workshop (DEWS 2005)*.
- de Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*(4), 1917–1930. doi:10.1121/1.1458024 PMID:12002874
- Duda, R. O., & Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, *15*(1), 11–15. doi:10.1145/361237.361242
- Gonzalez, R. C., & Richard, E. (2002). *Digital image processing*. Prentice Hall Press.
- Gooßen, A., Schlüter, M., Hensel, M., Pralow, T., & Grigat, R.-R. (2008). Ruler-based automatic stitching of spatially overlapping radiographs. In T. Tolxdorff, J. Braun, T. M. Deserno, A. Horsch, H. Handels, & H.-P. Meinzer (Eds.), *Bildverarbeitung für die Medizin 2008, Informatik aktuell* (pp. 192–196). Springer Berlin Heidelberg. doi:10.1007/978-3-540-78640-5_39
- Guru, D., Shekar, B., & Nagabhushan, P. (2004). A simple and robust line detection algorithm based on small eigenvalue analysis. *Pattern Recognition Letters*, *25*(1), 1–13. doi:10.1016/j.patrec.2003.08.007
- Harris, F. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, *66*(1), 51–83. doi:10.1109/PROC.1978.10837
- Herrmann, M., Zambanini, S., & Kampel, M. (2009). Image based measurement of ancient coins. In *Proc. of the 37th Conference on Computer Applications and Quantitative Methods in Archaeology* (pp. 117–121).
- Illingworth, J., & Kittler, J. (1988). A survey of the hough transform. *Computer Vision Graphics and Image Processing*, *44*(1), 87–116. doi:10.1016/S0734-189X(88)80033-1
- Johnson, H. J., & Christensen, G. E. (2002). Consistent landmark and intensity-based image registration. *Medical Imaging. IEEE Transactions on*, *21*(5), 450–461.
- Lahfi, Y., Lemoine, J., Raji, A., & Boudghene, F. (2003). An image matching approach applied in the control of the case of endovascular prosthesis positioning. *International Congress Series, 1256*, 1093–1101. doi:10.1016/S0531-5131(03)00274-7
- McLeod, P., & Wyvill, G. (2005). A smarter way to find pitch. In *Proc. International Computer Music Conference* (pp. 138–141).
- Nelson, R. (1994). Finding line segments by stick growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(5), 519–523. doi:10.1109/34.291445
- Oppenheim, A. V., Schaffer, R. W., & Buck, J. R. et al. (1989). *Discrete-time signal processing* (Vol. 2). Englewood Cliffs, NJ: Prentice Hall.
- Poliakow, E. V., Poliakow, V. V., Fedotova, L. A., & Tsvtkov, M. K. (2007). High-precision measuring scale rulers for flatbed scanners. *Bulgarian Journal of Physics*, *34*(2), 356–368.
- Raji, A., Lemoine, J., Bossu, Y. L. J. C., & Boudghene, F. (2002). Improved radiological control of endovascular prosthesis positioning, using a region-contour cooperation approach. In *Proceedings of the 16th International Congress and Exhibition in Computer Assisted Radiology and Surgery (CARS 2002)* (pp. 843–848).

Supakul, N., Newbrough, K., Cohen, M. D., & Jennings, S. G. (2012). Diagnostic errors from digital stitching of scoliosis images – the importance of evaluating the source images prior to making a final diagnosis. *Journal of Pediatric Radiology*, 42, 584–598. doi:10.1007/s00247-011-2293-y

Ueda, K., Baba, T., Nakagawa, Y., & Amano, K. (2005). Detection of scale intervals in digital images. In *Proceedings of the 21st International Conference on Data Engineering Workshops* (pp. 1232–1237).

Zambanini, S., & Kampel, M. (2009). Robust automatic segmentation of ancient coins. In *VISAPP*, 1, 273–276.

Abhir Bhalerao is an Associate Professor in Computer Science. He received his B.Sc. and Ph.D. degrees in 1986 and 1992 respectively. He joined as faculty at Warwick in 1998 having completed 5 years as a post-doctoral research scientist with the NHS and Kings Medical School, London and as a Research Fellow at Harvard Medical School. His current interests are in modelling chronic lung diseases from CT, multi-camera reconstruction, and forensic image analysis. He has published about 70 refereed articles in image analysis, medical imaging, graphics and computer vision. He was the general co-chair and local organizer of the British Machine Vision Conference, 2007. He was the co-founder Warwick Warp Ltd., which makes fingerprint software. He is currently a director of Patternanalytics Research Ltd., a company specializing in biometric access control and forensic imaging systems.

Greg Reynolds obtained his M.Eng. and Ph.D. degrees in Electronic Engineering from the University of Birmingham in 2004 and 2007 respectively. His Ph.D. was involved with the DSP and automated interpretation of NMR signals from brain tumours and has published several papers in this area. His research interests are in machine learning for image processing problems, particularly using neural networks and fuzzy logic techniques. In 2010 he founded Pattern Analytics Research Ltd, where he is currently the technical director.