Estimating the Orientation and Recovery of Text Planes in a Single Image

P. Clark and M. Mirmehdi
Department of Computer Science,
University of Bristol,
Bristol, BS8 1UB, UK.
{pclark/majid}@cs.bris.ac.uk

Abstract

A method for the fronto-parallel recovery of paragraphs of text under full perspective transformation is presented. The horizontal vanishing point of the text plane is found using an extension of 2D projection profiles. This allows the accurate segmentation of the lines of text. Analysis of the lines will then reveal the style of justification of the paragraph, and provide an estimate of the vertical vanishing point of the plane. The text is finally recovered to a fronto-parallel view suitable for OCR or other higher-level recognition.

1 Introduction

Optical character recognition (OCR) is a long-standing area of computer vision which in general deals with the problem of recognising text in skew-compensated face-on images. There has been little research however into the recognition of text in real scenes in which the text is oriented relative to the camera. Such research has applications in replacing the document scanner with a point-and-click camera to facilitate non-contact text capture, assisting the disabled and/or visually impaired, wearable computing tasks requiring knowledge of local text, and general automated tasks requiring the ability to read where it is not possible to use a scanner. In preparation to apply OCR to text from images of real scenes, a fronto-parallel view of a segmented region of text must be produced. This is the issue considered in this paper.

Previous work in estimating the orientation of planar surfaces in still images varies in the assumptions made to achieve this. Ribeiro and Hancock[8] and Criminisi and Zisserman[3] have both presented methods which use texture distortion to estimate the vanishing points of the text plane. Affine distortion in power spectra are found along straight lines in [8], and correlation measures are used in [3] to determine first the orientation of the vanishing line and then its position. Although text has repetitive elements (characters and lines) these elements do not match each other exactly, and sometimes may cover only a small area of the image. Rother[9] attempts to find orthogonal lines in architectural environments, which are assessed relative to the camera geometry. Murino and Foresti [7] use a 3D Hough transform to estimate the orientation of planar shapes with known rectilinear features. Gool et al.[10] and Yip[11] both find the skewed symmetry of 2D shapes which have an axis of symmetry in the plane, allowing for affine recovery. We

BMVC 2001 doi:10.5244/C.15.44

(a) Original Image (b) Located text regions (c) Thresholded

Figure 1: Preparation of paragraph for planar recovery

require recovery from perspective transformation, but as with these latter works we will use a priori information about the 2D shape we are observing.

Knowledge of the principal vanishing points of the plane on which text lies is sufficient to recover a fronto-parallel view. We observe that in a paragraph which is oriented relative to the camera, the lines of text all point towards the horizontal vanishing point of the text plane in the image. Also, paragraphs often exhibit some form of justification, either with straight margins on the left or right, or if the text is centred, a central vertical line around which the text is aligned. In such cases these vertical lines point toward the vertical vanishing point of the text plane. We have therefore concentrated our work on the recovery of paragraphs with three lines of text or more, with the reasonable assumption that at least some justification exists (left, right, centred or full).

To avoid the problems associated with bottom-up grouping of elements into a paragraph model, in this work we ensure the use of all of the global information about the paragraph at one time. The principle of 2D projection profiles are extended to the problem of locating the horizontal vanishing point by maximising the separation of the lines in the paragraph. The formation of the segmented lines of text will then reveal the style of justification or alignment of the paragraph, and provide an estimate of the vertical vanishing point.

The rest of the paper is structured as follows. In Section 2 we briefly review our previous work which provides the input to the work described here. Sections 3 and 4 discuss the paragraph model fitting stage, location of the horizontal vanishing point, separation of the lines of text, and estimation of the vertical vanishing point. In Section 5 the vanishing points of the text plane are employed to recover a fronto-parallel view of the paragraph suitable for higher level recognition. We conclude and consider future work in Section 6.

2 Finding Text Regions

In [2] we introduced a text segmentation algorithm which used localised texture measures to train a neural network to classify areas of an image as text or non-text. Figure 1(b) shows a large region of text which was found in Figure 1(a) using this approach. In this work we consider the output of the system presented in [2] and analyse each region individually to recognise the shape of the paragraph, recover the 3D orientation of the text plane, and generate a fronto-parallel view of the text.

In order to analyse the paragraph shape, we first require a classification of the text and background pixels. Since the region provided by the text segmentation algorithm will principally contain text, the background and foreground colours are easily separable through thresholding. We choose the average intensity of the image neighbourhood as an adaptive threshold for each pixel, in order to compensate for any variation in illumination across a text region. The use of partial sums[4] allow us to calculate these thresholds efficiently. To ensure the correct labelling of both dark-on-light and light-on-dark text, the proportion of pixels which fall above and below the thresholds is considered. Since in a block of text there is always a larger area of background than of text elements, the group of pixels with the lower proportion is labelled as text, and the other group as background. The example shown in Figure 1(c) demonstrates the correct labelling of some light text on a dark background and is typical of the input into the work presented here.

3 Locating the Horizontal Vanishing Point

In [5], Messelodi and Modena demonstrate a text location method on a database of images of book covers. They employ projection profiles to estimate the skew angle of the located text. A number of potential angles are found from pairs of components in the text, and a projection profile is generated for each angle. They observe that the projection profile with the minimum entropy corresponds to the correct skew angle. This guided 1D search is not directly applicable to our problem, which is to find a vanishing point in $\mathbb{R}^2$, with two degrees of freedom. In order to search this space, we will generate projection profiles from the point of view of vanishing points, rather than from skew angles.

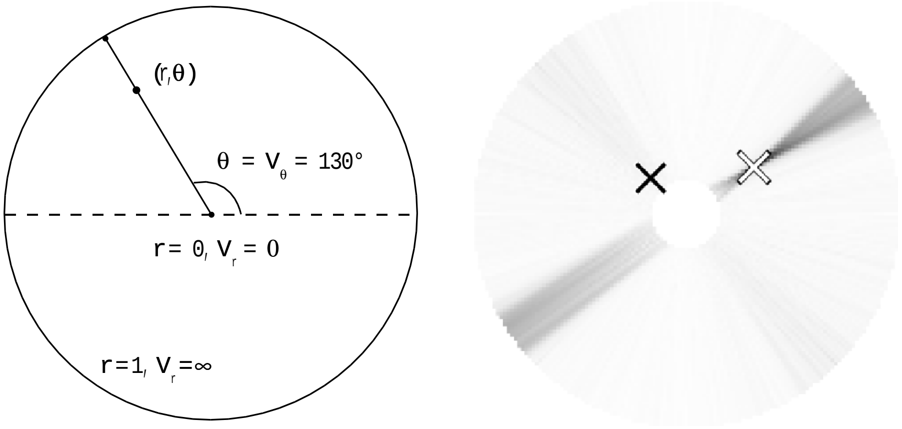
We use a circular search space C as illustrated in Figure 2(a). Each cell $c = (r, \theta)$, $r \in [0, 1)$ and $\theta \in [0, 2\pi)$, in the space C corresponds to a hypothesised vanishing point $\mathbf{V} = (V_r, V_\theta)$ on the image plane $\mathbb{R}^2$, with scalar distance $V_r = r/(1-r)$ from the centre of the image, and angle $V_\theta = \theta$. This maps the infinite plane $\mathbb{R}^2$ exponentially into the finite search space C. In our experiments, to ensure the accurate location of the vanishing point, the search space was populated with $10,000$ evenly positioned cells. A projection profile of the text is generated for every vanishing point in C, except those lying within the text region itself (the central hole in Figure 2(b)).

A projection profile B is a set of bins $\{B_i, i = 0, .., N\}$ into which image pixels are accumulated. In the classical 2D case, to generate the projection profile of a binary image from a particular angle ϕ, each positive pixel $\mathbf{p}$ is assigned to bin B_i, where i is dependent on $\mathbf{p}$ and ϕ according to the following equation:

$$i(\mathbf{p}, \phi) = \frac{\mathbf{p} \cdot \mathbf{U}}{s} \times N + \frac{N}{2} \tag{1}$$

where $\mathbf{U} = (\sin\phi, \cos\phi)$ is a normal vector describing the angle of the projection profile, and $s > N$ is the diagonal distance of the image. In this equation, the dot product $\mathbf{p} \cdot \mathbf{U}$ is the position of the pixel along the axis of the projection profile in the image defined by ϕ. Manipulation with s and N is then employed to map from this axis into the range of the bins of the projection profile.

In our case, instead of an angle ϕ, we have a point of projection $\mathbf{V}$ on the image plane, which has two degrees of freedom. Our bins, rather than representing parallel slices of the image along a particular direction, must represent angular slices projecting from $\mathbf{V}$. Hence, we refine (1) to map from an image pixel $\mathbf{p}$ into a bin B_i as follows:



(a) Relationship between search space and $\mathbb{R}^2$

(b) Scores for all projection profiles generated from Figure 1(c)

Figure 2: Search space

$$i(\mathbf{p}, \mathbf{V}) = \frac{\text{ang}(\mathbf{V}, \mathbf{V} - \mathbf{p})}{\Delta\theta} \times N + \frac{N}{2} \qquad (2)$$

where ang$(\mathbf{V}, \mathbf{V} - \mathbf{p})$ is the angle between pixel $\mathbf{p}$ and the centre of the image, relative to the vanishing point $\mathbf{V}$, and $\Delta\theta$ is the size of the angular range within which the text is contained, again relative to the vanishing point $\mathbf{V}$. $\Delta\theta$ is obtained from $\Delta\theta = \text{ang}(\mathbf{V} + \mathbf{t}, \mathbf{V} - \mathbf{t})$ where $\mathbf{t}$ is a vector perpendicular to $\mathbf{V}$ with magnitude equal to the radius of the bounding circle of the text region (shown in Figure 3). Unlike s in (1), it can be seen that $\Delta\theta$ is dependent on the point of projection $\mathbf{V}$. In fact $\Delta\theta \to 0$ as $\mathbf{V}_r \to \infty$ since more distant vanishing points view the text region through a smaller angular range. The use of $\mathbf{t}$ to find $\Delta\theta$ ensures that the angular range over which the text region is being analysed is as closely focused on the text as possible, without allowing any of the text pixels to fall outside the range of the projection profile's bins. This is vital in order for the generated profiles to be comparable, and also beneficial computationally since no bins need to be generated for the angular range $2\pi - \Delta\theta$ which is absent of text.

Having accumulated projection profiles for all the hypothesised vanishing points using (2), a simple measure of confidence is found for each projection profile B. The confidence measure was chosen to respond favourably to projection profiles with distinct peaks and troughs. Since straight lines are most clearly distinguishable from the point where they intersect, this horizontal vanishing point and its neighbourhood will be favoured by the measure. We found the squared-sum $\sum_{i=1}^{N} {B_i}^2$ to respond better than entropy or derivate-squared-sum measures, as well as being efficient to compute. The confidence of each of the vanishing points with regard to the binarised text in Figure 1(c) is plotted in Figure 2(b), where darker pixels represent a larger squared-sum, and a more likely vanishing point. The projection profile with the largest confidence is chosen as the horizontal vanishing point of the text plane. This winning projection profile and an example of a poor projection profile are shown in Figure 3, and marked in Figure 2(b) with a white cross and a black cross respectively. Despite general image noise and the resolution of the search

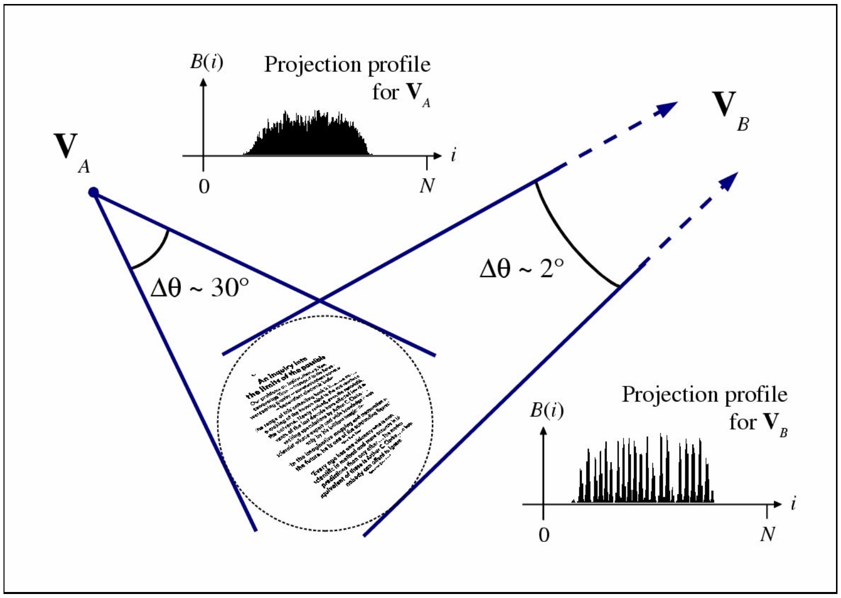


Figure 3: Two potential vanishing points $\mathbf{V}_A$ and $\mathbf{V}_B$, and their projection profiles.

space, this method has consistently provided a good estimate of the horizontal vanishing point in our experiments.

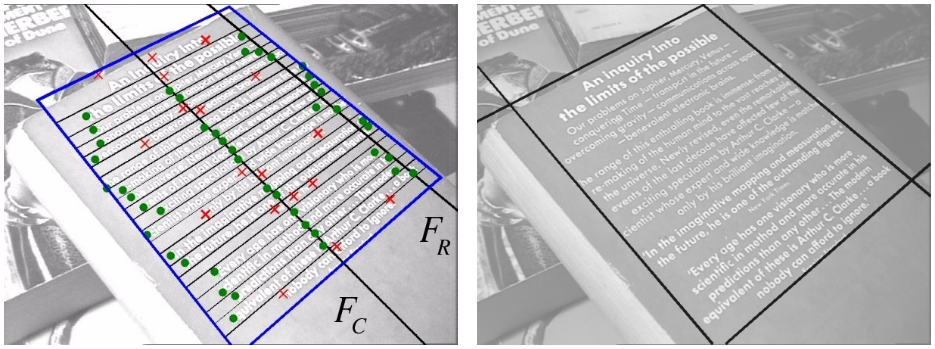
4 Locating the Vertical Vanishing Point

The location of the horizontal vanishing point, and the projection profile of the text from that position, now make it possible to separate the individual lines of text. This will allow the style of justification of the paragraph to be determined, and lead to the location of the vertical vanishing point.

We apply a simple algorithm to the winning projection profile to segment the lines. A *peak* is defined to be any range of angles over which all the projection profile's bins register more than K pixels, taken as the average height of the interesting part of the projection profile:

$$K = \frac{1}{y - x + 1} \sum_{i=x}^{y} B_i \tag{3}$$

where x and y are the indices of the first and last non-empty bins respectively. A *trough* is defined to be the range of angles between one peak and the next. The central angle of each trough is used to indicate the separating boundary of two adjacent lines in the paragraph. We project segmenting lines from the vanishing point through each angle in the range. All pixels in the binary image lying between two adjacent segmenting lines are collected together as one line of text. The result of this segmentation is shown in Figure 4(a). Both full and short lines of text are segmented accurately.



(a) Segmented paragraph

(b) Located paragraph frame

Figure 4: Paragraph recognition. (a) Line segmentation marked in black; points for line fitting in green (used) and red (rejected outliers); the rectangular frame on the text plane in blue. (b) The frame used for recovery.

We determine the left end, the centroid, and the right end of each of the segmented lines, to form three sets of points P_L, P_C, P_R respectively. Since we anticipate some justification in the paragraph, we will expect a straight line to fit well through at least one of these sets of points, representing the left or right margin, or the centre line of the paragraph. This will be the *baseline*, a line in the image upon which the vertical vanishing point must lie. To establish the line of best fit for each set of points, we use a RANSAC (random sampling concensus, [1]) algorithm to reject outliers caused for example by short lines, equations or headings. Given a set of points P, the line of best fit through a potential fit $F = \{\mathbf{p}_i, i = 1, .., L\} \subseteq P$ passes through $\mathbf{c}$, the average of the points, at an angle ψ found by minimising the following error function:

$$E_F(\psi) = \frac{1}{L^5} \sum_{i=1}^{L} ((\mathbf{p}_i - \mathbf{c}) \cdot \mathbf{n})^2 \tag{4}$$

where $\mathbf{n} = (-\sin\psi, \cos\psi)$ is the normal to the line, L^2 normalises the sum, and a further L^3 rewards the fit for using a large number of points. Hence for the three sets of points P_L, P_C, P_R we obtain three lines of best fit F_L, F_C, F_R with their respective errors E_L, E_C, E_R.

Condition	**Type of paragraph**
$E_L \simeq E_C \simeq E_R$	Fully justified.
$\min(E_L, E_C, E_R) = E_L$	Left justified.
$\min(E_L, E_C, E_R) = E_R$	Right justified.
$\min(E_L, E_C, E_R) = E_C$	Centrally justified.

Table 1: Classifying the type of paragraph

It is now possible to classify the style of justification of the paragraph using the rules in Table 1. Figure 4(a) shows the baseline F_C passing through the centre of the paragraph.

In this case $E_C < E_R$ and $E_C < E_L$, hence the last condition in Table 1 is satisfied and the paragraph is correctly identified as being centrally justified. The baseline represents a vertical line on the text plane, and is alone sufficient for weak perspective. However, for planes of text under full perspective, we need to find the distance along the baseline at which the vertical vanishing point lies. We proceed with a generic method, regardless of the style of paragraph: of the three lines of best fit, we take the two with the least error, and intersect them to estimate the position of the vertical vanishing point. In the example in Figure 4(a) the lines chosen were F_C and F_R. This method assumes that the two fitted lines accurately represent vertical margins of the paragraph. However for certain types of paragraph which are not fully justified, this assumption can break down (see Section 6).

Next, having found the vanishing points of the plane, we may project two lines from each to describe the left and right margins and the top and bottom limits of the paragraph. These lines are intersected to form a quadrilateral enclosing the text, as shown in Figure 4(b). This quadrilateral is then used to recover a fronto-parallel viewpoint of the paragraph of text as described in the next section.

5 Removing Perspective

Before mapping the image quadrilateral into a rectangle, a 3D model of the text block is desirable since it will provide us with the aspect ratio of the rectangle in the scene, and provide a good model of the origin of the image pixels on the text plane. We use the quadrilateral frame from the paragraph model to recover the 3D orientation of the text plane, and then fix the distance to obtain a scale-independent model. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be the vertices of the quadrilateral in the image plane, labelled clockwise from top-left. We wish to find $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, the world coordinates of the corners of the rectangle:

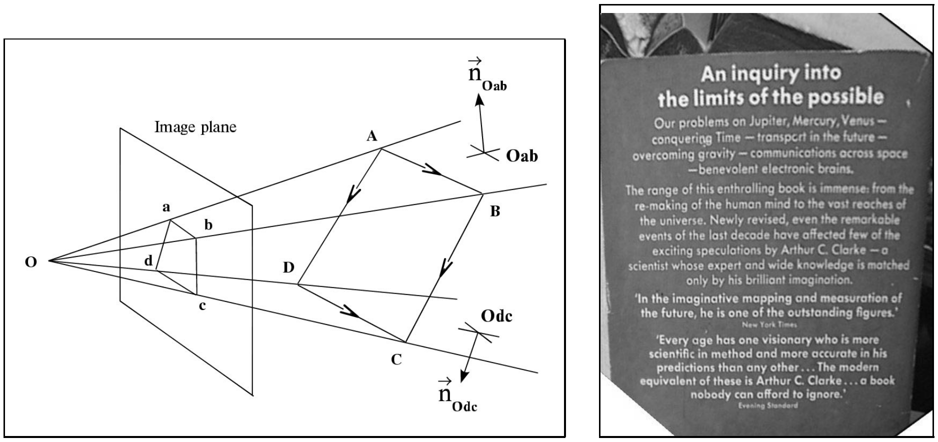
$$(\mathbf{A}\ \mathbf{B}\ \mathbf{C}\ \mathbf{D})^{\mathbf{t}} = \mathbf{O} + (\alpha\ \beta\ \gamma\ \delta)^{\mathbf{t}}(\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{d}) \tag{5}$$

where $\mathbf{O}$ is the centre of projection, and $\alpha, \beta, \gamma, \delta$ are the depths of the four points into the scene. In Figure 5(a), it can be seen that the projection from the origin $\mathbf{O}$ through the image line $\mathbf{ab}$ forms a plane $\mathbf{Oab}$ in the scene, upon which the top edge $\mathbf{AB}$ of the rectangle must lie. Similarly the projection through the bottom line of the quadrilateral $\mathbf{dc}$ forms a plane $\mathbf{Odc}$, upon which the bottom line $\mathbf{DC}$ of the rectangle must lie. Since the lines $\mathbf{AB}$ and $\mathbf{DC}$ are opposite edges of the rectangle in the scene, they must be parallel in the direction of the horizontal vector of the text plane $\vec{\mathbf{h}}$. Since this vector lies on both planes $\mathbf{Oab}$ and $\mathbf{Odc}$, it must be perpendicular to the normals of the two planes. Hence,

$$\vec{\mathbf{h}} \equiv \vec{\mathbf{AB}} \equiv \vec{\mathbf{DC}} = \vec{\mathbf{n}}_{\mathbf{Oab}} \times \vec{\mathbf{n}}_{\mathbf{Odc}} \tag{6}$$

$$\text{and } \vec{\mathbf{v}} \equiv \vec{\mathbf{AD}} \equiv \vec{\mathbf{BC}} = \vec{\mathbf{n}}_{\mathbf{Oad}} \times \vec{\mathbf{n}}_{\mathbf{Obc}} \tag{7}$$

where (7) applies the same principle to the left and right planes of the quadrilateral to recover the vertical direction $\vec{\mathbf{v}}$ of the rectangle in the scene. The two vectors $\vec{\mathbf{h}}$ and $\vec{\mathbf{v}}$ now describe the orientation of the text plane, but the depth of the plane into the scene is unknown. But since we are not interested in scale, we may fix $\alpha = 1, \mathbf{A} = \mathbf{a}$ and resolving the other three corners $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ relatively, obtain 3D coordinates for the text rectangle we wish to recover. We now use the aspect ratio of the rectangle in world space



(a) The geometry involved in planar recovery

An inquiry into
the limits of the possible

Our problems on Jupiter, Mercury, Venus —
conquering Time — transpcrt in the future —
overcoming gravity — communications across space
—benevolent electronic brains.

The range of this enthralling book is immense: from the re-making of the human mind to the vast reaches of the universe. Newly revised, even the remarkable events of the last decade have affected few of the exciting speculations by Arthur C. Clarke — a scientist whose expert and wide knowledge is matched only by his brilliant imagination.

'In the imaginative mapping and measuration of the future, he is one of the outstanding figures.'
New York Times

'Every age has one visionary who is more scientific in method and more accurate in his predictions than any other . . . The modern equivalent of these is Arthur C. Clarke . . . a book nobody can afford to ignore.'
Evening Standard

(b) Fronto-parallel recovery of example text

Figure 5: Recovery of text from image quadrilateral

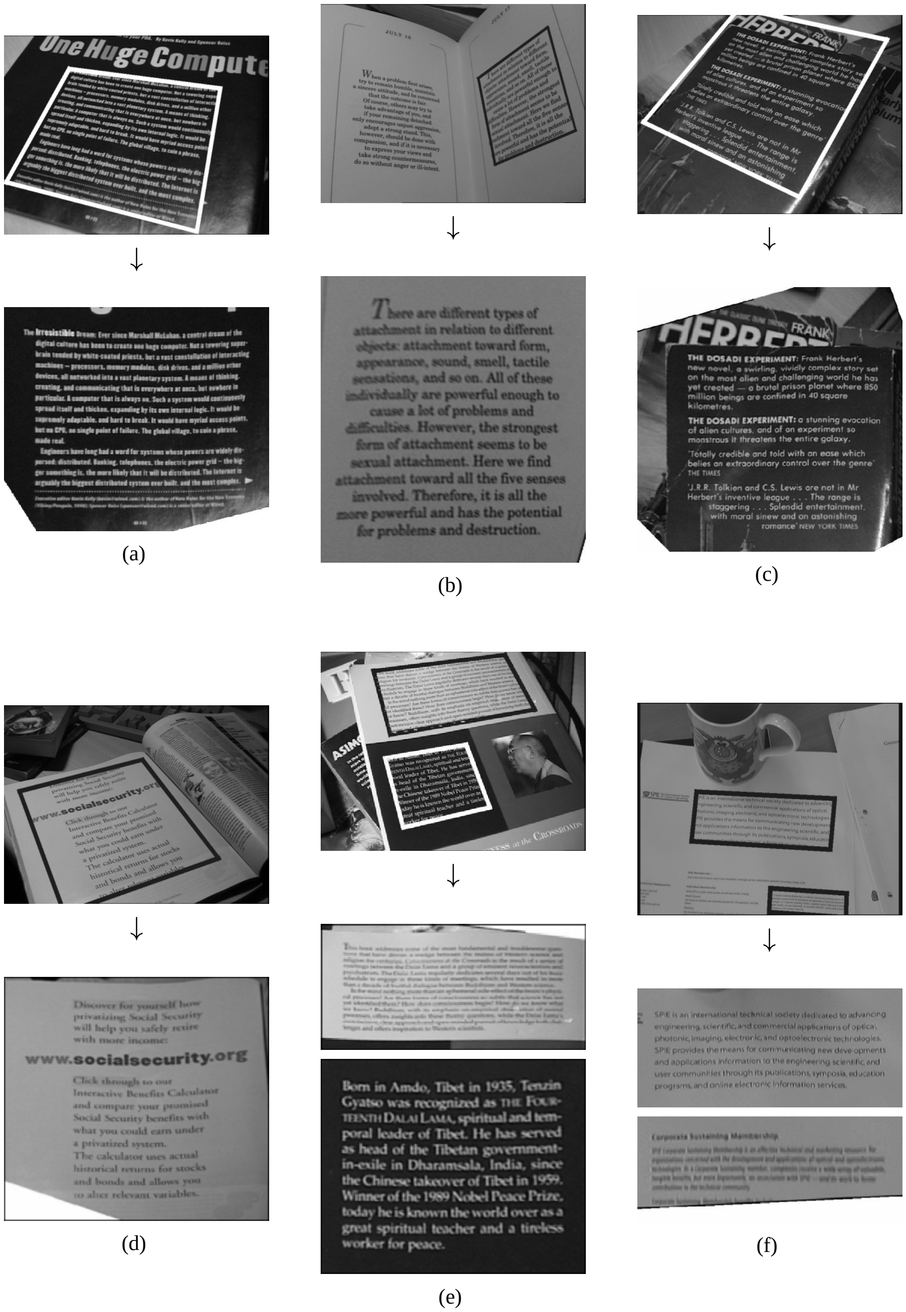
to construct a destination image, and with suitable interpolation generate a fronto-parallel view of the text.

The recovered page for the running example may be seen in Figure 5(b). Some further examples in Figure 6 show more cases of the recovery of paragraphs, with left-justified and centrally aligned text. Figures 6(a) and 6(b) present examples of paragraphs with light-coloured text on dark-coloured background and vice versa. Two examples of recovery with poorly aligned paragraphs are demonstrated in Figures 6(c) and 6(d). Despite the poor alignment, the proposed method has estimated the location of the horizontal and particularly the vertical vanishing points well and recovered the paragraphs correctly. Figures 6(e) and 6(f) present the recovery of multiple paragraphs in each image. Furthermore, in Figure 6(f) we note that good results are obtained even though there are only a few lines in each of the recovered paragraphs.

6 Discussion

We have presented a method for the fronto-parallel recovery of a paragraph under perspective transformation in a single image. Projection profiles from hypothesised vanishing points are used to robustly recover the horizontal vanishing point of the text plane, and segment the paragraph into its constituent lines. Line fitting on the margins and central line of the paragraph is then applied to estimate the vertical vanishing point. Using these principal vanishing points we find the orientation of the text plane and recover a fronto-parallel view. The algorithm performs well for a wide range of paragraphs, provided each paragraph has at least three or more full lines.

While generating $10,000$ projection profiles for potential vanishing points in Section 3 requires a large amount of processing, we have done this initially to reveal the nature of the search space. The results in Figure 2(b) demonstrate that the space has obvious large-scale features, which could direct a more efficient search. For example, an initial low



The Irresistible Dream: Ever since Marshall McLuhan, a central dream of the digital culture has been to create one huge computer. Not a towering superbrain tended by white-coated priests, but a vast constellation of interacting machines — processers, memory modules, disk drives, and a million other devices, all networked into a vast planetary system. A means of thinking, creating, and communicating that is everywhere at once, but nowhere in particular. A computer that is always on. Such a system would continuously spread itself and thicken, expanding by its own internal logic. It would be supremely adaptable, and hard to break. It would have myriad access points, but no CPU, no single point of failure. The global village, to coin a phrase, made real.

Engineers have long had a word for systems whose powers are widely dispersed: distributed. Banking, telephones, the electric power grid — the bigger something is, the more likely that it will be distributed. The Internet is arguably the biggest distributed system ever built, and the most complex. ▶

There are different types of attachment in relation to different objects: attachment toward form, appearance, sound, smell, tactile sensations, and so on. All of these individually are powerful enough to cause a lot of problems and difficulties. However, the strongest form of attachment seems to be sexual attachment. Here we find attachment toward all the five senses involved. Therefore, it is all the more powerful and has the potential for problems and destruction.

THE DOSADI EXPERIMENT: Frank Herbert's new novel, a swirling, vividly complex story set on the most alien and challenging world he has yet created — a brutal prison planet where 850 million beings are confined in 40 square kilometres.

THE DOSADI EXPERIMENT: a stunning evocation of alien cultures, and of an experiment so monstrous it threatens the entire galaxy.

'Totally credible and told with an ease which belies an extraordinary control over the genre' THE TIMES

'J.R.R. Tolkien and C.S. Lewis are not in Mr Herbert's inventive league . . . The range is staggering . . . Splendid entertainment, with moral sinew and an astonishing romance' NEW YORK TIMES

Discover for yourself how privatizing Social Security will help you safely retire with more income:

www.socialsecurity.org

Click through to our Interactive Benefits Calculator and compare your promised Social Security benefits with what you could earn under a privatized system. The calculator uses actual historical returns for stocks and bonds and allows you to alter relevant variables.

This book addresses some of the most fundamental and troublesome questions that have driven a wedge between the realms of Western science and religion for centuries. *Consciousness at the Crossroads* is the result of a series of meetings between the Dalai Lama and a group of eminent neuroscientists and psychiatrists. The Dalai Lama regularly dedicates several days out of his busy schedule to engage in these kinds of meetings, which have resulted in more than a decade of fruitful dialogue between Buddhism and Western science.

Is the mind nothing more than an ephemeral side-effect of the brain's physical processes? Are there forms of consciousness so subtle that science has not yet identified them? How does consciousness begin? How do we know what we know? Buddhism, with its emphasis on empirical observation of mental processes, offers insights into these thorny questions, while the Dalai Lama's own incisive, clear approach and open-minded pursuit of knowledge both challenges and offers inspiration to Western scientists.

Born in Amdo, Tibet in 1935, Tenzin Gyatso was recognized as THE FOURTEENTH DALAI LAMA, spiritual and temporal leader of Tibet. He has served as head of the Tibetan government-in-exile in Dharamsala, India, since the Chinese takeover of Tibet in 1959. Winner of the 1989 Nobel Peace Prize, today he is known the world over as a great spiritual teacher and a tireless worker for peace.

SPIE is an international technical society dedicated to advancing engineering, scientific, and commercial applications of optical, photonic, imaging, electronic, and optoelectronic technologies. SPIE provides the means for communicating new developments and applications information to the engineering, scientific, and user communities through its publications, symposia, education programs, and online electronic information services.

Corporate Sustaining Membership

SPIE Corporate Sustaining Membership is an effective technical and marketing resource for organizations concerned with the development and applications of optical and optoelectronic technologies. As a Corporate Sustaining member, companies receive a wide array of valuable, tangible benefits, but more importantly, an association with SPIE — and its work to foster contributions to the technical community.

Figure 6: Further examples of fronto-parallel recovery of paragraphs. In each case (a) to (f), the orignal image is shown above one or more recovered paragraphs.

resolution scan of C will reveal those angular regions which are likely to contain the correct vanishing point. These regions can then be searched thoroughly to find the precise angle and distance of the horizontal vanishing point.

In Section 4, we use two vertical lines of best fit from the paragraph to estimate the vertical vanishing point. For paragraphs which are not fully-justified, the accuracy of the fitted lines reduces when the number of lines in the paragraph is small, or the number of words per line is low (which will result in a poorly defined margin). We are exploring alternative indicators to the position of the vanishing point, for example line spacing.

Although the resulting images reproduced here are at low resolution, most of them are nevertheless suitable to be fed to an OCR system to interpret the text or to be read by a human observer. In another work [6] we have performed OCR on (already fronto-parallel) text in the scene, segmented using an active camera, with good results. In the near future we intend to integrate the work described here and in [6] towards an automatic system for text recognition in the environment.

References

[1] R. Bolles and M. Fischler. A RANSAC-based approach to model fitting and its application to finding cylinders in range data. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pages 637–643, 1981.

[2] P Clark and M Mirmehdi. Finding text regions using localised measures. In *Proc. 11th British Machine Vision Conference*, pages 675–684, 2000.

[3] A. Criminisi and A. Zisserman. Shape from texture: homogeneity revisited. In *Proc. 11th British Machine Vision Conference*, pages 82–91, 2000.

[4] S. Hodges and R. J. Richards. Faster spatial image processing using partial summation. Technical Report CUED/F-INFENG/TR.245, Cambridge University, 1996.

[5] S. Messelodi and C.M. Modena. Automatic identification and skew estimation of text lines in real scene images. *Pattern Recognition*, 32:791–810, November 1999.

[6] M. Mirmehdi, P. Clark, and J. Lam. Extracting low resolution text with an active camera for OCR. Accepted for SNRFAI'2001, 2001.

[7] V. Murino and G. Foresti. 2D into 3D Hough-space mapping for planar object pose estimation. *Image and Vision Computing*, 15:435–444, 1997.

[8] E. Ribeiro and E. Hancock. Detecting multiple texture planes using local spectral distortion. In *Proc. 11th British Machine Vision Conference*, pages 102–111, 2000.

[9] C Rother. A new approach for vanishing point detection in architectural environments. In *Proc. 11th British Machine Vision Conference*, pages 382–391, 2000.

[10] L Van Gool, T Moons, D Ungureanu, and A Oosterlinck. Characterization and detection of skewed symmetry. *CVIU*, 61(1):138–150, 1995.

[11] R. Yip. A Hough transform technique for the detection of reflectional symmetry and skew-symmetry. *Pattern Recognition Letters*, 21:117–130, 2000.