Anhang C

Auswertung des Entzerrungsalgorithmus

Die Auswertung erfolgte anhand eines Vergleichs der Erkennungsraten. Einmal wurde im Vorfeld der Texterkennung eine Entzerrung mit Hilfe des vorgestellten Algorithmus durchgeführt und einmal wurde die Entzerrung vom systemeigenen Entzerrungsmodul von OmniPage 16 übernommen. Das Ergebnis der Auswertung wird wie folgt visualisiert:

1. Entzerrte und binarisierte Version eines Dokuments
2. Das Ergebnis der Zeichenerkennung nach der Entzerrung mit dem vorgestellten Algorithmus
3. Das Ergebnis der Zeichenerkennung ohne Beteiligung vom Algorithmus

	Anzahl der Wörter	**Fehlerrate mit Algorithmus**	**Fehlerrate ohne Algorithmus**
Beispiel A	657	92,5%	91,7%
Beispiel B	640	95,4%	93,8%
Beispiel C	807	93%	91,6%
Beispiel D	91	87,9%	76,9%
Beispiel E	847	94,2%	91%
Beispiel F	799	91,6%	87,3%
Beispiel G	614	97,7%	96%
Beispiel H	489	94,5%	91,4%

Beispiel I	155	85,2%	87,7%
Beispiel J	708	94,8%	95%
Beispiel K	462	88,7%	86,6%
Beispiel L	277	85,2%	85,6%

Extraction of illusory linear clues in perspectively skewed documents

Maurizio Pilu
Hewlett-Packard Laboratori
Bristol, BS34 8QZ, UK
maurizio_pilu@hp.com

Abstract

This work deals with the recovery of illusory linear clues from perspectively skewed documents with the purpose of using them for rectification. The computational approach proposed implements the perceptual organization principles implicitly used in textual layouts. The numerous examples provided show that the method is robust and viewpoint and scale invariant.

1. Introduction

As sensor resolution increases and prices drop, cameras will one day be used to capture documents in lieu of flatbed scanners. One of the main disadvantages when capturing a document with a camera is that the non-contact image capture process causes geometric distortions dependent upon the camera orientation, in particular perspective skew, as typified in Figure 1-left.

This work addresses a fundamental problem that need to be faced when trying to passively deskew a captured document, namely the detection of linear clues that can be used as geometric primitives to determine the document plane orientation w.r.t. the camera and the rectifying homography.

Figure 1-left shows several kinds of linear clues that may arise in practice. Clue A is a vertical *illusory* [2] clue. It does not correspond to an actual linear feature but rather to a set of organized features arranged linearly. Clue B is a vertical *hard line* which is the projection of the actual document edge. Clues C are horizontal illusory lines, which have been inferred from the arrangement of characters into text lines. Similarly to clue B, clue D is a horizontal hard line. Clue E is a quadrilateral that can be either illusory or corresponding to an actual rectangular outline in the document (e.g. a figure box or the four document boundaries). Hard edges of the type B and D can be detected rather trivially in many ways using edge detectors (e.g. the Hough transform) and will not be addressed as such in this paper. Illusory edges such as C and A are rather difficult to find reliably in practice and most of the literature on document analysis has been only focusing on recovering groups of parallel text lines to compensate for rotation for OCR and other scanning applications.

In this paper we will be dealing with the problem of finding illusory clues of the type A and C without assumptions on the type of document, fonts or camera rotation and orientation other than the presence of some organized text.

Text had been designed long before perception studies but with the unspoken principles of perceptual saliency (see e.g. [2]) in mind. Give any organised text, even foreign and unintelligible to us, and words, lines and paragraph structures pop out preattentively. The approach presented in this paper is strongly based on a computational implementation of these perceptual organization principles and will be shown to be robust, fast and general

2. Related works

The subject of correcting perspective skew in documents has been largely neglected in the literature. Even recent works explicitly addressing camera-based document imaging such as [14] treat only rotation-induced skew.

The geometry of the rectification is well known [6] but the problem of passively and robustly detecting the geometric features needed is still open.

A substantial body of research has been dedicated to text and page segmentation in document images, but any distortion considered is again only rotation-induced. The main bottom-up methods used include many variations on projection profiles approaches, Hough-inspired techniques [12] and nearest-neighbour clustering [9][13]. Top-down methods seek to extract the high-level structure of the images, such as by using Manhattan layout analysis of the (white) background [1]. Methods that analyse connected components as we do are fairly common in document analysis. Besides the aforementioned [9][13] and others, an interesting approach that employs perceptual organization principles is [11], although assuming of parallel-lines. All the methods above assume that the text lines are still parallel in the image and could not work with perspective skew. The work

Beispiel A

/Beispiel A

Extraction ofillusory linear clues
in oerspectivelv skewed documents
Maurizio Pilu
Hewlett-Packard Laboraton
Bristol, BS34 SQL UK
maurizio_piluii bp.com
Abstract
This work deals with the rrcovery of illusory linear clues
from perspectively skewed documente with the purpose of
usinq them for mclification. The computational approach
proposed implements the perceptual orqanization prne.ipies
implicitlv used in tertual lavouts. The numerous clamoies
provided show that the method is robust and viempoint
and scale invariant.
1. Introduction
As sensor resolution increases and prices drop, cameras will
one dav be used to capture documents in lieu of flatbed
scanners. One of thc main disadvantaqes when captunnq a
document with a camera is that the non-contact image capture
process causes qeometric distortions dependent upon
the camera oncntation, in particular perspectiw skew, as
tvpified in Fiqure 1-left.
This work addresses a fundamental problem that need to
be faced when tryinq to passively dcskew a captured document,
namely the detection of linear clues that can be used
as qeometrtc primitives to determine the document plane
orientation w.r.t the camera and the rectifvinq homoara-
phy.
Fiqure I -left shows several kinds of linear clues that may
arise in practice. Clue A is a vertical illusory [2] clue. It
does not correspond to an actual linear feature but rather to
a set of oraanized features arranqed linearly. Clue B is a
vertical hard line which is the projection of the actual document
edqe. Clues C are horizontal illusory lines, which have
been inferred from the arranqement of characters into text
lines. Similarly to clue B, clue D is a horizontal hard line.
Clue E is a quadrilateral that can be either illusorv or curresoondinq
to an actual rectanqular outline in the document
(e.q. a fiqure box or the four document boundaries). Hard
edqes of the type 13 and D can be detected rather mvially in
many ways usinq edqe detectors (e.q. the Hough transform)
and will not be addressed as such in this paper. Illusory
ed. qes such as C and A arc rather difficult to find reliably
in practice and most of the literature on document analvsi's
has been only focusinq on recovennq qroups of parallel text
lines to compensate for mtation for OCR and other scanninq
applications.
In this paper we will be dealinq with the problem of findinq
illusory clues of the type A and C suthut assuinption.s
on the type of document, fonts or cam;...ra rotation and orientation
other than the presence of somc oraanized test.
Text had been desioned lonq before perception studies
but with the unspoken principles of perceptual saliency (sec
e.q. [2]) in mind. Give any oraanised text, even foreiqn
and unintelliqible to us, and words, lines and paraqraph
structures pop out preanentively. The approach presented
in this paaer is stronqlv based on a computational implementation
of these perceptual oraanization pnnciplcs and
will be shown to be robust, fast and qeneral
2. Reiated works
The subject of correctinq perspective skew in documents
has been largelv neqlected in the literature. Even recent
va.orks explicitly addressinq camera-based document imaqinq
such as 114] treat only rotation-induced skew,
The qeometry of thc rectification is well known (61 but
the problem ot passively and robustly detectinq the qeometne
features needed is still oaen.
A substantial body of research has been dedicated to text
and paqe seqmentation in document images, but any distortion
considered is aqain only rotation-induced. The main
bottom-up metluxis used include many variations on projection
orotiles approaches. Hough-inspired techniques [12]
and nearest-neighbour clustering [91[131. Top-down methods
seek to extract the high-level structure of the images,
such as by usinq Manhattan layout analysis of the (while)
backqround [11. Methods that analyse connected components
as we do are fairlv common in document analvsis. Besides
the aforementioned [9][13] and others, an interesting
approach that employs perceptual oraanization principles is
[111, although assuminq of parallel-lines. All the methods
above assume that the text lines are still parallel in the imaqe
and could not work with perspective skew. The work

/Beispiel A

Extraction of illusory linear clues
in perspectively skewed documents
Maurizio Pilu
HeWlert-Packard Laboratories
Bristol. BS34 SQL UK
maurizio piluiiho.com
Abstract
This work deals with the recovery of illusory linear clues
fmm perspectively skewed documents with the purpose of
using them .fOr rectification. The computational approach
oroposed imolements the oerceptual oroanization principles
implicitly used in textual layouts. The numerous eta',
pies provided show that the method is robust and viewpoint
and scale invariant.
1. Introduction
As sensor resolution increases and prices drop, cameras will
one day be used to capture documents in lieu of flatbed
scanners. One of the main disadvantages when capturing a
document with a camera is that the non-contact image capture
process causes geometric distortions dependent upon
the camera onentation, in particular oersoective skew, as
typified in Figure 1-left.
This work addresses a fundamental problem that need to
bc faced when trying to passively deskew a captured document,
namely the detection of linear clues that can be used
as geometnc primitives to deterniine the document plane
orientation w.r.t. the carnera and the rectifying hornography.
Figure I-left shows several kinds oflinear clues that may
arise in practice. Clue A is a vertical Wilson, [2] clue It
does not correspond to an actual linear feature but rather to
a set of oroanized features arranoed linearly. Clue B is a
vertical hard line which is the projection of the actual document
edge.
Clues C are horizontal illusory lines. which have
been inferred fron the arranoement of characters into text
lines. Similarly to clue B. clue D is a horizontal hard line.
Clue E is a quadrilateral that can he either illusory- or corresponding
to an actual rectanqular outline in the document
(e q. a figure box or the four document boundaries). Hard
the type B and D can he detected rather tnviallv in
erndaoi: 'Iwa.v. sino edoe detectors (e.o. the Hough transform)
and wilt not be addressed as such in this paper. Illusory
cdoes such as C and A. are rather difficult to find reliably
in practice and most of the literature on document analysis
has been only focusino on recoverino aroups of parallel text
lines to compensate for rotation for OCR arid other seanninc
application..
In this paper we will Se dealino w ith the problem of find
ino illusory clues of the type A and C without assumptions
on the type of document, fonts or eamera rOtatiOn and orientation
other than the oresence of some oroanized text .
Text had been desioned lonq before perception studies
but with the uns-poken pnnciples of perceptual saliency (:ce
e.q. [21 in mind. Give any orq,anised text, even foreion
and unintelliqible to us, and words, lines and paral._zraph
structure's pop out oreattentwelv The approach oresented
in this paper is stronolv based on a computational imple
mentation of these perceptual oroanization pnnciples and
will he shown to he robust, fast and oeneral.
2. Related works
The subiect of correctino oersoective skew in documents
has been largely neqlected in the literature. Even recent
works explicitly addressino camera-based document intaq
Mq such as 1141 treat only rotation-induced skew.
The oeometry of the rectification is well known ol hui
the oroblem ot passivelv and robustlv detectino the oeometric
features needed is still open.
A substantial body of research has been dedicated to text
and page seomentation in document images, but any distor
tion considered is aqain)nlv rotation-induced. The main
bottom-up methods used include manv variations on orojection
profiles approaches, Hough-inspired techniques [I 21
and nearest-neighbour clustering [91[131. Top-down methods
seek to extract the hiah-level structure of the images,
such as by usino Manhattan layout analysis of the (white)
backoround [1]. Methods that analvse connected components
as we do are fairly common in document analysis. Besides
the aforementioned [91113] and others, an interestino
approach that employs perceptual oroanization principles is
[111, although assumino of parallel-lines. All the methods
above assume that the text lines are still oarallel in the imaoe
and could not work with perspective skew. The work

Flattening Curved Documents in Images

Jian Liang, Daniel DeMenthon, David Doermann
Language And Media Processing Laboratory
University of Maryland
College Park, MD. 20770
{lj,daniel,doermann}@cfar.umd.edu

Abstract

Compared to scanned images, document pictures captured by camera can suffer from distortions due to perspective and page warping. It is necessary to restore a frontal planar view of the page before other OCR techniques can be applied. In this paper we describe a novel approach for flattening a curved document in a single picture captured by an uncalibrated camera. To our knowledge this is the first reported method able to process general curved documents in images without camera calibration. We propose to model the page surface by a developable surface, and exploit the properties (parallelism and equal line spacing) of the printed textual content on the page to recover the surface shape. Experiments show that the output images are much more OCR friendly than the original ones. While our method is designed to work with any general developable surfaces, it can be adapted for typical special cases including planar pages, scans of thick books, and opened books.

1. Introduction

Digital cameras have proliferated rapidly in recent years due to their small size, ease of use, fast response, rich set of features, and dropping price. For the OCR community, they present an attractive alternative to scanners as imaging devices for capturing documents because of their flexibility. However, compared to digital scans, camera captured document images often suffer from many degradations both from intrinsic limits of the devices and because of the unconstrained external environment. Among many new challenges, one of the most important is the distortion due to perspective and curved pages. Current OCR techniques are designed to work with scans of flat 2D documents, and cannot handle distortions involving 3D factors.

One way of dealing with these 3D factors is to use special equipments such as structured light to measure the 3D range data of the document, and recover the 2D plane of the page [1, 12]. The requirement for costly equipment, however, makes these approaches unattractive.

The problem of recovering planar surface orientations from images has been addressed by many researchers inside the general framework of shape estimation [5, 7, 10], and applied to the removal of perspective in images of flat documents [3, 4, 11]. However, page warping adds a non-linear, non-parametric process on top of this, making it much more difficult to recover the 3D shape. As a way out, people add in more domain knowledge and constraints. For example, when scanning thick books, the portion near the book spine forms a cylinder shape [8], and results in curved text lines in the image. Zhang and Tan [16] estimate the cylinder shape from the varying shade in the image, assuming that flatbed scanners have a fixed light projection direction. In terms of camera captured document images, Cao et al. [2] use a parametrical approach to estimate the cylinder shape of an opened book. Their method relies on text lines formed by bottom-up clustering of connected components. Apart from the cylinder shape assumption, they also have a restriction on the pose that requires the image plane to be parallel to the generatrix of the page cylinder. Gumerov et al. [6] present a method for shape estimation from single views of developable surfaces. They do not require cylinder shapes and special poses. However, they require correspondences between closed contours in the image and in the unrolled page. They propose to use the rectilinear page boundaries or margins in document images as contours. This may not be applicable when part of the page is occluded.

Another way out is to bypass the shape estimation step, and come up with an approximate flat view of the page, with what we call *shape-free* methods. For scans of thick bound volumes, Zhang and Tan [15] have an-

Beispiel B

/Beispiel B

Flattening Curved Documents in Images
Jian Liang, Daniel DeMenthon, David Doermann
Language And Media Processing Laboratory
University of Maryland
College Park, MD, 20770
{lIdaniel,doermann} gefar.urnd edu
Abstract
Compared to scanned images, document pictures
captured by camera can suffer from distortions due to
perspective and page warping. it is necessary to restore
a frontal planar view of the page before other
OCR techniques can be applied. In this paper we describe
a novel apprvach for flattening a curved document
in a single picture captured by an uncalibrated
camera. To our knowledge this is the first reported
method able to process general curved documents in images
without camera calibration. We prooose to model
the page surface by a developable surface, and exploit
the properties (parallelism and equal line spacing) of
the printed textual content on the page to recover the
surface shape. Experiments show that the output images
are much more OCR friendly than the 'riq:Tull
ones. While our method is designed to work with any
general developable surfaces, it can be adapted for typical
special cases incluuing pianar pages, scans of (hick
books, and opened books.
1. Introduction
Digital cameras have proliferated rapidly in recent
years due to their small size, ea.se of use, fast response,
rich set of features, and dropping price. For the OCR
community, they present an attractive alternative to
scanners as imaoing devices for capturing documents
because of their flexibility. However, compared to digital
sca.ns, camera captured document image often suffer
from many degradations both from intrinsic limits
of the device and because of the unconstrained external
environment. Among many new challenges, one of
the most important is the distortion due to perspective
and cuntd pages. Current OCR techniques are
designed to work with scans of fiat 2D documents, and
cannot handle distortions involving 3D factors.
One way of dealing with these 3D factors i.s to use
special equipments such as structured light to measure
the 3D range data of the document, and recover the 12D
plane of the page :1, 121. The requirement for eckstly
equipment, however, makes these approaches unattractive.

The problem of recovering planar surface orientations
from images has been addressed by many researchers
inside the general framework of shape estimation
[5, 7, 101, and applied to the removal of perspective
in images of flat documents 13, 4, 111. However, page
warping adds a non-linear, non-parametric process on
top of this, making it much more difficult to recover the
3D shape. As a way out, people add in more domain
knowledge and constraints. For example, when scanning
thick books, the portion near the book spine forms
a cylinder shape [81, and results in curved text lines in
the image. Zhang and Tan E161 estimate the cylinder
shape from the varying shade in the image, assuming
that flatbed scanners have a fixed light projection direction.
In terms of camera captured document images.
Cao et aJ. 121 use a parametrica.1 approach to estimate
the cylinder shape of an opened book. Their method
relies on text lines formed by bottom-up clustering of
connected components. Apart from the cylinder shape
assumption, they also have a restriction on the pose
that requires the image plane to be parallel to the generatrix
of the page cylinder. Gumerov et al. 161 present
a method for shape estimation from single views of developable
surfaces. They do not require cylinder shapes
and special poses. However, they require correspondences
between closed contours in the image and in
the unrolled page. They propose to use the rectilinear
page boundaries or margins in document images as
contours. This may not be applicable when part of the
page is occluded.
Another vbPay out is to bypass the shape estimation
step, and come up with an approximate flat view of the
page, with what we call shape-free methods. For sau-is
of thick bound volumes, Zhang and Tan 1151 have an-

/Beispiel B

Flattening Curved Documents in Images

Jian Liang, Daniel DeMenthon, David Doermann
Language .And Media Processing Laboratory
University of Maryland
College Park, MD, 20770
{lj,daniel,doermann}fl:cfar.umd edu

Abstract
Compared to scanned images, document pictures
captured by camera can suffer from distortions due to
perspective and page warping. it necessary to restore
a frontal planar view of the page before other
OCR techniques can be app/zed. In this paper we describe
a novel approach for flattening a curved document
in a single picture captured by an uncalibrated
camera. To our knowledge this is the first reported
method able to process general curved documents in images
without camera calibration. We propose to model
the page surface by a developable surface, and exploit
the properties (parallelism and equal line spacing) of
the printed textual content on the page to recover the
surface shape. Experiments show that the output 17/1ages
are much more OCR friendly than the original
ones. While our method is designed to work with any
general developable surfaces, st can be adapted for typical
special cases including p/anar pages, scans of thick
books, and opened books.

1. Introduction

years due to their small size, ease of use, fast respo e-nt
Digital cameras have proliferated rapidly in rec
rich et of features, and dropping price. For the OCR
community, they present an attractive alternative to
scanners as imaging devices for capturing documents
because of their flexibility. However, compared to digital
scans. camera (7aptured document images often suffer
from many degradations both from intrinsic limits
of the devices and because of the unconstrained external
en Among man)' new challenges, one of
the most impor pages distortion due to perspec
vironment. tant is the -
tive and curved . Current OCR techniques are
ded to work with scans of flat 2D documents, and
handle distortions
cannot involving 3D factors

One way of dealing with these 3D factors is to use
special equipments such as structured light to measure
the 3D range data of the document. and recover the 21)
plane of the page 11, 12. The requirement for costly
na
equipment. however, makes these approaches uttractive.

The problem of recovering planar surface orientations
from images has been addres.sed by many researchers
inside the general framework of shape estimation
15, 7, 101, and applied to the removal of perspective
in images of flat documents 13, -1, 111. Houever, pag.e
warping adds a non-linear, non-parametric process on
top of this. making it much more difficult to recover the
3D shape. As a way out people add in more domain
knowledge and constraints. For example, when scanning
thick book, the portion near the book spine forms
a cylinder shape 181, and results in curved text lines in
the image. Zhang a.nd Tan 1161 estimate the cylinder
shape from the van.ing shade in the image. zissurning
that flatbed sca.nners ha.ve a fixed light projection direction.
In terms of camera captured document images,
Cao et al. 121 use a pzu-ametrical approach to estimate
the cylinder shape of an opened book. Their method
relies on text lines formed by bottom-up clustering of
connected components. Apart from the cylinder shape
assumption, they also ha.ve a restriction on the pose
that requires the image plane to be parallel to the genera.trix
of the page cylinder. Gumerov et al. 161 present
a method for shape estimation from single views of developable
surfaces. They do not require cylinder sha.pes
and special poses. However, they require correspondences
between closed contours in the image and in
the unrolled page. They propose to use the rectilinear
page boundaries or margins in document ima.ges as
contours. This may not be applicable when part of the
pa.ge is occluded.
Another way out is to bypass the shape estimation
step, and come up with an approximate flat view of the
page. with what we call shape-free methods. For scans
of thick bound volumes, Zha.ng and Ta.n 1151 have an-

other method for straightening curved text lines. They find text line curves by clustering connected components, and move the components to restore straight horizontal baselines. The shape is still unknown but image can be OCRed. Under the same cylinder shape and parallel view assumptions as Cao et. al have, Tsoi et al. [14] flatten images of opened books by a bilinear morphing operation which maps the curved page boundaries to a rectangle. Their method is also shape-free. Although shape-free methods are simpler, they can only deal with small distortions and can not be applied when shape and pose are arbitrary.

Our goal is to restore a frontal planar image of a warped document page from a single picture captured by an uncalibrated digital camera. Our method is based on two key observations: 1) a curved document page can be modeled by a developable surface, and 2) printed textual content on the page forms *texture flow* fields that provide strong constraints on the underlying surface shape [9]. More specifically, we extract two texture flow fields from the textual area in the projected image, which represent the local orientations of projected text lines and vertical character strokes respectively. The intrinsic parallelism of the texture flow vectors on the curved page is used to detect the projected rulings, and the equal text line spacing property on the page is used to compute the vanishing points of the surface rulings. Then a developable surface is fitted to the rulings and texture flow fields, and the surface is unrolled to generate the flat page image.

Printed textual content provides the most prominent and stable visual features in document images [3, 11, 2, 15]. In real applications, other visual cues are not as reliable. For example, shade may be biased by multiple light sources; contours and edges may be occluded. In term of the way of using printed textual content in images, our work differs from [15, 2] in that we do not rely on connected component analysis which may have difficulty with figures or tables. The mixture of text and non-text elements will also make traditional shape-from-texture techniques difficult to apply, while our texture flow based method can still work. Overall, compared to others' work, our method does not require a flat page, does not require 3D range data, does not require camera calibration, does not require special shapes or poses, and can be applied to arbitrary developable document pages.

The remainder of this paper is organized into five sections. Section 2 introduces developable surfaces and describes the texture flow fields generated by printed text on document pages. Section 3 focuses on texture flow field extraction. We describe the details of surface estimation in Section 4, and discuss the experimental

Figure 1. Strip approximation of a developable surface

results in Section 5. Section 6 concludes the paper.

2. Problem Modeling

The shape of a smoothly rolled document page can be modeled by a developable surface. A developable surface can be mapped isometrically onto a Euclidean plane, or in plain English, can be *unrolled onto a plane* without tearing or stretching. This process is called *development*. Development does not change the intrinsic properties of the surface such as curve length or angle formed by curves.

Rulings play a very important role in defining developable surfaces. Through any point on the surface there is one and only one ruling, except for the degenerated case of a plane. Any two rulings do not intersect except for conic vertices. All points along a ruling share a common tangent plane. It is well known in elementary differential geometry that given sufficient differentiability a developable surface is either a plane, a cylinder, a cone, the collection of the tangents of a curve in space, or a composition of these types. On a cylindrical surface, all rulings are parallel; on a conic surface, all rulings intersect at the conic vertex; for the tangent surface case, the rulings are the tangent lines of the underlying space curve; only in the planar case are rulings not uniquely defined.

The fact that all points along a ruling of a developable surface share a common tangent plane to the surface leads to the result that the surface is the envelope of a one-parameter family of planes, which are its tangent planes. Therefore a developable surface can be piecewise approximated by planar strips that belong to the family of tangent planes (Fig. 1). Although this is only a first order approximation, it is sufficient for our application. The group of planar strips can be fully described by a set of reference points $\{P_i\}$ along a curve on the surface, and the surface normals $\{\mathbf{N}_i\}$ at these points.

Suppose that for every point on a developable surface we select a tangent vector; we say that the tangents are *parallel* with respect to the underlying surface

Beispiel C

other method for straightening curved text lines. They
find text line curves by clustering connected components.
and move the components to restore straight
horizontal baselines. The shape is still unknown but
image can be OCRed. Under the same cylinder shape
and parallel view assumptions a.s Cao et. al have Tsoi
et al. 1141 flatten images of opened books br a bilinear
morphing operation which mops the curved page
boundaries to a rectangle. Their method is also shapefree.

Although shape free methods arc simpler, they
can only deal with small distortions and can not be
applied when shape and pose are arbitrary.
Our goal is to rtore a frontal planar image of a
wa.rped document page from a single picture aptured
by an uncalihrated digital camera. Our method LS
based on two kev observations: I) a curved document
page can be modeled by a developable surface, and
2) printed text.ual content on the page forms texture
floe; fields that provide strong constraints on the underlying
surface shape :,91. More speeLcally, 9, extract
two texture flow fields from the textual area in the projected
image. which represent thc local orientations of
projected text lines and vertical character strokes respectiwly.
The intrinsic parallelism of the texture flow
VCCIAn on the curved page is used to detect the projeeted
raluios, and the equ.al text line spacir.g property
on the page is used to compute the vanishing points of
the surface rulings. Then a developable surface is fitted
to the rulings and texture flow fields, and the surface
is unrolled to generate the flat page image.
Printed textual content provides the most prominent
anel stable visual fcaturm in document images
;3, 11, 2, 151. In real applicatiora, other visual cues
are not as reliable. For exaw.ple. shade may be biased
by multiple light sources., contours and edges may be
occluded. In term of the way of using printed textual
content in images, our work differ, from ;15, 21 in that
we do not rely on connected component analysis which
may have difficulty with figures or tables. The mixture
of text and non-text element, will also make traditional
shape-from-texture techniques difficult to apply. while
our texture flow based method can still work. Overall,
compared to others work, our method does not
require a flat page, does not require 3D range data,
does not require camera calibration, does not require
special shapes or pages.
developable document pages.
The remainder of this paper is organized into five
sections. Section 2 introduce, developable surfaces and
describes the tcx-ture flow fields generated by printed
text on document pages. Section 3 focuses on texture
flow field extraction. We describe the details of surface
estimation in Section 4, and disciuss the experimental
results in Section 5. Section 6 conclud,
2. Problem Niodeling
The shape of a smoothly rolled document page can
be modeled by a developable surface.
A developable surface can be mapped isometrically onto a Euclidean
plane, or M plain English, can be unrolled onto a plane
without tearing or stretching. This process is called development.
Development does not chanoe the intrinsic
properties of the surface such as curve length or angle

Rulings play a very important role in defining deformed
by curves.
velopable surfaces. Through any point on the surface
there is one 3.nd only one ruling, except for the degenerated
case of a plane. Any two rulings do not intersect
except for conic vertices. All points along a ruling
share a common tangent plane. It is well known in
elementary differential geometry that given sufficient
differentiability a developable surface is either a plane,
a cylinder, a cone, the collection of the tangents of a
curve in space, or a composition of these types. On a
cylindrical surface, all ruling, are parallel; on a conic
surface. all rulinos intersect at the conic vertex: for the
tangent surf CMC , thc rulings are the tangent lines
of the underlying space curve; only in the planar case
are ruling,s not uniquely defined.
The fact that all point, along a ruling of a developable
surface share a common tangent plane to the
surface leads to the result that the surface is the env,
lope of a one-parameter family of planes, which are its
tangent, planes. Therefore a developable surface can be
piecewise approximated by planar strips that belong to
the family of tangent planes (Fig. 1). Although this is
only a fit order approximation, it is sufficient for our
application. The group of planar strips can be fully described
by a set of reference points kq} along a curve
on the surface, and the surface normals 1,1%:3 at these

Suppose that for every point on a developable surface
we select a tangent vector: and say that, the tanpoints.

gents arc parallel with respect to the underlying surface

Figure 1 Strip approximation of a developable surface

/Beispiel C

other method for straightening curved text lines. They
find text line curves by clustering connected components,
and move the components to restore straight
horizontal baselines. The shape is still unknown but
image can be OCRed. Under the same cylinder shape
and parallel view assumptions a.s Cao et. al have Tsoi
et al. 1141 flatten images of opened books br a bilinear
morphing operation which maps the curved page
boundaries to a rectangle. Their method is also shapefree.
Although shape-free methods are simpler, they
can only deal with small distortions and can not he
applied when shape and pose are arbitra.ry.
Our goal is to rtore a frontal planar image of a
wa.rped document page from a single picture aptured
by an uncalibrated digital camera. Our method LS
based on two key observations: I) a curved document
page can be modeled by a developable surface, and
2) printed text.ual content on the page forms texture
floe; fields that provide strong constraints on the underlying
surface shape :.91. More speeLcally, 9, extract
two texture flow fields from the textual area in the projected
image, which represent thc local orientations of
projected text lines and vertical character strokes respectiwly.
The intrinsic parallelism of the texture flow
VCCIAn on the curved page is used to detect the projeeted
raluios. and the equ.al text line spacir.o property
on the page is used to compute the vanishing points of
the surface rulings. Then a developable surface is fitted
to the rulings and texture flow fields, and the surface
is unrolled to generate the fiat page image.
Printed textual content provides the most prominent
anel stable visual fcaturm in document images
;3. 11; 2, 151. In real applicatiora, other visual cues
are not as reliable. For exaw.ple. shade may be biased
by multiple light sources., contours and edges may be
occluded. In term of the way of using printed textual
content in images, our work differ, from ;15, 21 in that
we do not rely on connected component analysis which
may have difficulty with figures or tables. The mixture
of text and non-text element, will also make traditional
shape-from-texture techniques difficult to apply. while
our texture flow based method can still work. Overall,
compared to others work, our method does not
require a flat page, does not require 3D range data,
does not require camera calibration, does not require
special shapes or pages. and can he applied to arbitrary
developable document pages.
The remainder of this paper is organized into five
sections. Section 2 introduce, developable surfaces and
describes the tcx-ture flow fields generated by printed
text on document pages. Section 3 focubes on texture
flow field extraction. We describe the details of surface
estimation in Section 4. and disciuss the experimental
Figure 1. Strip approximatmn of a dc.,kq).,b1, mil-fa,
results in Section 5. Section 6 conclud, tbc pdprr
2. Problem Modeling
The shape of a smoothly rolled document page can
be modeled lay a developable surface A developable
surface can be mapped Laometncally onto a Euclidean
plane, or in plain English, can be unrolled onto a plane
aithout tearing or stretching. This groem is called development.
Development does not change the intrinsic
properties of the surface such as curve length or angle
formed by curves.
Rulings play a very important role in defining developablesurfaces.
Through any point on the surface
there is one and only one ruling, except for the degenerated
case of a plane. Any two rulings do not intersect
except for conic vertices. All points wong a ruling
share a cornmon tangent plane. It is well known in
elementary differential geometry that given sufficient
differentiability a developable surface is either a plane,
a cylinder, a cone, the collection of the tangents of a
curve in space, or a composition of these type,. On a
cylindrical surface, all rulings are parallel; on a conic
surface. a rulings intersect at thc conic vertex; for the
tangent surface case, the rulings are the ungent lines
of the underlying space curve; only in the planar CA-Ne
tue raling not uniquely defined.
The fact that all points along a ruling of a developable
surface share a common truagent plane to the
surface leads to the result that the surface is the envelope
of a one-parameter family of planes, which are its
tangent planes. Therefore a developable surface can be
piecewise approximated by planar strips that belong to
the family of tangent planes (Fig. 1). Although this sa
only a first order approximation. it is sufficient for our
application. The group of planar strips can be fully described
by a set of reference points i),} along a curve
on the surface, and the surface normal., 14N,} at these
points.
Suppose that for every point on a developable surface
we select a tangent vector; we say that the tangent,
are parallel with respect to the underlying surface


Mach mit!
Die neue große
Sportspass-Aktion!
Für Erwachsene
Stepp-Aerobic · Body-Shape · Rücken-Fit
Präventivsport · Herzsport · Seniorensport
Basketball · Handball · Tischtennis · Volleyball
Für Schüler/Innen
Akrobatik (ab 6 J.) · Ballett (ab 6 J.)
Kindertanz (ab 6 J.) · Street-Dance (ab 9 J.)
Basketball (ab 9 J.) · Handball (ab 6 J.)
Turnen+Ballspiele (ab 6 Jahre)
Für Kids
Eltern-Kind-Turnen (1-2 J. · 3-4 J. · ab 5 J.)
Ballett (ab 4 J.) · Kindertanz (ab 4 J.)
HC Steglitz e.V.
Info-Telefon: 791 10 17
www.hcsteglitz.de

Beispiel D

/Beispiel D

Mach mit!
Die neue qrol3e
Sportspass-Aktion!
Für Erwachsene
Stepp-Aerobic Body-Shape Rücken-Fit
Praventivsport Herzsport Seniarensport
Basketball • Handball Tischtennis Volleyball
Fier Schillerlinnen
Akrobatik (ab 6 J.) Ballett (ab 6 J.)
Kindertanz (ab 6 J.) Street-Dance (ab 9 J.)
Basketball (ab 9 J.) Handball (ab 6 J.)
Turnen+Ballspiele (ab 6 Jahre)
Air Kids
Eltem-Kind-Turnen (1-2 J. 34 J. ab 5 J.)
Ballett (ab 4 J.) Kindertanz (ab 4 J.)

1-IC Steglitz a V.
InfosTelefon: 791 10 17
www,hcsteglitz.de

Mach mitI
Die neue qroI3c
Sportspass-AktioiiI
Far Erviachsene
Stepp-Aerobc • Body-Shape • RUcken-Fit
P(aventivsport • Herzsport • Seniorensport
Basketball halerl
• Handball • Tischtennis • Volleyball
Fiir ScInnen . .
Akrobatik (al) 6 J.) • Ballett (ab 6 J.)
tendertanz e (ab 6 Jahre)
(ab 6 J.) • Street-Dance (ab 9 J.) •.
f3asketball (ab 9 J.) • Handball (ab 6 J.)
Turnen+Ballspiel
Far Kids
Eltern-Kind .Turnen (1-2 J. - 3-4 J. • ab 5 J.)
Ballett (ab 4 J.) - Kindertanz (ab 4 J.)
Prob stunde qratis! Zentrale Laqel -
NC Steqlitz e.V. ,
Info-Telefon: 791 10 17
www.hcsteglitz.de

of ACIVS 2002 (Advanced Concepts for Intelligent Vision Systems), Ghent, Belgium, September 9-11, 200

ON THE RECOVERY OF ORIENTED DOCUMENTS FROM SINGLE IMAGES

Paul Clark and Majid Mirmehdi

{pclark|majid}@cs.bris.ac.uk
Department of Computer Science, Univerity of Bristol, Bristol BS8 1UB, UK.

ABSTRACT

A method is presented for the fronto-parallel recovery of text documents in images of real scenes. Initially an extension of the standard 2D projection profile, commonly used in document recognition, is introduced to locate the horizontal vanishing point of the text plane. This allows us to segment the lines of text, which are then analysed to reveal the style of justification of the paragraphs. The change in line spacings exhibited due to perspective is then used to recover the vertical vanishing point of the document. We do not assume any knowledge of the focal length of the camera. Finally, a fronto-parallel view is recovered, suitable for OCR or other high-level recognition. We provide results demonstrating the algorithm's performance on documents over a wide range of orientations.

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 fronto-parallel images. 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. The extraction of oriented documents from camera images is a new challenge in document processing, made possible by high resolution digital cameras, as well as recent developments and demands in the multimedia environment. There has been little research into the recognition of text in real scenes in which the text is oriented relative to the camera. Processing and compensating for such perspective skew 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.

Previous work in estimating the orientation of planar surfaces in still images varies in the assumptions made to achieve this. Ribeiro and Hancock [1] and Criminisi and Zisserman [2] both presented methods which used the distortion of repetitive planar texture to estimate the vanishing points of the plane. Affine transformations in power spectra were found along straight lines in [1], and correlation measures were used in [2] 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 [3] attempted to find orthogonal lines in architectural environments, which were assessed relative to the camera geometry. Murino and Foresti [4] used a 3D Hough transform to estimate the orientation of planar shapes with known rectilinear features. Gool et al. [5] and Yip [6] both found the skewed symmetry of 2D shapes which have an axis of symmetry in the plane, allowing for affine recovery. We require recovery from perspective transformation, and as with these latter works we will use a priori information about the 2D shape we are observing. Other than our recent work in [7], the only other work known to the authors on perspective recovery of text is [8]. The author seeks visual clues in the image which correspond to horizontal and vertical features on the document plane. Unfortunately, the vertical vanishing points of the text plane cannot be robustly estimated when only one vertical clue is present. Examples of this situation are when the document is single-column and when paragraphs are not fully justified.

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 and/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 towards 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 segmented lines of text are then analysed to reveal the style of justification or alignment of the paragraph. De-

Beispiel E

/Beispiel E

of ACIVS 2002 (Advanced Concepts for Intelligent Vision Systemut Ghent, Betztunt Septemher 9.4 j, 200
ON THE RECOVERY OF ORIENTED DOCUMENTS FROM SINGLE IMAGES
Paul Clark and Majid Mirmehdi
{pclarkJmajid}@cs.briseac.uk
Department of Computer Science, Univerity of Bristol, Bristol BS8 I UB, UK.
ABSTRACT
A method is presented for the frento-paratlel recovery of text documents in images of real scenc-s. Initially an extension of the standint 2D projection profile, commonly used in document recogni-
tion, is introduced to locate the horizontal vanishing point of the
text plane. This allows us to segment the lines of text, which arc
then analysed to reveal the style of iustification of the paragraphs.
The change in line spacings exhibited due to perspective is then used
to recover the vertical varushing point of the document. We do not
assume any knowledge of the focal length of the camera. Finally,
a fronto-parallel view is recovered, suitable for OCR or other high-
level recognition. We provide results demonstrating thc alogonthm's
performance on documents over a wide range of orientations.
I. INTRODUCTION
Optical character recopition (OCR) is a long-standing area
of computer vision which in general deals with the problem
of recognising text in skew-compensated fronto-parallel im-
ages. In preparation to apply OCR to text from images of real
scenes, a fronto-parallel view of a segmented region of text
mus-t be produced. This is the issue considered in this paper.
The extraction of oriented documents from camera images
is a new challenge in document processing, made possible
by high resolution digital cameras, as well as recent devel-
opments and demands in the multimedia environment. There
has been little research into thc recognition of text in real
scenes in which the text is oriented relative to the camera.
Processing and compensating for such perspective skew has
applications in replacing the document scanner with a point-
and-click camera to facilitate non-contact text capture, as-
sisting the disabled and/or visually impaired, wearable com-
puting tasks re-quiring knowledge of local text, and general
automated tasks requiring the ability to read where it is not
possible to use a scanner.
Previous work in estimating the orientation of planar surfaces
in still images varies in the assumptions made to achieve this.
Ribeiro and Hancock [1] and Criminisi and Ztsserman [2]
both presented methods which u.sed the distortion of repet-
itive planar texture to estimate the vanishing points of the
plane. Atline transformations in power spectra were found
along straight lines in [I J. and correlation measures were
used in [2] to determine first the orientation of the vanish-
ing 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 mall
area of the image. Rother [3] attempted to find orthogonal
lines in architectural environments, which were assessed rel-
ative to the camera geometry. Munno and Foresti [4] used
a 3D Hough transform to estimate the onentation of planar
shapes with known rectilinear features. Gool et al. [5] and
Yip [6] both found the skewed symmetry of 2D shapes which
have an axis of symmetry in the plane, allowing for atline
recovery. We require recovery from perspective transforma-
tion, and 35 with these latter works we will use a priori in-
formation about the 2D shape we are obsenang. Other thast
our recent work in [7], the only other work kriown to the au-
thors on perspt-ctive recovery of text is [8]. The author seeks
visual clues in the image which correspond to honzontal and
vertical features on the document plane. Unfortunately, the
vertical vanishing points of the text plane cannot be robustly
estunated when only one vertical clue is present. Examples
of this situation are NS, hen the document is single-tolumn and
when paragraphs are not fullv iustified.
Knowledge of the principal vanishing points of the plane on
which text lies is sufficient to recover a fronto-parallel view.
We obsent 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 itnage. Also, para-
graphs often exhibit some form of iustification, either with
stniight margins on the lefi andior right, or if the text is cen-
tred, a central vertical line around which the text is aligned.
In such eases these vertical lines point tovvards the vertical
vanishing point of the text plane. We have therefore con-
centrated our work on the recovery of paragraphs with tlute
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 paragaph model, in this work we ensure
the use of all of the global information about the parawaoh
at one time. The pnnciple of 2D projection profiles are ex-
tended to the problem of locating the horizontal vanishing
point hy maximising the separation of the lines in the para-
graph. The segmented lines of text are then analysed to reveal
tlie stvle of iustification or alignment of the paragraph. De-

/Beispiel E

```
Ro(..ccId of ACIVS 2002 (Advanced Concepts for Intelligent Vision System Ghent, Betztunt Septemher 9.4 j, 200
ON THE RECOVERY OF ORIENTED DOCCNIENTS FRONI SINGLE IMAGES
Paid Clark and .Maild Mirrnchdi
{pclarkJmajid}@cs.briseac.uk

Department of Computer Science, Univerity of Bristol, Bristol BS8 lUB, UK.
ABSTRACT
A method is presented for the fronto-parallel recosery of text doc-
uments in images of real scenes Initially an extension of the stan-
dard 20 projection profile, commonly used in document meat:m-
um. Is introduced to locate the honzontal vanishing point of the
text plane This allows us to segment the line. of text. which are
then analysed to reseal the style of justificanon of the paragraphs
The change in line spacings exhibited due to perspective is then used
to recover the sortical vanishing point ',if the document. q'e do not
ascume any knowledge of the focal length of the camera. Finally,
a fronto-ouallel view is recovered, suitable for OCR or other high-
level recoqution. We pro, ide results demonstrating the alqonthm's
penormance on documents over a wide range of onentations
1. I NTRODUCTION
Optical character recognition (OCR) is a long-standing area
of computer vision which in general deals w ith the problem
of recognising text in skew-compensated fronto-parallel im-
ages. In preparation to apply OCR to text front 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
The extraction of onented documents from ciunera images
is a new challenge in document processing, made possible
by high resolution digital cameras.. well as recent de' el-
opments and demands in the multimedia environment. 'There
has been little research into the recognition of text in real
scenes in which the text is oriented relative to the catnera.
Processing and compensating fOr such perspective skew has
applications in replacing the document scanner with a point-
and-click camera to facilitate non-contact text capture, as-
sisting the disabled antior visually impaired, wearable com-
outino tasks requiring know ledge of local text. and general
automated t.ks requiring the ability to read where it is not
possible to use a scanner.
Previous work in estimating the orientation of planar surfaces
in still images varies in the assumptions made to achieve this.
Ribeiro and Flancock III and Crirmnisi and Zissennan [2]
both presented methods which used the distortion of repe-
nts e planar texture to estimate the vanishing points of the
plane. Afline transformations in power spectra were found
along straight lines in [I], and correlation measures were
used in [2] to determine first the orientation of the vanish-
ing litte and then its positron..Although text has repetiti% e
elements (characters and lins( these elements do not math
c eich other exactly. and sometim s may cover onlriy oa msmall
area of the image. Rother [3] attempted to find orthogonal
lines in architectural environments. which were assessed rel-
ative to the camera geometry. Munno and Forest' [41 used
a 3D Hough transform to estimate the orientation of planar
shapes with known rectilinear features. Gool et al. [5] and
Yip [6] both found the skewed symmetry of 2D shap..es w
have an axis of symmetry in the plane. allowing tor atfine
recover. We require recovery from perspective transtbnna-
tion, and as with these latter work.s we will use a priori in-
fonnation about the 2D shape we are observ ing.. Other than
our recent work in [7], the only other work known to the au-
thors on persoective recovery of text is [8]. The author seek.
visual clues in the imace which correspond to horizontal and
vertical features on the document plane. Unfortunately, the
vertical vanishing points of the text plane not be robustly
estimated when only one .ertical clue is present. Examples
of this situation are when the document is single-column and
when paragraphs are not fully justified.
Knowledge of the principal vanishing points of the plane on
hich lest lies is sufficient to recover a fronto-parallel SiCss
We observe that in a paragraph which is oriented relative to
the camera, the lines of text all point towards the horizontal
vanishin..2 point of the text plane in the image. Also, para-
qraphs often exhibit some form of justification, either with
straight marains on the lett andior right, or lithe text is cen-
tred, a central vertical line around which the text is aliened.
In such eases these vertical lines point tovvards the vertical
vanishing point of the text plane. We have therefore con-
centrated our work on the recovery of paragraphs with tlute
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 paragaph model, in this work we ensure
the use of all of the global information about the parawaph
at one time. The onnciple of 2D projection profiles are ex-
tended to the problem of locating the horizontal vanishing
point hy maximising the separation of the lines in the para-
qraph. The segmented lines of text are then analysed to reveal
tlie style of justification or alignment of the paragraph. De-
```

	VP error	Angular error
HVP using projection profiles	0.129	2.16
VVP using margin intersection	0.0785	2.08°
VVP using line spacings	0.133	3.30°
VVP using margin intersection on left justified paragraphs	1.23	24.5°

Table 2: Average error for the various methods over 10° ≤ 80° in yaw and pitch.

Figure 9: Fronto-parallel recovery of example documents in Figure 6.

performance is poor due to the paragraph's jagged edge, as can be seen in Figure 8(b). Finally, Figure 8(c) shows the accuracy when the line spacings are employed on left-justified paragraphs. This method provides good results comparable to the first for all of the simulated images except those documents oriented beyond 80° in pitch, where the algorithm begins to fail. As with the horizontal vanishing point in Section 3, this may be explained by the orientation of the document becoming nearly perpendicular to the image plane. At such an extreme tilt, even if the lines of text are separated correctly, their proximity in the image means there is little accuracy in position and spacing for the curve fitting. In real world images, documents at such extreme angles cannot practically be read or used by OCR once recovered, hence this failure does not concern us. The advantage of the line spacings method is that it provides consistent results for paragraphs which are not fully justified.

The results for these experiments, and the location of the horizontal vanishing point in Section 3, are shown numerically in Table 2. The vanishing point (VP) error is calculated as the relative distance of the vanishing point from its ground truth, as described in Section 3. The angular error is derived from the final determined orientation of the horizontal and vertical vectors of the text plane. It can be seen that the accuracy of location of the vertical vanishing point is reasonable for both the margin intersection and the line spacings method. As the last row of Table 2 shows, intersecting margins is not suitable for documents with jagged edges.

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 6. This quadrilateral is then used to recover a fronto-parallel viewpoint of the paragraph of text.

The recovered page for the running example may be seen in Figure 9, alongside the second example introduced earlier. Further examples in Figure 10 show the recovery of different styles of paragraphs with left justified and centrally aligned text. Figure 10(a) shows the recovery of a segmented region of a book cover. Despite text of different sizes, and other image noise such as the specularity, the document is recovered robustly. Figure 10(b) shows a centrally justified paragraph which has been recovered at high resolution and is easily readable. In Figure 10(c) a left justified document was correctly identified and recovered, despite occlusion of part of the paragraph. When we applied commercial OCR software to the image in Figure 9(b), 87% of the characters and 94% of the words were recognised correctly.

6. DISCUSSION

We have presented a novel approach to the fronto-parallel recovery of a wide range of documents under perspective transformation in a single image, without knowledge of the focal length of the camera. Projection profiles from hypothesised vanishing points are used to robustly recover the horizontal vanishing point of the text plane, and to segment the paragraph into its constituent lines. Line fitting on the margins and central line of the document is then applied to deduce the style of justification of the paragraphs. To estimate the vertical vanishing point, for fully justified paragraphs the margin lines are intersected. For other styles of documents, the observed difference in the spacings of the lines of text are used to retrieve the tilt of the text plane, and hence the vertical vanishing point. Using the two 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 lines, or is fully justified. An optimisation algorithm for the search for the horizontal vanishing point was presented. In total the process takes around 10 seconds to recover a document, demonstrating its potential and applicability to realtime systems, as well as its suitability for normal scanning tasks.

Whilst the described method is suitable for documents or single paragraphs, due to the use of margins or line spacing to recover the vertical vanishing point, there must be at least two lines present for fully justified paragraphs, and at least

Beispiel F

KVP using projection profiles
VVP using margin intersection
VVP using line spacings
VVP using margin intersection
on left justified paragraphs
VP

error

0.129

0.0785

0.133
•23

Table 2: Average error for the vitious methods over 10v II

80° in yaw and pitch.
performance is poor due to the paragraph's jagged edge. as can be seen in Figure 8(b). Finally, Figure 11(c) shows the accuracy when the line spacings are employed on left-justified paragaphs. This method provides good results comparable to the first for all of the simulated images except those documents onented beyond 80° in pitch, where the algonthrn

begins to fail. As with the horizontal %wishing point ta section 3. this may be explained by the onentation of the document becoming nearly perpendicular to the image plane. N such an extreme tilt. even if the lines of text are separated correctly, their proxirrtity in the image means there s brat accuracy in position and spacing for the cunt fitting. In real world images, documents at such extreme angles cannot practically be read or used by OCR once recovered. hence this failure does not concern us. The advantage of the line spacing.s methodis that it provides consistent results for paragaphs which are not fully justified.
The results for these experiments. and the location of the horizontal vanishing point in section 3, are shown numetically in Table 2. The vanishing point (VP) error is calculated as the relative distance of the vanishing point from its ground truth, as described in section 3. The angular error is derived from the final determined orientation of the honzontal and vertical vectors of the text plane. It can be seen that the accuracy of location of the vertical vanishing point in reasonable for both the margin intersection and the line spacings method. As the last row of Table 2 shows, intersecting. margins is not MI itable for documents with jagged edges.
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. a.s shown in Figure 6. This quadrilateral is then used to recover a fromosparallel view-point of the paragraph of text.
The recovered page for the running example may be seen in Figure 9. alongside the second ex:ample intsoduced
Further examples in Figure 10 show the recovery of different styles of paragraphs with lel justified and centrally iriancd text. Figure 10(a) sh.ows the recovery of a segmented region of a hook covet. Despite text of different titzes. and other image noise such as the speccularity, the document is
Angular
error
2.16
LOU'
3.303
24r
recovered robustly. Figure 1 0(b) shortes a centrall:r tacilned paragraph which las been recovered at high reviluction and is easily readable in Figure let a left justified document was correctly identified and recovere.d, des-plie 'cc:4 Limon of part of the paragraph when we applied commercial OCR software to the image in figure 9(b). s711' o of the chractere and 94lc of the vo ords .doerc recoonhetl ceirrectiv.
6. DIM. I \NU)
bairt presented a nmet apprcat:h the 'ionia-parallel recovery. of a wide ranste tit documents under perspective transformation m a single image. without knolA ledoe of the focal length of the camera. Projection prattles from hypothesised sanishing points are used to robustly recover the horizontal vanishtng point of the text plane, iuid to segment the paragraph into its einstiruent lines. Line fitting on the margins and central line oldie document is then applied to deduce the style of justification of the paragraphs. To estimate the vertical vanishing point, for fully justified paragraphs the margin lines are intersected. For other styles of documents. the obsened difference in the spacings of the lines of text are used to retries e the tilt of the text plane. and hence the vertical vanishing point. Using the two principal vanishing points we find the onentation 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 lines, or is fully justified. An optimisation algorithm for the search for the horizontal vanishing point was presented. In total the process takes around 10 seconds to rccover a docwnent. demonstrating its potential and applicability to reaffirm systems, as well as its suitability for normal scanning tasks.
Whilst the described method is suitable for documents or single paragraphs. due to the use of margins or line spacing to recover the vertical vanishing point. there must be at least two lines present for fully justified parag.raphs. and at least
Fupure 9: Fronto-parallel rem erv of rumple dactun=l3en
Figure 6.

/Beispiel F

igs of ACIVS 2002 (Advanced concepts j'or imelligent nuon
Anoular .

error

2.16

2.08
VP
1ivp using
 vvP usi nroiection nrofiles 0.129
error
vvP no moult' intersection
using line spacings

vvP using margin intersection
on left iustified naraoraohs

Table 2: eraoe error for the various methods over 10 to

in Yaw and oitch.

performance is poor due to the paragraph's jagged edge, as
can be seen in Fioure 8(b). Finallv. Fioure 11(c) shows the ac-
curacy when the line spacings are employed on left-Justified
paraoanhs. This method provides oood results comparable
to the first for all of the simulated images except those doc-
uments onented bevond 80° in pitch. where the aloonthrn

beoins to fail. As with the horizontal %wishino point ta see-
ton 3, this may be explained by the onentation of the docu-
ment becomino nearly perpendicular to the un- ix i d'u-
such an extreme tilt, even if the lines of text jar' el:cal ckd
correctlv. Their proximitv in the unaoe means them ip. int.le
accuracy in position and spacing for the curve fitting in
real world imaoes. Documents tit such extreme .moles cannot
practically be read or used by OCR once recovered, hence
this failure does not concern us. The advantaoe of the line
spacings method is that it provides consistent results for para-
oraphs which are not fully iustified.
The results for these experiments, and the location of the hor-
izontal vanishino point in section 3. are shown numericallv
in Table 2. The vanishing point (VP) error is calculated . the
relative distance of the varnshino point tiom its oround truth.
As described in section 3 the angular error is denved from
thr final determined onentatton of the hontontal and vertical
vectors of the text plane. It can he seen that the accuracy of
location of the vertical vanishino point in reasonable for both
the margin intersection arid the line spacings method. As the
last row of Table 2 shows. intersectino maroins is not suitable
for documents with jagged edges.
Havino found the vanishino points (tithe plane. we ina^ project
two lines front each to describe the left and right marin
and the too and bottom limits of the paras.traph 1 hese lines
are intersected to form a quadrilateral enclosing the text as
shown in Fioure 6. This quadrilateral is then used to recover
a fronto-parallel viewpoint of the paragraph of text.
The recovered paoe for the runnino exa.mple mv be seen in
 figcr 9 alongside the second example intioduced earlier
Further examoles in Fioure 10 show the recoverv of different
styles of paragraphs with lel justified and centrally iiignco
text. Fioure 10(a) sh.ows the recoverv of a seomented re-
gion of a hook covet. Despite text of different titzes, and
other imaoe noise such as the snccularitv. the document is
recovered robustlv Fioure 10(b) shows a centrallv- oas:died
paragraph which has been rxovered a: high resoluta.' and
is easilv readable in Fioure 10fct a len iustitii.xl document
was correctly identified and rrcoserell despite occlusion of
part of the paraoraph. 14 hen we applied commercial (x:R
software to the image in FIL..we 9th), ST'. 01 me ,hanicters
and 94(!. of thc woe wore recoaibel correctl^
6. DISCUSSION
We have presented a novel approach to the fronto-parallel
recovery of a v. ide range of documents under perspective
transformation in a sinole imaoe. Without knowledoe of the
focal length of the camera. Projection profiles from hypoth-
esised vanishino points are used to robustlv reco.er the hor-
izontal vanishing point oldie text plane. andre sement the
paraoraph into its constituent imes. Line tinino on the mar-
gins and central line of the document is then applied to de-
duce the style of justification ot the paraoraphs. To estmute
the vertical vanishing point, for tully justified paragraphs the
maroin lines are intersected. For other stsle.s ot documents.
the obser, ed difference in the spacng.s of the lines of text
are used to retnese the tilt of the text plane. and hence the
vertical vanishing point. Using the two pnneipal vanishing
points we find the onenution of the text plane and rrsover
a fronto-parallel vicw. The algonthm pertomis well for a
wide ranoe f.if paraoraphs. provided each paraoraph has at
least three lines_ or is fully iustilicxl_ An optimisation algo-
rithm for the search for the horizontal vanishino point w.
presented. In total he process takes around 10 seconds to
recover a document demonstrating its potential and applica-
tar/ to tealtime systems, as well as its suitability for nunnal
scannino tasks.
Whilst the described method is suitable for documents or sin-
ole paraoraphs. due 10 the use of maroins or line spacino to
recover the vertical vanishing point, there mast bl at jesa_cst
two lines present for fully justified paragraphs. and . at t tal

Usage of continuous skeletal image representation for document images de-warping.

Anton Masalovitch, Leonid Mestetskiy
Moscow State University, Moscow, Russia
anton_m@abbyy.com, l.mest@ru.net

Abstract

In this paper application of continuous skeletal image representation to documents' image de-warping is described. A novel technique is presented that allows to approximate deformation of interlinear spaces of image based on elements of image's skeleton that lie between the text lines. A method for approximation of whole image deformation as combination of single interlinear spaces deformations is proposed and representation of it in the form of 2-dimensional cubic Bezier patch is suggested. Experimental results for batch of deformed document images are given that compare recognition quality of images before and after de-warping process. These results prove efficiency of the proposed algorithm.

1. Introduction

All the modern OCR systems assume that text lines in a document are straight and horizontal while in real images they are not. Image can be deformed before recognition in various ways. For example, if a thick book is scanned, text lines on the scan may be wrapped near the spine of book. If a digital camera is used to retrieve the image instead of a scanner, the text lines may be still wrapped because of low-quality optics of digital cameras. One important example of such deformation is the rounding of an image on borders as result of barrel distortion. Moreover, several types of deformation could be applied to the same image, making it impossible to build a precise model of image deformation. This is how the task of image de-warping appears.

The approach proposed in this paper is based on the construction of outer skeletons of text images. The main idea of the proposed algorithm is based on the fact that it is easy to mark up long continuous branches that define interlinear spaces of the document in outer skeletons. We approximate such branches by cubic Bezier curves to find a specific deformation model of each interlinear space of the document. On the basis of a set of such interlinear spaces' approximations, the whole approximation of the document is built in the form of a 2-dimensional cubic Bezier patch. After all this work is completed, we can de-warp an image using obtained approximation of image deformation.

This work is an extension of the article [1]. In this paper new method of automatic search for interlinear branches of skeleton is described. Also iteration method of image deformation approximation adjustment is given.

To test our algorithm we compare recognition results for a batch of images before and after the de-warping process.

2. Existing solutions

Algorithm of automatic image de-warping is needed nowadays for automatic OCR systems. Plenty of algorithms for image deformation approximation appeared in the last several years (see for example [7-11]). Unfortunately, most of these algorithms have some disadvantages that make them unusable for commercial OCR systems.

Existing solutions can be divided to three approaches:

First approach is to single out text lines by combining close black objects and then approximating each line shape using some characteristic points of line's black objects. For example, one can approximate text lines' shape by using middle points of black objects' bounding rectangles. Main disadvantage of this approach is that it is hard to define such characteristic points of black objects that can give a stable approximation of line shape.

Second approach is to build a model of possible deformation of an image and then try to apply this model for a specific image. Main disadvantage of this method is that it is almost impossible to build a complete model of image deformation. And if such a model describes only one type of deformation, one

Beispiel G

/Beispiel G

Usage of continuous skeletal image representation for document images de-warping.
Anton Masalovita, Leonid Mestetskiy
Moscow State University. Moscow. Russia
aeon m(tiabb %cam limest(drurnet
Abstract
In this paper application of continuous skeletal image representation to documents' image de-warping is described. A novel technique is presented that allows to approximate deformation of interlinear spaces of image based on elements of image's .skeleton that lie between the text lines. A method sfor approximation of whole image deformation as combination of single interlinear spaces deformations is proposed and representation of it in the form of 2-dimensional cubic Baier patch is &suggested. Experimental results for batch of deformed document images are given that compare recognition qudity of images before and after de-warping process. These results prove efficiency of the proposed algorithm.
I. Introduction
All the modern OCR systems assume that text lines in a document are straight and horizontal while in real images they are not. Image can be deformed before recognition in various w-ays. For example, if a thick book is scanned, text lines on the scan may bc wrapped near the spine of book. If a digital camera is used to retrieve the image instead of a scanner, the text lines may be still wrapped because of low-quality optics of digital cameras. One important example of such deformation is the rounding of an image on borders as result of bare! distortion. Moreover, several types of deformation could be applied to the same image, making it impossible to build a precise model of image deformation This is how the task of image de-warping appears.
Thc approach proposed in this paper is based on the constniction of outer skeletons of text images. The main idea of the proposed algorithm is based on the fact that it is easy to mark up long continuous branches that define interlinear spaces of the document in outer skeletons. We approximate such branches by cubic Bezier curves to find a specific deformation model of each interlinear space of the document On the basis of a set of such interlinear spaces' approximation_s, the whole approximation of the document is built in the form of a 2-dimensional cubic Bezier patch. After all this work is completed, we can de-warp an image using obtained approximation of image deformation. This work is an extension of the article [I]. In this paper new method of automatic search for interlinear branches of skeleton is described. Also iteration method of image deformation approximation adiustment is given.
To test our algorithm we compare recognition results for a batch of images before arid atler the de-warping process.
2. Existing solutions
Algorithm of automatic image de-warping is needed nowadays for automatic OCR systems. Plenty of algonthrns for image deformation approximation appeared in the last several years (see for example [7-H]). Unfortunately, most of these algorithms have some disadvantages that make them unusable for commercial OCR systems.
Existing solutions can be divtded to thn.we approaches:
First approach is to single out text lines by combining close black objects and then approximating each line shape using some characteristic points of line's black objects. For example, one can approximate text lines' shape by using middle points of black objects' hounding rectangles. Main disadvantage of this approach is that it is hard to define such characteristic points of black objects that can give a stable approximation of line shape.
Second approach is to build a model of possible deformation of an image and then try to apply this mo-del for a specific image. Main disadvantage of this method is that it is almost impossible to build a complete model of image defonnation. And if such a model describes only one type of deformation, one

/Beispiel G

Usage of continuous skeletal image representation for document images de-sarping.
Anton Masalovitch, Leonid Mestetsidy
Nosrow State Universe. Moscow. Russia
antnn nilitahhci %cam, Imectta ru net
Abstract
M this paper application of continuous skektal Image representation to documents' intage de-warpinq is described. A novel technique is prewnted that allows to approximate deformation of interlinear spaces of image based on elements of image's skeleton that lie between the text lines. A method .for approximation of whole image deformation as combination of single interlinear spaces defanations is proposed and representation of it in the _fonn of 2-dimensional cubic Sezier patch is suggested. Experimental results .for batch of deformed document images are given that conzpare recognition quality of images before and after de-warpinq process. These results prove efficiency of the proposed algorithm.
I. Introduction
All the modern OCR systems assume that text lines in a document are straight and horizontal while in real images thev are not. Image can be deformed before recognition in various ways. For example, if a thick book is scanned, text lines on the scan may be wrapped near the spine of book. If a digital camera is used to retrieve the image instead of a scanner, the text lines rnay be still wrapped because of low-qualitv optics of digital cameras. One important example of such deformation is the rounding of an image on borders as result of bare! distortion. Nioreover. several types of deformation could be applied to the same image, making it impossible to build a precise model of tillage deformation This is how the task of image de-warpinq appears.

The main idea of the proposed algorithm IN based on the fact that it is easy to mark up long continuous branches that define interlinear spaces of the docuinent in outer skeletons. We approximate such branches by cubic e.ezier curves to find a specific deformation model of 45
each interlinear space of the document On the basis of a set of such interlinear spaces' approximations, the whole approximation of the document is built in the form of a 2-dimensional cubic Bezier oatch. After all this work is completed, we can de-warp an imac.Ne usinq obtained approximation of imaae deformation. This work is an extension of the article [II In this paper new method of automatic search for interlinear branches of skeleton is descnbed. Also iteration method of image deformation approximation adjustment is qiven.
To test our algorithm we compare recognition results for a batch of images before and after the de-warpinq process.
2. Existinq solutions
Algorithm of automatic image de-warpinq is needed nowadays for automatic OCR systems. Plenty of algorithms for image deformation approximation appeared in the last several vears (see for example [7-I IO. Unfortunately, most of these algontlms have some disad'antages that make them unusable for commercial OCR systems.
Existinq solutions can he divided to three approaches:
- First approach is to singIe out text lines by combininq close black objects and then approximatinq each line shape usinq some characteristic points of line's black objects. For example, one can approximate text lines' shape bv using middle points of black objects' bounding rectanqles. Main disadvantage of this approach is that It is hard to define such characteristic points of black objects that can qive a stable approximation of line shape.
- Second approach is to build a model of possible deforniation of an Image and then try to apply this model for a specific image. Main disadvantaq,e of this method is that it is almost impossible to build a complete model of image deformation. And if such a model describes only one tvpe of deformation, one

Perspective Correction Methods for Camera-Based Document Analysis

L. Jagannathan and C. V. Jawahar
Center for Visual Information Technology
International Institute of Information Technology
Gachibowli, Hyderabad, India

Abstract

In this paper, we describe a spectrum of algorithms for rectification of document images for camera-based analysis and recognition. Clues like document boundaries, page layout information, organisation of text and graphics components, apriori knowledge of the script or selected symbols etc. are effectively used for removing the perspective effect and computing the frontal view needed for a typical document image analysis algorithm. Appropriate results from projective geometry of planar surfaces are exploited in these situations.

1. Introduction

Document images are omnipresent. Textual content in the form of books, newspapers and articles have been traditionally digitized using at-bed scanners and read with the help of OCRs. These reading systems may not be appropriate in situations where mobile, portable or non-contact reading systems are needed. Cameras, which can scan text without contact even on non-planar surfaces, is an emerging alternative to the conventional scanners. In general, cameras are small in size, lightweight and easy to use. Although many of the present day scanners outperform the popular cameras in resolution, the cameras remain attractive alternative especially in situations described above and for non-critical applications. Advances in sensor technology is expected to take the camera-based systems more favorable.

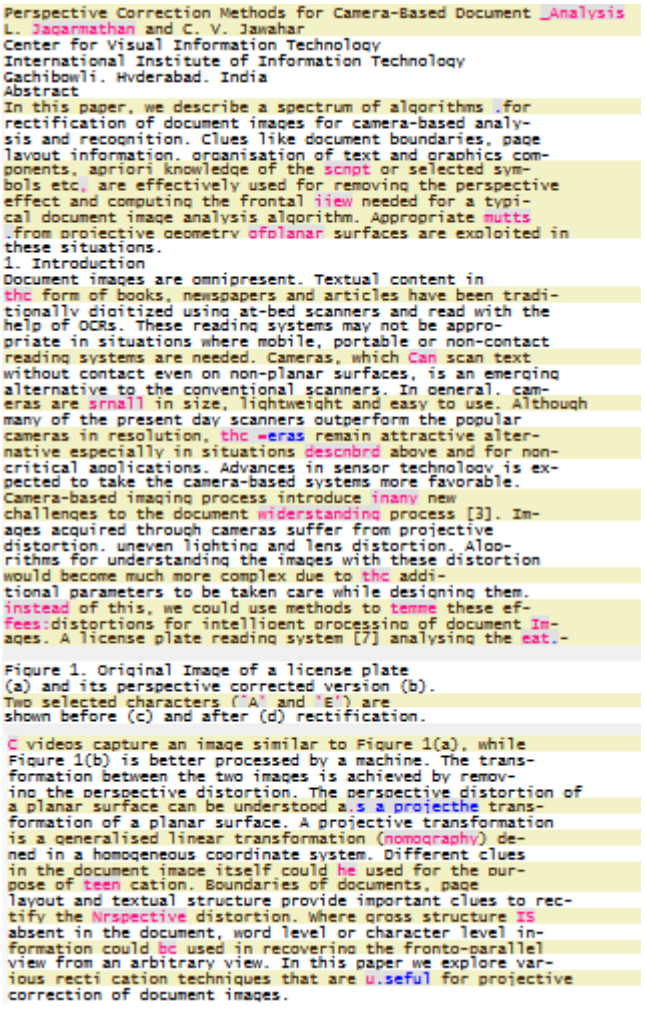
Camera-based imaging process introduce many new challenges to the document understanding process [3]. Images acquired through cameras suffer from projective distortion, uneven lighting and lens distortion. Algorithms for understanding the images with these distortion would become much more complex due to the additional parameters to be taken care while designing them. Instead of this, we could use methods to remove these effects/distortions for intelligent processing of document images. A license plate reading system [7] analysing the traf-

Figure 1. Original Image of a license plate (a) and its perspective corrected version (b). Two selected characters ('A' and 'E') are shown before (c) and after (d) rectification.

c videos capture an image similar to Figure 1(a), while Figure 1(b) is better processed by a machine. The transformation between the two images is achieved by removing the perspective distortion. The perspective distortion of a planar surface can be understood as a projective transformation of a planar surface. A projective transformation is a generalised linear transformation (Homography) defined in a homogeneous coordinate system. Different clues in the document image itself could be used for the purpose of rectification. Boundaries of documents, page layout and textual structure provide important clues to rectify the perspective distortion. Where gross structure is absent in the document, word level or character level information could be used in recovering the fronto-parallel view from an arbitrary view. In this paper we explore various rectification techniques that are useful for projective correction of document images.

Beispiel H

/Beispiel H



Perspective Correction Methods for Camera-Based Document _Analysis
L. Jagarmathan and C. V. Jawahar
Center for Visual Information Technology
International Institute of Information Technology
Gachibowli. Hyderabad. India
Abstract
In this paper, we describe a spectrum of algorithms .for rectification of document images for camera-based analysis and recognition. Clues like document boundaries, page layout information. organisation of text and graphics components, apriori knowledge of the scnpt or selected symbols etc, are effectively used for removing the perspective effect and computing the frontal iiew needed for a typical document image analysis algorithm. Appropriate mutts .from projective geometry ofplanar surfaces are exploited in these situations.
1. Introduction
Document images are omnipresent. Textual content in thc form of books, newspapers and articles have been traditionally digitized using at-bed scanners and read with the help of OCRs. These reading systems may not be appropriate in situations where mobile, portable or non-contact reading systems are needed. Cameras, which Can scan text without contact even on non-planar surfaces, is an emerging alternative to the conventional scanners. In general. cameras are srnall in size, lightweight and easy to use. Although many of the present day scanners outperform the popular cameras in resolution, thc =eras remain attractive alternative especially in situations descnbrd above and for non-critical applications. Advances in sensor technology is expected to take the camera-based systems more favorable.
Camera-based imaging process introduce inany new challenges to the document widerstanding process [3]. Images acquired through cameras suffer from projective distortion. uneven lighting and lens distortion. Aloorithms for understanding the images with these distortion would become much more complex due to thc additional parameters to be taken care while designing them. instead of this, we could use methods to temme these effees:distortions for intellioent processino of document Images. A license plate reading system [7] analysing the eat.-

Figure 1. Original Image of a license plate (a) and its perspective corrected version (b). Two selected characters ('A' and 'E') are shown before (c) and after (d) rectification.

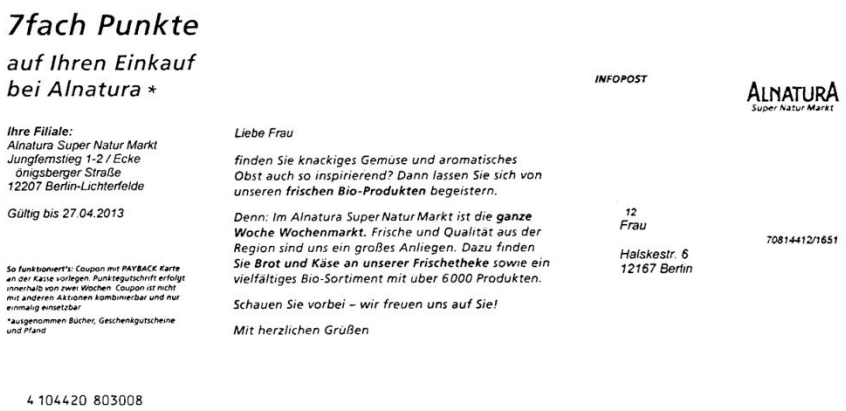
C videos capture an image similar to Figure 1(a), while Figure 1(b) is better processed by a machine. The transformation between the two images is achieved by removino the perspective distortion. The perspective distortion of a planar surface can be understood a.s a projecthe transformation of a planar surface. A projective transformation is a generalised linear transformation (nomography) defined in a homogeneous coordinate system. Different clues in the document image itself could he used for the purpose of teen cation. Boundaries of documents, page layout and textual structure provide important clues to rectify the Nrspective distortion. Where gross structure IS absent in the document, word level or character level information could bc used in recoverino the fronto-parallel view from an arbitrary view. In this paper we explore various recti cation techniques that are u.seful for projective correction of document images.

/Beispiel H

Perspective Correction Methods for Camera-Based Document Analysis
L. Jagannathan and C. V Jawahar
Center for Visual Information Technology
International Institute of Information Technology
Gachibowli. Hyderabad. India
Abstract
In this paper, we describe a spectnim of algorithms Ibr rectification of document images for camera-based analy- sis and recognition. Clues like document boundaries, page lavout information. organisation of tett and graphics com- ponents, apriori knowkdge of the script or selected sym- bols etc are effectively used for removing the perspective effect and computing the frontal •iew needed for a te- cal document image analysis algorithm. Appropriate results front projective geometry ofplanar surfaces an. exploited in these situations.
Document images are omnipresent. Textual content in
1. Introduction

Document images are omnipresent. Textual content in thc form of books, newspapers and articles have been tradi- tionally digitized using at-bed scanners and read with the help of OCRs. These reading systems may not be appro- priate in situations where mobile, portable or non-contact reading svstems are needed. Cameras. which Can scan text without contact even on non-planar surfaces, is an emerging alternative to the conventional scanners. In general, cam- eras are srnall in size, lightweight and easy to use. Although many of the present day scanners outperform the popular cameras in resolution. thc •eras remain attractive alter- native especially in situations descnbrd above and for non- critical applications. Advances in sensor technology is ex- pected to take the camera-based systems more favorable.
Camera-based imaging process introduce inany new challenges to the document •iderstanding process [3]. Im- ages acquired through cameras suffer from projective distortion, uneven lighting and lens distortion. Algo- rithms for understanding the images with these distortion would become much more complex due to thc addi- tional parameters to be taken care while desioning them. instead of this, we could use methods to temme these ef- fees:distortions for intelligent processing of document Im- ages. A license plate reading system [7] analysing the eat.-

Figure 1. Original Image of a license plate d version (b). (a) and its perspective correctected version
Two selected characters CA' and 'El are shown before (c) and after (d) rectation.
C videos capture .in image Figure ltai, while Figure I (bl is better processed by a machine trans- fOrmation between the two images Is achieved by remov- ing the perspective distortion The perspective distortion of surface can be understood as a projective trans- formation of a planar surface. A projec.tne transformation is a generalised linear transformation (Homography) de- nied in a homogeneous coordinate system. Ditterent clues in the document image itself could bc uscd the pur- pose of recti cation. Boundaries of documents. page layout and tcxtual structure provide important clues to rec- tify the perectox distortion. Where gross Structur ls absent in the document, word level or characCter level in formation could be used in recovering the fronto-parallel View' from an arbitrary viem. In this paper var- ious recIicaIion techniques that useful for Projective correction of document images.



7fach Punkte

auf Ihren Einkauf
bei Alnatura *

Ihre Filiale:
Alnatura Super Natur Markt
Jungfernstieg 1-2 / Ecke
önigsberger Straße
12207 Berlin-Lichterfelde

Gültig bis 27.04.2013

So funktioniert's: Coupon mit PAYBACK Karte an der Kasse vorlegen. Punktegutschrift erfolgt innerhalb von zwei Wochen. Coupon ist nicht mit anderen Aktionen kombinierbar und nur einmalig einsetzbar.
*ausgenommen Bücher, Geschenkgutscheine und Pfand

4 104420 803008

Liebe Frau

finden Sie knackiges Gemüse und aromatisches Obst auch so inspirierend? Dann lassen Sie sich von unseren **frischen Bio-Produkten** begeistern.

Denn: Im Alnatura Super Natur Markt ist die **ganze Woche Wochenmarkt.** Frische und Qualität aus der Region sind uns ein großes Anliegen. Dazu finden Sie **Brot und Käse an unserer Frischetheke** sowie ein vielfältiges Bio-Sortiment mit über 6000 Produkten.

Schauen Sie vorbei – wir freuen uns auf Sie!

Mit herzlichen Grüßen

INFOPOST

ALNATURA
Super Natur Markt

12
Frau

Halskestr. 6
12167 Berlin

70814412/1651

Beispiel I

/Beispiel I

```
7fach Punkte
auf Ihren Einkauf
bei Alnatura *
Ihre Filiale:
Alnatura Super Natur Markt
Jungtemstieg 1-2 / Ecke
onigsberger Strafie
12207 Berfin-Lichterfelde
&Bug bis 27.04.2013
So tunknonteres: Coupon mit PAYBACK Karte
an der Kane vorlegen. Punktegutschrift erfolgt
onerhalb von awes Wochen Coupon St nest
mot anderen Aktionen kombinierbar und nur
ernmahg ennsetzber
•ausgenommen Bücher, Geschenkgutscheine
und Pfand
4 104420 803008
Liebe Frau ,
finden Sie knackiges Gemine und aromatisches
Obst auch so inspirierend? Dann lassen Sie sich von
unseren frischen Bio-Produkten begeistern.
Denn: Im Alnatura Super Natur Markt ist die ganze
Woche Wochenmarkt. Frische und Qualitat aus der
Region sind uns ein grones Anliegen. Dazu finden
Sie Brot und Käse an unserer Frischetheke some ein
vielfaltiges Bio-Sortiment mit Ober 6000 Produkten_
Schauen Sie vorbei wir freuen uns auf Sie!
Mit herzlichen GrOßen

INFO POST


12
Frau
Halskestr. 6
12167 Berlin
ALNATURA
Super Natur Markt
70814412/1651



7fach Punkte
auf Ihren Einkauf
bei Alnatura*
Ihre Filiale:
Alnatura Super Natur Markt
Jungfernstieg 1-2 / Ecke
Kbnigsberger Straße
12207 Bedin-Lichterfelde
Gültig bis 27.04.2013
So funktioniert's: Coupon mit PAYBACK Karte
an der Kasse vorlegen. Punktegutschrift erfogt
]
innerhalbvon zwei Wochen. Coupon ist nicht
rot anderen Aktionen kombinierbar und nur
einrnalig in
eset:bar. .
'ausgenommen Bucher Geschenkgutscheine und Pfand

Liebe Frau ,
finden Sie knackiges GernUse und aromatisches
Obst auch so inspirierend? Dann lassen Sie sich von
unseren frischen Bio-Produkten begeistern.
Denn: Im Alnatura SuperNaturMarkt ist die ganze a
VVoche VVochenmarkt. Frische und Qualität aus der
Region sind uns ein grof3es Anliegen. Dazu finden
Sie Brot und Käse an unserer. Frischetheke sowie ein
vielfaltiges Bio-Sortiment mit über 6000 Produkten.
Schauen Sie vorbei - Wir freuen uns auf Sie
Mit herzlichen GrUßen
INFOPOST
12 *
Frau

Halskestr. 6
12167 Berlin

A LNATURA
Super Natur Markt
727081441211651
```

model of the discrete binary image. In the same Euclidean plane R^2, we define the polygonal figure μ as the set of the points formed by association of a finite number of non-overlapping bounded closed domains. This figure is then a model of the continuous binary image. There is a problem consists in the construction of the figure μ that adequately describes properties of the discrete image B. In mathematical terms this problem is posed as an approximation of a discrete object with a continuous object. Natural criteria of good approximations should satisfy the following natural criteria:

1) $B \subset \mu$, $W \subset R^2 \setminus \mu$, where means closure of a set;

2) Let $x, y \in I$ be a pair of adjacent nodes of the lattice and s_{xy} be a segment connecting these nodes. Then if $x, y \in \mu$, then $s_{xy} \in \mu$, and if $x, y \notin \mu$ then $s_{xy} \cap \mu = \emptyset$.

The first condition means that the figure covers all black points of a discrete image and all white points lie either outside of or on the boundary of the figure. The second condition can be reduced to the condition that the boundary of μ lies in the interface between white and black boundary points of the discrete image.

Let M be the set of all figures satisfying conditions 1 and 2. Any of them can be considered a continuous model of a binary image with acceptable accuracy. As we are going to build a skeleton of this figure, the most convenient representation for us is the figure with a piecewise linear boundary, since for such figures there are effective algorithms for construction of a skeleton. In this situation it is natural to choose from M a polygonal figure (PF) with minimal perimeter (see fig. 2). First, such PF exists and it is unique. Second, the number of its vertices is close to minimal among all PF satisfying conditions 1 and 2.

Figure 2. Representation of raster object with polygonal figure with minimal perimeter

The algorithm for solving this problem which requires a single pass over a raster image, has been described in [4].

6. Continuous skeletal representation of an image

The choice of the polygonal figure as a continuous model of the binary image reduces the problem of construction of a skeleton of the image to the well-known medial axis transform [5]. Contrary to discrete images for which the skeleton is determined ambiguously, the concept of a skeleton of a continuous figure has a strict mathematical formulation. The skeleton of a figure is the locus of points of centers of maximal empty circles. An empty circle does not contain any boundary points of the figure. The maximal empty circle is a circle which is not contained in any other empty circle, and which is not congruent to another. Note that empty circles can be thus either internal or external for the domains comprising the figure. Accordingly their centers form internal and external skeletons of the figure (see fig. 3).

Figure 3. Empty circles for polygonal figure and skeleton of polygonal figure.

This definition applies to any type of shape, not just a polygon. However there exist effective algorithms for construction of polygonal figures [4,6]. The algorithm used [2,3] is based on a generalization of Delauney triangulation for a system of sites of two types (points and segments) that comprise a PF boundary. It builds a skeleton in time O(n log n) where n is the number of PF vertices.

Skeleton of polygonal figure can represented as a planar graph, where nodes are points on a plane and bones are straight lines that connect the nodes. In such representation of a skeleton all nodes have no less than three nearest points on the border of the area and all bones lie between two linear fragments of the area border. Later in this article we will use only graph representation of a skeleton.

Let us also define a knot in skeleton as a node with more then two connected bones and final node as a node with only one connected node. And let us define a branch of skeleton as a consistent set of bones that has final node or knot node on each end and does not have knots in the middle of the branch. Later in this article we will operate only with branches of the skeleton and not with single bones.

47

Beispiel J

model of the discrete binary image In the same
Euclidean plane R. we define the polygonal
figure /../ as the set of the points formed by association
of a finite number of non-overlapping bounded closed
domains. This figure is then a model of the continuous
binary image. There is a problem consists in the
construction of the figure # that adequately describes
properties of the discrete image B. In mathenutical
terms this problem is posed as an approximation of a
discrete object with a continuous object. Natural
criteria of good approximations should satisfy the
following natural criteria:
1) Bcp, Wc R2 \/J , where means
closure of a set;
2) Let x.v E I be a pair of adiacent nodes of the
lattice and s. be a segment connecting these nodes.
Then if x,yEp, then s. Ep, and if .x,y p
then s. # = 0 .
The first condition means that the figure covers all
black points of a discrete image and all white points lie
either outside of or on the boundary of the figure. The
second condition can be reduced to the condition that
the boundary of p lies in the interface between white
and black boundary points of the discrete image.
Let M be the set of all figures satisfying conditions
1 and 2. Any of them can be considered a continuous
model of a binary image with acceptable accuracy. As
we are going to build a skeleton of this figure, the most
convenient representation for us is the figure with a
piecewise linear boundary, since for such figures there
are effective algorithms for con.struction of a skeleton.
In this situation it is natural to choose from M a
polygonal figure (PF) with minimal perimeter (see fig.
21, First, such PF exists and it is unique Second. the
number of its vertices is close to minimal among all PF
satisfying conditions 1 and 2.
Figure 2. Representation of raster object with
polygonal figure with minimal perimeter
The algorithm for solving this problem which
requires a single pass over a raster image. has fen
described in VII.
47
6. Continuous skeletal representation of an
image
The chruce of the polgonal figure as a continuous
model of the bine image reduces thc problem of
construction of a skeleton of the image to the w ell-
known medial axis transform 15) Contrary to ttiscrvte
'macs for which the skeleton is determined
ambiquously. the concept of a skeleton of a continuous
figure has a strict mathenatic.al formulation. The
skeleton of a figure is the locus of points of centers of
maximal empty circles. An empty circle does not
contain ass boundary points of the figure. fhe
maximal empty circle is a circle which is not contained
in an v other empty cm:le. and which is not cononieni
to another. Note that empty circles can be thus either
internal or external for the domains comonsinq the
figure. Accordinqly their centers form internal and
extemal skeletons of the figure (see fig. 3).

This definition applies to any type of shape, not just
a polygon. However there exist effective algorithms
for construction of polygonal figures [4,6]. The
algorithm used [2.3] is based on a qeneralization of
Delauney tnanqulation for a system of sites of two
types (points and segments) that comprise a PI'
boundary. It builds a skeleton in time O(n log n) whett
n is the number of PF vertices.
Skeleton of polygonal figure can represented as a
planar qraph, where nodes are points on a plane and
bones are straight lines that connect the nodes. In such
representation of a skeleton all nodes have no less than
thme nearest points on the border of the area and all
bones lie between two linear fraqments of the area
border. Later in this article we will use only q.raph
representation of a skeleton.
Let us also define a knot in skeleton as a node with
more then two connected bones and final node as a
node with only one connected node. And let us define
a branch of skeleton as a consistent set of bones that
has final node or knot node on each end and does not
have knots in the middle of the branch. Later in this
article we will operate only with branches of the
skeleton and not with single bones.
Figure 3. Empty circles for polygonal figure and
skeleton of polygonal figure.

model of the discrete binary image. In the same Euclidean plane R2 , we define the polygonal figure p as the set of the points formed by association of a finite number of non-overlapping bounded closed domains. This figure is then a model of the continuous binary image. There is a problem consists in the construction of the figure p thatadequately descnbcs properties of the discrete image B . In mathematical terms this problem is posed as an approximation of a discrete object with a continuous object. Natural criteria of good approximations should satisfy the following natural criteria:

I) Bcp, WC[R21,u1 . where [rnea, closure of a set;

2) Let x,y E I be a pair of adjacent nodes of the lattice and .s.,„ be a segment connecting these node.s. Then if X. v E .U. then su e p. and if X. v e p then s n p = 0 .

The first condition means that the figure covers all black points of a discrete image and all white points he either outside of or on the boundary of the figure. The second condition can be reduced to the condition that the boundary of p lies in the interface between white and black boundary points of the discrete image.

Let M bc the set of all figures satisfying conditions 1 and 2. Any of them can be considered a continuous model of a binary image with acceptable accuracy. As we are going to build a skeleton of this figure, the most convenient representation for us is the figure with a piccewise linear boundary, since for such figures there are effective algorithms for construction of a skeleton. In this situation it is natuni to choose frum M a polygonal figure (PF) with minimal perimeter (see fig. 21, First, such PF exists and it is unique Second, the number of its vertices is cloce to minimal among all PF satisfying conditions 1 and 2.

weis
.1.:' ::.1
s •
PI
ke

Figure 2. Representation of raster object with cr#

polygonal figure with minimal perimeter

h

The. algorithm for solving this problem which over requires a single pass a raster image, has been described in [4].

6. Continuous skeletal representation of an image

The choice of the polygonal figure as a continuous rnodel of the binary image reduces the problem of construction of a skeleton of thc image to the ,ell-kno.-n medial axis transform [5]. Contrary to discrete imaees for which the skeleton is determine,' ambizuously. the concept of a skeleton of a continuous ficure has a strict mathematical formulation. The skeleton of a figure is the locus of points of centers of maximal emotv circles. An empty circle des not contain an y boundary points if the figure. Hie maximal empty circle is a circle which is not contained in am other empty circle. and which is not conqnient to another Note that empt} circles can be thus either intemal or external for the domains comprisine the figure. Accordinqly their centers form internal ;Ind external skeletons of the figure (see liq 3).

Figure 3. Empty circles for polygonal figure and skeleton of polygonal figure.

This definition applies to anv tvge of shape. not iust a polygon. However there exist effective algonthms for construction of polygonal figures [4,6]. The algonthm used [2,3] is based on a generalization of Delauney triangulation for a system of sites of two tvpes (points and segments) that comprise a PF boundary. It builds a skeleton in time 0(n log n) where n is the tumber of PF vertices.

Skeleton of polygonal figure can represented a, a planar graph, where nodes are points on a plane and bones are straight lines that connect the nodes. In such representation of a skeleton all nodes have no less than three nearest points on the border of the area and all bones lie between two linear fragments of the area border. Later in this article we will use only graph representation of a skeleton.

Let us also define a knot in skeleton as a node with more then two connected bones and final node as a node with only one connected node. And let us define a branch of skeleton as a consistent set of bones that has final node or knot node on each end and does not have knots in the middle of the branch. Later in this article we will operate only with branches of the skeleton and not with single b-ones.

Ausgabe 10
27. April 2013

Vom Mustergut zum Ausflugsziel: Das Landgut A. Borsig

Ein Ausflugstipp der besonderen Art und seit Kurzem auch Referenzstandort für die Bundesgartenschau (BUGA) 2015 ist das Landgut A. Borsig in Groß Behnitz bei Nauen. Einst als landwirtschaftliches Mustergut der Eisenbahner-Dynastie Borsig angelegt und als Tagungsort des Kreisauer Kreises genutzt, erwartet die Besucher heute in den historischen Gebäuden in direkter Lage am Großen Behnitzer See ein einzigartiges Freizeit-Angebot. Für Tagesausflügler bieten die intakte Natur mit Wald, Wasserflächen sowie Parklandschaft und zwei informativen Dauerausstellungen ein spannendes Ausflugsziel. Für die kulinarische Attraktion sorgt Küchenchef Dirk Bismark, der seinen Gästen täglich saisonale märkische Gerichte anbietet, die mit marktfrischen regionalen oder biologisch angebauten Produkten zubereitet werden. Das neue, moderne Biohotel und das historischen Logierhaus bieten mit insgesamt 105 Doppelzimmern und 23 Suiten eine komfortable Unterkunft auf Zeit und laden zu erholsamen Stunden inmitten der Brandenburger Natur ein. Auch für Hochzeiten oder Familienfeiern setzt das Landgut A. Borsig mit seinen großzügigen Freiflächen und zahlreichen Räumen unterschiedlicher Größen und Charaktere den perfekten Rahmen. Egal ob unter freiem Himmel direkt am Seeufer, in der Dorfkirche oder im eigenen Standesamt, ob modern oder historisch – den »Ja-Sagern« sind keine Grenzen gesetzt. Und wer nach einem geeigneten Ort für Tagungen, Seminaren oder Incentives sucht, ist beim Landgut A. Borsig ebenfalls an der richtigen Adresse: Sieben helle Salons, die mit der neuesten Technik ausgestattet sind, bieten bis zu 700 Personen viel Raum für eine erfolgreiche Tagung, Veranstaltung oder Firmenfeier. Privatgäste wie auch Tagungsbesucher können gemeinsam Kochen, Filzen oder sich in der hauseigenen Lithografiewerkstatt inspirieren lassen. Technik-Fans versuchen sich an der hauseigenen, meterhohen Dampfmaschine oder erfahren bei einer persönlichen Führung durch den Gutsherrn Michael Stober mehr über diesen historischen Ort – nur 45 Fahrtminuten von Berlin-Mitte entfernt. Eine öffentliche Führung findet jeden ersten Sonntag im Monat um 11 Uhr statt, für 26,50 Euro gibt es dann neben dem Rundgang auch noch ein Drei-Gänge-Menü. Und zur BUGA 2015 präsentiert das Landgut als Referenzstandort rechtzeitig weitere spannende Angebote rund um Nachhaltigkeit und Natur. Weitere Infos unter www.landgut-aborsig.de

Ausflugsziel für die ganze Familie: intakte Natur mit Wald, Wasserflächen sowie Parklandschaft.

Restaurant Seeterrassen: Küchenchef Dirk Bismark bietet seinen Gästen täglich saisonale märkische Gerichte an.

Gewinnen Sie eine Übernachtung in einer Suite für zwei!

Schlossstraße *live* und der neue BUGA-Referenzstandort Landgut A. Borsig verlosen eine Übernachtung (nach Reservierung) in einer der großzügigen Suiten im Biohotel des Landgut A. Borsig für zwei Personen inkl. Frühstück und Drei-Gänge-Menü im Restaurant Seeterrassen. **Einsendeschluss: 11. Mai 2013**. Der Rechtsweg ist ausgeschlossen. Schreiben Sie an: zimmermann@abendblatt-berlin.de oder senden Sie eine frankierte Postkarte an Schlossstraße *live*, Postfach 350625, 10215 Berlin.

Viel Glück!

Beispiel K

/Beispiel K

Vom Mustergut zum Ausflugsziel:
Das Landgut A. Borsig
Im Ausflugstipp der beson-
deren Art und seit Kurzem
auch Referenzstandort für
die Bundesgartenschau (BUGA)
2015 ist das Landgut A. Bor-
sig in Groß Behnitz bei Nauen.
Einst als landwirtschaftliches
Mustergut der Eisenbahner-
Dynastie Borsig angelegt und
als Tagungsort des Kreisauer
Kreises genutzt, erwartet the Be-
sucher heute in den historischen
Gebäuden in direkter Lage am
Großen Behnitzer See
ein einzigartiges
Freizeit-Angebot.
Für Tagesausflügler
bieten die intakte
Natur mit Wald,
Wasserflächen
sowie Parkland-
schaft und zwei
informativen Dau-
erausstellungen
ein spannendes
Ausflugsziel. Für die kulinarische
Attraktion sorgt Küchenchef
Dirk Bismark, der seinen Gästen
täglich saisonale märkische Ge-
richte anbietet, die mit marktfri-
schen regionalen oder biologisch
angebauten Produkten zuberei-
tet werden. Das neue moderne
Biohotel und das historischen
Logierhaus bieten mit insgesamt
105 Doppelzimmern und 23
Suiten eine komfortable Unter-
kunft auf Zeit und laden zu er-
holsamen Stunden inmitten der
Brandenburger Natur ein. Auch
für Hochzeiten oder Familienfei-
em setzt das Landgut A. Borsig
mit seinen großzügigen Freiflä-
chen und zahlreichen Räumen
unterschiedlicher Größen und
Charaktere den
perfekten Rahmen.
Egal ob unter frei-
em Himmel direkt
am Seeufer, in der
Dorfkirche oder im
eigenen Standes-
amt, ob modern
oder historisch –
den »Ja-Sage«
sind keine Grenzen
gesetzt. Und wer
nach einem geeigneten Ort für
Tagungen, Seminaren oder In-
centives sucht, ist beim Landgut
A. Borsig ebenfalls an der rich-
tigen Adresse: Sieben helle Sa-
lons, die mit der neuesten Tech-
nik ausgestattet sind, bieten bis

zu 700 Personen viel Raum für
eine erfolgreiche Tagung, Veran-
staltung oder Firmenfeier:Privat-
gäste wie auch Tagungsbesucher
können gemeinsam Kochen, Fil-
zen oder sich in der hauseigenen
Lithografiewerkstatt inspirieren
lassen. Technik-Fans versuchen
sich an der hauseigenen, me-
terhohen Dampfmaschine oder
erfahren bei einer persönlichen
Führung durch den Gutsherrn
Michael Stober mehr über die-
sen historischen Ort nur 45
Fahrtminuten von Berlin-Mitte
entfernt. Eine öffentliche
rung findet jeden ersten Sonntag
im Monat um 11 Uhr statt, für
Ausgabe to
27. April 2013
26,50 Euro gibt es dann neben
dem Rundgang auch noch ein
Drei-Gänge-Menü. Und zur BUGA
2015 präsentiert das Landgut
als Referenzstandort rechtzeitig
weitere spannende Angebote
rund um Nachhaltigkeit und
Natur. Weitere Infos unter www.
landgut-borsigxxte
Gewinnen
Sie eine
Übernachtung
in einer Suite
für zwei!
Schlossstraße live und der
neue BUGA-Referenzstandort
Landgut A. Borsig verlosen
eine Übernachtung nach Re-
servierung) in einer der gräb-
züigen Suiten im Biohotel
des Landgut A. Borsig für zwei
Personen inkl. Frühstück und
Drei-Gänge-Menü im Restau-
rant Seeterrassen. Einsende-
schluss: 11. Mai 2013. Der
Rechtsweg ist ausgeschlos-
sen. Schreiben Sie an:
zimmermann@abendblatt-
berlin.de oder senden Sie
eine frankierte Postkarte an
Schlossstraße live, Postfach
350625, 10215 Berlin.
Viel Glück!

Restaurant Seeterrassen: nrchenchef Dirk Bismark bietet seinen Gästen täglich saisonale märkische Gerichte an.
Ausflugsziel für die ganze Familie: intakte Natur mit Wald, Wasserflächen sowie Parklandschatt.

/Beispiel K

Vom Mustergut zum Ausflugsziel:
Das Landgut A. Borsig

Ein Ausflugstipp der besonderen Art und seit Kurzem auch Referenzstandort für die Bundesgartenschau (BUGA) 2015 ist das Landgut A. Borsig in Groß Behnte bei Nauen. Einst als landwirtschaftliches Mustergut der Eisenbahner-Dynastie Borsig angelegt und als Tagungsort des Kreisauer Kreises genutzt, erwartet die Besucher heute in den historischen Gebäuden in direkter Lage am Groß Behnitzer See ein einzigartiges Freizeit-Angebot.
Für Tagesausflügler bieten die intakte Natur mit Wald, Wasserflächen sowie Parklandschaft und zwei informativen Dauerausstellungen ein spannendes Ausflugsziel. Für die kulinarische Attraktion sorgt Küchenchef Dirk Bismark, der seinen Gästen täglich saisonale märkische Gerichte anbietet, die mit marktfrischen regionalen oder biologisch angebauten Produkten zubereitet werden. Das neue, moderne Biohotel und das historischen Logierhaus bieten mit insgesamt 105 Doppelzimmern und 23 Suiten eine komfortable Unterkunft auf Zeit und laden zu erholsamen Stunden inmitten der Brandenburger Natur ein. Auch für Hochzeiten oder Familienfeiern setzt das Landgut A. Borsig mit seinen großzügigen Freiflächen und zahlreichen Räumen unterschiedlicher Größen und Charaktere den perfekten Rahmen. Egal ob unter freiem Himmel direkt am Seeufer, in der Dorfkirche oder im eigenen Standesamt, ob modern oder historisch – den "Ja-Sagern" sind keine Grenzen gesetzt. Und wer nach einem geeigneten Ort für Tagungen, Seminaren oder Incentives sucht, ist beim Landgut A. Borsig ebenfalls an der richtigen Adresse: Sieben helle Salons, die mit der neuesten Technik ausgestattet sind, bieten bis zu 700 Personen viel Raum für eine erfolgreiche Tagung, Veranstaltung oder Firmenfeier. Privatgäste wie auch Tagungsbesucher können gemeinsam Kochen, Filzen oder sich in der hauseigenen Lithografiewerkstatt inspirieren lassen. Technik-Fans versuchen sich an der hauseigenen, meterhohen Dampfmaschine oder erfahren bei einer persönlichen Führung durch den Gutsherrn Michael Stober mehr über diesen historischen Ort nur 45 Fahrtminuten von Berlin-Mitte entfernt. Eine öffentliche Führung findet jeden ersten Sonntag im Monat um 11 Uhr statt, für

Ausgabe 10
27. April 2013

26,50 Euro gibt es dann neben dem Rundgang auch noch ein Drei-Gänge-Menü. Und zur BUGA 2015 präsentiert das Landgut als Referenzstandort rechtzeitig weitere spannende Angebote rund um Nachhaltigkeit und Natur. Weitere Infos unter www.landgut-aborsig.de
Gewinnen Sie eine Übernachtung in einer Suite für zwei!
Schlossstraße live und der neue BUGA-Referenzstandort Landgut A. Borsig verlosen eine Übernachtung (nach Reservierung) in einer der großzügigen Suiten im Biohotel des Landgut A. Borsig für zwei Personen inkl. Frühstück und Drei-Gänge-Menü im Restaurant Seeterrassen. Einsendeschluss: 11. Mai 2013. Der Rechtsweg ist ausgeschlossen. Schreiben Sie an: zimmermann@abendblatt-berlin.de oder senden Sie eine frankierte Postkarte an Schlossstraße live, Postfach 350625, 10215 Berlin.
Viel Glück!

Restaurant Seeterrassen: Küchenchef Dirk Bismark bietet seinen Gästen täglich saisonale märkische Gerichte an.
Ausflugsziel für die ganze Familie: intakte Natur mit Wald, Wasserflächen sowie Parklandschaft.

Aufnahmephase

Einfachheit mit einer ganzen Reihe von Schwachstellen verbunden:

- die Notwendigkeit einer frühzeitigen Binarisierung des Bildes
- Empfindlichkeit gegenüber der Zeilenorientierung und -krümmung
- Empfindlichkeit gegenüber der Schriftgröße

Kantenbasierte Methoden stellen eine Möglichkeit dar, die Binarisierungsproblematik zu umgehen. Neben den klassischen Kantenextraktionsoperatoren (s. Abschnitt 2.3) [64][65] werden Methoden der Grauwert-Morphologie [57] [58] für eine Transformation in den Kantenraum eingesetzt, in dem bestimmte textspezifische Klassifizierungsmerkmale einfacher extrahiert werden können. Ein weiterer Vorteil der Transformation ist die dabei stattfindende Hochpass-Filterung der Aufnahme, wodurch die negativen Effekte einer inhomogenen Beleuchtung gemildert werden.

Auch für die Orientierungsproblematik wurden bereits zahlreiche Lösungsansätze vorgeschlagen. In [62] wird eine Weiterentwicklung des RLS-Algorithmus vorgestellt, wobei statt eines achsensymmetrischen strukturierenden Elements eine Familie von unterschiedlich ausgerichteten Filtern für die Schließen-Operation verwendet werden, um den Filter zu bestimmen, der mit der Zeilenorientierung im Dokument korrespondiert. Eine ähnliche Überlegung steht hinter den *Projektionsprofilen* – einer häufig eingesetzten Methode [66][67] zur Feststellung der Textflussrichtung durch die Akkumulierung der Pixelwerte entlang einer Kurve. Mit Hilfe der auf diese Weise produzierten Profilhistogramme lässt sich u.U. nicht nur die Zeilenorientierung, sondern auch die Schriftgröße eines Textblocks grob abschätzen. Anhand der globalen Musteränderungen in den Profilen kann man darüber hinaus die Grenzen eines Textblocks (engl. *recursive XY-Cuts* [68]) ablesen oder Aussagen über den Inhalt [69] machen. Zu den großen Schwächen der Projektionsprofile zählt ihre Anfälligkeit (s. Abb. 5.1.2) gegenüber Dokumentverzerrungen [66].

Die Skalierungsproblematik stellt eine große Herausforderung für die Textlokalisierung dar. Eine zuverlässige Merkmalsextraktion ist auf die Kenntnis der erwarteten Schriftgröße angewiesen, während man für die Schätzung der Schriftgröße aussagekräftige textspezifische Merkmale benötigt. Viele der schnellen Methoden [57][58][59][61] werden daher durch die Wahl geeigneter Parameter auf einen bestimmten anwendungsabhängigen Schriftgrößenbereich zugeschnitten, wohin-

60

Beispiel L

/Beispiel L

```
Aillizahmephase
Einfachheit mit einer ganzen Reihe von Schwachstellen verbunden:
• die Notwendigkeit einer fruhzeitigen Binarisierung des Bildes
• Empfindlichkeit qegenilber der Zeilenorientierung und -krummune
• Empfindlichkeit gegeniiber der SchriftgyaBe
Kantenbasierte Methoden stellen eine Moglichkeit dar, die Binarisierungsproble-
matik zu umgehen. Neben den klassischen Kantenextraktionsoperatoren (s. Ab-
schnitt 2.3) [64][65] werden Methoden der Grauwert-Morphologie [57] [58] tilr
eine Transformation in den Kantenraum eingesetzt, in dem bestimmte textspezifi-
sche Klassifizierungsmerkmale einfacher extrahiert werden können. Ein meiterer
Vorteil der Transformation ist die dabei stattfindende Hochpass-Filterung der
Aufnahme, wodurch die negativen Effekte einer inhomogenen Beleuchtung ge-
mildert werden.
Auch fi1r die Orientierungsproblematik wurden bereits zahlreiche LOsung.sansatze
vorgeschlagen. In [62] wird eine Weiterentwicklung des RLS-Algorithmus vorge-
stellt, wobei statt eines achsensymmetrischen strukturierenden Elements eine Fa-
milie von unterschiedlich ausgerichteten Filtern ftir die SchlieBen-Operation 'er-
wendet werden, um den Filter zu bestimmen, der mit der Zeilenorientierung im
Dokument korrespondiert. Eine ahnliche Oberlegung steht hinter den Prqielui-
onsprofilen – einer häufig eingesetzten Methode [661[67] zur Feststellung der
Textflussrichtung durch die Akkumulierung der Pixelwerte entlang einer Kurve.
Mit Hilfe der auf diese Weise produzierten Profilhistogramme lasst sich u.U.
nicht nur die Zeilenorientierung, sondern auch die SchriftgroBe eines Textblo'cr
grob abschätzen. Anhand der globalen Musteränderungen in den Profilen kann
man darilber hinaus die Grenzen eines Textblocks (engl. recursive XI-Cuts MI)
ablesen oder Aussagen uber den Inhalt [69] machen. Zu den groBen Schwachen
der Projektionsprofile zahlt ihre Anfalligkeit (s. Abb. 5.1.2) qegentiber Doku-
mentverzerrungen [66].
Die Skalierungsprobiematik stellt eine groBe Herausforderung für die Textlokali-
sierung dar. Eine zuverlässige Merkmalsextraktion ist auf die Kenntnis der erwar-
teten Schriftgraße angewiesen, wahrend man fur die Schatzung der Schritigroße
aussagekraftige textspezifische Merkmale benatigt. Viele der schnellen Methoden
[57][58][59][61] werden daher durch die Wahl geeigneter Parameter auf einen
bestimmten anwendungsabhangigen Schriftgral3enbereich zugeschnitten, wohin-
60
Aulhalunephas.e
Einfachheit mit einer ganzen Reihe von Schwachstellen verbunden:
• die Notwendigkeit einer friThzeitigen Binarisierung, des Bildes
• Empfindlichkeit gegenuber der Zeilenorientierung und -krummung
• Empfindlichkeit gegentiber der SchriftgroBe
Kantenbasierte Methoden stellen eine Moglichkeit dar, die Binarisicrungsproble-
matik zu umgehen. Neben den klassischen Kantenextraktionsoperatoren (s. Ab-
schnitt 2.3) [64][651 werden Methoden der Grauwert-Morpholoizie [57} {58} tilr
eine Transformation in den Kantenraum eingesetzt, in dem bestimmte textspeziti-
sche Klassifizierungsmerkmale einfacher extrahiert werden können. Ein weiterer
Vorteil der Transformation ist die dabei stattfindende Hochpass-Filterung der
Aufnahme, wodurch die negativen Effekte einer inhomogenen Beleuchtung ge-
mildert werden.
Auch fur die Orientierungsproblematik wurden bereits zahlreiche Lösungsansatie
vorgeschlagen. In [62] wird eine Weiterentwicklung, des RLS-Algorithmus vorge-
stellt, wobei statt eines achsensymmetrischen strukturierenden Elements eine Fa
milie von unterschiedlich ausgerichteten Filtern für die Sch1ief3en-Operation ver-
wendet werden, um den Filter zu bestimmen, der mit der Zeilenorienticrung im
Dokument korrespondiert. Eine ahnliche Uberlegung steht hinter den Projekti-
onsprofilen – einer häufig eingesetzten Methode [66][67] zur Feststellung der
Textflussrichtung durch die Akkumulierung der Pixelwerte entlang einer Kurvet.
Mit Hilfe der auf diese Weise produzierten Profilhistogramrne lasst sich u.U.
nicht nur die Zeilenorientierung, sondern auch die Schriftgrof3e eines Textblocks
grob abschätzen. Anhand der globalen Musteranderungen in den Profilen kann
man dariiber hinaus die Grenzen eines Textblocks (engl. recursive XY-cuts [68])
ablesen oder Aussagen über den Inhalt [69] machen. Zu den grof3en Schwächen
der Projektionsprofile zahlt ihre Anfalligkeit (s. Abb. 5.1.2) gegennüber Doku-
mentverzerrungen [66].
Die Skalierungsproblematik stellt eine große Herausforderung für die Textlokali-
sierung dar. Eine zuverlassige Merkmalsextraktion ist auf die Kenntnis der erwar
teten Schriftgrof3e angewiesen, wahrend man für die Schatzung der Schriftgroße
aussagekräftige textspezifische Merkmale benotigt. Viele der schnellen
[57][5811591161] werden daher durch die Wahl geeigneter Parameter auf einen
bestimmten anwendungsabhangigen Schriftgrof3enbereich zugeschnitten.
60
```