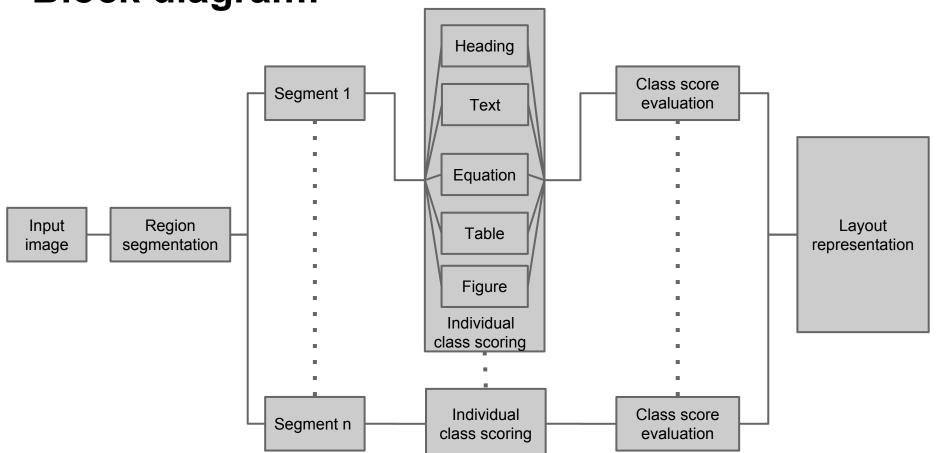
# **Document layout analysis**

Mallikarjun B R

# **Problem statement:**

- Aim of the project is to recognize the layout of technical documents.
- Layout is defined on the basis of location and class of segments within the document.
- Region segmentation is done using morphological processes.
- Classes considered are Headings, Text, Tables and Figures.
  - Headings are scored on the basis of location.
  - Text is scored on the basis of statistical analysis of connected components.
  - Tables are scored on the basis of number of parallel lines.
  - After thresholding, rest are considered as figures.
- Based on scores, classes are assigned.

**Block diagram:** 



# Segmentation

- Identification of columns using vertical SE.
- Identification of text lines and calculation of average space between lines.
- Division into sub-regions using SE of height equal to calculated line spacing.
- Identifying multi-column titles and figures.
- Improved segmentation of text blocks based on empty space at the end of paragraph.

## Bayesian Depth-from-Defocus with Shading Constraints

Then I in Streacher Sub Averyok, Mersishuad Kind Zisan Stephen I in San werterin (A. 101 Mersishuad), Mersish weard doe in Subject Property

#### Abstract

We present a remain that consistent the projection of story from foliations (PEPD) in much discuss and channel of placement. The object from the majoritant discussions is entirely in our proposed properties and properties of the object from the consistent of the object channel of the object of the object channel of the object of t

#### Introduction

happer mandefinies in Provide and dysocal techniques the influence in a contrept recover dup a, terms can be a fine influence in a contrept recover dup a, terms can be officially a provide and the finite and the contract of influence by the provide and the contract to that the diffusion is to see that the contract to adversariage behavior, or set to a Experiment (EED data of implications a function with a series with or level in this are consistent with a series with or level in this are consistent with a series with the first times, on the matter with a recovery when the contract times, on the matter with a recovery when the contract times, on the matter with a recovery when the contract times, on the matter with a contract when the contract times, to the matter with a contract when the contract times, to the matter with a contract time.

So has compensately a large some canes he sign resolution insight, 1912 any marker let appropriate resolution insight, 1912 any marker let appropriate technicistic sight of it. Informal classes. However, the continuing sight of the formal classes of the sight continuing sight of the formal classes of the sight and the sight of the sight of the sight of the sight lets, which can be seen clearly contain it in Million in the algebrasities as conference on the sight marker of the sight sight of the sight of the

We present in this perior a feelingue for it is to miigate the appropriate and construction. Description to of studies information, in contrast to redeem that a full in the influence the angel date of a surface.

but the recycle up a matery drips summore test filtred and a fine shall in (SBR), indirect terms a control for an old filtred in the second point of the filtred in the filtred in the second [1, 3, 3]. We determent used to appear so in Arrang cere to material context the control field in the second context and of the filtred in the second many process of a finantial field point the control many questions and shading in our manifest of the first process of the first and the control many questions and the first and the first and the control many first and first first in the control many the first and the first the control many first the first the first the first the first the first first the first the first first first the first fi

In others was published, so for jumpour houses membrane of BII that members strating contributes in man of that lovely and hashed visuality, axes in near that lovely and hashed visuality, axes in the so-of balling attractions in trained deaths, the frequent DII is removed to the training grade in, this mage accomposition that in a to space as both in form other than two frameworks, as before agreed in, this mage accomposition that in a to space as both in form other than two finest contributions that is a topic and the solution of the straining form of the most accomposition of the solution of the s

In the work, the true of a consideration of a table for many cell function extrement a regarded grings for a glabe with a known of the raw. The couper has become the first the performance of Reasons (24) which shall generate suppose that or matting 10.0 feet and they work of the couper for the states of the couper for t

#### Referred Work

Bagale foundations. The early to colors inflations to form a more III by primary good works that tende-force whose high tende-force whose high two colors or early step [1] is a [4]. Some first the early colors of more adjustant to the earlier form the edge. For the earlier the earlier form the edge, for the earlier two steps of the earlier form the edge of the earlier two more than earlier earlier to the earlier form the edge promitioned by a plane greatle, it are entire [13]. So, [20] and "in the edge plane interface and the dissipance," If it is seen to the edge plane in the edge plane in the edge plane in the edge plane.

1

# **Seamentation**

where  $P(\mathbf{d})$  is the prior distribution of depth map  $\mathbf{d}$ ,  $P(\mathbf{l}^{(1)}, \mathbf{l}^{(2)} | \mathbf{d})$  is the likelihood of observations  $\mathbf{l}^{(1)}, \mathbf{l}^{(2)}$ , and L is the log-likelihood of P, i.e.  $L = -\log P$ .

The likelihoost form can be profeted as the basic DPDI energy from Eq. (8), and the prior term as depth smoothness almost the links. "It

$$L(\Gamma^{(1)}, \Gamma^{(2)}, \mathbf{d}) = \sum_{d \in \mathbf{d}} (I^{(1)} \circ b(d, d) - T_i^{(2)})^d,$$
 (10)

$$L(\mathbf{d}) = \lambda \sum_{(G/D)} \langle J_1 - J_2 \rangle^2. \tag{11}$$

Hereafter, this particular formulation will be referred to as standard OFO.

To optimize the MRF model of Eqs. (10)-(11), we use the max product ground of the typical propagation algofithm [27], with a message update refriduct has propagates ressages in one direction and updates each male minight

## 3.3. Shading-based prior term

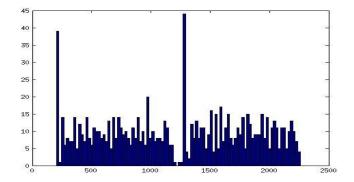
The smoothness piet of Eq. (11) can reduce note in the reconstructed depth, bit there is no provide any actitionals. Ensured the scene, we propose to take a refore infiliation of the scene of the standard of the propose to take a refore infiliation of the standard of the standard of the piet of the piet in the piet i

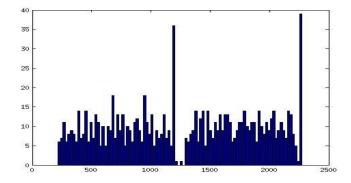
where n' = [w<sub>2</sub>, w<sub>3</sub>, w<sub>4</sub>], for surface normal is, and M is a synthetic ! I hand so that highly entropy a synthetic ! I hand so that highly entropy and the surface is a synthetic ! I hand so that highly entropy and the surface is a surface in the part in the grant part in the surface is control in each pare! I many the realizability [We absolute the Obtain the 4D coordinates for each point by re-projecting each part into the speak againfung to the mage construction [13, 3], depth value 2 from DPD, and the perspective projection manual in the speak of the perspective projection manual in the speak of the perspective projection manual in the properties of the perspective projection manual in the perspective projection manual in the perspective projection in the perspective projection manual in the perspective projection manual in the perspective projection in the perspective projection manual in the perspective projection manual in the perspective projection in the perspective projection manual in the perspective projection in the perspect

$$\left(\left(x-\frac{\omega}{2}\right)int,\left(y-\frac{h}{2}\right)int,\sigma\right)$$
.

where we was the resolution of the image, and was the pixell died.

For each pair of Buked pixels x,y in the MRP, we now have then depths  $a_{x,y}, y_{y}$ , xD positions  $p_{y}$ ,  $p_{y}$ , and normals  $y_{x,y}, y_{y}$ . Since the vector  $\vec{p}_{x}\vec{p}_{y}^{2}$  should be perpential far to





but also reveals high frequency shape variations that allow shape-from-shading (SFS) methods to match or exceed the level of detail obtainable by active sensing [10, 32]. We therefore seek to capitalize on shading data to refine and correct the coarse depth maps obtained from DFD. The utilization of shading in conjunction with DFD, however, passes a significant challenge in that the scene texture generally needed for DFD interferes with the operation of shape-from-shading, which requires surfaces to be free of albedo variations. Moreover, DFD and SFS may produce mean gruous depth estimates that need to be reconciled.

To address these problems, we first propose a Bayesian formulation of DFD that incorporates shading constraints in a manner that locally emphasizes shading cues in areas where there are ambiguities in DFD. To enable the use of shading constraints in textured scenes, this Bayesian DFD is combined in an iterative framework with a depth guided intrinsic image decomposition that aims to separate shading from texture. These two components mutually benefit each other in the iterative framework, as better depth estimates lead to improvements in depth-guided decomposition, while more accurate shading/texture decomposition amends the shading constraints and thus results in better depth estimates.

In this work, the object surface is assumed to be Lambertian, and the illumination environment is captured by imaging a sphere with a known reflectance. Our experiments demonstrate that the performance of Bayesian DFD with shading constraints surpasses that of existing DFD techniques over both course and tine scales. In addition, the use of shading information allows our Bayesian DFD to work effectively even for the case of untextured surfaces.

# **Table Feature Identification**

Tables can be identified using the following properties:

- Number of parallel lines
- Angles of lines (only horizontal or vertical)
- Start and end of parallel lines are aligned

## Work Done:

- Extracted edges using Canny Edge Detection
- Removed text edges using pre-processing (to avoid noise in line detection)
- Identified lines and their angles using Hough Transform
- Accept only if lines are horizontal or vertical

# Text & Heading Feature Identification

Text can be identified using average and variance of area of connected components.

## Work Done:

- Pre-processing to remove noise and improve connectivity within letters.
- Calculation of area of each of the connected components.
- Experimentation has been done based on the mean and variance of area of connected components.
- Headings and Subheadings are identified based on connected component area and position.
- Equations are identified based on statistical data of connected component area.

# **Results:**

Class	Label
Heading	1
Authors	2
Text	3
Sub Heading	4
Equations	5
Tables	6
Figures	7

## Bayesian Depth-from-Defocus with Shading Constraints

2						
Chen Li	Shuochen Su <sup>5</sup>	Yasuyuki M	fatsushita <sup>2</sup>	Kun Zh	ou'	Stephen Lin*
1 Sta	te Key Lab of CAD&CG, Zi	ejiang University	*Microsoft Rese	arel: Asia	Tsinghu	a Linwersity

Abstract

We present a natheal that enhances the performance of depth-from-deficies (DED) through the use of shading algorithms. DED suffers from important locations. DED suffers from important locations, another course slape reconstruction and pose accuracy on lexitaries is sufficient on the creep can with the help of shading. We integrate both forms of data within a Bayetian featurework that capitalities on their default within a Bayetian featurework that capitalities on their data within a Bayetian featurework that capitalities on the recover accurately from surfaces that romain texture. To address this mater, we propose an iterative technique that within shading information in the presence of textures. With this approach, we demonstrate improvements over existing 1901 techniques, as well as effective shape reconstruction of recurrents as the state least the state of textures.

## 1. Introduction

Depth-front-detocus (DPD) is a wisday-used technique that utilizes the relationship between depth, focus setting, and image biar to passively estimate a range map. A part of images is typically acquired with different focus settings, and the differences between their focul burn levels are thou ossel to infer the depth of each soone point. In common motive sensing techniques such as 1D scarming, DHD does not require direct interaction with the scene. Additionally, it offers the convenience of temploying a single stationary cantern, milks methods based on stereo vision.

With the rising popularity of large furnat leaves for high resolution imaging, DFD may increase in application due to fix shallow depth of field of such leaves. However, there exist imaging and some flictors that firmt the estimation accuracy of DFD. Among these is the limited size of lens apertures, which leads to coasse depth resolution. In addition to this, depth institutes can be severely dignorded in areas with insefficient sexue texture for measuring lead blor levels.

We present in this paper a technique that thins to imiigate the aforencemend drawbacks of DFD through the use of shading information. In contrest to defects blur, shading not only indicates the general shape of a sortice, But also reveals high-frequency shape variations that allow hape-inam-shading (SPS) inathods to match or exceed the level of detail obtainable by series cassing [10, 32]. We fluerefore such to capitalize on shading that to refine and airrest the course depth maps obtained from DFD. The utilization of shading in conjunction with DFD, however, possas a spirificant challenge in fair the secon texture generally needed for DFD inderfers with the quention of shape from-shading, which requires surfaces to be free of albeds varietions. Moreover, DFD and SFS may produce incongrues stepthes estimates that meed to be recorded.

To address these problems, we first propose a flayesture formulation of DFD that incorporates shading constraints in a manner that locally emphasizes shading constraints where there are ambiguities in DFD. To enable the use of shading constraints in textured scenes, this Reyesain DFD is combined in an iterative framework with a depth-guided intimise image decumpusation that aims to separate shading from texture. These two components methally benefit each other in the iterative framework, as better depth-scinares leaf or improvements in de sth-guided decomposition, while more accurate stacting/texture decomposition mends the shading constraints and thus results in better glopit scirimates.

In this work, the object surface is assumed to be I ambarna, and the illumination environment is captured by imaging a sphere with a known reflectance. Our experiments formostrate that the performance of Bayesian DFD with slading constitutes surpasses that of existing DFD ochtiques over both course and line scales. In addition, the use of shading information allows our Bayesian DFD to work effectively even for the case of untextured surfaces.

## 2. Related Work

Depth-from-defocus. There exists a substantial amount of literature on DFD, beginning with works that handle objects whose brightness consists of step edges [18, 25, 9]. Since his in-focus intensity profile of these edges is known, their lepth can be determined from the edge but. Letter methods use instead essured that object surfaces can be locally approximated by a place parallel to the sensor [33, 26, 30], but that local depth variations can be disregulated in the

Simulation. Some techniques utilize structured aluminatum utilized with testureless strafaces and improve but estimation [15, 14, 29]. DFD has been formulated as a Markov analom field (MRF) problems, which allows inclusion in constraints among neighboring points [21, 22, 20]. Define has also been modeled as a diffusion process that does for require recovery of the in-focus image when estimating legals [6].

Shape-from-shuding Considerable work has also been Jone on shape-from-shading. We refer the reader to the SFS surveys in [34, 5], and review only the most releant methods here. SFS has traditionally been applied un der restrictive settings (e.g., Lambertian surfaces, uniform Ibedo, directional lighting, orthographic projection), and everal works have aimed to broaden its applicability, such s to acciness perspective projection [19], non-Lambertian reflectance [16], and natural illumination [10, 8]. Nonuniform allocin has been particularly challenging to oversome, and has been approached using smoothness and enrropy priors on reflectance 131. Our work instead takes adrantage of defocus information to improve estimation for extured surfaces. Shape from shading has also been used o refine the depth data of uniform-albedo objects obtained by multi-view stereo [32]. In our method, SFS is used in the context of DFD with scenes containing albedo variations.

Intrinsic images - Infinisic image decomposition aims to separate an image into its reflectance and shading compunents. This is an ill-posed problem, since there are twice s many unknowns (reflectance, shading) as observations image intensities) per pixel. The various approaches that have been employed make this problem tractable through the inclusion of additional constraints, such as those derived from Retinex theory [11], trained classifiers [28], and multiple images under different lighting conditions [31]. Despite the existence of these different decomposition cues. he performance of intrinsic image absorithms has in general been rather limited [7]. Recently, range data has been exploited to provide strong constraints for decomposition. and this has led to state of the art results [12]. Insuired by his work, we also utilize depth information to aid intrinsic mage decomposition. However, our setting is considerably nore challenging, since the depth information we start with every rough, due to the coarse depth estimates of DFD and the problems of SFS when textures are present.

#### 4 B. Approach

In this section, we present our method for Bayesian DFD with shading constraints. We begin with a review of hostic DFD principles, followed by a description of our Bayesian DFD model, our shading-based prior term, the method for

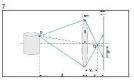


Figure 1. Imaging model used in depth-from-defocus.

fundling surface textures, and finally the iterative algorithm that integrates all of these commonents.

### 3.1. Basic principles of DFD

DIPD utilizes a pair of images taken with different focus estings. The effects of these focus estings on defocus blur still be described in terms of the quantities shown in Fig. 1, Let a seconder a secret point P located at a d-stance of from the camera lears. The light may candidated from P to the camrar are focused by the lens to a point Q according to the thin how mental terms.

$$\frac{1}{l} + \frac{1}{r_a} = \frac{1}{F},$$
(1)

Even  $v_{\rm ext}$  is the same of Q from the tens, and P is the local length. When the focus welling  $v_{\rm ext}$  which reposition the distance between the lens and sensor plane, is equal to  $v_{\rm ext}$  for expect  $p_{\rm ext}$  and  $p_{\rm ext}$  is equal to  $v_{\rm ext}$  for expect  $p_{\rm ext}$  in single boil to the length of the closest point  $v_{\rm ext}$  in the closest point  $p_{\rm ext}$  is the property blurce? Because in high  $v_{\rm ext}$  is distributed to different points in the surrow. Because of the intailizant symmetry of lenses, this blur is in the form of a circle. The radius  $v_{\rm ext}$  is the close to  $v_{\rm ext}$  is the contribution of  $v_{\rm ext}$  in the form of  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the form of  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle and  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  in the circle  $v_{\rm ext}$  is the circle  $v_{\rm ext}$  in the circle

$$b = \frac{R_{c}}{2} \begin{bmatrix} 1 & 1 & 1 \\ F & v & d \end{bmatrix}$$
, (2)

where It is the radius of tens. As seen from this equation, there is a direct relationship between depth of and blur radius b for a given set of comers parameters.

The light intensity of P within the blur circle can be exproved as a distribution function known as the point spread function (PSF), which we denote by h. In this paper, we proved the PSF h using a 2D Gaussian function [18]:

$$h(x, y; \sigma) = \frac{1}{2\pi \sigma^2} e^{-\frac{x^2-v^2}{2\sigma^2}}$$
(3)

with standard deviation  $\sigma = \gamma h$  where the constant  $\gamma$  can be determined by calibration [9]. Using the PSF, we can

3

# **Conclusion on results:**

- Approximately 80% accuracy in segmentation and classification.
- Errors include,
  - Equations get segmented as text because of similar line spacings.
  - Incorrect paragraph segmentation due to inconsistent line spacing. Solved in most of the cases.
  - Incorrect subheadings classification as text because of similar statistics.