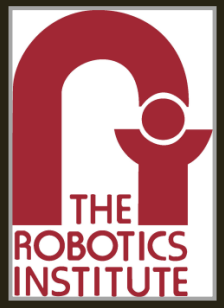
Rectification and 3D Reconstruction of Curved Document Images

Yuandong Tian Srinivasa Narasimhan



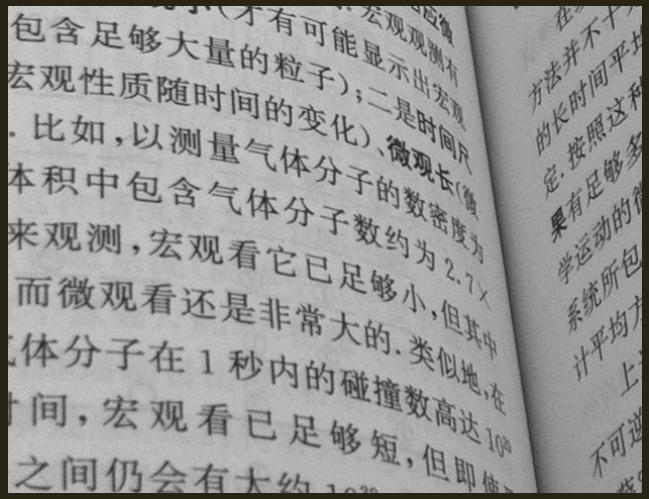
Carnegie Mellon University

Sponsors: NSF, ONR

Digitizing Documents Everywhere

Digitizing books

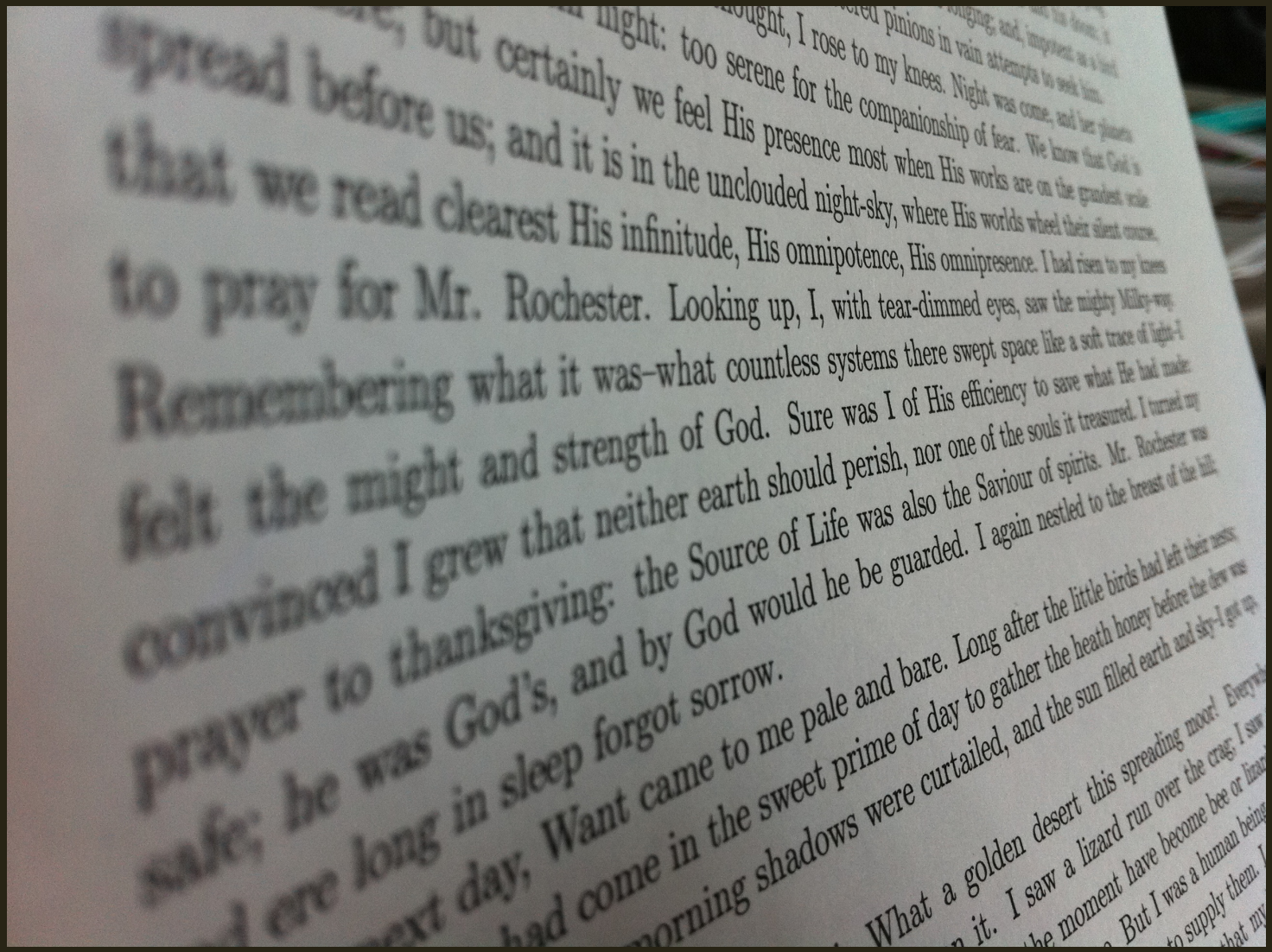
Rare manuscripts



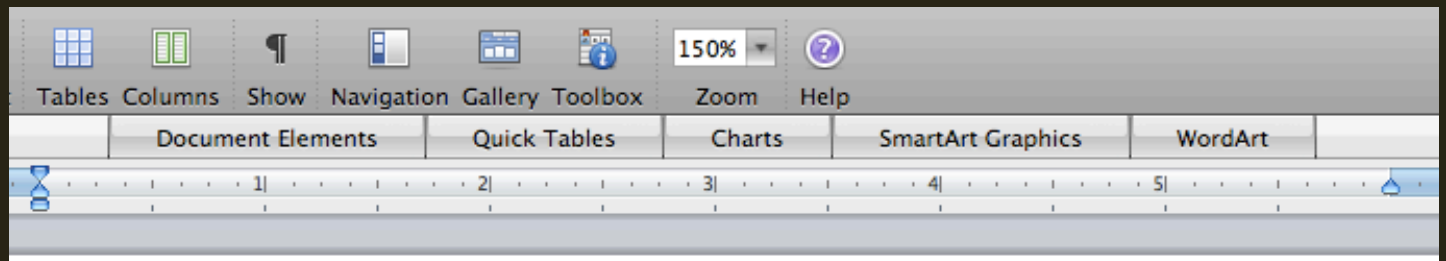
Foreign Language Translation

Bills and Receipts

A distorted image from cell-phone camera



OCR Performance under distortion



Ihe 'feel His preseixe most vltil Is In'ts lit Mt illityg t: _.4u %tit) I to;t:O'llz1;111\z'%1411\nit' k 0 Setelie fot tile comi)aAlots14 ikill It ' in the unclouded iittrk, '18AimIlls HiS infinitude, His oniipotelice, Is . '4411tiet:) orawott Vtiitostwasurt T. Looking up, I, w•lth, tem-IrciavA MS, SA it 4 /.11 at it was-what countless systems kloe 51 S\WAitt , It of God. Sure vas 1. () Els 6c,le'xi t ot itia ' the earth siloa pensiri, tOt Olt () SaIV't 1. la toibtoeyt sGoia:O e ve:0,\.iiee 1 0:e $::: t ::toe:t. . \ 16 ek I. Slo ac ;D:N v _le to It* Vt'\,ce:0 CP lesgee e c\° (), 'i wz '40 ..2146 ...'o o *110' 6, \"s \"" , ,0 ,,, ,,,, fi; ,§ZZC:%'CO\Q\ '\ V\ °) Z'S° 4 v- *i ti ill" 4,11:11At „Aro irer . i $izo, 4\\ to ,\\. \' A \k"' \.,■ttp,e1 0*
ktt

Imaging Documents using flat-bed scanner

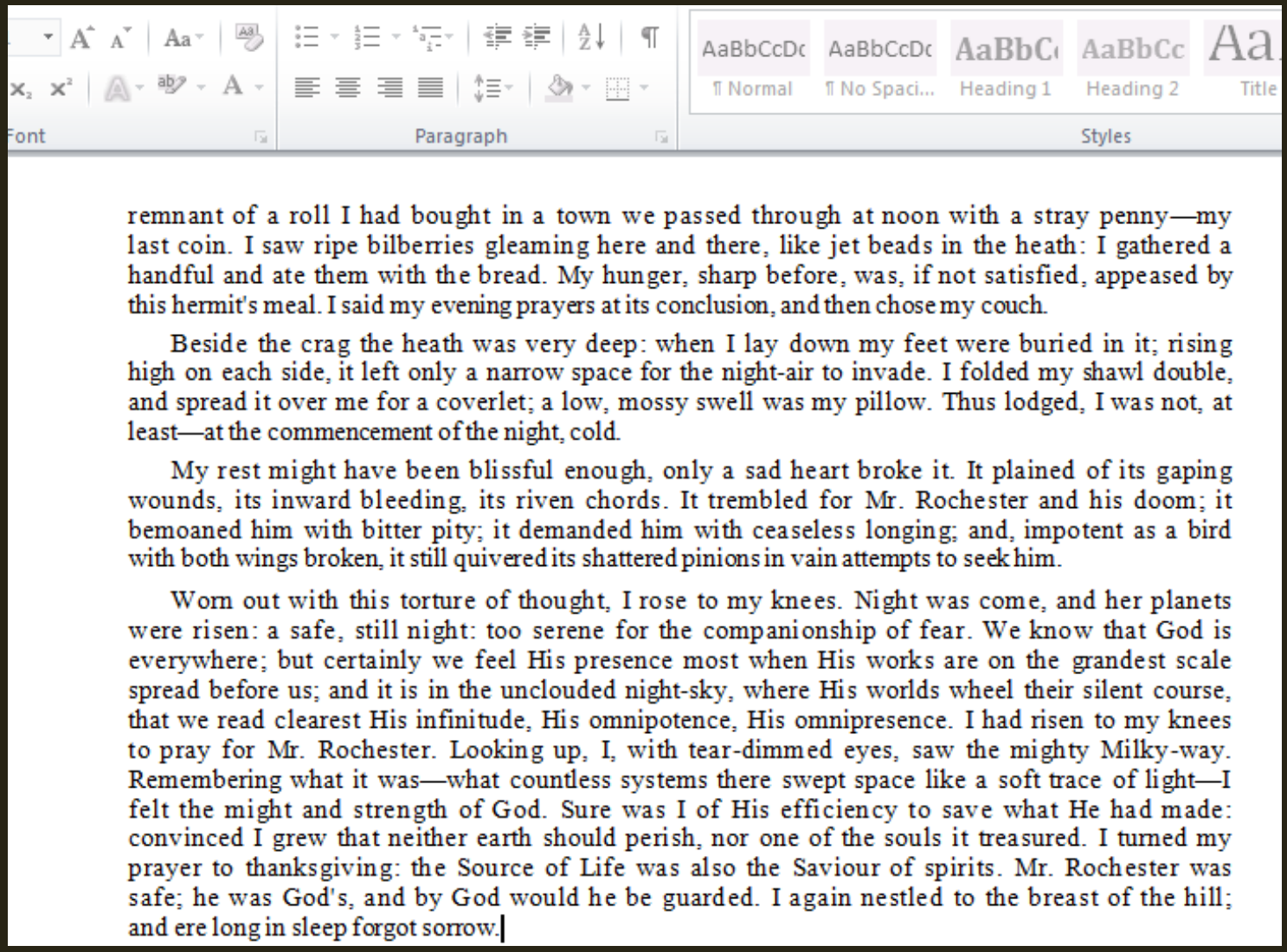
remnant of a roll I had bought in a town we passed through at noon with a stray penny–my last coin. I saw ripe bilberries gleaming here and there, like jet beads in the heath: I gathered a handful and ate them with the bread. My hunger, sharp before, was, if not satisfied, appeased by this hermit's meal. I said my evening prayers at its conclusion, and then chose my couch.

Beside the crag the heath was very deep: when I lay down my feet were buried in it; rising high on each side, it left only a narrow space for the night-air to invade. I folded my shawl double, and spread it over me for a coverlet; a low, mossy swell was my pillow. Thus lodged, I was not, at least–at the commencement of the night, cold.

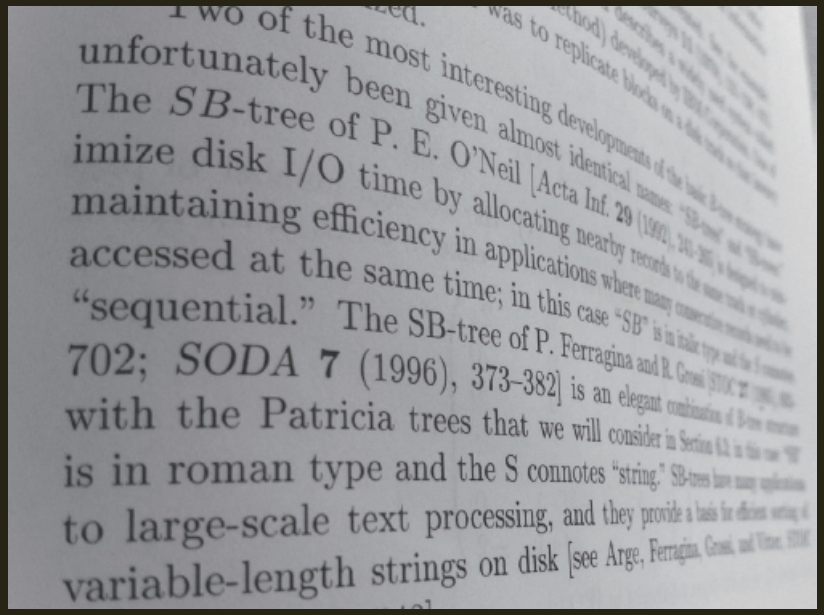
My rest might have been blissful enough, only a sad heart broke it. It plained of its gaping wounds, its inward bleeding, its riven chords. It trembled for Mr. Rochester and his doom; it bemoaned him with bitter pity; it demanded him with ceaseless longing; and, impotent as a bird with both wings broken, it still quivered its shattered pinions in vain attempts to seek him.

Worn out with this torture of thought, I rose to my knees. Night was come, and her planets were risen: a safe, still night: too serene for the companionship of fear. We know that God is everywhere; but certainly we feel His presence most when His works are on the grandest scale spread before us; and it is in the unclouded night-sky, where His worlds wheel their silent course, that we read clearest His infinitude, His omnipotence, His omnipresence. I had risen to my knees to pray for Mr. Rochester. Looking up, I, with tear-dimmed eyes, saw the mighty Milky-way. Remembering what it was–what countless systems there swept space like a soft trace of light–I felt the might and strength of God. Sure was I of His efficiency to save what He had made: convinced I grew that neither earth should perish, nor one of the souls it treasured. I turned my prayer to thanksgiving: the Source of Life was also the Saviour of spirits. Mr. Rochester was safe; he was God's, and by God would he be guarded. I again nestled to the breast of the hill; and ere long in sleep forgot sorrow.

OCR Performance on scanned image

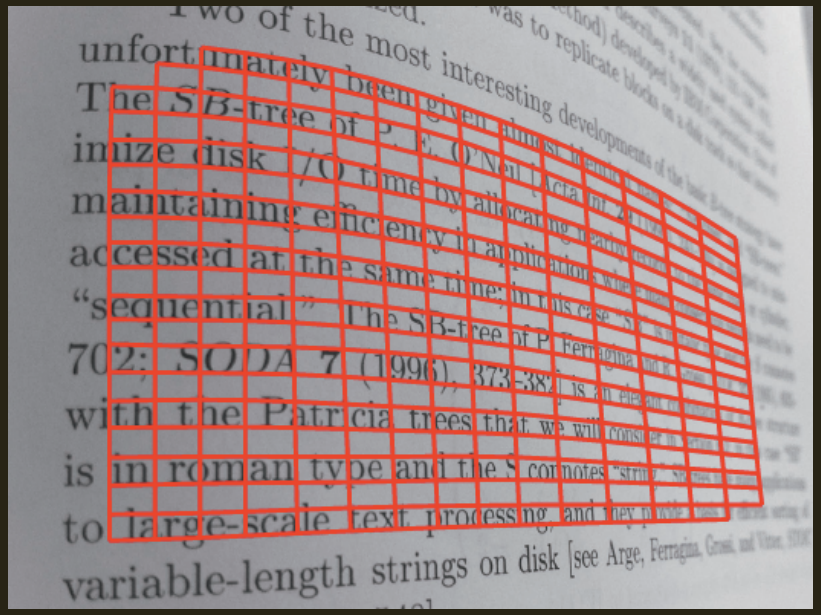


Contribution I: Text line analysis

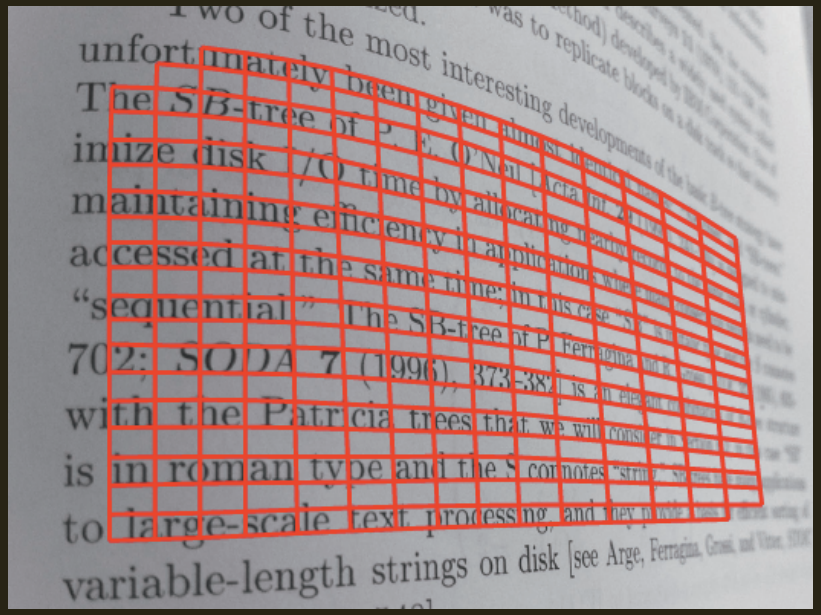


Two of the most interesting developments of the basic B-tree structure have unfortunately been given almost identical names: "SB-tree" and "SB-tree". The *SB*-tree of P. E. O'Neil [Acta Inf. **29** (1992), 241–265] is designed to minimize disk I/O time by allocating nearby records to the same track or cylinder, maintaining efficiency in applications where many consecutive records need to be accessed at the same time; in this case "*SB*" is in italic type and the S connotes "sequential." The SB-tree of P. Ferragina and R. Grossi [STOC **27** (1995), 693–702; *SODA* **7** (1996), 373–382] is an elegant combination of B-tree structure with the Patricia trees that we will consider in Section 6.2; in this case "SB" is in roman type and the S connotes "string." SB-trees have many applications to large-scale text processing, and they provide a basis for efficient sorting of variable-length strings on disk [see Arge, Ferragina, Grossi, and Vitter, STOC

Contribution I: Text line analysis



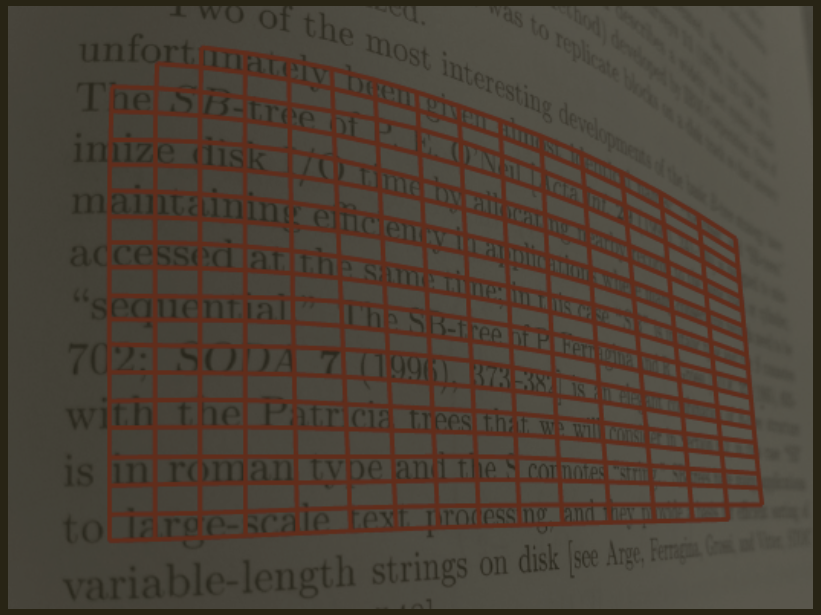
Contribution I: Text line analysis



Our contribution:

1. Works for multiple languages, font types and sizes
2. No global constraints on the shape of text lines.

Contribution I: Text line analysis



Our contribution:

1. Works on multiple languages, font types and sizes
2. No global constraints on the shape of text lines.

Previous works:

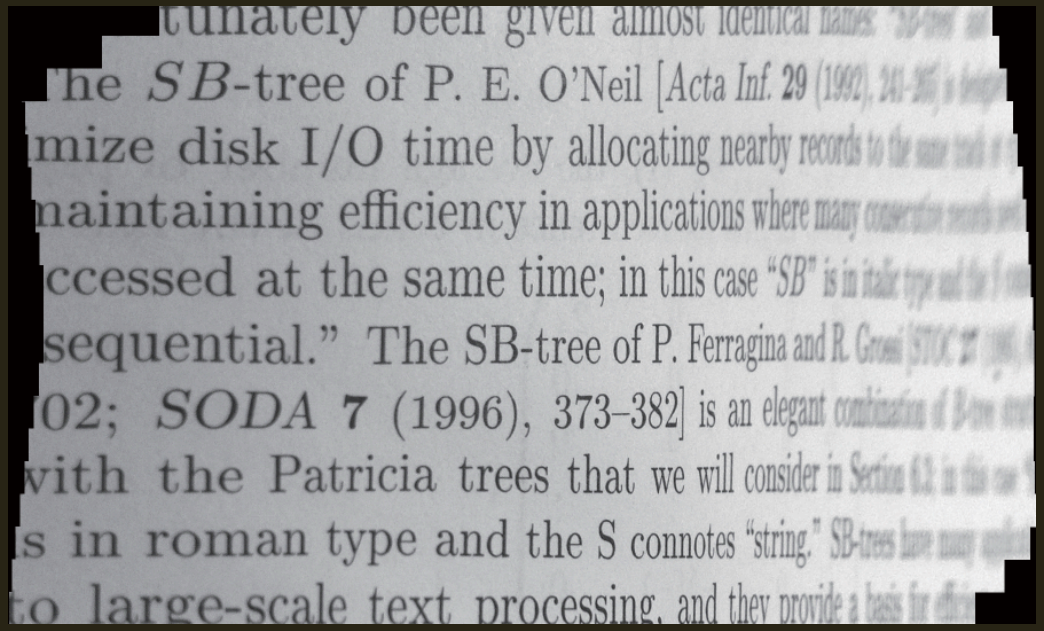
Connected Components/ Segmentation

[Zhang and Tan, ICDAR ’05]
[Koo and Cho, ECCV ’10]
[Zandifar, ICIAP ’07]

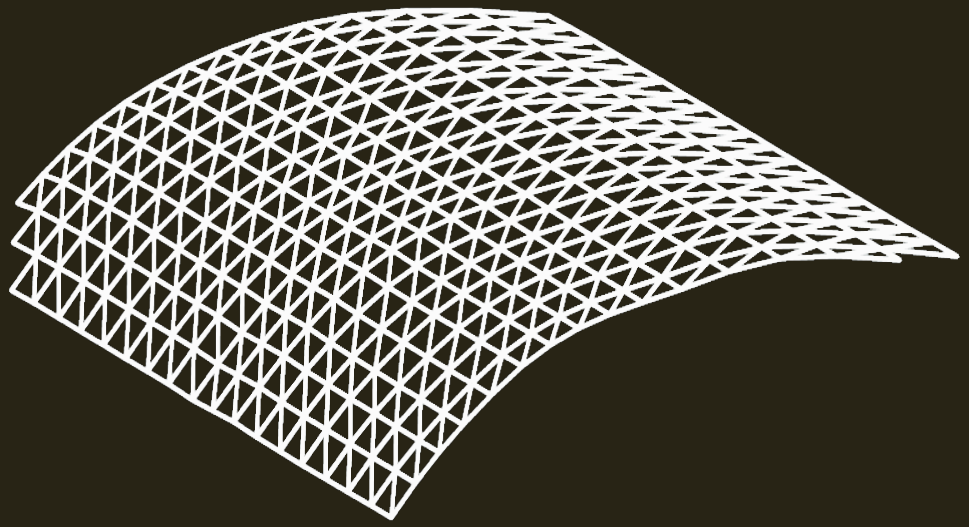
Image Binarization

[Cao et al, ICDAR ’03]

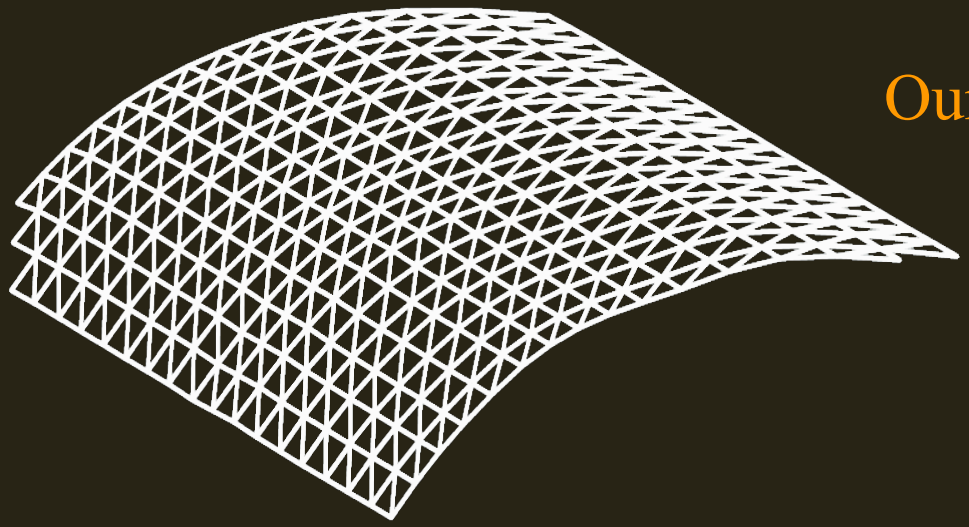
Contribution II: 3D Reconstruction



Contribution II: 3D Reconstruction



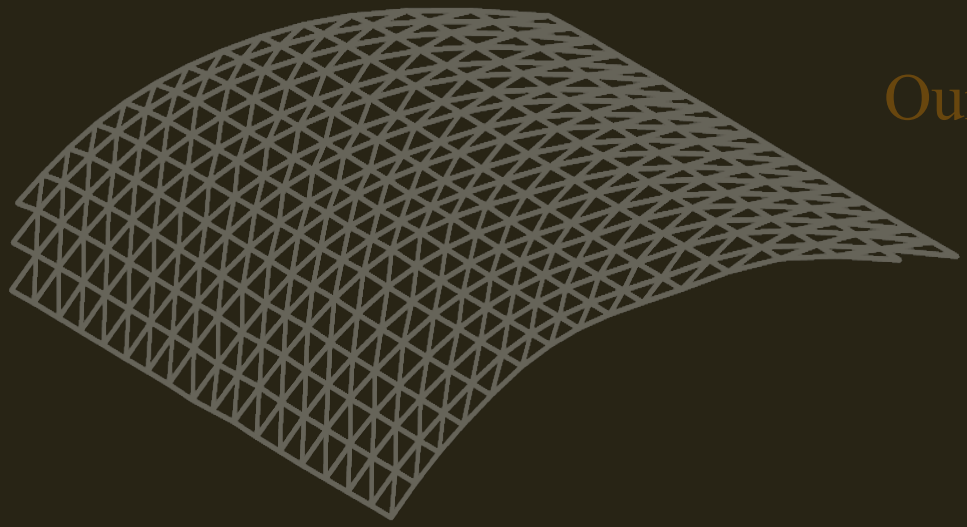
Contribution II: 3D Reconstruction



Our contribution:

1. From a single image
2. No global shape model
3. Globally optimal solution

Contribution II: 3D Reconstruction



Our contribution:

1. From a single image
2. No global shape model
3. Globally optimal solution

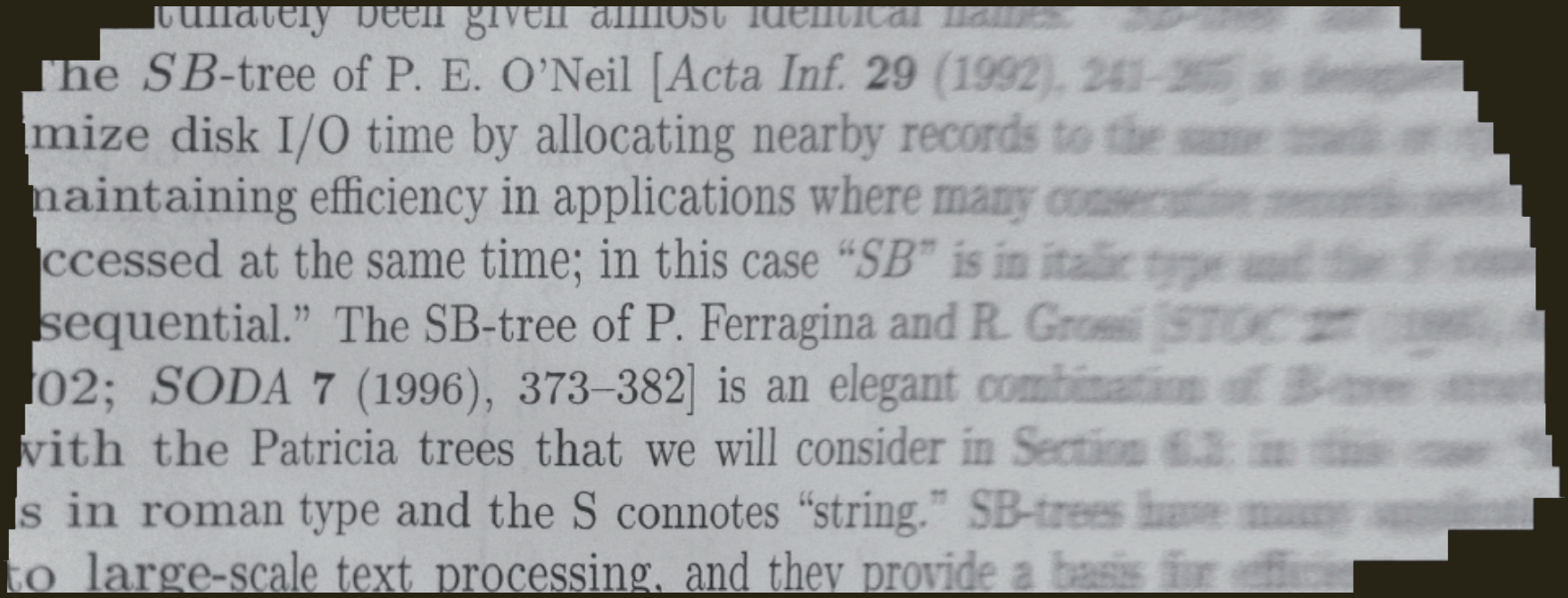
Previous works:

Laser scanner, stereo
[Brown and Seales, PAMI ’04]
[Based et al, HP White Paper ’01]
[Yamashita et al, ICPR ’04]

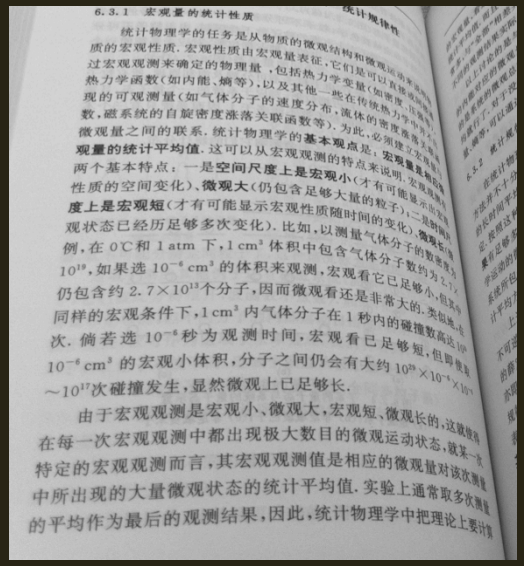
Cylinder / Developable
[Cao et al, ICDAR ’03]
[Zhang et al, CVPR ’04]
[Liang et al, PAMI ’07]

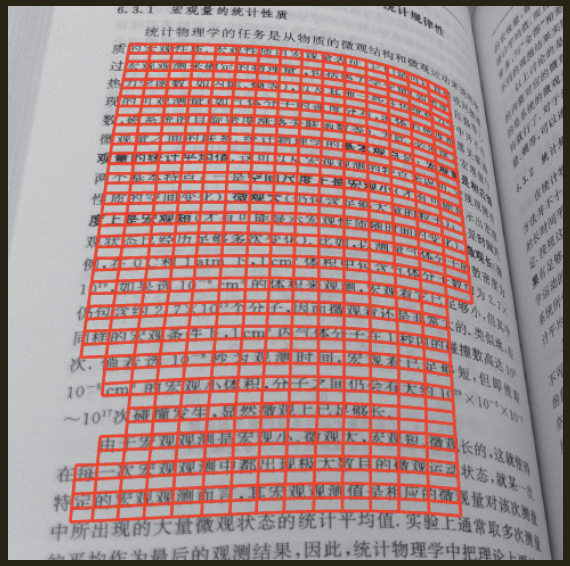
Shape-from-texture
[Forsyth, ICCV ’01]
[Witkin, Artificial Intelligence ’81]
[Malik and Rosenholtz, IJCV ’97]

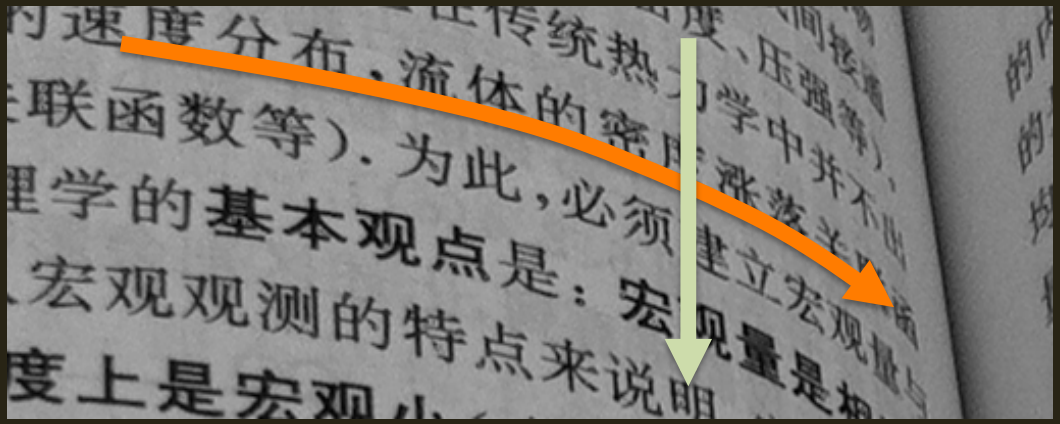
Final Rectification


tunately been given almost identical names.
The *SB*-tree of P. E. O'Neil [*Acta Inf.* **29** (1992), 241–265] is designed
mize disk I/O time by allocating nearby records to the same tracks of
naintaining efficiency in applications where many consecutive records need
ccessed at the same time; in this case "*SB*" is in italic type and the S con
sequential." The SB-tree of P. Ferragina and R. Grossi [*STOC* **27** (1995),
02; *SODA* **7** (1996), 373–382] is an elegant combination of B-tree struc
vith the Patricia trees that we will consider in Section 6.3; in this case "
s in roman type and the S connotes "string." SB-trees have many applica
to large-scale text processing, and they provide a basis for efficient

Horizontal and Vertical Text Directions

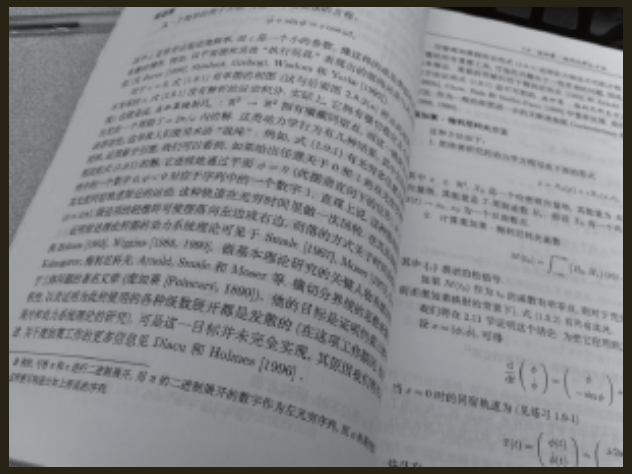


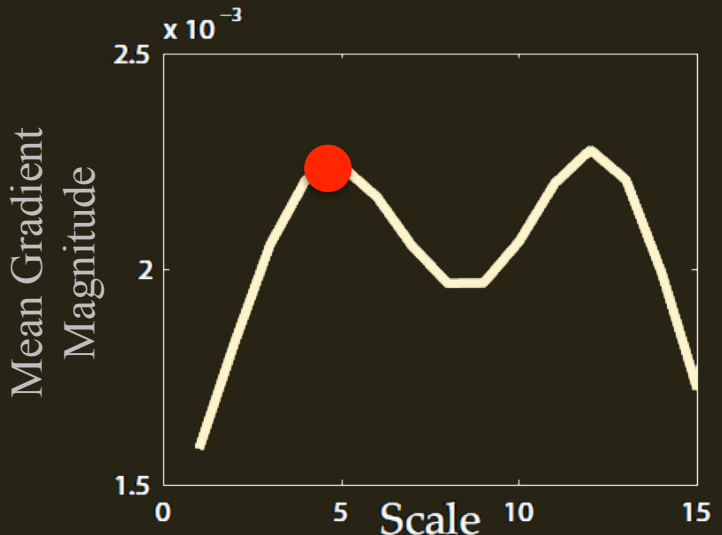


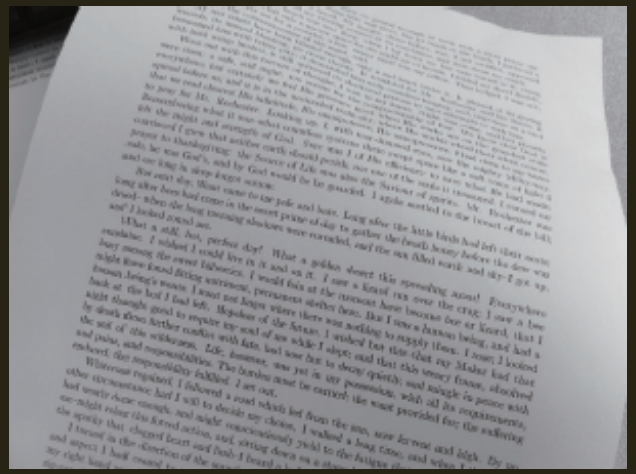


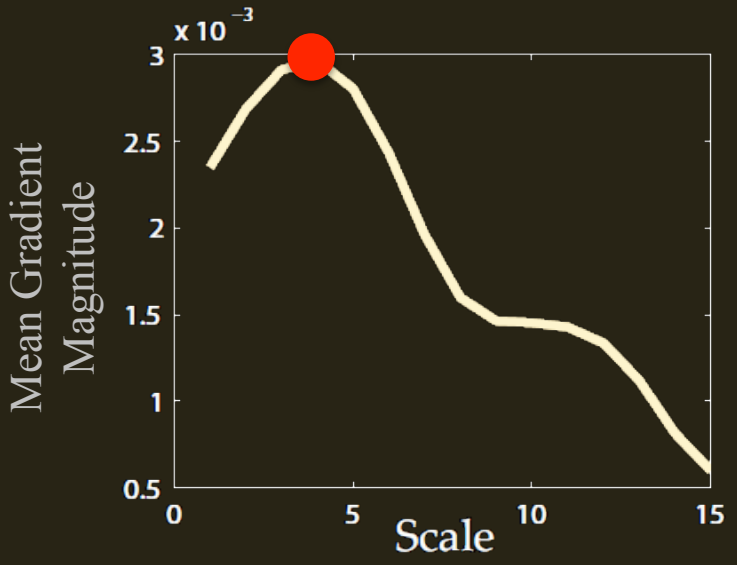
Horizontal Direction --- Rescaling

Horizontal Direction --- Rescaling



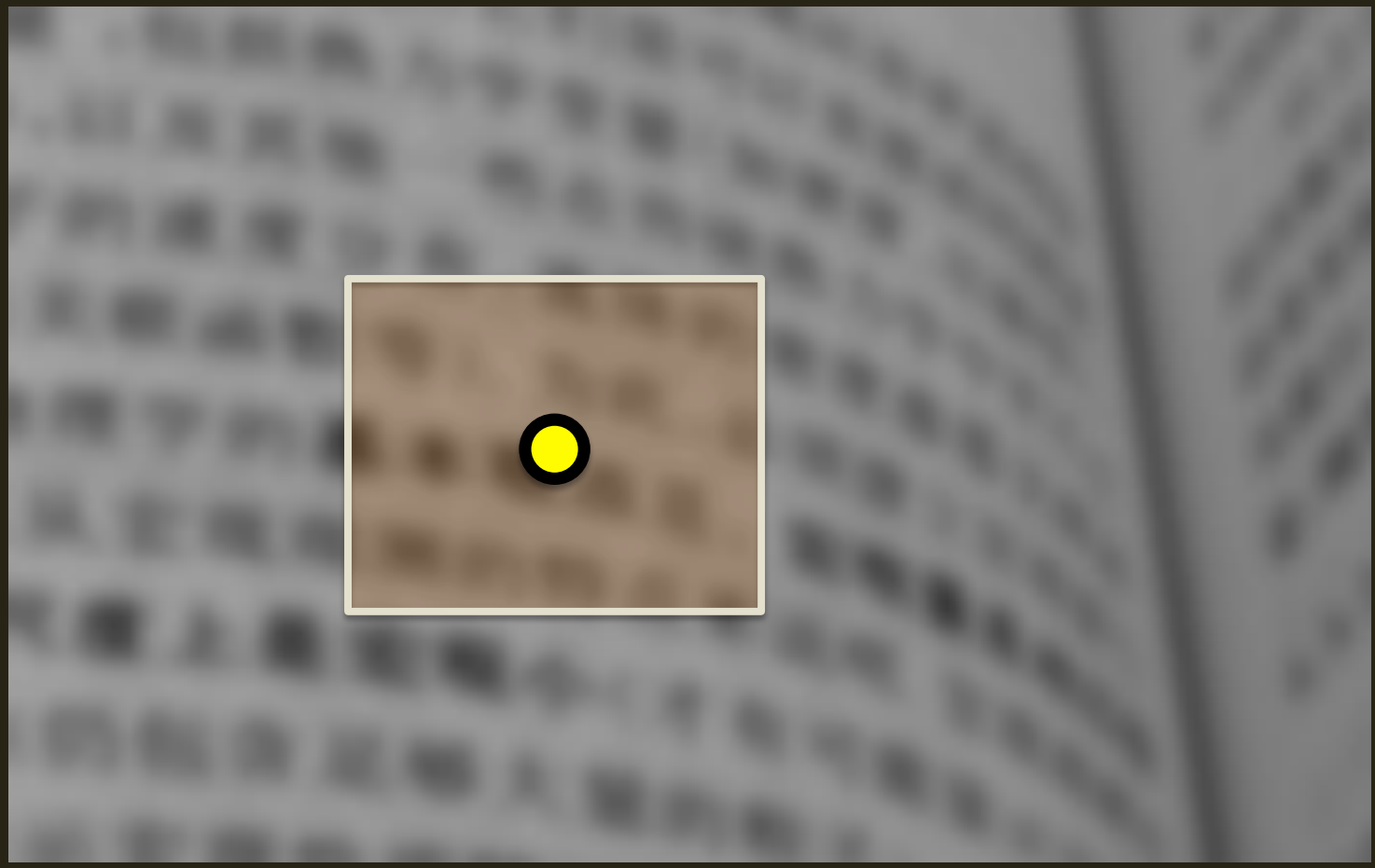




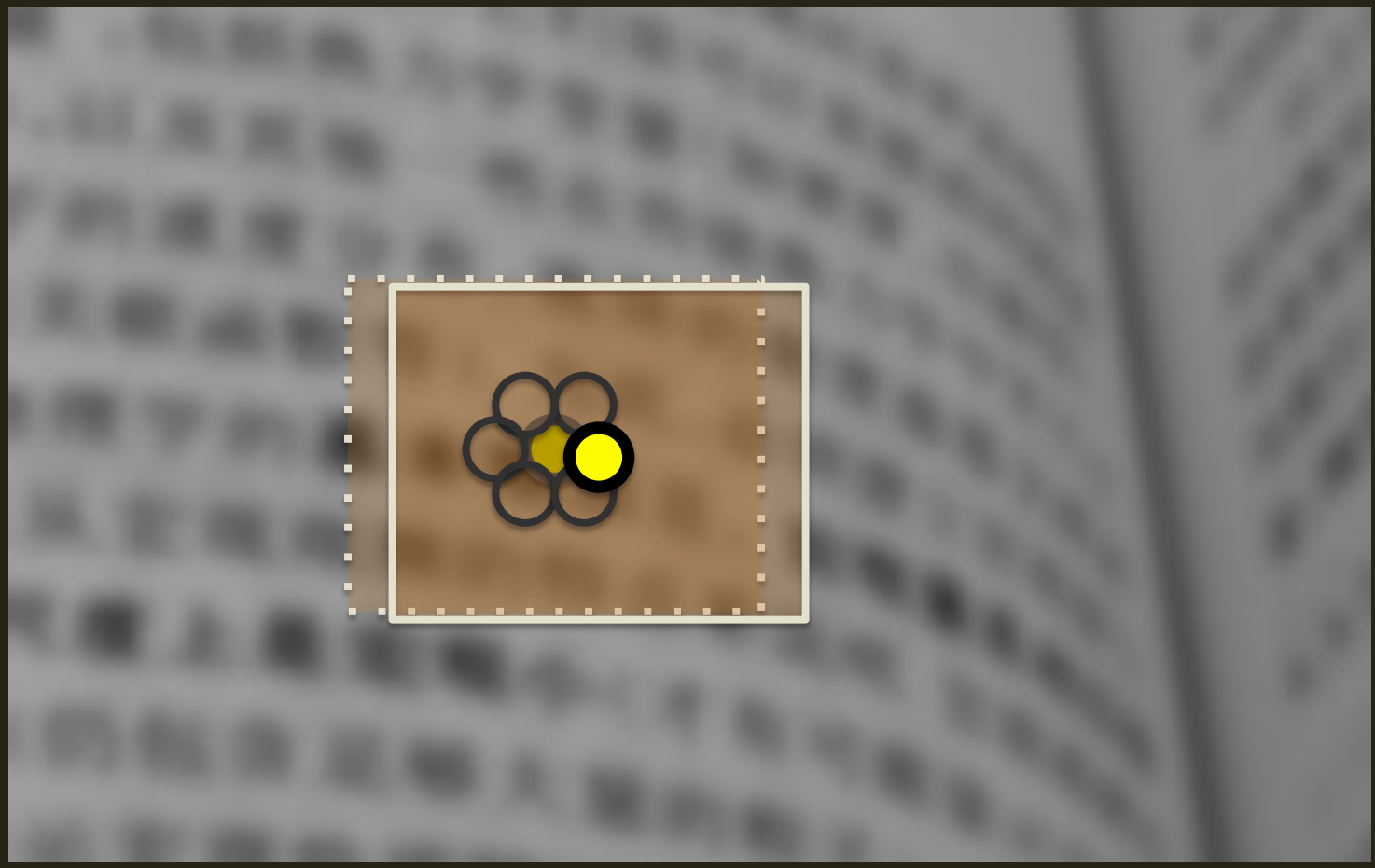


Horizontal Direction --- Line Tracing

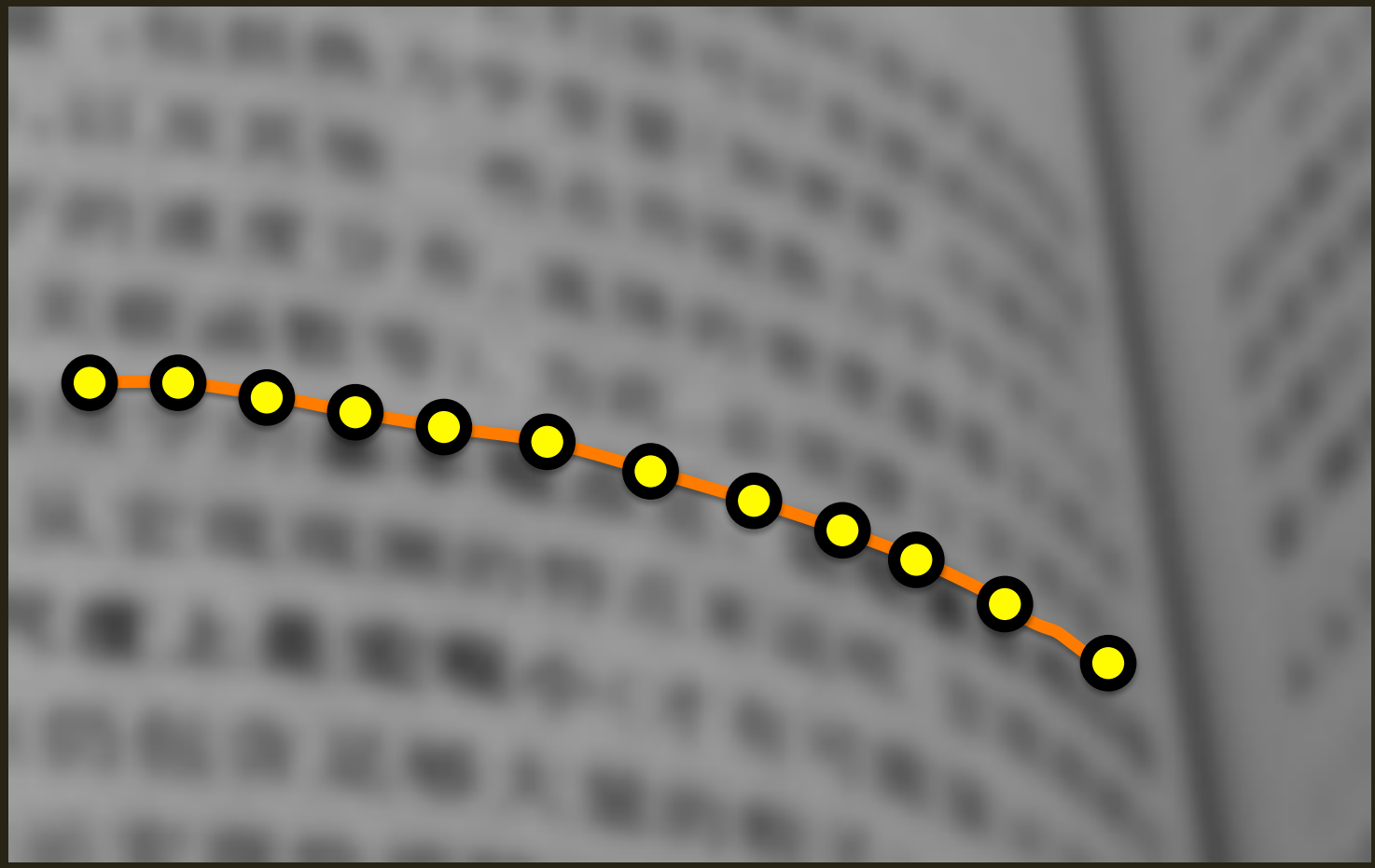
Horizontal Direction --- Line Tracing



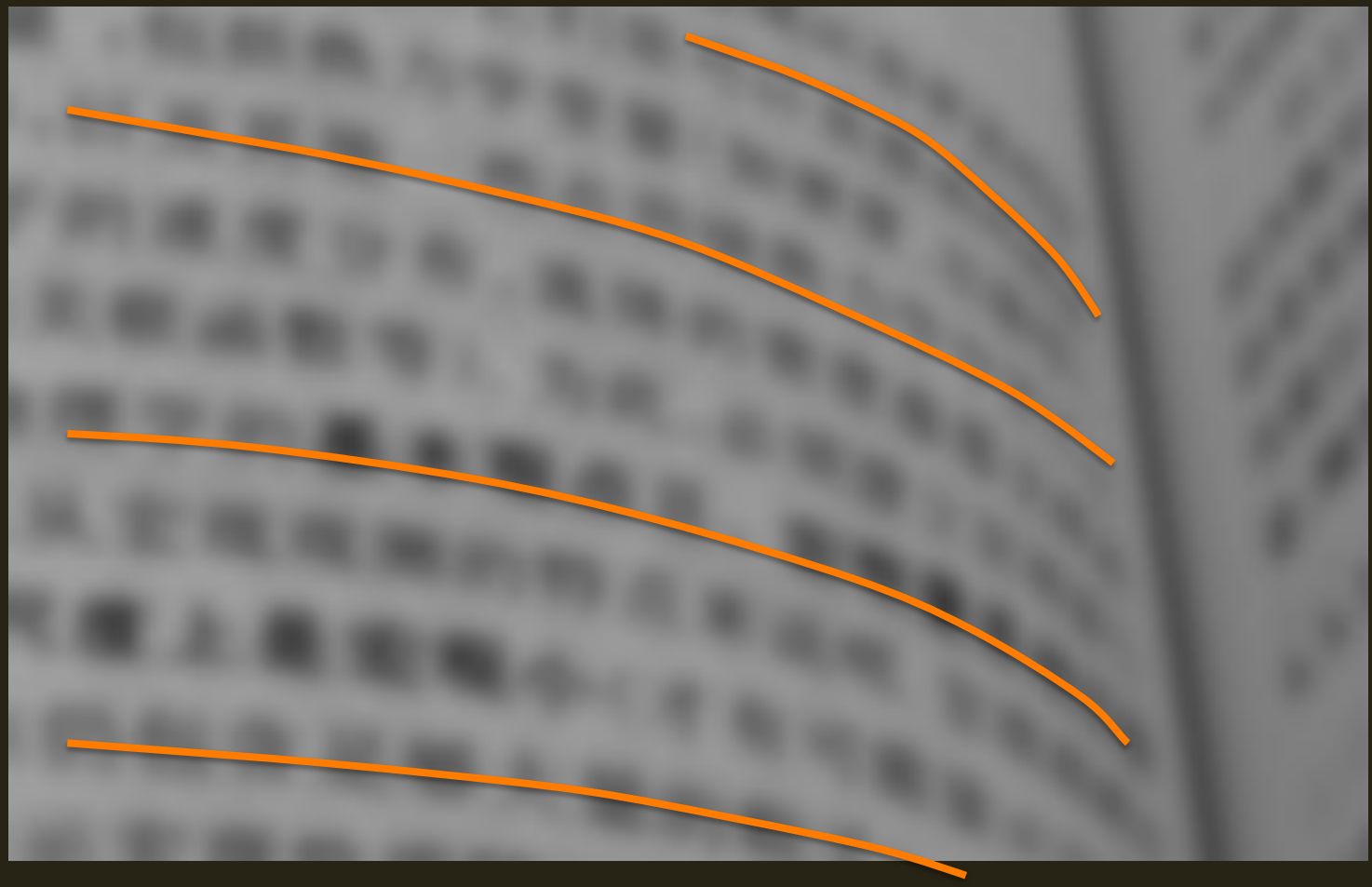
Horizontal Direction --- Line Tracing



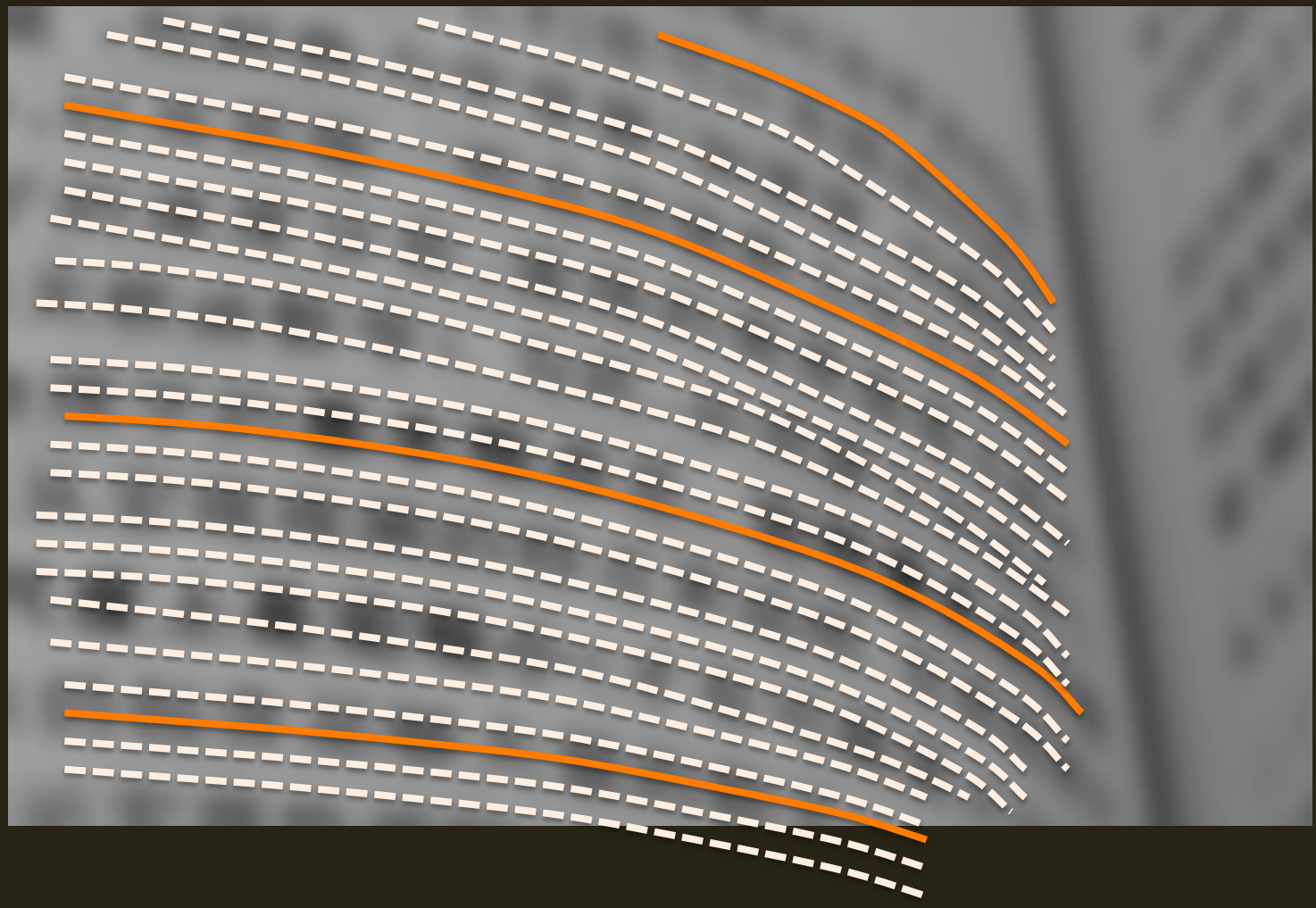
Horizontal Direction --- Line Tracing



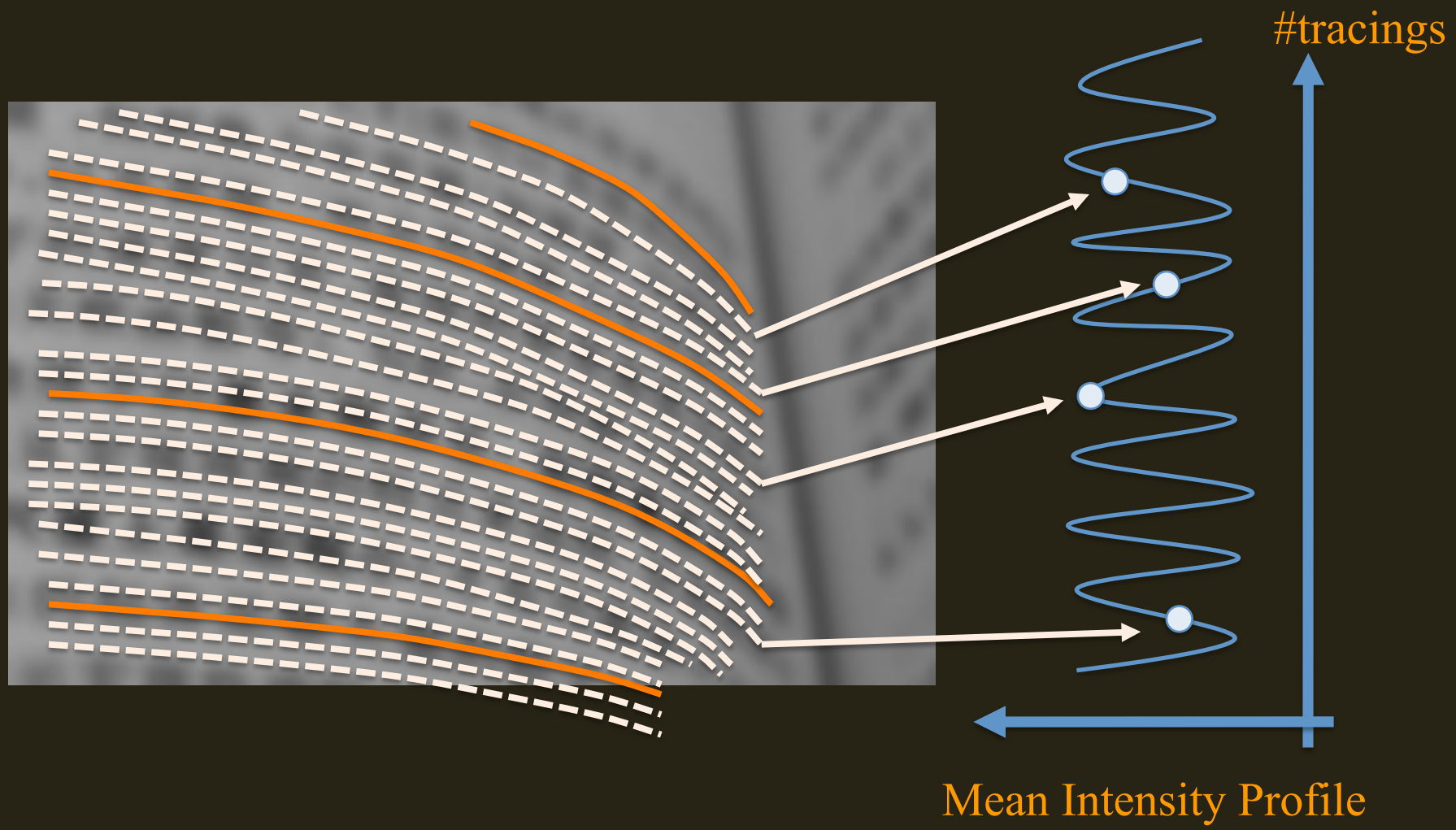
Horizontal Direction --- Line Re-sampling



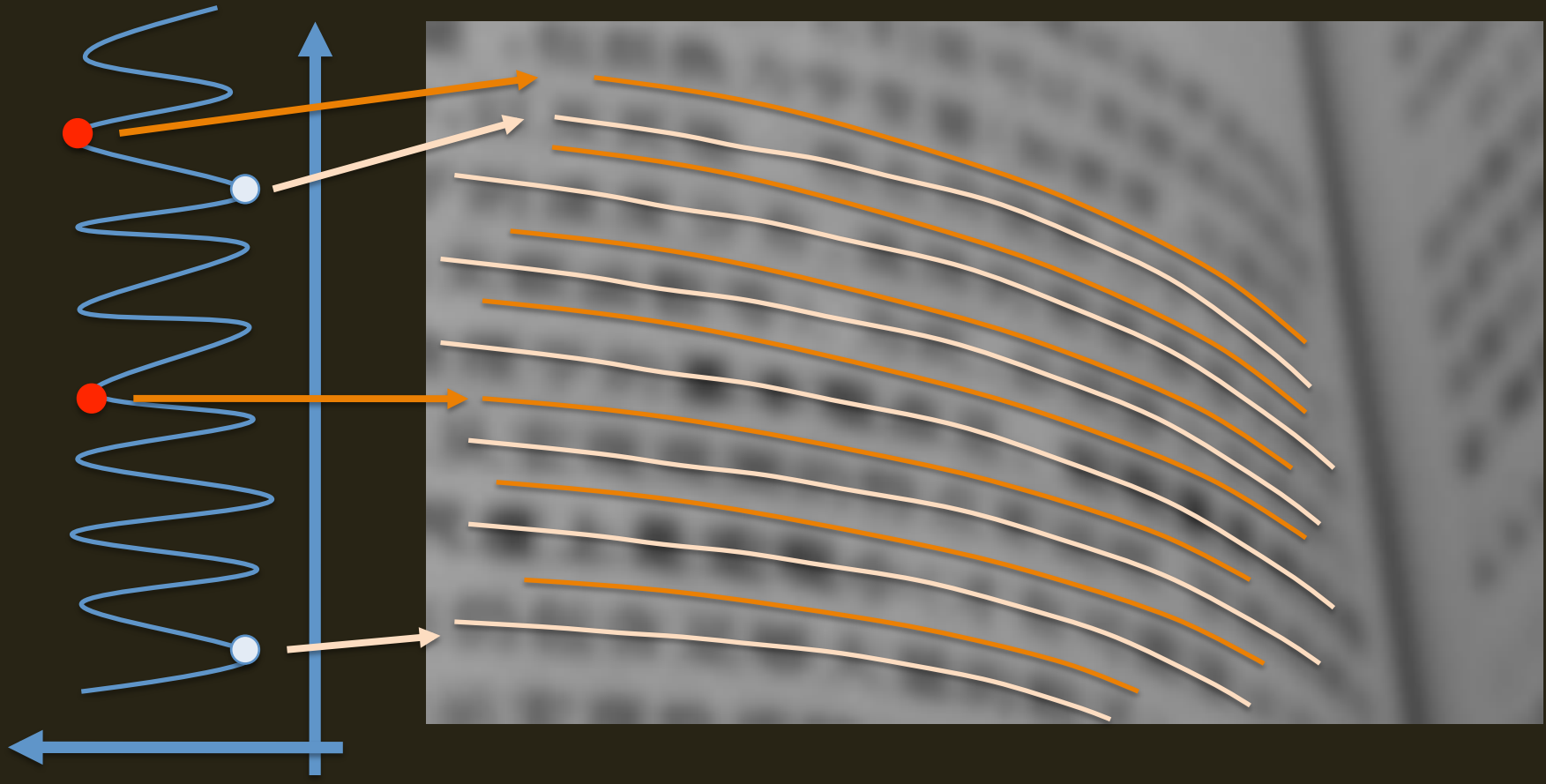
Horizontal Direction --- Line Re-sampling



Horizontal Direction --- Line Re-sampling

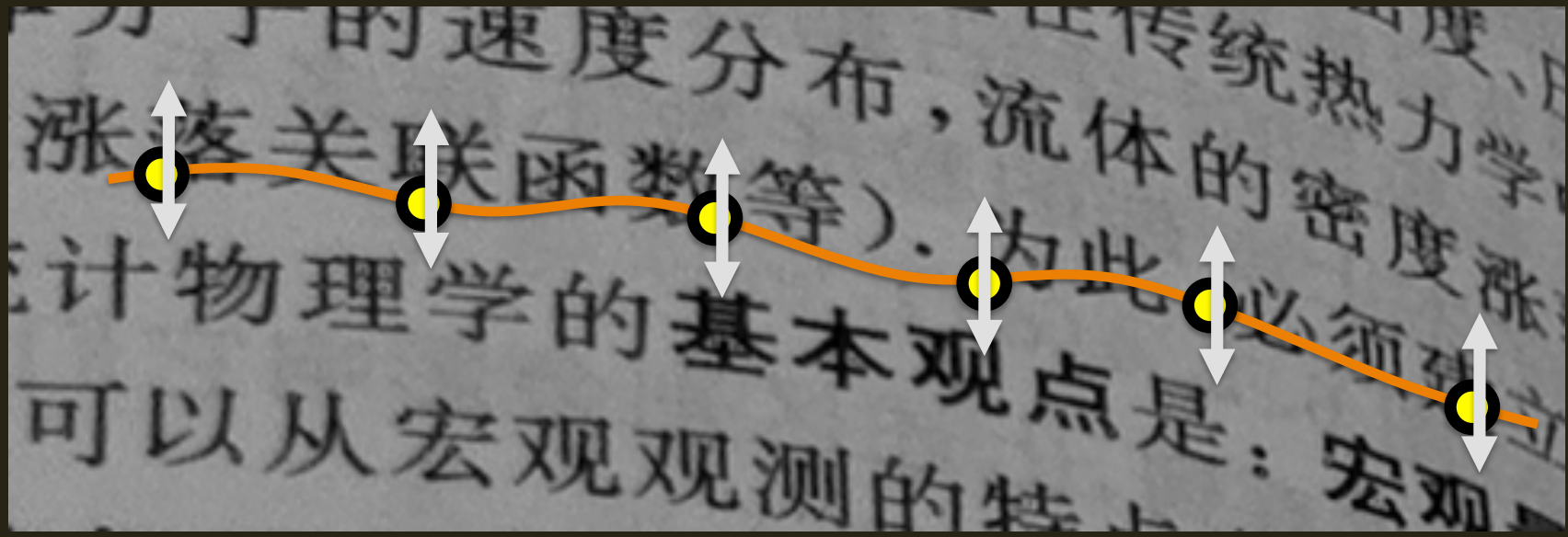


Horizontal Direction --- Line Re-sampling



Mean Intensity Profile

Horizontal Direction --- Line Refinement



Vertical Direction

b, d, l, k, f, B, E, F, L, D, M

English

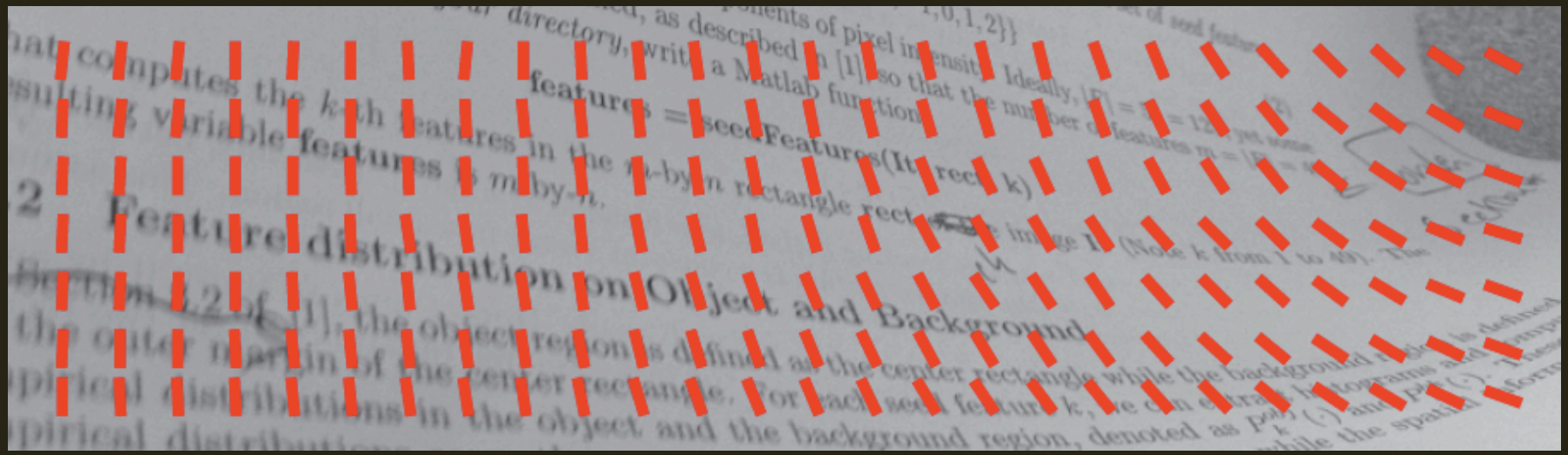
量子力学是描写微观物质的一个物理学理论，与相对论一起被认
理学的两大基本支柱，许多物理学理论和科学如原子物理学、固
学、核物理学和粒子物理学以及其它相关的学科都是以量子力学

Chinese

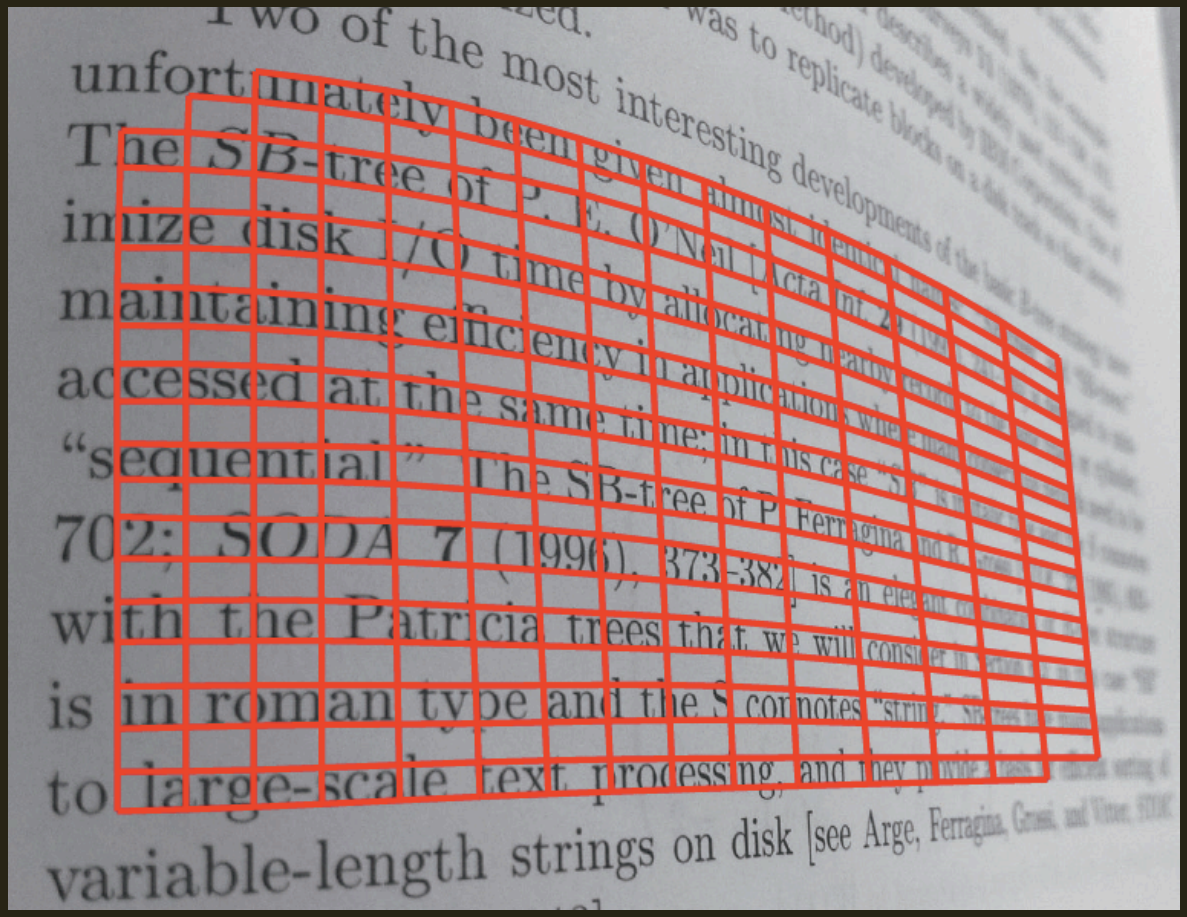
दिल्ली एक शहर है, लेकिन दिल्ली के दो केंद्र हैं।
दिल्ली का एक केंद्र चाँदनी चौक है और दिल्ली
का दूसरा केंद्र कनाट प्लेस है। लोग कहते

Hindi

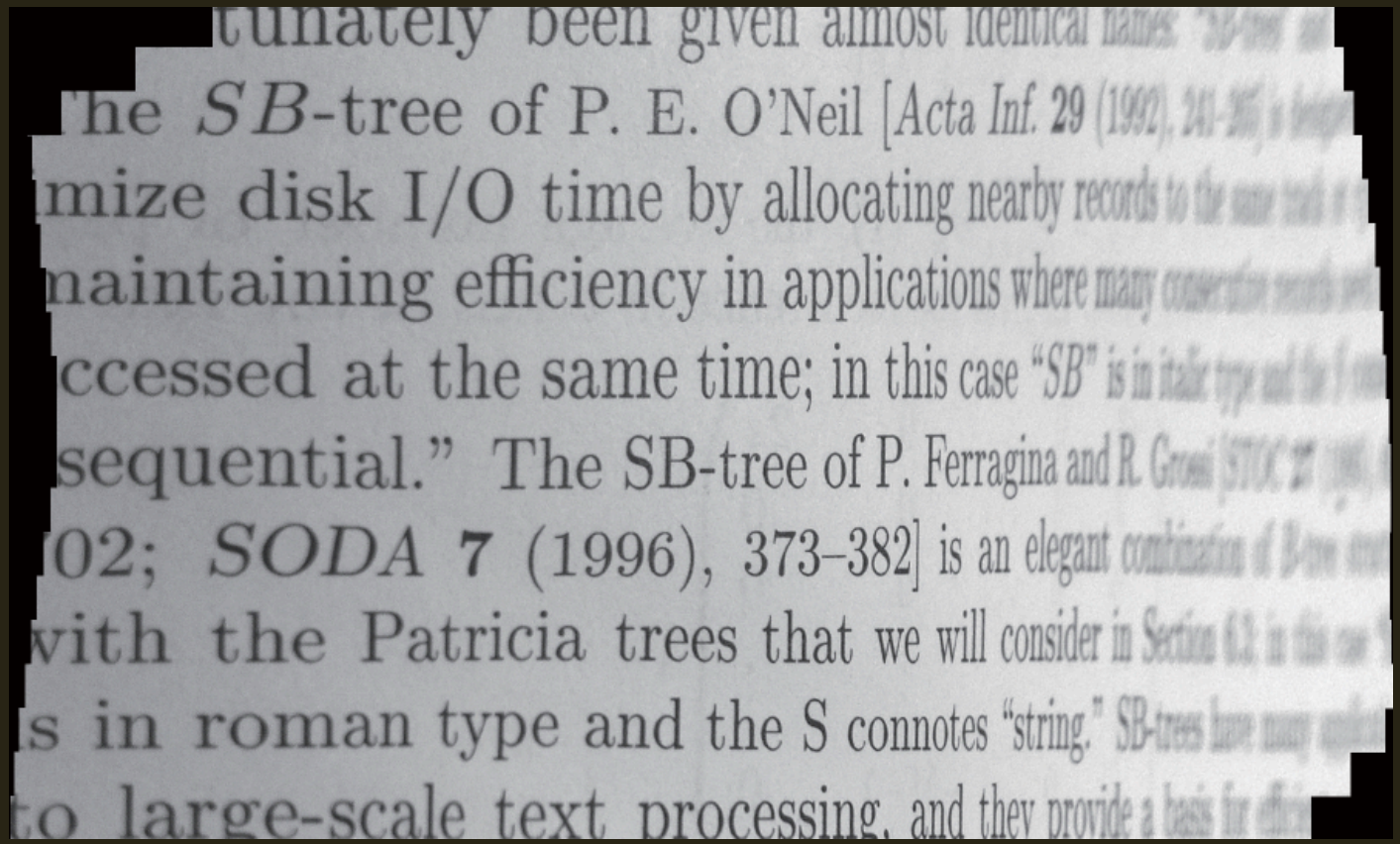
Vertical Direction



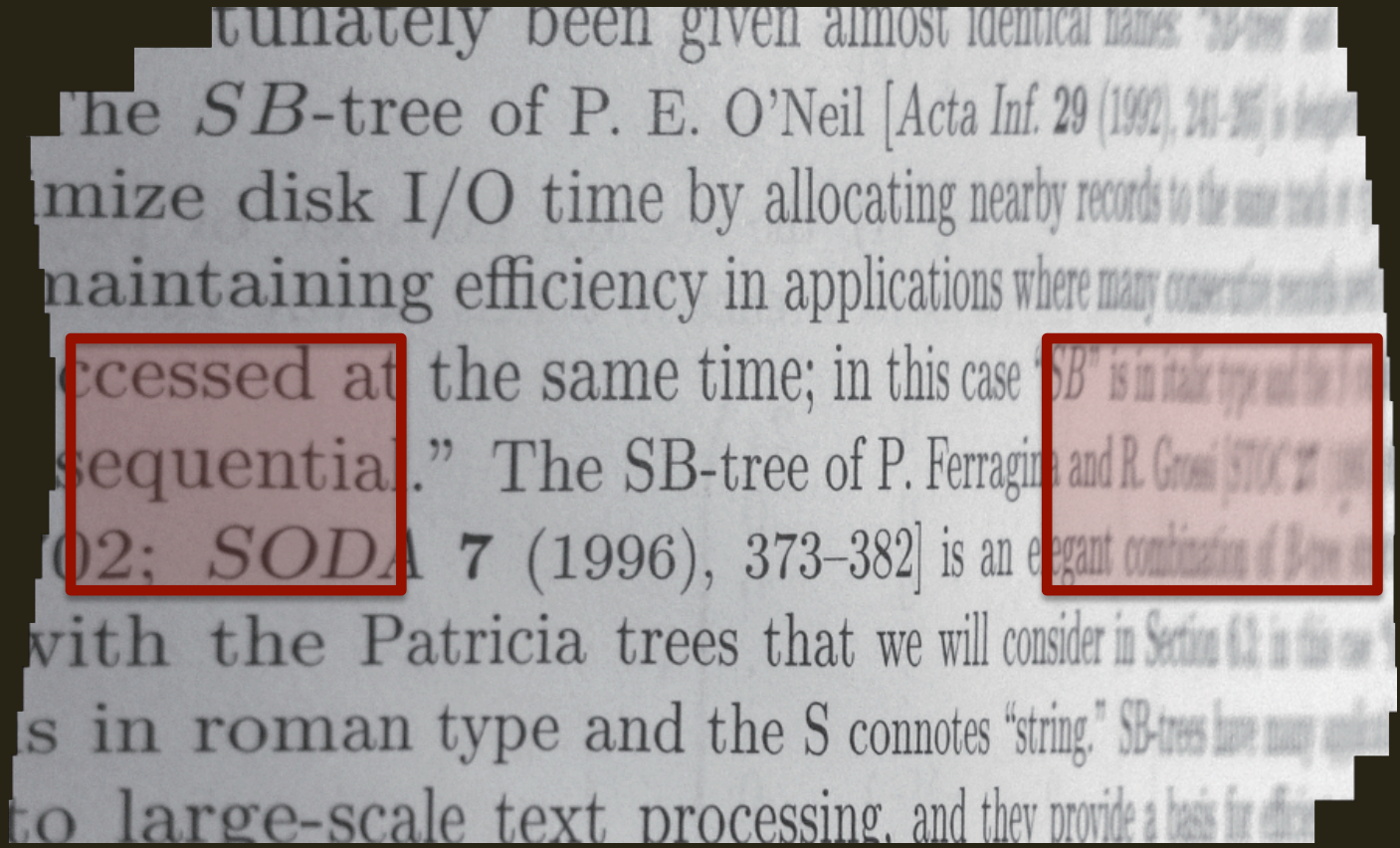
Rectification using 2D information



Rectification using 2D information



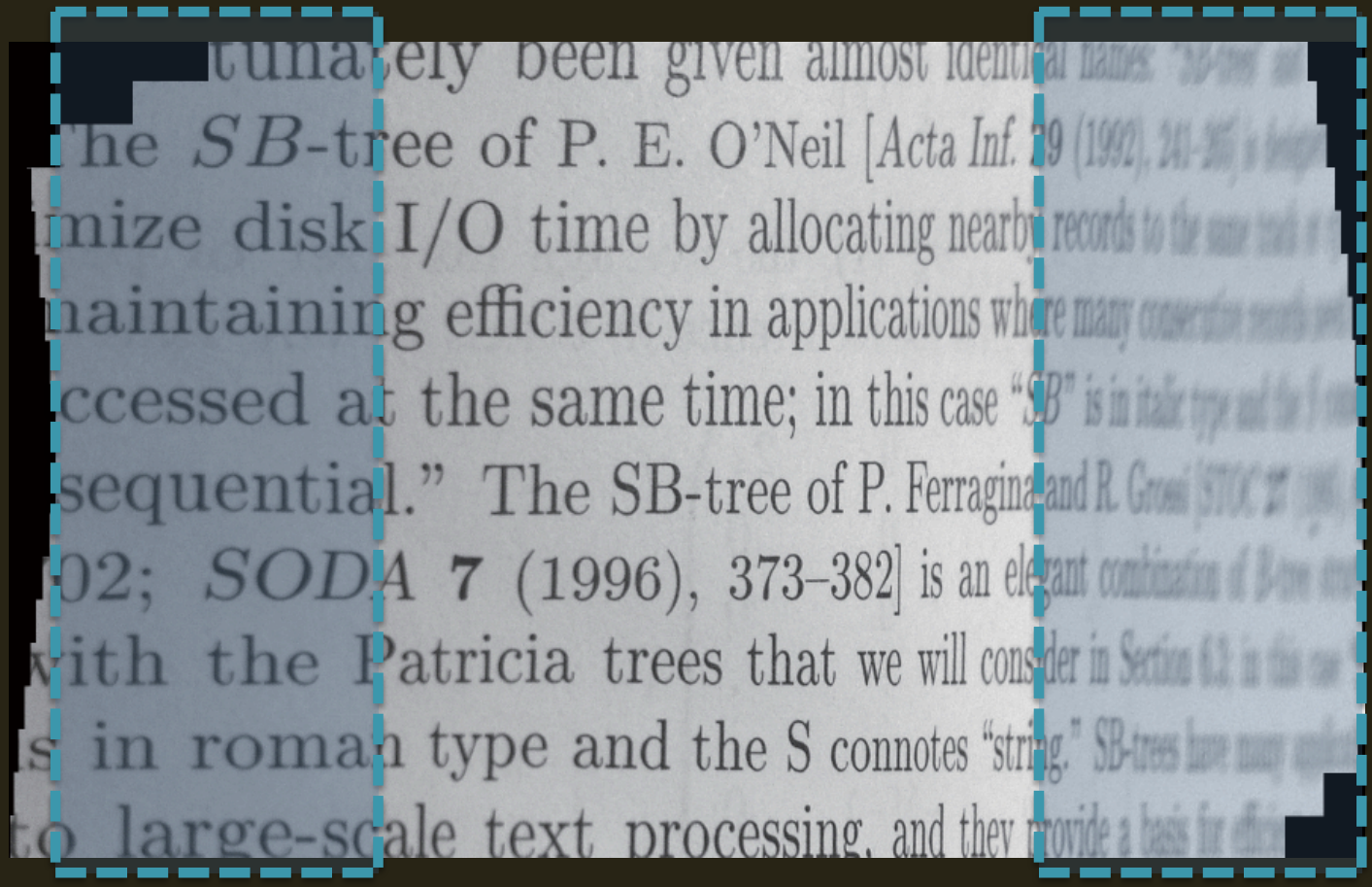
Rectification using 2D information



Close regions

Faraway regions

Rectification using 2D information

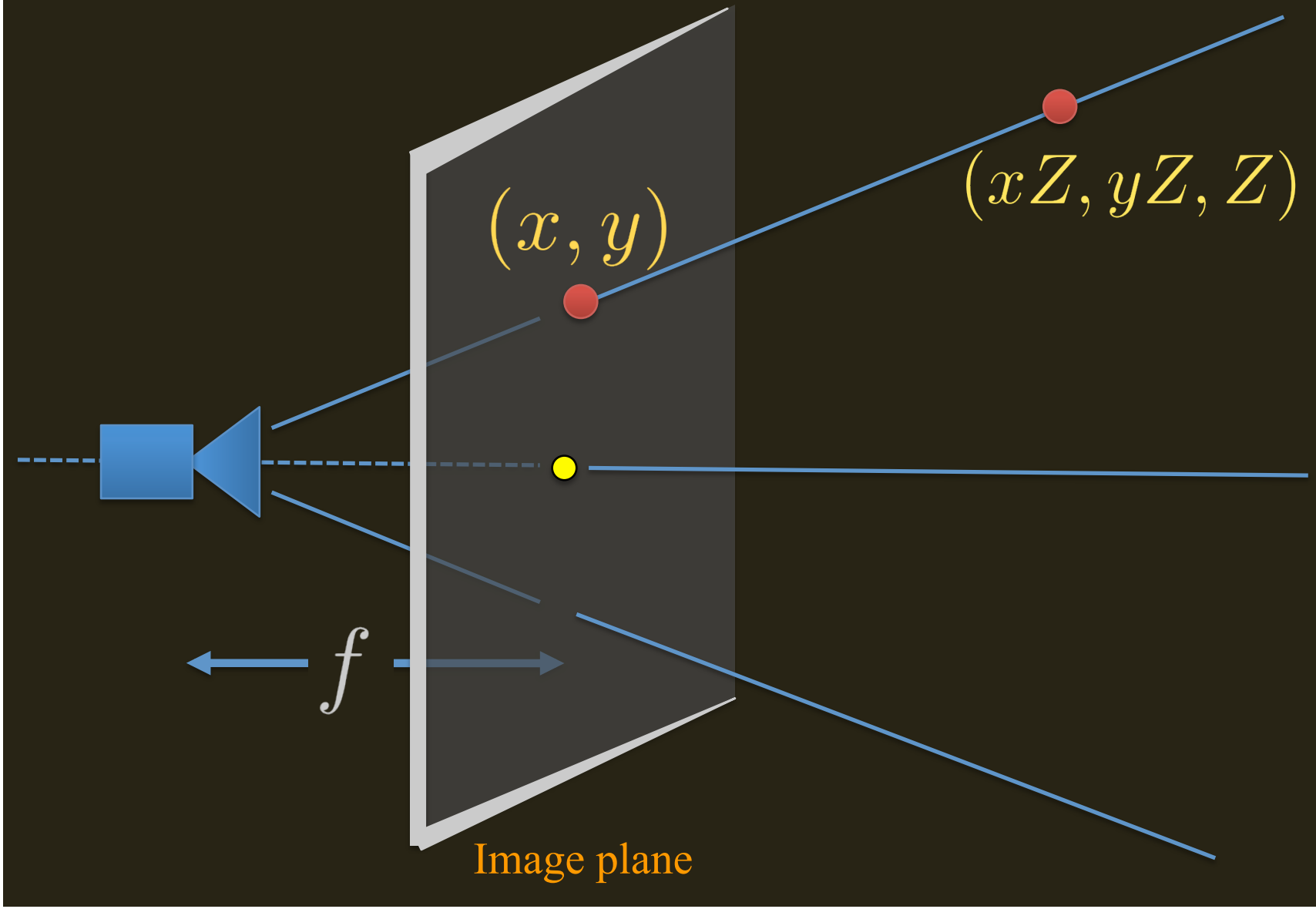


Darker regions

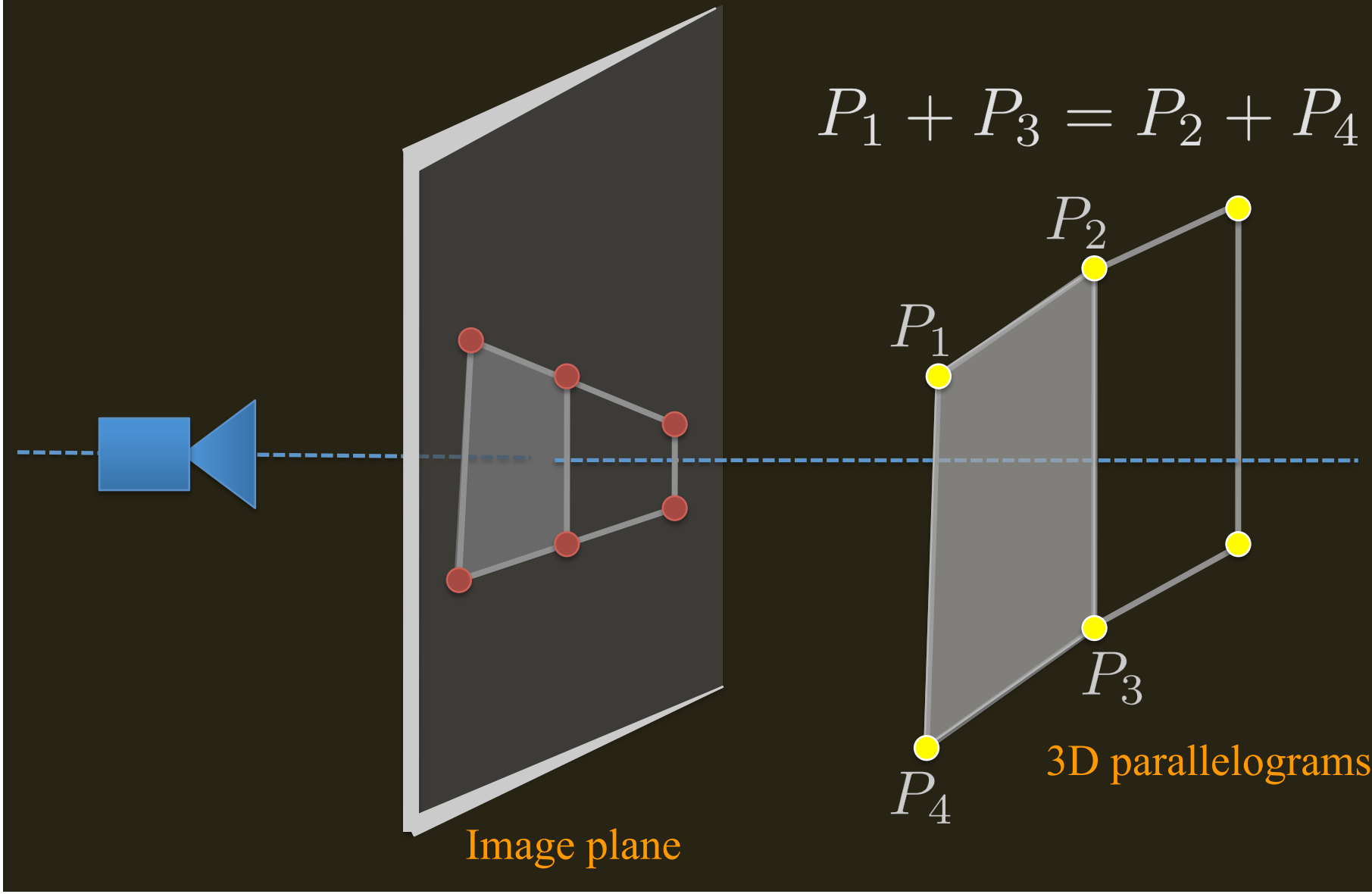
Brighter regions

3D reconstruction

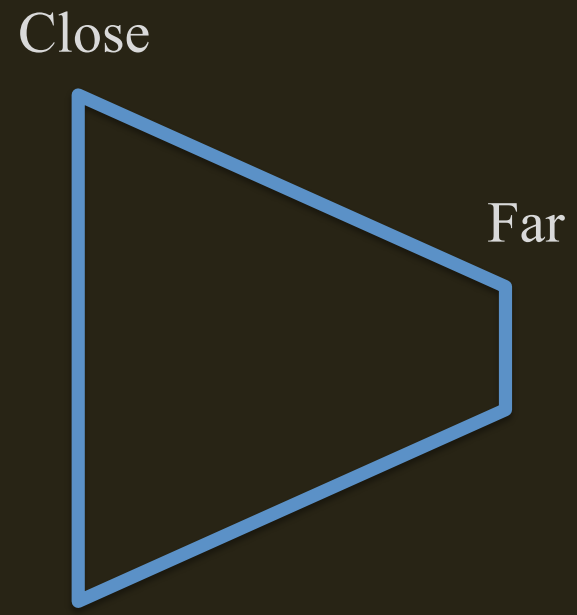
3D reconstruction --- Assumption I



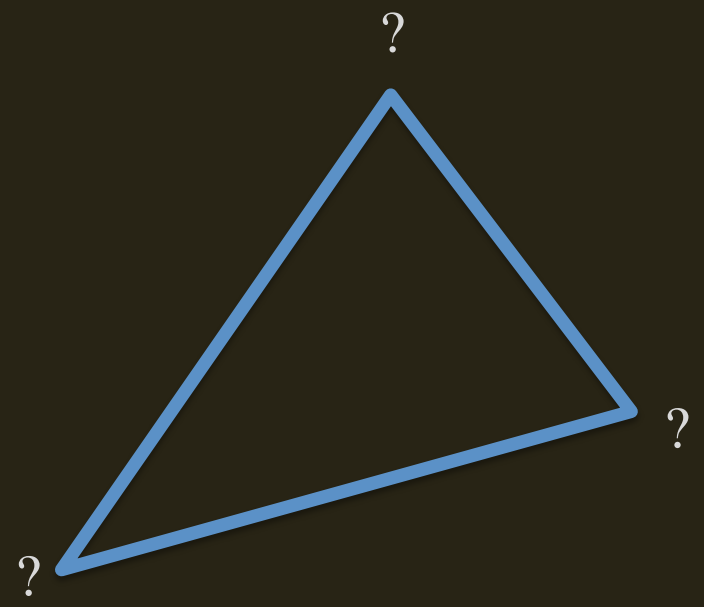
3D reconstruction --- Assumption II



Quadrilaterals versus Triangles



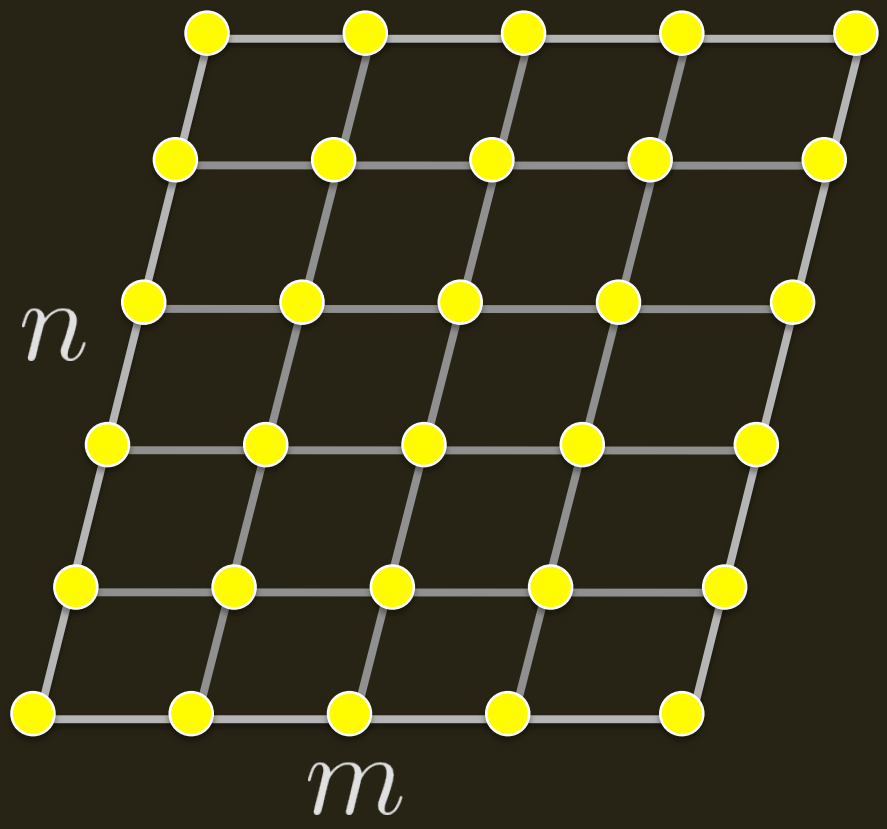
Our approach



Previous works

[J. Taylor et al, CVPR ’10]

Unknowns versus Constraints



$$mn \quad < \quad 3(m-1)(n-1)$$

unknowns constraints

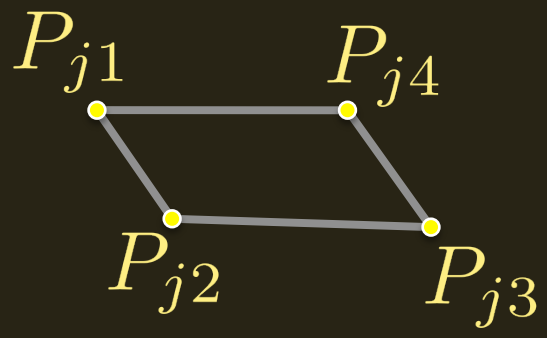
Formulation

Parameterization

$$P_i = (x_i Z_i, y_i Z_i, Z_i) \qquad \mathbf{Z} = [Z_1, Z_2, \ldots, Z_n]$$

Deviation from Parallelogram

$$\Delta_j = P_{j1} - P_{j2} + P_{j3} - P_{j4}$$

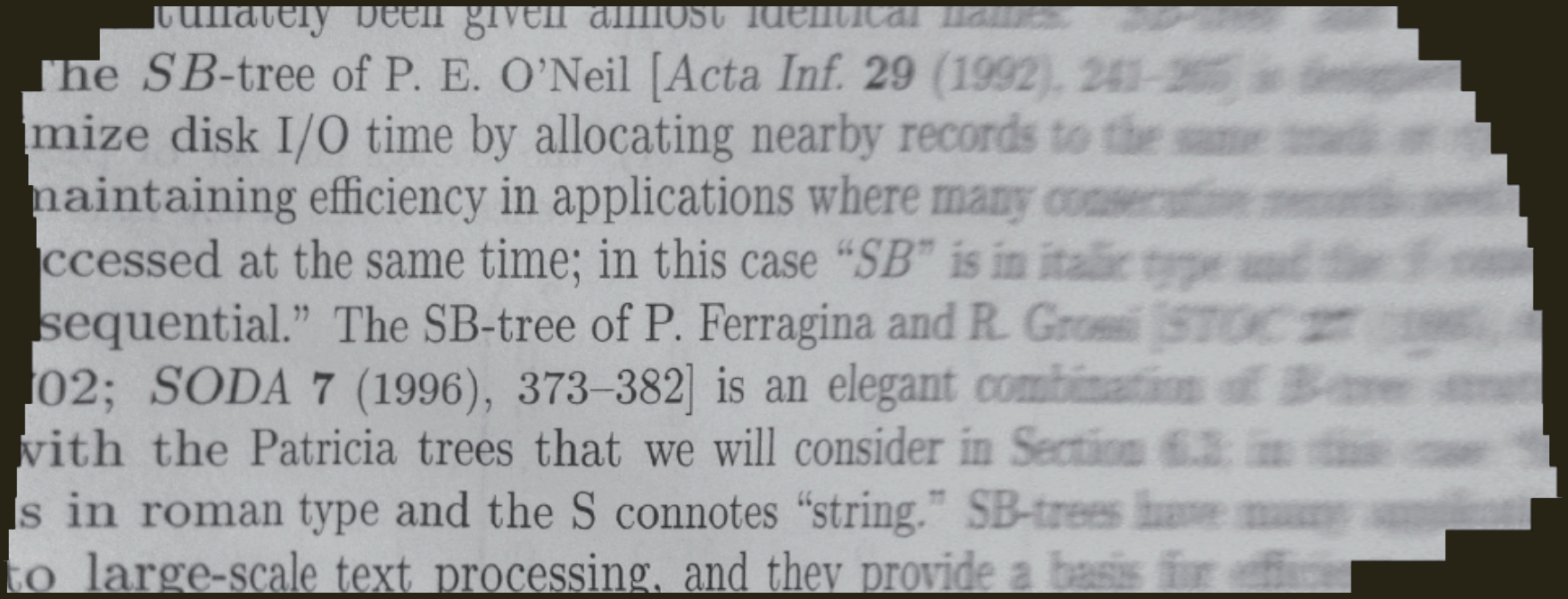


Minimization of sum of deviation:

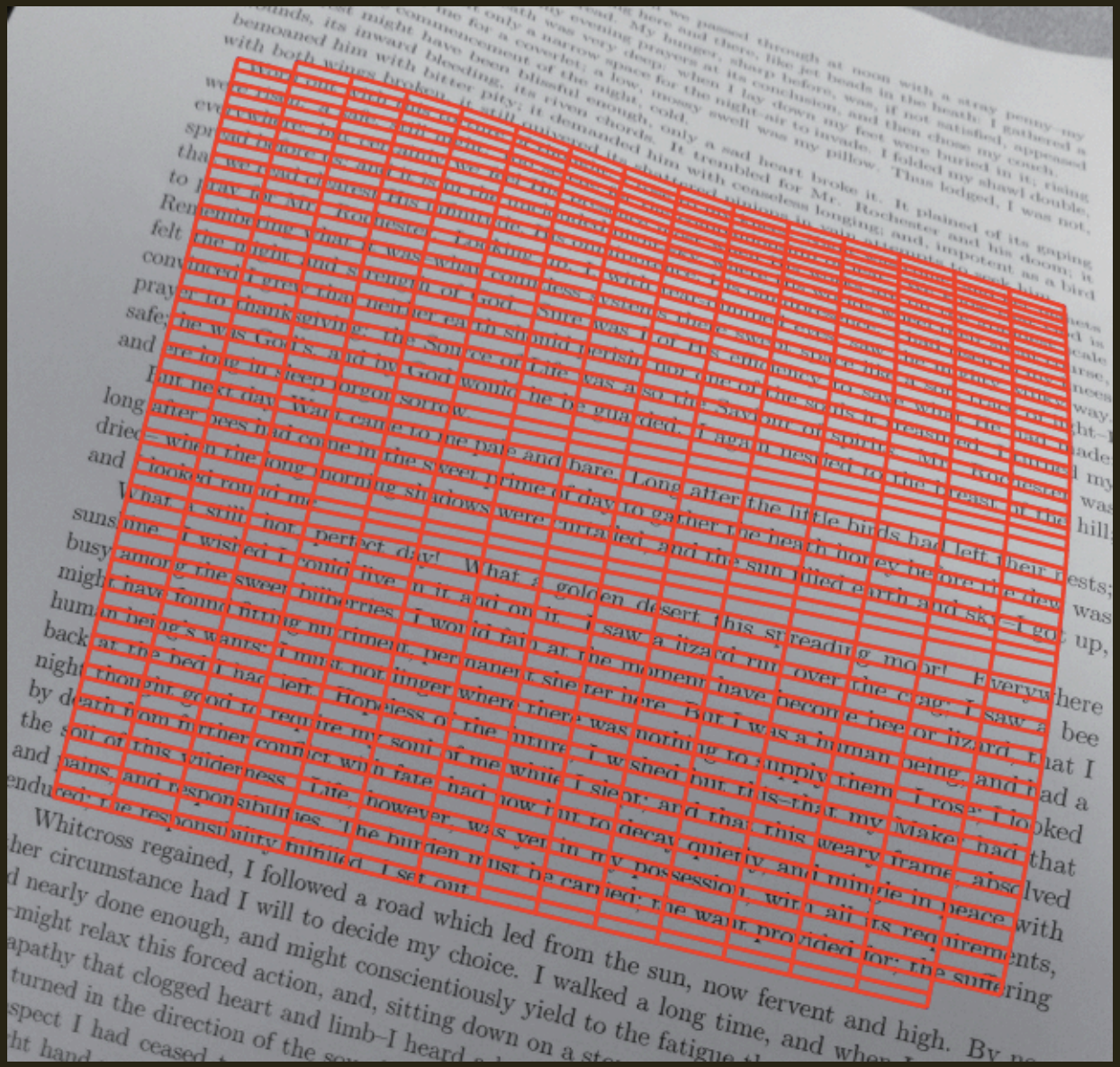
$$f(\mathbf{Z}) = \sum_{j=1}^{M} \|\Delta_j\|^2 \qquad \text{s.t.} \quad \|\mathbf{Z}\|^2 = 1$$

Homogenous Least Square → SVD

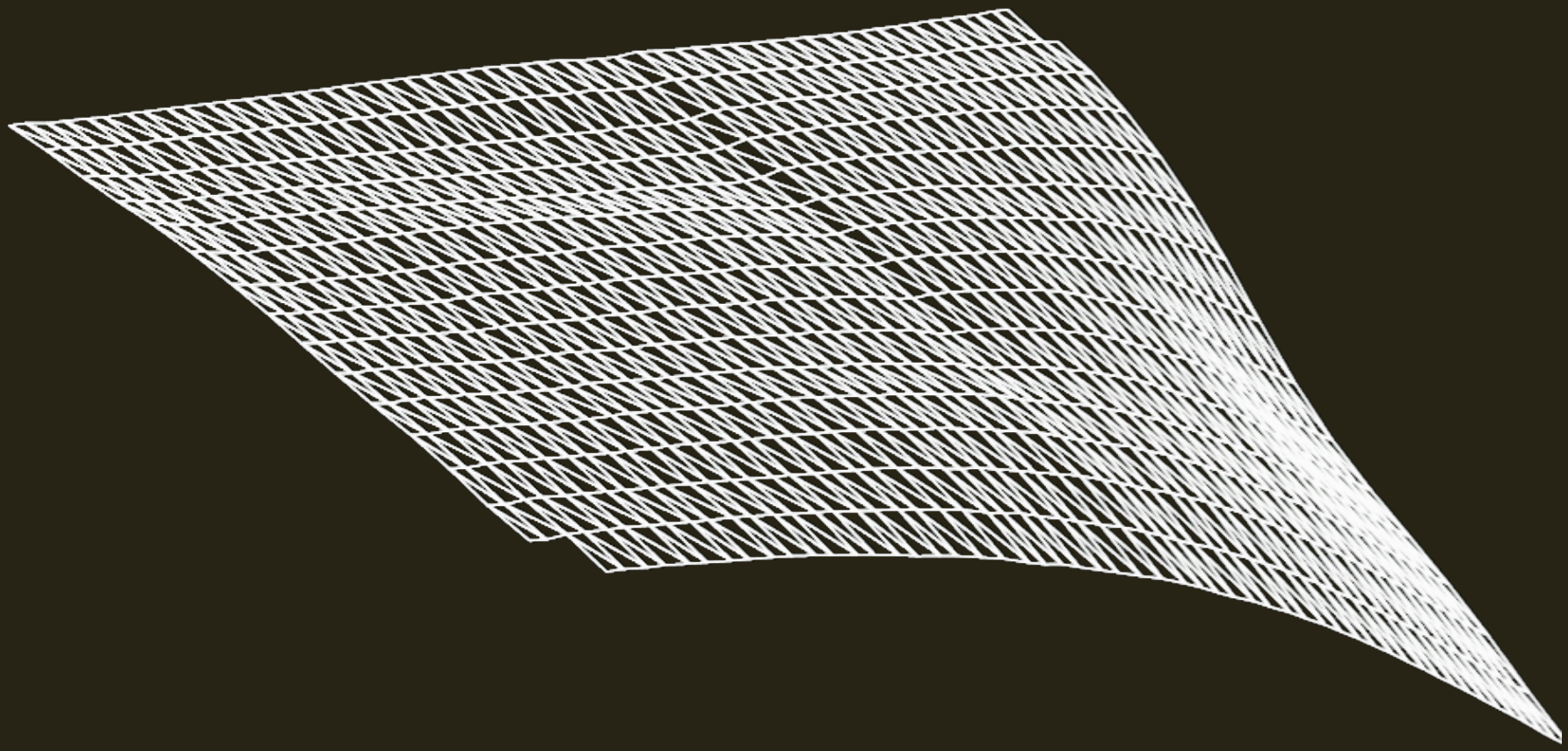
3D-aided Rectification



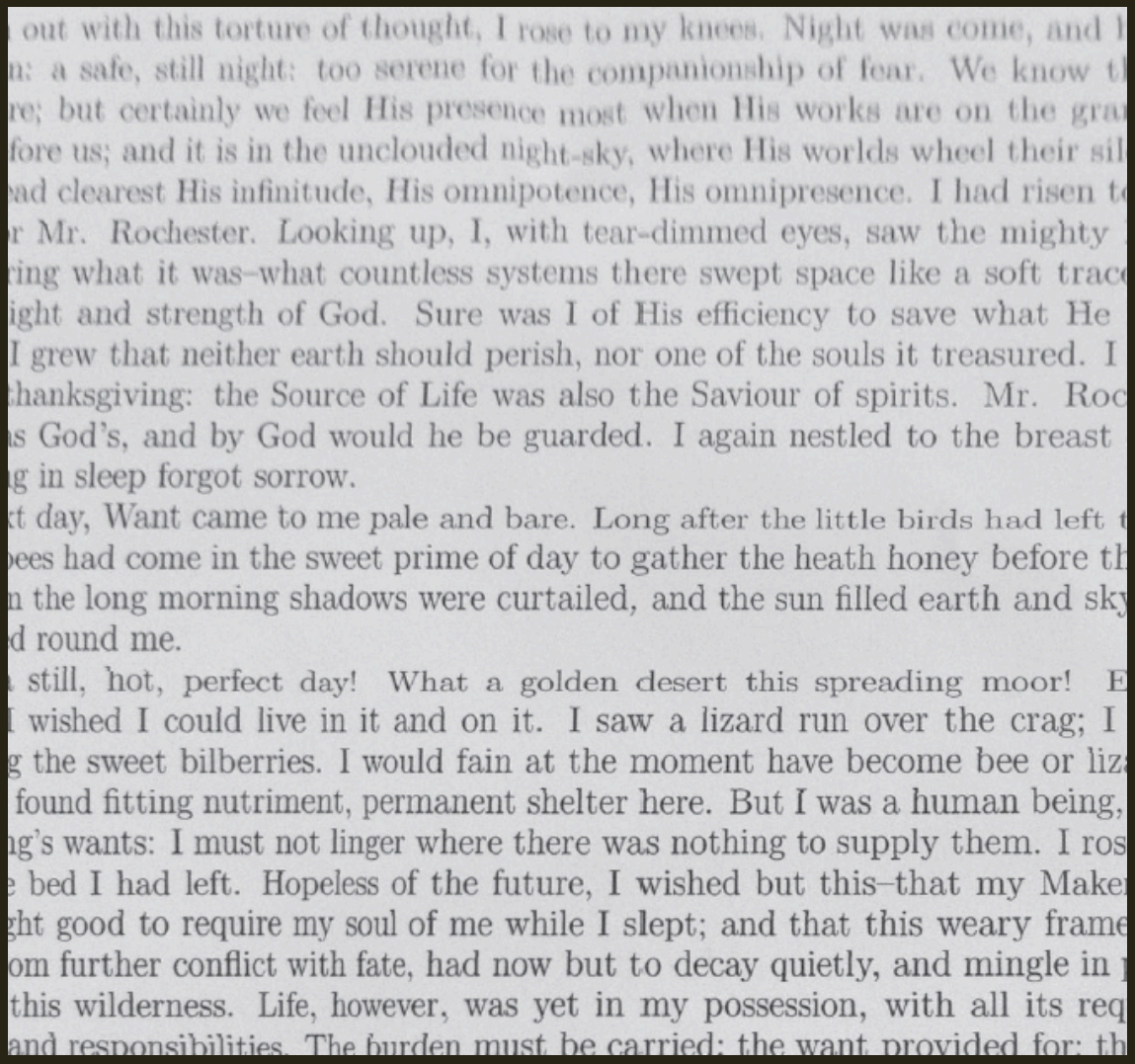
More results



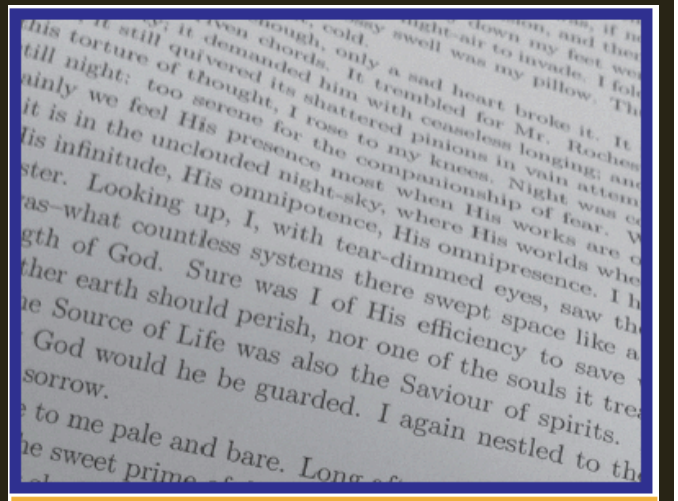
More results



More results

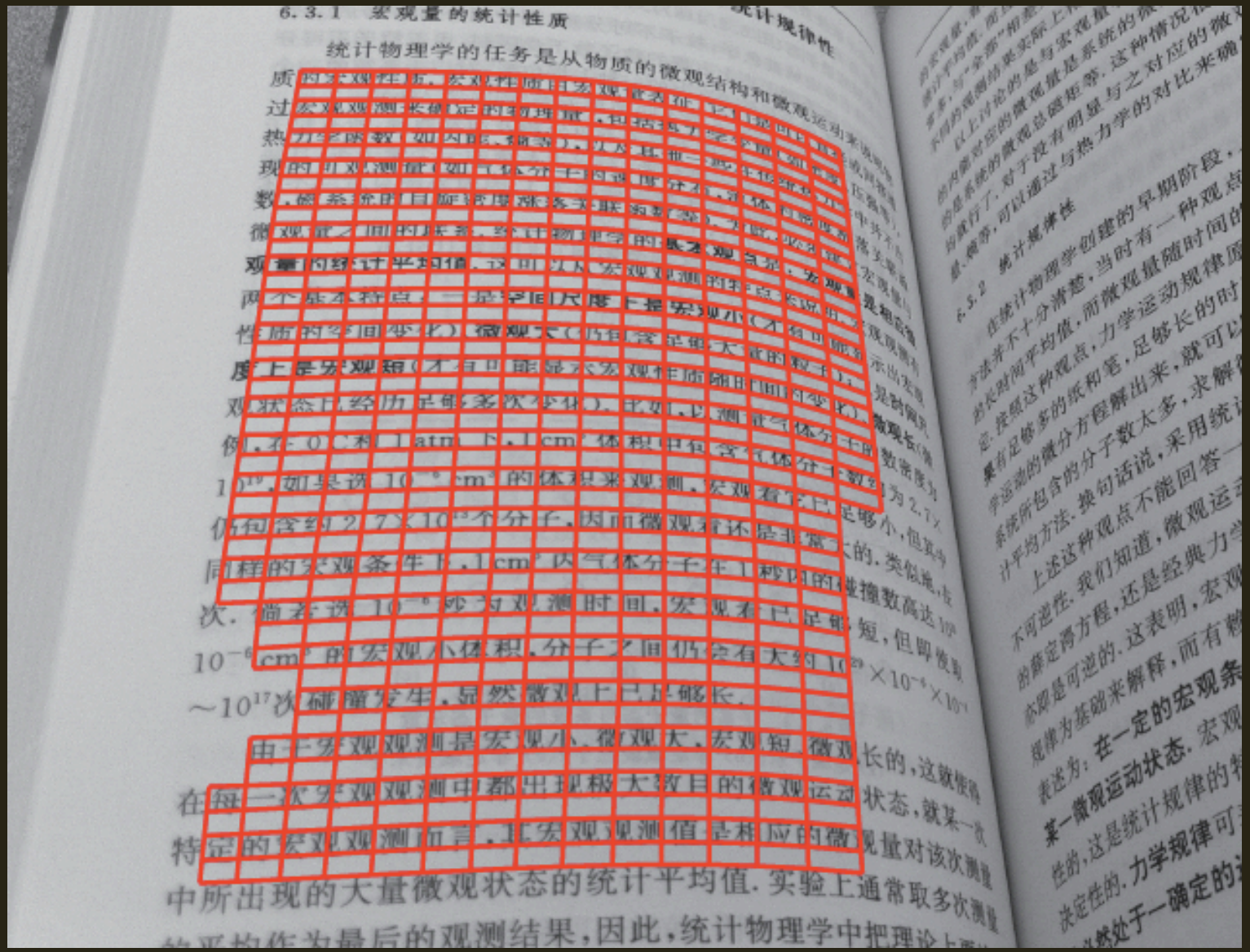


out with this torture of thought, I rose to my knees. Night was come, and
n: a safe, still night: too serene for the companionship of fear. We know t
re; but certainly we feel His presence most when His works are on the gra
fore us; and it is in the unclouded night-sky, where His worlds wheel their sil
ad clearest His infinitude, His omnipotence, His omnipresence. I had risen t
r Mr. Rochester. Looking up, I, with tear-dimmed eyes, saw the mighty
ring what it was–what countless systems there swept space like a soft trac
ight and strength of God. Sure was I of His efficiency to save what He
I grew that neither earth should perish, nor one of the souls it treasured. I
thanksgiving: the Source of Life was also the Saviour of spirits. Mr. Roc
s God's, and by God would he be guarded. I again nestled to the breast
g in sleep forgot sorrow.
t day, Want came to me pale and bare. Long after the little birds had left t
bees had come in the sweet prime of day to gather the heath honey before th
n the long morning shadows were curtailed, and the sun filled earth and sky
d round me.
still, hot, perfect day! What a golden desert this spreading moor! E
I wished I could live in it and on it. I saw a lizard run over the crag; I
g the sweet bilberries. I would fain at the moment have become bee or liz
found fitting nutriment, permanent shelter here. But I was a human being,
ng's wants: I must not linger where there was nothing to supply them. I ros
e bed I had left. Hopeless of the future, I wished but this–that my Make
ght good to require my soul of me while I slept; and that this weary frame
om further conflict with fate, had now but to decay quietly, and mingle in
this wilderness. Life, however, was yet in my possession, with all its req
and responsibilities. The burden must be carried; the want provided for; th

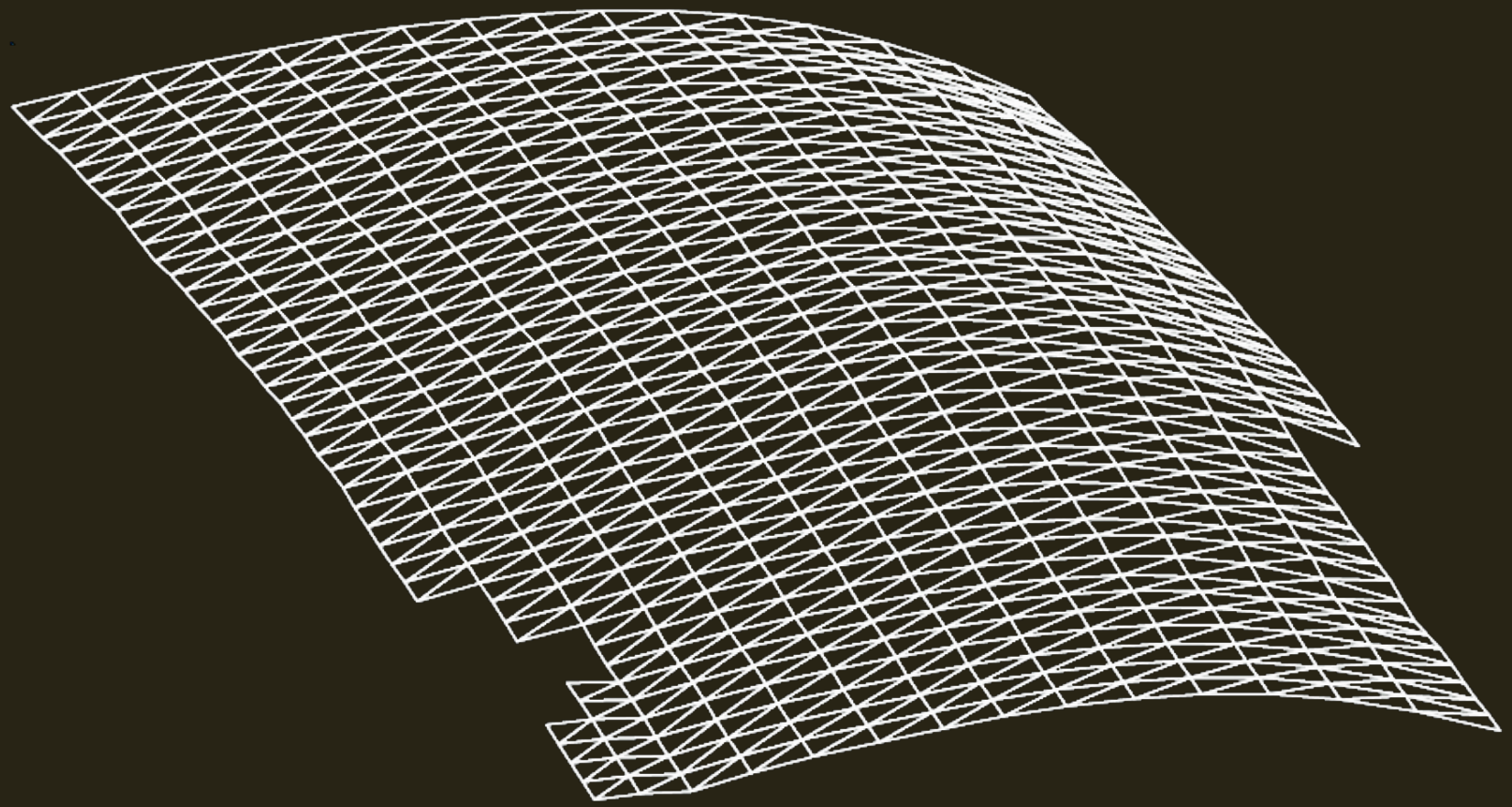


this torture of thought, I rose to my knee
still night: too serene for the companion
tainly we feel His presence most when H
d it is in the unclouded night-sky, where
t His infinitude, His omnipotence, His om
chester. Looking up, I, with tear-dimme
it was–what countless systems there swe
strength of God. Sure was I of His effi
t neither earth should perish, nor one of
ng: the Source of Life was also the Sav
nd by God would he be guarded. I aga
forgot sorrow.
nt came to me pale and bare. Long after

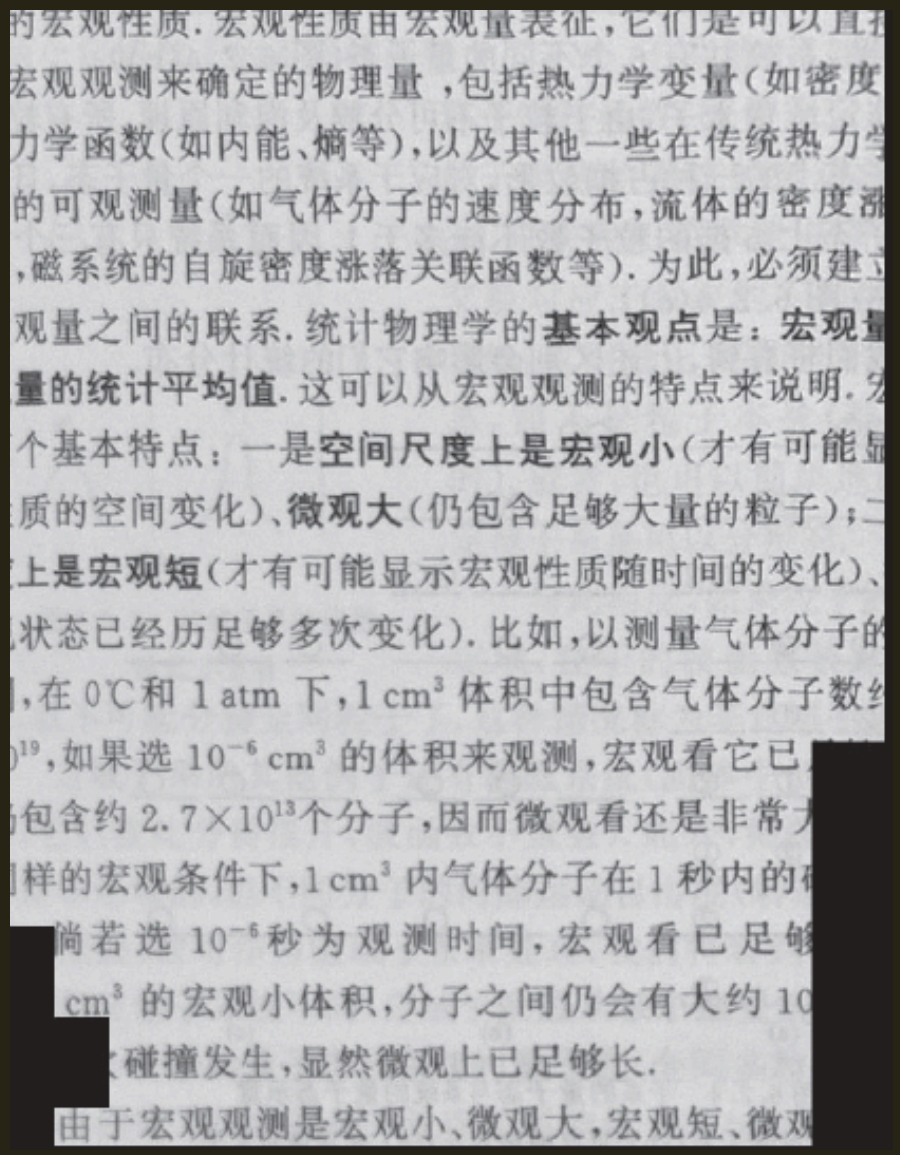
More results



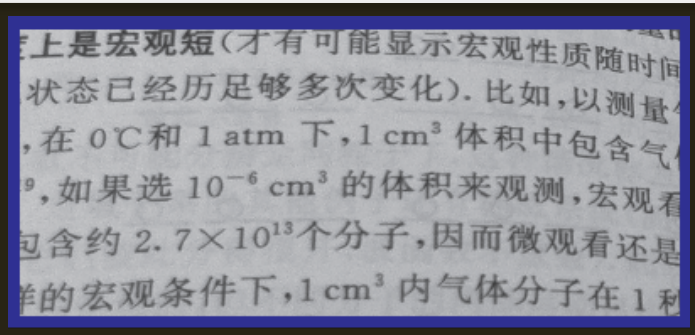
More results



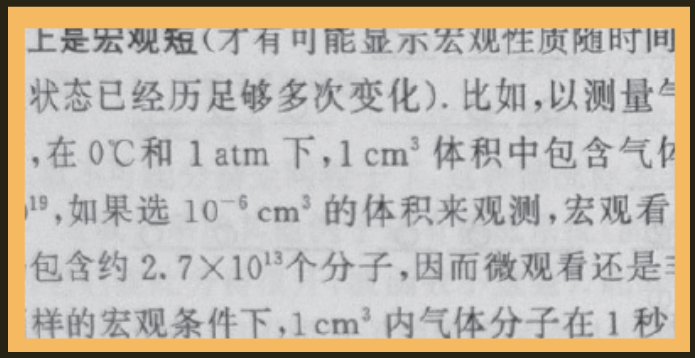
More results



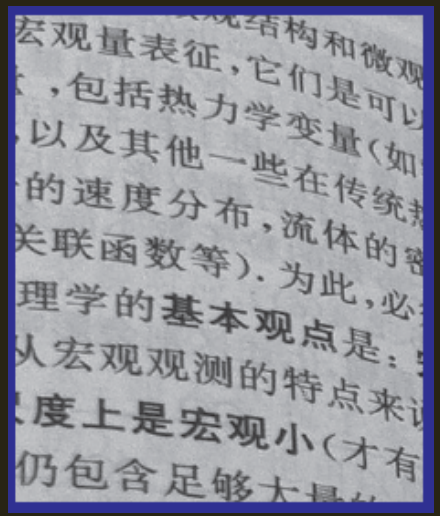
的宏观性质.宏观性质由宏观量表征,它们是可以直接
宏观观测来确定的物理量,包括热力学变量(如密度
力学函数(如内能、熵等),以及其他一些在传统热力学
的可观测量(如气体分子的速度分布,流体的密度涨
,磁系统的自旋密度涨落关联函数等).为此,必须建立
观量之间的联系.统计物理学的**基本观点**是:**宏观量**
量的统计平均值.这可以从宏观观测的特点来说明.宏
个基本特点:一是**空间尺度上是宏观小**(才有可能显
质的空间变化)、**微观大**(仍包含足够大量的粒子);二
上是宏观短(才有可能显示宏观性质随时间的变化)、
状态已经历足够多次变化).比如,以测量气体分子的
,在0℃和1 atm下,1 cm^3体积中包含气体分子数约
^{19},如果选10^{-6} cm^3的体积来观测,宏观看它已
包含约2.7×10^{13}个分子,因而微观看还是非常大
样的宏观条件下,1 cm^3内气体分子在1秒内的碰
倘若选10^{-6}秒为观测时间,宏观看已足够
cm^3的宏观小体积,分子之间仍会有大约10
碰撞发生,显然微观上已足够长.
由于宏观观测是宏观小、微观大,宏观短、微观



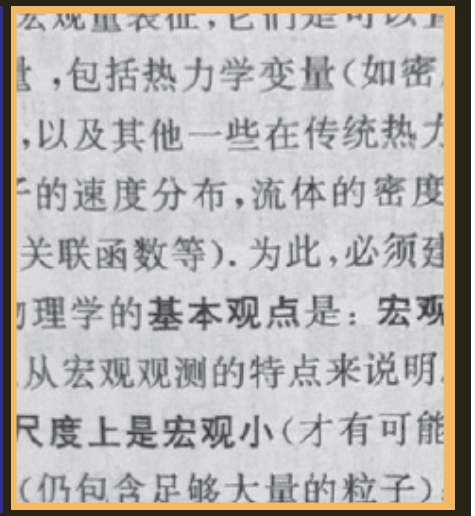
上是宏观短(才有可能显示宏观性质随时间
状态已经历足够多次变化).比如,以测量
,在0℃和1 atm下,1 cm^3体积中包含气
^{19},如果选10^{-6} cm^3的体积来观测,宏观看
包含约2.7×10^{13}个分子,因而微观看还是非常
样的宏观条件下,1 cm^3内气体分子在1秒



上是宏观短(才有可能显示宏观性质随时间
状态已经历足够多次变化).比如,以测量气
,在0℃和1 atm下,1 cm^3体积中包含气体
^{19},如果选10^{-6} cm^3的体积来观测,宏观看
包含约2.7×10^{13}个分子,因而微观看还是非常
样的宏观条件下,1 cm^3内气体分子在1秒

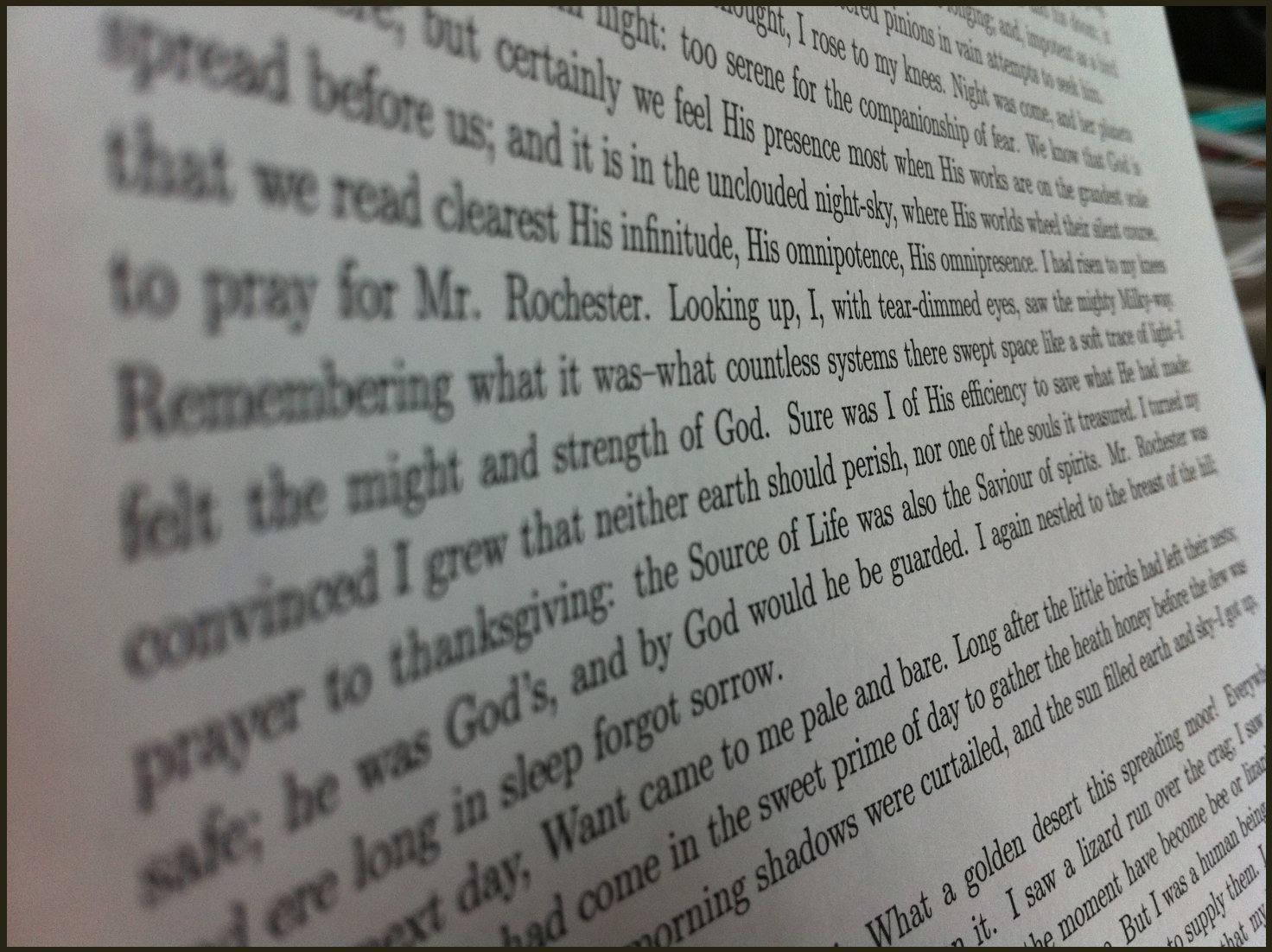


宏观结构和微观
宏观量表征,它们是可以
,包括热力学变量(如
以及其他一些在传统热
的速度分布,流体的密
关联函数等).为此,必
理学的**基本观点**是:
从宏观观测的特点来说
度上是宏观小(才有
仍包含足够大量

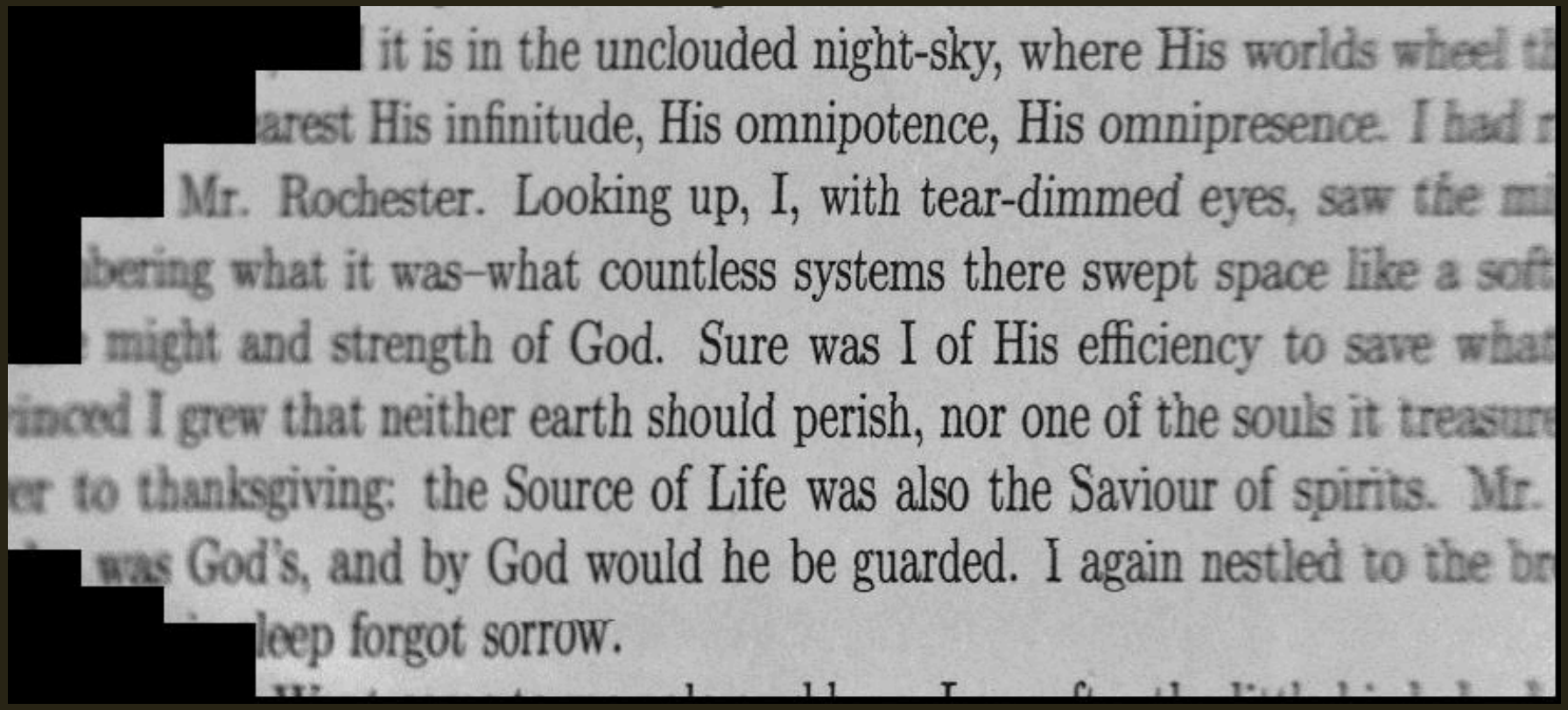


宏观量表征,它们是可以直
,包括热力学变量(如密
,以及其他一些在传统热力
的速度分布,流体的密度
关联函数等).为此,必须建
理学的**基本观点**是:**宏观**
从宏观观测的特点来说明
尺度上是宏观小(才有可能
(仍包含足够大量的粒子)

Did we achieve the goal?

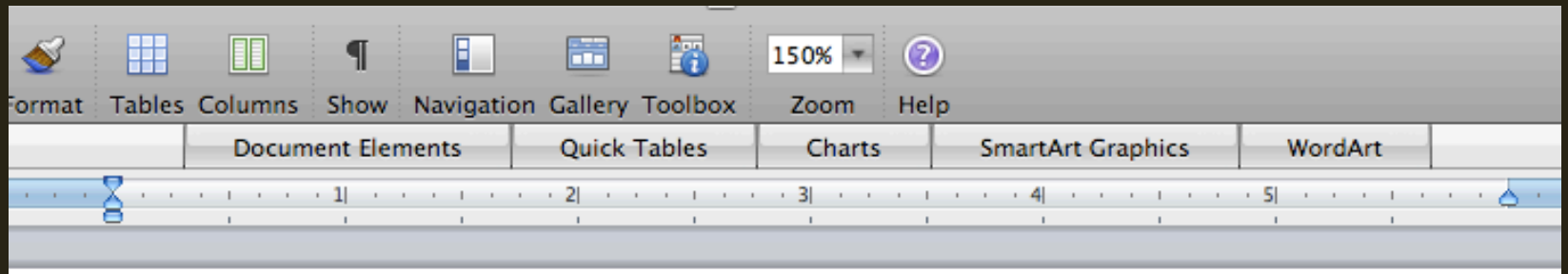


Did we achieve the goal?



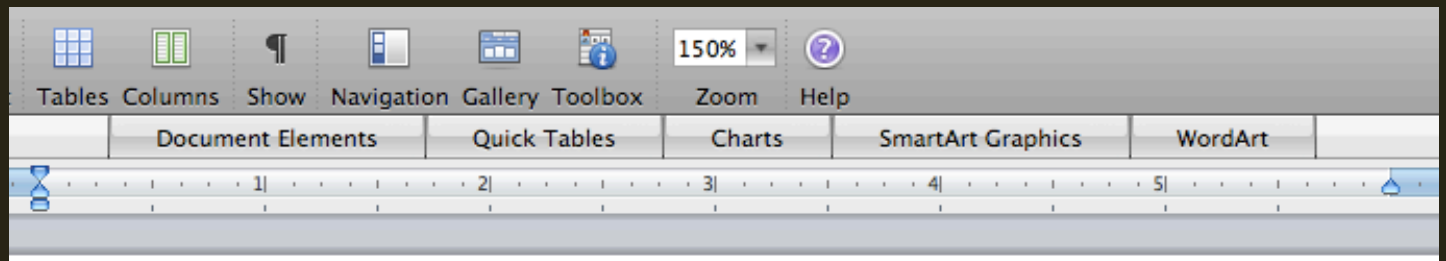
it is in the unclouded night-sky, where His worlds wheel th
arest His infinitude, His omnipotence, His omnipresence. I had n
Mr. Rochester. Looking up, I, with tear-dimmed eyes, saw the mi
bering what it was–what countless systems there swept space like a soft
might and strength of God. Sure was I of His efficiency to save what
inced I grew that neither earth should perish, nor one of the souls it treasure
er to thanksgiving: the Source of Life was also the Saviour of spirits. Mr.
was God's, and by God would he be guarded. I again nestled to the br
leep forgot sorrow.

Did we achieve the goal?



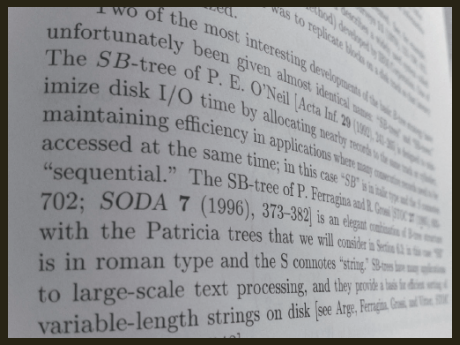
Fmight and strength of God. Sure was I of His efficiency to save Wh it is in the unclouded night-sky, where His wlarids whe6 r -, - est His infinitude, His omnipotence, His omnipresence. I hal . MI. Rochester. Looking up, I, with tear-dimmed eyes, saw the what it was-what countless systems there swept space like a - =ad I grew that neither earth should perish, nor one of the souls it rit to thanking: the Source of Life was also the Saviour of spiitts. Mr. 1 God's. and by God would he be guarded. I again nestled to the • Jeep forgot sorrow.

OCR Performance under distortion

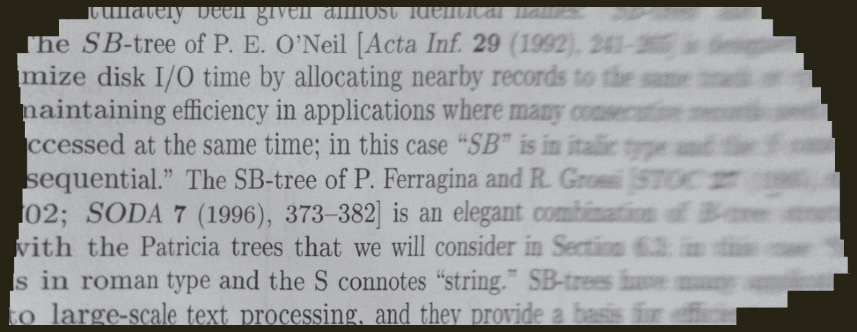


Ihe 'feel His preseixe most vltil Is In'ts lit Mt illityg t: _.4u %tit) I to;t:O'llz1;111\z'%1411\nit' k 0 Setelie fot tile comi)aAlots14 ikill It ' in the unclouded iittrk, '18AimIlls HiS infinitude, His oniipotelice, Is . '4411tiet:) orawott Vtiitostwasurt T. Looking up, I, w•lth, tem-IrciavA MS, SA it 4 /.11 at it was-what countless systems kloe 51 S\WAitt , It of God. Sure vas 1. () Els 6c,le'xi t ot itia ' the earth siloa pensiri, tOt Olt () SaIV't 1. la toibtoeyt sGoia:O e ve:0,\.iiee 1 0:e $::: t ::toe:t. . \ 16 ek I. Slo ac ;D:N v _le to It* Vt'\,ce:0 CP lesgee e c\° (), 'i wz '40 ..2146 ...'o o *110' 6, \"s \"" , ,0 ,,, ,,,, fi; ,§ZZC:%'CO\Q\ '\ V\ °) Z'S° 4 v- *i ti ill" 4,11:11At „Aro irer . i $izo, 4\\ to ,\\. \' A \k"' \.,■ttp,e1 0*
ktt

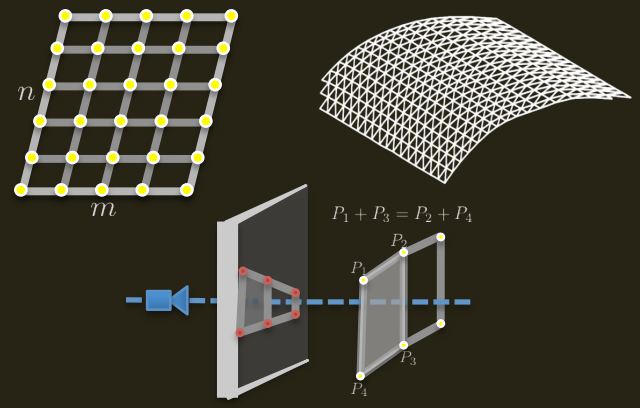
Conclusion



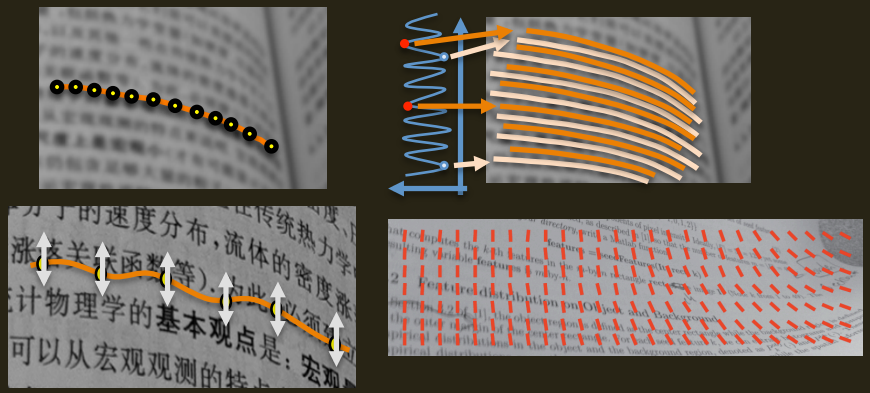
Rectification



Globally optimal 3D reconstruction

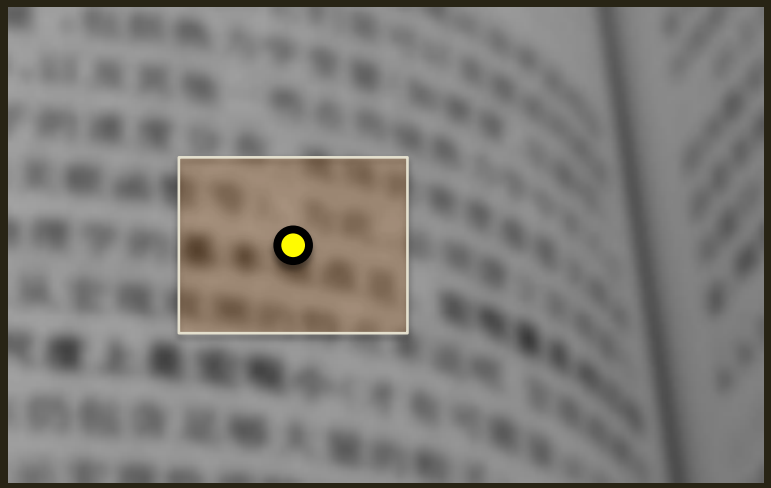


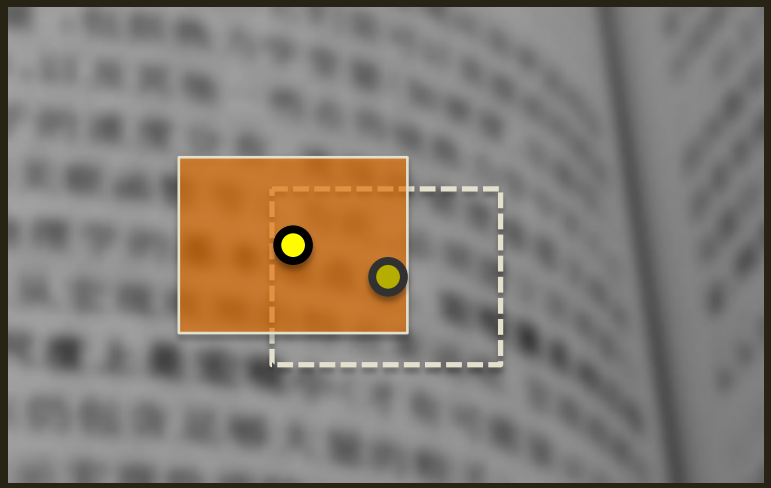
Text analysis on gray-scale image



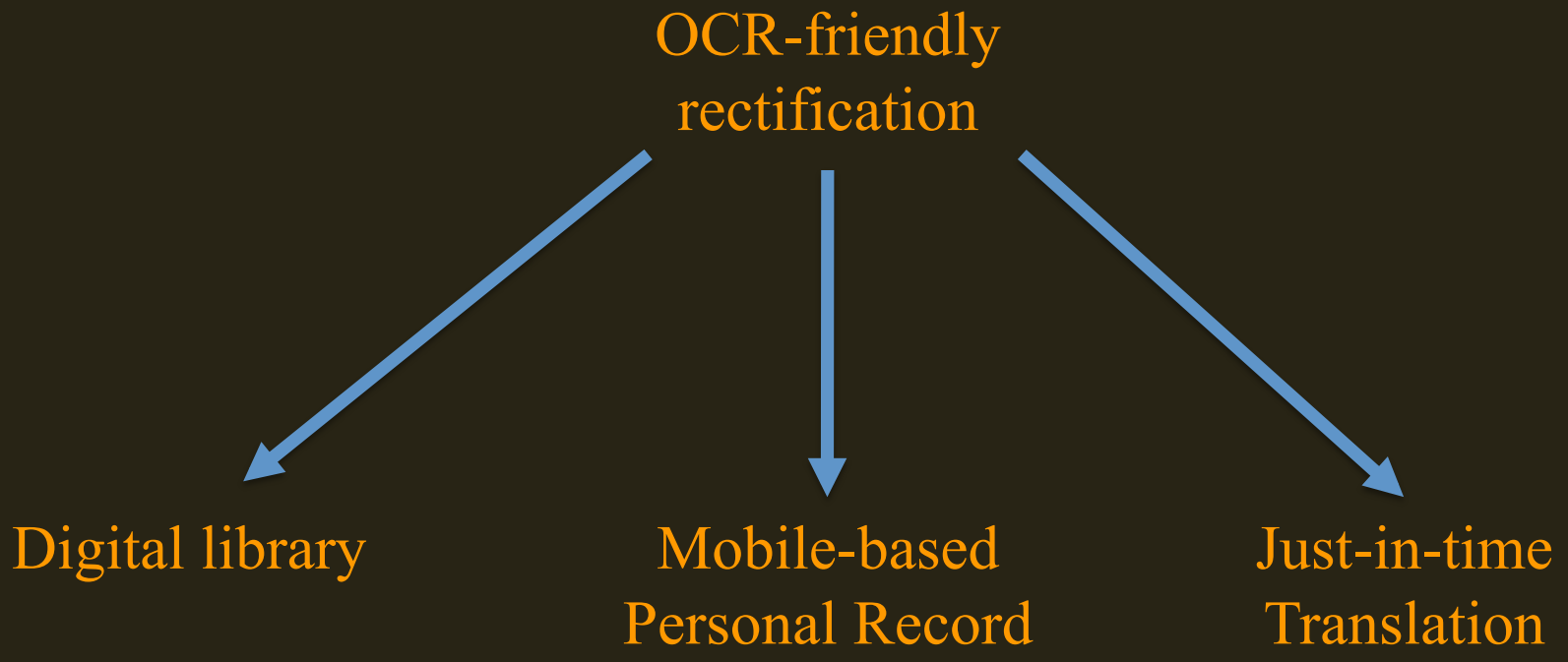
The Ultimate Goal

Thank you!





Motivation


OCR-friendly rectification
Digital library
Mobile-based Personal Record
Just-in-time Translation

Contribution

- Language-independent text analysis on gray-scale and low-resolution images
    - Previous works [CC, binarization]

- Reconstruction a broad range of 3D surface from a single image.
    - Previous works [Parametric forms, developable surface]
    - Previous works [Shape-from-X, complicated modeling, local solution.]

- Rectification of for-shortening and shading based on 3D surface.