

Computational Imaging VI

Charles A. Bouman
Eric L. Miller
Ilya Pollak
Editors

28–29 January 2008
San Jose, California, USA

Sponsored and Published by
IS&T—The Society for Imaging Science and Technology
SPIE

Cosponsored by
GE Healthcare (USA)

Volume 6814

The papers included in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. The papers published in these proceedings reflect the work and thoughts of the authors and are published herein as submitted. The publishers are not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from this book:

Author(s), "Title of Paper," in *Computational Imaging VI*, edited by Charles A. Bouman, Eric L. Miller, Ilya Pollak, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 6814, Article CID Number (2008).

ISSN 0277-786X

ISBN 9780819469861

Copublished by

SPIE

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time) · Fax +1 360 647 1445

SPIE.org

and

IS&T—The Society for Imaging Science and Technology

7003 Kilworth Lane, Springfield, Virginia, 22151 USA

Telephone +1 703 642 9090 (Eastern Time) · Fax +1 703 642 9094

imaging.org

Copyright © 2008, Society of Photo-Optical Instrumentation Engineers and The Society for Imaging Science and Technology.

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by the publishers subject to payment of copying fees. The Transactional Reporting Service base fee for this volume is \$18.00 per article (or portion thereof), which should be paid directly to the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Payment may also be made electronically through CCC Online at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher. The CCC fee code is 0277-786X/08/\$18.00.

Printed in the United States of America.

Paper Numbering: Proceedings of SPIE follow an e-First publication model, with papers published first online and then in print and on CD-ROM. Papers are published as they are submitted and meet publication criteria. A unique, consistent, permanent citation identifier (CID) number is assigned to each article at the time of the first publication. Utilization of CIDs allows articles to be fully citable as soon they are published online, and connects the same identifier to all online, print, and electronic versions of the publication. SPIE uses a six-digit CID article numbering system in which:

- The first four digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B ... 0Z, followed by 10-1Z, 20-2Z, etc.

The CID number appears on each page of the manuscript. The complete citation is used on the first page, and an abbreviated version on subsequent pages. Numbers in the index correspond to the last two digits of the six-digit CID number.

Contents

- vii *Conference Committee*
- ix *Stationary features and cat detection (Keynote Presentation Viewgraphs) [6814-21]*
D. Geman, Johns Hopkins Univ. (USA); F. Fleuret, INRIA Rocquencourt (France)

KEYNOTE PRESENTATION

- 6814 02 **Fast acquisition and reconstruction in imaging enabled by sampling theory (Invited Paper)** [6814-01]
Y. Bresler, Univ. of Illinois at Urbana-Champaign (USA)

IMAGE RECONSTRUCTION I

- 6814 03 **Regularized estimation of Stokes images from polarimetric measurements** [6814-02]
J. R. Valenzuela, J. A. Fessler, Univ. of Michigan (USA)
- 6814 04 **Non-homogeneous ICD optimization for targeted reconstruction of volumetric CT** [6814-03]
Z. Yu, Purdue Univ. (USA); J.-B. Thibault, GE Healthcare Technologies (USA); C. A. Bouman, Purdue Univ. (USA); K. D. Sauer, Univ. of Notre Dame (USA); J. Hsieh, GE Healthcare Technologies (USA)

GEOMETRY-BASED TECHNIQUES IN IMAGE ANALYSIS

- 6814 07 **MCMC curve sampling and geometric conditional simulation** [6814-06]
A. Fan, J. W. Fisher III, Massachusetts Institute of Technology (USA); J. Kane, Shell E&P (USA); A. S. Willsky, Massachusetts Institute of Technology (USA)
- 6814 08 **Mixing geometric and radiometric features for change classification** [6814-07]
A. Fournier, X. Descombes, J. Zerubia, INRIA (France)
- 6814 09 **3D object recognition using fully intrinsic skeletal graphs** [6814-08]
D. Aouada, H. Krim, North Carolina State Univ. (USA)
- 6814 0A **Content based image retrieval for matching images of improvised explosive devices in which snake initialization is viewed as an inverse problem** [6814-09]
S. T. Acton, A. D. Gilliam, B. Li, Univ. of Virginia (USA); A. Rossi, Platinum Solutions (USA)

SEGMENTATION

- 6814 0B **Segmentation of digital microscopy data for the analysis of defect structures in materials using nonlinear diffusions** [6814-10]
L. M. Huffman, Purdue Univ. (USA); J. Simmons, Air Force Research Lab. (USA); I. Pollak, Purdue Univ. (USA)

- 6814 0C **A new approach for joint estimation of local magnitude, decay, and frequency from single-shot MRI** [6814-11]
W. Tang, S. J. Reeves, Auburn Univ. (USA); D. B. Twieg, Univ. of Alabama at Birmingham (USA)
- 6814 0D **A novel image analysis method based on Bayesian segmentation for event-related functional MRI** [6814-12]
L. Huang, M. L. Comer, Purdue Univ. (USA); T. M. Talavage, Purdue Univ. (USA) and Indiana Univ. School of Medicine (USA)
- 6814 0E **Volumetric fMRI data analysis using an iterative classification method** [6814-13]
L. Liu, K. Han, Purdue Univ. (USA); T. M. Talavage, Purdue Univ. (USA) and Indiana Univ. School of Medicine (USA)
- 6814 0F **Rule-based fuzzy vector median filters for 3D phase contrast MRI segmentation** [6814-14]
K. S. Sundareswaran, Georgia Institute of Technology (USA); D. H. Frakes, 4D Imaging Inc. (USA); A. P. Yoganathan, Georgia Institute of Technology (USA)

SPARSE RECOVERY AND COMPRESSED SENSING

- 6814 0J **Greedy signal recovery and uncertainty principles** [6814-18]
D. Needell, R. Vershynin, Univ. of California, Davis (USA)
- 6814 0K **Blind reconstruction of sparse images with unknown point spread function** [6814-19]
K. Herrity, Univ. of Michigan (USA); R. Raich, Oregon State Univ. (USA); A. O. Hero III, Univ. of Michigan (USA)
- 6814 0L **Results in non-iterative MAP reconstruction for optical tomography** [6814-40]
G. Cao, C. A. Bouman, K. J. Webb, Purdue Univ. (USA)

IMAGE ANALYSIS I

- 6814 0P **A generalization of non-local means via kernel regression** [6814-23]
P. Chatterjee, P. Milanfar, Univ. of California, Santa Cruz (USA)
- 6814 0Q **Functional minimization problems in image processing** [6814-24]
Y. Kim, L. A. Vese, Univ. of California, Los Angeles (USA)
- 6814 0R **Determining canonical views of 3D object using minimum description length criterion and compressive sensing method** [6814-25]
P.-F. Chen, H. Krim, North Carolina State Univ. (USA)

IMAGE RECONSTRUCTION II

- 6814 0U **A least squares approach to estimating the probability distribution of unobserved data in multiphoton microscopy** [6814-27]
P. Salama, Indiana University (USA)

- 6814 0V **Progress in mesh based spatio-temporal reconstruction** [6814-28]
J. G. Brankov, R. Delgado Gonzalo, Y. Yang, M. Jin, M. N. Wernick, Illinois Institute of Technology (USA)
- 6814 0W **Diode laser absorption tomography using data compression techniques** [6814-29]
C. Lindstrom, Air Force Research Lab. (USA); C.-J. Tam, Taitech, Inc. (USA); R. Givens, Air Force Institute of Technology (USA); D. Davis, S. Williams, Air Force Research Lab. (USA)
- 6814 0X **3D macromolecule structure reconstruction from electron micrograph by exploiting symmetry and sparsity** [6814-30]
M. W. Kim, J. Choi, J. C. Ye, Korea Advanced Institute of Science and Technology (South Korea)
- 6814 0Y **Alternating minimization algorithm for quantitative differential-interference contrast (DIC) microscopy** [6814-42]
J. A. O'Sullivan, Washington Univ. in St. Louis (USA); C. Preza, The Univ. of Memphis (USA)

IMAGE ANALYSIS II

- 6814 0Z **Image filter effectiveness characterization based on HVS** [6814-31]
V. V. Lukin, N. N. Ponomarenko, S. S. Krivenko, National Aerospace Univ. (Ukraine); K. O. Egiazarian, J. T. Astola, Tampere Univ. of Technology (Finland)
- 6814 10 **Multimodal unbiased image matching via mutual information** [6814-32]
I. Yanovsky, Univ. of California, Los Angeles (USA); P. M. Thompson, UCLA School of Medicine (USA); S. J. Osher, Univ. of California, Los Angeles (USA); A. D. Leow, UCLA School of Medicine (USA)
- 6814 11 **Technology-assisted dietary assessment** [6814-33]
F. Zhu, A. Mariappan, C. J. Boushey, Purdue Univ. (USA); D. Kerr, Curtin Institute of Technology (Australia); K. D. Lutes, D. S. Ebert, E. J. Delp, Purdue Univ. (USA)
- 6814 12 **New methods for fMRI data processing based on locally linear embeddings** [6814-34]
L. Chaillou, I. S. Yetik, M. N. Wernick, Illinois Institute of Technology (USA)
- 6814 13 **Online consistency checking for AM-FM target tracks** [6814-35]
N. A. Mould, C. T. Nguyen, C. M. Johnston, J. P. Havlicek, Univ. of Oklahoma (USA)

POSTER SESSION

- 6814 14 **Image mosaicking based on feature points using color-invariant values** [6814-36]
D.-C. Lee, O.-S. Kwon, K.-W. Ko, Kyungpook National Univ. (South Korea); H.-Y. Lee, Samsung Advanced Institute of Technology (South Korea); Y.-H. Ha, Kyungpook National Univ. (South Korea)
- 6814 15 **Inverse perspective transformation for video surveillance** [6814-37]
T. E. Schouten, Radboud Univ. Nijmegen (Netherlands); E. L. van den Broek, Univ. of Twente (Netherlands)

6814 16 **Illumination normalization for face recognition using the census transform** [6814-38]
J. H. Kim, J. G. Park, C. Lee, Yonsei Univ. (South Korea)

6814 17 **Audio-video synchronization management in embedded multimedia applications**
[6814-41]
H.-U. Rehman, The Univ. of Texas at Austin (USA); T. Kim, Freescale Semiconductor, Inc.
(USA); N. Avadhanam, Novafora Inc. (USA); S. Subramanian, Freescale Semiconductor, Inc.
(USA)

Author Index

Conference Committee

Symposium Chair

Nitin Sampat, Rochester Institute of Technology (USA)

Conference Chairs

Charles A. Bouman, Purdue University (USA)

Eric L. Miller, Tufts University (USA)

Ilya Pollak, Purdue University (USA)

Program Committee

Samit Basu, GE Global Research (USA)

Thomas S. Denney, Auburn University (USA)

Peter C. Doerschuk, Purdue University (USA)

Peyman Milanfar, University of California, Santa Cruz (USA)

Joseph A. O'Sullivan, Washington University in St. Louis (USA)

Zygmunt Pizlo, Purdue University (USA)

Stanley J. Reeves, Auburn University (USA)

Yongyi Yang, Illinois Institute of Technology (USA)

Session Chairs

Keynote Presentation

Charles A. Bouman, Purdue University (USA)

Image Reconstruction I

Eric L. Miller, Tufts University (USA)

Geometry-based Techniques in Image Analysis

Mary L. Comer, Purdue University (USA)

Segmentation

Charles A. Bouman, Purdue University (USA)

Sparse Recovery and Compressed Sensing

Ilya Pollak, Purdue University (USA)

Image Analysis I

Paul Salama, Indiana University-Purdue University at Indianapolis (USA)

Image Reconstruction II

Edward J. Delp, Purdue University (USA)

Image Analysis II

Mireille Boutin, Purdue University (USA)

INTRODUCTION

DETECTION



STATIONARY FEATURES AND CAT DETECTION

DONALD GEMAN

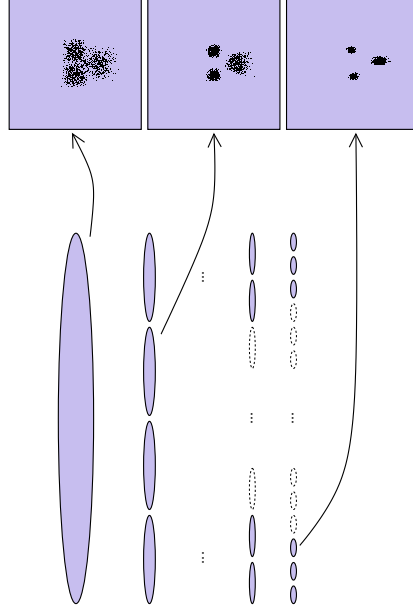
JOINT WORK WITH
FRANCOIS FLEURET

COMPUTATIONAL IMAGING VI

JANUARY 29, 2008

INTRODUCTION

DETECTION (CONT.)



HIERARCHY OF CLASSIFIERS

Advantages

- Highly efficient scene parsing;
- Organized focusing on hard examples.

Disadvantages

- Many classifiers to learn;
- Inefficient learning (unless training examples can be synthesized).

x

INTRODUCTION

RELATED WORK

Part-based Models

- Constellation (M. Weber, M. Welling, P. Perona),
- Composite (D. J. Crandall, D. P. Huttenlocher),
- Implicit Shape (E. Seemann, M. Fritz, B. Schiele),
- Patchwork of Parts (Y. Amit, A. Trouvé),
- Compositional (S. Geman).

Wholistic

- Convolutional Networks (Y. Lecun),
- Boosted Cascades (P. Viola, M. Jones),
- Bag-of-Features (A. Zisserman, FFLi).

5 / 36

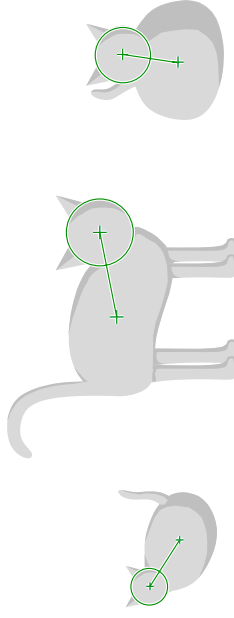
INTRODUCTION

POSE SPACE (CONT.)

For faces, typically $y = (u_c, v_c, \theta, s)$. Cats are more complex.



A coarse description is $y = (u_h, v_h, s_h, u_b, v_b)$.



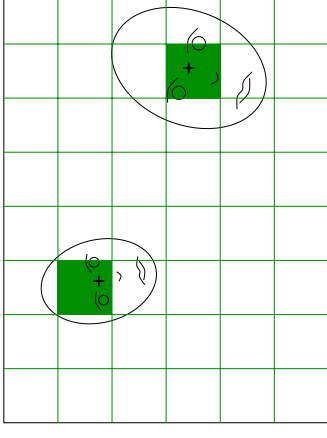
7 / 36

INTRODUCTION

POSE SPACE

Let \mathcal{Y} be the space of poses and $\mathcal{Y}_1, \dots, \mathcal{Y}_K$ a partition.

Let I be an image and $\mathbf{Y} = (Y_1, \dots, Y_K)$, where, for each k , Y_k is a Boolean variable stating if there is an object in I with pose in \mathcal{Y}_k .



6 / 36

TRAINING A DETECTOR

FRAGMENTATION

Given a training set

$$\left\{ \left(I^{(t)}, \mathbf{Y}^{(t)} \right) \right\}_t$$

we are interested in building, for each k , a classifier

$$f_k : \mathcal{I} \rightarrow \{0, 1\}$$

for predicting Y_k .

Without additional knowledge about the relationship between k , Y_k and I , we would train f_k with

$$\left\{ \left(I^{(t)}, Y_k^{(t)} \right) \right\}_t$$

Hence, each f_k is trained with a fragment ($\approx 1/K$) of the positive population.

8 / 36

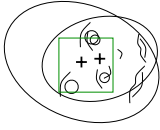
TRAINING A DETECTOR

AGGREGATION

To avoid fragmentation, samples are often normalized in pose. Let

$$\xi : \mathcal{I} \rightarrow \mathbb{R}^N$$

denote an N -dimensional vector of features. Let $\psi : \{1, \dots, K\} \times \mathcal{I} \rightarrow \mathcal{I}$ be a transformation such that the conditional distribution of $\xi(\psi(k, l))$ given $Y_k = 1$ does not depend on k .



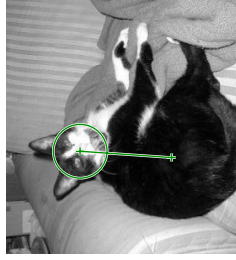
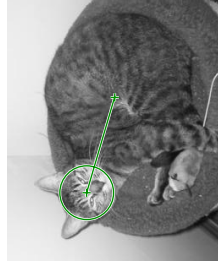
9 / 36

TRAINING A DETECTOR

AGGREGATION (CONT.)

However:

- Evaluating ψ is computationally intensive for any non-trivial transformation.
- The mapping ψ does not exist for a complex pose.



Hence, in practice, fragmentation is still the norm to deal with many deformations. But how does one deal with complex poses?

11 / 36

TRAINING A DETECTOR

AGGREGATION (CONT.)

Then train a single classifier g with the training set

$$\left\{ \left(\psi(k, l^{(t)}), Y_k^{(t)} \right) \right\}_{t,k}$$

and define

$$f_k(l) = g(\psi(k, l)).$$

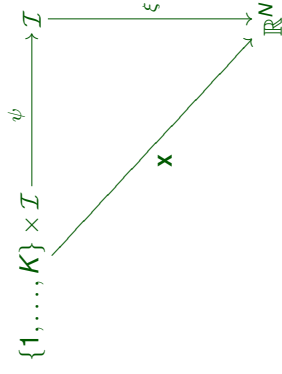
Here, samples are aggregated for training: all positive samples are used to build each f_k .

10 / 36

STATIONARY FEATURES

INTRODUCTION

We directly index the features with the pose.



12 / 36

STATIONARY FEATURES DEFINITION

Hence, we define a family of pose-indexed features as a mapping

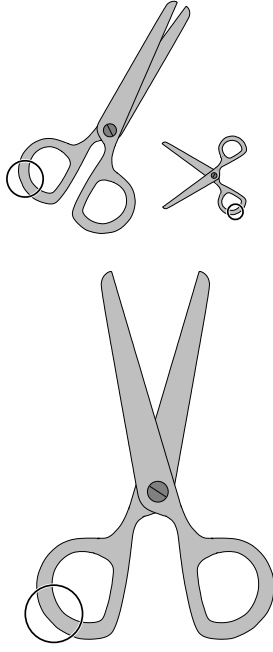
$$\mathbf{X} : \{1, \dots, K\} \times \mathcal{I} \rightarrow \mathbb{R}^N$$

such that they are **stationary** in the following sense: for every $k \in \{1, \dots, K\}$, the probability distribution

$$P(\mathbf{X}(k, l) = \mathbf{x} \mid Y_k = 1), \mathbf{x} \in \mathbb{R}^N$$

does not depend on k .

STATIONARY FEATURES DEFINITION (CONT.)



STATIONARY FEATURES INTERPRETATION

Suppose there is exactly one object in I ($\sum_k Y_k = 1$) and let $Z = Z(\mathbf{Y}) \in \{1, \dots, K\}$ be the corresponding pose cell. Then stationarity means that

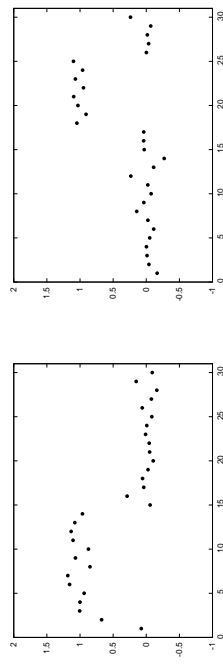
$$\mathbf{X}(Z) \perp Z$$

In other words, knowing $\mathbf{X}(Z)$, the feature values, tells you nothing about Z , the pose.

STATIONARY FEATURES TOY EXAMPLE, 1D SIGNAL

$$\mathcal{Y} = \{(\theta_1, \theta_2) \in \{1, \dots, N\}^2, 1 < \theta_1 < \theta_2 < N\}$$

$$P(I \mid \mathbf{Y} = (\theta_1, \theta_2)) = \prod_{n < \theta_1} \phi_0(I_n) \prod_{\theta_1 \leq n \leq \theta_2} \phi_1(I_n) \prod_{\theta_2 < n} \phi_0(I_n)$$



$$\mathbf{X}((\theta_1, \theta_2), I) = (I(\theta_1 - 1), I(\theta_1), I(\theta_2), I(\theta_2 + 1))$$

$$P(\mathbf{X} = (x_1, x_2, x_3, x_4) \mid \mathbf{Y} = (\theta_1, \theta_2)) = \phi_0(x_1)\phi_1(x_2)\phi_1(x_3)\phi_0(x_4)$$

STATIONARY FEATURES TRAINING

We then define

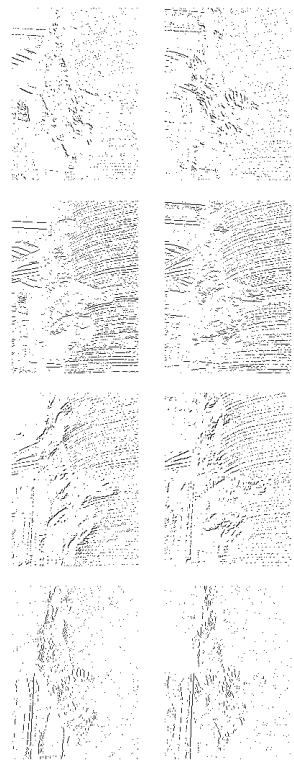
$$f_k(l) = g(\mathbf{X}(k, l)), \quad k \in \{1, \dots, K\},$$

where one single classifier $g: \mathbb{R}^N \rightarrow \{0, 1\}$ is trained with

$$\left\{ \left(\mathbf{x}(k, l^{(t)}), Y_k^{(t)} \right) \right\}_{t,k}.$$

Notice that stationarity is the main condition required to ensure that the training samples are identically distributed.

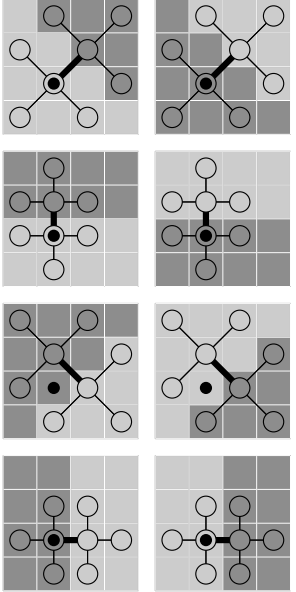
CAT DETECTION EDGE DETECTORS (CONT.)



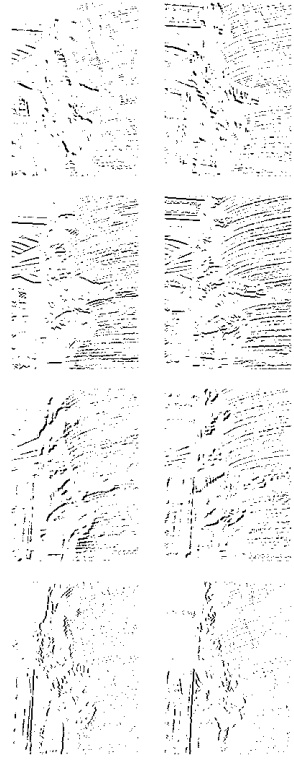
$\sigma = 1$

CAT DETECTION EDGE DETECTORS

Let $e_{\phi, \sigma}(l, u, v)$ be the presence on an edge in image l at location (u, v) with orientation $\phi \in \{0, \dots, 7\}$ at scale $\sigma \in \{1, 2, 4\}$.



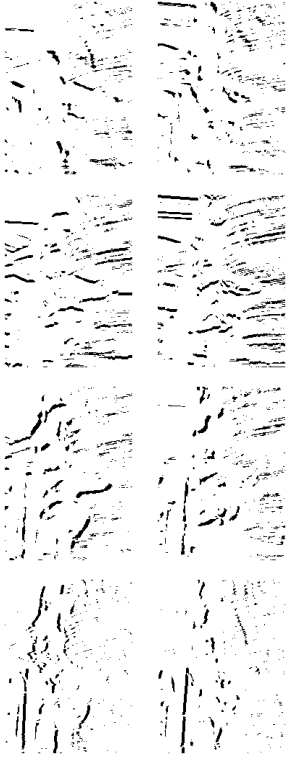
CAT DETECTION EDGE DETECTORS (CONT.)



$\sigma = 2$

CAT DETECTION EDGE DETECTORS (CONT.)

xiv

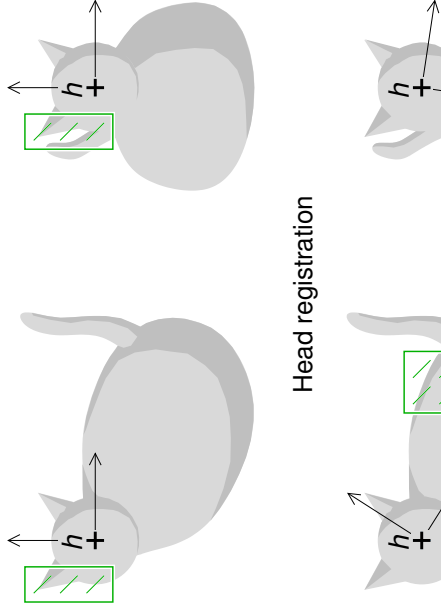


$\sigma = 4$

21 / 36

CAT DETECTION STATIONARY FEATURES

The windows W and W' are indexed either with respect to the head or with respect to the head-belly.



23 / 36

CAT DETECTION BASE FEATURES

We use three types of stationary features:

- Proportion of an edge (ϕ, σ) in a window W .
- L^1 -distance between the histograms of orientations at scale σ in two windows W and W' .
- L^1 -distance between the histograms of gray-levels in two windows W and W' .

22 / 36

CAT DETECTION CLASSIFIER

We build classifiers with Adaboost and an asymmetric weighting by sampling. If the weighted error rate is

$$L(h) = \sum_{t,k} \omega_{t,k} \mathbb{1}_{\{h(k, t^{(t)}) \neq z_k^{(t)}\}},$$

we use all the positive samples, and sub-sample negative ones with

$$\mu(k, t) \propto \omega_{t,k} \mathbb{1}_{\{y_k^{(t)}=0\}}.$$

Total of 2327 scenes containing 1683 cats, 85% for training.

24 / 36

FOLDED HIERARCHIES SUMMARY

Experiments (not shown) demonstrate:

- Fragmentation applied to complex poses is disastrous in practice;
- Naive, brute-force exploration of the pose space is computationally intractable.

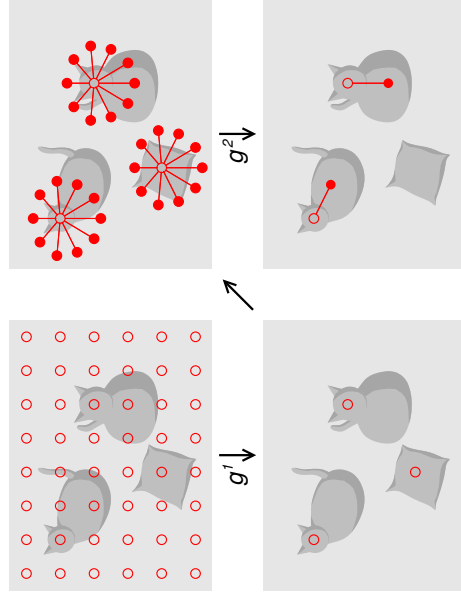
Alternatively:

- Stationary features largely avoid fragmentation;
- Hierarchical search concentrates computation on ambiguous regions.

We call this alternative a *folded hierarchy of classifiers*.

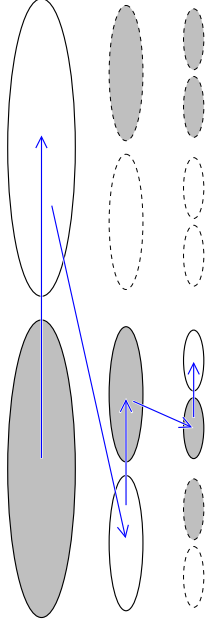
25 / 36

SCENE PARSING STRATEGY



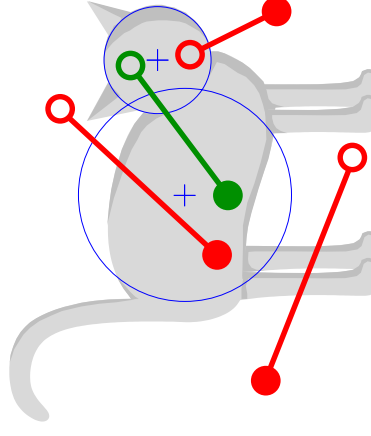
27 / 36

FOLDED HIERARCHIES SUMMARY (CONT.)



26 / 36

SCENE PARSING ERROR CRITERION



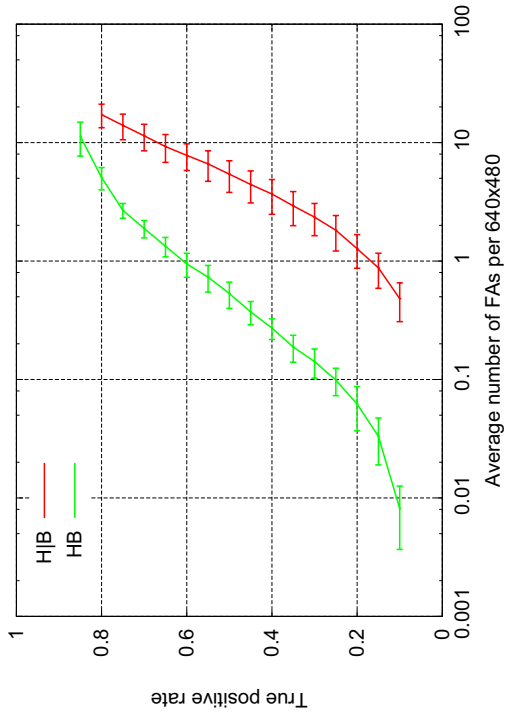
28 / 36

xvi **SCENE PARSING**
ERROR RATES

TP	H B	HB
85%	-	11.29 (3.61)
80%	17.23 (3.87)	5.07 (1.08)
70%	11.40 (2.89)	1.88 (0.32)
60%	7.80 (1.98)	0.95 (0.22)
50%	5.41 (1.62)	0.53 (0.13)

Average number of FAs per 640×480

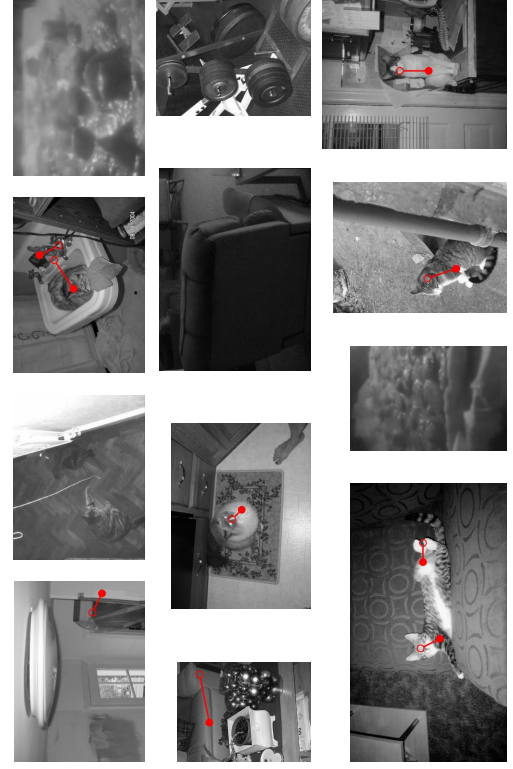
SCENE PARSING
ERROR RATES (CONT.)



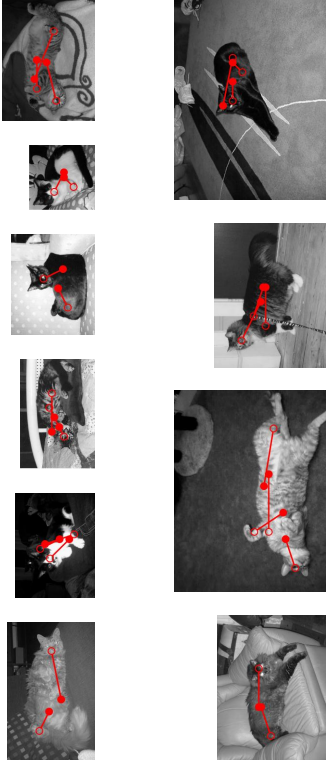
SCENE PARSING
RESULTS (PICKED AT RANDOM)



SCENE PARSING
RESULTS (PICKED AT RANDOM, CONT.)



SCENE PARSING RESULTS (SELECTED FALSE ALARMS)



33 / 36

CONCLUSION

A folded hierarchy of classifiers is highly efficient:

- For (offline) learning;
- For (online) scene processing.

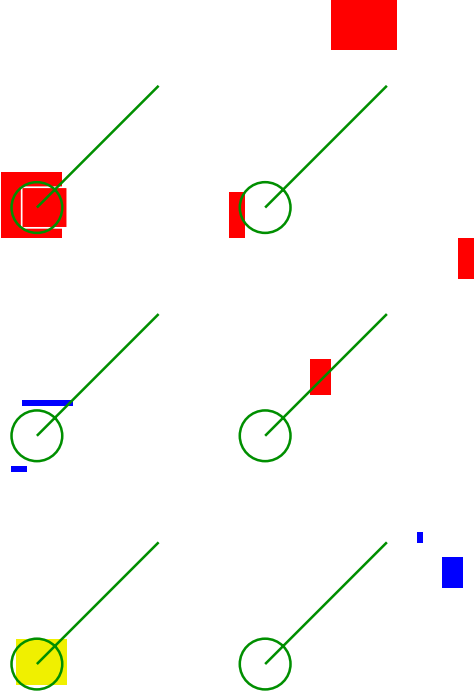
It combines the strengths of:

- Template matching;
- Powerful, wholistic machine learning;
- Hierarchical search.

However, this is achieved at the expense of

- Rich annotation of the training samples.
- Designing stationary features.

SCENE PARSING SELECTED STATIONARY FEATURES



34 / 36

ACKNOWLEDGMENT



35 / 36

36 / 36

