Spatio-Spectral Background Estimation in Remote Sensing Imagery

James Theiler and Brendt Wohlberg

Los Alamos National Laboratory

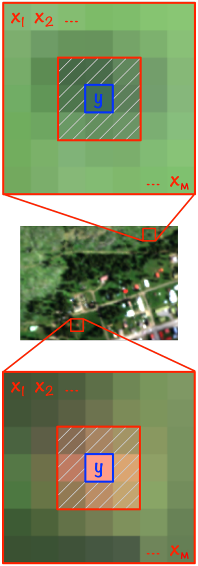
26 May 2016

Acknowledge: Stanley Rotman, Amanda Ziemann

Research supported by the United States Department of Energy
Hyperspectral Advanced Research and Development
for Solid materials project (HARD Solids).

Los Alamos
NATIONAL LABORATORY

Motivation: detect weak signals in cluttered backgrounds



Hunts Needle in a Haystack

How long does it take to find a needle in a haystack? Jim Moran, Washington, D. C., publicity man, recently dropped a needle into a convenient pile of hay, hopped in after it, and began an intensive search for (a) some publicity and (b) the needle. Having found the former, Moran abandoned the needle hunt.

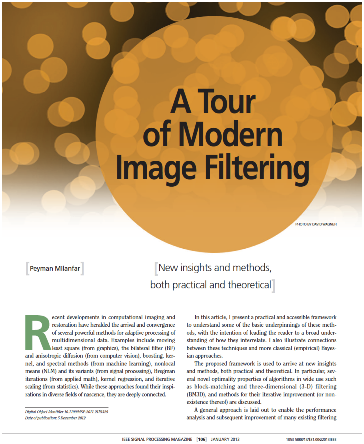
Motivation, Part II

- To solve **Remote Sensing** problems
  - To detect targets, first estimate the target-free background
- Can we use **Image Processing** tools?
  - To reduce noise, need to estimate noise-free pixel values
  - In-painting, need to estimate missing pixel values

Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using “non-local means”

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

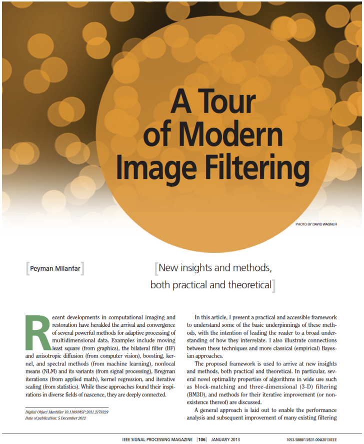
Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

IEEE SIGNAL PROCESSING MAGAZINE [106] JANUARY 2013 1053-5888/13/$31.00©2013IEEE

Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using “non-local means”

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

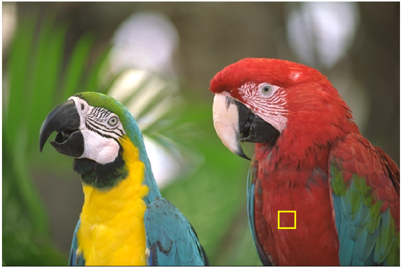
In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

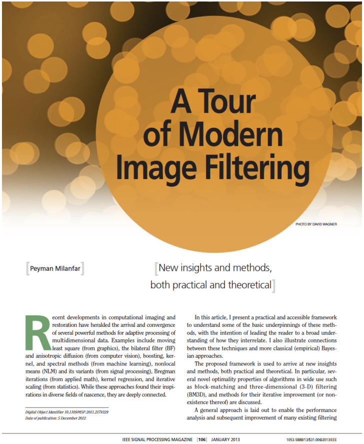
IEEE SIGNAL PROCESSING MAGAZINE [106] JANUARY 2013 1053-5888/13/$31.00©2013IEEE



Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using “non-local means”

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

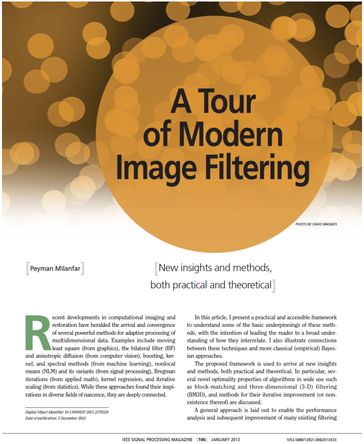
A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using "non-local means"

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

IEEE SIGNAL PROCESSING MAGAZINE [106] JANUARY 2013 1053-5888/13/$31.00©2013IEEE

Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using "non-local means"

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128

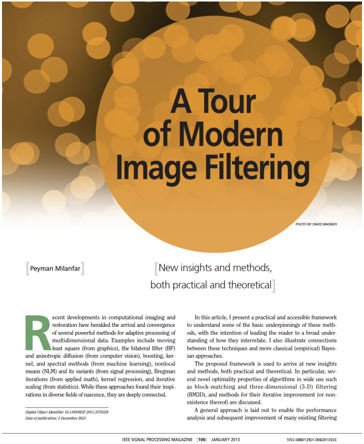

A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

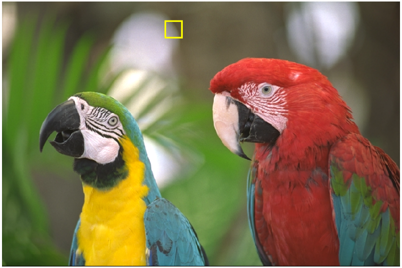
In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

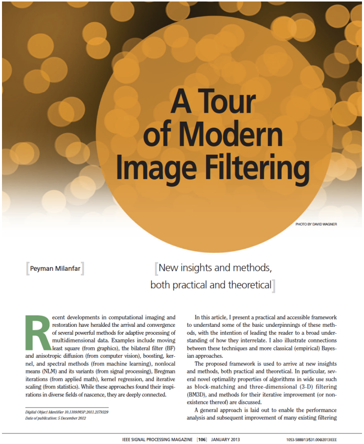
IEEE SIGNAL PROCESSING MAGAZINE [106] JANUARY 2013 1053-5888/13/$31.00©2013IEEE



Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using "non-local means"

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

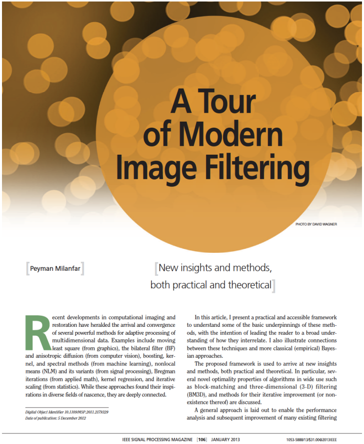
Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

IEEE SIGNAL PROCESSING MAGAZINE [106] JANUARY 2013 1053-5888/13/$31.00©2013IEEE

Inspired by recent developments in image processing

- Patched-based methods: filtering, regression, etc.
- eg, noise reduction using “non-local means”

P. Milanfar, IEEE Signal Processing Magazine (Jan 2013), 106–128



A Tour of Modern Image Filtering

PHOTO BY DAVID WAGNER

[Peyman Milanfar]

[New insights and methods, both practical and theoretical]

Recent developments in computational imaging and restoration have heralded the arrival and convergence of several powerful methods for adaptive processing of multidimensional data. Examples include moving least square (from graphics), the bilateral filter (BF) and anisotropic diffusion (from computer vision), boosting, kernel, and spectral methods (from machine learning), nonlocal means (NLM) and its variants (from signal processing), Bregman iterations (from applied math), kernel regression, and iterative scaling (from statistics). While these approaches found their inspirations in diverse fields of nascence, they are deeply connected.

In this article, I present a practical and accessible framework to understand some of the basic underpinnings of these methods, with the intention of leading the reader to a broad understanding of how they interrelate. I also illustrate connections between these techniques and more classical (empirical) Bayesian approaches.

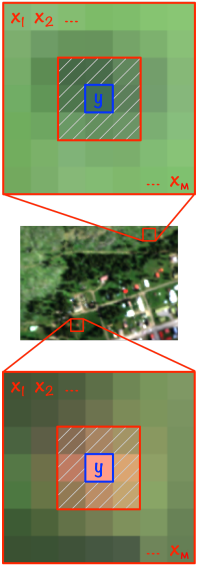
The proposed framework is used to arrive at new insights and methods, both practical and theoretical. In particular, several novel optimality properties of algorithms in wide use such as block-matching and three-dimensional (3-D) filtering (BM3D), and methods for their iterative improvement (or nonexistence thereof) are discussed.

A general approach is laid out to enable the performance analysis and subsequent improvement of many existing filtering

Digital Object Identifier 10.1109/MSP.2011.2179329
Date of publication: 5 December 2012

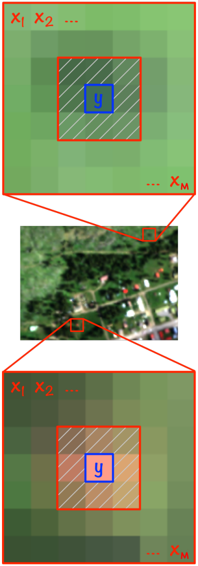
 1053-5888/13/$31.00©2013IEEE

Target and anomaly detection



- Detect weak signals in cluttered backgrounds
- Characterize cluttered backgrounds

Target and anomaly detection

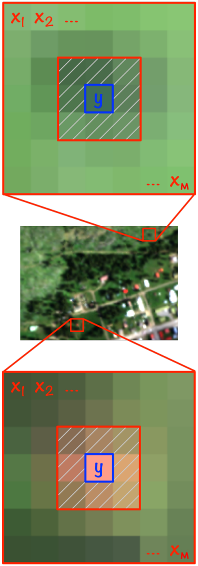


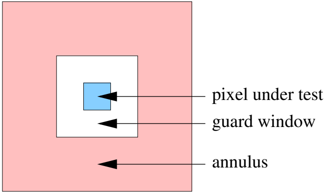
- Detect weak signals in cluttered backgrounds
- Characterize cluttered backgrounds

- Step 1: estimate the target-free background
  - eg, global Gaussian model: mean and covariance from data
  - or: mixtures, subspaces, endmembers, manifolds,. . .
  - or: local mean, local or global covariance
  - or: local regression

- Step 2: compare estimated to observed
  - disagreement implies anomalousness

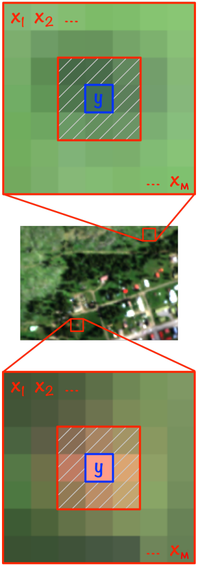
RX = Mahalanobis distance

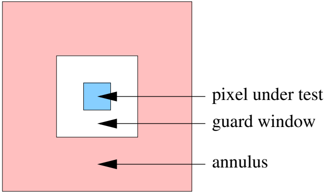




- Step 1: estimate the target-free background
  - average pixels in surrounding annulus
  - Estimate: $\hat{y} = (x_1 + x_2 + \cdots + x_M)/M$

RX = Mahalanobis distance





- Step 1: estimate the target-free background
  - average pixels in surrounding annulus
  - Estimate: $\hat{y} = (x_1 + x_2 + \cdots + x_M)/M$
- Step 2: compare estimated to observed
  - using Mahalanobis distance
  - Difference: $e = y - \hat{y}$
  - Covariance: $R = \langle ee^\mathsf{T} \rangle$
  - Anomalousness: $A = e^\mathsf{T} R^{-1} e$

Regression

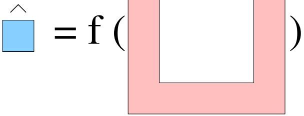
- Estimate target-free pixel value
  based on the pixels in the surrounding annulus
- eg, local mean:

$$\hat{y} = \frac{1}{M}(x_1 + \cdots + x_M)$$

- Seek function f, learned from the entire image, such that

$$\hat{y} = f(\mathbf{x}) = f(x_1, \ldots, x_M)$$

  provides a good estimate of y



Regression

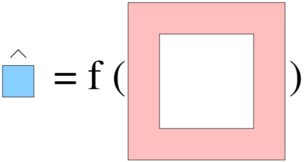
- Estimate target-free pixel value
  based on the pixels in the surrounding annulus
- eg, local mean:

$$\hat{y} = \frac{1}{M}(x_1 + \cdots + x_M)$$

- Seek function f, learned from the entire image, such that

$$\hat{y} = f(\mathbf{x}) = f(x_1, \ldots, x_M)$$

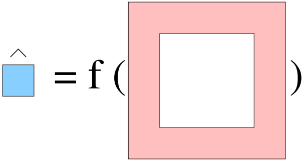
provides a good estimate of y



Aside: Linear $f(\mathbf{x})$ corresponds to convolutional filter

Regression: multiple (multi-/hyper-spectral) bands

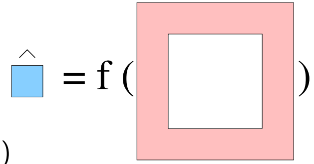
When we say:



- $\hat{y} = f(\mathbf{x})$

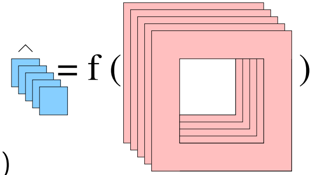
Regression: multiple (multi-/hyper-spectral) bands

When we say:



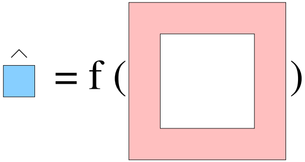
- $\hat{y} = f(\mathbf{x})$

What we mean is:



- $\hat{y} = f(\mathbf{x})$

Regression for target and anomaly detection



- This study:
  - local estimate of central pixel: $\hat{y} = f(\mathbf{x}) = f(x_1, \ldots, x_M)$
  - global estimate of covariance: $R = \langle\, (y - \hat{y})(y - \hat{y})^{\mathsf{T}}\, \rangle$
- Additive target detection (y vs. $y + \epsilon\mathbf{t}$):
  - local matched filter: $\mathcal{T}(\mathbf{x}, y) = \mathbf{t}^{\mathsf{T}} R^{-1}(y - f(\mathbf{x}))$
- Anomaly detection
  - RX-like: $\mathcal{A}(\mathbf{x}, y) = (y - f(\mathbf{x}))^{\mathsf{T}} R^{-1}(y - f(\mathbf{x}))$

Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$

- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test:
    Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$

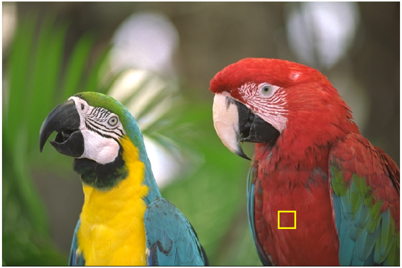
Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$

- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test: Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$

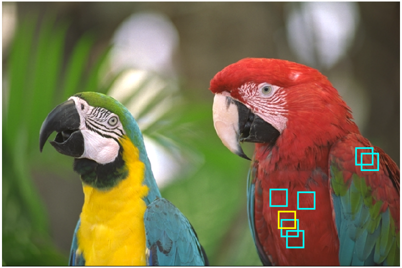
Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$
- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test:
    Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$



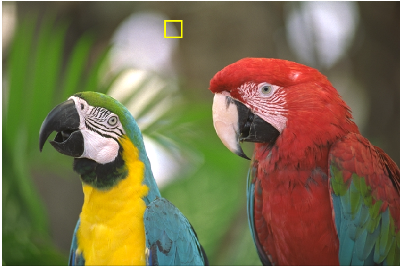
Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$
- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test:
    Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$



Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$
- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test: Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$



Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$
- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test: Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$

Regression algorithms

- Aim is to find f such that $y \approx \hat{y} = f(\mathbf{x})$
- kNN – (weighted) k-nearest neighbors (in $\mathbf{x}$-space)
  - Given annulus $\mathbf{x}$ around the pixel under test: Find locations $i_1, i_2, \ldots, i_k$ such that annuli $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_k}$ are close to $\mathbf{x}$. That is: $\|\mathbf{x} - \mathbf{x}_i\|$ is small.
  - Assign weights $w_1, \ldots, w_k$ so closer points have more weight: eg, $w_i \sim \exp(\|\mathbf{x} - \mathbf{x}_i\|^2/\sigma^2)$
  - Average the associated y values: $\hat{y} = f(\mathbf{x}) = \sum_{j=1}^{k} w_j y_{i_j}$
- Global least-squares linear fit
  - Let $f(\mathbf{x}) = f(x_1, \ldots, x_M) = \alpha_1 x_1 + \ldots \alpha_M x_M$
  - Fit α's globally over the whole image: minimize $\sum_i [y_i - f(\mathbf{x}_i)]^2$

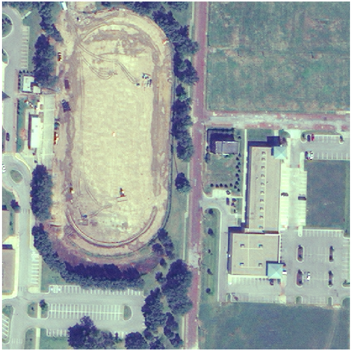
Images

Parrots

Boat

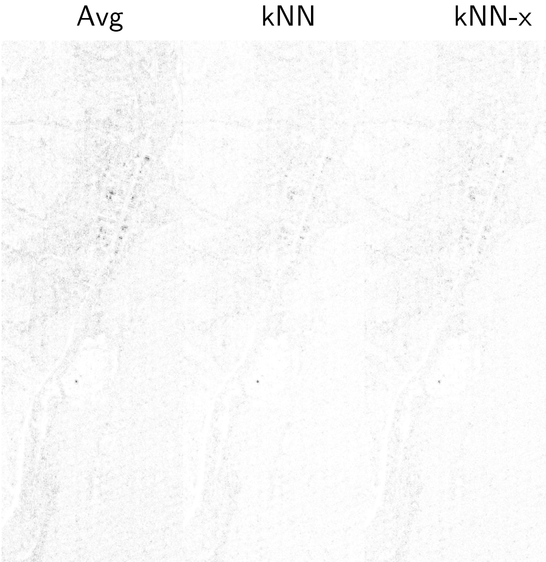
HyMap

WorldView-2



- 2 Photographs
- 2 Remote sensing
- Bands: 3,3,8,126

Where are the errors? (HyMap: one band)



3x3

Where are the errors? (HyMap: one band)

Avg kNN kNN-x

5x5

Where are the errors? (HyMap: one band)

Avg kNN kNN-x

11x11

Multispectral (3 bands)

- Two strategies
  - **Band-by-band:** for each band b, estimate bth component of y as function of bth components of annulus pixels $x_1, \ldots, x_M$.
  - **Bands-together:** treat y and $x_1, \ldots, x_M$ as vectors; this mostly affects choice of nearest neighbors
- For local mean, two strategies are equivalent
- Results: kNN outperforms local mean

SNR(dB)	Local Mean	kNN band-by-band	kNN bands together
Parrots	20.13	26.48	25.60
Boat	12.05	18.10	17.62
WV-2 (bands 1,2,3)	12.94	17.24	16.06
HyMap (bands 1,2,3)	16.04	19.66	19.49

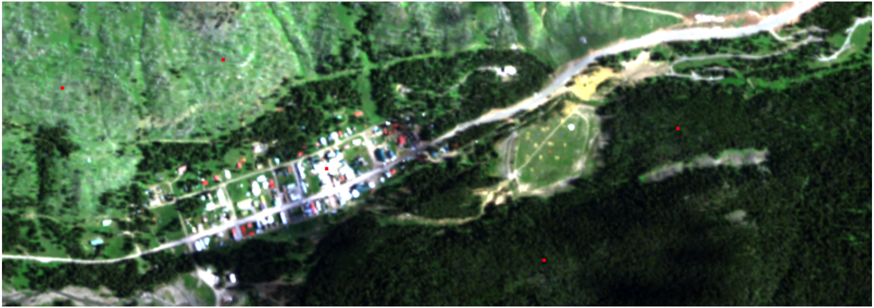
Higher SNR $\implies$ Improved target detection performance?

Higher SNR $\Longrightarrow$ Improved target detection performance?

N. Hasson, S. Asulin, S. R. Rotman, and D. Blumberg, "Evaluating backgrounds for subpixel target detection: **when closer isn't better**." *Proc. SPIE* **9472**, p. 94720R, 2015.

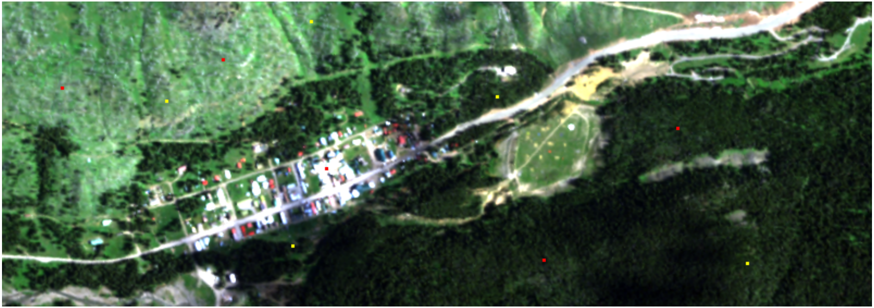
Implanting targets: misplaced pixel anomalies

Implanting targets: misplaced pixel anomalies



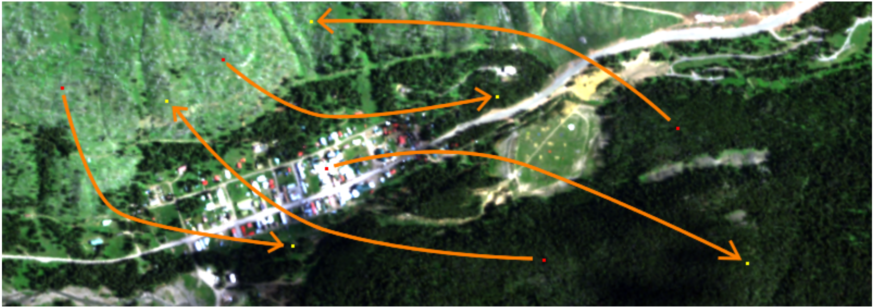
- Select some pixel locations at random (red squares)

Implanting targets: misplaced pixel anomalies



- Select some pixel locations at random (red squares)
- Select some more locations at random (yellow squares)

Implanting targets: misplaced pixel anomalies



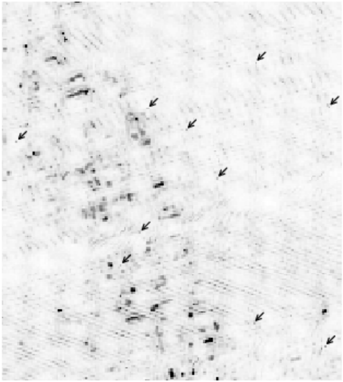
- Select some pixel locations at random (red squares)
- Select some more locations at random (yellow squares)
- Copy pixel values at red locations to yellow locations

Anomaly detection with implanted targets

- Choose a small number of locations at random
- Replace pixels with anomalous targets
  - eg, $\mathbf{y} \leftarrow (1-\epsilon)\mathbf{y} + \epsilon\mathbf{t}$
  - with $\mathbf{t} \in \mathcal{U}$
- Use regression to learn f
- Apply f to compute background estimates: $\hat{\mathbf{y}}_i = f(\mathbf{x}_i)$
- Make residual map: $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$
- Compute covariance of residuals: $\Phi = \langle \mathbf{e}\mathbf{e}^T \rangle$
- Use RX-like anomaly detection: $\mathcal{A}(\mathbf{e}) = \mathbf{e}^T\Phi^{-1}\mathbf{e}$
- Implanted targets should have high values of $\mathcal{A}$

Implanted targets (HyMap – Cooke City)

Mean

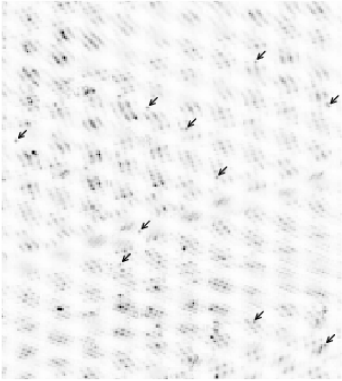


Anomalousness

744 false alarms

Implanted targets (HyMap – Cooke City)

$D_4\Sigma$ Regression

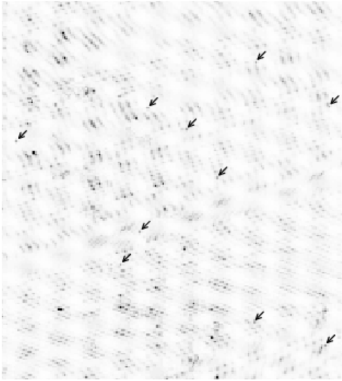


Anomalousness

233 false alarms

Implanted targets (HyMap – Cooke City)

$D_4\Sigma$-PCA Regression

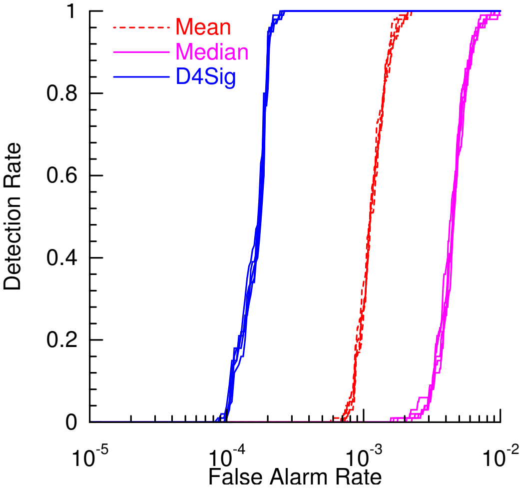


Anomalousness

167 false alarms

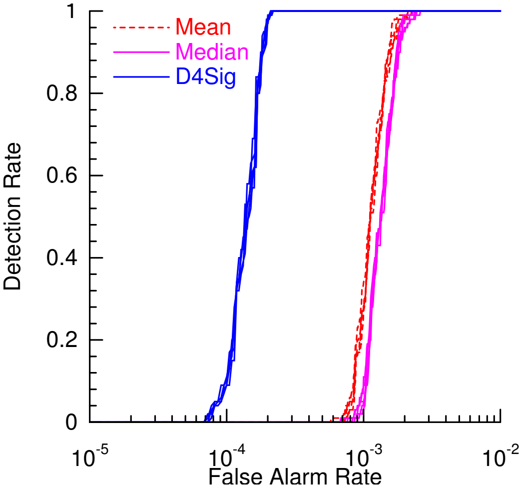
Implanted targets (HyMap – Cooke City)

Direct band-by-band



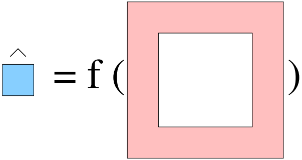
Implanted targets (HyMap – Cooke City)

PCA band-by-band



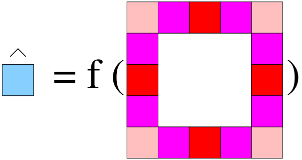
Aside: symmetrized regression

- $\hat{y} = f(\mathbf{x})$



Aside: symmetrized regression

- $\hat{y} = f(\mathbf{x})$

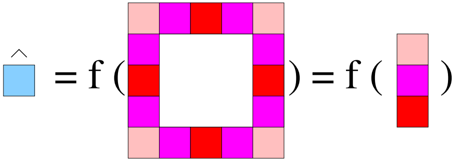


- Introduce symmetry-preserving features:

$$\Phi_p(x_1, x_2, \ldots, x_k)$$

Aside: symmetrized regression

- $\hat{y} = f(\mathbf{x})$



- Introduce symmetry-preserving features:

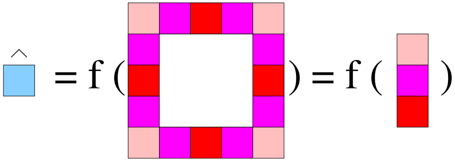
$$\Phi_p(x_1, x_2, \ldots, x_k)$$

- Make regression a function of features:

$$\hat{y} = f(\Phi_1, \ldots, \Phi_P)$$

Aside: symmetrized regression

- $\hat{y} = f(\mathbf{x})$



- Introduce symmetry-preserving features:

$$\Phi_p(x_1, x_2, \ldots, x_k)$$

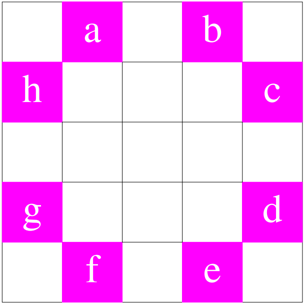
- Make regression a function of features:

$$\hat{y} = f(\Phi_1, \ldots, \Phi_P)$$

- Features preserve symmetry, so choice of function is arbitrary:
  - (weighted) k-nearest neighbors
  - Linear regression
  - Support vector regression, deep neural network, etc, etc.

Dihedral group (D_4) symmetry

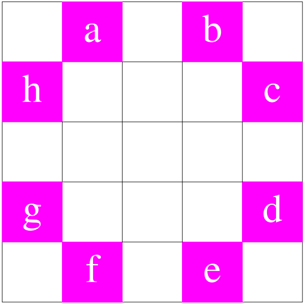
$$D_4 = \{1, r, r^2, r^3, s, sr, sr^2, sr^3\}$$



- Group has two generators: s, r
- s: reflection
  $sa = b$, $sb = a$, $sc = h$, $sd = g$, etc.
- r: rotation
  $ra = c$, $rb = d$, $rc = e$, $rd = f$, etc.
- Composition rules: $s^2 = 1$, $r^4 = 1$, $rs = sr^3$

$$e.g., \ r^2 \times sr^2 = r(rs)r^2 = r(sr^3)r^2 = (rs)r^5 = (rs)r = (sr^3)r = sr^4 = s.$$

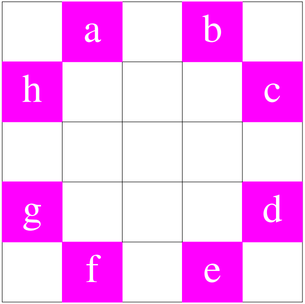
Sigma ($D_4\Sigma$) features



$$\Phi_{\Sigma\Sigma\Sigma} = a + b + e + f + c + d + g + h$$

- Features invariant to group operations: $g\Phi = \Phi$ for $g \in D_4$

Sigma-Delta ($D_4\Sigma\Delta$) features



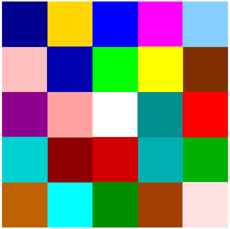
$$
\begin{aligned}
\Phi_{\Sigma\Sigma\Sigma} &= a+b+e+f+c+d+g+h \\
\Phi_{\Sigma\Sigma\Delta} &= \Big|(a+b+e+f)-(c+d+g+h)\Big| \\
\Phi_{\Sigma\Delta\Sigma} &= \big|(a+b)-(e+f)\big| + \big|(c+d)-(g+h)\big| \\
\Phi_{\Sigma\Delta\Delta} &= \Big||(a+b)-(e+f)| - |(c+d)-(g+h)|\Big| \\
\Phi_{\Delta\Sigma\Sigma} &= |a-b|+|e-f|+|c-d|+|g-h| \\
\Phi_{\Delta\Sigma\Delta} &= \Big|(|a-b|+|e-f|)-(|c-d|+|g-h|)\Big| \\
\Phi_{\Delta\Delta\Sigma} &= \big||a-b|-|e-f|\big| + \big||c-d|-|g-h|\big| \\
\Phi_{\Delta\Delta\Delta} &= \Big|\big||a-b|-|e-f|\big| - \big||c-d|-|g-h|\big|\Big|
\end{aligned}
$$

- Features invariant to group operations: $g\Phi = \Phi$ for $g \in D_4$

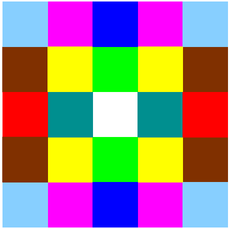
$$
\begin{aligned}
e.g.,\ r\Phi_{\Sigma\Sigma\Delta} &= \Big|(ra+rb+re+rf)-(rc+rd+rg+rh)\Big| \\
&= \Big|(c+d+g+h)-(a+b+e+f)\Big| = \Phi_{\Sigma\Sigma\Delta}
\end{aligned}
$$

Various symmetries

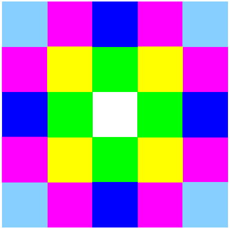
No symmetry



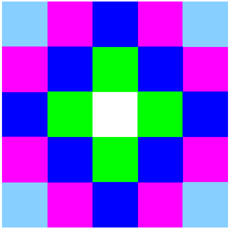
Klein K_4



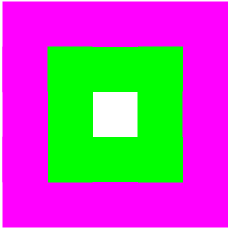
Dihedral D_4



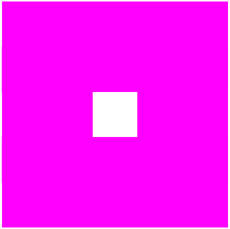
Diamond Rings



Square Rings

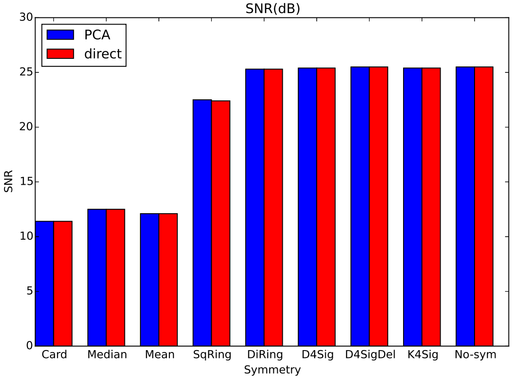


Mean



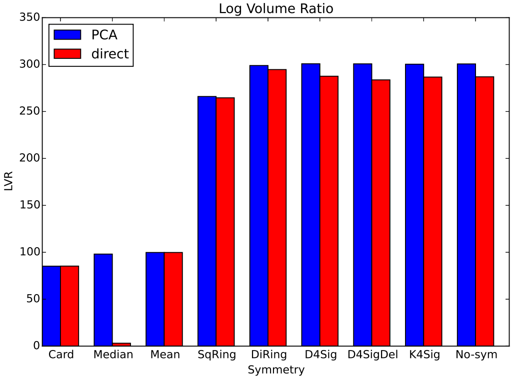
Using various symmetries

- Symmetry provides comparable performance with fewer features



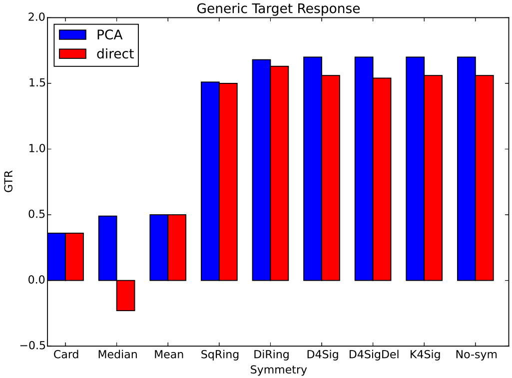
Using various symmetries

- Symmetry provides comparable performance with fewer features



Using various symmetries

- Symmetry provides comparable performance with fewer features



Conclusions

- Background non-stationarity is a great problem

Conclusions

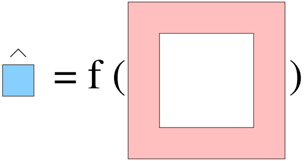
- Background non-stationarity is a great ~~problem~~ opportunity
  - Better background estimation $\rightarrow$ better target detection

Conclusions

- Background non-stationarity is a great ~~problem~~ opportunity
  - Better background estimation → better target detection
- Local mean is not the only way to estimate central pixel

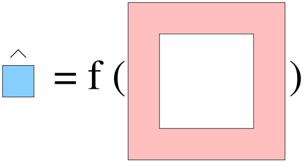
Conclusions

- Background non-stationarity is a great ~~problem~~ opportunity
  - Better background estimation → better target detection
- Local mean is not the only way to estimate central pixel: we can do better!



Conclusions

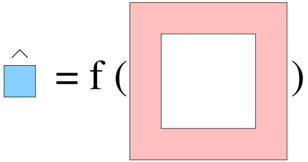
- Background non-stationarity is a great ~~problem~~ opportunity
  - Better background estimation → better target detection
- Local mean is not the only way to estimate central pixel: we can do better!



- Although motivated by modern image processing, we saw improvement just using ordinary regression
  - kNN, simple linear fit: both are better than local mean
  - and better still for larger annulus size

Conclusions

- Background non-stationarity is a great ~~problem~~ opportunity
  - Better background estimation → better target detection
- Local mean is not the only way to estimate central pixel: we can do better!



- Although motivated by modern image processing, we saw improvement just using ordinary regression
  - kNN, simple linear fit: both are better than local mean
  - and better still for larger annulus size
- What we did was simple – you can do better!

You can do better!

ie, Future Work

- Apply "real" machine learning tools

- Look at band-by-band variants
  - Consider transformations besides PCA

- Segment image; learn separate $f(\mathbf{x})$ for each segment
  - Use k-means-like approach to learn $f(\mathbf{x})$ and segments together
  - Use a separate covariance matrix R for each segment
  - Identify appropriate regularization of R

- Extend to known (and variable) target detection scenarios

- Optimize direct measures of performance (vs least squares)
  - Remember: closer isn't always better