Computational Strategies for Understanding Underwater Optical Image Datasets

Jeffrey W. Kaeli
MIT / WHOI Joint Program in Oceanographic Engineering

ADVISOR:
Hanumant Singh, Deep Submergence Lab, WHOI

COMMITTEE:
John Leonard, Marine Robotics Group, MIT
Ramesh Raskar, Camera Culture Group, MIT Media Lab
Antonio Torralba, CSAIL, MIT EECS
the “latency of understanding” paradigm

O(10-30 bytes/sec)  
Frietag et al. 2005

O(hours)  

O(days-weeks)  

O(100,000 images)
• compressed images *can* be transmitted acoustically
  *which images get sent?*

• classification serves as semantic compression

• real-time automated classification algorithms
  *correct for illumination/attenuation artifacts?*
motivation from related research

Loomis, 2011 PhD Thesis

• “the images were rigorously color corrected…”

• radically different approaches towards scene classification vs. object detection

• methods are computationally expensive
overview of thesis structure

Kaeli, 2013 PhD Thesis

0. introduction

1. underwater image correction

2. computational strategies

3. understanding underwater image datasets

4. conclusions
overview of thesis contributions

*Kaeli, 2013 PhD Thesis*

1. underwater image correction

- detailed model of underwater image formation

- review of broad range of correction techniques

- present novel method for correction for robotic imaging platforms based on estimating environmental and system parameters using multi-sensor fusion
overview of thesis contributions
Kaeli, 2013 PhD Thesis

2. computational strategies

- use Hierarchical Discrete Correlation [Burt 81] as basis for novel *octagonal* pyramid framework

- demonstrate efficient computation of gradient an

- explore design of invariant features for reduced computation overhead in situ
overview of thesis contributions

Kaeli, 2013 PhD Thesis

3. understanding underwater image datasets

- fast keypoint detection and description

- online navigation summaries [Girdhar 12] as basis for unsupervised mission-time low-bandwidth map

- supervised object detection: finding crabs

- building semantic maps
overview of thesis structure

1. underwater image correction

2. computational strategies

3. understanding underwater image datasets
1. underwater image correction

underwater image formation
1. underwater image correction

*robotic imaging platforms*

\[ c_\Lambda = BP_{\theta,\phi} r_\Lambda e^{-\alpha_\Lambda \ell} \]

\[ \log c_\Lambda = \log BP_{\theta,\phi} + \log r_\Lambda - \alpha_\Lambda \ell \]
1. underwater image correction

\[ \log c_A = \log BP_{\theta, \phi} + \log r_A - \alpha_A \ell \]
1. underwatet image correction

*diversity of approaches to correction*

- raw image
- frame averaging
- adaptive histogram equalization
- white balance
- homomorphic filtering
- a.h.e. + white balance

\[
\frac{1}{K} \sum_{k} c_{\lambda,k} \approx I_{\lambda} \quad \frac{1}{K} \sum_{k} r_{\lambda,k} = I_{\lambda} \bar{r}_{\lambda}
\]

\[
c_{\lambda} = I_{\lambda} r_{\lambda}
\]

\[
\log c_{\lambda} = \log I_{\lambda} + \log r_{\lambda}
\]
1. underwater image correction

*constrain light field equation using keypoints*
1. underwater image correction

\[ \alpha_{\Lambda} = \frac{\log c_{\Lambda,1} - \log c_{\Lambda,2}}{\ell_2 - \ell_1} \]

\[ \log BP(\theta_i, \phi_j) = \sum_{\Lambda} \frac{1}{|P|} \sum_P \log c_{\Lambda} + \alpha_{\Lambda} \ell \]
1. underwater image correction
   estimate beam pattern and correct
1. underwater image correction

sample corrected imagery
1. underwater image correction

corrected imagery
overview of thesis contributions

Kaeli, 2013 PhD Thesis

1. underwater image correction

- detailed model of underwater image formation
- review of broad range of correction techniques
- present novel method for correction for robotic imaging platforms based on estimating environmental and system parameters using multi-sensor fusion
overview of thesis structure

1. underwater image correction

2. computational strategies

3. understanding underwater image datasets
2. computational strategies

multi-scale image representations

convolution is still major bottleneck in many multi-scale image processing framework [Van 2011] even in fast keypoint description [Calonder 2010]

can we exploit pixel grid geometries that allow us to substitute adds and bit shifts for costly convolutions while still approximating a Gaussian? [Viola 2001]

applications on low-power robotic imaging platforms
2. computational strategies

*multi-scale image representations*

- continuous scale-space
  
  \[\text{[Lindeberg 1994]}\]

- discrete scale-space: pyramids
  
  \[\text{[Burt 1983]}\]
2. computational strategies

Hierarchical Discrete Correlation [Burt 1981]
2. computational strategies

the octagonal pyramid
2. computational strategies

the recursive octagonal kernel

\[ \mathcal{K} = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \]

Graph showing correlation with radius for different kernel weights (1, 2, 4, 8, 16) and comparison with Gaussian kernel.
2. computational strategies
the recursive octagonal kernel

only 3P adds + P bit shifts! (for sqrt(2) scale resolution)

compare with ~3.3P multiplies + ~2.7 adds (for 1 scale resolution)

must be vigilant about absolute orientation between levels
2. computational strategies

*efficient oriented gradient computation*

\[ M = \sqrt{I_x^2 + I_y^2} \]
\[ \theta = \tan^{-1}\left(\frac{I_y}{I_x}\right) \]

how fine angular resolution do we need if we’re binning?

\[ D_x = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \]
2. computational strategies

*efficient oriented gradient computation*

\[ M \approx \max(|I_x|, |I_y|) + \frac{1}{2} \min(|I_x|, |I_y|) \]

\[
\begin{cases}
I_x > 0 \\
I_y > 0 \\
|I_x| > |I_y| \\
||I_x| - |I_y|| > \frac{M}{2}
\end{cases}
\]

9% overestimate
RMS error only +3%
max bin diff ~6%

inspired by LBP [Ojala 2002]
2. computational strategies

*opponent color space*

\[ O_0 = \frac{1}{4} (R + 2G + B) \]
\[ O_1 = \sqrt{3} (R - B) \]
\[ O_2 = 2G - R - B \]
2. computational strategies

**color demosaicing**

\[
O_0 = \frac{1}{4} (R + 2G + B) \\
O_1 = \sqrt{3} (R - B) \\
O_2 = 2G - R - B
\]

compare \(O(16P)\) sums versus \(O(30P)\) multiplies!

[Malvar 04]
2. computational strategies

*underwater invariance*

\[
c_\Lambda = BP_{\theta,\phi} r_\Lambda e^{-\alpha_\Lambda \ell}
\]

\[
\log c_\Lambda = \log BP_{\theta,\phi} + \log r_\Lambda - \alpha_\Lambda \ell
\]

\[
\nabla \log c_\Lambda \approx \nabla \log r_\Lambda
\]

\[
\mathcal{O}_{L0} = \frac{1}{4} (\log R + 2\log G + \log B)
\]

\[
\mathcal{O}_{L1} = \sqrt{3} (\log R - \log B)
\]

\[
\mathcal{O}_{L2} = 2\log G - \log R - \log B
\]
3. understanding underwater image datasets

*illumination invariance – keypoint detection*
2. computational strategies

attenuation invariance

\[ s \approx \max \left( |O_{L1}^*|, |O_{L2}^*| \right) + \frac{1}{2} \min \left( |O_{L1}^*|, |O_{L2}^*| \right) \]

\[
\begin{align*}
\Omega_{L1}^* > 0 \\
\Omega_{L2}^* > 0 \\
|\Omega_{L1}^*| > |\Omega_{L2}^*| \\
| |\Omega_{L1}^*| - |\Omega_{L2}^*| | > \frac{s}{2}
\end{align*}
\]

\[ \tilde{\alpha} = \begin{bmatrix} \alpha_{L1} \\ \alpha_{L2} \end{bmatrix} = \begin{bmatrix} \sqrt{3} (\alpha_R - \alpha_B) \\ 2\alpha_G - \alpha_R - \alpha_B \end{bmatrix} \]
2. computational strategies

*underwater invariance*
2. computational strategies

*underwater invariance*

Other strategies of invariance:

- Single attenuation invariant axis [Finlayson 01]
- Gradients of log color [Funt 95]
- Comprehensive color image normalization [Finlayson 98]
overview of thesis contributions
Kaeli, 2013 PhD Thesis

2. computational strategies

- use Hierarchical Discrete Correlation \([Burt\ 81]\) as basis for novel octagonal pyramid framework

- demonstrate efficient computation of oriented gradients and color features

- explore design of invariant features for reduced computation overhead in situ
overview of thesis structure

1. underwater image correction

2. computational strategies

3. understanding underwater image datasets
3. understanding underwater image datasets

**keypoint detection**

- extrema in difference-of-Gaussian function across scale space make stable keypoints [Lowe 2004]

\[
D(\sigma) = (G(k\sigma) - G(\sigma)) \ast I = L(k\sigma) - L(\sigma)
\]

- however, for “homogeneous” kernels [Lindeberg 93]

\[
G(k\sigma) = G(\sigma) \ast G(\sigma)
\]

\[
D(\sigma) = G(\sigma) \ast (G(\sigma) - 1) \ast I
\]

- D can be accumulated up the scale space!
3. understanding underwater image datasets

*keypoint detection – compare pyramids*

\[ \ell = 2^{\frac{1}{2}} \]
\[ \ell = 2 \]
\[ \ldots \]
\[ \ell = 1^{\frac{1}{2}} \]
\[ \ell = 1 \]
\[ \ell = \frac{1}{2} \]
\[ \ell = 0 \]

downsample

\[ \ell \]
\[ \ell - 1 \]
\[ \ell - 2 \]
\[ \ell - 3 \]
\[ \ell - 4 \]
3. understanding underwater image datasets

*keypoint detection – compare pyramids*

\[
\sqrt{2} \text{ scale resolution ample for keypoint detection} \quad [\text{Lowe 04}]
\]

octagonal pyramid
- \(O(3)\) adds!
- 14 neighbors

traditional pyramid
- \(O(35)\) multiplies & \(O(27)\) adds
- 26 neighbors
3. understanding underwater image datasets

*keypoint detection – log intensity*
3. understanding underwater image datasets

Keypoint detection – SIFT (blue) OP (yellow)

OP detects same *kind* of keypoints in images (if not the same ones), appealing for bag-of-keypoints model
3. understanding underwater image datasets

keypoint description – QuAHOG

extract region, accumulate histogram of gradients
QUantize Accumulated Histogram or Oriented Gradients

analogous to LBP [Ojala 2002]
underwater imagery is largely redundant, how can we communicate “key” images?

- offline vs online approaches

- minimize our “surprise” at seeing the dataset

- use summary images as basis for semantic maps

[Girdhar & Dudek, 2012]
3. understanding underwater image datasets

Online Navigation Summaries

$\text{prior}$

$\text{posterior}$

$d_{KL}$

Kullback-Leibler divergence

$\text{[Girdhar & Dudek, 2012]}$
3. understanding underwater image datasets

*Online Navigation Summaries*
3. understanding underwater image datasets

*Online Navigation Summaries*

1 (1897)  
2 (101)  
3 (9)  
4 (35)  
5 (531)  
6 (5)  
7 (2)  
8 (1)
3. understanding underwater image datasets

*Online Navigation Summaries*
3. understanding underwater image datasets

*Online Navigation Summaries*
3. understanding underwater image datasets

Online Navigation Summaries

![Graph showing depth vs image number with images at different depths.]

- Depth [m]
- Image Number
3. understanding underwater image datasets

*Online Navigation Summaries*

- Conclusions
  - Decent summary of substrate types
  - Don’t need expensive features for bag of words model

- Further work
  - How to make robust to transmitting summary images partway through
3. understanding underwater image datasets

*supervised object detection: crabs*

- intuition from fish detection [Loomis 2011]
  1. color
     - saturated red
  2. shape
     - long thin legs
3. understanding underwater image datasets

*supervised object detection: crabs*

- white-balance log opponent color is simple subtraction
- compute hue and saturation using binary pattern method
- index hue by weight vector $w_\phi$ and multiply by saturation
3. understanding underwater image datasets

**supervised object detection: crabs**

- white-balance log opponent color is simple subtraction

- compute hue and saturation using binary pattern method

- index hue by weight vector $w_\phi$ and multiply by saturation

- accumulate up scale space and find local maxima

[Swain & Ballard, 1991]
3. understanding underwater image datasets

*supervised object detection: crabs*

---

1. flat  
5. corner  

2. edge  
6. thick bar  

3. misc.  
7. spit  

4. thin bar  
8. spot  

- gradients computed at lowest scale, accumulated, then threshold HOGs half their blurred mean gradient
overview of thesis contributions

Kaeli, 2013 PhD Thesis

3. understanding underwater image datasets

- fast keypoint detection and description

- online navigation summaries [Girdhar 12] as basis for unsupervised mission-time low-bandwidth map

- supervised object detection: finding crabs

- building semantic maps
conclusions

Advanced our ability to realistically process underwater images in mission time aboard robotic imaging platforms

Coupled with state-of-the-art image compression and acoustic transmission algorithms, reduce the latency of understanding paradigm for AUVs