Incremental Sparse Saliency Detection

Yin Li, *Yue Zhou*, Lei Xu, Xiaochao Yang, Jie Yang

Institute of Image Processing & Pattern Recognition

Shanghai Jiao Tong University, China

Outline

- Introduction
- Related Work
- Our Proposed Method
- Experiments and Analysis
- Conclusion and Future Work

Introduction

- **Motivation**

Everyone knows what attention is…

——William James

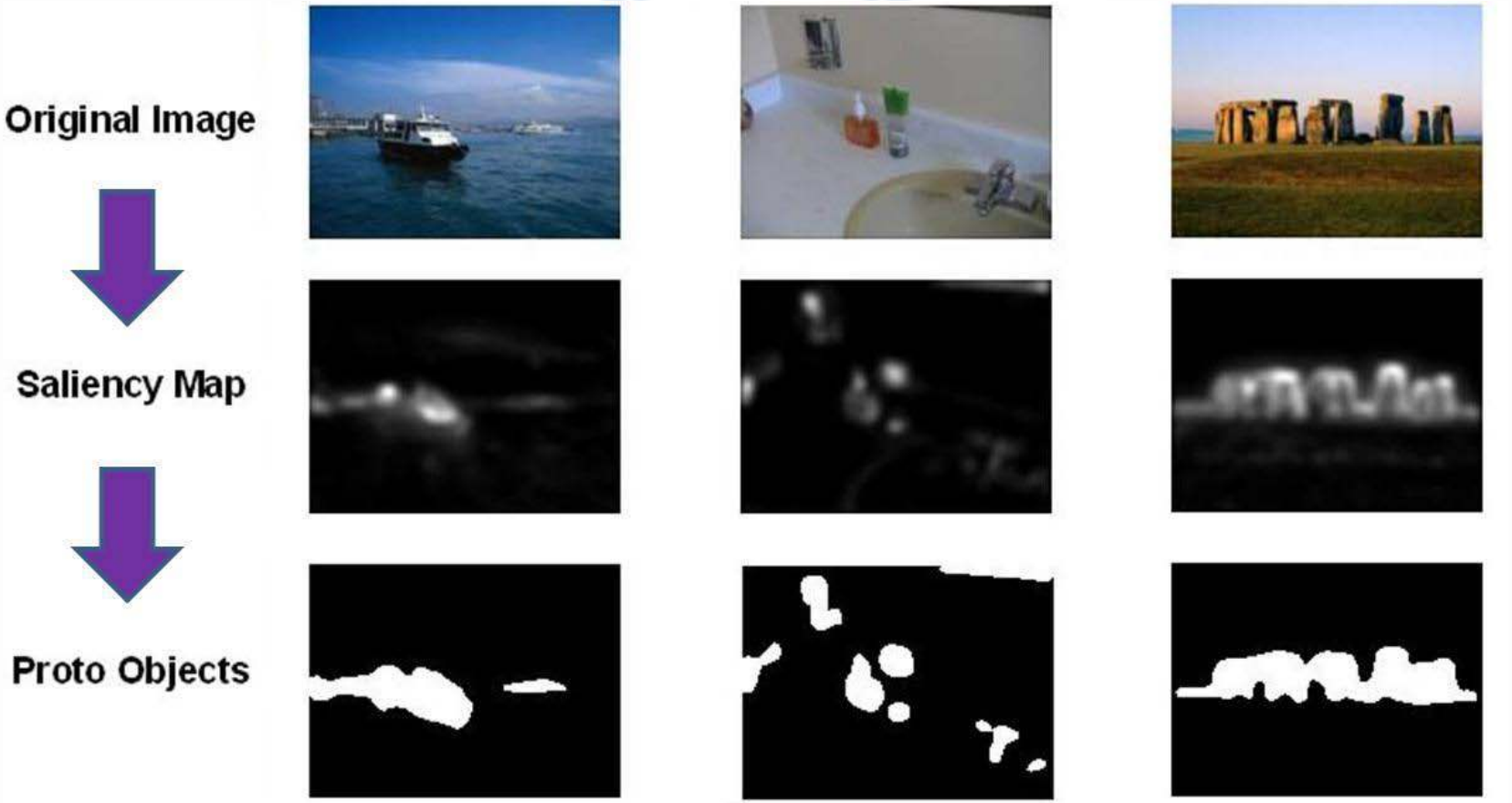
- **A computational approach to visual attention**
- **Fast selection for objects of interest in scenes**

Introduction

- **Difficulties**
  - **"Black box" problem**
    - Covert & overt attention
    - Biological plausible
  - **Difficulty in evaluation**
    - Quantitative analysis
    - The data set

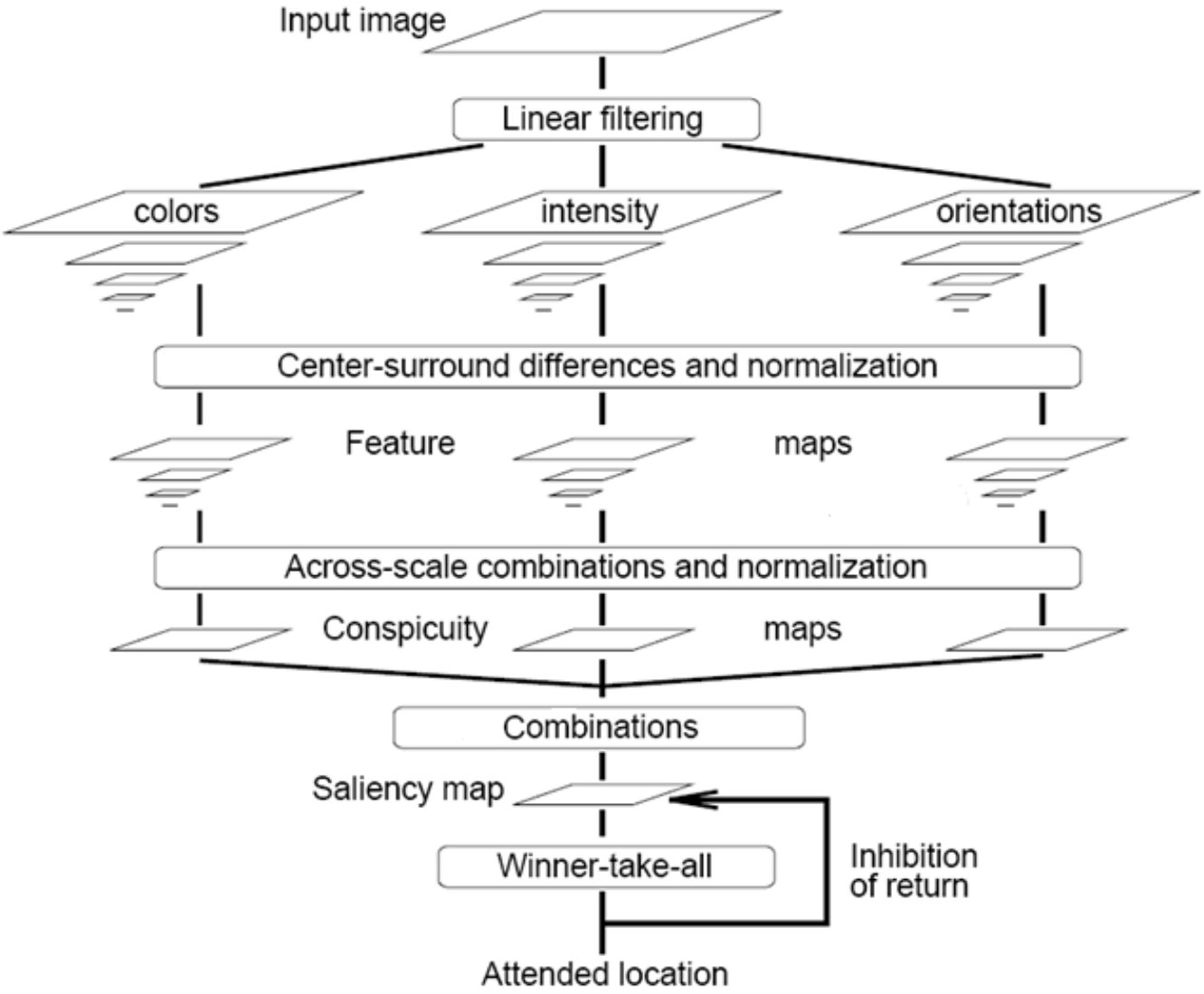
Introduction

Overview

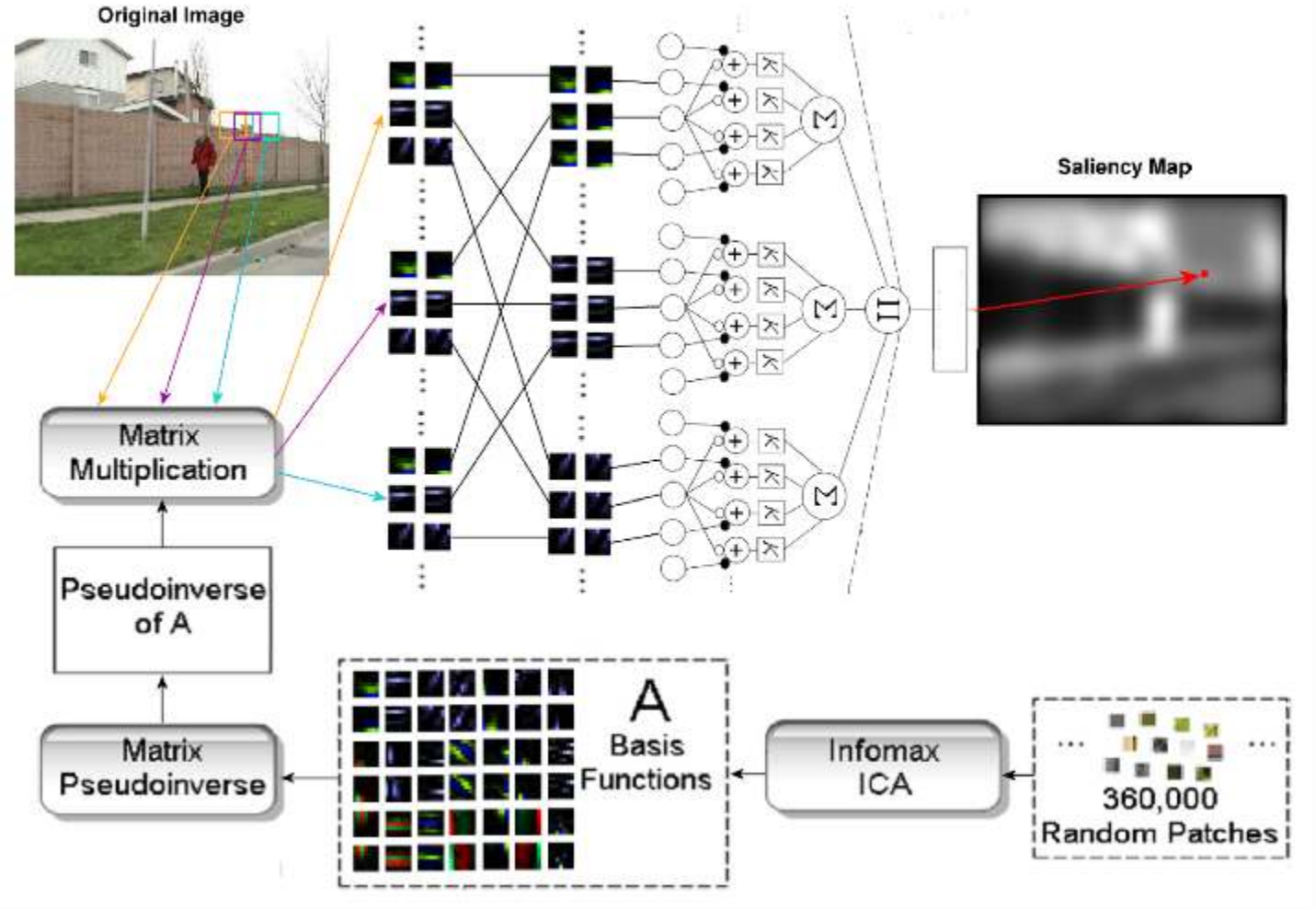


Related Work

Feature Integration: Itti1998, Itti2000, Itti2005, Gao2008…



Related Work



Bruce2004,

Hou2008,

…

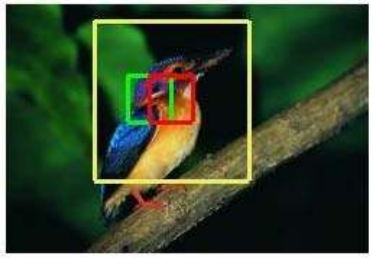
Related Work

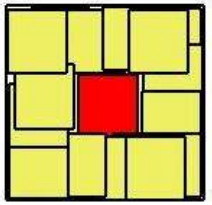
- **Other Method:**
  - Spectral Residual [Hou2007]
  - Contextual Guidance [Oliva2006]
  - Learning to Detect A Salient Object [Liu2007]
  - …

Our Proposed Model: Theory

Center-Surround Architecture

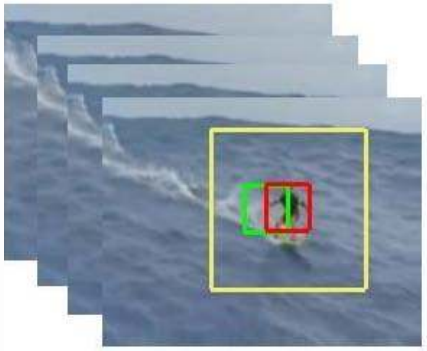
Image

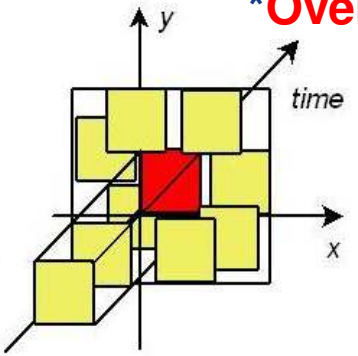




*Overlapping allowed

Video





Center

Surround

Feature

$$x = Fc \in R^n \qquad S = [Fs_1, Fs_2, \ldots, Fs_N] \in R^{n \times N}$$

Our Proposed Method: Theory

- **Saliency as Incremental Coding Length (ICL)**
  - For certain lossy coding scheme $L_\varepsilon(\bullet)$
    - ε — distortion tolerance
  - Saliency of the center is defined as ICL:

$$\delta L_\varepsilon(x) = L_\varepsilon(S \cup x) - L_\varepsilon(S) = L_\varepsilon(x \mid S)$$

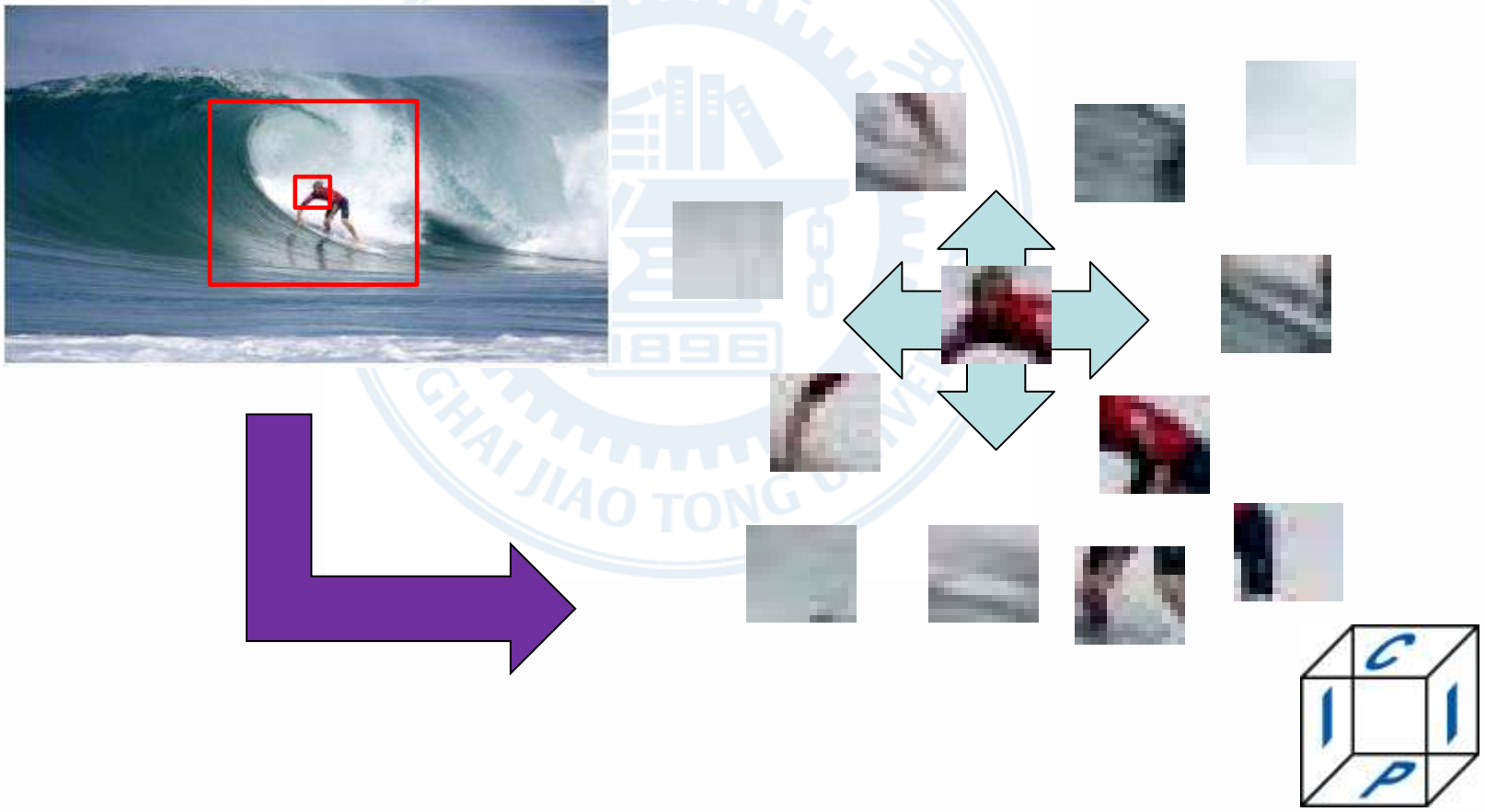
$$Sa(x) = \delta L_\varepsilon(x)$$

    - $x \mid S$ — encode x with S
    - **Optimum** coding scheme required

Our Proposed Method: Theory

- **Core Idea:**

  Saliency = Non-redundancy = Hard to encode



Our Proposed Method

- Sparse Coding Scheme
  - Center as the sparse linear representation of its surroundings

$$x \doteq \sum_{i=1}^{N} w_i F s_i = Sw \qquad w \in R^N$$

  - Traditional approach

$$w = \min_{w} \| x - Sw \|_2^2$$

Our Proposed Method

- **Sparse Coding Scheme**
  - **Our approach**

$$\min \| w \|_0 \quad s.t. \quad \| x - Sw \|_2^2 \leq \varepsilon$$

  - **Optimum coding length under distortion ε**
  - **Computational intractable — NP hard**

Our Proposed Method

- **Sparse Coding Scheme**
  - **Our approach (NP-hard)**

$$\min \| w \|_0 \quad s.t. \quad \| x - Sw \|_2^2 \leq \varepsilon$$

  - **Sparse assumption**

$$\| w \|_0 \ll N \quad given \quad n \ll N$$

*Feature invariance

(F is not important)

Solution (Polynomial)

$$\min \| w \|_1 \quad s.t. \quad \| x - Sw \|_2^2 \leq \varepsilon$$

Our Proposed Method

- **Sparse Coding Scheme**
  - **Our solution**

$$\min \| w \|_1 \quad s.t. \quad \| x - Sw \|_2^2 \le \varepsilon$$

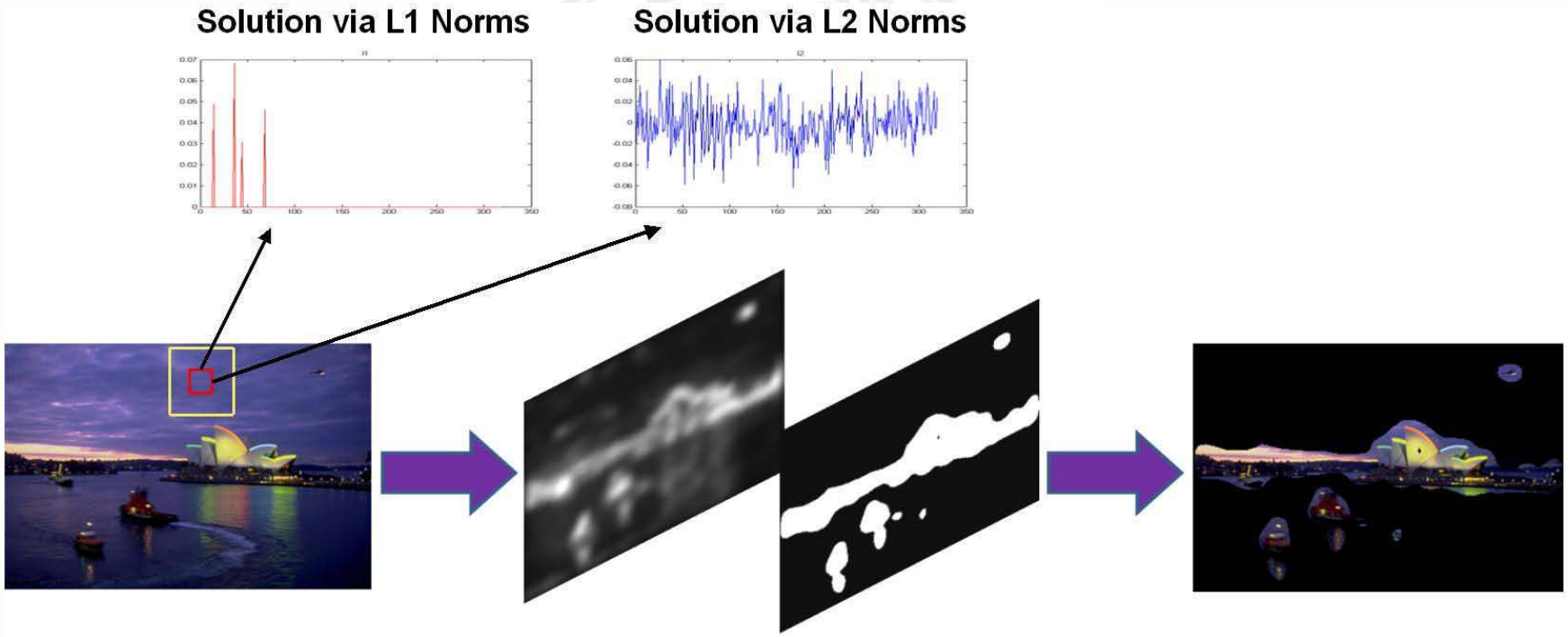
$$\min \lambda \| w \|_1 + \frac{1}{2} \| x - Sw \|_2^2 \qquad \lambda > 0$$

  - **Final saliency map by coding length**

$$Sa(c) = \delta L_\varepsilon(c) = \| w \|_0$$

Our Proposed Method

Sparse Coding Scheme



Our Proposed Method: Summary

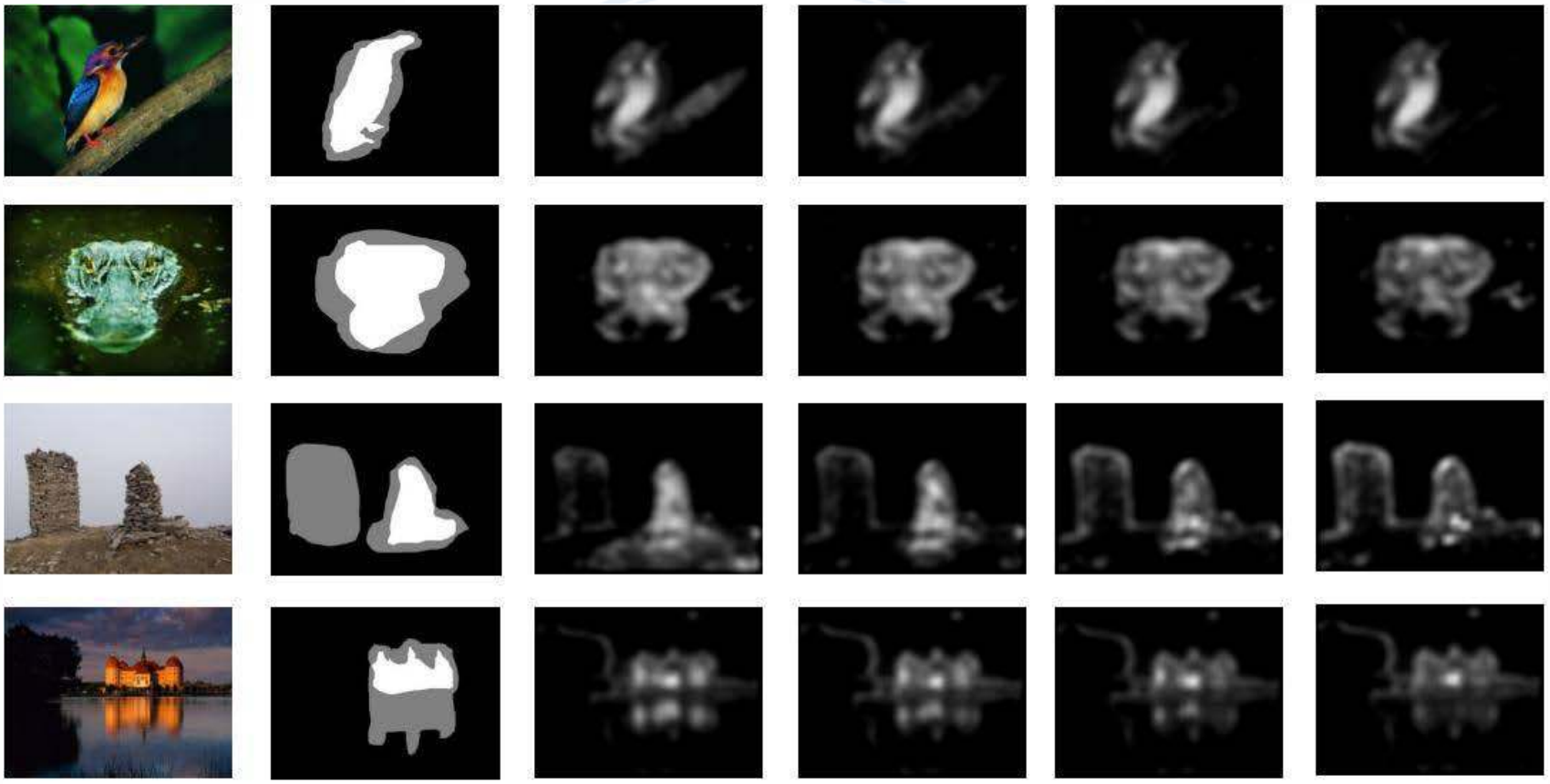
Summary

Algorithm1 (Incremental Sparse Saliency)

1. ***Input*** : given image I
2. ***for*** each patch c of the image I, calculate $\boldsymbol{x} = Fc$ and take patches from its surroundings to form $\boldsymbol{S}$
   - solve the optimization problem $min\ \lambda\|\boldsymbol{w}\|_1 + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{w}\|_2^2$
   - given the sparse solution $\boldsymbol{w}$, calculate the patch saliency $Sa(c)$ by $Sa(c) = \|w\|_0$, and accumulate the saliency by pixels
3. ***end***
4. ***Output*** : the saliency map of I

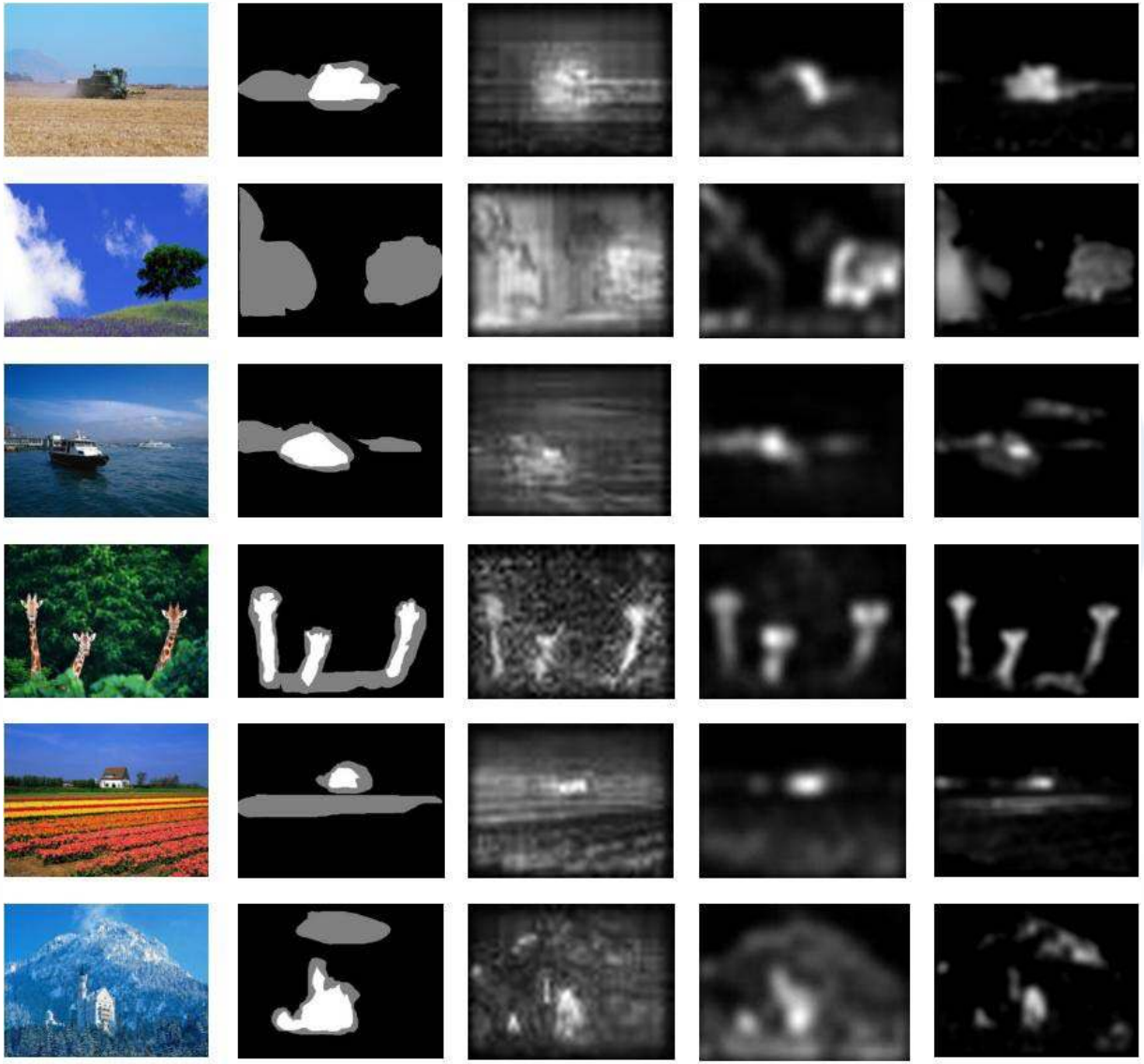
Experiment and Analysis

- One parameter: $\lambda > 0$



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

Experiment and Analysis: Images



From left to right

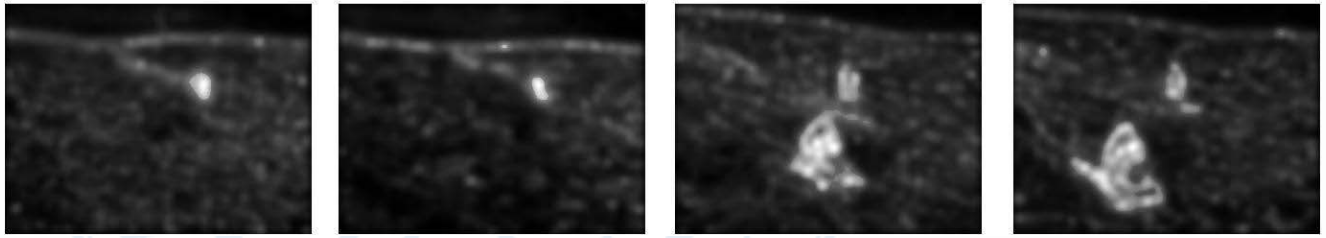
- Image
- Hand labeled
- Itti1998
- Hou2007
- Our Method

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Experiment and Analysis: Video

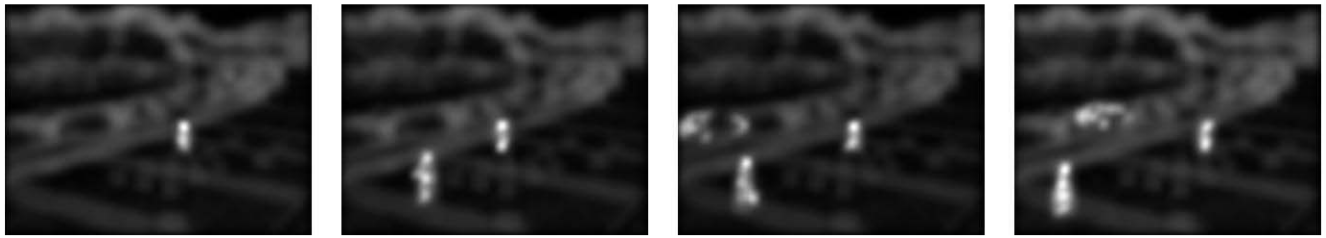
Video

Saliency Map



Video

Saliency Map



Conclusion and Future Work

Conclusion

- A visual saliency model by sparse coding
- Feature invariance
- Fairly good results

Future Work

- Quantitative evaluation of visual saliency
- Application of visual saliency in scene understanding

Thanks for your attention!