Photo-Quality Enhancement based on Visual Aesthetics

S. Bhattacharya*, R. Sukthankar**, M.Shah*

*University of Central Florida, **Intel labs & CMU
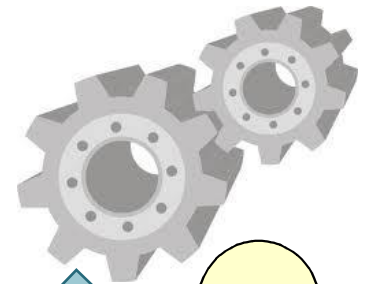
Motivation

Overview

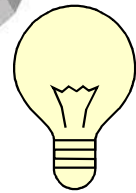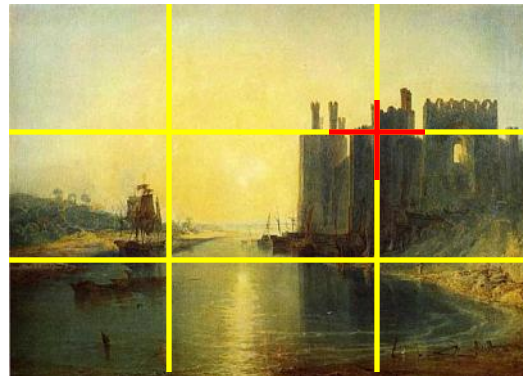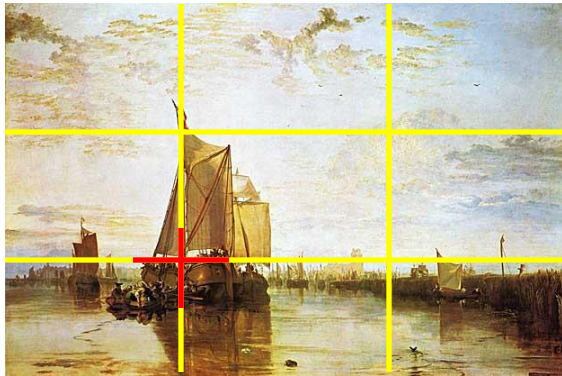# Visual Aesthetics: Rule of Thirds

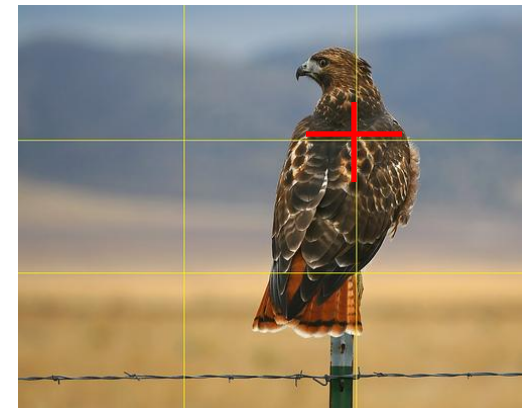Motivated by Renaissance Paintings…
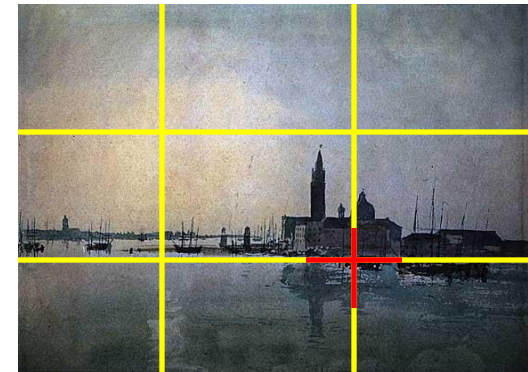
http://hoocher.com/Joseph_William_Turner/Joseph_William_Turner.htm

Rule of thirds: Subject of interest is aligned to one of the stress points

Professional photographs also abide this:

http://howtophotography.org/wp-content/uploads/2010/06/rule-of-thirds-photo2.jpg

Visual Aesthetics: Golden Ratio

Sky

Sea

~1.618k

k

Sky

Land

<u>Divine proportion</u>: Horizon divides sky and sea/land according to golden ratio.

An example professional photographic composition:

Modeling Aesthetics: Dataset
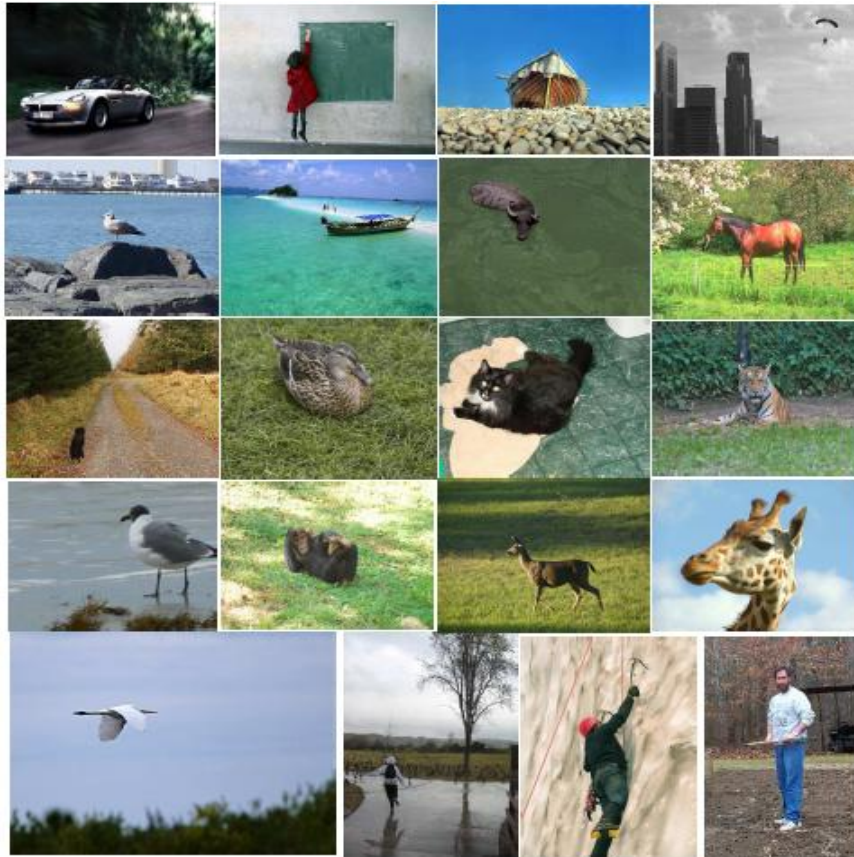
Single subject Compositions (384)   **Landscapes/Seascapes (248)**

Modeling Aesthetics: User study

1.76
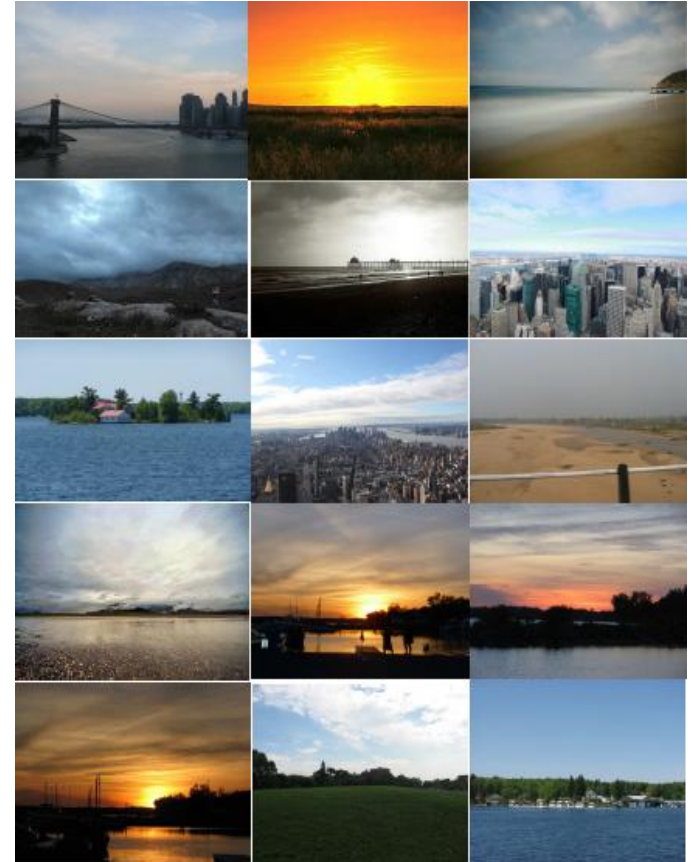
4.23

Best rated images

Poorly rated images

■ (dark red)	4<Fa<5
■ (yellow)	3<Fa<4
■ (cyan)	2<Fa<3
■ (dark blue)	1<Fa<2
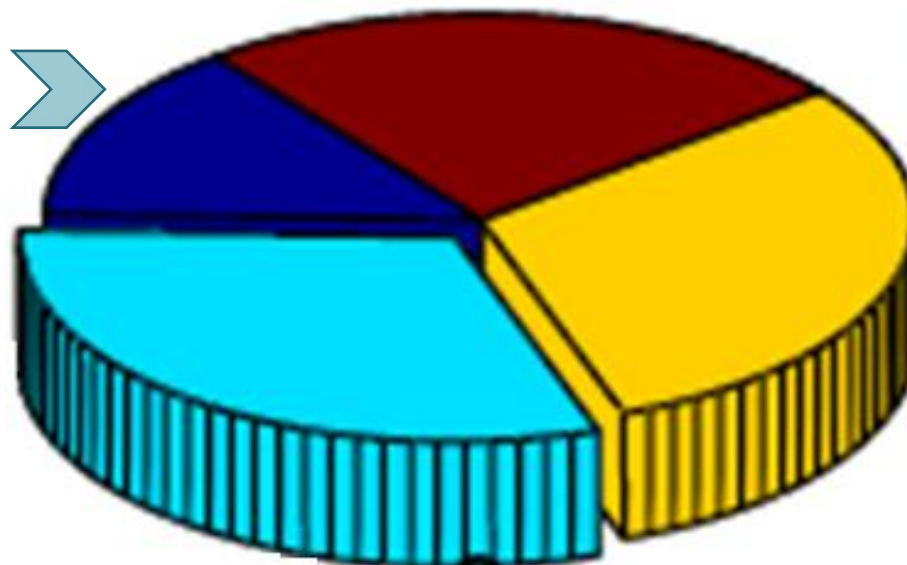
Modeling Aesthetics: Features

(a) Relative Foreground Location (Rule of Thirds)
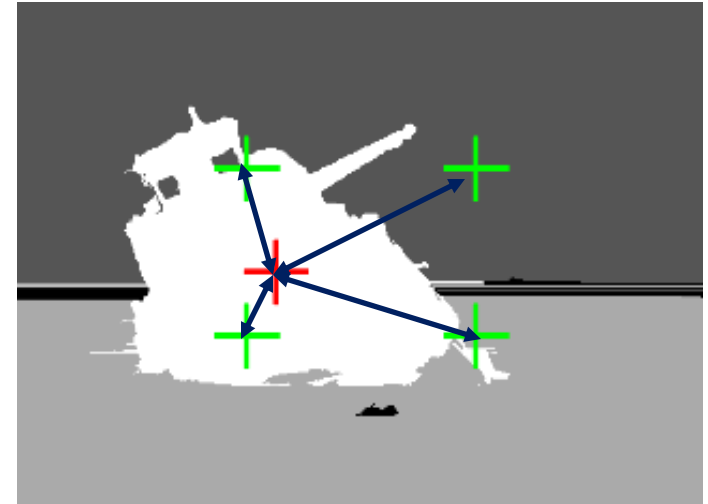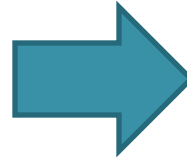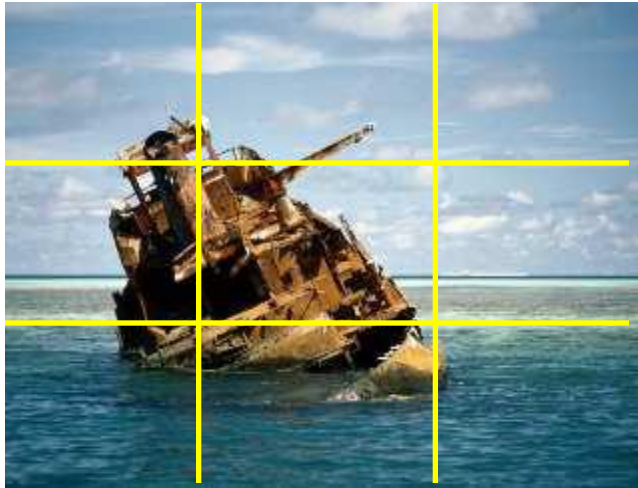
Visual Attention Center

Stress Point

$$\mathbf{F} = \frac{1}{h \times w} \left[||\mathbf{x}_0 - \mathbf{s}_1||_2, \ldots, ||\mathbf{x}_0 - \mathbf{s}_4||_2 \right]$$

Top-left	Top-Right	Bottom-Right	Bottom-Left
0.2940	0.4451	0.3365	0.0399

Modeling Aesthetics: Features

(b) Visual weight deviation from Golden Ratio (Divine Proportion)
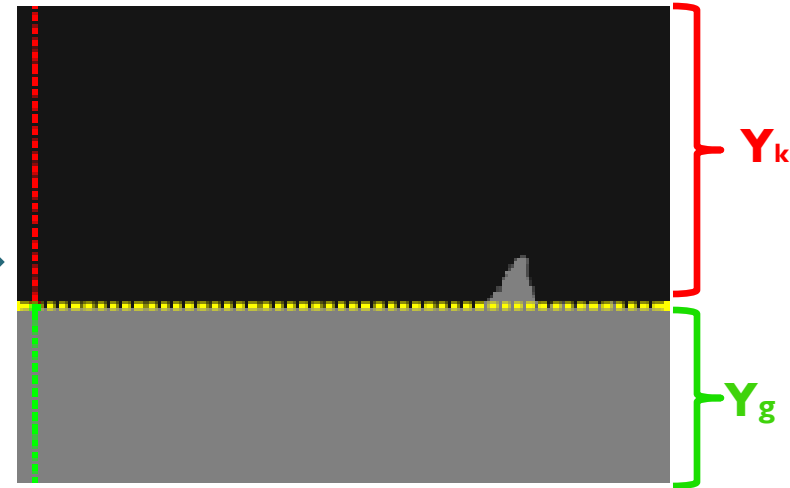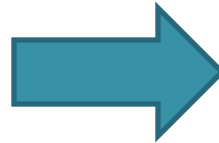
$$\mathbf{W} = \left[\left| \phi - \frac{Y_g}{Y_k} \right|, \left| \phi - \frac{Y_k}{Y_k + Y_g} \right| \right]$$

Experiments (Assessment)

Smooth mapping between Appeal factor and Aesthetic Features

Relative Foreground Location $\quad f_{rf}(F_a) : R^4 \to R, R \in \mathbf{F}$

Visual Weight Deviation $\quad f_{vw}(F_a) : R^2 \to R, R \in \mathbf{W}$

- Learn Support Vector Regression models
- Prediction accuracy:
  - Single subject compositions ~ 87%
  - Landscapes/Seascapes ~ 91%

Recomposition: Algorithm 1

Input Image

Optimal Object Placement

Spatial Recomposition

Semantic Segmentation

Single Subject?

In-painting

Labeled Elements
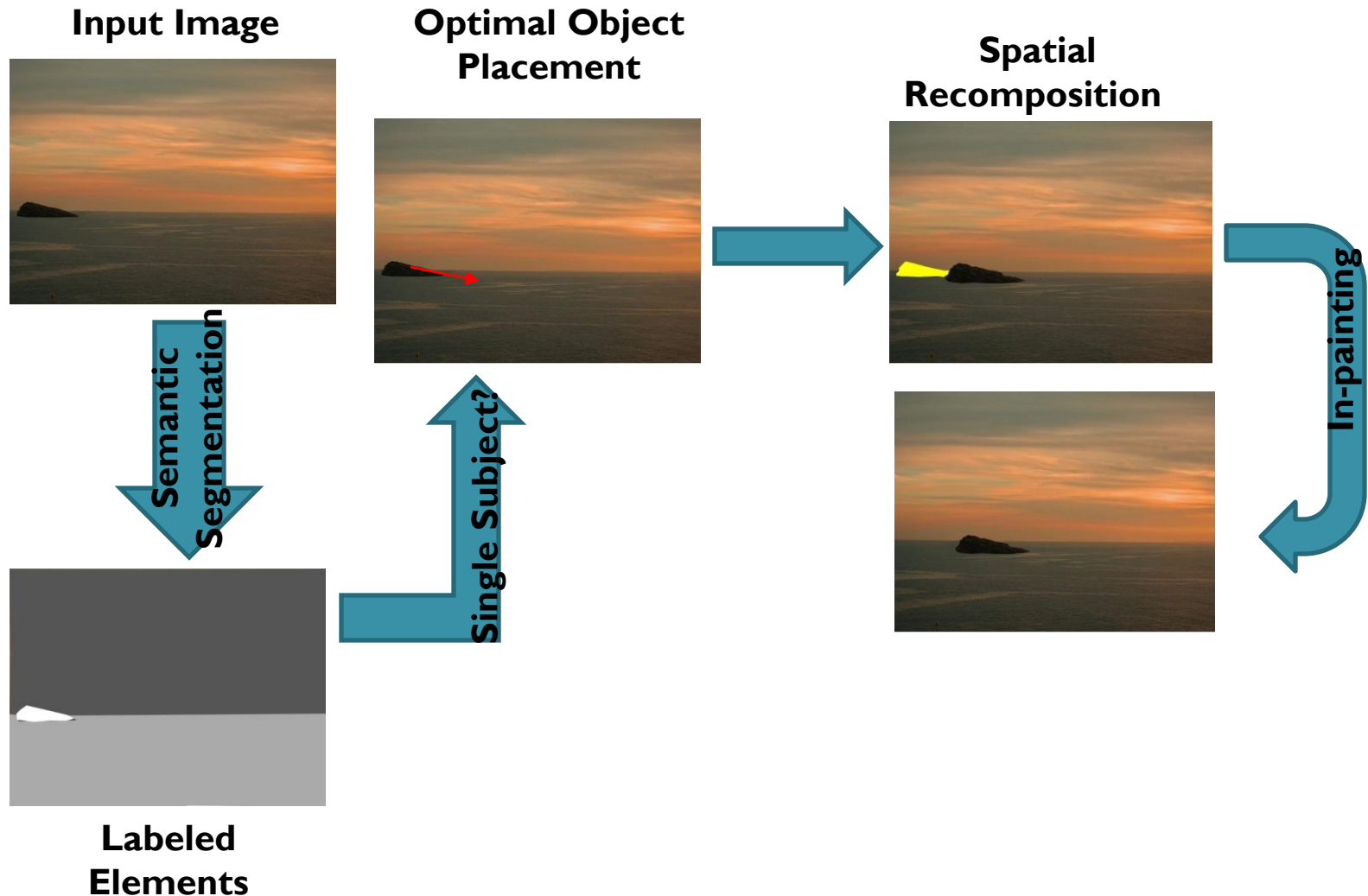
Optimal Object Placement

Labeled Image

$\psi_w(\mathbf{x_0})$

Support Neighborhood

$\mathbf{x_0}$
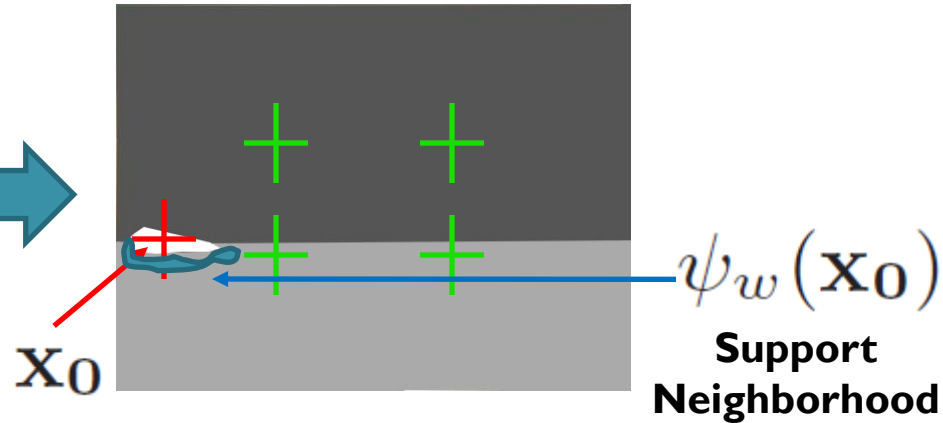
Find **x** that Maximizes Appeal

$$\arg \max_{\mathbf{x}} f_{rf}(F_a)$$ s.t. neighbors stay "like neighbors"

$$\sum_{\psi_w \forall \{R,G,B\}} ||I(\psi_w(\mathbf{x})) - I(\psi_w(\mathbf{x_0}))||_1 + \sum_{\psi_w \forall \{R,G,B\}} ||\nabla(\psi_w(\mathbf{x})) - \nabla(\psi_w(\mathbf{x_0}))||_1 < \delta$$

Intensity Term   **Gradient Term**

Optimization (Example)

PAF = 3.22

Original Image

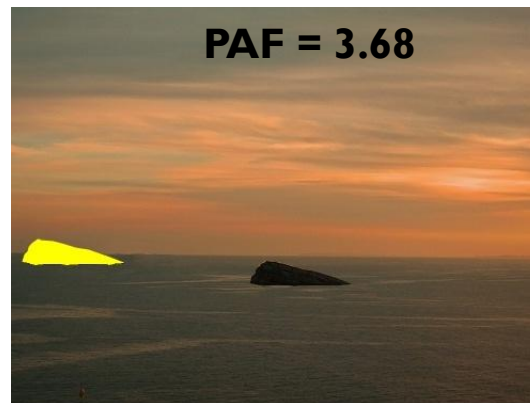Semantic constraint prevents this

Optimal Solution

PAF = 3.31

PAF = 3.68

PAF = 4.53
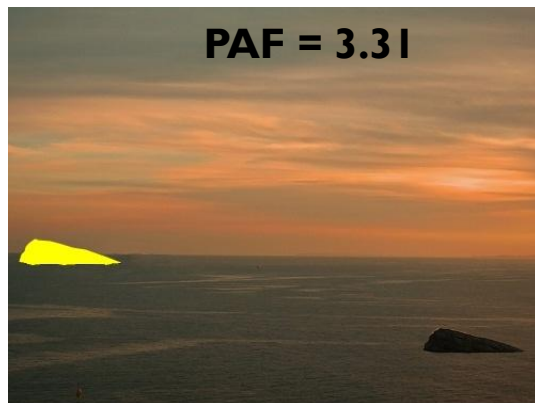
Perspective Scaling

Scaled Foreground
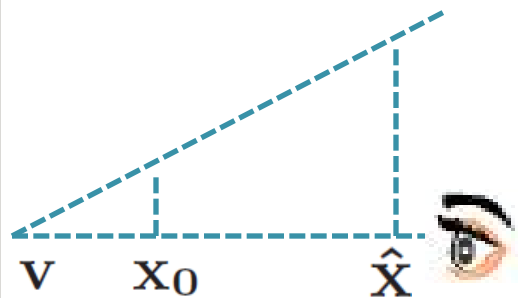
Scaling Factor $f_s = \dfrac{||\mathbf{v}, \mathbf{x_0}||_2}{||\mathbf{v}, \hat{\mathbf{x}}||_2}$ **Visual Attention Center**

Optimal location

$\mathbf{v}$ $\mathbf{x_0}$ $\hat{\mathbf{x}}$

Vanishing Point

Inpainting Foreground Hole

Inpaint Hole

Yunjun Zhang. Jiangjian Xiao. Mubarak Shah, "Region Completion in a Single Image", *EUROGRAPHICS 04*
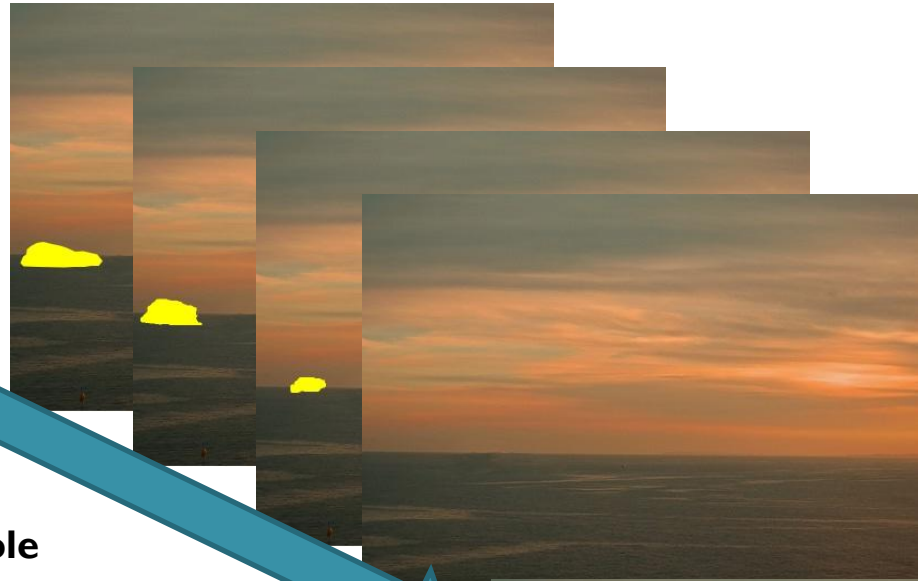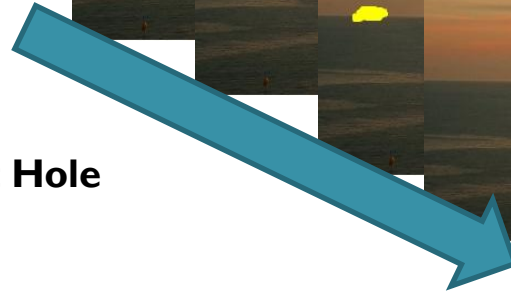
Recomposition: Algorithm 2

Input Image

Semantic Segmentation

Labeled Elements

Land/Sea scape?

Visual Weight Balancing

Optimally Crop/Expand

Balancing Visual Weights

- Ratio of Current extents $\dfrac{Y_k}{Y_g}$

- h = vertical extent of the balanced image

$$\frac{Y_k + h}{Y_g} = \frac{Y_g}{(Y_k + h) + Y_g}$$
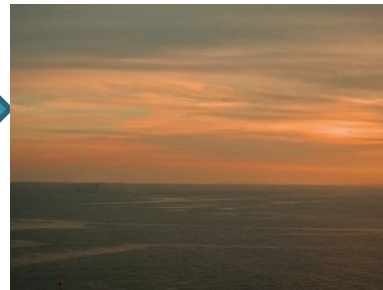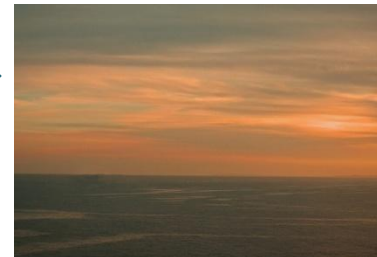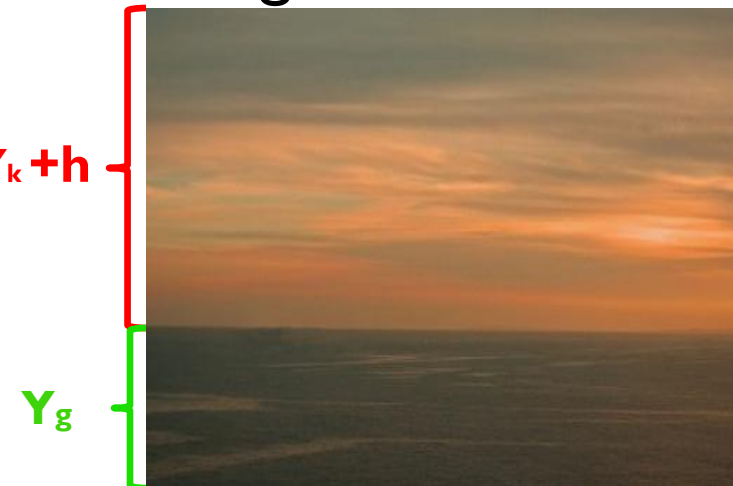
- Solve for h (sign of h determines crop/expansion)

Experimental Results

Single Subject Composition

Before Recomposition

After Recomposition

- Horse is moved to a more visually pleasing location
- Scaled appropriately
- Appeal increases by 64%

Results

PAF = 2.45

PAF = 4.29

PAF = 3.98

PAF = 4.46

Before

After

Results

PAF = 4.02

PAF = 4.34

Before

After

PAF = 3.13

PAF = 4.19

Results

Visual weight balancing

Before Recomposition

After Recomposition

- Optimally cropped support region to increase weights for sky
- Appeal factor increased by 51%

Balancing Visual weights

PAF = 4.02

PAF = 3.83

Before

After

PAF = 4.38

PAF = 3.92

Balancing Visual weights

PAF = 4.71

PAF = 4.02

Before

After

PAF = 4.49

PAF = 4.17

Not Perfect

Fa = 2.41 (Ground Truth)   PAF = 2.34
Before

Fa = 2.54 (Ground Truth)   PAF = 3.63
After

- Algorithm says nice, humans: otherwise

Summary: Optimal Placement

Increased # of Highly rated Images

Decreased # of Poorly rated Images

	4<Fa<5
	3<Fa<4
	2<Fa<3
	1<Fa<2

Before

After
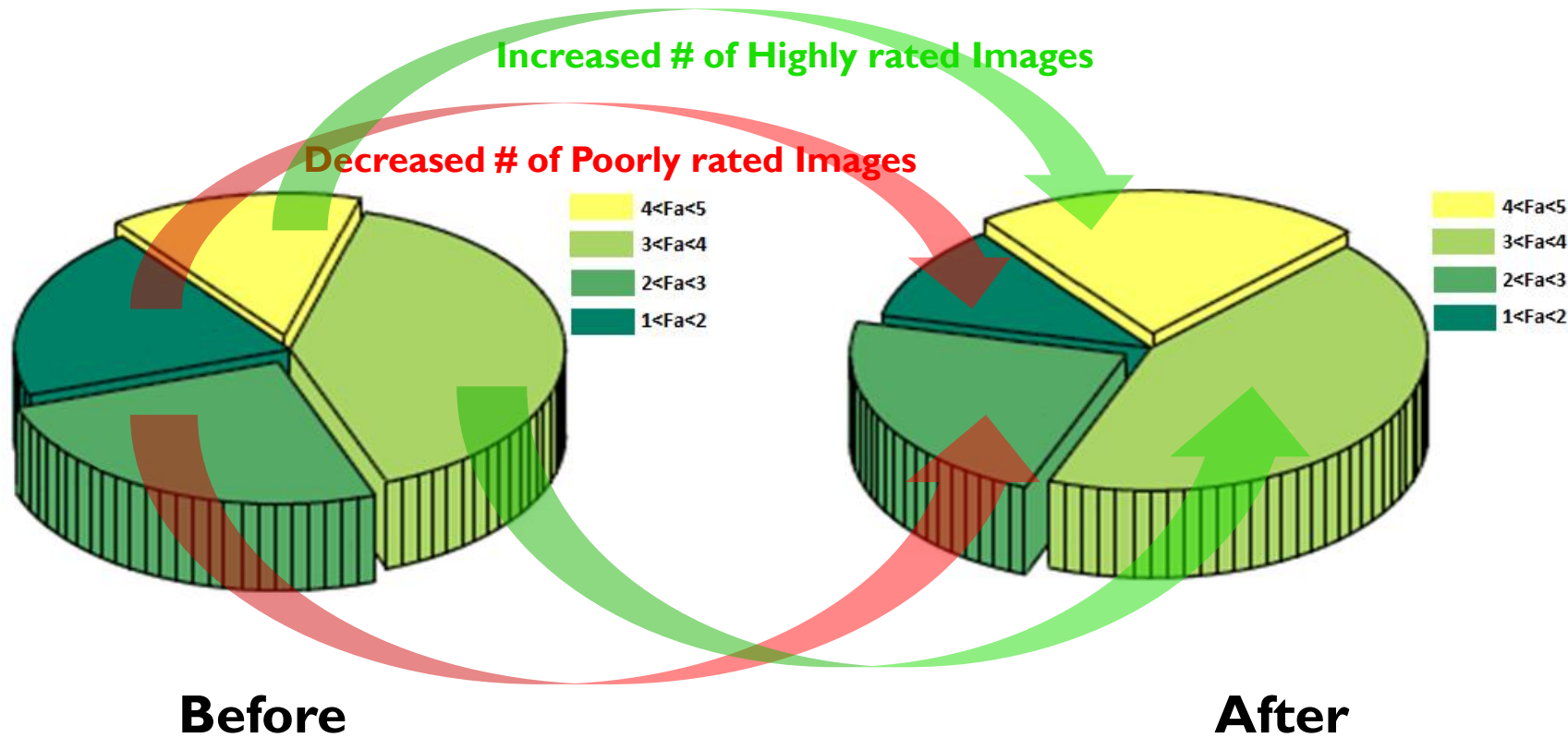
Conclusion

- Intelligent photo recomposition
- Can also be used for aesthetic filtering
- Easy to use practical tool

Questions?