ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Microsoft Research

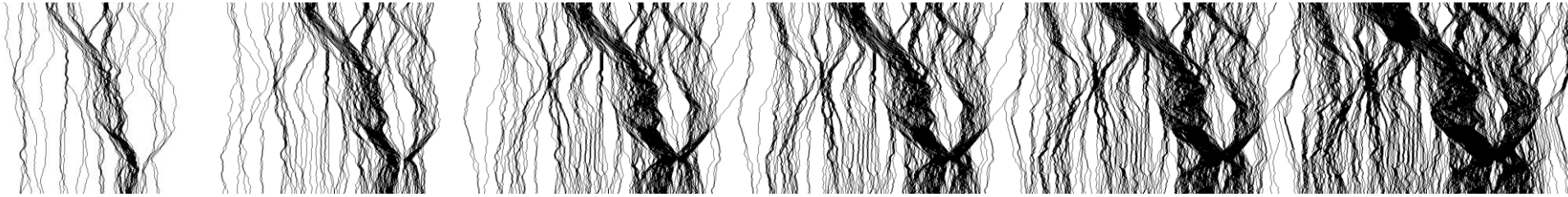
Visibility Maps for Improving Seam Carving

Alex Mansfield[1], Peter Gehler[1], Luc Van Gool[1,2] and Carsten Rother[3]

[1]Computer Vision Laboratory
ETH Zürich, Switzerland

[2]ESAT-PSI
KU Leuven, Belgium

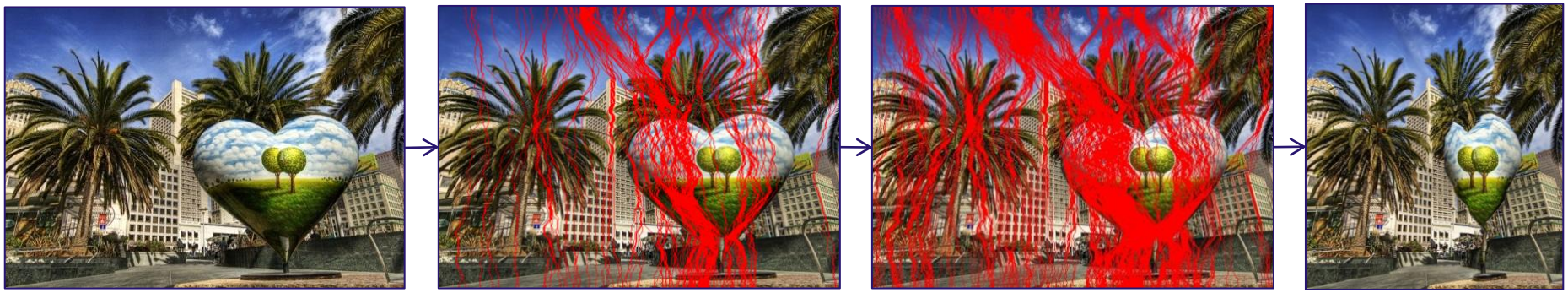
[3]Microsoft Research
Cambridge, UK



Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

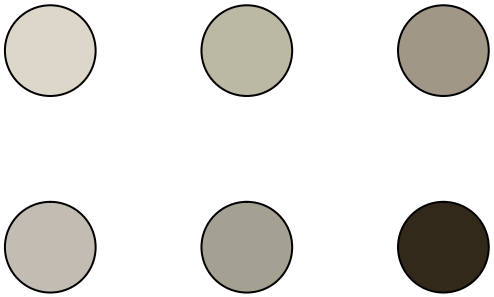
Seam carving



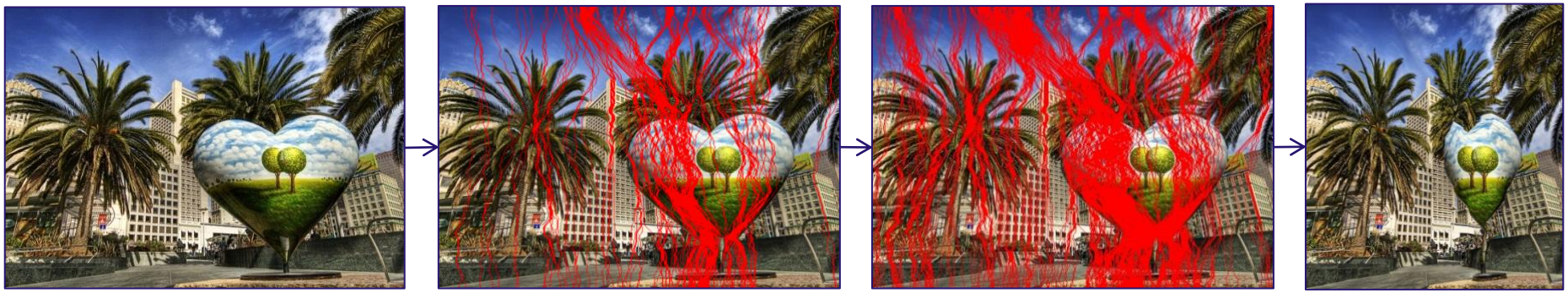
2007
Avidan and Shamir
2008
Rubinstein et al., Chen et al.
2009
Rubinstein et al., Dong et al., Han et al., Pritch et al.
2010
Grundmann et al., Zhang et al., Mansfield et al., …

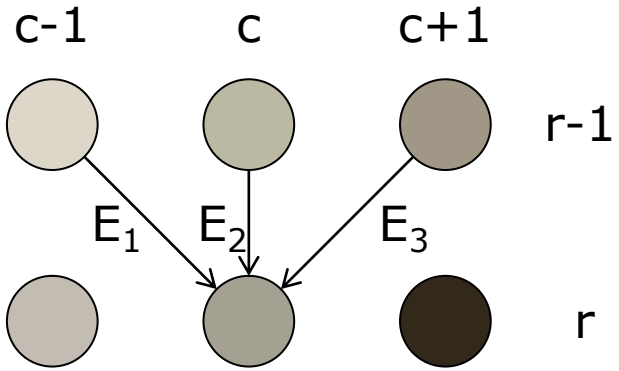
Seam carving

Seam carving



Seam carving





- **Dynamic programming**
- **Unary term**: cost of removing pixel, saliency (Avidan and Shamir 2007)
- **Contact term**: cost of bringing pixels into contact (Rubinstein et al. 2008)

Motivation

- Could seam carving be better?

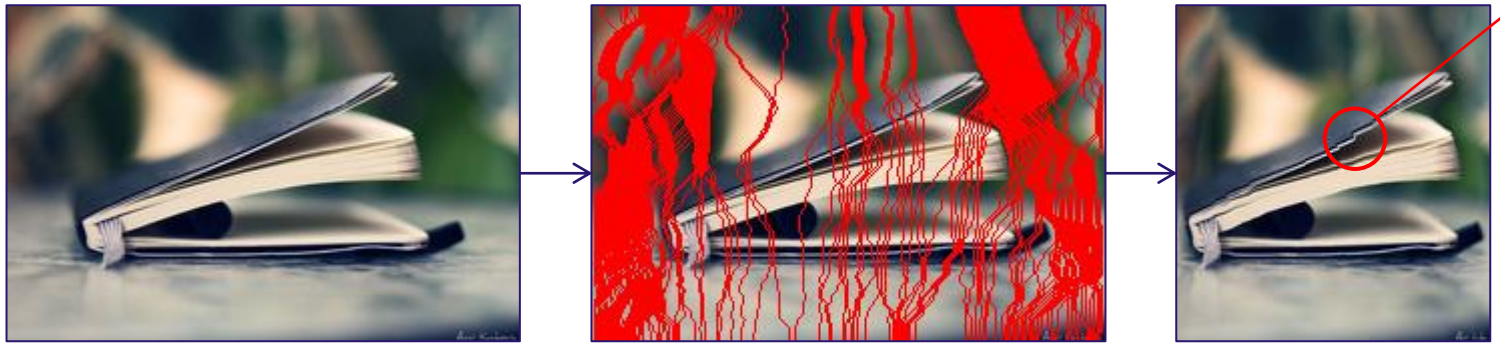
Pros

- Simple
- Processing efficient
- Memory efficient
- Produces range of retarget sizes

Cons

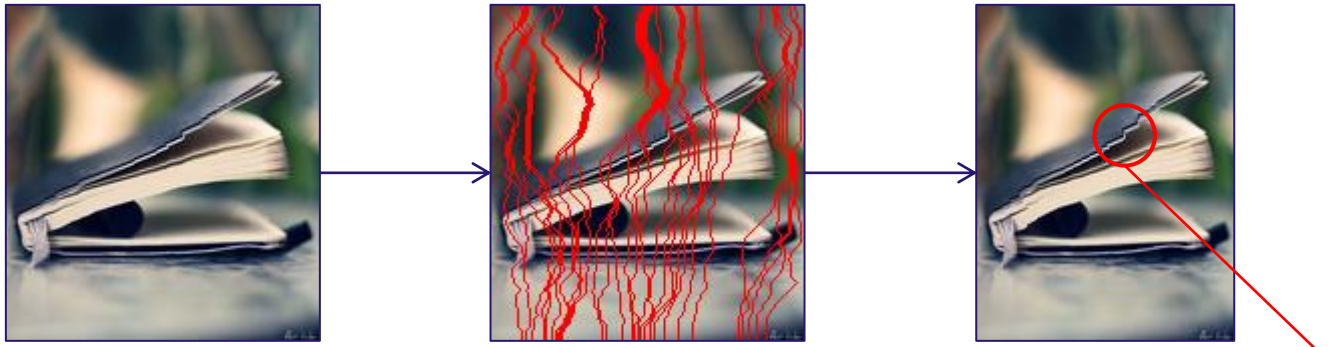
- Distortion of lines and structure
- Greedy optimization
  $\Rightarrow$ Han et al. 2009
- Seam carving with contact term does not optimize an energy defined directly between input and output

Energy minimized by seam carving



Artefact created

For next iterations, this is start image



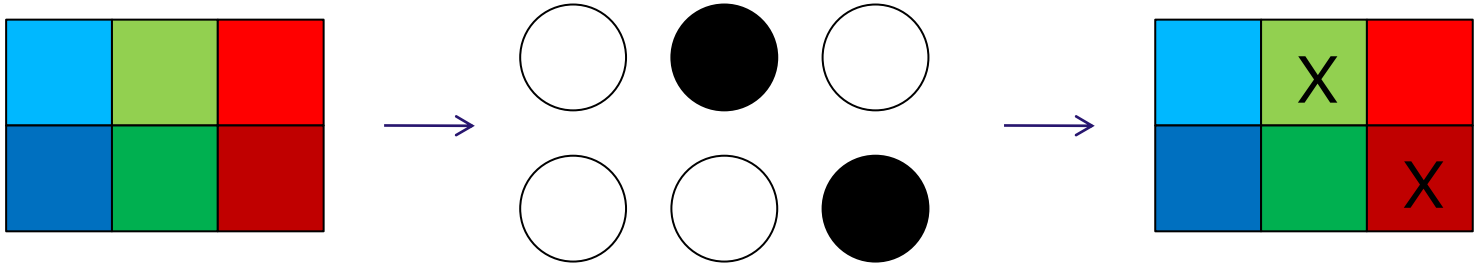
Artefact preserved

Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Visibility map

$$X_{r,c} \in \{0,1\}$$

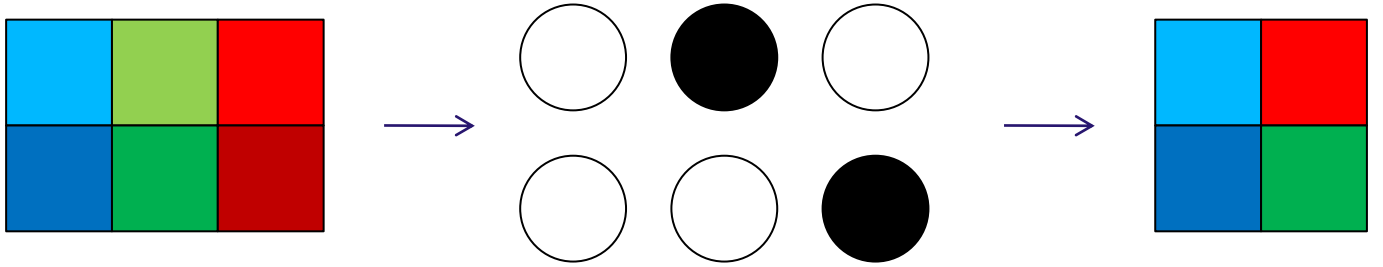


Input image

Visibility map

Visibility map

$$X_{r,c} \in \{0,1\}$$

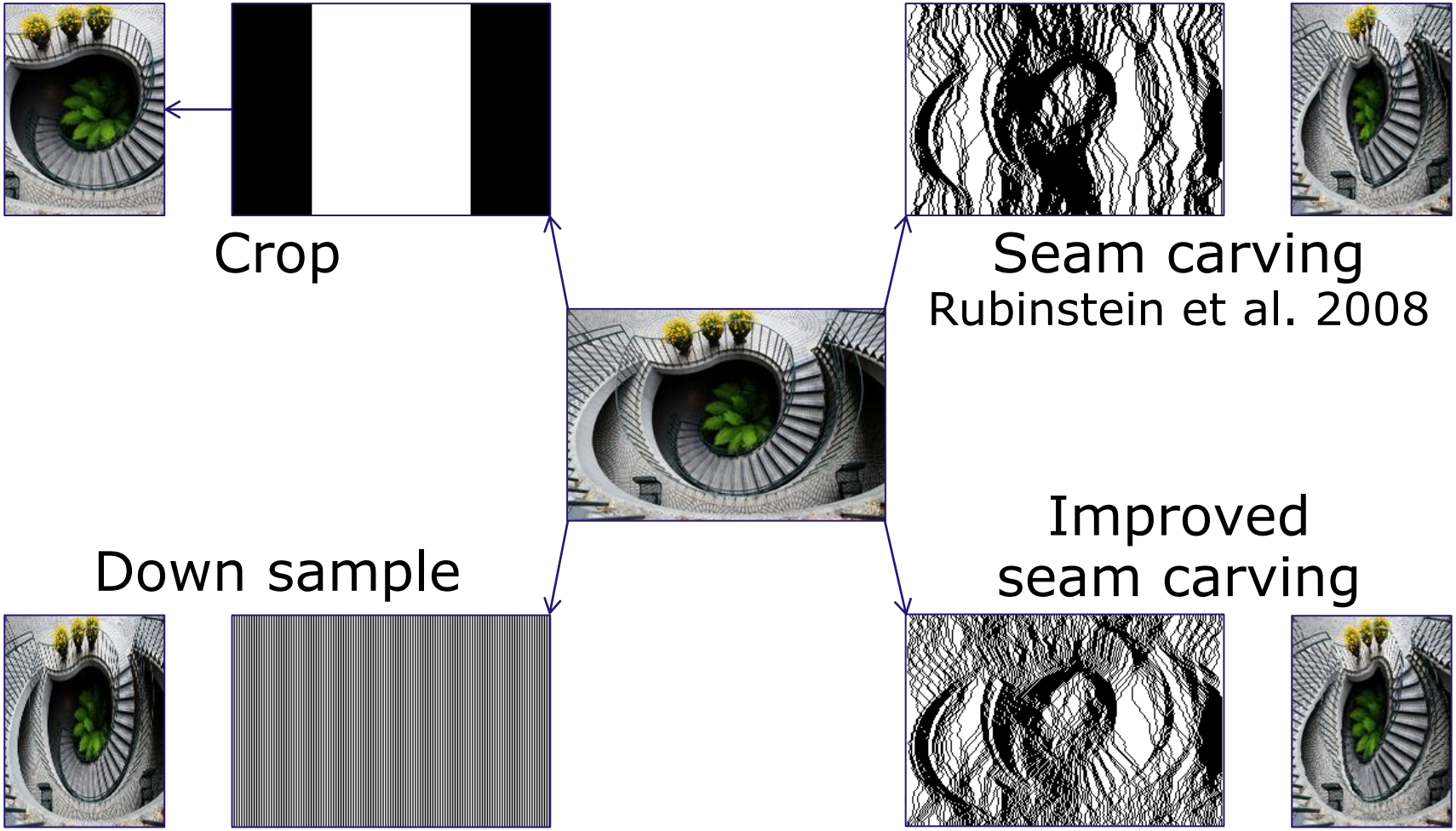


Input image

Visibility map

Output image

Visibility map



Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Energy model

Visibility map $X_{r,c} \in \{0,1\}$

$$E(\mathbf{X}) = \sum_{r,c} \psi^{\mathrm{U}}_{r,c}(\mathbf{X}) + \sum_{r,c_{\mathrm{l}}<c_{\mathrm{r}}} \psi^{\mathrm{H}}_{r,c_{\mathrm{l}},c_{\mathrm{r}}}(\mathbf{X}) + \sum_{r>1,c_{\mathrm{u}},c_{\mathrm{d}}} \psi^{\mathrm{V}}_{r,c_{\mathrm{u}},c_{\mathrm{d}}}(\mathbf{X})$$

Unary terms

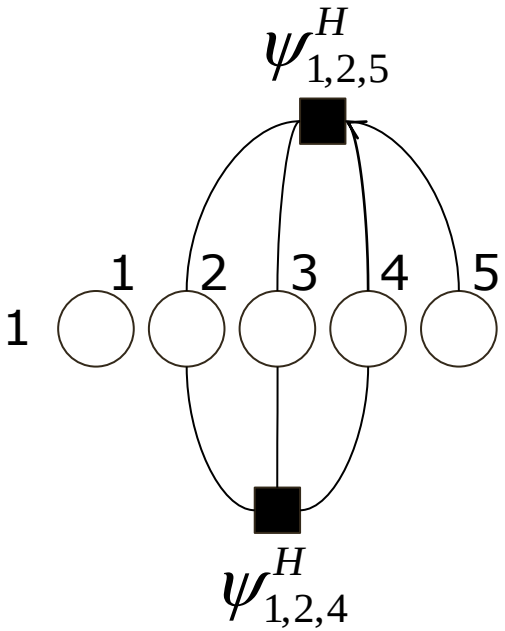
Higher order cliques

$$\psi^{\mathrm{U}}_{r,c}(X_{r,c}) = E^{\mathrm{U}}_{r,c}[X_{r,c} \neq 0]$$

Sparse structure

[.] is indicator function

Higher order cliques

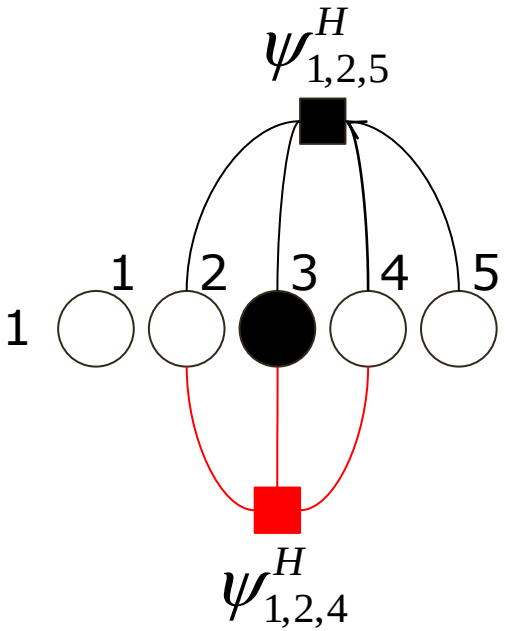


Horizontal contact term

$$\psi^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}(X_{r,c_\mathrm{l}},\ldots,X_{r,c_\mathrm{r}}) = \begin{cases} E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}, & X_{r,\{c_\mathrm{l},c_\mathrm{r}\}} = 1, X_{r,\{c_\mathrm{l}+1,\ldots,c_\mathrm{r}-1\}} = 0 \\ 0, & \text{otherwise} \end{cases}$$

E^H_{r,c_l,c_r} only a function of visibility map row r

Higher order cliques

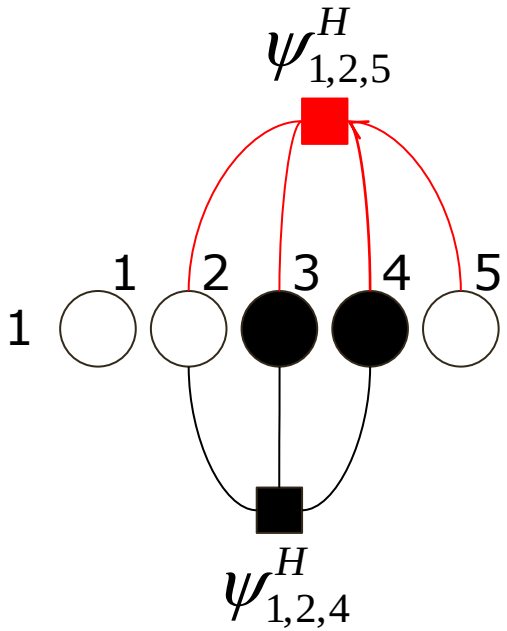


Horizontal contact term

$$\psi^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}(X_{r,c_\mathrm{l}},\ldots,X_{r,c_\mathrm{r}}) = \begin{cases} E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}, & X_{r,\{c_\mathrm{l},c_\mathrm{r}\}} = 1, X_{r,\{c_\mathrm{l}+1,\ldots,c_\mathrm{r}-1\}} = 0 \\ 0, & \text{otherwise} \end{cases}$$

E^H_{r,c_l,c_r} only a function of visibility map row r

Higher order cliques

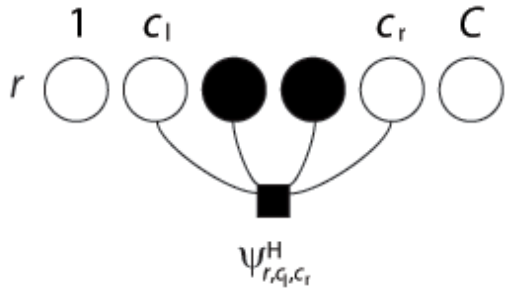


Horizontal contact term

$$\psi^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}(X_{r,c_\mathrm{l}},\ldots,X_{r,c_\mathrm{r}}) = \begin{cases} E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}, & X_{r,\{c_\mathrm{l},c_\mathrm{r}\}} = 1, X_{r,\{c_\mathrm{l}+1,\ldots,c_\mathrm{r}-1\}} = 0 \\ 0, & \text{otherwise} \end{cases}$$

E^H_{r,c_l,c_r} only a function of visibility map row r

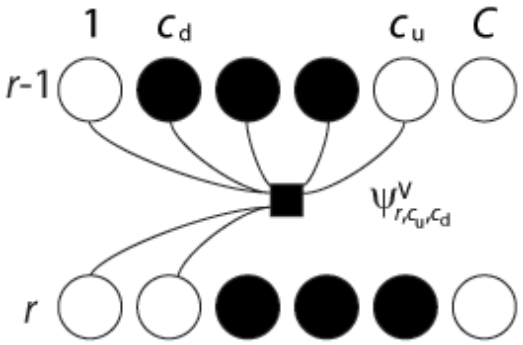
Higher order cliques



Horizontal contact term

$$\psi^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}(X_{r,c_\mathrm{l}},\ldots,X_{r,c_\mathrm{r}}) = \begin{cases} E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}, & X_{r,\{c_\mathrm{l},c_\mathrm{r}\}} = 1, X_{r,\{c_\mathrm{l}+1,\ldots,c_\mathrm{r}-1\}} = 0 \\ 0, & \text{otherwise} \end{cases}$$

E^H_{r,c_l,c_r} only a function of visibility map row r



Vertical contact term

$$\psi^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}}(X_{r-1,1},\ldots,X_{r-1,c_\mathrm{u}},X_{r,1},\ldots,X_{r,c_\mathrm{d}}) =$$

$$\begin{cases} E^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}}, & X_{r-1,c_\mathrm{u}} = 1, X_{r,c_\mathrm{d}} = 1, \sum_{c=1}^{c_\mathrm{u}-1} X_{r-1,c} = \sum_{c=1}^{c_\mathrm{d}-1} X_{r,c} \\ 0, & \text{otherwise} \end{cases}.$$

E^V_{r,c_u,c_d} only a function of visibility map rows $r-1$ and r

Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Energy terms

Unary term

$$E^{\mathrm{U}}_{r,c} = \left(\left| \left(\frac{\partial}{\partial x} \mathbf{I} \right)_{r,c} \right| + \left| \left(\frac{\partial}{\partial y} \mathbf{I} \right)_{r,c} \right| \right)^{n_{\mathrm{U}}}$$

Unary term of Avidan and Shamir 2007

Energy terms

$$E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} + S^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}$$
$$E^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} + S^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}}$$

Distortion terms

1. Magnitude distance

$$D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = \left| |I|_{r,c_\mathrm{l}} - |I|_{r,c_\mathrm{r}} \right|^{n_\mathrm{D}}$$
$$D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = \left| |I|_{r-1,c_\mathrm{u}} - |I|_{r,c_\mathrm{d}} \right|^{n_\mathrm{D}}$$

Measure created distortion

2. RGB distance

$$D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = \sum_{x\in\{\mathrm{R,G,B}\}} \left| I^x_{r,c_\mathrm{l}} - I^x_{r,c_\mathrm{r}} \right|^{n_\mathrm{D}}$$
$$D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = \sum_{x\in\{\mathrm{R,G,B}\}} \left| I^x_{r-1,c_\mathrm{u}} - I^x_{r,c_\mathrm{d}} \right|^{n_\mathrm{D}}$$

3. Relative RGB distance

$$D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = \sum_{x\in\{\mathrm{R,G,B}\}} \left| I^x_{r,c_\mathrm{l}} - I^x_{r,c_\mathrm{r}-1} \right|^{n_\mathrm{D}} + \left| I^x_{r,c_\mathrm{r}} - I^x_{r,c_\mathrm{l}+1} \right|^{n_\mathrm{D}}$$
$$D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = \sum_{x\in\{\mathrm{R,G,B}\}} \left| I^x_{r-1,c_\mathrm{u}} - I^x_{r-1,c_\mathrm{d}} \right|^{n_\mathrm{D}} + \left| I^x_{r,c_\mathrm{d}} - I^x_{r,c_\mathrm{u}} \right|^{n_\mathrm{D}}$$

Energy terms

$$E^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} + S^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}$$
$$E^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} + S^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}}$$

Seam terms

1. Repeat cost

$$S^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = (c_\mathrm{r} - c_\mathrm{l} - 1) D^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}}$$
$$S^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = (c_\mathrm{u} - c_\mathrm{d} - 1) D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}}$$

2. Average unary cost

$$S^{\mathrm{H}}_{r,c_\mathrm{l},c_\mathrm{r}} = \frac{(c_\mathrm{r} - 1) - (c_\mathrm{l} + 1)}{(c_\mathrm{r} - 1) - (c_\mathrm{l} + 1) + 1} \sum_{c=c_\mathrm{l}+1}^{c_\mathrm{r}-1} E^{\mathrm{U}}_{r,c}$$
$$S^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} = (c_\mathrm{u} - c_\mathrm{d} - 1) D^{\mathrm{V}}_{r,c_\mathrm{u},c_\mathrm{d}} \quad .$$

Regularize spatial distribution of seams to prevent ‘clumping’

Energy model

Unary terms

$$\psi^{\mathrm{U}}_{r,c}(X_{r,c}) = E^{\mathrm{U}}_{r,c}[X_{r,c} \neq 0]$$

$$E(\mathbf{X}) = \sum_{r,c} \psi^{\mathrm{U}}_{r,c}(\mathbf{X}) + \sum_{r,c_{\mathrm{l}}<c_{\mathrm{r}}} \psi^{\mathrm{H}}_{r,c_{\mathrm{l}},c_{\mathrm{r}}}(\mathbf{X}) + \sum_{r>1,c_{\mathrm{u}},c_{\mathrm{d}}} \psi^{\mathrm{V}}_{r,c_{\mathrm{u}},c_{\mathrm{d}}}(\mathbf{X})$$

Non-zero value $E^{H}_{r,c_l c_r}$

Non-zero value $E^{V}_{r,c_u c_d}$

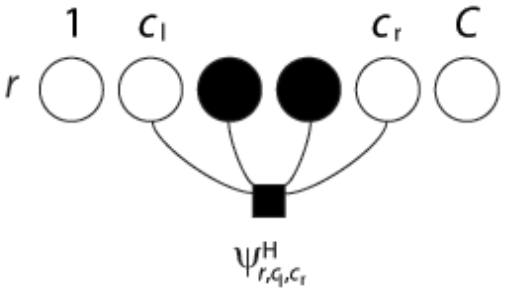
Sparse higher order cliques

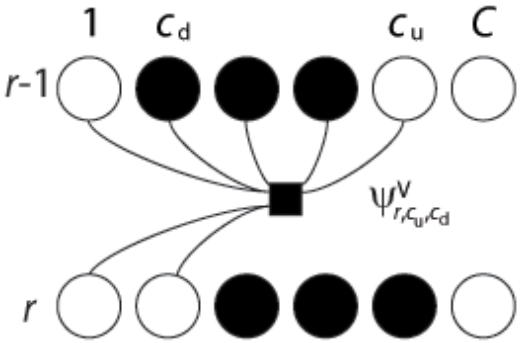
Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Optimization

- Higher order cliques built over pixels on one or two rows





Optimization

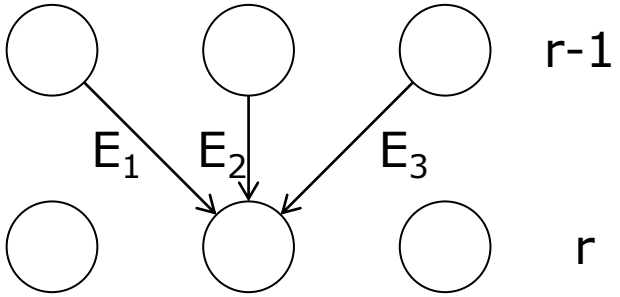
- Higher order cliques built over pixels on one or two rows
- Non-zero values of cliques depend on visibility map of same one or two rows

$E^H_{r,c_l c_r}$ only function of visibility map row *r*

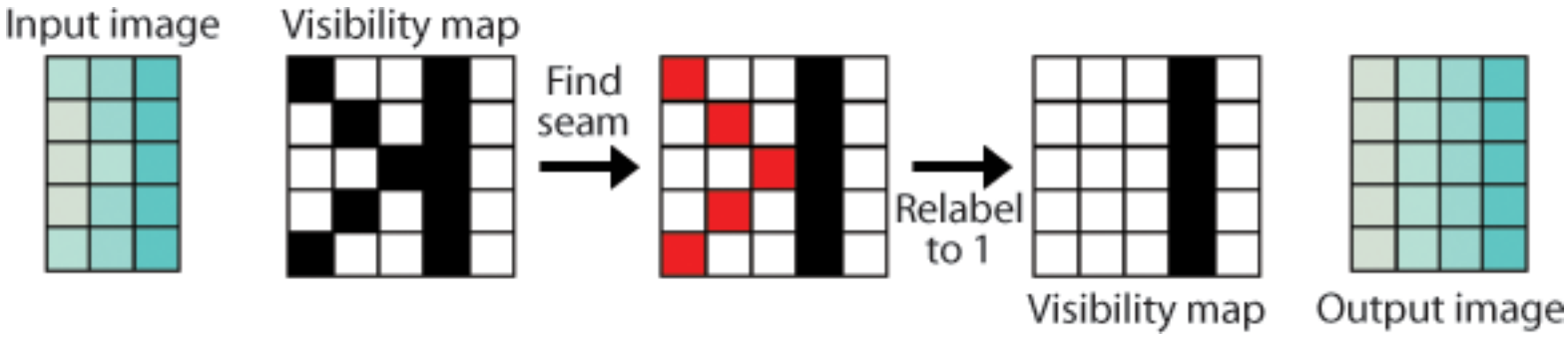
$E^V_{r,c_u c_d}$ only function of visibility map rows *r* and *r-1*

Optimization

- Higher order cliques built over pixels on one or two rows
- Non-zero values of cliques depend on visibility map of one or two rows
- ⇒ **can optimize with dynamic programming for seam to remove that minimizes this energy**



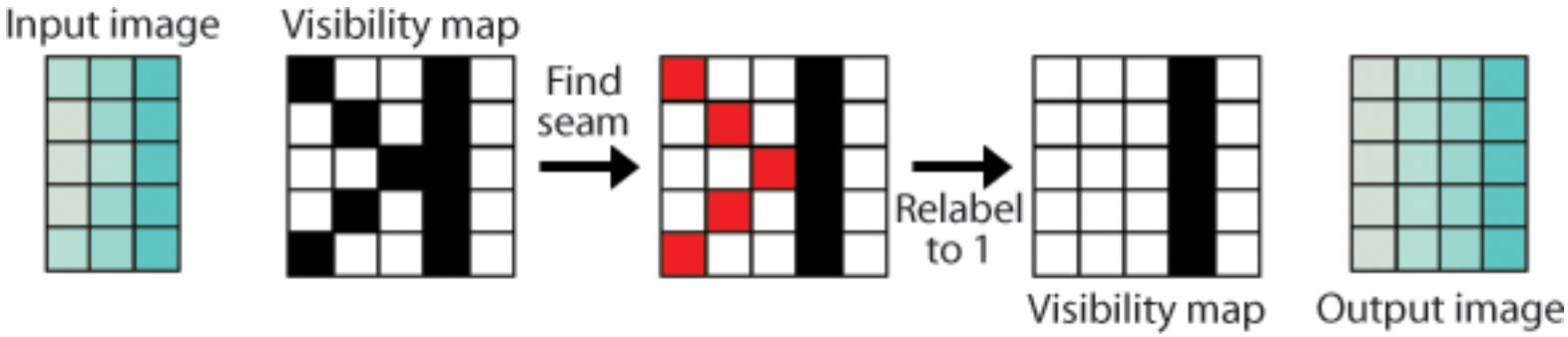
Refinement



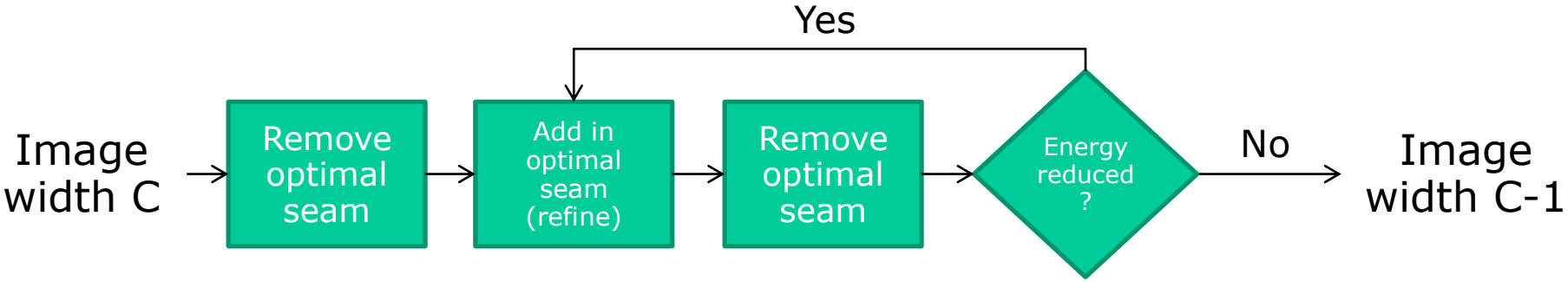
- Same algorithm, but labelling to 1 (visible) instead of labelling to 0 (non-visible)

Without refinement

Refinement



- Same algorithm, but labelling to 1 (visible) instead of labelling to 0 (non-visible)

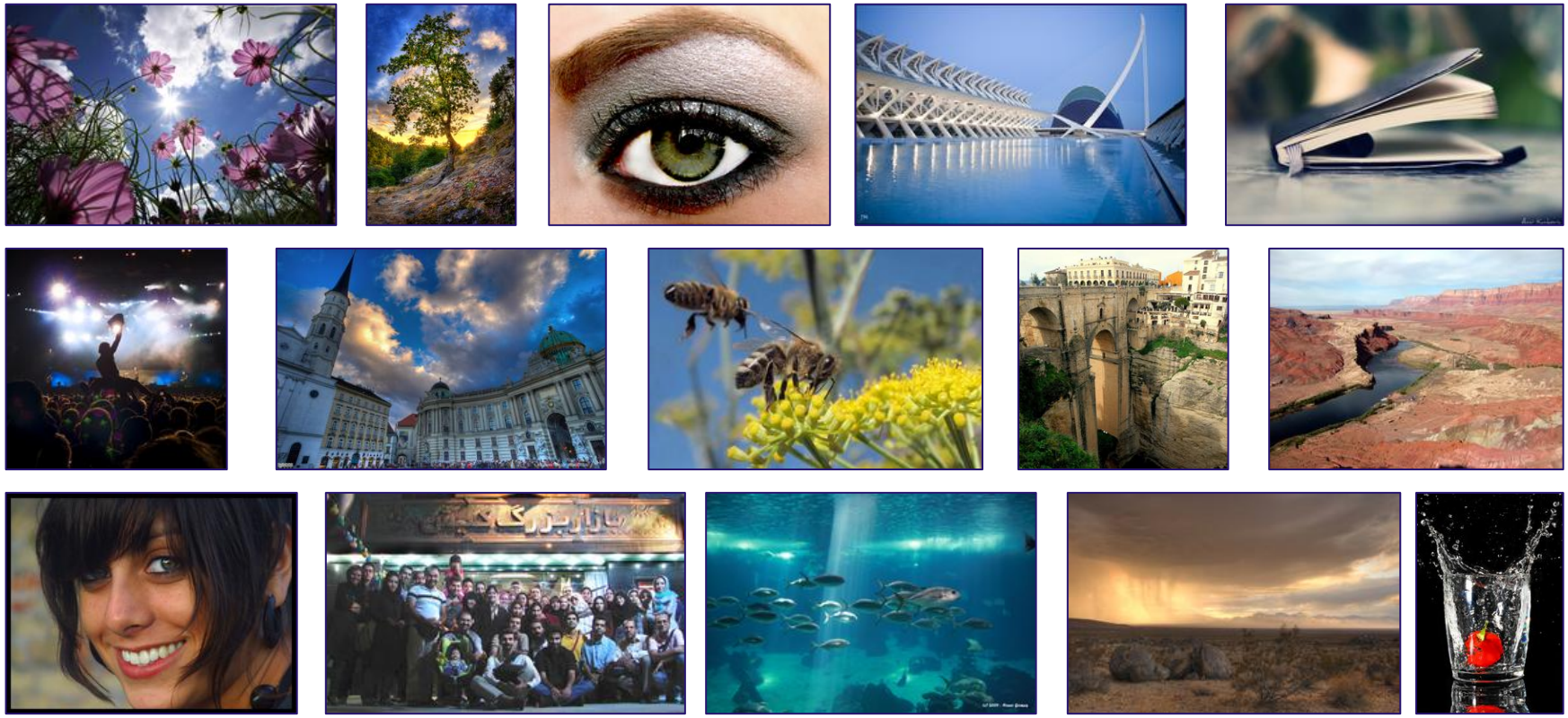


With refinement

Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Dataset



- 100 images from flickr
- Will be available online www.vision.ee.ethz.ch/~mansfiea/improvingsc/
- Performed tests with 288 parameter combinations

Improved Seam Carving

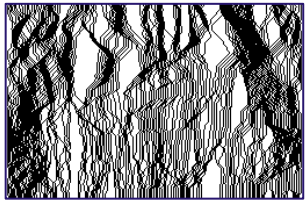
- We found bidirectional similarity measures to correlate poorly with human judgement
- By human judgement found the best parameters to be:
  - Unary of Avidan and Shamir 2007
  - Contact term of Rubinstein et al. 2008
  - Average unary cost seam term
  - No refinement

Results

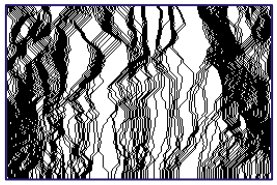
Input image

Improved SC

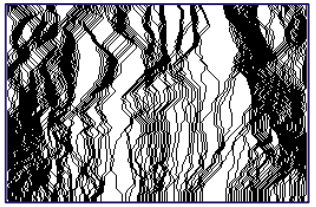
No unary



No seam term



+ refinement

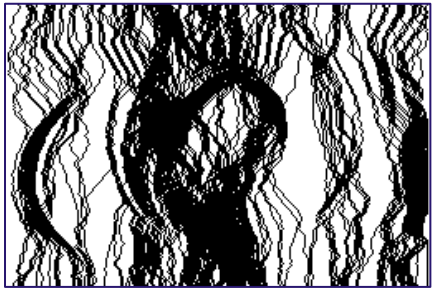


Reduced energy in 68.2% of results, average increase 2.3%

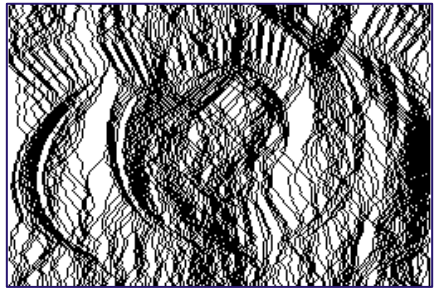
Results - Improvement

Input image

Seam carving
Rubinstein et al. 2008



Improved seam carving

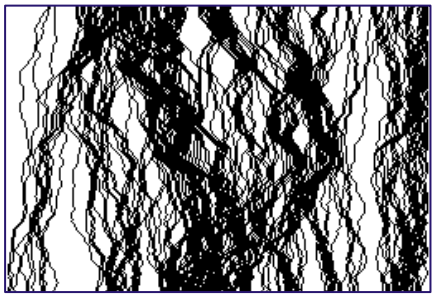


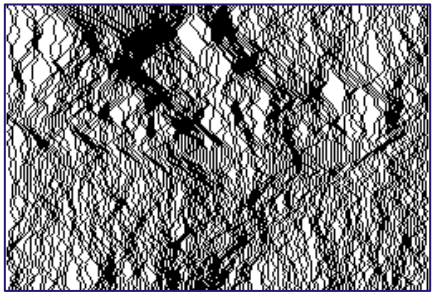
Results – No improvement

Input image

Seam carving
Rubinstein et al. 2008

Improved seam carving

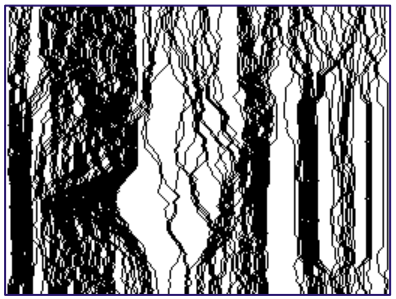




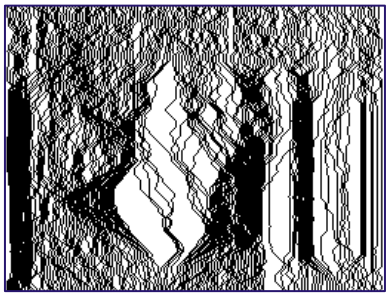
Results - Worse

Input image

Seam carving
Rubinstein et al. 2008



Improved seam carving



Outline

- Motivation
- Visibility map
- General energy model for visibility map
- Energy terms
- Optimization
- Results
- Conclusions

Conclusions

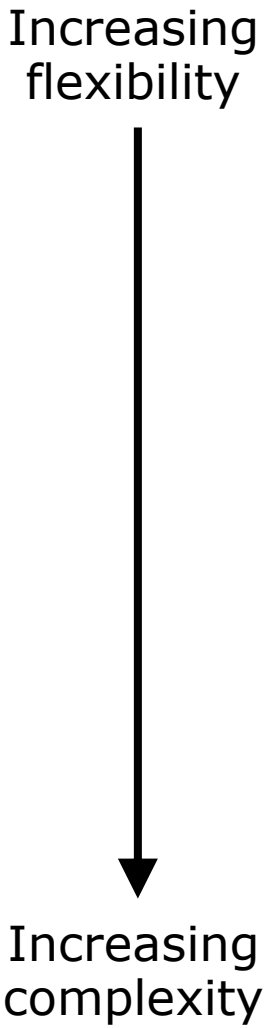
- Defined a general form of energy over a visibility map
- Increased flexibility of modeling and optimization
- Can be optimized with the same complexity as seam carving

Future directions

- Why is retargeting not yet ubiquitous?
  - No guarantees on quality of result
  - Speed
  - Insufficient user control
- Future approaches
  - User input
    (e.g. Barnes et al. 2009, Mansfield et al. 2010)
  - Structure detection
  - Lots of data
    (such as in Hays and Efros 2007)
  - Hierarchical approaches
    (e.g. Pritch et al. 2009)

Future directions

- Pixel removal
  - Avidan and Shamir 2007
- Pixel re-arrangement
  - Pritch et al. 2009
- Image warping
  - E.g. Wang et al. 2008, Krähenbühl et al. 2009
- Pixel estimation from generative model
  - Simakov et al. 2008



All code and images will be available at

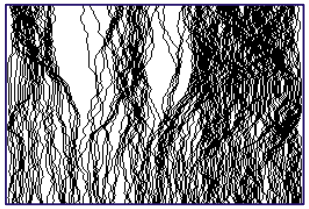
http://www.vision.ee.ethz.ch/~mansfiea/improvingsc/

ANY QUESTIONS?

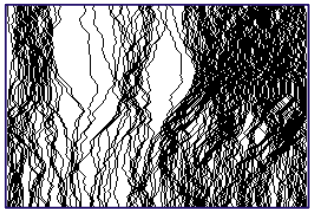
Acknowledgements
Images from Flickr users Stuck in Customs, Amir K., Telmo32, code poet, pheanix300, J, drurydrama, Dimit®I, druidabruxux, Sverrir Thor, Rikki Nadir, carlotardani, antonychammond, jimpg2_2010, amanky, Tambako the Jaguar, Reinante El Pintor de Fuego, papalars, *Zephyrance, Jenna Carver, Seattle Miles, Paulo Brandão, James Jordan, gin_able, Orangeya, pareeerica, Panoramas, stephcarter, Chuck "Caveman" Coker, Pierluigi Riccio, yamiq, Gianni D., Gabriela Camerotti, muha…, wolfpix, Nrbelex, papalars, Bill & Mavis, Elina Nilsson, Hamed Saber, visualpanic, mikebaird, joiseyshowaa, SF Brit, Mr Theklan, ~dolfi, Jule_Berlin, liber, chris bartnik photography, Claude Renault, yewenyi, Brandon Christopher Warren, Jaye, Kıvanç Niş, robokow, Al_HikesAZ, Trois Têtes, law_kewen, Michal Osmenda, Ben Chau, Dark L. Hacedor, mathias-erhart, N Gomes, millzero, aussiegall, pumpkincat210, Rev Dan Catt, Eneas, Eduardo Amorim, Cláudia*~Assad, .: sandman, Wagman_30, paul (dex), tochis, John Steven Fernandez, crsan, Anna Gay, SergioTudela, Thomas Hawk, ecstaticist, MorBCN, magdalar

Results

Input image

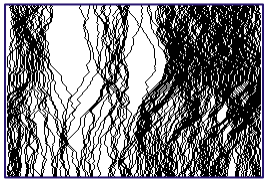


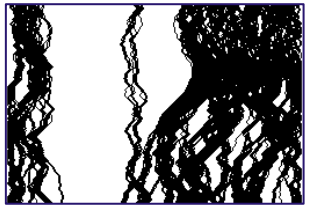
No unary



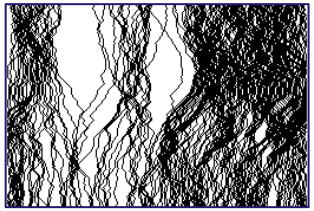
+ blending

Good terms



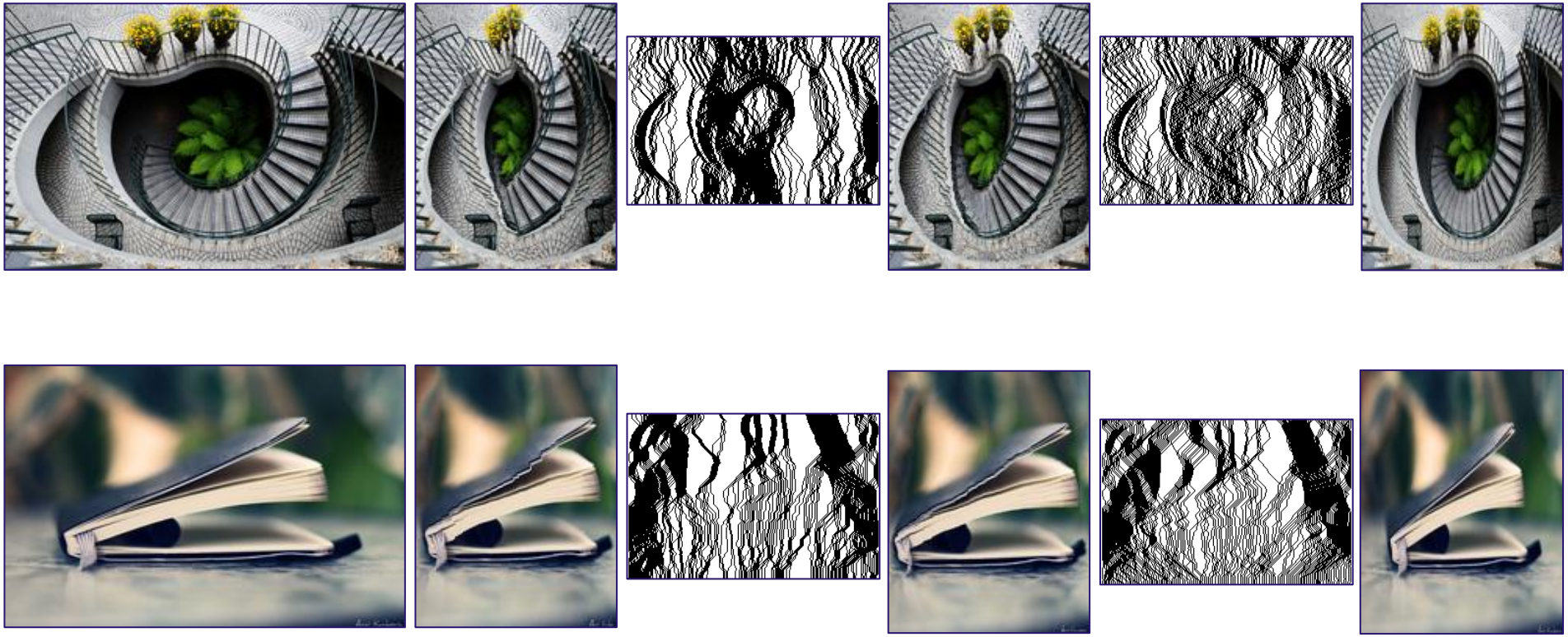


No shift term



+ refinement

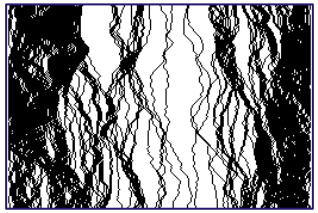
Results

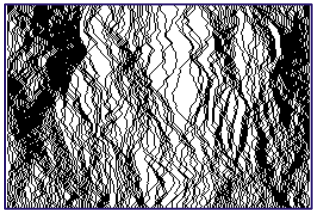


Input image | Seam carving | Seam carving visibility map | Improved seam carving | Improved seam carving visibility map | Scaling

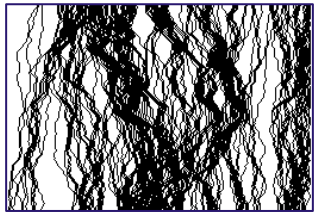
Results

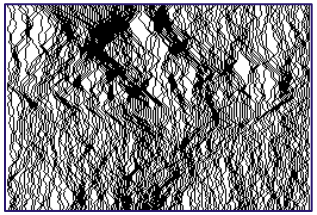











Input image

Seam carving

Seam carving visibility map

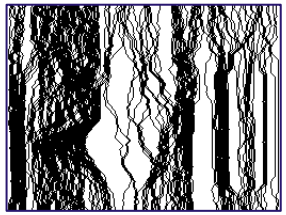
Improved seam carving

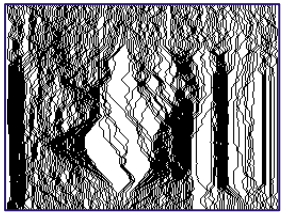
Improved seam carving visibility map

Scaling

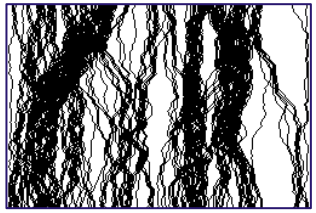
Results

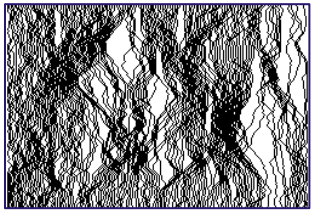











Input image

Seam carving

Seam carving visibility map

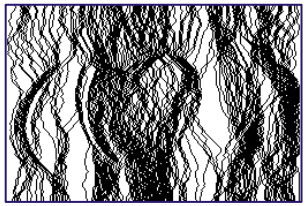
Improved seam carving

Improved seam carving visibility map

Scaling

Results

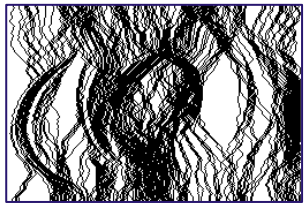
Input image



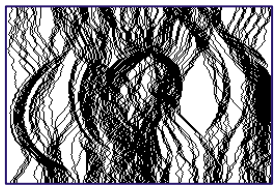
No unary

Good terms

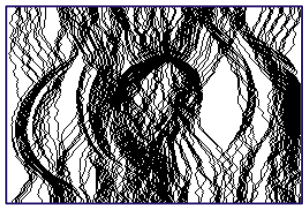
+ blending



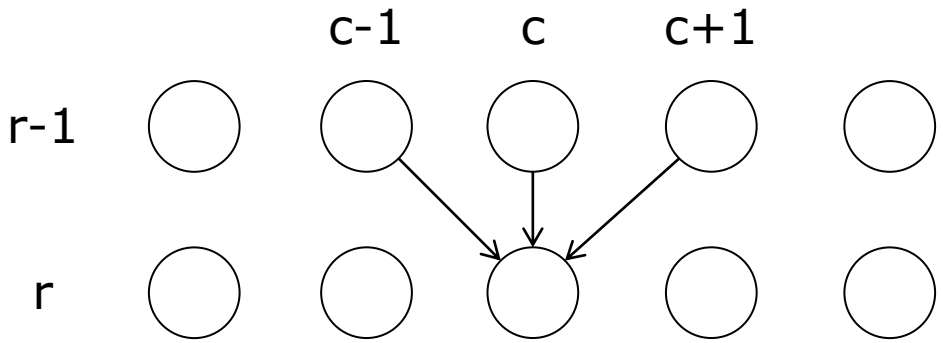
No shift term



+ refinement



Dynamic programming optimization



The energy term of each path is conditioned on the state of (r, c) and either $(r-1, c-1)$, $(r-1, c)$ or $(r-1, c+1)$.

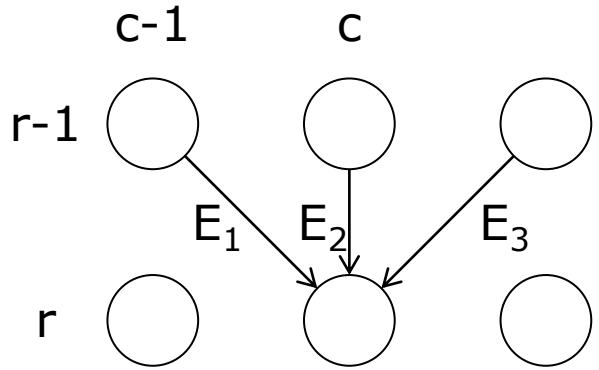
As only one pixel changes state per row, **the visibility map on rows $r-1$ and r is known**.

ψ^H_{r,c_l,c_r} depends on (r, c_l) to (r, c_r)
E^H_{r,c_l,c_r} depends on $(r, 1)$ to (r, R)
ψ^V_{r,c_u,c_d} depends on $(r, 1)$ to (r, c_d) and $(r-1, 1)$ to $(r-1, c_u)$
E^V_{r,c_u,c_d} depends on $(r, 1)$ to (r, R) and $(r-1, 1)$ to $(r-1, R)$

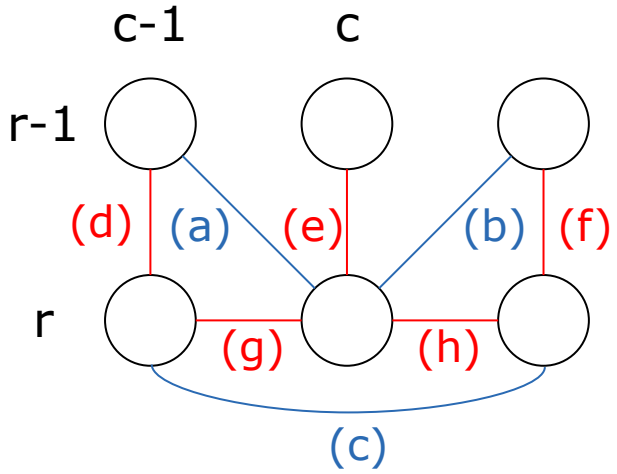
Hence all terms that are affected by this choice of path are known and the energy can be determined.

Dynamic programming optimization

Energy terms



Contact terms



$$E_1 = E_2 + E^{V\,(a)}_{r,c,c-1} - E^{V\,(d)}_{r,c-1,c-1}$$

$$E_2 = E^{U}_{r,c} + E^{H\,(c)}_{r,c-1,c+1} - E^{H\,(g)}_{r,c-1,c} - E^{H\,(h)}_{r,c,c+1} - E^{V\,(e)}_{r,c,c}$$

$$E_3 = E_2 + E^{V\,(b)}_{r,c,c+1} - E^{V\,(f)}_{r,c+1,c+1}$$

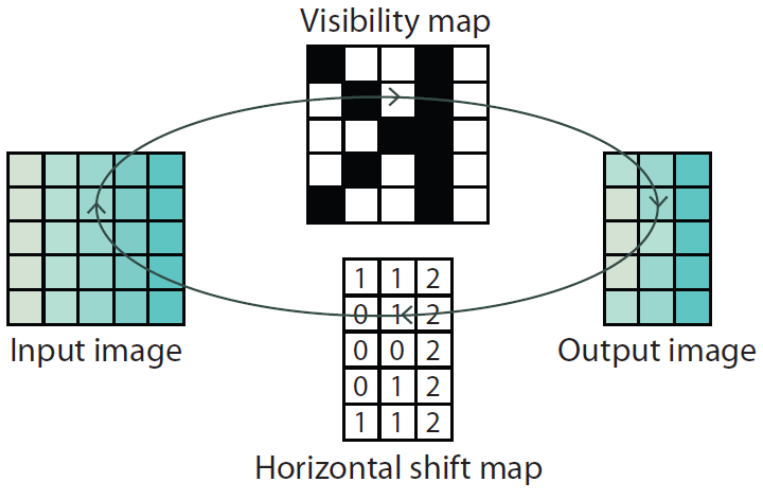
Contact terms turned on

Contact terms turned off

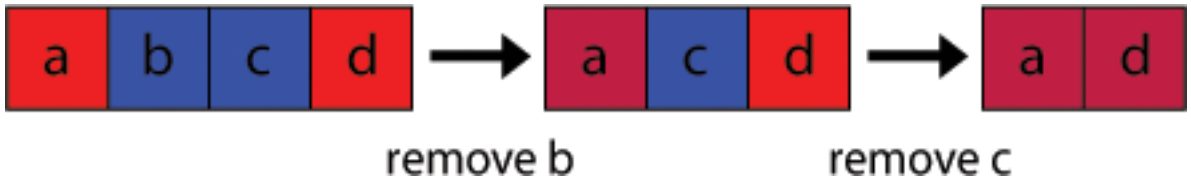
Relation to shift map

- Given a shift map M with an ordering constraint, the visibility map X is given by:

$$X_{r,c} = \begin{cases} 1, \exists(u,v) \text{ such that } ((u,v) + M_{u,v}) = (r,c) \\ 0, \text{otherwise} \end{cases}$$



Blending



- Visibility map assigns a weighting in a linear blending:

$$I^x_{r,c-1} = wI^x_{r,c} + (1-w)\, I^x_{r,c-1}$$
$$I^x_{r,c+1} = wI^x_{r,c} + (1-w)\, I^x_{r,c+1} \qquad \forall x \in \{\mathrm{R}, \mathrm{G}, \mathrm{B}\}$$

We use $w = 0.25$