Improved Seam Carving for Video Retargeting (2008)

&

Multi-operator Media Retargeting

Michael Rubinstein, Ariel Shamir, Shai Avidan. SIGGRAPH 2009.

Presenter: Nancy Wang
Discussant: Sally Ahn

Overview

- Purpose: Have method for supporting content aware resizing for videos as well as images.
- Key Ideas:
  - Improved Seam Carving
    - remove 2D seam manifolds from 3D-space-time volumes
  - Multi-operator Media Retargeting
    - Combine several operators together to resize image

Improved Seam Carving for Video Retargeting

- Improved Seam Carving -

Other purposed methods

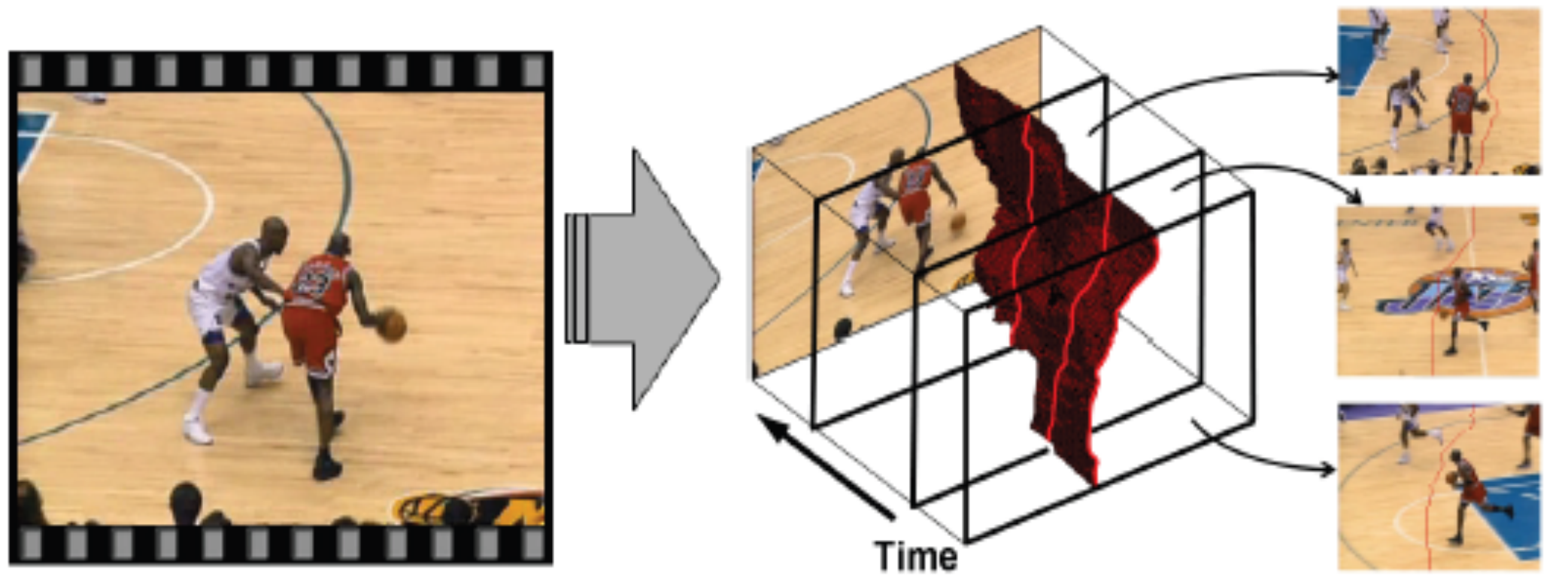
- Independently apply seam carving to each frame and resize it.
  - Lack of coherency

- Improved Seam Carving -

Other purposed method

- Static seams-find max energy projection to use to carve same set of seams for all the frames
  - doesn't work well for complex scenes
    - moving camera, multiple motions

Solution: treat video as 3D cube and extend seam carving to 2D manifolds in 3D volume



Seam Carving using Graph Cuts:

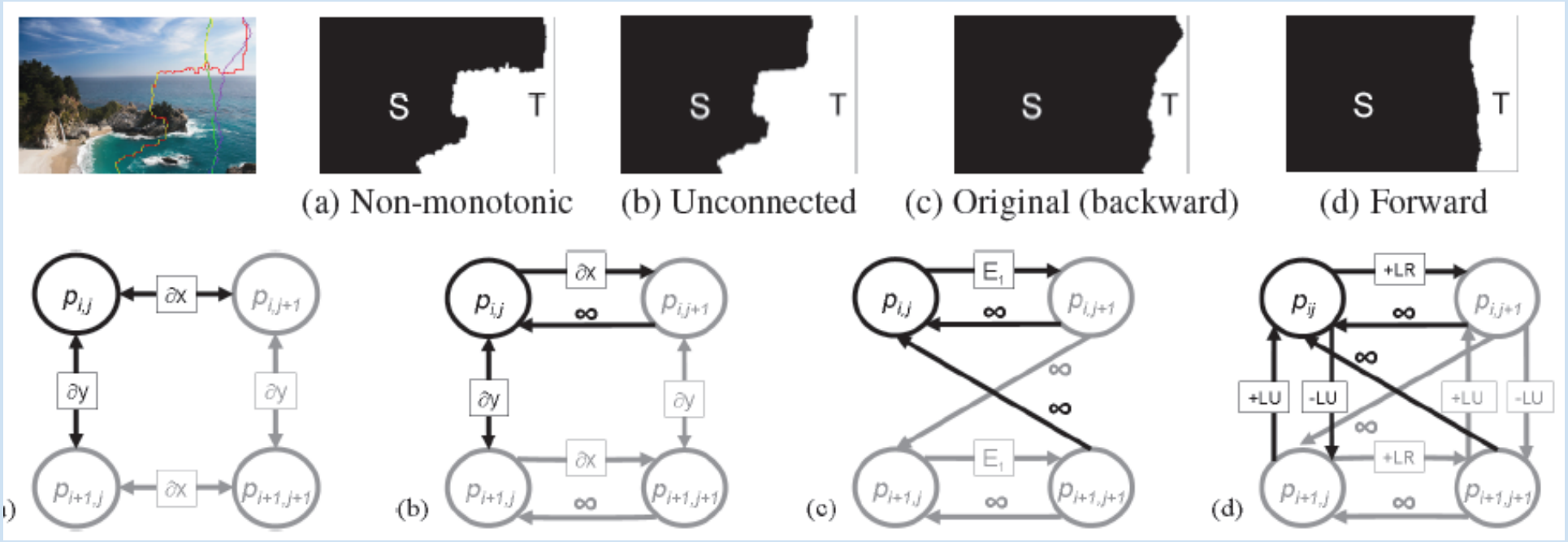
- For images

neighbors

$$Nbr(p_{i,j}) = \{p_{i-1,j}, p_{i+1,j}, p_{i,j-1}, p_{i,j+1}\}.$$

horizontal direction: $\partial x(i,j) = |I(i,j+1) - I(i,j)|$

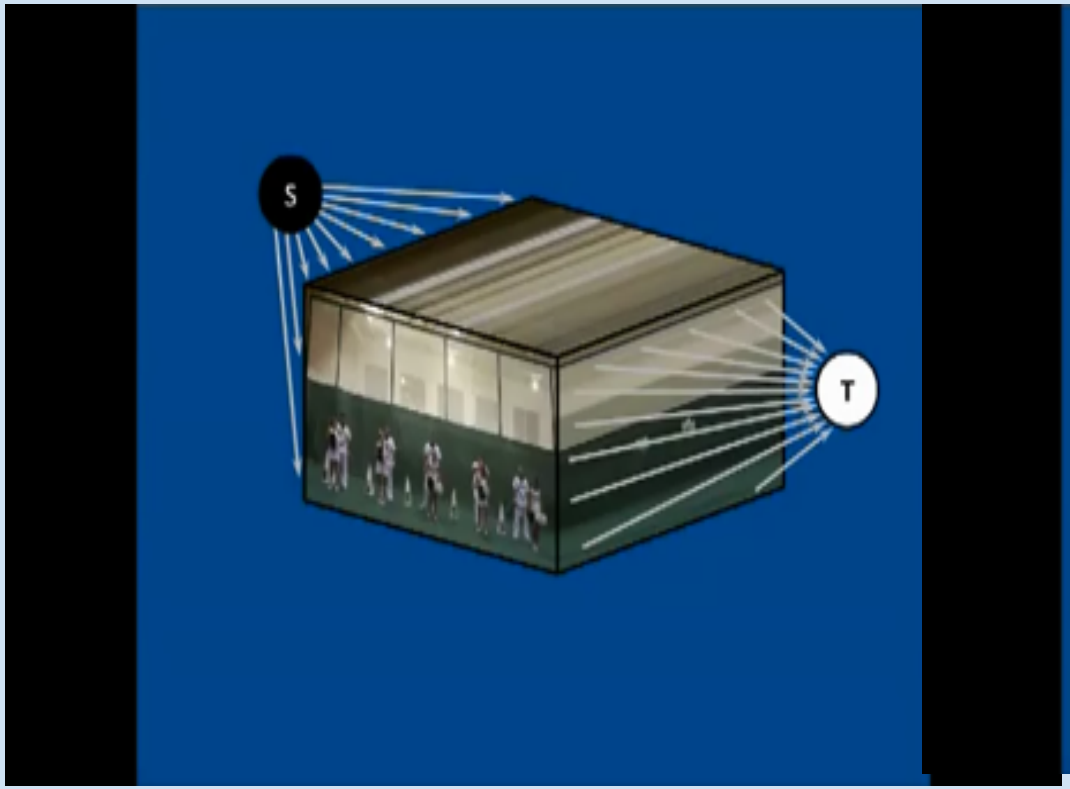
vertical: $\partial y(i,j) = |I(i+1,j) - I(i,j)|$

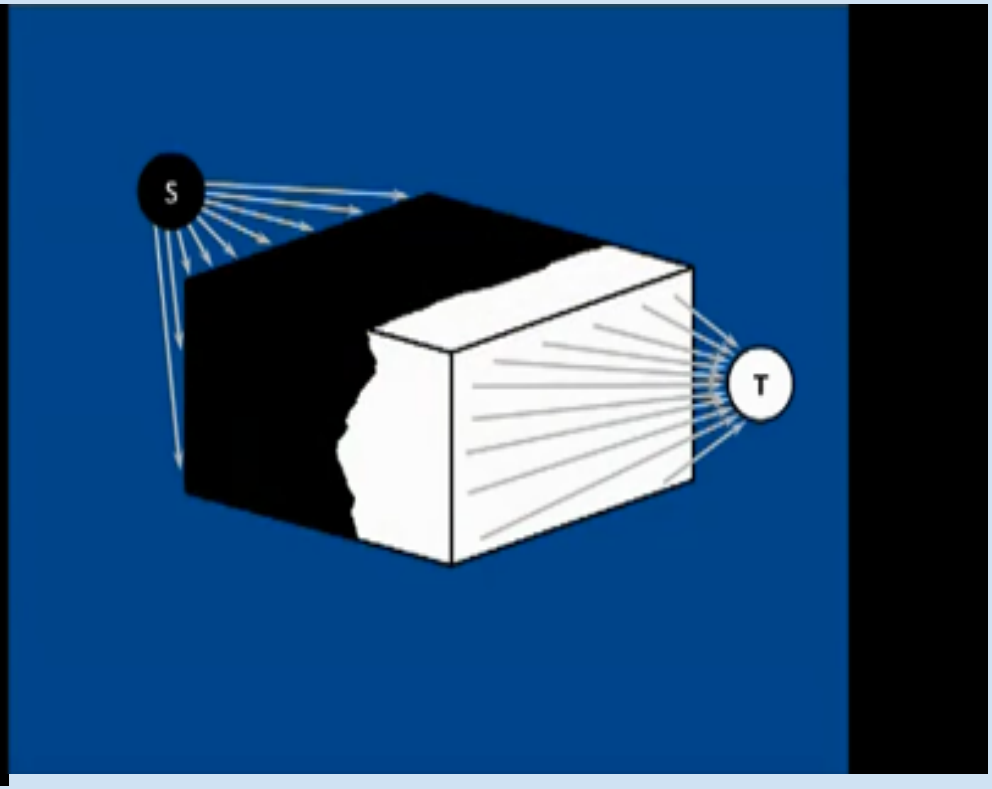


- Improved Seam Carving -

Seam Carving on Videos using Graph Cuts:

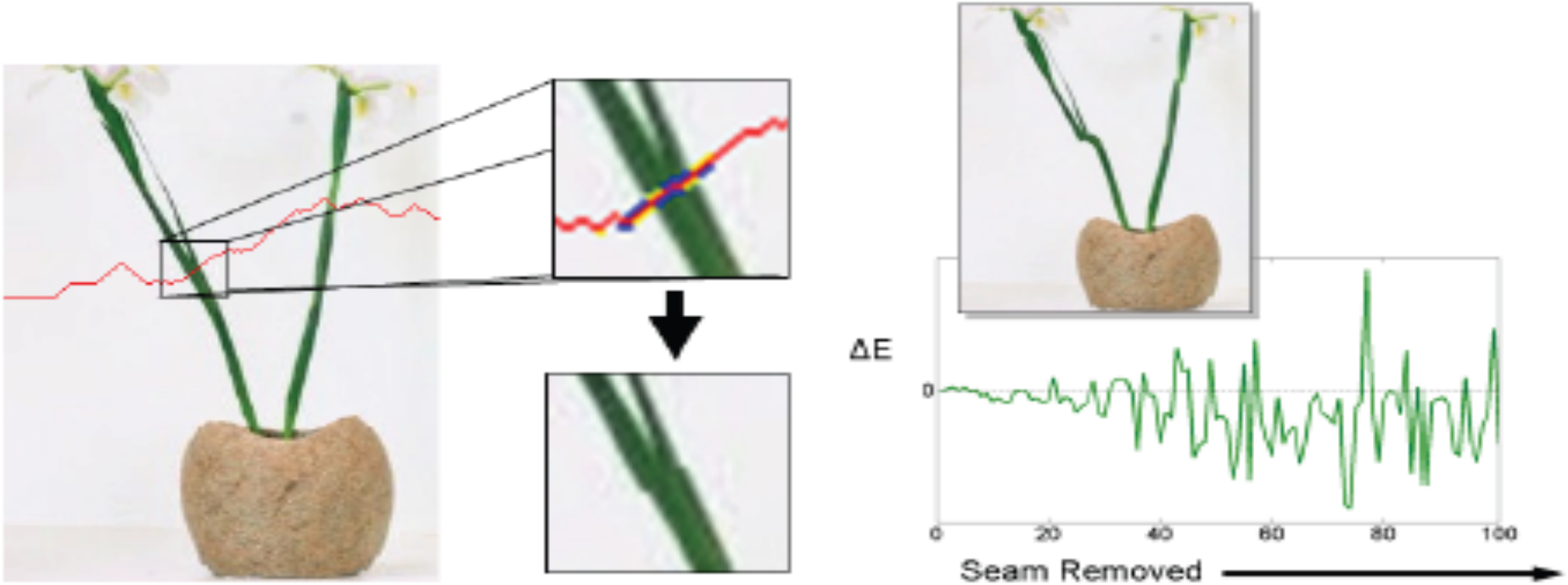
- Define grid-like graph
- consider the X x T planes in cube and use same graph construction as in X x Y for images include backward diagonal infinity arcs
  - S (source) and T (sink) nodes created and connected both sides
  - S/T cut - partitioning of the nodes in the graph





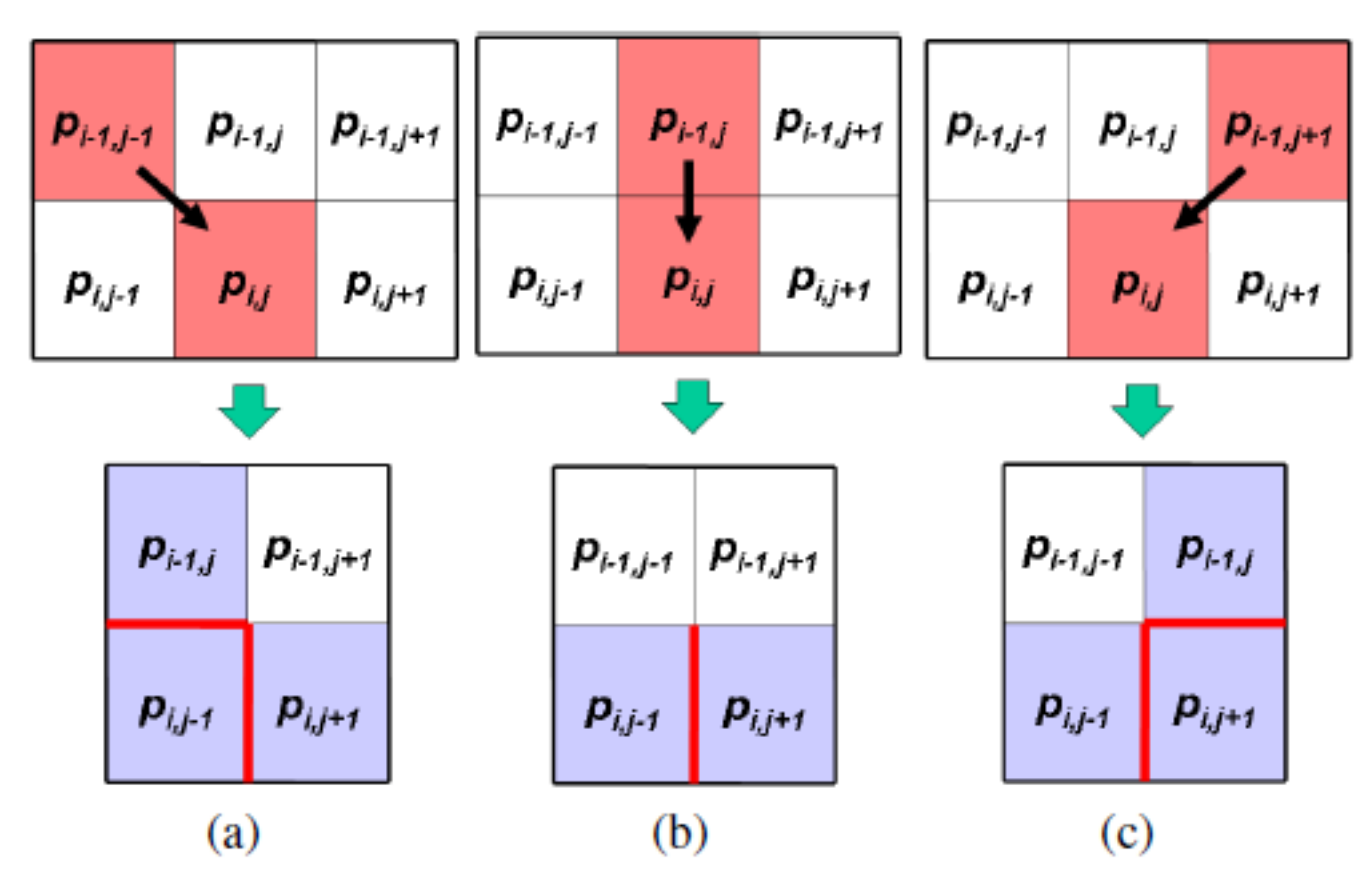
- Improved Seam Carving -

Artifacts still show



(a) (b)

Forward energy in Graph Cuts



(a) $C_L(i,j) = |I(i,j+1) - I(i,j-1)| + |I(i-1,j) - I(i,j-1)|$
(b) $C_U(i,j) = |I(i,j+1) - I(i,j-1)|$
(c) $C_R(i,j) = |I(i,j+1) - I(i,j-1)| + |I(i-1,j) - I(i,j+1)|$

Forward energy in Graph Cuts

- examine slices in 3D cube depending on seam direction
- define cost of every pixel removal as the new temporal pixel-edges created
- create arcs between nodes in the graph between time-steps with appropriate costs exactly as in X x Y domain

Results

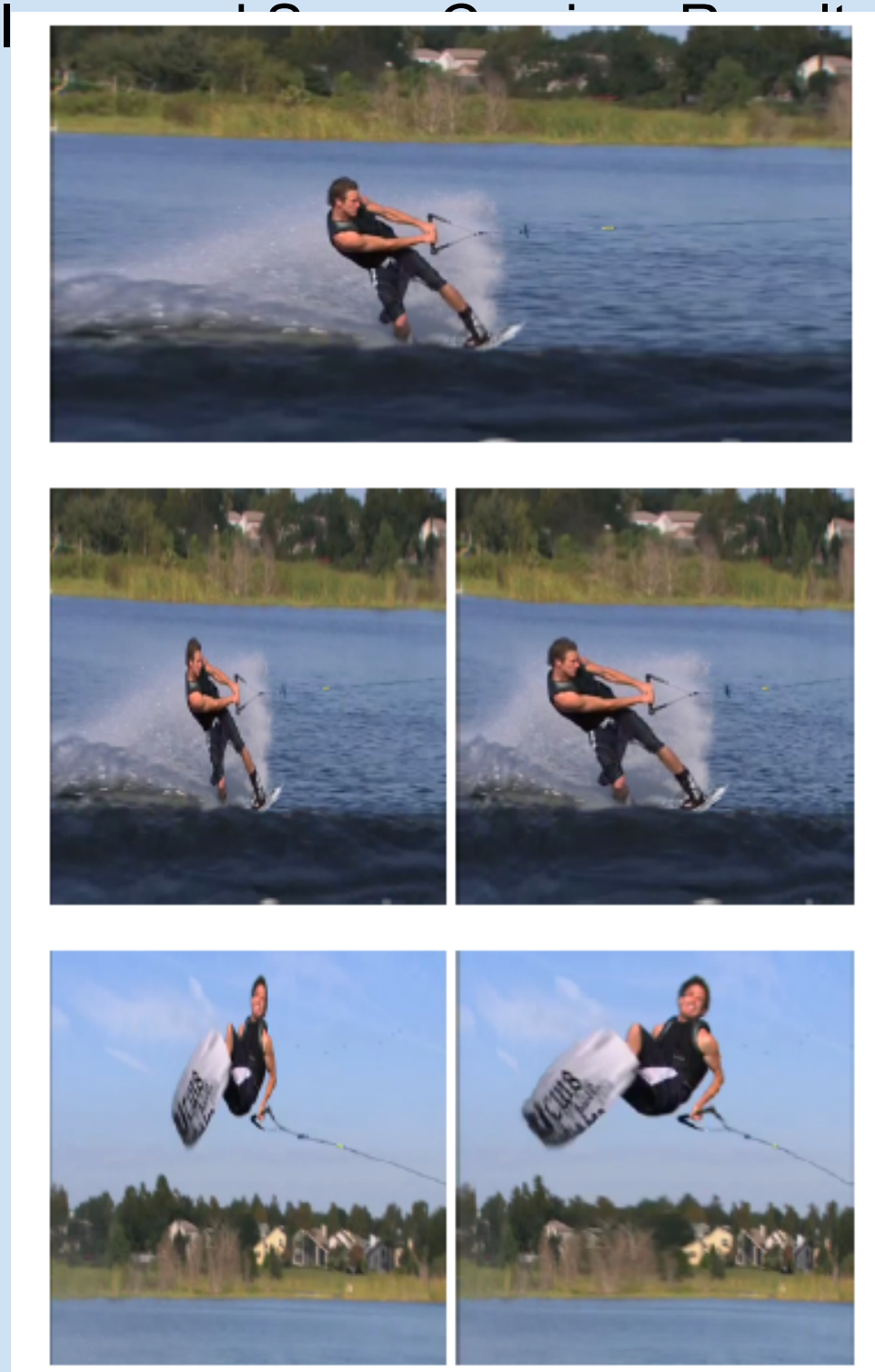


Figure 10

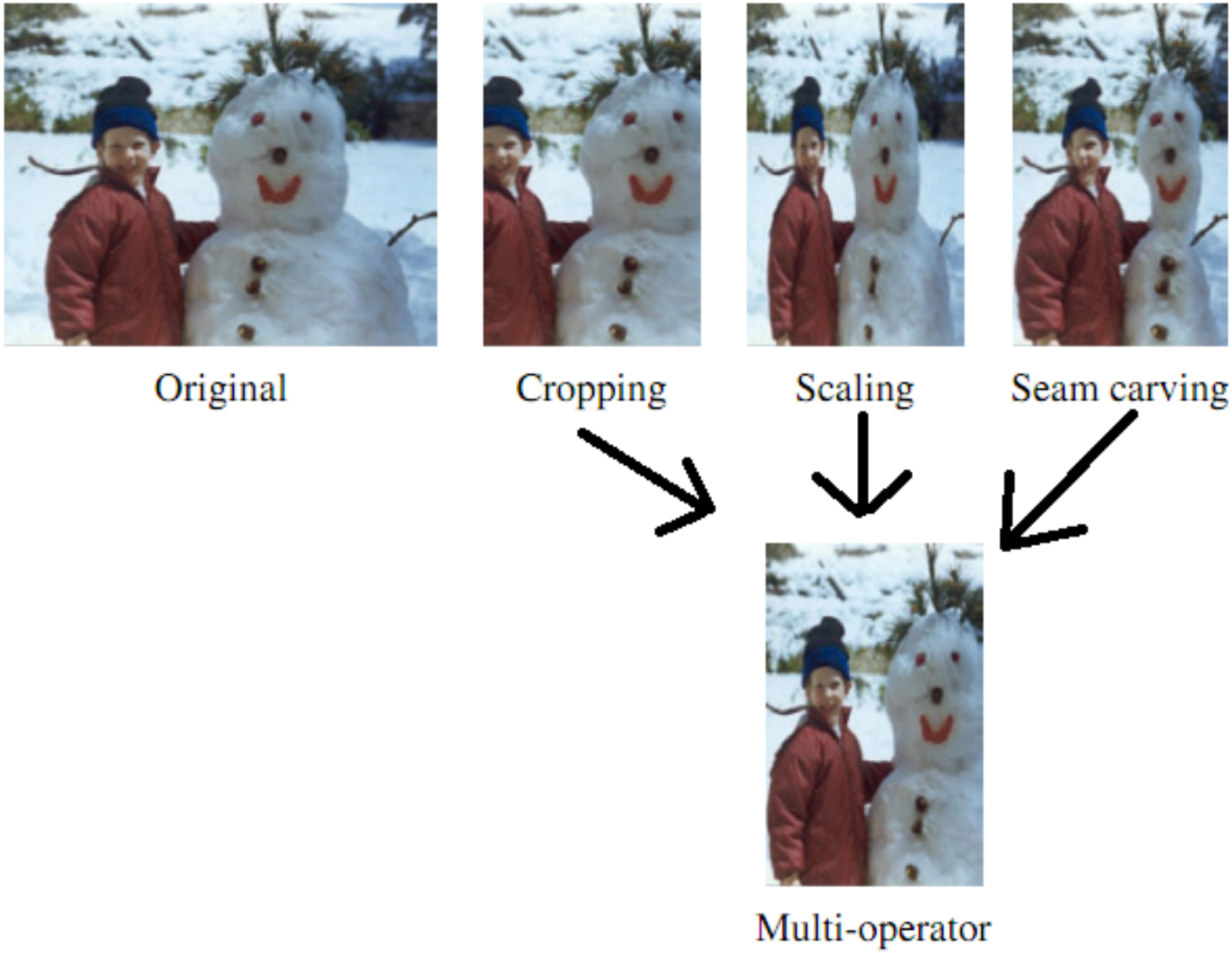
Figure 10

- Improved Seam Carving: Results -

Figure 14: Cases where forward energy fails.

Multi-operator Media Retargeting

- Multi-operator Media Retargeting-



- Multi-Operator Sequences
  - First define resizing space as conceptual multi-dimensional space combining several retargeting operators.
    - Given image I of size (w,h), define the resizing space as the space spanned by any subset of n types of retargeting operators, each one in width and height

- Multi-operator Media Retargeting-

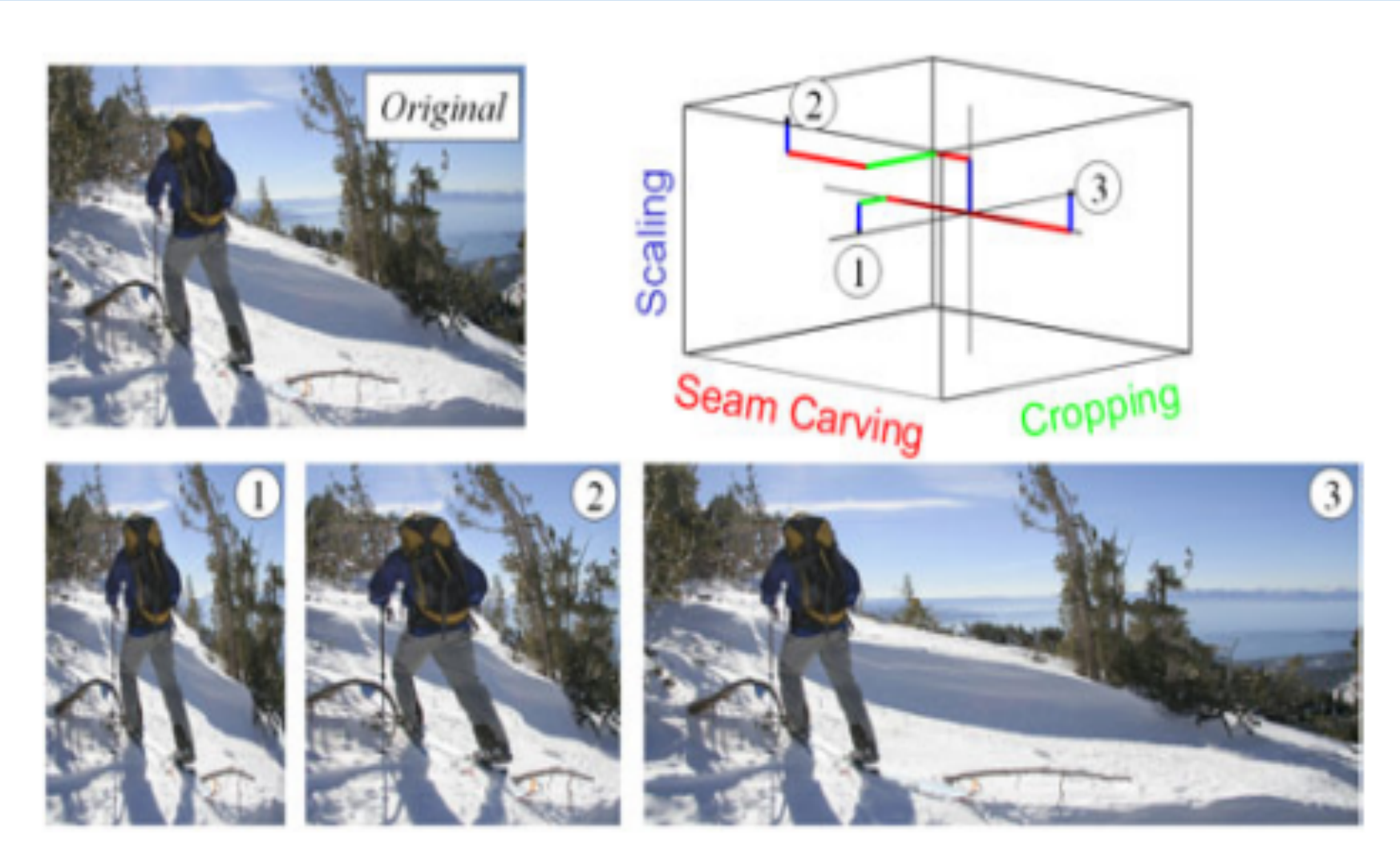


Figure 4: An example of a resizing space of an image using only changes in width by scaling, cropping and seam carving. Different retargeting results can be achieved using different multi-operator sequences represented by paths in the space

- Multi-operator Media Retargeting-

- There are many ways to retarget an image to a particular size.
- Note: resizing operators are **not** commutative
- Propose a novel similarity measure between images that is termed Bi-Directional Warping

- Multi-operator Media Retargeting-

- Bi-Directional Warping uses variant of DTW
  - Dynamic Time Warp
    - Finds optimal matching between two 1-D sequences s and t
    - Constraints
      - the first and last elements of t must be matched to that of s
      - all elements of t and s must be used in the warp path
      - warp must be monotonic
  - Bi-directional
    - allows algorithm to insert gaps in the warp
    - want single match that minimizes warping cost

- Multi-operator Media Retargeting-

Algorithm 1 Asymmetric-DTW($s[1..|s|], t[1..|t|]$)

```
allocate M[|s| + 1][|t| + 1]
M[0, 0] := 0
for i = 1 to |s| do
    M[i, 0] := ∞
for j := 1 to |t| do
    M[0, j] := 0
for i := 1 to |s| do
    for j := 1 to |t| do
        M[i, j] := min(M[i − 1, j − 1] + d(s[i], t[j]),
                       M[i, j − 1],
                       M[i − 1, j] + d(s[i], t[j]))
return M[|s|, |t|]
```

$$\mathrm{BDW}(S,T) = \frac{1}{N_S}\sum_{i=1}^{h} \text{A-DTW}(S_i, T_i) + \frac{1}{N_T}\sum_{i=1}^{h} \text{A-DTW}(T_i, S_i) \qquad (1)$$

- Multi-operator Media Retargeting-

(a) Source

(b) Seam Carving

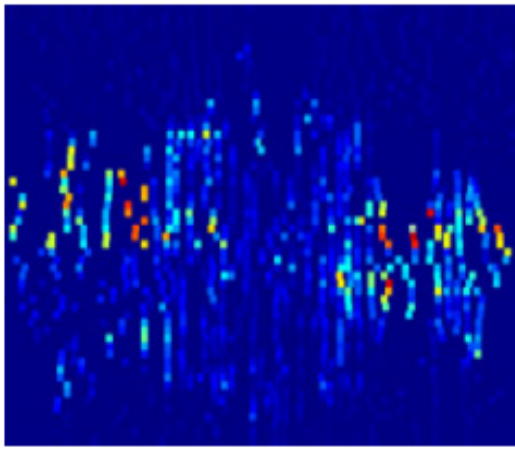
(c) Actual Seams

(d) Optimal matching, pixels

(e) 8×8 matching patches

(f) 16×16 matching patches

(g) 32×32 matching patches



(h) Distance map of 8×8 patches

Figure 6: Finding the optimal match using Asymmetric-DTW of image (b) to image (a) using different patch sizes (d)-(g). The element-wise distance function $d()$ was taken as the L_1-norm of RGB differences. Note that black pixels represent gaps in the matching. The distance itself is defined as the average or maximum of the cost of matching each element (h).

- Multi-operator Media Retargeting-

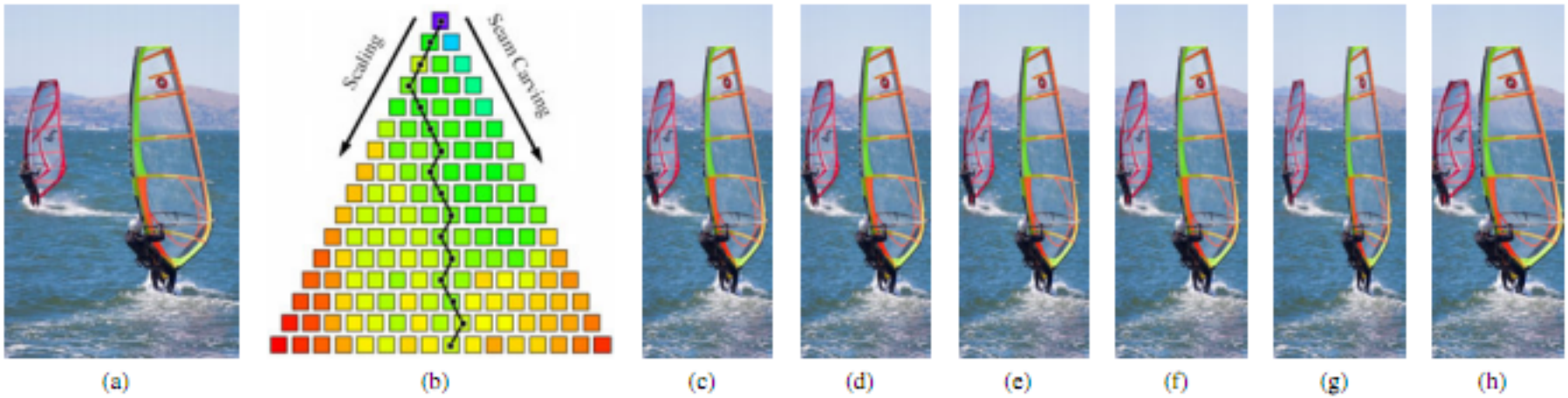
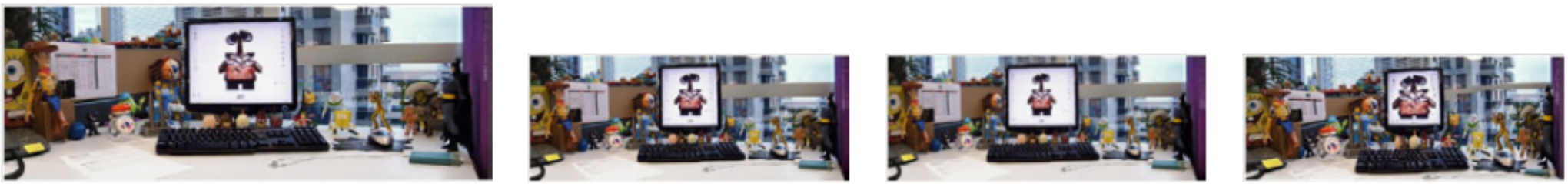


Figure 9: An illustration of the dynamic programming table used to optimize the search for the best *mixed* path using two operators only - seam carving (SC) and scaling (SL). The colors in table (b) indicate the BDW distance of the best image in each step. The original image is shown in (a) and the retargeted result is shown in (c) - this is the best result using a *mixed* path (i.e. the algorithm automatically determines the order of operators and how much each should contribute). The optimal operator sequence found is $\langle -30SL, -30SC, -10SL, -20SC, -10SL, -10SC, -10SL, -20SC, -10SL \rangle$. For comparison, we show the results of using two *regular* paths (d) $\langle -70SL, -80SC \rangle$ and (e) $\langle -80SC, -70SL \rangle$, and the optimal *regular* path (f) $\langle -90SC, -60SL \rangle$. (g) uses scaling and (h) seam carving.

- compute optimal path in n dimensional resizing space
- store in each step the cost of operator that best preserves image
- extend to videos using key frames

Results

- Multi-operator Media Retargeting: Results-



Input | Multi-op | Scaling | Seam Carving

Figure 14: Result of 2D retargeting. In this case we find the optimal *mixed* path using the bidirectional similarity measure and a combination of four retargeting operators (horizontal and vertical seam carving and scaling). The multi-operator result finds the best result by mainly applying seam carving in the horizontal dimension (compare the size of the monitor in the different methods) and scaling in the vertical dimension (look at the bottom of the desk and the face of Woody on the left). For comparison, we show the result of applying a uniform 2D scaling, or seam carving (running horizontal seam carving first, followed by vertical seam carving).

Seam Carving Multi-op Scaling

Figure 15: A comparison between seam carving (left), Multi-operator (center) and scaling (right). The multi-operator algorithm uses the BDW image similarity measure and finds the best *mixed* path using two image retargeting operators (seam carving and scaling).

The End

Extras

Energy Equations

For Static Seams

$$
\begin{aligned}
E_{\text{spatial}}(i,j) &= \max_{t=1}^{N}\{|\frac{\partial}{\partial x}I_t(i,j)| + |\frac{\partial}{\partial y}I_t(i,j)|\} \\
E_{\text{temporal}}(i,j) &= \max_{t=1}^{N}\{|\frac{\partial}{\partial t}I_t(i,j)|\} \\
E_{\text{global}}(i,j) &= \alpha \cdot E_{spatial} + (1-\alpha)E_{temporal}
\end{aligned}
$$

Discussion

What causes forward energy to fail in these examples? Would these fare better with the previous seam carving algorithm?

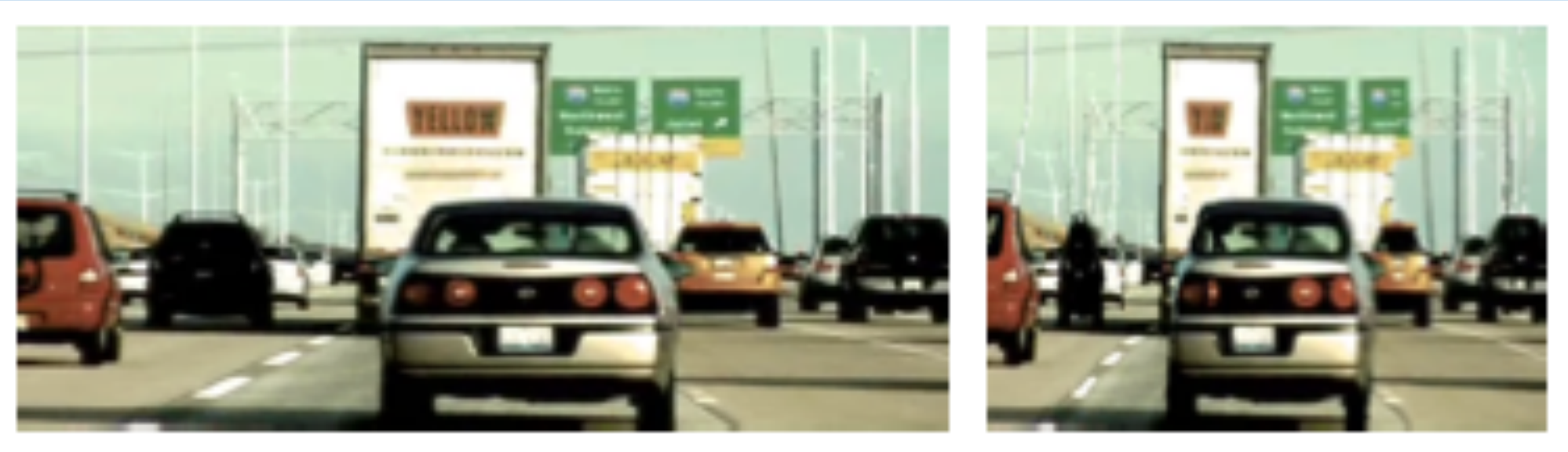
Discussion

What causes forward energy to fail in these examples? Would these fare better with the previous seam carving algorithm?

textured background more important than bicyclist/matchbox

Discussion

Another forward energy limitation example:



"due to the nature in which the camera and objects are moving"

What kind of camera/object movement would cause these artifacts?

Discussion

Source:

Discussion

Source:

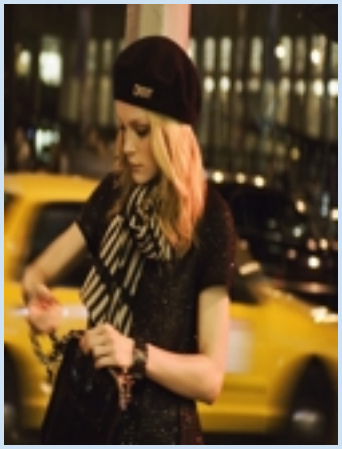
Multiop

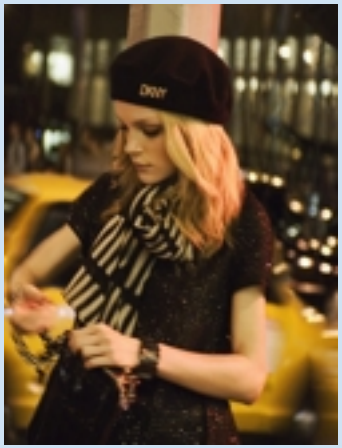
Cropping

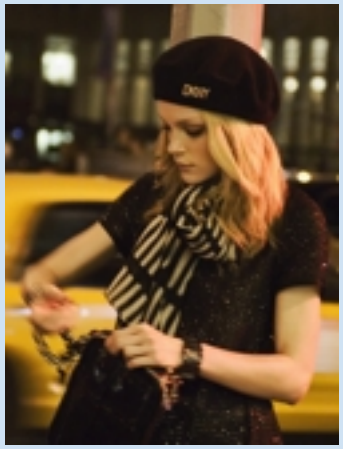
Improved Seam Carving

Discussion

Source:



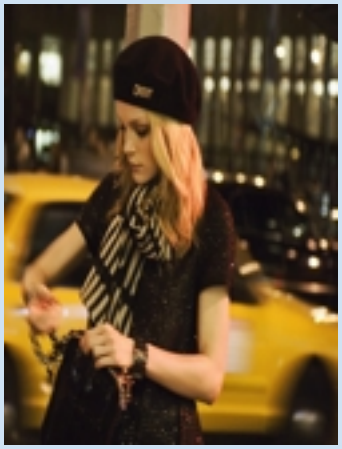


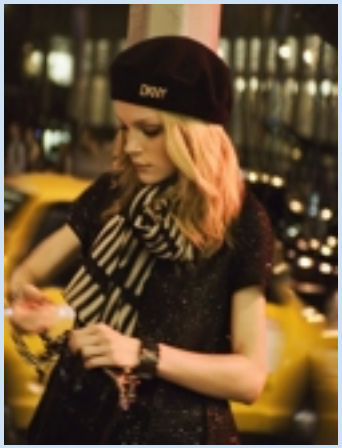


Discussion

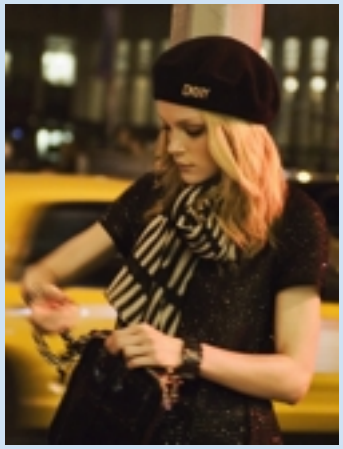
Source:

Multiop





Improved Seam Carving



Cropping

Discussion

The authors of "Multi-operator Media Retargeting" present user study data to back their claim that multiple operators for resizing is better than using a single one.

How valid are the conclusions drawn from this study?