Supplementary Material
Scene Carving:
Scene Consistent Image Retargeting

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 Ltd, Cambridge, UK
{mansfield, pgehler, vangool}@vision.ee.ethz.ch, carrot@microsoft.com

1 Introduction

In this supplementary material we first (Section 2) describe the details of a variant of scene carving (Sc. Carve-D.) not fully described in the paper but mentioned in Sect. 5.2 and Fig. 8(f) and Fig. 8(g). This algorithm aims to minimize the *visible* distortion in the next image, in contrast to the scene carving algorithm (Sc. Carve) we present in the paper in which the distortion in the background image, whether visible or not, is minimized. Then, in Section 3 we show results on more images for all algorithms, including Sc. Carve-D.

2 Scene Carving to Minimize Visible Distortion (Sc. Carve-D.)

In Section 5.2 of the paper, we described two choices of seam energies for scene carving. The first (1) calculated the energy of distortion in the "flattened" image, the image produced at the end of that iteration. This is an extension of the energy advocated in seam carving and used in S.C.+Obj. Occ. The second (2) calculated the energy of distortion in the background image alone. In the paper we present Sc. Carve, based on (2) and discuss why we choose it over (1) (Sc. Carve-D.) with reference to an empirical evidence. While not central to the paper, we now describe this algorithm in order to show how we produce the results.

Sc. Carve-D. follows the same strategy as scene carving, as outlined in Section 5.1. We consider all object positionings, and for each find an optimal seam in the background image. The only difference is in how we define "optimal", i.e. how we define the energy.

The energy contains three different types of terms. First we set the energy terms to account for the local energy i.e. for the distortion created directly by the removal of the pixels (Section 2.1). The standard forward energy of seam carving (1) is an example of this. Second, we add terms that account for energy in remote parts of the image caused by the background moving relative to the object positioning (Section 2.2). The hard constraint that all holes be covered by objects during scene carving is an example of this (6). Finally we can optimism for the seam in the background and add an energy term related to the object positioning (Section 2.3).

2.1 Local Energy

We start with the energy as in (1) in the paper from seam carving. This is the energy of created distortion in the background by the seam. Terms in certain locations are then changed from this when the distortion this energy encodes the cost for does not actually occur. For example, no visible distortion can be caused directly by pixels being removed at the edges of the image, as the pixels can consistently be considered to have become occluded. We therefore set energy term arcs pointing to the edges of the image to a small constant $u_{\mathrm{s}} = 6$,

$$E^{\mathrm{LR}}_{r,c} = E^{\mathrm{LU}}_{r,c} = E^{\mathrm{UR}}_{r,c} = u_{\mathrm{s}} \forall \, (r,c) \in \{ \, (r,c) : (r,c-1) \notin \mathcal{P} \vee (r,c+1) \notin \mathcal{P} \} \ . \tag{7}$$

A similar term applies to background pixels around the edges of objects. Here care is needed however. After the seam is removed, the pixels move to new positions. Distortion is created when pixels are put into contact that were not in contact before. If the locations of these pixels when in contact are both background pixels *before* the seam is removed, they will still be background pixels *after* the seam is removed, as the object positions are already fixed. The original energy is paid only if both of these pixel locations are background, as then the distortion is visible. Otherwise we again pay the small constant u_{s}. The exact condition for paying u_{s} around object boundaries and inside objects is:

$$E^{\mathrm{LR}}_{r,c} = u_{\mathrm{s}} \quad \forall (r,c) \in \{(r,c) : O_{r,c-1} > 0 \wedge O_{r,c} > 0\} \tag{8}$$

$$E^{\mathrm{LU}}_{r,c} = u_{\mathrm{s}} \quad \forall (r,c) \in \{(r,c) : O_{r,c-1} > 0 \wedge O_{r-1,c-1} > 0\} \tag{9}$$

$$E^{\mathrm{UR}}_{r,c} = u_{\mathrm{s}} \quad \forall (r,c) \in \{(r,c) : O_{r,c} > 0 \wedge O_{r-1,c} > 0\} \ . \tag{10}$$

The final local energy term is an additional parameter to control the encouragement of hole pixel removal. We add a negative constant $u_{\mathrm{h}} = -6$ to all energy terms for hole pixels to achieve this.

2.2 Remote Energy

Cost of Re-appearance. Terms are now added to account for re-appearance of background. Note that occlusion of background creates no distortion and so has no associated cost. In order to do this, we must first define the cost of re-appearing background. This must be given by the visible distortion this re-appearance causes. This is the *latent energy of distortion.* Whenever a seam passes through the background, distortion is created even if it is not visible. The energy of this distortion is calculated at each step and stored. Note that it has two components, horizontal and vertical, and is given by

$$E^{\text{lh}}_{r,c-1} = |I_{r,c-1} - I_{r,c+1}| \quad \forall (r,c) \in \mathcal{S} \tag{11}$$

$$E^{\text{lv}}_{r+1,c} = |I_{r+1,c} - I_{r,c+1}| \quad \forall (r,c) \in \mathcal{S} \ , \tag{12}$$

where $\mathcal{S}$ is the set of all pixels in the seam.

Objects That Moved. Now we can introduce this cost. First we consider objects that moved one pixel left from their position in the previous image. These objects have thus in general already occluded some background pixels on the left hand side and revealed some background pixels on the right hand side. If the seam passes to the right of revealed background pixels, they will remain revealed. We must pay a cost for this. Given the object positionings from the previous iteration in $O^{\text{old}}_{r,c}$, we have for each row r

$$\mathcal{C}^{\text{m}}_r = \{c : O_{r,c} = 0 \wedge O^{\text{old}}_{r,c} > 0\} \ . \tag{13}$$

For each $c_{\text{rev}} \in \mathcal{C}^{\text{m}}_r$, we must pay the cost for re-appearance. If the pixel to the right is a background pixel i.e. $O_{r,c_{\text{rev}}+1} = 0$, then the cost $E^{\text{lh}}_{r,c_{\text{rev}}}$ is added to all energies of all pixels $(r,c), c > c_{\text{rev}} + 1$. If the pixel above is a background pixel i.e. i.e. $O_{r-1,c_{\text{rev}}} = 0$, then the cost $E^{\text{lv}}_{r,c_{\text{rev}}}$ is added to $E^{\text{LR}}_{r,c_{\text{rev}}+1}$ and taken away from $E^{\text{LU}}_{r,c_{\text{rev}}+1}$ so that the cost is only paid if the background pixel above is not removed by the seam, and also added to all energies of all pixels $(r,c), c > c_{\text{rev}}+1$.

Objects That Did Not Move. A similar operation is carried out for objects that did not move from their position in the previous image. If the seam passes through or to the left of pixels next to this object which could be occluded, they will become occluded. If the seam passes to the left of pixels in this object which could be revealed, they will be revealed. In the case of re-appearance we must pay a cost. We have for each row r

$$\mathcal{C}^{\text{nm}}_r = \{c : O_{r,c} > 0 \wedge O_{r,c-1} = 0\} \ . \tag{14}$$

For each $c_{\text{rev}} \in \mathcal{C}^{\text{nm}}_r$, we must pay the cost for for re-appearance. The cost $E^{\text{lh}}_{r,c_{\text{rev}}-1}$ is added to all energies of all pixels $(r,c), c < c_{\text{rev}} - 1$. If the pixel above is a background pixel i.e. i.e. $O_{r-1,c_{\text{rev}}-1} = 0$, then the cost $E^{\text{lv}}_{r,c_{\text{rev}}}$ is added to $E^{\text{LR}}_{r,c_{\text{rev}}-1}$ and taken away from $E^{\text{UR}}_{r,c_{\text{rev}}-1}$ so that the cost is only paid if the background pixel above is not removed by the seam, and also added to all energies of all pixels $(r,c), c < c_{\text{rev}} - 1$.

Holes. In this algorithm, we are also subject to the constraint that no holes be made visible. This is enforced exactly as in (6) of scene carving.

2.3 Optimization and Object Positioning Energy

Given the above energy, the optimal seam for an object positioning can be found by dynamic programming. Finally, as in scene carving we compute the final energy by adding an energy term based on the object positioning: the negative of the unary used in S.C.+Obj. Occ. (see paper (4)) for all object pixels visible. At each iteration we take the joint object positioning and background image seam with the lowest energy.

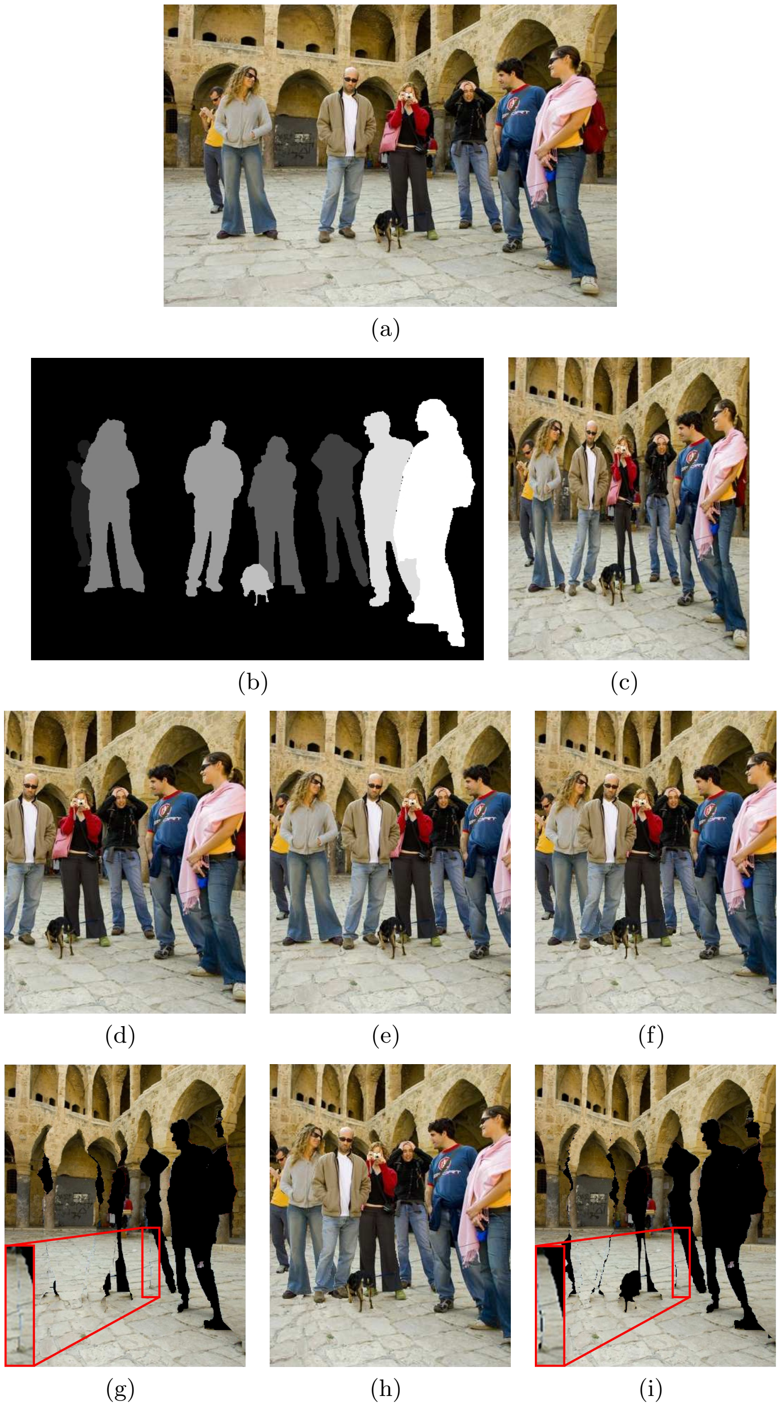
3 Additional Results

In the following figures, results including those from Sc. Carve-D. are shown, along with results for five additional images: the *Hot air balloons*, *Eiffel*, *Moscow*, *Penguins* and *Archway* images.

Note that in many cases the results produced by Sc. Carve-D. and Sc. Carve are very similar. When they are different it is only due to small artefacts, which are much less clear than the differences between Sc. Carve and the other algorithms. When small artefacts can clearly be seen in the images produced by Sc. Carve-D. but not those by Sc. Carve, due to the greediness of the former, these are shown highlighted in inlaid boxes with a red border in the background images.

Note that in general, these artefacts are due to the removal of all hole pixels from a region. This can lead to a highly distorted area of background for two reasons: (a) the background on either side of the object does not match or (b) small errors in the object segmentations label some object pixels as background, which are then left attached to the background here. In Sc. Carve-D., the hole pixels in these pixels can be removed and these areas of mismatched background brought together at no cost while they lie behind an object. In later iterations they have become visible in these examples. In Sc. Carve, the cost for bringing these areas of mismatched background together would always be paid, and so some hole pixels remain to protect this area of background from being uncovered.

Acknowledgements. We would like to thank the following users of Flickr for placing their work under the Creative Commons License, and therefore allowing us to use the images in our work: badkleinkirchheim (*Sledge*), William Hamon (*Dancers*), Alain Bachellier (*Boat*), wallyg (*London eye* and *Eiffel*), mario_mvc (*Hot air balloons*), Jean & Nathalie (*Moscow*), Wolfgang Staudt (*Archway*). The *People* and *Penguins* images were sourced from [11].



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 10. *People* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (300 seams removed). Red boxes highlight and show magnified versions of areas of background distortion visible in the Sc. Carve-D. result but not in Sc. Carve

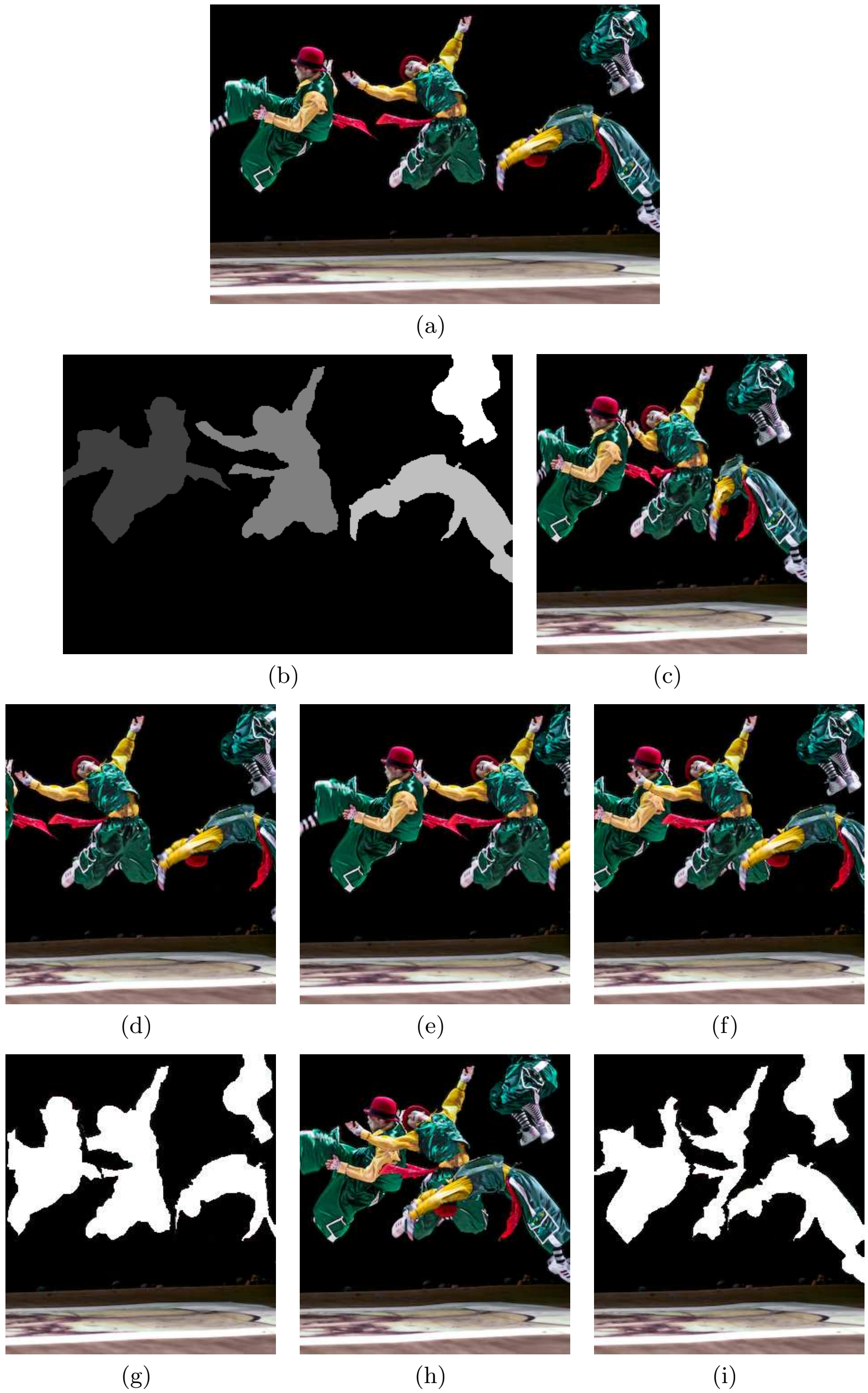


Fig. 11. *Dancers* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) with holes shown in white (200 seams removed)

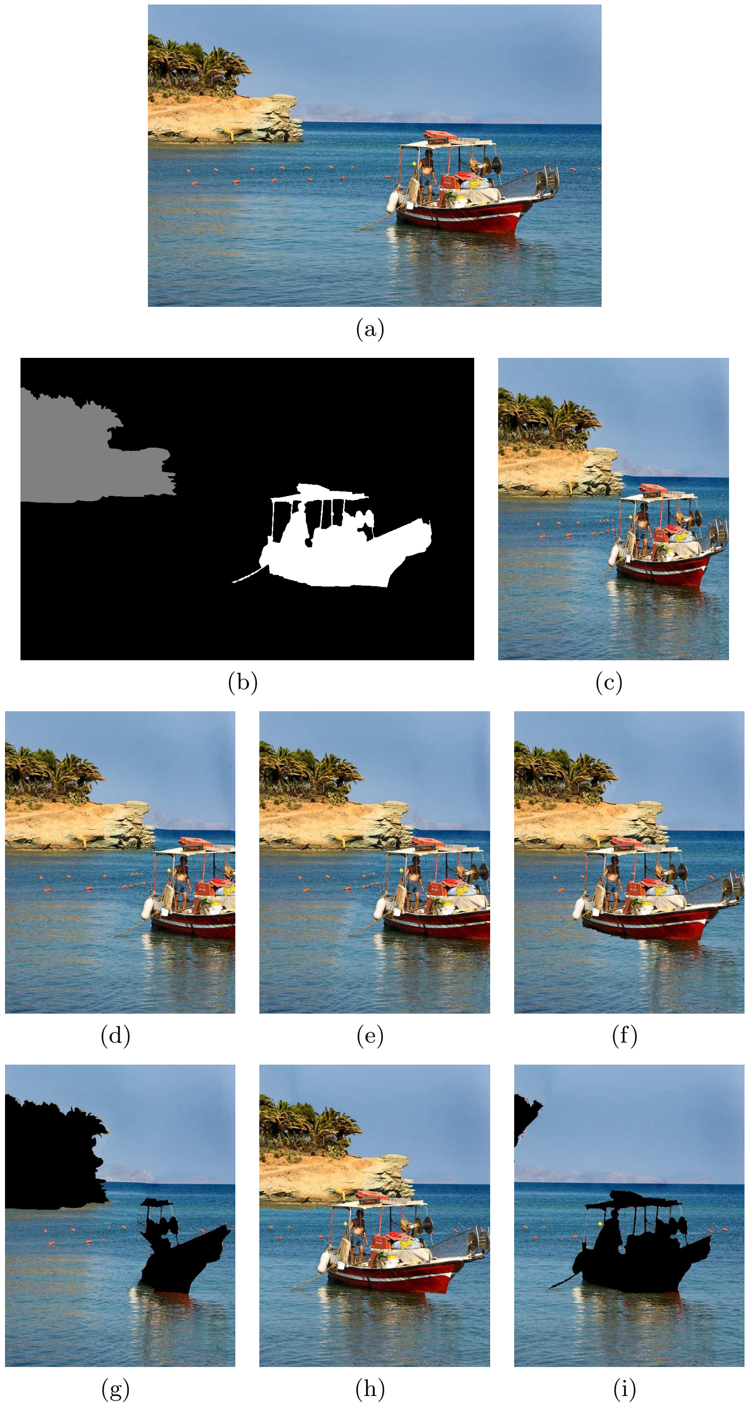


Fig. 12. *Boat* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (500 seams removed)

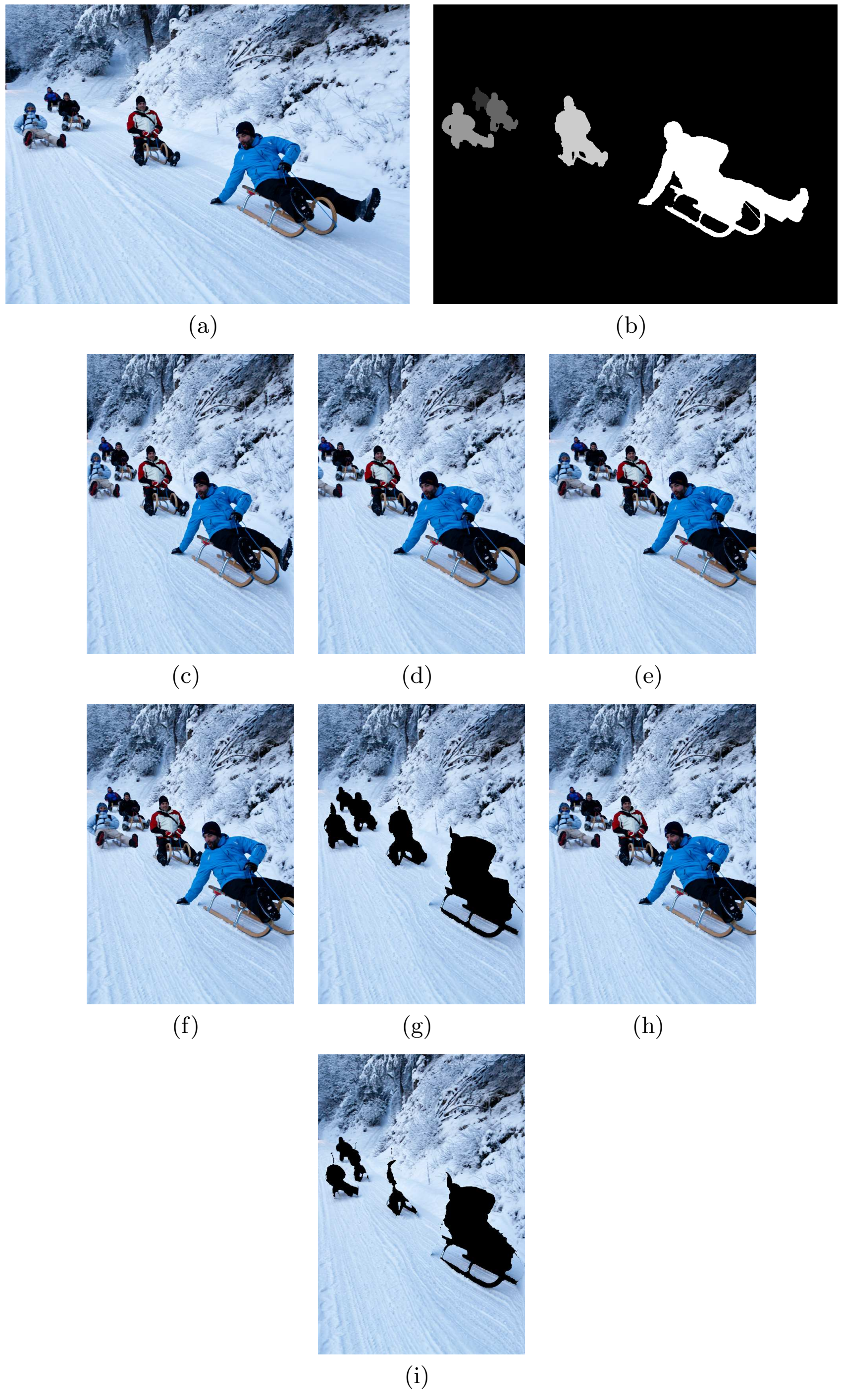


Fig. 13. *Sledge* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (500 seams removed)

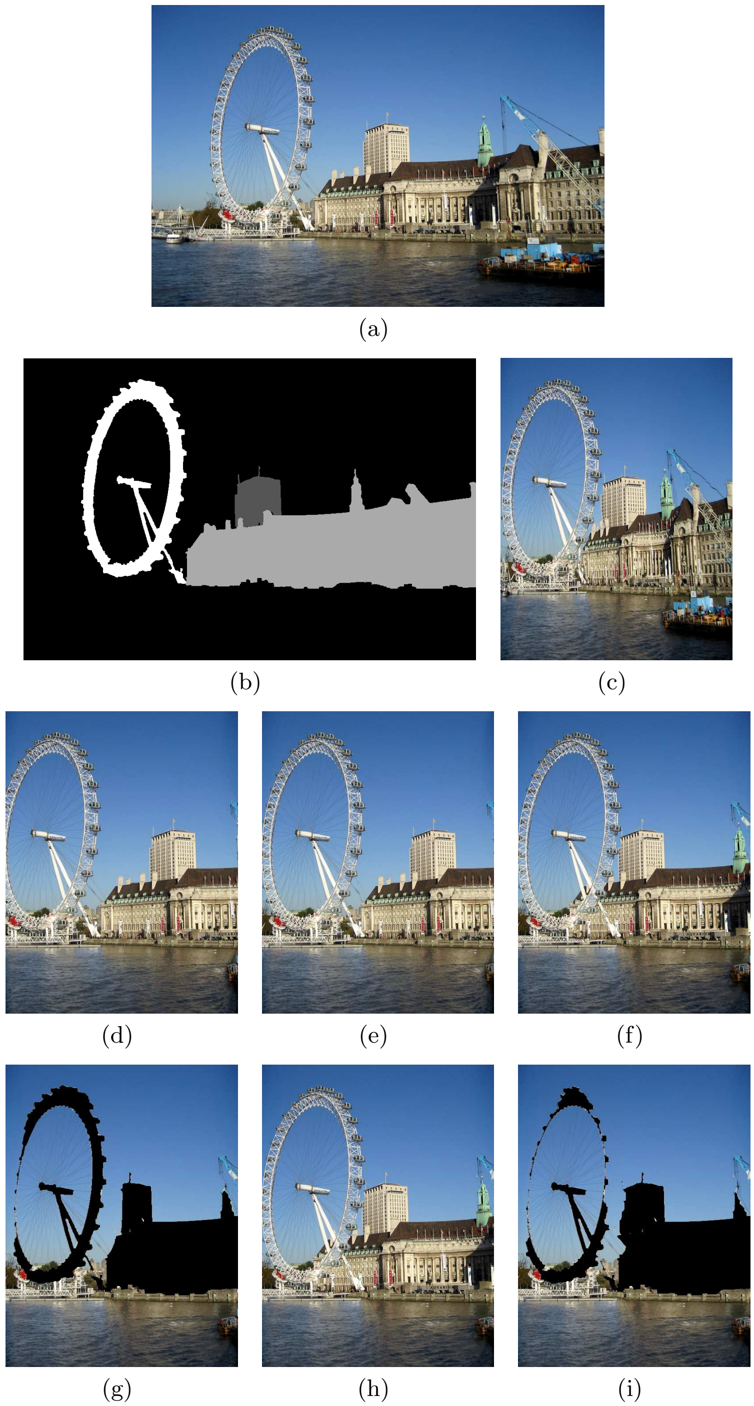
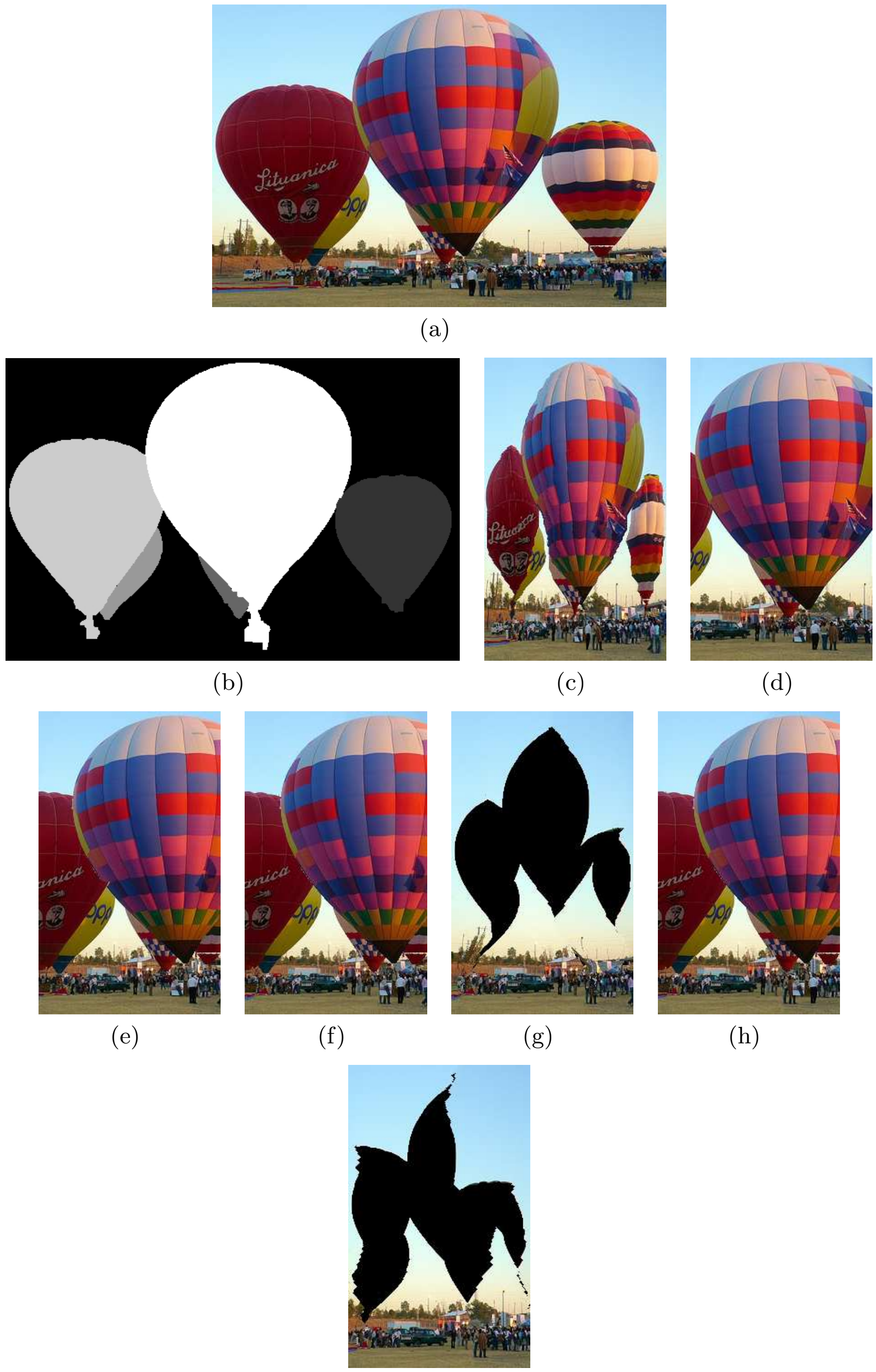


Fig. 14. *London eye* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (500 seams removed)



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 15. *Hot air balloons* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (300 seams removed)

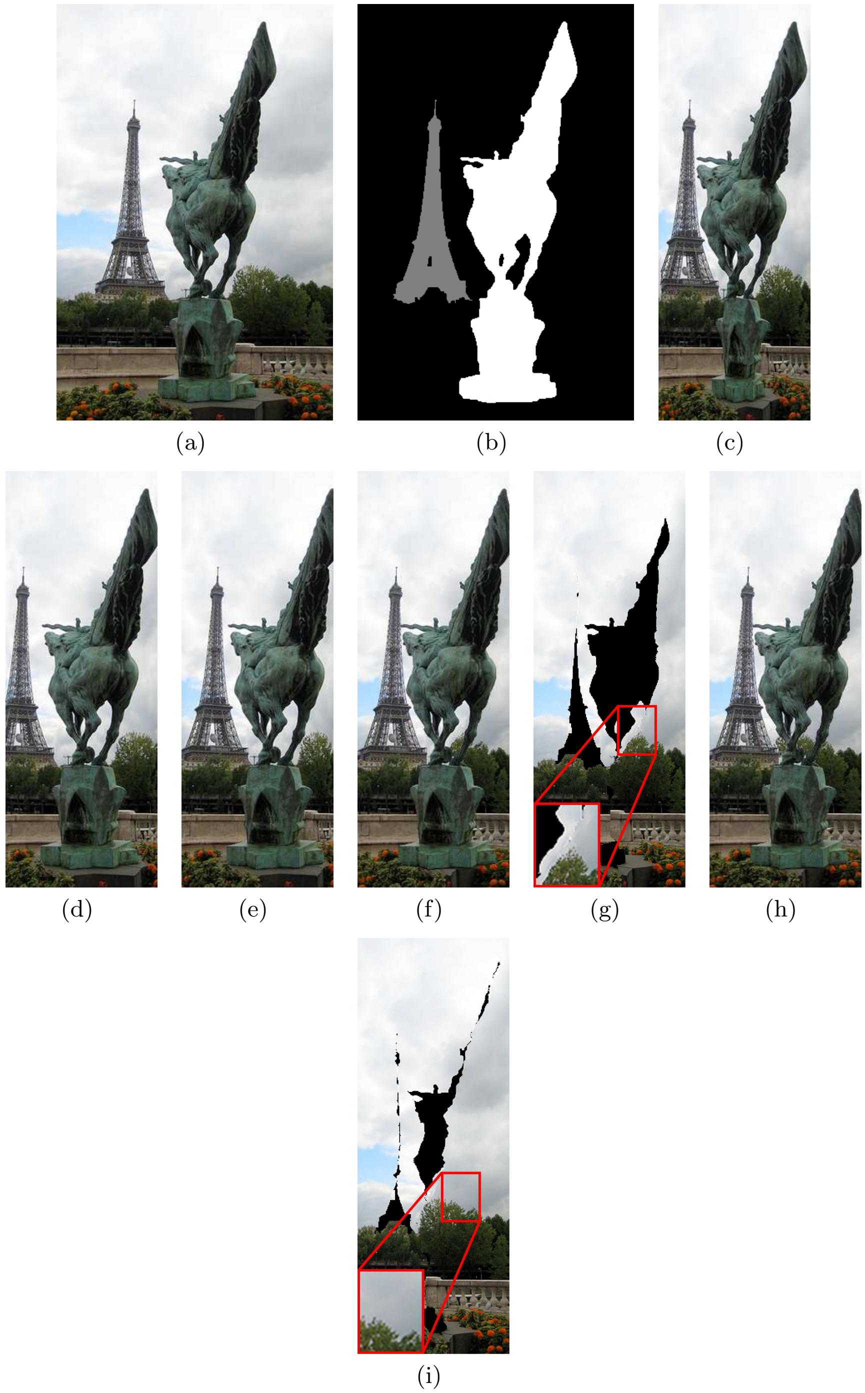
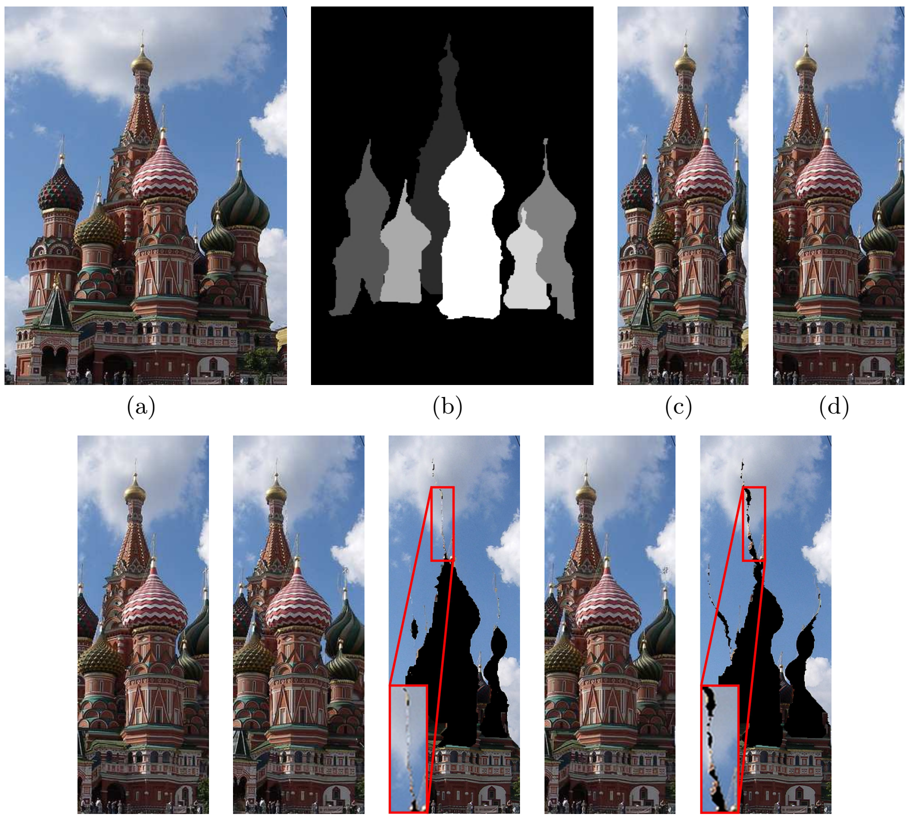


Fig. 16. *Eiffel* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (150 seams removed). Red boxes highlight and show magnified versions of areas of background distortion visible in the Sc. Carve-D. result but not in Sc. Carve



(a) (b) (c) (d)

(e) (f) (g) (h) (i)

Fig. 17. *Moscow* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (200 seams removed). Red boxes highlight and show magnified versions of areas of background distortion visible in the Sc. Carve-D. result but not in Sc. Carve

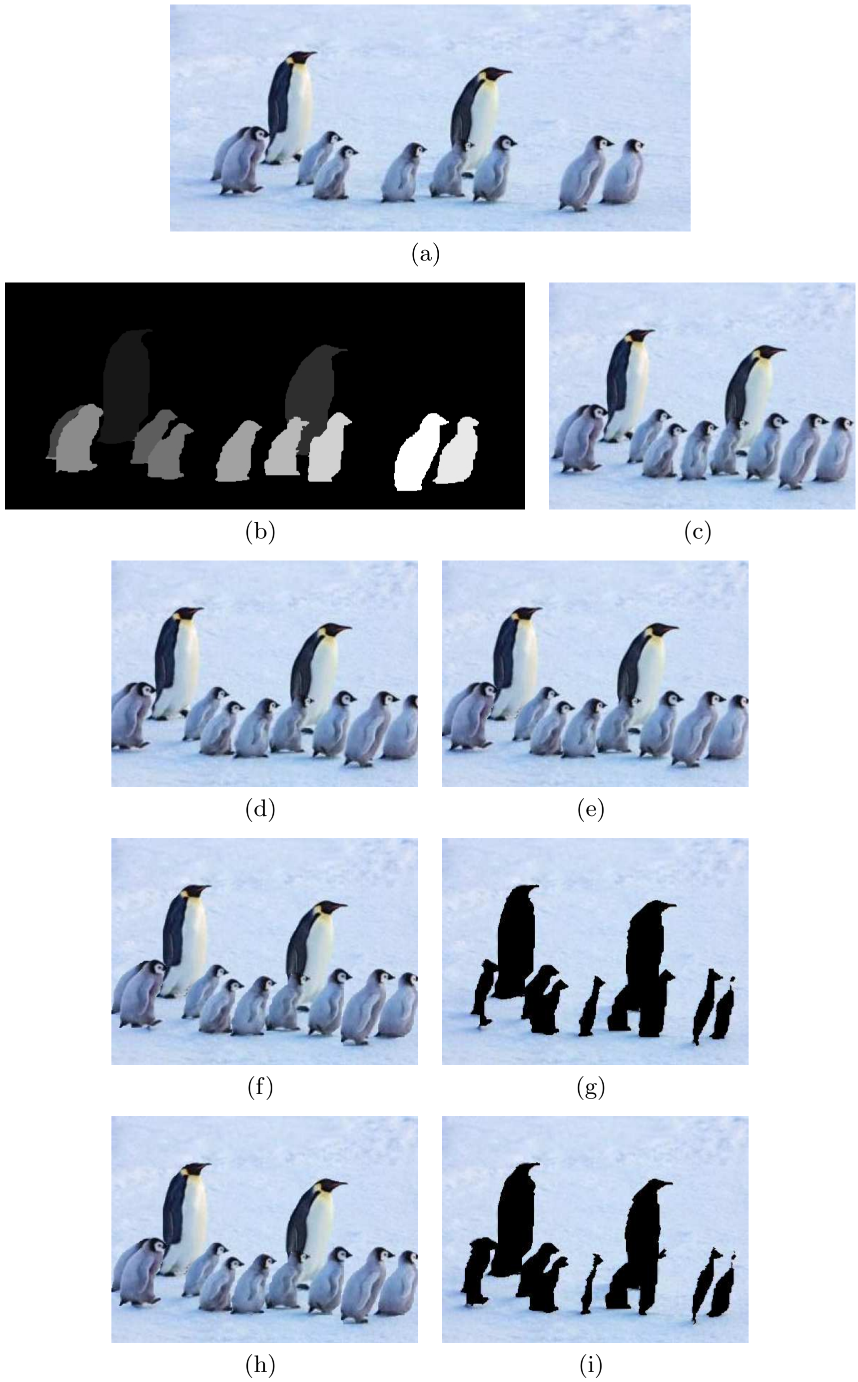


Fig. 18. *Penguins* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (300 seams removed)

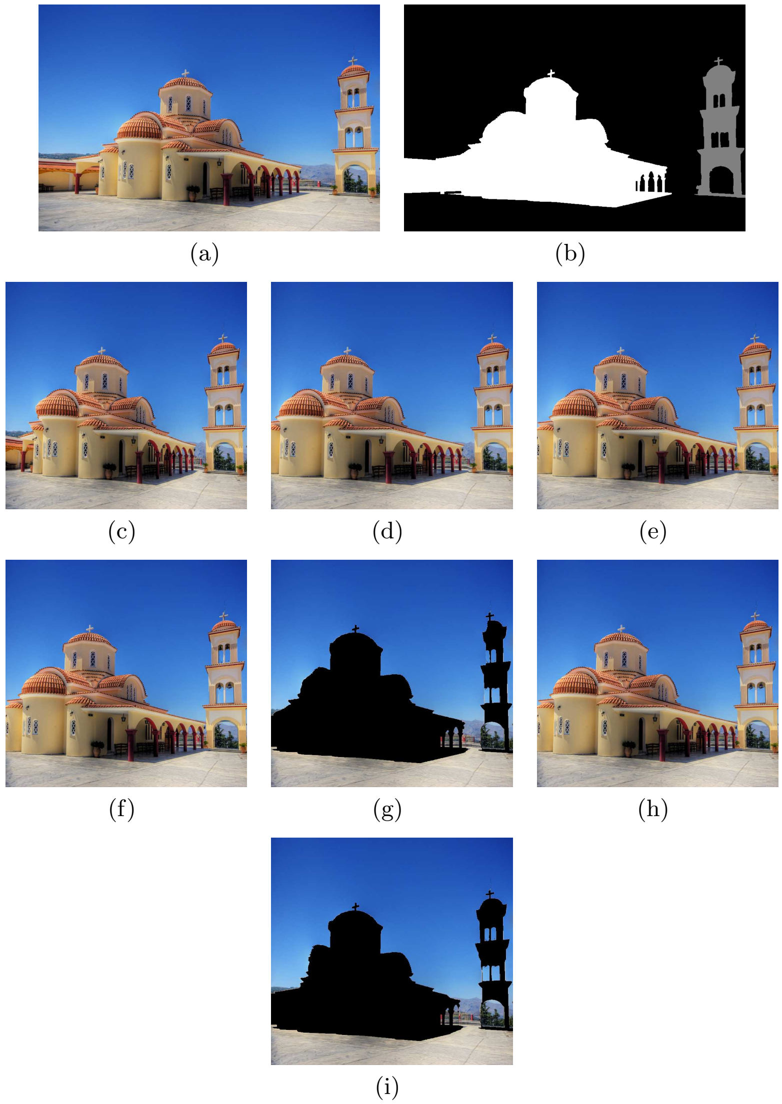


Fig. 19. *Archway* image (a) with depth map (b) retargeted by S.C. (c), S.C.+Obj. Prot. (d), S.C.+Obj. Occ. (e), Sc. Carve-D. (f) with background (g), Sc. Carve (h) with background (i) (300 seams removed)