

# Graph Cut Based Segmentation of Soft Shadows for Seamless Removal and Augmentation

Michael Nielsen and Claus B. Madsen

Laboratory of Computer Vision and Media Technology, Aalborg University, Denmark  
{mnielsen, cbm}@cvmt.dk  
<http://www.cvmt.dk>

**Abstract.** This paper introduces a new concept within shadow segmentation for usage in shadow removal and augmentation through construction of a multiplicity alpha overlay shadow model. Previously, an image was considered to consist of shadow and non-shadow regions. This makes it difficult to seamlessly remove shadows and insert augmented shadows that overlap real shadows. We construct a model that accounts for sunlit, umbra and penumbra regions by estimating the degree of shadow. The model is based on theories about color constancy, daylight, and the geometry that causes penumbra. A graph cut energy minimization is applied to estimate the alpha parameter. Overlapping shadow augmentation and removal is also demonstrated. The approach is demonstrated on natural complex image situations. The results are convincing, and the quality of augmented shadows overlapping real shadows and removed shadows depends on the quality of the estimated alpha gradient in penumbra.

**Keywords:** shadow segmentation, graph cuts, augmented reality.

## 1 Introduction

The methods that are investigated in this paper are part of an idea to augment real images with virtual objects. If these have to look believable their shadows must look like the real shadows. For this purpose the light sources must be known and they can be estimated by detecting the real shadows. It is also necessary to know the degree of shadow for any given pixel, otherwise augmented shadows will stack upon real shadows.

There are many applications that benefit from shadow detection. For example segmentation of foreground objects without obstruction from shadows, classification of e.g. faces with shadows that could make it difficult to find the best match, and extraction of illumination such as light source direction and color. We want to find a model that can be used for shadow segmentation as well as shadow synthesis.

Our aim is to make a purely pixel driven method which works on single images in un-augmented scenes with no geometric knowledge about the scene. However, we will assume outdoor illumination.

## 1.1 State of the Art

Salvador [1] distinguished between cast shadows (onto the ground plane) and self shadow. The detection relied on the edge image of a linearized chromaticity image. They considered dark pixels a-priori to be shadows and corrected this belief using heuristics concerning the edges of the real image and edges of the chromaticity image. This worked well in images with controlled simple geometry. It was tested with still images (of fruit) and video (moving people and cars).

Madsen [2] described shadows as an RGB alpha overlay. It is not just a black layer with an alpha channel, because the shadows are not only darker versions of the illuminated areas, but there is a change of hue, caused by the difference in hue between direct and ambient light. There is a fixed alpha for any given region.  $\alpha$  can be described as the degree of shadow and the overlay color relates to the tonal and intensity change of the shadow. Furthermore, shadows are characterized as full shadow, *umbra*, and half shadow *penumbra*, assuming only one light source. Multiple light sources would generate more complex grades of shadow regions.

Finlayson [3] takes advantage of planckian light and retinex theory. Assuming a single direct light source and another ambient light (different hue) computes a 1-d invariant image from the known path (illuminant direction) the shadow imposes on a 2-d log-ratio chromaticity plot. Note that ambient occlusion and surface normal direction is not taken into account in this model. The known path/offset is to be pre-calibrated. The edge maps of this invariant image can be used just like in [1] or to set a threshold in a retinex path. The results were images that looked flat and unsaturated with attenuated shadows and blur around the boundaries. The detected boundaries were high quality. Later an automatic initialization was developed using entropy minimization [4].

Cheng Lu [5] continued Finlayson's work using graph cuts for optimizing the shadow mask. Their method finds a binary shadow mask and use the illumination invariant chromaticity transform [3] as a static clue for computation of the capacities of the capacities in the graph model. They do not use any data term but considered the brightness changes in the means of windows around the supposed shadow edges as well as the chromaticity "shift" caused by the illumination color. It is not tested for difficult scenes and it does require knowledge of log illumination direction.

Previous work segmented shadows as a binary mask and used the edges as clues. They tested their algorithms in simplistic setups and had strict requirements to their cameras. We will use an  $\alpha$ -overlay shadow model based on Finlayson's color theory, which is the most versatile approach. Our model must be invertible so that it can be used to generate shadows and to remove shadows. The degree of shadow should be adjustable through the  $\alpha$ -parameter. The color theory does not explain what happens in the penumbra region. In the following the main theory is presented (for more detail refer to [3]) and the penumbra region is investigated. Followed by a graph cut algorithm for estimation of the  $\alpha$ -channel for a series of natural images.

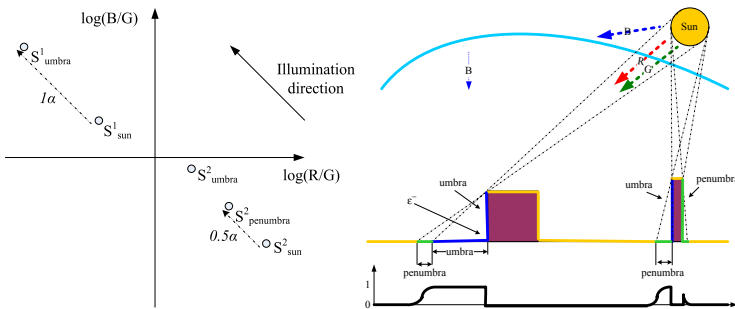
## 2 Methods

The color of a surface in full shadow is assumed to be a product of its color in sunlight and a fixed shading factor [6]:

$$\begin{bmatrix} R_{shad} \\ G_{shad} \\ B_{shad} \end{bmatrix} = \begin{bmatrix} \alpha R_{sun} \\ \beta G_{sun} \\ \gamma B_{sun} \end{bmatrix} \tag{1}$$

This relation holds for all pixels in the image. It follows from this generalization that in log chromaticity space there is an illumination direction vector added to the sunlit pixel, because  $\log(a * b) = \log(a) + \log(b)$ . See figure 1. However, it will be necessary to be able to weight the shading effect by an  $\alpha$  in the penumbra areas. The left side shows the illumination direction in log chromaticity space shows the direction that a surface color changes in the 2-d plot when the color temperature changes from sun to shadow. It follows a straight line. Surface 1 ( $S^1$ ) is plotted in sun and shadow (umbra). We extend the model to account for varying degrees of shadow. Surface 2 ( $S^2$ ) is plotted in sun, half shadow (penumbra), and shadow (umbra). However, tonal changes from ambient occlusions and inter-reflections do not follow the straight line. Ambient occlusions adjusts the color towards  $[0, 0]$ , while inter-reflections adjusts the color toward the chromaticities of the reflecting surfaces.

The right side (in fig. figure 1) shows the geometry that causes umbra and penumbra regions and how our  $\alpha$  overlay model should respond to those regions (shown at the bottom of the figure). In the sun  $\alpha = 0$  and in the shadow  $\alpha = 1$ , and the in penumbra region  $\alpha$  is proportional to the visible area of the sun, hence an S-shaped curve<sup>1</sup>. In the corner at the first box is a situation where the shadow becomes darker because the hemisphere is less accessible from those locations. This is not accounted for in the model.



**Fig. 1.** [Left] 2-d log chromaticity space and illumination direction. [Right] Geometry causing penumbra and umbra.  $\epsilon^-$  denotes ambient occlusion, which is not handled by the theory.

<sup>1</sup> For more detailed explanation [7]

The shadow model has two main tasks: 1. Shadow Augmentation for which the shadow region as a function of the sunlit region is needed. 2. Shadow Removal for which the sunlit region as a function of the shadow region is needed. The following model takes advantage of the relationship described in equation 1. It is adapted to control the degree of shadow with  $\alpha$  in equation 2 (for each pixel  $p$ ).

$$\rho_k^{shad} = (1 - S(\alpha)O_k)\rho_k^{sun}, \quad k = R, G, B \quad (2)$$

The model is easily inverted to represent  $\rho^{sun}$  (also called "shadow free image") as a function of  $\alpha$  and  $\rho^{shad}$ . The S-shaped profile is modeled by a sigmoid function ( $S$ ) in eq. 3.

$$S(\alpha) = \left( \frac{1}{(1 + e^{-(\alpha-0.5)*6.6})} - b \right) * a \quad (3)$$

Scaling factors  $a = 1.074$  and  $b = -0.0344$  were chosen such that  $S(\alpha) = 0$  when  $\alpha = 0$ , and  $S(\alpha) = 1$  when  $\alpha = 1$ .

The only initialization consists of finding the overlay color. The notation of the following investigation of optimal overlay color ( $O = \{o_r, o_g, o_b\}$ ) is simplified: The surface color (albedo) will be denoted  $A$ . The irradiance from the sky will be  $E_{sky}$  and the irradiance from the sun will be  $E_{sun}$ . We consider a sunlit pixel to be  $A(E_{sky} + E_{sun})$  and an umbra pixel to be  $AE_{sky}$ . Furthermore, ambient occlusion and sunlight direction is assumed to be fixed at zero.

$$AE_{sky} = (1 - \alpha_{max}O)A(E_{sky} + E_{sun}) \quad (4)$$

$$O = \frac{\left(1 - \frac{AE_{sky}}{A(E_{sky} + E_{sun})}\right)}{\alpha_{max}} = \frac{\left(1 - \frac{E_{sky}}{E_{sky} + E_{sun}}\right)}{\alpha_{max}}$$

where  $\alpha_{max} = 1$ . We conclude that  $O$  is an umbra pixel divided by its corresponding sunlit pixel.

## 2.1 Estimation of $\alpha$ Via Graph Cuts

We treated the  $\alpha$  estimation as a piecewise smooth labeling problem. An energy function would be minimized using graph cuts. If an energy function can be described as binary variables with regular energy terms with robust metrics, it is fast to find a strong local minimum [8] if the energy terms are regular.

The  $\alpha$  channel of the overlay was estimated through  $\alpha$ -expansions (thus reducing the problem to binary labels). For natural conditions where the illumination and camera settings are not fixed, it is not desirable to rely on pre-calibration of the camera that requires fixed settings for all images. We wish to find an initialization for the algorithm that is tailored for the image at hand without knowledge of the camera parameters. This process should be automatic, but at this stage the color of the overlay was given by manual initialization by handpicking a sunlit surface and its shadow counterpart. The mean red, green, and blue for each region was computed and the overlay color was given by  $1 - \mu_{shadow}/\mu_{sun}$  as in equation 4.

The graph construction is given in [8]. The energy terms for the total energy (equation 5) remains to be defined. There were a number of ideas how to use the data term and the simplest method is to use a constant, because there is no way to tell if a given pixel is shadow. A sunlit black matte surface is darker than most bright surfaces in shadow. We construct the energy implementing three assumptions: 1. Piecewise smoothness as the standard smoothness term ( $V$ ), 2. Attenuation. Reward  $\alpha$  cuts such that shadow edge attenuation occurs ( $A$ ), and 3. Legality. Punish illegal  $\alpha$  cuts that creates edges that were not present in the original photo ( $L$ ).

$$E(f) = w^- V^{p,q}(f) + w^+ A^{p,q}(f) + w^- L^{p,q}(f) \tag{5}$$

where  $E(f)$  is the total energy of the configuration  $f$  of all variables (pixels). The terms are weighted by  $w^+ = |p - q|$  or  $w^- = N_{max} + 1 - |p - q|$ .  $N_{max}$  is the maximum  $|p - q|$ .

In addition some heuristics about chromaticity edges was used. The overlay color ( $O$ ) corresponds to the illumination direction vector in [3] ( $\mathbf{d}$  eq. 6). All edges between pixels ( $p$  and their neighbors ( $q_b$ ) the direction from the log chromaticity plot ( $i^p$ , eq. 7) of the lightest pixel ( $i^+$ ) to the log chromaticity plot of the darkest pixel ( $i^-$ ) are found as chromaticity edges ( $\mathbf{d}^{pq} = i^- - i^+$ ). Then the difference between the illumination direction vector ( $\mathbf{d}$ ) and the chromaticity edge was found (eq. 8). If an intensity edge ( $\delta i = |i(p) - i(q)|$ ) between  $p$  and  $q$  was over a certain threshold ( $T_i$ ) and the difference from the illuminant direction was under a certain threshold ( $T_a$ ) then it was considered a shadow edge. This information was stored in a multi-layered shadow edge image. Each layer (bit) represented a neighborhood. Bit  $b$  for pixel  $p$  was set if pixel  $p$  in neighbor pixel  $q_b$  was labeled as a shadow edge.

$$\mathbf{d}_k = \log \left( \frac{1 - O_k}{1 - O_r} \right) \quad k \neq r \tag{6}$$

$$i^p_k = \log \left( \frac{\rho_{shad}(p)_k}{\rho_{shad}(p)_r} \right) \quad k \neq r \tag{7}$$

$$I_{SE}(p_b) = (\delta i > T_i) \& \left( \frac{\mathbf{d} \times \mathbf{d}^{pq}}{|\mathbf{d}| |\mathbf{d}^{pq}|} \right) < T_a \& (\mathbf{d} \bullet \mathbf{d}^{pq} > 0) \tag{8}$$

The resulting shadow edge image was morphologically dilated (bitwise) as to smoothen the penumbra regions. This allowed for a lower angle threshold ( $T_a$ ) and made the window based sampling used in [5] unnecessary.

Pott's energy term (eq. 9) was used as a global smoothness term ( $V$ ) and is applied for all  $p, q$  in the neighborhood of  $N$ .

$$V_{p,q \in N}(f(p), f(q)) = \begin{cases} 0, & \text{if } f(q) = f(p) \\ \lambda, & \text{if } f(p) \neq f(q) \end{cases} \tag{9}$$

The attenuation term was applied where shadow edges (in  $I_{SE}$ ) were detected. The energy was given by the absolute difference between the neighboring pixels

in the estimated shadow free image (inverted equation 2) without the use of window sampling unlike [5].

$$A_{p,q \in N \cap I_{SE}}(f(p), f(q)) = \min(0, |\rho_{sun}(p) - \rho_{sun}(q)|_{max} - K) \quad (10)$$

where  $\rho_{sun}(p)$  and  $\rho_{sun}(q)$  are the reconstructed shadow free image at pixels  $p$  and  $q$  given a labeling  $f(p)$  and  $f(q)$ .

This smoothness term is not a regular metric at the shadow edges, so it is made regular through truncation. When building the graph, regularity is tested ( $A(1,0) + A(0,1) \geq A(0,0) + A(1,1)$ ). When it is not regular, the energies are manipulated in such a way that the information given by the shadow edge is still maintained while  $A(1,0) + A(0,1) = A(0,0) + A(1,1)$ .

Improbable alpha cuts are punished by the mean gradient of the color bands  $|\rho_{sun}(p) - \rho_{sun}(q)|_{\mu}$  in eq. 11.

$$L_{p,q \in N}(f(p), f(q)) = \begin{cases} 0 & , \text{if } f(p) = f(q) \\ |\rho_{sun}(p) - \rho_{sun}(q)|_{\mu} & , \text{if } f(p) \neq f(q) \end{cases} \quad (11)$$

To avoid problems with soft edges a neighborhood system that resembles pyramid scale space jumps was used. It was based on a 4-connected neighborhood (having two neighborhood relations;  $x - 1$  and  $y - 1$ ) and was extended to 4-layers by adding neighborhood relations at  $-2$ ,  $-4$ , and  $-8$ . This should cover soft shadow edges 8 pixels wide. If the penumbras are wider than that, more layers can be added ( $-16$ , etc.).

$V$  and  $L$  are weighted inverse to the distance because the farther away a neighbor is, the more likely it is that it is not supposed to be smooth.  $A$  is weighted more if the neighbor is far away in order to maximize the inner shadow degree of the shadow region with a broad penumbra.

Neighboring surface colors that are mapped into the same constellation in the chromaticity plot will be a source for error in the method.

## 2.2 Augmentation

It is easy to augment shadows once the initial shadow degree ( $\alpha_0$ ) and its corresponding shadow free image is known. A new soft shadow region ( $\alpha_j, j > 0$ ) is constructed by choosing the geometric shape of the shadow. There should be a linear slope from 0 to 1 through penumbra. The new region should be added to the initial shadow region.

$$\alpha_{aug} = \min(1, \alpha_0 + \sum_{j \in A} \alpha_j) \quad (12)$$

where  $\alpha_{aug}$  is the augmented *alpha* map and  $A$  is the set of regions ( $A = \{a_1, a_2, \dots, a_N\}$ ) that should be included in the augmented *alpha* map.

To remove a region of shadow it can simply be subtracted.

$$\alpha_{aug} = \max(0, \alpha_0 - \sum_{l \in D} \alpha_l) \quad (13)$$

where  $D$  is the set of regions ( $D = \{d_1, d_2, \dots, d_N\}$ ) that should be deleted in the augmented  $\alpha$  map.

In the experiment the augmenting shadow regions were constructed by painting shadows using the mouse to add or subtract a pyramid shape where the mouse pointer was. It is easy to find connected shadow pixels using traditional connected pixels labeling algorithm on the alpha map and delete entire shadow regions by pointing out individual regions.



**Fig. 2.** Results trying to remove the shadow from a synthetic shadow image. [Left] Image (top) and detected shadow edges (bottom). [Middle] Estimated shadow free image (top) Estimated  $\alpha$  (bottom) Black is  $\alpha = 0$  and white is  $\alpha = 1$ . [Right] Augmented shadow image.

### 3 Results

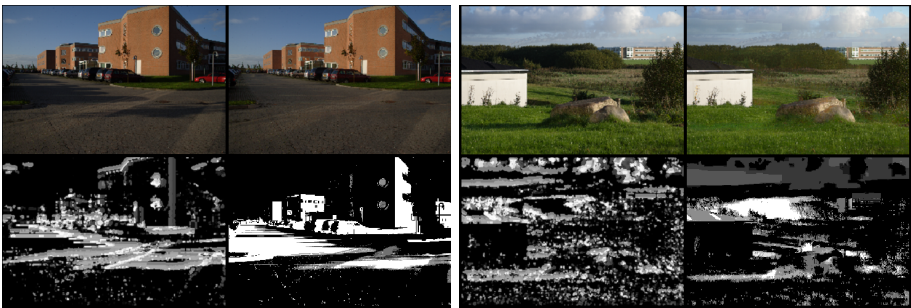
The test was performed using  $T_i = 10$ ,  $T_a = 0.5$ , dilation  $5 \times 5$  structural element.  $K = 255$  and the  $\alpha$  resolution is 10 degrees of shadow. Shadow estimation figures consists of four sections: (left-top ) The Image. (Left bottom) detected shadow edges. (right top) Estimated shadow free image. (right bottom) Estimated shadow degree. Black is no shadow and white is full shadow (umbra). Anything in between is penumbra. The linear professional camera was Canon EOS-1D Mark II. Images were stored in RAW format and converted into 24 bit bitmaps.

Figure 2 shows a synthetic image that resembles three squares of a Macbeth color checker pattern. The shadow (on top) is generated with the sigmoid shadow model. The overlay color was chosen from the mean of the shadow area and the mean of the sun area. The algorithm is applied to estimate the shadow degree and reconstruct the shadow free image. The algorithm is not cheated by the near-black albedo. When the shadow was estimated, some of the shadow was erased and new shadow was augmented on top of the old shadow.

Figures 3 through 4 shows natural examples captured by a linear professional camera in RAW format converted to 8bit per channel bitmap images. They are typical images with concrete and grass surfaces. The result is not very sensitive to lambda (fig. 3). The grass is not really diffuse surfaces, but the results are good anyway. Windows are very difficult. Especially when they reflect something in shadow, but is it really a false positive when this is estimated as shadow? The same question stands for the dark bottom of a cloud in fig. 4. The windows also cause gross inter-reflections, causing an increase of light in both shadow and sunlit regions.



**Fig. 3.** Professional camera. RAW converted into bitmap. Manual settings. [Left] Low lambda ( $\lambda = 1$ ). [Middle] Higher lambda ( $\lambda = 3$ ). The photographers leg is segmented as full shadow, but the shadow from the small thin pole in front of the blue container is deleted by the over-smoothed non shadow region. [Right] The shadow of the photographer has been removed.



**Fig. 4.** Professional camera. RAW converted into bitmap. Manual settings.

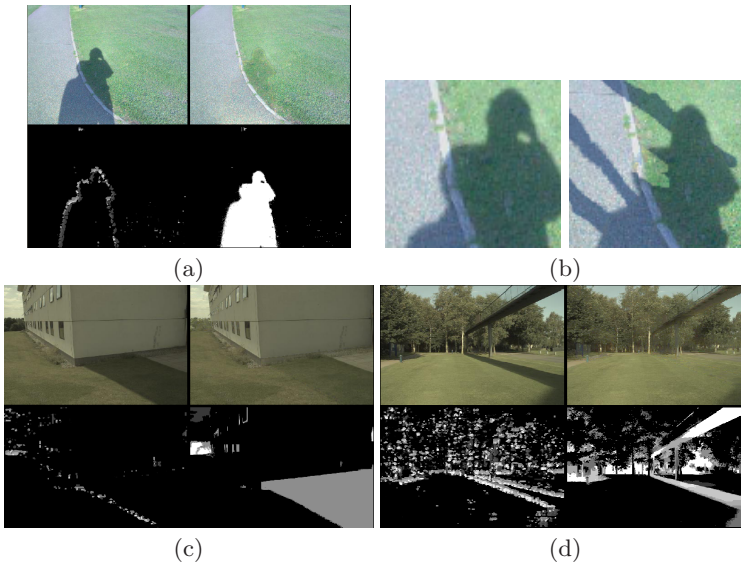


**Fig. 5.** Holiday photos. JPG compression. Automatic Settings.

The overlay color was optimally chosen from a neutral surface with little hue such as road or pavement when applicable. In fig. 4(right) the grass was used for initialization.

Figure 5 shows results from holiday photos captured by a consumer camera (Minolta Dimage 414). JPG compressed and automatic settings regarding white balance, focus, exposure, and shutter. Figure 5(right) was initialized differently





**Fig. 6.** (a,c,d) Our algorithm used on images from [3][6]. Screenshot from PDF papers, JPG compressed. (b) Seamless erased and augmented shadows.

from the others to see how it worked; the underexposed jacket was used. The result was good despite this bad initialization.

Figure 6 shows results on images from [3][6] for comparison. It shows clearly that our method yields more natural looking results; near perfection when the shadow degree is found correctly.  $T_i = 30$  in fig. 6(a) otherwise the high frequency texture was obstructive. The lower right corner of the bridge image shows that the shadow degree is underestimated when penumbra is too wide for the neighborhood system to cover.

## 4 Conclusions

Our contribution estimates a mask of arbitrary levels of shadow for a variety of natural conditions. The model was usable for augmentation and removal of shadows in natural images. The theoretical limitation of the model is that it is constrained to diffuse surfaces and takes no account for ambient occlusion and surface normals. However, qualitatively the method estimates the degree of shadow such that it efficiently minimizes the error from these limitations. Better approximation would require the overlay color not to be fixed. The results in this paper was based on manual initialization of overlay color. However, ongoing research using gaussian mixtures fitted to the log chromaticity plot in combination with the entropy minimization method in [4] shows promising results for future research. High frequency texture is a potential problem, which calls for chromaticity- and edge-preserving smoothing.

## Acknowledgments

This research is funded by the CoSPE project (26-04-0171) under the Danish Research Agency. This support is gratefully acknowledged.

## References

1. Salvador, E., Cavallaro, A., Ebrahimi, T.: Cast shadow segmentation using invariant color features. *Comput. Vis. Image Underst.* 95(2), 238–259 (2004)
2. Madsen, C.B.: Using real shadows to create virtual ones. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 820–827. Springer, Heidelberg (2003)
3. Finlayson, G.D., Hordley, S.D., Drew, M.S.: Removing shadows from images. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV(4) 2002. LNCS, vol. 2353, pp. 823–836. Springer, Heidelberg (2002)
4. Finlayson, G.D., Drew, M.S., Lu, C.: Intrinsic images by entropy minimization. In: Pajdla, T., Matas, J., eds.: In: Proc. 8th European Conf. on Computer Vision, Prague. pp. 582–595 (2004)
5. Lu, C., Drew, M.S.: Shadow segmentation and shadow-free chromaticity via markov random fields. In: IS&T/SID 13th Color Imaging Conference (2005)
6. Finlayson, G.D., Hordley, S.D., Drew, M.S.: Removing shadows from images using retinex. In: Color Imaging Conference, IS&T - The Society for Imaging Science and Technology pp. 73–79 (2002)
7. Nielsen, M., Madsen, C.B.: Segmentation of soft shadows based on a daylight- and penumbra model. In: Gagalowicz, A., Philips, W. (eds.) Proceedings of Mirage 2007, Springer, Heidelberg (2007)
8. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.* 26(2), 147–159 (2004)