EUROGRAPHICS 2014 / B. Lévy and J. Kautz
(Guest Editors)

Volume 33 (2014), Number 2

Light Montage for Perceptual Image Enhancement

Vlad Hosu[1], Mai Lan Ha[2] and Terence Sim[1]

[1]School of Computing, National University of Singapore
[2]School of Maths and Science, Singapore Polytechnic

Abstract

Recent photography techniques such as sculpting with light show great potential in compositing beautiful images from fixed-viewpoint photos under multiple illuminations. The process relies heavily on the artists' experience and skills using the available tools. An apparent trend in recent works is to facilitate the interaction making it less time-consuming and addressable not only to experts, but also novices. We propose a method that automatically creates enhanced light montages that are comparable to those produced by artists. It detects and emphasizes cues that are important for perception by introducing a technique to extract depth and shape edges from an unconstrained light stack. Studies show that these cues are associated with silhouettes and suggestive contours which artists use to sketch and construct the layout of paintings. Textures, due to perspective distortion, offer essential cues that depict shape and surface slant. We balance the emphasis between depth edges and reflectance textures to enhance the sense of both shape and reflectance properties. Our light montage technique works perfectly with a few to hundreds of illuminations for each scene. Experiments show great results for static scenes making it practical for small objects, interiors and small-scale outdoor scenes. Dynamic scenes may be captured using spatially distributed light setups such as light domes. The approach could also be applied to time-lapse photos, with the sun as the main light source.

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Image Processing and Computer Vision]: Enhancement— I.3.3 [Computer Graphics]: Picture/Image Generation —Viewing

1. Introduction

Sculpting with light is a collection of techniques applied by photographers to manually create pictures that reveal greater shape, depth and textures in finely controlled illumination conditions. Artists use Photoshop to composite images from different parts of photos taken under multiple illuminations from the same viewpoint, known as light stacks. This technique extends the freedom of their creativity, but it also raises challenges when the number of illuminations increases to more than a few dozens. Even for the most skillful artists, it is difficult to select the best patches out of many illuminations to create beautiful and realistic compositions.

Existing image-based lighting design methods have tried to simplify the composition process by offering user assisted tools that handle large light stacks. Reaching further we make the composition process fully automatic and produce rich and vibrant compositions. We achieve this in two stages. First we define criteria that evaluate how well each illumination reveals depth, shape and textures. Second we create a light montage technique which seamlessly composites parts of illuminations that best meet the enhancement criteria.

We consider the use of visual cues in creating enhanced montages. It has been shown that depth outlines [LC06, RiHL12] and suggestive contours [CGL08] improve the sense of depth and enhance shape perception. We adapt these features in our criteria for the composition to follow when it selects patches from different illuminations. We hypothesize that by improving the contrast of depth and shape contours we can create enhanced light montages. Moreover, an aspect that has not been considered in the previous rendering approaches is the effect of variable reflectance properties. The distortion of reflectance textures due to perspective strongly affects the perception of shape. Therefore, our method balances the emphasis between depth and shape contours versus reflectance textures. As a result, it produces realistic mon-

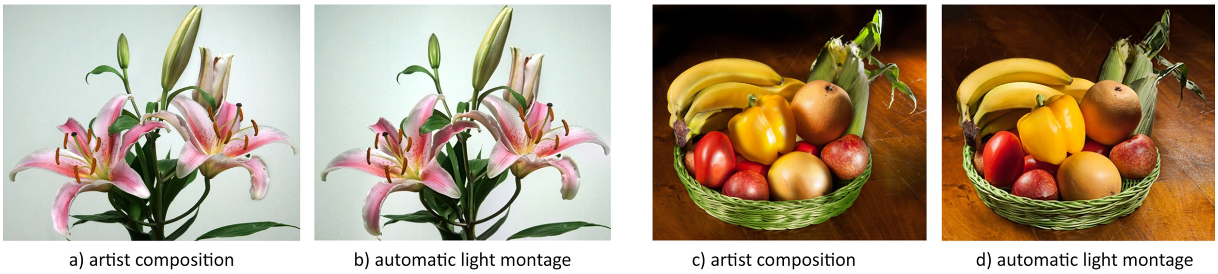

Figure 1: *Artist composition (a,c) vs our light montage (b,d). An experienced Photoshop artist was asked to create a beautiful montage for each light stack (a,b) 32 images and (c,d) 128 images. He took 20-30 minutes to carefully pick images and composite the montage. Our Matlab implementation takes 15 minutes to automatically generates 60 possible montages. We show compositions (b,d) that have a similar look to the one the artist created. The manual compositions have a definite artistic feel e.g. the fruit highlights and shading, whereas our montages appear more photorealistic.*

tages with correct emphasis on details, silhouettes and suggetive contours. In addition to the perceptual criteria, our method has the following characteristics:

- ease of use
  - require only an unconstrained light stack, without additional depth from scanners or RGBN information
  - extract essential edges that improve the perception of shape, depth and textures directly from the light stack
  - use a small set of parameters to automatically create montages that are similar to those that artists would edit manually
- photo realism: control the balance between realism versus the faithfulness of the emphasis
- correct exposure: taking care of artifacts due to incorrect exposure by removing distracting highlights and deep shadows

Existing enhancement methods involve answering the question “How do I make this image look good?”. This requires quantitative knowledge about how to design images. We want to offer photographers a way to create images qualitatively. They need only to answer “Which one of these images looks good?”. We provide users with multiple montages to choose from by emphasizing different aspects of the same scene. We limit the number of results to 60 distinctive montages to prioritize realism in compositions while still meeting perceptual criteria. To evaluate the method, we use the concept of Artistic Turing Test [Bod10]. Computer generated compositions can be compared against user generated artworks. If a program can create artworks that are on par with an artist, then it passes the test.

In summary, our main contribution is a new technique to automate sculpting with light by creating montages that emphasize depth contours and textures as enhancement criteria. In order to achieve this, we propose a method to extract depth edges from unconstrained light stacks. Building on an existing graph-cuts approach we introduce a blending method that seamlessly fades shading while preserving high-pass details. The method modulates the emphasis of depth and shape contours versus textures. It also removes incorrect shadows and distracting highlights to achieve proper exposure and reveal interesting details. As for the results, our light montage method can generate a small number of 60 montages with variable emphasis from which the end user can choose. The montage results are comparable to artist compositions using the same light stacks.

2. Related Work

Our work involves two main topics, one is light stack composition and the other is the extraction of essential contours that are relevant for the depiction of shape and depth. On top of that we develop a new two-band blending technique.

Regarding light stack composition, two directions have been studied. One is designing user-interactive tools to help select parts of illuminations and assemble them [ALK*03, ADA*04, BPB13]. The other direction involves fusing image details automatically [FAR06, CC03]. The former creates more realistic results, whereas fusion is easier to achieve automatically. Striving to make the best of both approaches our light montages are realistic, yet automatically assembled. Our method does not require user intervention due to its reliance on essential contours for perception. None of the existing works on light stack composition make use of visual cues algorithmically. For instance Akers et al. [ALK*03] create a user-assisted tool meant for perceptual enhancement. Users pick the illumination direction by selecting image parts with a feathered brush. They have the responsibility of choosing appropriate cues such as well contrasted shading, shadows for silhouettes, etc. Automatic image fusion based techniques maximize details [FAR06, CC03]. This produces sharp results, however photo-realism is reduced. Boyadzhiev et al. [BPB13] choose details more carefully using basis lights as linear combinations of the stack. They maximize stable orientations which reduces the amount of distracting shadows. Nevertheless, perceptual choices are delegated to the end user who is required to provide a rough manual segmentation that localizes basis lights. In addition to illumination stack methods, tone-mapping is applied on exposure stacks to bring out details. Mertens et al. [MKR07] and Paris et al. [PHK11] obtain very good results on HDR images, however their approaches are not compatible with light stacks.

Line drawings or sketches are the main precursor of paintings. Artists start by defining silhouettes and shape lines, then add shading, color and texture. These techniques have sparked the interest of researchers providing the inspiration for many NPR rendering techniques. Since the early work of Saito and Takahashi [ST90] on depiction of shape, many others have been proposed such as suggestive contours [DFRS03], silhouettes and creases [NKD06], apparent ridges [JDA07], etc. All these methods suggest a common principle. Most line drawings that enhance the depiction of shape follow object silhouettes and large gradients of the image intensity [CGL08]. They also strongly correlate with predictions of existing algorithms. We hypothesize that by emphasizing depth and shape edges through lighting we create an enhanced depiction of the scene structure. Moreover, our approach estimates essential contours based on light stacks alone, without using 3D models of RGBN images as required by all the previous methods.

The light stack composition involves assembling non-overlapping patches from different illuminations. Alpha-blending seamlessly transitions between segments however it also fades high-pass details. For fine segmentations it creates unintended side-effects such as reducing overall image sharpness. Burt et al. [BA83] introduce multi-band blending to handle these situations, however their approach produces color bleeding artifacts [BPB13]. To speed up the process and reduce artifacts, Brown and Lowe [BL03] propose a two-band blending technique. They decompose each image into a low and high pass component. The low pass is linearly cross-faded spanning the entire image, whereas the high pass is kept unchanged. The approach cannot be applied to our montages as our light stack images are completely overlapping unlike the partial overlap required by [BL03]. Thus, we propose a new two-band blending approach.

3. Overview

We automatically generate perceptually enhanced light montages from fixed viewpoint variable illumination photos, also known as light stacks. Our montages emphasize visual cues that are essential for perception, particularly those responsible for the depiction of depth, shape and material properties. Without the need for 3D reconstruction, we extract essential edges from unconstrained light stacks. These consist of silhouettes, suggestive outlines and textures.

Shading is strongly dependent on lighting conditions. Different illuminations reveal a full set of outlines and soft gradients that depict the boundaries and shapes of objects. Therefore, we first construct an intrinsic shading image for each light instance. Then we compute shape edges based on shadow boundaries. Textures are derived from intrinsic reflectance images [Wei01].

According to professional photographer Harold Ross [Ros] who is one of the promoters of sculpting with light techniques, a session of creating montage images involves the artist masking and blending multiple illuminations of a scene in Photoshop. For best results and to simplify the compositing process one usually picks several key images that best reveal shapes or details of different parts of objects. We take a similar approach, however the keen eye of the artist is replaced with indicators of perceptual enhancement such as essential edges.

Guided by essential edges the composition is left to our optimization technique that chooses and blends suitable patches without user assistance. We formulate our method as a Markov Random Field optimization. A fast two-band blending method seamlessly joins segmented light mattes into the final montage without decreasing sharpness at seam-transitions. To quickly generate multiple montage options with varying emphasis we run the optimization at a lower spatial resolution than the output images.

We generate several montage images with variable characteristics such as the emphasis on depth edges versus texture, level of overall contrast as well as the amount of shadows and highlight that are removed. Users need only to choose among several fully composited output results . We show that our compositions are similar to what artists create using the same photo-sets. The results are also compared with other methods that enhance images using light stacks. Our compositions are better in terms of clearer details, balanced contrast, correct exposure and most of all pleasant photorealistic montages.

4. Implementation

We first describe the setup required to collect light stacks, followed by the definition of essential details and corresponding methods to compute them automatically. After that, the required extension to the MRF energy function are presented together with two-band blending technique that smoothly transitions shading across segment boundaries. Lastly, we discuss about challenges in implementation efficiency and our solution to address the problem when the light stacks contain hundreds of images.

4.1. Input Acquisition

The usual approach to capture light stacks for the sculpting with light technique is to take long exposures while moving a hand-held light source to illuminate different parts of the scene from different directions. The photographer uses a

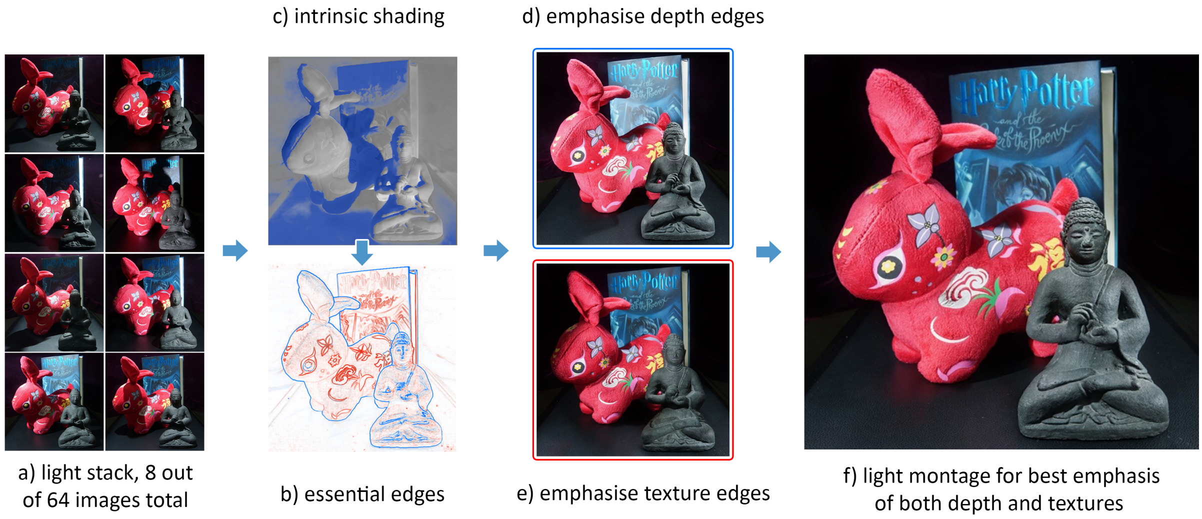

Figure 2: *Overview of our approach. We extract intrinsic shading images c) from the entire light stack, shadows are shown blue tinted. We detect essential details b) including depth outlines (blue) and reflectance textures (red). Two emphasis extremes are shown for depth d) and textures e). With a* $2:1$ *weighting of textures vs depth we obtain a well emphasized composite f).*

remote/timed trigger with the camera installed on a tripod. Because of the acquisition method, it is unavoidable that the photographer or the light source will be visible in some of the frames. Our technique ensures that transient objects or details are not used in the composition due to the stability of essential edges.

We have tested various lighting conditions such as using point light sources to create hard shadows or diffuse illumination to soften shading. In our experiments we have captured both indoor and outdoor scenes. Each light stack contains from 30 to over 100 illuminations.

4.2. Essential Edges

We enhance images by emphasizing perceptual cues responsible for depicting scene structure and material properties. Our light montage technique allows to composite illuminations to meet particular goals, such as creating local contrasts corresponding to a map of essential edges. Depth edges, which include silhouettes and shape suggestive contours are one type of essential edges. The other are reflectance textures which promote the perception of material properties.

For frontal illumination directions, cast shadows will mostly be attached to the corresponding occluding outline. Detached cast shadows outlines vary greatly in their position, whereas attached ones always overlap parts of silhouettes. Moreover, for lateral illumination, ridges and valleys act as silhouettes from the light source point of view. These are called suggestive contours. They represent cases when a true occluding contour would appear with relatively small changes in viewpoint. For instance in Fig. 3 the ridge of the nose is visible due to shadows produced by illumination from the right. Shadow outlines from lateral illuminations will frequently overlap suggestive contours. Hence, by detecting cast shadow outlines, and analyzing the cases where they frequently overlap we can detect both silhouettes and suggestive contours directly from a light stack.

Suppose we have n images in the light stack $S_1,\dots,S_n$, where S_l is the luminance channel of image l. From these, we first derive intrinsic shading W_l and reflectance images Q_l using Weiss' approach [Wei01]. Using the shading images, we propose a method to detect depth edges as follow:

Algorithm 1 Detecting depth edges J

$W \leftarrow IntrinsicShading(S)$
for all $p \in PixelLocations,\ l \in [1..n]$ **do**
  $H_l(p) \leftarrow W_l(p) - mean(W(p))$
end for
for all $p \in PixelLocations$ **do**
  for all $l \in [1..n]$ **do**
    $X_l \leftarrow 1 - |H_l(p)|$ ▻ shadow edge mask
    $Y_l \leftarrow \nabla H_l(p)$ ▻ shading edges
  end for
  $J(p) \leftarrow median(XY)$
end for

Regions with negative values $H_l(p) < 0$ are likely shadow candidates as shown in Fig. 3. Thus $1 - |H_l(p)|$ acts as a mask selecting shadow outlines which is applied on the edge map of each light stack instance. The reason for taking the median similar to [Wei01] is to extract stable edges. From what we have tested the approach is not overly sensitive to the accuracy of estimating shading images. However, the accuracy of the depth edges is lower when using the luminance channel instead of shading images. The shadow mask derived from H is computed in the same way.

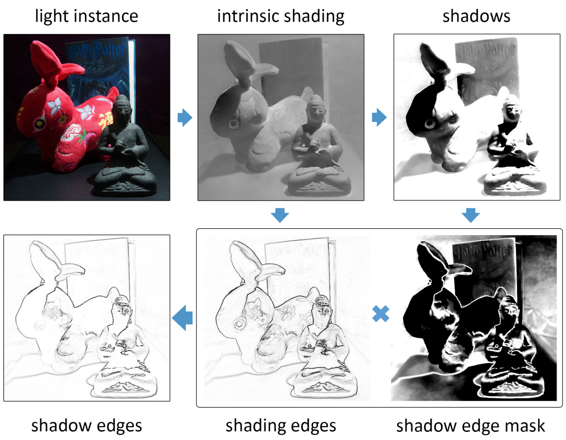

Figure 3: *Shadow edges are derived from intrinsic images: a lighting instance S_l form intrinsic shading H_l, from which a shadow contour mask is formed* $1-|H_l(p)|$.

To find the texture edges Q we multiply the gray level erosion [Hei91] of $1-J$ with the reflectance edges R which are obtained using Weiss' method [Wei01]. We linearly combine the texture and depth edges into the essential edge target to create light montages with different emphasis by varying η:

$$E = \eta J + (1-\eta)Q \tag{1}$$

The essential edges detection could be improved in several ways. First, many illuminations in the stack cover only a small part of the frame. Consequently, other methods could be used for detecting the most frequent edges overlaps, which are depth edges. For instance, the median calculation for depth edges may benefit from assigning a higher importance to strong edges relative to weaker ones. Second, we haven't considered how the HVS assigns importance to edges. Along these lines "Visually Significant Edges" [Av10] introduce a context-sensitive perceptual contrast evaluation method. This may prove to be a useful modulating factor in detecting depth edges.

4.3. MRF Energy Function

To create photo montages we composite parts of images in the light stack. We select non-overlapping patches from different lighting configurations by using an MRF optimisation. This is defined as a labeling problem where we have to choose a source image S_i for each pixel p in the output segmentation. $L(p)$ is the label assigned to pixel p. If two neighboring pixels in the composite have different labels assigned to them, we say that they are part of different segments and are divided by a seam.

MRF optimization problems are defined around an energy or cost function. This involves two terms: a pair-wise label compatibility (C_i) and a data-term or node potential (C_d). The total energy that is being minimized is the weighted sum of the two terms. For a labeling L, the total energy $C(L)$ between two neighboring points p,q is defined as:

$$C(L) = \delta^{-1}\sum_p C_d(p,L(p)) + \sum_{p,q} C_i(p,q,L(p),L(q)) \tag{2}$$

The relative importance of the data C_d and pairwise costs C_i is set by the scaling factor δ which in our implementation is fixed to 100. We use the Alpha-expansion Beta-shrink MRF solver proposed by Schmidt and Alahari [SA11].

4.3.1. Data Cost Function

Recall that E is the essential edge image computed in Sec. 4.2. We interpret $E(p) \in [0,1]$ as the desired level of contrast, with 1 corresponding to 100% gradient magnitude.

The stack of gradient magnitudes for each image in the light stack $G_l = |\nabla S_l|$ is globally normalized to span $[0,1]$. If $E(p) = 1$ we'd like to choose a lighting instance that has the strongest edges $G_{max}(p)$, whereas when $E(p) = 0$ select the weakest edges. We're looking for an E' that works for intermediate values. If we assume that $G(p)$ has a uniform distribution of magnitudes, then:

$$E'(p) = E(p)(G_{max}(p) - G_{min}(p)) + G_{min}(p)$$

The partial data cost function C'_d, excluding the shadows and highlights factor, for each point p with label $L(p)$ is:

$$C'_d(p) = |E'(p) - |\nabla S_{L(p)}(p)||^2$$

Minimizing C'_d will result in gradient magnitudes similar to those specified by the essential edges E'. The montage exposure may at times be less than optimal. It depends on the illumination of the light stack images. If the exposure is correct in most instances then the compositions will also be. In many cases the light source is too dim to cover the entire frame or too bright due to the proximity to objects. This creates montages with regions that are either completely dark or showing many specular highlights. In order to compensate we include a factor for shadows and highlights in the energy function.

We start with the threshold percentages $P_S = 30\%$ and $P_H = 1\%$. We derive binary masks M_S to select shadows and M_H highlights. Our intention is to reduce the deepest shadows and brightest specular reflections. In the case of shadows, we want to select the darkest 30% of all pixels in the stack:

Algorithm 2 Shadow Masks

```
P_S ← 0.3
A ← n × Width(S) × Height(S)            ▷ number of pixels
Find T_S such that count(H < T_S)/A = P_S
for all l ∈ [1..n], p ∈ PixelLocations do
    if H_l(p) < T_S then
        M_S(l)(p) ← 1    ▷ mask sub-threshold brightness
    else
        M_S(l)(p) ← 0
    end if
end for
```

Highlight masks are computed similarly for $P_H = 1\%$. In order to reduce the chance that areas in shadows or highlights are selected we define the final data cost as:

$$C_d(p) = (E'(p) - |\nabla S_{L(p)}(p)| \times M_S \times M_H)^2 \qquad (3)$$

4.3.2. Pair-Wise Compatibility Function

Similar to Kwatra et al. [KSE03] our compatibility function takes brightness similarity as a measure of seam correctness. We compute the compatibility and normalize it to $[0,1]$:

$$C_i = (|S_{L(p)}(p) - S_{L(q)}(p)| + |S_{L(p)}(q) - S_{L(q)}(q)|)^\gamma \quad (4)$$

The realism parameter γ allows to vary the number of generated segments. If γ is 0, the optimization will singularly minimize the data cost C_d without regard to the number of segments needed. When γ is several times greater than 1 the single lighting configuration that best optimized the data cost is selected. Intermediate values for the realism parameter will balance the coarseness of the segmentation with the accuracy in achieving the data cost objective.

4.4. Two-band Blending

We propose a two-band blending method with a shorter blending range for high pass details to maintain their sharpness and a wider one for the low pass that encourages smooth shading transitions. We start from the labeling L resulting from the graph-cut segmentation as shown in Algorithm 3.

We have considered gradient reconstruction to blend seams as implemented by Agarwala et al. [ADA*04]. Their approach produces hue shifts when incompatible gradients are encountered. Alpha-blending does not raise these issues and the results are more realistic. We have tried varying the blending radius by using content-based matting. The results are similar to fixed radius blending, but slower to process. An alternative would be to use the cross bilateral filter for faster anisotropic blur. We have not studied this direction yet.

Algorithm 3 Blending Montage Segments

```
for all p ∈ PixelLocations, l ∈ [1..n] do
    if L(p)=l then
        Z_l(p) ← 1                       ▷ mask each segment
    else
        Z_l(p) ← 0                       ▷ non-overlapping
    end if
end for
for all l ∈ [1..n] do
    Z_l ← ImageResize(Z_l, SizeOf(S_l))  ▷ Bilinear interp.
    Z'_l ← |Z_l ∗ K^σ_gauss|             ▷ diffuse mask for low pass
end for
Z ← Z / Σ_l Z_l                          ▷ normalize to unit sum
Z' ← Z' / Σ_l Z'_l
for all l ∈ [1..n], c ∈ [R,G,B] do
    X_l(c) ← |S_l(c) ∗ K^σ_gauss|        ▷ divide into low pass
    Y_l(c) ← S_l(c) − X_l(c)             ▷ and high pass
end for
for all c ∈ [R,G,B] do
    C(c) ← Σ_l (X_l(c) Z'_l + Y_l(c) Z_l)   ▷ blended montage
end for
```

4.5. Implementation Efficiency

We generate 60 montage options per light stack. Using the default settings, our Matlab implementation takes 15 minutes to complete the task. In order to achieve this goal we use a fast MRF solver and reduce the size of the optimization graph to 100×100 nodes. The initialization stage takes a few minutes to read the stack, build the graph, compute intrinsic images and essential edges. Running the optimization at a smaller spatial scale than the output images has two benefits: it reduces computation time, as well as considers shading for more realistic montages.

In our MATLAB implementation the optimizer takes 10 seconds to compose a light stack of 100 images for a graph size of 100×100. Going over 200×200 does not provide significant quality improvements however increases the computation time roughly linear with the graph size. The upscaling and blending takes about 5 seconds on 600×600px montages. The blending is fast and can use the same segmentation result to output higher resolution images. The segments are upscaled first and then blended. The blending range σ' is proportional to the ratio between the output size and the optimization size. The whole process to generate 60 montages, from light stacks consisting of 100 photos takes 15 minutes. The time required can be reduced significantly by first choosing a subset of the images that show interesting details rather than running it on the full stack.

5. Discussion and Results

The purpose of our work is to generate enhanced light montages that best depict the scene structure while augmenting details such as reflectance textures. The process is directed

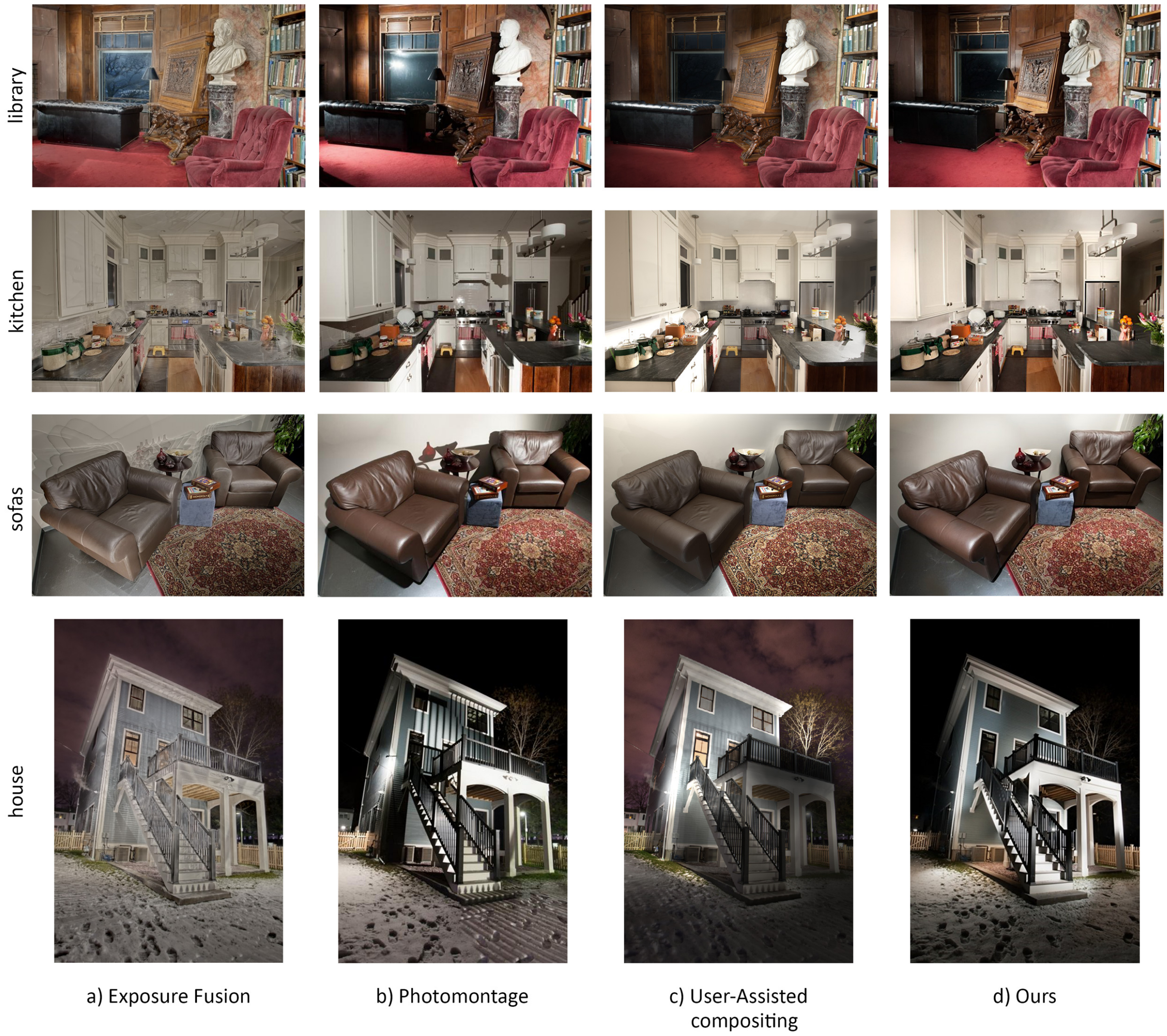

Figure 4: *Comparison of enhancement methods based on variable illumination or exposure. a) Exposure Fusion [MKR07] b) Photomontage [ADA*04] c) User-assisted Image Compositing [BPB13] d) Ours. All methods except for c) are automatic and do not require user assistance to generate the results. The light stacks used in this comparison can be downloaded from the official website courtesy of Boyadzhiev et al. [BPB13]. More results are available in Supplementary Files.*

Figure 5: *Results of multiscale shape and detail enhancement [FAR06] on the same light stack as the above.*

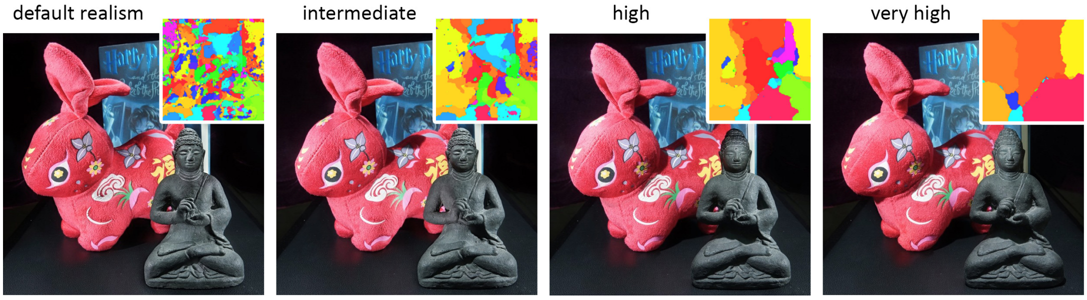

Figure 6: *Levels of realism achieved by increasing the parameter γ from the default value* 0.5 *to* 1, 2 *and* 4. *The optimal segments are shown in different hues, and change in the same sequence as the photos were acquired. Details are most clearly shown when many lighting conditions are used, whereas for fewer illuminations shadows become more apparent hiding some details.*

Figure 7: *Different emphasis criteria for the montages, ranging from textures only to depth only. From left to right, textures are over or under-exposed to favour bolder silhouetes and shape lines. The montage is not reducing either shadows or highlights.*

by simple edge features, yet it affects higher level visual cues. When emphasizing depth, object silhouettes are better defined using well-placed shadows or rim-lights. If textures or suggestive contours are highlighted, grazing illuminations are chosen which depict material properties and surface features. Whenever the essential edges are weak their corresponding contrast is reduced by selecting shadowed, under- or over-exposed regions.

To ensure a desired degree of realism, we've introduced the parameter γ. It controls the number of segments and thus lighting instances used in a montage. It increases realism by choosing a smaller number of seams to join highly compatible lighting conditions. The downside of increasing realism is limiting the composition options which in turn affects how well the enhancement goals are met. For instance, in Fig. 6 the details are clearer when realism is reduced, whereas they become obscured by shadows when realism is increased.

An exclusive emphasis on depth edges is preferred for many of the examples shown, in particular for the images in Fig. 4. In these examples texture is a less important cue for the depiction of shape and depth. Only the sofas example looks better when depth and textures are emphasized in a 2:1 ratio. On the other hand, for images that have many textures, such as the overview example in Fig. 2, a stronger emphasis on textures is preferred when selecting results from the available montages. In Fig. 2 we notice that if either depth or textures are exclusively emphasized the composition is not well exposed. We haven't used the shadow and highlight factors in this composition. For other light stacks, the preference is not clear, textures may be just as important as depth edges as it happens in the basket example, Fig. 7. Most montages will show best results when at least some depth edges are emphasized, so exclusive textures are rarely the case.

It is difficult to balance the contrast of different cues. For instance, when depth outlines are exaggerated, interior surface textures and shading may look flat. To achieve beautiful compositions we generate results that cover several emphasis options ranging from depth edges to textures only, for a total of five targets ($\eta \in [0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1]$), Fig. 7. We obtain better exposure with two levels of reduction for each of shadows and highlights: 30% ($P_S = 0.3$) of the deep shadows or 1% ($P_H = 0.01$) of strong highlights.The contrast factor κ is used to modulate the target edges (E^{κ}), taking values in $\{1, 0.7, 0.5\}$. It non-linearly changes the magnitude of the target edges $E \in [0, 1]$, and thus the overall contrast. For instance, if κ is 0.5, then mid-range edge magnitudes will be potentiated relative to the bottom range. The blending range σ is automatically set to $\max(SizeOut)/\max(SizeOptimization)$.

This brings the total number of parameter combinations to 60. After the user chooses one of the results, he has the options to adjust the realism parameter γ which changes the segmentation granularity. The example montages used in the comparisons do not require any parameter tuning, $\gamma = 0.5$. Montages rendered at $\gamma = 1$ or higher are more realistic, using fewer illumination conditions, see Fig. 6.

5.1. Comparisons

Regarding existing approaches, tone-mapping methods such as Mertens et al. [MKR07] in Fig. 4 column a) and detail fusion Fattal et al. [FAR06] Fig. 5 are not well suited for our light stacks. They produce persistent artifacts such as color bleeding, hue shifts, over-saturation and augment all contours equally including those of hard shadows and highlights resulting in non-photorealistic compositions.

Other methods such as Photomontage [ADA*04] Fig. 4 b) maximize all details equally. High contrast shadow outlines or highlights compete with milder, yet meaningful shape or texture cues. Consequently the results are not well exposed and in many cases do not look realistic. For instance shadows are often incorrectly segmented, house in Fig. 4 and over-exposed regions are selected such as in the library.

The authors of Photomontage do not recommend the method's automatic use. Their main goal is to create a user-assisted method. However, as their software offers the option we include it in the comparison. The default settings are not appropriate for our light stacks, see Supplementary Materials. Adjusting the parameters for best results, we set the compatibility to brightness and tune the application to produce a coarse segmentation, Fig.4 b). The examples show composites of at most 10 patches from different images. For finer segmentations the results quickly deteriorate, Fig. 8.

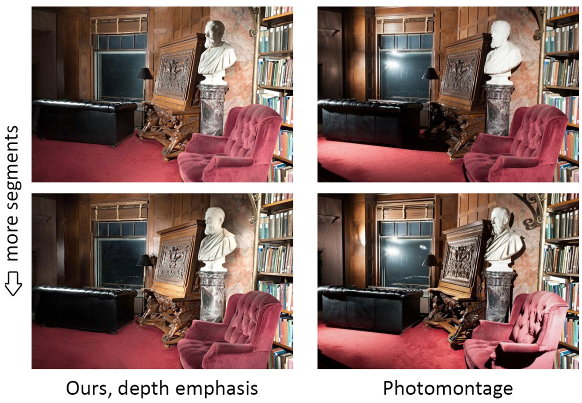

Figure 8: *Our method vs Photomontage [ADA*04]. As the number of seams is increased Photomontage produces more artifacts whereas our approach recreates the desired edge map more accurately. The realism parameter γ is 0.5 for the top montage and 0.2 for the bottom one.*

Boyadzhiev et al. [BPB13] create basis lights as linear combination of the images in the light stack. The approach requires user-assistance to create more compelling compositions by manually segmenting objects and assigning them to different basis lights. Due to the limited contrast and reduced details introduced by weighted averaging, some of the results are under-exposed e.g. the house and sofas in Fig. 4 c). In the kitchen example in Fig. 4 c), by combining images where the lights are both on and off give the top-right lamp an unrealistic look.

Artists are ultimately able to create any lighting design going beyond montages. Our method is limited by its goals and illumination diversity. It emphasizes the structure of the scene first, and then textures and material properties. In the house example in Fig. 4 our approach shows fairly contrasted depth edges, whereas alternative methods insist on random properties of the scene. For instance, users of method c) in Fig. 4 prefer to show the clouded sky, likely due to its emotional appeal. Our reliance on stable depth cues does not emphasize rare details of the light stack. For instance the view outside the window in the library example is visible in only 1 out of 83 total images.

Illustrators prefer to remove shadows in favour of showing more details [ALK*03]. The enhacement criteria including the shadow factor enable us to recreate such effects. Cast shadows are nonetheless reliable cues for object embedding and relationships. In our current approach we haven't explicitly considered this. The primary way to introduce more shadows in our compositions is by increasing the level of realism. This however does not allow to perfectly achieve the enhancement goals. To obtain similar emphasis, when working at high levels of realism, larger light stacks with diverse illuminations are required.

Another important criterium that would favour photorealism is introducing a globally consistent illumination direction. This can be easily incorporated into the data cost function. A brightness factor can keep the overall luminance of the montage close to a chosen lighting instance, balancing this with the other objectives.

More results and comparisons can be found in the Supplementary Material, available as a separate document.

6. Conclusion

We automatically generate a variety of enhanced light montages. This simplifies the user's role and shortens the time required to create beautiful images. The user's task is to pick the most pleasant from the available montage results without requiring the expertise to edit similar compositions.

Our approach stands as proof for the feasibility of automatic perceptual enhancement using light stacks. The dominance of depth cues relative to reflectance textures has a critical role in selecting where the emphasis falls. The re-

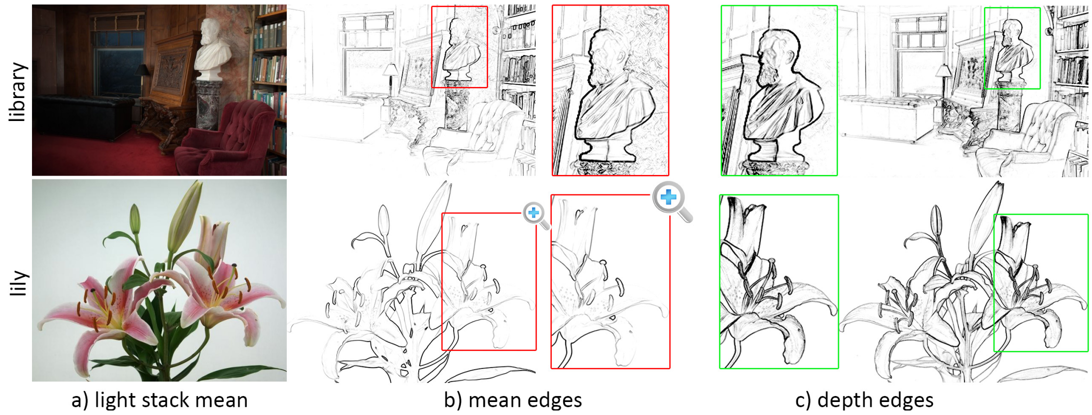

Figure 9: *Results of edge extraction methods. a) mean of the light stack b) edge map of the mean image c) depth edges using our method, gives better results on light stacks where most of the frame is illuminated, such as the lily example.*

sults compare favorably with existing methods and artist renderings, producing more realistic results. They enhance the structure of the scene and highlight material properties.

Beyond this work, the light montage technique can be easily extended to other types of light stacks. It could be used to composite variable exposure stacks, providing a new tone mapping approach. Moreover, it could handle joint exposure and illumination stacks. The technique is fast enough to be applied in video enhancement, frame by frame. A measure of montage consistency would be required for adjacent frames. A more immediate extension is its application to stereo light stacks, captured with stereo cameras. Stereoscopic displays can show two monocularly enhanced images for an ever better 3D experience.

References

[ADA*04] AGARWALA A., DONTCHEVA M., AGRAWALA M., DRUCKER S., COLBURN A., CURLESS B., SALESIN D., COHEN M.: Interactive Digital Photomontage. *ACM SIGGRAPH 23*, 3 (2004), 294–302. 2, 6, 7, 9

[ALK*03] AKERS D., LOSASSO F., KLINGNER J., AGRAWALA M., RICK J., HANRAHAN P.: Conveying shape and features with image-based relighting. *IEEE Visualization* (2003), 46. 2, 9

[Av10] AYDIN T., ČADÍK M.: Visually significant edges. *ACM SIGGRAPH* (2010). 5

[BA83] BURT P. J., ADELSON E. H.: A multiresolution spline with application to image mosaics. *ACM SIGGRAPH 2*, 4 (Oct. 1983), 217–236. 3

[BL03] BROWN M., LOWE D.: Recognising panoramas. *ICCV* (2003), 1218–1225 vol.2. 3

[Bod10] BODEN M. A.: The Turing test and artistic creativity. *Kybernetes 39*, 3 (2010), 409–413. 2

[BPB13] BOYADZHIEV I., PARIS S., BALA K.: User-assisted Image Compositing for Photographic Lighting. *ACM SIGGRAPH* (2013). 2, 3, 7, 9

[CC03] COHEN M. F., COLBURN A.: Image Stacks. 2

[CGL08] COLE F., GOLOVINSKIY A., LIMPAECHER A.: Where do people draw lines? *ACM SIGGRAPH* (2008). 1, 3

[DFRS03] DECARLO D., FINKELSTEIN A., RUSINKIEWICZ S., SANTELLA A.: Suggestive contours for conveying shape. *ACM SIGGRAPH* (2003), 848. 3

[FAR06] FATTAL R., AGRAWALA M., RUSINKIEWICZ S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM TOG 26*, 3 (2006), 51. 2, 7, 9

[Hei91] HEIJMANS H.: Theoretical aspects of gray-level morphology. *PAMI* (1991). 5

[JDA07] JUDD T., DURAND F., ADELSON E.: Apparent ridges for line drawing. *ACM SIGGRAPH* (2007). 3

[KSE03] KWATRA V., SCHODL A., ESSA I.: Graphcut textures: Image and video synthesis using graph cuts. *ACM SIGGRAPH 22*, 3 (2003), 2770286. 6

[LC06] LUFT T., COLDITZ C.: Image enhancement by unsharp masking the depth buffer. *ACM SIGGRAPH 1*, 212 (2006), 1206–1213. 1

[MKR07] MERTENS T., KAUTZ J., REETH F. V.: Exposure Fusion. *PCCGA* (Oct. 2007), 382–390. 3, 7, 9

[NKD06] NIENHAUS M., KIRSCH F., DOLLNER J.: Sketchy Illustrations for Presenting the Design of Interactive CSG. *Visualization* (2006). 3

[PHK11] PARIS S., HASINOFF S., KAUTZ J.: Local Laplacian filters: edge-aware image processing with a Laplacian pyramid. *ACM SIGGRAPH* (2011). 3

[RiHL12] RÖSS ING C., HANIKA J., LENSCH H.: Real-Time Disparity Map-Based Pictorial Depth Cue Enhancement. *CGF 31*, 2pt1 (May 2012), 275–284. 1

[Ros] ROSS H.: Harold Ross Sculpting with Light Photography. 3

[SA11] SCHMIDT M., ALAHARI K.: Generalized Fast Approximate Energy Minimization via Graph Cuts: Alpha-Expansion Beta-Shrink Moves. *Conference on Uncertainty in Artificial Intelligence* (2011). 5

[ST90] SAITO T., TAKAHASHI T.: Comprehensible rendering of 3-D shapes. *ACM SIGGRAPH 24*, 4 (Sept. 1990), 197–206. 3

[Wei01] WEISS Y.: Deriving intrinsic images from image sequences. *ICCV 2* (2001), 68–75. 3, 4, 5