

# Supplemental Material for User-guided White Balance for Mixed Lighting Conditions

Ivaylo Boyadzhiev  
Cornell University

Kavita Bala  
Cornell University

Sylvain Paris  
Adobe

Frédéric Durand  
MIT CSAIL

## 1 Importance of the single-light white balance step

The first step of our pipeline involves a standard single-light white balance iteration. This comes from our observations that after a single-light white balance correction we usually have a few natural looking regions. By fixing those regions, using the correct-color scribbles, users provide important boundary conditions for the later optimization.

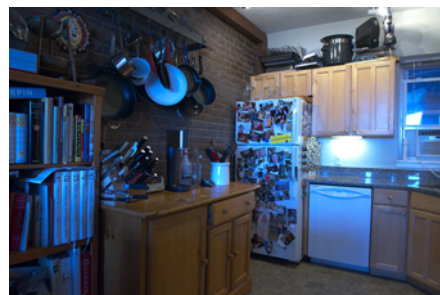
In this section we study how the final result of our system depends on the initial single-light white balance step. Note that we use the same set of scribbles (the one given in the main paper), what changes for each result is the input image on which the scribbles act.



(a) Input (original)



(b) Single-light wb (temperature = 3000)



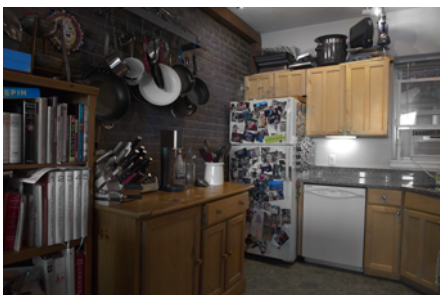
(c) Single-light wb (temperature = 2600)



(d) Our result (no single-light wb)



(e) Our result (warmer)

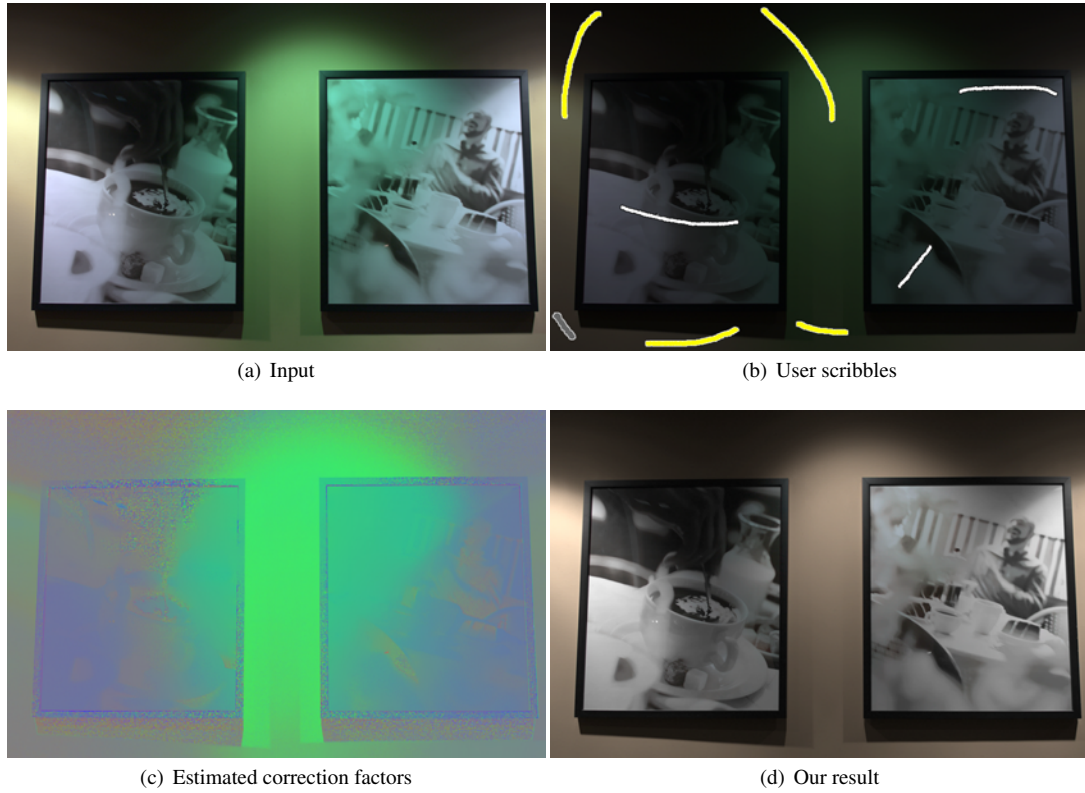


(f) Our result (cooler)

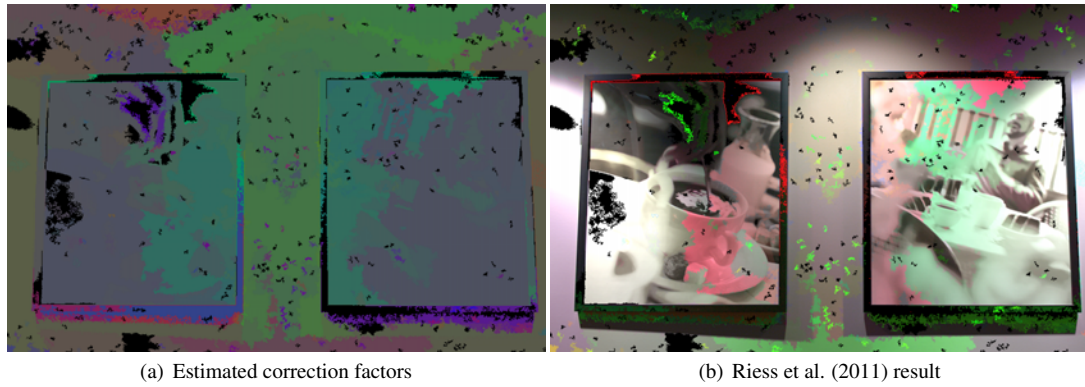
**Figure 1:** Placing the scribbles on the original image (a) affects the boundary conditions of our system. Although this leads to propagation of **non-natural looking colors**, our optimization is still able to remove unpleasant color casts along the way, image (d). However, the mixed lighting is observed through the unnatural looking colors (e.g. the bricks wall, the drawers). Next, we show comparison with two different single-light white balance settings. In the first one, image (b), we apply single-light white balance correction with higher temperature (3000K). This produces image (b), that has warmer colors. In the second one, we use white balance correction with lower temperature (2600K), which gives cooler result, image (c). Our optimization respects the initial data and correspondingly produces warmer result, image (e), and cooler result, image (f). In all cases, the results of our mixed-lighting white balance system looks preferable to the input image.

## 2 Results with comparison

### 2.1 Cafe

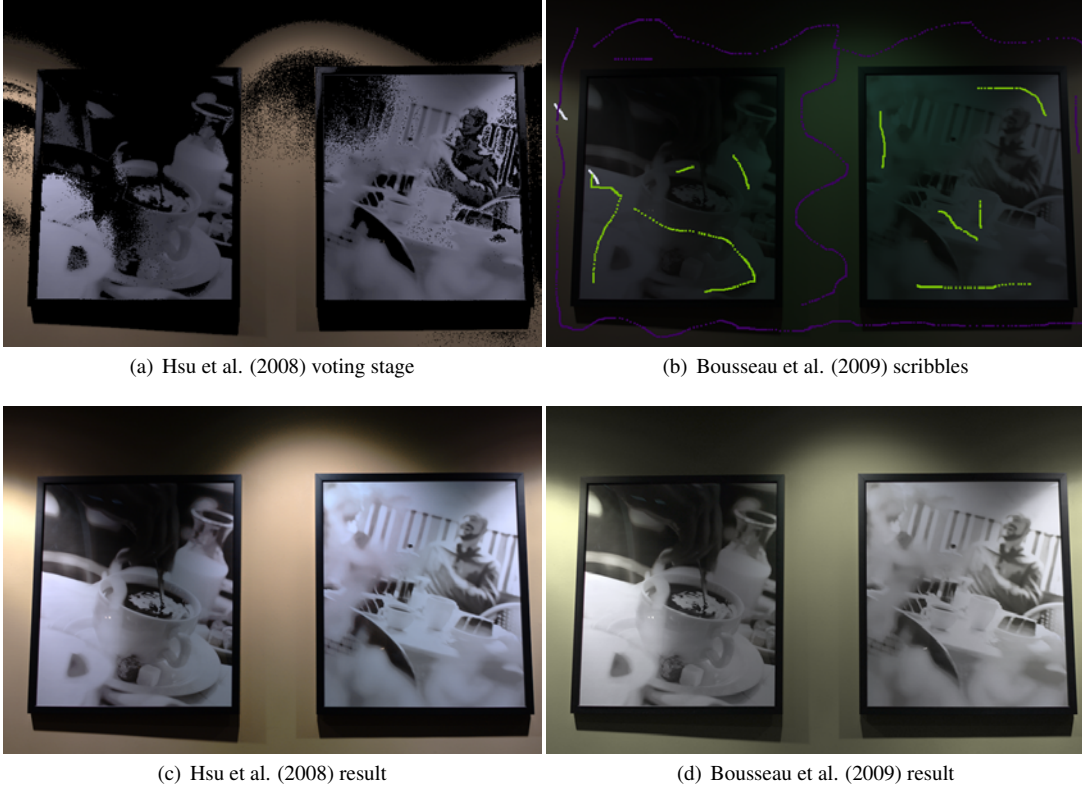


**Figure 2:** This is a fairly simple scene with a mixture of two lights and two major reflectances (the white color and the color of the wall). There is a tungsten light coming from both sides and a strong green light at the center. The camera white-balance settings were set to compensate for the tungsten light, so the input (a) already contains some natural looking regions. Using our systems, with only a few scribbles, image (b), users can produce proper white balance result **over the whole image** (d).



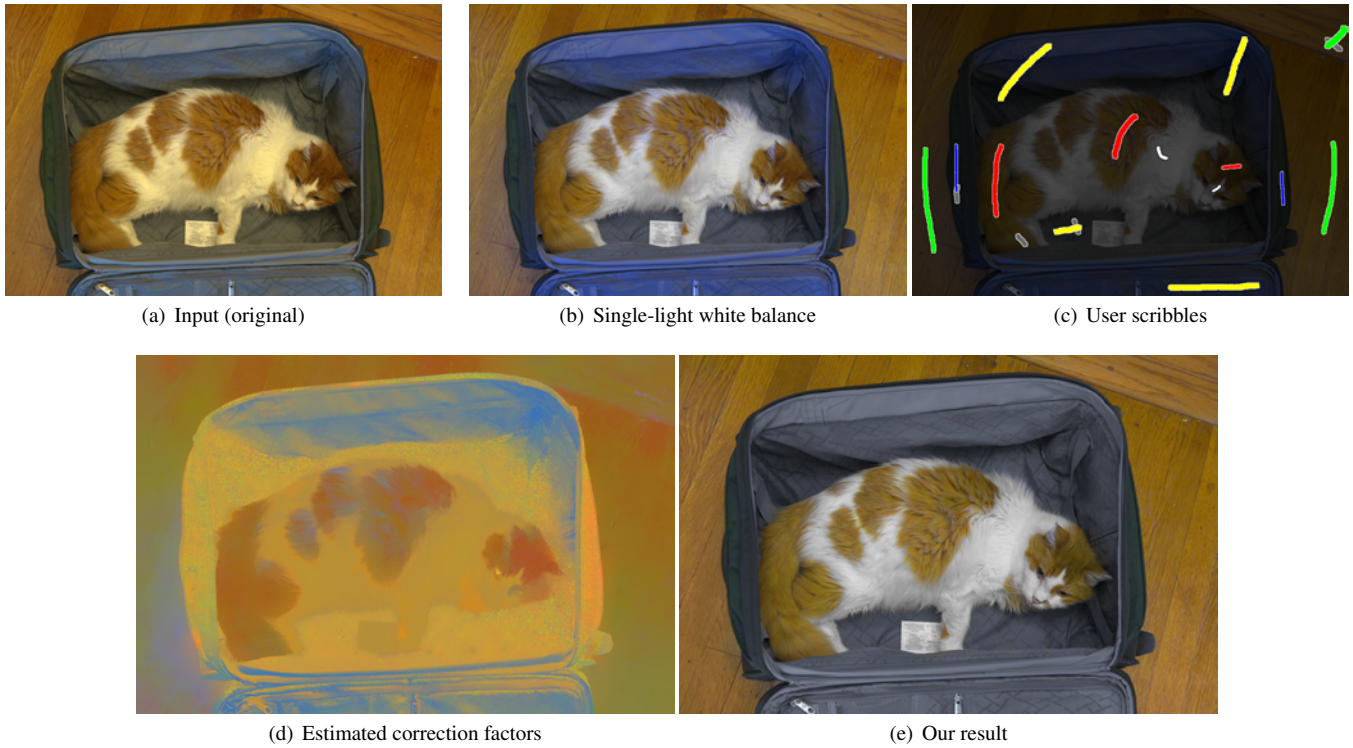
**Figure 3:** We compare this scene against the automatic system proposed by Riess et al. (2011). Notice that their assumption of locally dominant light sources fails to produce good enough (local) lights mixture estimation for the purpose of white-balance.





**Figure 4:** We compare against the automatic system of Hsu et al. (2008), which is restricted to two, known lights. For this scene it was particularly easy to guesstimate the lights, and their voting scheme produces plausible constraints, image (a), which are then successfully interpolated over the entire image (c). Note that selecting the colors of the two lights is not always as trivial as in this case. Our experience with their system shows that it often requires tedious trial-and-error process in order to estimate the two light colors. We also compare against the scribble-based system of Bousseau et al. (2009), which was originally designed for the purposes of intrinsic images decomposition. The result produced by their system is free of color casts, image (d). However, their set of constraints are not restrictive enough to explore the correct color of the wall.

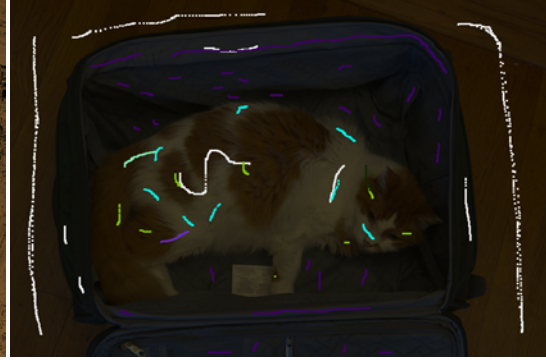
## 2.2 Cat



**Figure 5:** This is one of the images that has motivated our research in this topic. The mixture of outdoor and indoor lights has created unpleasant color casts through the entire image (a). In our system, users first apply global correction, using the available single-light white balance tools, image (b). This step fixes a few regions of the scene, but fails to remove spatially varying lights mixture. Next, users proceed with our scribble based system, image (c), to produce free of color casts result, image (e).



(a) Hsu et al. (2008) voting stage



(b) Bousseau et al. (2009) scribbles



(c) Hsu et al. (2008) result

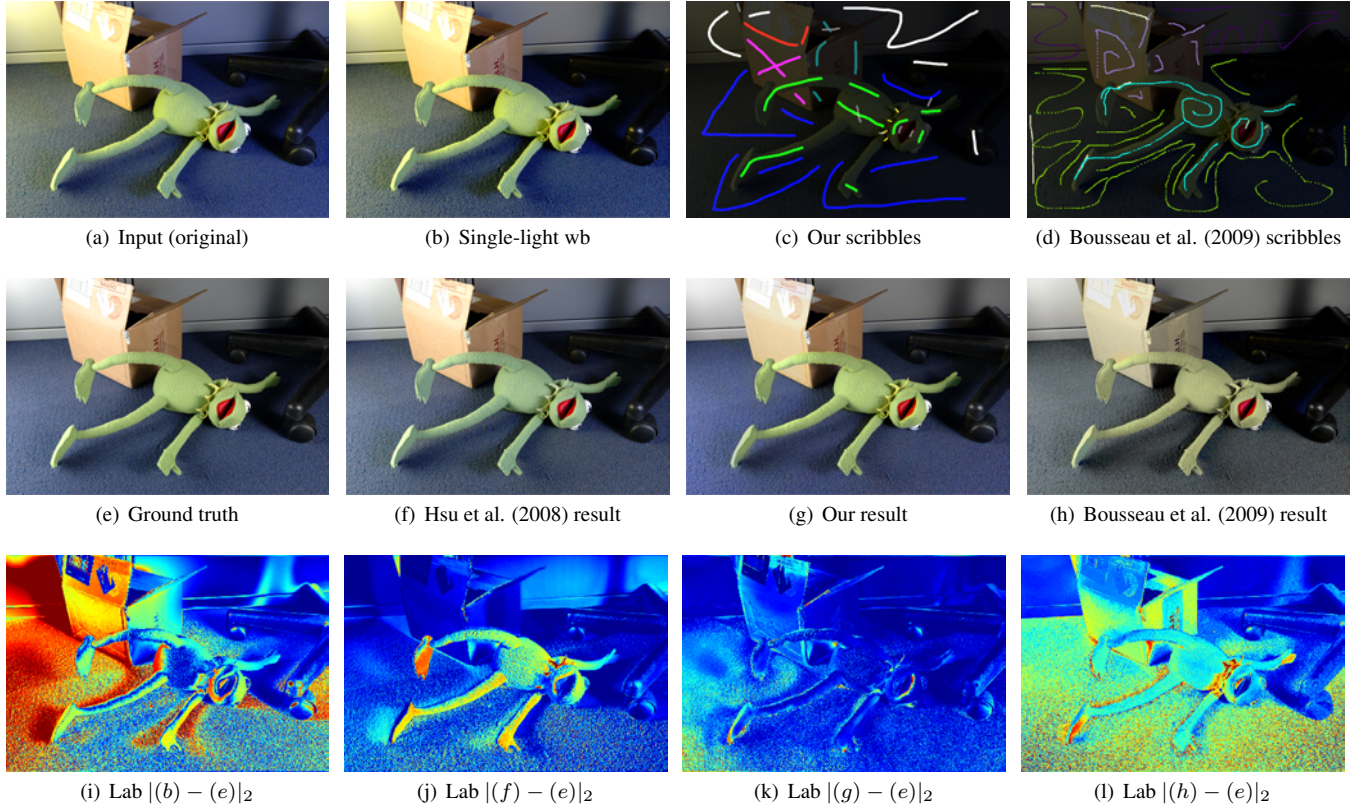


(d) Bousseau et al. (2009) result

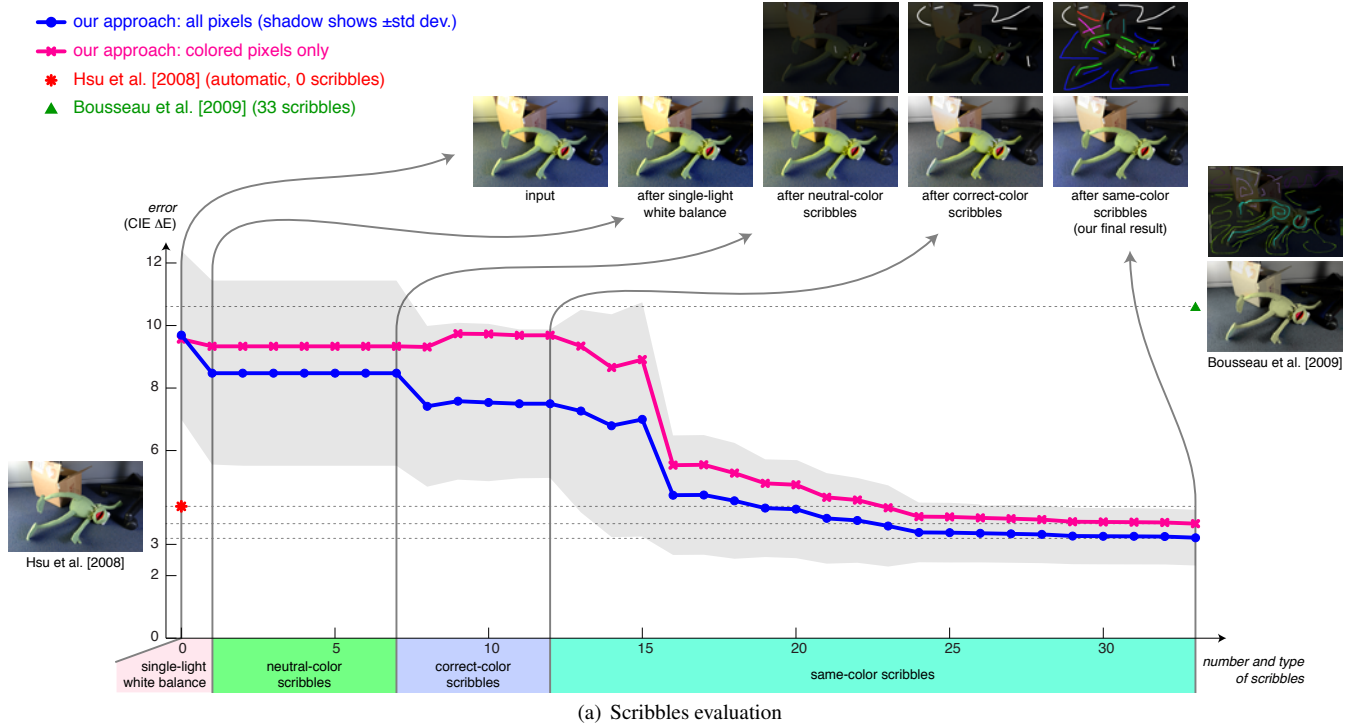
**Figure 6:** *Those competing systems fail to produce high-quality white-balance correction for this real-life scene. In particular, we have observed that the method of Hsu et al. (2008) is sensitive to the initially picked light colors. It was hard to find a pair of plausible lights that will not produce erroneous votes. For a concrete choice of daylight and tungsten light colors, their voting scheme tried to explain the brownish floor and the orangeish fur with the same reflectances. Furthermore, a few white fur pixels were incorrectly assigned to have a brownish reflectance, image (a). As a result of that, the cat’s fur has visual artifacts in their final result, image (c). The method of Bousseau et al. (2009) also fails to properly white-balance the cat, image (d). The reason is that their set of constraints depend on absolute value scribbles, related to the illumination at a pixel. Things like “fully lit” and “smooth illumination” become hard and ambiguous to specify, once you have complex geometry, such as fur or hair.*



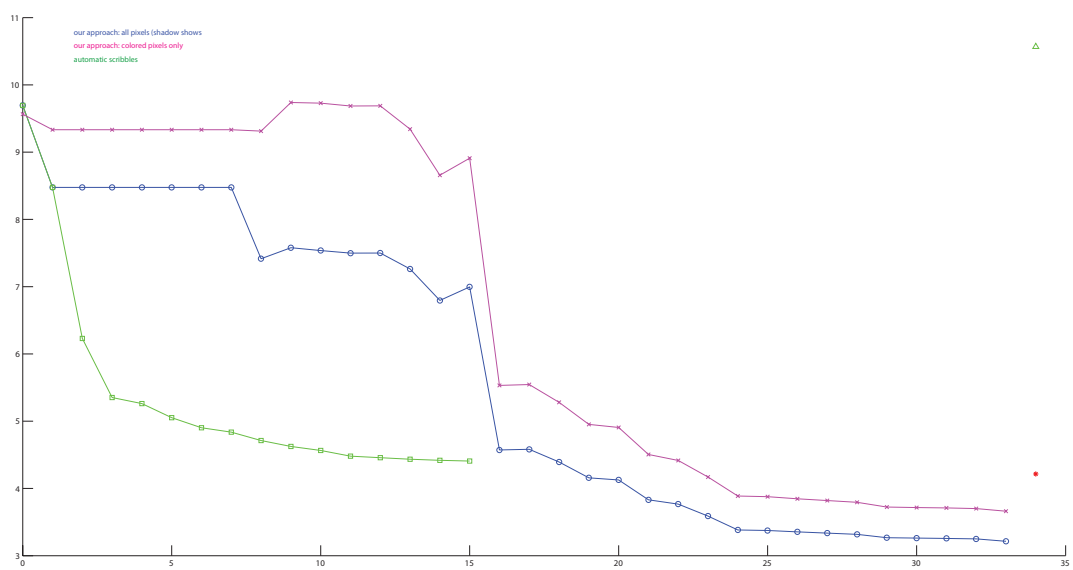
### 2.3 Kermit by Hsu et al. (2008)



**Figure 7:** We evaluate our approach on a ground truth data provided by Hsu et al. (2008). Input image (a) has a mixture of two lights. After the single-light white balance step (b), we obtain a few correct looking regions (e.g. on the bluish floor and on the body of Kermit). Next, users proceed with placing scribbles, image (c). Our method produces natural looking result (g), which is visually and numerically closer to the ground truth (e), compared to other competing approaches (f) and (h). More precisely, the system of Hsu et al. (2008), which is given the exact light colors, fails to vote for the correct green color of Kermit. Furthermore, their interpolation of the illumination mixtures leaves a few color casts on the floor. The system of Bousseau et al. (2009) achieves to remove most of the color casts (with the exception of the regions around Kermit’s leg and the box). However, the derived reflectances are far from the ground truth as their absolute value constraints, based on illumination properties, lack the expressiveness to explore the space of desired solutions for the purpose of white-balance. Furthermore, their results tend to look desaturated and reduce the variety of reflectances in the scene. The Lab L2-norm evaluation confirms that we have produced not only visually, but also numerically closer to the ground truth result.



**Figure 8:** We evaluate the importance of our scribbles on the Kermit ground truth data set. In this concrete example, the user starts by placing a few correct-color scribbles at regions that already look natural. Note that this step does not provide any new information for the image, and the result does not change for the first few iterations. However, those scribbles provide essential boundary conditions, which will push the solution towards a physically correct one, once more constraints are introduced into the system. Next, the user identifies a few neutral looking regions and places scribbles on them. This step fixes a few places (mainly around the neutral regions), but fails to produce globally consistent result. Finally, the user adds a couple of same-color scribbles, which introduce important non-local constraints between distant pixels of the image. The produced set of constraints, **based only on relative reflectance properties**, are enough for our system to produce good looking white-balanced result.

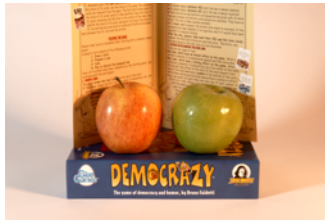


(a) Automatic scribbles placement

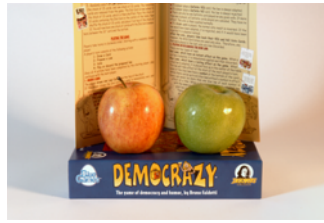
**Figure 9:** Automatic scribbles evaluation, given the same number of steps (15 in this case), explores better way of putting the scribbles and minimizing the overall image error.



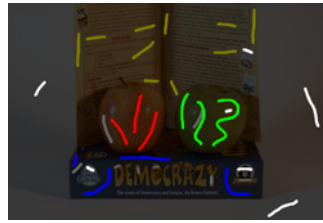
## 2.4 Apples by Hsu et al. (2008)



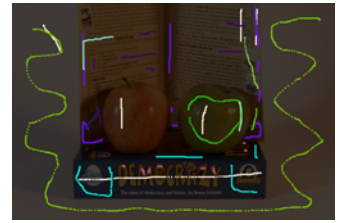
(a) Input (original)



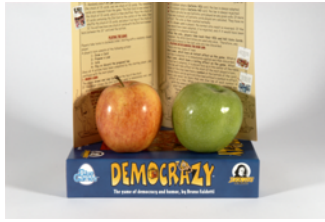
(b) Single-light wb



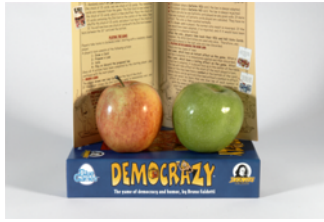
(c) Our scribbles



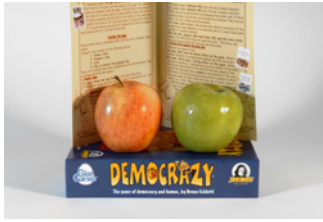
(d) Bousseau et al. (2009) scribbles



(e) Ground truth



(f) Hsu et al. (2008) result



(g) Our result



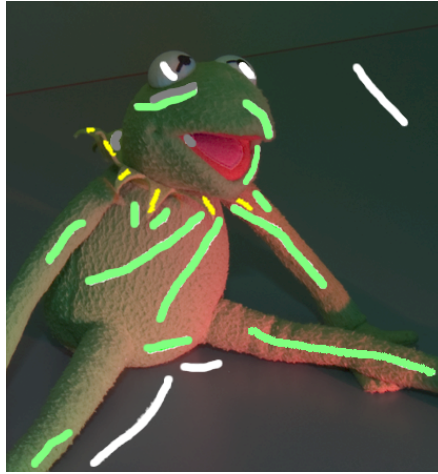
(h) Bousseau et al. (2009) result

**Figure 10:** We evaluate our approach on another ground truth data provided by Hsu et al. (2008). Our user-guided approach (c) produces good looking result (g), without assuming anything about the lights in the scene or their configurations. The method of Hsu et al. (2008), restricted to two known lights, achieves comparable result, image (f). The system of Bousseau et al. (2009), used in its white balance mode, as described in their paper, fails to reproduce the correct colors in this scene.

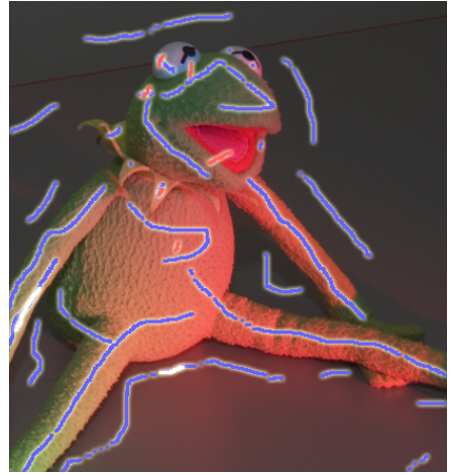
## 2.5 Kermit by Bousseau et al. (2008)



(a) Input (original)



(b) Our scribbles



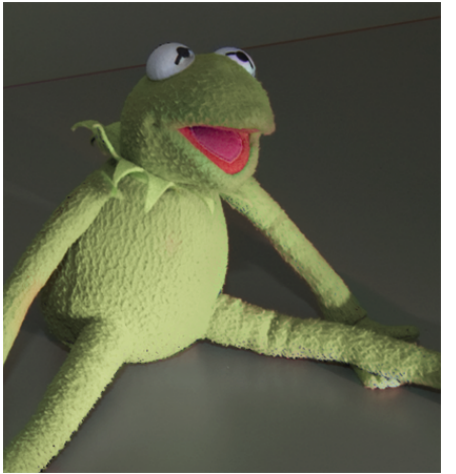
(c) Bousseau et al. (2009) scribbles



(d) Single-light wb



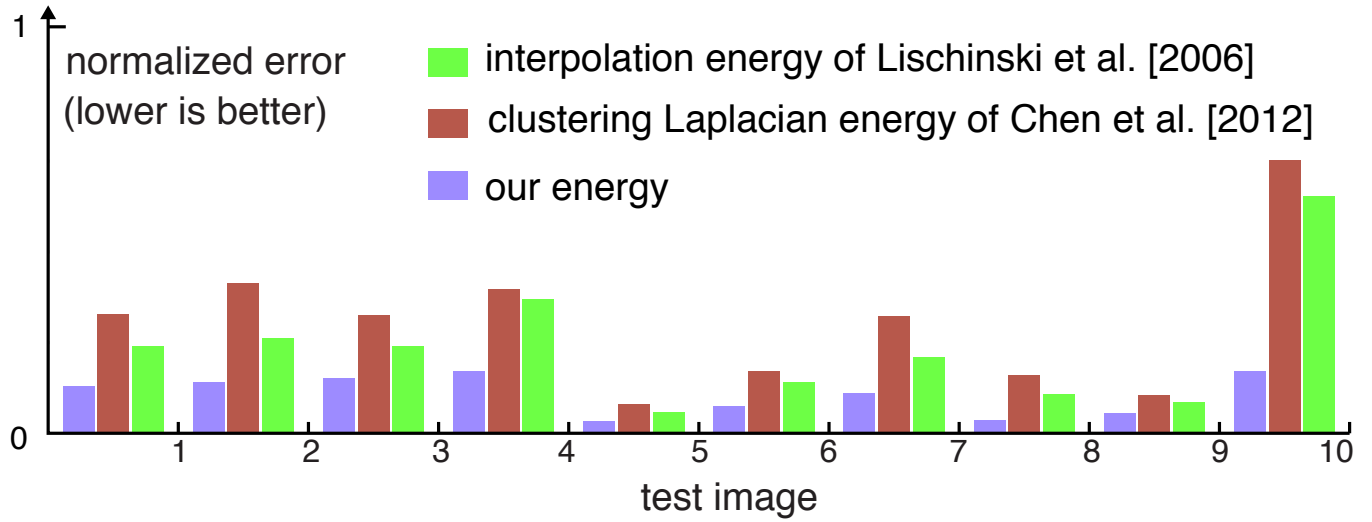
(e) Our result



(f) Bousseau et al. (2009) result

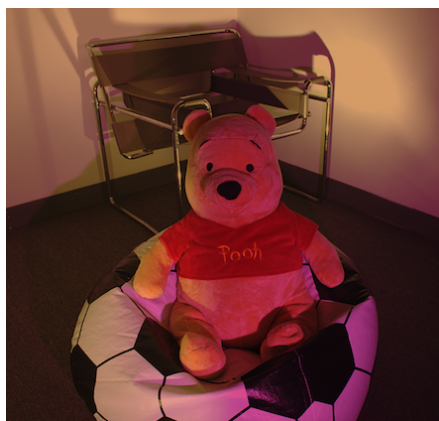
**Figure 11:** Comparing against scene provided by Bousseau et al. (2009). The input image (a) is lit by unknown for us mixture of lights. Using our system (b), we can produce more natural looking result (e), while successfully removing unpleasant color casts, even at hard places around edges. The system of Bousseau et al. (2009) produces result of comparable quality, image (f). However, closer look reveals that their interpolation scheme fails to remove all color casts around edges (e.g. zoom at the left elbow). Furthermore, their final result looks more desaturated. Note that it is not clear how to approach this scene, using the system of Hsu et al. (2008), as we do not know anything about the complex looking lighting conditions.

### 3 Energy comparison

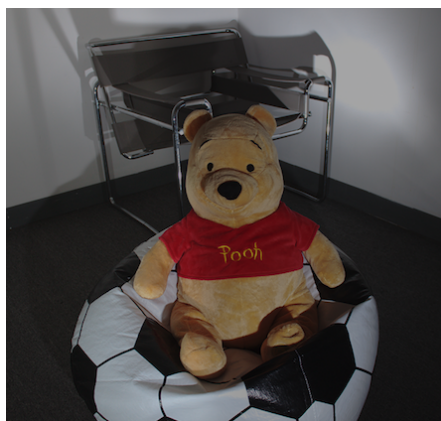


**Figure 12:** We compare our energy designed specifically for our task to the all-purpose interpolation energy proposed by Lischinski et al. (2005) and to the clustering Laplacian energy by Chen et al. (2012). With a sparse set of constraints, our energy interpolates missing values that are closer to the ground truth light mixtures. This corresponds to visually better white balanced results, as shown below.

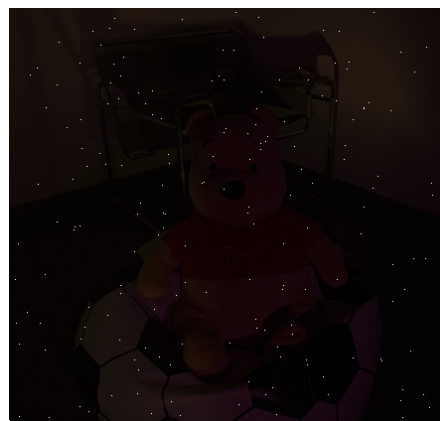




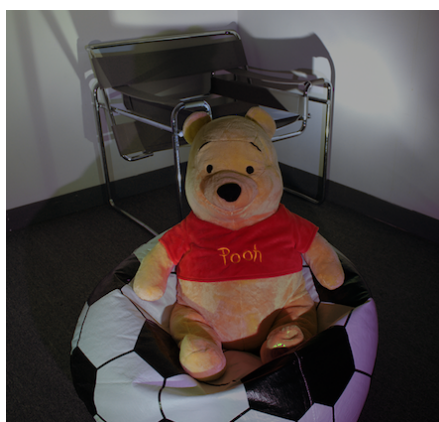
(a) Input (original)



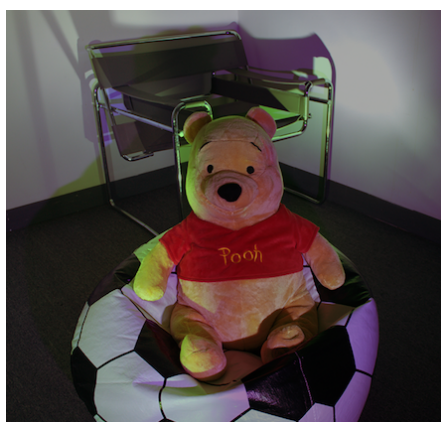
(b) Ground truth



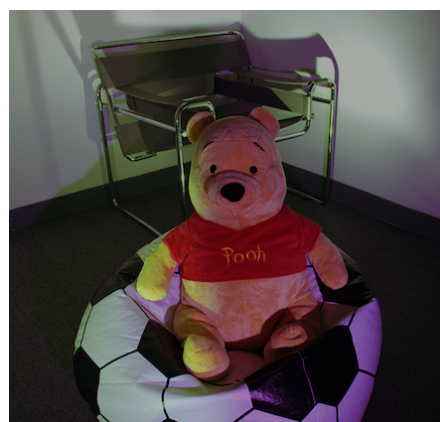
(c) Constraints



(d) Our result

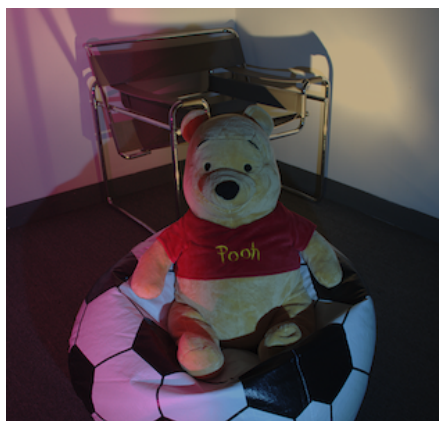


(e) Lischinski et al. (2005)

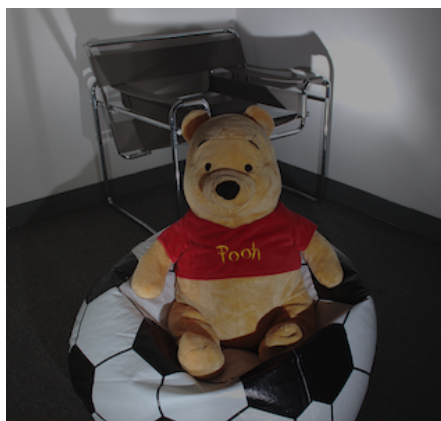


(f) Chen et al. (2012)

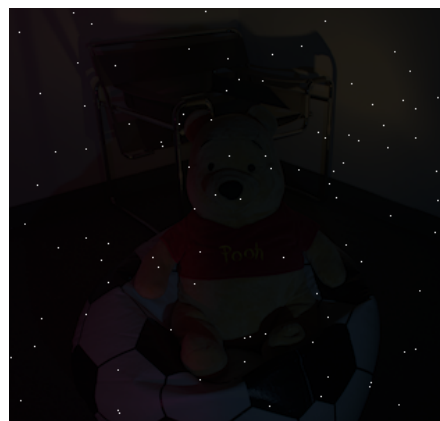
**Figure 13:** *Test 1. Pooh scene with a mixture of **two** lights.*



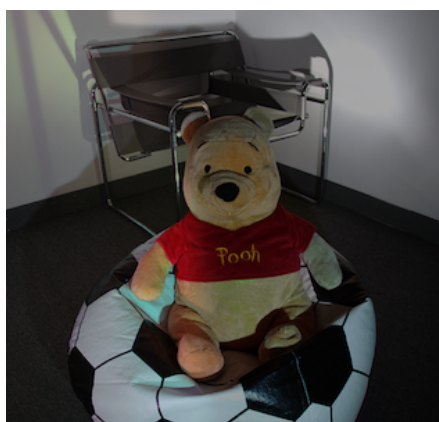
(a) Input (original)



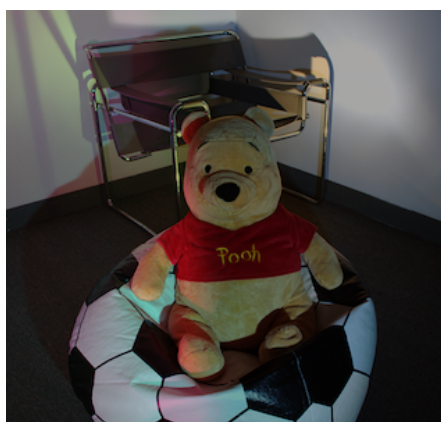
(b) Ground truth



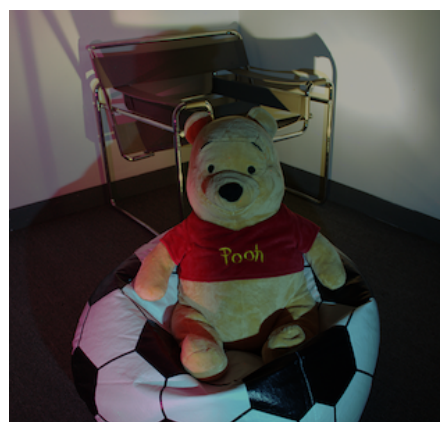
(c) Constraints



(d) Our result



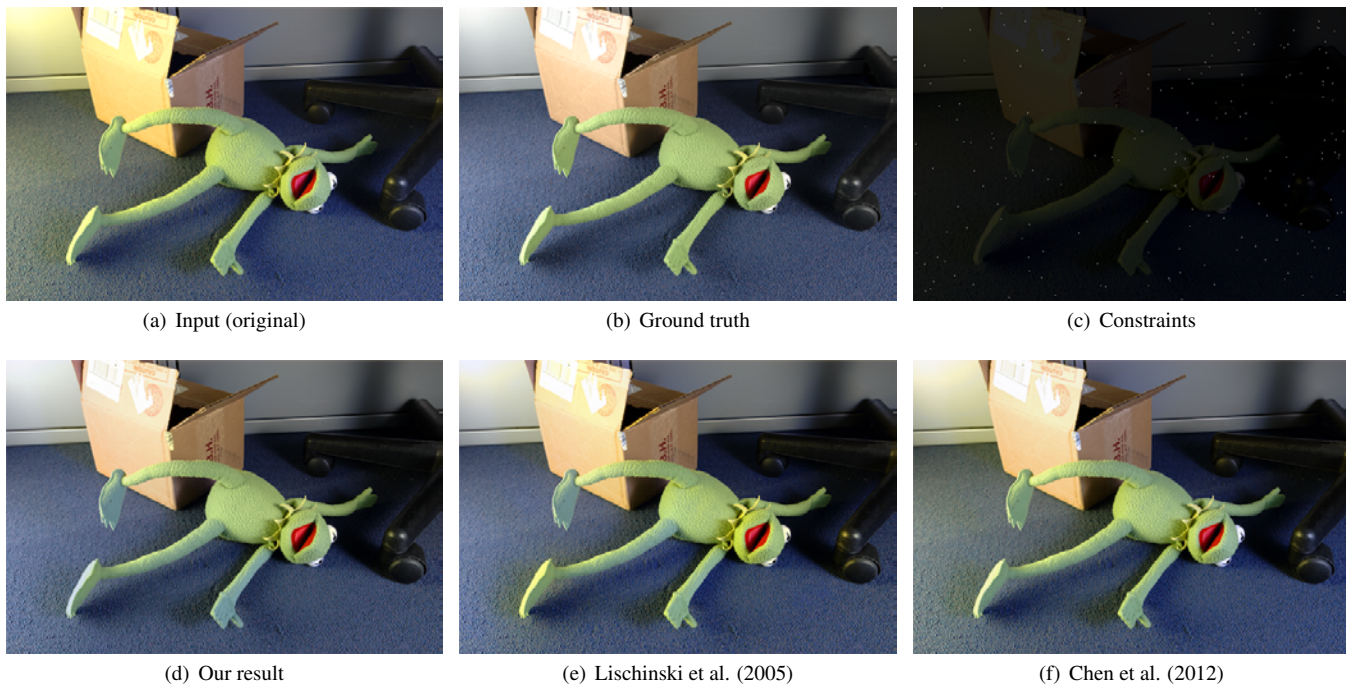
(e) Lischinski et al. (2005)



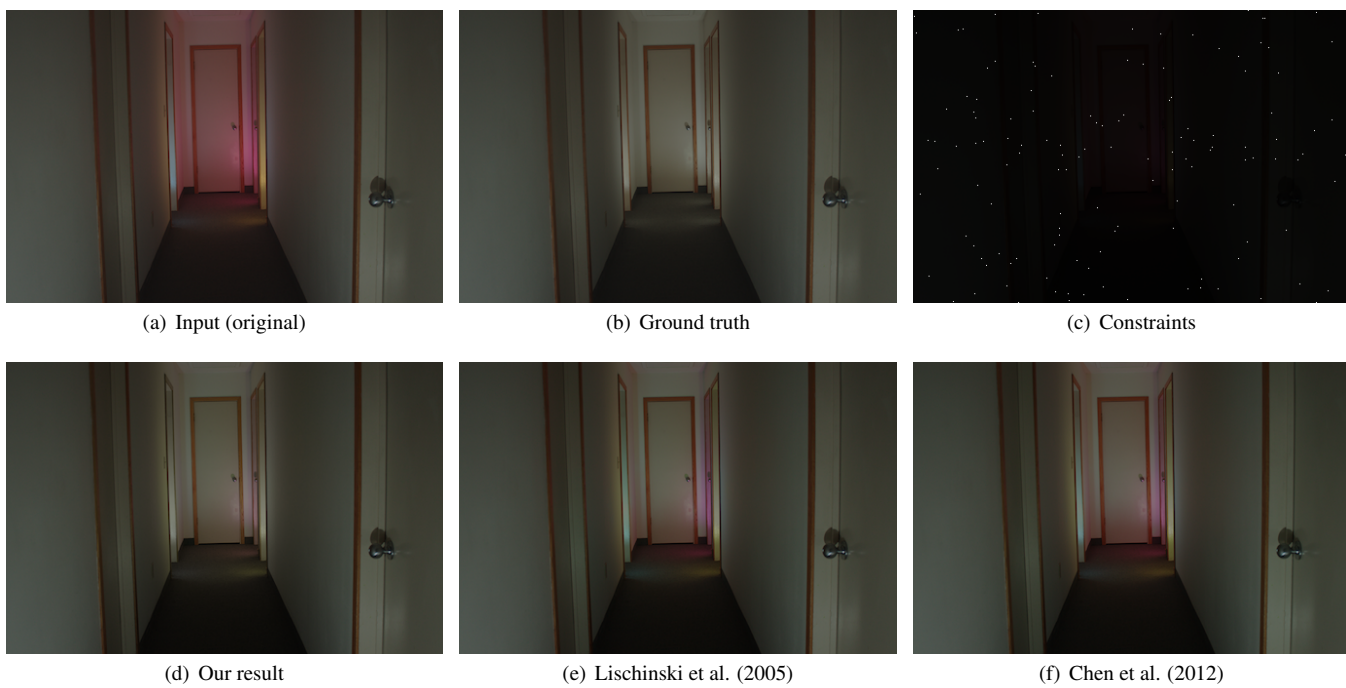
(f) Chen et al. (2012)

**Figure 14:** *Test 2. Pooh scene with a mixture of **three** lights.*





**Figure 15:** *Test 3. Kermit scene, mixture of two lights.*



**Figure 16:** *Test 5. Corridor scene, mixture of three lights.*





(a) Input (original)



(b) Ground truth



(c) Constraints



(d) Our result



(e) Lischinski et al. (2005)



(f) Chen et al. (2012)

**Figure 17:** *Test 7. Plant scene by Debevec's Light Stage data set. We simulate mixture of **three** lights, by composing three images with different color filters.*



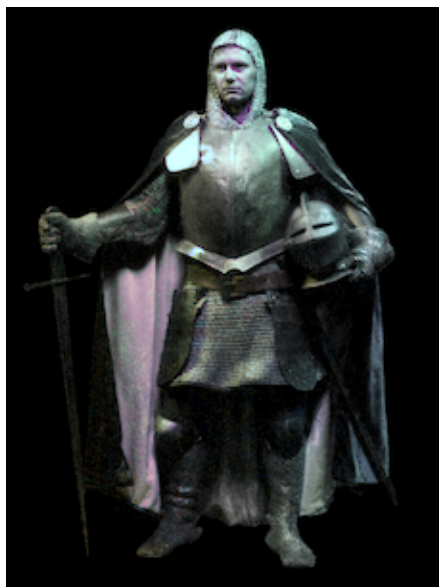
(a) Input (original)



(b) Ground truth



(c) Constraints



(d) Our result



(e) Lischinski et al. (2005)

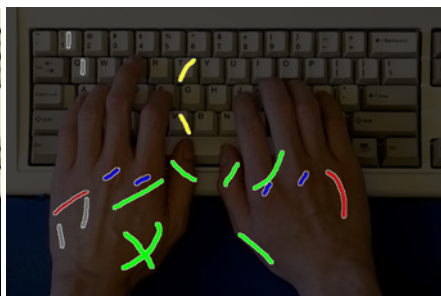


(f) Chen et al. (2012)

**Figure 18:** Test 10. Standing Knight scene by Debevec’s Light Stage data set. We simulate mixture of *two* lights, by composing two images with different color filters.



(a) Input



(b) User scribbles



(c) Estimated correction factors



(d) Our result

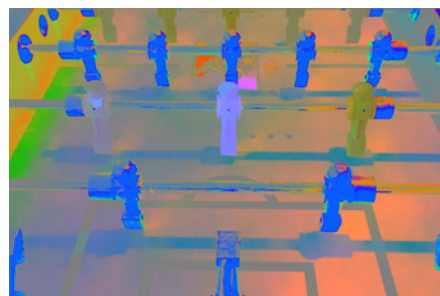
**Figure 19: Keyboard**



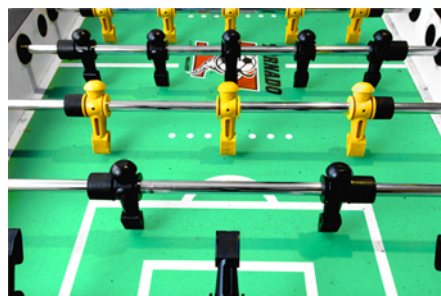
(a) Input



(b) User scribbles



(c) Estimated correction factors



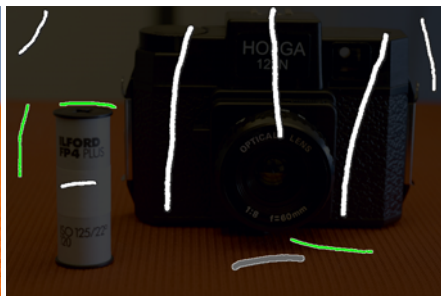
(d) Our result

**Figure 20: Foosball**

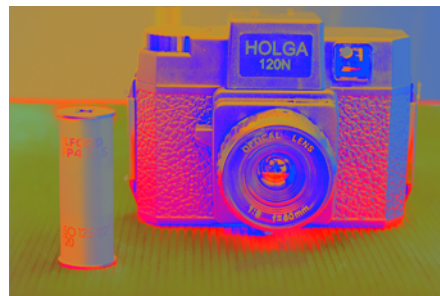




(a) Input



(b) User scribbles



(c) Estimated correction factors

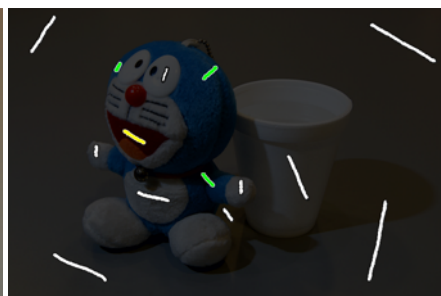


(d) Our result

**Figure 21: Camera**



(a) Input



(b) User scribbles



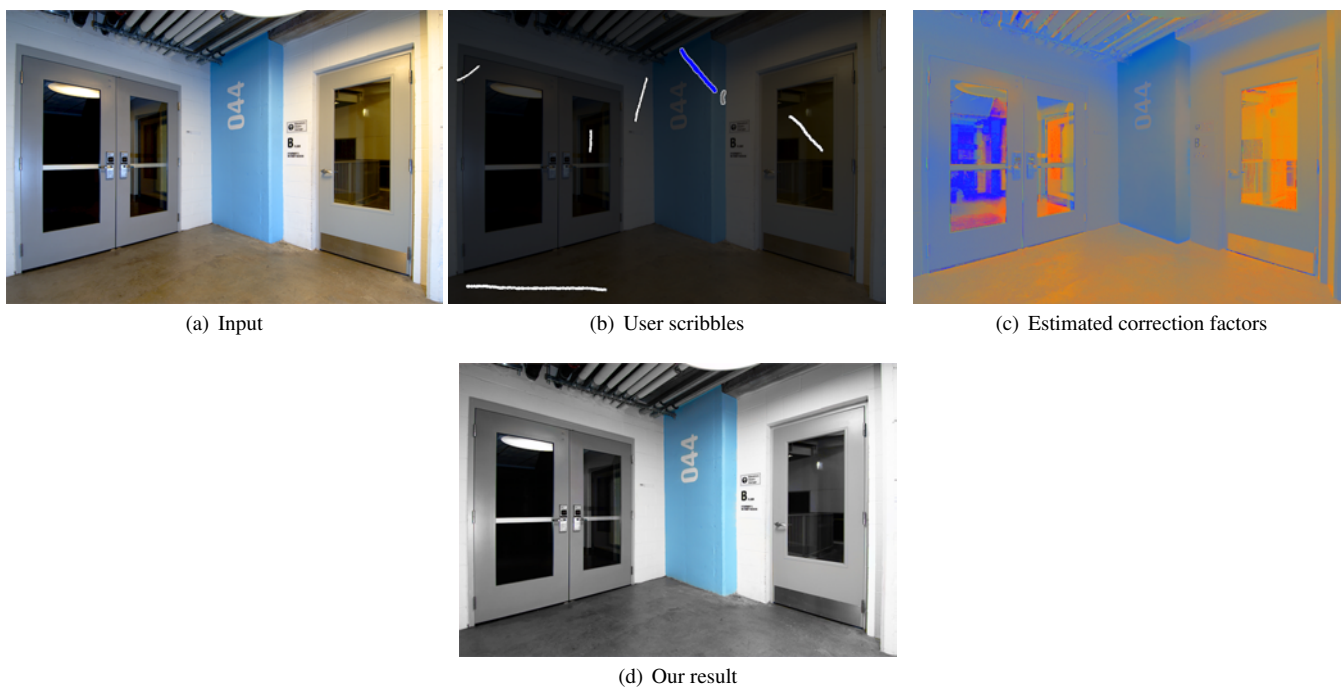
(c) Estimated correction factors



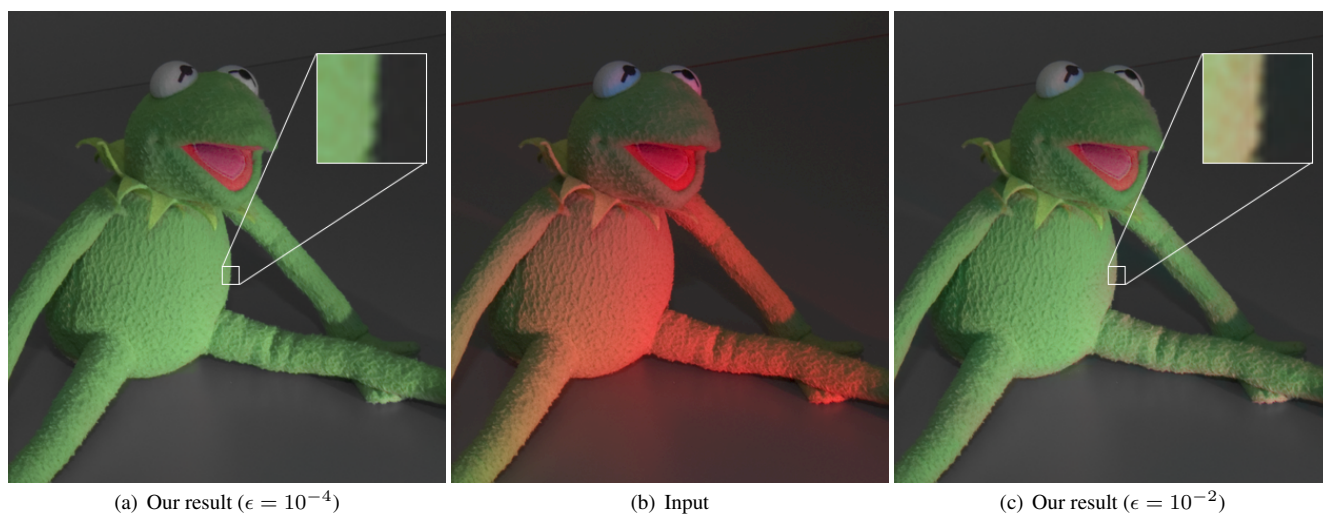
(d) Our result

**Figure 22: Blue creature**





**Figure 23: Hall**



**Figure 24: Kermit (effect of the regularization term):** smaller values of the regularization term (in the Matting Laplacian construction) give more importance to the model, rather than the smoothness of the solution. This allows our interpolation scheme to behave well around high-frequency details, such as edges.