

Greyscale Photograph Geometry Informed by Dodging and Burning

Carlos Phillips and Kaleem Siddiqi

School of Computer Science and Centre for Intelligent Machines, McGill University
{carlos, siddiqi}@cim.mcgill.ca

Abstract. Photographs are often used as input to image processing and computer vision tasks. Prints from the same negative may vary in intensity values due, in part, to the liberal use of dodging and burning in photography. Measurements which are invariant to these transformations can be used to extract information from photographs which is not sensitive to certain alterations in the development process. These measurements are explored through the construction of a differential geometry which is itself invariant to linear dodging and burning.

1 Introduction

Photographs are often used as test data in the computer vision literature. Properties of these photographs ranging from “Gaussian curvature” to edges, to statistical characterizations of the intensity values are commonly used to extract interpretations. In some cases, the method of acquisition of these photographs plays a central role. This is certainly the case in some of the work on the three-dimensional reconstruction of photographed scenes which is based on projective geometry (e.g. [1]). In other work, the mechanism by which the photograph is captured plays less of a role. It is sometimes overshadowed by the identification of features which are thought to be relevant to biological visual systems (e.g. [2]). Whether the camera optics are modeled or not, most approaches to computer vision ignore the variability of the prints resulting from identical scenes and optics. This observation also holds for digital photographs since their creation process closely mirrors that of photographs printed from film.

In this article we examine the process by which a black and white photograph is developed. This is followed by the construction of a geometry which is invariant to dodging and burning. We use this geometry to develop a set of invariant measures and illustrate their use in an example segmentation task.

1.1 Film

The active part of black and white film is an emulsion. It contains a uniform distribution of silver halide crystals suspended in a gelatin. Upon exposure to sufficient light, a small part of the crystal becomes reduced to metallic silver. The choice of a particular silver halide, such as silver bromide or silver iodide, together with the choice of gelatin determines the film’s response to light. The

film development process reduces crystals which have been sufficiently exposed to light entirely into metallic silver which makes the negative visibly dark in areas. The silver halide crystals do not store an intensity value. Their state is almost binary. Thus grey tones are represented by different densities of metallic silver simply referred to as the density on the film.

During the development process, the film is placed in a developer liquid. The developer reacts with the crystals which were hit by light. Crystals which have absorbed less light tend to react slower thus the amount of time that the film spends in the developer influences how dark the negative will be. This will in turn influence how bright the print will be. The term exposure is used to represent how affected the film is by the oncoming light. The exposure of a region on the negative is given by the product the average light intensity hitting that region and the amount of time during which the light has hit it, also known as the exposure time:

$$E = T \times I. \quad (1)$$

1.2 Printing

A photograph is printed by projecting a light through the negative onto photo paper and then developing the paper. The process by which photo paper captures information about the light hitting it is quite similar to that for the film. Here, as with the film, there are choices to be made about the chemical makeup of the light capturing membrane which will have profound effects on the final product. One of the key differences between photo paper and film is the number of stops. An average sheet of photo paper has no more than five stops of sensitivity to light intensity compared to the 15 stops which some negatives can represent. Thus film can record a range of intensities which is around 2^{15} greater than photo paper. There is therefore a decision to be made about what intensity range will be expressed on the photo paper. This decision will affect the contrast of the produced print as well as what information will be clearly represented in the printing.

Denser regions of the negative will absorb more of the light projected onto the negative than areas where there are few metallic silver spots. However, the spots are never so tightly packed that no light will pass at all. Every part of the photo paper will therefore be hit by light. Thus increasing exposure time will increase the exposure of the entire photo paper proportionally to the density of the relevant region on the negative. This process does not simply change the contrast of the photograph, it alters what information will be present in the final print. Since denser regions of the negative allow less light through, it takes longer for a sufficient amount of light to cross those regions to represent photograph details than it does for bright regions. Thus, the simplest way to display this information is to increase the exposure time of the photo paper. However, this in turn can saturate the dark regions of the print where light flows freely through the negative. This demonstrates that the way in which various densities on the print relate to each other and to the negative is not fixed. Figure 1(a) illustrates

this. Increasing the exposure time of the negative on the photo paper by a factor of k is equivalent to shifting the chemical reduction function to the left by $\ln(k)$. This follows directly from Equation 1.

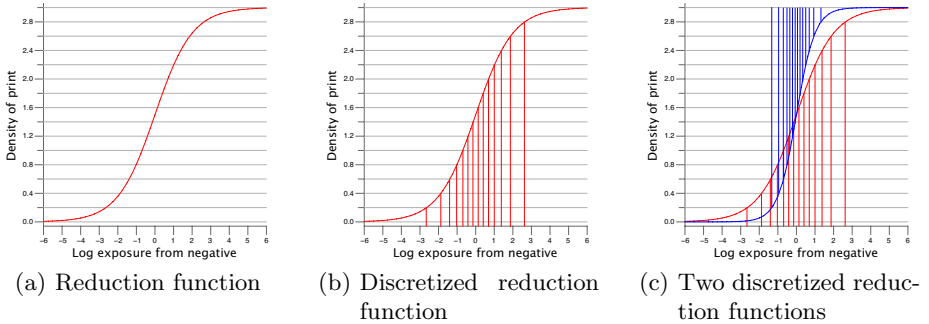


Fig. 1. Figure (a) shows an example chemical reduction function which specifies how exposure through the negative affects the density on the photo paper. The numbers in the graph only serve to show that the grid is regularly spaced and the standard direction of increasing values. Figure (b) adds bins representing a regular discretization of the photo paper density. The vertical lines represent the borders of the bins. Notice that the induced discretization of log exposures is not necessarily regular and depends on the reduction function. The example reduction function illustrates how exposure values are better captured in the mid-exposure range. Figure (c) demonstrates how a different choice of reduction function can change how exposure information is encoded.

1.3 Dodging and Burning

The previous section discussed the process by which a print is made directly from a negative. Dodging and burning are two commonly used techniques used during the printing of photographs which alter the exposure time of specific regions of the final print. Dodging involves completely blocking light from contacting certain regions of the photo paper for a portion of the exposure time. Burning involves doing a complete exposure followed by another exposure during which the light is restricted to contacting only a particular region of the photo paper. The process is often carried out using pieces of cardboard to block out the light. For dodging, the cardboard is affixed to a thin rigid rod so that it may be held over any part of the photograph.

For esthetic reasons, the dodging and burning tools must be kept in motion during their use. This allows the exposure times to vary “smoothly” between the regions being worked on and the rest of the photograph. Without the motion, the silhouette of the tool being used would become visible on the print. As indicated in [3] and in [4], photographers feel that they can apply these techniques liberally on photographs, especially when significant motion is used. Adams claims that most photographs can benefit from some dodging or burning [3]. Further, it is clear that he does not present these alterations as tools to change the subject,

but rather as tools to accentuate particular properties of the subject. Thus in processing photographs, we must be aware of the possibility that such transformations may have been applied and that they may be hard to detect or reverse. Figure 2 shows an extreme case of burning.

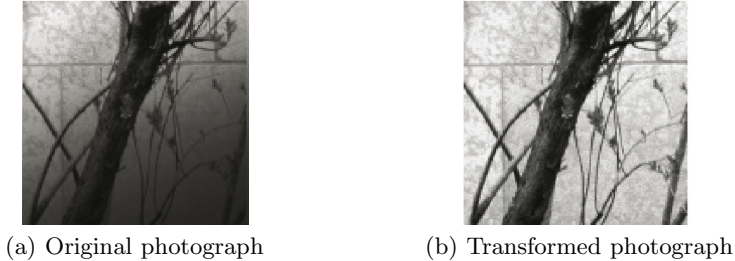


Fig. 2. Figure (a) shows a portion of the original photograph. Figure (b) shows the same photograph once burning has been applied. The burning is simulated using Adobe Photoshop™.

Dodging and burning selectively alter the exposure times of regions of the print. Their effect can therefore be explained using the standard notions of photographic development explored in §1.1 and §1.2. In particular, recall that changing exposure time is equivalent to shifting the chemical reduction function along the log exposure axis. This implies that dodging and burning can be seen as shifting the reduction function locally as long as every part of the print has a non-zero exposure time. In fact, assuming that when in motion the dodging or burning tool moves linearly and that the tool is close enough to the print so that the penumbra of the tool is not significant, this observation completely determines the impact of the techniques on the print. For example, a region which was hidden from exposure for exactly half the development time would have the same final development as if the chemical reduction function was shifted by $-\ln(2)$ for that region and there had been no dodging at all. Further, since dodging restricts the exposure time and burning increases it, the only difference between them is that burning shifts the reduction function to the left while dodging shifts it to the right.

2 Greyscale Photograph Geometry

As Florack argues in [5], we should not impose a topology or geometry directly on the space of photographs. Rather than impose properties on physical entities, we should instead impose them on the space of measurements and on measurement results. The measurements which we are interested in can be parameterized by their location, log-width and orientation. The log-width of a measurement distinguishes between the various scalings of a measurement while the orientation parameter distinguishes between its possible rotations. The set

of measurements and results for a given photograph is therefore a subset of the space $\mathcal{I} = \mathbb{M} \times \mathcal{U}$ which we call the image space and where $\mathbb{M} = \mathbb{R}^2 \times \mathbb{R} \times S^1$ is the measurement space and S^1 is the one dimensional sphere. The space \mathcal{U} contains all possible measurement results. A point in image space corresponds to a value being returned by a measurement at a given location, log-width and orientation. Throughout the rest of this article, we will refer to the set of image space points corresponding to a particular photograph as its *image*, or simply an image if the photograph is not specified.

Since \mathbb{M} is a four-dimensional space and the measurement values are a function of points in \mathbb{M} , we can assume that an image is a four-dimensional differential manifold in \mathcal{I} . Now that images have a manifold structure, we can define a geometry for these images based on the relationships between the components of their ambient space. We will begin by establishing a geometry on the measurement space \mathbb{M} and then proceed to extend it to the full image space \mathcal{I} .

2.1 Geometry of the Measurement Space

In the Riemannian geometry of a Euclidean space a (perhaps arbitrary) unit length is chosen. The notions of length and curvature are dependent on the chosen unit length in this situation. Our situation differs in that each measurement views its width as being of unit length. Thus if the locations of two measurements remain the same but their widths double, then their perceived distance will halve.

The measurement space with a fixed width, \mathbb{M}^w , looks like Euclidean space with the standard circle principal bundle. This has the effect of rotating the orientation values along with the location plane. Given the observations from this and the previous paragraph, we can define a Riemannian metric on \mathbb{M}^w as $\langle \cdot, \cdot \rangle_{(x,\theta)} = g_1 dx^1 \otimes dx^1 + g_2 dx^2 \otimes dx^2$. Since distances are inversely proportional to width as expressed above, the g_i coefficients must be $1/w^2$. Thus we get the following metric for \mathbb{M}^w

$$\langle \cdot, \cdot \rangle_{(x,\theta)} = \frac{dx^1 \otimes dx^1 + dx^2 \otimes dx^2}{w^2}. \quad (2)$$

The orientation circle and location parameters are closely related since the orientation circle represents the direction in which the measurement is taken. So any curve in measurement space must have its orientations and location tangents aligned. This can be achieved by creating a sub-Riemannian geometry defined by a distribution on \mathbb{M}^w using the following cotangent equation and then restricting curves to travel along this distribution:

$$\Omega = dx^2 - \tan \theta dx^1 = 0. \quad (3)$$

So far the proposed geometry is invariant to scaling of the visual signal since scaling the signal also scales measurements proportionally. However, in the full measurement space, width is allowed to vary. Since the width of a measurement

can be seen as a length from another measurement’s point of view, the measurement space metric must include a term to take into account variation in width which itself must be inversely proportional to current width:

$$\langle \cdot, \cdot \rangle_{(x,\theta)} = \frac{dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dw \otimes dw}{w^2}. \quad (4)$$

This makes the measurement space fully invariant to scaling as well as to rotation and translation. The distribution defined for \mathbb{M}^w can be used unchanged in \mathbb{M} to represent the relationship between location and orientation.

The constructed metric is well known as the metric for the upper-half plane model of three-dimensional hyperbolic geometry [6].

2.2 Geometry of the Image Space

Much as was the case for orientation, we cannot introduce measurement values into the metric. To do so would be to establish intrinsic location related properties in measurement values. The measurements we choose will be based on brightness information. We will therefore start by defining the measurements to collect average brightness over a unit disc. Let’s call these measurements $b_{x,\theta,w}$. We will examine how these measurements interact with a set of simple photographic transformations so as to derive better values to use when defining base measurements.

The image space \mathcal{I}_b for this measurement can be expressed by adding a brightness line $\mathcal{B} = \mathbb{R}$ to the measurement space: $\mathcal{I}_b = \mathbb{M} \times \mathcal{B}$. The argument for excluding measurement values from the metric applies to brightness. However, excluding brightness from the metric does not exclude it from the geometrical description. On the contrary, it defines \mathcal{I} geometrically up to acceptable transformations of the brightness line. We choose to include the linear dodging transformation, assuming linear reduction functions, introduced earlier along with varying exposure time in the set of transformations under which image content is invariant,

$$u \leftarrow u + a_1x + a_2y + t_u, \quad (5)$$

where a_1 , a_2 and t_u are arbitrary constants, the last one representing a translation of v . In this equation, u is a coordinate for the brightness line. We have chosen this transformation since it is a basic image correction transformation. As desired, a difference of brightness values at two distinct measurement points is meaningless (i.e. not invariant to the transformation group).

Our analysis bears some resemblance to that of Koenderink and van Doorn’s [4]. Our motivation is quite different. Most significantly, since measurements were not explicitly considered in [4] the notions of measurement width and orientation were not developed. In our work these ideas turn out to be key for the development of a methodology for hierarchical decomposition (moving from coarse measurement sizes to finer ones), which is the focus of §3. We note that in the context of metrics for greyscale photographs there is also related work by Eberly and his colleagues [7].

2.3 Invariant Measurements

It is clear from the previous section that average brightness is not an invariant measurement to the photographic transformations which we have discussed. There are many possible invariant measurements. In this section we focus on perhaps the simplest, a measurement type which is based on best fit planes.

For a given measurement point $v = (x_1, x_2, \theta, w)$, the associated measurement value for our example is a best fit plane p_v for the brightness of the photograph over a disk d described by location (x_1, x_2) and radius w . The fit minimizes the average integral of $(B(x, y) - p(x, y))^2$ over d , where B represents the brightness values while p represents the plane. The plane is recorded using three parameters. The first two are the slopes of the plane along orthonormal axes u_t and u_n where u_t is in line with θ and u_t forms a right-handed frame with u_n . The last parameter is simply the brightness value of the plane over point (x, y) . Thus the space of measurement values \mathcal{U} is composed of these planes p and for convenience we define functions $m_t, m_n, b : \mathcal{I} \rightarrow \mathbb{R}$ which return the u_t -slope, u_n -slope and elevation at a given image point respectively.

The plane parameters described above are clearly invariant to rotation and translation of the Euclidean plane. However, they are not invariant to transformations such as linear dodging or scaling. When a linear dodging is applied to a visual signal, it leaves the values collected by dm_t and dm_n along a curve invariant. The changes in plane elevations given by db , however, are not invariant to this subgroup of transformations as we have seen when examining average brightness measures. There we saw that though differences in brightness values were preserved along a brightness line, they are not preserved when taken across brightness lines. We can use this to define an invariant value given below, which is one way of representing how much the plane is increasing in height as you move along a curve:

$$\Omega_b = du - m\sqrt{dx_1^2 + dx_2^2}.$$

When a photograph is scaled by a scale factor s , all the slopes at corresponding points are scaled by $\frac{1}{s}$ and all the corresponding measurement widths are scaled by s . Therefore

$$\Omega_t = w \cdot dm_t \quad \text{and} \quad \Omega_n = w \cdot dm_n$$

are invariant under scaling of curves in \mathcal{I} . The values of $du - m\sqrt{dx_1^2 + dx_2^2}$ are also invariant for the same reasons. Thus we get three invariant values which represent the changes in slope and elevation of the planes along a given curve. In the next section, we use these measurements and relations to develop an algorithm for region grouping.

This concludes our construction of the geometric model and of our basic set of invariant measurements. This model can be used to create other measurement classes which are invariant to the photographic transformations covered in this paper. Further, expressions based on the values which are computed using the presented measurements can themselves be used as input into standard image processing algorithms. In the next section, we examine the properties which are represented in the model through an example task of segmentation.

3 Example Task

So as to illustrate aspects of the image space geometry, we define a naïve segmentation process based on the invariant measurements presented in §2.3. Algorithms 1 and 2 describe the two steps of the process in pseudo-code. First the invariant measurements are estimated so as to produce an image for the input photograph. The image manifold is collapsed into two-dimensional surfaces. Then the grouping algorithm segments the image points into separate surfaces based on local affinity. Note that the widths referred to in the algorithms are not scales as in [8] but simply widths of measurement discs. The presented algorithms are provided for the sole purpose of examining the structures created using the invariant measurements.

Figure 3(c) illustrates the level of detail collected at different measurement widths. Each layer image shows the averaging of approximation discs which have been kept after running the measurement algorithm. These layers can be

Algorithm 1. Measurement Algorithm

```

for  $w = max\_width$  down to 1 do
  for  $q$  a pixel in the photograph do
     $d \leftarrow$  the disk of radius  $w$  centered at  $q$ 
     $p_d \leftarrow$  plane fit to intensity over  $d$ 
     $f_{(q,w)} \leftarrow$  the mean squared error of the fit
  end for
end for
for  $w = max\_width$  down to 1 do
  for  $q$  a pixel in the photograph do
    if  $\frac{f_{(q,w)}}{w} > \epsilon$ 
    or  $\frac{f_{(q,w)}}{w} > \frac{f_{(q,w+1)}}{w}$ 
    or  $\frac{f_{(q,w)}}{w} > \frac{f_{(q,w-1)}}{w}$  then
      delete  $p_d$ 
    end if
  end for
end for
for  $w = 1$  to  $max\_width$  do
  for  $q$  a pixel in the photograph do
    for  $v \in \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  where the components represent directions
       $x_1, x_2$  and  $w$  resp. do
         $\Delta_b(q, w) \leftarrow \Omega'_b(v)$ 
         $\Delta_t(q, w) \leftarrow \Omega'_t(v)$ 
         $\Delta_n(q, w) \leftarrow \Omega'_n(v)$ 
      end for
    end for
  end for

```

This defines a discretized three-dimensional manifold in \mathcal{I} which depends on parameter ϵ . For our experiments, we set $\epsilon = 0.1$. We use symbols Δ' instead of Δ . to indicate that difference equations are used to approximate the infinitesimal measurements.

Algorithm 2. Grouping Algorithm

```

for  $w = 1$  to  $max\_radius$  do
  for  $q$  a pixel in the photograph do
    for  $v \in \{\gamma : \gamma \in \{-1, 0, 1\}^3\}$  do
      for  $i \in \{1, 2, 3\}$  do
        if  $v_i = -1$  then
           $q_i \leftarrow q_i - 1$ 
        end if
      end for
       $v = (|v_1|, |v_2|, |v_3|)$ 
      At location  $q$  with width  $w$  in direction  $v$ :
      if  $\max(\Delta_b, k\Delta_t, k\Delta_n) < \delta w$  then
         $(q_1, q_2, d)$  is adjacent to  $(q_1, q_2, d) + \gamma$ 
      end if
    end for
  end for
end for

```

This produces a grouping across measurement points which is dependent on parameters k and δ given the output from the measurement algorithm. For our experiments, we set $k = 1$ and $\delta = 0.5$.

recombined by simply drawing one over the other, from greatest width to smallest. Pixels for which there is no measurement in a given layer are left unaltered when the layer is drawn. This is shown in Figure 3(d). The important thing to note is that structures in the photographs appear at reasonable widths. For example the trunk of the tree appears around widths of 10 pixels but is not present around widths of 20 pixels. Figure 4 shows that in fact the layers presented in Figure 3 form two-dimensional surfaces. The results of applying the naïve grouping algorithm, shown in Figure 5(b), show that structural information about the photograph can easily be extracted from the image structure. Figures 5(a) and 5(b) confirm that the measurements are indeed unchanged by the application of a Adobe PhotoshopTM burning effect. Finally, Shi and Malik's graph-cut based algorithm [9] is applied on the original photograph and on the transformed photograph to show that this popular approach to segmentation is sensitive to dodging and burning transformations when intensity value measurements are used as inputs. This sensitivity can be problematic if the information from the original negative is being sought rather than information about the particular print being examined.

4 Conclusion

We have introduced a geometry which is invariant to certain forms of burning and dodging. We then used this geometry to create invariant measurements which represent information which would not change given a different development process.

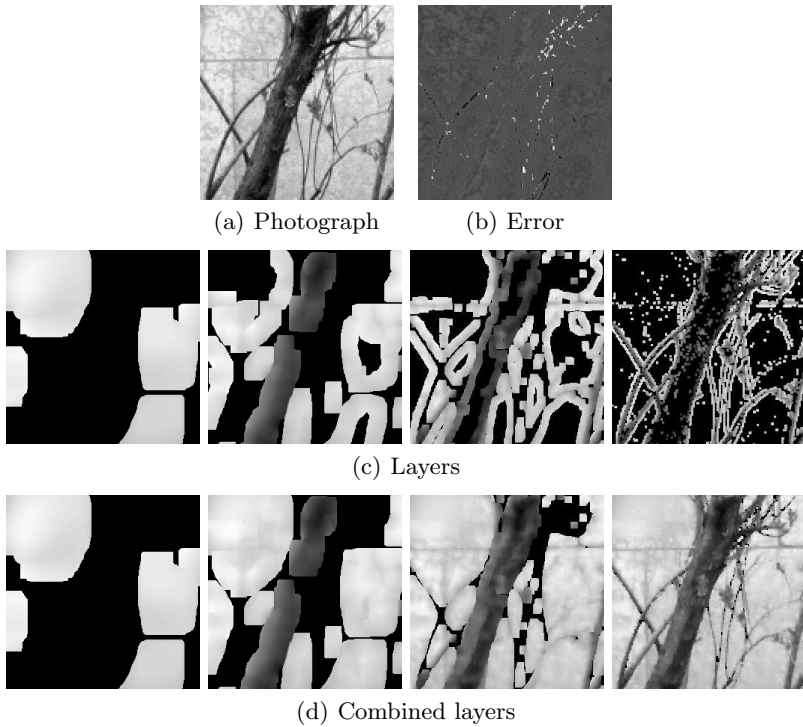


Fig. 3. (a): The original photograph. (b): The difference between the original photograph and the photograph reconstructed from the layers. (c): From left to right, the measurement size layers with widths 20, 10, 4 and 1 pixels, respectively. (d): From left to right, the reconstructed photographs for measurement size layers with widths down to 20, 10, 4 and 1 pixels, respectively.

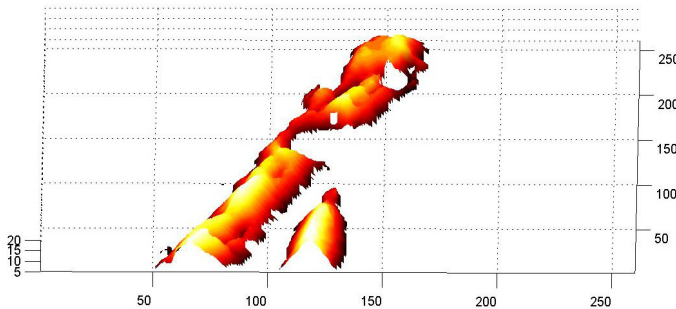


Fig. 4. Two of the surfaces obtained by the grouping algorithm, corresponding to the trunk of the tree and the region to its lower right. These surfaces are manifolds in 3D, which could in fact overlap when viewed from any fixed viewing direction.

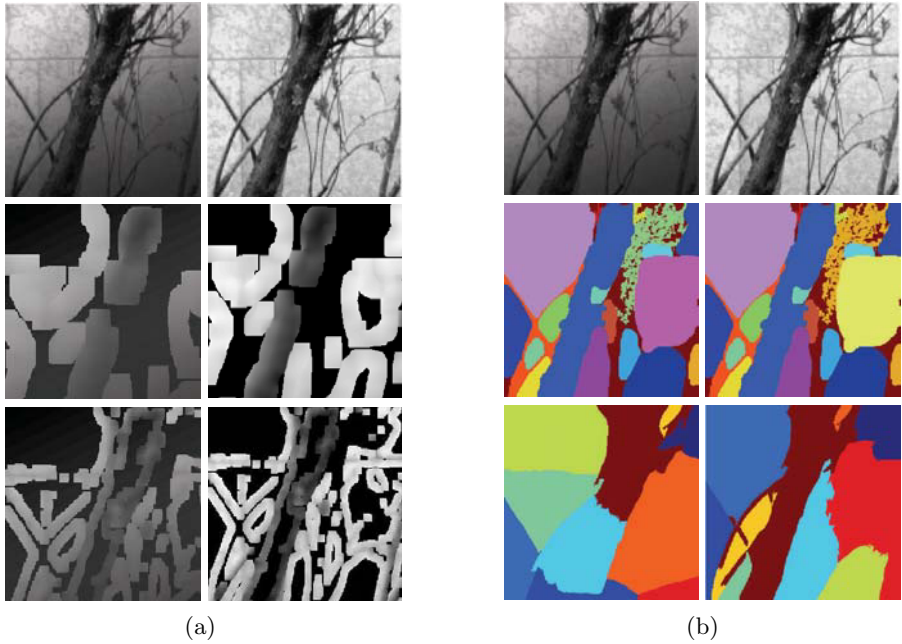


Fig. 5. These illustrations demonstrate that the presented approach is in fact invariant to the defined photographic transformations. (a): Left Column: from top to bottom, the original photograph and its layers with widths 10 and 4 pixels. Right Column: from top to bottom, the original photograph with a linear dodging transformation applied to it and its layers with widths 10 and 4 pixels. In fact the average grading is not represented in the image structure and both photographs are identically represented, as invariance to linear dodging requires. (b): Left Column: from top to bottom, the original photograph, our grouping and the grouping from [9]. Right Column: from top to bottom, the original photograph with a linear dodging transformation applied, our grouping of the transformed photograph and the grouping generated by [9]. The source code from <http://www.cis.upenn.edu/~jshi> with default parameters was used to compute the results of the algorithm presented in [9].

The presented algorithm used properties of best-fit planes to represent a photograph. There are many more possibilities for measurements which would fit this framework. Further, the photograph representation presented in this paper could be combined with existing vision algorithms for such tasks as object recognition within photographs for which we do not know the development process.

Acknowledgments

We would like to thank Jiambo Shi for making his source code available.

References

1. Faugeras, O.: Three-dimensional computer vision: a geometric viewpoint. MIT Press, Cambridge, MA, USA (1993)
2. Marr, D.: Vision: a computational investigation into the human representation and processing of visual information. W. H. Freeman, San Francisco (1982)
3. Adams, A.: The Print. Volume Book 3 of The New Ansel Adams Photography Series. Little, Brown and Company, Boston (1983)
4. Koenderink, J.J., van Doorn, A.J.: Image processing done right. European Conference On Computer Vision, Lecture Notes in Computer Science **2350** (2002) 158–172
5. Florack, L.: Image Structure. Kluwer Dorecht (1997)
6. Guggenheimer, H.W.: Differential Geometry. Dover Publications (1977)
7. Eberly, D.: Ridges In Image and Data Analysis. Kluwer Academic Publishers (1996)
8. Lindeberg, T.: Effective scale: A natural unit for measuring scale-space lifetime. IEEE Transactions On Pattern Analysis and Machine Intelligence **15** (1993) 1068–1074
9. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 888–905