Depth-Stretch: Enhancing Depth Perception Without Depth

Hagit Hel-Or
University of Haifa
Haifa, Israel
hagit@cs.haifa.ac.il

Yacov Hel-Or
The Interdisciplinary Center
Hertzeliya, Israel
toky@idc.ac.il

Renato Keshet
GE Global Research Center
Tirat Carmel, Israel
renatokeshet@gmail.com

Abstract

A simple and efficient method is presented to enhance the depth perception of an image. The approach termed Depth-Stretch (D-stretch) is a tone mapping operation that is applied to the shading component of the given image. Although re-rendering a scene under geometric transformations typically requires extracting the 3D model of the scene, we show that under very simple assumptions D-stretch can be implemented without 3D reconstruction, while still providing a convincing effect of depth enhancement.

1. Introduction

Improving depth impression in still images is a desired goal of many photographers and artists. For this purpose, a photographer may emphasize the scene perspective by composition tricks, she may carefully design the illumination setup to increase cast shadows, or amplify de-focus blur during acquisition or using post-production operations. Had the photographer the ability to actually stretch the 3D scene and increase its depth along the camera viewing direction, this would have strengthened the shading cues of the resulting photograph and thus increased its depth impression. However, such a manipulation requires 3D modeling of the scene, a demanding and still unresolved problem using a single RGB image.

In this work we present a novel method of post production depth enhancement termed Depth-Stretch or *D-stretch*. D-stretch is performed by manipulating the shading cues exaggerating the shading effects. We show that under very simple assumptions on the lighting setup and the the scene, D-stretch boils down to a simple tone mapping operation that is applied to the shading component of the given image. The proposed method is embarrassedly simple but efficient and effective, easily implemented using a single lookup table. The bottleneck of the approach is the necessity of computing the shading component of the image (or equivalently - the intrinsic image) which is known to be a complicated task [2]. Nevertheless, we show visually that the intrinsic image, and thus the shading component, need not be computed at a high quality level. In fact, crude shading computations with inferior quality are sufficient to produce very convincing D-stretch effects.

We further show the relation of the suggested approach with that of other known detail enhancing methods, including global or local tone mapping, unsharp masking and high-dynamic range (HDR) compression methods. Under specific conditions, these enhancing techniques can be explained under the D-stretch theory. However, in contrast to the other suggested methods, the proposed scheme is provably justified and derived directly from the image formation model. It does not rely on ad-hoc manipulations or common practice. In retrospect, the proposed approach may suggest an educated explanation why some of the ad-hoc manipulations, such as, edge-preserving details enhancement or high-dynamic range compression, emphasize in practice the depth perception.

2. Background and Previous Works

3D perception in a 2D image is captured by various visual cues [13]. Some cues are manifested through geometric properties such as linear perspective, texture gradient, or relation to objects in the scene whose real sizes are known. Other cues are based on atmospheric effects such as hazing where distant objects appear dull or de-saturated in color due to dust or particles in the atmosphere [24, 7]. Acquisition devices and their adjustment may also produce depth cues in the image, e.g. the loss of sharpness of objects distant from the plane of focus [18].

Another effective cue, which we exploit in this paper, is the shading effect [12]. Since the radiance reflected from any point in the scene depends, amongst others, on the object's geometry, the shading variations in the image provide a strong percept of depth and shape [12, 1]. In contrast to the above mentioned cues that provide qualitative depth perception, shading effects produce quantitative depth details which can be exploited for 3D reconstruction using shape from shading or photometric stereo techniques [12, 36, 35].

Given an already acquired image, one must resort to

post-production methods to enhance the depth percept. Various approaches have been suggested.In [8], shading and surface details are manipulated by exploiting a collection of images of a scene illuminated from different directions. The strongest shading effects in all images are integrated using edge-aware multi-scale decomposition. Similar input (although in a more controlled setup) is used in [22] to extract surface normals and manipulate surface photometric parameters (e.g. increasing the specular component). The goal of these methods is similar to that of this paper, however, they require a set of images obtained using a controlled setup with artificial illumination while this paper assumes a single input image acquired without controlled illumination. In [21], depth perception is emphasized by artificially introducing Gibbs artifacts (halos) near depth discontinuities.

Other approaches perform image enhancement from a single image by either requiring the depth or depth gradients at each pixel [27, 32], or estimating the depth or normal at each pixel of the input image [15, 20]. These methods directly rely on the depth information of the scene for enhancement. Finally, we mention studies in creation of basreliefs. These studies compress an original full depth structure into a restricted depth range, thus controlling the perceived depth of the final relief. However, these methods all assume as given, either 3D models, depth maps or z-buffer values on which height compression is performed [34, 29]. We emphasise that in our proposed approach, depth perception is enhanced with no depth, height or 3D information.

Other studies take a different approach and ignore the 3D structure of the image and apply detail enhancing operations irrespective of the 3D structure. Such approaches involve global or local contrast enhancement and high-frequency enhancement [6, 9, 4, 31]. In Section 6.3 we show that our proposed technique bears many similarities to these techniques but the suggested formulation is theoretically justified and is derived directly from the image formation model. Our approach performs precise mapping with no ad-hoc or qualitative operations.

In this paper we propose to enhance depth perception by manipulating the shading component of an image. Our contributions in relation to other detail enhancement approaches are:

- The manipulation we apply in this work is mathematically justified and derived directly from the image formation model. Thus, the approach suggested here is not an ad-hoc approach that looks good but a theoretically justified and precise method.

- The approach shows that, under reasonable assumptions on the image acquisition process, the exact depth-stretching reduces to a global tone-mapping with no need of depth information and relies only on the shading values. Although this is a known practice in some applications,this is the first paper that proves this point.

3. Depth Stretching without 3D Modeling

Assume an image is acquired of a general scene whose 3D model is represented in the viewer's coordinate system. That is, the Z axis coincides with the camera's optical axis and the X, Y axes are parallel to the x, y axes in the projection plane (image plane). Assuming the scene is illuminated by a point light source $\mathbf{l} = (l_x, l_y, l_z)$ where the direction of $\mathbf{l}$ indicates the light direction and the magnitude of $\mathbf{l}$ is the light intensity. The diffuse component reflected from a scene point, whose projection in the image plane is at (x, y)[1] is given by Lambert's law [12]:

$$I(x, y) = \rho(x, y)(\mathbf{l} \cdot \mathbf{n}(x, y)) \tag{1}$$

where $\rho(x, y)$ is the diffuse coefficient (albedo), $\mathbf{n}(x, y)$ represents the 3D normal of the physical point projected to (x, y) and $(\mathbf{l} \cdot \mathbf{n})$ is a scalar product between $\mathbf{l}$ and $\mathbf{n}$. Note, that the above equation depends on the light's spectral wavelength but due to the trichromatic model of the human visual system three independent equations, one for each color band of the image, suffices. For the sake of simplicity we express our scheme on a single band. Given the 3D surface model $Z(X, Y)$, the normal of a point at the surface whose projection on the image plane is at (x, y) is:

$$\mathbf{n}(x, y) = \frac{[p, q, -1]}{\sqrt{p^2 + q^2 + 1}} \tag{2}$$

where $p(x, y)$ and $q(x, y)$ are the surface derivatives: $p = \frac{dZ}{dX}$ and $q = \frac{dZ}{dY}$, and the denominator is a normalization factor to form a unit vector.

Assume for the moment that light direction is fronto-parallel and originates from the direction of the camera, thus $\mathbf{l} = k_p[0, 0, -1]$ where k_p is the light intensity. Substituting

[1] We assume the scene is distant from the camera and weak perspective can be used.

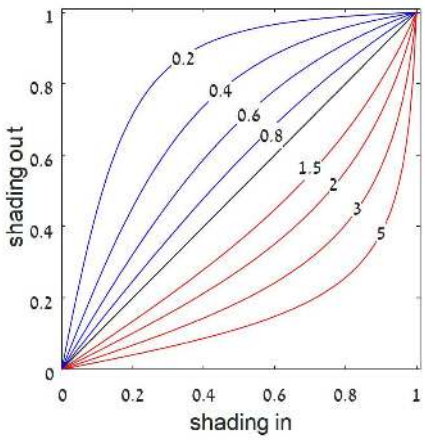
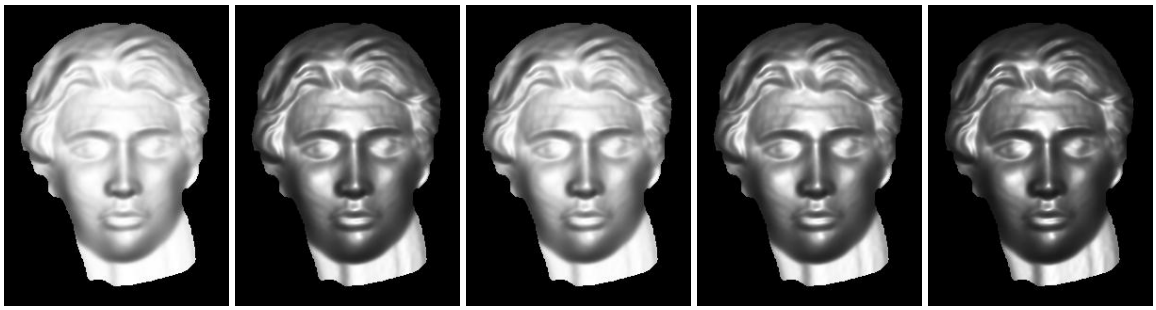


Figure 1. D-stretch tone curves for various α values.



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

Figure 2. Comparison of rendered 3D model vs D-stretch. a) Rendered Model b) Rendered 3D-scaled model c-e) D-stretch of (a) with $\alpha = 1.5, 2.5, 3.5$. Lighting position and camera are frontal-parallel.

$\mathbf{l}$ and $\mathbf{n}$ back into Equation 1, we obtain:

$$I(x,y) = \rho(x,y)k_p \frac{1}{\sqrt{p^2+q^2+1}} = R(x,y)S(x,y) \quad (3)$$

where $R(x,y) = \rho(x,y)k_p$ represents the *reflectance* component and $S(x,y) = \frac{1}{\sqrt{p^2+q^2+1}}$ is the *shading* component. Note, that the shading is in the range $[0,1]$ and depends only on the geometry of the scene.

Following the above notation, if the surface is stretched along the Z direction by a factor of α so that $Z'(X,Y) = \alpha Z(X,Y)$, the surface derivatives change accordingly: $p' = \frac{dZ'}{dX} = \alpha p$ and $q' = \frac{dZ'}{dY} = \alpha q$, and the shading component now reads:

$$S'(x,y;\alpha) = \frac{1}{\sqrt{\alpha^2 p^2 + \alpha^2 q^2 + 1}} \quad (4)$$

The reflectance component remains unchanged. Typically, in order to re-render the shading component under any geometrical transformation, the 3D structure of a scene must be reconstructed. Surprisingly, under the above assumptions re-rendering the scene under depth stretching can be applied easily using a mapping function on $S(x,y)$. Given that $S(x,y) = 1/\sqrt{p^2+q^2+1}$ (Equation 3), isolating p^2+q^2 we obtain:

$$p^2 + q^2 = 1/S^2(x,y) - 1$$

Substituting the above into Equation 4 we obtain:

$$S'(x,y;\alpha) = \frac{1}{\sqrt{\alpha^2\left(\frac{1}{S^2(x,y)} - 1\right) + 1}} \quad (5)$$

Equation 5, defines the D-stretch algorithm. If we are able to separate the shading component from the reflectance component, re-rendering the image with a stretched depth can be applied easily using a mapping function on $S(x,y)$ as given in Equation 5. Figure 1 shows the mapping function for various values of α. Note, that $\alpha > 1$ implies depth stretching while $\alpha < 1$ implies attenuation of depth or *D-shrinking*.

Figure 2 displays a comparison of the D-stretch result with true rendering of a 3D model. Figure 2a shows a rendering of a 3D model with fronto-parallel lighting and viewing direction ($\mathbf{l} = [0,0,-1]$). The 3D model was scaled by a factor of 2.5 along the z-axis and re-rendered as shown in Figure 2b. D-stretch was applied to the 2D *image* of Figure 2a in order to reproduce the effect of depth stretching in the model. Figures 2c-e show the results of applying D-stretch using stretch parameter $\alpha = 1.5, 2.5, 3.5$. It can be seen that using $\alpha = 2.5$, the effects of actual 3D stretching of the model is perfectly reproduced (MSE between Figures 2b and 2d is zero - see Figure 13 for 0°).

It should be emphasized that the above formulation was developed under the assumption that the light direction is $[0,0,-1]$. When light direction differs, the exact same formulation can be applied while representing the surface $Z(X,Y)$ in a coordinate system whose Z axis is parallel to the light direction. In this case, however, the depth is stretched towards the light direction and not towards the viewing direction. However, the perceived depth is still boosted as long as the light originates from the viewing hemisphere. Fortunately, this is the case in most images (otherwise cast and self shadows may cause interference in the image). See also Section 6.1 and Figure 13 for a discussion on change of lighting position.

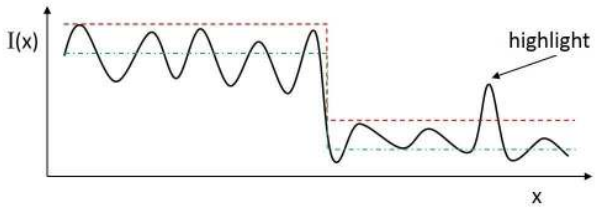


Figure 3. Extracting the reflectance component of an image. A 1d image is shown having a large reflectance change and smaller shading changes (solid black). The reflectance component can be viewed as the upper envelope bounding the image (dashed red). Applying Poisson reconstruction to the large derivatives results in a smooth profile of the image (double-dashed green). This profile is lifted to meet the upper envelope and produce the reflectance component of the image.

Figure 4. Image decomposition and D-stretch. a) Original. b) D-stretched with $\alpha = 1.5$ c) Reflectance component. d) Shading component.

4. Obtaining the Shading Image

To apply D-stretch on an image, the shading component must be extracted. Various approaches have been suggested to compute the shading image, typically associated and computed in combination with the intrinsic (reflectance) image [2] and sometimes with additional layers of image content such as specular map, diffuse map, shading, reflectance and illumination components. Reflectance and shading image decomposition have been proposed based on multiple images [33, 23, 16] or from a single image [30, 10, 28, 1].

In our testing, several techniques provided very good visual results, however, for the sake of consistency, in all the presented examples we have used shading extraction based on reflectance derivatives similar to [30]. The image is first converted into the log domain so that the reflectance and shading components are additive. Under the assumption that reflectance changes and shading changes rarely occur at the same location, Tappen et. al. suggested to classify each pixel derivative as originating either from changes in the reflectance or from changes in the shading. In our system, classification of pixel gradients is performed following the Retinex heuristics [17] which assumes that image derivatives with a large magnitude correspond to reflectance changes, while small derivatives are most likely caused by shading. Thus we define pixel gradients as reflectance gradients if they exceed a given threshold. Denote by $\mathcal{R}_x$ and $\mathcal{R}_y$ the images of the x and y derivatives of the pixels labelled as reflectance (and zeros elsewhere). The reflectance image $\mathcal{R}$ is then extracted by solving the normal Poisson equation:

$$\left[D_x^T, D_y^T\right] \begin{bmatrix} D_x \\ D_y \end{bmatrix} \mathcal{R} = \left[D_x^T, D_y^T\right] \begin{bmatrix} \mathcal{R}_x \\ \mathcal{R}_y \end{bmatrix}$$

where D_x and D_y are the x and y derivative operators defined in matrix form. The solution $\mathcal{R}$ is obtained using the conjugate gradient algorithm. The desired shading image is extracted by completion.

Consider the visualization in Figure 3. An image I(x,y) is visualized as a 1D function of luminance v.s. coordinate x. The image is depicted as having a single large reflectance change and smaller shading changes. Since the shading

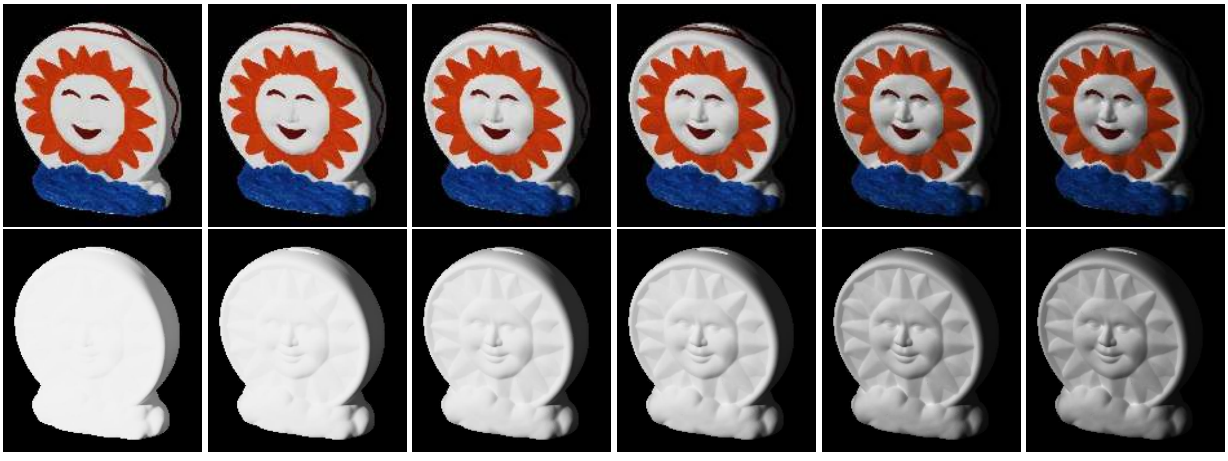


Figure 5. D-stretch of 'Sun' image from the MIT database [11]. Bottom: D-stretched shading image. Top: Resulting D-stretched image.

Figure 6. D-stretch applied to coin images. In each pair, original on left, D-stretched on right.

Figure 7. D-stretch applied to art reliefs images. In each pair, original on left, D-stretched on right.

Figure 8. D-stretch applied to map and textile images. In each pair, original on left, D-stretched on right.

component is bounded in the range [0..1], the reflectance component can be viewed as the upper envelope bounding the luminance image (dashed red line). However, the reflectance image computed as described above does not form an upper envelope (double-dashed green line). Thus values "lifting" is performed to the reflectance image so that its profile will bound the original image from above. In practice, care must be taken to avoid over-lifting due to image highlights, thus the lifting is performed so that 95% of the image values are below the lifted profile. Highlight pixels are removed and later added back into the final image (i.e. no D-stretch is applied to these pixels). Figure 4 shows example results of extracting the shading and reflectance components of an image.

5. Results and Applications

In this section we present results of applying D-stretch under various scenarios using different variations and towards a selection of different goals.

5.1. Stretching Depth

Figure 5 shows the effect of D-stretch on an image from the MIT Intrinsic Images dataset [11]. The Database provides the ground truth intrinsic image as well as the true shading image associated with the source image. D-stretch is applied to the shading image according to Equation 5 with different values of α. The change in perceived depth is significant as can be seen in the images.

In the case when the ground truth shading image is not provided, we extract the shading image (as well as the intrinsic image) using the method described in Section 4. D-stretch is then applied to the computed shading image and recombined with the intrinsic image as in Figure 4. Figures 6-8 display D-stretch examples of coins, art reliefs, maps and textiles. The original image in each pair is on the left and the D-stretched image is on the right. As can be seen, D-stretched coin images show enhanced depth of the imprinted figures and text (Figure 6) and the shading cues in the relief art are enhanced following D-stretch (Figure 7). These percepts are also confirmed in a user study as described in Section 6.4.

5.2. Enhancing Flash Photography

Flash photography is used in dark scenes to illuminate the objects of interest but also in situations where a directed light source forms displeasing shadows in the scene (see Figures 9a). In both cases it is well known that flash photography, although illuminating the scene well, often causes

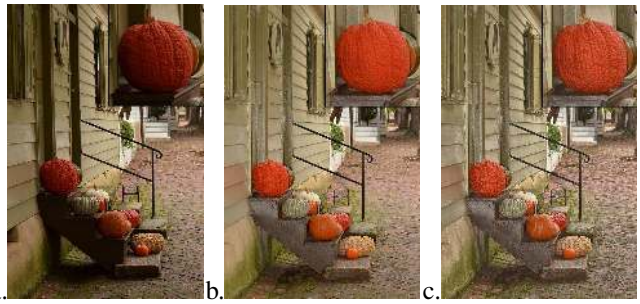


Figure 9. D-stretch applied to flash photography. a) No flash image has dark regions. b) Flash photography illuminates dark regions but attenuates shading cues. c) D-stretch enhances shading cues lost in flash. Pumpkin region enlarged at top for visualization.

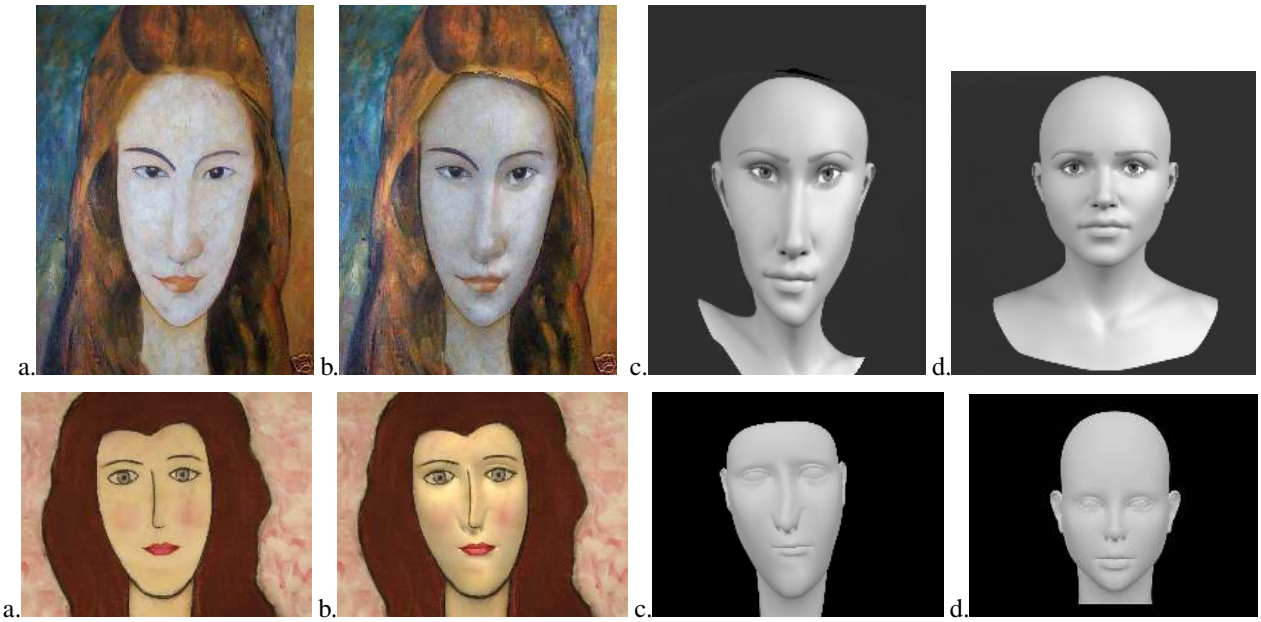


Figure 10. D-stretch using generic face shading image. a. Original image (Hbuterne by Modigliani - top) b. D-stretched face (alpha = 2.0). c. Shading image used (obtained by warping generic face shading image). d. Original generic face shading image.

"flattening" of the scene, namely attenuation of the shading cues [18] (see e.g. Figures 9b). Several studies attempt to recover the image details and accordingly, the shading cues, in the flash photograph by incorporating information from the non-flash image [25, 14, 26, 5]. These approaches require access to both flash and non-flash images of the same scene. In contrast, D-stretch can be applied directly to the flash images to re-introduce or enhance the shading cues without necessitating additional images. Figures 9c shows an example of D-stretch applied to a flashed photography. Notice the enhanced shading cues in the pumpkins and pavement tiles of Figure 9.

5.3. Using Generic Shading Models

In the case when the ground truth shading image is not provided or can not be computed from the image, another option is to exploit a generic shading image. Figure 10-top shows an example of D-stretching an image of Modigliani's Hébuterne painting. Modigliani's style of "flat" faces does not allow us to extract a shading image. Thus a generic face shading image is used instead (Figure 10d). The generic shading image is geometrically warped to align with the geometric features of the face in the original image (Figure 10c). The warped shading image is used to D-stretch the original Modigliani image shown in Figure 10a. Another example is given in Figure 10-bottom where a "flat painting" has been D-stretched.

5.4. Depth Shrinking: De-wrinkling & De-pimpling

The process of D-stretching can be modified to attenuate rather than enhance shading cues by setting $\alpha < 1$. This is useful in cases where shading is to be "flattened out". Classic examples are skin wrinkles and pimples. Figure 11 shows an example of D-shrinking an image of skin wrinkles. Skin images contain shading cues at different frequency scales. Figure 12-left shows an example of D-shrinking the low frequency content of the shading thus preserving the fine details of the skin texture. Figure 12-right shows an example where the skin image contains fine texture details as well as low frequency shading content that induces the percept of a curved arm. D-shrinking the shading image produces a "flattened" image where in addition to the attenuated pimples, both the fine skin features and the surface curvature have been attenuated if not eliminated. To attenuate the pimples alone, the shading image is filtered so that only mid frequency shading content is attenuated by D-shrink. The resulting image in Figure 12c preserves the fine skin features as well as the curved structure of the arm, while attenuating the pimples. Note, that in addition to the shading cues, the pimples are also distinguished by change in albedo, this however, is not removed by D-shrinking.

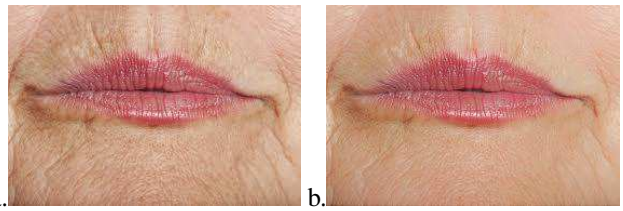


Figure 11. D-shrink applied to wrinkled skin. a) Original image. b) Applying D-shrink attenuates wrinkles.

a. b. c. a. b. c.

Figure 12. D-shrink applied to pimpled skin. a) Original image. b) Applying D-shrink to all shading content produces attenuated pimples however fine features as well as the global structural depth are attenuated as well. c) Applying D-shrink to low (left) or mid (right) frequency content of the shading image, pimples are attenuated yet the original skin's fine features as well as the global structure are preserved.

6. Discussion

6.1. Lighting Position

The D-stretch formulation developed in Section 3 assumes that the lighting direction coincides with the camera viewing direction $\mathbf{l} = [0, 0, -1]$. When lighting direction deviates from the assumed direction, D-stretch can be shown to provide shading effects as expected from stretching the scene along the light direction. This effect still provides enhancement in depth perception as long as the light direction does not deviate greatly from the viewing direction. Figure 13 plots the differences between a rendered 3D model and D-stretch results. The 3D model used in Figure 2 was scaled along its z-axis at various scale factors. The MSE between the scaled model's image and the D-stretched original model image is plotted as a function of the scale factor (D-stretch parameter α). Plots are shown for different lighting positions given as deviation from the frontal-parallel position. It can be seen that when lighting is fronto-parallel ($0°$) as assumed in the D-stretch derivation, indeed the resulting D-stretched image perfectly reproduces the image of the scaled model (MSE = 0) for all scale factors. As the lighting position deviates from $0°$, error increases with deviation angle. Error is greater for larger scaling factors. Visually, however the D-stretch still produces the percept of enhanced depth even for extreme lighting position, as can be seen in Figure 14 for a deviation of $40°$.

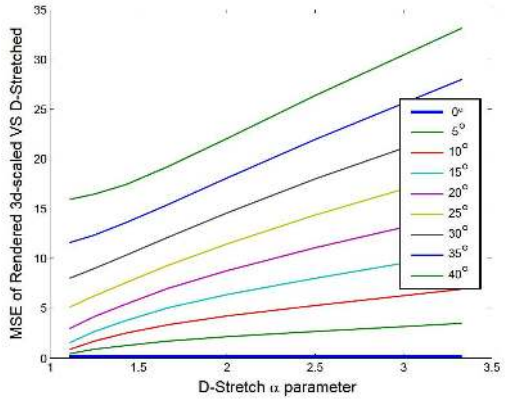


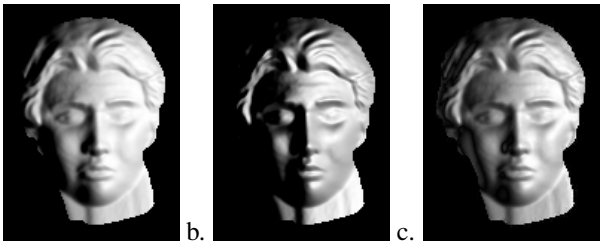
Figure 13. MSE difference between D-stretch results and rendered 3D model as a function of the D-stretch parameter alpha, for different lighting positions given as deviation from the frontal-parallel lighting position.

6.2. D-stretch vs. Gamma Tone Mapping

Both D-stretch and Gamma tone mapping are point operations with a non linear tone mapping curve. There are two main distinctions that we emphasize. First, Gamma manipulation is applied directly to pixel values whereas D-stretch is applied to the shading component. Indeed, D-stretching a scene with a uniform reflectance boils down to applying a global tone-mapping. In such cases, Gamma mapping (with $\gamma > 1$) bears similarity to the D-stretch operation. However, in other cases where chromatic edges exist in the image the two approach are conceptually different (see Figure 15). Second, Gamma Correction is typically applied to correct for the camera to monitor non-linearities, thus $\gamma < 1$ is used (typical value is around $\gamma = 1/2.2$). The D-stretch mapping curve enhances depth when $\alpha > 1$ having a mapping curve that significantly differs. Figure 16a shows the D-stretch tone mapping curve for various α values and Figure 16b shows the Gamma curve for different values of Gamma. Even for similar values of $\alpha > 1$, the tone curves of Gamma and D-stretch differ in mode (e.g. consider the $\alpha = 5$ curve): D-stretch more strongly attenuates the lighter tones compared to the Gamma Mapping and the Gamma Mapping affects the darker tones more strongly than D-stretch.

6.3. D-stretch v.s. Detail Enhancement

A significant body of work has been produced in the context of detail enhancement. These techniques were not de-



a. b. c.

Figure 14. Comparison of rendered 3D model vs D-stretch for extreme lighting angle. a) Rendered Model b) Rendered 3D scaled model c) D-stretch of (a) with $\alpha = 1.5$. Lighting position is at $40°$ deviation from frontal-parallel.

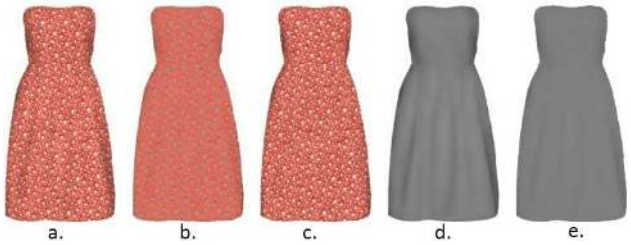


Figure 15. D-stretch v.s. Gamma Mapping. a) Original b) Gamma Mapping (D-shrinking the intensity image) c) D-shrinking (using the shading image) d) Original shading image e) Shading image after D-shrinking.

signed specifically to enhance 3D sensation, but rather produce an 'enhanced' and more pleasing image enabling an improved visibility of edges and details. However, the end results are often comparable. We refer here to two types of enhancements: The high dynamic range (HDR) compression techniques [31, 19, 4, 9] and the detail enhancement techniques [6, 31, 3]. The HDR methods, consider large changes in the image to be due to illumination changes. The image is decomposed into 'base' and 'details' components, where the 'base' is extracted using edge-aware smoothing operations in the log domain The illumination changes are assumed to be sharp and strong, thus are encoded in the 'base' component. For enhancement, either the variations in the 'base' component are attenuated, or the 'details' component is stretched, before recombining the two components. In both cases the 'details' are amplified relative to the 'base'. Similarly, in the D-stretch case, the goal is to separate the shading component from the reflectance component and stretching the shading component. Although the methods are similar in spirit they are conceptually different. The first distinction is that HDR methods attempt to separate the illumination from the reflectance components and do not consider the shading component per se. While detail enhancement methods do not care about the physical source of image gradients and simply follow the goal of amplifying small and mid range variations. The second distinction is in the shape of the mapping curve applied. In our method the curve is theoretically derived from the image formation model while in the enhancement methods the details are linearly multiplied in the log domain resulting in a tone mapping similar to a Gamma curve in the linear domain.

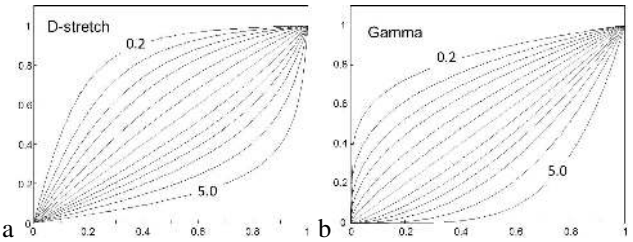


Figure 16. D-stretch vs Gamma. a) D-stretch tone curves for various α values ranging in $\{0.2 \ldots 5.0\}$. b) Gamma curves for various Gamma values ranging in $\{0.2 \ldots 5.0\}$.

6.4. User Study

Finally, we support our results with a user study that shows that indeed visual percept of the D-stretched images strongly correlates with stretching of scene depth. 37 subjects, of clear or corrected vision, ranging in ages 20-38, were recruited. Subjects were shown 20 pairs of images consisting of an original color image and its D-stretched version. Positioning of the original and the D-stretched image on the screen, was randomized. Examples of such images include figures shown above, in this paper. Subjects were asked to select the image perceived to have greater depth variation. Results show that subjects selected the D-stretched image as that showing greatest variation in depth, between 90-100% of the time for each image pair with a total average response of 96.44% over all image pairs.

Next, subjects were shown randomized sequences of 4 D-stretched images of the same original, using 4 different α values as in Figure 5. Subjects were asked to rank the images in order of increasing depth percept. The resulting ranking confusion matrix shows consistency between perceived ranking of depth and the α value. The Pearson Correlation coefficient for these results is 0.892.

Depth Order Ranking Confusion Matrix

depth order	1	2	3	4
1	66	5	1	2
2	5	65	3	1
3	1	3	64	6
4	2	1	6	65

In a final test, users were again shown pairs of images consisting of an original color image and its D-stretched version. Subjects were asked to select the image which is perceived as "better looking". This is an ambiguous request and many factors independent of D-stretch, may come into play affecting the subjects' responses. Notwithstanding, results show that 70% of the responses indicated a preference for the D-stretched images. The largest inconsistencies were found for images containing faces or human subjects, known to be a source of complexity in terms of perception. For the remaining test set, without faces, responses showed 80% preferences for the D-stretched images.

7. Conclusions

We have presented a simple tone mapping operation termed *D-stretch* that is shown to improve the depth percept of an image. The attractive feature about the proposed approach is that it does not require any 3D reconstruction and is applied directly to the shading component of an image. The approach is mathematically proven and simulation results demonstrate the visible effect of depth enhancement. Results are further supported by a user study.

References

[1] J. T. Barron and J. Malik. Shape, illumination, and reflectance from shading. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2015. 1, 4

[2] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. In A. Hanson and E. Riseman, editors, *Computer Vision Systems*, pages 3–26. Academic Press, New York, 1978. 1, 4

[3] J. Chen, S. Paris, and F. Durand. Real-time edge-aware image processing with the bilateral grid. In *ACM Trans. Graph.*, volume 26(3), page 103. ACM, 2007. 8

[4] F. Durand and J. Dorsey. Fast bilateral fltering for the display of high- dynamic-range images. In *ACM Trans. Graph.*, volume 21(3), pages 257–266, 2002. 2, 8

[5] E. Eisemann and F. Durand. Flash photography enhancement via intrinsic relighting. *ACM Trans. Graph.*, 23(3):673–678, 2004. 6

[6] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. In *ACM Trans. Graph.*, volume 27(3), page 67. ACM, 2008. 2, 8

[7] R. Fattal. Single image dehazing. *ACM Trans. Graph.*, 27(3):72:1–72:9, 2008. 1

[8] R. Fattal, M. Agrawala, and S. Rusinkiewicz. Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans. Graph.*, 26(3):51, 2007. 2

[9] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In *ACM Trans. Graph.*, volume 21(3), pages 249–256. ACM, 2002. 2, 8

[10] G. Finlayson, M. Drew, and C. Lu. Intrinsic images by entropy minimization. In *Euro. Conf. Computer Vision (ECCV)*, 2004. 4

[11] R. Grosse, M. Johnson, E. Adelson, and W. Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *Int. Conf. Comp. Vis.*, 2009. 4, 5

[12] B. K. Horn. *Obtaining shape from shading information*. MIT press, 1989. 1, 2

[13] I. Howard. *Perceiving in Depth*. Oxford University Press, 2012. 1

[14] J. Jia, J. Sun, C. Tang, and H. Shum. Bayesian correction of image intensity with spatial consideration. In *Euro. Conf. Computer Vision (ECCV)*, pages III: 342–354, 2004. 6

[15] S. Johnston. Lumo: Illumination for cel animation. In *Proceedings of the 2nd International Symposium on Non-photorealistic Animation and Rendering*, NPAR02, pages 45–52, 2002. 2

[16] P. Laffont, A. Bousseau, and G. Drettakis. Rich intrinsic image decomposition of outdoor scenes from multiple views. *IEEE Trans. Vis. and Comp. Graph.*, 19(2):210–224, 2012. 4

[17] E. H. Land and J. J. McCann. Lightness and retinex theory. *J. Opt. Soc. Am.*, 61(1):1–11, 1971. 4

[18] M. Langford, A. Fox, and R. Smith. *Langford's Basic Photography: The Guide for Serious Photographers*. Focal Press, 2010. 1, 6

[19] Y. Li, L. Sharan, and E. H. Adelson. Compressing and companding high dynamic range images with subband architectures. *ACM Trans. Graph.*, 24(3):836–844, 2005. 8

[20] J. Lopez-Moreno, J. Jimenez, S. Hadap, K. Anjyo, E. Reinhard, and D. Gutierrez. Non-photorealistic, depth-based image editing. *Computers and Graphics*, 35:99–111, 2011. 2

[21] T. Luft, C. Colditz, and O. Deussen. Image enhancement by unsharp masking the depth buffer. *ACM Trans. Graph.*, 25(3):1206–1213, jul 2006. 2

[22] T. Malzbender, T. Malzbender, D. Gelb, D. Gelb, H. Wolters, and H. Wolters. Polynomial texture maps. In *In Computer Graphics, SIGGRAPH 2001 Proceedings*, pages 519–528, 2001. 2

[23] Y. Matsushita, S. L. S. B. Kang, and H. Shum. Estimating intrinsic images from image sequences with biased illumination. In *Eur. Conf. Comput. Vis.*, pages 274–286, 2004. 4

[24] S. Narasimhan and S. Nayar. Vision and the atmosphere. *Int. Journal of Computer Vision*, 48(3):233–254, 2002. 1

[25] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. *ACM Trans. Graph.*, 23(3):664–672, 2004. 6

[26] R. Raskar, K. Tan, R. Feris, J. Yu, and M. Turk. Non-photorealistic camera: Depth edge detection and stylized rendering using multi-flash imaging. *ACM Trans. Graph.*, 23(3):679–688, 2004. 6

[27] T. Ritschel, K. Smith, M. Ihrke, T. Grosch, K. Myszkowski, and H. Seidel. 3D Unsharp Masking for Scene Coherent Enhancement. *ACM Trans. Graph.*, 27(3), 2008. 2

[28] L. Shen and C. Yeo. Intrinsic image decomposition using a local and global sparse representation of reflectance. In *IEEE Conf. Computer Vision and Pattern Recognition*, 2011. 4

[29] X. Sun, P. Rosin, R. Martin, and F. Langbein. Bas-relief generation using adaptive histogram equalization. *IEEE Trans. Vis. and Comp. Graphics*, 15(4):642–653, July 2009. 2

[30] M. Tappen, W. Freeman, and H. Adelson. Recovering intrinsic images from a single image. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(9):1459–1472, 2005. 4

[31] J. Tumblin and G. Turk. LCIS: A boundary hierarchy for detail-preserving contrast reduction. In *Proc. 26th annual conference on Computer Graphics and Interactive Techniques*, pages 83–90. ACM Press/Addison-Wesley Publishing Co., 1999. 2, 8

[32] R. Vergne, R. Pacanowski, P. Barla, X. Granier, and C. Shlick. Improving shape depiction under arbitrary rendering. *IEEE Trans. Vis. and Comp. Graphics*, 17(8):1071–1081, Aug 2011. 2

[33] Y. Weiss. Deriving intrinsic images from image sequences. In *Int. Conf. Computer Vision*, pages 68–75, 2001. 4

[34] T. Weyrich, J. Deng, C. Barnes, S.Rusinkiewicz, and A. Finkelstein. Digital bas-relief from 3D scenes. *ACM Trans. Graph.*, 26(3), Aug. 2007. 2

[35] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139–191139, 1980. 1

[36] R. Zhang, P.-S. Tsai, J. Cryer, and M. Shah. Shape-from-shading: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(8):690–706, 1999. 1