The Visual Computer manuscript No.
(will be inserted by the editor)

Perceptually Meaningful Image Editing

Manipulating perceived depth and creating the illusion of motion in 2D images

Reynold J. Bailey and Cindy M. Grimm⋆

Washington University in St. Louis

Received: date / Revised version: date

Abstract We introduce a novel concept called *perceptually meaningful image editing* and present techniques for manipulating the apparent depth of objects and creating the illusion of motion in 2D images. Our techniques combine principles of human visual perception with approaches developed by traditional artists. For our depth manipulation technique, the user loads an image, selects an object and specifies whether the object should appear closer or further away. The system automatically determines luminance or color temperature target values for the object and/or background that achieve the desired depth change. Our approach for creating the illusion of motion exploits the differences between our peripheral vision and our foveal vision by introducing spatial imprecision to the image.

1 Introduction

Traditional artists have developed numerous techniques for creating interesting visual effects. Although much of their work was done on flat, static surfaces, they were still able to create a sense of depth and motion in their paintings. Many of these artists had very little knowledge of the inner workings of the human visual system. Instead, they viewed the human visual system as a black box and through experimentation they learned to exploit its features. In a sense they have reverse engineered the human visual system to learn what type of inputs elicit certain responses in the brain.

Modern research from the fields of Biology, Physics, Psychology, Physiology, and Neuroscience has given us better insight into the functioning of the human visual system. Although the visual system is far from being fully understood, the knowledge we have gained, especially about the early stages of the visual pathway, is quite substantial. In fact, many of the visual effects created by traditional artists can now be explained in terms of the features of the human visual system. By tapping into this wealth of knowledge, a new class of image editing techniques which are *perceptually meaningful* can be developed. We define a perceptually meaningful editing technique as one designed to explicitly trigger certain visual cues. In this paper we present two such techniques - one designed to manipulate the apparent depth of objects in an image and the other designed to introduce the illusion of motion to an otherwise visually static image.

Fig. 1 Enhancing apparent depth using luminance (left) and color temperature (right). Input image (middle).

Our depth editing technique (see Fig. 1) is based on the fact that traditional artists generally manipulate the apparent depth of objects in a scene (without changing the scene geometry) either by changing the color temperature (warmth/coolness) or the luminance values in particular regions of the scene.

Colors that appear closer to the red end of the visible spectrum are said to be warm while the colors that appear closer to the blue end are said to be cool. In general, it has been observed in psychophysical experiments that warmer colors appear to advance toward the viewer while cooler colors appear to recede [22]. These perceived depth differences can be explained by the fact

⋆ rjb1@cse.wustl.edu, cmg@cse.wustl.edu

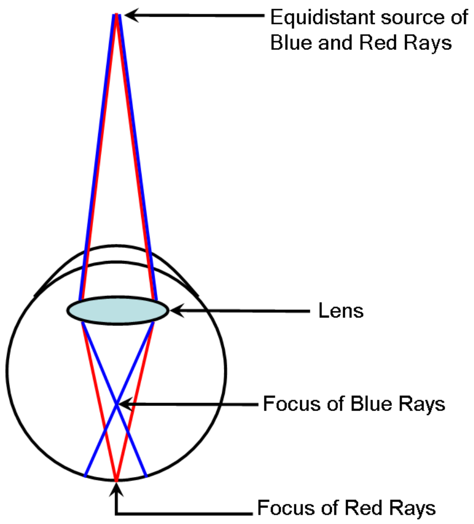

Fig. 2 Chromatic aberration. Assumes eye is red accommodated. Adapted from Sundet [22].

that shorter wavelengths of visible light are refracted more than longer wavelengths. As a result, equidistant sources of differing wavelengths cannot be simultaneously focused onto the retina by the eye's optical system. This is referred to as *chromatic aberration* and is illustrated in Fig. 2.

Artists also create apparent depth in their paintings by manipulating the luminance range. They do this by abruptly changing the luminance values across the boundary between two objects in specific regions of the painting. This is sometimes referred to as *haloing* and the resulting depth effect can be explained in terms of the center-surround organization [13] of the cells in our visual system. Center-surround cells respond more intensely to local discontinuities than to gradual changes in a scene. Hence, a sharp luminance change across the boundary between an object and the background causes the object to appear more advanced.

Techniques for manipulating luminance values and the warmth (or coolness) of colors are well established in existing editing systems. However, their use for adjusting apparent depth in an image is typically not a defined operation. Users are still left with the time consuming task of manually selecting and manipulating groups of pixels. Additionally, users have to ensure that any edits they make are seamless. Our depth editing technique eliminates these drawbacks. The user simply selects an object and specifies whether it should appear closer or further away. The system automatically determines luminance or color temperature target values for the object and/or background that achieve the desired depth change. Distance-weighted blending is used to ensure that the appearance of false edges is minimized.

Our technique for introducing the illusion of motion to a visually static scene is also motivated by the work of traditional artists. One common approach that was used to create the illusion of motion in their paintings is *spatial imprecision* (misalignment of brush strokes) as illustrated in Fig. 3 (left).

The illusion of motion occurs because of differences in our peripheral and foveal vision. The fovea is a small indentation in the retina where we have the highest visual acuity. Our peripheral vision, which occupies the majority of our visual field has much lower acuity. Ogden and Miller discovered that optic fibers that carry signals from the fovea to the lateral geniculate nucleus, where they are processed, are slow-conducting while optic fibers that carry signals from the peripheral regions are faster [18]. When we first look at paintings where spatial imprecision is used, our fast acting, low acuity peripheral vision gives us a rough idea of where the brush strokes are in the scene. Mentally, we join these brush strokes together to form a complete picture. This process is called *illusory conjunction*. It is only upon closer scrutiny with our slower, high acuity foveal vision that we notice that the strokes are misaligned. The illusion of motion is created because our visual system completes the picture differently with every glance (explanation adapted from Vision and Art: The Biology of Seeing [15]).

Our technique uses a two-step process to introduce spatial imprecision to an image. The input image is first segmented into regions of roughly uniform color and the resulting segments are then spatially perturbed. This technique can be applied over the entire image or to specific regions of the image.

2 Previous Work

2.1 Depth

Various sources of information in a scene contribute to our perception of depth. These are referred to as depth cues. Pictorial depth cues are ones we obtain from 2D images. These include perspective cues such as occlusion, relative size, distance to horizon, and shadows; and non-perspective cues such as color, relative brightness, shading, focus, and atmospheric effects [19].

Traditional computer graphics approaches for manipulating depth are typically scene-based (relying on 3D representations where depth information is readily available). These include the introduction of atmospheric effects like haze or fog [20], perspective changes such as the shape, size, or position of objects (or camera) [7], and level of detail (LOD) changes [11]. Very little work has been done in the area of image-based depth manipulation. The most commonly used approach is to simulate the depth-of-field effect from traditional photography. This effect can be achieved in commercially available image editing packages by applying a sharpening filter to the foreground and a blurring filter to the background. This has the effect of bringing different areas of an image in or out of focus. There has also been some recent work by Gooch and Gooch which suggests that the perceived

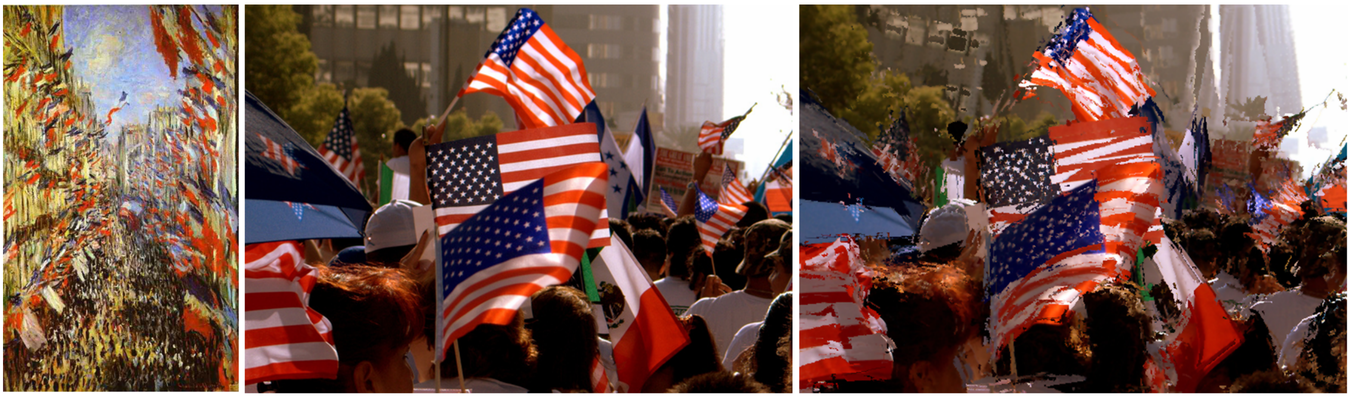

Fig. 3 Left: Example of spatial imprecision to create the illusion of motion. Claude Monet, Rue Montorgueil - Paris, Festival of June 30, 1878 (1878). Center: Input image. Right: Output image.

depth in an image can be enhanced by adding an artistic matting [10].

2.2 Motion

The human visual system is well equipped to detect actual motion in a scene. Certain types of cells along the visual pathway called *double-opponent cells* are directionally selective - responding to motion in a preferred direction. This type of motion perception enables us to determine the speed and direction of objects in the environment.

It is also possible for the visual system to infer motion from a static scene. Static repeated asymmetric patterns (RAPs), commonly used in optical illusions, are known to generate spurious motion signals. There is still a lot of debate over the exact neural mechanisms that cause this effect [1]. Patterns of adjacent equiluminant colors also trigger spurious motion signals [23] [15]. This occurs because our visual system cannot accurately detect the edges between these patterns hence they appear to float. Finally, spatial imprecision (as described in Section 1) also leads to the illusion of motion.

Computer graphics techniques for manipulating static images to trigger a motion response from the visual system typically rely on animation (changing the image over time). Color table animation [21], one of the earliest techniques developed, worked by swapping or cycling through colors in the frame buffer's color table. Freeman et al. [8] later demonstrated that applying local filters to images and continuously varying their phase over time creates the illusion of motion. More recently, Chuang et al. [4] created animated clips from single images by applying time varying 2D displacement maps to various segmented regions of the original scene.

Several static techniques, that do not rely on animation to trigger a motion response, have also been developed. The most common are non-photorealistic techniques for creating cartoon-style effects that suggest motion in a particular direction direction [16] [12] [17].

2.3 Object Selection/Image Segmentation

Object selection and (more generally) image segmentation play important roles in the techniques presented in this paper. Any existing object selection/image segmentation technique can be used (see [9] for a review). For our depth manipulation technique, we use an implementation of the *Lazy Snapping* method [14] to perform object selection, and for our technique for creating the illusion of motion, we use an approach proposed by Comaniciu and Meer [5] to perform image segmentation.

3 Depth Editing Technique

This section describes how luminance and color temperature target values are automatically generated for our depth editing technique. Target values are assigned to the pixels that lie on the object boundary (i.e. object pixels that lie closest to the edge) and to pixels that lie on the background boundary (i.e. background pixels that lie closest to the edge). These values are then propagated to every pixel in the image.

3.1 Notation

Let $\mathbf{B}$ be the set of background pixels, $\mathbf{O}$ be the set of object pixels, $\mathbf{I} = \mathbf{O} \cup \mathbf{B}$ be the set of all image pixels, T_{lum} be the set of luminance target values, and T_{col} be the set of color temperature target values. Let $p_o \in \mathbf{O}$ be an object boundary pixel and $p_b \in \mathbf{B}$ be the closest corresponding background boundary pixel. Let n be some user specified neighborhood size ($n = 9$ for the images in this paper) and let N_{p_o} and N_{p_b} be the n x n neighborhoods about p_o and p_b respectively.

3.2 Luminance Target Values

Let

$$Avg_o = \frac{1}{sizeof\,\{N_{p_o} \cap O\}} \sum_{p_i \in \{N_{p_o} \cap O\}} p_i$$

and

$$Avg_b = \frac{1}{sizeof\{N_{p_b} \cap B\}} \sum_{p_i \in \{N_{p_b} \cap B\}} p_i$$

be the average neighborhood pixel values of p_o and p_b respectively. Furthermore let

$$AvgLum_o = luminance(Avg_o)$$

and

$$AvgLum_b = luminance(Avg_b)$$

be the average neighborhood luminance values of p_o and p_b respectively. Target luminance values T_{lum} are assigned to p_o and p_b using the following rules:

- Case 1 - Reducing apparent depth by manipulating both object and background:
  - $T_{lum}(p_o) = T_{lum}(p_b) = \frac{1}{2}(AvgLum_o + AvgLum_b)$
- Case 2 - Reducing apparent depth by manipulating the object only:
  - $T_{lum}(p_o) = AvgLum_b$
- Case 3 - Reducing apparent depth by manipulating the background only:
  - $T_{lum}(p_b) = AvgLum_o$
- Case 4 - Enhancing apparent depth by manipulating either the object and/or background:
  - if $AvgLum_b \leq AvgLum_o$
    - $T_{lum}(p_o) = 1.0$ (white)
    - $T_{lum}(p_b) = 0.0$ (black)
  - if $AvgLum_o < AvgLum_b$
    - $T_{lum}(p_o) = 0.0$ (black)
    - $T_{lum}(p_b) = 1.0$ (white)

Once boundary target values have been assigned, they can be propagated to all pixels in the image using any flood-fill algorithm. We use a fast breadth-first approach. A Gaussian blur filter is applied to the resulting target image. This gives smoothly varying target values and hence a more natural looking result once the depth editing operation is complete. Target values for all cases can be precomputed and stored as images. This reduces lag time when performing depth manipulation, especially on large images.

When assigning target values for reducing apparent depth, it should be noted that we can substitute Avg_o and Avg_b for $AvgLum_o$ and $AvgLum_b$ respectively. This has the same effect of reducing luminance contrast between the object and the background, however it does not limit us to grayscale target values. We found that this generally gives more aesthetically pleasing results and is the approach used for the images in this paper. The top row of Fig. 4 shows the target luminance values generated for a hypothetical image.

3.3 Color Temperature Target Values

In order to assign target color temperature values, we simply modify the target luminance values T_{lum}. The user specifies a warm color w and a cool color c. Target color temperature values T_{col} are assigned to p_o and p_b using the following rules:

- Case 1 - Reducing apparent depth by manipulating both object and background:
  - $T_{col}(p_o) = T_{col}(p_b) = \frac{1}{2}(T_{lum}(p_o) + c)$
- Case 2 - Reducing apparent depth by manipulating the object only:
  - $T_{col}(p_o) = \frac{1}{2}(T_{lum}(p_o) + c)$
- Case 3 - Reducing apparent depth by manipulating the background only:
  - $T_{col}(p_b) = \frac{1}{2}(T_{lum}(p_b) + w)$
- Case 4 - Enhancing apparent depth by manipulating either the object and/or background:
  - if $AvgLum_b \leq AvgLum_o$
    - $T_{col}(p_o) = \frac{1}{2}(T_{lum}(p_o) + w)$
    - $T_{col}(p_b) = \frac{1}{2}(T_{lum}(p_b) + c)$
  - if $AvgLum_o < AvgLum_b$
    - $T_{col}(p_o) = \frac{1}{2}(T_{lum}(p_o) + c)$
    - $T_{col}(p_b) = \frac{1}{2}(T_{lum}(p_b) + w)$

We experimented with other approaches for assigning target color temperature values including estimating the color temperature (position along the spectrum) of values in the image then traversing the spectrum to the right or left to make the target values warmer or cooler. Such approaches are generally computationally expensive and do not offer any significant advantage over the approach presented here. The bottom row of Fig. 4 shows the target color temperature values generated for a hypothetical image.

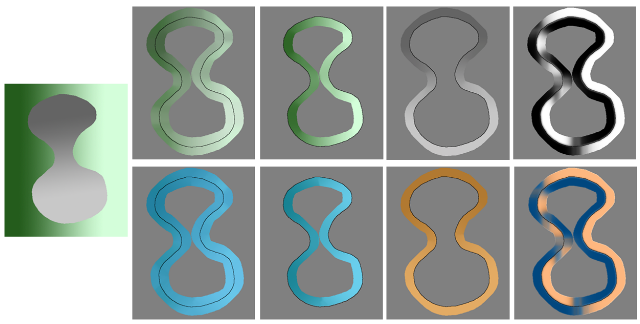

Fig. 4 Hypothetical input image (left). Target luminance values (top row). Target color temperature values (bottom row). Case 1, 2, 3, and 4 (left to right). The original edge is shown by a black line. Only a 60 pixel wide area of the target values are shown to better illustrate the relationship with the original image.

Fig. 5 Manipulating perceived depth by simple color replacement. Input image (left). Luminance target values (middle column). Color temperature target values (right column).

3.4 Results

The actual value V assigned to a pixel during depth manipulation is a weighted combination of the target pixel value T_{lum}(p) (or T_{col}(p)) and the original pixel value $\mathbf{I}$(p) at that point:

$$V(p) = (1 - f(d))\mathbf{I}(p) + f(d)T_{lum}(p)$$

where the parameter d, is a user specified (slider) depth value. f is some user specified falloff function of distance from every pixel to the edge between the object and background. We experimented with Gaussian, quadratic, and linear functions - all with pleasing results. Additionally, the user is allowed to select the distance over which blending occurs or have the system generate variable distances based on distance to the object skeleton.

With blending turned off and a zero falloff function, adjusting the apparent depth simply results in color replacement as illustrated in Fig. 5. Notice that the effect of luminance on perceived depth (middle column) is immediately obvious (brighter objects appear closer), however, the relationship between color and depth (right column) is more subtle. In a psychophysical study, conducted by Bailey et al. (using stimuli generated by our technique), they recorded the depth preferences of human observers presented with images of differently colored, equiluminant objects. They found that there is a clear relationship between the color temperature of an object and its perceived depth and that this relationship is especially pronounced when the objects are presented against darker backgrounds [2].

By varying the distances and falloff function, we can manipulate the perceived depth of objects by adding undertones and halos. Fig. 1 shows an example of enhancing perceived depth and Fig. 6 shows an example of reducing perceived depth. Notice that drastically reducing perceived depth results in an interesting ghosting effect. This occurs because target values are propagated from the boundary to every pixel in the image. In this manner we obtain an estimate of the pixel values behind the object. A similar effect can be achieved by performing a cutout of the object, using an image inpainting [3] or texture synthesis [6] approach to estimate the missing pixel values, and finally blending the object with the new pixels.

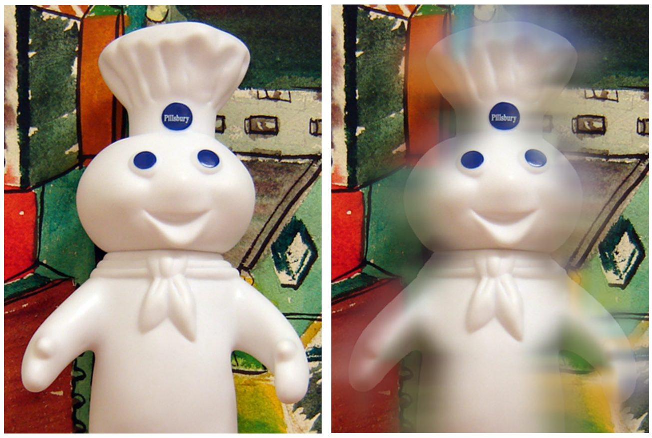

Fig. 6 Drastically reducing apparent depth results in a ghosting effect. Input image (left). Output image (right).

In the case of photorealistic editing (especially of human subjects), this type of image manipulation needs to be more subtle. Fig. 7 shows an example of combining luminance-based manipulation with color temperature-based manipulation in a very subtle manner to change the perceived depth of individuals in a photograph. In this manner, we can draw the viewer's focus to the individual who is furthest away from the camera. The difference image (top right) highlights just how subtle these manipulations have to be to result in a perceptual change.

4 Illusion of Motion Technique

In this section, we describe a simple technique for creating the illusion of motion by introducing spatial imprecision to a visually static image.

The input image is first segmented into regions of roughly uniform color. Each segment is assigned a unique ID and a local Cartesian coordinate system whose origin was arbitrarily chosen to correspond to the top-left corner of the bounding box of the segment. Different 2D spatial transformations are then applied to each segment with respect to its local coordinate system. For the images shown in this paper, the 2D spatial transformation comprised of a small random rotation (between -10 degrees and 10 degrees) and a small random translation (at most 20 pixels).

Let $p \in \mathbf{S}$ be a pixel in a segment $\mathbf{S}$. After undergoing the 2D spatial transformation, its new location p' is given by the standard composite transformation expression:

$$p' = \mathbf{T} * \mathbf{R} * p$$

where p and p' are expressed in homogeneous coordinates, $\mathbf{T}$ represents the translation matrix and $\mathbf{R}$ represents the rotation matrix.

Fig. 7 Combining luminance-based manipulation with color temperature-based manipulation to change the perceived depth of individuals in an image. Input image (left). Output image (right). Difference image (upper right) shows that a warm reddish undertone was added around the face of the girl on the left and that the luminance of her face was increased slightly. A cooler gray-green undertone was added around the face of the girl on the right. Scaled difference image (lower right) emphasizes these changes for display purposes.

Fig. 8 Creating the illusion of motion by introducing spatial imprecision. Input image (top). Output image (bottom).

Fig. 9 . Example of applying the illusion of motion technique to a specific region (foreground) of an image. Input image (top): Nicolas Poussin, The Rape of the Sabine Women (1634). Output image (bottom).

Perturbing segments in this manner introduces holes in the image plane. We experimented with various *inpainting* [3] approaches for filling these holes, but found that simply pasting the perturbed segments onto the original image gave better results. Fig. 3 and Fig. 8 show examples of applying this technique over entire images and Fig. 9 shows an example of applying this technique to a specific region of an image.

One drawback to naively performing random perturbations is that perceptually relevant segments (such as faces) may become (partially) occluded by other segments. To eliminate/minimize this problem, we plan to extend our framework to allow users to specify the relative importance of the various segments. By processing the most important segments last, we can ensure that they are not occluded.

5 Conclusion

This paper introduced the concept of *perceptually meaningful image editing* as a novel approach to image editing where the goal is to explicitly trigger certain visual cues. This paper also presented two perceptually meaningful techniques for manipulating perceived depth and creating the illusion of motion in 2D images.

Our technique for creating the illusion of motion can be applied over an entire image to create stylized renderings which simulate the use of misaligned brush strokes by traditional artists. It can also be applied to specific regions of an image to add a sense of motion to visually static objects in the scene.

Our depth manipulation technique will be useful in post-production film editing as part of a suite of digital grading tools. The digital grading process involves adding subtle color tones and smoothing effects to various scenes to further enhance the emotion that is being conveyed and also to ensure that there is coherence between shots taken from multiple cameras at different times. Grading techniques can also be applied to specific regions of a scene to emphasize/deemphasize certain objects or characters. Consider a case where the character closest to the camera is not the one performing some (important) action. By manipulating the perceived depth of each character, our technique can be used to draw the viewer's attention to the active character. The traditional approach for accomplishing this involved having an artist manually adjust the luminance range and paint subtle undertones and halos around the characters.

As part of our future work, we plan to conduct several psychophysical experiments to further evaluate the techniques presented in this paper (see [2] for the results of the first study, which looked at the impact of color replacement on perceived depth). We also plan to extend the notion of perceptual meaningful manipulation to 3D scenes.

References

1. Backus, B.T., Oruç, İ.: Illusory motion from change over time in the response to contrast and luminance. Journal of Vision **5**(11), 1055–1069 (2005)
2. Bailey, R., Grimm, C., Davoli, C.: The effect of warm and cool object colors on depth ordering. In: APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization, p. 161. ACM Press, New York, NY, USA (2006)
3. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 417–424. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (2000)
4. Chuang, Y.Y., Goldman, D.B., Zheng, K.C., Curless, B., Salesin, D.H., Szeliski, R.: Animating pictures with stochastic motion textures. ACM Transactions on Graphics **24**(3), 853–860 (2005). Special issue: Proceedings of ACM SIGGRAPH '05
5. Comaniciu, D., Meer, P.: Robust analysis of feature spaces: Color image segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 750–755 (1997)

6. Efros, A.A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 341–346. ACM Press, New York, NY, USA (2001)
7. Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F.: Computer Graphics — Principles and Practice, Second edition in C. The Systems Programming Series. Addison-Wesley (1996)
8. Freeman, W.T., Adelson, E.H., Heeger, D.J.: Motion without movement. In: SIGGRAPH '91: Proceedings of the 18th annual conference on Computer graphics and interactive techniques, pp. 27–30. ACM Press, New York, NY, USA (1991)
9. Freixenet, J., Muñoz, X., Raba, D., Martí, J., Cufí, X.: Yet another survey on image segmentation: Region and boundary information integration. In: ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III, pp. 408–422. Springer-Verlag, London, UK (2002)
10. Gooch, A.A., Gooch, B.: Enhancing perceived depth in images via artistic matting. In: Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization, p. 168. ACM Press, New York, NY, USA (2004)
11. Heckbert, P.S., Garland, M.: Multiresolution modeling for fast rendering. In: Proc. Graphics Interface '94, pp. 43–50. Canadian Inf. Proc. Soc., Banff, Canada (1994)
12. Kawagishi, Y., Hatsuyama, K., Kondo, K.: Cartoon blur: Non-photorealistic motion blur. Computer Graphics International p. 276 (2003)
13. Kuffler, S.W.: Discharge patterns and functional organization of the mammalian retina. Journal of Neurophysiology **16**, 37–68 (1953)
14. Li, Y., Sun, J., Tang, C.K., Shum, H.Y.: Lazy snapping. ACM Transactions on Graphics **23**(3), 303–308 (2004)
15. Livingstone, M.: Vision and Art: The Biology of Seeing. Harry N. Abrams, Inc. (2002)
16. Masuch, M.: Speedlines: Depicting motion in motionless pictures. In: ACM SIGGRAPH '99 Conference abstracts and applications, p. 277. ACM Press, New York, NY, USA (1999)
17. Moriya, T., Takahashi, T., Tanaka, A.: Non-photorealistic rendering technique for depicting 3D objects motion. In: SIGGRAPH '06: Research posters, p. 114. ACM Press, New York, NY, USA (2006)
18. Ogden, T.E., Miller, R.F.: Studies of the optic nerve of the rhesus monkey: Nerve fiber spectrum and physiological properties. Vision Research **6**, 485–506 (1966)
19. Pfautz, J.: Depth perception in computer graphics. Ph.D. thesis, University of Cambridge, Cambridge, UK, (2000)
20. Rogers, D.: Implementing Fog in Direct3D. NVIDIA Corporation (2000)
21. Shoup, R.G.: Color table animation. In: SIGGRAPH '79: Proceedings of the 6th annual conference on Computer graphics and interactive techniques, pp. 8–13. ACM Press, New York, NY, USA (1979)
22. Sundet, J.M.: Effects of colour on perceived depth: Review of experiments and evaluation of theories. Scandinavian Journal of Psychology **19**(2), 133–143 (1978)
23. Zimmerman, G.L., Nguyen, V.: Motion perception at equiluminance and the consequences for computational vision systems. In: D.P. Casasent (ed.) Proceedings SPIE Vol. 1825, Intelligent Robots and Computer Vision XI: Algorithms, Techniques, and Active Vision, pp. 401–412 (1992)

Reynold Bailey received the B.Sc. degree in Mathematics and Computer Science in 2001 from Midwestern State University and the M.S. degree in Computer Science in 2004 from Washington University in St. Louis, where he is currently a Ph.D. student. His research interests include applied perception in graphics and visualization and non-photorealistic rendering.

Cindy Grimm received the Bachelor's degree in Computer Science and Art from University of California, Berkeley in 1990 and the Master's (1992) and Ph.D. (1995) degrees from Brown University. She is a professor in the Department of Computer Science and Engineering at Washington University in St. Louis, where she works in the field of computer graphics.