The choices hidden in photography

Aaron Hertzmann

Adobe Research, San Francisco, CA, USA

Photography is often understood as an objective recording of light measurements, in contrast with the subjective nature of painting. This article argues that photography entails making the same kinds of choices of color, tone, and perspective as in painting, and surveys examples from film photography and smartphone cameras. Hence, understanding picture perception requires treating photography as just one way to make pictures. More research is needed to understand the effects of these choices on pictorial perception, which in turn could lead to the design of new imaging techniques.

Introduction

Do photographs convey objective information to a viewer? In common intuition, they do, being the result of an automatic, mechanical process (Costello, 2017), and so photography holds a special place in how we convey information, whether in journalism, legal proceedings, or social media, and many other areas. Indeed, current fears around the dangers of image manipulation imply a belief in the honesty of unmanipulated images. More nuanced discussions highlight choices made by photographers: a photographer chooses the subject and aims the camera, and selects zoom and exposure settings. These choices determine the content and the aesthetic qualities of an image (Palmer, Schloss, & Sammartino, 2013); much has been written about these subjective choices in documentary photography (Bersak, 2006; Morris, 2014) and social media (Hawley, 2022). But, one might think that, once the photographer presses the shutter-release button, the rest of the imaging process is an objective measurement and display of light. Indeed, in perception and art history, many texts define an ideal *picture* as one that displays light as though the viewer were looking through a window into the depicted scene (Gibson, 1971; Kemp, 1990; Pirenne, 1970; Yang & Kubovy, 1999). Many of these works describe linear perspective as "correct perspective." From this, one might conclude that pinhole cameras create correct pictures, which typical consumer cameras approximate with lenses.

This article describes how, rather than being objective measurements of light, all photographs display light according to hidden, subjective choices made by the photographer and camera manufacturer. These choices determine the depictions of tones, colors, and perspective. It is self-evident that a representational painter must make all of these choices, but, in photography, many of these choices are hidden in optical, mechanical, chemical, and/or software design choices. These choices embed both perceptual and aesthetic preferences, just as they do in painting. Such choices are mandatory. There is no single correct way to make a picture; conventional photographs almost never display light as though the viewer were looking through a window.

Hence, rather than being objective display of light measurements, photography is one type of visual depiction—it is a class of techniques for arranging colors and tones on a flat surface to convey information (Durand, 2007; Gibson, 1971). Hence, studies on pictorial space should begin from this assumption, rather than the assumption that linear perspective and linear tones are correct, but that artists sometimes deviate from them. Further research is needed to understand the nature of depiction of pictorial space with these choices. The computational photography techniques discussed in this article could enable systematic new research on depiction choices in art and photography that could, in turn, lead to new depiction techniques. Another implication is that, when used as experimental stimuli, the differences between photographs and real-world perception may affect the results, even when they are carefully calibrated for accuracy (Snow & Culham, 2021).

The reader is encouraged to try the following informal experiment: take a picture with a smartphone and compare its depictions against real-world appearances, including brightnesses, colors, and relative sizes of objects. On its own, the photograph may look very real, an accurate depiction of the scene. But, on close inspection, one may notice significant differences between the photo and the world—especially in large-scale scenes with significant lighting variations, like a sunlit mountain range or a nighttime city street. This difference is surprising for some viewers (Albert & Efros, 2016), and can help one to appreciate the subjective choices made, even in seemingly automatic,

realistic smartphone photography, and how such choices can seem to be correct if not inspected closely.

Tone and color choices in pictures

Painters and photographers adjust tones within an image to control emphasis and image aesthetics. Tonal adjustment is required; neither paintbrushes nor digital displays can reproduce the extreme range of lightnesses we experience in the real world. Sunlit objects indoors can reflect 30,000 times more light than shadowed objects (Figure 1), yet viewers can appreciate many gradations of light and shadow in such environments, owing to physiological and neural contrast adaptation. Neither print media nor consumer displays offer such absolute brightnesses or contrast ratios. This means that, under common lighting situations, artists and photographers must make choices about how to depict light intensities that the medium cannot reproduce.

In traditional film photography, the photographer directly controls how much light reaches the imaging surface by adjusting aperture and shutter speed, and choosing the film. If they do these steps poorly, the photo will be underexposed or overexposed (Figure 1a). However, professional film photographers do not just control exposure with global parameters. Famous photographs often came from laborious effort in the darkroom, where photographers added more or less light to different image regions (Figure 2), using a process called "dodging and burning." To a nonexpert, this effort does not appear in the final work, and considerable misunderstanding has historically arisen from belief that a photographer just points their camera and presses a button (Costello, 2017).

Consumer photographers do not normally make these choices. Instead, their exposures are determined in part by subjective processing decisions made by camera manufacturers, designed not just for accuracy, but to appeal to consumers' aesthetic preferences.

For much of the twentieth century, consumer photographers shot on film, and so the choices made by film and camera manufacturers played a large, hidden role in consumers' pictures. For example, starting in the 1950s, Kodak calibrated their photochemical formulas and processes by photographing professional models, all of whom were White women, and then asking nonexpert viewers which photos they liked more. Based on these kinds of experiments, an early paper from Kodak concluded that "optimum reproduction of skin color is not 'exact' reproduction" (MacAdam, 1951). According to Peter Hertzmann, who worked at Kodak in the 1970s, these viewers generally preferred photos that made skin look pinker than in real life, and so this is what Kodak film did for a long time. Kodak film worked poorly for darker skin tones; only in the 1970s did Kodak begin to include darker skin tones in their calibration processes (Roth, 2009). Likewise, in broadcast television: according to Emmy-winning video camera designer Jan Van Rooy, "Skin tone reproduction is not just science. It has to deal with the psychology of how people want to look." At least one Japanese manufacturer calibrated televisions according to skin color, with different calibration for TVs sold in the United States versus those sold in Japan (Roth, 2009).
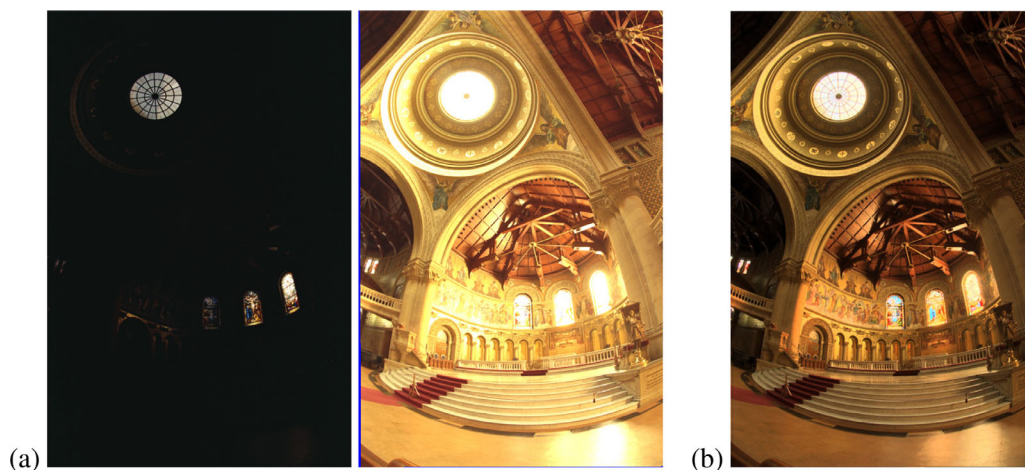

Figure 1. HDR imaging example of a scene with both bright sunlight and dark shadows, from (Debevec & Malik, 2008). (a) Two film photographs with different exposure times: 0.03 seconds and 16 seconds, respectively. The light coming through the stained glass windows is roughly 24,000 times brighter than the light reflecting from shadow regions. (b) HDR tone-mapped image, computationally generated from many different exposures, including the two on the left. The composite reveals fine detail in both the stained glass windows and the shadows, unlike the original exposures.

Figure 2. Professional film photography often entailed extensive local exposure adjustments, called dodging and burning, to manipulate contrast, emphasis, and detail. (a) James Dean, photographed by Dennis Stock (1955). (b) The original photographic proof, together with printing notes made by the printer, Pablo Inrio. © Dennis Stock/Magnum Photos.

Nowadays, the vast majority of photos are taken with smartphone cameras. Smartphone cameras avoid making exposure decisions at capture time, instead capturing high-dynamic range (HDR) light measurements (Chayka, 2022; Ernst & Wronski, 2021; Hasinoff et al., 2016). For display, a smartphone automatically converts HDR light measurements to an image using a process called tone mapping (Reinhard et al., 2010). Smartphones perform spatially varying tone mapping using deep neural networks, designed to produce aesthetically pleasing outputs. They also automatically detect different types of image elements, such as faces and clouds, and adjust each differently (Chayka, 2022; Konigsberger, 2021). These processes are directly analogous to the types of darkroom edits performed by photographers like Ansel Adams. Users may adjust color and exposure settings by adjusting sliders or selecting preset filters, but many users surely stick with the defaults. Most images we see from these phones are the product of these hidden aesthetic choices made by the camera designers and the training datasets they curated.

Night-time scenes provide a particularly vivid example. If depiction were simply about accurately reproducing light, nighttime pictures would just be dark. However, painters typically simulate physiological effects of darkness. For example, visual adaptation to darkness is simulated by bright highlights and dark shadows (chiaroscuro) or a blue-green tint (the Purkinje effect) (Figure 3). Professional film photographers use

Figure 3. Night-time paintings simulating physiological vision (a). Rather than depicting a night-time scene as dark, lit regions are depicted with bright tones, simulating visual adaptation (like pupil dilation) to the dark. *A Philosopher giving that Lecture on the Orrery in which a lamp is put in place of the Sun*, by Joseph Wright of Derby (1766). (b) An example of simulating the Purkinje effect, which causes very low-light real-world scenes to appear blue-green. *Nocturne in Blue and Gold* by James M. Whistler (1872–1875).

Figure 4. Low-light photography, tone-mapped with a baseline algorithm and with Night Sight (Liba et al., 2019), which enhances highlights and adds a blue tint.

similar effects, for example, as in Brasaï's night-time photos of Paris. Nowadays, smartphones automatically apply these techniques for low-light photos, in both the Google Pixel and iPhone. The Pixel uses a tone-mapping algorithm "designed to allow viewers of the photograph to see detail they could not have seen with their eyes, but to still know that the photograph conveys a dark scene" (Liba et al., 2019). To do so, the algorithm adds contrast and bluish tints to parts of an image (Figure 4), in a manner inspired by painting; the developers' explanations of their technique (Levoy, 2018; Liba et al., 2019) cite Hermann von Helmholtz and painters like Joseph Wright of Derby (Figure 3a).

Although painting and photography operate very differently, with different choices made in different ways, the same kinds of choices must be made in each, whether with film photography or smartphones.

Perspective choices in pictures

Geometric perspective describes the spatial arrangement and sizes of objects in an image. Most consumer and professional cameras approximate linear perspective, sometimes using very sophisticated optics. A photographer chooses the camera position, the focal length and the framing in linear perspective. Since its invention in the Renaissance, linear perspective has taken a favored place in the Western tradition, often treated as the correct perspective. Whereas cameras apply many nonlinear, spatially varying techniques to tone and color, linear perspective dominates photography to such a degree that one may not even recognize it as a choice.

Linear perspective is but one possible choice of perspective model. As eloquently pointed out by (Koenderink et al., 2016a), it has many shortcomings as either a model for how artists work or how best to depict real scenes. Linear perspective reproduces the light rays that would reach the viewer's eye, but

only if the viewer's eye is located at the image's focal center (and we ignore the effects of binocular vision, accommodation, focus, limited image resolution, limited dynamic range, and motion). Very often, we view perspective images from locations away from their focal centers, without much concern. Efforts to explain why this works have largely centered on the hypothesis that viewers subconsciously adjust for "incorrect" viewing position (Pirenne, 1970), which makes sense if human vision assumes linear perspective in pictures. Yet extensive experimental evidence shows that viewers do not adjust perspective perception based on typical pictorial cues (e.g., Cooper, Piazza, & Banks, 2012; Koenderink, Doorn, van Pepperell, & Pinna, 2016b; Todorović, 2008; Vishwanath, Girshick, & Banks 2005), even though a viewer can recognize distortions of familiar shapes.

Representational painters rarely follow linear perspective with rigid fidelity. From its invention in the Renaissance, artists observed problems with perspective, and violated it in various ways (Kemp, 1990; Kubovy, 1986); even artists with expert knowledge of linear perspective do not strictly follow it (Koenderink et al., 2016b; Pepperell & Haertel, 2014). Perhaps the most well-known problem with linear perspective is marginal distortion, where objects and faces become sheared in the periphery of wide-angle images (Figure 5a). Despite the many large-scale images of faces in art history, no classical painter appears to have sheared faces this way.

Given the limitations of linear perspective, some researchers have devised alternatives. One approach has been to describe a single family of projection formulae to be used for all images, but each of these has significant shortcomings. For example, Koenderink, Doorn, van Pepperell, and Pinna (2016a), following Helmholtz, suggest the use of stereographic projection, which removes distortion of objects in the periphery (Figure 5b) (Burleigh, Pepperell, & Ruta, 2018) propose a nonlinear perspective model designed to better capture the experience of visual space, but neither model preserves straight lines. In general, no single
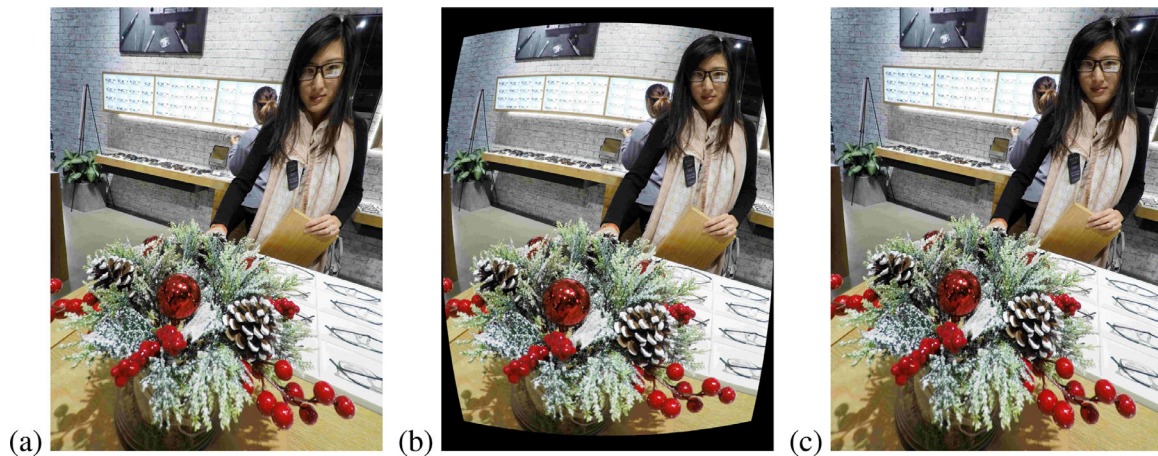
Figure 5. Comparison of photographic perspectives, from Shih et al. (2019). (a) Linear perspective. (b) Stereographic projection. (c) Automatic result from Shih et al. (2019), which uses face detection to apply stereographic projection for faces and linear projection elsewhere, thereby preserving straight lines.

two-dimensional projection of the world produces an image that satisfies all the sensible goals we might have for it.

Artists, in contrast, seem to implicitly use different projections for each image; that is, the projection function depends on the content of the image. Modern computational photography methods can now create such "content-aware" projections. These methods work by optimizing nonparametric projections, where the mapping from scene rays to image locations may be an arbitrary continuous function (Carroll, Agrawala, & Agarwala, 2009; Shih, Lai, & Liang, 2019). The optimization depends on the locations of faces, straight lines, and image texture (Figure 5c). Estimating scene depths allows even more complex projections that adjust relative object scales and locations (Badki, Gallo, Kautz, & Sen, 2017; Liu, Agrawala, DiVerdi, & Hertzmann, 2022).

As computational perspective methods mature and become available on smartphones, we may no longer think of linear perspective as the default mode for photography. This change will provide photographers techniques that have long been used by painters. However, there remain important perceptual and computational questions: how do we design new projection systems, and how do viewers perceive them?

Perception of pictures

Humans have drawn and painted pictures for at least 45,000 years, and pictures exist throughout human cultures, primarily functioning as tools for human communication and social behaviors (Dutton, 2009). Photography is a relatively recent tool for making pictures, and, throughout its history, it has both influenced and been inspired by painting (Scharf, 1968). Hence, research that truly aims to understand picture perception should not treat photography and representational painting as entirely distinct categories. Many of the same visual techniques can be used in either; both make choices about how to represent color, tone, and perspective, choosing from broadly the same palette of options.

Likewise, perception of photography should not be treated as equivalent to real-world vision. Compelling evidence shows that psychophysical and neurological responses differ for photos of objects versus for real objects (Pepperell, 2015; Snow & Culham, 2021). In important ways, viewing a photograph is more like viewing a painting than like viewing the real world, in that viewers recognize paintings and photographs as visual depictions (perhaps unconsciously) and interpret them accordingly.

Photographers—and camera manufacturers—make imaging choices driven in part by aesthetic goals around qualities like color palette and composition. One consequence of these choices are biases in image datasets, which in turn can limit the validity of research built on such datasets (Grauman et al., 2022; Pinto, Cox, & DiCarlo, 2008; Ponce et al., 2006; Torralba & Efros, 2011).

An extensive body of research has studied pictorial space for linear perspective (e.g., Cooper et al., 2012; Cutting, 2003; Vishwanath et al., 2005). Such rigor has not yet been applied to other sorts of projections, despite their advantages. The computational photography techniques surveyed in this article could provide tools for the systematic study of realistic depiction, beyond linear projection and tone. Conversely, researchers who study perception using photographs may need to understand the choices being made in their stimuli.

Discussions of pictures frequently discuss whether an image looks like the real scene (e.g., Chayka, 2022). But, given the fact that photographs cannot reproduce the light available to the viewer of a scene, what does it mean for a photograph to "look right?"

Keywords: picture perception, perspective projection, natural perspective, high-dynamic range (HDR) photography, tone mapping, computational photography

References

Albert, R., & Efros, A. A. (2016). *Post-post-modern photography: Capture-time perceptual matching for more faithful photographs* (Tech. Rep. No. EECS-2016-167). Berkeley, CA: UC Berkeley.

Badki, A., Gallo, O., Kautz, J., & Sen, P. (2017). Computational zoom: A framework for post-capture image composition. *ACM Transactions on Graphics, 36*(4).

Bersak, D. R. (2006). *Ethics in photojournalism: Past, present, and future*. Master's thesis, MIT.

Burleigh, A., Pepperell, R., & Ruta, N. (2018). Natural perspective: Mapping visual space with art and science. *Vision, 2*(2), 21.

Carroll, R., Agrawala, M., & Agarwala, A. (2009). Optimizing content-preserving projections for wide-angle images. *ACM Transactions on Graphics, 28*(3).

Chayka, K. (2022). Have iphone cameras become too smart? *The New Yorker*.

Cooper, E. A., Piazza, E. A., & Banks, M. S. (2012). The perceptual basis of common photographic practice. *Journal of Vision, 12*(5), 8–8.

Costello, D. (2017). *On photography: A philosophical inquiry*. New York: Routledge.

Cutting, J. E. (2003). Reconceiving perceptual space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space*. Cambridge, MA: MIT Press.

Debevec, P. E., & Malik, J. (2008). Recovering high dynamic range radiance maps from photographs. In *Proceedings of SIGGRAPH* (pp. 1–10).

Durand, F. (2007). An invitation to discuss computer depiction. In *Proceedings of NPAR*.

Dutton, D. (2009). *The art instinct: Beauty, pleasure, and human evolution*. London: Bloomsbury Press.

Ernst, M., & Wronski, B. (2021). Hdr+ with bracketing on pixel phones. Available at: https://ai.googleblog.com/2021/04/hdr-with-bracketing-on-pixel-phones.html. Accessed October 11, 2022.

Gibson, J. J. (1971). The information available in pictures. *Leonardo, 4*(1), 27–35.

Grauman, K., Westbury, A., Byrne, E., Chavis, Z., Furnari, A., Girdhar, R., . . . Malik, J. (2022). Ego4d: Around the world in 3,000 hours of egocentric video. In *Proceedings of IEEE CVPR*.

Hasinoff, S. W., Sharlet, D., Geiss, R., Adams, A., Barron, J. T., & Kainz, F. et al. (2016). Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics, 35*(6).

Hawley, R. E. (2022). BeReal and the fantasy of an authentic online life. *The New Yorker*.

Kemp, M. (1990). *The science of art: Optical themes in western art from brunelleschi to seurat*. New Haven: Yale University Press.

Koenderink, J., van Doorn, A., Pepperell, R., & Pinna, B. (2016a). On right and wrong drawings. *Art & Perception, 4*, 1–38.

Koenderink, J., Doorn, A. van, Pinna, B., & Pepperell, R. (2016b). Facing the spectator. *i-Perception, 7*(6), 2041669516675181.

Konigsberger, F. (2021). Image equity: Making image tools more fair for everyone. Available at: https://blog.google/products/pixel/image-equity-real-tone-pixel-6-photos/. Accessed October 11, 2022.

Kubovy, M. (1986). *The psychology of perspective and renaissance art*. Cambridge UK: Cambridge University Press.

Levoy, M. (2018). Night sight: Seeing in the dark on pixel phones. Available at: https://ai.googleblog.com/2018/11/night-sight-seeing-in-dark-on-pixel.html. Accessed October 11, 2022.

Liba, O., Murthy, K., Tsai, Y.-T., Brooks, T., Xue, T., Karnad, N., . . . Levoy, M. (2019). Handheld mobile photography in very low light. *ACM Transactions on Graphics (TOG), 38*(6), 1–16.

Liu, S. J., Agrawala, M., DiVerdi, S., & Hertzmann, A. (2022). Zoomshop: Depth-aware editing of

photographic composition. *Computer Graphics Forum, 41*(2), 57–70.

MacAdam, D. (1951). Quality of color reproduction. *Journal of the SMPTE, 39*(5), 468–485.

Morris, E. (2014). *Believing is seeing (observations on the mysteries of photography)*. New York: Penguin Books.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology, 64*, 77–107.

Pepperell, R. (2015). Artworks as dichotomous objects: Implications for the scientific study of aesthetic experience. *Frontiers in Human Neuroscience, 9*.

Pepperell, R., & Haertel, M. (2014). Do artists use linear perspective to depict visual space? *Perception, 43*(5), 395–416.

Pinto, N., Cox, D. D., & DiCarlo, J. J. (2008). Why is real-world visual object recognition hard? *PLoS Computational Biology, 4*(1).

Pirenne, M. H. (1970). *Optics, painting & photography*. Cambridge, UK: Cambridge University Press.

Ponce, J., & *et al.* (2006). Dataset issues in object recognition. In *Toward category-level object recognition*. Berlin, Heidelberg: Springer Berlin Heidelberg (pp. 29–48).

Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., & Myszkowski, K. (2010). *High dynamic range imaging: Acquisition, display, and image-based lighting* (2nd ed.). Burlington, MA: Morgan Kaufmann.

Roth, L. (2009). Looking at shirley, the ultimate norm: Colour balance, image technologies, and cognitive equity. *Canadian Journal of Communication, 34*(1), 111–136.

Scharf, A. (1968). *Art and photography*. New York: Penguin.

Shih, Y., Lai, W.-S., & Liang, C.-K. (2019). Distortion-free wide-angle portraits on camera phones. *ACM Transactions on Graphics, 38*(4).

Snow, J. C., & Culham, J. C. (2021). The treachery of images: How realism influences brain and behavior. *Trends in Cognitive Sciences, 25*(6), 506–519.

Todorović, D. (2008). Is pictorial perception robust? the effect of the observer vantage point on the perceived depth structure of linear-perspective images. *Perception, 37*, 106–125.

Torralba, A., & Efros, A. A. (2011). Unbiased look at dataset bias. In *Proceedings of IEEE CVPR*.

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature neuroscience, 8*(10), 1401–1410.

Yang, T., & Kubovy, M. (1999). Weakining the robustness of perspective: Evidence for a modified theory of compensation in picture perception. *Perception & Psychophysics, 61*(3), 456–467.