# sketcher: An R package for converting a photo into a sketch style image

#### Hiroyuki Tsuda

Keio Advanced Research Center, Keio University, Japan.



Figure 1. By using the sketcher package, you can convert a photo into a line drawing image. Drawing style (lineweight, texture smoothness, and the inclusion of shading) can be controlled.

*Abstract*—Line drawings have been useful stimuli for visual perception, cognition, and related fields, because edges play an important role in the visual system. Although there are many datasets of line drawing stimuli available for research, we often need a new set of stimuli for specific research purposes. Here I provide an R implementation of image processing effects where a photo is converted to a line drawing image. Drawing style, such as lineweight, texture smoothness, and the inclusion of shading can be controlled by adjusting the parameters of this algorithm. With this tool, line drawing stimuli can easily be created with just a few lines of script, and will be useful for various research applications.

**Keywords**: line drawing, edge detection, R package **Correspondence**: tsuda16k@gmail.com

#### Introduction

From the prehistoric cave art to charts and diagrams in scientific communication, line drawings are ubiquitous in the human culture. Although line drawings are not a feature of the natural world (Gibson, 1951), which lack many of the defining characteristics of objects (e.g., color, texture, shading), we nevertheless easily see shapes, surfaces, and spatial structures in line drawings. In fact, line drawings of visual scenes can be recognized as fast and accurately as photographs (Biederman and Ju, 1988). The depiction of objects and scenes in line drawings can be recognized by infants (Yonas and Arterberry, 1994), stone-age tribe members who have no pictorial art (Kennedy and Ross, 1975), a child without any experience

in picture-name associations since birth (Hochberg, 1962), and even chimpanzees (Itakura, 1994; Tanaka, 2007). These observations suggest that the ability of humans to recognize line drawings is not a literacy acquired through culture and convention (Goodman, 1976), but reflects the fundamental nature of how the biological vision system evolved to represent the visual world (Cavanagh, 2005; Hertzmann, 2020; Sayim and Cavanagh, 2011). Consequently, in psychology and neuroscience, line drawing stimuli have been a useful tool for studying the mechanisms of visual perception and cognition (Biederman and Ju, 1988; Hayes, 1988; Morgan, Petro, and Muckli, 2019; Snodgrass and Vanderwart, 1980; Walther et al., 2011).

There are many datasets of line drawing stimuli available for research (Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Brodeur, Dionne-Dostie, Montreuil1, & Lepage, 2010; Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; de Beeck and Wagemans, 2001; Nishimoto, Miyawaki, Ueda, Une, & Takahashi, 2005; Rossion and Pourtois, 2004; Saryazdi, Bannon, Rodrigues, Klammer, & Chambers, 2018; Snodgrass and Vanderwart, 1980). However, we often need a new set of stimuli for specific research purposes. In such cases, line drawing stimulus is usually created by turning a photo into a line drawing. This procedure can be performed by hand (tracing lines in a photo) or semi-automatically by using photo-editing softwares such as Photoshop and GIMP. The conversion process can be performed fully automatically by using image processing (computer vision) techniques. Image processing algorithms are typically implemented in



Figure 2. The procedure to photo-to-sketch conversion.

(a) Input image is first converted to grayscale, and (b) edge-preserving smoothing filter is applied (iteratively, if required), and then (c1) Gaussian blur or (c2) maximum filter is applied to the smoothed image from step b, and then (d1) image b is divided by image c1 or (d2) image b is divided by image c2, and finally (e1, e2) contrast enhancement is applied.

C, C++, or Python. Here, I provide an R implementation of photo-to-sketch image processing effects. I selected R language because many psychologists (including me) are familiar with R and therefore it is easy to use.

#### Method

There are many standard methods to detect edges in an image, such as Prewitt filter, Sobel filter, and Canny edge detector (for a review, see Russ, 2011). More sophisticated (artistic or stylized) sketch generation algorithms have also been proposed by using differenceof-Gaussians operator (Winnemöller, Kyprianidis, & Olsen, 2012) or deep neural networks (Yi, Liu, Lai, & Rosin, 2019). Standard edge detectors are suited for computer vision tasks, but the results of those detectors are usually look unnatural to the human eye. Elaborated edge detection and stylization algorithms have aesthetic appeal, but some distortions of edges and shapes can be introduced.

Here I employed a common procedure to achieve a pencil sketch effect, which can also be conducted easily by using photo editors such as Photoshop and GIMP (e.g., "Photo To Sketch", 2012; "Portrait Photo", 2012). This method is fairly simple but effectively produces a natural and aesthetically pleasing output that preserves original edge structure. The procedure is as follows: 1) Convert a photo to grayscale, and 2) apply a blur filter (Gaussian/maximum), then 3) divide the grayscale image (from step 1) by the blurred image (from step 2), which extracts edges in the original photo, and finally 4) adjust image contrast, if necessary. The choice of the blurring filter (Gaussian/maximum) affects the level of detail preserved

in the output image ("Photo To Sketch", 2012). In addition, the lineweight (strength of lines) within drawings can be altered by adjusting the kernel size of Gaussian/ maximum filter.

This procedure, however, is very sensitive to noise in the input image. Noise can be eliminated by pre-filtering the input image. The Gaussian filter and the median filter are commonly used for this purpose, but these filters can also degrade fine edges, which is problematic for producing sketch style images. Here I employed the guided filter proposed by He and colleagues (He, Sun, & Tang, 2013). This filter is a type of edge-preserving smoothing filter, which can serve as an effective preprocessing operator for edge extraction. In addition, by iteratively applying the guided filter with increasing kernel size, texture and shading can be eliminated, while keeping edges (e.g., Boyadzhiev, Bala, Paris, & Adelson, 2015). Here I applied this technique to control the level of detail (texture/shading) in the output sketch image. In figure 2, I summarized the photo-to-sketch conversion procedure employed in this study.

#### Results

Figure 3 shows how the choice of the blurring filter (Gaussian/maximum) affects the appearance of conversion results. While edges and textures are extracted in a similar manner by both procedures, shading is more prominent when the maximum filter is used. This is because smooth local gradients are emphasized by the use of maximum filter, while abrupt changes in gradient (edges) are mostly extracted if blurring is performed with the Gaussian smoothing filter.



**Figure 3**. The effect of blur filter choice on the sketch output. (a) Input image. (b) Gaussian filter. (c) maximum filter. Note that image b is the image at the center of Figure 4 (left), and image c is the center image of Figure 4 (right).



**Figure 4**. The effects of lineweight (kernel size of blur filter) and smoothing (the number of iteration of guided filter) on the resulting sketch. Images on the left were obtained with Gaussian filter, and images on the right were obtained with maximum filter.

Figure 4 shows the effects of the kernel size of blur filter (labeled as lineweight) and the number of times for pre-smoothing operation (labeled as smoothing) on the output sketch. If the lineweight value is L, then the kernel size of blur filter (Gaussian/maximum) is (2L + 1, 2L + 1). If smoothing value is zero, no pre-smoothing is applied. If smoothing value is K, then smoothing output  $I_k = f_{gf}(I_{k-1}, 2^k + 1)$ , where  $f_{gf}(I, D)$  means applying the guided filter to image I with kernel size of D, and  $I_0$  = the input image.

The results in Figure 4 demonstrate that the kernel

size of blur filter can be used to control the lineweight of the sketch. Increase in the kernel size also emphasizes shading, especially when the maximum filter was used (Figure 4, right). Applying the pre-smoothing operation significantly eliminates shading and fine textures (e.g., stitches of the hat). When smoothing parameter is small (1 or 2), only particle noise and fine texture are eliminated. If it is large, shading is also eliminated (Figure 4). This feature may be useful for studying the effect of shading on object recognition (Johnston, Hill, & Carman, 1992; Liu,



Figure 5. Sketch examples for images of objects and scenes.

Collin, & Chaudhuri, 2000; Ramachandran, 1998; Rossion and Pourtois, 2004).

## R package

The package is available on CRAN. It can be installed with: install.packages("sketcher"). A detailed introduction and usage of the package is described on my GitHub page: <u>https://github.com/tsuda16k/sketcher</u> and my webpage: <u>https://htsuda.net/sketcher</u>

### Discussion

In this study I implemented the photo-to-sketch image processing effect in R. The various parameters of the current procedure (blur kernel type, blur kernel size, and pre-smoothing operation) may be adjusted to allow for a variety of sketch appearances (Figure 1). I found that visually pleasing result can be obtained for most input images (Figure 5 shows examples for images of objects and scenes). Such drawing stimuli may be useful for various research applications.

#### References

- Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20(1), 38–64. https:// doi.org/10.1016/0010-0285(88)90024-2
- Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers, 35*(1), 158–167. <u>https://doi.org/10.3758/bf03195507</u>
- Boyadzhiev, I., Bala, K., Paris, S., & Adelson, E. (2015). Band-Sifting Decomposition for Image-Based Material Editing. ACM Transactions on Graphics, 34(5), 1–16. <u>https://doi.org/</u> 10.1145/2809796
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. *PLoS ONE*, 5(5), e10773. <u>https://doi.org/ 10.1371/journal.pone.0010773</u>

- Cavanagh, P. (2005). The artist as neuroscientist. *Nature*, 434(7031), 301–307. https://doi.org/10.1038/434301a
- Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture Naming by Young Children: Norms for Name Agreement, Familiarity, and Visual Complexity. *Journal of Experimental Child Psychology*, 65(2), 171–237. <u>https://doi.org/ 10.1006/jecp.1996.2356</u>
- de Beeck, H. O., & Wagemans, J. (2001). Visual Object Categorisation at Distinct Levels of Abstraction: A New Stimulus Set. *Perception*, 30(11), 1337–1361. <u>https://doi.org/10.1068/p3120</u>
- Gibson, J. J. (1951). What is a form? *Psychological Review*, 58(6), 403–412. <u>https://doi.org/10.1037/h0055232</u>
- Goodman, N. (1976). Languages of Art: An Approach to a Theory of Symbols, 2nd Edition. Indianapolis: Hackett Publishing Company.
- Hayes, A. (1988). Identification of Two-Tone Images; Some Implications for High- and Low-Spatial-Frequency Processes in Human Vision. *Perception*, 17(4), 429–436. <u>https://doi.org/10.1068/ p170429</u>
- He, K., Sun, J., & Tang, X. (2013). Guided Image Filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6), 1397–1409. https://doi.org/10.1109/tpami.2012.213
- Hertzmann, A. (2020). Why Do Line Drawings Work? A Realism Hypothesis. *Perception*, 49(4), 439–451. <u>https://doi.org/ 10.1177/0301006620908207</u>
- Hochberg, J., & Brooks, V. (1962). Pictorial Recognition as an Unlearned Ability: A Study of One Child's Performance. *The American Journal of Psychology*, 75(4), 624. <u>https://doi.org/</u> 10.2307/1420286
- Itakura, S. (1994). Recognition of Line-Drawing Representations by a Chimpanzee (Pan troglodytes). *The Journal of General Psychology*, 121(3), 189–197. https://doi.org/10.1080/00221309.1994.9921195
- Johnston, A., Hill, H., & Carman, N. (1992). Recognising Faces: Effects of Lighting Direction, Inversion, and Brightness Reversal. *Perception*, 21(3), 365–375. <u>https://doi.org/10.1068/p210365</u>
- Kennedy, J. M., & Ross, A. S. (1975). Outline Picture Perception by the Songe of Papua. *Perception*, 4(4), 391–406. <u>https://doi.org/10.1068/ p040391</u>
- Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does Face Recognition Rely on Encoding of 3-D Surface? Examining the Role of Shape-from-Shading and Shape-from-Stereo. *Perception*, 29(6), 729–743. <u>https://doi.org/10.1068/p3065</u>
- Morgan, A. T., Petro, L. S., & Muckli, L. (2019). Scene Representations Conveyed by Cortical Feedback to Early Visual Cortex Can Be Described by Line Drawings. *The Journal of Neuroscience*, 39(47), 9410–9423. https://doi.org/10.1523/

#### jneurosci.0852-19.2019

- Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods*, 37(3), 398–416. <u>https://doi.org/10.3758/bf03192709</u>
- Photo To Sketch With More Detail In Photoshop. (2012, October 02). Retrieved from <u>https://www.photoshopessentials.com/photo-effects/photo-to-sketch/</u>
- Portrait Photo To Color Sketch With Photoshop. (2012, October 02). Retrieved from <u>https://www.photoshopessentials.com/photo-effects/portrait-to-sketch/</u>
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature,* 331(6152), 163–166. https://doi.org/10.1038/331163a0
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's Object Pictorial Set: The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, 33(2), 217– 236. https://doi.org/10.1068/p5117
- Russ, J. (2011). *The Image Processing Handbook*. Sixth Edition, CRC Press.
- Saryazdi, R., Bannon, J., Rodrigues, A., Klammer, C., & Chambers, C. G. (2018). Picture perfect: A stimulus set of 225 pairs of matched clipart and photographic images normed by Mechanical Turk and laboratory participants. *Behavior Research Methods*, 50(6), 2498– 2510. https://doi.org/10.3758/s13428-018-1028-5
- Sayim, B., & Cavanagh, P. (2011). What Line Drawings Reveal About the Visual Brain. Frontiers in Human Neuroscience, 5, 118. https:// doi.org/10.3389/fnhum.2011.00118
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. <u>https://doi.org/ 10.1037/0278-7393.6.2.174</u>
- Tanaka, M. (2006). Recognition of pictorial representations by chimpanzees (Pan troglodytes). *Animal Cognition*, 10(2), 169– 179. <u>https://doi.org/10.1007/s10071-006-0056-1</u>
- Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy* of Sciences, 108(23), 9661–9666. <u>https://doi.org/10.1073/pnas.</u> 1015666108
- Winnemöller, H., Kyprianidis, J. E., & Olsen, S. C. (2012). XDoG: An eXtended difference-of-Gaussians compendium including advanced image stylization. *Computers & Graphics*, 36(6), 740–753. <u>https:// doi.org/10.1016/j.cag.2012.03.004</u>
- Yi, R., Liu, Y.-J., Lai, Y.-K., & Rosin, P. L. (2019). APDrawingGAN: Generating Artistic Portrait Drawings From Face Photos With Hierarchical GANs. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). <u>https://doi.org/10.1109/ cvpr.2019.01100</u>
- Yonas, A., & Arterberry, M. E. (1994). Infants Perceive Spatial Structure Specified by Line Junctions. *Perception*, 23(12), 1427– 1435. <u>https://doi.org/10.1068/p231427</u>