SmartShadow: Artistic Shadow Drawing Tool for Line Drawings

LVMIN ZHANG, Soochow University / Style2Paints Research
JINYUE JIANG, Style2Paints Research
YI JI, Soochow University
CHUNPING LIU, Soochow University

SmartShadow is a deep learning application for digital painting artists to draw shadows on line drawings, with three proposed tools. (1) Shadow brush: artists can draw scribbles to coarsely indicate the areas inside or outside their wanted shadows, and the application will generate the shadows in real-time. (2) Shadow boundary brush: this brush can precisely control the boundary of any specific shadow. (3) Global shadow generator: this tool can estimate the global shadow direction from input brush scribbles, and then consistently propagate local shadows to the entire image. These three tools can not only speed up the shadow drawing process (by 3.1× as experiments validate), but also allow for the flexibility to achieve various shadow effects and facilitate richer artistic creations. To this end, we train Convolutional Neural Networks (CNNs) with a collected large-scale dataset of both real and synthesized data, and especially, we collect 1670 shadow samples drawn by real artists. Both qualitative analysis and user study show that our approach can generate high-quality shadows that are practically usable in the daily works of digital painting artists. We present 30 additional results and 15 visual comparisons in the supplementary materiel. Finally, the dataset can also be used in related applications to further facilitate artistic creations.

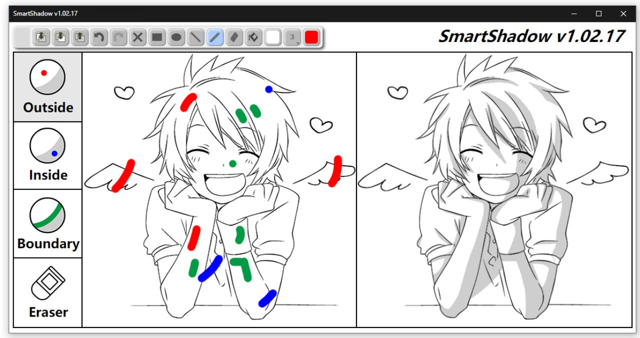


Fig. 1. **Screenshot of the SmartShadow.** The user gives scribbles as shadow indications (on the left) to obtain the high-quality shadow (on the right). *Smiling boy, used with artist permission.*

1 INTRODUCTION

"The purpose of drawing shadow is to show how we understand and feel about the objects and people."

— *The Art of Comic Book Drawing* [Aaseng et al. 2005]

Shadows in artworks are essentially different from that in photography or photorealistic fields of computer vision: the artwork shadows are *drawn* by artists. These shadows depicts the mood of characters and express the emotion of artists, without being constrained by physically correct light transmission laws or geometrically precise object structures. Artists adjust the location, scale, shape, density, and many other features of shadows to achieve diverse artistic purposes, *e.g.*, amplification, exaggeration, antithesis, silhouette, *etc.*

An application that can assist artists in drawing shadows for line drawings is highly desired. This is not only because creating shadows on line drawings is one of the most frequent and time-consuming tasks in the daily work of many digital painting artists, but also because shadow drawing is the foundation of a wide variety of further artistic creations, *e.g.*, hard shadows can be smoothed into soft shadings (with techniques like joint anisotropic diffusion [Weickert 1998]), shadows can be stylized with hatching or drafting effects [Zheng et al. 2020], sharp shadows can be used in cel-shading (see also the YouTube tutorial [Maga 2018]), *etc.*

Might we be able to achieve a deep learning approach that can quickly produce visually satisfying shadows given only a few user indications, saving the time and effort of digital painting artists, and simultaneously, facilitating more plentiful artistic creations? We present an interactive shadow drawing application (Fig. 1) to achieve these goals. This application consists of the following three proposed tools:

The first tool is the shadow brush. Users can draw blue or red scribbles (*e.g.*, Fig. 2-(a)) to coarsely indicate the areas inside or outside the shadows they want. This tool does not require users to have professional drawing skills, as it can "smartly" generate shadow shapes learned from large-scale artistic shadow data. This tool is well-suited for shadows without strict shape requirements or with low shape uncertainty, *e.g.*, inconspicuous background shadow, dense shadow of gathered small objects, *etc.*

The second tool is the shadow boundary brush. Users can use this brush to precisely control the shadow boundaries. They only need to scribble a small part of their wanted boundary (*e.g.*, the green scribbles in Fig. 2-(b)), and the tool will automatically estimate the boundary shape and generate the entire shadow. This tool is indispensable for professional use cases where the accurate shadow control is important, *e.g.*, character face shadows, salient object shadows, close-up shadows, *etc.*

The third tool is the global shadow generator. This tool can estimate the global shadow direction from input brush scribbles, and then propagate local shadows to the entire image consistently (*e.g.*, Fig. 2-(c)). This tool is user-friendly in that it is fully automatic and does not require artists to learn any extra technical knowledges, *e.g.*, managing screen-space shadow direction, world-space light orientation, *etc.* This tool is especially effective for complicated artworks, *e.g.*, drawings with multiple targets, artworks with complex structure, *etc.*

These three tools are designed in a data-driven way. To ensure the robustness and generalization, we learn hierarchical neural

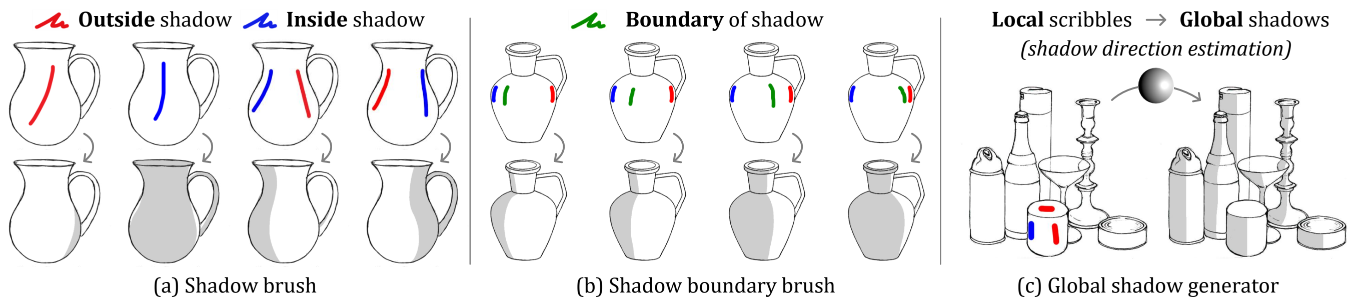


Fig. 2. **Objectives of our three proposed tools.** (a) The shadow brush allows users to coarsely control the areas inside or outside shadows. (b) The shadow boundary brush enables users to accurately control the shadow shapes. (c) The global shadow generator can estimate the global shadow direction and automatically produce globally consistent shadows. *Artworks used with artist permissions.*

networks with a large-scale dataset of both real-artist data and synthesized data. In particular, we collect 1670 line art and shadow pairs drawn by artists manually, 25,413 pairs synthesized by rendering engine, and 291,951 shadow pairs extracted from in-the-wild internet digital paintings. This dataset can also be used in related applications to further facilitate richer artistic creations.

Experiments show that the SmartShadow can speed up the shadow drawing process by 3.1×. User studies demonstrate that users can use this application to effectively achieve satisfactory shadows that are practically usable in their daily jobs. Besides, even if the users do not give any input edits, our approach can still generate plausible results that are more preferable than other fully-automatic shadow generating methods. Finally, we present 30 qualitative results and 15 additional comparisons in the supplementary materiel.

In summary, our contributions are: (1) We present the SmartShadow, a digital painting application to draw shadows on line drawings, including the tools of shadow brush, shadow boundary brush, and global shadow generator. (2) We collect a large-scale dataset of line drawing and shadow pairs drawn by real artists, as well as shadow data synthesized by rendering engines or extracted from in-the-wild digital paintings. (3) Perceptual user study and qualitative evaluations demonstrate that the SmartShadow is more preferable by actual end users when compared to other possible alternatives. (4) Results show that the SmartShadow can speed up the shadow drawing process by 3.1×.

2 RELATED WORK

Artistic shadow creation. Different from photography relighting or photorealistic rendering [Chen et al. 2010; Debevec et al. 2000; Matusik et al. 2004; Peers and Dutre 2005; Peers et al. 2009, 2007], the artistic creation of shadows is a perception-oriented process. ShadeSketch [Zheng et al. 2020] is the current state of the art in automatic artistic shadow generating. Sketch2Normal [Su et al. 2018] and DeepNormal [Hudon et al. 2018a] can generate normal maps from line drawings. Hudon *et al.* [Hudon et al. 2018b] also proposed a vectorgraph-based method for artistic shadow manipulation. Ink-and-Ray [Sykora et al. 2014] is a typical proxy-based method for illumination effects, and Dvorožňák *et al.* [Dvorožňák et al. 2018] extended this approach to a part-based high-relief proxy structure. PaintingLight [Zhang et al. 2020b] is a RGB geometry framework that converts artists' brush stroke history to lighting effects. Our approach allows users to intuitively manipulate the shadow with scribbles, *i.e.*, in a "what you see is what you get" manner.

Shadow synthesis and extraction. To ensure the robustness and generalization of our approach, we use shadow synthesis and extraction algorithms to increase the scale and diversity of our training data. A typical method is intrinsic imaging [Barrow and Tenenbaum 1978] in the field computational illumination. Optimizing-based approaches [Shen et al. 2011] solve the decomposition by optimizing an energy with specific constraints. Learning-based approaches [Barron and Malik 2012; Gehler et al. 2011; Serra et al. 2012] propose to learn the mapping between the input images and their albedo images from large amounts of data. Several in-the-wild datasets [Bell et al. 2014, 2013, 2015; Kovacs et al. 2017] and other synthetic or annotated datasets [Beigpour et al. 2013; Grosse et al. 2009] make intrinsic images scalable with deep learning methods.

Interactive creation and cartoon techniques. Scribble-based interactive tools are shown to be effective in creative fields like image colorization [Zhang et al. 2017] and sketch inking [Simo-Serra et al. 2018b]. Another closely related field is cartoon image processing. Manga structure extraction [Li et al. 2017], cartoon inking [Simo-Serra et al. 2018a,b, 2016], and line closure [Liu et al. 2018, 2015] methods analysis the lines in cartoon and digital paintings. A region-based composition method can be used in cartoon image animating [Sýkora et al. 2005]. Deep learning approaches [Chen et al. 2018; Wang and Yu 2020; Yi et al. 2019, 2020a,b] process artistic images or cartoon drawings in the domains of photographs and human portraits. Color filling applications [Sykora et al. 2009; TaiZan 2016; Zhang et al. 2018] colorize sketch or line drawings with optimization-based or learning-based approaches. Our approach generates shadow from line drawings, and can be used in digital painting and related artistic creation scenarios.

3 METHOD

We train a deep network to draw shadows given the line drawings and user input scribbles. In Section 3.1, we describe the objective of the neural architecture and the three proposed interactive tools: shadow brush, shadow boundary brush, and global shadow generator. We then describe our presented dataset and the customized training method in Section 3.2.

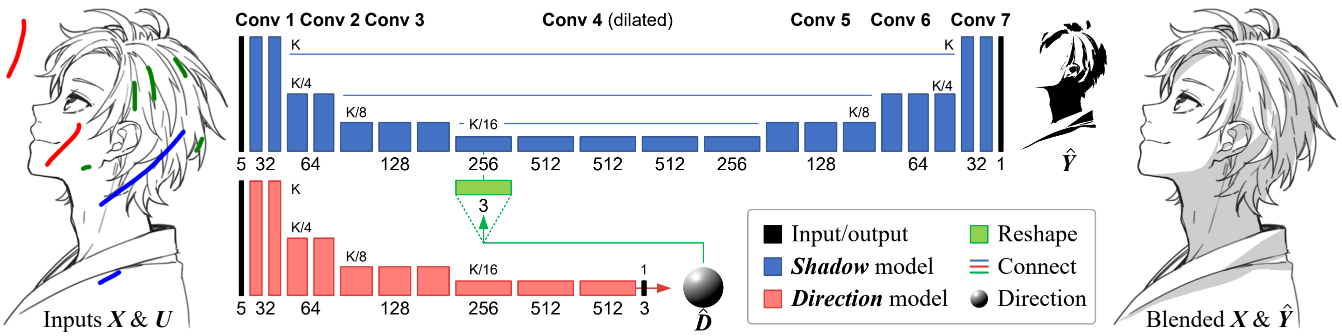


Fig. 3. **Network architecture.** We train two branches of the shadow drawing network. Both branches use the blue layers for predicting the shadow. The **direction model** branch uses red layers for predicting the global shadow direction. The **shadow model** branch uses blue layers for predicting the final output shadow. All convolutional layers use 3×3px kernels. We do not use any normalization layers. Shortcut connections are added to upsampling convolution layers. *Boy looking upside, used with artist permission.*

3.1 Interactive tools for shadow drawing

The inputs (Fig. 3-left) of our approach are the line drawing $\boldsymbol{X} \in \mathbb{R}^{H \times W \times 1}$ along with the RGBA user scribble canvas denoted by $\boldsymbol{U}^{H \times W \times 4}$. The output $\widehat{\boldsymbol{Y}} \in \mathbb{R}^{H \times W \times 1}$ is the estimation of pixel-wise shadow probability, which is binarized (threshold is 50% gray) and blended (multiplied) to the original line drawing for shadow effects (Fig. 3-right). The mapping is learned with the neural networks $\mathcal{F}(\cdot;\theta)$, parameterized by θ, with the architecture specified in Fig. 3. We train with the data distribution $\mathcal{D}$ with line arts, user inputs, and desired shadows. We minimize the objective with likelihood $\mathcal{L}$ describing the distances between the estimation and ground truth as

$$\theta^* = \arg\min_{\theta} \mathbb{E}_{\boldsymbol{X},\boldsymbol{U},\boldsymbol{Y}\sim\mathcal{D}}[\mathcal{L}(\mathcal{F}(\boldsymbol{X},\boldsymbol{U};\theta),\boldsymbol{Y})] \,. \tag{1}$$

We learn two network branches: the shadow model $\mathcal{F}_s(\cdot;\theta_s)$ and the shadow direction model $\mathcal{F}_d(\cdot;\theta_d)$. In inference, the direction model estimates the global shadow direction $\widehat{\boldsymbol{D}} \in \mathbb{R}^3$ for the shadow model to predict the shadow with

$$\widehat{\boldsymbol{Y}} = \mathcal{F}_s(\boldsymbol{X},\boldsymbol{U},\widehat{\boldsymbol{D}};\theta_s) \quad \text{and} \quad \widehat{\boldsymbol{D}} = \mathcal{F}_d(\boldsymbol{X},\boldsymbol{U};\theta_d) \,. \tag{2}$$

During training, the scribbles are synthesized for our tools by giving projections of the ground truth shadow $\boldsymbol{Y}$ with the projection function $\mathcal{P}_u$ as $\boldsymbol{U} = \mathcal{P}_u(\boldsymbol{Y})$. Because the training synthetically generates user inputs, our dataset only needs to contain line drawings, shadow directions, and our wanted shadows. In particular, we solve two sub-problems for the shadow model and shadow direction model with

$$\begin{aligned}\theta_d^* &= \arg\min_{\theta_d} \mathbb{E}_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{D}\sim\mathcal{D}}[\mathcal{L}_d(\mathcal{F}_d(\boldsymbol{X},\boldsymbol{U};\theta_d),\boldsymbol{D})] \,,\\ \theta_s^* &= \arg\min_{\theta_s} \mathbb{E}_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{D}\sim\mathcal{D}}[\mathcal{L}(\mathcal{F}_s(\boldsymbol{X},\boldsymbol{U},\boldsymbol{D};\theta_s),\boldsymbol{Y})] \,,\end{aligned} \tag{3}$$

where $\mathcal{L}_d$ is a likelihood function for the shadow direction estimation problem. The three proposed shadow drawing tools are detailed as follows.

Shadow brush. The shadow control is achieved by projecting $\mathcal{P}_u$ to sample pixels inside (*resp.*, outside) the ground truth shadows in $\boldsymbol{Y}$ as blue (*resp.*, red) scribbles. We observe that, unlike common pixel sampling problems (*e.g.*, [Sangkloy et al. 2017; Zhang et al. 2018, 2017]) where pixels are routinely distributed and sampled uniformly, shadow images are unique in their unbalanced pixel quantity inside and outside shadows. Based on this observation, we propose to balance the pixel sampling by introducing a Bivariate Normal Distribution (BND), with a Probability Density Function (PDF) denoted by $f_b(\cdot,\cdot)$. We sample n_i pixels inside the shadows and n_o pixels outsides, subjecting to the Bivariate Normal PDF [Wikipedia 2020] as

$$f_b(n_i,n_o) = \frac{\exp(-\frac{1}{2(1-\rho^2)}p_b(n_i,n_o))}{2\pi\sigma_i\sigma_o\sqrt{1-\rho^2}} \,, \tag{4}$$

where $p_b(\cdot,\cdot)$ is a bivariate Gaussian normal term

$$p_b(n_i,n_o) = (\frac{n_i-\mu_i}{\sigma_i})^2 - 2\rho(\frac{n_i-\mu_i}{\sigma_i})(\frac{n_o-\mu_o}{\sigma_o}) + (\frac{n_o-\mu_o}{\sigma_o})^2, \tag{5}$$

where $\{\mu_i,\mu_o,\sigma_i,\sigma_o,\rho\}$ are bivariate normal distribution values with $8, 8, 2, 2, 0.5$. Using these sampled pixels as starting positions, we synthesize small scribbles with line segments at random rotation $\theta \sim U(-\pi,\pi)$, length $l \sim U(5,15)$ pixels, and width $w \sim U(1,3)$ pixels.

Shadow boundary brush. The accurate shadow boundary control is achieved by projecting $\mathcal{P}_u$ to sample shadow edges in the ground truth $\boldsymbol{Y}$ as green scribbles. We randomly sample $n_b \sim U(0,16)$ pixels of these edges as scribble starting points, and then synthesize small solid circles at random radius of $r \sim U(5,15)$ pixels. Besides, we observe that an important characteristic of shadows drawn by artists is the smooth boundaries and sharp corners. We encourage such smoothness and sharpness by introducing an anisotropic penalty $\phi(\cdot)$ within the customized likelihood

$$\mathcal{L}(\widehat{\mathrm{Y}},\mathrm{Y}) = \lambda_a\phi(\widehat{\mathrm{Y}}) + \sum_p ||\widehat{\mathrm{Y}}_p - \mathrm{Y}_p||_2^2 \,, \tag{6}$$

where p is pixel position, $||\cdot||_2$ is Euclidean distance, λ_a is weighting parameter, and the penalty $\phi(\cdot)$ can be written as

$$\phi(\widehat{\mathrm{Y}}) = \sum_p \sum_{i\in w(p)} \sum_{j\in w(p)} \left(\delta(\boldsymbol{X})_{ij}||\widehat{\mathrm{Y}}_i - \widehat{\mathrm{Y}}_j||_2^2\right) \,, \tag{7}$$

where $w(p)$ is a 3×3 window centered at pixel position p, with $\delta(\cdot)$ being a Gaussian anisotropic term

$$\delta(\boldsymbol{X})_{ij} = \exp(-||\boldsymbol{X}_i - \boldsymbol{X}_j||_2^2/\kappa^2) \,, \tag{8}$$

where κ is an anisotropic weight. This term increases and encourages smoothness when $w(p)$ is located inside shadow areas with no steep line transitions in the line drawing $\boldsymbol{X}$, while decreases and allows for sharpness when $w(p)$ comes across salient line drawing patterns like corners or contours.

Global shadow generator. The global shadow generating is guided by the shadow direction $\boldsymbol{D} = [\alpha_x \ \ \alpha_y \ \ \alpha_z]^\intercal$ with α_x and α_y being in line with the axes of image-space width (right is positive) and height (upward is positive), and α_z facing out of the image panel. We use a customized likelihood for this global shadow direction as

$$\mathcal{L}_d(\widehat{\boldsymbol{D}}, \boldsymbol{D}) = \sum_p (\underbrace{-\widehat{\boldsymbol{D}}_p * \boldsymbol{D}_p}_{\text{cos}} + \underbrace{\lambda_n || \widehat{\boldsymbol{D}}_p - \frac{\widehat{\boldsymbol{D}}_p}{||\widehat{\boldsymbol{D}}_p||_2} ||_2^2}_{\text{norm}}), \tag{9}$$

where $*$ is dot product and λ_n is a penalizing weight. The "cos" term is a cosine likelihood between the estimated direction and the ground truth, and the "norm" term is a regulation to encourage the confidence — low-intensity estimation will be amplified to a norm unit scale. Note that (1) this tool is only a coarse recommendation of the shadow propagation, and more specified effects (*e.g.*, spot light, rim shadow, *etc.*) can be achieved with the other brush tools; and (2) this tool is fully automatic and does not require artists to learn any technical knowledges, *e.g.*, data structure for 3D space orientation, screen-to-world space conversion, *etc.*

3.2 Data preparation and training schedule

Ideally, we may invite professional artists to manually draw a sufficient number of line drawing and shadow pairs as the training dataset so as to capture their perceptual designs and artistic understandings. Nonetheless, the highly expensive and time-consuming artistic drawing process makes large-scale annotation impractical. Another choice is to synthesize a training dataset using algorithms. Although a synthetic dataset might be larger or more diverse than real data, their shadow appearance may not match the artists' wishes and demands. We propose a customized schedule method: we *pre-train* our models with large-scale and diverse synthesized/extracted data, and then *fine-tune* the models on high-quality real data drawn by artists, to simultaneously ensure the robustness and artistic faithfulness.

Data from real artists. We provide 1670 shadow samples drawn from actual artists (Fig. 4-(a)). Those data are from three resources: (1) We manually pick high-quality samples from a previous line drawing shadow dataset [Zheng et al. 2020]. (2) We search the key word "line drawing and shadow pairs" in internet illustration platforms Pixiv [pixiv.net 2007] and Danbooru [DanbooruCommunity 2018] to sample in-the-wild data pairs, and after that, artists are invited to refine the shadows into usable data format for our dataset. (3) We search the key word "line drawing" in Pixiv and Danbooru to sample 10,000 line drawings. We then invite the 12 artists to select their interested line drawings and choose their preferred shadow directions. Afterwards, they draw the target shadows according to their artistic decisions and perceptual understandings. In this way, we collect 1670 high-quality shadow samples that captures the perceptions and designs of artists.

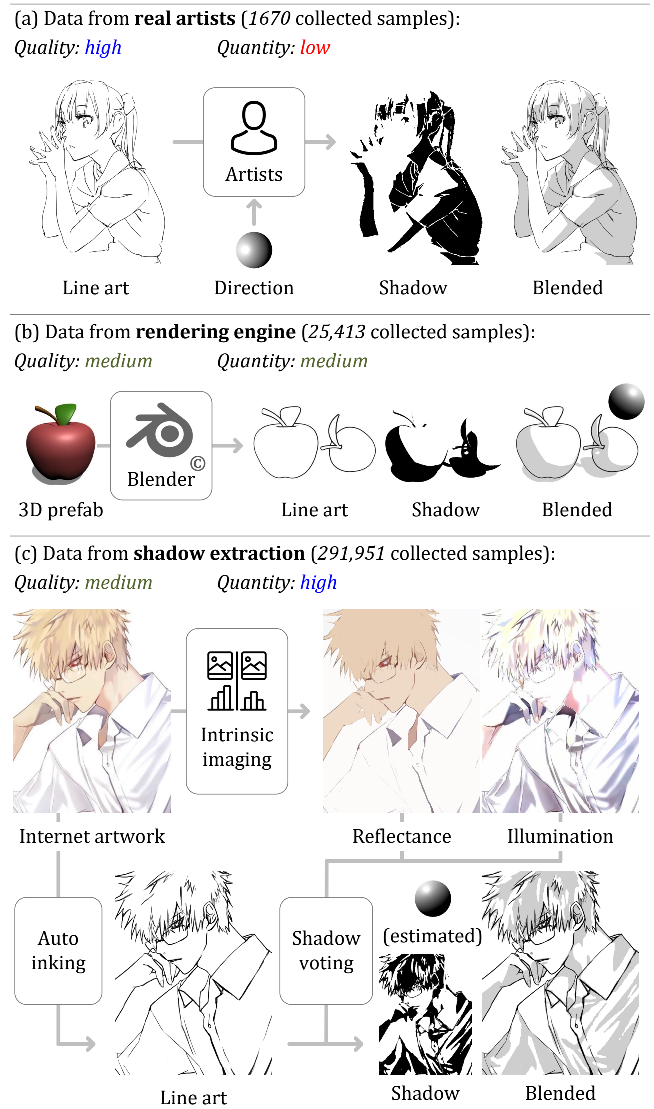


Fig. 4. **Dataset preparation.** We present a large-scale dataset with both real data drawn by artists manually and synthesized data obtained from rendering engines and shadow extraction algorithms.

Data from rendering engine. We use non-photorealistic rendering (NPR) techniques to obtain line art and shadow pairs. To be specific, we search the key word "free" in *Unity 3D Assets Store* and download 471 random 3D prefabs. We import them to the rendering engine Blender [Community 2018] and write a NPR script to generate 25,413 line art and shadow pairs at random shadow directions (Fig. 4-(b)).

Data from shadow extraction. We sample 300,000 random digital paintings from Danbooru dataset [DanbooruCommunity 2018] and Pixiv [pixiv.net 2007] (Fig. 4-(c)). We use auto inking method [Simo-Serra et al. 2016] to extract line arts, and use intrinsic imaging method [Bi et al. 2015] (enhanced with [Zhang et al. 2020a] and [Carroll et al. 2011]) to decompose reflectance and illumination

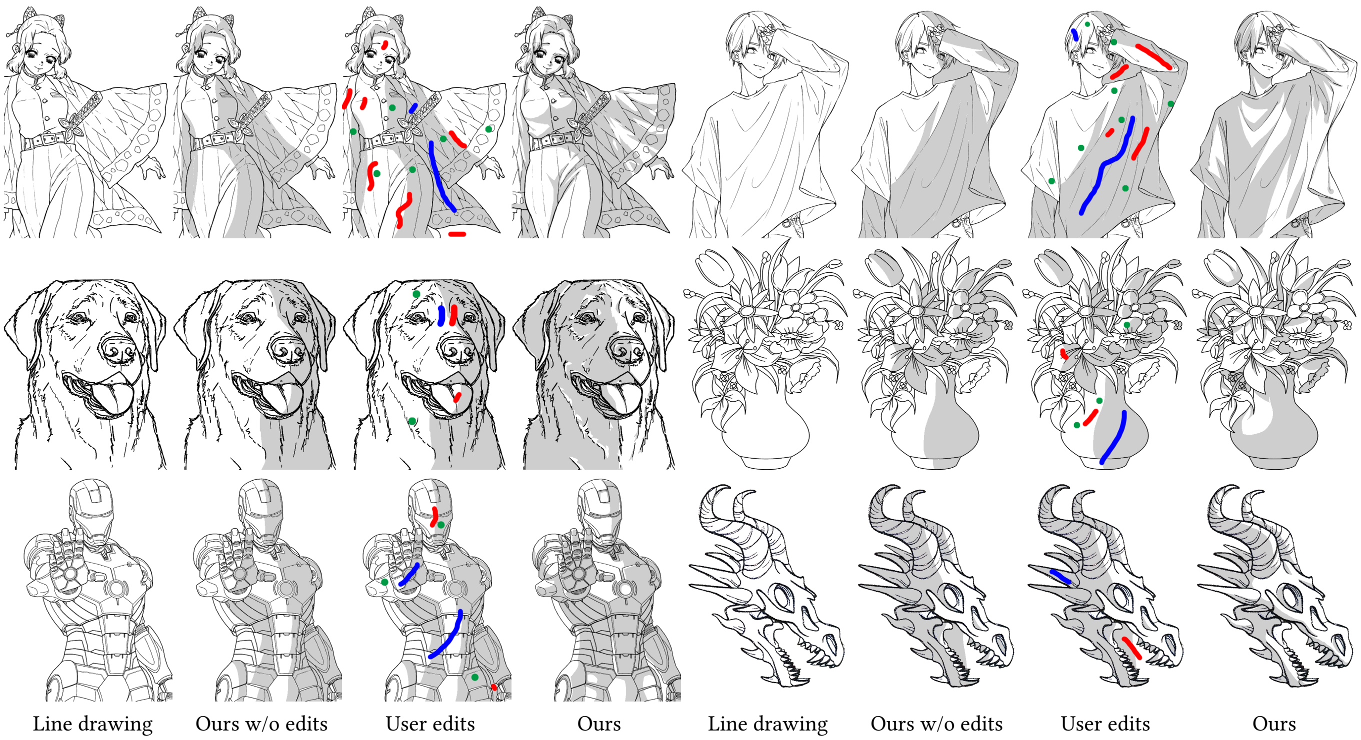


Fig. 5. **Examples of interactive shadow drawing.** Zoom in to see details of the shadows and user edits. 30 more results are presented in the supplement. The user scribbles are precisely one-pixel width and we dilated the scribbles for clearer presentation. *Artworks used with artist permissions.*

maps. We then perform a shadow voting using OTSU algorithm [Otsu 1979] to obtain the shadow, and use the Barron&Malik model [Barron and Malik 2015] to estimate the shadow direction. After that, we manually remove 8,049 pairs with obviously low quality, and acquire the remaining 291,951 qualified pairs.

Training schedule. Our proposed training schedule consists of two phases: (1) Firstly, we pre-train the models with the extracted large-scale shadows for 20 epochs and with the rendered shadows for 15 epochs. (2) Afterwards, as a fine-tuning, we train the models with the high-quality shadows from real artists for 10 epochs. In this way, we achieve a robust model that not only generalizes to diverse inputs but also learns from real-artist data to produce shadows that are faithful to the understanding and willingness of real artists.

4 EVALUATION

4.1 Experimental setting

Implementation details. Our framework is trained using the Adam optimizer [Kingma and Ba 2014] with a learning rate of $\mathrm{lr} = 10^{-5}$, $\beta = 0.5$, at batch size 8. Training samples are randomly cropped to be 256×256 pixels and augmented with random left-and-right flipping. As the shadow model is fully convolutional, it receives adjustable resolutions in inference.

Hyper-parameters. The proposed and recommended configuration is $\lambda_a = 1.0$, $\kappa = 0.1$, and $\lambda_n = 0.5$.

Compared methods. We test several shadow generation methods of (1) the generic model Pix2Pix [Isola et al. 2017] trained on our dataset with the same training schedule as ours; (2) the typical data-driven normal-based method DeepNormal [Hudon et al. 2018a] (official implementation); (3) the interactive method Sketch2Normal [Su et al. 2018] (official method trained with the same scribble shapes as ours); (4) the state-of-the-art shadow generating method ShadeSketch [Zheng et al. 2020] (official open-sourced codes); (5) our application without user edits (in this case we input same shadow directions as other methods when compared to them); and (6) our interactive application.

Testing samples. The tested images are Pixiv [pixiv.net 2007] line drawings and in-the-wild internet line arts. We make sure that all tested images are unseen from the training dataset.

4.2 Qualitative results

Interactive editing. We present examples of interactive shadow drawing in Fig. 5, and 30 additional results in the supplement. We can see that the users can work with our tools to achieve various shadow effects in diverse drawing topics, *e.g.*, human, animal, plant, robot, *etc.*

Comparison to previous methods. We present comparisons with both the automatic methods [Hudon et al. 2018a; Isola et al. 2017; Zheng et al. 2020] and the interactive method [Su et al. 2018] in Fig. 6, and 15 additional comparisons in the supplementary material. We can see that Pix2Pix [Isola et al. 2017] fails in achieving usable results, DeepNormal [Hudon et al. 2018a] tends to output shadows with severe distortions. The results of ShadeSketch [Zheng et al.

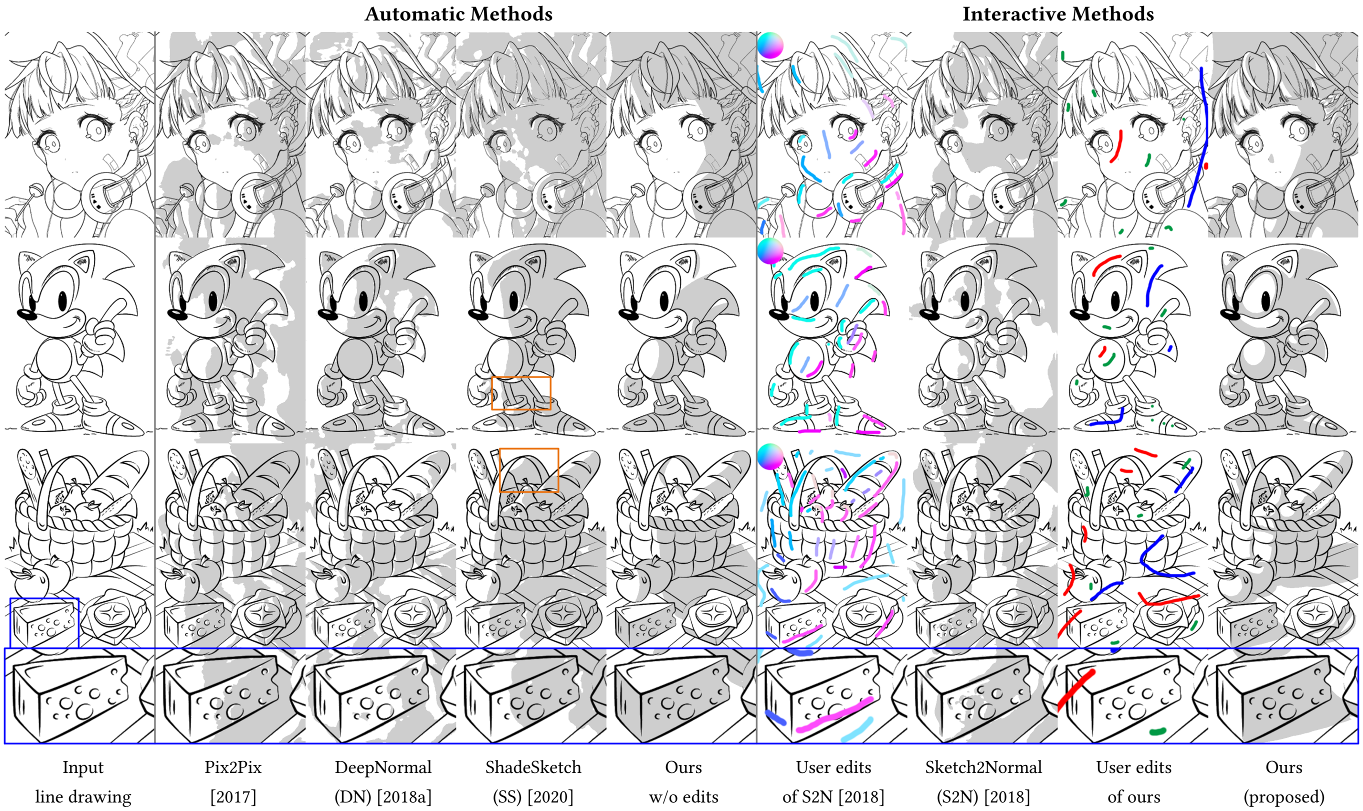


Fig. 6. **Comparisons to possible alternative methods.** 15 more full-resolution comparisons are provided in the supplementary material. The user scribbles are precisely one-pixel width and we dilated the scribbles for clearer presentation. *Artworks used with artist permissions.*

2020] is better than [Isola et al. 2017] and [Hudon et al. 2018a], but it has difficulty in addressing detailed areas, *e.g.*, the mouse legs and the handrails for baskets (as marked in orange rectangles in Fig. 6). Sketch2Normal [Su et al. 2018] yields low-quality shadows, despite the adequately given user scribbles. Our approach, regardless of whether to receive user edits or not, produces clean and practically usable shadows.

4.3 User study

Participant. The user study involves 15 persons: 10 non-artist amateurs and 5 professional artists. Each artist has at least two years of digital painting experience.

Setup. We sample 52 unseen line drawings from Pixiv [pixiv.net 2007], and then assign each line drawing to 3 random users targeted to 3 methods: a commercial tool (Adobe PhotoShop), our approach, and the baseline interactive method [Su et al. 2018]. We also use 4 fully-automatic methods [Hudon et al. 2018a; Isola et al. 2017; Su et al. 2018; Zheng et al. 2020] and the automatic mode of our method to generate shadows for each image. We ensure that any image is assigned to each user at most once to avoid users being trained for specific instances.

User guideline. When drawing shadows interactively, we inform the users that "your time consumption will be recorded and please draw at your normal speed". After they are finished, the users are also shuffled to rank the shadows of automatic methods [Hudon et al. 2018a; Isola et al. 2017; Zheng et al. 2020] and the automatic outputs of ours. We ask users the question — "Which of the following shadow do you prefer most to use in your daily digital painting? Please rank according to your preference."

Evaluation metric. We use the Time Consumption (TC) as speed metric. We record the precise drawing minutes, and split the time consumption into intervals of five minutes. We also use the Average Human Ranking (AHR) as preference metric. For each line drawing, the users rank the results of the 5 methods from 1 to 5 (lower is better). Afterwards, we calculate the average ranking obtained by each method.

Time consumption analysis. The time data are reported in Table 1. We can see that in a dominant majority of cases, our method consumes less than 10 minutes, while in most cases the commercial tool (Adobe PhotoShop) consumes more than 15 minutes. Besides, we report that the average time consuming of ours is 5.35 minutes while the commercial tool is 16.58 minutes, indicating a 3.1× speed up. See also the supplementary material for more detailed data.

Result. The user preferences are reported in Table 2. We have several interesting discoveries: (1) Our framework, even in automatic mode without any user edits, outperforms the secondly best method

(a) Input line drawing (b) W/o shadow brush (c) W/o shadow boundary brush (d) W/o global shadow generator (e) W/o balanced sampling $f_b(\cdot)$ (f) W/o anisotropic penalty λ_a (g) W/o norm regulation λ_n (h) Proposed full method

Fig. 7. **Ablative study.** We study the impact of each individual component within our framework by removing components one-by-one. The user scribbles are precisely one-pixel width and we dilated the scribbles for clearer presentation. *Artworks used with artist permissions.*

Table 1. **Time Consumption (TC).** We compare the time consuming of a typical commercial tool (Adobe PhotoShop) and ours. We visualize the time consumption of 52 shadow drawing cases, *e.g.*, in "ours" row and "$t < 5$" col, the "51.92%" means that the time consumption of our method is less than 5 minutes in 51.92% cases.

Time t (minutes)	$t < 5$	$5 \leq t < 10$	$10 \leq t < 15$	$15 \leq t < 20$	$t \geq 20$
Commercial tool	0.00%	3.84%	28.84%	53.84%	13.46%
Ours	51.92%	46.15%	1.92%	0.00%	0.00%

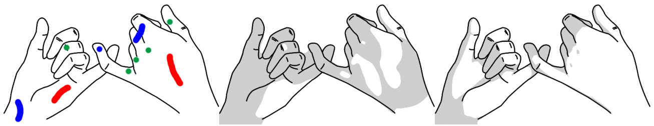
Table 2. **Average Human Ranking (AHR).** We present the ranking results of the user study. The arrow (↓) indicates that lower is better. Top 1 (or 2) score is marked in blue (or red). "*" indicates automatic processing without user hints.

Method	Pix2Pix [2017]	Hudon [2018a]	Su [2018]*	Zheng [2020]	Ours*
AHR ↓	4.53 ± 0.60	2.81 ± 0.76	4.19 ± 0.96	**2.44 ± 0.63**	**1.01 ± 0.13**

by a large margin of 1.43/5. (2) Zheng's approach [Zheng et al. 2020] reports the secondly best score. (3) The two normal-based methods [Hudon et al. 2018a; Su et al. 2018] reports similar perceptual quality, with [Hudon et al. 2018a] slightly better than [Su et al. 2018], despite that [Su et al. 2018] receives interactive edits while [Hudon et al. 2018a] is automatic.

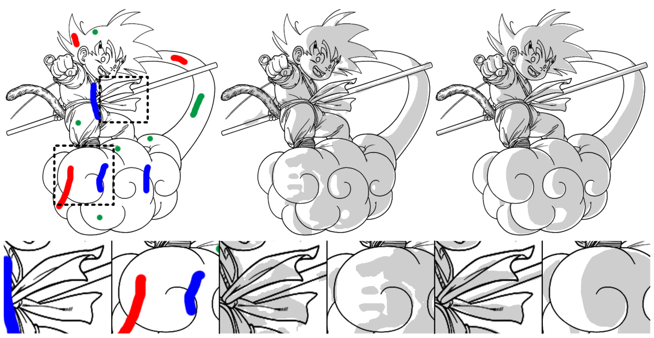
4.4 Ablative study

As shown in Fig. 7, our ablative study consists of the following experiments: (1) We remove the shadow brush and train our framework without red and blue scribbles. We can see that, in absence of the shadow brush, the shadow boundary brush cannot control the shadow locations by itself, resulting in many undesired shadows in the outputs (Fig. 7-(b)). (2) We remove the shadow boundary brush and train our framework without green scribbles. We can see that, without the help of shadow boundary brush, the shadow shape is out of control and users cannot implement their wanted shadow appearances (Fig. 7-(c)). (3) We remove the global shadow generator and train the shadow branch of our neural architecture without global shadow direction embedding. We can see that the global and local shadows becomes inconsistent and distorted (Fig. 7-(d)). (4) We train without the bivariate normal distribution sampling f_b, and instead, we simply sample random pixels as the starting position of training scribbles. We can see that the resulting shadows become severely unbalanced and defective (Fig. 7-(e)). (5) If trained without the anisotropic penalty λ_a, the neural networks fail in achieving sharp and smooth shadow boundaries, resulting in noisy outputs (Fig. 7-(f)). (6) If trained without the shadow direction norm regulation λ_n, the neural networks fail in recognizing appropriate shadow directions, and tends to output collapsed shadows surrounding input lines (Fig. 7-(g)). (7) The full framework suppresses these types



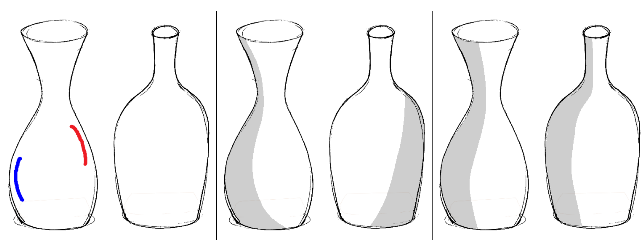
(a) Input (b) Uniform $f_b'(\cdot)$ (c) Proposed $f_b(\cdot)$

Fig. 8. **Influence of different sampling distribution for f_b.** We compare the proposed bivariate normal distribution sampling and a common alternative uniform random sampling. *Artwork used with artist permission.*



(a) Input (b) $\lambda_a = 0.05$ (c) $\lambda_a = 1.0$

Fig. 9. **Influence of the anisotropic penalty weight λ_a.** We visualize the outputs of our method with different anisotropic penalty weight λ_a. *Artwork used with artist permission.*



(a) Input (b) $\lambda_n = 0.01$ (c) $\lambda_n = 0.5$

Fig. 10. **Influence of the shadow direction norm weight λ_n.** We compare the output shadows from models trained with different norm weight λ_n. *Artwork used with artist permission.*

of artifacts and achieves a satisfactory balance over the shadow location, shape, and appearance (Fig. 7-(h)).

4.5 Influence of hyper-parameters

We further study the influence of the bivariate normal distribution sampling f_b by replacing it with a common uniform distribution $f_b'(n_i, n_o) \rightarrow n_i, n_o \sim U(0, 16)$. We can see in Fig. 8 that such uniform sampling causes shadow distortions while our customized sampling bypasses this artifact. We study different weights for the anisotropic λ_a and the norm λ_w in Fig. 9 and 10. We can see that a too small λ_a causes boundary distortions and a too small λ_n causes shadow direction defects.

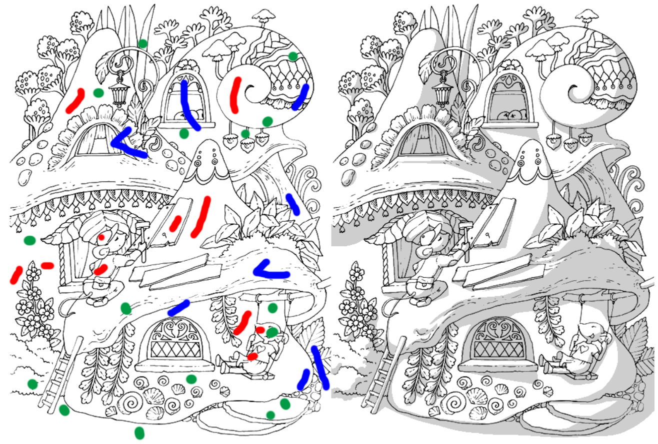


Fig. 11. **Robustness to complicated line drawing.** We present a challenging case where the input line drawing is complicated and detailed. The user scribbles are precisely one-pixel width and we dilated the scribbles for clearer presentation. *Artwork used with artist permission.*

(a) Input (b) User edit (c) Ours

Fig. 12. **Generalization to other art form.** We filter the left artwork to get the middle sketch and the user use our tools to achieve the right blended result. The user scribbles are precisely one-pixel width and we dilated the scribbles for clearer presentation. *Jardin de Paris, public domain.*

4.6 Robustness and generalization

We showcase the robustness in Fig. 11 with a challenging complicated line drawing. We also present a case where our framework is generalized to another art form in Fig. 12. See also the supplementary material for results with more diverse contents and topics.

5 CONCLUSION

We propose a digital painting application to generate shadows on line drawings, with three tools of the shadow brush, shadow boundary brush, and global shadow generator. We train hierarchical neural networks with a collected large-scale dataset of both synthesized data and real data drawn by artists. User study shows that our tools can speed up the shadow drawing process and can achieve practically usable shadows for the daily work of artists. Our dataset will be made publicly available to facilitate related techniques.

REFERENCES

Maury Aaseng, Bob Berry, Jim Campbell, Dana Muise, and Joe Oesterle. 2005. *The Art of Comic Book Drawing*. Walter Foster Publishing.

J.T. Barron and J. Malik. 2012. Color constancy and intrinsic images and shape estimation. *ECCV* (2012).

Jonathan T. Barron and Jitendra Malik. 2015. Shape, Illumination, and Reflectance from Shading. *TPAMI* (2015).

H. G. Barrow and J. M. Tenenbaum. 1978. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*, A. Hanson and E. Riseman (Eds.). Academic Press, 3–26.

S. Beigpour, M. Serra, J. van de Weijer, R. Benavente, M. Vanrell, O.Penacchio, and D.Samaras. 2013. Intrinsic Image Evaluation on Synthetic Complex Scenes. *ICIP* (2013).

Sean Bell, Kavita Bala, and Noah Snavely. 2014. Intrinsic Images in the Wild. *ACM Transactions on Graphics* 33, 4 (2014).

Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala. 2013. OpenSurfaces: A Richly Annotated Catalog of Surface Appearance. *ACM Transactions on Graphics* 32, 4 (2013).

Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala. 2015. Material Recognition in the Wild with the Materials in Context Database. *CVPR* (2015).

Sai Bi, Xiaoguang Han, and Yizhou Yu. 2015. An L1 Image Transform for Edge-Preserving Smoothing and Scene-Level Intrinsic Decomposition. *ACM Trans. Graph.* 34, 4, Article Article 78 (July 2015), 12 pages. https://doi.org/10.1145/2766946

Robert Carroll, Ravi Ramamoorthi, and Maneesh Agrawala. 2011. Illumination decomposition for material recoloring with consistent interreflections. In *ACM Transactions on Graphics*. ACM Press. https://doi.org/10.1145/1964921.1964938

Jiansheng Chen, Guangda Su, Jinping He, and Shenglan Ben. 2010. Face Image Relighting using Locally Constrained Global Optimization. *ECCV* (2010).

Yang Chen, Yu-Kun Lai, and Yong-Jin Liu. 2018. CartoonGAN: Generative Adversarial Networks for Photo Cartoonization. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE. https://doi.org/10.1109/cvpr.2018.00986

Blender Online Community. 2018. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam. http://www.blender.org

DanbooruCommunity. 2018. Danbooru2017: A Large-Scale Crowdsourced and Tagged Anime Illustration Dataset.

Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. 2000. Acquiring the reflectance field of a human face. *the 27th annual conference on Computer graphics and interactive techniques* (2000).

Marek Dvorožňák, Saman Sepehri Nejad, Ondřej Jamriška, Alec Jacobson, Ladislav Kavan, and Daniel Sýkora. 2018. Seamless Reconstruction of Part-Based High-Relief Models from Hand-Drawn Images, In Proceedings of International Symposium on Sketch-Based Interfaces and Modeling. *Proceedings of International Symposium on Sketch-Based Interfaces and Modeling*, Article 5.

P.V. Gehler, C. Rother, M. Kiefel, L. Zhang, and B. Scholkopf. 2011. Recovering intrinsic images with a global sparsity prior on reflectance. *NIPS* (2011).

Roger Grosse, Micah K Johnson, Edward H Adelson, and William T Freeman. 2009. Ground truth dataset and baseline evaluations for intrinsic image algorithms. *International Conference on Computer Vision* (2009).

Matis Hudon, Rafael Pages, Mairead Grogan, and Aljosa Smolic. 2018a. Deep Normal Estimation for Automatic Shading of Hand-Drawn Characters. *ECCV* (2018).

Matis Hudon, Rafael Pagés, Mairéad Grogan, Jan Ondřej, and Aljoša Smolić. 2018b. 2D Shading for Cel Animation, In Expressive: Computational Aesthetics, Sketch-Based Interfaces and Modeling, Non-Photorealistic Animation and Rendering, Tunç Aydın and Daniel Sýkora (Eds.). *Expressive: Computational Aesthetics, Sketch-Based Interfaces and Modeling, Non-Photorealistic Animation and Rendering*. https://doi.org/10.1145/3229147.3229148

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-Image Translation with Conditional Adversarial Networks. *CVPR* (2017).

Diederik P Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *Computer Science* (2014).

Balazs Kovacs, Sean Bell, Noah Snavely, and Kavita Bala. 2017. Shading Annotations in the Wild. *CVPR* (2017).

Chengze Li, Xueting Liu, and Tien-Tsin Wong. 2017. Deep Extraction of Manga Structural Lines. *ACM Transactions on Graphics* 36, 4 (2017).

Chenxi Liu, Enrique Rosales, and Alla Sheffer. 2018. StrokeAggregator: Consolidating Raw Sketches into Artist-Intended Curve Drawings. *ACM Transactions on Graphics* (2018).

Xueting Liu, Tien-Tsin Wong, and Pheng-Ann Heng. 2015. Closure-aware Sketch Simplification. *ACM Transactions on Graphics* 34, 6 (November 2015), 168:1–168:10.

Whyt Maga. 2018. Drawing and Coloring: A CEL-SHADING Tutorial. *https://www.youtube.com/watch?v=Xkek0JuorGE* (2018).

Wojciech Matusik, Matthew Loper, and Hanspeter Pfister. 2004. Progressively-Refined Reflectance Functions from Natural Illumination. *Eurographics Workshop on Rendering* (2004).

Nobuyuki Otsu. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 1 (jan 1979), 62–66. https://doi.org/10.1109/tsmc.1979.4310076

Pieter Peers and Philip Dutre. 2005. Inferring reflectance functions from wavelet noise. *the Sixteenth Eurographics conference on Rendering Techniques* (2005).

Pieter Peers, Dhruv K Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi, and Paul Debevec. 2009. Compressive light transport sensing. *ACM Transactions on Graphics* (2009).

Pieter Peers, Naoki Tamura, Wojciech Matusik, and Paul Debevec. 2007. Post-production Facial Performance Relighting using Reflectance Transfer. *ACM Transactions on Graphics* (2007).

pixiv.net. 2007. pixiv. *pixiv* (2007).

Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. 2017. Scribbler: Controlling Deep Image Synthesis with Sketch and Color. *CVPR* (2017).

M. Serra, O. Penacchio, R. Benavente, and M. Vanrell. 2012. Names and shades of color for intrinsic image estimation. *CVPR* (2012).

Jianbing Shen, Xiaoshan Yang, Yunde Jia, and Xuelong Li. 2011. Intrinsic Images Using Optimization. *CVPR* (2011).

Edgar Simo-Serra, Satoshi Iizuka, and Hiroshi Ishikawa. 2018a. Mastering Sketching: Adversarial Augmentation for Structured Prediction. *ACM Transactions on Graphics* 37, 1 (2018).

Edgar Simo-Serra, Satoshi Iizuka, and Hiroshi Ishikawa. 2018b. Real-Time Data-Driven Interactive Rough Sketch Inking. *ACM Transactions on Graphics* (2018).

Edgar Simo-Serra, Satoshi Iizuka, Kazuma Sasaki, and Hiroshi Ishikawa. 2016. Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup. *ACM Transactions on Graphics* 35, 4 (2016).

Wanchao Su, Dong Du, Xin Yang, Shizhe Zhou, and Hongbo Fu. 2018. Interactive Sketch-Based Normal Map Generation with Deep Neural Networks. *ACM on Computer Graphics and Interactive Techniques* 1, 1 (jul 2018), 1–17.

Daniel Sýkora, Jan Buriánek, and Jiří Žára. 2005. Sketching Cartoons by Example, In Proceedings of Eurographics Workshop on Sketch-Based Interfaces and Modeling. *Proceedings of Eurographics Workshop on Sketch-Based Interfaces and Modeling*, 27–34.

Daniel Sykora, John Dingliana, and Steven Collins. 2009. LazyBrush: Flexible Painting Tool for Hand-drawn Cartoons. *Computer Graphics Forum* 28, 2 (2009).

Daniel Sykora, Ladislav Kavan, Martin Cadik, Ondrej Jamriska, Alec Jacobson, Brian Whited, Maryann Simmons, and Olga Sorkine-Hornung. 2014. Ink-and-Ray: Bas-Relief Meshes for Adding Global Illumination Effects to Hand-Drawn Characters. *ACM Transactions on Graphics* 33, 2 (apr 2014), 1–15.

TaiZan. 2016. PaintsChainer Tanpopo. *PreferredNetwork* (2016).

Xinrui Wang and Jinze Yu. 2020. Learning to Cartoonize Using White-Box Cartoon Representations, In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *CVPR*.

J. Weickert. 1998. *Anisotropic Diffusion in Image Processing*. Uni-saarland.

Wikipedia. 2020. Bivariate Normal Distribution. *https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Bivariate_case* (2020).

Ran Yi, Yong-Jin Liu, Yu-Kun Lai, and Paul L. Rosin. 2019. APDrawingGAN: Generating Artistic Portrait Drawings From Face Photos With Hierarchical GANs. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. https://doi.org/10.1109/cvpr.2019.01100

Ran Yi, Yong-Jin Liu, Yu-Kun Lai, and Paul L Rosin. 2020a. Unpaired Portrait Drawing Generation via Asymmetric Cycle Mapping, In IEEE Conference on Computer Vision and Pattern Recognition (CVPR '20). *CVPR*.

Ran Yi, Mengfei Xia, Yong-Jin Liu, Yu-Kun Lai, and Paul L. Rosin. 2020b. Line Drawings for Face Portraits from Photos using Global and Local Structure based GANs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020), 1–1. https://doi.org/10.1109/tpami.2020.2987931

Lvmin Zhang, Chengze Li, Yi JI, Chunping Liu, and Tien tsin Wong. 2020a. Erasing Appearance Preservation in Optimization-based Smoothing, In European Conference on Computer Vision (ECCV). *European Conference on Computer Vision (ECCV)*.

Lvmin Zhang, Chengze Li, Tien-Tsin Wong, Yi Ji, and Chunping Liu. 2018. Two-stage sketch colorization. In *ACM Transactions on Graphics*.

Lvmin Zhang, Edgar Simo-Serra, Yi Ji, and Chunping Liu. 2020b. Generating Digital Painting Lighting Effects via RGB-space Geometry. *Transactions on Graphics (Presented at SIGGRAPH)* 39, 2 (2020).

Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. 2017. Real-Time User-Guided Image Colorization with Learned Deep Priors. *ACM Transactions on Graphics* 9, 4 (2017).

Qingyuan Zheng, Zhuoru Li, and Adam Bargteil. 2020. Learning to Shadow Hand-Drawn Sketches, In The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.