ColourAIze: AI-Driven Colourisation of Paper Drawings with Interactive Projection System

Fabrice Matulic
Preferred Networks Inc.
Japan
fmatulic@preferred.jp

ABSTRACT
ColourAIze is an interactive system that analyses black and white drawings on paper, automatically determines realistic colour fills using artificial intelligence (AI) and projects those colours onto the paper within the line art. In addition to selecting between multiple colouring styles, the user can specify local colour preferences to the AI via simple stylus strokes in desired areas of the drawing. This allows users to immediately and directly view potential colour fills for paper sketches or published black and white artwork such as comics. ColourAIze was demonstrated at the Winter 2017 Comic Market in Tokyo, where it was used by more than a thousand visitors. This short paper describes the design of the system and reports on usability observations gathered from demonstrators at the fair.

Author Keywords
Projection mapping; automatic colourization; interactive desk; AI art.

INTRODUCTION
Colouring a black and white sketch is a relatively painstaking task that requires a different skillset compared to just drawing line art, especially when the colouring is elaborate. In the comic industry, there are often two separate artists to produce coloured comics: a line artist and a colourist. At early design stages, the line artist may want to get an idea of how their sketches would look like in colour, without necessarily knowing what those colours should be. A modern approach to address this situation is to use an automatic colouring solution based on artificial intelligence (AI), such as one of the services that recently appeared on the web [2,24,26]. Those services use deep generative neural networks trained on thousands of images to automatically produce realistic colour fills for digitally drawn or scanned sketches. The process can be entirely automatic or the user can provide colour hints to the AI by simply adding strokes of the desired colour in particular regions of the drawing. The AI then takes those local hints into account to generate new fills.

While many users may be satisfied with results just appearing on a screen, artists who prefer working exclusively with pen and paper might be interested in seeing the colouring directly applied on their paper drawings. They can achieve that by repeatedly printing out digitised sketches, but this method is cumbersome and costly when experimenting with different colouring styles or drawing in iterations. Furthermore, there exists a large amount of published artwork only available in black and white, such as mangas, for which readers might be curious to see colourisations directly applied onto the paper.

Figure 1. The ColourAIze system with an example of a semi-automatically colourised line drawing

ColourAIze is an interactive colourising system that projects AI-determined colour fills directly onto the paper drawing with the colouring superimposed on the line art resulting in a novel analogue/digital blend (Figure 1). As with the web versions of the colouring services, the ability to specify colour hints is supported through a pen and touch interface. The system was deployed as twin booths at the Winter Comic Market in Tokyo at the end of 2017. Over the course of three days, more than a thousand visitors, including many amateur manga artists, experienced ColourAIze using self-drawn sketches or samples provided by demonstrators. This paper describes the contributed system and summarises the main observations on how it was used at the fair.

RELATED WORK
The idea of projecting dynamic digital content onto paper documents dates back to Wellner's DigitalDesk [22], a sys-

ISS '18, November 25–28, 2018, Tokyo, Japan

ACM ISBN 978-1-4503-5694-7/18/11…$15.00
https://doi.org/10.1145/3279778.3279785

tem that not only augmented documents with projected information, but also allowed users to interact with the content using touch, pens and other physical artefacts. Since this pioneering work, several prototypes building on that concept were created. They range from apparatuses with fixed overhead cameras and projectors that cover the entire desk [1,7,13,18,23], similar to the DigitalDesk, to mobile solutions, where the user can freely move a small projection device with a focused augmented view onto the paper content [10,20,21]. Typical demonstration prototypes developed for those systems include applications to capture portions of paper documents and link them to digital data. For instance, the static analogue content of a textbook can be augmented with projected multimedia teaching material [7].

With regard to automatic colourisation, there is a wealth of work in the computer vision literature on colourising natural monochrome images, with latest algorithms able to produce very realistic results thanks to convolutional neural networks [4,8,25]. Deep neural networks are also at the heart of modern techniques that can generate artistic colour fills [5,9,20,22] and photo-like images [4,14] from black and white drawings. These tools can be integrated in AI-guided systems, which may also include smart sketching aids [9,11], to assist artists in their creative work [16].

ColourAIze draws inspiration from those two categories of research work and aims to bring AI-driven colourisation to paper-based artwork using DigitalDesk-style projection and pen and touch interaction.

SYSTEM

ColourAIze was designed as an alternative interface to the web-based AI colourisation service PaintsChainer [24] that is closer to the pen and paper experience. One of the main goals was to expose PaintsChainer to a wider audience at the Tokyo Comic Market, especially manga artists who work mostly with analogue tools.

Following is a description of the software and hardware components that make up the system used at the fair.

PaintsChainer

PaintsChainer is a free colourisation web service based on deep convolutional generative adversarial networks [17] trained on several thousands of line drawings + coloured image pairs consisting mostly of manga artwork. Three colourisation styles are available, corresponding to different parametrisations of the underlying networks and training models: "Tanpopo" (gentle watercolour-like style), "Satsuki" (somewhat sharper colouring) and Canna (stronger highlights and shadows). Additionally, the user can provide local colour hints via simple colour line strokes. A single short stroke is generally sufficient to specify a colour preference for a particular subregion of the drawing, e.g. for the dominant hair or dress colour of a character. In most cases, the AI is able to automatically determine the (fuzzy) boundaries of the region to fill with appropriate shade variations of the chosen colour.

Hardware

The installation consists of a 2.1m tall box-shaped wooden booth with enclosures, platforms and fasteners for the sensing and display equipment. A short-throw full HD BENQ projector mounted within an enclosure at the back of the booth projects the colour fills onto a 60×80cm table at a 95cm height, i.e. for use while standing (Figure 1). Users place paper drawings to be colourised in a marked zone, hereafter referred to as the "capture area", located in the centre of the table. To prevent sheets of paper from slipping when tracing on them with the pen an A4-sized transparent non-slip mat is glued to the table. Drawings placed on the mat are captured by a 4K Logitech BRIO webcam fixed to the front edge of a small platform 60cm above the table. A 27 inch 4K display is placed on that platform to show content, in particular colourisation results, to onlookers standing behind the user (Figure 3).

Colour hints are input with a stylus, which is tracked by a Wacom Intuos Pro fixed to a plate located just underneath the capture area. The electromagnetic field of the Intuos is sufficiently strong for the sensors to detect the stylus through the 2cm-thick table. Colour selection and command execution are performed using touch input on a Dragon Touch X10 10.6 inch tablet, which transmits and receives data via Bluetooth. The tablet can of course be freely moved around and placed anywhere on the table to the user's liking. This bimodal design was chosen to promote bimanual use of the interface, as suggested by prior work [3,15], with the pen focusing exclusively on colour input on the paper drawing, while the non-dominant hand operates controls using touch. A further reason for not opting for an all-pen interface is the limited size of currently available commercial tablets for high-precision stylus input, which precludes pen use on large surfaces and unknown user media.

Finally, UI, input and image pre-processing are handled by a Windows PC placed in a compartment under the table. This PC is connected to a local network to access the PaintsChainer service running on a separate server.

User Interface

Calibration

For the projection mapping and pen and touch input to function properly, the system first needs to undergo two calibrations. Standard procedures are used for that. Visual calibration of the webcam with the projector is performed using four bright dots located around the capture area. Those dots are extracted from the image by computer vision to determine the affine transform needed to dewarp the images and correctly overlay the colour fills on the paper line art. The Intuos Pro is calibrated by successively tapping four crosshairs near the corners of its sensing area. Since the PC and its accessories are placed inside the booth and thus are hardly accessible during normal operation, a mechanism to trigger the calibration procedure from the front without exposing any administration controls and widgets

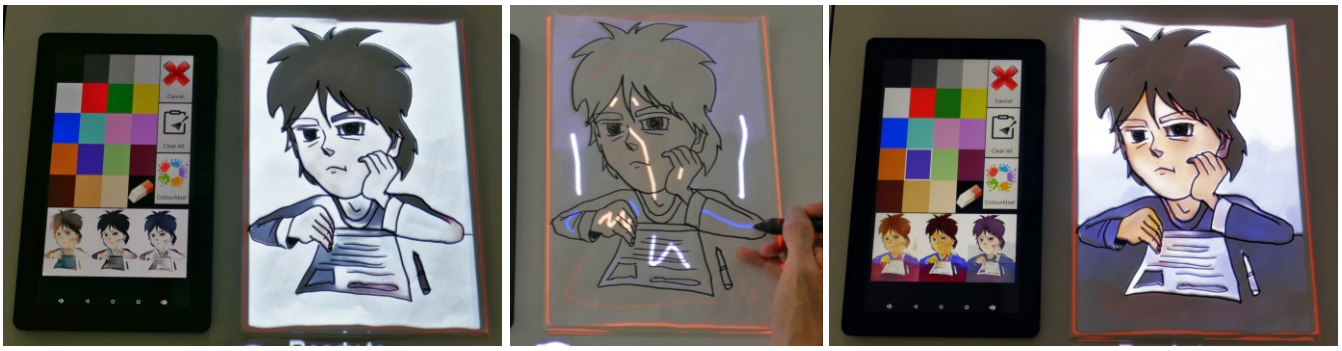

Figure 2. From left to right: initial colourisation; user adding local colour hints; result with the hints

to the user is included: It consists in fully blocking the 4K camera with the hand or an opaque object. The image processor detects the sudden brightness drop and thereupon launches the calibration pipeline.

User Interface

The user interface is composed of preview icons for each supported style, a colour palette and buttons to execute and clear the colourisation. Users apply colour hints for the AI by selecting a colour from the palette using touch and tracing with the pen on the desired regions of the paper drawing (Figure 2). To ensure the user can view clear, unmarred projections of either the colouring results or the hint strokes, the hovering detection capability of the Intuos is used to toggle between views depending on pen input. Specifically, when the pen comes into hovering range (indicating the intention to trace strokes), the opacity of the colour fill is reduced and the hint strokes become visible. When the pen moves out of detection range, the strokes disappear and the colouring is restored to full opacity. Hint strokes are also made visible when colours are selected in the palette.

The previews showing coloured drawings with the three different styles are shown at the bottom of the tablet screen. The user can tap a preview icon to project the selected colouring in full size onto the paper drawing. The previews are also displayed on the public monitor placed on the top platform with additional QR codes. These QR codes are linked to the digital images to allow the user and bystanders to easily download coloured drawings to their mobile device (Figure 3).

Processing Pipeline

When the button to execute the colourisation is pressed, the webcam captures an image frame, which is dewarped according to the affine transform obtained in the calibration step. The image is further cropped so as to keep only the line drawing and brightness. Contrast and sharpness are adjusted to yield a clear image with a high contrast between lines and background, since this is what the AI trained with images of professional artwork expects. The pre-processed image is then sent to PaintsChainer over the network along with the colour hints, if provided by the user. One to two seconds later the results are returned and the colour fill corresponding to the currently selected style is projected back onto the drawing. The results with the other styles are shown as preview icons on the tablet and as full images on the observation monitor.

The software was developed in Java with the help of the OpenCV library for image processing.

DEPLOYMENT AT THE COMIC MARKET

The Comic Market, also known as Comiket, is a large fair dedicated to self-published works, especially manga comics and art. It is held biannually at the Tokyo International Exhibition Center (Tokyo Big Sight). Two identical booths of the system were created and deployed at the 93rd Winter Comic Market, which took place between the 29th and the 31st of December 2017. In addition to the booths, a large palette-shaped table with paper and drawing utensils was installed for visitors, who wished to draw their own sketches. For the others, several different drawing samples to choose from were available. Visitors brought their self-drawn or chosen art to one of the available booths, where a demonstrator greeted them before giving them a brief explanation of the basic operations of the interface. The visitor was then free to experiment with the different colouring styles, apply various colour hints and even add (inked) lines to the sketch to see how it changed the colourisation. A glass panel was available to press against curved pages of thicker comics so that the projected colours were correctly superimposed on the line drawings. Each visitor was given roughly 5 minutes to use a booth, 3 minutes during busy periods, after which they were gifted with a thank-you plastic sleeve.

Observations

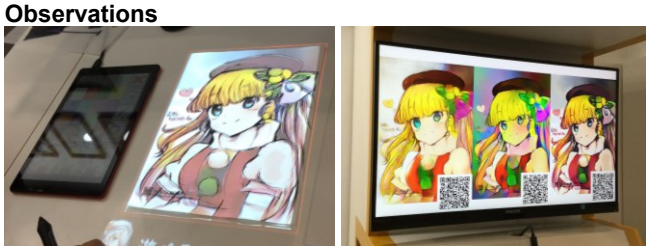

Figure 3. Colourised visitor drawing (left) with preview of all three colourisation styles on the observation monitor (right)

More than a thousand visitors came to use the booths over the three-day fair. As the goal was to get a maximum of

people to experience ColourAIze, it was not possible for demonstrators to conduct interviews or hand out questionnaires. The following observations are the result of assistants' recollection of visitors' behaviour and spontaneous comments while using the booths.

- Overwhelmingly, visitors found that ColourAIze delivered an enjoyable and novel experience. They were impressed that realistic and sometimes very professional-looking colour fills including shade variations matching the line art contours could be automatically produced and almost instantly projected back onto their sketches (see for example the image in Figure 3, which was drawn and coloured by a visitor). Roughly 35% expressed it was fun to see a colourisation that they had not imagined and that it could potentially inspire them when manually colouring their drawings.
- About 15% of the visitors drew their own sketches and 1% added lines to them while using the booth to see how it changed the colourisation. Generally, well-drawn character sketches were better colourised, but there were also some interesting instances where the AI was able to come up with convincing colour fills for sketches of objects, which were not represented in the training data. As a fun experiment, assistants also suggested to try to use physical objects as input, e.g. remote controls, hands, various mobile devices etc. to see what kind of colouring would be generated.
- 70% looked both at the projected colours on the drawing and the observation monitor, because it displayed large images with all three colouring styles and also possibly because the colours differed slightly (Figure 3). There was no modification of the colouring to attempt to correct for rendering differences between the projector and the screen or to mitigate the influence of the surrounding lighting conditions.
- Regarding the separation of input modalities for the interface, about 70% confused pen and touch at the very beginning, with most people wanting to operate the tablet with the stylus as well, despite the initial explanations and demonstrations of the assistants. Moreover, touch operations were overwhelmingly performed using the fingers of the pen-holding hand, i.e. the pen was temporarily tucked to tap buttons on the tablet. Those observations based on more than a thousand users confirm and complement findings from previous studies, which report that some participants want to use the pen not only for inking, but also for tool selection [3,14].
- Some inconsistencies in the colouring were observed when trying to colourise comic strips in which the same character appeared in multiple panels. The AI does not establish correspondences between panels and thus elements appearing in several frames may be colourised differently. This is a current limitation of PaintsChainer but hopefully future versions will be able to cope with comic strips.

Discussion

Judging from visitors' reactions both on site and on social media, projecting colour fills determined by AI onto paper drawings is entertaining for comic fans and potentially inspiring for amateur artists. The combination of projected colours with pencilled line art creates an interesting digital-analogue blend not unlike other projection mapping systems. In the case of ColourAIze, however, the contours for the mapping are created by the user, presumably resulting in a higher sense of personalisation and uniqueness. The ability to add simple colour hints also seems to be the right amount of customisation, giving users some level of control, without it becoming too tedious.

While many artists have gone completely digital, pen and paper remain the most simple and accessible instruments for regular people to quickly and casually sketch. Analogue colourisation beyond simple uniform fills typically requires more effort and tools, therefore ColourAIze might be useful to cartoonists, who would like to better express their art in colour. The current system, of course, only yields static projected results and while the colourisation can be influenced by local hints, it is not fully controllable and the results can be at times somewhat unexpected. Furthermore, there is no integration with creation and authoring processes to edit and further work with the colouring in order to produce publishable output. The first limitation is due to current AI colourisation engines and future algorithms might be able to provide more control and be more predictable with consistency across multiple images or strip panels. For now, for casual use, ColourAize's main appeal is perhaps its simplicity and immediacy. To that end, the interface can be further reduced to simple tools and input with a single modality, i.e. pen or touch for everything (with a preference for the pen as it is more precise). Regarding the UI, visitors' confusion with pen and touch and their tendency to use only a single hand for all input show that bimanual interaction for combined use of both modalities is not immediately intuitive, at least when it seems they can be used interchangeably. Getting used to such an interaction style may not require much time, but it is likely not suitable for quick walk-up-and-use systems.

CONCLUSION

This short paper introduced ColourAIze, a novel projection mapping system that scans black and white drawings on paper and projects back colour fills generated by AI onto the drawing. A simple pen and touch interface combining two tablets (one visible, one hidden) allows users to select colour styles and provide local colour hints to influence the colourisation. Two booths were deployed at the Winter 2017 Comic Market in Tokyo and the dominating impression from the visitors was that ColourAIze provides an original and playful experience. Amateur manga artists also expressed interest in such a system to inspire them during the early phases of the creative process. Observations from assistants further revealed an important issue about pen and touch interfaces –that people want to perform single-point

surface input with a single modality and a single hand– which designers of walk-up-and-use systems may want to take into consideration. Hopefully this work can inspire future mixed digital-analogue approaches to support art creation using AI.

REFERENCES

1. Raimund Dachselt and Sarmad Al-Saiegh. 2011. Interacting with printed books using digital pens and smart mobile projection. In *Proceedings of the Workshop on Mobile and Personal Projection (MP2)@ ACM CHI*.
2. Kevin Frans. Deepcolor. http://color.kvfrans.com
3. Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + touch = new tools. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, 27–36. https://doi.org/10.1145/1866029.1866036
4. Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. 2016. Let There Be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification. *ACM Transactions on Graphics* 35, 4: 110:1--110:11. https://doi.org/10.1145/2897824.2925974
5. P Isola, J Y Zhu, T Zhou, and A A Efros. 2017. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 5967–5976. https://doi.org/10.1109/CVPR.2017.632
6. Yuusuke Kataoka, Takashi Mastubara, and Kuniaki Uehara. 2017. Deep Manga Colorization with Color Style Extraction by Conditional Adversarially Learned Inference. *Information Engineering Express* 3, 4: 55–66.
7. Hideki Koike, Yoichi Sato, and Yoshinori Kobayashi. 2001. Integrating paper and digital information on EnhancedDesk: a method for realtime finger tracking on an augmented desk system. *ACM Transactions on Computer-Human Interaction* 8, 4: 307–322. https://doi.org/10.1145/504704.504706
8. Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. 2016. Learning Representations for Automatic Colorization. In *Proceedings of the European Conference on Computer Vision*, 577–593.
9. Yong Jae Lee, C Lawrence Zitnick, and Michael F Cohen. 2011. ShadowDraw: Real-time User Guidance for Freehand Drawing. In *ACM SIGGRAPH 2011 Papers*, 27:1--27:10. https://doi.org/10.1145/1964921.1964922
10. Chunyuan Liao, Hao Tang, Qiong Liu, Patrick Chiu, and Francine Chen. 2010. FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface. In *Proceedings of the international conference on Multimedia*, 361–370. https://doi.org/10.1145/1873951.1874001
11. Alex Limpaecher, Nicolas Feltman, Adrien Treuille, and Michael Cohen. 2013. Real-time Drawing Assistance Through Crowdsourcing. *ACM Transactions on Graphics* 32, 4: 54:1--54:8. https://doi.org/10.1145/2461912.2462016
12. Yifan Liu, Zengchang Qin, Tao Wan, and Zhenbo Luo. 2018. Auto-painter: Cartoon image generation from sketch by using conditional Wasserstein generative adversarial networks. *Neurocomputing*. https://doi.org/https://doi.org/10.1016/j.neucom.2018.05.045
13. George Margetis, Antonios Ntelidakis, Xenophon Zabulis, Stavroula Ntoa, Panagiotis Koutlemanis, and Constantine Stephanidis. 2013. Augmenting physical books towards education enhancement. In *User-Centered Computer Vision, 2013 1st IEEE Workshop on*, 43–49.
14. Fabrice Matulic. 2014. Towards Document Engineering on Pen and Touch-Operated Interactive Tabletops. *ETH Zurich Doctoral Dissertation*.
15. Fabrice Matulic and Moira Norrie. 2013. Pen and Touch Gestural Environment for Document Editing on Interactive Tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 41–50.
16. Changhoon Oh, Jungwoo Song, Jinhan Choi, Seonghyeon Kim, Sungwoo Lee, and Bongwon Suh. 2018. I Lead, You Help but Only with Enough Details: Understanding User Experience of Co-Creation with Artificial Intelligence. In *Proceedings of the Conference on Human Factors in Computing Systems*, 649:1--649:13. https://doi.org/10.1145/3173574.3174223
17. Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
18. John A Robinson and Charles Robertson. 2001. The LivePaper system: augmenting paper on an enhanced tabletop. *Computers & Graphics* 25, 5: 731–743.
19. Patsorn Sangkloy, Jingwan Lu, Chen Fang, FIsher Yu, and James Hays. 2017. Scribbler: Controlling Deep Image Synthesis with Sketch and Color. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*.
20. Hyunyoung Song, Tovi Grossman, George Fitzmaurice, François Guimbretière, Azam Khan, Ramtin Attar, and Gordon Kurtenbach. 2009. PenLight: combining a mobile projector and a digital pen for dynamic visual overlay. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems, 143–152. https://doi.org/10.1145/1518701.1518726

21. Hyunyoung Song, François Guimbretière, Tovi Grossman, and George Fitzmaurice. 2010. MouseLight: bimanual interactions on digital paper using a pen and a spatially-aware mobile projector. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2451–2460. https://doi.org/10.1145/1753326.1753697
22. Pierre Wellner. 1993. Interacting with paper on the DigitalDesk. *Communications of the ACM* 36, 7: 87–96. https://doi.org/10.1145/159544.159630
23. Chih-Sung Wu, Susan J Robinson, and Ali Mazalek. 2008. Turning a page on the digital annotation of physical books. In *Proceedings of the 2nd international conference on Tangible and Embedded Interaction*, 109–116. https://doi.org/10.1145/1347390.1347414
24. PaintsChainer. https://paintschainer.preferred.tech/index_en.html
25. Richard Zhang, Phillip Isola, and Alexei A Efros. 2016. Colorful Image Colorization. In *Proceedings of the European Conference on Computer Vision*, 649–666.
26. PaintsTransfer. http://paintstransfer.com