招待講演 | 特別セッション

[IKs] 画像感性 特別講演
Special Talk: Sensibility on Imaging

2018年6月19日(火) 14:30 ~ 15:20 レセプションホール (けやき会館 三階)

[IKs-01] Kansei Imaging: Holistic Image Enhancement
Kansei Imaging: Holistic Image Enhancement

*Carl Staelin[1], Boris Oicherman[2] （1. Google Inc., 2. University of Minnesota）

We describe and explore how Kansei Imaging should combine and integrate research and algorithms from a wide range of disciplines, such as image analysis, image processing, machine learning, printing, photography, painting and composition, color science, and beauty styling (makeup). We explore how the final application impacts nearly all aspects of the imaging solution: e.g. more user feedback and guidance permits stronger and "riskier" enhancements and alternatives. We also explore how integrating various analyses and segmentations into the solution in order to allow control over each enhancement on a per-pixel basis is vital. We finally raise topics for further research.

Kansei Imaging: Holistic image enhancement

Carl Staelin*, and Boris Oicherman**

*Google Inc.
**Weisman Art Museum, University of Minnesota

Kansei imaging, the development of algorithms to modify an existing image to reach a set of predefined emotional qualities, involves the complete range of image analysis and enhancement algorithms. Algorithms such as sharpening, denoising, and local contrast enhancement are part of a standard image enhancement repertoire, and these algorithms require matching analysis algorithms to evaluate the image and control the strength of the algorithms on a per-pixel basis. A complete Kansei imaging solution should include other analyses and algorithms, such as color and tone mapping, content and context detection, and image composition enhancement, all with matching analysis algorithms.

1. Introduction

Image enhancement algorithms, such as sharpening and denoising, have been around for decades, with new algorithms and applications, such as seam carving [3], regularly added to the field. Kansei Imaging [1] proposed that image enhancement algorithms might be tailored or controlled to yield a desired emotional affect, with the first steps focused on controlling color hue and saturation. The Automatic Image Enhancement Server [2,4] tackled other aspects such as sharpening and denoising. We propose that these and other algorithms should be knitted together to evaluate and enhance all the aspects of the image, including color, contrast, composition, and sharpness, to yield the desired effect.

2. Image Enhancement Algorithms

Many image enhancement algorithms have been developed to solve or address a wide range of image features. Standard and well-known classes of algorithms include sharpening and denoising [11], or contrast enhancement and local contrast enhancement [12]. There are other categories of image enhancements, such as memory color enhancement [13] and skin tone correction [2] that can positively affect the emotional impact of an image.

More recent enhancements include algorithms such as haze reduction [18], reflection removal [19], and object removal [29]. Many new algorithms and developments have been driven by consumer cellphone photography.

One element of the enhancement algorithms for Kansei imaging is defining the action in terms of emotion-related axes. [1] proposed a color emotion space for color activity, color weight, and color warmth. The enhancement algorithms may then process the images according to these aspects. Examples of less and more color active images:

Similar axes and definitions can be made for contrast, brightness, and sharpness. Other enhancements, such as memory colors do not have as obvious a choice of axes.

3. Image Composition

A vital aspect of each photograph is the subject and the image composition. Related aspects include things such as the framing of the subject within the image, and the relative illumination, color, and sharpness of the subject compared to the background.

There are various algorithms to restructure the image composition via cropping [6,7] or other means such as seam carving [3] or object placement [23]. In particular [6,23] define aesthetic measures that combine various important compositional aspects such as the rule-of-thirds, visual weight balance, and diagonal saliency, but ignore some important aspects, such as gaze direction or scan path.

Some aspects of image composition can convey emotional impact. A well-composed and balanced image is generally more relaxing and often preferred to an imbalanced image which can subtly provide tension or a feeling of disorder [17]. An example of poor composition where the subject's gaze direction is towards the near edge instead of across the image, giving a claustrophobic feeling as shown below. There are hundreds of books and

---

This work was done while the authors worked at Hewlett-Packard Laboratories Israel.

* Google, MATAM Building 30, P.O. Box 15096, Haifa 31905, Israel
e-mail: staelin@google.com
** Weisman Art Museum, 333 E. River Road, Minneapolis, MN 554355, USA. E-mail: boris@umn.edu

resources on photo composition such as [37,38,39,40], a good summary of photo composition advice is in [36].

Another common problem, particularly in consumer photographs, is the intersection of spurious and distracting background elements with a subject, such as a tree apparently sprouting out of a subject's head, as shown above, or distractions, commonly known as "photo bombing". Automatically identifying and potentially removing such distractions can dramatically improve the image aesthetics and appeal, such as this manually edited example from [41] shown below.

Finally, one should note that an important aspect of art is the element of "surprise". So, Kansei imaging should retain the ability to "surprise" the user. Few things are as boring as a "cookie cutter" composition with no uniqueness or "spark". These observations suggest the potential for composition-related Kansei emotive axes such as "relaxing" vs. "surprising".

4. Image Analysis Algorithms

Proper control and utilization of the various enhancement algorithms above relies on extensive image analyses. For example, sharpening and denoising algorithms require accurate noise and blur estimation on per-pixel basis to better separate signal from noise [14, 15]. Similarly, face and facial feature detection [16, 31] and skin detection [2] can be used to selectively sharpen eyes and mouth while largely leaving skin untouched. Blur estimation can be used to sharpen or refocus an image [30].

Obviously, significant low-level and semantic-level analyses related to image composition are also required. Subject or foreground segmentation may be done via saliency and matting [5,43], and myriad algorithms exist for object segmentation [42]. Other aspects of composition, such as horizon, leading lines, gaze direction, and so forth will also need to be analyzed and evaluated.

A partial list of various analyses and how they might used includes:

- Noise estimation: denoising and sharpening.
- Blur estimation: sharpening, depth estimation.
- Face and facial feature detection: sharpen eyes, enhance skin color, whiten and sharpen teeth, "make-up" color correction and contrast management. Foreground/subject detection and segmentation, and gaze detection.
- Gaze direction and scan path: image composition and foreground/background segmentation.
- Sky, foliage, skin segmentation: memory color correction, denoising and sharpening.
- Saliency: image composition and foreground background segmentation.
- Activity: sharpening and denoising.
- Foreground/background segmentation: color correction or enhancement, contrast enhancement or reduction, etc.
- Depth (distance from camera): haze reduction, color correction.
- Color weight, warmth, activity: color correction.

5. Artistic Intent

One key aspect is to infer any existing or latent artistic intent that may be in the input image. Recent work [21] creates a framework for recognizing various common artistic intents, such as "serene" or "romantic". Similarly [22] describes a framework to evaluate the emotional effect of an image, such as "amusement", "anger", or "contentment". If there is a clear subject or intent, then any enhancements should likely complement the existing subject or intent. For example, the subject might be subtly sharpened and the background blurred, or color scheme for the subject might be made more active and the background less active. Similarly, if there is a clear intent with respect to the existing colors, the enhancements might strengthen that intent rather than blithely optimizing for a generic intent.

6. Application

The context of the actual imaging application will have a significant impact on the set of available enhancements. For example, an application similar to the Automatic Image Enhancement Server [4] cannot modify the image composition and cannot afford to degrade images because there is no interactive user feedback or guidance and any imaging "failures" would result in the return of the whole photobook, so its set of enhancements is necessarily relatively limited and the enhancements must be applied conservatively.

Conversely, an application like Instagram has an explicit interactive image enhancement step in the image publication action. In this case the full set of enhancements would be available and the "best" or a set of alternative enhancements could be iteratively presented to the user, and the application may ask the user for explicit guidance as to the artistic and emotional intent for the image, such as "happy/active" or "broody/dark". In this way more extreme enhancement alternatives may be explored or selected.

One key element of any Kansei imaging application, particularly an interactive or user-configurable application, is providing the user with sensible, understandable controls that yield meaningful differences in the final output image. Recent work [28] uses a Bayesian aggregated user feedback-based method that converts multiple enhancement meta-parameters into a single user control slider. It may be possible to use similar approaches to help define a small set of user-understandable axes and controls that can guide and control the Kansei imaging's enhancements.

Other work [32, 33, 34] shows how machine learning can learn to predict or measure aspects of image aesthetics and quality assessment, which in turn might be used in a variety of ways, such as automatically choosing between enhancement alternatives. [35] shows how to extend these approaches to develop a personalized image quality measure based on relatively little user feedback.

7. Display and Print

The image enhancements should be done in a manner that is sensitive to the available output gamut and contrast. For displays, the input and output gamut and contrast will be similar, but printed images generally have much less contrast and a smaller and different color gamut than digital displays.

Since the gamut and contrast of print is generally much smaller than that of displays, using gamut mapping, hue rotation, tone mapping and other contrast and local contrast enhancement techniques is vital [26, 27]. In addition, the image processing and enhancements should be done in a fashion that is aware of the available output gamut so the enhancements may make use the available gamut, rather than post-processing an enhancement image to target it for print.

Additionally, print halftoning generally blurs images and the final print resolution is generally known. So, the entire output image should be sharpened to pre-compensate for half-toning blurring [24].

8. Development and Evaluation

Development of a Kansei imaging application is non-trivial, in particular because accurate measures and assessments of enhancement results can only be done by humans. In addition, since the problem is so complex and multi-dimensional, with composition, color, contrast, and myriad other features materially impacting the human emotional effect, any such application will need to be tested and evaluated on a development corpus containing at least thousands and likely tens of thousands of images. It is prohibitively expensive to regularly evaluate changes against such a large corpus because of the skilled human time investment.

One method that has proven useful is to create a large corpus containing a broad range of images, from high quality professional images to poor quality cellphone images by unskilled amateurs, with a broad range of subjects and content. The application may be run on each image, and the output compared to a "golden" result using an image quality assessment tool such as FSS [25]. The images may then be quickly sorted by visible difference to find images which were most significantly impacted by any change. "Golden" result images may be updated when the new result is an improvement.

References

1) Boris Oicherman and Carl Staelin, "Kansei Imaging: first steps", HPL-2009-12, February 2009.
2) Hila Nachlieli, Ruth Bergman, Darryl Greig, Carl Staelin, Boris Oicherman, Gitit Ruckenstein, and Doron Shaked, "Skin-Sensitive Automatic Color Correction". HPL-2009-13. February 2009.
3) Shai Avidan and Ariel Shamir, "Seam caring for content-aware image resizing", Proceedings SIGGRAPH'07, August 2007.
4) Renato Keshet, Carl Staelin, Boris Oicherman, Mani Fischer, Pavel Kisilev, Sagi Schein, Doron Shaked, Marie A. Vans, Hila Nachlieli, Ruth Bergman, Ron Maurer, Michal Aharon, and Alexander Berkovich, "Automatic Photo Enhancement Server (HIPIE 2)", Proceedings International Symposium on Technologies for Digital Photo Fulfillment. 2009.
5) Ming-Ming Cheng, Guo-Xin Zhang, Niloy J. Mitra, Xiaolei Huang, Shi-Min Hu, "Global contrast based salient region detection", Proceedings Conference on Computer Vision and Pattern Recognition, 2011.
6) Ligang Liu, Renjie Chen, Lior Wolf, Daniel Cohen-Or, "Optimizing Photo Composition", Eurographics 29(2), 2010.
7) Guanjun Guo, Hanzi Wang, Chunhua Shen, Yan Yan, Hong-Yuan Mark Liao, "Automatic image cropping for visual aesthetic enhancement using deep neural networks and cascaded regression", IEEE Transactions on

Multimedia, 2018.
8) Rafal Mantiuk, Scott Daly, and Louis Kerofsky, “Display adaptive tone mapping”, Proceedings ACM SIGGRAPH’08, 2008.
9) R. Mantiuk, R. Mantiuk, A. Tomaszewska, and W. Heidrich, “Color correction for tone mapping”, Eurographics 28(2), 2009.
10) Huanzhao Zeng, “Preferred skin colour reproduction”, PhD Thesis, Department of Colour Science, University of Leeds, September 2011.
11) Buyue Zhang and Jan P. Allebach, “Adaptive bilateral filter for sharpness enhancement and noise removal”, IEEE Transactions on Image Processing 15(5), May 2008.
12) Michael Elad, “Retinex by two bilateral filters”, Proceedins International Conference on Scale-Space Theories in Computer Vision, 2005.
13) Karen M. Braun, “Memory Color Enhancement Algorithm”, Proceedings International Congress of Imaging Science, 2006.
14) Pavel Kisilev, Doron Shaked, and Suk Hwan Lim, “Noise and signal activity maps for better imaging algorithms”, Proceedings International Conference on Image Processing, 2007.
15) Ce Liu, William T. Freeman, Richard Szeliski, and Sing Bing Kang, “Noise estimation from a single image”, Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2006.
16) Matthias Dantone, Juegen Gall, Gabriele Fanelli, and Luc Van Gool, “Real-time facial feature detection using conditional regression forests”, Proceedings Computer Vision and Pattern Recognition, 2012.
17) Ron Banner, “What makes good image composition?”, Proceedings Electronic Imaging, 2011.
18) Raanan Fattal, “Single image dehazing”, Proceedings SIGGRAPH’08, 2008.
19) YiChang Shih, Dilip Krishnan, Fredo Durand, William T. Freeman, “Reflection removal using ghosting cues”, Proceedings Computer Vision and Pattern Recognition, 2015.
20) Ruth Bergman, Hila Nachlieli, Gitit Ruckenstein, Mark Shaw, and Ranjit Bhaskar, “Perceptual segmentation: Combining image segmentation with object tagging”, HPL-2008-185. October 2008.
21) Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller, “Recognizing image style”, arxiv.org, July 2014.
22) Sicheng Zhao, Yue Gao, Xiaolei Jiang, Hongxun Yao, Tat-Seng Chua, and Xiaoshuai Sun, “Exploring principles-of-art features for image emotion recognition”, Proceedings International Conference on Multimedia, November 2014.
23) Subhabrata Bhattacharya, Rahul Sukthankar, and Mubarak Shah, “A Framework for photo-quality assessment and enhancement based on visual aesthetics”, Proceedings International Conference on Multimedia, October 2010.
24) Nicolas Bonnier and Albrecht J. Lindner, “Measurement and compensation of printer modulation transfer function”, Journal of Electronic Imaging 19(1), 2010.
25) Hua-wn Chang, Ming-hui Wang, Shu-qinq Chen, Hua Yang, and Zu-jian Huang, “Sparse feature fidelity for image quality assessment”, Proceedings International Conference on Pattern Recognition, November 2012.
26) Jan Morovic, "To develop a universal gamut mapping algorithm", PhD Thesis, University of Derby, October 1998.
27) Lihao Xu, Baiyue Zhao, and M. R. Luo, "Colour gamut mapping between small and large colour gamuts: Part 1. gamut compression", Optics Express 26(9), April 2018.
28) Yuki Koyama, Issei Sato, Daisuke Sakamoto, and Takeo Igarashi, "Sequential line search for efficient visual design optimization by crowds", Proceedings SIGGRAPH'17, 2017.
29) Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa, "Globally and locally consistent image completion", Proceedings SIGGRAPH'17, 2017.
30) Jinsun Park, Yu-Wing Tai, Donghyeon Cho, and In So Kweon, "A Unified approach of multi-scale deep and hand-crafted features for defocus estimation", Proceedings CVPR'17, 2017.
31) R. Ranjan, V. M. Patel and R. Chellappa, "HyperFace: A Deep Multi-task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence. pre-print December 2017.
32) Jongyoo Kin and Sanghoon Lee, "Deep learning of human visual sensitivity in image quality assessment framework", Proceedings CVPR'17, 2017.
33) Simone Bianco, Luigi Celona, Paolo Napoletano, Raimondo Schettini, “On the use of deep learning for blind image quality assessment”, arXiv:1602.05531v5, 2017.
34) Xin Lu, Hailin Jin, Jianchao Yang, and James Z. Wang, “Rating image aesthetics using deep learning”, IEEE Transactions on Multimedia 17(11), November 2015.
35) Jian Ren, Xiaohui Shen, Zhe Lin, Radomir Mech, and David J. Foran, “Personalized image aesthetics”, Proceedings ICCV’17, 2017.
36) Barry O Carroll, “20 Composition techniques that will improve your photos”, https://petapixel.com/2016/09/14/20-composition-techniques-will-improve-photos/, 2016.
37) Rudolf Arnhelm, “Visual Thinking”, University of California Press, 1969.
38) Bill Smith, “Designing a Photograph”, AMPHOTO Books, 2001.
39) Ian Roberts, “Mastering Composition”, Northlight Books, 2008.
40) The Editors of Eastman Kodak Company, “More Joy of Photography”, Addison-Wesley, 1981.
41) Tavis Leaf Glover, “How to edit photos like Steve McCurry: The art of removing distractions”, https://petapixel.com/2018/05/12/how-to-edit-photos-like-steve-mccurry-the-art-of-removing-distractions/, 2018.
42) Geda Bertasius, Lorenzo Torresani, Stella X. Yu, and Jianbo Shi, "Convolutional random walk networks for semantic image segmentation", Proceedings CVPR'17, 2017.
43) Yagiz Aksoy, Tunc Ozan Aydin, and Marc Pollefeys, "Designing effective inter-pixel information flow for natural image mattings", Proceedings CVPR'17, 2017.