



AESTHETIC-DRIVEN IMAGE COMPOSITION ENHANCEMENT : A SURVEY

Chao Fan

School of Computer Science and Technology, Tiangong University, China.

ABSTRACT

Composition has a great impact on image. With the development of smart phone, personal digital camera and the social sharing platforms such as Instagram, tons of photos are taken and shared by users. Everyone wants to take beautiful pictures. For professional photographers, due to their professional knowledge, such as rule of third, visual balance, diagonal dominance, they know how to adjust their camera and apply their knowledge to take aesthetically pleasing images. However, for amateurs, because of the lack of professional photographic composition knowledge, they usually don't know how to take a beautiful image. Many factors lead to aesthetically pleasing photos, one of them is composition. Good composition leads to visually pleasing image. In the recent years, image composition enhancement technics has drawn great attention to the researchers by employing the professional composition knowledge. Many methods have been proposed to improve the composition of the images, i.e., image retargeting, warping, seam carving, image cropping. Due to the popularity of deep learning, some AI technics have applied into the Image composition enhancement methods. In this paper, we present a survey of state-of-the-arts aesthetic-driven composition enhancement technics. In this paper, We first introduce the problem of the image composition enhancement problem, then we discuss the development of the computational aesthetics, next we present a comprehensive review of different image composition methods. At last, we give a conclusion after analyzing each method and present a research direction for the future. This survey intends to give the researchers a reference who wants to set foot in the field of image composition enhancement.

KEYWORDS: Image enhancement , Image re-composition, Image aesthetics, composition rules, Computer vision

1. INTRODUCTION:

With the development of smart phone, personal digital camera and the social sharing platforms such as Instagram, tons of photos are taken and shared by users. Gradually People seek visual aesthetics to take a good photograph from their digital advices. Many factors contribute to the aesthetics of one image such composition, light, contrast, one of them that contributes most to the image aesthetic is image composition. For professional photographers, due to their professional knowledge, such as rule of thirds, visual balance, diagonal dominance, they know how to adjust their camera and apply their knowledge to take an aesthetically-pleasing images. However, for amateurs, because of the lack of professional photographic composition knowledge, they usually don't know how to take a beautiful image. Thus, they often turn to some aesthetic-driven image composition tools for help hoping that these tools can enhance the aesthetics of their poorly-taken photos.

Aesthetic-driven image composition enhancement technics aims at enhancing the aesthetic of an image by improving the composition. It is a challenging problem because it needs to achieve better composition as well as preserve the important content of the original image. Moreover, image aesthetic is a very subject concept, everyone has his own opinion on whether this image is aesthetically pleasing. In this paper, We first introduce the problem of the image composition enhancement problem, then we discuss the development of the computational aesthetics, next we present a comprehensive review of different image composition methods. At last, we give a conclusion after analyzing each method and present a research direction for the future.

2. BACKGROUND:

2.1 Image composition enhancement:

The image composition enhancement problem which is also known as the image re-composition problem can be formulated as follows: given an image I with its important content and a set of visual composition constraints, the goal is to modify the image composition to produce an output image I' with better aesthetics quality (Islam, Lai-Kuan, & Chee-Onn, 2016).

2.2 Computational aesthetics:

Computational aesthetics is an emerging research direction that emerges when artificial intelligence, aesthetics, psychology and other disciplines have developed to a certain degree. Its purpose is to make computers have the ability to discover, transform, and create beauty similar to humans, and to study the internal mechanism of human perception of beauty. This research direction often includes these fields: image aesthetic assessment, image caption and image aesthetic enhancement.

2.2.1 Image aesthetic assessment:

Image aesthetic evaluation is usually classified as a classification or regression problem, where the image content is mapped to the aesthetic level provided by the human annotator (Kong, Shen, Lin, Mech, & Fowlkes, 2016; Lu, Lin, Jin, Yang, & Wang, 2014). Image aesthetic assessment has many applications, such as image retrieval system and image editing tools (Ni et al., 2013). Existing methods can be classified into two types: hand-crafted

features based methods and learning based methods. In the hand-crafted features based methods, these features are designed under the guidance of photography and psychological rules such as the rule of thirds composition, depth of field (DOF) and color (Datta, Joshi, Li, & Wang, 2006; Tong, Li, Zhang, He, & Zhang, 2004). In the learning-based methods, they apply deep learning in image aesthetic assessment (Jin et al., 2016; Lu et al., 2014; Ma et al., 2019; Z. Wang et al., 2017). Lu et al. (Lu et al., 2014) proposed the Rating Pictorial Aesthetics using Deep Learning (RAPID) model, with impressive accuracies on the Aesthetic Visual Analysis (AVA) dataset (Murray, Marchesotti, & Perronnin, 2012). Wang et al. (Z. Wang et al., 2017) develop a novel deep-learning based image aesthetics assessment model, called Deep Chatterjee's Machine (DCM), to learn multiple attributes on a variety of selected feature dimensions. Those attributes are then associated and transformed into the overall aesthetic rating, by a high-level synthesis network.

2.2.2 Image caption:

Image caption technics takes an image as an input then generates the comments of this picture. It can enable the computing devices the ability to understand the aesthetic of an image. The goal is to train an artificial intelligence (AI) system being able to perceive aesthetics as human (K. Chang, Lu, & Chen, 2017; W. Wang, Yang, Zhang, & Zhang, 2018). Chang et al. (K. Chang et al., 2017) proposed two approaches called the aspect oriented (AO) approach and the aspect fusion (AF) approach to solve the problem. Those two approaches use CNN-LSTM model to create a photo captioning system for each aspect. The first method uses the CNN model to select the most interesting aspect of the input image while the second method fuses the captions learned from the individual aspects to create a new caption.

2.2.3 Image aesthetic enhancement:

Image enhancement is considered a skillful artwork that involves transforming or altering a photograph using various methods and techniques to improve the aesthetics of a photo (Deng, Loy, & Tang, 2017).

It usually enhances the light, contrast and composition to make the input picture looks visually pleasing. Current works have shown some success in automating color enhancement (Bhattacharya, Sukthankar, & Shah, 2010; Lee, Sunkavalli, Lin, Shen, & Kweon, 2016; Sun, Chao, Kuo, & Hsu, 2017; Yan, Lin, Kang, & Tang, 2014), style transfer (Gupta, Johnson, Alahi, & Feifei, 2017; Huang & Belongie, 2017; Johnson, Alahi, & Feifei, 2016) and image composition enhancement.

3. IMAGE COMPOSITION ENHANCEMENT METHODS:

Aesthetic-driven image composition technics can be easily classified into three categories: rule-based methods, learning-based methods and example-based methods. The first category mainly applies the common composition rules such as rule of third, visual balance and diagonal dominance into the composition enhancement problem to achieve better composition. The second category usually models the composition features from large data and then applies these features to generate more visually pleasing images. The last category uses an image

as an example and the input image will be adjusted according to the composition of the reference image.

3.1 Rules-based methods:

Rules-based methods (Guo, Liu, Gu, & Wang, 2012; Jeong & Cho, 2015; D. S.-M. Liu & Huang, 2017; L. Liu, Chen, Wolf, & Cohenor, 2010; Riaz, Park, & Lee, 2015; W. N. Wang, Liu, Xu, Jiang, & Wang, 2015; Wong & Wong, 2012; F.-L. Zhang, Wang, & Hu, 2013) often apply different rules such as rule of thirds, diagonal dominance, visual balance and visual dominance to adjust the image composition. These methods can use cropping, warping or retargeting techniques to process the image.

Liu et al. (L. Liu et al., 2010) developed a novel computational means for evaluating the composition aesthetics of a given image based on measuring several well-grounded composition guidelines. A compound operator of crop-and-retarget is employed to change the relative position of salient regions in the image and thus to modify the composition aesthetics of the image. Fig. 1 shows the results. Wong et al. (Wong & Wong, 2012) presented a semi-automatic photographic re-composition approach that employs a semantics-preserving warp of the input image to enhance the visual dominance of the main subjects. Their method uses the image warping method to shift the subjects against the background (and vice versa), so that their visual dominance is improved, and yet preserve the desired spatial semantics between the subjects and the background. The re-composition is guided by a measure of the resulting visual dominance of the main subjects. Zhang et al. (F.-L. Zhang et al., 2013) presented a novel method to analyze the dependence between regions and objects in images, which considers both photographic and psychological impact; they take distance from power points, distance from diagonal lines, visual balance, relevance of objects into account. Guo et al. (Guo et al., 2012) presented a new algorithm which not only consider the rule-of-thirds into their methods but also visual similarity between the optimized image and the original one. Most methods can only optimize the image with single main object, while Liu et al. (D. S.-M. Liu & Huang, 2017) focused in group photo. Their approach can use suitable retargeting techniques, and coordinate the composition of original photos to make photos conform the aesthetic rules including rule-of-thirds and diagonal dominance.

However, existing rule-based methods cannot take all the composition rules into consideration and don't have an universal method to solve different categories of images.

3.2 Learning-based methods:

Due to the development of the deep learning technic, researchers start to use deep learning to achieve a better composition. Park et al. (Park, Lee, Tai, & Kweon, 2012) were the first to adopt the learning-based approach to learn a photo composition model and used it to find the optimal cropping window. they trained a normalized saliency map from visually pleasurable photos taken by professional photographers. They use Principal Component Analysis (PCA) to analyze training data and build a Gaussian mixture model (GMM) to describe the photo composition model. They can get different results by using different dataset. Fig.2 shows the experiment result.

Deng et al. (Deng et al., 2017) adopted the generative adversarial networks to solve the image enhancement problem. They introduced an adversarial learning-based model called EnhanceGAN that performs automatic image enhancement. Traditional image enhancement frameworks typically involve training models in a fully-supervised manner, which require expensive annotations in the form of aligned image pairs. In contrast to these approaches, their proposed EnhanceGAN only requires weak supervision (binary labels on image aesthetic

quality) and is able to learn enhancement operators for the task of aesthetic-based image enhancement.

Li et al. (Li, Wu, Zhang, & Huang, 2019) formulated the automatic image cropping problem as a sequential decision-making process and proposed a framework without the supervision of bounding box, namely Fast Aesthetics-Aware Adversarial Reinforcement Learning (FastA3RL). It contains the reinforcement learning (RL) based cropping network and the aesthetics evaluation network which is much more efficient than previous sliding window based methods.

Due to most of the learning-based methods use cropping to process images, they have the disadvantages of content loss. Another disadvantage is that it takes time to train a learned model because they need large number of images as training data.

3.3 Example-based methods:

Example-based methods take an image as a reference to guide the image composition enhancement. Chang et al. (H. Chang, Pan, Wang, & Chen, 2015) proposed a novel approach for performing joint re-composition and retargeting of photographic images (R2P). Given a reference image of interest, their method is able to automatically alter the composition of the input source image accordingly, while the recomposed output will be jointly retargeted to fit the reference. Fig.3 shows the result of this methods. This was achieved by recomposing the visual components of the source image via graph matching, followed by solving a constrained mesh-warping based optimization problem for retargeting to apply image composition enhancement to portrait, more aspect should be considered. Zhang et al. (X. Zhang, Constable, Chan, Yu, & Junyan, 2018) studied the composition in portrait paintings and proposed an algorithm to improve the composition of portrait photographs based on an example portrait painting. Fig. 4 shows the results. This method can achieve a better result for portrait than using rule-based methods.

These example-based methods often need to retrieve the reference image in database, it is also a challenging problem to retrieve an image which is composition-closely to the input images.

CONCLUSION:

In this paper, We first introduce the problem of the image composition enhancement problem, then we discuss the development of the computational aesthetics, next we present a comprehensive review of different image composition methods. Analysis shows that each method has its own advantages and disadvantages in the flexibility of reorganization and the ability to retain image features.

Since aesthetics is a very subjective concept, the image composition enhancement is a very difficult problem. In the current method, the rule-based method cannot comprehensively use multiple composition rules to optimize the image; in the learning-based method, there is a time-consuming problem and content loss problem; in the example-based method, it is challenging to retrieve an image which is composition-closely to the input images. In future work, we need to integrate multiple composition methods and use deep learning methods to learn the composition characteristics of the picture, so as to generate a better composition and preserve the semantic information and geometric information of the original picture as much as possible. We also need to develop an universal problem which can be applied to different kinds of images. At the same time, we should explore how to combine aesthetic composition with the problem of image generation in the field of image generation that is currently more popular, so as to be able to generate a more beautiful image in composition.

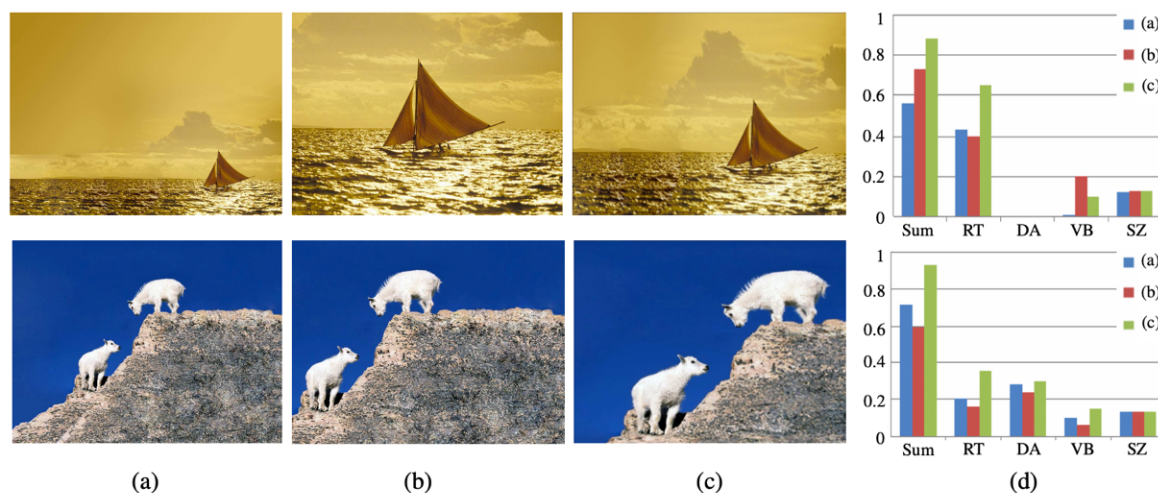


Fig.1: Results of aesthetic composition. (a) The original images; (b) an arbitrary cropping frame of (a); (c) the aesthetic composition result by our approach; (d) the aesthetic scores of (a), (b), and (c).



Fig. 2: Experimental results on image re-arrangement. First row: Input images. Second row: Corresponding saliency map. Third row: Results from Liu et al. (L. Liu et al.). Fourth row: Results using our approach. Fifth row: Saliency map of re-arranged photo

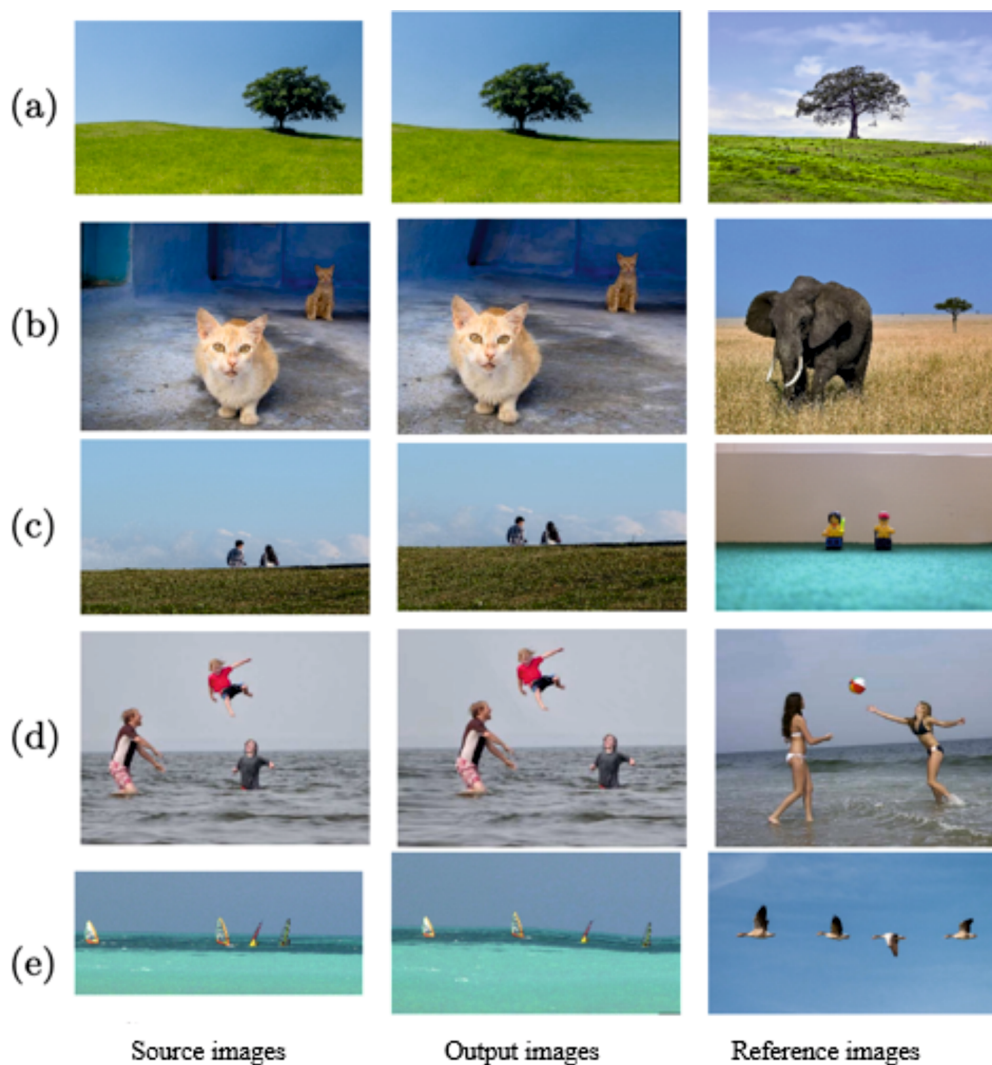


Fig. 3: Examples of image composition enhancement and retargeting with different foreground object numbers.

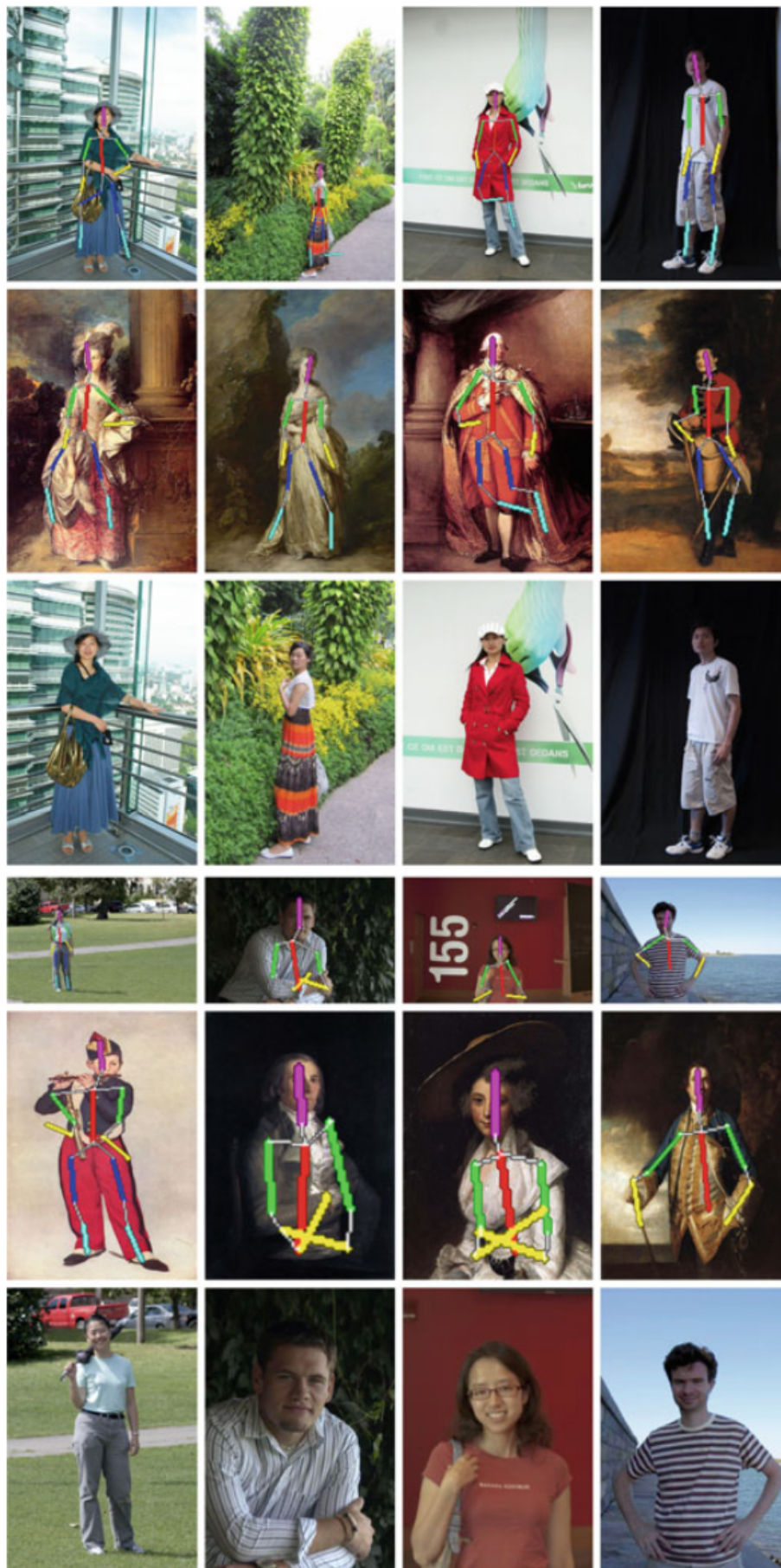


Fig. 4: Experimental results. First and fourth rows: input photographs with pose. Second and fifth rows: selected example portrait paintings with pose (CC Image courtesy The Athenaeum, Wikimedia Commons, <https://www.wikiart.org>, <http://www.artelista.com/>). Third and sixth rows: cropping results using the proposed method

REFERENCES:

- I. Bhattacharya, S., Sukthankar, R., & Shah, M. (2010). A framework for photo-quality assessment and enhancement based on visual aesthetics. Paper presented at the acm multimedia.
- II. Chang, H., Pan, P., Wang, Y. F., & Chen, M. (2015). R2P: Recomposition and Retargeting of Photographic Images. Paper presented at the acm multimedia.
- III. Chang, K., Lu, K., & Chen, C. (2017). Aesthetic Critiques Generation for Photos. Paper presented at the international conference on computer vision.
- IV. Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. Paper presented at the european conference on computer vision.
- V. Deng, Y., Loy, C. C., & Tang, X. (2017). Aesthetic-Driven Image Enhancement by Adversarial Learning. arXiv: Computer Vision and Pattern Recognition.
- VI. Guo, Y. W., Liu, M., Gu, T. T., & Wang, W. P. (2012). Improving Photo Composition Elegantly: Considering Image Similarity During Composition Optimization. *Computer Graphics Forum*, 31(7), 2193-2202. doi:10.1111/j.1467-8659.2012.03212.x
- VII. Gupta, A., Johnson, J., Alahi, A., & Feifei, L. (2017). Characterizing and Improving Stability in Neural Style Transfer. Paper presented at the international conference on computer vision.
- VIII. Huang, X., & Belongie, S. (2017). Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization. Paper presented at the international conference on computer vision.
- IX. Islam, M. B., Lai-Kuan, W., & Chee-Onn, W. (2016). A survey of aesthetics-driven image recomposition. *Multimedia Tools and Applications*, 76(7), 9517-9542. doi:10.1007/s11042-016-3561-5
- X. Jeong, K., & Cho, H.-J. (2015). A Digitalized Recomposition Technique Based on Photo Quality Evaluation Criteria. *Wireless Personal Communications*, 86(1), 301-314. doi:10.1007/s11277-015-2977-y
- XI. Jin, X., Chi, J., Peng, S., Tian, Y., Ye, C., & Li, X. (2016). Deep image aesthetics classification using inception modules and fine-tuning connected layer. Paper presented at the international conference on wireless communications and signal processing.
- XII. Johnson, J., Alahi, A., & Feifei, L. (2016). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. Paper presented at the european conference on computer vision.
- XIII. Kong, S., Shen, X., Lin, Z., Mech, R., & Fowlkes, C. C. (2016). Photo Aesthetics Ranking Network with Attributes and Content Adaptation. Paper presented at the european conference on computer vision.
- XIV. Lee, J., Sunkavalli, K., Lin, Z., Shen, X., & Kweon, I. S. (2016). Automatic Content-Aware Color and Tone Stylization. Paper presented at the computer vision and pattern recognition.
- XV. Li, D., Wu, H., Zhang, J., & Huang, K. (2019). Fast A3RL: Aesthetics-Aware Adversarial Reinforcement Learning for Image Cropping. *IEEE Transactions on Image Processing*, 28(10), 5105-5120.
- XVI. Liu, D. S.-M., & Huang, C.-C. (2017). Optimizing Aesthetic-Based Photo Retargeting. In *Smart Graphics* (pp. 41-60).
- XVII. Liu, L., Chen, R., Wolf, L., & Cohenor, D. (2010). Optimizing Photo Composition. *Computer Graphics Forum*, 29(2), 469-478.
- XVIII. Lu, X., Lin, Z., Jin, H., Yang, J., & Wang, J. Z. (2014). RAPID: Rating Pictorial Aesthetics using Deep Learning. Paper presented at the acm multimedia.
- XIX. Ma, N., Volkov, A., Livshits, A., Pietrusinski, P., Hu, H., & Bolin, M. (2019). An Universal Image Attractiveness Ranking Framework. Paper presented at the workshop on applications of computer vision.
- XX. Murray, N., Marchesotti, L., & Perronnin, F. (2012). AVA: A large-scale database for aesthetic visual analysis. Paper presented at the computer vision and pattern recognition.
- XXI. Ni, B., Xu, M., Cheng, B., Wang, M., Yan, S., & Tian, Q. (2013). Learning to Photograph: A Compositional Perspective. *IEEE Transactions on Multimedia*, 15(5), 1138-1151. doi:10.1109/tmm.2013.2241042
- XXII. Park, J., Lee, J., Tai, Y., & Kweon, I. S. (2012). Modeling photo composition and its application to photo re-arrangement. Paper presented at the international conference on image processing.
- XXIII. Riaz, S., Park, U., & Lee, S.-W. (2015). A photograph reconstruction by object retargeting for better composition. *Multimedia Tools and Applications*, 75(23), 16439-16460. doi:10.1007/s11042-015-3037-z
- XXIV. Sun, W., Chao, T., Kuo, Y., & Hsu, W. H. (2017). Photo Filter Recommendation by Category-Aware Aesthetic Learning. *IEEE Transactions on Multimedia*, 19(8), 1870-1880.
- XXV. Tong, H., Li, M., Zhang, H., He, J., & Zhang, C. (2004). Classification of digital photos taken by photographers or home users. Paper presented at the advances in multimedia.
- XXVI. Wang, W., Yang, S., Zhang, W., & Zhang, J. (2018). Neural Aesthetic Image Reviewer. arXiv: Computer Vision and Pattern Recognition.
- XXVII. Wang, W. N., Liu, J. C., Xu, X. M., Jiang, Y. Z., & Wang, L. (2015). Aesthetic Enhancement of Images Based on Photography Composition Guidelines. *Journal of South China University of Technology(Natural ence Edition)*, 43(5), 51-58.
- XXVIII. Wang, Z., Liu, D., Chang, S., Dolcos, F., Beck, D. M., & Huang, T. S. (2017). Image aesthetics assessment using Deep Chatterjee's machine. Paper presented at the international joint conference on neural network.
- XXIX. Wong, L., & Wong, K. (2012). Enhancing visual dominance by semantics-preserving image recomposition. Paper presented at the acm multimedia.
- XXX. Yan, J., Lin, S., Kang, S. B., & Tang, X. (2014). A Learning-to-Rank Approach for Image Color Enhancement. Paper presented at the computer vision and pattern recognition.
- XXXI. Zhang, F.-L., Wang, M., & Hu, S.-M. (2013). Aesthetic Image Enhancement by Dependence-Aware Object Recomposition. *IEEE Transactions on Multimedia*, 15(7), 1480-1490. doi:10.1109/tmm.2013.2268051
- XXXII. Zhang, X., Constable, M., Chan, K. L., Yu, J., & Junyan, W. (2018). Composition Improvement for Portrait Photographs. In *Computational Approaches in the Transfer of Aesthetic Values from Paintings to Photographs* (pp. 137-153).