Computational Visual Media
DOI 10.1007/s41095-xxx-xxxx-x Vol. x, No. x, month year, xx–xx

Research

Non-dominated sorting based multi-page photo collage

Yu Song[1,2,4], **Fan Tang**[3], **Weiming Dong**[1,2,4](✉), Changsheng Xu[1,2,4]

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The development of social networking services (SNSs) revealed a surge in image sharing. The sharing mode of multi-page photo collage (MPC), which posts several image collages at a time, can often be observed on many social network platforms, which enables uploading images and arrangement in a logical order. This study focuses on the construction of MPC for an image collection and its formulation as an issue of joint optimization, which involves not only the arrangement in a single collage but also the arrangement among different collages. Novel balance-aware measurements, which merge graphic features and psychological achievements, are introduced. Non-dominated sorting genetic algorithm is adopted to optimize the MPC guided by the measurements. Experiments demonstrate that the proposed method can lead to diverse, visually pleasant, and logically clear MPC results, which are comparable to manually designed MPC results.

Keywords siamese network; single object tracking; anchor-free; center point detection.

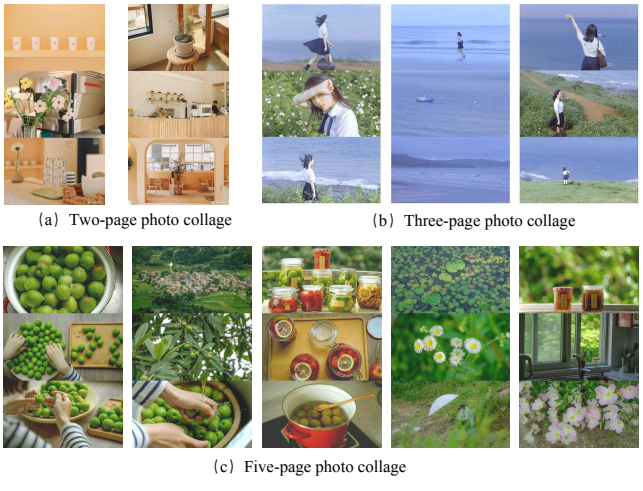

(a) Two-page photo collage (b) Three-page photo collage (c) Five-page photo collage

Fig. 1 Multi-page photo collage examples on social media.

1 Introduction

The popularity of digital cameras and mobile devices has provided people with the opportunity to take photos easily and record their lives. The limit on the number of uploaded images contributes to the difficulty in sharing a large image collection on social networks at once (for example, 10 for Facebook and 9 for WeChat). Moreover, viewing a set of images one by one may be time consuming and may contribute to the difficulty in helping the viewer obtain the main information of the images. Many users choose to post the images in a multi-page photo collage (MPC) mode to address the aforementioned problems by arranging the images into multiple collages with simple templates and posting the collages to the social media as shown in Fig. 1. Therefore, photos can be uploaded on social media within the limited numbers, and viewers can easily obtain information regarding the image collection.

Collages generally comprise images along the horizontal and vertical directions to enlarge the collage volume effectively and maintain the collage in a simple and straightforward organizational format as shown in Fig. 1. However, ensuring the visually pleasing appearance and effective information transmission of collages in a disorganized arrangement is difficult. The design of MPC is time and energy consuming for non-professional users.

Image collage methods are carefully designed to manage and explore the visualization of large image collections on a canvas [17, 18, 20, 33]. These methods focus on information aggregation and presentation to help users

1 NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, 100040, China. E-mail: {songyu2017, weiming.dong, changsheng.xu}@ia.ac.cn
2 School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100190, China.
3 School of Artificial Intelligence, Jilin University, Changchun, 130012, China. E-mail: tangfan@jlu.edu.cn
4 CASIA-LLvision Joint Lab, Beijing, 100190, China.
Manuscript received: 2020-XX-X; accepted: 2020-XX-XX.

analyze image collections. The layout of the collage is usually automatically generated. Each sub-image corresponds to a part of the canvas by assigning a position coordinate, a rotation angle, and a scale factor. However, these methods are used to allocate all images onto one canvas, and the layout among collages varies. Therefore, the final collage is complicated and suffers huge sub-module amounts and serious sub-module size variations. Moreover, collaging all images on one canvas may lead to an information explosion for social media sharing, and the frequent changing of collage layouts may add difficulty to the acquisition of information among different collages.

Album design is a visualization form that merges the image assignment and collage. These methods assign a certain number of images to each page and provide a configuration that can help decide the image appearance. However, existing research trends on the design of album pages have blank spaces and image rotations, which will decrease the delivery effectiveness and legibility of information. These techniques can hardly works well on the MPC.

This paper focuses on the issue of this application scenario. The largest challenge of the MPC is to determine the deep relationship between the entire collage choreography and user experience. Psychologists explored the inner connection between the element characters of the image and the visual experience of users. They revealed that people perceive elements gathered together as a whole rather than individuals [29]. Meaningful conclusions (e.g., the visual appreciation of a picture largely depends on the perceptual balance of its elements [10], and a positive relationship exists between balance and liking for a multi-element picture [11]) have been drawn. Most psychological literature refers to the assessment of preference for balance (APB) [30] and deviation of the center of mass (DCM) [10] to measure the character of balance. However, these measures cannot provide sufficient guidance for MPC due to the lack of consideration for high-level features for every single page and coordination of images among multi-pages, which are important for evaluating the MPC to achieve balance perception.

The visual data of social medias are characterized by their large scale, user data are of considerable difference, and the updating rhythm is of high speed. Therefore, the page numbers of MPC cannot be excessive, and the layout of each page cannot be substantially complicated. The task is split into the following three issues: assignment of images to separated pages, arrangement of the images in an exact page, and resolution of conflicts between the internal optimization of a page and optimization between pages.

Novel objective functions based on information gathering, which measure the compactness degree of image information within a page and the dispersion degree within pages, are proposed to solve the first issue. The engaged measure is formulated by refining the high-level semantic information of images and combining the image features of color, content, and object size. A novel uniqueness function is proposed to solve the second issue, and the collage per page is established in a balanced form at the feature level. As a process of multi-objective optimization, a non-dominated sorting genetic algorithm-II (NSGA-II) [5] is employed to solve the third issue, and the measurement of single and multiple collage pages is treated as equally important levels. A genetic algorithm is utilized to find a solution set that effectively balances all optimization objectives. Overall, the major contributions of this work are as follows:

- The formulation of MPC, which is a fresh new type of the multi-image joint presence task and a new form of photo sharing on social media.
- Novel image scatter measurements based on information aggregation, which formulate not only the evaluation within a collage page but also among multiple collage pages, are proposed.
- A new form of image balance at the image feature level and a uniqueness function is proposed, which can lead to visually pleasing image arrangements on a page.
- An optimal strategy based on NSGA-II , which jointly applies the two aforementioned measures to construct image distribution with clear outlines of information, is conducted.

2 Related Work

2.1 Image Collage

Image collage is a traditional technique for multi-image presence, which lays multiple images on one canvas. Highlighting salient information and speeding up browsing of an image set are the common core tasks. Tan et al. [27] arranged images globally by graph formulation and refined the final result by Voronoi tessellation. Cao et al. [1] focused on stylized comic layout generation through user-specified semantics artwork dataset. They further utilized a probabilistic graphical model to synthesize comic elements, which aims to guide the attention of the reader [2]. Chen et al. [3] leveraged the image collage techniques to visualize the video summarization in a static way. Yu et al. [33] took the photo collage as a circle packing issue to show the salient areas of each image. Jing et al. [12] extracted video key frames and arranged them with manga-style layout based on salience. Han et al. [9] collaged images into different interesting shapes via a tree-based approach.

Liu et al. [18] generated compact collages by splitting the canvas into irregular partitions. Liang et al. [17] proposed sample-based image arrangement and Voronoi tree map-based layout generation methods. Wu et al. [31] generated image correlation preserved collages by binary tree-based page analysis. Zheng et al. [35] used a deep generative model for sketching with user semantics information input. Pan et al. [20] visualized image collection summarization and generated the layout via tree-based image content analysis. Gan et al. [8] managed albums in a comic-like layout based on image classification.

These collage methods can have an outstanding effect on image collection information arrangement. However, placing images on one canvas led to the final result with a complex layout, thus lacking application in daily life. Some works focus on application scenarios with the popularity of mobile devices. Kong et al. [13] proposed a strategy to collage phone images into a centroidal Voronoi diagram. Song et al. [25] introduced a balance-based photo-posting strategy, which is tailored for the scenario in social media.

Overall, the collage research trends to design methods based on real application scenarios such as mobile devices and social platforms. This paper focuses on the task of MPC design, which is a comprehensive exploration of the presence of collages as well as a real form of daily image sharing with realistic application scenarios.

2.2 Multi-Objective Optimization

Multi-objective optimization (MOO) is a common problem for most research. Multi-angle analysis of an issue often introduces multiple optimization objectives under the constraint of conflicts. Optimizing the objectives simultaneously is difficult. Instead, some solutions, namely Pareto-optimal solutions, can optimize one objective without degrading the others. Many multi-objective evolutionary algorithms are used to search approximations of the Pareto-optimal solution and applied to several areas actively [5–7]. Li et al. [15] generated a realistic crowd model by considering both static and dynamic features. Lee et al. [14] modeled free-form surfaces by optimizing multiple objectives of design goal and cost efficiency to reduce the usage of curved panels. Rejeesh et al. [21] proposed a denoising algorithm based on optimal trilateral filtering. Su et al. [26] solved the super-resolution problem of a single image by merging the MOO strategy into GAN training. This paper employs the non-dominated sorting genetic algorithm-II (NSGA-II) to search for solutions (one multi-objective optimization method). The operator of the genetic algorithm is utilized to generate the solutions, and non-dominated sorting is utilized to select the excellent solution.

Fig. 2 Workflow of MPC generation.

3 Measurement Formulation

This paper proposed a new formulation tailored for MPC considering the demand for a clear information outline and visually pleasing experience. Numerous studies have worked on image quality [4, 16, 19, 34] or aesthetics assessment [22, 23, 32]. Meanwhile, all the methods are designed for single images, and the measurement of MPC is still lacking. The MPC is decomposed as a two-step task, namely image scatter to pages and image arrangement per page, as shown in Fig. 2. Specific measures, which can jointly contribute to the final generation of collages, have been designed for each step.

3.1 Feature Characterization

The feature item in this paper includes three kinds: color, content, and size. ***Color*** sets the tone of images and is characterized as dominant colors in RGB space. Each image is described by three dominant color descriptors and represented by the symbol fea_{col}. ***Content***, which is the core information carrier of an image, is characterized as the last layer output of VGG-16 [24] that is pre-trained on ImageNet.The content item is represented by the symbol fea_{con}. ***Size*** is an important descriptor for image objects and characterized as "Saliency Map" which is computed and represented by the symbol fea_{size}. Overall, three kinds of characterizations are utilized to describe image from a different angle (fea_{col}, fea_{con}, fea_{size}) and processed as image information in the following calculation.

3.2 Image Scatter to Pages

Image scattering is the chief task for a photo collage, and the information difference is the key factor. Gathering similar images into a page can illustrate a clear information outline among pages, which can ease viewer browsing, especially for the image set with serious information variation. The scattering quality is measured by determining

the similarity and dissimilarity degrees of images on a page and those between pages, respectively.

3.2.1 Measurement in a Page

The inner measurement of pages aims to determine if the images within a page have similar information. The value is calculated by feature subtraction among images, which is implemented as follows:

$$Mea_{col_in} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{P}} |fea_{col}(I_m) - fea_{col}(I_n)|, \quad (1)$$

$$Mea_{con_in} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{P}} |fea_{con}(I_m) - fea_{con}(I_n)|, \quad (2)$$

$$Mea_{size_in} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{P}} |fea_{size}(I_m) - fea_{size}(I_n)|, \quad (3)$$

where $\mathbb{P}$ is the set of images in a page; Mea_{col_in}, Mea_{con_in}, and Mea_{size_in} are the page inner measurements of color, content and size, respectively; $fea_{col}(I_m)$, $fea_{con}(I_m)$ and $fea_{size}(I_m)$ are the feature values of image I_m's color, content, and size, respectively. Overall, the entire page inner measurement can be written as follows:

$$Mea_{in} = Mea_{col_in} + Mea_{con_in} + Mea_{size_in}. \quad (4)$$

3.2.2 Measurement between Pages

The outer measurement of pages aims to determine the image difference among pages. The feature difference of all pictures that are not on the same page is summed as the page outer measurement, and the implement details are as follows:

$$Mea_{col_out} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{S}\cap\mathbb{P}} |fea_{col}(I_m) - fea_{col}(I_n)|, \quad (5)$$

$$Mea_{con_out} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{S}\cap\mathbb{P}} |fea_{con}(I_m) - fea_{con}(I_n)|, \quad (6)$$

$$Mea_{size_out} = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{S}\cap\mathbb{P}} |fea_{size}(I_m) - fea_{size}(I_n)|, \quad (7)$$

where $\mathbb{S}$ is the set of all images, and $\mathbb{P}$ is the set of images in a page; Mea_{col_out}, Mea_{con_out}, and Mea_{size_out} are the respective measurements of color, content and size between pages; $fea_{col}(I_m)$, $fea_{con}(I_m)$, and $fea_{size}(I_m)$ are the feature values of image I_m's color, content, and size,respectively. Overall, the entire measurement between pages can be written as follows:

$$Mea_{out} = Mea_{col_out} + Mea_{con_out} + Mea_{size_out} \quad (8)$$

The final objective functions can then be written as:

$$\arg\min Mea_{in}\,(\mathbb{S}, \mathbb{P}_1, ..., \mathbb{P}_M), \quad (9)$$

$$\arg\max Mea_{out}\,(\mathbb{S}, \mathbb{P}_1, ..., \mathbb{P}_M), \quad (10)$$

which is an issue of multi-objective optimization, and M is the collage page number.

3.3 Image Arrangement per Page

The image arrangement for each page is the final decision area when page scattering of images is decided. Researchers have found that users prefer the image with "balanced" characteristics [10, 28]. Thus, a balance formulation at the image level and a uniqueness mechanism are proposed to guide the image arrangement per page.

The most unique image is generally laid on the most evident location. The uniqueness mechanism describes the difference degree between one image and other images on a page, which is calculated as follows:

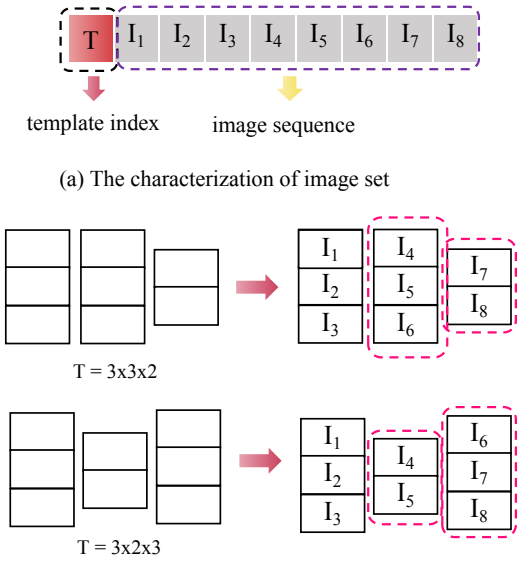
$$\begin{aligned} Uq(I_m) = \sum_{m}^{\mathbb{P}} \sum_{n}^{\mathbb{P}} & |\,(fea_{col}(I_m) - fea_{col}(I_n))\,| \\ & +|\,(fea_{con}(I_m) - fea_{con}(I_n))\,| \\ & +|\,(fea_{size}(I_m) - fea_{size}(I_n))\,|, \end{aligned} \quad (11)$$

Thus, each image for a determined page owns a uniqueness value, and the unique image owns the largest value. The unique image is presented in this paper to the page center location (if the image number is even, then this step is skipped due to the absence of center location). Then, the similarity between the remaining images is calculated. The two images are closer than any other images on the page and will be set to a pair.

Images on a page under the result of the unique image and image pairs are arranged as follows. First, the unique image is put in the center. Second, the paired image is placed at the symmetric position on the page. Therefore, the image arrangement results in a page can naturally be clear information, and the image-level balance can lead to a visually pleasing result.

4 Approach

The approach design faces two challenges. The first challenge is the solution space design. A solution must correspond to an MPC result, of which the page number, the image scatter result, and the image arrangement must be accurately expressed. The second challenge is the guidance design, which can help search for the final result in the solution space. For the first challenge, the MPC is represented by introducing predefined templates. By contrast, the MPC for the second challenge is half solved by Sec. 3, which can be used as the guidance. However, two objective functions are available, and dealing with the two associated and exclusive objective functions is the core link. The NSGA-II, which can directly leverage the solution representation form and improve the objective value without degrading the other objective values, is adopted for the aforementioned challenge. For clarity, this section will discuss template definition and NSGA-II.



(a) The characterization of image set

(b) The mapping rule between the characterization and actual collages

Fig. 3 Illustration of image set characterization and the mapping rule of templates. The collages in pink dash boxes are different.

4.1 Image Set Characterization and Templates Definition

This paper characterizes the image set first and equips it with an MPC template to achieve the detailed distribution of the image set. Each image set is characterized as an image sequence. The template is set in accordance with the total number of images in the image set, that is, the sum of the number of images on a single page in a template should meet the capacity of the image set. Given an exact image set, every possible configuration (how many pages and images per page) is seen as a template and all templates form the template space. Specifically, the template mentioned herein is sequential. That is, as shown in Fig. 3, the results generated by the same image sequence but different templates, namely $3 \times 3 \times 2$ and $3 \times 2 \times 3$ will be different. Notably, each page is in full-line collage (learned from the existing multi-page form Fig. 1), and all templates are designed under the assumption of the presence of upper limits for the total page number and the image number per page. Taking the image set containing eight images as an example, 47 templates in total are under the page number upper limit 4 and the image number in a page upper limit 4. An MPC is specific when the image set is sequentially characterized and a template is given. The final characterization for the MPC is presented as a sequence that starts with the template index and follows the image set characterization as shown in Fig. 3.

4.2 Non-dominated Sorting Genetic Algorithm-II

As previously mentioned, the MPC is an issue of multi-objective optimization, which is difficult to solve by common methods directly. This paper employs the NSGA-II to search the solution space. In this algorithm, the operator of the genetic algorithm generates candidate solutions, and the non-dominated sorting mechanism selects excellent solutions. Notably, the search space includes two parts: template and image space. The MPC can be represented according to Sec. 4.1, in which the sequential characterization naturally matches the gene appearance. The measurement can be calculated in accordance with Sec. 3, in which the value naturally matches the function of the adaption value.

4.2.1 Genetic Operators

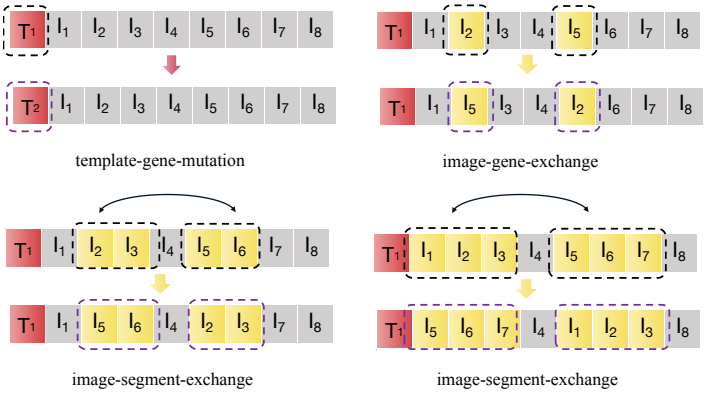


Fig. 4 Mechanism of genetic algorithm.

Gene mutations and the gene segment exchange can effectively search the existing solution space rounds. This paper presents two gene types, namely template and image genes, and the gene operator comprises three types: template gene mutation, image gene exchange, and image segment exchange as shown in Fig. 4.

Template-gene-mutation The template gene mutation is an operation that can test additional templates on the identical image arrangement. The template gene can mutate into another template index according to the probability.

Image-gene-exchange Instead of image gene mutation, the image gene exchange is adopted due to the feature of the image set, which indicates that images should be different from each other. The two image genes can exchange the location according to the probability.

Image-segment-exchange Exchange from one image exchange to image segment exchange is enlarged

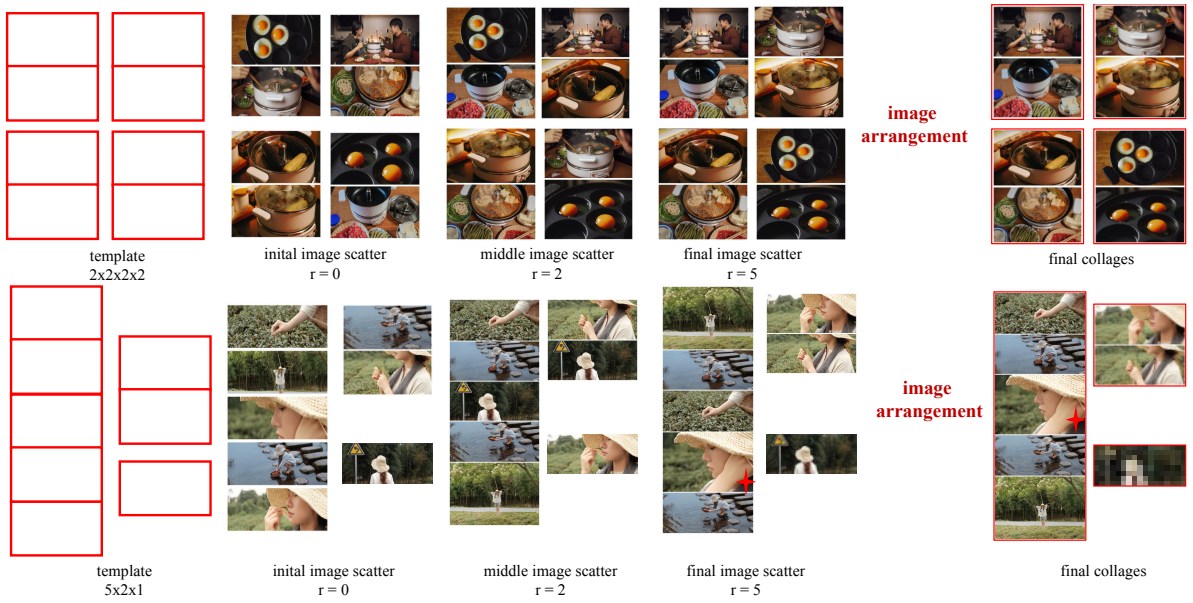


Fig. 5 Pipeline of template-fixed collage generation. The fixed template is sketched out with red boxes, and the uniqueness image is labeled with red star. As the number of iterations (r) increases, the collages gradually becomes clear and tidy information, which leads to a visually easy and pleasant experience.

considering the efficiency of operators. The two (or three) image gene segments can exchange the location according to the probability.

4.2.2 Non-dominated Sorting

The evolution of genetic algorithms is generally based on the determined fitness function values. Directly selecting generation by a determined value is impractical for the different optimization directions of different objective functions. The non-dominated sorting strategy is adopted in the select generated solutions [7]. Non-dominated sorting hierarchical solution space is conducted by judging the relationship between different solutions. If all the objective function values of a solution a is no more than the corresponding function values of the solution b, then solution a dominates solution b. Therefore, the non-dominated solution means that the solution is comprehensive, which demonstrates a good performance among all objectives.

5 Experiments

First, the previously proposed measurement is analyzed to discuss the effects of the coefficient on the result. Then, three experiments are designed to examine the performance of the proposed method in different settings. A user study is also taken to examine the differences in results between the proposed method and human beings. Each experiment has two results: image scatter to pages and image arrangements per page. The collage collection with eight images is used for the convenience of illustration. The image collection is randomly shuffled 10 times, and each shuffle result is equipped with a candidate template. Therefore, the initialization of the image sequence is obtained for the experiments.

5.1 Coefficient Analysis

Three items are involved in the definition of MPC measurement, namely content, size, and color, which are given equal importance. Experiments with item bias, which is unfolded into ablation study and enhancement experiments, are conducted for further insights.

5.1.1 Ablation Study

The ablation study comprises two steps: single and two items for the measurement. Particularly, number sequences are used as the representation of items; that is, 1, 0, and 0 respectively indicate content, size, and color. Only the content items are left, and the coefficient is 1. The collages in Fig. 5 are the example collages. These collages increasingly become visually pleasing as the number of iterations rises.

Single Item Only one item is left for the final measurement in this experiment. The measurement values are recorded in Tab. 1, and the outliers are marked in red. First, the measurement slightly varies and numerous same values appear when only one item exists. Second, the final iterate collages have many contradictory values. Take template $5 \times 2 \times 1$ for example, Mea_out value is almost unchanged in (0 0 1) mode, and the final value at $r = 5$ is smaller than

Tab. 1 Measurements of single item remaining experiment

Coefficient	**1 0 0**	Mea_in	Mea_out	**0 1 0**	Mea_in	Mea_out	**0 0 1**	Mea_in	Mea_out
Template $2 \times 2 \times 2$	0	1.13	0.78	0	0.37	0.57	0	0.82	0.5
	2	1.1	0.78	2	0.39	0.49	2	0.82	0.5
	5	0.98	0.8	5	0.38	0.61	5	0.73	0.52
Template $5 \times 2 \times 1$	0	1.5	0.65	0	1.13	0.54	0	0.84	0.37
	2	1.4	0.68	2	1.06	0.54	2	0.91	0.38
	5	1.36	0.67	5	0.63	0.72	5	0.77	0.37

Tab. 2 Measurements of two items remaining experiment

Coefficient	**1 1 0**	Mea_in	Mea_out	**0 1 1**	Mea_in	Mea_out	**1 0 1**	Mea_in	Mea_out
Template $2 \times 2 \times 2$	0	1.7	1.15	0	1.38	0.87	0	1.95	1.28
	2	1.59	1.17	2	1.31	0.89	2	1.92	1.28
	5	1.59	1.17	5	1.34	0.88	5	1.71	1.31
Template $5 \times 2 \times 1$	0	2.63	1.19	0	1.97	0.91	0	2.34	1.03
	2	2.46	1.23	2	1.97	0.93	2	2.3	1.06
	5	2	1.39	5	1.41	1.09	5	2.13	1.03

that at $r = 2$. Therefore, only one item can hardly describe the collages, and the exceptions in values and insensitivity to values are common.

Two Items Only two items are left for the final measurement in this experiment. The measurement values are recorded in Tab. 2, and the outliers are marked in red. The results show that the remaining and contradictory values still existed but to a lesser extent. Considering two items, template $5 \times 2 \times 1$ can almost provide a reasonable description, which may because of the difference in images. The item omission for template $2 \times 2 \times 2 \times 2$ still has serious effects.

5.1.2 Enhancement Experiment

All items are reserved in this experiment but with emphasized consideration for the final measurement. The measurement values are recorded in Tab. 3. We choose one item as an enhancement one and give it a higher coefficient. Outliers are also marked in red. One item is chosen as an enhancement and given a high coefficient. Outliers are also marked in red. The results reveal that template $5 \times 2 \times 1$ is seriously affected by the enhancement, and template $2 \times 2 \times 2 \times 2$ can maintain a normal level of collage description.

5.2 Template fixed MPC generation

The template is the direct factor that affects collage generation. Therefore, the generation result of the proposed method is first examined under fixed templates. Experiments are performed on the image collection by using uniform ($2 \times 2 \times 2 \times 2$) and a non-uniform template ($5 \times 2 \times 1$) templates. Fig. 5 shows the template sketch and the process

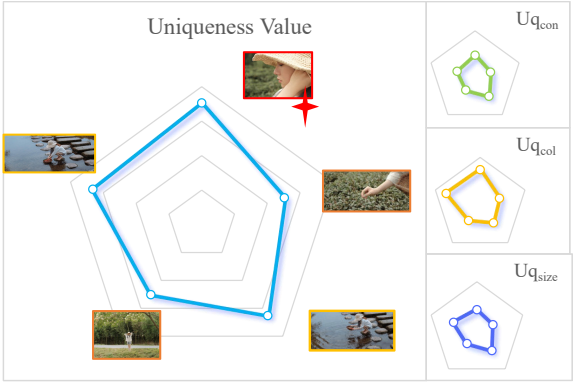


Fig. 6 Uniqueness value distribution of the page with five images. The left part is the distribution of comprehensive uniqueness value and the right part are the three sub-distributions of content (Uq_{con}), color (Uq_{col}), and size (Uq_{size}).

of collages generation. Starting with a disorganized images scatter ($r = 0$) initialization, the proposed method can reassign these images into pages in an information tidy form. For example, images with similar color information for templatefor template ($2 \times 2 \times 2 \times 2$) are close at $r = 2$ and the entire image scatter result converges at$r = 5$.

The final collage result must arrange images in a page based on the scatter result. All image number for the result of the template ($2 \times 2 \times 2 \times 2$) in a page is 2, of which the scatter result can be the final collage result. The page with five images needs the arrangement for the result of the template ($5 \times 2 \times 1$). Fig. 6 illustrates uniqueness value distribution, in which unique images and image pairs are respectively labeled by a red star and color boxes.

The final result in Fig. 5 is easy to read due to its

Tab. 3 Measurements of aspect enhancement experiment

Coefficient	**1 1 2**	Mea_in	Mea_out	**2 1 1**	Mea_in	Mea_out	**1 2 1**	Mea_in	Mea_out
Template $2 \times 2 \times 2$	0	3.34	2.16	0	2.44	3.65	0	3.08	2.03
	2	3.23	2.17	2	3.52	2.46	2	2.9	2.06
	5	3.05	2.19	5	3.3	2.48	5	2.94	2.05
Template $5 \times 2 \times 1$	0	4.3	1.94	0	4.97	2.22	0	4.6	2.1
	2	4.27	2	2	4.76	2.29	2	4.43	2.16
	5	3.54	2.13	5	4.13	2.43	5	3.4	2.48

clear information gathering effect, thus leading to a visually easy and pleasant experience. However, the generation of collages by a definite template is inflexible. Thus, the proposed method is further evaluated by presetting the collage page numbers.

5.3 Page Number fixed MPC Generation

Page number affects collage generation by indirectly limiting the template types. An exact page number setting generally corresponds to a template set in a limited number. For example, the collection of eight images reveal four template types, namely $((1 \times 7), (2 \times 6), (3 \times 5)$ and $(4 \times 4))$ under a setting of the deterministic page number 2. The experiments on two image collections ($collection_a$ and $collection_b$) are conducted under three different page number settings (2, 3 and 4).

5.3.1 Two-page collage generation

The template of two-page collage can be unfolded into seven individuals: $(1 \times 7), (7 \times 1), (2 \times 6), (6 \times 2), (3 \times 5), (5 \times 3), (4 \times 4)$. Instinctively, reading a collage page with a large image volume, which may be caused not only by the numerous information itself but also by the image volume variance among pages, is visually unfriendly. Fig. 7 shows the pipeline of two-page collage generation. With rough scatter ($r = 0$) initialization, the proposed method searches solutions in image and template spaces. The results in the optimization process vary, and the final scatter result can be convergent at a limited iteration number, that is, ($r = 10$) for $collection_a$ and ($r = 2$) for $collection_b$. The page volume of the final template is 3 and 4, which is an ideal configuration without any pre-designs. Images on a page are arranged similar to that in Sec. 5.2 based on the image scatter result. The final collage result is at the rightmost in Fig. 7, which is clear and visually pleasing.

5.3.2 Three-page collage generation

The total template number of a three-page collage is 21. Fig. 10 shows the pipeline of three-page collage generation. Compared with the configuration of two pages, the setting of three provides additional space to information gathering among pages. The final result of three pages demonstrates information aggregation and is easy to read whether for $collection_a$ or $collection_b$.

5.3.3 Four-page collage generation

The four-page configuration has a total of 20 templates. The total template number of a three-page collage is 21. Fig. 11 shows the pipeline of four-page collage generation, and the result is optimized step by step with iteration. However, Secs. 5.3.2 and 5.3.1, reveal that the information becomes specific per page as the page number increases. This situation has the opposite effect when the page number is 4. Fig. 11 shows that $collection_a$ must split two similar images into different pages to satisfy the page setting, and $collection_b$ must let two similar images become separate pages. Increasing the page number blindly is an ineffective means to improve the final quality. Thus, the proposed method is further evaluated by removing the page preset, and the solution is searched among all templates.

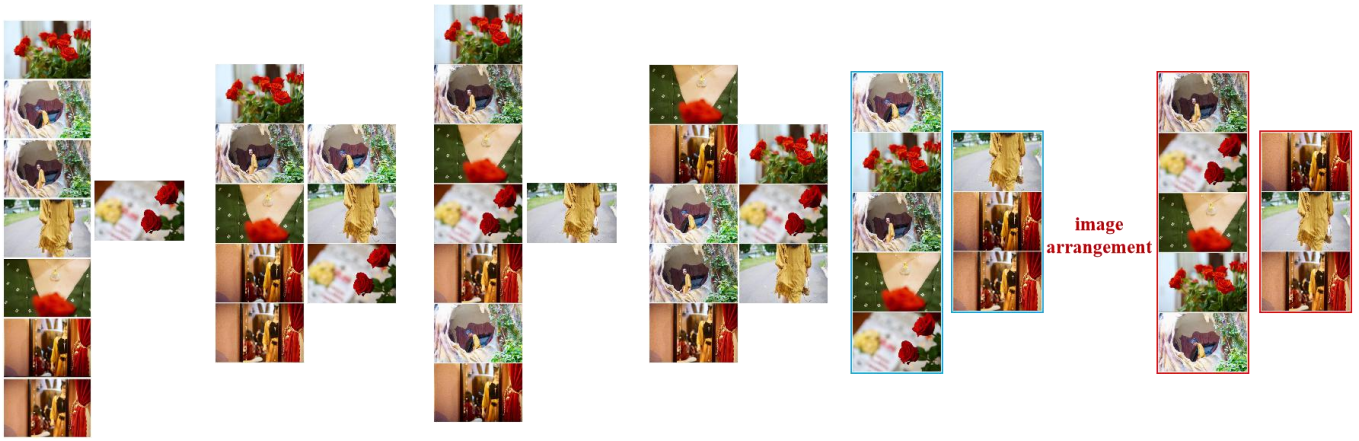
5.4 MPC generation without fixed condition

The template space is built under the assumption that the image page is no more than four to avoid an excessive number of templates for iteration. The size of the entire space is 47, which includes the two- and three-page templates. Fig. 12 compares the results generated by the proposed method (after 20 iterations) with random initialization and human results. The findings reveal that collages generated by the proposed method outperform random collages and are comparable to manual results.

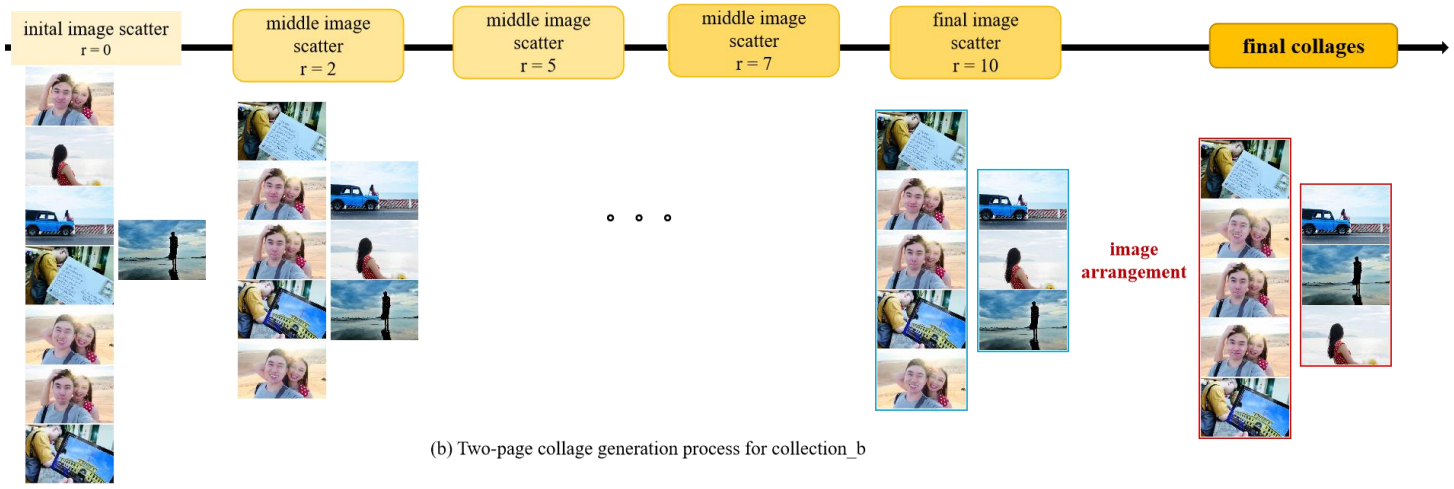
5.5 User Study

A total of 30 participants with different backgrounds (15 males and 15 females) were invited. Twenty new image collections, which were released on the website according to the designed arrangement, were obtained. Participants were shown three different results for each image set: 1) generated randomly, 2) generated by the proposed method, and 3) arranged by photographers. The participants were then asked to choose one from these results.

Participants were shown three different results for each image set, one is generated randomly, one is generated by our method, and the last one is arranged by photographers, and were asked to choose one from them.



(a) Two-page collage generation process for collection_a



(b) Two-page collage generation process for collection_b

Fig. 7 Pipeline of two-page collage generation. The entire generation follows the timeline, yellow blocks represent the iteration number, and collages nearby are the corresponding image scatter result. The image scatter result is surrounded by blue boxes. Then, the images in each page will be rearranged in accordance with the mechanism of uniqueness value computation and image similarity matching. The final collage result is labeled by red box.

Tab. 4 Statistics of the votes of different collage strategies

Strategy	Random	Our	Human
Votes	1.6%	47.3%	51.1%

Tab. 4 shows that the collages generated by the proposed method are superior to the random result and have votes close to the photographer result.

6 Limitation and Discussion

There are two main limitations of our work. One is that the image number of each template is fixed (8 in our experiments), thus it's hard to deal with the image set which has more or less images, which makes the method less of versatility. The other limitation is that our template form is fixed, so it's can not collage the images according to it's content flexibly. In the future, we will take advantage of template generated methods like [20] to enhance the generality of the algorithm.

Fig. 8 illustrates a failure case of our method. We can observe that our method prefers to scatter images that contain different visual information to individual pages (6 pages in Fig. 8(a)). However, images with different visual cues may contain similar scene semantics and should be scattered to one page (see Fig. 8(b)), so that the whole MPC can be more brief and clear. Moreover, Fig. 9 shows a special case. The result generated by our method is different from human result, while both results are visually pleasing and reasonable.

7 Conclusion

This study focuses on the issue of MPC generation, which is common in social networks. Compared with traditional

(a) Our result

(b) Human result

Fig. 8 Failure case of our method. Our result has low consistency with the human result.

(a) Our result

(b) Human result

Fig. 9 Special case of our method. Our result has low consistency with the human result.

collage studies, MPC must organize images into pages (scatter images into pages and arrange images per page) to generate visually pleasing results. Measurements based on information aggregation, which can guide image scattering, are proposed; color, content, and size of images are leveraged as basic items of image information. Furthermore, the measurements are optimized by searching for solutions with a non-dominated sorting algorithm-II (NSGA-II). A balanced strategy at the image-level and a uniqueness value computation mechanism are proposed to arrange images within a page. Experiments show that the proposed measurements, balance strategy, and uniqueness mechanism can lead to clear information and visually pleasing results. The results generated by the proposed method are compared with those of photographers. Merging different kinds of templates into MPC in a subtle manner will be considered in the future.

Acknowledgements

We thank the anonymous reviewers for their valuable comments. We thank Dr. Xiaowen Huang for providing some testing images. This work was supported by National Key R&D Program of China under no. 2020AAA0106200, and by National Natural Science Foundation of China under nos. 61832016 and U20B2070.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

[1] Y. Cao, A. B. Chan, and R. W. H. Lau. Automatic stylistic manga layout. *ACM Trans. Graph.*, 31(6), Nov. 2012. 2

[2] Y. Cao, R. W. H. Lau, and A. B. Chan. Look over here: Attention-directing composition of manga elements. *ACM Trans. Graph.*, 33(4), July 2014. 2

[3] T. Chen, A. Lu, and S.-M. Hu. Visual storylines: Semantic visualization of movie sequence. *Computers & Graphics*, 36(4):241–249, 2012. 2

[4] Z. Chen, J. Lin, N. Liao, and C. W. Chen. Full reference quality assessment for image retargeting based on natural scene statistics modeling and bi-directional saliency similarity. *IEEE Transactions on Image Processing*, 26(11):5138–5148, 2017. 3

[5] C. Coello Coello, G. B. Lamont, and D. A. van Veldhuizen. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007. 2, 3

[6] K. Deb. *Multi-objective optimization using evolutionary algorithms*, volume 16. John Wiley & Sons, 2001.

[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002. 3, 6

[8] Y. Gan, Y. Zhang, Z. Sun, and H. Zhang. Qualitative photo collage by quartet analysis and active learning. *Computers & Graphics*, 88:35–44, 2020. 3

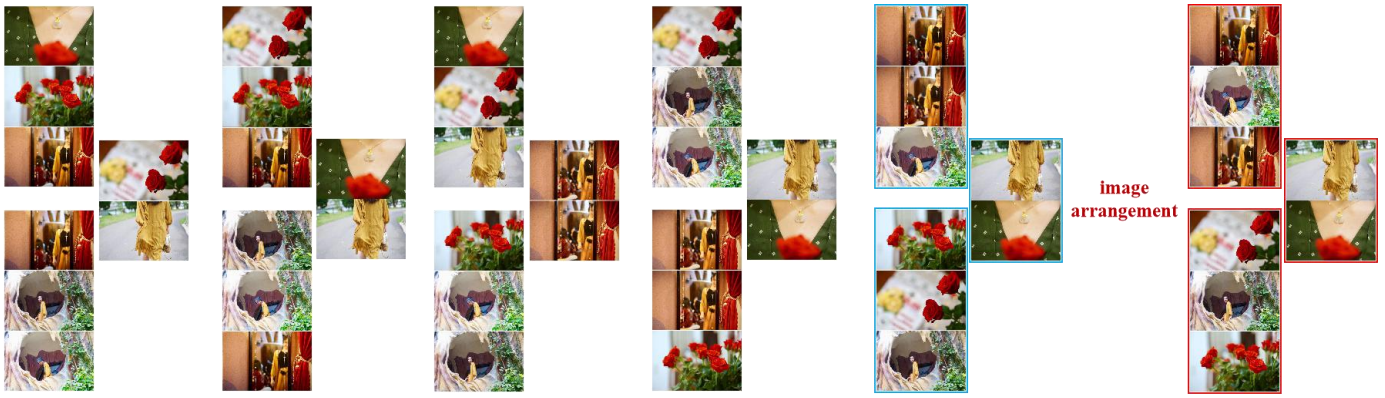
[9] X. Han, C. Zhang, W. Lin, M. Xu, B. Sheng, and T. Mei. Tree-based visualization and optimization for image collection. *IEEE Transactions on Cybernetics*, 46(6):1286–1300, 2016. 2

[10] R. Hübner and M. G. Fillinger. Comparison of objective measures for predicting perceptual balance and visual aesthetic preference. *Frontiers in psychology*, 7:335:1–335:15, 2016. 2, 4

[11] R. Hübner and M. G. Fillinger. Perceptual balance, stability, and aesthetic appreciation: Their relations depend on the picture type. *i-Perception*, 10(3):1–17, 2019. 2

[12] G. Jing, Y. Hu, Y. Guo, Y. Yu, and W. Wang. Content-aware video2comics with manga-style layout. *IEEE Transactions on Multimedia*, 17(12):2122–2133, 2015. 2

[13] S. Kong, X. Shen, Z. Lin, R. Mech, and C. Fowlkes. Photo aesthetics ranking network with attributes and content adaptation. In *European Conference on Computer Vision (ECCV)*, pages 662–679. Springer, 2016. 3



(a) Three-page collage generation process for collection_a

(b) Three-page collage generation process for collection_b

Fig. 10 Pipeline of three-page collage generation.

[14] C. Lee, S. Shin, and R. R. Issa. Multi-objective optimization of a free-form surface based on generative designs. In *International Conference on Computing in Civil and Building Engineering*, pages 1252–1261. Springer, 2020. 3

[15] D. Li and J. Zhong. Dimensionally aware multi-objective genetic programming for automatic crowd behavior modeling. *ACM Trans. Model. Comput. Simul.*, 2020. 3

[16] Y. Liang, Y.-J. Liu, and D. Gutierrez. Objective quality prediction of image retargeting algorithms. *IEEE Transactions on Visualization and Computer Graphics*, 23(2):1099–1110, 2016. 3

[17] Y. Liang, X. Wang, S. Zhang, S. Hu, and S. Liu. Photorecomposer: Interactive photo recomposition by cropping. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2728–2742, 2018. 1, 3

[18] L. Liu, H. Zhang, G. Jing, Y. Guo, Z. Chen, and W. Wang. Correlation-preserving photo collage. *IEEE Transactions on Visualization and Computer Graphics*, 24(6):1956–1968, 2018. 1, 3

[19] L. Ma, L. Xu, Y. Zhang, Y. Yan, and K. N. Ngan. No-reference retargeted image quality assessment based on pairwise rank learning. *IEEE Transactions on Multimedia*, 18(11):2228–2237, 2016. 3

[20] X. Pan, F. Tang, W. Dong, C. Ma, Y. Meng, F. Huang, T. Lee, and C. Xu. Content-based visual summarization for image collections. *IEEE Transactions on Visualization and Computer Graphics*, 2019. 1, 3, 9

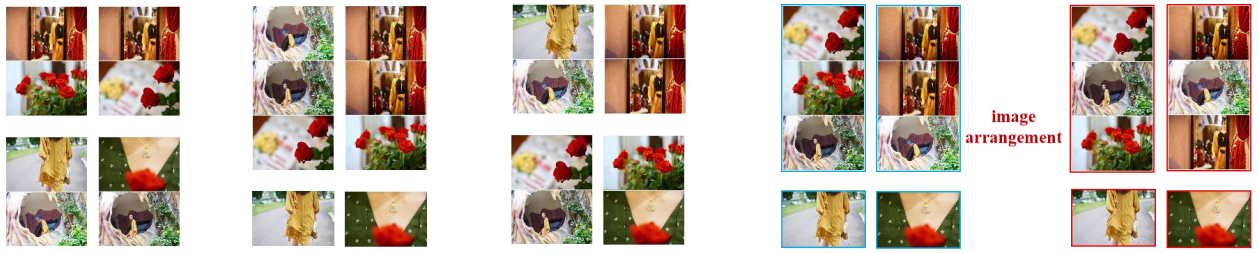
[21] M. Rejeesh and P. Thejaswini. Motf: Multi-objective optimal trilateral filtering based partial moving frame algorithm for image denoising. *Multimedia Tools and Applications*, 79(37):28411–28430, 2020. 3

[22] K. Sheng, W. Dong, H. Huang, M. Chai, Y. Zhang, C. Ma, and B.-G. Hu. Learning to assess visual aesthetics of food images. *Computational Visual Media*, 2020. 3

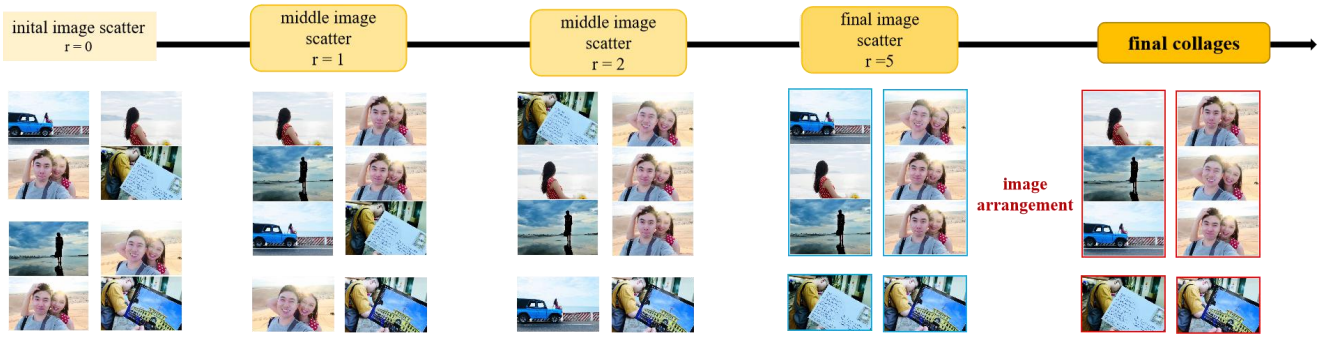
[23] K. Sheng, W. Dong, C. Ma, X. Mei, F. Huang, and B.-G. Hu. Attention-based multi-patch aggregation for image aesthetic assessment. In *Proceedings of the 26th ACM International Conference on Multimedia*, page 879–886, New York, NY, USA, 2018. Association for Computing Machinery. 3

[24] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, May 2015. 3

[25] Y. Song, F. Tang, W. Dong, F. Huang, and C. Xu. Balance-based photo posting. In *SIGGRAPH Asia 2019 Posters*,



(a) Four-page collage generation process for collection_a



(b) Four-page collage generation process for collection_b

Fig. 11 Pipeline of four-page collage generation.

pages 20:1–20:2, New York, NY, USA, 2019. Association for Computing Machinery. 3

[26] J. Su and H. Yin. Hypervolgan: An efficient approach for gan with multi-objective training function. *arXiv preprint arXiv:2006.15228*, 2020. 3

[27] L. Tan, Y. Song, S. Liu, and L. Xie. Imagehive: Interactive content-aware image summarization. *IEEE Computer Graphics and Applications*, 32(1):46–55, 2012. 2

[28] K. Thömmes and R. Hübner. Instagram likes for architectural photos can be predicted by quantitative balance measures and curvature. *Frontiers in Psychology*, 9:1050:1–1050:17, 2018. 4

[29] J. Wagemans, J. H. Elder, M. Kubovy, S. E. Palmer, M. A. Peterson, M. Singh, and R. von der Heydt. A century of gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization. *Psychological bulletin*, 138(6):1172, 2012. 2

[30] A. Wilson and A. Chatterjee. The assessment of preference for balance: Introducing a new test. *Empirical Studies of the Arts*, 23(2):165–180, 2005. 2

[31] Z. Wu and K. Aizawa. Very fast generation of content-preserved photo collage under canvas size constraint. *Multimedia Tools and Applications*, 75(4):1813–1841, Feb 2016. 3

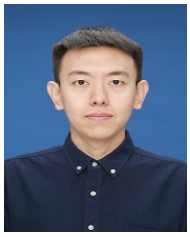
[32] M. Xu, J.-X. Zhong, Y. Ren, S. Liu, and G. Li. Context-aware attention network for predicting image aesthetic subjectivity. In *Proceedings of the 28th ACM International Conference on Multimedia*, page 798–806, New York, NY, USA, 2020. Association for Computing Machinery. 3

[33] Z. Yu, L. Lu, Y. Guo, R. Fan, M. Liu, and W. Wang. Content-aware photo collage using circle packing. *IEEE Transactions on Visualization and Computer graphics*, 20(2):182–195, 2014. 1, 2

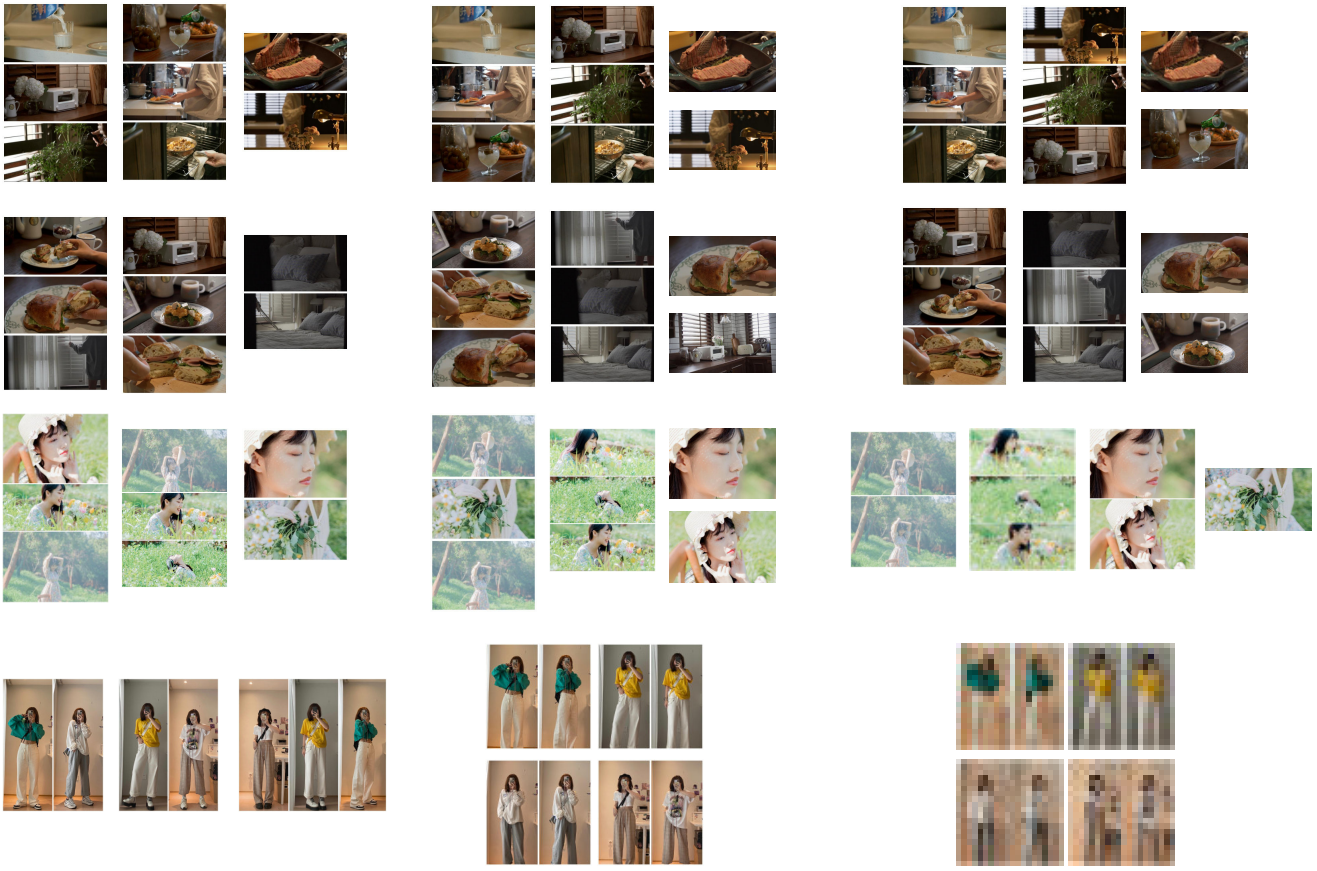
[34] Y. Zhang, W. Lin, Q. Li, W. Cheng, and X. Zhang. Multiple-level feature-based measure for retargeted image quality. *IEEE Transactions on Image Processing*, 27(1):451–463, 2017. 3

[35] X. Zheng, X. Qiao, Y. Cao, and R. W. H. Lau. Content-aware generative modeling of graphic design layouts. *ACM Trans. Graph.*, 38(4), July 2019. 3

Yu Song received the MSc degree in automation from Tianjin University in 2017. She is currently working toward the PhD degree in National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. Her research interests include image collage, image retargeting and machine learning.



Fan Tang is an assistant professor in School of Artificial Intelligence, Jilin University. He received the BSc degree in Computer Science from North China Electric Power University in 2013 and the Ph.D. degree from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, in 2019. His research interests include computer graphics, computer vision, and machine learning.



(a) Random (b) Ours (c) Human

Fig. 12 Results of the comparative experiment.

Weiming Dong is a Professor in the Sino-European Lab in Computer Science, Automation and Applied Mathematics (LIAMA) and National Laboratory of Pattern Recognition (NLPR) at Institute of Automation, Chinese Academy of Sciences. He received his BSc and MSc degrees in Computer Science in 2001 and 2004, both from Tsinghua University, China. He received his PhD in Computer Science from the University of Lorraine, France, in 2007. His research interests include computational visual media and computational creativity. Weiming Dong is a member of the ACM and IEEE.

Changsheng Xu is a Professor in National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and Executive Director of China-Singapore Institute of Digital Media. His research interests include multimedia content analysis/indexing/retrieval, pattern recognition and computer vision. He has hold 30 granted/pending patents and published over 200 refereed research papers in these areas. Dr. Xu is an Associate Editor of IEEE Trans. on Multimedia, ACM Trans. on Multimedia Computing, Communications and Applications and ACM/Springer Multimedia Systems Journal. He received the Best Associate Editor Award of ACM Trans. on Multimedia Computing, Communications and Applications in 2012 and the Best Editorial Member Award of ACM/Springer Multimedia Systems Journal in 2008. He served as Program Chair of ACM Multimedia 2009. He has served as associate editor, guest editor, general chair, program chair, area/track chair, special session organizer, session chair and TPC member for over 20 IEEE and ACM prestigious multimedia journals, conferences and workshops. He is IEEE Fellow, IAPR Fellow and ACM Distinguished Scientist.