Emotion Based Picture Collage

Chih-Hui Ho and Miao-Chiang Yen
PhD student: Kuan-Chuan Peng
Advisor: Prof. Tsuhan Chen
Undergraduate Honor Degree Program of EECS
National Chiao Tung University
Emails:bill83221@gmail.com, claude51315@gmail.com

Abstract

A picture collage application based on emotion ROI is proposed in the report. The main purpose is to utilize the concept of emotion ROIs and create picture collage with them. The collage is designed to show the most emotionally essential part of each input image and minimize the blank area in the canvas. The optimization is accomplished by greedy algorithm. Experimental results reveal the feasibility of the proposed method for picture collage.

1. Introduction

Picture collage is often seen in our daily life. For example, friends will post collage pictures on Facebook and those pictures carries a lot of memories. However, some picture collage system only displays the salience region of each image in the collage, which does not thoroughly convey the emotion in the image. Hence, our system utilize the results from Emotion6[1] and we proposed a novel method to combine the most emotionally essential part of each input image.

We have found that a previous work, Picture Collage[2], which has similar problem with us. The previous work addressed a method that could automatically generate a picture collage with several input images. Their picture collage has four features: salience region maximization, salience region ratio balance, blank space minimization and orientation diversity.

However, after reviewing previous work, we believe that maximizing the emotion ROI in each image is a better method to create a collage, because picture collage should be a composite image that contains the most essential part of different images without sacrificing the emotion that original images would like to express. It is proven in the paper [1], that emotion ROI is not identical to salience region. Hence, salience region maximization is replaced with emotion ROI in our picture collage.

Another main difference from the previous work is that each input image will be displayed in the collage picture with their original resolution, unlike the previous work. Every input image was cropped beforehand without discarding the most emotionally essential part for the sake of maintaining the original image's resolution. The cropping process will be discussed in Sec. 2.

The remainder of this report is organized as follows. In Sec. 2, the details of the cropping methods are proposed. In Sec. 3, problem formulation of our system will be covered. In Sec. 4, four proposed methods to accomplish our work are introduced. In Sec. 5, framework of our application will be shown. In Sec. 6, some experimental results are described. In Sec. 7, conclusions and suggestions for future works are presented.

2. Cropping method

We gather the 15 emotion ROIs from each input image from Emotion6, and apply Algorithm1 to generate the cropped images, or "Candidate ROIs".

Algorithm 1. Cropping and Candidate ROIs generating.

Input: 1) 15 emotion ROIs given from Emotion 6 and the original image

Output: Candidate ROIs

Step:

1. Set weighted value K
2. If pixel P is covered in N emotion ROIs, its pixel value will be K*N
3. Sum all the pixel value. Define the sum S
4. Set up a threshold T, and extract the pixels whose value are above threshold and set those pixels' value to 255
5. Find a minimum rectangle R that contains all the extracted pixels Fig.1(d).
6. If all the pixels value in R > 0.5*S, we put the cropped image into candidate ROIs.
7. Back to step 4 and set another smaller threshold T' < T
8. *Collect all the cropped images and take them as candidate ROIs*

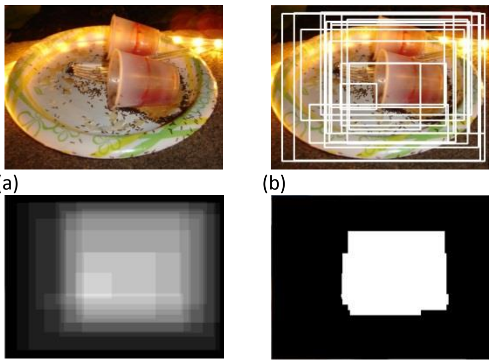

Figure 1 Illustration of stimuli map (a) Original image (b)Original image with emotion ROIs (c) stimuli map (d) Find a minimum rectangle contains the white area

By algorithm 1, we can assure that the candidate ROIs cropped from the original image contains most of the emotionally essential part of the image. These candidate ROIs are marked from level zero to level n. Level zero is the most essential part of the original image, because it is cropped with the highest threshold, compare to that of other level candidate ROIs. These cropped ROIs will not be resized during the rest of our algorithm.

3. Problem formulation

Our problem formulation is as follow. First, we introduce a new term, global blank ratio (GBR), and define it as

$$GBR = \frac{\sum Blank\_Area}{Total\_Area\_In\_Canvas}$$

Our goal is to fill the whole canvas, that is , to minimize the GBR of the canvas.

Second, we introduce another term, level zero block ratio (LZBR), and it is defined as

$$LZBR(x) = 1 - \frac{\text{Visible\_Area\_Of\_Level\_Zero}(x)}{\text{Total\_Area\_Of\_Level\_Zero}(x)}$$

where x is an input image index

Each input image has its distinct LZBR and our goal is to minimize LZBR, so that the image can be displayed without being overlapped by other images.

4. Proposals

In this section, we will propose four proposals, which are local move (LM), global move (GM), expand and rotate. The details of four proposals are illustrated in the following subsections.

4.1 Local Move (LM)

Local move is used to reduce GBR or LZBR. According to input image's LZBR, different mode of local move is applied. If LZBR < 0.2, local move of mode one is accepted. If LZBR >= 0.2, mode two is applied. Mode one is executed to reduce GBR and mode two is executed to decrease LZBR. There are 8 directions to try in local move, as shown in Fig.2. If a direction trial is accepted, next direction trial will be the same. Default direction is set to 'Up'. The detailed process is described in algorithm 2:

Algorithm 2. Local Move.

Step:

1. Define **L** be input image's LZBR.
2. If **L** < 0.2 then mode one is used.

Else mode two is used.

3. Try a direction to see whether GBR or LZBR decreases.
4. If a trial success, then try the same direction at next trial.

Else choose another direction at next trial.

5. Go to step3.

The loop from step 3. to step 5. breaks when one of the following two condition is satisfied. Condition one: If all the 8 direction trials are rejected. Condition two: If the number of successful trials is accumulated to ten.

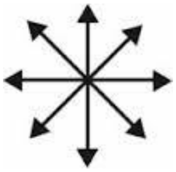

Fig.2 eight directions for local move

4.2 Global Move (GM)

The purpose of global move is to move the input image with a larger distance. Since local move can only move the input image in the adjacent region, it might not reduce the GBR and LZBR significantly in some situation. Hence, global move is proposed to solve this problem. Algorithm 3 describes the details of global move.

Algorithm 3. Global Move.

Step:

1. Define H as the height of current candidate ROI, and W as the width of current candidate ROI
2. Divide the canvas into grids, whose width is W and height is H
3. Scan grid by grid and choose the grids with largest blank area

From the experimental results, global move does successfully reduce LZBR and GBR, when local move does not work efficiently.

4.3 Expand

At the beginning, all the input images' levels are set to zero. The expand proposal is applied to increment input image's level, which can increase the size of candidate ROI. Thus, GBR can be reduced. Before expanding an image, we will check other images' LZBRs. If any image's LZBR increases, the expansion will not be applied. Therefore, some images in canvas expand and some do not in the expand proposal.

4.4 Rotate

The main purpose of rotation is to minimize the LZBR of all input image. Although through adjustments from LM and GM, the system can mostly avoid one image overlapping another. However, in some special situation, the image itself has to rotate in order to avoid overlapping other images. Our algorithm will try angle from -30 degree to 30 degree with 5 degree change at a time and select the best angle that reduce LZBR of all input images the most.

Fig. 3 the input image with web is rotated with certain degree to minimize the LZBR of its nearby input images

5. Framework of application

Our framework takes the advantage of the four proposals mentioned in the previous section. The flow chart in Fig 4 demonstrates the process of algorithm4

Algorithm 4. Framework of application.

Input*:* 1) N input images $\{I_1, I_2 \ldots I_N\}$with their candidate ROIs $\{\{I_{10}, I_{11} \ldots I_{1P}\}, \{I_{20}, I_{21} \ldots I_{2Q}\}, \ldots\ldots \{I_{N0}, I_{N1} \ldots I_{NR}\}\}$, where P, Q, R does not have to be identical.

Output: Collage picture

Step:

1. Randomly distribute N image in the canvas
2. Apply GM
3. The priority to apply step4 is according to the sorting result of N input images' LZBR. The input image that has higher LZBR will take step4 first.
4. If the selected input image's LZBR >= 0.2, apply GM and then apply LM mode2. Else, apply LM mode 1
5. Check if N input images' LZBR < 0.2. If satisfied, continue step 6, else back to step 3. The iteration between step 3 to step 5 will continue at most 3 times
6. Apply expand to each image
7. Apply rotate to each image

8. If GBR < 0.1, output the result. Else back to step 3. The iteration between step 8 and step 3 will continue at most 5 times.

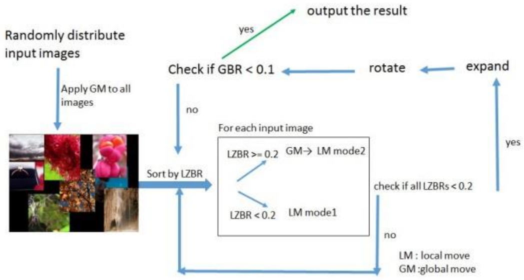

Figure 4 Flow chart of our application

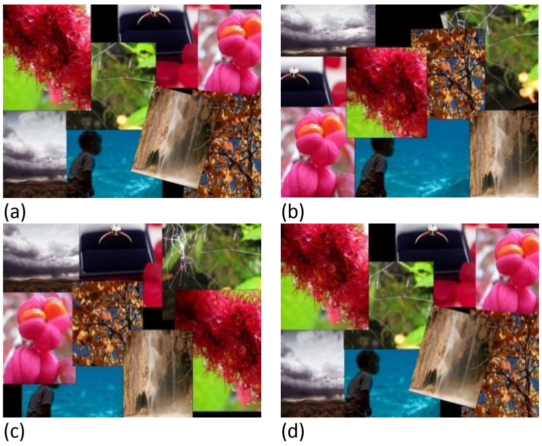

Figure 5 result of the picture collage

6. Experiment result

Some experiment results are shown in Table.1. Our test data are as following:

1. canvas size: 1024*768
2. number of input images : 8
3. size of input images :
   300*300 ~ 500*500

After executing our framework for 30 times, some statistical results can be computed. First, average execution time is 42.61 seconds with standard deviation is 12.42 seconds. Second, average GBR is 6.1% with standard deviation 3.1%. The last is about LZBR. Here we compute two statistic figures about LZBR. One is average LZBR and the other is average largest LZBR. Average LZBR is the average of overall LZBR. However, it does not show that all the LZBRs are balanced. Thus, we compute average largest LZBR instead. The average LZBR is 7.32% with standard deviation 6.2%. Largest and the average largest LZBR is 15.7% with standard deviation 2.3%. The result of largest average LZBR shows that the LZBRs are balanced.

	Avg.	Stdev.
Execution time	42.61s	12.42s
GBR	6.1%	3.1%
LZBR	7.32%	6.2%
Largest LZBR	15.7%	2.3%

Table.1 the statistical result of our experiment

7. Conclusion

In this report, an emotion based picture collage framework is proposed. This application aims to display the most emotionally essential region of input images on the canvas, without sacrificing its resolution.

The future work includes: 1) better algorithm to lower the execution time. 2) to reduce more blank area in the canvas. 3) Collage with animated layout

Reference

[1] Kuan-Chuan Peng, Amir Sadovnik, Andrew Gallagher, Tsuhan Chen,
Where do emotions come from: predicting the emotion stimuli map

[2] JingdongWang, Jian Sun, Long Quan, Xiaoou Tang, Heung-Yeung Shum
Picture Collage , Wang et al. ,CVPR06

[3] Xin Lu Poonam, Suryanarayan Reginald B. Adams, Jr. Jia Li Michelle G. Newman James Z. Wang
On Shape and the Computability of Emotions

[4] Kuan-Chuan Peng, Kolbeinn Karlsson, Tsuhan Chen
A framework of changing image emotion using emotion prediction

[5] Jana Machajdik, Allan Hanbury
Affective Image Classification using Features Inspired by Psychology and Art Theory