

A Novel Image Mosaic Method for Strip-Type Paper Fragments

Guo Cheng^a, Zhenshi Zhang^b, Tieliang Zhang^c

College of Basic Education, National University of Defense Technology, Changsha, China

^aguocheng@nudt.edu.cn, ^bCorresponding author: zhangzhenshi@nudt.edu.cn, ^ctieliangzhang@nudt.edu.cn

Abstract—Paper fragment mosaic is one of the applications of Image Mosaic, which plays an important role for Historical Document Re. Based on a series of modeling ideas, we propose a novel method to mosaic strip-type paper fragments. In this paper, Firstly, the boundaries are sampled and binarized from paper fragments. Secondly, a matching ratio model is built to calculate the possibility that two fragments are spatial adjacent. Thirdly, the mosaic problem is transformed into a Traveling Salesman Problem (TSP) based on graph theory, and then solved via Linear Programming. In the experiments, a Chinese and an English paper are vertically cut into pieces, each of which is scanned as a picture file. The program coded based on our method can automatically mosaic all the fragments into the complete original paper picture, which proves the effectiveness of our method.

Keywords—Image mosaic; Paper mosaic; Traveling salesman problem; Linear programming.

I. INTRODUCTION

An image mosaic uses a collection of small images to construct a larger image that suggests a whole image seen at a distance, but the individual images are visible when observed closely [1]. Image mosaic related technologies include spatial matching, data sampling, information fusion, and so on [2]. Image mosaic is one of the hottest research spots in digital image processing, and it has been widely used in various subjects such as Computer Graphics [3], Remote-sensing Image Processing [4], and so on.

Paper mosaic is one of the applications of image mosaic [5]. It is usually used to recovery historical documents [6]. The overlapping proportion of two fragments is the most important one of keys that matters the mosaic effect in Image Mosaic research area. If two fragments have no overlapping parts, the mosaic would be a problem because it is difficult to decide whether the two fragments are spatial adjacent. However, non-overlapping is a common situation for paper mosaic. Therefore, the method we propose for paper mosaic in this paper should mosaic fragments which are not overlapped to each other.

The paper is organized into three sections. Section I introduces the problem we try to solve. Section II gives the details of our method, including the idea, the models and the solution. Section III proves our method via experiments.

II. BOUNDARY-BASED MOSAIC METHOD

A. Boundary comparison

Since the fragments have no overlapping part, we can only use the boundary information to judge whether fragments are spatial adjacent. Therefore, the first step of our method is data sampling which generates a two-dimensional matrix from a given fragment shown in Figure 1. Equation (1) is an example of the matrix description, where a_{ij} indicates the gray value of the sampled pixel (i, j) ; q and p indicate the width and height of the fragment, respectively.

255	255
255	255
255	255
255	255
255	255
255	255
175	225
23	45
0	0
0	13
234	211
0	0
0	0
212	255
255	255
255	255

Figure 1. Data of two spatial adjacent boundaries.

$$A_n = \begin{bmatrix} a_{11}^n & a_{12}^n & a_{13}^n & \cdots & a_{1q}^n \\ a_{21}^n & a_{22}^n & a_{23}^n & \cdots & a_{2q}^n \\ a_{31}^n & a_{32}^n & a_{33}^n & \cdots & a_{3q}^n \\ \vdots & \vdots & \vdots & & \vdots \\ a_{p1}^n & a_{p2}^n & a_{p3}^n & \cdots & a_{pq}^n \end{bmatrix} \quad a_{ij}^n \in [0, 255] \quad (1)$$

Hypothetically, there are two boundaries existing in two fragments. In ideal conditions, the data of the two boundaries should be almost the same because they are spatial adjacent. In most conditions, however, the data is similar rather than same because of the damage. Therefore, we could compare

the data of the two boundaries and build a matching ratio model based on boundary data.

B. Binaryzation of boundary

The range of the grey value is from 0 to 255. Therefore, it needs a complicated model to calculate the match ratio of two boundaries. In paper mosaic problem, however, the grey value is not that complicated because the color is either white or black, i.e., the color of paper and word. Therefore, in order to simplify the model and decrease the cost time, we binarize the gray value. By setting a threshold, the gray value of boundary data is binarized into 0 and 255, shown in Figure 2 and Equation (2).

255	255	Binarize	255	255
255	255		255	255
255	255		255	255
255	255		255	255
255	255		255	255
175	225		255	225
23	45		0	0
0	0		0	0
0	13		0	0
234	211		255	255
0	0		0	0
0	0		0	0
212	255		255	255
255	255		255	255
255	255		255	255

Figure 2. Binarization of grey value of boundary data

$$a_{ij}^n = \begin{cases} 0 & | a_{ij}^n < k \\ 255 & | a_{ij}^n \geq k \end{cases} \quad (2)$$

C. Matching ratio model

A boundary of a strip-type fragment could be described as a one-dimensional vector. Equation (3) and Equation (4) are two vectors that indicate the left and right boundaries of two fragments, where i indicates the index of the fragment.

$$L_i = (a_1, a_2, \dots, a_p)^T \quad (3)$$

$$R_i = (b_1, b_2, \dots, b_p)^T \quad (4)$$

Therefore, we define d_{ij} as the matching ratio of fragment i and fragment j in Equation (5), where p indicates the length of the vector. The smaller d_{ij} is, the higher matching ratio is, and the easier for mosaic the two fragments are.

$$d(i, j) = \sum_{i=1}^p |a_i - b_j| \quad (5)$$

D. Graph model and TSP

The paper fragment mosaic is to reorder the disordered paper fragments according to the boundaries, and joint the spatial adjacent fragments two by two. Therefore, it could be described as a permutation and combination problem. If we could find the original order of the paper fragments, we could mosaic all the fragments correctly. In order to get the order, we should compare each two fragments and find whether these two are spatial adjacent. In our model, we use the abovementioned matching ratio d_{ij} to describe the

possibility of spatial adjacent. The smaller d_{ij} is, the larger the possibility is. Therefore, the Equation(6) indicates the final possibility when all the fragments have been compared two by two. In the equation, π indicates the order of fragments. Because the model is used for strip-type fragment mosaic, the order π is only one-dimensional and thus we only need to calculate the matching ratio between $A_n(i)$ and $A_n(i+1)$. When $D(\pi)$ is the smallest of all, π is the correct order of fragments, and the fragments could be perfectly mosaic.

$$D(\pi) = \sum_{i=0}^{n-1} d(A_\pi(i), A_\pi(i+1)) \quad (6)$$

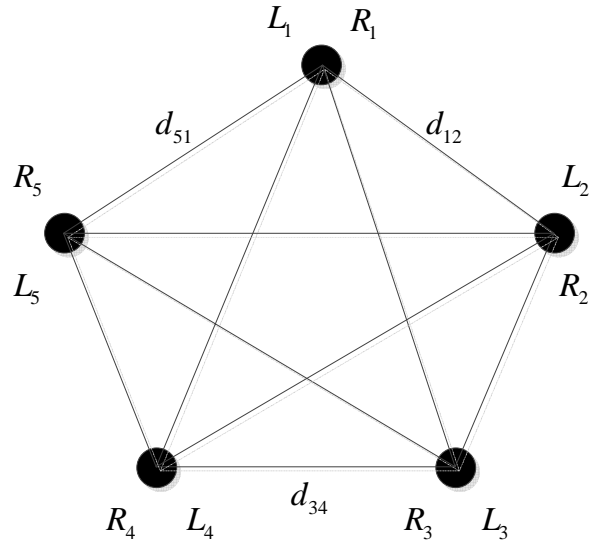


Figure 3. Graph for calculation of D

Next, to calculate the smallest value of D could be abstracted as a graph problem shown in Figure 3. The boundaries of fragments become the points of the graph. The boundary vectors L_i and R_i of each fragment become the

values of points of graph. The matching ratio d_{ij} becomes the length of the edge of graph connecting point i and point j . In this graph, if we find a path that pass each point only once, the sum of the length of the edges that consist the path will be $D(\pi)$ in Equation(6). The path passing each point only once in a graph is also known as Hamilton Path [7]. Therefore, the order problem is transformed into finding the shortest Hamilton path problem. The calculation problem becomes to calculate the total length of Hamilton path in a given graph, also known as the Traveling Salesman Problem (TSP) [8, 9].

E. Solution of TSP

Define $G(V, E)$ as the abovementioned graph, where V indicates the point collection and E indicates the edge collection.

Define a parameter

$$x_{ij} = \begin{cases} 1 & | \text{edge}(i, j) \in E' \\ 0 & | \text{edge}(i, j) \notin E' \end{cases} \quad (7)$$

Where E' indicates the shortest Hamilton path.

Therefore, the solution of TSP could be transformed into a linear programming problem [9] in Equation (8).

$$\begin{aligned} & \min \sum_{(i,j) \in E} d_{ij} x_{ij} \\ & \begin{cases} \sum_{(i,j) \in E} x_{ij} = 1 \\ \sum_{(j,i) \in E} x_{ji} = 1 \\ x_{ij} \in \{0,1\}, \{i,j\} \in E \end{cases} \end{aligned} \quad (8)$$

The solution contains the following five steps.

- (1) Sample and generate the boundary of all paper fragments $A(i)$, put them into the point collection.
- (2) Choose the point whose left boundary is null, i.e., the data is all 0, as the starting point. Remove this point out of the point collection.
- (3) Calculate the matching ratio between the starting point and any other point of the collection.
- (4) Find the smallest matching ratio, choose that point as the next one, and meanwhile remove it out the point collection.
- (5) Repeat step(3) and step(4), until the collection is empty.
- (6) The order of the choosing points is the order of the paper fragments.

III. EXPERIMENTS

In order to prove the effectiveness of our method, we have applied it to two experiments, one is for Chinese paper and the other one is for English paper. In each experiment, the paper is vertically cut into 19 pieces. Each piece is scanned as a picture file. We implement our method into a C++ program. The program can automatically mosaic all the fragments into the complete and original paper picture shown in Figure 4 and Figure 5. Therefore, the method proposed in this paper is effective.

In the future, we will focus on how to improve the efficiency of our method by optimizing the models and solutions.

城上层楼叠嶂。城下清淮古汴。举手揖吴云，人与暮天俱远。魂断。魂断。后夜松江月满。 簌簌衣巾莎枣花。村里村北响缦车。牛衣古柳卖黄瓜。海棠珠缀一重重。清晓近帘栊。胭脂谁与匀淡，偏向脸边浓。小郑非常强记，二南依旧能诗。更有鲈鱼堪切脍，儿辈莫教知。自古相从休务日，何妨低唱微吟。天垂云重作春阴。坐中人半醉，帘外雪将深。双鬟绿坠。娇眼横波眉黛翠。妙舞翩跹。掌上身轻意态妍。碧雾轻笼两凤，寒烟淡拂双鸦。为谁流睇不归家。错认门前过马。

我劝骀张归去好，从来自己忘情。尘心消尽道心平。江南与塞北，何处不堪行。闲离阻。谁念素襟襄王，何曾梦云雨。旧恨前欢，心事两无据。要知欲见无由，痴心犹自，倩人道、一声传语。风卷珠帘自上钩。萧萧乱叶报新秋。独携纤手上高楼。临水纵横回晚靥。归来转觉情怀动。梅笛烟中闻几弄。秋阴重。西山雪淡云凝冻。凭高眺远，见长空万里，云无留迹。桂魄飞来光射处，冷浸一天秋碧。玉宇琼楼，乘鸾来去，人在清凉国。江山如画，望中烟树历历。省可清言挥玉尘，真须保器全真。风流何似道家纯。不应同蜀客，惟爱卓文君。自惜风流云雨散。关山有限情无限。待君重见寻芳伴。为说相思，目断西楼燕。莫恨黄花未吐。且教红粉相扶。酒阑不必看茱萸。俯仰人间今古。玉骨那愁瘴雾，冰姿自有仙风。海仙时遣探芳丛。倒挂绿毛么凤。

俎豆庚桑真过矣，凭君说与南荣。愿闻吴越报丰登。君王如有问，结袜赖王生。师唱谁家曲，宗风嗣阿谁。借君拍板与门槌。我也逢场作戏，莫相疑。 鞞髯嫌枕印。印枕嫌髯。闲照晚妆残。残妆晚照闲。可恨相逢能几日，不知重会是何年。茱萸仔细更重看。午夜风翻幔，三更月到床。 簾纹如水玉肌凉。何物与依归去、有残妆。金炉犹暖麝煤残。惜香更把宝钗翻。重闻处，余熏在，这一番、气味胜从前。菊暗荷枯一夜霜。新苞绿叶照林光。竹篱茅舍出青黄。霜降水痕收。浅碧鳞鳞露远洲。酒力渐消风力软，飏飏。破帽多情却恋头。 烛影摇风，一枕伤春绪。归不去。凤楼何处。芳草迷归路。汤发云腴酽白，盏浮花乳轻圆。人间谁敢更争妍。斗取红窗粉面。炙手无人傍屋头。萧萧晚雨脱梧楸。谁怜季子敝貂裘。

：下： 一。 手： 魂断。古汴。簌簌。衣巾。莎枣花。村里。村北。响缦车。牛衣。古柳。卖。黄。瓜。海棠。珠。缀。一。重。重。清晓。近。帘。栊。胭脂。谁。与。匀。淡。偏。向。脸。边。浓。小。郑。非。常。强。记。二。南。依。旧。能。诗。更。有。鲈。鱼。堪。切。脍。儿。辈。莫。教。知。自。古。相。从。休。务。日。何。妨。低。唱。微。吟。天。垂。云。重。作。春。阴。坐。中。人。半。醉。帘。外。雪。将。深。双。鬟。绿。坠。娇。眼。横。波。眉。黛。翠。妙。舞。翩。跹。掌。上。身。轻。意。态。妍。碧。雾。轻。笼。两。凤。寒。烟。淡。拂。双。鸦。为。谁。流。睇。不。归。家。错。认。门。前。过。马。 我劝骀张归去好，从来自己忘情。尘心消尽道心平。江南与塞北，何处不堪行。闲离阻。谁念素襟襄王，何曾梦云雨。旧恨前欢，心事两无据。要知欲见无由，痴心犹自，倩人道、一声传语。风卷珠帘自上钩。萧萧乱叶报新秋。独携纤手上高楼。临水纵横回晚靥。归来转觉情怀动。梅笛烟中闻几弄。秋阴重。西山雪淡云凝冻。凭高眺远，见长空万里，云无留迹。桂魄飞来光射处，冷浸一天秋碧。玉宇琼楼，乘鸾来去，人在清凉国。江山如画，望中烟树历历。省可清言挥玉尘，真须保器全真。风流何似道家纯。不应同蜀客，惟爱卓文君。自惜风流云雨散。关山有限情无限。待君重见寻芳伴。为说相思，目断西楼燕。莫恨黄花未吐。且教红粉相扶。酒阑不必看茱萸。俯仰人间今古。玉骨那愁瘴雾，冰姿自有仙风。海仙时遣探芳丛。倒挂绿毛么凤。 俎豆庚桑真过矣，凭君说与南荣。愿闻吴越报丰登。君王如有问，结袜赖王生。师唱谁家曲，宗风嗣阿谁。借君拍板与门槌。我也逢场作戏，莫相疑。 鞞髯嫌枕印。印枕嫌髯。闲照晚妆残。残妆晚照闲。可恨相逢能几日，不知重会是何年。茱萸仔细更重看。午夜风翻幔，三更月到床。 簾纹如水玉肌凉。何物与依归去、有残妆。金炉犹暖麝煤残。惜香更把宝钗翻。重闻处，余熏在，这一番、气味胜从前。菊暗荷枯一夜霜。新苞绿叶照林光。竹篱茅舍出青黄。霜降水痕收。浅碧鳞鳞露远洲。酒力渐消风力软，飏飏。破帽多情却恋头。 烛影摇风，一枕伤春绪。归不去。凤楼何处。芳草迷归路。汤发云腴酽白，盏浮花乳轻圆。人间谁敢更争妍。斗取红窗粉面。炙手无人傍屋头。萧萧晚雨脱梧楸。谁怜季子敝貂裘。

Figure 4. The experiment of Chinese paper mosaic

fair of face.

The customer is always right. East, west, home's best. Life's not all beer and skittles. The devil looks after his own. Manners maketh man. Many a mickle makes a muckle. A man who is his own lawyer has a fool for his client.

You can't make a silk purse from a sow's ear. As thick as thieves. Clothes make the man. All that glitters is not gold. The pen is mightier than sword. Is fair and wise and good and gay. Make love not war. Devil take the hindmost. The female of the species is more deadly than the male. A place for everything and everything in its place. Hell hath no fury like a woman scorned. When in Rome, do as the Romans do. To err is human; to forgive divine. Enough is as good as a feast. People who live in glass houses shouldn't throw stones. Nature abhors a vacuum. Moderation in all things.

Everything comes to him who waits. Tomorrow is another day. Better to light a candle than to curse the darkness.

Two is company, but three's a crowd. It's the squeaky wheel that gets the grease. Please enjoy the pain which is unable to avoid. Don't teach your Grandma to suck eggs. He who lives by the sword shall die by the sword. Don't meet troubles half-way. Oil and water don't mix. All work and no play makes Jack a dull boy.

The best things in life are free. Finders keepers, losers weepers. There's no place like home. Speak softly and carry a big stick. Music has charms to soothe the savage breast. Ne'er cast a clout till May be out. There's no such thing as a free lunch. Nothing venture, nothing gain. He who can does, he who cannot, teaches. A stitch in time saves nine. The child is the father of the man. And a child that's born on the Sab-

...e.
right face. The customer is always right. East, west, home's best. Life's not all beer and skittles. The devil looks after his own. Manners maketh man. Many a mickle makes a muckle. A man who is his own lawyer has a fool for his client.
You can't make a silk purse from a sow's ear. As thick as thieves. Clothes make the man. All that glitters is not gold. The pen is mightier than sword. Is fair and wise and good and gay. Make love not war. Devil take the hindmost. The female of the species is more deadly than the male. A place for everything and everything in its place. Hell hath no fury like a woman scorned. When in Rome, do as the Romans do. To err is human; to forgive divine. Enough is as good as a feast. People who live in glass houses shouldn't throw stones. Nature abhors a vacuum. Moderation in all things.
Everything comes to him who waits. Tomorrow is another day. Better to light a candle than to curse the darkness.
Two is company, but three's a crowd. It's the squeaky wheel that gets the grease. Please enjoy the pain which is unable to avoid. Don't teach your Grandma to suck eggs. He who lives by the sword shall die by the sword. Don't meet troubles half-way. Oil and water don't mix. All work and no play makes Jack a dull boy.
The best things in life are free. Finders keepers, losers weepers. There's no place like home. Speak softly and carry a big stick. Music has charms to soothe the savage breast. Ne'er cast a clout till May be out. There's no such thing as a free lunch. Nothing venture, nothing gain. He who can does, he who cannot, teaches. A stitch in time saves nine. The child is the father of the man. And a child that's born on the Sab-

Figure 5. The experiment of English paper mosaic

ACKNOWLEDGMENT

The research in this paper is supported by the High School Doctoral Program Research Project in China (No.20134307120039) and Nature Science Foundation of Hunan Province in China (No.14JJ3003).

REFERENCES

- [1] Liu, Qiuliang, "Image mosaic algorithms and optimization," M.E.E., UNIVERSITY OF DELAWARE, pp58, 2007.
- [2] Dongyan Guo, Jinhui Tang, Ying Cui, Jundi Ding, Chunxia Zhao, Saliency-based content-aware lifestyle image mosaics, Journal of Visual Communication and Image Representation, Vol.26, pp.192-199, 2015.
- [3] Malcolm Richardson, Modern Computer Graphics, pp.210, 1989.
- [4] Jaewon Choi, Hyung-Sup Jung, Sang-Ho Yun, An Efficient Mosaic Algorithm Considering Seasonal Variation: Application to KOMPSAT-2 Satellite Images, Sensors, Vol.15(3), pp.5649-5665, 2015.
- [5] Myounghun Han, Dongwann Kang, Kyunghyun Yoon, Efficient paper mosaic rendering on mobile devices based on position-based tiling, Journal of Real-Time Image Processing, vol.9(3), pp.549-556, 2014.
- [6] Oksana Samko, Yu-Kun Lai, David Marshall, Paul L. Rosin, Virtual unrolling and information recovery from scanned scrolled historical documents, Pattern Recognition, Vol.47(1), pp. 248-259, 2014.
- [7] Richter R. Bruce, Hamilton path in generalized Petersen graphs, Discrete Math, Vol.313(12), pp.1338-1341, 2013.
- [8] Jornsten, Kurt and Kalcsics, Jurg, A New Lagrangean Approach for the Travelling Salesman Problem, NHH Dept. of Business and Management Science Discussion Paper No. 2015/4, 2015.
- [9] Wang Yong, A Genetic Algorithm with the Mixed Heuristics for Traveling Salesman Problem, International Journal of Computational Intelligence & Applications, Vol.14(1), pp.1-13, 2015.
- [10] Ankony Robert C., Linear programming with the simplex algorithm: Parallelization and other optimizations, California State University, M.S.C.S., 2012.