PICWORDS: AN AUTOMATIC IMAGE RENDERING APPROACH BY PACKING KEYWORDS

MONIKA Y S[1], SRINIVASA G[2]

[1]Student, Dept. of CSE, Sri Venkateshwara College of Engineering, *monikays15@gmail.com*

[2]Asst. Professor, Dept. of CSE, Sri Venkateshwara College of Engineering, *gsrinivasa11@gmail.com*

ABSTRACT

PicWords is one of the novel-text art system: where the source picture and keyword are taken as the input to fuse together forming a target picture called PicWords (picture packed with keyword/s).Unlike traditional text-graphics that are created manually with highly skilled artists involving huge amount of tedious manual work, PicWords is an automatic non-photorealistic rendering (NPR) packing system. Given any source image, its silhouette has been generated initially and segmented into patches. Also keywords are chosen randomly that relates to the image. Next, the warping of keyword into the image is obtained using mean value coordinates method. Finally post-processing methods are proposed for the better view of an image.

Keywords: *Calligrams, non-photorealistic rendering, NPR packing, PicWords.*

1. INTRODUCTION

Whenever people see any picture, they will be curious to know who that is or what actually the picture means. Since, the picture alone cannot predict or give complete or the minimum information about it, people surf on the internet to know who that particular person or what the particular picture represents. In another case, people might know few information about any particular person and never know how they look like. Thus, instead of going with this dilemma, this paper presents a novel text-art system to overcome the dilemma called PicWords.

PicWords is a text-art system which is a kind of non-photorealistic rendering [1] (NPR), particularly it comes under NPR packing; where the picture is packed with the keyword(s) that relates the picture (depicting the image by arranging a collection of small pictorial elements automatically). NPR is an area of computer graphics that focuses on enabling a wide variety of expensive styles for digital art. In contrast to the traditional computer graphics, which has focused on photorealism, NPR is inspired by artistic styles such as painting, drawing, technical illustration and animated cartoons. NPR possesses various applications such as toon shading and scientific visualization in the movies and video games. The following figure 1.1 represents the NPR art style.

The most fascinating part of this approach is that the human visual system (HVS) is attracted towards the two modalities: the picture and the keyword, sequentially during the human perception process. This approach or text-art was inspired by the advertisement posters that contain actor portraits and the introduction words. These posters of attractive look are hand-made and require highly imagination skills i.e., manually. Therefore, we explore an automatic approach to generate this kind of posters that is PicWords.

The earliest and traditional NPR packing art consist of Roman mosaics, where small squares of colored glass merge to form a detailed scene or image. If the small cooperating elements that are used to form an image are limited to only words or letters, it is known as Calligrams. Calligrams are another type of the NPR packing where it entirely depends upon the artists' imagination that has rich tradition within it. PicWords is a kind of Calligram.

The whole generation process of the PicWords contains: Picture-only module, Keyword-only module, Picture and Keyword module and post-processing module. The output of the each module is fed as input to the next module i.e., the output of the picture-only module is fed as input to the keyword only module (ranked image patches

that serves as input to the keyword module and ranked keyword respectively). In the third picture and keyword module, the two modalities are fused together by warping each other to its corresponding image patch. Finally, post-processing module is used to further refine the PicWords. All these modules are explained in the further sections.

Figure 1.1: Non-photorealistic rendering (NPR) digital art system

2. RELATED WORK

NPR packing addresses the arrangement of a multitude of small tiles to form artistic representation to enhance multimedia representations and has multiple applications such as scrapbook. These non-photorealistic rendering has been further classified into two types of art: (1) Mosaicking (2) Calligram.

2.1. MOSAICKING: In image processing Mosaics refer to the art where an image formed by the collection of small tiles together. Mosaicking aims in forming image regions by packing the multitude of atomic rendering elements. There are many applications with mosaics. Some of the representative works of mosaics are as follows:

(1).*Roman Mosaics*: This was the first approach presented by Hausner et al.[4] which addresses the irregular image tile shapes through an energy minimization scheme for shape packing. Here for the Voronoi diagrams, Lloyd's method was applied giving a rectangular orientation mosaic tiles.

(2).*Jigsaw image mosaics*[3] (JIM): this is another packing system proposed by Kim et al., which minimizes the measure of packing's quality by using the active contour based optimization scheme. The pictures are formed by the tiles that are arbitrary in shape.

Another representative work on Mosaicking was proposed by orchard and kalpan where the image is formed by the tiles of irregular shapes. It also has the capability of cropping the regions from the image database to use as tiles. One of the disadvantages with this packing method is that the viewer cannot read the content related with the target image.

2.2. CALLIGRAMS: Calligram [2] is a poem, a phrase or a word in which the typeface, calligraphy or handwriting is arranged in a way that creates a visual image. It is a computer-aided design of text-based art-forms has been explored in a number of different contexts. One well-known example is ASCII-art.

(1).*ASCII art:* The textual and numerical characters form the building blocks for forming and image in the ASCII art [5], i.e., instead of single characters many characters are arranged to form an image. Basically, ASCII art is a technique of composing a picture with printable text characters.

ASCII art is a graphic design technique that uses computers for presentation and consists of pictures pieced together from the 95 printable (from a total of 128) characters defined by the ASCII Standard from 1963 and ASCII compliant character sets with proprietary extended characters (beyond the 128 characters of standard 7-bit ASCII).

(2).*Fatfonts***:** This is the approach developed by Nacenta [6] which is based on the Arabic numerals. While preserving the overall visual context, this method enables the accurate reading of the numerical data.

The main disadvantage among the Calligrams approaches is that there is no relation between the existing image and the target image

3. PROPOSED SYSTEM

The proposed system consists of four modules which are (1) Picture module (2) Keyword module (3) Picture and Keyword module (4) Post-processing module. The system architecture or the overview of this approach is as depicted in the figure 3.1 below.
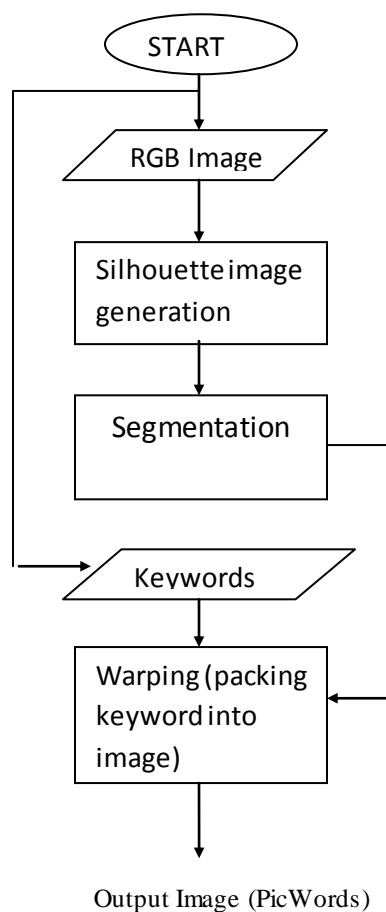

Figure 3: General Flow Diagram

In this approach, the source image is fed into the picture module to generate a segmentation map and a silhouette image, both of which collaboratively generate some patches. A keyword list is obtained from the keywords module. Then we construct a correspondence between the patch list and symbol filling techniques are adopted to improve the visual effects of PicWords. The figure 3 illustrates the flow of the following modules.

1) *Picture Module*:

Silhouette image generation: First task in the image module is to segregate the image into yang part (foreground image) and yen part (background image). This segregation is achieved through the mean shift segmentation method, where the yang part is kept safely for the further process and yen part has been neglected or discarded. The silhouette image generation takes various forms as shown in figure 3.2 below. Once the final silhouette image has been obtained, this is fed into the segmentation process for the patch generation. Basically, the original image can be directly fed into the segmentation process.
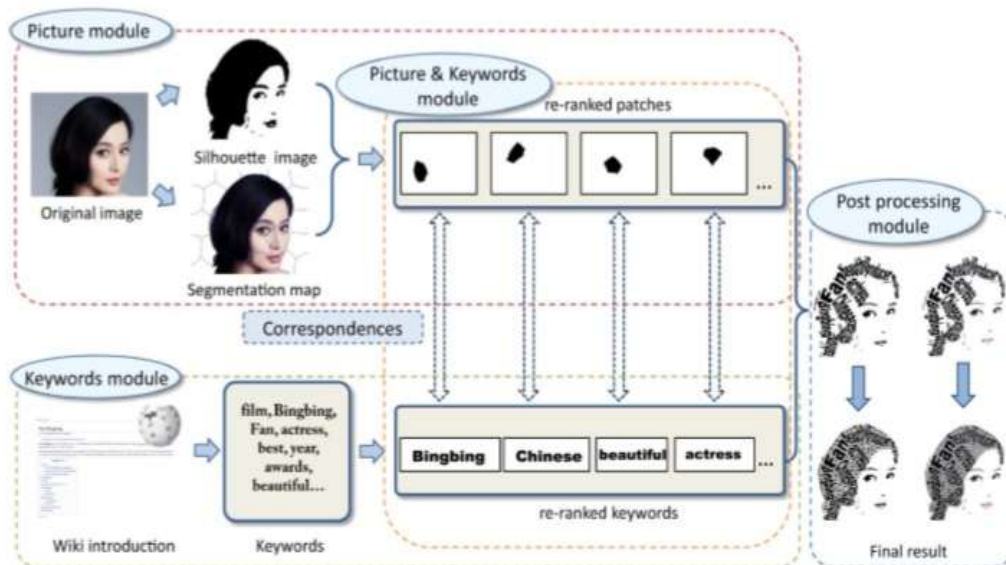
Figure 3.1: System Overview.

(a)   (b)   (c)
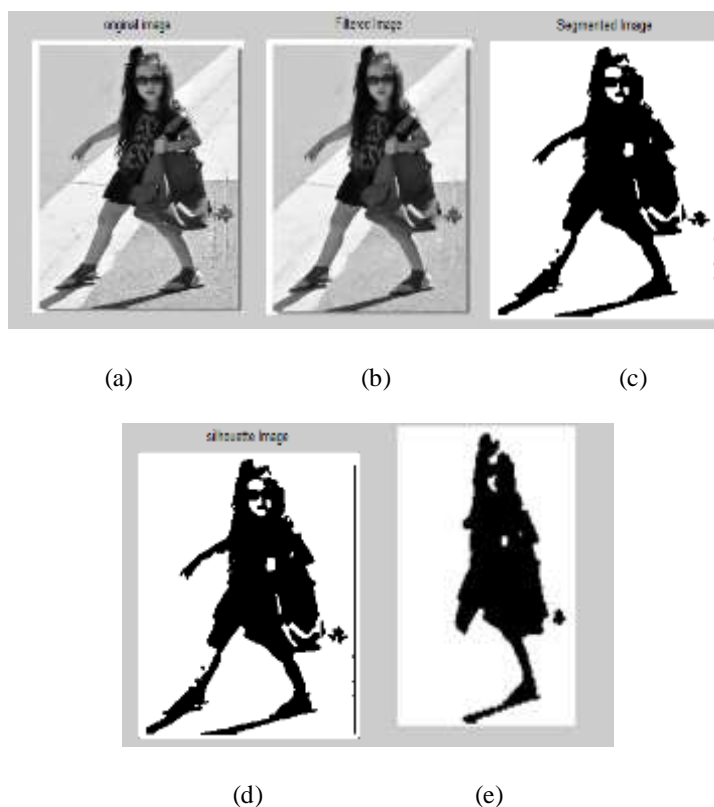

(d)   (e)

Figure 3.2: Silhouette image generation (a) original image (b) filtering image (c) Segmented foreground (d) Binary map (e) final silhouette image after smoothing.

But in this paper the silhouette image is taken as the input to the segmentation process so that the quality of the image is high comparatively, since smoothing and filtering are undertaken during the silhouette image generation.

Patch Generation: The super-pixel segmentation method is imposed on the silhouette image to obtain the patches. This proposed method is comparatively good than original mean-shift segmentation [8] because this provides near rectangular patches which is useful in fitting the keywords. Usually, only yang part is used for this purpose.

Convex Decomposition [10]: If the target image is apparently different from the original image they produce the distortion. To avoid distortion, the polygons should be near convex to avoid it, for this purpose the concave index has been estimated for the each pixel based on the minimum near convex decomposition algorithm used for the robust shape representation.

(2)*Keyword Module*:

This module contains two phases: keyword collection and keyword ranking. Keyword Collection: the keywords are extracted or gathered manually by the viewer. This obtained keyword is further processed in the next module.

(3)*Picture and Keyword Module*:

Correspondence and warping are the important tasks performed in this module. Initially, the correspondence between the patch and keyword are verified such that the appropriate patch contains the keyword that well suits it as described in the keyword ranking module. The ranked patches and the ranked keywords are sequentially mapped from high weight to the low weight such that; the important keyword and the highest frequency keyword is placed in the centre and the longer keyword is mapped to the longest patch and so on.

Warping: warping [12] text to the image patches means that the particular keyword should fit into the patch which is in arbitrary shape. This is the challenging task as the patches vary in their shapes and sizesFor this purpose warping is important and it contains three processes they are: size and alignment direction alignment, Correspondence points localization and Warping.

Size and Principal Direction alignment: As described, the image patch is of arbitrary shape where the keywords are always in upright direction and untilted. Thus the first step is to align the principal orientation of the word so that the size of the keyword is adjusted according to the size of the patch container.

Correspondence points localization: Warping aims at providing the explicit correspondence between the source and target image shapes. Other points are mapped according to their relative positions with respect to the correspondence points. This correspondence is achieved through computer automatically. First the contour of the word has been extracted, these contours 'C' [7] is sampled as 'n' points { $c_1,......c_n$ } and are imposed as edge points, this edge or outline of the irregular shaped image is approximated by the piecewise-linear paths with vertices. This forms the boundary of the patch so that the keyword fits into it approximately.

Warping is processes based on the correspondences constructed, that is the word is warped into the new position of the patch boundary. Based on the generalized notions of barycentric coordinates, Hormann et., al presented the mean value coordinates for warping between two arbitrary planar polygons. Barycentric coordinates provides a convenient way to linearly interpolate data that is given at the corners for the triangles. For a given planar triangle [v_1, v_2, v_3] any point in v has three masses w_1, w_2, and w_3. If placed at the corresponding vertices of the triangle, their barycentric will coincide with v as follows:

$$\frac{w_1 v_1 + w_2 v_2 + w_3 v_3}{w_1 + w_2 + w_3} = v$$

Thus w_1, w_2, w_3 are defined as the barycentric coordinates of the v.

(4)*Post-processing module***:**

After warping of the keyword into the patch if there is any blank area remaining, those areas can be filled with any special symbols for a special ornamental effect. This approach is for the binary image and for the better visualization effects grayscale images and colored images are used for PicWords.

4. CONCLUSION

PicWords is one of the best approaches in the image processing field that can illustrate the image informatively by fusing both the source picture and the provided keyword seamlessly into one target image. This is completely automatic thus reducing the manual work of the artist. PicWords s has great market potentials; they can be used in various applications for replacing the original profile pictures with the PicWords and also in some of the advertisement.

ACKNOWLEDGEMENT

REFERENCES

[1]. J. Gomes, Warping and Morphing of Graphical Objects. San Francisco, CA, USA: Morgan Kaufmann, 1999, vol. 1.

[2]. T. Strothotte and S. Schlechtweg, Non-Photorealistic Computer Graphics: Modeling, Rendering, and Animation.   San Francisco, CA, USA: Morgan Kaufmann, 2002.

[3]. J. Kim and F. Pellacini, "Jigsaw image mosaics," ACM Trans. Graph., vol. 21, no. 3, pp. 657–664, 2002.

[4]. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intel., vol. 24, no. 5, pp. 603–619, 2002.

[5]. K. Hormann and M. S. Floater, "Mean value coordinates for arbitrary planar polygons," ACM Trans. Graph., vol. 25, no. 4, pp. 1424–1441, 2006.

[6]. J. Xu and C. S. Kaplan, "Calligraphic packing," in Proc. Graphics Interface, 2007, pp. 43–50.

[7]. J. Orchard and C. S. Kaplan, "Cut-out image mosaics", in Proc. 6th Int. Symp. Non-Photorealistic Animation and rendering, 2008, pp. 79–87.

[8]. X. Xu, L. Zhang, and T. T. Wong, "Structure-based ascii art," ACM Trans. Graph., vol. 29, no. 4, pp. 52:152:9, 2010.

[9]. P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," IEEE Trans. Pattern Anal. Mach. Intel., vol. 33, no. 5, pp. 898–916, 2011.

[10]. Z. Ren, J. Yuan, C. Li, and W. Liu, "Minimum near-convex decomposition for robust shape representation," in Proc. 2011 IEEE Int. Conf. Computer Vision (ICCV), 2011, pp. 303–310.

[11]. M. Nacenta, U. Hinrichs and S. Carpendale, "Fatfonts: Combining the symbolic, and visual aspects of numbers," in Proc. Int. Working Conf. Advanced Visual Interfaces, 2012, pp. 407–414.