

A New Technique for Adding Scribbles in Video Matting

Neven galal el gamel¹, Hossam El-Din Moustafa² and F. E.Z. Abou-Chadi³

¹ Department of Electronics & Communications Engineering
Faculty of Engineering, Mansoura University, Mansoura Egypt

² Department of Electronics & Communications Engineering
Faculty of Engineering, Mansoura University, Mansoura Egypt

³ Department of Electronics & Communications Engineering
Faculty of Engineering, Mansoura University, Mansoura Egypt

Summary

In this paper, a new algorithm for adding scribbles in video matting without any user interaction is introduced, to improve the matting results depending on a combination of different background subtraction algorithms. The performance of the proposed technique is evaluated and compared with previous techniques and the results have shown that it has good performance and minimizes the error.

1. Introduction

Video matting is a critical operation in commercial television and film production, giving a director the power to insert new elements seamlessly into a scene or to transport an actor into a completely new location [1]. In the matting or matte extraction process, a foreground element of arbitrary shape is extracted, or pulled, from a background image. The matte extracted by this process describes the opacity of the foreground element at every pixel. Combined with the foreground color, the matte allows an artist to modify the background or to composite (i.e., transfer) the foreground into a new background. For image matting most existing methods require the input image to be accompanied by a trimap. A trimap is the process of labeling each pixel as foreground, background, and unknown [2]. Most statistical matting approaches require the user to supply a trimap manually. Instead of requiring a carefully specified trimap, some recently proposed matting approaches allow the user to specify a few foreground and background scribbles (use two color black and white; black scribbles for background and white scribbles for foreground) as user input to extract a matte [3].

This paper proposes a new method for adding scribbles for video matting algorithms. It relies on using the output of a new combination method of background subtraction

algorithms, and the decision fusion algorithm output [4] to the proposed method for adding the scribbles. The organization of the rest of the paper is as follows: Section 2 describes the combination method for background subtraction algorithms. Section 3 describes the matting problem. The proposed algorithm is described in Section 4. Sec. 5 presents the suggested method of adding the scribbles. Sec. 6 presents the error measurements. Section 7 is the conclusion.

2. A Combination Method for Background Subtraction in Video Sequences

The combination algorithm consists mainly of two steps [4]. The first step is to apply some image processing procedure to the output of three algorithms: the frame difference method, the Mixture of Gaussian method, and the approximated median method. The second step is to apply a data fusion procedure to the output of the three algorithms. As shown in Figure (1)

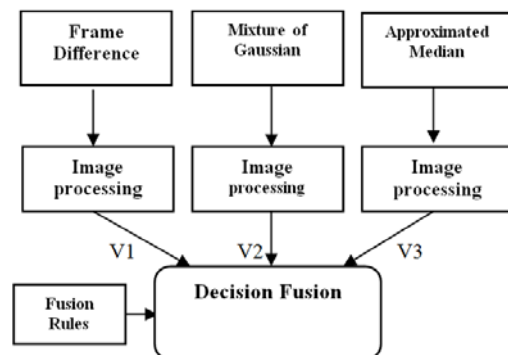


Fig.(1) A schematic diagram of the combination Algorithm[4].

The decision level fusion procedure was performed using the following Fusion Rules:

If $((V1 \& V2) \text{ OR } (V1 \& V3) \text{ OR } (V2 \& V3))$ is FG then $O(i,j) = FG$ (1)

If $((V1 \& V2) \text{ OR } (V1 \& V3) \text{ OR } (V2 \& V3))$ is BG then $O(i,j) = BG$ (2)

where:

$V1$: the output of the modified frame difference.

$V2$: the output of the modified Mixture of Gaussian.

$V3$: the output of the modified approximated median.

$O(i,j)$: is the pixel (i,j) of the output.

The result of the combination algorithm performed well compared to the performance of each algorithm individually as shown in Figure (2, 3).

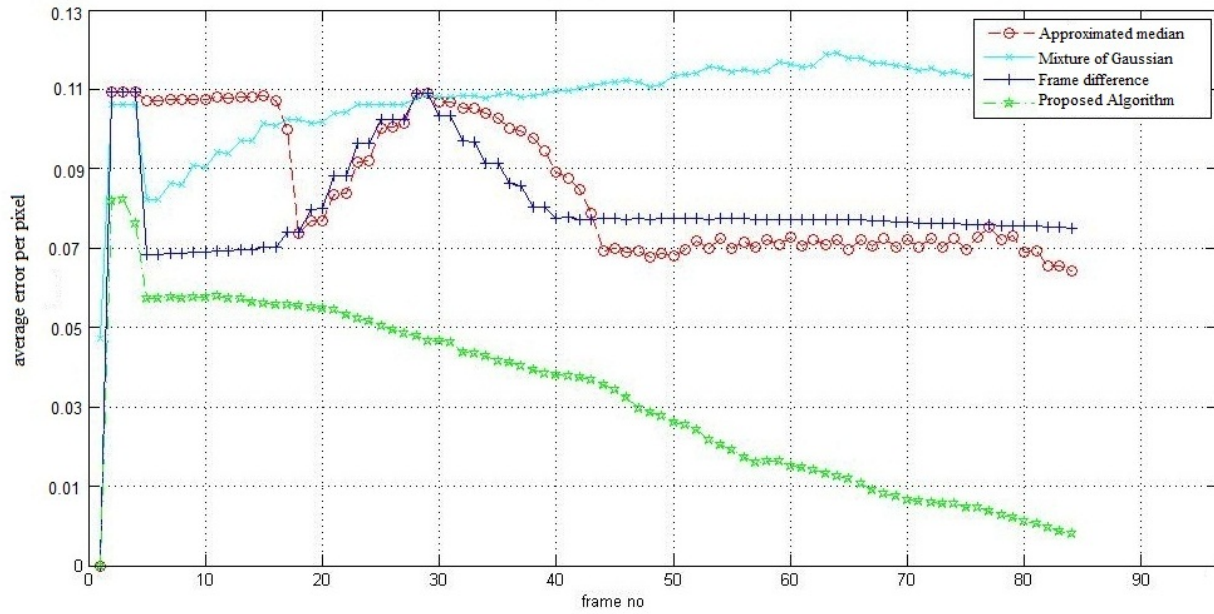


Fig. (2) The average error per pixel as a function of frame number (Images taken from [4])

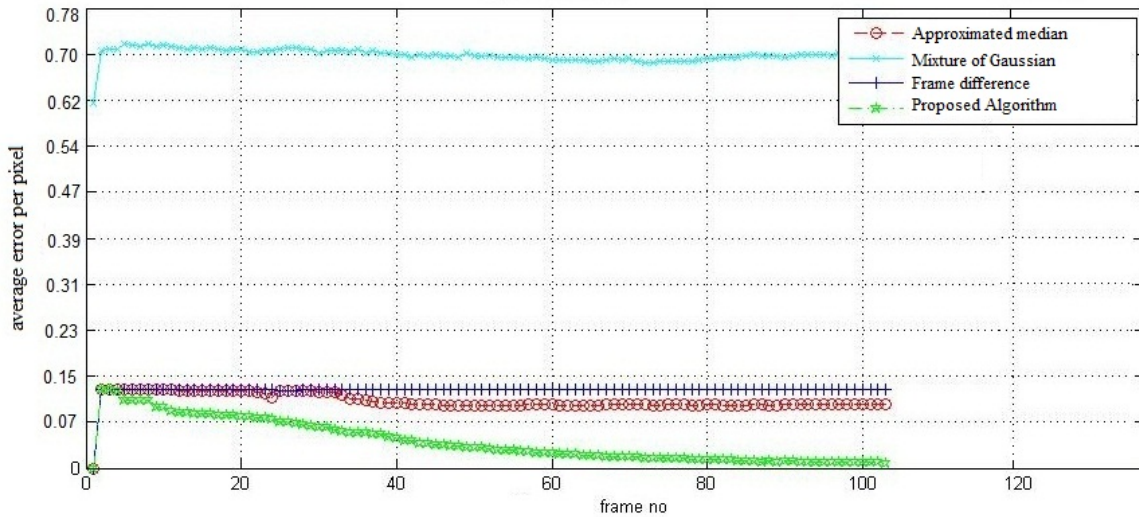


Fig. (3) The average error per pixel as a function of frame number (Images taken from [4])

3. Matting Problem

Extracting foreground objects from still images or video sequences plays an important role, thus it has been extensively studied for more than twenty years. This problem was mathematically established by Porter and Duff in 1984 [5]. Mathematically, the observed image $I_z(z = (x, y))$ is modeled as a convex combination of foreground image F_z and background image B_z by using the alpha matte α_z [2]:

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z \quad (3)$$

where α_z can be any value in $[0,1]$. If $\alpha_z = 1$ or 0 , pixel z is called definite foreground or definite background, respectively. Otherwise pixel z is called mixed. In most natural images, accurately estimating alpha values for mixed pixels is essential for fully separating the foreground from the background.

Given only a single input image, all three values α , F and B are unknown and need to be determined at every pixel location. The known information we have for a pixel are the three dimensional color vector I_z (assuming it is represented in some 3D color space), and the unknown variables are the three dimensional color vectors F_z and B_z , and the scalar alpha value α_z . Matting is thus inherently an under-constrained problem, since 7 unknown variables need to be solved from 3 known values. Most matting approaches rely on user guidance and prior assumptions on image statistics to constrain the problem to obtain good estimates of the unknown variables. To solve this problem a trimap was used; a trimap is a three color. It is a rough (typically hand-drawn) segmentation of the image into Three regions: white represents the foreground, black represents the background, and gray represents the unknown region [2]. Most recent methods expect the user to provide a trimap as a starting point. An example of a trimap is shown in Figure (4).

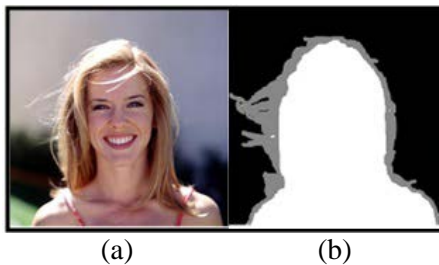


Fig. (4) (a): Original Frame, and (b) The Trimap Defined by Three Regions (Images taken from [6]).

Instead of requiring a carefully specified trimap, some recently proposed matting approaches allow the user to specify a few foreground and background scribbles as user input to extract a matte [2]. This intrinsically defines a very coarse trimap by marking the majority pixels (pixels haven't been marked by the user) as unknowns. The difficulty in using this technique for video matting, however, is the need for the user to create a trimap or adding scribbles at each frame.

4. The Proposed Algorithm

Constructing the scribbles depends mainly on the output of the decision fusion algorithm of the proposed algorithm and a fabricated hand drawing picture (fabricated with a simple paint program); which is constant for all the video frames as shown in Figure(6).

The steps of the new algorithm of adding the scribbles automatically are as follows:

- 1- Consider the output frame from the proposed decision fusion algorithm result taken from [4]; as shown in Figure (5).

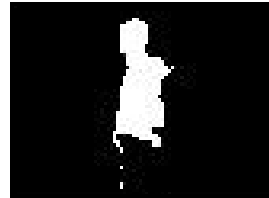


Fig.(5) The output of the decision fusion algorithm

- 2- Use a fabricated picture with two color blue and white the background is blue and there is small white rectangle on it (we can use any background color it does not effect on the performance) ;as shown in Figure(6).

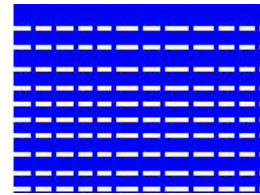


Fig.(6) The Fabricated Picture.

- 3- Transform the fabricated picture to black and white as shown in Figure (7).

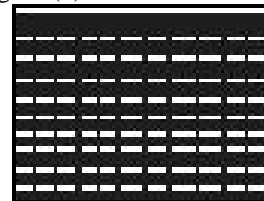


Fig.(7) Transforming the fabricated picture into black and white.

- 4- Multiply the processed frame of the output of the proposed decision fusion algorithm by the black and white fabricated picture, as shown in Figure (8).

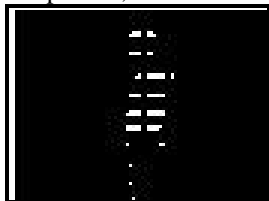


Fig.(8) The output of the Multiplication step

- 5- Compare the resultant frame from step 4 with the blue and white picture obtained in step 2 to obtain the final frame; as shown in Figure (9).

Let

I_2 : the output of step no.2

I_4 : the output of step no.4

I_5 : the output of step no.5

If (I_4 =white & I_2 =white) then I_5 =white

If (I_4 =black & I_2 =white) then I_5 =black

If (I_4 =black & I_2 =blue) then I_5 =blue

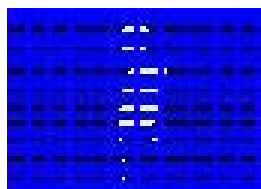


Fig.(9) The Output of the Comparison Step.

- 6- Replace the blue pixels by the original frame pixels. This results the image with the added scribbles, as shown in Figure (10).

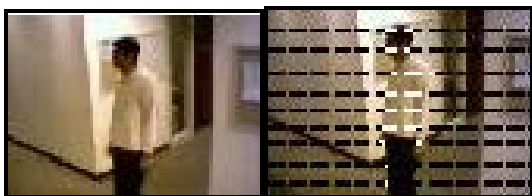


Fig.(10)The Resultant Image after Scribbles Addition process .

5. Improving the Process of Adding Scribbles

Having added the scribbles, it has been found that there is some error in the scribble addition process such that there is a black pixel on the foreground and a white pixel in the background Fig. (11). In order to circumvent this

problem; a number of steps was added to prevent or actually to reduce these errors.

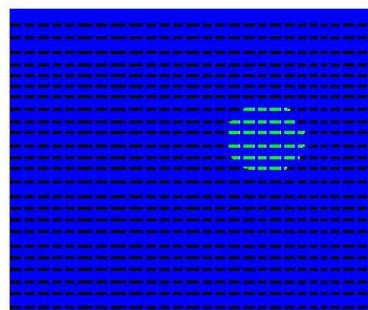


Fig. (11) Image with Scribbles and the Errors Appear in Red Circle.

It can be noted from Figure (9) that removing each black pixel adjacent to a white one in the same rectangle minimizes greatly the error.

This process is outlined as follows:

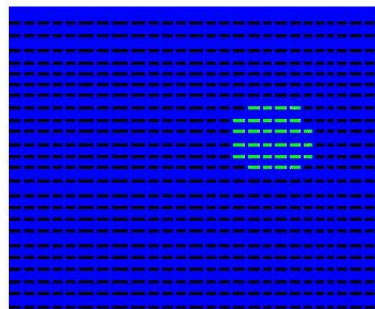
- 1- Test the 8-neighbours of each white pixel and replace the black pixel by a green one, as shown in Figure (12).



50

Fig (12) The Result of Replacing Pixels

- 2- Repeat this code many times corresponding to the width of the rectangle (number of pixels), as shown in Figure (13).



50

Fig.(13)The Output of the Repeated Code

- 3- Replace each green pixel by a blue one, as shown in Figure (14).

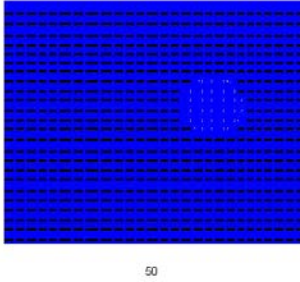


Fig. (14) The Output after Replacing Green Pixel by Blue One.

- 4- Replace each blue pixel by the original frame; this produces the picture after adding scribbles, as shown in Figure (15).

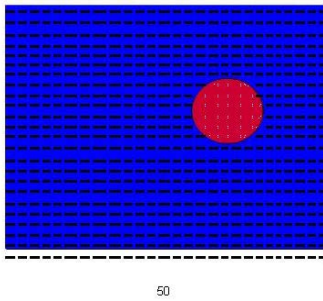


Fig. (15) The video Frame After Adding the Improved Scribbles on it.

Figure (16) shows the results after applying the video matting e procedures.

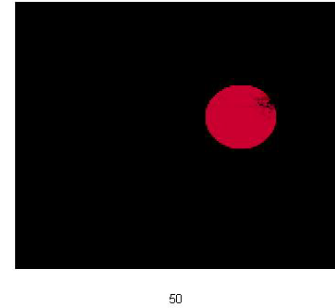


Fig. (16) The Matting Output After the Process of Adding Scribbles.

Figure (17) shows the result of applying the above modifications to another video sequences.

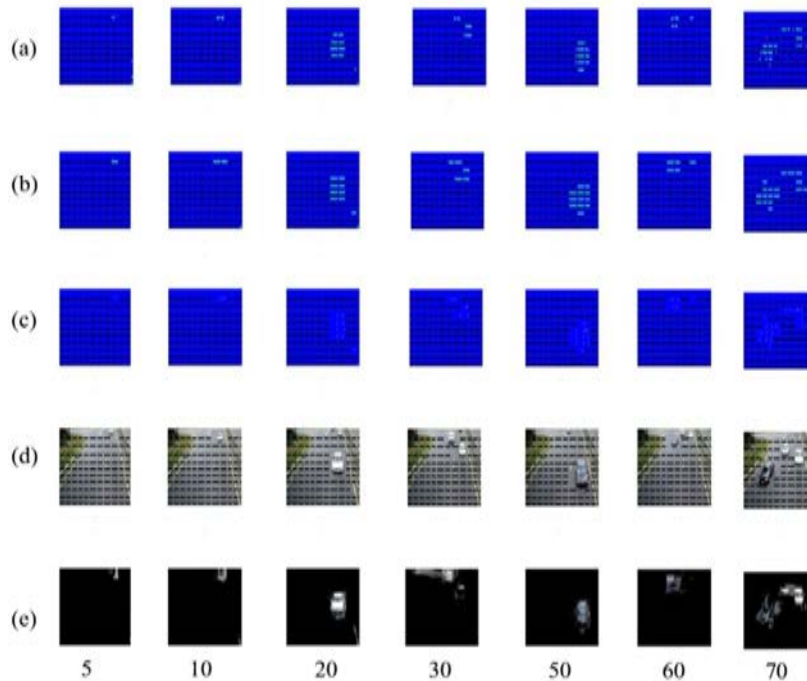


Fig.(17) The Steps and the Result of Adding Scribbles in Video Sequence Number2. Where (a) the output of adding scribbles; b: the picture after removing black pixel near the white one; c: replace green pixel by blue one; d: the frames after adding scribbles; and e:digital matting output.

6. Performance Measurements

The Average error per pixel in each frame was calculated by using Eq.(4). Using two fabricated videos with their corresponding separate foreground video (fabricated by flash program; and then converted to a video sequence) [7].

$$AEP(q) = \frac{1}{m \cdot n} \sum_{i=1}^n \sum_{j=1}^m |O(i, j) - f(i, j)| \quad (4)$$

where:

O(i,j): The pixel (i,j) of the output of each algorithm after transforming it to black and white frames.

f(i,j): The pixel(i,j) of the fabricated video accurate foreground.

n: Number of pixels in each row of the frame

m: Number of pixels in each column of the frame

q: Frame number.

AEP: Average error per pixel in each frame

Figures (18, 19) show the calculated average error per pixel in each frame.

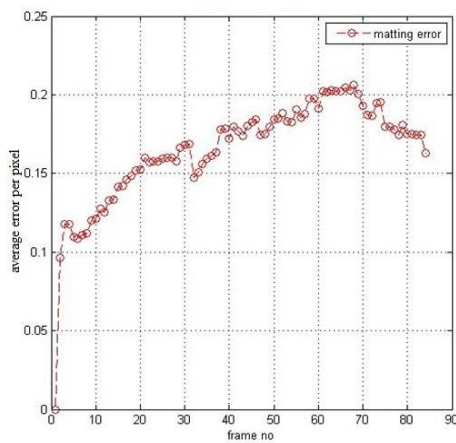


Fig.(18) The Matting average error per pixel in each frame for a fabricated video number1

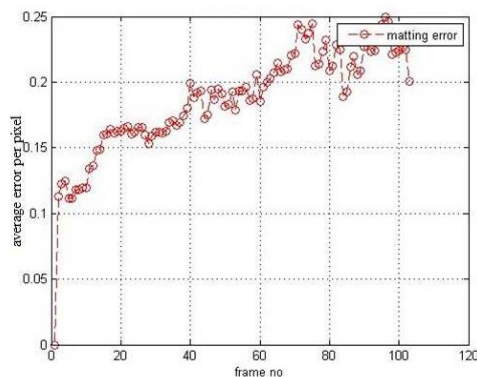


Fig.(19) The Matting average error per pixel in each frame for a fabricated video number 2

7. Conclusion

In this paper a new algorithm was introduced to add scribbles in video matting algorithms. The proposed algorithm use the output of a previously developed algorithm that uses data fusion criteria[4],using a fabricated picture to all the video frames of the video sequence, applying the matting algorithm the visual inspection shows good results. It is argued that the proposed algorithm of adding the scribbles shows good performance and minimizes the error.

References

- [1] Yung-Yu Chuang, Aseem Agarwala, Brian Curless, David H. Salesin, and Richard Szeliski, "Video Matting of Complex Scenes," <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.164.911&rep=rep1&type=pdf>
- [2] Jue Wang and Michael F. Cohen, "Image and Video Matting: A Survey," <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.190.1825&rep=rep1&type=pdf>
- [3] Anat Levin, Dani Lischinski, and Yair Weiss, "A Closed Form Solution to Natural Image Matting," School of Computer Science and Engineering The Hebrew University of Jerusalem <http://wenku.baidu.com/view/9d0bb062011ca300a6c39041.html> Access date 19/03/2013
- [4] Neven Galal El-Gamal, Hossam El-Din Moustafa, F. E.Z. Abou-Chadi "A New Combination Method for Background Subtraction in Video Sequences," proc. Of the 8th International Conference on Informatics and Systems (INFOS 2012), No.4, pp.23-28, 2012.
- [5] T. Porter and T. Duff, "Compositing digital images," Proc. of A special interest group of the Association for Computing Machinery (ACM SIGGRAPH), vol.18, no.3, pp. 253–259, July 1984.
- [6] Chuang, Y.-Y., Curless, B., Salesin, D., and Szelisiki, R., "A Bayesian approach to digital matting," Proc. of Computer Vision and Pattern Recognition (CVPR 2001), vol. 2, pp. 264 – 271, 2001.
- [7] "Convert SWF to AVI with Powerful Features", <http://www.avi-swf-convert.com>.