Unsupervised Automation of
Photographic Composition Rules in Digital Still Cameras

Serene Banerjee and Brian L. Evans

Embedded Signal Processing Laboratory, Center for Perceptual Systems
The University of Texas at Austin, Austin, TX 78712-1084 USA
{serene,bevans}@ece.utexas.edu

ABSTRACT

When taking pictures, professional photographers apply photographic composition rules, e.g. rule of thirds. The rule of thirds says to place the main subject's center at one of four places: at 1/3 or 2/3 of the picture width from left edge, and 1/3 or 2/3 of the picture height from the top edge. This paper develops low-complexity unsupervised methods for digital still cameras to (1) segment the main subject and (2) realize the rule-of-thirds.

The main subject segmentation method uses the auto-focus filter, opens the shutter aperture fully, and segments the resulting image. These camera settings place the main subject in focus and blur the rest of the image by diffused light. The segmentation utilizes the difference in frequency content between the main subject and blurred background. The segmentation does not depend on prior knowledge of the indoor/outdoor setting or scene content.

The rule-of-thirds method moves the centroid of the main subject to the closest of the four rule-of-thirds locations. We first define an objective function that measures how close the main subject placement obeys the rule-of-thirds, and then reposition the main subject in order to optimize the objective function. For multiple main subjects, the proposed algorithm could be extended to use rule-of-triangles by adding an appropriate constraint.

1. INTRODUCTION

To make a photograph more appealing, professional photographers apply a wealth of photograph composition rules. This research focuses on automating selected photograph composition rules to improve the quality of pictures taken by amateur photographers. So, the camera provides alternatives obeying the photograph composition rules to the user while acquiring pictures. The applications are good for personal use or for professionals who need to document, such as realtors and architects.

The two parts to solving the proposed problem are (i) identifying the main subject(s) in the photograph, and (ii) placing the main subject(s) and postprocessing the background by following the photograph composition rules. This paper proposes an unsupervised automated method for identifying the main subject in the photograph that leverages camera settings. After segmentation of the main subject the paper also automates two photograph composition rules of *rule-of-thirds* and *background blurring*. The proposed work could be extended to two or more main subjects.

To detect the main subject, the user should point the camera at the main subject. The auto-focus filter then focuses the main subject. Before or after the user takes the picture, the shutter aperture is opened completely to allow the lens optics to blur anything not in the plane of focus (where the main subject is). This complementary picture taken by the camera has the main subject in focus, and the rest of the picture blurred with diffused light. The difference in frequency content between the main subject and the background is exploited by using appropriate filtering to segment the main subject.

To place the main subject, the *rule-of-thirds* can be followed. Here the canvas is divided into three parts along the width and height, respectively. The center of the main subject should be placed at one of the four places: at 1/3 or 2/3 of the picture width from left edge, and 1/3 or 2/3 of the picture height from the top edge. After segmenting the main subject, the main subject is relocated on the canvas, following the rule-of-thirds.

Background blurring is either introduced to enhance the sense of motion where the main subject is moving or decrease the depth-of-field of the picture where the main subject is stationary. After main subject segmentation background blurring is simulated with region-of-interest filtering.

The paper organization is as follows. Section 2 discusses previous research in main subject detection, and highlights the importance for near real-time main subject detection. Section 3 describes the main subject detection process. Section 4 illustrates automation of rule-of-thirds. Section 5 describes simulated background blurring. Section 6 discusses the implementation complexity of the proposed algorithm. Section 7 concludes the paper.

2. BACKGROUND

Previous work on main subject detection was targeted towards detecting the main subject in offline settings. Luo, Etz, Singhal, and Gray[1,2] proposed a Bayes net based framework to detect the main subject. The network is trained on a training set and tested on a test set. Their method involves (a) region segmentation, (b) perceptual grouping, (c) feature extraction, and (d) probabilistic reasoning and training. Their method requires training time, which limits its application to detect the main subject on the fly.

In a wavelet-domain approach, Wang, Li, Gray, and Wiederhold[3,4] analyze the statistics of the high-frequency wavelet coefficients to segment the focused regions in an image. The method computes a multi-level wavelet transform, feature extraction, and postprocessing step for boundary smoothing. The method uses Haar wavelets, which have transfer functions that are scaled versions of $1 + z^{-1}$ and $1 - z^{-1}$ for the lowpass and highpass filters, respectively. The Haar wavelets and feature extraction can be implemented in fixed-point arithmetic. Nonetheless, the wavelet-domain method is computationally intensive.

In a spatial-domain approach, Won, Pyan, and Gray[5] develop an iterative algorithm based on variance maps. Their method yields more accurate segmentation when compared to the aforementioned wavelet-based approach.[3,4] The proposed spatial-domain method is an iterative approach that further refines the answer using the watershed algorithm. The method is also computationally intensive.

3. MAIN SUBJECT SEGMENTATION

To detect the main subject on the fly, Banerjee and Evans[6] propose an in-camera main subject segmentation algorithm. The proposed approach utilizes digital still camera controls, such as the auto-focus filter and the software-controlled shutter aperture. Assuming the user points to the the main subject, the auto-focus filter puts the main subject in focus.[7,8] The shutter aperture is then opened all the way, so that the light from the out-of-focus objects do not converge as sharply as from the objects in focus. The resulting blur in the out-of-focus objects is used to detect the focused main subject using filtering, edge detection, and contour smoothing.

3.1. Mathematical Formulation

Let i be an intensity value in an n-dimensional Euclidean space, $\Re^n$, such that $i \in F$, where F is the *image domain*. The object and the background classes are represented by F_o and F_b, respectively, with $F_o \subset F$ and $F_b \subset F$. The objective is to separate F_o from F_b.

Suppose $i \in \Re^n$ is mapped to an n-dimensional space, Ω^n, induced by ∇i, where ∇ is the gradient operator. An n-tuple s in Ω^n is mapped such that $s \in G$, where G is the *gradient domain*. The gradient domain, G can be further partitioned as $G_H(s)$ and $G_L(s)$ domains:

$$G_H(s) = \{s | s \geq \delta \text{ and } s \in G\}$$

$$G_L(s) = \{s | 0 \leq s < \delta \text{ and } s \in G\}$$

Here, δ is the threshold. The high and the low frequency domains are represented by $G_H(s)$ and $G_L(s)$, respectively.

Now $G_H(s) \mapsto F_H(i)$ and $G_L(s) \mapsto F_L(i)$, where subscripts H and L associated with F correspond to transformation from high and low frequency domains, respectively. The mapping of G_H to F_H and G_L to F_L requires similar set of transformation and $G(s) = G_H(s) + G_L(s)$ in the intensity domain:

$$F(i) = aF_H(i) + bF_L(i) \tag{1}$$

Here, $b = ka$ and k is a constant.

Now, $i \in F$ lies in $F_H(i)$ or $F_L(i)$, with probabilities a and b, respectively, so $a + b = 1$. Possible choices of a and b could be $\frac{1}{(k+1)}$ and $\frac{k}{(k+1)}$, respectively. Then,

$$F_H(i) - F(i) = \frac{k}{k+1} \ (F_H(i) - F_L(i)) \tag{2}$$

For the digital still camera application, the main subject class, F_o, is in focus, and the background class, F_b, is blurred by widening the shutter aperture. The spatial-domain difference $F_H(i) - F(i)$ will have sharper gradients around F_o and smoother gradients around F_b. Segmentation of F_o is induced by this difference of gradient information.

To generate $F_H(i)$ and $F_L(i)$ in the $\Re^n$ domain, high and lowpass filters can be designed, respectively. For the highpass filter, the criterion will be to select the frequencies, so that $\nabla i > \delta$. Similarly, the lowpass filter will have frequencies so that $\nabla i < \delta$. The filter can be designed adaptively, and the choice of the filter coefficients will determine its characteristics.

3.2. Implementation

For the 2-D case, the above conditions are met with an image sharpening filter as modeled in Fig. 1. So,

$$g(x,y) = I(x,y) - I_{smooth}(x,y) \tag{3}$$

and

$$I_{sharp}(x,y) = I(x,y) + k \, g(x,y) \tag{4}$$

Thereby,

$$I(x,y) = \frac{1}{k+1} \, I_{sharp}(x,y) + \frac{k}{k+1} \, I_{smooth}(x,y) \tag{5}$$

and

$$I_{sharp}(x,y) - I(x,y) = \frac{k}{k+1} \left(I_{sharp}(x,y) - I_{smooth}(x,y)\right) \tag{6}$$

Subtracting the user-intended image from the sharpened image generates an edge map in which the edges around the main subject are sharper than the background edges. Hence, the problem of segmenting the main subject reduces to separating the regions with the sharper edges from the regions with smeared ones. Banerjee and Evans[6] show that by combining the sharpening operation, difference calculation, and the edge detection to detect the main subject, the implementation complexity is similar to that of a 5×5 filter.

In Figs. 3(a), 4(a) and 5(a), the main subjects are in focus, while the background blur is achieved by a wider shutter aperture. Figs. 3(b), 4(b) and 5(b), show that after sharpening the image and taking the difference, the main subject edges are stronger. With further edge detection, the results of locating the main subjects before contour closing are shown in Figs. 3(c), 4(c) and 5(c), respectively. Figs. 3(d), 4(d) and 5(d) show the detected main subject mask.

4. AUTOMATED PLACEMENT OF THE MAIN SUBJECT BY RULE-OF-THIRDS

The post-segmentation objective is to automatically place the main subject following the rule-of-thirds. The rule-of-thirds says to place the main subject in one of four places: at 1/3 or 2/3 of the picture width from left edge, and 1/3 or 2/3 of the picture height from the top edge. A mathematical measure is defined to check how close the picture follows the rule-of-thirds, and this measure is optimized to reposition the main subject.
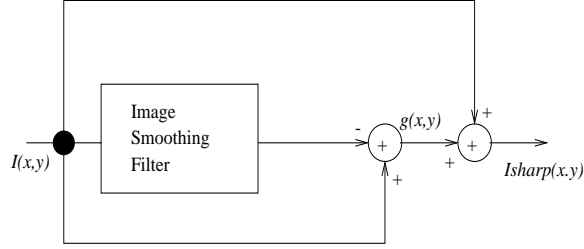
Figure 1. Model for an image sharpening filter

4.1. Mathematical Formulation

Let S be the scene domain of the main subject where $S \in \{\mathbf{v} | \mathbf{v} \in \text{Main subject}\}$ and $\mathbf{v} = (x_1, y_1), (x_2, y_2), ..., (x_i, y_i)$ are the pixel positions. Then, the center of mass is defined as the weighted sum of the components and cardinality of the scene domain. Consider that there are n main subjects. The center of mass for each of them is computed independently. A 2-dimensional function, $f(x, y)$ is defined such that it reaches a minimum when a center of mass is at the one-third position in the canvas both along the x and y axis. The objective will be to minimize the summation of the value of the function generated by the center of mass positions (x'_n, y'_n) of the n main subjects.

4.2. Implementation

For the current implementation, we chose $n = 1$, and the function $f(x, y)$ is product of the Euclidean distance from the four one-third corners in the canvas. Let (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) be the four one third corners. And, (x, y) is the position of the center of the mass of the main subject. Then,

$$f(x, y) = ((x - x_1)^2 + (y - y_1)^2)((x - x_2)^2 + (y - y_2)^2)((x - x_3)^2 + (y - y_3)^2)((x - x_4)^2 + (y - y_4)^2) \quad (7)$$

So, $f(x, y) \geq 0$ with $f_{min}(x, y) = 0$, and the minimum is attained when the center of mass is at one of the one-third corners. Thus, after computation of the center of mass, the image pixels are shifted so that they fall at a one-third corner.

The center of mass is computed along the rows and columns respectively. For each row (or column) if w_n is the number of "ON" pixels in the detected main subject mask , then the center of mass is defined as:

$$center = \frac{w_n * \text{row (or column) location}}{\Sigma w_n} \quad (8)$$

After computing the center of mass, a comparison is made as to which of the four one-third corners is closest to the current position of the center of mass. The picture is then shifted, so that the center of mass falls at the closest one-third corner.

The original pictures, Figs. 3(a), 4(a) and 5(a) have the main subject closer to the canvas center. Figs. 3(d), 4(d) and 5(d) show the detected main subject masks, the 1/3 and 2/3 lines on the canvas along the height and width, respectively, and the position of the center of mass of the detected main subject. Figs. 3(e), 4(e) and 5(e) show the main subjects repositioned following the rule-of-thirds.

For multiple main subjects in the photograph, the proposed algorithm can be extended for automation of the *rule-of-triangles*. The rule-of-triangles states that when there are more than one main subjects in the picture, then their center of mass should not lie on the same line in the canvas, but should form triangles on the canvas. This can be automated by adding an additional constraint during minimization, that no two center of masses lie on the same row in the canvas.
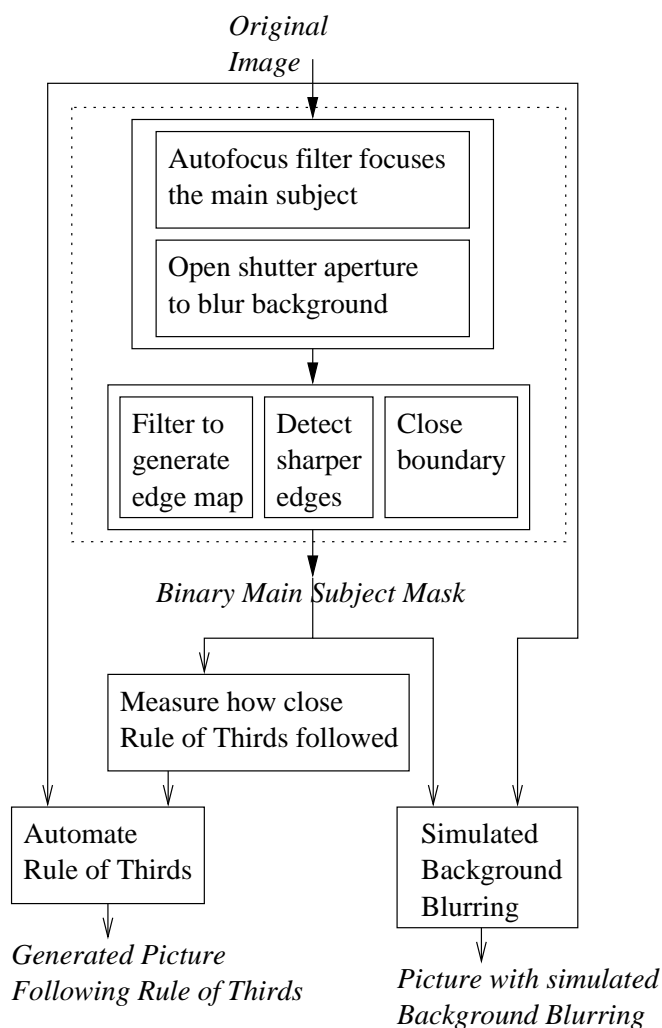
Original
Image

Autofocus filter focuses
the main subject

Open shutter aperture
to blur background

Filter to
generate
edge map

Detect
sharper
edges

Close
boundary

Binary Main Subject Mask

Measure how close
Rule of Thirds followed

Automate
Rule of Thirds

Simulated
Background
Blurring

Generated Picture
Following Rule of Thirds

Picture with simulated
Background Blurring

Figure 2. Proposed automation of selected photograph composition rules for digital still camera

5. SIMULATED BACKGROUND BLUR

For background blur simulation, the original image is first masked with the detected main subject mask. Then region of interest filtering is performed on the masked image, so that the main subject pixels remain unaltered, whereas artistic effects can be added to the background.

For this paper, the images are convolved with a motion blur filter that simulates linear and radial blurs produced by horizontal and rotational movement of the camera. The filtering involves convolving the image with a series of filters and compositing the filtered images. Figs. 3(f), 4(f) and 5(f) show simulated background blurring that could have resulted from camera panning. The current example simulates a linear motion of the camera by 10 pixels. Other values of linear, radial, or zoomed motion blurs can also be simulated.

6. IMPLEMENTATION COMPLEXITY

The proposed algorithm is shown in Fig. 2. The main subject mask can be detected with 18 multiply-accumulates, 4 comparisons and 6 memory accesses per pixel.[6] The post-segmentation complexity will depend on the number of rules that are automated.

Using the main subject mask, background blurring realized by a 3×3 filter takes 9 multiply-accumulates and 4 memory accesses per pixel. Using the main subject mask, the rule-of-thirds algorithm requires 2 multiply-accumulates, 1 comparison, and 1 or 3 memory access per pixel, as well as four comparisons and one division (explained below). One memory access per pixel is needed to calculate the center of mass. An additional two memory accesses per pixels is neded only if the picture is shifted instead of cropped.

For automated placement of the main subject following the rule-of-thirds, the center of mass for the detected main subject mask is computed with 2 multiply-accumulates and 1 memory read per pixel, and one division. The closest one-third corner is computed with four comparisons. The next step is to alter the picture so that the center of mass of the main subject is at one of the four one-third corners.

One approach is to crop the picture so that the center of mass of the main subject falls on one of the one-third corners. This is computationally very simple. Instead of cropping the picture, every pixel in the entire image could be shifted by the same amount so that the center of mass of the main subject occurs at one of the one-third corners. After shifting the image, many pixel values along two of the edges of the image would be undefined. These pixels could be given values through pixel replication along the boundary of known pixel values. The shifting approach requires one memory read and one memory write per pixel.

In the best case, the center of mass falls at one of the one-third corners so that the image does not have to be altered. In the worst case, the center of mass is at one of the corners of the picture, so that one-third of the rows and one-third of columns would be cropped or need to be given new values. In the average case, e.g. if the main subject were originally in the middle of the picture, one-sixth of the rows and one-sixth of the columns would be cropped or be given new values.

A digital still camera uses approximately 160 digital signal processor instruction cycles per pixel. Automating background blurring or rule-of-thirds requires fewer digital signal processing cycles. The proposed algorithms are amenable for implementation completely in fixed-point data types and arithmetic.

7. CONCLUSION

This paper proposes to automate selected photograph composition rules to provide feedback to the amateur photographer when taking pictures. The first step is to detect the main subject in the photograph. Either before or after the user takes a picture, the shutter aperture is opened fully to blur the objects that are out-of-focus using lens optics. The resulting frequency difference between the main subject and out-of-focus regions is used to detect and segment the main subject. The detection and segmentation of the main subject is similar in implementation complexity to a 5×5 filter. The algorithm could be extended to pictures with multiple main subjects.

After detecting the main subject, this paper automates two of many photographic composition rules. One rule places the main subject following the rule-of-thirds and the other simulates background blur. The proposed algorithms to automate the rules have computational requirements less than current digital still cameras use to compress and decompress an image using JPEG.
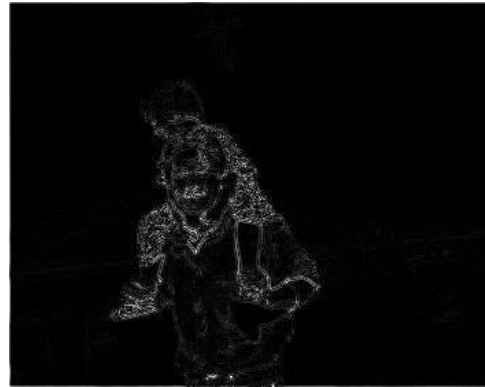
The proposed algorithms are amenable to implementation in fixed-point data arithmetic in digital signal processors, which are commonly used in digital still cameras. The algorithms avoid any *à priori* training information as does the Bayes net approach to detect the main subject.[1,2] The algorithms also avoid the iterations or computationally intensity of the hierarchical wavelet transform.[3,4] The algorithms do not depend on the scene setting or content.

REFERENCES

1. S. P. Etz and J. Luo, "Ground Truth for Training and Evaluation of Automatic Main Subject Detection," in *Proc. SPIE Conf. on Human Vision and Electronic Imaging*, vol. 3959, pp. 434–442, Jan. 2000.
2. J. Luo, S. P. Etz, A. Singhal, and R. T. Gray, "Performance-Scalable Computational Approach to Main Subject Detection in Photographs," in *Proc. SPIE Conf. on Human Vision and Electronic Imaging*, vol. 4299, pp. 494–505, Jan. 2001.
3. J. Li, J. Z. Wang, R. M. Gray, and G. Wiederhold, "Multiresolution Object-of-Interest Detection for Images with Low Depth of Field," in *Proc. IEEE Int. Conf. on Image Analysis and Processing*, pp. 32–37, Sept. 1999.
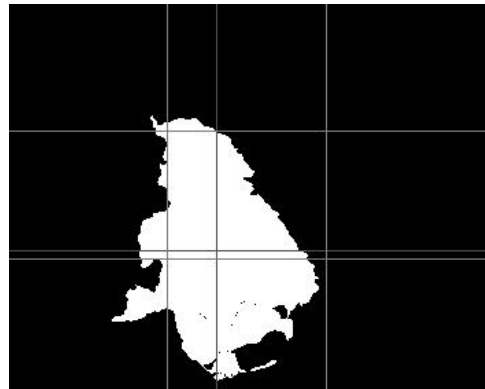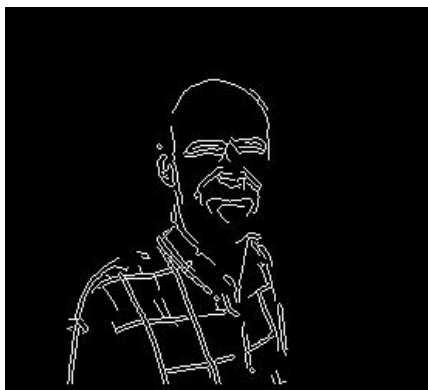
Figure 3. Automation of photographic composition rules by detecting the main subject, the man and the child, which are in focus: (a) Digital image with background blur from large shutter aperture; (b) Sharper edges are prominent in the filtered image; (c) Rough outline of main subject; (d) Detected main subject mask, with center of mass not following the rule-of-thirds; (e) Generated picture obeying rule-of-thirds; (f) Simulated background blur which could result from camera panning.
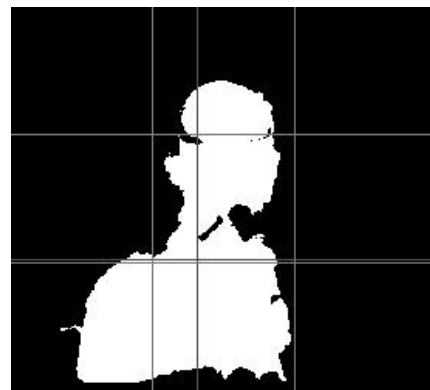
Figure 4. Automation of photographic composition rules by detecting the main subject, the man, which is in focus: (a) Digital image with background blur from large shutter aperture; (b) Sharper edges are prominent in the filtered image; (c) Rough outline of main subject; (d) Detected main subject mask, with center of mass not following the rule-of-thirds; (e) Generated picture obeying rule-of-thirds; (f) Simulated background blur which could result from camera panning.

Figure 5. Automation of photographic composition rules by detecting the main subject, the doll, which is in focus: (a) Digital image with background blur from large shutter aperture; (b) Sharper edges are prominent in the filtered image; (c) Rough outline of main subject; (d) Detected main subject mask, with center of mass not following the rule-of-thirds; (e) Generated picture obeying rule-of-thirds; (f) Simulated background blur which could result from camera panning.

4. J. Z. Wang, J. Li, R. M. Gray, and G. Wiederhold, "Unsupervised Multiresolution Segmentation for Images with Low Depth of Field," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 85–90, Jan. 2001.

5. C. S. Won, K. Pyan, and R. M. Gray, "Automatic Object Segmentation in Images with Low Depth of Field," in *Proc. IEEE Int. Conf. on Image Proc.*, pp. 805–808, Sept. 2002.

6. S. Banerjee and B. L. Evans, "A Novel Gradient Induced Main Subject Segmentation Algorithm for Digital Still Cameras," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, 2003.

7. N. N. K. Chern, P. A. Neow, and J. M. H. Ang, "Practical Issues in Pixel-Based Autofocusing for Machine Vision," in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 3, pp. 2791–2796, May 2001.

8. C. H. Park, J. H. Paik, Y. H. You, H. K. Song, and Y. S. Cho, "Auto Focus Filter Design and Implementation Using Correlation between Filter and Auto Focus Criterion," in *Proc. IEEE Int. Conf. on Consumer Electronics*, pp. 250–251, June 2000.