

Evaluating the Rule of Thirds in Photographs and Paintings

Seyed Ali Amirshahi^{1,2}, Gregor Uwe Hayn-Leichsenring²,
Joachim Denzler¹ and Christoph Redies^{2,*}

¹ Computer Vision Group, Department of Computer Science,
Friedrich Schiller University, Jena, Germany

² Experimental Aesthetics Group, Institute of Anatomy I, University of Jena School of
Medicine, Jena University Hospital, Jena, Germany

Received 28 April 2013; accepted 16 November 2013

Abstract

The rule of thirds (ROT) is one of the best-known composition rules used in painting and photography. According to this rule, the focus point of an image should be placed along one of the third lines or on one of the four intersections of the third lines, to give aesthetically pleasing results. Recently, calculated saliency maps have been used in an attempt to predict whether or not images obey the rule of thirds. In the present study, we challenged this computer-based approach by comparing calculated ROT values with behavioral (subjective) ROT scores obtained from 30 participants in a psychological experiment. For photographs that did not follow the rule of thirds, subjective ROT scores matched calculated ROT values reasonably well. For photographs that followed the rule of thirds, we found a moderate correlation between subjective scores and calculated values. However, aesthetic rating scores correlated only weakly with subjective ROT scores and not at all with calculated ROT values. Moreover, for photographs that were rated as highly aesthetic and for a large set of paintings, calculated ROT values were about as low as in photographs that did not follow the rule of thirds. In conclusion, the computer-based ROT metrics can predict the behavioral data, but not completely. Despite its proclaimed importance in artistic composition, the rule of thirds seems to play only a minor, if any, role in large sets of high-quality photographs and paintings.

Keywords

Experimental aesthetics, computational aesthetics, aesthetic quality assessment, rule of thirds, saliency map, photography, painting

* To whom correspondence should be addressed. E-mail: christoph.redies@med.uni-jena.de

1. Introduction

In 1797, John Thomas Smith proclaimed that the rule of thirds represents a more harmonizing proportion to follow in paintings of rural scenery than any other proportion (Smith, 1797). Ever since, the rule of thirds has been referred to as an important principle governing the spatial composition of aesthetic images. Although different descriptions have been given for the rule of thirds, most if not all of them suggest that, in order to create a photograph or painting of high aesthetic quality, the main object or focus point should be placed along one of the two imaginary horizontal or the two imaginary vertical lines that divide the image into nine equal parts (the third lines). For example, in his treatise, Smith (1797) advocated to fill the painting area by “two thirds of one element (as of water) to one third of another element (as of land)”, rather than placing their border in the center of the painting.

According to art critics and experts in photography, the rule of thirds is one of the most important composition rules used in painting and photography (Gooch *et al.*, 2001; Mai *et al.*, 2011; Meech, 2004; Peterson, 2011). In one version of the rule of thirds, the focus point in an image is placed along one of the third lines (Fig. 1a). In another version, the focus point is placed on one of the four intersections of the third lines. Figure 2a–e shows photographs that follow the rule of thirds (from the dataset of Mai *et al.*, 2011). In these photographs, the main focus point or object is on or near the third lines or on their intersections. From the same dataset (Mai *et al.*, 2011), photographs that do not follow the rule of thirds are shown in Fig. 2f–j.

Others, for example Arnheim in his study of compositional balance in images of simple geometric forms (Arnheim, 1982), stressed the perceptual

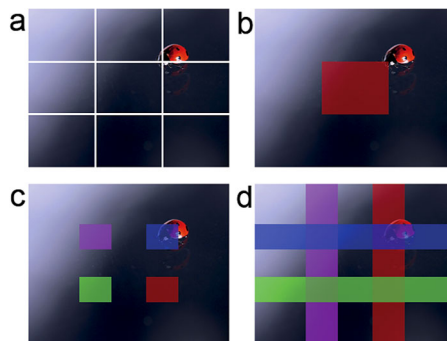


Figure 1. Illustration of the rule of thirds (ROT). The third lines (a) and the regions used in the calculation of different ROT measures (b–d) are marked on the photograph. Previously, Datta *et al.* (2006) and Li and Chen (2009) used a central rectangle covering the third lines to compute their ROT measure (b). Following the study by Mai *et al.* (2011), we calculated a saliency-based ROT measure either for four boxes overlying the intersections of the third lines (c) or for stripes overlying the third lines (d).

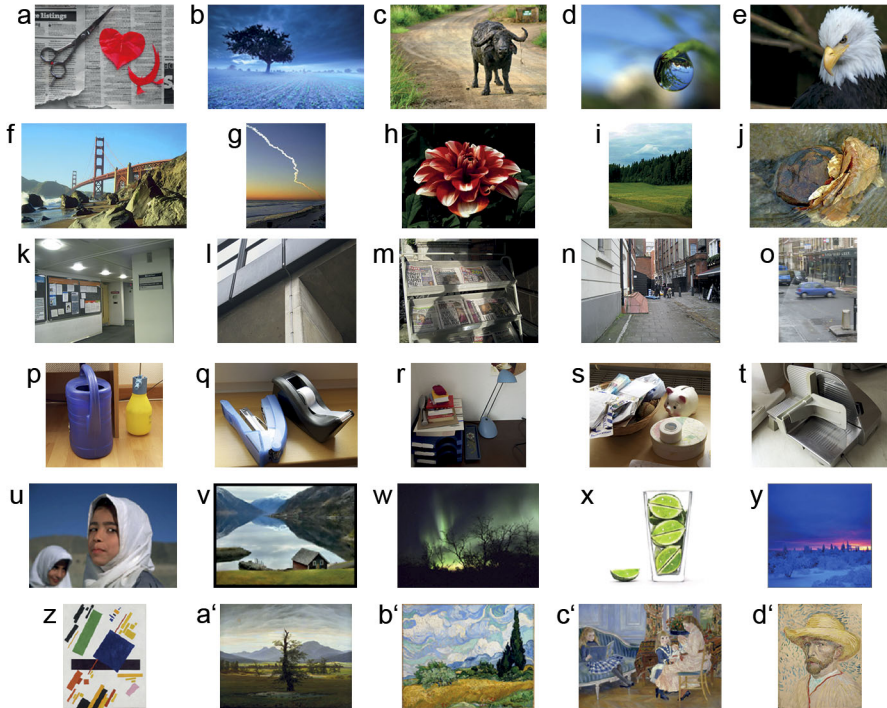


Figure 2. Examples of photographs analyzed in the present study. (a–e) Photographs that follow the rule of thirds (Mai *et al.*, 2011). (f–j) Photographs that do not follow the rule of thirds (Mai *et al.*, 2011). (k–o) Photographs of scenes taken almost randomly (McManus *et al.*, 2011a). (p–t) Photographs of simple scenes with one or a few objects (Redies *et al.*, 2012). (u–y) High-quality photographs downloaded from the Photo.net website. (z–d′) Paintings similar to the ones analyzed in the present study. (z) Abstract painting by Kazimir Malevich, 1916, Stedelijk Museum. (a′) Painting of complex natural scene by Caspar David Friedrich, 1822, Alte Nationalgalerie. (b′) Complex natural scene by Vincent van Gogh, 1889, Metropolitan Museum of Art. (c′) Complex scene with persons by Pierre-Auguste Renoir, 1884, Alte Nationalgalerie. (d′) Portrait painting by Vincent van Gogh, 1887, Van Gogh Museum. The paintings shown in (z–d′) are in the public domain and were downloaded from the Wikimedia Commons database (Google Art Project).

importance of the center of the image and the objects that it contains. Tyler (1998) reported a general tendency in art portraits that one of the eyes of the person depicted is centered along the midline in rectangular paintings (Tyler, 1998). However, other researchers have found no evidence for such a bias (McManus and Thomas, 2007). Palmer *et al.* (2008) reviewed some of the discrepancies between the on-center and off-center theories. They conducted a series of psychological experiments that confirmed the on-center bias, but only for forward-facing symmetrical objects; left-facing and right-facing objects tended to be located off-center. For vertically positioned objects, people preferred positions in the image that reflected spatial asymmetries in their

functional properties and the typical position of the object relative to the observer (Sammartino and Palmer, 2012). Moreover, Leyssen *et al.* (2012) demonstrated that semantics plays a role in images containing two meaningful objects. Observers preferred images in which semantically related objects were close together and unrelated objects were far apart.

McManus *et al.* (2011a) extended this type of investigation to other major axes in the rectangular frame of photographs. They critically evaluated Arnheim's theory of visual balance and found no evidence to support it (McManus *et al.*, 2011a). Nevertheless, this and a previous study by the authors (McManus *et al.*, 2011b) demonstrated that there are reproducible rules, according to which naive participants and experts reach aesthetic judgments on photographs.

Some recent psychological studies provided precise physical definitions for concepts such as balance in images and investigated large numbers of images (e.g., see McManus *et al.*, 2011a, b), but some of the earlier studies were less clear on how they measure image properties, or they based their results on a rather limited number of images. In the field of computational aesthetics, well-defined local and global physical properties were measured in large sets of images by modern digital image processing tools (Hoenig, 2005). However, the computational studies focused largely on simple, low-level features that can hardly explain the variability of aesthetic responses amongst humans or top-down cognitive processes involved in aesthetic judgments (Leder *et al.*, 2004). Also, claims in computational aesthetics that the statistical properties relate to perceptual phenomena were sometimes not supported by data from rigorous psychological experiments. Therefore, a combination of both approaches is warranted to study the determinants of individual preference.

In computational aesthetics, an increasing amount of studies on the quality assessment of images and videos was published in the last few years (Amirshahi and Larabi, 2011; Eskicioglu, 2000; Luo and Tang, 2008). Recently, there has been a shift of interest towards the aesthetic quality assessment of paintings and photographs (Amirshahi *et al.*, 2012, 2013; Bhattacharya *et al.*, 2010; Cavalcanti *et al.*, 2006; Datta *et al.*, 2006; Li and Chen, 2009; Wu *et al.*, 2010; Xue *et al.*, 2012). Extracting multiple features from paintings and photographs is one of the common approaches taken in computational aesthetics (Bhattacharya *et al.*, 2010; Cavalcanti *et al.*, 2006; Datta *et al.*, 2006; Li and Chen, 2009; Xue *et al.*, 2012). Most metrics share common features, for example, exposure of light, colorfulness, saturation, hue, aspect ratio, shape convexity, etc. Interestingly, the rule-of-thirds feature is amongst the features commonly used to assess the composition in paintings and photographs. For example, Datta *et al.* (2006) extracted 56 features and Li and Chen (2009) extracted 40 features from the image and used these features in their pooling

system to come up with a score to assess the aesthetic quality of photographs and paintings, respectively, based on subjective high quality or low quality ratings. In both metrics, the rule of thirds is used as a feature. An aspect common to both metrics is that the computations are restricted to the central window outlined by the third lines (Fig. 1b). The arguments behind this selection were that, normally, the object positioned on one of the third lines tends to stretch towards the central point (Datta *et al.*, 2006) and that the observer tends to focus on the center part of the painting (Li and Chen, 2009).

Mai and collaborators (2011) developed a computer-assisted method to predict whether a photograph respects the rule of thirds or not. They based their method on maps of saliency for each image. Saliency maps highlight image regions that are different from their surroundings and are calculated with the aim of predicting average looking behavior. Mai *et al.* (2011) built a classifier that achieved around 80% accuracy in automatically detecting the rule of thirds in a photograph. Similar or related methods have been used in unsupervised cropping methods for digital photography (Banerjee and Evans, 2004; Wong and Low, 2011; Zhang *et al.*, 2005). However, at present, the computer-based approaches cannot take into account semantic meaning, which also contributes to the appearance of image content as salient for human observers. Such top-down processing has an effect on looking behavior and, consequently, on the subjective assessment of whether an image obeys the rule of thirds (Borji *et al.*, 2013).

There are many different ways of defining what may stick out in an image, and, consequently, there are many metrics to calculate saliency maps. In a recent survey of 35 of such methods, Borji *et al.* (2013) systematically studied their accuracy in predicting eye movements. Accuracy varied greatly between the different methods, depending on how accuracy was defined and the perceptual task involved (e.g., looking at synthetic *versus* natural images), but some methods consistently performed better than others. In the realm of art, Wallraven *et al.* (2009) recorded eye movements in participants who viewed paintings from different Western art period, and compared the fixation maps with two different types of saliency maps. They found a high, but not perfect correlation between the saliency maps and the behavioral data. Fuchs *et al.* (2011) obtained similar results with a dataset of abstract and figurative paintings and observed that the effect of saliency on eye fixations is short-lived.

In the study by Mai *et al.* (2011), the authors did not correlate the performance of their automatic rule-of-thirds (ROT) classifier with subjective ROT ratings. It also remained unclear whether the photographs that followed the rule of thirds were perceived as more aesthetic. These issues will be addressed in the present study by comparing computed ROT values and subjective ROT rating scores. We also propose an algorithm that calculates ROT values in real

time without the need of training a classifier. Three different methods for computing saliency were compared (Achanta *et al.*, 2009; Harel *et al.*, 2007; Itti *et al.*, 1998). As controls, we studied two additional datasets that do not follow the rule of thirds. Finally, we studied photographs that were rated as highly aesthetic by the users of a web-based photography forum, and artworks of Western provenance. We asked the following questions:

- (1) How well do the calculated ROT values correlate with subjective ROT scores?
- (2) How high are calculated ROT values for high-quality photographs and artists' paintings of Western provenance? Do paintings differ in their values, depending on the content that is depicted in the paintings (abstract art, portraits, natural scenes or complex scenes with persons)?
- (3) Are photographs that follow the rule of thirds perceived as more aesthetic than those that do not?

2. Material and Methods

2.1. Image Datasets Used

We used five different image datasets in our experiments to assess the performance of the proposed metrics and for evaluating the rule of thirds in paintings and photographs. Sample images from each image category are shown in Fig. 2.

2.1.1. Rule-of-Thirds Photographs

This dataset consists of 679 photographs, which were randomly selected from a dataset of 2089 photographs that follow the rule of thirds and were collected by Mai *et al.* (2011). Examples are shown in Fig. 2a–e. As a control, we randomly selected 403 photographs from a dataset of images by the same authors (Mai *et al.*, 2011) that do not follow the rule of thirds (Fig. 2f–j). However, images were selected so that the mean values for the two groups matched approximately in terms of self-similarity, complexity and anisotropy (Redies *et al.*, 2012). Both types of images were gathered from the photo-sharing websites Flickr (<http://www.flickr.com>) and Photo.net. Selection was based on the question of whether they did or did not follow the rule of thirds (Mai *et al.*, 2011); whether the images were aesthetic or not, did not play a role in their selection, especially in the case of the photographs downloaded from the Flickr website. The two datasets do not share any images with the dataset to be introduced in Section 2.1.4.

2.1.2. *Photographs Taken Almost Randomly*

As another control, we analyzed a dataset of 606 photographs kindly provided by Prof. Chris McManus, University College, London (McManus *et al.*, 2011a). Image size was 2048 × 1536 pixels. The photographs were taken with a Canon Ixus 82 IS digital camera while the photographer was walking down streets or parks, sitting on buses or trains, in buildings, or other locations. The photographer made an explicit attempt to avoid pointing the camera at objects. Moreover, where possible, photographs were sampled at a regular interval to avoid a selection bias for particular objects or scenes. As a result, the photographs are examples of images that do not follow the rule of thirds. Sample photographs from the dataset are shown in Fig. 2k–o.

2.1.3. *Photographs of Simple Scenes of Objects*

This dataset consists of 200 photographs of simple (non-complex) scenes with one or only a few household and laboratory objects. The images were taken using a 15.1 megapixel digital camera (EOS 500D with EF-85 mm f/3.5–5.6 IS USM lens; Canon, Tokyo, Japan), as described previously (Redies *et al.*, 2012). The main focus points of the images were the objects shown; the rule of thirds or other aesthetic criteria were not intentionally followed. Figure 2p–t represents sample images from this dataset. This database is available for public use through the website of our research group (<http://www.inf-cv.uni-jena.de/en/aesthetics>).

2.1.4. *High-Quality Photographs (Photo.net)*

This dataset consists of 200 photographs downloaded from the photo-sharing website Photo.net (<http://www.photo.net>). The images were randomly selected from photographs that have an aesthetic rating of more than 5.5 out of a scale of 1 to 7. The ratings were given by the members of the website, that is, they were peer-reviewed by professional and amateur photographers. Figure 2u–y represents sample images from this dataset of high-quality photographs.

2.1.5. *Paintings*

This dataset consists of 727 paintings and comprises artworks of around 200 Western painters from a wide range of different centuries and a large variety of diverse art styles. The paintings were produced by well-known artists and collected by prestigious museums. The paintings were scanned from high-quality art books by members of our group using a calibrated digital scanner (Perfection 3200 Photo, Epson). Figure 2z–d' represents sample images that are similar to paintings in the dataset. For copyright reasons, we cannot reproduce the exact paintings in this publication.

To test whether the subject matter depicted in the paintings had an effect on the calculated ROT values, we classified the majority of the paintings based on their content. To reach reliable results, four categories of subject matters

that contained large numbers of paintings were selected from the dataset. The categories were: abstract artworks (188 paintings; Fig. 2z), natural scenes (54 paintings, Fig. 2a', b'), complex scenes with persons (151 paintings; Fig. 2c'), and portraits (191 paintings; Fig. 2d').

2.2. Image Calculations

Computational methods to study whether image composition complies with the rule of thirds have been proposed previously. As examples, we mentioned the studies by Datta *et al.* (2006) and Li and Chen (2009) in the Introduction section. In the two studies, rule-of-thirds-related features are used as an input value in a pooling system that is designed to evaluate the aesthetic quality of a painting or photograph. By themselves, these features cannot predict the overall aesthetic quality of an artwork.

Mai *et al.* (2011) calculated different features based on saliency maps and used them in a classifier to determine whether an image followed the rule of thirds or not. Saliency maps have been used in different image and video quality metrics as well as in the field of computational aesthetics (Amirshahi and Larabi, 2011; Borji *et al.*, 2013; Fuchs *et al.*, 2011; Mai *et al.*, 2011; Wallraven *et al.*, 2009; Wong and Low, 2011). Extending the work by Mai and colleagues, we propose a simple, robust and fast method to assess the rule of thirds. Unlike the previous approach, which is a supervised method, our method is unsupervised and does not require prior training of a classifier.

Over the last few years, a number of different ways to calculate visual saliency maps have been proposed (Borji *et al.*, 2013). In the present work, we employ three well-established metrics, which we refer to as the frequency-tuned (FT) method (Achanta *et al.*, 2009), the ITTI method (Itti *et al.*, 1998), and the graph-based visual saliency (GBVS) method (Harel *et al.*, 2007). In the FT method, the distance between the Lab pixel vector in a Gaussian filtered image and the mean Lab vector for the image is calculated. The ITTI method is based on the use of the Gaussian blur filter. This filter is applied on the image in a pyramid manner and the difference between the original image and level 4 of the pyramid is calculated. The GBVS method takes the same approach as the ITTI method, but the calculations are done at higher levels (Harel *et al.*, 2007).

In the proposed approach, we first calculated the saliency map for each image. The previous methods by Datta *et al.* (2006) and Li and Chen (2009) were based on calculating mean values over a central region in the original image (Fig. 1b) in separate color channels in the HSV and HSL color space. In the present study, we determined the maximum sum of saliency over four different regions that were related to the rule of thirds. Two approaches for selecting the regions of interest were followed. In the first approach, we used four boxes that were each centered on one of the intersections of the third lines

(Fig. 1c) as the regions of interest. In the second approach, we introduced four stripes that were each centered on one of the third lines (Fig. 1d). The stripes and the boxes had a width and height that corresponded to 10%, 16% or 20% of the width and height of the image, respectively. Next, we calculated the mean saliency value for each of the four regions and then took the maximum value as a measure for the saliency placed on the third lines. For details of the calculations, see the Supplementary Appendix.

Initially, we calculated stripes and boxes with different widths for all three saliency methods (FT, ITTI and GBVS). Because the GBVS method with boxes of a 16% width gave the best correlations with subjective ratings of the degree, to which an image followed the rule of thirds (Supplementary Table 1), this combination of parameters was used for the rest of the study.

To investigate whether saliency maps can be used to evaluate the existence of the rule of thirds, we measured the mean saliency over all images in each dataset introduced above. For each image, saliency maps were normalized to sum up to 1. To calculate the mean saliency map for each image category, we first resized each image in the category to 1024×1024 pixels so that the saliency maps of all images could be added to one another on a pixel-by-pixel basis. The result was then divided by the number of images in each category.

2.3. Rating Experiments

2.3.1. Participants

Twenty participants between 19 and 41 years (mean age: 26.8 ± 5.9 SD, nine males) rated the aesthetics of presented images (Experiment 1). Ten other participants between 22 and 33 years (mean age: 26.1 ± 3.4 SD, four males) estimated the degree to which the images followed the rule of thirds.

2.3.2. Stimuli

We used the dataset described in Mai *et al.* (2011) who separated the dataset into a subset of photographs that follow the rule of thirds, and a subset that does not obey the rule of thirds. From each subset, we randomly selected 100 images (200 images in total). The majority of photographs were color images, but grayscale images were also used. To achieve a consistent display on the screen, we reduced the size of all images to 800 pixels on the longer side. Smaller images were shown at their original size. In Experiment 2, we drew thin white lines on the images to indicate the third lines (Fig. 1a). All images were presented with a maximum size of 20.5 cm (longest side) on a computer screen. Participants viewed stimuli from a distance of 90 cm (assured by a chin rest). Hence, images covered a maximum of 13° of the visual angle.

2.3.3. Experiment 1: Rating of Aesthetic Appearance

First, we showed ten example images for 2 s each to get the participants acquainted with the database. Then, participants were asked to rate the images

according to their aesthetics on a mouse-based scale. The scale for the aesthetics rating was from 0 to 1 and continuous (100 steps, not visible to the participants) and endpoints were labeled with ‘not aesthetic’ and ‘aesthetic’, respectively (or the German equivalent to those terms). There was no time restriction. After response by clicking on the scale, the next image appeared. The 200 stimuli were presented separately and in random order on a black screen (Color Edge CG241W LCD monitor, EIZO Europe, Germany). Before each image appeared on the screen, the cursor of the mouse was set to the midpoint of the scale.

2.3.4. *Experiment 2: Rating of the Presence of the Rule of Thirds*

First, we showed the participants three sample images. One of the images followed the rule of thirds based on the fact that the focus of interest was placed along one of the third lines. Another image followed the rule of thirds based on the fact that the focus of interest was on one of the intersection of the third lines. The third image did not follow the rule of thirds. In the experiment, ten example images were shown for 2 s to introduce the database. Then, participants were asked to evaluate by mouse click whether or not the focus point of the images was on one of the third lines or, in another round, on one of the four intersections of the third lines. The participants were asked to choose between the left (‘no’) and the right side (‘yes’) of a scale. They also had the option of selecting intermediate positions on the scale so that they could indicate the degree to which an image followed the rule of thirds. There was no time restriction. After the response, the next image appeared. The 200 stimuli were presented separately and in random order on the black screen. Five participants evaluated the third lines first and the other five participants evaluated the intersection points first.

3. Results

3.1. *Subjective Rule-of-Thirds (ROT) Rating*

As a baseline for this study, we first asked ten participants to rate the degree to which 200 photographs followed the ROT. In separate sessions, participants evaluated two different ROT criteria that are commonly used. In one session, they assessed whether the focus of interest was on one of the third lines. In another session, they assessed whether the focus of interest was on one of the intersections of the third lines (Fig. 1). For the rating, we selected 100 images that followed the ROT (here called ROT+ images) and 100 images that did not follow it (ROT– images) from the dataset of Mai *et al.* (2011; see Section 2.1.1).

For the rating based on the intersections, the ROT+ photographs received an average score of 0.76 ± 0.17 SD on a continuous scale from 0 to 1. As

expected, the average score of ROT– photographs was much lower (0.23 ± 0.22 SD; $p < 0.0001$, two-tailed Mann–Whitney test; Cohen’s $d = 2.71$). The corresponding scores for the rating based on the third lines were 0.62 ± 0.22 SD and 0.35 ± 0.23 SD, respectively ($p < 0.0001$; $d = 1.21$).

3.2. Calculation of Saliency-Based ROT Measure

Next, we calculated saliency-based ROT values for the three methods (GBVS, ITTI and FS) with different parameters and asked which paradigm yielded the highest correlations with the subjective ROT ratings. The same 200 photographs were used as described in Section 3.1. Results in Supplementary Table 1 indicate that the highest correlations were obtained for the GBVS method with boxes of 16% width placed on the intersections of the third lines (Fig. 1c). For this metric, Fig. 3 shows a dot plot of the subjective ROT rating *versus* the calculated ROT measure. Largely confirming the initial classification of Mai *et al.* (2011), the majority of ROT– images (blue dots in Fig. 3) received not only low scores for subjective ROT ratings (see above) but also low values calculated for ROT (*mean* 0.069 ± 0.016 SD). In contrast, the ROT values calculated for the ROT+ images (red squares in Fig. 3), which received high subjective ROT scores, scattered more widely at a much higher level (*mean* 0.109 ± 0.058 SD; $p < 0.0001$, two-tailed Mann–Whitney test; $d = 0.94$). While the subjective ROT scores did not correlate with the calculated ROT values for the ROT– images, we found a moderate correlation between the two measures for the ROT+ images (red regression line in Fig. 3).

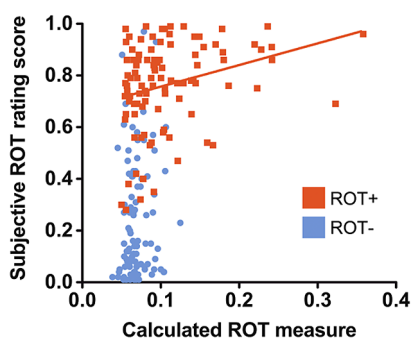


Figure 3. Rating scores for the degree to which the rule of thirds (ROT) was present in photographs, plotted as a function of the calculated ROT measure. Red squares represent results for the database of photographs that follow the ROT (ROT+). Blue dots represent results for the database of photographs that do not follow the ROT (ROT–). The red line represents the regression line for the ROT+ data points (Spearman correlation coefficient $\rho: 0.31$, $p < 0.002$). This figure is published in color in the online version.

Because the GBVS metric resulted in the highest overall correlation between calculated ROT values and subjective ROT scores (Supplementary Table 1), we used this metric in the rest of the study. This choice was in agreement with a recent survey of 35 saliency metrics, in which the GBVS method outperformed most of the other methods in eye movement prediction accuracy (Borji *et al.*, 2013).

We also compared the subjective ROT scores with the ROT values that were calculated with the methods introduced by Datta *et al.* (2006) and Li and Chen (2009). However, no correlation was found. We conclude that the two metrics do not help us in distinguishing between images that follow or do not follow the rule of thirds in our dataset of images.

Figure 4 displays rainbow-colored saliency maps for the images shown in Fig. 2 (same spatial arrangement). In the majority of images, the positions of high saliency values (yellow to red color) correspond to subjective focus points of interest.

Normalized average saliency maps for each image category are shown in Fig. 5. Confirming similar results by Mai *et al.* (2011), the average saliency

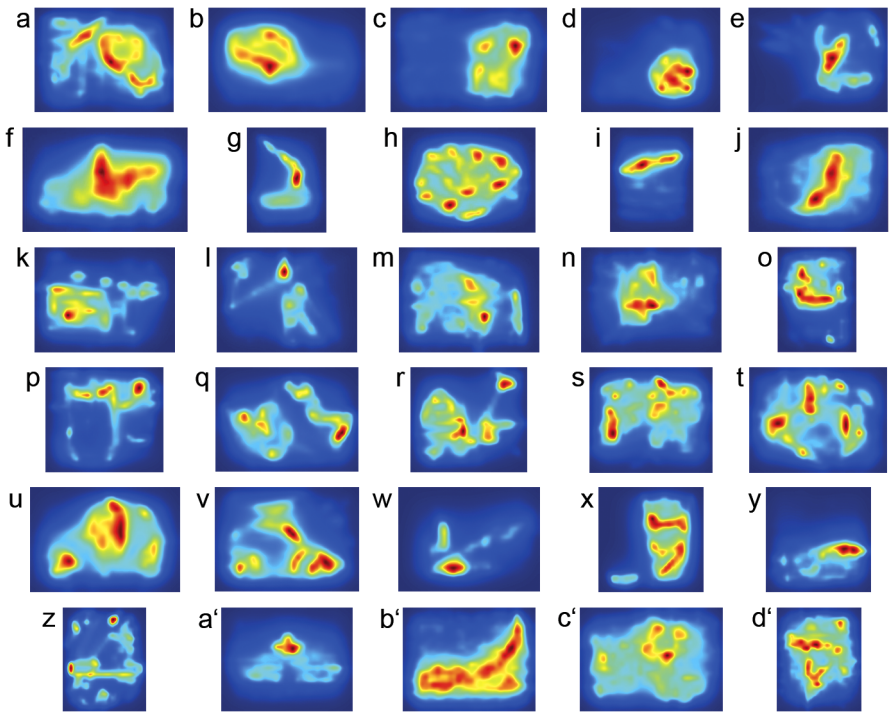


Figure 4. Saliency maps for the photographs displayed in Fig. 2 (same spatial arrangement). The maps were calculated by the GBVS saliency method introduced by Harel *et al.* (2007). The saliency values are rainbow-pseudocolor coded.

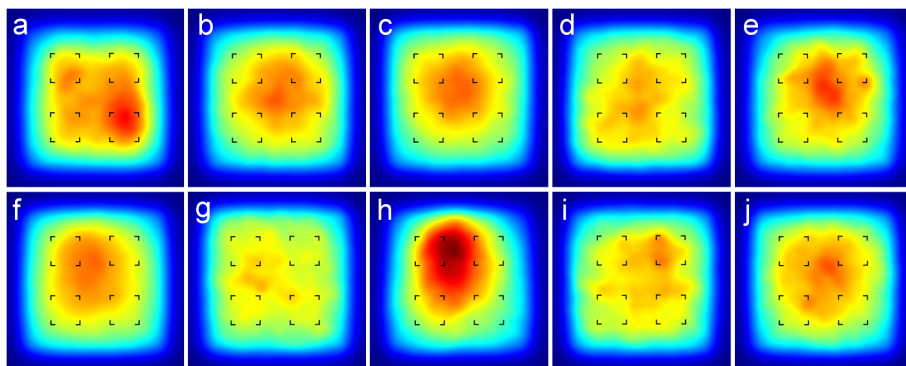


Figure 5. Mean saliency maps for the image categories analyzed in the present study. (a) 679 photographs that follow the rule of thirds (Mai *et al.*, 2011). (b) 403 photographs that do not follow the rule of thirds (Mai *et al.*, 2011). (c) 606 photographs of scenes taken almost randomly (McManus *et al.*, 2011a). (d) 200 photographs of simple scenes with one or a few objects (Redies *et al.*, 2012). (e) 200 high-quality photographs downloaded from the Photo.net website. (f) 727 paintings of Western provenance. From these paintings, mean saliency maps for 188 abstract artworks are shown in (g), for 191 portraits paintings in (h), for 54 natural scenes paintings in (i) and for 151 paintings of complex scenes with persons in (j). The black corners in each map delineate the boxes used for the calculation of saliency.

map for the ROT+ images (Fig. 5a) shows highest saliency values in the box that is centered on the lower right intersection of the third lines, but also relatively high values at the other intersections. By contrast, saliency is highest in the center of the ROT– photographs (Fig. 5b). A similar central peak of saliency is observed for the photographs of scenes taken almost randomly (Fig. 5c), the simple scenes of objects (Fig. 5d), and the Photo.net dataset of high-quality photographs (Fig. 5e). For the entire painting dataset (Fig. 5f), there is a central tendency for high saliency values, but peak saliency is shifted slightly towards the upper left corner. A separate analysis for the different categories of paintings reveals that this shift is especially prominent in portraits (Fig. 5h). The position of this peak may relate to the tendency of artists to place the faces, which are highly salient, into the upper regions of the paintings where they are close to or covered by the third lines. Average saliency in the other categories of painting categories (abstract art, Fig. 5g; natural scenes, Fig. 5i; and complex scenes with persons, Fig. 5j) is more evenly distributed across the images.

The above results were quantified by calculating the saliency values with the GBVS method (boxes overlying intersections, 16% width; see Fig. 1c) for each image separately. Average values for each image category are shown in Fig. 6. The results confirm that the ROT+ images have higher mean saliency values (0.102 ± 0.053 SD) than the ROT– photographs (0.067 ± 0.016 SD);

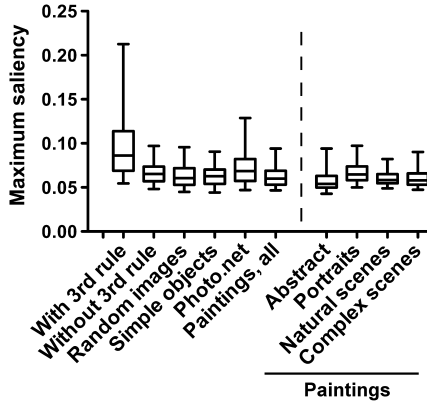


Figure 6. Mean calculated saliency values for all image categories analyzed. Maximum values for the four boxes shown in Fig. 1c were calculated for each image with the method by Harel *et al.* (2007) and averaged for each image category. ‘With 3rd rule’, photographs that follow the rule of thirds (Mai *et al.*, 2011); ‘Without 3rd rule’, photographs that do not follow the rule of thirds (Mai *et al.*, 2011); ‘Random images’, photographs of scenes taken almost randomly (McManus *et al.*, 2011a); ‘Simple objects’, photographs of simple scenes with one or a few objects (Redies *et al.*, 2012); ‘Photo.net’, photographs downloaded from the Photo.net website; ‘Paintings, all’, paintings of Western provenance. From these paintings, values are marked ‘Abstract’ for abstract artworks, ‘Portraits’ for portraits paintings, ‘Natural scenes’ for natural scene paintings, and ‘Complex scenes’ for paintings of complex scenes with persons. Whiskers represent the 5–95 percentiles.

$p < 0.001$; $d = 0.90$) as well as all the other image categories (for all other comparisons, $p < 0.001$, Kruskal–Wallis test with Dunn’s multiple comparison test; $d = 0.73$ to 1.73). Among the paintings, portraits have higher values (0.068 ± 0.016 SD) than abstract images (0.059 ± 0.018 SD; $p < 0.001$, $d = 0.46$) and complex scenes with persons (0.061 ± 0.013 SD; $p < 0.001$, $d = 0.55$).

3.3. Subjective Rating of Beauty

In the last part of this study, participants rated the aesthetics of the 100 ROT+ photographs and 100 ROT– photographs that had been previously assessed for the degree to which they complied with the rule of thirds (Section 3.1). The mean aesthetic rating score for the ROT+ images (0.59 ± 0.14 SD) was slightly higher than that of the ROT– images (0.54 ± 0.14 SD; $p < 0.01$, two-tailed Mann–Whitney test, $d = 0.36$). In Fig. 7, aesthetic rating scores were plotted as a function of the subjective ROT scores and the calculated ROT values. There was a weak overall correlation between the aesthetic rating score and the subjective ROT rating score (Spearman $\rho = 0.17$, $p < 0.05$; Fig. 7a). However, the aesthetic rating score did not correlate significantly with the calculated ROT measure (Fig. 7b).

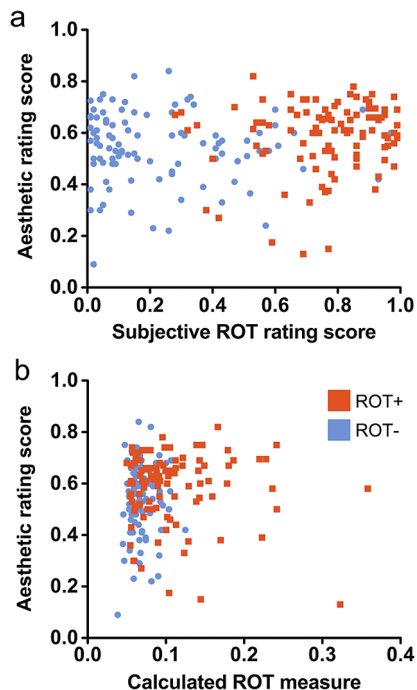


Figure 7. Aesthetic rating scores plotted as a function of the subjective rule-of-thirds (ROT) rating score (a) and the calculated ROT measure (b). Red squares represent results for the database of photographs that follow the ROT (ROT+). Blue dots represent results for the database of photographs that do not follow the ROT (ROT−). This figure is published in color in the online version.

4. Discussion

4.1. Computational versus Behavioral Measures for the Rule of Thirds

We used three different saliency metrics to calculate two rule-of-third (ROT) measures and validated the results in comparison with subjective data from participants who rated the degree to which photographs followed the rule of thirds in a psychological experiment. Compared to the other methods, the GBVS metrics (Harel *et al.*, 2007) with boxes centered on the intersections of the third lines, yielded the strongest correlations between the computed data and the subjective scores. The GBVS method was shown previously to be an outstanding metric to predict looking behavior in a comparative study of 35 saliency methods (Borji *et al.*, 2013).

Overall, the correlation between computed ROT values and behavioral scores was of intermediate strength (Spearman $\rho = 0.47$; Supplementary Table 1). For photographs that did not follow the rule of thirds on subjective grounds (ROT− images), low calculated ROT values were obtained on av-

erage (Fig. 6). The two control databases, which had not been generated to comply with the rule of thirds (almost random photographs, McManus *et al.*, 2011a; and simple scenes of objects, Redies *et al.*, 2012) also yielded low calculated ROT values. By contrast, average calculated ROT values were much higher for photographs that were rated as concordant with the rule of thirds in the present study (ROT+ images). However, computed values scattered widely and the correlation with the subjective ROT scores was relatively weak (Spearman $\rho = 0.31$; Figs 3, 6). Therefore, we conclude that factors other than saliency are likely to contribute to subjective ROT ratings. Presumably, top-down mechanisms that take the semantics of a photograph into account are amongst the most relevant factors. A similar combination of bottom-up (saliency-based) and top-down (cognitive) mechanisms has been invoked in the prediction of eye movements (Borji *et al.*, 2013). In conclusion, we demonstrate that low-level visual saliency contributes to the perception of the rule of thirds in photographs in a bottom-up fashion.

Strikingly, the mean saliency maps (Fig. 5) demonstrate that peak saliency in the rule-of-thirds (ROT+) photographs (Fig. 5a) is not evenly distributed across the boxes overlying the four intersections of the third lines. Instead, they are concentrated in the lower left box. It remains to be investigated whether this observation is unique to the dataset analyzed in the present study, corresponds to a more general tendency by photographers, or reflects other compositional rules.

The present results corroborate the study by Mai *et al.* (2011), which achieved a 72–80% classification rate of photographs that followed the rule of thirds *versus* photographs that did not. However, unlike the study by Mai *et al.* (2011), the method proposed by us does not require any prior training of a classifier. It represents an automated tool that instantly delivers a single measure, which can contribute to the question of whether a photograph complies with the rule of thirds. As such, practicing artists or photographers may use this method in isolation during their work. A robust, fast and reliable method for measuring the ROT feature might be valuable also for research in the field of experimental aesthetics.

4.2. Correlation of Aesthetic Ratings with Rule-of-Thirds Judgments and Measures

The rule of thirds is a feature that is widely used to assess the aesthetic quality of paintings and photographs (see Introduction). In the present study, we challenged the validity of this approach by relating aesthetic rating scores to ROT rating scores for a dataset of 200 photographs (Fig. 7a). The two scores correlated only weakly (Spearman $\rho = 0.17$). Moreover, there was no correlation between the aesthetic rating score and the calculated ROT measure (Fig. 7b). The mean calculated ROT values for aesthetic photographs (Photo.net dataset)

were only slightly higher than those of control photographs that do not follow the rule of thirds (Fig. 6). These results suggest that the rule of thirds plays only a minor, if any, role in the aesthetic evaluation of photographs.

4.3. ROT Values Computed for Paintings

Low calculated ROT values were also observed for the dataset of paintings. However, we did not obtain subjective rating scores for the aesthetic value of the paintings. Nevertheless, curators of prestigious art museums considered their acquisition worthwhile, probably because of the superior artistic quality of the paintings. Also, we did not rate the degree to which the paintings followed the rule of thirds. Nevertheless, the low calculated ROT values for paintings suggest that the rule of thirds, at least as far as it is based on saliency-based mechanisms, is not a decisive factor for determining the visual quality of the paintings.

Mean saliency maps for portrait paintings (Fig. 5h) showed a prominent peak close to the upper left box. This peak likely corresponds to the average position of the faces in portrait paintings. It results in relatively high saliency values close to the upper left ROT box and, consequently, leads to slightly higher calculated ROT values (Fig. 6). Average saliency in the other categories of paintings is more evenly distributed (Fig. 5g, i, j) and mean saliency values are about as low as in photographs that do not follow the rule of thirds.

4.4. Conclusions

In summary, our findings suggest that the rule of thirds might not be as important for the evaluation of the visual quality in photographs and artworks as previously assumed (see Introduction). Evidently, not following this rule does not necessarily result in images of low visual quality. We can only speculate why the rule of thirds plays such an important role in textbooks on photography and art. Perhaps, like the golden section, the rule of thirds mirrors the desire of artists and photographers to comprehend rules of artistic composition. Therefore, it might have become a normative aspect of creating artworks rather than a qualitative one. The rule of thirds may also help beginners to endow the products of their creativity with a particular visual structure under conscious control. Eventually, as artists gain intuitive expertise in artistic composition, they may drop the rule, which might be the reason why we did not find it in high-quality photographs and artworks.

Acknowledgements

We thank Steve Palmer, Chris McManus and an anonymous reviewer for highly constructive criticism and suggestions on a previous version of the manuscript. We are grateful to members of our groups for critical feedback

and discussion, and to Chris McManus for sharing a dataset of randomly taken photographs.

References

- Achanta, R., Hemami, S., Estrada, F. and Susstrunk, S. (2009). Frequency-tuned salient region detection, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, USA, pp. 1597–1604.
- Amirshahi, S. A., Koch, M., Denzler, J. and Redies, C. (2012). PHOG analysis of self-similarity in esthetic images, in: *Proceedings of SPIE (Human Vision and Electronic Imaging XVII)*, San Francisco, CA, USA, pp. 8291, 82911J.
- Amirshahi, S. A. and Larabi, M. (2011). Spatial-temporal video quality metric based on an estimation of QoE, in: *Proceedings of the Third International Workshop on Quality of Multimedia Experience (QoMEX)*, Menechem, Belgium, pp. 84–89.
- Amirshahi, S. A., Redies, C. and Denzler, J. (2013). How self-similar are artworks at different levels of spatial resolution?, in: *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, Anaheim, CA, USA, pp. 93–100.
- Arnheim, R. (1982). *The Power of the Center: A Study of Composition in the Visual Arts*. University of California Press, Berkeley, CA, USA.
- Banerjee, S. and Evans, B. L. (2004). Unsupervised automation of photographic composition rules in digital still cameras, in: *Proceedings of SPIE 5301*, San Jose, CA, USA, pp. 364–373.
- Bhattacharya, S., Sukthankar, R. and Shah, M. (2010). A framework for photo-quality assessment and enhancement based on visual aesthetics, in: *Proceedings of the International Conference on Multimedia*, Florence, Italy, pp. 271–280.
- Borji, A., Sihite, D. N. and Itti, L. (2013). Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study, *IEEE Trans. Image Proc.* **22**, 55–69.
- Cavalcanti, C., Gomes, H., Meireles, R. and Guerra, W. (2006). Towards automating photographic composition of people, in: *Proceedings of the IASTED International Conference on Visualization, Imaging, and Image Processing*, Palma de Mallorca, Spain, pp. 25–30.
- Datta, R., Joshi, D., Li, J. and Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach, in: *Proceedings of Computer Vision-ECCV 2006*, Graz, Austria, pp. 288–301.
- Eskicioglu, A. M. (2000). Quality measurement for monochrome compressed images in the past 25 years, in: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, TX, USA, pp. 1907–1910.
- Fuchs, I., Ansoerge, U., Redies, C. and Leder, H. (2011). Saliency in paintings: bottom-up influences on eye fixations, *Cogn. Comp.* **3**, 25–36.
- Gooch, B., Reinhard, E., Moulding, C. and Shirley, P. (2001). Artistic composition for image creation, in: *Proceedings of Rendering Techniques 2001 Eurographics*, London, UK, pp. 83–88.
- Harel, J., Koch, C. and Perona, P. (2007). Graph-based visual saliency, *Adv. Neural Inf. Process. Syst.* **19**, 545–552.

- Hoening, F. (2005). Defining computational aesthetics, in: *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Aire-la-Ville, Switzerland, pp. 13–18.
- Itti, L., Koch, C. and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis, in: *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1254–1259.
- Leder, H., Belke, B., Oeberst, A. and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments, *Br. J. Psychol.* **95**, 489–508.
- Leysen, M. H. R., Linsen, S., Sammartino, J. and Palmer, S. E. (2012). Aesthetics of spatial composition: Semantic effects in two-object pictures, *i-Perception* **3**, 25–49.
- Li, C. and Chen, T. (2009). Aesthetic visual quality assessment of paintings, *IEEE J. Sel. Top. Signal Process.* **3**, 236–252.
- Luo, Y. and Tang, X. (2008). Photo and video quality evaluation: Focusing on the subject, in: *Proceedings of Computer Vision — ECCV 2008*, Marseille, France, *Lecture Notes in Computer Science* **5304**, pp. 386–399.
- Mai, L., Le, H., Niu, Y. and Liu, F. (2011). Rule of thirds detection from photograph, in: *Proceedings of the IEEE International Symposium on Multimedia (ISM)*, Dana Point, CA, USA, pp. 91–96.
- McManus, I. C., Stöver, K. and Kim, D. (2011a). Arnheim's Gestalt theory of visual balance: Examining the compositional structure of art photographs and abstract images, *i-Perception* **2**, 615–647.
- McManus, I. C. and Thomas, P. (2007). Eye centering in portraits: A theoretical and empirical evaluation, *Perception* **36**, 167–182.
- McManus, I. C., Zhou, F. A., l'Anson, S., Waterfield, L., Stover, K. and Cook, R. (2011b). The psychometrics of photographic cropping: the influence of colour, meaning, and expertise, *Perception* **40**, 332–357.
- Meech, S. (2004). *Contemporary Quilts: Design, Surface and Stitch*. Batsford, London, UK.
- Palmer, S. E., Gardner, J. S. and Wickens, T. D. (2008). Aesthetic issues in spatial composition: Effects of position and direction on framing single objects, *Spat. Vis.* **21**, 421–449.
- Peterson, B. (2011). *Learning to See Creatively*. Crown Publishing, New York, NY, USA.
- Redies, C., Amirshahi, S. A., Koch, M. and Denzler, J. (2012). PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects, in: *Proceedings of the ECCV 2012 Ws/Demos*, Florence, Italy, *Part I, Lecture Notes in Computer Science* **7583**, pp. 522–531.
- Sammartino, J. and Palmer, S. E. (2012). Aesthetic issues in spatial composition: Effects of vertical position and perspective on framing single objects, *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 865–879.
- Smith, J. T. (1797). *Remarks on Rural Scenery*. Nathaniel Smith, London, UK.
- Tyler, C. W. (1998). Painters centre one eye in portraits, *Nature* **392**, 877–878.
- Wallraven, C., Cunningham, D., Rigau, J., Feixas, M. and Sbert, M. (2009). Aesthetic appraisal of art — from eye movements to computers, in: *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging (CAe09)*, Victoria, BC, Canada, pp. 137–144.
- Wong, L. K. and Low, K. L. (2011). Saliency retargeting: An approach to enhance image aesthetics, in: *IEEE Workshop on Applications of Computer Vision*, Kona, HI, USA, pp. 73–80.

- Wu, Y., Bauckhage, C. and Thureau, C. (2010). The good, the bad, and the ugly: Predicting aesthetic image labels, in: *20th International Conference on Pattern Recognition (ICPR)*, Istanbul, Turkey, pp. 1586–1589.
- Xue, S. F., Lin, Q., Tretter, D. R., Lee, S., Pizlo, Z. and Allebach, J. (2012). Investigation of the role of aesthetics in differentiating between photographs taken by amateur and professional photographers, in: *Electronic Imaging*, San Francisco, CA, USA, *Proc. SPIE 8302*, 83020D, DOI:10.1117/12.914686
- Zhang, M., Zhang, L., Sun, Y., Feng, L. and Ma, W. (2005). Auto cropping for digital photographs, in: *Proc. IEEE International Conference on Multimedia and Expo*, Amsterdam, The Netherlands, DOI:10.1109/ICME.2005.1521454.