Towards Segmentation from Multiple Cues: Symmetry and Color

R. Shor and N. Kiryati*
Department of Electrical Engineering–Systems
Tel Aviv University
Tel Aviv 69978, Israel

nk@eng.tau.ac.il

Abstract

Towards segmentation from multiple cues, this paper demonstrates the combined use of color and symmetry for detecting regions of interest (ROI), using the detection of man-made wooden objects and the detection of faces as working examples. A functional that unifies color compatibility and color-symmetry within elliptic supports is defined. Using this functional, the ROI detection problem becomes a five-dimensional global optimization problem. Exhaustive-search is inapplicable due its prohibitive computational cost. An adaptive random search method rapidly converges to the correct solution. The added value obtained by combining color and symmetry is demonstrated.

Index terms: Image segmentation, color symmetry, integration of multiple cues, global optimization, object recognition

1 Introduction

Most segmentation methods associate a scalar measurement or a vector of measurements with each pixel, characterizing its grey level, color or the texture to which it belongs. Once the measurement space has been defined, segmentation can be viewed as an optimization problem: regions should be uniform internally, different from adjacent ones and "reasonable" in their

*Corresponding author.

number and shapes. Formalizing these vague concepts in the form of an objective function and devising an efficient way to perform the search are both difficult problems.

The gestalt school suggested grouping principles that guide human perceptual organization. They include similarity, proximity, continuation, symmetry, simplicity and closure. Incorporating the gestalt principles in machine vision is an attractive idea [23]. In particular, the gestalt principles relate to *global* shape properties and represent *a-priori* visual preferences, issues that a successful segmentation method should address. Note, however, that studies in human perceptual organization are often limited to binary images, commonly to dot patterns. In computer vision, using binarized images requires successful edge detection, which is largely equivalent to segmentation, thus relying on the unknown. Application of the gestalt principles to image segmentation is thus desirable, but not straightforward.

Symmetry is one of the gestalt grouping principles. It appears in man made objects and is also common in nature [15, 27]. The omnipresence of symmetry has motivated many studies on symmetry in images, see e.g., [19, 20, 24, 29]. The possibility of rapidly finding symmetric areas in raw gray level images, as shown in [20], encourages the use of symmetry as a cue for segmentation. However, symmetry is not always maximal where expected. One example in [20] shows greater symmetry between a tree and its shadow than the symmetry of the tree itself. This indicates that symmetry alone is insufficient and that additional cues should be used. Our long term goal is to develop a unified low-level vision module, in which several basic visual tasks, each difficult when carried out separately, assist one another towards accomplishing their missions. This will simplify vision system design, require less top-down feedback and lead to more stable and robust performance.

This paper is a step towards computationally efficient symmetry-aided-segmentation. The idea of carrying out segmentation in conjunction with symmetry detection is not new in itself [16, 19, 20, 26, 28], but the concept is still in its infancy and much remains to be studied. Our approach is quite general, but is presented here in the context of two specific vision tasks: the detection of man-made wooden objects, and the well-studied frontal-view face detection task [2, 10]. Starting from a color image, the similarity of each pixel color to wood color or to skin color can be quantified. Many algorithms for face detection based on skin color are available, e.g. [1, 9, 11, 12]. Face detection using symmetry has also been considered, e.g. [3, 4, 9]. Our interest is in the *added value* obtained by performing segmentation *in conjunction* with symmetry detection, rather than as separate processes.

To maintain the generality of the method presented, we intentionally ignore highly specific and application-dependent cues that can be very useful in particular applications, such as the position and exact shape of facial features in the case of face detection. The algorithm can thus be applied, with minimal changes, to other computer vision problems, where roughly symmetric objects, characterized by some uniformity in an arbitrary measurement space, have to be rapidly detected.

Let D denote an elliptical domain at any location, orientation and scale within a color image I. D is thus characterized by five parameters. Let $S(D)$ denote a measure of the symmetry of the image within D, with respect to the major axis of D, taking color into account. Let $C(D)$ quantify the color-compatibility of D, i.e., the dominance of wood-colored (or skin-colored) pixels within D. Define a functional $F(D) \equiv \mathcal{F}\{S(D), C(D), D\}$ that combines symmetry, color compatibility and size. The global maximum of $F(D)$ corresponds to some elliptical domain D^* in which the image is highly symmetric and exhibits high color compatibility. The operational goal is to efficiently find D^*.

2 Color Compatibility

Identifying the "best" color-space for grouping tasks such as skin segmentation is still controversial. We carried out a modest performance evaluation, using images taken locally and some images from the University of Stirling face database [8]. We compared the following color-spaces: *RGB*, *YES* [9], *TSV* [11], *rgb* (normalized RGB: *NRGB*) [30], *HSV*, *XYZ*, $L^*U^*V^*$ and *xyz* [5]. It turned out that *YES*, *rgb* and *TSV* were most useful for skin segmentation, *rgb* yielding the best results in our tests. The *rgb* color space is defined by

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \tag{1}$$

The transformation from *RGB* to *rgb* is nonlinear. All values are normalized by intensity $(R + G + B)$ and $b = 1 - r - g$. Note that skin colors typical to people of different origins, including Asian, African American and Caucasian cluster in the *rgb* color space [30]. Furthermore, the *rgb* color space is insensitive to surface orientation and illumination direction [6]. The red points in Fig. 1 depict the scatter in the *rgb* color-space of skin-pixels from 35 face images. Colors of different wood types also cluster in the *rgb* space (blue points). Note the overlap between skin and wood colors (magenta).

Let $\mathbf{f}(i,j) = [r(i,j), g(i,j)]^T$ denote the normalized color vector at a specific pixel (i,j). We wish to obtain $c(i,j)$, a measure of the similarity of $\mathbf{f}(i,j)$ to a given family of colors, e.g., wood colors or skin colors. The method used is inspired by a skin detection algorithm presented in [9], but we use the *rgb* color-space while [9] uses *YES*. More important, in [9] $c(i,j)$ is a binary function, classifying pixels as either members or non-members of the family, while here intermediate similarity values are accommodated.

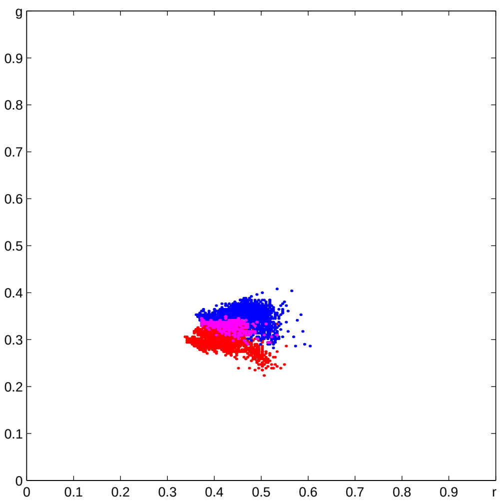

Figure 1: *Red:* A scatter-diagram showing the position in the *rgb* space (normalized RGB) of the skin pixels in 35 face images. Some of the images were imported from the Stirling database [8], a few were taken locally. *Blue:* The clustering of wood colors, taken from 7 images of objects made of dark and light wood. *Magenta:* Positions in which skin and wood colors overlap.

The class-conditional probability density function of skin-colored pixels can be reasonably modeled by a two dimensional Gaussian [30], where the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{K}$ are estimated from an appropriate training set. Our small-scale experiments indicate that a 2-D Gaussian model is suitable for wood-color as well. Equal-height contours of the color-family probability density function are ellipses in the r-g plane, whose centers are all at $\boldsymbol{\mu}$ and whose principal axes depend on $\mathbf{K}$. Table 1 shows the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{K}$ obtained for skin and wood from Fig. 1.

Family	$\boldsymbol{\mu}$	$\mathbf{K}$
Skin	$\begin{pmatrix} 0.41 \\ 0.31 \end{pmatrix}$	$\begin{pmatrix} 9.00 & 0.67 \\ 0.67 & 2.51 \end{pmatrix} \cdot 10^{-4}$
Wood	$\begin{pmatrix} 0.46 \\ 0.35 \end{pmatrix}$	$\begin{pmatrix} 10.2 & -0.6 \\ -0.6 & 2.7 \end{pmatrix} \cdot 10^{-4}$

Table 1: The mean vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{K}$ for skin and wood.

The similarity measure $c(i,j)$ is taken as the color-family probability density of $\mathbf{f}(i,j)$, i.e.

$$c(i,j) = \exp\left\{-\frac{1}{2}\left[\mathbf{f}(i,j)-\boldsymbol{\mu}\right]^T \mathbf{K}^{-1}\left[\mathbf{f}(i,j)-\boldsymbol{\mu}\right]\right\} \tag{2}$$

Visualizations of $c(i,j)$ are shown in Fig. 2. Given an image domain D, its color compatibility is quantified as

$$C(D) = \frac{1}{||D||}\sum_{(i,j)\in D} c(i,j). \tag{3}$$

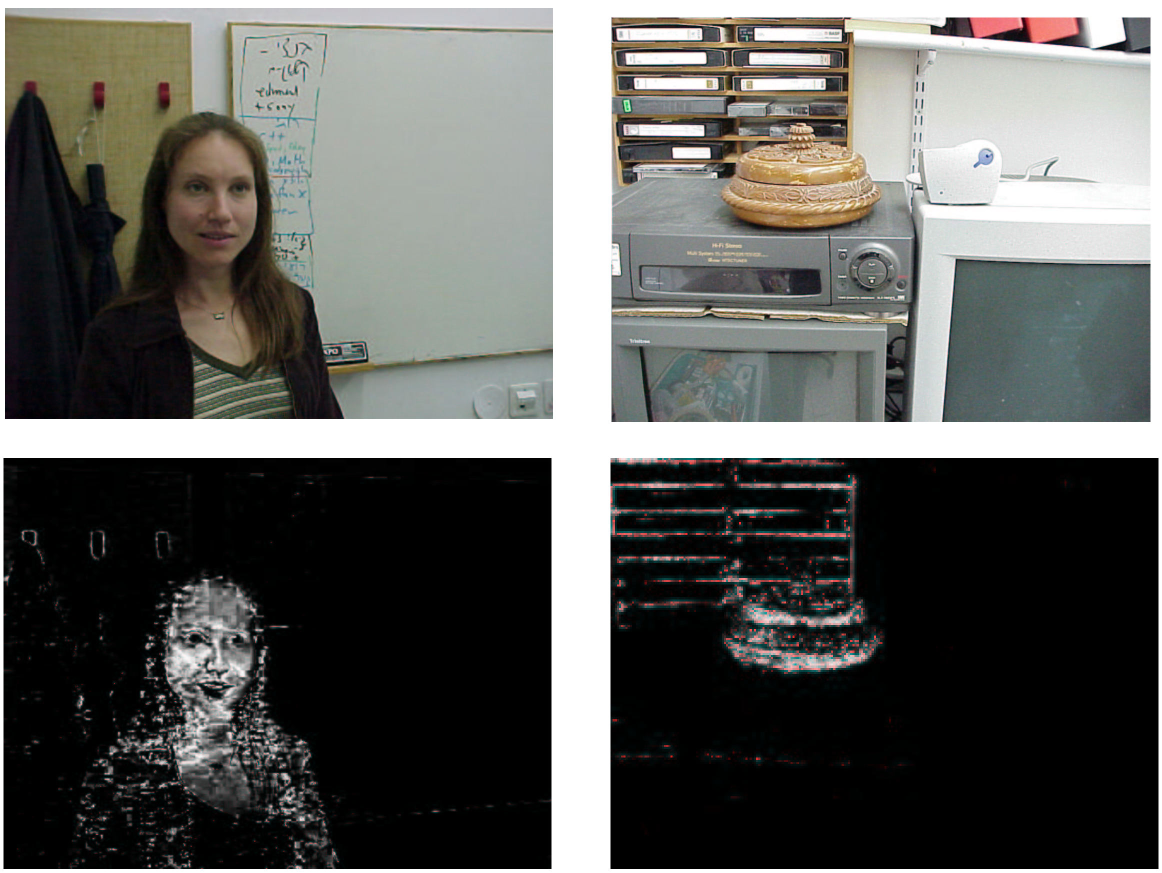

Figure 2: *Top:* Two color images. *Bottom:* The corresponding skin (left) or wood (right) similarity functions $c(i,j)$.

3 Color Symmetry Measurement

The mirror-symmetry of a continuous scalar function $f(x, y)$ with respect to the x-axis can be measured [20, 21, 31] by a reflectional correlation coefficient

$$S_f = \frac{\int\int f(x,y) f(x,-y) dx\, dy}{\int\int f^2(x,y) dx\, dy} \quad . \tag{4}$$

Note that any function f can be expressed as the sum of a fully symmetric component f_s and a perfectly anti-symmetric component f_{as}; it is easy to show that

$$S_f = \frac{||f_s||^2 - ||f_{as}||^2}{||f_s||^2 + ||f_{as}||^2} \ . \tag{5}$$

For non-negative functions f, $S_f \in [0, 1]$.

Symmetry measurement with respect to an arbitrary axis t, in a translated and rotated coordinate system (t, s), can be implemented by alignment of the (t, s) system with the (x, y) system, i.e., translation and rotation of the relevant sub-image to a standard position. In this research, local symmetry is measured in *elliptic* domains. This is accomplished by multiplying the relevant sub-image, in the standard position, by an elliptical Gaussian window

$$G(x, y, r_x, r_y) = \frac{1}{2\pi r_x r_y} \ e^{-x^2/2r_x^2} \cdot e^{-y^2/2r_y^2} \ , \tag{6}$$

where (r_x, r_y) are referred to as the effective radii of the elliptic support.

Variations in illumination intensity over the scene distort symmetry measurements based on grey levels. This phenomenon must be taken into consideration in face image analysis, since most face images are taken indoors, with great spatial variability in the illumination intensity. One novel aspect of this research is symmetry measurement of *color* images. By using the *rgb* (normalized *RGB*) color-space, the symmetry measured is that of a vector field of normalized color components. This compensates for spatial intensity variations.

The reflectional color symmetry measure of a color image $\mathbf{f}$ in a domain D is defined as

$$S(D) = \frac{||r||^2 S_r(D) + \||g||^2 S_g(D) + ||b||^2 S_b(D)}{||r||^2 + ||g||^2 + ||b||^2} \tag{7}$$

i.e., the weighted average of the 2-D symmetry values measured in the r, g and b normalized color components with respect to the energy in each component (within D).

Given the image $\mathbf{f}(i,j)$, the measure $S(D)$ of color symmetry in an elliptical domain D is a function of five parameters: the center coordinates of D, the effective radii corresponding to D and the orientation of D, i.e., the angle between its major axis and the x-axis. Observe that $S(D)$ is in itself scale-invariant, reflecting the fact that magnification has no effect on symmetry.

4 Objective Function

As defined, the measures of color compatibility $C(D)$ and color symmetry $S(D)$ within an elliptic domain D are scale invariant. Therefore, the symmetry and color compatibility associated with a tiny symmetric area of the right color will be higher than in a larger support, in which both symmetry are color compatibility are not as perfect. This is an undesirable state of affairs, since in any relevant image one may find many unimportant tiny symmetric regions of the right color, and at the limit each single pixel is perfectly symmetric and uniform in color. Regions of interest for image segmentation are usually much larger. Scale dependence should therefore be built into the objective function. The objective function is also the point of choice for imposing application-specific a-priori knowledge and preferences, possibly expressed via a function $P(D)$. Thus, the objective function takes the general form $F(D) = \mathcal{F}\{S(D), C(D), D, P(D)\}$. The objective function used in this research is of the form

$$F(D) = S^k(D) \cdot C^l(D) \cdot ||D|| \cdot P(D) \tag{8}$$

where k and l are positive integers. $P(D)$ was used, for example, to limit the ratio between the length of the major and minor axes in an elliptic approximation of a human face.

5 Global Optimization

Given an image, the objective function is a highly complex, multimodal function of five parameters: the coordinates of the centroid of the supporting region, its effective radii and its orientation. Maximizing the objective function is an elaborate global optimization problem. Solving this problem by exhaustive search is computationally prohibitive.

We implemented both a conventional genetic search algorithm [18] and a variation of the probabilistic genetic algorithm (PGA) described in [20]. Both algorithms perform quite well. The PGA tends to be faster than the standard genetic algorithm in the initial stages, but its final convergence seems to be slower. Typically, only about 3000 evaluations of the objective function are needed to reach the global maximum with either algorithm. Considering that 32 bits are used to describe each hypothesis (7 bits for each of the the location parameters x, y and 6 bits for each of the other three parameters), computing time is reduced by 6 orders of magnitude with respect to exhaustive search ($3 \cdot 10^3$ vs. 2^{32}). The images in Figs. 3-4 demonstrate the performance of the algorithm. Some of its limitations are shown in Fig. 5.

6 Discussion

Towards segmentation from multiple cues, this paper demonstrates the combined use of symmetry (global feature) and color (local property) for detecting regions of interest. Frontal-view face detection and the detection of man-made wooden objects have been used as working examples, but generality has been carefully maintained and the approach is not limited to those applications. One novel aspect of the suggested approach is the measurement of *chromatic* symmetry, thus compensating for illumination variations in the scene. Note that, unlike previous works, symmetry is analyzed *in conjunction* with skin detection rather than sequentially. The integration of the symmetry and color cues takes place in a unified objective function.

Great computational savings are obtained by avoiding exhaustive search. The global optimization method used can be extended [13, 17, 22] to detect more than one region of interest in the image. Further computational gains can be achieved by caching values of the objective function, thus avoiding unnecessary recomputation. An interesting direction for future research is the extraction of smoothness properties of the objective function. These could lead to global optimization with *guaranteed* convergence [25].

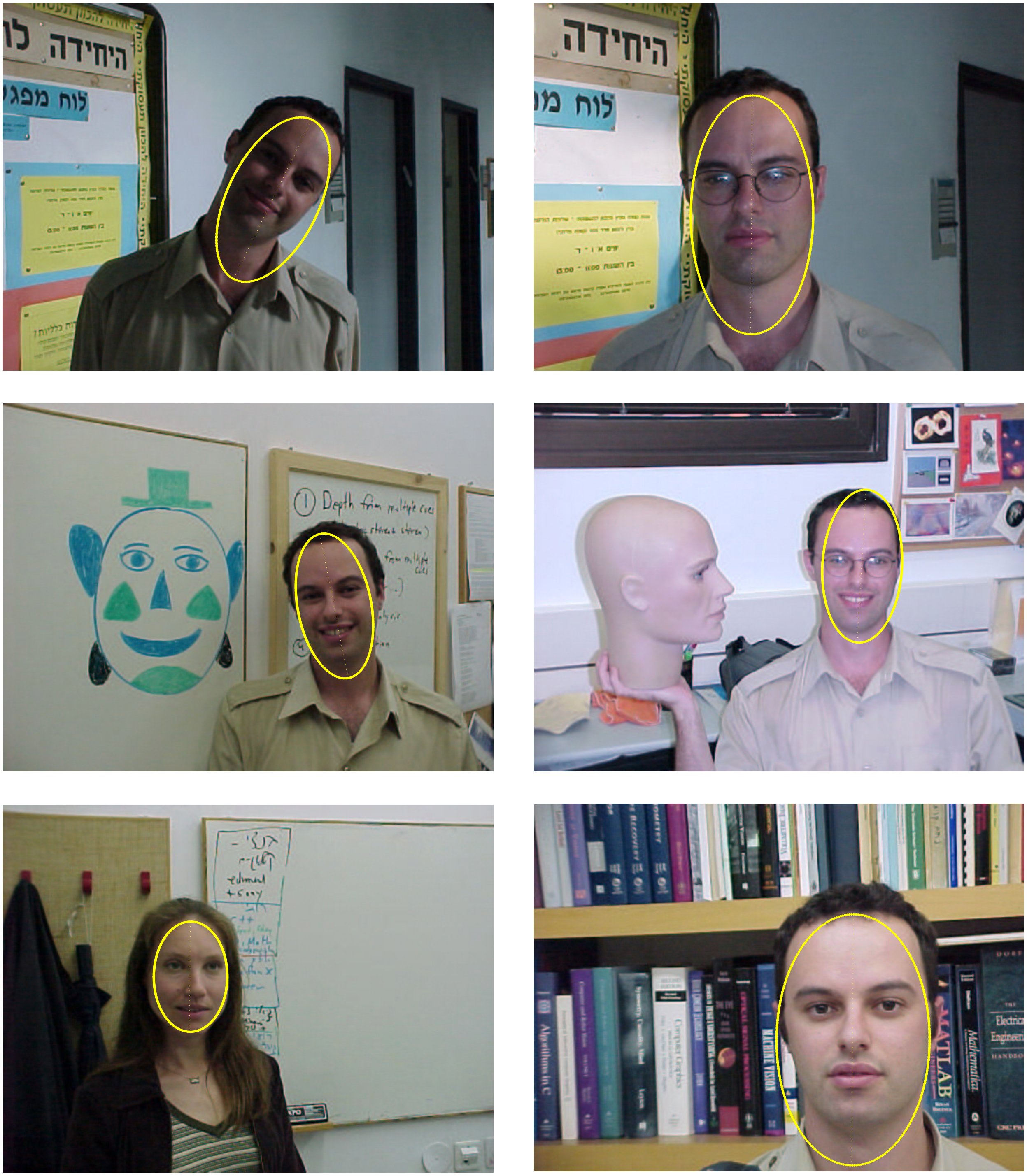

Figure 3: The suggested algorithm applied as a face detector. Note that, to maintain generality, the method relies only on symmetry and color-fitness: facial features are not used. *Top-left:* By measuring symmetry in the normalized color channels, large illumination variations can be accommodated. *Top-right:* Since facial features are not used, glasses pose no difficulty other than a slight reduction in color compatibility. *Middle-left:* Added value provided by the color cue. *Middle-right:* Added value brought by the symmetry cue. *Bottom-left:* Asymmetric background. *Bottom-right:* Cluttered background.

Figure 4: The suggested algorithm as a detector of man-made wooden objects.

References

[1] T.C. Chang and T.S. Huang, "Facial feature extraction from color images", *Proc. International Conference on Pattern Recognition*, Vol. 1, pp. 39-43, Jerusalem, 1994.

[2] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and machine recognition of faces: A survey", *IEEE Proceedings*, Vol. 83, pp. 705-740, 1995.

[3] A.J. Colmenarez and T.S. Huang, "Frontal view face detection", *SPIE* Vol. 2501, pp. 90-98, 1995.

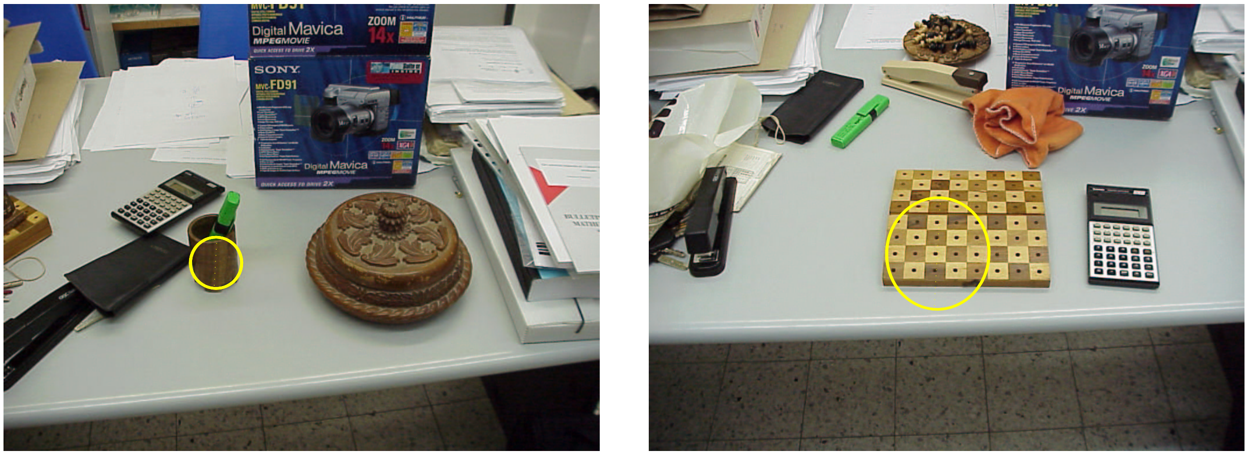

Figure 5: Limitations of the suggested algorithm. *Left:* Convergence to a local maximum. *Right:* Perspective projection leads to skewed symmetry. The chess board is an extreme case in which skewing turns symmetry to anti-symmetry. In this example, parts of the chessboard are sufficiently symmetric locally, but (the image of) the whole chessboard is not symmetric. Accommodating skewed symmetry in the method is straightforward, but would require higher dimensional search.

[4] T. Kondo and H. Yan, "Automatic human face detection and recognition under non-uniform illumination", *Pattern Recognition*, Vol. 32, pp. 1707-1718, 1999.

[5] Y. Ohta, T. Kanade and T. Sakai, "Color information for region segmentation", *Computer Graphics and Image Processing*, Vol. 13, pp. 222-241, 1980.

[6] T. Gevers and A.W.M. Smeulders, "Color-based object recognition", *Pattern Recognition*, Vol. 32, pp. 453-464, 1999.

[7] S.J. McKenna, S. Gong and Y. Raja, "Modelling facial colour and identity with gaussian mixtures", *Pattern Recognition*, Vol. 31, pp. 1883-1892, 1998.

[8] The Psychological Image Collection at Stirling (PICS), University of Stirling Psychology Department, `http://pics.psych.stir.ac.uk` .

[9] E. Saber and A.M. Tekalp, "Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost function", *Pattern Recognition Letters*, Vol. 19, pp. 669-680, 1998.

[10] A. Samal and P.A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey", *Pattern Recognition*, Vol. 25, pp. 65-77, 1992.

[11] J.C. Terrillon, M. David and S. Akamatsu, "Detection of human faces in complex scene images by use of a skin color model and of invariant fourier-mellin moments", *Proc. 14th International Conference on Pattern Recognition*, pp. 1350-1356, Brisbane, 1998.

[12] H. Wu, Q. Chen and M. Yachida, "Face detection from color images using a fuzzy pattern matching method", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, pp. 557-563, 1999.

[13] D. Beasley, D.R. Bull and R.R. Martin, "A sequential niche technique for multimodal function optimization", *Evolutionary Computation*, Vol. 1, pp. 101-125, 1993.

[14] J. Bigün, "Recognition of local symmetries in gray value images by harmonic functions", *Proc. International Conference on Pattern Recognition*, pp. 345-347, Rome, 1988.

[15] M. Gardner, *The New Ambidextrous Universe - Symmetry and Asymmetry from Mirror Reflections to Superstrings*, Freeman, New York, 1979.

[16] J.M. Gauch and S.M. Pizer, "The intensity axis of symmetry and its application to image segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, pp. 753-770, 1993.

[17] D.E. Goldberg and J. Richardson, "Genetic algorithms with sharing for multimodal function optimization", *Proc. 2nd Int. Conf. on Genetic Algorithms*, pp. 41-49, Cambridge, Mass., 1987.

[18] J.H. Holland, "Genetic algorithms", *Scientific American*, pp. 44–50, 1992.

[19] M.F. Kelly and M.D. Levine, "Annular symmetry operators: a method for locating and describing objects", *Proc. Int. Conf. on Computer Vision (ICCV)*, pp. 1016-1021, Cambridge, Mass., 1995.

[20] N. Kiryati and Y. Gofman, "Detecting symmetry in grey level images: the global optimization approach", *International Journal of Computer Vision*, Vol. 29, pp. 29-45, 1998.

[21] G. Marola, "On the detection of the axes of symmetry of symmetric and almost symmetric planar images", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, p.. 104-108, 1989.

[22] B.L. Miller and M.J. Shaw, "Genetic algorithms with dynamic niche sharing for multimodal function optimization", IlliGAL Report No. 95010, University of Illinois, department of general engineering, 1995. Available At *http://gal4.ge.uiuc.edu*.

[23] T.R. Reed and H. Wechsler, "Segmentation of textured images and gestalt organization using spatial/spatial-frequency representations", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, pp. 1-12, 1990.

[24] D. Reisfeld, H. Wolfson and Y. Yeshurun, "Context free attentional operators: the generalized symmetry transform", *Int. J. Computer Vision*, Vol. 14, pp. 119-130, 1995.

[25] A. Törn and A. Žilinskas, *Global Optimization*, Lecture Notes in Computer Science #350, Springer-Verlag, 1989.

[26] L. Van Gool, T. Moons, D. Ungureanu and E. Pauwels, "Symmetry from shape and shape from symmetry", *Int. J. Robotics Research*, Vol. 14, pp. 407-424, 1995.

[27] H. Weyl, *Symmetry*, Princeton University Press, 1952.

[28] A. Ylä-Jääski and F. Ade, "Grouping symmetrical structures for object segmentation and description", *Computer Vision and Image Understanding*, Vol. 63, pp. 399-417, 1996.

[29] H. Zabrodsky, S. Peleg and D. Avnir, "Symmetry as a continuous feature", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 17, pp. 1154-1166, 1995.

[30] J. Yang and A. Weibel, "Tracking human faces in real-time", *Technical Report CMU-CS-95-210*, Carnegie Mellon University, 1995.

[31] T. Zielke, M. Brauckmann and W. Von Seelen, "Intensity and edge based symmetry detection with application to car following", *CVGIP: Image Understanding*, Vol. 58, pp. 177-190, 1993.