A VARIATIONAL APPROACH TO GEOMETRIC-OPTICAL ILLUSIONS MODELING

Werner Ehm
Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany
<ehm@igpp.de>

Abstract

A straight line may appear slightly curved when placed in a suitable context consisting, for example, of an array of curves intersecting it. To date there seems to be no commonly accepted theory for such geometric-optical illusions. Starting from the assumption that only the local target-context interactions matter, we propose a purely geometric approach which predicts some of the observed curvatures via a variational principle.

Simple geometric elements such as a straight line, circle, or rectangle may appear distorted in the eye of the beholder when displayed within a suitable geometric pattern. The study of such geometric-optical illusions [GOI] dates back to the 1850s, at least, and has produced a vast literature, e.g. [1, 7, 10]. Nevertheless, a consensus about possible explanations nor even their underlying principles has not been achieved.

This contribution takes a phenomenological—rather than explanatory—approach [9] leaving aside neurophysiology and focusing at the geometry of the visual field in the first place. We consider a simple type of GOI that demonstrates illusory curvature by means of a straight line (the *target*) intersecting an array of other lines (the *context*): depending on the configuration of the context lines the target is perceived as bent or curved in some way rather than as straight ("Hering type" illusion [1]). Specifically, we are looking for a mathematical principle of a purely geometric nature that would allow us to predict the concrete shape of the illusory percept. The development relies on the assumption that context features distant from the target are largely irrelevant, and only the local target-context interactions should matter [9]. As for the latter we resort to an explanation of the Hering type illusions which surmises a misperception of angles sometimes referred to as *regression to right angles:* that acute angles are overestimated while obtuse angles are underestimated (e.g., [4, 5, 8]). Invoking this finding as a hypothesis we derive candidate illusory percepts as solution curves of a minimization problem reminiscent of Fermat's principle.

1. Context lines and vector fields

We focus on the case where the context consists of a family of planar curves that do not intersect each other. Such curves may be imagined as the contour lines of an undulated landscape. For definiteness, we will represent the context by means of a real-valued smooth function $(u, \theta) \mapsto c(u, \theta)$ defined on a rectangle $\mathcal{U} \times \Theta$ of $\mathbb{R}^2$, where θ is a parameter indexing the context line $u \mapsto C_\theta(u) := (u, c(u, \theta))$.

The context lines may be conceived as the stream lines of a planar flow given by a vector field v. Let S denote the set of all points $\xi = (\xi_1, \xi_2)$ such that $\xi = (u, c(u, \theta))$ for some $(u, \theta) \in \mathcal{U} \times \Theta$. The gradient of C_θ at any such ξ is given by the vector $(1, \partial_1 c(\xi_1, \theta))$, where ∂_j denotes the partial derivative w.r.t. the j-th argument. Since the context curves do not intersect, the gradient is a function of ξ alone, and $v(\xi) = (1, \partial_1 c(\xi_1, \theta))$ represents a vector field on the set S; that is, to any point $\xi \in S$ is attached a vector, $v(\xi)$, considered to be emanating from ξ, that indicates the "velocity" of the flow at the point ξ. Conversely, starting out from a vector field v on a domain $S \subset \mathbb{R}^2$ one can, under weak conditions, construct a family of non-intersecting curves whose gradients are given by v. This provides us with an equivalent representation of the context by means of a vector field instead of a family of context lines. In the following we shall work with both representations: the former is more convenient for theoretical developments; the latter is useful when it comes to examples.

The context forms a continuum in $\mathbb{R}^2$ if, as supposed here, the parameter set Θ is an interval. For graphical representations as in GOI illustrations it is necessary to select a finite sample out of the continuum of context curves and to exhibit this discrete subset only. The full set of context lines may then be seen as continuously interpolating the sample. This raises questions regarding the density (or sparsity) of the sample, and how a continuum model could possibly account for discrete sampling effects. We briefly comment on these issues in the discussion.

2. A variational principle for Hering type illusions

In the following, the context is represented by a smooth (continuously differentiable) vector field v on some planar domain S. For simplicity we assume $S = \mathbb{R}^2$. In view of the static, purely geometric character of the context it is natural to assume that $|v(\xi)| = 1$ for all $\xi \in \mathbb{R}^2$ ("constant velocity"). Here $|y| = \sqrt{\langle y, y\rangle}$ and $\langle y, z\rangle = y_1 z_1 + y_2 z_2$ denote the Euclidean norm (length) and inner product, respectively, of vectors $y, z \in \mathbb{R}^2$. An important geometric fact used below is that the "correlation" $\langle y, z\rangle / |y||z|$ equals the cosine of the angle between y and z; we denote this angle as $\angle(y, z)$. The target τ is assumed to be the straight line between the two points $x_0,\ x_1 \in \mathbb{R}^2$, parametrized as $\tau(t) = (1-t)x_0 + t x_1$ $(0 \leq t \leq 1)$.

The percept of the target τ in a Hering type illusion is not a straight line but appears slightly curved. The basic idea of our approach is to model the percept as a perturbation of τ that is characterized by a minimum principle. Setting up the principle involves three ingredients: (a) the local interactions hypothesis; (b) the regression to right angles hypothesis; (c) the fact that the straight line is the shortest path between two points. By (a), the context v is "recognized" only along candidate paths (in the vicinity of the target) while global aspects are disregarded. Observing (b) and (c) we then take the principle to posit that *given the context vector field v, the straight line target τ is distorted such that (i) the stream lines of v (the context lines) are intersected "as orthogonally as possible," and (ii) the distorted line is as short as possible.*

This could be stated mathematically as an optimization problem under side conditions. Since there is no obvious criterion suggesting length or orthogonality as the primary or the side condition, we propose to optimize a weigthed mixture of the two. Concretely, we consider the following *variational problem*: *minimize the functional*

$$x \mapsto \int_0^1 |\dot{x}(t)|\, dt + \alpha \int_0^1 \frac{\langle \dot{x}(t), v(x(t))\rangle^2}{|\dot{x}(t)|}\, dt \tag{1}$$

over the set $\mathcal{X}$ of all continuously differentiable planar curves $x(t)$, $t \in [0, 1]$, with given endpoints $x(0) = x_0$ and $x(1) = x_1$.

In this setting, $\mathcal{X}$ comprises the possible candidates for the actual percept. The first integral in (1), $\int_0^1 |\dot{x}(t)|\, dt = \int_0^1 |dx(t)|$, represents the length of x. (The superscript dot denotes the derivative w.r.t. the parameter t.) The integrand of the second integral is proportional to the square of the cosine of the angle subtended by v and the curve x at the point $x(t)$, thus measures the deviation from orthogonality along the curve. The division by $|\dot{x}(t)|$ is to make the whole expression homogeneous in $\dot{x}$ (and thus invariant under reparametrizations of t.) The number $\alpha \geq 0$, finally, accounts for the strength of the illusory effect. Note that when $\alpha = 0$ only length is being minimized, and the solution of the problem reduces to the straight line between x_0 and x_1, that is, to τ. Since the actual percept deviates only sligthly from the straight line target, one may anticipate that α should be small.

The above minimum principle is formally related to Fermat's principle, which characterizes the path of a light ray through an inhomogeneous medium. Indeed, on rewriting the functional (1) in the form $x \mapsto \int_0^1 F(x(t), \dot{x}(t))\, dt$ with integrand

$$F(x(t), \dot{x}(t)) = |\dot{x}(t)| + \alpha \langle \dot{x}(t), v(x(t)) \rangle^2 / |\dot{x}(t)| = |\dot{x}(t)| \left(1 + \alpha \frac{\langle \dot{x}(t), v(x(t)) \rangle^2}{|\dot{x}(t)|^2} \right), \qquad (2)$$

one sees that the variational problem amounts to minimizing the functional $x \mapsto \int_0^1 n(t)\, |dx(t)|$ where

$$n(t) = 1 + \alpha\, \langle \dot{x}(t), v(x(t)) \rangle^2 / |\dot{x}(t)|^2 = 1 + \alpha \cos^2 \angle(\dot{x}(t), v(x(t)))$$

is the "refraction index"—which in our case depends not only on the "medium" (here: the context) as traversed by the path, via $v(x(t))$, but also on the gradients of the path, $\dot{x}(t)$.

3. Approximate solution of the variational problem for small α

We apply the machinery of the calculus of variations [2]. A solution curve $x \in \mathcal{X}$ of the minimization problem necessarily satisfies the *Euler equation*

$$\frac{d}{dt} \nabla_{\dot{x}} F(x(t), \dot{x}(t)) - \nabla_x F(x(t), \dot{x}(t)) = 0 \qquad \text{(for all } t\text{)}. \qquad (3)$$

Here $\nabla_x F$, $\nabla_{\dot{x}} F$ denote the partial gradients of F w.r.t. the (vector) arguments x, $\dot{x}$, respectively. Let $v'(\xi)$ denote the 2 by 2 matrix of partial derivatives of v at the point $\xi \in \mathbb{R}^2$ ("Jacobian"), and $v'(\xi)^t$ its transpose. In our special case the Euler equation becomes

$$\big(1 - \alpha \langle \rho, v(x) \rangle^2\big)\, \dot{\rho} = -2\alpha \Big[[v(x) - \langle \rho, v(x) \rangle \rho]\, \frac{d}{dt} \langle \rho, v(x) \rangle + \langle \rho, v(x) \rangle\, \big(v'(x) - v'(x)^t\big)\, \dot{x} \Big], \quad (4)$$

where for compactness of notation we omitted the argument t and wrote $\rho = \dot{x}/|\dot{x}|$.

Initially, (4) is a system of two nonlinear, 2nd-order differential equations. However, because both sides of (4) can be shown to be throughout orthogonal to the direction ρ of the curve x, the tangential component of (4) vanishes along x and only the component orthogonal to it matters. For each t, let $\rho^\perp = \rho^\perp(t)$ denote the unit vector making $\{\rho, \rho^\perp\}$ a positively oriented orthonormal basis. Taking inner products with $\rho^\perp$ and assuming $\alpha < 1$ one finds after some calculation that the nontrivial component of the Euler equation is given by

$$\langle \dot{\rho}, \rho^\perp \rangle = -2\alpha |\dot{x}|\, \frac{\langle v(x), \rho^\perp \rangle \langle v'(x)\rho, \rho \rangle + \langle v(x), \rho \rangle\, (\partial_1 v_2(x) - \partial_2 v_1(x))}{1 - \alpha\, \langle v(x), \rho \rangle^2 + 2\alpha\, \langle v(x), \rho^\perp \rangle^2}\,. \qquad (5)$$

The sign of the left-hand side of (5) tells us whether the curve x running from x_0 to x_1 turns to the left (sign $= 1$) or to the right (sign $= -1$). Since for $\alpha < 1$ the denominator of the fraction at the right-hand side is positive, this qualitative information can be read off from the numerator.

Explicit solutions to (5) are generally not available. However, when the perceived curvature is weak, which is the case of primary interest, approximate solutions can fortunately be obtained. In this case α will be small, and the (exact) solution x_α can be assumed to be close to the straight line τ. This suggests to make an ansatz $x_\alpha \doteq \tau + \alpha\gamma$ with γ representing an approximation to the (rescaled) deviation of x_α from τ. Then

$$\dot{x}_\alpha \doteq \ell \rho_0 + \alpha \dot{\gamma}, \qquad \ddot{x}_\alpha \doteq \alpha \ddot{\gamma},$$

where $\ell = |x_1 - x_0|$ denotes the length of τ and $\rho_0 = (x_1 - x_0)/\ell$ its direction. As α tends to 0, the left-hand side of (5), $\langle \dot\rho_\alpha, \rho_\alpha^\perp \rangle$, can be approximated by $\ell^{-1}\alpha \langle \ddot\gamma, \rho_0^\perp \rangle$ up to terms of a smaller order than α. Thus after division by α/ℓ the left-hand side of (5) tends to $\langle \ddot\gamma, \rho_0^\perp \rangle$, while the right-hand side approaches the limit

$$-2\ell^2\mathfrak{C}, \qquad \mathfrak{C} = \langle w, \rho_0^\perp \rangle \langle w' \rho_0, \rho_0 \rangle + \langle w, \rho_0 \rangle \left(\partial_1 w_2 - \partial_2 w_1\right), \tag{6}$$

wherein

$$w = w(t) = v(\tau(t)), \quad w' = w'(t) = v'(\tau(t))$$

stand for the vector field and its Jacobian, respectively, evaluated at the target line τ. Thus in the limit $\alpha \to 0$ one gets

$$\langle \ddot\gamma, \rho_0^\perp \rangle \; = \; -2\ell^2\mathfrak{C}. \tag{7}$$

The expression on the right-hand side of this equation depends only on τ and v, hence it is known. Therefore, integrating (7) twice gives $\langle \gamma, \rho_0^\perp \rangle$ up to an affine function that can be determined from the boundary conditions for γ, which is a loop starting and ending at the origin. This finally yields our approximation to the lateral deviation of x_α from τ, namely the expression $\alpha\langle \gamma, \rho_0^\perp \rangle$.

To summarize: An approximate solution of the Euler equations for small $\alpha > 0$ is given by the curve

$$t \mapsto \tau(t) + \alpha\langle \gamma(t), \rho_0^\perp \rangle \rho_0^\perp, \tag{8}$$

where the second term is obtained as described above. This curve represents our guess for the percept of the target line. It is uniquely determined by the context and the target except for the parameter α, which has to be estimated experimentally.

4. Examples

Let us return to the initial representation of the context discussed in Section 1. Suppose we are given a real function $c(u, \theta)$ such that c is strictly increasing in θ for each fixed u. The context curves $u \mapsto C_\theta(u) = (u, c(u, \theta))$ then do not intersect for different parameters θ, and we may assume that for each $\xi = (\xi_1, \xi_2)$ there exists $\theta = \vartheta(\xi_1, \xi_2)$ such that $c(\xi_1, \vartheta(\xi_1, \xi_2)) = \xi_2$. The curves C_θ generate a vector field v of unit directions given by

$$v(u, c(u, \theta)) = \frac{(1, \partial_1 c(u, \theta))}{\sqrt{1 + (\partial_1 c(u, \theta))^2}}$$

which in terms of the parameters $\xi_1 = u$ and ξ_2 can be written as

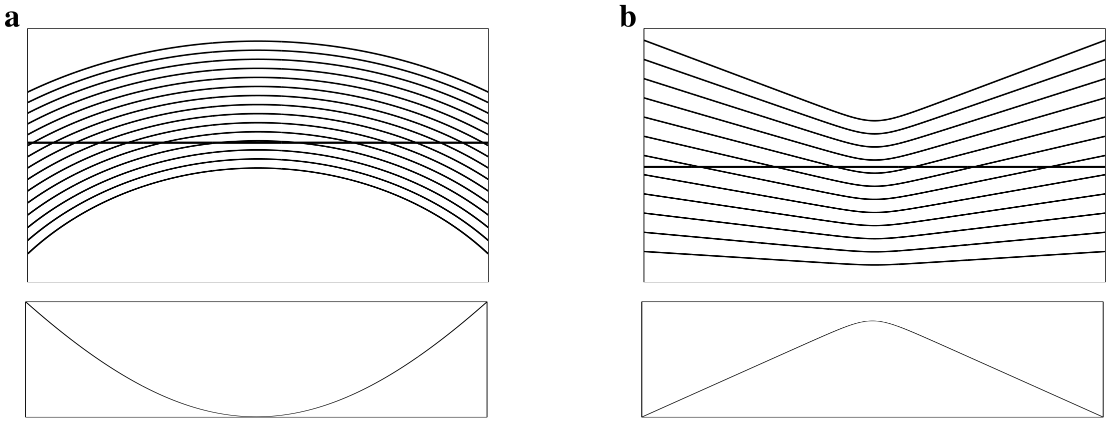
$$v(\xi) = v(\xi_1, c(\xi_1, \vartheta(\xi_1, \xi_2))) = \Gamma^{-1/2}(1, \partial_1 c), \qquad \text{with} \qquad \Gamma = 1 + (\partial_1 c)^2\,. \tag{9}$$

By the invariance of the variational problem under rotations and translations we may assume without loss of generality that the target τ is the horizontal line segment between $x_0 = (-\ell/2, 0)$ and $x_1 = (\ell/2, 0)$ $(\ell > 0)$. It is then convenient to parametrize all curves by the variable $\xi_1 = u = \ell(t - 1/2)$. Within this setting, one can calculate the crucial quantity $\mathfrak{C}$ from (6), (7) explicitly in terms of partial derivatives of c. The result is

$$\mathfrak{C} = K/\Gamma^2, \qquad K = \left(1 - (\partial_1 c)^2\right)\partial_{11}^2 c \, + \, (\partial_1 c)^2 \, \frac{\partial_1 c}{\partial_2 c}\, \partial_{12}^2 c\,, \tag{10}$$

where both K and Γ (cf. (9)) are evaluated at the arguments $(\xi_1, \vartheta(\xi_1, 0))$. (Note that one has $c(\xi_1, \vartheta(\xi_1, 0)) = 0$ by the definition of ϑ, so the parameter θ for which the curve C_θ crosses the target at the point $(\xi_1, 0)$ is $\theta = \vartheta(\xi_1, 0)$.)

Figure 1 presents three examples. In each case, the curves C_θ are obtained from a single given function by translation (panel **a**) or by dilation (panels **b**, **c**), respectively, these operations being parametrized by θ.



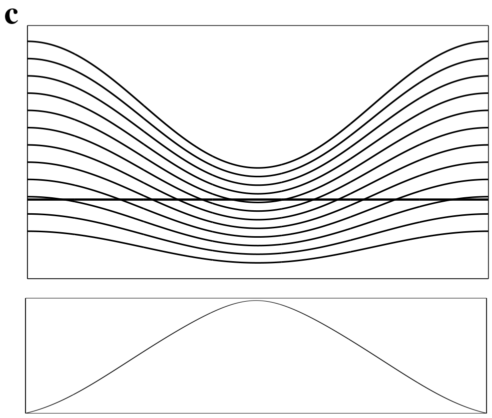


Figure 1.
Straight line target embedded in 3 different contexts (upper part of each panel).
Lower part: the curve $\langle \gamma, \rho_0^\perp \rangle$ constructed in Section 3 (lateral deviation of the percept from the target).
Note: This gives only the (putative) shape of the distortion, up to some factor reflecting the strength of the illusory effect.

5. Discussion

In this contribution we have used the calculus of variations for modeling illusory percepts in GOIs. Of course, such an approach cannot explain the origin of the perceptual phenomenon. Even as a purely descriptive tool it has a number of limitations and needs further development.

As noted earlier, a continuous vector field can only serve as an approximation to a context of finitely many lines. That may suffice with regard to the shape of the illusory percept, but is blind to the possible impact of the number (or density) of context lines on the magnitude of the perceptual distortion. In the extreme of a very sparse context as in the Poggendorff illusion, one would expect an essentially symmetric context-target interaction, wheras a massive context should remain unaffected by the target. But still, although the present approach does not yield any prediction regarding the relation between context density and effect size, it allows to account for it by means of the parameter α. Experimentally, α can be found, and the said relation be studied, by compensating for the illusory effect [8]: one presents the target τ *minus* α times the lateral deviation (cf. (8)) and lets the subject adjust α such that she perceives a straight line. An experiment along these lines is on the way. As an additional step, one also could discretize the criterion (1) by passing from the integral to the Riemann sums corresponding to the intersection of finitely many context lines with the target.

Another limitation of the present approach comes from the boundary conditions of the minimization problem which pin down the endpoints of all candidate percepts at those of the target. This precludes allowing for an illusory tilt as it occurs in the Zöllner illusion. We currently investigate variants with free endpoints complemented by additional criteria enforcing closeness of the candidate percepts to the target. Ultimately, the goal is to find a functional on a set of candidate percepts the minima of which provide “good” predictors of the actually perceived targets *under broad conditions.* Indeed, considering the sheer difficulty of a subjective discrimination of similar percepts, it will be important to work with a multitude of differently shaped targets exhibiting less symmetry than the commonly studied GOI examples when predictive goodness is to be assessed. An encouraging observation telling us that we may not be too far from a “geometry of the visual field” [9, 10] is the following purely geometrical interpretation of the variational principle (1): the approximate solutions for small α are exact geodesics in a suitable Riemannian geometry [3]. Details will be given elsewhere.

Finally, it has to be pointed out that vector fields permeate much of the literature on GOIs, with substantially varying usage and meaning. Also, variational principles are widely used in computer vision. Two examples close in spirit to the present work are the search for a “generating” vector field explaining the length distortion in a T-shaped figure [6] and the modeling of illusory contours in a Kanisza triangle [11]. However, despite apparent similarities there are important differences in type of criteria as well as objective.

Acknowledgments

The author is indebted to Steffen Heinze and Jiří Wackermann for inspiring discussions and valuable comments on earlier versions of the manuscript.

References

[1] S Coren, JS Girgus (1978). *Seeing is deceiving: The psychology of visual illusions.* Hillsdale, NJ: Lawrence Erlbaum Assoc.

[2] B Dacorogna (2004). *Introduction to the calculus of variations.* London: Imperial College Press

[3] S Heinze (2010). *Private communication*

[4] H Helmholtz (1867). *Handbuch der physiologischen Optik.* Leipzig: Voss

[5] RI Horrell (1971). The angle of intersection of contours as the determinant of a geometric illusion. *Percept. Psychophys.* **10**, 208-210

[6] G Lehmann (1967). Die Analyse geometrisch-optischer Täuschungen durch Vektorfelder. *Z. Exp. Angew. Psychol.* **14**, 442-462

[7] JO Robinson (1998). *The psychology of visual illusions.* New York: Dover

[8] DA Smith (1978). A descriptive model for perception of optical illusions. *J. Math. Psychol.* **17**, 64-85

[9] J Wackermann (2010). Geometric-optical illusions: A pedestrian’s view of the phenomenal landscape. In: A Bastianelli, G Vidotto (Eds.), *Fechner Day 2010.* Padua: International Society for Psychophysics, pp. 171-176

[10] G Westheimer (2008). Illusions in the spatial sense of the eye: Geometric-optical illusions and the neural representation of space. *Vision Res.* **48**, 2128-2142

[11] W Zhu, T Chan (2007). A variational model for capturing illusory contours using curvature. *J. Math. Imag. Vis.* **27**, 29-40