Scene from
"Lola rennt"

(Optically guided) sensorimotor action is ***automatic*** (Lola's mind is otherwise occupied!). In this respect Lola is a *zombie*.

This is *physiology* & space-time is generated by the Galilean transformation group of classical kinematics.

No problem! Science will handle this fine!

Umberto Boccioni
"La strada entra nella casa" (1911).

This is *Visual Space,* which is a *mental entity*.

It has to do with *conscious* perception.

The *exact sciences* deal with *zombies*.

This is very important and progress has been spectacular!

Not so with *experience*, that is *conscious perception*. The exact sciences have *nothing* to say about this.

A "science" of *conscousness* is forever out of reach!

Phenomenology so far hasn't recovered from behaviourism.

Science has to be silent. One needs novel headways …

What is “SHAPE”?

Definition:
Two spatial configurations have *the same shape* if they are equivalent *modulo an “irrelevance” transformation*.

Spatial configurations: You ***know*** (a mental thing!). Think of “pictorial objects”*, e.g., La Gioconda* in 3D.

Irrelevance transformations: Again, you ***know*** (mental things again!). Apparently they are *similarities & congruences (movements),* but here we’re talking *mental* movements!

"Default Visual Space"

Consider a *stationary*, *monocular* observer. Don't assume any prior knowledge.

Since no point or direction of visual space is "special", visual space has to be *homogeneous* and *isotropic*.

Since no distance or direction of physical space is considered "special", the structure of visual space is *invariant against arbitrary rotation–dilations about the vantage point.*

One concludes that the space is Riemannian with metric (Cartesian coordinates $\{x, y, z\}$, polar $\{\varrho, \vartheta, \varphi\}$)

$$
\begin{aligned}
\mathrm{d}s^2 &= \frac{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}{x^2 + y^2 + z^2} = \frac{\mathrm{d}\varrho^2 + \varrho^2 \mathrm{d}\vartheta^2 + \varrho^2 \sin^2 \vartheta \mathrm{d}\,\varphi^2}{\varrho^2} = \\
&= \mathrm{d}(\log \frac{\varrho}{\varrho_0})^2 + \mathrm{d}\vartheta^2 + \sin^2 \vartheta\, \mathrm{d}\varphi^2.
\end{aligned}
$$

Limited Field of View

For a limited field of view (*of indefinite extent!*) centered on the direction $\vartheta = 0$ one introduces **Riemann normal coordinates**

$$\{u, v\} = \vartheta \, \{\cos\varphi, \sin\varphi\},$$

and the metric reduces to

$$\mathrm{d}s^2 = \mathrm{d}u^2 + \mathrm{d}v^2 + \mathrm{d}w^2,$$

where $w = \log \frac{z}{z_0}$.

However, this cannot be a *Euclidean* space because the uv and the w–dimensions are *incommensurable* and are not permitted to “mix”, thus Euclidean rotations are vetoed!

Graph Spaces

Given the Euclidean plane $\mathbb{E}^2$ as "base space", consider the trivial fiber bundle $\mathbb{E}^2 \times \mathbb{A}^1$ whose fibers are copies of the affine line $\mathbb{A}^1$. Its cross–sections are called "graphs", they can be visualized as surfaces with only "frontal parts", *i.e.*, "reliefs".

The bundle $\mathbb{E}^2$ can be treated as a Cayley–Klein homogeneous geometry if one considers the fibers to be "isotropic": Then the group of movements does not allow the fibers and the base space dimensions to "mix". The metric may be taken as

$$\mathrm{d}s^2 = \mathrm{d}u^2 + \mathrm{d}v^2 + \varepsilon^2\,\mathrm{d}w^2,$$

where ε is a non–trivial solution of the quadratic equation $x^2 = 0$, a "nil–squared number".

A simple subspace

A simple subspace is $\mathbb{E}^1 \times \mathbb{A}^1$ (a "line" of the visual field, "depth section"). It is conveniently modeled by the *dual number plane.*

A "dual number" is a number like $z = x+\varepsilon y$, where x, $y \in \mathbb{R}$ and the "imaginary unit" ε. (Although $\varepsilon \neq 0$ neither $\varepsilon > 0$, nor $\varepsilon < 0$! The *logic* in intuistionistic! Moreover $\varepsilon \notin \mathbb{R}$.)

Intuitively, dual numbers are "infinitesimal environments" of the real numbers.

Notice that the linear transformation $z' = az + b$ ($a = a_1 + \varepsilon a_2$, $b = b_1 + \varepsilon b_2$, $z = z_1 + \varepsilon z_2$ dual numbers) is

$$z' = (a_1 z_1 + b_1) + \varepsilon(a_2 z_1 + a_1 z_2 + b_2),$$

thus the imaginary parts don't "mix into" the real part!

The group of similarities

Writing dual numbers as $x + \varepsilon y$, the group of similarities is

$$\begin{aligned} x' &= \sigma x + \tau, \\ y' &= \rho x + \gamma y + \delta, \end{aligned}$$

where:

σ is a *scaling in the visual field*,
τ is a *translation in the visual field*,
ρ is an "*isotropic rotation*",
γ is a *scaling in depth*,
δ is a *shift in depth*.

For $\sigma = \gamma = 1$ you have *proper movements*, For $\sigma =$, $\tau = 0$ the visual field is invariant.

Length and angle metrics

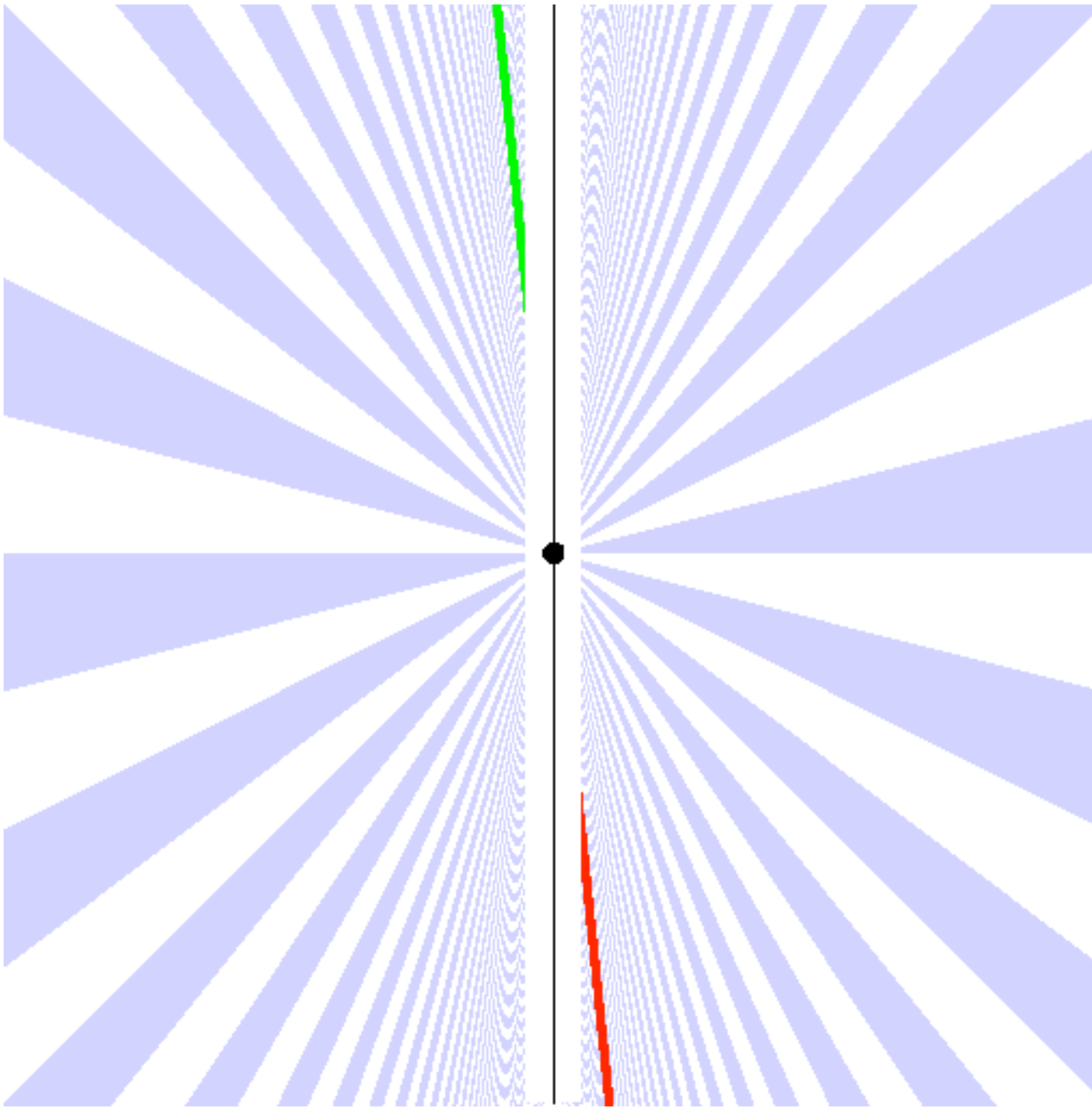
The real part of $(x + \varepsilon y) - (u + \varepsilon v) = x - u$ is invariant under proper movements, hence one defines $|x + \varepsilon y| = x$. In case $|z - w| = 0$ it usually happens that $z \neq w$: such points are "parallel". Parallel points differ by an imaginary quantity that is conserved under proper motions.

For $z = x + \varepsilon y$, $w = u + \varepsilon v$ the ratio $\frac{v-y}{u-x}$ is invariant under linear transformations. It is interpreted as the *angular difference* of z and w.

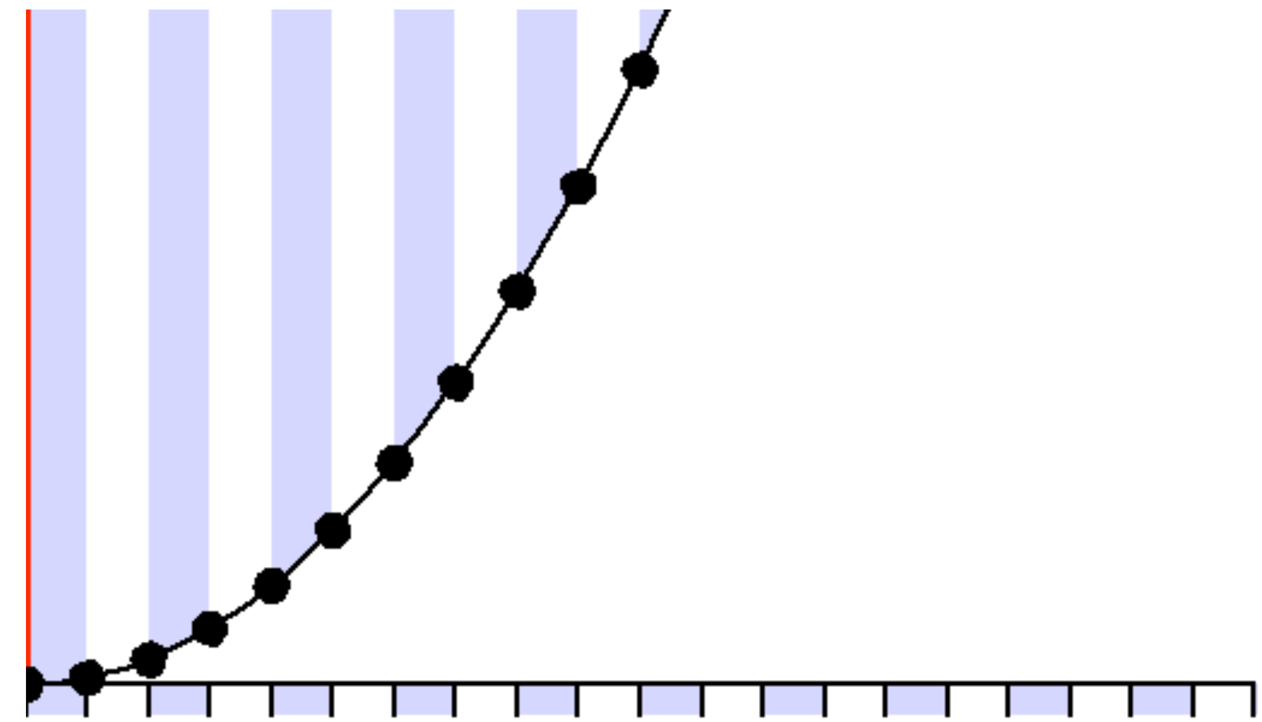
One writes $z = x + \varepsilon y = x\mathrm{e}^{\varepsilon \frac{y}{x}}$, or $|z| = x$, $\varepsilon \angle z = \arctan \varepsilon \frac{y}{x}$ (thus $\angle z = \frac{y}{x}$). This leads to

$$\begin{aligned} z + w &= (x + u) + \varepsilon(y + v), \\ z \,.\, w &= xy\,\mathrm{e}^{\varepsilon(\angle z + \angle w)}. \end{aligned}$$

Both distance and angle metrics are *parabolic*.



A "rolling wheel" looks like a parabola with isotropic axis ("hub" at infinity!). A "unit circle of the 2nd kind of the dual plane is $u + \frac{1}{2}\varepsilon\, u^2$.



The angle metric is (like the distance metric!) *parabolic*. Angles are *not periodic*, but range from $-\infty$ to ∞. The protractor at left shows a direction (red line) that rotates at *uniform speed*. There is nothing like a *"pirouette"* in this space!

Relief space $\mathbb{E}^2 \times \mathbb{A}^1$

The group of visual field conserving similarities of relief space is

$$\begin{aligned} x' &= x, \\ y' &= y, \\ z' &= \alpha x + \beta y + \gamma z + \delta. \end{aligned}$$

The depth shift δ is always irrelevant.

Thus you have an isotropic rotation $\{\alpha, \beta\}$ ("additive plane") and a depth scaling γ ("relief scaling"). These transformations are familiar from the "bas–relief ambiguity" of SFS. The relief scaling was identified by the German sculptor Hildebrand in the 1890's.

These transformations are part of the ambiguity group of any "purely monocular depth cue". The are due to the fact that rotations about and dilations centered on the vantage point leave the optical structure at the eye invariant.

PROBLEM OF FORM

IN

PAINTING AND SCULPTURE

BY

ADOLF HILDEBRAND

TRANSLATED AND REVISED WITH THE AUTHOR'S CO-OPERATION

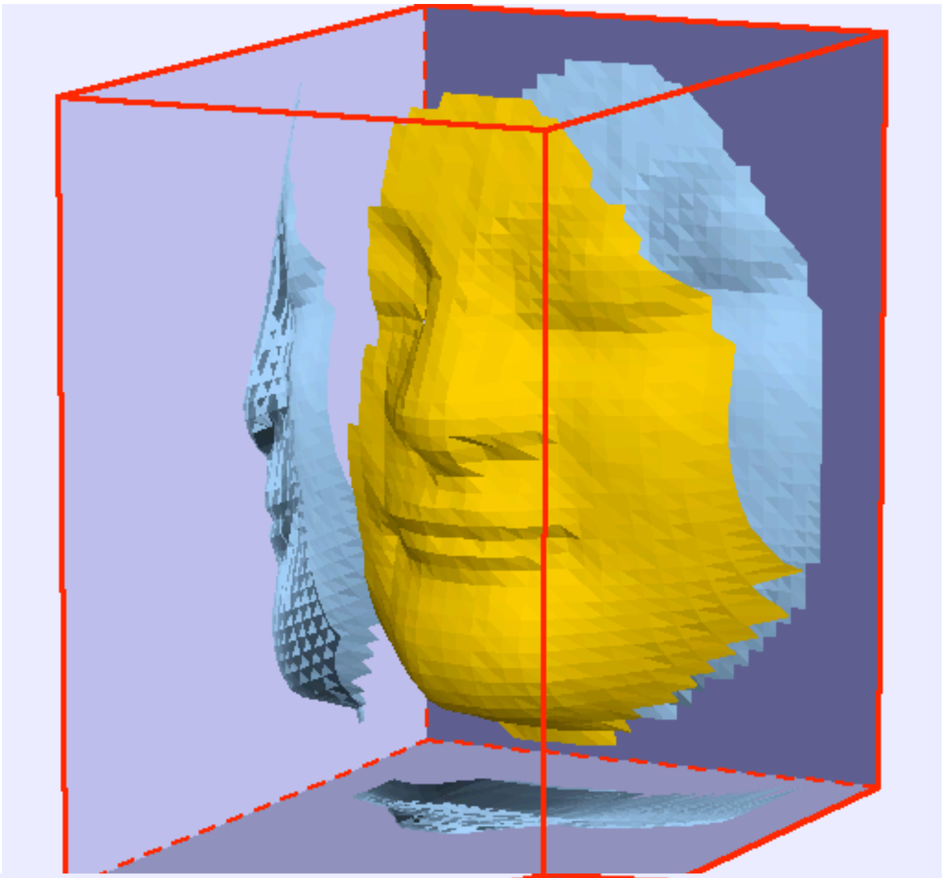
BY

MAX MEYER AND ROBERT MORRIS OGDEN
(*University of Missouri*) (*University of Tennessee*)

First edition 1893

WITH THIRTY ILLUSTRATIONS AND A PORTRAIT OF THE AUTHOR

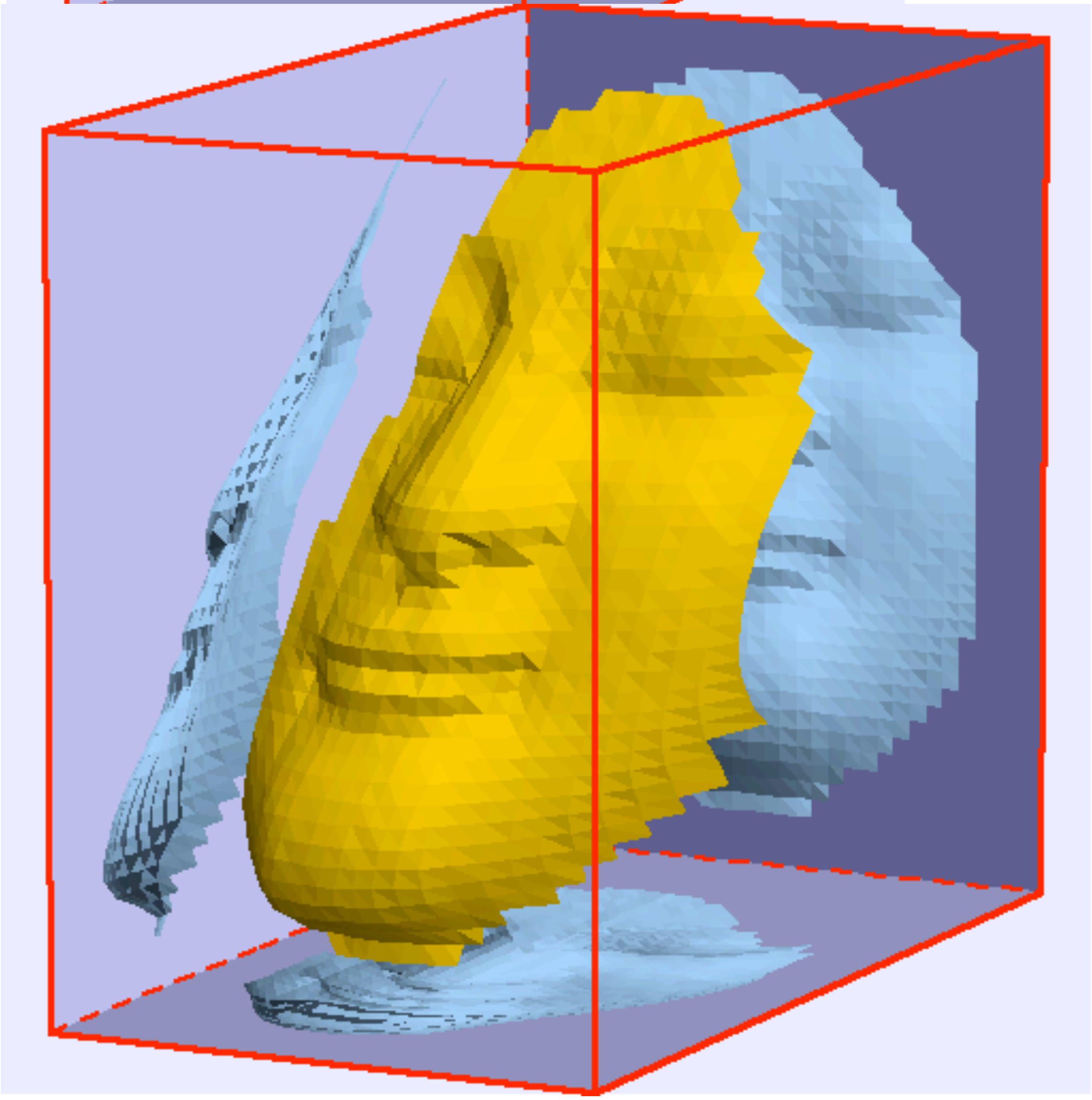
G. E. STECHERT & CO.
NEW YORK
1945

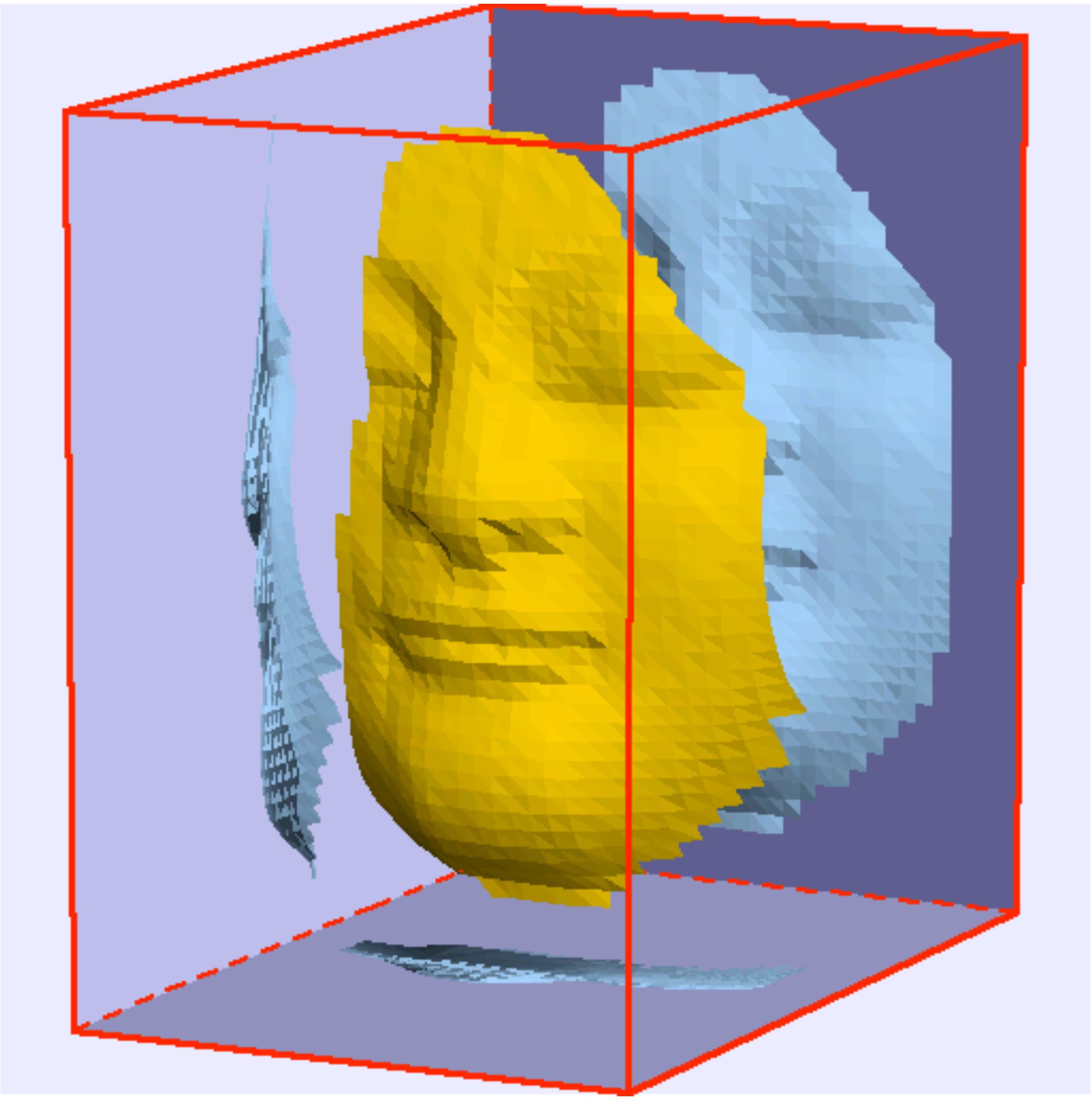


The "Bas–relief Ambiguity":

$$z' = \omega_x x + \omega_y y + \sigma z + \tau_z,$$

for rotation $\boldsymbol{\omega}$, scaling σ and z–translation τ_z.





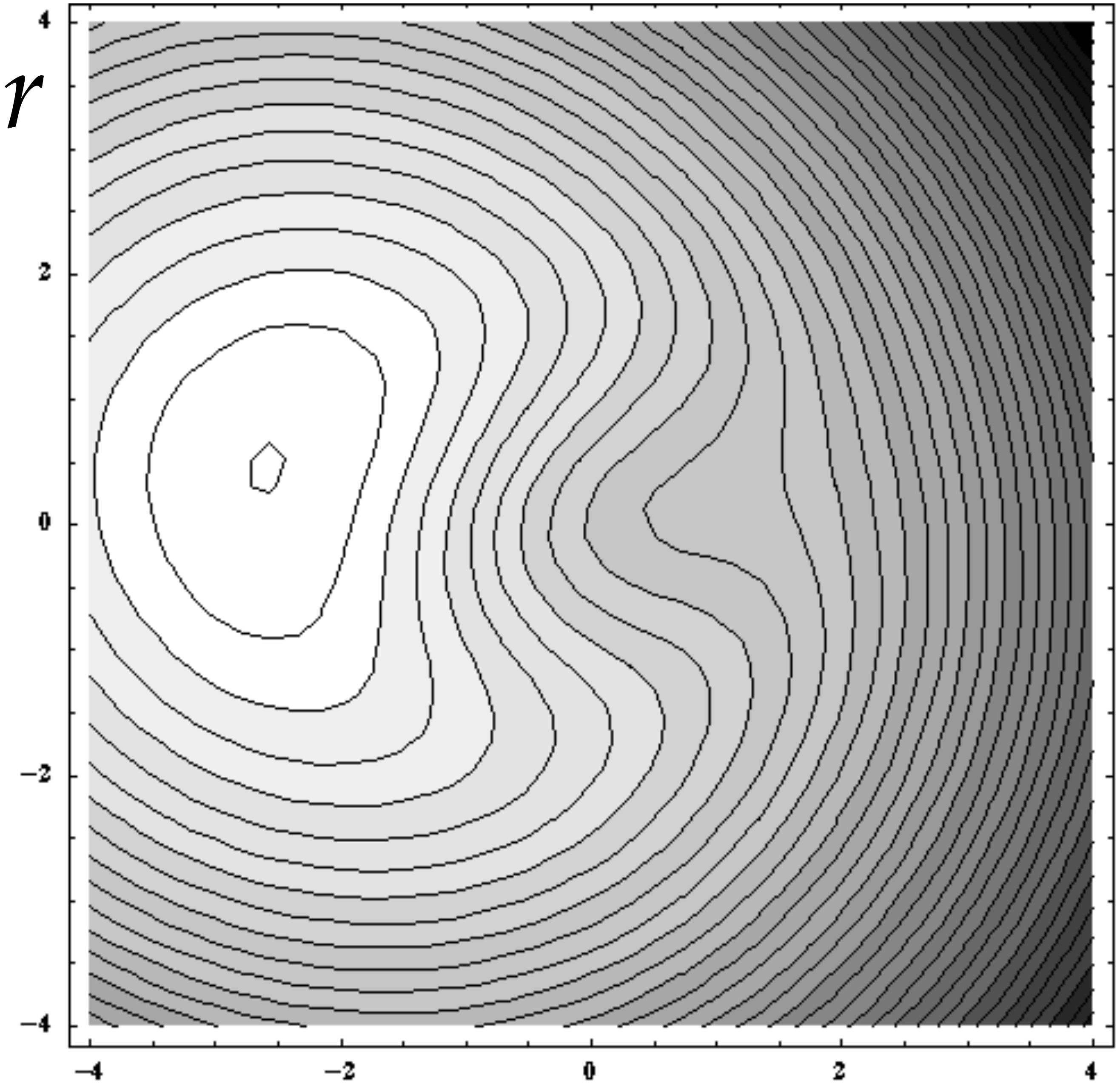
Differential Geometry in Relief Space

The normals of a relief are all parallel because isotropic! Thus one studies surfaces via tangent plane variations. A surface $\{u, v, \varepsilon w(u, v)\}$ has tangent planes spanned by $\{1, 0, \varepsilon\, w_u\}$ and $\{0, 1, \varepsilon\, w_v\}$. These may be mapped by parallelity on the surface $\{u, v, \frac{1}{2}\varepsilon\,(u^2 + v^2)\}$: the *"spherical image" of the surface*. Projected upon the base space (an isometric projection!) one has the map

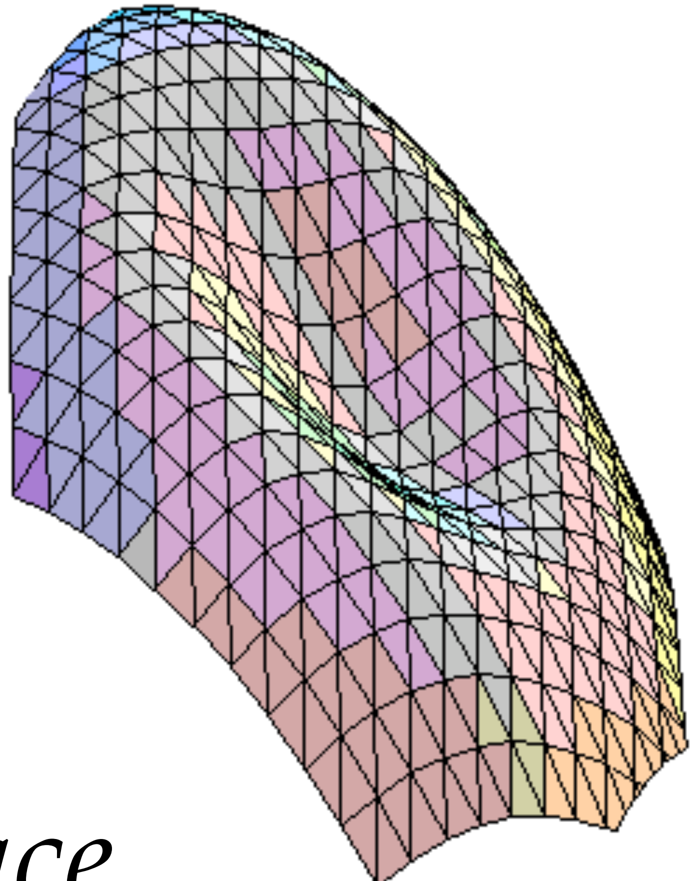
$$\{u, v, \varepsilon\, w(u, v)\} \mapsto \{w_u(u, v), w_v(u, v)\},$$

which may double (because isometric) to the spherical image. (It is nothing but the familiar "gradient space" of machine vision!)

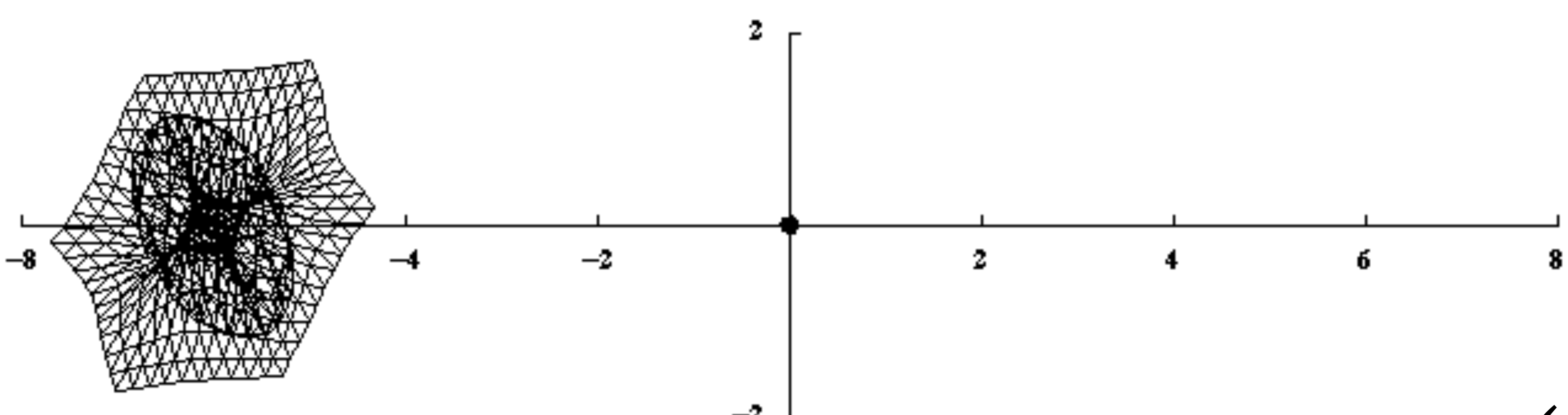
The spherical image is invariant under translations, translates under rotations, and uniformly scales under similarities.



contour plot



surface



spherical image (“gradient space”)

“Local Shape” (curvature, *etc.*)

Locally one approximates relief by a Taylor expansion

$$w(u,v) = \sum_{k=0}^{\infty}\left[\frac{1}{k!}\sum_{i=0}^{k} a_{k-i,\,i}\binom{k}{i}u^{k-i}v^{i}\right].$$

The 0^{th} order can be annulled by a depth shift, the 1^{st} order by a rotation. A rotation about the isotropic direction brings the 2^{nd} order to canonical form, a similarity even settles the magnitude of the curvature. The remaining terms define the *pure local shape*:

$$w(u,v) = \frac{C}{\sqrt{2}}\left(\cos(S+\frac{\pi}{2})u^2 + \cos(S+\frac{\pi}{2})v^2\right) + \ldots$$

C the “Casorati curvature”, S the “shape index”.

“Local Shape” — II

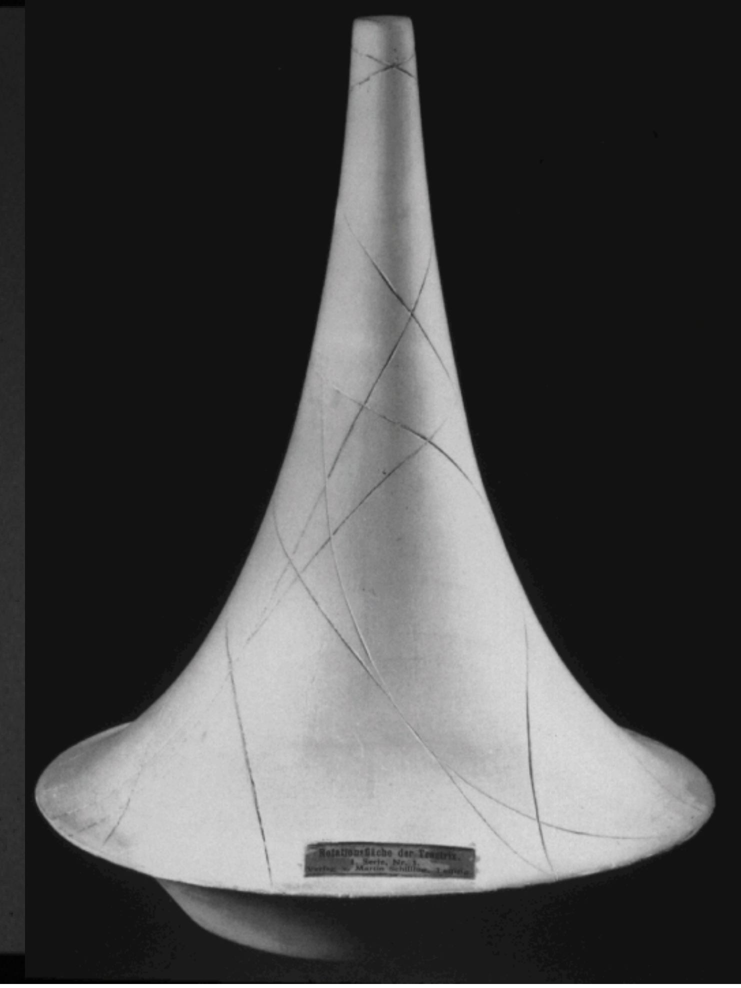
The “Casorati curvature” is $C = \sqrt{\frac{1}{2}(w_{xx}^2 + 2w_{xy}^2 + w_{yy}^2)}$, which equals $\sqrt{\frac{1}{2}(\kappa_{\max}^2 + \kappa_{\min}^2)}$. (Where $\kappa_{\max}$, $\kappa_{\min}$ are the principal curvatures.)

The “shape index” is

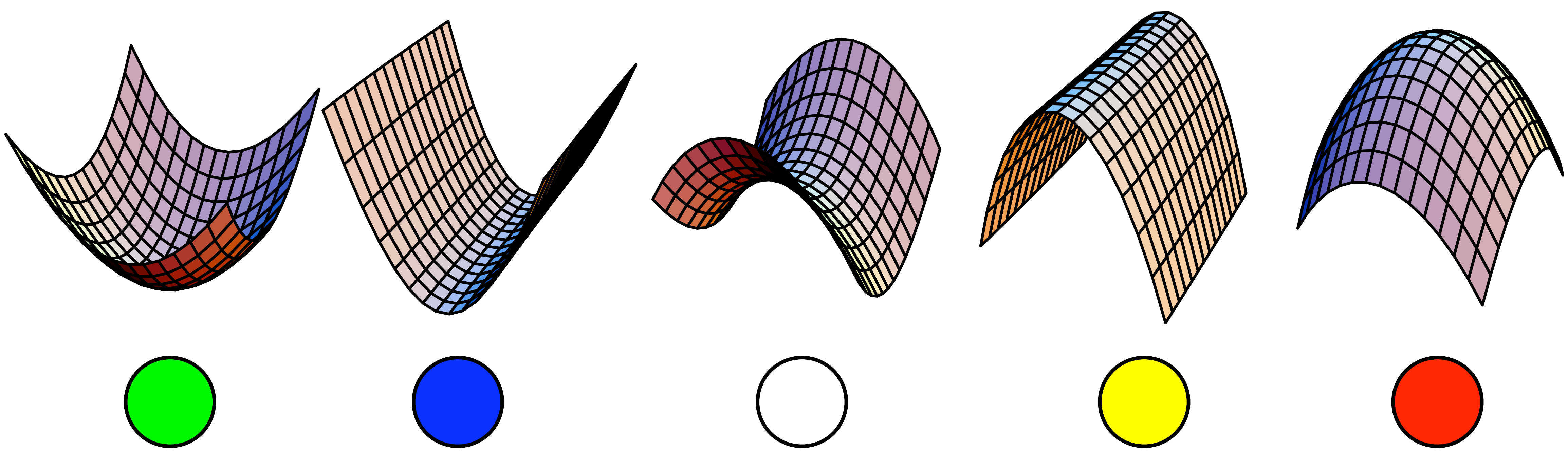
$$S = \arctan\left(\frac{\kappa_{\min} + \kappa_{\max}}{\kappa_{\min} - \kappa_{\max}}\right),$$

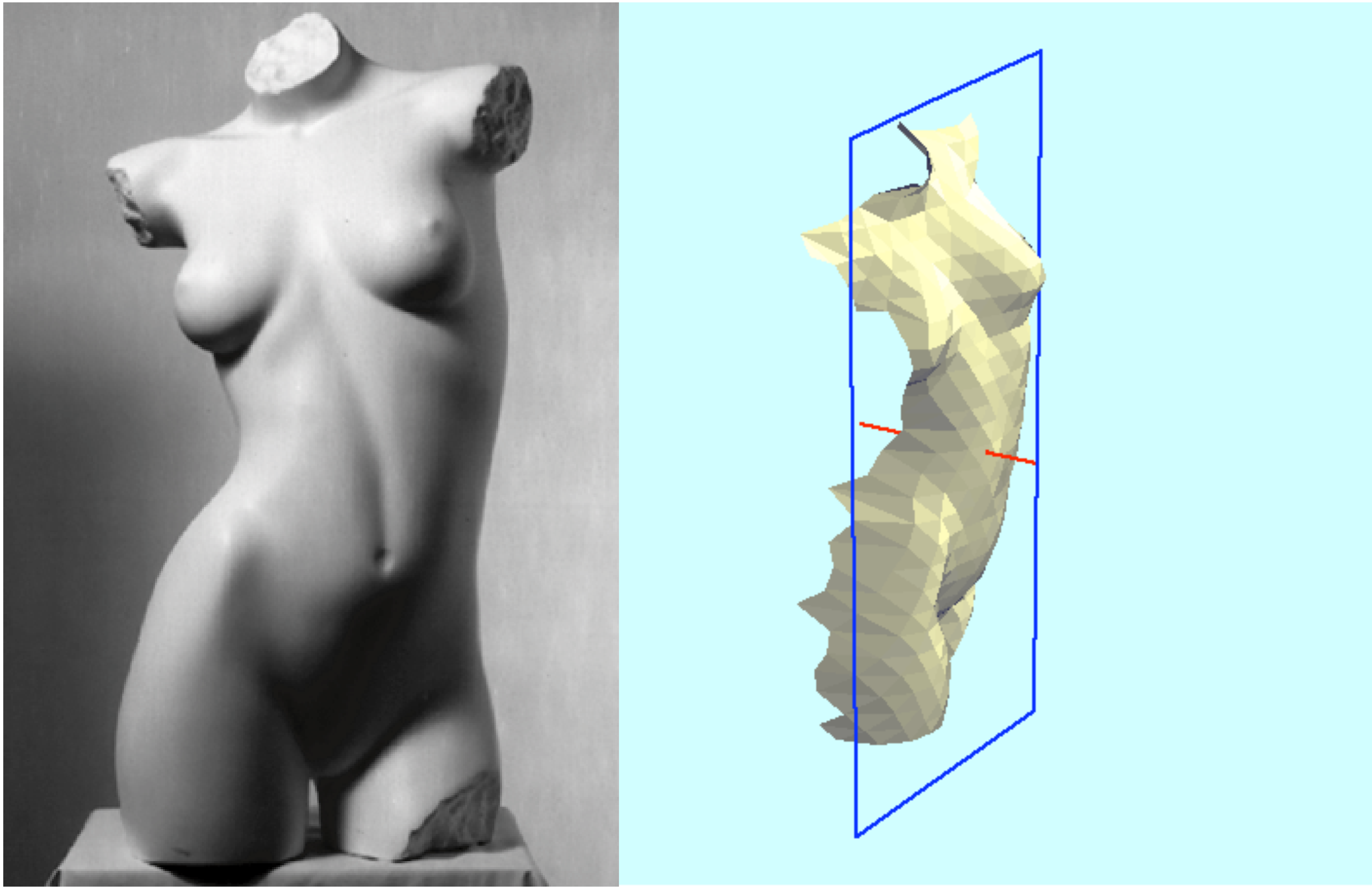
taking values on the interval $[-\pi/2, +\pi/2]$. Shapes are related as “object & mould”, shapes $\pm\pi/2$ are umbilicals, $s = \pm\pi/2$ cylinders and $S = 0$ symmetrical saddles.

Notice that the Casorati curvature is a *size*, the local *shape* is fully characterized by a single number, the shape index.

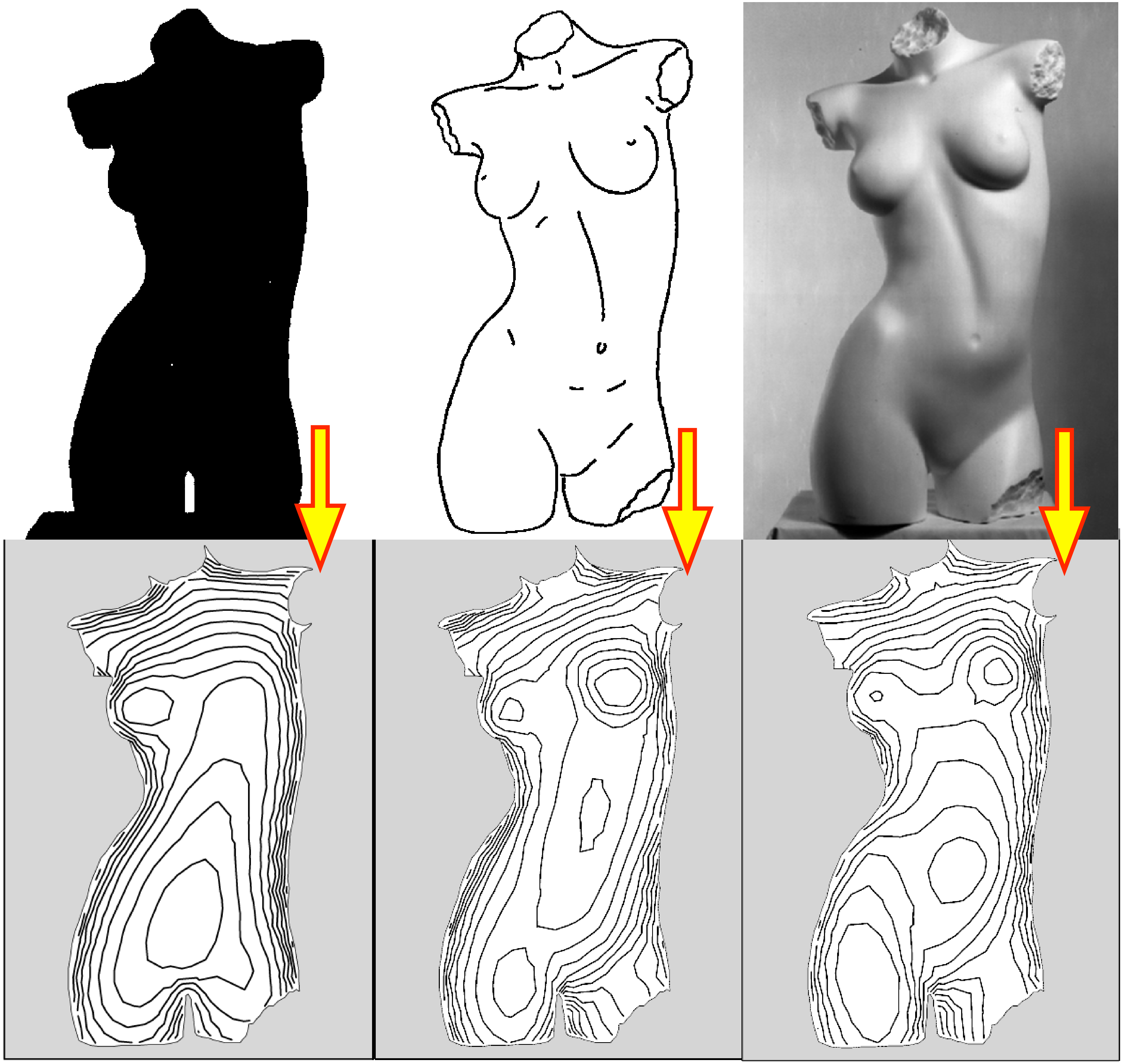


lowest order local shape





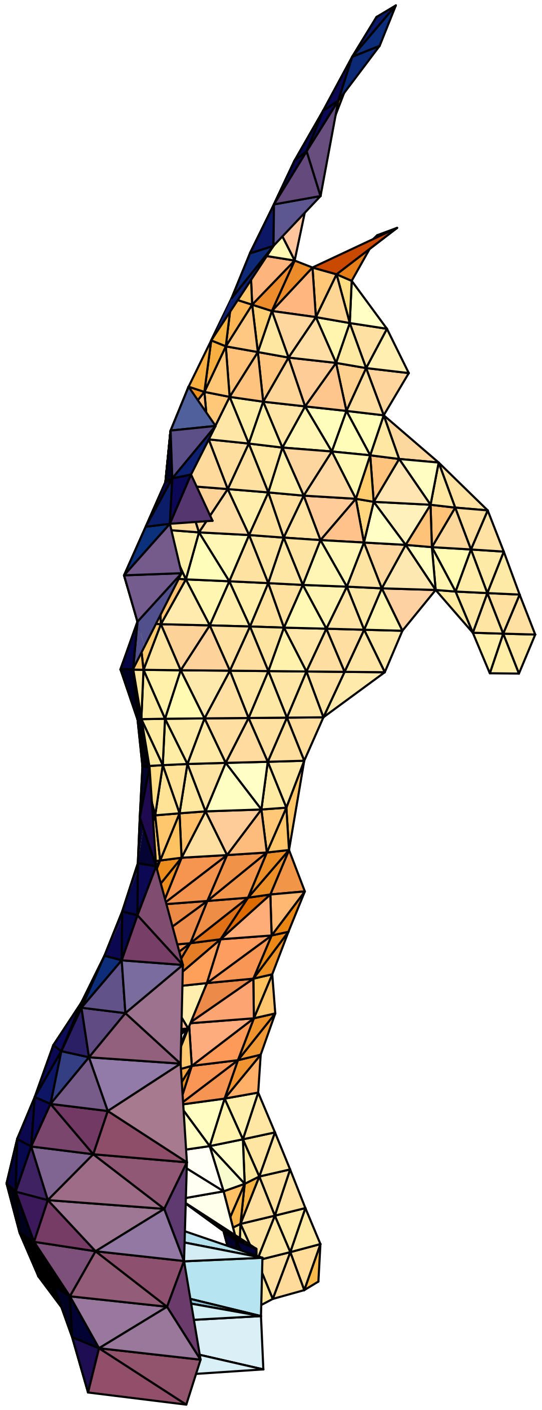
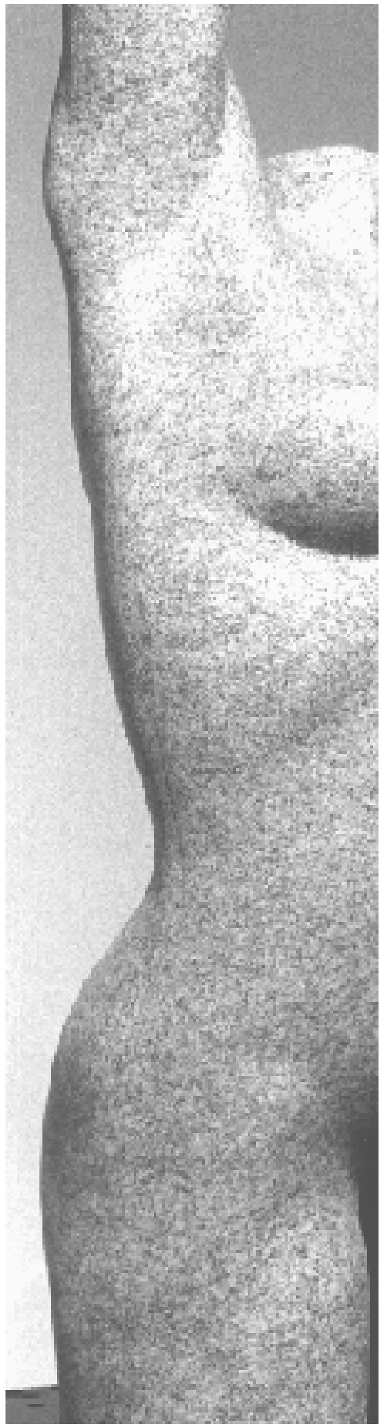
Various psychophysical methods allow one to study "pictorial relief" of human observers as a function of viewing conditions, pictorial cues, and various psychological factors. The "responses" are *relief*, *i.e.*, graphs in pictorial space and can be analyzed by the machinery of the geometry of isotropic space.



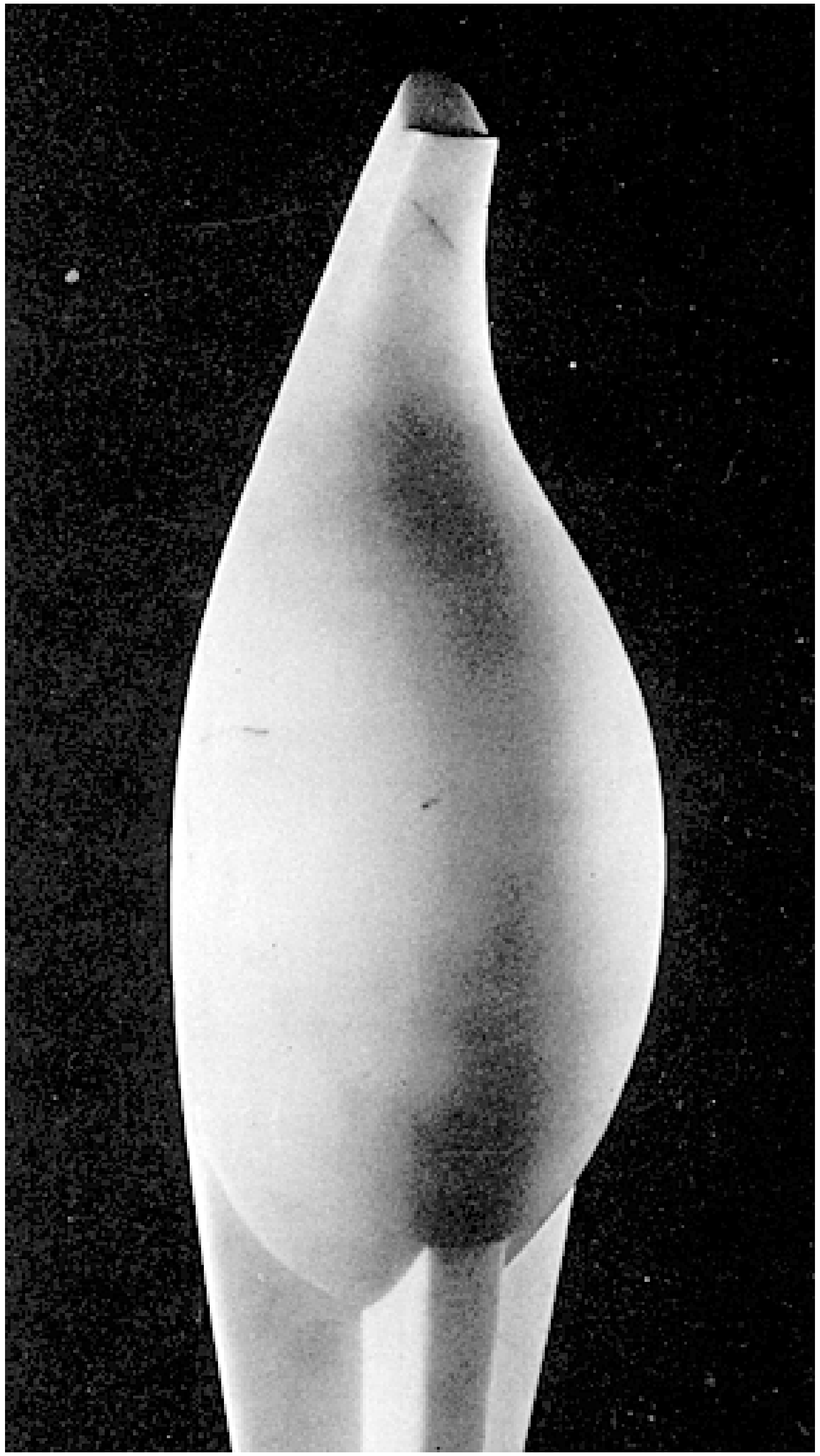
The relief depends upon the richesse of the bouquet of cues.

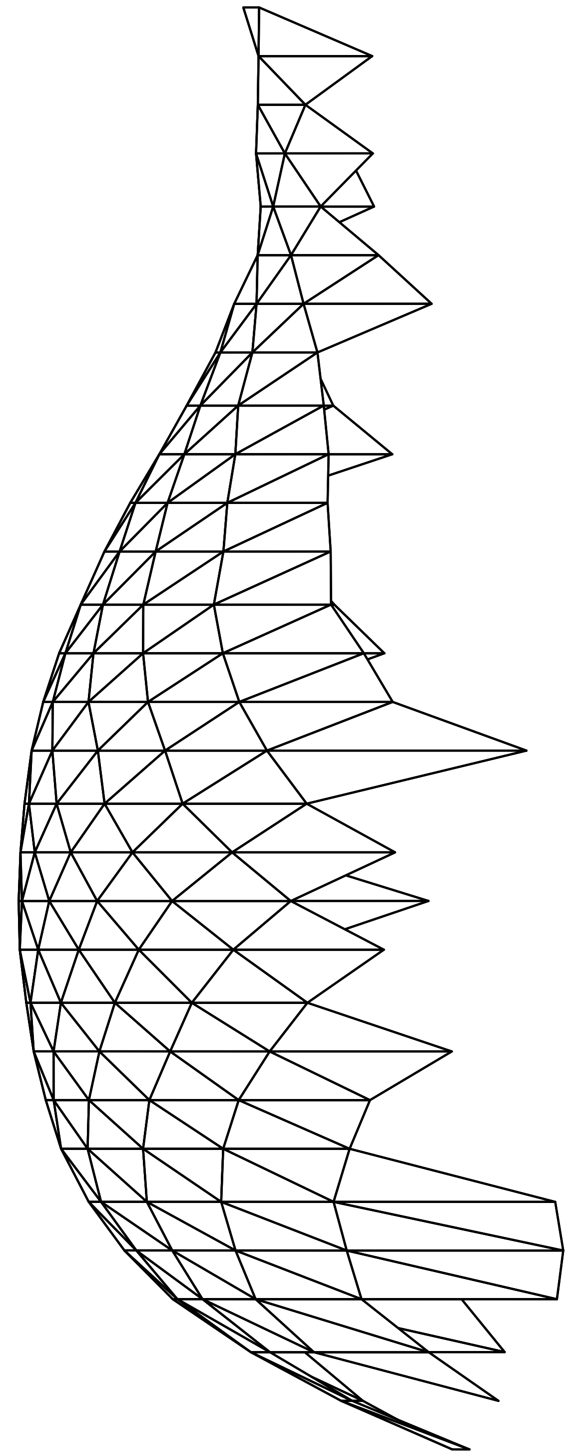
Observers are ***similar*** if there are plenty of cues *(then they are "driven" by the optical structure)*.

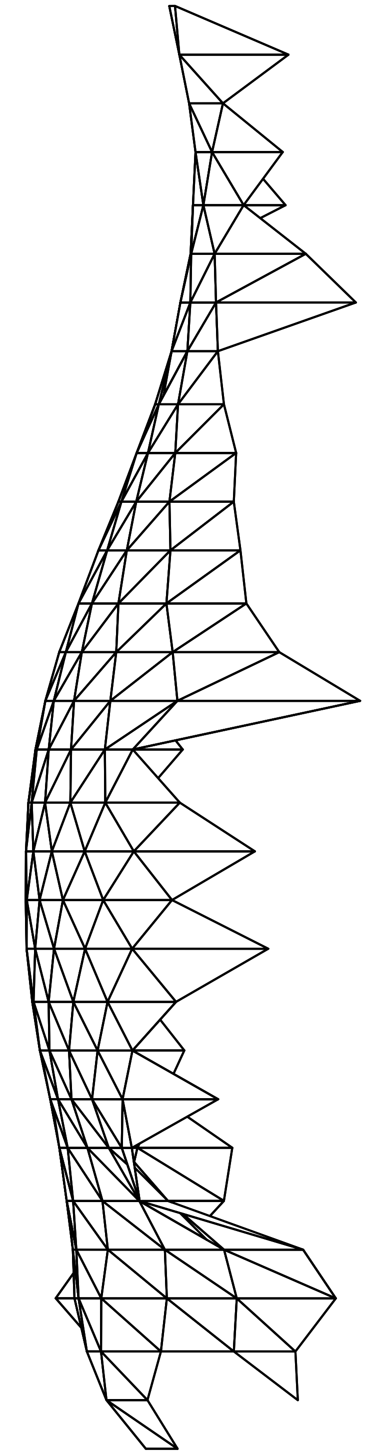
Observers are ***idiosyncratic*** if cues are hard to come by *(then they are driven by their constructive imagination)*.

Responses are never “veridical” in the naïve sense, nor are they expected to be!

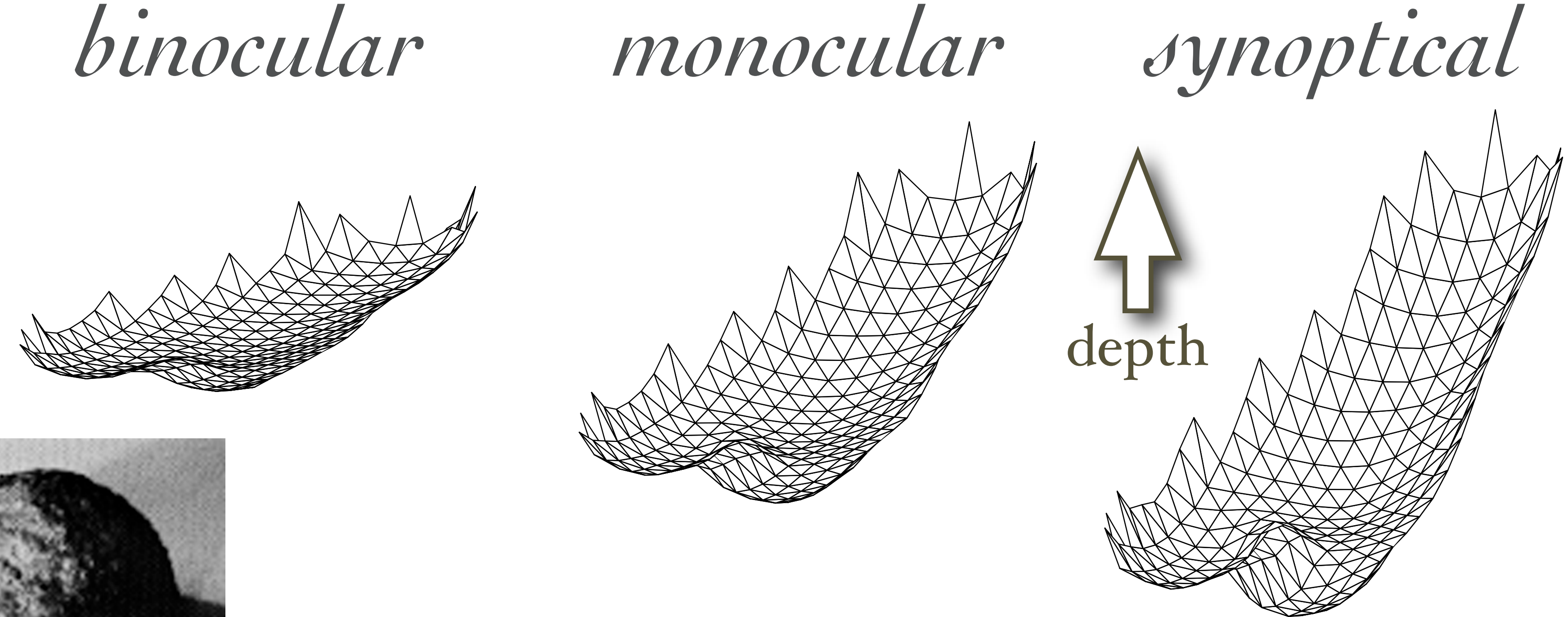


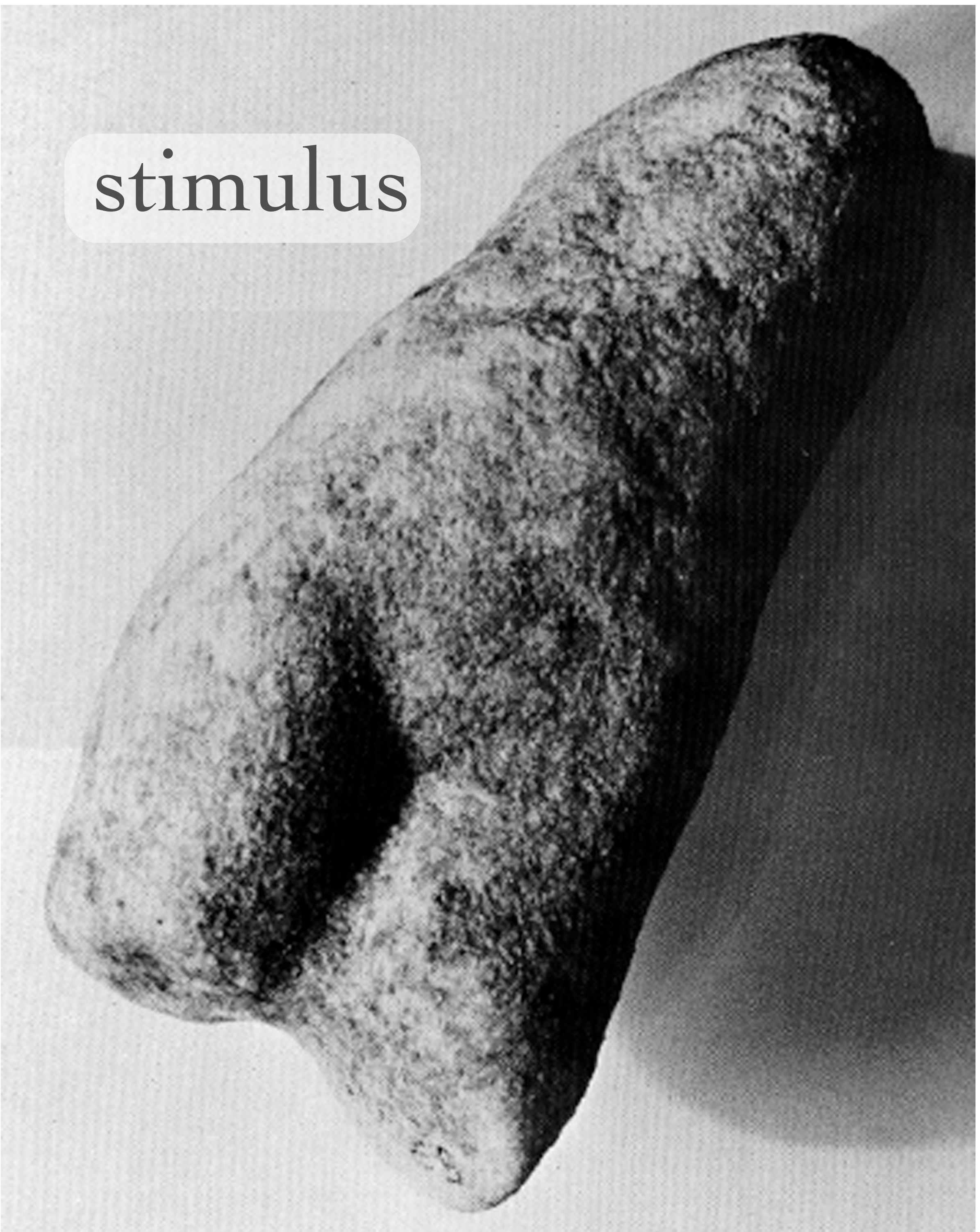


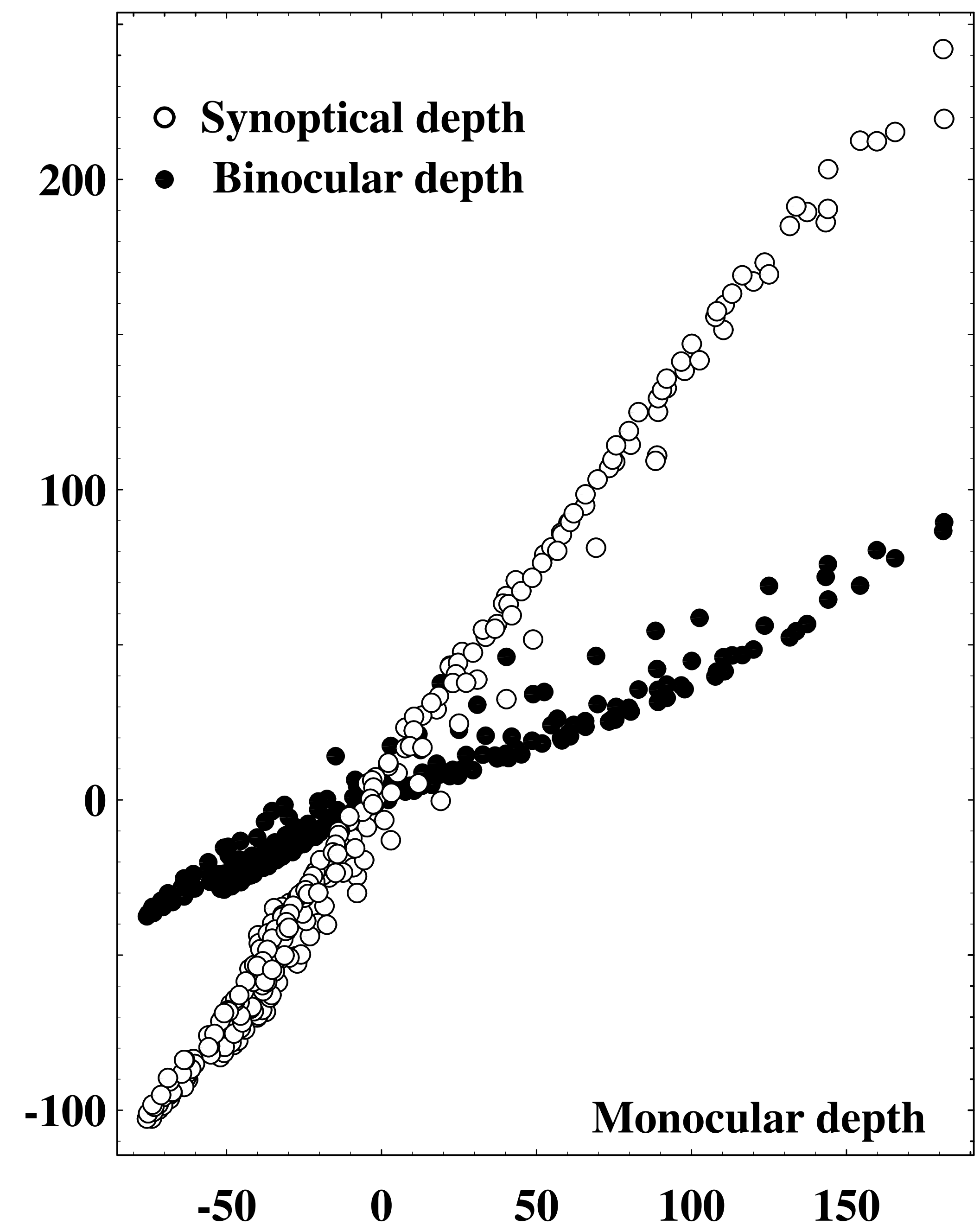


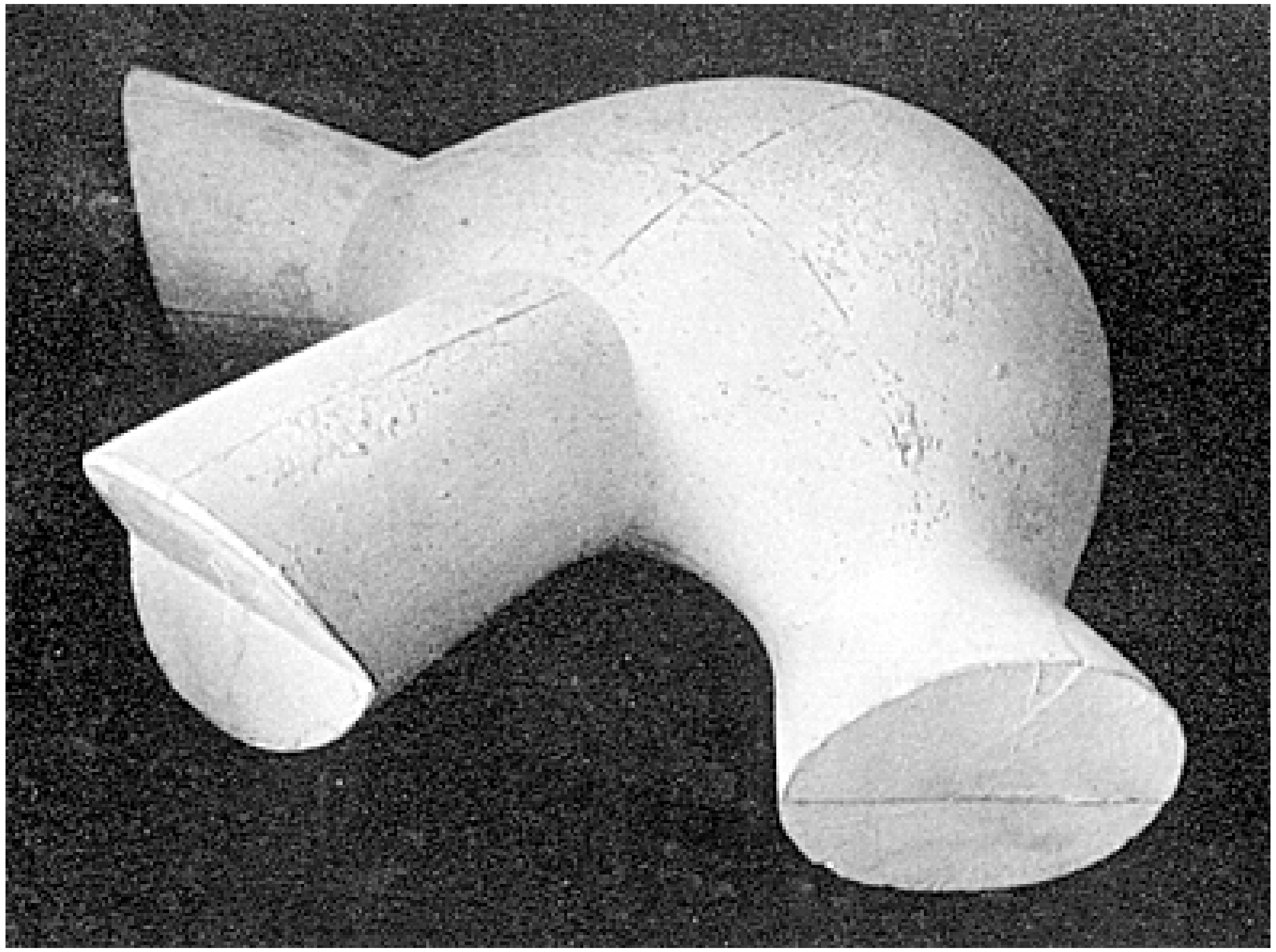
Observers use idiosyncratic depth scalings.

viewing modes affect depth of relief

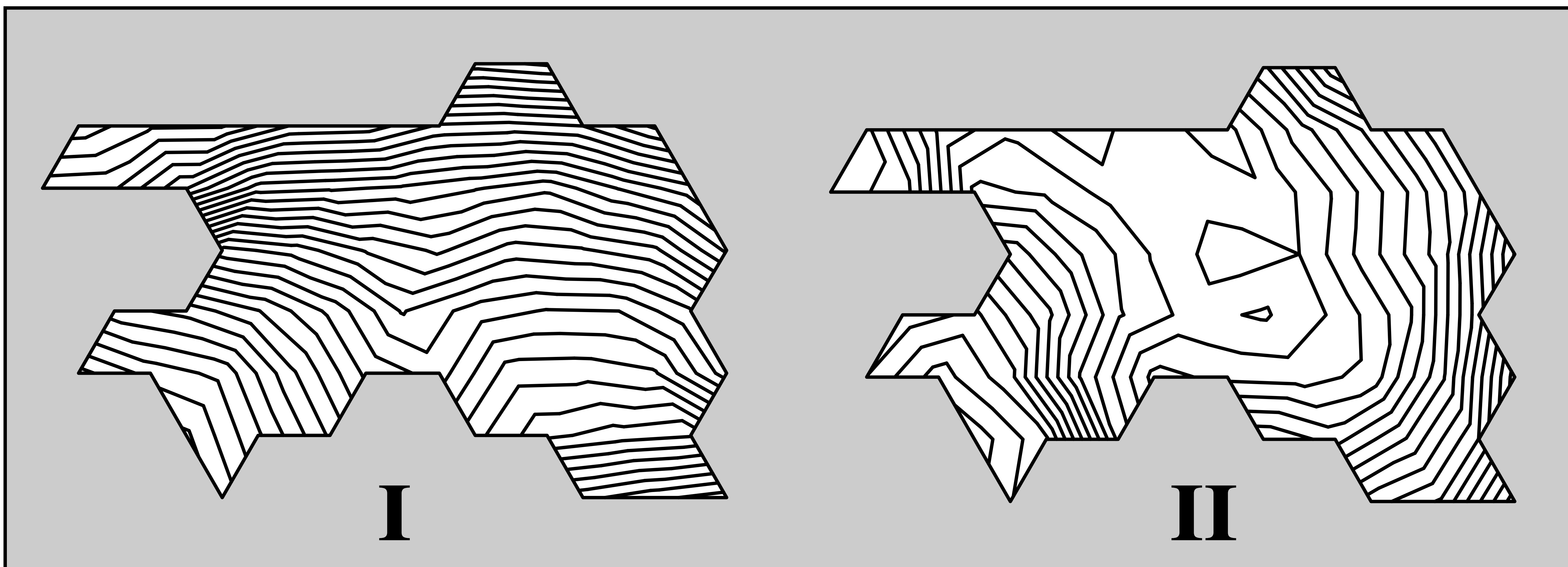


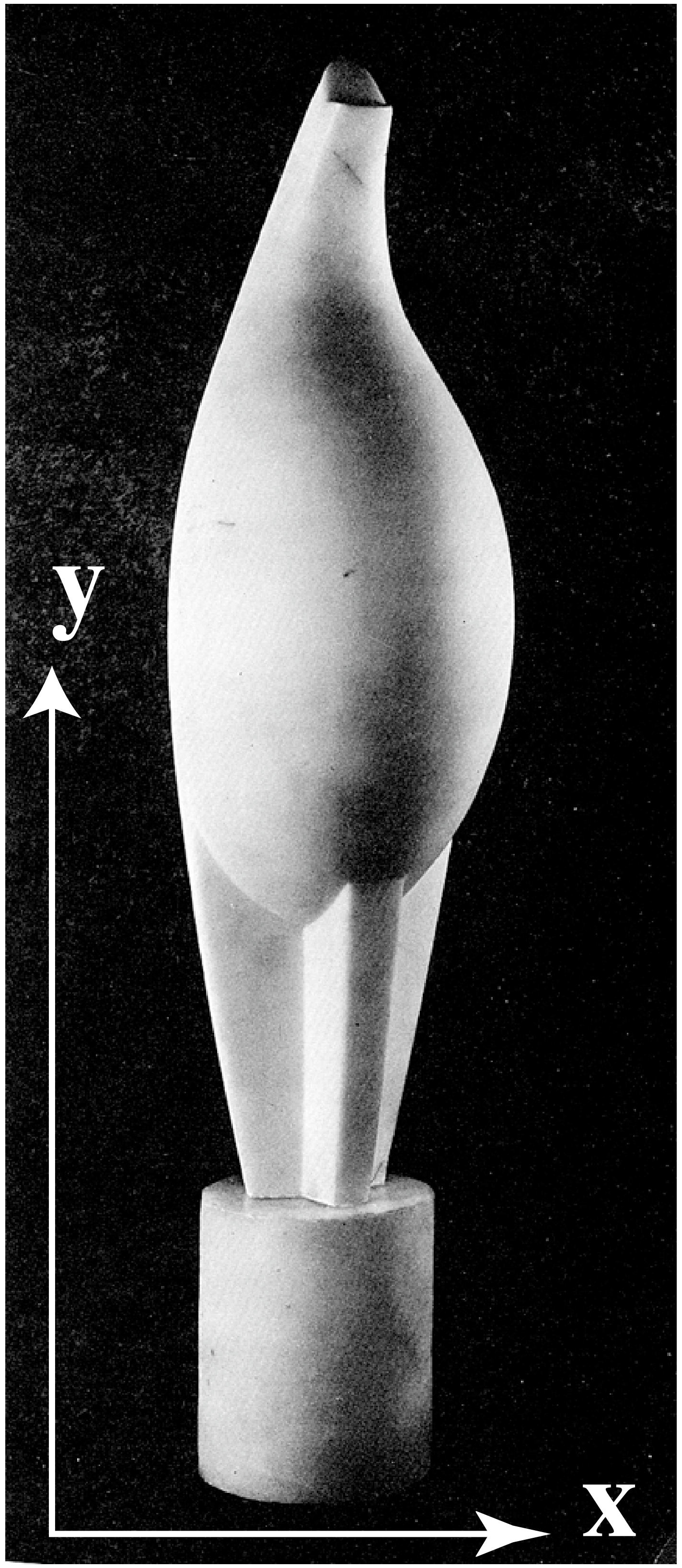




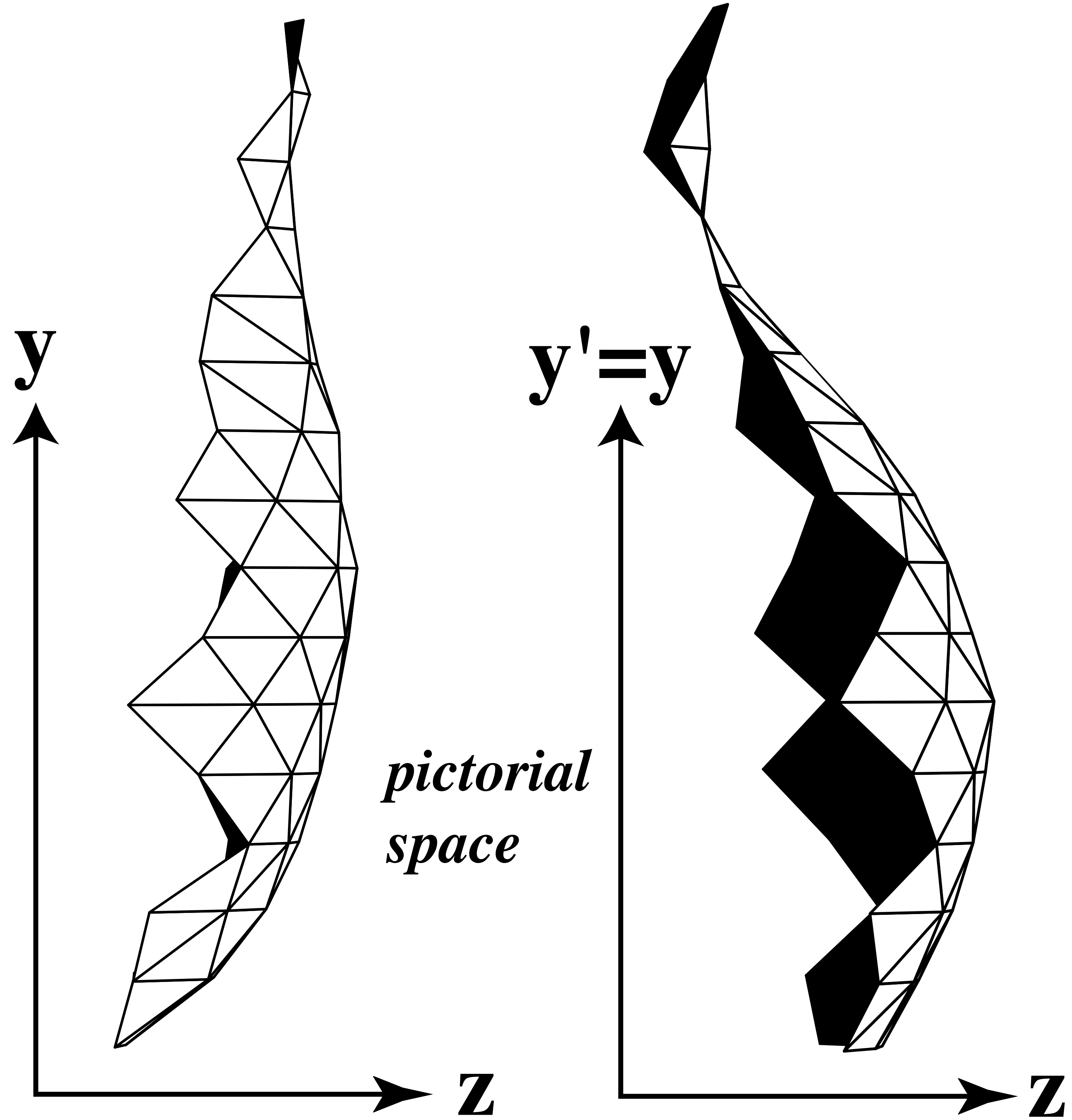


Observers use idiosyncratic (isotropic) rotations.





picture plane



$$z' = \alpha\, x + \beta\, y + \zeta\, z$$

Conclusions

• the default structure of the "visual space" of a monocular, stationary observer is the vector bundle $\mathbb{S}^2 \times \mathbb{A}^1$, the "visual field" times the *isotropic* log–distance dimension;

• "pictorial space" is $\mathbb{E}^2 \times \mathbb{A}^1$, an "infinitesimal" (indefinite size!) patch of the visual field. Its "proper motions" coincide with the intersection of the ambiguity groups of the "monocular cues";

• the differential geometry of "reliefs" is non–Euclidean. *E.g.*, visual things have only "front", but no "back" sides;

• human psychophysics of "pictorial relief" (that is "pictorial SHAPE") is well described by this geometry. The (idiosyncratic) "mental movements" are just the proper motions of singly isotropic space.

thank you for your attention

jan.koenderink@telfort.nl

TUDelft