Measurement Duality*

Luc Florack

Introduction

It is a public fact that "shape" can only be defined in operational terms.

> "Shape" is only operationally defined. Thus things do not "have a shape" the way Santa Claus has a red suit.
>
> —JAN KOENDERINK (1990)

A shape can be *observed* or *manufactured*, and even its mere thought entails the possibility of a *potential* realization of such an act (an imagination in your mind's eye, a gesticulation). Shape as an attribute of an observable puts the very role of observation—generically the physical interaction with some kind of source—into focus.

> Ich werde [. . .] in der transzendentalen Überlegung meine Begriffe jederzeit nur unter den Bedingungen der Sinnlichkeit vergleichen müssen, und so werden Raum und Zeit nicht Bestimmungen der Dinge an sich, sondern der Erscheinungen sein: was die Dinge an sich sein mögen, weiß ich nicht, und brauche es auch nicht zu wissen, weil mir doch niemals ein Ding anders, als in der Erscheinung vorkommen kann.
>
> —IMMANUEL KANT (1787)

One branch of mathematics, known as *distribution theory*, and in particular its central concept of *duality*, is particularly suited for the purpose of operational shape description, because it provides a mechanism for probing shapes. That is, it explicitly accounts for the indispensable role of a sensorium—or motorium, the analogy is somewhat whimsical—in the representation of shape.

Just for the gist of it I consider a very specific kind of duality and a very specific type of source fields, viz. *topological duality* in the context of the structural description of *digital images*. "Raw images"—a collection of pixels on a lattice—are much like "die Dinge an sich", not because they are not operationally represented (on the contrary!), but because their form reflects computer architecture rather than "relevant" structure. To abstract from the machine

*This work has been presented at the Spring School "Biophysics of Shape", organised by the Helmholtz Institute, Utrecht, The Netherlands, April 2–4, 1997.

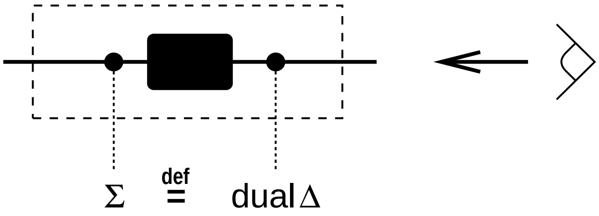

Figure 1: Duality amounts to a "what-you-see-is-all-you've-got" philosophy and boils down to straightforward linear filtering.

implies that one declines from the representation dictated by machine technicalities, such as discretisation and quantisation details, in return for an alternative, more sophisticated format ("die Erscheinung").

Theory

Let us denote the class of all possible raw images by Σ, and call it *state space* for ease of reference. Instead of directly accessing pixels, we postulate a *device space* of filters, Δ say—sensorium or motorium—and monitor the *joint response* triggered by an input image: Figure 1. Formalised:

Paradigm 1 *State space is the topological dual of device space:* $\Sigma \stackrel{\text{def}}{=} \Delta'$.

This paradigm entails that we conceive of a grey-value as the output $F[\phi] \in \mathbb{R}$ of a linear filter $\phi \in \Delta$, not as a function or pixel value $f(x) \in \mathbb{R}$. (I use the latter to model a "raw image".)

In the context of duality it is natural to consider *equivalence classes*: input images that trigger identical responses are equivalent. Example: suppose we have only one filter at our disposal, which reads its input f and returns its pixel average $\overline{f}$. Then $g \sim f$ iff $\overline{g} = \overline{f}$. Clearly lots of images map to the same mean value. One cannot expect to solve a sophisticated image analysis task on the basis of such a poor "resolution". In view of a *generic* image processing framework, we cannot content ourselves with this state of affairs. A conceptual image F must somehow be one-to-one related to the given pixel data f, except for non-measurable nitty-gritty details. That is, if $g(x) = f(x)$ almost everywhere, then $G[\phi] = F[\phi]$ for all physically plausible filters ϕ, *and vice versa*.

This leads us to the question: what are "physically plausible" filters? Without loss of generality—and by virtue of various consistency considerations beyond the scope of this summary—one can adopt the class proposed by Laurent Schwartz.

Definition 1 *Let* $\mathcal{S}(\mathbb{R}^n)$ *be the class of smooth functions of rapid decay, then* $\Delta \stackrel{\text{def}}{=} \mathcal{S}(\mathbb{R}^n)$*, whence* $\Sigma \stackrel{\text{def}}{=} \mathcal{S}'(\mathbb{R}^n)$*. The latter is also known as the class of tempered distributions.*

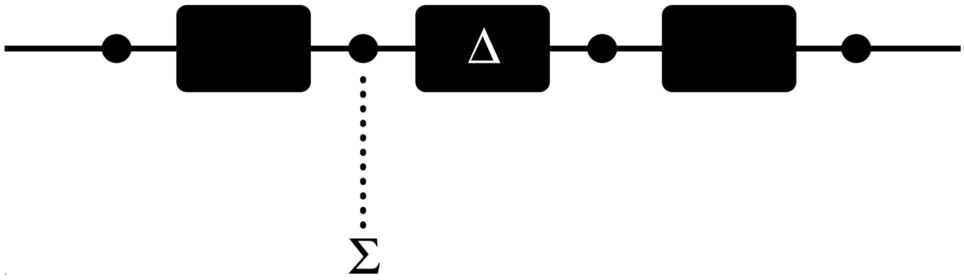

Figure 2: Image processing calls for an algebraic structure on device space.

The majority of tempered distributions is "regular", meaning that they can be written by an integral formula known as the *Riesz representation formula*:

$$F[\phi] = \int dz\, f(z)\, \phi(z)\,,$$

for some function f of polynomial growth. Tempered distributions which are not regular always involve the *Dirac distribution*: $\delta[\phi] \equiv \phi(0)$. In practice one uses the integral formula even in these cases, associating the Dirac distribution with the "function-under-the-integral" $\delta(z)$. Such distributions are not at all pathological, and share virtually all nice properties characteristic for distributions in general, e.g. infinite differentiability! "Point stimuli" lie at the core of "reverse engineering" disciplines, in which one aims to establish filter profiles of a black box system. In image analysis the stimulus is given (a raw image f), and filters are defined *ad libitum*.

Rapid decay reflects filter confinement. Despite this filters may have *any* size; a suitable measure of size is the usual normalised second order central momentum. Smoothness of $\mathcal{S}(\mathbb{R}^n)$ is hardly a demand. This follows from the fact that $\mathcal{S}'(\mathbb{R}^n)$ is larger than any of the function spaces typically employed in non-dualistic models. In other words, the filter class $\mathcal{S}(\mathbb{R}^n)$ guarantees a more-than-sufficient segregation of quality. In fact, if $F[\phi] = G[\phi]$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$, then f and g differ by at most a non-measurable function, exactly as desired.

One *image processing consistency* demand should be mentioned. If output is interpreted as potential input to yet another filtering stage, then it can be shown that Δ must constitute a *convolution algebra*: Figure 2. Fortunately this is the case for $\mathcal{S}(\mathbb{R}^n)$!

"Scale-space theory" is basically Schwartz theory equipped with a point concept: a positive "zeroth order" filter consistent with the image processing demand, *i.e.* one generating an *autoconvolution algebra*. There exists only one point operator in $\mathcal{S}(\mathbb{R}^n)$, *viz.* the normalised Gaussian (of arbitrary width and base point). For later use I will collectively denote these operators, together with all their derivatives, by $\mathcal{G}(\mathbb{R}^n)$.

Definition 2 *A scale-space representation is obtained by subjecting a raw image to the Gaussian family* $\Delta \stackrel{\text{def}}{=} \mathcal{G}(\mathbb{R}^n)$*, in other words, it is an element of* $\Sigma \stackrel{\text{def}}{=} \mathcal{G}'(\mathbb{R}^n)$*.*

If you look at the integral formula, you will appreciate that there are always dual interpretations explaining a given change in grey-value: either your filter, or your input image has changed.

This trivial observation has far-reaching consequences, and is formalised as follows. When we talk of a "spatial transformation" (shift, rotation, scaling, *etc.*), we are in reality thinking of relative spatial relationships between *physical objects*.

Definition 3 (Push Forward) *Let $\theta : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto \theta(x)$ be a spatial transformation. The push forward of a filter is then defined as the mapping*

$$\theta_* : \Delta_x \to \Delta_{\theta(x)} : \phi \mapsto \theta_*\phi \stackrel{\text{def}}{=} |\det \nabla\theta^{\text{inv}}|\, \phi \circ \theta^{\text{inv}} .$$

Subscripts attached to Δ indicate what happens to a filter's centre of gravity (whence the terminology). One naturally "pulls back" the source field in the dual view.

Definition 4 (Pull Back) *With $\theta(x)$ and its push forward $\theta_*\phi$ as defined in the previous definition, the pull back of the input image is defined as the mapping*

$$\theta^* : \Sigma_{\theta(x)} \to \Sigma_x : F \mapsto \theta^* F \quad \textit{defined by} \quad \theta^* F[\phi] \stackrel{\text{def}}{=} F[\theta_*\phi] .$$

Note that the base point ("focus of attention") now moves in opposite direction! These definitions may seem a bit abstract on first sight, but both $\theta_*\phi$ (filter transformation) as well as θ^*F (transformation of input image) should make more sense than $\theta(x)$ (transformation of a "void"!), and in a way it is better to say that one of the former two defines the others. Indeed, it is easy to think of dual filter/image transformations without existence of an underlying spatial transformation, but impossible to make sense of the latter without any manifestation on physical objects:

> Space and time are not measurable in themselves: they only form a framework into which we arrange physical events.
>
> —MORITZ SCHLICK (1920)

If you write Definition 4 in integral form you will see that it is basically a change of integration dummies:

$$\theta^* F[\phi] = \int dz\, f(\theta(z))\, \phi(z) = [\text{subst. } y = \theta(z)] = \int dy\, |\det \nabla\theta^{\text{inv}}(y)|\, f(y)\, \phi(\theta^{\text{inv}}(y)) = F[\theta_*\phi] .$$

Example: one can acquire an image either by shifting a patient underneath a scanner, or by moving the scanner in the opposite sense. This generalises to any not necessarily rigid transformation (at least conceptually: in this example one option will likely have legal implications…). The relevant formulas in this case are: $\theta(z) = z + x$, $\theta^* f(z) = f(z + x)$, $\theta_*\phi(z) = \phi(z - x)$, and the above equality can be rewritten as the correlation $f \star \phi(x)$ in two equivalent ways.

Definitions 3 and 4 form the basis for all image manipulations. I will mention a few important ones.

From Samples to Images

The step from a mere grey-value sample $F[\phi]$ to an actual output image $F \star \phi(x)$ in the patient example relies on the push forward/pull back principle, applied to spatial translations over all vectors x within the relevant field of view.

A consistent image model takes into account all symmetries of space and time, *i.e.* not only translations (homogeneity), but also spatial rotations (isotropy) and spacetime scalings (scale invariance). The technique remains the same.

Derivatives

If you want to take a *derivative* of your image, you implement the *dual derivative* of your basic filter. The relevant formula is

$$\nabla F[\phi] \stackrel{\mathrm{def}}{=} F[-\nabla \phi] \,.$$

(Why the minus sign? Hint: use the integral formula and your secondary school math.) Note that the r.h.s. is both well-defined as well as operationally realizable (*viz.* by plain linear filtering). This generalises to any order!

Temporal Causality

It is sometimes argued that the Gaussian filter family is not suited in case *manifest temporal causality* is a prerequisite, as in active vision. However, the following "Koenderink trick" can be applied (consider only 1D temporal sequences $f(t)$ for simplicity).

The basic observation is that there must be *some* time domain in which the Gaussian family makes sense (by virtue of its uniqueness there is no alternative), say parametrised by a parameter $s \in \mathbb{R}$. The physically reasonable domain for a visual system actively participating in the world is of course the history part of the time axis. Therefore discard the unknown future by introducing a time horizon (the present moment) and mapping the past semi-axis onto the s-domain: Figure 3 (left). Once such an isomorphism has been established, say $t = \tau(s; a)$—note that it depends on the present moment a—follow your nose: the isomorphism acts on a filter $\phi(s)$ by the recipe of "push forward", yielding a filter $\tau_* \phi(t; a)$—which again depends on a—producing the desired filter profile in the t-domain: Figure 3 (middle and right).

The construction reminds us of Kant's remark, but note that the "causal world" picture obtained by the dual action of pull back,

$$\tau^* F[\phi] = F[\tau_* \phi] \,,$$

cannot be disqualified on objective grounds! The F (with Riesz representation $f(t)$) on the r.h.s. could be a causally processed but fully recorded video tape, *i.e.* causality is introduced by the act of observation as Kant conjectures (note that the content of the tape does *not* depend on the present moment a). The $\tau^* F$ on the l.h.s. is a signal in the s-domain—which does depend on a—and could reflect the signal history as a function of fiducial acquisition time a (if history is what you recall of it then it indeed changes over time!).

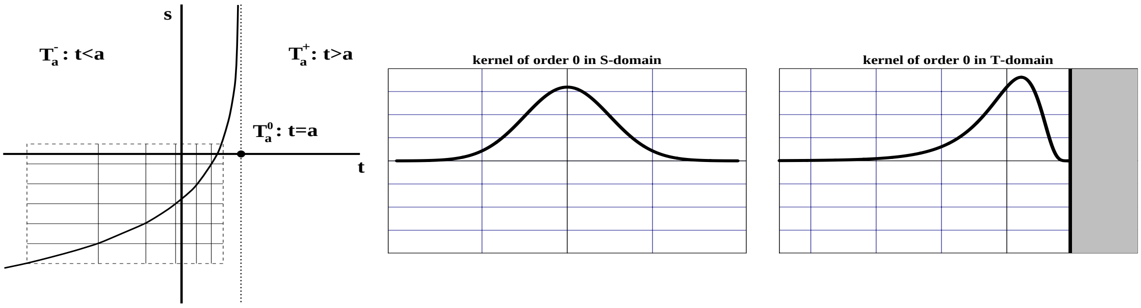

Figure 3: Left: the isomorphism $t = \tau(s; a)$ discards the unknown future. The asymptote indicates the time horizon. The box delimits a typical time window for a real-time system. Note that uniform sampling of the **S**-domain implies a *graded resolution history*. Middle and right: Comparison of causal point operator profiles in s and t domains. The map τ_*—which depends on a—brings us from there to the t-representation on the right, which corresponds to a fixed delay $a - t$. Time progresses from left to right; the causal filter vanishes rapidly but smoothly towards the time horizon. (Of course there is no such moment in the left graph.) The shaded region indicates the unrevealed future.

Optic Flow

The appropriate differential tool in the context of optic flow is the *Lie derivative*. It expresses the rate of change of a quantity when moving in the direction of a vector field $\boldsymbol{v}$, say, and is proportional to that vector field. The following definition therefore suggests itself:

$$(\boldsymbol{v} \cdot \nabla)F[\phi] \stackrel{\text{def}}{=} F[-\nabla \cdot (\boldsymbol{v}\,\phi)]\,.$$

The l.h.s. has Riesz representation $(\boldsymbol{v} \cdot \nabla)f(x)$, the usual directional derivative for a scalar function. Transposed to filter space on observes that the dual Lie derivative of a filter ϕ must be $-\nabla \cdot (\boldsymbol{v}\,\phi)$. Apart from the minus sign we see that the gradient operator acts on both filter as well as vector field, *i.e.* contains an additional divergence term $\phi\,\text{div}\,\boldsymbol{v}$ not present in the scalar case.

Optic flow is a velocity pattern $\boldsymbol{v}$ such that if one comoves with the induced flow *some measurable entity* is preserved. The simplest instance of such entity is $F[\phi]$, *i.e.*

$$\frac{d}{dt}F[\phi] = (\boldsymbol{v} \cdot \nabla)F[\phi] + \frac{\partial}{\partial t}F[\phi] = 0\,,$$

or, in dual form (omitting the overall minus),

$$F[\frac{d}{dt}\phi] = F[\nabla \cdot (\boldsymbol{v}\,\phi)] + F[\frac{\partial}{\partial t}\phi] = 0\,.$$

A suggestion for solving the latter will appear in the International Journal of Computer Vision in the near future.

References

[1] L. M. J. Florack. Data, models, and images. In P. Delogne, editor, *Proceedings of the International Conference on Image Processing 1996*, volume 1, pages 469–472, Lausanne, Switzerland, September 16–19 1996. IEEE.

[2] L. M. J. Florack, W. J. Niessen, and M. Nielsen. The intrinsic structure of optic flow incorporating measurement duality. Accepted for publication in the International Journal of Computer Vision, 1996.

[3] I. Kant. *Critik der Reinen Vernunft*. Johann Friedrich Hartknoch, Riga, second edition, 1787.

[4] J. J. Koenderink. Scale-time. *Biological Cybernetics*, 58:159–162, 1988.

[5] J. J. Koenderink. *Solid Shape*. MIT Press, Cambridge, 1990.

[6] M. Schlick. *Space and Time in Contemporary Physics*. Oxford University Press, London, 1920.

[7] L. Schwartz. *Théorie des Distributions*. Hermann, Paris, second edition, 1966.