

What is an image?

Laurent Lafforgue

(Huawei Paris Research Center, Boulogne-Billancourt, France)

Université du Luxembourg, December 5th, 2022

Some fundamental questions:

- **Starting remarks:**

- Image processing is an important part of computing and of AI.
- Any type of computing is based at least implicitly on a type of mathematical modelling.

- **The derived problems of image modelling:**

- Which type of mathematical objects can model images?
- Which type of mathematical relation can model the relationship of images with the real things of which they are the images?

- **The derived problems of image classification:**

- Are there spaces of images, i.e. geometric objects whose points parametrize possible images of some type?
- Are there theories of images (in the mathematical sense)?

Numerical images:

- **Basic facts:**

- Cameras usually capture images in the form of numerical images.
- Numerical images consist in collections of "picture elements" (= pixels).
- Each pixel consists in a measure (i.e. a number) of intensity of light in one or several colors.
- Pixels are parametrized by a pair of finite discrete sets of plane coordinates.

- **Mathematical interpretation:**

- Each collection of measures of intensity indexed by a finite discrete set of plane coordinates is a discrete approximation of a numerical function defined on a plane domain (usually a rectangle).

Which key mathematical notions are involved?

- **The notion of number:**

Measures of intensity of light are given as numbers.

- **The notion of point:**

- Pixels are indexed by discrete finite sets of points of some plane domain.
- These discrete finite sets are thought as approximations of continuous domains which are made of infinitely many points.

- **The notion of (numerical) function:**

- Numerical images are thought as more or less accurate approximations (depending on the chosen number of pixels) of numerical functions defined on a chosen plane domain.

Two foundational problems raised by numerical images:

- **For our mind, images do not consist in numerical functions:**

- Our mind has the ability to feel intensity of light, but without any perceived relation to numbers.
- Our mind distinguishes colors but without any perceived relation to frequency measure.
- For our mind, points do not exist and images do not consist in points.
- For our mind, images appear as patchworks of subdomains to which we are able to give names and which are tied by relations, especially relations of
 - { relative positions,
 - { inclusion,
 - { decomposition.

- **The theory of numerical functions is not a theory of images:**

- Almost all numerical functions (or discrete families of pixels) do not correspond to images that may appear.
- The linear spaces of functions (or of collections of pixels) are **not** spaces of images.

An induced natural problem:

- **Question:**

Is it possible to propose mathematical modellings for images which would correspond more closely to the way our mind sees images?

- **Two a priori requirements:**

- Such modellings of images (and of their relations with the real things of which they are the images) should be “geometric” in a mathematical sense.
- Such modellings of images as our mind sees them should also be related to “language” in a mathematical sense.
Indeed, we have the ability to describe images with words.

Classical geometry: embedded geometric shapes

- **Definition by equations in families of variables:**

→ Classically, a geometric shape is defined by a family of equations

$$f_1(x_1, \dots, x_n) = 0,$$

...

$$f_k(x_1, \dots, x_n) = 0$$

in variables x_1, \dots, x_n .

→ These equations may be

{ algebraic,
analytic,
differential,
...

→ It is understood that such equations define subsets (consisting in elements called its “points”) of n -dimensional affine spaces endowed with coordinate functions x_1, \dots, x_n .

- **Definition by inequalities:**

→ One can also define subdomains by a combination of

equations $f_i(x_1, \dots, x_n) = 0, 1 \leq i \leq k,$

and inequalities $g_j(x_1, \dots, x_n) \geq 0, 1 \leq j \leq \ell.$

→ Such subdomains have stratified boundaries.

→ They can be glued along boundary components.

Classical geometry: induced inner topological structures

Any geometric shape defined in some n -dimensional affine space by a family of equations (and possibly inequalities) may be endowed with structures induced from the ambient space:

- An underlying set of “points”:

This is the subset of n coordinates points (x_1, \dots, x_n) which verify the prescribed equations (and possibly inequalities).

- Possibly, a metric:

Any “distance” function on the ambient affine space restricts to its embedded geometric shapes.

- Most often, a topology:

→ It consists in a family of subsets of the underlying set, stable under finite intersection and arbitrary union, called the “open subsets” U .

→ The family of open subsets is ordered by inclusion

$$V \subseteq U.$$

→ It is endowed with an induced notion of coverings

$$U = \bigcup_{i \in I} V_i$$

of open subsets U by smaller open subsets V_i .

Classical geometry: induced inner geometric structures

Any geometric shape X defined in some n -dimensional affine space \mathbb{A} may also be endowed with a more refined geometric structure of the form:

$$\begin{array}{ccc}
 U & \longmapsto & \mathcal{O}(U) \\
 \parallel & & \parallel \\
 \text{arbitrary open subset} & & \text{ring of functions on } U \\
 \text{of } X & & \text{which verify some "local" condition} \\
 & & \text{[e.g.: } C^\infty, \text{ analytic, algebraic, ...]}
 \end{array}$$

- Such a structure \mathcal{O} is a "presheaf" (of rings) in the sense that any inclusion $V \subseteq U$ defines a restriction map $\mathcal{O}(U) \rightarrow \mathcal{O}(V)$.
- It is even a "sheaf" (of rings) in the sense that, for any covering

$$U = \bigcup_{i \in I} V_i,$$

the family of restriction maps

$$\mathcal{O}(U) \rightarrow \prod_{i \in I} \mathcal{O}(V_i)$$

induces a bijection to

$$\left\{ (f_i \in \mathcal{O}(V_i))_{i \in I} \mid \left. \begin{array}{l} \text{restriction of } f_i \text{ to } W = \text{restriction of } f_j \text{ to } W \\ \text{for any } W \text{ with } W \subseteq V_i, W \subseteq V_j \end{array} \right\} .$$

Classical geometry: notions of intrinsic geometric structures

A geometric object of some type [e.g. differential, analytic, algebraic, ...] can be defined as consisting in:

- An underlying set of “points” X .
- A topological structure on X
(possibly refined by a “metric” = distance function)
consisting in a family of “open subsets” $U \subseteq X$,
stable under finite intersection and arbitrary union,
ordered by inclusion $V \subseteq X$
and endowed with the induced notion of covering

$$U = \bigcup_{i \in I} V_i.$$

- A sheaf of rings of “coordinate functions”

$$U \longmapsto \mathcal{O}_X(U)$$

which allow X to be locally embeddable in affine spaces \mathbb{A} such that

- $$\left\{ \begin{array}{l} - \text{ the topology of } X \text{ is induced by that of } \mathbb{A}, \\ - \text{ the sheaf of rings } \mathcal{O}_X \text{ on the topological space } X \\ \text{ is induced by the sheaf of functions of the chosen type \\ \text{ [differential, analytic, algebraic, ...] on affine spaces } \mathbb{A}. \end{array} \right.$$

Classical geometry: notions of geometric maps

Consider two intrinsic geometric objects (X, \mathcal{O}_X) and (Y, \mathcal{O}_Y) of a given chosen type [e.g. differential, analytic, algebraic, ...].

- A geometric map (or “morphism”) of this chosen type is a map

which

$$u: X \longrightarrow Y$$

- is “continuous” in the sense that

V	\longmapsto	$u^{-1}(V)$
\parallel		\parallel
open subset of Y		open subset of X
- respects the geometric structures, in the sense that, for any open subset $V \subseteq Y$, composition with $u: u^{-1}(V) \rightarrow V$ induces a ring homomorphism

$$\mathcal{O}_Y(V) \longrightarrow \mathcal{O}_X(u^{-1}(V)).$$

- Geometric maps (or morphisms) naturally compose

$$(X \xrightarrow{u} Y \xrightarrow{v} Z) \longmapsto (X \xrightarrow{v \circ u} Z).$$

The composition law is associative and has “unit” morphisms id_X .

- This defines “categories”
 = collection of “objects” + “morphisms” between the objects
 + associative composition law of morphisms (with units).

Categories as natural contexts for classification problems:

Consider a category \mathcal{C} consisting in

- a collection of objects X ,
- sets $\text{Hom}(X, Y)$ of morphisms $X \rightarrow Y$ between objects,
- an associative composition law of morphisms
 $\text{Hom}(X, Y) \times \text{Hom}(Y, Z) \longrightarrow \text{Hom}(X, Z)$ with units id_X .

Definition. – An object M of \mathcal{C} is called “classifying” for some type \mathbb{T} of mathematical beings if

- for any object X of \mathcal{C} , there is a set $P_{\mathbb{T}}(X)$ of mathematical beings of type \mathbb{T} parametrized by X ,
- any morphism $X \xrightarrow{u} Y$ of \mathcal{C} defines a map of change of parameters $u^* : P_{\mathbb{T}}(Y) \longrightarrow P_{\mathbb{T}}(X)$,
- there are identifications

$P_{\mathbb{T}}(X) \xrightarrow{\sim} \text{Hom}(X, M) = \{X \xrightarrow{m} M\}$, $X = \text{object of } \mathcal{C}$,
such that the change of parameters maps $u^* : P_{\mathbb{T}}(Y) \longrightarrow P_{\mathbb{T}}(X)$
are transformed into the composition maps

$$\begin{aligned} \text{Hom}(Y, M) &\longrightarrow \text{Hom}(X, M), \\ (Y \xrightarrow{m} M) &\longmapsto (X \xrightarrow{m \circ u} M). \end{aligned}$$

Classifying spaces:

Yoneda's lemma. –

If an object $M = M_{\mathbb{T}}$ of a category \mathcal{C} is “classifying” for some “type \mathbb{T} of parametrized mathematical beings”, it is entirely characterized (up to unique isomorphism) by \mathbb{T} , as an object of \mathcal{C} .

Consequences. –

- Usually, the definition of a “type \mathbb{T} of mathematical beings” is linguistic. Proving that an object M of \mathcal{C} is “classifying” for such a \mathbb{T} yields a linguistic description of M .
- If \mathcal{C} is a geometric category with a point object $\{\bullet\}$, and $M \cong M_{\mathbb{T}}$ is classifying, the underlying set of points $\text{Hom}(\{\bullet\}, M)$ of M identifies with the set of mathematical beings of type \mathbb{T} parametrized by $\{\bullet\}$, i.e. the set of “absolute” mathematical beings of type \mathbb{T} .

Consequences for the interpretation and the classification of images:

- If the ability of our human minds to formulate linguistic descriptions of what we see could inspire mathematical modellings, it could take the form of a category of images (or of geometric objects of which they are the images) whose objects could be presented as “classifying objects” for some types of mathematical beings.
- In order to possibly define “spaces of images”, we first need to consider and study notions of “parametrized families of images”. For that purpose, it is not enough to consider discrete families of images. We need to wonder about “continuous families of images” and their possible natural parametrizing objects, related by morphisms of “change of parameters”.

Linguistic descriptions of points of topological spaces:

Consider

- a topological space X ,
- an arbitrary base \mathcal{B} of open subsets of X
(in the sense that any open subset of X is a union of elements of \mathcal{B}).

Proposition. – Suppose X is “sober” (a technical hypothesis almost always verified in practice). Then the underlying set of points of X can be recovered abstractly from \mathcal{B} as the set of maps

$$\begin{array}{ccc} U & \mapsto & MU = \text{subset of } \{\bullet\} \\ \parallel & & \parallel \\ \text{element of } \mathcal{B} & & \text{set consisting in one element} \end{array}$$

such that, for any elements $U, U', (V_i)_{i \in I}$ of \mathcal{B} verifying $U \cap U' = \bigcup_{i \in I} V_i$,

we have $MU \cap MU' = \bigcup_{i \in I} MV_i$.

Remarks. –

- In this linguistic description, an element V of \mathcal{B} is the set of maps $U \mapsto (MU \subseteq \{\bullet\})$ of this type such that $MV = \{\bullet\}$.
- In general, a topological space has infinitely many bases \mathcal{B} . They provide as many different linguistic descriptions.

Sheaves and linguistic descriptions of continuous maps:

Definition. –

- (i) A (set-valued) presheaf on a topological space T is a map P
- $$\left\{ \begin{array}{l} \text{open subset } (U \subseteq T) \mapsto \text{set } P(U), \\ \text{inclusion } (V \subseteq U) \mapsto \text{restriction map } P(U) \rightarrow P(V), \end{array} \right.$$
- compatible with all compositions $W \subseteq V \subseteq U$.
- (ii) A (set-valued) sheaf F is a presheaf such that, for any covering $U = \bigcup_{i \in I} V_i$, the restriction maps induce a bijection
- $$F(U) \xrightarrow{\sim} \left\{ (f_i \in F(V_i))_{i \in I} \mid \begin{array}{l} \text{restriction of } F_i \text{ to } W = \text{restriction of } f_j \text{ to } W \\ \text{for any } W \text{ with } W \subseteq V_i, W \subseteq V_j \end{array} \right\}.$$

Example. – The “terminal” sheaf $1 : U \mapsto \{\bullet\}$ = set with a unique element.

Proposition. – Suppose X is a “sober” topological space, endowed with a base \mathcal{B} . Then continuous maps $T \rightarrow X$ correspond to maps

$$U \longmapsto MU = \text{subsheaf of the sheaf } 1 \text{ on } T$$

\parallel

element of \mathcal{B}

such that for any elements $U, U', (V_i)_{i \in I}$ of \mathcal{B} verifying $U \cap U' = \bigcup_{i \in I} V_i$, we have

$$MU \cap MU' = \bigcup_{i \in I} MV_i.$$

The most general abstract context for the notion of sheaf:

Definition (Grothendieck). – Consider a “site” (\mathcal{C}, J) consisting in

$$\mathcal{C} = \text{category} = \left\{ \begin{array}{l} \text{collection } \text{Ob}(\mathcal{C}) \text{ of “objects”,} \\ \text{set } \text{Hom}(X, Y) \text{ of “morphisms” } X \rightarrow Y \text{ between objects } X, Y, \\ \text{associative composition law of morphisms} \\ (X \xrightarrow{f} Y, Y \xrightarrow{g} Z) \longmapsto (X \xrightarrow{g \circ f} Z), \\ \text{unit morphisms } \text{id}_X : X \rightarrow X, \end{array} \right.$$

$J = \text{“topology” on } \mathcal{C}$

$$= \left\{ \begin{array}{l} \text{collection of families of morphisms on the same target } X \\ (X_i \xrightarrow{u_i} X)_{i \in I} \\ \text{which are called “coverings” of the objects } X \text{ of } \mathcal{C}. \end{array} \right.$$

Then:

- (i) A (set-valued) presheaf on \mathcal{C} is a map P objects X of $\mathcal{C} \mapsto \underline{\text{set}} P(X)$,
 morphisms $(X \xrightarrow{u} Y) \mapsto$ “change of parameters” map
 $P(u) : P(Y) \rightarrow P(X)$,
 compatible with all compositions $X \xrightarrow{u} Y \xrightarrow{v} Z$.
- (ii) A sheaf F is a presheaf such that, for any covering
 $(X_i \xrightarrow{u_i} X)_{i \in I}$
 the “change of parameters” maps induce a bijection

$$F(X) \xrightarrow{\sim} \left\{ (f_i \in F(X_i))_{i \in I} \left| \begin{array}{l} \text{for any “commutative” square of morphisms of } \mathcal{C} \\ \begin{array}{ccc} W & \xrightarrow{w} & X_i \\ w' \downarrow & & \downarrow u_i \\ X_j & \xrightarrow{u_j} & X \end{array} & \text{with } u_i \circ w = u_j \circ w', \\ \text{we have } F(w)(f_i) = F(w')(f_j). \end{array} \right. \right\}$$

Categories of sheaves:

Proposition. –

(i) For any category \mathcal{C} , presheaves on \mathcal{C} make up a category

$$\widehat{\mathcal{C}}.$$

It is endowed with the “Yoneda embedding”

$$\mathcal{C} \xrightarrow{y} \widehat{\mathcal{C}}$$

$$X \mapsto y(X) = \begin{cases} Y & \mapsto \text{Hom}(Y, X), \\ (Y_1 \xrightarrow{u} Y_2) & \mapsto (\text{Hom}(Y_2, X) \xrightarrow{\bullet \circ u} \text{Hom}(Y_1, X)). \end{cases}$$

(ii) For any topology J on \mathcal{C} , J -sheaves make up a subcategory

$$\widehat{\mathcal{C}}_J \hookrightarrow \widehat{\mathcal{C}}.$$

It is endowed with a natural “sheafification” operation

$$\widehat{\mathcal{C}} \longrightarrow \widehat{\mathcal{C}}_J$$

such that the composite $\widehat{\mathcal{C}}_J \hookrightarrow \widehat{\mathcal{C}} \rightarrow \widehat{\mathcal{C}}_J$ is identity.

(iii) The composite

$$\ell : \mathcal{C} \xrightarrow{y} \widehat{\mathcal{C}} \xrightarrow{\text{sheafification}} \widehat{\mathcal{C}}_J$$

allows to see $\widehat{\mathcal{C}}_J$ as some kind of “completion” of \mathcal{C} .

It depends on the choice of the topology J on \mathcal{C} .

Toposes and their sketches:

The construction $(\mathcal{C}, J) \mapsto \widehat{\mathcal{C}}_J$ provides an extremely general mathematical framework for the natural human mind operations to interpolate from a given amount of information (here \mathcal{C}) using inference rules (here J).

One can go in the other direction:

Definition. –

- (i) A topos is a category which is equivalent to some $\widehat{\mathcal{C}}_J$.
- (ii) A geometric presentation of a topos \mathcal{E} is a pair (\mathcal{C}, J) endowed with an equivalence $\widehat{\mathcal{C}}_J \xrightarrow{\sim} \mathcal{E}$.

Remarks. –

- Such a (\mathcal{C}, J) can be seen as a sketch of \mathcal{E} .
- Any topos \mathcal{E} has infinitely many presentation sketches (\mathcal{C}, J) .
- In order to draw a sketch of a given topos \mathcal{E} , it suffices to choose (i.e. to name) a “rich enough” family of “objects” of \mathcal{E} . Then the
$$\left\{ \begin{array}{l} \text{morphisms between pairs of objects} \\ \text{composition law of these morphisms} \\ \text{topology } J \end{array} \right.$$
 will be induced from \mathcal{E} .

Morphisms of toposes and their linguistic descriptions:

Fact. – One can define a general notion of “morphism” of toposes $\mathcal{E}' \xrightarrow{f} \mathcal{E}$ such that, if $\begin{cases} \mathcal{E} = \text{topos of sheaves on a sober topological space } X, \\ \mathcal{E}' = \text{topos of sheaves on a topological space } T, \end{cases}$ they correspond to continuous maps $f : T \rightarrow X$.

Remarks. –

- In general, the morphisms $\mathcal{E}' \rightarrow \mathcal{E}$ between two toposes $\mathcal{E}', \mathcal{E}$ make up a category $\text{Geom}(\mathcal{E}', \mathcal{E})$.
- As $\text{Set} = \text{category of sheaves on } \{\bullet\}$, one defines for any topos \mathcal{E} $\text{Geom}(\text{Set}, \mathcal{E}) = \text{pt}(\mathcal{E}) = \text{“category of points” of } \mathcal{E}$.

Theorem (Diaconescu). – One can associate to any (\mathcal{C}, J) a “first-order (geometric) theory” $\mathbb{T}_{\mathcal{C}, J}$ such that:

- $\text{pt}(\widehat{\mathcal{C}}_J)$ naturally identifies with $\mathbb{T}_{\mathcal{C}, J\text{-mod}}(\text{Set}) = \text{category of set-valued models of } \mathbb{T}_{\mathcal{C}, J}$,
- more generally, for any topos \mathcal{E} , $\text{Geom}(\mathcal{E}, \widehat{\mathcal{C}}_J)$ naturally identifies with $\mathbb{T}_{\mathcal{C}, J\text{-mod}}(\mathcal{E}) = \text{category of } \mathcal{E}\text{-valued models of } \mathbb{T}_{\mathcal{C}, J}$.

Remark. – The vocabulary of the theory $\mathbb{T}_{\mathcal{C}, J}$ consists in the names of the objects and of the morphisms in \mathcal{C} .

From language descriptions to geometry:

Theorem (Grothendieck, Hakim, Lawvere, Joyal, Makkai, Reyes, ...). –

Let \mathbb{T} be an arbitrary “first-order (geometric) theory”.

Then there exists a topos $\mathcal{E}_{\mathbb{T}}$ (unique up to equivalence) such that:

- (i) for any topos \mathcal{E} , the category of \mathcal{E} -valued models of \mathbb{T}
 $\mathbb{T}\text{-mod}(\mathcal{E})$ identifies with $\text{Geom}(\mathcal{E}, \mathcal{E}_{\mathbb{T}})$,
- (ii) in particular, the category of set-valued models of \mathbb{T}
 $\mathbb{T}\text{-mod}(\text{Set})$ identifies with $\text{pt}(\mathcal{E}_{\mathbb{T}})$.

Remarks. –

- $\mathcal{E}_{\mathbb{T}}$ is called the “classifying topos” of \mathbb{T} .
It is a geometric incarnation of the semantics of \mathbb{T} .
- For any topos \mathcal{E} , there are infinitely many theories \mathbb{T} such that
$$\mathcal{E} \cong \mathcal{E}_{\mathbb{T}},$$

i.e. infinitely many language descriptions of \mathcal{E} .
- The topos $\mathcal{E}_{\mathbb{T}}$ may be constructed as
 $\mathcal{E}_{\mathbb{T}} = \text{topos of sheaves on } (\mathcal{C}_{\mathbb{T}}, \mathcal{J}_{\mathbb{T}})$ where

$\left\{ \begin{array}{l} \mathcal{C}_{\mathbb{T}} = \text{category whose objects and morphisms are formulas in the vocabulary of } \mathbb{T}, \\ \mathcal{J}_{\mathbb{T}} = \text{topology defined by the axioms of } \mathbb{T}. \end{array} \right.$

Consequence for images:

- Any language description of an image I formulated as a “first-order geometric theory” \mathbb{T}_I would define a topos $\mathcal{E}_{\mathbb{T}_I}$ incarnating the interpretation of the image I as a geometric being.
- Such a topos $\mathcal{E}_{\mathbb{T}_I}$ of an image I would also be presentable in the form

$$\mathcal{E}_{\mathbb{T}_I} \cong \widehat{(\mathcal{C}_I)_{J_I}}$$

where the pair (\mathcal{C}_I, J_I) would be a schematic drawing of the image I :

- the objects of \mathcal{C}_I could be pieces of the image I that can be given a name,
 - the morphisms of \mathcal{C}_I could consist in relations between these pieces,
 - the topology J_I could formalize decomposition rules of large complex pieces of I into smaller more elementary pieces.
- More quantitative descriptions of images (e.g. with coordinates) could be incarnated in extra structures, such as inner rings of the toposes $\mathcal{E}_{\mathbb{T}_I}$.

Consequence for the classification of images:

- To look for a theory (or for theories) of images formulated in “first-order geometric logic” is equivalent to look for a “space of images” in the (extremely general) form of a topos.
- For such a theory \mathbb{T}_{im} of images, particular images would appear as set-valued “models” of \mathbb{T}_{im} or, equivalently, as points of the “classifying topos” $\mathcal{E}_{\mathbb{T}_{\text{im}}}$ of \mathbb{T}_{im} .
- The topos associated to a particular image I could be the smallest subtopos of $\mathcal{E}_{\mathbb{T}_{\text{im}}}$ which contains the point corresponding to I . As a consequence of Caramello’s “duality theorem”, this would be the classifying topos of the theory \mathbb{T}_I deduced from the theory of images \mathbb{T}_{im} by adding to the axioms of \mathbb{T}_{im} all the properties that can be expressed in the language of \mathbb{T}_{im} and which are verified by the particular image I .