22. Abstraction versus Realism: Not the Real Question

Jörg R.J. Schirra and Martin Scholz

When browsing through a book on computer graphics, one. usually finds a lot of more or less interesting pictures that are produced by means of computers. These pictures are embedded in pages of technical texts describing how this image generation was performed and why it provides a better way to do so than other methods. Less space is usually given to the methodological background and the motivation that base the occupation with computer visualizations. In this chapter, we want to add to the more technically oriented rest of this book some reflections as to why such techniques can be interesting not only for computer graphics researchers, and where, from a communication-theoretic point of view, they might be of use in our society.

22.1. The Naïve Opposition of Abstraction and Realism

As was already mentioned in Chapter 1, most approaches in model-based computer visualizations carry a more or less implicit dedication to naturalism, often named the ideal of "photorealism". Realism is seen as the ultimate goal of any effort of creating pictorial representations of reality as it is or could be, and naturalism as the single stylistic method to achieve realism. In photography, for example, realism is seemingly guaranteed by a particular kind of causal relation holding between the scene represented and the image produced. By means of this causal relation, any aspect relevant for realistic representation seems to be automatically covered.

Abstraction appears in contrast as the intentional omitting of aspects that would have to be present for a realistic representation; or, in an extended sense, also as the pictorial representation of aspects that are actually not visible at all. Some applications, however, call for images that include some abstraction, images that are usually seen as less realistic in consequence. For example in a textbook on anatomy, what is mostly used is not a photo, which could be taken quite easily, but a sketch that has been produced with much effort (cf. also Section 1.1.1): The effort is not only a manual one, but essentially one to decide whether the aspects to be communicated can really be "read" from the sketch in an unambiguous, clear manner that does simultaneously fit to the variants in appearance the anatomist might see in reality. The sketch is to refer to all instances of what it shows by means of a typical instance that is stripped to the essentials: this "stripping" is far from being trivial.

In the following, we refer essentially to purposeful pictures. Interrogating the fine arts in this respect is less fertile since, first, such an investigation would have to refer to the same basic distinctions as discussed here. Second, those distinctions are, however, played with in the fine arts in a complicated manner and on much more reflected levels that are seldom involved in computer visualization. The central theme of the American 'fotorealism' of the 1960's and 70's, for example, is an indirect critique of the visual access to reality in the modern industrial societies: an access that is almost totally mediated by technical reproductions, and thus open to all kinds of hidden manipulations (cf. [Hel75]). The images of artists, like Close, Bell, Morley, do not try to show reality in a photo-like realism; their subject is the mediated access to what is believed to be reality by media that are assumed to present realistically. The classical "abstract art" — of Kandinsky, Duchamps, Arp, Grosz, to name just a few — can be understood as a reflection about the reference relation of pictures, too: a reaction on the earlier, academized forms of naturalism. People interested in computer visualizations usually have less ambitious goals: their field is basically the one of "useful graphics" with more or less straight-forward purposes.

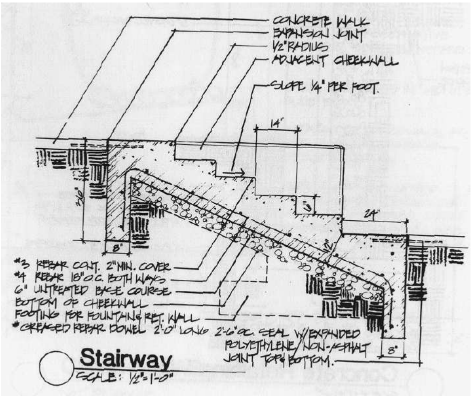
In the context of the traditional view of a strict opposition, we have to ask whether or in what degree abstraction and realism exclude each other mutually; or alternatively, whether they mark two different dimensions of pictorial information presentation, and hence may co-occur. A clear analysis of the relation between abstracted and realistic depiction is to provide essential hints and restrictions for the construction and application of useful computational visualization methods.

22.2. Three Examples of Purposeful Pictures

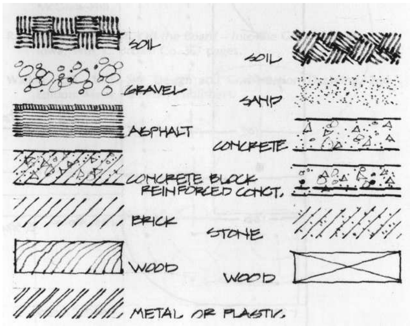
Let us have a look at some typical contexts of using purposeful images in our culture: architecture, engineering, and advertising:

First example: architecture

As a rule, the images drawn by architects have two different purposes (at least): either to give the client an impression of the object under planning by means of an *architectural sketch* (cf. again Fig. **??** on page **??** (??)); or as a draft used for the realisation, i.e., to impart the correct data to



(a) **Construction plan for a stairway**



(b) **Conventions used in (a)**

Figure 22.1.: Images in architecture: Examples of construction plans

the constructional engineers by means of an *architect's plan*, a *construction drawing*, etc. (cf. Fig. 22.1). While architectural sketches are used as a relatively unrestricted medium of communication between the architect and his client (drawn ideas, so to speak), building plans are legally binding for everybody involved. The purposes clearly affect the kind of rendering:

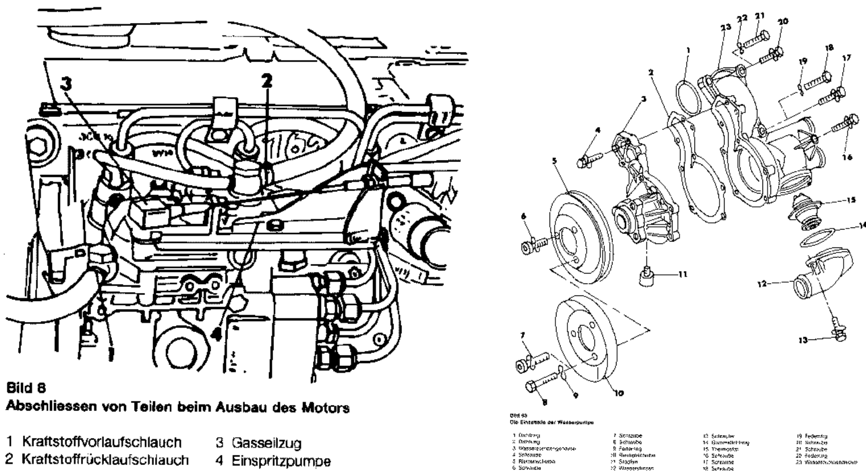
- Architectural sketches are notorioulsy changed and redrawn, starting from the first rough lines up to the final refined version. Apart from the fact that a refinement takes place, no generally applicable rules can be ascertained for the way of representation to be used. The final version, then, bases the rendering of a building plan.

- In the building plan, strict representational conventions between the users have to be applied.[1] The single tasks of representation are again devided into clearly defined drawing typs, e.g., overview, sectional view, site plan, ground plan, etc.

Thus, although strong conventions of representation may be useful, they do not yet exist in an explicit form for architectural sketches. On the other hand, the building plan is a legally binding basis for the realisation exactly because those conventions of representation are codified as a result of the necessity of clear preconditions for realizing a building, the educational standard in our society, and the tradition of this trade.

[1] For example, the kind of lines is regulated by the German Industrial Norm DIN 15, the kind of specifying the measurements on a blueprint in DIN 1356, the legibility of the numbers, the uniformity of the scale, and the arrangement of the views in DIN 6.

Second example: repair instructions in engineering
In engineering, two major tasks of pictures used in repair instructions can be observed: one is to help identify objects in complex environments that are difficult to comprehend or overlook. Or the function of an object or part is to be clarified in dependence of the spatial configuration of the whole system. Designers of repair instructions follow certain heuristics:



(a) **Identifying objects in their context** (b) **A *blowing-up* sketch**

Figure 22.2.: Pictures in repair instructions

- A naturalistic representation like a photo of an engine with lots of parts does usually not enhance the clear and easy identification of objects in a complex environment. A schematic representation of the objects' places and their rough forms is often enough: all irrelevant aspects that may take someone's mind off are eliminated (cf. Fig. 22.2). The point of view is determined by the perceptual situation of the one who is to perform the actual repair: he must be able to easily map the sketch onto his perception.
- For the proper identification, the parts or objects are labeled with numbers or letters that are explained by noun phrases in a legend (cf. Fig. 22.2 (a)). The text of a repair instruction refers to the arrangement of the parts in the representations by using directly those noun phrases.
- For clarifying the function of a part of an engine, documents for repair and operation include isometric sketches and illustrations of objects

simultaneously in two ways: on the one hand, the objects are shown in isolation; on the other hand, the spatial relations between the parts are illustrated with the help of *blowing-up sketches* where the single parts are connected by dash lines (cf. Fig. 22.2 (b)). Understanding these pictures presupposes a lot of knowledge about those parts and their typical shapes. The point of view must be carefully determined in order to present the typical outlines, not irritating the observer by an unusual view of an object.

Engineering has to use some conventions of representation when functional connections are to be shown. If the pictorial representation is employed to make clear the function of the objects within the whole system, it is not important how exactly the objects look like, i.e., what color or texture their surfaces have, or which light produces what reflections. Those aspects can be abstracted — even to such a degree that mere icons ('pictograms') are used (as in electrical engineering, e.g., the symbols for a condensator or a transistor).

Third example: product advertising
As a third example, purposeful *photos* are examined, i.e., a kind of picture usually rated naturalistic in contrast to the former examples. Typical situations for using purposeful photos are the realistic representations of products in advertising or on the outside of the sales package, e.g., the photo of a coffee maker.[2] The task of the photos of a product in this context is (a) promoting sales of the object depicted, and (b) cultivating the image of the trade name. Both purposes are usually intertwined and show in the presentation:

- The photos printed on the packages of consumer goods are standardized by the specification of the particular line of products given by the manufacturer. All products of one such product line are exemplified by one common 'image' regardless of actual differences in form, size, color or function of the concrete object. Thus, a type is constituted that allows the customer to distinguish the products of that producer from comparable appliances of competitors.

- A photographic image is conceived of as the only presentation achieving sufficient quality (with relatively low costs) for showing clearly the colors, forms, and features that distinguish this product from its rivals.

- Over and above showing the pure visual appearance, the photographic naturalism gives the observer two important pieces of information. The photographic image is read as a proof of the existence of the object shown. And the image is seen as a first visual indication of the functional potential of the product. The latter point is visualized by emphasizing the differences in form between this product and other products.

[2]Similar considerations hold, by the way, for portraits as used by casting agencies.

- The photographer can chose among several types of parameters to enhance the customers' ability to "read" the picture: for example, he or she may set a clearly directed light and use the resulting shades to demonstrate the spatial extension of the object in question — a wrong or illogical shadow destroys the intention of the image. Reflections can be employed for visualizing the condition of the surfaces; the addition of well-known objects into the set — e.g., a cup, a book, etc. — helps to clarify the proportions of the product, as well as its purpose.

The communication of the correct spatial extension of an object is only one special case of pictorial information: in such purposeful photos, the mental representations of object types and of the producer's company image or product line are to be evoked by means of the typical visual appearance of an individual object. Moreover, to a significant degree the latter is originally constituted by such photos.

In general, the producers of pictures have tried for a long time to create "realistic" and "naturalistic" images, for example in photography. But from the artists' point of view, the intention guiding this kind of "realism" was rather to elaborate a "construction of (our perception of) reality" than to "represent reality *per se*", as can be clearly observed in the development of the perspective in the Renaissance (cf., e.g., [KL89]) As images were used in mass media more and more inflationarily, the creators, who were no longer artists but designers, technicians, and marketing specialists, ignored the restrictions inherent to images. Though in the end, we have to realize that the communicative intention attempted originally, namely to present "the reality" and the essence of the objects as given by means of a naturalistic representation, has failed — indeed had to fail due to the very character of pictures and their role in communication.

22.3. Several Kinds of Realism

Speaking of realism in singular seems to induce that there is exactly one kind of realism, the climax of which is reached in photography and similar representations (cf. Fig. 22.3).[3] But why should we rate examples, like Figures 22.4 and 22.5, as being less realistic? Certainly, visual aspects have been omitted, like color, or added, like lines as the edges of objects. But those pictures can very well serve a lot of purposes related to corresponding real scenes: for instance, recognizing the objects represented (Otto v. Guericke's vacuum experiment of 1659) or classifying something as of the same type (a bison).[4] The preceding section gave some further examples.

[3] Actually, a color photo was meant to be used here; we resigned from that plan due to reasons of technical effort, now showing a digitized black-and-white version of a photo as a mere hint to full photographic naturalism.

[4] We do not state that the function of the paleolithic painting given was to classify; but at least, it could be used to that purpose, as well.

Figure 22.3.: Example Photography

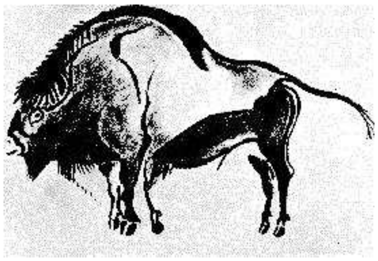


Figure 22.4.: Paleolithic drawing of a bison [Altamira, Spain]

Figure 22.5.: Engraving illustrating one of v. Guericke's vacuum experiments

Putting it simply, 'realism', as we understand the expression here, is the property of a representation to give an impression of a configuration of spatial objects that is or could be in the world. 'Naturalism' in our sense refers to the quality of a pictorial representation to evoke a visual impression as close as possible to the one of the scene depicted. The naïve opinion is that realistic images have to be naturalistic. In a reverse deduction, it was often assumed that non-naturalistic pictures are therefore not realistic, nurturing the traditional opposition of the latter to abstraction.[5]

Figures 22.4 and 22.5 indicate, however, that there is a (potentially unlimited) number of distinct kinds of realism; only some of them are naturalistic presentations, as well. The clue to the choice must basically lie in the particular communicative intention and the restrictions of the situation at hand. Where do computational visualizations fit in here?

The construction process of the traditional "photorealistic" model-based computer visualization consists essentially of two phases:

1. providing a three-dimensional geometric model (modeling)
2. projection of the model onto a two-dimensional image plane (rendering).

[5] There are, of course, a number of different determinations of the meaning of 'realism' and 'naturalism' dependig on the context of discussion (e.g., in literature or epistemology). Those used here are particularly fertile for considerations of purposeful pictures.

In the ordinary understanding, the projection phase imitates photography. The geometric model stands for the photographic motif "in reality." Pictures produced in that manner can be easily changed and redone. In contrast to manipulations of the real world, changes in the geometric model or projection parameters are essentially reversible without problems: an advantage over and above photography that makes computational visualistics attractive for the production of naturalistic quasi-photos, as it seems (or even of artificial cinema). Though, we have no hint so far that naturalism is the only possible goal, or that it can be achieved particularly well in contrast to other styles by model-based computational visualization.

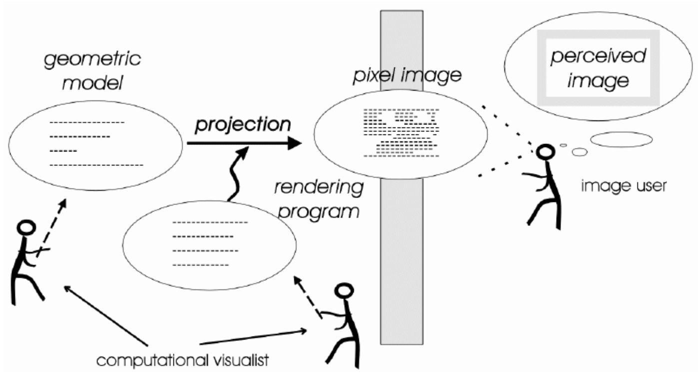


Figure 22.6.: Phases and interfaces of standard model-based computer graphics

In fact, the geometric model is a (formalized) description based on a data structure that allows the computer scientist to describe three-dimensional geometric objects. In contrast to a real object, the object model, i.e., the description of the object's geometric and optical properties, already includes an abstraction; it concentrates on certain aspects of the object described. Its descriptive character is also the reason for the model to be changeable that easily and reversibly. From that description, the projection creates another description on the basis of a data structure that allows us to represent two-dimensional matrices of points with color attributes. It is a certain presentation of that second description that, finally, can be perceived as an image (cf. Fig. 22.6). The projection algorithm, too, is primarily given as a description of what is to do in which cases, thus focusing on merely those aspects of the pseudo-optical projection that were rated as relevant. Again, it can be changed so easily because it is originally just a description.

The abstractions underlying the model and the projection method suggest once more that the relation between realism and abstraction cannot be a pure opposition. Abstracting as we understand it in this book is the process by which an extract of all the information available for some theme or scenario is refined so as to reflect the importance of certain aspects for the communicative situation at hand (cf. again section **??**). Descriptions of the geometric and optical aspects of a scene are just one kind of information that can be considered by computational visualists, and this type of information is as abstract as any other description. Moreover, constructing the model could focus on different aspects as well, and variants of the rendering program could "translate" those aspects into visible features of pixel images.

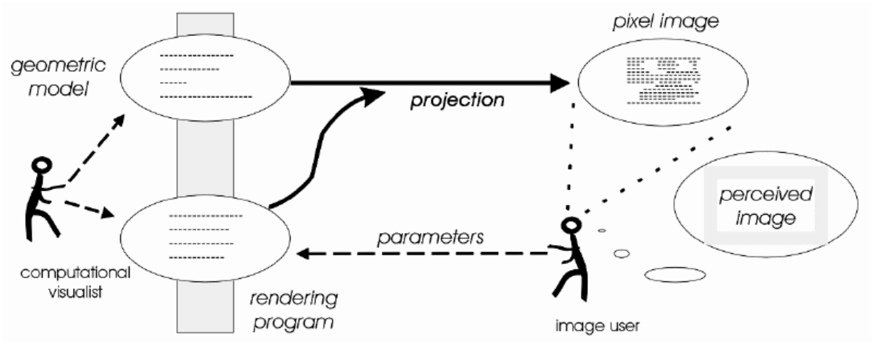


Figure 22.7.: Phases and interfaces of tele-rendering

When dealing with computational visualisations in interactive systems, the schema of Fig. 22.6 has to be changed in a particular manner without altering the characterizations of the data structures given above: the picture is in fact produced by means of the rendering algorithm at some point in time and place apart from the original image creator (cf. Fig. 22.7). Again, it is the descriptive nature of geometric model and rendering program that makes possible this *tele-rendering*, as we may call the situational separation of the computational visualist's designing activities and the actual image production finally induced by the image user. However, we have to expect consequences on the communicative function of the picture created by tele-rendering.

In the light of these considerations, we investigate the following theses:

Thesis 1: Realism and abstraction are not incompatible; they are relatively independent dimensions of the characterization of a picture; in particular, realistic images include abstractions, and vice versa.

Thesis 2: Naturalism provides merely one form of realistic presentation; it is an abstracted one, as well: the abstraction in it focuses on the aspect of visual appearance of a scene. Other aspects necessary for the interpretation of such an image as realistic are not represented and appear only indirectly as inferred.

Thesis 3: Using naturalistic representation for the communication of other aspects (in particular: non-visual aspects like function) is dysfunctional:

> the focus of attention is directed to the visual appearance instead of the aspects to be actually communicated. The bodily basis of those non-visual aspects must not form the foreground of the presentation; it has merely to serve as an anchor point for the communicative intention.

In the following three sections, we examine the communicative conditions of images with respect to the three theses, and add particular conclusions as for the case of tele-rendering.

22.4. Images as Signs: Considerations from Communication Theory

Any serious effort of discussing our theses has to consider that images are signs, i.e., are used in communication. Determinations of abstraction or realism depend essentially on that characterization:

> A sign is a material object that is used in a communicative act by somebody in order to direct the awareness of somebody to something — the signified thing ('thing' in a broad sense). (cf. REF!!!)

An important consequence of this determination is that signs do not refer to anything by themselves. It is always the sign user who refers *with* them. This holds for images as a particular kind of signs as well; that is, we must not consider a picture alone, but have to include those who use the picture in our considerations.[6] To ask why a picture "works" is to ask why the communicative act works, i.e., why somebody can use the picture to communicate with somebody else. Obviously, this question is of particular interest in the case of tele-rendering, i.e., if the actual rendering process is performed autonomously by the computer, and the original image creator is not present any more, as in the ZOOM ILLUSTRATOR, for example (cf. Chapter 13). The computational visualist can be conceived in this case as providing indeed a large set of pictures together with options for the image user to chose one or the other. That the user in fact changes parameters of the image production within certain limits by his choice, does not primarily matter.

The link between abstracting and communicating already appeared at the very beginning of picture production, as the paleologist A. Leroi-Gourhan, for example, observed:

> *Apparently, paleolithic art comes off a real written language, so to speak, and follows a path on which, starting from the abstract, step by step the manners of representing shape and motion are*

[6] In particular, this has effects on the discussion of "resemblance" as the primary mechanism of reference for images, which we do not want to discuss further in this paper.

> *elaborated; and at the end of that path, art finds to realism* [in the sense of 'naturalism'], *and finally dies away.*
>
> [LG84, p. 243] (Translated by the first author)

Visual art and written language have — as far as we understand today — a common communicative beginning, and due to the intimate relation to (oral) language, that starting point is closer to abstraction than to naturalism. The relation between pictures and language is interesting for us also for another reason: How could we rate the quality of purposeful pictures, as those given in section 22.2? An obvious way is to use the verbal description of what purpose the image has, what is to be communicated by its use; that description would have to be compared with another verbal description of what an observer (image user) — possibly some kind of "standardized user" — can recognize as the picture's communicative function (*vulgo*: what it "says" to him).

In the following, we summarize some basics of verbal communication, and try to transfer them to pictorial communication. The fundamental element of descriptions is the utterance of a singular declarative sentence about concrete things, sentences like (S 1) and (S 2):

S 1 *The on/off switch is situated at the lower left backside of the device, close to the power line.*

S 2 *In Figure 22.3 on page 373, the person I mean is sitting below a large tree.*

Logically, singular assertive utterances are divided into several functional components (cf. [KL73] and [TW83]): one is called the involved set of singular terms, or *nominators*, which refer to some given (i.e., already mutually known) individual objects. Nominators, like 'the on/off switch', 'the device', 'the power line', and 'the person I mean' in the examples above, are used for identifying something already known to the other interlocutors that is to serve as an anchor point for further information. The second component is called the general term, or *predicator*. The function of predicators is to introduce a standard gauge with respect to which the objects considered are rated: a dimension of distinction to be communicated (i.e., not known before to the other interlocutors). In the examples above, we employ the predicators 'to be sitting below a large tree', and 'to be situated at the lower left backside of something and close to something'. Nominators and predicators cannot stand alone — they only make sense when combined in an assertion.[7]

[7] More precisely speaking: predicators and nominators are only defined in that context. One-word utterances of infants ('holophrases') belong to a different, more elementary category of communication, cf. [?]. In comparing the logical and the linguistic terms, the nominators correspond mainly to the definite noun phrases of the sentence; predicators are given mostly by means of the predicates (verbs, adjectives, adverbs, indefinite noun phrases, and prepositional phrases); however, logical and grammatical functions do not necessarily map one-to-one.

A *logical copula* is the third component of any assertion with the function of binding the nominators as arguments of the predicator, and, even more important, of performing the act of assertion.[8]

Where can we associate, after this short detour to the fundamental logic of assertions, the function of pictures? Are pictures equivalent to nominators, i.e., used as acts of reference in order to identify things? Are they equivalent to predicators, i.e., used to communicate dimensions of distinctions? Or are they equivalent to assertions, i.e., whole propositions, especially compound propositions (conjunctions), as those used in verbal descriptions of an image?

All three possibilities mentioned above are apparently realized: We can use a picture of the Golden Gate Bridge to identify that object to somebody and then utter "Built in four years and finished in 1937". We also can use a picture in a book for mushroom determination to tell the doctor what kind of toadstool was in the lethal ragout: "This one". And finally, a picture can be employed to tell a whole story, as in a comic, thus being equivalent to a proposition. Note that in the first two cases, it is not merely the image use that mediates the communicative act: a predicative act, or a nominative act respectively, must be added to complete the pictorial proposition.

We here run into a particular characteristic of images: they are open to interpretations and do not contain an unambiguous clue as to how to understand them; captions, for example, have to be employed to clarify the interpretation (cf. also Chapter 12). A picture showing a large red suspension bridge could be used to identify a certain single object, or the place or even the place at a particular time; or it may serve as a communicative tool for classifying suspension bridges in general, or red large suspension bridges in particular, or being red, etc. And in the right context, the very same image may even tell a story, like "It was a beautiful evening without fog, when they reached the Golden Gate, and on their bike, they crossed the bridge and finally entered the city of their dreams". Thus at first sight, a picture is in general used as a proposition — sometimes additional nominators or predicators are applied, and overwrite, as it seems, the corresponding component of the image.

The question is whether the three cases of use mentioned above are possibly derived, secondary applications of pictures. The "true nature" of pictorial representation, it seems, is not really grasped yet on that level. Indeed, that compound propositions can be used to describe what a picture shows does not already mean that this picture is logically equivalent to the proposition, although in some applications the relation to the proposition may dominate its

[8] cf. [KL73, I.4 and p. 90]. The logical copulae have to be distinguished from the linguistic copula, which essentially has the function of a non-specific verb binding adjectives as predicates to a subject; the logical copulae correspond to FREGE's '$\vdash$', but carry in addition a polarity correlated to truth values; cf. [Sch94a, Chap. 6]. Traditionally, two copulae are studied in logic, corresponding to the two sides of a (binary) distinction: one for ascribing the predicator to that set of nominators, and the other for denying it ("internal negation"); cf. also [SW75, p. 239].

use. When a verbal image description is produced, the image, more precisely speaking, takes the part of the referential context, it provides the situation to which the verbal description refers.

In fact, the analysis of the logical form of assertive utterances given above is not complete: assertions are necessarily uttered in a certain *situational context* that determines their meaning, and hence has to be conceived as a fourth functional component (cf. [Sch94a, Sect. 3.4]). The nominators in particular cannot be understood without the context: following Strawson's explications of the reference relation, a nominator does not simply represent an object by means of a one-to-one relation, i.e., like a name tag; but it *picks out* an object from a certain given finite contextual domain of discourse objects, therefore conceptually requiring a one-to-many relation (cf. [Str71, p. 17ff] and [Tug82, Sec. 21ff]). In an utterance, the context usually remains implicit, but as in example sentence S 2 above (p. 377), it can be explicitly mentioned: the expression 'in Figure 22.3 on page 373' refers to the (nested) context in which the rest of the proposition has to be interpreted.[9]

Two types of contexts have to be distinguished: the *lexical* context and the *referential* context. The lexical context is used in explanations of the connection between several assertions uttered in sequence: it is mainly viewed as the comprehension of the text communicated up to that moment. Understanding an assertion is conceived as transforming its context to the context that has to be used for any subsequent utterance (cf. [Kam90]). A novel is a perfect example for a fictional lexical context: the assertions of the text refer essentially to objects introduced earlier in the context of the novel.[10] The transformation of contexts by understanding an assertion is not simply the addition of the proposition uttered literally. The logical interaction of the new information with the contextual information changes the latter and leads to information not explicitly mentioned. There is a transputed message beside the one transmitted, as STROTHOTTE and STROTHOTTE have put it in [SS97, p. 98ff]. Any speaker has to be aware of these implicatures of what she utters (cf. [Gri74]): although she did not explicitly say that p her interlocutor has understood that she told him that p by what she said, and in the following will act and interpret correspondingly.

[9] More precisely speaking: it establishes a connection between the current context of the discourse and the context to be used to understand the proposition; cf. [Fau85]. Already the distinction between nominators and predicators indicates the existence of contexts, since the nominators' function, namely to identify objects already known as anchor points for a new dimension of distinction communicated by the predicators, is particularly needed in cases where the logical context differs from the actual situation of the interlocutors, i.e., when they speak of something not present, past, hypothetical or fictional (cf. [Tug82, Sec. 26.I]).

[10] A special kind of assertions can be used to introduce an object for the first time (e.g., existential assertions); we do not deal with them here any further, although they may play an important role for image interpretation, as well, since this function can be integrated in the general function of nominators as examined here (cf. [Tug82, Sec. 22]).

In contrast to lexical contexts, the referential context of an assertion appears in explanations that anchor propositions in "what they actually mean", i.e., reality (or perception of reality). With respect to its referential context, an assertion can be true or false: for example, one has to look in the world in order to determine whether an assertion that ascribes some color to some object is true or not. Therefore, questions of realism involve referential contexts. A sketchy analysis of referential contexts is given in the next section. At this point, it may suffice that the referential context most important for us is the visual context, i.e., what is visually perceived by the interlocutors. With respect to the immediate visual context, pictures, in particular naturalistic pictures, are similar to novels and their relation to the immediate dialog history: they are *fictional visual contexts*. Though, in contrast to novels they are referential, not lexical.

Due to spatial restrictions, we do not want to elaborate further the relation between referential and lexical contexts (cf. [Sch93] and [Sch94a, Sec. 5.4]). But it should be clear that the relation between an image and a verbal description of that image provides the essential information if our goal is to improve the communicative quality of a purposeful picture: the description of what is to be communicated. Unfortunately, this relation is not unique: only if conventions of pictorial representation are used that are sufficiently agreed upon and well-known, the contents of an image could be deciphered in a correct way by the viewer, that is, in the way intended by the producer. Therefore, the producers must be as clear as possible about what they want to communicate, and how their interlocutors in the pictorial communication act might interpret — and possibly misunderstand — the picture.

This is also true in the case of tele-rendering: then, techniques anticipating or tracing the end-user's actual interpetation of the image can be employed to adapt the tele-rendering's parameters. Mechanisms for controlling generalized fisheye presentations are prominent examples given in this book (cf. the corresponding entry in the subject index); information used primarily for the rendering, like that encoded by G-buffers, may also be useful for the image producer when he wants to take into consideration what a distant observer might see in a picture (cf. Chapter ?2? and "viewpoint descriptions" as in [SS97, Chapter 9]). In any case: it is the original image producer who has to decide what picture the end-user gets in which situation.

22.5. Abstraction in Realism

With this theoretical background, let us now investigate the aspects of abstractness that appear in realistic pictures.

The essential aspect of abstraction in a realistic depiction — a successful photo like Fig. 22.3, for example — becomes clear as soon as we ask what is actually seen there: spots of color? Or geometric entities? Or spatio-temporally extended material objects ('spatial objects' for short)? Or persons? etc. Although we usually do not deny the first two answers, the latter two are primarily expected. But how can that be? Spatial objects have more sides than are *de facto* depicted — therefore the technique of perspective is important for realistic depiction. And 'person' is a concept with a structure that is even much richer and more complicated: a lot of aspects of a person have to be omitted in the pictorial representation. That is: most features necessary to identify something as a spatial object (and even more so for a person) are actually not transmitted by the pictorial communication, but transputed in the sense of the STROTHOTTEs mentioned above: they are added by the observer when looking at and interpreting the picture. Since the step to spatial objects is the crucial one, we ignore in the following the case of persons and deal with pictures of humans only as instances of images of a particular kind of spatial objects.

What we usually see on/in [*sic!*] a realistic image — the image content — are spatial objects, i.e., entities that are quite clearly separated spatially and temporally from their surrounding; entities that have parts[11] and distinguishable visual aspects: one can view them from several perspectives, and thus, one can recognize the very same object even if it looks different at different times (from several viewpoints). However, there cannot be any doubt that the image shows basically a colored surface only. The expression covering the important dependencies between the image surface and the image content is *'object constitution'*: usually it refers to the "invention" of spatial, persistent, identifiable, and countable objects in the mental development of infants (cf. [Pia37] and [Tug82, Sect. 25]) — the growth of the ability of recognizing individuated objects in the colloquial meaning, that is; but it is also used in a more general sense for naming the relation between geometric Gestalts (behaving similar to shadows) and spatial objects in the proper sense.

The relation of object constitution may best be explained by means of the current understanding of vision in AI: we here present as an example the perception of spatial objects by motion as demonstrated by the systems ACTIONS and XTRACK.[12]

In most computational approaches to visual object recognition, the signal of a video camera, i.e., essentially a temporal sequence of matrices of intensity values similar to those pixel images mentioned in Section 22.3 (cf. also Fig. 22.6 on p. 374), stands at the beginning. On a general level, two structurally different phases of subsequent processing can be distinguished: in the lower phase, the primary data is processed "bottom up" (data-driven): the results depend essentially on the original data and Gestalt factors alone (cf. also [Kof35]).

[11] At least, they have geometric parts, like the left part, the middle part, etc., and material parts, i.e., they are distinguished from their material.

[12] cf. [Kol92]; an elaborated description and further references are to be found in [Sch94a, Chap. 9]. Considering motion in this context emphasizes particularly the crucial aspect of object constitution: that such an object is identified in several situational contexts where it may appear quite differently.

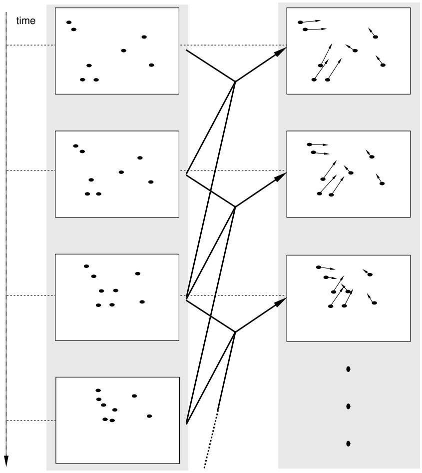


Figure 22.8.: First grouping

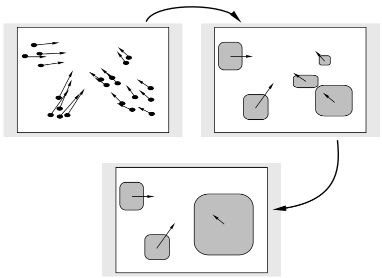


Figure 22.9.: Second grouping

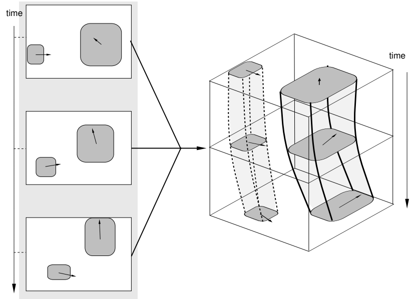


Figure 22.10.: Third grouping

The "bottom-up" part of ACTIONS: First, extrema of the intensity field at consecutive instants are grouped to instantaneous (velocity) vectors, if they are of the same kind and at almost the same position (Fig. 22.8); then, closely positioned similar vectors at one instant are grouped to spatially extended entities; if several of these still instantaneous entities happen to be close to each other and have similar velocity vectors, they are merged (Fig. 22.9); finally, another temporal grouping of the development of those entities is applied, resulting in entities that are also extended in time (Fig. 22.10).

In the higher phase, intermediate data is related to other sources of information, an integration, which is usually performed "top-down" (goal- or expectation-driven). On the lower level of our particular system (cf. Fig. 22.8 to 22.10), candidates of spatio-temporally extended objects are calculated by means of several layers of grouping criteria depending on similarity of the entities grouped (including spatial and temporal closeness).

However, these candidates are not yet spatial objects in the usual sense: this can be clearly seen in situations like that shown in Figure 22.11: do object candidates A and D belong to the same spatial object, or is it A and E that have to be associated? It is impossible by means of Gestalt factors alone to identify the corresponding object candidates before and after the "melting" — and thus, to find the correct two spatial objects. The pure "bottom up" grouping has to be complemented by additional "teleological" knowledge providing the conditions under which deformations, loss or exchange of parts and substance, occlusions, etc. do or do not alter the identity of an instance in the course of time.

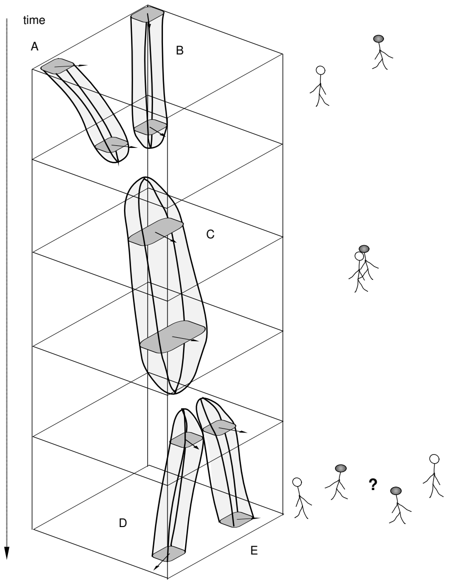


Figure 22.11.: Problem of identity for bottom-up processing: five candidates for two objects

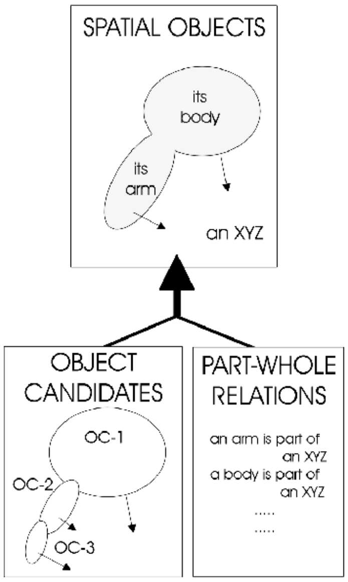


Figure 22.12.: Sketch of object constitution: relating perceptual geometric and functional part-whole relations

The established way to do this is using "object models": they essentially describe which geometric configurations of parts form an instance of a particular type of object.[13] The projection of the object models to the object candidates perceived finally establishes the perception of spatio-temporally extended, persistent, and localizable entities — exactly the type of objects involved in spatial descriptions and realistic pictures.[14] The largely geometric information about the *actual* configurations from the bottom-up phase is combined with information about part-whole relations governing the *possible* range of configurations (cf. Fig. 22.12). In general, the constructed descriptions of what is seen form distinct contexts on different levels; perception can be conceived of as the systematic relation between these contexts: the

[13] cf., e.g., [Mar82]. The geometric models used in model-based computer visualization are usually more closely related to the geometric descriptions resulting the bottom-up phase than to the object models in the sense of Marr: information about part-whole structures are only seldom integrated, the models are "flat" (cf. Chapter 3). It should be obvious that meronomical information, i.e., the hierarchies of model components in the geometric models of computer graphics, play an essential role in computer animation, cf. Chapter 15.

[14] This projection is, in a way, analogous to the projection described by the rendering program in model-based computer visualization, cf. again Fig. 22.6: here however, the projection is used to decide whether a given picture could or could not be the picture of a certain object model.

description of one level is used as the referential context for the description of the next higher level in the sense mentioned in the preceding section (cf. Fig. 22.13).

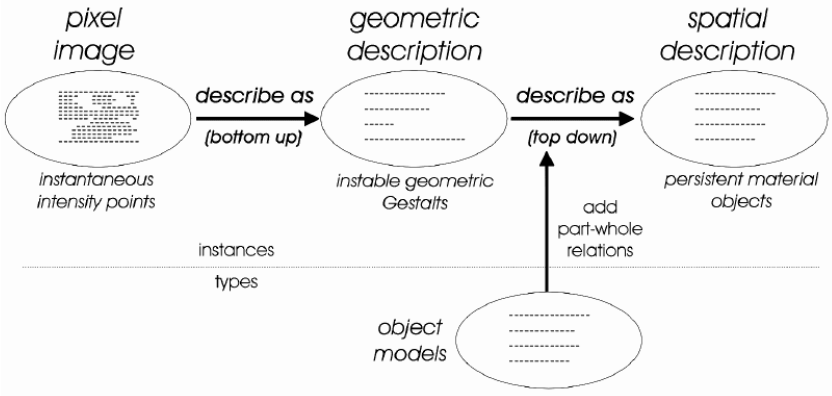


Figure 22.13.: Conceiving perception as a cascade of descriptions

It is an important logical characteristics of those calculi, that geometric Gestalt principles allow us to formulate descriptions that organize individual geometric shapes ('object candidates') within a coordinate system of locations of *one* perceptual situation. But it is not possible in general to associate those shapes as the same individual in different contexts: if an object candidate gets out of sight, it "dies away" and can never come back. If the corresponding spatial object again enters the field of sight, a completely new object candidate is born (cf. also [Vie91]). Among other effects, the transformation to a description based on spatial objects enables us to originally integrate several perceptual situations and their corresponding coordinate systems of places and times, and thus, to speak of persistent objects, at all, that is of objects that do not disappear when we do not look at them (cf. [Tug82, Sect. 26]).

After all, what an image contains primarily is, at best — i.e., reached merely data-driven — a configuration of visual Gestalts that may serve as the referential context for a description involving spatial objects: that is, the spatial objects are never contained explicitly. Any picture user has to infer (more or less implicitly) that spatial objects are meant at all with the configuration of light intensities presented. This is true for naturalistic images, as well. In contrast to other (i.e., non-naturalistic) realistic pictures, a naturalistic depiction concentrates particularly on the accuracy of the geometric/optical aspects of spatial objects, thus directing the focus of attention of any image user primarily to those aspects. Furthermore, an image user usually assumes in that case that this focusing is intentional: therefore he assumes, too, that it is exactly this aspect — the visual appearance of a configuration of objects — that the image producer wants to tell him.

Since we have determined realistic representations as representations that deal essentially with a configuration of spatial objects as it is or could be in the world, we now have to concede that realistic pictures indeed are never immediately realistic in that sense. They are linked to a (verbal) realistic representation by means of an abstraction: some crucial aspects of spatial objects that are not geometric are in fact omitted in any "realistic" image (and hidden surface removal is merely the tip of that iceberg, cf. also [Sch94a, Note 64]). Only the automatism of our perceptual capacities enables us to "undo" that abstraction and see there spatial objects, like trees or houses.

22.6. Realism in Abstraction

At this place, an interesting question arises: how much "geometric naturalism" (just for having an expression) is involved in the first four example pictures of section 22.2, pictures that are clearly rated to be abstract, i.e., architectural sketches, construction plans or drawings in repair instructions? How much of the visual appearance can (and should) be ommitted? And what exactly is the communicative function of the remaining naturalistic parts, a function that makes it necessary to keep those visual aspects in the depiction? This section presents only some first considerations concerning those questions.[15] In order to have a handy term in the following, let us call that rest of the visual appearance of te scene depicted 'the picture's *naturalistic residue.*'

Of course, the function of the naturalistic residue depends on the general communicative intention guiding the picture's use. In particular, it depends on what is to be actually communicated, and also on what the "sender" can do to ascertain that the "receiver" gets that message correctly? As a first observation in the light of the distinction between nominators and predicators in verbal communication as mentioned in section 22.4 we find that the naturalistic residue is essentially associated to something analogous to a nominative function: it anchors some new aspects to be shared in those aspects that are already known, for example, by means of immediate visual perception, as in the example of a repair instruction (cf. again Fig. 22.2 (a) on p. 370): such a picture is used in face of the corresponding real situation. The objects' identity with respect to the repair instruction is to be communicated. That information is given by means of the legende and labels anchored in a rather simplified sketch that shows essentially the very scene the reader sees. Without the reader's ability to match the sketch with the objects the picture would be not really helpful. However, it is not necessary that the corresponding predicative part of the picture is as abstract as in that example: the focus may be on merely one particular part of the visual appearance. Though in most cases, the picture is not immediately confronted

[15] We leave it to the reader to consider these questions for the other abstracted pictures in this book.

with what it is assumed to depict. The following list is a first collection of such cases without any pretension of completeness:

- if the **shape of one object** (already known otherwise) is to be communicated, the visual context of the object may be completely reduced: In some cases however, some clues for the situational setting must be included, e.g., a horizont line for the vertical orientation of the object. As for the representation of the object, the main outline from a certain (typical) perspective is often sufficient for communicating the shape.

- if the **spatial configuration** (of objects already introduced as theme of the conversation in some way) is the aspect in focus, representations of the objects forming that configuration is secondary and may be oppressed to a degree just enough for identifying the objects' identity, and thus serve as an anchor for the information to be actually transmitted. Often, as in a subway connection map, the objects forming the configuration (the subway stations in the example) are pictorially reduced to a mimimum, e.g., an icon, that is actually explained by some piece of text.

  In addition, a non-visual aspect may be closely tied by convention to the spatial configuration in certain communicative settings: a mechanical engineer may indeed "see" more or less effortlessly the function of some part of an engine just by its spatial position with respect to the whole configuration.

- if the **spatial distribution of one particular visual aspect** in some context is what the image producer wants to tranfer by means of the image, e.g., the places of red things in a certain scene, it is usually not sufficient to just put some red pigment at about the places of the objects in the image plane: the contours rather should be given as well, as an anchor for the interpetation of the color spots.

  Due to the very nature of our conception of spatial objects (which also forms the basis for colored objects in this context), the colors have to be spatially bound to an object carrying them. Again, the geometric aspects of spatial objects serve not as the main content of the message, but as a necessary condition for the communication of that content.

- similarly, the **numerical distribution of visual aspects** in some context could be concerned, e.g., the number of red things compared to the one of blue things in a certain scene; again, the naturalistic residue serves as an anchor, which is an idea related to the nominators of assertions.

- if the **spatial distribution of a non-visual attribute** in a scene is to be communicated, rudiments of the shapes of objects can carry the function of anchoring the distribution, again similar to the nominators' function in an assertion. Minimal requirement is that the image observer must be able to identify the objects in question: a precise repropduction of the visual appearance of those objects is not necessary, and may even interfer with the visual encoding (e.g., by color) of the non-visual attributes concerned.

A typical example is provided by communicating air temperature in a weather map; some geographic shapes are visible in order to identify the locations. Partially, the sketches in repair instructions for identifying the relevant objects as mentioned above also belong partially to that class.

Note that in the latter cases, it is possible to first show the reduced shapes, and then add the (possibly visually encoded) attributes in question: but it would usually not work the other way round: first presenting the mere attributes and then merging in the geometric base. In these cases, only the first type of presentation supports the interpretation as intended: an assertion with nominators that refer to objects already known in order to anchor new information communicated by predicators.

Obviously, spatial objects are a central category for a lot of non-visual aspects. By means of representations of the geometric component of spatial objects it is possible to communicate those other aspects with pictures using the particular scheme of organization inherent to geometric Gestalts — namely that they are integrated in the coordinate system of perception. There is still, quite obviously, lot of room for future research.

22.7. Abstraction *and* Realism: Conclusions for Computational Visualization

Where do we stand at the end of our examination?

First: literally each of the pictures presented in this book is in fact an example of a realistic depiction as we understand this term: it gives an impression of the configuration of spatial objects. Most of the pictures in this book are also the result of abstraction: namely the process of generating an extract of information according to the communicative situation.[16] They are, thus, illustrations of our *first thesis*: realism and abstraction in the senses considered here are not incompatible concepts (cf. p. 375).

Second: in most cases, the representation of the spatial configuration is used as the basic scheme of organization in which the most relevant information to be communicated is placed. Sometimes, the spatial configuration *per*

[16] In some cases (as in the photo in Fig. 22.3 on p. 373), it is not clear whether the extract of information was intentionally produced for a particular communicative situation; therefore, we hesitate to ascribe our concept of abstraction to all the pictures given in this book.

se is the content to be transferred to the image user; the spatial objects in that configuration are, then, depicted so that they could be identified — however usually not naturalistically. Only if the knowledge of the visual appearance of the objects is the central goal of the communication, the naturalistic style is appropriate. In this case, too, merely one aspect of spatial objects is the focus of an abstraction, others are ommited: this is our *thesis 2*.

Third: in the framework of intentional communication, *thesis 3* is in fact an immediate consequence of thesis 2: if a naturalistic presentation directs the focus of attention of an image user particularly to the visual appearance of spatial objects as stated by thesis 2, it should rather be avoided if other aspects are to be communicated, particularly if the aspects intended are not related to the visual appearances of the objects presented. Stylistic elements of naturalism should be used only to the extent in which aspects of the visual appearance are necessary for identifing those objects and their spatial configuration of which some other aspects are to be communicated.

Abstraction and realism, thus, are definitely not just opposite concepts, or form the end points of a more or less continuous linear scale: the naturalistic and non-naturalistic representations we have observed are simultaneously abstract and realistic. They form a loose structure of clusters that are more or less closely related to each other. Only if we understand 'abstract' in a *secondary* sense as 'non-naturalistic', we may project those clusters onto a linear scale, with the naturalistic pictures at the zero point, and the other forms more or less distant to that point. This is, of course, the more colloquial meaning of 'abstraction' with respect to pictures, but it is less productive in discussions on computer visualizations, as we have seen.

For somebody applying computational visualizations, basically two questions have to be answered in advance:

1. What aspects of the scene or configuration of objects have to be communicated, i.e., should be easily and clearly "readable" by the interlocutor / end-user?

2. How much of the naturalistic residue, i.e., the set of geometry and visual properties of the objects, has to be present in order to anchor sufficiently the (potentially non-visual) aspects to be originally communicated, without putting too much stress on the mere anchor?

The computer cannot and should not take over the task of determining what is important, what has to be communicated in a particular situation. Even in the case of tele-rendering, the computational visualist who has designed the program has to determinate under what circumstances the computer has to show which aspects as the important ones.[17] Of course, the

[17]Techniques like the application of a *degree of interest* (DOI), *focus points* (FP) or *aspect of interest* (AOI) as defined in Chapter 2 allow the computer visualist to have some influence of what a user of his system is going to see. Saying that the system choses by means of DOI etc. is only a shortened way of speaking.

machine also cannot add the creative potential of an artist, who succeeds to direct the attention of his public to some hitherto ignored aspects (be they visual or not) of something by means of an innovative form.

What computers can provide, and that essentially is, what computational visualists should concentrate on with their work, is the following:

1. a broad set of methods for rendering pictures that are not naturalistic;

2. rules for assisting the decision of how much of the geometry and the pure visual appearance is enough for the purposes at hand and the rendering methods applicable;

3. strategies for testing whether the communicative intention could be fulfilled by means of an image in the actual situation.

The latter point directs toward an analogy to listener models or, more generally, partner models as they are commonly used in systems for generating natural language utterances: an 'observer model'. As it is so easy to misinterpret pictorial communication, it is even more necessary to keep control of what potential observers in the context of the image observation will presumably understand immediately or only after several false tries. For the image producer, it is crucial to adopt the perspective of the potential image observer, which is usually not that easy since the background and the present context of action may be pretty different for those two. Algorithmic strategies will accordingly be hard to find out and implement.[18] FURNAS's technique of generalized fisheye presentations and their extensions and applications in this book (cf. [Fur86], and the subject index) can be conceived of as a simplified, but computationally more treatable approach. Of course, most interesting results are to be expected of those techniques for the enhancement of tele-rendering.

As for the second point mentioned, few solutions have been brought forward yet: at least, some effort has been taken for finding rules that suggest whether some part of a compound structure to be communicated is to be transferred as text, and which other part should be given as a picture (cf. Section 13.2). Rules governing the decision between several graphic styles were not considered. Though, designers have proposed heuristics for their work that could be adapted to computer visualizations, as well.

At last, inventing rendering algorithms for automatically producing corresponding images are certainly the task most directly associated to the work of computational visualists. Of course, enriching or differentiating the geometric models basing the computer rendering is one aspect of that task (as already mentioned in Chapter 1 and exemplified in Chapter 3). As for the

[18] As a matter of fact, one of us has earlier dealt with a listener model for generating spatial descriptions in German that involves the production of sketch-like graphics as an intermediate representation, cf. [Sch94a, Part II] or [Sch95a]; the 'observer model' projected here could, in a way, be seen as the inversion of that project.

rendering part in the light of the above considerations, the trend might be to divide that part of computer visualization in two phases: one is for generating the geometric basis of the objects' configuration as far as it is needed to serve as an anchor for the information to be mainly communicated (the nominatoric component of that picture, so to speak); and another one is for adding that latter information in an appropriate manner (in a way: the predicative component of the pictorial assertive act intended). Whereas the first phase is to find a minimum of representation for the aspects concerned, the second is to give a maximum correspondingly.

What we might expect from computer visualists, after all, is essentially a skillful toolbox, an *organon* of visualization: it contains as its core a set of methods to be combined for the actual production of an image or sequence of images. It also contains assistance for the correct and even an advantageous use of those methods. And finally, there ought to be available some instruments for rating the success. There is a vast field waiting.