Computer Image Processing: An Epistemological Aid in Scientific Investigation

Vincent Israel-Jost

Université Catholique de Louvain

In recent years, the production of scientific images has been changed in a drastic way by the combination of instruments delivering numeric data and computers. The latter permit scientists to display numeric data in image form, but also to develop a variety of computational procedures to enhance image data. This paper aims to discuss these practices of image processing and to defend several theses regarding the epistemological role of data processing, especially with respect to the interpretation of images produced by instruments.

1. Introduction

In many scientific fields, today's practices of empirical enquiry rely heavily on the production of images that display the investigated phenomena. And while scientific images of phenomena have been important for a long time, what is striking now is that scientists have found ways to visualize such widely different types of phenomena. In the past twenty or thirty years, we have become accustomed to seeing images of galaxies, of cells, of the human brain but also of blood flow or of turbulent fluids. This success in making images relies first on the set of imaging devices that have flourished in the twentieth century, instruments such as radio-telescopes, electronic microscopes, MRI and many more. Without these, many phenomena would be inaccessible to the human eye, either because they are too remote, too small, too deep, or because they can't be probed using visible light. Imaging, then, has become a transversal field in itself, the goal of which

is to provide scientists with the means to make images of even more phenomena or somewhat better images than previously produced ones.

In a way, using an instrument makes for a seemingly easier way to produce images, when this is compared to the pre-instrument era, when scientists would collaborate with artists who had to carry their vision while drawing nature. However, even though an instrument simplifies the production of images by automating it, there is much more to using an imaging instrument than just pressing a button. Researchers who carry out empirical investigations using images have to develop a variety of skills and to know much about the instrument that they use. This knowledge is important to ensure that the image is produced in an optimal way, using the right instrument settings and also optimizing the external conditions such as temperature, humidity, light, preparation of the investigated object, etc. This knowledge is also important later, when it comes to interpreting images. Indeed, with more and more instruments available, based on different technologies, phenomena can be pictured in ways that share very little with photography. Interpreting scientific images sometimes amounts to deciphering, based on what is known about the imaging technique and the particular instrument one is using. As a result, while imaging techniques have indeed made the production of images easier, they are also demanding for scientists if they are to be used correctly and to play a concrete and well-identified role in empirical enquiries. In fact, they might be so demanding for users as to create new epistemological problems. This is what I will argue in section 2, where I present the challenges imaging technologies pose to scientists, especially with respect to image interpretation. In section 3, I consider the main possible reactions to this problem, and I argue that, though serious, they can be overcome by enquiring *about the instrument*, using this acquired knowledge to better use instruments and better interpret the images that they produce.

In the remainder of the paper, I will discuss one particularly salient aspect of contemporary scientific imaging, which is the widespread use of digital data processing. Indeed, in addition to the instrument, imaging devices also rely on computers, not just as display devices but also as machines that permit data storage and retrieval as well as mathematical transformation. I will especially focus on the latter, and I will argue that the possibility to numerically process the data has a strong impact on both the production of images and on their interpretation. After introducing computer image processing in section 4 by presenting its general principle and a few examples of what can be achieved, I will defend several theses in the subsequent sections that aim to cast light on the ways that image processing helps scientists to keep up with the increasing amount of knowledge and skills that instrumented techniques require from them.

The first thesis, defended in section 5, is that processing algorithms embeds knowledge that is required for image interpretation and that the interpreter of the image does not have to possess this knowledge in great detail. This results in an economy of knowledge for interpretation by a human agent. In section 6, I focus on the tacit knowledge that is involved in scientific image interpretation, that is, the type of knowledge that scientists are required to possess but that is not, or is not easily, communicable. This is generally taken to be an undesirable feature of a scientific activity because it renders its results somewhat opaque and ungraspable by people who are not experts in the field. The thesis defended here is that image processing results in less tacit knowledge in scientific investigations because some tacit knowledge is turned into explicit knowledge. The third thesis (section 7) is twofold. I defend (i) that computer processing extends the computational capacities of investigators and (ii) that this extension of computational capacities also results in an extension of their observational capacities. Together, these theses form an argument in favor of the use of image processing as these new computational practices permit scientists to respond to the epistemological challenge posed by imaging technologies.

2. Imaging Instruments in the Digital Age and the Problem of Image Interpretation

Before focusing on image processing, it is important to first state what we expect from imaging instruments, namely, obtaining images from which we can reliably know about certain aspects of the imaged phenomenon. This will permit us to establish in the next sections what computer image processing does in order to help achieve this goal. Indeed, it appears that when an investigator (or a group of investigators) is using an instrument with no further computer treatment, it is not always very easy to conclude that such and such state of affairs is the case from looking at the various types of images that are produced nowadays. Some imaging instruments used in the right circumstances certainly provide researchers with images that are perfectly clear and that can be interpreted in a straightforward way. But there are also many difficulties that investigators can face in front of a scientific image, which can render the interpretation shaky, if not impossible. The aim of this section is to state what an imaging instrument is, what the general principles that underlie the imaging process are, and to identify the conditions that guarantee that knowledge can be obtained from images. Given that these conditions are not always met, this will lead to analyzing, in the remainder of the paper, the role played by image processing to supplement instruments.

Making images of a phenomenon always requires detection of a certain type of wave.[1] The waves that permit the making of images are often electromagnetic but mechanical waves such as ultrasound can also be used for that purpose, as is the case in echography, for instance.[2] The waves carry information about the object of interest, either because they interact with the object (they are reflected, diffused, absorbed, etc.) or because they are emitted by this object, as with stars that we see because they emit light (as well as other radiations). When the waves travel to us, we can detect them with the appropriate device. An imaging instrument is equipped with a detector of a type of wave, which registers events of detection with a certain spatial organization: an event of detection is associated with the location where it reached the detector plane. This is what permits formation of an image, as opposed to when events are simply counted, e.g., with a Geiger counter that does not give further spatial information. At this point, however, a condition is still missing for the image to be really informative about the phenomenon. Indeed, the spatial organization of the image is supposed to also relate to that of the phenomenon. For instance, a photograph shows a scene in the same spatial organization as that which we experience from the same viewpoint with our eyes. An instrument is not just a detector, it also selects only those events whose origin we can trace back in order to know about the phenomenon of interest. More, then, must be said about the conditions that guarantee that such relationship between image space and object space holds.

In his influential paper on observation, Shapere (1982) posed a condition on the transmission of information so that what is detected by the instrument can actually be linked to a precise origin in the object of interest. Comparing the trajectory of neutrinos with that of photons, both emitted from the core of the sun, Shapere noted that while the former travel in straight lines, due to the extremely weak chance of interaction with matter, the latter interact so much that for a photon to travel from the core to the surface of the sun will typically take 100,000 to 1,000,000 years. The temporal and spatial uncertainty at this point is such that all information is lost regarding the origin of the photon. By contrast,

1. This choice of words should not suggest that the wave aspects always play a significant role in the imaging process. Very often, images are formed from *particle* detection, though mechanical waves do not admit of another description, say in echography or elastography. With that in mind, I will often use the wave description in what follows.

2. Other radiations can be used as well. The case of the observation of the sun advocated in Shapere (1982) is based on the detection of neutrinos that are emitted from the core of the sun. Though Shapere doesn't describe this case as one involving images, it is nonetheless relevant to discuss one aspect of Shapere's analysis in a discussion about imaging instruments—see just below.

detected neutrinos have directly reached the detector and still carry information about their origin. Shapere's conditions of observation are then:

> (1) information is received (can be received) by an appropriate receptor; and (2) that information is (can be) transmitted directly, i.e., without interference, to the receptor from the entity x (which is the source of information).

I take Shapere's comparison between neutrinos and photons to be important from an epistemological point of view, since in one case we can know the origin of the radiation while in the other case we cannot. However, his condition that no interference should perturb the transmission of information is too strong, even if we change it with a probabilistic phrasing such as "the transmission of information should be direct, i.e., with only an extremely weak chance of interaction." Since what matters is that we can trace back an event of detection to the origin of the emission or interaction of the radiation in the investigated object, I argue that the direct path that Shapere takes to be necessary is only the simplest condition under which we can indeed locate the origin of the radiation. For instance, it doesn't really matter if light is deflected by, say, a mirror between an object and a roll of film inside a camera. It doesn't prevent us from obtaining an image that is in clear correspondence with a photographed scene or object. In order to characterize the successful detection that preserves the information about the source, my suggestion is that the way information is transmitted, rather than being direct, should simply be knowable.

To establish a knowable relationship between image and investigated object, it is not enough to focus on how information is transmitted between the object and the instrument. The instrument's design plays an important role in order to achieve the establishment of this relationship. Indeed, the instrument is not limited to a detector; it is also a device that aims to select waves so that the detected ones can in fact be associated with a precise enough location of their origin. I will give two examples of ways that this can be accomplished. The first example is the geometric selection of the incoming waves, as realized by example in the *camera obscura* with pinhole collimation. The pinhole—a small aperture at the entrance of the box—selects the angle of the incoming rays of light, so that a given point of the image only receives light from a certain direction. Without the pinhole, each point of the image would receive light coming from any direction and there would be no one-to-one correspondence between the image space and the surface of the imaged object. This one-to-one correspondence permits the observer to avoid the uncertainty as to the origin of the incoming light. The same idea lies behind the use of lenses for visible light, or of collimation for electromagnetic waves of higher energy.

My second example is that of the selection of gamma photons in scintigraphy according to their energy. In scintigraphy, a radioactive tracer is administered and takes a certain route in the organism. This tracer emits gamma photons of a characteristic energy, which can in turn be detected by a gamma-camera, revealing information about the inner structures of the organism and its physiology or possible pathologies. However, a number of photons are scattered by the tissues before they reach the detector and since this changes their direction in an unpredictable manner, these detected photons could not be associated with a precise location of their origin. Since the change of direction is correlated with a loss of energy, gamma-cameras are also equipped with an energy selection device, which permits experimenters to not count the detected photons whose energy is outside of the typical energy window of the radioactive tracer. As a consequence, the remaining signal is one that has been transmitted in approximately straight lines from the organism to the detector, and the resulting image is in much clearer correspondence with the source object.

In light of the previous analyses, I can now give a definition of an imaging instrument. An imaging instrument is a device that can detect a certain type of waves or radiations and that attributes a position to a detected event *on the image space* (the surface of the captor). For an instrumented technique based on such imaging instrument to be successful—that is, for images to have some representational value—it has to be possible to trace back the origin of the values attached to a position on the detector to more or less precise regions of the physical space where the object of interest lies. The key condition for this is that the waves that are detected have a trajectory between the object and the detector that is knowable. I will get back to this condition at the end of the current section but first, I will detail additional elements about the digital aspects of instruments.

In recent years major technological developments have transformed most imaging instruments that deliver spatial information in a radical way, turning them into digital instruments. A digital instrument is an instrument whose outcome is a list of numbers that can be saved in a computer file. In contemporary practices of scientific investigation, digital instruments have become as widespread as digital photography for non-scientific use. In a digital instrument, an electronic detector transforms an event of detection into an electric signal. The surface of the detector is decomposed into a discrete array so that each event of detection is associated with a small surface area. For example, taking a very rough detector of gamma rays whose surface is decomposed into a 2×2 array, if one obtains the ordered list of values (52, 34, 45, 37), the first value, 52, corresponds to the number of gamma photons that have been detected in the first surface area of the detector, say the upper-left, then 34 is the

number of gamma photons detected in the upper-right surface area, etc. A rudimentary 2×2 image can then be formed on a computer screen by attributing a color to each value. As I already stated with non-digital imaging instruments, it is crucial to establish a relationship between the features of the resulting image and the object of investigation so that the image reveals something about this object. For instance, it is not so much that 34 photons were detected in the upper-right corner of the detector and the resulting shade of grey that is interesting for scientists. They expect to relate this to a certain localized phenomenon.

There are several reasons for the success of digital instruments in contemporary empirical investigations. The instruments become easier to use, similarly to the way digital photography has greatly decreased the number of steps required to produce an image compared with analog photography. They are combined with computers and images are immediately displayed on screens instead of being printed. This goes faster, it reduces costs and permits the viewer to change certain display parameters very quickly and easily. For instance, the colors used to display the quantitative information can be modified using a number of predefined options that are called "lookup tables"; or the zoom factor and framing of the image can equally be changed immediately, with no effort. Then, image files can be stored, retrieved and circulated with much more ease than printed images. Finally and most importantly for the topic of this paper, the outcome of digital imaging instruments is a list of numeric values that can be manipulated with computer algorithms. Indeed, the numeric format allows for consideration of images as mathematical objects – lists or arrays of numbers, that is, vectors or matrices on which operations can be performed. It is this feature that I aim to explore in sections 4–7, as a response to the problem that emerges from the use of instruments that I now present.

I stated as a general condition for images to be epistemically valuable that the way information is transmitted (from the investigated object to the detector) should be knowable. This permits investigators to potentially trace back this information to its origin, and to associate features of the object with features of the image. This association is what I refer to as the interpretation process. However, this condition of knowability must be spelled out a bit. Indeed, scientists do not face just the extreme cases of no or little complexity of information transmission, as in a clear photograph, or of such complexity that the relationship between image and investigated object is impossible to figure out, as with photons interacting for more than 100,000 years inside the sun before they can be detected. Very often, information can be traced back in principle, but it is hard to do so in practice. The reason is that many factors and parameters contribute to the making of an image. First, there are geometric considerations. The

spatial organization of the image must somehow relate to that of the investigated object. For this geometric component alone, several factors contribute: what is the trajectory of the detected waves? Are they at all scattered, deflected or reflected before they reach the detector? Does the instrument select the incoming direction of the rays and how does this impact on the geometry of the image? Secondly, the knowledge about the imaging technique permits one to relate quantitative variations of the image (variations of values of pixels or of the corresponding shades of colors) to variations of some property of the object of investigation. For example, a region of an X-ray image with a relatively high count of rays is interpreted as showing a region of the human body with low attenuation rate, that is, relatively low in density compared with bony structures. But those are only the most general features of a given type of image. Knowing them amounts to knowing what is often referred to as the "theory of the instrument," but this is not enough to guarantee a correct interpretation of images. Every instrument departs from the generic, ideal, instrument that would instantiate such a theory, so that one also has to take many other factors into account, typically including various types of artifacts, noise, blur, geometric aberrations and other confusing effects. Though these factors may be just as knowable as the theory of the instrument, through calibration, more often than not, interpreting a scientific image is a very complex operation that is not so easily performed. So the problem that I am stating here is that of the complexity of the interpretation process, even when a correct interpretation is possible in principle. Scientists must take many factors into account: they must know a lot and also be able to use their knowledge to make the right interpretation and both aspects of knowledge possession and its correct use may lead to serious difficulties.

It is way beyond the scope of this paper to list exhaustively what the investigator should know in order to make good use of an imaging technique. What I want to emphasize, however, is that this knowledge is crucial to correctly interpret images and learn about phenomena, and that the amount of required knowledge tends to become greater and greater as experiments rely on more and more sophisticated instruments or chains of instruments. The worry is that the amount of required knowledge about the experimental setup might increase to a point at which it is difficult for most scientists to avoid errors. I will refer to this as the epistemological problem of image interpretation; an insistence on the amount of knowledge that is required for the investigator to make no mistakes. This is a problem that is distinct from that of the "theory ladenness of observation," which states that an observation loaded with theory (as is, arguably, every single observation in science) shall remain inconclusive/subjective/unfit to debunk a theory. I do not believe this to be the case, though I won't offer

justification for this here.[3] I shall then assume that there is such a thing as a correct image interpretation, that it is often attainable in principle, but that it is also so demanding in many cases that it poses a threat to the success of empirical investigation carried out with imaging instruments.

The epistemological problem of image interpretation can be linked to more general worries about the role of instruments (not necessarily imaging instruments) in empirical investigations, because instruments are demanding for investigators who might have to possess a lot of knowledge in order to make good use of them. This has led philosophers to various positions as to how to account for the role of instruments in enquiries. The next section aims to quickly survey these philosophical positions about instruments and to argue in favor of a very important place for them in successful investigations. The rest of this paper will then be dedicated to showing that computer image processing has become key to secure investigations based on imaging techniques.

3. The Role of Instruments in Scientific Investigation

When looking at contemporary practices of scientific investigation, there is no doubt that instruments of all sorts (measuring instruments, counters, detectors, imaging instruments, etc.) are used extensively and that they must contribute in no small part to the establishment of scientific knowledge. The results that they bring, whether they be single values, graphs or images must then somehow permit scientists to conclude that a certain state of affair is the case, for example that the local temperature in a room is 20°C, that the soil in a certain region is contaminated with radioactive substances, or that a patient has a brain tumor. However, instruments used by scientists cannot be thought of as machines that deliver truth. There are conditions that must be met in order to guarantee that the conclusions reached by scientists using the outcome of an instrument are valid. In particular, investigators are in charge of testing, controlling and calibrating their instrument so as to know how they work. The oft-used terminology is that they must know and be able to use the "theory of the instrument" (Adam 2004; Chalmers 2003; Franklin 1994; Schindler 2013) to correctly interpret the outcome.[4] It is also possible that, by learning about the instrument, investigators come to know that it is not functioning well and

3. For arguments against the inconclusiveness of theory-laden observation, see for instance Franklin et al. 1989 or Franklin 1994.

4. In the previous section, I drew a distinction between knowledge regarding the general aspects of an instrument (e.g., the "theory of the microscope," referring to the technique in general) and knowledge about the specifics of a given instrument, its particular defects, etc. In the remainder of this paper, I will use the term "theory of the instrument" as

that some material part of the device must be changed, but I will only consider the situation where an instrument works as expected, with no major problem.

The theory of the instrument has often been discussed in philosophy in relation with the concept of observation. In this very wide literature on observation, one important line of argument to discuss whether instruments (or some instrument) can be used as observational means considers the theory of the instrument as the central point. But from there, three different positions emerge. The first one is that, contrary to when we use mere perception (e.g., the naked eye), using an instrument creates an intermediate between the subject and the investigated object. Therefore, we have to know about this intermediate in order to secure the conclusions of the investigation. Since there is no theory of the instrument to be known when looking directly at a certain object, there is a strong difference between the two types of investigations, with and without instrument. Proponents of this first position conclude that the knowledge acquired with an instrument cannot be observational; it is inferred from the theory of the instrument and depends on the subject's background knowledge. These philosophers are of course not advocating for the avoidance of instruments in scientific investigation, but they define observation with an epistemological agenda, trying to ensure that observational knowledge is as independent as possible from already established knowledge. In their view, only by doing so could our observations serve to test our hypotheses and theories. Carnap, for example, found the use of instruments legitimate and admitted that physicists could have their own liberal usage of the term "observation" that included instrumented observations. He nonetheless kept a rigorous stance in order to define observation for philosophical purposes, accepting only instances of unassisted perception (Carnap 1966, pp. 225–26).

The second position denies that the use of an instrument requires much knowledge about the theory of the instrument, and that, in the same way that most of us do not know how our eyes function while being perfectly apt to see with them, an experimenter can ignore most of the physics of her instrument while being very good at using it. While such an externalist position with instruments is often found in epistemological writings, it is much less defended by philosophers who focus on experimental practice. One such philosopher, Ian Hacking, famously made that point about the

---

applying to both general and particular aspects of an instrument. Thus, claiming that one needs to know the theory of the instrument to successfully use it amounts to saying that both types of knowledge, general and particular, must be possessed in great enough detail by the investigator(s).

microscope (Hacking 1981), and it would arguably be possible to defend the same claim for all imaging instruments.[5] But this particular point of Hacking's position has been criticized, for instance by Morrison (1990), drawing also from Galison and Assmus (1989). For them, background theoretical knowledge is required to some significant extent, thus calling for a third position.

This third position admits that a good amount of knowledge about the instrument must be possessed in order to successfully use it in an investigation, but it asserts that this doesn't prevent scientists from reaching conclusive results. This position is held by a majority of philosophers of experimentation, led by Franklin in his essay on experimentation (Franklin 1986). The emphasis here is less on the defense of a particular conception of scientific observation. What is most important is that—observation or not—experimental conclusions based on the use of well-controlled and well-studied instruments are nonetheless about the best-justified claims that scientists can make. Instruments can be understood and controlled in a way that it is rational to accept experimental results, even when they conflict with already well-established hypotheses.

These three positions, though they are reconstructed from debates on observation, are helpful to discuss more generally of the validity of results obtained with instruments. The first one amounts to being skeptical to some extent about the possibility of ever securing the background knowledge required to draw valid experimental conclusions. As experiments rely on more and more sophisticated instruments or chains of instruments, the worry is that the amount of required knowledge about the experimental setup might increase to a point when it is impossible for most scientists to avoid errors. The second position amounts to denying that valid experimental results can be obtained only after proper theoretical training and asserts that looking through or with an instrument is akin to looking with the naked eye with enough habit. The third position takes the worry that underlies the first position seriously but it draws a more optimistic conclusion from there. Together, the three positions cover the three logical reactions to the problem raised at the end of the previous section, the epistemological problem of image interpretation, that is, the problem of the

---

5. I don't mean to imply here that it is Hacking's intention to defend the claim that one doesn't need to know about the instrument (the microscope in his case) in order to make good use of it to reach valid conclusions. In fact, Hacking also claims that microscope experimenters nonetheless learn about their instruments, if only empirically and that this knowledge is key to making them good experimenters. I am therefore just noting Hacking's ambivalence on the question, when he writes: "One needs theory to make a microscope. You do not need theory to use one" (1981, p. 309). This ambivalence is mostly due to the fact that the "theory of the instrument" remains unanalyzed.

amount of knowledge one must possess to reach valid conclusions from scientific images. Either this problem exists and it threatens the epistemological value of scientific imaging (position 1), or it does not exist, it states something that is not the case (position 2). Or it exists but one can find solutions to counter it (position 3).

When we look at the field of scientific imaging, it appears that only position 3 is viable as a response to the problem of image interpretation. Indeed, the skeptical conclusions of the first position are contradicted by the more and more important place taken by imaging instruments in empirical investigations and by the unquestionable success that using these instruments brings (e.g. early detection of many sorts of physical conditions). To state that the knowledge acquired by using instruments would be somehow inferior or less dependable than that obtained without them would I think go too brutally against the facts of scientific practice. However, what also appears when paying attention to practices is that scientists make a good deal of efforts in order to learn about their instruments, making position 2 also irrelevant. So while the sophistication of contemporary imaging techniques arguably creates a problem, namely that the amount of knowledge required about the experimental setup might just be too much to handle for experimenters, the practical successes obtained with these instruments call for an explanation instead of a contestation. This is what position 3 amounts to, and the elements of explanation that will be spelled out in the rest of this paper focus on the increasing role of image processing. In the next section, I present the general principles of image processing and the way computers are used in combination with imaging instruments.

4. Imaging System in the Digital Age: The Computer Part

In the previous sections, I have argued that imaging instruments are epistemically demanding for investigators who have to know elements of the theory of the instrument in order to properly use one and to correctly interpret images. I have also noted that today's instruments are for the most part digital, which has led to a much wider use of imaging instruments. Indeed, a major consequence of entering the digital age is that instruments such as microscopes and telescopes are now used in combination with computers, and that instead of looking through them, one now looks at an image formed on a computer screen that can be saved, stored, retrieved and circulated.[6] There are therefore many more instruments that qualify as imaging

6. Images could be formed with telescopes long before they became digital, by using an analog camera, but the primary way of using telescopes was to look through them. Scientists could also produce microphotographic images by combining a microscope with an analog camera.

instruments now, and they are also used intensively because computers render their use so much easier. In this section, I introduce the notion of image processing or of data processing for image enhancement to present another aspect of the combination of imaging instruments with computers. This aspect goes way beyond the usual understanding of digital as permitting investigators to produce and display images more easily, without printing. Computers also serve to modify the numeric data produced by the instruments alone. While I will elaborate in the next sections on how data processing constitutes a response to the problem of image interpretation by rendering images somehow more accessible to investigators, this section aims to present the principles of data processing.

An important terminological point is in order to distinguish between the outcome of the instrument and what is done with this outcome through computer processing. The former, that is, the list of values that obtains as the output of the instrument is referred to as the raw data. This term shouldn't be understood too literally, as meaning that these data are free of any sort of treatment or filtering. Rather, they are the least processed data that are accessible to scientists, even if they have already been transformed by some built-in procedure of the instrument. However, these data are only rarely used as the ultimate evidence from which scientists judge the investigated phenomenon. Most of the time, the raw data are processed with computers before they are used as evidence (and displayed as images onto computer screens). Computers then serve to turn the raw data into processed data. As a result, some of the efforts required from the investigator to interpret images can be taken charge of by computer programs. In order to understand how this is done, it is first important to remember that the raw data have a numeric format. They can therefore be considered as mathematical objects on which operations can be performed.

As a first, trivial, example, I will take a step away from imaging instruments and turn to a measuring instrument whose outcome is a single numeric value. Suppose, then, that an investigator has a digital thermometer that gives an estimate of the temperature. Suppose in addition that the investigator has learned through some calibration experiments that the thermometer systematically underestimates the temperature by one degree Celsius, delivering in fact a not very good estimate (as already stated in section 2, it is not the goal of this paper to discuss the means by which knowledge about the instrument and its defects is collected). This defect of the instrument is not very demanding for the investigator when she knows about it. All she needs to do is to add one degree Celsius to the obtained value. But there exists another solution to the problem (other also than changing the instrument or improving the one she has). Indeed, the piece of knowledge on which the correction is based is easy to formalize.

The relationship between the measured temperature **x** and the actual temperature **y** is x=y−1. One thus obtains a model of how the datum **x** is affected by the inaccuracy of this instrument. Now, since the experimenter is interested in accessing the actual temperature from the raw datum that she has, she needs to reverse this model and use instead z=x+1=T(x). Here, the processed datum **z** is supposed to be identical with, or at least significantly closer to, the actual temperature **y**. This example becomes of course less trivial in case the defect varies according to the temperature range, but the solution would be similar: Find out about the error of the measuring instrument for all the relevant ranges of temperature to create a model of data acquisition, and then use this model in reverse to process the data and obtain a much more accurate temperature estimate.

How do we go from this trivial single-valued example to an understanding of data processing for imaging instruments? The raw data produced by an imaging instrument can suffer from what is identified as a degrading factor. A very similar problem to the one just described with the thermometer could happen (and often does happen) with the detector of an imaging instrument. An instrument can sometimes detect only a certain proportion of the events it is supposed to detect, due to physical limitations of its sensitivity. If, after calibration, scientists determine that each of the surface areas of the detector only counts fifty percent of the incoming rays, they can establish the formal relationship between the detected signal **x** and the (supposedly) actual amount of rays that reaches the surface of the detector **y** as: x=y/2. This is a model of the degrading factor, and one that can be easily used in reverse to obtain a much better estimate of the signal: z=2*x. And again, a more complicated problem can be modeled the same way. If for example the sensitivity of the detector varies spatially, this can be taken into account to rectify the data point-by-point, for each of the elementary surface elements of the detector that deliver individual data points.

The general structure of such data processing procedures can then be spelled out. Suppose that a degrading factor can be modeled as a mathematical relationship between the raw data **x** and what the data would be without this degrading factor **y**: x=D(y). If **D** is invertible,[7] its inverse **T** permits one to obtain processed data that are supposedly no longer, or

---

7. Not all operators are mathematically invertible. An important part of the work of mathematicians working in image processing is to analyze the conditions under which a given operator is invertible. Then, with larger and larger datasets, even when an operator is invertible, its inverse can be extremely hard to compute due to computational costs. It is then another important task for mathematicians to optimize computation in order to adapt to the available computer resources, paying attention to both computation times and memory/data storage limitations.

much less, affected by the degrading factor: $\mathbf{y} = \mathbf{T(x)}$. In order to make things more concrete, and to illustrate some of the theses I will be defending in the following sections, I will now present two other examples of image processing algorithms that are frequently used in scientific laboratories. Following the example that I just gave of an algorithm that corrects for spatial inhomogeneity of detection rate, the two next ones deal with blurry images and image reconstruction.

The first example is that of an imaging instrument that delivers blurry images. In image processing, this type of defect is well known and understood, and it is fairly easy to at least improve the result. What characterizes the blur of an image is that the detection (of light, of x-rays, etc.) is not spatially accurate enough. Therefore, the image of a point, instead of being a point, is a spot, the diameter of which determines how blurry the image will be. If a small, almost punctual, object appears on the image as a rather wide spot, images taken with this same device will appear very blurry. Again, standard calibration operations permit scientists to know about this problem in a detailed way. They make an image of a very small object that approximates a point and they measure the diameter and shape of its image. The function that approximates this image, called the point-spread function (or PSF), is characteristic of an instrument and can serve to create a model of this particular degrading factor (see figure 1 for details). Indeed, the relationship between the blurry image $\mathbf{g}$ and the non-blurry image $\mathbf{f}$ is expressed by a convolution with the PSF $\mathbf{p}$ as follows: $\mathbf{g} = \mathbf{f}*\mathbf{p}$ where * is
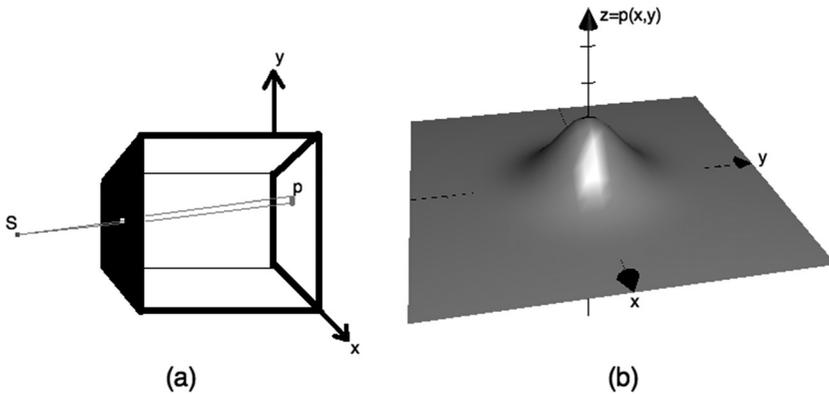
Figure 1. In (a), a small source S is used to measure the response of an imaging instrument schematized as a *camera obscura*. The image of S can be blurry for various reasons (non-punctual aperture, imprecision of the detector in the plane (x,y), etc.). The image of S can be modeled as a function z=p(x,y) to be used in a deconvolution algorithm (b).

the convolution operator. Then, this model is used again in reverse to obtain the processed data $h=T(g)$, where T is the inverse of the convolution with the measured PSF. The processed dataset h – or rather the corresponding processed image when h is displayed – is significantly sharper than the raw data and therefore much closer to f. This is known as the deconvolution method, which is a very common post-treatment in scientific image production.

The second example requires more technical details. Many imaging techniques rely on types of ways that can interact, not just with the surface of the investigated object, but that also traverse it and interact with its inner structure. As a result, the image carries information about a three-dimensional structure, but this information is projected onto a two-dimensional image. The most familiar images of this type are the radiographs that we see for example at the dentist's, which show the teeth, their nerves, the gum and the jaw all displayed in one image. Similarly, lung radiographs do not show just the lungs, but also the rib cage, the diaphragm, etc. In some cases, the fact that a lot of information is projected and structures that appear superimposed in the image is not a problem. If the angle is well-chosen, the organ of interest will be displayed clearly enough, without being obstructed by the presence of other structures. But in some other cases, projective images prove to be too messy, with too many overlapping details to be helpful to answer certain questions, even when taken from the best angle. There is simply too much information for one single image.

In order to obtain a clearer and more usable result with the same instrument, scientists have developed methods of tomographic reconstruction. These methods aim to produce images that show the investigated object in three dimension, slice by slice. Each reconstructed image represents a thin slice of the object as if it were cut-open. Instead of having the whole depth of the object projected and superimposed on one image, one can then navigate along the depth and see inner structures in a much clearer way. Tomographic methods rely on a mathematical result obtained by Radon in 1917 that proves that a 3D function can be exactly reconstructed from its 2D projections, provided that these projections are known for all directions around some axis of the object (Radon 1917). The application of this mathematical result to medical imaging led to the development of X-ray scanners or CAT-scans (Computed Axial Tomography), as well as to other techniques such as Single Photon Emission Computed Tomography (SPECT) in nuclear medicine for instance. These new types of instruments simply consist in mounting an already existing instrument (a radiograph, a gamma-camera, etc.) onto a rotating part so as to make projection images of the body from various angles. One can then obtain enough projection images, under enough different directions, to obtain a

good sampling of all the projections. The conditions for the application of Radon's method thus obtain and one can reconstruct the 3D distribution of the relevant physical property: attenuation of X-rays, rate of emission of gamma photons, etc.

The reconstruction process that leads from projection images to a slice-by-slice representation of the inner body has the same structure as the previous examples that I gave. Firstly, the degrading factor is identified as the inherent projective nature of the technique: Each point of the image carries information about the whole depth of the imaged object. To counter this problem, a projection operator $\mathbf{P}$ must model how a 3D distribution of the given physical property $\mathbf{f}$ is projected and gives rise to projection data (the raw data that are collected with the instrument) $\mathbf{g}$, so that $\mathbf{g=P(f)}$. Secondly, the operator $\mathbf{P}$ must be inverted so as to obtain a good approximate of the original distribution $\mathbf{f}$ from the raw data $\mathbf{g}$. Radon's result is precisely about this part. It states that $\mathbf{P}$ is invertible if and only if one has collected projections under enough different angles. Let $\mathbf{N}$ be the sufficient number of these projections, then $\mathbf{P}$ is the aggregation of $\mathbf{N}$ projection operators $\mathbf{P=}$ $\mathbf{(P^1,P^2,\ldots,P^N)}$ that model the projection for each of the $\mathbf{N}$ successive positions taken by the instrument (see figure 2). Similarly, g is the aggregation of projection images for each of the $\mathbf{N}$ successive positions taken by the instrument: $\mathbf{g=(g^1,g^2,\ldots,g^N)}$. Under these conditions, one can calculate the inverse $\mathbf{T}$ of $\mathbf{P}$ and use it to reconstruct $\mathbf{f=T(g)}$.

The few examples that I have given only cover a small part of what computer scientists can do now to process images so as to render them somehow more epistemically valuable or more accessible to investigators. Image processing (or data processing for image enhancement) has become as ubiquitous as the use of digital images itself, since most of the time, digital images in science are processed. Perhaps the most visible sign of the importance of image processing is the growing community of researchers who are dedicated to this field and the impressive number of journals.[8]

Computers now play a very important, and sometimes indispensable, part in the production of images. This is the reason why the term "imaging system," which refers to the combination of a digital instrument that produces raw data with a computer on which processing algorithms are implemented, is better suited to capture the way that images are actually produced. In the three remaining sections of this paper, I analyze the novel

---

8. These journals either address general issues (e.g., pattern recognition, computer vision, etc.) or focus on one domain of application (e.g., medical, satellite imaging, etc.) Examples include the *International Journal of Computer Vision, Journal of Mathematical Imaging and Vision, Image and Vision Computing, Pattern Recognition, IEEE Transactions on Image Processing, IEEE Transactions on Medical Imaging* and many more.
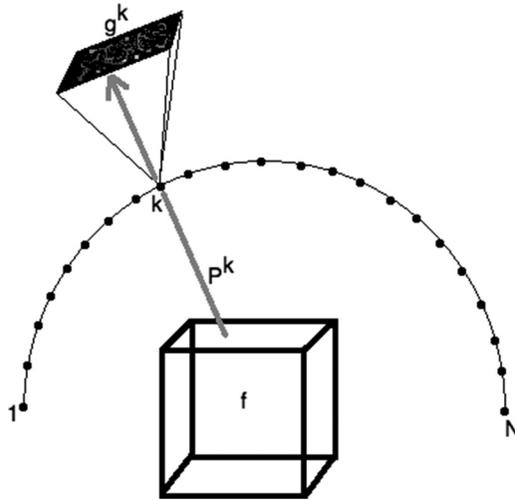
Figure 2. An instrument revolves around an unknown distribution of a physical property f. For one position k of the instrument, one obtains a projection image g^k. The mathematical operator $P=(P^1,\ldots,P^k\ldots,P^N)$ models the projection for each of the N positions taken successively by the instrument. If N is great enough, the inverse P^{-1} of P exists and f can be computed.

roles that computer image processing plays with respect to knowledge and image interpretation.

5. Economy of Knowledge and Skills For Investigators

In his 1981 essay "Do We See through Microscopes?," Hacking portrays two widely different types of microscopists. The first type, that appears at earlier times, when microscopes are still subject to many aberrations ("bad microscopes"), is that of experimental geniuses such as Leeuwenhoek, who could figure out ways to compensate for many of the degrading factors that affected what was seen through a microscope. The second type, today's investigators, are not required to be more than technicians having virtually no knowledge about the physics of the instrument. Hacking's characterization of today's microscopist may be a bit exaggerated: it is hard to imagine that a professional microscopist would at no point receive some training at least in optics and that this experimenter would not know a thing in physics that would be relevant at times for his scientific practice. It remains true, however, that as instruments are improved to produce better, clearer, images of phenomena, they relieve experimenters from the necessity to develop some of the skills associated with the figure of the experimental genius.

An extraordinary experimenter is characterized by two sorts of abilities. Firstly, she knows how to *use* her instrument in a way that maximizes the relevant information regarding a given investigation. In microscopy for example, there is a certain way to prepare the sample, to set the light and to tune the instrument very finely so that what is seen through the microscope is not just the big blur that an amateur would obtain. From there, her second ability is to be able to correctly interpret what she sees, separating artifacts from features that are informative for the investigation and ignoring the former while focusing on the latter. But even a well-prepared experiment using an imaging instrument can lead to images that are not interpretable in a straightforward way. There can remain various difficulties such as aberrations, deformations, overlapping of objects, etc. This is where extraordinary experimenters stand out, as they find ways to rectify appearances to correctly identify phenomena.

In the previous section, I presented three examples of data processing for image enhancement that aim to solve typical difficulties that experimenters encounter with scientific images produced by instruments alone. The first one was the possibly spatially variable rate of detection of rays that leads to inhomogeneous images. A correct interpretation of such images is one that takes into account the known inhomogeneity and doesn't associate it with a property of the investigated object but with a defect of the instrument. The second example was that of blurry images that one obtains with a detector that suffers from spatial inaccuracy. The contours of imaged objects are less sharp and the surface of blurry images of objects is bigger than it should be. This may be a problem, for example, for physicians who evaluate the size of a tumor in medical imaging, to check whether some treatment led to the reduction of this size. Here too, good experimenters are those who know about the spatial inaccuracy of the detection and who can take it into account for their interpretation. Finally, the third example was that of projection images with the inherent overlapping between various structures. The challenge here is for the investigator to be able to nonetheless discriminate distinct objects.

I presented these examples as being representative of the problems faced with raw data. More often than not, an instrument presents several such limitations all at once. For example, images obtained with a gamma-camera in nuclear medicine have a low spatial resolution, the detector is often not perfectly homogeneous and one can only access projection images that are therefore affected by the overlapping problem. The raw image (the image that one obtains by displaying the raw data without further processing) can be extremely difficult to interpret in this case, with so many degrading factors at once, thus requiring extraordinary capacities from the experimenter.

However, processing the data to counter these three problems renders the image much more easily interpretable. It does so by showing the object of interest much more clearly and by no longer demanding that the investigator know much about the instrument and its defects or limitations. Those have been corrected for and while most investigators know at least the type of defects that apply to the technique and, roughly, how they are countered by data processing, they need not know about the details—either of the problems or of the solutions. They no longer have to become specialists of the instrument to be good investigators in their domain of inquiry: a proper training with the theoretical and practical aspects of a technique is enough. This is the first identified role of data processing in imaging techniques: It relieves investigators from the requirement to know about the general and specifics of an instrument in very much detail. The experimenter can rely on data processing that already incorporates and uses knowledge to make the interpretation of images less epistemically demanding.

To claim that algorithms render images sometimes much more easily interpretable doesn't amount to saying that investigating phenomena with images becomes a somewhat trivial activity with image processing. Investigators are experts of their field: there remains much knowledge and know-how that must be possessed regarding the protocol of investigation (how to prepare the sample/patient, how to best set the instrument, etc.) as well as for the interpretation of images: an image on which artifacts and noise have been removed is certainly easier to interpret than when one still has to separate those from the phenomenon, but the task of interpreting, say, a clean radiograph isn't for everyone.[9]

6. Knowledge Made Explicit, Less Tacit Knowledge in Experimental Science

Since the emergence of a post-positivist philosophy of science in the late 1950s, one relativist worry has concerned the various individual background knowledge, capacities and skills that render scientific results dependent on subjects. In particular, Polanyi and others have insisted on the tacit knowledge that experts possess and make use of in a given field (Polanyi

---

9. Though I'm discussing image processing algorithms as help for interpretation, which they are most of the time, it's worth noting that an increasing amount of the work done by computer scientists in the domain of scientific imaging is about fully automated interpretation, especially in medicine, where Computer Assisted Diagnosis (CAD) aims to deliver an interpretation under propositional form from an image. Still, even in that case, there remains a task for an expert (human) investigator who is in charge of controlling the automated diagnosis.

1958; Collins 2001). It is a kind of knowledge that is particularly difficult for non-experts to reach because it is hardly communicable. In the same way that explaining to someone how you recognize a face—from which facial features, at which level of details, etc.—could be particularly difficult, an expert of a certain imaging technique would also have trouble explaining exactly in which cases he or she recognizes a certain pattern of interest in the presence of noise, aberrations, artifacts, blur and other degrading factors. Making the right judgment from images that are affected by these factors is a hard-gained skill that may rely on tacit knowledge.

Possessing the tacit knowledge required to correctly interpret images arguably amounts to having learned empirically about the instrument in a way that is quite distinct from theoretical training about the instrument. For instance, one could learn in the classroom about the type of noise that affects a certain technique, its theoretical distribution that follows a certain probability law and its theoretical reason, but this person would probably still miss the experience of having looked at thousands of images and being able to subtract noise immediately to see the phenomenon. It's a type of knowing about the instrument, about the technique and the way data are produced, which seems to be required for proper image interpretation, while not being really teachable. What kind of knowledge is at stake and how is it used? Instead of answering this question, one is stuck with the idea of an unanalyzable skill or "know-how" that is just gained through prolonged immersion in experimental activity. Implementing data processing algorithms precisely addresses the problem of finding out explicit ways of compensating for degrading factors.

By rendering images more easily interpretable, data processing relieves image interpreters from the obligation to possess a good part of such tacit knowledge. They do not have to gain the skills that are required to correct for aberrations and to separate phenomena from artifacts. But beyond this economy already discussed in the previous point, data processing also helps to render tacit knowledge much less mysterious. Indeed, one of the reasons why tacit knowledge has been deemed incommunicable is that it is not clear what exactly a trained subject does when she relies on it, for example to interpret images correctly in face of the various degrading factors.

This claim does not amount to saying that deblurring or denoising algorithms are models for actual cognitive tasks that are performed by skilled agents who possess tacit knowledge. When such skilled agents know how to interpret noisy, blurry or otherwise degraded images, they don't necessarily do the same mental operations as those implemented in algorithms. However, algorithms do exhibit ways to deal with degrading factors in an entirely explicit way, thus avoiding the problem of dependence of experimental activities upon tacit knowledge. Correcting procedures involve

the two phases that I have described in section 3. The first one consists in gathering knowledge about the degrading factors so as to create a formal model. For example, one measures the point spread function of an imaging device or the distribution of noise. The formal model implements those features by determining which mathematical operator applies best. As seen previously, the point spread function is convolved with ideal (non-blurry) data, leading to the formation of a blurry image, but in the case of noise, the operator is often a simple addition. The sensitivity of the detector is modeled using a multiplication factor, etc. The second step consists in using the formal model of the degrading factors in reverse, so as to improve images, as in the deblurring example, where the point spread function is used in a deconvolution algorithm. Both steps are made equally explicit in algorithms. By reading the computer program that implements for example a deblurring algorithm, one will find the equation of a function that describes the point spread function as well as the lines of code that perform a mathematical deconvolution. The same goes of course with any other type of data processing: image reconstruction, denoising, geometric correction etc.

The claim that computer processing permits scientists to make investigation less dependent upon tacit knowledge could hardly be sustained if only a handful of diehard computer scientists knew how to explore and decipher data processing computer programs. So is it really the case that the tacit knowledge that is arguably turned into explicit knowledge in principle is actually accessible to scientists in practice? I can see at least two difficulties that threaten my characterization of data processing as a way to turn tacit into explicit and accessible knowledge. The first one is that computer programs can be developed by private companies that protect their software so that they can be used in combination with an instrument, but functioning as black boxes for experimenters, with only the programmers of the company knowing about the details of algorithms. The second one is that professional coding can be extremely technical, with complex data structures and various sub-programs that require an immense amount of effort to be understood by an external user.

In response to these worries, as I mentioned earlier, there are many dedicated journals that publish methods, algorithms, or models and whose goal is to display and discuss the tools of data processing thoroughly and in a form that is completely explicit. Data processing is structured as a proper scientific field, much more than it is the private product of industrial companies. The vast majority of published work in this field is achieved in research labs of universities, though the private sector also contributes with research papers, as well as with scientific collaborations and funding. But the main point is that the methods that are ultimately implemented by the industry are those that have been published. Therefore,

investigators in a given field that involves images do have access to the details of the various transformations of the raw data. Even if the code of the software that comes from the instrument manufacturer is not directly accessible to the scientists that use them, the mathematical methods that they rely on are documented and easily accessible since they have been published. Also, scientists are often asked to enter certain parameters to run algorithms, which encourages them, and even forces them to know how their data are processed.

Data processing thus achieves the goal of rendering scientific imaging less dependent upon the skills of individual researchers, first, by making images more easily interpretable, and secondly, by having become a scientific field on its own with methods that are being almost systematically published. As a consequence, the amount of tacit knowledge that is involved in empirical investigation is significantly reduced after the development of computer methods for data processing.

7. Extension of Computational Capacities and Subsequent Extension of Observational Capacities

The skills involved in interpreting images can require more or less time and efforts to be acquired. Even though some images seem at first quite obscure, they still permit specialists to recognize patterns that are known to reliably indicate certain phenomena with proper training. But some data require more than training to be useful for correct detection of phenomena. One example is MRI because the detected signal can only be obtained in the Fourier (frequency) space. Displaying these data in image form thus gives no direct indication about the organs and structures that are being explored; the features of an image in Fourier space are simply unreadable even by the most talented MRI experimenter. In the case of MRI, data must be "reconstructed," that is, transformed to be displayed in a different space, namely the 3D Cartesian space in place of the Fourier space.

With MRI raw data, then, phenomena cannot be correlated with any recognizable pattern in the Fourier space. MRI radiologists are also unable to mentally do an operation similar to reconstruction from Fourier to Cartesian space. This requires an inverse Fourier transform on thousands or millions of points, which goes way beyond the computing capacities of the human brain. Therefore, in this case, computer algorithms serve to extend these computing capacities so as to render MRI data usable.

The extension of computational capacities that is rendered possible with computers has been discussed by some philosophers, mostly in relation with models and simulations (Galison 1996; Winsberg 2001; Fox-Keller 2003; Frigg and Reiss 2008; Humphreys 2009). In the context of this

paper, that doesn't address issues about simulations,[10] fewer philosophers have proposed to account for the now widespread use of computers to supposedly improve the quality of empirical data obtained with instruments. Humphreys is a notable exception to this when he argues that "computational devices are the numerical analogues of empirical instruments: They extend our limited computational powers in ways that are similar to the ways in which empirical instruments extend our limited observational powers" (Humphreys 2004, p. 116). This metaphor is particularly well suited to the context of this paper because it appeals to two important elements of the contemporary way to carry out empirical investigations: the use of instruments to supplement our senses and the use of computers to supplement our computational capacities. In image processing, these two elements combine so as to give scientists even greater observational[11] powers than with instruments alone. Humphreys is aware of this and gives an example with MRI:

> An external magnetic field lines up the spins of the protons in hydrogen nuclei. A brief radio signal is used to distort that uniformity, and as the hydrogen atoms return to the uniform lineup, they emit characteristic frequencies of their own. Computers process this information to form an image suitable for display to the human visual apparatus, this image representing various densities of hydrogen and its interaction with surrounding tissue. (2004, p. 38)

The computer processing that Humphreys describes is precisely the inverse Fourier transform, which makes the image suitable for display to the human visual apparatus because it is no longer a representation of frequencies in the Fourier space, but of densities of hydrogen in tissues in the Cartesian 3D space. Therefore, it doesn't just extend the computational powers of human agents, it also is crucial to extend observational powers when used in combination with the MRI apparatus that collects data in the frequency

10. One could argue that data processing such as image reconstruction is similar to a simulation or perhaps even is a form of simulation. It is at least closely related since it relies on models that appear on equations and these equations must be solved. I don't address the question of the similarities and differences between data processing and simulations here.

11. My use of the term "observation" is not in line with most of the conceptions defended in philosophy, since observation is often associated with the use of unassisted perception: naked-eye seeing, unaided hearing, etc. I am not here making an argument in favor of instruments as legitimate means of observation as this would require a whole paper specifically on observation. I am using the term in a more casual way, as an equivalent of "valid experiment" (Franklin 1986) or "detection of phenomena" (Bogen and Woodward 1988).

domain. The inverse Fourier transform algorithm serves as a sort of virtual lens that corrects otherwise unreadable data, but a lens that has no equivalent as an actual material device.

Of the several examples of data processing that I have given so far (correction of sensitivity, deblurring, 3D reconstruction from 2D projections), MRI best supports the claim that data processing increases our observational capacities.[12] Indeed, while these other examples illustrated how data processing helped scientists by producing data that were somehow *easier* to interpret, reconstructed MR images are straightforwardly interpretable while raw MR images in the Fourier space are not interpretable at all, by anyone. This demonstrates that data processing is more than a tool for helping scientists to see better and clearer in scientific images. While some types of processing aim to do "image enhancement," that is, improving an already readable image, some other types really are indispensable to gain any relevant information with a given technique. In addition to MRI, echography (ultrasound imaging), or optical coherence tomography (OCT) are other techniques for which no image is obtained other than one reconstructed from the raw data, using algorithms.

Another way to understand such methods as MRI is to get back to the condition that I gave in section 2 for an imaging technique to be successful. There, I argued—against Shapere's condition that (carriers of) information be directly transmitted—that the trajectory of waves should be knowable. I proposed this condition precisely in response to the computational possibilities that are offered nowadays, that lead to so many more imaging techniques being successfully developed. Before the computer era, it was legitimate to posit that only data that were in a straightforward enough relationship with the object of investigation could be correctly interpreted. But with computers comes an extraordinary increase of computational capacities, and therefore, possibilities to decipher data that would be unworkable for human agents. From the moment that a formal model can be validated to describe the processes of imaging, this model, as complex as it is, just has to be invertible for scientists to gain information about the imaged object. Hence, inverse Fourier transform or deconvolution algorithms have proved successful to obtain accurate representations of investigated objects.

12. I am focusing here on the augmentation of observational capacities that is based on mathematical transforms applied to otherwise unreadable data. This is distinct from the case where visualization alone makes data about invisible phenomena (e.g., telescopic data in the infrared part of the spectrum) visible by just displaying them on a computer screen, that is, by attributing colors to values—but without changing the values.

I argue, then, that data processing permits scientists to extend their observational capacities to a whole new level, because information about the investigated object can now be encoded in almost any arbitrary fashion by an instrument and yet be transformed into a completely interpretable image. While imaging instruments used to be considered correctly working only in case they would deliver interpretable images, they can now serve the unique function of collecting data whose epistemological value depends on further processing.

8. Conclusion

The goal of this paper was to present the practices of data processing for scientific imaging and to evaluate the roles that they play with respect to investigators. Three such roles were identified: facilitating interpretation, turning tacit into explicit knowledge, and extending computational capacities in order to also increase observational power. This study will trigger, I hope, an interest for refined philosophical analyses of data processing, as it has become such an important transversal field for scientific investigation. The massive use of data processing in scientific imaging, the ever growing number of researchers involved in this field, and the massive number of published works in dedicated journals certainly suggest that more attention should be dedicated to these new aspects of scientific practice.

References

Adam, Matthias. 2004. "Why Worry About Theory-Dependence? Circularity, Minimal Empiricality and Reliability." *International Studies in the Philosophy of Science* 18: 117–132.

Bogen, Jim and Jim Woodward. 1988. "Saving the Phenomena." *The Philosophical Review* 97: 303–352.

Carnap, Rudolf. 1966. *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*. New York: Basic Books.

Chalmers, Alan. 2003. "The Theory-Dependence of the Use of Instruments in Science." *Philosophy of Science* 70: 493–509.

Collins, Harry. 2001. "Tacit Knowledge, Trust, and the Q of Sapphire." *Social Studies of Science* 31: 71–85.

Fox Keller, Evelyn. 2003. "Models, Simulation, and Computer Experiments." Pp. 198–215 in *The Philosophy of Scientific Experimentation*. Edited by Hans Radder. Pittsburgh: University of Pittsburgh Press.

Franklin, Alan. 1986. *The Neglect of Experiment*. Cambridge: Cambridge University Press.

Franklin, Alan, M. Anderson, D. Brock, S. Coleman, J. Downing, A. Gruvander, J. Lilly, J. Neal, D. Peterson, M. Price, R. Rice, L. Smith,

S. Speirer and D. Toering. 1989. "Can a Theory-Laden Observation Test the Theory?" *British Journal for the Philosophy of Science* 40: 229–231.

Franklin, Alan. 1994. "How to Avoid the Experimenter's Regress." *Studies in the History and Philosophy of Science* 25: 463–491.

Frigg, Roman and Julian Reiss. 2008. "The Philosophy of Simulation: Hot New Issues or Same Old Stew?" *Synthese* 169: 593–613.

Galison, Peter and Assmus, Alexi. 1989. "Artificial Clouds, Real Particles." Pp. 225–274 in *The Uses of Experiment: Studies in the Natural Sciences*. Edited by David Gooding, Trevor Pinch, and Simon Schaffer. Cambridge: Cambridge University Press.

Galison, Peter. 1996. "Computer Simulations and the Trading Zone." Pp. 118–157 in *The Disunity of Science: Boundaries, Contexts, and Power*. Edited by Peter Galison and David J. Stump. Stanford: Stanford University Press.

Hacking, Ian. 1981. "Do We See Through a Microscope?" *Pacific Philosophical Quarterly* 62: 305–22.

Humphreys, Paul. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.

Humphreys, Paul. 2009. "The Philosophical Novelty of Computer Simulation Methods." *Synthese* 169: 615–626.

Morrison, Margaret. 1990. "Theory, Intervention and Realism." *Synthese* 82: 1–22.

Polanyi, Michael. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. University of Chicago Press.

Radon, Johann. 1917. "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse [Reports on the proceedings of the Royal Saxonian Academy of Sciences at Leipzig, mathematical and physical section] (Leipzig: Teubner) 69: 262–277; Translation: Radon, J., trans. P. C. Parks. 1986. "On the Determination of Functions from Their Integral Values along Certain Manifolds." *IEEE Transactions on Medical Imaging* 5: 170–176.

Schindler, Samuel. 2013. "Theory-Laden Experimentation." *Studies in History and Philosophy of Science Part A* 44: 89–101.

Shapere, Dudley. 1982. "The Concept of Observation in Science and Philosophy." *Philosophy of Science* 49: 485–525.

Winsberg, Eric. 2001. "Simulations, Models, and Theories: Complex Physical Systems and their Representations." *Philosophy of Science* 68: S442–S454.