The Digital Secret of the Moving Image

Enrico Terrone[*]

University of Torino

First, I analyze the five conditions that constitute Noël Carroll's definition of the moving image, and I argue that the third condition – the type-template-performance condition – raises a problem concerning the relationship between the cinematic type and the digital type. Solving this problem allows us to explain away the other conditions introduced by Carroll, and to give a simpler formulation of the definition of the moving image. Then, I compare this new definition to a similar amendment of Carroll's definition proposed by Trevor Ponech. Finally, I test my definition by considering two issues that all attempts to define the moving image have to deal with: the historical issue and the depiction issue.

1. Carroll's Definition

In order to define films, and more generally moving images, Noël Carroll (1996, p.70) proposed a set of five necessary conditions:

> (1) Detached display: the space of the viewer is spatially disconnected from the content of the images.

> (2) Technical possibility of the impression of movement.

> (3-4) Type-template-performance ontological structure:

>> (3) the performance-token (i.e. what we see on the detached display) is generated from a template (e.g. the film strip, the DVD, the television signal) which is, in its turn, the token of a type (i.e. the moving image as such);

>> (4) moreover, the performance-token (i.e. the projection) is not an artwork in its own right.

---

[*] Email: enriterr@gmail.com

(5) Two-dimensionality.

In a later text, Carroll (2008) strengthens his definition by arguing that the five necessary conditions also count as jointly sufficient. In his old account, Carroll characterized the conditions as necessary but not jointly sufficient, since he did not intend to include among the moving images some artefacts (as for instance flip books and the zoetrope) that however satisfy all the necessary conditions. Yet, in his new account, Carroll considers those artefacts as appropriate instances of the concept of a moving image, and so these same necessary conditions can be now considered as jointly sufficient:

> So, x is a moving image if and only if (1) x is a detached display or a series thereof; (2) x belongs to the class of things from which the promotion of the impression of movement is technically possible; (3) performance tokens of x are generated by templates that are themselves tokens; (4) performance tokens of x are not artworks in their own right; and (5) x is two-dimensional. Notice that each of these five conditions is alleged to be necessary and to be conjointly sufficient. (2008, p.73)

I what follows, I will argue that Carroll's third condition – the type-template-performance condition – is both the most problematic and the most important one. It is the most problematic one because it raises a problem that Carroll does not consider but that I claim is crucial for a definition of cinema. At the same time, the third condition is the most important one in the sense that the other conditions can be explained just by using the third condition and the solution to the problem it raises.

2. The Digital Type Problem

The problem is the following. Let us consider the case in which the cinematic template is not a film strip or an electronic signal, but a digital copy of the movie. In this case, the digital copy is not only the token of a cinematic type C (the movie as a type), but also the token of a digital type D (the sequence of the frames considered as bitmaps). So, the question is: are those two types the same entity or two different entities?

If we assume that they are different entities, then we have to find either something that is a token of D but not a token of C, or something that is a token of C but not a token of D. The first alternative is not viable: if, as we have assumed, D is a digital version of C, then every token of the digital type D is also a token of the film type C.

So we have to explore the second alternative, and suppose that the analogical tokens of the film type C are not tokens of the digital type D. This seems correct if we consider the digital type D as a series of numerical digits: the film strip or the electronic signals have nothing to do with numerical digits. But the digital type is not a series of numerical digits. The numerical digits are just placeholders for the chromatic values of the pixels (that is, the points that constitute the moving pictures). Therefore the digital type D has to be considered as a temporal series of frames, each of which is composed by a spatial distribution of pixels. Given that, nothing prevents us from considering the film strip or the electronic signal as tokens of the digital type D: they are spatio-temporal distributions of points of colour that instantiate the temporal series of frames individuated by the digital type, and so they can be considered as its tokens. Although in these cases the pixels are not separable one from another, the visual properties specified by the pixels are however instantiated.

We can also formulate our argument as a reductio ad absurdum:

> *If*
>
> (I) D is a digital type whose tokens ds are digital copies of the cinematic type C
>
> *and*
>
> (II) the cinematic type C has at least one token c that is not a token of the digital type D,
>
> *then*
>
> (III) the cinematic type C has some constitutive visual properties different from those specified by D (namely, some properties such that c is a token of C but not of D),
>
> *and so*
>
> (IV) the tokens ds of D are not digital copies of C.

The conclusion (IV) contradicts the hypothesis (I) which must be considered true (unless we want to endorse a counterintuitive "purist ontology" claiming that there can not be digital copies of films), so the hypothesis (II) must be false.

2. *The Explanatory Power of the Digital Type*

Technology reveals the essence. The digital copies of a film are not just mere tokens: they individuate a type structure that is exactly the film structure. The cinematic type C and the digital type D are ontologically the same entity, but the digital type has an epistemological advantage: it gives us access to the structure of the moving image. Therefore we can use this structure in order to derive, one by one, the other conditions that articulate Carroll's definition.

(1) The first condition proposed by Carroll – the detached display condition – claims that the content of the moving image "is disconnected phenomenologically from the space in which I dwell bodily as a physical being" (2008, p.57). That is to say that the moving image supports an experience that emulates allocentric vision (having coordinates independent from the viewer's body, and allowing her to recognize *what* there is) but not egocentric vision (a kind of vision that has coordinates related to viewer's body, and allows the viewer to recognize *where* she is).

I argue that this condition can be derived from the type-token-template condition. The cinematic display is necessarily detached from the space of the viewer because the content of the moving images is individuated by a type that can be instantiated in different spaces of vision and that has no special relation with a singular space of vision. All the vision systems that Carroll attempts to distinguish from cinema by introducing the detached display condition (e. g. mirrors, telescopes) can be distinguished from cinema simply by considering that they do not have a type-template-performance ontological structure. On the contrary, the cinematic display has a structure that is specified by the digital type, and it is detached because the type itself is detached in virtue of its abstractness. There could be just one token whose display is not detached from the space of the viewer: for instance, the monitor of a digital camera, in case it were used

as a mirror. Yet, given that the identity of the moving image is individuated by the type, not by the token, we must conclude that also in this case the cinematic display is detached, that is, it is "discontinuous from the space we inhabit" (Carroll 2008, p.58): although a single token (the token that instantiates the movie during its production) could be not detached, the type is detached because of its being an abstract type, and all the other tokens will be detached too, because of their instantiating an abstract type.

(2) In formulating his second condition – the possibility of the impression of movement – Carroll has to talk about the "possibilities of the impression of movement" and not simply about "the impression of movement," for he wants to take into account works such as Chris Marker's *La jetée*, Hollis Frampton's *Poetic Justice* and Michael Snow's *One Second in Montreal*, made, partly or wholly, by still images. According to Carroll, these works are different from a mere slideshow because the viewers can legitimately expect (at least at the first viewing) that sooner or later there will be some movement in the pictures: "it is always justifiable to entertain the possibility that the image *might* move" (2008, p.60).

   This condition is in its turn derivable from the structure of the digital type. "To entertain the possibility that the image *might* move" is indeed "always justifiable" because the film as a type is composed by a series of frames whose temporal rate is capable of determining the impression of movement in our perception. The epistemological possibility (we know that there could be movement) is based upon an ontological possibility (the structure of the type, namely: its being constituted by a series of frames, supports the impression of movement). The content of paintings and photographs can not move since they are constituted by a *spatial distribution* of colored points, whereas the content of the moving images can move since they are constituted by a *spatio-temporal distribution* of colored points.

   In short, the movie's content can move since the movie not only occupies a surface, but also has a duration. The moving image is not constituted by a series of frames because it is moving: it is moving because it is constituted by a series of frames. That is why "movement is a permanent possibility in cinema" (Carroll 2008, p.60). Even in the cases in which the

moving image *does not really move*, it *might move*, because the cinematic type carries this possibility in its structure.

(3, 4) The third and the fourth conditions proposed by Carroll are strictly related. The third condition claims that the moving image has a distinctive type-template-performance ontological structure: the movies are abstract structures (here the type) that are instantiated by material objects (here the templates) that can be screened giving rise to visual displays (here the performances). I am arguing that this condition alone can ground the entire definition of the moving image.

The fourth condition adds that "the performance of a motion picture – a film showing – is not an artwork in its own right" (Carroll 2008, p.68), unlike the performance of theatrical or musical artworks. In other words, "performances of motion pictures are not objects of artistic assessment, whereas theatrical performances are" (Carroll 2008, p.70). That is because the theatrical performance is an interpretation that involves intentional acts, while the cinematic performance is just a mechanical process that involves nothing more than physical causation.

In my account, Carroll's fourth condition can be explained as a mere consequence of the third one. The cinematic type, as a spatio-temporal distribution of colored points, specifies all the relevant visual properties that must be actualized by the screening. Therefore the cinematic performance (that is, the screening process) leaves no space for interpretation and intentional determinations, and so it can not be evaluated as a work of art in itself, unlike the musical or the theatrical performances. The structure of the cinematic type, conceived as a series of frames composed by pixels, is such that, at the token level, it only remains to make those frames accessible to the viewers: there is nothing more to add. Surely, we can distinguish between better or worse templates (that is, better or worse copies), and also between better or worse performances (that is, better or worse screenings). Yet this is just a technical distinction, which means that there can be copies or performances that exactly reproduce almost all the pixels specified by the type (as it happens with a high definition copy and a high quality projector) and others tokens that miss a relevant number of pixels (as for instance a VHS copy or a performance made by an old and worn projector). It is not a matter of interpretation at all, it is only

a matter of better or worse approximation to the visual appearance fully specified by the type.

(5) The fifth condition – two-dimensionality – is introduced by Carroll in order to exclude the moving sculptures from the definition of the moving image:

> Consider what might be called moving sculptures of the sort that are exemplified by the moving figurines on various antique clocks [...] Inasmuch as these sorts of figures are multiples, they meet the conditions stated above for membership in the category of the moving image, as do the statuettes of ballerinas that once cavorted on the tops of music boxes. All these examples are detached displays in the same sense that the locale of Burnham Wood on a theater stage is literally discontinuous with the space we inhabit. These ballerinas are not dancing in our space, but some ideal, ethereal space. Moreover, the aforesaid sculptures, we are assuming, are manufactured from a template and come in multiples. And the mechanical movement in the semblance of a pirouette of the ballerina — her performance, if you will — is not an artwork in its own right. Yet, although she satisfies the criteria, clearly, our ballerina does not seem to be the sort of thing that we have in mind when we speak of motion pictures or the moving image. (2008, p.72-73)

Carroll claims that two-dimensionality is necessary in order to exclude the moving sculptures, yet he has to specify that two-dimensionality alone can not cleave motion pictures from theatre, because "there is, in fact, theater that is two-dimensional, for example the shadow-puppet shows of Bali, Java, and China" (2008, p.73). So, in order to cleave motion pictures from theatre, we still need the type-template-performance condition. But this solution raises a question that Carroll does not consider: what about a shadow-puppet show made by means of a moving sculpture? In this case, we have both two-dimensionality and the type-template-performance structure, and so we have to conclude – against our intuitions – that these images *are* motion pictures. So the problem of the moving sculptures has not really been solved by the two-dimensionality condition, because it suffices to use the moving sculptures as an input for a shadow-puppet show and the problem arises again.

The solution of the moving sculptures problem is not to be found in the fifth condition, that is, two-dimensionality, but once again in the third condition: the type-template-performance structure. Moving sculptures differ from motion pictures because they have different structures at the type level. The type of the moving sculpture is not made of a temporal series of frames and by a spatial distribution of pixels, but it rather comprehends properties as height, weight, chemical composition. So the moving sculpture is excluded from the cinematic domain because of the different ontological structure of its type, and not because of its three-dimensionality.

Such a conclusion allows us to take into account the possibility of holography and of three-dimensional cinema, whose intuitive cinematic nature is admitted by Carroll himself:

> Suppose that it became possible to remake motion pictures, like the HBO television series Rome, in the round by means of holography. Imagine that we could project a scene of mortal combat in the Coliseum three-dimensionally with the audience seated around the virtual arena like ancient Romans. Would not such a spectacle be rightfully categorized as a moving image? (2008, p.73)

We need a criterion that allows us to distinguish those three-dimensional screenings, like holography, that our intuition is prone to consider as moving images, from those ones, like moving sculptures, that we aim to exclude from the cinematic domain. The two-dimensionality condition fails to support this distinction and so Carroll is forced to exclude both moving sculpture and holography from the cinematic domain. Yet, if we pose as a criterion the ontological structure of the type, then we can consider holograms as having a peculiar cinematic type whose frames are three-dimensional spatio-temporal distributions of pixels, instead of two-dimensional ones like in ordinary movies. In this way, we can exclude from the cinematic domain only the moving sculptures, whose types are not spatio-temporal distributions of pixels, while preserving the cinematic nature of holography despite their three-dimensionality.

4. Ponech's Definition and the Type-Token Confusion

So far, we have shown that the five conditions proposed by Carroll can be reduced to the third one – the type-template-performance condition – in virtue of the properties of the digital type, which allow us to make the structure of the cinematic type explicit . A similar strategy is adopted by Trevor Ponech (2006), and now it is time to compare our account with his. Ponech tries to reduce Carroll's five conditions to the first one – the detached display condition – that he reformulates as a "stroboscopic display condition":

> I agree that such displays are 'detached.' My reasons go a bit beyond Carroll's, though [...] I identify cinema with the visual display [...] Very generally, then, the visual display is a delimited area of illumination [...] Pixels and stroboscopic-motion are two narrow, intrinsic features it possesses essentially. 'Pixel' usually denotes 'picture element.' I use it in a slightly adjusted but related technical sense. By 'pixels' I intend points of light. This usage converges with descriptions of movie images as constructed from separate regions varying independently in spectral distribution. At a basic level of physical description, visual displays are composed of pixels [...] A display is more than a field of pixels. It is also an event. At a nonmicrophysical level of description, the display and the pixels therein undergo a distinctive type of motion. Pixels flash, periods of illumination alternate with periods of nonillumination, the pixels' brightness and intensity fluctuates [...] This is the display's stroboscopic motion. (2006, p.191-192)

In such an account, as in mine, the moving image is conceived as a spatio-temporal distribution of pixels, yet Ponech situates the distribution at the level of the tokens, not at the level of the type. Ponech defines the moving image as a stroboscopic display, but stroboscopy is a particular technique, that is, a particular way of constructing cinematic tokens by means of a regular succession of phases of light and phases of obscurity. In this sense, the best argument that Ponech could bring in defense of his definition is the following: "To date, all movies derive from stroboscopic motion" (2006, p.193).

In his criticism of Ponech's paper, Jonathan Walley observes that "this is merely an assertion that something has historically being the case, which does not exactly make for a strong essentialist argument meant to transcend historical variations in cinematic technologies" (2007, p.408). In such a perspective, the main objection to Ponech's definition is the following: how can we take into account past, present and future techniques that do not use stroboscopy but can however produce moving images very similar to those produced by the stroboscopic display? As Walley points up:

> Most film projection system are stroboscopic, employing a rotating shutter to alternately block and reveal light, but not of them are [...] Phosphors in video monitors do not exhibit true stroboscopic motion [...] LCD displays also do not require stroboscopy. (2007, p.408)

In his reply to Walley, Ponech is forced to revise his definition of a stroboscopic display, in order to accommodate phosphors and LCD displays among the cinematic displays:

> By 'stroboscopic visual display' (SVD) I intend a spatially-temporally delimited field of resolving elements, comprising points of light I call 'pixels', which undergoes a rapid cycle of phase changes when it is activated. (2007, p. 412)

And just below Ponech specifies that:

> [t]he resolving elements' onset/offset phase need not alternate between illumination and extinction; the onset/offset dynamics might comprise a stream of high frequency, continuous changes to the elements' brightness, intensity, and spectral distribution. (2007, p.413-414)

Yet, so reformulated, the stroboscopic display is no more really stroboscopic: it is just a spatio-temporal distribution of points of light changing in their color values. And in such an account the cinematic display covers a too wide domain: even a rainbow, as a spatio-temporal distribution of points of light changing in their values, could belong to the domain of moving images.

The problems for Ponech's definition arise from the fact that he defines the moving image by focusing on its tokens, instead of its type, and so he promotes an aspect that is indeed just a historical contingency – stroboscopy – to the rank of an essential feature. In this sense his definition of cinema is affected by a "type-token confusion", and a definition that remains at the token level, instead of taking into account the type level, can not be disengaged from historical contingencies.

I propose to solve this problem by shifting the spatio-temporal distribution of pixels from the token level to the type level. A movie really is, as Ponech points out, a "spatially-temporally delimited field of pixels", but in order to individuate the ontological specificity of cinema, this visual field has not to be considered as a physical phenomenon, but rather as a type that can be iterated by a variety of tokens.

5. The Historical Issue

Carroll's definition of the moving image presupposes that movies belong to the ontological kind of abstract types, and so does our definition of a movie as a spatio-temporal distribution of pixels. Yet it is doubtful whether artifacts and especially artworks can be considered as abstract types. Rohrbaugh (2003) has argued that conceiving artworks as abstract types does not take into account not only their being created instead of discovered (cf. Levinson), but also their modal and temporal flexibility, that is, the fact that an artwork could have had different properties (here the modal flexibility) and could change its properties in time (here the temporal flexibility). Rohrbaugh has therefore proposed to give up the type model and to conceive artworks as "historical individuals", namely: "continuants which stand in a relation of ontological dependence to a causally-connected series of physical (sometimes mental) particulars" (2003, p.197).

According to Rohrbaugh, the identity of a historical individual is not determined by a structure (that is, by an abstract type), but depends on its belonging to a causal-historical chain that starts with an event of creation and ends when there is more than one link that continues the chain in time. Nevertheless, he admits that historical individuals *have* structures that are transmitted through the chain, which therefore works as

"a structure-preserving and transmitting mechanism which undergirds the continued life of an historical individual" (2003, p.184).

That being the case, also in an account of movies as historical individuals, the cinematic type has to be recovered in order to specify which structure is transmitted along the chain that constitutes the historical individual. We need to specify this structure if we want to distinguish movies from other genres of historical individuals. Ultimately, even if the abstract type as a spatio-temporal distribution of pixels does not show us what a movie *is*, it however shows us what a movie *transmits* along the chain that constitutes it as a historical individual.

6. The Depiction Issue

Another objection can be raised against Carroll's definition, and derivatively against Ponech's and mine as amendments of Carroll's one. It claims that these definitions do not take into account depiction as an essential feature of movies. According to some perceptual theories of depiction, pictures are essentially surfaces that sustain the distinctive visual experience of *seeing-in* (cf. Wollheim 1987; Hopkins 2009), or that trigger the same recognitional abilities that would be triggered by face-to-face experiences of the depicted objects (cf. Schier 1986; Lopes 1996). If moving pictures are pictures, as their name suggests, then the possibility of depiction has to be taken into account in their definition.

There are two different strategies in order to face this objection. First, we could claim that there are moving images that do not depict (for instance Brackage's or McLaren's abstract films), and so the definition of the moving image concerns a more general issue than the depiction issue. Depicting movies are just a subgenre of a cinematic genre that should be defined independently of the features of its subgenre. In order to make this point clear, we should use the expression "moving image" to designate the most general cinematic genre, and "moving picture" to designate the subgenre of the depicting moving images. A similar solution is endorsed by both Carroll and Ponech.

Yet there is another possible solution to the depiction issue, which can be obtained from my account of the movie as a type that specifies a spatio-temporal distribution of pixels. This solution does not need to introduce a

543

distinction between the moving image and the moving picture and to claim that only the latter depicts. In my account all movies basically depict in virtue of their being constituted by a type that specifies a spatio-temporal distribution of pixels. That is because the spatio-temporal distribution of colored points that individuates the content of a movie also characterizes the content of our visual experience – as suggested by Wittgestein's *Tractatus* (2.0251): "Space, time, color (being colored) are forms of objects", and as explained by Goodman's *The Structure of Appearance*:

> If we divide the stream of experience into its smallest concrete parts and then go on to divide these concreta into sense qualia, we arrive at entities suitable as atoms for a realistic system. A visual concretum might be divided, for example, into three constituent parts: a time, a visual-field place, and a color. (1967, p.189)

Every movie, as constituted by a spatio-temporal distribution of pixels, individuates the content of a possible visual experience. Moreover, the movie overcomes the absolute singularity of a possible visual experience, and transforms it in a repeatable type that can have multiple instantiations. That being the case, the distinction between figurative movies and abstract movies only depends on our recognitional abilities and on our conceptual apparatus. It does not matter whether or not a movie triggers our recognitional abilities: for any putative abstract movie, we can always conceive a possible world in which there is a light-energy distribution in an environment that could determine a visual experience whose content corresponds to the spatio-temporal distribution of colored points specified by this very movie.

Ultimately, the depiction issue and the historical issue turn out to be connected. In principle, every movie can work as a depiction, but we need to know the movie's history of production in order to establish *whether* it is used as a depiction and *of which environment* it is a depiction. Provided that a movie can always work in principle as a depiction, in order to effectively work as a depiction the movie has to be embodied in a causal-historical chain that allows the viewer not only to experience the visual structure, but also to infer which is the environment that the movie is intended to represent.

To sum up, both the perceptual accounts of movies' depiction and the pragmatic accounts of movies as historical individuals rely on the structure of the cinematic type. The pragmatic accounts need the type in order to explain what is transmitted along the chain that constitutes the historical individual, whereas the perceptual accounts need the type as a structure sustaining the pictorial experience – in Haugeland (1991) and Kulvicki (2006) terms: a structural "bare bones content" sustaining the recognitional "fleshing out".

In order to establish what a movie effectively depicts we need to consider its history of production and the perceptual skills of its viewers. But in principle a movie depicts independently of its history of production and of its viewers' perceptual skills, since the cinematic type always individuates a spatio-temporal distribution of colored points that corresponds to the content of a possible visual experience.

References

Carroll, Noël (1996): *Theorizing the Moving Image*, New York: Cambridge University Press.

— (2008): *The Philosophy of Motion Pictures*, Oxford: Blackwell.

Davies, David (2009): "Ontology", in Livingston P. and Plantinga C. (eds.), *Routledge Companion to Philosophy and Film*, London: Routledge, 217-226.

Goodman, Nelson (1967): *The Structure of Appearance* (second edition), Cambridge: Harvard University Press.

Haugeland, John (1991): *Representational genera,* in Ramsey W., Stich S. and Rumelhart D. (eds.), *Philosophy and Connectionist Theory*, Hillsdale: Lawrence Erlbaum, 61-90.

Kulvicki, John (2006): *On Images: Their Structure and Content*, London: Clarendon.

Hopkins, Robert (2009): "Depiction", in Livingston P. and Plantinga C. (eds.), *Routledge Companion to Philosophy and Film*, London: Routledge, 64-74.

Lopes, Dominic (1996): *Understanding Pictures*, London: Clarendon.

Ponech, Trevor (2006): "The Substance of Cinema," *The Journal of Aesthetics and Art Criticism*, 64, 187-198.

— (2007): "Cinema Again: A Reply to Walley," *The Journal of Aesthetics and Art Criticism*, 65, 412-416.

— (2009): "Definition of 'Cinema' ", in Livingston P. and Plantinga C., *Routledge Companion to Philosophy and Film*, New York, Routledge, 52-63.

Rohrbaugh, Guy (2003): "Artworks as Historical Individuals", *European Journal of Philosophy*, 11-2, 177–205.

Schier, Flint (1986): *Deeper into Pictures*, Cambridge: Cambridge University Press.

Walley, Jonathan (2007): "On Ponech on the Essence of Cinema," *The Journal of Aesthetics and Art Criticism*, 65, 408-412.

Wittgenstein, Ludwig (1922): *Tractatus logico-philosophicus*, London: Kegan Paul, Trench, Trubner & Co.

Wollheim, Richard (1987): *Painting as an Art*, London: Thames and Hudson.