THE DIGITAL SECRET OF THE MOVING IMAGE

STARTING NOTE

I am grateful to the reviewers of Estetika for their insightful comments. I also wish to thank Trevor Ponech and John Zeimbekis for their detailed answers to my questions about their theories. This paper was originally presented as a talk at the European Society for Aesthetics Conference, Universidade do Minho, Braga, June 26th 2012. In that circumstance, Noël Carroll, David Davies, and Robert Hopkins asked me questions that gave me helpful insights in order to improve my work. An earlier and much shorter version of this paper was published online in the *Proceedings of the European Society for Aesthetics* 4 (2012): 532-546.

I. INTRODUCTION: ONTOLOGY AND DEFINITION OF CINEMA

Cinema, arguably like any other art form, raises two main metaphysical issues. First, an *ontological issue*, concerning the basic ontological category to which a given cinematic work belongs, and second, a *definition issue*, concerning the criteria whereby we can establish whether a given entity is or is not a cinematic work. Addressing the ontological issue for a certain form of art leads us to establish a necessary condition that contributes to addressing the definition of this very form of art. Indeed, if cinematic works belong to a certain ontological category, then a given entity, in order to be a cinematic work, must belong to that same category.

Most authors who have tried to define or characterize cinema have overlooked such an ontological issue. For example, André Bazin argues that cinematic work is a special "mummy," which preserves events instead of bodies.[1] Yet, he does not specify which ontological category such "mummy" should belong to. Likewise, Roman Ingarden characterized the film as "a unique visible music of the transformation of things and of living persons in a spatial world," while Stanley Cavell claimed that the film is "a moving image of skepticism;"[2] yet, neither Ingarden nor Cavell specify what ontological category things like "visible music" or "moving image of skepticism" should belong to.

On the other hand, Walter Benjamin provides us with a useful insight for an ontological characterization of cinema, suggesting that a key feature of cinematic works is their "technical reproducibility."[3] Still, Benjamin aims to highlight social and cultural consequences of technical reproducibility rather than to investigate its ontological underpinnings.

[1] André Bazin, "Ontologie de l'image photographique," in *Qu'est-ce que le cinéma?* (Paris: Les Éditions du Cerf, 1958), 14.
[2] Roman Ingarden, *Untersuchungen zur Ontologie der Kunst: Musikwerk. Bild. Architektur. Film* (Tübingen: Max Niemeyer, 1962), 338. Stanley Cavell, *The World Viewed* (enlarged edition), (Boston: Harvard University Press, 1979), 188.
[3] Walter Benjamin, "Das Kunstwerk im Zeitalter seiner technischen Reproduzierbakeit" in *Schriften* (Frankfurt am Main: Suhrkamp Verlag, 1955).

In the history of the attempts to define or characterize cinema, Noël Carroll was, it appears, the first scholar who made the dependence of definition on the ontological issue explicit.[4] Carroll's definition of what he calls "the moving image" includes an insightful ontological account of cinema. But I believe that the connection between ontology and the definition of the moving image requires further investigation.

With this aim in mind, I shall start by introducing the five conditions that constitute Carroll's definition of the moving image. Some of these conditions treat the moving image as a particular display, while others treat it as a type, a non-particular entity that can be instantiated by particulars (§II). The latter conditions raise an ontological puzzle in the case in which the moving image is a digital movie. In this case, the moving image as a type is instantiated by a digital encoding which in turn is a type; therefore, the digital encoding of a digital movie enigmatically ends up in being both a type and a token. Solving such puzzle leads us to to conceive of the moving image as *a type that specifies a spatiotemporal distribution of pixels* (§III). I shall argue that this new definition can autonomously take into account all those specifically cinematic features which Carroll accounted for by means of his five conditions (§IV, §V, §VI, §VII, VIII). Finally, I shall investigate the key notions of pixel and type in more depth, thereby establishing to what extent an account of the moving image as a type involves an account of cinema as a Platonic form of art (§IX, §X).

II. CARROLL'S DEFINITION OF THE MOVING IMAGE

Carroll addresses the problem of definition by describing five conditions an entity *x* must satisfy in order to qualify as a moving image:

1) "*x* is a detached display."[5] More specifically, the cinematic display consists of a "visual array," and it is "detached" since it provides the spectator with the visual experience of a space which is not connected to her body.[6] The space S presented by the display does not allow the spectator to orient her body with respect to S through an experienced connection between S and her body. The display provides the spectator with a perspective on a space, but this perspective is reduced to a "disembodied viewpoint."[7] Through the display, the spectator experiences a space that is not her space.

[4] Noël Carroll, *Theorizing the Moving Image* (New York: Cambridge University Press, 1996).
[5] Ibid., 70.
[6] Ibid., 61.
[7] Ibid., 63.

2) "*x* belongs to the class of things from which the impression of movement is technically possible."[8] That is to say, the cinematic display is technically produced in such a way that it can provide the spectator with a visual experience of movement.

3) "Performance tokens of *x* are generated by a template that is a token."[9] In claiming that *x* has tokens, Carroll is presupposing that *x* is a type, a non-particular entity that can be instantiated by particulars. Furthermore, Carroll calls "templates" those particulars that instantiate *x* in virtue of their being objects (e.g., film print, videotape, DVD, computer file), while he calls "performance tokens" those particulars that instantiate *x* in virtue of their being events, namely screenings.

4) "Performance tokens of *x* are not artworks in their own right."[10] That is, the screening of a movie, unlike the execution of a symphony or the staging of a play, is not artistically evaluable in its own right. What one can artistically assess is nothing but the movie as a type.

5) "*x* is [...] two-dimensional."[11] That is, the visual array constituting the cinematic display is just a flat surface.

In a later text, Carroll strengthens his definition by arguing that the five necessary conditions are also jointly sufficient.[12] In the previous account, Carroll characterized the conditions as necessary but not jointly sufficient, since he did not intend to include among the moving images some artifacts (such as flip books and the zoetrope) that nonetheless satisfy all the necessary conditions. Yet, in the new account, Carroll changes his mind, thereby treating those artifacts as full-fledged moving images. Thus, he turns his necessary conditions into jointly sufficient ones:

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.[13]

[8] Ibid., 70.
[9] Ibid.
[10] Ibid.
[11] Ibid.
[12] Noël Carroll, *The Philosophy of Motion Pictures* (Oxford: Blackwell, 2008), 53-79.
[13] Ibid., 73.

It is worth noting that conditions (3) and (4) are ontological requirements that establish *what kind of entity* a moving image is, while conditions (1), (2), and (5) rather specify *what further features an entity of that kind* must possess in order to be a moving image. Let us call the former *type-conditions* and the latter *display-conditions*.

Carroll's definition has been criticized both as *too* essentialist by Trevor Ponech and as *not* essentialist *enough* by Thomas Wartenberg. Still, both such criticisms focus mainly on display-conditions rather than on type-conditions. Ponech focuses on condition (1) and argues that the essence of the moving image can be made explicit by revealing the structure of its displays.[14] Wartenberg focuses on condition (5) by arguing that the requirement of two-dimensionality makes the nature of the moving image strongly dependent on our current historical context in which technologies for the production holograms are still not available.[15]

Unlike Ponech and Wartenberg, I shall criticize Carroll's definition by focusing on type-conditions and their relationship to display-conditions. For this purpose, I start by questioning whether the term *x* really refers to the same kind of entity in these two groups of conditions.

On the one hand, in display-conditions, *x* seems to refer precisely to a display, that is, a two-dimensional visual array that portrays a detached space, and may trigger the impression of movement. On the other hand, in conditions (3) and (4) *x* designates the moving image as a type having templates and performances as its tokens. Yet the display, as a visual array, should be a *particular* entity, whereas the type, as such, is a *non-particular* entity.[16] How can the putative moving image *x* be both particular and non-particular?

Of course, it cannot. Since the type unquestionably is a non-particular, the only way to reconcile the moving image as type with the cinematic display is to conceive of the display as the last step in the instantiation of such type. In other words, the display should not be identified with the moving image as such, but rather with the instance of the moving image that Carroll calls "performance token."

Thus, we can rephrase Carroll's definition by claiming that the moving image is a type that can be instantiated by a display that should be (1) detached; (2) capable of triggering the impression of movement; (3) produced by means of a template; (4) non-artistically-evaluable as such; and (5)

[14] Trevor Ponech, "The Substance of Cinema," *Journal of Aesthetics and Art Criticism*, 64 (2006), 191: "I agree that such displays are 'detached.' My reasons go a bit beyond Carroll's, though. […] I identify cinema with the visual display."

[15] Thomas Wartenberg, "Carroll on the Moving Image," *Cinema: Journal of Philosophy and the Moving Image*, 1 (2010), 78, http://www4.fcsh.unl.pt:8000/~pkpojs/index.php/cinema/index: "How do we know now that future developments in the moving image will not affect our willingness to call something a moving image in such a way that the necessary conditions Carroll has laid down will be violated?"

[16] Peter Strawson, *Individuals* (London: Methuen 1959), 231-233. On the one hand, "[p]articulars have their place in the spatio-temporal system, or, if they have no place of their own there, are identified by reference to other particulars which do have such a place." On the other hand, the type is an entity "of which there are many particular instances but which is itself a non-particular."

two-dimensional. Nevertheless, once we turn to the special case of the *digital* moving image, then Carroll's definition, even in such different guise, raises a puzzle, which I will focus on in the following section.

III. THE PUZZLE OF THE DIGITAL TYPE

Carroll claims that the main difference between theater and cinema is that "the play performance is generated by an interpretation that is a type, whereas the performance of the motion picture is generated by a template that is a token."[17] Furthermore, in his 1996 book, he claims that the cinematic template "is a film print, but it might also be a videotape, a laser disk, or a *computer program*."[18] And, in his 2008 book, he specifies that the cinematic template "was a film print, but in recent decades it might be a videotape, a laser disk, a DVD, or an *instantiated computer program*."[19]

I see Carroll's adding of the adjective "instantiated" to the term "computer program" as the clue of a peeping puzzle. If the cinematic template is *a computer program*, as Carroll writes in 1996, then the evidence that a computer program is a *type* which is made of digital symbols contradicts Carroll's aforementioned claim that the cinematic template is a *token*. In his 2008 book, Carroll tries to avoid such contradiction by specifying that, in the digital case, the cinematic template is not a computer program but an *instantiated* computer program. Yet, in the latter case, a new puzzle pops up.

Consider a moving image whose template is an instantiated computer program, or better yet, an instantiated digital file, which consists of traces or circuits. Such a concrete particular is both the template token of a cinematic type C (i.e., the moving image as type) and the token of a digital type D (i.e., the file as sequence of digits). We thus have two types at play. Are those C and D the same type or two different types? In the latter case, what exactly is the relation between C and D?

At a first sight, C and D seem to be completely different types. As pointed out by Nicholas Wolterstorff and Julian Dodd, the cinematic type C specifies what visual qualities ought to be instantiated by showings.[20] By contrast, the digital type D specifies a sequence of digital symbols, so that D is not instantiated by visual showings but rather by physical representations of this sequence of symbols.

---

[17] Carroll, *Theorizing the Moving Image*, 70.
[18] Carroll, *Theorizing the Moving Image*, 67, my emphasis.
[19] Carroll, *The Philosophy of Motion Pictures*, 66, my emphasis.
[20] Nicholas Wolterstorff, *Works and Worlds of Art* (Oxford: Clarendon, 1980), 94. Julian Dodd, *Works of Music: An Essay in Ontology* (Oxford: Oxford University Press, 2007), 16. According to Wolterstorff, the moving image is a type instantiated by "an occurrence of a sequence of illuminated colour-patterns (counting black and white as colours)." According to Dodd, "a film, after all, is just a type whose tokens are datable, locatable showings."

Still, C and D are related to one another, since D specifies the sequence of symbols that allows C to be instantiated by a visual array. A token of D, if coupled with an appropriate device capable of translating the digital symbols into chromatic values, behaves as a cinematic template whereby C is instantiated by a showing.

To sum up, the shift from an analog template to a digital template involves a change in the instantiation of the moving image. In the analog case, the whole instantiation required a two-stage process: first, the moving image was embodied by a template; second, the template was used to generate a screening. In the digital case, the structure of instantiation is more complicated. The moving image is not directly embodied by a physical template, but rather encoded by a digital template which is in turn a type, namely, the digital type. The sequence of symbols specified by such digital type is embodied by a physical particular (made of traces or circuits) which is finally used to generate a screening. Therefore, in the digital case, the whole instantiation of the moving image requires a three-stages process: first, the cinematic type is encoded by a digital type; second, such digital type is embodied by a physical particular; third, such a particular is used to generate a screening.

The mediation of the digital type D between the cinematic type C and its final showing allows us to understand exactly what C itself is. The structure of C, indeed, has to be such that it can be encoded by means of the structure of D. In the digital type D, symbols are placeholders for light values, which correspond to chromatic qualities. I will call such values *pixels*. A pixel, so understood, is not the digital encoding of a light value, but the light value itself.

Furthermore, D is structured in a temporal series of frames, each of which is made of a spatial distribution of pixels. Since the structure of D is aimed to encode the constitutive features of C, should we infer that C consists of a temporal series of frames which are made of spatial distributions of pixels?

Such a question forces us to face the two horns of a dilemma: either a cinematic type C cannot be wholly encoded by a digital type D, or C should have constitutive features that can all be encoded by D. Choosing the first option amounts to claiming that there are movies that cannot be digitally encoded, but this claim seems to contradict our practices concerning movies and their appreciation. For example, the practice of digital restoration of early films would no longer make any sense if old analog films have constitutive features that cannot be encoded by the structure of the digital type.[21]

[21] Scholars like Rodowick or Aumont suggest that analog movies have features that digital movies necessarily lack. I cannot analyze their arguments in this paper, so I limit myself to observing that endorsing such arguments, in the current era in which almost all cinema is becoming digital, amounts to claiming that a relevant portion of the history of cinema is about to disappear. I do not believe so. Digital technology currently guarantees high definition in such a way that it is hard to see how even very fine texture could not be captured by it. See David Rodowick, *The Virtual Life of Film* (Boston: Harvard University Press, 2007). Jacques Aumont, "Que rest-t-il du cinéma?," *Rivista di estetica*, 46

Thus, we are left with the second option. Since the only features that can be encoded by the digital type concern temporal series of frames and spatial distributions of pixels, choosing the second horn of the dilemma amounts to acknowledging that the cinematic type is wholly characterized by temporal series of frames and spatial distributions of pixels. Here is the new definition of cinema. *The moving image is nothing but a type specifying a temporal series of frames that are made of spatial distributions of pixels*. In short, the moving image is a type that specifies a spatiotemporal distribution of pixels.

IV. RETHINKING THE MOVING IMAGE

The digital encoding of a moving image is not a token template, that is, a concrete particular. It is rather an abstract notational structure that reveals the ontological structure of the moving image as a type. With respect to concrete templates like film strips or videotapes, indeed, the digital encoding has an epistemological advantage, that is, it makes the ontological structure of the moving image explicit. In this sense, technology reveals the essence.

The unpacking of Carroll's condition (3), which concerns the cinematic types and its tokens, leads us to a thorough definition of the moving image. But what about the other four conditions of Carroll's definition? How can they be traced back to the ontological structure that is made explicit by the digital type?

As previously pointed out, conditions (1), (2) and (5) are display-conditions in which *x* refers to the display that instantiate the moving image, whereas conditions (3) and (4) are type-conditions in which *x* refers to the moving image as a type that can be instantiated by displays. Addressing the puzzle of the digital type has led us to develop condition (3) so as to define the moving image as a type that specifies a spatiotemporal distribution of pixels, which can be instantiated by displays.

Still, according to Carroll's other conditions, the display that instantiates the moving image has further necessary features; (1) it is detached; (2) it is capable of producing the impression of movement; (4) it is not artistically evaluable; (5) it is two-dimensional. In the following sections I shall argue that all these features of the display can be taken into account in terms of the cinematic type introduced in condition (3). More specifically, I shall consider detachment in section §V, impression of movement in section §VI, two-dimensionality in section §VII, and non-evaluability in section §VIII.

(2011): 17-32.

V. DETACHMENT

Condition (1) of Carroll's definition states that the cinematic display is detached, that is, in watching the display, the spectator experiences things that are not localizable in the spatiotemporal system that has her body as its center. In other words, the moving image supports an experience that allows the spectator to recognize *what* there is, but not *where* she is with regard to what there is.

Indeed, a particular showing of a moving image is not necessarily detached. I can use my web-cam as a mirror while shaving. In this case I can recognized where I am with regard to the displayed scene. The display is necessarily detached only if it is considered as a token of the moving image as a type. Treating the display as a token requires that the scene displayed be able to be shown in a multiplicity of different spaces, without any special connection to the spectators who inhabit those spaces. If we conceive of the display as a token, then it does not matter whether I can shave by looking at my web-cam. From this perspective, the current display of my web-cam is just the token of a type that can be displayed in a multiplicity of other circumstances. What matters is that the generic spectator of a showing of the moving image produced by my web-cam cannot shave by looking at it. In this sense the cinematic display is detached in virtue of its being the token of a type.

Since the repeatability of the moving image as a type allows this image to be replicated in several different spaces, the displayed space cannot have any special connection to the bodies placed in all those spaces. Repeatability necessarily breaks the spatial connection between the displayed space and the space of the audience. If we consider the display as a particular event, nothing prevents the displayed space from being connected to the space that the beholder inhabits. What makes the displayed space necessarily detached is the repeatability of the moving image as a type. The only token that is connected to the spectator's own space is arguably the particular display that instantiates the moving image during its production, as in the case of a web-cam used as a mirror. But if the display is considered as any token of an existing moving image, then it is necessarily detached from the space of the spectator.

From this perspective, all the "prosthetic devices"[22] (e. g. mirrors, microscopes, telescopes) that Carroll attempts to distinguish from the moving image by means of the detached display condition can be distinguished much more simply by considering that they do not have a type-token ontological structure. Mirrors, as well as other glass-based prosthetic devices, cannot be repeated. They are nothing but visible particulars. They are not tokens of types that specify visual features. There is no type of which several mirrors, being displays, are all tokens. Any mirror displays only its own space. The basic ontological difference between mirrors and moving images is that the

[22] Carroll, *Theorizing the Moving Image*, 57.

former are just particulars whereas the latter are tokens of types. The fact that moving images are detached display whereas mirrors are not simply follows from such difference.

VI. IMPRESSION OF MOVEMENT

In formulating condition (2), Carroll considers the spectator's disposition to believe that what the display presents might move, instead of giving the mere impression of movement, because he wants to take into account works such as Marker's *La jetée*, Frampton's *Poetic Justice* and Snow's *One Second in Montreal*, made, partly or wholly, by still images. According to Carroll, these works are different from a mere slide show because spectators of the former can legitimately expect (at least at the first viewing) that sooner or later there will be some movement in the pictures. As Carroll puts is, "it is always justifiable to entertain the possibility that the image *might* move."[23]

I also argue that this condition can derive from the structure of the cinematic type that the digital encoding makes explicit. "To entertain the possibility that the image *might* move" is indeed "always justifiable" because the moving image as a type consists of a series of frames whose temporal rate is capable of affording the impression of movement to our perceptual system. The epistemological possibility (we know that there could be movement) is based upon an ontological possibility (the type consists of a series of frames, so it is capable of affording the impression of movement to us). Paintings and photographs can not move since they consist of a *spatial* distribution of colored points, whereas the moving image can move (and spectators believe it can) since it consists of a *spatiotemporal* 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 does not consist of a series of frames because it can move: it can move because it consists of a series of frames. That is why "movement is a permanent possibility in cinema."[24] 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.

From this perspective, sound can also be treated as a permanent possibility in cinema, for the same reason as movement.[25] Both sound and movement unfold in time. Therefore, in order to have auditory features, a work must unfold in time. Since the moving image, as a spatiotemporal distribution of pixels, takes place in time, it carries the possibility of sound in its own structure. As the moving image can afford the impression of movement to spectators, so this image can afford the impression that some sound is synchronized with (or, at least, somehow connected to) what is displayed. From this perspective, silent films such as Kaurismäki's *Juha* or Hazanavicius' *The*

[23] Carroll, *The Philosophy of Motion Pictures*, 60.
[24] Ibid., 60.
[25] I owe this insight to one of my referees.

Artist, which lack *sound* because of stylistic choices (instead of technical limitations), function similarly to "static films"[26] such as Marker's *La jetée* or Godard and Gorin's *Letter to Jane*, which lack *movement* because of stylistic choices. In this sense, the unexpected movement of an eye in a scene of *La jetée* exploits the spectator's attitude to entertain the possibility that the image *might move* even in a static film, as well as the unexpected sound of a glass in a scene of *The Artist* exploits the spectator's attitude to entertain the possibility that the image *might sound* even in a silent film.

Still, if the moving image ultimately is a spatiotemporal distribution of pixels, how can we distinguish moving images from slide shows? Slide shows are also constituted by series of frames, that is, spatiotemporal distributions of pixels. How do they differ from moving images? I argue that slide shows belong to an ontological category that is in between still images and moving images. More specifically, the slide show provides us with an experience that is akin to the experience of still images, but rests upon an ontological structure that is akin to the structure of moving images.

The difference between the slide show and the moving image is basically a matter of frame rate. Below a certain threshold rate R1, the series of frames is experienced as a series of still images. Above a certain threshold rate R2, the series of frames is experienced as a moving image (likewise, a series of musical notes can be heard as a continuous melody only if such notes are played at a rate that is above a certain threshold). In between R1 and R2, the series of frame is experienced as a jerky image, that is, an image that is no longer still but not yet moving.

In spite of lacking the possibility of the impression of movement, slide shows exhibit a distinctive temporal mood. Like moving images, and unlike mere books of images, slide shows can be synchronized with sounds. Yet, the frame rate of a slide show, unlike that of a moving image, is not necessarily established by the maker or by the practice, but can be up to the presenter.[27]

To sum up, Carroll's conditions (1) and (2), which describe the relation between the cinematic display and the experience of the spectator, can be explained in terms of structural features of the moving image as a type. Cinematic displays afford detachment and impression of movement to the spectator because they are tokens of the moving image, which is a type constituted by a series of frames made of pixels. Detachment and impression of movement are phenomenal and epistemic consequences of the ontological structure of the cinematic type, as well as being transparent and drinkable are phenomenal and epistemic consequences of the chemical structure of water.

[26] Ibid., 61.

[27] A quite impressive example of a slide show whose frame rate is up to the performer can be found in the so-called "Kodak Carousel scene" of the TV series *Mad Men* (Season 1, episode 13; the scene is available on YouTube with the title "Mad Men – The Carousel").

VII. TWO-DIMENSIONALITY

Carroll introduces the requirement of two-dimensionality in order to exclude "moving sculptures of the sort that are exemplified by the moving figurines on various antique clocks" from the domain of the moving image.[28] This sort of moving sculpture is akin to the moving image to the extent that they both present a detached space in which we can see movement. Furthermore, they have both multiple instances that are produced from a template.

Still, moving sculptures are three dimensional whereas moving images are two-dimensional.[29] The requirement of two-dimensionality is thus sufficient to exclude moving sculptures from the cinematic domain. On the other hand, Carroll acknowledges that this requirement is not sufficient to cleave the moving image from theater. That is because "there is, in fact, theater that is two-dimensional, for example the shadow-puppet shows of Bali, Java, and China."[30]

In order to cleave motion pictures from shadow-puppet shows, Carroll exploits his condition (3), namely the type-template-performance condition. But his solution raises a question that he does not explicitly consider. What about a shadow-puppet show made by means of a moving sculpture? Since this sort of shadow-puppet show satisfies both the two-dimensionality requirement and the type requirement, we should conclude – against our intuitions – that it *is* a moving image. Thus, the case of the moving sculpture has not really been explained away by the two-dimensionality condition. If we use moving sculptures in a shadow-puppet show, the problem pops up again.

A supporter of Carroll's definition might by replying such a manufactured shadow-puppet show *is* in fact a moving image since it satisfies all Carroll's conditions. Indeed, Carroll himself seems to invoke a similar argument in his 2008 book, when he claims that manufactured flip books *are* moving images, whereas handmade flip books *are not*.[31]

I think that this is an *ad hoc* reply, and a quite unsound one. Indeed, we normally conceive of both handmade images like paintings and manufactured images like photographs or prints as *still images*. Why should we behave differently in the case of *moving images*? Why the handmade/manufactured divide, which is not relevant in order to establish whether something is or is not a still image, should become relevant when we establish which shadow-puppet shows (or flip books) are moving images and which are not? I see no reason to treat manufactured shadow-puppet shows (or manufactured flip books) as moving images while treating handmade shadow-puppet

[28] Ibid., 72.
[29] By claiming that moving images are two-dimensional, Carroll means that their screenings take place onto flat surfaces. Of course, what is seen by spectators in such two-dimensional surfaces can (and usually does) consist of three-dimensional scenes.
[30] Ibid., 73.
[31] Ibid., 75.

shows (or handmade flip books) as images of a different kind. The only reason I can see is the defense of Carroll's definition of the moving image.

If you want to include shadow-puppet shows and flip books in the domain of the moving image, you should include *all* shadow-puppet shows and *all* flip books in this domain. Thus, by accepting that manufactured shadow-puppet shows are moving images, Carroll would put his definition on a slippery slope leading to the conclusion that *all* shadow-puppet shows are moving images. In this way, Carroll's definition would be reduced to a variant of Berys Gaut's account, according to which any object-generated image that exhibits movement counts as a moving image, so that even "Plato's parable of the cave in the Republic would also count as a kind of object-generated cinema."[32] Yet Gaut's account of cinema raises two problems that, I argue, are much more puzzling than the problem of manufactured shadow-puppet shows raised by Carroll's definition.

First, if you want to preserve the intuition according to which "still movies" like *Letter to Jane* are moving images, then you should accept any object-generated image that *might* exhibit movement into the domain of the moving image. Such a domain, which in Gaut's account is already immense, is further extended. Even the shadow on the wall that is generated by my table is a moving image, since my table might move, and therefore its shadow *might* move.

Second, in spite of our shared understandings, in Gaut's account cinema is no longer something that has been invented towards the end of the 19th century, but rather something almost as old as the wheel or the knife. From such a perspective, to treat the Lumière Brothers as the inventors of cinema would be as wrong as to treat Gutemberg as the inventor of writing. Indeed, I agree that it is arguable whether the Lumière Brothers really invented cinema. In cinema museums, you can find many screening devices made in the 19th century that somehow anticipate the Lumières' *Cinématographe*. Yet I do not know of any cinema museum in which shadows of tables are exhibited as examples of cinema. In this sense, Gaut's ontology of cinema is too hospitable, even up to the point of contradicting established practices, intuitions, and judgments concerning cinema.

If you want to avoid such a slippery slope towards too hospitable an ontology of cinema, you should find a safer way than Carroll's to prevent moving sculptures from counting as moving images. For this purpose, it is worth setting Carroll's condition (5) (i.e. the two-dimensionality requirement) aside, and looking once again to his condition (3), that is, to the structure of the moving image as a type. 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 spatiotemporal

[32] Cf. Berys Gaut, *A Philosophy of Cinematic Art* (Cambridge: Cambridge University Press, 2010), 6.

distribution of visual qualities, but it rather includes properties such as height, weight, and chemical composition. 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 difference in the ontological structure also explains why the moving image, which is just made of visual (and possibly auditory) qualities, provides us with the *impression of movement* whereas the moving sculpture, which has further physical and chemical features, provides us with *true movement*. Furthermore, the distinctive ontological structure of the moving image as a type allows us to take into account the possibility of holography, whose affinity with the moving image is suggested by Carroll himself: "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?"[33]

We need a criterion to distinguish holographic screenings, which our intuition is prone to consider as moving images, from moving sculptures, which we aim to exclude from the cinematic domain. The two-dimensionality condition fails to support this distinction, since Carroll is forced to exclude both moving sculpture and holography from the cinematic domain. But if we pose as a criterion the ontological structure of the type, then we can treat holograms as peculiar cinematic types whose frames are *three-dimensional* distributions of pixels, instead of two-dimensional ones like in ordinary movies. In this way, we can substantially exclude from the cinematic domain only the moving sculptures, whose types are not spatiotemporal distributions of pixels at all, while preserving the cinematic nature of holograms in spite of their three-dimensionality. In short, the ontological structures of types explains why the hologram is not different from the moving image as the moving sculpture is.[34]

VIII. NON-EVALUABILITY

Carroll's condition (4) claims that cinematic tokens lack artistic value. I argue that such lack follows from the fact that all you need in order to instantiate a moving image is already specified by the corresponding type, and no further contributions are required. You only need to instantiate the visual (and auditory) patterns constituting a moving image as a type, that is, the patterns established by the makers of the image. Such instantiation can be done with merely automatic processes, in which there is no room for human intentionality and creativity, and, *a fortiori*, for artistic value. By

[33] Carroll, *The Philosophy of Motion Pictures*, 73.

[34] Shadow-puppet shows of any kind also differ substantially from moving images since the former are not individuated by discrete spatiotemporal distributions of frames and pixels but rather by continuous spatiotemporal distributions of light and darkness.

virtue of the cinematic type's ontological structure, once the makers of a moving image constitute it as a type, the subsequent displays will be just a matter of automatic technical processes.

On the other hand, theatrical performance is an interpretation that involves intentional acts. The reason for the sharp difference between cinema and theater lies precisely in the structure of the type. The theatrical work as a type is just a written text, so in principle it cannot specify all the perceptible properties that should constitute an experienceable instance of the work. In order to turn such a text into a theatrical production you need some creative act of interpretation. By contrast, the cinematic type, as a spatiotemporal distribution of visual qualities, specifies all the relevant experienceable features that constitute an instance of the moving image, leaving no room for interpretation and preventing the screening of a moving image, unlike the presentation of a theatrical performance, from being evaluated as a work of art in its own right.

The structure of the cinematic type, conceived as a spatiotemporal distribution of pixels, is such that, at the token level, it only remains to make those distributions accessible to the viewers. Of course, we can distinguish between better or worse screenings. But such normative distinction only concerns technical procedures. Some displays may correctly instantiate the visual pattern specified by the type (as in the case of high definition copies and high quality projectors), while others may have poor quality (as in the case of VHS copies or old and worn projectors). But the difference between a good and a bad screening of a moving image is not a matter of interpretation at all, let alone of artistic value. It is only a matter of technical approximation to the visual appearance wholly specified by the moving image as a type.

IX. PIXELS

So far, I have argued that the moving image is *a type specifying a spatiotemporal distribution of pixels*. In these final sections, I shall focus on the fact that not only digital movies but moving images in general are types specifying spatiotemporal distributions of pixels. For this purpose, I shall analyze in more depth the key notions of pixel and type.

According to Ponech, pixels are particular points of light:

> '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.[35]

[35] Ponech, "The Substance of Cinema," 191-92.

Pixels, so understood, are constitutive elements of both digital and analog displays. The only difference is that the points of lights constituting a digital display are arrayed in a grid whereas in analog displays they are not. In the latter case, indeed, the spatial distribution of pixels matches the irregular – but nevertheless discrete – distribution of individual grains on the film strip.

The cinematic display is essentially *discrete*, that is, consisting of particular points of lights separated by temporal and spatial interstices. Discreteness is sharply exhibited by the digital display, whose pixels are regularly distributed in a grid. Discreteness also characterizes analog displays, whose pixels are separated by spatial interstices in spite of the lack of a regular grid, since every frame consists of individual grains of color. Furthermore, both analog and digital pixels belong to frames that are separated from each other by temporal interstices, since analog and digital projectors both show a limited number of frames per second, and since the film of analog projectors alternately blocks and reveals light.

The limit of Ponech's account is the conception of the pixel as a *particular* point of light. The pixel, so understood, can only concern a *particular* display, that is, a *particular* showing of a moving image. It follows that any new showing involves a completely *new* series of pixels on the screen. Therefore, Ponech's characterization of the moving image in terms of pixels cannot take into account the moving image as a repeatable work that can be instantiated by a multiplicity of showings. If the moving image is made of pixels that are no more than particulars, then there is no way to appropriately relate the screening of *Behind the Candelabra* that I watched in London to another screening of the *same* movie that took place in Los Angeles. Therefore, I agree with Ponech that pixels are "the substance of cinema,"[36] but I argue that in order to take repeatability of cinematic works into account we should conceive of the pixel not as *a particular point of light*, but rather as *a value of light* which can be instantiated by a multiplicity of particular points of light.

If all of this is right, both digital and analog moving images are types that specify pixels. The only difference is in the way in which the cinematic type specifies the pixels that should be instantiated by displays. In the analog case, pixels are implicitly specified by means of concrete templates, as for example reels of celluloid that allow us to instantiate roughly the same spatiotemporal distribution of light values at each showing of a given moving image. By contrast, in the digital case pixels can be explicitly specified as light values, because the template is no longer a concrete object but a series of numbers denoting light values.

In a similar vein, Gaut accounts for digital cinema by conceiving of the pixel as a discrete unit that measures "the light intensity […] as a discrete integer."[37] Yet he challenges the claim that, in digital pictures, the pixel is a "minimal denotative unit" by claiming that "the parts of a pixel

[36] Ibid., 187.
[37] Gaut, *A Philosophy of Cinematic Art*, 57.

denote the parts of the area of the object that the pixel denotes […] The denotation relation still holds at the sub-pixel level. The parts of a pixel do denote, unlike the parts of a word."[38] That is to say, if we look closely at a pixel on the screen, then we can see a small colored area whose colored parts in their turn denote.

I argue that Gaut's argument is wrong, since what we truly see in looking closely at the screen is not the pixel itself, but the particular spot of light that instantiates the light value constituting the pixel. This spot of light is seen as a small colored area having colored parts, but the pixel instantiated by this token is a light value having no parts at all. By claiming that in digital pictures the pixel is not a minimal denotative unit, Gaut seems to mistake the ontological nature of the pixel with the empirical fact that the particular spots of light instantiating pixels are not usually recognized by viewers as minimal denotative units. But, in digital pictures, the pixel *is* a *minimal denotative unit* since it is not a particular spot of light but rather a light value that *denotes* the light intensity in a precise spatiotemporal location.

In challenging the claim that in digital pictures the pixel is a minimal unit, Gaut also argues that:

> The digital photograph is not […] different from a traditional photograph. For the latter is comprised of sometimes billions of individual grains […] In this respect there is also an array of picture elements in the traditional photograph, albeit one with vastly more elements than is usual in digital photographs, and which are not arrayed in a grid. Keep on enlarging such a photograph, and in the end one will see individual grains, from which the object is not recognizable, even though the grains denote parts of the object.[39]

Still, those considerations do not necessarily show that pixels are not minimal denotative units. Rather, they seem to show that also traditional photographic pictures have minimal denotative units, namely grains, which play the role of pixels, in spite of the lack of a notation capable of representing them. Let us consider, in this sense, the example proposed by Gaut:

> This is the lesson of Michelangelo Antonioni's *Blow-Up* (1966): as Thomas, the photographer played by David Hemmings, keeps enlarging the image that he thinks shows a murder, the grains of film become more prominent and it becomes impossible in the end to tell what they denote.[40]

If, as Gaut writes, "the grain become more prominent," then the grain has not to be identified with the area it occupies on the paper (or on the screen, if we move from the photographic

[38] Ibid., 58.
[39] Ibid., 59.
[40] Ibid., 59.

example to a hypothetical cinematic counterpart of it). An area that becomes *more prominent* is no longer the same area. The grain has rather to be identified with a light value, which could be instantiated by *more or less prominent* areas on the paper (or on the screen). Moreover, in developing the *Blow-Up* case, Gaut claims that, after enlarging the image, "it becomes impossible in the end to tell what [the grains] denote." Yet, in so doing, Gaut conflates two issues: what a grain denotes, and what the picture depicts. Indeed, after enlarging the image, it becomes impossible in the end to tell what the picture depicts, but not what the grains denote. Even enlarged, a grain of a photograph still denotes the light intensity in a precise spatiotemporal location, and that is why the *Blow-Up* photographer keeps on analyzing this photograph with the purpose of understanding what the picture depicts. Grains in analog pictures have to be ultimately identified with light values, or pixels, whereby these pictures depict their subjects.

X. TYPES

In sections from §III to §VIII I have argued that the moving image is *a type* specifying a spatiotemporal distribution of *pixels*. In section §IX I have investigated what a pixel is. In order to seal this definition of the moving image, it only remains to investigate what a type is.

Following Wolterstorff and Dodd, I argue that the moving image is a *normative* type, that is, a type that establishes what visual features a *correct* instance of such image *ought to* exhibit.[41] As normative types, moving images can have two sorts of instances; *incorrect* ones (which only possess some relevant subset of the normative features), and *correct* ones (which possess all the normative features). Conceiving of the moving image as a normative type allows us to take into account an indispensable aspect of our cultural practices, that is, the fact that we usually assess not only cinematic works ("this is a good movie, that is a bad one...") but also instantiations of these works ("this is a faithful screening, that is a flawed one."). Carroll's condition (4) points out that the former is an artistic assessment whereas the latter is rather a technical assessment. Still, they both are assessments, which rest upon some form of normativity. More specifically, the artistic assessment rests upon some standard of taste, whereas the technical assessment rest upon the work itself, understood as a normative type that establishes the standard for its correct screenings.

According to Wolterstorff and Dodd, treating a type as normative obliges us to treat this type as a Platonic universal. Therefore, movies or symphonies, as *normative* types, are "abstract, fixed, unchanging, and eternally existent entities."[42] Yet, David Davies challenges the claim that conceiving of moving images as normative types commit us to such a counterintuitive Platonic

[41] Wolterstorff, *Works and Worlds of Art*, 94. Dodd, *Works of Music*, 16.

[42] Dodd, *Works of Music*, 36.

view, according to which movies are not created but discovered.[43] He argues that we can conceive of the moving image as a normative type by considering the Wittgensteinian account of normativity developed by Robert Brandom: "a *pragmatist* conception of norms – a notion of primitive correctnesses of performance *implicit* in *practice* that precede and are presupposed by their *explicit* formulation in *rules* and *principles*"[44].

In such an account of normative types, what establishes whether a particular display D is a correct instance, a flawed instance or a non-instance of a given cinematic work W is not an explicit list of light values residing in the abstract space of the universals where they are grasped and made normative by the filmmaker. Instead, the status of D as an instance of W depends upon an implicit negotiation between two parties: what the filmmaker specifies in making his or her work W public in a given cultural context, and the practices that implicitly establish which displays are fully qualified to play the experiential role in the appreciation of W. According to Davies, there are no explicit rules of correctness for the instantiation of a given moving image W. Practices sanction conditions of appreciation of W, and cinematic displays are technically assessed with respect to the role they play in appreciation of W, but nothing more.

I agree with Davies that normativity of the moving image as a type rests upon cultural practices rather than upon some abstract, fixed, unchanging, and eternally existent entity. Yet, I think that, in this respect, the case of digital cinema requires special treatment, since digital technology allows us to produce instances of a given image (regardless of its being static or moving) that are all "phenomenally identical in respect of color, shape, and size."[45] That is because digital technology provides us with a sort of notation whereby we can represent an image W by means of a *script* S of discrete symbols which denote the points of light constituting W. Thus, in order to instantiate W, we only need some device capable of translating the discrete symbols constituting S into the points of light constituting W. If these devices are properly functioning, then all instances of W produced by coupling such devices with S are phenomenally identical.

In principle, digital cinema enables the filmmaker to establish once and for all the only way in which correct instances of her work should appear. Specifying a moving image W by means of a digital script, indeed, amounts to unequivocally establishing the appearance of any correct instance of W. In this way, the standard of correctness for the instances of W is no longer implicit in practice but made explicit by means of the digital script. As a thought experiment, we can even conceive of

[43] David Davies, "What Type of 'Type' is a Film?," in *Art and Abstract Objects*, ed. C. Mag Uidhir (Oxford: Oxford University Press, 2013), 263-83.
[44] Robert Brandom, *Making it Explicit: Reasoning, Representing, and Discursive Commitment* (Cambridge: Harvard University Press, 1994), 21.
[45] John Zeimbekis, "Digital Pictures, Sampling, and Vagueness: The Ontology of Digital Pictures," *Journal of Aesthetics and Art Criticism*, 70 (2012), 51. It is worth noting that, according to Zeimbekis, the digital encoding of pictures is a notational schema but it is not a full-fledged Goodmanian notational system since it lacks Goodman's semantic requirement of finite differentiation.

special digital devices (perhaps embedded in mobile phones) with which moviegoers can check whether the movie they are watching is shown correctly. Such devices could measure light values on the screen and compare them with the original light values approved by the filmmaker and stored in some online database.

If all of this is right, should we conclude that digital technology turns moving images into everlasting Platonic entities? I do not think so. Digital technology is something that was created in the context of our cultural practices. For this reason, what a moving image becomes thanks to digital technology still rests upon our practices.

Nevertheless, digital technology seems capable of supporting what we could call a *Platonic practice*, that is, a way of univocally establishing the appearance of a work that no longer depends on metaphysical virtues, but rather on technical devices. Such devices are a necessary condition for the rise of a Platonic practice, but not a sufficient one. Agreement is also required. Therefore, cinema can become a Platonic practice, namely *Platonic cinema*, only if practitioners agree that all correct screenings of a moving image W ought to be phenomenally identical by complying with the pixels specified once and for all by the maker of W.

Currently, we have a technique allowing for Platonic cinema, but we do not yet have a practice establishing Platonic cinema as the cinematic medium in force. In spite of the fact that digital technology enables us to make moving images explicit in terms of unique series of pixels, we keep relying on the implicit normativity of cultural practices in order to constitute movies as spatiotemporal distributions of light values. Yet, if we want to be sure of transmitting not just instances but *correct instances* of our movies to future generations, then Platonic cinema, which our technology already enables in principle, is the right way forward.

BIBLIOGRAPHY

Aumont, Jacques. “Que rest-t-il du cinéma?” *Rivista di estetica*, 46 (2011): 17-32.

Bazin, André. “Ontologie de l’image photographique,” in *Qu’est-ce que le cinéma?* Paris: Les Éditions du Cerf, 1958. English translation: “The Ontology of the Photographic Image.” *Film Quarterly* 13, 4 (1960): 4-9.

Benjamin, Walter. “Das Kunstwerk im Zeitalter seiner technischen Reproduzierbakeit” in *Schriften*. Frankfurt am Main: Suhrkamp Verlag, 1955. English translation in *The Work of Art in the Age of its Technological Reproducibility, and Other Writings on Media*. Boston: Harvard University Press, 2008.

Brandom, Robert. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard University Press, 1994.

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

Carroll, Noël. *The Philosophy of Motion Pictures*. Oxford: Blackwell, 2008.

Cavell, Stanley. *The World Viewed* (enlarged edition). Boston: Harvard University Press, 1979.

Davies, David. "What type of 'type' is a film?" In *Art and Abstract Objects*, edited by Christy Mag Uidhir, 263-83, Oxford: Oxford University Press, 2013.

Dodd, Julian. *Works of Music: An Essay in Ontology*. Oxford: Oxford University Press, 2007.

Gaut, Berys. *A Philosophy of Cinematic Art*. Cambridge: Cambridge University Press, 2010.

Ingarden, Roman. *Untersuchungen zur Ontologie der Kunst: Musikwerk. Bild. Architektur. Film*. Tübingen: Max Niemeyer, 1962. English translation: *The Ontology of the Work of Art*. Athens: Ohio University Press, 1989.

Ponech, Trevor. "The Substance of Cinema." *The Journal of Aesthetics and Art Criticism*, 64 (2006): 187-98.

Rodowick, David. *The Virtual Life of Film*. Boston: Harvard University Press, 2007.

Strawson, Peter. *Individuals*. London. Methuen 1959.

Wartenberg, Thomas. "Carroll on the Moving Image." *Cinema: Journal of Philosophy and the Moving Image*, 1 (2010): 69-80, http://www4.fcsh.unl.pt:8000/~pkpojs/index.php/cinema/index.

Wolterstorff, Nicholas. *Works and Worlds of Art*. Oxford: Clarendon, 1980.

Zeimbekis, John. "Digital Pictures, Sampling, and Vagueness: The Ontology of Digital Pictures." *Journal of Aesthetics and Art Criticism*, 70 (2012): 43-53.