Aisthesis

Firenze University Press
www.fupress.com/aisthesis

OPEN ACCESS

Citation: W.-L. Moon (2018) Some Photographic Images Are Transparent. *Aisthesis* 11(2): 23-33. doi: 10.13128/Aisthesis-23113

Copyright: © 2018 W.-L. Moon.This is an open access, peer-reviewed article published by Firenze University Press (http://www.fupress.com/aisthesis) and distribuited under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The authors have declared that no competing interests exist.

Some Photographic Images Are Transparent

Won-Leep Moon
(Dongguk University)
wlm@dongguk.edu

Abstract. Kendall Walton argued that photographs are transparent, that we literally see things through them. This claim provoked many objections, and one line of argument has focused on the fact that when we see objects in ordinary situations we see their approximate location with respect to us, whereas in typical photographs we do not. The author argues, however, that this egocentric spatial information is not what distinguishes literal seeing from typical photograph seeing. Instead of it, the author proposes two conditions for normal, literal seeing. One is that the seeing be real-time, and the other is that the image be "empty". Some photographic images meet these conditions.

Keywords. Kendall Walton, photographic Images, transparency.

1. INTRODUCTION

In his well-known article *The Ontology of the Photographic Image* André Bazin asserted: «The photographic image is the object itself» (Bazin [1971]: 14). In a similar vein, Kendall Walton famously declared that photographs are *transparent*. What he meant is that «we *see*, quite literally, our dead relatives themselves when we look at photographs of them» (Walton [1984]: 252). For Walton, photographs are on a par with mirrors and other prosthetic devices like telescopes and microscopes in their ability to help us see things. This claim was a rather bold one and provoked many objections. Among them one line of arguments has been prominent. It focuses on the fact that when we see an object in ordinary situations we also perceive its approximate location (direction and distance) with respect to us, whereas in typical photographs we do not. Gregory Currie ([1995]: 48-78) and also Noël Carroll ([1996]: 55-63) have argued that for this reason we cannot say we literally see things through photographs. Knowing the egocentric location of the objects is a necessary condition of seeing, they say.

Walton (1997) gave a counterexample to this argument, which I think is a legitimate one. If you see a (reflected) flower in a room

full of mirrors, you would not know where it actually is. But isn't it still true that you see, literally, the flower? I can offer a variation of this. Think of a boy who sees a fish in a creek. If he does not know that the fish appears shallower than it actually is, Currie and Carroll would have to say that he does not see it, which seems counterintuitive. They may reply that he knows its location at least roughly. Perhaps he does, but the misperception in this case is not due to the inevitable imprecision of perception which we must tolerate. Instead, it is due to a consistent natural phenomenon, which the boy can and should learn.

Jonathan Cohen and Aaron Meskin (2004) avoid Walton's counter by proposing a nondoxastic solution. They retain Currie's and Carroll's original insight that spatial information is the key, but instead of the viewer's subjective belief about the egocentric location of the object, the image's objectively carrying the information was taken as the criterion of seeing. Bence Nanay (2010) has also proposed a related theory. In this paper, however, I try to show that spatial relation between viewer and object is irrelevant for seeing[1]. My argument will be primarily directed to Cohen and Meskin and to Nanay. After that, I argue for alternative necessary conditions. One is that the perception be real-time; another is that the image be «empty». Finally I show that some photographic images meet these conditions.

One thing needs clarification. Walton (2008) said that his aim is not conceptual analysis, but that is what I will do here. Insofar as he uses an existing word («see»), it cannot be immune to conceptual analysis. True, he said that he does not worry about whether what he proposes is «a new sense of the word», and does not mind using a new word like *schmee* instead of *see* ([2008]: 111), but I doubt he can be so indifferent. If it *is* a new sense or word, as Berys Gaut says, «the claim that we see through photographs would then lose most of its interest» ([2010]: 90). Just think of the sentence «We literally schmee our dead relatives through photographs», where *schmee* is similar to, but not the same as, *see*. I will take the transparency claim as involving the ordinary sense of *see*, as most disputants did. With his transparency claim, Walton wanted to explain photographic realism—the sense of transparency—and to that extent his indifference to the ordinary sense might be justified in a way. But the subjective realism is irrelevant in this paper[2]. I am interested in the (objective) transparency claim itself, and will try to resolve the controversy surrounding it.

2. SPATIAL RELATION IS NOT THE KEY

Cohen and Meskin give an example of X's carrying information about Y: «The state of the room's thermometer carries information about the temperature in the room insofar as there is an objective probabilistic connection between the two» ([2004]: 200). Here, the carrier of information is the «state» of the thermometer, specifically, the thermometer's reading. As for the egocentric spatial information (ESI), although they sometimes say that the whole device (mirror, telescope, etc.) carries it, what actually changes with changing information seems to be the «image». They write, «change in an object's egocentric location would bring about change in the (mirror-produced) image» ([2004]: 203). I will call this condition MCV (Movement Changes View)[3]. This condition is not met for photographs: «As I move around the world with the photograph, the egocentric location of the depictum changes, but the photographic image does not» ([2004]: 201).

[1] Helen Yetter-Chappell (2017) recently made a persuasive argument against this spatial relation, but her focus is on the detachability of eyes. My paper may be regarded as offering additional arguments.

[2] I think the strong transparency claim is based on a misunderstanding of this realism which, I believe, is just a psychological effect. Scott Walden (2016) is correct to argue that the realism, the «contact phenomenon», is not necessary for seeing. Also see Moon (2018).

[3] Cohen and Meskin use «image» ambiguously. If I move in front of a mirror, the mirror image changes in a sense, but in another it does not. What changes with my movement is the perspective view of the mirror image. Hence «view» is more appropriate here.

I do not think that requiring ESI is equivalent to requiring MCV, but for now I concentrate on MCV. Nanay's thesis demands a weaker variant of MCV. For him, the view does not need to change with every possible movement. What is necessary for seeing is that «there is at least one way for the perceiver to move such that if she were to move that way, her view of the perceived object would change continuously» ([2010]: 468). For my purpose, though, this difference is not significant.

It is true that if you move in front of an object, normally your view of it changes. But the same is true for photographs. Think of a photograph of an apple whose image is a circle on the surface. If you move in front of it, your view of it will change. For instance, the apple (the circle) may look elliptical if seen obliquely. But perhaps this is not what is meant by MCV. Nanay argues that what changes in this case is the view of the *photograph itself*, not of the apple. That is, you do not get to see a different side of the apple. I can grant that, but what if the image is 3D? Think of a holographic image of the apple. In this case, if you move, you do get to see a different side of it. But as Nanay himself admits ([2010]: 472, 474-475), this does not seem to make the image transparent. I think we agree that holographic images in general are not transparent, but even if they were, it would not be due to their three-dimensionality. What this shows may be only that MCV is not *sufficient* for seeing, but still it is a weakness of the theory. Mirror images' transparency is not explained by their satisfaction of MCV.

Now I will argue that MCV is not *necessary* for seeing. Imagine an object that is invisible («transparent» in the ordinary sense) except from exactly one viewpoint. In this case there would be *no* movement that produces a different view, but still we would have to say that we see it when we are at that point. It seems that seeing a view of an object is sufficient for seeing it, and MCV is not required. Proponents of MCV may object that disappearance of view *is* a change in view. Perhaps it is, but it is certainly not a *continuous* change which is important for Nanay. Besides, the «change» is not sensitive to the direction or amount of movement. Whatever way you move, the «change» is the same. Presumably this is not what is meant by MCV. Nanay considers a related case, where the view remains the same after movement (rather than disappears), and says that this is like an afterimage and so is not seeing. But this does not necessarily mean that you did not see the object *before* the movement, before you see an afterimage. Moreover, suppose that the image keeps updating, that is, it is not an *after-* but *current*-image of the same side. Then I find no reason why we cannot say we keep seeing it. (This points to the importance of real-time seeing which I will discuss later).

There is an interesting difference between the two counterexamples I have presented, however. Both, as I believe, are counterexamples to MCV, but only the second is clearly one to ESI. In the first case of visible-only-from-a-point object, if I am at that point I would normally see the object's rough egocentric location. It would be odd to say that I do not see the location simply because other sides are invisible. When I am not at the point, on the other hand, I would not need to worry about my view's lacking ESI, for there would be *no* view of the object. Thus this case is not a clear counterexample to the necessity of my view's having ESI. (In the second case where my view remains after movement, though updating, I do have a view that does not carry ESI. So this can be a counterexample). Hence my earlier remark that ESI is not equivalent to MCV.

The two examples mentioned are all very imaginary. In ordinary cases of direct seeing, probably there is no counterexample to either MCV or ESI. For ordinary cases of *indirect* seeing, however, necessity of ESI can be easily refuted. (The same is not quite true for MCV). Take the case of mirror images. These images can seem to carry ESI. For example, one may point to the part «mov[ing] around the world with the photograph» quoted before from Cohen and Meskin, and say that we cannot do the same with a mirror image. But we *can* do it. If the mirror moves away, the image in it moves away at twice its speed, so if you move toward the receding mirror at the same

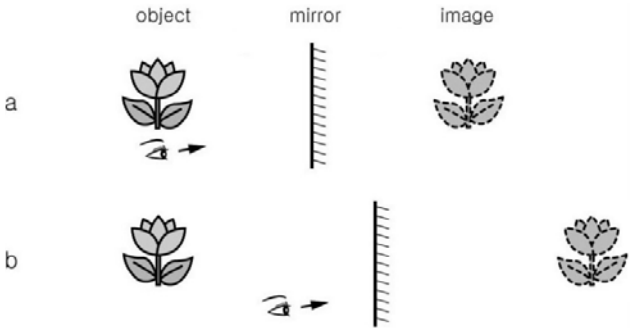

Figure 1. Viewer-object distance changes, but viewer-image distance is the same.

twofold speed, you can keep the image in constant perspective (Figure 1). This amounts to «moving around with the image», and shows that mirror images do not carry ESI. What they normally do carry is egocentric spatial information of the *image itself*.

Why then do mirror images feel so informative spatially? I offer two main reasons. One is that in many cases of mirror viewing, you see a reflection of yourself, and if the mirror image includes your image, naturally it carries ESI. The other is that in most cases of mirror viewing, you know the egocentric position and angle of the device, the mirror. Such knowledge is what enables you to know the location of the objects seen in it. In the case of photographs, we do not usually have that knowledge (about the egocentric position and angle of the camera), but if we do have it, photographs can be as spatially informative as the usual mirror image.

One may object that we do not just *know* the egocentric position of the mirror when we do know it, but we *perceive* it, whereas we do not normally perceive the egocentric past position of the camera. But that is not true. We do not usually perceive the mirror surface which is what matters. When we see a mirror, what we normally see is its frame and the *image* in it. It is true that we can usually guess the surface correctly from the position of the frame, but the frame can be blocked from view, or the mirror can be so big that we may not see its boundary. And anyway the frame does not determine the surface, for not all mirror surfaces are plane. Thus, to know an object's location via a mirror, normally we need to have prior *knowledge* about its surface. And if knowledge about the device is admitted, as I said, photographs too can be spatially informative. There is no essential difference between mirror images and photographs (and holograms too) in terms of ESI. From the next section I will put forth my own theory of transparency. And in the course of that, I will also present more arguments against the importance of ESI.

3. REAL-TIME SEEING

Everyday perception is real-time: it is sufficiently immediate. Ordinary mirror images are certainly real-time, and this can explain the case of Walton's flower: that we can say we see the flower despite our ignorance of its location can be explained by the fact that we see it in real time. The real-time contact condition can explain other phenomena. Most people, I believe, have a stronger sense of seeing in a surveillance monitor than in a movie. This is not only because the former may call for our immediate action. Even if the image is banal, like the interior of your house with just the sunlight in, it will elicit a sense of seeing if it is real-time.

Although it takes time for light to travel and for our neurons to process the visual information, the delay is usually very short. In almost all everyday situations our visual perception is sufficiently immediate[4]. The problem, of course, is star watching. As Walton writes, «We also find ourselves speaking of observing through a telescope the explosion of a star which occurred millions of years ago» ([1984]: 252). I argue, however, that star watching is not a clear counterexample. Think of stars' «twinkling». It is a common expression, but the stars themselves do not twinkle, do not turn on and off regularly, so it must be a perspec-

[4] Nigel Warburton (1988) also emphasized the importance of real-time («virtually simultaneous» for him) perception, but crucially he did not think that star seeing is real-time.

tival description. That is, it means that they *look* like twinkling from here, and it is real-time in that it refers to a current event: the twinkling happens now. And if a star twinkles now, it implies that the «star» is there now, regardless of its objective existence. Thus insofar as we take stars as sometimes-twinkling things, watching them is immediate. «Bright star» is another case in point. It means that the star is bright now, not that it was bright millennia ago. We see a *current* state of the «star». Yet another example is stars' direction. In a legitimate sense, Polaris is in the north now. When we call it the North Star, we do not mean that the star was in the north millennia ago. It *is* in the north. It gives us light from that direction *now*. What matters is the light itself, not some information it carries about the distant object. This is not to say that we experience the light just as sensations. The light's properties are perceived as belonging to a distant point (or distant disk of light)[5]. This is star, at least in its traditional sense.

For most of human history, we have been confined to our planet both physically and mentally, so it is natural that many of our descriptions of stellar objects and events are perspectival. American Heritage Dictionary lists two definitions of «star» that are relevant here. One is more scientific and mentions things like «mass of gas» and «energy generated by nuclear reactions». The other is traditional and includes many viewpoint-relative terms: «Any of the celestial bodies visible at night from Earth as relatively stationary, usually twinkling points of light». It appears that in this second definition, star watching is real-time.

In the first definition, of course, star watching is *not* real-time. But it is not obvious that the possibility of statements like «I see a star as it was millennia ago» poses a problem. For this to be a counterexample to my theory, first it must be a case of normal, genuine seeing, and, second, it must involve delayed seeing. But I think its apparent normality is due to the use of the *second*, traditional meaning of star, and if so, there should be no delayed perception, despite the clause «as it was». Just try adding «twinkling» such as «I see a twinkling star as it was millennia ago». We see that «as it was» clause plays no real role and there is no actual delayed perception. Even if we drop «twinkling», it is not certain what properties, exactly, of the star of the past we see belatedly. Its brightness? But what we normally see in the night sky is the *apparent*, not *absolute*, brightness of the stars, and there is no delayed seeing of apparent brightness. I am not arguing that star watching never involves delay. My point is that there is an ambiguity which we had better avoid if we are to find out whether delayed-but-normal seeing exists. So now I present an example that does not involve stars.

Ordinary mirror viewing is real-time, but what if the mirror gets farther and farther away? Suppose I place a mirror at a 12-light-hour distance from here, and train a gigantic telescope on it so that I can see a reflection of the earth via the combination. I will then be able to see this place that I am at *as it was yesterday*, and it is obvious that if I see, say, my son playing there/here, I do not see him in real time. Is this telescope-mirror transparent? Do I see, genuinely, my son playing yesterday? I do not deny an intuition that says yes, but there is a legitimate sense in which this is not a normal case of seeing. Normally, if I say to people, «I see my son», they will take it for granted that I see his *current* state. And if I say, instead, «I see, literally, my son of yesterday», they probably will not understand what I mean. This suggests that normally, seeing implies real-time seeing.

It can be objected that the above telescope-mirror case is subnormal only because we have no actual experience of it, and that if such devices become ubiquitous, statements like «I see my son of yesterday» will have no problem. That may be true, but there is still a sense in which only real-time seeing is normal. Suppose that all my visual perceptions are delayed by a significant amount of time, say, one hour. Then I will be practically blind. In a sense, even in normal situations I am

[5] What I have said of distant stars applies to our sun, too. Take sunset, for example. We would say «The sun is setting now», even if the star itself had died suddenly 5 minutes ago for some reason.

always blind to something. For example, if I am looking forward, I am blind to my back. But we do not normally say I am blind for this reason. If I see *nothing* in real time, however, it seems appropriate to say that I am blind. Even if the delay is only one minute, if I am hit by a car, I will have a visual experience of the approaching car probably only after I fell down to the ground. Here I would be reluctant to say that I genuinely see the car. It would be like an afterimage. As another example, compare seeing an apple, in the telescope-mirror, sitting on your table yesterday (but eaten by now) and seeing an apple on your table now (which you can eat if you want). In a legitimate sense, you really see an apple only in the latter case.

I do not deny that delayed seeing, too, is a kind of seeing. But an important part of the transparency theory has been that the road from no delay to delay is a «slippery slope» (so delay cannot be a reason to deny the status of genuine seeing) as Walton (1984) and Jonathan Friday (1996) have argued, and I think I have showed in this section that delay does matter, that there is a significant break stopping the slide.

In passing I would like to think about the importance of real-time seeing in comparison with that of ESI. What is the importance of real-time seeing? Why would it be necessary for normal seeing? The obvious answer is because it helps our survival. If so, we may compare its importance with that of ESI which is also helpful for survival. For this, I compare two cases. In the first, ESI is present but it is not real-time. In the second, ESI is not present but it is real-time. Suppose that in the telescope-mirror, my son is (was) at a location that is one meter to the left of my present location. Does this egocentric information have any significance? Hardly. Of what use is it to know his past location *with respect to my present location*? Knowing his absolute location at the time can be surely useful, but knowing his *egocentric* location is useless when he is not there now. Now compare this with the surveillance monitor where perception is real-time but the spatial information is absent. It is evident that real-time seeing is much more important for the purpose of survival. If I see an object in the monitor, even if I do not know where it is relative to me, there can still be time enough to find out. If real-time contact is lost, however, nothing can return it to us[6].

4. EMPTY IMAGE

I claim that genuine seeing via image is possible only when the image is empty. «Empty image» is to convey the intuitive sense that *nothing is really there* at the image location. The canonical example is a virtual image, such as a mirror image. We clearly have the sense that nothing is really there, and hence it is an empty image. But imagine a magic mirror, which we can actually step into and touch the «image» in it. This is not empty: Something is really there. This magic mirror is so different from real mirrors, however, that we had better drop the name «mirror» and devise a simpler example. The example and argument that I am going to make here is actually almost identical with that of Gaut ([2010]: 89), although his main thesis is different from mine[7]. Imagine a room which is divided into two equal sections, left and right, by a line on the ground. On the left, there is

[6] Some comments are in order regarding Currie's view that knowing egocentric temporal location of objects is necessary for seeing. The view has some affinity with mine, but is different in that the perceiver's knowledge is not a requirement for me. Real-time contact is an objective relation between the perceiver and the event perceived. Besides, it is a matter of the relation itself, not an information about it. Another difference is that Currie does not explicitly require the temporal relation be immediate. Even if there is a significant gap between the perception and the perceived event, it seems OK for him if the perceiver recognizes the gap somehow.

[7] His main thesis there is that «we see an object only if rays of light pass uninterruptedly from it to our eyes» (Gaut [2010]: 91). But I do not agree with that. He thinks Currie's ([1995]: 60) transducer is not a counterexample, but it is. The screen is *not* like a television screen. Perceptually it is like a window, as Currie says. As I understand it, it receives a ray at one side and temporarily changes it to a different form of energy and then emits another ray from the other side as if it were continuation of the former.

a real person, and on the right, a wax figure which looks and moves exactly like the person on the left. Visually the two are indistinguishable, except that they are mirror-symmetrical. In fact, if you stand on one side, you could mistake the figure on the other side as a mirror image. Another important fact of this room is that the duplication in appearance and movement is produced through a mind-independent mechanism, in real time. Now, would we say that the figure on the right is transparent? Suppose the person on the left is your son. Do you literally see your son via the wax figure? I believe, with Gaut, that the intuitive answer is no. The image should be empty so that you can see right through it, so to speak. Hence my claim that only empty images are transparent.

I do not deny that there is a sense in which you see your son via the wax figure. After all, it provides reliable visual information about him in real time, just like a mirror image, and Walton might say that this is a case of literal seeing without problem. However, it is also undeniable that there is a strong intuition *against* taking it as literal seeing. My aim in this paper is not to deny that there can be a consistent concept of «see» that encompasses all normal photographs and also the wax figure, but to find if there is a narrower one. If there is such a narrower one, we may take it as the literal, or more literal, concept of «see». Our common, I believe, intuition regarding the wax figure indicates that there *is* such a concept. Another indicator is, of course, the intuition regarding real-time seeing I discussed in the last section. I will try to explain what connects these intuitions soon, but before that I need to elaborate further on my concept of empty image.

I have considered a virtual image which is empty, and a physical figure which is not. But there is a middle ground. It is (optical) real images, examples of which are some holographic images and the images in some microscopes. There is an intuition that these real images are less transparent than virtual ones, and indeed Edwin Martin (1986) once suggested that only virtual images may be transparent. Walton (1986) responded to this that real images, too, can be transparent because microscopes are transparent. But the real images in microscopes are three-dimensional and can look like virtual ones. If you realize that they are in fact real, you might think again. Suppose you enter a «mirror» – something you thought was an ordinary mirror at first – and get as close as you want to the images, although they are still immaterial. Then you would certainly think again (about their transparency). Real images definitely feel less transparent, and I think this can be explained by our sense that something *is* there. They are called «real», after all. On the other hand, however, they feel *more* transparent than the wax figure, and this can be explained by the sense that nothing is really there because we pass right through them. So I am undecided on whether real images are empty or not, and I leave the question at that, for this will not matter in my discussion of photographic images. My bottom line is that if an image is virtual it is empty, and if it is composed of matter it is not.

There is a little ambiguity in my use of «image». I said that for transparency the *image* should be empty. But is the wax figure an image? I think it can be called one. We may call it a «physically real image», but even if this is not appropriate it should not matter. If it is not an image, I can just change my requirement from «image be empty» to «representation (or some other suitable term) be an empty image».

I have proposed two conditions for (transparent) seeing, real-time contact and empty image, but at first glance they seem unrelated to each other. Why the two? It would be good if we had a common theme connecting them. I think we can find it in the fact that we are embodied creatures *living* in the world. Take the necessity of real-time seeing. It is obvious that real-time perception is important because we have bodily needs, like getting food and avoiding injury. Even if we were disembodied (so that we cannot physically interact with the world, though we can see and visually move around), we may suppose that we would still have *intellectual* needs and have to acquire new knowledge regularly to survive as minds. Even so, real-time seeing is not necessary. We can learn

new things from the past world, and besides, even if there is a delay of one year in our perception, each day will be a new day for us.

It also appears that the distinction between empty and non-empty images makes sense only because we are embodied. If an image is empty, we are not able to interact with it. In fact, I take it that an image is empty if and only if there is no potential for interaction. In other words, the sense that *nothing is really there* at the image location is equivalent to the implicit belief that *there is no potential for interaction* with the image. In fact this can explain the ambiguous status of optical real image. The ambiguity can be ascribed to the uncertainty of whether we can say we interact with it. In one sense we do not interact, for we pass right through it, but in another sense we do, for when we pass we probably disrupt the image.

When we do not interact even with «solid» objects due to disembodiment, then, the distinction between empty and non-empty images will have no significance. The entire world will be empty, in effect. My proposed concept of literal seeing is based on our embodied existence, and it entails that if we were disembodied we would see nothing literally (except in the sense in which we can say we literally see mirror images). Carroll has hinted at the same idea. He wrote, «But if we call what we see on the silver screen a "view", then it is a disembodied view» ([1996]: 60). Although we are not literally disembodied, given that the seeing is not real-time, it is not really different from disembodied seeing.

5. SOME PICTURES ARE EMPTY IMAGES

Photographs are physically real and so it may seem they are not empty images. But there is a legitimate sense in which the image we actually see is an empty one. It is customary to say that pictures are 2D images, but in another sense they are 3D, and for this latter I will use the notation **image**. This **image** is empty. What is important in this notion is our *perception of depth* in pictures. See Figure 2. We can see it as a 2D array of figures, but we can also see it as a 3D scene with, say, two near and two far balls of the same size. What is crucial here is that in the latter case we perceive, feel, depth. We do not just know that the picture can be interpreted as a scene in depth. A pattern-recognizing computer can do that, too. We experience depth, not just understand that something is in front of another. I believe that this distinction has not been made clearly enough in the previous literature on picture perception.

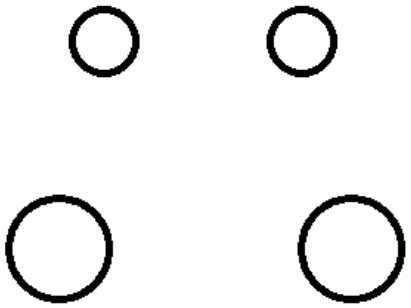

Figure 2.

For instance, Richard Wollheim spoke of our discerning «something standing out in front of […] something else» in pictures (Wollheim [1998]: 46), but Malcolm Budd says, «It is unclear what Wollheim supposes this visual awareness of depth to be» (2008: 203). In fact, if Wollheim had meant what I am referring to here, the feeling of depth, he would not have argued that we see a picture's surface and its subject matter at the same time—his twofoldness thesis. As Figure 2 shows, I believe, we cannot feel depth and see surface at the same time. The perceived depth is a sort of illusion, and having it seems to preclude seeing the surface. For us to see both at the same time, the picture elements would have to divide roles between them. That is, it would have to be that some of the line segments cause us to have the illusion, while the other segments let us see the surface. But I strongly doubt that depth illusion works that way. At any rate this paper requires only that *sometimes* we can concentrate on the depicted scene, which I think is not a strong claim. Think of an audience engrossed in a movie at a theater. I doubt anyone among them is aware of the screen surface.

(Nanay has defended twofoldness thesis by supposing that we see, but not attend to, the surface when we see pictures. He says that «while we do simultaneously see both the surface and the depicted scene, we do not simultaneously attend to both» ([2018]: 168). For the purpose of this paper, it is fine if we can solely attend to the scene. But his defense of twofoldness feels strained to me. While unconscious perception is certainly possible, do we have to suppose this difficult-to-confirm process in *every* picture perception? Here I will briefly propose a different theory. It is possible that when we see pictures our attention switches between the surface and the content so effortlessly, at least in most cases, that the two experiences feel simultaneous. Moreover, we may hold in memory what we perceived split-second before (the surface, say) and compare it with what we currently perceive (the content, say), and this can contribute to our sense of simultaneous awareness. This hypothesis reconciles Wollheim with E.H. Gombrich. It admits the impossibility of simultaneous perception of surface and content, but at the same time explains, I believe, what Wollheim wanted to explain with his twofoldness claim, namely, seeing-in and aesthetic appreciation of pictures. Twofoldness does not seem to require literal simultaneity of perception).

I return to the problem of depth perception. Although everyone will agree that the depicted scenes in pictures are 3D, in a way their three-dimensionality is not very obvious. As Budd says, «[The picture-surface] does not seem to the viewer that he is seeing a three-dimensional state of affairs». Thus his proposition is that «the only relevant sense in which a picture, seen as a depiction of its subject, can look like its subject is with respect to the two-dimensional aspect of the subject's visual appearance» ([2008]: 219). But what I am trying to convey is that we do see a three-dimensional state of affairs in a picture, and that is what I call **image**.

One might think that Figure 2 is a special case in which perspective – apparently converging parallel lines – produces a sense of depth. But there are many depth cues, and even in a very simple line drawing of a face, say, we perceive the nose as protruding relative to the face surface. Computers can recognize the spatial relation, but they would not have the phenomenology of perceiving a protrusion. Actually, if we did not have the depth «illusion», we would not see it as a picture of a (normal) face. Depth illusion *is* picture perception, in the sense that it is the result of the brain's construction of a 3D scene out of the 2D array of marks. We need three axes to represent 3D space, and the left-right (x) and up-down (y) axes are easily represented, but how do we represent the to-and-fro (z) axis? The «illusion» should be understood as our visual system's way of representing the z axis. Without it, there is no scene perception. The very fact that a picture is intelligible as a depiction of 3D scene implies that we already have depth illusion in a particular way.

A corollary is that there is no such thing as a 2D image, at least in one sense. There are only *2D arrays of marks* in which people see **images**. One may say that images in pictures are 2D in that they show, to use John Hyman's (2006) term, «occlusion shapes» of objects. But suppose there is a picture of two balls, one partly blocked by the other. Most people will see it as I have just described, as two balls, but it is not impossible to see it as a single strange object. The point is that *object individuation* or *figure/ground distinction* should come first before we can say anything about shapes, 2D or not, and this individuation involves depth perception in most cases. In order for us to see the ellipse on a surface as representing the occlusion shape of a table, for example, we need first to see the ellipse as an individual figure against some background, and this normally involves depth perception. It is not that the concept of occlusion shape is useless. It is that calling it a 2D image, as if it does not depend on 3D perception, is misleading[8]. Less misleadingly, it is *an aspect of a perceived 3D object*, of an **image**. In the strict absence of depth perception, pictures of 3D

[8] Anthony Derksen (2004) made a related argument against Hyman, but my approach is at a more basic level.

scene are just meaningless arrays of color patches, at least in most cases.

I do not argue that all pictures require depth perception. As Hyman ([2006]: 134-136) shows, some pictures do not seem to depict depth. What I do say is that in the majority of pictures what we see in them are (3D) **images**[9]**,** and also that they are empty. In this respect, picture perception is no different from seeing a mirror image. In both cases we see an **image** of the subject, although in the mirror case I do not need to use the notation **image**, for its three-dimensionality is unambiguous. Both **images** are also empty: nothing is really there at the **image** location. The mirror image is virtual, of course, but the **image** in a picture is virtual, too, in the sense that the light does not come from its apparent location[10].

Their difference, of course, is that in pictures we also see their two-dimensionality, or «planarity». This planarity is difficult to ignore[11]. It is always ready to be perceived, ready to remind us that we are looking at something extraneous to the representation (the planarity does not represent anything). Still, we can evade it in various ways. We may try concentrating on the **image**, or if that is difficult we may see the photograph through a hole with one eye. Or we may enlarge the picture greatly and see it from a great distance. In this way the planarity can be imperceptible (accommodation and binocular depth cues have little effect if the scene is distant). In all these cases, the fact remains that we see a photograph. Thus, given that photographic images can be produced in real time, we have the conclusion of this paper: some photographic images are transparent[12].

Before closing I want to mention an interesting kind of empty image that is based on photography. That is stereoscopic 3D images. Here we have a very strong 3D illusion and barely notice the planarity, so I am inclined to say that these are transparent (if they are real-time). The problem is that it seems uncertain whether we can say we see *photographic images*. If we can say we do, we have got additional examples of transparent photographic images.

REFERENCES

Bazin, A., 1971: *What is Cinema?*, vol. 1, transl. by H. Gray, University of California Press, Berkeley.

Budd, M., 2008: *Aesthetic Essays*, Oxford University Press, Cambridge (Ma).

Carroll, N., 1996: *Theorizing the Moving Image*, Cambridge University Press, Cambridge.

Cohen, J., Meskin, A., 2004: *On the Epistemic Value of Photographs*, "Journal of Aesthetics and Art Criticism" 62, pp. 197-210.

Currie, G., 1995: *Image and Mind: Film, Philosophy, and Cognitive Science*, Cambridge University Press, Cambridge.

Derksen, A., 2004: *Occlusion Shapes and Sizes in a Theory of Depiction*, "British Journal of Aesthetics" 44, pp. 319-341.

Friday, J., 1996: *Transparency and the Photographic Image*, "British Journal of Aesthetics" 36, pp. 30-42.

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

[9] It seems to me that authors are at a loss when they want to denote the thing the viewer sees in a picture. It cannot be the subject itself, but it cannot be an «image» either, for the word is already taken up to refer to the picture itself (we cannot say, «I see an image in an image»). Take the expression «depicted scene», for example. Readers usually understand what it refers to in the context, but the expression itself could easily mean the subject, say, some historical event.

[10] Virtual image is «made by rays that do not actually come from where the image seems to be» (*Encyclopædia Britannica*).

[11] I avoid using the word «surface» which invokes things like texture and surface reflection. Even without those, we can perceive the planarity which is a more fundamental feature.

[12] Real-time contact and empty image may not be sufficient for transparency. Mind-independency (MI) of the photographic process is probably also necessary, but I took it as given. One may doubt MI for ordinary photographs, but I am concerned with real-time photographic images, about whose MI there will be little controversy.

Gombrich, E.H., 1969: *Art and Illusion*, 2nd ed., Princeton University Press, Princeton.
Hyman, J., 2006: *The Objective Eye*, University of Chicago Press, Chicago.
Martin, E., 1986: *On Seeing Walton's Great-Grandfather*, "Critical Inquiry" 12, pp. 796-800.
Moon, W.L., 2018: *Documentary and its realism*, "Studies in Documentary Film" 11, pp. 43-55.
Nanay, B., 2010: *Transparency and Sensorimotor Contingencies: Do We See Through Photographs?*, "Pacific Philosophical Quarterly" 91, pp. 463-480.
Nanay, B., 2018: *Threefoldness*, "Philosophical Studies" 175, pp. 163-182.
Walden, S., 2016: *Transparency and Two-Factor Photographic Appreciation*, "British Journal of Aesthetics" 56, pp. 33-51.
Walton, K., 1984: *Transparent Pictures: On the Nature of Photographic Realism*, "Critical Inquiry" 11, pp. 246-277.
Walton, K., 1986: *Looking Again through Photographs: A response to Edwin Martin*, "Critical Inquiry" 12, pp. 801-808.
Walton, K., 1997: *On Pictures and Photographs: Objections Answered*, in Allen, R., Smith, M. (eds.), *Film Theory and Philosophy*, Clarendon Press, Cambridge, pp. 60-75.
Walton, K., 2008: *Postscripts to* «Transparent Pictures»*: Clarifications and To Do's*, in *Marvelous Images*, Oxford University Press, Cambridge, pp. 110-116.
Warburton, N., 1988: *Seeing Through «Seeing Through Photographs»*, "Ratio" 1 (1), pp. 64-74.
Wollheim, R., 1987: *Painting as an Art*, Thames and Hudson, London.
Yetter-Chappell, H., 2017: *Seeing through eyes, mirrors, shadows and pictures*, "Philosophical Studies" 175 (8) (https://doi.org/10.1007/s11098-017-0948-8).