

# A Semiology of Video [1]

Takahiko iimura  
Tokyo Institute of Polytechnics  
4-50-4 Yamato-cho, Nakano-ku, Tokyo, 165-0034, Japan  
+813-3310-3281  
[iimura@gol.com](mailto:iimura@gol.com)

I made a video trilogy during 1975-76: *Camera, Monitor, Frame* (1976); *Observer / Observed* (1975); and *Observer / Observed / Observer* (1976). The aim of this project was to create a Semiology of Video as a video work rather than a written text; it was completed after my series *Self Identity* (1972-74). The relationship between language and image, which was initially explored using "I" as the subject in *Self Identity*, is developed according to the logical structure of a video system dealing with the grammar of language at the same time.

I have presented these works along with an accompanying lecture at a number of universities and museums in the United States and Europe. Now I have a strong urge to put the logic of these video works into words. Since they were originally produced in English, I feel it is necessary to write in Japanese, recognizing differences between English and Japanese for both a Japanese audience and myself. Although my format is that of notes on a particular work, I hope its broader theoretical implication will be equally valid.

This essay primarily focuses upon "This is a camera 1" and "This is a camera 2," two segments of *Camera, Monitor, Frame*. [2] Several important differences between a Semiology of Video and a Semiology of Film are already apparent in these works. The reference point of "This is a camera 1" is *Kino Eye* by Dziga Vertov; for "This is a camera 2" it is the propositions of Sergei Eisenstein's "word" theory along with Christian Metz's "sentence" theory. In the latter video, I have proposed a third alternative by comparing the structure of the Japanese sentence to my video work. This proposition may be hypothetical, but it is also a Semiology of Video analyzing the structure of a Japanese sentence. It differs from Eisenstein's montage theory which he also applied to the analysis of the Japanese ideogram (*kanji*). When I have presented this alternate proposition in the West, it has generated considerable discussion.

*Observer / Observed* was made before the other two works in the trilogy; *Camera, Monitor, Frame* followed; and *Observer / Observed / Observer* completed it.  
*Camera, Monitor, Frame* nonetheless comes first thematically and

First published at COSIGN-2002,  
02 – 04 September 2002, University of Augsburg,  
Lehrstuhl für Multimedia-Konzepte und Anwendungen,  
Germany

served the role of introducing the other two works. Its title correctly implies that it is concerned with the basic objects of a video system. It is not intended to focus upon the mechanical or technical aspects of the camera or the monitor, but rather their role in the video system and their definition in terms of language. Therefore the limits of language are also examined relative to the image. As the intertitles "This is a camera" and "This is a monitor" indicate, this work deals with the logic of language and image in the definition of the objects.

Based upon an understanding of these objects, the *Observer...* works can then be located. The viewpoint has been shifted from that of the individual object to the structural problem of the video system in the interrelationship between the observer and the observed. It thus presents a structural viewpoint rather than dealing with the definition of an object. This may be demonstrated by the titles which depict relationships concerning both the observer and the one who is seen through the usage of the diagonal line (" / ").

I have utilized a linguistic approach throughout the trilogy. Compared to a series of images in which a picture takes on the character of a sign, a language possesses a much more logical structure. This is not to subordinate a language to a picture or vice versa, but instead to distinguish between two signs which have different characteristics. In actuality this is also explained by language, but I am seeking an approach which refers to both language and picture simultaneously. The way to do this is to articulate the picture as well as the language while regarding the picture syntactically like a language. One should not examine an individual picture by itself, but "read" the syntax and the morphology of the image. For language, I am likewise concerned with the sentence, not the individual word. (In the video there are word phrases which are not sentences in a grammatical sense; these are instances where the subject or the predicate, which constitute the sentence, is either hidden or omitted.)

The main themes of the video trilogy can be expressed in grammatical terms:

- (1) A simple sentence that uses demonstrative and personal pronouns as its subject and an auxiliary verb as its predicate (*Camera, Monitor, Frame*)
- (2) A complex sentence with relative pronouns (*Camera, Monitor, Frame*)
- (3) The indefinite article, the definite article, and the infinitive (*Camera, Monitor, Frame*)
- (4) The active voice and the passive voice of the verb (*Observer / Observed*)
- (5) The positive and the negative of the gerund (*Observer / Observed*)
- (6) The active and the passive voices of a simple sentence that uses a demonstrative pronoun as the subject and a verb as the

predicate. Also, four combinations of the positive and negative forms (*Observer / Observed*)

(7) The complex sentence with a relative pronoun combined with a sentence which uses a personal pronoun as the subject and a verb as the predicate (*Observer / Observed / Observer*)

(8) A complex sentence with two relative pronouns and the active and the passive voices of the verb (*Observer / Observed / Observer*).

I will be discussing only the first two themes here. [2]

The above themes are all parts of a grammar. Another interesting approach would be to show grammatical differences within the same image. In other words, there are plural signifieds of the picture to the signifier in the same sentence.

As far as I know, there is no paper or work in film or video which attempts to analyze the relationship between language and picture in terms of individual cases of grammar. There are many theoretical works, including books by Christian Metz, which posit semiologies for film; video is still a young medium and has not yet received such analysis. I have read some of these theoretical treatises. Although they have enhanced my understanding of the topic, my video productions have taken quite a different approach. Metz, for example, analyzes the narrative of dramatic films. I am not concerned with drama. Instead, I have attempted to develop a semiological text within the framework of my video work, one which differs radically from the discussion of the sign in words. This is another approach to semiology, one which serves as an experiment with video as well as with semiology.

While Metz's approach to semiology is concerned with (dramatic) film, I deal with video, which has certain elements in common with film, yet has its own unique system. I am particularly interested in the structure of video functioning as a system. My work should not be analyzed simply for the pictures which appear on the screen and their accompanying sound. Instead it should be considered within the context of the image being manipulated through an entire system. In this way the structure of video as a closed circuit can then be comprehended.

## Camera, Monitor, Frame (1976-1998)

### This is a camera 1

The main theme of this work is the dialectic between the visual and the language. There are two propositions. One is "This is a camera"; the other, "I am Takahiko Imura." The former is a definition for the object; the latter defines the subject. The theme of this work is to synthesize and combine two propositions into the dialectics of the visual and the language. In the first shot the sentence "This is a camera" is spoken to Camera 1 (No.1A; the shot references are all to Program No. of the chart). The spoken text is accomplished by a voiceover (unsynchronized sound). The next shot features the face of a person saying "I am Takahiko Imura" (No.2C); it has a synchronized voice with lip movement. In the first shot, the speaker does not appear in the picture, so it is not known who spoke the words. It is technically possible that the second shot also uses voiceover; however, since it appears to be synchronized, an audience would generally assume that the speaker is identical to the person in the picture. In terms of identifying the voice, the only difference between the two shots is the existence of lips whose movement corresponds with the voice. Clearly there is a

discrepancy between the picture and the voice. This phenomenon is often ignored in the narration of documentary films, but the identity of image and sound play an important role in my video work.

Let's return to our analysis. After presenting two propositions, Camera 1 is again accompanied by the voice which was heard in the first shot, again stating "This is a camera" (No.3A). Then the camera pans to white wall (no object), and the voice over states "This is not a camera" (No.3B). This is the opposite of the first proposition. A white wall which shown no object (no picture) may be defined in many ways. However, when presented in conjunction with the first proposition, its definition must negate the original. The visual relationship between the two is shown by the panning process. Panning, remaining within the same shot, retains the continuity of the same space / time; cutting disconnects the two. The shot itself moves from the presence of a camera to its absence. The text "This is not a camera" must therefore negate the original proposition. Normally the pronoun "this" is supplemented by the presence of an object, but there is no object in an empty space. One could also say "This is not a camera" when a face or another object appears; however, in an empty space the demonstrative pronoun loses the object. Yet no object - no picture is itself an object for "this" since there is no other reference. Thus the pronoun "this" signifies an object as well as no object.

In the next picture Camera 2 is accompanied by the voice once again stating, "This is a camera" (No.4D). Following a pan, the face reappears with a voice now saying, "This is not a camera" (No.4C). With the introduction of Camera 2 the setting becomes clear: two cameras face each other, and the person is seated next to Camera 2. The two cameras are panning each other. This setting is retained throughout the work; it is used to show the inter-relationship of the back-and-forth movement in the video. Camera 2 is identified by a number which has been attached to it, but is accompanied by the "This is a camera" voice which described Camera 1. The identical sentence has been voiced although it is referring to different cameras because it refers to the generality of the object, a camera; the difference is determined by visually reading the number. Thus the generality of language and the individuality of the object, or the abstraction of language and the concreteness of the object, are shown simply by the numbers "1" and "2" written on the cards. In addition to watching the picture and listening to the voice, the viewer must perform the third function of reading the card. Following the pan, the face with the synchronized voice says "This is not a camera," the same sentence which accompanied the empty space. This time, however, the speaker's subject is shown. In English, unlike Japanese, "this" can be used to identify a person as well as a thing. Unlike the instance of the empty space, "this" signifies the speaker; and unlike the "I" of "I am Takahiko Imura," "this" objectifies the self. Thus far the positive statement "This is a camera" has been spoken three times, its negative twice. The picture, however, has differed every time except for shot 1A and 3A. This means that the identical sentence can accompany different pictures. In the case of the negative sentence there are endless verbal possibilities, but two propositions are set up initially to limit the scope of the object. The face saying "This is not a camera" completes the cycle of the first proposition concerning the positive and the negative.

Next the proposition "I am Takahiko Imura" will be similarly interrogated. As in the previous case, the subject uses "this" instead of "I." Accompanying the picture of Camera 1, the voice says "This is not Takahiko Imura" (No.5A). After a cut to the face, the voice says "This is Takahiko Imura" (No.6C; sync sound). Like

the cutting between the first two propositions, this editing visually suggests each independent statement. The statement “This is Takahiko Iimura” combines the two propositions “I am Takahiko Iimura” and “This is a camera.” The demonstrative pronoun which is used for the camera becomes the subject in this sentence with the object being the proper name. “This” is spoken by the person in synchronized voice, and thus becomes an objectified statement.

The two propositions are then repeated, but now a picture of Camera 2 (No.7D) is followed by a pan to the face (No.7C). Unlike the earlier cutting, both propositions are located in the same space / time.

In the final shot (No.8C+D) the face appears in the left half of the screen with Camera 1 occupying the right part. Here the voice says (sync sound) “I am a camera.” This statement is the logical conclusion from the propositions developed up to this point. Both the person and the camera are in the same picture, and linguistically the subject “I” identifies with the object of the camera. If “I am a camera” were spoken as an independent

- (1) the establishment of the proposition (No.1,2);
- (2) the positive and the negative of the proposition and the exchange of the pronoun (No.3, 4, 5, 6);
- (3) the representation of the proposition (No.7);
- (4) the conclusion (No.8).

The two propositions could have been expressed in eighth different ways, including the positive and the negative and the exchange of the pronoun, but not all were used.

- (1) This is a camera
- (2) This is not a camera
- (3) I am Takahiko Iimura
- (4) I am not Takahiko Iimura
- (5) This is Takahiko Iimura
- (6) This is not Takahiko Iimura
- (7) I am a camera
- (8) I am not a camera

Expressions 1, 2, 3, 5, 6 and 7 occurred in “This Is a camera 1”; 4 and 8 did not.

**Camera, Monitor, Frame**

This is a camera 1					b/w, sound, 2'00"	
picture plan						
	A	B	C	D		
picture						
description	camera 1	wall (white)	face	camera 2		
cable connections	C.2-VTR 2	C.2-VTR 2	C.1-VTR 1	C.1-VTR 1		

program						title 5" total 2'00"	
No	picture	description	voice		min.sec		
1	A	camera 1	This is a camera	a.d.	10		
2	C	face	I am Takahiko Iimura	sync	20		
3	A	camera 1 (pan to B)	This is a camera	a.d.	30		
	B	wall (white)	This is not a camera	a.d.		45	
4	D	camera 2 (pan to C)	This is a camera	a.d.	50		
	C	face	This is not a camera	sync		1.05	
5	A	camera 1	This is not Takahiko Iimura	a.d.	1.15		
6	C	face	This is Takahiko Iimura	sync	1.25		
7	D	camera 2 (pan to C)	This is a camera	a.d.	1.33		
	C	face	I am Takahiko Iimura	a.d.		1.45	
8	C+D	camera 1&face	I am a camera	sync	1.55		

← left pan      sync =synchronized      a.d. =audio dubbed

**Chart 1. This is a camera 1**

sentence, it would be a metaphor. Beyond this, moreover, the person who is looking through the camera is speaking, so the person and the camera (or, more accurately, the eye of the person has been replaced by the lens of the camera) have been combined and can be identified together.

This work consists of a set of two pictures of cameras along with the face of a person plus a final shot in which the person and a camera appear together. Altogether, there are eleven pictures, including three panning, in eight actual shots.

On the other hand, the language of the work consists of four stages:

The conclusion “I am a camera” is close to what Dziga Vertov suggested in his *kino-eye* writings [3] and his film *The Man With a Movie Camera* (1929). While Vertov’s essays are rather metaphorical, my conclusion is drawn from the dialectic between the two propositions. In *The Man With a Movie Camera* the cameraman often appears in the frame with the camera. There are also several double exposure shots featuring lens and the eye. On the metaphorical level, “This is a camera 1” owes much to him, but it follows the logical process of two propositions in the relation between language and image. In other words, my works attempt to achieve a dialectic between picture and sound recording of both synchronized and asynchronized voices against the picture. This was not possible in Dziga Vertov’s silent pictures.

**This is a camera 2**

This work is similar to “This is a camera 1” as two cameras are facing each other, but there is no person. The cameras pan each other, and the picture consists only of Camera1, Camera 2, and empty space. There is no synchronized sound, only voiceover.

The narrative of “This is a camera 2” begins with the sentence “This is a camera which shoots this,” The second “this” then becomes the subject of the next sentence, “is a camera which shoots this,” a repetition of the original sentence. It repeats itself endlessly like a loop. Such repetition is possible because “this” in English has the same form in both the subjective and objective cases.

In Japanese one says “Kore wa kore o satsueisuru kamera dearu kore wa kore o satsueisuru kamera dearu” (which translates very roughly as “This this shoot camera is this this shoot camera is”). Since the Japanese language does not have the cases of pronoun or relative pronoun, this has been translated into Japanese using the particles (words which define the role of a noun in a sentence) “wa” and “o” after “kore” (“kore wa” means “this” in the subjective case; “kore o” is “this” used as an object). A Japanese equivalent for the word “this” could not be switched from object to subject, so “kore” would need to be used twice. Another structural difference between Japanese and English is that in Japanese the subordinate clause comes first, followed by the main clause.

The structure of “This is a camera 2” is cyclic. The first shot begins with a picture of Camera 2 accompanied by the voiceover “This is a camera which shoots” (No.1B). There is a pause in the voiceover, and the object “this” is not spoken until the next image appears. The next shot begins with Camera 1; the same sentence is spoken (No.2D), beginning with the “this,” the object of the first sentence, which also serves as the subject of the following sentence. Next is an empty space; the object “this” of the second sentence is spoken (No.3A). Panning from the empty space to Camera 2 is accompanied by the sentence “is a camera which

from the empty space to the picture of the third camera. In fact the picture (No.3B) is defined by the sentence without the subject (in Japanese “satueisuru kamera dearu” or “shooting camera is”). Here, as often occurs in Japanese, a sentence is able to communicate without a subject. I believe this sentence (“is a camera which shoots this”) is readily understandable in English as well as Japanese. Since the sentence provides an adequate definition of the picture, this suggests something interesting concerning the relationship between the image and language.

Before entering into a general discussion, I want to examine this relationship from the viewpoint of the image. Camera 1 is the object of Camera 2. Then Camera 1 becomes the subject and shoots Camera 2 as its object. This is how the two cameras are set up. The object becomes the subject and vice versa through the workings of the word “this.” This is how feedback is structured in video; the cyclic sentence is its verbal equivalent. Now the empty space between the pictures of the two cameras cannot become the subject of a camera (an empty space can not shoot). As a consequence of the logic concerning the picture, therefore, an empty space as the object can not turn into the subject. Applying such logic to verbal concerns, the object “this” applied to the empty space would turn into the subject. But in relation to the picture there follows a discrepancy between the language and the picture. The picture does not necessarily follow the logic of language. Accordingly, it would be correct to regard the sentence spoken simultaneously with the picture (No.3B), which follows the empty space, as an “object sentence,” a sentence without a subject.

If a picture can be equivalent to an “object sentence,” this posits a new theory concerning the relationship between language and the image. Two prominent theories are the montage theory of Sergei Eisenstein and the film semiology of Christian Metz. My concern here is with a picture (shot) equivalent to language. Eisenstein regarded a picture (shot) as a word, and a montage made out of pictures (shots) as a sentence. [4] This comes from his analysis of the Japanese character in which an ideogram consists of different characters (*hieroglyphs*). [5] On the other hand, Metz considers a picture (shot) as a “sentence.” [6] Metz’s idea is based on traditional French film theory in which a shot has more weight than a montage. (The word “picture” should be regarded as “shot” in reference to Eisenstein and Metz.)

What I am suggesting with the “object sentence” is that the subject is not necessarily required for the definition of a picture. Even without the subject, the sentence still provides an equivalent for the picture. This point of view differs from both Eisenstein’s theory concerning the “word” and Metz’s “sentence” theory. It is a third theory, in between the other two. In my work I have attempted to fully prove that the “object sentence” provides an alternative. Although this idea was derived from the form of a sentence without the subject which exists in the Japanese language,[7] I have shown that it can be applied to English. This may be a specialized case, yet is it not possible to generalize it? After all, a picture taken as a shot has no subject, but exists like an object with a predicate, an “object sentence.”

(Revised and translated by the author with the assistance of Bill Thompson from his book *Eizo Jikken no Tameni* ((*For Visual Experimentation*)), Seido-sya, Tokyo, Japan, 1987, pp.243-262.)

Since I first wrote the above paper “A Semiology of Video” [8] in

**Camera, Monitor, Frame**

**This is a camera 2** b/w, sound, 1'04"

picture plan				
	A	B	C	D
picture				
description	wall (white)	camera 2	wall (white)	camera 1
cable connections	C.1-VTR 1	C.1-VTR 1	C.2-VTR 2	C.2-VTR 2

program					title 5" total 1'04"	
No	picture	description	voice (dubbed)	min.sec		
1	B		camera 2	This is a camera which shoots	11	
2	D		camera 1	This is a camera which shoots	16	
3	A	wall (white) (pan to B)	This		20	
	B		camera 2	is a camera which shoots		
4	C	wall (white) (pan to D)	This		27	
	D		camera 1	is a camera which shoots		
5	B		camera 2	This is a camera which shoots	37	
6	C	wall (white) (pan to D)	This		40	
	D		camera 1	is a camera which shoots		
7	A	wall (white) (pan to B)	This		47	
8	B		camera 2	is a camera which shoots	53	
9	D		camera 1	This is a camera which shoots	58	

→ right pan

**Chart 2. This is a camera 2**

shoots” (No.3B). Following this are the empty space (No.4C), pan to Camera 1 (No.4D), Camera 2 (No.5B), empty space (No.6C), pan to Camera 1 (No.6D), empty space (No.7A), all with corresponding narration. This concludes the first cycle. Up to this point, there have been 10 pictures, 7 shots and 3 pannings, altogether.

So far there have been three types of connections between the images: from one camera to the other (No.1B - 2D); from the second camera to empty space (No.2D- 3A); and from the empty space back to the first camera (No.3A - 3B). In the first case, the object “this” is spoken (directed) toward the other camera; however, the second object “this” is spoken toward the empty space. The first object “this” then simultaneously turns into the subject of the next sentence, but the second must wait until the third picture of a camera appears through the panning process. When this third image appears, “is a camera which shoots” is spoken starting from the predicate “is,” therefore, corresponding to the second object, one hears “this,” yet it appears to be like a sentence without a subject. This is necessitated by the time interval during the pan

Japanese, 1983, 15 years have passed. During this period video semiology has neither been discussed as much as film semiology nor has it developed distinctively as a study. What is the meaning of this absence of video semiology in comparison to film? I do not think that film semiology in its present state can be considered adequate for video, and I do think that a study of semiology unique to video should be established. I hope "A Semiology of Video" can serve as a basis for further discussion.

This year (1998) I have had an opportunity to remake the videotapes which I consider the foundation works for "A Semiology of Video": *Camera, Monitor, Frame* (1976), *Observer / Observed* (1975), and *Observer / Observed / Observer* (1976).[9] Without altering their basic concept, I remade them more simply than the originals, reducing their length drastically (three tapes altogether 22min. from 59min.) and cutting the redundant repeating parts of the original.

The discussion hereafter is based on this remade version, but the writing previously published (on "This is a camera 1" and "This is a camera 2") remains unchanged, since those pieces were remade without fundamentally altering their concept. For instance, in the case of "This is a camera 1," nothing in the script has changed, only the timing has been shortened from 3minutes 30seconds in the first version to 2minutes 45seconds in the revised one; and "This is a camera 2" has been reduced from 3minutes 30seconds to 2minutes by giving up the repetition.

### This is a monitor 1

On the work "This is a monitor," I recently wrote the following:

After the already mentioned "This is a camera," there comes "This is a monitor." (The work *Camera, Monitor, Frame* includes the five small pieces "This is a camera 1," "This is a camera 2," "This is a monitor 1," "This is a monitor 2," and "To See the Frame"). "This is a monitor 1" shows the various states of the monitor.

First we see a monitor (a product) without any image (No.1A); next, by switching on the monitor, which is connected to the camera, the monitor is made to show reduplicated feedback monitors in the form of a tunnel (No.1B). Then the camera goes off, and on the monitor appears no image but the scan lines (No.2C). Throughout these images, a voice utters only "This is a monitor," repeated over and over. For this (and every other) piece there is a chart of picture plan and program, the former provides picture, description and cable connections and the latter, picture, description, voice, and minutes and seconds. What essentially constitutes this piece is the state of the image (and its absence) on the monitor. The first monitor, as an object, shows clearly the existence of another monitor within the monitor; but after we see the feedback monitor, the voice saying "This is a monitor" indicates that the actual monitor which the audience is watching is also a part of the work. In this case the actual monitor is not only to present a work on screen, but also becomes a work itself.

Therefore this piece could not be achieved as a work in the form of a book, which is a printed medium.[10]

The above quotation is long but I think it fully explains the work. So called "reality" and "fiction" for a monitor are related, in that one can come and go between them according to the image itself as it is shown on that same monitor. A hand which invades the space for switching on the monitor act as a catalyst.

### Camera, Monitor, Frame

This is a monitor 1 b/w, sound, 1'15"

picture plan			
	A	B	C
picture			
description	monitor off within monitor	monitor in feedback	camera off picture noise
cable connections	C-VTR	C-VTR-M	M-VTR

program				title 5" total 1'15"	
No	picture	description	voice (dubbed)		min.sec
1	A	monitor off within monitor	This is a monitor		10
	B	hand in frame switches on monitor monitor in feedback	This is a monitor		20
2	C	camera off picture noise	This is a monitor		39
3	B	camera on monitor in feedback	This is a monitor		49
	A	hand in frame switches off monitor	This is a monitor		1.00
4	C	camera off picture noise	This is a monitor		1.10

Chart 3. This is a monitor 1

The sentence, "This is a monitor," is a definition and a statement. In English it has an article "a" which indicates the singular (there is no article in Japanese), but the monitor displays not only multiple states of a monitor (though they are derived from a single one) but also, in the feedback state, one sees numerous monitors in the form of tunnel. These images are plural; however, since they are mirror images of the same monitor, one may nevertheless regard them as a singular one.

### This is a monitor 2

This piece is also trying to formulate a definition of the monitor, not through use of a voice but through reading (silent) letters. Unlike the voice, reading is a self-conscious act of the audience and is a more active perceptual activity than receiving the voice message, which "is heard" automatically. "This is a monitor" as a sentence is self referential as long as it is seen on a monitor, where the act of reading a sentence on a piece of paper pinned on the wall (No.1C) is transferred to looking at the monitor. That one would undoubtedly become aware that the sentence one is reading on paper is in fact on a monitor, is due to the subsequent image of a monitor within the real monitor (No.1D), a feedback image in which multiple monitors are seen in a configuration like a tunnel (No.2B). Here also the word "monitor" comes and goes between reality and image, but since the word itself is also an image, the relation of words and image, which is different from the case of words and voice, becomes even closer. In fact if one considers words on paper as "real," then when those words are transferred onto a monitor within a real monitor, one can not distinguish the "reality" of the words from the image of the words. The "reality"

**Camera, Monitor, Frame**

This is a monitor 2 b/w, silent, 1'17"

picture plan				
	A	B	C	D
picture				
description	title on paper "This is a monitor"	monitor in feedback	title on paper "This is a monitor"	title within monitor
cable connections	C.1-VTR	C.1-VTR 1-M	C.2-VTR 2	C.2-VTR 2 (C.1-M)

program		title 5" total 1'17"	
No	picture	description	min.sec
1	C	title on paper (pan to D)	10
	D	title within monitor	21
2	A	title on paper (pan to B)	27
	B	monitor in feedback	35
3	D	title within monitor (pan to C)	42
	C	title on paper (pan to B)	58
4	B	monitor in feedback	1.12

**Chart 4. This is a monitor 2**

of letters as they are seen pinned on the wall is also an image, as long as one sees them on the monitor. To be accurate, therefore, one should say "This is an image of a monitor," but in that case one would exclude the monitor which the audience is actually watching. In this video the spoken word "monitor" is referring to both the image as well as to reality. The words and the image seem to form a tautology, but they verify that the sentence: "This is a monitor" could have plural images.

If I apply this relationship of words and image to the signifier and signified of Linguistics, I could say that in the case of the sentence, "This is a monitor," the signifier has plural images of the signified, and the signification is polysemic. This can be also said of the relation of the voice (signifier) and the image (signified) of "This is a camera 1."

The big difference between this piece (#2) and the first piece (#1) is not only the difference in the delivery of the text by voice or by letters, but also the editing, in that #1 is edited by cutting, and more than half of #2 by panning the camera. Using two cameras, one camera pans from the letters on paper to the monitor in which the letters are seen (No.1D); the other camera pans similarly from the letters on paper to the image of feedback on the monitor (No.2B). Further, both use backward panning as well (No.3D). Panning shows two objects, words and image, in same time axis, and effects a relationship which is the equivalent of "and" or "then" in words. Panning is unlike the disjunction which occurs in the case of cutting. In particular, the panning from the words to the monitor in which the words are seen shows this relationship literally and as a self-verifying image.

On the other hand, the pan from the words to the feedback image seems to be a jump. However, in the system of video, feedback is simply the image of self-referentiality which is formed by a closed circuit camera with the monitor. When the camera and the monitor

are in the same circuit, the monitor displays images of feedback. The first panning, to the words on the monitor, involves two cameras. One shoots the words, the other the monitor, so that the words on the monitor have been shot by a second camera, not but the same camera as in the case of the feedback. Though it jumps visually from the words to the feedback image, the signifier "This is a monitor" delivers a signified which is unique to video.

**To See The Frame**

"To See the Frame," the last piece in *Camera, Monitor, Frame*, deals with the issue of "frame" which is common to both camera and monitor.

The piece begins with big letters: "TO SEE THE FRAME (No.1A);" then a bright white frame appears inside of the monitor frame (No.2B). After the big letters appear again a black silhouette of the monitor is shown in outline (No.4C), and then the letters again, this time inside a monitor's frame (No.5D). Next frames in the form of a tunnel are generated by feedback (No.6E), followed by a monitor frame with letters (No.7D), and at the end there is a picture with only noise (No.8F), so that in every case the frame, which is referred to by the words, is different.

Unlike the case of "Seeing" in "Observer / Observed #2," "To See the Frame" designates the object precisely. By using the infinitive ("To see"), the words specify the act of "seeing;" in this case the infinitive points out the objective. By contrast, the gerund (Seeing) indicates a certain *state*. In "To See the Frame" the monitor frame, which is restricted by the definite article "the," falls within the general category of frame, but this category does not discriminate the real one which the audience is watching from a frame as an image; nor does it discriminate the image as the outline of the CRT(Cathode Ray Tube) from the outline of the monitor as a box through which the frame image of a tunnel formed by feedback. In spite of the definiteness of the article "the," this article supports the indefiniteness of the image. The indefinite article "a," in this case, would not adequately specify the frame; hence the definite article was used to fix the object of "to see." In the case of "This is a monitor," the monitor must be regarded as a total object, whereas "frame" has a generality of signification (as a common noun) but at the same time denotes a specific part of the monitor. The frame as a general attribute can be referenced by any of the various frames included in the image. Any one individual frame matters, yet there is always a most exterior outline which is common to them all. The frame is a boundary against the inside and at the same time against the outside. The frame in general functions as a container of the image and comprises its own image by itself as well. In the instruction "to see the frame," "frame" indicates not only the external outline but also the frame as an image boundary against the inside.

I would like to think about not only the frame of the monitor but also the frame of the video signal, which scans at 30 frames per second. This frame is usually invisible, but it can be made visible if one runs the videotape backward or fast-forward artificially, even though one can not handle videotape by hand like film. In

## Camera, Monitor, Frame

To See the Frame b/w, silent, 1'25"

picture plan

	A	B	C	D	E	F
picture						
description	title on black B.G.	monitor on within monitor	monitor off within monitor	title within monitor	monitor in feedback	camera off picture noise
cable connections	C-1-VTR	C-1-M C2-VTR	C-1-M C2-VTR	C-1-M C2-VTR	C-2-VTR-M	M-VTR

program		title 5" total 1'25"		
No	picture	description	min.sec	
1	A	title on black B.G: "TO SEE THE FRAME"	10	
2	B	monitor on within monitor	20	
3	A	title on black B.G: "TO SEE THE FRAME"	30	
4	C	monitor off within monitor	40	
5	D	title within monitor: "TO SEE THE FRAME"	50	
6	E	monitor in feedback	1.00	
7	D	title within monitor: "TO SEE THE FRAME"	1.10	
8	F	camera off picture noise	1.20	

Chart 5. To See the Frame

film I have a piece called "To See the Frame, Not To See the Frame"[11], in which deal with the frame in its double meaning of the frame of film and the screen onto which the film is projected.

Actually the frame of the video signal scans 60 frames per second because of the interlaced scanning of odd and even raster line numbers, alternately, at 30 FPS. So the frame functions with respect not only to the concept of space but also with respect to the concept of time as well. Therefore "To See the Frame" should also be regarded as making visible the manipulation of time. In this piece the picture noise (No.8F) which appears at the end implies, by its scan lines, the passing of time at certain intervals.

Notes:

[1] This is added and completed version to a paper first published in English under the same title in "Takahiko Imura: Film and Video," Anthology Film Archives, New York, 1990, pp. 44-50. See my remarks, just before (after a line) "This is a monitor 1", connecting the first and current versions.

[2] As stated in note [1], the present paper is a complete version dealing with all the works.

[3] Dziga Vertov wrote in one of his *Kino-Eye* manifestos (1923), "I am kino-eye, I am a mechanical eye. I, a machine, show you the world as only I can see it." On the same page he also wrote, "Now I, a camera, flying myself along their resultant..." *Kino-Eye: The Writing of Dziga Vertov*, translated by Keven O' Brien, University of California Press, Berkeley, CA, 1984, p.17. Emphasis (Italics) is mine.

[4] "It is interesting to watch such a process moving also through the history of language in relation to the word (the 'shot') and the sentence (the 'montage phrase'), and to see just such a primitive stage of 'word-sentences' later 'foliating' into the sentence, made up of separately independent words." Sergei Eisenstein, *Film Form and Film Sense*, edited and translated by Jay Leyda, The World

Publishing Company, Cleveland, OH, 1957, pp.236-237. Emphasis is mine.

[5] After several examples of the ideogram, Eisenstein wrote, "But this is montage! Yes. It is exactly what we do in the cinema, combining shots that are depictive, single in meaning, neutral in content - into intellectual contexts and series." *Ibid.*, pp.30.

[6] "I am not suggesting that each shot equals a single sentence. That is why I have placed the word 'sentence' between quotation marks through this passage. The 'correspondence' between shot and sentence is on a global scale and is derived from the fact that a shot is an actualized unit, a unit of discourse, and is inherently dissimilar to the word. The filmic shot is of the magnitude of the sentence, so to speak." Christian Metz, *Film Language, A Semiotic of the Cinema*, translated by Michael Taylor, Oxford University Press, New York City, NY, 1974, p.86. Emphasis is mine.

[7] Refer to Takahiko Imura, "The Visuality in the Structure of the Japanese Language," *Art & Cinema*, New York City, NY, Dec., 1978, pp.16-22. Reprinted in *Takahiko Imura, Film and Video*, Anthology Film Archives, New York, 1990, pp.40-43, both in English.

[8] *Eizogaku (ICONICS)*, The Japan Society of Image Arts and Sciences, Tokyo, 1983, Vol.2-7, No.27, pp.34-45. Also reprinted in Takahiko Imura, *Eizo Jikken no Tameni (For Visual Experimentation)*, Seido-Sha, Tokyo, 1986, pp.243-262, and Takahiko Imura, CD-ROM, *Eizo Jikken no Tameni (For Visual Experimentation)*, Euphonic Co., Tokyo, 1998. The English translation by the author in *Takahiko Imura, Film and Video*, Anthology Film Archives, New York, 1990 pp. 44-50. The German translation by Michael Glasmeire in *Takahiko Imura, Film und Video*, DAAD Galerie, Berlin, 1992, pp.44-50.

[9] This was co-produced with the Banff Centre for the Arts, Banff, Canada in 1998 during an artist residency, and is planned for publication in a CD-ROM version, including CG and text, in conjunction with Euphonic Co., Tokyo in English and Japanese.

[10] Takahiko Imura, "Electronic Publishing and Multimedia Art," *Kiyoku*, (the university magazine), Nagoya University of Arts and Design, Nagoya, 1998, No.4, pp.25-34, in Japanese.

[11] Takahiko Imura, film, a piece in "Models, Reel 2," 1972.

## Observer / Observed (1975-1998)

*Observer / Observed* and *Observer / Observed / Observer* are works which capture the viewpoint of a mediator situated in the same video set up as in *Camera, Monitor, Frame*, and which deal with the issues that were discussed in relation to that work.

Though *Observer / Observed* was produced earlier, as I stated before, than *Camera, Monitor, Frame* (1976), it is presented after *Camera, Monitor, Frame*, together with *Observer / Observed / Observer*, with which it comprises a pair.

The mediator I am referring to is an integral part of the set up, and not only defines the image relationship as seen from his / her viewpoint but also is often a camera operator as well. Since no complete sentence is uttered (except #3), his / her speech neither specify a sentence object, nor indicates a sentence subject clearly, but is a statement concerning "seeing" using mainly the isolated verb or gerund.

The video set up is basically a relationship of interfaces. Facing each other, a pair comprising a camera and monitor are mediated by the mediator (*Observer / Observed* #1 and *Observer / Observed / Observer* #1). Or a monitor and a mediator are positioned to face two cameras set side by side (*Observer / Observed* #2, #3). These

facing structures effect a relationship of a “round trip” of the image through the simultaneity of video. In other words, this again is the structure of “seeing” in the relation of the observer and the observed. The relation of simultaneity between the observer and the observed is indicated in the title by “/.”

## #1 Observe / Observed

Among the overall trilogy this group, *Observer / Observed*, occupies the central role. There are three pieces under this title. First comes *Observe / Observed*.

In this piece the relationship between the observer and the observed is transferred to a pair of cameras and a monitor which face each other. A person stands beside the camera. This set up was already used in “This is a camera 1,” but in that case the person at first stands independently, separated from the camera; then, after the stages of dialectic between the words and images, at the end he says, “I am a camera,” looking through the camera.

In the present piece, the person stands beside the camera from the beginning, and responds to the camera. At first we see a close up of both eyes of the person (No. 1A); she looks up and down, and a female voice (the same person’s voice) utters “observe” corresponding with the movements of the eyeballs. In the 1975 version, “observe” was uttered by a male voice, but this time it is said by the person herself. The use of the male voice not only gave the viewer an order to “Observe” but inappropriately also gave an order to the person in the picture to “Observe,” through its having been spoken by the male voice.

This time, with the order being given by the person herself, this misconception can be avoided. Of course the voice can also mean an order this time as well, but it can be understood that the order is not issued by a third person toward the person in the picture, and that she herself utters the order to a third person outside of the picture.

Next a camera is seen, then the image pans down to a person with a camera shown on a monitor (No.2B). And again, it pans up to the camera, repeating the up and down movement. This up and down movement is a response to the movement of the eyeballs at the beginning. At the shot of the camera, “Observe” is uttered; at the shot of the pan-down to the person with a camera shown on a monitor, “Observed,” the past form, is uttered. This is not intended to mean the past tense, but indicates the passive voice by using the past perfect, as in “to be observed.” The relation of *Observe / Observed*, therefore, is repeated with the movement of the image panning up and down. One might wonder why the image of the person with a camera, in the monitor, is regarded as “Observed.” But if one considers that all images are subjects for “observing,” then the reason behind the word “Observed” for the image inside of the monitor is following.

## Observer / Observed

#1 Observe / Observed b/w, sound, 2'40"

picture plan

	A	B	C	D
picture				
description	eyes	camera & face with camera	camera & monitor in feedback	face with camera monitor in feedback
cable connections	C.2-VTR 2	C.2-M.1 C.1-VTR	C.1-VTR-M.1	C.2-M.2 C.1-VTR-M.1

program		title 5" total 2'40"		
No	picture	description	voice (dubbed)	min.sec
1	A	eyes (look up)	Observe	30
		eyes (look down) (repeated)	Observe (repeated)	
2	B	camera (pan up & down)	Observe	1.00
		face with camera within monitor (repeated)	Observed (repeated)	
3	C	camera (pan up & down)	Observe	1.30
		monitor in feedback (repeated)	Observed (repeated)	
4	D	face with camera (pan right & left)	Observed	2.00
		monitor in feedback (repeated)	Observed (repeated)	
5	A	eyes (look to right)	Observe	2.35
		eyes (look to left) (repeated)	Observe (repeated)	

← pan right & left | pan up & down

Chart 6. Observer / Observed

Consider here the context of the image. At first naked eyes are seen saying “Observe;” next only a camera is seen again, saying “Observe;” and then a monitor which is shot by the camera shows a person with a camera. The camera seen as an object, therefore, is the camera which shoots (observes) the person who is seen next, and as a result, one sees the image of a monitor in which there is a person with a camera. Accordingly if one says “Observe” to the camera, one could say “Observed” to the image of a monitor which is being shot by the camera. This is because the photographer and the one who is photographed stand in the relationship of “Observe / Observed.” In other words, video is a system which is capable of visualizing as its object both photographer and photographed. (Obviously this is not a capability of film). Usually the audience watches only the object, and it is very rare to observe the photographer, as an object, at the same time.[1] Further, in this case the recording camera is not the one which is “observing” but is the one that is “Observed” on the monitor. One can notice that the “observing” camera is fixed, but the “Observed” camera is moving up and down, handled by the person. This is of course exactly the up and down panning which the audience is really watching. Therefore, it is not the normal equation of “Observe” = record, but the equation “Observed” = record that is formed here.

Next the picture again shows the camera, and when it pans down this time the image it turns into feedback (No.3C). Here again one hears the voice saying “Observe” and “Observed.” This is because the monitor is connected not to the “observing” camera, but rather to the “Observed” (recorded) camera, which causes the phenomenon of feedback. When the camera and the monitor are in the same circuit, the camera shoots (observes) the monitor, and the “Observed” image becomes amplified images of the monitor. In



contrast to the previous shot, which switched between images of human eyes and the camera, it is a quite interesting phenomenon that here in the present shot the “Observed” monitor is amplified by the “observing” camera.

Next the picture repeats horizontal panning between a monitor with the image of feedback and a monitor with the image of a person holding a camera (No.4D), and the word “Observed” is heard for both. In this case the image was also recorded by the camera which is panning. Here one hears only “Observed” for both, and one notices the lack of “Observe,” but since both images are shown within a monitor, as I have already explained, one can understand the rationale for this case.

At the end we come back to the close up of both eyes as in the first image, and the eyeballs move from right to left horizontally, in response to the previous shot (No. 5A).

This work deals with the conjugation of the verb in active and passive voices. For the active voice we have the direct image of the close-up of both eyes, and for the passive voice we have the indirect image with a monitor. The monitor, especially, as a medium, provides the image which corresponds to the passive voice. It is interesting to consider whether another medium, in the case for instance of “Listen” (ear) and “Listened” (speaker), could function in the same way.[2]

## #2 Seeing / Not Seeing

“Seeing / Not Seeing” is a silent piece with superimposed letters. Basically the work consists of images on a monitor which contrast the closing and opening of eyes shown in close-up and feedback image of the monitor, which replaces the eyes.

As I mentioned already at “To See the Frame,” “Seeing” is a gerund, not an infinitive. This is an important point, because the gerund, with “ing” which is similar to the progressive form of the verb, is considered here as a state or condition. Without any specific object, “Seeing” and “Not Seeing,” which are seen as images in the monitor, are themselves contrasted.

At first a hand comes in, switching on the monitor (No.1A); as soon as both eyes are seen the letters “Seeing” are superimposed (No.1B), only to disappear when the image is switched off. Then closed eyes come on with the letters “Not Seeing” superimposed (No.2C); with the opening of the eyes the message changes to “Seeing (No.2B);” and again when the image is switched off “Not Seeing” is superimposed (No.2A). In these sequences, for “Seeing” opened eyes are shown, and for “Not Seeing,” either closed eyes or no image, the empty monitor, are shown.

Though in two cases we are seeing “Not Seeing,” in one case it is the eyes in the monitor which are “Not Seeing” and in the second case it is we who are “not Seeing” the image. In the former case, the subject is the eyes in the image; that is, “She (the eyes) is not seeing.” In the latter case “We are not seeing the image.” Even though the words are the same, “Not Seeing,” the difference is whether the subject is in the image or is us (the audience), and whether the object is in the image or not.

Recently I showed this work and an audience member commented that it seemed to resemble the painting by Rene Magritte, “This is

## Observer / Observed

#2 Seeing/Not Seeing		b/w, silent, 1'58"			
picture plan					
	A	B	C	D	
picture					
description	monitor off within monitor	opened eyes within monitor	closed eyes within monitor	monitor in feedback	
cable connections	C.2-VTR	C.1-M1 C.2-VTR 2	C.1-M1 C.2-VTR 2	C.2-VTR 2-M2	
program title 5" total 1'58"					
No	picture	description @ @			min.sec
1	A		a hand in frame. switches on monitor		10
	B		opened eyes on screen. "Seeing" superimposed		20
	B		a hand in frame. switches off monitor		30
	A		no picture on screen		35
2	A		a hand in frame. switches on monitor		40
	C		closed eyes on screen. "Not Seeing" superimposed		50
	B		closed eyes opened. "Seeing" superimposed		58
	B		a hand in frame. switches off monitor		1.11
	A		no picture on screen. "Not Seeing" superimposed		1.20
3	A		a hand in frame. switches on monitor		1.25
	D		monitor in feedback. "Seeing" superimposed		1.34
	D		a hand in frame. switches off monitor		1.42
	A		no picture on screen. "Not Seeing" superimposed		1.53

Chart 7. Seeing / Not Seeing

not a pipe.” The painting by Magritte is a denial of the picture (two pipes appear, one on a canvas painted within the painting, the other painted as though floating in the air above the other). The sentence, “This is not a pipe,” has a specific (if doubled) object. In my video no object for denial is shown, only a denial for sight: “Not Seeing.” There is even a self-contradiction in simply seeing (reading) “Not Seeing.”

In that sense this phrase is close to John Cage’s statement, “I have nothing to say and I am saying it.”[3] Further, the subject of “Not Seeing,” as I mentioned already, changes depending on the image in the monitor. If I had said in the video, “This is not a monitor,” one might have pointed out the similarity with the painting by Magritte; but what I am concerned with is the relation of the gaze between the image and the audience.

At the end, instead of the gaze of the person, “Seeing” is superimposed over an image of feedback which is caused by the camera seeing the monitor (No.3D), and then after switching off the monitor, “Not Seeing” is seen over no image (No.3A). This is the gaze of the camera eye, which has replaced the human eyes, and so the piece demonstrates that the video medium can not only substitute for human eyes but also has the capacity to produce its own image.

## #3 She Sees / Seen

This piece consists of the opening / closing of eyes, the eyes shown on a monitor, and the absence of an image. The voices speak four sentences: “She sees,” “She is seen,” “She does not see,” and “She

is not seen.” These are affirmative or negative sentences, in active or passive voices, and each has the subject “She.” The narration is by a male voice to emphasize that it was spoken by a third person. Though it is not intentional, the pronunciation of “She” and “See” is confusing, and hard to distinguish in the inflection of the Japanese speaker. So there are cases which sound like “She shes” for “She sees,” and which emerge as unexpected puns on the words.[4]

Back to the work, the four spoken sentences have corresponding images attached: “She sees” = Opened eyes, “She does not see” = Closed eyes, “She is seen” = Opened and closed eyes on a monitor, and “She is not seen” = No image on the monitor and picture noise only, without a monitor.

In the active voice each word has a corresponding image, one to one, but the passive voice has plural images. Thus regardless of whether her eyes are open or not, there is the affirmative sentence “She is seen,” and regardless of whether the monitor is seen or not, there is the negative sentence, “She is not seen.”

Particularly in this piece, a question may be raised concerning the circumstance that “She is seen” is heard regardless of whether the subject’s eyes are opened or closed. But the presence of the monitor, which contains her image, makes her the object “to be seen.” This does not affirm the traditional view of woman as an object “to be seen,” in spite of the fact that the monitor imprisons the subject inside of the frame. As an image “she” claims herself to be not only one who “is seen” but also as one who “sees,” by the opening / closing of her eyes. At the same time, the image “She sees” is in fact the one which “is seen” by the audience. The constitution of the subject by the statement, “She sees,” makes the object “She is seen” work in this way as well.

The number of affirmative and negative sentences and active and passive voices is not balanced in relation to the images. For the four kinds of sentences, there are six images altogether, among which the three images each of affirmative and negative sentences are balanced; but there are two active and four passive voices, whose numbers consequently are not balanced. The passive voice has a greater variety of images.

Yet another argument to be accounted for is that an image has no negation, only affirmative properties. Even the condition of no image can be regarded as an image. In this case, “having no image” on the monitor is nevertheless a condition in which the monitor is an image, and the picture noise itself is an image. The negation exists only in words. This argument is familiar, but what I have attempted is not to determine only whether an image in general has a negative form or not. One can hypothesize an image corresponding to negative form-needless to say using words-and following the logic of words, and relying upon a conjectured relationship between images and words and this image, of the negative form corresponding to “She sees” (opened eyes), must be the image of closed eyes, wherein of course the audience can recognize the image of “She does not see.”

In general all of the images are to “be seen” (passive voice) by the audience. But if one supposes that the image can have a subject, then an image for “She sees” in the active voice can be established. Shouldn’t this work be regarded as giving the active position of a sentence subject to an image (the female) which is normally “seen” as an object, thereby also empowering the (female) image to see?

### Observer / Observed

#3 She Sees/Seen b/w, sound, 3'15"

picture plan

	A	B	C	D	E	F
picture						
description	opened eyes	opened eyes within monitor	closed eyes	closed eyes within monitor	monitor off within monitor	camera off picture noise
cable connections	C.1-VTR	C.1-M.1 C.2-VTR	C.1-VTR	C.1-M.1 C.2-VTR	C.1-M.1 (off) C.2-VTR	C.2 (off)-VTR

program title 5" total 3'15"

No	picture	description @ @	voice (dubbed)	min.sec @
1	A	opened eyes	She sees	10
2	B	opened eyes within monitor	She is seen	20
3	C	closed eyes	She dose not see	30
4	D	closed eyes within monitor	She is seen	40
5	E	monitor off within monitor	She is not seen	42
6	B	opened eyes within monitor	She is seen	50
	D	closed eyes within monitor	She is seen	1.00
7	A	opened eyes	She sees	1.12
	C	closed eyes	She does not see	1.26
8	F	camera off picture noise	She is not seen	1.30
9	B	opened eyes within monitor	She is seen	1.40
10	A	opened eyes	She sees	1.51
	C	closed eyes	She dose not see	2.00
11	D	closed eyes within monitor	She is seen	2.11
12	E	monitor off within monitor	She is not seen	2.20
13	D	closed eyes within monitor	She is seen	2.30
14	C	closed eyes	She dose not see	2.40
	A	opened eyes	She sees	2.50
15	B	opened eyes within monitor	She is seen	3.00
16	F	camera off picture noise	She is not seen	3.10

Chart 8. She Sees / Seen

Notes:

[1] *The Man With a Movie Camera* (1926) by Dziga Vertov is a pioneering example. In this film, not only does “The Man With a Movie Camera” appear, but also, through the double exposure of camera lens and human eye, the film metaphorically captures the camera as an object.

[2] If one were to use a microphone and a speaker as the media for a work, “Listen,” there would remain a shortfall in its expression by the visual medium of video. “Listen” is not as perceptually selective as “See,” and the difference between “to listen” and “to be heard” is not so clear as that between “to see” and “to be seen.” Sound has the property that it is “to be heard” indiscriminately. The ear does not have the directionality of the eyeball, and the difference between hearing (the first information type) and reproduced sound (the second information type) is not so great as the difference between viewing (the first information type) and image (the second information type).

[3] John Cage *I have nothing to say and I am saying it*, video tape published in series of American Masters, The Music Project for Television, Inc. and American masters, 1990.

[4] Professor Peter D'Agostino of Temple University remarked that "This is a feminist piece."

## Observer / Observed / Observer (1976-1998)

This work is an extension of the previous tape, *Observer / Observed*. In *Observer / Observed / Observer* the issues which were dealt with previously are further structured through observing of the observer of the observed; and the observed, this time, is also observing the observer realizing the "round trip" structure of the video image. In other words, the one who sees observes the one who is seen at the same time as the one who is seen observes the one who sees, thereby switching over the viewpoint. Philosophically speaking, the position of subject and object circulate freely and as interchangeable entities with each other. This economy of the image is a mechanism unique to video, one which film doesn't have, and it means that video operates as a system.

What makes this possible is the operation of double feedback, in which two sets of cameras and monitors are facing each other. In piece #1 a person is mediated between the camera and the monitor, and in pieces #2 and #3, which consist of only the camera and the monitor without the person, the image pans repeatedly between two cameras facing each other. In this work three media, image, letters, and voice, deal respectively with exactly the same phenomena. However, the voice, which is uttered, defines the structure most completely; next are the letters, which are abbreviations; and last is the image, which "represents." The three media simultaneously supplement each other and make the complicated structure perceptible. In the letter abbreviations, "C" means "camera," "M" means "monitor," and "." between two letters indicates the shooting by the former of the latter. "/ between two letters indicates the relation of shooting and being shot, in which the former is shot by the latter. The spoken parts are either simple sentences or complex sentences in which more than two simple sentences are conjoined by relative pronouns; the simple sentences correspond to the image shot directly, and the complex sentences to the feedback image, which includes both shooting and being shot.

### #1 I See You / Myself

This piece comprises four pictures: two faces (male and female) each next to a camera are seen directly; another two faces (the same ones) are each seen on a monitor within another monitor (which was shot by the camera), facing each other. These pictures are shot by two cameras which face one another, with the male and female faces placed next to each camera. Two monitors are also there, facing one another, and each person repeatedly pans between the opposite person and the opposite monitor. The voice utters two sentences; "I see you who is shooting me" and "I see myself who is shooting you." The former sentence accompanies the directly-shot and the latter accompanied the image on a monitor within another monitor, shot by the other camera. The sentences establish "I" as the subject and make clear that there are the "I" who is normally not known (the "I" of the camera's point of view). The "I" who has

"you" or "myself" as its object of seeing and which also has "me" or "you" as its object of shooting, therefore, "I" has two objects which make up the complex sentence.

The first picture is the direct shot of a man with a camera, accompanied by the sentence (spoken by a woman), "I see you who is shooting me (No.1A)." In this case, "I" is a woman who is outside of the picture, and is the object of the shooting; and "you" is the man who is inside the picture. At the bottom of the picture, the pronouns which are uttered, "I-You-Me," are superimposed. These pronouns are in the order which they are uttered in English; in Japanese, by the way, the order of the words "I - You - Me" is different from that in English. At this moment who the "I" is, who is apparently the object of the shooting, and who will appear in the next picture, is not clear. The "I" as a voice is invisible, and the "I" as a letter is abstract, so that in this situation "I" as a substitute for the audience is also conceivable. Next, through a pan, a woman with a camera is seen on the monitor and the voice, which is synchronized with the picture, says, "I see myself who is shooting you (No.1B)." The superimposed letters are "I - Myself - You;" since the pan limits the shot to objects which are physically adjacent within the same space, "you" can only mean the one who is located opposite, on the other side.

Since the picture is synchronized with the sound, the viewer knows that the camera in the monitor image is actually the one which is being shot / recorded, and what we see on the monitor inside the monitor, shot by the other camera, is what can be called a "reflexive" image, in the same sense that the word "myself" is a reflexive term. The third picture is of a woman with a camera and is narrated by a man this time: "I see you who is shooting me (No.2C)." This is the opposite of the first picture; here the subject "I" is a man who is outside of the picture. The superimposed letters are same, "I - You - Me." There is a pan to the fourth picture, which is of a man with a camera who says, "I see myself who is shooting you (No.2D)." The superimposed letters are "I - Myself - You." The picture and the sound are synchronized.

So far I have explained four kinds of pictures. From this point forward the four pictures repeat, but this time the direction of the pan is reversed, now going from the picture in the monitor to the image which is shot directly. This means that the piece now moves from the reflexive image to the direct image, and presents the relationship of the two in the passage back and forth between them (No.3B, 3A, No. 4D, 4C).

This piece realizes a structure of simultaneity in which the two persons influence each other through the images and texts; that is, through their functions as the observer and the observed, and as the photographer and the photographed. These four standpoints of the two persons overlap each other, so that in practice two standpoints, the observer / the photographed and the observed / the photographer, are articulated through the two kinds of sentence and are identified by the sources of the sound, which is delivered accordingly either inside or outside of the picture.

**Observer / Observed / Observer**

**#1 I See You/Myself** b/w, sound, 1'57"

picture plan

	A	B	C	D
picture				
description	male with camera "I-You-Me" superimposed	female with camera within monitor "I-Myself-You" superimposed	female with camera "I-You-Me" superimposed	male with camera within monitor "I-Myself-You" superimposed
cable connections	C.1-M.1 C.2-M.2-VTR 2	C.1-M.1 C.2-M.2-VTR 2	C.1-M.1-VTR 1 C.2-M.2	C.1-M.1-VTR 1 C.2-M.2

program title 5" total 1'57"

No	picture	description	voice (sync)	min.sec
1	A	male with camera "I-You-Me" superimposed (pan to B)	(♂) I see you who is shooting me	13
	B	female with camera within monitor "I-Myself-You" superimposed	(♀) I see myself who is shooting you	27
2	C	female with camera "I-You-Me" superimposed (pan to D)	(♀) I see you who is shooting me	43
	E	male with camera in monitor "I-Myself-You" superimposed	(♂) I see myself who is shooting you	57
3	B	female with camera within monitor "I-Myself-You" superimposed (pan to A)	(♀) I see myself who is shooting you	1.11
	A	male with camera "I-You-Me" superimposed	(♂) I see you who is shooting me	1.25
4	D	male with camera within monitor "I-Myself-You" superimposed (pan to C)	(♂) I see myself who is shooting you	1.39
	C	female with camera "I-You-Me" superimposed	(♀) I see you who is shooting me	1.52

← left pan      ♂ male voice      ♀ female voice  
→ right pan

**Chart 9. I See You/Myself**

**#2 Camera 1/2 - Monitor 1/2**

This piece was produced with the same video set up as # 1, with two facing pairs of cameras and monitors, and realized the interrelational structure of the cameras and monitors but without the two mediators. Though the basic structure is the same as #1, the description (the spoken text and the superimposed abbreviation letters) is different. Here the description involves the relation of "camera 1 and 2" and "monitor 1 and 2" instead of the pronouns "I" and "you." By repeating the pan, the peculiar condition of "double feedback" is created. When one person's camera pans between the other person's camera and their monitor, then (depending on whether the other person's camera is shooting one's own camera or the adjacent monitor) the image will be either one's own camera in the monitor within the monitor, or a feedback image of the monitor.

At first camera 2 shoots camera 1 directly and the voice states, "Camera 2 shoots camera 1(No.1A)." The superimposed letters are "C.2 - C.1." After the pan, when camera 1 shoots camera 2, camera 2 shoots monitor 1 which was shot by camera 1. In monitor 1, we see camera 2, while camera 2 shoots its own image (No.1B). The voiceover is "Camera 2 shoots monitor 1 which is shot by camera 1 which shoots camera 2." The superimposed letters are "C.2 - M.1 / C.1-C.2." Next camera 2 again shoots camera 1 directly, but the shot is cut off. Then, when camera 1 shoots camera 2, camera 2 shoots monitor 1, which was shot by camera 1. The voiceover is "Camera 2 shoots monitor 1 which is shot by camera 1 which shoots monitor 2 (No.1B)." The superimposed letters are "C.2 - M.1 / C.1 - M.2." In the image of the monitor, the numbers 1 and 2 are seen alternating within the form of a tunnel (double feed-

**Observer / Observed / Observer**

**#2 Camera 1/2-Monitor 1/2** b/w, sound, 1'53"

picture plan

	A (A')	B	B'	C(C')	D	D'
picture						
description	camera 1	camera 2 within monitor 1	monitor 1 & 2 in feedback	camera 2	camera 2 within monitor 1	monitor 1 & 2 in feedback
cable connections	C.2-VTR 2	C.2-VTR 2 C.1-M.1	C.2-M.2-VTR 2 C.1-M.1	C.1-VTR	C.1-VTR 1 C.2-M.2	C.1-M.1-VTR 1 C.2-M.2

program title 5" total 1'53"

No	picture	description	voice (dubbed)	min.sec
1	A	camera 1 "C.2-C.1" superimposed (pan to B)	Camera 2 shoots camera 1	15
	B	camera 2 within monitor 1 "C.2-M.1 / C.1-C.2" superimposed (pan to B')	Camera 2 shoots monitor 1 which is shot by camera 1 which shoots camera 2	32
	B'	monitor 1 & 2 in feedback "C.2-M.1 / C.1-M.2" superimposed	Camera 2 shoots monitor 1 which is shot by camera 1 which shoots monitor 2	52
2	C	camera 2 "C.1-C.2" superimposed (pan to D)	Camera 1 shoots camera 2	1.03
	D	camera 2 within monitor 1 "C.1-M.2 / C.2-C.1" superimposed (pan to D')	Camera 1 shoots monitor 2 which is shot by camera 2 which shoots monitor 1	1.21
	D'	monitor 1 & 2 in feedback "C.1-M.2 / C.2-M.1" superimposed	Camera 1 shoots monitor 2 which is shot by camera 2 which shoots monitor 1	1.39
3	A	camera 1 "C.2-C.1" superimposed	Camera 1 shoots camera 2	1.47

← left pan  
**Chart 10. Camera 1/2 – Monitor 1/2**

back). The spoken sentence here is not easy to understand unless one is quite attentive. Though it is possible that a listener would find it incomprehensible during one reading, the abbreviation in letters does help to make the meaning clear.

The above has described the images which were shot by camera 2; next seen are the images which were shot by camera 1. Here cameras 1 and 2, and monitors 1 and 2, are reversed from their previous roles, and in the case of the double feedback on monitor 2, the order of the numbers is 2 and then 1, as they are seen alternating within the feedback tunnel shape. This interchange of the numbers is also carried over to the voiceover and the superimposed letters.

Thus the complex relationships between shooting and the images which are effected by two cameras and two monitors have been examined, and through their carefully articulated relationship they reveal for the first time the structure of the interaction between "seeing" and video. Not only do we find a deep relationship between "seeing" and video, but, since the modalities of the "reading" of text and the "listening" to a voice are also included, one could say that here the structure of perception is also being revealed. What I have attempted is to challenge the limits of perception in these "complex relationships."

**#3 Camera 2 - Camera 1 / Monitor 1/2**

As in #2, here two cameras and two monitors are used; however, this time two monitors are placed beside camera 1, facing camera 2, and all the shots are taken by camera 2 only. The piece deals mainly with the movement relationships created by panning between camera 1, monitor 1, and monitor 2, rather than articulating the various structural relationships, as in #2. Therefore the voiceover here takes the form, "pan from - to -"




At first we see camera 1, which is shot by camera 2 (No.1A); then the shot pans to monitor 1, on which we make out “our own” image (our camera, camera 2), which is being shot by camera 1. The voiceover, which is spoken at the beginning of the pan, is “Camera 2 pans from camera 1 to monitor 1 which is shot by camera 1 (No. 1A . 1B).” The superimposed letters are “C.2 - C.1 . M.1 / C.1.” (where . means pan). Next the shot pans farther to monitor 2, which is being shot by camera 2. In this case not double feedback but simply only the everyday kind of feedback, between camera 2 and monitor 2, is seen. The voiceover states, “Camera 2 pans from monitor 1 to monitor 2 which is shot by camera 2 (No.1B . 1C) .” The superimposed letters are “C.2 - M.1.M.2 / C.2.” Altogether there are three different images here, along with the two pans between them: camera 1, camera 2 seen on monitor 1, and monitor 2’s feedback. Through the pans, which restrict the shot to one space, but which nevertheless move between the opposite camera, the monitor on which our own camera is seen, and the feedback image, a “round trip” between other and self is effected. The piece is not just a group of arbitrary movements of the two cameras, but is a circle-like coordinated movement in which pans by one camera are reflected in movements by the other. Visually there are only repeating pans, but these movements convey a deeper structure, in traversing from the reflexive images to the feedback images. The piece thus develops from a static structure into structured movement. The question here is whether as a viewer one is able to capture the meanings of this movement fully and in depth.

(Translated from Japanese by the author with the assistance of Tony Conrad, Professor of the State University of New York, Buffalo, New York).

**Observer / Observed / Observer**

#3 Camera 2-Camera 1/Monitor 1/2 b/w, sound, 1'33"

picture plan

	A	B	C
picture			
description	camera 1	camera 2 within monitor 1	monitor 2 in feedback
cable connections	C.2-VTR 2	C.2-VTR 2 C.1-M.1	C.2-VTR 2-M.2

program title 5" total 1'33"

No	picture @	description	voice (dubbed) @	min.sec
1	A	camera 1 (pan to B) *C.2-C.1 = M.1/C.1* superimposed	Camera 2 pans from camera 1 to monitor 1 which is shot by camera 1	6
	B	camera 2 within monitor 1 (pan to C) *C.2-M.1 = M.2/C.2* superimposed		20
	C	monitor 2 in feedback		28
2	B	camera 2 within monitor 1 (pan to A) *C.2-M.1 = C.1-C.2* superimposed	Camera 2 pans from monitor 1 to monitor 2 which is shot by camera 2	36
	A	camera 1 (pan to C) *C.2-C.1 = M.2/C.2* superimposed		48
	C	monitor 2 in feedback		58
3	C	monitor 2 in feedback (pan to B) *C.2-M.2 = M.1/C.1* superimposed	Camera 2 pans from monitor 2 to monitor 1 which is shot by camera 1	1.05
	B	camera 2 within monitor 1 (pan to A) *C.2-M.1 = C.1-C.2* superimposed		1.19
	A	camera 1		1.28

→ right pan

**Chart 11. Camera 2 – Camera 1 / Monitor 1/2**