*Proceedings of 5th International Conference on Computational and Cognitive Models of Creative Design*. Gero, J. S. and Maher, M. L. (eds), Sydney University, December 2001, pp 173-185

# CONSTRUCTING INTER-RELATIONSHIPS

*Computations for Interaction in Art*

E.A.EDMONDS, J. DIXON
*Creativity & Cognition Research Studios*
*Department of Computer Science*
*Loughborough University*
*Leicestershire, LE11 3TU, UK.*

**Abstract.** The thinking behind the author's time-based and interactive art is briefly reviewed from the early inspirations of Malevich to current logic based work employing image recognition. In this work one can think about the implications of underlying structures in ways that were not possible without computers. Generating time-based and interactive work of this kind was transformed by the computer. It enables the effective construction of inter-relationships that provide computational models for building correspondences between forms and systems. The paper describes the computational methods used to introduce interaction to the video construct works in which the body movements of the observer influence the visual behavior of the work.

## 1. Background

In Malevich's essay of 1919, "On new systems in art" (Malevich, 1968), he introduced the notion of making art with the help of "a law for the constructional inter-relationships of forms". Patricia Railing, in "From Science to Systems of Art" (Railing, 1989) shows how Malevich was looking towards science to give a formal basis to his art. Developments of such notions were widespread in the twentieth century For example, in a discussion of the British Systems movement, Bann relates this view to a linguistic one:

*"It is not the recurrence of the rectangle in Van Doesburg's work which needs to engage us, but the series of relationships between rectangles and the extent to which those relationships can be adequately formalised."* (Bann, 1972).

The interest of artists in systems was a natural pre-cursor to an interest in computers. After all, a computer is basically a logical machine that manipulates systems of symbols. When a computer manipulates such a system, it, as it were, automatically moves the pieces and so searches for some result, such as checkmate.

The output that a computer generates, be it text, sounds, drawings or movement, for example, is simply a mapping from a set of internal symbols. For all of its ability to calculate quickly or put images on screens it is the operation of formal systems that is special about the computer. Thus, for some artists, the computer was a natural and irresistible medium to explore. A significant exhibition in which much of this early exploration was shown was Cybernetic Serendipity held at the Institute for Contemporary Art in London in 1968 (Riechardt, 1968). Cybernetic Serendipity and other related activities of that era all showed the coming together of technologists and artists in collaborations of one form or another.

Malevich's insight about science and systems took an important step forward with the artists who adopted programming in the sixties – including Manfred Mohr, Harold Cohen and myself, for example. A program *is* a system as mentioned above and so here was a new way of applying systems. Each artist came at the issue from his or her personal point of view and not all of them would accept the idea of a system as used in art; but, if they programmed they certainly used a system in practice.

At the first Computational Modelling of Creativity Conference, held on Heron Island in 1988, I presented a case that put the use of formal systems at the core of a certain kind of creativity in art practice.

*"Knowledge-based systems have introduced a new mechanism for the support of creativity. Where experts interact with a knowledge base in order to externalize and refine their knowledge they are involved in a process of coming to new understandings. It is that process that is a significant trigger for creative thought."* (Edmonds, 1993).

This paper follows the earlier one and brings the developments up to date with the extension into the construction of interactive works.

## 2. Using Computers in Art Systems

### 2.1 EARLY WORK WITH COMPUTERS

More than thirty years ago I used a computer for the first time to perform an art task for the construction of a work. I was making a constructed relief, but I had a problem. I had many bits and pieces and I wanted to arrange them according to certain rules. I found it very hard. Always, when I made an arrangement, it broke one of the rules I wanted to satisfy. I had access to a computer and I managed to obtain three hours of computer time. I had to switch the computer program off after three hours because someone else needed it, but I had not quite solved the problem. However, my program had reduced it into something I could solve myself. So I finished the job off by hand (Cornock and Edmonds, 1973). That was good but did not excite me very much in terms of using computers. Much later it became clear that many of the structural issues that had concerned me, such as the co-existence of two colour in the same space, could be tackled by moving from static to time-based work. By this time, such work was made a possibility by the fact that computing technology had moved on and could support it in a unique way.

2.2 VIDEO CONSTRUCTS

The work that I am concerned with now is what I have come to call ‘Video Constructs’ (Edmonds, 1988, 1989). These pieces are time-based, that is, they exist in time just as music and film do. The concrete and final destination of the images is a video monitor. However, the fact that it is generated through a computer system allows considerable attention to be paid to the structures that underlie the images, and their movement in time. In video constructs, the logic in the computer provides the underlying structure that leads to the form of the work. The image on the screen is the concrete reality of the work that holds a determined correspondence with the underlying time-based structure.

The most exciting element of the constructive video is, perhaps, the careful and very terse way in which a specification of what occurs in time is possible. The brevity of the specification is extremely important in the development of ideas. The inevitable exploration is so strongly supported by this aspect of the use of the computer that new ways of thinking about work emerge in their very construction (figure 1).

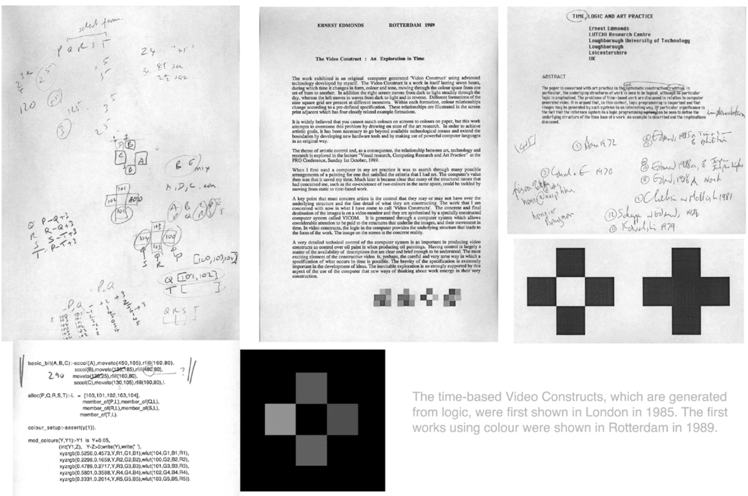


*Figure 1: Note on Video Constructs (1985/90), 2000*

The exploration of time-based constructive work made possible by modern computer technology is more than a new way of doing something. The conceptual development that goes along with the art practice is something new that has implications beyond video constructs. The new understandings will inevitably influence drawing, painting and construction as much as they are influencing the video.

## 3. Interactive Video Constructs

Most recently, I have returned to my early concern for the use of computers in the development of participation. The time-based video constructs have become *interactive* video constructs. It is not hard to understand how the structures in time can be so constructed as to react to events detected by sensor systems. In the example of *Kyoto* (2001), A real time image analysis system is incorporated. The work includes a digital video camera that points towards the space in front of the work: it looks at the audience. The system uses motion analysis designed to detect human activity and so reports to the video construct computer about the presence of people and about the degree of motion (e.g. “fast waving of a left arm”). The behaviour of the piece is then reactive to what participants are doing. The generative system at the heart of the work has a real-time correspondence with the participant’s movement and, in turn, a visualised correspondence with the images displayed (figure 2).

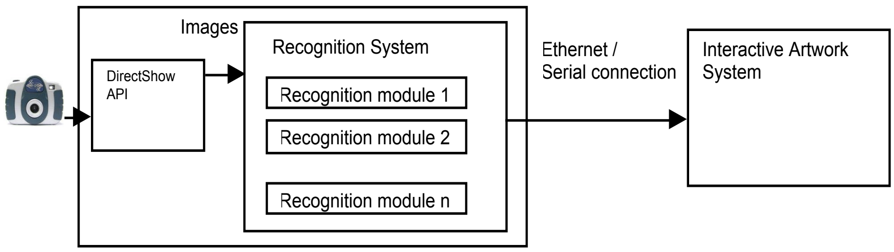


*Figure 2: The image recognition integration with the video construt*

### 3.1. INTERACTIVE SYSTEM: MODELLING USER BEHAVIOR

*Kyoto*, and similar work, incorporate a human motion recognition system designed to operate with a single person in front of a camera (Dixon, 2001). A video stream from the camera is taken as input and processed in real time. Continuous interpretation of motion is provided from the images in the video stream. The system builds a model of human movement in front of the work and passes it, in real time, to the computer controlling the images. The model can represent the following:-

- Presence/absence of a person
- Distance from the camera

- Quantity of motion
- General direction of motion
- Number of separate motions
- Which body part caused each motion
- Position of each motion
- Recognition of specific motions (e.g. waving)

In order to create this data, the human in front of the camera needs to be differentiated from the background. This is achieved by thresholding the input images. Thesholding involves ignoring pixels which have one attribute above or below a certain value. In this case the attribute used is light intensity. The background and the person must be of a significantly different intensity. The system can operate in two modes: When the background is lighter than the person and when the background is darker than the person.

For simplicity, the system assumes that a person is standing upright, facing the camera and that their full figure is within the frame (figure 3). Recognition information can still be provided if up to half of the person is off the edge of the image. A subset of information is still available if there is more than one person in front of the camera. The quantity of

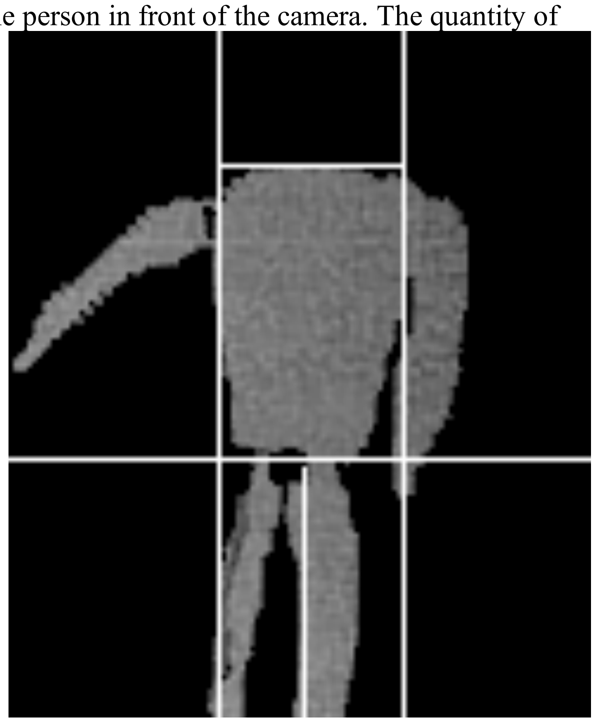

*Figure 3: A figure segmented into meaningful regions*

motion is available and the distance from the camera can be provided, but this becomes distorted if the foreground becomes particularly cluttered.

The system works most reliably in constant light levels. Calibration is required for each installation. A reference image of the unobscured background must be taken so that the foreground can be isolated during normal running.

## 3.2 PROGRAM STRUCTURE

The core of the program is a loop which continually iterates in order to capture images from the camera and to recognise motions in them. Before this processing can commence, image acquisition needs to be initialised and any calibration for the surrounding environment executed. Whenever the system is set-up in new environment the background is potentially different. In order for the background to be removed from the image, a reference picture of the background is taken. A flow diagram of the program structure is shown in figure 4.

The frame rate at which the images are captured and processed at is significant. The time taken to process the images and recognise the motion is a limiting factor, but the optimum time difference between frames needs to be determined. Assuming that the processing will not take longer than the desired frame rate, a delay must be introduced so the system waits until the correct time to capture the next image. This is normally the case.

## 3.3 INTERACTION: INFORMATION FOR THE ARTIST TO USE

The design of the program produced some user interaction requirements. In the context of this system, user interaction has two meanings: the user who configures and sets up the system, e.g. in preparation for an exhibition, and the user who interacts with the art work once the system is running. The former is the artist and the system enables this user to experiment and explore different settings. This is used in combination with the explorations involved in building the image part: the video construct. This, then, extends the scope of the creative computational investigations, described in the earlier papers, to include interaction.

With regard to the operator, a dialogue will occur during the initialisation phase. The system needs to be supplied with an unobscured picture of the background. This must be taken once the camera is in position. The camera must not be moved after the reference background picture has been taken. If the artwork and camera are required to be moved to another location, then the system must be restarted and a new background reference picture must be taken as part of initialisation. The system will require a delay between frames to be specified so that images are processed at configurable regular intervals.

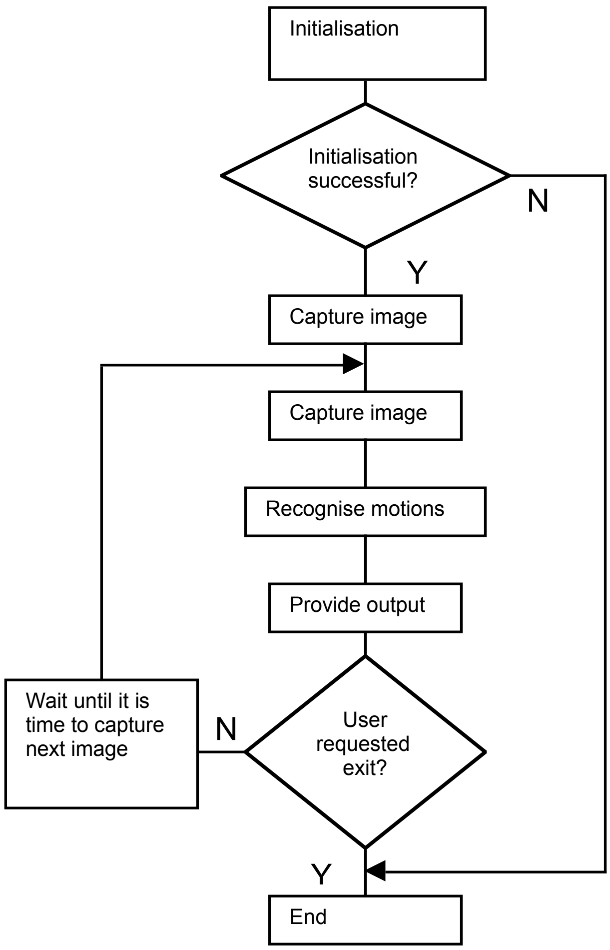


*Figure 4: The structure of the recognition program*

3.4 RECOGNISING SPECIFIC MOTIONS

Each specific motion, (e.g. waving, flapping arms up and down) produces and arc or arcs of motion around the joint or joints. Various kinds of motion are specifically identified and modelled in the recognition system.

*Figure 5: An arc of motion*

### *3.4.1 Waving*

Fractions of the arcs of motion are produced between any two frames of this complete motion. These correspond to regions of motion which have been found. Templates were created by dividing up the full arcs of the joints into smaller sections that can be matched.

*Figure 6: Regions of motion in the arc*

### *3.4.2 Fractions of Waving*

The size of the regions of motion created is affected by the speed of motion and time difference between frames. A smaller time difference means more accurate matching, but requires many more templates. A Larger time difference means that arcs must be split into fewer separate regions rendering less templates, but entire motions of the human subject must be more accurate and disjoint from other movements i.e. arm must move cleanly between two places within a given time to provide a large, clean region of motion.

Each region of motion is compared against all of the known templates which represent *parts* of motions for the body part which caused the motion. A top level function was designed which takes in the structure that stores information about the motions and the motion image itself. The motion image along with the information about a single motion is then passed on to a function which matches the appropriate templates for the body part which caused it. This function loops around for each appropriate template and compares it to the region of motion. The template that matched most closely is then returned as the recognised template. This results in the *type* of motion being known as well as which part of the body caused it.

Before a template is compared with region of motion, it is scaled to be the same size in terms of resolution, as the region of motion. This is because the template matching compares the two buffers on a pixel-by-pixel basis. If they did not have the same number of pixels, the results would be meaningless. The first design of the function which matches appropriate templates to regions of motion, matched all of the templates for the specified body part. Unpredictable results were produced, because regardless of the *aspect ratio* of a motion template it was scaled to the same resolution as the region of motion which was being matched.

Often a distorted template was matching better than the correct template. This was resolved by a function, which checked the compatibility of the aspect ratio. Testing yielded that 0.7 and 1.3 were the correct lower and upper limits for horizontal and vertical resolution scale factors.

### *3.4.3 Quantity of Motion*

Originally the design for calculating the amount of motion in the image took the motion image directly after the subtraction operation. All non-black pixels were counted and the proportion of the image which is motion was calculated using the count, and the total number of pixels in the image. Along with the non-black pixels that were created by motions from the person, this method also included non-black pixels that were noise created slight variations in background.

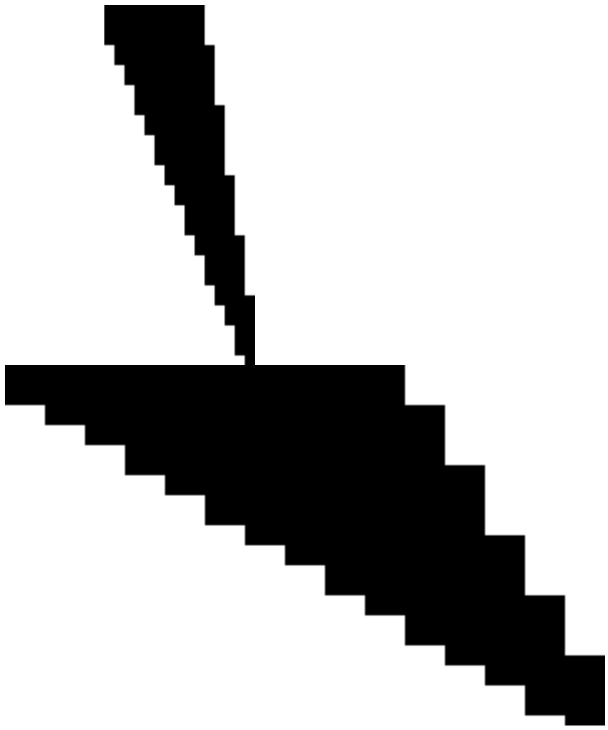

*Figure 7: A template and its distorted version*

The design was revised to process the motion image that has been processed by the region labelling function. An additional parameter was required to determine how many *colours* to count up from zero (how many regions of motion there are). All pixels that are a colour which depicts a region, are counted. Similarly to the first design, the proportion of the motion image, is calculated using the count or region pixels and the number of pixels in the image. this algorithm produces a proportion which truly represents the amount of intended motion.

*3.4.4 Distance from Camera*

No elegant design is apparent for determining a person's distance from a single camera. The design that was produced uses the premiss that the person's distance from the camera is proportional to the amount of the background that is obscured from vision. The Image Arithmetic subtraction function is used to remove the background from the input image leaving a silhouette of the person as non-black and the rest of the image as black. The number of non-black pixels are counted, and identically to the quantity of motion function, the proportion is calculated using the total number of pixels in the image. The higher the proportion of the background that is obscured, the closer the person is to the camera.

### 3.5 INTERACTION: THE ARTWORK'S RESPONSE

Having discussed how the presence and movement of people is detected, we can now consider how, in a work such as *Kyoto*, the visual artwork is reactive to what the people are doing. The interactive artwork system (figure 2) is searching through a set of rules and, as it does so, generating the sequence of images that form the output of the work. Each image represents the state of the search at that moment. In the earlier systems (Edmonds, 1988) the sequence of states was entirely determined by the search strategy used by the software to explore the rules. In the interactive case, however, the search engine has available to it a stream of data that is a coded representation of the behavior of the viewer. The system allows considerable detail to be influential. But in the case of *Kyoto*, only the nearness of the person and the overall level of movement are used. The images that represent the state of the search are then modified, in real time, according to a small set of interaction rules.

In this case, the image is a set of coloured stripes and the nearer the person is to the piece the narrower the stripes become. This creates a sensation of the work retreating as the viewer approaches it. In addition, the rate of change is, up to a point, directly proportional to the amount of movement (e.g. waiving) that is detected. However, too simple a relationship is not particularly appropriate. One point is that there is always some movement (using y=Mx+C rather than y=Mx to relate image rate of change, y, to person movement, x). In addition, when the degree of person movement reaches a particular level, the images revert to the slowest level (**if** x>Max **then** y=C). In effect, the piece does not "like" wild articulation. This notion is borrowed from Edward Ihmatovich, whose piece SAM, moved in relation to sound but stopped if things became too loud (Riechardt, 1968).

The rules for interaction in *Kyoto* are, therefore, very simple. However, simple rules do not imply simple interaction or a simple aesthetic experience. The situation turns out to be rather the reverse. The viewer of the work has no instructions and is engaged by the changing images without the interaction, which adds a layer of complexity to an already quite rich situation. Experience showed that the use of complex and elaborate interaction rules could lead to an appearance of random behavior by the image generation system. Only too easily, the experience of *interaction* can be lost all together. The simplicity of the basic forms used combined with the simplicity of the interaction rules seems to have led to a more engaging work that might have been produced using all of the information provided by the image analysis component. Future experiments, however, will test this point out much more fully than was possible in this one case.

*Figure 8: A suspended video construct at Kettle's Yard Cambridge, UK (2001)*

## 4. Conclusion

The art practice that provides the context for the paper is concerned with computer controlled work, for example as seen in figure 8. The thinking behind the author's time-based and interactive art was briefly reviewed from the early inspirations of Malevich to current logic based work employing image recognition. In this work one can think about the implications of underlying structures in ways that were not possible without computers. Generating time-based and interactive work of this kind was transformed by the computer. It enables the effective construction of inter-relationships that provide computational models for building correspondences between forms and systems. The methods used to model the behavior of people looking at a work were described and it was shown how these are used in interactive art.

The key point is that what the computer enabled me to do was to express, at a much higher level than I was used to, what I was after: so I could talk about structures. What one could do with the system was express structure and have the system generate the implications of that structure which one can then look at and think about and evaluate. What that means is that one can start to think about the implications of the structures in ways that were not possible without computers. Generating time-based and interactive work of this kind was transformed by using the computer. It was not just a matter of a speeded up process but one that was changed in kind.

It is interesting to speculate on the implications of art work of this kind for interactive computing more generally. Perhaps the most intriguing thought is that the creative engagement of the user might be enhanced by an interactive situation that is NOT as clear as we might be encouraged to provide by the textbooks in human-computer interaction. Perhaps by carefully breaking the rules of design we might help and encourage more creativity in our users? The work provides a method for developing a whole class of art works. Experimenting with them can point us to the benefits that could lead to new classes of creative interactive systems that encourage and develop the user's creativity. The *"significant trigger for creative thought"* in the artist, reported in the 1988 Heron Island meeting, might here be provided to everyone.

## Acknowledgements

With thanks to Linda Candy, who has strongly encouraged the work described in this paper. The anonymous referees also made helpful comments on an earlier draft.

## References

Bann, S.: 1972, Introduction, *Systems Arts Council of Great Britain Catalogue*. pp 4-14.

Cornock, S and Edmonds, E. A.: 1973, The Creative Process where the Artist is Amplified or Superseded by the Computer, *Leonardo* **6**, pp 11-15.

Dixon, J. 2001. *Waving to the Computer*. Mcomp. Project Report, Department of Computer Science, Loughborough University, UK.

Edmonds, E. A.: 1988, Logic and Time-Based Art Practice, *Leonardo, Electronic Art Supplemental Issue*, pp 19-20.

Edmonds, E. A.: 1989, Vers "Vidéo Constructs" *Mesures Art International 3,* Mesures, Liège.

Edmonds, E. A. 1993: Knowledge-Based Systems for Creativity *Modeling Creativity and Knowledge-Based Creative* Design, Gero & Maher (editors), Erlbaum, New Jersey, USA, pp.259-271.

Malevich, K. S.: 1968, On New Systems in Art*, Essays on Art 1925-1933 Vol. 1*, T Andersen (editor). Rapp & Whiting, London, pp 83-119.

Railing, P.: 1989, *From Science to Systems of Art*. Artists Bookworks.

Riechardt, J. (editor) 1968. *Cybernetic Serendipity: special issue of Studio International*, London.