

## THE ENTROPIC ENVELOPE

Hector Rodriguez, School of Creative Media, City University of Hong Kong. E-mail: <smhect@cityu.edu.hk>.

See <[www.mitpressjournals.org/toc/leon/XX/X](http://www.mitpressjournals.org/toc/leon/XX/X)> for supplemental files associated with this issue. <Editor will fill in the X's>

Submitted: June 10, 2016.

### Abstract

This essay introduces a computational model of the temporal experience of moving image sequences. The heart of the model, the concept of an entropic envelope, is based on information theoretic ideas. The concept is first described informally and then explained in a more mathematically precise manner. The essay finally concludes with some tentative reflections on the methodology of computation model-building in cinema studies.

### Keywords

Entropic envelope; computational film theory; cinematic time consciousness; image flows; epistemology of computer vision.

### Introduction

Computational technologies afford new methods for the precise analysis of moving images, opening up new possibilities for the intersection of science and cinema studies. There are at least two ways in which computational methods might impact film scholarship. First of all, the scholar might use computational methods that in some sense produce results analogous to concepts and capacities that the scholar already has, for instance cut detection, face recognition, motion detection, etc. A film researcher might, for instance, use a cut detection algorithm to segment a digitized film into shots or sequences. This approach might be used to count the number of cuts in a given movie or a segment of a movie and so to describe its editing style. Another scholar might employ a computer vision algorithm to determine what how many shots in a movie are close-ups. This kind of investigator relies on computational tools to realize more efficiently and rapidly certain tasks that she could in theory perform without the use of computers.

Secondly, the cinema scholar might develop computational methods that generate new critical concepts. Instead of using familiar notions like “cut”, “close-up”, “fade to black”, or “tracking shot”, the researcher might discover, through the process of designing or refining customized algorithms, fresh concepts that were not previously part of the standard critical repertoire. In this second approach, the scholar’s encounter with algorithmic technologies has the potential to expand and reconfigure the conceptual stock of cinema analysis [1]. These new concepts are essentially bound up with the use of computer technologies.

One of the main strategies of concept formation in cinema studies is the construction of mathematical models. The elaboration of a mathematical model can serve the function of constructing a concept. I shall speak here of *concept formation through model building*. The content of the concept is the model. From this point of view, the introduction of a new mathematical model in the humanities can reshape the way we think about a certain domain of critical analysis [2].

As an instance of the formation of a concept through the construction of a mathematical model, this essay introduces the concept of an *entropic envelope*. The concept is meant to capture certain properties of cinematic sequences that can be formalized through the application of information theoretic techniques. This concept encapsulates a mathematical model

intended to reconfigure how we think about the temporal properties of image sequences. In particular, I want to argue that models are technologies that, among other uses, serve to guide or channel our attention. By creating a new concept through the construction of a model, we learn to pay attention to certain properties of moving image sequences that we might otherwise have overlooked.

My discussion is divided into three parts. The first introduces the entropic envelope notion without giving any mathematical background. The second part, which is intended for the more technically inclined reader, explains the algorithm in greater detail. The third and final section, more self-reflexive and methodological in nature, considers the process of concept formation through model building as a complex task influenced by philosophical, artistic, and technical factors. My aim here is to encourage cinema researchers to investigate the nature of model building from this multidimensional perspective.

### 1. The Entropic Envelope

I propose to characterize the temporal flow of a video sequence as a pattern of equilibria and disequilibria. Equilibrium occurs when a frame in some sense resembles immediately preceding frames. Disequilibrium occurs when a frame contains substantially new information relative to immediately preceding frames in a temporal image sequence. The model I am presenting measures the informational novelty of frames with reference to their shape or contour properties, the visual arrangement of edges and corners. Corners and edges are pixels with reasonably high brightness gradients in one or both spatial coordinate directions. The detection of edges and corners therefore depends only on the distribution of brightness values throughout an image. Figure 1 illustrates this idea by displaying a movie frame alongside its contour representation.

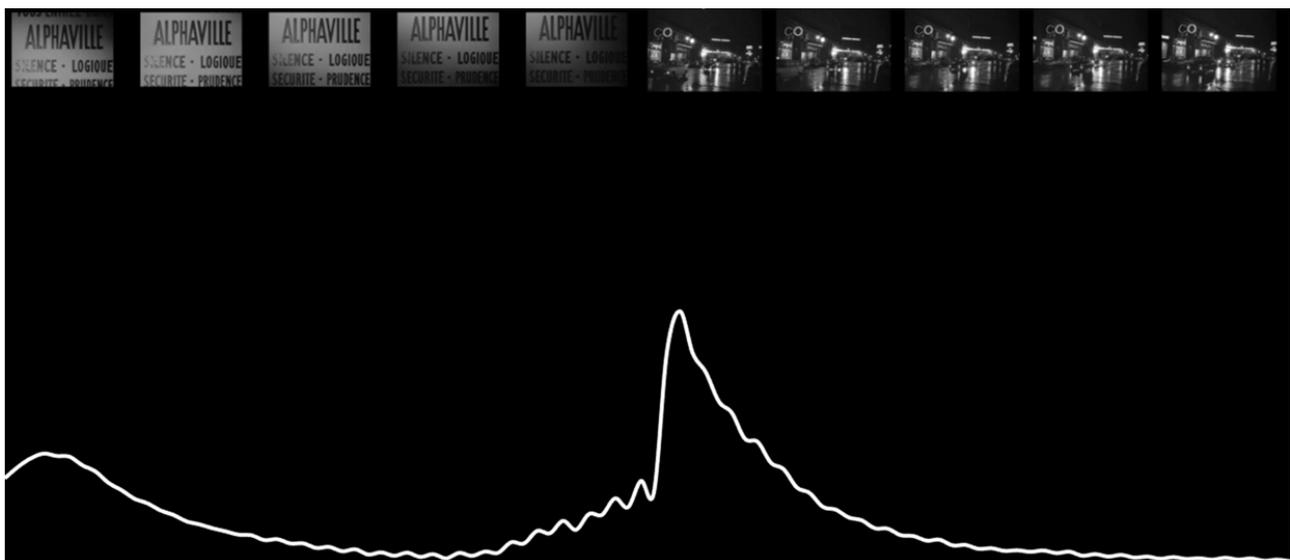


Fig.1. Contour structure of a single frame in *Alphaville* (Jean-Luc Godard, 1965).

My model brackets other aspects of the image sequence, such as for instance color, dialogue, object and event recognition, camera movement, etc. This essay therefore suspends any questions of narrative content and treats a video sequence as a stream of graphical structures.

Figure 2 shows the entropic curve for a sequence from the film *Alphaville* (Jean-Luc Godard, 1965). Notice that the curve has a relatively high value at a moment when there is a cut, since cuts can be viewed as moments where the contour structure of the new image differs sharply from that of immediately

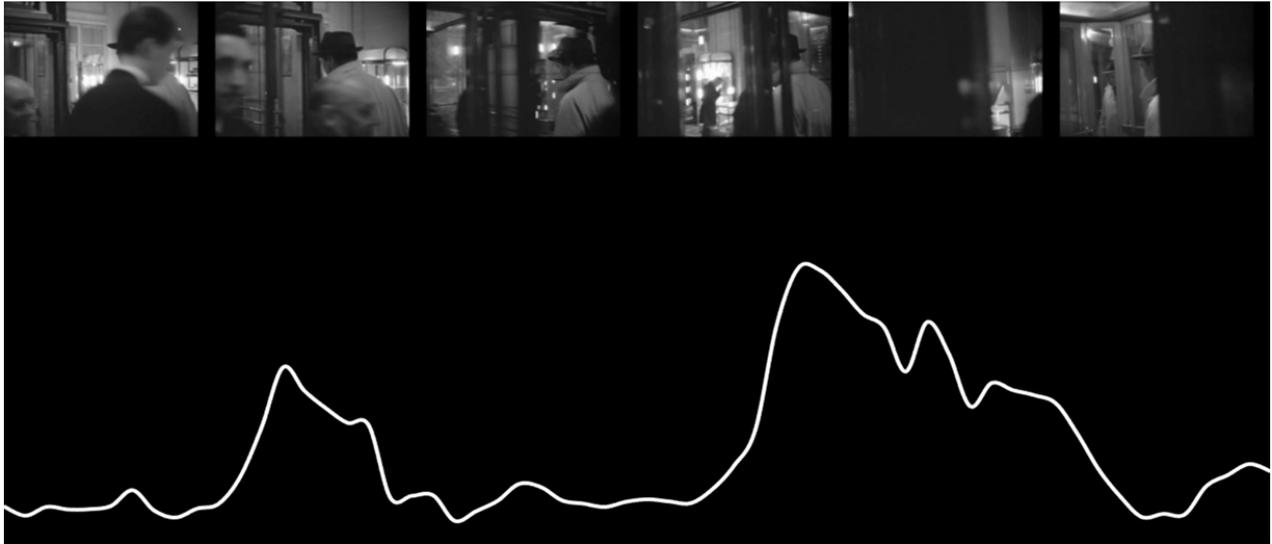
Fig. 2. Entropic envelope of a sequence from *Alphaville*.



I propose to employ a concept from information theory, the concept of relative entropy, to measure the degree of disequilibrium induced by a new frame against the background of immediately preceding frames. To put it very roughly, the degree of equilibrium or disequilibrium of a frame is the quantity of information contained in its contour representation relative to immediately preceding frames. The precise computation of this measure will be explained in the more formal part (section 2) of this paper. The temporal organization of a sequence can be graphed as a curve where the abscissa represents the time, measured in frames, and the ordinate represents the disequilibrium of the corresponding frame. The entropic envelope of a video sequence between frames A and B is the portion of the relative entropy curve between those frames.

preceding images. The entropic curve is, in general, a reasonable cut detection algorithm. Cut detection is not, however, the main aim of this project. The curve represents what could be called the internal melodic line of a video sequence. Even within a continuous shot, there are dynamic patterns of equilibria and disequilibria. The entropic envelope captures those patterns.

Fig. 3. Entropic envelope of another sequence from *Alphaville*.



Consider for instance a sequence from *Alphaville* as shown in Figure 3. The curve rises whenever certain elements enter the frame with disrupt the stability of a composition. For instance, the arrival of detective Lemmy Caution in the hotel in *Alphaville* involves a moving camera shot that follows the main character in a fairly stable composition, but this stability is punctuated by vertical architectural elements that introduce carefully patterned moments of disequilibrium.

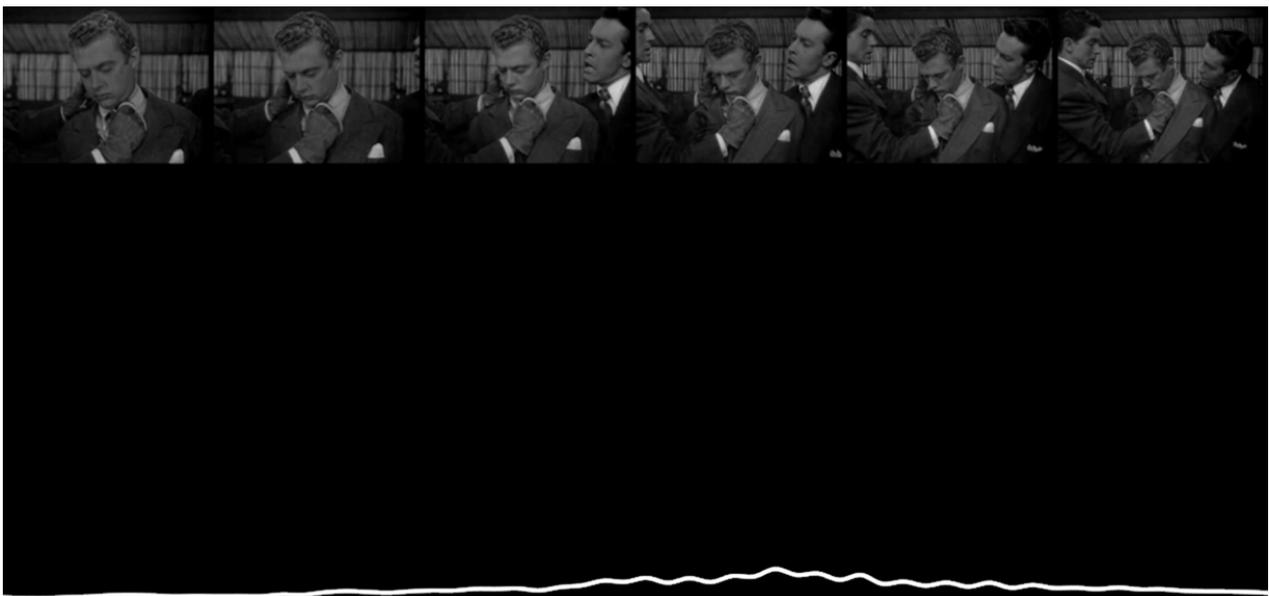
The sequence exemplifies how the entropic envelope captures the temporal rhythm of a video sequence. It identifies and represents rising and falling patterns of equilibria and disequilibria. This model can be used, for instance, to compare and contrast the temporal organization of image sequences from different countries, periods, artistic movements, etc.

Compare the entropic envelope for the above scenes from *Alphaville* with this one from Hitchcock's *Rope* as shown in Figure 4. It is clear that the entropic envelope in Hitchcock's

film is far more stable than that of Godard's work. Godard's mise-en-scene, at least during the hotel arrival sequence, generates moments of disequilibrium that punctuate the underlying equilibrium of the composition. These two examples sharply differ from the case of Stan Brakhage's 1963 experimental film *Mothlight* (Figure 5). That film was made without a camera, by collecting and collaging insect wings, blades of grass, and other bits of organic matter. The result is visually far more radical than either *Alphaville* or *Rope*, as is evident from the shape of the entropic envelope from this sequence.

We can think of the entropic curve, very loosely, as a representation of the experience of an ideal spectator who pays attention only to the pattern of changing contour structures in a video sequence, and who anticipates that every coming frame will resemble recent frames. This ideal spectator makes an assumption of continuity, and is ready to register moments of disequilibrium, in which the new frame is informationally rich

Fig. 4. Entropic envelope of a sequence from *Rope* (Alfred Hitchcock, 1948).



relative to its predecessors in the sequence. The ideal spectator is, in other words endowed with what Ernst Gombrich calls a “break-spotter”, i.e., the ability to detect “points of maximal information content” in a temporal sequence [3].

The successive values of the envelope, roughly speaking, measure the divergence between the ideal viewer’s expectations, formed out of her experience of previous frames, and the incoming frame.

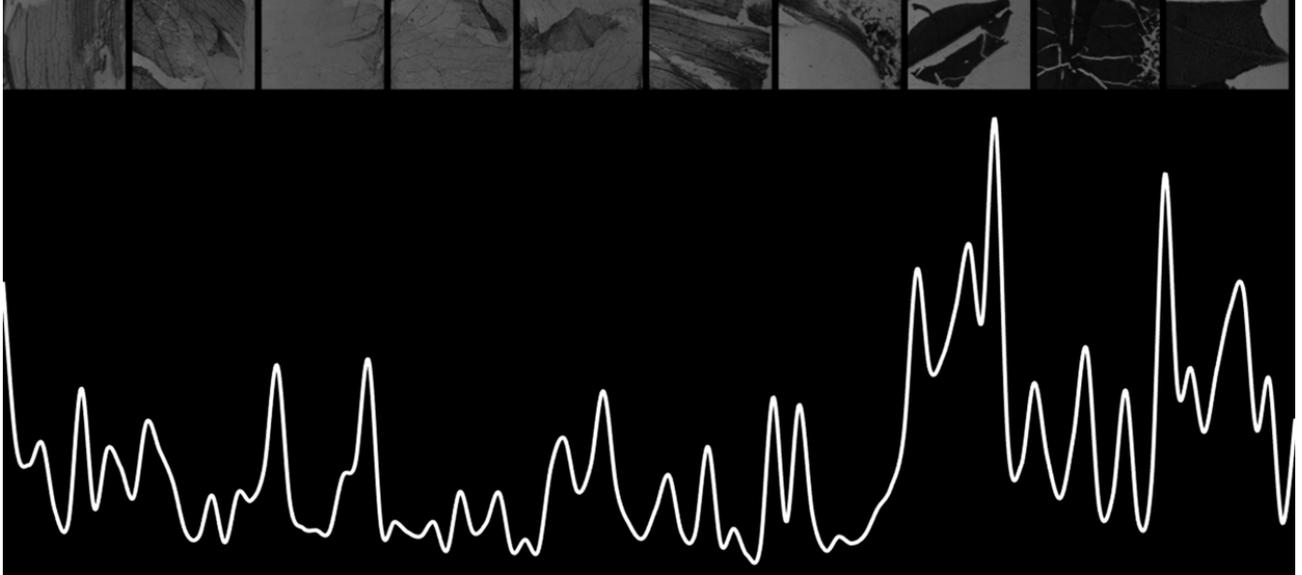


Fig. 5. Entropic envelope of another sequence from *Mothlight*. (Stan Brakhage, 1963).

I claim that this model can deepen our apprehension of the visual rhythm of image sequences. My main aim is to enhance our perception and appreciation of certain temporal properties of image-sequences. I would suggest that most works of cinematic art, even those that possess a rich narrative content, depend on having some form of temporal organization at the graphic level, which can be analyzed in its own terms.

The following video clips illustrate the entropic envelope of the movies described here:

*Alphaville*: <https://vimeo.com/169828109>

*Rope*: <https://vimeo.com/170119475>

*Mothlight*: <https://vimeo.com/170119958>

## 2. Technical Description

Non-technical readers may choose to skip this section, which describes the mathematical content of the model in more precise detail. I have designed this model so that the mathematical level required for its comprehension and implementation is relatively simple. My aim is to encourage humanists and artists who might wish to write their own source code to test the ideas presented here, and to develop more fine-grained models of cinematic style.

My approach connects with an ongoing research tradition that brings information theoretic concepts to bear on video sequences [4]. Researchers have shown that information theoretic concepts like the Shannon Entropy, Mutual Information, or the Jensen Divergence provide reasonable metrics for image similarity or difference, and so can be fruitfully used to summarize video sequences and to detect cuts and fades [5]. Encouraged by these results, I propose to use a well-known information theoretic metric, the so-called Relative Entropy or

Kullback-Leibler Divergence, to formalize a model of cinematic spectatorship.

The model involves the following basic notions.

A movie is a temporal sequence of images. Every image is a brightness function of two arguments. I shall express the image function at time  $t$  as  $I_t(x, y)$ , where the values of  $x$  and  $y$  represent the horizontal and vertical location of a pixel. Every pixel consists of a single grayscale value, represented as a real

number in the interval  $[0,255]$ . The model can be augmented in the future by adding color information.

The model I am proposing emphasizes the shape properties of images. The notion of “shape” can be represented in terms of image gradients. The gradient vector of image  $I_t(x, y)$  is

$$G_t(x, y) = \left[ \frac{\partial I_t(x, y)}{\partial x}, \frac{\partial I_t(x, y)}{\partial y} \right], \quad (1)$$

with magnitude

$$M_t(x, y) = \sqrt{\left[ \frac{\partial I_t(x, y)}{\partial x} \right]^2 + \left[ \frac{\partial I_t(x, y)}{\partial y} \right]^2}, \quad (2)$$

and angle

$$A_t(x, y) = \tan^{-1} \left[ \frac{\partial I_t(x, y) / \partial x}{\partial I_t(x, y) / \partial y} \right]. \quad (3)$$

The image is represented in digital form. The data is discretized in two ways. First of all, the brightness of every pixel is approximated by a finite-precision (floating-point) number. Secondly, the spatial coordinates of every pixel are also approximated by floating-point numbers. A discrete approximation to the gradients is computed using the well-established Sobel operator. This project then uses gradient vectors to detect edges and corners. The algorithm, which is due to Harris, is not described here [6]. I will use the expression *contour pixel* to denote any pixel that is either an edge or a corner. Contour pixels for a frame from *Alphaville* are shown in black in Figure 1.

An image can be divided into  $M$  non-overlapping rectangular sub-regions. For each sub-region, my procedure forms an

angular histogram, containing  $K + 1$  bins, by counting the number of contour pixels whose gradient angle falls within one of  $K$  intervals. The  $(K + 1)^{\text{th}}$  bin contains the total number of non-contour pixels in that sub-region. This last bin tends to have a very high value relative to the other bins. To compensate for this higher value, the weight of the other bins is doubled. The histograms for the  $M$  sub-regions of each frame are then combined into one feature vector  $H_t$  of  $M \times (K + 1)$  dimensions. This vector supplies a statistical characterization of what might be called the shape structure of the image. I shall refer to it as the *shape descriptor* or *contour structure* of image  $t$ .

In this project, I set  $M = 4$  and  $K = 8$ . In other words, the image is divided into four rectangles, and every angular gradient can fall into one of eight possible intervals.

The feature vector  $H_t$  can be normalized so that all entries add up to 1, thus specifying a discrete probability distribution for image  $I_t(x, y)$ . Every image can in this context be understood as a discrete random variable with  $M \times (K + 1)$  possible values.

I now use these basic ideas to construct an abstract model of cinematic temporality.

The temporality of the cinema is here modeled by assuming that the ideal spectator experiences a movie  $F$  one frame at a time. The current image  $I_T(x, y)$  divides all other frames of  $F$  into those belonging to the past, i.e., the frames  $I_t(x, y)$  where  $t < T$ , and those belonging to the future, where  $t > T$ .

My model assumes that the ideal spectator does not simply apprehend every frame as disconnected from preceding frames. The shape descriptors associated with recent frames are not simply forgotten. Rather, the spectator viewing frame  $T$  retains the shape information from previous frames. Recent frames influence the spectator's expectation concerning the new frame. The influence of each immediately past frame on the viewer's expectation decays as a function of its distance from the current frame. Let the retention at current frame  $T$  be

$$R_T = \sum_{i=1}^{T-1} w_i H_i \quad (4)$$

where  $w_i \in [0,1]$  is the weight of frame  $i$  as a function of its distance from the current frame  $T$  and  $H_i$  is the histogram of frame  $i$ . I have set the value of  $w_i = 0.95^{T-i}$ . This setting implies that the impact of past frames gradually fades over time.

I model the temporal disequilibrium induced by frame  $T$  as the divergence of the normalized vector  $H_T$  from the retention  $R_T$ . I quantify this divergence by using the information-theoretic notion of relative entropy [7].

The Relative Entropy, or Kullback-Leibler Divergence, of two discrete distributions  $p$  and  $q$  is defined as follows:

$$E(p||q) = \sum p(x_i) \log_2[p(x_i)/q(x_i)] \quad (5)$$

If we consider  $q$  as a hypothesized distribution and  $p$  as the actual distribution, then the relative entropy of  $p$  with respect to  $q$  measures the error induced by the assumption that reality has distribution  $q$  when its actual distribution is in fact  $p$ . More precisely, the relative entropy  $E(p||q)$  measures the cost, i.e., the number of surplus bits, incurred by any coding scheme which assumes that a given variable has probability distribution  $q$  when its actual distribution is in fact  $p$ .

The disequilibrium associated with frame  $T$  is given by

$$S_T = E(R_T||H_T) \quad (6.1)$$

where  $R_T$  and  $H_T$  are normalized (i.e., they represent probability distributions).

We can think of  $S_T$  as a measure of the informational novelty contained in the  $T$ th frame. If it is very high, the present image is highly informative relative to the frames in the recent past. In other words, (6.1) measures the degree of disequilibrium that an ideal spectator experiences when perceiving the structure of the current frame in relation to the expectation formed on the basis of the recent frames of the movie.

It is possible to avoid the assumption that the ideal viewer perceives individual frames by replacing equation (6.1) with this alternative measure:

$$S_T = E(R_T||R_{T+1}) \quad (6.2)$$

This measure compares the retention of the current frame with the retention of the coming frame. Since retention is always taken over a segment of frames, (6.2) does not use isolated images directly. Other modifications of (6.1) are possible, in line with the phenomenological notion, expressed by Husserl, that the present is not a point-like instant but an extended segment of time.

The changing informational content of successive frames in a video sequence can be graphed as a curve in  $\mathbb{R}^2$  where every abscissa represents frame  $T$  and the corresponding ordinate represents the disequilibrium  $S_T$ . Rising and falling values of the graph denote the disequilibrium produced by unexpected frames. The entropic envelope of a film from frame  $A$  to frame  $B$ , where  $A < B$ , comprises the sequence of values  $S_T$  where  $T$  runs from  $A$  to  $B$ .

Note that the relative entropy is not symmetric. The relative entropy of  $P$  with respect to  $Q$  is not in general the same as that of  $Q$  with respect to  $P$ . If we reverse the direction of time for a given sequence, the shock associated with each frame need not remain invariant. The entropic envelope need not even have the same shape, only reversed. Rather, it will be a qualitatively different curve, with a different pattern of peaks and valleys. The relative entropy is thus sensitive to the temporal direction of a moving image sequence. Because the relative entropy is asymmetric, it is not a genuine distance metric in the mathematical sense [8]. This asymmetric property is important, for conceptual reasons. It would be inappropriate to use a true mathematical distance to represent temporal processes, which are *asymmetrically* directed from the past towards the future. The entropic envelope concept is designed to take into account the arrow of time.

### 3. Model Building

A concept is not, so to speak, found readymade in nature. It must be constructed. I have already noted that one of the main functions of mathematical models is the construction of concepts. I have constructed the concept of an entropic envelope by specifying a mathematical model of cinematic temporality. The construction of this model was a fairly complex affair that brought together philosophical, aesthetic, and technical factors. A major goal of this paper is to encourage researchers in the humanities to pay attention to and reflect on the possible role of mathematical models in cinema studies from these various perspectives.

The design of a mathematical model in the humanities is often related to the researcher's theoretical or philosophical commitments. The model presented here was influenced by a philosophical tradition inaugurated by phenomenologist Edmund Husserl [9]. This tradition, which has been recently developed by Bernard Stiegler, regards a film as an essentially temporal object. A temporal object "constitutes itself temporally" as "that which appears in passing, as that which passes, as that which manifests itself in disappearing" [10]. I take it as a given that there exist temporal objects like melodies or videos. I also take it as a given that we apprehend (perceive, grasp) those objects. The question arises as to what is essentially involved in our apprehension of a temporal object. To apprehend a temporal object demands a threefold experience of attention, retention, and anticipation. The viewer's attention to the present moment is intertwined with her anticipation of the immediately upcoming future and her retention of the immediate past. This threefold structure of conscious experience affords the viewer's apprehension of temporal objects like melodies or films. Husserl proposed that the consciousness of the present sound in a musical sequence, for instance, is entwined with the retention not only of the immediately preceding note but also of previous notes as well as the anticipation of the future notes. The influence of any specific tone on the present fades gradually over time, like the tail of a comet.

The concept of an entropic envelope was constructed with reference to Husserl's a priori characterization of time consciousness. Equation (4), in particular, formalizes the impact of past frames on the present experience as gradually fading over time. The model I am proposing is not, however, simply a mechanical implementation of the Husserl's thesis concerning the threefold experience of time. Rather, the construction of the model also transforms the Husserlian conception. No mathematical model implements a phenomenological thesis in a mechanical way. The design of the model involves an ineliminably creative element. The philosophical framework does not determine a unique mathematical formalization. For instance, nothing in Husserl's doctrine demands the use of information theoretic measures like the relative entropy. The philosophical conception guides, but does not completely determine, the construction of the model. Its designer must make certain choices, and those choices will reconfigure the concept of a temporal object. In addition, note that the model captures only the general idea of Husserl's thesis rather than any of its specific details.

The model conceptualizes the notion of a temporal object as an organized pattern of equilibria and disequilibria. This conceptualization depends essentially on computational methods. It would not be possible to describe video sequences in this way without the use of information technologies.

In particular, the formal model is directly applicable only to video sequences available to us in a computationally tractable form. The films mentioned here were originally produced in analog form, and the computation of the entropic envelope requires digitized versions of them. Digitization presupposes that the films in the corpus under investigation have measurable attributes, i.e., attributes with quantitative structure [11]. In the case of a movie, the measurable attributes are the brightness levels at specific locations of every two-dimensional frame. The digital version of a single frame consists of measurements of these light intensities, represented by an array of numbers. The digital frame is in this sense already a formal object. The formalization of every frame in a movie is a condition for the application of my model. A movie is a succession

of two-dimensional matrices that record the outcomes of brightness measurements. My model operates not on any brute or unmediated reality but on data that has already been conceptualized in a way that affords computational processing.

Numerical pixel data has to be processed even further to be usable in the model. A digital frame is an array of localized brightness values, but this structure is not stable enough from frame to frame to sustain the detection of temporal patterns. A character might, for instance, move her head only slightly, yet this subtle motion would result in very different brightness values for each individual pixel across two or more successive frames. But from the point of view of the temporal pattern of the film, this change must count as a relatively subtle one. We must associate each frame with some structure that captures the global contour structure in a way that is persistent across small changes. For this purpose, I construct a histogram of gradient orientations out of the pixel data, as described in the technical section of this paper. This histogram is a model of the data. In other words, I work directly with a model of the data, not with the actual pixel data [12]. Any mathematical model in the arts will often be constructed not directly on the data that has actually been measured but on a highly abstract model of that data, or even on several layers of increasingly abstract models.

The data model constrains what the entropic curve model can represent. More specifically, the data for each frame is modeled as a contour structure, and the successive values of the entropic curve represent the relative information contained in the contour structure of frames within the unfolding sequence. I use a model of the data for each frame to construct a model of the temporal sequence.

In addition to this representational function, the entropic envelope model also serves what may be described as a perceptual or revelatory function. By giving us a means for representing certain properties of image sequences, this model also gives us a means for noticing them. In this somewhat loose sense, a mathematical model resembles a microscope or a telescope, i.e., a technology that is both representational and perceptual. The entropic envelope guides our perceptual attention to certain features of the moving image sequences. Researchers can use it to learn how to pay attention to certain aspects of the visual rhythm of cinematic sequences that they might not otherwise detect.

In this context, it is essential to bear in mind that every mathematical model embodies a norm for its use. To be a competent user of this model is to know that the entropic envelope does not represent any information regarding the recognition of objects or events. To be able to read the entropic curve of a cinematic sequence is in part to know what the curve can represent or denote. The curve is a graphical object, but it only conveys information to one who knows how it is intended to be understood. The user of a perceptual model is expected to understand and pay attention to those properties of the moving image sequence that the model can represent.

The competent user of this model is capable of looking at movies *according to* the model. The viewer comes to understand what the model represents and then pays attention to precisely those properties of the image sequence represented by the model. The model represents the relative contour information of a moving image sequence. The viewer who understands this point is able to perceive image sequences in accordance with the model, which makes the revelatory function possible.

The entropic envelope can be viewed as a model of the film being analyzed and as a model of the user of the model. It is a model of the films, in that it represents and reveals certain properties of those films. It is a model of the viewer, in that it supplies a norm for viewing. The model mediates between the viewer and the moving images.

The point of this model is not to reproduce or represent what “average” or “normal” viewers actually experience when viewing a film. Rather, one of the main uses of this model is to *reshape* how we look at the film, so that we might come to detect aspects of the temporal organization of images for which we currently lack critical concepts.

This model is oriented towards the formal analysis of cinematic art through the application of computational methods. A crucial assumption is that the temporal patterns detected by the entropic envelope capture aesthetically relevant properties of image sequences. These properties pertain to the level of analysis that neoformalist film theory refers to as the “surface texture” or “film style” level [13]. I would suggest that moviemakers organize the stream of images as patterns of equilibria and disequilibria that comprise the distinctive rhythm of the film. In this sense, the research method applied here accords with certain aspects of the creative process whose outcomes are being researched. Without the possibility of the temporal patterning of graphical forms dynamically arranged on a two-dimensional image surface, cinematic art would not be what it is.

A final observation. An important advantage of the application of mathematics is that formal techniques enables further processing by the application of other mathematical methods. For instance, segments of the entropic curve can be analyzed in the frequency domain using the Discrete Fourier Transform method. The entropic curve of a film is a formal object, which can in turn be further processed by means of more sophisticated formal methods. Successive frequency domain representations of the entropic envelope for two of the films mentioned in this essay can be viewed here:

*Alphaville*: <https://vimeo.com/183599541>

*Mothlight*: <https://vimeo.com/183600460>

## Conclusion

This essay has presented the concept of an entropic envelope, which encapsulates a mathematical model of the temporal organization of video sequences. My aim in presenting this model is to highlight the importance, for cinema scholars and other humanists, of undertaking research on the nature of model-building as a multidimensional process that brings together philosophical, aesthetic, and technological factors. Future work on this model will involve its extension to the analysis of sound and narrative content as well as its application in domains like games and performances that involve some degree of temporal patterning [14].

## References and Notes

1. I believe that this tactic works best when the cinema scholar is herself involved in the production of the software in question, and so has come to understand at least some of the mathematical concepts on which the algorithm depends. There are many scholars and artists working on the computational analysis of existing films. A relatively recent example is Benjamin Grosser’s *Computers Watching Movies*: <http://bengrosser.com/projects/computers-watching-movies/>. Another example is the Cinemetrics project of Frederic

Brodbeck: <http://cinemetrics.fredericbrodbeck.de/>. An earlier work of mine that also depends on computational analysis is the Gestus project: <http://concept-script.com/gestus/index.html>.

2. My understanding of models and concept formation is inspired by research in the philosophy of mathematical models conducted by Mary S. Morgan and Margaret Morrison. See for instance: M. S. Morgan, “Imagination and Imaging in Model Building,” *Philosophy of Science* 71, No. 5 (2004), pp. 753-766, Morgan, *The World in the Model: How Economists Work and Think* (Cambridge: Cambridge University Press, 2012); M. Morrison, “Scientific Understanding and Mathematical Abstraction,” *Philosophia* 34, No. 3 (2006), pp. 337-353. Morrison, *Reconstructing Reality: Models, Mathematics, and Simulations* (Oxford: Oxford University Press, 2015). See also the essays contained in the extremely helpful anthology: Morrison and Morgan, eds., *Models as Mediators* (Cambridge: Cambridge University Press, 1999). The methodological remarks on section 3 of this paper are strongly influenced by the “models as mediators” conception proposed by Morgan and Morrison.

3. E. H. Gombrich, *The Sense of Order: A Study in the Psychology of Decorative Art* (Oxford: Phaidon, 1979), p. 121. F. Cucker, *Manifold Mirrors: The Crossing Paths of Art and Mathematics* (Cambridge: Cambridge University Press, 2013), pp. 120-124.

4. F. Escolano, P. Suau and B. Bonev, *Information Theory in Computer Vision and Pattern Recognition* (London: Springer, 2009).

5. Qing Xu, Yu Liu, Xiu Li, Zhen Yang, Jie Wang, Mateu Sibert, and Riccardo Scopigno, “Browsing and exploration of video sequences: A new scheme for key frame extraction and 3D visualization using entropy based Jensen divergence,” *Information Sciences* Volume 278, (2014) pp. 736-756; Z. Černeková, I. Pitas and C. Nikou, “Information theory-based shot cut/fade detection and video summarization,” *IEEE Trans. Circ. Syst. Video Technol.* 16, (2006) pp. 82–91; M. Mentzelopoulos and A. Psarrou, “Key-frame extraction algorithm using entropy difference,” *Proc. ACM SIGMM Int. Workshop on Multimedia Information Retrieval, New York, NY, USA, (2004)* pp. 39–45.

6. C. Harris and M. Stephens, “A combined corner and edge detector,” *Proceedings of the 4th Alvey Vision Conference* (1988) pp. 147–151, <<http://www.bmva.org/bmvc/1988/avc-88-023.pdf>>.

7. S. Kullback and R.A. Leibler, (1951). “On information and sufficiency,” *Annals of Mathematical Statistics* 22, No. 1 (1951) pp. 79–86.

8. Moreover, it does not satisfy the triangle inequality.

9. E. Husserl, *The Phenomenology of Internal Time-Consciousness*, J. Churchill, trans. (The Hague: Marinus Nijhoff, 1964).

10. B. Stiegler, *Technics and Time, 3: Cinematic Time and the Question of Malaise*, R. Beardsworth and G. Collins, trans. (Stanford: Stanford University Press, 2011) p. 36. A clear discussion of Stiegler’s appropriation of Husserl can be found in: B. Roberts, “Cinema as Mnemotechnics: Bernard Stiegler and the Industrialization of Memory,” *Angelaki* 11, No. 1 (2006) pp. 55-63.

11. For a related discussion of questions about the philosophy of measurement, see J. Michell, “History and Philosophy of Measurement: A Realist View”, in *Proceedings of the 10th IMEKO TC7 International symposium on advances of measurement science*, June 30-July 2, 2004. Like Michell, I endorse a realist view of measurement according to which measurements are discovered in the things themselves rather than assigned to them.

12. Anyone who reads the technical section of this paper will appreciate that my approach is even more complicated than described here, in the sense that it consists of layers of models of the data. The notion of a model of data has been stressed by Patrick Suppes. See in particular: “Models of Data”, *Studies in Logic and the Foundations of Mathematics* vol. 44 1966, pp. 252-61.

13. D. Bordwell, *Poetics of Cinema* (London: Routledge, 2007), chapter 3.

**14.** The author gratefully acknowledges the contribution of his research assistant, Hugo Yeung. He is also extremely grateful to Felipe Cucker for his guidance and encouragement during the writing of this essay. Comments from Darko Strajn, Barbara Jpaelj Pavesic, Bryan Chung, Richard Allen, Hui Yuk, and Daniel Howe spurred important changes to the manuscript.