Every thing thinks: sub-representative differences in digital video codecs

Adrian Mackenzie

Institute for Cultural Research

Lancaster University, LA1 4YD, UK

a.mackenzie@lancaster.ac.uk

> Every body, every thing, thinks and is a thought to the extent that, reduced to its intensive reasons, it expresses an Idea the actualisation of which it determines (Deleuze, 2001, 254).

What would it mean for anthropologies of technology if we took seriously Deleuze's claim that every thing thinks? This claim is hard to stomach for several reasons. Technologies are generally seen as the expression of much highly organised thinking (scientific, design, engineering, artistic, financial, political, etc). Moreover, many technologies today are very much designed to think in specific albeit limited ways. This is particularly the case in so-called 'intelligent systems,' but it holds for any designed or made thing. Smartness, intelligence, sophistication, cleverness: are not all of these prized qualities in technology actually expressions of thought and of much mental effort? What does the affirmation '*every* thing thinks and is a thought' add? Somewhat counter-intuitively, or at least, contrary to commonsense, Deleuze's more specific claim that things determine the actualisation of Ideas, I would argue, points in a different direction towards a much more problematic mode of existence of things. Intelligent or smart technologies may in their ingenuity, render this problematic mode less visible. They may detract from the Problem-Ideas, from the singular problem-setting imperatives, actually at work in technologies.

From Deleuze's perspective, thinking would not merely refer to things through the mental work of developing concepts that represent them. It would actually be part of things. In other words, Deleuze can be read as bringing a radically constructivism to the fore. Some years after Gilles Deleuze published *Difference and Repetition* in 1968, the US Patent Office awarded the early patents on compressed digital image transmission. Two lineages of patents began to emerge. Both sought to isolate and intensify certain forms of repetition in moving images. In the first, the patents described some ideas for the application of signal processing techniques known as fast Fourier transforms (FFT) to video images (Means, 1974, Speiser, 1975). These transforms extracted more compact encodings of video images and sounds so that information networks and transmission systems could store and move them more easily. In the second, the patents described how to computationally predict future video images from past images (Haskell and Limb, 1972, Haskell and Puri, 1990). The 'motion estimation' or 'motion compensation' techniques analysed patterns of movement of objects and figures across successive video or television frames. The transformations of images through the fast Fourier transform (later by the Discrete Cosine Transform (DCT) and motion estimation still stand at the very centre of compression techniques used in digital media technologies (used in JPEG images, MPEG, DVDs, etc.), and in digital video *codecs* such as MPEG-2 in particular. The entwining of these two lineages turns out to be important in very many domains of contemporary culture.

Just as Deleuze's *Difference and Repetition* concerned the rumble of *sub-representative* differences in philosophical thought, these patents lineages expressed the envelopment of sub-representative differences in technologically mediated perception. Just as Deleuze's book advanced an account of how identities and sameness stem from 'a more profound game of difference and repetition' (Deleuze, 2001, xvix), these patents announced an alteration in the micro-perceptual and infrastructural supports of representation and repetition of images and sound. The image or the frame, which in media such as photography, film and even television could still be regarded as a the support of representation, began to dissolve or at least re-distribute itself into thresholds of brightness, colour and motion vectors, into what in Deleuze's terms, could be understood as differentials. Today, the tremendous architectural complexity of a single technological instance, the MPEG-2 video codec, might serve as an index to certain spatio-temporal dynamics (problems or desires, depending on your viewpoint) in media technologies.

What is at stake in developing a sub-representative account of media technologies, or in developing radical constructivist accounts of technology more generally? The 'sub-representative,' in Deleuze's thought, refers to that aspect of things that cannot be consciously thought, or reduced to the presence of an object to a subject mediated by a concept or category. While the sub-representative cannot be identified, measured, or calculated as such, it is felt, and in some cases, felt intensely. So while the proliferation of digital video might on the one hand be seen as a paroxysm of representation, an unbounded expansion of the power to represent, from the sub-representative standpoint, it could also be seen as 'imbued with a presentiment of groundlessness' (Deleuze, 2001, 276), alterity and differences. In making sense of this claim, the discussion that follows makes a big jump from Deleuze's account of Ideas, difference and repetition to a specific, but hardly very well-known technology. One can easily speak of technology in general without acknowledging the specificity of a technology. Deleuze's claim that 'every thing thinks,' however, might bring to an anthropology of technology a re-freshed conception of specificity, and the grounds of specificity.

Stated rather baldly, specificity and actual differentiations, as seen in sensibilities, social institutions, global organisations, local work-arounds, civic epistemologies, political economies and forms of personhood associated with audiovisual media, would all represent affirmations of the problem-setting imperatives that move through codecs. As Constantin Boundas writes, in his commentary on *Difference and Repetition,*

> Just like Kant, Deleuze believes that Ideas are problem-setting imperatives. But unlike Kant, Deleuze believes that the ability of a problem to be solved must be made to depend on the form that the problem takes. (Boundas, 1996, 88)

The growth of video material culture can seen then as an affirmation of a problem-Idea. In the labyrinthine plenitude of digital video, one paths leads to a key technical component: *codecs.* Software and hardware codecs transform images and sound. Transformed images move through communication networks much more quickly than uncompressed audiovisual materials. Without codecs, an hour of raw video footage would need 165 CD-ROMs or take roughly 24 hours to move across a standard computer network (10Mbit/sec Ethernet). Instead of 165 CDs, we take a single DVD on which a film has been encoded by a codec. We play it on a DVD player that also has a codec, usually implemented in hardware. Instead of 32Mbyte/sec, between 1-10 MByte/sec streams from the DVD into the player and then onto the television screen.

The economic and technical value of codecs can hardly be overstated. DVD, the transmission formats for satellite and cable digital television (DVB and ATSC), HDTV as well as many internet streaming formats such as RealMedia and Windows Media, third generation mobile phones and voice-over-ip (VoIP) all depend on video and audio codecs. They form a primary technical component of contemporary audiovisual culture in many of its most global dimensions. Physically, codecs take many forms, in software and hardware. Today, codecs nestle in set-top boxes, mobile phones, camcorders, video cameras and webcams, personal computers, media players and other gizmos. Codecs perform encoding and decoding on a digital data stream or signal, mainly in the interest of finding what is different in a signal and what is mere repetition. They scale, reorder, decompose and reconstitute perceptible images and sounds. They only move the differences that matter through information networks and electronic media. This performance of difference and repetition of video comes at a cost. Enormous complication must be compressed in the codec itself.

Much is at stake in this infrastructure and image logistics from the perspective of cultural studies of technology and media. On the one hand, codecs analyse, compress and transmit images that fascinate, bore, fixate, horrify and entertain billions of spectators. Most of these images are repetitive or clichéd. They reinforce or shore existing orderings of difference, identity and power. There are many re-runs of old television series or Hollywood classics. Youtube.com, a video upload site, currently offers 13,500 wedding videos. Yet the spatio-temporal dynamism of these images matters deeply. They open new patterns of circulation and permit new symbolic differences to appear. To understand that circulation matters deeply, we could think of something we don't want to see, for instance, the many executions of hostages (Daniel Perl, Nick Berg, and others) in Jihadist videos since 2002. Islamist and 'shock-site' web servers streamed these videos across the internet using the low-bitrate Windows Media Video codec, a proprietary variant of the industry-standard MPEG-4. The shock of such events – the sight of a beheading, the sight of a journalist pleading for her life – depends on circulation through online and broadcast media. A video beheading lies at the outer limit of the visual pleasures and excitations attached to video cultures. Would a beheading, a corporeal event that takes video material culture to its limits, occur without codecs and networked media? We need to understand how media infrastructures such as codecs envelop differences such that differences, modifications and variations in sensibility, subjectivities and institutions arise.

From eye to infrastructure: envisioning centres of calculation

One way to glimpse the different path opened up by Deleuze's account of sub-representative differences is to contrast it with the now canonical approaches to technology developed by science studies in the 1980s and 1990s. From a science and technology studies (STS or SCOT) perspective, we could say that the compression techniques presented in 1970s patents began to anchor centres of calculation in the midst of the chaotic, surging flows of cinematic, televisual and video images. A centre of calculation is, according to Bruno Latour,

> any site where inscriptions are combined and make possible a type of calculation. It can be a laboratory, a statistical institution, the files of a geographer, a data bank, and so forth. This expression locates in specific sites an ability to calculate that is too often placed in the mind (Latour, 1999, 304).

A codec is a site where calculations are done on images. Why should images need calculation? Ostensibly, the patents addressed the logistics of moving images. They proposed different ways of doing this: an 'apparatus capable of performing a discrete cosine transform with lightweight, low-cost, high-speed hardware suitable for real-time television image-processing' (Means, 1974, 1); 'a linear transform device capable of the rapid generation of linear transforms of a spatial or temporal signal' (Speiser, 1975, 2); or a 'system for encoding a present frame of video signals comprising, means for dividing the picture elements of the present frame into moving and nonmoving regions' (Haskell and Limb, 1972)

Moving images and, to a lesser extent, sound, indeed pose logistical problems. People in electronic media cultures have constantly imagined images circulating everywhere. Millions of images flicker across TV and cinema screens. Analogue television broadcasting solved the logistics problem in a Fordist fashion: images produced in studios passed through electromagnetic waves transmitted from central stations to many identical receivers. With few exceptions, the central platform of studio and transmitter came under State and/or large corporate ownership and control. As many studies of television and radio have shown, the forms of identity, representation and consumption associated with broadcast television align closely with political, economic and cultural forms of nation-states. However, despite all its recent attempts to offer interactivity, broadcast television cannot keep up with the information age's kaleidoscopic imagining of images flowing in many directions at once.

The calculations that codecs perform on images are not purely logistical. Put differently, any change in image logistics does not only concern media infrastructures. It affects embodied habits of perception, sometimes at a micro-perceptual level (sensations of brightness, colour and movement alter), sometimes at the level of media-historical habits (where, when and how images are made and seen), sometimes at the level of affect and temporality. Today, the calculations done by codecs take many forms and occur on many platforms: set-top boxes, mobile phones, cameras, computers, media players, and other gizmos cradle codecs.

Those arrangements are indeed fascinating. However, here I want to concentrate on the problem of how to connect embodied vision and media infrastructure. To comprehend this connection, I suggest, we need to shift attention from centres of calculation to *centres of envelopment*, a concept that Gilles Deleuze proposes late in *Difference and Repetition.* Centres of envelopment, according to Deleuze, interiorise differences in situations already structured by settled orders of representation, resemblance, extension and qualities perceived by human subjects. Such centres crop up in 'complex systems' where series of differences come into relation:

> to the extent that every phenomenon finds its reason in a difference of intensity which frames it, as though this constituted the boundaries between which it flashes, we claim that complex systems increasingly tend to interiorise their constitutive differences: the centres of envelopment carry out this interiorisation of the individuating factors (256).

Here, I will not be able to develop a full explanation of this concept. It comes late in *Difference and Repetition,* after long consideration of the constitution of time, memory, habit, difference and intensity. Importantly, the notion of a centre of envelopment is only one of a series of notions Deleuze developed to describe how differences relate to differences. Instead I concentrate on the basic idea that a centre of envelopment interiorises differences since this is almost a refrain in Deleuze's thought. Sub-representative differences are essentially interior since they differ intensively rather than extensively. If we take the rather risky step of understanding codecs as a centre of envelopment, we could ask how codecs interiorise 'constitutive differences' in ways that lead to a proliferation of moving images. The functioning of codecs, their capacity to compress and move moving images in space and time, and to generate sensations of qualities of colour, depth, light and form, certainly relies on calculation. However, calculation itself must derive in principle, for Deleuze, from 'a difference of intensity.' By tracking how intensities (or differences of difference) matter in codecs, we might begin to get a feel for how video material cultures eventuate. In substituting envelopment for calculation, Deleuze allows us to shift focus from extension to intensity, and thereby to describe how altered feelings, expectations and transformed sensibilities occur in and around technologies. If every thing expresses an Idea, if 'our experience of things, if you will, *can be conceptual*' (Henare, et al., 2006, 13), if every thing is a partial traversal of a field of affirmation or 'solutions' to the problem-setting imperative of an Idea, then feelings, sensibilities, and expectations of 'more to come' might be seen as symptoms of the impersonal individuations set in train by differences.

Spatio-temporal dynamisms and differences in repetition

The concept of centre of envelopment addresses something quite elementary: the persistence of intensities, singularities, and pure-spatial dynamisms in worlds largely organised by systems that order, calculate and represent. A centre of envelopment is like a micro-structural black hole peppering the fabric of the everyday, constantly interiorising elementary differences that give rise to phenomena, and serving as a site of actualisation for problem-setting Ideas. According to Deleuze, 'disparity ... difference or intensity ... is the sufficient reason of all phenomena' (Deleuze, 2001, 222). The problem in developing this philosophical perspective into an analytical engine for an anthropology of technology is that the technology of video codecs seems thoroughly actual. How do the densely interwoven calculations that codecs perform on images and sounds participate in what Deleuze called the 'asymmetrical synthesis of the sensible' (Deleuze, 2001, 222), or the ordinary everyday sensibilities of electronic media as they play out on information networks today?

Spatio-temporal dynamisms are critical here. As mentioned above, the primary characteristic of codecs is to allow digital images to move around in many forms. Video iPods, digital cameras, mobile phones, media players, dvds, satellite broadcasts, or internet media encode and decode video streams. Starting from the 1970s patent lineages, several decades of heavily funded public and private research has gone into liberating moving images from the bulkiness of film stock and projects, and the fixity of television transmitters and television sets. Tens of thousands of patents litter the wake of that research begun in the early 1970s. Today, probably the most widespread video or moving image codec is MPEG-2 stands at the confluence of several lineages of image encoding. DVD's, for instance, consist of MPEG-2 files. The International Standards Organisation (ISO formalised the MPEG-2 (a.k.a. H.262) encoding and decoding procedures as a standard (ISO/IEC 13818-1, 1999) in the early-1990s. The standard calls itself a 'transport system' (ISO/IEC 13818-1, 1999). The standards documents grew from the work of several thousand engineers and software designers meeting in Europe and North America. Other engineers implement the arrangements described in the documents in software or hardware codecs (coder-decoder). Video codecs for different standards (MPEG-1, MPEG-2, MPEG-4, H.261, H.263, the important H.264, theora, dirac, DivX, MJPEG, WMV, RealVideo, etch) litter electronic media networks.

Any codec that implements the MPEG-2 coding standard incarnates extraordinarily complicated calculations. Algorithmically it must draw on several distinct compression techniques: converting signals from time domain to frequency domain using Discrete Cosine Transforms, quantisation, Huffman and Run Length Encoding, motion compensation, timing and multiplexing mechanisms, retrieval and sequencing techniques. The standard borrows from the earlier, lower resolution video standard, MPEG-1 (ISO/IEC 11172-1:1993) and from other image standards such as JPEG. Legally, it imposes relations with many intellectual property claims (700 patents held by entertainment, telecommunications, government, academic and military owners administered through the MPEG-LA patent pool). Finally, it competes with many other codecs (e.g. China's AVC – Advanced Video Codec - versus the increasingly popular H.264 versus other versions such as Microsoft Windows VC-1 – Windows Media 9).

At the intersection of technical, legal and economic forces, codecs display a mosaic, composite character. They make sensible compromises between extensity and quality, between how images circulate (online, in media materials and transmission formats) and the sensations (brightness, detail, luminance, etc) associated with them. Their composite character reflects a constant and dynamic negotiation between the political economy of telecommunications and the media-historical perceptual habits of visual cultures. Telecommunications and media companies create cable, satellite and wireless network bandwidth as a market commodity. They sell bandwidth to anyone who will pay for information and images to move. At a very deep level, the architecture of an MPEG-2 codec reflects the assumption that all movement costs something in time, computation or bandwidth. Reducing the cost of movement means that more people can pay for that movement. If a codec compresses images, it makes their movement more likely. However any reduction in image size has to take into account human eyes, just as every reduction in the space between airline seats should take into account the postures and shapes of passengers' bodies. Eyes and ears do not have universal, timeless physiological properties. They have media-historical habits. Electronically mediated visual culture shapes eyes and ears, and creates perceptual habits at many levels. For instance, the conventions of the rectangular 4:3 ratio TV screen, the 16:9 ratio cinema screen, the number of scan lines, or the colour models of PAL/NTSC television broadcasts go deep into visual habits. Sensations of colour, texture, brightness and level of detail all feed into habits of viewing. The video codecs behind DVDs, High Definition Television, mobileTV for 3G cellular telephones, RealPlayer, or satellite digital video broadcasts attempts to take those expectations into account and meld them with the limited channel capacities of networks, broadcast spectrum or cables.

This outline of the situation of MPEG-2 codecs suggests that the spatio-temporal dynamisms found in a centre of envelopment link very different scales, levels and orders of movement and difference. Codecs make trade-offs between micro-perceptual sensations of brightness, colour, resolution and movement in trying to meet constraints concerning the cost of bandwidth on satellite or cable infrastructures (for instance, SkyChannel, a UK-based digital satellite TV broadcaster uses MPEG-2 compression to transmit many channels from one satellite). An analysis of trade-offs between image quality and media logistics could determine how the codec mangles or blends different interests (Pickering, 1995). Undoubtedly interesting, such an analysis would show how social relations have been displaced into the codecs. At base, it would indicate how '[m]en *[sic]* and things exchange properties and replace one another' (Latour, 1996, 61). The sheer volume of moving images, their extension, repetition and multiplication, would correlate with changes in their quality.

This approach to technological projects as a series of compromises or exchanges can yield rich results. However, any analysis of trade-offs between extension (or distribution) and image-quality remains fundamentally conservative in relation to differences. Here Deleuze's thought opens a very different path. Deleuze links repetition to irreducible differences. For every instance of repetition, Deleuze suggests, we need to look for the hidden repetition or resonances between differences. The system of resonances between differences comprises an Idea. 'Ideas have the power to affirm divergence; they establish a kind of resonance between divergent series' (Deleuze, 2001, 278). It is ‘pure movement’ (24) or the dynamic of an Idea in process of becoming, organising and unfolding a time and space in which repetition occurs. ‘Every thing thinks’ (254) because every thing, no matter how banal, ordinary, repetitive or singular, occurs within spatial-temporal processes that unfurl from movement inherent to an Idea. Movements of becoming 'are' the mode of existence of the dissimilar, the different or the unequal (Deleuze, 2001, 128).

Transformation into tendencies

In contrast to the notion of *problematization* that Paul Rabinow has distilled from the work of Michel Foucault (Rabinow, 2003), the problem-imperative of an Idea is not primarily concerned with truth and falsehood. It is based on a distribution of the singular and the ordinary. This can be illustrated by one of the two main ways video codecs handle differences and repetition in images. It comes from the first lineage of patents. This lineage treats human vision as a sub-representative process that detects differences in brightness, illumination and shadow rather than seeing things. For the purposes of image transmission, variations in brightness and colour count more than the typical analysis of representations in terms of figures, figure-ground relations, forms and contents.

In MPEG-2, the Discrete Cosine Transform (DCT) treats a video frame (or field) as a spatially extended distribution of brightness and colour. This treatment bears no resemblance to the figures and forms found in a given frame. It slices each frame into three separate planes: one of luminance (brightness) and two of chrominance (colour). The most detailed spatial calculation done by the codec on video images begins by analysing luminance of different areas in the picture. (It handles colour planes as larger blotches or patches.) The transform coding process seeks to elicit forms of repetition or redundancy from the variations of luminance across the plane of the image: the luminance or chrominance of a pixel in an image mostly exhibits a high level of correlation with neighbouring pixels. The colour and brightness of a pixel usually predicts those of its neighbours. So, a pixel belongs horizontally and vertically in a *sequence,* a well-defined order of increasing or decreasing values of luminance or chrominance. The codec defines a transformation that summarises or *contracts* the relations between the individual adjacent pixels as a *series* (the sum of the sequence) or as a periodic function that expresses variations of luminance and chrominance. It transforms a spatially-extended distribution of pixels into a function that can then be expressed as a series (for reasons explained below). The transformation pivots on the fact that the information content of an individual pixel is relatively small because by far the majority of adjacent pixels in a given image are identical. Tendencies or variations of colour or brightness across the plane of the image have more value that any particular element of the image.

The Discrete Cosine Transform treats each image as a set of periodic signal or waveform. Once transformed into a periodic signal, it can be broken down into a series of component cosine waves of different frequencies. For a given signal, some of the component waves contribute more energy to the overall signal than others. The transform coding selects only the most energetic or high amplitude components, and discards the rest. In more technical terms, the transform coding extracts the components of the signal with greatest *spectral density*. It reduces repetition (and hence storage space or transmission) by extracting the differentiating or individuating traits of luminance and chrominance in the image. When MPEG-encoded images are displayed on some screen, the decoding process reconstructs an image from the series. It assign values of luminance and chrominance to pixels on the screen on the basis of the coefficients in the series. It unfolds the displayed image by putting parts of it back together.

A great deal more could be said about the provenance and development of the Discrete Cosine Transform. It has a rich and continuing history of development coming out of natural sciences, and communications engineering. The important point is that via such techniques, codecs in a certain sense perceive the image. The spatio-temporal dynamisms they introduce in electronic media cultures concerns how far this movement of perception can go. These dynamisms do not only *exchange* properties between human and things. Deleuze's thought diverges from any exchange-based account of these dynamisms. Deleuze writes:

> Every spatio-temporal dynamism is accompanied by the emergence of an elementary consciousness which itself traces directions, doubles movements and migrations, and is born on the threshold of the condensed singularities of the body or object whose consciousness it is. (220)

In its analysis of spectral density, and selection of the most energetic component of the signal, the codec isolates tendencies or emphases. Perhaps this is something that also occurs in bodily perception. It treats what can be seen in an image as composed of tendencies and emphases that can be seized in a contractile movement. The codec embodies an 'elementary consciousness,' an awareness of transitions, of variable distributions of light and colour. It contracts variations of luminance and chrominance distributed across the plane of the image. Transform compression addresses the differences that spatially extend patches of brightness and colour in the image. This redundancy organises the image as an extended field of sensations of light and colour. The transform compression turns spatially distributed or extended repetition into transient differences expressed by coefficients of the different component frequencies. It synthesises space and time differently. The movement of contraction, and the elementary consciousness it presupposes, no longer occurs in the bodies of seeing subjects, but also in the technical apparatus of the codec and hence in assorted media technologies. Already here, we see how a thing expresses an Idea, if an Idea can be understood as a problem-setting system of differential elements (eyes, infrastructures, screens, images, calculations, etc) that form centres of envelopment around singularities.

Intensities and differences in sameness

By rendering an image as a set of (digitised or 'quantised') waveforms, MPEG-2 deeply fissures the objecthood of visual representation. Digital signal processing intimately concerns sensation rather than representation (of objects or figure). It processes brightness and colour, without concern for form or figure in an image. In its analysis of spectral density, and selection of the most energetic component of the signal, it isolates tendencies or emphases. It treats an image as composed of tendencies and emphases that can be seized in a contractile movement and summed as a series.

The codec connects two apparently ontologically distant entities, eyes and media infrastructures. They articulate embodied sensations of light and colour with the economically valuable markets for bandwidth of information and communication infrastructures. However, everything discussed so far concerns the extension of images in the world, and how to move images around more often, in greater numbers. Where are the interiorised differences characteristic of centres of envelopment?

One could say that codecs devote themselves to the reduction of difference (and this would be very much in line with critical theory and phenomenology's general treatment of media). They multiply the repetition of the same. In extending the reach of images, they constitute an *extensity* for video. In particular, we could say that the compression of the image responds to a demand for sameness (Terranova, 2004, 136). The demand is for the same qualities of brightness and chrominance wherever the image goes. Yet something more is stake in codec than massive reproduction of sameness. In Deleuze's account of the synthesis of the sensible, perceptions of extension and quality derive from differences, and particularly from *intensity*:

> The extension and 'extensity '(the result of the process of extending) of phenomena in space-time and qualities of sensation (*qualia* such as 'redness') attached to phenomena flow from a 'deeper disparateness' or 'difference in intensity' (236).

At core, intensities bind differences to each other. When differences come in relation to each other, intensities arise. Deleuze defines intensity as 'a difference which refers to other differences' (117). Intensities differ from representations. Such second order differences can be physical: differences such as pressure, temperature and density 'drive fluxes of matter or energy' (Delanda, 2002, 159); they can be also biological, psychic, social, aesthetic or philosophical (118), or some mixture of these. Intensities inhabit sensation.

An account of technology in terms of intensities would be somewhat novel. Intensities arise when differences come into relation. Just as the DCT turns the spatial distribution of brightness and colour of an image into series of frequency components, in *Difference and Repetition* Deleuze developed a series-based explanation of intensity. Deleuze abstractly understands differences as series:

> The first characteristic seems to us to be organisation in series. A system must be constituted on the basis of two or more series, each series being defined by the differences between the terms which compose it (117).

According to the quasi-mathematical notion of series, each term in a series differs from preceding and following terms: A' – A'' – A''' – A'''' – etc. Deleuze says that intensity arise from two or more series of differences in relation. Within the codec, DCT literally generates one set of series or one set of differential relations. What is the other series?

Motion estimation: 'the embedding of presents within themselves'

MPEG video never flickers. The second lineage of patents starting in the 1970s took movement as its problem. Using techniques they inherit from that lineage, MPEG codecs perform a second major calculation called *motion compensation*. Transform coding treats individual pictures themselves as spatial distributions of luminance and chrominance values to be reorganised in series. Motion compensation, by contrast, treats the relation between successive pictures in terms of vectors of movement. MPEG video never flickers because it calculates and predicts transitions between pictures. It dismantles the temporally discrete recording, storing and transmitting of pictures put in place by film strip or television. Film frames and analogue video 'fields' have fixed boundaries. They leave the habits of human perception to bridge between frames. In video codecs, calculation fills in the gap between frames. Again, the technical apparatus of codec takes on part of the work of embodied perception in the interests of changing the relation between body and media infrastructure.

In order to do this, components of the codec involved in motion estimation assume that nothing much happens between successive frames apart from spatial transformations (translation, rotation, skewing, etch) of parts (macroblocks) of the image. In the process of encoding a video sequence, the MPEG-2 codec analyses each picture in relation to a previous and a future reference picture. It calculates and transmits series of motion vectors describing how different parts of the frame move in relation to their position in the reference pictures. Motion compensation does not distinguish types of movement. Film scholars often distinguish camera movements (pans, tilts, zoom in) and then ascribe different motivations to movement. MPEG-2 decomposes every movement into the directions and rate of movement of macroblocks. A typical PAL DVD image contains roughly 800 macroblocks. At 30 frames/sec, the codec tracks the movement of roughly 24000 macroblocks. The 'pictures' streamed on the internet, down-linked via satellite or burned on DVD mostly comprise long series of vectors describing blocks in motion. Decoding the MPEG stream means turning these vectors back into patterns of blocks moving around in frames on the screen. The decoding side of a codec frenetically recomposes images from blocks moving in all directions.

We could view motion compensation as a second series of differences. Motion compensation alters the temporality of moving images. We have already seen that, in a first centre of envelopment, transform coding contracts the image into a series of differences of brightness and colour. In motion compensation, the image or picture itself is no longer the elementary component of motion perception. Motion compensation re-organises the picture into series of motion vectors describing relative movements of blocks in time.

Images in overflow

What happens when these two series come into communication in the codec? According to Deleuze, when series of differences communicate with each other, spatio-temporal dynamisms emerge:

> Once communication between heterogeneous series is established, all sorts of consequences follow within the system. Something passes between the borders, events explode, phenomena flash, like thunder and lightning. Spatio-temporal dynamisms fill the system, expressing simultaneously the resonance of the coupled series and the amplitude of the forced movement which exceeds them (118).

As centres of calculation, codecs repeat and render images with a strong concern for accurate repetition. The MPEG-2 bitstream and most other contemporary video codecs puts two different series in relation. First they generate series of values that express a key frame, and then they generate series of values that express how elements of that key frame move. However, as centres of envelopment, they also do something that lies at the heart of technological repetition. Deleuze refers to this in terms of events, explosions, excess and flashes. Where do we see these? In a sense, the consequences can be seen everywhere today in the growth of video material culture. A preliminary analysis of the spatio-temporal dynamisms could examine ways in which images have become extended and new qualities of images as they proliferate.

The first facet concerns the spatial-temporal dynamism of calculation done within the codecs themselves. The actual ratio of transform-coded or 'intra-frame' pictures (I-pictures) and 'inter-frame' pictures (P- and B-pictures) in a given MPEG-2 bitstream varies. It depends on where the encoding is done, the bandwidth of the expected transmission channel and the size of the display screen. In an MPEG-2 data-stream, the *Group of Pictures* (GOP) structure defines the precise mixture of different picture-types at encoding time. A GOP usually has 12 or 15 pictures in a sequence such as I_BB_P_BB_P_BB_P_BB_P_BB_. The order of the pictures in a GOP does not correspond to their viewing sequence. A dozen or so block motion-compensation frames follow one transform coded I-picture. The ratio of different picture types in a bitstream directly affects the encoding time, the transmission time and the decoding time. Calculating motion compensation is much slower than the highly optimised block transforms. Yet motion vectors take much less time to transmit than transform-coded pictures. (Some DVD players offer an option that displays the bitrate of the images on screen. Variations in this bitrate indicate different ratios of DCT and motion estimation being done by the player's codec.)

Even the drastically simplified description I have given should indicate that both transform coding and motion compensation entail much comparison, sorting and shifting of numerical values around in arrays. The time and cost of this calculation can be high, and every reduction of it matters. Practically, codecs must make direct trade-offs between computational time and bandwidth/storage space. A highly compressed image will take longer to generate but require less storage space or network bandwidth. The trade-offs made in encoding sometimes result in artefacts visible on screen as such as blocking and mosaic effects. At times, motion prediction cannot work smoothly. A change in camera shot or an edit breaks the flow of movement between adjacent frames. In that case, the codec falls back on transform coding. A momentary but visible splintering of the images – so called 'motion blocking' – occurs. Whereas film flickers, digital video 'motion-blocks.' Motion-blocking appears as horizontal and vertical edges where the MPEG-2 motion compensation algorithm has sliced the image into macroblocks. (We could also look at 'reordering delay,' a 'delay in the decoding process that is caused by frame reordering' (ISO/IEC 13818-1, 1999). All of these fringe-effects and the technical designs processes that trade-off between different forms of compression and rendering of images come from the space-time of calculation itself.

The second facet concerns the spatio-temporal dynamism of video materials on contemporary mediascapes. If we ignore all the physicalities of spectatorship, video encoded and decoded by codecs probably looks much the same. Much effort has gone into making them look the same or almost the same. In fact, researchers and standards organizations such as the ITU (International Telecommunications Union) have developed complicated testing regimes that distinguish objective versus subjective video quality. Are these visible artefacts, themselves the effect of technical compromises made in the name of cost, the only sign of an event for a codec, the only phenomena that flash and explode? Videos might look the same but circulate very differently. As Deleuze writes, 'difference pursues its subterranean life while its image reflected by the surface is scattered' (240). The production of sameness seeks to cement the relation between eye and screen image through intensified sensation. It does detailed work on framing, resolution, brightness, colour and movement. Yet this sameness envelopes a very different relation to infrastructure. It generates powerful spatio-temporal dynamics in the relation between eye and infrastructure. Video churns on the internet. Video streams and broadcasts surge (satellite TV, digital high definition, Web). New ecologies of spectatorship, consumption and occasionally citizenship both mimic and differ from cinema and television spectatorship. As formats, platforms, and products mushroom and new forms of making, viewing, and moving populate flows of images.

Finally, the two lineages of patents I mentioned at the beginning of the paper suggest another spatio-temporal dynamism or 'forced movement' around codecs. Almost 700 patents apply to the MPEG-2 standard (http://www.mpegla.com/m2/m2-patentlist.cfm). Tremendously tight intellectual property arrangements bind the codecs. Each aspect of the codecs calculations we have been discussing undergoes intensive variation as slight reductions in computational costs, and detours around existing intellectual properties, are sought out.

Centres of envelopment interiorise differences: 'we claim that complex systems increasingly tend to interiorise their constitutive differences' writes Deleuze (256), and at the same time give rise to spatio-temporal dynamisms, patterns of extension and qualitative differences on multiple scales. As things that think, how do codecs interiorise 'constitutive differences'? The MPEG-2 codec, I have suggested, can be understood as a composite process of change in the relations between between eye and media infrastructure. An intensity inhabits these relations. Deleuze describes an eye as 'bound light':

> An animal forms an eye for itself by causing scattered and luminous excitations to be reproduced on a privileged surfaces of its body. The eye binds light, it is itself a bound light (96).

The codec also binds light, but not just to 'privileged surfaces' of the body. It binds light to the movement of images in media infrastructures. It causes some forms of luminance, chrominance and movement to be reproduced on 'privileged surfaces' within media technologies. In order to do this, different temporal syntheses must come into relation: the habit-based contraction of perception in DCT; the embedding of presents in each other in motion estimation.

The relation between transform coding and motion estimation, between the technical treatments of luminance-chrominance and movement, is however unstable. It envelops or comprehends some aspects of the movement of images, but not all of them:

> '[E]ach intensity *clearly* expresses only certain relations or certain degrees of variation. Those that it expresses clearly are precisely those on which it is focused when it has the *enveloping* role. In its role as the *enveloped,* it still expresses all relations and all degrees, but confusedly' (252)

The MPEG-2 codec only clearly expresses certain relations between eye and infrastructure. It focuses on brightness, colour, and movement of images as framed by the history of photography, cinema and television but puts this in relation to the political economy of telecommunications (with its constraints on bandwidth, memory, processing power, etc). A primary intensity comes from the light-binding relations between eye and infrastructure. This relation frames the dynamics. The two lineages of patents – transform coding and motion estimation - represent different treatments of repetition in digital video. When transform coding and motion estimation come together in MPEG-2, different spatio-temporal dynamics and forms of interiorisation result. Interiorisation occurs when the spatio-temporal phenomena begin to cover over the intensities that gave rise to them. A proliferation of patents around codecs occurs. At the same, the trade-offs made between computation and bandwidth in the MPEG-2 codecs mean that the physical forms of codecs proliferate in chips, software, gadgets, and boxes. Finally, the circulation of video itself changes. New ecologies of images burgeon.

The codecs, I have been arguing, envelop relations between eye and infrastructure. Does it make any difference that the streaming digital video of decapitation never flickers? It could be argued that the intensive paths generated by codecs in the extension of images make no difference to the viewers. In other words, viewers might see straight through the codecs. They might be not be seen. For viewers, however, execution and hostage videos and vast pools of pornography open up alongside the streams of wedding, baptism and graduation videos. Viewers may not be highly conscious of how brightness, chrominance and movement have been minutely altered by the codec. These differences can be easily cancelled out or remain almost imperceptible. This does not mean that they make no difference. On the contrary, the proliferation of video materials and the degrees of variation opening up around video streams today suggest that viewers are caught up in the spatio-temporal dynamisms of video material culture.

The Idea that thinks the thing

How far have we come in thinking technologies as thinking things? No thing is a pure expression of an Idea. An Idea cannot be through as such. Moreover, things inhabit already structured worlds. Hence, Deleuze's notion of a centre of envelopment, a site for actualisation of an Idea in worlds organised by systems of extension and quality, reflects the need to accommodate in any radically constructivist account of things the existing orders of representation, personhood, and identity. Certainly, nothing of what I have said of codecs belongs solely to advanced technologies. It is not as if codecs think more than, say, photographs, rock carvings or oil-paintings. The singularity of codecs, the different worlds to be found in them, come from the intensities they put in relation. Those differences actualise in the spatio-temporal, scale-transforming dynamics of video material culture.

Why privilege the two patent lineages, and the series of transformations they produce in the moving images as the locus of differences? The patent lineages stand in the discussion above as centres of envelopment, as tendencies that attract much mental effort to create new variants, versions and modifications as well as property claims. Centres of envelopment mark the surface of things with zones of infolding and ingression, with regions where an Idea or a problem is in actualisation, where 'intensive reasons' play out.

Crucially, to make sense of Deleuze's claim that everything things, we cannot treat an Idea as a concept, a construct that represents or holds difference in identity. It is a system of 'positive, differential multiplicity' (288) that needs to be thought from a radical constructivism. The effort to think oriented by a radically constructivist alignment is particular relevant to composite things that exist in a flux of documents, international standards, versions, software and hardware implementations, and diverse and constantly decentered applications.

References

Boundas, Constantin. "Deleuze-Bergson: An Ontology of the Virtual." In *Deleuze : A Critical Reader*, edited by Paul Patton, x, 316 p. Oxford ; Cambridge, Mass.: Blackwell, 1996.

Delanda, Manuel. *Intensive Science and Virtual Philosophy*. (London & New York: Continuum, 2002).

Deleuze, Gilles. *Difference and Repetition*. Translated by Paul Patton, *Athlone Contemporary European Thinkers*. (London ; New York: Continuum, 2001).

Haskell, Barin G., and John O. Limb. "Predictive Video Encoding Using Measured Subject Velocity." USA: Bell Telephone Laboratories, 1972.

Haskell, Barin G., and Atul. Puri. " Conditional Motion Compensated Interpolation of Digital Motion Video." USA: AT&T Bell Telephone Laboratories, 1990.

Henare, Amiria J. M., Martin Holbraad, and Sari Wastell. *Thinking through Things : Theorising Artefacts in Ethnographic Perspective*. (Milton Park, Abingdon, Oxon ; New York: Routledge, 2006).

ISO/IEC 13818-1. "Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Systems, Pictures and Associated Audio Information: Video." 1999.

Latour, Bruno. *Aramis, or the Love of Technology*. Translated by Catherine Porter. (Cambridge, MA & London: Harvard University Press, 1996).

———. *Pandora's Hope : Essays on the Reality of Science Studies*. (Cambridge, Mass.: Harvard University Press, 1999).

Means, Robert. "Discrete Cosine Signal Transform Processor." USA: Secretary of the Navy, United States Government, 1974.

Pickering, Andrew. *The Mangle of Practice : Time, Agency, and Science*. (Chicago: University of Chicago Press, 1995).

Rabinow, Paul. *Anthropos Today. Reflections on Modern Equipment*. (Princeton and Oxford: Princeton University Press, 2003).

Speiser. "Serial Access Linear Transform." USA: Secretary of the Navy, US Government, 1975.

Terranova, Tiziana. *Network Culture. Politics for the Information Age*. (London: Pluto Press, 2004).