Surveillance (oversight), Sousveillance (undersight), and Metaveillance (seeing sight itself)

Steve Mann
Humanistic Intelligence Institute, Veillance Foundation
330 Dundas Street West, Toronto, Ontario, Canada, M5T 1G5
http://www.eyetap.org

Abstract

Surveillance is an established practice that generally involves fixed cameras attached to fixed inanimate objects, or PTZ (Pan Tilt Zoom) cameras at a fixed position. Sur-veillance only provides part of the veillance story, and often only captures a partial truth. Further advances in miniaturization, together with wireless communication technologies, are giving rise to kine-matic veillance (“kineveillance”): wearable, portable, and mobile cameras, as well as unpersoned aerial vehicles (UAVs). These additional veillances give us a more complete picture: multiple viewpoints from multiple entities bring us closer to the truth. In contrast to the extensive mathematical and conceptual framework developed around surveillance (e.g. background subtraction, frame-differencing, etc.), now that surveillance is no longer the only veillance, we need new mathematical and conceptual understandings of imaging and image processing. More importantly we need new tools for understanding the many veillances and how they interact. Therefore this paper introduces metaveillance and the veillance wavefunction for metasensing: the sensing of sensors and the sensing of their capacity to sense.

1. Surveillance is not the only Veillance

Surveillance is an established field of research and practice[62, 61, 15, 21, 34] in which sensors are affixed to stationary objects. “Surveillance” is a French word that means “oversight” (“sur” means “over”, and “veillance” means “sight”), and suggests an omniscient authority’s “God’s eye view” from above [19, 40].

A common surveillance camera housing is the dome, in which the camera is concealed behind a transparent but darkly smoked plastic hemisphere. Inside the hemisphere there is also usually an additional opaque black shroud to further conceal the camera. A common design objective in surveillance systems is concealment. In addition to camera domes, concealment is also achieved by hiding cameras inside other objects such as streetlights (Fig. 1), which can be placed throughout entire neighbourhoods or cities, for complete surveillance, while themselves remaining hidden.

Figure 1. “Stealth” streetlight camera by Apollo, and “Embedded Invisible PTZ IP Camera” by OWLS AG. Here surveillance cameras are concealed inside ordinary objects like streetlights, which can be placed throughout entire neighbourhoods or cities, for complete surveillance, while themselves remaining hidden.

1.1. Truth and the need for a scientific understaning of Veillance

The opposite of hypocrisy is integrity. In Greek mythology, Aletheia is the goddess of truth and the Greek word “alethia” means “disclosure”, “not concealed”, “truth”, or “sincerity”. Alethia’s opposites are Ape (deception), pseudologoi (lies), and Dolos (trickery)[57, 56].

Science is a human endeavour in which we attempt to go wherever the truth may lead us, in pursuit of new
discoveries[31]. The term OpenScience was coined by S. Mann in 1998 to emphasize this need for truth and open disclosure [48], and if surveillance cannot deliver on that promise, we need to look at other veillances!

1.2. The many Veillances

Recent advances in miniaturization and wireless communication have made mobile sensing practical, giving rise to a transition from static veillances to kinematic veillance (“kineveillance”), i.e. sensing from a moving/kinematic frame-of-reference.

More generally, surveillance (oversight) is no longer the only veillance. We also have sousveillance (undersight)[30, 60, 7, 39, 4, 25, 63, 51, 6, 22, 59, 13, 5, 55, 3, 65, 54, 58, 35], in the form of wearable computing and self-sensing [38, 49] (now known as “Quantified Self”), and the wearable face-recognizer [64, 37], and computer vision systems to help people see and remember.

Surveillance happens when we’re being watched and sousveillance happens when we do the watching. And with social networking, we now have co-veillance (side-to-side watching, e.g. when people watch each other) [50]. Thus we must think of Veillance in a broader sense beyond merely Surveillance. (See Fig. .)

1.3. Surveillance and AI (Artificial Intelligence)

Surveillance is a broad concept. It goes beyond visual (camera-based) sensing, to include audio surveillance, data surveillance, and many other forms of surveillance. When we say “we’re being watched” we often mean it in a broader sense than just the visual. When police are listening in on our phone conversations, that is still surveillance. So, more generally, “surveillance” refers to the condition of being sensed. We often don’t know who or what is doing the sensing. When a machine learning algorithm is sensing us, that is also surveillance. AI (Artificial Intelligence) often involves surveillance. Some surveillance is harmful (“malveillance”). Some is beneficial (“bienveillance”), like when a machine senses our presence to automate a task (flushing a toilet, turning on a light, or adjusting the position of a video game avatar).

HI (Humanistic Intelligence) is a new form of intelligence that harnesses beneficial veillance. HI is defined by Kurzweil, Minsky, and Mann as follows:

“Humanistic Intelligence [HI] is intelligence that arises because of a human being in the feedback loop of a computational process, where the human and computer are inextricably intertwined. When a wearable computer embodies HI and becomes so technologically advanced that its intelligence matches our own biological brain, something much more powerful emerges from this synergy that gives rise to superhuman intelligence within the single ‘cyborg’ being.” [52]

HI involves an intertwining of human and machine in a way that the human can sense the machine and vice versa, as illustrated in Fig. 4.

HI is based on modern control-theory and cybernetics, and as such, requires both controllability (being watched) and observability (watching), in order to complete the feedback loop of any kind of HI. In this way, surveillance (being watched) and sousveillance (watching) are both required in proper balance
Figure 4. **The Six Signal Flow Paths of HI**: A human (denoted symbolically by the circle) has senses and effects (informatic inputs and outputs). A machine (denoted by the square) has sensors and effectors as its informatic inputs and outputs. But most importantly, HI involves intertwining of human and machine by the signal flow paths that they share in common. Therefore, **these two special paths of information flow are separated out**, giving a total of six signal flow paths.

Figure 5. **HI requires both Veillances**: machines must be able to sense us, and we must be able to sense them! **Thus veillance is at the core of HI**. In this sense Surveillance is a half-truth without sousveillance. Surveillance alone does not serve humanity.

for the effective functioning of the feedback between human and machine. Thus Veillance (Surveillance AND Sousveillance) is at the core of HI. (See Fig. 5.)

Poorly designed human-computer interaction systems often fail to provide transparency and immediacy of user-feedback, i.e. they fail to provide sousveillance. As an example of such a “Machine of Malice”, an art installation was created by author S. Mann to exemplify this common problem. The piece, entitled “Digital Lightswitch” consists of a pushbutton (push-on/push-off) light switch where a random 3 to 5 second delay is inserted along with a 10 percent packet loss. The parameters were adjusted to maximize frustration in order to show a negative example of what happens when we fail to properly implement a balanced veillance.

2. **Sequential Wave Imprinting Machine**

This section describes some unpublished aspects of a wearable computing and augmented reality invention by author S. Mann for making visible various otherwise invisible physical phenomena, and displaying the phenomena in near-perfect alignment with the reality to which they pertain. As an embodiment of HI (Humanistic Intelligence), the alignment between displayed content and physical reality occurs in the feedback loop of a computational or electric process. In this way, alignment errors approach zero as the feedforward gain increases without bound. In practice, extremely high gain is possible with a special kind of phenomenological amplifier (ALIA = Alethioscopic/Arbitrary Lock-In Amplifier / “PHENOMENAmplifier™”) designed and built by the author to visualize veillance.

An example use-case is for measuring the speed of wave propagation (e.g. the speed of light, speed of sound, etc.), and, more importantly, for canceling the propagatory effects of waves by sampling them in physical space with an apparatus to which there is affixed an augmented reality display. Whereas standing waves, as proposed by Melde in 1860, are well-known, and can be modeled as a sum of waves traveling in opposite directions, we shall now come to understand a new concept that the author calls “sitting waves”, arising from a product of waves traveling in the same direction, as observed through a phenomenological augmented reality amplifier, in a time-integrated yet sparsely-sampled works 90 percent of the time, and, combined with the delay, users would often press it once, see no immediate effect, and then press it again (e.g. turning it back off before it had time to come on). See Fig. 6.
Cameras and microphones and other sensors don’t EMIT energy, but, rather, they sense energy. Cameras sense light energy (photons). Microphones sense sound energy.

Thus $\langle \psi_\mu | \psi_\mu \rangle$ does not correspond to any real or actual measurement of any energy like sound or light, but, rather, it is a metaquantity, i.e. a sensing of a sensor, or a sensing of the capacity of a sensor to sense!

The word “meta” is a Greek word that means “beyond”, and, by way of examples, a meta-conversation is a conversation about conversations. A meta joke is a joke about jokes. Metadata (like the size of an image or the date and time at which it was taken) is data about data. Likewise metaveillance (metasensing) is the seeing of sight, or, more generally, the sensing of sensing (e.g. sensing sensors and sensing their capacity to sense).

Thus the space around a video surveillance (or sousveillance) camera, or a hidden microphone, can have, associated with it, a metawavefunction, $\psi_\mu$, in which $\langle \psi_\mu | \psi_\mu \rangle$ increases as we get closer to the camera or microphone, and, for a fixed distance from the camera or microphone, $\langle \psi_\mu | \psi_\mu \rangle$ typically increases when we’re right in front of it, and falls off toward the edges (e.g. as many cameras have lens aberrations near the edges of their fields of view, and microphones “hear” best when facing directly toward their subject).

If we are a long way away from a camera, our face may occupy less than 1 pixel of its resolution, and be unrecognized by it. By this, I mean that a person looking through the camera remotely, or a machine learning algorithm, may not be able to recognize the subject or perhaps even identify it is human. As we get closer to the extent that we occupy a few pixels, the camera may begin to recognize that there is a human present, and as we get even closer still, there may be a point where the camera can identify the subject, and aspects of the subject’s activities.

Likewise with a microphone. From far away it might not be able to “hear” us. By this I mean that a remote person or AI listening to a recording or live feed from the microphone might not be able to hear us through the microphone.

Thus $\langle \psi_\mu | \psi_\mu \rangle$ gives us a probability of being recognized or heard, or the like.

Let’s begin with the simplest example of a metaveillance wave function, namely that from a microphone, in the one-dimensional case, where we move further from, or closer to, the microphone along one degree-of-freedom.

The subscript, $\mu$, is dropped when it is clear by context that we are referring to a metawavefunction rather than an ordinary wavefunction as we might find.
in quantum mechanics, or the like.

Consider an arbitrary traveling metawave function \( \psi(x, t) \) whose shape remains constant as it travels to the right or left in one spatial dimension (analogous to the BCCE of optical flow[27]). The constancy-of-shape simply means that at some future time, \( t + \Delta t \), the wave has moved some distance along, say, to \( x + \Delta x \). Thus:

\[
\psi(x, t) = \psi(x + \Delta x, t + \Delta t).
\] (2)

Expanding the right hand side in a Taylor series, we have:

\[
\psi(x + \Delta x, t + \Delta t) = \psi(x, t) + \psi_x \Delta x + \psi_t \Delta t + h.o.t.,
\] (3)

where h.o.t. denotes (higher order terms). Putting the above two equations together, we have:

\[
\psi_x \Delta x + \psi_t \Delta t + h.o.t. = 0.
\] (4)

If we neglect higher order terms, we have:

\[
\frac{\Delta x}{\Delta t} \psi_x + \psi_t = 0, \forall \Delta t \neq 0,
\] (5)

where the change in distance, divided by the change in time, \( \frac{\Delta x}{\Delta t} \) is the speed, \( c \) of the traveling wave.

In the case of a surveillance camera, or a microwave motion sensor (microwave burglar alarm), \( c \) is the speed of light. In the case of a microphone (or hydrophone), \( c \) is the speed of sound in air (or water).

More generally, waves may travel to the left, or to the right, so we have:

\[
\psi_t \pm c\psi_x = 0.
\] (6)

Multiplying these solutions together, we have:

\[
\left( \frac{\partial}{\partial t} - c \frac{\partial}{\partial x} \right) \left( \frac{\partial}{\partial t} + c \frac{\partial}{\partial x} \right) \psi = 0,
\] (7)

which gives:

\[
\frac{\partial^2 \psi}{\partial t^2} = c^2 \frac{\partial^2 \psi}{\partial x^2}.
\] (8)

This is the wave equation in one spatial dimension, as discovered by Jean-Baptiste le Rond d’Alembert in 1746, due to his fascination with stringed musical instruments such as the harpsichord[16], which Euler generalized to multiple dimensions:

\[
\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = 0,
\] (9)

where \( \nabla^2 \) is the Laplacian (Laplace operator, named after Pierre-Simon de Laplace, who applied it to studying gravitational potential, much like earlier work by Euler on velocity potentials of fluids[20]). This further generalizes to the Klein-Gordon generalization of the Schrödinger wave equation:

\[
\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi + \left( \frac{mc}{h} \right)^2 \psi = 0,
\] (10)

for a particle of mass \( m \), where \( h = h/2\pi \) is Planck’s constant.

More generally, we can apply a wide range of wave theories, wave mechanics, wave analysis, and other contemporary mathematical tools, to metawaves and veillance, and in particular, to understanding veillance through phenomenological augmented reality[43].

2.2. Broken timebase leads to spacebase

Waves in electrical systems are commonly viewed on a device called an “oscillograph”[26, 32]. The word originates from the Latin word “oscillare” which means “to swing” (oscillate), and the Greek word “graph” which means drawing or painting. A more modern word for such an apparatus is “oscilloscope”[33, 24] from the Latin word “scopium” which derives from the Greek word “skopion” which means “to look at or view carefully” (as in the English word “skeptic” or “skeptical”). The oscillograph or oscilloscope is a device for displaying electric waves such as periodic electrical alternating current signals.

In 1974 author S. Mann came into possession of an RCA Cathode Ray Oscillograph, Type TMV-122, which was, at the time, approximately 40 years old, and had a defective sweep generator (timebase oscillator). Since it had no timebase, the dot on the screen only moved up-and-down, not left-to-right, thus it could not draw a graph of any electrical signal, but for the fact that Mann decided to wave the oscillograph back and forth left-to-right to be able to see a two-dimensional graph. In certain situations, this proved to be a very useful way of viewing certain kinds of physical phenomena, when the phenomena could be associated with the position of the oscilloscope. This was done by mounting a sensor or effector to the oscilloscope. In one such experiment, a microphone was mounted to the oscilloscope while it was waved back and forth in front of a speaker, or vice-versa. In another experiment, an antenna was mounted to the oscilloscope while it was waved back and forth toward and away from another antenna. With the appropriate electrical circuit, something very interesting happened: traveling electric waves appeared to “sit still”. A simple superheterodyne receiver was implemented by frequency mixing with the carrier wave, e.g. \( \cos(\omega t) \) of the transmitter. When one of the two antennae (either one) is attached to an oscilloscope with no sweep (no timebase), while the other remains stationary, the oscilloscope traces out the radio wave as a function of space rather than of time.

If the transmitted wave is a pure unmodulated carrier, the situation is very simple, and we can visualize
the carrier as if “sitting” still, i.e. as if we’re moving at the speed of light, in our coordinate frame of reference, and the wave becomes a function of only space, not time. The wave begins as a function of spacetime:

\[ \psi(x, t) = \cos(\omega t - kx); \text{ wavenumber } k = \omega/c. \]  

(11)

In this case the received signal, \( r(x, t) \) is given by:

\[ \cos(\omega t - kx) \cos(\omega t) = \frac{1}{2} \cos(2\omega t - kx) + \frac{1}{2} \cos(kx). \]  

(12)

Half the received signal, \( r \), comes out at about twice the carrier frequency, and the other half comes out in the neighbourhood of DC (near zero frequency). The signal we’re interested in is the one that is not a function of time, i.e. the “sitting wave”, which we can recover by lowpass filtering the received signal to get:

\[ s(x) = \frac{1}{2} \cos(kx). \]  

(13)

This operation of multiplication by a wave function was performed at audio frequencies using a General Radio GR736A wave analyzer, and at other times, using a lock-in amplifier, and at radio frequencies using four diodes in a ring configuration, and two center-tapped transformers, as is commonly done, and at other times using modified superheterodyne radio receiving equipment.

A drawback of some of these methods is their inability to visualize more than one frequency component of the transmitted wave.

2.3. Metawaves

When the transmitter is stationary (whether it be an antenna, or a speaker, or the like) and the receiver (e.g. a receiving antenna, or a microphone) is attached to the oscilloscope, the device merely makes visible the otherwise invisible sound waves or radio waves. But when these two roles are reversed, something very interesting happens: the apparatus becomes a device that senses sensors, and makes visible their sensory receptive fields. In the audio case, this functions like a bug sweeper, in which a speaker is moved through the space to sense microphones, but unlike other bug sweepers the apparatus returns the actual underlying veillance wavefunction, as a form of augmented reality sensory field, and not just an indication that a bug is present.

Now consider the case in which the transmitted signal is being modulated, or is otherwise a signal other than a pure wave \( \cos(\omega t) \). As an example, let’s consider \( \psi(x, t) = \cos(\omega t - x) + \cos(5\omega t - x) \), so the received signal is:

\[ r(x, t) = \frac{1}{2} \cos(x) + \frac{1}{2} \cos(x - 2\omega t) \]

\[ + \frac{1}{2} \cos(5x - 4\omega t) + \frac{1}{2} \cos(5x - 6\omega t), \]  

(14)

which when lowpass filtered, only gives us the fundamental. Thus a wave analyzer or modern lock-in amplifier such as Stanford Research Systems SR510 cannot be used to visualize such a wave. A more traditional lock-in amplifier, such as Princeton Applied Research PAR124A, will visualize harmonics, but in the wrong proportion, i.e. since the reference signal is a square wave, higher harmonics are under-represented (note that the Fourier series of a square wave falls off as \( 1/n \), e.g. the fifth harmonic comes in at only 20 percent of its proper strength).

Thus existing lock-in amplifiers are not ideal for this kind of visualization in general.

2.4. A Lock-in amplifier designed for Metaveillance

The approach of Mann was therefore to invent a new kind of lock-in amplifier specifically designed for augmented reality visualizations of waves and metawaves.

Whereas a common ideal of lock-in amplifier design is the ability to ignore harmonics, in our application we wish to not only embrace harmonics, but to embrace them equally. If we were to turn on our sensing of harmonics, one at a time, we would be witnessing a buildup of the Fourier series of our reference signal. For the square wave, each harmonic we add to our reference signal, allows more and more of the measured signal harmonics through, but colored by the coefficients of the Fourier series representation of the square wave. Figure 8 illustrates a comparison of the the reference signal waveforms of the PAR124A lock-in amplifier with the modified lock-in amplifier of the Sequential Wave Imprinting Machine (SWIM)[36].

2.5. From timebase to spacebase

A more recent re-production of this early experiment is illustrated in Figure 9, with an oscilloscope-based implementation. An LED-implementation is shown in Fig. 10. An Android-based version was also created.

2.6. Wearable SWIM

Oscillographs were too heavy to swing back-and-forth quickly (RCA Type TMV-122 weighs 40 pounds or approx. 18kg). So in 1974, Mann invented the SWIM (Sequential Wave Imprinting Machine). The SWIM, waved back-and-forth quickly by hand or robot, visualizes waves, wave packets (wavelets), chirps, chirplets, and metawaves, through PoE (Persistence of Exposure) [36]. See Fig. 11 and http://wearcam.org/swim

3. Phenomenological AR bots and drones

Constrained to linear travel, SWIM is useful as a measurement instrument (Fig. 12). Over the years the
Figure 8. Left: A modern LIA (Lock In Amplifier) ignores all but the fundamental. Older LIAs use polarity reversal and are thus sensitive to increasing harmonics on a 1/n basis where n = 1, 3, 5, ... This is why older LIAs often work better with the SWIM (Sequential Wave Imprinting Machine)[36], as long as they're modified to compensate for weaker higher frequency components of the waveform being visualized. Right: Reference waveforms of Mann’s “Alethioscope” have equal weightings of all harmonics. As we include more harmonics, instead of approaching a square wave, we approach a pulse train. Early SWIM used a pulse train as its reference signal. This made time “sit still” (like a strobe light on a fan blade) for a true and accurate AR (Augmented Reality) visualization without distortion.

Figure 9. We’re often being watched by motion sensors like the microwave sensor of Fig. 2. Left: When we try to look at the received baseband signal from the sensor, as a function of time (artificially spatialized), it is difficult to understand and has little meaning other than a jumble of lines on the screen. Center: When we shut off the timebase of the oscilloscope, and wave it back and forth, we see the very same waveform but displayed naturally as a function of space rather than time. Right: Stephanie, Age 9, builds a robot to move SWIM back-and-forth in front of the sensor. As a function of space the displayed overlay is now in perfect alignment with the reality that generated it. This alignment makes physical phenomena like electromagnetic fields more comprehensible and easy to see and understand.

Figure 10. The SWIM’s multicomponent/arbitrary waveform lock-in amplifier. Square wave visualized using multi-component reference signal $\cos(\omega t) + \cos(3\omega t)$ making visible the first two terms of its Fourier series expansion, resulting in phenomenological augmented reality display of $\cos(\omega t) + 1/3\cos(3\omega t)$ on 600 LEDs in a linear array rapidly swept back and forth on the railcar of an optical table. This suggests expanding the principles of compressed sensing[8, 18] to metaveillance. Inset image: use beyond veillance, e.g. atlas of musical instruments (trumpet pictured) and their waveforms. Amplifiers in picture: SR865 (drifty with screen malfunction); SR510; and PAR124, not used at this time.

Figure 11. Miniaturized wristworn SWIM: Metaveillance for everyday life. Left: Invented and worn by author S. Mann. Wristworn SWIM makes visible the otherwise invisible electromagnetic radio waves from a smartphone (heterodyned 4x/8x as if 20,000MCPS). Right: array of LEDs on circuitboard made in collaboration with Sarang Nerkar. We find a listening device concealed inside a toy stuffed animal (Okapi). Visualized quantities are the real part of measured veillance wave functions. Magnitudes of these indicate relative veillance probability functions.

embodiment an X-Y plotter is connected to the real and imaginary (in-phase and quadrature) components of the author’s special flatband lock-in amplifier and pushed through space to trace out a complex waveform in 3D while a light bulb is attached where the pen normally would go on the plotter.

More recently, Mann and his students reproduced this result using a spinning SWIM on a sliderail to reproduce gravitational waves – making visible an otherwise hidden world of physics. See Fig. 13.

ARbotics (AR robotics) can also be applied to vision (Fig. 14). Here we map out the magnitude of the
Figure 12. Left: Sliderail SWIM to teach veillance wave principles. A speaker emitting a 10050 cycles/sec tone. The microphone’s metawave has 11 cycles in a 15 inch run. Teaching speed-of-sound calculation: 15 in. * 10050 cycles/sec / 11 cyles = 13704.54... in./sec. At 25deg. C, theoretical speed of sound = 346.23 m/s (0.5% measurement err.). The real part of the veillance wavefunction is shown but SWIM can also display magnitude (steady increase toward microphone). Right: “Bugbot” (bug-sweeping robot) finds live microphone hidden in bookshelf and visualizes its veillance waves in a 7-dimensional (3 spatial + RGB color +time) spacetime continuum. (Green=strongest; redshift=toward; blueshift=away).

Figure 13. Complex-valued “gravlet” wavefunction visualized on a robotic SWIM that spins while moving back and forth. Data[1] from LIGO[2] was used with its Hilbert transform, noting the result is a chirplet[44, 45, 46, 47, 9, 53, 23, 10] (“gravitational signal” rather than “gravitational wave”). SWIM explores periodic realtime data at any scale from atomic to cosmic, as well as displays arbitrary data.

metawave function, where the phase can be estimated using Phase Retrieval via Wirtinger Flow [11].

4. Sparsity in the Spacetime Continuum

In conclusion, Metaveillance and Veillance Wavefunctions show great promise as new tools for understanding the complex world where surveillance meets moving cameras (wearables, drones, etc.). Further research is required in the area of Compressed Sensing [8, 18] to fully utilize this work, e.g. to build completely filled-in high-dimensional spacetime maps of Veillance (“Compressed Metasensing”).

Moreover, just as oscillography was the predecessor to modern television, the alethioscope and robotic SWIM (“ARbotics”) could be the future that replaces television, VR, and AR.

Finally, surveillance, AI, and security are half-truths without sousveillance, HI, and suicurity (self care). To write the veillance/cyborg code-of-ethics we need to fully understand all veillances and how they interplay. Metaveillance gives us the tools to accomplish this understanding, in a multi,cross,inter/intra,trans,metadisciplinary/passionary mix of design, art, science, technology, engineering, and mathematics.
References


