

## Special Section on Wavelet Applications

**Harold Szu**

Naval Surface Warfare Center  
Dahlgren Code B40  
17320 Dahlgren Road  
Dahlgren, Virginia 22448-5100  
E-mail: [hszu@nswc.navy.mil](mailto:hszu@nswc.navy.mil)

In covering a field such as video compression, it would be logical that we begin with the compression lessons learned in (1) and (2) and then define wavelets in (3). Afterwards, we can better delineate the contribution of wavelets to the image and video compression.

(1) Syntactic and Semantic Singularity Maps: Lossless data compression is at most a few factor of 2 and is not addressed here. Image compression is at a factor of several hundreds and is generally lossy. Therefore, a better compression strategy must begin with an image preprocessing that preserves image fidelity. As opposed to working directly at the traditional pixel representation domain, preprocessing creates an intermediate representation, edge maps, or, in general, the singularity map defined by points, steps and tent top edges, etc., where the values of  $n$ th derivatives obtained over pixels are not unique at different approaches. Then we can talk about lossy compression to either preserve the syntactic geometry of the singularity map or the semantic ATR (aided target recognition) map.

(2) Temporal Redundancy Reduction: Video compression strategy requires a tradeoff of spatial redundancy reduction with temporal redundancy reduction. We do not change the common sense strategy of video compression, established ever since the satellite surveillance days, that only the change needs to be sent. However, if a new transform can preserve the locality better, then we should not try to compress too much of each individual frame; rather, we should keep some redundancy for a more robust tracking of the correspondence between singularity maps, ensuring that change is more efficiently tracked. Furthermore, as we

work with the reduced and intermediate singularity maps rather than with the full pixel map, we become even more efficient. In other words, the tracking of the singularity throughout the video sequence is efficient and cost effective, and the mapping between two representations is kept and sent through the channel for syntactic-preserving video compression.

(3) Wavelet Transforms (WT): Wavelets have become popular in the last decade due to the freedom they have afforded in choosing an efficient representation for any class of local wideband transient signals (WT). This is possible because WT is defined as a bank of self-matched filters at a constant relative bandwidth  $\Delta f$  at the central frequency  $f$  (the so-called constant fidelity  $Q = \Delta f / f$  representation). One can choose any appropriate transform kernel called a mother wavelet satisfying a mild admissibility condition: namely, a bounded total power spectral density when it is inversely linear frequency-weighted. This kernel can be anything other than the periodic sinusoidal basis.

To harvest the wavelet compression turns out to be not as simple and straightforward as some of us may have experienced as simple tool users. If the signal-processing world is perfect in the sense of being linear and noiseless, then any complete ortho-normal (CON) basis including periodic Fourier basis will do the job. However, it is not, and the freedom of choosing a transform kernel in general other than the periodic sinusoidal basis, e.g., a general set of localized bases called wavelets, becomes important. For example, features of images and signals consist mainly of localized singularities. Further, it is well known that Fourier trans-

form (FT) spreads the singularity energy all over Fourier frequency representation. Consequently, any compression strategy based on a finite selection of FT coefficients will create infidelity: overshoot and undershoot ringing effects known as the Gibb's phenomenon. On the other hand, the wavelet kernel such as the one Alfred Haar proposed in 1910 consists of step singularity functions that can capture the class of step jumps efficiently. Recently, Ingrid Daubechies successfully generalized Haar WT to include a family of self-dual CON Wavelets. Haar WT has one unknown coefficient (1, -1), when it conserves the zeroth moment of signal amplitude (namely, a zero area of wavelet). Daubechies' WT filters have more unknowns filter coefficients, 4, 6, 8, etc., which can be progressively determined by higher and higher orders of moment-conservation. Furthermore, the whole industry of digital signal processing (DSP) has contributed significantly to the maturity of discrete WT. Particularly, DSP has the non-self-dual or Bi-ortho-normal (BON) generalization of the self-dual CON. This is possible in terms of the sub-band filter bank: quadrature mirror filter (QMF means the symmetry at  $\pi/2 = 2\pi/4$ ). QMF has symmetric filter coefficients, e.g., 1, 2, 3, 2, 1, which have only cosine FT and no sine FT, i.e., the complex exponential FT has a zero phase. Such QMF filters are most suited for the image processing that must preserve the location of the digital pixels. For example, should a filter have a nonzero phase, then its recursive iteration may result in a false phase accumulation which, when it exceeds  $2\pi$  value, results in the pixel location being mistakenly shifted by one. Also, a symmetric DWT filter bank to-

gether with a mirror-symmetric image boundary condition can minimize the aforementioned discontinuity (Gibb's phenomenon) effect occurring at the image boundary.

Recently, Szu *et al.*, in *SPIE Proceedings of Wavelet Applications Conferences* held in Orlando in 1996, 1997, and 1998, have demonstrated an adaptive superposition of Haar wavelets as non-self-dual and symmetric DWT, called super-Haar, which can match singular signal classes better with few terms in capturing the signal singularity energy.

WT frees us from the bondage of conventional Fourier transforms, designed efficiently by Fourier a century ago for global periodic signals, but we have used it and abused it since. For pedagogical reasons, I have used the identical acronym WT for representing both the wideband transients and wavelet transform itself, so that we know for certain when to use WT; otherwise, WT should not be used for narrow-band stationary and periodic signals.

There have been six annual SPIE Wavelet Applications Conferences at Orlando in the last six years in April and two special sections in *Optical Engineering* (September 1992 and July 1994). This special section of the *Journal of Electronic Imaging* is confined to wavelet video compression related issues. This is partly due to the success of the recent FBI fingerprint WT compression program AFIS (Automatic Fingerprint Identification System) for wireless single frame of image compression and transmission. The next challenge seems to be wireless video transmission for man-in-the-loop applications. In this regard, the Congressional Mandate to the 1997 Department of Defense funding is called

WaveNet, combining wavelet transform technology and perceptual neural net. The goal of WaveNet is to reduce the industry risk of adopting WT for information processing to supplement the standard discrete cosine transform (DCT)-related efforts such as the Joint Photographic Expert Group (JPEG) and the Moving Photographic Expert Group (MPEG). Accepting that challenge, we must approximate the WT real number mathematics to integer and binary representations in order to design and build wavelet chip sets for image/video compression and wireless transmission through the Army Single Channel Ground-Air Radio System (SINCGARS, VHF at 16 kbits per second bandwidth) for man-in-loop digital battlefield applications. In short, WaveNet combines wavelet transform technology and perceptual neural net to squeeze an elephant video through a SINCGARS mouse hole while retaining the visual integrity of the elephant.

This task is possible because we have achieved a proper tradeoff between spatial redundancy reduction and temporal redundancy reduction using both continuous WT and discrete WT in tandem. This is demonstrated to be good for human spatiotemporal contrast sensitivity response function, defined as CWT in "Video Johnson-like criterion for moving vehicle identification" by Szu *et al.* We know that video transmission at the camera frame rate of 30 Hz is an overkill for digital TV but not so for analog TV, minimizing the movie flicking and jittering effect. In the computer vision community, another challenge is to solve the so-called Stereovision Correspondence Principle from one image to another. This challenge happens also in a sequence of video. This is overcome by the reduced intermediate representation—the sin-

gularity map derived in terms of Mexican hat wavelet, which adequately represents the human visual system and is conveniently implemented in terms of the discrete artificial neural networks. If we do not overly compress and distort the singularity map of a single frame extracted by CWT, we can achieve a better alignment of spatial singularities in the so-called optical flow of edge maps (not the inefficient pixel optical flow).

Eight papers are presented in this special section. The introductory paper by Daubechies points to the future of wavelet applications. Kuon, Venkataraman, and Nasrabadi review the traditional video compressions by vector quantization; Grim and Szu evaluate the video quality measure for control, and Mujica *et al.* introduce an object tracking algorithm. Hsu and Szu describe the WaveNet brassboard; Noel and Szu introduce wavelet applications for smarter sensors. Szu, Cox, and Hsu discuss video Johnson-like performance criterion, and Szu and Do-Duc present image error-resilient and video error-concealment codes.

Looking at the future, I would encourage a proper combination of both syntactic-preserving and semantic-preserving compression strategy which shall be optimal for wide-area surveillance ATR. More efficient motion estimation of DWT in terms of singularity maps may be developing with the help of the second generation wavelets as advocated by Daubechies, among other pioneers, in this special section.

**Harold Szu:** Biography and photograph appear with the paper "Video compression quality metrics correlation with aided target recognition (ATR) applications" in this issue.