WTSOM: A ROBUST STILL-IMAGE TRANSMISSION SCHEME SUITABLE FOR FADING CHANNELS

Hervé Boeglen[1], Christian Chatellier[2], Olivier Haeberle[1] , Pascal Bourdon[2] and Christian Olivier[2]

[1]Laboratory MIPS, IUT Mulhouse, 61 rue Albert Camus, 68093 Mulhouse Cedex France
[2]Laboratory SIC, Téléport 2, Bd Marie et Pierre Curie, 86962 Futuroscope Cedex France

ABSTRACT

The aim of this work is to present a robust still-image transmission scheme and to show its performance when used on a mobile fading channel. After a review of the main physical effects induced by mobile fading channels and how to model them, we describe the WTSOM (Wavelet Transform Self-Organizing Map) image transmission scheme. In section 2, we describe a two-stage strategy in order to use WTSOM on fading channels. The first stage consists in classical Error Control Coding (ECC) whereas the second stage is based on an inpainting model for image reconstruction. The combination of these techniques yields a very good visual quality of the received images over a wide range of Doppler frequencies.

1. THE CHALLENGE OF A MOBILE FADING CHANNEL

1.1. The physics

In mobile radio communications, the emitted electromagnetic waves often do not reach the receiving antenna directly due to obstacles blocking the line-of-sight (LOS) path. In fact, the receiving waves are a superposition of waves coming from all directions due to reflection, diffraction and scattering caused by buildings, trees and other obstacles. This effect is known as *multipath propagation* [1]. As a result, the received signal consists of an infinite sum of attenuated, delayed, and phase-shifted replicas of the transmitted signal, each influencing each other. Depending of the phase of each partial wave, the superposition can be constructive or destructive. Apart from that, when transmitting digital signals, the form of the transmitted impulse can be distorted during transmission and often several individually distinguishable impulses occur at the receiver due to multipath propagation. This effect is called *impulse dispersion*. The value of the impulse dispersion depends on the propagation delay differences and the amplitude relations of the partial waves. The consequence of this effect is a distortion of the frequency response characteristic of the transmitted signal.
Besides the multipath propagation, also the *Doppler effect* [1] has a negative influence on the transmission characteristics of the mobile radio channel. Because of the mobility of the receiver, the Doppler effect causes a frequency shift of each of the partial waves. As a result, the spectrum of the transmitted signal undergoes a frequency expansion during transmission. This effect is called *frequency dispersion*. The value of the frequency dispersion mainly depends on the maximum Doppler frequency and the amplitudes of the received partial waves. In the time domain, the Doppler effect implicates that the impulse response of the channel becomes time-variant.
Multipath propagation in connection with the movement of the receiver and/or the transmitter leads to drastic and random fluctuations of the received signal. Fades of 30 to 40dB and more below the mean value of the received signal level can occur several times per second as can be seen on the following figure:

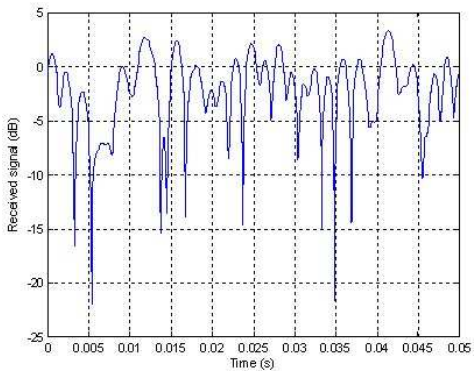


Fig.1: Fading channel received signal amplitude

In digital data transmission, the momentary fading of the received signal causes *burst errors* i.e. errors with strong connections with each others. Therefore, a fading interval produces burst errors, where the burst length is determined by the duration of the fading. Corresponding to this a connecting interval produces a bit sequence almost free of errors. From this, one can easily understand that developing efficient error protection strategies for digital transmission systems implies a good knowledge of the fading channel statistics.

1.2. Statistical fading channel models

For the choice of a transmission technique and the design of digital receivers it is vital to be able to simulate the effects of a fading channel. Since the pioneering work of Clarke [2] in the late sixties there has been a lot of theoretical work completed by measurement campaigns showing that in the case of non frequency-selective channels, the fluctuations of

the received signal can be modelled by multiplying the transmitted signal with an appropriate stochastic model process. In urban and suburban areas, where there is no LOS component, the transmission can be modelled by a Rayleigh process. In rural regions, however, the LOS component is often part of the received signal, so that the Rice process is more appropriate. These two processes can be grouped in the following expression for the amplitude z of the received signal:

$$p_Z(z) = \frac{2z(K+1)}{\mathrm{Pr}} \exp\left[-K - \frac{(K+1)z^2}{\mathrm{Pr}}\right] I_0\left(2z\sqrt{\frac{K(K+1)}{\mathrm{Pr}}}\right), z \geq 0 \quad (1)$$

where Pr is the power of the received signal, $I_0(.)$ is the 0^{th}-order modified Bessel function of the first kind and K is the so-called Rice factor. The value of this K-factor indicates the importance of fading. If K = 0, the fading is maximum and equation (1) models a Rayleigh process. For K > 0 we get a Rice process and for K = ∞ there is no fading. For the phase distribution one can show that the uniform distribution is accurate.

Besides the probability density of the amplitude and the phase, the shape of the Power Spectral Density (PSD) of the previously presented processes models the Doppler effect of the channel. Measurements have shown that it can be described by the so-called Jakes PSD (although it was first derived by Clarke). Figure 2 shows a Jakes PSD and the corresponding autocorrelation function.

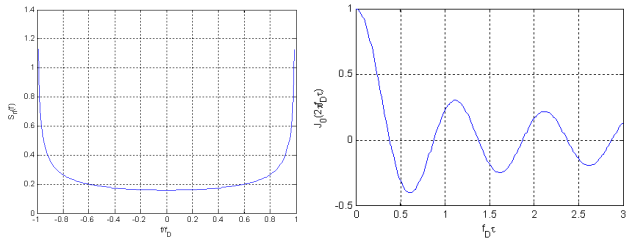


Fig.2: Jakes PSD and autocorrelation function

For the modelling of frequency-selective channels, one can use the Wide Sense Stationary Uncorrelated Scattering (WSSUS) assumption introduced by Bello [3]. In that case, the channel impulse response is a superposition of several independent delayed non-frequency selective impulse responses (individual paths).

When it comes into practical implementation, a wide range of standardized statistical channel models, each suitable for a particular application (e.g. indoor WLAN, outdoor mobile telephony) exists. These models use the WSSUS assumption. One well-known example is the set of COST207 channel models established for the deployment of second generation mobile telephony [4]. In the following, we will use the COST207 TU (Typical Urban) channel model.

2. TRANSMITTING COMPRESSED IMAGES ON NOISY CHANNELS: THE WTSOM SCHEME

The motivation of our work comes from the fact that the transmission of JPEG or JPEG2000 images over insecure, noisy channels even with low Bit Error Rates (BER) can lead to severe consequences on decoded visual information. The reason for such high error sensitivity is the adoption of Variable Length Coding (VLC) schemes. Although VLC are known to achieve better compression rates than Fixed-Length Codes (FLC), they also introduce dependence among codewords. Figure 3 below shows the result of the transmission of a JPEG2000 image over a Rayleigh channel with an SNR of 26dB using a Star 16QAM modulation (header not corrupted):

a) Transmitted image
PSNR = 30.35dB

b) Received image
BER=2.5e-3, PSNR = 13.1dB

Fig.3: JPEG 2000 image transmission

The WTSOM image coding system combines two well-known techniques in the image processing community: Wavelet Transformation (WT) and Vector Quantization (VQ). The first stage of our compression scheme consists in applying a wavelet decomposition up to level 3 using a Daubechies 9/7 filter bank. After this operation, we discard the less important subbands keeping only the most significant ones (see Figure 4).

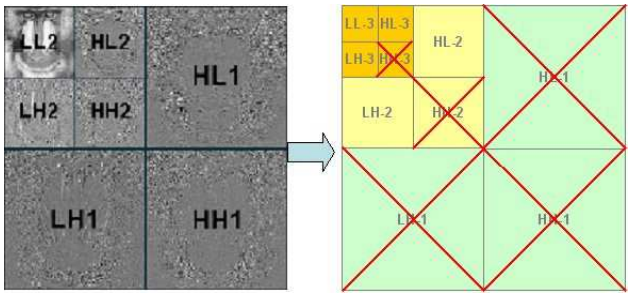


Fig.4: WTSOM Wavelet decomposition

The second stage is to apply VQ on the remaining subbands of the wavelet decomposition. To do this, we use Kohonen's Self-Organizing Map (SOM) algorithm, because of its interesting property of auto-organization: close vectors in terms of Euclidian distance correspond to close indices in terms of location in the self-organized codebook. This makes it really appropriate for transmission, since we can fit such a codebook on digital modulation points, and minimize noise-related distortion of codebook elements [5]. The corresponding transmission chain is represented on Figure 5. Each codebook contains 256 vectors and only the vectors

indices are transmitted. Thanks to these techniques we obtain a compression rate in the order of 25 for grayscale images.

The first WTSOM scheme used to transmit only grayscale images so we decided to go a step further by applying the techniques presented previously on a color image. Taking into account works showing that human beings are more sensitive to fast luminosity changes than to fast color and saturation changes, we decided to adopt a YC_bC_r color space rather than the traditional RGB color space. This representation has also the advantage of decorrelating the three color planes of the image. Finally, the preceding techniques are applied on the three color planes of the image except that in the case of the C_b and C_r planes we only keep the LL3 subband. This yields a compression rate in the order of 42.

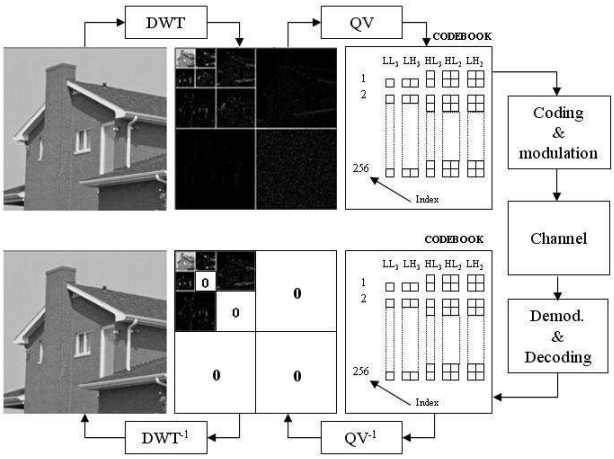


Fig.5: WTSOM transmission chain

3. TRANSMITTING WTSOM IMAGES ON FADING CHANNELS: A TWO-STEP STRATEGY

Although WTSOM's very good performance has been demonstrated even on very noisy ionospheric channels [6], a more challenging task appears when trying to transmit images over mobile fading channels. In the following we describe a two-stage strategy. The first stage is a classical, yet simple, Unequal Error Protection (UEP) ECC scheme. The second stage deals with a post processing inpainting reconstruction algorithm. The two stages can be combined to give better results.

3.1. Star 16QAM and UEP EEC

Differential modulation/demodulation techniques have proved to be very robust due to the fact that they can mitigate the phase distortions caused by fading. Moreover, as no channel estimation is needed, the system complexity is reduced significantly.

In order to keep the same bandwidth efficiency as in the AWGN case [6], we use a 16-point constellation that may be differentially encoded and decoded. This scheme is known as Star 16QAM [7] and uses a combination of independent 2DASK and 8DPSK. Three of the four bits are modulated by a Gray encoded 8DPSK. The remaining bit is amplitude modulated by a 2DASK scheme: a "0" doesn't change the amplitude while a "1" changes it. Figure 6 shows the signal constellation of the 16DASPK modulation. For our application we have optimised the Star 16QAM key parameters i.e. the ring ratio and the bit mapping. More details about this can be found in [8].

In the WTSOM scheme, the quantized low-frequency subbands (LL3) contain the most important information about the image. This is the reason why we apply a (255,205) Reed-Solomon (RS) block code on these subbands. In the case of slow fading, this ECC can be completed by a symbol block interleaver, which spreads out burst errors over a large number of symbols so as not to exceed the RS code burst error correction capability. Figure 7 gives an example of a transmission using these techniques where the right image shows artefacts due to uncorrectable errors.

Figure 8 summarizes the behaviour of our compression scheme in terms of Peak Signal-to-Noise Ratio (PSNR) over three decades of the normalized Doppler frequency f_DT_S (f_D = Doppler frequency in Herz and T_S = symbol time in seconds). One can notice that the average PSNR (blue curve; averaged over 50 transmissions per f_DT_S value) of the transmitted "fruits" image is quite insensitive to the Doppler frequency. The red curve (second from the top) represents the best case PSNR over 50 transmissions whereas the black curve represents the worst case. One can also notice that the difference between these two curves is larger for low f_DT_S values. This is a classical characteristic of fading channels: low f_DT_S values imply longer fade intervals thus producing long error bursts but also long connecting intervals meaning that one can get no errors at all during an image transmission.

Another interesting feature to underline is WTSOM's very good tolerance to high BER values (between 2 and 3%). These high BER values are to be compared to those required by JPEG or JPEG2000 images which have to be in the order of 10^{-3}.

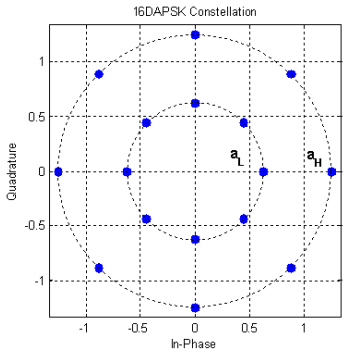


Fig.6: Star 16QAM constellation

In order to be able to use our reconstruction algorithm (described in the next section), we calculate a Cyclic Redundancy Check (CRC) on each of the indices of the

three low frequency subbands. The parity words obtained are appended to the image data and are protected by a (255,187) RS code. The combination of these techniques yields a compression rate of Rc = 29 for color images.

a) Transmitted WTSOM image, PSNR=26.33dB

b) Received WTSOM image after ECC correction $f_D T_s$ =0.001, SNR=17dB, BER=2.2%, PSNR=22.8dB

Fig. 7: Example of a WTSOM transmission

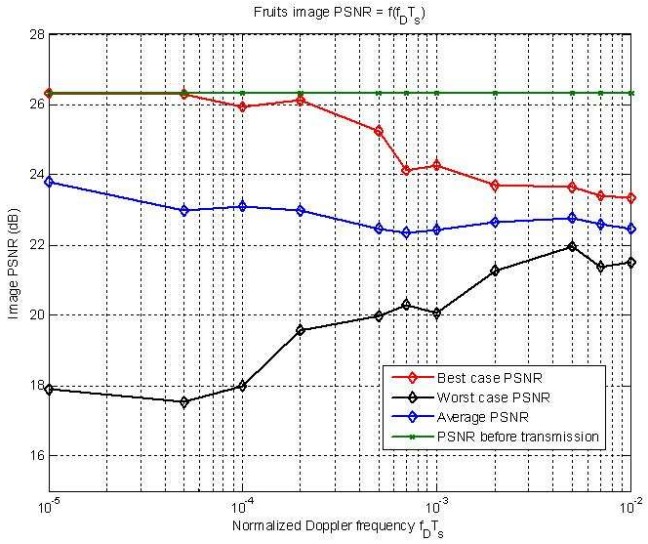


Fig. 8: WTSOM "Fruits" image PSNR versus $f_D T_S$

3.2. Inpainting reconstruction algorithm

As pointed out in the previous sections of this document, one major advantage of the combined WT/VQ approach in the WTSOM algorithm is the fact that key elements of image geometry are preserved throughout the compression process. Not only does this feature allow us to reduce the impact of transmission errors on original/received image distortions, it also provides us with a file structure being actually able to handle image processing-based error concealment techniques.

As image processing applications require higher levels of reliability and efficiency, mathematical image processing, especially Partial Differential Equations (PDE)-based approaches, have become very popular over the last two decades. Let $I(x,y)=I$ be a still, gray-level image, represented by a function of $\Omega \subset \Re^2 \rightarrow \Re$ that associates to a pixel $(x,y) \in \Omega$ its gray-level value $I(x,y)$; Ω is the support of the image. In [9], Alvarez *et al.* show that Perona-Malik's well known anisotropic diffusion model for edge-preserving filtering [10] can also be written as follows:

$$\begin{cases} \dfrac{\partial I(x,y,t)}{\partial t} = c_\xi I_{\xi\xi} + c_\eta I_{\eta\eta} \\ I(x,y,0) = I_0(x,y) \end{cases} \quad (2)$$

where I_0 denotes a noisy version of I, $I_{\xi\xi}$ and $I_{\eta\eta}$ the second derivatives of I in orthogonal directions ξ and η, and η and ξ the gradient and orthogonal-gradient directions.

For c_ξ=1 and c_η=1, this equation actually is the heat equation, hence the concept of *diffusion*: each gray-level intensity $I(x,y)$ happens to diffuse over neighbouring pixels $(x \pm \partial x, y \pm \partial y)$ for a given time t, creating a (Gaussian) smoothing effect on the overall image. Knowing that unit vector ξ is actually tangent to the edge curve for every point of the image, as shown on Figure 9, we can easily understand how by describing a c_ξ-weighted (with c_η=0), tangential diffusion of $I(x,y)$ along local edges, Alvarez *et al.*'s formulation can allow us to interpolate existing structures into missing areas.

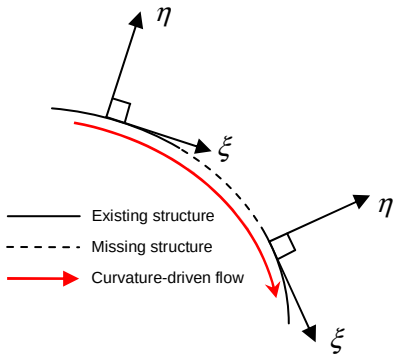


Fig. 9: Interpolation using geometry-driven flows

In [11] a 2nd order digital inpainting model for error concealment in color JPEG images was proposed. This model is based on Tschumperlé's tensor-based diffusion model, and can be written as:

$$\begin{cases} \dfrac{\partial I_i}{\partial t} = \text{trace}(\mathbf{T}\mathbf{H}_i) + \lambda_e (I_i - I_{0_i}) \\ I_i(x,y,0) = I_{0_i}(x,y) \end{cases} \text{ for } i=1,2,...,m \quad (3)$$

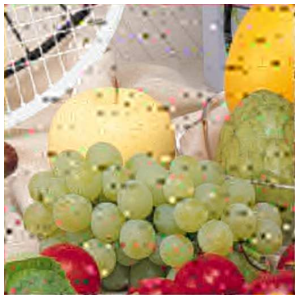
where $\mathbf{I}=[I_1 I_2 .. I_m]^T$ denotes a vector-valued function (*e.g.* a RGB image with m=3), $\mathbf{H}_i$ the Hessian matrix of I_i, $\lambda_e=\lambda(1-\chi_D)$ an extended Lagrange multiplier using the characteristic function (mask) $\chi_D(x,y)$ of the inpainting domain D (χ_D=1 if inside, 0 if not), and where matrix $\mathbf{T}$ is referred to as a *diffusion tensor*, and defines a direction and

strength for the smoothing effect based on local geometry. More details about this inpainting model can be found in [11].

Using the PDE inpainting model described previously we apply the following algorithm to the received image:

- An error map $\chi_D(x,y)$ is built from CRC verification,
- First inpainting phase based on Tschumperlé's diffusion model is applied on all LL3 streams,
- WT level LL2 is reconstructed using the inpainted LL3, the error map $\chi_D(x,y)$ is up-sampled to match the LL2 size, the second inpainting phase is applied to LL2,
- Previous step is repeated up to the last level which represents the decoded and error-concealed image.

Figure 10 shows an example of the reconstruction process applied on a distorted image where the PSNR gain reaches 2.24dB. Figure 11 shows the PSNR gain of our inpainting model versus f_DT_S when applied to the most degraded images of our transmission series. This gain is constant over the three decades of f_DT_S and is in the order of 2.8dB.



a) Received WTSOM image after ECC correction
f_DT_s=0.0005, SNR=17dB
BER=2.8%, PSNR=22.59dB

b) Received WTSOM image after inpainting
PSNR = 24.83dB

Fig. 10: WTSOM transmission with reconstruction

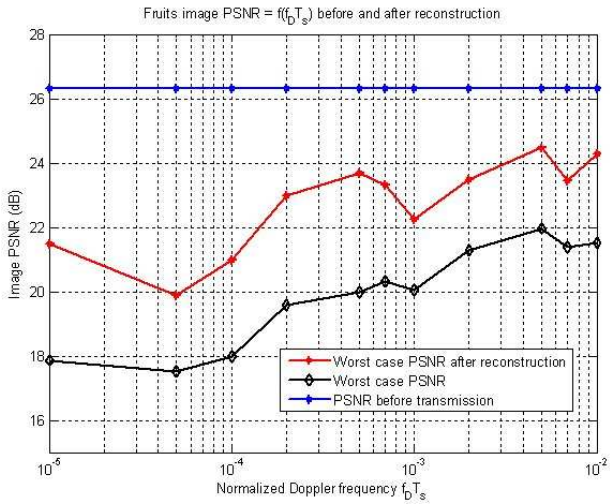


Fig. 11: WTSOM "Fruits" image PSNR versus f_DT_S

4. CONCLUSION

We have presented a robust image compression scheme working on fading channels with BER as high a 2.5%. Our results show a constant visual quality over three decades of the normalized Doppler frequency f_DTs. This image coding system works in two stages which can be combined to give better results. The first stage is a classical ECC scheme, yet simple and efficient, combining a bandwidth-efficient differential modulation and an RS block code. The second stage is based on a digital inpainting algorithm that is able to get rid of most visual artefacts induced by the channel yielding a minimum PSNR gain of 2dB. This algorithm can therefore be a good alternative to high complexity channel coding strategies.

5. REFERENCES

[1] M. Paetzold, *Mobile fading channels*, Wiley, 2002.

[2] R. H. Clarke, "A statistical theory of mobile-radio reception", *Bell Syst. Tech. Journal*, vol 47, pp. 957-1000, July/Aug. 1968.

[3] P.A. Bello, "Characterization of randomly time-variant linear channels", *IEEE Trans. Comm. Syst.*, Bd. CS-11, no. 4, 1963, pp. 360-393.

[4] COST 207, "Digital land mobile radio communications", *Office for Official Publications of the European Communities*, Final Report, Luxembourg, 1989.

[5] O. Aitsab, R. Pyndiah, B. Solaiman, "Joint optimization of multidimensional SOFM codebooks with QAM modulations for vector quantized image transmission", *3rd International workshop in signal/image processing, pp. 3-6, Manchester, UK,* November 1996.

[6] C. Chatellier, P. Bourdon, B. Souhard, C. Olivier, "An efficient joint source channel coding scheme for image transmission through the ionospheric channel", *European Conference Propagation and Systems, Brest France,* March 2005.

[7] Svensson, N.A.B. On differentially encoded star 16QAM with differential detection and diversity *Vehicular Technology, IEEE Transactions on* Volume 44, Issue 3, Aug. 1995 Page(s): 586-593

[8] H. Boeglen, C. Chatellier, O. Haeberle, "On the robustness of a joint source-channel coding scheme for image transmission over non frequency selective Rayleigh fading channels", *ICTTA'06: 2nd IEEE International Conference on Information & Communication Technologies: From theory to applications*, Damascus Syria, April 2006.

[9] L. Alvarez, P-L. Lions, J-M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion II", *SIAM Journal on Numerical Analysis 29 (3)*, pp. 845-866, 1992.

[10] P. Perona, J. Malik, "Scale-space and edge detection using anisotropic diffusion", *IEEE transactions on Pattern Analysis and Machine Intelligence 12 (7)*, pp. 629-639, 1990.

[11] P. Bourdon, C. Chatellier, B. Augereau, and C. Olivier, "A multi-resolution, geometry-driven error concealment method for corrupted JPEG color images", *EURASIP Signal processing: Image Communication 20 (7)*, pp. 681-694, August 2005.