NTNU

Norwegian University of
Science and Technology

Power Efficient Video Communication for Mini Helicopter

Ola Naalsund Ingvaldsen

Master of Science in Communication Technology
Submission date: June 2010
Supervisor: Tor Audun Ramstad, IET

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

Problem Description

To prolong the operating time of a remotely controlled micro helicopter, the processing must be kept simple and the video communication power efficient. We will assume that the processing of the video takes considerable less energy than the actual signal transfer. Thus, compression is an important factor to reduce the transfer rate. Further we assume there are negligible power limitations on the feedback channel. We can then use iterative transmissions with a feedback channel, where the signal to noise ratio on the feedback channel is much higher than that of the feed forward. When operating, the transmission will experience Rice or Rayleigh fading depending on whether or not a direct line of sight is available between the operator and the helicopter. To reduce the problems associated with fading, channel diversity should be obtained in some way.

The task is to simulate a complete system with video compression and transmission. The compression should be based on a three dimensional DPCM-codec with motion compensation. Realistic channel estimation must be included. The carrier frequency should be chosen based on the system specifications and what the telecommunications authorities permit.

Assignment given: 18. January 2010
Supervisor: Tor Audun Ramstad, IET

Abstract

In this thesis, a video communication system for use in a mini helicopter is reviewed. The transmitter is located in a small battery powered mini helicopter, weighing about 15 grams, and the receiver is the helicopter's remote control. The operator controls the helicopter only based on the video feed sent from the helicopter, hence it is critical that the delay is kept to a minimum. Due to the helicopter's small size, the energy available is very limited, and both coding and transmission should be efficient and have low complexity. Since processing takes considerable less energy than transmission, compressing the video is an important factor to reduce the needed transmission rate. Video compression is done using three-dimensional DPCM with motion compensation. The prediction is done around the channel noise; hence, the system uses joint source channel coding. The analog system utilizes the feedback channel already needed for closed loop DPCM to transmit every sample in an iterative fashion which expands the bandwidth to reduce transmission power. The simulations shows that for an ideal feedback channel and with a Gaussian distributed input, the system acts according to OPTA.

The system is simulated in different environments and components are added to the system to make the simulations more realistic and the system more robust. The sampling period of the system with a bandwidth expansion of three is 288 ns, and with a carrier frequency of 2.39 GHz, it is reasoned that this would give flat fading in an indoor office environment.

When flat fading is introduced, channel estimation and diversity are implemented to reduce the degradation caused by fading, and by adding three-dimensional DPCM, the system becomes very robust against the fast channel fluctuations caused by multipath fading. Further, in fading the iterative system experiences graceful degradation, meaning the system does not break down below a certain power but has a steady decrease in performance for decreasing E_p/N_0. The system shows good results (received PSNR greater than 40 dB) for an average E_p/N_0 of 3.5 dB in the receiver, averaged over all test inputs.

Simulations shows a substantial gain in lowering of the variance of the transmitted signal by video compression. For high E_p/N_0, the transmission power can be reduced by 17 dB to yield the same result as with no compression.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AWGN – Additive White Gaussian Noise

BS – Base Station or remote control

CSNR – Channel Signal to Noise Ratio

CSI – Channel Side Information

DPCM – Differential Pulse Code Modulation

EGC – Equal Gain Combining

FIR – Finite Impulse Response

FPS – Frames Per Second

HC – Helicopter

ISI – Inter-Symbol-Interference

JSCC – Joint Source Channel Coding

MSE – Mean Squared Error

MRC – Maximal Ratio Combining

MV – Motion Vector

OPTA – Optimal Performance Theoretically Attainable

PDF – Probability Density Function

PDP – Power Delay Profile

PSNR – Peak Signal to Noise Ratio

QVGA – Quarter VGA (320×240 pixels)

RMS – Root Mean Square

SNR – Signal to Noise Ratio

TV1 – Test Video 1

TV2 – Test Video 2

TV3 – Test Video 3

UAV – Unmanned Air Vehicle

CHAPTER 1

INTRODUCTION

This thesis reviews a complete system used for transmitting a wireless video feed. The transmitter is located in a small battery powered mini helicopter, weighing about 15 grams, and the receiver is the helicopter's remote control. The operator controls the helicopter only based on the video feed sent from the helicopter, hence it is critical that the delay is kept to a minimum. Further, since the weight of the helicopter is only about 15 grams, it has a very small battery where most of the power is used for thrust, implying that transmission complexity and power should be kept as low as possible. Since the communication is expected to operate at variable SNRs, it is important that the system has graceful degradation, meaning it does not break down but shows steady decreasing performance for decreasing input power. This motivates the use of a low-complexity joint source channel coding transmission scheme, that utilize bandwidth expansion to reduce the energy consumption.

This thesis will review a low complexity video communication system with time discrete continuous amplitude for the use described above. It will simulate the system in a wireless environment with video compression, channel estimation and multiple antennas on the remote control and review its performance. The video feed has a frame resolution of 320×240 pixels and transmits 15 frames per second.

The communication and processing of video and images differ from normal signal processing in several ways. Images of natural events, in contrast to text and numbers, can be compressed with loss of measured information, but with low visual degradation. Thus, in transmission of images, distortion should not

Figure 1.1: *A prototype of the helicopter that the communication system in this thesis has been designed for. Image is taken from* `www.proxdynamics.com`

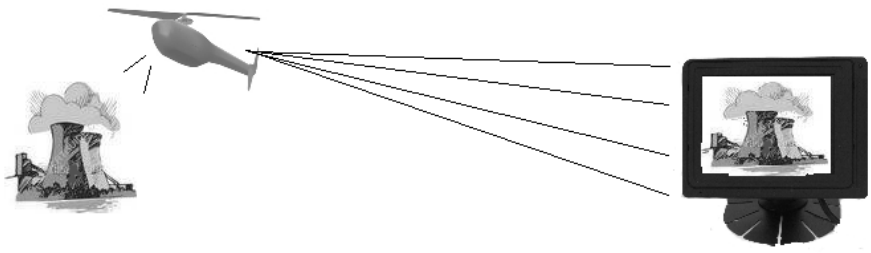


Figure 1.2: *A typical use of the helicopter can be overseeing hazard zones where it is dangerous or inconvenient for people to get. The transmission of the video stream is the part this thesis will look into*

be measured in bit errors but as the observer's ability to see the degradation. Since humans cannot see small degradations in images, a natural step towards minimizing resources is image compression.

Shannon proved that when optimizing the transmission of a source over a channel, the source and channel coding could always be split into two independent entities. Thus optimizing the source coding can be done without considering the channel and vice versa. However, the separation theorem requires infinite complexity and delay to obtain full separability. This motivates for the investigation of systems using joint source channel coding (JSCC), where the signal is coded in accordance with both channel and source simultaneously.

According to Shannon[21], there are two ways to increase capacity over a channel: Increasing power or increasing bandwidth. For a system, transmitting at high rates with limited energy, this motivates for the investigation of bandwidth expanding transmission schemes.

The problem to be addressed in this thesis is the aspect of image communication from a mini helicopter. The helicopter has limited energy, as it is battery driven, and thus low energy consumption is of the most importance. Due to secrecy, the current video transmission scheme used in a preliminary version of the helicopter is not known. Therefore, a direct comparison in energy consumption cannot be done in this paper. It is assumed that the company can do this when they receive a copy of this thesis.

1.1 Scenarios

The goal of the helicopter is that it can be used in a number of different environments and still have a reliable communication with the base station over an operating time of about 30 minutes[1]. It is beyond the scope of this thesis to test for all different wireless environments so this section will describe one typical operating scenario for the helicopter, which will be referred throughout the thesis.

Scenario 1
The helicopter is operating inside a typical office building, with many small rooms, and there exists no line of sight between the helicopter and its operator. Therefore, the signal is based on reflections and signal components going through walls or bending around corners. The helicopter is moving at less than 5 m/s and the straight-line distance between helicopter and operator is about 100 meters.

Even though the scenario will be referred, much of the theory and the results obtained are general and will hold for any scenario.

1.2 This Thesis

As stated above, this thesis deals with the problem of energy and complexity limitations in a video transmission system. It is a continuation of an earlier work done by the author together with Brage Høyland Ellingsæter[7]. The

[1] See http://www.proxdynamics.com/

previous work was a project that considered two different systems for image communications for an unmanned air vehicle(UAV), where one of the systems was the same as will be reviewed in this thesis, only now in more detail. The other system that was considered used closed loop DPCM around the quantizer and transmitted using pulse position modulation (PPM). It was shown that the analog system outperformed PPM concerning performance, bandwidth and robustness. Fading, diversity, channel estimation and video compression are typical topics that will be more emphasized in this work.

The thesis' main contributions are summarized in the following:

- Transmission - The modulation and actual transfer of information from helicopter to base station and back.
- Video compression - Decrease necessary transmission rate or lower total energy consumption by decreasing the amount of information to transfer.
- The wireless environment - Fading and path loss.
- Means to combat fading - Channel estimation, power allocation and introduction of channel diversity.
- Simulations - Both of the system as a whole and separate parts.

The thesis is organized as follows: Chapter 2 discusses the theory and background information behind this thesis and gives some indications as to what can theoretically be expected. Chapter 3 explains the simulations, methods applied in implementation and assumptions and approximations that are made. Chapter 4 presents the results obtained through simulations and discuss them with focus on power and complexity. Chapter 5 concludes the thesis.

CHAPTER 2
THEORY

This chapter will describe the system reviewed in this thesis. How it works and how it is expected to perform in wireless communications where power, complexity and use of bandwidth are vital. It will explain concepts and theoretical background used and referenced later. First, some mathematics related to information theory will be presented, before the transmission scheme is described. Then some aspects of transmitting over a wireless channel are reviewed and measures taken to combat the impact of the channel. At last video compression is explained.

2.1 Capacity, Rate Distortion and OPTA

Channel capacity, rate distortion and OPTA are all expressions that all have their origin in the work of Claude Shannon[21, 22]. This section will give a brief introduction to these expressions, their connection to the thesis and each other. In Shannon's groundbreaking 1948 article he laid out the basic elements of communications, including the limit for how much information that could be transferred over a noisy channel without errors. This limit is called the channel capacity and for an additive white Gaussian noise (AWGN) channel, he obtained the following:[21]

$$C = B \log \left(1 + \frac{P}{N_0 B} \right) \tag{2.1}$$

where P is signal power, B is the single sided baseband bandwidth and $N_0 B$ is the noise power. If the base of the logarithm is 2, then the capacity is measured

in bits per second. It is interesting to note that the maximum capacity that can be achieved by a power-limited system with infinite bandwidth is:[16]

$$\lim_{B\to\infty} C = \lim_{B\to\infty} \left(1 + \frac{P}{N_0 B}\right) = B\frac{P}{N_0 B} = \frac{P}{N_0} \tag{2.2}$$

Toby Berger defines rate distortion theory as "The theoretical discipline that treats data compression from the viewpoint of information theory"[3]. The aim of data compression is to extract the relevant information of a signal and thus remove redundancy so that a signal can be transferred at a lower rate. If information that cannot be restored is removed, distortion will be applied to the signal during compression. How much distortion to expect can be expressed as a function of the desired rate and is called the rate distortion function. Distortion is often very subjective and will be different in many applications. For instance, in image communication, the distortion should be a measure of intelligibility, meaning that distortion should be a measure of how well the receiver can interpret the received distorted signal. This is difficult to express mathematically and in this thesis, following [3] and [4], the distortion is measured as the mean squared error for simpler calculations.

By equating the desired rate-distortion function and the channel capacity, the optimal performance theoretically attainable (abbreviated OPTA) from a system with a given distortion can be obtained. The signal distortion can be found as a function of channel degradation, channel bandwidth and source bandwidth. Thus, if a scheme reaches OPTA, there is no other scheme that is better because it fully exploits the combination of source and channel.

The rate-distortion function calculates the minimum rate required to achieve a particular distortion given a distortion measure, here by mean squared error[4]. Given the transmission of an uncorrelated Gaussian source over a channel with additive white Gaussian noise, the rate-distortion function derived in [3] is given by:

$$R_s = \max\left(0, W\log_2(\frac{\sigma_X^2}{\sigma_D^2})\right) \tag{2.3}$$

Where W is the bandwidth of the source, σ_X^2 is the signal variance and σ_D^2 is the allowable distortion. If Equation 2.3 and Equation 2.1 are equated and solved with respect to SNR, the following expression for OPTA is obtained, assuming $\sigma_X^2 > \sigma_D^2$

$$\frac{\sigma_X^2}{\sigma_D^2} = \left(1 + \frac{P}{N_0 B}\right)^{\frac{B}{W}} \tag{2.4}$$

Notice that by increasing the ratio between the channel bandwidth B and the signal bandwidth W, the input power can be reduced for the given quality. From Equation 2.4, the benefit from expanding the bandwidth can be seen intuitively: If the bandwidth of the channel is larger than that of the signal, the signal bandwidth could be expanded in some fashion to reduce power. For instance, if the bandwidth is expanded by a factor of three, $B = 3W$, one can reduce the power by a factor of eight to get a $10dB$ SNR.

Equation 2.4 is only valid if the source is uncorrelated and Gaussian distributed. This is rarely the case, but the principle showed by the equation is generally valid: Bandwidth can be expanded to reduce power and still perform in an optimal manner.

2.1.1 OPTA for Correlated Gaussian Sources

Now, the transmission of correlated Gaussian sources will be reviewed. Assuming the source has a power spectral density $S_{XX}(F)$ and bandwidth W, the rate-distortion function can be found parametrically [3]. Given MSE distortion, the rate in bits per second, in terms of a parameter μ, is given by

$$R_S(\mu) = \int_0^W \max\left(0, \log_2 \frac{S_{XX}(F)}{\mu}\right) dF. \tag{2.5}$$

Intuitively, this means that only the region where $S_{XX}(F)$ is above μ is preserved. Hence below μ, the distortion is larger than the fidelity criterion, see Figure 2.1.

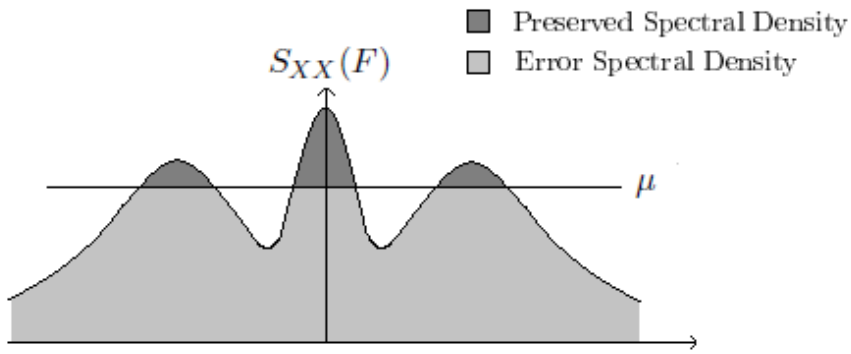


Figure 2.1: *Spectral densities of source, error and reproduction for optimum encoding, given MSE distortion*

Assuming a channel with noise power spectral density $S_{NN}(F)$, and that the channel is power constrained with bandwidth B, the channel capacity can be described parametrically in terms of a parameter θ [3]. The capacity in bits per second is given by

$$C_S(\theta) = \int_0^B \log_2 \left(1 + \frac{(\theta - S_{NN}(F))^+}{S_{NN}(F)}\right) dF \tag{2.6}$$

with power constraint

$$2\int_0^B (\theta - S_{NN}(F))^+ dF \leq P \tag{2.7}$$

where θ is maximum average power that can be sent on the channel and can be found using the water filling principle and

$$x^+ = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \tag{2.8}$$

As in the previous case of uncorrelated source and noise, OPTA can be obtained by setting the rate distortion function equal to the capacity, and solve for the obtainable SNR as a function of the CSNR. However, by further assuming the channel noise and reconstruction noise is low, meaning that $S_{NN}(F) < \theta$ and $S_{XX}(F) > \mu$ for all F, the expression reduces to

$$\int_0^W \log \left(\frac{S_{XX}(F)}{\mu}\right) dF = \int_0^B \log \left(\frac{\theta}{S_{NN}(F)}\right) dF \tag{2.9}$$

By solving for μ and θ on one side, introducing the bandwidth relation between channel and source $r = \frac{B}{W}$ and making a change of integration the following equation is obtained:

$$\log(\mu\theta) = \frac{1}{W}\int_0^W \log \left(S_{XX}(F) S_{NN}^r(rF)\right) dF \tag{2.10}$$

For Equation 2.10 to be valid for all W and with $r = 1$, the integrand must be constant so that

$$\log(\mu\theta) = \frac{1}{W}\int_0^W \log \left(S_{XX}(F) S_{NN}(F)\right) dF \tag{2.11}$$

$$= \log \left(S_{XX}(F) S_{NN}(F)\right) \int_0^W W dF \tag{2.12}$$

and

$$\mu\theta = S_{XX}(F) S_{NN}(F) \tag{2.13}$$

Thus the only case when OPTA is reached for analog transmission is when the signal and channel noise spectra are inverse functions. Even though the noise in a transmission system is often modeled by AWGN, which is often close to reality, not all sources are Gaussian. It is important to realize that transmitting some random variable $\{X(t)\}$ to the user when the average channel input power is bounded by P is most difficult when both $\{X(t)\}$ and $\{N(t)\}$ are Gaussian. This is because this maximizes the rate distortion function and minimizes the capacity. OPTA is often computed for the joint Gaussian case, simply because there are simple equations for the capacity and rate distortion function in this case, but OPTA improves in non-Gaussian cases.

2.1.2 OPTA for Flat Fading Channel

When a signal is transmitted wirelessly, the signal will be affected by fading. What fading is and why it occur can be seen in Section 2.3.3. If the transmission is performed over a flat fading channel, the channel capacity is reduced in comparison to no fading. In this section, the fading will be modeled as flat Rayleigh fading, meaning that the fading factor in the channel model in Figure 2.2 is governed by the Rayleigh probability density function so that

$$\gamma = \frac{\bar{P}\alpha}{N_0 B} \tag{2.14}$$

where $\bar{P}$ is the average transmission power and γ has the following probability density function

$$f_\Gamma(\gamma) = \frac{1}{\bar{\gamma}} e^{-\gamma/\bar{\gamma}} \tag{2.15}$$

If P is the instantaneous power inserted into the channel, the instantaneous

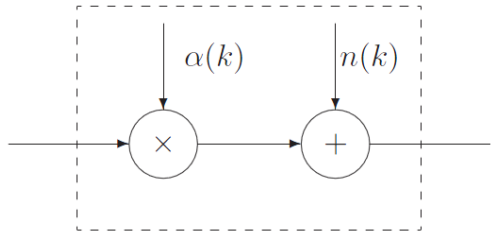


Figure 2.2: *Model of a flat fading channel with fading gain $\alpha(k)$ and white noise $n(k)$*

CSNR is equal to $\gamma P/\bar{P}$. If the source signal is Gaussian distributed, the capacity per channel use, given the stated CSNR, is

$$C = \frac{1}{2}\log\left(1 + \gamma\frac{P}{\bar{P}}\right) \tag{2.16}$$

and the average channel capacity is given in [16] and is

$$C_S = B \int_0^\infty f_\Gamma(\gamma) \log\left(1 + \gamma \frac{P}{\bar{P}}\right) d\gamma \tag{2.17}$$

The instantaneous power is of course a function of γ and can be related to the average power using the probability density function of γ.

$$\bar{P} = \int_0^\infty f_\Gamma(\gamma) P(\gamma) d\gamma \tag{2.18}$$

By using a Gateaux variation and a Lagrange multiplier, the optimal distribution of power can be found to be [16]

$$P(\gamma) = \bar{P}\left[\frac{1}{\gamma_0} - \frac{1}{\gamma}\right] \tag{2.19}$$

where

$$\frac{1}{\gamma_0} = 1 + \int_{\gamma_0}^\infty \frac{1}{\zeta} f_\Gamma(\zeta) d\zeta \tag{2.20}$$

By using the obtained power allocation and inserting it into Equation 2.17 the final average capacity for a flat fading channel is obtained

$$C_s = B \int_{\gamma_0}^\infty f_\Gamma(\gamma) \log\left(\frac{\gamma}{\gamma_0}\right) d\gamma \tag{2.21}$$

Since power must be positive, transmission can only occur when $\gamma > \gamma_0$ so the lower integration limit in the equation above has been changed to γ_0.

The rate-distortion function for a colored Gaussian signal, meaning the signal spectra is not constant, is given in [16] as

$$R_S = W \log_2\left(\kappa_x^2 \frac{\sigma_x^2}{\sigma_D^2}\right) \tag{2.22}$$

where κ_x^2 is the spectral flatness measure. The spectral flatness measure will always have a value in the interval $0 \leq \kappa_x^2 \leq 1$ and is given by

$$\kappa_x^2 = \frac{\exp\left\{\frac{1}{W}\int_0^W \log\left(S_{XX}(F)\right) dF\right\}}{\frac{1}{W}\int_0^W S_{XX}(F) dF} \tag{2.23}$$

OPTA for a channel experiencing flat fading can be found by equating the rate distortion function in Equation 2.22, with the average capacity for a fading channel in Equation 2.21.

$$W \log_2 \left(\kappa_x^2 \frac{\sigma_x^2}{\sigma_D^2} \right) = B \int_{\gamma_0}^{\infty} f_\Gamma(\gamma) \log \left(\frac{\gamma}{\gamma_0} \right) d\gamma \tag{2.24}$$

Solving for $SNR = \frac{\sigma_x^2}{\sigma_D^2}$, the following is obtained:

$$\frac{\sigma_x^2}{\sigma_D^2} = \kappa_x^{-2} \exp \left\{ r_{avg} \int_{\gamma_0}^{\infty} f_\Gamma(\gamma) \log \left(\frac{\gamma}{\gamma_0} \right) d\gamma \right\} \tag{2.25}$$

where r_{avg} is the average ratio between B and W. Since the rate distortion function used to find this expression for OPTA is for correlated Gaussian sources, and correlated sources require less resources for transmission than uncorrelated, there will be an improvement in SNR given by the inverse of the spectral flatness measure. At the same time the SNR is lowered due to the fading channel by the integral in the exponent times the bandwidth relation. The OPTA formula derived here is the optimal for a flat fading channel. In this thesis however, the input power is adjusted to keep the capacity or quality as constant as possible. To keep the capacity in Equation 2.16 constant, it is obvious that P must be inversely proportional to γ.

$$P = \frac{k}{\gamma} \tag{2.26}$$

And the average power is

$$\bar{P} = \int_{\gamma_0}^{\infty} f_\Gamma(\gamma) \frac{k}{\gamma} d\gamma \tag{2.27}$$

The lower limit of the integral is set to γ_0 to keep the integrand from approaching infinity, and should be the value of the deepest fades the power will compensate for. By inserting the value for $f_\Gamma(\gamma)$ into the equation and set $z = \gamma/\bar{\gamma}$, the following expression is obtained:

$$\bar{P} = \frac{k}{\bar{\gamma}} \int_{\gamma_0/\bar{\gamma}}^{\infty} \frac{1}{z} e^{-z} dz \tag{2.28}$$

The integral is called the exponential integral and is represented here by the function $E_1(\gamma_0/\bar{\gamma})$

$$\bar{P} = \frac{k}{\bar{\gamma}} E_1(\gamma_0/\bar{\gamma}) \tag{2.29}$$

An expression for the capacity when the rate is kept constant can be found by inserting the expression for $\bar{P}$ into Equation 2.16 and multiplying it with $2B$:

$$C_s = B \log \left(1 + \frac{\bar{\gamma}}{E_1(\gamma_0/\bar{\gamma})} \right) \tag{2.30}$$

In [1], Alouni and Goldsmith derived the following expression for the capacity using channel inversion and fixed rate:

$$C = B \log \left(1 + \frac{\bar{\gamma}}{E_1(\gamma_0/\bar{\gamma})} \right) e^{-\gamma_0/\bar{\gamma}} \tag{2.31}$$

This is in accordance with the expression in Equation 2.30 except for the last exponential which represents $1 - P_{out}$ which is the probability of $\gamma > \gamma_0$, i.e. the channel gain is large enough for transmission. This is included since no transmission can be completed when the channel is in a very deep fade. In Equation 2.30, this probability was excluded because the system used in this thesis has to transmit all the time to avoid delay, and in deep fades, it is up to the receiver to handle the poor data received. By implementing diversity, P_{out} will increase and the expression in Equation 2.31 will approach that derived here in Equation 2.30.

2.2 Transmission Scheme

Based on the theory explained in the previous section, this section will explain the transmission scheme studied on this thesis. The system utilizes a feedback channel to send more than one sample per source sample, and in this way expand the bandwidth. It will be shown that for the optimal case the system performs according to OPTA. First, a general version of the system will be described, and then an ideal system with bandwidth expansion, and at last the system with noise on the feedback channel as well will be presented.

If the signal and noise are both white, Equation 2.13 is clearly satisfied since then both S_{XX} and S_{NN} are constants. In this case the signal manipulator in the transmitter will just be a constant adapting the input signal to the channel, denoted α. This is intuitive since the input signal is white and thus each sample must be equally weighted in order for it to maintain whiteness. The receiver filter tries to recreate the input signal, i.e. minimizing the mean squared error. Thus, the receiver-shaping filter will be a Wiener filter, which attenuates the signal plus noise in an optimal way. According to Equation 12.7.9 in [11], this yields the optimum Wiener factor

$$\widehat{\beta} = \frac{\sigma_X^2}{\alpha^2 \sigma_X^2 + \sigma_N^2} \tag{2.32}$$

where σ_X^2 is the input variance and σ_N^2 is the noise variance. By setting $\frac{1}{\alpha}$ outside of the equation and introducing $\sigma_C^2 = \alpha^2\sigma_X^2$ yields

$$\widehat{\beta} = \frac{1}{\alpha}\frac{1}{\alpha(1+\frac{\sigma_N^2}{\sigma_C^2})} = \frac{1}{\alpha}\beta \tag{2.33}$$

This system is shown in Figure 2.3, and combined with an ideal feedback channel it makes up the basis for finding a bandwidth expanding system with performance given by OPTA.

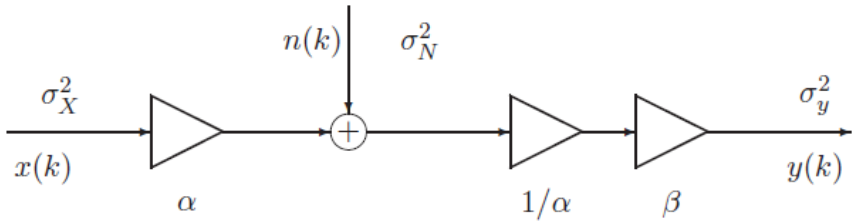


Figure 2.3: *Optimal system with white source and noise signal using no bandwidth alteration*

As stated in Section 2.1, bandwidth expansion is one way to increase the channel capacity. Below, a linear system using incremental transmission with an ideal feedback channel, as reviewed in [3] and depicted in Figure 2.4, is analyzed. Previously, it was shown that if both the input signal and noise were white, the system in Figure 2.3 reached OPTA. By adding a noiseless feedback channel, the transmission scheme is not altered and the system will still act according to OPTA.

The feedback channel can now be used to expand the bandwidth of the source signal transmitted over the channel. This can be done by transmitting several channel samples for each source sample. As long as the channel samples are uncorrelated, the bandwidth will be expanded in an optimal way. Since one transmission per source sample reaches OPTA, then M transmissions per source sample will intuitively also reach OPTA as long as the transmissions are independent of each other. The only difference is that transmissions occur at a higher frequency. If the energy for transmitting one source sample $x(n)$ over the channel is constrained by some variable, E_S, and M channel samples are transmitted per source sample, the energy per channel sample is constrained by $E_1 = \frac{E_S}{M}$. The first channel sample is $x(n)$, and the receiver receives a

noisy version of $x(n)$ and sends it back to transmitter over the ideal feedback channel. The transmitter, now in possession of the real sample and the noise-corrupted sample held by the receiver, sends the difference between the two samples. The receiver receives a noisy version of the difference and adds it to what was received earlier and can now be expected to be closer to the real value than after just one transmission. Notice that the sampling frequency of the channel is MF_S while for the input signal it is F_S.

For example, if $M = 3$ and the transmitter wants to send the source sample $x(n) = 100$ to the receiver. Since the feedback channel is ideal, the transmitter always knows exactly what was received by the receiver and transmits the error. $y_m(n)$ is the received channel sample for iteration m and $\hat{y}(n)$ is the accumulated channel samples making up the total received source sample. Since

Sent at transmitter		Received
$x_1(n) = 100$	$\longrightarrow$	$y_1(n) = 110,\ \hat{y}(n) = 110$
$x_2(n) = -10$	$\longrightarrow$	$y_2(n) = -11,\ \hat{y}(n) = 99$
$x_3(n) = 1$	$\longrightarrow$	$y_3(n) = 1,\ \hat{y}(n) = 100$

each of the channel samples are made with less energy than one would use if there were no bandwidth expansion, one might think that the same result may be accomplished by sending one sample with all the energy. This is not the case, as shown in earlier sections.

The system with bandwidth expansion and an ideal feedback channel is depicted in Figure 2.4. If the system parameters A_i and B_i are optimized so that for each transmission, A_i provides correct power level and B_i provides optimal reception. This system will reach OPTA for uncorrelated Gaussian sources and noise.[16] The optimized parameters will not be derived here and

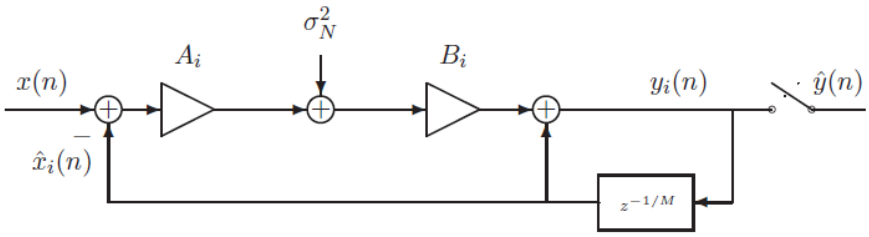


Figure 2.4: *System for incremental transmission of Gaussian samples with ideal feedback channel*

the full derivation can be found on page 165 in [3]. The parameters are given

as:

$$\sigma_{D_0}^2 = \sigma_X^2 \tag{2.34}$$

$$A_i = \sqrt{\frac{E_1}{\sigma_{D_{i-1}}^2}} \tag{2.35}$$

$$B_i = \frac{\sqrt{\sigma_{D_{i-1}}^2 E_1}}{E_1 + \sigma_N^2} \tag{2.36}$$

$$\sigma_{D_i}^2 = \frac{\sigma_{D_{i-1}}^2 \sigma_N^2}{E_1 + \sigma_N^2} \tag{2.37}$$

for $i = 1, 2, ..., M$. The values $\sigma_{D_i}^2$ are the variances of the input signals at each iteration, which is equal to the remaining error after i iterations. It is important to see that a system with feedback does not increase the capacity, but reduces the necessary power level by expanding the bandwidth [4].

2.2.1 System With Noisy Feedback Channel

In this section, the optimal system with noisy feedback will be presented. This is the foundation for the transmission scheme used in this thesis.

In all real wireless communication systems, the feedback channel described will also be corrupted by noise. Nevertheless, as stated in [3], the results obtained in the previous section are of use in systems where a feedback channel is much more reliable than the forward channel. Thus, if one assumes the base station can deliver much more power than the transmitter (20-30 dB over that of the transmitter) a feedback channel with noise will be of great help. This is exactly the case in this thesis, where it is assumed that the helicopter can transmit with very low power compared to the base station. The problem then is to find the optimal power levels on the feedback channel to come as close to OPTA as possible.

It is assumed that the noise level in the feedback channel is equal to that of the feed forward channel. The parameters A_i, $i = 1, 2, ..., M$ still adjust the signal level in the feed forward channel to the same average level for each iteration, while C_i does the same for the feedback channel, although to a different level. D_i is the optimal Wiener factor that will minimize the total distortion between d_i and $\hat{d}_i$ for each iteration. The accumulator updates the feedback level for each iteration of a sample. d_i will, as with ideal feedback, be optimal although now with an equivalent noise resulting from both noise contributions

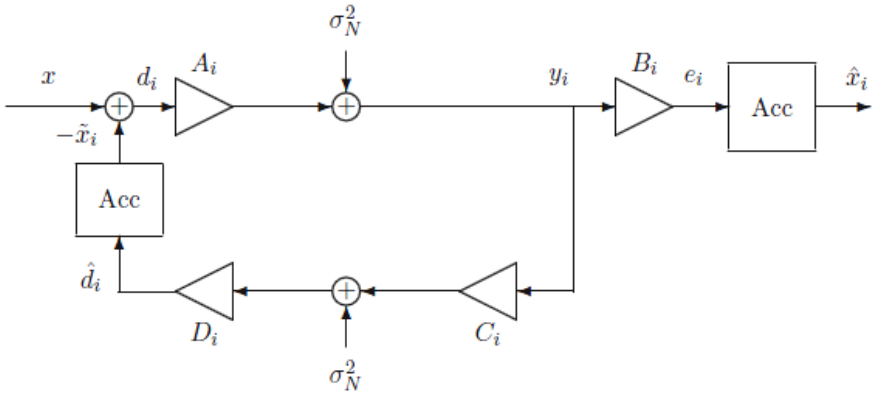


Figure 2.5: *System for incremental transmission of Gaussian samples with noisy feedback channel*

weighted by the multipliers. The problem is now that the accumulators in the receiver and the transmitter do not see the same signal, which implies that $\widehat{x}_i$ is not equal to $\widetilde{x}_i$. What can be done is to select B_i, $i = 1, 2, ..., M$ to make e_i as close to d_i as possible. This should make $x_{out} = \widehat{x}_M$ as close to x as possible.

The calculations of the optimal parameters will not be done here, but results obtained from simulations done by Anna Kim and Tor Ramstad at IET, NTNU gave the formulas below. As before, for the first transmission $\sigma_{D_0}^2 = \sigma_X^2$ and

$$A_i = \sqrt{\frac{E_1}{\sigma_{D_{i-1}}^2}} \tag{2.38}$$

$$C_i = \sqrt{\frac{E_2}{E_1 + \sigma_N^2}} \tag{2.39}$$

$$D_i = \frac{\sqrt{\sigma_{D_{i-1}}^2 E_2 E_1}}{(E_2 + \sigma_N^2)\sqrt{E_1 + \sigma_N^2}} \tag{2.40}$$

$$B_i = \frac{\sqrt{\sigma_{D_{i-1}}^2 E_1}}{E_1 + \sigma_N^2} \tag{2.41}$$

$$\sigma_{D_i}^2 = \frac{\sigma_{D_{i-1}}^2 \sigma_N^2}{E_1 + \sigma_N^2} \tag{2.42}$$

for $i = 1, 2, ..., M$ [16].

2.3 The Wireless Channel

The wireless operation of the helicopter leads to several different degradations of the signal, and more complex models for path loss and multipath fading. This section will describe the implications of transferring a signal through the air and how this affects the received signal.

2.3.1 Carrier Frequency

The carrier frequency determines where in the frequency domain the band-limited signal will position itself. According to the Norwegian Post and Telecommunication Authority's[1] *Forskrift om generelle tillatelser til bruk av frekvenser §8*[2]:

> *§8. Mobile video links*
> *The center frequencies 2327 MHz and 2390 MHz are allowed to be used for mobile video link with maximum radiated power 2 W e.i.r.p. Maximum occupied bandwidth is 8 MHz.*

This is a part of the frequency band where no registrations or permits are necessary as long as the use comply with the regulations. The two different frequencies do not differ much and there is no practical advantage in choosing one over the other. However, 2390 MHz is closest to the well-known and examined 2.4 GHz-band which is an advantage when looking at theoretical performance. Many communication systems utilize the unlicensed industrial and scientific 2.4GHz-band where IEEE 802.11 wireless LAN and Bluetooth are the most prominent. Note that the maximum power in any direction cannot exceed 2 W. This means that the power on the feedback channel cannot exceed 2 W. Neither can the feed-forward channel, but it is assumed that it will not exceed this limit simply because of energy constraints.

The bandwidth needed for transmission without bandwidth expansion, assuming transmission of one pixel per sample is 240×320 pixels $\times 15$ fps $= 1.152$ MHz. The system can thus maximum expand the bandwidth 6 times the initial bandwidth and still act according to regulations. The implications of the use of this carrier frequency will not be further explained here, instead the implications will be addressed in the sections most fitting.

[1] Post- og teletilsynet
[2] http://www.lovdata.no/cgi-wift/ldles?doc=/sf/sf/sf-20090602-0580.html

2.3.2 Path Loss

When transferring a signal, by either cable or wirelessly, the signal will lose power with increasing distance between transmitter and receiver. When the information lies in the amplitude of the signal, the signal level must be adjusted for this loss. Since path loss can be considered constant over several samples, the samples can be amplified in the receiver to obtain the expected mean value.

It is interesting to model path loss to find what transmission power is needed to reach a certain SNR in the receiver. In a wireless system, this path loss is difficult and complex to model accurately due to the large variation of channels and environments. The transmission path between the transmitter and the receiver can vary from a simple line-of-sight path and one that is obstructed by large, small and even moving objects as cars and people.

The field of radio propagation is sometimes divided into large and small-scale path loss; where large-scale path loss determines the loss in received signal power due to the large distance between transmitter and receiver, while small-scale path loss refers to the rapid change in signal power due to small changes in the distance between the transmitter and receiver. In this thesis, large-scale path loss will be referred as path loss, and small-scale will be referred as multipath fading or just fading. The difference between these is illustrated in Figure 2.6 and fading is further examined in Section 2.3.3. In this thesis, the importance of estimating path loss for the scenario, presented in Section 1, is that it gives a realistic pointer as to what SNR's that can be expected in the receiver with a given transmission power.

The mechanism behind path loss in a wireless system are diverse, but can generally be attributed to reflection, diffraction and scattering[17]. If a signal hits a smooth object with much larger dimensions than the wavelength of the signal, a part of the signal will be reflected of the surface perpendicular to the angle of incidence. The rest of the signal will be absorbed by the object or go through it with a shift in direction. How much that will be absorbed and how much that will be reflected depends on the material of the object, angle of incidence and wavelength. Typical objects that reflect signals are walls, floors and the surface of the earth. Diffraction occurs when the path of the signal is obstructed by objects with sharp edges or with small openings in it. Based on Huygens' Principle, secondary waves are created after impact and will form behind the objects even though there is no line of sight. How the signal will "bend" around the object is dependent on the geometrics of the object and the frequency of the signal. Diffraction is illustrated in Figure 2.7. Scattering

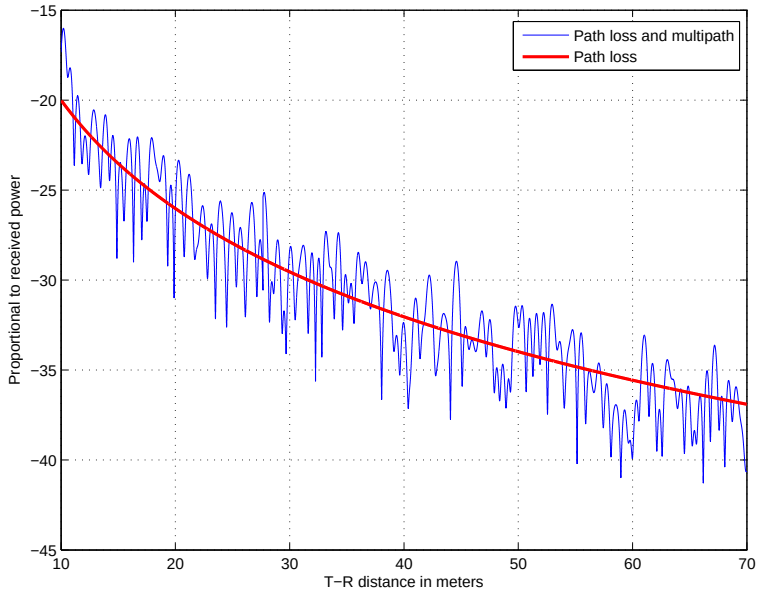


Figure 2.6: *Large and small scale path loss. The signal will fluctuate rapidly over time and decrease steadily in average signal power as a function of the distance between transmitter and receiver.*

is what happens when a signal hits objects that are small compared to the wavelength of the signal. The signal will then spread out in all directions. This typically happens when the signal hits a rough surface.

When the helicopter is visible for the operator, the path is called a line-of-sight-path. If there are no obstructions, the signal will propagate along a straight line between transmitter and receiver. If this is the case, the received signal power will decrease proportional to the square of the distance between them. The path loss in dB can then be described by:

$$PL(dB) = 10\log_{10}\frac{P_t}{P_r} = -10\log_{10}\left(\frac{K}{d_{T-R}^2}\right) \tag{2.43}$$

where

$$K = \frac{G\lambda^2}{16\pi^2} \tag{2.44}$$

P_t and P_r are the transmitted and received power respectively, G is the total antenna gain from both receiver and transmission antennas, λ is the wavelength

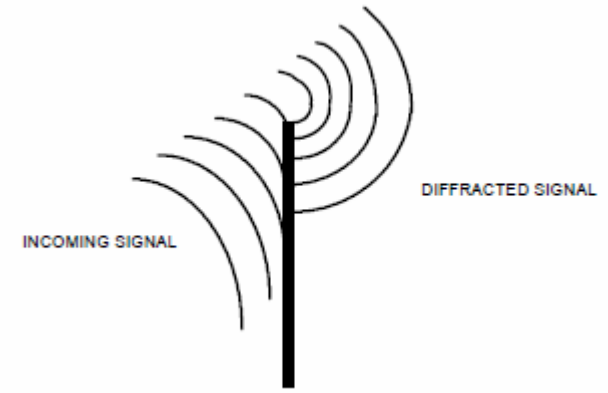


Figure 2.7: *Diffraction of a signal. A diffracted wave front is formed when the impinging transmitted signal is obstructed by sharp edges within the path [24]*

of the signal and d_{T-R} is the distance between transmitter and receiver.[5]

When there is no line of sight between the transmitter and receiver, the model becomes more complex. Many actual models for path loss are derived from a combination of empirical and analytical approaches. Actual field measurements are combined with curve fitting to give a more general model of the environment. Most of these models agree that path loss is a logarithmic function of the relative distance between transmitter and receiver and a path loss exponent, n.[17]

$$PL(d) \propto \left(\frac{d}{d_0}\right)^n \tag{2.45}$$

$$PL(dB) = PL(d_0) + 10n \log\left(\frac{d}{d_0}\right) \tag{2.46}$$

$PL(d_0)$ in Equation 2.46 is the path loss at a reference distance d_0 measured in dB. It is often estimated using Equation 2.43 for line-of-sight path loss. n is dependent on the environment in question and the carrier frequency and typical values ranges from 2 in free space up to $\sim$ 5 in a shadowed urban area[17]. This equation does not give an accurate path loss models for all environments, but gives a simple valuable estimate that can be used to analyze how a system will perform under substantial path loss, which is the intention of estimating path loss in this thesis. Since path loss is a slow varying feature of the wireless channel, it does not corrupt the signal alone; it makes the signal more vulnerable to noise. Since path loss is slowly changing one can assume that it will be the same for several consecutive frames and the pixel values can

be scaled to have the expected mean value. This is in contrast to fading that alters the received signal even when noise is not considered.

When the helicopter is operating inside a building, the path loss becomes more complicated. Even when there is a line of sight between transmitter and receiver, the signal will still be corrupted by many reflections and objects since the number of objects per unit volume can be assumed much higher inside than outside. An indoor environment is also more variable than an outdoor environment and path loss may vary a lot from building to building depending on construction material, antenna placement and other variables as whether or not doors are open.

Indoor path loss can be modeled in the same way as outdoor path loss (Equation 2.46), but will now be determined by a random parameter as well.

$$PL(dB) = PL(d_0) + 10n \log\left(\frac{d}{d_0}\right) + X_\sigma \tag{2.47}$$

where X_σ is a normal distributed random variable with zero mean and σ standard deviation. In [2], Andersen et al. created models for several indoor environments and they found that n is in the range ~ 2 - ~ 3 and σ in the range ~ 5 - ~ 10.

2.3.3 Fading

Small-scale path loss or fading is used to describe the rapid changes in amplitude or phase of a signal over very short periods of time or distance. As previously described in Section 2.3.2, the signal will take many different paths to reach the receiver. Therefore, the signal visible to the receiver is a superposition of many realizations with a shift in amplitude, phase and/or time. In addition to the signal strength decreases as a function of the distance between transmitter and receiver, the signal strength will also change very rapidly over very short periods of time due to multipath. The causes of fading are also related to reflection, diffraction and scattering as large-scale path loss is.

There are in general four types of fading that can occur and it is important to know what type can be expected so its effects can be counteracted. It is also important to know what type to expect when the system's performance should be simulated on a computer as is the case in this thesis. What type of fading one can expect is dependent on several factors, including the relative movement of transmitter and receiver, the environment in question and the

carrier frequency. This section will give a short description of the different types of fading and explain in what scenarios each of them can be expected.

A power delay profile is a measurement taken in the environment of operation and gives rise to many parameters used to describe and compare different channels. It is a spatial or temporal average of several consecutive impulse response measurements collected and averaged over a local area. A typical power delay profile from an indoor environment is shown in Figure 2.8. The mean excess delay ($\bar{\tau}$) and rms delay spread (σ_τ) are multipath parameters that can be calculated from the power delay profile (abbreviated PDP) and is used to compare different multipath channels. The mean excess delay spread is the first moment of the PDP while the rms delay spread is the square root of the second moment of the PDP. Both the mean excess delay and rms delay spread says something about how much time dispersion there is in the channel, meaning the average time it takes from the first signal realization is received to the last multipath component is received. Typical values are on the order of nanoseconds in indoor channels, and microseconds in outdoor radio channels[17]. It is important to note that the parameters are highly dependent on the noise threshold used in the measurements of the PDP. The noise threshold is used to differentiate received multipath components from thermal noise in the receiver. If the threshold is too low, noise will be mistaken as multipath components, hence giving an artificially high delay spread. If it is too high, small multipath components may be interpreted as noise which gives an artificially low delay spread.

In the same way delay spread parameters operate in the time domain, coherence bandwidth (B_c) is used to characterize a channel in the frequency domain. The coherence bandwidth is proportional to the inverse of the rms delay spread, but the exact relationship depends on the structure of the environment in question. Coherence bandwidth is defined as the range of frequencies of which the channel passes the signal with equal gain and linear phase. Since an exact relationship between coherence bandwidth and rms delay spread is a function of the specific channel impulse response, exact general values cannot be obtained as they have to be measured in the environment. According to [15], there exists several estimates which range between:

$$\frac{1}{50\sigma_\tau} \leq B_c \leq \frac{1}{5\sigma_\tau} \tag{2.48}$$

While delay spread is a parameter that describes the time dispersive nature of the channel in a local area, it says nothing about the time varying nature

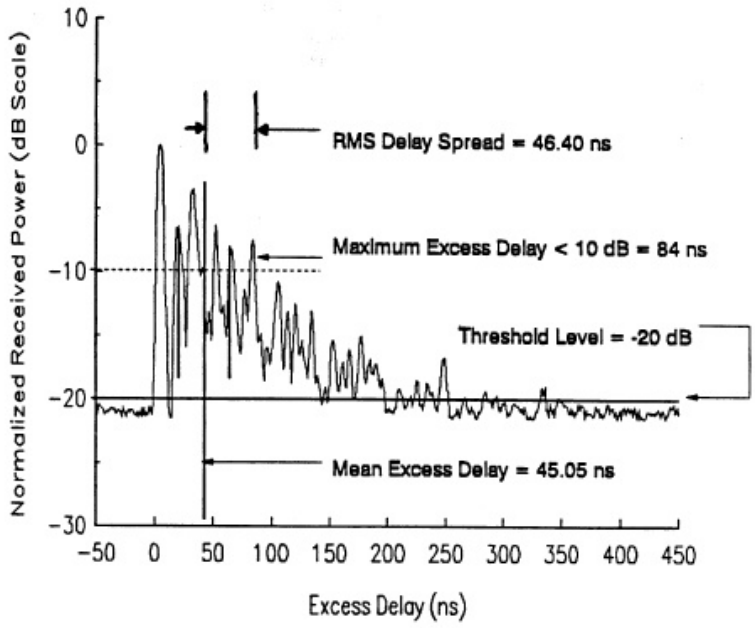


Figure 2.8: *An example of an indoor power delay profile with rms delay spread, mean excess delay, maximum excess delay(10dB) and thresholds shown.[17]*

of the channel caused by the relative movement between the transmitter and receiver. The Doppler spread is a measure of how much the frequency is changed due to movement. At a certain angle α between the transmitter and the receiver and a relative speed of v between them, the Doppler shift in frequency is:[17]

$$f_d = \frac{v f_c}{c} \cos(\alpha) \tag{2.49}$$

where c is the speed of light and f_c is the carrier frequency. The Doppler spread is the maximum Doppler shift, which is when $\cos(\alpha)$ is equal to one.

$$B_D = \frac{v f_c}{c} \tag{2.50}$$

The coherence time is the time domain dual of the Doppler spread. It is used to characterize the time varying nature of the frequency dispersiveness of the channel in the time domain and is approximately the inverse of the Doppler spread.

$$T_C \approx \frac{1}{B_D} \tag{2.51}$$

Coherence time can be seen as the time duration over which two received signals have a strong correlation in channel gains[17]. In relation to channel

estimation, which is explained in Section 3.3, the coherence time says something about how often the channel needs to be estimated to have a reliable estimation.

If one defines the coherence time to be the time where the channel gains have a correlation greater than 0.5, it is approximately[23]

$$T_C \approx \frac{9}{16\pi B_D} \tag{2.52}$$

The mean between Equation 2.51 and 2.52 is often used as a "rule of thumb" for estimation of coherence time since Equation 2.52 is often too strict and Equation 2.52 implies a time duration where the fading may fluctuate wildly[17].

The parameters described above can be used together with the sampling period T_s and the signal bandwidth B to estimate which of the four different types of fading that can be expected. The fading can be either flat or frequency selective, meaning that the channel can either affect all spectral components in the same way or apply different gain to different frequencies. If the fading is frequency selective, the channel can be viewed as a discrete FIR filter which will give rise to inter-symbol-interference (ISI) in the receiver. Flat fading is the most common and the spectral characteristics of the signal are preserved, but the strength of the received signal will fluctuate due to multipath. A signal undergoes flat fading if

$$B \ll B_c \tag{2.53}$$

and

$$T_s \gg \sigma_\tau \tag{2.54}$$

and frequency selective fading if

$$B > B_c \tag{2.55}$$

and

$$T_s < \sigma_\tau \tag{2.56}$$

In addition to being either flat or frequency selective, the signal can undergo either fast fading or slow fading, relating to how fast the channel changes. If the fading is fast, the channel gains from one symbol to the other have little to no correlation while if it is slow, the channel gains are expected to be very much the same from one symbol to the next. Slow fading implies that the Doppler spread of the channel is much less than the bandwidth of the signal. The

relative velocity between the transmitter and the receiver determines whether the channel will be a fast or slow fading channel. Fast fading is expected if

$$T_s > T_c \tag{2.57}$$

and

$$B < B_d \tag{2.58}$$

and slow fading is expected if

$$T_s \ll T_c \tag{2.59}$$

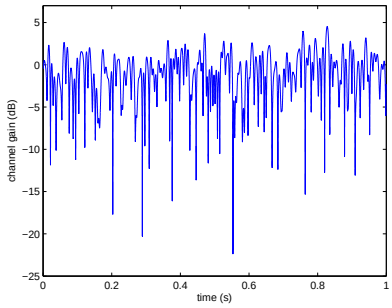
and

$$B_s \gg B_d \tag{2.60}$$

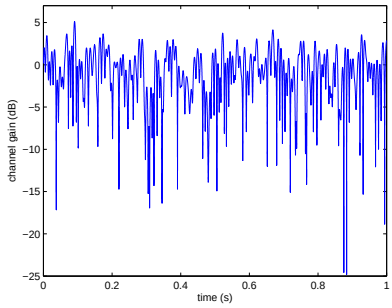
In digital transmission, flat fading will affect the signal so that the signal is closer to the error-threshold and hence more sensitive to noise. If the signal level is above this threshold, the signal will be perfectly reconstructed. In analog transmission on the other hand, the amplitude level is continuous and the fading creates error even without noise. Therefore, when the information lies in the amplitude, the channel has to be estimated in either the transmitter, the receiver or both, so that the change in amplitude can be compensated.

2.4 Channel Diversity

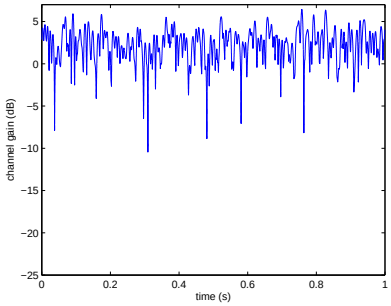
As explained in Section 2.3.3, fading will take its toll on the system. By introducing channel diversity, this degradation of the signal can be significantly reduced at a relatively low cost. Almost all wireless communication devices made now a days utilize one or more forms of diversity to combat fading[17]. The theory is that using more than one uncorrelated channel, the channels will experience different fading characteristics, so the combination of the signals will have less fading than in a single channel. If one have M independent Rayleigh fading channels in the receiver, and p is the probability that one of them will have an instantaneous SNR below a certain limit, the probability that all of them will have an instantaneous SNR below the same limit is of course p^M. Hence, it is obvious that with only two different channels, the error rate will decrease significantly. When one channel is in a deep fade, there is only a small chance that the other channel will also experience this fade. See Figure 2.9 for an illustration of this. This section will shortly describe some different techniques for applying diversity on the channel. The focus will be on utilizing two different channels but the principle can easily be expanded to hold for more than two.



(a) A Rayleigh fading channel over a time period of 1 second



(b) A Rayleigh fading channel over a time period of 1 second



(c) The sum of the two fading channels in (a) and (b)

Figure 2.9: *Two uncorrelated Rayleigh fading channels and the sum of the two. Notice that the sum has much less deep fades than the two individually*

Channel diversity can be obtained in several ways, but the principle for all methods is to add some redundancy to the channel. This can be done by transmitting the same signal over two different carrier frequencies. If the frequencies are separated by much more than the coherence bandwidth, the two signals will experience independent fading over the channel. This is called frequency diversity.

Time diversity can be achieved by dividing the signal into time slots longer than the coherence time of the channel and each slot is sent two or more times. Since the blocks have a longer duration than the coherence time, the blocks will experience independent fading. When received, one must store the first time slot, wait for the next one and then use the best one or combine them in

some manner.

A third way and the original diversity technique is called antenna diversity or space diversity. Space diversity exploits that a signal sent from two antennas separated in space will experience different channel gains. To have independent or nearly uncorrelated channels, the separation needs only to be about half a wavelength or more[9] and simulations performed in [14], showed that at 2.4 GHz a separation down to 0.16λ gave essentially uncorrelated signals. For a signal with a carrier frequency of 2390 MHz, which is the carrier frequency decided upon for this thesis in Section 2.3.1, one wavelength is $\lambda = 12.6$ cm, which is a realistic separation on the base station. Note that this small separation will not have any impact on path loss, just the rapid fluctuation in signal power due to multipath fading. The signal received on two antennas will also introduce channel diversity. The small separation is normally used on the mobile device while the base station needs a much larger separation to obtain diversity. This is because the base station is normally placed at a higher ground than the mobile (the helicopter in this case) and experiences much less scatter so that the channel does not change that much by movement close to the base station. This is not the case in this scenario where the base station and the helicopter are assumed to be positioned in the same environment.

If the signal is transmitted with two antennas and received with one, the received signal is the superposition of the two signals with different fading and will give a total channel gain similar to that in Figure 2.9(c). If the signal is transmitted using one antenna and received using two, the receiver has more options and how the signals are combined is called combining scheme. The receiver can either choose the signal with the highest amplitude, hence assuming that the signal with most power has the smallest fade. This will give the best result if one of the signals is experiencing a deep fade but in many cases neither of the signals will experience a very deep fade and this scheme waste power. Maximal ratio combining (MRC) is a scheme first proposed by Kahn[12][17] which weighs each of the received signals according to their individual SNR and then adds them together. Maximal ratio combining produces a SNR equal to that of the sum of all incoming signals, thus it can get an acceptable SNR even though none of the signals are themselves acceptable. Maximal ratio combining requires channel knowledge at the receiver. As will be explained in Section 3.3, this is not feasible in this thesis so another combining scheme must be used. Equal gain combining (EGC) is almost the same as MRC but the branches are all weighted equally. This can still give acceptable SNR even though none of the branches are themselves acceptable but the total gain will be smaller than

for MRC. Both EGC and MRC require a coherent phase reference to co-phase the different branch signals before they are summed.

2.4.1 Diversity Gain

The gain from diversity can be measured as how much power can be reduced to get the same performance as without diversity. If the receiver always picks the signal with the best instantaneous SNR and sends that to output, the gain can be showed to be[17]:

$$\bar{\gamma} = \Gamma \sum_{k=1}^{M} \frac{1}{k} \tag{2.61}$$

where $\bar{\gamma}$ is the average received output SNR, Γ is the average SNR on each of the uncorrelated channels and M is the number of channels or diversity branches.

If maximum ratio combining is used, the total SNR is the sum of SNR's on each of the channels. Meaning the total received SNR can be acceptable even though none of the received SNR on any of the branches is. The gain from using EGC is only marginally inferior to that of MRC and the total received SNR can be written as:

$$\bar{\gamma} = \Gamma \left[1 + \frac{(M-1)\pi}{4} \right] \tag{2.62}$$

The diversity gains from using different combining schemes can be seen in Figure 2.10.

2.5 Video Compression

In videos picturing natural events, with a reasonable high frame rate, two consecutive frames contain much of the same information. Some parts of the image may have been moved in space and other parts may have small variations in light, but overall to the human eye they look very much the same. Further, pixels close to each other in one frame have a very strong correlation between them. Information that is repeated in many pixels, either in space (in the same frame) or over time (consecutive frames), is called redundant information, i.e. information that is unnecessary or repeated. Examples of this can be seen in Figure 2.11 and 2.12. When transmitting over a channel, one want to be able to transmit as little information as possible to reduce the total energy consumption. The receiver should still be able to interpret the signal with as

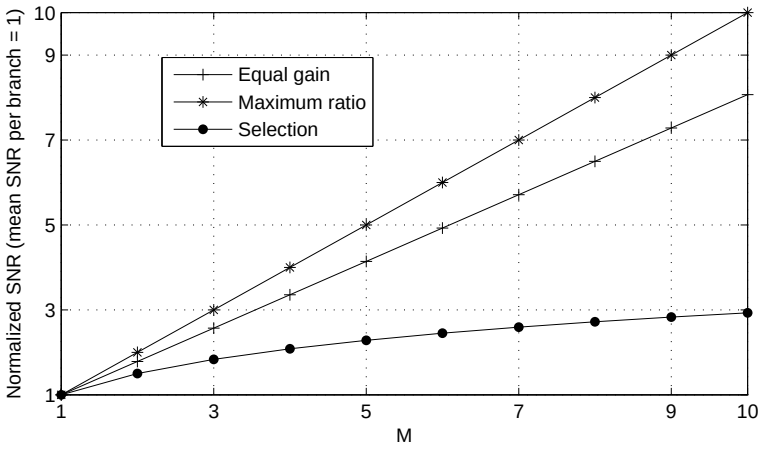


Figure 2.10: *Diversity gain from using different combining schemes. M is the number of uncorrelated signal branches and Γ has been set equal to 1*

(a) *Frame 1* (b) *Frame 2* (c) *Difference between Frame 1 and 2*

Figure 2.11: *Two consecutive frames in a video of people walking down a street and the difference between them. The entropy in both the frames is about 6.75 bits/pixel while it is about 2.5 bits/pixel in the difference image*

little distortion as possible, hence the aim of video and image compression is to remove as much as possible of the redundant information. Further, since humans do not perceive small distortions in the video, this too is irrelevant information.

One of the objectives of this thesis is to simulate a complete system with video compression and transmission, where the compression is based on a three dimensional DPCM-codec with motion compensation. This section will explain three-dimensional DPCM with motion compensation. First, prediction in time is explained, then prediction in space.

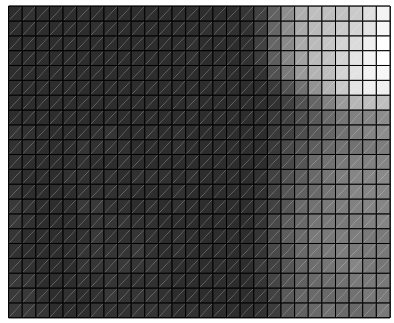


Figure 2.12: *A small (21x28 pixels) section of the video frame depicted in Figure 2.11(a). Notice that neighboring pixels are very much the same*

Differential coding in three dimensions is based on the assumption that each pixel has a large component consisting of a weighted sum of the closest spatial neighbors and the pixel at the previous frame shifted in a certain direction due to motion. As seen in Figure 2.11, two consecutive frames look very much the same but they still have a small difference between them due to motion. If the image is divided into smaller blocks there is a good chance that a block in the current frame will exist in the previous frame, although at a different position. Therefore, instead of transmitting the original block, the difference between the block and the previous frame's corresponding block is transmitted together with a motion vector describing which part of the previous frame was subtracted. The receiver, which has the previous frame, can then add the block given by the motion vector to the received block to get the decoded block.

The motion vector (abbreviated MV) needs to be transmitted as well, and it is reasonable to believe that the system is more vulnerable to errors in the MVs than in the rest of the image. Hence, they have to be transmitted with more power. Each MV consists of two numbers where the range is depending on the search window. They can be transmitted using the same transmission scheme as the rest of the image, but with a slightly higher power. The search for the best MV is a time consuming operation and many algorithms have been proposed to reduce the search time. Many of them are based on the principle that the block that minimizes the squared error is a local minimum, so that one should always search in the direction that reduces the error[13]. This assumption does not always hold, but very often gives good results[10]. To reduce the search time further, the frame can be decimated in the beginning of the search to roughly locate the area where the minimum is, and then search

with a higher resolution in this area. Many motion estimation algorithms also use half-pixels or even quarter-pixels, meaning the image is interpolated to give a higher accuracy for the motion vector. Interpolation and decimation can be combined to get both an efficient and accurate algorithm to find the MV. The MV can be written as $[a, b]$ and the compression in time can be written as:

$$d(p, m, n) = x(p, m, n) - \eta \times x(p - 1, m + a, n + b) \tag{2.63}$$

where p is the frame number and m and n is the row and column number respectively. η is the weight applied to the predicted signal, and should be less than one to keep the system stable.

Image Compression

As seen in Figure 2.12, there is much redundancy between pixels in the same frame in addition to between frames. After the motion estimated pixels from the previous frame have been subtracted, there is still some spatial correlation left in the signal. To remove some of this redundancy, spatial differential coding can be used. This attempts to remove the redundancy by predicting each pixel based on previously transmitted pixels in the same frame, and just process the difference between the predicted pixel and the actual value. If the pixels are correlated, the variance of the difference will be smaller than that of the original signal. The differentiation in space can be written as:

$$\hat{d}(p, m, n) = d(p, m, n) - \alpha d(p, m - 1, n) - \beta d(p, m, n - 1) - \rho d(p, m - 1, n - 1) \tag{2.64}$$

where $d(p, m, n)$ is the result from Equation 2.63. The parameters α, β and ρ can be calculated to minimize the variance of the prediction error, assuming that $d(p, m, n)$ has zero mean:

$$\begin{aligned} E[\hat{d}^2(p, m, n)] = & E[(d(p, m, n) - \alpha d(p, m - 1, n) - \\ & \beta d(p, m, n - 1) - \rho d(p, m - 1, n - 1))^2] \end{aligned} \tag{2.65}$$

The zero mean assumption is valid since $d(p, m, n)$ is the original pixel value minus a pixel from the closest frame. So if two consecutive frames have the same mean and η is close to 1, the expected value of d is zero. To find the optimal weights, Equation 2.65 is differentiated with respect to each coefficient, equated to zero and solved to obtain the following set of simultaneous equations:

$$\begin{bmatrix} R_d(0,0) & R_d(1,1) & R_d(0,1) \\ R_d(1,1) & R_d(0,0) & R_d(1,0) \\ R_d(0,1) & R_d(1,0) & R_d(0,0) \end{bmatrix} \times \begin{bmatrix} \alpha \\ \beta \\ \rho \end{bmatrix} = \begin{bmatrix} R_d(1,0) \\ R_d(0,1) \\ R_d(1,1) \end{bmatrix} \tag{2.66}$$

Note that the optimal parameters are defined solely by the two-dimensional autocorrelation function (R) of the image and can be obtained by some elementary matrix operations. When deriving the parameters, an ergodic process was assumed, meaning the autocorrelation function is only dependent on the distance between the two actual pixels and not their absolute position in the image. This is of course not accurate since how much two pixels, a given distance from each other, differ, will vary at every position in the image. To avoid this, the autocorrelation has to be calculated many times during the processing of one single frame to ensure optimal weights.

Differential coding can be implemented in two different ways: closed loop and open loop. In closed loop, there is only one predictor which is placed after the main noise source, and the predicted signal is sent back to the source coder via a feedback channel. Closed loop DPCM is an example of joint source channel coding since noise from transmission is taken into account when the signal is coded or decoded. In open loop, there are two identical predictors. One placed at the source coder and one at the decoder, making the channel or transmission independent of the coding. This saves some complexity but will add a loss in performance compared to closed loop since the additional variance from the additive channel noise are not considered in the prediction.

CHAPTER 3

SIMULATIONS AND METHODS

This chapter will describe how the individual parts and the system as a whole was implemented and tested, and how different parameters were evaluated. Assumptions and approximations made in the simulations will be thoroughly explained to point out potential weaknesses of the simulations.

3.1 Input and Tools

All simulations were done in MATLAB[1], and the input for the transmission and video/image compression simulations was chosen from three different video-clips. The video clips are all from The Laboratory for Image and Video Engineering (LIVE) at The University of Texas' video quality database[19, 20]. Frames from each of the three videos are depicted in Figure 3.1. All of them were resized to QVGA resolution (320x240 pixels) and converted to gray scale. The resizing and color conversion were done using the open source program Virtual Dub[2]. The videos are chosen so that they all have some movement between consecutive frames in different magnitude. They are not chosen to fit the scenarios described in Section 1, but rather cover as many situations with different motion as possible. A description of each of the videos follows.

Test Video 1
A frame from test video 1 (TV1) is depicted in Figure 3.1(a). It shows a crowd

[1]See http://www.mathworks.com/products/matlab/
[2]See http://www.virtualdub.org for details

(a) *Test video 1. People crossing the street*

(b) *Test video 2. A stationary camera zooms out looking over a build site*

(c) *Test video 3. The camera follows a tractor driving on a field*

Figure 3.1: *One frame from each of the three test videos used in the simulations and testing*

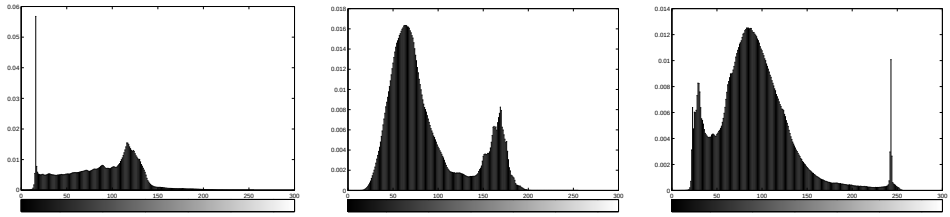
of people crossing an intersection in the street. There are some people close to the camera as seen in the frame and other in the background giving a different speed of movement. Some people as crossing the street horizontally while other diagonally, but most of the movements are in the horizontal direction. There are also stationary objects in the video as the building in the background and the ground. The density function for TV1 is displayed in Figure 3.2(a) and there is a notable peak for the pixel value 16. This peak corresponds to the pants and backpack that the woman crossing close to the camera is wearing. The camera is stationary during the entire sequence and it is the surroundings that move.

Test Video 2

Test video 2 (TV2) starts with a close-up look at some cranes above the intersection of many railway tracks, and immediately starts to zoom out to give a far-away view of the tracks and the cranes with some buildings in the background. There is very little motion in the video except from the zooming and one person crossing the tracks in the middle of the frame. The motion is mostly very even and the zooming is slow so two consecutive frames are expected to be very much alike. A frame from TV2 can be seen in Figure 3.1(b) and the density function is displayed in Figure 3.2(b). The density function shows that the image is divided in two main peaks; one representing the ground and the tracks, while the other represents the sky.

Test Video 3

Figure 3.1(c) depicts a frame from test video 3 (TV3). TV3 shows a tractor crossing a field. The camera follows the tractor so it is always in the center of the frame. The sharp peak in the density function, displayed in Figure 3.2(c),



(a) Average density function for test video 1. The average variance is 0.0240

(b) Average density function for test video 2. The average variance is 0.0275

(c) Average density function for test video 2. The average variance is 0.0276

Figure 3.2: *Density function for all of the three test videos. The density functions displayed are the average of the density function for all the frames in each of the videos*

represents the parts in the image where the sun is reflected and the dark peak is the shadow under the tractor and inside the wheels. The tractor stays in almost the same place for the whole video so there is not much motion in that part but the grass in the background is changing fast over both time and space, meaning it is expected to be difficult to compress with normal prediction. The motion in the wheels is circular so the block size has to be very small for the whole block to shift in the same direction.

To evaluate the results from the total system including both compression and transmission, peak signal to noise ratio (PSNR) was used as displayed in Equation 3.1. It is more common to use SNR or bit error ratio (BER) for evaluation of transmission schemes, but since this is a continuous amplitude modulation system, no bit errors can be determined. PSNR is preferred over SNR because it is more applicable to image quality. PSNR is one of the simplest measurements of objective image quality, but does not always give reliable results. For instance, two images that look almost the same to the human eye can have large differences in PSNR, while two images that are perceived very different can have almost the same PSNR. In spite of these weaknesses PSNR is much used due to its simplicity. Images with PSNR above 35 dB is known to have little or no degradation visible to the human eye [16]. However, since no objective measurement of image quality completely renders the human assessment of an image, subjective inspection of the resulting images should be performed as well. PSNR is given by:

$$PSNR(dB) = 10\log_{10}\frac{\max(\text{pixel value})^2}{MSE} \tag{3.1}$$

where MSE is the mean squared error of the received image.

3.2 Transmission Scheme

The iterative transmission scheme presented in Section 2.2 was simulated in several different settings; first as the ideal case, where there is no noise on the feedback channel. Then, components were added to make the simulations more realistic and to increase the performance of the system: Noise on all channels including the feedback channel, fading with channel estimation, diversity and video compression using motion compensation. In all cases, white Gaussian noise was used to model additive receiver noise. When the feedback channel was not ideal, feedback power was set to 20 or 30 dB over feed forward power in the initial simulations and 20 dB later.

The variance of the coded signal to be transmitted is a most important parameter because it is the initial value of $\sigma^2_{D_i}$ in Equation 2.38, which adjust the power outage on the channel to be that specified by E. If this variance is set incorrectly, the system might use more or less than the power specified and the simulations becomes invalid. Since the quality of the coding is dependent on the quality of the channel, this variance will not be the same for different power levels and must be found in an iterative manner for each power level. This is due to joint source channel coding. To adjust for this and get reliable results, the simulations were done at least 10 times for each sub result to let the variance converge to the correct level.

The same principle as to having the correct variance holds for the previous frame as well. When differential coding using time prediction as well is imposed on the system, the last transmitted and the last received frame is used for coding and decoding. If the system operates with a high received SNR, there will be little to no noise on the received frame and there is no problem. If the SNR is low, the previous frame has to be about the same PSNR as the one that is transmitted to get valid results. To obtain this, the simulations are repeated for several frames for each sub result. In a practical realization of the system, the previous frame and the variance will not lead to any problems since if the initial values are set relatively close their true value, they will adjust to the correct levels in a short time.

The space prediction is done around the channel noise, as can be seen in Figure 3.9. This is very important with regards to performance and is the whole clue regarding JSCC. This is the same principle as with closed loop DPCM with quantization, where the prediction is done around the quantizer so that the quantization noise is included in the prediction. When the image

is predicted in time with motion compensation, the prediction has to be done around the noise as well. This is normally more difficult since the transmitter and receiver usually have a different image. The transmitter has transmitted a frame of some sort and the receiver has received a noisy version of this frame. If the transmitter uses this perfect original to deduce the motion and the receiver uses a noisy version of the same image, the noise from the previous frame will impose itself into the next in addition to transmission noise for the frame in question. The noise will then accumulate over several frames and after transmission of some frames the only result left after decoding is noise. Since, in this system, the receiver always transmit what it received back to the transmitter over the feedback channel, the transmitter have a good idea of what image the receiver have and can use this as the previous frame to estimate the motion. The only difference between the two frames will now be the noise introduced on the feedback channel in the last iteration. Since the base station utilizes a much higher power level on the feedback, this noise is assumed to be low. The time prediction now takes noise on the channel into account for coding, thus the coding can be regarded as JSCC.

The transmission scheme was implemented so that for each sample that is sent from the helicopter, the corresponding feedback sample is received right away with no delay. If the helicopter is 1 km away from the base station, which is the target range of the company[3], the propagation delay will be

$$\begin{aligned} \tau &= \frac{2\text{km}}{c} && (3.2)\\ &= 3.33\mu\text{sec} && (3.3)\\ &\approx 23T_s && (3.4) \end{aligned}$$

where c is the speed of light and T_s is the sampling period when a bandwidth expansion of three is used. So in a practical system, the feedback sample is not received for at least $23T_s$ since there will also be a small hardware delay on both sides of the transmission. This approximation is done to make the simulations a bit less complex, but the results will still be valid since a real implementation could just transmit the first iteration for every pixel in the frame and then the next iteration for the entire frame and so on. In theory, since this is a form of time diversity, this would make the system more robust against fading. It is assumed in all iterations that the feed forward and the feedback channel are separated in frequency so there is no interference between them.

[3]See http://www.proxdynamics.com/

Since this thesis is in some parts a continuation of an earlier project, some results regarding the performance of the system have already been deduced. These results will be repeated in this work but in a shorter manner[4].

3.2.1 Fading

As argued earlier, fading will take its toll on the system as well. In most of the simulations there was assumed flat fading. This fading was modeled as Rayleigh fading using the MATLAB function *rayleighchan* with sampling period and Doppler spread as input parameters. The Doppler spread was calculated with Equation 2.50 with a specified velocity as input and sampling period T_s when transmitting 15 fps with a resolution of 240×320 equal to:

$$T_s = \frac{1}{M \times 240 \times 320 \times 15} = \frac{868\text{ns}}{M} \tag{3.5}$$

The next part will, based on the theory outlined in Section 2.3.3, describe what type of fading to expect when the helicopter is operating in the scenario described in Section 1.

In Scenario 1, the helicopter is moving at a speed of 5 m/s in an enclosed office building. In [18], Valenzuela et al. found, after many measurements, that the median rms delay spread σ_τ for an office building was 25 ns and that it very rarely exceeded 50 ns. For reasonable values of $M (M < 10)$, the signal fulfills the first requirement to expect flat fading, being that the sampling period is greater than the rms delay spread of the channel. The other requirement is that the coherence bandwidth should be greater than the bandwidth of the signal. If Equation 2.48 is used as an estimate for coherence bandwidth, the conclusion is unclear, but a rule of thumb from [17] states that if $T_s \geq 10\sigma_\tau$ flat fading can be considered and if $T_s < 10\sigma_\tau$ frequency selective fading can be expected independent of the coherence bandwidth. If the bandwidth expansion is $M = 3$ so that $T_s = 289$ ns and $\sigma_{tau} = 25$ ns, this simplification of the problem gives flat fading for Scenario 1, given a maximum bandwidth expansion of 3. In this thesis, the fading for Scenario 1 is assumed to be flat, but in a practical implementation of the system, frequency selective fading may be expected as well under some circumstances. A velocity of 5 m/s gives a Doppler spread of about 40 Hz when the sampling frequency is 2.390 GHz. By using the arithmetic mean of Equation 2.51 and 2.52, a coherence time of $T_c \approx 30ms$ is obtained. Using this together with the theory in Section 2.3.3, the conclusion is that the signal will undergo slow fading. This is a much more certain assumption than that the signal will undergo flat fading.

[4]For more details, see [7]

It is for all simulations assumed that the fading experienced on the feedback channel and the feed forward are the same. Therefore, even though the feedback and feed forward channel are separated in frequency, they have to be close enough to each other that they have the same fading. The fading will for all simulations be simulated as Rayleigh fading. Rice fading may occur as well, especially in an open environment where there is a line of sight between the transmitter and receiver. Rice fading has one strong signal component corresponding to the line of sight path, so the fades will not be as deep as Rayleigh. Rice fading is thus disregarded because Rayleigh fading gives a harsher environment and if the system performs well under Rayleigh fading, it will do so under Rice fading as well.

3.3 Channel Estimation

Small-scale path loss or multipath fading will rapidly change the amplitude of the received signal in an unpredictable fashion and cause errors even without the presence of noise. To counteract this the amount of degradation applied to the signal from the channel has to be estimated and adjusted for, in either the receiver, the transmitter or in both.

Channel estimation is normally based on the transmission of a control-signal known to the receiver. The receiver can then, based on what is received and what is actually sent, estimate the degradation the channel is applying to the signal. In this thesis, channel estimation is based on the transmission of high-powered pulses from the base station (BS) to the helicopter (HC) over the feedback channel. It is done in this order because it is assumed that BS has much more energy available than HC and can thus send pulses that are more powerful, resulting in a more accurate estimation. Since the pulses are of known amplitude to HC, HC can receive them and deduce the amount of alteration the channel is applying to the amplitude of the signal. This type of channel estimation does only estimate the alteration the channel is doing on the amplitude of the signal, and not the frequency degradation, ISI or time delay. Since the information of this transmission scheme lies in the amplitude of the signal, this is sufficient. If the ISI or the alteration in frequency should be estimated as well, the channel would have to be sampled over a longer time each time so a PDP could be determined. This would make the channel estimation more expensive since it imposes a larger delay on the system. Figure 3.3 shows the fading channel modeled as a fading factor followed by addition of noise. The amplifier G is multiplied with the signal before transmission to adjust for the fluctuations imposed by C. It is obvious that if G should counteract C, it

has to be the inverse of the channel, $G = \frac{1}{C}$, hence the channel must be known to set G correctly.

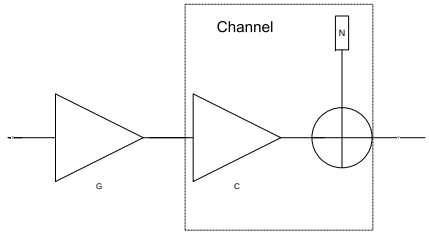


Figure 3.3: *The optimal value for G is the inverse of the fading factor C, hence the channel must be estimated to find G*

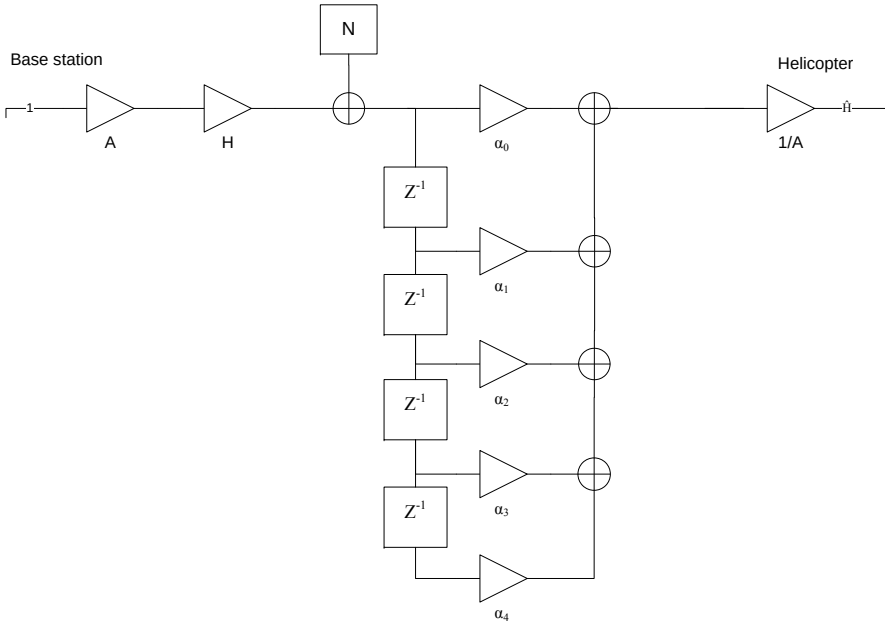


Figure 3.4: *Channel estimation with $M = 5$*

To estimate the channel a weighted sum of the $M - 1$ previous pulses and the one received was used. The previously received samples are included to get a more reliable estimate in the deep fades where the estimation is more vulnerable to noise. If a weighted sum is used, the samples must be so close

to each other that they are in the same fade so that the sum does not remove information about the channel but only noise. Between samples, the channel is predicted to avoid a step-function.

$$\hat{H}(n) = \frac{1}{A}\left[\sum_{m=0}^{M-1} \alpha_m h(n-m)\right] \tag{3.6}$$

The success of the estimation is based especially on three parameters:

- The power of the pulses. Noise will be added to the signal in HC, and the higher transmission power, the lesser impact will noise have on the estimation. This is represented by the amplitude, A in Equation 3.6
- How often the pulses are sent from BS to HC. Are they sent at a high frequency, the estimation becomes more accurate and HC can average some of the pulses to reduce the amount of random noise. Lower frequency means less use of energy, and since the rest of the system has to wait while the channel-estimation-pulses are sent, a high frequency may introduce a delay on the rest of the transmission.
- Upon receiving the pulses, HC will calculate a weighted average of several pulses to remove noise and give a more accurate estimation. If too many pulses are averaged, the large fluctuations of the channel may be lost. If too few are used, the noise may have a higher impact on the estimation. It is important that the total gain on the channel is one, $\sum_m^M \alpha_m = 1$.

The setting of these parameters is highly dependent on what kind of fading one may expect in the system. For instance, if it is slow fading the pulses can be sent a quite low frequency compared to the rest of the system. One typically only needs a single pulse to estimate the channel over the same duration as several hundred pixels can be transmitted. Since the pulses are transmitted so rarely, the power use can be quite high and one does not need to average over more than one or two pulses. On the other hand, if the channel to be estimated is a fast fading channel, the pulses has to be sent more often, which means less energy per pulse which again means more noise. In this thesis, all fading are expected to be slow, as reasoned in Section 3.2.1. Channel estimation was simulated individually to find and optimize these parameters.

It is important to note that these parameters are not separable, i.e. M cannot be found independently of A and the sampling frequency. Since this complicates the finding of the optimal values, the degrees of freedom was reduced by setting the sampling frequency first, and then the power and at last finding the best M and α_m given power and sampling frequency.

To find the sampling period of the channel, the Fourier transform was used to find the Nyquist sample frequency of the channel, while the other parameters were found by iterating over many different options to see which give the best result, meaning the least MSE between the estimation and original channel after upsampling. Upsampling was done so that the same gain should not be used on many consecutive transmissions to avoid a stepped function. This was done by predicting the values in between samples based on earlier samples and was for most values a straight line as can be seen in Figure 3.5. Because the values in the deep fades have a large difference between them, a straight line prediction between samples would give enormous errors in the deep fades, giving them values 30 dB lower than the actual value. This is because it is difficult to predict where the fade will "turn". For this reason the slope of the straight line was weighted with the value of the channel sample with 1 as maximum weight. This makes the prediction in the deep fades much more cautious.

Since the system has to stop all transmissions while the channel is estimated to avoid interference, there will be a delay on the transmission. To avoid this delay, the bandwidth has to be increased for normal transmission. It is important to note that this does not cost any energy, but decrease the capacity of the system.

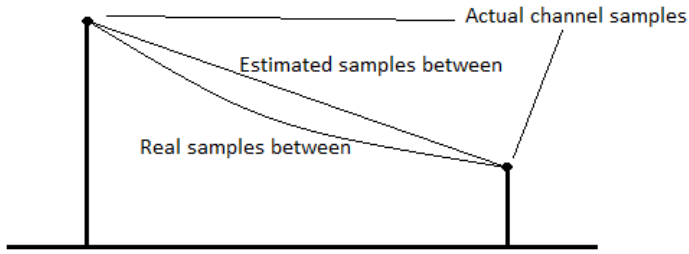


Figure 3.5: *The channel is sampled rarely, and in between the actual channel values, the channel is estimated as a straight line even though this is not entirely correct, the estimation is very close.*

Since HC does not have any CSI, the feedback channel cannot always operate exactly 20 or 30 dB higher than the feed forward channel. This is because HC adjust its transmission power based on the channel gains. In the implementations, the feedback operated with 20 or 30 dB above that of the initial power on the feed forward channel.

3.4 Channel Diversity

As explained in Section 2.4, the introduction of channel diversity will significantly reduce the impact of multipath fading on the system. Since energy and complexity are very limited resources in this system, both frequency and time diversity are not fit for use since they require twice the power. In this thesis, it is unfeasible to have more than one antenna on the helicopter due to the small size. For all of these reasons, channel diversity was obtained by simulating two antennas on the base station. Two antennas require twice the power and more space and can therefore only be fitted on the base station.

Since there are two antennas on the base station and only one at the helicopter, the transmission will be slightly different depending on whether the helicopter or the base station is transmitting. It is assumed that only the helicopter has full knowledge of how the channel alters the amplitude of the received signal so adjustment for fluctuations has to be done on the helicopter-side for both transmission and receiving. Why only the helicopter has channel knowledge and not the base station was explained in Section 3.3. The diversity scheme can be seen in Figure 3.6 and Figure 3.7.

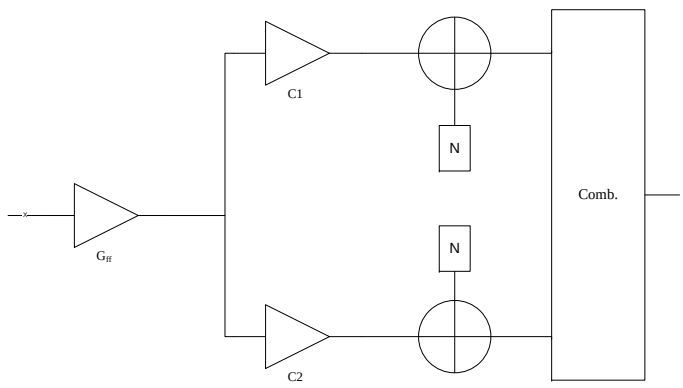


Figure 3.6: *How diversity is obtained when transmitting a signal from the helicopter to the base station. A gain is first applied to the signal to reverse the fading. Then the signal is transmitted over two independent fading channels with separate noise in each channel and a combination of the signals are received.*

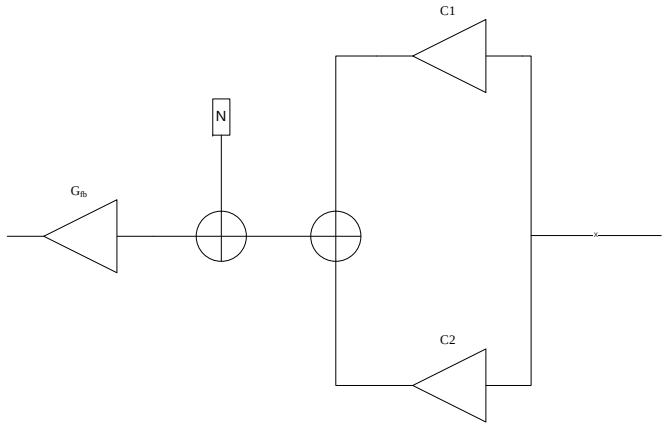


Figure 3.7: *How diversity is obtained when transmitting a signal from the base station to the helicopter. The helicopter receives the sum of the two transmitted signals. A gain is then applied to adjust for fading.*

At the base station, two separate signals are received, and how to combine those two is up to the receiver. As explained in Section 2.4, the optimal way to combine them is Maximal ratio combining, but since this requires channel knowledge at the receiver, it cannot be used in this system since only the helicopter has channel knowledge. The combining scheme used in the simulations is instead the sum of the two signals with equal gain. This is called equal gain combining and has a performance marginally inferior to maximum ratio combining[17]. It was assumed that the two signal branches were co-phased, so they could be added directly without a synchronization signal.

When transmitting from the base station to the helicopter, the gain to adjust for fading is applied in the helicopter after noise has been added. This will enhance the noise in the same manner as the signal so that the SNR is not changed after the gain, but the total PSNR for the whole frame is changed to the better because fluctuations from fading will be removed. Because of this, in the deep fades, the SNR on the feedback channel can decrease to a value below that of the feed forward. When this happens, the helicopter should not take the signal on the feedback channel into account, and only transmit zeros after the first iteration. This is because, in this case, the feedback channel is less reliable than the feed forward and should therefore not be used. The two gains G_{ff} and G_{fb} (ff for feed forward and fb for feedback), are calculated so that they try to cancel the fading imposed by the channel. The calculations

are done without the impact of noise, and since the combining scheme is the sum, the two gains will be the same. If $\hat{x}$ is the received signal, and x is the transmitted signal so that

$$\hat{x} = G(xC_1 + xC_2) \tag{3.7}$$

To find G, equate $\hat{x}$ and x to obtain

$$G = \frac{1}{C_1 + C_2} \tag{3.8}$$

From Equation 3.8, the gain from introducing channel diversity can be seen intuitively. If there was only one channel, the gain would be $G = 1/C$, which would one average be much higher than the one in Equation 3.8.

If both C_1 and C_2 are small, G will be very large. Since the helicopter has a very limited amount of energy, a maximum limit for the size of G_{ff} has to be set. This limit should be large enough to compensate for the small fluctuations in the amplitude but not so high that is tries to compensate when both the channels are in a deep fade and consume all the energy. In the simulations, this limit was 3 times the input energy and was chosen so as a trade-off between energy consumption and performance.

$$G_{ff} = \max\left[\frac{1}{C_1 + C_2}, 3E_s\right] \tag{3.9}$$

In the deepest fades, the transmitted signal is reduced to nearly zero even when the helicopter is transmitting at $3E_s$, and if the power level is reduced sufficiently, the helicopter wastes energy by transmitting and should not transmit anything at all. G_{fb} does not need this limitation because the signal can be adjusted in the receiver with negligible use of energy. Since both G_{ff} and G_{fb} are dependent on the channel gains C_1 and C_2, the channel has to be estimated. The helicopter does not need to estimate the two different channels separately. Since both the gains needed (G_{ff} and G_{fb}) are based on the sum of the two channels, this is all that needs to be estimated. Therefore, the channel estimation can still be done as described in Section 3.3, only difference is that the base station now transmits on two antennas instead of one and the energy is doubled.

3.5 Video Compression

The parameter η in Equation 2.63 was in the simulations set to 0.98. This is because the correlation between consecutive frames are expected to be high

when they are adjusted for motion, but the factor should still be less than 1 to keep the system stable.

The frames are sent pixel by pixel, row by row. For each pixel, if there does not exist a MV for that block, it has to be found before continuing transmission. The MV's can be found by searching through the previous frame around the position of the block in the original frame to find the position that yields the least squared error between them. The difference between the positions of the top left pixels in the two blocks are then stored as the motion vector for that block and transmitted to the base station. In the simulations, a "brute-force" motion estimation algorithm was used. Within a search window of 15 pixels in any direction from the original block position, the algorithm tries every possibility to find the shift that minimizes the MSE between two blocks. This is time consuming, and in a practical implementation, another algorithm should be used. Even though the algorithm used is unrealistic, the simulations will be valid since a faster algorithm is still expected to give good results. The block size for all simulations was set to 8×8 pixels.

Equation 2.66 showed that the optimal coefficients for space prediction depend solely on the autocorrelation function of the image. Since the autocorrelation has to be calculated many times for one single frame to give reliable results, it is unfeasible for a system that requires low complexity to use these optimal values. So in this system static values must be used. In [8], Jain found that these values gave good results for prediction of natural images:

$$\alpha = \beta = 0.95 \tag{3.10}$$

and

$$\rho = -0.9 \tag{3.11}$$

where the parameters are in correspondence with Equation 2.64. These values are for images picturing natural events. The space prediction in this thesis is predicting an image that is already compressed via time prediction and can no longer be considered a natural image. Therefore, new static values had to be estimated. To do this, time prediction was done on all frames in the three test videos and then the autocorrelation was calculated and optimal coefficients for every frame, was calculated and averaged for every video to get new parameters. Since the image is not considered ergodic will the parameters calculated with an ergodic assumption not yield the optimal parameters for every part of the image, but on average they will be the best.

The simulations of video compression were first done with only prediction in space. Here too was optimal parameters calculated for every video and the results was expected to be close to Equation 3.10 and 3.11. Then prediction in time only was briefly examined before prediction, done in space and time together, was reviewed. The space prediction coefficients used after the time predicted signal had been subtracted are expected to be smaller than to one in Equation 3.10 and 3.11 since there is less redundant information left in the image.

3.6 Path Loss

This section will calculate an estimated transmission power needed for the scenario presented in Section 1, by using the theory explained in Section 2.3.2. The calculation is based on many assumptions and will work as a pointer to what transmission power is needed to reach a certain SNR in the receiver.

In Scenario 1, the helicopter is operating indoor in an office building. By using Equation 2.47 with $n = 3$ and $\sigma = 7$, which is the values for n and σ for an office building[2], together with Equation 2.43, one can calculate a realistic path loss at distance d and reference distance $d_0 = 1\text{m}$. The carrier frequency is set to 2390 MHz and antenna gains G equal to one.

$$\begin{aligned} PL(\text{dB}) &= -10\log\left(\frac{\lambda^2}{16\pi^2 d_0^2}\right) + 10n\log\left(\frac{d}{d_0}\right) + X_\sigma && (3.12)\\ &= -10\log\left(\frac{126\times 10^{-3}}{16\pi^2}\right) + 30\log(d) + X_7 && (3.13)\\ &= 31 + 30\log(d) + X_7 && (3.14) \end{aligned}$$

Using this, one can continue to deduce the average needed transmission power to reach a certain SNR at a distance d meters between transmitter and receiver. Since about 98% of the values from X_7 will be less than $2\sigma = 14\text{dB}$, and since this calculation is interesting for the worst case of X_7, $X_7 = 14\text{dB}$. Since

$$P_r(\text{dB}) = P_t - PL \tag{3.15}$$

and

$$N(\text{dB}) = 10\log(4k_bTB) \tag{3.16}$$

where k_b is Boltzmann's constant $k_b \approx 1.38\times 10^{-23}$, T is the temperature in the receiver and B is the bandwidth of the signal, receiver SNR(dB) can be

calculated. Using that $B = 3.456$ MHz and that the receiver operates at room temperature, $T = 300$K:

$$\begin{aligned} SNR(D)(\text{dB}) =& P_r - N && (3.17)\\ =& P_t - PL - N && (3.18)\\ =& P_t - 45 - 30\log(d) - \\ & 10\log\left(4k_b 300 \times 3.456M\right) && (3.19)\\ =& P_t - 30\log(d) + 87 && (3.20) \end{aligned}$$

hence,

$$P_t = SNR + 30\log(d) - 87 \tag{3.21}$$

By inserting a distance and the wanted SNR into Equation 3.21, one can calculate the expected needed transmission power. Transmission power versus receiver SNR can be seen in Figure 3.8. Any lowering of the received SNR caused by fading is not included in this derivation.

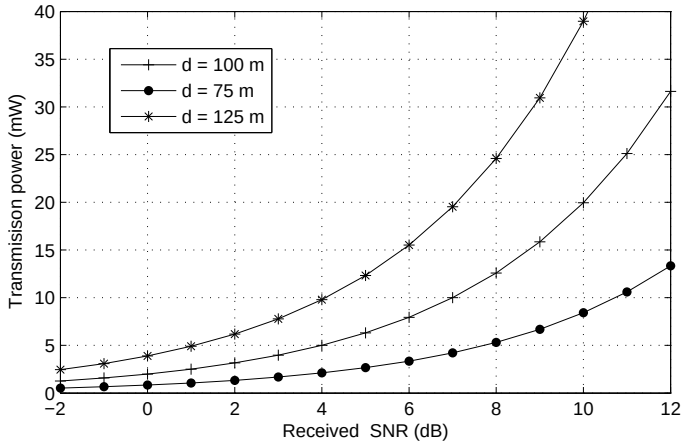


Figure 3.8: *Transmission power needed to reach a certain receiver SNR according to Equation 3.21. The transmission power is in milliwatts and the SNR is in dB*

3.7 Total System

Since the system is made up from many different components, this section will clarify how they work together. The total system with fading, channel diversity and video compression with motion compensation is viewed in Figure 3.9. Note that the sampling frequency is different on the different branches. This is how the system was implemented.

The gains A', B', C' and D' are calculated in the same manner as A, B, C, D for the transmission system described in Section 2.2, but with different energy E and variance σ_D^2. It is assumed that the channels are the same for all the transmissions. Also note there is not actually more than one feed forward or feedback channel. It is the same channel used every time. So when the motion vectors are transmitted in the outer loop, the rest of the system has to wait. This will increase the actual bandwidth of the system some so capacity is lost, but no delay is experienced. The channel estimation will also increase the bandwidth, but this is not included in Figure 3.9.

After the motion vectors are received by the base station the result are quantized to get integer numbers. In the simulations, this was only a rounding operation, rounding the received signal to nearest integer.

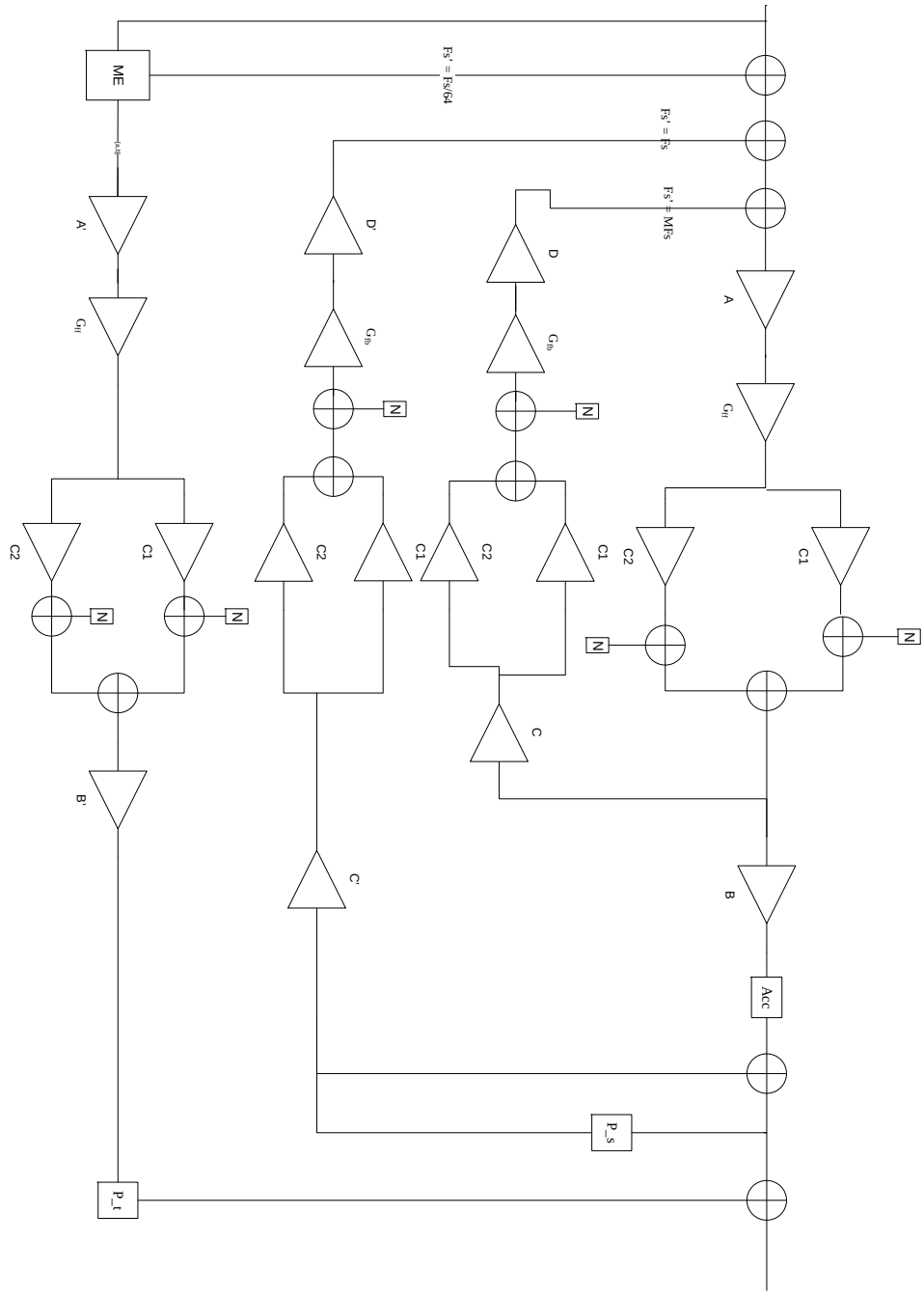


Figure 3.9: *Total system reviewed in this thesis. Note that the sampling frequencies are different for different parts of the sketch*

CHAPTER 4

RESULTS AND DISCUSSION

This chapter will present the results obtained in the simulations and discuss them with emphasis on power efficiency, use of bandwidth, complexity and performance. First, the results regarding channel estimation will be presented followed by video compression. At last the transmission scheme and whole system is simulated. In these simulations, the parameters obtained from both channel estimation and video compression were used.

4.1 Channel Estimation

The channel estimation is done in accordance with Figure 3.4 and Equation 3.6. The aim is to find fitting parameters for A, M, α_m and how often the channel must be sampled to yield good results. Good results are results that require a small amount of energy (low amplitude) and low complexity (imposes only a small delay on the system, meaning the channel should be estimated as rarely as possible) and still give a good estimation of the channel state. Since these parameters are assumed to be static in the operation of the helicopter, the simulations are only done for one channel, assuming the helicopter is traveling at max speed equal to 10 m/s[1], use a bandwidth expansion of 3 and is experiencing flat fading. The fading is simulated as Rayleigh fading. The channel that was used in the simulations is viewed in Figure 4.1.

[1]This is the maximum speed the helicopter can achieve and will thus give the fastest fading. See `http://www.proxdynamics.com/products/` for product details.

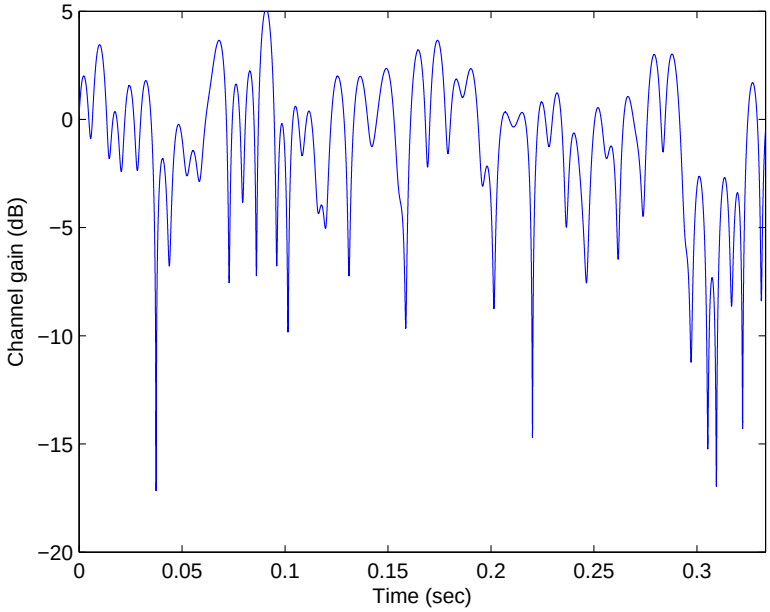


Figure 4.1: *The flat Rayleigh fading channel used in the channel estimation simulations. It was created assuming a helicopter velocity of 10 m/s and a sampling frequency corresponding to a bandwidth expansion of 3, with transmission of* $15 \times 320 \times 240$ *pixels per second. It is showed over a period of 5 frames*

When deciding how many times the channel should be estimated or sampled per second, the trade-off between power, delay and estimation accuracy should be considered. Since the pulses are sent from the base station, which is assumed to have enough power available, the pulses are simulated with a receiver SNR of 20 dB. A receiver SNR of 20 dB is only expected when the channel gain is zero dB, and will be smaller when the channel is in a fade, degrading the estimation for deeper fades. To find how many times per second the channel has to be estimated or sampled, the Fourier transform was applied to the channel as a whole and examined. The frequency response of the channel in Figure 4.1 can be seen in Figure 4.2. It has been normalized so that the maximum value is at zero dB. The Nyquist sampling theorem states that a sampling frequency of at least $F_s = 2f_m$ is needed for perfect reconstruction of the signal, where f_m is the maximum frequency of the signal to be sampled. The maximum frequency is the highest frequency where the signal still has energy. In practical matters, -30dB or -40dB is as good as zero and if the signal was assumed to have zero energy when it never again crossed this limit,

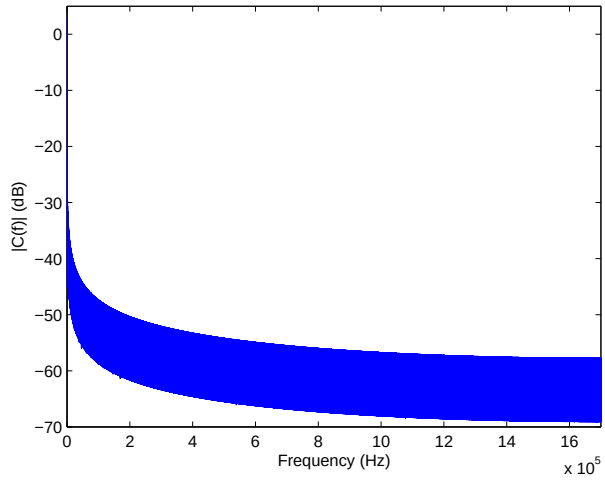


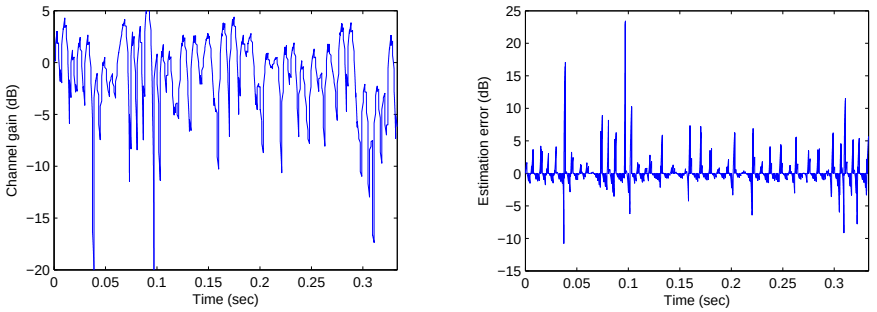
Figure 4.2: *The normalized single sided amplitude spectrum of the fading channel in Figure 4.1*

the results in Table 4.1 was obtained. The interval column holds the most interesting numbers, showing how many pixels that can be transmitted before the channel has to be estimated again.

Table 4.1: *Maximum frequency obtained from the frequency response in Figure 4.2, when* limit *was used as zero. Interval is how many pixels that can be transmitted before the channel must be estimated again.*

Limit	f_m	Interval
−20 dB	335 Hz	1722 pixels
−25 dB	766 Hz	752 pixels
−30 dB	1.9 kHz	305 pixels
−40 dB	18 kHz	31 pixels

Simple simulations showed that when only estimating the channel for every 1722 pixel, as was the result from using -20 dB as zero, the channel is significantly undersampled. The accuracy is lost for large parts of the channel and the deep fades are sometimes removed completely. This can be seen in Figure 4.3, especially that the deep fades disappear completely. The estimation error plot shows that there is many large errors throughout the channel. With a sampling frequency of 1532 Hz, corresponding to using -25 dB as zero, there is still some small visible artifacts, but they are mostly lost in the very deep



(a) *The channel in Figure 4.1 sampled with a sampling frequency of 670 Hz*

(b) *Estimation error in dB*

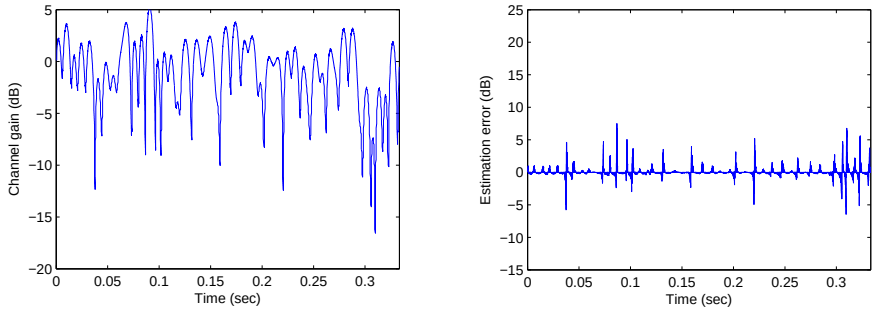
Figure 4.3: *The channel estimation with a sampling frequency of 670 Hz and the estimation error compared to the original channel. The MSE of the estimation was 0.04*

fades, and far from as much as is the case with a sampling frequency of 670 Hz. Since the system cannot compensate for the loss in power in the deepest fades anyway, and the amount of delay that needs to be added to get a good estimate in the deepest fades as well are considered high; this sampling frequency is deemed appropriate and is used for the rest of the simulations.

The next two parameters that affect the accuracy of the channel estimation is how many samples that are averaged, and the value of the weights in the averaging. This is M and α_m in Figure 3.4. A script that tested many different values for both M and α was implemented and the result was that the best approach is to not use any averaging at all, meaning $M = 1$ and $\alpha_m = 1, \forall m$. This is due to the fact that the channel is still a bit undersampled and two or more consecutive samples are not very correlated. It can also be explained by the high use of power, removing much degradation imposed by noise so the actual received sample has less artifacts than the combination.

After each sample is received, the channel gains for the next 752 pixels are predicted to give a smoother and more correct estimation. The final results after simulating transmission of a high powered pulses 1532 times per second of known amplitude with additive white Gaussian noise, followed by prediction of the next samples are depicted in Figure 4.4. The mean squared error between the estimate and the actual channel is 0.0025. Because of the noise, the deep fades have a much smaller SNR in the receiver and are affected more severely than the rest of the channel. Notice that the estimation error is significantly

reduced in comparison to the one in Figure 4.3, even for the deep fades.



(a) *The channel in Figure 4.1 sampled with a sampling frequency of 1532 Hz*

(b) *Estimation error in dB*

Figure 4.4: *The estimated channel based on the original in Figure 4.1 with no averaging and a sampling frequency of 1532 samples per second. The channel has been interpolated between samples to have the same resolution as the original. Note that noise has been added to the transmission as opposed to the one in Figure 4.3. The MSE of the estimation is 0.0025. The estimation error is also included for comparison to the one in Figure 4.3*

4.2 Video Compression

First the performance of the compression by applying only prediction in space is reviewed. This is done for the reader to note the difference when motion estimation and time prediction is added to the system. It is also important to note that it can be useful to have a single-frame-compression scheme in the system if this is wanted for synchronization.

Many of the results are viewed using histograms or density functions showing how the pixel values in the coded image are spread out. The aim of compressions in this case is to lower the variance so one can use less power for transmission, since use of power is directly related to the variance (see Equation 2.38). The width of the density functions are directly related to the variance of the system. This is because the density functions can be seen as a low resolution PDF of the image.

According to Equation 2.66, the prediction coefficients, α, β and ρ, are solely based on the autocorrelation of the image. Table 4.2 shows the calculated averaged optimum prediction coefficients from each of the three test videos,

Table 4.2: *Optimal space prediction parameters for each of the three test videos and the mean*

	TV1	TV2	TV3	MEAN
α	0.89	0.72	0.63	0.75
β	0.87	0.78	0.70	0.78
ρ	-0.77	-0.51	-0.35	-0.54

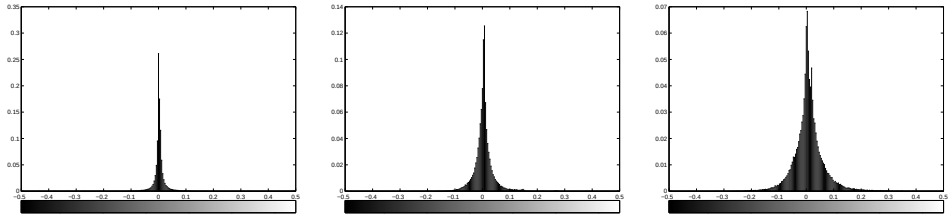
and the density functions retrieved when applying these parameters can be seen in Figure 4.5. The coefficients for TV3 are significantly smaller than that of TV1 and TV2, which means there is less correlation between neighboring pixels in TV3, and thus, will be more difficult to compress to the same degree compared to the other two. This is related to the background consisting of grass and earth, which vary much more than the monotone background of TV1 and TV2. This can also be seen in the density functions where TV3 has a much larger spread of the values after coding than TV1 and TV2. Averaged variances of the prediction error after using different coefficients can be seen in Table 4.3. From both the table and the density functions it is evident that the compression has the least effect on TV3.

The transmission scheme is optimal when the source input is Gaussian distributed, but as seen in the density functions, this is not the case. The density functions look much more like they are Laplacian distributed, which often is used to model prediction error [6].

Table 4.3: *Variance of prediction error when different parameters were used for space prediction.*

	TV1	TV2	TV3
Optimal parameters from Table 4.2	5.91×10^{-4}	1.05×10^{-3}	2.9×10^{-3}
Mean parameters from Table 4.2	9.07×10^{-4}	1.49×10^{-3}	3.56×10^{-3}
Parameters usually used for natural images from Equation 3.10 and 3.11	6.45×10^{-4}	1.10×10^{-3}	2.9×10^{-3}

From the results in Table 4.3, it it obvious that the parameters found in [8] for normal images give almost as good results as the ones calculated via the autocorrelation function, and on average they give a better prediction than is obtained using the mean of the optimal parameters.



(a) *Average density function of prediction error for TV1 with optimal space prediction coefficients. The average variance of the prediction error was* 5.91×10^{-4}

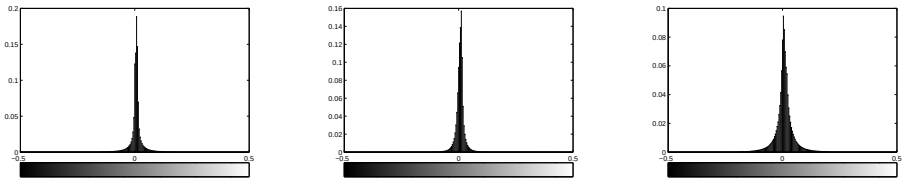
(b) *Average density function of prediction error for TV2 with optimal space prediction coefficients. The average variance of the prediction error was* 1.05×10^{-3}

(c) *Average density function of prediction error for TV3 with optimal space prediction coefficients. The average variance of the prediction error was* 2.9×10^{-3}

Figure 4.5: *Density function of the prediction error whit only space prediction and parameters are the optimal from Table 4.2*

Now, compression using only prediction between consecutive frames will be studied, i.e. time prediction alone, without any prediction in space. In Figure 4.6, the average density functions of the prediction error for each of the test videos are depicted. The differences between the histograms are not so clearly visible as was the case when prediction was done only in space. TV3 has higher variance after compression than the other two but TV2 has the lowest in contrast to when prediction was done in space only, where TV1 had the lowest variance. This is due to the fact that TV3, picturing a moving tractor, has a circular motion on the large wheels. This circular motion makes parts of the wheel move faster than others and since the block size is not very small compared to the frame resolution, the system cannot compensate for the motion. TV1 has large parts in the same level of gray and very smooth motion. A significant part of the prediction error for TV1 is on the edges of the large objects and the blocks containing components from both background and foreground which has different motion parameters. TV2 experiences a very slow motion where new parts of the image is slowly added by zooming out and the difference between two frames even without motion estimation is minimal.

If prediction is done in both time and space, which is the main compression method used in the system, space prediction parameters has to be found. As stated in Section 3.5, after time prediction, the images cannot be regarded as natural images so standard parameters cannot be used. The parameters found, using the same method as when only space prediction was regarded, can be



(a) *The average variance of the prediction error for TV1 was* 7.83×10^{-4} (b) *The average variance of the prediction error for TV2 was* 3.44×10^{-4} (c) *The average variance of the prediction error for TV3 was* 1.52×10^{-3}

Figure 4.6: *Average density functions of prediction error when only time prediction was used*

found in Table 4.4.

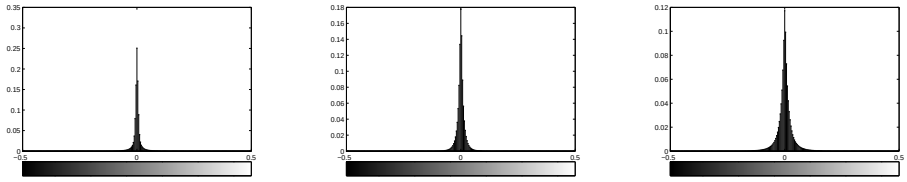
Table 4.4: *Optimal space prediction parameters for each of the three test videos and the mean. These parameters should be used when the time predicted image has been subtracted*

	TV1	TV2	TV3	MEAN
α	0.79	0.69	0.58	0.69
β	0.72	0.64	0.65	0.67
ρ	-0.54	-0.37	-0.32	-0.41

The mean parameters in Table 4.4 are the parameters used for simulations regarding the whole system, but the density functions of the videos coded with their individual optimums are also included here and can be seen in Figure 4.7.

Table 4.5: *Variance of the prediction error when different prediction techniques were used. The prediction coefficients for time was 0.98 and for space it was the mean values from Table 4.2 and Table 4.4. The variances for the test videos without any coding are also included*

	TV1	TV2	TV3	MEAN
No coding	24×10^{-3}	27.5×10^{-3}	27.6×10^{-3}	26.4×10^{-3}
Space only	6.41×10^{-4}	1.10×10^{-3}	2.90×10^{-3}	1.55×10^{-3}
Time only	7.83×10^{-4}	3.44×10^{-4}	1.52×10^{-3}	8.82×10^{-4}
Both space and time	2.97×10^{-4}	2.92×10^{-4}	1×10^{-3}	5.3×10^{-4}

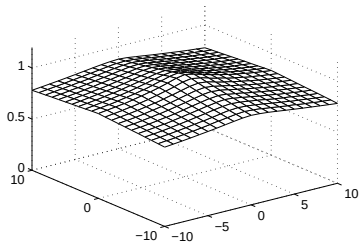


(a) *The average variance of the prediction error for TV1 was* 3.42×10^{-4}

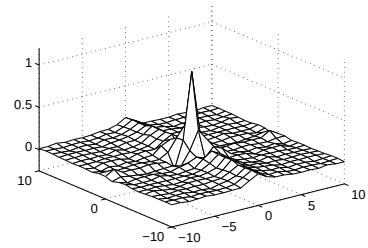
(b) *The average variance of the prediction error for TV2 was* 2.89×10^{-4}

(c) *The average variance of the prediction error for TV3 was* 9.74×10^{-4}

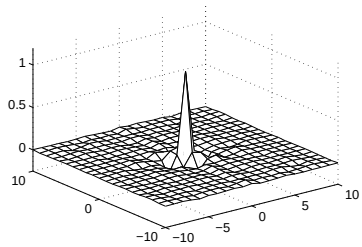
Figure 4.7: *Average density functions of prediction error when prediction was done in both time and space. The space prediction coefficients that were used was in correspondence with the individual optimums from Table 4.4*



(a) *The uncompressed frame*



(b) *Prediction in space only by using the mean parameters from Table 4.2*



(c) *Prediction in time and space by using the space prediction parameters from Table 4.4. Time prediction parameter was 0.98*

Figure 4.8: *The two-dimensional autocorrelation functions for the 10 closets samples in space. The input was one frame from TV1.*

In Table 4.5, the results from using different prediction variants are summarized. Time prediction lower the average variance of test videos compared to just space prediction except for TV1 where it is slightly increased. Time prediction is a much more complex compression than space prediction due to the search to find the correct motion vectors. This can, as stated earlier, be done relatively fast, but compared to prediction in space with static parameters it is a time consuming operation, which in addition is being done on the helicopter's side who has limited computer power. So if the gain from applying time prediction in the coder as well did not reduce the variance much, it could be argued that the gain in lower variance is lost due to high complexity. Using only time prediction makes no sense in this system since there is already an established feedback channel and the extra cost of applying prediction in space as well is minimal. So the interesting variances to compare are those with only space prediction and those where both time and space prediction is used. Of course it is important to compare with the variance of the uncoded videos as well, but this is very much the same for all videos. The gain of using prediction in both time and space compared to in just space is 4.67 dB. This implies that E_p/N_0 can be reduced by almost 5 dB if time prediction is applied as well and still get the same PSNR. Since these simulations are performed without the presence of noise, this gain is the best-case and will only hold for high values of E_p/N_0 where there is only a small amount of noise on the prediction error. This gain is significant and makes up for the extra complexity added by the motion estimation. On average, the gain from compressing the videos compared to transmitting them uncoded is 17 dB. To see how much correlation is removed from the signal, the two-dimensional autocorrelation functions were calculated for the uncompressed frame, the compressed signal using space prediction and the compressed signal using prediction in both time and space and are depicted in Figure 4.8. The frame has a very even autocorrelation function, almost resembling a uniform function, suggesting there is very much correlation between pixels even as much as 10 pixels away from each other which is the range the autocorrelation is plotted over. When space prediction is applied to the frame, the correlation is reduced significantly and there is one peak which quickly decreases in the center. There is still some correlation left in the image even several pixels away, mostly in horizontal direction, but also some in vertical direction. When time prediction is added as well, there is almost just one peak and the rest is zero. If the signal was completely uncorrelated there would just be one peak resembling the dirac delta function at position (0,0).

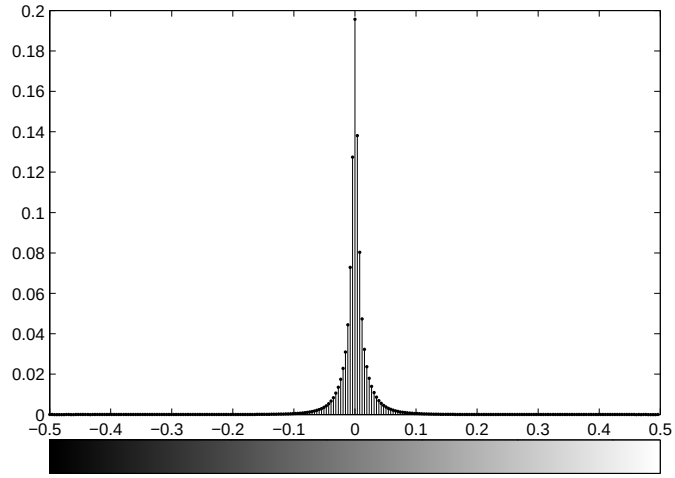


Figure 4.9: *The combined average density function of the prediction error for all three test videos. The parameters used for space prediction was the mean values in Table 4.4*

In Figure 4.9, the average density function of the prediction error for all three test videos is plotted. The space prediction parameters used was the same that will be used for the rest of the system simulations. It is interesting to note that about 20 percent of the samples are exactly zero, which implies no power use at all.

4.3 Total System and Transmission Scheme

The results obtained in the simulations of the transmission scheme and the total system will be presented in this section. A PSNR of 35 dB is known to have little or no degradation visible to the human eye and is thus used as a reference on many of the plots[16]. This limit is for single images and may not hold for videos since even small amounts of noise that is replaced 15 times per second may be visible and cause flickering even though it is difficult to spot in a single frame. For coding on the single-frame "Lena", which is the input in some of the first plots, 35 dB PSNR is used as reference for almost error-free transmission and is confirmed by the subjective opinion of the author.

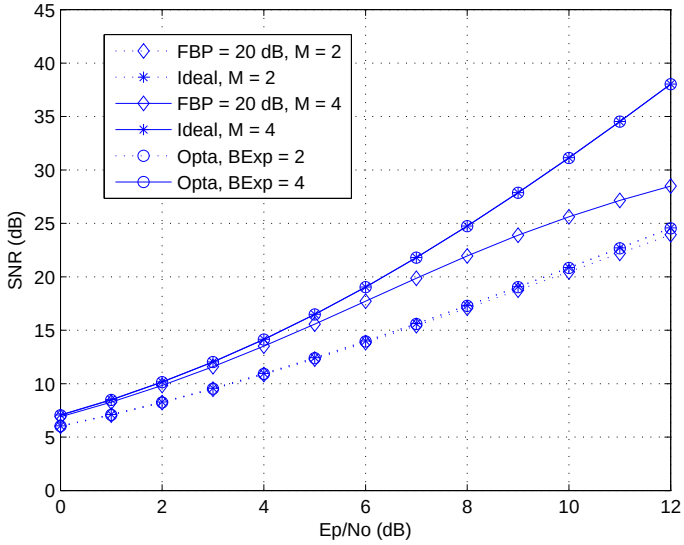


Figure 4.10: *Performance of the transmission scheme with ideal feedback with bandwidth expansion 2 and 4 is compared to OPTA and the transmission scheme with noise on the feedback channel. The transmission power on the feedback channel was 20 dB higher than that on the feed forward and the input was Gaussian distributed*

In Figure 4.10, the iterative transmission system with a Gaussian source as input and without the differential coder is shown. As stated in Section 2.2, the iterative analog system is optimal when the input and noise distributions are Gaussian and the feedback is ideal. The OPTA curve is indistinguishable from the curve from the system with ideal feedback, verifying the optimality of the system. It is also noticeable that at low power, the performance with a non-ideal feedback channel is close to OPTA, but diverges from the OPTA curve as power increases.

Results from the system with feedback power 20 dB above feed forward power is shown in Figure 4.12. A PSNR of 35 dB is reached at $E_p/N_0 = 6$ dB with 5 iterations. This is 3 dB higher than in the ideal feedback case viewed in Figure 4.11. It is seen from the plot that the gain from having 5 iterations compared to 4 is minimal and when $E_p/N_0 > 10$ dB, 4 iterations is actually better than 5 and when $E_p/N_0 > 13$ dB, it seems 3 iterations is better than 4. An explanation for this behavior is that when an ideal feedback channel is used, every transmission in the system is optimal, but as the feedback signal is corrupted with more and more noise, i.e. decreasing feedback power, high

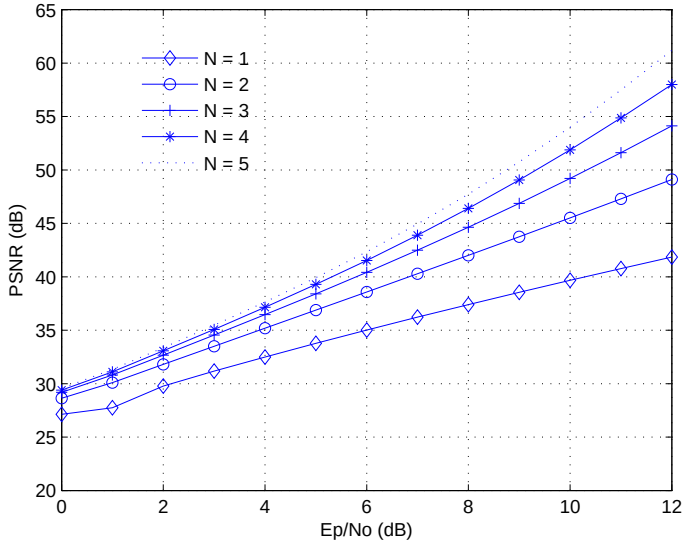


Figure 4.11: *PSNR (dB) as a function of E_p/N_0 (dB) with different numbers of iterations(N) for the image "Lena" with ideal feedback channel, i.e. no noise on feedback channel*

numbers of iterations actually accumulates noise to the signal. Thus, with a non-ideal feedback channel a number of iterations will have a corresponding maximum obtainable PSNR, depending on the feedback power. This thesis is primarily concerned with results obtained for low power, and at low E_p/N_0 it is clear that increasing number of iterations increases PSNR.

In Figure 4.13, the results for the system is shown, now with a feedback power 30 dB above feed forward power. In this case a PSNR of 35 dB is reached at $E_p/N_0 \approx 5$ dB for 3, 4 and 5 iterations. This is 1 dB lower than when the feedback channel had a power of 20 dB more than the feed forward channel. It is clear that the results from using 30 dB more power on the feedback channel resembles the results from using the ideal feedback channel more closely than 20 dB, which is expected. As feedback power goes to infinity the result would be equal to the ideal feedback channel case. It takes as much power as $E_p/N_0 = 16$dB for 4 iterations to exceed 5 in respect to PSNR in this case.

The results in Figure 4.11, 4.12 and 4.13 are all shown with a curve corresponding to no bandwidth expansion, i.e. $N = 1$. From these results, it is clear

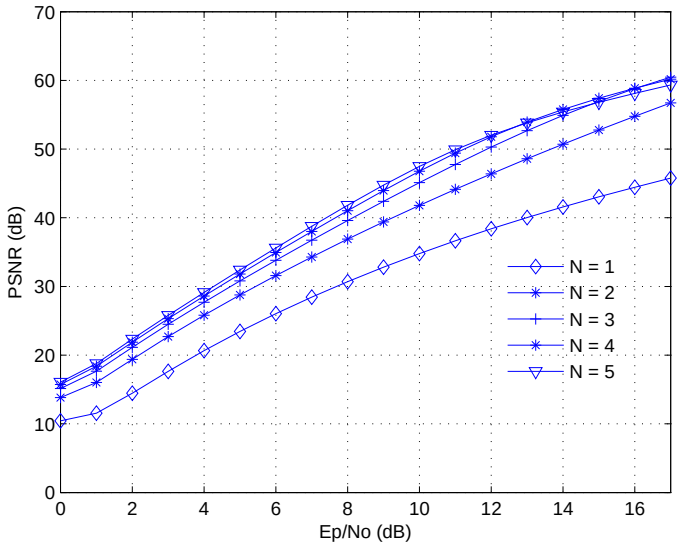


Figure 4.12: *PSNR (dB) as a function of E_p/N_0 (dB) with different numbers of iterations(N) for the image "Lena", with noisy feedback and feedback power 20 dB above feed forward power*

that bandwidth expansion offers a considerable performance gain, at a certain power input, as was predicted and which actually was the motivation behind the implementation of the system in this thesis. Since power is a parameter of the exponential function describing SNR (Equation 2.4), the difference in PSNR between no bandwidth expansion and bandwidth expansion will only increase with increasing power input, when an ideal feedback channel is used. As mentioned above, a certain number of iterations have a corresponding maximum obtainable PSNR, when using a non-ideal feedback channel. Thus, it is expected that at high power inputs the performance of no bandwidth expansion will catch up with the performance of bandwidth expansion. But at what power level this happens is dependent on the power of the feedback channel. In the case of a feedback channel with power 5 dB above feed forward power, it was found that this occurs at $E_p/N_0 = 120$ dB.

It is clear that for both cases with non-ideal feedback channel that has been reviewed, the gain from using 4 or 5 iterations, at low power, compared to using 3 is minimal. But the gain from using 3 compared to 1 or 2 is high. Since each iteration increases the bandwidth of the signal with $B = 240 \times 320 \times 15 = 1.152$ MHz, this gain has a cost, and in the rest of the results a bandwidth expansion

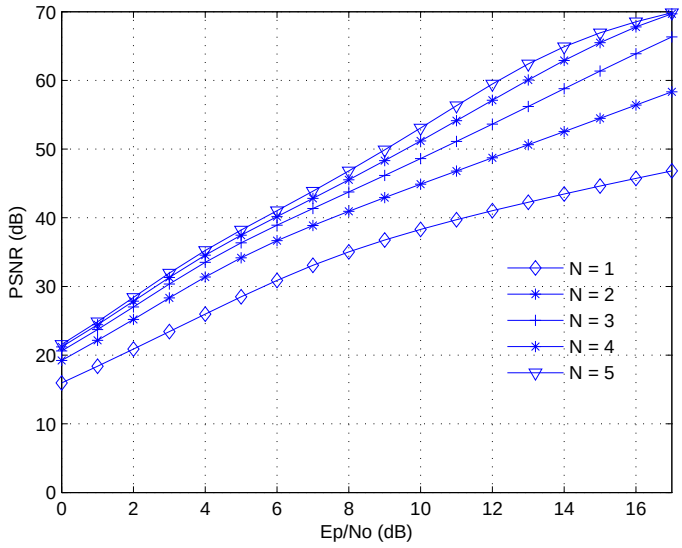


Figure 4.13: *PSNR (dB) as a function of E_p/N_0 (dB) with different numbers of iterations(N) for the image "Lena", with noisy feedback and feedback power 30 dB above feed forward power*

of 3 was used. The gain from having feedback power 30 dB above feed forward compared to having 20 dB is as stated 1 dB in E_p/N_0 to reach a PSNR of 35 dB. This gain is small compared to extra amount of energy used by the base station and therefore, the feedback channel is set to be 20 dB higher than the feed forward in the majority of the remaining simulations.

When the three test videos described in Section 3.1 are input to the communication system, the simulations are done for several different cases, where components are added to the system gradually to either make the system more realistic or to increase performance. The system is reviewed and compared for the four following cases. The name of the case are listed in **bold** and is used to reference that case for the remaining discussion. Feedback power are for all cases 20 dB above that of the feed forward and bandwidth expansion is 3.

1. **Noise**. Prediction is only done in space using the mean parameters from Table 4.2 and the channel is modeled as an AWGN channel on both feed forward and feedback channel.

2. **Fading**. Flat fading and channel estimation is added to the the system to make it more realistic. Except from fading and channel estimation it

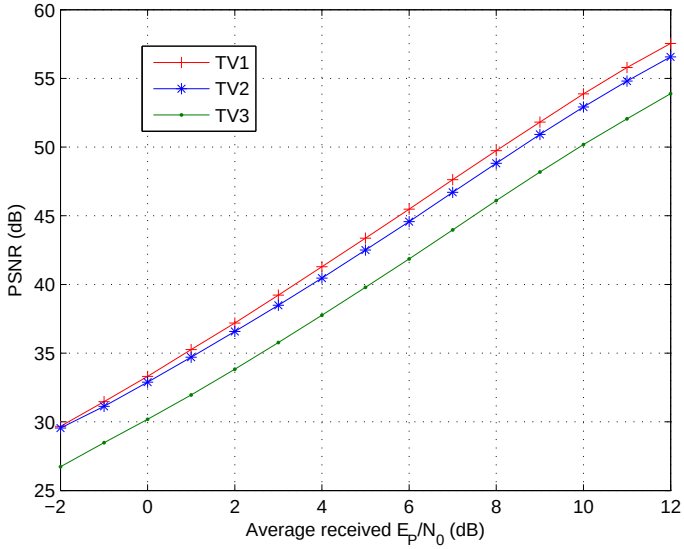


Figure 4.14: *PSNR vs* $\frac{E_p}{N_0}$ *for the transmission scheme for all test videos for the* ***Noise****-case. The mean space prediction parameters from Table 4.2 was used for prediction. The feedback channel had a power 20 dB higher than the feed forward and there was no fading on the system*

is the same as the previous case.

3. **Diversity**. Two antennas are implemented on the base station side of the system to make it more robust against the fading introduced in the previous case.

4. **3d-dpcm**. Compression is based on prediction in both space and time. The space prediction parameters are the mean parameters from Table 4.4. Diversity, fading and channel estimation are still implemented.

Since the input now is a video and not a single-frame-image, 35 dB PSNR may no longer be suitable to reference acceptable quality, because, as argued earlier, small amounts of noise that shifts 15 times per second are much more visible than the same noise density in a single-frame. Further, the 35 dB limit should be confirmed subjectively even for images and may give different subjectively results depending on the input. This assumption is confirmed subjectively by the author by examination of the video outputs. This can also be, to some degree, seen in the example images depicted in Appendix A. For the simulations where the test-videos are input, 40 dB PSNR is used as reference to when the received video quality is acceptable.

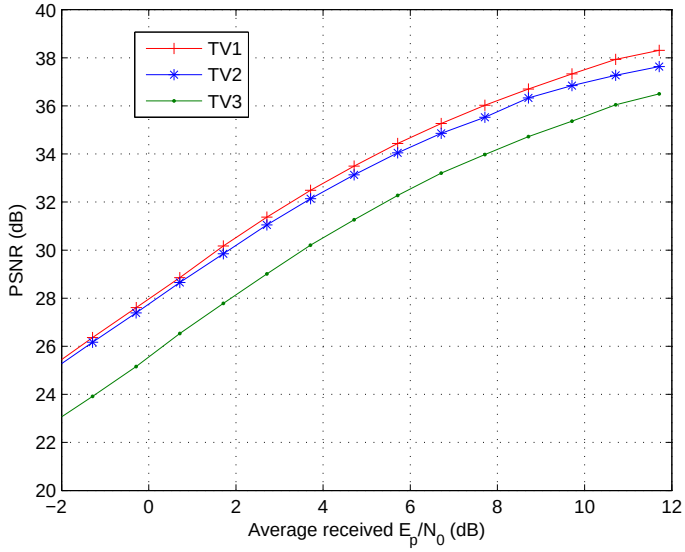
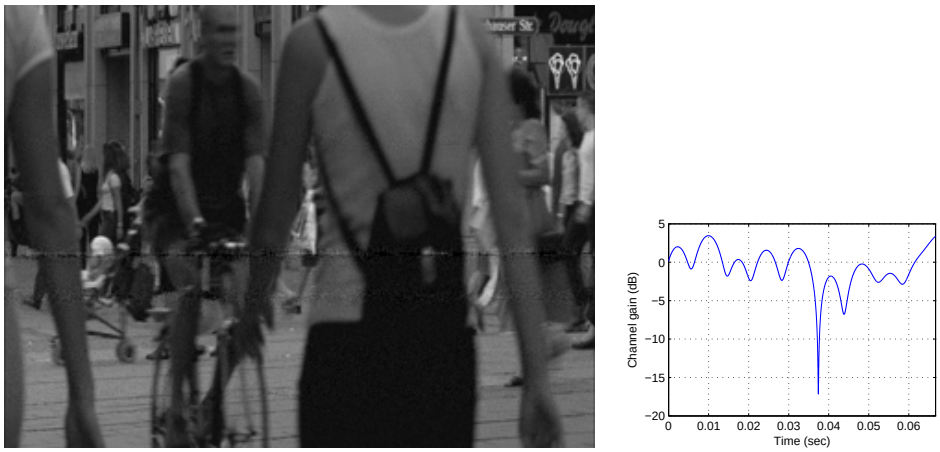


Figure 4.15: *PSNR vs* $\frac{E_p}{N_0}$ *for the total system using for the case* ***Fading***

In Figure 4.14, the results when the three test videos described in Section 3.1 were transmitted using the iterative transmission scheme with 3 iterations and feedback power 20 dB above feed forward are plotted (**Noise**-case). The results are significantly better than when "Lena" was the input. This can be due to two reasons: Now, the mean space prediction parameters from Table 4.2 were used instead of the ones normally used for natural images, which was the case when "Lena" was input. This gives better prediction, hence a lower variance of the transmitted signal, and better results. An other explanation can be that "Lena" is more difficult to code. This reason alone seems unlikely because all of the three test videos shows significantly better results than "Lena" did and the reason why "Lena" has lower PSNR is probably due to a combination of the the two arguments above. TV1 has the best PSNR for all input powers, then TV2 and at last TV3 has the lowest. This is expected and in accordance with earlier results from Section 4.2, where it was shown that TV3 was the most difficult, and TV1 the easiest test video to code when prediction was done only in space. In Figure 4.15, the same scenario is plotted, only here fading is introduced on the channel as well, corresponding to the **Fading**-case described earlier. The order of the test videos are the same, but now they all seem to converge towards a maximum PSNR between 38 and 40 dB. This is because the system cannot compensate for the deepest fades and even though the quality of the image is perfect everywhere else, the rows influenced by the

deep fades consists almost exclusively of noise. When fading is introduced, neither of the test-videos reach a PSNR of 40 dB, but at 11 or 12 db E_p/N_0, where the curves start to flatten out, the image quality in most of the image can bee assumed to be very good except where the deep fades occur. This is in contrast to when there is no fading where TV1 and TV2 reach a PSNR of 40 dB at E_p/N_0 equal to 3.5 and 4 dB respectively. TV3 does not reach the same limit before E_p/N_0 is larger than 5 dB. It is clear that the introduction of flat fading on the channel deteriorate the system performance significantly for all values of E_p/N_0 and in the **fading**-case, a very high transmission power is needed to combat the effects of fading effectively. A frame from TV1 received with $E_p/N_0 = 6$ dB and its corresponding channel are depicted in Figure 4.16. The received image looks flawless except where the fades occur where the noise are stronger. Where the deepest fade occur, there are a few rows consisting of almost only noise. In a single image, these variations in image quality within the image may seem manageable but in a video they will change place and jump around 15 times each second and the video will seem to flicker heavily.



(a) (b)

Figure 4.16: *A frame from TV1, received with average $E_p/N_0 = 6dB$ and parameters corresponding to the case **Fading**. There is fading on the channel, which can be seen in the image. There are several rows that consists of just noise, but where the channel is not in a deep fade, the image look almost perfect. The corresponding channel is also plotted. Notice how the fades correspond to noise in the image. The PSNR is 37 dB so the channel is "nicer" than for the average performance plotted in Figure 4.15 where $E_p/N_0 = 6$ dB corresponds to a PSNR of just above 34 dB*

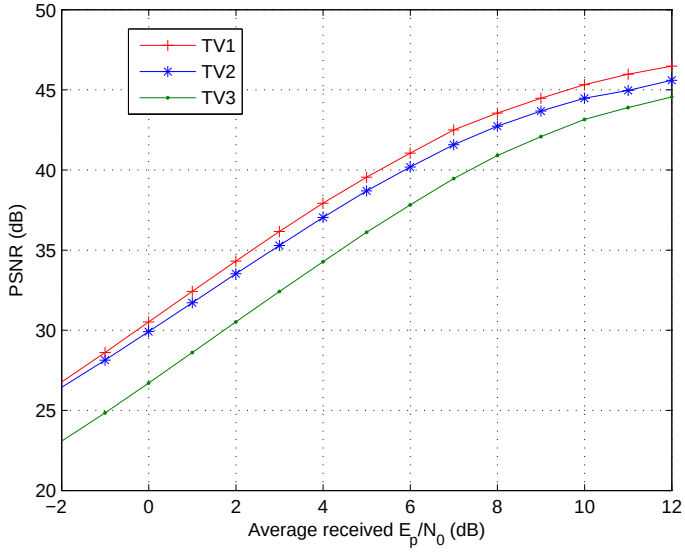


Figure 4.17: *PSNR vs $\frac{E_p}{N_0}$ for the total system with fading and channel diversity but compression is only based on prediction in space and not in time. This corresponds to the case* ***Diversity*** *described earlier*

When channel diversity is added to the system, the system becomes more stable and has less flickering within an image, and overall PSNR is better. The results when diversity was added can be seen in Figure 4.17. The PSNR of the received image is better for all E_p/N_0 than it was without diversity. This is of course expected since the system would now become more robust against fading. The curve converges towards a maximum PSNR in this case as well, but now for a higher PSNR. This is because even though the fading has been reduced due to channel diversity there is still some fades that cannot be compensated for. The gain from applying channel diversity on the system can be seen in Figure 4.18. The gain almost resembles a linear function and the gain increase with increasing E_p/N_0. This is because this diversity scheme uses a form of equal gain combining, meaning every branch is weighted equally, independent of the instantaneous SNR on the branch. The gain from using 2-branch channel diversity with equal gain combining was derived in Section 2.4, and was found to be 1.8, meaning the performance from applying diversity is that E_p/N_0 can be divided by 1.8 to get the same performance as without diversity. This can be seen in the two Figures, the one depicting the **Fading**-case and the one depicting the **Diversity**-case. When there is no diversity TV3 has a PSNR of about 30 when $E_p/N_0 = 4$ dB, while when diversity is

added, 30 dB PSNR is reached for about $E_p/N_0 = 2$ dB. The gain flattens out at around 11 dB. This is because neither the **fading** nor the **diversity** case can compensate for the remaining deep fades without highly increasing power.

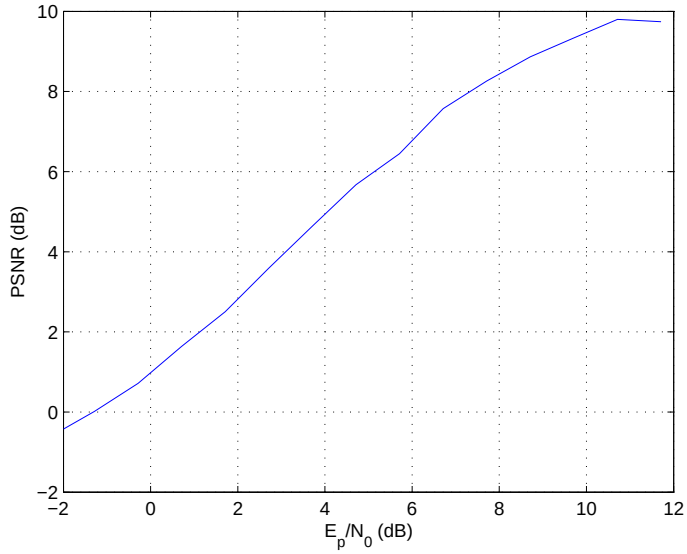


Figure 4.18: *The difference in PSNR as a function of E_p/N_0 for TV2 when diversity is added to the system.*

In Figure 4.19, the results for the total system including video compression with motion estimation are depicted. The performance is notably increased compared to the same system without prediction in time in Figure 4.17. The order of which test video that has the best results is turned around. When prediction is done in time as well, TV2 shows marginally better results than TV1, as opposed to earlier results when TV1 has the best performance. This is in accordance with the results found in Section 4.2 where it was shown that after 3d-dpcm compression, TV2 had a marginally lower variance than TV2. TV3 is still the test video with lowest PSNR at any E_p/N_0. This is in correspondence with the results in Table 4.5, which showed that TV3 had the highest variance after the predicted signal had been subtracted. TV1 and TV2 reaches a PSNR of 40 dB for about $E_p/N_0 = 3$ and 2.5 dB, which is respectively a 2 and 3.5 dB (40% and 58%) reduction compared to the **Diversity**-case in Figure 4.17. When TV3 is input, E_p/N_0 can be reduced from 7.5 dB to 5 dB to reach a PSNR of 40 dB, a reduction of 36%. A big difference from the

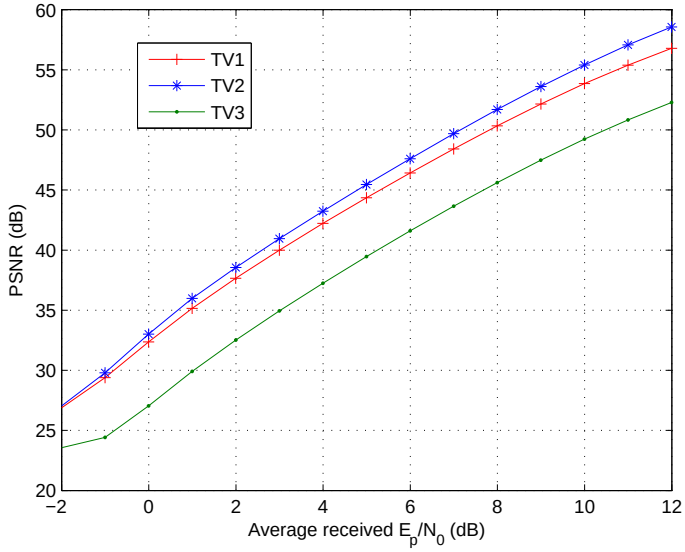


Figure 4.19: *PSNR vs* $\frac{E_p}{N_0}$ *for the case* ***3d-dpcm****. With diversity, channel estimation, fading and noise on all channels and video compression with motion estimation*

earlier cases is that the curves does not seem to flatten out for any of the test videos, which they probably will for a higher E_p/N_0.

Since the main transmission on the feed forward channel in the system has to wait while the channel is estimated and when the motion vectors are transmitted, the bandwidth is increased somewhat and the sampling period decreased accordingly. The new and total bandwidth of the system is:

$$B = 15 \times 320 \times 240 \times 3 + \frac{15 \times 320 \times 240}{64} + \frac{15 \times 320 \times 240}{752} \quad (4.1)$$

$$= 3.476 \times 10^6 \text{Hz} \quad (4.2)$$

and T_s is correspondingly

$$T_s = \frac{1}{B} = 287.7\text{ns} \quad (4.3)$$

where the first term represents the transmission of pixel values, the second is the motion vector with a block size of 64 pixels and the last is the channel estimation being estimated for every 752 pixels. The previous calculated T_s from equation 3.5 with a bandwidth expansion of 3 was 289.3 ns, so the sampling period is only marginally altered and the assumptions for what type of fading to be expected derived in Section 2.3.3 are still valid.

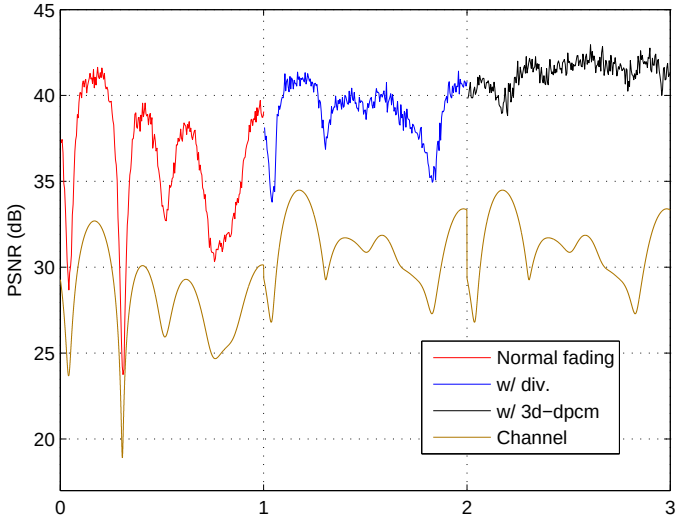


Figure 4.20: *Instantaneous PSNR for one frame from TV1 transmitted with three different cases:* ***Fading*** *(red),* ***Diversity*** *(blue) and* ***3d-dpcm*** *(black). The channel is included (brown) for correlation comparison. Average E_p/N_0 in the receiver was 5 dB and the feedback power was 20 dB above feed forward*

To see how the different methods (with and without diversity and time prediction) reacts to fading, the PSNR was calculated for each row in the received image and plotted over the channel for comparison. This is done with TV1 and TV2 as input in Figure 4.20 and 4.21 respectively. The plot consists of four curves: One showing the instantaneous PSNR for a frame when there is fading on the channel (**Fading**-case), one showing instantaneous PSNR for the received frame when diversity is added as well (**Diversity**-case) and one showing instantaneous PSNR when prediction in time is added in addition to diversity (**3d-dpcm**-case). The last curve shows how the channel varies under each frame. The channel used is the same for all three cases, but when diversity is added, it is the sum of the two channels. The exact value of the channel gains cannot be seen from the plot since the y-axis is PSNR and the channel is included solely for comparison with the fluctuations of the PSNR within a frame. For both the fading and diversity case, for both TV1 and TV2, the instantaneous PSNR resembles the channel almost perfectly. The fading case follows every fade and seem to break down in the deepest fades. The diversity case does the same, but in this case the fades are not as deep. When three dimensional DPCM is applied to the system, there is no visible correlation

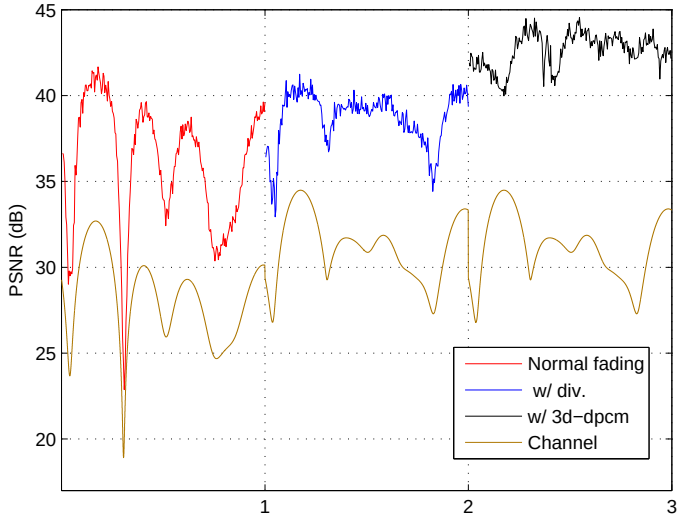


Figure 4.21: *Instantaneous PSNR for one frame from TV2 transmitted with three different cases:* ***Fading*** *(red),* ***Diversity*** *(blue) and* ***3d-dpcm*** *(black). The channel is included (brown) for correlation comparison. Average E_p/N_0 in the receiver was 5 dB and the feedback power was 20 dB above feed forward*

between the PSNR and the channel, and it fluctuate much less than the case without time prediction. PSNR was calculated for the frame when there was no fading on the channel and in this case (not included in the figure), the PSNR was overall nearly constant. It fluctuated rapidly within the range of ±0.5dB, so the larger changes in the 3D-DPCM case is due to channel fades, just not the channel plotted below. One could think that since it uses the previous image to predict block values, it has correlation with earlier channel gains, but this is not the case. The changes are probably due to a combination of the current and some previous channels.

Since the system tries to compensate for the fading before transmitting on the feed forward channel and after receiving on the feedback channel, the strong correlation between the channel and the PSNR for the fading and diversity case is unexpected. The expected form would be more constant but break down in the deepest fades since it cannot compensate for that low channel gains. Since the signal received on the feedback channel is amplified together with noise in the helicopter, the SNR is not changed because the noise is amplified in the same manner as the signal. Therefore, the consequence of the fading

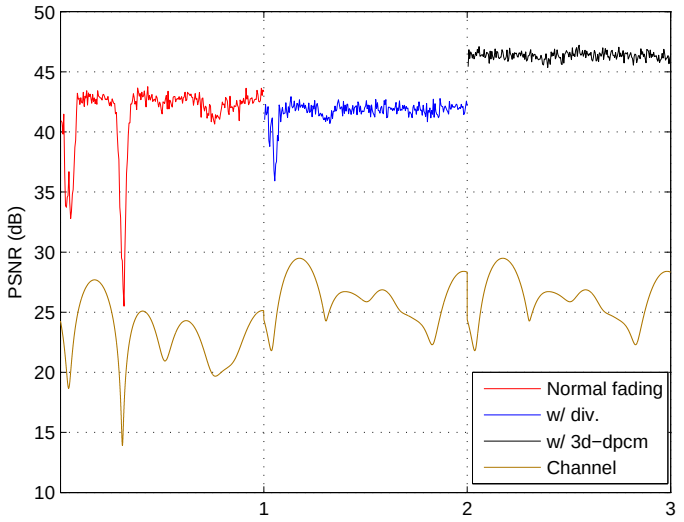


Figure 4.22: *Instantaneous PSNR for one frame from TV2 transmitted with three different cases:* ***Fading*** *(red),* ***Diversity*** *(blue) and* ***3d-dpcm*** *(black). The channel is included (brown) for correlation comparison. Average E_p/N_0 in the receiver was 5 dB and the feedback power was 40 dB above feed forward. Compare with Figure 4.21 where feedback power is only 20 dB above feed forward*

on the feedback channel is that the noise is amplified as the inverse of the channel, leading to a poor reception in the helicopter in the deep fades. Even though the helicopter only takes samples that have a higher SNR than there is one the feed forward channel into account, the feedback SNR will change with the channel. This is the reason for the strong correlation between channel and PSNR in Figure 4.20 and 4.21 for the **Fading** and **Diversity** case. In Figure 4.22 the same Figure is repeated for TV1 but now the feedback power is increased from 20 dB above feed forward to 40 dB. Here, the PSNR is more what to be expected. It is much more constant but still breaks down in the in the deepest fades except for the **3D-dpcm** case where is is almost the same as the no-fading case.

It is obvious from Figure 4.19 that adding time prediction to the system increases the performance significantly. Depending on the input, the performance is increased and on average 40 dB PSNR is reached for $E_p/N_0 = 3.5$ dB instead of 6.17 dB compared to the **Diversity**-case, which is an average reduction of 43%. The complexity added to the helicopter is also significant

Table 4.6: *Average E_p/N_0 (dB), for all test videos, needed in the receiver to reach a PSNR of 40 dB for the different cases described earlier*

	TV1	TV2	TV3	MEAN
Noise-case	3.5	4	5	4.17
Fading-case	> 12	> 12	> 12	–
Diversity-case	5	6	7.5	6.17
3d-dpcm-case	3	2.5	5	3.5

since the search for motion vectors is a heavy operation, but since this can be optimized based on the extensive research that can be found in the literature, this complexity can be reduced. It can also be seen from the figures showing instantaneous PSNR that the instability of the system due to fading is reduced considerably by adding prediction in time. This is due to the fact that the large reduction in signal variance increases the SNR of the received signal and it becomes more resistant against channel fluctuations. The results from the transmission scheme when different components were added are summarized in Table 4.6.

By using Equation 3.21, one can calculate an estimated transmission power needed to reach a certain E_p/N_0 in the receiver. For Scenario 1, the distance between the transmitter and receiver is about 100 meters. This gives the average transmission power:

$$P_t = 3.5 + 30\log(100) - 87 = -23.5 \text{ dB} \quad (4.4)$$

$$= 4.5 \text{ mW} \quad (4.5)$$

Of course, fading will impact the necessary transmission power as well, implying that the gain to neutralize fading will increase the needed transmission power. P_t is the averaged needed transmission power to reach a 40 dB PSNR. This, together with the restrictions imposed by the telecommunications authorities, described in Section 2.3.1, gives an upper limit for how much higher power than the feed forward channel can be used on the feedback channel. If transmission power from the base station should be 30 dB greater than that of the helicopter, it would need an average transmission power of 4.5 W, which is more than twice the limit imposed by the government. Maximum radiated power cannot exceed 2 W, which gives a maximum feedback power 26 dB above that of the feed forward.

CHAPTER 5

MAIN FINDINGS AND CONCLUSION

Based on the results and discussion in Chapter 4, this chapter will conclude the thesis. Simulations have been performed on channel estimation, video compression and transmission both separately and together to fully understand their inherent properties and how they work together. Components were added to the system gradually to the see the improvement caused by every component and to weigh the added complexity versus the gain in performance.

The results from the channel estimation showed that the flat fading channel, given a bandwidth expansion factor of 3 and a top velocity of the helicopter of 10 m/s, could be estimated accurately by transmitting one single pulse with receiver SNR of 20 dB for every 750 pixel. The mean squared error of the total channel estimation was 0.0025.

The video compression was done by subtracting a predicted pixel from the original before transmission, hence only processing the difference. The receiver would predict the pixel in the same manner based on earlier received pixels and add this to get the decoded pixel. First, prediction was based solely on pixels within the same frame as the one to be transmitted. The receiver did the predicting and transmitted the result to the transmitter via a feedback channel. This reduced the variance and hence the power needed to transmit by 12 dB on average for all the test inputs. When the transmitter first subtracted a pixel from the previous frame shifted due to motion, and then subtracted the predicted value received on the feedback channel, the variance was lowered by additionally 4.7 dB. These simulations were done without the presence of

noise, making 16.7 dB the maximum average gain from compression with these test videos. For low values of E_p/N_0, the gain was smaller.

The iterative amplitude continuous system acted according to OPTA when the input was Gaussian distributed and the feedback channel ideal and marginally less for a non-ideal feedback channel with a 20 dB higher power than the feed forward channel. Simulations showed that the resulting signal after differential coding is not Gaussian but rather Laplacian distributed. Since the system only is optimal for Gaussian input, the system will not reach OPTA with a Laplacian source. After reviewing the system for different levels of bandwidth expansion and feedback power, it was reasoned that a bandwidth expansion of 3 and a feedback power 20 dB above feed forward yielded the best performance weighted against the added cost by increasing the bandwidth and power further. These values were used for the remaining simulations.

When both the feed forward channel and the feedback channel was deteriorated by flat Rayleigh fading, a 40 dB PSNR, which is deemed acceptable video quality, was not reached for E_p/N_0 smaller than 12 dB. Here, the channel was estimated in the helicopter to counteract the fading using the parameters obtained under the separated channel estimation simulations. But due to power limitations and harsh channel conditions it could not compensate for all the fades. Channel diversity was obtained by adding an extra antenna on the base station, and a PSNR equal to 40 dB was reached for an average $E_p/N_0 = 6.2$ dB. Motion was estimated in the transmitter and the motion vector was transmitted to the receiver for every block (block size of 8×8 pixels). When the motion compensated predicted signal was subtracted before the spatial prediction was subtracted, the performance increased further. E_p/N_0 could on average be reduced with 2.7 dB to reach the same target PSNR, a reduction by 43%. Three-dimensional DPCM together with diversity made the system much more robust against fading and instantaneous PSNR is only barely reduced even in the deepest fades. Even though motion detection introduces a higher complexity to the system and imposes a higher need for computer power in the small helicopter, the added gain in performance is considered adequate to make up for the added complexity.

The video communication system is heavily dependent on the stability of the feedback channel, both for compression and transmission, and if the feedback channel should break down or the SNR decrease rapidly, the performance of the system would suffer accordingly. In a practical implementation of the system, one of the largest challenges would be to guarantee the stability of the

feedback channel independent of the environment and conditions the helicopter is operating. The system uses a bandwidth of 3.5 MHz to reduce transmission power and it was argued that this bandwidth and carrier frequency (2.39 GHz) would give flat fading in an indoor office environment, but in a more open environment like outdoor, the fading will be frequency selective. Since the channel estimation only estimates the channel degradation on the amplitude of the signal, the estimation has to be changed to measure alterations in frequency as well. This takes more time and will result in a longer delay.

Disregarding these challenges, it is the authors opinion that this scheme has great potential in systems where the low energy consumption and complexity trump the large use of bandwidth.

BIBLIOGRAPHY

[1] Mohamed-Slim Alouni and Andrea J Goldsmith. Capacity of rayleigh fading channel under different adaptive transmission and diversity-combining techniques. *IEEE Transactions on vehicular technology*, 48(4):1165–1181, July 1999.

[2] J. B. Andersen, T. S. Rappaport, and S. Yoshida. Propagation measurements and models for wireless communications channels. *Communications Magazine, IEEE*, 33(1):42–49, Jan 1995.

[3] Toby Berger. *Rate distortion theory.* Prentice-Hall, Inc, Englewood Cliffs, New Jersey, USA, 1971.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* John Wiley & Sons, Inc, Hoboken, New Jersey, USA, 2006.

[5] Andrea Goldsmith. *Wireless Communications.* Cambridge University Press, 40 West 20th Street, New York, USA, 2005.

[6] Rafael Gonzalez and Richard Woods. *Digital Image Processing.* Pearson Educationl, Inc, Upper Saddle River, New Jersey, USA, 2008.

[7] Ola Naalsund Ingvaldsen and Brage Høyland Ellingsæter. Power efficient image communication for uav. *Main project as a part of a Master degree at the Norwegian University of Science and Technology*, December 2009. Email for copy: ola.ingvaldsen@gmail.com.

[8] A.K. Jain. Image data compression: A review. *Proceedings of the IEEE*, 69(3):349–389, March 1981.

[9] W. C. Jakes. New techniques for mobile radio. *ell Laboratory Rec.*, 48(11):326–330, 1970.

[10] Xuan Jing and Lap-Pui Chau. An efficient three-step search algorithm for block motion estimation. *IEEE Transactions on Multimedia*, 6(3):435–438, 2004.

[11] Dimitris G. Manolakis John G. Proakis. *Digital Signal Processing.* Pearson Education, Inc, Upper Saddle River, New Jersey, USA, 2007.

[12] L. Kahn. Ratio squarer. *Proceedings of IRE(Correspondence)*, 42:1074, November 1954.

[13] T. KOGA, K. IINUMA, A. HIRANO, Y. IIJIMA, and T. ISHIGURO. Motion-compensated interframe coding for video conferencing proceedings. *NTC'81 (IEEE)*, pages G.5.3.1 – G.5.3.4, 1981.

[14] Wim A. Th. Kotterman, Gert F. Pedersen, and Kim Olesen. Diversity properties of multiantenna small handheld terminals. *EURASIP J. Appl. Signal Process.*, 2004:1340–1353, 2004.

[15] William C. Y. Lee. *Mobile Cellular Telecommunications Systems.* McGraw-Hill, Inc., New York, NY, USA, 1990.

[16] Tor A. Ramstad. *Image Communication.* Department of Electronics and Telecommuncations, NTNU, NO-7491 Trondheim, Norway, 2009.

[17] Theodore S. Rappaport. *Wireless Communications, Principles and Practice.* Prentice Hall, Inc, Upper Saddle River, New Jersey, USA, 2002.

[18] Adel A. M. Saleh and Reinaldo A. Valenzuela. A statistical model for indoor multipath propagation. *Selected Areas in Communications, IEEE Journal on*, 5(2):128–137, 1987.

[19] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack. Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, Jan 2010.

[20] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack. A subjective study to evaluate video quality assessment algorithms. *SPIE Proceedings Human Vision and Electronic Imaging*, Jan 2010.

[21] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[22] Claude E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec., pt. 4*, pages 142–163, 1959.

[23] Raymond Steele. *Mobile Radio Communications*. IEEE Press, Piscataway, NJ, USA, 1994.

[24] John C. Stein. Indoor radio wlan performance part ii: Range performance in a dense office environment. *Technical report, Intersil Corporation, 2401 Palm Bay, Florida*.

APPENDIX A

EXAMPLE IMAGES

This section will show some example frames transmitted with the total system for the different cases described in Section 4.3. There is no discussion of the images, since they are included for the reader to compare the image quality received using the different cases with different SNRs.

Figure A.1: *From from TV1 for the* ***Noise****-case.* $E_p/N_0 = 1$ *dB,* $PSNR = 35.5$ *dB*

Figure A.2: *From from TV1 for the* ***Noise****-case.* $E_p/N_0 = 2$ *dB,* $PSNR = 37.4$ *dB*

Figure A.3: *From from TV1 for the* ***Noise****-case.* $E_p/N_0 = 3$ *dB,* $PSNR = 39.4$ *dB*

Figure A.4: *From from TV1 for the* ***Fading****-case.* $E_p/N_0 = 1$ *dB,* $PSNR = 28.9$ *dB*

Figure A.5: *From from TV1 for the* ***Fading****-case.* $E_p/N_0 = 2$ *dB,* $PSNR = 30.2$ *dB*

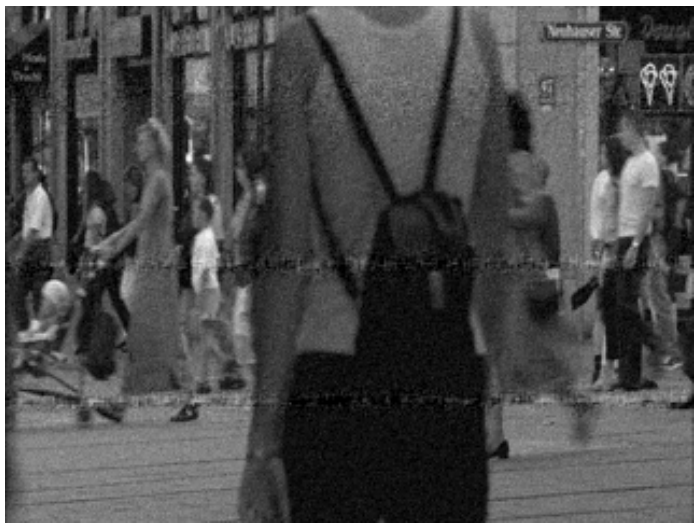


Figure A.6: *From from TV1 for the* ***Fading****-case.* $E_p/N_0 = 3$ *dB,* $PSNR = 31.4$ *dB*

Figure A.7: *From from TV1 for the* ***Fading****-case.* $E_p/N_0 = 5$ *dB,* $PSNR = 33.5$ *dB*

Figure A.8: *From from TV1 for the* ***Fading****-case.* $E_p/N_0 = 6$ *dB,* $PSNR = 34.6$ *dB*

Figure A.9: *From from TV1 for the* ***Diversity****-case.* $E_p/N_0 = 1$ *dB,* $PSNR = 33.7$ *dB*

Figure A.10: *From from TV1 for the* ***Diversity****-case.* $E_p/N_0 = 2$ *dB,* $PSNR = 35.7$ *dB*

Figure A.11: *From from TV1 for the* ***Diversity****-case.* $E_p/N_0 = 3$ *dB,* $PSNR =$ 37.8 *dB*

Figure A.12: *From from TV1 for the* ***Diversity****-case.* $E_p/N_0 = 4$ *dB,* $PSNR =$ 39.8 *dB*

Figure A.13: *From from TV1 for the* ***Diversity****-case.* $E_p/N_0 = 5$ *dB,* $PSNR = 41.9$ *dB*

Figure A.14: *From from TV1 for the* ***3d-dpcm****-case.* $E_p/N_0 = 0$ *dB,* $PSNR = 32.4$ *dB*

Figure A.15: *From from TV1 for the* ***3d-dpcm****-case.* $E_p/N_0 = 1$ *dB,* $PSNR = 35.2$ *dB*

Figure A.16: *From from TV1 for the* ***3d-dpcm****-case.* $E_p/N_0 = 2$ *dB,* $PSNR = 37.7$ *dB*

Figure A.17: *From from TV1 for the* ***3d-dpcm****-case.* $E_p/N_0 = 3$ *dB,* $PSNR = 40$ *dB*

Figure A.18: *From from TV1 for the* ***3d-dpcm****-case.* $E_p/N_0 = 4$ *dB,* $PSNR = 42.2$ *dB*