Visual Communications and Image Processing IV

Part 2

PROCEEDINGS
SPIE—The International Society for Optical Engineering

Visual Communications and Image Processing IV

William A. Pearlman
Chair/Editor

8–10 November 1989
Philadelphia, Pennsylvania

Sponsored by
SPIE—The International Society for Optical Engineering

Cooperating Organizations
IEEE Circuits and Systems Society
IEEE Acoustics, Speech, and Signal Processing Society
European Association for Signal Processing (EURASIP)

SPIE
P
PROCEEDINGS
SERIES

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 206/676-3290 (Pacific Time) • Telex 46-7053

Volume 1199
Part Two of Three Parts

SPIE (The Society of Photo-Optical Instrumentation Engineers) is a nonprofit society dedicated to advancing engineering and scientific applications of optical, electro-optical, and optoelectronic instrumentation, systems, and technology.

P

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," *Visual Communications and Image Processing IV*, William A. Pearlman, Proc. SPIE 1199, page numbers (1989).

Library of Congress Catalog Card No. 89-63405
ISBN 0-8194-0238-9

Address inquiries and notices to Director of Publications, SPIE, P.O. Box 10, Bellingham, WA 98227-0010 USA.

CONTENTS

(continued)

(continued)

Part Two

(continued)

(continued)

POSTER SESSION II

Image Restoration, Enhancement, and Analysis

Chair
Toshio Koga
NEC Corporation (Japan)

VIDEO SIGNAL ENHANCEMENT USING FINE STRUCTURE PRE-EMPHASIS WITH MODULO LIMITING

Chung H. Lu

Department of Electrical Engineering
Center for Communications and Signal Processing Research
New Jersey Institute of Technology
Newark, New Jersey 07102

Abstract

Companding is required in temporal or vertical pre-emphasis to realize the potential SNR improvement. A modulo limiter (ML) is a nonlinear device which maps the input signal to a predetermined output range. If the input falls within the predetermined output range, no change is made to the input signal and the output equals the input. If the input falls outside the predetermined output range, a DC offset equal to the length of the predetermined output range is repeatedly added to or subtracted from the input until the result falls inside the predetermined output range. Analysis shows that a modulo limiter can be incorporated in the pre-emphasis filters to limit the pre-emphasized signal amplitute to within the range of the input signal from $-\frac{1}{2}$ to $\frac{1}{2}$ volts. The reduced dynamic range will alleviate stringent requirement on companding.

1. Introduction

Baseband noise in FM system has a parabolic power spectrum. Pre-emphasis is a well known technique in FM transmission system to improve received signal quality by suppressing the effect of the larger noise at high frequencies [1]. In video signal transmission with frequency modulation, a CCIR (Comité Consultatif International des Radiocommunications) pre-emphasis/de-emphasis standard is commonly used. The standard specifies the characteristics of the pre-emphasis and de-emphasis processors in terms of frequency response. The pre-emphasis and the de-emphasis processors are complementary to each other so as not to introduce distortion to the video signal, assuming the channel bandwidth is wide enough.

In general, the pre-emphasis processor is a filter whose transfer function tends to feature higher gain at higher frequencies. On the other hand, the de-emphasis processor is a filter featuring lower gain at higher frequencies. The processors are matched so that the end-to-end connection produces a flat frequency response. The FM noise at higher frequencies, which could have been otherwise unacceptably too large, especially at frequencies around color subcarrier, are consequently suppressed to an acceptable level. The high-pass nature of the pre-emphasis filter has the effect of differentiating or accentuating vertical edges and undoubtedly increases the peak-to-peak amplitude of the video signal. The FM carrier may be over deviated by the increased amplitude, leading to the phenomenon of image distortion known as vertical edge tearing.

Video signal is highly redundant when the signal does not contain much motion or spatial variation. In such cases, the video signal manifest its redundancy with a highly structured power spectrum. Furthermore, the structured spectrum exhibits empty spaces (in frequency domain) where there is no signal energy. It is such structured spectrum which an interline or interframe process attempts to exploit.

Unfortunately, when there is motion or detailed spatial variation, the spectral gaps start to fill up with signal energy of unexpected magnitude and location. Such unforeseeable events can be tracked with an adaptive process. An alternative approach is to predistort the signal at the transmitter and then complementarily restore it at the receiver. This paper describes an interline pre-emphasis process (Section 2) and an interframe pre-emphasis process (Section 3). Collectively, these pre-emphasis processes are called fine structure pre-emphasis processes. Section 4 introduces a modulo limiter (ML) and Section 5 applies the ML in the fine structure pre-emphasis process to ease the requirement of companding, which is needed to realize the SNR improvement capability of fine structure pre-emphasis. Section 6 summarizes the paper.

2. Interline Processing and Vertical Pre-Emphasis

The highly structured spectrum of a video signal makes it possible to design an interline process for video signal enhancement. In a TV signal, strong correlation exists horizontally between adjacent samples and vertically between adjacent lines. Correlation also exists between adjacent fields and adjacent frames. Because of the strong line-to-line correlation, the spectrum of a TV signal consists of spectral components centered around the harmonics of line scanning frequency. The effect of temporal (field-to-field and frame-to-frame) correlation, on the other hand, is to produce around each scanning line harmonic multiple spectral lines. These spectral lines are spaced at a frequency equal to the frame repetition rate.

As the line-to-line correlation increases, *i.e.*, the interline variation decreases, the size of such empty space increases. At the extreme when the image becomes constant vertically, there is no interline variation and the video signal becomes periodic with a period equal to the line scanning period. All spectral lines disappear except the canning line harmonics. Strictly speaking, because of the vertical blanking and synchronization pulses, the video signal can never become periodic with a period equal to the line scanning period. However, since we are concerned only with active video signal, the above scenario is very much likely. To be sure, the above above observation refers to short time spectrum.

For an NTSC (National Television Standard Committee) format signal, the line scanning frequency is 15.73426 KHz and the frame repetition rate is approximately 30 Hz. Thus one sees in the spectrum of an NTSC signal a fine structure which corresponds to the line scanning rate of 15.734264 KHz and a superfine structure which corresponds to the frame rate of 30 Hz. The spectrum consists of pickets of spectral lines separated by 30 Hz. The pickets are separated by 15.734 KHz. For color signals, the chrominance component exhibits a similar fine structure spectrum, which is interleaved with the luminance component. Thus, in a composite NTSC signal, which contains both luminance and chrominance components, the spectrum consists of spectral pickets spaced by 7.867 KHz. The luminance component has a peak at low frequencies around DC and the chrominance component has a peak centered at the color subcarrier frequency f_{sc}.

High correlation between consecutive scanning lines keeps video signal energy confined around the scanning line harmonics. Vertical variation in the scenes has the effect of spreading the energy spectrum around the scanning line harmonics. The sharper the variation, the broader the spectrum spreads. In most video signals, there are empty spaces between the scanning line harmonics. Transmission channel induced noise does not have such structured spectrum, on the other hand. The spaces are therefore mostly occupied by noise. In principle, a comb filter can be used to suppress such noise. In practice, however, occasional sharp horizontal edges will place energy in these spaces. Simply suppressing energy in these spaces tends to blur the horizontal edges or detailed vertical variation, an effect which is not desirable.

Adaptive schemes have been studied which include a horizontal edge detector and switch the comb filter on and off depending on the absence and presence of horizontal edges. An alternative approach is to incorporate a complementary comb filter at the transmitter [2,3]. The complementary transmitter comb filter acts as a predistortion processor and the receiver comb filter restores the predistorted signal. The end-to-end cascade of the transmitter and receiver comb filters has a flat frequency response. The video signal is therefore not distorted, but the noise between scanning line harmonics is suppressed by the receiver comb filter.

Because of its close resemblance to CCIR pre-emphasis, and because of the structured spectrum of video signals, such process for reducing the inter-harmonic noise is called fine structure pre-emphasis or vertical pre-emphasis. In comparison, CCIR pre-emphasis can be called horizontal pre-emphasis. The transmitter comb filter is called the pre-emphasis filter and the receiver comb filter is called de-emphasis filter.

3. Interframe Processing and Temporal Pre-Emphasis

The structured spectrum of a video signal can be observed also in a smaller (in frequency) scale. Such superfine structure reflects the strong frame-to-frame correlation in a video signal. For still images, the video signal is strictly periodic with a repetition rate of approximately 30 Hz. The signal power spectrum thus consists of spectral lines spaced 30 Hz apart. Because of the strong line-to-line correlation, these spectral lines cluster around each of the scanning line harmonics, with a void between the scanning line harmonics.

Interframe averaging has been studied for applications in noise reduction [4,5]. The averaging process is essentially a (low pass) comb filter which suppresses noise between the spectral lines. As motion or scene change occurs, the spectral lines broaden to fill the void between the spectral lines with signal energy. The broadening of spectral lines increases as motion becomes faster, or scene changes become more drastic, *i.e.*, changing to a totally different scene. Suppressing noise in these regions by direct filtering consequently distorts the signal.

An adaptive scheme incorporates a motion detection mechanism so that averaging can be regulated by the degree of motion. Whether it is adaptive or not, interframe averaging is done only at the receiver. Alternatively, a complementary (high pass) comb filter can be employed at the transmitter to predistort the video signal. The receiver (low pass) comb filter then restores the predistorted signal and suppresses the noise at the same time. Such noise suppression process is called superfine structure pre-emphasis, or interframe or temporal pre-emphasis. The transmitter (high pass) comb filter is called the pre-emphasis processor and the receiver (low pass) comb filter is called the de-emphasis processor.

Interframe process in general and temporal pre-emphasis in particular can be understood both in the frequency domain and in the time domain. First consider the frequency domain point of view. In a video signal of still images, there is no signal power in the frequency ranges between the spectral lines spaced at the frame repetition rate of approximately 30 Hz. Substantial power distribution in these frequency ranges can be found only when there are large interframe differences such as due to scene changes or motion. In most video sequences, there is very little signal power in the regions between the 30 Hz spectral lines. Random channel noises, on the other hand, do not have interframe correlation and therefore do not have a line structured spectrum. As a matter of practice, the noise power spectrum in each of the regions between the 30 Hz spectral lines can be assumed white.

At the transmitter, the low signal energy regions between the spectral lines can be boosted by a frequency selective pre-emphasis amplifier. At the receiver, a complementary de-emphasis filter is then applied to restore the pre-distorted signal and in the process reduce the noise in the regions between the spectral lines at the same time. Such interframe processing can be called superfine structure pre-emphasis/de-emphasis or temporal pre-emphasis/de-emphasis. Therefore, from a frequency domain point of view, temporal pre-emphasis or interframe processing in general reduces the noise energy in the spectral region where there is no signal energy.

From the time domain point of view, temporal pre-emphasis reduces random noise using the principle of averaging. The pre-emphasis is basically a frame differentiator. It takes samples from two consecutive frames, removes the interframe redundancy, and outputs the weighted difference signal. The complementary de-emphasis processor at the receiver is a leaky frame integrator. The leaky coefficient of the frame integrator is equal to the weighting coefficient in the frame differentiator so that the integrator matches the differentiator to obtain a unity transfer coefficient. With such arrangement, the pre-emphasis de-emphasis combination is transparent to the input video signal. On the other hand, the noise is reduced through integration of the de-emphasis.

Interframe processing produces a signal which, on the average, is smaller in dynamic range and requires smaller baseband bandwidth. Only rarely does an extreme signal swing occur. Such extreme signal swing corresponds to motion and scene changes. The original signal is relatively uniformly distributed, but the weighted frame difference tends to be concentrated around zero. In an analog system, a smaller dynamic range and baseband bandwidth facilitates more efficient use of channel bandwidth and improves the received signal quality. In a digital system, a more efficient