ON THE RELATIONSHIPS BETWEEN MOTION AND ALIASING IN TYPICAL VIDEO RASTERS

G. Cortelazzo, R. Manduchi, C. Monti

Dipartimento di Elettronica e Informatica
Via Gradenigo 6/A - 35131 Padova - Italy

Abstract

This work applies frequency domain motion analysis to the explanation of the visual artifacts and of the aliasing of video signals. Non-standard perspectives of motion consideration, possibly leading to a better understanding or to a quantification of the effects of motion, find their interest in the relevance of motion issues in video coding and signal processing.

1. Introduction

Motion rendition without artifacts is a central issue in video coding and signal processing. A motivation for variable-rate image coding is the attempt of keeping constant the perceptual quality of the coded signal by increasing the bit-rate when the scene's motion increases (It is well known that if the bit-rate is fixed the perceptual quality of the coded signal becomes worse as the motion content of the signal increases). Motion-adaptive filtering is one of the most effective solutions for luminance/chrominance separation and scanning rate conversion. Within such a perspective the exploration of the ways motion leads to information loss and/or to visual artifacts appears quite interesting.

This work examines the effects of motion on the spatio-temporal discretization relative to the two lattices most commonly encountered with television signals: the 3-D orthogonal lattice and the lattice orthogonal in the spatial coordinates and hexagonal in any of the two spatio-temporal directions [1]. Such lattices are related to the discretization of the signals defined on the most typical video rasters, i.e., the progressive and the 2:1 interlaced rasters.

The adopted motion model refers to translations on planes parallel to the image plane, and it includes velocity and acceleration.

In first approximation the operation of visual perception can be modelled in first approximation as a spatio-temporal low-pass filtering [2,3]. If the frequency response of the visual system is denoted by $E(k_x,k_y,f)$ the visual artifacts due to spatio-temporal discretization can be interpreted as the difference between the output of $E(k_x,k_y,f)$ relative to a continuous image and the output of $E(k_x,k_y,f)$ relative to the spatio-temporal discretization of the continuous image. The latter output may differ from the former because of the aliasing corrupting some energy of the continuous signal and/or because of the presence of spectral repetitions in the pass-band of $E(k_x,k_y,f)$.

It is important to note that a spatio-temporal discretization suitable for a perfect reconstruction of the continuous signal is not necessarily free of perceptual artifacts. In order to simply justify such an apparent discrepancy between information theory and vision, consider that the recovery of the continuous signal from its discrete version uses as reconstruction instrument a low-pass filter having the fundamental cell of the dual lattice of the discrete signal as pass-band, instead the perception of a continuous image from its discrete version uses as reconstruction instrument the eye's frequency response.

In television applications there are occurrences where the most appropriate signal processing target is perfect signal reconstruction and others where the most appropriate target is artifact-free visual perception.

The former situation is common to all the signal manipulations not directly related to image display: subsampling for information compression, which is sometimes done before transmission, is a good example. The second situation pertains the processing of the signal to be direct by displayed.

This work considers the effects of motion with respect to both signal information loss and visual artifacts introduction.

Next section reviews some fundamental notions of frequency domain analysis of motion. Section 3 considers the effects of velocity and acceleration with respect to the introduction of visual artifacts. Section 4 examines the loss of signal information produced by motion. Section 5 has the conclusions.

2. Frequency domain analysis of motion

Let $u_s(\mathbf{x})$, $\mathbf{x}^T = (x,y) \in R^2$ be the image of a still object and $U_s(\mathbf{k})$, $\mathbf{k}^T=(k_x,k_y)\in R^2$ be the spatial Fourier transform of $u_s(\mathbf{x})$. If the object of the scene translates complanarily with the image plane with constant velocity $\mathbf{v}$ for time

$-\infty<t<\infty$, it is well known [4-9] that the spatio-temporal Fourier transform of $u(\mathbf{x},t)=u_s(\mathbf{x}-vt)$, denoted as $U(\mathbf{k},f)$, (where f is the temporal frequency) is

$$U(\mathbf{k},f) = U_s(\mathbf{k})\delta(f+\mathbf{k}^T v) \tag{1}$$

Also if constant velocity motion is confined to a finite time-interval, the signal to consider is $u(\mathbf{x},t)=u_s(\mathbf{x}-vt)w(t)$, with

$$w(t) = \begin{cases} 1 & 0 \le t \le T \\ 0 & \text{elsewhere} \end{cases} \tag{2}$$

and the Fourier transform of $u(\mathbf{x},t)$ is

$$U(\mathbf{k},f) = U_s(\mathbf{k})\ W(f+\mathbf{k}^T v) \tag{3}$$

with

$$W(f) = e^{-j2\pi fT/2}\text{sinc}(fT). \tag{4}$$

(The definition of $w(t)=0$, $t\notin[0,T]$ assumes that 0 is the background signal value. The modification of $w(t)$ in order to suit other background values is straightforward.)

If the object of the scene translates complanarily with the image plane for an infinite time interval with constant acceleration $\mathbf{a}$, it can be shown [10] that the spatio-temporal Fourier transform of $u(\mathbf{x},t) = u_s(\mathbf{x}-\frac{a}{2}t^2)$ is

$$U(\mathbf{k},f) = U_s(\mathbf{k})C_2(\mathbf{k},f) \tag{5}$$

with

$$C_2(\mathbf{k},f) = \begin{cases} \delta(f) & \underline{k}^T\underline{a} = 0 \\ \dfrac{e^{-j\pi/4\,\mathrm{sgn}\,\mathbf{k}^T\mathbf{a}}\ e^{j\pi f^2/\mathbf{k}^2\mathbf{a}}}{\sqrt{|\mathbf{k}^T\mathbf{a}|}} & \underline{k}^T\underline{a} \neq 0 \end{cases} \tag{6}$$

If the constantly accelerated motion has non-zero initial velocity, the Fourier transform of $u(\mathbf{x},t)=u_s(\mathbf{x}-vt-\frac{a}{2}t^2)$ becomes

$$U(\mathbf{k},f) = U_s(\mathbf{k})C_2(\mathbf{k},f+\mathbf{k}^T v) \tag{7}$$

A finite motion duration interval implies the convolution with respect to f of the right-hand sides of (7) by $W(f)$ of (4).

It is worth noting that if $U_s(\mathbf{k})$ is spatially band-limited, $U(\mathbf{k},f)$ given by (1) is also band-limited in the temporal frequency, while $U(\mathbf{k},f)$ given by (3) is not band-limited in the temporal frequency.

If $u(\mathbf{x},t)$ is sampled on a lattice, and $p(\mathbf{x},t)=u(\mathbf{x},t)$ at $(\mathbf{x},t)\in\Lambda$, i.e., $p(\mathbf{x},t)$ denotes the discretization of $u(\mathbf{x},t)$ on the spatio-temporal lattice Λ, it is well known [1] that the discrete Fourier transform of $p(\mathbf{x},t)$, denoted as $P(\mathbf{k},f)$ is the repetition of $U(\mathbf{k},f)$ on the points of Λ^*, the lattice dual of Λ.

The signal processing targeted to perfect signal reconstruction should make $P(\mathbf{k},f)$ as close as possible to $U(\mathbf{k},f)$ in any fundamental cell of Λ^* including the support of $U(\mathbf{k},f)$. The signal processing targeted to perceptual reconstruction should make $P(\mathbf{k},f)$ so that $E(\mathbf{k},f)P(\mathbf{k},f)$ is as close as possible to $E(\mathbf{k},f)U(\mathbf{k},f)$.

3. Motion-dependent visual artifacts of 2:1 interlaced and progressive scanning

This section compares interlaced versus progressive scanning with respect to the motion-dependent introduction of energy in the pass-band of $E(\mathbf{k},f)$. In order to compare equal-density lattices the data pertain to scanning standard 1250/2:1/50 and 625/1:1/50. Furthermore the analysis concerns only vertical motion as the two lattices differe only in their vertico-temporal geometry.

Call Λ_1 and Λ_2 the vertico-temporal lattices corresponding to the sampling of the signals produced by the two scanning standards 1250/2:1/50 and 625/1:1/50 respectively. Fig. 1 concerns scanning standard 1250/2:1/50. It shows the spectral repetitions on Λ_1^* of the Fourier transform of the image produced by a still object (solid line) and those of the image produced by an object moving with constant velocity (dashed line). The dimond, according to [2,3], represents the pass-band of the first order linear model of the eye's vertico-temporal response at a viewing distance of four picture heights.

Fig. 1 shows the maxima spectral extensions, a still scene and and an object moving with vertical velocity 1/12.5 Ph/s, can have in order not to give visual artifacts after the discretization on Λ_1. Spectral extensions corresponding to longer segments would bring to different $E(\mathbf{k},f)$ outputs for the continuous and the discrete signal.

Fig. 2 shows a similar situation for vertico-temporal lattice Λ_2, associated to scanning standard 625/1:1/50. It can be seen that the maximum allowable vertical bandwidth for the still case is much smaller than that of the previous case, whilst the maximum extension for the moving object is greater.

Fig. 1 shows that interlaced scanning is likely to be more affected than progressive scanning by the acceleration, as the repetition centered at (25,625) is closer to the eye's frequency response pass-band than any spectral repetition of Λ_2^* of Fig. 2.

4. Motion-dependent aliasing of 2:1 interlaced and progressive scanning

This section considers the signal information loss produced by velocity and acceleration on interlaced and progressive scanning. In order to appreciate the relationship between motion parameters and lattice geometry it is appropriate to compare equal-density lattices. As previously noted a good example of such an application is given by subsampling which offers a simple information compression technique. For example's sake consider the subsampling of index two [1] of the orthogonal lattice associated to scanning standard 625/1:1/50. Three subsampling strategies are of interest: i) the skipping of a line every two on every field (associated to scanning standard 313/1:1/50); ii) the skipping of a field every two (associated to scanning standard 625/1:1/25); iii) the skipping of even lines on even fields and of odd lines on odd fields (associated to scanning standard 625/2:1/50).

Let call the vertico-temporal lattices associated to the three subsampling strategies Λ_3, Λ_4 and Λ_5 respectively; the spectral repetitions relative to their dual lattices and their fundamental cells are shown in Fig. 3, 4 and 5.

It should be noted that the role played in the previous section by the eye's frequency response, in this section is taken by the fundamental cells of the dual lattices which are marked in Fig. 3, 4 and 5. The intuitive expectations related to the subsampling procedures are confirmed by the frequency domain reasoning based on the motion model. Fig. 3 shows that subsampling strategy i), as expected, has poor vertical resolution but motion rendition comparable with that of the original lattice. Fig. 4 shows instead that subsampling strategy ii) keeps the vertical resolution of the original lattice at the expenses of the temporal resolution, which is exactely halved. The dashed segements of Fig. 3 and 4 indicate the maximum alias-free spectral extension for an object moving at 1/12.5 Ph/s. It should be noted that in Fig. 3 such an extension is greater than that of the still object, instead in Fig. 4 it is smaller.

Subsampling strategy iii), related to interlaced 2:1 scanning, as well known, has resolution characteristics intermediate with respect to those of the previous subsampling strategies. Fig. 5 shows that the maximum vertical resolution a moving object can have without aliasing decreases with vertical velocity. The mechanism by which acceleration generates aliasing, is the spectral expansion of the repetitions at (±25, ±625) into the fundamental cell. It should be noted that the spectral geometry of interlaced 2:1 scanning suits the eye's frequency response better than the geometry of progressive scanning, however with respect to signal reconstruction purposes such a characteristic is not necessarily as valuable as in the context of perceptual reconstruction.

In other words, if the subsampling concerns a signal to be displayed, strategy iii) is to be preferred; but if no direct display is involved strategy i) or ii) may be more appropriate than strategy iii) depending on the signal characteristics.

5. Conclusion

This work has applied frequency domain motion analysis to the examination of the visual artifacts and of the aliasing of typical video imagery. It has been shown how typical television notions can be explained in frequency domain by suitable motion models.
The relevance of motion issues in time-varying image processing makes the approach worth further investigation. More accurate motion models and more general spatio-temporal discretization structures are presently under study.

References

[1] E. Dubois, "The Sampling and Reconstruction of Time-Varying Imagery with Application in Video Systems", IEEE Proceedings, Vol. 73, no. 4, Apr. 1985, pp. 502-522.

[2] Z.L. Budrikis, "Model Approximation to Visual Spatio-Temporal Sine-Wave Threshold Data", Bell System Technical Journal, Vol. 52, No. 9, Nov. 73, pp. 1643-1667.

[3] D.H. Kelly, "Adaptation Effects on Spatio-Temporal Sine-Wave Thresholding", Vision Res., 12, 1972, pp. 89-101.

[4] J.O. Drewery, "The filtering of luminance and chrominance signals to avoid cross-colour in a PAL colour system", BBC Techn. Report RD 1975/91, Dec. 1975.

[5] T.S. Huang, "Image Sequence Analysis", Springer-Verlag, Berlin, 1983.

[6] D.J. Heeger, "Optical Flow From Spatio Temporal Filters", Proc. IEEE Int. Conf. Computer Vision, London, June 87, 181-190.

[7] D.J. Fleet, A.D. Jepson, "A Cascaded Filter Approach to the Construction of Velocity Selective Mechanisms", IEEE Trans. on Patt. Anal. and Mach. Int. Vol. 11, no. 3, pp. 315-325 March 89.

[8] A.B. Watson, A.J. Ahumada Jr., "A Look at Motion in the Frequency Domain", NASA Techn. Memorandum 843522, NASA-Ames Research Center, 1983.

[9] E.H. Adelson, J.R. Bergen, "Spatiotemporal Energy Models for the Perception of Motion", J. Opt. Soc. Amer., Vol. 2, pp. 284-299, 1985.

[10] M. Balanza, G. Cortelazzo, "Frequency Domain Analysis of Accelerated Motion", Proc. SPIE's 33rd Int. Symposium on Optical, Optoelectronic Applied Science and Engineering, San Diego, 1989, pp. 1153-1158.

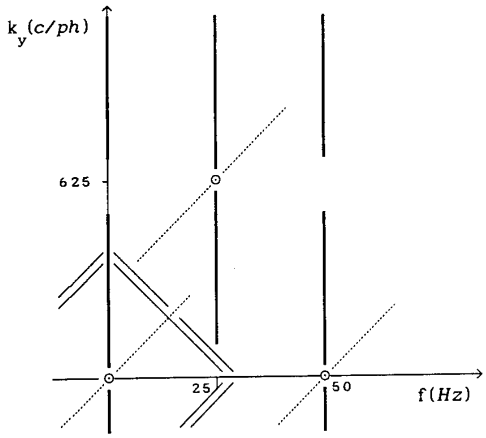

Fig. 1 Spectral vertico-temporal repetitions relative to standard 1250/2:1/50.

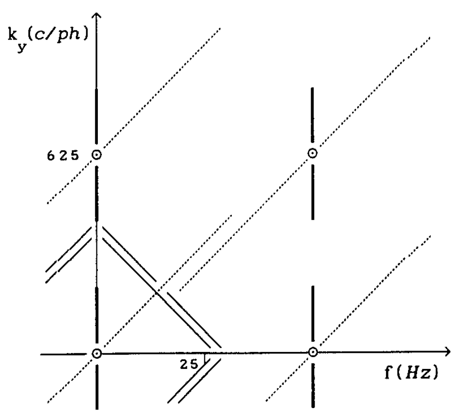

Fig. 2 Spectral vertico-temporal repetitions relative to standard 625/1:1/50.

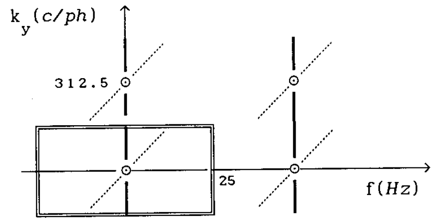

Fig. 3 Spectral vertico-temporal repetitions relative to standard 313/1:1/50.

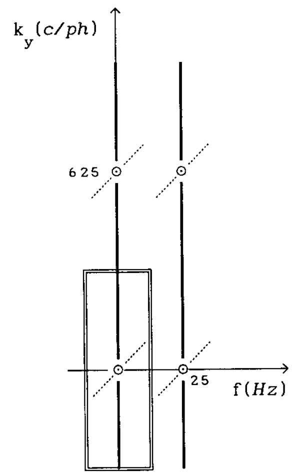

Fig. 4 Spectral vertico-temporal repetitions relative to standard 625/1:1/25.

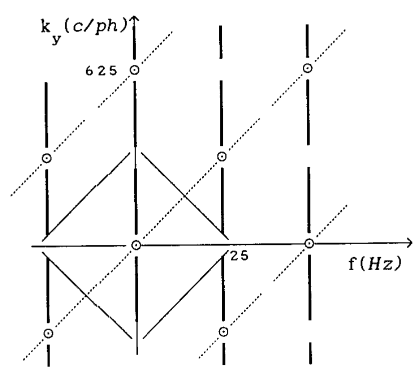

Fig. 5 Spectral vertico-temporal repetitions relative to standard 625/2:1/50.