Int. J. Electron. Commun. (AEÜ) 69 (2015) 487–491

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Mathematical modeling of coding gain and rate-distortion function in multihypothesis motion compensation for video signals

Andreja Samčović*

University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11040 Belgrade, Serbia

ARTICLE INFO

Article history:
Received 23 January 2013
Accepted 10 October 2014

Keywords:
Video compression
Coding gain
Rate-distortion function
Bit allocation
Lagrange method

ABSTRACT

This paper presents analytical model on motion-compensated predictive processing of the video signals. Multihypothesis prediction extends motion compensation with one prediction signal to the linear superposition of several motion-compensated prediction signals with the result of increased coding efficiency. The prediction error variances of the multihypothesis pictures are derived and minimized in this paper in order to find optimal bit allocations. Then, coding gains of the multihypothesis structures are calculated and compared, which is based on traditional prediction theories. Also, the corresponding rate-distortion functions are obtained in the closed form. Some experimental results show that more multihypotheses bring the smaller distortion of signals and hence improve the efficiency of video coding.

1. Introduction

For a transmission of video signals, sophisticated coding algorithms that utilize motion-compensation are required for a data compression of the video signal that results in an acceptable picture quality at the receiver. Motion-compensated coding schemes achieve data compression by exploiting the similarities between successive frames of a video signal [1]. Often, with such schemes, motion-compensated prediction (MCP) is combined with intra-frame coding of the prediction error [2–4]. Successful applications range from digital video broadcasting to low rate videoconferencing.

Many coders employ more than one motion-compensated prediction signal simultaneously to predict the current frame [5]. The term "multihypothesis motion compensation" has been coined for this approach. A linear combination of multiple prediction hypotheses is formed to arrive at the actual prediction signal. The theoretical motivations for multihypothesis motion compensation have been presented in [6,7]. Block-based multihypothesis MCP as introduced in [8] was a further effort to model the temporal correlation more efficiently. The approach is based on the idea that a combination of several prediction hypotheses is better suited to model complex temporal correlation. This is accomplished by selecting and combining several hypotheses from many previously decoded frames. Multihypothesis MCP has also the property to reduce the prediction noise which is introduced by the MCP model. This property was investigated theoretically in [3].

Our motivation is to compute performance bounds on coding gains while using multihypothesis structure over traditional intra-frame video coding scheme and to compare obtained results to the established algorithms. In this paper we introduce an analytical, block-based model for multihypothesis MCP. We propose a video coding scheme utilizing block-based multihypothesis MCP and discuss the coding efficiency of different multihypotheses structures with respect to the number of pictures in the group as well as time distance between pictures.

This paper is organized as follows. The structure of the multihypothesis is given in Section 2. The coding gains of the multihypotheses are derived and compared analytically and numerically in Section 3. The rate-distortion functions are obtained in Section 4. Finally, some concluding remarks are pointed out at the end of the paper.

2. Structure of the multihypothesis

In contrast to short-term or long-term MCP, we use N blocks from previous frames in order to predict one block in the current frame [8]. All blocks which are available for prediction are called hypotheses. The N hypotheses that predict a block are grouped to a multihypothesis or N-hypothesis. Constant predictor coefficients are used to combine linearly hypotheses of a multihypothesis. Fig. 1 shows the process of a prediction. The predicted block is determined by linear combination of the individual components.

* Tel.: +381 113091217; fax: +381 113096704.
E-mail address: andrej@sf.bg.ac.rs

http://dx.doi.org/10.1016/j.aeue.2014.10.007

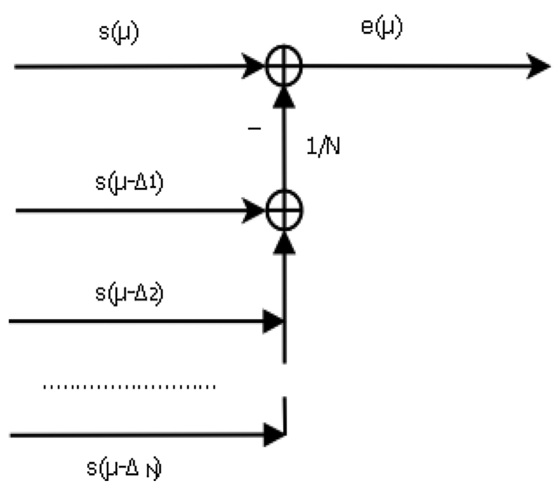

Fig. 1. Prediction of the multihypothesis.

In the above figure s is the input signal, e is the error signal, $\mu = \mu(x, y, t)$ represents a vector with three components: two spatial, statistically independent components x and y, and one temporal component t, while each hypothesis is addressed by a spatial-temporal displacement $\Delta = \Delta(\Delta_x, \Delta_y, \Delta_t)$, which is normally distributed. This address is relative to the position of the predicted block. One can see that

$$e(\mu) = s(\mu) - \frac{1}{N}\sum_{i=1}^{N} s(\mu - \Delta_i) \tag{1}$$

The variance of the prediction error signal $\sigma_e{}^2$ is given by the conditional expectation of the prediction error $E\left|e^2(\mu)\right|$:

$$E\left|e^2(\mu)\right| = \sigma_e^2; \quad E\left|e(\mu)\right| = 0 \tag{2}$$

$$\sigma_e^2 = \sigma_s^2 - \frac{2}{N}\sum_{i=1}^{N}\varphi_{SS}(\Delta_i) + \frac{1}{N^2}\sum_i\sum_j \varphi_{SS}(\Delta_i - \Delta_j) \tag{3}$$

where φ_{SS} is the autocorrelation function of the input signal s. It is connected with the input signal variance $\sigma_s{}^2$ and the correlation factor ρ:

$$\varphi_{SS}(\Delta) = \sigma_s^2 \rho(\Delta); \quad \rho(0) = 1 \tag{4}$$

Taking into account this formula, we get the following:

$$\sigma_e^2 = \left[1 - \frac{2}{N}\sum_{i=1}^{N}\rho(\Delta_i) + \frac{1}{N^2}\sum_i\sum_j \rho(\Delta_i - \Delta_j)\right]\sigma_s^2 \tag{5}$$

The correlation coefficient of the distance between two hypotheses is given by $\rho(\Delta_1 - \Delta_2) = \rho(\Delta_2 - \Delta_1) = \rho_s^{\Delta_h}$, where ρ_s is the spatial correlation coefficient and Δ_h is the distance between two hypotheses, respectively. The distance between two hypotheses means the physical distance between two blocks [9]. The equality between correlation coefficients holds because the components of the displacements Δ_1 and Δ_2 for both hypotheses are identical and independently Gaussian distributed [3].

When $N = 1$, we obtain the case with only predictive P-frames. For $N = 2$ hypotheses, the variance of the prediction error is given by

$$\sigma_e^2 = \left[\frac{3}{2} - \rho(\Delta_1) - \rho(\Delta_2) + \frac{1}{2}\rho(\Delta_1 - \Delta_2)\right]\sigma_s^2 \tag{6}$$

Fig. 2 represents the prediction with two hypotheses H_1 and H_2 taken from the previous frame, where Δ_r is the radial displacement of each hypothesis that means the Euclidian distance to the zero displacement error vector. We can see that $0 \le \Delta_h \le 2\Delta_r$. The displacements of the two hypotheses are given by $\Delta_1 = \Delta(\Delta_{x1}, \Delta_{y1}, \Delta_t)$ and $\Delta_2 = \Delta(\Delta_{x2}, \Delta_{y2}, \Delta_t)$, where $\sqrt{\Delta_{x1}^2 + \Delta_{y1}^2} = \sqrt{\Delta_{x2}^2 + \Delta_{y2}^2} = \Delta_r$.

Therefore, the correlation coefficients are equal: $\rho(\Delta_1) = \rho(\Delta_2) = \rho_s^{\Delta_r}\rho_t^{\Delta_t}$, where ρ_t is the temporal correlation coefficient. Taking that all into account, we obtained the variance of the prediction error for $N = 2$ hypotheses in the form

$$\sigma_e^2 = \left[\frac{3}{2} - 2\rho_S^{\Delta_r}\rho_t^{\Delta_t} + \frac{1}{2}\rho_S^{\Delta_h}\right]\sigma_s^2 \tag{7}$$

The prediction error variances are also calculated for the cases of $N = 3$, 4 and 8 hypotheses as well:

$$\sigma_e^2 = \left[\frac{4}{3} - 2\rho_S^{\Delta_r}\rho_t^{\Delta_t} + \frac{2}{3}\rho_S^{\sqrt{3}\Delta_r}\right]\sigma_s^2 \tag{8}$$

$$\sigma_e^2 = \left[\frac{5}{4} - 2\rho_S^{\Delta_r}\rho_t^{\Delta_t} + \frac{1}{2}\rho_S^{\sqrt{2}\Delta_r} + \frac{1}{4}\rho_S^{2\Delta_r}\right]\sigma_s^2 \tag{9}$$

$$\sigma_e^2 = \left[\frac{9}{8} - 2\rho_S^{\Delta_r}\rho_t^{\Delta_t} + \frac{1}{4}\rho_S^{\sqrt{2}\Delta_r} + \frac{1}{8}\rho_S^{2\Delta_r} + \frac{1}{4}\rho_S^{\sqrt{2-\sqrt{2}}\Delta_r} + \frac{1}{4}\rho_S^{\sqrt{2+\sqrt{2}}\Delta_r}\right]\sigma_s^2 \tag{10}$$

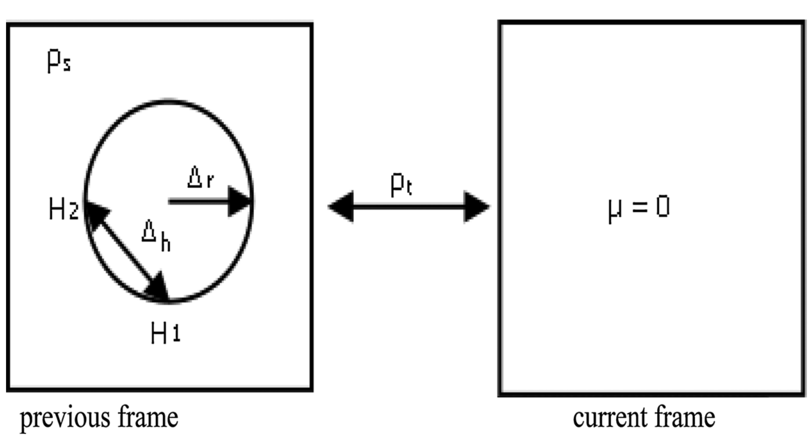

Fig. 2. Two-hypotheses prediction.