MULTIHYPOTHESIS MOTION-COMPENSATED PREDICTION WITH FORWARD-ADAPTIVE HYPOTHESIS SWITCHING

Markus Flierl

Telecommunications Laboratory
University of Erlangen-Nuremberg
Erlangen, Germany
flierl@LNT.de

Bernd Girod

Information Systems Laboratory
Stanford University
Stanford, CA
girod@ee.stanford.edu

ABSTRACT

Multihypothesis motion-compensating predictors combine several motion-compensated signals to predict the current frame signal. More than one motion-compensated signal, or hypothesis, is selected for transmission. Long-term memory motion-compensated prediction is a further concept for efficient video compression and is an example for forward-adaptive hypothesis switching. One motion-compensated signal is selected from multiple reference frames for transmission.

This paper extends the theory of multihypothesis motion-compensated prediction to forward-adaptive hypothesis switching. Assume, that we combine N hypotheses. Each hypothesis that is used for the combination is selected from a set of motion-compensated signals of size M. We study the influence of the hypothesis set size M on both the accuracy of motion compensation of forward-adaptive hypothesis switching and the efficiency of multihypothesis motion-compensated prediction. In both cases, we examine the noise-free limiting case. That is, we neglect signal components that are not predictable by motion compensation. Selecting one hypothesis from a set of motion-compensated signals of size M, that is, switching among M hypotheses, will reduce the displacement error variance by factor M when we assume statistically independent displacement errors. Integrating forward-adaptive hypothesis switching into multihypothesis motion-compensated prediction, that is, allowing a combination of switched hypotheses, increases the gain of multihypothesis motion-compensated prediction over the single hypothesis case for growing hypothesis set size M.

1. INTRODUCTION

Efficient video compression algorithms employ more than one motion-compensated signal simultaneously to predict the current frame of a video signal. The term "multihypothesis motion compensation" has been coined for this approach [1]. Theoretical investigations in [2] show that a linear combination of multiple prediction hypotheses can improve the performance of motion-compensated prediction. It is reported in [3] that an optimal multihypothesis motion estimation algorithm selects hypothesis such that their displacement error correlation coefficient is maximally negative. This optimal multihypothesis predictor exhibits the property that its gain over single hypothesis prediction increases for decreasing hypothesis displacement error variance.

Experimental results in [4] suggest that long-term memory motion compensation enhances the efficiency of multihypothesis prediction. Long-term memory motion compensation, as introduced in [5], extends each motion vector by a variable picture reference parameter which is able to address the reference frames in the long-term memory buffer. For video compression, the encoder has to select one reference frame per motion vector for transmission. In the following, choosing one signal from a set of motion-compensated signals will be called forward-adaptive hypothesis switching.

To obtain insight into the above mentioned experimental results, the theory of multihypothesis motion-compensated prediction is extended by forward-adaptive hypothesis switching. We combine both predictors and superimpose N hypotheses where each hypothesis is obtained by switching among M motion-compensated signals.

The paper is organized as follows: Section 2 summarizes briefly the signal model and the results for multihypothesis motion-compensated prediction. Section 3 introduces forward-adaptive hypothesis switching. We discuss the minimization of the radial displacement error and propose an equivalent predictor with just one hypothesis. Section 4 combines both predictors and investigates the overall performance.

2. MULTIHYPOTHESIS MOTION-COMPENSATED PREDICTION

Multihypothesis motion-compensated prediction is analyzed in [2] and optimal multihypothesis motion estimation is investigated in [3, 6]. This section briefly summarizes previous work and will help us to extend multihypothesis motion-compensated prediction by forward-adaptive hypothesis switching.

2.1. Signal Model

A signal model for motion-compensated prediction with two linearly combined hypotheses is depicted in Fig. 1.

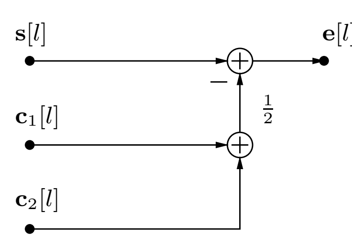

Fig. 1. Multihypothesis motion-compensated prediction with two linearly combined hypotheses.

The current frame signal $\mathbf{s}[l]$ at discrete location $l = (x, y)$ is predicted by averaging N hypotheses $\mathbf{c}_\nu[l]$ with $\nu = 1, ..., N$.

Obviously, motion-compensated prediction should work best if we compensate the true displacement of the scene exactly for each candidate prediction signal. Less accurate compensation will degrade the performance. To capture the limited accuracy of motion compensation, we associate a vector valued displacement error $\mathbf{\Delta}_\mu$ with the μ-th hypothesis $\mathbf{c}_\mu$. The displacement error reflects the inaccuracy of the displacement vector used for motion compensation. We assume a 2-D stationary normal distribution with variance $\sigma^2_{\mathbf{\Delta}}$ and zero mean where x- and y-components are statistically independent. The displacement error variance is the same for all N hypotheses. This is reasonable because all hypotheses are compensated with the same accuracy. Further, the pairs $(\mathbf{\Delta}_\mu, \mathbf{\Delta}_\nu)$ are assumed to be jointly Gaussian random variables. As there is no preference among the N hypotheses, the correlation coefficient $\rho_{\mathbf{\Delta}}$ between two displacement error components $\mathbf{\Delta}_{x\mu}$ and $\mathbf{\Delta}_{x\nu}$ is the same for all pairs of hypotheses.

For simplicity, we assume that all hypotheses $\mathbf{c}_\mu$ are shifted versions of the current frame signal $\mathbf{s}$. The shift is determined by the displacement error $\mathbf{\Delta}_\mu$ of the μ-th hypotheses. For that, the ideal reconstruction of the band-limited signal $\mathbf{s}[l]$ is shifted by the continuous valued displacement error and re-sampled on the original orthogonal grid.

The proposed model neglects "noisy" signal components and assumes that motion accuracy is basically the decision criterion for motion estimation.

2.2. Summary of Results

The 2-D power spectrum of the prediction error $\Phi_{\mathbf{ee}}$ normalized to the 2-D power spectrum of the current frame signal $\Phi_{\mathbf{ss}}$ is given in [3] as a function of the number of linearly combined hypotheses N, the 2-D Fourier transform P of the continuous 2-D pdf of the displacement error $\mathbf{\Delta}$, and the displacement error correlation coefficient $\rho_{\mathbf{\Delta}}$ of the linearly combined hypotheses.

$$\frac{\Phi_{\mathbf{ee}}(\omega)}{\Phi_{\mathbf{ss}}(\omega)} = \frac{N+1}{N} - 2P(\omega, \sigma^2_{\mathbf{\Delta}}) + \frac{N-1}{N} P\left(\omega, 2\sigma^2_{\mathbf{\Delta}}(1-\rho_{\mathbf{\Delta}})\right) \quad (1)$$

$$P(\omega, \sigma^2_{\mathbf{\Delta}}) = e^{-\frac{1}{2}\omega^T \omega \sigma^2_{\mathbf{\Delta}}} \quad \text{with} \quad \omega = (\omega_x, \omega_y)^T \quad (2)$$

It is shown in [6] that for optimal multihypothesis motion estimation the displacement error correlation coefficient $\rho_{\mathbf{\Delta}}$ is maximally negative.

$$\rho_{\mathbf{\Delta}} = \frac{1}{1-N} \quad \text{for} \quad N = 2, 3, 4, \ldots \quad (3)$$

The rate difference ΔR is used as a performance measure.

$$\Delta R = \frac{1}{4\pi^2} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{1}{2}\log_2\left(\frac{\Phi_{\mathbf{ee}}(\omega)}{\Phi_{\mathbf{ss}}(\omega)}\right) d\omega \quad (4)$$

It represents the maximum bit-rate reduction (in bit/sample) possible by optimum encoding of the prediction error $\mathbf{e}$, compared to optimum intra-frame encoding of the signal $\mathbf{s}$ for Gaussian wide-sense stationary signals for the same mean squared reconstruction error. A negative ΔR corresponds to a reduced bit-rate compared to optimum intra-frame coding.

It is assumed that the displacement error is entirely due to rounding and is uniformly distributed in the interval $[-2^{\beta-1}, 2^{\beta-1}] \times [-2^{\beta-1}, 2^{\beta-1}]$, where $\beta = 0$ for integer-pel accuracy, $\beta = -1$ for half-pel accuracy, $\beta = -2$ for quarter-pel accuracy, etc. The displacement error variance is

$$\sigma^2_{\mathbf{\Delta}} = \frac{2^{2\beta}}{12}. \quad (5)$$

One important result is that the gain of multihypothesis motion-compensated prediction with jointly optimal motion estimation over motion-compensated prediction increases by reducing the displacement error variance of each hypothesis.

3. FORWARD-ADAPTIVE HYPOTHESIS SWITCHING

Long-term memory motion-compensated prediction is very useful for efficient video compression [5]. The technique extends motion-compensated prediction such that previously decoded frames are utilized. This is achieved by permitting a variable reference picture selection for each block, where each reference picture is a previously decoded frame. The encoder has to select one reference frame per motion vector for transmission. In the following, choosing one signal from a set of motion-compensated signals is called forward-adaptive hypothesis switching.

3.1. Signal Model

A signal model for hypothesis switching is depicted in Fig. 2 for two hypotheses. The current frame signal $\mathbf{s}[l]$ at discrete location $l = (x, y)$ is predicted by switching among M hypotheses $\mathbf{c}_\mu[l]$ with $\mu = 1, ..., M$. The resulting prediction error is denoted by $\mathbf{e}[l]$.

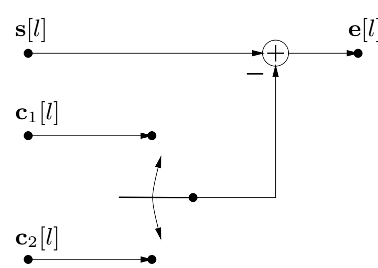

Fig. 2. Forward-adaptive hypothesis switching for motion-compensated prediction.

The assumptions in Section 2.1 also apply for this model. In addition, we assume statistically independent, spatially constant displacement errors for all hypotheses. It is further assumed that motion accuracy is basically the decision criterion for switching. In consequence, the hypothesis with the smallest displacement error is selected for prediction.

3.2. Minimizing the Radial Displacement Error

How does hypothesis switching improve the accuracy of motion-compensated prediction? Let us assume that the components of the displacement error for each hypothesis are i.i.d. Gaussian [2]. The Euclidean distance to the zero displacement error vector defines the *radial displacement error* for each hypothesis.

$$\mathbf{\Delta_r}_\mu = \sqrt{\mathbf{\Delta_x}_\mu^2 + \mathbf{\Delta_y}_\mu^2} \quad (6)$$

The hypothesis with minimum radial displacement error

$$\mathbf{\Delta_r}^M = \min_\mu(\mathbf{\Delta_r}_1, \ldots, \mathbf{\Delta_r}_\mu, \ldots, \mathbf{\Delta_r}_M) \quad (7)$$

is used to predict the signal.

In the following, hypothesis switching is described by means of the reliability function of the minimum radial displacement error. The reliability function of a random variable is closely related to the distribution function and is defined as the probability of the event $\{\mathbf{\Delta_r}^M > r\}$.

$$R_{\mathbf{\Delta_r}^M}(r) = \Pr\{\mathbf{\Delta_r}^M > r\} \quad (8)$$

The reliability function of the minimum radial displacement error can be expressed in terms of the reliability function of the set of M hypotheses.

$$\begin{aligned} R_{\mathbf{\Delta_r}^M}(r) &= \Pr\{\min_\mu(\mathbf{\Delta_r}_1, \ldots, \mathbf{\Delta_r}_\mu, \ldots, \mathbf{\Delta_r}_M) > r\} \\ &= \Pr\{\mathbf{\Delta_r}_1 > r, \ldots, \mathbf{\Delta_r}_M > r\} \\ &= R_{\mathbf{\Delta_r}_1 \ldots \mathbf{\Delta_r}_M}(r, \ldots, r) \end{aligned} \quad (9)$$

For switching two hypotheses, the probability of the event that the switched radial displacement error is larger than r is equal to the probability of the event that both radial displacement errors are larger than r.

Each displacement error is drawn from a 2-D normal distribution with zero mean and variance $\sigma^2_{\mathbf{\Delta}_\mu}$ [2]. The displacement errors of the M hypotheses are assumed to be statistically independent. The x- and y-components of the displacement errors are arranged to vectors $\mathbf{\Delta_x}$ and $\mathbf{\Delta_y}$, respectively.

$$p_{\mathbf{\Delta_x}\mathbf{\Delta_y}}(\Delta_x, \Delta_y) = \frac{1}{(2\pi)^M |C_{\mathbf{\Delta_x}}|} e^{-\frac{1}{2}\left[\Delta_x^T C_{\mathbf{\Delta_x}}^{-1} \Delta_x + \Delta_y^T C_{\mathbf{\Delta_x}}^{-1} \Delta_y\right]} \quad (10)$$

$$C_{\mathbf{\Delta_x}} = C_{\mathbf{\Delta_y}} = \begin{pmatrix} \sigma^2_{\mathbf{\Delta_x}1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\mathbf{\Delta_x}2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\mathbf{\Delta_x}M} \end{pmatrix} \quad (11)$$

The criterion for switching is the radial displacement error. To obtain a closed-form expression of the radial displacement error pdf, it is assumed that the variances in x- and y-direction are identical. With this assumption, we can easily determine the probability density function of $\mathbf{\Delta_r}$ [7]:

$$p_{\mathbf{\Delta_r}}(\Delta_r) = \left(\prod_{\mu=1}^{M} \Delta_{r\mu}\right) \frac{1}{|C_{\mathbf{\Delta_x}}|} e^{-\frac{1}{2}\Delta_r^T C_{\mathbf{\Delta_x}}^{-1} \Delta_r} \quad (12)$$

An M-dimensional Rayleigh pdf is obtained describing M independent radial displacement errors.

In order to minimize the radial displacement error, the M-dimensional reliability function of the displacement error has to be determined.

$$\begin{aligned} R_{\mathbf{\Delta_r}_1 \ldots \mathbf{\Delta_r}_M}(\Delta_{r1} \ldots \Delta_{rM}) &= \int_{\Delta_{r1}}^{\infty} \cdots \int_{\Delta_{rM}}^{\infty} p_{\mathbf{\Delta_r}}(u) du \\ &= e^{-\frac{1}{2}\Delta_r^T C_{\mathbf{\Delta_x}}^{-1} \Delta_r} \end{aligned} \quad (13)$$

The reliability function of the minimum radial displacement error $R_{\mathbf{\Delta_r}^M}(r)$ is obtained by evaluating the M-dimensional reliability function at the same value r for all dimensions: $R_{\mathbf{\Delta_r}^M}(r) = R_{\mathbf{\Delta_r}_1 \ldots \mathbf{\Delta_r}_M}(\mathbf{1}r)$. The vector $\mathbf{1}$ contains for each dimension the value one. The minimum radial displacement error is also Rayleigh distributed. It is noted that a one-dimensional pdf is given by the negative derivative of the reliability function.

$$R_{\mathbf{\Delta_r}^M}(r) = e^{-\frac{1}{2}\frac{r^2}{\alpha^2}} \quad (14)$$

$$p_{\mathbf{\Delta_r}^M}(r) = \frac{r}{\alpha^2} e^{-\frac{1}{2}\frac{r^2}{\alpha^2}} \quad (15)$$

$$\alpha^2 = \frac{1}{\mathbf{1}^T C_{\mathbf{\Delta_x}}^{-1} \mathbf{1}} \quad (16)$$

The variance of the minimum radial displacement error is of interest. The covariance matrix of the Rayleigh pdf in (12) is $C_{\mathbf{\Delta_r}} = (2 - \frac{\pi}{2}) C_{\mathbf{\Delta_x}}$ and the variance of the switched radial displacement error is given by $\sigma^2_{\mathbf{\Delta_r}^M} = (2 - \frac{\pi}{2})\alpha^2$ [8]. In order to omit the constant factor, the variance of the minimum radial displacement error is stated as a function of the covariance matrix $C_{\mathbf{\Delta_r}}$.

$$\sigma^2_{\mathbf{\Delta_r}^M} = \frac{1}{\mathbf{1}^T C_{\mathbf{\Delta_r}}^{-1} \mathbf{1}} \quad (17)$$

For example, the variances of the radial displacement errors might be identical for all M hypotheses. (17) implies that switching of independent Rayleigh distributed radial displacement errors reduces the variance by factor M.

$$\sigma^2_{\mathbf{\Delta_r}^M} = \frac{\sigma^2_{\mathbf{\Delta_r}}}{M}. \quad (18)$$

3.3. Equivalent Predictor

Section 3.2 shows that both the individual radial displacement errors and the minimum radial displacement error are Rayleigh distributed. This suggests to define an equivalent motion-compensating predictor for switched prediction. This predictor uses just one hypothesis but the variance of its displacement error is much smaller. The distribution of the switched displacement error is assumed to be separable and normal with zero mean and variance

$$\sigma^2_{\mathbf{\Delta_x}^M} = \frac{1}{\mathbf{1}^T C_{\mathbf{\Delta_x}}^{-1} \mathbf{1}}. \quad (19)$$

The equivalent predictor with reduced displacement error variance represents the more accurate motion compensation achieved by switched prediction. Consequently, forward-adaptive hypothesis switching lowers the energy of the motion-compensated prediction error.

4. MULTIHYPOTHESIS PREDICTION WITH FORWARD-ADAPTIVE HYPOTHESIS SWITCHING

It has been demonstrated in [4] that the linear combination of hypotheses is more efficient when the hypotheses are obtained by long-term memory motion compensation. We know from Section 3 that choosing among M reference frames can reduce the displacement error variance by as much as factor M. This lower bound is obtained for switching among M hypotheses with statistically independent displacement errors.

Forward-adaptive hypothesis switching reduces the variance of the displacement error $\sigma_{\mathbf{\Delta}}^2$ in (1). When we assume that motion compensation is performed with half-pel accuracy ($\beta = -1$), Fig. 3 depicts the rate difference over the size M of the motion-compensated signal set. The performance of $N = 2, 4, 8$, and ∞ linearly combined hypotheses is compared to motion-compensated prediction with forward-adaptive hypothesis switching ($N = 1$).

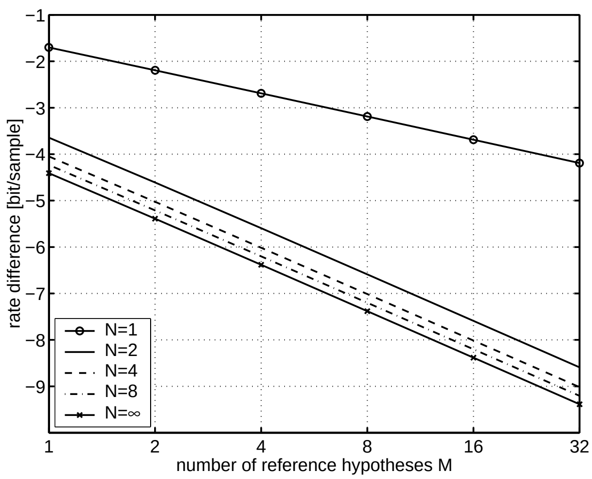

Fig. 3. Rate difference over the number of reference hypotheses M for multihypothesis motion-compensated prediction with forward-adaptive hypothesis switching. The switched hypotheses are just averaged and no residual noise is assumed. The displacement inaccuracy β is set to -1.

We can observe in Fig. 3 that doubling the number of reference hypotheses decreases the bit-rate for motion-compensated prediction by 0.5 bit/sample and for multihypothesis motion-compensated prediction by 1 bit/sample. The gain going from $N = 1$ to $N = 2$ is the largest, independent of the number of reference hypotheses M. In addition, this gain increases for a larger number of available motion-compensated signals M. This theoretical result supports our experimental finding reported in [4].

5. CONCLUSIONS

This paper investigates the performance of multihypothesis motion-compensated prediction with forward-adaptive hypothesis switching. For that we linearly combine N hypotheses. Each hypothesis that is used for the combination is selected from a set of motion-compensated signals of size M. We study the influence of the hypothesis set size M on both the accuracy of motion compensation of forward-adaptive hypothesis switching and the efficiency of multihypothesis motion-compensated prediction. In both cases, we examine the noise-free limiting case.

It is shown that forward-adaptive switching among M hypotheses with statistically independent displacement errors reduces the displacement error variance by factor M. Reducing the displacement error variance of each linearly combined hypothesis increases the gain by multihypothesis motion-compensated prediction with jointly optimal motion estimation over motion-compensated prediction.

We combine the two predictors, allow a superposition of switched hypotheses and improve the gain by multihypothesis motion-compensated prediction over single hypotheses prediction. Experimental results in [4] support this theoretical finding. For the ideal predictors we observe that doubling the number of reference hypotheses M decreases the bit-rate for motion-compensated prediction by 0.5 bit/sample and for multihypothesis motion-compensated prediction by 1 bit/sample.

6. ACKNOWLEDGMENT

The authors would like to thank Thomas Wiegand for helpful discussions and useful comments on the theory.

7. REFERENCES

[1] G.J. Sullivan, "Multi-Hypothesis Motion Compensation for Low Bit-Rate Video Coding," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Minneapolis, Apr. 1993, vol. 5, pp. 437–440.

[2] B. Girod, "Efficiency Analysis of Multihypothesis Motion-Compensated Prediction for Video Coding," *IEEE Transactions on Image Processing*, vol. 9, no. 2, pp. 173–183, Feb. 2000.

[3] M. Flierl, T. Wiegand, and B. Girod, "A Video Codec Incorporating Block-Based Multi-Hypothesis Motion-Compensated Prediction," in *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, Perth, Australia, June 2000, vol. 4067, pp. 238–249.

[4] M. Flierl, T. Wiegand, and B. Girod, "Rate-Constrained Multi-Hypothesis Motion-Compensated Prediction for Video Coding," in *Proceedings of the IEEE International Conference on Image Processing*, Vancouver, Canada, Sept. 2000, vol. III, pp. 150–153.

[5] T. Wiegand, X. Zhang, and B. Girod, "Long-Term Memory Motion-Compensated Prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 70–84, Feb. 1999.

[6] M. Flierl and B. Girod, "Multihypothesis Motion Estimation for Video Coding," in *Proceedings of the Data Compression Conference*, Snowbird, Utha, Mar. 2001.

[7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1991.

[8] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1994.