Block Adaptive Quantization of Multiple Frame Motion Field

F. Lavagetto †‡and R. Leonardi †*
†AT&T Bell Labs/ Visual Comm. Res. Dept.
Holmdel, NJ 07733, U.S.A.
‡DIST, University of Genova
Genova, I-16145 Italy
* DE/LTS, Swiss Federal Institute of Technology
Lausanne, CH-1015 Switzerland

Abstract

This paper presents a method for the adaptive quantization of the motion field obtained from multiple reference frames. The motion estimation is obtained through a block-matching technique, with the assumptions of pure translational motion and uniform motion within a block. The regular assumption of constant intensity along the motion trajectory in the spatio-temporal path is relaxed. On one hand, this allows for illumination changes between the reference frames and the frame to be motion-compensated. On the other hand, it allows for additional freedom to reduce the prediction error after motion-compensation when the translational and rigid body motion assumptions are violated. The matching function represents the difference for each block of the luminance signal in the frame to be encoded with respect to a linear combination of 2 displaced blocks of same size in the reference frames. The adaptive quantization mechanism is based on evaluating, on a block basis, the local sensitivity of the displaced frame difference signal to a quantization of the motion field parameters. It is shown how such sensitivity depends only on the reference frame signals, which allows to keep it below a desired threshold without additional information to be sent to a receiver. Simulations are carried out on standard CIF (240x360) source material provided to the ISO/MPEG. Results are discussed to show the improvement with respect to the strategy suggested in the draft recommendation of the ISO MPEG for interactive video at 1.5 Mbps.

1 Introduction

The most promising algorithms for video coding employ motion-compensated prediction for exploiting interframe redundancy [1]. Two nearby frames typically exhibit high correlation, and the entropy of the frame difference is lower than that of the original frames. This frame difference entropy is further reduced if motion compensation is applied between the two frames. Fig. 1 shows a prototype video coder to encode a sequence of images.

Motion compensation works as follows. The motion between an image frame to be encoded and a reference frame already encoded is estimated. The estimation of the motion is encoded and employed to obtain a representation of the frame to be encoded (the motion-compensated predicted frame), if the reference frame is available. To this end all frames of the video image are partitioned into a set of blocks comprised of $Q \times R$ picture elements (pels), typically 16×16. Each block B of a frame to be encoded in turn is assigned a motion vector, i.e., a displacement $\mathbf{d}$ relative to the location of this block B with respect to a same size block in the reference frame that best matches the pel values in B.

The reference frame is typically a frame in the past relative to the frame to be encoded. The time interval separating both frames is chosen according to the available channel bandwidth. The closer this frame, the better the motion rendition and the more accurate the estimation of the motion vector $\mathbf{d}$. This motion estimate describes therefore the displacement of a block of pixels from the time instant of the reference frame until the time instant of the frame to be encoded.

For this purpose, it is often assumed that there is no illumination change along the motion trajectory of the block B. Let us call $I(\mathbf{x}, t)$ the image signal where $\mathbf{x}$ represent the spatial coordinates and t the time variable. The constant illumination assumption becomes on all differentiable points of the motion trajectory (e.g. non boundary points of moving objects):

$$\frac{dI(\mathbf{x},t)}{dt} = \frac{\partial I}{\partial \mathbf{x}} \cdot \frac{d\mathbf{x}}{dt} + \frac{\partial I}{\partial t} = 0 \qquad (1)$$

where $\frac{\partial I}{\partial \mathbf{x}} = \nabla_{\mathbf{x}} I$ denotes the spatial gradient and $\frac{d\mathbf{x}}{dt} = \mathbf{v}$ denotes the velocity along the motion trajectory. Equation (1) contrains $\mathbf{v}$ to lie on a straight line in the (v_x, v_y) space and it is perpendicular to $\nabla_{\mathbf{x}} I$. Several points having the same velocity will intersect at a single point uniquely specifying the uniform velocity for the desired block B. This approach was originally suggested by Horn and Schunk [2]. Further to this assumption, it is often considered that the displacement d corresponds to a pure translation and that the motion remains uniform within each $Q \times R$ block. Given the nature of image sequences, where a 3-D scene gets projected on an image plane through the camera optics, occlusion of objects is unavoidable, which results in incorrect motion estimates. The following reasons account for this bias in the motion estimation:

- Between 2 successive frames, a correspondence between all $Q \times R$ blocks of the frame to be encoded and blocks of the same size in the reference frame is never guaranteed.
- B blocks may contain different moving parts whereas a constant uniform displacement is assumed.

Moreover, projections of 3-D moving objects on the image plane will only correspond to a pure translation on very limited cases. Finally deformation of objects cannot be handled with such simplifying assumptions. Nevertheless, this model is a trade-off between estimating the exact motion field and algorithmic complexity. It remains sufficiently accurate in most cases especially if the time interval separating both frames is small, and if the objective of video compression is to minimize the entropy of the prediction, rather than finding the exact motion.

To select a displacement as the motion vector for block B, a search among a set of candidate displacements is performed. Typically, the candidate set is comprised of all the displacements of a predetermined step size (e.g. half pel, single pel) that are within a fixed range of the location of block B. We shall denote this candidate set by D. Most often an exhautive search strategy is applied to find the optimal candidate to a certain matching criterion. As a matching criterion, the integral of the absolute error signal is many times used as it is computationally efficient and it is easily implementable in VLSI. Let us call n the time instant of the frame to be encoded and $n-1$ the time instant of the reference frame. The optimum displacement $\mathrm{d}^*(\mathrm{x})$ will then satisfy the following condition

$$\mathrm{d}^*(\mathrm{x}) = \arg[\min_{\mathrm{d} \in D} \sum_{\mathrm{x}' \in N(\mathrm{x})} |I(\mathrm{x}', n) - I(\mathrm{x}' + \mathrm{d}, n-1)|] \qquad (2)$$

where x' is the location of an individual pel that is a member of $N(\mathrm{x})$, the set of all pel locations in block B. $I(\mathrm{x}, t)$ is a function that typically yield the luminance value of the image sequence at the location and time specified by its arguments. In some implementations, a good estimate of the optimal displacement is determined through a more limited search effort (i.e. time of search). Several well known methods are the 3-step search, the logarithmic search, and the conjugate direction search. Among other matching functions, the L_2 norm of the error signal has also been selected. As an alternative approach, a good motion estimate can be obtained by searching for the maximum of the normalized cross correlation function (also called phase correlation) between the block B and the search area in the reference frame [3]. $\mathrm{d}^*(\mathrm{x})$ is then expressed by

$$\mathrm{d}^*(\mathrm{x}) = \arg[\max_{\mathrm{d} \in D} \frac{\sum_{\mathrm{x}' \in N(\mathrm{x})} I(\mathrm{x}', n) I(\mathrm{x}' + \mathrm{d}, n-1)}{\sqrt{\sum_{\mathrm{x}' \in N(\mathrm{x})} I^2(\mathrm{x}', n)} \times \sqrt{\sum_{\mathrm{x}' \in N(\mathrm{x}+\mathrm{d})} I^2(\mathrm{x}', n-1)}}] \qquad (3)$$

The displacement selected as the motion vector and the corresponding error signal for each B block are employed as a representation and can thus be further quantized and encoded as appropriate for transmission and storage. Little effort has been made to jointly estimate a desired quantization for such motion and prediction information. Some attempt will be made here in the case of multiple reference frames as will be seen later.

Better compression of the data that represent the video images can be achieved by motion-compensated interpolation [4], a motion estimation technique which incorporates an additional reference frame, which is located in the future relative to the frame to be encoded. An interpolative system is used to predict any frames that are temporally between the reference frames. Motion-compensated interpolation has demonstrated superiority compared to frame repeat that produces undesirable jerkiness or frame averaging that creates bluriness, at little additional cost in bandwidth. In effect, interpolation errors do not feed back in the DPCM prediction loop of a motion compensated predictive scheme, so that this information may be quantized more coarsely than prediction errors of motion-compensated predicted frames. In addition, if the interpolation error is significant only for blocks containing several moving objects, spatio-temporal masking can be efficiently used to lower the quality of the reconstruction on these blocks without making it visually perceptible.

The interpolation determines for each block B a block in the past reference frame denoted by $I(\mathbf{x}, m)$ and a block in the future reference frame denoted by $I(\mathbf{x}, p)$, that when combined yields to a best approximation of the block B. The combination is typically a weighted sum of the values of the pels of the selected blocks. The displacements denoted by $\mathbf{d}_m$ and $\mathbf{d}_p$ respectively from the block to be encoded to each of the determined blocks are employed as estimates of the motion of the block relative to each of the reference frames and are taken as motion vectors. From the weighted sum of the 2 reference blocks, an interpolation error is obtained by subtracting the weigthed sum from the values of the pels forming block B. Weights, motion vectors and interpolation error signal serve as a representation for each block. They may be quantized and encoded as appropriate for transmission and storage.

The determination of the weights for the weighted sum is generally performed by employing predetermined limitations on either the weights or the displacement candidates. For example, ISO MPEG draft proposal [5] suggests to take on a block by block basis either equal contribution $(1/2, 1/2)$ from both displaced blocks or selecting to use for a block a contribution from just one of the reference frames $(0, 1)$ and $(1, 0)$. Assuming that the uniform translational motion assumption is correct, each block will have a correspondence either in the past or the future or both reference frames, depending whether this block was occluded, or uncovered during the time interval that separates both reference frames. The criteria to select which configuration to use is often based on the one that leads to the minimum energy of the interpolation error signal.

The purpose of this paper is to present a more general formalism to optimally select the set of displacements and weights without imposing any predetermined limitation on the weight sample values (continuous optimization). This way the regular assumption of constant intensity along the motion trajectory in the spatio-temporal path is relaxed. On one hand, this allows for illumination changes between the reference frames and the frame to be motion-compensated. On the other hand, it allows for additional freedom to reduce the interpolation error after motion-compensation when the translational and rigid body motion assumptions are violated.

The formalism for the joint optimization of the weight/displacement set is presented with multiple reference frames in section 2. This section demonstrates as well how the proposed solution will converge to a minimum of the matching criterion. It discusses also how to impose certain constraints on the optimization problem to take into account reasonable assumptions on the displacements joint distribution and on the weights joint distribution, based on measurements obtained from real video source material. Section 3 discusses implementation aspects and fast search procedures showing computational advantages. Section 4 analyses quantization aspects of the weights: it shows how the quantization step size affects the interpolation error only based on the reference frame signals. Section 5 presents some simulation results on standard CIF (240x360) source material provided to the ISO/MPEG. Results are discussed to show the improvement with respect to the strategy suggested in the draft recommendation of the ISO MPEG for interactive video at 1.5 Mbps.

2 Multiple frame motion estimation

2.1 The two reference frame case

Fig. 2 shows 3 sets of pels a, b, and c from the same representative area of 3 frames $I(\mathbf{x}, m)$, $I(\mathbf{x}, n)$, and $I(\mathbf{x}, p)$ of an input video signal. Frame $I(\mathbf{x}, n)$, is to be encoded via motion compensated interpolation. $I(\mathbf{x}, m)$ is a reference frame in the past relative to $I(\mathbf{x}, n)$ while $I(\mathbf{x}, p)$ is a reference frame in the future relative to frame $I(\mathbf{x}, n)$. $Q \times R$ block of pels B is a subset of pels contained in a, and corresponds to the block to be encoded for which a search for motion vectors and corresponding weights is to be performed. Although the optimization procedure is presented with 2 reference frames $I(\mathbf{x}, m)$, and $I(\mathbf{x}, p)$, it can easily be extended to more reference frames. As we shall see, it is unnecessary to impose the causality of the order in which the frames are taken. In other words, nothing imposes that $I(\mathbf{x}, n)$ has to stay within both reference frames, but this is done for the clarity of the presentation and in accordance with the interpolation terminology.

To determine the motion vectors and weights for block B, the displacements pointing to a block of pels in each of frames $I(\mathbf{x}, m)$, and $I(\mathbf{x}, p)$ must be found such that the weighted sum of the pel values forming the pointed blocks approximates the signal in B with a minimum error signal E. In other words, we look for the solution to

$$\min_{\mathbf{u}, \mathbf{d}_m \in D_m, \mathbf{d}_p \in D_p} \sum_{\mathbf{x}' \in N(\mathbf{x})} E^2(\mathbf{x}') \quad (4)$$

where $u^T = [\alpha_m \alpha_p]$ is the weight vector accorded to each pel value of the respective block in both reference frames; d_m and d_p represent the displacements for frames $I(x, m)$, and $I(x, p)$, respectively; D_m and D_p represent the sets of candidate displacements in both reference frames; and the interpolation error signal is given by

$$E(x) = I(x, n) - (\alpha_m I(x + d_m, m) + \alpha_p I(x + d_p, p)) \quad (5)$$

A candidate displacement pair (d_m, d_p) is graphically depicted in Fig. 2 by the 2 arrows starting from block B in frame $I(x, n)$. Also shown in Fig. 2 are the search areas corresponding to the sets D_m and D_p. These are illustrated by the shaded areas in both reference frames.

The solution to (4) is obtained in two stages. In the first one it is imposed that the partial derivatives of $\sum_{x' \in N(x)} E^2(x')$ with respect to α_m and α_p are equal to zero. This leads to the following system of two linear equations

$$M.u = v \quad (6)$$

where u is the weight vector as expressed above and vector v is given by

$$v = \begin{bmatrix} \sum_{x' \in N(x)} I(x' + d_m, m) I(x', n) \\ \sum_{x' \in N(x)} I(x', n) I(x' + d_p, p) \end{bmatrix} \quad (7)$$

and where M is given by

$$M = \begin{bmatrix} \sum_{x' \in N(x)} I^2(x' + d_m, m) & \sum_{x' \in N(x)} I(x' + d_m, m) I(x' + d_p, p) \\ \sum_{x' \in N(x)} I(x' + d_m, m) I(x' + d_p, p) & \sum_{x' \in N(x)} I^2(x' + d_p, p) \end{bmatrix} \quad (8)$$

In Appendix 1, we show that matrix M can be shown remains definite positive for all values of x, d_m and d_p, thereby guaranteeing that the corresponding optimal values of the weights $(u^*)^T = \alpha_m^*$ and α_p^* will lead to a minimum of the interpolation error energy for that particular pair d_m, d_p. Singular values of M result in an infinite set of solutions u^* that satisfy (4). Such degenerated cases occur, for example, when there exists a perfect match (no error) between the B block and a same size block in reference frames $I(x, m)$ and $I(x, p)$. For such degenerate cases, any arbitrary value may be chosen as the optimum value for one of the weights and the optimum value for the remaining weight is determined by solving either of the two linear equations represented by (6). For a more complete discussion of such cases refer to Appendix 1.

In non degenerate cases, the optimal solution u^* is given by

$$u^* = M^{-1}.v \quad (9)$$

As can be seen, ideal weights are functions dependent on the displacement pair d_m, d_p. Therefore, the optimum solution requires the selection of a pair that will satisfy equation (6) and has the lowest E energy. To find such an optimum pair d_m, d_p, requires a search among a set of candidate displacement pairs represented by the cartesian product $D_m \times D_p$. For each candidate displacement pairs, matrices M, M^{-1} and vector v are computed so that equation (9) yields particular values for u^*. The values of u^* and their corresponding displacement pair d_m, d_p are then employed to compute the energy of E. The triplet (u^*, d_m, d_p) of a particular displacement pair is stored together with its associated energy value if it yields a smaller energy than any other displacement pair already evaluated has yielded. The triplet that remains stored when all pair of candidate displacements have been examined is comprised of the optimum weights and displacements for the corresponding block B.

The advantage of modeling continuously the weighting factors the uncovered/covered area problem is automatically taken into account. As we shall see in section 5, the distribution of weight pairs α_m, α_p is centered around the axis $\alpha_m + \alpha_p = 1$. There is no need any longer to impose the predetermined pairs (0, 1) and (1, 0) when there exists only one reference frame in the past or in the future. In addition, this approach handles illumination changes that may exist between the 3 frames, which is very adequate when the time interval separating both reference frames gets large.

2.2 Generalization of the formalism

The generalization of this approach to more than two reference frames is straightforward. Let us assumed, we considered N reference frames occuring at time instant $m_1, m_2, ..., m_N$. The minimization task is equivalent to find the pair $(\mathbf{u}, \mathbf{D})$ that is given by

$$(\mathbf{u}, \mathbf{D})(\mathbf{x}) = \arg\left[\min_{\mathbf{u}\in R, \mathbf{D}\in\prod_{i=1}^{N} D_i} \sum_{\mathbf{x}'\in N(\mathbf{x})} E^2(\mathbf{x}')\right]$$

$$= \arg\left[\min_{\mathbf{u}\in R, \mathbf{D}\in\prod_{i=1}^{N} D_i} \sum_{\mathbf{x}'\in N(\mathbf{x})} (I(\mathbf{x}', n) - \mathbf{u}^T \mathbf{I}(\mathbf{x}'))^2\right] \quad (10)$$

where the weight vector is $\mathbf{u}^T = [\alpha_{m_1} \alpha_{m_2} \ldots \alpha_{m_n}]$. $\mathbf{I}(\mathbf{x})$ represents the reference frame vector and is defined by $\mathbf{I}^T = [I(\mathbf{x}, m_1) I(\mathbf{x}, m_2) \ldots I(\mathbf{x}, m_N)]$; $\mathbf{D}$ represents the set of N displacement vectors and is expressed in a matrix form by $\mathbf{D}^T = [\mathbf{d}_{m_1} \mathbf{d}_{m_2} \ldots \mathbf{d}_{m_N}]$. In this approach it is obvious that in each block B the signal in the frame to be encoded is seen as a linear combination of displaced blocks in N reference frames.

Using the same two step approach as before by setting first the partial derivative of the energy of the interpolation error signal to 0 with respect to vector $\mathbf{u}$ leads to a system of N linear equations expressed by the matrix vector relation of equation (6). $\mathbf{M}$ is an $N \times N$ matrix is given by

$$\mathbf{M} = \begin{bmatrix} \sum_{\mathbf{x}'} I^2(\mathbf{x}' + \mathbf{d}_{m_1}, m_1) & M_{2,1} & \ldots & M_{N,1} \\ \sum_{\mathbf{x}'} I(\mathbf{x}' + \mathbf{d}_{m_1}, m_1) I(\mathbf{x}' + \mathbf{d}_{m_2}, m_2) & \sum_{\mathbf{x}'} I^2(\mathbf{x}' + \mathbf{d}_{m_2}, m_2) & \ldots & M_{N,2} \\ \ldots & \ldots & \ldots & \\ \sum_{\mathbf{x}'} I(\mathbf{x}' + \mathbf{d}_{m_1}, m_1) I(\mathbf{x}' + \mathbf{d}_{m_N}, m_N) & \ldots & \ldots & \sum_{\mathbf{x}'} I^2(\mathbf{x}' + \mathbf{d}_{m_N}, m_N) \end{bmatrix} \quad (11)$$

The N-dimensional vector $\mathbf{v}$ becomes now

$$\mathbf{v} = \begin{bmatrix} \sum_{\mathbf{x}'\in N(\mathbf{x})} I(\mathbf{x}' + \mathbf{d}_{m_1}, m_1) I(\mathbf{x}', n) \\ \sum_{\mathbf{x}'\in N(\mathbf{x})} I(\mathbf{x}' + \mathbf{d}_{m_2}, m_2) I(\mathbf{x}', n) \\ \ldots \\ \sum_{\mathbf{x}'\in N(\mathbf{x})} I(\mathbf{x}' + \mathbf{d}_{m_N}, m_N) I(\mathbf{x}', n) \end{bmatrix} \quad (12)$$

As long as the matrix is non singular, an optimal weight $\mathbf{u}^*$ can be found by solving the system as described by equation (9) corresponding to a candidate displacement vector $\mathbf{D}$. The strategy to follow thereafter is similar to the 2 reference frame case, where all sets of candidate displacements are searched till it is found which one lead to a minimum of the E energy.

2.3 Imposing constraints to the model

The formalism presented in the previous subsections is relatively general; it can handle complex phenomena such as multiple occlusions and illumination changes between frames. It however increases significantly the computational complexity as will be presented in section 4 and does not include a series of assumptions that have shown to be very powerful in the context of computing motion fields. We present now different alternatives that lead to possible improvements in the estimation process, by imposing a number of reasonable constraints to the set of displacement vectors, and weight vector. We present also how smoothness in the motion field on non boundary points of moving objects can be guaranteed through the use of specific a priori information. Results are written for the 2 reference frame case but can easily be extended to the N reference frame case.

2.3.1 Weight vector constraint

If the assumption of constant translational velocity for all points within a block B is correct, it is reasonable to believe that for every block there will be always a match in at least one reference frame. In section 5, we shall show that this hypothesis is correct by looking at the distribution of weight values α_m, α_p. It turns out that these values are distributed around the straight line

$$\alpha_m + \alpha_p = 1 \quad (13)$$

By setting $\alpha = \alpha_m$, the interpolation error expressed in (5) becomes

$$E(\mathrm{x}) = I(\mathrm{x}, n) - I(\mathrm{x} + \mathrm{d}_p, p) - \alpha\,(I(\mathrm{x} + \mathrm{d}_m, m) - I(\mathrm{x} + \mathrm{d}_p, p)\,) = X - \alpha Y \tag{14}$$

The optimal weight corresponding to a minimum of (4), is obtained by taking the partial derivative of the interpolation error energy with respect to α and setting it to 0. This leads to

$$\sum_{\mathrm{x}' \in N(\mathrm{x})} 2Y.(X - \alpha Y) = 0 \tag{15}$$

which implies the solution

$$\alpha = \frac{\sum_{\mathrm{x}' \in N(\mathrm{x})} XY}{\sum_{\mathrm{x}' \in N(\mathrm{x})} Y^2} = \frac{\sum_{\mathrm{x}' \in N(\mathrm{x})}[I(\mathrm{x}', n) - I(\mathrm{x}' + \mathrm{d}_p, p)].[I(\mathrm{x}' + \mathrm{d}_m, m) - I(\mathrm{x}' + \mathrm{d}_p, p)]}{\sum_{\mathrm{x}' \in N(\mathrm{x})}[I(\mathrm{x}' + \mathrm{d}_m, m) - \sum_{\mathrm{x}' \in N(\mathrm{x})} I(\mathrm{x}' + \mathrm{d}_p, p)]^2} \tag{16}$$

as long as $\sum_{\mathrm{x}' \in N(\mathrm{x})} I(\mathrm{x}' + \mathrm{d}_m, m) - I(\mathrm{x}' + \mathrm{d}_p, p) \neq 0$. If this last inequality does not hold, the interpolation error enery is independent of α, as can be seen from the espression of the interpolation error (14). By replacing this expression into (4), the interpolation error energy can be written in this case as

$$\epsilon = \sum_{\mathrm{x}' \in N(\mathrm{x})} (X^2 - (\alpha Y)^2) = \sum_{\mathrm{x}' \in N(\mathrm{x})} X^2 - \frac{[\sum_{\mathrm{x}' \in N(\mathrm{x})}(XY)]^2}{\sum_{\mathrm{x}' \in N(\mathrm{x})} Y^2} \tag{17}$$

Finding the optimal displacement pair corresponds to finding the candidate one that will result in a minimum expression for (17).

2.3.2 Velocity constraint

It is also possible to impose a certain constraint on the velocity field due to the inertia of most moving bodies in space. Assuming each object (block) is present in each considered frame ($I(\mathrm{x}, m)$, $I(\mathrm{x}, n)$ and $I(\mathrm{x}, p)$) so that there are no occlusions, and given that the time instant at which the frames are sampled is known, it is possible to establish a model for the temporal evolution of the displacement vectors. In other words, we can write

$$\mathrm{d} = \mathrm{d}(t) \tag{18}$$

with the initial condition $\mathrm{d}(n) = 0$ as the displacement is always supposed to be measured with respect to the frame to be motion-compensated, i.e. I(x, n). As an example, we can have a quadratic model for the displacement vectors, e.g.

$$\mathrm{d}(t) = \mathrm{A}(t - n)^2 + \mathrm{L}(t - n) \tag{19}$$

where A is the acceleration vector and L is another parameter vector. We will defer this kind of approach to a later publication. A more simple model is obtained by imposing no acceleration so that we have,

$$\mathrm{d}(t) = \mathrm{L}(t - n) \tag{20}$$

This corresponds to assume a constant velocity L of objects during the time interval separating the two reference frames. Considering this property at time instants m and p, we get

$$\mathrm{d}_m = \frac{m - n}{p - n}\mathrm{d}_p \tag{21}$$

Instead of searching all candidate pairs belonging to $D_m \times D_p$, only one set (D_m or D_p) can be considered. the candidate for the other reference frame following the expression given by (21). If the cardinality of each set of candidate displacement is assumed to be C^2, the computational complexity remains limited to $O(C^2 \times T_e)$ where T_e is the time required to compute the energy of the interpolation error for a given candidate pair (instead of $O(C^4 \times T_e)$ when (21) does not apply).

3 Implementation aspects and computational complexity

3.1 Fast search methods

Several methods for symplifying the search process have been invetigated. To limit the number of possible configurations the search procedure can be interrupted whenever the DFD signal gets below a certain threshold T. The goal is to limit as much as possible the number of displacement pairs $(\mathrm{d}_m, \mathrm{d}_p)$ to find the minimum of the DFD signal.

The philosophy that is used in what follows is built on the fact that such signal decreases monotically as the displacement moves away from the direction of minimum distortion.

A 3-step search procedure can be derived in a similar fashion as for the one reference frame block-matching technique. In other words, the strategy is the following:

- Compute the DFD minimum corresponding to the $(0,0)$ location and the 8 ones centered around $(0,0)$ and distant by $E(|\mathrm{D}_m|/2), E(|\mathrm{D}_p|/2)$, respectively, where $E(x)$ denotes the smallest integer larger than x. Call $(\mathrm{d}_m^{(0)}, \mathrm{d}_p^{(0)})$, such optimal pair out of the 3^4 possible ones.
- Compute the DFD minimum corresponding to $(\mathrm{d}_m^{(0)}, \mathrm{d}_p^{(0)})$ and the 8 ones centered around it distant by $E(|\mathrm{D}_m|/4), E(|\mathrm{D}_p|/4)$, respectively. Denote the corresponding location by $(\mathrm{d}_m^{(1)}, \mathrm{d}_p^{(1)})$.
- Compute the DFD minimum corresponding to $(\mathrm{d}_m^{(1)}, \mathrm{d}_p^{(1)})$ and the 8 ones centered around it distant by $E(|\mathrm{D}_m|/8), E(|\mathrm{D}_p|/8)$, respectively. The associated pair is chosen as the optimal one.

In case it is decided to terminate the searching process whenever a threshold T is crossed, it is chosen to facilitate the occurence of small displacements by scanning those pairs that are the closest to the center location $(0,0)$. This overall 3-step procedure is outlined in Fig. 3.

As a second general concept to reduce the number of search steps, we use neighboring spatial information to establish a scanning strategy for picking the displacement pairs. Let us call $\mathrm{x}_{(01)}$ and $\mathrm{x}_{(10)}$ the blocks to the left and above the current block located at $\mathrm{x}_{(00)}$, respectively. Let assume that the optimal set of displacement pairs are denoted for these two blocks by $A = (\mathrm{d}_m^{(01)}, \mathrm{d}_p^{(01)})$ and $B = (\mathrm{d}_m^{(10)}, \mathrm{d}_p^{(10)})$, respectively. As it is supposed that in natural images, the motion field is smooth, we shall choose as initial displacement pair either A or B for block $\mathrm{x}_{(00)}$, that we call $C = (\mathrm{d}_m^{(00)}, \mathrm{d}_p^{(00)})$. The choice is made according to which pair will have minimum DFD energy. Together with the threshold strategy, future pairs are scanned so as to choose them close to C. A ping-pong approach is used to switch between the two reference frames. In other words, if d_m is fixed, d_p is changed, and vice-verso. The reference frame that is chosen fixed first is the closest one to frame I_n, as there is more confidence in the neighboring block initial estimate.

For an exhautive search procedure with thresholding, the scanning order uses one of the two following strategies. In the first one, assuming d_m is fixed first, all values of d_p are scanned in a circular fashion around $\mathrm{d}_p^{(00)}$ as depicted in Fig.4. The sense whether counter-clockwise or clockwise is arbitrarily decided. As soon as the DFD energy is less than T, the process is stopped. In case all values of d_p have been scanned, the one leading to the minimum of the DFD energy is made the fixed reference in frame I_p. From there, the scanning occurs in a dual fashion for d_m in frame I_m. This ping-pong procedure is iterated until the DFD energy gets below T or it stops to decrease. In the switching between frame I_m and I_p, the DFD energy is monotically decreasing, which ensures that the search will converge.

The second scanning strategy is obtained by operating the ping-pong from one frame to the other only by comparing a subset of displacements in the possible search range of each frame. In other words, it is decided to only scan, let say 9 d_m when d_p is fixed, then keep constant the one value of d_m that minimize the DFD over the 9 possible configurations and proceed similarly on frame I_p in a ping-pong fashion as for the first strategy.

To speed up the searching strategy, we also extend the concept of conjugate direction search. The approach used here can also be considered with threshold, eventually starting with initial displacements chosen in accordance with neighboring block displacements. Given the pair represented by C, we decide to choose the minimum of the DFD energy on the set of all displacements having same row value than $\mathrm{d}_m^{(00)}$. To further limit the searching strategy, the search for d_p values is performed in a conjugate direction manner on frame I_p. Starting with values on the same row as $d_p^{(00)}$, a minimum of the DFD energy is found let say at location $d_{p,1}^{(00)}$; the search for the next minimum is computed on the same column of $d_{p,1}^{(00)}$. The process is iterated until a local minimum is reached and the corresponding value is assigned to $d_m^{(00)}$. All displacements with same row value are similarly assigned a DFD energy value. The one resulting to the lowest minimum over these displacements, that we call $d_{m,1}^{(00)}$ is selected. The search proceeds then on the same column as $d_{m,1}^{(00)}$ in the same way as per $d_m^{(00)}$. The process stops when the DFD energy stops decreasing. As an alternative, a limited set of displacements, say 3 or 5, could be compared on each direction, rather than $[2 \times |D_m| + 1]^2, [2 \times |D_p| + 1]^2$ respectively for each frame.

4 Weight quantization

Once the n-tuple $(\alpha_m, \alpha_n, d_m, d_p)$ has been estimated with the procedure described previously, it is interesting to see how to minimize the increment of the DFD energy if the pair (α_m, α_p) is quantized to a different pair $(\alpha_m^{(q)}, \alpha_p^{(q)})$. Setting a limit on the distortion to the DFD energy, it will be then possible to choose an appropriate quantization strategy.

It can be shown that introducing a distortion pair $(\Delta\alpha_m, \Delta\alpha_p)$ the Δ distortion on the DFD energy is expressed by

$$\Delta = m_{11}(\Delta\alpha_m)^2 + m_{22}(\Delta\alpha_p)^2 + 2m_{12}(\Delta\alpha_m)(\Delta\alpha_p) \tag{22}$$

where m_{11} and m_{22} are the diagonal elements of matrix M, and m_{12} is the anti-diagonal element of matrix M. It is important to notice that equation (23) holds independently of values taken by frame I_n. This ensures that a decoder can estimate the quantization step sizes for α_m and α_p directly from a reconstruction of frames I_m and I_p, as long as he knows the tolerable distortion Δ_{max}.

For non-degenerate cases ($|M| \neq 0$), it is possible to establish a quantization strategy for (α_m, α_p) that will ensure to keep Δ below a threshold Δ_{max}. This is achieved by solving for $(\Delta\alpha_m, \Delta\alpha_p)$ in equation (22). With these assumptions, the quantization step have to be lower than

$$\Delta\alpha_m = \frac{m_{22}}{m_{11}}\Delta\alpha_p \tag{23}$$

$$\Delta\alpha_p = \sqrt{m_{11}\Delta_{max}/|M|} \tag{24}$$

It was observed that imposing the constraint of equation (13), the DFD energy is much less sensitive to the quantization of (α_m, α_p). In this case, the maximum step size for α_m is defined by

$$\Delta\alpha = \sqrt{\Delta_{max}/(m_{11} + m_{22} - 2m_{12})} \tag{25}$$

It is obvious that the pair of equations (24)-(25) results on imposing larger step sizes to $\Delta\alpha_m$ and $\Delta\alpha_p$ due to the large value of m_{11} or m_{22}.

On the contrary, the term $m_{11} + m_{22} - 2m_{12}$ of equation (26) is usually very small as the non negative elements of matrix M have similar magnitude which results in a low sensitivness to the quantization of α.

5 Simulation results

Simulations have been carried out on the standard CIF(240x360) sequence Table-Tennis employing blocks of 16x16 pels and assuming a maximal displacement of 10 pels/frame. The previous frame $I(x, +d_m, m)$ and the following frame $I(x, +d_p, p)$, have been used as references for the interpolation of the in-between frame. A three-step fast algorithm has been initially employed to estimate the block displacements together with a threshold on the interpolation error in order to interrupt the search procedure as soon as a good interpolation is obtained. The joint distribution of α_m and α_p has been computed to justify the introduction of the constraint on the weights vector as explained in subsection 2.3.1. The distribution exhibits a high peak in correspondence of the weight configuration ($\alpha_m = 0.5$, $\alpha_p = 0.5$) and significant values along the line $\alpha_m + \alpha_p = 1$ while being almost zero elsewhere. In fig. 5 the distribution of $\alpha_m + \alpha_p$ is plotted showing the above described behaviour. The joint distribution of the displacement vectors d_m *and* d_p has also proven the validity of the simmetry constraint on the motion field outlined in subsection 2.3.2. As an example, the Y component of the displacement vectors with reference to the past reference and to the future one has been added yielding the peaked distribution plotted in fig. 6. In fig. 7 the curves of the Peak Signal to Noise Ratio (PSNR) of the interpolated Tennis Table sequence are plotted in case of unconstrained continuous optimization of the weights α_m *and* α_p and with the progressive introduction of the constraints. In fig. 8 the PSNR curve obtained with the three-step search is compared with the one obtained applying an exhaustive search procedure exploiting the estimated displacement of spatially neighboring blocks. The interpolation algorithm employing the continuous optimization of the weights α_m *and* α_p outperforms the algorithm employing the MPEG coarse quantization as shown by the PSNR curves plotted in fig. 9. In this last simulation both the constrains on weights and velocity have been introduced and the continuous weights have been quantized by means of a uniform quantizer whose stepsize is adapted block by block according to equation (26). The gain in the PSNR obtained through the use of the continuous optimization algorithm is of only 0.4 dB considering the frame averaged data of fig. 9. Plotting the PSNR gain versus block temporal activity, as reported in fig. 10, a peak of 2.0 dB can be measured in correspondence of moving blocks thus justifying the increased computational complexity. In the specific simulation carried out on Tennis Table sequence, the blocks getting the maximal benefit, in terms of PSNR gain, are those containing the ball where motion is high and where the MPEG coarse quantization yields a very poor interpolation.

6 Conclusion

The continuous optimization of the interpolation error energy obtained by employing multiple reference frames jointly with the adaptive quantization of the motion field parameters has proven to outperform the ISO MPEG strategy recommended for interactive video at 1.5 Mbps. The improvements, in terms of Signal to Noise Ratio, are significant in correspondence of moving blocks where a finer quantization of the interpolation weights allows a far better approximation. This suggests the introduction of an additional feature in the scheme of a video coder allowing the possibility to switch between a coarse interpolation algorithm (according to MPEG recommendations) for static blocks and a finer interpolation algorithm for moving blocks. As far as the test sequence Table-Tennis is concerned, the bitrate increment due to the finer interpolation of moving blocks is not significant thanks to the adaptive quantization of the interpolation weights.

A Appendix 1: Existence of minima and singular cases

As discussed in section 2 the motion vectors and weights of a block B_i in case of two reference frames, are estimated by solving

$$\min_{\mathrm{u}, \mathrm{d}_m \in D_m, \mathrm{d}_p \in D_p} \sum_{\mathrm{x}' \in N(\mathrm{x})} E^2(\mathrm{x}') \tag{26}$$

where $\mathrm{u}^T = [\alpha_m \alpha_p]$ is the weight vector, d_m and d_p represent the displacements for frames $I(\mathrm{x}, m)$ and $I(\mathrm{x}, p)$ searched within the sets D_m and D_p, respectively, and the interpolation error is given by

$$E(\mathrm{x}) = I(\mathrm{x}, n) - (\alpha_m I(\mathrm{x} + \mathrm{d}_m, m) + \alpha_p I(\mathrm{x} + \mathrm{d}_p, p)) \tag{27}$$

The solution to (27) is obtained by imposing that the partial derivatives of $\sum_{\mathrm{x}' \in N(\mathrm{x})} E^2(\mathrm{x}')$ with respect to α_m and α_p are equal to zero, leading to a two linear equations system

$$\mathrm{M}.\mathrm{u} = \mathrm{v} \tag{28}$$

with v and matrix M being

$$\mathrm{v} = \begin{bmatrix} \sum_{\mathrm{x}' \in N(\mathrm{x})} I(\mathrm{x}' + \mathrm{d}_m, m) I(\mathrm{x}', n) \\ \sum_{\mathrm{x}' \in N(\mathrm{x})} I(\mathrm{x}', n) I(\mathrm{x}' + \mathrm{d}_p, p) \end{bmatrix} \tag{29}$$

$$M = \begin{bmatrix} \sum_{x' \in N(x)} I^2(x' + d_m, m) & \sum_{x' \in N(x)} I(x' + d_m, m) I(x' + d_p, p) \\ \sum_{x' \in N(x)} I(x' + d_m, m) I(x' + d_p, p) & \sum_{x' \in N(x)} I^2(x' + d_p, p) \end{bmatrix} \quad (30)$$

The determinant Δ of the matrix M is proven to be non negative and, consequently, M to be semidefinite positive guaranteeing the minimization of the energy term reported in (28). The mathematical demonstration is given in case of two reference frames and then extended to the general case of N references.

Let us define the determinant

$$\Delta = \sum_{x' \in N(x)} I^2(x' + d_m, m) \times \sum_{x' \in N(x)} I^2(x' + d_p, p) - \left[\sum_{x' \in N(x)} I(x' + d_m, m) I(x' + d_p, p) \right]^2 \quad (31)$$

and express the set N(x) as $\{x_1, x_2, \cdots x_K\}$ where $K = Q \times R$ and equals the total number of pels in a block.

As an aid to comprehension let us introduce the following notations: $I(x_i + d_m, m)$ is called I_{mi}, and $I(x_i + d_p, p)$ is denoted by I_{pi}. The determinant Δ can then be expressed as

$$\Delta = \sum_{i=1}^{K} I_{mi}^2 \times \sum_{i=1}^{K} I_{pi}^2 - \left[\sum_{i=1}^{K} I_{mi} I_{pi} \right]^2 \quad (32)$$

Expanding each term of (33) as

$$\begin{aligned} \Delta &= (I_{m1}^2 + I_{m2}^2 + I_{m3}^2 + \cdots + I_{mK}^2) \times (I_{p1}^2 + I_{p2}^2 + I_{p3}^2 + \cdots + I_{pK}^2) - \\ &\quad (I_{m1} I_{p1} + I_{m2} I_{p2} + I_{m3} I_{p3} + \cdots + I_{mK} I_{pK})^2 \\ &= (I_{m1}^2 I_{p1}^2 + I_{m2}^2 I_{p2}^2 + I_{m3}^2 I_{p3}^2 + \cdots + I_{mK}^2 I_{pK}^2) + (I_{m1}^2 I_{p2}^2) + (I_{m1}^2 I_{p3}^2) + (I_{m2}^2 I_{p3}^2) + \cdots - \\ &\quad 2 (I_{m1} I_{m2} I_{p1} I_{p2} + I_{m1} I_{m3} I_{p1} I_{p3} + I_{m2} I_{m3} I_{p2} I_{p3} + \cdots) \\ &= (I_{m1} I_{p2} - I_{m2} I_{p1})^2 + (I_{m1} I_{p3} - I_{m3} I_{p1})^2 + (I_{m2} I_{p3} - I_{m3} I_{p2})^2 + \cdots \\ &= \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} (I_{mi} I_{pj} - I_{mj} I_{pi})^2 \end{aligned} \quad (33)$$

References

- [1] A.N. Netravali and B.G. Haskell, *Digital Pictures Representation and Compression*, Plenum Press, 1988.
- [2] B.K.P. Horn and B.G. Schunk, "Determining Optical Flow", Artificial Intelligence, Vol. 17 (1-3), pp. 185-203, Aug. 1981.
- [3] C.D. Kuglin and D.C. Hines, "The Phase Correlation Image Alignment Method", *Proc. of the IEEE Conf. on Cybern.*, Sept. 1975, pp. 163-165.
- [4] A. Puri, R. Aravind, B.G. Haskell and R. Leonardi, "Video Coding with Motion-Compensated Interpolation for CD-ROM Applications", *Signal Processing: Image Communication*, Vol. 2, No. 2, August 1990.
- [5] "MPEG Video Simulation Model Three (SM3)", Doc. ISO-IEC/JTC1/SC2/WG8 N/ MPEG90.