Optimal Unified Approach to Warping and Overlapped Block Motion Estimation in Video Coding

Aria Nosratinia and Michael T. Orchard

Department of Electrical Engineering,
Princeton University, Princeton, NJ 08544
aria@ee.princeton.edu orchard@ee.princeton.edu

ABSTRACT

Among the class of block-based motion estimators, warping prediction and overlapped block motion estimation have emerged recently as two of the most effective inter-frame estimation algorithms. Warping estimation is based on linear operations in the motion field, while overlapped block estimators are based on a linear sum of motion-based estimators in the intensity field. In this paper, we propose a unified framework for estimation of pixel intensities based jointly on warping and overlapped blocks. We motivate our estimator through a discussion of ambiguities in an incomplete (sparsely sampled) motion field; and that different object motions call for resolution of motion ambiguities in either the motion or intensity domain. We offer a means of optimizing the joint estimator *simultaneously* in intensity and motion domains, thus guaranteeing improved performance compared to warping and overlapped block estimators, which the joint model contains as special cases. Furthermore, the joint framework provides an excellent vehicle for studying the interactions and relative merits of warping and overlapped block estimators in the presense of various motion scenarios.

Keywords: Motion estimation, video coding, overlapped block, warping.

1 Introduction

Block-based motion estimation methods have played a central role in the development of video coding algorithms. These methods include block-matching, overlapped block motion compensation [1–3], warping [4], and others [5]. [1] Block-matching is employed in many video coding standards such as H.261, MPEG-1 and MPEG-2. The two other major categories, namely warping – also known as *control grid interpolation (CGI)* – and overlapped block motion compensation (OBMC), are more recent developments that have achieved significant improvements over conventional block matching. Warping was first designed to solve perspective correction problems in computer graphics, and was later applied to video coding [4]. Improved versions of it have been successfully applied to a number of different inter-frame estimation problems [6,7]. Overlapped block motion compensation (OBMC) was at first explored with hopes of reducing block artifacts [1], and was later motivated directly as a means of

[1] For the lack of a better terminology, we refer to the aggregate of these algorithms as *block-based methods*. Standard block-matching is thus only a subset of block-based methods.

minimizing error energy [3]. In particular, overlapped block motion compensation has come into prominence through its introduction into the H.263 video coding standard.

Motion compensation models the inter-dependence of consecutive frames through a motion field. Although a densely sampled motion field almost always carries enough information to reconstruct the present frame completely from the past frame, transmission of all motions often would require a higher bitrate than the intensity frame itself. Therefore, block-based methods transmit only a subset of motion vectors, and reconstruct the present frame through this partial information. This partial motion information leaves some uncertainty at the pixels where a motion vector is not given. The structure of each block-based method reflects the level of complexity at which it is willing to address these uncertainties, as well as what it considers to be the dominant mechanism, through which motion uncertainties assert themselves. According to their computational sophistication, and how well their assumptions match the realities of particular video sequences, block-based methods have varying degrees of success, in terms of estimation accuracy, on various video sequences.

This paper proposes a unifying framework for block-based motion estimation methods. While motion vectors and their relationship with the intensity frames are often characterized through the relationships of the underlying continuous motion and intensity fields, we prefer to view sampled intensity and motions mainly as two arrays of data, with the aim of using one data set (motions) to estimate the other (intensity), through operations on both motion and intensity data sets. From this perspective, various block-based methods differ only in their approach to using the available (but incomplete) motion information. Some concentrate their efforts in the motion domain – e.g. warping – while others perform their operations in the intensity domain – e.g. overlapped block estimation. Our proposed method, on the other hand, considers linear operations in both intensity and motion domains. We offer a means of optimizing the parameters of a joint estimator *simultaneously* in intensity and motion domains, thus guaranteeing improved performance compared to other known block-based methods. This optimization is achieved through a descent method on the energy function of a representative set of sequences, and is performed off-line.

In Section 2, we briefly discuss overlapped block and warping motion estimation and how they address motion field uncertainties. Based on resulting insights, we offer a joint framework for block-based motion estimation. The resulting motion estimator contains known block-based motion estimation methods as special cases. Section 3, we develop an iterative algorithm to optimize the parameters of the joint estimator. Section 4 presents experimental results, and Section 5 closes with some concluding discussions.

2 Motion Estimation Structures and Motion Field Uncertainties

Given that, due to practical considerations, one cannot afford to transmit a motion vector for every pixel, estimation of intensities requires a method of assigning motion vectors to pixels where the motion array does not *directly* provide us with information. In other words, one has to resolve the uncertainty resulting from the absence of a densely sampled motion field. Different block-based algorithms deal with this uncertainty in different ways. The structure of each block-based method reflects the approach chosen in characterizing this uncertainty, and gives each block-based method its distinct flavor and properties.

Among block-based methods, block-matching implicitly assumes that each block of pixels in the past frame moves with a uniform translational motion. In other words, it allocates each motion vector to the set of pixels that are closer to it, than to any other motion vector. This resolves the motion ambiguity in a trivial way: although it is reasonable for a predictor within a block to rely heavily on the motion vector which, in a sense, has been optimized for that block, motion vectors assigned to neighboring blocks can also contain useful information. More advanced block based methods resolve motion ambiguities more intelligently, through utilizing the information of neighboring motion vectors.

As an example of how neighboring motion vectors can be helpful, consider the signals shown in Figure 1. This simplified one-dimensional example – which we shall refer to as Example 1 – shows $I_{k-1}(t)$ and $I_k(t)$, representing the intensities of the previous and current frames respectively. This example represents cases where a rotating object is viewed at different angles in consecutive frames. We cannot afford to send many motion vectors, and in this case only the two motion vectors v_1 and v_2 are available to the decoder. However, the decoder can make a very good (in this case errorless) intensity estimate by linearly interpolating nearby motion vectors wherever a transmitted motion vector is not available. (This is a consequence of the fact that rigid body rotation gives rise to a linearly varying motion field.) This process essentially defines warping motion estimation.

Now consider the signals shown in Figure 2. In this example – which we shall refer to as Example 2 – frame $k-1$ contains two objects moving with constant velocities v_1 and v_2; with the left object occluding the right. In the absence of information on t^* (the occlusion boundary), which could be anywhere between t_1 and t_2 (sampling points of block motion field) one cannot make an exact intensity estimate. However, it is possible at any point t to form two estimates of the intensity based on v_1 and v_2. One can then find an optimal weighting of resulting intensities based on the statistics of t^*, which can be extracted from a training sequence. This defines the principles of overlapped block motion estimation.

Returning to the two-dimensional case, we now introduce the needed notation and formalize an expression for our estimator. Each frame of the image sequence is defined on a rectangular grid of N pixels. The lattice is denoted by $\mathcal{S}$, and its members $s \in \mathcal{S}$ are denoted $s = [i, j]^t$, where i and j denote the row and column indices respectively. $I_k(s)$ denotes the intensity at pixel s at frame k of the sequence to be coded, $\hat{I}_k(s)$ is the *estimated* intensity, and $\tilde{I}_k(s)$ is the pixel intensity of the corresponding *decoded* frame. We omit the pixel argument when referring to the whole frame, e.g. I_k. The set $\mathcal{V}(s) = \{v_n(s)\}$ represents motion vectors assigned to blocks neighboring pixel s, with respect to some fixed definition of block neighbors of any pixel. (Note: the definition of block neighbors of a pixel s will, in general, depend on the pixel position within a block.) A motion vector is obviously shared among the sets corresponding to many different pixels. Once again, we often omit the dependence on s for simplicity of expression, and the relationship is implied by context. Finally, $g_k(s)$ represents the intensity gradient of frame k at location s.

Overlapped block motion estimators are characterized by

$$\hat{I}_k(s) = \sum_{v_n \in \mathcal{V}(s)} w_n(s)\, \tilde{I}_{k-1}(s - v_n) \;, \tag{1}$$

where the set of weights $\{w_n(s)\}$ are determined through an optimization that results in the solution of a linear system of equations [3].

Warping motion estimators are defined as

$$\hat{I}_k(s) = \tilde{I}_{k-1}(s - \sum_{v_n \in \mathcal{V}(s)} a_n(s)\, v_n) \;, \tag{2}$$

where now $\{a_n(s)\}$ are the set of estimator parameters. Warping estimators perform a linear operation in the motion domain, returning one motion which is used to estimate intensities, whereas overlapped block estimators use multiple intensity estimates, each with one motion vector only.

We propose the following generalization of the above estimators

$$\hat{I}_k(s) = \sum_m w_m(s)\; \tilde{I}_{k-1}(s - \sum_{v_n \in \mathcal{V}(s)} a_{m,n}(s)\, v_n) \;. \tag{3}$$

In the case where $[a_{m,n}] = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, this estimator reduces to the overlapped block motion estimator. When $[w_m] = [1\ 0\ \ldots\ 0]^t$, it reduces to the warping motion estimator. In its general case, it offers a more flexible way of resolving motion field ambiguities, and thus reduces the estimation error energy.

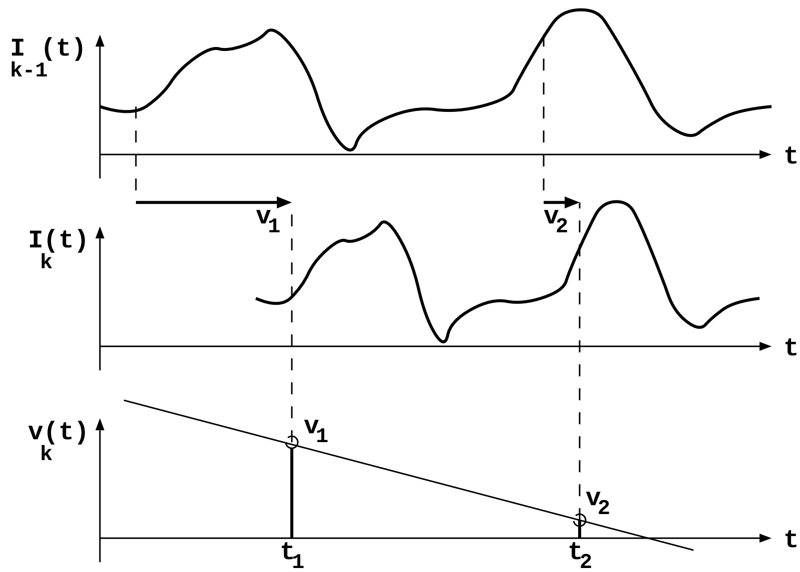


Figure 1: *Example 1: motion uncertainties are best resolved in motion domain.* v_1 *and* v_2 *are "block" motion vectors.*

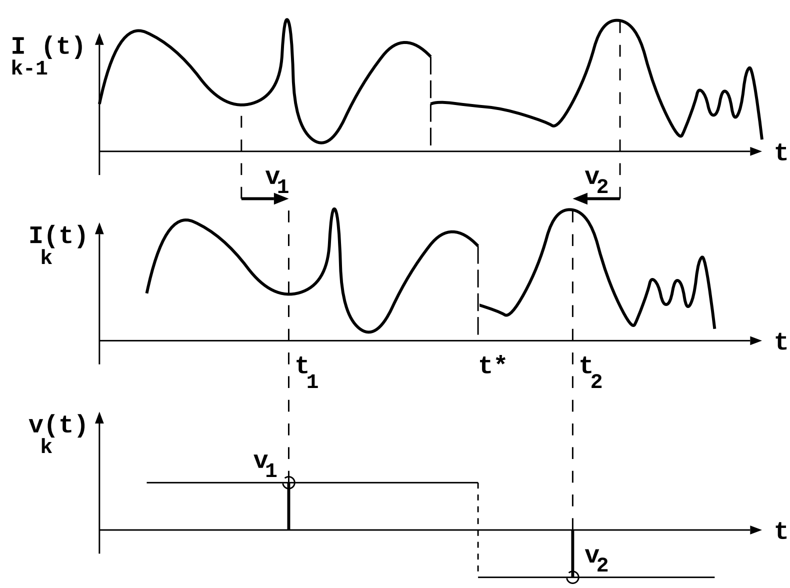


Figure 2: *Example 2: motion uncertainties are best resolved in time (space) domain.* v_1 *and* v_2 *are "block" motion vectors.*

One may ask: is there any reason to believe that introduction of new (or additional) terms into the estimator is to any advantage? Note that in a fair comparison, the new estimator should use the same number of motion vectors as the overlapped block or warping counterpart. In that case, it is perfectly reasonable to ask whether, in the absence of new information, one can hope for better performance.

Assuming that we used M motion vectors in each of the estimators, overlapped block and warping estimators will each have M terms, whereas the joint estimator can have up to M^2 terms. Linear estimation theory asserts that given M sources of information, any linear sum of them cannot make any further contribution to estimation accuracy. However, our estimator is *not linear in terms of the motion vectors.* The gains that we anticipate are afforded through the non-linear relationship of motion vectors and intensities, which is itself a result of the complex shape of the intensity terrain.

Returning to the Examples 1 and 2, one can see that a linear operation in either the intensity or motion domain *alone* cannot perform well in both cases. In Figure 1, each of motion vectors v_1 and v_2 are unsuitable for areas close to $(t_1 + t_2)/2$. Each of v_1 and v_2 would produce a grossly mistaken estimate at such points, and no weighted sum of those estimates can be correct. However, a single estimate based on the correct motion interpolation can result in errorless estimation. Conversely, in the example of Figure 2, we know that at any point on the horizontal axis, one of v_1 or v_2 is correct. Therefore, one would like to choose between resulting estimates $I_{k-1}(t - v_1)$ or $I_{k-1}(t - v_2)$, according to the chances of the point being associated with v_1 or v_2. But in this case, choosing a motion that is interpolated between v_1 and v_2 will return meaningless results. Since each of the motion and intensity approaches fails for the other case, one is motivated for a joint approach.

The astute reader recognizes that these examples have been deliberately constructed so that in each of them, motion field ambiguities are resolved nicely – in fact optimally – by one of the two approaches (motion or intensity domain). Practical cases are often more complicated, and the solution is never so obvious or clear cut. Often times constituting elements of both these examples co-exist, in which case neither one nor the other of the approaches is ideally and uniformly suitable, and a joint approach is warranted. The coefficients of this joint approach – $\{a_{m,n}\}$ and $\{w_m\}$ in (3) – should reflect the frequency and degree of occurrence of the phenomena that call for operations in motion vs. intensity domains. In the following section, we present a method of optimizing these coefficients from intensity and motion data.

3 Optimal Windows for Joint Overlapped Block and Warping

Recall that the proposed estimator has the form

$$\hat{I}_k(s) = \sum_m w_m \, \tilde{I}_{k-1}(s - \sum_{v_n \in \mathcal{V}(s)} a_{m,n} v_n) \; .$$

We present the following result as an optimality condition. These are zero-gradient conditions imposed on the expected value of squared error over the training sequence.

Fact: *Given the distortion objective function*

$$D(s) = E\left[|I_k(s) - \hat{I}_k(s)|^2 \right] \tag{4}$$

a set of necessary conditions for the optimality of the joint warping and overlapped block estimator, given above, is provided by

$$\sum_\ell w_\ell E\left[\tilde{I}_{k-1}(s - v_\ell^*) \; v_m^t \; g_{k-1}(s - v_n^*) \right] - E\left[I_k(s) \; v_m^t \; g_{k-1}(s - v_n^*) \right] = 0 \qquad \forall \, m, n \;\; , \tag{5}$$

$$\sum_{\ell} w_\ell \; E\left[\tilde{I}_{k-1}(s-v_\ell^*)\; \tilde{I}_{k-1}(s-v_n^*)\right] - E\left[I_k(s)\; \tilde{I}_{k-1}(s-v_n^*)\right] = 0 \qquad \forall\, n \;, \tag{6}$$

where

$$v_\ell^* \;\triangleq \sum_{v_i \in \mathcal{V}(s)} a_{\ell,i} v_i \;\; . \tag{7}$$

While we present this result as an optimality *condition*, note that it is nothing but a zero-gradient condition on a squared error cost function, where the cost is written as a function of estimator parameters. Therefore, using the gradients as indicated above, the estimator parameters can be optimized through an iterative descent algorithm. We now proceed to derive equations (5) and (6).

The optimality conditions are derived by setting partial derivatives of the distortion cost function to zero.

$$\begin{aligned} \frac{\partial D}{\partial a_{m,n}} &= 2E\left[\left(\hat{I}_k(s) - I_k(s)\right) \frac{\partial \hat{I}_k(s)}{\partial a_{m,n}}\right] \;, \\ &= 2E\left[\left(\hat{I}_k(s) - I_k(s)\right)\; \frac{\partial}{\partial a_{m,n}} \sum_{\ell} w_\ell \; \tilde{I}_{k-1}(s-v_\ell^*)\right] = 0 \;\; . \end{aligned} \tag{8}$$

But we have

$$\begin{aligned} \frac{\partial \tilde{I}_{k-1}(s-v_\ell^*(s))}{\partial a_{m,n}} &= \frac{\partial \tilde{I}_{k-1}(s-v_\ell^*(s))}{\partial v_x^*}\; \frac{\partial v_x^*}{\partial a_{m,n}} + \frac{\partial \tilde{I}_{k-1}(s-v_\ell^*(s))}{\partial v_y^*}\; \frac{\partial v_y^*}{\partial a_{m,n}} \;, \\ &= -g_{k-1}^t(s-v_\ell^*(s))\; \frac{\partial v_\ell^*(s)}{\partial a_{m,n}} \;, \end{aligned} \tag{9}$$

where we have denoted $v_\ell^* = [v_x^* \;\; v_y^*]^t$. Direct substitution gives

$$\frac{\partial D}{\partial a_{m,n}} = 2E\left[\left(I_k(s) - \hat{I}_k(s)\right)\; \sum_{\ell} w_\ell \; g_{k-1}^t(s-v_\ell^*)\; \frac{\partial v_\ell^*}{\partial a_{m,n}}\right] \tag{10}$$

Observing that

$$\frac{\partial v_\ell^*}{\partial a_{m,n}} = \begin{cases} 0 & \text{if } \ell \neq m \\ v_n & \text{if } \ell = m \end{cases} \;, \tag{11}$$

we have

$$\frac{\partial D}{\partial a_{m,n}} = 2\, w_m \; E\left[\left(I_k(s) - \hat{I}_k(s)\right)\; g_{k-1}^t(s-v_m^*)\;\; v_n\right] = 0 \;\; , \tag{12}$$

Assuming that w_m is non-zero and substituting from the definition of $\hat{I}_k(s)$,

$$\sum_{\ell} w_\ell \; E\left[\tilde{I}_{k-1}(s-v_\ell^*)\; v_m^t \; g_{k-1}(s-v_n^*)\right] = E\left[I_k(s)\; v_m^t \; g_{k-1}(s-v_n^*)\right] \tag{13}$$

This establishes the first set of equations. For the second set of conditions:

$$\frac{\partial D}{\partial w_m} = 2E\left[\left(\hat{I}_k(s) - I_k(s)\right)\; \frac{\partial}{\partial w_m} \sum_{\ell} w_\ell \; \tilde{I}_{k-1}(s-v_\ell^*)\right] \;\; ,$$

$$
\begin{aligned}
&= 2E\left[\left(\sum_{\ell} w_\ell\, \tilde{I}_{k-1}(s-v_\ell^*) - I_k(s)\right)\tilde{I}_{k-1}(s-v_n^*)\right] \quad, \\
&= 2\sum_{\ell} w_\ell\; E\left[\tilde{I}_{k-1}(s-v_\ell^*)\,\tilde{I}_{k-1}(s-v_n^*)\right] - 2E\left[I_k(s)\,\tilde{I}_{k-1}(s-v_n^*)\right] \quad, \\
&= 0\ .
\end{aligned} \tag{14}
$$

Two sets of equations (5) and (6) reflect optimization with respect to $\{a_{m,n}\}$ and $\{w_m\}$. Unfortunately, the equations are coupled and cannot be solved sequentially. The gradient condition (5) is a nonlinear set of equations and requires an iterative solution. But given $\{v_\ell^*\}$, the gradient condition (6) constitutes a set of linear equations, which can be solved directly through L-U factorization or other known and efficient techniques. The final solution is thus achieved through a two phase process: in phase one, we start with fixed $\{a_{m,n}\}$ (hence $\{v_n^*\}$), and solve the system of equations for $\{w_m\}$. Phase two fixes $\{w_m\}$ and solves (5) iteratively for $\{a_{m,n}\}$. The final solution is obtained by starting with an initial guess for $\{a_{m,n}\}$ and iterating between these two phases. In our experiments, equilibrium was achieved in all cases after two or three iterations between phases.

We note that when $[a_{m,n}] = \mathbf{I}$, the first set of optimality equations is trivially true, and the second set reduces to the optimality condition of the overlapped block estimator, as expected.

4 Experimental Results

In this work, we consider estimators that operate on four motion vectors (i.e. $m, n, \ell, i \in \{0, 1, 2, 3\}$). Each pixel s chooses the four motion vectors that are closest to it, resulting in a neighborhood structure shown by the dotted line in Figure 3. Please note that all following figures are illustrated on the array of pixels within the dotted rectangle.

The most general joint estimator, as characterized in (3), is over-parameterized (under-determined). Aside from the issue of wasted computation, we found that it offers very little in improved performance or insight compared to the reduced order model below, which we chose for our experiments.

$$
\hat{I}_k(s) = \sum_{m=0}^{3} w_m\; \tilde{I}_{k-1}(s - v_i)\; +\; w_4\, \tilde{I}_{k-1}(s - \sum_{n=0}^{3} a_n v_n)\ . \tag{15}
$$

We also believe that this model is more appropriate for our investigations in that it contains warping and overlapped block estimators simply as additive terms. It reduces to overlapped block with $w_4 = 0$ and to warping with $[w_0 ... w_3] = [0 ... 0]$. This model can potentially offer insights into the interactions of its warping and overlapped block constituent parts, mainly through coefficient w_4. The magnitude of w_4 in the optimum estimator provides some measure of the relative importance of warping and overlapped block in various sequences.

We applied this model to the first ten frames of two standard sequences "football" (352×240) and "claire" (352×288). The average mean squared error (MSE) results, using 16×16 motion blocks, are shown in Table 1. Figures 4, 5, and 6 demonstrate the optimum coefficients for "football". Figures 7, 8, and 9 do the same for "claire". The coefficients were taken to be quadrantally symmetric, hence the other, unshown coefficients can be obtained by 90^o rotations of the given plots, i.e.

$$
\begin{aligned}
a_0(i,j) = a_1(N-i-1, j) = a_2(i, N-j-1) = a_3(N-i-1, N-j-1) \\
w_0(i,j) = w_1(N-i-1, j) = w_2(i, N-j-1) = w_3(N-i-1, N-j-1)
\end{aligned} \tag{16}
$$

where $N = 16$ in our case. To clarify this symmetry, Figure 4 shows all four coefficients $\{a_0, \ldots, a_3\}$. For economy of presentation, we show only a_0 and w_0 in the rest of the experiments.

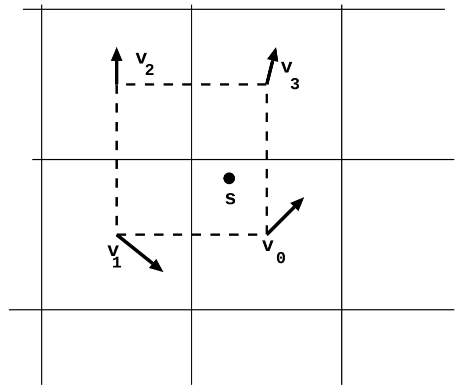


Figure 3: *Block motion grid and motion vectors used in predicting pixel s. Solid squares specify motion blocks, and the dotted rectangle delineates pixels that use the four motion vectors v_0, v_1, v_2 and v_3 for estimation.*

Observe the difference between the profile of "cross-coefficient" w_4 in the two cases. In "football", this coefficient has average 0.29, which means that optimally warping (CGI) provides roughly %30 of the estimate and overlapped block (OBMC) the other %70. In "claire", however, w_4 has an average of 0.80. This shows a large variation in the relationship of warping and overlapped block estimators from one source to another.

In "football", going from overlapped block to the joint approach over the first 10 frames yields a gain of less than 2% (0.07dB) in MSE. In "claire", the same comparison shows a 7.4% (0.3 dB) improvement in MSE. The situation is reversed when comparing CGI with the joint estimator (of course the jointly optimal estimator is always superior; we are only comparing the gains).

5 Conclusion

This paper presented a unified framework for block-based motion estimation and compensation. In a sense the joint framework is a generalization of warping and overlapped block motion estimators. The joint estimator is motivated by motion ambiguities that are inherent in any block-based motion estimation algorithm. Some situations call for resolving these ambiguities in the intensity domain, while others necessitate doing so in the motion domain. The joint estimator recognizes this diversity of motion scenarios through allowing linear operations in both motion and intensity domains. Since there exists a method of optimizing – according to the statistics of representative sequences – the coefficients of the joint estimator, one can tailor the joint estimator to match different motion mechanisms, without the need of explicitly enumerating and accounting for all such mechanisms. The joint estimator provides a framework in which one can investigate the cross-dependencies and relative merits of warping and overlapped block motion estimators, in various motion situations.

Average MSE	block matching	OBMC	CGI	joint approach
"football" 0-10	195.71	139.30	166.15	136.87
"claire" 0-10	4.39	2.97	3.39	2.75

Table 1: *Experimental results with overlapped block (OBMC), warping (CGI) and joint estimators*

We would like to mention briefly that the computations in the joint framework need not necessarily be heavy. For example, the motion interpolation part can be accomplished through quantizing the interpolator parameters to sums of powers of two, much like the coefficients of the overlapped blocks in H.263. Since interpolated motions are no longer necessarily integer valued, interpolation becomes necessary in the intensity domain; but computations can be substantially reduced through half-pixel interpolation of the past frame, and quantizing motions to half-pixel accuracy. This does not have a substantial effect on the quality of estimates, and can be performed in parallel with other coder/decoder operations. Another of major areas of concern in developing video coding hardware is memory bandwidth. The joint framework needs hardly any more memory access cycles than the overlapped block algorithm. Also, the savings in memory transfer offered in [8] apply equally well to the joint estimator.

Current research includes a search for heuristics to either replace optimization, or ease its computational burden; aiming at an adaptive joint estimator.

6 REFERENCES

[1] H. Watanabe and S. Singhal. Windowed motion compensation. In *Proc. SPIE Visual Comm. Image Processing*, volume 1509, pages 582–589, 1991.

[2] C. Auyeung, J. Kosmach, M. T. Orchard, and T. Kalafatis. Overlapped block motion compensation. In *Proc. SPIE Visual Comm. Image Processing*, volume 1605, pages 561–571, November 1992.

[3] M. T. Orchard and G. Sullivan. Overlapped block motion compensation: an estimation-theoretic approach. *IEEE Trans. Image Processing*, 3(5):693–699, September 1994.

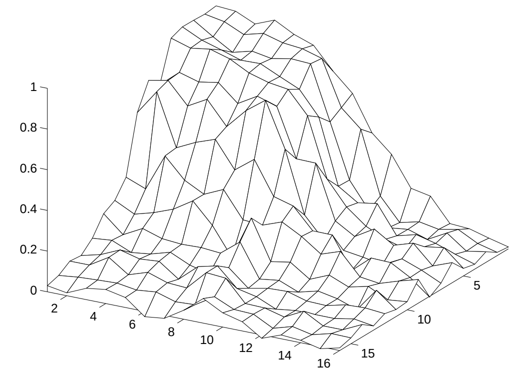
[4] G. J. Sullivan and R. L. Baker. Motion compensation for video compression using control grid interpolation. In *Proc. IEEE ICASSP*, volume 4, pages 2713–2716, May 1991.

[5] Jiro Katto and Mutsumi Ohta. An analytical framework for overlapped block motion compensation. In *Proc. IEEE ICIP*, Austin,TX, November 1994.

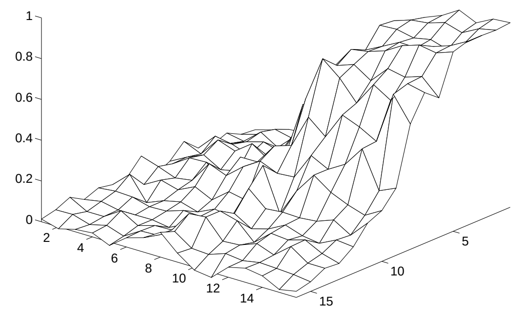
[6] Y. Nakaya and H. Harashima. An iterative motion estimation method using triangular patches for motion compensation. In *Proc. SPIE Visual Comm. Image Processing*, volume 1605, pages 546–557, November 1991.

[7] A. Nosratinia, N. Mohsenian, M. T. Orchard, and B. Liu. Interframe coding of magnetic resonance images. *IEEE Trans. Medical Imaging*, to appear in 1996.

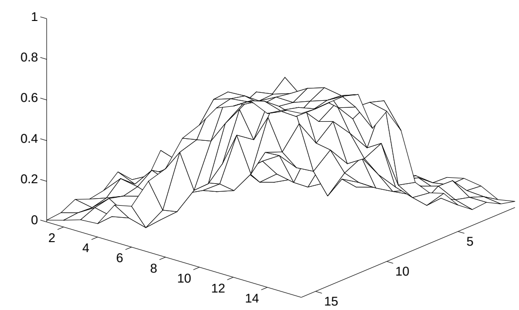
[8] Gary J. Sullivan and Michael T. Orchard. Methods of reduced-complexity overlapped block motion compensation. In *Proc. IEEE ICIP*, volume II, pages 957–961, Austin, TX, November 94.



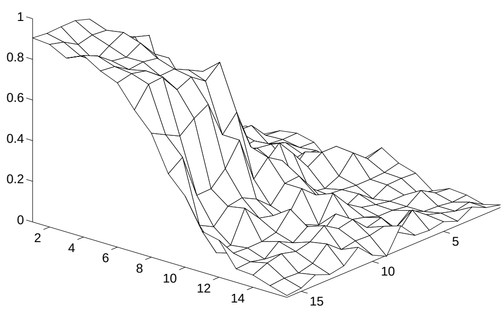
coefficient a_0



coefficient a_1



coefficient a_2



coefficient a_3

Figure 4: *Warping coefficients on frames 0-10 of "football"*

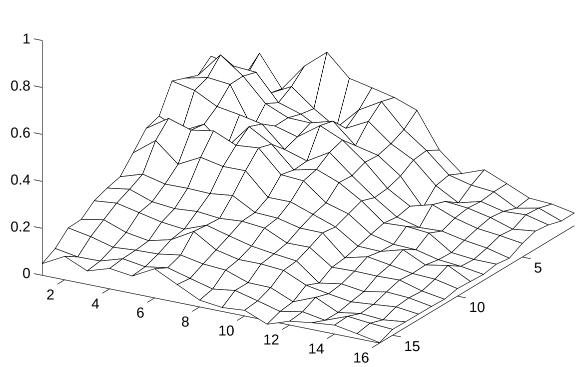


Figure 5: *Overlapped block coefficient* w_0 *on frames 0-10 of "football"*

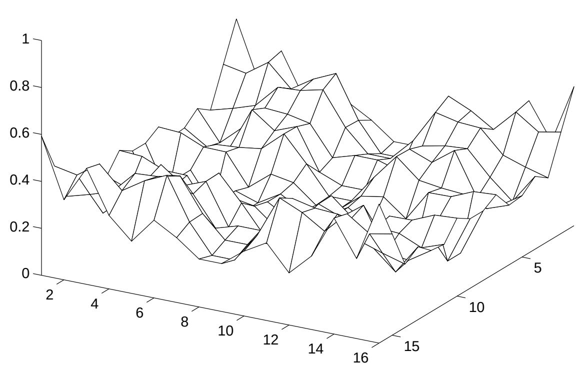


Figure 6: *Cross-coefficient* w_4 *on frames 0-10 of "football"*

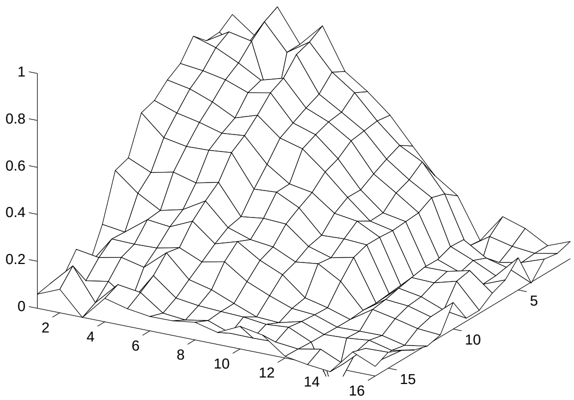


Figure 7: *Warping coefficient* a_0 *on frames 0-10 of "claire"*

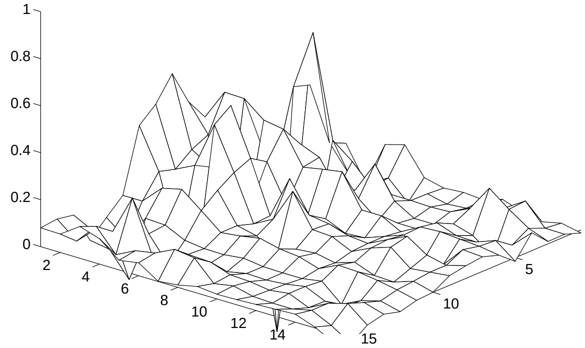


Figure 8: *Overlapped block coefficient* w_0 *on frames 0-10 of "claire"*

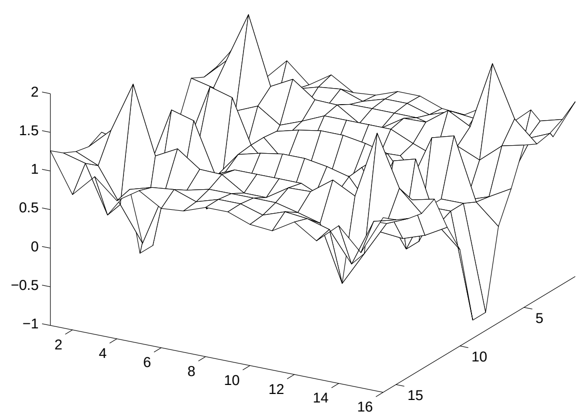


Figure 9: *Cross-coefficient* w_4 *on frames 0-10 of "claire"*