Variational Motion Compensated Deinterlacing

Sune Keller, François Lauze and Mads Nielsen

The IT University of Copenhagen
The Image Group
Rued Langgaardsvej 7, 2300 Copenhagen S, Denmark
{sunebio,francois,malte}@itu.dk
www.itu.dk/image

Abstract. In this paper we present a novel motion compensated deinterlacing scheme based on a joint variational framework of motion estimation and creation of new pixel values. We propose a specific energy formulation for that purpose, optimized by solving corresponding Euler-Lagrange equations. The difficult issue of calculating motion on interlaced image sequences is addressed and the presented algorithm proves to solve what we call *The Interlacing Problem*: depicting highly detailed regions in motion without artifacts to the human visual system (HVS).

1 Introduction and Background

Since the birth of television in the 1930's interlacing have been the dominant scanning format and even today it is used in HDTV in Japan, USA and Europe. In interlaced scan only half the horizontal lines of each frame are recorded and displayed as illustrated in Fig. 1 alternating between the even and odd numbered lines and the recorded partial frames are called even and odd fields.

Most modern displays are progressive, i.e. display full frames and thus conversion from interlaced to progressive scan, *deinterlacing*, is needed to create every other line of the sequence (white lines in Fig. 1).

Deinterlacing can be done very simple using methods like line averaging with reasonably acceptable output quality, but interlacing artifacts are generally not removed. Two types of interlacing artifacts, serration and line crawl, are shown in Fig. 1, while a third type, line flicker appears when a detail is only seen in one horizontal line and thus only displayed at half the intended screen refresh rate. For further details on interlacing artifacts see [1] and [6]. The output of motion adaptive deinterlacers like the ones described in [6] and state of the art chips sets such as the DCDi® by Faroudja has almost no interlacing artifacts left.

Motion adaptive deinterlacers are however unable to solve what we like to call *The Interlacing Problem*: the problem of correctly handling vertical high frequency image content in motion, illustrated in Fig. 2.

A way to handle the problem may be to smooth the region in order to avoid any interlacing artifacts, but it will also blur out visually important details. The spatially missing information is often present across several frames and if the motion is known, it could be propagated along the motion trajectories over time

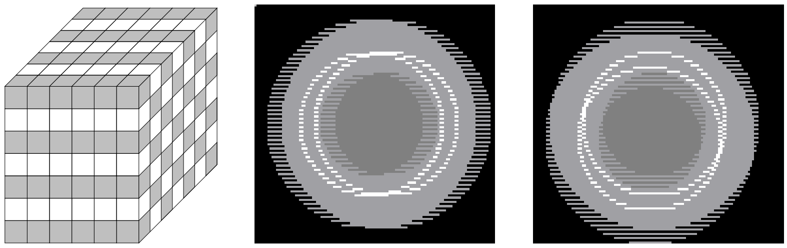

Fig. 1. Left: The interlacing (width, height, time)-volume. Interlacing artifacts shown by a merge of a time t field with its neighboring time $t+1$ field: In the middle serration due to horizontal motion and to the right it is line crawl due to vertical motion

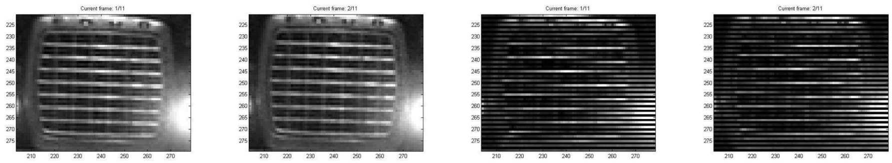

Fig. 2. The Interlacing Problem. Still 1&2: Zoom in on two consecutive progressive frames of the sequence `Truck` of a truck driving towards the camera. Still 3&4: the same two frames interlaced: knowledge of the correct spatial structure is almost gone

to new pixel positions, and this is the principle of motion compensated deinterlacing (MCDI). Fig. 2 clearly illustrates how difficult it would be to to recreate the correct structure using only spatial information. The video corresponding to these frames shows that the perforated structure of interlaced image volumes in combination with the highly detailed content, makes motion computation a non-trivial task.

Motion compensated deinterlacing (MCDI) use temporally coherent information from a *motion field / optical flow* estimate (we will use the two terms interchangeably) as input to an interpolation of the missing lines with maximized output quality. It must also be robust to inaccuarete motion estimation and switch to pure spatial interpolation, for instance. This is not a simple task. For instance, Biswas and Nguyen propose in [2] a motion compensated algorithm which produce results that may still contain serration artifacts. Bellers and de Haan has done extensive work on developing and testing motion compensated deinterlacing and a selection of 11 algorithms are presented and tested in [1]. They show that the presented motion compensated deinterlacers are better than simple deinterlacers, although motion adaptive methods are left out of the test.

A major problem with motion estimation in interlaced sequences is that standard motion estimation algorithms are developed for progressive sequences and thus not directly applicable.Typically a simple deinterlacer is used to create an initial progressive sequence [1]. This influences the accuracy of the resulting obtained flow due to the interlacing artifacts introduced to the sequence. The problem of obtaining an accurate flow will be most severe in the regions troubled by The Interlacing Problem and where an accurate flow is needed the most.

Deinterlacing can be seen as a form of *Image Sequence Inpainting* and in this paper we propose a new motion compensated scheme based on the inpainting variational formulation of Lauze and Nielsen [8] which we briefly discuss in Section 2. Due to the very nature of interlaced sequences, a straightforward application of this algorithm is problematic. In Section 3 we propose instead a three steps algorithm for motion estimation and deinterlacing. We present our test results in Section 4 and conclude in Section 5.

2 Motion Compensated Inpainting

In [8], Lauze and Nielsen proposed a generic Bayesian formulation for image sequence inpainting and motion recovery. The observed damaged image sequence is denoted u_0 and we define it to be on a spatiotemporal domain denoted Ω with D being the locus of the missing data in Ω, that is $D \subset \Omega$. $\boldsymbol{v}$ is the motion field of the de-noised and restored image sequence u, the desired result of the inpainting process. They show that under mild assumptions, the conditional probability $p(u, \boldsymbol{v}|u_0, D)$ can be factorized as

$$p(u, \boldsymbol{v}|u_0, D) \propto \underbrace{p(u_0|u, D)}_{P_0}\ \underbrace{p(u_s)}_{P_1}\ \underbrace{p(u_t|u_s, \boldsymbol{v})}_{P_2}\ \underbrace{p(\boldsymbol{v})}_{P_3}. \tag{1}$$

where u_s and u_t are the spatial and temporal distribution of intensities respectively. P_0 is the image sequence likelihood, P_1 the spatial prior on image sequences, P_3 the prior on motion fields and P_2 acts as both a likelihood for the motion field and a temporal prior on image sequences. The Maximum a posteriori (MAP) is then sought for in order to reconstruct the image sequence and recover the motion field. Using the Bayesian to Variational rationale of Mumford ([9]), the problem is transformed into this continuous minimization problem

$$E(u, \boldsymbol{v}) = E_0(u, u_0) + E_1(u_s) + E_2(u_s, u_t, \boldsymbol{v}) + E_3(\boldsymbol{v}). \tag{2}$$

Using Calculus of Variations, a minimizer for the above energy is characterized by the coupled system of equations, $\partial E(u, \boldsymbol{v})/\partial u = 0$ and $\partial E(u, \boldsymbol{v})/\partial \boldsymbol{v} = 0$.

Selecting the distributions to use for each of the E_i-terms in (2) is a tradeoff between tractability and modelling accuracy. Discussions on the matter can be found in [7] and in [8], the above energy (2) is instantiated as

$$E(u, \boldsymbol{v}) = \lambda_0 \int_{\Omega \setminus D} (u - u_0)^2 dx + \lambda_1 \int_{\Omega} \psi(|\nabla u|^2) dx +$$
$$\lambda_2 \int_{\Omega} \psi\Big(\left(\frac{\partial u}{\partial \boldsymbol{V}}\right)^2 + \gamma \left|\frac{\partial \nabla u}{\partial \boldsymbol{V}}\right|^2\Big) dx + \lambda_3 \int_{\Omega} \big(\psi(|\nabla v_1|^2) + \psi(|\nabla v_2|^2)\big) dx \tag{3}$$

where ∇ is the spatial gradient operator, λ_i and γ are some constants and v_1 and v_2 are the x- and y-components of the flow field, i.e. $\boldsymbol{v} = (v_1, v_2)^T$ and

$\boldsymbol{V} = (\boldsymbol{v}^T, 1)^T$. The two $\partial \cdot / \partial \boldsymbol{V}$ terms are the $\boldsymbol{V}$-directional derivatives of u and ∇u respectively[1]. Instead of using the $|\cdot|$ function which is non differentiable at the origin, we replace it by the approximation $\psi(s^2) = \sqrt{s^2 + \varepsilon^2}$, where ε is a small positive constant. The spatial gradient term is thus a variation of the term introduced by Rudin et al. in [10], while the flow part is a slight rewriting of the energy proposed by Brox et al. in [3]. The resulting system of Euler-Lagrange equations are solved within a multiresolution framework, interleaving motion recovery and motion compensated equations, starting with a basic inpainting at coarsest scale, assuming null motion. For further details see [7].

By taking as missing data locus D the set of missing lines in an interlaced sequence, this would provide a deinterlacing algorithm. However, each 'hole' in the sequence is only one line high and thus does not scale down and a multiresolution inpainting cannot be used. In the next section we propose an algorithm for deinterlacing based on inpainting methodology presented in this section.

3 Deinterlacing Algorithm

As mentioned in section 1, some care must be taken in motion calculations on the interlaced input. We do it by computing two separate flows respectively for even and odd fields, then remoulding the flow to cover the full progressive output domain and finally computation of the intensities – the actual deinterlacing.

3.1 Motion Estimation Steps

The motion estimation part that we propose is a two steps process. In order to be as close as possible to the known data, we proceed the following way. We extract the odd field sequence, i.e. the sequence of the known lines of odd fields, and 'compress' it vertically by removing the unknown lines. In the same way, we produce the even field sequence.

1. Compute the flow $\boldsymbol{v}_o$ for the odd field sequence and the flow $\boldsymbol{v}_e$ for the even field sequence.
2. "Fusion" $\boldsymbol{v}_o$ and $\boldsymbol{v}_e$ in order to produce a candidate flow field for the full progressive sequence.

The flow recovery algorithm that we have used is the already mentioned one of [3]. Generally variational methods produce the best and most accurate flow field of any methods available today, as the survey in [4] shows.

We have two forward flow fields, one for the odd original fields and one for the even original fields, both containing flows from fields at time t to time $t+2$. To get flows for all frames in the output we have to *fuse* these two flows together: we start by a simple interleaving (and find initial values in the new lines by line averaging), we *scale* the resulting flow from the compressed fields used in the motion estimation and halve the motion vectors to be from time t to time $t+1$

[1] $\partial f(x)/\partial \boldsymbol{V} = \lim_{\epsilon \to 0} \left(f(x + \epsilon \boldsymbol{V}) - f(x) \right) / \epsilon = \nabla f(x) \cdot \boldsymbol{V}$

and not to time $t+2$ (assuming linear flow from t to $t+2$). Finally we *smooth* the new flow field to get it consistent, correct errors due to the assumptions made in the two first steps and improve the estimates in the new lines. This all ends up being a *fuse-scale-smoothing* process as expressed by this energy minimization

$$\boldsymbol{v} = \operatorname*{Arg\,min}_{\boldsymbol{v}} \int_{\Omega \setminus D} (\boldsymbol{v} - \boldsymbol{v}_{o/e})^2 dx + \lambda \int_{\Omega} \psi(|\nabla \boldsymbol{v}|^2) dx \tag{4}$$

where composing $\boldsymbol{v}_{o/e}$ is the fusion-scaling part and the full minimization is the smoothing. As before $\psi(s^2) = \sqrt{s^2 + \varepsilon^2}$. After deriving an Euler-Lagrange equation of (4), discretization is done using standard methods; The first term is treated similar to the E_0 in section 3.2 and the second term is the same as the smoothing term of the flow algorithm (see [3]) and thus treated similarly.

3.2 Motion Compensated Deinterlacing Step

The actual deinterlacing is done by minimizing the following energy

$$E(u) = \lambda_0 \int_{\Omega \setminus D} (u - u_0)^2 dx + \lambda_1 \int_{\Omega} \psi_1(|\nabla u|^2) dx + \lambda_2 \int_{\Omega} \psi_2(|\nabla u \cdot \boldsymbol{v} + u_t|^2) dx \tag{5}$$

λ_0 looks superfluous, but serves an important purpose; set to zero it makes the scheme an implementation of the energy without de-noising. most deinterlacers do not de-noise and leave all the original intensities of u_0 untouched. λ_0 allows to test our method in this setting as well. As argued in [1] de-noising can help remove some temporal flickering noise. We agree on this, but stress the need to carefully control the de-noising to avoid over-smoothing. The Euler-Lagrange equation corresponding to the energy formulation in (5) is

$$\begin{aligned} \frac{\partial E}{\partial u} = \lambda_0 \chi(u - u_0) - \lambda_1 \mathrm{div}_2 \left(\frac{\nabla u}{\psi_1(|\nabla u|^2)} \right) & \\ -\lambda_2 \mathrm{div}_3 \left(\frac{\psi_2'(|\nabla u \cdot \boldsymbol{v} + u_t|^2)}{|\nabla u \cdot \boldsymbol{v} + u_t|} (\nabla u \cdot \boldsymbol{v} + u_t) \boldsymbol{V} \right) & = 0 \end{aligned} \tag{6}$$

where λ_0 is two times λ_0 in (5) and χ is the characteristic function, which takes on the value 1 in $\Omega \setminus D$ and 0 elsewhere. The discretization of the gradient descent equation corresponding to (6) is detailed in [7] and references therein.

4 Experiments

It is often discussed whether to use objective or subjective evaluation of the performance of deinterlacers, e.g. in [1] and [6]. We have chosen subjective as the human visual system will be the final judge and also it seems to be the choice in industry, e.g. at Bang & Olufsen. We have used three experts and two laymen to judge the quality in an A/B comparisons with Faroudja's DCDi®

Table 1. Optimal settings for our three step motion compensated deinterlacing scheme

	fixed point iterations	inner iterations	convergence threshold	weights
Motion estimation (3)	5	20	10^{-7}	$\gamma = 100$, $\lambda_3 = 70$
Fuse-scale-smoothing (4)	10	100	10^{-6}	$\lambda = 1$
Deinterlacing (6)	5	50	10^{-7}	$\lambda_0 = 1$, $\lambda_1 = 0.1$ $\lambda_2 = 1, \mathbf{5}, 10$

(combined motion and spatial edge adaptive deinterlacing) in the DVP-1010 Video Processor, which is considered state of the art in the world of home cinema.

In our algorithm there are 17 free parameters and after initial testing we chose a set of standard settings and tested at least two alternative values to each parameter besides the values tested initially. This of course far from covers all possibilities, but the amount of output data created in such a test and the time needed to evaluate it would require resources beyond our reach. The optimal settings found can be seen in Table 1. Two of the 17 parameters are not in the table: For the motion estimation step we use 60-80 levels with a scale factor of 1.04 in the multiresolution pyramid on PAL resolution material (720×576). For the weight λ_2 in the deinterlacing step three choices are given. Our tests showed $\lambda_2 = 5$ to be the best choice.

Generally our three step algorithm has proven quite robust to changes in parameter settings. The flow calculations can give very detailed (or controlled poorly, over-segmented) flows, but they are some degree smoothed out in the fuse-scale-smoothing and thus working on improved motion estimation on interlaced material for deinterlacing will most likely prove worthwhile.

On seven different test sequences with The Interlacing Problem – details in motion – our algorithm clearly outperformed the Faroudja DCDi® according to all five test persons. Each of the three test sequences having a progressive original was judged almost indistinguishable from their original. On an eighth sequence with very complex motion (hand held camera in a close-up of fingers typing on a keyboard) the motion was not recovered correctly and some serration were observed by one person in the test panel. In Fig. 3 some examples of the high quality outputs are shown. Due to copy protection we are unable to produce stills or record video of the output from the Faroudja processor. On the `Copenhagen Pan` Faroudja has some flicker on the buildings and also produces quite a lot of serration on the roof windows of the castle, where our deinterlacer produces a perfect result. On `BBC3` our deinterlacer recovers the details and the motion is smooth, while Faroudja has staggering motion and the details are hard to see. On `Truck`, which was used to illustrate The Interlacing Problem in Fig. 2, we are able to (re)create the correct structure of the grill while Faroudja fails at this and creates some dark diagonal lines instead. The test sequence `Credits` (not shown) has rolling titles, a type of content on which most deinterlacers including the Faroudja in test perform horribly, but the output of our algorithm

Fig. 3. Three examples of the high quality output of our Motion Compensated Deinterlacer. Top: `Copenhagen Pan` (the motion is a right-to-left pan), notice the details on the house in the bottom middle and the roof window of the castle. Left: `Truck`, compare to original in Fig. 2. Right: In this still from `BBC3` (rotational motion on an axis in the lower left corner) a few artifacts are seen in the still but not during playback

is indistinguishable from the progressive original. The general performance of the Faroudja is typical for motion adaptive deinterlacers: It is on the same level as the variational motion adaptive deinterlacer in [6], which we compared it to.

In the test of motion compensated deinterlacers in [1] there is to few subjective results given in [1] to really compare to our scheme. Given the relatively small improvement in objectively evaluated performance compared to simple line averaging in [1], we believe to outperform all the schemes in [1] based on the subjective performance of our deinterlacer presented above.

5 Conclusion and Future Work

As the previous section revealed we have a very good deinterlacer, but it is also very computationally heavy. We have shown that a technique so far used for inpainting (see [8]) can be redeveloped to do deinterlacing and further on our Total Variation Motion Compensated Deinterlacing outperforms Faroudja's state of the art deinterlacer. We produce high quality outputs when the input flow is correct and precise, and generally solves The Interlacing Problem, although improvements can be done.

In terms of running time, we are far from realtime, but to get closer to the desired realtime version we plan to replace the SOR solver with multigrid solvers shown to run realtime for variational flow computations in [5].

All though the multiresolution scheme used in [8] cannot be used for deinterlacing the idea to simultaneous update flow and intensities could be used to improve the flow obtained and the final output quality of our deinterlacer. Other options are further tweaking of the parameters of the existing algorithm, the use of other models than total variation.

References

1. E.B. Bellers and G. de Haan. *De-interlacing. A Key Technology for Scan Rate Conversion.* Elsevier Sciences Publishers, Amsterdam, 2000.
2. M. Biswas and T.Q Nguyen. A novel de-interlacing technique based on phase plane correlation motion estimation. *International Symposium on Circuits and Systems, ISCAS*, 2:604–607, 2003.
3. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *Proceedings of the 8th European Conference on Computer Vision*, volume 4, pages 25–36, Prague, Czech Republic, 2004. Springer–Verlag.
4. A. Bruhn, J.Weickert, and C.Schnörr. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005.
5. A. Bruhn, J. Weickert, C. Feddern, T. Kohlberger, and C. Schnörr. Variational optic flow computation in real-time. *IEEE Trans. Image Proc.*, 14(5):608–615, 2005.
6. S. Keller, F. Lauze, and M. Nielsen. A total variation motion adaptive deinterlacing scheme. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Procedings of the 5th Scale-Space Conf.*, volume 3459 of *LNCS*, pages 408–419, Berlin, 2005. Springer.
7. F. Lauze. *Computational Methods For Motion Recovery, Motion Compensated Inpainting and Applications.* PhD thesis, IT University of Copenhagen, 2004.
8. F. Lauze and M. Nielsen. A variational algorithm for motion compensated inpainting. In S. Barman A. Hoppe and T. Ellis, editors, *British Machine Vision Conference*, volume 2, pages 777–787. BMVA, 2004.
9. D. Mumford. Bayesian rationale for the variational formulation. In *Geometry-Driven Diffusion In Computer Vision*, pages 135–146. Kluwer Academic Publishers, 1994. Computational Imaging And Vision.
10. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.