Joint Framework for Motion Validity and Estimation Using Block Overlap

Michael Santoro, *Member, IEEE*, Ghassan AlRegib, *Senior Member, IEEE*, and Yucel Altunbasak, *Fellow, IEEE*

***Abstract*— This paper presents a block-overlap-based validity metric for use as a measure of motion vector (MV) validity and to improve the quality of the motion field. In contrast to other validity metrics in the literature, the proposed metric is not sensitive to image features and does not require the use of neighboring MVs or manual thresholds. Using a hybrid de-interlacer, it is shown that the proposed metric outperforms other block-based validity metrics in the literature. To help regularize the ill-posed nature of motion estimation, the proposed validity metric is also used as a regularizer in an energy minimization framework to determine the optimal MV. Experimental results show that the proposed energy minimization framework outperforms several existing motion estimation methods in the literature in terms of MV and interpolation quality. For interpolation quality, our algorithm outperforms all other block-based methods as well as several complex optical flow methods. In addition, it is one of the fastest implementations at the time of this writing.**

***Index Terms*— Block matching, de-interlacing, motion vector (MV), reliability, true motion estimation, validity.**

I. Introduction

MOTION estimation allows for a wide variety of tasks such as object tracking, frame-rate conversion, super-resolution, and depth estimation in 3-D TV. In such applications, the true motion is desired; however, finding the true motion is an ill-posed problem since the 3-D scene must be projected onto a 2-D image plane. In addition to the problem of estimating the motion of pixels whose true motion is in three dimensions; untextured regions, small structures, occlusions, deformations, and others types of complex motion further contribute to the ill-posedness of motion estimation. Although state-of-the-art algorithms have made remarkable progress in handling many of these types of motion, there are still several cases in which algorithms fail. For motion estimation to be useful in practical applications, it is necessary to detect failures and assign each motion vector (MV) a confidence value. In this paper, we introduce a joint framework which characterizes the validity of MVs and uses the validity to improve the quality of the motion field.

The majority of motion estimation algorithms in the literature use either block matching or optical flow to determine a dense motion field. Although optical-flow-based algorithms provide superior MV quality over block-based algorithms [1], they do so at the expense of high computational complexity and long run times. In the interest of real-time applications, block matching algorithms provide a flexible trade-off between complexity and MV quality. Therefore, we present our motion estimation framework in the context of block matching and compare its performance to state-of-the-art optical flow methods.

Block matching requires the use of the translational-motion model and brightness-constancy assumption to estimate the motion of blocks between image pairs. Unfortunately, these two requirements are often violated for real images; the actual motion can only be approximated as a translation for small displacements, and the brightness-constancy assumption does not hold for illumination changes due to non-uniform lighting, shadows, etc [2]. Block matching is also sensitive to block size. Large blocks are needed to avoid local minima; however, large blocks produce poor matches compared to small blocks. To a large degree, the block size problem is minimized by using a hierarchical block matching framework [3]. In a hierarchical framework, large blocks are used in downsampled image pairs to provide an initial estimate of the motion, and the block size is successively reduced as the resolution of the images increases.

To determine the best matching block, correlation-based approaches are generally favored due to their robustness and low complexity [4]. Several correlation metrics such as the mean-squared error (MSE), mean absolute difference (MAD), and sum of absolute deviations (SAD) have been introduced in the literature. However, most block matching methods use the SAD correlation metric, which computes the ℓ_1 norm of pixel differences for all pixels in a given block. The SAD correlation metric is a non-convex function; as the image resolution increases and the block size decreases in a hierarchical framework, the number of local minima increases.

To regularize the results of block matching, i.e., to choose among the local minima, low-complexity smoothness constraints have been introduced [5], [6]. These constraints limit the solution based on the assumption that the motion field should be locally constant. It is desired that the smoothness constraints penalize deviations among MVs while preserving natural discontinuities in the motion field. The effects of

Manuscript received May 17, 2012; revised November 27, 2012; accepted December 2, 2012. Date of publication December 20, 2012; date of current version February 12, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Zhou Wang.

M. Santoro and G. AlRegib are with the Center for Signal and Image Processing (CSIP), School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: msantoro@gatech.edu; alregib@gatech.edu).

Y. Altunbasak is with Tubitak, Ankara 06540, Turkey (e-mail: yucel.altunbasak@tubitak.gov.tr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2012.2235452

different low-complexity smoothness constraints on the quality of the motion field were also evaluated in [5], [6].

Although regularizing the motion estimation problem using smoothness constraints reduces the number of possible solutions, it does not force a unique solution. In fact, unless the regularization function is strictly convex, multiple solutions may exist. In previous approaches [5]–[7] the chosen solution depends on the block matching search order and on the order in which the smoothness constraints are applied. To help overcome this, we introduce a new block overlap term as both a regularizer and validity metric to improve the quality of the motion field.

In the next sections, we introduce the proposed validity metric which will be used in subsequent sections for improving the quality of the motion field. In section II, we provide an overview of existing validity metrics and introduce the proposed validity metric. We also compare our validity metric to existing metrics in the context of hybrid de-interlacing. In section III, we present the energy minimization framework and discuss the shortcomings of previous work. In section IV, we combine the block overlap regularizer into the existing framework and give our block-overlap-based algorithm. Experimental results comparing our algorithm with the state-of-the-art are given in section V, and the conclusions and future works are presented in section VI.

II. Validity Metrics

In this section, we provide an overview of existing validity metrics that have been applied to block-based motion estimation algorithms and introduce the proposed block-overlap-based validity metric. We divide the existing validity metrics into two categories: smoothness-based validity and gradient/variance-based validity. The content in this section was first introduced in [8].

A. Smoothness-Based Validity

1) Smoothness Metric 1: Wang *et al.* introduced smoothness-based validity in the context of de-interlacing [9]. A confidence value was assigned to each MV based on block correlation and MV smoothness. Specifically, the validity expression was given as

$$R(\mathbf{v}) \approx \frac{Corr(\mathbf{x}, \mathbf{x}+\mathbf{v}) + Smooth(\mathbf{v})}{\sum_{k \in \mathbf{V}^4} \left[Corr(\mathbf{x}, \mathbf{x}+\mathbf{v}_k) + Smooth(\mathbf{v}_k) \right]}, \quad (1)$$

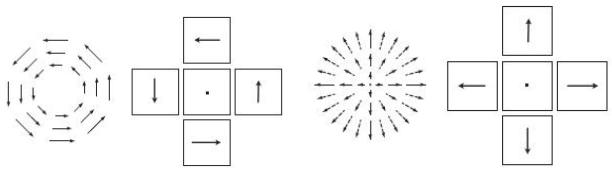
where $\mathbf{v}$ is the current MV and $\mathbf{x}$ is the position in a block B of pixels. In (1), the correlation and smoothness values are calculated for all of the MVs in a four-connected neighborhood, $\mathbf{V}^4$. The $Corr$ and $Smooth$ terms are given as

$$Corr(\mathbf{x}, \mathbf{x}+\mathbf{v}) \approx \sum_{\mathbf{x} \in B} [I_1(\mathbf{x}) - I_0(\mathbf{x}+\mathbf{v})]^2 \quad (2)$$

$$Smooth(\mathbf{v}) \approx \min \left\{ \|\mathbf{v} - \mathbf{v}_k\|^2 \,\middle|\, 0 < k \leq 4 \right\}, \quad (3)$$

where I_1 and I_0 are the current and previous images, respectively.

The $Smooth$ term in (3) assumes that a true MV must be similar to at least one of its neighboring MVs in order to be valid. However, as shown in Fig. 1, this assumption fails in the event of a rotation or scaling, where the neighboring MVs are significantly different from the center MV.



(a) (b)

Fig. 1. Center and neighboring MVs for rotation and scaling. (a) MVs for rotation. (b) MVs for scaling.

2) Smoothness Metric 2: A similar smoothness-based validity metric was introduced by Liu and Shen in the context of super resolution [10]. Liu and Shen made the assumption that a MV is valid if it results in a small normalized SAD and has consistent neighboring MVs. However, in contrast to the work of Wang *et al.*, thresholds were applied to both the normalized SAD and smoothness in order to classify a MV as valid or invalid. The decision rule was given as

$$\begin{aligned} &If \quad (Corr(\mathbf{x}, \mathbf{x}+\mathbf{v}) < T_E) \mapsto R(\mathbf{v}) = 1 \\ &Else \\ &\qquad If \ (Smooth(\mathbf{v}) < T_C) \mapsto R(\mathbf{v}) = 1 \\ &\qquad Else \mapsto R(\mathbf{v}) = 0 \\ &End \\ &End, \end{aligned} \quad (4)$$

where T_E and T_C were determined experimentally. The first If statement indicates that a MV is valid if the normalized SAD is small, and the second If statement indicates that a MV is valid if it is similar to neighboring MVs. More details can be found in [10]. The $Corr$ and $Smooth$ terms were given as follows:

$$Corr(\mathbf{x}, \mathbf{x}+\mathbf{v}) = \frac{\sum_{\mathbf{x} \in B} \left| I_1(\mathbf{x}) - I_0(\mathbf{x}+\mathbf{v}) \right|}{\epsilon + \sum_{\mathbf{x} \in B} I_1(\mathbf{x})} \quad (5)$$

$$Smooth(\mathbf{v}) = \sum_{k \in \mathbf{V}^8} \mathbf{a}(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_k), \quad (6)$$

where $\mathbf{a}(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_k)$ is an indicator function for the dot product of unit vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{v}}_k$, and ϵ is a small scalar to prevent division by zero. The neighborhood size was increased from $\mathbf{V}^4$ to $\mathbf{V}^8$, which includes horizontal, vertical, and diagonal neighbors. The indicator function was given as

$$\mathbf{a}(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_k) = \begin{Bmatrix} 1 & \hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_k > T_S \\ 0 & \text{otherwise} \end{Bmatrix}, \quad (7)$$

and the threshold T_S was determined experimentally.

Aside from the difficulty in setting thresholds, this metric also assumes that a MV is valid if the $Corr$ term is below a given threshold, which fails for uniform regions. In addition, the $Smooth$ term in (6) is not a useful measure of validity for motion such as rotation and scaling, which was shown in Fig. 1.

Fig. 2. Multiple matches with a large gradient and small DFD. (a) Matches in adjacent image. (b) Desired block in current image.

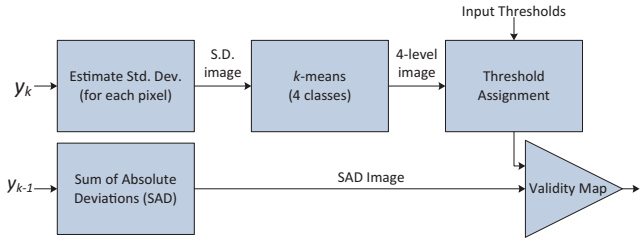


Fig. 3. Generation of validity map for the metric of [13].

B. Gradient/Variance-Based Validity

The two metrics in this section rely on block texture in addition to a correlation metric (SAD or square displaced frame difference (DFD)). While the SAD or square DFD rely on the intensity differences among pixels in a block, the texture of a block provides a characterization of more intuitive qualities such as coarseness, contrast, directionality, line-likeness, and roughness [11]. In the context of validity, properly discerned texture provides an additional discriminator for determining if image blocks are a good match.

1) Gradient Metric: François *et al.* calculated a confidence value for a given direction which depends on the spatial gradient in that direction and on the DFD [12]. The spatial gradient, which is commonly approximated using a Sobel operator, is a computationally inexpensive measure of texture. In the work of [12], the validity was given separately for the $x-$ and $y-$directions. The validity for the x-direction was given as

$$R_x(\mathbf{v}) = \frac{1}{1 + \dfrac{1 + Corr(\mathbf{x}, \mathbf{x}+\mathbf{v})}{2\,Text_x(\mathbf{x})}}. \tag{8}$$

In (8), the $Corr$ term is the same as that given in (2), and the $Text$ term was given as follows:

$$Text_x(\mathbf{x}) = \sum_{\mathbf{x}\in B} G_x^2(\mathbf{x}), \tag{9}$$

where $G_x(\mathbf{x})$ is the gradient in the x-direction for each position $\mathbf{x}$ within block B. Similarly, $R_y(\mathbf{v})$ can be found by replacing x with y in (8) and (9). However, since a single confidence value for each MV is desired, we make a slight modification to the metric of François *et al.* by bounding the validity such that

$$R(\mathbf{v}_i) = \min(R_x(\mathbf{v}), R_y(\mathbf{v})). \tag{10}$$

Therefore, the validity of any given MV will be bounded by the least valid component of the MV.

The expression given in (8) implies that a large confidence value should be assigned to a MV that has a large gradient and small DFD. However, this is not true around edges, where the DFD may be small and the gradient may be large. As an example, consider the two images with a vertical edge as shown in Fig 2. In Fig. 2, the DFD will be small for all of the multiple block matches in Fig. 2(a), and the block in Fig. 2(b) contains an edge (and hence large gradient). Therefore, this metric will assign a high confidence value to a block with an edge, even if the block has multiple matches in the adjacent image. This violates the validity assumption, i.e., a block with multiple matches should be marked as invalid.

2) Variance Metric: Patti *et al.* introduced a variance-based validity metric in the context of de-interlacing [13]. An adaptive threshold-based metric was used that assumes regions of low local variance should have a low SAD threshold value, whereas regions of high local variance should have a high SAD threshold value.

The entire process is shown in Fig. 3. Starting from the image pair $\{I_1, I_0\}$, the standard deviation is estimated for each pixel in image I_1, which produces the standard deviation (SD) image. To remove high SD values, the SD image is quantized into four classes using the k-means algorithm [14]. Next, a set of predetermined image thresholds are assigned to each of the four classes. The validity map is determined by comparing the SAD value to the standard deviation thresholded value. If the SAD value is below its corresponding deviation threshold, the MV is labeled as valid. Otherwise, the MV is labeled invalid.

The main difficulty with the process shown in Fig. 3 is the determination of the 'Input Thresholds'. These thresholds were determined experimentally in [13], and were shown to be sensitive to image content. In the results of [13], it was demonstrated that this metric fails around stationary edges in the image.

C. Proposed Validity Metric

A block-overlap-based validity metric is proposed to overcome the weaknesses of the metrics in sections II-A and II-B, where it was shown that the validity was sensitive to neighboring MVs, image features, or thresholds.

The motivation behind using block overlap as a validity metric can be understood if we consider that between two sets of images whose motion is to be determined, the MVs should map each pixel in the reference image to a unique position in the adjacent image. In images where there is occlusion, scaling, or other types of complex motion, there will not be a unique pixel mapping between images. These types of motion will produce areas of block overlap in the motion-compensated image, and we wish to penalize such overlaps (and hence MVs). It should be noted that the block overlap metric will be constrained by the minimum block size since every pixel within a given block is assumed to have the same MV.

To develop our block overlap approach, we begin by considering a pair of images whose motion we wish to estimate.

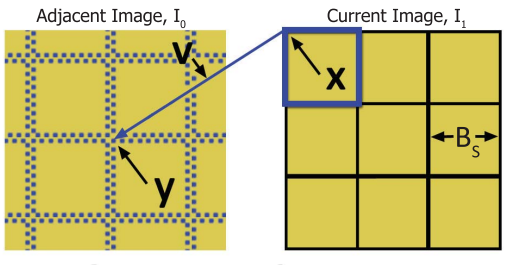


Fig. 4. Blocks of current image I_1 and MC blocks in adjacent image I_0.

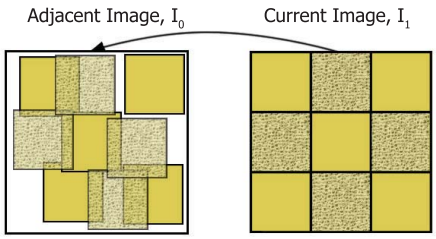


Fig. 5. Block mappings for real images.

Let the current image I_1 be divided into a grid of square blocks of size B_S, and let the grid of the adjacent image I_0 be determined by the MC blocks, as shown in Fig. 4. The position of a block in I_1 is denoted as $\mathbf{x}$ and the position of the MC block in I_0 as $\mathbf{y} = \mathbf{x} + \mathbf{v}$, where $\mathbf{v}$ is the MV. To simplify the analysis, we assume a square grid; however, the analysis also holds for non-square grids.

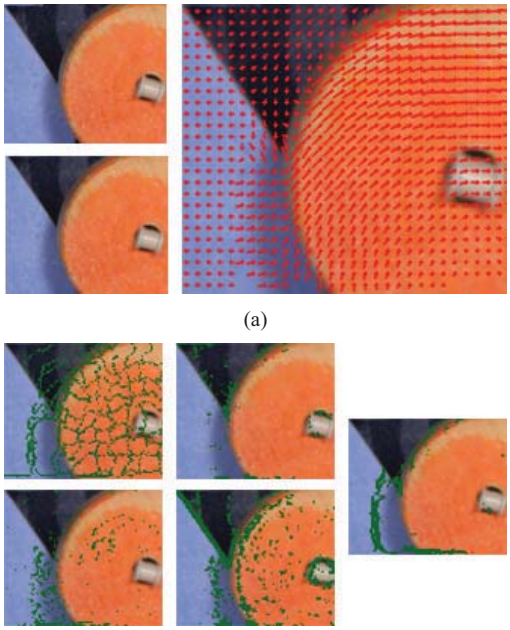
In the ideal case, there exists an injective function $f : I_1 \mapsto I_0$ such that each block in I_1 is mapped to a unique block in I_0. However, for real images with various motion types, the process of mapping blocks from I_1 to I_0 is non-surjective, i.e., blocks in I_1 may be mapped to same block position in I_0, and the whole of I_0 is not necessarily filled. An example of the block mapping for real images is given in Fig. 5. As shown in Fig. 5, the mapped MC blocks for real images may overlap each other to different degrees. We wish to characterize the degree of overlap as the uncertainty in the motion estimation decision, i.e., the validity of a given MV will depend on the amount to which its MC block overlaps with other MC blocks.

We now derive the proposed validity metric. An MC block at position $\mathbf{y}$ will cover an area of B_S^2 in I_0 if no overlap occurs. However, when blocks in I_0 overlap, a volume (perhaps nonuniform) is generated, which we denote as $\mathbf{L_x}(\mathbf{y})$, the overlap at position $\mathbf{y}$ mapped from position $\mathbf{x}$. The degree of overlap for the block at position $\mathbf{y}$ is given by the ratio $B_S^2/\mathbf{L_x}(\mathbf{y})$, where a large ratio indicates less/no overlap and therefore a higher confidence in MV $\mathbf{v}$. Note that the minimum value of $\mathbf{L_x}(\mathbf{y})$ is B_S^2 since a MC block will cover an area of B_S^2 even if there is no overlap with other MC blocks.

However, in some cases the block at position $\mathbf{y}$ does not generate a large degree of overlap in I_0 even though the block has a large SAD value. Therefore, it is necessary to weight the degree of overlap by the SAD value. The validity metric

Algorithm 1 Calculate Volume for Each $\mathbf{L_x}(\mathbf{y})$

For all blocks in I_1, set $\mathbf{y} = \mathbf{x} + \mathbf{v}$
for $k = y_i,\ k < y_i + B_S,\ k$++ **do**
  for $l = y_j,\ l < y_j + B_S,\ l$++ **do**
    $\mathbf{L_x}\begin{bmatrix} k \\ l \end{bmatrix} \mathrel{+}= 1$
  end for
end for



(a)

(b)

Fig. 6. Image Pair, MVs, and invalid pixels for army sequence. (a) Current/adjacent images and corresponding MVs. (b) Pixels with R(v) < 0.5 for different validity metrics.

for MV $\mathbf{v}$ is given as

$$R(\mathbf{v}) = \frac{B_S^2}{\left(1 + \dfrac{\mathrm{SAD}(\mathbf{x}, \mathbf{y})}{\mu}\right)\mathbf{L_x}(\mathbf{y})}, \tag{11}$$

where SAD($\mathbf{x}$, $\mathbf{y}$) is the SAD between blocks $\mathbf{x}$ and $\mathbf{y}$, and μ is the mean of the SAD value over all MVs. The mean μ is included to make the validity metric more robust to brightness variations between images. Equation (11) states that a MV (and hence MC block) which does not produce any overlap in I_0 and has a small SAD value will result in a large confidence value. The algorithm used to determine $\mathbf{L_x}(\mathbf{y})$ is given in Algorithm 1, where y_i and y_j represent the vertical and horizontal positions of the block at position $\mathbf{y}$, respectively.

We demonstrate the strength of the proposed validity metric using an image pair from [1] (top left/bottom left) and the corresponding MV field (right), as shown in Fig. 6(a). The images in Fig. 6(b) show the MVs in I_1 with low confidence values ($R(\mathbf{v}) < 0.5$) for the different validity metrics. The values of $R(\mathbf{v})$ were normalized to fall between 0 and 1, where $R(\mathbf{v}) = 1$ indicates the highest confidence in the validity of a MV. For the smoothness-based metric of Liu, we used the thresholds given in [10], i.e., $T_E = 0.05$, $T_C = 5$, and

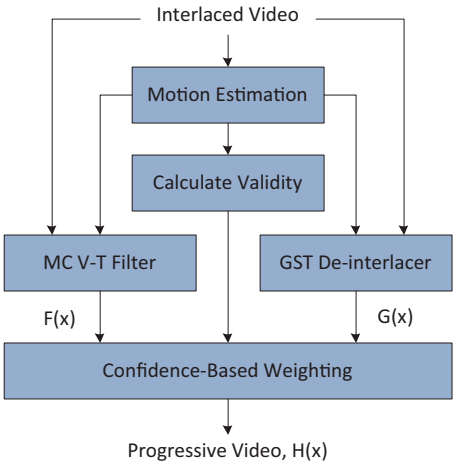


Fig. 7. Hybrid de-interlacing algorithm of [9].

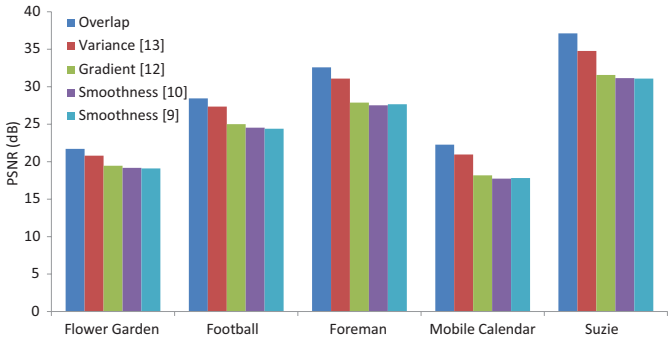


Fig. 8. Average PSNRs of hybrid results for different validity metrics. To visualize the results more easily. The top to bottom ordering in the legend matches the left to right ordering of the bars for each sequence.

$T_S = 0.9$. The four k-means thresholds of (20, 14, 12, and 4) given in [13] were used for the variance-based method.

The smoothness-based metric of Wang (top left image of Fig. 6(b)) penalizes the minimum MV deviation, which can be seen to mark valid MVs on the rotating disk as invalid. The smoothness-based metric of Liu (top right) shows an improvement for the MVs on the rotating disk, but it does not detect the occlusion for the blue surface to the left of the disk, where there is a covering of the background. The gradient-based (bottom left) and variance-based (bottom right) metrics also do a poor job of classifying MVs as valid or invalid. The proposed block overlap metric (far right), however, correctly detects occlusion and invalid MVs surrounding the disk.

D. Validity Metric Comparison Using Hybrid De-Interlacing

To quantitatively compare our validity metric with the metrics introduced in the previous sections, we implemented the hybrid de-interlacing algorithm of [9]. The algorithm discussed in [9] chooses between two types of de-interlacers based on the validity of MVs. The first de-interlacer is based on the generalized sampling theorem (GST) [15], [16], which is theoretically the optimal method for generating a de-interlaced image. However, this method is very sensitive to MV errors and often produces artifacts in the de-interlaced image. The second de-interlacer is based on the vertical-temporal filter with motion compensation (MC V-T filter) [16], which is very robust to MV errors. Although this method is less prone to generating artifacts, it produces lower resolution de-interlaced images than those of the GST de-interlacer.

The choice between the two de-interlacing methods is made based on the MV confidence value, which is illustrated in Fig. 7. The confidence value is used to weight pixels of both image $G(\mathbf{x})$ and $F(\mathbf{x})$ in order to produce the final progressive image, $H(\mathbf{x})$.

Using the algorithm shown in Fig. 7, the proposed validity metric was quantitatively compared with the four other validity metrics. The MVs were estimated to quarter-pixel accuracy using a hierarchical block matching algorithm, and a 3×3 block size was used for all validity metrics in accordance with the details given in the papers. The 3×3 block size was applied as the final block size for the highest-resolution level of the hierarchy.

Interlaced test sequences were created from the following five progressive video sequences: *Flower Garden (CIF, 30 fps), Football (CIF, 30 fps), Foreman (CIF, 30 fps), Mobile Calendar (CIF, 30 fps),* and *Suzie (QCIF, 30 fps)*. To avoid aliasing, local pixel averaging was used to generate the interlaced images from the original progressive images. The PSNR of the de-interlaced image was calculated with respect to the original progressive image for all frames of the five sequences. The PSNR (H($\mathbf{x}$) in Fig. 7) of each sequence for the different validity metrics is shown in Fig. 8. It can be seen that the proposed validity metric outperforms the four other validity metrics for all of the video sequences. The smoothness-based metrics performed similarly, and the gradient-based metric only provided a slight improvement over the smoothness-based metrics. The variance-based metric was a significant improvement; however, when averaged over all of the video sequences, the proposed metric provided an additional 1.4 dB of improvement over the variance-based metric.

The superior performance of the proposed validity metric can be explained by its ability to quantify the MV validity independently of the image content or the behavior of neighboring MVs. The performance of the hybrid de-interlacer also benefits from the more conservative nature of our block overlap validity metric. For example, it may be the case that two blocks are mapped to the same position in the motion-compensated image although only one block is mapped correctly. With our block overlap validity metric, the MVs of both blocks will be penalized, and depending on the amount of overlap and SAD error, the pixels from these blocks will be combined using the more conservative MC V-T de-interlacer.

Although the hybrid de-interlacer introduced in this section was shown to outperform the GST and MC V-T filter in [9], it should be noted that hybrid de-interlacers suffer from high-computational complexity. In addition, switching between different de-interlacers may produce blur in the de-interlaced image. The interested reader is referred to [17] for more background on de-interlacing methods.

III. Energy Minimization for Motion Estimation

As discussed in section I, the ill-posed nature of motion estimation requires the use of regularization to reduce the number of possible solutions. In low-complexity motion estimation algorithms, regularization is performed through explicit smoothness constraints. With the addition of smoothness constraints, the motion estimation problem becomes an energy minimization problem, and it is necessary to choose the optimal MV among several MV candidates. However, as we discuss in this section, the energy minimization functional is not strictly convex, i.e., there may be MVs which produce the same minimum energy. Therefore, we wish to introduce the MV validity metric from the previous section into the energy minimization framework to further constrain the number of solutions.

A. Problem Formulation

The motion estimation problem can be formulated as the following energy minimization problem:

$$E = \min_i \{\mathcal{D}(I_0, I_1, v_i) + \lambda \mathcal{R}(v_i)\}, \tag{12}$$

where $\mathcal{D}(I_0, I_1, v_i)$ is a data term that measures the similarity of blocks in images I_0 and I_1 for a given MV v_i, $\mathcal{R}(v_i)$ is a regularization term which penalizes deviations in the smoothness of the motion field, and the Lagrange multiplier λ is used to weight the regularization term over the data term. The derivation of (12) can be found in [18]. The goal of the motion estimation problem is to choose a MV v_i such that the energy in (12) is minimized.

B. Data Term

The majority of the block matching methods in the literature use the SAD error metric for the data term in (12) because of its robustness and low-complexity [19]. The SAD error is given as follows:

$$\mathcal{D}(I_0, I_1, v_i) = \sum_{\mathbf{x} \in B} |I_0(\mathbf{x}) - I_1(\mathbf{x} + v_i)|, \tag{13}$$

where $\mathbf{x}$ is the pixel position in a block B of pixels.

C. Regularization Term

Robust potential functions such as the ℓ_1 and ℓ_2 norms [20] are often used for the regularization term [5], [6]. In the work of [5], Bartels and de Haan analyzed several different potential functions and candidate sets. A candidate set refers to the spatial and/or temporal MVs of neighboring blocks. Bartels and de Haan found that quality of the motion field was not highly sensitive to the number of MVs in the candidate set. In contrast to the work of [5], [21] and others, we do not assume that temporal MVs are available in previous frames (e.g., the first two frames of a video sequence). Instead, we focus only on the spatial MVs available in the current frame. It should be noted that the quality of future temporal MVs will depend on the quality of the initial MVs determined from the first two frames of the video sequence.

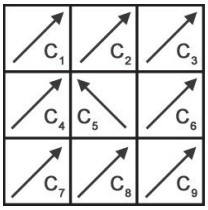


Fig. 9. Candidate set of eight-connected spatial MVs.

In a previous work by the authors, we found that using eight spatial candidates provided the best quality motion field [18]. An example candidate set of eight-connected spatial MVs is shown in Fig. 9, where C_5 is the current MV and the spatial MVs are the eight-connected neighbors. Similar to the work of [5], we chose the optimal potential function as follows:

$$\mathcal{R}(v_i) = \sum_{j \in \mathcal{C}^s} \|v_i - v_j\|_1, \tag{14}$$

where v_i and v_j are spatial MVs in candidate set $\mathcal{C}^s$ and $v_i \neq v_j$.

In order to use eight spatial candidates in $\mathcal{R}(v_i)$, it is first necessary to determine all of the MVs for the current frame by minimizing the SAD only. While this incurs additional overhead, it is only necessary for the first pair of frames in a video sequence since additional frame pairs can take advantage of temporal MVs. The initial MVs are determined by minimizing the data term in (13).

D. Lagrange Multiplier

The Lagrange multiplier λ in (12) should be chosen to weight the regularization term over the data term. Choosing a larger value of λ speeds up the convergence of the MVs; however, a large value of lambda will oversmooth the motion field. To overcome this, we initialized λ to a small value ($\lambda = \frac{3}{4}$ block size) and increased its value in proportion to the iteration number of (12), e.g., for the second and third iterations, the value of λ is multiplied by two and three, respectively. We show the effect of the initial λ value on the quality of the motion field in Fig. 10. For the results on the eight Middlebury [1] image sequences shown in Fig. 10, we iterated (12) until the MVs converged (three iterations). With an initial block size of 32×32, a λ value of 24 ($\frac{3}{4} \times 32$) can be seen to minimize the endpoint error in Fig. 10, where the endpoint error is given as

$$EE = \sqrt{(u - u_{GT})^2 + (v - v_{GT})^2}. \tag{15}$$

In (15), (u, v) is the computed MV and (u_{GT}, v_{GT}) is the ground-truth MV provided from the Middlebury database.

E. Shortcomings of Previous Work

As previously discussed, there will not necessarily be a unique minimum for a given MV since neither of the terms in (12) is strictly convex. Therefore, it is necessary to choose between MVs which produce the same overall energy. Without an explicit way to discriminate between MVs that produce the same overall energy, the chosen MV will depend on the

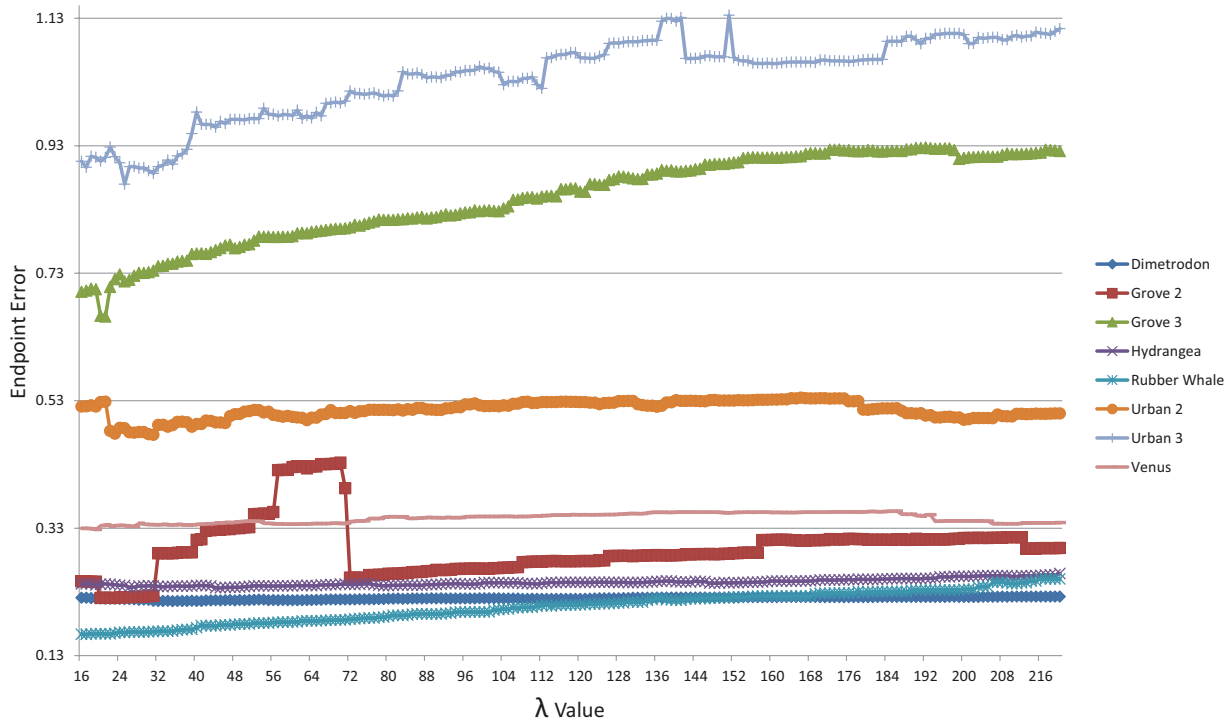


Fig. 10. Endpoint error for different values of λ. Eight sequences shown were taken from the Middlebury [1] database.

order in which the spatial candidates $\mathcal{C}^s$ are tested in (12). Fortunately, as we will show in section IV, a block overlap regularizer helps to choose the optimal MV in such cases.

IV. Block Overlap Minimization Framework

Block overlap introduces an additional error metric for determining which MV minimizes the overall energy. We modify the energy expression of (12) as follows:

$$E = \min_i \{\mathcal{D}(I_0, I_1, v_i) + \lambda \mathcal{R}(v_i) + \mathcal{O}(I_0, I_1, v_i)\}, \quad (16)$$

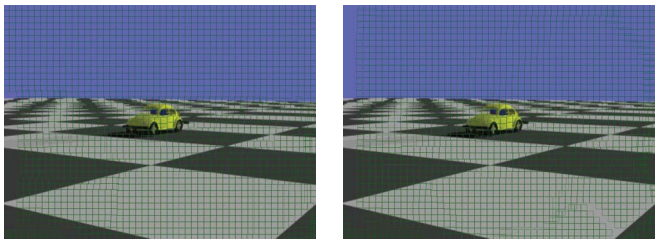
where $\mathcal{O}(I_0, I_1, v_i)$ represents the overlap regularizer for the MC block in I_0 given by v_i. Similar to the overlap-based validity metric in section II-C, the overlap regularizer is given as

$$\mathcal{O}(I_0, I_1, v_i) = (1 + \mathcal{D}(I_0, I_1, v_i)) \frac{\mathbf{L_x(y)}}{B_S^2}, \quad (17)$$

where $\mathbf{L_x(y)}$ is the volume given in Algorithm 1 and B_S is the block size. In (17), the overlap volume $\mathbf{L_x(y)}/B_S^2$ is multiplied by the data term $\mathcal{D}(I_0, I_1, v_i)$, which allows the overlap volume to contribute more to the energy for large SAD values and less for small SAD values. We add the constant '1' to the data term so that the overlap volume still contributes to the energy when the SAD is zero, which improves the performance in uniform regions.

Using the SAD in (13), the smoothness constraints of (14), and the overlap regularizer (17), the energy expression of (16) can be rewritten as

$$E = \min_i \left\{ \sum_{\mathbf{x}\in B} |I_0(\mathbf{x}) - I_1(\mathbf{x}+v_i)| + \lambda \sum_{j\in\mathcal{C}^s} \|v_i - v_j\|_1 + \left(1 + \sum_{\mathbf{x}\in B} |I_0(\mathbf{x}) - I_1(\mathbf{x}+v_i)|\right) \frac{\mathbf{L_x(x+}v_i\mathbf{)}}{B_S^2} \right\}. \quad (18)$$



(a) (b)

Fig. 11. Visual comparison of MC blocks for overlap and non-overlap versions of energy equation. (a) MC blocks using energy of (19). (b) MC blocks using energy of (12).

We re-write (18) in a more compact form by adding the constant '1' to the overall expression inside the brackets of (18), which does not affect the minimization. The new, compact expression is given as

$$E = \min_i \left\{ \left(\sum_{\mathbf{x}\in B} |I_0(\mathbf{x}) - I_1(\mathbf{x}+v_i)| + 1 \right) \times \left(\frac{\mathbf{L_x}(\mathbf{x}+v_i)}{B_S^2} + 1 \right) + \lambda \sum_{j\in\mathcal{C}^s} \|v_i - v_j\|_1 \right\}. \quad (19)$$

In the next two sections, we demonstrate some of the advantages of the new energy minimization framework of (16) compared to the framework of (12).

A. Uniform Block Distribution in Textureless Regions

One of the advantages of the overlap regularizer is that it provides a more uniform distribution of blocks in regions with little texture (smooth regions). In general, smooth regions will have multiple minima, and without the block overlap regularizer, the value of λ should be large to force the MVs to

Algorithm 2 Proposed Motion Estimation Algorithm

1: Form image hierarchy and begin at lowest-resolution level.
2: For all blocks in image I_1, find the corresponding blocks in image I_0 with the lowest SAD error.
3: Using the MVs from Line 2, find the MV with the minimum energy by applying (19).
4: Iterate Line 3 until MVs converge, reduce the block size, and repeat until block size is 1×1.
5: Pass converged MVs to next level of hierarchy and return to Line 2. Repeat until highest-resolution level of hierarchy is reached.

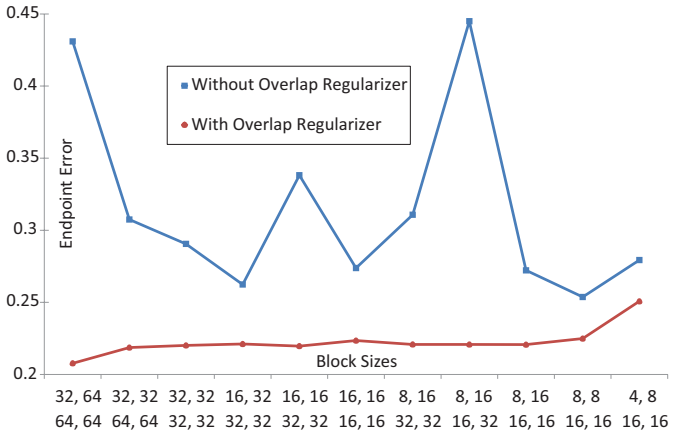


Fig. 12. MV error for different block sizes.

converge. However, without knowing the segmentation of the objects in the image, a large value of λ will also oversmooth the MVs in other regions.

If any two MVs have the same combined SAD and smoothness, the overlap regularizer in (19) will choose the MV (and hence MC block) which results in the least amount of overlap. As shown in Fig. 11, the new energy framework of (19) provides a more uniform distribution of blocks for the smooth regions than the energy framework of (12).

B. Reduced Sensitivity to Block Size

Another advantage of the overlap regularizer is reduced sensitivity to different block sizes. The reduced sensitivity is closely related to the distribution of blocks; the overlap regularizer attempts to keep the distribution of blocks uniform regardless of block size. To demonstrate the reduced sensitivity, we chose the "Grove 2" image sequence from the Middlebury database [1] and varied the block size for the energy minimization frameworks of (19) and (12). The change in MV error for different block sizes is shown in Fig. 12. In Fig. 12, the block sizes are given on the x-axis for a four-level hierarchy, where the top left block size is for the highest-resolution level of the hierarchy, and the bottom right block size is for the lowest-resolution level. The endpoint error was given in (15).

C. Progression of Algorithm

The progression of the algorithm used to minimize the energy of (19) is given in Algorithm 2. The block diagram for the steps in Algorithm 2 is given in Fig. 13.

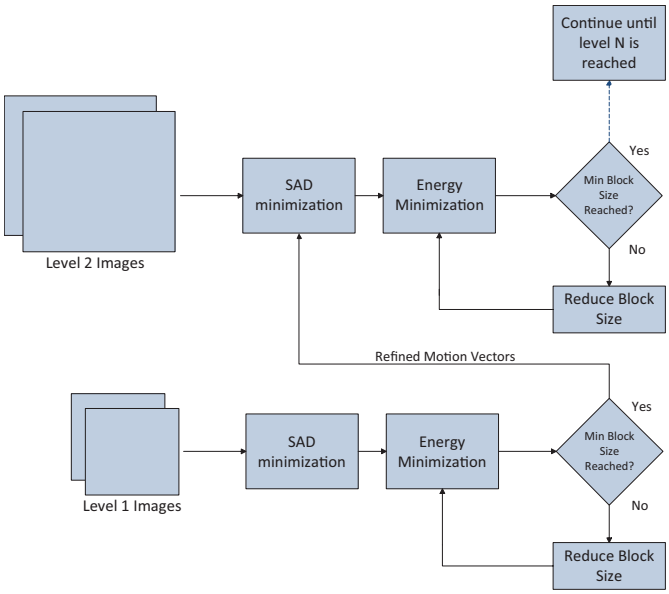
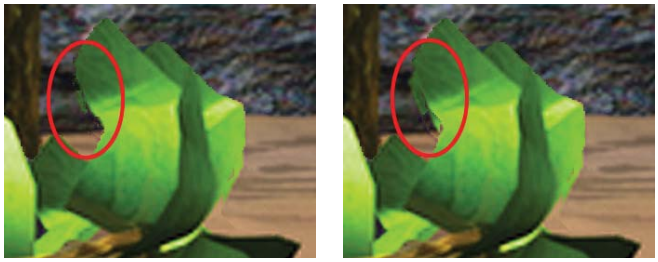


Fig. 13. Block diagram showing progression of motion estimation algorithm.



(a) (b)

Fig. 14. Visual comparison of MC frames for Grove 2 sequence. (a) Section of MC frame using (19). (b) Section of MC frame using (12).

TABLE I
IMPROVEMENT OF NEW ENERGY MINIMIZATION (19) OVER (12)

Sequence	Endpoint Error for (19)	Endpoint Error for (12)	Improv. in dB
Dimetrodon	0.215	0.215	0.00 dB
Grove 2	0.202	0.254	0.98 dB
Grove 3	0.618	0.683	0.43 dB
Hydrangea	0.230	0.230	0.00 dB
Rubber Whale	0.161	0.161	0.00 dB
Urban 2	0.418	0.472	0.53 dB
Urban 3	0.662	0.897	1.32 dB
Venus	0.315	0.330	0.20 dB

V. RESULTS

In all the results that follow, we used a four-level hierarchy and the algorithm given in Algorithm 2 to obtain quarter-pixel MVs. For images with a VGA resolution, the run time of our method is approximately two seconds using unoptimized code. The experiments were carried out on an Intel Core i7 875 K with 12 GB RAM using only a single execution thread.

We demonstrate the effectiveness of the new energy minimization framework using the eight ground truth test sequences from Middlebury University [1]. In Table I, we show comparisons of endpoint error for the MVs of

Average endpoint error	avg. rank	Army (Hidden texture) GT im0 im1 all	Army disc	Army untext	Mequon (Hidden texture) GT im0 im1 all	Mequon disc	Mequon untext	Schefflera (Hidden texture) GT im0 im1 all	Schefflera disc	Schefflera untext	Wooden (Hidden texture) GT im0 im1 all	Wooden disc	Wooden untext	Grove (Synthetic) GT im0 im1 all	Grove disc	Grove untext	Urban (Synthetic) GT im0 im1 all	Urban disc	Urban untext	Yosemite (Synthetic) GT im0 im1 all	Yosemite disc	Yosemite untext	Teddy (Stereo) GT im0 im1 all	Teddy disc	Teddy untext
BlockOverlap [61]	59.0	0.17 60	0.35 52	0.16 62	0.48 59	1.02 53	0.46 59	0.75 60	1.31 58	0.59 58	0.40 60	1.47 60	0.33 61	0.96 47	1.26 36	1.14 60	1.40 59	1.47 38	0.86 62	0.31 76	0.22 71	0.86 77	1.20 64	1.78 51	2.19 72
HBpMotionGpu [43]	59.5	0.17 60	0.41 64	0.13 57	0.61 65	1.34 67	0.59 66	0.95 66	1.68 68	0.76 66	0.38 59	1.63 65	0.27 60	1.11 62	1.49 64	1.27 63	0.66 39	1.53 42	0.45 38	0.20 60	0.18 63	0.28 49	1.12 62	2.04 60	1.67 63
2D-CLG [1]	61.4	0.28 71	0.62 73	0.21 68	0.67 68	1.21 62	0.70 68	1.12 70	1.80 71	0.99 73	1.07 73	2.06 71	1.12 73	1.23 70	1.52 68	1.62 74	1.54 65	2.15 67	0.96 66	0.10 4	**0.11** 1	0.16 8	1.38 71	2.26 71	1.83 68
Nguyen [33]	61.7	0.22 68	0.47 68	0.19 66	0.87 70	1.29 64	0.97 71	1.17 71	1.81 72	0.92 71	0.99 71	1.82 68	1.07 71	1.17 67	1.49 64	1.46 69	0.72 41	2.09 64	0.60 51	0.14 31	0.14 35	0.20 19	1.37 69	2.18 70	1.86 69
Horn & Schunck [3]	65.4	0.22 68	0.55 70	0.22 70	0.61 65	1.53 70	0.52 64	1.01 68	1.73 69	0.80 68	0.78 68	2.02 69	0.77 68	1.26 72	1.58 71	1.55 72	1.43 60	2.59 75	1.00 69	0.16 49	0.18 63	0.15 7	1.51 72	2.50 73	1.88 70
TI-DOFE [24]	67.3	0.38 76	0.64 74	0.47 75	1.16 74	1.72 73	1.26 76	1.39 77	2.06 78	1.17 76	1.29 74	2.21 73	1.41 76	1.27 73	1.61 72	1.57 73	1.28 56	2.57 74	1.01 70	0.13 23	0.15 44	0.16 8	1.87 74	2.71 74	2.53 73
Adaptive flow [45]	70.8	0.36 74	0.59 71	0.37 74	1.21 75	1.60 71	1.23 75	1.21 73	1.77 70	1.18 77	0.94 70	2.03 70	0.97 69	1.20 69	1.57 70	1.08 56	1.73 69	1.90 56	1.12 72	0.59 78	0.37 78	1.37 78	1.37 69	2.16 68	1.81 67
SLK [47]	71.8	0.30 73	0.70 75	0.36 73	1.09 73	1.77 74	1.21 74	1.25 75	1.98 76	1.03 75	1.56 77	2.26 74	1.71 77	1.54 77	1.82 76	2.14 77	2.02 74	2.79 77	1.36 74	0.17 54	0.16 51	0.26 40	2.43 76	3.18 76	3.31 76
PGAM+LK [55]	74.0	0.37 75	0.70 75	0.59 77	1.08 72	1.89 76	1.15 72	0.94 65	1.59 64	0.88 70	1.40 76	3.28 78	1.33 75	1.37 75	1.70 74	1.67 75	2.10 75	2.53 73	1.39 76	0.36 77	0.28 77	0.65 74	1.89 75	2.72 75	2.71 74
FOLKI [16]	75.0	0.29 72	0.73 77	0.33 72	1.52 77	1.96 77	1.80 77	1.23 74	2.04 77	0.95 72	0.99 71	2.20 72	1.08 72	1.53 76	1.85 77	2.07 76	2.14 76	3.23 78	1.60 77	0.26 74	0.21 70	0.68 75	2.67 77	3.27 77	4.32 77
Pyramid LK [2]	76.9	0.39 77	0.61 72	0.61 78	1.67 78	1.78 75	2.00 78	1.50 78	1.97 75	1.38 78	1.57 78	2.39 75	1.78 78	2.94 78	3.72 78	2.98 78	3.33 78	2.74 76	2.43 78	0.30 75	0.24 75	0.73 76	3.80 78	5.08 78	4.88 78

Fig. 15. Screenshots taken from Middlebury benchmark [1] with our endpoint error results highlighted (and shown by arrow). The full tables are available at http://vision.middlebury.edu/flow/eval/results/.

Average interpolation error	avg. rank	Mequon (Hidden texture) im0 GT im1 all	Mequon disc	Mequon untext	Schefflera (Hidden texture) im0 GT im1 all	Schefflera disc	Schefflera untext	Urban (Synthetic) im0 GT im1 all	Urban disc	Urban untext	Teddy (Stereo) im0 GT im1 all	Teddy disc	Teddy untext	Backyard (High-speed camera) im0 GT im1 all	Backyard disc	Backyard untext	Basketball (High-speed camera) im0 GT im1 all	Basketball disc	Basketball untext	Dumptruck (High-speed camera) im0 GT im1 all	Dumptruck disc	Dumptruck untext	Evergreen (High-speed camera) im0 GT im1 all	Evergreen disc	Evergreen untext
DPOF [18]	32.1	3.34 58	6.82 64	1.29 60	3.40 4	4.93 4	1.29 6	5.00 52	6.36 29	3.40 56	5.86 44	8.94 58	3.51 54	11.0 20	13.8 20	3.59 8	6.56 39	12.7 38	2.28 6	7.99 27	18.2 28	1.55 15	8.24 36	12.9 35	1.70 10
TC-Flow [46]	32.3	3.31 56	6.70 61	1.22 21	3.91 30	5.95 31	1.45 26	3.64 6	5.84 14	**1.28** 1	5.70 35	8.50 48	3.22 19	11.2 33	14.1 36	4.44 60	6.34 24	12.3 25	2.41 40	7.79 24	17.9 24	1.55 15	8.42 52	13.2 54	1.74 40
BlockOverlap [61]	32.9	2.98 11	5.47 9	1.33 67	4.38 46	6.09 35	1.88 63	4.26 33	5.57 5	3.14 51	5.56 25	7.32 4	4.14 70	11.1 25	13.9 22	3.77 36	6.41 30	12.3 25	2.54 65	7.75 20	17.4 14	3.02 76	**7.32** 1	**11.4** 1	1.78 56
Sparse-NonSparse [56]	33.0	3.07 24	5.88 27	1.21 11	3.61 10	5.33 10	1.33 9	4.29 34	7.47 57	2.19 34	5.37 11	7.74 15	3.21 16	11.5 50	14.5 55	4.36 56	6.66 50	12.9 51	2.41 40	8.69 51	20.1 53	1.67 35	8.27 40	13.0 42	1.70 10
Sparse Occlusion [54]	34.0	3.16 38	6.18 45	1.23 32	4.14 37	6.24 48	1.45 26	3.67 8	5.84 14	1.52 8	5.61 32	8.26 36	3.15 8	11.5 50	14.4 46	4.48 65	6.26 19	12.1 19	2.46 49	8.52 46	19.6 50	1.54 12	8.28 41	13.0 42	1.75 45

Fig. 16. Screenshots taken from Middlebury benchmark [1] with our interpolation error results highlighted (and shown by arrow). The full tables are available at http://vision.middlebury.edu/flow/eval/results/.

(19) and (12). As shown in Table I, the new energy minimization framework results in an improvement of 0.43 dB when averaged over all of the sequences. For the sequences in Table I where no improvement was reported, the energy framework of (12) does a sufficient job of minimizing the block overlap through the use of the SAD and smoothness constraints, i.e., it is not necessary to incorporate an overlap regularizer. For the "Grove 2" sequence, the large improvement can be attributed to the ability of the overlap regularizer to reduce errors around the occluded edges of the leaves. The large improvement for the "Urban 3" sequence is a result of the improvement in MVs around the edges of the image; the overlap regularizer prevents a large overlap of MC blocks at the edges. For a small region from the "Grove 2" sequence, we show the visual improvement of (19) over that of (12) using the motion-compensated frames in Figs. 14(a) and (b), respectively. In Fig. 14, improvements generated by (19) can be seen in the top half of the images near the left edges of the leaves (indicated by red ellipses). Similar improvements throughout the "Grove 2" sequence are responsible for the large improvement reported in Table I.

In addition to the results shown in Table I, we also submitted our results to the Middlebury online benchmark for comparison with other motion estimation algorithms. The image database (without ground truth MVs) provided by Middlebury University contains 12 sets of two-frame images at resolutions of 640 × 480, 584 × 388, 420 × 360, and 316 × 252. The online benchmark results are shown in Figs. 15 and 16.

All of the algorithms shown in Figs. 15 and 16 are based on optical flow, with the exception of "Adaptive Flow [47]," which is based on block matching. The optical-flow-based algorithms try to infer a low-level segmentation of the image through complex regularizers or explicit image segmentation. However, these algorithms generally have high computational complexity and perform poorly for regions with small structures. In contrast, the results in Figs. 15 and 16 show that our algorithm is capable of outperforming several optical-flow-based without the need for complex regularization or segmentation, and it does so while keeping computational complexity low.

The rankings in Figs. 15 and 16 show how our algorithm performs in terms of endpoint and interpolation error, respectively. For the interpolation error, our algorithm outperforms all other block-based methods as well as optical-flow-based methods with long run times, and it produces the smallest interpolation error for the "Evergreen" sequence compared to all 67 algorithms currently in the database. At the time of this writing, our algorithm has one of the fastest run-times on the "Urban" sequence.

VI. Conclusion

The motion validity and estimation framework proposed in this paper uses a block-overlap-based validity metric to assign confidence values to each MV and improve the quality of the motion field. In section II, it was shown that the proposed validity metric does not depend on neighboring MVs, image features, or manual thresholds, and it outperforms other block-based validity metrics in the literature. In addition, as shown in section II-D, the proposed validity metric consistently produces de-interlaced images with higher PSNR than other metrics.

In addition to characterizing the confidence of MVs, the proposed validity metric was also used as an additional regularizer in the energy function. By introducing the block-overlap-based regularizer, we were able to provide a more uniform distribution of blocks, reduce the dependence on block size, and improve the quality of the motion field in terms of endpoint error. The published results for the Middlebury

sequences in section V show that our method performs well compared to other state-of-the-art methods in the literature, and it is well-suited for applications which require a real-time approach.

In future works, we wish to address the potential application of the proposed method in video compression systems. In video compression applications, the degradation due to inaccurate MVs is not as significant as for applications such as MC frame interpolation. The interpolation error results from the Middlebury benchmark shown in section V indicate that the proposed method performs similarly to the best optical-flow-based methods, and outperforms all of the algorithms for the "Evergreen" sequence. More research is needed to determine how these small differences in interpolation error relate to video compression efficiency and quality.

REFERENCES

[1] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 1–31, 2011.

[2] J. Weickert, A. Bruhn, T. Brox, and N. Papenberg, "A survey on variational optic flow methods for small displacements," in *Mathematical Models for Registration and Applications to Medical Imaging* (Mathematics in Industry), vol. 10, H.-G. Bock, F. Hoog, A. Friedman, A. Gupta, H. Neunzert, W. R. Pulleyblank, T. Rusten, F. Santosa, A.-K. Tornberg, V. Capasso, R. Mattheij, H. Neunzert, O. Scherzer, and O. Scherzer, Eds. Berlin, Germany: Springer-Verlag, 2006, pp. 103–136.

[3] M. Bierling, "Displacement estimation by hierarchical block matching," in *Proc. SPIE, Visual Commun. Image Process.*, vol. 1001, pp. 942–2087, Oct. 1988.

[4] I. Richardson, *H.264 and MPEG-4 Video Compress: Video Coding for Next-Generation Multimedia*. New York: Wiley, 2003.

[5] C. Bartels and G. de Haan, "Smoothness constraints in recursive search motion estimation for picture rate conversion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1310–1319, Oct. 2010.

[6] G. de Haan, P. Biezen, H. Huijgen, and O. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368–379, Oct. 1993.

[7] H. B. Yin, X. Z. Fang, H. Yang, S. Y. Yu, and X. K. Yang, "Motion vector smoothing for true motion estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 2. May 2006, p. 2.

[8] M. Santoro, G. AlRegib, and Y. Altunbasak, "Block-overlap-based validity metric for hybrid de-interlacing," in *Proc. Int. Conf. Image Process.*, 2012, pp. 1–4.

[9] D. Wang, A. Vincent, and P. Blanchfield, "Hybrid de-interlacing algorithm based on motion vector reliability," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 1019–1025, Aug. 2005.

[10] M. Liu and Y. Shen, "Multiframe super resolution based on block motion vector processing and kernel constrained convex set projection," *Proc. SPIE*, vol. 7257, p. 72571J, Jan. 2009.

[11] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst. Man Cybern.*, vol. 8, no. 6, pp. 460–473, Jun. 1978.

[12] E. Francois, J.-F. Vial, and B. Chupeau, "Coding algorithm with region-based motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 97–108, Feb. 1997.

[13] A. Patti, M. Sezan, and A. Tekalp, "Robust methods for high-quality stills from interlaced video in the presence of dominant motion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 328–342, Apr. 1997.

[14] C. W. Therrien, *Decision Estimation and Classification: An Introduction to Pattern Recognition and Related Topics*. New York: Wiley, 1989.

[15] L. Vandendorpe, L. Cuvilier, B. Maison, P. Queluz, and P. Delogne, "Motion-compensated conversion from interlaced to progressive formats," *Signal Process.: Image Commun.*, vol. 6, no. 3, pp. 193–211, 1994.

[16] G. Thomas, "A comparison of motion-compensated interlace-to-progressive conversion methods," *Signal Process., Image Commun.*, vol. 12, no. 3, pp. 209–229, 1998.

[17] G. De Haan and E. Bellers, "Deinterlacing-an overview," *Proc. IEEE*, vol. 86, no. 9, pp. 1839–1857, Sep. 1998.

[18] M. Santoro, G. AlRegib, and Y. Altunbasak, "Adaptive search-based hierarchical motion estimation using spatial priors," in *Proc. Int. Conf. Comput. Vis. Theory Appl.*, 2012, pp. 1–5.

[19] A. Wedel and D. Cremers, *Stereo Scene Flow for 3D Motion Analysis*. New York: Springer-Verlag, 2011.

[20] M. J. Black, "Robust dynamic motion estimation over time," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1991, pp. 296–302.

[21] J. Wang, D. Wang, and W. Zhang, "Temporal compensated motion estimation with simple block-based prediction," *IEEE Trans. Broadcast.*, vol. 49, no. 3, pp. 241–248, Sep. 2003.

Michael Santoro (M'02) received the B.S. degrees in both electrical engineering and computer engineering from the University of Florida, Gainesville, in 2005, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2007 and 2012, respectively.

He is currently a Post-Doctoral Researcher with Pontifica Universidad Católica de Chile, Santiago, Chile. His current research interests include resolution enhancement of images and video, video coding, and signal processing applications for magnetic resonance imaging.

Dr. Santoro was a recipient of the Top 10% Paper Award at the 2012 International Workshop on Multimedia Signal Processing, the Outstanding Service Award at the Center for Signal and Image Processing in 2009, the Outstanding Graduate Teaching Assistant Award from the Georgia Institute of Technology in 2007 and 2009, the Best Use of Custom Designed Chips Award at the IEEE 2005 Robotics Competition, and the Best Senior Design Project Award from the University of Florida in 2005.

Ghassan AlRegib (SM'10) is currently an Associate Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, where his research group is involved in research on projects related to image and video processing and communications, immersive communications, seismic processing, quality of images and videos, and 3-D video processing.

Prof. AlRegib was the recipient of the ECE Outstanding Graduate Teaching Award in 2001, the Center for Signal and Image Processing (CSIP) Research Award, and the CSIP Service Award, both in 2003, and the ECE Outstanding Junior Faculty Member Award at Georgia Tech in 2008. He was the Chair of the Special Sessions Program at the IEEE International Conference on Image Processing (ICIP) in 2006. He was the Area Editor of Columns and Forums in the *IEEE Signal Processing Magazine* from 2009 to 2012, an Associate Editor of the *IEEE Signal Processing Magazine* from 2007 to 2009, and the Tutorials Co-Chair of the IEEE ICIP in 2009. He was on a number of technical program committees of several international workshops and conferences. He has been a member on the Editorial Board of the *Wireless Networks Journal* since 2009. He is also the Chair of the Speech and Video Processing Track at Asilomar 2012. He is a consultant for a number of companies and organizations.

Yucel Altunbasak (F'12) is currently the President of the Scientific and Technological Research Council of Turkey. In 1999, he was an Associate Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, where he became a Professor in 2009. He was involved in research with 19 Ph.D. students. He has authored or co-authored more than 170 papers in journals and conferences, and holds 40 patents, including the patents in pending.

Dr. Altunbasak was the recipient of the National Science Foundation CAREER Award in 2002 and the 2003 Outstanding Junior Faculty Member Award from the School of Electrical and Computer Engineering, Georgia Institute of Technology. He was an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Signal Processing: Image Communication*, and the *Journal of Circuits, Systems and Signal Processing*. He was the President of TOBB in Turkey from 2009 to 2011.