Optical Flow Estimation Using Temporally Oversampled Video

SukHwan Lim, *Member, IEEE,* John G. Apostolopoulos, *Member, IEEE,* and Abbas El Gamal, *Fellow, IEEE*

***Abstract*— Recent advances in imaging sensor technology make high frame rate video capture practical. As demonstrated in previous work, this capability can be used to enhance the performance of many image and video processing applications. The idea is to use the high frame rate capability to temporally oversample the scene and thus to obtain more accurate information about scene motion and illumination. This information is then used to improve the performance of image and standard frame-rate video applications. The paper investigates the use of temporal oversampling to improve the accuracy of optical flow estimation (OFE). A method for obtaining high accuracy optical flow estimates at a conventional standard frame rate, e.g. 30 frames/s, by first capturing and processing a high frame rate version of the video is presented. The method uses the Lucas-Kanade algorithm to obtain optical flow estimates at a high frame rate, which are then accumulated and refined to estimate the optical flow at the desired standard frame rate. The method demonstrates significant improvements in optical flow estimation accuracy both on synthetically generated video sequences and on a real video sequence captured using an experimental high-speed imaging system. It is then shown that a key benefit of using temporal oversampling to estimate optical flow is the reduction in motion aliasing. Using sinusoidal input sequences, the reduction in motion aliasing is identified and the desired minimum sampling rate as a function of the velocity and spatial bandwidth of the scene is determined. Using both synthetic and real video sequences it is shown that temporal oversampling improves OFE accuracy by reducing motion aliasing not only for areas with large displacements but also for areas with small displacements and high spatial frequencies. The use of other OFE algorithms with temporally oversampled video is then discussed. In particular the Haussecker algorithm is extended to work with high frame rate sequences. This extension demonstrates yet another important benefit of temporal oversampling, which is improving OFE accuracy when brightness varies with time.**

***Index Terms*— Optical flow estimation, Motion estimation, High speed imaging, CMOS image sensor, Temporal oversampling**

I. Introduction

A Key problem in the processing of video sequences is estimating the motion between video frames, often referred to as optical flow estimation (OFE). Once estimated, optical flow can be used in performing a wide variety of tasks such as video compression, 3-D surface structure estimation, super-resolution, motion-based segmentation and image registration. Optical flow estimation based on standard frame rate video sequences, such as 30 frames/s, has been extensively researched with several classes of methods developed including gradient-based, region-based matching, energy-based, Bayesian, and phase-based [1], [2], [3]. Many of these methods require a large number of operations to achieve acceptable estimation accuracy and perform poorly for large displacements [4], [5], [6], [7], [8]. Moreover, certain applications require more accurate and dense velocity measurements of optical flow than can be achieved by these methods.

Oversampling has been previously applied in many 1D signal processing applications such as analog-to-digital conversion, audio processing, and communications. Oversampling has also been applied to still imaging applications such as extending dynamic range by capturing multiple images [9], [10]. However, until recent advances in the design of CMOS image sensors and digital signal processors, it was considered too costly to capture sequences at a high frame rate and use them in standard rate video processing applications. On the capture side, several researchers have recently demonstrated implementations of high frame rate capture up to several thousand frames per second [11], [12], [13]. Krymski *et al.* [11] describe a 1024×1024 Active Pixel Sensor (APS) with column level ADC achieving frame rate of 500 frames/s. Stevanovic *et al.* [12] describe 256×256 APS with 4 analog outputs achieving frame rate of 1000 frames/s. Kleinfelder *et al.* [13] describe a 352×288 Digital Pixel Sensor (DPS) with per pixel bit parallel ADC achieving 10,000 frames/s or 1 Giga-pixels/s.

There are several benefits of using high frame rate sequences for OFE. First, the brightness constancy assumption [1], [2], [3] made implicitly or explicitly in most OFE algorithms becomes more valid as frame rate increases. Thus it is expected that using high frame rate sequences can enhance the estimation accuracy of these algorithms. Another important benefit is that as frame rate is increased the captured sequence exhibits less motion aliasing. Indeed large errors due to motion aliasing can occur even when using the best optical flow estimators. For example, when motion aliasing occurs a wagon wheel might appear to rotate backwards even to a human observer. This specific example is discussed in more detail in Section III. There are many instances when the standard frame rate of 30 frames/s is not sufficient to avoid motion aliasing and thus incorrect optical flow estimates [8], [14]. Note that motion aliasing not only depends on the velocities but also on the spatial bandwidths. Thus, capturing sequences at a high frame rate not only helps when velocities are large but also for complex images with low velocities but high spatial bandwidths. In [15], [16], Handoko describes a method using high frame rate video sequence for block-based motion vector estimation at standard frame rate, commonly used in video compression standards such as MPEG. Their paper proposed an iterative block matching algorithm utilizing a high frame rate sequence to generate motion vectors at 30 frames/s. The main focus was to reduce computational complexity and hence reduce power consumption. The reduction in computational complexity was achieved by utilizing high frame rate sequences to effectively

SukHwan Lim and John Apostolopoulos are with Hewlett-Packard Laboratories (sukhwan@hpl.hp.com, japos@hpl.hp.com) and Abbas El Gamal is with Stanford University (abbas@isl.stanford.edu). The research was done when the first author was at Stanford University.

reduce the search area between two consecutive frames.

In [17], [18] a method for obtaining accurate optical flow estimates at standard frame rate from a high frame rate sequence was described. This paper presents a more detailed and complete study of the initial work reported in [17], [18]. In Section II we present a method based on the well-known Lucas-Kanade algorithm for computing accurate optical flow estimates at a standard frame rate using a temporally oversampled (high frame rate) version of the video sequence. Using synthetic input sequences generated by warping of a still image, we show that the proposed method provides significant improvements in accuracy. In addition, using real sequences captured from a high-speed camera, we demonstrate that the proposed method provides motion fields that more accurately represent the motion that occurred in the video. In Section III we briefly review 3-D spatio-temporal sampling theory, and analyze the effects of temporal sampling rate and motion aliasing on OFE accuracy. We present simulation results using sinusoidal input sequences showing that the minimum frame rate needed to achieve high accuracy is largely determined by the minimum frame rate necessary to avoid motion aliasing (which may be produced by large displacements, or small displacements with high spatial bandwidths). In Section IV we discuss how the proposed method can be used with OFE algorithms other than the Lucas-Kanade algorithm. In particular, we extend the Haussecker algorithm [19] to work with high frame rate sequences. Furthermore, with this extension we show that another important benefit of temporal oversampling is improved optical flow estimation accuracy even when the brightness varies with time.

II. OFE Using Temporally Oversampled Video Sequences

A. Proposed Method

In this subsection we present a method for obtaining high accuracy optical flow estimates at a standard frame rate by capturing and processing a high frame rate version of the video. The idea is to estimate optical flow at a high frame rate and then carefully integrate it temporally to estimate the optical flow between frames at the slower standard frame rate. Temporal integration, however, must be performed without losing the accuracy gained by using the high frame rate sequence. Obviously, if the temporal integration does not preserve the accuracy provided by the high frame rate sequence, then this approach would lose many of its benefits.

The block diagram of our proposed method is shown in Figure 1 for the case when the frame rate is 3 times the standard frame rate. We define OV as the *oversampling factor* (i.e., the ratio of the capture frame rate to the standard frame rate) and thus $OV = 3$ in the block diagram. Consider the sequence of high-speed frames beginning with a standard-speed frame (shaded frame in the figure) and ending with the following standard-speed frame. We first obtain high accuracy optical flow estimates between consecutive high-speed frames. These estimates are then used to obtain an accurate estimate of the optical flow between the two standard-speed frames.

We first describe how optical flow at a high frame rate is estimated. Although virtually any OFE method can be employed

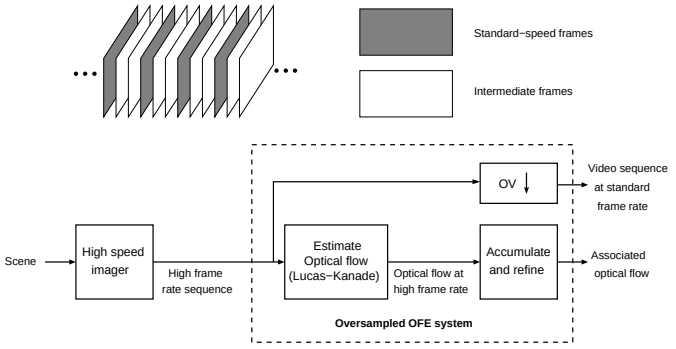


Fig. 1. The block diagram of the proposed method ($OV = 3$).

for this stage, we decided to use a gradient-based method since higher frame rate leads to reduced motion aliasing and better estimation of temporal derivatives, which directly improve the performance of such methods. In addition, because of the smaller displacements between consecutive frames in a high-speed sequence, smaller kernel sizes for smoothing and computing gradients can be used, which reduces the memory and computational requirements of the method.

Of the gradient-based methods, we chose the well known Lucas-Kanade's algorithm [20], which was shown to be among the most accurate and computationally efficient methods for optical flow estimation [1]. Each frame is first pre-filtered using a spatio-temporal low pass filter to reduce aliasing and systematic error in the gradient estimates. The gradients i_x, i_y, and i_t are typically computed using a 5-tap filter [1]. Assuming the optical flow is constant for pixels in the neighborhood, the displacement is calculated by computing a least squares estimate of the brightness constraint equations ($i_x d_x + i_y d_y + i_t = 0$) for the pixels in the neighborhood. Optionally, the brightness constraints near the center of the neighborhood can be given higher weight than those farther from the center. The weighted least-squares estimate of the optical flow $(d_x(x,y), d_y(x,y))$ can be found by solving the 2×2 linear equation

$$\begin{bmatrix} \sum_{u,v} w i_x^2 & \sum_{u,v} w i_x i_y \\ \sum_{u,v} w i_x i_y & \sum_{u,v} w i_y^2 \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix} = - \begin{bmatrix} \sum_{u,v} w i_x i_t \\ \sum_{u,v} w i_y i_t \end{bmatrix},$$

where w is the weighting function that assigns higher weight to the center of the neighborhood and the summations are over the neighborhood whose sizes are typically 5×5 pixels. Note that we have omitted the spatial parameters (u,v) in $w(u,v)$ and $(x-u, y-v)$ in $i_x(x-u,y-v), i_y(x-u,y-v)$ and $i_t(x-u, y-v)$ to simplify the notation. The smallest eigenvalue of the 2×2 matrix in the equation can be used as a confidence measure [1] such that we can discard the optical flow estimates whose confidence measure is lower than a threshold.

After optical flow has been estimated at the high frame rate, we use it to estimate the optical flow at the standard frame rate. This is the third block of the block diagram in Figure 1. The key in this stage is to integrate optical flow temporally without losing the accuracy gained using the high frame rate sequences. A straightforward approach would be to simply accumulate the optical flow estimates between consecutive

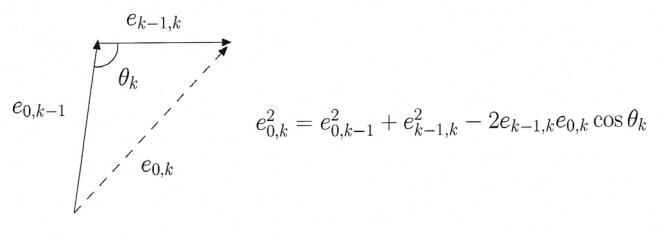


Fig. 2. The accumulation of error vectors when accumulating optical flow without using refinement.

high-speed frames along the motion trajectories. The problem with this approach is that errors can accumulate with the accumulation of the optical flow estimates. To understand how errors can accumulate for a pixel, consider the diagram in Figure 2, where $e_{k,l}$ is the magnitude of the OFE error vector between frames k and l. Assuming that θ_k, the angles between the error vectors in the figure are random and uniformly distributed and that the mean squared magnitude of the OFE error between consecutive high-speed frames are equal, i.e., $E[e_{j-1,j}^2] = E[e_{0,1}^2]$ for $j = 1, \ldots, k$, the total mean-squared error is given by

$$\begin{aligned} E[e_{0,k}^2] &= E[e_{0,k-1}^2 + e_{k-1,k}^2 - 2e_{k-1,k}e_{0,k-1}\cos\theta_k] \\ &= \sum_{j=1}^{k} E[e_{j-1,j}^2] - 2\sum_{j=1}^{k} E[e_{j-1,j}e_{0,j-1}\cos\theta_j] \\ &= \sum_{j=1}^{k} E[e_{j-1,j}^2] = kE[e_{0,1}^2], \end{aligned}$$

which grows linearly with k. On the other hand, if the optical flow estimation errors are systematic, i.e., line up from one frame to the next, and their magnitudes are temporally independent, which yields $E[e_{j-1,j}e_{l-1,l}] = E[e_{j-1,j}]E[e_{l-1,l}]$, then the total mean-squared error is given by

$$\begin{aligned} E[e_{0,k}^2] &= E[e_{0,k-1}^2 + e_{k-1,k}^2 + 2e_{k-1,k}e_{0,k}] \\ &= E[(e_{0,k-1} + e_{k-1,k})^2] \\ &= E[(\sum_{j=1}^{k} e_{j-1,j})^2] = k^2E[e_{0,1}^2], \end{aligned}$$

which grows quadratically with k. In practice, the optical flow estimation error was shown to have a random component and a non-zero systematic component by several researchers [5], [7], [21], [22], and as a result, the mean-squared error $E[e_{0,k}^2]$ is expected to grow faster than linear but slower than quadratic in k.

To prevent this error accumulation, we add a refinement (or correction) stage after each iteration (see Figure 4). We obtain frame $\hat{k}$ by warping frame 0 according to our accumulated optical flow estimate $\tilde{d}$, and assume that frame k is obtained by warping frame 0 according to the true motion between the two frames, (which we do not know). By estimating the displacement between frames k and $\hat{k}$, we can estimate the error between the true flow and the initial estimate $\tilde{d}$. In the refinement stage, we estimate this error and add it to the accumulated optical flow estimate. Although the estimate

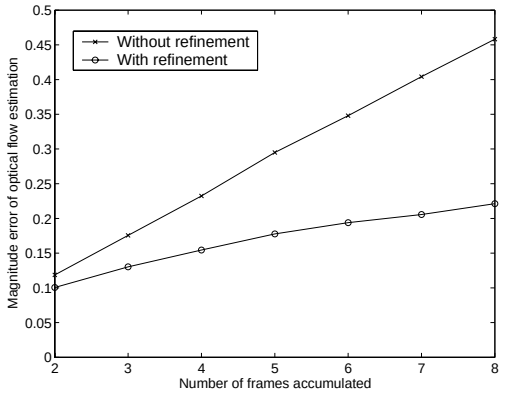


Fig. 3. The average magnitude of the OFE error with and without refinement.

of the error is not perfect, we found that it significantly reduces error accumulation. Figure 3 illustrates this reduction by comparing the average magnitude of the OFE error with and without the refinement stage. This figure was generated by performing OFE on the synthetic video sequence of scene 1 in Table 1, but operating at different frame rates. The generation of synthetic video sequences is described in more detail in Section II-B.

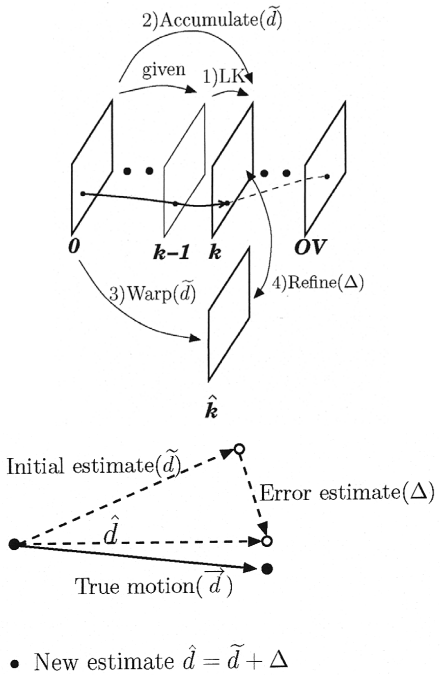


Fig. 4. Accumulate and refine stage.

A description of the proposed method is given below. Consider $OV + 1$ high-speed frames beginning with a standard-speed output frame and ending with the following one. Number the frames $0, 1, \ldots, OV$ and let $\hat{d}_{k,l}$ be the estimated optical flow (displacement) from frame k to frame l, where $0 \leq k \leq l \leq OV$. The end goal is to estimate the optical flow between frames 0 and OV, i.e, $\hat{d}_{0,OV}$.

Proposed method:

1) Capture a standard-speed frame, set $k = 0$.
2) Capture the next high-speed frame and set $k = k + 1$.
3) Estimate $\hat{d}_{k-1,k}$ using Lucas-Kanade method.
4) $\tilde{d}_{0,k} = \hat{d}_{0,k-1} + \hat{d}_{k-1,k}$ where addition of optical flow estimates are along the motion trajectories.
5) Estimate Δ_k, the displacement between frame k and $\hat{k}$.
6) Set refined estimate $\hat{d}_{0,k} = \tilde{d}_{0,k} + \Delta_k$.
7) Repeat steps 2 through 6 until $k = OV$
8) Output $\hat{d}_{0,OV}$ the final estimate of optical flow at the standard frame rate

The number of operations required to compute the optical flow using our method is roughly $2OV$ times that required to compute the optical flow using standard frame rate sequences. It is obtained by assuming that the computational complexity of 2)Warp and 3)Accumulate (in Figure 4) is much lower than that of 1)LK and 4)Refine. For example, when $OV = 4$, the computational complexity is increased by approximately a factor of 8. Liu et. al. [4] discussed the trade off between accuracy and complexity in some well-known OFE methods and showed that the range of the computational complexity for those methods span several orders of magnitude. By examining Table 1 and Figure 2 in [4], our method with $OV = 4$ would be less compute intensive than methods such as Fleet & Jepson, Horn & Schunk, Bober and Anandan's method. This is because the standard Lucas-Kanade method is one of the least compute intensive OFE methods. Also, note that since our method is iterative, its memory requirement is independent of the frame rate. Furthermore, since the proposed method uses 2-tap temporal filter for smoothing and estimating temporal gradients, its memory requirement is less than that of the conventional Lucas-Kanade method, which typically uses a 5-tap temporal filter [1]. Thus, we believe that it is feasible to implement the proposed method when OV is not prohibitively large.

Although our method does not require real-time operation to perform successful OFE, it may be beneficial to briefly consider the feasibility of real-time operation. Liu et. al. [4] and Bober et. al. [23] projected the number of years it would take to perform real-time operation of certain OFE algorithms. They used the execution time estimates of various OFE methods and assumed that computational power doubles each year [24]. Applying their approach, we project that it should be possible to perform OFE with our method ($OV = 4$) using a generic workstation in 2005. In addition, we believe that it would be even more feasible to implement our method in dedicated VLSI circuits since they typically have higher computational capabilities than generic CPUs. One potential problem in implementing our method in real-time is the high data rate requirement associated with high frame rate. It may be costly to implement our method in real-time because of the high inter-chip data rate between the sensor, the memory and processing chips. This problem can be alleviated by integrating the memory and the processing with the CMOS image sensor on the same chip. The idea is to (i) operate the sensor at a higher frame rate than the standard frame rate, (ii) process the high frame rate data on-chip, and (iii) only output the video frames and associated optical flow estimates at the standard frame rate [18], [25]. In this case, only the standard frame rate video frames and associated optical flow estimates need to be transferred off-chip because all the data transfers at high frame rate would occur inside the chip. Note that intra-chip data transfers can support much higher bandwidths and lower power consumption as compared to off-chip transfers. Overall, a single-chip solution would result in lower system cost and power consumption [18], [25].

B. Simulation and Results

In this subsection, we describe the simulations we performed using synthetically generated natural image sequences to test our optical flow estimation method. To evaluate the performance of the proposed method and compare with methods using standard frame rate sequences, we need to compute the optical flow using both the standard and high frame rate versions of the same sequence, and then compare the estimated optical flow in each case to the true optical flow. We use synthetically generated video sequences obtained by warping of a natural image. The reason for using synthetic sequences, instead of real video sequences, is that the amount of displacement between consecutive frames can be controlled and the true optical flow can be easily computed from the warping parameters. Also, in addition to results obtained from synthetic sequences, we provide experimental results using real sequences captured from a high-speed camera. We demonstrate that the proposed method provides motion fields which more accurately represent the motion that occurred in the video.

We use a realistic image sensor model [26] that incorporates motion blur and noise in the generation of the synthetic sequences, since these effects can vary significantly as a function of frame rate, and can thus affect the performance of optical flow estimation. In particular, high frame rate sequences have less motion blur but suffer from lower SNR, which adversely affects the accuracy of optical flow estimation. The image sensor in a digital camera comprises a 2-D array of pixels. During capture, each pixel converts incident photon flux into photocurrent. Since the photocurrent density $j(x, y, t)$ A/cm^2 is too small to measure directly, it is spatially and temporally integrated onto a capacitor in each pixel and the charge $q(m, n)$ is read out at the end of exposure time T. Ignoring dark current, the output charge from a pixel can be expressed as

$$q(m,n) = \int_0^T \int_{ny_0}^{ny_0+Y} \int_{mx_0}^{mx_0+X} j(x,y,t)dxdydt + N(m,n), \quad (1)$$

where x_0 and y_0 are the pixel dimensions, X and Y are the photodiode dimensions, (m, n) is the pixel index, and $N(m, n)$ is the noise charge. The noise is the sum of two independent components, shot noise and readout noise. The spatial and temporal integration results in low pass filtering that can cause motion blur. Note that the pixel intensity $i(m, n)$ commonly used in image processing literature is directly proportional to the charge $q(m, n)$.

The steps of generating a synthetic sequence are as follows.

1) Warp a high resolution (1312×2000) image using perspective warping to create a high resolution sequence.
2) Spatially and temporally integrate (according to Equation (1)) and subsample the high resolution sequence to obtain a low resolution sequence. In our example, we subsampled by factors of 4×4 spatially and 10 temporally to obtain each high-speed frame.
3) Add readout noise and shot noise according to the model.
4) Quantize the sequence to 8 bits/pixel.

Three different scenes derived from a natural image (Figure 5) were used to generate the synthetic sequences. For each scene, two versions of each video, one captured at a standard frame rate ($OV = 1$) and the other captured at four times the standard frame rate ($OV = 4$), are generated as described above. The maximum displacements were between 3 and 4 pixels/frame at the standard frame rate. We performed optical flow estimation on the ($OV = 1$) sequences using the standard Lucas-Kanade method as implemented by Barron *et al.* [1] and on the ($OV = 4$) sequences using the proposed method. Both methods generate optical flow estimates at a standard frame rate of 30 frames/s. Note that the standard Lucas-Kanade method was implemented using 5-tap temporal filters for smoothing and estimating temporal gradients while the proposed method used 2-tap temporal filters. The resulting average angular errors between the true and the estimated optical flows are given in Table I. The densities of all estimated optical flows are close to 50%, where the density is defined as the percentage of pixels that have confidence measures higher than a set threshold. The average optical flow estimation error was obtained by averaging over those pixels.

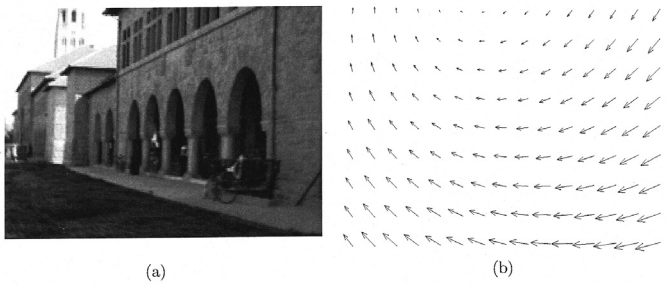


Fig. 5. (a) One frame of a test sequence and (b) its known optical flow.

Scene	Lucas-Kanade ($OV = 1$)		Proposed ($OV = 4$)	
	Angular	Magnitude	Angular	Magnitude
1	4.43°	0.24	3.43°	0.14
2	3.94°	0.24	2.91°	0.17
3	4.56°	0.32	2.67°	0.17

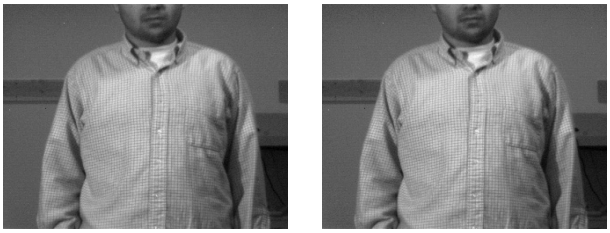
TABLE I
AVERAGE ANGULAR ERROR AND MAGNITUDE ERROR USING LUCAS-KANADE METHOD WITH STANDARD FRAME RATE SEQUENCES VERSUS THE PROPOSED METHOD USING HIGH FRAME RATE SEQUENCES.

The results demonstrate that using the proposed method in conjunction with the high frame rate sequence can achieve higher accuracy. Note that the displacements were kept relatively small (as measured at the standard frame rate) to make comparison between the two methods more fair. As displacements increase, the accuracy of the standard Lucas-Kanade method deteriorates rapidly and hierarchical methods should be used in the comparison instead. On the other hand, the proposed method is much more robust to large displacements because of the higher sampling rate.

To investigate the gain in accuracy of the proposed method for large displacements, we applied the Lucas-Kanade method, our proposed method with $OV = 10$, and the hierarchical matching-based method by Anandan [6] as implemented by Barron [1] to a synthetic sequence. The maximum displacement was 10 pixels/frame at the standard frame rate. The average angular errors and magnitude errors of the estimated optical flows are given in Table II. For comparison, we calculated average errors for Anandan's method at locations where Lucas-Kanade method gave valid optical flow, although Anandan's method can provide 100% density. Thus, values in the table were calculated where the densities of all estimated optical flows are close to 50%.

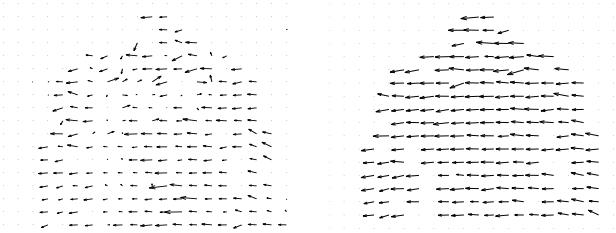
	Angular	Magnitude
Lucas-Kanade	9.18°	1.49
Anandan's	4.72°	0.53
Proposed ($OV = 10$)	1.82°	0.21

TABLE II
AVERAGE ANGULAR AND MAGNITUDE ERROR USING LUCAS-KANADE, ANANDAN'S AND PROPOSED METHOD.



(a) Frame 9 (b) Frame 13

Fig. 6. Two frames from the real video sequence captured at 120 frames/s.



(a) Standard Lucas-Kanade (b) Proposed method

Fig. 7. Optical flow estimation results for a real video sequence.

We also applied our method to real video sequences captured using an experimental high speed imaging system [27]. The system is based on the DPS chip described in [13] and can operate at frame rates of up to 1400 frames/s. Although

we cannot measure the error quantitatively, we can make qualitative comparisons. Figure 6 shows frames 9 and 13 of the real video sequence. It was captured at 120 frames/s ($OV = 4$) when a person was moving horizontally from right to left. The horizontal optical flow was near 2 pixels/frame for standard frame rate (30 frames/s) sequences. We tried to keep the horizontal velocity within the acceptable range for the standard Lucas-Kanade method. To estimate the optical flow between frame 9 and 13, our method used all the frames between frames 9 and 13, while the standard Lucas-Kanade method just used frames 9 and 13. The resulting optical flows are shown in Figure 7. The effect of aliasing can be seen in some areas in Figure 7 (a) where the optical flow estimates point in the *opposite direction* of the true motion. Note that this experiment with a real video sequence was performed in a favorable situation for the standard Lucas-Kanade method. We carried out similar experiments for larger displacements and the difference in the performance is significantly larger than what can be seen in Figure 7. The experiment also serves the purpose of examining the effect of aliasing in optical flow estimation when high spatial frequency is present with moderate displacements. It shows that the standard frame rate of, e.g., 30 frames/s is not sufficient to avoid motion aliasing and thus incorrect optical flow estimates. Note that the shirt has high spatial frequency although the displacements were small enough to be within general acceptable range of the standard Lucas-Kanade method. This illustrates that capturing sequences at a high frame rate not only helps when velocities are large but also for complex images with low velocities but high spatial bandwidths. Another benefit that is illustrated in Figure 7 is that the estimated motion field when using temporal oversampling exhibits significant coherence (i.e. it is smoothly varying) similar to the true motion of the object, and this property is very useful for compression as well as many other video signal processing applications that utilize motion fields.

III. Effect of Motion Aliasing on Optical Flow Estimation

This section reviews 3-D spatio-temporal sampling theory and investigates the effect of motion aliasing on the accuracy of optical flow estimation. Readers familiar with 3-D spatio-temporal sampling and motion aliasing may wish to skip the review in III-A, and continue with III-B. Motion aliasing can produce large errors even with the best optical flow estimators. Perhaps the most well known example is that of the wagon wheel in a Western movie (shot at 24 frames/s) which appears to be moving in the opposite direction from what physically makes sense given the wagon's motion. This section briefly reviews 3-D spatio-temporal sampling theory and describes how motion aliasing occurs not only for large displacements but also for small displacements with high spatial frequency. Our experiments also demonstrate that the minimum frame rate necessary to achieve good OFE performance is largely determined by the minimum frame rate necessary to prevent motion aliasing in the sequence.

A. Review of Spatio-temporal Sampling Theory

A simplified but insightful model of motion is that of global motion with constant velocity in the image plane. The pixel intensity, assuming this model, is given by

$$\begin{aligned} i(x,y,t) &= i(x-v_x, y-v_y, 0) \\ &= i_0(x-v_x, y-v_y), \end{aligned}$$

where $i_0(x,y)$ denotes the 2-D pixel intensity for $t=0$ and v_x and v_y are the global velocities in the x and y directions, respectively. This model is commonly assumed either globally or locally in many applications such as motion-compensated standards conversion and video compression. After taking the Fourier transform, we obtain

$$I(f_x, f_y, f_t) = I_0(f_x, f_y) \cdot \delta(f_x v_x + f_y v_y + f_t),$$

where $I_0(f_x, f_y)$ is the 2-D Fourier transform of $i_0(x,y)$ and $\delta(\cdot)$ is the 1-D Dirac delta function. Thus, it is clear that the energy of $I(f_x, f_y, f_t)$ is confined to a plane given by $f_x v_x + f_y v_y + f_t = 0$. If we assume that $i_0(x,y)$ is bandlimited such that $I(f_x, f_y) = 0$ for $|f_x| > B_x$ and $|f_y| > B_y$, then $i(x,y,t)$ is bandlimited temporally as well, i.e, $I(f_x, f_y, f_t) = 0$ for $|f_t| > B_t$ where $B_t = B_x v_x + B_y v_y$. Note that the temporal bandwidth depends on *both the spatial bandwidths and the spatial velocities*. To simplify our discussion, we assume in the following that sampling is performed only along the temporal direction and that the spatial variables are taken as continuous variables (no sampling along the spatial directions). This assumption simplifies the analysis, and interestingly is not entirely unrealistic, since it is analogous to the shooting of motion picture film, where each film frame corresponds to a temporal sample of the video. Figure 8 shows the spatio-temporal spectrum of video when sampled only in the temporal direction. For simplicity of illustration, we consider its projection onto the (f_x, f_t)-plane, where the support can be simplified to $f_x v_x + f_t = 0$. Each line represents the spatio-temporal support of the sampled video sequence.

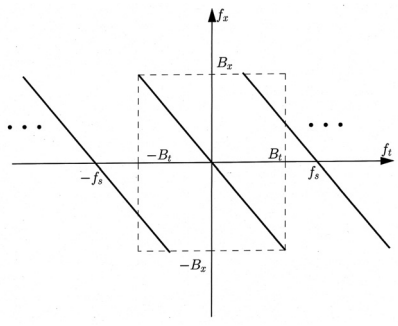


Fig. 8. Spatio-temporal spectrum of a temporally sampled video.

Let us consider the problem of how fast we should sample the original continuous video signal along the temporal dimension such that it can be perfectly recovered from its samples. Assume that an ideal low-pass filter with rectangular support in the 3-D frequency domain is used for reconstruction, although in certain ideal cases, a sub-Nyquist sampled signal can also be reconstructed by an ideal motion-compensated reconstruction

filter assuming the replicated spectra do not overlap (see [3] for details). To recover the original continuous spatio-temporal video signal from its temporally sampled version, it is clear from the figure that the temporal sampling frequency (or frame rate) f_s must be greater than $2B_t$ in order to avoid aliasing in the temporal direction. If we assume global motion with constant velocity v_x and v_y (in pixels per standard-speed frame) and spatially bandlimited image with B_x and B_y as the horizontal and vertical spatial bandwidths (in cycles per pixel), the minimum temporal sampling frequency $f_{\mathrm{s,Nyq}}$ to avoid motion aliasing is given by

$$f_{\mathrm{s,Nyq}} = 2B_t = 2B_x v_x + 2B_y v_y, \tag{2}$$

where $f_{\mathrm{s,Nyq}}$ is in cycles per standard-speed frame. (e.g. When standard frame rate is 30 frames/s, the 60 frames/s correspond to 2 cycles per standard-speed frame.) Note that the temporal sampling frequency in cycles per standard-speed frame is the oversampling factor OV. Moreover, since OV is an integer in our framework to ensure that standard-speed frames correspond to a captured high-speed frame (see Figure 1), the minimum oversampling factor to avoid motion aliasing, OV_{theo}, is

$$\begin{aligned} OV_{\mathrm{theo}} &= \lceil f_{\mathrm{s,Nyq}} \rceil \\ &= \lceil 2B_x v_x + 2B_y v_y \rceil. \end{aligned}$$

To illustrate this relationship consider the simple case of a sequence with only global motion in the horizontal direction (i.e., with $v_y = 0$). Figure 9 plots $OV_{\mathrm{theo}} = \lceil 2B_x v_x \rceil$ versus horizontal velocity and spatial bandwidth for this case.

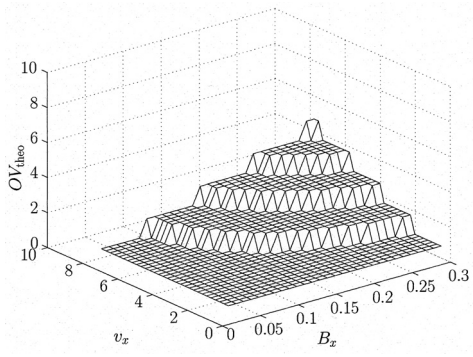


Fig. 9. Minimum OV to avoid motion aliasing, as a function of horizontal velocity v_x and horizontal spatial bandwidth B_x.

Motion aliasing can produce large errors even with the best optical flow estimators. The classic example is that of the wagon wheel in a Western movie (shot at 24 frames/s) which appears to be moving in the opposite direction from what physically makes sense given the wagon's motion (See Figure 10). In this example the wagon wheel is rotating counter-clockwise and we wish to estimate its motion from two frames captured at times $t = 0$ and $t = 1$. The solid lines represent positions of the wheel and spokes at $t = 0$ and the dashed lines represent the positions at $t = 1$. Optical flow is locally estimated for the two shaded regions of the image in Figure 10. As can be seen, the optical flow estimates are in the *opposite* direction of the true motion (as often experienced by a human observer). The wheel is rotating counter-clockwise, while the optical flow estimates from the local image regions would suggest that it is rotating clockwise. This error is caused by insufficient temporal sampling and the fact that optical flow estimation (and the human visual system) implicitly assume the smallest possible displacements (corresponding to a low-pass filtering of the possible motions). This example shows that motion aliasing can cause incorrect motion estimates for any OFE algorithm. To overcome motion aliasing, one must either sample sufficiently fast, or have a prior information about the possible motions as in the case of the moving wagon wheel, where the human observer makes use of the direction of motion of the wagon itself to correct the misperception about the rotation direction of the wheel.

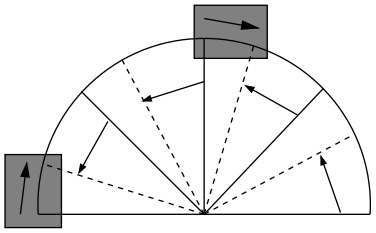


Fig. 10. Wagon wheel rotating counter-clockwise illustrating motion aliasing from insufficient temporal sampling: the local image regions (gray boxes) appear to move clockwise.

Let us consider the spatio-temporal frequency content of the local image regions in Figure 10. Since each shaded region has a dominant spatial frequency component and the assumption of global velocity for each small image region holds, its spatio-temporal frequency diagram can be plotted as shown in Figure 11 (A). The circles represent the frequency content of a sinusoid and the dashed lines represent the plane where most of the energy resides. Note that the slope of the plane is inversely proportional to the negative of the velocity. The spatio-temporal frequency content of the baseband signal after reconstruction by the OFE algorithm is plotted in Figure 11 (B). As can be seen aliasing causes the slope at which most of the energy resides to not only be different in magnitude, but also to have a different sign, corresponding to motion in the opposite direction.

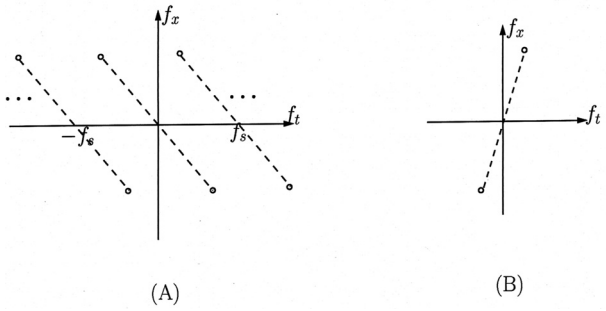


Fig. 11. Spatio-temporal diagrams of, (A) the shaded region in Figure 10 and (B) its baseband signal.

Before continuing, it is useful to review other popular

approaches for improving OFE performance by varying the sampling rate. It is well known that motion aliasing or temporal aliasing adversely affects the accuracy of OFE and many researchers pointed out that large systematic errors arise when displacements are large [5], [6], [7], [8]. Multi-resolution algorithms applied spatially help overcome these problems as the spatial subsampling in effect reduces the motion between frames by the subsampling factor [6], [14]. Although temporal aliasing can be caused by large displacements, it can also be caused by high spatial frequencies with low-to-moderate displacements. In this case, the spatial low-pass filtering used as part of the coarse-to-fine estimation can partially overcome the aliasing in the high spatial frequencies [7], [8], [14]. Note that these multi-resolution approaches adapt the spatial sampling, but do not modify the temporal sampling: the frame rate for processing is given by the frame rate of the desired output OFE which is also equal to the frame rate of the video acquisition.

B. Simulation and Results

In this subsection we discuss simulation results using sinusoidal test sequences and the synthetically generated natural image sequence used in Subsection II-B. The reason for using sinusoidal sequences is to assess the performance of the proposed method as spatial frequency and velocity are varied in a controlled manner. As discussed in the previous subsection, motion aliasing depends on both the spatial frequency and the velocity and can have a detrimental effect on optical flow estimation. Using a natural sequence, it would be difficult to understand the behavior of the proposed method with respect to spatial frequency, since in such a sequence, each local region is likely to have different spatial frequency content and the Lucas-Kanade method estimates optical flow by performing spatially local operations. In addition, typical figures of merit, such as average angular error and average magnitude error, would be averaged out across the frame. The use of sinusoidal test sequences can overcome these problems and can enable us to find the minimum OV needed to obtain a desired accuracy, which can then be used to select the minimum high-speed frame rate for a natural scene.

We considered a family of 2-D sinusoidal sequences with equal horizontal and vertical frequencies $f_x = f_y$ moving only in the horizontal direction at speed v_x (i.e., $v_y = 0$). For each f_x and v_x, we generated a sequence with $OV = 1$ and performed optical flow estimation using the proposed method. We then incremented OV by 1 and repeated the simulation. We noticed that the average error drops rapidly beyond a certain value of OV and that it remained relatively constant for OVs higher than that value. Based on this observation we defined the minimum oversampling ratio OV_{exp} as the OV value at which the magnitude error drops below a certain threshold. In particular, we chose the threshold to be 0.1 pixels/frame. Once we found the minimum value of OV, we repeated the experiment for different spatial frequencies and velocities. The results are plotted in Figure 12.

Recall the discussion in the previous subsection (including Figure 9) on the minimum oversampling factor as a function of spatial bandwidth and velocity needed to avoid motion aliasing. Note the similarity between the theoretical results in Figure 9 and their experimental counterpart in Figure 12. This is further illustrated by the plot of their difference and its histogram in Figure 13. This similarity suggests that reduction in motion aliasing is one of the most important benefits of using high frame rate sequences. The difference in Figure 13 can be further reduced by sampling at a higher rate than $\lceil f_{\text{s,Nyq}} \rceil$ to better approximate brightness constancy and improve the estimation of temporal gradients. It has been shown that gradient estimators using a small number of taps suffer from poor accuracy when high frequency content is present [28], [29]. In our implementation, we used a 2-tap temporal gradient estimator, which performs accurately for temporal frequencies $f_t < \frac{1}{3}$ as suggested in [28]. Thus we need to sample at a rate higher than 1.5 times the Nyquist temporal sampling rate. Choosing an OV curve that is 1.55 times the Nyquist rate (i.e., $\lceil 1.55 f_{\text{s,Nyq}} \rceil$), in Figure 14 we plot the difference between the OV_{exp} curve in Figure 12 and the OV curve. Note the reduction in the difference achieved by the increase in frame rate.

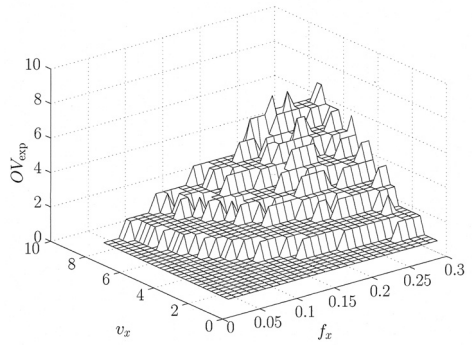


Fig. 12. Minimum OV as a function of horizontal velocity v_x and horizontal spatial frequency f_x.

We also investigated the effect of varying OV and motion aliasing on the accuracy using the synthetically generated image sequences presented in Subsection II-B. Figure 15 plots the average angular error of the optical flow estimates using the proposed method for OV between 1 and 14. The synthetic test sequence had a global displacement of 5 pixels/frame at $OV = 1$. As OV was increased, motion aliasing and the error due to temporal gradient estimation decreased, leading to higher accuracy. The accuracy gain resulting from increasing OV, however, levels off as OV is further increased. This is caused by the decrease in sensor SNR due to the decrease in exposure time and the levelling off of the reduction in motion aliasing. For this example sequence, the minimum error is achieved at $OV = 6$, where displacements between consecutive high-speed frames are approximately 1 pixel/frame.

To investigate the effect of motion aliasing, we also estimated the energy in the image that leads to motion aliasing. Note that since the sequence has global motion with constant velocity, the temporal bandwidth of the sequence can be estimated as $B_t = 5B_x + 5B_y$ by assuming the knowledge of initial estimates of $v_x = v_y = 5$ pixels/frame. Thus, motion

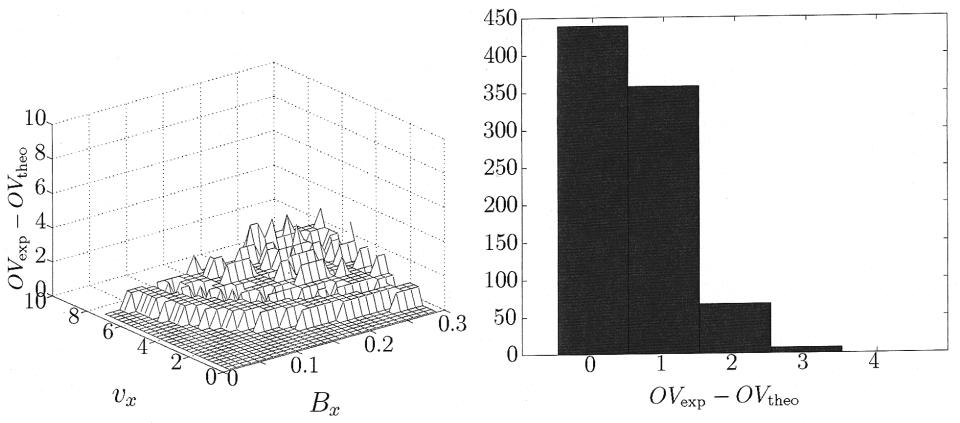


Fig. 13. Difference between empirical minimum OV for good optical flow estimation performance and OV corresponding to the Nyquist rate.

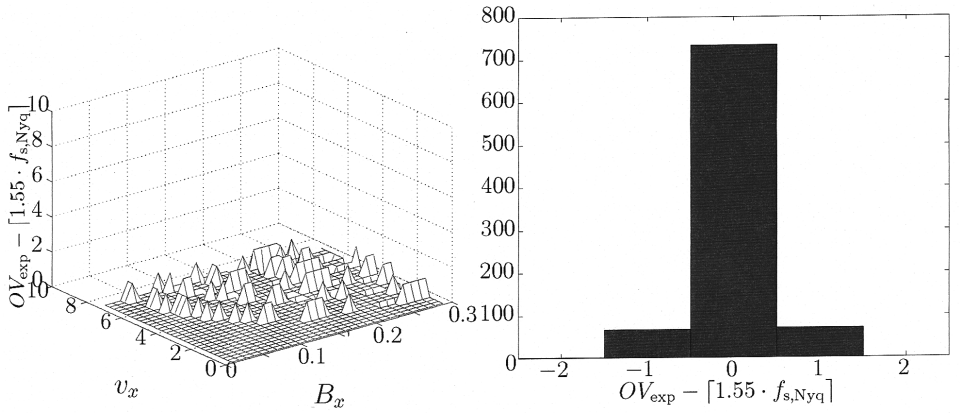


Fig. 14. Difference between empirical minimum OV for good optical flow estimation performance and OV corresponding to 1.55 times the Nyquist rate.

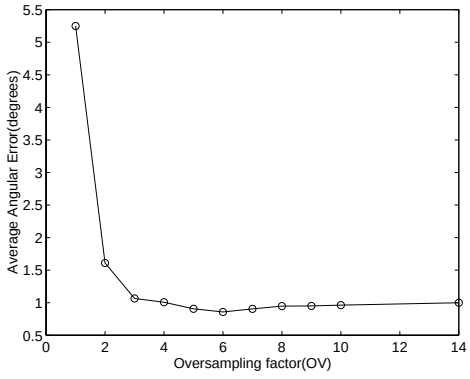


Fig. 15. Average angular error versus oversampling factor (OV).

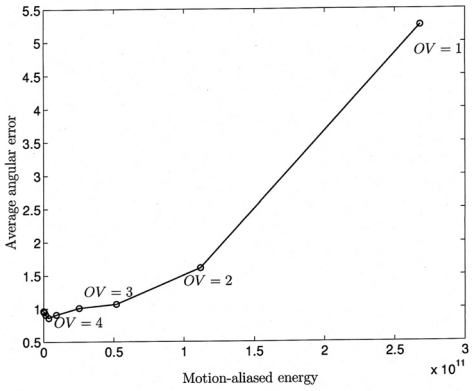


Fig. 16. Average angular error versus energy in the image that leads to motion aliasing.

aliasing occurs for spatial frequencies $\{f_x, f_y\}$ that satisfy the constraint $f_x + f_y > OV/10$. By using 2D-DFT of the first frame and this constraint, we calculated the energy in the sequence that is motion aliased for different OVs. Figure 16 plots the average angular error versus the energy that is motion aliased. Each point corresponds to an OV value and it is clear that the performance of the proposed OFE method is largely influenced by the presence of motion aliasing. This confirms that temporal oversampling leads to reduction of motion aliasing and increased accuracy for a practical OFE algorithm. Thus, a key advantage of high frame rate is the reduction of motion aliasing. Also, this example shows that with initial estimates of velocities, we can predict the amount of energy in the image that will be aliased. This can be used to identify the necessary frame rate to achieve high accuracy optical flow estimation for a specific scene.

IV. Extension to Handle Brightness Variation

In the previous sections we described and tested a method for obtaining high accuracy optical flow at a standard frame

rate using a high frame rate sequence. We used Lucas-Kanade method to estimate optical flow at high frame rate and then accumulated and refined the estimates to obtain optical flow at standard frame rate. The Lucas-Kanade method assumes brightness constancy, and although high frame rate makes this assumption more valid, in this section we show that brightness variations can be handled more effectively using other estimation methods. Specifically, we show that by using an extension of the Haussecker [19] method, temporal oversampling can benefit optical flow estimation even when brightness constancy assumption does not hold.

Several researchers have investigated the problem of how to handle the case when the brightness constancy assumption does not hold [19], [30], [31], [32], [33], [34], [35]. It has been shown that a linear model with offset is sufficient to model brightness variation in most cases [30], [31], [35]. For example, Negahdaripour *et al.* developed an OFE algorithm based on this assumption and demonstrated good performance [30], [31]. Haussecker *et al.* developed models for several cases of brightness variation and described a method for coping with them [19]. We will use Haussecker's framework with the assumption of linear brightness variation for estimating optical flow at high frame rate.

A. Review of Models for Brightness Variation

We begin with a brief summary of the framework described in [19]. The brightness change is modelled as a parameterized function h, i.e.,

$$i(\mathbf{x}(t), t) = h(i_0, t, \mathbf{a}),$$

where $\mathbf{x}(t)$ denotes the path along which brightness varies, $i_0 = i(\mathbf{x}(0), 0)$ denotes the image at time 0, and $\mathbf{a}$ denotes a Q-dimensional parameter vector for the brightness change model. The total derivative of both sides of this equation yields

$$(\nabla i)^T \mathbf{v} + i_t = f(i_0, t, \mathbf{a}), \tag{3}$$

where f is defined as

$$f(i_0, t, \mathbf{a}) = \frac{\mathrm{d}}{\mathrm{d}t}[h(i_0, t, \mathbf{a})].$$

Note that when brightness is constant, $f = 0$ and Equation 3 simplifies to the conventional brightness constancy constraint. The goal is to estimate the parameters of the optical flow field $\mathbf{v}$ and the parameter vector $\mathbf{a}$ of the model f. Remembering that $h(i_0, t, \mathbf{a} = \mathbf{0}) = i_0$, we can expand h using the Taylor series around $\mathbf{a} = \mathbf{0}$ to obtain

$$h(i_0, t, \mathbf{a}) \approx i_0 + \sum_{k=1}^{Q} a_k \frac{\partial h}{\partial a_k}.$$

Thus, f can be written as a scalar product of the parameter vector $\mathbf{a}$ and a vector containing the partial derivatives of f with respect to the parameters a_k, i.e.,

$$f(i_0, t, \mathbf{a}) = \sum_{k=1}^{Q} a_k \frac{\partial f}{\partial a_k} = (\nabla_{\mathbf{a}} f)^T \mathbf{a}. \tag{4}$$

Using Equation 4, Equation 3 can be expressed as

$$\mathbf{c}^T \mathbf{p}_h = 0,$$

where

$$\begin{aligned} \mathbf{c} &= [(\nabla_{\mathbf{a}} f)^T, (\nabla i)^T, i_t]^T \\ \mathbf{p}_h &= [-\mathbf{a}^T, \mathbf{v}^T, 1]^T. \end{aligned}$$

Here, the $(Q+3)$-dimensional vector $\mathbf{p}_h$ contains the flow field parameters and the brightness parameters of h. The vector $\mathbf{c}$ combines the image derivative measurements and the gradient of f with respect to $\mathbf{a}$. To solve for $\mathbf{p}_h$, we assume that $\mathbf{p}_h$ remains constant within a local space-time neighborhood of N pixels. The constraints from the N pixels in the neighborhood can be expressed as

$$\mathbf{G}\mathbf{p}_h = 0,$$

where $\mathbf{G} = [\mathbf{c}_1, ..., \mathbf{c}_N]^T$. The estimate of $\mathbf{p}_h$ can be obtained by a total least squares (TLS) solution.

B. Using Haussecker Method with High Frame Rate

We assume a linear model with offset for brightness variation which yields $f = a_1 + a_2 i_0$. We use Haussecker's method to estimate v_x, v_y, a_1 and a_2 for every high-speed frame. We then accumulate and refine v_x, v_y, a_1 and a_2 in a similar manner to the method described in Section II to obtain optical flow estimates at a standard frame rate.

The parameters v_x and v_y are accumulated and refined exactly as before, and we now describe how to accumulate and refine a_1 and a_2 along the motion trajectories. To accumulate a_1 and a_2, we first define $\hat{a}_{1(k,l)}$ and $\hat{a}_{2(k,l)}$ to be the estimated brightness variation parameters between frames k and l along the motion trajectory. We estimate $\hat{a}_{1(k-1,k)}$ and $\hat{a}_{2(k-1,k)}$ and assume that $\hat{a}_{1(0,k-1)}$ and $\hat{a}_{2(0,k-1)}$ are available from the previous iteration. Since $f = a_1 + a_2 i_0$, we model the brightness variation such that

$$\begin{aligned} i_{k-1} - i_0 &= \hat{a}_{1(0,k-1)} + \hat{a}_{2(0,k-1)} i_0 \\ i_k - i_{k-1} &= \hat{a}_{1(k-1,k)} + \hat{a}_{2(k-1,k)} i_{k-1}, \end{aligned}$$

for each pixel in frame 0, where i_k is the intensity value for frame k along the motion trajectory. By arranging the terms and eliminating i_{k-1}, we can express i_k in terms of i_0 such that

$$i_k = \hat{a}_{1(k-1,k)} + (1 + \hat{a}_{2(k-1,k)})(\hat{a}_{1(0,k-1)} + (1 + \hat{a}_{2(0,k-1)}) i_0). \tag{5}$$

Let $\tilde{a}_{1(0,k)}$ and $\tilde{a}_{2(0,k)}$ denote the accumulated brightness variation parameters between frames 0 and k along the motion trajectory. Therefore, by definition, $i_k = \tilde{a}_{1(0,k)} + (1 + \tilde{a}_{2(0,k)}) i_0$ and by comparing this equation with Equation 5, accumulated brightness variation parameters are obtained by

$$\begin{aligned} \tilde{a}_{1(0,k)} &= \hat{a}_{1(k-1,k)} + (1 + \hat{a}_{2(k-1,k)}) \hat{a}_{1(0,k-1)} \\ \tilde{a}_{2(0,k)} &= \hat{a}_{2(k-1,k)} + (1 + \hat{a}_{2(k-1,k)}) \hat{a}_{2(0,k-1)}. \end{aligned}$$

Frame $\hat{k}$ is obtained by warping frame 0 according to our initial estimate of optical flow between frames 0 and k and changing the brightness according to $\tilde{a}_{1(0,k)}$ and $\tilde{a}_{2(0,k)}$, i.e.,

$$\text{Frame } \hat{k} = (1 + \tilde{a}_{2(0,k)}) i_k (x - \tilde{v}_{x(0,k)}, y - \tilde{v}_{y(0,k)}) + \tilde{a}_{1(0,k)},$$

where $\tilde{v}_{x(0,k)}$ and $\tilde{v}_{y(0,k)}$ are the accumulated optical flow estimates between frames 0 and k. By estimating the optical

flow and brightness variation parameters between original frame k and motion-compensated frame $\hat{k}$, we can estimate the error between the true values and the initial estimates obtained by accumulating. For the optical flow, we estimate the error and add it to our initial estimate, whereas for the brightness variation parameters, we perform the refinement as

$$\begin{aligned}\hat{a}_{1(0,k)} &= a_{1\Delta} + (1 + a_{2\Delta})\tilde{a}_{1(0,k)} \\ \hat{a}_{2(0,k)} &= a_{2\Delta} + (1 + a_{2\Delta})\tilde{a}_{2(0,k)},\end{aligned}$$

where $a_{1\Delta}$ and $a_{2\Delta}$ are the brightness variation parameters between frames k and $\hat{k}$. The accumulation and refinement stage is repeated until we have the parameters between frames 0 and OV.

We tested this method using the sequences described in Subsection II-B but with global brightness variations. In these sequences, however, the global brightness changed with $a_{1(0,OV)} = 5$ and $a_{2(0,OV)} = 0.1$. We performed optical flow estimation on the $OV = 1$ sequences using the Haussecker's method and on the $OV = 4$ sequences using our extended method. The resulting average angular errors and magnitude errors between the true and the estimated optical flows are given in Table III.

Scene	Haussecker ($OV = 1$)		Proposed ($OV = 4$)	
	Angular	Magnitude	Angular	Magnitude
1	5.12°	0.25	3.33°	0.15
2	6.10°	0.32	2.99°	0.18
3	7.72°	0.54	2.82°	0.18

TABLE III

AVERAGE ANGULAR ERROR AND MAGNITUDE ERROR USING HAUSSECKER'S METHOD WITH $OV = 1$ SEQUENCES VERSUS PROPOSED EXTENDED METHOD WITH $OV = 4$ SEQUENCES.

These results demonstrate that using high frame rate, high accuracy optical flow estimates can be obtained even when brightness varies with time, i.e., when brightness constancy assumption does not hold. Furthermore, with this extension, we have also demonstrated that our proposed method can be used with OFE algorithms other than the Lucas-Kanade algorithm.

V. SUMMARY

In this paper, we proposed a method for providing improved optical flow estimation accuracy for video at a conventional standard frame rate, by initially capturing and processing the video at a higher frame rate. The method begins by estimating the optical flow between frames at the high frame rate, and then accumulates and refines these estimates to produce accurate estimates of the optical flow at the desired standard frame rate. The method was tested on synthetically generated video sequences and the results demonstrate significant improvements in OFE accuracy. We also validated our algorithm with real video sequences captured at high frame rate by showing that the proposed method provides motion fields which more accurately represent the motion that occurred in the video. Also, with sinusoidal input sequences, we showed that reduction of motion aliasing is an important potential benefit of using high frame rate sequences. We also described methods to estimate the required oversampling rate to improve the optical flow accuracy, as a function of the velocity and spatial bandwidth of the scene. The proposed method can be used with other OFE algorithms besides the Lucas-Kanade algorithm. For example, we began with the Haussecker algorithm, designed specifically for optical flow estimation when the brightness varies with time, and extended it with the proposed method to work on high frame rate sequences. Furthermore, we demonstrated that our extended version provides improved accuracy in optical flow estimation as compared to the original Haussecker algorithm operating on video captured at the standard frame rate. Finally, we believe that temporal oversampling of video is a promising and soon-to-be-practical approach for enhancing the performance of many image and video applications.

ACKNOWLEDGMENTS

The work in the paper is partially supported under the DARPA Grant No. N66001-02-1-8940 and under Programmable Digital Camera Program by Agilent, Canon, HP, Interval Research and Kodak. The authors would like to thank Ali Ozer Ercan for helping with high-speed video capture and would also like to thank the reviewers for valuable comments that greatly improved the paper.

REFERENCES

[1] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of Optical Flow Techniques," *International Journal of Computer Vision*, vol. 12(1), pp. 43–77, February 1994.

[2] C. Stiller and J. Konrad, "Estimating Motion in Image Sequences," *IEEE Signal Processing Magazine*, pp. 70–91, July 1999.

[3] A. Murat Tekalp, *Digital Video Processing*, Prentice-Hall, New Jersey, USA, 1995.

[4] H. Liu, T.-H. Hong, M. Herman, and R. Chellappa, "Accuracy vs. Efficiency Trade-offs in Optical Flow Algorithms," in *Proceedings of the 4th European Conference on Computer Vision, Cambridge, UK*, April 1996, vol. 2, pp. 174–183.

[5] R. Battiti, E. Amaldi, and C. Koch, "Computing Optical Flow Across Multiple Scales: An Adaptive Coarse-to-Fine Strategy," *International Journal of Computer Vision*, vol. 6(2), pp. 133–145, 1991.

[6] P. Anandan, "A computational framework and an algorithm for the measurement of visual motion," *International Journal of Computer Vision*, vol. 2, no. 3, pp. 283–310, January 1989.

[7] J. Weber and J. Malik, "Robust Computation of Optical Flow in Multi-Scale Differential Framework," *International Journal of Computer Vision*, vol. 14, no. 1, pp. 67–81, 1995.

[8] W. J. Christmas, "Spatial filtering requirements for gradient optical flow measurement," in *Proceedings of the Ninth British Machine Vision Conference*, September 1998, vol. 1, pp. 185–194.

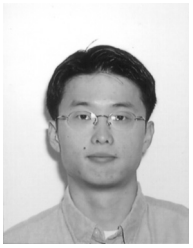
[9] D. Yang, A. El Gamal, B. Fowler, and H. Tian, "A 640×512 CMOS Image Sensor with Ultra Wide Dynamic Range Floating-Point Pixel-Level ADC," in *Digest of Technical Papers, IEEE International Solid-State Circuits Conference*, February 1999, pp. 308–309.

[10] X. Liu and A. El Gamal, "Photocurrent estimation from multiple nondestructive samples in CMOS image sensors," in *Proceedings of the SPIE Electronic Imaging Conference*, January 2001, vol. 4306.

[11] A. Krymski, D. Van Blerkom, A. Andersson, N. Block, B. Mansoorian, and E.R. Fossum, "A High Speed, 500 Frames/s, 1024×1024 CMOS Active Pixel Sensor," in *Proceedings of Symposium on VLSI Circuits*, 1999, pp. 137–138.

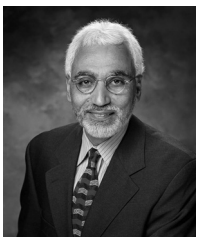
[12] N. Stevanovic, M. Hillegrand, B. J. Hostica, and A. Teuner, "A CMOS Image Sensor for High Speed Imaging," in *Digest of Technical Papers, IEEE International Solid-State Circuits Conference*, February 2000, number 104-105.

[13] S. Kleinfelder, S.H. Lim, X. Liu, and A. El Gamal, "A $10,000$ Frame/s 0.18μm CMOS Digital Pixel Sensor with Pixel-Level Memory," in *Digest of Technical Papers, IEEE International Solid-State Circuits Conference*, February 2001, pp. 88–89.
[14] Eero Simoncelli, *Ph.D. Thesis: Distributed Analysis and Representation of Visual Motion*, MIT, Cambridge, MA, USA, January 1993.
[15] D. Handoko, S. Kawahito, Y. Takokoro, M. Kumahara, and A. Matsuzawa, "A CMOS Image Sensor for Focal-plane Low-power Motion Vector Estimation," in *Proceedings of Symposium on VLSI Circuits*, June 2000, pp. 28–29.
[16] D. Handoko, S. Kawahito, Y. Takokoro, M. Kumahara, and A. Matsuzawa, "On Sensor Motion Vector Estimation with Iterative Block Matching and Non-Destructive Image Sensing," *IEICE Transactions on Electronics*, vol. E82-C, no. 9, pp. 1755–1763, September 1999.
[17] S.H. Lim and A. El Gamal, "Optical Flow Estimation Using High Frame Rate Sequences," in *Proceedings of the 2001 International Conference on Image Processing*, October 2001, vol. 2, pp. 925–928.
[18] Suk Hwan Lim, *Ph.D. Thesis: Video Processing Applications of High Speed CMOS Image Sensors*, Stanford University, Stanford, CA, USA, March 2003.
[19] H. W. Haussecker and D. J. Fleet, "Computing Optical Flow with Physical Models of Brightness Variation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 661–673, 2001.
[20] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of DARPA Image Understanding*, 1981, pp. 121–130.
[21] J. K. Kearney, W. B. Thompson, and D. L. Boley, "Optical flow estimation: An error analysis of gradient-based methods with local optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9(2), pp. 229–244, March 1987.
[22] J. W. Brandt, "Analysis of Bias in Gradient-Based Optical Flow Estimation," in *Proceedings of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers*, 1994, vol. 1, pp. 721–725.
[23] M. Bober and J. Kittler, "Robust Motion Analysis," in *Proceedings of IEEE Conference on COmputer Vision and Pattern Recognition Seattle, WA*, June 1994, pp. 947–952.
[24] D. Patterson and J. Hennessy, *Computer Organization and Design: the Hardware/Software Interface*, Morgan Kaufman, San Mateo, USA, 1994.
[25] S.H. Lim and A. El Gamal, "Integration of Image Capture and Processing – Beyond Single Chip Digital Camera," in *Proceedings of the SPIE Electronic Imaging Conference*, January 2001, vol. 4306, pp. 219–226.
[26] A. El Gamal, *EE392B Classnotes: Introduction to Image Sensors and Digital Cameras*, http://www.stanford.edu/class/ ee392b, Stanford University, 2001.
[27] A. Ercan, F. Xiao, X. Liu, S.H. Lim, A. El Gamal, and B. Wandell, "Experimental High Speed CMOS Image Sensor System and Applications," in *Proceedings of the First IEEE International Conference on Sensors*, June 2002.
[28] E. P. Simoncelli, "Design of Multi-dimensional Derivative Filters," in *Proceedings of IEEE International Conference on Image Processing*, 1994, vol. 1, pp. 790–794.
[29] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing*, Prentice-Hall, New Jersey, USA, 1999.
[30] S. Negahdaripour and C.-H. Yu, "A Generalized Brightness Change Model for Computing Optical Flow," in *Proceedings of the Fourth International Conference on Computer Vision*, 1993, pp. 2–11.
[31] S. Negahdaripour, "Revised Definition of Optical Flow: Integration of Radiometric and Geometric Cues for Dynamic Scene Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 9, pp. 961–979, 1998.
[32] G. D. Hager and P. N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 10, pp. 1025–1039, 1998.
[33] D. W. Murray and B. F. Buxton, *Experiments in the Machine Interpretation of Visual Motion*, MIT Press, Cambridge, Mass., 1990.
[34] A. Verri and T. Poggio, "Motion Field and Optical Flow: Qualitative Properties," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 5, pp. 490–498, 1989.
[35] M. Mattavelli and A. Nicoulin, "Motion estimation relaxing the constancy brightness constraint," in *Proceedings of the IEEE International Conference on Image Processing*, 1994, vol. 2, pp. 770–774.



SukHwan Lim received the B.S. degree with honors in Electrical Engineering from Seoul National University, Seoul, Korea, in 1996, the M.S. degree and Ph.D. degree in Electrical Engineering from Stanford University in 1998 and 2003, respectively. He is currently a member of technical staff at Hewlett-Packard Laboratories. At Stanford University, he worked on programmable digital camera project and his research focused on high-speed CMOS image sensor and its video processing applications. He co-designed a 10,000 frames/sec CMOS image sensor and developed video processing applications that exploit the high-speed imaging capability. His research interests include image sensors, digital cameras and digital signal processing architectures for smart imaging devices. He is also interested in image/video processing algorithms such as optical flow estimation, denoising, deblurring and superresolution algorithms.

John Apostolopoulos received his B.S., M.S., and Ph.D. degrees from MIT, and he joined HP Labs in 1997 where he is currently a principal research scientist. He also teaches at Stanford where he is a Consulting Assistant Professor in EE. He received a best student paper award for part of his Ph.D. thesis, the Young Investigator Award (best paper award) at VCIP 2001 for his paper on multiple description video coding and path diversity for reliable video communication over lossy packet networks, and in 2003 was named "one of the world's top 100 young (under 35) innovators in science and technology" (TR100) by Technology Review. John contributed to the U.S. Digital Television and JPEG-2000 Security standards. He is a member of IEEE, and also servers as an Associate Editor of IEEE Transactions on Image Processing (2002-present) and of IEEE Signal Processing Letters (2000-2003), and a member of the IEEE Image and Multidimensional Digital Signal Processing (IMDSP) technical committee. His research interests include improving the reliability, fidelity, scalability, and security of media communication over wired and wireless packet networks.



Abbas El Gamal received his B.Sc. degree in Electrical Engineering from Cairo University in 1972, the M.S. in Statistics and the PhD in Electrical Engineering from Stanford in 1977 and 1978, respectively. From 1978 to 1980 he was an Assistant Professor of Electrical Engineering at USC . He has been on the faculty of of the Department of Electrical Engineering at Stanford since 1981, where he is currently a Professor and the Director of the Information Systems Laboratory. He was on leave from Stanford from 1984 to 1988 first as Director of LSI Logic Research Lab, then as cofounder and Chief Scientist of Actel corporation. In 1990 he cofounded Silicon Architects, which is currently part of Synopsys. His research interests include: digital imaging, network information theory, and integrated circuit design. He has authored or coauthored 150 papers and 25 patents in these areas. He has served on the board of directors and advisory boards of several IC and CAD companies. He is a Fellow of the IEEE and a member of the ISSCC Technical Program Committee.