An Empirical Comparison of High Definition Video and Regular Video in Optical Flow Computation

J. P. Adkins-Hill, J. M. Fortunato, Y. Zhang, and J. R. Sullins
Youngstown State University
{jadkins, jefortunato, yzhang, john}@cis.ysu.edu

Abstract

This paper presents a comparative study of high definition videos and regular videos in the context of optical flow estimation. The hypothesis is that, because of its higher resolution, a high definition video can yield more accurate optical flow data, which is critical for many motion-based researches. The experiments were carried out using videos that captured a wide variety of motions in both natural and indoor settings, which ensures a statistically sound comparison. Analysis through visual examinations and quantitative violation measures indicates that, in general, high definition videos are superior to regular videos, but the presence of fast moving objects could complicate a specific application.

1. Introduction

Extracting accurate motion data from videos is a fundamental topic in computer vision and image science and has found applications in medical imaging, robotic navigation, deformable model tracking, as well as facial expression analysis [1, 2, 3, 4, 5]. The quality of estimated motion is dependent upon many factors, such as the robustness of an algorithm, the complexity of motion, illumination variation and image resolution. In theory, one can improve the accuracy of motion data by merely increasing the image resolution with a high definition (HD) video. However, no experimental results are available to support such a hypothesis. A study that compares the performances of HD videos and regular videos is of strong theoretical interest.

Another motivation of conducting such a study is that more and more applications call for the use of high resolution images, especially in security surveillance and face biometrics [6]. However, before any large scale deployment, the benefit of using more expensive HD camcorders over regular surveillance cameras must be fully justified, usually starting with a performance evaluation analysis.

The objective of this study is to investigate whether a HD video will produce better motion results than a regular video. Optical flow is chosen over other motion estimation methods because it generates a pixel-level *dense* motion field, which is more sensitive to image resolution change. The resolution of HD videos is 4.5 times higher than that of regular videos. To ensure an objective comparison, several measures were taken: (i) HD and regular videos were acquired simultaneously under the same settings; (ii) Various types of motions were considered (rigid, non-rigid, elastic and fluid); (iii) Flows were generated using the same computational parameters; (iv) Two normalized violation ratios were used to quantify the flow results.

Many evaluation studies of optical flow algorithms have been reported, as represented by the work of Barron *et al* [7], and more recently, by the work of Baker *et al* using a new database [8]. But we are not aware of prior work on evaluating HD and regular videos in the context of optical flow computation. This study provides valuable preliminary information upon which more thorough investigations can be conducted.

2. Background

2.1. Optical Flow

Optical flow is a velocity field that represents the 3D motion of an object across a 2D image plane [9]. The governing equation uses two basic assumptions: (i) the observed brightness of a point remains constant over the time interval between two frames; (ii) the points in a small image window move at similar speeds and directions (spatial smoothness constraint).

Given the image brightness, $E(x,y,t)$, as a function of image coordinates (x,y) and time (t), the brightness constancy between two frames can be expressed as a total derivative by the conservation principle:

$$\frac{dE}{dt} = 0. \tag{1}$$

With an assumption that the brightness function is differentiable in both spatial and temporal domains, the above equation can be expanded as:

$$\frac{dE(x(t),y(t),t)}{dt} = \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial y}\frac{dy}{dt} + \frac{dE}{dt} = 0. \quad (2)$$

With the notation of $U = [u = dx/dt,\ v = dy/dt]^T$ being the motion vector, ∇E and E_t being the spatial and temporal derivatives of the brightness function, the brightness constancy equation can be written as:

$$(\nabla E)^T U + E_t = 0. \quad (3)$$

Eq. (3) does not guarantee a complete solution of the motion field (the well-known aperture problem), and must be regularized by a smoothness constraint:

$$obj\,(u,v) = (E_x u + E_y v + E_t)^2 + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2), \quad (4)$$

Where E_x, E_y, u_x, u_y, v_x and v_y denote the partial derivatives of the corresponding variables, and λ is the Lagrange multiplier (regularization coefficient). An optical flow solution (u, v) obtained by minimizing the objective function represents a compromise between the observed motion and the smoothness constraint.

2.2. Robust Algorithm

In applications, the two fundamental assumptions underlying the optical flow equation are often violated, especially in the presence of multiple motions that involve physical boundaries and depth discontinuities. Various algorithms have been developed to give a more reliable solution. Detailed discussions on those issues can be found in [7, 8].

In this study, a robust algorithm using a piecewise smoothing and multi-resolution strategy is used [10]. The violations of brightness constancy and spatial smoothness conditions are addressed in a framework that is similar to the "line-process" model, where over-smoothing of discontinuities is handled by a local adaptive method [11].

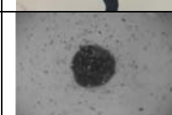
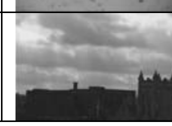
We used the implementation by Black and Anandan that is available on line [12]. In all of the experiments, only the default parameter values were used. This ensures that the outcome of this investigation is of reference value to similar comparison studies in the future.

3. Experiments

3.1. Data Set

To have a representative data set, a total of 35 scenes were recorded that include basic motion types: rigid, non-rigid and fluid motions (Table 1). A SONY SR1 camcorder (HD) and a Cannon XL1S camcorder (regular) were used in video acquisition. The resulting image resolutions are 1440 by 1080 pixels and 720 by 480 pixels, respectively. In each scene, two camcorders were operated with the same or similar parameter values: distance to objects, view angle, focus, recording time and length, and capture speed (30 frames per sec.).

Table 1. A few selected video samples from 35 scenes.

Motion	Setting	Descriptions	Sample
Rigid	Indoor	A spinning fan.	
	Outdoor	The traffic scene at an intersection.	
Non-rigid	Indoor	A hand moving at a steady speed. (articulated + elastic).	
	Outdoor	A person walking and skating on ice. (articulated).	
Fluid	Indoor	Water rotating in a bucket with particles.	
	Outdoor	Slow moving clouds.	

3.2. Computational Procedure

The procedure has three steps: (1) in each video, four sub-sequences were randomly selected (each has 200-500 frames); (2) both HD and regular video frames were geometrically normalized to the landmarks in the scene to ensure a fair comparison; (3) for each sub-sequence, a series of optical flow results were generated using pairs of adjacent frames. For example, a test run with a 100-frame sub-sequence would use these frame pairs: (1, 2), (2, 3) … (99, 100).

4. Results and Discussions

Given a frame pair, the program generates four outputs: (1) vertical flow, (2) horizontal flow, (3) spatial smoothness violation, and (4) brightness constancy violation. In flow images, an object's motion is normalized to a 0-255 grayscale, with brighter pixels for upward or rightward motion and darker pixels for downward or leftward motion.

4.1. Visual Examination

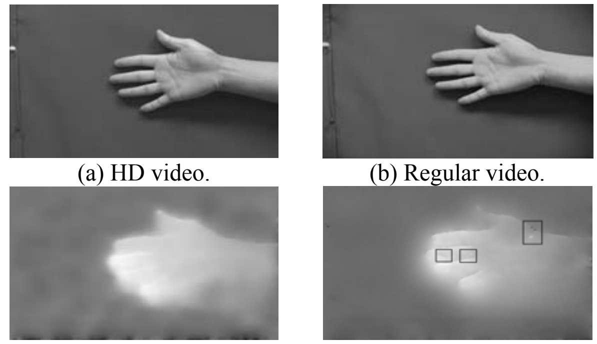
Figure 1 shows the vertical flows of a hand moving at a steady speed. Both HD and regular videos captured the overall hand motion quite well. However, a closer examination reveals that the flow image of HD videos is smoother and more continuous than that of regular videos. More importantly, the *blocking effect* is common in flow images of regular videos, as marked by the rectangles in Figure 1 (d). By blocking effect we mean that a block of pixels has very different intensities than its surrounding pixels, indicating an interruption to the normal flow field. The blocking effect occurred frequently along the boundary of a moving object, but could also show up inside the object. About 20-30% of regular frames had the blocking effect, while less than 8% of HD frames were affected. One possible explanation is that HD videos can pick up movements on much finer scales that are missed by regular videos.

Figure 2 shows the horizontal flows of a busy traffic scene of cars and pedestrians moving at varying speeds. Again, HD videos yielded better results than regular videos as reflected in the smoother and well-defined vehicle shapes in flow images. But unlike the HD videos of "moving hand" that delivered good results consistently, the quality of flow images from the HD videos of "intersection traffic" fluctuated over the entire sequence, probably caused by the changing speeds of vehicles and many occlusions. This raises the question of how to quantify the performance of HD and regular videos, even at the complex motion scenes.

4.2. Quantitative Measure: Violation Ratios

We computed two flow violation ratios (brightness constancy violation and spatial smoothness violation) by dividing the number of pixels of violations to the total number of pixels. Details about the definition and computation of the two violations can be found in [10].

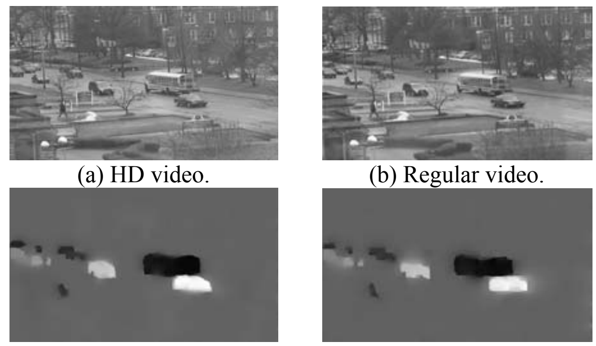
Violation ratios of the two aforementioned scenes using 50 randomly selected frame pairs are plotted in Figure 3 and Figure 4. In the first scene, HD videos show much less violations than regular videos in both ratios. In the second scene, HD videos have better performances as measured by the brightness constancy violation ratio, but do not show clear advantages over regular videos in terms of spatial smoothness violation ratio, suggesting that vehicles' motions between some frame pairs are too large that break the small displacement assumption of optical flow algorithm. It is apparent that the quantitative violation measures agree with the visual examinations.



(a) HD video. (b) Regular video.

(c) HD vertical flow. (d) Regular vertical flow.

Figure 1. Optical flows of the "moving hand" scene.



(a) HD video. (b) Regular video.

(c) HD horizontal flow. (d) Regular horizontal flow.

Figure 2. Optical flows of the "intersection traffic" scene.

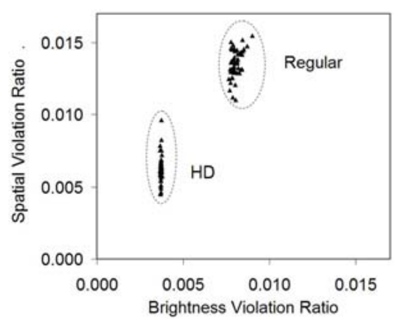


Figure 3. Violation ratios of the "moving hand" scene.

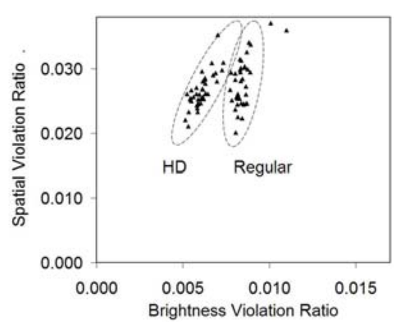


Figure 4. Violation ratios of "intersection traffic" scene.

4.3. Motion Speed

To have a better understanding of how an object's motion speed affects the flow results, we computed a series of violation ratios of a spinning fan. The fan's motion speed is measured by its rotation degrees per second. The average results of four sub-sequences are plotted in Figure 5 and Figure 6. Several observations can be made: (i) HD videos show better performances consistently as indicated by their lower violation ratios, regardless of the motion speed; (ii) a quasi-monotonic relationship exists between the violation ratios and the motion speed. In other words, the quality of flow data deteriorates as objects' motion speed increases; (iii) violation ratios stabilize after certain speed thresholds, though flow data become unusable far before the motion speed reaches those thresholds; (iv) the spatial smoothness constraint is more sensitive to motion changes than the brightness constancy assumption. It should be stressed that similar observations have been obtained from scenes such as rotating water, people skating and objects swinging in pendulum motion, etc.

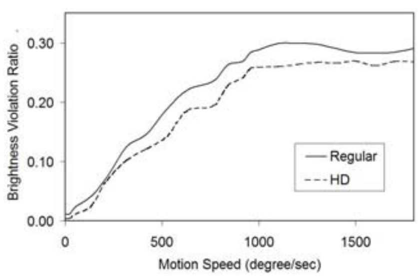


Figure 5. The relationship between the fan's motion speed and the brightness constancy violation.

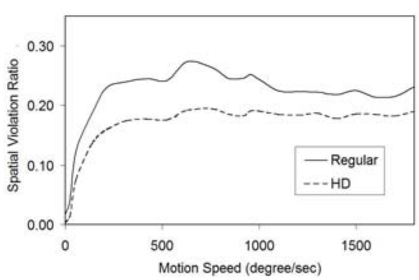


Figure 6. The relationship between the fan's motion speed and the spatial smoothness violation.

5. Conclusions

High quality motion data extracted from videos is essential for a wide variety of applications. This study investigated the feasibility of improving the quality of optical flow using high definition videos. Experiments were carried out by comparing the outputs of a robust optical flow algorithm using HD and regular videos of 35 scenes. Both visual examinations and quantitative analysis of two violation ratios suggested that HD videos outperformed regular videos, though cautions should be taken when dealing with more complex motion scenes.

6. Acknowledgements

This work was supported in part by the NSF-REU grant No. 0453463 at Univ. of South Florida and the URC grant 08-#8 at Youngstown State University. The authors thank Dr. Labrador for a summer REU support.

7. References

[1] L. Dougherty, J. C. Asmuth, A. S. Blom, L. Axel, and R. Kumar, "Validation of an optical flow method for tag displacement estimation", *IEEE Trans. on Medical Imaging,* 18(4), 1999, pp. 359-363.

[2] G. N. DeSouza and A. C. Kak, "Vision for mobile robot navigation: a survey", *IEEE Trans. on Pattern Analysis and Machine Intel.*, 24(2), 2002, pp. 237-267.

[3] D. Decarlo and D. Metaxas, "Optical flow constraints on deformable models with applications to face tracking", *Inter. Journal of Computer Vision*, 38(2), 2000, pp. 99-127.

[4] I. Essa and A. Pentland, "Coding, analysis, interpretation and recognition of facial expressions", *IEEE Trans. on Pattern Anal. and Machine Intel.*, 19(7), 1997, pp. 757-763.

[5] V. Manohar, D. B. Goldgof, S. Sarkar, and Y. Zhang, "Facial strain pattern as a soft forensic evidence", *The 8th IEEE Workshop on Appl. of Comp. Vision*, 2007, pp. 42-45.

[6] P. J. Phillips, P. J. Flynn, W. T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. J. Worek, "Overview of the face recognition grand challenge", *Computer Vision and Pattern Recog.,* 2005, pp. 947-954.

[7] J. L. Barron, D. J. Fleet, and S. Beauchemin, "Performance of optical flow techniques", *International Journal of Computer Vision*, 12(1), 1994, pp. 43-77.

[8] S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow", *International Conference on Computer Vision*, Rio de Janeiro, Brazil, 2007.

[9] B. K. P. Horn and B. G. Schunck, "Determining optical flow", *AI Memo 572*, Massachusetts Institute of Tech., 1980.

[10] M. J. Black, P. and Anandan, "The robust estimation of multiple motions: parametric and piecewise-smooth flow fields", *Computer Vision and Image Understanding*, 63(1), 1996, pp. 75-104.

[11] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and Bayesian restoration of images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(5), 1991, pp. 401-412.

[12] http://www.cs.brown.edu/~black/