

Estimate Large Motions using Reliability-based Dynamic Programming

Minglun Gong¹ and Yee-Hong Yang²

¹Laurentian University, Department of Math & Computer Science, Sudbury, ON, Canada

²University of Alberta, Department of Computing Science, Edmonton, AB, Canada

Abstract

Detecting and estimating motions of fast moving objects has many important applications. However, most existing motion estimation techniques have difficulties in handling large motions in the scene. In this paper, the reliability-based dynamic programming technique proposed by Gong and Yang is extended and applied to large motion estimation problem. Compared with the Gong and Yang approach, the extended algorithm removes the constant penalty assumption and also explicitly enforces the inter-scanline consistency constraint. The experimental results indicate that the new algorithm can effectively estimate velocities for fast moving objects. The algorithm can also be configured to produce sparse but reliable flow fields.

Keywords: Motion estimation, Optical flow, Dynamic programming, Reliability-based Dynamic Programming.

1 Introduction

Motion estimation has many important applications, such as structure from motion and video compression. Previous works in motion estimation are nicely surveyed by Barron et al. [2]. As they suggested, different techniques can be classified into four categories: gradient-based, matching-based, energy-based, and phase-based approaches.

When fast moving objects exist in the scene, being able to detect and estimate their motions is certainly important. However, most of the existing approaches have difficulty in handling large motions. For example, gradient-based approaches rely on the calculation of spatiotemporal derivatives. It is almost impossible to accurately estimate temporal derivatives for fast moving objects, since we do not have dense enough samples in the temporal domain. As well, large motions introduce challenges to matching-based approaches. Basically, the searching range has to be increased, which will increase the possibility of false matches.

The motivation of this paper is, therefore, to propose a reliable matching-based approach that can effectively estimate large motions. Since our approach is based on the reliability-based dynamic programming algorithm [3], we call it the reliability-based large motion estimation algorithm, or RME for short.

1.1 Related work

Dynamic programming (DP) is an efficient technique for solving certain optimization problems. Several approaches have applied DP techniques to the motion

estimation problems [7, 9]. In the orthogonal DP approach [7], the matching is performed alternatively on horizontal and vertical image strips. The spacing and width of the strips are gradually reduced to refine the matching results. In Sun's approach [9], the DP optimization is conducted for different scanlines separately. The best path found is limited to pixel-level accuracy, and an additional surface fitting process is used to provide sub-pixel accuracy. Both approaches are able to generate smooth flow fields. However, neither approach tries to solve the large motion estimation problem.

The reliability-based DP algorithm is first proposed by Gong and Yang for solving stereo vision problems [3]. Instead of tracing the best path only, their approach simultaneously traces the two best paths. The cost difference between these two paths is used to evaluate the reliability of solutions on the best path. Solutions with low reliability values are filtered out.

However, Gong and Yang's approach assumes a constant penalty term is applied whenever neighboring disparity values are different; no matter how large is the difference. This assumption is suitable for stereo vision application only, but not so for motion estimation. This is because, in motion estimation, we should treat large velocity changes and small velocity changes differently. In addition, their approach does not explicitly enforce inter-scanline consistency. The well-known "horizontal streaks" problem of DP algorithms is alleviated through enforcing the strong and weak consistency constraints, which do not hold for motion estimation.

The presented RME algorithm is an extension of the original reliability-based DP algorithm and can be applied to large motion estimation. Different from the original approach, the new algorithm does not require the constant penalty assumption. Furthermore, the inter-scanline consistency is also explicitly enforced through performing DP along both horizontal and vertical scanlines and then cross-validating the results.

2 The RME algorithm

Figure 1 shows the overall structure of the RME algorithm. Like other matching-based approaches, the algorithm starts with the matching cost calculation process, which generates a 4D matching cost volume. Then, the algorithm undergoes an iterative process. During the process, the reliability threshold t is fixed to a constant value, while the discontinuity cost λ is increased each time. Reliable velocities may be assigned to some pixels during a given iteration. These pixels will be treated as ground

control points (GCPs) in the following iterations. As a result, the number of pixels with known velocities will always increase.

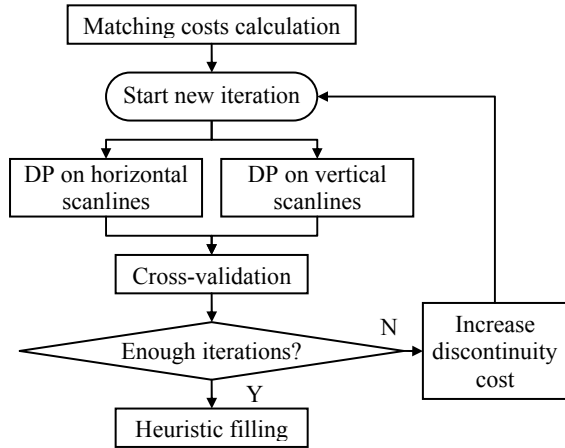


Figure 1: Flowchart for the RME algorithm.

Within each iteration, the reliability-based DP is used to search for candidate velocities with reliability values larger than t . The search is conducted along both horizontal and vertical scanlines. The solutions found under different searching directions are cross-validated. That is, a candidate velocity is assigned to a pixel *iff* it is selected in both searching processes.

Performing DP along both directions involves some redundant computation. However, it helps to produce more accurate results since:

- Potential mismatches can be detected and removed. Even though the reliability-based DP can filter out many mismatches through reliability thresholding, it certainly cannot remove all of them. Cross-validating between two solutions can further remove possible mismatches.
- The inter-scanline consistency is explicitly enforced. Most DP-based algorithms suffer from the “horizontal streak” problem since the solution for different scanlines are calculated independently. Obviously, calculating solutions along both horizontal and vertical scanlines and then cross-validating between the two solutions can effectively remove these streaks.

Depending on how dense or how accurate the flow field is needed, the user can choose the number of iterations to use and the rate at which the value of λ to grow. If a dense flow field is needed, an additional heuristic filling process can be used to assign velocities to pixels without reliable solutions. In the RME algorithm, the median filter is applied to the horizontal and vertical velocity channel separately.

In the following subsections, we discuss the matching cost calculation process and the extended reliability-based DP algorithm in detail.

2.1 Matching cost calculation

The matching cost calculation process computes the costs of assigning different candidate velocities to different

pixels. In order to accommodate for potential large velocity changes at adjacent frames, the RME algorithm uses only two frames: the central frame at time k and a reference frame at time $k-1$. The sum of absolute differences (SAD) is used in this paper for efficiency consideration. Other dissimilarity/similarity measures, such as normalized cross correlation and sum of squared differences, can also be used under the same framework. Using SAD, the matching cost of assigning velocity \mathbf{d} to pixel (p,q) is calculated by:

$$C(p, q, \mathbf{d}) = \sum_{-w \leq i, j \leq w} |I(p+i, q+j) - I^1(p+i-\mathbf{d}_x, q+j-\mathbf{d}_y)|$$

where, w controls the window size $(2w+1)$ used for matching. $I(x,y)$ and $I^1(x,y)$ are the intensities of pixels at (x,y) in the central and reference frame, respectively.

2.2 Reliability-based DP for motion estimation

Most DP algorithms proceed as follows: The accumulated costs for all the candidates at pixel $p-1$ are calculated first, and then they are used to calculate the accumulated costs for candidates at pixel p . After we have the accumulated costs for all the candidates at the last pixel in the scanline, we can find the best path through tracing back from the candidate that gives the smallest accumulated cost.

In Gong and Yang’s approach, under the constant penalty assumption, at most two possible paths need to be considered when searching for the best path that starts from a given candidate. Hence, the accumulated cost \mathbf{S} for candidate (u,v) at pixel p is calculated by:

$$\begin{cases} \mathbf{S}(0, u, v) = C(0, u, v) \\ \mathbf{S}(p, u, v) = C(p, u, v) + \min[\mathbf{S}(p-1, \mathbf{b}[p-1]) + \lambda, \mathbf{S}(p-1, u, v)] \end{cases}$$

where $\mathbf{b}[p-1]$ is the candidate that gives the smallest accumulated cost at pixel $p-1$.

In order to treat large and small velocity changes differently, in the RME algorithm, the penalty is defined as a function of the velocity difference between two candidates (u,v) and (s,t) :

$$\Lambda(u, v, s, t) = \lambda(|u-s| + |v-t|)$$

As a result, when searching for the best path that starts from candidate (u,v) at pixel p , we have to consider all candidates at pixel $p-1$. That is:

$$\mathbf{S}(p, u, v) = C(p, u, v) + \min_{(s,t)} [\mathbf{S}(p-1, s, t) + \Lambda(u, v, s, t)]$$

The difference between the reliability-based DP and the traditional DP is that, instead of tracing the best path only, the former approach simultaneously traces the two best paths. The cost difference between these two paths is defined to be the approximate reliability of solutions on the best path. Using a threshold, solutions with approximate reliabilities lower than the threshold are filtered out so that only reliable solutions remain.

As we trace the two best paths, it is possible that the two paths merge into one. Under the constant penalty

assumption, it can be proven that when two paths merge at pixel p , they will merge at $\mathbf{b}[p]$. Hence, in Gong and Yang’s approach, a new alternative path can start from candidate $\mathbf{b}'[p]$, which has the second-to-the-smallest accumulated cost at pixel p .

However, without the constant penalty assumption, it is possible that the two best paths merge at a candidate other than $\mathbf{b}[p]$. Hence, we have to use a different strategy to select where the alternative path should start.

Whenever the two paths merge, the RME algorithm will first determine whether or not the two paths merge at $\mathbf{b}[p]$. If they do, the alternative path will start from $\mathbf{b}'[p]$, and the approximate reliability of the corresponding solution will be $r = S[p, \mathbf{b}'[p]] - S[p, \mathbf{b}[p]]$. Otherwise, the alternative path will start from $\mathbf{b}[p]$, and the corresponding solution will have a negative approximate reliability value. The actual reliability value is therefore unimportant since we always

set reliability threshold ≥ 0 .

3 Experimental results

Different sequences are used to test the new RME algorithm. However, due to space limitations, only the “tennis” sequence (downloaded from <http://sampl.eng.ohio-state.edu/~sampl/database.htm>) is discussed in detail. In our experiment the 40th frame is used as the central frame.

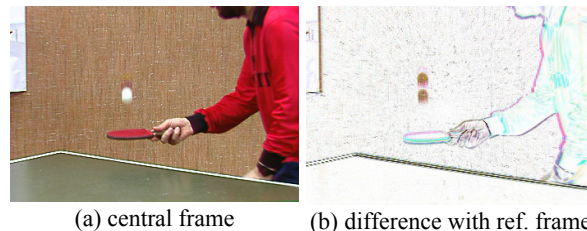


Figure 2: The “tennis” sequence.

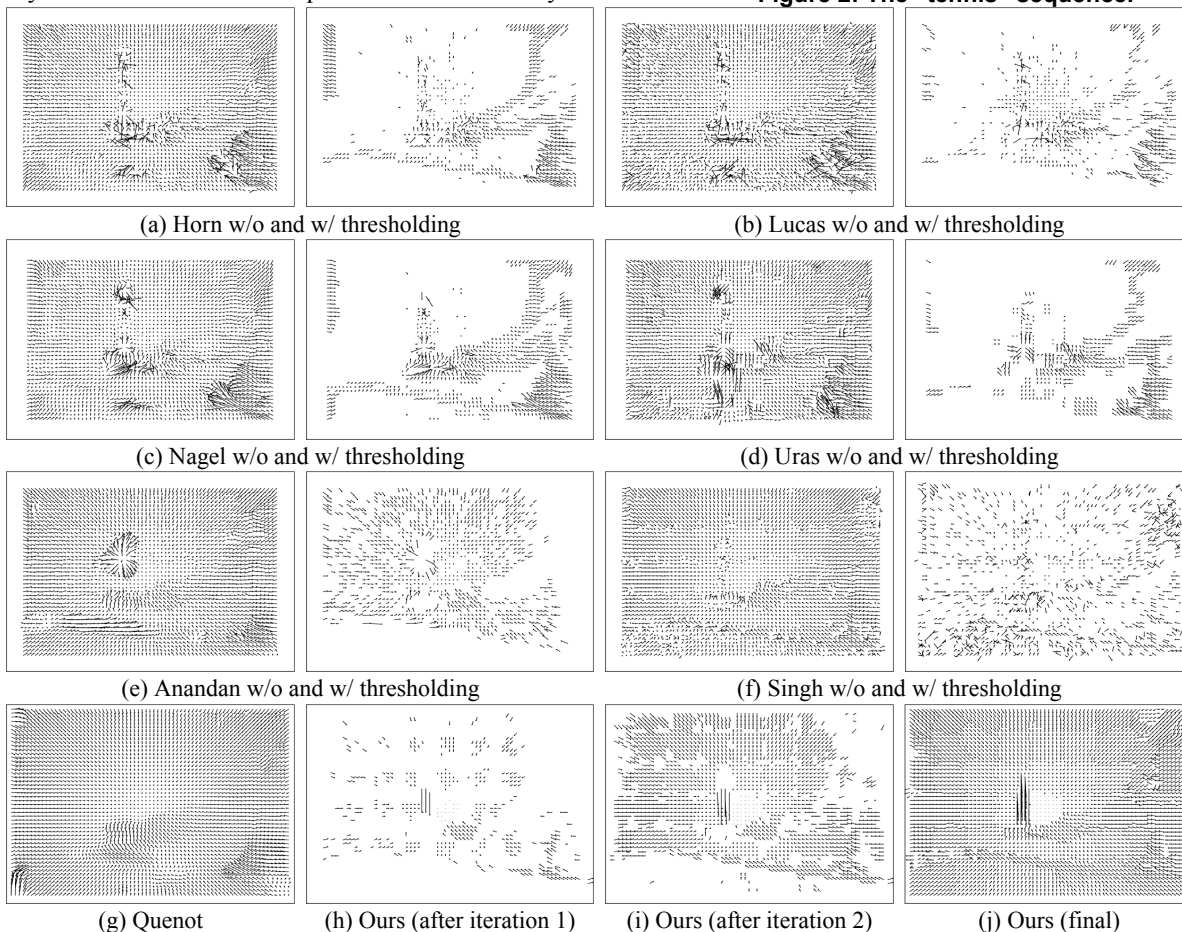


Figure 3: Flow fields estimated using different algorithms.

Figure 2 shows the central frame and the difference between the center frame and the reference frame (the darker the color, the larger the difference). We can see that that this dataset is challenging due to the following facts:

- Large velocity range: the displacement of the ping-pong ball in adjacent frames is over 20 pixels.

- Complex motions: the motion is caused by the combinations of the zoom camera, the falling ping-pong ball, and the moving player.
- Presence of non-Lambertian surface: the table tennis table has a glossy surface, and therefore, the brightness constancy does not hold.

Figure 3 shows the optical flow results estimated using

different algorithms. The implementations for algorithms proposed by Horn [4], Lucas [5], Nagel [6], Uras [10], Anandan [1], and Singh [8] were downloaded from <ftp://csd.uwo.ca/pub/vision/>. For these algorithms, both the dense flow field and the flow field filtered using a confidence threshold are shown. The implementation for Quenot's algorithm [7] was downloaded from <ftp://ftp.limsi.fr/pub/quenot/>.

It is interesting to observe from the filtered flow fields that, gradient-based approaches [4, 6, 10] have more confidence in areas around the objects' boundaries, while matching-based approaches [1] are more confident in areas with rich textures, such as the background wall. However, all the above algorithms failed to estimate reasonable motions for the fast moving ping-pong ball.

The results of our approach are generated by running the RME algorithm for three iterations. During the iterative process, the reliability threshold t is fixed at 5, and the discontinuity cost λ steps through 0, 2, and 4.

As shown in Figure 3(h), after the first iteration, only the most reliable candidates are selected. However, even though the estimated optical flow is very sparse, the movement of distinct features, such as the ping-pong ball and the player's hand, are correctly estimated. After we increase λ to 2 and run for another iteration, the optical flow generated become quite dense. The areas left unassigned include the region occluded by the ping-pong ball in the reference frame and the glossy surface of the table tennis table.

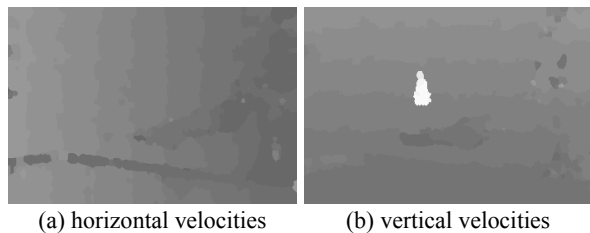


Figure 4: Horizontal and vertical velocities.

As shown in Figure 3(j), all the missing velocities can be filled up through running the program for another iteration with $\lambda=4$, followed by the heuristic filling process. To better illustrate the details of the flow field generated, the horizontal and vertical velocities are also encoded as grayscale images and are shown in Figure 4. The results show that reasonable velocities are estimated for almost the whole scene, such as the background, the ping-pong ball, the player's bat and hand, etc. However, obvious errors can still be found in some areas, such as the area occluded by the ping-pong ball in the reference frame. In addition, the result is also limited to pixel level accuracy.

4 Conclusions

In this paper, the reliability-based DP algorithm proposed by Gong and Yang is extended and applied to large motion estimation problem. Compared with the

original approach, the new algorithm makes the following two major extensions:

- The constant penalty assumption is removed, and accordingly, a new strategy is proposed to select where the best alternative path should start. As a result, we have the freedom to treat large velocity changes and small velocity changes differently.
- The inter-scanline consistency is explicitly enforced though performing reliability-based DP along both horizontal and vertical scanlines. The results generated under the two settings are cross-validated so that potential mismatches and horizontal streaks can be eliminated.

It is noteworthy that our approach is different from the orthogonal DP algorithm [7], even though we both involve matching along horizontal and vertical directions. In the orthogonal DP, the objective is to transform the 2D alignment problem into a sequence of 1D alignment problems; while in our approach, the objective is to detect potential mismatches and to enforce inter-scanline consistency.

The experimental results show that the new RME algorithm can effectively estimate motions for fast moving objects in the scene. The algorithm can also be configured to produce sparse but reliable flow fields. In these sparse flow fields, areas that are error prone, such as those invisible in the reference image, will be left unassigned.

Currently, the RME algorithm can only provide motions with up to pixel-level accuracy. Its extension to provide sub-pixel accuracy will be investigated in the future.

References

- [1] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *IJCV*. 2(3):283-310. 1989.
- [2] J. L. Barron, D. J. Fleet, & S. S. Beauchemin. Performance of optical flow techniques. *IJCV*. 12(1):43-77. 1994.
- [3] M. Gong & Y.-H. Yang. Fast stereo matching using reliability-based dynamic programming and consistency constraints. *Proc. of ICCV*. pp. 610-617. 2003.
- [4] B. K. Horn & B. G. Schunck. Determining optical flow. *Artificial Intelligence*. 17(1-3):185-203. 1981.
- [5] B. D. Lucas & T. Kanade. An iterative image registration technique with an application to stereo vision. *Proc. of IJCAI*. pp. 674-679. 1981.
- [6] H.-H. Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*. 33(3):299-324. 1987.
- [7] G. M. Quenot. Computation of optical flow using dynamic programming. *Proc. of IAPR Workshop on Machine Vision Applications*. pp. 249-252. 1996.
- [8] A. Singh. An estimation-theoretic framework for image-flow computation. *Proc. of ICCV*. pp. 168-177. 1990.
- [9] C. Sun. Fast optical flow using 3D shortest path techniques. *Image and Vision Computing*. 20(13-14):981-991. 2002.
- [10] S. Uras, F. Girosi, A. Verri, & V. Torre. A computational approach to motion perception. *Biological Cybernetics*. 60(1):79-87. 1988.