A Comprehensive Evaluation on Non-deterministic Motion Estimation
Changzhu Wu, Qing Wang
School of Computer Science and Engineering,
Northwestern Polytechnical University, Xian 710072, P. R. China
qwang@nwpu.edu.cn

Abstract
When computing optical flow with region-based matching, very few of them can be reliably obtained, especially for the high-contrast areas or those with little texture. Instead of using a single pixel from the reference frame, non-deterministic motion utilizes multiple pixels within a neighborhood to represent the corresponding pixel in the current frame. Although remarkable improvement has been made with this method, the weight associated to each reference pixel is quite sensitive to the selection of its standard deviation. To address this issue, a dual probability is presented in this paper. Intuitively, it enhances those weights of pixels that are more similar to its counterpart in the current frame, while suppressing the rest of them. Experimental results show that the proposed method is effective to deal with intense motion and occlusion, especially in the case of reducing the adverse impact of noise.

1. Introduction
Motion estimation plays an important role in many computer vision and image processing tasks, and has been one of the most active hotspots in the research area. Although a great number of approaches have been proposed in the literatures over last decades, extra efforts still need to be made to deal with more intricate cases such as occlusion and acute motion frequently encountered in motion analysis.

Existed methods can mainly be divided into three groups: gradient based approaches [1,2], block matching based approaches [3,4,6,7] and wavelet domain based approaches [5,12,13]. Although the first group of methods is remarkable for their simplicity and computational efficiency, they cannot be used reliably for high-contrast areas or those with little or no texture. Moreover, they can hardly be used to deal with occlusion, especially in the presence of noise. Wavelet domain based methods may obtain optical flows with great accuracy, but the algorithms are usually more complicated.

Moreover, sub-pixel accuracy is required in some typical application such as AVS and super-resolution reconstruction, etc. The accuracy of block matching based approach can be improved by either interpolation filtering [6] or quadratic fitting [8]. However, block matching with sub-pixel accuracy is extremely computationally intensive. Besides, non-deterministic motion is proposed in [8] with application in super resolution reconstruction (SRR). All the pixels within a neighborhood are employed to represent their corresponding pixels in the current frame. Although the computational burden for calculating the associated weights seems prohibitive, it can be substantially reduced as discussed in [14].

Recently, much attention has been drawn to tackle the loss of information under different circumstances [11-14]. Occluded regions were directly detected either to prevent over-smoothing of motion [10] or to infer motion vectors with bilateral diffusion [11]. In [12], LBS (Low Band Shift) was combined with symmetric padding to minimize the effects of boundary discontinuities. However, all the above issues can be properly addressed when the successive frame (called backward reference frame) is considered as shown in [13] in the cases of acute motion, and occlusion.

In the paper, we estimate the current frame from both the forward and backward reference frames using non-deterministic motion estimation, which is extended by introducing the dual probability. The main advantages of the work are: Firstly, non-deterministic motion can improve the accuracy to sub-pixel and effectively deal with noise; secondly, performance can be further improved together with robustness to the selection of parameter; and thirdly, frequently occurred issues, such as occlusion and acute motion, can be properly dealt with ascribed to the introduction of the backward reference frame.

The rest of the paper is organized as follows: Section 2 describes the concept of non-deterministic motion estimation. In Section 3, some concerns for non-deterministic motion estimation are discussed and then dual probability is presented in motion estimation. Experimental results and discussions are shown in Section 4 and conclusion is drawn in Section 5.
2. Non-deterministic motion estimation

Typical optical flow estimation with region-based matching is confronted with the following two issues: Firstly, optical flow of most locations cannot be reliably obtained, especially for those with little or no texture; Secondly, optical flow cannot be well dealt with in the case of noisy images.

To overcome the above-mentioned problems, non-deterministic motion can be employed by using multiple optical flows despite of the single one obtained by typical optical flow estimation, and each location within the neighborhood is associated with a weight, i.e.,

$$W(x, y; dx, dy) = \exp \left( -\frac{D(x, y; dx, dy)}{2\delta^2} \right)$$

with

$$D(x, y; dx, dy) = \sum_{s,t} (I_{ref}(x+s, y+t) - I_{ref}(x+dx(x, y)+s, y+dy(x, y)+t))^2$$

where $D(x, y; dx, dy)$ is Euclidean distance of two image patches. $I_{ref}(x, y)$ and $I_{ref}(x, y)$ are the pixels of the current and reference frames at $(x, y)$, $-W \leq dx, dy \leq W$ with $[-W, W]$ being search ranges; $(dx, dy)$ is the displacement from $(x, y)$, and $-P \leq s, t \leq P$ with $(2P+1)$ being the sizes of image patches.

Subjective quality of the two methods is shown in Figure 1, from which we can see that noise can be properly dealt with by non-deterministic motion.

![Figure 1](image_url)

(a) Original Image of the 69th frame of Foreman; (b) Motion-compensated image with deterministic motion (PSNR=31.3240), and (c) Motion-compensated image from non-deterministic motion (search window = 7x7 and STD = 0.05, PSNR=34.9866).

3. Dual-weighted non-deterministic motion estimation

3.1 Analysis on sensitivity of standard deviation

Ascribed to the introduction of spatial redundancy, non-deterministic motion can overcome the adverse factors, e.g. noise. However, parameter $\delta$ needs to be delicately designed in order to obtain excellent performance as shown in Figure 2. Smaller value will cause most of the weights close to zero, thus only very few neighbor pixels remain in effect; on the other hand, larger value will lead to over-smoothed result.

Therefore, extra efforts need to be made to reduce the sensitivity of the selection of standard deviation $\delta$.

![Figure 2](image_url)

Figure2 mean PSNR of Foreman sequence with different $\delta$.

![Figure 3](image_url)

Figure 3 mean PSNR of Foreman sequence under varying $\delta$ and $C$.

3.2 Dual probability

For non-deterministic motion, each pixel in the reference frame is associated with multiple pixels in the current frame, called dual index, which is denoted in Equation (3). As the pixels between the reference and current frames form a many-to-many relationship, a dual probability may be introduced for each pixel in the reference frame. Intuitively, if a pixel is more close to its counterpart in the current frame, it should have greater impact on it than the rest of them.

Thus for the $q^{th}$ element $(k,l,s,t)$ in $D(k,l,m,n)$, its dual probability can be defined as the following way,

$$DW(k,l,s,t) = \exp \left( -\frac{D(k,l,s,t)}{2\delta^2} \right)$$

For the sake of simplicity, $\delta_2$ is set to be a function of $\delta_1$, i.e., $\delta_2 = f(\delta_1)$. Here $\delta_1$ is the standard deviation $\delta$ in equation (1). In order to reduce the sensitivity of weight to the selection of the standard deviations, $\delta_2$ should be inversely proportional to $\delta_1$.

Therefore, $f(\cdot)$ is defined in equation (5).

$$f(x) = C/x$$

where $C$ is a positive constant, which controls the smoothness of the dual probability.

From Figure 2 we can see that the application of dual probability can greatly reduce the sensitivity of the parameters. Furthermore, Figure 3 shows the fact...
that, although selection of $C$ may lead to the loss of performance, it still outperforms the traditional non-deterministic motion estimation by at least 3dB.

3.3 Computational flow of the proposed method

According to the above-mentioned analysis, we can conclude the computation process of the dual-weighted non-deterministic motion estimation as follows:

**Step 1. Motion pre-estimation.** For the given current frame $I_{cur}$, block matching algorithm is applied to the forward- and backward-reference frames $I_{R_f}$ and $I_{R_b}$ to obtain the initial base motion vectors, respectively.

**Step 2. Computation of dual weight.** For each pixel in the current frame, we carry out the following sub-steps,

**Sub-step 2.1. Computation of Euclidean distance.** Given the search window of $(2P+1)(2P+1)$ around $I_{R_f}(x+bx(x,y),y+by(x,y))$ in the forward reference frame, the corresponding Euclidean distance between the two image patches within the window is calculated using equation (2), here $(bx(x,y),by(x,y))$ is the pre-estimated basis motion vector at point $(x,y)$.

**Sub-step 2.2. Mark dual pixels.** For each pixel at $(x+bx(x,y)+m,y+by(x,y)+n)$ within the search window in the reference frame, mark the pixel at $(x,y)$ in the current frame as its **dual index**, abbreviated as $DI$.

**Sub-step 2.3 Computation of dual weight.** First, the forward weight is computed using equation (1), and denoted as $FW_{R_f}$.

Then, with the $DI$ obtained from sub-step 2.2, the corresponding dual probability of pixel to the corresponding pixel in the current frame patches can be easily extracted by equation (4), and abbreviated as $BW_{R_f}$.

Finally, the dual weight can now be calculated as

$$DW_{R_f}(i,j;m,n) = FW_{R_f}(i,j;m,n) \cdot BW_{R_f}(i,j;m,n)$$

**Sub-step 2.4. Backward motion estimation.** For the backward reference frame, performs step 2.1-2.3 to obtain backward dual weight with the current frame, and abbreviated as $DW_{R_b}$.

**Step 3. Motion compensation.** With all the weights calculated in the last step, motion compensation can now be performed by summing all weighted pixels within the neighborhood (See Equation (7)).

$$I_{wc}(i,j) = \frac{\sum_{(r,P,R_f)} \sum_{m,n \in (-P,P)} DW_{R_f}(i,j;m,n) \cdot I_{r}(i+bx(i,j)+m,j+by(i,j)+n)}{\sum_{(r,P,R_b)} \sum_{m,n \in (-P,P)} DW_{R_b}(i,j;m,n)}$$

where $I_{wc}(i,j)$ is $(i,j)$ pixel of the motion-compensated image. $I_{r,r} \in \{R_f,R_b\}$ are the reference images. $(bx(i,j),by(i,j))$ is the pre-estimated base motion vector of the corresponding pixel, and $m$ and $n$ are the offsets in the horizontal and vertical directions, respectively.

4. Experimental results and discussion

In order to verify the effectiveness of the proposed dual-weighted non-deterministic motion estimation method (DW-NME), several experiments were performed together with normal non-deterministic based approach (NME) and dual-tree complex wavelet transformation based method (DTCWT). Four standard video sequences were used for performance evaluation, including Football (acute motion 90 frames), Foreman sequence (smooth motion, 300 frames), Vectra with occlusion (142 frames) and Coastguard sequence with complicated motion (300 frames) (all in CIF format).

In experiments, some important parameters were preset as follows: the search window is $7 \times 7$ with image patch size to be $3 \times 3$, the STD is set to 0.20 and $C=0.025$. For the DTCWT method, the parameters are set as described in [13]. Besides, in order to deal with noise, hard-threshold filtering is applied before reconstructing the motion-compensated image in the wavelet domain.

Figure 4(a) shows the PSNR of Foreman sequence, which is degraded by additive white Gaussian noise with STD=8.0. From the figure we can see that, comparing to DTCWT approach, non-deterministic methods obviously improve the overall performance. However, it is worth to note that, when dealing with more complicated scenes (200-300 frames), normal non-deterministic method degrades seriously and the proposed method shows more robustness ascribed to the introduction of the dual probability.

Next, the PSNR of the more challenging Football sequence is shown in Figure 4(b). The whole sequence can mainly be divided into three stages: several moving objects with acute motion and occlusion (1-30 frames), no moving object with steady background (31-40 frames) and one or two moving objects under more complicated background (42-90 frames). The result shows that the proposed method outperforms the other two methods up to 4dB on average under any of the circumstances.

To illustrate the robustness of the proposed method to occlusion, the PSNR chart for the Vectra sequence is also shown in Figure 4(c). The result shows that our approach not only can properly deal with the loss of information, but also gives more stable performance.
over the whole sequence. Besides, Figure 4(d) gives the PSNR for the Coastguard sequence. From the chart we can see that, the proposed method becomes more stable when confronting with complicated motion (50-90 frames).

Finally, in order to illustrate how noise affects the performances of different methods, mean PSNR of Foreman sequence under varying noise level is shown in Figure 5. The DTCWT method is quite sensitive in the presence of noise, whilst both non-deterministic methods demonstrate great robustness when the STD of noise is smaller than 6.0. However, with increasing STD, the performance of NME approach degrades severely, while remarkable results can still be obtained by the proposed DW-NME method.

5. Conclusion

Non-deterministic motion can effectively reduce the impact of noise on the optical flow estimation. In this paper, dual probability is presented to improve the robustness of non-deterministic motion to its standard deviation. Meanwhile, the backward reference frame is considered to further improve the motion vector accuracy in the case of occlusion and acute motion. Experimental results have shown that, comparing to DTCWT based approach, PSNR can be improved nearly 14% (4dB) on average with non-deterministic method in the cases of acute motion and 3dB for the case of steady sequences. Moreover, performance and robustness can be further significantly improved by the introduction of dual probability.

Considering the importance of motion estimation in super resolution reconstruction, the future work is the application of the proposed non-deterministic motion estimation in resolution restoration to further validate its effectiveness.

6. Acknowledgement

The work was supported by Program for New Century Excellent Talents in University (NCET-06-0882) and Aviation Research Fund (2008ZC53017), P. R. China.

7. References


2336