MOTION ESTIMATION USING ADAPTIVE BLOCKSIZE OBSERVATION MODEL AND EFFICIENT MULTISCALE REGULARIZATION

Stephanus Suryadarma Tandjung, Teddy Surya Gunawan, Chong Man Nang

School of Applied Science
Nanyang Technological University
Singapore 639798
[steph, teddy, cmn]@sentosa.sas.ntu.edu.sg

ABSTRACT

Bayesian motion estimation requires two pdf models: observation model and motion field (prior) model. The optimization process for this method uses sequential approach, e.g. simulated annealing. This paper proposes adaptive blocksize observation model and multiscale regularization for the prior model and the optimization process. The purposes are to increase the speed and to improve the result. The proposed framework can initialize the bayesian method. The result in this paper shows one of the possibility of its usage. Many strategies can be derived from this framework to work for itself or to support the Markov random field modeling for motion estimation.

1. INTRODUCTION

This paper introduces a strategy to solve motion estimation problems, such as occlusion and aperture problems. The strategy involves the distinction of type of processes/stages in estimation algorithm, a sensing stage and a regularization stage. In bayesian method, sensing stage is called observation or likelihood model and the regularizing stage is called prior model. The designers of such system in Fig. 1 can put any types of sensing algorithm and any types of regularization algorithm to satisfy their objectivity.

Adaptive blocksize observation model [1], as a sensing stage, analyses the unique match from a pixel or a block of pixels of a frame to the subsequence frame. Then, the result is regularized by an efficient multiscale regularization framework [2].

The purposes of the above strategy are to increase the speed of optimization processes [3] and to achieve a more global result [4] than sequential optimization processes. The result of this approach can be followed by Bayesian method [5] if the process is found to be useful.

The regularization stage is able to receive several types of information that are produced by sensing algorithm. The
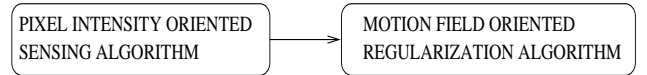
Fig. 1. The motion estimation stages
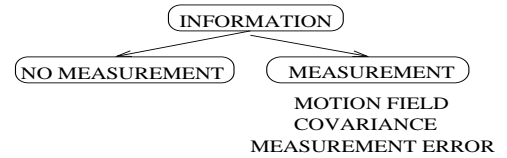
Fig. 2. The types of information

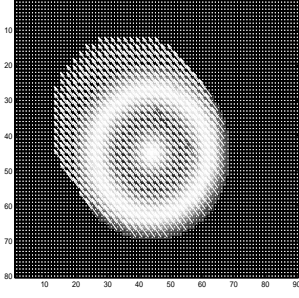informations are motion field, the measurement error and covariance matrix.

Another useful information is "no measurement". The experiment for this type of information is given in the last section.

2. ADAPTIVE MATCHING ALGORITHM

A suitable initialization motion field can reduce the computational cost needed by the stochastic process in MRF based motion estimation. The initialization process is commonly based on the intensity assumption, known as the implicit constraint. An objective in this model is to add the implicit constraint into the observation model. This model yields a motion vector that has been tested as the only best candidate for a given point. Equation 1 is the observation model on the left term and a block model on the right term. Both models can have a minimum energy value according to d, the displacement vector and h, the block size. λ are weighting functions, g is the image, l is line field.

$$\min_{d,h} \left\{ \lambda_g U_g(g_t|d,l,h,g_{t-1}) + \lambda_h U_h(h|g_{t-1}) \right\} \quad (1)$$

Equation 2 is the functions of the model.

Fig. 3. The result of the adaptive blocksize observation model

Fig. 4. A smaller blocksize threshold reduces false motion vectors in occlusion sites

$$\min_{d,h} \left\{ \lambda_g \sum_{x=1}^{h} \frac{[g_t(x) - g_{t-1}(x + d(x))]^2}{2\sigma^2} + \lambda_g \Theta(h) \right\}$$
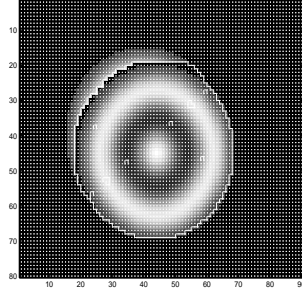(2)

where h denotes the order of MRF neighborhood system and σ^2 denotes the variance of pixel intensities and x is a pixel site. The auxiliary variable for "indiscriminate texture" variable is

$$\Theta = \begin{cases} \Theta + 1; & U_g(g_t|\widehat{d}_{(i,j,k)}, l, h, g_{t-1}) - \\ & U_g(g_t|\widehat{d}_{(i,j,k-1)}, l, h, g_{t-1}) = 0 \\ 0; & else \end{cases}$$
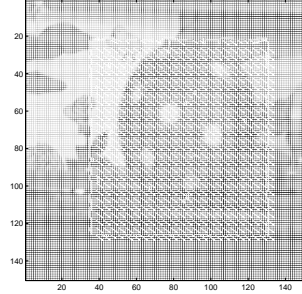(3)

This adaptive observation model is used for local motion estimation. The algorithm in this paper has a purpose to exploit the implicit information on the beginning of optimization process. The matching based motion estimation is named implicit approach. The proces is continued by the regularization technique (MRF based). In this technique, the explicit information, as the prior knowledge, is exploited to minimize the matching error on only one motion vector. Therefore, the aperture problem is solved and the occlusion locations can be detected. Fig. 3 shows the motion vectors and Fig. 4 shows the line field. From those figures, several sites in the middle of object that has different motion vectors can be identified. Fig. 5 is motion vectors result from a shifted block of a part of calendar-train sequence. This figure shows the result that cannot be achieved by a fixed blocksize [3].

3. MULTISCALE SMOOTHING ALGORITHM[6]

The motion fields on the finest resolution can be smoothed with this multiscale smoothing. For block matching, the finest resolution is not motion-vector/pixel but motion- vector/block. The motion-vector/block resolution can be processed by this algorithm. Remember that block matching

Fig. 5. The motion vector for synthetic motion from real image

employes implicit smoothing constraint for the pixels, in this case, the motion vector result can not be smoothed anymore.

The algorithm produces the best estimate of the motion estimation process given all available measurements. The measurement can consist of different measurements error or no measurement on specific points. No measurement information can help to reduce the speed of measurement if can be detected earlier. Example, the matching is done around the edge and higly texture object only, the rest is no measurement. The multiscale process:

$$x(s) = A(s)x(s\overline{\gamma}) + B(s)w(s)$$
(4)

is the process from finest scale s and propagate to higher scale $s\overline{\gamma}$ in quadtree until the lowest scale. $w(s)$ is a zero-mean unit-variance white noise process that is added to the model with a gain β. The measurement equations:

$$y(s) = C(s)x(s) + v(s)$$
(5)

where , $v(s)$ is a zero-mean white noise process with covariance $R(s)$. This covariance can be related to the matching algorithm, the value of its error or the size of the block. A further research is needed in this strategy.

To use this algorithm, it must give the states $x(s)$ at the root node: zero mean with covariance $P(0)$. To compute the prior covariances of all states at individual nodes on the tree via Lyapunov equation:

$$P(s) = A(s)P(s\overline{\gamma})A^T(s) + B(s)B^T(s) \qquad (6)$$

The core of multiscale algorithm consists of an upward estimation sweep and a downward smoothing sweep. Define the following quantities:

- $Y(s) = \{y(\sigma)|\sigma$ is a descendant of $s\}$ is the collection of measurements at all nodes below s but not including s.

- $\widehat{x}(\sigma|s) = E[x(\sigma)|\sigma \in Y_s \bigcup y(s)]$ is the best estimate of $x(s)$ given measurement at node s and all nodes below s

- $\widehat{x}(\sigma|s+) = E[x(\sigma)|\sigma \in Y_s]$ is the best estimate of $x(s)$ given measurement at all nodes below s

- $\widetilde{P}(\sigma|s) = Cov[x(\sigma) - \widehat{x}(\sigma|s)]$

- $\widetilde{P}(\sigma|s+) = Cov[x(\sigma) - \widehat{x}(\sigma|s+)]$

The upward sweep initializes at the finest level from the prior covariances: $\widehat{x}(s|s+) = 0$, $P(s|s+) = P(s)$ and the upward model

$$x(s\overline{\gamma}) = F(s)x(s) + \overline{w}(s) \qquad (7)$$

$$y(s) = C(s)x(s) + v(s) \qquad (8)$$

where $F(s) = P(s\overline{\gamma})A^T P(s)^{-1}$, $E[\overline{w}(s)\overline{w}(s)^T] = P(s\overline{\gamma}) - F(s)A(s)P(s\overline{\gamma}) = Q(s)$

The upward sweep computes the best estimate of the state at a node given all measurements at or below that node. It consists of three steps at each scale:

- *The update step*:

$$\widehat{x}(s|s) = \widehat{x}(s|s+) + K(s)[y(s) - C(s)\widehat{x}(s|s+)] \qquad (9)$$

$$P(s|s) = [I - K(s)C(s)]P(s|s+) \qquad (10)$$

$$K(s) = P(s|s+)C^T(s)[C(s)P(s|s+)C^T(s) + R(s)]^{-1} \qquad (11)$$

The updated estimate is the best estimate of $x(s)$ given all measurements at or below s.

- *The prediction step*:

$$\widehat{x}(s|s\alpha_i) = F(s\alpha_i)\widehat{x}(s\alpha_i|s\alpha_i) \qquad (12)$$

$$P(s|s\alpha_i) = F(s\alpha_i)P(s\alpha_i|s\alpha_i)F^T(s\alpha_i) + Q(s\alpha_i) \qquad (13)$$

The predicted estimate is the best estimate of $x(s)$ given all measurements at node $s\alpha_i(i = 1\ldots q)$ or below.

- *The merge step:*

$$\widehat{x}(s|s+) = P(s|s+)\sum_{i=1}^{q} P^{-1}(s|s\alpha_i)\widehat{x}(s|s\alpha_i) \qquad (14)$$

$$P(s|s+) = [(1-q)P(s)^{-1} + \sum_{i=1}^{q} P^{-1}(s|s\alpha_i)]^{-1} \qquad (15)$$

The merge step combines the predicted estimate of $x(s)$ given each of its child subtrees. The merged estimate is the best estimate of $x(s)$ given all measurements below node s.

The downward sweep computes the best estimate of the states at a node given all available measurements everywhere on the tree:

$$\widehat{x}(s|0) = \widehat{x}(s|s) + J(s)[\widehat{x}(s\overline{\gamma}|0) - \widehat{x}(s\overline{\gamma}|s)] \qquad (16)$$

$$P(s|0) = P(s|s) + J(s)[P(s\overline{\gamma}|0) - P(s\overline{\gamma}|s)]J(s) \qquad (17)$$

$$J(s) = P(s|s)F^T P^{-1}(s\overline{\gamma}|s) \qquad (18)$$

The equation 16 is the result, the estimate motion vectors.

4. MATCHING ALGORITHM-MULTISCALE REGULARIZATION INTEGRATION

Motion estimation can be divided into two parts, the sensor and the process. Other advance processes that can be put into the whole process are the intelligence interpretation or any other higher level algorithms. The sensor part in this case can be optical flow result, block matching result or adaptive matching result. The adaptive matching result can be put into the focus of this paper, because this algorithm has the ability to enlarge the blocksize. The enlarging process is recursively computed on some degree until a motion vector with a unique measurement error value is achieved. This approach can be thought as a variable size aperture for
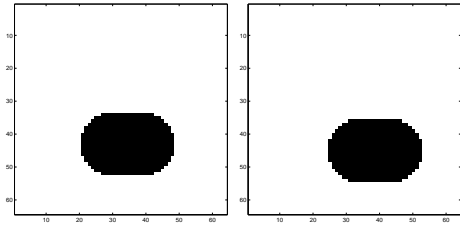
Fig. 6. The synthetic image

(a) The difference   (b) The motion vector

Fig. 7. The middle sites cannot be detected as moving pixels.

the sensor. It is also possible to put the result on different resolution on quadtree of multiscale smoothing algorithm. Meanwhile, all the measurement is put on the finest resolution.

After the adaptive matching has been executed, the most possible matching for the motion fields based on intensity is achieved. The smoothing process can follow the sequence. The strategy to separate matching and smoothing is proposed in this paper. After multiscale smoothing, the characteristic of the result is the motion field that follows Markov random fields modeling in scale.

When fixed block matching is used to supply this multiscale smoothing, the finest resolution cannot be put into the algorithm, but few level above the finest scale according to block size can be supplied. This is because the implicit smoothness has played in the development of motion field.
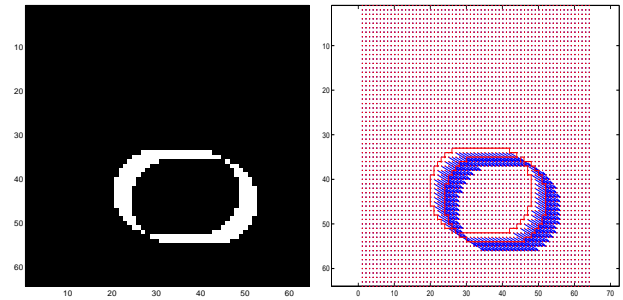
5. EXPERIMENTAL RESULTS

This experiment uses a pair of synthetic image that has no texture on the black object. The object in Fig. 6 is black ellipse and is shifted several pixels apart. Adaptive blocksize observation model can produce motion vector (Fig. 7(b)) on sites where the difference exist as illustrated in Fig. 7(a). The human perception can not agree with such a result. This framework can solve this problem by adding segmentation information to the multiscale regularization and gives "no measurement information" into the sites that have zero motion vector but belongs to the object.

The result from the first stage with the "no measurement information" on the specific sites becomes the input for the efficient multiscale regularization algorithm. Then, the result is inserted into bayesian motion estimation to make it more smooth. Figure 8 is the output of bayesian motion estimation that works with observation model, displacement model for smoothness and line model for discontinuity.
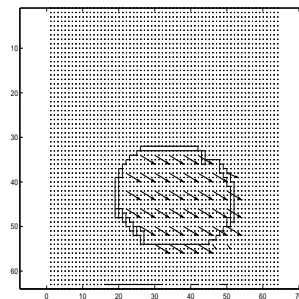
6. REFERENCES

[1] Stephanus S. T., M. N. Chong, and K. K. Wong, "Adaptive blocksize observation model for mrf modeling of motion field," *IEEE Int. Conf. on Information, Communications and Signal Processing*, vol. CD ROM ICICS'99, December 1999.

[2] K. C. Chou, A. S. Willsky, and A. Benveniste, "Multiscale recursive estimation, data fusion, and regularization," *IEEE Trans. on Automatic Control*, vol. 39, pp. 464–478, March 1994.

[3] Stephanus S. T., T. S. Gunawan, and M. N. Chong, "Motion estimation using adaptive matching and multiscale method," *SPIE Conf. on Visual Communication and Image Processing*, p. will be presented, June 2000.

[4] Stephanus S. T., T. S. Gunawan, and M. N. Chong, "Multiscale motion estimation," *IECI Conf. on Multimedia and Network*, pp. 260–264, March 2000.

[5] E. Dubois and J. Konrad, "Bayesian estimation of motion vector fields," *IEEE Trans. on Pattern Analysis and Machine Intelligent*, vol. 14, pp. 910–927, September 1992.

[6] T. T.-J. Ho, *Multiscale Modeling and Estimation of Large-Scale Dynamic Systems*, Ph.D. thesis, Massachusetts Institute of Technology, September 1998.

Fig. 8. After multiscale smoothing