An Heuristic Search based Approach for Moving Objects Tracking

[1]Elena Sánchez-Nielsen, [2]Mario Hernández-Tejera
[1]Department of S.O.R. and Computer Science
University of La Laguna, 38271 La Laguna, Spain
enielsen@ull.es
[2]Institute of Intelligent Systems and Numerical Applications in Engineering,
University Campus of Tafira, 35017 Gran Canaria, Spain
mhernandez@isusiani.ulpgc.es

Abstract

Fast and accurate tracking of moving objects in video streams is a critical process in computer vision. This problem can be formulated as exploration problems and thus can be expressed as a search into a state space based representation approach. However, these search problems are hard to solve because they involve search through a high dimensional space. In this paper, we describe an A* heuristic search for computing efficient search through a space of transformations corresponding to the 2D motion of the object, where most promising search alternatives are computed by means of integrating target dynamics into the search process, and ideas from information theory are used to guide the search. The paper includes evaluations with video streams that illustrate the efficiency and suitability for real-time vision tasks on general purpose hardware. Moreover, the computational cost to carry out the tracking task is smaller than real time requirements (40 ms).

1 Problem Formulation

Let $T = \{t_1, \ldots, t_r\} \subseteq \Re^2$ be a set of points that represent a template, let $I = \{i_1, \ldots, i_s\} \subseteq \Re^2$ be another set of points that denote an input image, let a bounded set of translational transformations **G(.)** be a set of transformations $G : \Re^2 \to \Re^2$ consisting of translations, which are parameterized by a vector $(\Delta x, \Delta y)^T \in [\Delta_{\min}, \Delta_{\max}] \times [\Delta_{\min}, \Delta_{\max}] \subseteq \Re^2$ and let a bounded error notion of quality of match $Q(G; T, I, \varepsilon)$ be a measurement for computing the degree of match between a template T and a current input image I. The quality of match assigned to a transformation G is represented by the allowed error bound ε, when template points are brought to point's image using a transformation G:

$$Q(G) = \sum_{t \in T} \max_{i \in I} \| G(t) - i \| < \varepsilon \qquad (1)$$

Given a template T, an input image I and an error bound ε, template position matching problem can be viewed as a search process in the space of transformations in order to find the transformation, G_{max}, that maximizes the quality of match $Q(G)$:

$$G_{\max}(T, I, \varepsilon) = \arg\max_{G \in \mathbf{G}} Q(G; T, I, \varepsilon) \qquad (2)$$

2 A* Search Algorithm

Problem solving through the use of heuristic search strategies is expressed through a *state space based representation approach* [Pearl, 1984]. According to the heuristic search framework, template position matching problem is formulated as: *the search process oriented to find the transformation parameters $G_{max} \in$ **G(.)** that maximizes the quality of function match Q(G) between the transformed template G(T) and the current image I.* Next, each one of the elements of the problem is described:

- *State:* each state n is associated with a subset $G_k \subseteq \mathbf{G}(.)$, which is represented by the transformation corresponding to the centre that is referred to as G_c.
- *Initial state:* is represented by a bounded set of translational transformations **G(.)**, which allowing matching the current template into the current scene.
- *Final state:* is the transformation that best matches the current template points into the current scene, according to the quality of function *Q(G)*. The quality of function match assigned to a transformation G is expressed in terms of the partial directed Hausdorff distance between the transformed template T and the current image I:

  $$Q(G) = h_k(G(T), I) < \varepsilon \qquad (3)$$

  That is, the final state corresponds to the state with the translation G_{max} which verifies that the distance between 80% of translated template points and current image points are less than 2.0.
- *Operators:* for each current state n, the operators A and B are computed:
  - A. Each partial set of transformations from the current state is partitioned into four regions by vertical and horizontal bisections (that is, four new states).
  - B. The quality of function match (equation 3) is computed for each one of the four new states.

Splitting operation generates an expanded quaternary tree where with each node a 2^i x 2^j region is associated. The splitting operation is finished when the quadrisection process computes a translational motion according to the quality of function match *Q(G)*. The best node to expand is computed through the use of an A* approach [Pearl, 1984]. The corresponding cost value assigned to each state *n* is defined as:

$$f(n) = c(n) + h^*(n) \tag{4}$$

h(n)* is formulated in terms of expanding the best state of the search space based on measuring the similarity among the distribution functions that characterize the current state *n* and the objective state, denoted respectively by *P* and *Q*. The distribution function *P* is represented by a histogram distances $\{H_{Gc}\}_{i=1\ldots r}$, which contains the number of template points *T* at distance d_j with respect to the points of the input image *I*, when the transformation G_c of the current state *n* is applied on the current template *T*. The distribution function *Q* is modelled by approximating $\{H_{Gc}\}_{i=1\ldots r}$ by an exponential distribution function $f(n) = ke^{-an}$.

The similarity between the distributions *P* and *Q* is evaluated through the relative entropy or Kullback-Leibler distance measurement from information theory [Cover and Thomas, 1991]:

$$D(P \parallel Q) = \sum_{i=1}^{R} p_i \log \frac{p_i}{q_i} \tag{5}$$

An estimated cost function *c(n)* is employed to lead the search process towards promising solutions. With this aim, a depth term is added to the *f(n)* function (expression 4). This term is based on the number of operators of type *A* applied from the initial search node to the current node *n*.

An alpha-beta predictive filtering [Shalom and Rong, 1993] is incorporated into the search algorithm with the purpose of predicting target position and reducing the initial size of the search space of transformations, focused on the assumption that there is a relationship between the size of the search area and the resulting uncertainty of the alpha-beta predictive filtering. Figure 1 illustrates the integration of the stages that compose the alpha-beta filtering with the computation of the dimension *M x N* of **G(.)** and their respective equations.

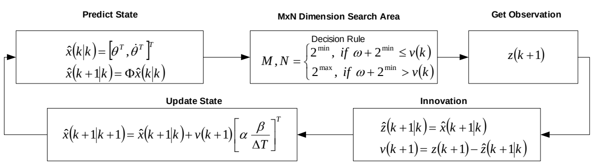

Figure 1. Computing adjustable-size based search area.

The algorithm for template position matching in our approach based on an A* search is as follows.

Input

G(.): initial set of transformations.

ε: distance error bound allowed when template points are brought to point's image using a transformation *G*.

v(k) : uncertainty measurement computed by alpha-beta filtering approach.

ω: estimate which determines from what value should be selected 2^{max} or 2^{min} bound value for computing *MxN* dimension of **G(.)**.

2^{max} and 2^{min}: upper and lower nearest values to the innovation factor *v(k)*.

η: number of operators of type *A* applied from the initial search state to the current state *n*.

Output

MxN dimension of initial state set **G(.)**

$$G_{max} = \left[g_x^{max}, g_y^{max}\right]$$

Algorithm

Step 1) Compute adjustable *MxN* dimension of initial state set **G(.)**:

$$M, N = \begin{cases} 2^{min}, \ if \ \omega + 2^{min} \le v(k) \\ 2^{max}, \ if \ \omega + 2^{min} > v(k) \end{cases}$$

Step 2) Find G_{max} such the quality of function match $Q(G) = h_k(G(T), I) < \varepsilon$ is verified:

While ($Q(G) > \varepsilon$) **Do**

2.1) Split current state *n* into four new states $\{n\}_{i=1\ldots4}$

2.2) Compute $Q(G_c) \leftarrow h_k(G_c(T), I)$ for each new n_i

2.3) Expand the best state n_i according the evaluation function $f(n) = c(n) + h^*(n)$:

2.3.1) $h^*(n) \leftarrow D(P \| Q)$

2.3.2) $c(n) \leftarrow c(n-1) + \eta$

End While

Step 3) Output $G_{max} = \left[g_x^{max}, g_y^{max}\right]$

3 Comparative Analysis and Runtime

Diverse experiments have been carried out with 24 different sequences, achieving the same behaviour for all of them on a P-IV 2.4 GHz. Comparing the results achieved from the proposed A* heuristic search that uses an adjustable-size search area approach with the blind search strategy that does not use information to guide the search process and does not estimate the dimension appropriated for initial search state such as the conventional search strategy described by [Rucklidge, 1996], demonstrates that the proposed A* heuristic search framework is computationally lighter and as consequence, faster than the blind search strategy, in an average rate of three times better. Average time measured in seconds for processing each frame for the 24 different sequences is 0.01, confirming the adaptation of the search strategy proposed to real-time restrictions under unrestricted environments for arbitrary shapes.

References

[Cover and Thomas, 1991] T. Cover and J. Thomas. Elements of Information Theory. *John Wiley & Sons Incs*, 1991.

[Pearl, 1984] J. Pearl. Heuristics. Intelligent Search Strategies for Computer Problem Solving. *Addison-Wesley Series in Artificial Intelligence*, 1984.

[Rucklidge, 1996] W. Rucklidge. Efficient computation of the minimum Hausdorff distance for visual recognition. *Lecture Notes in Computer Science, 1173, Springer-Verlag*, 1996.

[Shalom and Rong, 1993] Estimation and Tracking: Principles, Techniques, and Software. Artech House, Boston, 1993.