РОССИЙСКАЯ АКАДЕМИЯ НАУК

СИБИРСКОЕ ОТДЕЛЕНИЕ

АВТОМЕТРИЯ

№ 2 1994

УДК 629.705 : 621.391.268

Б. А. Алпатов

(Рязань)

ОПТИМАЛЬНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ ДВИЖУЩЕГОСЯ ОБЪЕКТА В ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ

Задача слежения за изображением объекта сформулирована как задача оценки параметров. Решение, полученное путем последовательной оптимизации критерия максимума апостериорной вероятности, включает процедуры совмещения изображений, классификации участка наблюдаемого изображения на точки фона и объекта, межкадровой фильтрации яркости изображений и траекторной фильтрации координат центра объекта.

Введение. В ряде технических приложений возникает задача автоматического слежения за протяженным движущимся объектом, с течением времени изменяющим свои параметры (яркость, конфигурацию) и наблюдаемым на неоднородном фоне. В литературе при решении этой проблемы доминируют эвристические подходы. Необходимо отметить также, что эта задача, имея много общего с задачей навигации летательных аппаратов по картам местности [1], имеет тем не менее существенные отличия, связанные с необходимостью выделения множества точек, принадлежащих объекту слежения. В работе [2] изложен подход к решению данной задачи, однако примененная процедура классификации не является наилучшей. В то же время линеаризованный алгоритм, описанный в [3], работоспособен только в установившемся режиме.

В данной статье на основе относительно несложных моделей состояния и наблюдения дается более корректное решение поставленной задачи.

Модель наблюдения. Будем использовать следующую модель наблюдения:

$$\mathbf{L}(n) = \mathbf{R}(n)\mathbf{H}(n) + (\mathbf{I} - \mathbf{R}(n))\mathbf{G}(n) + \Xi(n), \tag{1}$$

где $\mathbf{L}(n)$ — наблюдаемое изображение L_n, представленное в форме вектора; $\mathbf{H}(n)$ — вектор, соответствующий изображению объекта H_n; $\mathbf{G}(n)$ — вектор, соответствующий изображению фона G_n; $\Xi(n)$ — вектор аддитивных шумов; $\mathbf{R}(n)$ — квадратная диагональная матрица, диагональные элементы которой $r_{ii}(n) = \{1, 0\}$; $\mathbf{I}$ — единичная матрица; n — номер кадра изображения.

Удобно в дальнейшем для каждой точки $(i, j) \in L_n$ переписать (1) в виде

$$l(i, j, n) = s(i, j, n) + \xi(i, j, n), \tag{2}$$

где $s(i, j, n) = r(i, j, n)h(i, j, n) + (1 - r(i, j, n))g(i, j, n)$; $h(i, j, n)$, $g(i, j, n)$ — соответственно яркости точек объекта и фона; $\xi(i, j, n)$ — нормальный, некоррелированный по пространству и времени шум с нулевым средним и дисперсией D; $r(i, j, n) = 1$, если в точке (i, j) в n-м кадре объект закрывает фон, и $r(i, j, n) = 0$ в противном случае.

В дальнейшем предполагается, что изображение объекта H_n представляет собой связную совокупность точек. Модель (1) нетрудно распространить на случай наблюдения за несколькими объектами и наблюдения с помощью нескольких различных датчиков.

Модели состояния. Полагая фон неподвижным относительно некоторой системы отсчета и пренебрегая эффектами дискретизации изображения, примем модель состояния фона в виде

$$g(i, j, n) = g(i, j, n-1) + \omega(i, j, n-1), \tag{3}$$

где $g(i, j, 0)$ — гауссова случайная величина с известным математическим ожиданием и дисперсией; $\omega(i, j, n) \sim N(0, D_\omega)$ — процесс, некоррелированный во времени и протекающий независимо в каждой точке изображения. По существу, такое утверждение означает отказ от использования в алгоритме оценки реально существующей пространственной корреляции изображений. Платой за это является некоторое снижение потенциально возможной точности решаемых задач. Однако выигрыш в уменьшении сложности используемых алгоритмов является куда более важным с точки зрения возможности решения задач в реальном масштабе времени.

Таким образом, модель фона задается как марковский процесс, независимо протекающий в каждой точке, что формирует неоднородное по пространству изображение. Реальные реализации процессов $g(i, j, n)$ не могут быть отрицательными и представляют собой подмножество реализаций, моделируемых с помощью (3).

Для описания объекта введем дополнительную систему координат (ν, μ), начало которой совпадает с центром изображения объекта. Для определенности под центром объекта слежения будем понимать среднеарифметическое координат составляющих его точек. С использованием этой системы координат, пренебрегая эффектами пространственной дискретизации, можно записать, что

$$h(\nu, \mu, n) = h(\nu, \mu, n-1) + \eta(\nu, \mu, n-1), \tag{4}$$

где в отношении состояния $h(\nu, \mu, 0)$ и процессов $\eta(\nu, \mu, n-1)$ принимаются ранее сделанные предположения о фоне.

Для обеспечения возможности отслеживания изменений размеров и конфигурации объекта слежения (при отсутствии трехмерной модели объекта) введем в рассмотрение множество точек $\Psi_n \supset H_n$. Будем предполагать, что геометрический центр множества Ψ_n совпадает с точкой $\lambda(n) = [\lambda_x(n)\, \lambda_y(n)]^T$, принимаемой за центр объекта. Конфигурация множества Ψ_n может быть различной и, в принципе, может изменяться с изменением формы объекта.

Зададим на множестве Ψ_n параметры $r(\nu, \mu, n)$, где по-прежнему $r(\nu, \mu, n) = 1$, если точка (ν, μ) в n-м кадре принадлежит объекту, и $r(\nu, \mu, n) = 0$ в противном случае. Введем также для каждой точки Ψ_n распределение $W(r(\nu, \mu, n)) = \{p(\nu, \mu, n), 1 - p(\nu, \mu, n)\}$, где $p(\nu, \mu, n)$ — вероятность принадлежности точки (ν, μ) к объекту в n-м кадре. Вероятности $p(\nu, \mu, n)$ могут быть представлены как

$$p(\nu, \mu, n) = p(\nu, \mu, n-1)p(\mathrm{O}/\mathrm{O}) + (1 - p(\nu, \mu, n))p(\mathrm{O}/\Phi), \tag{5}$$

где $p(\mathrm{O}/\mathrm{O})$ — вероятность того, что точка объекта с координатами (ν, μ) не перейдет в точку фона за время кадра на множестве Ψ_n; $p(\mathrm{O}/\Phi)$ — вероятность появления новой точки, принадлежащей объекту. Вновь включаемым точкам в Ψ_n (при его расширении) назначается небольшая априорная вероятность $p_a(\nu, \mu, n)$ принадлежности к объекту.

Перейдем к описанию модели движения объекта. Полагая, что изображение объекта — это связная совокупность точек, которая смещается вдоль осей координат от кадра к кадру, запишем уравнение состояния в виде

$$\Lambda(n) = \mathrm{C}\Lambda(n-1) + \Theta(n-1), \tag{6}$$

где $\Lambda(n) = [\lambda_x(n)\dot{\lambda}_x(n)\,\lambda_y(n)\dot{\lambda}_y(n)]^T$ — вектор координат и их производных точки нирование наблюдения L_n).

Предположим, что в результате предыдущей обработки известны априорное гауссово распределение для $\lambda(n)$, априорное распределение $\overline{W}(r(\nu, \mu, n))$ на множестве Ψ_n, априорные гауссовы распределения для $h(\nu, \mu, n)$ и $g(\nu, \mu, n)$. Эта информация может быть сформирована на этапе начального обнаружения объекта. Таким образом, можно считать, что имеется наблюдаемое изображение с центром в точке $\overline{\lambda}(n)$. Используя ранее введенные модели состояния и наблюдения и применяя метод максимума апостериорной вероятности, после логарифмирования приходим к необходимости минимизации критерия

$$J_n = (\lambda(n) - \overline{\lambda}(n))^T \mathbf{K}_\lambda^{-1}(n)(\lambda(n) - \overline{\lambda}(n)) - 2 \sum_{\nu, \mu \in \Psi_n} \ln \overline{W}(r(\nu, \mu, n)) +$$

$$+ \sum_{\nu, \mu \in H_n} 1/\overline{D}_h(\nu, \mu, n)(h(\nu, \mu, n) - \overline{h}(\nu, \mu, n))^2 +$$

$$+ \sum_{i, j \in G_n} 1/\overline{D}_g(i, j, n)(g(i, j, n) - \overline{g}(i, j, n))^2 + 1/D \sum_{i, j \in L_n} (l(i, j, n) - s(i, j, n))^2, \quad (7)$$

где $\overline{\lambda}(n)$ — прогнозируемые координаты объекта; $\mathbf{K}_\lambda(n)$ — ковариационная матрица вектора $\lambda(n)$; $\overline{h}(\nu, \mu, n)$ и $\overline{g}(i, j, n)$ — прогнозируемые яркости точек объекта и фона; $\overline{D}_h(\nu, \mu, n)$ и $\overline{D}_g(i, j, n)$ — прогнозируемые дисперсии яркостей.

Отыскание глобального минимума J обычными методами в пространстве всех параметров представляет чрезвычайно трудоемкую задачу. Поэтому воспользуемся тем обстоятельством, что к n-му кадру имеются некоторые оценки параметров $r(\nu, \mu, n)$, $h(\nu, \mu, n)$, $g(i, j, n)$, и, следовательно, эти параметры вначале могут быть приняты равными прогнозируемым. Таким образом, оптимизацию (7) предлагается проводить в несколько этапов по группам параметров. На первом этапе определяется измерение $\widetilde{\lambda}(n)$ путем оптимизации критерия

$$J_{n1} = (\lambda(n) - \overline{\lambda}(n)) \mathbf{K}_\lambda^{-1}(n)(\lambda(n) - \overline{\lambda}(n)) + 1/D \sum_{i, j \in L_n} (l(i, j, n) - \overline{s}_1(i, j, n))^2, \quad (8)$$

где $\overline{s}_1(i, j, n) = \overline{r}(i, j, n)\overline{h}(i, j, n) + (1 - \overline{r}(i, j, n))\overline{g}(i, j, n)$. Заметим, что $\overline{r}(i, j, n)$ и $\overline{h}(i, j, n)$ являются функциями от $\lambda(n)$, а сумма $\sum_{\nu, \mu \in \Psi_n} \ln \overline{W}(r(\nu, \mu, n))$ не влияет на оптимизацию значения $\lambda(n)$.

Поиск $\widetilde{\lambda}(n)$, при котором достигается абсолютный минимум J_{n1}, сводится к наложению прогнозируемого изображения объекта на прогнозируемое изображение фона для каждой точки исследуемой области и вычислению J_{n1}. Если пренебречь априорным распределением $\lambda(n)$ и принять фон нулевым, нетрудно показать, что первый этап оптимизации сведется к классическому поиску максимума взаимной корреляционной функции наблюдаемого изображения и прогнозируемого (эталонного) изображения объекта. В результате первого этапа отыскивается $\widetilde{\lambda}(n)$, обеспечивающее выход в окрестность глобального минимума (7), и определяется множество точек Ψ_n на изображении L_n, содержащее объект.

На втором этапе с учетом найденного значения $\hat{\lambda}(n)$ оцениваются параметры $r(\nu, \mu, n)$ (причем по-прежнему полагается, что $h(\nu, \mu, n) = h(\nu, \mu, n)$ и $g(i, j, n) = \bar{g}(i, j, n)$) путем оптимизации критерия

$$J_{n2} = -2 \sum_{\nu, \mu \in \Psi_n} \ln \overline{W}(r(\nu, \mu, n)) + 1/D \sum_{i, j \in L_n} (l(i, j, n) - \bar{s}_2(i, j, n))^2, \tag{9}$$

где $\bar{s}_2(i, j, n) = r(i, j, n)h(i, j, n) + (1 - r(i, j, n))\bar{g}(i, j, n)$.

Так как предполагается, что обнаружение новых объектов не ведется, то можно считать, что $r(i, j, n) = 0$ для $(i, j) \notin \Psi_n$, и оптимизацию (9) в этом случае следует проводить на множестве Ψ_n, т. е.

$$J_{n2} = -2 \sum_{\nu, \mu \in \Psi_n} \ln \overline{W}(r(\nu, \mu, n)) +$$

$$+ 1/D \sum_{\nu, \mu \in \Psi_n} (l(\nu, \mu, n) - r(\nu, \mu, n)\bar{h}(\nu, \mu, n) - (1 - r(\nu, \mu, n)) \cdot \bar{g}(\nu, \mu, n))^2. \tag{10}$$

Оптимизация (10) сводится к поточечной классификации множества Ψ_n на точки объекта, для которых оценка $\hat{r}(\nu, \mu, n) = 1$, и точки фона, для которых $\hat{r}(\nu, \mu, n) = 0$. Значение критерия в точке (ν, μ) при $r(\nu, \mu, n) = 1$ составит:

$$J_{n2}(r(\nu, \mu, n) = 1) = -2\ln\bar{p}(\nu, \mu, n) + 1/D(l(\nu, \mu, n) - \bar{h}(\nu, \mu, n))^2, \tag{11}$$

а при $r(\nu, \mu, n) = 0$ соответственно

$$J_{n2}(r(\nu, \mu, n) = 0) = -2\ln(1 - \bar{p}(\nu, \mu, n)) + 1/D(l(\nu, \mu, n) - \bar{g}(\nu, \mu, n))^2, \tag{12}$$

где $\bar{p}(\nu, \mu, n)$ — прогнозируемая вероятность, причем в соответствии с (5)

$$\bar{p}(\nu, \mu, n) = \hat{p}(\nu, \mu, n - 1)p(О/О) + (1 - \hat{p}(\nu, \mu, n - 1))p(О/\Phi),$$

а $\hat{p}(\nu, \mu, n - 1)$ — апостериорная вероятность принадлежности точки (ν, μ) к обьекту в $(n - 1)$ кадре изображения. Очевидно, необходимо выбрать то значение $r(\nu, \mu, n)$, для которого величина критерия $J_{n2}(r(\nu, \mu, n))$ меньше, и, следовательно, $\hat{r}(\nu, \mu, n) = 1$ и точка (ν, μ) принадлежит обьекту, если

$$(l(\nu, \mu, n) - \bar{g}(\nu, \mu, n))^2 - (l(\nu, \mu, n) - \bar{h}(\nu, \mu, n))^2 > 2D\ln \frac{1 - \bar{p}(\nu, \mu, n)}{\bar{p}(\nu, \mu, n)}. \tag{13}$$

В качестве прогнозируемых яркостей точек обьекта в точках, принадлежащих фону, до решения задачи классификации ($\bar{r}(\nu, \mu, n) = 0$) выбираются яркости $\bar{h}(\nu, \mu, n)$ близлежащих точек обьекта.

Апостериорная вероятность принадлежности точки к обьекту может быть вычислена с помощью формулы Байеса

$$\hat{p}(\nu, \mu, n) = \frac{\bar{p}(\nu, \mu, n)W(l(\nu, \mu, n)/\bar{h}(\nu, \mu, n))}{\bar{p}(\nu, \mu, n)W(l(\nu, \mu, n)/\bar{h}(\nu, \mu, n)) + (1 - \bar{p}(\nu, \mu, n)W(l(\nu, \mu, n)/\bar{g}(\nu, \mu, n)))}, \tag{14}$$

где $W(l(\nu, \mu, n)/h(\nu, \mu, n))$ и $W(l(\nu, \mu, n)/\bar{g}(\nu, \mu, n))$ — условные гауссовы распределения. Вероятности $\hat{p}(\nu, \mu, n)$ могут быть непосредственно использованы для классификации.

Если принять $\bar{p}(\nu, \mu, n) = \hat{p}(\nu, \mu, n - 1)$, то классификация точек по критерию максимума апостериорной вероятности может выполняться несколько иначе, а именно $\hat{r}(\nu, \mu, n) = 1$, если

$$\ln\left(\frac{\hat p(\nu,\mu,n)}{1-\hat p(\nu,\mu,n)}\right) = \ln\left(\frac{\hat p(\nu,\mu,n-1)}{1-\hat p(\nu,\mu,n-1)}\right) +$$

$$+\frac{(l(\nu,\mu,n)-\bar g(\nu,\mu,n))^2-(l(\nu,\mu,n)-h(\nu,\mu,n))^2}{2D} > 0. \qquad (15)$$

При необходимости, после классификации, может быть проведена логическая фильтрация для удаления одиночных точек и заполнения «дыр» на объекте, являющихся результатом ошибочной классификации. После выделения множества точек, принадлежащих объекту, целесообразно уточнить $\tilde\lambda(n)$ путем вычисления среднеарифметического координат выделенных точек.

Третий этап оптимизации критерия (7) состоит в нахождении оценок $\hat h(\nu,\mu,n)$ и $\hat g(i,j,n)$ при условии, что определены $\tilde\lambda(n)$ и $\hat r(\nu,\mu,n)$. Очевидно, что задача оценки $h(\nu,\mu,n)$ и $g(i,j,n)$ теперь может считаться линейной и дальнейшая оптимизация (7) сводится к раздельной оптимизации по точкам фона и объекта. Так, по отношению к точкам объекта необходимо минимизировать критерий

$$J_{n3} = \sum_{\nu,\mu\in H_n}(1/\bar D_h(\nu,\mu,n)(h(\nu,\mu,n)-\bar h(\nu,\mu,n))^2 + 1/D(l(\nu,\mu,n)-h(\nu,\mu,n))^2), \qquad (16)$$

а по отношению к точкам фона

$$J_{n4} = \sum_{i,j\in G_n}(1/\bar D_g(i,j,n)(g(i,j,n)-\bar g(i,j,n))^2 + 1/D(l(i,j,n)-g(i,j,n))^2). \qquad (17)$$

Из выражений (16), (17) следуют обычные уравнения фильтра Калмана [4]:

$$\hat h(\nu,\mu,n) = \bar h(\nu,\mu,n) + k(\nu,\mu,n,h)(l(\nu,\mu,n)-\bar h(\nu,\mu,n)), \qquad (18)$$

$$\hat g(i,j,n) = \bar g(i,j,n) + k(i,j,n,g)(l(i,j,n)-\bar g(i,j,n)), \qquad (19)$$

где $k(\nu,\mu,n,h) = \dfrac{\bar D_h(\nu,\mu,n)}{D+\bar D_h(\nu,\mu,n)}$; $\bar D_h(\nu,\mu,n) = \hat D_h(\nu,\mu,n-1) + D_\eta$; $\bar h(\nu,\mu,n) = \hat h(\nu,\mu,n-1)$; $\hat D_h(\nu,\mu,n) = \dfrac{D\bar D_h(\nu,\mu,n)}{D+\bar D_h(\nu,\mu,n)}$. Аналогично вычисляются $k(i,j,n,g)$. Для точек фона, закрытых в данном кадре объектом, в соответствии с используемой моделью $\hat g(i,j,n) = \bar g(i,j,n) = \hat g(i,j,n-1)$.

Если считать, что измерение координат сопровождается белым шумом $\delta\lambda(n) = [\delta\lambda_x(n)\,\delta\lambda_y(n)]^T$ с гауссовым распределением, некоррелированным с $\lambda(n)$, т. е.

$$\tilde\lambda(n) = \mathrm{A}\Lambda(n) + \delta\lambda(n), \qquad (20)$$

где $\mathrm{A} = \begin{bmatrix}1\,0\,0\,0\\0\,0\,1\,0\end{bmatrix}$, то с учетом принятой модели состояния (6) алгоритм оптимальной оценки вектора $\Lambda(n)$ представляет также линейный фильтр Калмана:

$$\hat\Lambda(n) = \bar\Lambda(n) + \mathrm{K}(n)(\tilde\lambda(n) - \mathrm{A}\bar\Lambda(n)), \qquad (21)$$

где $\mathrm{K}(n)$ вычисляется обычным образом, а из $\mathrm{K}(n)$, в свою очередь, нетрудно получить $\mathrm{K}_1(n)$, используемую в (7).

Исследование полученного алгоритма проводилось путем статистического моделирования на ЭВМ с использованием искусственных изображений обстановки. При этом установлено, что ошибка измерения координат объекта не превышает десятых долей интервала пространственной дискретизации вплоть до отношений контраста объекта к среднеквадратическому отклонению шума два к одному. Результаты моделирования подтвердили способность алгоритма отслеживать изменения, возникающие в объекте в процессе наблюдения, при наличии сложного фонового изображения.

Заключение. Задача слежения за изображением объекта, движущимся на сложном фоне, сформулирована как задача оценки параметров. Решение, полученное путем последовательной оптимизации критерия максимума апостериорной вероятности, включает несколько этапов: совмещение прогнозируемого и наблюдаемого изображений; классификацию точек (сегментацию) участка наблюдаемого изображения на точки фона и объекта; межкадровую фильтрацию яркостей точек фона и объекта; траекторную фильтрацию координат центра объекта.

СПИСОК ЛИТЕРАТУРЫ

1. Баклицкий В. К. Применение метода фильтрации Калмана к синтезу корреляционно-экстремальных систем // Изв. вузов. Радиоэлектроника.—1982.—№ 3.
2. Алпатов Б. А. Оценивание параметров движущегося объекта в последовательности изменяющихся двумерных изображений // Автометрия.—1991.—№ 3.
3. Алпатов Б. А. Оценивание параметров движущегося фрагмента изображения // Техника средств связи. Сер. Техника телевидения.—1991.—№ 2.
4. Брайсон А., Хо Ю Ши. Прикладная теория оптимального управления.—М.: Мир, 1972.

Поступила в редакцию 17 февраля 1993 г.