

Capítulo 3

Aproximación probabilística al seguimiento visual

En este capítulo se establecen las bases del problema del seguimiento visual de objetos. En primer lugar se muestra como el seguimiento se puede expresar como un problema de estimación, donde se desea encontrar los valores que describen el movimiento del objeto pero sólo se dispone de observaciones de estas variables. Se presenta la solución del problema con un enfoque probabilístico que pretende encontrar los valores más probables condicionados a las observaciones obtenidas hasta ese momento. Para encontrar estos valores condicionados se utiliza el Teorema de Bayes, el cuál se presenta en forma recursiva para poder ser aplicado a series temporales. Para poder completar la definición es necesaria la formalización en términos probabilísticos del conocimiento previo por medio de modelos dinámicos y la definición de la relación existente entre variables y observaciones por medio de la función de *likelihood*. Una vez definido completamente el problema, se repasa la forma de poder realizar los cálculos que permiten la obtención de la estimación de las variables por medio de la representación muestral de funciones de densidad de probabilidad. Finalmente, se introduce el problema de la estimación de múltiples objetos.

3.1 Introducción.

La finalidad de los algoritmos de seguimiento (*tracking*) es la estimación del número y el estado de los objetos que están dentro de una región de interés. El término **estado** se utiliza para caracterizar al objeto que está siendo seguido. Normalmente, el estado es un vector de cantidades que representa al objeto en el espacio de estados. Sin embargo, en las aplicaciones del mundo real, no se dispone directamente del estado del objeto, en su lugar, se obtienen un conjunto de medidas del estado del objeto proporcionadas por uno o más sensores, estas medidas reciben el nombre de **observaciones**. De esta forma, el seguimiento se convierte en un problema de **estimación**, donde se desea estimar el estado del objeto a partir de las observaciones.

Los primeros algoritmos de estimación se basaban en el hecho de que es posible modelar de forma única el sistema que genera los estados y el proceso de medida que genera las observaciones. Utilizando el modelo del sistema es posible predecir el estado del objeto antes de que éste ocurra¹. El proceso de medida se utiliza para enlazar las observaciones y las predicciones. Así, utilizando ambos modelos, los algoritmos clásicos de estimación intentan encontrar la combinación óptima² de observaciones y predicciones para estimar el estado actual del objeto, ver Fig. 3.1.

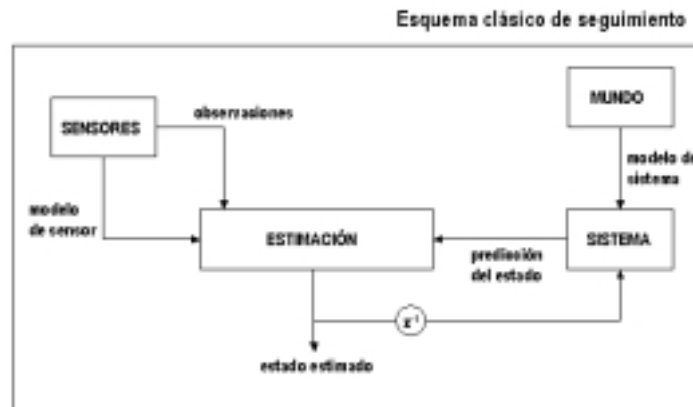


Figura 3.1: Esquema clásico de un sistema de seguimiento.

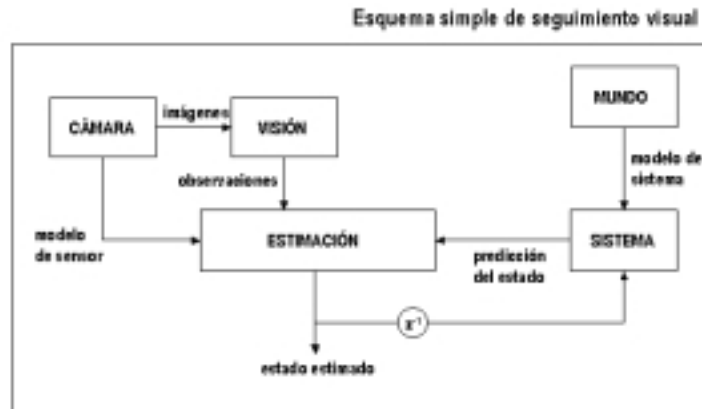
El modelo del sistema puede hallarse a partir de un sistema de aprendizaje a partir de ejemplos, o ser modelado de forma matemática si se dispone de un conocimiento profundo del problema. Hay gran cantidad de algoritmos de aprendizaje a partir de ejemplos que pueden utilizarse para este propósito [6, 76, 11, 65].

El modelado del proceso de medida depende del tipo de sensor utilizado para medir el estado del objeto. Algunas veces es posible modelar el sensor de forma matemática y encontrar la ecuación que describe el proceso de medida. Por ejemplo, es posible tener el modelo de un radar[10]. El término **visual** se incluye en el problema del seguimiento si se utiliza un sistema de adquisición de imágenes como sensor. Utilizar las imágenes como medida añade complejidad al problema por la gran cantidad de información que contienen. Información que puede ser redundante para resolver el problema del seguimiento visual. Por este motivo, en la mayoría de problemas de seguimiento visual, hay una etapa de preprocesamiento que extrae características de interés de las imágenes. Estas características son utilizadas como las observaciones del sensor, ver Fig. 3.2 (a). En visión por computador, esta es la aproximación más utilizada[22, 96, 40, 90, 31, 67, 81].

¹ Anticiparse a los hechos es otra de las finalidades importantes del seguimiento.

² En el sentido de alguna medida de error.

El problema que hay con esta aproximación es que si hay errores en la fase de preprocesamiento no pueden ser recuperados por el sistema de estimación, provocando que el sistema de seguimiento no funcione de forma adecuada. Por tanto, es necesario cambiar el planteamiento del problema para incluir la fase de preprocesamiento en el problema de estimación, ver Fig. 3.2 (b).



(a) Esquema básico de seguimiento visual.



(b) Esquema actual de seguimiento visual.

Figura 3.2: Esquemas de seguimiento visual.

Los trabajos de Terzopoulos y Szeliski [86], Baumberg y Hogg [5], e Isard y Blake [42] son ejemplos de algoritmos de seguimiento visual que siguen el esquema de visión y estimación conjunta. Todos estos trabajos utilizan como observación la forma visual (contorno) del objeto. Es posible modelar la forma del objeto e incluir este modelo en los algoritmos de estimación. Finalmente, señalar que el modelo de movimiento del contorno es posible aprenderlo a partir de secuencias de imágenes simples, previamente

a la utilización del esquema completo. Sin embargo, una desventaja de estos sistemas es que se componen de una gran cantidad de parámetros que es necesario sintonizar y que este aprendizaje ha de ser realizado previamente, con lo que no se permite un reajuste del sistema si éste cambia durante la secuencia de interés. Por otro lado, la tendencia actual es utilizar la representación basada en *blobs*, como demuestra el reciente trabajo de Isard[45], en la que intenta adaptar su algoritmo basado en contornos para poder trabajar directamente con *blobs*.

3.2 Aproximación probabilística.

3.2.1 La función de *likelihood* y el Teorema de Bayes.

Representaremos el estado de un objeto con una variable aleatoria, S , cuyos valores pertenecen al espacio de estados \mathcal{S} , y de la cuál no conocemos su valor. Se dispone de una observación que se representa con la variable aleatoria Z , que toma sus valores, z , de un espacio de medida \mathcal{M} . La relación entre S y Z se encuentra en la función de *likelihood*.

Definición: La función de *likelihood*, para una variable aleatoria S que representa el estado del objeto y una observación $Z = z$ se define como:

$$p(z|s) = P(Z = z|S = s) \quad s \in \mathcal{S}, \quad z \in \mathcal{M} , \quad (3.1)$$

donde $p(\cdot|s)$ es una función de densidad de probabilidad en \mathcal{M} . La función de *likelihood* puede ser interpretada como una medida de probabilidad, la probabilidad de que un estado s genere una observación z , es decir, es la función que modela el proceso de medida. Sin embargo, lo que se persigue en los problemas de seguimiento es la estimación del estado actual del objeto a partir de las medidas obtenidas. La forma de obtener esta estimación a través de la función de *likelihood* es utilizando el Teorema de Bayes. La densidad de probabilidad condicionada para S dada la observación $Z = z$ es:

$$p(s|z) = \frac{p(z|s)p(s)}{\int p(z|s)p(s) ds} . \quad (3.2)$$

La expresión de la Ec. (3.2) es conocida como Teorema de Bayes. La densidad descrita por la función p se denomina densidad **previa** (en adelante *prior*), la densidad dada por $p(\cdot|z)$ se denomina **posterior** de S dada la observación $Z = z$, por tanto:

$$\text{posterior} \propto \text{likelihood} \times \text{prior} .$$

El Teorema de Bayes proporciona una formulación matemática de un método cuyo conocimiento previo puede combinarse con el nuevo conocimiento (proporcionado por

las observaciones)³.

Suponer que tenemos una observación inicial $Z_1 = z_1$, por el Teorema de Bayes tenemos:

$$p(s|z_1) \propto p(z_1|s) p(s) . \quad (3.3)$$

Ahora podemos suponer que se dispone de una segunda observación $Z_2 = z_2$, donde Z_1 e Z_2 son independientes dado $S = s$, entonces:

$$\begin{aligned} p(s|z_1, z_2) &\propto p(z_2|s) p(z_1|s) p(s) \\ &\propto p(z_2|s) p(s|z_1) . \end{aligned} \quad (3.4)$$

La expresión de la segunda línea de la Ec. (3.4) tiene la misma forma que la de la Ec. (3.3), excepto que la distribución posterior de S dado $Z_1 = z_1$ juega el papel de distribución *prior* para la segunda observación. Este proceso puede repetirse n veces si disponemos de n observaciones independientes.

3.2.2 Filtraje Bayesiano.

La versión recursiva del Teorema de Bayes se denomina Filtraje Bayesiano [36, 3, 82, 11]. A continuación, se describe el proceso completo del Filtraje Bayesiano sin asumir ninguna forma funcional de las densidades de probabilidad.

Se expresará el Filtraje Bayesiano en términos de tiempo discreto, t , por motivos computacionales. Sea \mathcal{S} el espacio de estados del objeto. El estado del objeto en tiempo t se denota como \mathbf{s}_t . Normalmente, el estado del objeto es un vector de escalares. En los problemas clásicos de seguimiento, los componentes escalares del espacio de estados son por ejemplo: posición, velocidad y aceleración. Puede haber componentes adicionales que estarán relacionados con la identificación del objeto.

La información previa del objeto se representa por medio de un proceso estocástico $\{\mathbf{s}_t; t \geq 0\}$, donde \mathbf{s}_t es el estado del objeto en tiempo t . Este proceso estocástico es la descripción probabilística del modelo dinámico del movimiento del objeto. El proceso comienza en tiempo $t = 0$ y el interés es la estimación de \mathbf{s}_t para $t > 0$. Los posibles estados de este proceso corresponden a los posibles estados del objeto dentro del espacio de estados, \mathcal{S} .

Suponer que en tiempo t se ha obtenido un conjunto de observaciones en tiempos $t_1 \leq \dots \leq t$. Para permitir la posibilidad de que puedan recibirse más de una observación en un instante de tiempo, se representará el conjunto de observaciones en tiempo t por medio del vector \mathbf{z}_t . Utilizando el modelo estocástico del espacio de estados del objeto y la función de *likelihood* es posible calcular la distribución posterior como:

³Otra posible interpretación es describir una relación “causa-consecuencia”. Es decir, el posterior se puede ver como la causa más probable que explica la consecuencia ocurrida.

$$p(\mathbf{s}_t | \mathbf{z}_t, \dots, \mathbf{z}_1) = \frac{p(\mathbf{s}_t, \mathbf{z}_t, \dots, \mathbf{z}_1)}{p(\mathbf{z}_t, \dots, \mathbf{z}_1)} . \quad (3.5)$$

Son necesarias dos asunciones sobre el proceso estocástico y la función de *likelihood* para permitir el cálculo de la distribución posterior. Primero se asume que la función de densidad de probabilidad para el sistema estocástico forma una cadena temporal de Markov, es decir, que la probabilidad de estar en el estado \mathbf{s}_t en tiempo t sólo depende del estado en tiempo $t - 1$:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}, \dots, \mathbf{s}_1) = p(\mathbf{s}_t | \mathbf{s}_{t-1}) . \quad (3.6)$$

Y debido a la asunción de Markov:

$$p(\mathbf{s}_t, \dots, \mathbf{s}_1) = \int \prod_{i=1}^t p(\mathbf{s}_i | \mathbf{s}_{i-1}) p(\mathbf{s}_0) d\mathbf{s}_0 , \quad (3.7)$$

donde $p(\mathbf{s}_0)$ es la densidad de probabilidad del estado del objeto para $t = 0$.

La segunda asunción se realiza sobre las observaciones

$$p(\mathbf{z}_t, \dots, \mathbf{z}_1 | \mathbf{s}_t, \dots, \mathbf{s}_1) = \prod_{i=1}^t p(\mathbf{z}_i | \mathbf{s}_i) . \quad (3.8)$$

Esta asunción significa que las observaciones en tiempo t sólo dependen del estado del sistema en tiempo t . Esto no es cierto en la mayoría de casos reales, por ejemplo, si el espacio de estados del objeto sólo incluye su posición espacial y la observación es una medida de velocidad, esta observación dependerá del estado del objeto en un intervalo de tiempo. La solución en este caso es añadir la velocidad al espacio de estados del objeto. Existen otro tipo de observaciones como el error en la detección de un objeto durante un determinado periodo de tiempo, para el cual la solución no es tan directa. Este tipo de observaciones pueden depender de toda la historia anterior de las posiciones del objeto.

Teniendo en cuenta estas dos asunciones es posible desarrollar completamente la expresión de la Ec. 3.5 como:

$$\begin{aligned}
p(\mathbf{s}_t | \mathbf{z}_t, \dots, \mathbf{z}_1) &= \frac{p(\mathbf{s}_t, \mathbf{z}_t, \dots, \mathbf{z}_1)}{p(\mathbf{z}_t, \dots, \mathbf{z}_1)} \\
&= \frac{p(\mathbf{z}_t | \mathbf{s}_t, \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) p(\mathbf{z}_{t-1}, \dots, \mathbf{z}_1)}{p(\mathbf{z}_t, \dots, \mathbf{z}_1)} \\
&= p(\mathbf{z}_t | \mathbf{s}_t) p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) \frac{p(\mathbf{z}_{t-1}, \dots, \mathbf{z}_1)}{p(\mathbf{z}_t, \dots, \mathbf{z}_1)} \\
&= \frac{p(\mathbf{z}_t | \mathbf{s}_t) p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1)}{\int p(\mathbf{z}_t | \mathbf{s}_t) p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_t},
\end{aligned} \tag{3.9}$$

y como el denominador no depende de \mathbf{s}_t :

$$p(\mathbf{s}_t | \mathbf{z}_t, \dots, \mathbf{z}_1) \propto p(\mathbf{z}_t | \mathbf{s}_t) p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1), \tag{3.10}$$

donde $p(\mathbf{z}_t | \mathbf{s}_t)$ es la función de *likelihood* y $p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1)$ el *prior* para la estimación del estado del objetivo en tiempo t . Este *prior* se puede interpretar como un paso de **predicción**, es decir, cuál será el estado más probable del objeto antes de que se obtengan las observaciones en tiempo t , \mathbf{z}_t . La función de *likelihood* se puede interpretar como un paso de **corrección** de la predicción a partir de los valores de las observaciones. Para hallar la forma recursiva del Filtraje Bayesiano, se desarrolla la expresión del *prior* utilizando la asunción de independencia condicionada entre estados y observaciones:

$$\begin{aligned}
p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) &= \int p(\mathbf{s}_t, \mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_{t-1} \\
&= \int p(\mathbf{s}_t | \mathbf{s}_{t-1}, \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_{t-1} \\
&= \int p(\mathbf{s}_t | \mathbf{s}_{t-1}) p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_{t-1}.
\end{aligned} \tag{3.11}$$

La expresión final que describe el proceso del Filtraje Bayesiano es la siguiente:

$$p(\mathbf{s}_t | \mathbf{z}_t, \dots, \mathbf{z}_1) \propto p(\mathbf{z}_t | \mathbf{s}_t) \int p(\mathbf{s}_t | \mathbf{s}_{t-1}) p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_{t-1}. \tag{3.12}$$

El algoritmo completo del Filtro Bayesiano se muestra en la Fig. 3.3. Para poder implementar el Filtro Bayesiano en una aplicación de seguimiento es necesaria la especificación de las funciones de densidad del proceso estocástico (modelo dinámico),

Filtro Bayesiano	
Prior: $p(\mathbf{s}_0)$.	
Para $t = 1 \dots T$:	
Predicción:	
$p(\mathbf{s}_t \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) = \int p(\mathbf{s}_t \mathbf{s}_{t-1}) p(\mathbf{s}_{t-1} \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) d\mathbf{s}_{t-1} \quad . \quad (3.13)$	
Corrección:	
$p(\mathbf{s}_t \mathbf{z}_t, \dots, \mathbf{z}_1) \propto p(\mathbf{z}_t \mathbf{s}_t) p(\mathbf{s}_t \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) \quad . \quad (3.14)$	

Figura 3.3: Algoritmo del Filtro Bayesiano.

$p(\mathbf{s}_t | \mathbf{s}_{t-1})$ y de la función de *likelihood* (modelo de medida), $p(\mathbf{z}_t | \mathbf{s}_t)$. El caso más simple es cuando las densidades de probabilidad son Gaussianas. Entonces es posible desarrollar la Ec. (3.12) de forma simple con un conjunto de ecuaciones lineales recursivas. Para modelos dinámicos no lineales, seguimiento de múltiples objetos y funciones de *likelihood* complejas es necesaria la utilización de métodos de Monte Carlo para poder calcular de forma eficiente el posterior.

3.2.3 Filtro de Kalman.

Sin duda alguna el filtro de estimación más utilizado en la práctica es el Filtro de Kalman [49]. A continuación se mostrará como este filtro es un caso particular del Filtro Bayesiano, en el cual se asumen densidades de probabilidad Gaussianas para el modelo dinámico y la función de *likelihood*.

Se asume el siguiente modelo dinámico:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}) \propto \exp \left\{ -\frac{1}{2} ((\mathbf{s}_t - \mathbf{A} \mathbf{s}_{t-1})^T \mathbf{Q}_t^{-1} (\mathbf{s}_t - \mathbf{A} \mathbf{s}_{t-1})) \right\} \quad , \quad (3.15)$$

donde \mathbf{A} es una matriz que define un movimiento determinista del objeto, y \mathbf{Q}_t es la matriz de covarianza de un término de movimiento aleatorio. Si el estado inicial del proceso dinámico, \mathbf{s}_0 , también es Gaussiano tenemos que el estado del objeto a lo largo del tiempo sigue una distribución Normal:

$$\mathbf{s}_t \sim N(\boldsymbol{\mu}_t, \mathbf{P}_t) \quad , \quad (3.16)$$

donde la media y la covarianza del estado del objeto, $\boldsymbol{\mu}_t$ y \mathbf{P}_t , pueden ser calculadas de forma recursiva como:

$$\begin{aligned}\boldsymbol{\mu}_t &= \mathbf{A}\boldsymbol{\mu}_{t-1} , \\ \mathbf{P}_t &= \mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T + \mathbf{Q}_t .\end{aligned}\quad (3.17)$$

Las observaciones son una combinación lineal del estado del objeto más un error de medida con densidad de probabilidad Gaussiana:

$$\mathbf{z}_t = \mathbf{M}\mathbf{s}_t + \boldsymbol{\omega}_t , \quad (3.18)$$

donde $\boldsymbol{\omega}_t \sim N(0, \mathbf{R}_t)$. Es posible ver que la función de *likelihood*, $p(\mathbf{z}_t|\mathbf{s}_t)$, también es Gaussiana ya que vuelve a ser una combinación lineal de variables Gaussianas.

A partir de las Ec. (3.15) y (3.18) se puede derivar el filtro de Kalman siguiendo las expresiones de la recursión del filtro Bayesiano:

$$p(\mathbf{s}_{t-1}|\mathbf{z}_{t-1}, \dots, \mathbf{z}_1) = \eta(\mathbf{s}_{t-1}; \boldsymbol{\mu}_{t-1}, \mathbf{P}_{t-1}) , \quad (3.19)$$

donde $\eta(\cdot)$ es una densidad de probabilidad Gaussiana. Aplicando la predicción de la Ec. (3.13) a la expresión (3.19) y utilizando la densidad del proceso dinámico Ec. (3.16) tenemos:

$$p(\mathbf{s}_t|\mathbf{z}_{t-1}, \dots, \mathbf{z}_1) = \eta(\mathbf{s}_t^-; \boldsymbol{\mu}_t^-, \mathbf{P}_t^-) , \quad (3.20)$$

donde:

$$\begin{aligned}\boldsymbol{\mu}_t^- &= \mathbf{A}\boldsymbol{\mu}_{t-1} , \\ \mathbf{P}_t^- &= \mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T + \mathbf{Q}_t .\end{aligned}\quad (3.21)$$

Y una vez recogida la nueva observación, se obtiene:

$$p(\mathbf{s}_t|\mathbf{z}_t, \dots, \mathbf{z}_1) = \eta(\mathbf{s}_t; \boldsymbol{\mu}_t, \mathbf{P}_t) , \quad (3.22)$$

donde:

$$\begin{aligned}\boldsymbol{\mu}_t &= \boldsymbol{\mu}_t^- + \mathbf{P}_t^- \mathbf{M}^T (\mathbf{M}\mathbf{P}_t^- \mathbf{M}^T + \mathbf{R}_t)^{-1} (\mathbf{z}_t - \mathbf{M}\boldsymbol{\mu}_t^-) , \\ \mathbf{P}_t &= \mathbf{P}_t^- - \mathbf{P}_t^- \mathbf{M}^T (\mathbf{M}\mathbf{P}_t^- \mathbf{M}^T + \mathbf{R}_t)^{-1} \mathbf{M}\mathbf{P}_t^- .\end{aligned}\quad (3.23)$$

Este resultado es el mismo que el obtenido en la definición original del Filtro de Kalman, ver Fig. 3.4, y muestra como es un caso particular del Filtraje Bayesiano.

La restricción de que la forma de las densidades de probabilidad sean Gaussianas impide que el Filtro de Kalman pueda ser utilizado directamente en el seguimiento de múltiples objetos. Como veremos más adelante, en el caso de múltiples objetos es necesaria la utilización de tantos filtros como objetos se desea seguir y un proceso previo de asociación de datos cuyo objetivo básico es la puesta en correspondencia

Filtro de Kalman	
Prior: $s_0 \sim N(\boldsymbol{\mu}_0, P_0)$.	
Para $t = 1 \dots T$:	
Predicción:	$\begin{aligned} \boldsymbol{\mu}_t^- &= A\boldsymbol{\mu}_{t-1} , \\ P_t^- &= AP_{t-1}A^T + Q_t . \end{aligned} \tag{3.24}$
Corrección:	$\begin{aligned} K_t &= P_t^- M^T (MP_t^- M^T + R_t)^{-1} , \\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_t^- + K_t(z_t - M\boldsymbol{\mu}_t^-) , \\ P_t &= P_t^- - K_t M P_t^- . \end{aligned} \tag{3.25}$

Figura 3.4: Algoritmo del Filtro de Kalman.

de cada medida con el filtro al que es más probable que pertenezca. Además, la unimodalidad de la densidad Gaussiana no permite la generación de múltiples hipótesis del estado de un objeto, por lo que en el caso de realizar el seguimiento en un entorno complejo, si ocurre un error grave en el proceso de medida el Filtro de Kalman no puede recuperarse.

3.3 Modelos dinámicos.

Consideremos el problema de seguir la posición de una persona en una secuencia exterior. Existen zonas de la imagen por las que seguro que no pasará la persona, depende del entorno de la secuencia. Este *prior* restringe el conjunto de posibles posiciones de la imagen en las que puede estar la persona. Sin embargo, esta condición es muy pobre y reduce poco el conjunto de casos posibles en los que puede estar la persona. Además no se utiliza el hecho de que sabemos que la persona realiza un tipo de movimiento característico para disminuir el conjunto de posiciones en las que se puede encontrar si conocemos la posición en el instante de tiempo anterior.

Por tanto, en un sistema de seguimiento visual, la utilización de un modelo del movimiento del objeto acota el conjunto de estados posibles, lo que no permite que un error grave en el proceso de observación provoque que el sistema falle. Además, proporciona una probabilidad previa que actúa a nivel de predicción del nuevo estado del objeto.

Dentro del entorno probabilístico descrito con anterioridad, el modelo dinámico se representa por la densidad de probabilidad:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}, \dots, \mathbf{s}_1) ,$$

donde debido a la asunción de Markov queda:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}) .$$

En esta sección veremos los diferentes modelos dinámicos que se utilizan en la mayoría de aplicaciones de visión por computador. En general, se utilizan modelos dinámicos lineales, lo que implica que es posible expresarlos de forma probabilística por medio de densidades de probabilidad Gaussianas. Estos procesos estocásticos se denominan procesos de Gauss-Markov. De esta forma es posible utilizar el Filtro de Kalman para realizar el seguimiento visual. Como se muestra en la Fig. 3.2 estos modelos han de sintonizar sus parámetros para adecuarlos al tipo de movimiento del objeto dentro de la escena. Por tanto, también se repasará como es posible encontrar estos parámetros a partir de un aprendizaje previo del movimiento del objeto.

3.3.1 Procesos de Gauss-Markov.

El modelo dinámico más simple para un proceso estocástico de Markov es un movimiento aleatorio:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}) \propto \exp \left\{ -\frac{1}{2} ((\mathbf{s}_t - \mathbf{s}_{t-1})^T \mathbf{Q}_t^{-1} (\mathbf{s}_t - \mathbf{s}_{t-1})) \right\} . \quad (3.26)$$

Por ejemplo, en un conjunto de estados que representa posiciones 2D, este proceso representa una trayectoria aleatoria bidimensional de un objeto. En esta trayectoria el paso entre frames es isotrópico (posible en cualquier dirección), ver Fig. 3.5 (a). Esta forma particular de movimiento aleatorio se conoce como **movimiento Browniano**⁴. La representación en forma generativa del movimiento aleatorio sería:

$$\mathbf{s}_t = \mathbf{s}_{t-1} + \boldsymbol{\xi}_t ,$$

donde $\boldsymbol{\xi}_t \sim N(0, \mathbf{Q}_t)$. Este modelo suele emplearse cuando el movimiento es totalmente impredecible. Por ejemplo, en aplicaciones de seguimiento de rasgos faciales para reconocimiento de expresiones [66].

Sin embargo, es más usual observar un desplazamiento determinista, \mathbf{d} , en el movimiento aleatorio:

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}) \propto \exp \left\{ -\frac{1}{2} ((\mathbf{s}_t - \mathbf{s}_{t-1} - \mathbf{d})^T \mathbf{Q}_t^{-1} (\mathbf{s}_t - \mathbf{s}_{t-1} - \mathbf{d})) \right\} . \quad (3.27)$$

En el caso de traslaciones en el espacio, \mathbf{d} es un vector que representa el desplazamiento medio en cada instante de tiempo, ver Fig. 3.5 (b). Después de T instantes de

⁴Conocido como modelo físico de la termodinámica de partículas microscópicas.

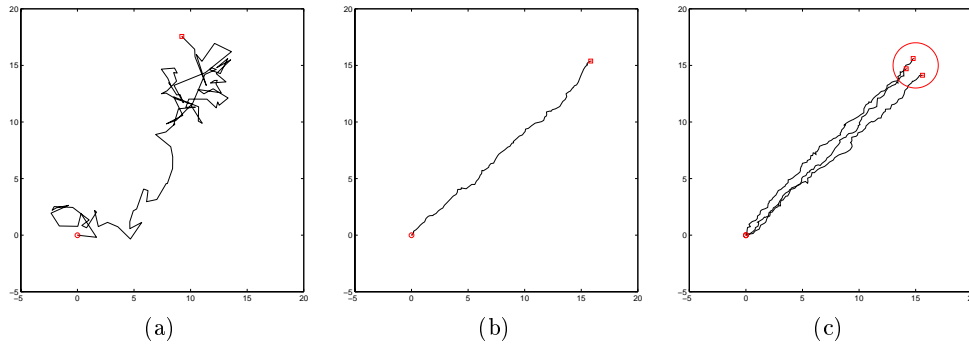


Figura 3.5: Movimiento aleatorio en un plano. (a) Movimiento completamente aleatorio ($T = 100$, $Q = \mathbf{I}$ y $\sigma = 1$). (b) Movimiento aleatorio con un desplazamiento determinista ($T = 100$, $\sigma = 0.1$, $\mathbf{d} = (0.15, 0.15)$). (c) Varias trayectorias del movimiento aleatorio con desplazamiento determinista; los puntos finales de las trayectorias tienen una densidad Gaussiana (en la gráfica se muestra un intervalo de confianza de la densidad Gaussiana del 95 %).

tiempo, el punto final de las trayectorias de un modelo aleatorio, comenzando desde el mismo origen, queda distribuido en el plano de forma Gaussiana, ver Fig. 3.5 (c). Su distribución es la suma de T Gaussianas, cuya covarianza crece en cada paso de tiempo en una proporción de \sqrt{T} . Esto provoca una incerteza mayor en la posición del objeto a medida que avanza el tiempo.

Este proceso se puede ver como un modelo dinámico de velocidad constante, simplemente se aumenta la representación del estado, $\mathbf{s} = (\mathbf{x}, \mathbf{u})^T$, donde \mathbf{x} es la posición del objeto y \mathbf{u} su velocidad. En este caso: $\mathbf{x}_t = \mathbf{x}_{t-1} + (\Delta t)\mathbf{u}_{t-1}$ y $\mathbf{u}_t = \mathbf{u}_{t-1}$. Si se sustituye la velocidad en la ecuación de la posición tenemos que $\mathbf{d} = (\Delta t)\mathbf{u}_{t-1}$ y $\mathbf{u}_t = \mathbf{d}(\Delta t)^{-1}$. El sistema final quedaría:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{d} + \boldsymbol{\xi}_t ,$$

y la estimación de la velocidad no tendría sentido porque se supone constante. El problema de este modelo es que no considera cambios de velocidad y que no tiene en cuenta posibles movimientos oscilatorios. Sin embargo, es un modelo útil cuando se tiene una frecuencia alta de adquisición de observaciones ya que es posible suponer que la componente de aceleración queda incluida en el movimiento aleatorio, $\boldsymbol{\xi}_t$.

El modelo dinámico de Gauss-Markov general que incluye dinámicas con variación de velocidad y movimientos oscilatorios es un proceso Auto-regresivo de segundo orden, cuya expresión es:

$$\mathbf{s}_t - \bar{\mathbf{s}} = \mathbf{A}_1(\mathbf{s}_{t-1} - \bar{\mathbf{s}}) + \mathbf{A}_2(\mathbf{s}_{t-2} - \bar{\mathbf{s}}) + \boldsymbol{\xi}_t , \quad (3.28)$$

donde para cumplir la asunción de Markov:

$$\mathbf{S}_t = \begin{pmatrix} s_t \\ s_{t-1} \end{pmatrix} ,$$

y se reescribe como un proceso de primer orden:

$$\mathbf{S}_t - \bar{\mathbf{S}} = \mathbf{A}(\mathbf{S}_{t-1} - \bar{\mathbf{S}}) + \boldsymbol{\Sigma}_t , \quad (3.29)$$

donde

$$\mathbf{A} = \begin{pmatrix} A_1 & A_2 \\ \mathbf{I} & 0 \end{pmatrix} , \quad \bar{\mathbf{S}} = \begin{pmatrix} \bar{s} \\ \bar{s} \end{pmatrix} , \quad \text{y} \quad \boldsymbol{\Sigma}_t = \begin{pmatrix} \xi_t \\ 0 \end{pmatrix} .$$

El parámetro $\bar{\mathbf{S}}$ se interpreta como el valor estacionario de la media del estado del objeto. La función de este parámetro se puede ver de forma clara en el caso de seguimiento visual de formas [11], ya que representa la forma media del objeto. Unido a la condición de que $\|\mathbf{A}\|^2 < 1$, que garantiza la estabilidad del sistema, hace que no aumente la incerteza de la forma como ocurría con el movimiento aleatorio determinista conforme avanza el tiempo. Por tanto, se divide el modelo dinámico de la forma en una parte con movimiento aleatorio determinista para los desplazamientos y una parte de deformación dinámica de la forma controlada por los parámetros \mathbf{A} y $\bar{\mathbf{S}}$.

Por otro lado, si el estado se representa por un escalar y la matriz \mathbf{A} tiene la forma:

$$\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} ,$$

nos encontramos con un modelo de aceleración constante. Modelo dinámico más utilizado en aplicaciones de visión que utilizan el esquema simple de seguimiento visual[16, 40, 34, 69].

Finalmente, es posible expresar la Ec. (3.29) como:

$$\mathbf{S}_t = \mathbf{A}\mathbf{S}_{t-1} + \mathbf{D} + \boldsymbol{\Sigma}_t , \quad (3.30)$$

donde

$$\mathbf{D} = (\mathbf{I} - \mathbf{A})\bar{\mathbf{S}} ,$$

lo que significa que también podemos representar el movimiento aleatorio determinista con la expresión del modelo Auto-regresivo de orden 2.

Por último, lo único que nos queda por comentar es como determinar la forma de las matrices \mathbf{A} , \mathbf{D} y \mathbf{Q} . La más sencilla es conocer el tipo de movimiento del objetivo. Si no es posible saberlo a priori, necesitamos aprender el movimiento a partir de secuencias de ejemplo.

3.3.2 Aprendizaje de movimiento.

En la práctica, los parámetros de los modelos dinámicos suelen establecerse inicialmente de forma manual. Después, durante una fase de pruebas, se refinan estos parámetros para mejorar los resultados del seguimiento. Sin embargo, si se dispone de datos del movimiento típico del objeto (por ejemplo marcados de forma manual) o de secuencias donde el objeto de interés es fácil de localizar (si se dispone de un entorno controlado de prueba), es posible realizar un aprendizaje de estos parámetros.

Los métodos de aprendizaje más utilizados se basan en el principio de *maximum likelihood* [6]. Este principio se basa en buscar los valores óptimos de los parámetros que maximizan una función de *likelihood*. Suponer que queremos encontrar los parámetros, θ , de la densidad de probabilidad que ha generado los datos, $p(\mathbf{s})$. La función de *likelihood* es la encargada de hacer explícita la dependencia de los parámetros:

$$\mathcal{L}(\theta) \equiv p(\mathbf{s}|\theta) .$$

Por ejemplo, considerar que se dispone de una secuencia de M posiciones correctas del movimiento de un objeto, $\{\mathbf{s}_1, \dots, \mathbf{s}_M\}$, y que el modelo dinámico puede describirse como un movimiento aleatorio determinista, Ec. (3.27). Entonces es posible escribir el proceso dinámico como:

$$\mathbf{s}_t = \mathbf{s}_{t-1} + \mathbf{d} + \boldsymbol{\xi}_t ,$$

e incluir el desplazamiento determinista dentro de la distribución del movimiento aleatorio, para obtener la siguiente densidad de probabilidad:

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}) = \eta(\mathbf{s}_t - \mathbf{s}_{t-1}; \mathbf{d}, \mathbf{Q}) ,$$

donde $\eta(\cdot)$ es una densidad Gaussiana de media \mathbf{d} y matriz de covarianza \mathbf{Q} . De esta forma si expresamos las muestras como $\mathbf{u}_t = \mathbf{s}_t - \mathbf{s}_{t-1}$ tenemos que $\mathbf{u}_t \sim N(\mathbf{d}, \mathbf{Q})$, y la función de *likelihood* queda:

$$\mathcal{L}(\mathbf{d}, \mathbf{Q}) = p(\mathbf{u}_t|\mathbf{d}, \mathbf{Q}) . \quad (3.31)$$

Como $p(\mathbf{u}_t|\mathbf{d}, \mathbf{Q})$ es una densidad Gaussiana es posible encontrar los parámetros que maximizan la función de *likelihood* de forma analítica:

$$\hat{\mathbf{d}} = \frac{1}{M-1} \sum_{i=1}^{M-1} \mathbf{u}_i , \quad (3.32)$$

$$\hat{\mathbf{Q}} = \frac{1}{M-1} \sum_{i=1}^{M-1} (\mathbf{u}_i - \hat{\mathbf{d}})(\mathbf{u}_i - \hat{\mathbf{d}})^T . \quad (3.33)$$

Aprendizaje de modelos dinámicos

Dado un conjunto de M muestras del estado de un objeto, $\{s_1, s_2, \dots, s_M\}$, encontrar los parámetros A_1 , A_2 , D y Q para un proceso Auto-regresivo de segundo orden que describe la dinámica del objeto.

Algoritmo

1. Calcular los siguientes coeficientes para $i, j = 0, 1, 2$:

$$R_i = \sum_{k=3}^M s_{k-i}, \quad R_{ij} = \sum_{k=3}^M s_{k-i} s_{k-j}^T, \quad R'_{ij} = R_{ij} - \frac{1}{M-2} R_i R_j^T .$$

2. La estimación de los parámetros A_1 , A_2 y D viene dada por:

$$\begin{aligned} \hat{A}_2 &= (R'_{02} - R'_{01}(R'_{11})^{-1}R'_{12})(R'_{22} - R'_{21}(R'_{11})^{-1}R'_{12})^{-1} . \\ \hat{A}_1 &= (R'_{01} - \hat{A}_2 R'_{21})(R'_{11})^{-1} . \\ \hat{D} &= \frac{1}{M-2} (R_0 - \hat{A}_2 R_2 - \hat{A}_1 R_1) . \end{aligned}$$

3. Si se requiere la forma estándar del proceso auto-regresivo, el parámetro \bar{s} es estimado como:

$$\hat{\bar{s}} = (I - \hat{A}_2 - \hat{A}_1)^{-1} \hat{D} .$$

4. Por último, la estimación del parámetro Q es:

$$\hat{Q} = \frac{1}{M-2} (R_{00} - \hat{A}_2 R_{20} - \hat{A}_1 R_{10} - \hat{D} R_0^T) .$$

Figura 3.6: Algoritmo de *maximum likelihood* para el aprendizaje de los parámetros de un proceso AR de orden 2 [11].

También es posible utilizar el método de *maximum likelihood* para encontrar los parámetros del modelo general de un proceso de Gauss-Markov, ver Fig. 3.6. El desarrollo completo y la demostración del algoritmo se puede encontrar en [11]. Baumberg y Hogg desarrollaron un algoritmo relacionado pero que sólo encuentra los parámetros de movimiento determinista de la forma [5].

Todos estos métodos se basan en una obtención previa de las muestras del estado del objeto. Sin embargo, en la práctica, estas muestras son observaciones obtenidas de secuencias de ejemplo más sencillas. Entonces el problema se puede replantear para hacer el aprendizaje del movimiento teniendo en cuenta que las muestras son observaciones [64, 65]. Tratar las muestras como observaciones permite aprender el movimiento a la vez que se realiza el seguimiento visual.

Existen otro tipo de modelos de movimiento más complejos, utilizados en aplicaciones de captura de movimiento humano[48, 17, 91, 1, 26] o reconocimiento de actividades[92]. En estos casos se buscan las restricciones del movimiento de partes del cuerpo humano para acotar el conjunto de estados posibles.

3.4 Modelado de la función de *likelihood*.

Para poder aplicar el Filtraje Bayesiano, además de especificar un modelo estocástico del movimiento del objeto es necesario el modelado del proceso de medida, es decir, de la función de *likelihood*, $p(\mathbf{z}_t|\mathbf{s}_t)$. El modelado de la función de *likelihood* es muy dependiente de la representación del objeto a seguir[9, 44, 43, 7, 57, 80, 45]. Lo que es importante resaltar es que en la práctica esta función suele ser multimodal debido a la presencia de falsos positivos. En estos casos, el Filtro de Kalman no funcionaría correctamente. A continuación se muestra un ejemplo sencillo de modelado de la función de *likelihood* que es suficiente para explicar esta característica.

En el trabajo original de Isard y Blake [42], se asume un proceso de observación estacionario⁵, por tanto:

$$p(\mathbf{z}_t|\mathbf{s}_t) = p(\mathbf{z}|\mathbf{s}) .$$

Se definirá el modelo para variables unidimensionales, es decir, las medidas obtenidas en tiempo t se reducen a un conjunto de posiciones escalares: $\mathbf{z}_t = \{z_1, z_2, \dots, z_M\}$, donde M es el número de medidas. La densidad de probabilidad del proceso de medida tiene la forma $p(\mathbf{z}|\mathbf{s})$, donde \mathbf{s} es una posición 1-dimensional. Las múltiples medidas se deben a la presencia de falsos positivos, de manera que sólo una de las medidas es la correspondiente al objeto de interés:

$$\phi_m = \{\text{la verdadera medida es } z_m\}, \quad m = 1, \dots, M .$$

⁵Aunque el algoritmo del filtraje Bayesiano no lo requiere.

Por otro lado también es posible que el objeto no sea visible, es decir, que ninguna de las medidas obtenidas sea la correcta:

$$q = 1 - \sum_{m=1}^M p(\phi_m) . \quad (3.34)$$

Este razonamiento es común en el seguimiento de objetos [82, 10]. Ahora la función de *likelihood* se puede expresar como:

$$p(z|s) = q p(z|\text{falso positivo}) + \sum_{m=1}^M p(z|s, \phi_m) p(\phi_m) . \quad (3.35)$$

Una forma funcional razonable para esta expresión puede obtenerse haciendo algunas asunciones: la primera es que todas las medidas tienen la misma probabilidad de ser correctas, es decir, $p(\phi_m) = r, \forall m$; la segunda, que los falsos positivos siguen una distribución de Poisson con densidad espacial igual a λ ; y finalmente, que cualquier medida no tiene tendencia y tiene una distribución normal con varianza σ^2 . Esto nos lleva a la siguiente expresión:

$$p(z|s) \propto 1 + \frac{1}{\sqrt{2\pi\sigma\alpha}} \sum_{m=1}^M \exp \left\{ -\frac{(z_m - s)^2}{2\sigma^2} \right\} , \quad (3.36)$$

donde $\alpha = q\lambda$. Los picos en esta función de densidad corresponderán a las características medidas, y la densidad del posterior tenderá a reforzarse en estos puntos, ver Fig. 3.7.

El nivel de falsos positivos refleja la probabilidad de que el verdadero objeto no sea detectado. El efecto en el seguimiento es que se permite la posibilidad de “*tunneling*”: una buena hipótesis podría sobrevivir a un fallo transitorio de las observaciones, por ejemplo debido a una desaparición momentánea. Los parámetros σ , en unidades de distancia, y α , en unidades de distancia inversa, pueden sintonizarse a partir de la observación del error de medida, la densidad de los falsos positivos, λ , y la probabilidad de no detección, q .

En la práctica, dada una hipotética posición (del proceso de predicción), no es necesario evaluar todas las medidas encontradas, podemos eliminar aquellas para las cuales:

$$\frac{1}{\sqrt{2\pi\sigma\alpha}} \exp \left\{ -\frac{(z_m - s)^2}{2\sigma^2} \right\} \ll 1 .$$

Es como si se establece una ventana de búsqueda⁶ alrededor de la posición s fuera de la cuál las medidas pueden ser ignoradas.

⁶Esta ventana se denomina *gate* [10].

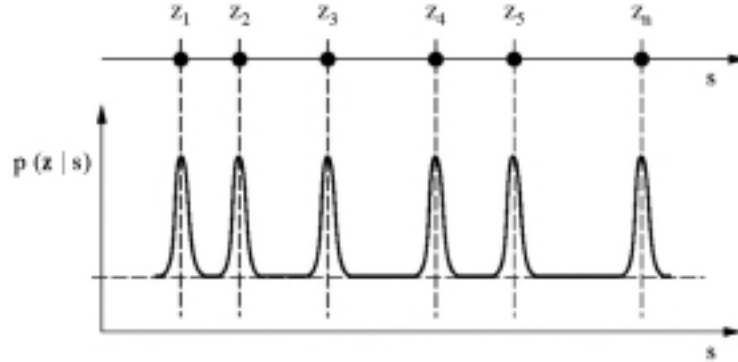


Figura 3.7: Ejemplo de función de *likelihood* para variables unidimensionales.

En muchas aplicaciones prácticas, los falsos positivos suelen estar bastante dispersos y σ es suficientemente pequeña para no incluir más de una medida en la ventana de búsqueda, con lo que la densidad se convierte en unimodal y es posible la utilización del Filtro de Kalman. Sin embargo en entornos complejos estas dos asunciones no son correctas.

Por último, un aspecto a tener en cuenta es el estudio del evento “hay M medidas”, ψ_M . Si se asume que s siempre permanece en la región de búsqueda (región cubierta por los sensores), el evento ψ_M no proporciona información adicional al estado del objeto. Sin embargo, si se permite que el objeto esté fuera de la región de interés, el evento ψ_M sí es informativo: un valor de M por encima del valor medio de la probabilidad de falsos positivos aumenta la probabilidad de que s esté dentro de la región de búsqueda.

3.5 Filtraje Bayesiano con partículas.

La principal dificultad del Filtraje Bayesiano es su implementación, en el sentido de mantener una buena representación del posterior, $p(s_t|z_t, \dots, z_1)$, en presencia de funciones de *likelihood* multimodales y dinámicas no lineales, ya que en estos casos no es posible utilizar la simplificación que se asume en el Filtro de Kalman. Esta representación debe ser capaz de manejar distribuciones multimodales y también ha de poder trabajar con un vector de estados multidimensional. No existe una solución satisfactoria a este problema. En esta sección se muestra la aproximación más utilizada en la mayoría de aplicaciones: la representación muestral de densidades de probabilidad[77, 28, 42].

3.5.1 Representación muestral de densidades de probabilidad.

Considerar que se desea encontrar la representación de la siguiente densidad de probabilidad:

$$p(s) = p_1(s)p_2(s) , \quad (3.37)$$

donde $p_2(s)$ es una densidad de probabilidad que es posible simular, es decir, que se puede representar por medio de un conjunto de N muestras:

$$\{s^i\} \quad i = 1, \dots, N .$$

Si $p(s)$ no se puede aproximar directamente por un conjunto de muestras es posible utilizar el esquema del *factored sampling*[28] para encontrar su representación muestral como:

$$\{(s^i, \pi^i)\} \quad i = 1, \dots, N ,$$

donde cada π^i es el **peso** de cada muestra, s^i , y se encuentra evaluando la primera densidad de probabilidad, $p_1(s)$, para cada muestra como sigue:

$$\pi^i = \frac{p_1(s^i)}{\sum_{j=1}^N p_1(s^j)} , \quad (3.38)$$

Finalmente, se realiza un proceso de selección de muestras, s^i , que representarán a la densidad $p(s)$ más correctamente cuanto mayor sea N . Para el proceso de selección, se utilizan los pesos como probabilidad de escoger una muestra.

La aplicación directa de este método se encuentra en la evaluación del Teorema de Bayes:

$$p(s|z) = kp(z|s)p(s) , \quad (3.39)$$

donde para encontrar la representación del posterior se necesita una densidad *prior* que sea posible muestrear y una densidad $p(z|s)$ que sea posible evaluar en s^i para encontrar los valores de los pesos:

$$p(s|z) \approx \{(s^i, \pi^i)\} \quad i = 1, \dots, N \quad \text{donde} \quad \pi^i = \frac{p(z|s = s^i)}{\sum_{j=1}^N p(z|s = s^j)} .$$

El Filtro de Partículas es la extensión temporal del método de *factored sampling*.

3.5.2 Extensión temporal: Filtro de Partículas

En el problema del cálculo de la expresión del Filtraje Bayesiano, Ec. (3.12), el objetivo es hallar la densidad $p(\mathbf{s}_t | \mathbf{z}_t, \dots, \mathbf{z}_1)$ a partir de la densidad en el instante de tiempo anterior, $p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1)$. Para realizar la explicación del algoritmo seguiremos los pasos típicos de los filtros de estimación, es decir, predicción y corrección.

Suponer que se dispone de una representación muestral de esta última densidad:

$$\{\mathbf{s}_{t-1}^i\} \quad i = 1, \dots, N ,$$

donde el superíndice indica el número de muestra, y el subíndice el instante de tiempo. Se puede ver cada muestra de $p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1)$ como un posible estado del objeto en tiempo $t - 1$.

Predicción

Para cualquier muestra, \mathbf{s}_{t-1}^i , pueden obtenerse muestras de $p(\mathbf{s}_t | \mathbf{s}_{t-1} = \mathbf{s}_{t-1}^i)$ ya que se supone que se dispone del modelo dinámico:

$$\mathbf{s}_t = \mathbf{f}(\mathbf{s}_{t-1}) + \boldsymbol{\xi}_i ,$$

donde $\mathbf{f}(\cdot)$ es la componente determinista y $\boldsymbol{\xi}_i$ la componente aleatoria del modelo dinámico. Para obtener las nuevas muestras de $p(\mathbf{s}_t | \mathbf{s}_{t-1} = \mathbf{s}_{t-1}^i)$ simplemente tenemos que aplicar el modelo dinámico:

$$\{(\mathbf{f}(\mathbf{s}_{t-1}^i) + \boldsymbol{\xi}_i^l)\} ,$$

el superíndice l indica que podemos obtener varias muestras para cada \mathbf{s}_{t-1}^i , debido a la componente aleatoria del modelo dinámico.

Por otro lado, se cumple que:

$$p(\mathbf{s}_t, \mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) = p(\mathbf{s}_t | \mathbf{s}_{t-1}) p(\mathbf{s}_{t-1} | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1) ,$$

que podemos representar por:

$$\{(\mathbf{f}(\mathbf{s}_{t-1}^i) + \boldsymbol{\xi}_i^l, \mathbf{s}_{t-1}^i)\} .$$

Como podemos marginalizar la densidad conjunta eliminando elementos de la representación muestral, la representación de $p(\mathbf{s}_t | \mathbf{z}_{t-1}, \dots, \mathbf{z}_1)$ viene dada por:

$$\{(\mathbf{f}(\mathbf{s}_{t-1}^i) + \boldsymbol{\xi}_i^l)\} .$$

Si reindexamos todas las muestras que tenemos⁷, la representación del *prior* queda:

$$\{\mathbf{s}_t^{i,-}\} ,$$

donde el superíndice '-' indica que es nuestra representación del estado del proceso antes de que se tomen las medidas.

Corrección

La corrección es simple, la predicción actúa como *prior* y tenemos que convertirlo en posterior. Esto se realiza escogiendo un peso apropiado para cada muestra:

$$\pi_t^i = p(z_t | \mathbf{s}_t = \mathbf{s}_t^{i,-}) ,$$

quedando la representación del posterior como:

$$\{(\mathbf{s}_t^{i,-}, \pi_t^i)\} \quad i = 1, \dots, N .$$

Finalmente, se seleccionan N muestras, utilizando como probabilidad de selección el peso de cada muestra, para hallar la representación final del posterior:

$$\{\mathbf{s}_t^i\} \quad i = 1, \dots, N .$$

Para acabar de especificar el algoritmo completo del Filtro de Partículas, ver Fig. 3.8, sólo queda especificar la forma del proceso de selección de muestras. La idea es utilizar la suma de los pesos:

$$\begin{aligned} c_t^0 &= 0 , \\ c_t^i &= c_t^{i-1} + \pi_t^i \quad \forall i = 1, \dots, N . \end{aligned}$$

como función de distribución para generar las nuevas muestras:

1. Generar un número aleatorio $r \in [0, 1]$ con distribución uniforme.
2. Encontrar, por medio de subdivisión binaria, el número de muestra j más pequeño para el cual $c_t^j \geq r$.
3. Seleccionar $\mathbf{s}^i = \mathbf{s}^{j,-}$.

Este algoritmo causa que el algoritmo completo tenga una complejidad $O(n \log n)$.

⁷Que en este momento pueden ser más de N .

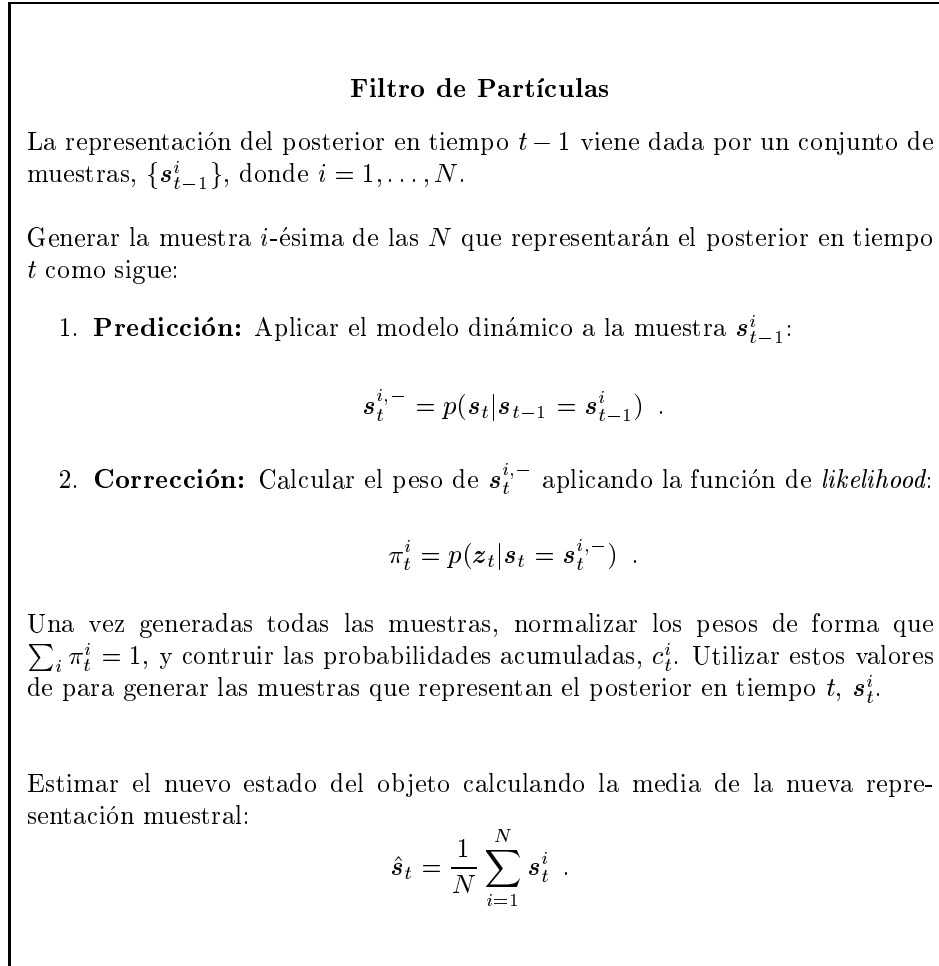


Figura 3.8: Algoritmo genérico del Filtro de Partículas.

Para disminuir la complejidad del algoritmo es posible utilizar un algoritmo determinista de selección de muestras, ver Fig. 3.9. El algoritmo determinista tiene complejidad $O(n)$.

3.5.3 Aspectos prácticos.

Los Filtros de Partículas han sido probados con éxito en varias aplicaciones prácticas de visión [42, 44, 43, 9, 95, 7, 57], pero también pueden producir resultados no deseados.

Un aspecto práctico importante es escoger el número de partículas; mientras que el valor esperado de una densidad estimada con una representación muestral es el

Inicializar: $j = 1$. **Para** $i = 1 \dots N$:
Mientras $(\tilde{c}_t^j < i)$ $j++$;
 Generar $\mathbf{s}_t^i = \mathbf{s}_t^{j,-}$
 donde:

$$c_t^j = \sum_{n=1}^j \pi_t^n \quad \tilde{c}_t^j = \lfloor c_t^j N \rfloor .$$

Figura 3.9: Algoritmo determinista para la selección de las muestras utilizadas para representar la densidad del posterior[43].

valor verdadero de la integral, puede ser necesaria una gran cantidad de muestras antes de que la varianza del estimador sea suficientemente pequeña. En la práctica, la elección del número de muestras sólo puede resolverse de forma experimental.

Es posible ver el Filtro de Partículas como un algoritmo de búsqueda, donde tenemos una serie de estimaciones del verdadero estado de un proceso, que ajustamos utilizando un modelo del proceso estocástico y que finalmente comparamos con observaciones. La dificultad está en que la función de *likelihood* en los problemas de visión suele tener muchos picos estrechos. Las predicciones de las muestras del estado pueden entrar en algunos de ellos pero no en todos, perdiendo hipótesis que pueden provocar el error en la estimación del estado. Este problema aún se hace más importante en espacios de dimensiones grandes.

3.6 Seguimiento de múltiples objetos.

La Fig. 3.10 muestra una representación de los elementos funcionales de un sistema típico de seguimiento de múltiples objetos. En esta representación aparecen dos nuevos elementos:

- La asociación entre cada observación y los objetos que están siendo seguidos.
- La gestión de eventos como aparición y desaparición de objetos.

El proceso de asociación de datos utiliza la predicción de los estados de los objetos que están siendo seguidos para definir una ventana de búsqueda de observaciones, *gate*, para cada objeto. Esta ventana se utiliza para eliminar aquellas observaciones que son menos probables que pertenezcan al objeto que está siendo seguido. Una vez escogidas las observaciones que pertenecen a la ventana de cada objeto, existen varios métodos que podemos aplicar para utilizarlas en su estimación.

Normalmente, es útil definir una distancia entre las observaciones y la predicción del estado de un objeto. Por ejemplo, para el método más simple, el vecino más cercano (NN). Esta aproximación determina una única asignación de forma que como máximo se utiliza una observación para corregir un estado. Existe una variación de

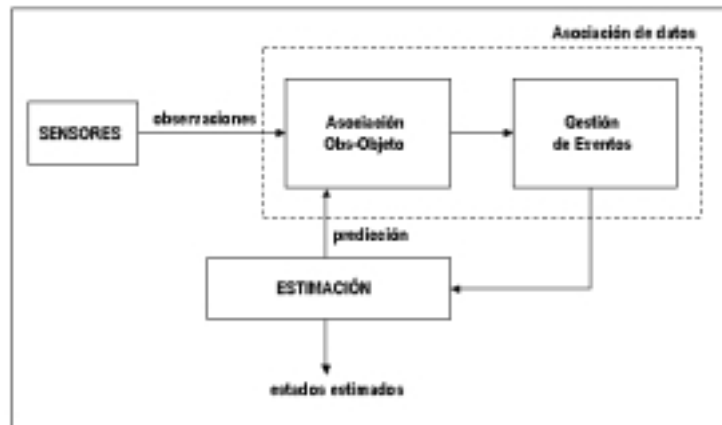


Figura 3.10: Esquema básico de un sistema de seguimiento de múltiples objetos.

este método, el vecino más cercano global (GNN). En el vecino más cercano se considera cada predicción por separado y se busca la observación con mínima distancia, esto puede provocar que la misma observación se utilice para ajustar dos predicciones. En el método GNN una observación sólo puede utilizarse en la estimación de un sólo objeto.

La alternativa a estos dos métodos es utilizar todas las observaciones de la ventana de búsqueda definida por la predicción para estimar el estado del objeto. De esta manera, una observación puede utilizarse en la estimación de múltiples objetos. La idea es utilizar una suma ponderada, determinada de forma probabilística, de todas las observaciones de la ventana de búsqueda. Este método de asociación de datos probabilístico (PDA) se desarrolló originalmente en el seguimiento de un sólo objeto en presencia de falsos positivos. La extensión a múltiples objetos recibe el nombre de asociación de datos probabilística conjunta (JPDA, *Joint Probabilistic Data Association*).

Una última posibilidad es la utilización de métodos de seguimiento múltiple de objetos sin una asociación explícita entre observaciones y predicciones. En este caso el filtro de estimación ha de ser capaz de propagar toda la densidad de probabilidad que representa la estimación del estado de todos los objetos para cada paso del algoritmo. Como hemos visto anteriormente, el Filtro Bayesiano no asume una forma explícita de densidad de probabilidad. Además, se ha visto como es posible implementarlo de forma computacional por medio del Filtro de Partículas. Por tanto, consideramos razonable plantearnos el objetivo de utilizar este filtro en el seguimiento visual de múltiples objetos.