

APÉNDICE L

L. FUNDAMENTOS DEL PROCESO DE ETIQUETADO MEDIANTE RELAJACIÓN

La teoría de relajación probabilista es aplicable a cualquier problema de correspondencia en general, por lo que en particular y en el campo de la visión, se puede aplicar tanto a la visión estereoscópica para correspondencia de características como al reconocimiento de patrones entre los modelos de una base de datos y los objetos a reconocer.

El problema se formula en el marco de la teoría de la probabilidad Bayesiana, siendo el objetivo la asignación de etiquetas a los objetos procedentes de la escena. Dichas etiquetas están asociadas a los modelos almacenados en una base de datos. En el proceso están involucradas tanto medidas de tipo unitario, relacionadas con las entidades simples, como medidas de tipo binario, que relacionan pares de objetos.

Vamos a definir y estructurar el procedimiento para llevar a cabo el objetivo. Sea $A = \{a_1, a_2, \dots, a_N\}$ el conjunto de N objetos de la escena que queremos etiquetar, es decir determinar cuál es el modelo de la base de datos más similar a cada uno de dichos objetos. Por tanto, a cada objeto a_i se le debe asignar una etiqueta θ_i , cuyo valor es cualquiera de las M etiquetas del conjunto de modelos Ω : $\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$. La notación ω_{θ_i} indica que una determinada etiqueta del conjunto de modelos se asocia con una etiqueta θ_i concreta de la escena. Al final del proceso se pretende que cada objeto tenga asignada una única etiqueta. Por conveniencia se definen dos conjuntos de índices:

$N_0 \equiv \{1, 2, \dots, N\}$ y $N_i \equiv \{1, 2, \dots, i-1, i+1, \dots, N\}$. Para cada objeto a_i se dispone de un conjunto de m_1 medidas \mathbf{x}_i , que corresponden a atributos asociados a cada objeto, son las medidas de tipo unitario: $\mathbf{x}_i = \{x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(m_1)}\}$. La abreviatura $\mathbf{x}_{i, i \in N_0}$ indica los vectores de todas las medidas de tipo unitario \mathbf{x}_i hechas en el conjunto de objetos A , por tanto, $\mathbf{x}_{i, i \in N_0} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

Por otro lado, para cada par de objetos a_i y a_j se dispone de un conjunto m_2 de medidas binarias A_{ij} : $A_{ij} = \{A_{ij}^{(1)}, A_{ij}^{(2)}, \dots, A_{ij}^{(m_2)}\}$. Los mismos tipos de medidas unitarias y binarias se obtienen sobre los modelos.

Con lo anterior y en el marco de la probabilidad Bayesiana la relajación probabilista para el etiquetado de objetos se define como sigue: la etiqueta θ_i de un objeto a_i tendrá el valor ω_{θ_i} , dado que es la etiqueta más probable en base a los valores de las medidas unitarias y los valores de las relaciones binarias entre los objetos. Mediante inducción matemática se puede demostrar que no hay necesidad de incluir todas las relaciones binarias de todos los objetos a la hora de identificar un objeto y la etiqueta más apropiada del objeto a_i es ω_{θ_i} y viene dada por,

$$P(\theta_i = \omega_{\theta_i} \mid \mathbf{x}_{j, j \in N_0}, A_{ij, j \in N_i}) = \max_{\omega_\lambda \in \Omega} P(\theta_i = \omega_\lambda \mid \mathbf{x}_{j, j \in N_0}, A_{ij, j \in N_i}) \quad (L.1)$$

donde la letra mayúscula P representa la probabilidad de un evento, entonces $P(\theta_i = \omega_\lambda)$ y $P(\theta_i = \omega_{\theta_i})$ expresan las probabilidades de que la etiqueta de la escena θ_i se corresponda con las etiquetas del conjunto de modelos ω_λ y ω_{θ_i} respectivamente. Bajo ciertas suposiciones, las probabilidades condicionales en esta ecuación pueden expresarse de forma que indica que el método más apropiado para encontrar el máximo es un procedimiento de actualización mediante relajación. Utilizando la regla de Bayes y tras un exhaustivo tratamiento que puede encontrarse desarrollado en Christmas y col. (1995) la solución deseada al problema de etiquetado tal y como se define en (L.1) puede obtenerse utilizando el esquema iterativo siguiente,

$$P^{(n+1)}(\theta_i = \omega_{\theta_i}) = \frac{P^{(n)}(\theta_i = \omega_{\theta_i})Q^{(n)}(\theta_i = \omega_{\theta_i})}{\sum_{\omega_\lambda \in \Omega} P^{(n)}(\theta_i = \omega_\lambda)Q^{(n)}(\theta_i = \omega_\lambda)} \quad (L.2)$$

donde

$$Q^{(n)}(\theta_i = \omega_\alpha) = \prod_{j \in N_i} \sum_{\omega_\beta \in \Omega} P^{(n)}(\theta_j = \omega_\beta) p(A_{ij} \mid \theta_i = \omega_\alpha, \theta_j = \omega_\beta) \quad (L.3)$$

donde $P^{(n)}(\theta_i = \omega_{\theta_i})$ es la probabilidad de correspondencia en el nivel (iteración) n de que la etiqueta de la escena θ_i se corresponda con la etiqueta del modelo ω_{θ_i} y $P^{(n+1)}(\theta_i = \omega_{\theta_i})$ es la probabilidad actualizada de la citada correspondencia en el nivel $n+1$. La cantidad $Q^{(n)}(\theta_i = \omega_{\alpha})$ expresa el soporte o refuerzo que la correspondencia $\theta_i = \omega_{\alpha}$ recibe en la iteración n^{th} procedente de los otros objetos de la escena, teniendo en cuenta las relaciones binarias que existen entre ellos y el objeto a_i . La función de densidad de probabilidad $p(A_{ij} | \theta_i = \omega_{\alpha}, \theta_j = \omega_{\beta})$ aparece en otros métodos, tales como un coeficiente de compatibilidad, éste es el caso de los trabajos de Rosenfeld y col. (1976) o Hummel y Zucker (1983), dicho coeficiente cuantifica la compatibilidad entre la correspondencia $\theta_i = \omega_{\alpha}$ y una correspondencia vecina $\theta_j = \omega_{\beta}$.

Algunos avances en la teoría de los procesos de etiquetado por relajación los encontramos en Christmas y col. (1995) o Fu y Yan (1997) si bien con anterioridad se pueden destacar los siguientes trabajos: Rosenfeld y col. (1976), Hummel y Zucker (1983), Peleg (1980), Ranade y Rosenfeld (1980), Wang y col. (1983).