

# Estimación del Movimiento Coherente: Computación Evolutiva como Alternativa al Annealing Determinista

Francisco Escolano y Miguel Cazorla  
Grupo i3a: Informática Industrial e Inteligencia Artificial  
Departamento de Tecnología Informática y Computación  
Universidad de Alicante  
E-03690, San Vicente, Spain  
Fax/Phone: 346-5903681  
e-mail: sco@i3a.dtic.ua.es

28 de noviembre de 1997

## Resumen

En este trabajo teórico formulamos soluciones al problema de la Estimación del Movimiento Coherente tomando como referencia funciones de coste que integran variables binarias de decisión y parámetros continuos. Ello se plantea por un lado desde técnicas basadas en *Annealing* Determinista, y por otro, desde enfoques propios de la Computación Evolutiva. En ambos casos se propone la integración de la búsqueda de variables binarias con el descenso por gradiente en el espacio de parámetros.

*Palabras Clave.* Problema de Correspondencia, Minimización de Energía, Computación Evolutiva, *Annealing* Determinista.

## 1 Introducción

La progresiva formulación de funciones de coste en las cuales se integran variables de decisión binarias y parámetros continuos, ha permitido aportar soluciones a problemas de naturaleza combinatoria en el área de la Visión Artificial y el Reconocimiento de Patrones. Dicha integración repercute beneficiosamente, en la mayoría de los casos, en mejora de la eficiencia del procedimiento de búsqueda, ya que la toma de decisiones binarias permite acotar la búsqueda en el espacio de parámetros. En esta contribución analizaremos dicho efecto tomando como problema de referencia la Estimación del Movimiento Coherente. Propondremos soluciones desde dos puntos de vista. En primer lugar aplicaremos los

resultados recientes en el ámbito del *Annealing* Determinista y, a continuación plantearemos como alternativa la Computación Evolutiva para la gestión de las variables de decisión.

## 2 Formulación del Problema

La *Teoría del Movimiento Coherente*<sup>1</sup> [9] [10] [11] aporta una formulación del *Problema de la Correspondencia* entre dos conjuntos de puntos o de características  $\{x_i\}$  y  $\{y_j\}$ , con  $i, j \in \{1, 2, \dots, N\}$  extraídas de imágenes sucesivas, en términos de la minimización restringida de la siguiente función de coste o energía:

$$E[V, v(x)] = \sum_{i,j=1}^N V_{ij} |v(x_i) - (y_j - x_i)|^2 + \lambda \sum_{n=0}^{\infty} \frac{\sigma^{2n}}{n!2^n} \int_{\Omega} |D^n v(x)|^2 dx$$

El primer término incorpora una métrica de similitud entre la variación discreta de las características y el campo de movimiento continuo  $v(x)$  sobre  $x_i$ . El conjunto  $\{V_{ij}\}$  define una *matriz de correspondencia*  $V$  cuyos elementos son variables binarias de decisión de manera que  $V_{ij} = 1$  si el punto  $x_i$  se corresponde con  $y_j$ , y  $V_{ij} = 0$  en caso contrario. La decisión de asignación óptima es aquella que minimiza la energía global al tiempo que satisface las restricciones de asignación *uno-a-uno*:

$$\sum_{i=1}^N V_{ij} = 1, \sum_{j=1}^N V_{ij} = 1, \forall i, j \in \{1, 2, \dots, N\}$$

es decir, la matriz de correspondencia es una *permutación* (matriz cero-uno cuadrada cuyas filas y columnas suman la unidad). Con respecto a la *Teoría de la Función Mínima*<sup>2</sup> [7] se incorpora como objetivo la extracción de un campo de movimiento a partir de la interpolación del campo generado por la correspondencia. Por este motivo el segundo término penaliza campos no-uniformes suavizando el resultado al restringir, al igual que en [4], su variabilidad local siendo  $D^{2n}v(x) = \nabla^{2n}v(x)$  y  $D^{2n+1}v(x) = \nabla(\nabla^{2n}v(x))$ . El rango de dispersión local permitido está modulado por la desviación  $\sigma$  de manera que si  $\sigma \rightarrow 0$  obtenemos la teoría de Ullman. La principal aportación es la integración de la toma de decisiones locales de correspondencia, usualmente binarias, y la estimación de parámetros de movimiento en toda la región.

<sup>1</sup>Long Range Motion Coherence Theory.

<sup>2</sup>Minimal Mapping Theory

### 3 Annealing Determinista

La aplicación de métodos inspirados en la Física Estadística, tales como el *Annealing Determinista* [2], mejora la robustez del proceso de minimización frente a mínimos locales. Se trata de transformar el problema restringido discreto-continuo original en un problema no-restringido continuo. Si aplicamos los resultados obtenidos en [12] y [3] al problema que estamos tratando, la nueva función de referencia es

$$E[V, v(x), \mu, \nu] = \sum_{i,j=1}^N V_{ij} |v(x_i) - (y_j - x_i)|^2 + \lambda \sum_{n=0}^{\infty} \frac{\sigma^{2n}}{n! 2^n} \int_{\Omega} |D^n v(x)|^2 dx \\ + \sum_{i=1}^N \mu_i (\sum_{j=1}^N V_{ij} - 1) + \sum_{j=1}^N \nu_j (\sum_{i=1}^N V_{ij} - 1) + \frac{1}{\beta} \sum_{i,j=1}^N V_{ij} \log V_{ij}$$

La matriz de correspondencia se ha transformado en una *matriz doblemente estocástica* (matriz cuadrada cuyos elementos son positivos y la suma de filas y columnas es la unidad) lo cual implica imponer restricciones de no-negatividad por un lado, y de asignación por otro. Lo primero se garantiza incluyendo una *función barrera* o término de entropía  $\sum_{i,j=1}^N V_{ij} \log V_{ij}$  controlado por el parámetro  $\beta > 0$  que modula el grado de convexidad de la función. El cumplimiento de las restricciones de asignación se garantiza mediante los dos términos siguientes que están influenciados por los multiplicadores de Lagrange  $\mu_i$  y  $\nu_j$ .

Por tanto el objetivo es encontrar un punto de silla de la función de coste (minimizar con respecto a  $V$  y  $v(x)$  al tiempo que se maximiza dicha función con respecto a  $\mu$  y  $\nu$ ) y todo ello para cada valor de  $\beta$ . Se ha establecido [5] una relación directa entre la búsqueda, usualmente ineficiente, de los multiplicadores mediante ascenso por gradiente y el procedimiento debido al teorema [6] según el cual *una matriz doblemente estocástica puede obtenerse a partir de cualquier matriz cuadrada cuyos elementos son positivos, mediante un proceso iterativo de normalización alternativa de filas y columnas*.

El algoritmo resultante consta de un proceso global de *Annealing Determinista* en el que se incrementa progresivamente el parámetro  $\beta$  (inversa de la temperatura). Para cada valor de  $\beta$  se actualizan en primer lugar las variables de correspondencia (algoritmo *SoftAssign*) mediante una asignación probabilística (la exponenciación garantiza la no-negatividad de las variables de decisión).

El proceso iterativo de normalización de filas y columnas de complejidad cuadrática con respecto a  $N$  (*proceso de Sinkhorn*) permite obtener la correspondencia óptima para dicho valor de  $\beta$ . A continuación, una vez obtenidas las nuevas variables se procede a actualizar el campo de movimiento mediante métodos variacionales. En este sentido, en la práctica el término de suavizado se reduce a  $\lambda \frac{\sigma^2}{2} \int_{\Omega} |Dv(x)|^2 dx$  con lo cual el proceso iterativo se basa en la aplicación de las ecuaciones de Euler sobre la función de coste original.

## Movimiento Coherente: *Annealing* Determinista

### Inicialización:

$$\beta \leftarrow \beta_0, v(x) \leftarrow v_0(x), V_{ij} \leftarrow 1 + \epsilon$$

### Annealing Determinista: Mientras $\beta < \beta_f$

#### SoftAssign: Mientras No-Convergencia ( $v(x)$ )

$$\forall i, j \in \{1, 2, \dots, N\}$$

$$Q_{ij} \leftarrow -\frac{\partial E[\{V_{ij}\}, v(x)]}{\partial V_{ij}} = -|v(x_i) - (y_j - x_i)|^2$$

$$V_{ij} \leftarrow \exp(\beta Q_{ij})$$

#### Sinkhorn: Mientras No-Convergencia ( $V$ )

Actualizar  $V$  Normalizando Filas

$$V_{ij} \leftarrow \frac{V_{ij}}{\sum_{j=1}^N V_{ij}}, \forall i, j \in \{1, 2, \dots, N\}$$

Actualizar  $V$  Normalizando Columnas

$$V_{ij} \leftarrow \frac{V_{ij}}{\sum_{i=1}^N V_{ij}}, \forall i, j \in \{1, 2, \dots, N\}$$

### Actualización del Campo de Movimiento

$$v(x) \leftarrow v(x) + \frac{dE[V, v(x)]}{dv(x)}$$

$$\beta \leftarrow \beta_r \beta$$

Un aspecto central en este método es el impacto del proceso de *annealing* sobre los elementos de la matriz de correspondencia. A medida que se incrementa  $\beta$  (con un factor  $\beta_r$ ) las diferencias entre la mejor entrada y las demás se acentúan de manera que en el límite ( $\beta \rightarrow \infty$ ) obtenemos matriz de correspondencia de tipo permutación. A medida que progresa el método el algoritmo de correspondencia (*SoftAssign*) obtiene mejores propuestas del campo de movimiento entre características lo cual facilita la interpolación del campo y resulta de gran utilidad para aplicar el paso variacional.

## 4 Computación Evolutiva

Como planteamiento alternativo proponemos sustituir el esquema de *annealing*, concretamente la parte dedicada a *SoftAssign*, para gestionar la generación y evolución de una población de soluciones candidatas hasta alcanzar un mínimo global [1]. En este sentido pensamos que resulta interesante incorporar los resultados formales aportados para problemas similares tales como el *Problema del Transporte* [8]. El principio básico de este método es mantener, en paralelo, una población  $P$  de soluciones candidatas (matrices de correspondencia binarias) que son evaluadas en términos de su coste con respecto a la función original  $E[V, v(x)]$ , y sobre la que se aplican operadores de variación (mutación y recombinación) y selección.

La población inicial,  $P_0$ , de tamaño  $\mu$  constante durante el proceso, viene dada por un procedimiento orientado a introducir el máximo número de ceros posible y cuyo funcionamiento comentaremos junto con el del operador de mutación. Este procedimiento genera un vértice de la frontera convexa delimitada por las restricciones de asignación. A partir de este punto, fijamos un campo de movimiento inicial  $v_0(x)$ , evaluamos  $P_0$  y entramos en un proceso de minimización evolutiva.

A partir de una población  $P_t$ , con  $t > 0$ , el *operador de recombinación* está basado en el *cruce aritmético*: a partir de dos padres  $V^k$  y  $V^l$  (que satisfacen las restricciones y son seleccionados al azar) se generan dos hijos  $X^{kl}$  e  $Y^{kl}$ , de manera que:  $X^{kl} = \alpha V^k + (1 - \alpha)V^l$  e  $Y^{kl} = \alpha V^l + (1 - \alpha)V^k$  de manera que  $\alpha \in [0, 1]$ . El efecto de este cruce es generar dos hijos que también satisfacen las restricciones ya que estas definen un conjunto convexo. El resultado es la población  $P_t^r$ .

A continuación se aplica un *operador de mutación* sobre  $P_t^r$  mediante un *procedimiento de reparación de admisibilidad*. Dicho procedimiento consiste en extraer de cada individuo  $V_t^k$  una submatriz  $W_t^k$  a partir de dos conjuntos de índices (relativos a filas y columnas). Se calculan los nuevos límites para las restricciones de asignación (sumas de filas y de columnas en la submatriz). Sobre  $W_t^k$  se aplica un procedimiento de *inicialización* que consiste en visitar en un orden aleatorio (distribución uniforme) todos los elementos de  $W_t^k$  y ajustar su valor en función del valor actual de la suma de elementos en la fila y columna considerada. El ajuste de un elemento restringe el de los elementos que figuran en su misma fila y columna. Como resultado, las restricciones locales de asignación se ajustan a los límites especificados para  $W_t^k$ . La reinsertión de los elementos de la nueva submatriz en la matriz original garantiza que la mutación preserva las restricciones globales.

La selección se realiza mediante un criterio de proporcionalidad con respecto al coste de la decisión una vez fijado el campo de movimiento. Seleccionamos  $\mu$  individuos de  $P_t \cup P_t^m$  de manera que la población re-

sultante permanezca constante.

Finalmente en cuanto al *criterio de convergencia* detendremos el proceso cuando no observemos mejoras significativas en la calidad de las soluciones generadas.

Una vez que converge el proceso de decisión actualizamos el campo de movimiento. Para ello tomamos como referencia el mejor de los individuos.

### **Movimiento Coherente: Computación Evolutiva**

#### **Inicialización:**

$$P_0 \leftarrow \{V_0^k\}, k \in \{1, 2, \dots, \mu\}, v(x) \leftarrow v_0(x), t \leftarrow 0$$

Mientras No-Convergencia( $v(x)$ )

**Computación Evolutiva:** Mientras No-Convergencia( $P_t$ )

Recombinación en  $P_t$

$$X_t^{kl} \leftarrow \alpha V_t^k + (1 - \alpha) V_t^l$$

$$Y_t^{kl} \leftarrow \alpha V_t^l + (1 - \alpha) V_t^k$$

$$P_t^r \leftarrow P_t^r \cup (X_t^{kl}, Y_t^{kl})$$

Mutación en  $P_t^r$

$$W_t^k \leftarrow \text{SubMatriz}(V_t^k)$$

$$P_t^m \leftarrow \text{Reparar}(W_t^k, P_t^r)$$

Selección en  $P_t \cup P_t^m$

$$P_{t+1} \leftarrow \text{Seleccionar} \left( P_t \cup P_t^m, \frac{\exp(-E[\{V_t^k\}, v(x)])}{\sum_{k=1}^{\mu} \exp(-E[\{V_t^k\}, v(x)])} \right)$$

$$t \leftarrow t + 1$$

#### **Actualización del Campo de Movimiento**

$$v(x) \leftarrow v(x) + \frac{dE[V, v(x)]}{dv(x)}$$

Con respecto al planteamiento previo de *Annealing* optamos por la gestión de una población de matrices de correspondencia. En términos genéricos ese esquema es trasladable a otros problemas (p.e. emparejamiento de grafos) caracterizados por funciones de coste con similares características siempre y cuando las restricciones determinen un conjunto convexo como en este caso.

## 5 Conclusiones

En cuanto al problema de Estimación del Movimiento Coherente aportamos dos planteamientos de solución, *Annealing* Determinista y Computación Evolutiva. Ambos esquemas se diferencian en la gestión de las variables/matrices de correspondencia pero comparten el uso de dichos procesos de búsqueda para simplificar la estimación/interpolación del campo de movimiento continuo. En futuros trabajos realizaremos estudios comparativos acerca de la eficiencia relativa de ambos esquemas. En este sentido cabe destacar el interés de incorporar, en ambos casos, un mecanismo de multirresolución que complemente la acotación de la búsqueda de parámetros de movimiento.

## Referencias

- [1] Bäck, T., Fogel, D., Michalewicz, Z., Eds.: Handbook of Evolutionary Computation. Institute of Physics Publishing and Oxford University Press (1997).
- [2] Geiger, D., Girosi, F.: Parallel and Deterministic Annealing from MRFs: Surface Reconstruction. IEEE Tran. PAMI **13**(5) (1991) 402-412.
- [3] Gold, S., Lu, C.P., Rangarajan, A., Pappu, S., Mjolsness, E.: New Algorithms for 2D and 3D Point Matching: Pose Estimation and Correspondence. In G. Tesauro, D. Touretzky, J. Alspector, Eds. Advances in Neural Information Processing Systems **7** 957-964. MIT Press Cambridge MA (1995).
- [4] Horn, B.K.P.: Robot Vision. MIT Press. Cambridge MA (1986).
- [5] Rangarajan, A., Gold, S., Mjolsness, E.: A Novel Optimizing Network Architecture with Applications. Neural Computation **8**(5) (1996) 1041-1060.
- [6] Sinkhorn, R.: A Relationship between Arbitrary Positive Matrices and Doubly Stochastic Matrices. Ann. Math. Statist. **35** (1964) 876-879.
- [7] Ullman, S.: The Interpretation of Visual Motion. MIT Press, Cambridge MA. (1979).
- [8] Vignaux, G.A., Michalewicz, Z.: A Genetic Algorithm for the Transportation Problem. IEEE Trans. SMC. **21** (1964) 445-452.
- [9] Yuille, A.L., Grzywacz, N.M.: A Computational Theory for the Perception of Coherent Visual Motion. Nature **2** (1988) 71-74.
- [10] Yuille, A.L., Grzywacz, N.M.: A Mathematical Analysis of the Motion Coherence Theory. Int. J. Computer Vision **3** (1989) 155-175.
- [11] Yuille, A.L.: Generalized Deformable Models, Statistical Physics and Matching Problems. Neural Computation **2**(1) (1990) 1-24.
- [12] Yuille, A.L., Kosowsky, J.J.: Statistical Physics Algorithms that Converge. Neural Computation **6**(3) (1994) 341-356.