

Discriminative Random Fields

Análisis del artículo - Ricardo Omar Chávez García

Snajiv Kumar, Martial Hebert

15 de junio de 2009

- Trata de resolver el problema de clasificación y etiquetado de regiones dada una imagen natural.
- Se presentan los Campos aleatorios discriminativos (DRF) para modelar interacciones de dependencia espacial en imágenes.

Introducción (1)

- Un problema importante en visión es la interpretación semántica de una escena.
- Para esta tarea es importante clasificar varias regiones y objetos en categorías significativas.
 - Un pixel de la imagen.
 - Un bloque de una cuadrícula regular.
 - Una forma irregular.
 - Un objeto de la imagen.

- Es importante utilizar la información contextual en la forma de dependencias espaciales en las imágenes naturales.
- *La idea es cómo mantener la consistencia espacial global utilizando modelos que sólo necesiten considerar dependencias locales relativas.*

Relaciones espaciales (1)

- Hay dos tipos de interacciones espaciales utilizadas para propósitos de clasificación y etiquetado.
 - Suavizado espacial de las etiquetas en imágenes naturales.
 - Interacciones complejas observadas en los datos (p.e. textura).
- La incertidumbre en el etiquetado, además de la capacidad de hacer predicciones globales con restricciones locales, hace a los modelos gráficos un candidato para solucionar el problema.

Relaciones espaciales (2)

- Existen dos principales enfoques que tratan de incorporar dependencias espaciales: probabilistas y no-probabilistas.
- En los esquemas probabilistas dos modelos han sido utilizados para incorporar restricciones contextuales espaciales.
 - Causal. Son distribuciones de probabilidad globales definidas en grafos dirigidos utilizando probabilidades de transición locales (p.e. árbol causal).
 - No-causal. Son distribuciones de probabilidad globales definidas en un grafo no dirigido utilizando potenciales locales de cliques.

- Son los modelos gráficos no dirigidos más populares en visión computacional.
- Permiten incorporar restricciones contextuales locales en problemas de etiquetado.
- Son utilizados generalmente en un marco de trabajo probabilístico generativo.

$$P(x|y) \propto p(x, y) = P(x)p(y|x)$$

- Por viabilidad computacional $p(y|x) = \prod_{i \in S} p(y_i|x_i)$.
- La interacción entre etiquetas en un MRF se modela con el término $P(x)$, que se puede ver como la probabilidad a priori.

Campos aleatorios condicionales (1)

- Modelan la distribución posterior $P(x|y)$ como una campo de Gibbs.
- Permite capturar dependencias arbitrarias entre las observaciones sin recurrir a aproximaciones del modelo.
- CRF han demostrado superar a los modelos ocultos de Markov tradicionales en el etiquetado de secuencias de texto.

Sea $G = (S, E)$ un grafo tal que x es indexado por los vértices de G . Entonces se dice que (x, y) es un campo aleatorio condicional, cuando al condicionar y , las variables aleatorias x_i obedecen la propiedad de Markov con respecto a el grafo: $P(x_i|y, x_{S-\{i\}}) = P(x_i|y, x_{N_i})$.

Campos aleatorios condicionales (2)

- Si se asume sólo potenciales de cliques hasta de un par y utilizando el teorema de Hammersley-Clifford, la distribución conjunta sobre las etiquetas x dadas las observaciones y puede ser escrita como,

$$P(x|y) = \frac{1}{Z} \exp\left(\sum_{i \in S} A_i(x_i, y) + \sum_{i \in S} \sum_{j \in N_i} l_{i,j}(x_i, x_j, y)\right)$$

- Donde A_i es el potencial de asociación y $l_{i,j}$ el potencial de interacción.

Campos aleatorios discriminativos (1)

- Se basa en los campos aleatorios condicionales.
- Extiende el CRFs de 1-D utilizando modelos discriminativos locales para capturar las asociaciones de clases para sitios individuales, como también las interacciones en sitios vecinos en una estructura regular e irregular 2-D.
 - Los potenciales unarios y por pares son diseñados utilizando clasificadores discriminativos locales arbitrarios.
 - Son definidos sobre estructuras 2-D que generalmente inducen grafos con ciclos.
- Permite las interacciones tanto en los datos observados como en las etiquetas.

Campos aleatorios discriminativos (2)

- El potencial de asociación $A(x_i, y)$ se puede ver como una medida de que tan probable es que el sitio i tome la etiqueta x_i dada la imagen y ignorando los efectos de otros sitios en la imagen.
 - Actúa como un clasificador complejo no lineal para sitios individuales.
- El potencial de interacción se puede ver como la medida de cómo las etiquetas en las vecindades de i y j podrían interactuar dada una imagen y .
 - Interacción discriminativa de etiquetas dependiente de los datos.

Campos aleatorios discriminativos (3)

- Los parámetros $\theta = \{w, v, \beta, K\}$ son aprendidos simultáneamente.
- Donde w y v son los parámetros para los potenciales de asociación e interacción, β es un coeficiente que controla el grado de suavidad y K le da flexibilidad al modelo para permitirle al algoritmo de aprendizaje ajustarse a contribuciones relativas en el potencial de interacción.
- Los parámetros son aprendidos utilizando un enfoque estándar de máxima probabilidad.

Campos aleatorios discriminativos (4)

- Dada una imagen de prueba y , el objetivo es encontrar la configuración de etiquetas óptima sobre los sitios de la imagen, dónde la optimalidad está definida con una función de costo.
- La función de costo utilizada es la Maximum A Posterior (MAP) definida como $C(x, x^*) = 1 - \delta(x_i - x^*_i)$.
- Se utilizó el algoritmo Iterated Conditional Modes (ICM) para estimar MAP locales.

Detección de estructuras hechas por el hombre (1)

- El conjunto de entrenamiento (108) y el de prueba (129) contienen imágenes de tamaño 256×384 de la base de datos Corel.
- Cada imagen fue dividida en 16×16 bloques.
- Se utilizaron como características el histograma pesado y la orientación.

Detección de estructuras hechas por el hombre (2)



(a) Input image



(b) Logistic

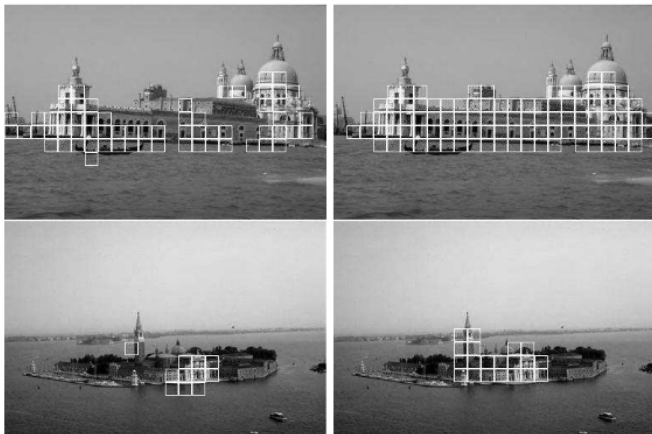


(c) MRF



(d) DRF

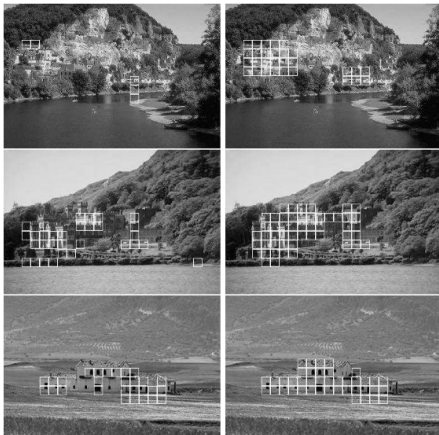
Detección de estructuras hechas por el hombre (3)



Detección de estructuras hechas por el hombre (4)



Detección de estructuras hechas por el hombre (4)



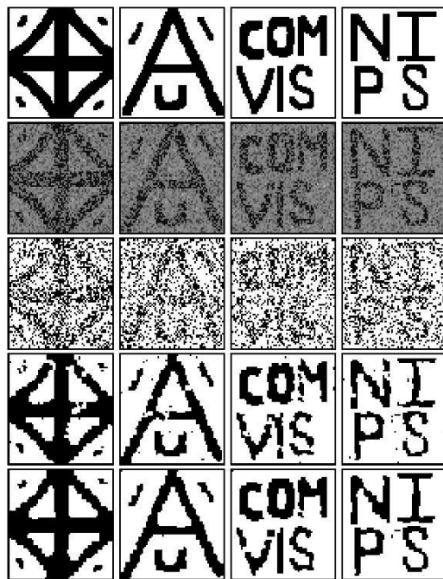
- Para una imagen de 256x384 pixeles al enfoque DRF le toma 2.42s, mientras que para MRF 2.33s y al clasificador logístico 2.18.
- El enfoque DRF supera al MRF en cuanto a evitar falsos positivos y a obtener un mejor porcentaje de detección de estructuras.

	MRF	Logistic ⁻	DRF ⁻	Logistic	DRF ($K = 0$)	DRF
DR (%)	57.2	45.5	60.9	55.4	68.6	70.5
FP (per image)	2.36	2.24	2.24	1.37	1.21	1.37

- La forma del potencial de interacción provoca redundancia en los efectos producidos por los términos de dependencia e independencia de los datos.
- El método de aprendizaje de los parámetros sobre estima los coeficientes de interacción (β).

	MRF	Logistic ⁻	DRF ⁻	Logistic	DRF
DR (%)	58.35	47.50	61.79	60.80	72.54
FP (per image)	2.44	2.28	2.28	1.76	1.76

Eliminación del ruido



Ventajas del DRF sobre MRF (1)

- DRF permite relajar la fuerte suposición de independencia condicional de los datos observados tomada en cuenta en los MRF por viabilidad.
- Los DRFs derivan su poder de clasificación explotando los modelos probabilísticos discriminativos, en vez de los modelos generativos utilizados para modelar observaciones en un MRF.

Ventajas del DRF sobre MRF (2)

- La interacción entre etiquetas en un DRF está basada en la idea de discriminar pares de datos observados, haciéndolo adaptativo a los datos en vez de ser fijado a priori como en los MRFs.
- Todos los parámetros en el modelo DRF son estimados simultáneamente de los datos de entrenamiento. En los MRFs, usualmente, se separa el aprendizaje de los parámetros de probabilidad, del aprendizaje de los parámetros del campo.

Problemas con el método

- 1 La diferencia en tiempo es mayor al considerar la dependencia entre los vecinos y las etiquetas.
- 2 El método no es robusto porque requiere de varias modificaciones según la aplicación que se le quiera dar.
- 3 Surgen problemas como el seleccionar el mejor clasificador local para la aplicación.
- 4 El aprendizaje de los parámetros es una tarea costosa para el método propuesto.
- 5 Su desempeño fue aceptable tomando en consideración 2 clases.

- 1 El enfoque DRF supero en cuanto a porcentaje de detección y disminución de falsos positivos al enfoque MRF.
- 2 No considerar las suposiciones hechas, generalmente, en un MRF demostró ofrecer mejores resultados.
- 3 Como trabajo a futuro se propone utilizar mapeos generales de kernels para mejorar la exactitud de la clasificación.

¿Alguna pregunta?