

03/2010

"Blurred Shape Model": innovador reconocimiento automático de objetos



Una de las dificultades que encontramos en el reconocimiento automático de imágenes proviene de las alteraciones que sufren los objetos en su representación. El fundamento de los actuales sistemas inteligentes consiste en extraer información relevante del objeto, como podría ser su contorno, y aprender que ciertas combinaciones de información corresponden a objetos determinados. Investigadores del Centre de Visió per Computador han propuesto una nueva metodología para la descripción de objetos y su aprendizaje: el premiado descriptor "Blurred Shape Model" (BSM) considera la relación entre el nivel de difuminación del contorno y la facilidad para ser aprendido como un objeto determinado. Los usos de esta nueva metodología van desde el análisis de textos para ser procesados, hasta la robótica, pasando por el análisis de contenido de imágenes en internet.

El reconocimiento automático de objetos en imágenes lleva cerca de 50 años siendo estudiado por las áreas de Inteligencia y Visión Artificial. No obstante, aunque para nosotros reconocer un objeto en una imagen es algo trivial, aún no hemos conseguido desarrollar un proceso

computacional que sea capaz de obtener unos resultados mínimamente parecidos a nuestro sistema visual.

Entre muchos otros motivos, el principal problema proviene de las alteraciones que sufren los objetos cuando aparecen en imágenes. Aunque nosotros somos robustos a la detección de objetos aunque las imágenes contengan ruido, cambios en la iluminación, sombras, occlusiones, diferentes puntos de vista, etc., incluir este conocimiento a un sistema computacional aún es una tarea compleja. Podríamos decir que la mayoría de los sistemas inteligentes actuales intentan imitar nuestro comportamiento visual a partir de dos pasos básicos: descripción y aprendizaje. Primero tenemos que extraer la información relevante de las imágenes (contornos por ejemplo) y después aprender que una combinación determinada de información corresponde a un objeto en particular.

En el trabajo Blurred Shape Model (BSM), hemos propuesto una nueva metodología para la descripción de objetos así como para su aprendizaje. En particular, nos hemos centrado en imágenes binarias y en escalas de grises para detectar diferentes tipos de símbolos, tanto en imágenes en entornos abiertos como en documentos manuscritos. En el primer escenario hay que tratar la problemática de los cambios en el ambiente así como sobre los puntos de vista. En el segundo escenario, el principal problema es tratar la variación que los símbolos sufren debido a los diferentes tipos de escritura de diferentes autores.

La primera propuesta de este trabajo, el descriptor Blurred Shape Model (que recibió el premio al mejor trabajo en el Iberian Conference on Pattern Recognition and Image Analysis de 2007), representa los objetos a partir de la codificación de las relaciones espaciales entre sus partes más representativas. Por ejemplo, supongamos como puntos representativos del objeto aquellos píxeles que forman parte de los contornos. Entonces, el BSM define una rejilla de regiones y para cada punto de contorno calcula la inversa de las distancias al resto de puntos de contorno próximos a él. De esta forma, se codifican relaciones espaciales y el tamaño de la rejilla sobre el objeto define el nivel de difuminación que queremos que aprenda el sistema inteligente. En general, difuminaciones grandes hacen que se aprendan bien los objetos pero se podrían llegar a confundir con otros. Por otra parte, difuminaciones reducidas serían demasiado específicas y no podríamos encontrar objetos parecidos que hayan sufrido algún tipo de deformación. Por este motivo, el método testea todos los valores de difuminación posibles sobre unos datos de aprendizaje para adaptar los parámetros a las necesidades de cada problema. Sería como enseñar una imagen detrás de otra a un niño de un año, con el fin de incluir nueva información en su fuente de conocimiento. La Figura 1 muestra un ejemplo de los modelos difuminados para los contornos de un objeto de tipo manzana.

Una vez el sistema detecta los niveles de desenfoque o difuminación de los objetos a aprender, técnicas de Inteligencia Artificial son usadas para aprender a diferenciar entre los objetos. Por ejemplo, utilizamos un clasificador estadístico que aprende a diferenciar modelos difuminados de manzanas de modelos difuminados de peras, otro que separe manzanas de plátanos, etc. Finalmente, si combinamos las respuestas de estos aprendizajes parciales, obtenemos una metodología multi-clase, que permite detectar de forma automática un conjunto de objetos en imágenes. En particular, la metodología aplicada usa métodos estadísticos que se centran en encontrar aquellas partes que mejor separan unos objetos de otros. Además, el sistema se beneficia de corrección de errores mediante el uso de una clasificación que se basa en la

metodología seguida en el campo de teoría de la información. La figura de la izquierda muestra un ejemplo de cómo un robot programado con un sistema inteligente y con una fuente de conocimiento de descriptores difuminados de fruta, puede detectar las zonas de interés en una imagen y clasificar el objeto de entre un conjunto de posibilidades.

Los resultados de esta metodología sobre datos capturados manualmente y de fuentes públicas han resultado ser altamente satisfactorios. Comparados con los resultados de los métodos que compiten en este campo, se han encontrado mejoras estadísticamente significativas. La Figura 2 muestra los resultados comparando diferentes descriptores del estado del arte con el BSM y aplicados sobre dos metodologías de aprendizaje diferentes. Los datos consisten en 70 objetos diferentes en imágenes de la base de datos pública MPEG7. Se puede ver que la mejora es considerable.

El futuro de este trabajo se centra en incrementar el número de objetos a diferenciar y en encontrar características más relevantes que sean codificadas dentro del modelo difuminado, tales como el color o la textura. La aplicabilidad de esta nueva metodología va desde el análisis automático de textos para su procesado, hasta la robótica, permitiendo la navegación e interactividad de sistemas inteligentes, así como para el análisis eficaz de contenido entre todos los millones de imágenes de los que disponemos en internet.

Sergio Escalera
sescalera@cvc.uab.es

Referencias

"Blurred Shape Model for Binary and Grey-level Symbol Recognition". Sergio Escalera, Alicia Fornés, Oriol Pujol, Petia Radeva, Gemma Sánchez, and Josep Lladós. Pattern Recognition Letters, doi:10.1016/j.patrec.2009.08.001, 2009.

[View low-bandwidth version](#)