

Detección de objetos en imágenes a color utilizando redes de RBF

A. Fernández

Postgrado en Instrumentación, Facultad de Ciencias, Universidad Central de Venezuela

J. A. Moreno

LACE, Facultad de Ingeniería, Universidad Central de Venezuela

Abstract

A simple computational system for a fast and reliable object detection in color images is proposed in this paper. The method consist in training a RBF neural network using suitable preprocessing data from objects contours. Probability densities functions are obtained for each class and using Bayes Theorem, the probability for a detected object is computed. The proposed method is tested using dissimilar and similar objects categories, with high intra-class variability but invariant pose and scale. Promising results are obtained since the neural network achieves a good separation of the categories with a simple structure.

Resumen

Se propone un sistema computacional sencillo para la detección rápida y confiable de objetos en imágenes a color. El procedimiento consiste en entrenar una red neuronal del tipo RBF utilizando la data proveniente de los contornos de los objetos convenientemente preprocesada. Se obtienen las funciones densidad de probabilidad para cada clase y con el Teorema de Bayes, se obtiene la probabilidad con la que un objeto pueda ser detectado. Se prueba el sistema entrenando la red con categorías de objetos disímiles y semejantes, con gran variabilidad intra-clase pero con orientación espacial y escala constantes. Se obtienen resultados prometedores ya que la red logra una buena separación de las clases con una estructura bien sencilla.

Introducción

La detección de objetos en una imagen es el procedimiento mediante el cual se ubican en la imagen él o los objetos de interés. Constituye una tarea complicada debido a la cantidad de variables que pueden afectar la apariencia del objeto, tales como: su posición y orientación en la imagen, la presencia de obstáculos que provoquen su ocultamiento parcial, las diferencias debidas a condiciones de iluminación y la variabilidad intra-clase, es decir, objetos de una misma clase que posean importantes variaciones entre sí (Zhang y Zelinsky, 2004). Entre las aplicaciones se encuentra el desarrollo de sistemas de visión artificial para robótica, por medio de los cuales otorgar a un robot la capacidad de detectar diferentes tipos de objetos en su ambiente de trabajo que le facilite la ejecución de las tareas para las cuales haya sido diseñado.

El mecanismo tradicional de detección se fundamenta en representar los objetos en función de uno o varios rasgos invariantes y robustos (color, textura, forma, contexto, otros) para luego entrenar un clasificador que logre distinguirlos del fondo. Aunque las estrategias diseñadas han evolucionado en rapidez y confiabilidad, todavía no hay una respuesta definitiva al problema que logre superar la gran cantidad de variables que afectan la apariencia de los objetos (Zhang y Zelinsky, 2004). Con el objetivo de contribuir a la solución, se presenta una propuesta para la detección de objetos en imágenes a color, simple, de respuesta rápida y con la posibilidad de separar varias clases diferentes ofreciendo resultados con alta confiabilidad.

Procedimiento

Las etapas que conforman el sistema de detección se mencionan a continuación:

→ Partiendo de las imágenes de los objetos a color, se extrae la componente de intensidad o luminancia Y , por lo que las imágenes originales se transforman en imágenes en escala de grises.

→ Cada imagen de intensidad se considera como estado inicial de un Autómata Celular Bidimensional que, bajo una regla de transición simple aplicada sobre una vecindad de Von Neumann, evoluciona a una configuración donde los contornos son resaltados (Scarioni y Moreno, 1998). Las imágenes de los contornos se presentan como imágenes binarias donde los pixeles en negro representan los bordes. En la figura 1 se observa la obtención de bordes con esta técnica.

→ Se calculan los momentos centrados hasta orden 4 asociados a cada imagen de contorno, con lo cual se obtienen patrones de 16 componentes para representar cada objeto.

→ Patrones de 2 clases (2 tipos de objetos) se introducen a una Red Neuronal tipo RBF con una capa oculta y una unidad de salida. El esquema de una red de este tipo se observa en la figura 2. $h(x)$ son funciones de base radial gaussianas multidimensionales, w_i son los pesos sinápticos de la unidad de salida y $f(x)$ es la combinación lineal de las salidas de la capa oculta.

→ Se efectúa el entrenamiento de la red utilizando un esquema de aprendizaje supervisado en el cual se emplea un Programa Evolutivo para hallar los vectores centroides y las varianzas de las funciones gaussianas y un algoritmo tipo Adaline para encontrar los pesos sinápticos de la unidad de salida.

→ Para cada conjunto de entrenamiento y una vez finalizada la fase de aprendizaje, se estiman las densidades de probabilidad de las salidas de la red (d) utilizando núcleos gaussianos.

→ Empleando el Método de Levenberg-Marquardt, se obtienen las funciones densidad de probabilidad (FDP) de cada clase, ajustando las densidades obtenidas del paso anterior con una combinación lineal de N funciones gaussianas,

$$f(d) = \sum_{i=1}^N A_i e^{-\left(\frac{d-B_i}{C_i}\right)^2}$$

donde A_i , B_i y C_i son los parámetros de ajuste.

→ Para la etapa final del sistema, se utilizan las FDP de las 2 clases y el Teorema de Bayes, con el fin de determinar la clase a la que pertenece un objeto. Si tenemos 2 clases mutuamente excluyentes Ω_1 , Ω_2 y un evento dado por la aparición de un objeto O , O será de la clase Ω_1 si $p(\Omega_1 | O) > p(\Omega_2 | O)$, es decir, si $p(O | \Omega_1)p(\Omega_1) > p(O | \Omega_2)p(\Omega_2)$, utilizando el Teorema de Bayes. $p(O | \Omega_1)$ y $p(O | \Omega_2)$ son las FDP del paso anterior y $p(\Omega_1)$ junto con $p(\Omega_2)$, son las probabilidades a priori para cada clase.

Experimentos

Para determinar si el sistema funciona, se procede al entrenamiento de la red utilizando primero, 2 conjuntos de imágenes de objetos disímiles (sillas, tazas) y luego, dos conjuntos de imágenes de objetos semejantes (tazas, teteras). Parte de las imágenes, tomadas de la base de datos de Caltech (<http://www.vision.caltech.edu/archive.htm>), se muestran en la figura 3.

Se utilizan 80 imágenes de entrenamiento por clase las cuales presentan fondo no uniforme, diferentes condiciones de iluminación y una importante variabilidad intra-clase, manteniendo aproximadamente constantes, la escala y orientación espacial de los objetos.

Resultados y Discusión

En la figura 4 se muestran las FDP para sillas y tazas mientras que en la figura 5 se presentan las FDP para teteras y tazas.

Las gráficas muestran como el sistema logra una buena separación de las clases, lo cual le confiere capacidad de efectuar la detección de los objetos involucrados. La estructura de la red neuronal, que es el núcleo del sistema, es bien sencilla comparada con otras estrategias de detección reportadas en la literatura (Viola y Jones, 2001), (Torralba et al, 2003), (Ali y Shah, 2005); se logran separar sillas de tazas con 4 unidades ocultas, mientras que con 7 se logra el

objetivo entre tetras y tazas. Esto beneficia la capacidad de generalización del sistema y su rapidez de respuesta, dado que el número de parámetros necesario es relativamente pequeño.

Debido a la manera en la cual se obtienen los patrones de rasgos, el sistema sería invariante, por lo menos, a los cambios de iluminación y a las rotaciones en el plano de la imagen. Por otro lado, disponer de las FDP de las clases resulta ventajoso, ya que permite definir un criterio probabilístico para realizar una detección, a través del Teorema de Bayes.

Referencias

Ali S, Shah M, (2005), A Supervised Learning Framework for Generic Object Detection in Images, *Tenth IEEE ICCV, Vol. 2*, 1347-1354.

Scarioni A, Moreno JA, (1998), Border Detection in Digital Images with a Simple Cellular Automata Rule, *Proceedings of the Third Conference on Cellular Automata for Research and Industry ACRI'98 in Cellular Automata Research Towards Industry*, 146-156.

Torralba A, Murphy KP, Freeman WT, Rubin MA, (2003), Context-based vision system for place and object recognition, *ICCV03*, 273-280.

Viola P, Jones M, (2001), Rapid Object Detection using a Boosted Cascade of Simple Features, *Conference CVPR01*, 1:511-518.

Zhang W, Zelinsky G, (2004), Current Advances in Computer-based Object Detection and Target Acquisition, *Technical Report EYECOG-04-01*.

Lista de figuras:

figura 1 : Muestra de la detección de bordes.

figura 2: Arquitectura de la Red Neuronal tipo RBF.

figura 3: Muestra de las imágenes de entrenamiento.

figura 4: FDP para las categorías sillas y tazas.

figura 5: FDP para las categorías teteras y tazas.

figura 1:

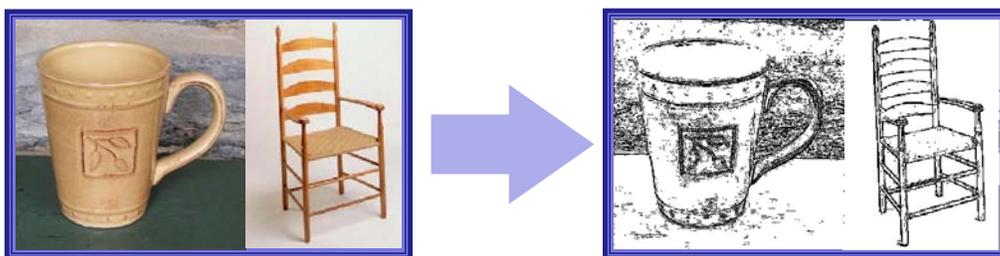


figura 2:

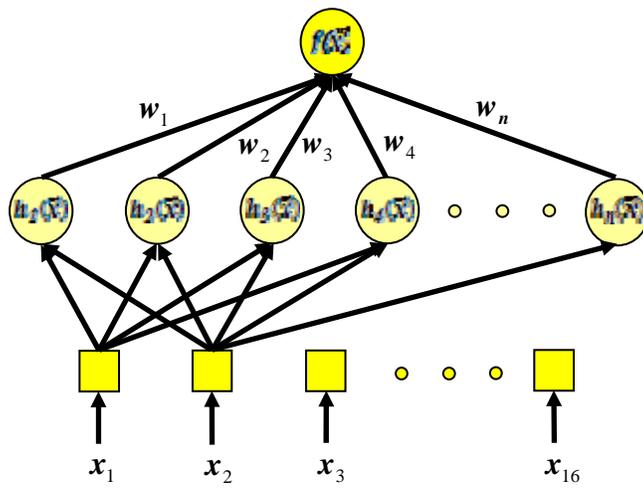


figura 3:



figura 4:

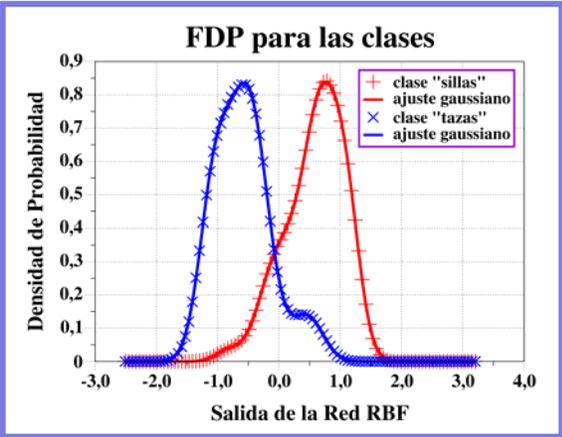


figura 5:

