

Clasificación de imágenes de textura por metodo de Bosques de decisión y Nearest Neighbour

Cristian Alejandro Vergara Perico
Universidad de los Andes, Bogotá, Colombia.

Keywords—*Clasificación, Textura, Arboles de decisión, Nearest Neighbour,*

I. DESCRIPCIÓN DE LA BASE DE DATOS

LA base de datos contiene 1000 imágenes en escala de grises que muestran diferentes texturas. Estas imágenes se clasifican en 25 categorías como madera, piso, madera entre otros. De las mil categorías se tomaron 30 muestras de cada categoría para el entrenamiento y 10 para realizar la validación. De esta forma la base de datos se subdividió en 750 imágenes de entrenamiento y 250 de test.

II. FILTROS Y METODOS USADOS

II-A. Creación del diccionario

- Se creo un banco de filtros de 2 tipos hilbert y segunda derivada de gaussiana. Para cada una de estas categorías se seleccionaron 8 direcciones cada 45° hasta completar los 360°.
- Se realizo una cros-correlación de cada uno de los filtros con un arreglo, el cual contiene la concatenación de 25 imágenes aleatorias representativas de cada una de las categorías de la base de test.
- Se reorganizo el arreglo para obtener una matriz de dimensiones del número de pixeles por el número de filtros. Debido al alto número de pixeles, se seleccionó una muestra representativa de 5000 de toda la población de pixeles.
- A la muestra seleccionada se le aplico el método de agrupación de k-means con 50 grupos seleccionados.
- Finalmente se obtienen los centroides de cada uno de los 50 grupos, los cuales serán usados como referencia para medir la distancia euclidiana de cada punto de las imágenes de evaluación y de entrenamiento.

II-B. Número de textones empleados

El número de textones que se debe emplear para representar las texturas debe ser alto, con el fin de discriminar mejor características de las texturas, sin embargo, la selección de números muy elevados genera un incremento en el tiempo y el costo computacional empleado. De esta manera se emplearon 50 textones.

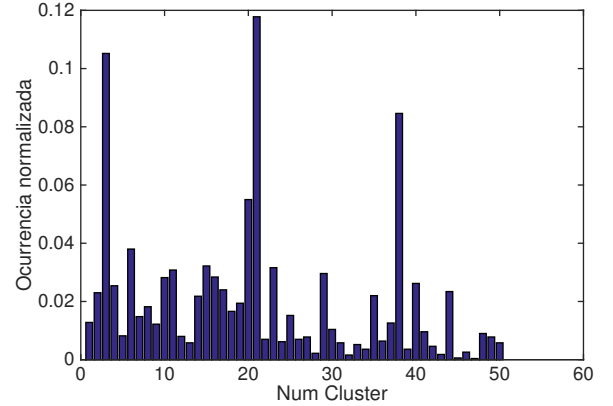


Figura 1: Ocurrencia normalizada de pixeles en cada texton

II-C. Discriminación filtros

La figura 1 muestra el texton que mas pxeles discrimino, formando el diccionario de textones

Se puede ver que los grupos que mas discriminaron fueron el 3, 21 y 38. Los centroides de los grupos corresponden a una combinación lineal de las respuestas de los 32 filtros, por lo cual no se puede indicar que exista un filtro que sea mas relevante que el resto.

III. CLASIFICACIÓN

Para la clasificación se encuentra el histograma de las ocurrencias de cada uno de los textones en cada imagen. Este histograma se normalizo (ecuación 1), de esta forma al sumar todas las categorías del histograma normalizado el resultado es 1.

$$H1_n = H1/N_{pix} \quad (1)$$

III-A. Intersección

El método se basa en maximizar la ecuación 2

$$D = \sum_{i=1}^{N_{grupos}} \min(H1_i, H2_i) \quad (2)$$

¹ Este trabajo fue realizado para el curso de maestría Visión Artificial dictado el primer semestre de 2015 en la Universidad de los Andes, Colombia

El máximo valor que se puede obtener es 1, lo cual correspondería a la intersección completa en cada clase del histograma. Ver referencia [1].

III-B. Chi-Cuadrado

La distancia chi cuadrado se basa en encontrar una distancia mediante la ecuación 3. Ver referencia [1].

$$D = \sum_{i=1}^{N_{grupos}} \frac{(H1_i - H2_i)^2}{H1_i} \quad (3)$$

Una mayor correspondencia se da al tener una distancia D menor, lo cual indica que la diferencias en las clases de ambos histogramas es menor.

III-C. Árboles de decisión

En este laboratorio se empleo una versión de bosque de decisión llamada treebagger [2]. En el entrenamiento esta recibe como parametros:

- El número de arboles del bosque
- Las imágenes de entrenamiento con cada uno de as características que corresponde al histograma de textones
- Las imágenes de entrenamiento con cada uno de as características que corresponde al histograma de textones
- La clase de arboles, para este caso de clasificación
- Número de variables en las hojas

El algoritmo se basa en la generación de arboles aleatoriamente a partir de la selección de una muestra de características para la formación de cada árbol. Esta aleatoriedad genera que no exista una sobre-especialización en la base de datos de entrenamiento del algoritmo en la base de entrenamiento. Este sigue el siguiente procedimiento.

Se encuentra el tamaño de la muestra de la población de características mediante la ecuación 4

$$N_m = \sqrt[2]{N_{grupos}} \quad (4)$$

Para la formación de cada árbol se escogen un grupo de N_m características seleccionadas aleatoriamente. Teniendo como entrada estas características se realizan las preguntas de los nodos de manera que el árbol optimice el proceso de discriminación de las etiquetas.

El número de arboles se escogió basándose en la Figura 2.

La figura 2 describe el error existente por dejar por fuera del bosque arboles que clasifiquen categorías no seleccionadas en la muestra tomada para la formación del arbo. Es evidente que con muy pocos arboles la probabilidad de no tener en cuenta una categoría aumenta y conforme se aumenta el número de arboles esta probabilidad disminuye.

Con el de no sobre-especializar los arboles de decisión el numero mínimo de observaciones en cada hoja se estableció en 12. De esta manera se obtuvo un numero de nodos promedio en cada árbol de 91. En la figura 3 se muestra las ramificaciones d uno de los arboles del bosque.

Se obtiene que la profundidad del arbol no es muy grande lo que indica que la probabilidad de sobre-especialización decrece.

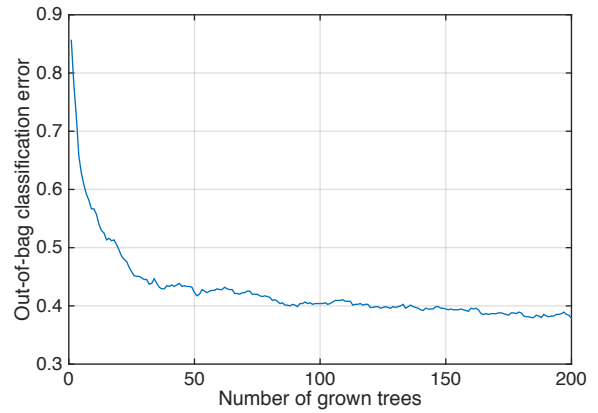


Figura 2: Error de no incluir una característica

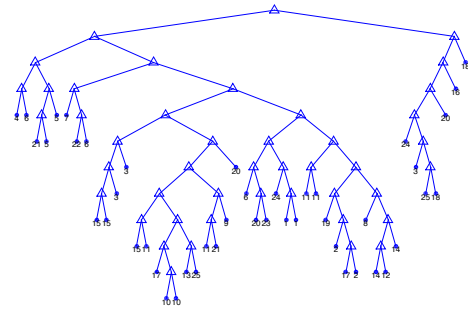


Figura 3: Muestra de arbol de decisión

IV. RESULTADOS

Para evaluar cada uno de los metodos a nivel de las categorías se realizo la matriz de confusión, la cual tiene en su ejes horizontales las categorías a las cuales pertenece la imagen y en sus ejes verticales las categorías en las cuales se clasifico cada imagen. Las matrices de confusión se encuentran en la figura 4

V. DISCUSIÓN DE RESULTADOS

V-A. Mejor metodo

Para encontrar el mejor metodo de clasificación se realizo la suma de los verdaderos positivos con respecto al numero de verdaderos totales. El resumen se encuentra en la tabla ??.

Metodo	Porcentaje [%]
Entrenamiento Chi-Cuadrado	100
Evaluación Chi-Cuadrado	62.8
Entrenamiento Intersección	100
Evaluación Intersección	66
Entrenamiento Bosque de decisión	91.6
Evaluación Bosque de decisión	66

Cuadro I: Porcentaje de acierto en todas las categorías

De la tabla I se puede concluir que los tres metodos obtuvieron resultados similares a nivel global, sin embargo,

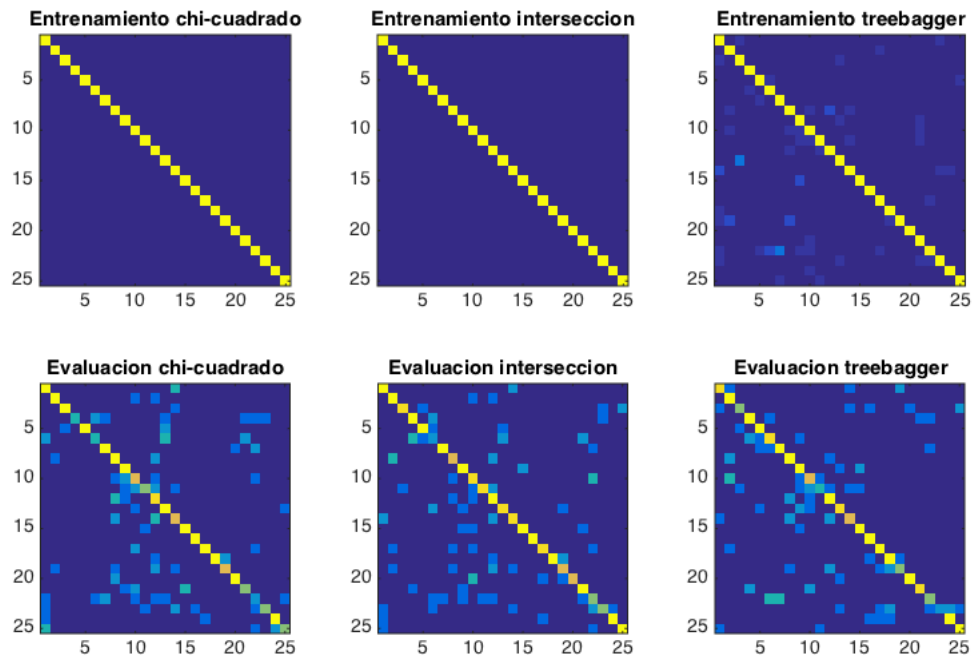


Figura 4: Matrices de confusión, las intensidades mas altas indican un mayor valor

el metodo chi-cuadrado tiene un desempeño levemente mas bajo. Otra característica importante es que en la etapa de entrenamiento los 2 metodos de nearest neighbour tienen un desempeño de 100 [%] como era de esperarse, mientras, que el bosque de decisión tubo un alto desempeño de 91 [%]. Esto indica que el clasificador no tiene una sobre-especialización en la etapa de entrenamiento. En la subsección V-B se describirán los tiempos de procesamiento.

V-B. Tiempos

El resumen de los tiempos se encuentra en la tabla.

Metodo	Tiempo [s]
Chi-Cuadrado	1.67
Intersección	1.56
Arboles de decisión	190.68

Cuadro II: Tiempos obtenidos en el procesamiento con cada metodo

V-C. Categorías mayor confusión

El cuadro III resume las categorías de mayor confusion para cada uno de los metodos

Metodo	Categorías	Categorías con que se confundieron
Chi-Cuadrado	4, 22, 6	4, 13, 7
Intersección	6	4
Arboles de decisión	23, 11	24, 3

Cuadro III: Categorías confusas, score ¡40 [%]

Del cuadro III se obtiene que las categorías 6 y 4 generarón una mayor confusión en el metodo de nearest neighbourhood. Ambas son subcategorías de madera. En la figura ?? se muestra la similitud entre algunas de las imagenes de ambas categorías.

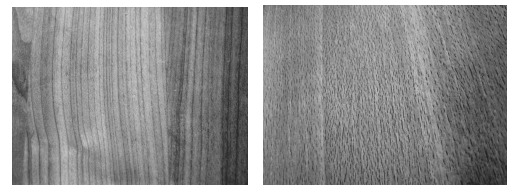


Figura 5: Comparación imagen categoría 4 madera 1 (izquierda) y categoría 6 madera 3 (derecha)

En la figura 5 se observa la semejanza en la orientación de las texturas, y de la forma de la textura.

V-D. Limitaciones del metodo

Para imágenes donde las fronteras se encuentran muy cercanas, como en el caso del pelo, los filtros empleados no

consiguen caracterizar bien los contornos locales de la textura

Solo se puede tomar una muestra para la formación del diccionario de textones, al seleccionar esta muestra aleatoriamente se disminuye la posibilidad de no clasificar texturas .

La clasificación por el método de clasificación de vecinos mas cercanos es muy sensible a ruido cuando las distancias calculadas para la clasificación son muy cercanas entre ellas. Es decir para este caso, cuando la diferencia entre histogramas de la misma clase es similar a la diferencia entre histogramas entre clases diferentes.

Para el método de treebagger la selección aleatoria de pequeñas muestras para la formación de los arboles reduce el error de sobre-especialización, sin embargo, el número de arboles incluidos en el bosque debe ser lo suficientemente grande para evitar incurrir en errores de no incluir características dentro del bosque.

V-E. Limitaciones base de datos

Solo contiene 25 categorías, con lo cual no se podría clasificar texturas que no se encuentren en la base de datos.

No son imágenes aleatorias de texturas, es decir las imágenes fueron tomadas con el fin de extraer características relevantes de las texturas. Esto genera que los algoritmos no sean lo suficientemente genéricos para detectar estas texturas en imágenes de la vida cotidiana.

VI. CONCLUSIONES

Tanto el metodo de intersección como el de los arboles de decisión arrojan resultados similares, sin embargo en la figura 4 se observa que la variación en el metodo de arboles tiene una mayor varianza en las categorías que el metodo de intersección. Es importante resaltar que el tiempo de procesamiento descrito en el cuadro II del metodo de intersección es significativamente menor que el metodo de treebagger. Sin embargo, con el aumento de categorías, de imagenes de entrenamiento e imagenes de evaluación, el rendimiento del clasificador de intersección disminuye ver ref [2]

Se aconseja aumentar la base de datos se puede generalizar mas el entrenamiento y mejorar la clasificación de las imágenes. De la misma manera es aconsejable seleccionar imagenes mas aleatorias.

Se puede tener un mayor banco de filtros con el fin de poder extraer respuestas a contornos mas delgados como el caso del pelo y a su vez contornos mas gruesos. Se recomienda incluir los filtros circulares para tener altas respuestas en características como poros.

REFERENCIAS

- [1] P. Arvelaez, "Computer Vision. IBIO4490. Lecture 5, Clasification."
- [2] MATLAB, *version 8.4.0 (R2014b)*. Natick, Massachusetts: The Math-Works Inc., 2014.