

Identificación de las principales tipologías de nubes utilizando HOG

Carmelina Ierardi
carmelina.ierardi@gmail.com

Manuel R. Arahal
arahal@us.es

Resumen

Este artículo aborda la clasificación automática de nubes a partir de imágenes tomadas desde el suelo mediante técnicas que utilizan histogramas de vectores gradiente. La identificación de nubes tiene una importancia creciente en la industria solar como medio para producir predicciones locales de nubosidad a muy corto plazo. En este contexto se presentan los histogramas de vectores gradiente para extracción de características útiles para la clasificación. Se discuten y exponen ciertas propiedades de este tipo de características que son posteriormente puestas a prueba en la clasificación de nubes. El clasificador empleado es una máquina de vectores soporte SVM. Se consideran en la clasificación las principales tipologías de nubes: cúmulos, cirros y estratos.

Palabras clave: HOG, Identificación nubes, Clasificación nubes

1 Introducción

Uno de los objetivos de la visión automática es interpretar el contenido de imágenes, siendo la clasificación de formas un paso para realizar esta tarea. Las aplicaciones prácticas han ido creciendo en las últimas décadas y se encuentran en todos los sectores. En particular se aborda aquí en el contexto de la industria solar para conseguir identificar las nubes a partir de imágenes tomadas desde el suelo. Esta identificación se puede utilizar para predecir la energía disponible para generación en plantas solares [15] permitiendo el uso de estrategias predictivas robustas [10] y también es útil en otros sectores como la gestión energética de edificios [5, 12].

De acuerdo con el Atlas Internacional de Nubes, publicado en 1956 por la Organización Meteorológica Mundial (OMM), las nubes se clasifican en diez géneros (clasificaciones básicas), que se excluyen mutuamente, como muestra la figura 1, donde se representa la altura en que se forman y su apariencia típica [14].

Las formas nubosas fundamentales son tres: cir-

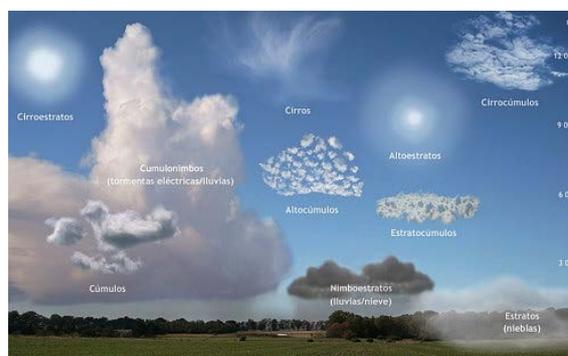


Figura 1: Tipologías de nubes según la OMM

ros, cúmulos y estratos; todos los restantes tipos corresponden o bien a estos tipos puros o son modificaciones y combinaciones de los mismos, a diferentes alturas, donde la variación de las condiciones del aire y humedad son responsables de las diversas formas que presentan.

Como es sabido, la identificación de formas se compone principalmente de dos pasos: la extracción de características que se agrupan en un vector y la clasificación del vector de características mediante alguna técnica [3]. Las características que han sido empleadas anteriormente en la clasificación de nubes son muy variadas. En este artículo sólo se considera el espectro visible y no el caso de sensores especiales (IR, UV, etc). En la literatura, los primeros estudios se basaban en la intensidad luminosa de cada píxel en imágenes monocromas [8]. Posteriormente se han usado imágenes en color, lo cual abre la puerta a usar características basadas directamente en los niveles de intensidad de los canales verde, rojo y azul [11] y también a variables derivadas de éstas, por ejemplo: la crominancia, la relación azul/rojo, la diferencia azul-rojo, la saturación, etc. En [2] se presenta un estudio de la aplicación de este tipo de variables cromáticas.

Las características anteriormente citadas suelen ser insuficientes para la correcta clasificación de nubes debido a que la información que transmiten es necesariamente puntual. Para distinguir tipos de nubes es preciso tomar en consideración distribuciones de medidas adquiridas en una región

de la imagen. Se abre así la puerta a la utilización de características provenientes del análisis de texturas. En este grupo se van a destacar por un lado los estudios clásicos basados en la matriz de co-ocurrencia de niveles de gris [4] y por otro técnicas más actuales basadas en diseños binarios locales (LBP) [13]. Existen también intentos de combinar características de tipos variados, por ejemplo texturales y colorimétricas, en un mismo clasificador. Por ejemplo en [6] se propone la validación cruzada sobre clasificadores de k-vecinos-más-cercanos (kNN) para realizar la selección de características. La selección no contempla todas las características sino un grupo de candidatas previamente seleccionadas. Para la pre-selección se utiliza una medida de la distancia entre las funciones de densidad probabilística estimadas para cada característica evitándose la carga de cálculo que conlleva la exploración exhaustiva.

El uso de los histogramas de orientación tienen muchos precursores. Los más citados son Dalal y Triggs que presentaron al algoritmo de detección de peatones con excelentes resultados. En [1] se describe con detalle un descriptor de peatones basado en el gradiente. Este trabajo se ha convertido en la base de muchos otros métodos posteriores, no solo para la detección de peatones, sino también para detectar cualquier otro tipo de objeto en la imagen, como por ejemplo nubes [7]. En este artículo se utilizan características calculadas a partir del descriptor HOG (Histogram of Oriented Gradients) y son usadas con un clasificador basado en SVM (Support Vector Machines). En particular se considera el histograma del gradiente que usa la dirección del gradiente de orientación que es invariante a cambios de iluminación. Estos gradientes no están representados con valores puntuales, sino más bien en forma de un histograma para permitir invariancia a pequeñas deformaciones locales [17].

Este trabajo se organiza de la siguiente manera: en la *Sección 2* se explica en detalle el descriptor HOG. A continuación, en la *Sección 3* se introduce la base de datos utilizadas, así como las tipologías de nubes. Se analizan y evalúan los resultados obtenidos en la *Sección 4*. Finalmente, en la *Sección 5* se resumen las aportaciones y conclusiones de este trabajo.

2 Descriptor HOG

El descriptor HOG utiliza la información básica del gradiente en cada uno de los píxeles de la imagen sea para la detección que para el reconocimiento de los objetos en una imagen. La definición propia de gradiente representa el máximo cambio en la intensidad de una imagen

I en una cierta dirección. El gradiente se calcula para cada uno de los píxeles de la imagen y queda definido para cada píxel (x,y) por dos valores:

- dirección donde el cambio de intensidad es máximo;
- magnitud del cambio en esta dirección.

Dependiendo de estos dos parámetros para cada píxel se pueden distinguir diferentes situaciones de la configuración local, alrededor del píxel en relación al cambio de contraste y a la forma local.

Existen diferentes formas para poder calcular el gradiente. En el contexto del descriptor HOG se realiza a partir de la diferencia de intensidad que hay entre los píxeles adyacentes en dirección tanto horizontal (ecuación 1), como vertical (ecuación 2).

$$dx = I(x+1, y) - I(x-1, y) \quad (1)$$

$$dy = I(x, y+1) - I(x, y-1) \quad (2)$$

A partir de estas diferencias se pueden calcular los dos parámetros fundamentales, que son la orientación θ (ecuación 3) y la magnitud g (ecuación 4) del gradiente en el punto (x,y) :

$$\theta(x, y) = \arctan \frac{dy}{dx} \quad (3)$$

$$g(x, y) = \sqrt{dx^2 + dy^2} \quad (4)$$

Una vez calculados la orientación y la magnitud del gradiente para cada uno de los píxeles, estos nos darán una información global de la imagen. En la figura 2 están representadas dos diferentes tipologías de nubes, en particular un cirro y un cúmulo que vamos a utilizar para dar un ejemplo de aplicación del método del gradiente, como se muestra en figura 3.

Como resultado se obtendrá la magnitud global del gradiente para cada uno de los píxeles de estas imágenes, que nos están indicando todos los píxeles donde hay un alto cambio de intensidad en cualquier dirección. Básicamente estos cambios se concentran en la silueta de las nubes y dan informaciones acerca de las formas de las mismas. Los datos obtenidos se pueden usar para poder distinguir entre los varios tipos de nubes y efectuar una clasificación (figura 3).



Figura 2: Diferentes tipologías de nubes: cirro y cúmulo

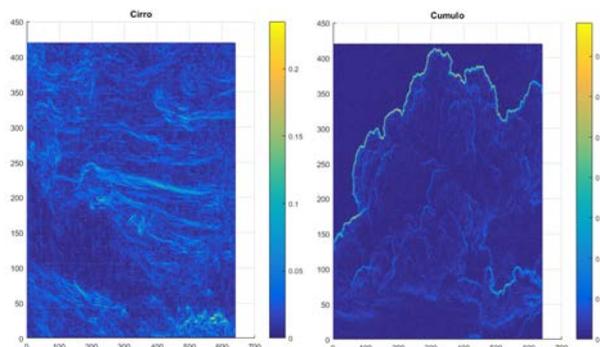


Figura 3: Magnitud del gradiente respectivamente del cirro y del cúmulo

En este ejemplo se ha utilizado la magnitud, pero igualmente se puede utilizar la orientación y puede ser también información útil para poder distinguir la forma de las nubes (figura 4).

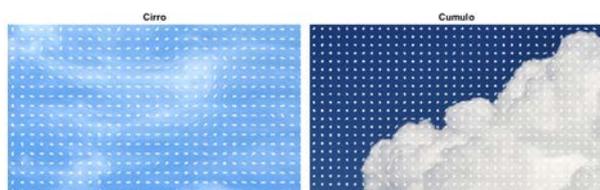


Figura 4: Orientación del gradiente respectivamente del cirro y del cúmulo

Para imágenes en color, se suele dar prioridad al canal dominante de la imagen en lugar de pasar a escala de grises. Esto se consigue calculando el gradiente para cada uno de los canales: Red, Green, Blue (o HSV o YCbCr) [16] y escogiendo el gradiente del canal con magnitud mayor.

La información del gradiente local se puede agregar en forma de histogramas, los cuales se calculan en diferentes áreas de la imagen; es decir se va a convertir la información de cada píxel en una representación global de toda la imagen en forma de vector de características que capture la forma global de la nube: esto es lo que se define con el nombre de histograma del gradiente orientado (se mantendrá la denominación HOG).

2.1 Parámetros fundamentales

La construcción del descriptor de HOG actúa principalmente en dos pasos:

- División de la imagen en un número fijo de celdas y para cada una se calcula un histograma de las orientaciones de los gradientes.
- Cálculo de los histogramas para todas las celdas que se combinan en bloques para obtener la representación global de toda la imagen en forma de vector de características.

A partir de estos pasos cada uno de los histogramas permite capturar información de la orientación dominante de la imagen. Esto se corresponderán con los valores altos en cada uno de los histogramas y al mismo tiempo proporciona información espacial de como están distribuidos estos gradientes en la imagen.

Hay que hacer referencia a una serie de parámetros que van a influir en el descriptor final:

- Tamaño de las celdas;
- Signo del gradiente;
- Número de intervalos en el histograma de orientación;
- Número de celdas en cada bloque.

Todos estos parámetros irán a determinar el número de dimensiones final del descriptor HOG es decir, el número de componentes del vector (en general suele tener dimensiones muy altas). El tamaño total del vector de características, utilizando los parámetros propuesto en este artículo para cada imagen, es 3780 unidades.

Seguidamente se explicarán en detalle cada uno de estos parámetros. En primer lugar se tiene que fijar el tamaño de las celdas, que va a ser fijo para todas las imágenes del set. En general se eligen valores entre 6×6 y 8×8 píxeles. Esta última es la que se utilizará a lo largo de este artículo.

Otro parámetro que se tiene que tener en cuenta es la división del rango de orientaciones en un número de intervalos fijos. En este trabajo se considera la orientación del gradiente sin signo, donde el rango de orientación va desde 0° hasta 180° , con esta suposición dos gradientes con la misma dirección pero sentido inverso se consideran equivalentes y quedan asignados al mismo intervalo (figura 5).

Estos son suposiciones que se adaptan bien al presente caso, pero claramente se pueden cambiar los parámetros ajustándolos cada vez a su caso [9].

Un último parámetro que hay que fijar es en cuantos subintervalos se va a dividir el conjunto de orientaciones ($\delta\theta$). En este trabajo se han usado 9 subintervalos, que es un valor bastante habitual, entonces cada subintervalo tiene un grupo de orientaciones de 20° (figura 5).

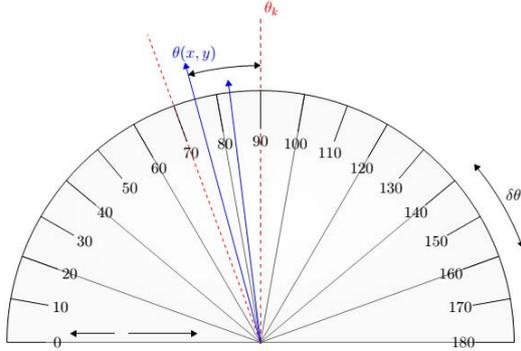


Figura 5: Detalles Parámetros

Una vez que se han fijados todos los parámetros necesarios para el cálculo del histograma, cada uno de los gradientes de la celda, que tiene su propia orientación y magnitud, quedará asignado a uno de los intervalos en función de la orientación del gradiente. Así haciendo cada intervalo ($\delta\theta$), tiene asignado todos aquellos gradientes. cada uno con su respectiva magnitud, que respecta las condiciones antes citadas. Finalmente el valor de uno de estos intervalos en el histograma final se obtiene acumulando la magnitud de todos estos gradientes asignados al intervalo. La expresión matemática del concepto está representada por la ecuación 5.

$$h(k) = \sum_{(x,y) \in C} \omega_k(x,y) \cdot g(x,y) \quad (5)$$

Considerando una celda que contiene un conjunto de píxeles, cada uno con su orientación y magnitud, para calcular el valor del histograma en una determinada posición k , este valor se obtiene como la suma de las magnitudes para todos los píxeles de la celda, ponderados por un factor que determina la asociación del gradiente a dicho intervalo k .

Para que el modelo no sea sensible a pequeñas variaciones del gradiente, vamos a aplicar una interpolación en orientación, es decir se asigna cada píxel a los dos intervalos mas cercanos con un peso proporcional a la distancia de la orientación a cada uno de los intervalos (figura 5).

Esta distancia es la que se utilizará para encontrar el factor de ponderación de un determinado gradiente al intervalo, que se normalizará por el rango

($\delta\theta$) de cada uno de los intervalos del histograma (ecuación 6).

$$\omega_k(x,y) = \max\left(0, 1 - \frac{\theta(x,y) - \theta_k}{\delta\theta}\right) \quad (6)$$

El cálculo del histograma se repite para todas las celdas en que se divide la imagen, de forma que para cada una de las celdas, se tiene su correspondiente histograma, donde se acumulas para cada uno de los intervalos la magnitud de los gradientes ponderados por el factor de asignación que están en la celda.

Similarmente a la interpolación en orientación, se va a considerar la interpolación espacial, donde píxeles muy cercanos en la imagen pueden quedar asignados a celdas diferentes, por lo tanto pequeños cambios en la forma o en la localización de la nube no darán lugar a variaciones significativas en la representación final. Para hacer esto se asigna cada píxel a las cuatro celdas mas cercanas con un peso proporcional a la distancia del píxel al centro de la celda. Se calcularán las distancias del píxel en dirección x ($d_{i,j}^x$) y en dirección y ($d_{i,j}^y$). Similarmente a la ecuación 6, se obtienen las ecuaciones 7 y 8.

$$\omega_{ij}^x(x,y) = \max\left(0, 1 - \frac{d_{ij}^x}{\delta x}\right) \quad (7)$$

$$\omega_{ij}^y(x,y) = \max\left(0, 1 - \frac{d_{ij}^y}{\delta y}\right) \quad (8)$$

Estos factores de asignación de un píxel a una celda en las direcciones x e y se combinan con el factor de asignación de cada gradiente a uno de los intervalos del histograma. Así se obtiene la ponderación final que se aplica a cada una de las magnitudes de los gradientes de todos los píxeles de la imagen para acabar obteniendo el valor concreto del histograma para cada uno de los intervalos, como representado en la ecuación 9.

$$h_{ij}(k) = \sum_{(x,y)} \omega_{ij}^x(x,y) \omega_{ij}^y(x,y) \omega_k(x,y) g(x,y) \quad (9)$$

El próximo paso será obtener el vector de características con la representación final del descriptor a partir de los histogramas de orientación calculados en cada una de las celdas en que se divide la imagen. Esta representación final se va a basar en

normalizar y agrupar estos histogramas en forma de bloques.

Uno de los objetivos de cualquier descriptor debe ser conseguir la máxima invariancia posible a todas aquellas variaciones que se pueden producir en la imagen: iluminación, posición, escala, etc. La necesidad de normalizar localmente cada una de las zonas de la imagen tiene como objetivo, conseguir que la magnitud global del gradiente sea siempre similar, también ante algunos cambios, como por ejemplo el contraste. Para hacer esto vamos a normalizar por bloque, es decir por agrupaciones de celdas vecinas y en general se eligen bloques de 2×2 celdas, obteniendo un único vector por bloque concatenando los histogramas de todas las celdas.

La normalización de los histogramas se va a realizar en el nivel de bloques, es decir se va a normalizar el vector resultante de la concatenación. La normalización se obtiene dividiendo cada uno de los componentes del vector por su norma L2. Esta normalización contribuye a reducir las diferencias en la representación final entre imágenes similares y por lo tanto permite obtener un descriptor final mucho más robusto. Habitualmente los bloques se colocan con una separación de una sola celda entre ellos, tanto en horizontal, como en vertical y la representación final del descriptor HOG se obtiene simplemente concatenando la representación normalizada de todos estos bloques solapados.

3 Base de datos de imágenes utilizadas

Las imágenes de la base de datos que han sido utilizadas para los experimentos tienen características y parámetros (que contribuyen para una buena calidad de la imagen) entre sí, muy diferentes: nitidez, contraste, ruido, distorsión, etc. Estas han sido originadas con diferentes cámaras y consecuentemente con sensores y lentes diferentes, por esto tienen diferentes perspectivas. Todas las muestras utilizadas para el entrenamiento fueron redimensionadas antes de extraer las características. En el momento de la detección, las imágenes de prueba se redimensionaron para adaptarse y permitir la extracción de características HOG en la misma longitud que el vector de características del clasificador. Los parámetros utilizados para los experimentos están representados en la tabla 1.

Han sido utilizados dos conjuntos de imágenes para cada grupo de experimentos. El primer conjunto ha sido usado para el entrenamiento del clasificador y el otro como conjunto de validación. El conjunto total constó de 200 muestras para

cada tipología, es decir 600 muestras para grupo de experimento. En particular el 50% se ha usado para el entrenamiento, y el restante por efectuar las pruebas necesarias.

Tabla 1: Parámetros utilizados para los experimentos

Tamaño celda	Signo gradiente	Orientaciones	Celdas por bloque
8×8 píxel	sin signo	9 de 20°	2×2 celdas

3.1 Tipologías de nubes

En los experimentos, se utilizaron conjuntos de datos incluyendo las varias tipologías de nubes, en particular cúmulos, cirros y estratos. Cada imagen incluye una sola tipología de nube sin superponerse. El aspecto de las nubes como la forma, la densidad y el color, dependen de la naturaleza de sus componentes y las condiciones atmosféricas.

Los contornos netos generalmente indican la presencia de componentes de gran tamaño, con buena densidad, como pueden ser los cúmulos. La transparencia es índice de elementos poco condensados, como los cirros. Los estratos son caracterizados por la opacidad.

4 Resultados obtenidos

Los experimentos se llevaron a cabo usando la herramienta de software matemático, Matlab.

Con el fin de verificar si el método de HOG es válido o no, se han realizado experimentos mediante el uso de imágenes de pruebas, con 100 muestras para cada tipología de nubes. Como se ve en la tabla 2 los resultados van desde un 89% hasta un 96% de detección correcta del tipo de nube.

Tabla 2: Resultados experimentos

Tipología	Muestras usadas	Precisión
Cúmulo	100	90%
Cirro	100	96%
Estrato	100	89%

Entre los resultados, la clase que ha tenido la peor detección ha sido el estrato. Esto está debido principalmente a su forma muy uniforme y con pocos contornos. Contrariamente a cúmulo y cirro donde las fronteras están bien delineadas. En la tabla 3 se presentan los detalles de los experimentos.

Solo en el caso de identificación de los estratos se han tenidos muestras que no han sido clasificadas en ninguna de las clases (NaN).

Tabla 3: Detalles Resultados experimentos

	Cúmulo	Cirro	Estrato	NaN
Cúmulo	92%	8%	0%	—
Cirro	3%	96%	1%	—
Estrato	0%	2%	89%	9%

Se ha observado en algún conjunto de imágenes que los resultados de clasificación mejoran si se ponen algunas restricciones a las imágenes. La restricción más significativa es utilizar imágenes que contienen solo la nube en exámen, sin considerar las partes "vacías" ocupadas por el cielo.

El método es invariante a los diferentes tipos de iluminación, obviamente considerando que para fotos demasiadas oscuras algunas partes de cielo y nube no se podrían distinguir.

5 Conclusiones

Se ha presentado la metodología de histogramas de vectores gradiente HOG par el cálculo de características para la clasificación automática de nubes a partir de imágenes tomadas desde el suelo. Se han presentado las principales propiedades del método en el contexto de la clasificación de nubes.

Los resultados obtenidos son satisfactorios (tabla 3), por lo tanto esta técnica es muy prometedora para la identificación de nubes.

El siguiente paso es mejorar la parte de clasificación usando un algoritmo de ventana deslizante para detectar y usar solo la parte del fotograma que contiene la nube. A diferencia de los algoritmos de ventana deslizante para detectar peatones, en este caso también es importante saber y calcular la posición relativa de la nube en la foto. Eso permite calcular su posición con respecto a la estación de la cámara y estimar la zona de sombra que producirá en tierra según el tipo de nube.

Referencias

- [1] Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection. 2005.
- [2] Soumyabrata Dev, Yee Hui Lee, and Stefan Winkler. Systematic study of color spaces and components for the segmentation of sky/cloud images. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 5102–5106. IEEE, 2014.
- [3] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

- [4] Robert M Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, 1979.
- [5] Tamir Hegazy and Mohamed Hefeeda. Industrial Automation as a Cloud Service. 2015.
- [6] Anna Heinle, Andreas Macke, and Anand Srivastav. Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques*, 3(3):557–567, 2010.
- [7] Junhyuk Hyun and Jeonghyun Baek. Proposing a Fast Circular HOG Descriptor for Detecting Rotated Objects. 2015.
- [8] WP Kegelmeyer Jr. Extraction of cloud statistics from whole sky imaging cameras. Technical report, Sandia National Labs., Livermore, CA (United States), 1994.
- [9] Qing Li and Guangzhou Qu. Matching Between SAR Images and Optical Images based on HOG descriptor. 2010.
- [10] Daniel Limon, Ignacio Alvarado, Teodoro Alamo, Manuel R Arahal, and Eduardo F. Camacho. Robust control of the distributed solar collector field ACUREX using MPC for tracking. In *Proceedings of the 17th World Congress The International Federation of Automatic Control*, pages 958–963, 2008.
- [11] Charles N Long, Jeff M Sabburg, Josep Calbó, and David Pagés. Retrieving cloud characteristics from ground-based daytime color all-sky images. *Journal of Atmospheric and Oceanic Technology*, 23(5):633–652, 2006.
- [12] R Mena, F Rodríguez, M Castilla, and Manuel R Arahal. A prediction model based on neural networks for the energy consumption of a bioclimatic building. *Energy and Buildings*, 82:142–155, 2014.
- [13] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.
- [14] Organización Meteorológica Mundial. <https://http://public.wmo.int/es>, 01 de junio de 2016.
- [15] Asanka S Rodrigo and K H E Perera. Cloud Images Capturing System for Solar Power Level Prediction. 2015.
- [16] Banerji Sugata and Sinha Atreyee. HaarHOH: Improving the HOG Descriptor for Image Classification. 2013.
- [17] Aditya Tatu and Benjamin Kimia. Exploring the Representation Capabilities of the HOG Descriptor. 2011.