CARACTERIZACIÓN CONJUNTA CROMÁTICA-TEXTURAL EN LA IDENTIFICACIÓN DE NUBES

Eduardo Fernández-Cantalejo Padial

Escuela Técnica Superior de Ingeniería de Sevilla, eduardofpadial@gmail.com

Manuel R. Arahál

Escuela Técnica Superior de Ingeniería de Sevilla, arahal@us.es

Resumen

La identificación automática de nubes es un problema que se ha abordado desde diferentes puntos de vista en los últimos años. En particular mediante visión por ordenador usando clasificadores de formas que utilizan características espectrales, texturales y/o cromáticas. El presente estudio afronta el problema sin limitarse a un tipo particular de características. Se realiza una fusión de técnicas buscando el mejor emparejamiento de características abordando cuestiones cruciales como la invariabilidad respecto al ángulo y la diversidad de las muestras.

Palabras clave: GLCM, características cromáticas, clasificación, nubes.

1. INTRODUCCIÓN

La visión artificial es un valor añadido en auge en la industria y sociedad actuales. Sus campos de aplicación se amplían cada día hasta el punto de ser una parte vital en muchas áreas de la automatización.

El caso de la identificación de nubes tiene especial relevancia en diversas aplicaciones relacionadas con la industria solar (producción de energía, edificios, etc.) y en meteorología. En estas aplicaciones la identificación de nubes es un paso previo para la predicción del recurso solar a corto plazo. Al predecir posibles situaciones adversas se puede reaccionar con antelación para evitar o paliar los efectos más perniciosos.

En la literatura se han empleado diversas técnicas para la clasificación automática de nubes, que abarcan desde la distribución de colores de cada píxel individualmente en el espacio RGB [7], de manera sistemática [3] hasta el análisis de texturas mediante matrices de co-ocurrencia de nivel de gris [10]. Todos estos estudios realizan un análisis de cada una de cada técnica individualmente.

El presente artículo estudia la posibilidad de emplear técnicas diferentes y fusionar los resultados para la consecución de mejores resultados a los

obtenidas de manera independiente.

Este nuevo enfoque se centra en la caracterización híbrida de características cromáticas conjuntamente con características texturales. Para ello se ha realizado un estudio individual de cada una de ellas para seleccionar las mejores características de cada técnica y posteriormente aplicar la fusión en la técnica mixta cromático-textural.

El objetivo final consiste en la descripción y clasificación de nubes desde una estación con base en la superficie terrestre, en contraposición a las imágenes tomadas desde los satélites. El tipo de nube será determinante para la estimación de la altura en que éstas se forman y desarrollan y de su posible evolución, en dirección y en altura.

2. ANTECEDENTES

Durante las últimas décadas se han desarrollado numerosas técnicas para la clasificación de patrones basados en diferentes características como la forma, el color y la textura. Cada una de estas aproximaciones muestra su conveniencia en función del problema y el tipo de patrón que se aborde. En el caso del reconocimiento e identificación de nubes la forma no será determinante debido a su alta variabilidad, por lo que se descartará este tipo de características. El problema se afronta, por tanto, como la búsqueda de las mejores características cromáticas y texturales que, primero de forma independiente y posteriormente de forma conjunta, mejor describa cada tipo de nube.

El interés de la detección y clasificación de nubes es más notorio en algunas aplicaciones que en otras. Por ejemplo, en el caso de plantas fotovoltaicas las nubes pueden producir una caída instantánea de la radiación solar directa (RSD) que se traduce en una gran perturbación en la producción que ha de ser absorbida por el sistema eléctrico [6]. En el caso de plantas termosolares el efecto puede ser aliviado por la inercia térmica del sistema, sobre todo si se usa acumulación de calor en aceite, sales u otro medio [8]. Para otras aplicaciones la radiación solar indirecta (RSI) juega un papel importante por lo que el paso de nubes dispersas puede tener escasas consecuencias [9], por

Tabla 1: Clasificación de nubes SWIMCAT conforme WMO.

Etiqueta SWIMCAT	Nro. imágenes	Tipo de nube conforme WMO	Descripción
A	224	Cielo despejado	Sin nubes o nubosidad < 10 %
B	89	Cirrocúmulos y altocúmulos	Patrón de pequeñas nubes
C	251	Cúmulonimbus y nimboestratos	Nubes oscuras, cubiertas y densas
D	135	Cúmulos	Nubes bajas e hinchadas con bordes definidos, blancas o gris muy claras
E	85	Estrato y altoestratos	Capa de nubes medias y bajas uniformes usualmente cubierta y gris
		Estratocúmulos	Capa de nubes medias y bajas, rota o completa, blanca o gris
		Cirros y cirroestratos	Nubes altas y finas de humo o cubriendo el cielo

lo que basta con conocer la nubosidad general. En todos los casos la predicción de la nubosidad y el estudio de su impacto tienen un gran interés para optimizar la operación de la planta [4].

3. METODOLOGÍA

El análisis se planteará como una búsqueda exhaustiva y sistemática de las características de manera que tengan el mayor grado de independencia del tipo de imagen y las condiciones de luminosidad en el momento de la obtención de éstas. Por ello se empleará un método basado en los siguientes principios:

- Colección de imágenes extensa, de forma que se incluya el mayor número de escenarios posibles en el estudio, dentro de una normalidad de funcionamiento, esto es, en los momentos entre el amanecer y el atardecer, pero sin incluir éstos.
- Normalización estadística mediante la eliminación de la media y ponderado con la varianza del conjunto de características obtenidas.
- Clasificador, se aplicará únicamente el clasificador de mínima distancia, con un desempeño aceptable para la evaluación del potencial de mejora del método presentado. Estudios posteriores evaluarán el mejor clasificador a emplear.

Una vez definidos los elementos comunes se presentarán tres técnicas para la clasificación de nubes: las dos primeras conforme a técnicas incluidas en estudios pasados basados en tratar las características cromáticas y texturales de forma separada y posteriormente se presentará un método que integra ambos tipos de características para su uso en un mismo clasificador.

La evaluación de los resultados se ha realizado mediante el procesamiento de las imágenes de la base de datos SWIMCAT que previamente se han empleado para su entrenamiento, ya que el objetivo de este trabajo es el estudio de viabilidad del método de clasificación con características de diferente naturaleza y se considera suficiente con este tipo de validación.

A continuación se profundiza en cada uno de estos aspectos del análisis.

3.1. COLECCIÓN DE IMÁGENES

La calidad y dimensión de la muestra de imágenes debe ser suficientemente extensa y significativa para que los resultados puedan considerarse válidos y reproducibles. Se busca la universalización de los resultados mediante la exhaustividad en el tipo y cantidad muestras en el conjunto de estudio.

Asimismo debe escogerse un conjunto de tipos de nube que sea capaz de satisfacer la funcionalidad para la que se realiza el estudio. Los resultados de este estudio pretende la estimación de la altura de las nubes en el cielo por lo que la clasificación de la que se parte responde a aquellas nubes que tengan un rango de altura específico.

Este estudio ha empleado la base de datos SWIMCAT de nubes (Singapore Whole sky IMaging Categories), empleada en [2]. Esta colección de fotografías contiene 784 imágenes clasificadas por personal experto del Servicio Meteorológico de Singapur en cinco categorías como las mostradas en la Figura 1 y conforme a lo recogido en la Tabla 1. Este conjunto se considera como muestra válida para un estudio en profundidad.

El conjunto de nubes que forman parte de la base de datos SWIMCAT es suficientemente extensa y se divide en cinco categorías. Sin embargo, esta base de datos no considera todas las categorías de

Figura 1: Tipos de nube A, B, C, D y E

finidas por la World Meteorological Organization (WMO) en su Atlas Internacional de Nubes.

Mediante la identificación de estas clases de nubes se podrá realizar una estimación de la altura de las nubes.

3.2. NORMALIZACIÓN

El tratamiento de estas distancias se realiza con sus valores en el intervalo $[0,255]$ [3]. Sin embargo, para poder realizar la fusión posterior con las características texturales se necesitará la normalización de los valores entre $[0,1]$. Sin embargo, también se desea la obtención de características robustas e independientes de las condiciones externas durante la toma de imágenes, por lo que en este punto se realizarán dos tipos de normalización para estimar cuál resultará con un mejor clustering. Normalización estadística: $Dist_{est} = \frac{Dist - \mu}{\sigma}$

3.3. CLASIFICADOR DE MÍNIMA DISTANCIA

Se ha optado por el clasificador más sencillo para validar los resultados ya que su funcionamiento será tanto más correcto cuanto más separados se encuentren los clusters entre sí. De esta manera se está positivizando el distanciamiento entre las categorías de nubes.

Sea E un conjunto dotado de una (pseudo)métrica $d : E \times E \rightarrow R$. Supongamos que se dan m prototipos $p_1, p_2, \dots, p_m \in E$, tales que $p_i \in c_i$ representa a la clase c_i ($i = 1, 2, \dots, m$), es decir, un prototipo por clase. Una clasificación por distancia mínima se define para cada $x \in E$ como $x \in c_i \Leftrightarrow d(x, p_i) \leq d(x, p_j); j = 1, 2, \dots, m$ (1.4) donde d es la (pseudo)métrica definida en E .

En este tipo de clasificador, la fase de aprendizaje consiste únicamente en la elección de un buen representante o prototipo de cada clase en el conjunto de prototipos que pertenecen a la misma. Normalmente el representante elegido suele ser aquel que se encuentra más centrado dentro de la distribución de los prototipos en la clase. La Figura 2 muestra la clasificación de un punto en un espacio con tres clases diferentes.

Figura 2: Clasificador de mínima distancia

3.4. FAMILIAS DE CARACTERÍSTICAS

3.4.1. Características cromáticas

Las características cromáticas son a menudo consideradas como poco robustas debido que son sensibles a cambios en las condiciones lumínicas y de escala, entre otros aspectos. Sin embargo, hay estudios que emplean estas técnicas para la identificación de nubes con éxito y que se revisará en este estudio con una base de datos de imágenes de gran población y diversidad.

Toda imagen en color puede expresarse en diversos espacios de color de forma que el color se distribuye de forma diferente en cada uno de ellos. Cada uno de los planos de un espacio de color se expresa como escala de grises.

Un estudio exhaustivo debe cubrir todas las posibilidades. Por ello, se emplearán todos los espacios de color descritos y empleados en estudios similares, esto es, se considerarán los siguientes planos de color: RGB, HSV, YIQ, $L^*a^*b^*$ y los derivados (R-B), (R/B), (B-R/B+R) y croma, donde éste último se calcula como $C = \max(R, G, B) - \min(R, G, B)$. Cada uno de estos 16 planos de color constituye una de las 16 características cromáticas de píxel en cada imagen.

Entre las diferentes técnicas (índice de Pearson, Análisis de Componentes Principales, ...) se ha empleado el análisis de la Distancia Geométrica Euclídea (DGE) a la diagonal del espacio de color como descriptor cromático, pues se busca evitar generalizaciones que puedan distorsionar las características individuales de cada punto de la imagen.

El análisis de la DGE consiste en calcular la distancia de un píxel para cada tupla de las 16 características previamente definidas a la diagonal de un cubo formado por el máximo valor en cada uno de estas características, tal como muestra la Figura 3.

Tabla 2: Relación y fuente de las características texturales.

Etiqueta	Abrev.	Característica	Haralick 1973	Soh 1999	Clausi 2002
1	ENE	Segundo momento angular (Energía) (Uniformidad)	1	1	2
2	CON	Contraste (Inercia)	2	2	5
3	COR	Correlación	3	3	8
4	SCV	Suma de cuadrados: Varianza	4		
5	HOM	Momento de la diferencia inversa (Homogeneidad)	5	4	7
6	MDS	Media de la suma	6		
7	VDS	Varianza de la suma	7		
8	ENS	Entropía de la suma	8		
9	ENT	Entropía	9	5	3
10	VDD	Varianza de la diferencia	10		
11	EDD	Entropía de la diferencia	11		
12	MIC1	Medidas informativas de correlación (1)	12		
13	MIC2	Medidas informativas de correlación (2)	13		
14	CMC	Coefficiente de máxima correlación	14		
15	ACR	Autocorrelación		6	
16	DIS	Disimilaridad		7	4
17	TDG	Tono de grupo (Cluster shadow)		8	
18	RDG	Relevancia de grupo (Cluster Prominence)		9	
19	PDM	Probabilidad máxima		10	1
20	DIN	Diferencia inversa			6
21	INN	Diferencia inversa normalizada			9
22	IDN	Momento de la diferencia inversa normalizada			10

Figura 3: Distancia de un píxel en el espacio RGB

El cálculo de la distancia se basa en sencillas reglas geométricas, obteniendo la fórmula siguiente[7]:

$$Dist = |\overline{D}| \sin \left(\arccos \left(\frac{|\overline{C} - \overline{D}| - |\overline{D}| - |\overline{C}|}{-2|\overline{D}||\overline{C}|} \right) \right) \quad (1)$$

donde de la Figura 4 se obtiene que los módulos:

$$|\overline{D}| = \sqrt{255^2 + 255^2 + 255^2} = 441,67$$

$$|\overline{C}| = \sqrt{R^2 + G^2 + B^2}$$

La distancia de cada píxel a la línea de gris para cada trupla de planos de color puede tratarse como una característica independiente.

Las distancias normalizadas constituirán las ca-

racterísticas del vector de características que servirán de entrada al clasificador.

Figura 4: Distancia de un píxel en el espacio RGB

$$Vect = [Dist_{RGB}, Dist_{RGH}, \dots, Dist_{HSV}, \dots] \quad (2)$$

Cuando se emplean gran cantidad de datos estadísticamente distribuidos, es usual descartar los datos inferiores al cuantil 25% y superiores al 75%, sobre todo cuando se emplearán estos datos para generar los prototipos de un clasificador basado en distancias. Los límites intercuantiles se han calculado como:

- Distancia intercuantil: $IQD = Q_3 - Q_1$
- Límite inferior: $IQI = Q_1 - 1,5 \cdot IQD$
- Límite superior: $IQS = Q_3 + 1,5 \cdot IQD$

En este estudio se han considerado las 560 tuplas de características cromáticas posibles para una muestra de 213.600 píxeles extraídos de las imágenes de la base de datos SWIMCAT. Por tanto, cada tupla supone una característica diferente en el clasificador de mínima distancia. Agrupando varios pares de tuplas se mejoran los resultados de clasificación. La Figura 5 muestra las regiones de clasificación generadas por una sola tupla y agrupadas en parejas de tuplas.

Figura 5: Distribución de clases en 1D y 2D

3.4.2. Características texturales

Las características texturales son magnitudes estadísticas extraídas del valor del entorno de un punto, a diferencia de las características cromáticas que sólo depende del valor tonal de cada píxel de manera individual.

Así, se asume que la información de contexto textural en una imagen se encuentra definida por la relación espacial general o “media” con la que los tonos de gris tienen unos con otros. Esta información textural se considera adecuadamente especificada por la matriz de frecuencias relativas P_{ij} con la cual dos celdas vecinas separadas una distancia d tienen lugar en la imagen, una con un tono gris i y la otra con un tono de gris j . Tales matrices de frecuencias de tonos de gris con dependencia espacial son función de la relación angular entre las celdas vecinas así como de la distancia entre ellas.

Las características texturales que se han considerado son las definidas por Haralick (1973)[5], Soh (1999) [10] y Clausi (2002) [1] que se construyen como Matriz de Coocurrencia de Niveles de Gris (GLCM), presentadas en la Tabla 2 y cuya formulación matemática se recoge en la Tabla 6.

Como se ha indicado, el resultado durante el tratamiento de textura mediante la matriz de coocurrencia depende de dos parámetros: distancia entre vecinos y ángulo entre vecinos. Este estudio considerará las vecindades de 1 píxel y cuatro ángulos: 0° , 90° , 135° y 270° , como se indica en la Figura 6.

En este estudio se han considerado todas las imágenes de la base de datos, siendo el entorno

Figura 6: Ángulos de vecindad en la matriz de coocurrencia

del punto la imagen completa, de tamaño 192x192 píxeles.

3.5. TÉCNICA MIXTA DE CLASIFICACIÓN

Las características texturales se han calculado a partir de 534 muestras de imágenes de la base de datos de SWIMCAT, lo que suponen 2136 características si tenemos en cuenta los cuatro ángulos, tal como se muestra en la Figura 6, y extraído cada una de las 22 características texturales, de manera que se obtiene un vector de 2136×22 características texturales.

Las características cromáticas se han extraídos de las mismas muestras que las características texturales, a razón de 1.000 píxeles por cada tipo de nube, por lo que se obtiene una muestra final de 213.600 píxeles y constituyendo un vector de características de dimensiones 213600×560 (filas x columnas). Se puede identificar, por tanto, el valor cromático de cada píxel y el valor correspondiente de la textura asociada.

A partir de esta información se pueden crear matrices de características de tal manera que cada columna corresponde a un píxel y cada fila puede significar indistintamente una característica cromática o textural indistintamente. A partir de estas matrices se pueden generar los vectores prototipos de cada que servirán para evaluar cada píxel de cualquier imagen de forma individual.

4. RESULTADOS

4.1. CARACTERÍSTICAS CROMÁTICAS

La aplicación del método de la Distancia Geométrica Euclídea a las 16 características cromáticas definidas anteriormente ofrece 560 combinaciones de distancias posibles, que es el número de características que se han estudiado

Tabla 3: Resultado de la caracterización cromática.

Características		Tasa de error en clasificación					
Trupla 1	Trupla 2	A	B	C	D	E	Global
	R.b*_R/B	0,00 %	82,90 %	12,74 %	13,83 %	42,77 %	26,59 %
G.V.R/B	H.Q.R/B	0,00 %	22,64 %	5,17 %	12,54 %	22,53 %	11,34 %

Tabla 4: Resultado de la caracterización textural.

Características				Tasa de error en clasificación					
T#1	T#2	T#3	T#4	A	B	C	D	E	Global
-	-	ENT	ACR	2,00 %	16,29 %	15,54 %	38,70 %	25,88 %	16,87 %
-	EDD	ENT	VDS	2,79 %	8,71 %	17,93 %	25,74 %	17,65 %	13,87 %
CON	RDG	ENS	MDS	2,01 %	2,25 %	15,94 %	24,07 %	7,35 %	10,87 %

por cada píxel. Debido al elevado número de variables y la dificultad de segmentar manualmente cada imagen se han tomado muestras significativas de cada tipo de nube para entrenar el clasificador.

En esta investigación se han calculado de forma exhaustiva todas las posibles combinaciones para el clasificador de mínima distancia y se han obtenido tasas de error en la clasificación aceptables para 1 sola dimensión y buenos en caracterizaciones con dos distancias, como se observa en la Tabla 3.

La Figura 5 muestra la distribución de los píxeles caracterizados por lo que se puede observar las distancias que existen entre las cinco clases de nubes. Debe notarse que, debido al solapamiento se obtienen tasas de error altas para las clases B y E.

Algunas de las figuras obtenidas apuntan a la existencia de subclases de nubes, lo cual es positivo para clasificaciones más exigentes que requieras de categorías de nubes no incluidas en este estudio.

4.2. CARACTERÍSTICAS TEXTURALES

Un componente esencial del marco conceptual de la textura es una medida, o más preciso, cuatro medidas muy relacionadas de las que todas las características derivan. Estas medidas son matrices denominadas matrices de tono de gris de vecinos angulares dependientes del espacio. Estas matrices son función de una relación angular entre las celdas adyacentes así como de la distancia entre ellas. El análisis se ha realizado para valores de distancia $d=1$ en sus cuatro direcciones (0° , 45° , 90° y 135°).

Las características texturales empleadas son 22 variables que se han estudiado en emparejamientos de hasta cuatro variables, mejorando el rendimiento global con cada nueva característica que se empleaba en el clasificador. La Tabla 4 muestra los

resultados para los mejores grupos de dos, tres y cuatro características texturales.

4.3. TÉCNICA MIXTA CROMÁTICO-TEXTURAL

Finalmente se han realizado clasificaciones empleando conjuntamente una selección de características cromáticas y texturales para estudiar si una aplicación conjunta podría obtener mejores resultados que cada una de estas técnicas por separado.

Por ello se ha aplicado un vector de características en el que por cada subimagen I de tamaño $N_x \times N_y$ píxeles, en que se divide una imagen a segmentar, se ha asignado a cada uno de estos píxeles el mismo valor de característica textural correspondiente a la muestra completa.

Las características que se han seleccionado son las que obtuvieron mejor resultado de forma independiente. Este emparejamiento es un candidato firme a obtener un buen resultado, si bien la suma de dos resultados óptimos no garantiza el óptimo global. Del grupo cromáticas: B, L y R/B y del grupo texturales: Contraste, Relevancia de grupo, Media de la Suma y Entropía de la Suma. El resultado es la obtención de una tasa de error global del 8.70 %, con un buen rendimiento de clasificación para todas las clases salvo para los cúmulos que quedaría en un 21,44 %, tal como se muestra en la Tabla 5.

5. CONCLUSIONES

Este estudio pone de manifiesto que la aplicación de características descriptivas de diferentes familias, como son las cromáticas y las texturales, pueden fusionarse para obtener mejores resultados durante la clasificación.

La caracterización cromáticas depende únicamente del valor del píxel, mientras que la caracteriza-

Tabla 5: Resultado de la técnica mixta.

Características				Tasa de error en clasificación					
T#1	T#2	T#3	T#4	A	B	C	D	E	Global
CON	RDG	ENS	MDS	0,25 %	6,05 %	5,05 %	21,44 %	10,61 %	8,70 %
B.L.R/B		H.I.R/B							

ción textural se obtiene de magnitudes estadísticas aplicadas a subimágenes de $N_x \times N_y$ píxeles. La integración de ambos tipos en una misma matriz de características ha posibilitado esta clasificación conjunta.

En este análisis se han revisado técnicas de manera exhaustiva conocidas a una base de datos extensa de imágenes como validación de los resultados previo a la aplicación de la técnica mixta.

Referencias

- [1] Clausi, D. A. (2002). “An analysis of co-occurrence texture statistics as a function of grey level quantization”, *Canadian Journal of remote sensing*, 28(1), pp 45-62.
- [2] Dev, S., Lee, Y. H. y Winkler, S. (2015) “Categorization of cloud image patches using an improved texton-based approach”, *Image Processing (ICIP), 2015 IEEE International Conference on*, pp 422-426.
- [3] Dev, S., Lee, Y. H., y Winkler, S. (2014) “Systematic study of color spaces and components for the segmentation of sky/cloud images”, *IEEE International Conference on Image Processing (ICIP)*, pp. 5102-5106.
- [4] Gallego, A.J., Fele, F., Camacho, E.F. y Yebra, L. (2013) “Observer-based Model Predictive Control of a parabolic-trough field”, *Solar Energy*, 97, pp 426-435.
- [5] Haralick, R. M. (1979) “Statistical and structural approaches to texture”, *Proceedings of the IEEE*, 67(5), pp 786-804.
- [6] Limón, D., Alvarado, I., Álamo, T. Arahall, M.R., y Camacho, E.F., (2008) “Robust control of the distributed solar collector field ACUREX using MPC for tracking”, *IFAC Proceedings Volumes*, 41(2), pp 958-963.
- [7] Mantelli Neto, S. L., von Wangenheim, A., Pereira, E. B., y Comunello, E. (2010). “The use of Euclidean geometric distance on RGB color space for the classification of sky and cloud patterns”. *Journal of Atmospheric and Oceanic Technology*, 27(9), pp 1504-1517.
- [8] Medrano, M., Gil, A., y Martorell, I., y Potau, X., y Cabeza, L.F., (2010) “State of the art on high-temperature thermal energy storage for power generation. Part 2-Case studies”, *Renewable and Sustainable Energy Reviews*, 14 (1), pp 56-72. Elsevier.
- [9] Pawlowski, A., Mendoza, J.L., y Guzmán, J.L., y Berenguel, M., y Ación, F.G., y Dormido, S., (2015) “Selective pH and dissolved oxygen control strategy for a raceway reactor within an event-based approach”, *Control Engineering Practice*, 44, pp 209-218.
- [10] Soh, L. K., y Tsatsoulis, C. (1999). “Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices”, *Geoscience and Remote Sensing, IEEE Transactions on*, 37(2), pp 780-795.

Tabla 6: Fórmula de características texturales.

Característica	Fórmula
Energía	$f_1 = \sum_i \sum_j p(i, j)^2$
Contraste	$f_2 = \sum_{i=1}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}_{ i-j =n}$
Correlación	$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Suma de cuadrados: Varianza	$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j)$
Homogeneidad	$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j)$
Media de la suma	$f_6 = \sum_{i=2}^{2N_g} i \cdot p_{x+y}(i)$
Varianza de la suma	$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i)$
Entropía de la suma	$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i))$
Entropía	$f_9 = - \sum_i \sum_j p(i, j) \log(p(i, j))$
Varianza de la diferencia	$f_{10} = \text{variance}(p_{x-y})$
Entropía de la diferencia	$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i))$
Medidas informativas de correlación (1)	$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$ donde: $HXY = - \sum_i \sum_j p(i, j) \log(p(i, j))$ $HXY1 = - \sum_i \sum_j p(i, j) \log(p_x(i) p_y(j))$ $HXY2 = - \sum_i \sum_j p_x(i) p_y(j) \log(p(i, j))$
Medidas informativas de correlación (2)	$f_{13} = [1 - e^{-2 \cdot (HXY2 - HXY)}]^{1/2}$
Coeficiente de máxima correlación	$f_{14} = [\text{Segundo mayor autovalor de } Q]^{1/2}$ donde: $Q(i, j) = \sum_k \frac{p(i, k) \cdot p(j, k)}{p_x(i) \cdot p_y(j)}$
Autocorrelación	$f_{15} = \sum_i \sum_j (ij) p(i, j)$
Disimilaridad	$f_{16} = \sum_i \sum_j i - j \cdot p(i, j)$
Tono de grupo (Cluster shadow)	$f_{17} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j)$
Relevancia de grupo (Cluster Prominence)	$f_{18} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j)$
Probabilidad máxima	$f_{19} = \sum_{i,j} \text{MAX}[p(i, j)]$
Diferencia inversa	$f_{20} = \sum_{i,j=1}^G \frac{C_{ij}}{1+ i-j }$
Diferencia inversa normalizada	$f_{21} = \sum_{i,j=1}^G \frac{C_{ij}}{1+ i-j ^2/G^2}$
Momento de la diferencia inversa normalizada	$f_{22} = \sum_{i,j=1}^G \frac{C_{ij}}{1+(i-j)^2/G^2}$