Rev.Acad.Canar.Cienc., III (Núm. 2),105-118 (1991)

SOBRE LA GENERACION DE NUEVOS ESPACIOS DE COEFICIENTES DE REPRESENTACION EN VISION ARTIFICIAL

J.A. Muñoz Blanco, O.Bolivar Toledo, S. Candela Solá, R. Moreno Díaz
Departamento de Informática y Sistemas
Universidad de Las Plamas de Gran Canaria
Apartado de Correos 322,Las Palmas de G.C.

ABSTRACT

The objectives are as follows: Given an image, it is aimed to obtain an alternative description, from where to proceed to generate a new image, according to certain selection rules, which, in essence, consists in selecting coefficients in a "complete" representation. This is performed in a hierarchycal structure by "computation layers", which generate new spaces (higher level tranformated images), the last of which permits a decision layer for visual recognition and/or classification.

RESUMEN

Los objetivos que se pretenden alcanzar son los siguientes: Dada una imagen pretendemos obtener una representación alternativa de forma que, a partir de ella se pueda generar una nueva imagen de acuerdo a ciertas reglas de selección que, en esencia, consisten en extraer coeficientes de una representación completa. Esto se lleva

a cabo a traves de una estructura jerarquica, mediante "computacion por capas", que genera nuevos espacios (imagenes transformadas de mas alto nivel), permitiendonos la última de ellas tomar una decision para reconocimiento y/o clasificacion visual.

1. INTRODUCCION

En anteriores trabajos [3,5] los autores desarrollaron ideas sobre nuevas arquitecturas para Visión. En dichos trabajos se expone que cualquier sistema se define concretando sus espacios de entrada y salida y la estructura relacional que los engarza, describiendo ésta última el nivel de descripción y la función del sistema.

La semantica propia del sistema puede repartirse entre los espacios de entrada y salida y la estructura relacional. Para un mismo comportamiento global, cuanto mayor sea la capacidad operativa de los simbolos codificados sobre variables de entrada, menor será la complejidad del cálculo necesario para explicar la conducta observada

Los párrafos anteriores tienen una inmediata interpretación en el sentido de que ha de existir una especie de **principio de conservación** de la cantidad de complejidad **contenida** en las salidas o decisiones del sistema cuando esta se reparte entre los espacios de entrada y salida, y de forma que al aumentar la complejidad y nivel de los datos de entrada, se reduce consecuentemente la complejidad de las reglas de decisión u operación. Y viceversa.

En otras palabras, si los datos del espacio de entrada tienen una carga simbólica elevada, las reglas de decisión pueden ser sencillas, para obtener una carga alta de significado o complejidad en el espacio de salida. Y viceversa.

2. GENERACIÓN DE ESPACIOS DE COEFICIENTES DE REPRESENTACION

El objetivo del presente trabajo puede ser resumido de la siguiente forma: Dada una imagen , se trata de obtener una descripcion alternativa de la misma, sobre la que proceder para generar un nuevo patron segun las normas de selección que se exponen en 3. Los métodos de 3 son en esencia de selección de coeficientes en una representación completa. Por lo tanto, la función de la capa computacional es la de generar un espacio de las mismas dimensiones que el de salida, construido a partir de **coeficientes** de representación, de manera que tal representación sea completa.

Por el momento, nos restringimos a procedimientos lineales de generación de coeficientes de representación.

Si f(x,y) es una función de buen comportamiento y definida en los intervalos (0,a), (0,a), siempre es posible encontrar un conjunto (infinito) de coeficientes descriptores en ese intervalo. Si f(x,y) está discretizada (limitada en banda), el conjunto es finito, y, por supuesto, igual al formado por el número de elementos discretos de f(x,y), puesto que los coeficientes son independientes y no admitimos redundancia.

La forma típica de plantear el problema es el **expandir** f(x,y) en términos de ciertos coeficientes F(w1,w2)

$$F(w1,w2) = \int f(x,y)\, \alpha(x,y,w1,w2)\, dxdy$$

tal que si los α son linealmente independientes, es posible **resolver** f(x,y), es decir, computar

$$f(x,y) = \int F(w1,w2)\, \beta(w1,w2,x,y)\, dw1dw2$$

Resolver aquí es utilizado en el sentido algorítmico, y no en el sentido analítico estricto.

Un problema típico de análisis es el de la **transformada inversa**, es decir, dado α determinar analíticamente β, que es un problema clásico en ecuaciones integrales.

Desde el punto de vista de la matemática finita, a la que se está avocado a priori en proceso de imágenes, el problema anterior tiene un planteamiento real mucho más transparente y sencillo, aunque quizás frustrante por lo poco sofisticado.

Para este problema, una imagen es un conjunto de NxN pixels, con dos índices de orden i,j, y un valor de luminancia f. Es decir, una especie de matriz {fij}. La ordenación de i,j es elegida a priori, y de hecho, en todos los sistemas prácticos, el sistema de barrido es tal que un solo número de orden (entre 0 y NxN) basta para determinar unívocamente la situación del pixel y leer su intensidad. Es decir, una imagen es un conjunto ordenado de pixels, del 1 al NxN, con valores de luminancia Ii. Cuestión diferente será que queramos hacer corresponder las operaciones sobre los pixels a operaciones realizadas por sistemas con arquitectura paralela, en cuyo caso se deberá deshacer y rehacer, la ordenación. Para nuestros propósitos, basta considerar el caso unidimensional.

De acuerdo con ésto, una imagen es un conjunto de valores fi ($i=1....N^2$). Admitimos, a priori, que no hay posiciones privilegiadas ni minusvalorables (si las hubiese, adjudicaríamos más pixels -números i- aumentando la resolución de las zonas privilegiadas, y viceversa para las de minus-valor).

Generar un espacio de coeficientes consiste en buscar unos valores Fj

$$Fj = \sum_{i} \alpha ji \; fi$$

tales que los αji son linealmente independientes, ésto es

$$|\alpha ji| \# 0$$

para que el problema inverso, de **resolución de fi**, es decir

$$fi = \sum_j \beta ij\, Fj \quad \text{con} \quad \beta = \alpha^{-1}$$

tenga solución única. Entonces diremos que Fj es una representación **completa** y no redundante, en coeficientes, de fi.

Lo primero que salta a la vista en este planteamiento es que los αij pueden ser cualesquiera, con la única restricción de independencia lineal, o sea $|\alpha ij| \# 0$. Si son ortogonales, es decir, si

$$\Sigma\, \alpha il \alpha lj = \delta ij$$

entonces, la analítica se simplifica considerablemente, aunque no necesariamente la algorítmica.

La justificación de la expansión de las fi en términos de los coeficientes Fj, para el caso de proceso de imágenes y reconocimiento visual, se basa en una conjetura, que aún no ha sido reducida a teoremas formales, y por lo tanto, no tiene ningún sustento teórico. La conjetura es la siguiente:

Dado un conjunto de posibles imágenes, C(f), éste ha de ser dividido o clasificado en clases Ck, en general de forma probabilística, o borrosa. Entonces: la nitidez de clasificación, en el sentido de mínima probabilidad de error en asignación a clases, o mínima borrosidad, se puede conseguir cuando C(f) se representa en el espacio alternativo C(F). No sólo la conjetura no ha sido formalizada en términos de sus adecuados teoremas, sino que, en general, se procede al inverso. Es decir, se admite a priori que la conjetura es válida, se define el conjunto C(f) y se buscan representaciones alternativas de f, F, para generar C(F) tal que, experimentalmente se pueda mostrar que las tasa de error en la tarea de clasificación (o reconocimiento) están por debajo de lo inaceptable. Es decir, que el sistema funcione aceptablemente.

Sin embargo, mientras que la conjetura no haya sido reducida a teoremas probables o pendientes de prueba, el método sigue siendo válido, en cuanto proporciona pistas para consolidar o eliminar tal conjetura, así como para construir sistemas que resuelvan situaciones prácticas concretas.

Existe un caso particular de la conjetura anterior que ha dado un juego considerable en los procesos visuales y que fue importada, en cierto modo, por el papel que jugó y juega en los procesos auditivos y en los problemas de transmisión y comunicaciones. Es la transformada de Fourier y sus variantes tipo Fourier-Bessel. Otro caso es el de los momentos, donde a partir del espacio f se llega al F por potencias de las coordenadas que han de ser mantenidas dentro de cierto rango. Esta técnica de representación alternativa tiene un origen de importación no perceptual, más relacionado con la física, pero que, en esencia, se enmarca dentro de la conjetura general ya descrita. Por otro lado [6], la representación o cálculo de momentos de bajo orden, para una amplia clase de imágenes, permite atacar racionalmente el problema de los universales, es decir, el de generación por máquina de los parámetros que han de caracterizar las transformaciones afines que reduzcan los descriptores de una imagen a la aparición standard, con vistas al reconocimiento o clasificación posterior.

Los experimentos que se llevaron a cabo [4] fueron en la línea de la citada metodología: poner a prueba la validez de la idea de la representación en espacios de coeficientes de Fourier y Momentos, aunque, consecuentemente, ahora se hace a dos niveles con estructuración jerárquica.analizado en las

3 GENERACIÓN DE PATRONES EN ESPACIOS TRANSFORMADOS

La generación de un patrón en un espacio **transformado** consiste en la representación, como nueva imagen, de los **pixels**, Fij:

$$Fij = \mathcal{T}ij(Fkl)$$

El patrón obtenido depende lógicamente de la regla de ordenación seguida para la representación por los subíndices i, j, aunque en la mayoría de los casos (sobre todo cuando $\mathcal{T}$ es una transformación lineal) existe una **ordenación natural**, indicada por el caracter de la transformación.

Nos interesan los casos de espacios de representación por coeficientes, es decir, los casos en que Fij son los coeficientes de una expansión de Fkl.

Un patrón en el espacio transformado es una nueva imagen, Γ'ij obtenida de Fij por una regla de selección

$$\Gamma'ij = R\,(Fij)$$

Consideraremos que la regla de seleccion, R, una vez definida, actúa siempre localmente (aunque la aplicación de la regla puede requerir una operación previa global). La regla de selección tiene normalmente la naturaleza de una decisión del tipo IF...THEN...ELSE.

La generación de un patrón en el espacio transformado actúa por consiguiente, en dos pasos:

a) Definir la ordenación de los coeficientes Fij

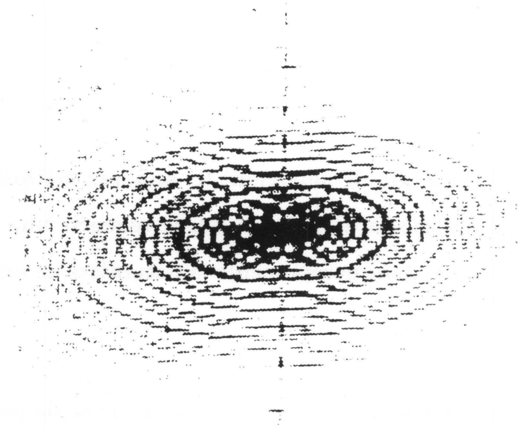
b) Aplicar la regla de selección R.

Consideremos, por ejemplo, que el espacio transformado sea el del módulo de la transformada de Fourier, es decir:

$$F(w1,w2) = |\mathcal{T}[f(x,y)]|$$

donde $\mathcal{T}$ es ahora la transformación bidimensional cartesiana de Fourier.

Existe una ordenación natural obvia de w1 y w2 en el espacio bidimensional y coordenadas cartesianas, que, con la regla de selección $f'12=F(w1,w2)$ nos da la representación típica de la transformada de Fourier en cuatro cuadrantes, dos de los cuales son redundantes.



Nótese que f' puede obtenerse despues de aplicar cualquiera de los procedimientos adicionales de preproceso tratados en [2]. Una regla sencilla de generación del patrón puede ser simplemente a través del umbralizado de las **imágenes de coeficientes** preprocesadas. Es decir, de f' se pasa a Γ tal que

$$\Gamma = \text{preproceso}\,(f')$$

y el patrón Γ' se obtiene por la regla

$$\Gamma' = \begin{cases} \Gamma & \text{si } \Gamma \geqq \Theta \\ 0 & \text{si } \Gamma < \Theta \end{cases}$$

donde Θ es, en general, un umbral local adaptativo.

La regla anterior puede elaborarse adecuadamente, con condiciones diversas, $\mathbb{C}$, tales que

$$\Gamma' = \Gamma\,[\text{si } \mathbb{C}\,(\Gamma) = 1, \text{ else } \Gamma = 0] \approx \Gamma.\,\mathbb{C}(\Gamma)$$

Por ejemplo, $\mathbb{C}$ puede ser que Γ esté comprendida en cierto rango de frecuencias y además supere cierto umbral adaptivo. O bien (lo cual será una regla que utilizaremos en la definición de un sistema práctico y en las aplicaciones):

$\mathbb{C}\,(\Gamma) = 1$ para Γ un miembro de los n máximos locales de valor superior.

Como se deduce, la generación de un patrón en el espacio transformado está justificada, para el reconocimiento visual, si:

a) El nuevo patrón es formal y/o prácticamente más apropiado para tareas de reconocimiento de una clase de imágenes.

b) El patrón generado presenta ya propiedades de invarianza cuya inclusión es deseable en el proceso final de reconocimiento.

Para el caso de las transformaciones lineales en la generación de coeficientes, la única invarianza afín que es estrictamente lograble es la de las traslaciones, por lo que ella sóla no justificaría la generación de un patrón en el espacio de coeficientes, en cuyo caso tendría que darse la situación a) para tal justificación.

4 GENERACIÓN DE ESPACIOS DE RECONOCIMIENTO Y CLASIFICACIÓN

A partir de la generación del patrón Γ' en el espacio de coeficientes, el proceso continúa según el camino usual, con técnicas globales, locales, o mixtas de reconocimiento de formas a varios niveles, es decir, en general, se definen **átomos** descriptores, o alfabetos, basados en propiedades semilocales o globales, se construye una frase o sentencia descriptora de la situación y se procede a una decisión en base a la estructura de tal frase. El proceso de decisión supone la consulta a una base de datos o memoria local del sistema, que ha sido construida durante una etapa de aprendizaje, o directamente inyectada, o ambas a la vez.

El proceso anterior puede realizarse, obviamente, a niveles muy distintos y se admite, usualmente, una jerarquización de los mismos.

A nivel analítico, los pasos para caracterización global (es decir, generación de descriptores analíticos globales) suponen:

a) La realización de una nueva transformación del tipo de las descritas en la sección 2 para la generación de descriptores, tras los cual se calculan funciones descriptoras nuevas que sean invariantes frente a las transformaciones afines.

b) Por reducción del patrón a uno nuevo, que es invariante frente a transformaciones afines, y el cálculo posterior de descriptores por una transformada o por otros métodos.

c) Por realizacion de **medidas invariantes** directas sobre el patrón original.

Sea $\Gamma(x,y)$ el patrón original. Los procesos anteriores son formulables de la siguiente forma:

a) Dado $\Gamma(x,y)$, se obtienen Dij por transformación directa del mismo y de ellos se pasa a nuevos $D'ij = F(Dkl)$, con las condiciones de invarianza afín. Los D'ij han de ser, en número, muy reducidos respecto de los grados de libertad iniciales de $\Gamma(x,y)$. D'ij genera la **frase desciptora** de la situación, y es la base para la toma de decisiones en la clasificacion.

b) Dado $\Gamma(x,y)$, se obtienen los parámetros de invarianza afín,(xo,yo) (centro de gravedad), K (razón de homotecia) y Φ (ángulo standard de giro), tales que se obtiene un nuevo patrón:

$$\Gamma'(x,y) \text{-----}> \Gamma(x',y')$$

$\Gamma(x',y')$ se transforma para obtener descriptores Dij que ya son invariantes afines y con ellos se construye la frase descriptora de la situación, tal como antes. Nótese, que en muchos casos [1], los Díj más convenientes son precisamente los propios $\Gamma(x',y')$.

c) Dado $\Gamma(x,y)$, se realizan medidas sobre el mismo, medidas que pueden ser uni o bidimensionales para los casos que nos ocupan. Combinaciones lineales o no lineales de estas medidas proporcionan nuevas medidas invariantes afines, que forman la frase descriptora. Por ejemplo, para imágenes preprocesadas, umbralizadas y reducidas a dos niveles, unas medidas que son invariantes frente a traslaciones y giros son el perimetro Pe, los perimetros interiores Pi (como conjunto no ordenado), el area A, y las areas interiores Ai (no ordenadas). Las medidas no lineales

$$\frac{P^2}{A^2} \quad \frac{P^2 i}{A^2 i}$$

son, en adición, invariantes frente a homotecias. Asimismo, el vector V [centro de gravedad, punto mas alejado] y otros vectores semejantes, son tal que, las medidas no lineales:

$$\frac{V}{P}$$

son invariantes afines.

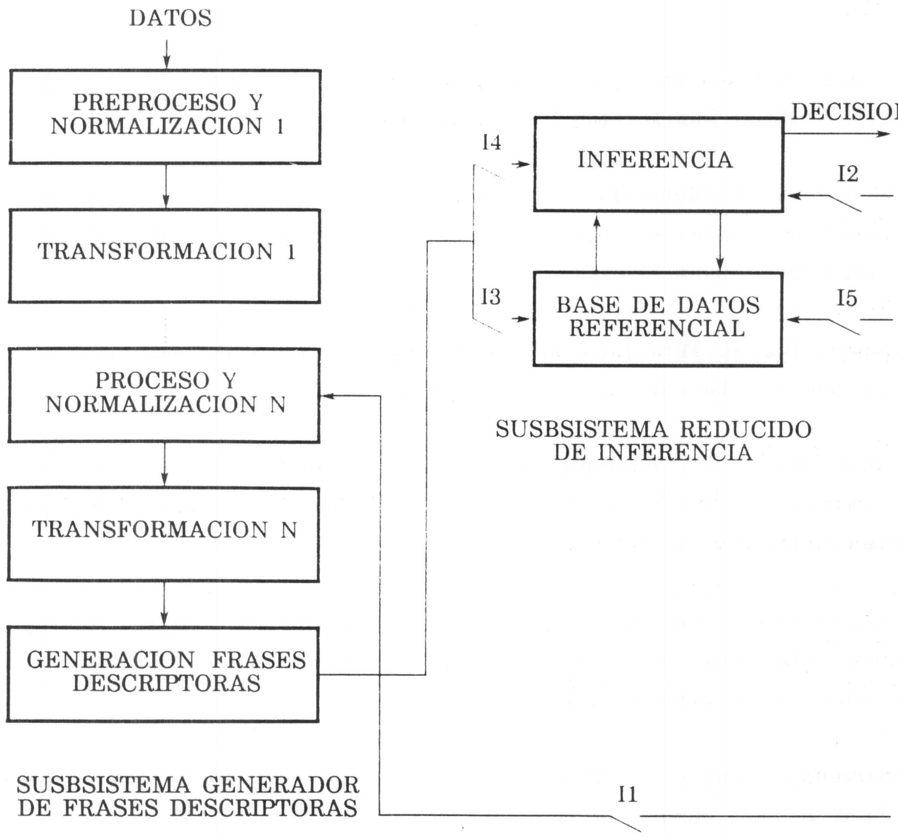
Tales medidas son suficientes, en muchos casos practicos, para formar frases descriptoras completas que resuelven univocamente el problema de decision.

El problema de decision, para las representaciones analiticas, se reduce usualmente al cálculo del error minimo o distancia minima entre la frase descriptora incógnita y un conjunto de frases descriptoras de referencia, que forman el banco de datos. La relación entre las frases descriptoras de referencia y el de las clases de decisión puede no ser biunívoca (que seria la situación teórica deseable) sino de muchos a uno. En esta fase, se procede heuristicamente y el procedimiento de reconocimiento o clasificación se da por válido si la tasa de errores en la asignación de etiquetas -o clase- está por debajo de un valor prefijado.

5 SELECCION DE UNA ESTRUCTURA COMPUTACIONAL

Pasamos ahora a describir, de acuerdo con lo expuesto en [5], la estructura computacional concreta desarrollada para llevar a cabo los objetivos que nos hemos propuesto en el desarrollo del presente trabajo.

La estructura se corresponde con un sistema visual predeterminado como el indicado en [5], que en este caso es tal que el preproceso y la generación de frases descriptoras se realiza en dos fases de naturaleza jerarquizada. Esto se ilustra en la siguiente figura



Sistema Visual Predeterminado

Como puede deducirse de esta figura, el esquema se corresponde con un sistema jerarquizado de procesos de nivel bajo y medio en cascada, diferenciandose dos grandes bloques. El primero de ellos correspondiente al subsistema generador de frases descriptoras, cuya filosofia se basa en la realizacion, en cascada, de una operación de proceso y normalizacion seguida de una transformacion, de forma que a medida que vamos profundizando en ella va aumentado la carga de significado de las posibles frases descriptoras generadas.

6. REFERENCIAS

[1] Candela, S. Transformaciones de Campo Receptivo Variable en Proceso de Imágenes y Vision Artificial. Tesis Doctoral. Universidad Politécnica de Canarias. 1987.

[2] Moreno-Díaz, R., Muñoz-Blanco, J.A. Aspectos analíticos del preproceso de imágenes y su implicación en teoría retinal. Libro homenaje al Profesor Nacere Hayek Calil. pp 473-482. Universidad de La Laguna. 1990.

[3] Moreno-Díaz, R., Mira, J. Un marco teórico para interpretar la función neuronal a altos niveles. Biocibernética. pp 149-171. Ed. Siglo XXI. 1984.

[4] Muñoz-Blanco, J.A. Jerarquización de Estructuras de Nivel Bajo y Medio para Reconocimiento Visual. Aplicaciones a Formas y Texturas. Tesis Doctoral. Universidad Politécnica de Canarias. 1987

[5] Muñoz-Blanco, J.A., Bolivar-Toledo, O., Moreno-Díaz, R. Reduced Inference Systems: Alternative Architectures for Vision. Publicado en Lectures Notes on Computer Sciences. Springer Verlag (en imprenta). 1992.

[6] Santana,o., Mendez, J.A., Moreno-Díaz, R. Momentos normalizados para el proceso de datos visuales. Revista de la Real Academia de Ciencias Exactas Físicas y Matemáticas. Tomo LXXV. Cuaderno 1. pp 287-289.1981.

Recibido: 10 de Diciembre de 1991