

Obtención del Icono y del Contorno de la Imagen por Medio de **Métodos de Topología Digital**

I. Introducción

Sabemos que nuestros ojos están equipados para ver el color de los objetos; pero es menos sabido que también lo están para percibir los límites espaciales entre los objetos, sus contornos. La noción de borde visual designa la frontera entre dos superficies de diferente luminancia, cualquiera que sea la causa de esta diferencia de luminancia (diferentes iluminaciones, diferentes propiedades reflectantes, etc.) para un punto de vista determinado (por ejemplo, entre dos superficies, una de las cuales está tras la otra, hay un borde visual, pero si cambia el punto de vista, el borde ya no estará en el mismo lugar).

El sistema visual está equipado por construcción con instrumentos capaces de reconocer un borde visual y su orientación, una ranura, una línea, un ángulo, un segmento: estos preceptos son como las unidades elementales de nuestra percepción de los objetos y del espacio. El mecanismo extremadamente complejo de esta percepción ha dado buenos resultados, puesto que somos capaces de distinguir bordes visuales de muy pequeña dimensión.

Las propiedades físicas del mundo no dependen de nuestra mirada sobre él. El mundo, a grosso modo, tiene siempre la misma apariencia o, al menos, esperamos encontrar en él, cada día, cierto número de elementos invariantes. La percepción de esos aspectos invariantes del mundo (tamaño de los objetos, formas, orientaciones, propiedades de las superficies, etc.) es lo que se designa mediante la noción de constancia perceptiva: a pesar de la variedad de las percepciones, siempre se encontrarán algunos aspectos constantes.

Estas constantes perceptivas invariantes nos facilitan la vida, permitiéndonos crear esquemas mentales del modelo del mundo. En el presente trabajo se hace referencia a la importancia que tiene el icono y el contorno de la imagen en diferentes áreas. Se presenta la metodología para obtener

tanto el icono representativo de una imagen real, así como el contorno de la misma. Para lograr esto, primero la imagen debe de ser preparada, utilizando para ello los filtros espaciales correspondientes (procesamiento de la imagen por medio de máscaras) que nos permitan obtener la imagen binaria. A partir de la imagen binaria se le aplican los métodos de topología digital, logrando con ello la obtención del icono o del contorno deseado.

II. Sensación y percepción de la imagen

Existe una gran diferencia entre la sensación y la percepción [Matlin y Foley¹], la primera se refiere a experiencias inmediatas y básicas, generadas por estímulos aislados simples, en cambio, la segunda incluye la interpretación de esas sensaciones, dándoles significado y organización. Por ejemplo, cuando leemos un libro los procesos sensoriales y preceptuales demuestran sus extraordinarias habilidades. Los ojos recorren cada página a paso firme, identificando letras y palabras tan rápido que desafían las explicaciones. Así mismo, cada mañana al despertar y levantarnos ejecutamos varias funciones de sensación y de percepción como es: ver, escuchar, tocar, oler y degustar.

Cuando en la radio escuchamos algunas notas musicales, sus características de volumen y tono son sensaciones. En caso de reconocer las primeras notas que forman parte de una tonada, experimentaremos una percepción.

Matlin y Foley también comentan que cuando tenemos frente a nuestros ojos una imagen, cada segmento es frag

mentado por receptores sensoriales en el sistema nervioso en una serie de impulsos. Sin embargo, nuestro sistema nervioso se las ingenia para reconstruir el mundo real de escaleras, pilares y puertas a partir de esta serie de impulsos eléctricos. Nuestras percepciones son organizadas con elegancia y son un espejo razonable del mundo real.

Es por ello, que existen varias razones básicas, en donde una de ellas es filosófica, que es el reto de recrear las cualidades de objetos y poder trasladar éstas del mundo exterior al interior de la mente. Es por esto que el icono y el contorno de una imagen juegan un papel muy importante.

II.a Forma e información

Desde hace algunas décadas las tentativas para superar la teoría de la forma se concentraron en gran parte alrededor de la noción de información [Rudolf Arnheim²], en el sentido técnico que le dieron las teorías surgidas de los célebres trabajos de Hannon y Weaver. La idea directriz es que, en una figura dada, existen partes que proporcionan mucha información y otras que proporcionan poca: estas últimas son las que apenas dicen más de lo que dice ya su vecindad, las enteramente previsibles; acerca de ellas se habla de redundancia.

La noción de información ha permitido reformular los principios guesálticos de modo más general, englobándolos en el principio del mínimo: de dos organizaciones informacionales posibles de una figura dada, la más sencilla es la que será percibida, la que implica más redundancia o, lo que viene a ser lo mismo, aquella cuya descripción moviliza menos información.

Se puede decir, que la imagen es arbitraria, inventada, plenamente cultural, su visión por su parte, es casi inmediata. El estudio cultural de la percepción ha demostrado abundantemente que unos sujetos que nunca han sido expuestos a ella tienen una capacidad innata para percibir los objetos figurados en una imagen y su organización de conjunto, con tal de que se les den los medios de utilizar esta capacidad, explicándoles qué es una imagen.

II.b La relación de la imagen con lo real

Rudolf Arnheim propuso una sugestiva y cómoda tricotomía entre valores de la imagen en su relación con lo real:

a) Un valor de representación: la imagen representativa es la que representa cosas concretas (de un nivel de abstracción inferior al de las imágenes mismas).

b) Un valor de símbolo: la imagen simbólica es la que representa cosas abstractas (de un valor de abstracción inferior al de las imágenes mismas).

c) Un valor simbólico: una imagen puede servir de signo cuando representa un contenido cuyos caracteres no refleja visualmente.

Una de las razones esenciales de que se produzcan imágenes es la que se deriva de la pertenencia de la imagen en general al campo simbólico y que, en consecuencia, la sitúa como mediación entre el espectador y la realidad.

Reconocer algo en una imagen es identificar, al menos parcialmente lo que se ve en ella con algo que se ve o podría ver en la realidad. En general, es un proceso, un trabajo, que utiliza las propiedades del sistema visual, es por ello que el icono y hasta cierto punto el contorno de una imagen son muy relevantes.

El trabajo de reconocimiento: Es un hecho, que un gran número de las características visuales del mundo real se reencuentran, tal cual, en las imágenes y que, hasta cierto punto, se ve en éstas lo mismo que en la realidad: bordes visuales, colores, gradientes de tamaño y de textura, etc. Con mayor amplitud puede decirse que la noción de constancia perceptiva, base de nuestra aprehensión del mundo visual que nos permite atribuir cualidades constantes a los objetos y al espacio, es también el fundamento de nuestra percepción de las imágenes.

J. Aumont menciona que el trabajo de reconocimiento, en la misma medida en que se trata de re-conocer, se apoya en la memoria, más exactamente, en una reserva de formas de objetos y de disposiciones espaciales memorizadas: la constancia perceptiva es la comparación incesante que hacemos entre lo que vemos y lo que ya hemos visto.

La constancia perceptiva es, pues, el resultado de un trabajo psicofísico complejo. Pero esta inestabilidad de reconocimiento llega más lejos aún, puesto que somos capaces no sólo de reconocer, sino de identificar los objetos, a pesar de las distorsiones eventuales que les hace sufrir su reproducción por la imagen. El ejemplo, más llamativo es el del rostro: si reconocemos fácilmente el modelo de un retrato fotográfico (o de un retrato pintado, si está suficientemente

de acuerdo con los códigos naturalistas), es gracias a la constancia perceptiva; pero si reconocemos también al modelo de una caricatura, hay que suponer que hacemos intervenir además, otros criterios (nadie se parece literalmente a su caricatura).

El caricaturista capta, unas invariantes del rostro, que antes no habíamos necesariamente observado [E. H. Gombrich³], pero que podrán en adelante desempeñar el papel de indicadores de reconocimiento (con otro vocabulario se encuentra la misma idea en los primeros trabajos semiológicos de Humberto Eco⁴). Del mismo modo, si encontramos de nuevo a alguien a quien habíamos perdido de vista durante mucho tiempo, lo reconoceremos gracias a invariantes del mismo orden, difíciles a menudo, además, de precisar analíticamente. Son estas invariantes de las imágenes las que se tratan de obtener con el método de topología digital.

Dicho de otro modo: el trabajo del reconocimiento utiliza, en general, no sólo las propiedades elementales del sistema visual, sino también capacidades de codificación ya bastante abstractas: reconocer no es comprobar una similitud punto por punto, es localizar invariantes de la visión, algunas ya estructuradas, como una especie de grandes formas.

Haciendo intervenir su saber previo, el espectador de la imagen suple, pues, lo no representado, disminuyendo así las lagunas de la representación. Esta complementación interviene en todos los niveles, del más elemental al más complejo, siendo el principio de base establecido por Gombrich³ el de que una imagen nunca puede representarlo todo.

Son innumerables los ejemplos de aplicación de este principio de base. Este principio, interviene tanto para permitirnos ver una escena realista en un grabado en blanco y negro (cuya percepción completamos añadiendo al menos todo lo que falta entre los lazos grabados y, a veces, una idea sobre los colores ausentes), como para restituir las partes ausentes y ocultas de los objetos representados (en particular de los personajes).

En el fondo, el espectador puede llegar, en cierta medida, hasta inventar, total o parcialmente, la imagen; Gombrich recuerda, además, que ciertos pintores se sirvieron deliberadamente de esta facultad proyectiva para inventar imágenes, buscándolas en formas aleatorias como unas manchas de

tinta hechas al azar. La imagen es, pues, tanto desde el punto de vista de su autor como de su espectador, un fenómeno ligado, también a la imaginación.

En los humanos, la intuición funciona como un sentido extra, que, más allá de estar subordinados a la razón, puede complementarla e incluso estar por encima de ella.

La facultad de proyección del espectador descansa en la existencia de esquemas perceptivos. Exactamente como en la percepción corriente, la actividad del espectador ante la imagen consiste en utilizar todas las capacidades del sistema visual y, en especial, sus capacidades de organización de la realidad, y en confrontarlas con los datos icónicos precedentemente encontrados y almacenados en la memoria de forma esquemática. Dicho de otro modo, el papel del espectador, en este enfoque, es una combinación constante de reconocimiento y rememoración.

Existen enfoques que ponen sobre todo el acento en los procesos intelectuales que participan en la percepción de la imagen y entrañados por la imagen, y no conceden tanta importancia al estadio puramente perceptivo.

El enfoque cognitivo: La psicología cognitiva es una rama de la psicología cuyo desarrollo ha sido espectacularmente rápido durante el último decenio. Como su nombre lo indica, pretende esclarecer los procesos intelectuales del conocimiento, entendido en un sentido muy amplio, que incluye, por ejemplo, la actividad del lenguaje, pero también, más recientemente, la actividad de fabricación y de consumo de las imágenes.

III. Preprocesamiento de la imagen

III.a Preprocesamiento para obtener la imagen binarizada

Las imágenes con las que se trabajan son del mundo real y también de algunos dibujos sencillos como caricaturas. Cabe mencionar que las imágenes del mundo real son mucho más difíciles de tratar, ya que la variedad de colores es muy grande, por ello se recomienda comenzar con imágenes sencillas.

Fue implementado un software con los algoritmos necesarios para obtener, de las imágenes tratadas las características deseadas, buscando además la eliminación del ruido que regularmente tienen muchas imágenes.

El software puede cargar tanto imágenes a color como imágenes en escala de grises. El método para obtener la escala de grises usado es obtener el promedio de las cantidades de rojo, verde y azul, siendo éste el valor de la escala.

Posteriormente, de acuerdo a la imagen en escala de grises se le puede aplicar alguno de los siguientes filtros [Rafael González⁵].

- Inversión de la imagen: Consiste en cambiar el orden del color de blanco a negro mediante la siguiente función: $f(x)=255 - x$ donde x es la tonalidad original de la imagen y $f(x)$ la nueva tonalidad
- Ecuación automática [Awco G.⁶]: su función es ayudar a distribuir uniformemente en 256 tonalidades diferentes (se trabaja con 8 bits) los niveles de gris originales de la imagen.
- Cambio de contraste: tiene la función de incrementar el rango dinámico de los niveles de gris de la imagen procesada con la finalidad de distinguir más claramente un objeto de otro (o de su contorno). Para este algoritmo se usa una función (en forma de s) que modifica una función en línea recta (las tonalidades originales) para hacer que los niveles más claros lo sean todavía más y los más oscuros lo sean todavía más.
- Cambio de contraste utilizando logaritmos: para este filtro el nuevo valor en escala de grises, $f(x)$ del píxel x , se obtiene usando la siguiente relación:

$$f(x) = \log_{10}(x+1) \cdot 105.96124$$

que permite distribuir el valor de los píxeles logarítmicamente, ayudando a realzar las partes oscuras de la imagen.

- Filtro personalizado: aquí se puede ingresar una matriz misma que usará para hacer que el valor de los valores vecinos intervenga en el valor de un cierto píxel.
- Modificación del histograma: una variante a la ecuación automática, sólo que en este filtro se ingresan dos valores (inferior y superior), y solamente los píxeles que tengan valores entre estos dos valores se distribuyen en las 256 tonalidades diferentes.

- Binarización (o umbralización): se ingresa un valor a partir del cual el valor de un píxel solamente tomará uno de dos valores posibles (negro o blanco).

III.b Preprocesamiento de la imagen binarizada apoyado por conceptos de topología digital

Después de obtener la imagen binarizada según los filtros de la sección III.a existen todavía objetos que no son útiles y que pueden ser eliminados de la imagen tomando en consideración su tamaño, es decir se pueden eliminar los objetos que no cumplan con un tamaño de píxeles establecido.

Los objetos que pueden ser eliminados pueden ser de dos tipos: hoyos dentro de las imágenes y objetos que no cumplan con un cierto tamaño establecido que generalmente el preprocesamiento (sección III.a) no puede eliminar y que la mayoría de las veces son grupos de píxeles de tamaño pequeño.

Para poder eliminar estos grupos de píxeles se utilizan principios de topología digital [Rose Rosenfeld⁷] a continuación se muestran las cuatro definiciones principales usados:

Definición 1. Si a cada entrada de una imagen binaria $M(i,j)$ se le asocia un punto de coordenadas enteras en el plano (puntos lattice). Si $M(i, j) = 0$ entonces el punto asociado será llamado blanco; mientras que si $M(i, j) = 1$, el punto será llamado negro. Con tal regla de asociación se obtiene una imagen binaria formada por puntos lattice (figura 1.b).

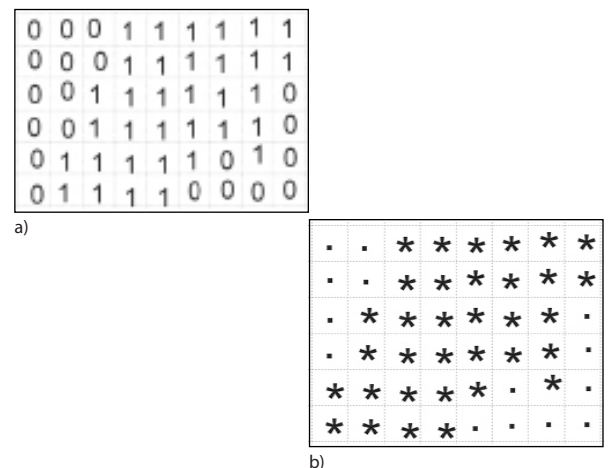


Figura 1. (a) Imagen binaria como matriz de 0's o 1's, (b) Imagen binaria en puntos lattice.

Definición 2. Una imagen digital bidimensional p es una tetrada

$$P = (Z^2, m, n, B)$$

donde

$$B \subseteq Z^2$$

con $(m,n) = (8,4)$ o $(4,8)$. Los puntos de B se llaman puntos negros, mientras que los puntos de Z^2/B se llaman puntos blancos de la imagen, constituyen el fondo de la misma. Cuando B es un conjunto finito la imagen P se denomina finita. En lugar de Z^2 , vamos a considerar un conjunto finito

$$S \subset Z^2$$

el cual contiene a todas las imágenes. En la práctica, S puede considerarse como la colección de todos los puntos de la pantalla de un monitor de tal manera que, nos referimos a la imagen

$$P = (S, m, n, B)$$

Definición 3. Dos puntos lattice

$$p, q \in Z^2$$

se dicen 8-adyacentes si son distintos, y cada coordenada de uno de ellos difiere en a lo más uno, de la respectiva coordenada del otro, es decir, para

$$p = (x_1, x_2)$$

y $q = (y_1, y_2)$ se tiene $|x_i - y_i| \leq 1$ $|x_i - y_i| = 1$, $i=1,2$.

Dos puntos lattice se llaman 4-adyacentes si ellos son 8-adyacentes y difieren en a lo más una de sus coordenadas. Para $n=4$ u 8 , un n-vecino del punto p, es un punto q que es n-adyacente a p (fig. 2).

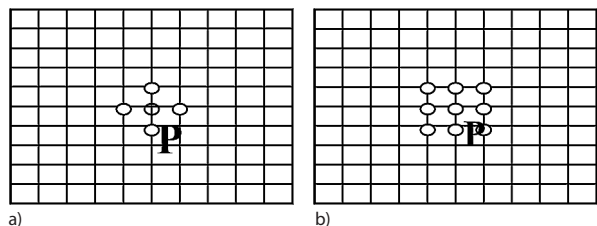


Figura 2. a) p y sus 4 vecinos; b) p y sus 8 vecinos

Definición 4. Dos puntos negros en una imagen digital, $P = (S, m, n, B)$ se denominan adyacentes si

son m-adyacentes, y dos puntos blancos o bien dos puntos blancos y uno negro se dicen adyacentes si ellos son n-adyacentes. Una imagen digital (S, m,n,B) también se llamará una (m,n) imagen digital (fig. 3)

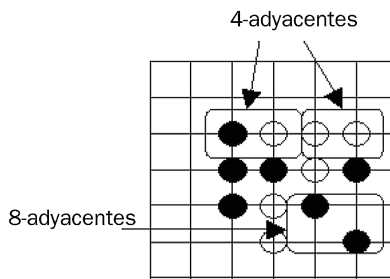


Figura 3. Adyacencia entre puntos

Como se puede apreciar en la figura 3, los puntos blancos son 4-adyacentes y los negros son 8-adyacentes, este principio se usa para el algoritmo de etiquetación de componentes que se describe a continuación:

Algoritmo de etiquetación de componentes [Rafael González⁵]

Hágase un barrido de la imagen píxel a píxel de izquierda a derecha y de arriba hacia abajo.

Sea p el píxel en cada paso del proceso de barrido y sean r, s, t y u los vecinos superior, superior derecho, superior izquierdo e izquierdo, respectivamente. Las etiquetas asignadas son valores mayores que 1, el valor 0 representa un punto blanco y el valor 1 un punto negro.

Paso 1. Buscar un píxel p con valor 1.

Paso 2. Examinar los valores de r, s, t y u, si al menos uno de ellos tiene etiqueta:

Si la etiqueta de todos los vecinos es la misma, asignar a p dicha etiqueta.

Si existe más de una etiqueta diferente entre los vecinos de p, se elige una etiqueta de dichos vecinos y se procede a reemplazar en la imagen todas las coincidencias de dichas etiquetas por la etiqueta seleccionada.

Fin del algoritmo.

Este algoritmo se usa para etiquetar los píxeles que son 8-adyacentes (puntos negros), para etiquetar los píxeles que son 4-adyacentes se usan solamente los vecinos r y u para cumplir con dicha definición.

Después de etiquetar los componentes se procede a realizar la eliminación de los componentes que no cumplen con un tamaño definido por el usuario.

del árbol, el cual lo podemos representar como un icono con rasgos representativos del mismo o bien tan sólo su borde. Esto va a depender del uso que le queramos dar al icono o al borde.

Los filtros espaciales jugaron un papel muy importante en esta metodología. Sin embargo, viendo hacia el futuro, este sistema deberá de ser más autónomo, un sistema basado en inteligencia artificial. Finalmente, a la imagen binaria se le aplican los métodos de topología digital, logrando con ello la obtención del icono o del contorno deseado.

Las imágenes con las que se trabajaron, son imágenes del mundo real. Los objetos que pueden ser eliminados pueden ser de dos tipos: hoyos dentro de las imágenes y objetos que no cumplan con un cierto tamaño establecido que generalmente el preprocesamiento no puede eliminar y que la mayoría de las veces son grupos de píxeles de tamaño pequeño. **T**

Referencias

- [1] MARGARET W. MATLIN Y HUGH J. FOLEY
1996 Sensación y Percepción, Prentice may.
- [2] RUDOLF ARNHEIN
1962 Arte y Percepción Visual, Eudeba, Buenos Aires.
- [3] E. H. GOMBRICH
1959 Arte e ilusión, Paidós.
- [4] HUMBERTO ECO
1999 La estructura ausente: Introducción a la semiótica, Edición Lumen, 5ª. Edición, España.
- [5] GONZÁLEZ C. RAFAEL, WOODS, RICHARD
Tratamiento digital de imágenes, Addison-Wesley/
Díaz de Santos
- [6] AWCOCK G. J.
Applied image processing, Editorial McGraw Hill,
106-108
- [7] ROSENFELD A., KONG T. Y.
1989 Digital Topology: Introduction and Survey, Computer Vision, Graphics and Image Processing, 48, 357-393.
- [8] SHCHEPIN E. V., NEPOMNYASHCHII G. M.
1990 On Topology Analysis of Images, Mezvuzovskii Abornik nauchinh trudo, Moscú, MIR.
- [9] RODRÍGUEZ P. MARIO, BYKOV A., ZERKALOV L.
1999 Index of a point of 3-D digital binary image and algorithm for computing its Euler characteristic, Pattern Recognition, V. 32., N.5, pp. 845-850.

* Rodolfo Martínez

** Moises Ramírez

* Investigador del CIIDIR Oaxaca, IPN

** Profesor-Investigador de la UTM