

IDENTIFICACIÓN DE UNA CARTA DENTAL EMPLEANDO ESTRUCTURAS DE DATOS MULTIDIMENSIONALES

JOHN WILLIAM BRANCH

Escuela de Sistemas, Universidad Nacional de Colombia Sede Medellín, Medellín

LUIS HERNANDO SILVA

Departamento Ingeniería de Sistemas, Universidad de Antioquia, Medellín

ALEJANDRO MONTOYA

Departamento de Ingeniería de Sistemas, Universidad de San Buenaventura Seccional Medellín, Medellín

Recibido para revisión 18 de diciembre de 2000, aceptado 10 septiembre 2001; versión final recibida 24 de Septiembre 2001

RESUMEN: El presente artículo, propone una alternativa de solución al problema del reconocimiento de cadáveres N.N. en nuestro medio, haciendo uso de la carta dental asumida en la legislación colombiana (ley 38 de 1993) y utilizando para ello estructuras multidimensionales y una interfaz gráfica de usuario que le permita a éste, interactuar con el sistema en forma visual.

PALABRAS CLAVES: Estructuras multidimensionales, Odontograma, Rectángulo de Mínima Frontera (MBR), Árbol KD, Métodos de Acceso a Puntos (PAM's), Métodos de Acceso a Rangos (SAM's).

ABSTRACT: This paper shows a solution for identification of corpses without name in our city. We are using dental card in according to colombian law 38 of 1993. The solution uses multidimensional structures and a user graphical component that permits easy man-machine interaction using visual format.

KEY WORDS. Multidimensional Structures, Dental Cards, Minimal Bounding Rectangle (MBR), R Tree, KD Tree, PAM Method (PAM's), Spatial Access Method (SAM's).

1 INTRODUCCIÓN

El presente trabajo pretende proponer un método que permita en forma ágil y confiable reconocer a una persona a través de su carta dental. También es posible que se pueda lograr una mejora a la carta dental promulgada por la ley 38, de tal manera que sea más sencillo el trabajo con esta ley, por parte de todo aquel que tenga la necesidad de usarla, ya sea en el período premorten como en el posmorten.

La masificación del uso de estructuras de datos multidimensionales y sus métodos de acceso en nuestro medio, al igual que realizar nuevas investigaciones y aplicaciones de estas estructuras que permiten el manejo de grandes volúmenes de datos en medias y altas dimensiones en forma eficiente, son prioridades que se consideraron al emprender este proyecto.

La implementación de una interfaz gráfica es uno de los objetivos a desarrollar, con ésta se busca dar al usuario, el odontólogo forense, un medio natural para que interactúe con el sistema que se disponga para su uso.

2 EL PROBLEMA

Una de las consecuencias generadas por las guerras, es la dificultad para el reconocimiento de los muertos que ellas producen y, la existente en Colombia no es ajena a esta problemática, especialmente la que se dio por causa de las acciones del narcotráfico, a finales de los años 80 y principios de los 90, donde los ajusticiamientos no terminaban con la muerte de la víctima, además a los cuerpos sin vida, se les mutilaban sus características físicas, ya fuese por incineración del cuerpo, por amputación de manos o por quemaduras de huellas y rostro. Esto hace imposible la identificación de las personas por el método de la dactiloscopia. En busca de resolver este problema, se trabajó la idea de proponer una ley que permitiese construir un banco de datos con base en las cartas dentales de los ciudadanos colombianos y así lograr a través de estas en una forma eficaz y económica, el reconocimiento de los cadáveres.

Se ha logrado desarrollar un método científico para el reconocimiento de personas, a través de prue-

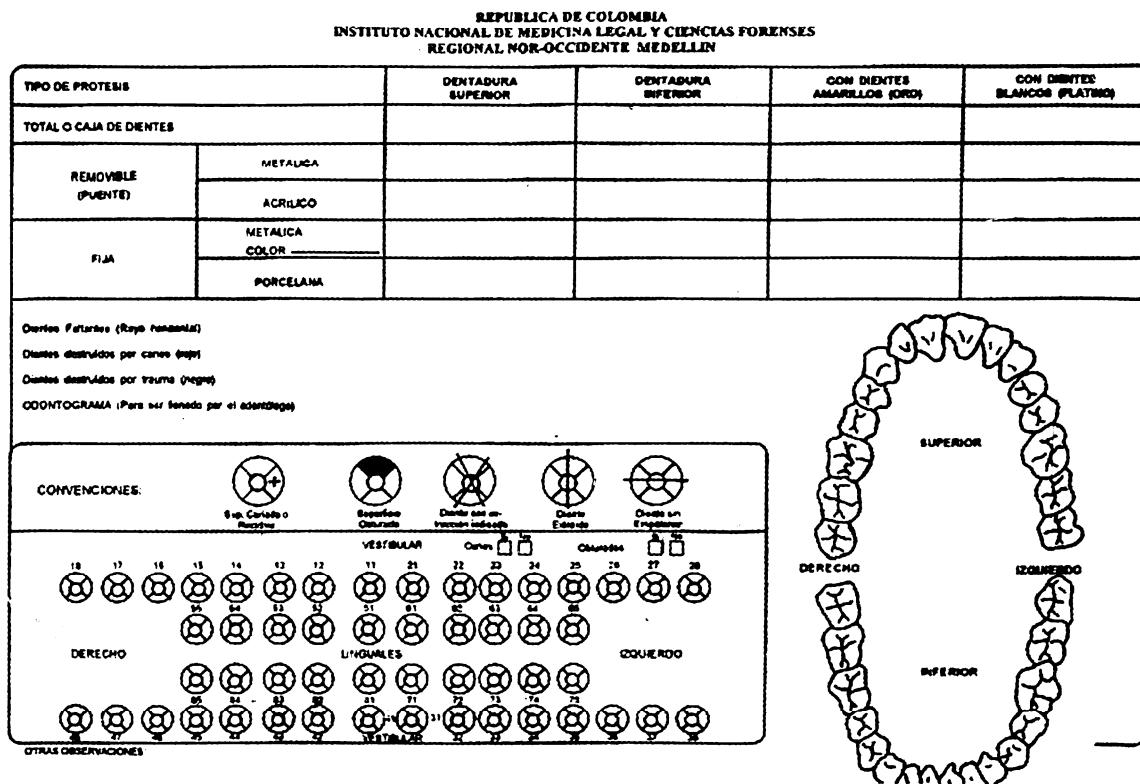


Figura 1. Carta dental

bas de ADN, el cual es un método muy eficaz pero altamente costoso, pues su proceso contempla el hacer múltiples pruebas a todos aquellos que se presume son familiares de sangre con el individuo que se pretende identificar, como también a éste último, lo cual limita su uso a países industrializados.

Estas son las razones para que en enero de 1993 se promulgara la Ley 38, la cual dice que se unifica el sistema de dactiloscopia y se adopta la carta dental con fines de identificación y para lo cual se adopta un modelo de carta dental (ver Figura 1).

Una técnica básica para realizar este reconocimiento consiste en capturar las convenciones asociadas a la carta dental a ser evaluada utilizando una herramienta que se desarrolle para este propósito. Con estas convenciones, convertidas en un conjunto de datos con componentes en varias dimensiones, se pretende proponer un modelo por medio del cual, con base en las características inherentes a los diferentes datos y a través de estructuras de datos multidimensionales, se logre el reconocimiento de la carta dental capturada.

3 GENERALIDADES

3.1 Clases de Dentición

En los humanos hay dos tipos de dentición:

- (i) Una primaria, temporal, decidua o dientes de leche, conformada por veinte dientes, así: diez (10) dientes superiores y diez (10) dientes inferiores, los cuales se distribuyen en cuatro (4) incisivos centrales, cuatro (4) incisivos laterales, cuatro (4) caninos y ocho (8) molares.
- (ii) Una dentición definitiva o permanente, compuesta esta por 16 superiores y 16 inferiores, los cuales se distribuyen en: cuatro incisivos centrales, cuatro incisivos laterales, cuatro caninos, ocho premolares o bicuspides y doce molares.

3.2 Nomenclatura Dentaria

En la odontología hay muchas clases de nomenclaturas, pero para el efecto de este trabajo, se tomará

la nomenclatura internacional. La cual se refiere al nombre y número del diente.

Número: Todo diente se representa por un número de dos dígitos. El primer dígito indica el cuadrante donde está ubicado el diente, izquierdo, derecho, superior o inferior e indica si es temporal (cuadrantes 5 a 8) o permanente (cuadrantes 1 a 4), el segundo dígito indica la posición que ocupa el diente dentro del cuadrante.

Nombre: En la dentición temporal, los nombres de los dientes son los mismos en los diferentes cuadrantes así: Central, Lateral, Canino, Primer Molar, Segundo Molar.

En el caso de la dentición permanente se tiene por cuadrante los siguientes nombres: Incisivo Central, Incisivo Lateral, Canino, Primer Bicúspide (Premolar), Segundo Bicúspide (Premolar), Primer Molar Permanente, Segundo Molar Permanente, Tercer Molar Permanente.

4 ESTRUCTURAS DE DATOS

4.1 Definición de un Objeto de Dato Espacial

Se define como objeto de dato espacial a un objeto compuesto por un conjunto de atributos no espaciales y un atributo de tipo de dato espacial, el cual se encuentra en un espacio euclíadiano cualquiera d -dimensional, su localización es única y está definida por el valor de sus d coordenadas (Ver Figura 2.)

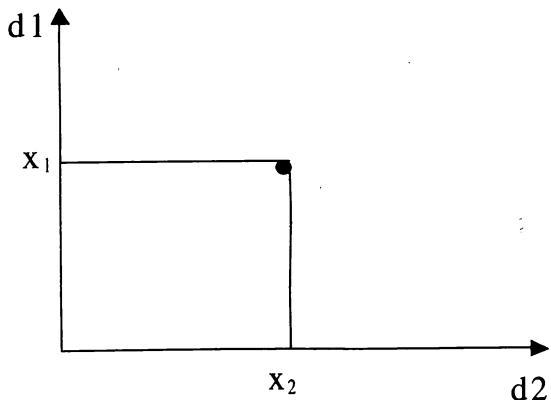


Figura 2. Objeto de Dato Espacial

4.2 Métodos de Acceso Multidimensionales

Son los métodos que soportan las búsquedas en las bases de datos espaciales y se clasifican en Métodos de Acceso a Puntos (PAM's) y Métodos de Acceso a Rangos (SAM's). Los PAM's han sido diseñados principalmente para ejecutar búsquedas sobre bases de datos

puntuales, esto es, bases de datos que solamente almacenan puntos. Los puntos pueden encontrarse en dos o más dimensiones, pero no tienen una extensión espacial.

Los SAM's, se encuentran disponibles para manejar objetos extendidos tales como líneas, polígonos o poliedros en grandes dimensiones.

Dado que un objeto se encuentra en el espacio euclíadiano d -dimensional y su localización es única, se puede construir un índice que permita localizarlo a través de los valores de sus coordenadas.

Por lo general, los índices pueden manejar entradas simples de tamaño aproximadamente igual en una forma eficiente. Esto puede ser logrado por medio de una aproximación del objeto de dato original a una figura simple, tal como un rectángulo o una esfera. Un índice puede administrar solamente el MBR de cada objeto -Minimal Bounding Rectangle o Rectángulo de Mínima Frontera-, esto es, el rectángulo con la más pequeña esquina superior derecha que contiene al objeto, junto con un apuntador a la entrada del objeto en la base de datos.

Donde las búsquedas de atributos múltiples son la regla y la búsqueda de atributos simples son la excepción, existen ventajas en usar un índice de múltiples atributos comparado con varios índices de atributos simples. Primero, el ablocamiento de términos índice y datos en disco puede reducir dramáticamente el número de accesos de I/O necesarios para la búsqueda. Segundo, cuando se insertan nuevos registros, una sola actualización de su índice de atributos múltiples es necesaria, en su defecto, múltiples índices de atributos simples requieren de múltiples actualizaciones.

4.3 Árbol R

4.3.1 Definición

El árbol R, propuesto por Antonin Guttman, es una estructura índice que se utiliza para manejar datos de puntos y datos espaciales al mismo tiempo. En donde los procedimientos de inserción, borrado y búsqueda son intermezclados sin reorganización periódica. Usa una tupla para representar un dato espacial en la base de datos. En orden a recuperar el dato, cada tupla tiene un único identificador, el identificador de tupla. (Ver Figura 3).

En un nodo hoja de un árbol R, se tiene un registro índice que hace referencia al dato espacial. El registro índice está compuesto por I (Identificador de tupla). I es un rectángulo d -dimensional y es la frontera del rectángulo del dato espacial indexado, y cada entrada en un identificador de tupla, corresponde a

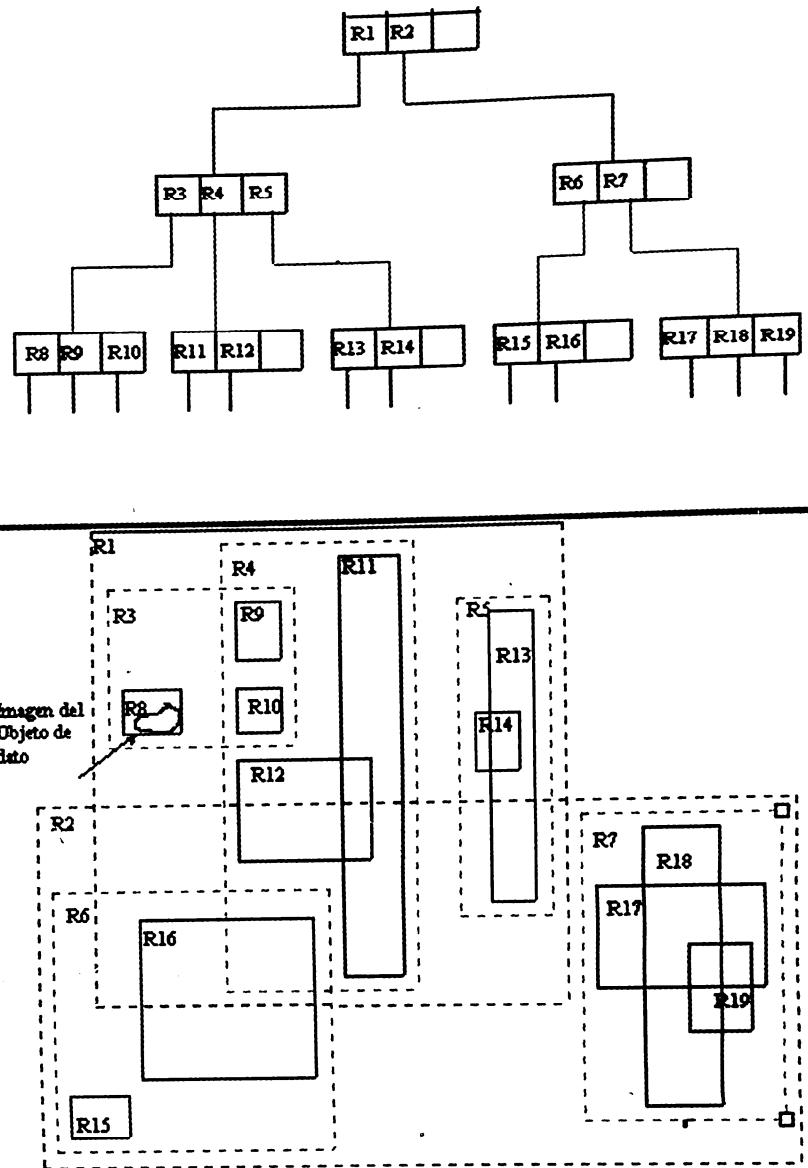


Figura 3. Árbol R

los límites superior e inferior en cada una de las dimensiones del rectángulo. Los nodos corresponden a páginas de disco, así, el índice reside en el disco, y la estructura está diseñada de tal forma que una búsqueda espacial requiera visitar solamente un pequeño número de nodos.

Una ruta hacia abajo en el árbol corresponde a una secuencia de rectángulos anidados, el último de los cuales contiene el objeto de datos.

Los nodos no hojas contienen entradas (I, apuntador a nodo hijo), donde I es el MBR que contiene

a todos los rectángulos en las entradas de los nodos inferiores. El apuntador al nodo hijo es el apuntador a un nodo inferior en el árbol R.

4.3.2 Propiedades

Un árbol R satisface las siguientes propiedades:

- (i) Un árbol R es un árbol de altura balanceada y todas sus hojas están en el mismo nivel.
- (ii) El nodo Raíz tiene al menos dos hijos, a no ser que sea un nodo hoja.

- (iii) Sea M el máximo número de entradas en un nodo completo y sea $m < M/2$, el mínimo número de entradas en un nodo. Cada nodo no hoja contiene entre m y M hijos registros de índice, a menos que sea la raíz.
- (iv) Por cada entrada (I , Apuntador a nodo hijo) en un nodo no hoja, I es el más pequeño rectángulo que espacialmente contiene todos los rectángulos en el nodo hijo.
- (v) Toda hoja contiene entre m y M registros de índice, al menos que sea la raíz.

La altura de un árbol R que contiene N registros de índice es a lo sumo $\lceil \log_m N \rceil$, porque el factor de bifurcación de cada nodo es al menos m .

El máximo número de nodos es $\lceil \frac{N}{m} \rceil + \lceil \frac{N}{m^2} \rceil + \dots + 1$. En el peor caso, la utilización de espacio de todos los nodos, excepto la raíz está entre m y M .

5 ALTERNATIVAS DE SOLUCION

La meta principal de un método de indexado es minimizar el tiempo requerido, en el caso promedio y en el peor de los casos, en las operaciones de consulta (White *et al.*, 1997). Además, estructuras que soporten actualizaciones dinámicas (inserciones y borrados) y que posean una eficiente implementación basada en disco, para así, con base en estas propiedades escalar a bases de datos gigantes.

5.1 El Vector de Características

En el proceso de asignación de índices, los objetos a ser indizados se representan por su vector de características en medias o altas dimensiones.

Dado que para la identificación de los diferentes objetos existentes en la dentadura de un individuo se dividió cada diente (pieza dentaria) en cinco superficies, serán las convenciones asociadas a estas cinco superficies por cada diente las que determinarán el vector de características de la carta dental.

5.2 Las Diferentes Alternativas

Para el desarrollo de la solución se puede, como primera propuesta, tomar los 32 dientes permanentes, en ellos se incluyen los veinte temporales y se diferencian con un atributo en la base de datos, como las dimensiones de la carta dental. Esto hace muy simple el problema, en cuanto a número de dimensiones, pues genera un índice muy pequeño por altura, optimizando el tiempo de búsqueda en el índice, pero lleva a que los procedimientos de búsqueda encuentren un gran número de registros que cumplan las condiciones

de la búsqueda, esto, debido a que muchas de las convenciones se asocian a superficies del diente y no a todo el diente. Por lo tanto, una convención asignada a una superficie se asociaría al diente, lo que implica que varios registros en la base de datos pueden tener asociada esta convención en un mismo diente, aunque en diferente superficie, obligando al procedimiento a acceder a varios bloques de disco y luego a hacer la selección de los que cumplan las condiciones asociadas a las superficies del diente, con base en atributos almacenados en la base de datos.

Como una segunda propuesta, asociar todas las posibles convenciones a dimensiones, así: ya que son 32 dientes permanentes y 20 dientes temporales, esto implica que se tendrán $260 = (32 * 5) + (20 * 5)$ posibles dimensiones para cada carta dental, si se opta por esta propuesta.

La forma como se asocian estas 260 dimensiones es:

Las primeras 160 dimensiones se asocian a los 32 dientes permanentes, asignando la dimensión de acuerdo a la siguiente función:

Número de Dimensión = ((Cuadrante - 1) * 40) + Incremento + Superficie

Incremento = (Número de Diente MÓDULO (Cuadrante * 10 + 1)) * 5

Número de Diente = Número asignado de acuerdo a la convención definida arriba.

Cuadrante = Primer dígito del Diente.

Superficie = Se asigna de acuerdo a la siguiente nomenclatura:

1 = Vestibular

2 = Mesial

3 = Lingual

4 = Distal

5 = Oclusal

Las siguientes 100 dimensiones se asocian a los 20 dientes temporales o deciduos, asignando la dimensión de acuerdo a la siguiente función:

Número de Dimensión = ((Cuadrante - 5) * 25) + Incremento + Superficie

Incremento = $160 + (\text{Número de Diente MÓDULO} (\text{Cuadrante} * 10 + 1)) * 5$

Número de Diente = Número asignado de acuerdo a la convención mostrada arriba.

Cuadrante = Primer dígito del Diente

Superficie = Se asigna de acuerdo a la nomenclatura arriba definida.

Ahora bien, dado que los dientes temporales tienen la misma estructura que veinte de los dientes permanentes y que un individuo no puede tener en un momento el mismo diente temporal y su correspondiente permanente, se puede utilizar únicamente el

vector de características que se obtenga por los dientes permanentes, disminuyendo a 160 características las dimensiones asociadas a la carta dental.

Para saber si el diente es temporal o permanente, se le asigna a éste un atributo en la base de datos que indique esta propiedad; así, al disminuir el número de dimensiones, se minimiza la complejidad de la solución. Para la asignación de las características a las dimensiones, se utilizan 160 características, dadas por los dientes permanentes y sus respectivas cinco superficies ($32 * 5$), de igual forma como se explicó anteriormente, esto es, con la misma función denotada arriba. Serán por lo tanto, 160 las diferentes dimensiones que se deberán tener en cuenta en el manejo de los datos asociados a la carta dental, si se opta por esta última propuesta.

En todas las alternativas, la edad, determinada por la fecha de nacimiento o por el rango de edades que determine el perito, y el sexo del individuo serán atributos asociados a las claves.

6 IMPLEMENTACION DE LAS ALTERNATIVAS

Para el problema en cuestión, los procesos de búsqueda se realizan únicamente con base en las características médica odontológicas asociadas a los dientes. Los datos referentes a la identificación del individuo no tienen una importancia significativa para el caso y por lo tanto no los utilizaremos como claves de búsqueda.

Dada esta premisa, se genera el siguiente problema: si se opta por manejar treinta y dos dimensiones, lo que se presenta es un problema de acceso a claves secundarias con repetición, esto es, a un valor de clave, se pueden asociar múltiples registros en la base de datos.

Cuando se asocian todas las características médica odontológicas a las superficies de los dientes en una carta dental, se conforma una clave primaria, esto es, sus valores asociados identifican inequívocamente a un registro. Ahora bien, si asociamos las características

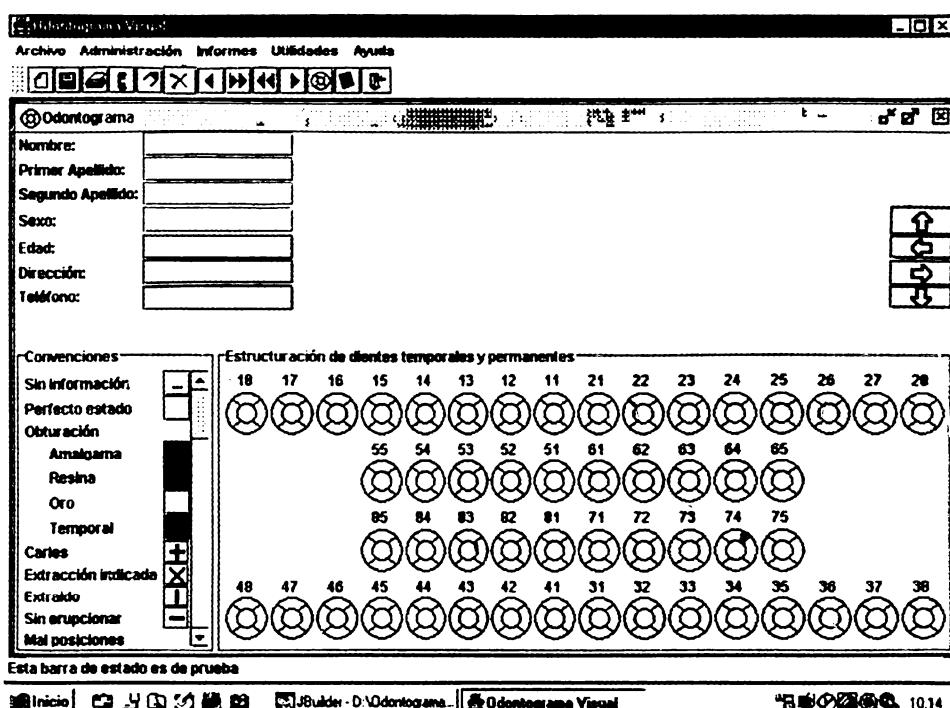


Figura 4. Interfaz Gráfica

Se puede observar que existe una gran diferencia entre las tres alternativas propuestas y que esto conlleva a realizar implementaciones de estructuras con algunos cambios que nos permitan evaluar su desempeño para tomar una decisión veraz en el sentido de optar por la que demuestre mejor desempeño.

médico odontológicas a un diente y no a una superficie, se puede dar el siguiente caso: dos cartas dentales diferentes a las cuales lo único que las diferencia es el hecho de poseer una misma característica médica odontológica en el mismo diente, pero en superficies diferentes, al asignarle dicha característica al diente,

se genera una clave secundaria con repetición, esto es, para un mismo valor de clave, determinada esta por todas las convenciones asociadas, existen dos registros en la base de datos que tienen asociados ese valor de clave, dando como resultado un manejo de claves repetidas.

Lo mismo ocurre si se opta por manejar 160 dimensiones.

Las estructuras de datos multidimensionales creadas hasta ahora, sirven para el manejo de claves, dadas estas por la componente espacial, lo que implica que se manejan solamente claves primarias.

Para resolver el problema que se plantea, se opta por la siguiente propuesta: implementar una estructura de datos multidimensional reformada (árbol R), para el manejo de claves secundarias en múltiples dimensiones. Con esta premisa, a lo que se llega, es a comparar el desempeño de una estructura de datos multidimensional (árbol R) que se implementa para manejar claves secundarias, con dos números de dimensiones diferentes, treinta y dos y ciento sesenta, contra la misma estructura de datos pero que maneja solamente datos para claves primarias, doscientas sesenta dimensiones.

6.1 La Interfaz Gráfica

Con la interfaz gráfica definida para el usuario se puede, en forma sencilla, definir los criterios para la búsqueda de un individuo o grupo de individuos existentes en la base de datos y que cumplan con las condiciones acá previstas - ver Figura 4 -. Se empieza con el despliegue de un odontograma, activando el ícono que identifica a un diente, el cual ubica los dientes en la misma posición en que se encuentran definidas por la carta dental, luego que se despliega éste, se procede a asociar las diferentes convenciones, haciendo *clic* en el respectivo elemento (convención) denotado en la columna de la izquierda y luego *clic* en la superficie del diente a la que se requiere asociar la citada convención.

El sistema le permite al usuario, el odontólogo forense, adicionar y modificar, convenciones sobre cualquier superficie de cualquier diente, los cuales inician con un valor por defecto equivalente a "Perfecto Estado". El sistema controla que dos convenciones incompatibles sobre una superficie, no puedan ser asignadas por el usuario.

Luego que el usuario ha terminado de entrar los datos, genera un archivo, utilizando para ello la opción "Crear Archivo de Salida" en el menú de "Utilidades"; con este archivo de salida, que realmente es una interfaz con los métodos de búsqueda, el programa de con-

sultas realizará el proceso que permitirá al usuario a través de la opción "Cargar Archivo de Entrada" en el menú de "Utilidades" ver los demás datos y las cartas dentales de los individuos que el proceso de búsqueda halló con base en los parámetros entrados por éste.

El sistema tiene incorporado además, funciones como *nuevo*, *abrir*, *imprimir* etc., establecidas como estándar para un programa visual; están disponibles para el usuario a través de un menú definido para ello.

6.2 La Interfaz de Consulta

Para la interfaz que se requiere construir entre el editor gráfico y el método de acceso a los datos, se utilizó como base la función que se definió para determinar la dimensión de acuerdo al diente y a la superficie a la cual se necesita asociar una convención. La función se modificó quedando la interfaz conformada así:

Los primeros 520 bytes equivalen a las 260 posiciones de las superficies de los 52 dientes, y cuya función de direccionamiento es:

1. Si el diente está entre 11 y 48 (dientes permanentes):

Posición en bytes = $((\text{Cuadrante} - 1) * 40) + \text{Incremento} + (\text{Superficie} - 1) * 2$

Incremento = (Número de Diente MÓDULO (Cuadrante * 10 + 1)) * 5

Número de Diente = Número asignado de acuerdo a la convención definida.

Cuadrante = Primer dígito del Diente.

Superficie = Se asigna de acuerdo a la nomenclatura arriba definida.

2. Si el diente está entre 51 y 85 (dientes temporales o deciduos):

Posición en bytes = $((\text{Cuadrante} - 5) * 25) + \text{Incremento} + (\text{Superficie} - 1) * 2$.

Incremento = $320 + (\text{Número de Diente MÓDULO} (\text{Cuadrante} * 10 + 1)) * 5$.

Número de Diente = Número asignado de acuerdo a la convención mostrada arriba.

Cuadrante = Primer dígito del Diente.

Superficie = Se asigna de acuerdo a la nomenclatura ya definida.

A continuación se detallan los bytes utilizados por los otros campos que se asocian al registro de datos:

Los siguientes 20 bytes para el nombre.

Los siguientes 20 bytes para el apellido.

Los siguientes 8 bytes para la fecha de nacimiento.

El siguiente byte para el sexo (1 = masculino, 2 = femenino).

Los siguientes 30 bytes para la dirección.

Los siguientes 20 bytes para la ciudad.

7 CONCLUSIONES

- (i) El uso de estructuras de datos multidimensionales no solo se circumscribe al ámbito geográfico, estas permiten atacar y resolver muchos tipos de problemas, acá, se utilizan como medio de consulta para resolver el problema de la identificación de cadáveres N.N.
- (ii) La técnica del análisis del comportamiento de los métodos de acceso con base en muestras de los datos y evaluando el número de accesos a disco, es muy útil, cuando se requiere tomar decisiones en cuanto a cual estructura de datos utilizar en un caso particular.
- (iii) Al construir algoritmos para comparar el desempeño de los diferentes métodos de acceso, se pueden ver las bondades y defectos que ellos tienen, de tal manera que en trabajos futuros se tengan bases en la toma de decisiones sobre cual de ellos utilizar.

REFERENCIAS

- Chorafas D. N. *Realidad Virtual*, Ed. Prentice Hall. 1995
- Gonzalez R. y Woods, R. *Tratamiento Digital de Imágenes*. Ed. Addison-Wesley/Díaz de Santos. 1992
- Hearn, D. y Baker, M., *Gráficas por Computadora*. Ed. Prentice Hall Hispanoamericana, S.A. 1994

- Kreveld, M., Nievergelt, J., Roos, T. and Widmayer, P., *Algorithmic Foundations of Geographic Information Systems*. Ed. Springer. 1997
- Marr, D., *VISION: And “Computational Investigation into Human Representation and Processing of Visual Information”*. 1982
- Papadias, D., Theodoridis, Y., Stefanakis, E., *Multi-Dimensional Range Query Processing with Spatial Relations*. Geographical Systems, Vol. 4, No. 4, pp. 343-365. URL. www.dbnet.ece.ntua.gr/~stefanak/publications.html
- Rojas, Y., *Implantación de fórmula para un archivo sistematizado de carta dental*. Fiscalía General de la Nación. Dirección Seccional Cuerpo Técnico de Investigación. Sección Criminalística. Medellín. Julio de 1996.
- Theodoridis, Y. and Sellis, T., *A Model for the Prediction of R-tree Performance*. Proceedings of the 15th ACM Symposium on Principles of Database Systems (PODS'96). URL. www.dbnet.ece.ntua.gr/publications/technical-reports.html
- Vànegas, R., *Módulo de Odontología Forense*. Diploma en Ciencias Forenses. Universidad de Antioquia. 1999
- Vélez, A. y Silva, H., *Evaluación del Desempeño de Algunos Métodos de Acceso*. Universidad de Antioquia. 2000
- White, D. and Jain, A., *Similarity Indexing with the SS-tree*. Visual Computing Laboratory. University of California, San Diego.