



**UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA
COORDINACIÓN DE LA CARRERA DE INGENIERÍA
EN CIENCIAS Y SISTEMAS**

**CAVERNA VIRTUAL PARA LA
EXPLORACIÓN DE DATOS EN RED**

NYDIA LISSET CORNEJO LOPEZ

Asesorada por Inga. Elizabeth Domínguez Alvarado

GUATEMALA, AGOSTO DE 2003

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**CAVERNA VIRTUAL PARA LA
EXPLORACIÓN DE DATOS EN RED**

TRABAJO DE GRADUACIÓN

PRESENTADO A JUNTA DIRECTIVA DE LA

FACULTAD DE INGENIERÍA

POR

NYDIA LISSET CORNEJO LÓPEZ

Asesorada por: Inga. Elizabeth Domínguez Alvarado

AL CONFERÍRSELE EL TÍTULO DE

INGENIERA EN CIENCIAS Y SISTEMAS

GUATEMALA, AGOSTO DE 2003

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Sydney Alexander Samuels Milson
VOCAL I	Ing. Murphy Olimpo Paiz Recinos
VOCAL II	Lic. Amahán Sánchez Alvarez
VOCAL III	Ing. Julio David Galicia Celada
VOCAL IV	Br. Kenneth Issur Estrada Ruiz
VOCAL V	Br. Elisa Yazmina Vides Leiva
SECRETARIO	Ing. Pedro Antonio Aguilar Polando

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Sydney Alexander Samuels Milson
EXAMINADOR	Ing. Marlon Antonio Pérez Turk
EXAMINADORA	Inga. Elizabeth Domínguez Alvarado
EXAMINADOR	Ing. Luis Alberto Vettorazzi España
SECRETARIO	Ing. Pedro Antonio Aguilar Polando

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

CAVERNA VIRTUAL PARA LA EXPLORACIÓN DE DATOS EN RED

Tema que me fuera asignado por la coordinación de la carrera de Ingeniería en Ciencias y Sistemas con fecha 23 de febrero de 2002

Nydia Lisset Cornejo Lopez

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	v
GLOSARIO	vii
RESUMEN	xv
OBJETIVOS	xvii
INTRODUCCIÓN	xix
1. AMBIENTES VIRTUALES	
1.1 ¿Qué son los ambientes virtuales?.....	1
1.2 Componentes de los ambientes virtuales.....	2
1.3 Características de los ambientes virtuales.....	3
1.4 Tipo de ambientes virtuales.....	5
1.5 Requerimientos de un ambiente virtual.....	7
1.5.1 Tiempo de respuesta.....	8
1.5.2 Comunicaciones.....	8
1.5.3 Autonomía.....	8
1.5.4 Modelación basado en física.....	9
2. AMBIENTE VIRTUAL AUTOMÁTICO EN CAVERNA	
2.1 ¿Qué es exactamente una caverna virtual?.....	11
2.2 Origen de la caverna virtual.....	14
2.3 Componentes de la caverna virtual.....	14
2.3.1 Computadoras.....	15
2.3.2 Sistemas gráficos.....	16
2.3.3 Sistema de sonido.....	19
2.3.4 Sistema de seguimiento.....	20
2.4 Ventajas de la utilización de la caverna virtual.....	22

3. SISTEMA DE EXPLORACIÓN DE DATOS	
3.1 ¿Qué es un sistema de exploración de datos?.....	23
3.2 <i>Data Warehouse</i>	24
3.2.1 Dimensiones y medidas.....	24
3.2.2 Jerarquía y abstracción de datos.....	25
Proceso del descubrimiento del conocimiento en las bases de datos	26
3.3.1 Fase de proyecto previo.....	27
3.3.2 Fase de integración de datos.....	27
3.3.2.1 Constructor del cubo de datos CCD.....	28
3.3.2.2 Modelo jerárquico de datos de n-dimensiones...	30
3.3.3 <i>Data Mining</i>	31
3.3.3.1 Fundamentos del <i>Data Mining</i>	31
3.3.3.2 Alcance del <i>Data Mining</i>	33
3.4 Procesamiento de transacción en línea (<i>OLPT</i>).....	33
3.5 Procesamiento analítico en línea (<i>OLAP</i>).....	34
3.5.1 Característica de la tecnología <i>OLAP</i>	35
3.5.2 Arquitectura de la tecnología <i>OLAP</i>	36
3.5.3 Operaciones <i>OLAP</i>	36
3.5.3.1 Operaciones <i>OLAP</i> de selección de datos.....	37
3.5.3.2 Operaciones <i>OLAP</i> de agregado de datos.....	38
3.5.3.3 Operaciones <i>OLAP</i> de agregado de datos.....	38
3.6 Métodos para computar el cubo de datos.....	39
3.6.1 Modelo de procesamiento analítico relacional en línea (<i>ROLAP</i>).....	39
3.6.2 Modelo de procesamiento analítico multidimensional en línea (<i>MOLAP</i>).....	42
3.7 <i>Data Warehouse</i> virtual DWV.....	44

4.	INTERACCIÓN EN EL AMBIENTE VIRTUAL	
4.1	Visualización de datos.....	47
4.1.1	Arquitectura de la visualización de datos.....	49
4.1.2	Unidad de control visual (UCV).....	49
4.1.3	Creación de los objetos virtuales.....	54
4.1.4	Administrador de la interfaz del usuario (AIU).....	57
4.1.4.1	Manipulación del punto de vista.....	60
4.1.4.2	Control de la aplicación.....	61
4.2	Interacción con el cubo de datos virtual.....	61
4.2.1	Perspectiva y tamaño.....	61
4.2.2	Orientación.....	62
4.2.3	Interacción continua.....	63
4.2.4	Interacción discreta.....	64
4.2.4.1	Menús flotantes.....	65
4.2.4.2	Movimiento de los cubos de datos en tres Dimensiones.....	66
4.3	Interacción basada en el sistema.....	68
4.4	Interacción basada en los datos agregados.....	69
4.5	Interacción basada en el ambiente.....	70
5.	SISTEMA DE RED DE LA CAVERNA VIRTUAL	
5.1	Capa de comunicación.....	73
5.2	Implementación de las interfases <i>CORBA</i>	75
5.2.1	Capa de comunicación para el servidor.....	75
5.2.2	Capa de comunicación para el cliente.....	77
5.3	Subsistemas de comunicación.....	78
5.4	Protocolo de UCV y DWV.....	79

6. APLICACIÓN DE LA CAVERNA VIRTUAL	
6.1 ¿Qué es <i>Crumbs</i> ?	82
6.1.1 Ventajas de la utilización de la aplicación	82
6.2 Trazado de fibras	83
6.3 Análisis de la distribución de imagen	85
6.4 Navegación a través de la caverna virtual con <i>Crumbs</i>	86
6.4.1 Rotación	87
6.4.2 Análisis de un cuerpo por medio de rebanadas	88
6.5 Aplicación real de <i>Crumbs</i>	89
6.6 Otras aplicaciones de la caverna virtual	91
6.6.1 Arquitectura y construcción	91
6.6.2 <i>Crayoland</i>	92
6.6.3 Meteorología y astronomía	94
CONCLUSIONES	97
RECOMENDACIONES	99
BIBLIOGRAFÍA	101

ÍNDICE DE ILUSTRACIONES

FIGURAS

1	Ambiente normal y ambiente virtual	1
2	Sistema de realidad virtual de escritorio	5
3	Casco estereofónico	6
4	Caverna virtual	12
5	Esquema de la caverna virtual	12
6	Interacción de un usuario dentro del ambiente virtual automático en caverna	13
7	Monitores de la computadora principal	15
8	Proyector de imágenes	17
9	Espejos de la caverna virtual	17
10	Lentes especiales	18
11	Panel de mezclas	20
12	Unidad de control de seguimiento	21
13	Lentes de seguimiento y apuntador tridimensional	22
14	Dimensiones y medidas	25
15	Fases del proceso de descubrimiento del conocimiento en las bases de datos	26
16	A. Concepto de jerarquía. B. Esquema de jerarquía.	30
17	Operaciones OLAP de selección de datos	37
18	El sistema ROLAP utiliza tablas relacionadas para el almacenamiento de los resultados de las funciones GROUP BY	40
19	Arquitectura del sistema ROLAP	41

20	El sistema MOLAP utiliza arreglos contiguos para el almacenamiento de los resultados de las funciones GROUP BY	42
21	Arquitectura del sistema MOLAP	43
22	Componentes de la unidad de control visual	51
23	Arquitectura de la UCV	53
24	Usuarios sumergidos en un sistema de realidad virtual, discuten el cubo de datos.	55
25	Paleta de Color	56
26	Atributos de datos, utilizando cubos y esferas desde un mismo punto de vista. (a) Cubos como agregado. (b) Las flechas señalan algunos de los datos que han sido ocultados.	57
27	Vista a través del cubo de datos	58
28	Esquema del administrador de la interfaz del usuario	59
29	Marco trazado a cuadrícula	62
30	Usuarios posicionan los menús en cualquier punto del ambiente virtual	66
31	Movimiento de objetos entre órbitas	67
32	Interacción basada en datos agregados	70
33	Trazo de fibras	84
34	Plano ortogonal	85
35	Huevo fertilizado de drosophilia	87
36	Planos ortogonales para el corte en rebanadas	88
37	Imagen de un cartílago obstruido	90
38	Usuario visualizando un proyecto de construcción	92
39	Imágenes en dos dimensiones, proyectadas en un ambiente de tres dimensiones	93
40	Usuarios interactuando con el paisaje de <i>Crayoland</i>	94
41	Usuarios visualizando una imagen del espacio	95

GLOSARIO

<i>API</i>	<i>Application Programming Interface</i> . Es un conjunto de funciones que proporcionan al programador la oportunidad de incluir su propio código.
Algoritmo de <i>Data Mining</i>	Es un algoritmo que ingresa un conjunto de datos y que su salida, es un conjunto de observaciones.
Ancho de Banda	El ancho de banda es la máxima cantidad de datos que pueden pasar por un camino de comunicación en un momento dado, normalmente medido en segundos. Cuanto mayor sea el ancho de banda, más datos podrán circular por ella al segundo.
Árbol de decisión	Estructura en forma de árbol que representa un conjunto de decisiones.
Cámara sintética	Es una proyección en perspectiva, la cual acentúa cuando el observador se acerca al objeto.
CCD – Shell	Consola del Constructor del Cubo de Datos.

Cliente / servidor	La arquitectura cliente/servidor es un modelo para el desarrollo de sistemas de información, en el que las transacciones se dividen en procesos independientes que cooperan entre sí para intercambiar información, servicios o recursos. Se denomina cliente al proceso que inicia el diálogo o solicita los recursos y servidor, al proceso que responde a las solicitudes.
Cognoscitivo	Es el conocimiento y la manera de encarar a un objeto.
CORBA API	<i>Common Object Request Broker Architecture - Application Programming Interface</i> . Es una clase desarrollada en C++ que se encargan del manejo del ORB.
Cubo de memoria de imágenes	Es un modelo de datos multidimensional, encargado de almacenar temporalmente un conjunto de imágenes.
Data Mart	Subconjunto de <i>Data Warehouse</i> , en el cual, una parte resumida o altamente enfocada de los datos, se ubica en una base de datos aislada, para darle uso particular.
DBMS	<i>Database Management System</i> . Sistema encargado de la administración de Bases de Datos.

Dominio	Es un rango de información a la cual se tiene acceso.
Emular	Simular.
Espacio de datos	Es la dimensión de un arreglo de tres dimensiones que contiene medidas de datos para que se puedan visualizar.
Espacio virtual	Las dimensiones de una caja, que contiene todas las coordenadas de cada objeto virtual en el mundo virtual.
Estado de máquina	Es un autómata finito determinístico.
Estereoscopia	Se refiere a las técnicas que nos permiten que cada ojo vea una imagen distinta para poder ver el efecto de profundidad, llamado 3d.
<i>Group By</i>	Es una cláusula de <i>SQL</i> , utilizada para separar los registros seleccionados en grupos específicos
<i>Hand Coupled</i>	Es un término que se utiliza para hacer referencia al mapeo directo entre la mano del usuario y el objeto virtual.

Ícono	Un icono es una pequeña imagen, normalmente un símbolo, utilizado para representar gráficamente un programa, un fichero, o una función en la pantalla del ordenador, para facilitar su localización.
Iconografía	Rama de la historia del arte que tiene entre sus tareas determinar el contenido de las obras de arte, los programas de ciclos y la investigación del surgimiento, difusión y transformación de los tipos de representaciones.
IDL	Interfaz que permite invocar de manera transparente las operaciones desde un servicio en una red remota.
Inmersión	Experimentar un intercambio de solicitudes y respuestas dentro de un ambiente de realidad virtual
Interfases MIDI	Interfaz digital de sonido.
Java	Java es un lenguaje de programación por objetos creado por Sun Microsystems, Inc. que permite crear programas que funcionan en cualquier tipo de ordenador y sistema operativo. Se usa el Java para crear programas especiales denominados <u>applets</u> , que pueden ser incorporados en páginas web para hacerlas interactivas.

Join	Es una cláusula de <i>SQL</i> , utilizada para unir los registros seleccionados en una base de datos.
Latencia	Es el tiempo de retardo de un movimiento del usuario, y su respectiva actualización en las pantallas del mundo virtual.
MOLAP	<i>Multidimensional On-line Analytical Processing.</i>
MDDB	<i>Multi-Dimensional Data Base</i>
Normalización	Es un proceso que clasifica relaciones, objetos, formas de relación y demás elementos en grupos, en base a las características que cada uno posee.
Octree	Es una estructura de árbol que indexa sistemas de tres dimensiones. Cada nodo del árbol posee ocho nodos hijos.
OLAP	<i>On-line Analytical Processing.</i>
OpenGL	Es un conjunto de librerías desarrolladas en el lenguaje de programación Visual C, para construir y visualizar imágenes en tres dimensiones a partir de un conjunto de coordenadas.
ORB	<i>Object Request Broker.</i> Clase que permite la comunicación entre los objetos <i>CORBA</i> , y poder responder a requerimientos hechos por clientes a los servidores.

Ordenador	Es un administrador de los recursos del computador.
Ortogonal	Lo que resulta de trazar todas las líneas proyectantes perpendiculares a un plano.
Protocolo	Un protocolo es una serie de reglas que utilizan dos ordenadores para comunicar entre sí.
Redes neuronales artificiales	En una simulación computacional del comportamiento de partes del cerebro humano mediante la réplica en pequeña escala de los patrones que éste desempeña para la formación de resultados a partir de los sucesos percibidos.
Regla de inducción	Es una afirmación general en la cual se fundamenta un número de observaciones particulares.
<i>ROLAP</i>	<i>Relational On-line Analytical Processing.</i>
<i>Scroll</i>	Barra de desplazamiento.
Sensomotriz	Habilidad de capturar ondas sensoriales y de movilidad.
Simulación	Es realizar una actividad en el mundo virtual, muy similar a una actividad de la vida real.

Sistema transparente	Es un sistema cuyas actualizaciones de imágenes no son percibidas por el usuario.
SOAP API	<i>Simple Object Access Protocol - Application Programming Interface</i>
SQL	<i>Structured Query Language</i> o Lenguaje Estructurado de Consulta. Es un lenguaje orientado al manejo de consultas sobre las bases de datos.
Taza de marco	Es el número de veces que el motor de gráficos, es capaz de recargar un marco en el <i>buffer</i> por segundo.
TCP/IP	<i>TCP/IP</i> son las siglas de <i>Transmission Control Protocol/Internet Protocol</i> , el lenguaje que rige todas las comunicaciones entre todos los ordenadores en Internet. <i>TCP/IP</i> es un conjunto de instrucciones que dictan cómo se han de enviar paquetes de información por distintas redes.
Tiempo real de entrega	Es la habilidad de refrescar un marco en el <i>buffer</i> , correspondiente a un evento real.
Todo	Es un mundo reservado, el cual indica el número total de agregaciones de atributos correspondientes a una dimensión en particular.

RESUMEN

Un ambiente virtual es un sistema que se construye por medio de una computadora con el fin de simular un ambiente real en donde el usuario, por medio de interfases; interactúa con él. Los ambientes virtuales cumplen con varias características tal como la forma de interacción, la inmersión y la tridimensionalidad; que básicamente se enfocan en lograr la atención y la manipulación del mundo virtual con objetos en tres dimensiones.

La caverna es un sistema de ambiente virtual, en cuyas paredes se proyectan imágenes en tres dimensiones, con las cuales los usuarios interactúan experimentando la exploración de datos de forma visual.

Dentro de la caverna virtual, el usuario utiliza dispositivos que le permiten interactuar con el mundo que le rodea. Para poder visualizar las imágenes en tres dimensiones y con efectos especiales, el usuario utiliza lentes especiales. La manipulación de los objetos en las imágenes proyectadas, el usuario utiliza un apuntador que a su vez, también proporciona constantemente su posición dentro de la caverna.

La caverna virtual está dotada físicamente de sistemas de proyección visual, sistemas de sonido, sistemas de seguimiento y de procesamiento de datos; los cuales trabajan en conjunto para crear un ambiente muy similar al mundo real.

Tras la manipulación de objetos dentro de la caverna virtual, existe un sistema de exploración de datos, el cual se encarga de la extracción y análisis de información desconocida, almacenada en una base de datos. Todo el

esquema para disponer de la información necesaria para la toma de decisiones la brinda un sistema de *Data Warehouse*.

La caverna virtual utiliza un sistema de visualización, el cual está compuesto por tres módulos: Constructor del cubo de datos, Cubo de memoria de imágenes y la Unidad de control visual; siendo esta última la responsable de generar y manejar el ambiente virtual para la visualización y la exploración de los datos.

Hasta este punto se ha mencionado aspectos de construcción, comunicación y exploración entre los componentes de la caverna virtual. Es importante hacer mención de las múltiples aplicaciones que la caverna virtual puede tener. El campo de la medicina y la biología hacen uso de la aplicación *Crumbs*, que ha sido construida para poder visualizar, analizar y explorar estructuras biológicas por medio de fibras volumétricas.

La construcción y arquitectura también hacen uso de este tipo de aplicaciones, ya que con la caverna virtual es posible construir, visualizar e interactuar con un proyecto de construcción habitacional. Otra aplicación lleva como nombre *Crayoland*, la cual crea un mundo virtual dibujado con crayones, basado en imágenes de paisajes naturales. La caverna virtual también tiene aplicación en las áreas de la meteorología y la astronomía, ya que permite la interacción del usuario con imágenes del espacio exterior y de fenómenos naturales.

OBJETIVOS

General

- Brindar a las personas que están involucradas en el área de la tecnología de la información, un documento que les ayude a conocer y aprender sobre las Cavernas Virtuales para la exploración de datos en red.

Específicos

- Explicar sobre la inmersión de la realidad virtual, y la relación con la exploración de datos.
- Describir de manera comprensible como es la interacción entre el ambiente virtual de las cavernas, con la disposición de la información por medio de la red.
- Explicar los mecanismos y tecnología que este tipo de sistema utiliza para cumplir con su objetivo.

INTRODUCCIÓN

Desde sus inicios la realidad virtual ha sido desarrollada para fines de entretenimiento y espectáculo, tal como las películas de ciencia ficción; en donde los usuarios experimentan una interacción con el mundo virtual que no llega más allá de simular la ubicación en ambientes reales y sus respectivos movimientos.

Hoy la realidad virtual se utiliza no solo para simular ambientes reales sino también para la exploración interactiva y visual de los datos, en donde los usuarios pueden extraer de forma visual, grandes cantidades de información de las bases de datos definidas.

Se presenta un sistema, que utiliza los avances y propiedades de los ambientes de la realidad virtual, todos los requerimientos que un ambiente virtual necesita así como también el proceso de inmersión, utilizando una caverna la cual está dotada con los componentes físicos y tecnológicos necesarios para que por medio de ella, el usuario experimente libremente una realidad transparente de interacción con el medio.

El proceso de inmersión involucra la exploración de datos, lo cual permite realizar una navegación dentro de la información, para que tan solo con un movimiento de la mano, se pueda seleccionar el tema de interés. La exploración de datos utiliza un proceso de descubrimiento del conocimiento en las bases de datos, el cual permite obtener los datos solicitados. El modelo de bases de datos es jerárquico de n-dimensiones, llamado también cubo de datos, esta propiedad contribuye a que la información se encuentre en distintos niveles de abstracción.

El sistema de exploración de datos en un ambiente virtual, consiste en tres módulos que permiten administrar cada elemento importante en el sistema. El primer módulo se encarga de la administración de la interfaz con el usuario, preocupándose de que dicho usuario tenga una manipulación correcta y una interacción real con una visualización totalmente clara. El segundo módulo enfatiza su trabajo en la interacción con el cubo de datos virtual, y el tercer módulo se encarga de la construcción del cubo de datos, es decir, de toda la administración de las bases de datos que almacena la información.

Es muy importante mencionar el esquema de la arquitectura principal del sistema, ya que es esta parte en donde se integran tanto los componentes remotos, como los componentes locales, con el objetivo de proporcionar un ambiente de inmersión transparente en cuanto al análisis y la exploración.

Los profesionales de la medicina y la biología, se auxilian de este tipo de sistemas para analizar interactiva y visualmente determinada deficiencia física que presenta su paciente. Con el sistema de la caverna virtual, ellos experimentan el análisis visual, más cercano a la realidad, que incrementa la certeza del diagnóstico.

1. AMBIENTES VIRTUALES

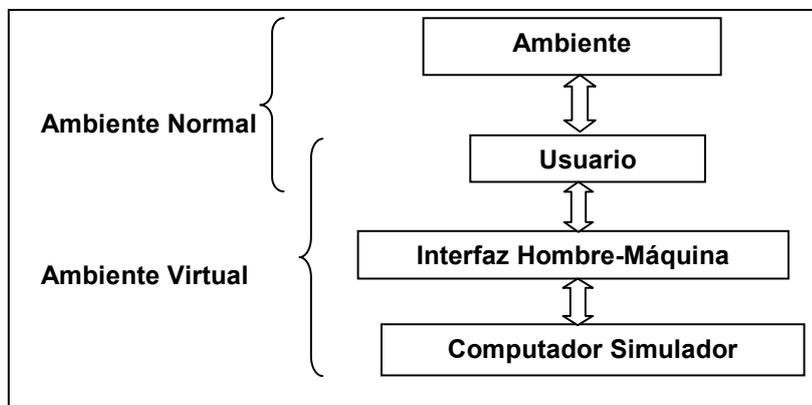
En este capítulo se presentan las características y requerimientos que deben poseer los ambientes, para brindar un ambiente de inmersión virtual muy similar al mundo real.

Existen distintos tipos de ambientes virtuales, dependiendo del tipo de aplicación que se desea proyectar. Muchas veces es necesario utilizar diversos dispositivos para realizar una inmersión en el ambiente virtual.

¿Qué son los ambientes virtuales?

Los ambientes virtuales son sistemas que permiten la interacción entre el usuario y un ambiente construido por una computadora, mediante una interfaz. El objetivo del uso de la interfaz, es reemplazar la percepción del mundo real del usuario por la de un mundo virtual, es decir, hacer creer al usuario que se encuentra en un lugar distinto al físico.

Figura 1. Ambiente normal y ambiente virtual



Componentes de los ambientes virtuales

Los ambientes virtuales tienen cuatro componentes básicos los cuales son:

- **Procesador de entrada.** Tiene como función principal controlar las fuentes de datos del sistema, tal como los dispositivos o periféricos que sirven de interfaz con el usuario. Pero no solo los dispositivos y periféricos sirven como fuentes de datos, ya que en los sistemas virtuales multiusuario se recibe continuamente información de la red. La interfaz con el usuario puede ser por medio de comandos de voz o de texto. O bien, ejecutando una manipulación directa por medio de un localizador. El propósito de la interfaz es apoyar la actividad del usuario para que las tareas se realicen con menos errores, mayor calidad y mejor rendimiento. Es necesario que la interfaz sea lo más natural posible para que el usuario se enfoque en la interacción con el ambiente y no en aprender a manipularla.
- **Procesador de simulación.** Este procesamiento constituye el núcleo del sistema virtual ya que es el encargado de analizar el entorno a partir de datos de entrada, tal como la posición del usuario y el estado de las tareas, de allí determina las nuevas acciones a tomar.

- **Generación del ambiente virtual.** Se encarga de generar la salida que percibe del usuario, grafica las tareas o acciones que puede realizar un objeto, tomando en consideración la posición del ambiente actual. La generación del ambiente virtual puede ser dividida en varios subprocesos dependiendo del grado de complejidad y diversidad de los dispositivos de salida, cuyas entradas son las salidas del procesador de simulación y los datos de la base de datos correspondiente.

La generación del ambiente virtual puede dividirse en dos tipos de sistemas, el primero en el que las leyes de la física no se obedecen o se alteran, y el segundo en donde se simula un ambiente real siguiendo las leyes de la física.

Características de los ambientes virtuales

Para que los ambientes virtuales puedan cumplir con su objetivo, es necesario que cumplan con las siguientes características:

- **Interacción.** Como se mencionó anteriormente, la interacción permite que el usuario manipule el curso de la acción dentro del ambiente virtual, que el sistema responda a los estímulos del usuario que lo utiliza y que se cree una interdependencia entre ellos.

Existen dos puntos importantes en la interacción con el ambiente virtual. El primero de ellos tiene que ver con la *navegación*, que es la habilidad del usuario para moverse independientemente dentro del ambiente virtual, y el *posicionamiento del punto de vista físico del usuario*, ya que él puede verse a sí mismo a través de los ojos de alguien más o puede moverse dentro del ambiente virtual. Las restricciones de la navegación son establecidas por el inventor del software, el cual permite varios grados de libertad, como por ejemplo el poder volar, caminar, nadar, oler, etc.

El segundo punto importante es la dinámica del ambiente, que constituye las reglas de cómo los componentes del ambiente virtual interactúan con el usuario para intercambiar energía o información.

- **Inmersión.** Es involucrarse y enfocarse completamente en el ambiente donde se trabaja. La inmersión posee dos características importantes, la primera es la *habilidad para enfocar la atención del usuario* y la segunda es *convertir la base de datos en experiencias personales*, que contribuye con el sistema natural del aprendizaje humano.

- **Tridimensionalidad.** Es una de las características más importantes para todo sistema de realidad virtual, ya que es el encargado de la manipulación de los sentidos del usuario para dar forma al ambiente virtual; principalmente al sentido de la vista. Todos los objetos que forman parte del ambiente virtual se muestran al usuario en las tres dimensiones del mundo real, en el sentido del espacio que ocupan y con efectos del sonido necesarios para crear un ambiente lo más real posible.

Tipos de ambientes virtuales

De acuerdo a la forma de interacción entre el usuario y el sistema virtual, los ambientes virtuales se clasifican en:

- **Realidad virtual de escritorio.** Es el tipo de realidad virtual en el que el sistema se grafica en el monitor de la computadora o en un sistema de proyección. Tiene un costo bajo y excelente resolución.

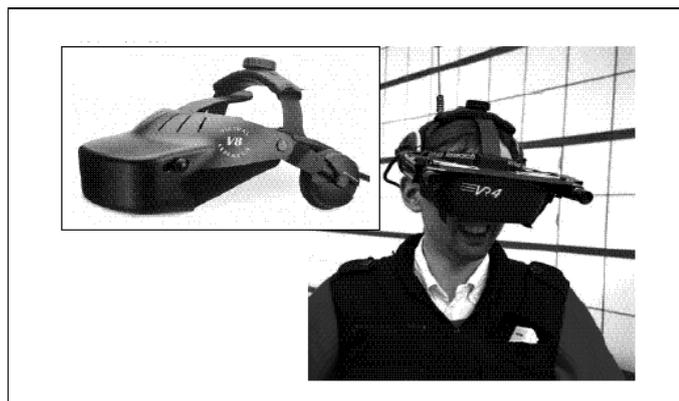
Figura 2. Sistema de realidad virtual de escritorio



Fuente: http://mailweb.udlap.mx/~tesis/vonraesfeld_p_pa/capitulo1.html

- **Realidad proyectada.** En este tipo de realidad virtual, una imagen del usuario en movimiento y del ambiente virtual se proyectan en una pantalla. El usuario puede verse a sí mismo, como si realmente estuviese en la escena, es decir, él podrá observar la interacción de su cuerpo con el mundo de manera real, ya que el sistema se comportará y reaccionará en tiempo real.
- **Sistemas de inmersión.** El sistema de inmersión se basa en una coordinación entre los movimientos del usuario por medio de la utilización de un casco estereofónico, dotado de sensores de posición y de orientación; y de imágenes proyectadas. El casco informa continuamente la posición del usuario y hacia dónde está observando el usuario. En ocasiones dicho casco puede estar conectado a un traje de sensores especial u otra clase de dispositivo para lograr más realismo en el ambiente.

Figura 3. Casco estereofónico



Fuente: http://mailweb.udlap.mx/~tesis/vonraesfeld_p_pa/capitulo1.html

- **Tele-presencia.** Significa presencia remota, ya que proporciona a la persona la sensación de estar físicamente en la escena creada por la computadora. Esta tecnología de simulación es tan avanzada que convence al usuario de estar en un ambiente de realidad virtual. Estos sistemas administran sensores remotos que están instalados en el mundo real, y que adicionalmente se unen con los sentidos de un operador humano.

Requerimientos de un ambiente virtual

Un ambiente virtual presenta varios requerimientos, debido a que posee características de interacción, múltiples usuarios y dinamismo. Cuando se habla de múltiples usuarios, se refiere al continuo intercambio de audio, video y gráficas que se emiten en tres dimensiones y en tiempo real. Hoy en día tanto la interacción y el dinamismo establecen restricciones de ancho de banda, latencia, confiabilidad y distribución; es por ello que se deben cubrir con los requerimientos de tiempo de respuesta, comunicaciones, autonomía y modelación basado en física.

Tiempo de respuesta

En un ambiente virtual continuamente se producen retardos entre la transmisión de las acciones del usuario y la retroalimentación; todo ello debido a la distribución y, al hardware y software de red. Si estos retardos son muy grandes, el usuario no conoce el estado real del sistema en un momento específico, no conoce el resultado real de las acciones que realiza y no puede tomar decisiones efectivamente. Es necesario entonces, controlar el tiempo de procesamiento de los comandos, de la visualización del mundo y de la transmisión de mensajes para minimizar el tiempo de respuesta.

Comunicaciones

Debido a la importancia que tiene la transmisión de los estados del ambiente virtual a los usuarios que están conectados mediante una red, es necesario tomar muy en cuenta el software de comunicación a utilizar, enfocarse en un protocolo estándar de mensajes y en un mecanismo para asegurar que la percepción del ambiente virtual sea igual para todos los usuarios. También es necesario minimizar la transmisión de datos y la manera de responder a los datos que se pierden para evitar sobrecargar la red con información irrelevante.

Autonomía

En el ambiente virtual no solo existen usuarios reales, sino también agentes virtuales autónomos, los cuales dependen del contexto del ambiente y que además, se puede interactuar con ellos.

Es necesario un mecanismo que permita grabar la interacción entre el usuario y los agentes, para que se pueda tener retroalimentación y análisis de las acciones hechas y procurar así un sistema autónomo e inteligente.

Modelación basada en física

La modelación basada en física permite que el usuario explore un ambiente virtual en el que los objetos se comportan tal como los objetos de un ambiente real. Esto es de mucha importancia en aplicaciones cuyo objetivo es obtener conocimiento del ambiente real, como por ejemplo, aplicaciones de entrenamiento, simulación, diseño y tele-operación. Existen dos requerimientos mínimos de este tipo de modelos, la detección de colisiones y el terreno dinámico.

- **Detección de colisiones.** Aumenta la credibilidad del ambiente virtual, debido a que los usuarios no tendrían una visión real si se les permitiera atravesar los objetos del ambiente. Este factor es de mucha importancia para simulaciones de planeación y entrenamiento.
- **Terreno dinámico.** Tiene una estrecha relación con los efectos sobre el ambiente. Si un usuario realiza una acción de transformación sobre un objeto, dicha transformación se debe representar como tal en el ambiente virtual, para que la misma pueda ser actualizada y que los usuarios puedan ver el producto de esa interacción en tiempo real.

2. AMBIENTE VIRTUAL AUTOMÁTICO EN CAVERNA

El ambiente virtual automático en caverna o caverna virtual, es un sistema sofisticado de realidad virtual; el cual permite que el usuario experimente la sensación real de estar presente en el ambiente proyectado, por medio de la inmersión.

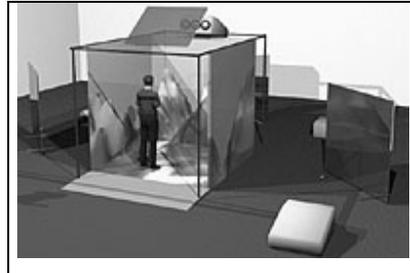
Se dice que es *virtual* debido a que las imágenes son proyectadas por computadora e *inmersión*, debido a que el sistema proporciona al usuario un sentido de presencia.

Para causar esa sensación real de presencia, se utilizan varios dispositivos que no solamente abarcan el área visual, sino también el sonido, la administración y el procesamiento de datos.

2.1 ¿Qué es la exactamente una caverna virtual?

Una caverna virtual es una habitación pequeña, formada por tres paredes y el piso; con dimensiones aproximadamente de diez pies de ancho por diez pies de largo en cada lado. Está compuesta por varios proyectores de alta resolución que procesan imágenes con alta velocidad y que se proyectan en sus paredes para producir vistas estereoscópicas. También existen otros dispositivos que emiten sonidos de acuerdo a la imagen proyectada para causar una sensación real.

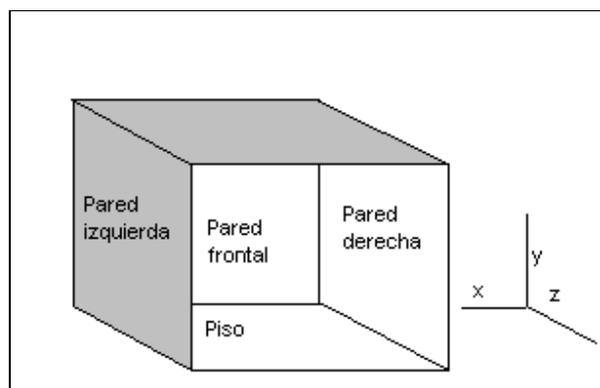
Figura 4. Caverna virtual



Fuente: http://www.indiana.edu/~ids/archives/tech_cave.html

Las imágenes son vistas por el usuario por medio de lentes livianos especiales que se encargan de causar el efecto de tres dimensiones y que controlan continuamente su orientación física dentro de la caverna, interpretándolas tanto en coordenadas cartesianas en los ejes x, y y z; como en coordenadas polares. El procesamiento de la captación de información del dispositivo, se realiza con una velocidad alta y se almacena en estructuras de datos que alimentan el motor de gráficos una vez por ciclo.

Figura 5. Esquema de la caverna virtual



Los usuarios de la caverna virtual también utilizan un dispositivo muy parecido al mouse denominado apuntador tridimensional, que es sensible al movimiento de la mano del usuario y se utiliza para varias funciones, tal como el de proveer de un apuntador que guíe dentro del ambiente virtual y manipular los objetos del ambiente.

Figura 6. Interacción de un usuario dentro del ambiente virtual automático en caverna



Fuente: http://www2.gratisweb.com/lhorena_ortiz/pag-virtual.htm

2.2 Origen de la caverna virtual

Platón realizó una serie de estudios sobre la realidad y percepción humana, por medio de una analogía en la cual un hombre define el fundamento de su realidad, por medio de la danza de sombras de fuego sobre las paredes de su caverna. Tomando como base esa analogía, el Doctor Thomas DeFanti y Dan Spandin tuvieron la inquietud de construir el primer prototipo de la caverna virtual, que fue llevada a cabo en 1992, en el Laboratorio de Visualización Electrónica de la Universidad de Illinois en Chicago, Estados Unidos. El segundo prototipo se construyó en el Instituto Beckman para la Tecnología Avanzada por el Centro Nacional para aplicaciones de computación NCSA, con el objetivo de desarrollar campos relacionados con la realidad virtual y visualización científica.

2.3 Componentes de la caverna virtual

La caverna virtual es una colección de diferentes sistemas complejos que trabajan en conjunto. Si alguna pieza llega a fallar, puede afectar el buen funcionamiento del sistema completo.

2.3.1 Computadoras

- **Computadora principal.** La función de la computadora principal es coordinar todo lo que sucede en la caverna, verificar que todos los componentes del sistema trabajen correctamente y controlar el acceso de los usuarios al sistema. La ubicación física de la computadora principal no es la misma a la ubicación de la caverna, debido a la seguridad; únicamente sus dos monitores, el teclado y el mouse se encuentran junto a la caverna.

Figura 7. Monitores de la computadora principal



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 5

- **Computadora de soporte.** La función de la computadora principal es coordinar todo lo que sucede en la caverna, verificar que todos los componentes del sistema trabajen correctamente y controlar el acceso de los usuarios al sistema.

2.3.2 Sistema de gráficos

La caverna virtual es principalmente un sistema gráfico, que por lo tanto requiere de hardware especial. Dicho sistema está dividido en dos partes; las imágenes proyectadas en las paredes de la caverna y los lentes de seguimiento que se encargan de hacer que las imágenes se observen en tres dimensiones.

- **Proyectores.** Aproximadamente sesenta imágenes por segundo deben ser proyectadas en cada una de las tres paredes y el piso de la caverna; las cuales son desplegadas por cuatro proyectores de alta resolución, que están ubicados detrás de cada una de las tres paredes de la caverna. Para proyectar las imágenes del piso, existe un proyector suspendido en el cielo de la caverna.

Figura 8. Proyector de imágenes



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 8

- **Espejos.** Los proyectores necesitan estar a cierta distancia de las paredes de la caverna, pero muchas veces el área donde se encuentra la caverna es muy reducida, por lo que se utilizan espejos, para que las imágenes se proyecten hacia ellos, y las transmitan a las paredes.

Figura 9. Espejos de la caverna virtual



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 9

- **Lentes especiales.** Como ya se anotó, los lentes, son dispositivos que se encargan entre otras cosas, de hacer que las imágenes proyectadas en las paredes de la caverna, se vean con el efecto de tres dimensiones.

En la caverna virtual se proyectan dos imágenes diferentes, una para el ojo izquierdo y la otra para el ojo derecho. Los lentes se encargan de bloquear la parte del ojo derecho para que el usuario no pueda ver nada, mientras se despliegan las imágenes de la parte del ojo izquierdo, y vise-versa.

Con los lentes, el usuario puede ver un objeto desde diversos ángulos que él desee, como si lo estuviera viendo en la vida real.

Figura 10. Lentes especiales



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 10

- **Emisores estéreo.** Los emisores estéreo son pequeños dispositivos en forma de cajas, localizadas en las esquinas de la caverna y encargadas de sincronizar los lentes del usuario con las actualizaciones de las pantallas.

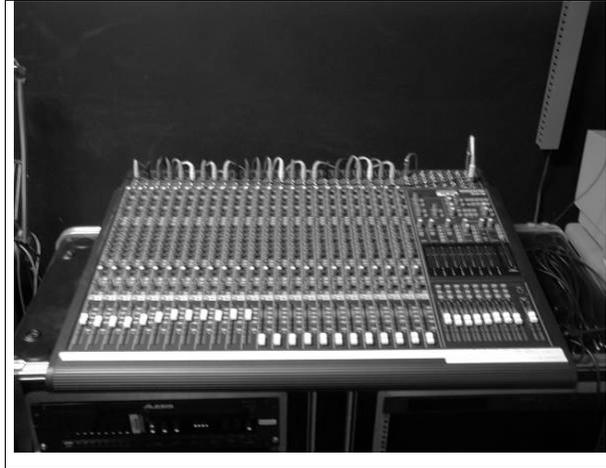
2.3.3 Sistema de sonido

Además de la gran capacidad de gráficos que la caverna virtual posee, también cuenta con un sistema sofisticado de sonido; el cual a su vez, posee interfaces MIDI, bocinas, estaciones de trabajo INDY, sintetizadores y un panel de mezclas.

Las estaciones de trabajo INDY, son servidores de sonido que trabajan internamente. Todos los dispositivos y comandos de sonido son administrados por los sintetizadores. Tanto la interfase MIDI como el sintetizador están colocados a la par de la caverna.

- **Panel de mezclas.** El panel de mezclas controla un sistema de audio de ocho canales con sonido digital.

Figura 11. Panel de mezclas



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 11

2.3.4 Sistema de seguimiento

El sistema de seguimiento de la caverna virtual, es extremadamente compleja y se compone de varios dispositivos que trabajan juntos para monitorear la localización y orientación de los lentes y el apuntador tridimensional.

- **Unidad de control de seguimiento.** La unidad de control de seguimiento es el dispositivo más importante en este sistema, ya que es el que se encarga del monitoreo de la ubicación del usuario.

Figura 12. Unidad de control de seguimiento



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 12

- **Lentes de seguimiento.** Esta clase de lentes son muy parecidos a los lentes especiales, con la diferencia de que tienen un cable adherido, para poder monitorear su localización.
- **Apuntador tridimensional.** El apuntador tridimensional, es la interfase primaria entre el usuario y la caverna. Se utiliza para manipular los objetos e interactuar con el mundo proyectado. Permite movilizarse dentro del área de la caverna.

Figura 13. Lentes de seguimiento y apuntador tridimensional



Fuente: Kalev Leetaru. Easy Demo Manual. Página No. 13

2.4 Ventajas de la utilización de una caverna virtual

En comparación con otros sistemas virtuales, el ambiente virtual automático en caverna posee varias ventajas sobre los demás, las cuales se citan a continuación:

- Los usuarios tienen libertad para movilizarse.
- Se tiene la oportunidad de aumentar el realismo del ambiente debido a las proyecciones en las pantallas derecha e izquierda de la caverna.
- El sistema provee comodidad e higiene, ya que el usuario no necesita de trajes ni dispositivos complejos para interactuar con el ambiente.

3. SISTEMA DE EXPLORACIÓN DE DATOS

En esta parte del trabajo se presenta el sistema de exploración de datos, como parte fundamental de la administración de la información en el ambiente virtual; la estructura y la tecnología que se utiliza para la obtención y posteriormente, el análisis, de la información que se requiera en la interacción con el usuario.

3.1 ¿Qué es un sistema de exploración de datos?

El sistema de exploración de datos es un entorno el cual está dedicado a la extracción y análisis de información desconocida, almacenada en una base de datos, que representa un alto grado de importancia en la toma de decisiones. La provisión de visualizaciones, o representaciones gráficas de conjuntos de datos, es de gran utilidad para facilitar el proceso de comprensión de la información contenida en dichos datos. Con el desarrollo de los sistemas de visualización por medio de ordenadores, las capacidades provistas por las técnicas de visualización convencionales han tenido un gran desarrollo, principalmente a través de la posibilidad de la interacción directa del usuario con la información visualizada.

3.2 Data Warehouse

El *Data Warehouse* es una recopilación de datos de varias bases de datos operacionales, que integra la información en un modelo lógico, agrupada por tópicos comunes y almacenada de tal forma que resulta accesible a los tomadores de decisiones. El *Data Warehouse*, no facilita las operaciones de una organización, pero si provee de predicciones, anomalías o evidencias sobre las cuales pueden ser el punto de partida para la toma de decisiones.

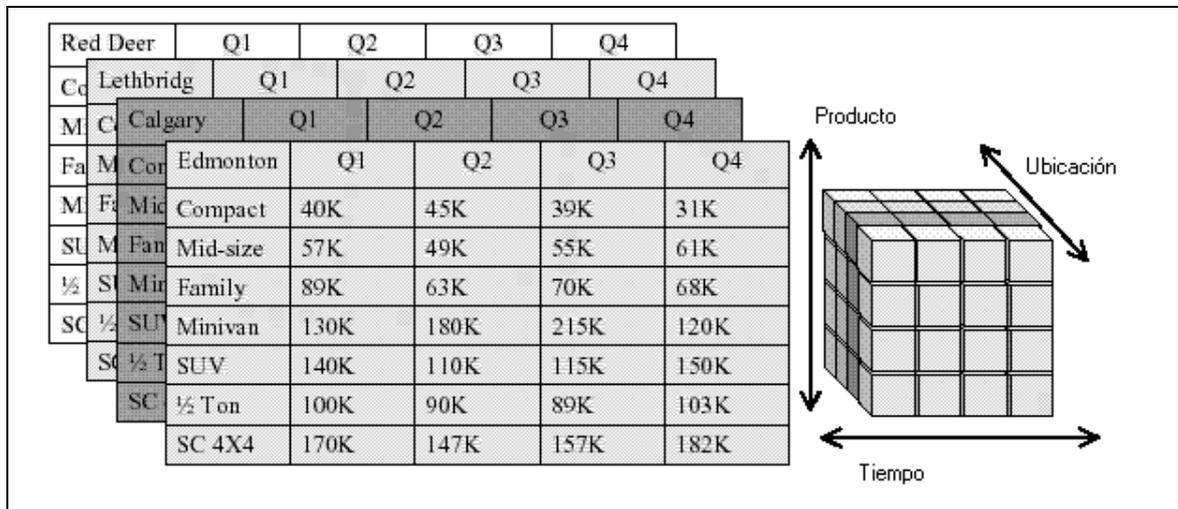
La construcción de un *Data Warehouse* conlleva a la transformación de los modelos de datos tradicionales, en modelos de datos multidimensionales con diferentes niveles de abstracción.

3.2.1 Dimensiones y medidas

El principio fundamental en el diseño de un *Data Warehouse*, es la identificación de las dimensiones de los datos. La información se presenta en una matriz de dos dimensiones de n columnas por m filas. Usualmente los atributos que se encuentran en la primera columna y primera fila representan los dominios de las dimensiones de los datos. Al añadir una nueva matriz que está relacionada con respecto a otro concepto, se origina una tercera dimensión.

En cada una de las celdas de cada matriz, se encuentra los valores respectivos a los que se les llama medidas. Se pueden añadir cuantas matrices sea necesario, construyendo un cubo de datos, tal y como se muestra en la figura 14.

Figura 14. Dimensiones y medidas



3.2.2 Jerarquía y abstracción de los datos

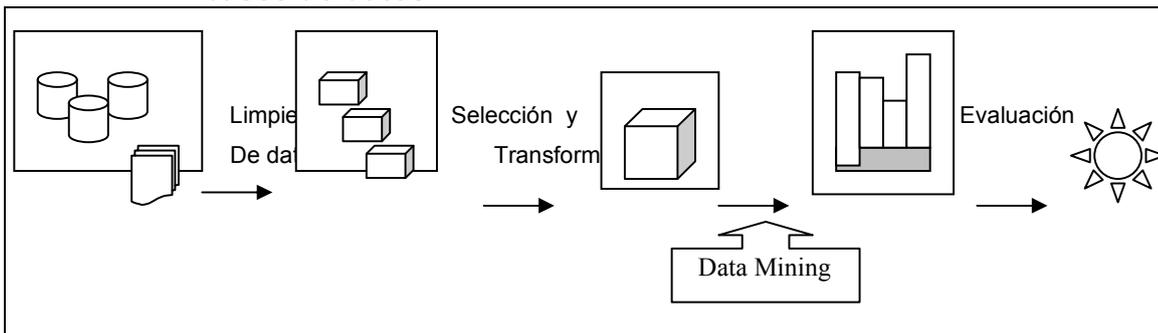
La definición de una dimensión podría incluir el concepto de jerarquía, que describe a la dimensión en términos de secuencia de matrices desde un nivel bajo hasta un nivel alto de dominios. Por medio de la jerarquía se puede interpretar, cómo las dimensiones pueden ser abstraídas en varios niveles; por lo tanto cada nivel representa un nivel de abstracción.

Cada esquema que describe un nivel conceptual, se llama esquema de jerarquía. Con la palabra *todo*, se denota al más alto nivel de jerarquía del sistema con el máximo grado de abstracción de una dimensión.

3.3 Proceso de descubrimiento del conocimiento en las bases de datos

Los términos de *Data Mining* y descubrimiento del conocimiento, normalmente se utilizan como sinónimos, pero realmente el *Data Mining* es solo una fase del proceso de descubrimiento del conocimiento de las bases de datos. Dicho proceso consiste en la identificación de esquemas válidos y potencialmente utilizables en los datos.

Figura 15. Fases del proceso de descubrimiento del conocimiento en las bases de datos



El proceso de descubrimiento del conocimiento de las bases de datos, comprende las tres fases que se muestran en la figura 15 y que a continuación se describen:

3.3.1 Fase de proceso previo

También conocida como fase de depuración, debido a que en esta fase se eliminan datos que están incompletos y que son inaplicables. En este proceso previo se transforma la información en bruto encontradas en la base de datos, en una recolección de ítemes con datos completos y relevantes para un determinado tema. Por ejemplo, para analizar los modelos de ventas de una compañía internacional, se construiría un cubo de datos que se centrara solamente en figuras de venta sin hacer caso a la información inaplicable que pueda existir en los archivos planos.

3.3.2 Fase de integración de datos

En esta fase se integran los datos y la consolidación que combina diferentes fuentes, previamente procesados en una fuente homogénea que sea apta para explorar. El almacén de los datos o el cubo de n-dimensiones, se construye normalmente en esta etapa. El propósito de construir un *Data Warehouse* es crear una vista consolidada de datos importantes que serán el punto de partida para el proceso de toma de decisiones.

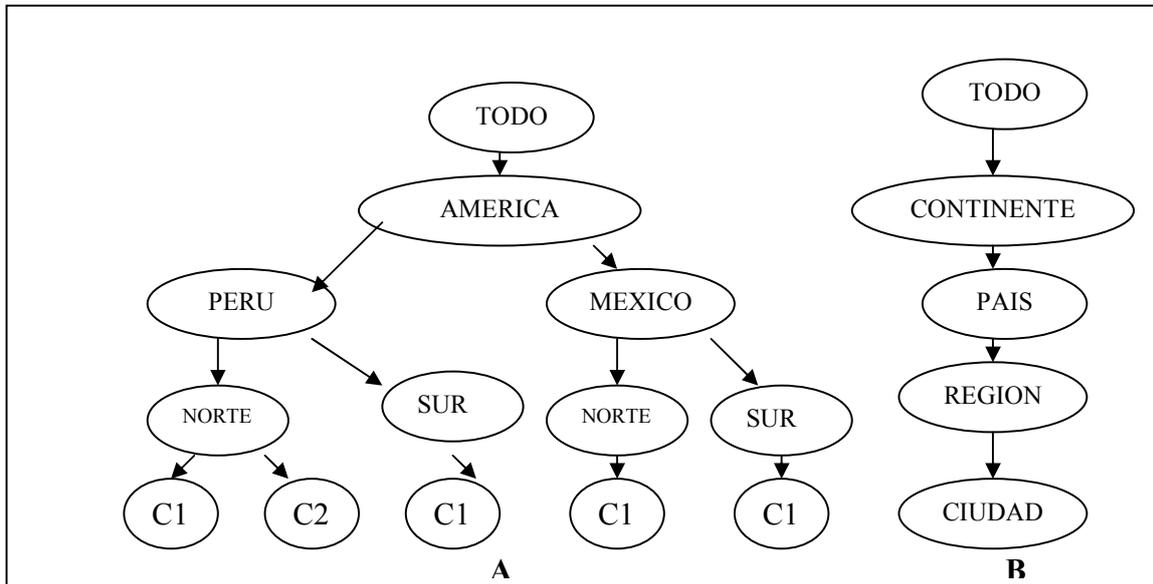
3.3.2.1 Constructor del cubo de datos CCD

El constructor del cubo de datos es un módulo de la caverna virtual, responsable de la culminación del procesamiento previo de los datos y de la consolidación de estos o mejor dicho de la construcción del cubo de datos. El constructor del cubo de datos esta instalado en un sistema remoto, y consiste en un servidor que posee un modelo de distribución de tipo cliente-servidor. En la información en bruto proveniente de fuentes simples o múltiples se busca primero según los criterios dados a los ítemes incompletos e inaplicables durante el proceso previo de aislar los datos. En el caso de existir fuentes múltiples, el proceso de búsqueda se ejecuta en un *servidor principal* que utiliza métodos de acceso remotos capaces de ejecutar las averiguaciones apropiadas del *SQL*. Con esta información entonces, se recopilan por relevancia y de manera homogénea, los datos necesarios para la creación del cubo de datos. El constructor del cubo de datos también puede ser instruido para crear más de un cubo de datos en casos en los cuales se considere más de un tema central a la vez. Éstos cubos de n-dimensiones de datos se llaman *Conjunto de Almacén de Datos de N-dimensiones*. Para que el Constructor del cubo de datos pueda crear el cubo de datos, primero se debe extraer la información estructural sobre las fuentes de datos; esta información se utiliza para definir cada una de las n-dimensiones del cubo.

Cuando se habla de una dimensión o variable, se debe asociar inmediatamente con una perspectiva organización lógica de entidades según las cuales la empresa desee una visión de los datos. En la definición de una dimensión también debe incluir la jerarquía conceptual, que describe lo más amplio posible el concepto de dimensión en términos de una secuencia de datos entre los conceptos de alto y bajo nivel. Cada nivel de esa jerarquía conceptual define un nivel de campo o abstracción. Un ejemplo clásico de una dimensión del cubo de datos es el *tiempo*, que se asocia generalmente a la jerarquía conceptual (día, mes, año, trimestre). La presentación ordenada en diferentes niveles de acuerdo a la jerarquía de la información ayuda a mejorar la comprensión de la información a la hora de la toma de decisiones.

Al subir de jerarquía de *día* a *mes*, los datos pueden ser vistos en un nivel más alto del campo abstracto que proporciona cada vez menos detalle (lista ascendente). Es importante precisar que este proceso no puede automatizarse completamente, ya que el concepto de dimensión es definido por el usuario y no se considera el tema principal. Por ejemplo, para una dimensión típica, como la *localización*, descrita, por ejemplo, por una jerarquía conceptual (continente, país, región, ciudad), la información en bruto dentro del *DBMS*, puede describir localizaciones usando solamente números de identificación del almacén de datos. En tales casos, el experto del sistema debe definir la *ciudad*, qué ciudades corresponden a una determinada *región*, y qué regiones pertenecen a un *país* y así sucesivamente.

Figura 16. A. Concepto de jerarquía. B. Esquema de jerarquía.



3.3.2.2 Modelo jerárquico de datos de n-dimensiones

El modelo de datos jerárquico también recibe el nombre de cubo de datos, cuya representación lógica es una estructura geométrica que generalmente se define en tres dimensiones. Un ejemplo de datos de tres dimensiones podría ser el interés de una empresa de negocios, por saber el número de *productos* vendidos en un *mercado* específico, durante un periodo de *tiempo*. Las dimensiones estarían representadas por productos, mercado y tiempo, a los cuales se les asocia un eje cartesiano y que forman un cubo de datos. Se debe de tomar en cuenta que un cubo de datos no necesariamente debe tener tres dimensiones ya que depende del nivel de detalle.

Las dimensiones de un cubo de datos son hereditarias por naturaleza, permitiendo la visualización de los datos desde distintas perspectivas. Ejemplo de ello podría ser la dimensión del tiempo que puede ser representada en años, semestres, meses y días. Los distintos niveles de abstracción de una dimensión pueden ser representados por el concepto de herencia.

3.3.3 Data Mining

Es el paso final del proceso de descubrimiento del conocimiento de las bases de datos, en el cual se realiza un análisis de los patrones por medio de algoritmos especiales. El *Data Mining* es una poderosa tecnología que ayuda a concentrarse en la información más importante, debido a que se basa en la extracción de información oculta y predecible que permite la toma de decisiones pro-activas y conducidas.

En esta etapa la inmersión en el ambiente juega un papel muy importante debido a la necesidad de visualización de datos.

3.3.3.1 Fundamentos del Data Mining

Las técnicas de *Data Mining* son los resultados de un largo proceso de investigación y desarrollo de productos. Esta evolución comenzó cuando los datos del negocio se almacenaron por primera vez en computadoras, y continuó con mejoras en el acceso de los datos y, más recientemente con tecnologías generadas para permitir a los usuarios navegar a través de ellos en tiempo real. El *Data Mining* esta soportado por tres tecnologías que han sido desarrolladas de una manera robusta:

- Recolección masiva de datos
- Potentes computadoras con multiprocesadores
- Algoritmos de *Data Mining*

Las bases de datos crecen a un ritmo sin medida, la necesidad paralela de motores computacionales mejorados, se alcanza de forma más efectiva con tecnología de computadora con multiprocesamiento paralelo.

Los algoritmos de *Data Mining* utilizan técnicas que pueden redituar los beneficios de automatización en las plataformas de *hardware* y *software* que utilizan procesamiento paralelo de alto rendimiento y que permite el análisis de bases de datos masivas en minutos. Procesamiento más rápido significa que los usuarios pueden automáticamente experimentar con más modelos y con datos más complejos. La alta velocidad hace que sea práctico para los usuarios analizar inmensas cantidades de datos y a su vez, grandes bases de datos producen mejores predicciones.

Los algoritmos de *Data Mining* utilizan más comúnmente las siguientes técnicas:

- Redes neuronales artificiales
- Árboles de decisión
- Algoritmos genéticos
- Métodos del vecino más cercano
- Regla de inducción

3.3.3.2 Alcance del *Data Mining*

Al proveer bases de datos de suficientemente tamaño y calidad, la tecnología de *Data Mining* puede generar nuevas capacidades:

- **Predicción automatizada de tendencia y comportamiento.** El *Data Mining* automatiza el proceso de encontrar información predecible en grandes bases de datos. Las preguntas que tradicionalmente requieren un intenso análisis manual, se pueden contestar directa y rápidamente desde los datos.
- **Descubrimiento automatizado de modelos previamente desconocidos.** Las herramientas de *Data Mining* barren las bases de datos e identifican modelos previamente escondidos en un solo paso. Otros problemas de descubrimiento de modelos incluye identificar datos anormales que pueden representar errores de ingreso en la carga de datos.

3.4 Procesamiento de transacciones en línea (*OLPT*)

Es un procesamiento en tiempo real diseñado para facilitar las aplicaciones orientadas al manejo de transacciones. A diferencia del procesamiento en lote, que procesa datos solo en momentos específicos, el procesamiento de transacción en línea permite optimizar la administración y proveer de información exacta instantáneamente.

El procesamiento de transacciones en línea, es uno de los métodos más utilizados para la implantación de grandes sistemas, como por ejemplo, los cajeros electrónicos, que se conectan a un computador central cada vez que un tarjeta-habiente realiza una transacción en el. La filosofía de dicho procesamiento, es la actualización inmediata de los datos para que se refleje en todos los puntos que hacen parte de la red de información y principalmente en el ambiente virtual.

En un *Data Warehouse* se depositan datos para consulta, análisis y divulgación, a diferencia del *OLPT*, en donde los datos se reúnen y almacenan para operación y control.

3.5 Procesamiento analítico en línea (*OLAP*)

El procesamiento analítico en línea, también llamado procesamiento analítico o análisis multidimensional; se apoya en una visión multidimensional de datos como el *Data Warehouse*.

El procesamiento analítico en línea es una tecnología de procesamiento y análisis de datos, que crea nueva información a partir de los datos existentes, por medio de un conjunto de transformaciones y cálculos numéricos.

3.5.1 Características de la tecnología OLAP

Dentro de las características que la tecnología OLAP presenta, están las siguientes:

- Visión multidimensional y lógica de los datos en el *Data Warehouse*.
- La visión multidimensional es independiente de la forma en que se almacenan los datos.
- Ofrece una consulta interactiva y análisis de los datos. Usualmente la interacción se realiza en varias pasadas, lo cual provoca la profundización en niveles cada vez más detallados o el ascenso a niveles superiores.
- Ofrece opciones de modelado analítico, incluyendo un motor de cálculo para poder realizar mediciones de datos numéricos a través de muchas dimensiones.
- Maneja modelos de pronóstico, análisis de tendencias y análisis estadísticos.
- Recupera y presenta gráficas y datos tabulares en dos o tres dimensiones.
- Responde con rapidez a las consultas, de modo que el proceso de análisis no se interrumpe y la información no se des-actualiza.

- Tiene un motor de depósito de datos multidimensional que almacena los datos en arreglos.

3.5.2 Arquitectura de la tecnología OLAP

La tecnología *OLAP* (*On-line Analytical Processing*), es un componente importante del bloque de acceso y uso de la arquitectura de referencia del *Data Warehouse*. Los componentes de la tecnología *OLAP*, se capturan en sub-bloques de los bloques de acceso y uso. El componente *OLAP* del sub-bloque de análisis y reporte, representa las opciones de análisis y reporte de los servicios *OLAP* requeridos, mientras que la transformación a la estructura multidimensional, así como el acceso a los componentes del *Data Warehouse* y de los *Data Marts*, son partes del sub-bloque de acceso y recuperación.

3.5.3 Operaciones OLAP

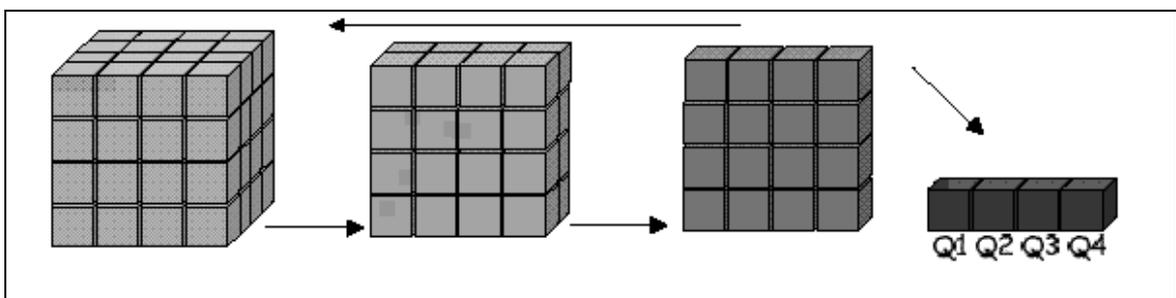
En la implementación de la exploración multidimensional jerárquica y específica de los datos, es necesaria la ejecución de un grupo de operaciones llamadas analíticas en línea. Las operaciones analíticas en línea permiten la exploración interactiva de un *Data Warehouse* y la manipulación del nivel de abstracción de la información que se presenta. Dichas operaciones se clasifican en operaciones de selección de datos y operaciones de agregados y resumen de datos.

3.5.3.1 Operaciones OLAP de selección de datos

La función de las operaciones de selección de datos es enfocarse en el análisis de un subconjunto de datos sin cambiar su integridad. Existen tres tipos de operaciones de selección; de rebanada, de datos y de pivote. Las dos primeras operaciones permiten establecer restricciones en los dominios de las dimensiones de datos; y la operación de pivote, se encarga de intercambiar las uniones de las dimensiones de los datos.

- **Operación de rebanada.** Rebanar un cubo de datos consiste en establecer un valor a lo largo de una dimensión.
- **Operación de datos.** Esta operación permite que la visión sea restringida en más de una dimensión, dando origen a un sub-cubo de los datos originales.
- **Operación de pivote.** También llamada operación de *rotación*. Se utiliza con propósitos de visualización ya que al ejecutar dicha operación, el usuario puede observar distintos ángulos del objeto virtual.

Figura 17. Operaciones OLAP de selección de datos



3.5.3.2 Operaciones OLAP de agregado de datos

Las operaciones de agregado, aplican funciones matemáticas a un conjunto de datos, que da como resultado diferentes vistas del *Data Warehouse*.

- **Operaciones distributivas.** Estas operaciones son las más sencillas ya que se basan en la división de los datos de entrada, en conjuntos de datos que pueden ser agregados o combinados con otros.
- **Operaciones algebraicas.** Se basan en las características del conjunto de datos para procesar sus valores.

3.5.3.3 Operaciones OLAP de resumen de datos

Las operaciones de resumen de datos se basan en la modificación de una vista de datos dada, para transformarla en una vista más general o más detallada.

- **Operación de subida.** Esta operación agrega datos en una dimensión superior; es decir, sube de un nivel específico a un nivel más general.
- **Operación de descenso.** Esta operación es opuesta a la anterior, ya que consiste en agregar una dimensión inferior; es decir, baja de un nivel general a un nivel más específico.

3.6 Métodos para computar el cubo de datos

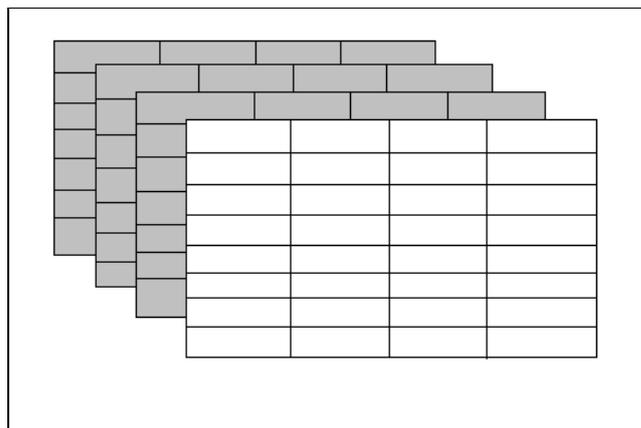
El siguiente paso después de la definición de las dimensiones de los datos y la aplicación del concepto de jerarquía, es procesar la operación de agregado *GROUP BY*, sobre todos los subconjuntos de dimensiones de datos posibles que existan. Para un cubo de datos de n-dimensiones, el número de operaciones *GROUP BY* que se tendrían que aplicar, aumenta en considerable cantidad, y se tiene la limitante que en la Unidad de Control Visual (UCV) de la caverna virtual, solo un número pequeño de dimensiones de datos se pueden visualizar al mismo tiempo. Por tal motivo existen dos tipos de métodos que permiten realizar el análisis de las dimensiones de datos de forma relacional y multidimensionalmente.

3.6.1 Modelo de procesamiento analítico relacional en línea (ROLAP)

ROLAP (Relational On-line Analytical Processing), es un tipo de herramienta que provee a los usuarios la capacidad de realizar análisis de grandes volúmenes de datos almacenados en bases de datos relacionales. Esto quiere decir que una celda en el espacio multidimensional, se representa en el sistema como una tupla, con algunos atributos para identificar su valor y la ubicación en el espacio. Por ello el método *ROLAP*, es llamado *Basado en el valor*.

La arquitectura *ROLAP*, accesa a los datos almacenados en un *Data Warehouse* para proporcionar los análisis *OLAP*. La premisa de los sistemas *ROLAP* es que las capacidades *OLAP* se soportan mejor contra las bases de datos relacionales.

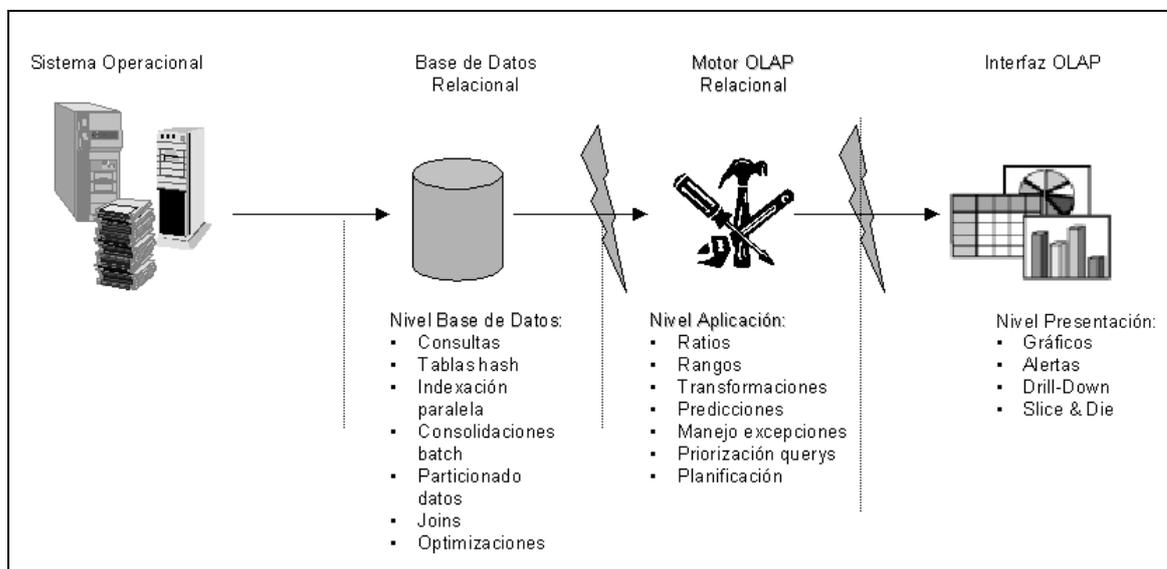
Figura 18. El sistema *ROLAP* utiliza tablas relacionadas para el almacenamiento de los resultados de las funciones *GROUP BY*



El sistema *ROLAP* utiliza una arquitectura de tres niveles.

- **El nivel de base de datos.** Utiliza bases de datos relacionales para el manejo, acceso y obtención de los datos.
- **El nivel de aplicación.** Es el motor que ejecuta las consultas multidimensionales de los usuarios.
- **El motor *ROLAP*.** Se integra con los niveles de presentación, a través de los cuales los usuarios realizan los análisis *OLAP*.

Figura 19. Arquitectura del sistema ROLAP



Fuente: <http://www.map.es/csi/silice/DW2251.html>

Después de que el modelo de datos se ha definido para el *Data Warehouse*, los datos se cargan desde el sistema operacional; luego se ejecutan rutinas de bases de datos para agregar el dato, y entonces se crean los índices para optimizar los tiempos de acceso a las consultas.

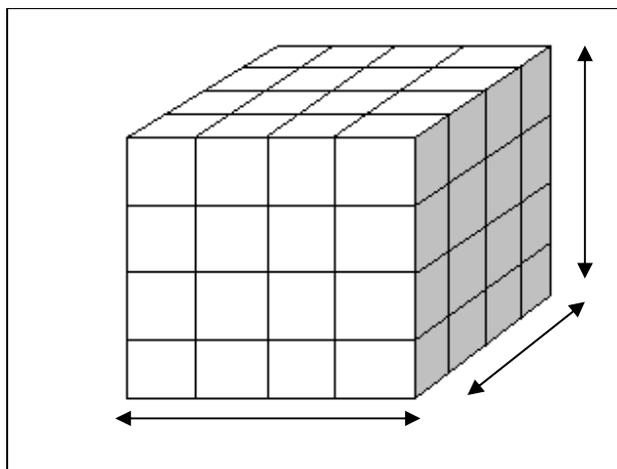
Los usuarios finales ejecutan los análisis multidimensionales, a través del motor *ROLAP*, que se encarga de transformar dinámicamente las consultas a consultas *SQL*. Se ejecutan estas consultas *SQL* en las bases de datos relacionales, y sus resultados se relacionan mediante tablas cruzadas y conjuntos multidimensionales para devolver los resultados a los usuarios.

La arquitectura *ROLAP* es capaz de usar datos precalculados si estos están disponibles, o de generar dinámicamente los resultados desde los datos elementales. Esta arquitectura accede directamente a los datos del *Data Warehouse*, y soporta técnicas de optimización de accesos para acelerar las consultas. Estas optimizaciones son, *particionado de los datos a nivel de aplicación, soporte a la desnormalización y joins múltiples*.

3.6.2 Modelo de procesamiento analítico multidimensional en línea (*MOLAP*)

La arquitectura *MOLAP* utiliza bases de datos multidimensionales para proporcionar el análisis, su principal premisa es que el *OLAP* está mejor implantado ya que almacena los datos multidimensionalmente. Por el contrario, la arquitectura *ROLAP* cree que las capacidades *OLAP* están perfectamente implantadas sobre bases de datos relacionales.

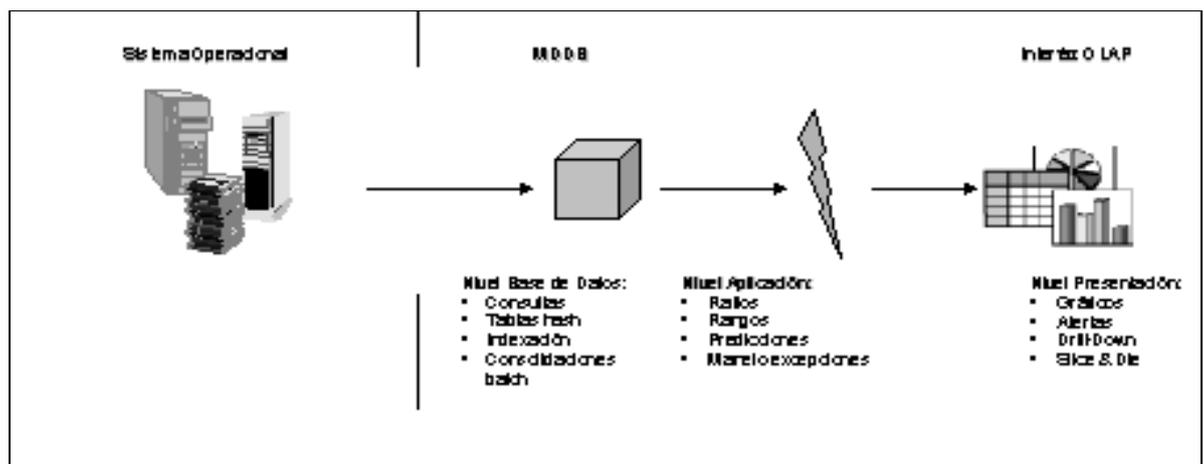
Figura 20. El sistema *MOLAP* utiliza arreglos contiguos para el almacenamiento de los resultados de las funciones *GROUP BY*



Un sistema *MOLAP* usa una base de datos multidimensional llamada también *Basada en la posición*; en la que la información se almacena y se visualiza multidimensionalmente, y utiliza una arquitectura de dos niveles como se muestra en la figura 21.

- **El nivel de la base de datos multidimensional.** Es el encargado del manejo, acceso y obtención de datos.
- **El nivel de aplicación.** Es el responsable de la ejecución de los requerimientos *OLAP*. El nivel de presentación se integra con el de aplicación y proporciona una interfaz a través del cual los usuarios finales visualizan los análisis *OLAP*. Una arquitectura cliente / servidor permite a varios usuarios acceder a la misma base de datos multidimensional.

Figura 21. Arquitectura del sistema *MOLAP*



Fuente: <http://www.map.es/csi/silice/DW2251.html>

La información procedente de los sistemas operacionales, se carga en el sistema *MOLAP*, mediante una serie de rutinas batch. Una vez cargado el dato elemental en la Base de Datos multidimensional (*MDDB*), se realiza una serie de cálculos en batch, para calcular los datos agregados a través de las dimensiones del negocio y rellenar la estructura *MDDB*.

Tras rellenar dicha estructura, se generan índices y algoritmos de tablas *hash* para mejorar los tiempos de accesos a las consultas. Una vez que el proceso de compilación se ha terminado, la *MDDB* está lista para su uso, en el cual, los usuarios solicitan informes a través de la interfaz; y la lógica de aplicación de la *MDDB* obtiene los datos.

La arquitectura *MOLAP* requiere de cálculos intensivos de compilación. Lee los datos precompilados, y tiene capacidades limitadas de crear agregaciones dinámicamente o de hallar ratios que no se hayan precalculados y almacenados previamente.

3.7 Data Warehouse Virtual (DWV)

El *Data Warehouse Virtual* es una conceptualización de un *Data Warehouse* centralizado que está formado por un conjunto de datos distribuidos y del *CCD-Shell*, que se encarga del manejo y consulta de esos datos.

El CCD-Shell construye y actualiza un modelo multidimensional. El CCD-Shell no almacena ninguna transacción localmente, pero mantiene un repositorio de meta-data que se sincroniza con todos los recursos para mantener actualizada la información, para que la distribución de datos sea lo más transparente posible y así evitar inconsistencias; por ello el nombre de *Warehouse Virtual*

La meta-data del CCD-Shell mantiene una tabla con la información de la ubicación y organización de los datos. Toda la meta-data y los datos de ubicación tienen un formato *XML*, para facilitar la interpretación, el mantenimiento, y hacer que todo el sistema sea extensible y flexible.

Similar al *Warehouse* tradicional, el *Warehouse* virtual brinda una serie de tipos de consultas para suplir los requerimientos de los clientes. Existen tres clases de consultas; la primera llamada *consulta Warehouse*, la cual provee una estructura básica del DWV, que incluye el número de cubo de datos y el tema principal de cada cubo. La segunda llamada *consulta del esquema del cubo* que muestra la meta-data de un cubo de datos específico en el DWV. Esta meta-data incluye una descripción de todas las dimensiones válidas y sus respectivas jerarquías. Y finalmente, la *consulta del cubo de datos*, que se utiliza para obtener un cubo entero o un subconjunto de él mismo. Esto es particularmente útil en las aplicaciones de la UCV, que requiere de cubos con dimensiones pequeñas para las sesiones de visualización. Todas las solicitudes de consulta y sus respuestas se realizan en el formato *XML*.

4. INTERACCIÓN EN EL AMBIENTE VIRTUAL

La interacción en el ambiente virtual conlleva a una serie de procesos que se encargan, por medio de interfaces; de la transformación de datos multidimensionales, a la presentación de los objetos virtuales ante el usuario.

Con técnicas y funciones especiales, dicha presentación se realiza de manera que el usuario experimenta una inmersión en el mundo virtual totalmente transparente y real. Aspectos como la localización de la posición del usuario, son indispensables para la transparencia de la actualización de imágenes y del movimiento de menús.

4.1 Visualización de datos

Existen varias técnicas de visualización que se emplean en los ambientes virtuales, tal como la visualización iconográfica de datos, la cual es una técnica que emplea los atributos visuales de íconos para ligarlos a los datos; o bien técnicas que producen contraste de texturas correspondientes a estructuras de datos. Sin embargo, la técnica de visualización que se emplea en el ambiente de la caverna virtual está orientada a un nivel conceptual.

Los objetos del ambiente virtual no están diseñados para transmitir al usuario un resultado específico, sino para transmitir el significado de ese resultado.

El sistema visual humano, es bastante complejo debido a que no solo es un procesador de datos, sino también es un reconocido de patrones. Actualmente se han introducido varias herramientas gráficas que producen efectos especiales, muy similares a los que se ven en la realidad y que utilizan dispositivos comunes tal como el mouse o el teclado para interactuar con el mundo virtual; sin embargo, se ven limitados debido a que los dispositivos comunes no facilitan la interacción y los resultados de un proceso de *Data Mining* en una representación visual natural, ya que implica un ambiente de n-dimensiones.

En el sistema de visualización de datos de la caverna virtual, el usuario puede caminar, inspeccionar y evaluar las notas visuales que son producto de un proceso de *Data Mining*. El usuario utiliza un puntero virtual, el cual le permite interactuar con el mundo, y proporcionar su ubicación en el área constantemente.

La interacción del usuario por medio de la utilización del puntero virtual, consiste por ejemplo, si se tiene alguna duda sobre un objeto en tres dimensiones proyectado en el ambiente virtual, en presionar el botón del puntero, para que se despliegue una ventana con la información correspondiente al objeto seleccionado.

4.1.1 Arquitectura de la visualización de datos

La arquitectura del sistema de visualización está formado por tres módulos:

- Constructor del cubo de datos (CCD)
- Cubo de memoria de imágenes
- Unidad de control visual

No es necesario que los tres módulos estén ubicados físicamente en el mismo lugar, ya que su comunicación se realiza por medio de la red utilizando la arquitectura *CORBA*, en donde el constructor del cubo de datos actúa como el servidor y tanto el cubo de memoria de imágenes como la unidad de control visual actúan como clientes.

4.1.2 Unidad de control visual (UCV)

La unidad de control visual cuyas siglas en inglés son *VCU (Visualization Control Unit)*, es el módulo responsable de generar y de manejar el ambiente virtual para la visualización y la exploración de los datos.

La unidad de control visual, esta formada por seis sub-sistemas:

- Interfaz de *CORBA*
- Cubo de Datos
- Control de gráficos
- Motor de *OpenGL*
- Administrador de la interfaz del usuario
- Administrador de *Octree*

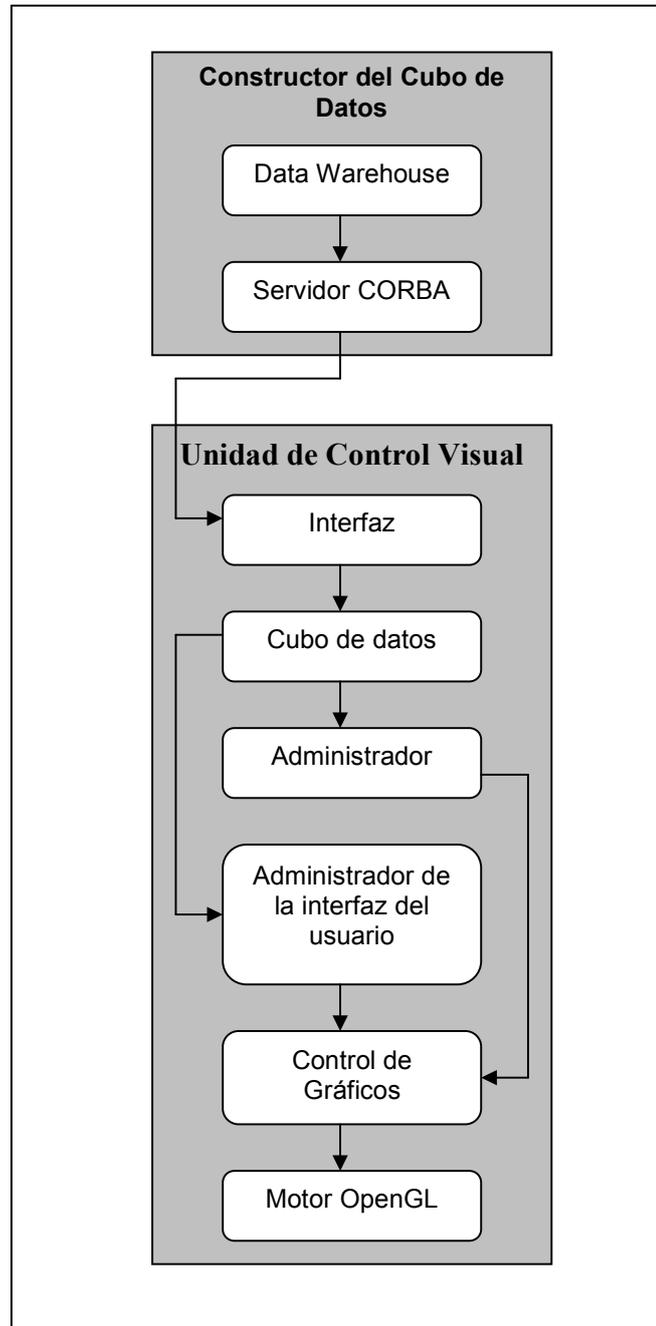
Cuando la Unidad de Control Visual inicia su funcionamiento, se presenta al usuario, el mundo virtual. En este punto el término *mundo* se refiere al conjunto de cubos de datos que pasan por el administrador *Octree* hacia el subsistema de control de gráficas. El usuario debe especificar el dominio sobre el cual desea trabajar, ya que con esta información la interfaz de *CORBA* solicita la información que necesita el constructor del cubo de datos, para la construcción del cubo de datos en su respectiva plataforma. Dicha información incluye los atributos de los cubos y el concepto de jerarquía que define los niveles de abstracción de datos.

Existen tres variables del control ambiental que se deben especificar en el momento de la creación del mundo virtual:

- Atributos
- Rango
- Nivel de abstracción

El cubo de datos se representa en una estructura simple de datos en memoria; los datos de entrada del usuario; se obtienen por medio del administrador de la interfaz del usuario, el cual determina los tres atributos que constituirán las dimensiones *x*, *y* y *z*.

Figura 22. Componentes de la unidad de control visual

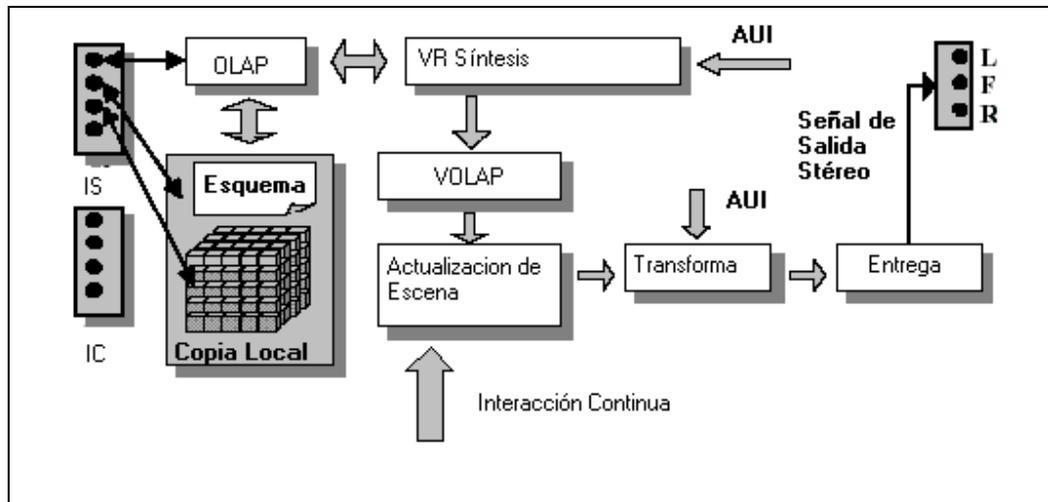


El CCD debe ser independiente de la UCV, y viceversa. El CCD y la UCV deben construirse dentro de una capa intermedia que proporcione los medios para retransmitir mensajes entre los dos subsistemas. En el sistema de exploración de datos se implementa un protocolo de comunicación simple, que define un conjunto de peticiones, UCV a CCD; y sus correspondientes respuestas, CCD a UCV. Una vez que el CCD completa la creación del cubo N-dimensional de datos, envía una señal a la UCV por medio del CCD-Shell; puesto que se genera un mundo virtual en tres dimensiones, solo un subconjunto de esas dimensiones se pueden ver en cualquier momento.

El administrador *Octree* recibe los objetos y los normaliza para crear una estructura apropiada. El proceso de normalización es muy importante debido a que las magnitudes no son el punto central en las operaciones de *Data Mining*, sino la correlación de los datos que se quieren explorar. Con la utilización de los tres atributos, cada normalización se representa por medio de un cubo proporcional.

Debido a que se ubicará al usuario dentro de un ambiente de inmersión virtual, es imprescindible que los retrasos sean mínimos. Por esta razón los sistemas de inmersión virtual confían en la UCV, para agregar los datos menos detallados. Esto reduce con eficacia la dependencia de la red a un mínimo.

Figura 23. Arquitectura de la UCV



La UCV, está compuesta por un módulo llamado *VOLAP*, el cual se utiliza para proveer escalabilidad en la orientación de grandes cantidades de datos. Así como la cantidad de objetos que se entregan incrementa, el marco de las transformaciones de la imagen disminuye. Esto efectivamente reduce el sentido de presencia real, debido al retardo entre el movimiento del usuario y la actualización de ese movimiento en la imagen. Para prevenir estas situaciones, las imágenes que se proyectan se dividen en porciones, y el usuario capta una porción a la vez.

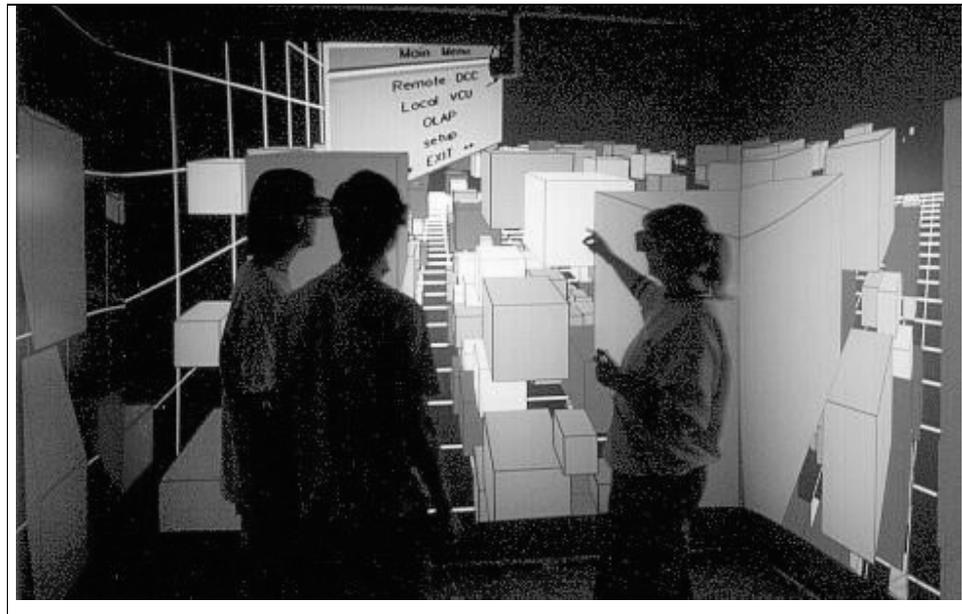
La UCV posee una entrada a la vía de trayectoria en donde se realiza la interacción de parámetros, la cual se alimenta por el Administrador de la Interfaz del Usuario (AIU). La interacción discreta de parámetros describe al comando de usuario, como si le perteneciera los datos que se están visualizando. La interacción continua de parámetros describe la ubicación del usuario en el ambiente.

4.1.3 Creación de los objetos virtuales

Los datos que se presentan al usuario durante la inmersión en el ambiente virtual, se transforman en objetos gráficos que unidos forman parte del mundo virtual. La UCV visualiza los cubos de datos tridimensionalmente, y los recibe del *Data Warehouse Virtual* por medio de una función de tres variables. Cada variable representa una dimensión de un dato determinado, que se asocia a una dimensión física x , y y z .

Cada entrada al cubo de datos, es una estructura que contiene dos medidas $M1$ y $M2$; la UCV las traza en dos funciones $M1(x, y, z)$ y $M2(x, y, z)$, en un plano de tres dimensiones. Por ejemplo, si se analizara el tema de “productos vendidos”, la decisión de un procesamiento *OLAP*, no estaría interesada tanto en el hecho de que un determinado año a , el total de ventas de un producto p , fue de una cantidad c ; sino lo importante es el contexto en el cual las cantidades o medidas ocurren. La UCV expresa dicho contexto, por medio de la asociación de las cantidades con características visuales que se ligan a objetos virtuales; la primera característica, es el tamaño t , con la cual se asocia a la medida $M1$ (productos vendidos). Después de la normalización, $M1(Xa, Yp, Zc)$ se utiliza para formar un cubo o esfera con esa longitud o radio respectivamente, ubicado en la posición Xa , Yp y Zc ; para el año a , producto p y la cantidad c dentro de un rango establecido. En la figura 24, se muestra a los usuarios discutiendo el tema de productos vendidos.

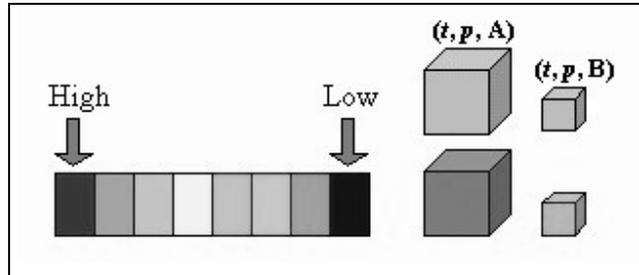
Figura 24. Usuarios sumergidos en un sistema de realidad virtual, discuten el cubo de datos



Fuente: <http://www.cs.ualberta.ca/~ayman/dive-on.htm>

La segunda característica es el color del objeto, el cual se escoge entre ocho colores que se encuentra en la paleta de colores, como se muestra en la figura 25; y el rango de los valores normalizados para la medida $M2$ se individualiza y asocia a la paleta. El segmento de color rojo se utiliza para indicar valores mas altos de un determinado concepto asociado, disminuyendo hacia la derecha en donde el azul representa el valor mas bajo. El uso de colores para codificar los resultados de la exploración de datos ha mostrado resultados significativos. Por ejemplo, en el nivel más bajo equivalente a información específica, el color puede representar una herramienta efectiva que ayude a localizar anomalías y representar la desviación a lo largo de las dimensiones. Esto es particularmente útil para el análisis de fluctuación de mercado.

Figura 25. Paleta de color

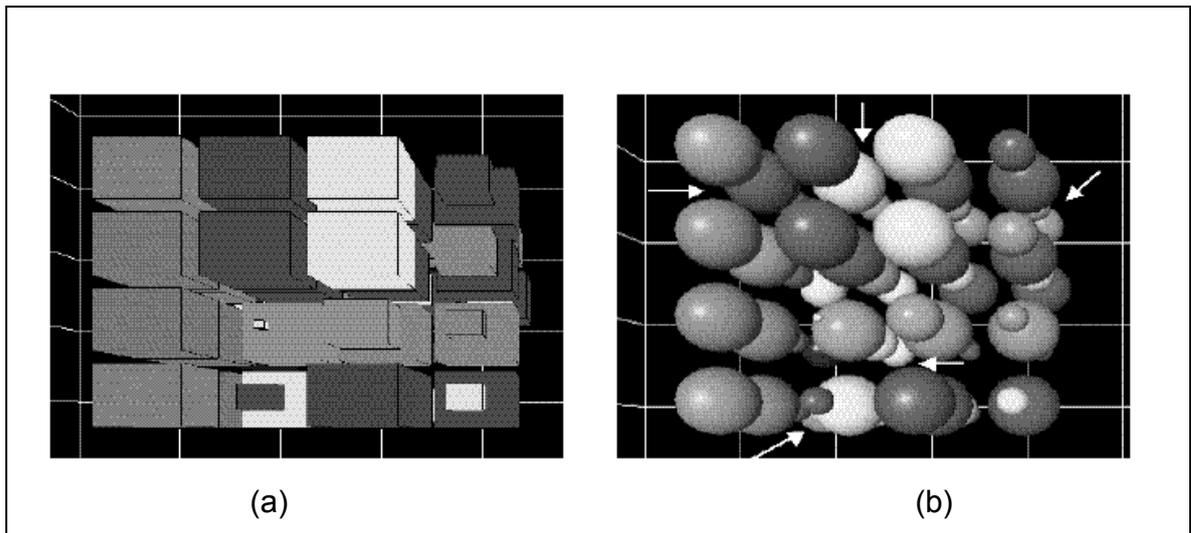


Fuente: <http://www.acm.org/crossroads/espanol/xrds7-3/diveon.html>

Los cubos como objetos geométricos incorporan información gráfica, cada uno con la talla y el color según lo descrito anteriormente. Sin embargo, con la utilización de esferas en lugar de cubos, se puede presentar la misma cantidad de información, obstaculizando la visualización de menos objetos. Es posible observar una esfera más pequeña detrás de otra más grande; lo cual esto no es posible al tratar con los cubos; sin embargo, es importante precisar que un sistema virtual con esferas es mucho más complejo que con los cubos.

Para crear una esfera tridimensional, el sistema debe realizar simulaciones de la fuente de luz y sus reflejos, cálculos normales del vector y la especificación del material. Ninguno de estos cálculos se requieren en el caso de los cubos, debido a que los cálculos son más sencillos.

Figura 26. Atributos de datos, utilizando cubos y esferas desde un mismo punto de vista. (a) Cubos como Agregado. (b) Las flechas señalan algunos de los datos que han sido ocultados



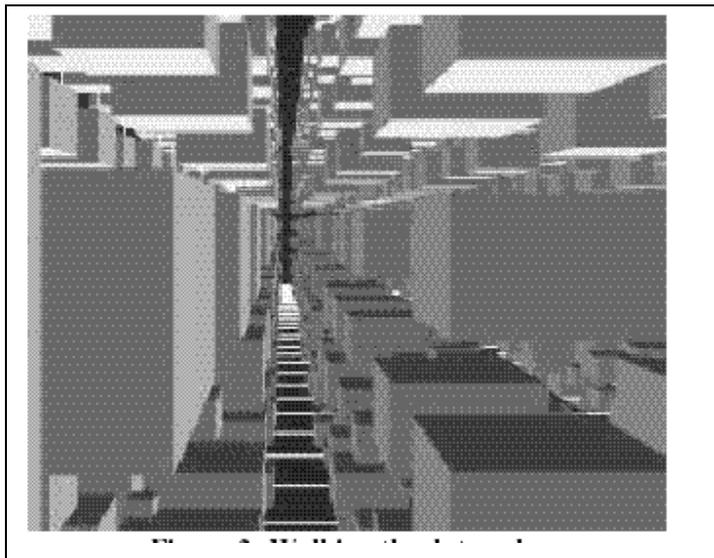
Fuente: Ayman Ammoura, Osmar R. Zaïyane, y Yuan Ji. Immersed Visual Data Mining: Walking The Walk Página No. 6

4.1.4 Administrador de la interfaz del usuario (AIU)

El AIU es el componente de la caverna virtual que maneja todos los aspectos de la reacción del ser humano con el ordenador. Todas las señales e imágenes son examinadas por el AIU, las cuales se alimentan de la UCV para mantener el funcionamiento del ambiente virtual. La actualización constante de la localización del usuario es necesaria para determinar la localización inicial del menú flotante, del número activo del menú y las opciones actuales del mismo.

El AIU utiliza esta información para trasladarlos a gráficas estéreo y simular un mundo real. Por ejemplo, si el usuario camina en el ambiente, las imágenes deben cambiar, para dar la sensación de que se introduce en el ambiente. El flujo de información emitida por la mano del usuario, se utiliza para fijar la posición del menú tridimensional en el mundo virtual. Esta es la razón por la cual se llaman menús flotantes, ya que cambian de posición conforme cambia de posición del usuario.

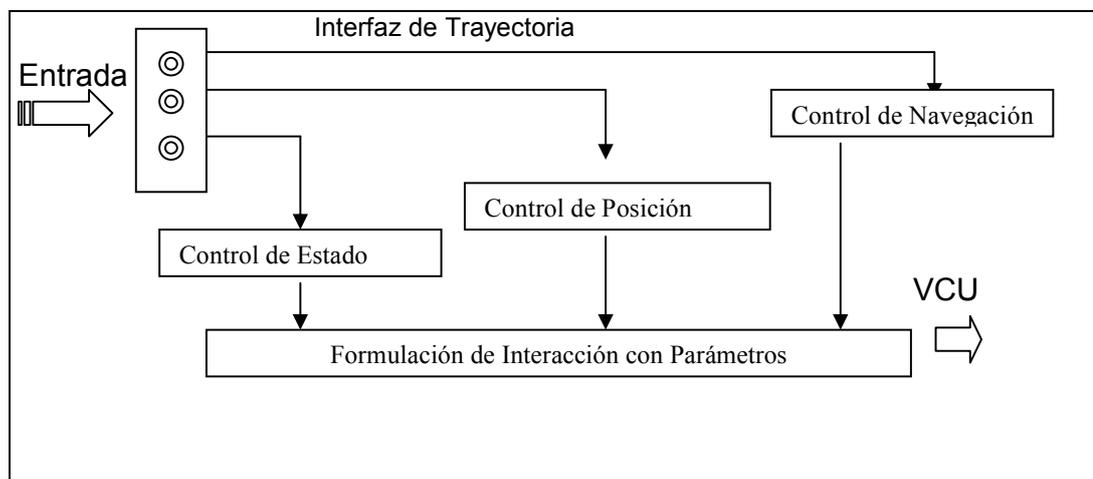
Figura 27. Vista a través del cubo de datos



Fuente: Ayman Ammoura, Marc Perron y Osmar Zaïane. *Walking the Data Cube: An Immersed Virtual Environment for Visualising Data Mining Results.* Página No. 4

En cada aplicación de inmersión en ambiente virtual, la esencia de la interacción hombre-computador se puede clasificar en tres categorías que son manejadas por el AIU como manipulación del objeto, a la manipulación del punto de vista, o control de la aplicación. La razón de esta taxonomía es el hecho de que toda la simulación de la realidad es cerca de la simulación de visiones que cambian alrededor de los usuarios junto con la capacidad de interactuar con los objetos a su alcance. Como resultado de esto, debe usarse una interfaz altamente transparente para el sistema sensomotriz del ser humano, que proporcione una visión acorde al movimiento realizado por el usuario.

Figura 28. Esquema del administrador de la interfaz del usuario



4.1.4.1 Manipulación del punto de vista

Existen ciertas características necesarias que deben tener las aplicaciones para maximizar el sentido de la realidad y que son utilizadas para crear vistas apropiadas de los datos generados en nuestro sistema. El sistema de inmersión virtual manipula el punto de vista para producir dos vistas naturales distintas que representa los datos durante la inmersión en el ambiente virtual.

- **Punto de vista del usuario.** Para simular experiencias diarias, el ambiente virtual se construye desde el punto de vista del usuario como si él estuviera en el centro del mundo virtual, para crear eficacia un marco egocéntrico de referencia.
- **Punto de vista exocéntrico.** El segundo marco de referencia esta disponible para proveer al usuario los medios para substraerse fuera de sí, del mundo virtual. Las visiones egocéntricas son esenciales para la exploración local de los datos donde el usuario puede examinar las relaciones que existen entre las características de los datos consecutivos y tener acceso a todos los atributos disponibles que pertenezcan a un objeto específico. Inversamente, el acercamiento egocéntrico permite al usuario examinar y detectar los modelos globales mirando los datos de un punto de vista exterior.

4.1.4.2 Control de la aplicación

La razón por la cual se ha adoptado esta división, es debido a que la simulación real abarca los cambios de vista del usuario y la habilidad de interactuar con los objetos plasmados en las vistas.

4.2 Interacción con el cubo de datos virtual

La realidad virtual debe simular al mundo real. Esto se aplica también a la visualización de datos. El ser humano se siente más confortable en un ambiente familiar, por lo tanto el mundo virtual debería cumplir las leyes físicas. Pero eso no siempre se puede cumplir debido a las limitaciones de los sistemas gráficos computacionales.

4.2.1 Perspectiva y tamaño

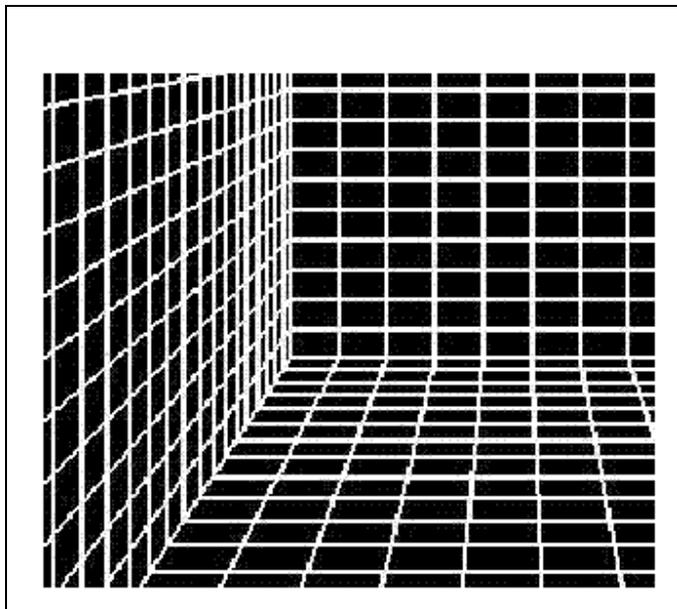
Computadoras modernas capaces de procesar gráficas en tres dimensiones, utilizan un modelo de cámara sintética; que trata de emular la operación de una cámara que posee lentes al frente y detrás de ella. En esta configuración las imágenes son proyectadas a través de lentes sobre el plano de la película. Sistemas de computadoras pueden emular estos sistemas utilizando simples cálculos trigonométricos.

Uno de los problemas más grandes que se ha dado en este tipo de sistemas, es la imposibilidad de poder proyectar una imagen que excede del volumen establecido y que muchas veces no resulta con el mismo efecto si se divide en fracciones. La solución para este problema es la graduación a escala de la imagen que se quiere proyectar, para lograr ver su totalidad y proporcionar una visualización natural.

4.2.2 Orientación

Cuando un usuario está inmerso en un mundo virtual, es importante mantener el concepto del origen. Los usuarios deben saber donde están localizados con respecto al dominio en tres dimensiones. Cuando el usuario ve un cubo, luego experimenta la curiosidad de saber que recurso de información representa; por ello se adjunta un menú en donde se menciona los atributos representados y su concordancia con los ejes x , y y z . Al mismo tiempo el usuario necesita un marco de referencia y perspectiva para determinar el tamaño de los objetos en el mundo. Para ello se utiliza como base un marco trazado a cuadrícula como se muestra en la figura 29. Esto sirve tanto para dar la sensación de perspectiva como para tener la referencia de donde se deben plasmar los cubos.

Figura 29. Marco trazado a cuadrícula



Fuente: Ayman Ammoura, Marc Perron y Osmar Zaïane. Caminado en el Cubo de datos: Inmersión en el ambiente virtual para la visualización de resultados de Data Mining . Página No. 5

4.2.3 Interacción continua

El término *continuo* se refiere a la salida continua del flujo de datos que se reporta al AIU y su respectiva actualización. En este proceso de interacción, el usuario desempeña un rol pasivo, ya que la interacción continua se debe dar de tal manera que sea completamente transparente al usuario. Por tal motivo, la UCV debe tener retroalimentación constante del estado del proceso de inmersión.

El usuario provee información transparente e importante con la utilización del casco, ya que se crea una estructura de datos que es recibida por el AIU, con la cual se calcula las coordenadas relativas x , y y z del dispositivo. Estos datos también son recibidos por la UCV para tener la ubicación y orientación de su cabeza. Este proceso es importante ya que se calculan todas las perspectivas del punto de vista del usuario, con lo cual se obtiene las transformaciones de imágenes conforme al movimiento del usuario y se crea realismo en la inmersión.

El apuntador que el usuario utiliza en su mano, también es una fuente de información para ubicar su posición. Dicho dispositivo es una herramienta que permite al usuario señalar, consultar y manipular los objetos en el ambiente.

4.2.4 Interacción discreta

La interacción discreta se refiere a la interacción, hombre-computadora, que se inicia por el usuario al cumplir con una necesidad, solicitud o acción. Esta clase de interacción incluye todas las funciones de navegación, consulta de objetos y llamadas a menús. A diferencia de la interacción discreta, que es totalmente transparente al usuario, la iteración puede inicializarse y terminarse en cualquier instante que se desee.

El apuntador posee tres botones con los cuales el usuario puede seleccionar las funciones principales del sistema. El AIU es una máquina de estados, esto significa que la interpretación de un botón presionado depende del estado actual de la interacción discreta. Por ejemplo, si el primer botón se presionara mientras el menú esta activo, el AIU podría interpretarlo como *scroll* a través de las opciones. Pero si el menú todavía no esta activado, entonces el primer botón podría activar el sistema de menús. La interacción del menú es proporcionada por un tipo de menús en tres dimensiones, que se integran al mundo virtual.

4.2.4.1 Menús flotantes

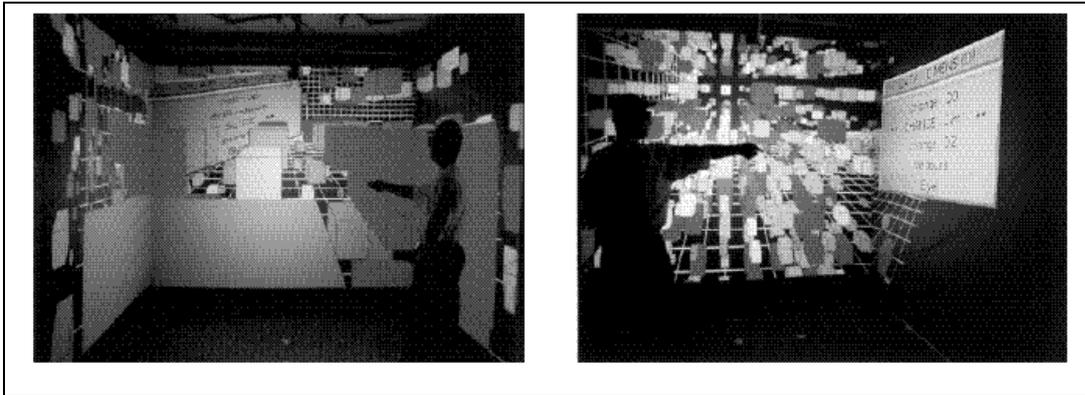
Al explorar un menú en el ambiente, seguramente se ocultan partes de la escena virtual. Para solucionar este problema existen menús flotantes con la propiedad *hand-coupled*, que garantizan legibilidad del texto y evita que se oculten escenas virtuales. El apuntador posee tres botones los cuales se utilizan para mostrar un menú, *scroll* a través de él y ejecutar una opción determinada. Al seleccionar un menú, este se ubica de acuerdo a la posición del puntero y se mueve en relación al movimiento del usuario, creando el efecto de un menú flotante en el ambiente virtual. El menú puede ser mostrado en cualquier instante y también puede suspenderse mientras el usuario continúa examinando los datos antes de hacer una selección.

El sistema de menús permite que el usuario ejecute las funciones *OLAP*, de ambiente, remotas y locales, por medio de la localización de una acción.

Como ya se mencionó, la posición del puntero se localiza constantemente por el mundo virtual, ya que es necesario que el puntero virtual tenga sincronía en el movimiento sobre las paredes de la caverna. Es decir que cuando el usuario mueva físicamente el puntero, el puntero virtual sobre las pantallas, se deberá mover en la misma dirección.

El usuario puede realizar dos operaciones *OLAP* como por ejemplo, la operación de rebanada y la operación de dado, por medio de la utilización del puntero virtual que identifica los parámetros que se necesitan para dichas operaciones. Para obtener una rebanada de los datos, el usuario debe restringir el rango a lo largo de uno o dos ejes, y en la operación de dado el usuario debe especificar el rango a lo largo de los ejes x, y y z.

Figura 30. Usuarios posicionan los menús en cualquier punto del ambiente virtual

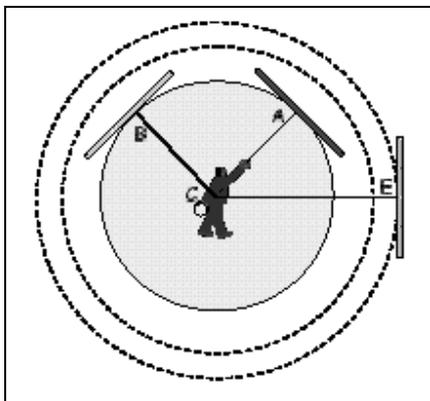


Fuente: <http://www.cs.ualberta.ca/~ayman/dive-on.htm>

4.2.4.2 Movimiento de los cubos de datos en tres dimensiones

En el ambiente virtual el usuario puede mover un cubo de datos de una posición lejana hacia una posición cercana, ante la vista del usuario; o bien a una posición al mismo nivel. Para realizar dichos movimientos, el usuario debe tomar el cubo, realizar ese movimiento con su brazo y colocarlo en el lugar deseado dentro del ambiente. Este tipo de movimientos contribuye a que el sentido de presencia sea aún más real y transparente.

Figura 31. Movimiento de objetos entre órbitas



Fuente: Ayman Ammoura, Marc Perron y Osmar Zaiane. Caminado en el Cubo de datos: Inmersión en el ambiente virtual para la visualización de resultados de Data Mining . Página No. 3

Como se muestra en la figura 31, el usuario ubicado en el centro de la caverna, los círculos alrededor de él, simbolizan las distintas posiciones ya sean lejanas o cercanas a la vista del usuario. El puede mover un objeto que se encuentra en la posición E hacia la posición A, tomando en cuenta el movimiento de su brazo. Al mismo tiempo puede mover el objeto de la posición A, hacia una posición B.

4.3 Interacción basada en el sistema

La interacción basada en el sistema se refiere a la necesidad de mantener el control de la aplicación, para mandar instrucciones al CCD y a la UCV; con respecto a qué y cómo se observan los datos. Como se menciona en la sección 4.1.2; durante el inicio, el sistema envía un mensaje al constructor del cubo de datos, quien solicita el conjunto de todas las N-dimensiones o variables disponibles en el almacén de datos. El correspondiente documento *XML* es recibido por la UCV y presentan al usuario una lista de las cuales él especifica las tres dimensiones de datos que se visualizarán (uno para cada eje *x*, *y* y *z*). En ese punto, el CCD construye el correspondiente cubo de datos tridimensional, en el nivel más bajo del campo y lo envía a la UCV junto con otro documento *XML* conteniendo las tres jerarquías conceptuales respectivas. Las operaciones *OLAP* se realizan localmente, dentro de la UCV, para reducir al mínimo el uso de una red posiblemente congestionada.

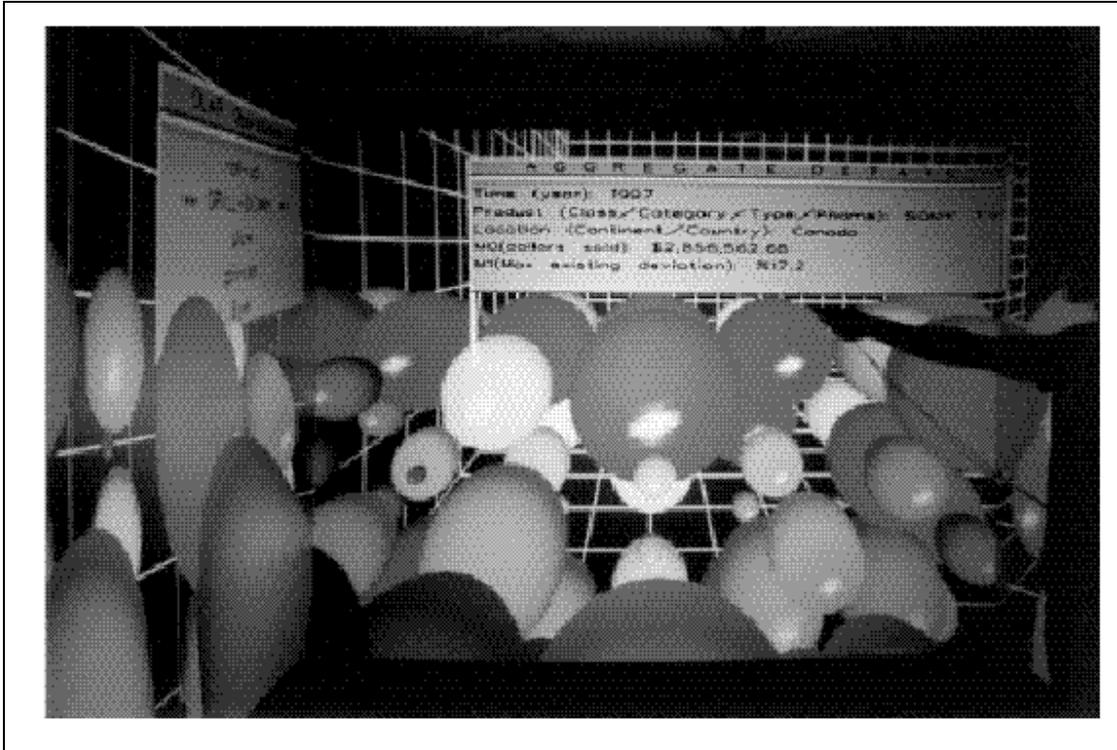
Toda la interacción basada en el sistema es proporcionada a través del uso de la jerarquía flotante del menú tridimensional, que se inicia presionando el primer botón del apuntador. El AIU implementa estos menús con seis grados de libertad para permitir que fluya libremente en la realidad virtual en total sincronización con la mano del usuario. La secuencia de la estructura de datos de la orientación, se transmite desde el apuntador que fue usado para rotar la superficie del menú para el usuario.

4.4 Interacción basada en los datos agregados

Las metas más importantes de un usuario durante la exploración visual de los datos, son identificar lo que está buscando y conocer su contenido. De acuerdo con este resultado, la interacción basada en los datos agregados se ha concentrado sobre todo en proporcionar instantáneamente el linaje asociado a cada señal visual para cualquier objeto dado. Según lo discutido en la sección 4.1.3, la talla y el color de cada objeto es el resultado de medidas normalizadas en el almacén de datos. Puesto que la normalización es irreversible, lo importante es que el sistema mantenga los datos originales usados antes de ser presentados. La localización obtenida del apuntador se utiliza para trazar un puntero en el ambiente de realidad virtual.

Cuando el puntero está cerca de un objeto, el segundo botón activa un pequeño panel de texto que visualiza toda la información correspondiente a ese campo o agregado. Esta información es suficiente como para identificar todos los aspectos de ese agregado, que incluye el nivel actual del campo a lo largo de cada una de las tres variables o dimensiones; y el valor real de la primera y segunda medida. El panel de texto aparece como un rectángulo tridimensional que no es ocultado por ningún otro objeto y esta hecho para apuntar un objeto seleccionado.

Figura 32. Interacción basada en datos agregados



Fuente: <http://www.cs.ualberta.ca/~ayman/dive-on.htm>

4.5 Interacción basada en el ambiente

Este tipo de interacción es el responsable de la manipulación del punto de vista del usuario, y de la construcción y uso de un mapa cognoscitivo del ambiente virtual. Según lo considerado anteriormente, la capacidad de localizar un determinado dato agregado dentro del ambiente, es la segunda búsqueda después de la operación de exploración de datos en la realidad virtual. Proporcionar medios de navegación eficaces es esencial ya que dependiendo del nivel de agregación, sólo una parte de los datos se pueden visualizar.

El tercer botón del puntero se utiliza para el control de la navegación y de la localización. Cuando el usuario presiona dicho botón, tiene dos opciones de movimiento, el movimiento de coordinación y el movimiento de la trayectoria. El primer tipo de movimiento provee al usuario una correspondencia de los planos de los ejes x , y y z ; que indican la celda de su localización actual. Con el segundo movimiento, el usuario puede señalar la dirección a la cual desea realizar el traslado y basándose en la trayectoria de los punteros, la imagen se transporta para simular la sensación de volar a través del ambiente.

La velocidad del vuelo se controla por la distancia entre la mano del usuario y su cabeza; así, al estirar el brazo dará lugar a una transición más rápida, en cambio, al ponerla cerca del cuerpo habrá más retardo.

5. SISTEMA DE RED DE LA CAVERNA VIRTUAL

En los capítulos anteriores se presentaron los diferentes procesos que trabajan en conjunto para crear un ambiente en el cual, se realice el análisis de datos remotos y distribuidos, y se proyecten imágenes de forma transparente y real al usuario; quien es el que experimenta la inmersión en la caverna virtual. La existencia de un sistema de comunicaciones está implícita a través de dichos procesos.

En este capítulo se describe la capa de comunicación del sistema de red y la forma en que los tres componentes principales UCV, AIU y DWV; interactúan entre sí.

5.1 Capa de comunicación

Existen dos interfases de aplicación que se utilizan para propósitos de comunicación, los cuales son:

- *SOAP (Protocolo Simple de Acceso de Objetos)*
- *CORBA (Arquitectura Común para Requerimientos de Objetos)*

El *SOAP API* se utiliza sobre *TCP/IP* para proveer accesibilidad a varias aplicaciones basadas en la web. *SOAP* es un protocolo ligero para el intercambio de información en un ambiente descentralizado y distribuido, que se basa en *XML* y que consiste en tres partes: un *encabezado* que describe la estructura del contenido del mensaje y de cómo se procesa; un *conjunto de reglas de codificación* para expresar instancias de tipo de datos en determinada aplicación; y *funciones* para poder realizar llamadas a procedimientos remotos y sus respectivas respuestas.

El *CORBA API* consiste en un conjunto de métodos que proveen transparencia al usuario, para invocar procedimientos remotos y esperar los resultados de retorno en forma de un objeto. Cada cliente y servidor poseen un *ORB (Object Request Broker)* que recibe y procesa distintos requerimientos. *CORBA* proporciona una solución para el desarrollo de aplicaciones distribuidas y heterogéneas. En muchas aplicaciones los datos que se utilizan están distribuidos en varias localidades o sistemas, y es necesario actualizarlos constantemente. Otra razón para distribuir una aplicación es la necesidad de ejecutar procesos independientes y de no sobrecargar de trabajo a los subsistemas. Ambos factores se aplican directamente al sistema de la caverna virtual. Así como el sistema solicita datos distribuidos, el DWV desempeña sus propios procesos para proveer datos transparentes a la UCV.

5.2 Implementación de las interfases CORBA

Una de las principales características del sistema de la caverna virtual es la habilidad de explorar visualmente conjuntos de datos remotos y distribuidos. Dicha habilidad se logra con la implementación de un *CORBA API*. Esta característica también se necesita tanto para la UCV como para la DWV, ya que tienen que tomar en cuenta plataformas de red y de sistema operativo. La manera de visualizar los datos en este modelo es por medio de un objeto que puede ser requerido por un cliente; el servidor devuelve dicho requerimiento en forma de objeto, que contiene la información necesaria.

5.2.1 Capa de comunicación para el servidor

La estrategia de implementación que se utiliza en el desarrollo de la capa de comunicación para el servidor, emplea un sub-constructor de cubo de datos ya que solo un dato se puede visualizar a la vez.

Todos los objetos del servidor son temporales, lo que significa que si el sistema se apaga, los cambios del estado se pierden y los objetos referenciados, vienen a ser inválidos. El CCD a través del sub-constructor, provee el servicio limitado de solo lectura, la cual es suficiente para que las aplicaciones del *Data Mining* no realicen actualizaciones regulares sobre el *Data Warehouse*.

Para extraer datos del *Data Warehouse*, se necesita un *API*, el cual provee acceso de administrador a la base de datos por medio del CCD. El sistema de la caverna virtual debe ser diseñado de tal manera que durante el cambio de un sistema de base de datos a otra, no se necesite reconstruir el sistema completo. El *API* provee operaciones que permiten al constructor, obtener un arreglo de datos y un conjunto de esquemas de jerarquía, para facilitar el nivel de abstracción actual.

Las operaciones del CCD, que también son válidas para el cliente; utilizan una interfase del sub-constructor. Este último se implementa de acuerdo a *IDL* (*interface Description Language*). Cada definición *IDL* se mapea a estructuras de clases en C++. Dichas clases proveen todas las operaciones básicas del sub-constructor en forma virtual; por tal razón el sub-constructor debe proporcionar al menos estas implementaciones para esta clase de operaciones básicas.

La función del servidor principal es inicializar las operaciones del sub-constructor. Dicha función es la responsable de inicializar el *ORB* en tiempo de ejecución, obtener la ruta del manejador del objeto, crear y activar el administrador del sub-constructor. Un manejador de objetos se utiliza para facilitar la comunicación de los objetos *CORBA*. Una vez que este paso de completa, se conecta el sub-constructor con el módulo del CCD para tener acceso a la base de datos.

5.2.2 Capa de comunicación para el cliente

Dentro de la estructura de la capa de comunicación, se encuentra un cliente de la UCV; el cual es más simple que el sub-constructor. Parte de ello es el factor en el cual *CORBA* provee plantillas de clase para objetos conocidos por tipo *_var*. Un objeto de tipo *_var* se crea para cada interfase definida en la especificación IDL. Estos son similares a los punteros de objetos, pero están designados para proporcionar toda la administración de memoria que no utilizan los punteros. Esto significa que un programador puede crear objetos apilados asignándoles su respectivo tipo *_var* y olvidarse de la administración de su ubicación en la memoria dinámica.

El cliente debe obtener una referencia de un objeto remoto, antes de tener la oportunidad de utilizar alguna de las operaciones del sub-constructor. Esto se realiza por medio de la inicialización del *ORB* local, en la cual se traspasa la referencia al cliente, quien después puede utilizar los servidores *ORB* para convertir dicha referencia a una referencia actual de un objeto remoto. La reducción es similar a disminuir de intensidad, pero brinda mayor seguridad ya que un objeto de referencia no puede ser reducido a menos que el objeto referenciado sea una instancia de la clase reducida. Una vez que el constructor obtiene su objeto de referencia ya reducido, el cliente puede solicitar operaciones de la misma manera en que los métodos son llamados por objetos locales. *CORBA* oculta todos los detalles de la ubicación de los objetos y de su implementación.

5.3 Subsistemas de comunicación

Todos los mensajes de comunicación internos del sistema de la caverna virtual entre la UCV y el DWV están formados por documentos *XML*. El DWV esta designado para ser un servidor independiente, capaz de responder a cualquier solicitud de cualquier cliente. El software del servidor se construye por medio de la utilización de una capa *API* para Java, en la cual se interpreta solicitudes de entrada y se codifican para enviarlos desde los módulos del servidor, hacia el CCD. Estos módulos de los servidores, son capaces de proporcionar los servicios por medio de la construcción de un conjunto de objetos que el cliente puede llamar remotamente, y esperar un documento *XML* con la respectiva respuesta. Además, la UCV manipula la inmersión en la realidad virtual por medio de la utilización de un módulo del cliente, para llamar al objeto CCD-Shell, responsable de servicios particulares.

La comunicación internamente del DWV también se implementa como una aplicación cliente/servidor que utiliza interfases *CORBA* Y *SOAP*. Para permitir que el CCD-Shell controle todos los recursos de los datos distribuidos, cada recurso se construye sobre una plataforma de trabajo como un servidor *CORBA* O *SOAP*.

El *CORBA* CCD-Shell y el servidor *SOAP*, proporcionan funciones necesarias para la administración del *Data Warehouse* y las consultas. En una sesión un cliente establece una conexión al DWV y después solicita cierta información iniciando un ciclo de solicitud y respuesta.

5.4 Protocolo de UCV y DWV

La sesión de visualización se inicia cuando el usuario camina dentro de la caverna virtual, con lo cual se indica al CCD que asigna las variables a utilizar. Luego el usuario utiliza un menú para investigar el tipo de datos que se almacenan en el DWV.

El DWV se construye sobre miles de *Data Marts*, que contienen más de un cubo de datos de n-dimensiones, y que se construye sobre un proceso de análisis. Esta información se traslada a la UCV en forma de un documento *XML* que contiene el nombre de cada uno de los cubos y su respectivo tamaño.

Una vez que un cubo de datos de n-dimensiones se escoge, la UCV necesita conocer las especificaciones de ese cubo; esto incluye información tal como el número de dimensiones de datos, el número de cantidades y sus respectivos nombres. Dicha información se recopila por el servidor y se presenta al usuario.

6. APLICACIÓN DE LA CAVERNA VIRTUAL

Hasta este punto se han mencionado los procesos, metodologías, tecnologías y reglas que se deben tomar en cuenta para el funcionamiento de la caverna virtual, y así cumplir con el objetivo de crear un ambiente virtual muy similar al ambiente real.

La aplicación de la caverna virtual puede enfocarse en distintas áreas de estudio y análisis, e incluso para objetivos turísticos y sociales. Este capítulo se enfoca en la descripción de una aplicación en el área biomédica cuyo nombre en inglés es *Crumbs*.

Mediante la inmersión en un ambiente de realidad virtual, como el que se ha descrito hasta este punto; los profesionales de la medicina pueden entender importantes principios fisiológicos y de anatomía; y no solo utilizarlo para fines educativos, si no observar y estudiar algún problema que presente un paciente, y poder dar su diagnóstico al respecto.

Una ventaja de la utilización de la caverna virtual para fines biomédicos, es que es posible navegar alrededor de los órganos o a través de los mismos.

6.1 ¿Qué es *Crumbs*?

Crumbs es una aplicación de la caverna virtual, que facilita la visualización y el trazado de fibras volumétricas que forman parte de estructuras biológicas. Precisamente del término *fibras*, se origina el nombre *Crumbs*, cuyo significado en español es *migajas* y cuya interpretación es que con dicha herramienta se puede desmenuzar un cuerpo para su respectivo estudio. En otras palabras *Crumbs* combina las ventajas de la visualización del volumen de un cuerpo y la inmersión en un ambiente de realidad virtual, con el objetivo de facilitar la identificación y el análisis de estructuras biológicas complejas.

6.1.1 Ventajas de la utilización de la aplicación

Un grave problema en el área médica y biológica con respecto al análisis científico, es el trazado de cuerpos fibrosos en tres dimensiones. Existen métodos tradicionales que involucran el análisis de porciones de fibras en dos dimensiones, y cuyo volumen se construye a partir de la información de esas dos primeras dimensiones, lo cual provoca inexactitud y retardo.

Otros métodos como las imágenes por medio de resonancia magnética, no son eficientes cuando se trata del análisis de una pequeña fibra, en un material con bajo contraste y alta densidad.

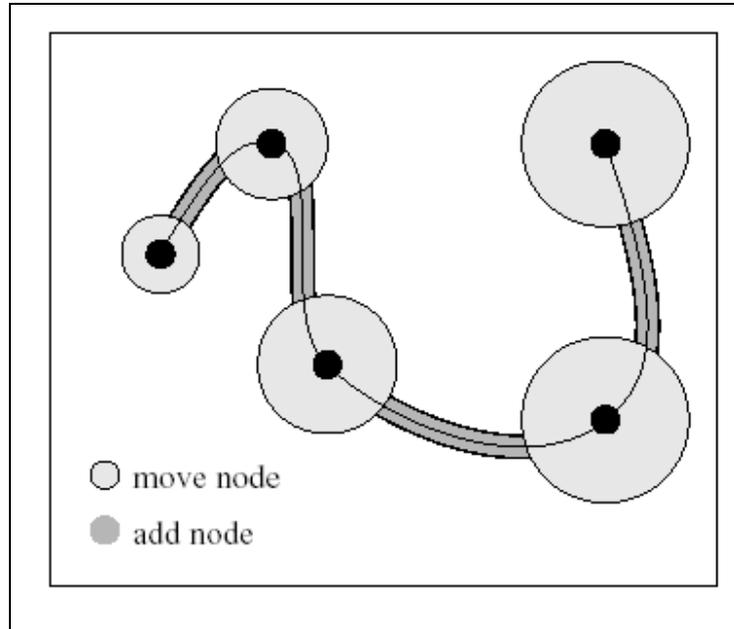
Crumbs brinda una solución a dichos problemas ya que permite:

- Trazar y visualizar la más pequeña fibra de un material determinado por medio de métodos volumétricos.
- Solucionar el problema de la navegación en un ambiente de tres dimensiones.

6.2 Trazado de fibras

Crumbs permite trazar fibras manualmente, en conjuntos de datos volumétricos. El usuario especifica el direccionamiento por medio de la localización del puntero a través de la fibra. *Crumbs* mantiene una lista ordenada de puntos y los une por medio de una línea curva, almacenando el largo de su trayectoria. La alteración del direccionamiento se puede registrar por medio del movimiento, la eliminación o agregación de puntos. La acción que se realiza cuando un botón se presiona depende de la ubicación espacial del puntero relativamente con la existencia de los objetos. Como se muestra en la figura 33, cuando un botón se presiona, se compara la ubicación del puntero con la ubicación de los los puntos existentes y se dibuja una línea uniendo dichos puntos. Si el puntero se encuentra adentro de una partícula, dicha partícula se adhiere al puntero hasta que el botón se suelta. Si el puntero se encuentra cerca de una curva pero no dentro de una partícula, se crea una nueva partícula por medio de la inserción de un punto a la lista, transformándose en un nuevo nodo.

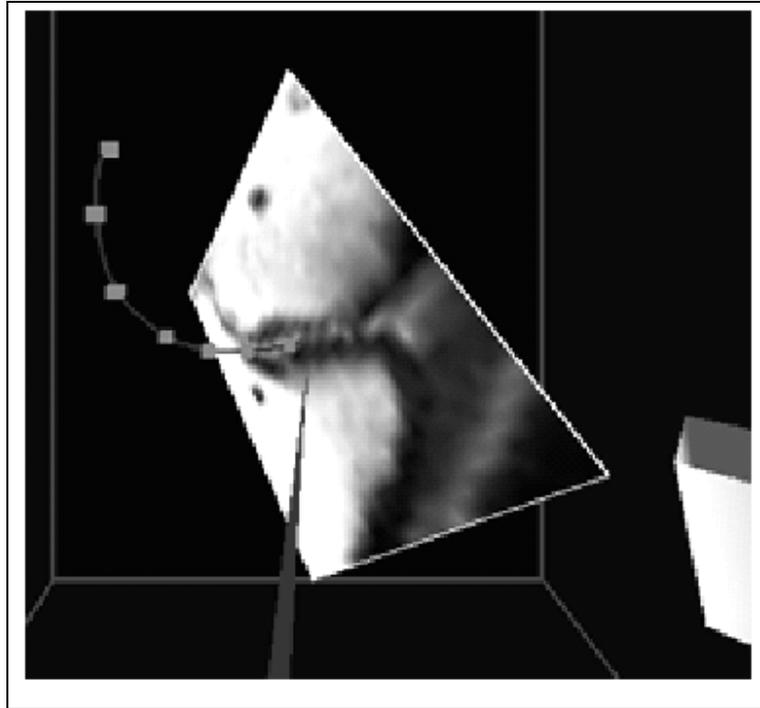
Figura 33. Trazo de fibras



Fuente: Brady, Rachael; y otros. **CRUMBS: una herramienta de ambiente virtual, para la exploración de imágenes biológicas. No. de página: 3**

Crums permite posicionar partículas muy pequeñas con la utilización de un plano especial. Dicho plano posee una resolución ortogonal a una curva determinada, tal como se muestra en la figura 34. La ubicación del plano a lo largo de la curva se determina por la posición del puntero. El plano especial sirve para tres propósitos; primero, cuando la posición de un nuevo punto permite dibujar una nueva dimensión. Segundo, es una herramienta práctica para examinar el direccionamiento de los nodos. Y tercero, el plano especial se utiliza para colocar un nuevo puntero cuando la curva avanza. La lista ordenada de puntos definida por un subconjunto de datos físicos, se almacena para que puedan ser examinados por otros programas o para cargarlos en una sesión siguiente.

Figura 34. Plano ortogonal



Fuente: Brady, Rachael; y otros. CRUMBS: a virtual environment tracking tool for biological imaging. No. de página: 4

6.3 Análisis de la distribución de imagen

Herramientas de visualización han sido convencionalmente designadas como aplicaciones independientes, que operan con datos de archivos que están almacenados en los clientes, mientras la aplicación se ejecuta. *Crumbs* comunica las herramientas independientes por medio de procesos remotos, que podrían ser representados como un dispositivo de almacenamiento de datos, o como un recurso de procesamiento de datos. Todo esto con la finalidad que la inmersión en el mundo virtual sea totalmente transparente al usuario.

6.4 Navegación a través de la caverna virtual con *Crumbs*

Una vez que la estructura fibrosa a trazar se despliega en las paredes de la caverna virtual, aparecen etiquetas marcando su ubicación con respecto del volumen de datos. Para optimizar la vista de una región en particular, el usuario ajusta el punto de vista conforme a su posición y la orientación del volumen de datos.

La habilidad de determinar la ubicación actual del usuario, la orientación y los movimientos, crea el potencial para métodos inteligentes de captura de esta información. *Crumbs* explota dicha información para su sistema de menús y la administración de las partículas.

La aplicación *Crumbs* utiliza el sistema de menús flotantes que permite activar herramientas de visualización con sus parámetros asociados, desarrollar análisis de los datos y especificar las posiciones de las fibras. Esto permite a cada objeto tener su propio menú con sus propios parámetros.

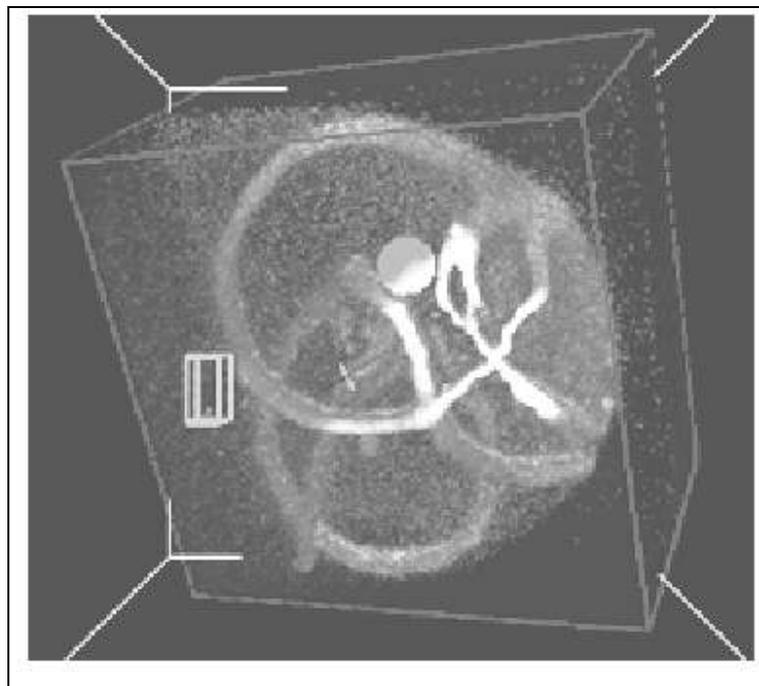
La navegación en la caverna virtual en donde se proyectan imágenes de carácter biológico, es compleja ya que la inmersión es similar a la inmersión en una nube, el usuario no puede ver muy lejos ya que no tiene una comprensión clara sobre hacia donde quiere dirigirse. La navegación en *Crumbs* ha sido modelada de acuerdo al manejo de un carro o avión. También posee la capacidad de traslación por medio del punto en donde se encuentra el usuario, para que al presionar un botón el puntero navegue a través del objeto muy similar al proceso de aceleración de un carro. Para que el usuario pueda orientarse en el ambiente, se ubica un mapa global en una de las esquinas físicas de la caverna virtual.

6.4.1 Rotación

Otra característica de *Crumbs* es la capacidad de rotación por medio del segundo botón del puntero, el cual consiste en el cambio de posición de una de las partes del objeto. En un espacio de tres dimensiones, el centro de rotación es difícil de escoger, los puntos naturales a considerar son la ubicación del puntero y el centro del conjunto de datos.

Por medio de la rotación se puede analizar con mayor detalle un objeto desde cualquier punto de vista.

Figura 35. Huevo fertilizado de drosophila

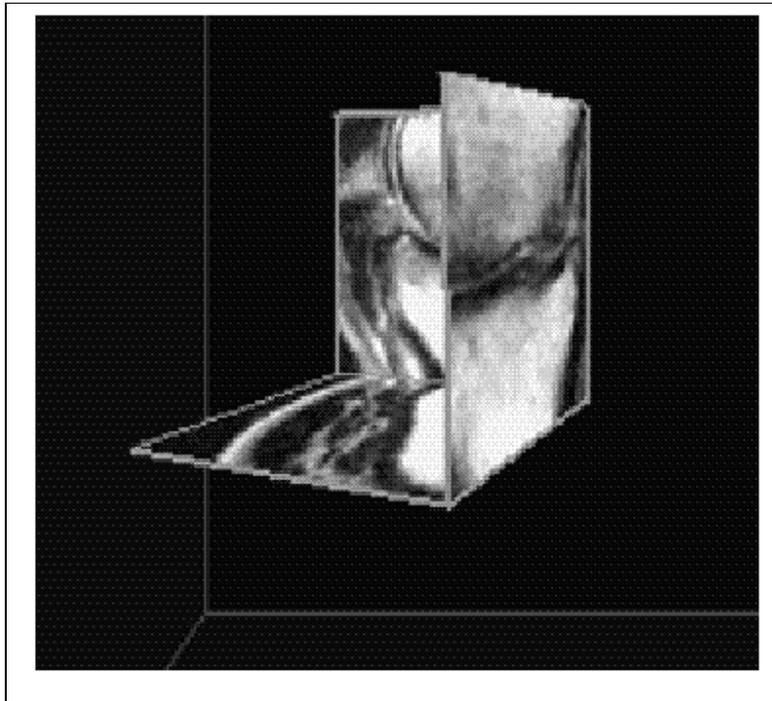


Fuente: Brady, Rachael; y otros. CRUMBS: a virtual environment tracking tool for biological imaging. No. de página: 6

6.4.2 Análisis de un cuerpo por medio de rebanadas

Para poder dividir en rebanadas a un objeto, se utilizan tres planos ubicados ortogonalmente, como se muestra en la figura 36. En los planos se muestra la imagen proyectada y los cortes que se pueden realizar, según su posición; por medio de la selección de cualquiera de los tres planos.

Figura 36. Planos Ortogonales para el corte en rebanadas



Fuente: Brady, Rachael; y otros. CRUMBS: a virtual environment tracking tool for biological imaging. No. de página: 4

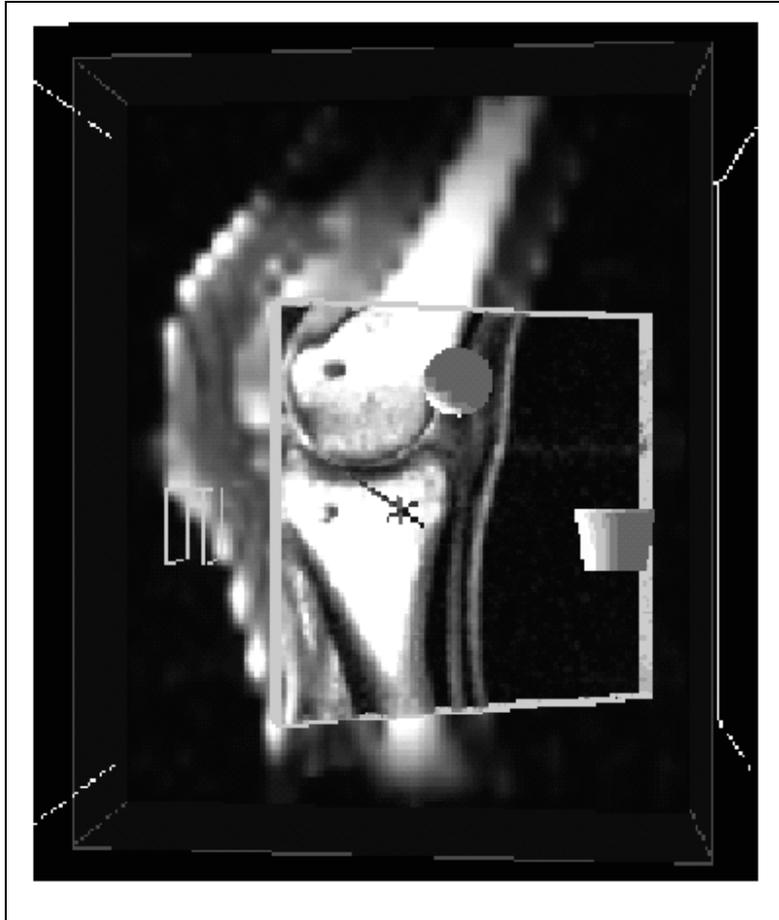
6.5 Aplicación real de *Crumbs*

Hasta el momento se ha construido un marco teórico sobre las características, el funcionamiento y las distintas tecnologías que la caverna virtual utiliza para brindar un ambiente de inmersión muy similar al ambiente de la realidad. Es importante señalar que la caverna virtual junto con la aplicación *Crumbs*, forman un recurso de mucha utilidad para el análisis biológico de los distintos componentes y órganos del cuerpo humano.

La implementación del sistema de la caverna virtual sería un recurso valioso para las ciencias médicas en Guatemala, ya que es una herramienta para el análisis, estudio y aprendizaje de los componentes y órganos del cuerpo humano, por medio de la visualización de imágenes en tres dimensiones y la interacción que la caverna virtual ofrece.

Por su misma naturaleza, la aplicación tiene un amplio campo de acción, ya que podría ser utilizada tanto por un cardiólogo como por un traumatólogo, e inclusive por un ortodoncista. Es decir, que la caverna virtual y la aplicación *Crumbs* contribuiría en el campo de la salud del ser humano.

Figura 37. Imagen de un cartílago obstruido



Fuente: Brady, Rachael; y otros. CRUMBS: a virtual environment tracking tool for biological imaging. No. de página: 7

Como se muestra en la figura 37, en donde se analiza un cartílago dañado de un paciente, de la misma manera se podría proyectar imágenes de cualquier órgano del ser humano, y facilitar su estudio y diagnóstico.

6.6 Otras aplicaciones de la caverna virtual

Debido a la capacidad estereoscópica de la caverna virtual y a su diseño físico; permite a científicos e investigadores interactuar con los datos de una forma jamás antes posible. Como se mencionó al inicio, el contenido de este capítulo se enfoca a la aplicación biomédica de la caverna virtual; sin embargo existen varias áreas en las que se ha utilizado, y que a continuación se mencionan.

6.6.1 Arquitectura y construcción

Existe una aplicación de la caverna virtual que está relacionada con el área de la arquitectura y la construcción. Debido a la forma de inmersión e interacción con el mundo virtual que en la caverna se puede llevar a cabo, es posible visualizar, construir y modificar un proyecto de arquitectura.

Una maqueta a escala de un proyecto residencial, por ejemplo, permite visualizar de diferentes puntos de vista y de forma real, el panorama terminado. Con este recurso, las personas pueden hacerse una idea bastante clara sobre dicho proyecto. Sin embargo como se menciona, está hecho a escala, y por esa razón hay características que pueden escaparse de la percepción humana, tal como la distribución interna de ambientes de una residencia, que aunque hay planos en donde se especifiquen las medidas exactas, una persona solo tendrá una idea de su tamaño.

Es en este punto, en donde la caverna virtual proporciona una ventaja, ya que la persona puede visualizar la residencia, por ejemplo, con las medidas reales. Dicha persona podría caminar por el pasillo de la casa construida con las medidas reales, y verificar que dicho tamaño es el apropiado y que se está aprovechando correctamente el área a construir.

Figura 38. Usuario visualizando un proyecto de construcción



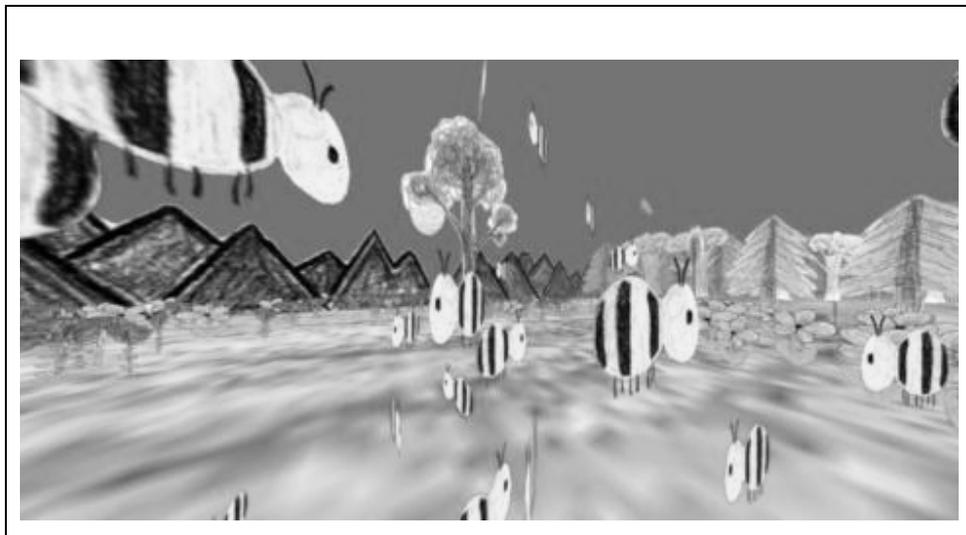
Fuente: [http://www.cs.ualberta.ca/~graphics/vizroom/Visualization Facility \(VizRoom\).htm](http://www.cs.ualberta.ca/~graphics/vizroom/Visualization%20Facility%20(VizRoom).htm)

6.6.2 Crayoland

Crayoland es una de las primeras aplicaciones creada para la navegación a través del espacio virtual de la caverna, orientado específicamente a la diversión y entretenimiento.

El mundo creado por *Crayoland*, comprende un paisaje natural con árboles, montañas, flores, casas, lagos, etc.; dibujados y coloreados a base de crayones de cera. Las imágenes están dibujadas en dos dimensiones, pero proyectadas en un ambiente virtual en tres dimensiones.

Figura 39. Imágenes en dos dimensiones, proyectadas en un ambiente de tres dimensiones



Fuente: <http://www.evl.uic.edu/pape/projects/crayoland/>

El usuario puede interactuar con el mundo virtual cambiando de posición las rocas, cortando las flores y caminando a través del paisaje. *Crayoland* proporciona un conjunto de efectos especiales de sonido y movimiento, ya que se puede escuchar el sonido del lago, el canto de las aves e inclusive se escuchan los pasos del usuario sobre el llano o sobre cualquier superficie. *Crayoland* también construye las imágenes y los sonidos de una serie de insectos como abejas y mariposas, que a su vez reaccionan al percibir la presencia de una persona.

Figura 40. Usuarios interactuando con el paisaje de *Crayoland*

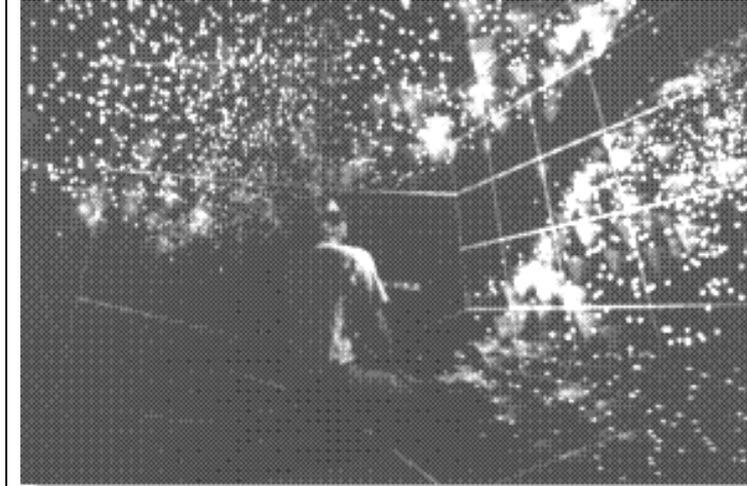


Fuente: <http://www.evl.uic.edu/pape/projects/crayoland/>

6.6.3 Meteorología y astronomía

La caverna virtual también tiene aplicación en las áreas de la meteorología y la astronomía, ya que permite la interacción del usuario con imágenes del espacio exterior y de fenómenos naturales. Dichas imágenes proyectadas en un ambiente de tres dimensiones, permiten la exploración de objetos como meteoros, estrellas, planetas, huracanes o cualquier otro fenómeno. Debido a las propiedades de la caverna virtual, el usuario puede acercarse, alejarse o caminar alrededor del objeto. En la figura 38, se muestra a un usuario visualizando constelaciones en el espacio.

Figura 41. Usuario visualizando una imagen del espacio



Fuente: [http://cave.ncsa.uiuc.edu/Welcome to the NCSA CAVE \(Cave Automatic Virtual Environment\).htm](http://cave.ncsa.uiuc.edu/Welcome%20to%20the%20NCSA%20CAVE%20(Cave%20Automatic%20Virtual%20Environment).htm)

CONCLUSIONES

1. El sistema de exploración de datos, inmerso en un ambiente virtual, utiliza los principales avances de la realidad virtual, teniendo una relación estrecha con la exploración de datos en bases de datos multidimensionales desarrollado en cavernas.
2. La estructura de datos multidimensional, posee las características esenciales para el almacenamiento de datos que servirán en la construcción de un cubo de datos, de acuerdo a la sesión de inmersión virtual.
3. La estructura de un sistema distribuido que posee la caverna virtual, contribuye a la transparencia del funcionamiento del sistema virtual, ante el usuario.
4. La continua obtención de las posiciones, tanto de la cabeza como de la mano del usuario, permiten la actualización y coordinación de movimientos con los objetos virtuales en las pantallas de la caverna.
5. De acuerdo con las características, funcionalidades y recursos que presenta, la caverna virtual puede ser utilizada para análisis científico de datos; tomando en cuenta que no sólo permite la visualización de objetos en tres dimensiones, sino también, la interacción con ellos; a tal grado de poder dividirlo en rodajas y examinar cada una de las partes.

RECOMENDACIONES

1. Al realizar una inmersión en el ambiente virtual, es necesario considerar lo siguiente:
 - Colocación adecuada tanto del puntero, como de los lentes; ya que de ello depende la sincronía entre los movimientos del usuario con las imágenes proyectadas en las pantallas.
 - El usuario debe tener una inducción previa, de la forma en que debe interactuar con el sistema.

2. En el proceso de almacenamiento de datos en las estructuras multidimensionales, se debe tomar en cuenta lo siguiente:
 - Definir los datos que representarán las dimensiones de los objetos, para que sea un factor que esté implícito, en la construcción del cubo de datos.

BIBLIOGRAFÍA

1. Amoura, Ayman. **DIVE-ON: De la base de datos a la realidad virtual**
Agosto 2001. <http://www.acm.org/crossroads/espanol/xrds7-3/diveon.html>
2. Amoura, Ayman. **Publications**. Agosto 2001
<http://www.cs.ualberta.ca/~ayman/publications.htm>
3. Amoura, Ayman. **Towards a Novel OLAP Interface for Distributed Data Warehouse**. Octubre 2001.
<http://searchwebservices.techtarget.com/sDefinition.htm>
4. Amoura, Ayman. **Walking the Data Cube: An Inmersed Virtual Environment for Visualizing Data Mining Results**. Octubre 2001
<http://www.uvirtual.ean.edu.co/uvirtual/sections.htm>
5. Baker, Pauline. **Human Factors in Virtual Environments for the Visual Analysis of Scientific Data**. Septiembre 2001.
<http://archive.ncsa.uiuc.edu/Vis/Publications/humanFactors.html>
6. Basgall, Monte. **Rachael Brady Explores New Realm of Scientific Analysis**. Abril 2002. <http://www.egr.duke.edu/News/Stories/38.html>
7. Brady, Ráchale, y otros. **Crumbs: a virtual environment tracking for Imagine**. Abril 2002.
8. Criner, Amanda. **Climbing inside the cave**. Agosto 2003.
<http://www.dailyillini.com/aug00/aug25/news/feature.html>
9. **Data mining & Analysis**. Marzo 2002. <http://www.domeyer.com/art3.html>

10. **Data Warehousing & OLAP.** Agosto 2001.
<http://gaston.snu.ac.kr/~swlee/dw-dm.html>
11. Kalev Leetaru. NCSA, Visualization and Virtual Enviroments Group. **Easy Demo Manual.** Febrero 2002. <http://easydemo.ncsa.uiuc.edu>
12. NCSA. **About The CAVE.** Marzo 2002. <http://cave.ncsa.uiuc.edu>
13. **OLAP, ROLAP, MOLAP.** Febrero 2002.
<http://www.map.es/csi/silice/DW2251.html>
14. Pape, Dave. **The CAVE Virtual Reality System.** Agosto 2001
<http://www.evl.uic.edu/pape/CAVE/>
15. Pape, Dave. **Crayoland.** Agosto 2003
<http://www.evl.uic.edu/pape/projects/crayoland/>
16. Porras, Raesfeld. **Aplicación de mapas cognitivos difusos a VRML en un ambiente virtual.** Agosto 2001.
http://mailweb.udlap.mx/~tesis/vonraesfeld_p_pa/capitulo1.html
17. **Realidad Virtual.** Marzo 2002.
http://www2.gratisweb.com/lhorena_ortiz/pag-virtual.htm
18. **Relational OLAP.** Marzo 2002.
<http://www.logosphere.com/OLAP/ROLAP/rolap.html>
19. **The CAVE.** Agosto 2001.
http://www.indiana.edu/~ids/archives/tech_cave.html
20. **The CAVE at NCSA.** Agosto 2002
[http://cave.ncsa.uiuc.edu/Welcome to the NCSA CAVE \(Cave Automatic Virtual Environment\).htm](http://cave.ncsa.uiuc.edu/Welcome%20to%20the%20NCSA%20CAVE%20(Cave%20Automatic%20Virtual%20Environment).htm)