



AVANCES EN SISTEMAS INTERACTIVOS PARA PERSONAS CON PARÁLISIS CEREBRAL

César Mauri Loba

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

Cesar MAURI LOBA

AVANCES EN SISTEMAS
INTERACTIVOS PARA PERSONAS
CON PARÁLISIS CEREBRAL

TESIS DOCTORAL

dirigida por el Dr. Agusti SOLANAS

Departament d'Enginyeria Informàtica i Matemàtiques



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2014



DEPARTAMENT D'ENGINYERIA INFORMÀTICA
I MATEMÀTIQUES

Av. Països Catalans, 26
43007 Tarragona
Tel. 34 977 559 703
Fax 34 977 559 710
secdeim@urv.cat

Agustí Solanas Gómez, profesor Agregado del Departamento de Ingeniería Informática y Matemáticas de la Universidad Rovira i Virgili,

HAGO CONSTAR:

Que el presente trabajo titulado "*Avances en Sistemas Interactivos para Personas con Parálisis Cerebral*", presentado por César Mauri Loba para la obtención del título de Doctor, ha sido realizado bajo mi supervisión en el Departamento de Ingeniería Informática y Matemáticas de esta universidad, y que cumple con todos los requisitos.

Tarragona, 9 de Mayo de 2014

Agradecimientos

Al Dr. Agustí Solanas, director de esta tesis, por toda su ayuda y sabios consejos.

Al Dr. Toni Granollers por toda su ayuda y colaboración.

A los usuarios y profesionales de la Associació Provincial de Parálisis Cerebral (APPC) de Tarragona sin la colaboración de los cuales esta tesis no hubiera sido posible.

A mi familia, por su apoyo incondicional a lo largo de este trayecto.

Por último, dar las gracias a mis amigos y compañeros de trabajo que, de un modo u otro, han respaldado este esfuerzo.

Resumen

La Parálisis Cerebral es la discapacidad motora más común durante la infancia y consiste en una alteración de la movilidad debida a una lesión en el sistema nervioso central producida durante la etapa prenatal, perinatal o infantil. Estas lesiones, que no son una enfermedad, no evolucionan, pero sus consecuencias sí que interfieren en el desarrollo de la persona. Las dificultades motoras, que suelen ir acompañadas en muchas ocasiones por problemas sensoriales, alteraciones en la atención, cognición y otros, suponen un hándicap para la exploración del entorno, la participación en las actividades cotidianas y las relaciones con los iguales, que son la base del aprendizaje y el desarrollo.

Si consideramos que las principales limitaciones de las personas con parálisis cerebral se encuentran en el desplazamiento, la manipulación, la comunicación y el acceso a la información, podemos comprender el papel que las tecnologías de la información y la comunicación pueden desempeñar para la mejora de la calidad de vida del colectivo.

Esta tesis, que se encuadra dentro de la disciplina de la Interacción Persona-Ordenador (IPO), se centra en la concepción, aplicación y evaluación de sistemas interactivos para personas con parálisis cerebral. En primer lugar hacemos una revisión del estado del arte en sistemas interactivos audiovisuales para personas con parálisis cerebral. A continuación realizamos una propuesta de distintos métodos para el análisis de la calidad de la interacción. Después abordamos varias propuestas de sistemas interactivos basados en visión artificial diseñados para personas con parálisis cerebral. De entre éstas, hacemos especial énfasis en una propuesta de sistema interactivo audiovisual orientado a personas con trastornos físicos, cognitivos y sensoriales en grados moderado y severo que ha demostrado ser asequible para muchas de estas personas y les proporciona una experiencia lúdica motivadora, que facilita el trabajo de ciertas competencias por parte de los profesionales.

Se trata de una investigación aplicada que, además de resultados científicos, también ha buscado los prácticos y que ha contado con la estrecha colaboración de los usuarios y de los profesionales de una entidad de atención a personas con parálisis cerebral. Como parte del trabajo realizado se han desarrollado distintas herramientas que se han puesto a disposición de la comunidad en forma de software libre.

Resum

La Paràlisi Cerebral és la discapacitat motora més comuna durant la infància i consisteix en una alteració de la mobilitat deguda a una lesió en el sistema nerviós central produïda durant l'etapa prenatal, perinatal o infantil. Aquestes lesions, que no són una malaltia, no evolucionen, però les seves conseqüències sí que interfereixen en el desenvolupament de la persona. Les dificultats motores, que solen anar acompanyades en moltes ocasions per problemes sensorials, alteracions en l'atenció, cognició i altres, suposen un handicap per a l'exploració de l'entorn, la participació en les activitats quotidianes i les relacions amb els iguals, que són la base de l'aprenentatge i el desenvolupament.

Si considerem que les principals limitacions de les persones amb paràlisi cerebral es troben en el desplaçament, la manipulació, la comunicació i l'accés a la informació, podem comprendre el paper que les tecnologies de la informació i la comunicació poden exercir en la millora de la qualitat de vida del col·lectiu.

Aquesta tesi, que s'enquadra dins de la disciplina de la Interacció Persona-Ordinador (IPO), se centra en la concepció, aplicació i avaluació de sistemes interactius per a persones amb paràlisi cerebral. En primer lloc fem una revisió de l'estat de l'art en sistemes interactius audiovisuals per a persones amb paràlisi cerebral. A continuació realitzem una proposta de diferents mètodes per a l'anàlisi de la qualitat de la interacció. Després abordem diverses propostes de sistemes interactius basats en visió artificial dissenyats per a persones amb paràlisi cerebral. D'entre aquestes, fem especial èmfasi en una proposta de sistema interactiu audiovisual orientat a persones amb trastorns físics, cognitius i sensorials en graus moderat i sever que ha demostrat ser assequible per a moltes d'aquestes persones i els proporciona una experiència lúdica motivadora, que facilita el treball de certes competències per part dels professionals.

Es tracta d'una investigació aplicada que, a més de resultats científics, també ha buscat els pràctics i que ha comptat amb l'estreta col·laboració dels usuaris i dels professionals d'una entitat d'atenció a persones amb paràlisi cerebral. Com a part del treball realitzat s'han desenvolupat diferents eines que s'han posat a disposició de la comunitat en forma de programari lliure.

Abstract

Cerebral Palsy is the most common motor disability in childhood and entails impaired mobility due to damage to the central nervous system just before or during the birth, or during the early childhood. These lesions, which are not a disease, are not progressive, but its consequences interfere with the development of the child. Motor impairments often come with other difficulties such as alterations in attention, cognition, and others. Such deficiencies are a handicap for the exploration of the environment, for the participation in daily activities and for the relationship with peers which are the basis for learning and development.

If we consider that the main limitations of people with cerebral palsy are in displacement, manipulation, communication and access to information, we can understand the role that information and communication technologies can play in improving the quality of life of this group of people.

This thesis, which falls within the Human-Computer Interaction (HCI) scope, focuses on the design, implementation and evaluation of interactive systems for people with cerebral palsy. Firstly, we review the state of the art in audio-visual interactive systems for people with cerebral palsy. Secondly, we propose different methods for analysing the quality of the interaction. Finally, we propose several interactive systems for people with cerebral palsy based on computer vision, among which we put special emphasis on an interactive audio-visual proposal aimed at people with physical, cognitive and sensory impairments in moderate and severe degrees. This system proved to be attainable for many of these people and to provide a motivating playing experience that allows professionals to develop some skills on the users.

This is an applied research which, in addition to scientific results, also has led to practical outcomes and has been carried out in close collaboration with users and professionals of a cerebral palsy care centre. As part of this work, several free and open source tools have been developed and released to the community.

Índice general

1. Introducción	1
1.1. Parálisis cerebral	2
1.2. Motivaciones y objetivos	4
1.3. Publicaciones	7
1.4. Premios	8
1.5. Organización de la tesis	9
2. Sistemas interactivos audiovisuales	11
2.1. Clasificación	12
2.2. Principales propuestas	13
2.2.1. El sistema Soundbeam	13
2.2.2. Soundscapes	15
2.2.3. Proyecto MEDiate	18
2.2.4. Otros proyectos	22
2.3. Sistemas musicales interactivos	25
2.3.1. Breve reseña histórica	26
2.3.2. Definición	27
2.3.3. Elementos de los SMI	28
2.4. Dispositivos de entrada	29
2.4.1. Sensor de ultrasonidos: Soundbeam	30
2.4.2. Sensores de infrarrojos	31
2.4.3. Análisis de sonido y voz	32
2.4.4. Visión artificial	32
2.5. Material audiovisual y correspondencia	35
2.5.1. Correspondencia en las propuestas estudiadas	36
2.5.2. Contenido	39
2.5.3. Contenido visual	40
2.5.4. Contenido sonoro	41
2.5.5. Contenido táctil	42
2.6. Efectos terapéuticos	43
2.6.1. Efectos observados	44
2.6.2. Desarrollo de las sesiones experimentales	47
2.6.3. Técnicas de análisis	49

3. Nuestras propuestas de técnicas y sistemas interactivos	57
3.1. Evaluación de la interacción	58
3.1.1. Interacción basada en conmutadores	59
3.1.2. Producción de texto	63
3.1.3. Control del dispositivo apuntador	63
3.1.4. Rendimiento de la interacción	65
3.1.5. Método de evaluación propuesto	66
3.2. Ratón Facial y Enable Viacam	73
3.2.1. Funcionamiento	73
3.2.2. Estado actual del proyecto	74
3.3. SITPLUS	75
3.3.1. Objetivos y criterios de diseño	75
3.3.2. Arquitectura del sistema	76
3.3.3. Módulos de visión artificial	82
3.3.4. Módulos de audio	82
3.3.5. Controles remotos Wii	83
3.3.6. Salida gráfica	86
3.3.7. Actividades	89
3.3.8. Estado actual del proyecto	92
3.4. Switch Viacam	93
3.5. Herramienta on-line de anotación de vídeo	95
3.5.1. Motivaciones	95
3.5.2. Requisitos	96
3.5.3. Arquitectura	97
3.5.4. Interfaz de usuario	98
3.5.5. Integración con ANVIL	100
4. Resultados	101
4.1. Resultados evaluación de la interacción	102
4.1.1. Caso 1: A.A.	103
4.1.2. Caso 2: C.G.	105
4.2. Resultados proyecto Switch Viacam	107
4.2.1. Caso 1: Y.P.	108
4.2.2. Caso 2: Q.H.	108
4.2.3. Ventajas	108
4.3. Resultados proyecto SITPLUS	109
4.3.1. Entorno de experimentación	109
4.3.2. Procedimiento	110
4.3.3. Resultados cuantitativos	111
4.3.4. Resultados cualitativos	113

Índice general	XI
<hr/>	
5. Conclusiones	125
5.1. Conclusiones	125
5.2. Contribuciones	127
5.3. Trabajo futuro	128
Bibliografía	131

CAPÍTULO 1

Introducción

En este capítulo se presentan el enfoque, las motivaciones y el contexto de esta tesis, así como los retos que abordamos a lo largo de este trabajo. Además, se describen brevemente las soluciones que proponemos para abordar esas cuestiones y se incluye una recopilación de las principales publicaciones y premios conseguidos. Finalmente se exponen la estructura y la organización del documento.

Índice

1.1. Parálisis cerebral	2
1.2. Motivaciones y objetivos	4
1.3. Publicaciones	7
1.4. Premios	8
1.5. Organización de la tesis	9

Esta tesis se centra en la concepción, aplicación y evaluación de sistemas interactivos para personas con parálisis cerebral. Concretamente, en la generación de herramientas que permitan evaluar dicha interacción y, especialmente, en la concepción y aplicación de nuevos sistemas interactivos, principalmente aquellos basados en visión artificial.

La investigación de este tipo de sistemas se encuadra dentro de la disciplina de la Interacción Persona-Ordenador (IPO) atendiendo a la definición dada por el *Special Interest Group in Computer Human Interaction (SIGCHI)* de la *Association for Computer Machinery (ACM)*:

« Es la disciplina relacionada con el diseño, evaluación e implementación de sistemas informáticos interactivos para el uso de seres humanos, y con el estudio de los fenómenos más importantes con los que está relacionada. »

Puesto que se trata de una investigación aplicada, todo el trabajo realizado se ha hecho contando con la estrecha colaboración de los usuarios y de los profesionales de una entidad de atención a personas con parálisis cerebral.

En nuestro caso, la Asociación Provincial de Parálisis Cerebral (APPC) de Tarragona.

Especialmente se ha centrado el interés en aquellas personas con parálisis cerebral y etiologías afines en grados de afectación moderado y severo. Es decir, personas con graves deficiencias intelectuales y cognitivas unidas a importantes limitaciones sensoriales, de movilidad y de comunicación. Ello obedece a que éste es el subgrupo más numeroso en la entidad antes citada, además de ser el colectivo para el cual, según nuestra experiencia, existe un menor número de recursos disponible.

Además de la investigación que se aborda en esta tesis, se ha propiciado en todo momento que los esfuerzos realizados tuvieran aplicación en el contexto real, es decir, en un centro de atención donde los profesionales y usuarios no son expertos en informática tomando, por ello, especial relevancia el diseño centrado en el usuario y las pruebas reales llevadas a cabo *in-situ*. Además, las características de los usuarios finales hacen que haya que establecer planteamientos realistas en cuanto al tipo y complejidad de las actividades que pueden desarrollar.

En muchos casos el diseño de estos sistemas debe tener en cuenta el uso tanto por parte del profesional, que actúa como facilitador de la tecnología, como del usuario final. Además, las especiales características de estos últimos invalidan en muchos casos el uso de dispositivos estándar para la interacción (*ej.* teclado, ratón y pantalla táctil) y la aplicación de la mayoría de métodos comunes de evaluación de la Interacción Persona-Ordenador (IPO) como por ejemplo pensar en voz alta, entrevistas, cuestionarios, etc.

El trabajo realizado incluye un importante esfuerzo de desarrollo de distintas herramientas software que, publicadas bajo licencias de código abierto, se usan de forma regular en ésta y otras instituciones similares, además de por numerosos usuarios en todo el mundo.

1.1. Parálisis cerebral

La parálisis cerebral es la discapacidad motora más común durante la infancia [89]. En todo el mundo se estima una prevalencia de entre el 1,5 y más del 4 por 1.000 de los niños nacidos con vida [4, 7, 98, 61].

Según Confederación ASPACE¹ la definición de parálisis cerebral (en adelante PC) es:

« La Parálisis Cerebral es un trastorno global de la persona consistente en un desorden permanente (irreversible y persistente a lo largo de toda

¹Confederación Española de Federaciones y Asociaciones de Atención a las Personas con Parálisis Cerebral y Afines. <http://www.aspace.org/>

la vida) y no inmutable (no quiere decir que las consecuencias no cambien involutiva o evolutivamente) del tono, la postura y el movimiento (trastorno neuromotor), debido a una lesión no progresiva (no aumenta ni disminuye, es decir, no es un trastorno degenerativo) en el cerebro antes de que su desarrollo y crecimiento sean completos (se produce en un período de tiempo en el cual el sistema nervioso central está en plena maduración). Esta lesión puede suceder durante la gestación, el parto o durante los primeros años de vida, y puede deberse a diferentes causas, como una infección intrauterina, malformaciones cerebrales, nacimiento prematuro, asistencia incorrecta en el parto, etc.

Puede generar la alteración de otras funciones superiores (atención, percepción, memoria, lenguaje y razonamiento) en función del tipo, localización, amplitud y disfunción de la lesión neurológica y el nivel de maduración anatómico en que se encuentra el encéfalo cuando esta lesión se produce, e interferir en el desarrollo del Sistema Nervioso Central (va a repercutir en el proceso madurativo del cerebro y por lo tanto en el desarrollo del niño).

A estos problemas se pueden asociar otros de diversa índole y no menos importantes. Se trata de problemas clínicos, sensoriales, perceptivos y de comunicación. Esto hace que exista una enorme variedad de situaciones personales, no generalizables, que dependen del tipo, localización, amplitud y difusión de la lesión neurológica. Así, en algunas personas la PC es apenas apreciable, mientras que otras pueden estar muy afectadas y necesitar de terceras personas para su vida diaria.

Las bases esenciales utilizadas para el diagnóstico de la PC son cinco, e incluyen desde la localización de la lesión en los tejidos nerviosos de la cavidad craneal, la comprobación de la lesión permanente y no progresiva, la detección de la lesión precoz (durante la maduración cerebral), el cuadro clínico dominado por un trastorno motor (de postura, movimiento y tono muscular), y la variabilidad del cuadro clínico a lo largo de la vida.

La PC no se puede curar, pero si la persona afectada recibe una atención adecuada que le ayude a mejorar sus movimientos, que le estimule su desarrollo intelectual, que le permita desarrollar el mejor nivel de comunicación posible y que estimule su relación social, podrá llevar una vida plena. »

Es decir, el término Parálisis Cerebral no denota de forma inequívoca un trastorno, sino que más bien es un paraguas bajo el cual se aglutinan múltiples perfiles de personas con discapacidad. De hecho, en España es bastante común que a los centros de ASPACE acudan personas que, a pesar de que no han sido diagnosticadas de parálisis cerebral, presentan una sintomatología similar

(de ahí el término “afines” en las siglas de ASPACE).

Por todo ello se trata de un colectivo muy heterogéneo, donde son múltiples las dimensiones que caracterizan a un individuo concreto. De ahí que normalmente se hable de atención personalizada ya que, normalmente, no se pueden equiparar dos personas con parálisis cerebral.

1.2. Motivaciones y objetivos

Nuestra relación con el mundo de la discapacidad empieza en el momento de presentar el Proyecto Final de Carrera [87] de los estudios de Ingeniería Informática. Se trataba de un sistema que, mediante los movimientos de la cabeza, permitía manejar el puntero de ratón. Enseguida se vio el potencial que tenía como herramienta de acceso alternativo al ordenador para personas con discapacidad. Año y medio más tarde se fundó la empresa CREA Sistemes Informàtics² para comercializar una evolución de dicho proyecto con el nombre de Ratón Facial [88] y otros sistemas de visión artificial para el acceso alternativo al ordenador como WebColor Detector³. Con el tiempo, el proyecto Ratón Facial derivó en una versión de código abierto del mismo bajo el nombre de Enable Viacam [85].

En 2004 la Asociación Provincial de Parálisis Cerebral de Tarragona (APPC) se interesó en estos proyectos y se inició una colaboración para el desarrollo de un programa de I+D que continua actualmente en vigor. Durante los tres primeros años se llevaron a cabo muchas actividades para mejorar estos sistemas y fomentar el uso de las nuevas tecnologías en general entre el colectivo de usuarios del centro. Parte de esta labor consistía en evaluar cuales eran las herramientas y los productos de apoyo más indicados para cada caso individual. Estas valoraciones normalmente se realizaban a partir del ensayo de distintos productos de apoyo con los usuarios, de sus propias opiniones (siempre que fuera posible) y, finalmente, de la intuición y criterio subjetivo de los profesionales. Si bien en la mayoría de ocasiones esta estrategia era la más adecuada, aun así pensábamos que resultaría útil disponer de información cuantitativa que complementase esta valoración. Por ello propusimos nuestro propio sistema de evaluación pensando en las particularidades del colectivo de parálisis cerebral.

Sin embargo, a lo largo de este tiempo y a pesar de los esfuerzos antes relatados, observábamos que sólo una pequeña parte de los usuarios se podían beneficiar directamente del uso de estas tecnologías, principalmente por sus limitaciones cognitivas, ya que se da la circunstancia de que la mayor parte de

²<http://www.crea-si.com>

³<http://www.crea-si.com/esp/webcolor.php>

usuarios de esta entidad tienen limitaciones moderadas o severas, especialmente entre la población infantil. Es decir, en muchos casos no comprendían como usar las más sencillas aplicaciones lúdicas o educativas que les proporcionaban sus profesores y terapeutas a pesar de tener acceso físico a ellas.

A la experiencia anterior, basada en el trabajo directo con los usuarios del centro, se unió el conocimiento de la musicoterapia. La terapia musical o musicoterapia es una disciplina que consiste en el uso de la música dentro de una estructura de trabajo organizada con fines terapéuticos. La musicoterapia apareció en los años 50 y es una profesión reconocida en muchos países (en España todavía no). En la APPC se imparten de forma regular sesiones de musicoterapia por varios músicos especialistas. Profesionales del centro reconocen los efectos beneficiosos de estas sesiones, especialmente con respecto a las habilidades comunicativas (muchas veces no verbales), psicomotrices y de autoestima.

La aplicación de la musicoterapia se basa en la realización de sesiones dirigidas por el musicoterapeuta y que, a través de la utilización de instrumentos acústicos, permiten obtener beneficios para la salud de la(s) persona(s) a quien(es) va dirigida la terapia. Cuanto más severas son las discapacidades físicas o cognitivas de los participantes, más grande es la responsabilidad del profesional respecto a la empatía e interpretación. En los casos en los que el usuario se ve más limitado, por ejemplo si no puede articular sonidos con la boca ni tocar un instrumento, menos participación activa puede tener y, en un caso extremo, llega a resultar una experiencia totalmente pasiva (musicoterapia receptiva).

Estas experiencias motivaron la realización de pruebas combinando el uso de Ratón Facial con el sistema DanceMusic⁴ (éste último genera sonidos MIDI y colores en la pantalla al mover el ratón). Un usuario del centro, que no participaba activamente en las sesiones de musicoterapia, lo probó y pudimos observar inmediatamente la motivación que sentía al interactuar con el sistema. Otras pruebas con niños que no eran capaces de usar el ordenador, ahora respondían ante este sistema. Estas evidencias unidas al conocimiento de otros proyectos ya existentes (que se estudian más adelante) motivaron el abordar la idea como proyecto de investigación y desarrollo.

El objetivo era crear un sistema interactivo, preferentemente sin contacto, que proporcionase una salida sonora y/o visual relacionada con el gesto del usuario. Gesto (voluntario) que, con tecnologías como la visión artificial, puede extraerse de casi cualquier parte del cuerpo y por pequeño que sea; lo que lo hace especialmente apropiado para personas con limitaciones físicas severas.

Estas ideas se materializan en el proyecto SITPLUS (Sistema InTeractivo

⁴<http://www.xtec.cat/dnee/udc/>

para Personas con pLUridiScapacidad)⁵ que consiste en un sistema capaz de analizar a distancia las acciones del usuario (en nuestro caso movimientos físicos, mediante visión artificial y otros sensores de movimiento, y emisiones orales) para transformarlo después en imágenes y sonido acordes con éstas.

Esto permite crear un espacio sensible donde existe una libertad gestual configurable, que se adapta y se caracteriza por la ausencia de constantes inamovibles⁶ y, en general, no hay reglas a obedecer. Creemos que la generación de sonido, música e imágenes basada en el gesto, proporciona una actividad altamente interactiva y asequible a la gran mayoría de usuarios.

Se trata pues de construir un sistema que pueda servir como instrumento musical accesible, que permita la libre exploración y expresión creativa (a través de la percepción y la manipulación del sonido y la imagen), que permita a un/a chico/a jugar y divertirse o que, en casos donde existen limitaciones cognitivas más severas, permita que las acciones sucedan por uno mismo; la experiencia vital del “yo soy!”, que cuando el usuario realiza algún movimiento, algo pasa. Se espera que, a través de esta actividad, sea posible –a veces de forma inconsciente– desarrollar distintas capacidades como la participación, la comunicación o la creatividad, y contribuya, por tanto, a la mejora en la calidad de vida.

Además del reto que supone la construcción de un sistema de estas características, se une la voluntad de conocer la utilidad como herramienta de trabajo para los profesionales que atienden a personas con parálisis cerebral, así como el impacto que tiene su uso en el usuario final. Para ello, esta investigación se aborda desde un enfoque interdisciplinar, por lo que en la APPC se ha formado un equipo permanente compuesto de ingenieros y psicopedagogos contando con el asesoramiento puntual de otros especialistas como logopedas, terapeutas ocupaciones, fisioterapeutas, artistas sonoros y musicoterapeutas.

Por último, también se marca como objetivo el que este tipo de tecnología sea económica. Gracias a la potencia de los ordenadores actuales, el abaratamiento de dispositivos de visualización de gran tamaño (pantallas LCD o proyectores) y a periféricos de bajo coste como webcams, micrófonos y mandos de juego, es posible construir sistemas virtuales interactivos enteramente basados en software. Ello, permite reducir costes y la complejidad necesaria respecto a una instalación audiovisual interactiva profesional.

⁵Este proyecto se desarrolla, investiga y aplica de forma sistemática en el laboratorio de la Asociación Provincial de Parálisis Cerebral de Tarragona (APPC).

⁶Por ejemplo un violín, un piano, etc. son instrumentos donde su construcción física impone una gestualidad y un continuo energético único y fijo entre gesto y resultado perceptivo [27].

1.3. Publicaciones

Revistas

- Mauri, C.; Solanas, A. y Granollers, T. “A Nonformal Interactive Therapeutic Multisensory Environment for People With Cerebral Palsy”, en *International Journal of Human-Computer Interaction*, volumen 28, número 3, pp. 202-212, 2012. DOI:10.1080/10447318.2011.581894
- Mauri, C.; Granollers, T.; Lores, J. y García, M. “Computer vision interaction for people with severe movement restrictions.”, en *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, número 2, pp. 38-54, Agora Center, University of Jyväskylä, Finlandia, 2004.

Congresos internacionales

- Mauri, C.; Solanas, A.; Granollers, T.; Bagés, J. y García, M. “Interactive Therapeutic Multi-sensory Environment for Cerebral Palsy People”, en *INTERACT 2009*, parte II, LNCS 5727, pp. 696-699, Uppsala, Suecia, 2009. ISBN:978-3-642-03657-6
- Ponsa, P.; Mauri, C.; Amante, B. y Díaz, M. “An Experimental Study on Computer Vision Interaction in the Use of Graphical Display in Home Systems”, en *The 21st Symposium on Human Factors in Telecommunication (HFT 2008)*, Kuala Lumpur, Malasia, 2008.
- Mauri, C.; Granollers, T. y Solanas, A. “On the Assessment of the Interaction Quality of Users with Cerebral Palsy”, en *2nd Intl. Conf. on Availability, Reliability and SEcurity (ARES)*, pp. 799-805, Viena, Austria, 2007. ISBN:0-7696-2775-2

Capítulos de libro

- Mauri, C.; Solanas, A. y Granollers, T. “Towards the Introduction of Accessibility Resources in Public Telecentres”, en *Engineering the User Interface: From Research to Practice*, Springer, EEUU, 2008. ISBN:978-1-84800-135-0

Congresos nacionales

- Mauri, C., Solanas, A., Garcia, M., Granollers, T., Parés, N., y Bagés, J. “SATI: Sistema Audiovisual Terapèutic Interactiu”, en *TECNODISCAT. Workshop de R+D+I en tecnologies per a l'autonomia de les persones amb discapacitat*, pp. 94-100, Barcelona, 2009. ISBN:978-84-7653-506-6

- Mauri, C.; Solanas, A.; Granollers, T.; Parés, N.; Bagés, J.; y García, M. “Entorno interactivo multimodal para personas con parálisis cerebral”, en *IV Jornadas Iberoamericanas de Tecnologías de Apoyo a la Discapacidad - LAS TECNOLOGÍAS DE APOYO EN PARÁLISIS CEREBRAL*, Madrid, Octubre 2009.
- Mauri, C.; García, M. and Bagés, J. “Terapia musical y multisensorial mediante las nuevas tecnologías: proyecto SATI (Sistema Audiovisual Terapéutico Interactivo)”, en *II Jornades de Paràlisi Cerebral*, APPC, Tarragona, Noviembre 2008.
- Mauri, C.; Granollers, T.; Lorés, J. y Solanas, A. “La Accesibilidad en Telecentros Públicos: Una Experiencia Real”, en *VII INTERACCION 2006. Congreso Internacional Interacción persona Ordenador*, pp. 177-185, Puertollano (Ciudad Real), 2006. ISBN:84-690-1613-X
- García, M.; Mauri, C. “Parálisis cerebral y nuevas tecnologías: Ayudas técnicas basadas en visión artificial.”, en *IV Jornadas Onubenses sobre Parálisis Cerebral*, Huelva. 2005.
- García, M. y Mauri, C. “Experiencia de interacción persona-ordenador a través de webcam con usuarios con discapacidad motriz grave y moderadamente afectados.”, en *Tecnología, Educación y Diversidad. Retos y realidades de la inclusión digital TECNONEET 2004*, pp. 259-264, Murcia, 2004.
- Mauri, C. “Interacción persona-ordenador mediante cámaras webcam.”, en *Interacción 2004: V Congreso de Interacción Persona-Ordenador*, pp. 366-367, Lleida, España, 2004.

1.4. Premios

- Accésit Feria Minusval 2004 – Lleida
- Nominación a los premios AT AWARD 2004 - Dusseldorf (Alemania)
- “Premio a la excelencia en innovación tecnológica 2004” - Colegio Oficial de Ingenieros Técnicos de Telecomunicación de Cataluña (COETTC) – Barcelona
- “Premi Accés 2007” en categoría investigación - Fundación Salas – Barcelona
- Accésit III Premio Anna-Bonomi Talita 2008 – Barcelona

- Accésit VI Premio Vodafone a la Innovación en Telecomunicaciones (2012) - Fundación Vodafone España – Sevilla

1.5. Organización de la tesis

El resto de la tesis se organiza en cuatro capítulos. El segundo capítulo recoge el estado del arte en sistemas interactivos audiovisuales para personas con parálisis cerebral. En el tercer capítulo se exponen las distintas propuestas en cuanto a técnicas de evaluación y sistemas interactivos. En el cuarto capítulo se presentan y discuten los resultados obtenidos durante nuestra investigación. Por último, en el quinto capítulo se comentan las conclusiones y se resumen las aportaciones y las propuestas de trabajo futuro.

CAPÍTULO 2

Sistemas interactivos audiovisuales

En este capítulo se resume el estado del arte en relación a los sistemas interactivos audiovisuales orientados a personas con discapacidad para uso expresivo, lúdico y terapéutico.

Índice

2.1. Clasificación	12
2.2. Principales propuestas	13
2.2.1. El sistema Soundbeam	13
2.2.2. Soundscapes	15
2.2.3. Proyecto MEDIANE	18
2.2.4. Otros proyectos	22
2.3. Sistemas musicales interactivos	25
2.3.1. Breve reseña histórica	26
2.3.2. Definición	27
2.3.3. Elementos de los SMI	28
2.4. Dispositivos de entrada	29
2.4.1. Sensor de ultrasonidos: Soundbeam	30
2.4.2. Sensores de infrarrojos	31
2.4.3. Análisis de sonido y voz	32
2.4.4. Visión artificial	32
2.5. Material audiovisual y correspondencia	35
2.5.1. Correspondencia en las propuestas estudiadas	36
2.5.2. Contenido	39
2.5.3. Contenido visual	40
2.5.4. Contenido sonoro	41
2.5.5. Contenido táctil	42
2.6. Efectos terapéuticos	43

2.6.1. Efectos observados	44
2.6.2. Desarrollo de las sesiones experimentales	47
2.6.3. Técnicas de análisis	49

En la literatura consultada existen bastantes trabajos orientados a la rehabilitación física o mental, pero no tantos en la línea que es de nuestro interés. Esta escasez se debe, en parte, a que la investigación de este tipo de sistemas es relativamente reciente (principios de los 90). Por ello, tras presentar la taxonomía, ofrecemos un recorrido por las principales propuestas encontradas en la literatura. Más adelante se proporciona una visión general de la tecnología empleada, el tipo de contenidos audiovisuales utilizados así como la metodología de aplicación para finalmente acabar hablando de los efectos terapéuticos y su análisis.

Hasta donde nosotros sabemos, no existe ningún otro trabajo que haga una recopilación tan extensa sobre este tipo de sistemas, por lo que consideramos la elaboración y estudio de este estado del arte como una contribución de esta tesis en sí misma.

2.1. Clasificación

Establecer una clasificación de los sistemas interactivos terapéuticos orientados a personas con discapacidad no es una tarea fácil puesto que en la literatura existen numerosas propuestas que combinan tecnologías dispares para gran diversidad de aplicaciones [32], aun así distinguimos varias familias en función de su aplicación terapéutica:

- Los destinados a la **rehabilitación física**, es decir que asisten al usuario o lo motivan para que realice ejercicios de rehabilitación físicos. Su estudio se enmarca en disciplinas como la neurología, ingeniería biomédica y la medicina física y rehabilitadora. Se enfoca, por ejemplo, a usuarios en fase de rehabilitación post-traumática [131] o con enfermedades como Parkinson [28].
- Los destinados la **rehabilitación mental**, por ejemplo para el entrenamiento de ciertas actividades de la vida diaria como: conducir una silla de ruedas [56], comprar en un supermercado [119] o cruzar una calle [130], para el tratamiento de fobias [94, 104, 95, 116], para el tratamiento de desordenes post-traumáticos [110] o para el tratamiento del

autismo [2, 121, 120]. Estos sistemas han estado diseñados con el objetivo de simular la exposición del usuario a una situación con el fin de realizar un aprendizaje o desensibilizarlo.

- Los destinados a fomentar una **interacción creativa o expresiva**, cuyos objetivos terapéuticos son más amplios e intentan hacer participar al usuario de una forma que, por sus limitaciones físicas o cognitivas, difícilmente podría hacerlo de otra forma. Estos sistemas intentan abarcar un espectro más amplio de usuarios y en particular aquellas personas con trastornos múltiples y profundos. **Esta es la línea en la que se centra principalmente esta investigación, aunque también prestaremos atención a aquellos sistemas que, pudiendo estar es esta categoría, también puedan ser empleados como instrumentos musicales accesibles.**

Dentro de esta última categoría, la mayoría de propuestas encontradas en la literatura, se basan en la utilización de los movimientos físicos y las emisiones orales de los participantes, para proporcionar una retroacción en forma de imágenes, sonido y/o vibración. Las propuestas son muy variopintas yendo desde aplicaciones informáticas hasta grandes instalaciones interactivas multimodales.

2.2. Principales propuestas

2.2.1. El sistema Soundbeam

2.2.1.1. Terapia sonora

A principios de los años 90 Phil Ellis empieza a aplicar lo que denomina terapia sonora utilizando el sistema Soundbeam en niños con parálisis cerebral, autismo y dificultades de aprendizaje múltiples y profundas (PMLD, *profound and multiple learning difficulties*).

La terapia sonora es una propuesta que combina la potencia de las nuevas tecnologías con una respuesta estética al sonido, lo que puede fomentar la interacción y el desarrollo de las habilidades comunicativas. Todo desde un enfoque no intervencionista (o mínimamente intervencionista) [36].

El sistema Soundbeam consiste en un sensor de ultrasonidos capaz de medir la distancia desde éste hasta el obstáculo más cercano. Esta zona del espacio se divide (virtualmente) en pequeñas regiones a cada una de las cuales se le hace corresponder una nota o acorde musical. El usuario, a través de sus gestos, activa el sensor a diferentes distancias generando sonidos y melodías y creando, de esta forma, un espacio virtual (véase la Sección 2.4.1 para más

información sobre este sistema). Esta herramienta permitió, en palabras de Ellis, que niños que no eran conscientes de su entorno exterior o no eran capaces de tomar el control o expresarse, pudieran ir tomando consciencia de su entorno, ir aprendiendo a controlar los movimientos físicos y responder y comunicarse con el mundo exterior. Ellis afirma [38] que este dispositivo puede ayudar a desarrollar el control físico, mejorar el rango de audición, despertar la curiosidad a través de la exploración y activar la auto-expresión. Las claves de la eficacia parecen ser la percepción de la relación causa-efecto y el descubrimiento del control (*agency*) a través de la acción.

Para describir la motivación interna que siente el participante al experimentar y explorar el entorno sonoro, Ellis acuña el término “resonancia estética” [37] que define como:

« Aquellos instantes durante los cuales se puede observar al niño totalmente inmerso en sus experiencias sonoras gracias a la motivación interna que produce el resultado sonoro obtenido a través de sus movimientos y gestos. »

A finales de los 90 se inicia el proyecto CARESS (*Creating Aesthetically REsonant Environments in Sound*) [53] alrededor de la tecnología de Sound-Beam. En este proyecto también se ensayan, al margen de Soundbeam, otros dispositivos de entrada como sensores de actividad muscular y sensores de torsión de antebrazo.

2.2.1.2. Terapia vibro-acústica

A partir del año 1998 Ellis hace evolucionar la terapia sonora a lo que denomina terapia vibro-acústica (*Vibroacoustic Sound Therapy (VAST)*) y comienza a experimentar con personas ancianas en residencias de la tercera edad [38, 39]. La terapia vibro-acústica incluye el uso de micrófonos y dispositivos vibratorios (concretamente una plataforma vibratoria del mismo fabricante que Soundbeam).

En las terapias con micrófonos se emplea como fuente sonora la voz y los sonidos vocales emitidos por el participante, así como los sonidos que se producen al frotar el micrófono contra superficies de distintos materiales y texturas. Los sonidos son amplificados y tratados digitalmente para añadirles efectos como reverberación o eco. Este tipo de terapia sonora se emplea para trabajar aspectos como [38]: la inflexión vocal, la enunciación, el repertorio fonológico, el uso expresivo de la voz y la escucha. El interés principal se centra en la comunicación expresiva (no verbal) a través de cambios de tono, volumen y timbre. En [41] pueden encontrarse algunas propuestas para las sesiones de terapia usando micrófonos.

El dispositivo vibratorio se emplea para reforzar los sonidos creados mediante Soundbeam o a través del micrófono y también para facilitar la relajación del participante al final de la sesión promoviendo una sensación general de bienestar físico y mental. Para esto último se emplean grabaciones especialmente diseñadas que combinan música clásica o relajante, con otro tipo de música de contenido variado con tonos sinusoidales de baja frecuencia (entre 20 y 75 Hz). Según las experiencias de Ellis este tipo de terapia es más apropiada para gente de la tercera edad y, en el caso de jóvenes, sólo cuando se observa mucho estrés o agitación [38].

2.2.1.3. iMUSE

A partir del año 2004 Ellis, ya desde el centro de investigación iMUSE (*interactive Multi-Sensory Environments*)¹ en la Universidad de Sunderland en el Reino Unido, se centra exclusivamente en la experimentación con personas ancianas combinando las técnicas antes mencionadas con una parte visual. Con ello pretende centrar la atención a través de la inmersión y fomentar la creatividad a través de un acoplamiento intuitivo entre percepción y acción [40]. Para ello emplea los sistemas G-Force², visualizador musical matemático, y Arkaos³, visualizador de bucles de vídeo y efectos en tiempo real vía MIDI.

2.2.2. Soundscapes

Hacia 1987 Anthony Brooks [10, 18, 17] inicia un proyecto auto-financiado de investigación y desarrollo que denominó HANDI-MIDI. Este proyecto estaba motivado por las experiencias con un miembro de su familia con discapacidad severa. Brooks utilizó una guitarra bajo MIDI y un pedal de efectos, de manera que él tocaba el bajo mientras su tío discapacitado accionaba el pedal. Con ello se dio cuenta que, a pesar de la limitada capacidad de expresión del sistema, permitía abrir vías de comunicación y proporcionar una enorme satisfacción.

A partir de ahí empezó a buscar nuevas formas de expresión más ricas. Después de mucha experimentación seleccionó un sensor de movimiento mediante infrarrojos. Este sistema permitía detectar desde movimientos muy pequeños (un pestañeo por ejemplo) a grandes movimientos corporales. El dispositivo constaba de tres sensores que, mediante triangulación, permitían detectar la posición en el espacio. El proyecto pasó a denominarse Soundscapes.

¹<http://my.sunderland.ac.uk/web/projects/imuse/ahome>

²<http://www.soundspectrum.com/g-force/>

³<http://www.arkaos.net/>

Soundscapes está formado por una librería de herramientas para la captura de la función corporal y por distintos programas para crear la respuesta a partir del gesto. Entre las herramientas de captura cuenta con sensores de actividad corporal (actividad cerebral, tensión muscular, resistencia eléctrica de la piel, ritmo cardíaco, etc.), de infrarrojos, de ultrasonidos y de visión artificial. La mayoría de las veces la transmisión de datos se realiza a través del protocolo MIDI y el sistema MAX/MSP⁴ se utiliza para la manipulación de los datos y creación de contenido. Soundscapes también hace referencia al concepto y a la metodología de trabajo [10].

Uno de los usos del sistema Soundscapes es el de “amplificador de expresión” para personas con discapacidad. Proporciona la posibilidad de generar sonidos e imágenes a partir de cualquier movimiento. Se trata pues de crear un espacio sensible donde es divertido permanecer en él y donde se estimula la motivación y creatividad a través del juego y la diversión. Además debe ser adaptable con el objetivo que sea accesible para todos, independientemente de las habilidades y limitaciones, preferencias o deseos, género o credo [19].

El uso terapéutico de Soundscapes queda ilustrado en proyectos europeos como Twi-aysi (*The world is - as you see it*) [54] y CAREHERE (*Creating Aesthetically Resonant Environments for the Handicapped, Elderly and REhabilitation*) [12, 55].

Twi-aysi es heredero del proyecto CARESS [53] iniciado por Ellis y en él se trabajó con niños con pluridiscapacidad de 3 a 6 años en el centro CAVI en Dinamarca⁵. En este proyecto se trabajó la manipulación de sonidos, imágenes coloridas y robots (focos con cabeza móvil motorizada). También se exploró la posibilidad de usar los sensores para permitir una navegación 3D y, aunque los usuarios rechazaron el uso de gafas de visión estereoscópica, la experiencia tuvo éxito. Una de las conclusiones fue que el uso combinando de audio y vídeo proporciona un mayor grado de inmersión que usando estos canales por separado. Este trabajo derivó en el proyecto CAREHERE.

En el proyecto CAREHERE [12] se amplió el abanico de usuarios incluyendo personas ancianas, enfermos de Parkinson, jóvenes con daño cerebral severo y otras discapacidades (todos usuarios de silla de ruedas y sin comunicación verbal) y adultos con discapacidad cognitiva (con capacidades comunicativas y motrices). Se adoptó un enfoque más personalizado intentando potenciar en cada participante aspectos concretos como: asociación causa efecto, expresión, movilidad, actividad verbal, etc. Se introdujo el uso de la visión artificial como medio de entrada del gesto a través del sistema EyesWebTM[70]. Algunas de las actividades que se propusieron a los participantes fueron: *cross modal*

⁴<http://www.cycling74.com/>

⁵Centre for Advanced Visualisation and Interactivity, <http://www.cavi.dk>

painting (música y pintura mediante gestos), *Shadow-ation* (creación de sombra chinescas empleando un proyector) entre otras propuestas. Las principales conclusiones del proyecto fueron el éxito de la aplicación de pintura mediante EyesWebTM[28] y el establecimiento de una metodología de evaluación cualitativa y cuantitativa.

Brooks ha explorado el uso de la plataforma Soundscapes con niños y adultos con discapacidad moderada y severa empleando sensores sin contacto [18, 50]. También ha empleado robots como medio para proporcionar retroacción [16, 105], ha investigado el uso de Soundscapes como herramienta de rehabilitación para personas con daño cerebral adquirido [14, 19] y también ha explorado el uso de la tecnología EyeToy© de la consola Sony Playstation 2 como herramienta de rehabilitación para niños en hospitales [20]. Ha experimentado con diversos sistemas de captación, principalmente con los sensores de infrarrojos antes referidos (véase la Sección 2.4.2 para más detalles sobre esta tecnología), pero también con Soundbeam [14, 50] y visión artificial [12]. En alguna ocasión también ha empleado micrófonos combinados con generadores de efectos digitales y dispositivos vibradores [50].

Brooks redefine el concepto de “resonancia estética” previamente acuñado por Ellis como [19]:

« La situación en la cual la respuesta a la intención del usuario es tan inmediata y estéticamente placentera como para hacerle olvidar el movimiento físico (y a menudo el esfuerzo) necesario para llevar a cabo la intención. »

En el pasado, estos movimientos o ejercicios repetitivos eran tediosos (en el campo de la rehabilitación) o incluso imposibles de conseguir con otros medios (en el campo de las necesidades especiales con niños con discapacidades severas).

En sus artículos más recientes [105, 23] nos habla de los procesos no formales de aprendizaje dada su naturaleza autónoma inherente a las experiencias de exploración, juego y diversión. Estas teorías se fundamenta en el trabajo de autores como Vygotsky [127] que afirma que el juego es la fuente para el desarrollo del niño o Csikszentmihalyi [33] que afirma que puede motivar el aprendizaje si proporciona experiencias complejas y desafiantes con retroacción inmediata, aunque el balance entre la incapacidad y la maestría es crucial en lo relativo a la curiosidad y la motivación. Es decir, las propuestas no deben ser ni demasiado fáciles ni demasiado difíciles.

2.2.3. Proyecto MEDiate

El proyecto MEDiate (*Multisensory Environment Design for an Interface between Autistic and Typical Expressiveness*) [102, 99, 100, 124, 106] es fruto del trabajo de un consorcio formado por el Responsive Environments Centre de la Universidad de Portsmouth (Reino Unido), la Faculteit Kunst, Media and Technologie del Hogeschool voor de Kunsten Utrecht (Países Bajos), el Grupo de Experimentación en Comunicación Interactiva de la Universidad Pompeu Fabra, el Social, Genetic and Developmental Psychiatry Research Centre del Kings College de Londres (Reino Unido) y la empresa Show Connections Limited (Reino Unido).

MEDIATE se dirige exclusivamente al colectivo de personas con trastornos del espectro autista que presenta grandes diferencias respecto al colectivo con parálisis cerebral, especialmente en los aspectos motriz y cognitivo. Aun así se estudia en este trabajo por sus interesantes aportaciones.

MEDIATE consiste en una instalación multisensorial de grandes dimensiones cuyo principal objetivo es [99] que permita al niño/a con autismo divertirse y tener la posibilidad de jugar, explorar y ser creativo dentro de un espacio controlable y seguro. También se pretende que puedan llevarse a cabo estudios de carácter psicológico y que los padres de los niños puedan observar a sus hijos jugando en este entorno.

2.2.3.1. Estrategias de diseño

En el diseño del sistema se tuvieron en cuenta las siguientes estrategias:

- Proporcionar a los usuarios sensación de control. Esto viene motivado por el hecho que normalmente los niños con autismo raras veces experimentan sensación de control sobre su entorno.
- Potenciar las acciones no repetitivas. Muchas veces los niños autistas se suelen aislar de los estímulos exteriores, que los saturan, realizando movimientos repetitivos denominados estereotipias. Éstos les permiten relajarse y sentirse más seguros. También pueden ser muy repetitivos cuando se obsesionan haciendo algo que les gusta una y otra vez. Ambos comportamientos no son deseables puesto que aíslan al niño del mundo. Por eso el sistema se diseñó para detectar estos patrones repetitivos y modificar la respuesta del sistema en función de estos.
- Adaptable a cada usuario. Como existe gran diversidad entre los niños con autismo, el comportamiento del sistema no es fijo ya que varía en función del patrón de interacción del niño. Con ello se intenta captar su

2.2. Principales propuestas

19

atención en todo momento a través de un diálogo creativo y exploratorio con el entorno.

- Hacer partícipes a los usuarios del proceso de análisis y diseño. Se escogieron a varios niños autistas con capacidad comunicativa para que mostraran sus opiniones acerca del tipo de estímulos y modos de interacción propuestos.

2.2.3.2. Entorno y diseño

En cuanto al tipo de estímulos, se prefirieron las formas abstractas frente a las concretas para evitar el riesgo que las preferencias de los diseñadores no gustaran a los niños o les produjeran reacciones adversas. Existe una probabilidad realmente alta de que un estímulo concreto (imagen de un perro, un coche, un familiar, etc.) acapare totalmente la atención e impida percibir otros elementos del entorno.

Para proporcionar sensación de control a los usuarios más afectados, el sistema se diseñó de forma que intentara establecer un diálogo basado en las acciones del usuario con reacciones muy directamente relacionadas con éstas; pensando en que las acciones del usuario serían necesariamente simples y sin metas específicas. Por ello se pensó en los gestos corporales.

En general la idea era imponer las mínimas restricciones posibles al usuario y crear un entorno físico interactivo de forma complementa no invasiva. La interacción a través de objetos podría haber llevado a los niños a una de las actitudes no deseadas: la pérdida de interés en el objeto o un interés obsesivo por el mismo.

2.2.3.3. La instalación

La instalación tiene forma de hexágono (véase la Figura 2.1) de aproximadamente 6 metros de diámetro. En su interior, distintos elementos sirven de interfaz de usuario (véase la Figura 2.2):

- Las **baldosas del suelo** incorporan sensores que detectan las pisadas y se envían al controlador central del sistema.
- **Tune Fork**. Es una estructura en forma de tubo con variedad de texturas a lo largo de su extensión. Incorpora sensores que detectan la presión ejercida y envían esta información a través de MIDI.
- **Impression Wall**. Es una pared con estructuras acolchadas con sensores para medir la presión además de dispositivos vibratorios.

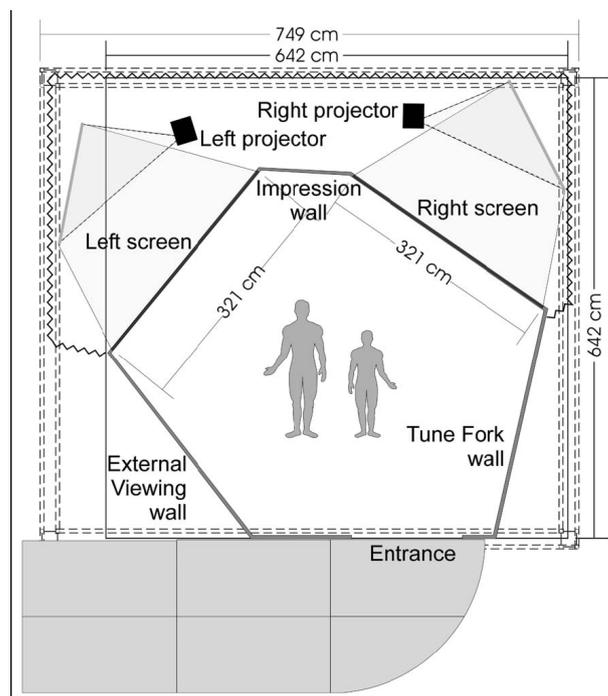


Figura 2.1: Vista de la planta del espacio físico de MEDATE [101].

- Dos grandes **pantallas interactivas** de retro-proyección (300 x 225 cm) se usan para mostrar los gráficos y sirven como enormes pantallas táctiles. El motor gráfico se basa en un sistema de partículas que permite generar un contenido abstracto.
- Sistema de **visión artificial** compuesto por siete cámaras e iluminación por infrarrojos permiten localizar la posición del usuario dentro del entorno y extraer su silueta (véase la Sección 2.4.4.1 para más información sobre siluetas).
- Sistema de **captura de sonido**. Un conjunto de micrófonos distribuido al largo del entorno del entorno capta cualquier sonido emitido por el participante.
- Sistema de **generación sonora** compuesto por ocho altavoces distribuidos a lo largo del entorno capaces de generar sonido 3D.

2.2.3.4. Funcionamiento

Cuando el usuario entra por primera vez en el entorno, la respuesta del sistema es simple y literal. La pantalla muestra una representación del usuario análoga a su silueta. Ésta imita la forma del usuario y sus movimientos de

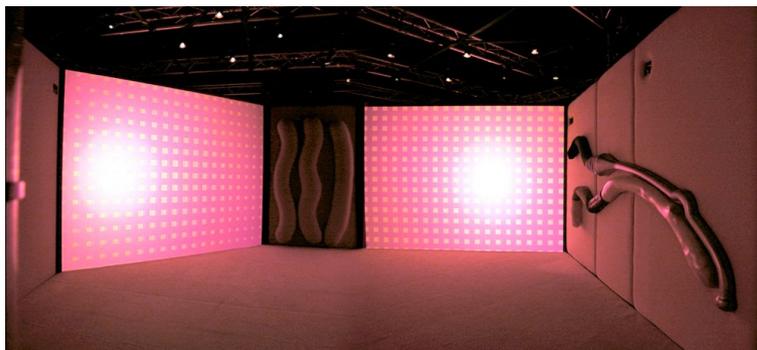


Figura 2.2: Vista panorámica del interior de la instalación de MEDIATE [101].

una manera obvia e inmediata. De forma similar los sonidos producidos por el usuario, como sus pasos, son amplificados y presentados al usuario y la presión sobre el *Impression Wall* resulta en una respuesta vibratoria proporcional.

El entorno de MEDIATE proporciona un espacio donde los estímulos son limitados, enfocados y sólo responden al usuario. De esta manera, el sistema intenta establecer un diálogo con el usuario. En efecto, el usuario controla la respuesta del sistema, pero al igual que en un diálogo real, el sistema no permanece pasivo y, después de un periodo inicial para que el usuario se acostumbre al entorno, éste empieza a adaptar su comportamiento. Para ello el sistema dispone de un módulo funcional encargado de proporcionar una respuesta inmediata a las acciones del usuario y otro módulo que va adaptando el comportamiento del anterior. El rango de posibilidades va desde respuestas inmediatas y miméticas hasta respuestas más complejas, de carácter artístico. La modalidad puede ser propia (por ejemplo, si el usuario emite un sonido, MEDIATE responde con un sonido) o cruzada (por ejemplo, un gesto del usuario se traduce en una respuesta sonora). Las diferentes variables que controlan la complejidad de la respuesta son gestionadas de forma automática por el controlador del sistema, que analiza los comportamientos repetitivos del usuario.

2.2.3.5. Adaptabilidad del sistema

MEDIATE adapta automáticamente la respuesta del sistema a los comportamientos del usuario. Concretamente, se focaliza en las acciones repetitivas (no deseables) y la no actividad. Cuando no hay actividad, el sistema genera de manera automática algunos estímulos para invitar a la participación, mientras que si se detectan comportamientos repetitivos se modifica la salida para romper estas dinámicas y hacer que el participante vuelva a interactuar de forma creativa. En la Sección 2.5.1 se amplía la información a este respecto.

2.2.3.6. Resultados y discusión

Por la instalación de MEDiate pasaron (entre 2003 y 2004) más de 90 niños con autismo moderado y severo. El tiempo que estuvieron en el entorno va de 5 a los 35 minutos y todos utilizaron, como mínimo, uno de los elementos interactivos. La mayoría de ellos entraron de forma voluntaria y espontánea. Todas las sesiones fueron registradas en vídeo y analizadas por los psicólogos. De acuerdo con los resultados, ninguno de los participantes se mostró inquieto o incómodo (únicamente tuvo que interrumpirse una de las sesiones debido a la sobre-excitación del niño) y la mayoría presentó la deseada sensación de control y acción (*agency*).

2.2.4. Otros proyectos

2.2.4.1. IntelliVision

IntelliVision [25] –fruto de una colaboración entre la Universidad de Nueva York y el New York Foundling Hospital– es un entorno multisensorial dirigido inicialmente a niños hospitalizados con discapacidad visual neurológica (también llamada discapacidad visual cortical)⁶. Algunos de estos niños, presentan discapacidades cognitivas asociadas que, conducen a comportamientos violentos o auto-destructivos. Habitualmente estos niños presentan restos de visión y suelen emplear la visión periférica para el seguimiento de estímulos visuales.

El sistema consiste en una gran pantalla de vídeo montada sobre la cama o silla de ruedas. Una cámara de vídeo captura la imagen del usuario y convierte sus movimientos en cambios en las imágenes y sonidos presentados en función del módulo de configuración seleccionado de entre los tres disponibles. Dentro de cada módulo es posible ajustar otros parámetros como: sensibilidad de movimiento, paletas de colores, género musical y volumen. Incluye además un módulo táctil que vibra en respuesta a las bajas frecuencias de sonido⁷. Adicionalmente se dispone de una cámara de vídeo para registrar la sesión. La programación se llevó a cabo con MAX/MSP/Jitter.

Uno de los programas que incorpora IntelliVision muestra la imagen del usuario en pantalla empleando colores vibrantes. Sobre la imagen se muestra una barra vertical. Cualquier movimiento se convierte en notas musicales que

⁶Ocurre cuando una parte del cerebro que es la responsable de la visión está lesionada. En otras palabras, el ojo por sí mismo es normal, pero el cerebro no procesa la información correctamente. Tiene varias causas, que incluyen, pero no están limitadas a, la falta de oxígeno antes, durante o después del parto, enfermedades producidas por virus o bacterias como la meningitis y el citomegalovirus, o una lesión traumática.

⁷En este caso se utilizó un transductor táctil o *bass shaker* parecido a los empleados en los equipos de audio para automóvil.

van apareciendo en la pantalla. Cuando la barra se desplaza sobre las notas, éstas son reproducidas junto con el fondo musical. Mientras el usuario se continua moviendo van apareciendo nuevas notas.

El proceso de desarrollo incluyó una primera fase de observación donde se mostraba a los usuarios distintos materiales visuales (imágenes de naturaleza, urbanas y formas abstractas). Durante esta fase se vio que lo que más atraía la atención eran los cambios en las imágenes y los movimientos de éstas más que el propio contenido. En la siguiente fase se incorporó la imagen del propio usuario y el movimiento de éste disparaba efectos visuales coloridos de carácter abstracto. Se vio que algunos niños se mostraban más atraídos al verse en la pantalla. También se probó con distintos tipos de música y sonidos y nuevamente se observó que lo más atracción causaba era los cambios. Por lo que respecta a los estímulos táctiles (vibración) vieron que contribuían a calmar al usuario y centrar su atención. Véase la Sección 2.5.1 para más información sobre la propuesta.

2.2.4.2. Virtual Music Instrument (VMI)

Se trata de una propuesta del Dr. Tom Chau y sus colaboradores del Prism Lab⁸, organización vinculada al hospital Bloorview Kids Rehab y la Universidad de Toronto en Canadá. El sistema –compuesto por elementos de uso común como un ordenador, cámara USB, pantalla y altavoces– muestra la imagen del usuario en la pantalla y transforma cualquier movimiento en música [71, 65]. Está pensado, especialmente, como instrumento musical alternativo para su uso en sesiones de musicoterapia.

El proyecto se fue completando por fases donde se exploraron distintas propuestas:

- En primer lugar, el ordenador mostraba una imagen del usuario en pantalla y se disponía de dos modos de operación: el primero iniciaba sonidos aleatorios de percusión mientras el usuario se movía y el segundo iniciaba la reproducción de una pieza MIDI. La reproducción se detenía al detenerse el movimiento, y se reiniciaba cuando el usuario volvía a moverse. Este experimento mostró la necesidad de poder ajustar la sensibilidad del sistema.
- Posteriormente se incluyó imagen en color empleando el sistema “Building Blocks”⁹. En este caso el movimiento iniciaba eventos sonoros como

⁸Paediatric Rehabilitation Intelligent Systems Multidisciplinary.
<http://www.prismlab.org>

⁹Entorno de programación MIDI orientado a objetos.
<http://aureality.midiworld.org/products/buildingblocks/buildingblocks.html>

arpeggios, acordes o breves frases musicales. Además este tipo de eventos musicales no había que definirlos previamente ya que el propio software los generaba automáticamente dentro de unos márgenes (de estructura rítmica, escala musical, etc.). En las pruebas –realizadas con 5 niños con parálisis cerebral con edades comprendidas entre 4 y 6 años– se vio que esta propuesta se prestaba a la confusión y dificultaba la distinción entre causa y efecto.

- La tercera fase pasó por simplificar la respuesta sonora creando algo muy parecido a un instrumento musical. Concretamente, se definen algunas regiones del espacio alrededor del niño a cada de las cuales se le asignó distintas propiedades musicales (tono, volumen y timbre) y parámetros de configuración (sensibilidad y tiempo de respuesta). Cada una de estas regiones se representa en pantalla mediante una forma coloreada o la imagen de un instrumento (el terapeuta o facilitador podía ajustar convenientemente el número de zonas, sonidos asociados, sensibilidad, etc.). Cuando se detectaba movimiento dentro de alguna de estas regiones se iniciaba el evento musical apropiado. En el estudio participaron 7 niños con edades comprendidas entre los 2.5 y 7 años con diagnóstico mixto (parálisis cerebral y/o atrofia muscular espinal)
- La cuarta fase incluye la idea de asignar a cada región sensible una nota musical distinta e impartir lecciones de música (ejercicios de técnica, escucha y aprendizaje de la melodía de, al menos, una canción). En el estudio participaron los niños más mayores de la escuela con diagnóstico mixto.

Las conclusiones generales de las últimas investigaciones llevadas a cabo con el VMI en niños con parálisis cerebral revelan que éste ofrece amplias oportunidades a los niños en relación a la participación y el desarrollo creativo. VMI potencia el sentimiento de competencia y además mejora su condición física, cognitiva, comunicativa, emocional, conductual y de competencia social.

Sus autores también proponen [1] la utilización de VMI como herramienta para fisioterapia, terapia ocupacional, logopedia y psicoterapia para la rehabilitación psicosocial y fisiológica en niños con parálisis cerebral.

El sistema está disponible para el público en forma de producto comercial¹⁰.

¹⁰<http://www.bloorview.ca/research/commercialization/vmi.php>

2.2.4.3. Adaptive Use Musical Instruments

Se trata de un sistema de visión artificial con el cual sus usuarios pueden tocar un teclado virtual mediante pequeños movimientos de la cabeza. Concretamente los movimientos a derecha e izquierda activan las teclas de forma secuencial con lo que es posible crear melodías escalares. Es posible seleccionar distintas escalas y tipos de sonido. También es posible emplear sonidos de percusión. En este caso la cámara detecta los movimientos de la cabeza y los sonidos se hacen corresponder a cada uno de los cuatro cuadrantes en que se divide el área de captura. Los usuarios pueden tocar solos o con otros percusionistas en conjunto. El sistema ha sido probado con éxito durante una semana en el Programs Inc. School de Poughkeepsie NY. El software está programado en MAX/MSP y es operativo para Mac y PC [96]¹¹.

2.2.4.4. Sistemas comerciales

La empresa GestureTek¹² de Toronto (Canadá) está especializada en el desarrollo de sistemas interactivos basados en visión artificial. Una de sus líneas de productos es la que denominan “*motion activated multi-sensory products*” (productos multisensoriales activados por movimiento)¹³ donde la interacción se realiza a través de una cámara, mostrando al usuario en un entorno virtual, y permite a éste interactuar de forma independiente con la imagen de la pantalla o proyección, así como explorar su creatividad e imaginación a través de efectos visuales y auditivos. Todos los productos de esta línea incluyen efectos estimulantes y relajantes [45].

2.3. Sistemas musicales interactivos

Los sistemas virtuales interactivos que nos ocupan en este trabajo heredan la tradición del mundo de la música y, concretamente, de los sistemas musicales interactivos. Por ello este campo ha sido una referencia para este tipo de sistemas, hasta el punto que muchos de ellos están propuestos por músicos profesionales. Es por todo esto que, antes de presentar una revisión de las distintas tecnologías más comúnmente empleadas, ofrecemos una breve descripción de los sistemas musicales interactivos.

¹¹Véase vídeo en: <http://www.youtube.com/watch?v=ABqABv7MnmM>

¹²<http://www.gesturetekhealth.com>

¹³En la página web existe un apartado donde pueden descargarse artículos de investigación donde se ha empleado la tecnología de la empresa.

2.3.1. Breve reseña histórica

Posiblemente uno de los primeros sistemas musicales informatizados que puede considerarse un instrumento musical es el programa *Conduct* para GROOVE creado en 1969 por Max Mathews [82, 81, 83]. Este sistema fijaba la secuencia de tonos a interpretar (la partitura) ya que normalmente es el parámetro que el compositor determina de forma más rígida. De esta forma el intérprete podía concentrarse en controlar otros parámetros como la amplitud, el tempo o el balance.

Si un instrumento es capaz de conocer de antemano que nota viene a continuación, se dispone de cierto conocimiento para poder tomar decisiones sobre la marcha. En lugar de emplear una secuencia de notas predeterminada, cada nota que se debe tocar, por ejemplo, puede resultar de una decisión combinada entre el intérprete y el instrumento. Es lo que Chadabe [29] denomina el paso de la “automatización de la memoria” a la “automatización del proceso”.

Con un instrumento tradicional, es sólo el intérprete quien toma decisiones en tiempo real durante el diálogo que ocurre entre éste y su instrumento acústico. Sin embargo, los nuevos instrumentos digitales pueden tomar millones de decisiones por segundo, lo que les convierte en instrumentos inteligentes [118, 31]. Esta nueva “inteligencia” sugiere un continuo entre los instrumentos convencionales pasivos, hasta una máquina de interpretación musical totalmente autónoma (sin intervención humana). Entre estos dos extremos, el instrumento inteligente puede responder de muy variadas formas, no siempre predecibles, forzando a su vez al intérprete a reaccionar. Es lo que Chadabe denomina “composición interactiva” [30] donde el intérprete comparte el control de la música con el ordenador.

Este tipo de sistemas, donde se mezcla la música, la interactividad y la informática se han venido llamando durante las dos últimas décadas [62] **sistemas musicales interactivos**. Gracias al aumento de la potencia de cálculo de los ordenadores y al abaratamiento de éstos, así como a la aparición del protocolo MIDI, popularizaron este tipo de sistemas durante los años 80.

Antes de cerrar esta reseña queremos hacer referencia al que, probablemente, es el primer instrumento musical de la historia que se toca sin contacto. En rigor no se trata de un sistema musical interactivo ya que no incorpora inteligencia de ningún tipo, pero resulta interesante por el modo de interacción. Se trata del Theremin (véase la Figura 2.3), inventado en 1920 por Leon Theremin [123]. Para tocar este instrumento, el intérprete acerca o aleja las manos de unas antenas. Una de estas antenas controla el volumen y la otra la frecuencia; parámetros que están en función de la distancia a la que se sitúa la mano.



Figura 2.3: Leon Theremin tocando el Theremin [6].

2.3.2. Definición

Un sistema musical interactivo [62] reúne las siguientes propiedades fundamentales:

- Incorpora una unidad central de proceso (UCP).
- Es interactivo.
- Genera un resultado musical en tiempo de interpretación¹⁴, bajo el control de uno o más intérpretes.

Esta definición es muy amplia y dentro de ella cabrían dispositivos como sintetizadores digitales que normalmente no se consideran sistemas musicales interactivos. Por ello es necesario refinar el término interactivo en este contexto: los sistemas musicales interactivos deben ser lo suficientemente interactivos para afectar y modificar las acciones del intérprete(s), provocando un diálogo entre el intérprete(s) y el ordenador.

¹⁴Lo que es más restrictivo que en tiempo real

2.3.3. Elementos de los SMI

Un sistema de estas características se compone de los siguiente elementos [128]:

1. Los gestos. Es el gesto instrumental. El movimiento de la cabeza, o la flexión de una articulación, la articulación de sonidos con la boca, etc.
2. El controlador gestual. Interfaz del controlador. Como se capturan y analizan los gestos, es decir, la conversión del mundo analógico al digital y el análisis de los datos.
3. Correspondencia (*mapping*). Una vez obtenidos y analizados todos los datos a tiempo real de entrada del gesto hay que asociarlos a parámetros de control. Es decir, hay que poner en relación el gesto con el control de los parámetros sonoros o audiovisuales en función de un resultado deseado o a explorar. ¿Qué repercusión sensible tiene el gesto?, ¿A qué lo asociamos? ¿A la frecuencia de un sonido?, ¿A la intensidad?, etc. Existen distintos tipos de correspondencia en función de la complejidad del sistema.
 - Correspondencia uno-a-uno, pocos-a-muchos. El tipo de correspondencia más obvia es aquella que asocia un solo parámetro de la salida (por ejemplo tono, volumen, color, etc.) con una dimensión de control independiente. Ésta es la correspondencia uno-a-uno. Ocurre muchas veces que la salida no se compone de un único parámetro. La situación en la cual las dimensiones de la salidas son muchas más que las de la entrada se denomina correspondencia pocos-a-muchos [72]. Por ello, muchos autores han propuesto descomponer la cadena de correspondencia en varios niveles, del gestual al perceptivo y del perceptivo a los parámetros de síntesis [129, 128]; reduciendo de esta forma la dimensionalidad de la salida. Esta definición que, en principio, se aplica únicamente a la generación de sonido, puede extenderse a la generación de cualquier otro contenido.
 - Correspondencia uno-a-muchos, muchos-a-uno. El significado de la uno-a-muchos (correspondencia divergente [113]) se puede entender más fácilmente con la pregunta “¿Qué controla el arco del violín?” El arco de un violín controla muchos aspectos de sonido como el volumen, el timbre, la articulación y el tono. La correspondencia muchos-a-uno (correspondencia convergente) puede comprenderse con la cuestión: “¿Dónde esta el control de volumen del violín?”.

No existe un único “control de volumen”. El volumen depende de la velocidad del arco, de la presión del arco, de la cuerda elegida y de la posición de los dedos sobre las cuerdas [52]. Los instrumentos musicales reales presentan una gran complejidad en el control que implica relaciones uno-a-muchos y muchos-a-uno.

Además en los instrumentos musicales acústicos la correspondencia no es del todo lineal. Por ejemplo, soplar más fuerte en un instrumento de viento no sólo afecta al volumen, sino también al tono. Notemos que las funciones de correspondencia no lineales sólo pueden ser definidas de forma explícita en un contexto uno-a-uno. Una correspondencia uno-a-muchos o muchos-a-uno es inherentemente no lineal.

- Correspondencia compleja: memoria, no-linealidad y caos. Generalmente las funciones de correspondencia son instantáneas, por lo que, en un momento dado la salida del sistema sólo depende de su entrada en ese instante. Si se tiene en cuenta el historial de la interacción pueden generarse más comportamientos en la salida. Por ejemplo:
 - Filtros digitales de sonido (pasa-bajos, pasa-altos, etc). Necesitan de las n últimas muestras de sonido generado para calcular el nuevo valor.
 - Velocidad de la entrada (con un diferenciador).
 - Variando el resultado en función del historial de gestos, como ocurre con MEDIANTE.

4. Diseño del motor. El algoritmo que genera la respuesta sonora y/o visual. Hay que contemplar distintos tipos de síntesis sonora.
5. Retorno primario o secundario. Estos son los retornos ligados a la manipulación del interfaz, los retornos producidos por el instrumento.

2.4. Dispositivos de entrada

En este apartado se resumen algunos de los dispositivos de entrada utilizados en las propuestas estudiadas. Se hace especial hincapié en los dispositivos sin contacto por el interés que tienen para el colectivo de personas con parálisis cerebral.

2.4.1. Sensor de ultrasonidos: Soundbeam

SoundBeam [107], que ya se ha introducido en la Sección 2.2.1, es uno de los sistemas más utilizados con personas con discapacidad, especialmente niños. Por ejemplo Phil Ellis lo utiliza en sus sesiones de terapia desde principios de los 90 [36, 38, 39, 40]; aunque también ha sido empleado por otros autores [71, 12, 50, 11, 105, 15].

Los detalles de funcionamiento del sistema Soundbeam son los siguientes (según los manuales de uso del fabricante)

- El sistema puede tener conectados simultáneamente hasta 12 sensores de ultrasonidos, aunque parece ser que en las sesiones de terapia sólo se emplea uno.
- El rango de trabajo del sensor es configurable entre 0,6 y 6 metros de distancia. Se le puede indicar una distancia mínima y otra máxima dentro de la cual se producirá el sonido, es el rango de trabajo.
- El rango de trabajo se puede dividir hasta en 64 particiones.
- La base para la composición son las llamadas secuencias de tonos (*pitch sequences*) que son conjuntos de hasta 64 notas o acordes (hasta cuatro notas de polifonía).
- Cada acorde se corresponde con una división y es posible definir un desplazamiento (*offset*) respecto a la primera nota de la secuencia. Cada secuencia de tonos se puede vincular a uno o varios sensores.
- La activación de los tonos se puede realizar de distintas maneras que se configuran con el parámetro modo de activación (*trigger mode*) que admite las siguientes opciones:
 - *Single shot*. Reproduce un tono de la duración de la interrupción del haz. Se reproduce un nuevo tono cuando cesa la primera interrupción y se inicia otra. El movimiento dentro del haz no produce cambios.
 - *Re-trigger*. Se reproduce una nota cada vez que el haz se interrumpe y cuando se mueve el obstáculo dentro del haz. Permite generar *glissandos* y según el manual es el modo más empleado.
 - *Poly sustain*. Como el modo *single-shot*, pero cada nueva interrupción genera una nueva nota que sigue sonando para siempre. Existe una polifonía máxima y se puede silenciar interrumpiendo fuera del rango de trabajo.

- *Re-trigger sustain*. Como el modo *re-trigger* pero cada nueva interrupción o movimiento dentro del haz añade una nueva nota al acorde.
 - *Gated single shot*. Como el modo *single-shot* pero se sostiene cada nota durante un periodo configurable de tiempo. Si se elimina la interrupción la nota sigue sonando y si se genera una nueva interrupción se silencia la nota anterior y se reproduce una nueva.
 - *Gated Re-Trigger*. Como el *re-trigger* pero cada nota suena sostenida durante un periodo configurable de tiempo.
- El dispositivo es capaz de almacenar hasta un total de 128 configuraciones, 30 de las cuales vienen de fábrica.

Soundbeam es meramente un controlador gestual (dispositivo de entrada) y por lo tanto es necesario conectarlo vía MIDI a un módulo de sonido u ordenador para poder obtener un resultado sonoro o visual.

2.4.2. Sensores de infrarrojos

Uno de los sistemas de entrada por infrarrojos más popular en este ámbito ha sido Dimension Beam (DB) fabricado por la empresa (ya desaparecida) Interactive Light Inc durante los 90. Este dispositivo emite un haz invisible de luz infrarroja que adquiere en el espacio un volumen en forma de óvalo. Cualquier movimiento (no detecta objetos estáticos) dentro de este espacio es detectado y transformado en mensajes MIDI [109]. El sensor detecta la distancia del objeto en movimiento hasta la cabeza del sensor (eje Y) y respecto (perpendicularmente) al centro del haz (eje X). Actualmente el fabricante de instrumentos musicales Roland¹⁵ incorpora esta tecnología en algunos de sus productos.

Brooks ha utilizado de forma extensiva los sensores de infrarrojos en su propuesta Soundscapes. Empezó empleando un modelo de tres cabezales [17] que, por medio de triangulación, era capaz de retornar una posición en 3D del objeto o parte del cuerpo utilizada para la interacción. También ha empleado sensores del tipo DB e incluso ha combinado varios de ellos para conseguir mayor resolución [13] y facilitar la configuración y puesta en marcha [18].

Una de las principales limitaciones de la tecnología de sensores infrarrojos es que es sensible a los cambios de luz [15], lo que se tiene que tener en cuenta, especialmente si se desea trabajar con efectos de luz o en exteriores.

¹⁵<http://www.roland.com/>

2.4.3. Análisis de sonido y voz

La captación de la voz y otros sonidos mediante micrófonos, no sólo se ha utilizado como fuente de material sonoro [41], sino también como elemento de control a partir del análisis del audio. Por ejemplo, en el proyecto MEDIANTE [106] se utilizan hasta 5 micrófonos para captar cualquier sonido emitido por el usuario. De la fuente sonora se analizan¹⁶ la amplitud, el tono, el contenido armónico y los formantes. De todos estos datos sólo se utilizan aquellos que son relevantes atendiendo a parámetros como estabilidad y coherencia por ejemplo.

2.4.4. Visión artificial

La idea básica de la interacción basada en visión artificial es la utilización de una o más cámaras para observar una escena y analizar las fuentes de vídeo en tiempo real para entender los gestos del usuario, expresiones u otros mensajes del canal visible.

2.4.4.1. Siluetas

Generalmente, en la literatura estudiada, la mayoría de las técnicas de visión artificial empleadas se basan en la extracción y procesado de la silueta del participante.

Uno de los primeros sistemas data de mediados de los 70 cuando Krueger [68] desarrolló un sistema llamado VideoPlace. Detrás del usuario se situaba un potente sistema de iluminación, con lo que la cámara capturaba una imagen muy contrastada, casi binarizada, del participante. De esta forma era sencillo separar la silueta del usuario del fondo permitiendo la integración con los gráficos generados por ordenador. Es decir, integraba la silueta del usuario en el escenario virtual con el que se permitía interactuar. Krueger notó que los usuarios estaban muy motivados cuando utilizaban el sistema y sugirió que el sistema podría hacer más amenas las sesiones de rehabilitación física.

El sistema ideado por Krueger fue perfeccionado e introducido en el mercado por la empresa Vivid Group de Toronto (Canadá) (actualmente GestureTek y antiguamente Mandala) [103] con el nombre de Vivid GX. Este sistema, que se ha usado en algunas investigaciones clínicas como herramienta de soporte a la rehabilitación [108], funciona con un fondo de color negro ante el cual se sitúa el usuario con ropa clara. Mediante segmentación por color (croma) se consigue separar fácilmente la imagen del usuario en primer plano del fondo.

¹⁶Empleando el objeto de MAX/MSP fiddle~

La imagen del participante, a todo color, se superpone sobre un escenario generado por ordenador con el que el usuario puede interactuar. Para ello se emplean unos controles que comparten el espacio perceptual entre el usuario y el ordenador. Cada control se representa como una región localizada en la imagen que normalmente el usuario puede accionar. Por ejemplo, disponía de un módulo que presentaba al usuario un kit de percusión que el usuario podía hacer sonar tocando virtualmente los distintos elementos [65].

Estos dos sistemas no eliminan la necesidad de un entorno controlado (iluminación especial, fondo homogéneo, etc.), lo que complica y encarece la instalación. Uno de los primeros sistemas que eliminó esta necesidad, y además se concibió como instrumento musical, fue “Very Nervous System” (VNS) [111, 132]. Éste analizaba los cambios espacio-temporales de la secuencia de vídeo para determinar aquellos píxeles que habían cambiado respecto a un determinado patrón. Este patrón podía ser fijo (por ejemplo la escena que ve la cámara antes de la entrada del participante) con lo que se conseguía detección de presencia, o bien los n últimos fotogramas capturados, con lo cual era posible detectar cualquier objeto en movimiento. La interacción se llevaba a cabo de forma similar a Vivid GX, es decir, ciertas regiones de píxeles pre-definidas desencadenaban un evento al detectar presencia o movimiento. A este evento se le hacía corresponder una nota MIDI o la variación de un parámetro como el tempo o el timbre. La literatura recoge además que VNS fue empleado con personas con discapacidad como instrumento musical accesible.

Dada la época en que fue concebido el sistema VNS (principios de los 90) el procesamiento de imagen se realizaba con hardware especializado. Con el aumento de la potencia de cálculo de los ordenadores este proceso hace tiempo que es posible hacerlo por software. Un ejemplo de ello es la EyeToy¹⁷ de la consola Sony PlayStation 2. Este sistema, concebido con finalidades lúdicas, emplea el análisis de movimiento para incluir la imagen del participante en el escenario del juego y permitirle la interacción a través del movimiento de brazos, piernas, cabeza, etc. EyeToy se ha empleado en diversos estudios [115, 108, 60, 20] para valorar sus propiedades como herramienta de rehabilitación física. Estos estudios revelan que los participantes –especialmente los niños– experimentan EyeToy como una actividad divertida, motivadora y fácil de usar. Por el contrario, al tratarse de una plataforma cerrada dispone de un catálogo de aplicaciones limitado y diseñado únicamente con finalidades lúdicas, por lo que no puede ser ajustado convenientemente para usuarios con movilidad limitada o tiempos de reacción más largos.

En este apartado también nos referimos al sistema EyesWeb¹⁸ [28]. Se tra-

¹⁷<http://www.eyetoy.com>

¹⁸<http://www.eyesweb.org>

ta de un entorno de programación visual que permite interconectar distintos bloques para crear un *pipeline* de procesamiento de imágenes. Una de las técnicas que incorpora es precisamente la de cálculo de *Silhouette Motion Images* (SMI). Una SMI es una imagen con información sobre las variaciones de la forma de la silueta y posición de los últimos fotogramas del vídeo. Las SMI están inspiradas en las técnicas denominadas *motion-energy images* (MEIs) y *motion-histoty images* (MHIs) [9, 8].

Una imagen MEI es el resultado de acumular la suma de los cuadrados de la diferencia entre fotogramas consecutivos, lo que proporciona una distribución espacial robusta del movimiento. Además al tratarse de operaciones sencillas –diferencia, suma y producto– es apropiada para aplicaciones de tiempo real.

La imagen MHI es similar a la imagen MEI sólo que al acumular las diferencias entre fotogramas consecutivos se aplica un factor de decaimiento lineal de forma que los últimos cambios tienen mayor peso que los primeros y el resultado que se obtiene es que el movimiento que ha sucedido en los fotogramas más recientes aparece con un valor mayor que el movimiento más antiguo.

Una imagen SMI se diferencia de una imagen MEI en que no se tiene en cuenta el fotograma más reciente. De esta forma sólo se considera el movimiento, mientras que la postura actual se descarta. Además, el sistema EyesWeb, permite extraer una serie de parámetros útiles para la interacción:

- Cantidad de movimiento. Consiste en el cálculo del área de la imagen SMI, da una idea aproximada de la cantidad de masa que se está moviendo en un instante determinado. Para hacer esta medida independiente de la distancia del sujeto respecto de la cámara, el área de la SMI se divide entre el área de la silueta.
- Índice de contracción. Valor entre 0 y 1 que da una idea de como el usuario está utilizando el espacio de alrededor. Las variaciones de esta medida pueden usarse para saber si el sujeto se está contrayendo o expandiendo y, por ejemplo, para generar la retroacción adecuada en cada caso.
- Índice de direccionalidad. Indica en que medida una trayectoria es directa o flexible. Se calcula como la relación entre la longitud de la trayectoria rectilínea entre dos puntos y la suma de las longitudes de los segmentos de la trayectoria real.
- Segmentación de movimiento. Estudiando la evolución de la cantidad de movimiento a lo largo del tiempo es posible determinar en que intervalos el sujeto está inmóvil o se está moviendo.

- **Fluidez e impulsividad.** Estudiando la secuencia de pausas y movimientos es posible evaluar si se trata de un movimiento fluido y regular, o más bien vacilante. Un movimiento será poco fluido si una acción se realiza con pausas frecuentes y será más fluido si la acción se realiza de una forma continua. La impulsividad está relacionada con movimientos de corta duración a gran velocidad (picos).

Este tipo de parámetros se han utilizado tanto para el control de la retroalimentación [28], como para obtener descripciones simbólicas de la actividad del usuario para su posterior estudio [12].

En el proyecto MEDiate [106], por ejemplo, se emplea la técnica de detección de presencia antes referida. Al tratarse de una instalación, esta técnica es apropiada, pues la calibración del sistema de cámaras se realiza por el personal técnico durante el montaje. Esta técnica requiere que la iluminación de la escena sea estable, por lo que en lugar de utilizar el canal visible se empleó el de infrarrojos (iluminando la escena con lámparas de este tipo y colocando filtros en los objetivos de las cámaras). De esta forma el contenido proyectado en las pantallas o las variaciones de la iluminación convencional del entorno son invisibles para el sistema de visión. Con un total de 9 cámaras se recuperaba la posición del participante en el plano de la sala, las siluetas que eran proyectadas en las pantallas y la posición que tocaba el usuario en la pantalla.

2.4.4.2. Otras técnicas de visión

En ámbitos como los deportes, el cine y la rehabilitación se utilizan sistemas de seguimiento de marcas reflectantes mediante visión artificial. Sistemas como Vicon, Qualisys o SIMI además de ser muy caros requieren de personal experto para su manejo [14] por lo que los hace impracticables para el ámbito que nos ocupa. Esta misma idea se ha empleado, en el ámbito de la rehabilitación física, con equipamiento de bajo coste basado en catadióptricos adheridos con velcro y cámaras web con filtro de infrarrojos [46].

Por último, merece la pena destacar el uso del seguimiento facial para el control de un sistema musical adaptativo [96].

2.5. Material audiovisual y correspondencia

El contenido hace referencia al tipo de material visual y auditivo, mientras que la correspondencia hace referencia al tipo de relación que se establecerá entre la entrada (el gesto: movimiento, sonidos vocales, manipulación directa) y el resultado (en forma de imagen, sonido y vibración), lo que determinará en gran medida el tipo de interacción que se llevará a cabo. Es muy importante

realizar correctamente esta elección, de ella dependerá que seamos capaces de captar y mantener la atención del participante, especialmente en personas con afectaciones severas.

2.5.1. Correspondencia en las propuestas estudiadas

En general la correspondencia uno-a-uno es la más empleada (véase la Sección 2.3.3 para tipos de correspondencia). Por ejemplo, en todos los experimentos llevados a cabo por Phil Ellis se ha empleado este tipo de correspondencia. Soundbeam mide la distancia desde el obstáculo más cercano (una dimensión), a esta distancia se le hace corresponder una nota MIDI que, en todos sus experimentos, envía a un módulo de sonido que sintetiza la nota reproducir. No deja tan claro, en el caso de la salida visual, que tipo de correspondencia emplea.

Brooks, al emplear un abanico más amplio de dispositivos y técnicas de interacción, no siempre da todos los detalles acerca de la correspondencia empleada. Algunas de configuraciones empleadas con más éxito [18, 12] incluyen la figura de un cohete espacial que rotaba a partir de los movimientos laterales del niño y el control de un foco con cabezal motorizado cuya orientación también respondía a los movimientos laterales de la cabeza del participante (ambas configuraciones con correspondencia uno-a-uno). Esto le lleva a observar que una respuesta que genera movimiento a partir de movimiento resulta muy estimulante (en este caso gesto y respuesta comparten la modalidad visual). Otra propuesta incluye una aplicación de pintura interactiva [12] (véase la Figura 2.4), donde el usuario pinta en la pantalla a partir de un umbral de velocidad (ajustable) a partir de la velocidad absoluta de la parte del cuerpo en movimiento (mediante la técnica SMI antes comentada), el rango dinámico del movimiento corresponde directamente con el cambio de color y la saturación [50].

En otros proyectos como IntelliVision (véase la Sección 2.2.4.1) la correspondencia es más compleja: incluye una respuesta inmediata a la hora de hacer aparecer notas en la pantalla que más tarde (memoria) serán interpretadas. Instrumentos musicales virtuales estudiados como VMI (véase la Sección 2.2.4.2) o el “Adaptative Musical Instruments” (véase la Sección 2.2.4.3) que se basan en zonas de píxeles activas hacen corresponder un posición en el espacio (2D) a un evento sonoro.

En todos los sistemas mencionados la correspondencia es estática, en el sentido que una vez determinados los parámetros de configuración (manualmente) el comportamiento no varía de forma automática en respuesta a las acciones ejecutadas por el usuario. Sin embargo se ha visto [24] la necesidad de adaptar esta configuración a lo largo de una sesión. Por lo que, sería in-



Figura 2.4: Aplicación de pintura interactiva [12].

interesante poder disponer de técnicas que realizaran esta adaptación de forma autónoma. Esto es precisamente lo que hace el proyecto MEDIANTE (véase la Sección 2.2.3).

En este proyecto se intentan evitar los comportamientos repetitivos y obsesivos de los niños autistas modificando el comportamiento en función de las acciones realizadas por el usuario durante el transcurso de una sesión. Para hacerlo posible no sólo es necesario registrar las acciones del usuario, sino que además hace falta inferir en que estado se encuentra éste. De ello se encarga el módulo *Signature Analyzer*. Éste construye un perfil del usuario y reconoce sus patrones repetitivos. El análisis, que se realiza en tiempo real, tiene en cuenta la información de los múltiples sensores y, además de permitir actuar sobre el comportamiento del sistema, almacena estos datos para su posterior estudio.

El funcionamiento consiste en que la primera vez que el usuario hace un gesto o sonido particular, el sistema lo almacena como ruido. La segunda vez que aparece la misma secuencia ésta es detectada e identificada como un patrón. Las veces siguientes la secuencia se reconoce directamente. Es decir, acaba generando una representación del repertorio del comportamiento del usuario o una firma. Por ejemplo, si el usuario camina en círculos, los sensores bajo las baldosa y las cámaras detectarán esta actividad e informarán al *Signature Analyzer* el cual reconocerá el movimiento circular al observar múltiples ocurrencias del mismo patrón (o similar) en las secuencias de datos.

Otro módulo denominado *Decision Maker* se encarga de modificar el comportamiento global de MEDIANTE con el objetivo de establecer, mantener y desarrollar un diálogo con el usuario. En términos generales, este módulo modela el comportamiento del usuario y aplica distintas reglas en función del estado de éste, del estado del sistema y de las tendencias del usuario y del sistema. Se basa en los cambios de actividad y niveles de repetición observados a medio plazo y hacia donde se dirige dicha repetición (modalidades visual, sonora y táctil) para alterar de forma gradual los parámetros de control.

El estado del usuario se basa en sus niveles de actividad y las ocurrencias de los patrones más repetitivos. Se distinguen los siguientes estados: repetición, repetición sostenida, actividad e inactividad sostenida.

La complejidad de la respuesta se modula ajustando distintos parámetros en los módulos de estímulo. Para cada modalidad (visual, sonora y táctil) se definen cuatro tipos de correspondencia ordenadas de forma creciente según su complejidad. Empleando notación musical, éstas se han denominado: imitación, eco, comentario y contrapunto. En todas ellas la respuesta depende de la acción del usuario, pero mientras contrapunto produce una respuesta más mediada, débilmente acoplada a la entrada; la imitación directamente imita o amplifica las acciones del usuario. También es posible ajustar la complejidad a través del parámetro riqueza. Éste afecta a la intensidad percibida y proporciona variaciones más sutiles que el cambio de tipo de correspondencia. El tipo de correspondencia y la riqueza constituyen una dimensión de la complejidad. La otra viene dada por como se interrelacionan las modalidades, yendo desde una configuración en modalidad propia (por ejemplo los sensores de visión provocan respuestas visuales) hasta modalidades cruzadas (la emisión de un sonido provoca una respuesta en forma de imagen). En teoría, en el nivel más elevado de complejidad, cualquier tipo de estímulo puede provocar respuesta en cualquiera de las modalidades.

Se definen, además, varios estados de funcionamiento para encargados de mantener o hacer evolucionar el tipo de respuesta. Éstos son:

- *Plateau*. Estado más simple. Se utiliza en la fase inicial para ayudar al

usuario a familiarizarse con el entorno. La transición a otro estado se efectúa tras observar cambios significativos y transcurrido un lapso de tiempo mínimo.

- *Regulate*. Regula el nivel de complejidad que el sistema exhibe. Se activa cuando el usuario muestra un comportamiento significativamente repetitivo. Su propósito es reducir la complejidad de la respuesta hasta que el usuario se sienta cómodo.
- *Explore*. Permite que nuevos elementos sean introducidos, como por ejemplo correspondencias más complejas. Se activa cuando el usuario muestra una actividad sostenida que no es repetitiva.
- *Tease Primary*. Se activa cuando el usuario lleva un rato inactivo (inactividad sostenida) e intenta captar nuevamente su atención generado estímulos no solicitados. Éstos se envían a través de la modalidad en la que se ha observado más activo al usuario (modalidad primaria).
- *Tease Secondary*. Se activa cuando el usuario esta activo pero obsesionado generando un tipo de respuesta o repitiendo un gesto o una vocalización una y otra vez (repetición sostenida). El sistema activa una propuesta a través de una modalidad distinta a la primaria. Así se intenta distraer al usuario de su obsesión y ganar su interés mediante otro modo de interacción.

Para concluir este apartado, indicar que el silencio es también una de las respuestas posibles del sistema y suele corresponder a la no actividad del participante o a cierta ubicación en el espacio. Según Brooks [24] es importante que exista esta zona neutral para que los participantes con discapacidades profundas aumenten su comunicación con el facilitador y puedan tener momentos de descanso.

2.5.2. Contenido

La elección del contenido en las distintas modalidades –visual, sonora y táctil– es crucial. En nuestra experiencia hemos constatado la motivación que ejercen ciertas imágenes o sonidos sobre niños y adultos con parálisis cerebral. Normalmente se trata de motivos que resultan familiares al individuo como imágenes de los miembros de su familia o entorno donde viven; música o sonidos que suelen escuchar habitualmente en su casa o bien aficiones que han desarrollado hacia objetos (juguetes, automóviles, animales, etc.) o personajes (protagonistas de cuentos infantiles, cantantes, actores, etc.). Estos elementos son empleados habitualmente en centros de educación especial para facilitar

la motivación del estudiante a la hora de realizar una tarea. Son las personas cercanas al individuo las que conocen y por tanto pueden asesorar en la elección de éstos.

En este sentido Brooks [24] propone la colaboración del facilitador en esta labor para identificar el tipo de contenido que puede ser apto para el participante; la idea es que éste proporcione diversión, juego y/o experiencias creativas. Esta información debe recogerse en un perfil personal, junto con otros parámetros de configuración del sistema, que evoluciona con el tiempo para optimizar el interés y disponer de una medida de evaluación [50, 17].

La modalidad también es importante puesto que algunos individuos puede tener deficiencias a nivel sensorial. Van Leeuwen, tras realizar pruebas con ancianos [125], hipotetiza que el valor añadido del contenido visual está relacionado con factores motivacionales y afectivos que, a su vez, activan la curiosidad cognitiva. Aunque este rol puede cambiar en función de la capacidad perceptiva y motriz del participante. Sí parece estar más clara la relación entre la presencia de imágenes y la belleza percibida. Brooks indica en [17] que para niños con discapacidades múltiples y profundas el nivel de respuesta ante los un estímulo visual o auditivo depende de cada participante, aunque también afirma [54] que la inmersión en un entorno visual puede tener un potencial similar que en un entorno sonoro. En el ámbito de la estimulación, nos indican [79] que los sonidos unidos a los movimientos o desplazamientos que sufren las imágenes dentro del campo visual del niño desempeñan un papel importante en la estimulación.

2.5.3. Contenido visual

En general parece ser que las imágenes contrastadas y cambiantes son las que mejor son captadas por las personas gravemente afectadas [25, 79]. Es la contraposición de dos tonos de color saturados y contrastados o el movimiento de la imagen lo que consigue captar la atención¹⁹. Si el contenido tiene que ser contrastado, será necesario atender a las restricciones que existen a la hora de combinar colores para que estos sean claramente visibles y el contraste apreciable. En este sentido hay que tener en cuenta el ojo humano tiene algunas limitaciones a la hora de ver ciertas combinaciones de color, por ello es recomendable emplear colores complementarios puesto que poseen alto nivel de contraste.

En relación a la semántica de la imagen, parece ser que las formas abstractas son más apropiadas en ciertas discapacidades como el autismo [100], donde una imagen concreta (familiar, perro, coche, etc.) puede llevar asociadas

¹⁹De hecho esto es algo que muchos hemos experimentado, ya que por ejemplo, es fácil detectar una luz intermitente cuando entra dentro de nuestro campo de visión.

ciertas connotaciones. En este caso, si al usuario le sugieren algo agradable, puede que ayude a motivarle y se muestre más partícipe, aunque también podría llegar a ser contraproducente si el usuario se muestra sobreexcitado ante cierto contenido. En el otro extremo también puede ocurrir que cierta imagen provoque emociones negativas en el participante y por tanto sea reticente a participar o incluso sienta miedo o malestar. Por ello, en la mayoría de literatura estudiada, y especialmente cuando se tratan participantes autistas, el contenido abstracto suele ser la elección más común. En ámbitos como en el caso de niños con plurideficiencias también se han obtenido resultados [80] empleando imágenes concretas de personas y objetos del entorno del participante. En personas sin discapacidad de edad avanzada se ha visto [12] que existe preferencia hacia los contenidos audiovisuales naturales o reales.

La imagen del propio usuario o la silueta de éste también se ha empleado como contenido en algunos experimentos [25, 50, 1, 106] a veces combinada con efectos visuales. Bien por necesidad de tener una referencia para poder realizar la interacción cuando se emplea visión artificial y zonas sensibles que el usuario debe accionar (como por ejemplo el caso de los juegos de EyeToy), o bien como material visual algunas veces tratado mediante efectos de vídeo [50, 25].

Por último también existen experiencias en el uso de visualizadores musicales con personas ancianas [40] y también para visualizar el sonido para personas sordas [11]. Empleando el sistema Soundbeam como dispositivo de entrada se ha visto que estos gráficos tienden a fomentar la variabilidad de los movimientos, la exploración y la sensación de pintar con sonido.

2.5.4. Contenido sonoro

En experimentos con niños con discapacidad visual cortical y parálisis cerebral [25] se ha visto que lo que más atrae la atención son los cambios más que el contenido en sí. En [79] se propone, para estimular el área auditiva, la reproducción de secuencias que alternen sonido con silencio, así como pulsaciones sonoras rítmicas y monótonas, sonidos impactantes, breves sonidos ambientales, etc.

El sonido instrumental también se ha venido utilizando extensamente, especialmente con el sistema Soundbeam. Recordemos que éste genera notas musicales que se envían a través del protocolo MIDI, por lo que normalmente debe conectarse a un módulo de sonido²⁰ constituyendo un sintetizador mu-

²⁰Un módulo de sonido es un instrumento musical electrónico que no posee interfaz y que debe ser tocado mediante un controlador externo. La comunicación con el módulo de sonido se suele hacer mediante protocolo MIDI. Muchos de estos módulos respetan la especificación General MIDI que, entre otras cosas, define un conjunto de 127 instrumentos estándares.

sical, con la particularidad que en lugar de teclado se toca en el aire. De esta forma queda bastante acotado el tipo de sonidos empleados en este caso. Por ejemplo, en una experiencia con Soundbeam [125] se indica que se utilizaron sonidos de campanas, de orquesta o de fantasía, además, todos ellos contenían bajas frecuencias para que fueran más efectivos a través del dispositivo táctil (véase la Sección 2.5.5). Otros sistemas no basados en Soundbeam como IntelliVision [25], VMI [71] y “Adaptive Use Musical Instruments” [96] también emplean sonidos instrumentales.

La escala musical pentatónica se ha empleado en instrumentos para musicoterapia [1, 65] ya que proporciona una estética apropiada para el desarrollo, confortable y estable; además asegura que no haya errores armónicos, proporcionando al participante una experiencia orientada al éxito.

Ellis [41] propone la utilización de un micrófono frotándolo sobre distintas superficies para obtener diferentes resultados en función de la textura, el tipo de movimiento (velocidad, dirección, etc.) y los efectos sonoros digitales que se apliquen. Ellis también promueve [40] el uso de la voz y los sonidos vocales como fuente sonora combinada con el uso de efectos sonoros digitales. Concretamente, propone emplear la reverberación para desinhibir al participante y fomentar la vocalización, y el eco para promover la escucha atenta. Además, el eco de larga duración permite crear masas sonoras complejas compuestas de varias emisiones sucesivas que se van acumulando.

El tipo de sonidos empleados en el proyecto MEDiate depende del nivel de complejidad que el sistema decide aplicar en cada momento (en función de la actividad anterior del niño como se ha visto en la Sección 2.5.1 sobre la correspondencia). Por ejemplo, al detectarse pisadas en las baldosas en el nivel de complejidad más bajo, produce un sonido similar al de pisadas sobre gravilla –sonido concreto–, a medida que la complejidad aumenta los sonidos van adquiriendo un carácter más abstracto que pueden ir desde chapoteos en el agua hasta voces cantando. En el caso del dispositivo *tune fork* (véase la Figura 2.5) los sonidos intentan recordar las diferentes texturas de que está compuesto este dispositivo; desde sonidos naturales a sonidos de percusión. También se utilizó efecto de eco con la voz del participante. Todo el sonido se emitía a través de un sistema 3D compuesto por 8 altavoces capaz de distribuir y localizar el sonido.

2.5.5. Contenido táctil

En el contexto de este documento, cuando hablamos de contenido táctil hacemos referencia únicamente a aquel que se puede transmitir a través de transductores de vibración, quedando excluidos otros sistemas como dispositivos hápticos u otros para presentar imágenes y contenido gráfico en relieve



Figura 2.5: Dispositivo *Tune Fork*. Proyecto MEDiate [101].

(normalmente orientados a la población invidente).

La mayoría de estos transductores táctiles funcionan de forma similar a un altavoz, sólo que trabajan en rangos de frecuencia muy bajos (entre 20 y 150Hz dependiendo del modelo). Para que el usuario pueda notar estas vibraciones, se acoplan a la silla, cama u otro mueble donde se encuentre el usuario.

Este tipo de dispositivo se ha empleado generalmente para fomentar la relajación y la atención [25, 39], empleando en algunos casos [125] música calmada mezclada con tonos sinusoidales de baja frecuencia (entre 40 y 120Hz), o como estímulo adicional para reforzar la asociación de causa-efecto [38].

2.6. Efectos terapéuticos

No cabe duda que si conseguimos fomentar la participación en una persona con discapacidad mental severa –que normalmente permanece pasiva– esto pondrá en marcha una serie de mecanismos que favorecerán su desarrollo. Este tipo de personas suelen estar retraídas, con muy pocas capacidades para la comunicación, lo que las aísla de su entorno. Poner a su disposición un canal que permita de forma efectiva la comunicación, la expresión de sentimientos, de estados de ánimo e incluso la creatividad contribuirá, sin lugar a dudas, en la mejora de la calidad de vida de esta persona.

A la hora de hacer una propuesta concreta es necesario determinar una serie de parámetros como el entorno, la tecnología empleada, la configuración del sistema, el material audiovisual, etc. Lo ideal sería poder pedir la opinión del usuario. Por desgracia, en nuestro caso muchas veces no resulta posible. En efecto, la mayoría de las personas a las que van dirigidas estas propuestas no poseen una capacidad comunicativa suficiente: no pueden expresar cuales son sus gustos, ni siquiera si aquello cierta propuesta les ha gustado más respecto

a otra. Por lo tanto, las respuestas las deberemos obtener por medio de la observación ligada a un conocimiento muy cercano de la persona.

Veamos en primer lugar que efectos beneficiosos se han observado en distintos experimentos para luego mostrar de que formas se desarrollará una sesión y finalmente adentrarnos en las técnicas de análisis empleadas.

2.6.1. Efectos observados

Kathryn Russell [114] ha identificado algunas áreas de desarrollo experimentando con niños con discapacidad utilizando Soundbeam:

- **Consciencia estética.** Incluye la capacidad de elegir y juzgar que sonidos o movimientos seleccionar y manipular.
- **Imaginación.** Anticipando sonidos y movimientos mucho tiempo antes, quizá una semana antes de las sesiones de terapia.
- **Habilidades de escucha.** Escuchando lo que ocurre al moverse o quedarse quieto.
- **Habilidades para escoger.** ¿Qué sonido escogeré? ¿Qué parte de mi cuerpo moveré? ¿Dónde lo moveré?
- **Habilidades conceptuales.** Especialmente “inicio, mitad y final”. ¿Cómo empezaré? ¿Qué haré entonces? ¿Cómo finalizaré? Conceptos musicales específicos como fuerte, flojo, rápido, lento, variación; los conceptos de unión, palabras con movimiento, notando el color del tono, etc.
- **Habilidades de planificación motora.** ¿Qué movimiento debo hacer ahora para producir...?
- **Cognición reflexiva.** ¿Qué sensaciones me transmitía la pieza que creé? ¿Qué podría haber hecho diferente? ¿A qué me recuerda?
- **Habilidades memorísticas.** ¿Puedo recordar qué sonido me gustó la última vez?, ¿quiero volver a usarlo?
- **Orientación espacial.** ¿De dónde viene ese ladrido?
- **Habilidades lingüísticas.** Describiendo qué hice y cómo me siento, poniéndole título a mi trabajo.
- **Explorando una hipótesis.** Recuerdo que si me muevo de esta forma, se produce un tipo de sonido. Si me vuelvo a mover de esta forma, ¿obtendré el mismo resultado?

- **Habilidades sociales.** Esperando turno, compartiendo en espacio sensible para producir improvisaciones en grupo.
- **Confianza en uno mismo.** Esto es algo que yo puedo hacer.

Ellis [122] enumera una serie de respuestas en niños con discapacidad profunda:

- Interpretando, escuchando, verbalizando, componiendo con sonido.
- Mostrando a menudo “resonancia estética” a través de expresiones faciales.
- Estando involucrado activamente durante largos períodos de tiempo.
- Revelando una habilidad para concentrarse nunca vista antes.
- Empezando a descubrir, explorar, expresar y comunicar las propias sensaciones.
- Realizando respuestas físicas significativas - movimientos y gestos que hasta el momento no habían sido vistos o no se habían realizado antes de forma independiente.

Ellis propone nueve categorías de progreso a identificar [36]:

1. De involuntario a voluntario.
2. De accidental a intencionado.
3. De indiferente a interesado.
4. De confinado a expresivo.
5. De aleatorio a con propósito.
6. De grueso a fino.
7. De exploratorio a preconcebido.
8. De aislado a integrado.
9. De solitario a individual.

Según Ellis [37] el objetivo ha de ser de que forma cada individuo puede desarrollar mejor su potencial y que esto puede incidir en el desarrollo físico y comunicativo del paciente. También afirma que gracias a la terapia sonora se pueden producir adelantos en el desarrollo psicomotriz, pero que estos no pueden ser la base de ningún programa.

En otro estudio con Soundbeam [133] se ha observado un incremento de los movimientos exploratorios en dos niños de 5 y 6 años con parálisis cerebral severa.

Desde un punto de vista de la musicoterapia, Soundbeam puede permitir [71] entrenar la asociación causa-efecto, potenciar la capacidad del niño de actuar sobre su entorno y mejorar el rango de movimiento. Además, se le supone la habilidad de mejorar la atención, la auto-expresión (comunicación expresiva), la imitación (comunicación receptiva) y la habilidad social.

Experiencias con VMI (véase la Sección 2.2.4.2) sugieren que este sistema crea un entorno que es adecuado para el desarrollo y potencia la exploración activa y la participación, que es clave para potenciar las habilidades sociales y comunicativas, motrices y cinestéticas, el desarrollo cognitivo y el crecimiento emocional. Se ha observado también como el sistema contribuye a favorecer las sensaciones de orgullo y maestría que son poderosos elementos motivadores. La posibilidad de compartir su trabajo les proporciona la sensación de sentirse valorados. Las principales categorías de desarrollo se resumen con las siguientes frases:

- Desarrollo de la consciencia visual: ver y ser visto a través de una experiencia reflectiva. El sistema proporciona una experiencia de relación (permite establecer contacto visual con uno mismo –a través de la pantalla– y con el musicoterapeuta) creativa y basada en el juego. Proporciona además una oportunidad de desarrollo, controlando objetos del entorno, los cuales son normalmente inaccesibles a los niños con parálisis cerebral.
- Desarrollo de la consciencia auditiva: aprendiendo y construyendo conceptos a través de una estructura musical. El sistema proporciona experiencias musicales positivas (gracias al uso de la escala pentatónica) que se centran en las capacidades del niño y no en sus discapacidades.
- Desarrollo de la consciencia cinestética: fomentando la participación corporal a través de una experiencia expresiva musical multisensorial. El sistema proporciona oportunidades para la integración sensorial necesaria para el desarrollo normal de un niño. Concretamente contribuye a proporcionar una estimulación adecuada y clara, a promover la actividad física, a proporcionar un desafío de dificultad apropiada, a fomentar

un comportamiento adaptativo y participativo y está orientado a que el participante cree la música él mismo.

- Desarrollo de la auto-consciencia: restaurando la auto-imagen y expresando sentimientos a través de una experiencia musical. Los niños se (re)descubren a ellos mismos a través de esta experiencia, donde tienen la posibilidad de mostrar su verdadero yo, para ser vistos y oídos.

Un caso de estudio [50] con un chico con retraso mental profundo²¹, tras un total de 28 sesiones combinando distintos dispositivos como Soundbeam, cámaras, pulsadores y micrófonos hizo emerger las siguientes cuestiones: “mi expresión causa efectos divertidos” y “necesito que me guíen, por lo menos al principio”. Según las personas cercanas al chico, éste gana en concentración y autoestima.

En otro caso de estudio [105] –llevado a cabo con 4 niños con discapacidades múltiples y profundas de edades entre 4 y 6 años empleando sensores de infrarrojos y focos de luz motorizados– emergieron las siguientes cuestiones: “sensación de estar haciendo” y “aquí yo tengo el control”. Todos los niños mostraron mucha concentración e intencionalidad.

Según la National Autistic Society del Reino Unido el proyecto MEDiate (véase la Sección 2.2.3) tiene potenciales en las siguientes áreas:

- Reducción de estrés y ansiedad. Alivio de depresión.
- Dar voz a personas con autismo y posibilitar el diálogo (no verbal).
- Promoción de la creatividad y del control físico.
- Diversión
- Incremento de la sensación de bienestar y de la calidad de vida.
- Aumentar el conocimiento y entendimiento de los niños con espectro autista.
- Fomento de la independencia.

2.6.2. Desarrollo de las sesiones experimentales

Distintos son los escenarios empleados para llevar a cabo la experimentación. Desde instalaciones a medida como en el proyecto MEDiate, pasando

²¹16 años reales, con una edad mental estimada de 2,5 años, hiperactivo y diagnosticado con los síndromes de Angelman y de Pader Willy

por entornos inmersivos de realidad virtual [18] hasta las instalaciones ad hoc en centros hospitalarios [25], de atención a personas con discapacidad [38], residencias de ancianos [39] o domicilios privados [14].

Es también habitual en este tipo de experimentos la participación de, como mínimo, el investigador principal y un facilitador. El facilitador es una persona cercana al participante (maestro, cuidador, tutor, padre, etc.) que está al tanto de las necesidades del usuario, se encarga de facilitar la comunicación con él, de seleccionar el material multimedia más apropiado y de observar e interpretar sus respuestas. Esta figura es esencial, debido al conocimiento que tiene del usuario con el que trabaja, lo que le permite interpretar las respuestas de éste, cuestión muy importante en el caso de afectaciones cognitivas moderadas o severas.

El nivel de intervención del investigador, facilitador y otros implicados en el desarrollo de la sesión puede tener distintos grados:

- **No intervención.** Un ejemplo lo constituye el proyecto MEDiate (véase la Sección 2.2.3) donde al participante se le sitúa frente a la puerta de entrada de la instalación y se espera que éste, de forma autónoma, entre y empiece a interactuar.
- **Intervención ligera.** Muchas veces, los trastornos mentales severos de los participantes hacen que su capacidad de iniciativa y comprensión sean muy reducidas, por lo que el facilitador debe guiar al usuario [21] en el inicio del uso del sistema para que éste poco a poco entre en el juego. Es decir, la exploración del entorno virtual en muchas ocasiones comienza por una acción motriz (eferente) involuntaria o propiciada por el facilitador que proporciona un estímulo (aferente) sonoro y/o visual. Éste, a su vez, genera una reacción intencional (consciente) resultando en una motivación interactiva que cierra el bucle [18]. Después se deja al participante para que continúe la exploración por su cuenta. En esta situación se busca un estado similar al estado de flujo [33] que puede observarse, por ejemplo, cuando un personas está completamente concentrada jugando a un videojuego [12].

Ellis, por ejemplo, indica [40] que la intervención debe ser lo menos invasiva posible y que el papel del investigador es el de optimizar el entorno y seleccionar el material sonoro más apropiado. Este mismo autor, también propone un itinerario a realizar en sus sesiones con el sistema Soundbeam [125]. Concretamente:

1. Empezar la interacción social o solitaria con el micrófono empleando efectos como la reverberación, el eco y el eco con cambio tonal.

2. Interacción social o solitaria con el sistema Soundbeam a partir de movimientos laterales de brazos y manos.
 3. Relajación empleando música calmante a la que se le añaden bajas frecuencias.
- **Sesión guiada.** Al trabajar con herramientas como VMI, pensada para las sesiones de musicoterapia, la investigación se ha realizado desde el punto de vista de esta disciplina [1]. Esto implica que las sesiones son lideradas por el musicoterapeuta quien, en sesiones individuales, intenta motivar al usuario para que participe empleando técnicas como la improvisación clínica o actividades orientadas a la tarea.

Las sesiones suelen finalizar cuando el usuario se muestra cansado [71], empieza a perder el interés [105] o porque expira el tiempo asignado [22]. En general, la duración de las sesiones suele estar entre los 5 y los 45 minutos, aunque la franja más usual está entre los 15 y 30 minutos.

2.6.3. Técnicas de análisis

La literatura estudiada recoge, principalmente, la valoración de tipo cualitativo realizada a partir de la observación de las sesiones directamente y a través de grabaciones en vídeo. En general estas valoraciones son llevadas a cabo por investigadores en psicología, pero contando, normalmente con la ayuda de los facilitadores que ayudan a interpretar las observaciones y a sugerir preferencias de los usuarios [105].

La grabación de la sesión normalmente se efectúa con una cámara enfocando al rostro del usuario para registrar los gestos faciales. En los casos en los que existe retroacción visual suele ser habitual que ésta también quede registrada, por lo que existen experiencias con el uso de varias cámaras [18] de modo que puedan tenerse distintos planos de la interacción y del contenido y también espejos situados tras el usuario para recoger, con una sola cámara, contenido y rostro [5].

Resumimos a continuación algunos enfoques de análisis encontrados en la literatura relacionada:

2.6.3.1. Observación directa

La observación directa ha sido la principal estrategia para recopilar datos durante décadas. Se ha utilizado por investigadores de disciplinas como la etnografía, la antropología, la psicología y otras. A raíz de la generalización de los grabadores de vídeo, las sesiones experimentales podían ser registradas y visionadas las veces que fuera necesario con el fin de anotar las conductas y

permitiendo que los investigadores consultaran de nuevo los acontecimientos reales en vez de confiar sólo en las anotaciones y la propia memoria [112].

2.6.3.2. Cuestionarios y entrevistas

Dada la tipología de usuarios –con trastornos mentales severos– no suele resultar posible emplear técnicas que requieran de la colaboración de éste para responder preguntas, dar una opinión subjetiva sobre una experiencia ni responder un formulario. Es cierto que se han empleado cuestionarios o entrevistas con usuarios con problemas de conducta [12], ancianos [38], niños hospitalizados [20] o personas con derrame cerebral [108], pero no con personas con discapacidades mentales severas.

2.6.3.3. Análisis por capas (*layered analysis*)

Esta técnica, utilizada por Ellis [37, 125], básicamente consiste en, a partir de la grabación de cada sesión, hacer un montaje en episodios para cada individuo. Con ello se pretende introducir un cierto grado de imparcialidad, objetividad y racionalidad a las observaciones. La idea es recoger una serie de indicadores que desvelen múltiples facetas cualitativas de fenómenos potencialmente significativos desde un enfoque etnográfico. Consta de varias etapas:

1. Grabación original. La sesión de terapia sonora se graba íntegramente en vídeo. Normalmente la cámara se sitúa ante el usuario para recoger sus gestos y expresiones faciales. La técnica original indicaba que una vez seleccionadas las secuencias más relevantes, la cinta se podía reutilizar.
2. Vídeo maestro. En la grabación original se identifican los comportamientos más significativos y se copian al vídeo maestro que crece de sesión en sesión creando, de forma gradual, un retrato del comportamiento y respuesta de cada individuo. En cada corte se anotan los tiempos y las observaciones de carácter general y específicas (por ejemplo el tipo de sonido empleado, información específica acerca del movimiento físico, un largo periodo de silencio, un gesto o expresión facial, observaciones que se repiten en cada sesión, etc.). Habitualmente se extraen entre 6 y 20 cortes del vídeo original.
3. Visionado detallado. Cuando en el vídeo maestro se han acumulado unas 10 o 12 sesiones, se procede a un visionado detallado y se completan las anotaciones con los comportamientos significativos o repetitivos que se observen. Se incluyen las notas de campo tomadas por el investigador o profesor/cuidador. También se pueden incluir las aportaciones de otros

colaboradores profesores, cuidadores, etc. Estas observaciones pueden influenciar la estructura y contenido de sesiones futuras.

4. Vídeo por capas. Se separan y se graban independientemente (en capas) los diferentes componentes de actividad (respuestas a distintos sonidos, ciertos tipos de movimiento físico, respuesta auditiva, expresión facial, etc.) para cada una de las tres secciones de la sesión. De esta forma es posible estudiar de forma cronológica la respuesta a cada uno de estos componentes de actividad. Esto permite observar el comportamiento y proporciona un retrato detallado y exhaustivo. Propone distinguir entre 6 o 7 capas.
5. Vídeo de capturas. A partir del vídeo maestro se seleccionan ejemplos significativos a intervalos periódicos (por ejemplo cada mes) y se graban de forma secuencial para proporcionar una secuencia cronológica comprimida que facilite observar y mostrar los progresos conseguidos. Este material puede ser muy ilustrativo y útil para mostrarlo a profesores, a padres, en un congreso, etc.

Opcionalmente puede elaborarse un vídeo resumen a partir de pequeños fragmentos de algunas sesiones hasta completar una duración máxima de unos 10 minutos. Los cortes escogidos no tienen porque incluir todas las sesiones. De esta forma se puede ver en un espacio corto de tiempo la evolución a lo largo de varios meses (similar a cuando se graba una flor a intervalos largos de tiempo y es posible observar como eclosiona). Suele ser muy ilustrativo de la evolución del individuo.

La técnica descrita puede ser costosa en tiempo, puesto que requiere la edición y el visionado detallado de los vídeos. Según el autor las observaciones son de tres tipos y siguen el siguiente orden cronológico: **descriptivas, enfocadas y selectivas**. Al principio es difícil identificar que es significativo y que no lo es a la hora de escoger que cortes se añaden al vídeo maestro. A lo largo del tiempo resulta posible identificar secuencias de comportamiento repetitivas, por lo cual la observación acontece más enfocada. Concretamente cuando el autor observaba el vídeo a gran velocidad se hacían evidentes algunos patrones de movimiento relacionados con ciertos sonidos. Al cabo de unos meses de observación es posible ver líneas de desarrollo en el usuario, es cuando la observación se ha vuelto selectiva. Ello posibilita predecir la actitud del usuario ante diferentes tipos de sonidos.

Inicialmente esta técnica basada en vídeo implicaba una selección del material grabado y una preservación de parte de sólo éste por razones de economía. De un tiempo a esta parte, gracias a la proliferación de las cámaras digitales y los medios de almacenamiento masivos, es posible conservar todo el material

grabado. Esto, que Brooks denominó [22] *recursive reflection*, lo que permite poder analizarlo en un futuro –y volver a reflexionar sobre ello– en el caso que aparecieran nuevas técnicas de análisis u otras personas implicadas que pudieran aportar más datos.

2.6.3.4. Estrategias participativas

Intentan disminuir las limitaciones de los investigadores basadas, normalmente, en sus propios instintos y habilidades para comprender las interacciones entre el usuario y el sistema [105]. Para ello estas estrategias integran al facilitador y otros participantes (padres, tutores, otros cuidadores, etc.) en todo el proceso experimental; desde la selección del material, modo de interacción, etc., pasando por el ajuste en tiempo real del sistema, a la interpretación de resultados y la sugerencia de mejoras.

Por ejemplo en las investigaciones llevadas a cabo en [21] y [105] con niños con trastornos múltiples y profundos²² los facilitadores colaboraron en las siguientes tareas:

- Creación de un perfil básico del usuario basado en las habilidades, necesidades, preferencias y otras características. Información que es de gran ayuda a la hora de seleccionar contenidos y modo de interacción.
- Participación en la sesión para asegurar el bienestar de los participantes y determinar su finalización.
- Redacción de notas sobre la percepción del participante antes, durante y después de la sesión. Esto mejora la validez del estudio reduciendo errores debidos a interpretaciones precipitadas, simplistas e ingenuas [105].
- Visionado participativo del vídeo, a veces [50] incluyendo a los padres u otras personas cercanas.
- Contrastado de las anotaciones del vídeo. En este caso los investigadores realizaron, cada uno por separado, anotaciones detalladas sobre el vídeo basándose en criterios como: interacción del participante, configuración y respuesta del sistema a partir de expresiones faciales y otros gestos no verbales. Las anotaciones, realizadas con el software Anvil [64], fueron revisadas por los facilitadores.

²²En esta experiencia se utilizó un enfoque inductivo exploratorio y las sesiones se enfocaron a la percepción del entorno por parte del usuario, intereses específicos, auto-orientación y logros

Además en este estudio intentan generalizar aquellas expresiones comunes relacionando, normalmente, expresiones faciales con gestos de la cabeza o las extremidades.

En el caso del proyecto MEDIANTE (véase la Sección 2.2.3) también se adoptó un enfoque similar. Las sesiones se grababan en DVD y tras las mismas se entregaban copias a los padres y cuidadores de los niños así como a los psicólogos. A partir de las grabaciones varios psicólogos valoraron el nivel agencialidad (*agency*) a partir del nivel de complejidad de la exploración.

A los padres también se les pidió la colaboración valorando antes, durante y después de las sesiones cuestiones como: humor, disfrute anticipado y actual, agencialidad, comportamiento repetitivo, exploración, reacción general y respuesta sensorial (mediante una versión adaptada del *Sensory Profile Test* [35]). Otras variables que fueron analizadas mediante cuestionarios incluyen el comportamiento creativo (con técnicas que examinan la reconfiguración de elementos como colores y tonos sonoros), los comportamientos libres en el juego libre antes y después de las sesiones (con evaluadores que desconocían las hipótesis y el orden de las sesiones). Sin embargo se descartó codificar en detalle los movimientos estereotipados (muy comunes entre el colectivo autista), puesto que es complejo definir que se entiende por repetición, ya que puede ocurrir a niveles de comportamiento micro y macro.

Relacionado con el proyecto VMI, como herramienta pensada para las sesiones de musicoterapia que es, la investigación se ha realizado desde el punto de vista de esta disciplina [1]. Concretamente la investigación se basa en el paradigma cualitativo y su naturaleza es abductiva y descriptiva. Durante las sesiones experimentales se recogen grabaciones en vídeo y notas de campo. Del vídeo se seleccionan aquellos momentos más interesantes –bajo criterio de dos musicoterapeutas– que son posteriormente analizados. El análisis consiste en una descripción narrativa acompañada de reflexiones adicionales a la cual se la añaden las notas de campo. Empleando un software de análisis de contenidos²³ se delimitaron algunas categorías beneficio terapéutico y se extrajeron algunas conclusiones.

2.6.3.5. Anotación de vídeo

Durante la última década y debido a la cada vez mayor potencia y disponibilidad de almacenamiento de los ordenadores, las aplicaciones de anotación de vídeo han proliferado y actualmente están disponibles varias soluciones comerciales y libres [26, 48, 64, 59].

La funcionalidad básica de estas herramientas incluye la posibilidad de cargar un fichero de vídeo y anotar los acontecimientos a lo largo de la línea

²³QSR N6 <http://www.qsrinternational.com/products.aspx>

temporal. Los eventos pueden ser momentáneos (es decir, se producen en un período de tiempo corto) o duracionales (es decir, que tiene un principio y un final claramente separados en el tiempo).

El análisis y la anotación de una grabación de vídeo es una tarea que consume mucho tiempo [48] ya que requiere, como mínimo, visionar el vídeo una vez y suele ser frecuente tener que repetir una secuencia en repetidas ocasiones para asegurarse de los fenómenos observados. Además, para intentar limitar la subjetividad de las observaciones realizadas, suele ser conveniente que sean varias las personas que realicen la anotación de un mismo vídeo.

Todo lo anterior implica demoras para los investigadores y carga de trabajo adicional para los facilitadores (notemos que su intervención es esencial en algunas ocasiones). Además, en muchos casos, las limitaciones presupuestarias impiden que los facilitadores puedan dedicar tiempo a actividades de investigación (por ejemplo, anotando vídeos fuera de línea). De hecho, en nuestro contexto los facilitadores tiene que ocuparse de otras muchas tareas, por lo tanto, esta carga de trabajo adicional derivada de la anotación de vídeo puede llegar a ser prohibitiva.

Algunas de estas herramientas de anotación de vídeo también incluyen varias mejoras como la capacidad de mostrar varias secuencias de vídeo al mismo tiempo u otras que permiten acelerar el análisis y anotación de los vídeos como, por ejemplo, la integración de algunos datos recogidos automáticamente (*ej.* pistas de audio y datos de sensores) o derivados de análisis automáticos (*ej.* reconocimiento de expresiones faciales o de voz), e incluso la posibilidad de tomar anotaciones en línea [59]²⁴.

Otro ejemplo que permite acelerar el proceso de anotación son las herramientas VCode y VData de la Universidad de Illinois [48]. VCode es una herramienta de anotación, como Anvil, que está disponible para entornos Mac OS. Permite mostrar varios ficheros de vídeo simultáneamente y añadir tantos eventos (momentáneos y duracionales) como se deseen pudiendo indexarlos con una clave. Durante el proceso de anotación permite que varios anotadores trabajen simultáneamente en tiempo real y proporciona el nivel de acuerdo entre ellos mediante al cálculo del coeficiente kappa.

2.6.3.6. Grabación del uso (análisis de *logs*)

La técnica de grabación de uso, más conocida como análisis de *logs* o simplemente *logging*, se basa en registrar todas las actividades realizadas por el usuario con el sistema para su posterior análisis. Para ello es preciso de

²⁴Por desgracia, esta última mejora sólo está disponible como parte de un sistema propietario y no se puede obtener libremente. Además, no ha sido diseñado para ser utilizado mientras se está al cargo de personas con parálisis cerebral.

una aplicación secundaria que realice automáticamente esta labor que pase totalmente desapercibida por el usuario [47].

Por ejemplo en [12] se sugiere la utilización de medidas extraídas del análisis de movimiento mediante visión artificial (véase la Sección 2.4.4) para producir descripciones simbólicas de los movimientos realizados por el usuario; aunque los intentos de aplicación han sido infructuosos hasta el momento [105].

En el proyecto MEDiate (véase la Sección 2.2.3) no sólo se guardan los registros de actividad del usuario, sino que además se infiere un estado de actividad en el que éste puede estar. A pesar de esto, el autor no ha sido capaz de encontrar información publicada sobre los resultados en este sentido.

CAPÍTULO 3

Nuestras propuestas de técnicas y sistemas interactivos

En este capítulo se describen las técnicas y los sistemas interactivos para personas con parálisis cerebral desarrollados como parte de esta tesis.

Índice

3.1. Evaluación de la interacción	58
3.1.1. Interacción basada en conmutadores	59
3.1.2. Producción de texto	63
3.1.3. Control del dispositivo apuntador	63
3.1.4. Rendimiento de la interacción	65
3.1.5. Método de evaluación propuesto	66
3.2. Ratón Facial y Enable Viacam	73
3.2.1. Funcionamiento	73
3.2.2. Estado actual del proyecto	74
3.3. SITPLUS	75
3.3.1. Objetivos y criterios de diseño	75
3.3.2. Arquitectura del sistema	76
3.3.3. Módulos de visión artificial	82
3.3.4. Módulos de audio	82
3.3.5. Controles remotos Wii	83
3.3.6. Salida gráfica	86
3.3.7. Actividades	89
3.3.8. Estado actual del proyecto	92
3.4. Switch Viacam	93
3.5. Herramienta on-line de anotación de vídeo	95
3.5.1. Motivaciones	95
3.5.2. Requisitos	96
3.5.3. Arquitectura	97

Capítulo 3. Nuestras propuestas de técnicas y sistemas

58

interactivos

3.5.4. Interfaz de usuario	98
3.5.5. Integración con ANVIL	100

La primera contribución que se expone en este capítulo son algunas herramientas y técnicas para evaluar la calidad de la interacción de las personas con parálisis cerebral.

La segunda contribución que se incluye en este capítulo son los resultados del desarrollo de un sistema audiovisual interactivo creado específicamente para personas con parálisis cerebral. Estos resultados incluyen el propio software (llamado SITPLUS), que está disponible libremente y bajo licencia de código abierto, y la aplicación de diferentes técnicas de visión por computador para la interacción persona ordenador.

El proyecto SITPLUS se utiliza de forma regular en la APPC como una actividad integrada en la dinámica de funcionamiento del centro, lo que también da cuenta de su utilidad. Gracias a tratarse de una solución software, es posible tenerla disponible en varias salas distintas y muchos profesionales son capaces de ponerla en práctica.

En este capítulo se describe en primer lugar una propuesta de evaluación de la interacción persona-ordenador para personas con parálisis cerebral. A continuación el proyecto Enable Viacam del que se deriva el proyecto SITPLUS. Después se describe el proyecto SITPLUS y también se presenta el proyecto Switch Viacam que aprovecha parte de los esfuerzos dedicados al proyecto SITPLUS. Por último, se describe una herramienta para facilitar el análisis de las observaciones de las sesiones experimentales.

3.1. Evaluación de la interacción

Las personas con alguna discapacidad, y en particular las personas con parálisis cerebral, interaccionan con los ordenadores, bien mediante dispositivos estándar (los cuales pueden requerir ciertos ajustes), o bien mediante dispositivos adaptados o especialmente diseñados. Es lo que en conjunto se conoce como ayudas técnicas o tecnologías de apoyo.

Dada la diversidad de necesidades personales y tecnologías de apoyo, es necesario que exista un proceso de evaluación y ajuste de forma que estas tecnologías se adapten lo mejor posible al usuario. Este proceso de adaptación se realiza habitualmente basándose en el criterio cualitativo de los profesionales y, siempre que sea posible, de los propios usuarios. En general, este enfoque es suficiente en muchas ocasiones, aunque esta evaluación puede enriquecerse con métricas que ofrezcan unos criterios cuantitativos. El objetivo es que di-

chas métricas complementen la opinión subjetiva del profesional y ayuden a la elección de los dispositivos, aplicaciones y ajustes más adecuados para cada usuario.

Dada la diversidad de tecnologías de apoyo antes mencionada, se presentan distintas modalidades de entrada y dispositivos que son de uso común entre el colectivo de personas con parálisis cerebral y que tendremos en cuenta de cara a nuestra propuesta.

3.1.1. Interacción basada en conmutadores

Un conmutador, en su concepción más simple, se basa en un dispositivo mecánico que cierra un contacto eléctrico cuando es accionado por el usuario. Este evento, que consiste en el cambio de estado de un solo bit, se transmite al ordenador donde generalmente se utiliza para desencadenar una acción simple (*ej.* para una actividad de causa-efecto) o para permitir una interacción más rica a través de métodos de selección por barrido.

La simplicidad de este método de interacción permite que, prácticamente, cualquier parte del cuerpo que conserve movilidad voluntaria pueda utilizarse para realizar la interacción. No obstante, el acceso mediante conmutadores y sistemas de barrido es una de las últimas alternativas que habría que considerar puesto que este estilo de interacción limita enormemente la velocidad.

Dadas las severas limitaciones motrices que tienen muchas personas con parálisis cerebral y su simplicidad, el uso de conmutadores suele ser un recurso de uso bastante extendido.

En el mercado existe una gran variedad de diseños de conmutadores y dispositivos y sistemas que los emulan, entre los cuales destacamos:

- **Mecánico.** Consiste en un contacto eléctrico que se cierra cuando el usuario ejerce cierta presión. Los hay en forma de botón (véase la Figura 3.1), de pedal, de varilla, por torsión, etc.
- **De soplo/absorción.** Diseñado para personas con limitaciones severas de movilidad que conservan un buen control de los labios. Consiste en un tubo que el usuario puede activar o desactivar sorbiendo o soplando [90, p. 275].
- **De inclinación.** Está diseñado para cerrar el circuito cuando cambia de posición en el espacio. Usualmente está formado por un recipiente de mercurio que se sitúa sobre la cabeza o brazo de la persona que lo utiliza [90, p. 275].
- **Basado en sonido o voz.** La orden se desencadena cuando el usuario emite un sonido cualquiera o un fonema concreto.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

60



Figura 3.1: Conmutador mecánico de botón [69].

- Basado en visión artificial. Una cámara conectada al ordenador está pendiente de una escena concreta; cuando esta escena sufre algún cambio concreto o el usuario realiza un gesto concreto, se activa la orden. En el marco de esta tesis se ha desarrollado un sistema (véase la Sección 3.4) que encaja en esta categoría.

3.1.1.1. Actividades de causa-efecto

Un objetivo fundamental para las personas con parálisis cerebral es la realización de actividades que permitan establecer la relación causa-efecto y así poder asentar que la realización de determinados actos o movimientos tienen consecuencias sobre el entorno.

El aprendizaje de la relación causa-efecto emerge de forma natural en los niños sin discapacidad durante los primeros meses de vida. Para las personas con parálisis cerebral, puede ser necesaria la utilización de recursos específicos para su adquisición.

Los profesionales suelen emplear conmutadores en combinación con aplicaciones sencillas (ej. presentaciones hechas en Power Point) que responden de alguna manera cuando se presiona el conmutador (por ejemplo, reproduciendo un sonido o vídeo). También es común la adaptación de juguetes eléctricos o electrodomésticos para ser usados con un conmutador con el mismo propósito.

La secuencia de interacción para trabajar la relación causa-efecto es básicamente la siguiente:

1. **Sistema:** solicita al usuario que presione el conmutador. A veces esto significa mostrar algún tipo de invitación o simplemente una pantalla va-

cía. Normalmente los profesionales suelen ofrecer algún tipo de refuerzo adicional.

2. **Usuario:** presiona el conmutador. Dependiendo del caso, esto puede tardar unos segundos, minutos o no llegar a suceder jamás.
3. **Sistema:** reproduce una acción predefinida en forma de sonido, imagen, vídeo o incluso vibración.

3.1.1.2. Selección por barrido

La selección por barrido es una técnica que permite, mediante uno o más conmutadores, seleccionar elementos entre un conjunto (*ej.* opciones de un menú, carácter en un teclado virtual o casilla de un tablero de comunicación). En su modalidad más simple, las distintas opciones se van recorriendo a una velocidad determinada llamada velocidad de barrido. En el momento en que la opción deseada está resaltada, el usuario activa el conmutador para seleccionarla.

La velocidad de barrido es un parámetro crítico que debe ser ajustado cuidadosamente de manera que sea la mayor posible sin que ello implique una tasa inaceptable de errores. La regulación de esta velocidad se hace generalmente de manera experimental, aunque también se han propuesto técnicas para ajustarla automáticamente a medida que el usuario interacciona con el sistema [73].

Generalmente un conmutador genera dos tipos de eventos: cuando se pulsa y cuando se libera. Esta distinción permite dar mayor flexibilidad a la de construir sistemas de barrido ya que permite asociar comportamientos específicos a la condición de mantener el disparador activo (por ejemplo, distinguir entre una pulsación corta o larga o hacer que el barrido avance mientras el conmutador se mantiene pulsado).

Los métodos de barrido más comunes son:

- Barrido secuencial por tiempo (1 conmutador). Se establece una relación de orden entre los elementos del conjunto y se van recorriendo en ese orden esperando el tiempo de barrido hasta pasar a la siguiente. Cuando se encuentra seleccionado un elemento determinado, éste se señala convenientemente mediante vídeo inverso, un recuadro alrededor, etc. El usuario pulsa el conmutador para seleccionar el elemento activo en ese momento.
- Barrido secuencial dirigido (1 conmutador). Como el anterior pero el avance se hace mientras el conmutador se mantiene pulsado y la selección

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

62

al liberarlo. Otra variante consiste en avanzar a base de pulsaciones cortas y seleccionar mediante pulsación larga.

- Barrido secuencial con aceptación por tiempo (1 conmutador), al accionar el conmutador el barrido avanza, al soltarlo y pasado un tiempo determinado se activa el elemento seleccionado.
- Barrido secuencial dirigido (2 conmutadores). El primer conmutador sirve para avanzar al siguiente elemento, mientras que el segundo sirve para seleccionar el elemento deseado.

Además, los métodos anteriores admiten ciertas variantes:

- Barrido con retroacción auditiva. Cada vez que se selecciona un elemento reproduce un sonido asociado a éste (usualmente mediante un sistema de texto a voz). Este tipo de barrido facilita el acceso cuando la pantalla no se puede ver bien (por ubicación, posicionamiento del usuario o por deficiencia visual), cuando el usuario carece de lecto-escritura o cuando no es capaz de discriminar visualmente (*ej.* mediante el uso de iconos o pictogramas) los elementos.
- Barrido por bloques. En lugar de recorrer el conjunto final de elementos, éstos se agrupan en conjuntos (cada conjunto, a su vez, puede incluir otros subconjuntos) de forma que, en general, se acelera la selección de un elemento al no tener que recorrer todas las opciones posibles una a una hasta llegar a la deseada. Los barridos por bloques más comunes son:

- Por filas o columnas.
- Por grupos definidos de forma explícita (estableciendo previamente los elementos que forman parte de un mismo grupo) o implícita (*ej.* porque forman parte de un mismo contenedor de la interfaz gráfica de usuario).
- Binario. Muestra en todo momento dos grupos de elementos que se pueden seleccionar mediante los mecanismos de barrido vistos anteriormente o bien directamente si se dispone de dos conmutadores. Una vez seleccionado uno de los grupos éste se subdivide en otros dos subgrupos y así sucesivamente hasta llegar a los elementos individuales.

Las técnicas de barrido por bloques tienen la ventaja de acelerar el proceso de selección pero resultan más abstractos para el usuario final y es más complicado poder disponer de una retroacción auditiva eficaz, por lo que no es recomendable para personas con un nivel cognitivo bajo.

3.1.2. Producción de texto

Entre los dispositivos y productos de apoyo que utilizan las personas con parálisis cerebral para la producción de texto, destacamos:

- Teclados y accesorios para teclado. Incluye teclados de diferentes tamaños y elementos como cobertores. Un cobertor es una placa de metacrilato que se instala sobre el teclado y a la cual se le han practicado agujeros que coinciden con la posición de las teclas. El cobertor evita que se realicen pulsaciones involuntarias y permite que las manos descansen sobre el teclado.
- Teclados en pantalla. Consiste en un software que muestra en la pantalla del ordenador la representación de un teclado, generalmente con la disposición convencional. El usuario selecciona las teclas utilizando el dispositivo apuntador, mediante acceso directo (si se trata de una pantalla táctil) o mediante métodos de barrido. Para acelerar la producción de texto y reducir el número de errores, los teclados en pantalla suelen incorporar sistemas de predicción de palabras.



Figura 3.2: Teclado en pantalla con predicción de palabras.

3.1.3. Control del dispositivo apuntador

El ratón convencional, a pesar de ser el dispositivo más extendido entre la población general, resulta muy difícil de usar para muchas personas con parálisis cerebral. Por ello, se enumeran a continuación algunas de las alternativas más extendidas al ratón que emulan u ofrecen una funcionalidad similar a la de éste:

- *Trackball*. Es un dispositivo apuntador que consiste en una bola que está sujeta dentro de un receptáculo y que puede rotar libremente. Unos sensores detectan la rotación relativa de los ejes X e Y, lo que permite

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

64

mover el puntero por la pantalla. Junto a la bola, suele haber dos o más botones cuya funcionalidad suele coincidir con la de sus homólogos en el caso del ratón. El *trackball* resulta de utilidad en el caso de personas que conservan cierta precisión en el movimiento de sus manos, aunque su rango sea reducido, y facilita el accionamiento de los botones una vez se ha detenido en puntero en la posición deseada puesto que, a diferencia del ratón, el *trackball* no se desplaza sobre una superficie.

- *Touchpad*. Consiste en un pequeño panel táctil que permite mover el dispositivo apuntador con ligeros movimientos de un dedo. Incorpora dos o tres botones y se incluye de serie en muchos ordenadores portátiles. Resulta de utilidad a personas con movimiento en sus dedos de poca amplitud aunque preciso.
- Pantalla táctil. Parecido al anterior, aunque en este caso se realiza la manipulación directa de los elementos en pantalla prescindiendo de puntero visible y suele admitir la posibilidad de usar gestos con uno o más dedos. Requiere cierta precisión y amplitud de movimientos para su manejo. Se puede combinar con un cobertor de metacrilato con idénticas ventajas que en caso del teclado (véase la Figura 3.3).



Figura 3.3: Pantalla táctil con cobertor.

- *Joystick* isométrico. Consiste en una palanca montada sobre una base donde un sistema de muelles o similar se encarga de mantenerla en posición vertical. Generalmente, el ángulo de desplazamiento de la palanca respecto a la posición de reposo se hace corresponder a la velocidad de desplazamiento del puntero de pantalla. Una variante de este tipo de *joystick* es el *point stick* y consiste en una pequeña palanca recubierta por un capuchón antideslizante y funciona detectando la fuerza aplicada,

típicamente midiendo la resistencia eléctrica de un material. También es conocido por su nombre comercial TrackPoint de IBM. Otra variante es el *joystick* de contactos (digital). Consta de una palanca en cuya base se han montado cuatro microinterruptores. Permite hasta 8 direcciones distintas. Una vez fijada la dirección el puntero incrementa la velocidad hasta alcanzar un máximo prefijado.

- Emulación teclado. La mayoría de sistemas operativos permiten desplazar el puntero en pantalla usando el teclado numérico, siendo su funcionamiento equivalente al del caso del *joystick* de contactos. Una variante de éste consiste en un banco de conmutadores montados sobre una plataforma. Cada conmutador tiene una función específica, usualmente cuatro para la dirección y cuatro más para acciones concretas.
- Control con la cabeza. Bien indirectamente a través de un cabezal li-cornio y emulación mediante teclado, o bien directamente a través del seguimiento de la posición de la cabeza (en el marco de esta tesis se describe un sistema (véase la Sección 3.2) que encaja en esta categoría). Indicado para aquellas personas con buen control cefálico.
- Control con la mirada. Permiten detectar donde está mirando el usuario haciendo que el puntero se desplace a dicha ubicación. En general se suele reservar para los casos más extremos.

3.1.4. Rendimiento de la interacción

La idea de extraer valores cuantitativos a partir de la interacción entre una persona y una máquina no es nueva. De principios del siglo veinte data un estudio sobre análisis de errores a la hora de escribir a máquina usando un teclado QWERTY [74]. Cuando se creó el primer ratón durante los años 60 también se realizaron diferentes pruebas donde se analizaban parámetros como la tasa de errores o el tiempo que el usuario tardaba en seleccionar un texto en pantalla [42].

Más recientemente, y ya en el ámbito de la accesibilidad, existen algunos trabajos orientados a facilitar la elección y ajuste de las tecnologías de apoyo basándose en el análisis de valores cuantitativos que ayuden en la decisión (para un resumen véase [67]).

Algunos de estos trabajos se basan en cuestionarios y otros proponen la realización de tareas interactivas durante las cuales se registra la actividad del usuario para posteriormente ser analizada. Entre estos últimos destacamos Single Switch Performance Test [76], el proyecto Compass [66] y el sistema ERGOLAB [93, 92].

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

66

Single Switch Performance Test (SSPT) es una aplicación que permite medir la capacidad de activar un solo conmutador. SSPT permite medir el tiempo medio requerido para activar o liberar un conmutador y también para realizar varias activaciones sucesivas. Está disponible gratuitamente en Internet.

El proyecto Compass es una aplicación que mide la habilidad del usuario en varios tipos de interacción con el ordenador. Propone pruebas para realizar con conmutador (activación simple y selección por barrido), introduciendo texto (letra, palabra y frase) y con un dispositivo apuntador (selección, arrastre y selección en menú). Compass recopila velocidad y precisión durante las pruebas y genera un informe con los resultados. Compass está disponible como *shareware*.

ERGOLAB es una aplicación que también mide varias variables de rendimiento de la interacción. Consiste en tres etapas de dificultad creciente con diversos ejercicios en cada una de ellas. La primera etapa consta de ejercicios para realizar con conmutador (activación simple, mantener la pulsación y selección por barrido), con un dispositivo apuntador (selección y arrastre) y ejercicios basados en el manejo de información oculta (paso de páginas de un libro y uso de barras de desplazamiento entre otros). Los resultados que arroja la aplicación después de cada ejercicio consisten en medidas de tiempo (duración del entrenamiento, duración global del ejercicio, tiempo de reacción medio para obtener el resultado correcto y tiempo de apoyo medio sobre el conmutador) y medidas de error (errores de tiempo y de localización). ERGOLAB ha sido concebido especialmente para usuarios con parálisis cerebral por lo que parte de nuestro trabajo se basará en este sistema.

La Norma ISO 9241 [57] es una norma que abarca distintos aspectos sobre la ergonomía en la interacción persona-ordenador. El estándar consiste en 17 partes. La parte 4, cuyo título es “Requisitos del teclado”, define las características físicas del teclado y las métricas de rendimiento para la entrada de texto. La parte 9, titulada “Requisitos para dispositivos de entrada diferentes al teclado” también define procedimientos y métricas para evaluar la eficiencia y eficacia de los dispositivos apuntadores. Ninguno de los trabajos antes mencionados utiliza estas normas ISO.

3.1.5. Método de evaluación propuesto

El método de evaluación propuesto intenta simplificar los métodos existentes reduciendo la complejidad al mínimo. Dadas las limitadas capacidades de interacción de las personas con parálisis cerebral, las métricas de evaluación deberían tener en cuenta los siguientes factores:

- Capacidad cognitiva. Según el lugar y grado de la afectación neurológica,

pueden asociarse otras alteraciones como son la cognición, la atención, la memoria, etc. lo que hay que tener en cuenta para los métodos de evaluación empleados.

- Capacidades sensoriales y motrices. También existen limitaciones en este sentido que pueden invalidar ciertos métodos de evaluación y que acotan el rango de dispositivos y técnicas de interacción que pueden usarse.
- Experiencia previa. Es posible que el usuario tenga adquiridos algunos patrones de interacción basados en la experiencia previa y que estos no sean los más óptimos para el individuo. Por ejemplo, algunas personas con parálisis cerebral manejan una silla de ruedas mediante un *joystick* o un conmutador; no siendo estos dos dispositivos los más eficaces para el control de un sistema de ventanas. El método tendría que tener en cuenta esta experiencia previa, compensarla de alguna forma y no descartar de entrada otras alternativas que, con el tiempo, podrían llegar a ser mejores.
- Evaluación del progreso. La métrica debería ser útil también para poder evaluar el progreso del usuario que puede adquirir a base de práctica. Dado que, en general, los avances suelen ser muy lentos, el método debería poder medir con cierta precisión.

Nos centramos en simples tareas interactivas, en el sentido de que se trata de partes de otras más complejas. El hecho de focalizar nuestra atención en estas tareas interactivas de bajo nivel está justificado en nuestro contexto.

En primer lugar, las tareas simples son más apropiadas para los usuarios con parálisis cerebral, ya que son fáciles de aprender y reproducir incluso cuando existen limitaciones cognitivas severas. En segundo lugar, los indicadores son simples y claros y las tareas son fácilmente reproducibles, por lo que son útiles para comparar los dispositivos interactivos. Por último, estas medidas establecen la base para el desarrollo de otras más complejas.

A continuación se describen las cinco tareas interactivas propuestas que conforman la prueba. Los datos de rendimiento se recogerán durante las sesiones interactivas con usuarios reales. Esta información se almacena para su posterior análisis.

Cada tarea puede ser configurada para ejecutarse en modo de demostración o en producción. El modo de demostración se utiliza para indicar al usuario cómo realizar la tarea y no hay recogida de datos. En el modo de producción, el usuario debe realizar la tarea por sí mismo sin asistencia externa y se recogen datos de rendimiento que se almacenan en la base de datos.

68 **Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos**

Las únicas limitaciones que deberían impedir que el usuario pueda ejecutar una tarea son su capacidad cognitiva o etapa educativa en la que se encuentra (por ejemplo, si todavía no ha aprendido a leer y escribir). Las limitaciones físicas deberían poderse superar con la tecnología de apoyo adecuada.

3.1.5.1. Tarea 1. Respuesta simple a un estímulo

Al usuario se le presentará un estímulo en forma de imagen y/o sonido. El estímulo podrá ser concreto (fotografía de una persona cercana, dibujo de juguete, voz familiar, etc.) o abstracto (figura geométrica, cambio de color o tonalidad, sonido sintético, etc.). La elección entre una modalidad de entrada (imagen o sonido) y el tipo de estímulo (concreto o abstracto) dependerá de factores como la motivación, las capacidades visuales y auditivas y el nivel cognitivo del usuario. Para aquellos usuarios con una afectación más profunda puede ser interesante prescindir del uso del color pues en ocasiones puede distraer [43]. Los dispositivos necesarios para este ejercicio son los conmutadores.

Un vez presentado el estímulo el usuario debe dar una respuesta lo antes posible (i.e. presionar el conmutador). Opcionalmente, tras la respuesta, el sistema puede presentar un refuerzo en forma de estímulo visual y/o auditivo para indicar al usuario que lo ha hecho correctamente. Este proceso se repite un número finito de veces. El evaluador debe configurar el número de iteraciones apropiadas para el usuario. El tiempo transcurrido entre la respuesta del usuario y la presentación de nuevos estímulos se elige aleatoriamente a partir de un intervalo de valores posibles.

Las medidas que se recogen y almacenan son:

- Tiempo de respuesta. Es el tiempo que transcurre entre la presentación del estímulo y la respuesta del usuario.
- Número de iteraciones.
- El tiempo total para completar la tarea.
- Número de errores. Se entiende por error el hecho de que el usuario dé una respuesta sin que previamente se haya presentado el estímulo.

3.1.5.2. Tarea 2. Mantener una respuesta discreta

Este ejercicio es similar al anterior, pero con la diferencia que ahora hay que mantener la respuesta (presionando el conmutador) mientras no desaparece el estímulo. El tiempo que dura el estímulo varía de forma aleatoria dentro de un intervalo configurable. Opcionalmente, el propio estímulo puede indicar el

tiempo que queda de alguna manera, como una barra de progreso o escala de sonido.

Esta tarea puede ser útil para tareas más complejas como por ejemplo la interacción mediante código Morse (el usuario ha de ser capaz de realizar pulsaciones cortas y largas), la función de arrastrar y soltar un elemento de la interfaz gráfica o el uso de algunos métodos de barrido.

Las medidas que se recogen y almacenan son:

- Tiempo de respuesta inicial. Es el tiempo de respuesta de la tarea 1. De hecho, los resultados deberían ser los mismos que en la tarea 1.
- Tiempo de respuesta final. Es el tiempo que transcurre entre la desaparición del estímulo y la desaparición de la respuesta del usuario (cuando se deja de presionar el conmutador). Normalmente, este valor debería ser muy similar a la anterior.
- Tiempo de pulsación. Es el tiempo que ha durado cada respuesta no iniciada erróneamente.
- Número de errores iniciales. Corresponde al número de errores de la tarea 1.
- Número de errores final. Número de veces que el usuario ha interrumpido la respuesta de manera temprana.
- El tiempo total para completar la tarea.

Estas medidas son similares o se pueden derivar de las encontradas en ERGOLAB [93] y Compass [66].

3.1.5.3. Tarea 3. Selección de un elemento mediante barrido

Aunque, en esencia, es una respuesta simple a un estímulo con una mayor carga cognitiva, creemos que esta tarea, debido a su importancia, debe ser incluida en el método de evaluación propuesto. Antes de iniciar el ejercicio, es necesario establecer una velocidad de barrido determinada, de manera que los resultados no se vean alterados por las variaciones que un sistema de regulación automática pueda introducir y que, por lo tanto, sea posible hacer comparaciones entre diferentes dispositivos.

Esta tarea le presentará al usuario una serie de opciones a lo largo de una dimensión, entre las cuales deberá elegir una (la opción correcta se indicará por medio de un estímulo visual y/o acústico), estas opciones se deberán alternar consecutivamente, creando un estímulo visual y/o acústico cada vez. En el momento apropiado, el usuario enviará una respuesta. La elección correcta o

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

70

errónea puede estar reforzada con un estímulo visual y/o acústico. Tanto si la persona realiza la tarea con éxito o como si no, el sistema esperará un tiempo aleatorio, que se puede configurar, antes de presentarle un nuevo test.

Además de los valores registrados en la tarea 1, tendremos que añadir la velocidad de exploración utilizada y los errores cometidos por respuestas tardías.

3.1.5.4. Tarea 4. Posicionar el puntero y seleccionar

Las mediciones de evaluación más comunes para dispositivos de puntero son la velocidad y exactitud; hasta el punto que existe una norma ISO para la evaluación de este tipo de dispositivos. Concretamente dentro de la norma ISO 9241 “Requisitos ergonómicos para trabajos de oficina con pantallas de visualización de datos (PVD).”, en la parte 9 “Requisitos de los dispositivos de entrada que sean teclados” [57] propone una sola métrica: el rendimiento. Las unidades de rendimiento son bytes por segundo y se obtienen basándose en la velocidad y la precisión.

$$Performance = \frac{ID_e}{MT} \quad (3.1)$$

Donde el término ID_e indica la dificultad de la tarea (en bytes) y MT el tiempo necesario para llevar a cabo el movimiento. La dificultad se calcula como:

$$ID_e = \log_2 \left(\frac{D}{W_e} + 1 \right) \quad (3.2)$$

Donde el término D se refiere a la distancia al objetivo y W_e al tamaño efectivo de éste, a diferencia de la ley de Fitt [77], se tiene en cuenta el tamaño efectivo en lugar del tamaño real del objeto W . El tamaño efectivo se obtiene a partir de la distribución de las coordenadas elegidas durante varias pruebas, y por lo tanto refleja la precisión con la que se ha llevado a cabo la prueba. Además, el tamaño efectivo puede ayudar a determinar el tamaño que los elementos de la interfaz gráfica deberían tener. La fórmula para obtener el tamaño efectivo es la siguiente:

$$W_e = 4,133 \times SD_x \quad (3.3)$$

Donde SD_x es la desviación estándar de las coordenadas elegidas durante el ejercicio. Para más información acerca de estas métricas, véase [34].

Además de las métricas estándar, algunos estudios [78, 63] proponen una serie de métricas adicionales que nos permitan caracterizar la tarea de señalar con el puntero. Entre éstas, vamos a usar la tasa de re-entrada al objetivo

(*Target Re-entry [TRE]*) y la desviación respecto a la trayectoria ideal (*Movement Offset [MO]*). La trayectoria ideal se considera que es una línea recta entre el origen y el destino. La desviación se calcula como la media de la distancia entre las diferentes muestras de la trayectoria real y la trayectoria ideal.

Vamos a definir un ejercicio basado en la norma, pero con algunas modificaciones para adaptarlo al tipo de usuarios con los que estamos tratando. Éste consistirá en una ventana, el centro de la cual será el punto de partida para todas las pruebas de ejercicio. Este punto puede estar opcionalmente indicado con un estímulo visual. Los estímulos de destino aparecerá alrededor de este punto en forma de imágenes que pueden estar acompañadas de sonido. La distancia y el tamaño de los estímulos será configurable; y el ángulo se hará variar de forma aleatoria. Opcionalmente, se pueden usar otros estímulos visuales y acústicos para reforzar la ejecución correcta o incorrecta.

Una vez que la tarea se ha iniciado, el puntero se coloca en el punto central de esta ventana y se mantiene en esta posición hasta el instante en que aparece el estímulo objetivo. Esto se hace para facilitar la tarea en el caso del uso de dispositivos en los que sea más difícil mantener la posición fija (por ejemplo, un sistema de seguimiento de la cabeza o los ojos). El usuario debe mover el puntero y hacer clic en el objetivo, tan al centro como sea posible, para ello el estímulo puede estar diseñado de manera que anima a señalar en el centro (por ejemplo forma de diana o espiral). La prueba se repite un número finito de veces a intervalos aleatorios dentro de un margen configurable.

Las medidas que se recogen y almacenan son:

- Tiempo de reacción. Es el tiempo que transcurre entre que se presenta el estímulo objetivo y el usuario comienza a mover el puntero. Hay que tener en cuenta que este valor puede no tener validez dependiendo del tipo de dispositivo utilizado.
- Tiempo de ejecución. Es el tiempo que transcurre entre que se presenta el estímulo objetivo al usuario y éste es seleccionado. El valor del tiempo de ejecución es el que se utilizará para el cálculo del rendimiento.
- Rendimiento.
- Número de re-entradas.
- Tasa de re-entrada al objetivo.
- Desviación respecto a la trayectoria ideal.
- Tamaño efectivo del objetivo (calculado para todo el ejercicio).

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

72

- Número de errores. Se considera error hacer clic fuera del objetivo.
- Tiempo total para completar el ejercicio.

Puede resultar interesante registrar todos los eventos enviados por el dispositivo apuntador para más tarde poder visualizar en pantalla las trayectorias realizadas.

3.1.5.5. Tarea 5. Producción de caracteres y texto

Para el análisis de la producción de caracteres y texto se emplearán las métricas definidas en el estándar ISO 9241-11 [58], aunque con los matices pertinentes para ser usada con personas que pueden presentar graves alteraciones cognitivas o aquellas que todavía están en proceso de aprendizaje de la lecto-escritura. Para ello el usuario deberá transcribir textos prefijados cuya extensión podrá estar comprendida entre un solo carácter o varias líneas de texto. Se permitirá por configuración dos modos de funcionamiento, uno sin edición (los caracteres introducidos incorrectamente se descartan) y con edición. El dispositivo empleado para producir texto no tiene por que ser forzosamente el teclado convencional.

El sistema presentará una ventana (de fondo preferentemente oscuro con caracteres en blanco con un tamaño suficiente). Dentro de la ventana existirán dos áreas, una donde se mostrará el texto que hay que transcribir y la otra donde el usuario deberá copiar el texto. El texto de muestra se presentará íntegro. Opcionalmente, se podrá reforzar mediante un estímulo visual el siguiente carácter que debe introducirse. La introducción de cada carácter o de toda la frase, correcto o incorrecto, podrá ir acompañada de un refuerzo visual y/o acústico. El ejercicio constará de varios test donde aparecerán en orden aleatorio las palabras de una lista.

Las medidas que se recogen y almacenan son:

- Carácter por segundo (CPS).

$$CPS = \frac{\text{Caracteres introducidos}}{\text{Tiempo empleado}} \quad (3.4)$$

- Tasa de errores de carácter (*Character Error Rate* [CER]).

$$CER = \frac{\text{errores}}{n} \times 100 \quad (3.5)$$

En caso de realizar el ejercicio con modo edición $\text{errores} = s + d + i$ donde $s = \text{número de sustituciones}$, $d = \text{número de borrados}$, $i = \text{número de inserciones}$ y $n = \text{número de caracteres}$. En modo sin edición los errores se consideran los caracteres introducidos incorrectamente.

- Longitud del texto de muestra.
- Tiempo empleado por texto.
- Tiempo total para completar el ejercicio.

3.2. Ratón Facial y Enable Viacam

Publicado a finales de 2003 como herramienta comercial, Ratón Facial [88] es un sistema que permite emular las funciones de un ratón a partir de los movimientos de la cabeza del usuario captados con una webcam. El producto fue pionero dentro del mercado español y uno de los primeros en aparecer en todo el mundo. A finales de 2008, se libera una versión de código abierto del mismo sistema que se publica bajo el nombre de Enable Viacam (eViacam) [85].

3.2.1. Funcionamiento

Se captura la imagen del usuario con una cámara (generalmente una cámara web), se localiza la zona donde se encuentra el rostro y se analiza el movimiento dentro de esta área. La cantidad de movimiento detectada se hace pasar a través de un filtro pasa bajos para eliminar el ruido de alta frecuencia, se multiplica por un factor que el usuario puede ajustar y si finalmente el resultado supera cierto umbral se generan las órdenes apropiadas para mover el puntero del ratón por la pantalla.

La aplicación también permite emular los clics de ratón deteniendo el puntero sobre una posición determinada durante un instante (método al que nos referimos como “clic por espera”). Si es necesario, el clic por espera se puede desactivar y realizar el clic por medios externos, por ejemplo, mediante un conmutador u otro software. Cuando el clic por espera está activado, se muestra una pequeña ventana flotante que permite seleccionar entre diferentes acciones clic: botón izquierdo, botón derecho, botón central, arrastrar y soltar, y doble clic.

Para localizar la presencia y la posición del rostro, se utiliza un clasificador en cascada basado en características Haar que fue propuesto inicialmente por Paul Viola [126] y mejorado por Rainer Lienhart [75]. Este clasificador está implementado en la biblioteca OpenCV [97], que también ofrece varias cascadas entrenadas para detectar rostros.

Una vez localizado el rostro, se realiza una estimación del movimiento en el interior del área alrededor de éste mediante el algoritmo de flujo óptico de Horn y Schunck [51]. Como resultado se obtienen dos campos de velocidad

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

74

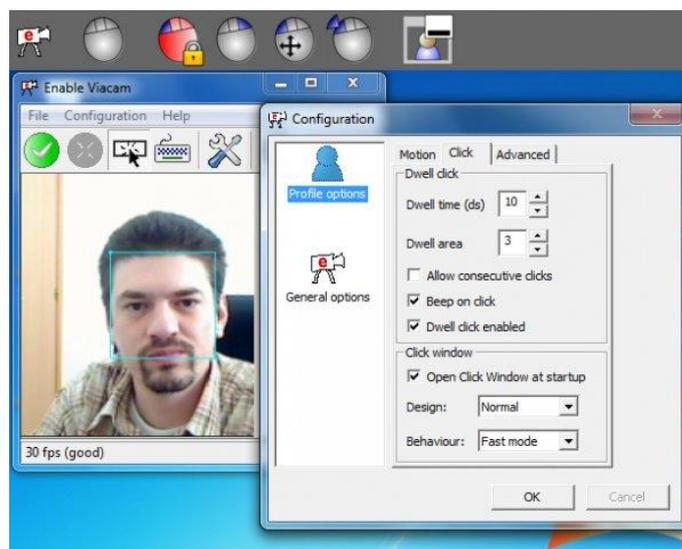


Figura 3.4: Captura de pantalla de Enable Viacam.

(uno por cada eje) de la misma resolución que las imágenes de entrada. Ambos campos son submuestreados a una matriz de 15×15 y se calcula el módulo para cada celda de dicha matriz. Para descartar zonas estáticas (por ejemplo, partes del fondo) sólo las celdas con suficiente cantidad de movimiento se promedian para calcular el vector de movimiento resultante (el que acabará haciendo desplazar el puntero del ratón). En nuestro caso, hemos ajustado empíricamente el sistema para que seleccione sólo las celdas cuyo módulo sea mayor que $0,05 * MAX_MODULUS$, donde $MAX_MODULUS$ es el módulo máximo de todas las celdas.

3.2.2. Estado actual del proyecto

Desde su lanzamiento se han publicado 20 revisiones de Enable Viacam y sólo desde el sitio de SourceForge ha sido descargado más de 325.000 veces. Recibe regularmente peticiones de nuevas características y reportes de errores por parte de sus usuarios. Está disponible en 13 idiomas (inglés, español, catalán, gallego, francés, italiano, alemán, turco, portugués, ruso, árabe, hebreo y japonés), siendo la mayoría de traductores los propios usuarios de la herramienta.

A nivel español nos consta que se está usando en entidades de referencia como el Hospital Nacional de Paraplégicos (Toledo) o el Institut Guttmann (Badalona). A nivel mundial tenemos constancia de uso por parte de numerosos usuarios particulares.

Ha recibido el apoyo económico de entidades como la Generalitat de Catalunya y la Junta de Andalucía incluyéndolo, esta última, en la distribución Guadalinux para su red de Telecentros. A fecha de hoy, la inversión total realizada en el proyecto supera los 100.000 euros.

3.3. SITPLUS

3.3.1. Objetivos y criterios de diseño

El objetivo de SITPLUS es crear una zona sensible donde el usuario, a pesar de sus limitaciones físicas, cognitivas y/o sensoriales, sea capaz de interactuar. La intención es promover la interacción a través del juego y la diversión con el fin de mejorar la relación causa-efecto y la sensación de control. Este punto es importante teniendo en cuenta el hecho de que, por lo general, muchas de estas personas tienen una interacción muy limitada con su entorno. Por lo tanto, permite a los usuarios actuar sobre su entorno cercano lo que puede promover efectos positivos en su calidad de vida.

Para hacer frente a esta tarea, es necesario tener en cuenta las deficiencias y las habilidades conservadas de los participantes para proporcionar una solución accesible.

- **Accesibilidad física.** Nuestros participantes tienen una amplia variedad de limitaciones físicas. Por lo tanto, los dispositivos convencionales de entrada (teclado y ratón) no se pueden utilizar en la mayoría de los casos. En nuestro caso, ofrecemos diferentes canales de entrada basados en gestos físicos y en las emisiones orales (no necesariamente habla) de los usuarios. De esta manera podemos aprovechar las habilidades físicas que nuestros participantes todavía conservan. Los gestos físicos se recogen a través de una cámara web empleando técnicas de visión artificial o mediante el uso de otros sensores como acelerómetros o giroscopios, mientras que para las emisiones orales se utiliza un micrófono.
- **Accesibilidad sensorial.** Algunos participantes también tienen problemas de audición y/o de visión. En nuestro caso asumimos que el participante es capaz de, al menos, ver o escuchar aunque presente algunas dificultades. Por lo tanto, proporcionamos estímulos altamente contrastados y multimodales (en nuestro caso, imagen y sonido).
- **Accesibilidad cognitiva.** El sistema debe ser accesible a pesar de las deficiencias cognitivas de los participantes o, al menos, a aquellos que están en el proceso de aprendizaje de la relación causa-efecto. Para ello, ofrecemos una relación directa, casi literal, entre las entradas y las salidas.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

76

Por otra parte, también hemos tenido en cuenta el hecho de que muchos centros que atienden a personas con parálisis cerebral no pueden invertir grandes cantidades en equipos y, por ello, nuestra solución está completamente basada en software libre que se ejecuta en ordenadores convencionales equipados con dispositivos de uso común como como cámaras web, altavoces y micrófonos.

Este enfoque reduce drásticamente los costes, en contraste con otros sistemas basados en elementos hardware específicos tales como los expuestos en el Capítulo 2, lo que además requiere de la intervención de profesionales cualificados. Esto es crucial para la mayoría de los centros de atención que, por lo general, disponen de presupuestos ajustado y prefieren invertir en equipamiento versátil como, por ejemplo, ordenadores personales.

3.3.2. Arquitectura del sistema

La arquitectura del proyecto SITPLUS ha evolucionado a lo largo del tiempo, desde los primeros prototipos construidos sobre versiones modificadas de eViacam y aplicaciones externas para el procesamiento de sonido y generación de imagen, a una aplicación integrada que proporciona soporte para dispositivos de entrada adicionales (por ejemplo, el mando de Wii) y simplifica su uso e instalación.



Figura 3.5: Captura de una primera versión adaptada de Enable Viacam.

Como ya se ha comentado, los primeros prototipos utilizaban una versión modificada de eViacam (véase la Figura 3.5). Dicha versión en lugar de mover

3.3. SITPLUS

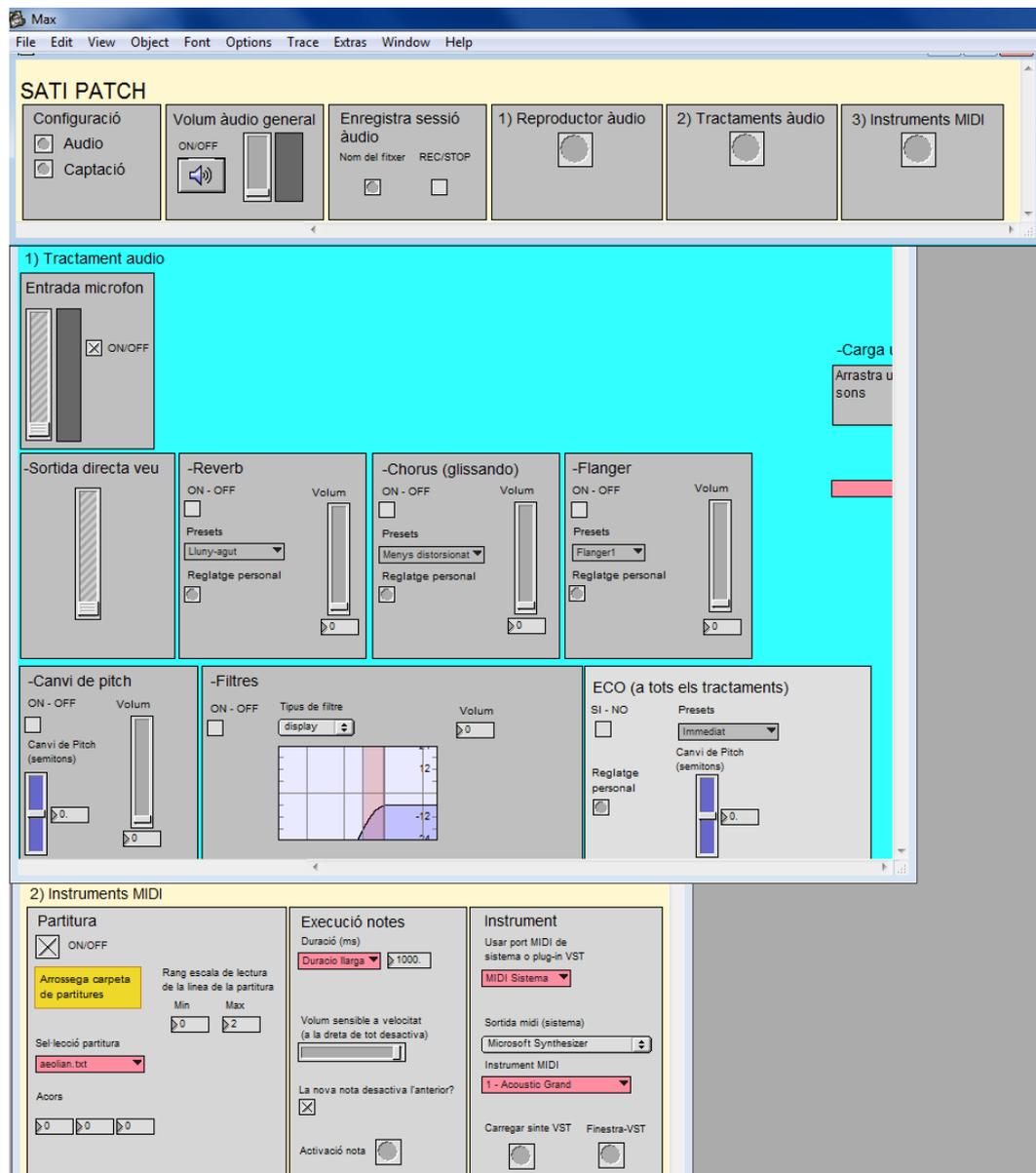


Figura 3.6: Captura de un patch Max/MSP para la generación de sonido.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

78

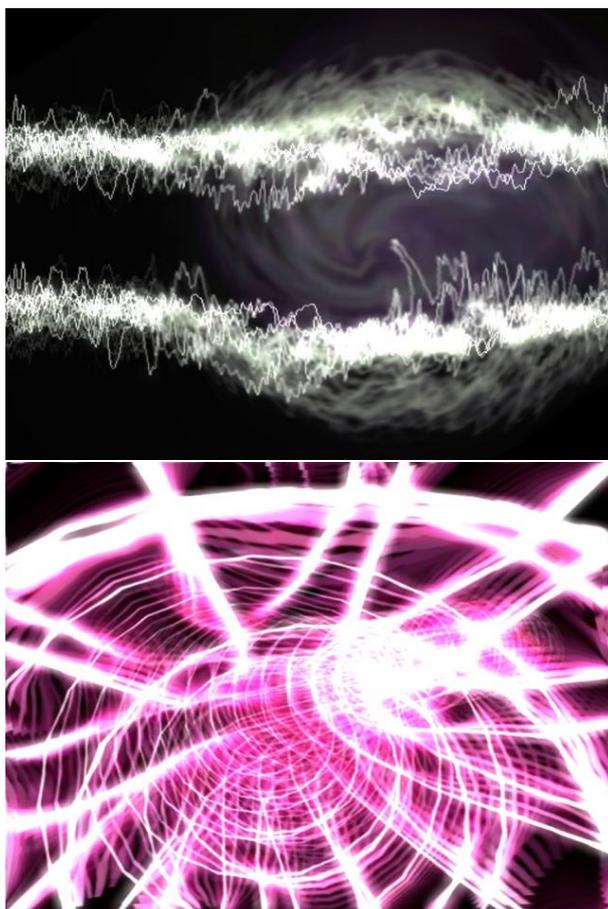


Figura 3.7: Captura de pantalla R4.

el puntero el ratón, enviaba a una aplicación externa la cantidad de movimiento detectado. La generación de sonido se realizaba a través de una colección de aplicaciones externas (también llamados *patches*, véase la Figura 3.6) construida utilizando los lenguajes de programación PureData y Max/MSP.

Por último, otra aplicación externa llamada R4¹ (véase la Figura 3.7) se utilizó inicialmente para proporcionar una salida gráfica. Esta aplicación genera una serie de animaciones visuales de acuerdo con el patrón de sonido que está siendo reproducido. Este sistema se puede personalizar eligiendo entre aproximadamente 100 escenas diferentes. Sin embargo, después de algunos ensayos con diferentes escenas, decidimos dejar de usarlo porque la relación entre el sonido y la animación no era evidente, incluso para músicos profesionales.

El escenario anterior demostró ser flexible pero no demasiado usable. Es

¹<http://www.rabidhamster.org/R4/>

decir, las aplicaciones debían ponerse en marcha en un orden determinado, los controles estaban dispersos en diferentes ventanas, y la instalación de todo el conjunto requería habilidades avanzadas. Como estábamos preocupados por su aplicabilidad en entornos reales, se decidió diseñar un nuevo sistema para hacer frente a estas cuestiones teniendo en cuenta los siguientes requisitos:

- El sistema debe dar soporte a dos actores principales, a saber, los facilitadores y los usuarios finales. Las sesiones de trabajo suelen ser individuales y en ellas está presente la figura de un profesional que llamamos facilitador (para más información véase la Sección 2.6.2).

En contraste con las actividades causa-efecto convencionales en las que, una vez iniciadas, se transfiere el control al usuario, nuestro sistema debe permitir el facilitador ajustar algunos parámetros a medida que el usuario avanza. En nuestro caso, la progresión no se mide por el hecho de alcanzar una meta concreta (por ejemplo, el usuario pulsa un conmutador), sino que el objetivo principal es interaccionar, disfrutar y ser conscientes de esa interacción. Por lo tanto, la progresión es una medida subjetiva y a menudo evaluada por el facilitador a partir de la observación o, en el mejor de los casos, de la comunicación que pueda establecer con el usuario. A veces, la progresión es muy lenta o incluso no se observa progresión en absoluto. En otras ocasiones el usuario hace una progresión rápida, incluso llegando a solicitar cambios en los ajustes continuamente.

- Aparte de ajustarse a las capacidades de los usuarios finales, el sistema debe ser usable para un profesional de un centro de atención sin que sea un experto.
- El sistema debe ser capaz de manejar múltiples canales de entrada simultáneamente, algunos de los cuales requieren un procesamiento en forma de *pipeline* (por ejemplo, procesamiento de imágenes). Además es deseable poder sacar partido al paralelismo que ofrecen los múltiples núcleos de los procesadores actuales.
- El sistema debe ser capaz de proporcionar salidas en distintas modalidades como visual y auditiva.

Finalmente, se optó por desarrollar una aplicación desde cero. Esta aplicación ofrece varias actividades (actualmente dos, para más información véase la Sección 3.3.7). Una vez iniciada, cada actividad proporciona una ventana con el panel de control (véase, por ejemplo, la Figura 3.8) para el facilitador (de forma que pueda ajustar los parámetros de la actividad a medida que

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

80

progresar la sesión) y otra ventana con la salida gráfica para el usuario final. Cada ventana puede colocarse en un monitor diferente para evitar distraer al usuario cuando sea necesario realizar algún ajuste. También es posible utilizar un solo monitor y cambiar entre estas dos ventanas cuando sea necesario.

Desde el punto de vista del usuario final, el sistema es compatible con los siguientes canales de entrada: seguimiento de movimiento basado en visión artificial, análisis de sonido y controladores de la consola Nintendo Wii. Las salidas del sistema pueden ser en forma de gráficos y/o sonido. Se baraja la posibilidad de incluir otras modalidades de interacción basadas en conmutadores y dispositivos vibratorios.

Por último, la aplicación proporciona algunos diálogos para configurar la cámara, los controles de Wii y la salida de sonido.

A nivel interno, en lugar de desarrollar una aplicación monolítica, se optó por un diseño modular con un enfoque orientado al flujo de datos que, de acuerdo con nuestra experiencia con los sistemas PureData y Max/MSP, ha demostrado ser muy flexible y muy apropiado para el procesamiento en *pipeline* en sistemas con múltiples procesadores.

Con el objetivo de hacer que los esfuerzos realizados puedan ser reutilizados en el futuro, el sistema se basa en dos elementos distintos: una biblioteca (*libsitplus*) y la propia aplicación SITPLUS.

3.3.2.1. *libsitplus*

SITPLUS está construido alrededor de un núcleo desarrollado en C++ (*spcore*), que proporciona servicios básicos como la carga de módulos, la instanciación de componentes, tipos básicos y la internacionalización, entre otros, y las clases base e interfaces para el desarrollo de nuevos componentes.

El bloque de construcción básico es el componente (interfaz *IComponent*). Un componente es una unidad de procesamiento que proporciona pines de entrada y/o de salida y, opcionalmente, un panel de interfaz gráfica de usuario. Los componentes se agrupan en un objeto compuesto (patrón *composite*) y sus pines se pueden conectar para construir aplicaciones orientadas a flujo de datos.

Cada componente se identifica por su nombre de tipo y se pueden crear múltiples instancias del mismo. Cada pin tiene un tipo asociado, que puede ser un tipo específico (por ejemplo, entero, carácter o número en coma flotante) o genérico. Sólo los pines cuyos tipos coinciden pueden conectarse entre sí.

Para hacer extensible esta aplicación marco, se pueden desarrollar nuevos componentes y tipos en forma de bibliotecas de carga dinámica (DLLs).

Junto con el núcleo, se proporcionan dos bibliotecas adicionales de soporte:

- *sphost*: proporciona algunas funciones de ayuda y clases para desarrollar

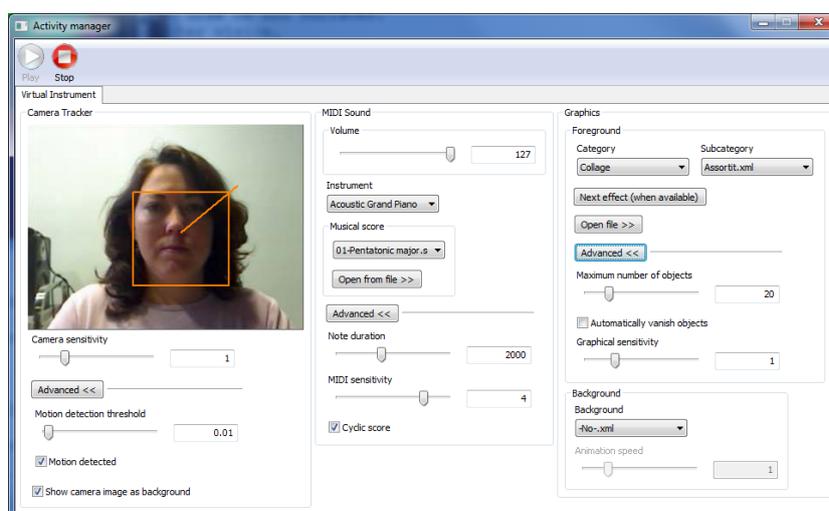


Figura 3.8: Captura de la actividad “Instrumento musical virtual”.

aplicaciones finales. Por ejemplo, un intérprete de secuencia de órdenes para facilitar la creación de nuevos componentes.

- `widgets_base`: proporciona clases base para el desarrollo de elementos de interfaces de usuario.

Finalmente, también se proporcionan otros módulos adicionales como por ejemplo:

- `mod_camera`: componentes y tipos para capturar imagen desde una cámara.
- `mod_io`: componentes para entrada/salida de archivo.
- `mod_midi`: componentes y tipos MIDI.
- `mod_sdl`: componentes y tipos para dibujar sobre superficies SDL.
- `mod_vision`: componentes con tratamientos de visión artificial.
- `mod_widgets`: componentes de interfaz de usuario como botones, cuadros combinados, etc
- `mod_wiimotes`: componentes y tipos para el mando de la consola Nintendo Wii

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

82

3.3.3. Módulos de visión artificial

El objetivo de estos módulos es detectar el movimiento de los usuarios usando técnicas de visión artificial. Actualmente, este módulo implementa dos tipos de detectores de movimiento:

- Detector de movimiento fino. Este detector permite a los facilitadores elegir la región de interés (cuadro naranja de la Figura 3.8) donde se extrae el movimiento usando flujo óptico usando la misma técnica que la descrita en la Sección 3.2. Este detector retorna un vector en coordenadas polares con la dirección principal del movimiento y el módulo del mismo. Es decir, extrae el desplazamiento relativo de los píxeles dentro de la zona escogida. Mediante la rotación de la flecha dentro del cuadro naranja, es posible cambiar la orientación del vector devuelto. Debido a su alta sensibilidad, que es configurable, este detector es muy apropiado para las personas que tienen un rango de movimientos reducido. Sin embargo, en las pruebas realizadas hemos visto que también resulta útil para las personas con mayor movilidad, ya que no requieren el uso de elementos adicionales, tales como ropa especial, pegatinas, etc.
- Detector de movimiento grueso. Se coloca sobre el cuerpo del participante (por lo general en una extremidad superior) un marcador de color como por ejemplo, un guante o un adhesivo. Mediante el uso de técnicas de segmentación de color [88], se localiza y devuelve la posición 2D de dicho marcador. La región de la imagen dentro de la cual se realiza la detección también se puede acotar. Este detector es adecuado para personas con un rango de movimiento amplio en sus extremidades. Sin embargo, apenas lo usamos en la práctica porque es muy sensible a los cambios de luz ambiental (es decir, requiere de una cuidadosa calibración) y la mayoría de nuestros usuarios suelen tener deficiencias graves que les impiden la realización de movimientos de gran amplitud.

3.3.4. Módulos de audio

SITPLUS procesa el audio de dos maneras diferentes. Por un lado incluye un secuenciador MIDI para reproducir partituras musicales, y por otro, un subsistema de captura, procesamiento y reproducción de audio digital en tiempo real.

3.3.4.1. Secuenciador MIDI

El secuenciador MIDI es un componente de la biblioteca libsitplus que permite reproducir las notas musicales de una partitura a través de un sinte-

tizador MIDI. El sintetizador MIDI puede ser un hardware externo conectado al ordenador o, más comúnmente, un sintetizador software que normalmente forma parte del sistema operativo.

A diferencia de los secuenciadores MIDI convencionales que reproducen una partitura a una velocidad fija (o tempo) y siempre avanzando en la misma dirección a lo largo de la partitura, nuestro secuenciador permite tocar notas de las dos maneras siguientes:

- **Modo absoluto.** A partir de un valor entre 0 y 1 se selecciona directamente una nota (o acorde) de la partitura. Para ello, la partitura se divide en tantas particiones como notas contiene. A cada partición se le asigna un índice entre 0 y 1. La nota seleccionada es aquella cuyo índice está más cerca del valor de entrada.
- **Modo relativo.** Permite recorrer las notas de la partitura al tiempo que las reproduce ya sea hacia adelante o hacia atrás. El valor de la entrada se acumula y cada vez que esta suma supera, en valor absoluto, el umbral de 1 se reproduce una nota. Opcionalmente, es posible ajustar una opción que permite reproducir la partitura de forma continua de manera que cuando se alcanza un extremo de la misma se continua reproduciendo por el otro extremo. Es decir, siempre que se detecte movimiento se van reproduciendo notas. El secuenciador permite personalizar otros parámetros como la partitura, el volumen, el instrumento y duración de las notas.

3.3.4.2. Procesamiento en tiempo real de audio

El módulo de procesamiento en tiempo real de audio es una aplicación externa desarrollada con PureData y maneja la entrada de audio (a través de un micrófono) y el procesamiento en tiempo real de efectos digitales. Actualmente, incluye los siguientes efectos de audio: reverberación, coro, distorsión, cambio de tono, eco y eco con cambio de tono.

Optamos por utilizar PureData porque es muy apropiado para el procesamiento de audio en tiempo real, si bien nuestro sistema oculta complementemente este hecho y ejecuta PureData en segundo plano.

3.3.5. Controles remotos Wii

Desde su introducción en 2006, la consola de juegos Wii se ha caracterizado por sus innovadores dispositivos inalámbricos de interfaz humana, entre los que destacan:

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

84



Figura 3.9: Controladores Wii Remote y Nunchuck.

- Wii Remote (también conocido como wiimote). Es el controlador principal de la consola. Su versión original incluye varios botones, un acelerómetro de 3 ejes y un sensor óptico para el dispositivo apuntador. Existen diferentes extensiones externas que se pueden conectar al wiimote para mejorar las capacidades del controlador, siendo la más común el llamado Nunchuck (incluido con la consola), que añade otro acelerómetro de 3 ejes. La Figura 3.9 se muestra el mando Wii Remote y el Nunchuck.

Las nuevas versiones del mando de Wii también incluye un giroscopio de 3 ejes (también disponible como una extensión acoplable llamada Motion Plus).

- Wii Balance Board. Se trata de un periférico en forma de báscula de baño y está diseñado para ser usado estando de pie encima de éste e inclinándose en una dirección determinada.

La Wii Balance Board contiene cuatro sensores de presión en las patas que se utilizan para medir el centro de equilibrio del usuario.

Ambos controladores utilizan tecnología Bluetooth para conectarse a la consola, lo que ha permitido el desarrollo de controladores de terceros para otros dispositivos distintos de la consola (por ejemplo, un ordenador). Nuestro sistema soporta los controladores antes mencionados.

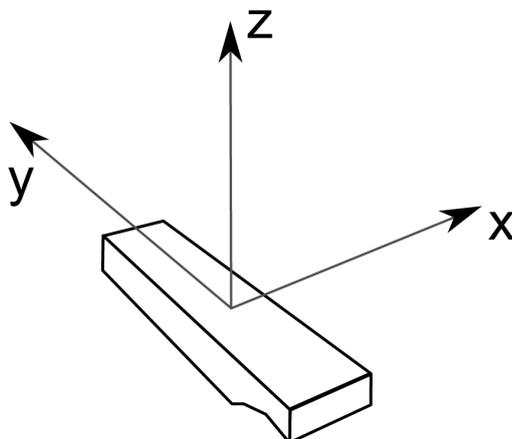


Figura 3.10: Ejes Wii Remote [3].

Para el Wii Remote, SITPLUS detecta si éste integra giróscopos (es decir, la extensión Motion Plus) y se la ofrece al usuario en caso afirmativo. La utilización de los giroscopios es preferible a los acelerómetros dado su mayor precisión para pequeños movimientos.

Cuando se utilizan los acelerómetros, el valor detectado para cada eje (medido en g) se acumula (sólo necesitamos una estimación aproximada del movimiento global) y el resultado se escala dentro del rango de trabajo. Después se calcula el valor absoluto y, por último, se eliminan (se reducen a 0) aquellos valores que estén por debajo de cierto umbral. Dicho umbral se determina automáticamente cuando el mando está en reposo, lo que permite eliminar el ruido del sensor en estas condiciones. La frecuencia de actualización se reduce de 90Hz a 30Hz calculando el promedio de cada 3 actualizaciones.

El facilitador puede seleccionar de cuales de los ejes se extraerá movimiento. Por defecto se utilizan los ejes X y Z (véase la Figura 3.10). El resultado final es una estimación de la cantidad de movimiento, aceleración es este caso, de forma que cuanto mayor es este, mayores son también los valores obtenidos.

Cuando se utilizan los giroscopios el proceso es muy similar al descrito para los acelerómetros, pero debido a su mayor precisión, no se calcula el valor absoluto y simplemente se acumulan las actualizaciones. El resultado es que, cuando el usuario gira el controlador alrededor de un eje, se obtiene un valor positivo o negativo, dependiendo del sentido del giro, cuyo valor absoluto es proporcional a la velocidad de rotación.

Para la Balance Board se calcula el centro de equilibrio a partir de los valores de los cuatro sensores y sus coordenadas se normalizan entre -1 y 1. 0 significa que la masa está completamente centrada a lo largo del eje, y 1

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

86

(o -1) significa que la masa está completamente desplazada al borde del eje. El carácter absoluto del valor resultante hace que funcione mejor con el modo absoluto del secuenciador MIDI. Al igual que en el caso del Wii Remote, la Balance Board ofrece lecturas a una velocidad de 90Hz que se reduce a 30Hz utilizando la técnica antes descrita.

3.3.6. Salida gráfica

La salida gráfica se proporciona a través de un módulo que hace evolucionar una animación 2D basada en imágenes predefinidas de acuerdo con un valor de entrada. Cuanto mayor es este valor, más rápida es la animación. La velocidad se puede ajustar a través de un multiplicador lo que permite al facilitador modificar interactivamente la sensibilidad durante la sesión. Este módulo muestra una ventana independiente y redimensionable lo que permite ubicarla en un monitor secundario (si está disponible).

Se admiten los dos tipos de animaciones siguientes:

- Animación cíclica. Permite definir una animación basada en un bucle de imágenes en movimiento. A cada imagen se le asigna una transición que la hace evolucionar cada vez que se recibe un nuevo valor de entrada. Cada imagen cuenta con los siguientes atributos: ruta al archivo (están soportados los formatos JPEG, TIFF y PNG), y rotación, escalado y ubicación iniciales. El escalado y la ubicación se expresan en valores normalizados relativos al tamaño de la ventana gráfica. De esta forma cuando se cambia el tamaño de ésta, todos los elementos en su interior se escalan en la misma proporción. Están disponibles las siguientes transiciones:
 - IDENTITY. No tiene ningún efecto. Útil para mostrar imágenes estáticas.
 - ALPHA. La imagen aparece y se desvanece modificando el valor del canal alfa.
 - SCALE. La imagen se escala hasta alcanzar su tamaño máximo (determinado por el atributo de escalado) y luego se reduce paulatinamente hasta desaparecer.
 - ROTATE. La imagen rota en un sentido y luego en el opuesto. Se puede seleccionar el arco total de giro y el sentido inicial de rotación.
 - CHANGE. La imagen se cambia por otra.
 - VIBRATE. La imagen vibra cambiando ligeramente su tamaño.

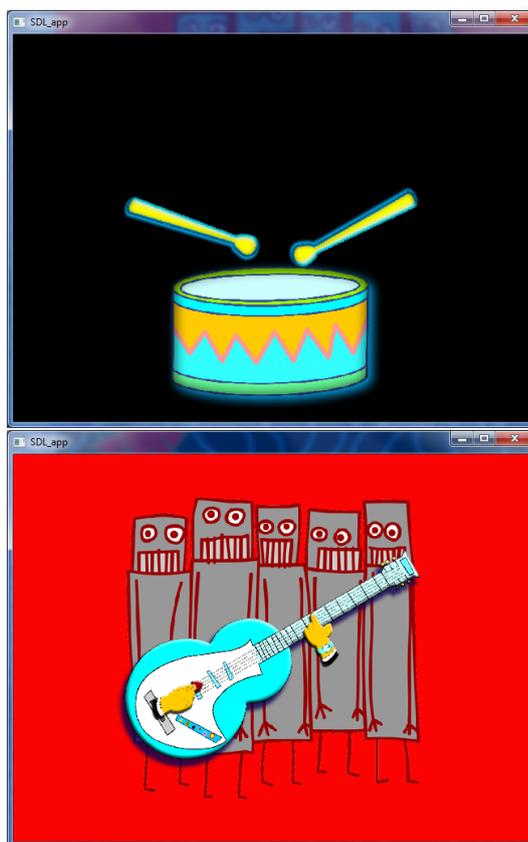


Figura 3.11: Capturas de animaciones cíclicas en SITPLUS.

- TRANSLATE. La imagen se desplaza de una posición inicial a otra final y viceversa.

La Figura 3.11 muestra dos animaciones proporcionadas por nuestro sistema que imitan instrumentos musicales.

- Animación tipo collage. El objetivo es permitir crear una escena animada compuesta de múltiples imágenes, de forma semejante a un collage (véase la Figura 3.12). Funciona de manera similar a la animación cíclica pero con algunas diferencias:
 - Las imágenes aparecen, realizan una transición completa y desaparecen.
 - Se introduce cierta aleatoriedad, por ejemplo, la rotación, el escaldado y la ubicación iniciales son elegidos al azar. Además, el número de imágenes simultáneas en la pantalla no está limitado por el número

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

88

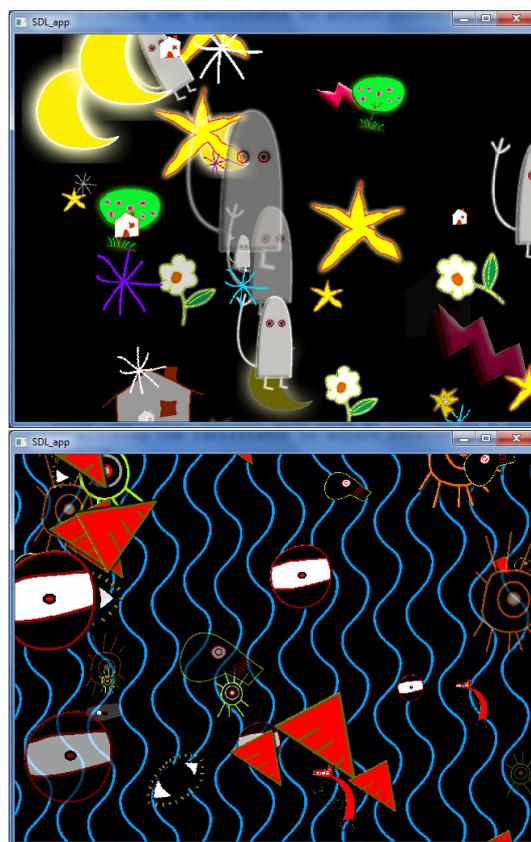


Figura 3.12: Capturas de animación collage en SITPLUS.

de imágenes disponibles, es decir, una imagen dada puede aparecer varias veces en la pantalla. Actualmente, es posible seleccionar entre 1 y 100 imágenes simultáneas, siendo 20 el valor predeterminado. Finalmente, hay un tipo de transición adicional, RANDOM, que selecciona al azar uno de los tipos de transición disponibles enumerados anteriormente.

- Se puede establecer una opción que hace desaparecer todas las imágenes después de unos segundos de inactividad.

Además, es posible definir un fondo para completar la escena. Este fondo puede ser una imagen estática, una secuencia animada de imágenes (cada imagen se muestra durante un cierto intervalo de tiempo que es independiente del valor de entrada) o la imagen de la cámara.

La descripción de una animación se realiza a través de un archivo XML. SITPLUS ofrece una colección de imágenes y animaciones predefinidas, pero se pueden definir nuevas animaciones que se adapten a las necesidades de los

usuarios. Por ejemplo, un collage de fotos con imágenes que resulten familiares o motivadoras al usuario.

3.3.7. Actividades

Actualmente nuestro sistema ofrece dos tipos de actividad distintas. En ambos casos el objetivo principal es dejar claro que, cuando hay actividad en la entrada (gestos físicos o sonidos), el sistema responde generando una respuesta sonora y gráfica. Esta respuesta, en lugar de ser una reacción preprogramada que se dispara al pulsar un conmutador, consiste en un bucle continuo de gesto-retroacción [18].

3.3.7.1. Jugando con la voz

Usando un micrófono, se recogen las emisiones orales del usuario u otros sonidos, como la respiración o pequeños golpes sobre este dispositivo. En los experimentos, se han probado micrófonos inalámbricos, tanto de mano, de diadema como de solapa, aunque los micrófonos convencionales de sobremesa también funcionaron adecuadamente. La selección de un determinado tipo de micrófono depende de las capacidades del usuario (por ejemplo, si el usuario es capaz de sujetarlo) y sus preferencias personales. El micrófono más utilizado es el de solapa ya que es el que menos molestias supone. Muchas veces también es el propio facilitador quien sujeta el micrófono a la altura de boca del usuario.

El flujo de entrada de audio se analiza para extraer su amplitud y luego se procesa a través de diferentes efectos digitales de audio como se describe en la Sección 3.3.4.2. El audio resultante es entonces amplificado y enviado a los altavoces. El valor de la amplitud se utiliza para alimentar el módulo gráfico, de manera que, cuando se detecta algún sonido, se hace evolucionar la animación gráfica. Con el fin de crear una experiencia inmersiva, se utiliza un sistema con cuatro altavoces que envuelven al usuario, no obstante hemos probado con éxito el sistema usando unos altavoces de escritorio con *subwoofer*.

La actividad cuenta con un panel de control (véase la Figura 3.13) que permite ajustar varios parámetros durante la sesión. Este panel de control se divide en dos áreas. El área izquierda contiene los ajustes de procesamiento de audio como el volumen principal (con una opción para amplificar la entrada de micrófono cuando la señal es muy débil), la intensidad de los diferentes efectos de audio, la sensibilidad (multiplica la amplitud de la envolvente del sonido en la entrada por el valor seleccionado mediante el control deslizante) y el umbral (corta a cero la amplitud resultante cuando es inferior a este valor). La casilla de verificación “Sonido detectado” indica si la señal está por encima del umbral y por lo tanto sirve para ajustar este valor. Sin la opción de um-

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

90

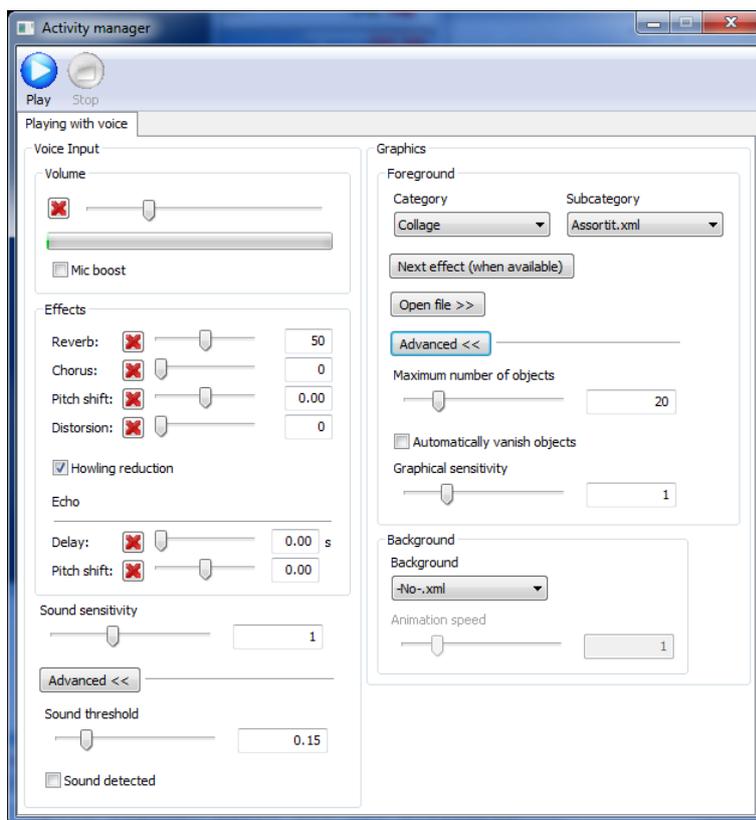


Figura 3.13: Captura del panel de control de la actividad “Jugando con la voz”.

bral, el ruido de fondo haría evolucionar la animación gráfica continuamente, lo que podría confundir al usuario cuando permaneciese en silencio. El área de la derecha contiene la configuración de la salida gráfica que incluye dos menús desplegables para seleccionar la animación específica (opcionalmente, el facilitador puede abrir un archivo que contiene otra animación específica), el número máximo de objetos simultáneos en pantalla, una casilla de verificación para permitir que las imágenes desaparezcan cuando cuando no se detecta actividad durante unos instantes (estas dos opciones sólo se aplican para las animaciones tipo collage) y la sensibilidad gráfica (otro multiplicador para ajustar el grado en el que evolucionará la animación en función del valor de entrada). Bajo las opciones anteriormente descritas, están los ajustes de selección de fondo que permite indicar si se mostrará un fondo y cuál (el sistema proporciona varios fondos de muestra tanto estáticos como animados). Una vez activado, la velocidad de la animación también puede ajustarse.

La selección de un efecto de audio concreto depende de la experiencia previa con el usuario y la evaluación del facilitador durante la sesión. Por lo

general, se comienza con una reverberación suave. Este efecto introduce cierta persistencia en el sonido que permite al usuario escuchar y reconocerse a sí mismo. A veces, si el usuario tiene una producción oral pobre, utilizamos un eco de período corto lo que repite varias veces el sonido y muestra al usuario que, incluso el más leve ruido, produce resultados. Cuando la relación causal entre los sonidos emitidos y escuchados se entiende (esto puede durar unos pocos segundos, minutos, varias sesiones o no llegar a suceder), se introducen otros efectos, que son menos evidentes como el cambio de tono o distorsión. El eco de período largo (entre 2 y 10 segundos), por ejemplo, permite al usuario generar un resultado complejo que mezcla el audio de sonidos anteriores con los nuevos.

En cuanto a la salida gráfica se refiere, por lo general se utiliza una animación basada en collage o una animación cíclica abstracta. A veces, y si está disponible, se utiliza una animación personalizada específicamente para el usuario (por ejemplo, tenemos un usuario que está encantado de ver imágenes de sí mismo en la pantalla y por ello le hemos creado un collage personalizado con fotografías suyas).

3.3.7.2. Instrumento musical virtual

Esta actividad permite, usando el detector de movimiento basado en visión artificial o un mando de Wii, reproducir sonidos que imitan diferentes instrumentos musicales de acuerdo a los gestos del usuario.

La relación entre los eventos de movimiento y las notas musicales interpretadas depende del sistema de entrada empleado. Cuando se utiliza un modo de entrada de tipo absoluto (es decir, el detector de movimiento grueso basado en visión o la Wii Balance Board), la posición absoluta a lo largo de un eje (se puede escoger entre X [horizontal] o Y [vertical]) se asigna directamente a una nota de la partitura. Cuando se utiliza un modo de entrada relativo, es decir, el detector de movimiento fino basado en visión o el mando de Wii, el movimiento calculado se acumula y el resultado se utiliza para seleccionar una nota de la partitura. De esta manera, cualquier movimiento a lo largo de un eje hace sonar las notas. El módulo del movimiento detectado se utiliza para hacer evolucionar la animación gráfica.

La actividad cuenta con un panel de control (véase la Figura 3.8) que permite ajustar interactivamente varios parámetros durante la sesión. Este panel de control está dividido en tres áreas. El área izquierda contiene la configuración de la entrada gestual, tanto la basada en visión artificial como en los mandos de Wii. Contiene el visor de la cámara que permite ajustar la región de interés (recuadro naranja) y la dirección principal de movimiento que se extraerá. Bajo éste se muestra el control deslizante de la sensibilidad, el

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

92

del umbral (que permite al usuario detener completamente la generación de sonido y la animación gráfica cuando no hay actividad) y una opción para mostrar la imagen de la cámara como un fondo en la ventana gráfica. El área central contiene controles para el secuenciador MIDI, incluyendo, el control de volumen, el de selección del instrumento (se muestra un subconjunto de los más comunes entre los 128 posibles de la especificación General MIDI), la partitura musical (que también puede ser cargada desde archivo) y, como ajustes avanzados, la duración de la nota (3 segundos por defecto), la sensibilidad (permite ajustar la sensibilidad del secuenciador independientemente del resto, aunque no se utiliza normalmente) y si la partitura se considerará cíclica (activado por defecto). Finalmente, el panel de la derecha es el mismo que para la actividad “Jugando con la voz”.

Antes de comenzar, el facilitador decide el área de donde se extraerá el movimiento teniendo en cuenta las capacidades físicas del usuario y la experiencia en sesiones anteriores. La situación más común es cuando se utiliza el movimiento de la cabeza, pero también trabajamos con las extremidades superiores o movimiento de las manos, incluso movimientos más leves de los labios o de los párpados en casos extremos.

Hay varias partituras predefinidas, desde escalas naturales o cromáticas, a escalas pentatónicas o más complejas. Por lo general, se comienza con una escala simple, de forma que las notas aumentan de tono cuando se mueve en una dirección, y decaen cuando se mueve en la dirección opuesta. Cuando el usuario comprende bien esta mecánica, se cambia a escalas más complejas. El instrumento seleccionado depende de las preferencias personales (en el caso de que se conozcan) y la experiencia previa con el usuario. Entre los instrumentos más utilizados están el piano, el saxofón, la batería, la guitarra y la caja de música.

En cuanto a la representación gráfica, el razonamiento es prácticamente el mismo que para la actividad anterior, pero para algunos instrumentos musicales específicos se selecciona una animación concreta del instrumento en cuestión. Por ejemplo, si se selecciona el sonido de guitarra se suele acompañar con una animación cíclica de una guitarra.

Una vez iniciada la actividad, el facilitador ajusta la sensibilidad empíricamente de acuerdo con el rango de movimiento del usuario.

3.3.8. Estado actual del proyecto

El proyecto se empezó a poner en práctica a finales de 2007 y partir del curso 2008-09 se convierte en una actividad que se lleva a cabo con carácter regular en la APPC hasta hoy.

En el año 2010 se pone a disposición del público el software, lo que ha

permitido que se esté utilizando en otras organizaciones como:

- Dentro de la red ASPACE:
 - Fundación Numen (Madrid)
 - ASPACE Barcelona
 - ASPACE Granada
 - ASPACE Baleares
- Escuela E.E. Jean Piaget (Granada)
- Fundació el Maresme (Mataró)
- Música a l'abast (Barcelona)

Además aparece en catálogos especializados como Open Source Assistive Technology Software (OATS) [117] y desde mediados de 2011 también se incluye en la distribución Debian GNU/Linux.

Ha recibido el apoyo económico de entidades la Generalitat de Catalunya, la Fundació .CAT, Caixa Tarragona y Fundación Repsol, acumulando un inversión total que supera los 150.000 euros.

Por último, el proyecto SITPLUS ha inspirado el proyecto GAME-ABLING [44] que pretende crear una plataforma de juegos para personas con parálisis cerebral. Dichos juegos estarán diseñados teniendo en cuenta las especiales características de estas personas y podrán usarse mediante dispositivos como micrófonos, cámaras web, mandos de la consola Wii y Kinect. Se trata de un proyecto a dos años financiado por la Comisión Europea. En el momento de escribir estas líneas, GAME-ABLING se encuentra en su segundo año de desarrollo.

3.4. Switch Viacam

La experiencia y los esfuerzos dedicados a los proyectos antes mencionados y, en particular, el desarrollo de la biblioteca libsitplus (véase la Sección 3.3.2.1), nos ha permitido desarrollar otras aplicaciones relacionadas. Este es el caso del proyecto Switch Viacam (sViacam) [86].

Switch Viacam es un emulador de conmutador basado en cámara web. Es decir, realiza la misma función que un conmutador convencional (véase la Sección 3.1.1.1) pero usando una cámara web y técnicas de visión artificial. Al igual que el resto de proyectos, también se distribuye como software libre.

Su funcionamiento se basa en que, utilizando una cámara web, permite configurar una área concreta dentro de la cual se detecta movimiento, lo que

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

94

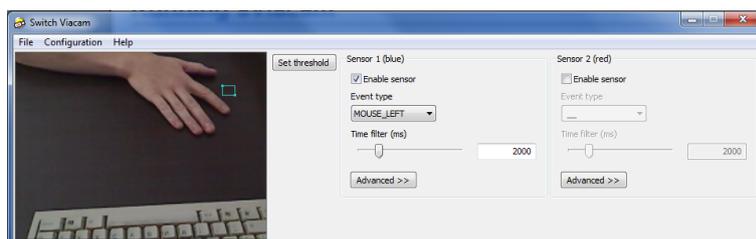


Figura 3.14: Captura de la pantalla principal de Switch Viacam.

dispara un clic del ratón o la pulsación de una tecla. En nuestro caso, podemos seleccionar una región rectangular de la imagen capturada donde se realizará la detección (véase la Sección 3.2 para una descripción detallada de la técnica utilizada para extraer movimiento). Una vez que la cantidad del movimiento (específicamente, el módulo, sin importar la dirección) supera cierto umbral, se activa la acción predefinida.

La interfaz de usuario de Switch Viacam se representa en la Figura 3.14. Su uso es bastante simple y consta de los siguientes pasos:

1. El facilitador coloca la cámara de manera adecuada, de modo que el movimiento que el usuario puede realizar aparece dentro de la imagen. Generalmente es buena idea enfocar una superficie (por ejemplo una mesa o pared) para evitar eventos de espurios provocados por el ruido de fondo, esto también permite colocar un adhesivo de modo que el usuario tenga una referencia física de donde disparar la acción.
2. El facilitador activa uno de los sensores haciendo clic en la casilla de verificación “Activar sensor”. Un cuadrado de color (sensor) aparece sobre la imagen. Dentro de esta área es donde se detecta movimiento.
3. El facilitador coloca el sensor y cambia el tamaño de éste para seleccionar el área donde se quiere detectar el movimiento. Si es necesario, el control deslizante “Filtro de tiempo (ms)” permite ajustar el intervalo (en milisegundos) que sViacam esperará tras la generación de un evento hasta que permita generar el siguiente. Por defecto este tiempo es de 2 segundos.
4. El facilitador hace clic en el botón “Establecer umbral”. Esto ajustará automáticamente el umbral de movimiento necesario por encima del cual se dispara un evento. Opcionalmente, este umbral se puede ajustar manualmente. Para ello, el facilitador tiene que abrir las opciones avanzadas haciendo clic en el botón “Avanzado” y luego ajustar el valor de “Umbral detección movimiento” al valor deseado.

5. Por último, el facilitador selecciona el tipo de evento que se generará a través de la lista desplegable “Tipo de evento”. El valor adecuado depende de la aplicación externa que vaya a utilizarse, permitiendo escoger entre varios eventos de teclado y ratón.

Switch Viacam está destinado a personas que encuentran muy difícil o imposible el uso de dispositivos convencionales (por ejemplo, teclados, ratones, *joysticks*, etc.) debido a sus limitaciones físicas y, a menudo, cognitivas.

El objetivo es enriquecer la gama de herramientas ya existentes y su uso resulta muy conveniente bajo ciertas condiciones de trabajo. Cuenta con una serie de ventajas tales como la simplicidad, ya que no se requiere cables u otros elementos, la flexibilidad, ya que permite elegir prácticamente cualquier parte del cuerpo con la que trabajar, y el bajo coste, porque sólo requiere un ordenador con cámara.

3.5. Herramienta on-line de anotación de vídeo

Como vimos en la Sección 2.6.2, la anotación de vídeos es una técnica comúnmente utilizada para estudiar, entre otros, el comportamiento de niños con parálisis cerebral bajo diferentes estímulos. Tomar estas notas requiere ver el vídeo varias veces en un proceso costoso, que consume mucho tiempo y que, a veces, resulta bastante frustrante.

En este apartado presentamos una aplicación² para anotar vídeo en línea, es decir, mientras se produce la acción. Nuestra aplicación puede ejecutarse en cualquier dispositivo portátil con iOS (por ejemplo, un iPhone), y de forma inalámbrica almacena la información en un equipo remoto. Gracias a su sencilla interfaz, especialmente diseñada para capturar las principales acciones de los niños con parálisis cerebral, nuestra aplicación permite a los cuidadores/facilitadores para prestar atención a los niños y tomar notas al mismo tiempo. Como resultado, se reduce drásticamente el tiempo necesario para anotar sesiones (fuera de línea) y los costes económicos asociados.

3.5.1. Motivaciones

Tras varios años de experimentación, los beneficios cualitativos del proyecto SITPLUS son evidentes. Sin embargo, los datos cualitativos y cuantitativos son necesarios para apoyar nuestras observaciones diarias y, por desgracia, la recogida de estos datos ha demostrado ser una tarea costosa, compleja y que consume mucho tiempo.

²Desarrollada por Pablo A. Pérez Martínez, colaborador en nuestra investigación.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

96

De acuerdo con nuestra experiencia trabajando con personas con parálisis cerebral, la mayoría de ellas no pueden expresar sus sentimientos oralmente o por medio de expresiones faciales o gestos. En tales casos, son necesarios expertos con un conocimiento profundo acerca de los participantes para interpretar correctamente sus reacciones. Como se ha dicho antes, nos referimos a estos expertos como facilitadores.

Al evaluar el comportamiento de las personas con parálisis cerebral con el proyecto SITPLUS (o cualquier otro sistema informático), las técnicas tradicionales de la Interacción Persona Ordenador (IPO) (por ejemplo, pensar en voz alta, entrevistas, cuestionarios, etc.) no se puede utilizar. Por lo tanto, los investigadores se ven obligados a depender de las observaciones de comportamiento que, en la actualidad, por lo general implica grabar las sesiones experimentales en vídeo para su posterior análisis, a menudo con la ayuda de herramientas de anotación de vídeo (para nuestra investigación hemos utilizado ANVIL [64]).

El análisis y la codificación de una grabación de vídeo es una tarea que consume tiempo [49] lo que implica retrasos para los investigadores y una carga de trabajo adicional para los facilitadores (hay que tener en cuenta que su intervención es esencial en algunos casos). Además, en muchos casos, los recursos disponibles impiden que los facilitadores dediquen su tiempo a actividades de investigación (por ejemplo, a anotar vídeos fuera de línea).

Con el objetivo de recoger las evaluaciones de los facilitadores y, al mismo tiempo, reducir su carga de trabajo, hemos desarrollado una aplicación que se ejecuta en un dispositivo móvil (por ejemplo, el iPhone) y permite la anotación de varias variables durante el transcurso de una sesión (es decir, mientras el facilitador ya está participando en la sesión).

3.5.2. Requisitos

En el diseño de nuestra aplicación hemos tenido en cuenta los siguientes requisitos:

- **Fácil de usar.** La aplicación debe ser simple e intuitiva. Los facilitadores deben ser capaces de utilizarla al mismo tiempo que prestan atención al participante. Además, aprender a usarla no debería requerir mucho esfuerzo.
- **Inalámbrica.** La aplicación no debe limitar los movimientos físicos del facilitador. Así, utilizar un dispositivo móvil parece ser la mejor opción.
- **En línea/interactiva.** La aplicación debería permitir a los facilitadores registrar fácilmente eventos predefinidos, reacciones y estados de los

3.5. Herramienta on-line de anotación de vídeo

97

participantes, mientras se les está presentando atención.

- **Compatible.** La aplicación debe ser interoperable con herramientas comunes de anotación de vídeo como ANVIL.
- **Orientada la parálisis cerebral.** La aplicación debe tener en cuenta las principales acciones que las personas con parálisis cerebral pueden llevar a cabo.

3.5.3. Arquitectura

Nuestra propuesta se basa en la clásica arquitectura cliente-servidor. Por un lado, tenemos un ordenador personal que actúa como servidor, esto es, acepta conexiones entrantes TCP³. En el otro lado, tenemos un dispositivo móvil (el iPhone en nuestro caso) que es capaz de conectarse de forma inalámbrica con el servidor y enviar información al mismo.

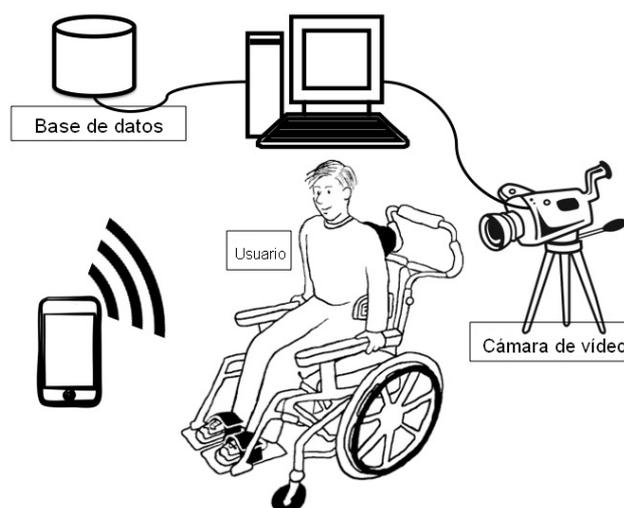


Figura 3.15: Esquema general de la arquitectura.

El esquema general de la arquitectura se representa en la Figura 3.15. El participante se coloca en el centro del proceso, mientras que una cámara de vídeo registra y envía las sesiones a un ordenador (es decir, al servidor) que los almacena en una base de datos local. Al mismo tiempo, un facilitador utiliza el iPhone (es decir, el cliente) para registrar la información sobre el

³TCP corresponde a las siglas en inglés de *Transmission Control Protocol*, es decir, Protocolo de Control de Transmisión. Es un protocolo estándar ampliamente utilizado que se define en la RFC 793.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

98

estado del participante. Esta información se transmite de forma inalámbrica a la computadora y se almacena en la base de datos local en forma de XML⁴. Estos archivos XML se pueden cargar fácilmente en ANVIL junto con los vídeos grabados.

3.5.4. Interfaz de usuario

Hemos diseñado una interfaz simple que permite al facilitador tomar notas fácilmente de los principales eventos mientras que presta atención al participante. Nuestra interfaz, que se representa en la Figura 3.16, utiliza varios iconos para ilustrar las acciones más comunes realizadas por los participantes. (Hay que tener en cuenta que el rango de acción de una persona con parálisis cerebral es muy limitado.)



Figura 3.16: Capturas de pantalla de la herramienta de anotación.

De izquierda a derecha y de arriba a abajo, las imágenes de la interfaz de usuario propuesta representan (i) “Sonrisa”, (ii) “Sonrisa larga/carcajada”, (iii)

⁴XML corresponde a las siglas en inglés de *Extensible Markup Language* y es comúnmente utilizado para el intercambio de información estructurada.

3.5. Herramienta on-line de anotación de vídeo

99

“Participación activa”, (iv) “Distracción/Pérdida de atención”, y (v) “Emisión oral”.

“Risa” y “Emisión oral” son variables momentáneas que se registran como un evento instantáneo que sucede en un momento dado. Por el contrario, “Risa larga/carcajada”, “Participación activa” y “Distracción/Pérdida de atención” son variables duracionales, que se registran como un intervalo de tiempo (es decir, con un evento de *inicio* y de *fin*). Los eventos duracionales están etiquetados con el texto “[D]”, además, el icono cambia de color mientras está sucediendo la acción. Gracias al resaltado de color, es muy fácil para el facilitador ver si la acción está siendo registrada por el dispositivo. La Figura 3.17 muestra una imagen detallada de un icono enfatizado.



Figura 3.17: Eventos duracionales enfatizados.

Cuando se inicializa la interfaz de usuario todos los iconos se muestran en blanco y negro. Cuando se presiona un icono que representa una variable momentáneo, se muestra un efecto de parpadeo para confirmar la acción⁵. Si el facilitador pulsa un icono que representa una variable duracional, el icono cambia de color, lo que indica que la acción, evento o estado está en curso. Si el icono se presiona de nuevo (al finalizar la acción), el icono vuelve a su forma original en blanco y negro. Mediante el uso de esta técnica, el facilitador no tiene que recordar si ha presionado el icono, simplemente tiene que observar la presencia o ausencia de color para saber si la acción o estado está en curso. Es evidente que algunas acciones o estados pueden ocurrir simultáneamente, mientras que otros son incompatibles. Por ejemplo, es imposible que un participante se muestre participativo y distraído al mismo tiempo. Por

⁵No hemos incluido sonido para evitar cualquier interferencia con los participantes.

Capítulo 3. Nuestras propuestas de técnicas y sistemas interactivos

100

lo tanto, para simplificar la tarea del facilitador, todas estas limitaciones han sido programados en la interfaz.

Por último, en la parte inferior de la interfaz, se ha añadido un botón de “Arranque/Parada” que permite al facilitador sincronizar el envío de información al servidor.

3.5.5. Integración con ANVIL

Como se indicó anteriormente, el dispositivo móvil envía la información al servidor. A continuación, el servidor convierte esta información en un archivo XML que se puede cargar en ANVIL y ya sincronizado con la grabación en vídeo. Como resultado, ANVIL se puede utilizar para visualizar las notas tomadas por el facilitador durante la sesión y, si es necesario, se pueden añadir comentarios adicionales fuera de línea. La Figura 3.18 muestra una captura de pantalla de ANVIL donde se destacan los eventos anotados.



Figura 3.18: Herramienta de anotación integrada con ANVIL.

CAPÍTULO 4

Resultados

En este capítulo se presentan y se discuten los resultados de los experimentos llevados a cabo con distintas personas con parálisis cerebral.

Índice

4.1. Resultados evaluación de la interacción	102
4.1.1. Caso 1: A.A.	103
4.1.2. Caso 2: C.G.	105
4.2. Resultados proyecto Switch Viacam	107
4.2.1. Caso 1: Y.P.	108
4.2.2. Caso 2: Q.H.	108
4.2.3. Ventajas	108
4.3. Resultados proyecto SITPLUS	109
4.3.1. Entorno de experimentación	109
4.3.2. Procedimiento	110
4.3.3. Resultados cuantitativos	111
4.3.4. Resultados cualitativos	113

El análisis del uso de nuestros sistemas por personas con parálisis cerebral supone un importante reto. En primer lugar por la interdisciplinariedad de la tarea, y en segundo lugar porque trabajar con personas con parálisis cerebral en grado moderado y severo no resulta sencillo: no siempre se muestran cooperativas, la comunicación (cuando es posible) es a menudo muy complicada, muchos de los usuarios tienen una salud delicada, presentando en ocasiones crisis epilépticas u otro tipo de convulsiones y generalmente los usuarios no son capaces de expresar claramente sus sentimientos, preferencias y opiniones.

Las especiales características de nuestros participantes invalidan el uso de la mayoría de los métodos clásicos para evaluar la experiencia de usuario, lo que obliga a recurrir a otros métodos de análisis que permitan prescindir de, o al menos reducir, la colaboración activa por parte del usuario.

La obtención de resultados cuantitativos supone también un reto complejo, en primer lugar por la necesidad de interpretar las respuestas de los usuarios,

lo que obliga a realizar observaciones detalladas de la situación interactiva (*ej.* a través de grabaciones en vídeo). Además, la heterogeneidad y el número reducido de individuos con parálisis cerebral hacen difícil la realización de experimentos a gran escala, el uso de grupos control y la generalización de resultados. Por ello, creemos imprescindible completar los datos cuantitativos con valoraciones de carácter cualitativo en base al estudio de casos particulares.

En efecto, dada la naturaleza de nuestra propuesta, estamos convencidos de la importancia de combinar el análisis cuantitativo y cualitativo. En [91] se señalan las características que distinguen a la investigación cualitativa de la cuantitativa (que son aplicables a nuestro trabajo). Por otra parte, el análisis cualitativo adquiere todavía más relevancia en los experimentos aquí expuestos. Las condiciones personales de los participantes no pueden ser tomadas en cuenta sólo con el análisis de datos numéricos. Los facilitadores, maestros y cuidadores comparten con ellos en todas sus actividades, por lo que pueden aportar un punto de vista (cualitativo), sin el cual el experimento carece de sentido.

4.1. Resultados evaluación de la interacción

Inicialmente se realizaron una serie de pruebas preliminares con un grupo de cuatro usuarios, dos hombres y dos mujeres, con edades comprendidas entre los 7 y los 24 años, con limitaciones motrices desiguales y distintos grados de retardo mental.

Se probó con la actividad de respuesta simple a un estímulo, concretamente 4 rondas de 5 iteraciones cada una. Excepto para uno de los usuarios, el resto fue incapaz de completar más allá del 35 % de la prueba debido a su falta de atención y cansancio. El usuario que mejor lo hizo fue capaz de completar el 90 % de la prueba. No obstante, para éste, los resultados obtenidos presentaron una desviación estándar muy elevada ($m = 1,2s, \sigma = 1,05s$).

Por lo tanto, estos resultados iniciales nos hacen descartar aquellos candidatos con mayores afectaciones cognitivas que tienden a prestar menor atención y cansarse antes (lo que introduce retardos variables que distorsiona la medida del tiempo para acceder al dispositivo físico). Para el resto de usuarios, habría que considerar también la desviación estándar como un indicador relevante, especialmente para personas con espasticidad, donde habría que preferir aquellas opciones de acceso con menor desviación ya que suelen indicar un acceso más estable (aunque sea más lento).

En pruebas posteriores se ha aplicado este método de análisis en los dos casos siguientes:

4.1.1. Caso 1: A.A.

A.A. es un joven de unos 20 años con parálisis cerebral y nivel cognitivo conservado. Tiene severas limitaciones físicas que le impiden usar sus manos para cualquier tipo de manipulación. Se desplaza en una silla de ruedas motorizada que controla mediante 3 conmutadores montados en el cabecero de la misma y accede al ordenador mediante el sistema SmartNav¹ que consiste en una cámara de infrarrojos que hace el seguimiento de una pegatina reflectante que se adhiere, usualmente, sobre la frente del usuario. También dispone de un teléfono con pantalla táctil que maneja con la punta de su nariz.

A.A. expresa que está satisfecho con su sistema de acceso al ordenador, aunque en situaciones como en exteriores con mucha iluminación, nos indica que no le funciona correctamente. Por ello le hemos ofrecido Enable Viacam como alternativa en esos casos ya que no tiene esta restricción de funcionamiento.

Para comparar el funcionamiento de ambos sistemas de acceso, se le ha solicitado realizar la misma prueba de rendimiento con cada uno de los dispositivos antes mencionados, concretamente la tarea de posicionar y seleccionar con el puntero (véase la Sección 3.1.5.4). Se ha empleado un ordenador portátil con pantalla de 17 pulgadas. En ambos casos se ha pedido que haga clic a un total de 9 objetivos circulares de 50 píxeles de diámetro separados entre ellos por una distancia de 300 píxeles.

	Tiempo camino (s)	Nº errores	Nº re-entradas	Rendimiento (bps)	Desviación trayectoria (px)
	3,02	0	0	0,89	190,62
	3,55	0	0	0,75	259,93
	2,51	0	0	1,08	123,07
	2,50	0	0	1,09	127,92
	2,37	0	0	1,19	38,91
	2,48	0	0	1,12	25,73
	2,41	0	0	1,17	91,63
	3,06	0	0	0,88	269,42
Media	2,74	0	0	1,02	140,90
σ	0,42	0	0	0,16	92,38

Tabla 4.1: Resultados con SmartNav.

En la Tabla 4.1 se presentan los resultados obtenidos con SmartNav. Se han omitido los datos de tiempo de reacción puesto que no existía retardo entre

¹<https://www.naturalpoint.com/smartnav/>

las distintas iteraciones (es decir, una vez se había hecho clic sobre uno de los objetivos inmediatamente se resaltaba el siguiente). En cuanto al tiempo de camino (o tiempo de ejecución) es poco más de 2,7s en promedio y se mantiene bastante estable. Hay que tener en cuenta que una parte de este tiempo, un segundo concretamente, es el consumido por el mecanismo de clic por espera. En esta prueba el usuario no ha cometido ningún error y no ha necesitado re-entrar al objetivo en ninguna ocasión, dato que se ve corroborado por la desviación relativamente reducida de las trayectorias respecto de las ideales (rectilíneas).

	Tiempo camino (s)	Nº errores	Nº re-entradas	Rendimiento (bps)	Desviación trayectoria (px)
	3,73	0	1	0,75	917,93
	3,92	0	1	0,69	602,56
	2,39	0	0	1,18	166,57
	5,74	1	2	0,49	958,77
	2,353	0	0	1,23	91,60
	15,47	3	6	0,18	2571,01
	2,52	0	0	1,11	133,52
	3,14	0	0	0,86	32,13
media	4,91	0,50	1,25	0,81	684,26
σ	4,41	1,07	2,05	0,36	848,89

Tabla 4.2: Resultados con Enable Viacam.

Al realizar la misma prueba con Enable Viacam (véase la Tabla 4.2). Vemos claramente que los resultados no son tan buenos. En general, e incluso eliminando la fila 6 que presenta valores claramente atípicos, todos los registros son peores incluyendo tiempos de ejecución más elevados, presencia de algunos errores y re-entradas al objetivo, y trayectorias que se apartan bastante de las ideales. Estos resultados se corresponden con el criterio del usuario que manifiesta que prefiere el primer sistema.

Las diferencias de funcionamiento entre ambos sistemas se pueden deber principalmente a dos razones:

- Falta de entrenamiento. Este usuario maneja con asiduidad el primero de los sistemas desde hace varios años y ha adquirido un nivel elevado de destreza. Sin embargo, antes de estas pruebas, A.A. sólo ha tenido ocasión de manejar Enable Viacam durante unos minutos. Por ello, creemos que con entrenamiento estos registros mejorarían notablemente.

- Limitaciones en el funcionamiento. Es posible que el segundo sistema presente limitaciones en el funcionamiento y, específicamente, una respuesta menos lineal que el primero lo que también incide negativamente en los resultados obtenidos (y la impresión subjetiva del usuario).

Como trabajo futuro nos hemos propuesto mejorar la linealidad de la respuesta del sistema Enable Viacam haciendo detección y seguimiento de características concretas de la imagen en lugar de calcular el flujo óptico de toda una región. Además, hemos sugerido al usuario empiece a usar nuestro sistemas en aquellas situaciones (por ejemplo, en exteriores con mucha luz solar) donde SmartNav no funcione correctamente. Con ello esperamos mejorar también el entrenamiento del usuario.

4.1.2. Caso 2: C.G.

C.G. es una mujer de 38 años con parálisis cerebral tipo tetraparesia espástica debido a una anoxia perinatal y presenta una afectación severa de la movilidad global y funcional. Es completamente dependiente para las actividades de la vida diaria. Se desplaza en una silla de ruedas a motor que, de momento, ella no es capaz de conducir. Dado su elevado nivel de disartria y atendiendo a la demanda que ella formuló, desde hace un año aproximadamente se ha introducido un sistema de comunicación alternativo que consiste en un Tablet PC en el cual se ejecuta un software donde ella puede seleccionar distintos pictogramas organizados en forma de cuadrícula. Para ello utiliza un método de barrido secuencial dirigido con dos conmutadores. El primer conmutador, que permite avanzar al siguiente elemento, lo tiene montado junto a su mano derecha y lo acciona con un ligero movimiento. El segundo conmutador, que le permite seleccionar el elemento resaltado, lo tiene montado en el cabecero de su silla y lo acciona con un ligero movimiento hacia atrás de la cabeza.

Si bien el sistema de acceso le resulta funcional y ella no expresa rechazo ni demanda de mejora, hemos querido analizar hasta que punto podría resultar más ventajoso trasladar el conmutador del cabecero de la silla para que lo tuviera al alcance de su mano izquierda. Para ello hemos estudiado la respuesta a un estímulo simple (véase la Sección 3.1.5.1) en 8 iteraciones con cada uno de los pulsadores. Se ha fijado un tiempo aleatorio entre 1 y 4 segundos entre la respuesta del usuario y la presentación del estímulo siguiente. Las tablas 4.3, 4.4 y 4.5 presentan los datos obtenidos con el conmutador en la mano derecha, la cabeza y la mano izquierda respectivamente.

	tiempo de respuesta (s)	nº de errores
	3,67	0
	1,67	0
	2,36	0
	2,81	0
	4,75	1
	2,44	1
	2,58	0
	1,94	0
media	3,13	
σ	1,18	
Total errores		2
Duración total tarea	25 s	

Tabla 4.3: Resultados conmutador mano derecha.

	tiempo de respuesta (s)	nº de errores
	3,58	0
	3	0
	3,02	0
	2,31	0
	3,41	0
	1,53	0
	2,02	0
	1,55	0
media	2,55	
σ	0,81	
Total errores		0
Duración total tarea	23,85 s	

Tabla 4.4: Resultados conmutador cabeza.

4.2. Resultados proyecto Switch Viacam

107

	tiempo de respuesta (s)	nº de errores
	2,58	1
	1,78	0
	0,97	1
	1,19	0
	0,62	0
	1,78	1
	1,25	0
	0,84	0
media	1,38	
σ	0,64	
Total errores		3
Duración total tarea	14,38 s	

Tabla 4.5: Resultados conmutador mano izquierda.

A la hora de interpretar los resultados, parece que el conmutador del cabecero es el que mejores resultados da en términos de número de errores (no cometió ninguno en esta ronda de pruebas) y de estabilidad (menor desviación estándar respecto la media en términos relativos). En cuanto a tiempo de respuesta, puede dar la sensación de que el conmutador en la mano izquierda es el más rápido, no obstante el resultado se explica porque la usuaria tenía tendencia a presionar en repetidas ocasiones sin esperar a que el estímulo hubiera aparecido, de ahí el número de errores cometido que es el más elevado de todas las pruebas realizadas.

A pesar de que los datos no son concluyentes, sí existen indicios de que el funcionamiento del conmutador del cabecero es conveniente y que podría haber duda en cuanto que mano a utilizar. Ante tal disyuntiva, el equipo de fisioterapeutas de C.G. nos indica que es preferible potenciar el uso del lado derecho de su cuerpo dado su condición física, por lo que finalmente se desecha la posibilidad de que use un conmutador con su mano izquierda.

Además, dado que las pruebas dan indicios de que tiene mejor control de cabeza que de extremidades, nos planteamos como trabajo futuro el situar el conmutador derecho a la altura de su sien derecha.

4.2. Resultados proyecto Switch Viacam

Los primeros ensayos llevados a cabo en la APPC han mostrado que la herramienta resulta útil y eficaz, especialmente para actividades de rehabilitación física. A continuación se reportan dos casos:

4.2.1. Caso 1: Y.P.

Y.P. es una joven de 28 años con hemiplejía izquierda. Sandra, su fisioterapeuta, quiere que integre su extremidad superior izquierda en su esquema corporal, que la acepte, porque por lo general no desea utilizarla ni le gusta que nadie se la coja. Como a Y.P. le encanta la música, especialmente las canciones infantiles, Sandra decidió sacar provecho de ello y propuso una actividad en la que Y.P., con su mano sana, mueve su mano pléjica para hacer que se reproduzca una canción. Después de reproducirse durante unos segundos, Sandra detiene la canción y le pide a Y.P. que vuelva a activarla para continuar escuchándola.

4.2.2. Caso 2: Q.H.

Q.H. es un joven de 26 años con hemiplejía izquierda. Ha sido operado recientemente de la columna vertebral y sus músculos erectores siguen estando débiles. Isabel, su fisioterapeuta, quiere que permanezca de pie. Q.H. ama la música de su cantante favorito y por ello Isabel le propone una actividad en la que tiene que iniciar el reproductor de música levantando su mano izquierda mientras está de pie. Gracias a esta actividad, Q.H. se mantiene de pie y también usa su brazo enfermo.

4.2.3. Ventajas

En comparación con un conmutador mecánico, destacamos las siguientes fortalezas que han sido corroboradas por las fisioterapeutas antes referidas:

- **Robustez física y seguridad.** Algunos usuarios tienen muy poco control de sus movimientos voluntarios y de la fuerza que pueden ejercer. Por lo tanto, cuando se utilizan conmutadores convencionales existe el riesgo de romperlos o de hacerse daño a uno mismo.
- **Flexibilidad.** Permite definir un área de trabajo de una manera muy flexible (se puede desplazar, ampliar y reducir de forma simple). Se puede trabajar con movimiento tanto grueso como fino. Con algunos usuarios es mejor no colocar ninguna referencia física (por ejemplo, una pegatina) ya que puede suponer una distracción, mientras que con otros usuarios es justamente lo contrario. Además, se puede utilizar con cualquier parte móvil del cuerpo.
- **Simplicidad.** No hay necesidad de cables, brazos de montaje, etc. Es rápido de instalar y permite empezar a trabajar de inmediato, lo cual

es crucial en un entorno donde el tiempo que se puede dedicar a cada usuario es escaso.

- **Compatibilidad.** Permite trabajar con cualquier aplicación externa que acepte como entrada eventos del ratón o pulsaciones de teclas.

Para más información véase el vídeo [84] donde se ilustra el funcionamiento de este sistema.

4.3. Resultados proyecto SITPLUS

4.3.1. Entorno de experimentación

Desde principios de 2008, se ha habilitado de forma permanente una sala (véase la Figura 4.1) en la Associació Provincial de Paràlisi Cerebral (APPC) dedicada al proyecto SITPLUS. Desde entonces se llevan a cabo sesiones experimentales con carácter regular habiéndose convertido en una actividad rutinaria del centro.



Figura 4.1: Laboratorio SITPLUS en la APPC.

El laboratorio está equipado con materiales de absorción acústica, un ordenador personal, un monitor de 42 pulgadas, un sistema de sonido con cuatro

altavoces, una cámara de vídeo, varios micrófonos inalámbricos y una cámara web.

El hecho disponer de esta sala permite trabajar de forma individual con los usuarios, reduciendo distracciones y centrando la atención en la actividad, al tiempo que evita molestias a otros usuarios. Además, es importante destacar que el uso de una pantalla de gran formato y un equipo de sonido potente y de calidad proporcionan una experiencia mucho más envolvente.

4.3.2. Procedimiento

Por lo general, hay tres participantes en una sesión experimental: un investigador (tomando anotaciones y ajustando los parámetros del sistema), un facilitador y el usuario. Cada sesión dura aproximadamente de 15 a 30 minutos.

Cuando el usuario llega, se reproduce un sonido característico de una campana y el facilitador explica al usuario (cuando sea necesario) lo que va a hacer allí. Esto ayuda al usuario a concentrarse en la tarea y recordar sesiones anteriores. A continuación, el facilitador sugiere una tarea para el usuario, ya sea jugar con la voz o con el instrumento virtual. La propuesta concreta y los ajustes del sistema dependen de las habilidades del usuario y de su tasa de éxito observada en sesiones anteriores. Por lo general, los usuarios comienzan con la actividad de instrumento virtual y luego se pasa a jugar con su voz.

Normalmente, los usuarios interaccionan desde su silla de ruedas, pero si el fisioterapeuta lo recomienda, se utiliza un bipedestador o una colchoneta. Dependiendo del usuario, el facilitador alienta a que éste interaccione sugiriéndole que realice un gesto físico (por ejemplo, mover la cabeza) o una emisión oral. Una vez que el usuario empieza a interaccionar por su cuenta, el facilitador trata de intervenir lo menos posible.

El facilitador observa detenidamente al usuario, trata de interpretar su comportamiento y vela para que todo transcurra con normalidad. A veces, el facilitador puede pedir al investigador que tome nota de algunas observaciones que realiza o que ajuste algunos parámetros del sistema (por ejemplo, cambiar el instrumento). La sesión finaliza cuando expira el período de tiempo asignado o cuando el usuario lo pide o muestra signos de cansancio. En ese momento se detiene el software, vuelve a sonar la campana y el facilitador explica al usuario que la sesión ha terminado.

Toda la sesión se graba en vídeo, incluyendo los comentarios finales del investigador, el facilitador y el usuario (cuando sea posible) para su posterior análisis. Algunas de las anotaciones tomadas se tienen en cuenta de forma inmediata para las futuras sesiones (*ej.* actividad favorita, los instrumentos preferidos, efectos digitales que han propiciado mayor grado de interacción,

etc) y para mejorar el software.

Inicialmente, se trató de analizar los vídeos siguiendo un procedimiento muy estricto. En primer lugar, los facilitadores y otros profesionales vinculados al usuario realizaban un primer visionado para tener una idea general de la situación y las reacciones de los usuarios. En segundo lugar, visionaban el vídeo dos o tres veces más de forma individual tomando notas de diferentes variables, como la expresividad, sonrisas, repertorio fonológico, número de emisiones orales, nivel de atención, el nivel de comprensión de la actividad, el grado de la satisfacción, la interacción voluntaria, el número de refuerzos necesarios, el grado de cansancio, postura, etc. En tercer lugar, los vídeos eran visionados de nuevo por, al menos, dos personas simultáneamente para contrastar diferentes puntos de vista y las evaluaciones subjetivas. Por último, los evaluadores debatían hasta alcanzar un acuerdo sobre la evaluación.

Este procedimiento requería la intervención de, al menos, 3 profesionales (es decir, 1 técnico y 2, o más, facilitadores o cuidadores) durante un período de tiempo de entre 2 y 4 horas (cada uno) para analizar una única muestra de vídeo de 25 minutos. Debido al hecho de que todos los evaluadores realizaban esta tarea voluntaria y que tienen que ocuparse de muchas otras tareas para las que están contratados, esta metodología de análisis acabó resultando impracticable.

Con el objetivo de ofrecer algunos datos cuantitativos, se ha relajado el procedimiento de evaluación descrito anteriormente y nos hemos centrado en el análisis de la participación de los usuarios teniendo en cuenta el porcentaje de tiempo durante el cual están interaccionando voluntariamente y el número de sonrisas durante las sesiones (esta simplificación reduce significativamente el tiempo requerido, pero el procedimiento sigue tomando mucho tiempo). Ambas variables (interacción voluntaria y el número de sonrisas) han sido evaluados por facilitadores que trabajan frecuentemente con los participantes. Las sonrisas se determinan en función de las expresiones faciales (concretamente de la boca). La interacción voluntaria se considera que ocurre cuando el usuario mira a la pantalla y realiza movimientos no estereotipados o emisiones orales al mismo tiempo.

El sistema que se propone tiene por objeto el fomento de la interacción de los usuarios a través del juego, por lo tanto, contar el número de sonrisas parece ser un método razonable para determinar si realmente están disfrutando de las sesiones.

4.3.3. Resultados cuantitativos

Los usuarios que hemos elegido para nuestro estudio tienen diferentes perfiles de discapacidad y de personalidad. Esto puede dar lugar a un conjunto

Usuario	Edad	Sexo	% tiempo interacción voluntaria	Sonrisas min.
A.M.	8	F	81,9 %	1,80
M.M.	10	M	100,0 %	9,70
H.O.	12	M	63,8 %	3,39
K.G.	5	M	94,1 %	3,71
S.F.	24	F	73,9 %	1,52
E.P.	8	F	41,9 %	0,73
A.F.	4	M	68,7 %	0,17
Media	10,14		74,90 %	3,00
Desviación estándar	6,69		19,58 %	3,22

Tabla 4.6: Resultados cuantitativos.

incomparable de resultados (hay que tener en cuenta que esta es una situación común debido a la naturaleza altamente idiosincrásica de estos usuarios). Hemos grabado vídeos de más de 200 sesiones de usuarios con parálisis cerebral interactuando con nuestro sistema, donde cada sesión tiene una duración de unos 25 minutos. En la actualidad, se han analizado cuantitativamente unos 25 vídeos de 7 usuarios diferentes, es decir, 12,5 % de los datos recogidos.

Como se puede observar en la Tabla 4.6, los resultados obtenidos son bastante homogéneos. El porcentaje promedio de tiempo durante el cual los usuarios interactúan voluntariamente es casi el 75 %, lo que es un valor muy elevado si tenemos en cuenta que las personas con parálisis cerebral suelen tener dificultades para mantener la atención en una determinada tarea. De la Tabla 4.6, se puede derivar que el usuario M.M. es un valor atípico, porque está interactuando voluntariamente todo el tiempo, por lo tanto, no se puede considerar para calcular las estadísticas. Si este usuario se elimina de los análisis, se obtiene una notable 70,7 % de la media de tiempo de interacción voluntaria.

El número de sonrisas por minuto también es un resultado muy positivo. En promedio, los usuarios sonríen 3 veces por minuto durante las sesiones. Este resultado parece indicar que los usuarios realmente disfrutaban de la interacción con nuestro sistema.

Con el fin de aclarar la evidencia estadística de la relación entre la interacción voluntaria y el número de sonrisas, hemos calculado el coeficiente de correlación de Pearson entre el porcentaje de tiempo en que los usuarios interactúan voluntariamente y el número de sonrisas por minuto de una sesión.

Resulta que el coeficiente de correlación de Pearson entre las dos varia-

bles estudiadas es de 0,69, lo que indica una correlación positiva entre ellos. Esto podría indicar que los participantes que interaccionaron voluntariamente también sonrieron más, lo que puede indicar que están disfrutando de la actividad que están haciendo. Sin embargo, reconocemos que este resultado podría también ser influenciada por las diferencias individuales. Para mostrar los resultados gráficamente hemos trazado la gráfica de la Figura 4.2 junto con la línea que resulta de una regresión lineal.

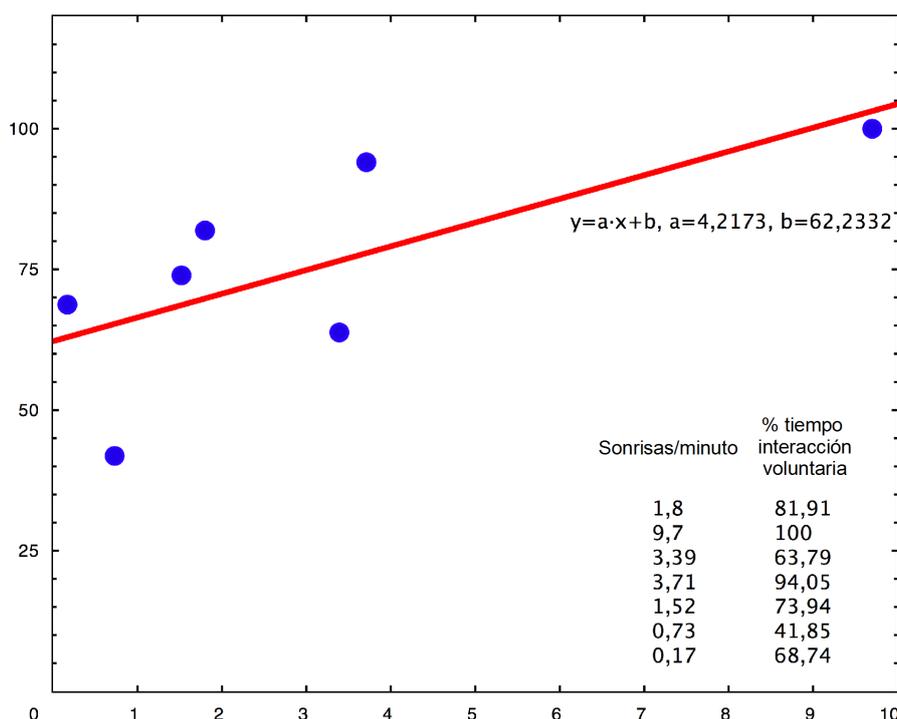


Figura 4.2: Recta de regresión y datos.

4.3.4. Resultados cualitativos

Debido al número limitado de usuarios y recursos, los resultados cuantitativos que hemos obtenido no son exhaustivos. Con el objetivo de ofrecer una visión más completa de los resultados de nuestro estudio, hemos incluido algunos análisis cualitativos del uso de nuestro sistema por parte de distintos usuarios de la APPC durante tres cursos (de 2010 a 2013). Estos resultados se basan en nuestra propia observación completada con el punto de vista de los facilitadores que han participado en las sesiones.

Además de alumnos de la escuela de educación especial, también han participado usuarios del taller ocupacional y de la residencia de la APPC. En total, más de 44 personas con parálisis cerebral.

De todo este grupo, poco más de la mitad se ha descartado por los siguientes motivos:

- **El usuario no interacciona con el sistema (8 casos).** Si tras varias sesiones no se obtiene evidencia de que el usuario interacciona de forma consciente con el sistema, se descarta. Como en algunos casos esto no resulta aparente, para evaluar si la interacción es consciente se han realizado pruebas como desconectar la salida gráfica y sonora del sistema, volverla a conectar al cabo de unos segundos y observar los cambios en el rostro y los gestos del usuario. El perfil de los usuarios que pertenecen a este grupo corresponde a personas con afectaciones cognitivas severas y profundas y edades mentales que se estiman por debajo de los 6 meses.
- **El usuario tiene tendencia a sobreexcitarse, acrecentar sus estereotipias o tiene problemas de conducta (5 casos).** Hay usuarios que ante ciertas situaciones, como participar en nuestra actividad, tienden a sobreexcitarse lo que en general no suele ser deseable. Otros, con rasgos del trastorno del espectro autista, tienden a realizar estereotipias (movimientos repetitivos). En algunos casos, como se verá, nuestro sistema ha servido para reducir dichas estereotipias, pero en otros el sistema tiende a acrecentarlas (lo que también se puede interpretar como un rechazo hacia la actividad). Por último, hay personas con problemas de conducta (*ej.* agresividad), lo que desaconseja su participación en la actividad.
- **El usuario tiene un estado de salud muy delicado (3 casos).** Algunas personas que acuden a la APPC tienen un estado de salud muy precario, lo que implica que requieren atenciones más específicas (*ej.* administración de oxígeno), causan baja frecuentemente y prefieren descansar en lugar de participar en ciertas actividades. Además de los 3 casos contabilizados, la baja participación de A.V. (véase la Tabla 4.7) durante el último curso se debe a un deterioro importante de su estado de salud que acabó desembocando en su fallecimiento.
- **El usuario tiene buena destreza manipulativa y/o buen nivel cognitivo (3 casos).** Desde un principio el sistema se ha concebido para aquellas personas con deficiencias sensoriales y, sobre todo, cognitivas y físicas. Por ello se ha tendido a descartar a aquellas personas con parálisis cerebral que presentan un nivel cognitivo normal (o prácticamente

normal) o aquellas que, a pesar de tener más limitaciones cognitivas, no presentan deficiencias (o éstas son muy leves) en las extremidades superiores. Se tomó esta decisión para dar preferencia a otros usuarios que tienen su participación más limitada en otras actividades.

- **El usuario expresa que la actividad no le gusta (3 casos).** Como es natural, hay algunos usuarios a los cuales no les gusta la actividad y así lo expresan. Generalmente es algo que resulta obvio al cabo de pocas sesiones, aunque en el caso de E.P., que tampoco se ha contabilizado en el recuento (véase la Tabla 4.7) manifestó que la actividad no le gustaba tras varios años de haber participado en ella. Los motivos para este cambio de actitud no están claros y podrían obedecer a la edad (dejó de participar justo al inicio de la pubertad), al cambio de facilitadores y tutores de aula, a cambios de medicación, etc.
- **Otros motivos (6 casos).** Entre los cuales el hecho de que ya utilizan el sistema en otros espacios (*ej.* en el aula), por decisión de padres o tutores, por previsión de abandonar la institución en un breve plazo o defunción.

Si bien resulta complejo establecer unos criterios generales de inclusión y exclusión, atendiendo a lo anterior, podemos extraer como conclusiones que SITPLUS no resulta funcional para personas con deficiencias cognitivas profundas para las cuales no es posible observar una interacción activa con el sistema. En el otro extremo, para los usuarios que tienen una capacidad cognitiva conservada, hemos tendido a excluirlos de la actividad en favor de aquellos con deficiencias cognitivas moderadas dado que disponen de recursos para participar en otras muchas actividades. En el caso de que el usuario tenga una buena capacidad manipulativa en las EESS, también se tiende a excluirlo de la actividad. Otro criterio de exclusión sería el estado de salud si éste es precario. Por último, si tras unas sesiones de prueba el usuario muestra patrones de comportamiento disruptivos o rechaza la actividad, también sería descartado. En cualquier caso, es importante poder realizar algunas sesiones de prueba antes de excluir definitivamente a un usuario de la actividad.

Quitando los casos antes mencionados, la Tabla 4.7 muestra a los que han participado de forma regular durante estos tres cursos.

Son un total de 15 participantes distintos, 7 hombres y 8 mujeres, de edades comprendidas entre los 4 y los 41 años. A lo largo de estos tres cursos se han realizado un total de 256 sesiones. La duración de las mismas oscila entre los 10 y 30 minutos, con un promedio aproximado de 20 minutos por sesión. Se puede observar un descenso del número de sesiones a lo largo del tiempo que se debe a cambios organizativos internos de la APPC ajenos al proyecto.

Usuario	Sexo	Año nacimiento	Servicio	Nº	Nº	Nº
				sesiones	sesiones	sesiones
				10-11	11-12	12-13
A.P.	M	1987	Taller	11	5	4
A.V.	M	1988	Taller	5	5	1
B.V.	M	2005	Escuela	8	6	6
E.C.	F	2000	Escuela	8	5	4
E.P.	F	2000	Escuela	9	6	3
I.O.	F	2006	Escuela	9	6	5
K.G.	M	2003	Escuela	10	9	4
L.C.	F	1977	Residencia	7	5	-
L.R.	F	2006	Escuela	10	9	1
P.O.	F	2006	Escuela	7	8	4
S.E.	M	1987	Residencia	6	4	-
S.F.	F	1984	Taller	6	4	5
S.P.	F	1982	Taller	4	3	1
X.S.	M	1971	Residencia	10	5	-
Y.S.	M	1977	Residencia	8	4	4
Total:				126	84	46

Tabla 4.7: Participantes SITPLUS durante tres cursos académicos.

El trabajo realizado durante estas sesiones se enmarca dentro de un programa de intervención más amplio donde se persiguen los siguientes objetivos generales:

- Participar activamente en la actividad.
- Ser consciente de la relación causa-efecto. Para más información véase la Sección 3.1.1.1.
- Focalizar la atención en la tarea y mantenerla durante el máximo tiempo posible.
- Establecer vínculos fluidos de relación y confianza con los profesionales.
- Enriquecer y diversificar los propios recursos expresivos.
- Explorar utilizando los sentidos del oído y la vista, y asimilar las informaciones recibidas.
- Fomentar la comunicación.
- Potenciar la toma de decisiones.

- Disfrutar de la actividad.

A continuación se resumen los resultados cualitativos de algunos usuarios que hemos considerado más significativos:

- **A.P.** es un joven con parálisis cerebral que acude al taller ocupacional, presenta una afectación cognitiva moderada, restricciones severas en cuanto a su capacidad manipulativa con las EESS y expresa claramente cuando algo le gusta o no a través de comportamientos disruptivos.

Cuando se le va a buscar al taller, sabe perfectamente a dónde va y lo recibe con gran sonrisa. En ese momento se le recuerdan las normas básicas para ir: comportarse, sentarse bien y controlar su sialorrea². Lo acata y su motivación es tan grande que lo cumple completamente. Durante toda la sesión se le retira el babero (con la excusa de que aparece en pantalla) y apenas se mancha de saliva.

Una vez en la sala, sabe perfectamente lo que tiene que hacer y lo hace encantado. Su participación es muy activa y se le ve disfrutar y desinhibirse mucho. La relación con los distintos facilitadores es muy fluida y se hace entender perfectamente. Es por ello que, en las pocas ocasiones que ha mostrado disgusto o cansancio, se ha respetado su decisión y la sesión se ha terminado en ese momento.

Con la actividad “Instrumento musical virtual” interacciona realizando movimientos arribar y abajo con su mano izquierda. Para fomentar su capacidad de decisión, se ha preparado un panel en papel (véase la Figura 4.3) a través del cual puede escoger el instrumento deseado. Durante las sesiones hace un uso intensivo de éste solicitando entre 8 y 15 cambios de instrumento. Le gusta explorar los sonidos que éstos producen.

Hay que decir A.P. presenta ciertos rasgos narcisistas por lo que se ha configurado que la salida visual del sistema sea su propia imagen o un collage de fotos propias. Hemos observado que así centra mucho más la atención que si usamos imágenes abstractas.

Con la actividad “Jugando con la voz” le cuesta emitir sonidos (no tiene habla) pero con el tiempo se ha ido mostrando más participativo y ha empezado a disfrutar con los efectos y las producciones que él mismo realiza.

En general, valoramos muy positivamente la participación de A.P. y consideramos que ha sido (y está siendo) una valiosa experiencia para él, habiéndose alcanzado todos los objetivos antes enumerados, además

²Babeo, excesiva producción de saliva.



Figura 4.3: Panel para escoger instrumento musical (en catalán).

de otros como el control del babeo y la adopción de una postura correcta a la hora de sentarse.

- **A.V.** era un joven con parálisis cerebral muy afectado físicamente y con una afectación cognitiva bastante importante.

Reconocía la actividad desde el momento en que se le anticipaba y se mostraba contento de participar.

Interaccionaba desde su silla, ya que le proporcionaba el posicionamiento más cómodo para él y, dada su limitada capacidad física, solía hacerlo a través de ligeros movimientos de rotación de la cabeza y pequeños movimientos de sus manos (que el sistema era capaz de recoger a pesar de su reducida amplitud).

La actividad “Jugando con la voz” le suponía mayor esfuerzo pero le gustaba, se sorprendía y reía escuchando y explorando los sonidos que emitía.

Si bien era capaz de entender la relación causa-efecto, le costaba focalizar su atención, lo que se traducía en una baja participación activa. Cuando se le pedía que interaccionara con el sistema, lo hacía y acababa respondiendo con una sonrisa. Cuando se le preguntaba si la actividad le

gustaba, respondía afirmativamente, aunque probablemente para complacer al facilitador. No obstante, con el tiempo se hizo cada vez más aparente que la actividad no era completamente de su agrado lo que, unido al deterioro de su estado de salud, hizo que le acabáramos descartando como participante.

- **B.V.** es un niño con parálisis cerebral tipo encefalopatía epiléptica y presenta un retardo mental moderado.

Cuando se le va a buscar al aula parece saber a dónde va y nos recibe con alegría. Se ha adaptado rápidamente a la sala, a los profesionales y a la actividad. Una vez en la sala, su participación es muy activa y sabe y entiende lo que debe hacer. Se le ve disfrutar y desinhibirse.

Las sesiones las dividimos en dos partes: primero B.V. hace un trabajo con la actividad “Instrumento musical virtual” y después con la actividad “Jugando con la voz”.

Con el instrumento virtual suele trabajar unos 15 minutos a través del movimiento de sus manos. A B.V. le gustan los instrumentos de percusión y nos lo indica con su patrón de actividad, ya que es más activo con instrumentos de este tipo. Es totalmente consciente de que es él quien realiza la actividad. Si apagamos el sistema se enoja y cuando lo volvemos a activar se alegra. Cuando le ponemos otros instrumentos que no son de percusión continua participando en la actividad, pero se cansa antes. Se aprovecha su motivación para trabajar el movimiento controlado de las EESS y la relación causa-efecto (*ej.* si no se mueve, no suena la música). A través del software también se intenta reconducir su conducta disruptiva de llevarse las manos a la boca (si se mete las manos en la boca, la música no suena) y lo está entendiendo perfectamente.

En la actividad “Jugando con la voz” se le coloca un micrófono de solapa y con varios efectos modificamos su voz para intentar que se muestre atento a los cambios producidos. Esta actividad le resulta más difícil. Tiene tendencia a ponerse el micrófono en la boca y no realiza muchas producciones vocales. Normalmente esta actividad dura unos 5 minutos y no disfruta tanto como con la anterior.

En una ocasión, después de una sesión en la que había participado y disfrutado, al volver al aula se puso a gritar. Desde el aula lo volvieron a llevar a la sala, su enojo cesó y continuó interaccionando con nuestro sistema y riendo.

En resumen B.V. participa activamente y disfruta de la actividad, entiende lo que debe hacer, comprende la relación causa-efecto y prefiere la

actividad “Instrumento musical virtual” a “Jugando con la voz”, siendo los instrumentos de percusión sus preferidos. En todos los casos y hasta tres profesionales distintos, han valorado positivamente continuar con la actividad.

- **E.C.** es una niña con parálisis cerebral secundaria con extirpación de astrocitoma cerebral. A diferencia de la mayoría de usuarios de la APPC, su lesión cerebral le sobrevino a la edad de 7 años por causa de un tumor.

Cuando se la va a buscar al aula y se le anticipa a dónde va lo recibe con alegría. A la pregunta de si quiere ir, contesta con una gran sonrisa. Al principio del curso 2010-11 y durante las primeras sesiones, E.C. estaba más seria y retraída, como si estuviera expectante e insegura. Pero a lo largo del tiempo, podemos afirmar que la progresión ha sido positiva. En la actualidad su participación en la sala es muy activa y se la ve disfrutar y desinhibirse mucho.

Lo que más le gusta, aunque le produce cansancio, es mover las extremidades superiores y la cabeza para tocar el instrumento virtual. Sobre todo se le hace trabajar el brazo izquierdo, aunque a menudo le pedimos que mueva ambos. Normalmente esta actividad la combinamos con la de “Jugando con la voz” para evitar que se extenúe y para utilizarlo como refuerzo positivo. Puede permanecer 15 minutos seguidos tocando el instrumento. Con la actividad “Jugando con la voz” se la ve emitir sonidos y explorar con curiosidad, deteniéndose a escuchar y recreándose en los resultados sonoros. Hay que motivarla un poco al principio, pero una vez ya se ha animado, no hay forma de pararla. Hemos observado que el nivel de atención ha ido aumentando a lo largo del tiempo.

Se aprovechan las sesiones para trabajar la comunicación y la toma de decisiones. Concretamente, con el panel de instrumentos (véase la Figura 4.3) ella puede escoger el que desea con la mirada. Su preferencia es el piano, probablemente porque solía tocarlo antes de su lesión cerebral, pero a menudo también elige el violín o la trompeta.

A lo largo de estos tres cursos, tres facilitadores distintos han valorado positivamente su participación en la actividad y han recomendado su continuidad.

- **E.P.** es una niña con parálisis cerebral de tipo tetraparesia espástica. Su afectación física y cognitiva es moderada.

Reconoce la sala, los profesionales y la actividad. Su participación durante el primer curso era fluctuante: había días que estaba muy activa y otros cansada. En cambio, durante el segundo curso su participación

fue mucho más estable. Sabe y entiende lo que debe hacer. Cuando está activa, se la ve disfrutar mucho y estar atenta a todo. La relación con el adulto es muy fluida y se hace entender si está disfrutando o no.

Se aprovecha el día que está activa para trabajar el movimiento controlado de la cabeza, de las extremidades superiores y la relación causa-efecto (si no sube la cabeza o no mueve los brazos, no suena la música) y lo hace bien.

Habitualmente interacciona desde su silla de ruedas, aunque también hemos probado que lo haga tumbada sobre una colchoneta. En esos casos hemos observado que tiene más libertad de movimiento y disfruta más.

Con la actividad “Jugando con la voz”, le cuesta emitir sonidos, pero cuando lo hace, es como un gran descubrimiento: se queda parada escuchándose a sí misma y, posteriormente, esboza una sonrisa.

Durante los dos primeros cursos se valoró positivamente su participación en la actividad y se recomendó su continuidad. Pero durante el último sólo se llevaron a cabo tres sesiones de trabajo durante las cuales mostró no estar a gusto. Tras estas tres sesiones los profesionales decidieron interrumpir su participación en la actividad dejando pendiente la posibilidad de continuarla en el aula.

En resumen, E.P. durante los dos primeros años participaba activamente, disfrutaba de la actividad y la anticipaba con alegría. Entendía lo que había que hacer y comprendía la relación causa-efecto. Al tercer curso dejó de participar en la actividad puesto que expresaba estar a disgusto. Se desconocen los motivos de este cambio de actitud.

- **S.E.** es un joven que vive en la residencia de la APPC. Tiene parálisis cerebral con afectaciones físicas y cognitivas en grado moderado.

S.E. tiene un temperamento nervioso y tiende a realizar estereotipias consistentes en llevarse las manos a la boca y lamérselas. Dado que esta conducta acababa por producirle lesiones en las manos, el equipo de la residencia optó por colocarle unas férulas que le cubren ambas manos y parte de los antebrazos. Con las férulas sigue manteniendo esta conducta pero sus manos no resultan dañadas.

Teniendo en cuenta lo anterior, se marcó como uno de los principales objetivos de trabajo el reducir las estereotipias. Para ello se ha empleado la actividad “Instrumento musical virtual” situando el área de seguimiento de movimiento por encima de los hombros del sujeto de tal forma que deba levantar sus brazos para activar el instrumento.

En general podemos decir que se relaja mucho durante las sesiones con el sonido que él mismo produce. Normalmente llega nervioso y lamiéndose las férulas de las manos. A lo largo de la sesión se relaja y se le ve sonriente y expectante hacia los sonidos, lo que le conduce a reducir el número de estereotipias. Con el tiempo ha ido entendiendo la relación causa-efecto y su participación ha sido más sostenida.

- **K.G.** es un niño con parálisis cerebral con afectación física moderada y cognitiva muy leve. Es muy alegre y colaborativo, tiene un buen nivel de atención e interés por aprender cosas nuevas y participa en muchas actividades.

Cuando se le va a buscar al aula, sabe perfectamente a dónde va y lo recibe con gran emoción. Conoce el espacio y la ubicación y sabe llegar solo desde su clase (bajo supervisión del adulto). Las sesiones se han realizado desde su silla o tumbado en la colchoneta. Puntualmente también se han realizado desde un bipedestador.

Una vez en la sala rápidamente pide iniciar actividad, sabe perfectamente lo que tiene que hacer y lo hace encantado. Su participación es muy activa y se le ve disfrutar y desinhibirse mucho. La relación con el adulto es muy fluida y se hace entender perfectamente. Es por ello que, en las pocas ocasiones que ha mostrado disgusto o cansancio, se ha respetado su decisión y la sesión ha terminado en ese preciso momento.

Respecto a la actividad “Instrumento musical virtual”, participa de forma activa y disfruta mucho, se hace entender perfectamente y pide diferentes instrumentos, parece que prefiere la percusión, aunque le gusta ir cambiando. Principalmente se trabaja el movimiento de las EESS y la relación causa-efecto, que comprende perfectamente.

Con la actividad “Jugando con la voz” es cuando más disfruta. Le gusta explorar con su voz, emitir sonidos y escucharse. Mantiene la atención durante mucho tiempo, ríe mucho y se nota que está animado y disfrutando.

Cabe destacar que inicialmente K.G. tenía problemas para esperar su turno cuando estaba a punto de participar en algunas de las actividades que le gustan, como la interacción con nuestro sistema. Por ello, su maestra decidió aprovechar esa motivación para trabajar la espera de turno. Al principio, cuando preveía que la actividad estaba a punto de comenzar, empezaba a llorar. Lo mismo ocurría cuando la sesión estaba a punto de terminar. Después de algunas sesiones, dejó de llorar al finalizar la sesión, y ahora, incluso devuelve el micrófono al facilitador cuando lo pide, como una señal de aceptación de la finalización de la sesión.

4.3. Resultados proyecto SITPLUS

123

A pesar de que encajaría en el perfil de usuario que tiene buen nivel cognitivo, se optó por que continuara participando en la actividad dada la elevada motivación que le produce. Durante estos tres cursos, los facilitadores han valorado positivamente su participación y han recomendado su continuidad.

CAPÍTULO 5

Conclusiones

Este capítulo resume las conclusiones y contribuciones de la presente tesis y describe posibles líneas de trabajo futuro que surgen a partir de posibles mejoras de las propuestas detalladas y aquellos objetivos alcanzados parcialmente.

Índice

5.1. Conclusiones	125
5.2. Contribuciones	127
5.3. Trabajo futuro	128

5.1. Conclusiones

En esta tesis se ha presentado una parte del trabajo realizado a lo largo de casi una década en relación a sistemas interactivos para personas con parálisis cerebral, con un énfasis en aquellas que presentan deficiencias en grados moderados y severos, grupo para el cual existen menos recursos y herramientas.

Todo el trabajo de campo se ha llevado a cabo gracias a la colaboración Associació Provincial de Paràlisi Cerebral (APPC) de Tarragona, entidad colaboradora necesaria para esta investigación. Cabe decir que la APPC viene realizando, desde hace una década, una apuesta decidida por la introducción de las tecnologías de la información y la comunicación en el día a día del centro y de la vida de las personas con parálisis cerebral. Esto incluye también el fomento de las labores de investigación y desarrollo como lo ha demostrado financiando gran parte de la tesis aquí presentada.

Tesis que, además de resultados científicos, también persigue resultados prácticos que emanan del conocimiento cercano de las condiciones en las que viven estas personas y los recursos que tienen a su disposición. Por ello se ha puesto un especial énfasis en que producir herramientas, de software libre y que no requieren hardware especializado, que se ha procurado que se apliquen de una forma sistemática en un contexto real y que ha trascendido los muros de la propia institución donde se inició el proyecto.

Como se ha comentado, el abordaje de esta tesis ha supuesto retos importantes, más si se tiene en cuenta que se ha hecho desde la perspectiva de la ingeniería dentro de la disciplina de la interacción persona-ordenador. Por ello, ha resultado fundamental la colaboración de los profesionales que trabajan día a día con las personas con parálisis cerebral, ya que, cuando la interlocución directa con los usuarios finales no es posible (o muy compleja), son ellos quienes tienen un conocimiento directo de estas personas, nos permiten interpretar sus respuestas e indicarnos cuando una experiencia resulta relevante para la persona. Esta visión de los profesionales no se puede obviar en un trabajo como este, a pesar del nivel de subjetividad inherente. Asimismo, los profesionales también son prescriptores y, en parte, usuarios de los sistemas aquí propuestos, por lo que su colaboración también ha sido fundamental para el diseño y desarrollo de los mismos.

El planteamiento de la tesis ha tenido siempre presente la especificidad de las características individuales de los participantes. Si bien es cierto que se ha intentado generalizar los resultados, incluso establecer comparaciones entre distintos individuos, también lo es que resulta imprescindible aportar los detalles de casos concretos como también se ha hecho.

De entre las propuestas y resultados presentados, la que refiere a la evaluación de la calidad de la interacción, fue uno de los primeros trabajos que se abordó. A partir de nuestra experiencia afirmamos que dichas métricas, si bien útiles, son sólo aplicables en usuarios con limitaciones cognitivas leves. Ello se debe a que para realizar las pruebas son necesarios ciertos niveles de comprensión y atención de los que carecen las personas más afectadas. En tales casos se desvirtúan completamente los resultados debido a la elevada variabilidad de los mismos. También se ha visto que, de entre los indicadores, no sólo habría que fijarse en los tiempos para completar las tareas propuestas, también en la estabilidad de los mismos (*ej.* a través de la desviación estándar) ya que los resultados muestran indicios que están relacionados con la calidad del acceso. Otro factor clave a tener en cuenta es el aprendizaje ya que de éste depende la destreza que tenga el usuario con un determinado sistema de acceso. A su vez, la optimización del acceso es algo que sólo tiene sentido abordarse una vez el usuario ha adoptado un sistema en concreto y ha adquirido cierta destreza. Por último y no por ello menos importante, la opinión que expresa el usuario debería ser tenida en cuenta en primera instancia. Cabe decir también que esta línea de trabajo quedó relegada a un segundo plano hasta que hace relativamente poco tiempo han concurrido dos nuevos candidatos para los cuales tenía sentido llevarlas a cabo.

En relación al proyecto Switch Viacam y aun disponiendo únicamente de resultados preliminares, es fácil ver que, sin sustituir al conmutador clásico, permite desarrollar de una manera práctica y efectiva otros escenarios de uso,

por ejemplo, como herramienta destinada a la rehabilitación física de personas con parálisis cerebral.

En relación al proyecto SITPLUS, este ha supuesto el foco principal de nuestros esfuerzos, teniendo en cuenta que era la propuesta con mayor número de beneficiarios potenciales en la APPC y otras entidades similares. Inicialmente se enfocó hacia aquellas personas con trastornos cognitivos moderados y severos. Con el tiempo hemos visto que aquellas personas con trastornos más profundos no son capaces de interactuar con nuestro sistema de forma consciente (ni tampoco con otros sistemas de causa-efecto). Por otra parte, hemos constatado que algunas personas con afectaciones cognitivas más leves también acaban disfrutando de la actividad. Éste constituye el amplio abanico de posibles candidatos a usar nuestro sistema, dentro del cual habría que eliminar a aquellas personas que presentan conductas y respuestas no deseadas, aquellas que tienen un estado de salud delicado y a las que expresan que no están a gusto con la actividad. Con todo, SITPLUS se ha convertido en una actividad que se desarrolla regularmente en la APPC y en otras entidades similares.

Se han presentado resultados cuantitativos y cualitativos en relación a la aplicación de nuestro sistema en personas con parálisis cerebral. Los primeros se han basado en el análisis de parte de las grabaciones en vídeo realizadas durante las sesiones experimentales y muestran indicios claros de que los usuarios que participaron en el estudio focalizan su atención y disfrutan con nuestra actividad. De hecho, se ha encontrado una correlación entre la participación y el número de veces que el usuario sonríe.

Los resultados cuantitativos se han complementado con nuestras propias observaciones cualitativas y las de los profesionales a lo largo de tres cursos académicos durante los cuales nuestro sistema se ha utilizado como parte del plan de intervención en usuarios de la APPC. A partir de dichas observaciones afirmamos que nuestro sistema resulta útil para trabajar de forma individual distintos objetivos como el refuerzo de la relación causa-efecto, la atención, la relación usuario-profesional, el trabajo de los propios recursos expresivos, la comunicación y la toma de decisiones. Adicionalmente y para ciertos casos particulares, la motivación del usuario por participar en la actividad, también ha resultado útil de cara a corregir (por lo menos durante el transcurso de las sesiones), ciertos comportamientos (*ej.* control de la sialorrea y respeto del turno) y estereotipias (*ej.* llevarse las manos a la boca).

5.2. Contribuciones

A lo largo de esta tesis se han realizado las siguientes contribuciones:

- En el Capítulo 2, una recopilación de sistemas interactivos audiovisuales orientados a personas con discapacidad, y parálisis cerebral en concreto, para uso expresivo, lúdico y terapéutico.
- Técnicas para la evaluación de la calidad de la interacción para personas con parálisis cerebral, una parte de las cuales ha mostrado ser de utilidad.
- La aplicación sistemática y análisis de resultados del proyecto SITPLUS en personas con parálisis cerebral durante varios años.
- El uso de ciertos métodos y algoritmos de visión artificial con finalidades interactivas.
- Los propios sistemas Switch Viacam y SITPLUS, desarrollados como parte del trabajo de esta tesis y que se han puesto a disposición de la comunidad.

5.3. Trabajo futuro

Esta tesis recoge una serie de esfuerzos y resultados que se han realizado hasta la fecha, no obstante nuestro compromiso con la línea de trabajo iniciada continua y por ello esperamos poder aplicar nuestros esfuerzos en los siguientes frentes:

- Ampliar los resultados cuantitativos de los experimentos con el proyecto SITPLUS. El enfoque adoptado para el análisis de las sesiones experimentales, basado en el visionado y anotación de vídeos, requiere de un considerable esfuerzo. Para ampliar los resultados resulta imprescindible acelerar el análisis de las sesiones. Para hacerlo plantemos la introducción de sistemas que permitan la anotación de observaciones en tiempo real, es decir, durante las sesiones (*ej.* como el presentado en la Sección 3.5), la introducción de sistema automáticos de recogida de datos, tanto en nuestro propio sistema como externamente (*ej.* sistemas que permitan el análisis automático de la expresión facial del usuario y sistemas que permitan la recogida de otras variables fisiológicas y físicas) y la adopción de herramientas más potentes que permitan la integración de todos los datos y el análisis de los mismos (*ej.* recientemente tuvimos conocimiento de las herramientas VCode and VData de la Universidad de Illinois [48]).

En el momento de escribir estas líneas, ya se ha iniciado un trabajo donde se está evaluando el sensor Afectiva Q Sensor¹ en distintas personas con parálisis cerebral de la APPC. Este sensor, en forma de reloj de muñeca, permite registrar de forma continuada la conductividad y temperatura de la piel, así como la aceleración (incorpora un acelerómetro de 3 dimensiones). Una de las situaciones en las que se plantea usar este sensor es durante las sesiones experimentales con el proyecto SITPLUS.

- Mejorar la descripción del perfil de los participantes en los experimentos. Teniendo en cuenta la heterogeneidad de las personas con parálisis cerebral y a fin de adoptar las mejores estrategias posibles de interacción, debemos conocer las características individuales de los usuarios con cierta precisión. También, cuando se documentan estudios con este tipo de personas, debería aportarse un perfil detallado para cada caso de estudio.

En el futuro planeamos utilizar la Clasificación del Funcionamiento, de la Discapacidad y de la Salud (CIF). La CIF pertenece a la familia de clasificaciones internacionales desarrolladas por la Organización Mundial de la Salud (OMS) y permite describir el funcionamiento y la discapacidad de la persona asociados con las condiciones de salud. Desde su publicación en 2001, la CIF y ha sido considerada como referencia en las nuevas iniciativas a la atención y promoción a personas con discapacidad.

Preveamos que la adopción de este sistema de descripción resultará útil para:

1. Documentar de forma más precisa el perfil de los participantes en futuros experimentos y facilitar la generalización de resultados.
2. Permitir que el sistema proponga tareas específicas o ajuste automáticamente algunos parámetros según el perfil de usuario.

En el momento de escribir estas líneas, ya está en marcha un proyecto promovido por ASPACE y en el cual participará la APPC para adoptar internamente la CIF como sistema de descripción, por lo que una vez implantado ya se dispondrá de facto de estos perfiles detallados de los participantes.

- Aunque actualmente algunas de las sesiones con nuestro proyecto SITPLUS ya se están llevando a cabo con sólo dos personas, usuario y facilitador, tenemos previsto añadir una extensión del sistema que permita su manejo de forma remota (*ej.* utilizando un Tablet PC) de forma que

¹<http://www.qsensortech.com/resources/>

permita al facilitador realizar los ajustes sin tener que estar físicamente frente al ordenador y también permita a ciertos usuarios realizar por sí mismos tareas como la selección de instrumentos.

- Aunque como se ha dicho, el interés de esta investigación y del proyecto SITPLUS se enfoca a las personas con parálisis cerebral, no se descarta que también pueda beneficiar a otros colectivos con discapacidad cognitiva (daño cerebral, autismo, síndrome de Down, retardo mental, etc.).

Bibliografía

- [1] Ahonen-Eerikäinen, Heidi, Andrea Lamont y Roger Knox: *Rehabilitation for Children with Cerebral Palsy: Seeing Through the Looking Glass: Enhancing Participation and Restoring Self-Image through the Virtual Music Instrument*. International Journal of Psychosocial Rehabilitation, 12(2):41–66, 2008.
- [2] Alcantud, Francisco, Gerardo Herrera, Gabriel Labajo, I. Dolz, C. Gaya, V. Avila, A. Blanquer, J. L. Cuesta y J. Arnaiz: *Assessing Virtual Reality as a Tool for Support Imagination*. En *Computers Helping People with Special Needs*, volumen 2398 de *Lecture Notes in Computer Science*, páginas 185–191. Springer, Berlin, 2002.
- [3] Andjelkovic, Michael: *Ejes Wii Remote*. http://commons.wikimedia.org/wiki/File:Wii_Remote_accelerometer_axes.svg, 2009. Último acceso: 05/05/2014.
- [4] Arneson, Carrie L, Maureen S Durkin, Ruth E Benedict, Russell S Kirby, Marshalyn Yeargin-Allsopp, Kim Van Naarden Braun y Nancy S Doernberg: *Prevalence of cerebral palsy: Autism and developmental disabilities monitoring network, three sites, United States, 2004*. Disability and health journal, 2(1):45–48, 2009.
- [5] Azeredo, Maria: *Real-time composition of image and sound in the (re)habilitation of children with special needs: a case study of a child with cerebral palsy*. Digital Creativity, 18(2):115–120, 2007.
- [6] Bettmann, Corbis: *Leon Theremin tocando el Theremin*. http://en.wikipedia.org/wiki/File:Lev_Termien_playing_-_cropped.jpg, 1927. Último acceso: 05/05/2014.
- [7] Bhasin, Tanya Karapurkar, Sally Brocksen, Rachel Nonkin Avchen y Kim Van Naarden Braun: *Prevalence of four developmental disabilities among children aged 8 years: Metropolitan Atlanta Developmental Disabilities Surveillance Program, 1996 and 2000*, volumen 55, páginas 1–9. US Department of Health and Human Services, Centers for Disease Control and Prevention, 2006.
- [8] Bobick, Aaron F. y James Davis: *The recognition of human movement using temporal templates*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(3):257–267, 2001.

- [9] Bradski, Gary R. y James Davis: *Motion segmentation and pose recognition with motion history gradients*. Machine Vision and Applications, 13(3):174–184, 2002.
- [10] Brooks, Anthony: *Virtual interactive space(V.I.S.) as a movement capture interface tool giving multimedia feedback for treatment and analysis*. En *Int Congr World Confed Phys Ther*, página 289, Japón, 1999.
- [11] Brooks, Anthony: *Body Electric and Reality Feedback loops: Virtual Interactive Space & Entertainment Potentials*. En *International Conference on Artificial Reality and Telexistence*, páginas 93–98, Japón, 2004. KAIST / VRSJ.
- [12] Brooks, Anthony: *CAREHERE: Creating Aesthetically Resonant Environments for the Handicapped, Elderly and Rehabilitation*. En *Intl Conf. Disability, Virtual Reality & Assoc. Tech*, páginas 191–198, Oxford, Reino Unido, 2004.
- [13] Brooks, Anthony: *Enhanced Gesture Capture in Virtual Interactive Space*. En *Computers in Art and Design Education Conference*, Suecia, 2004. Denmark and Malmö University.
- [14] Brooks, Anthony: *HUMANICS 1: a feasibility study to create a home internet based telehealth product to supplement acquired brain injury therapy*. En *Intl Conf. Disability, Virtual Reality & Assoc. Tech*, Oxford, Reino Unido, 2004. University of Reading.
- [15] Brooks, Anthony: *Interactive Painting: an evolving study to facilitate reduced exclusion from classical music concerts for the deaf community*. En *Intl Conf. Disability, Virtual Reality & Assoc. Tech.*, páginas 101–108, Reino Unido, 2004.
- [16] Brooks, Anthony: *Robotic Synchronized to Human Gesture as a Virtual Coach in (Re)habilitation Therapy*. En *International Workshop on Virtual Rehabilitation*, Lausanne, Suiza, 2004. VRlab, EPFL.
- [17] Brooks, Anthony: *Soundscapes*. En *Inhabited Information Spaces: Living with your Data*, páginas 89–99. Springer, Nueva York, 2004.
- [18] Brooks, Anthony, Antonio Camurri, Nishan Canagarajah y Stefan Hasselblad: *Interaction with shapes and sounds as a therapy for special needs and rehabilitation*. En *4th Int. Conference On Disability, Virtual Reality, and Associated Technologies*, páginas 205–212, Veszprém, Hungría, 2002. P. Sharkey, C. Sik Lányi, P. Standen.

- [19] Brooks, Anthony y Eva Petersson: *Humanics 2: Human-computer interaction in acquired brain injury*. En *International Conference on Human-Computer Interaction*, volumen 8 de *Universal access in HCI: Exploring new dimensions of diversity*, Las Vegas, EEUU, 2005. Lawrence Erlbaum Associates.
- [20] Brooks, Anthony y Eva Petersson: *Play Therapy Utilizing the Sony Eye-Toy*. En *Eighth International Workshop on Presence*, páginas 305–314, Londres, Reino Unido, 2005. Slater, M.
- [21] Brooks, Anthony y Eva Petersson: *Raw emotional signalling, via expressive behaviour*. En *International Conference on Artificial Reality and Telexistence*, volumen 157 de *ACM International Conference Proceeding Series*, páginas 133–141, Christchurch, Nueva Zelanda, 2005. ACM.
- [22] Brooks, Anthony y Eva Petersson: *Recursive reflection and learning in raw data video analysis of interactive 'play' environments for special needs health care*. En *Enterprise networking and Computing in Healthcare Industry, 2005*, páginas 83–87, 2005.
- [23] Brooks, Anthony y Eva Petersson: *SoundScapes: non-formal learning potentials from interactive VEs*. En *SIGGRAPH '07: ACM SIGGRAPH 2007 educators program*, página 18, San Diego, California, EEUU, 2007. ACM.
- [24] Brooks, Anthony y Eva Petersson: *Stillness design attributes in non-formal rehabilitation*. En *Computers in Art Design and Education*, páginas 36–44, Perth, Australia, 2007. Curtin University of Technology.
- [25] Brumback, Christine, Liubo Borrisov, Jeffrey Galusha y Annette DiIorio: *IntelliVision*. En *The Sixth International Conference Ubiquitous Computing*, Nottingham, Reino Unido, 2004. University of Nottingham.
- [26] Burr, B.: *VACA: a tool for qualitative video analysis*. En *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, 2006.
- [27] Cadoz, Claude: *Continuum énergétique du gest au son - simulation multisensorielle interactive d'objets physiques*. En *Interfaces homme-machine et création musicale. Proceedings of a conference organized by IRCAM and Ina-GRMm*. Hermes, París, 1999.

- [28] Camurri, Antonio, Barbara Mazzarino, Gualtiero Volpe, Pietro Morasso, Federica Priano y Cristina Re: *Application of multimedia techniques in the physical rehabilitation of Parkinson's patients*. The Journal of Visualization and Computer Animation, 14:269–278, 2003.
- [29] Chadabe, Joel: *Some Reflections on the Nature of the Landscape within which Computer Music Systems are Designed*. Computer Music Journal, 1(3):5–11, 1977.
- [30] Chadabe, Joel: *Interactive Composing: An Overview*. Computer Music Journal, 8(1):22–28, 1984.
- [31] Chadabe, Joel: *Electric Sound: The Past and Promise of Electronic Music*. Prentice Hall, Nueva Jersey, EEUU, 1997.
- [32] Cobb, Sue y P.M. Sharkey: *A decade of research and development in disability, virtual reality and associated technologies: review of ICDVRAT 1996-2006*. International Journal of Virtual Reality, 6(2):51–68, 2007.
- [33] Csikszentmihalyi, Mihaly: *Flow: The Psychology of Optimal Experience*. Harper Perennial, 1991.
- [34] Douglas, Sarah A., Arthur E. Kirkpatrick y I. Scott MacKenzie: *Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard*. En *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, páginas 215–222, Nueva York, EEUU, 1999. ACM Press.
- [35] Dunn, L.M., E.R. Padilla, D.E. Lugo y L.M. Dunn: *Test De Vocabulario en Imagenes Peabody*. American Guidance Service (AGS), Inc., Circle Pines, MN, EEUU, 1986.
- [36] Ellis, Phil: *Incidental Music; a case study in the development of sound therapy*. The British Journal of Music Education, 12:59–70, 1995.
- [37] Ellis, Phil: *Layered Analysis: A Video-based Qualitative Research Tool to Support the Development of a New Approach for Children with Special Needs*. Bulletin of the Council for Research in Music Education, 130, 1996.
- [38] Ellis, Phil: *Improving Quality of Life and Well-Being for Children and the Elderly through Vibroacoustic Sound Therapy*. En *Computers Helping People with Special Needs*, volumen 3118 de *Lecture Notes in Computer Science*, páginas 416–422. Springer-Verlag, 2004.

- [39] Ellis, Phil: *Vibroacoustic Sound Therapy: Case Studies with Children with Profound and Multiple Learning Difficulties and the Elderly in Long-Term Residential Care*. Studies in Health Technology and Informatics, 103:36–42, 2004.
- [40] Ellis, Phil: *The Development of Interactive Multisensory Environments for Expression, 1992 - 2007*. Keynote, Luxembourg Society for Music Therapy, 2006.
- [41] Ellis, Phil, Joanne McElligott, Aoife Power y Lieselotte Van Leeuwen: *Touching Sound: Suggestions for Activities to support perceptual learning and expression*. Informe técnico, University of Sunderland & Creative Partnerships, 2007.
- [42] English, William K., Douglas C. Engelbart y Melvyn L. Berman: *Display selection techniques for text manipulation*. IEEE Transactions on Human Factors in Electronics, HFE-8:5–15, 1967.
- [43] F. L. Van-Nes, J. F. Juola, R. J. A. M. Moonen: *Attraction and Distraction by Text Colours on Displays*. En Hans-Jorg Bullinger, Brian Shackel (editor): *INTERACT 87 - 2nd IFIP International Conference on Human-Computer Interaction*, páginas 625–630. North-Holland, 1987.
- [44] GAME-ABLING, Consorcio: *Proyecto GAME-ABLING*. <http://game-abling.eu/>, 2012. Último acceso: 05/05/2013.
- [45] Gesturetek: *Gesturetek Health | Sensory Solutions*. <http://www.gesturetekhealth.com/products-sensory.php>, 2013. Último acceso: 13/8/2013.
- [46] Gil, Jose A., Mariano Alcaniz, Javier Montesa, Moises Ferrer, Javier Chirivella, Enrique Noe, Carolina Colomer y Joan Ferri: *Low-cost Virtual Motor Rehabilitation System for Standing Exercises*. En *Virtual Rehabilitation, 2007*, páginas 34–38, Venecia, Italia, 2007.
- [47] Granollers, Toni: *MPIu+a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares*. Tesis de Doctorado, Universitat de Lleida, Lleida, 2004.
- [48] Hagedorn, J., J. Hailpern y K.G. Karahalios: *VCode and VData: illustrating a new framework for supporting the video annotation workflow*. En *Working Conference on Advanced Visual interfaces*, páginas 317–321. ACM, 2008.

- [49] Hailpern, Joshua, Karrie Karahalios, James Halle, Laura Dethorne y Mary Kelsey Coletto: *A3: HCI Coding Guideline for Research Using Video Annotation to Assess Behavior of Nonverbal Subjects with Computer-Based Intervention*. ACM Transactions on Accessible Computing, 2(1), 2009.
- [50] Hasselblad, Stefan, Eva Petersson y Anthony Brooks: *Empowered interaction through creativity*. Digital Creativity, 18(2):89–98, 2007.
- [51] Horn, B.K.P. y B.G. Schunck: *Determining optical flow*. Artificial Intelligence, 17:185–203, 1981.
- [52] Hunt, A. y R. Kirk: *Mapping Strategies for Musical Performance*. En Wanderley, Marcello y M. Battier (editores): *Trends in Gestural Control of Music*. IRCAM - Centre Pompidou, París, 2000.
- [53] i3net: *CARESS. Creating Aesthetically REsonant EnvironmentS in Sound*. <http://www.bristol.ac.uk/caress/>, 2000. Último acceso: 05/05/2014.
- [54] i3net: *Tw-aysi. The world is - as you see it*. <http://www.bristol.ac.uk/Twi-aysi/>, 2001. Último acceso: 05/05/2014.
- [55] i3net: *CAREHERE. Creating Aesthetically Resonant Environments for the Handicapped, Elderly and Rehabilitation*. <http://www.bristol.ac.uk/carehere/>, 2003. Último acceso: 05/05/2014.
- [56] Inman, Dean P. y Ken Loge: *Teaching motorized wheelchair operation in virtual reality*. En *1995 CSUN Virtual Reality Conference*, Northridge, 1995. California State University.
- [57] Internacional, Estándar: *Ergonomics of human-system interaction—Guidance on accessibility for human-computer interface*. ISO/TS 16071, 2003.
- [58] ISO, WD: *9241-11. Ergonomic requirements for office work with visual display terminals (VDTs)*. The international organization for standardization, 1998.
- [59] IT, Noldus: *The observer and Pocket Observer*. <http://www.noldus.com>, 2011. Último acceso: 1/2/2011.

- [60] Jannink, Michiel J. A., Gelske J. van der Wilden, Dorine W. Navis, Gerben Visser, Jeanine Gussinklo y Maarten Ijzerman: *A Low-Cost Video Game Applied for Training of Upper Extremity Function in Children with Cerebral Palsy: A Pilot Study*. *CyberPsychology & Behavior*, 11(1):27–32, 2008.
- [61] Johnson, Ann: *Prevalence and characteristics of children with cerebral palsy in Europe*. *Developmental Medicine & Child Neurology*, 44(09):633–640, 2002.
- [62] Jordà Puig, Sergi: *Digital Lutherie: Crafting musical computers for new music's performance and improvisation*. Tesis de Doctorado, Universitat Pompeu Fabra, Barcelona, 2005.
- [63] Keates, Simeon, Faustina Hwang, Patrick Langdon, P. John Clarkson y Peter Robinson: *Cursor measures for motion-impaired computer users*. En *Assets '02: Proceedings of the fifth international ACM conference on Assistive technologies*, páginas 135–142, Nueva York, EEUU, 2002. ACM Press.
- [64] Kipp, Michael: *Anvil - A Generic Annotation Tool for Multimodal Dialogue*. En *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, páginas 1367–1370, 2001.
- [65] Knox, Roger, Andrea Lamont, Tom Chau, Yani Hamdani, Heidi Schwellnus, Ceilidh Eaton, Cynthia Tam y Patricia Johnson: *Movement-to-music: Designing and implementing a Virtual Music Instrument for young people with physical disabilities*, 2003.
- [66] Koester, H.H., E. LoPresti, G. Ashlock, W. McMillan, P. Moore y R. Simpson: *Compass: Software for Computer Skills Assessment*. En *CSUN 2003 International Conference on Technology and Persons with Disabilities*, Los Angeles, CA, March 2003.
- [67] Koester, H.H. y E.F. LoPresti: *Tools and Methods for Assessment in Computer Access*. En *RESNA 2006 Annual Conference*, Atlanta, GA. Arlington, VA, 2006. RESNA Press.
- [68] Krueger, Myron W.: *Environmental technology: making the real world virtual*. *Commun. ACM*, 36(7):36–37, 1993.
- [69] L OU, Debbie: *Conmutador mecánico*. http://en.wikipedia.org/wiki/File:Access_device.JPG, 2008. Último acceso: 05/05/2014.

- [70] Lab, InfoMus: *EyesWeb*. <http://www.infomus.org/EywMain.html>, Último acceso: 30/07/2008.
- [71] Lamont, Andrea, Roger Knox, Tom Chau, Yani Hamdani, Heidi Schwellnus, Cynthia Tam y Patricia Johnson: *Converting Movements to Music New Musical Exploration Opportunities for Children in Rehabilitation*. En *Canadian Association for Music Therapy 29 th Annual Conference*, páginas 26–30, Regina, Saskatchewan, Canadá, 2002.
- [72] Lee, M. y D. Wessel: *Connectionist models for real-time control of synthesis and compositional algorithms*. En *1992 Int. Computer Music Conference*, páginas 277–280, San Francisco, CA, 1992. International Computer Music Association.
- [73] Leshner, Gregory W., Jeffery D. Higginbotham y Bryan J. Moulton: *Techniques for Automatically Updating Scanning Delays*. En *Proceedings of the RESNA 2000*, páginas 85–87. Resna Press, VA, 2000.
- [74] Lessenberry, D.D.: *Analysis of Errors*. Syracuse, NY: L.C. Smith and Corona Typewriters., 1928.
- [75] Lienhart, Rainer y Jochen Maydt: *An Extended Set of Haar-like Features for Rapid Object Detection*. En *IEEE ICIP*, páginas 900–903, 2002.
- [76] Liffick, B., B. Romich y K. Hill: *A single-switch performance evaluation tool*. En *11th International Conference on Human-Computer Interaction.*, Las Vegas, NV., 2005.
- [77] MacKenzie, I. S.: *Fitts' law as a research and design tool in human-computer interaction*. En *Human-Computer Interaction*, 7, páginas 91–139, 1992.
- [78] Mackenzie, Scott I., Tatu Kauppinen y Miika Silfverberg: *Accuracy measures for evaluating computer pointing devices*. En *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, páginas 9–16, Nueva York, EEUU, 2001. ACM Press.
- [79] Martínez Segura, María José: *Tecnologías de la Información y la Comunicación y estimulación sensoriomotriz en niños con plurideficiencias*. En *III CONGRESO NACIONAL DE TECNOLOGÍA, EDUCACIÓN Y DIVERSIDAD*, Murcia, 923–25 2005.

- [80] Martínez Segura, María José y Francisco Alberto García Sánchez: *Influencia de las TIC en la ejecución motriz de niños con grave discapacidad*. En *I Congreso Internacional de Psicopedagogía: Ámbitos de intervención del psicopedagogo*, 53–45 2006.
- [81] Mathews, Max, Joan E. Miller, F. Richard Moore, John R. Pierce y J.C. Risset: *The Technology of Computer Music*. The MIT Press, Cambridge, MA, EEUU, 1969.
- [82] Mathews, Max y F. Richard Moore: *GROOVE, a Program for Realtime Control of a Sound Synthesizer by a Computer*. En *4th Annual Conference of the American Society of University Composers*, páginas 22–31. Columbia University, 1969.
- [83] Mathews, Max y F. Richard Moore: *Groove - A Program to Compose, Store, and Edit Functions of Time*. Communications of the ACM, 13(12):715–721, 1970.
- [84] Mauri, C.: *Vídeo de Switch Viacam*. <http://youtu.be/u5J5WIo5Hyc>, 2012. Último acceso: 23/10/2012.
- [85] Mauri, Cesar: *Sitio web del proyecto Enable Viacam (eViacam)*. <http://viacam.org>. Último acceso: 05/05/2014.
- [86] Mauri, Cesar: *Sitio web del proyecto Switch Viacam*. <http://sviacam.sf.net>. Último acceso: 05/05/2012.
- [87] Mauri, Cesar: *Sistema de control en temps real del punter del ratolí mitjançant visió artificial*. Tesis de Licenciatura, URV, Tarragona, 2002.
- [88] Mauri, Cesar, Toni Granollers, Jesus Lores y Mabel García: *Computer Vision Interaction for People with Severe Movement Restrictions*. Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, 2(1):38–54, 2006.
- [89] M.D., Pasquale Accardo, Pasquale Accardo M.D., Larry Desch M.D., Louis Pellegrino M.D. y Nancy Roizen M.D. (editores): *Capute & Accardo's Neurodevelopmental Disabilities in Infancy and Childhood: Neurodevelopmental Diagnosis and Treatment*, página 17. Paul H Brookes Pub Co, 3ª edición, 2007.
- [90] Montoya, Rafael Sánchez: *Ordenador y Discapacidad*. CEPE, S.L., 2ª edición, 2002.

- [91] Neill, James: *Qualitative versus Quantitative Research: Key points in a classic debate*. <http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html>, 2007. Último acceso: 23/03/2011.
- [92] Noirhomme-Fraiture, Monique, Clairette Charriere y Jean Vanderdonck: *A Laboratory of Ergonomic Analyses for Children Suffering from Cerebral Palsy*. En *Tools for Working with Guidelines.*, páginas 35–49. TFWWG, 2000.
- [93] Noirhomme-Fraiture, Monique, Clairette Charriere, Jean M. Vanderdonck y Claudy Bernard: *ERGOLAB: a screen usability evaluation tool for children with cerebral palsy*. En *CHI '93: INTERACT '93 and CHI '93 conference companion on Human factors in computing systems*, páginas 23–24, Nueva York, EEUU, 1993. ACM Press.
- [94] North, Max M., Sarah M. North y Joseph R. Coble: *Virtual Environment Psychotherapy: A Case Study of Fear of Flying Disorder*. *Presence: Teleoperators and Virtual Environments*, 6(1):127–132, 1997.
- [95] North, Max M., Sarah M. North y Joseph R. Coble: *Virtual Reality Therapy: An Effective Treatment for Psychological Disorders*. En *Handbook of Virtual Environments: Design, Implementation, and Applications*, páginas 1065–1078. Lawrence Erlbaum Associates. Mahwah, NJ, 2002.
- [96] Oliveros, Pauline, Leaf Miller, Zevin Polzin y Zane Van Dusen: *Adaptive Use Musical Instruments: Software Musical Instrument*. <http://bsjeon.net/thisAbility/en/pauline.html>. Último acceso: 30/09/2008.
- [97] OpenCV, Equipo de: *OpenCV*. <http://opencv.org/>, 2013. Último acceso: 21/02/2013.
- [98] Paneth, Nigel, Ting Hong y Steven Korzeniewski: *The descriptive epidemiology of cerebral palsy*. *Clinics in perinatology*, 33(2):251–267, 2006.
- [99] Pares, Narcis, Anna Carreras, Jaume Durany, Jaume Ferrer, Pere Freixa, David Gomez, Orit Kruglanski, Roc Pares, Ignasi J. Ribas, Miquel Soler y Alex Sanjurjo: *Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment*. En *IDC '05: Proceedings of the 2005 conference on Interaction design and children*, páginas 110–116, Boulder, Colorado, 2005. ACM.

- [100] Pares, Narcis, Anna Carreras, Jaume Durany, Jaume Ferrer, Pere Freixa, David Gomez, Orit Kruglanski, Roc Pares, J. Ignasi Ribas, Miquel Soler y Alex Sanjurjo: *MEDIATE: An Interactive Multisensory Environment for Children with Severe Autism and No Verbal Communication*. En *Third Int'l Workshop Virtual Rehabilitation*, páginas 43–52, 2004.
- [101] Pares, Narcis, Paul Masri, Gerard van Wolferen y Chris Creed: *Fotos del Consorcio Europeo Proyecto MEDIATE; FP5/IST/Systems and Services for the Citizen/Persons with special needs (IST-2000-26307)*.
- [102] Pares, Narcis, Paul Masri, Gerard van Wolferen y Chris Creed: *Achieving Dialogue with Children with Severe Autism in an Adaptive Multisensory Interaction: The MEDIATE Project*. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):734–743, 2005.
- [103] Pengilly, Sylvia: *Integrating performance, live electronics, and interactive video*. *Computers & Mathematics with Applications*, 32(1):75–77, 1996.
- [104] Pertaub, David Paul, Mel Slater y Chris Barker: *An experiment on public speaking anxiety in response to three different types of virtual audience*. *Presence: Teleoperators and Virtual Environments*, 11(1):68–78, 2002.
- [105] Petersson, Eva y Anthony Brooks: *Non-formal therapy and learning potentials through human gesture synchronised to robotic gesture*. *Universal Access in the Information Society*, 6(2):167–177, 2007.
- [106] Portsmouth, University of: *MEDIATE: Multisensory Environment Design for an Interface. Final report*. Informe técnico, University of Portsmouth, 2004.
- [107] Project, Soundbeam: *Soundbeam for special needs and music therapy*. <http://www.soundbeam.co.uk>, 2013. Último acceso: 13/8/2013.
- [108] Rand, Debbie, Rachel Kizony y Patrice L. Weiss: *Virtual Reality Rehabilitation for All: Vivid GX versus Sony PlayStation II EyeToy*. En *5th International Conference on Disability, Virtual Reality and Associated Technologies*, páginas 87–94, Reino Unido, 2004. University of Reading.
- [109] Rich, R.: *Interactive Light Dimension Beam*. *Electronic Musician*, 12(7):154–159, 1996.
- [110] Rizzo, Albert, Jarrell Pair, Peter J. McNerney, Ernie Eastlund, Brian Manson, Jon Gratch, Randy Hill y Bill Swartout: *Development of a VR*

- Therapy Application for Iraq War Military Personnel with PTSD*. En *Medicine Meets Virtual Reality 13: The Magical Next Becomes the Medical Now*, volumen 111 de *Studies in Health Technology and Informatics*, páginas 407–413. IOS Press, 2005.
- [111] Rokeby, David: *Very Nervous System (VNS) 'Canadier omsætter video til musik'*. Ingeniøren, 1990.
- [112] Rosenblum, K.L., C. Zeanah, S. McDonough y M. Muzik.: *Video-taped coding of working model of the child interviews: a viable and useful alternative to verbatim transcripts?* *Infant Behavior and Development*, 27(4):544–549, 2004.
- [113] Rovan, J. B., Marcello Wanderley, S. Dubnov y Philippe Depalle: *Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance*. En *Kansei - The Technology of Emotion Workshop*, Génova, Italia, 1997.
- [114] Russell, Kathryn: *Imagining the Music, Exploring the Movement: Soundbeam in the Sunshine State*. *Queensland Journal of Music Education*, 4(1):41–48, 1996.
- [115] Sandlund, Marlene, Eva L Waterworth, Suzanne McDonough y Charlotte Häger-Ross: *Interactive Games in Motor Rehabilitation for Children with Sensorimotor Disorders*. En *Virtual Rehabilitation, 2007*, página 80, Venecia, Italia, 2007. IEEE.
- [116] Sik-Lanyi, Cecilia, V. Simon, L. Simon y V. Laky: *Using virtual public transport for treating phobias*. En *The Fifth International Conference on Disability, Virtual Reality and Associated Technologies*, Reino Unido, 2004. University of Reading.
- [117] Software, OATS: *SITPLUS*. <http://www.oatsoft.org/Software/sitplus>, 2013. Último acceso: 27/3/2013.
- [118] Spiegel, Laurie: *A Short History of Intelligent Instruments*. *Computer Music Journal*, 11(3):7–9, 1987.
- [119] Standen, P.J. y J. J. Cromby: *Can students with developmental disability use virtual reality to learn skills which will transfer to the real world?* En *1995 CSUN Virtual Reality Conference*, Northridge, 1995. California State University.

- [120] Strickland, Dorothy: *A Virtual Reality Application with Autistic Children*. Presence: Teleoperators and Virtual Environments, 5(3):319–329, 1996.
- [121] Strickland, Dorothy: *Virtual Reality for the Treatment of Autism*. En *Virtual Reality in Neuro-Psycho-Physiology*, páginas 81–86. IOS Press, Amsterdam, 1997.
- [122] Swingler, Tim: *'That Was Me!': Applications of the Soundbeam MIDI controller as a key to creative communication, learning, independence and joy.*, 1998.
- [123] Theremin, Leon S.: *The Design of a Musical Instrument Based on Cathode Relays*. Leonardo Music Journal, 6:49–50, 1996.
- [124] Timmermans, Hans, Gerard van Wolferen, Paul Newland y Simon Kunath: *MEDIATE: Key sonic developments in an interactive installation for children with autism*. En *2004 International Computer Music Conference*, páginas 201–204. ICMA, 2004.
- [125] Van Leeuwen, Lieselotte y Phil Ellis: *Facilitating the experience of agency through an intersensory interactive environment*. Digital Creativity, 18(2):99–105, 2007.
- [126] Viola, Paul y Michael J. Jones: *Rapid Object Detection using a Boosted Cascade of Simple Features*. En *IEEE CVPR*, 2001.
- [127] Vygotsky, L. S.: *The Genesis of Higher Mental Functions*. En Wertsch, J. V. (editor): *The Concept of Activity in Soviet Psychology*, páginas 144–188. Sharpe, New York, 1981.
- [128] Wanderley, Marcello y Philippe Depalle: *Contrôle gestuel de la synthèse sonore*. En *Interfaces homme-machine et création musicale. Proceedings of a conference organized by IRCAM and Ina-GRMm.*, páginas 145–163. HERMES Science Publications, París, 1999.
- [129] Wanderley, Marcello, N. Schnell y J. B. Rovan: *Escher - Modeling and Performing Composed Instruments in Real-Time*. En *1998 IEEE International Conference on Systems, Man and Cybernetics*, páginas 1080–1084, 1998.
- [130] Weiss, Patrice L., Yuval Naveh y Noomi Katz: *Design and testing of a virtual environment to train CVA patients with unilateral spatial neglect to cross a street safely*. Occupational Therapy International, 10:39–55, 2003.

- [131] Weiss, Patrice L., Debbie Rand, Noomi Katz y Rachel Kizony: *Video capture virtual reality as a flexible and effective rehabilitation tool*. Journal of NeuroEngineering and Rehabilitation, 1, 2004.
- [132] Winkler, Todd: *Creating Interactive Dance with the Very Nervous System*. En *Connecticut College Symposium on Arts and Technology*, páginas 212–217, New London, CT, 1997.
- [133] Woo, A., P. Rigby y R. Knox: *Applications of Soundbeam Music Technology with Children with Physical Disabilities*. Technical report, Bloorview MacMillan Centre, 1997.

