

La Usabilidad de las Interfaces Basadas en Visión

Cristina Manresa-Yee, Esperança Amengual y Pere Ponsa

Resumen

Las interfaces basadas en visión hacen uso de gestos para la comunicación del usuario con el sistema interactivo sin la necesidad de dispositivos que requieran contacto físico. El modelado de gestos suelen realizarse en laboratorios y es importante que se lleve a cabo la evaluación de su usabilidad, pero la gran diversidad de

usos y aplicaciones de estas interfaces hace que resulte difícil decidir qué factores tener en cuenta cuando se mide la usabilidad. En este artículo se presenta una revisión de la literatura cuyo objetivo es clasificar y recopilar los factores y las métricas de usabilidad que se utilizan para validar las interfaces basadas en visión.

Palabras clave

Interfaces basadas en visión, usabilidad, gestos, validación

1 Introducción

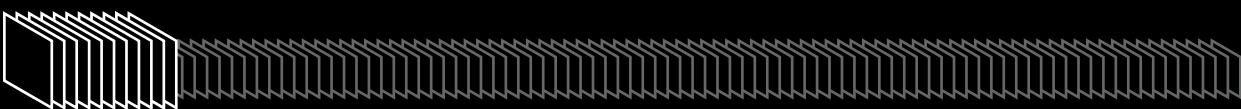
La Interacción Persona-Ordenador (IPO) es la disciplina que estudia el intercambio de información entre las personas y los ordenadores. Se encarga del diseño, la evaluación y la implementación de sistemas informáticos interactivos, y del estudio de los fenómenos más importantes con los que está relacionada (Hewett et al., 1992).

Cuando a un ordenador se le capacita con el sentido de la visión a través de una cámara, éste es capaz de adquirir información de sus usuarios y del entorno en el que opera (Porta, 2002). Esta información puede ser utilizada para comunicar al usuario con el ordenador, dando lugar a lo que se denomina una interfaz basada en visión (VBI – Vision-Based Interface en inglés).

A través de la visión por computador, los sistemas pueden percibir y entender a sus usuarios, así como las acciones que éstos llevan a cabo dentro de un contexto interactivo. En la actualidad este tipo de interfaces está ganando adeptos por su capacidad de identificar y reconocer los gestos y movimientos de los usuarios en tiempo real. Estos gestos pueden ser utilizados como mecanismo de entrada para interactuar con el ordenador: desde su uso dentro de un quirófano para no tener contacto físico con el ordenador (Graetzel et al., 2003) hasta su uso para navegar en mundos de realidad virtual (Tollmar et al., 2004).

Un gesto puede definirse como una configuración estática de una parte del cuerpo, como por ejemplo el gesto de señalar, o puede ser un conjunto de movimientos del cuerpo, o de partes del cuerpo, como el gesto de asentir con la cabeza. La utilización de los gestos propios del usuario para interactuar con un sistema de comunicación puede resultar natural, directo e intuitivo (Cabral et al., 2005). Sin embargo, algunos autores también critican la naturalidad y la facilidad de intuición de estas interfaces (Norman, 2010, Bowman et al., 2012) y su usabilidad (Norman y Nielsen, 2010).

Independientemente de las aplicaciones concretas que las VBI pueden ofrecer, este tipo de interfaces normalmente incluye tres fases: la detección, la identificación y el seguimiento (Porta, 2002). La detección determina la presencia o la ausencia de un elemento, la identificación busca el reconocimiento de un elemento en particular y el seguimiento tiene como objetivo seguir un elemento identificado durante un periodo de tiempo. Por ejemplo, cuando una VBI actúa de dispositivo de entrada (ej. el seguimiento de los ojos o de partes del cuerpo para reemplazar



un ratón (Manresa-Yee et al., 2008)), la interfaz detecta, reconoce y realiza el seguimiento de la parte del cuerpo que se utiliza para interactuar con el ordenador. Del mismo modo, existen otros sistemas interactivos que utilizan gestos como mecanismos de entrada para interactuar con el ordenador en los que no es necesario informar sobre la posición de la parte del cuerpo involucrada en la interacción, sino que detectar y reconocer el gesto es suficiente (Manresa-Yee et al., 2005).

Uno de los retos más difíciles en el diseño de las VBI es la identificación o definición de los gestos más apropiados para una aplicación en particular. A menudo los gestos se definen en un entorno de laboratorio en el que se deberían llevar a cabo evaluaciones de usabilidad para medir el grado en que los usuarios pueden utilizar el sistema diseñado en un determinado contexto de uso y para conseguir unos objetivos determinados de manera efectiva, eficiente y satisfactoria (ISO 9241-11, 1998). Sin embargo, resulta difícil decidir los factores y las métricas que deberían considerarse en una evaluación de usabilidad.

Wickeroth et al. (2009) propusieron una escala de usabilidad de gestos GUS (GUS – Gesture Usability Scale), basada en la escala de usabilidad de un sistema SUS (SUS – System Usability Scale) (Brooke, 1996) y la completaron añadiendo cinco preguntas relacionadas con la interfaz gestual. Estos elementos pretenden medir la fiabilidad, el rendimiento y la conformidad que el usuario percibe en base a sus expectativas. Además, con el objetivo de definir un modelo cuantitativo que considere la usabilidad de una interfaz basada en gestos, Barclay et al. (2011) presentaron un modelo que consideraba otros factores como la precisión, la fatiga, la naturalidad y la duración del gesto. Asimismo, analizando la literatura, se han encontrado otros autores que tienen en cuenta factores no considerados por estos dos trabajos.

El propósito de este artículo es revisar las prácticas de evaluación de la usabilidad que se llevan a cabo actualmente para evaluar las interfaces basadas en visión. Para ello se ha realizado una revisión de la literatura cuyo objetivo es clasificar los factores y métricas que podrían ser utilizados a la hora de evaluar la usabilidad de las VBI.

La estructura de este artículo es la siguiente: en el apartado 2 se describen los factores que se consideran en las evaluaciones de las VBI, así como las métricas utilizadas para medir estos factores. Además, se presenta una clasificación de estos factores en base a los atributos expuestos en ISO

9241-11: efectividad, eficiencia y satisfacción de usuario. En el último apartado se resumen los hallazgos más importantes y se presentan las conclusiones de este artículo.

2 Factores de La Usabilidad

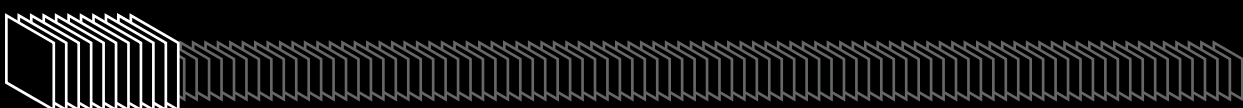
El término ‘usabilidad’ ha sido definido de manera diferente por distintos autores y diversos estándares (Abran et al., 2003). Con el objetivo de clasificar los factores de usabilidad de las VBI hallados en la literatura nos hemos basado en los tres atributos definidos en ISO 9241–11 Requisitos ergonómicos para trabajos de oficina con pantallas de visualización de datos (PVD). Parte 11: Guía sobre usabilidad:

- Efectividad se refiere al rendimiento de las tareas, a la precisión con la que el usuario cumple con sus objetivos.
- Eficiencia es la cantidad de esfuerzo necesario para conseguir el nivel de efectividad en el cumplimiento de los objetivos. La eficiencia es la relación entre el nivel de efectividad y el consumo de recursos.
- Satisfacción se refiere al confort del usuario en la utilización del sistema.

Como punto de partida de la búsqueda hemos utilizado las siguientes palabras clave y sinónimos: usability AND (“vision based interface” OR “vision based user interface” OR “gesture based interface” OR “gestural interface” OR “perceptual user interface” OR “natural user interface” OR “natural interface”)

En esta búsqueda de trabajos de investigación se excluyen aquellos trabajos que no utilizan técnicas de visión por computador o no han descrito la fase de evaluación. El motivo es que este trabajo se centra en compilar información sobre factores de evaluación de la usabilidad y métricas.

En el diseño de un gesto para una interfaz es importante considerar dos aspectos diferentes. En primer lugar, se deben tener en cuenta el contexto de la aplicación y los requisitos del usuario (Saffer, 2009). En segundo lugar, de acuerdo con Lenman et al. (2002), los gestos se caracterizan a partir de tres dimensiones diferentes: cognitiva, tecnológica y articularia:



- Los aspectos cognitivos están relacionados con la facilidad con la que se aprende a utilizar la interfaz.
- El aspecto tecnológico se refiere a la necesidad de tener en cuenta el estado del arte de la tecnología, tanto en el momento en el que se diseñe la VBI como en el futuro. Para explotar el uso de los gestos en el diseño de interfaces persona-ordenador es necesario proporcionar los mecanismos que harán posible que estas interfaces sean interpretadas por los ordenadores (Pavlovic et al., 1997).
- El conocimiento de la antropometría humana permite considerar aspectos articulatorios para que un gesto resulte cómodo y se eviten los gestos forzados (Nielsen et al., 2004).

Estas consideraciones han dado lugar a un conjunto de factores a tener en cuenta en la evaluación de la usabilidad de una VBI. A continuación se muestra una clasificación de los factores hallados en la literatura clasificados en base a los atributos de efectividad, eficiencia o satisfacción y las métricas usadas para medirlos. Llegados a este punto es preciso subrayar que algunos autores, en lugar de utilizar la definición proporcionada por la ISO 9241-11, en la que se basa este artículo, han utilizado otras definiciones del término usabilidad no contempladas en este trabajo.

2.1 Efectividad

La efectividad de una VBI se suele evaluar en base a la precisión y el ratio de errores, siendo ambas métricas cuantitativas.

2.1.1 Precisión y Ratio de Errores

La precisión es el grado de exactitud en el reconocimiento de los gestos realizados por el usuario. Este factor está relacionado con la robustez y la precisión de las técnicas de visión por ordenador. Se trata de un factor que suele utilizarse frecuentemente en las evaluaciones y es un indicador de la unicidad del gesto (Barclay et al., 2011), es decir, si un gesto es parecido a otro, el sistema puede malinterpretarlo y llevar a cabo una acción equivocada.

Los errores están relacionados con la precisión. Se pueden distinguir dos tipos de errores:

- Gestos malinterpretados dentro del conjunto de gestos posibles debido a la falta de unicidad.
- Gestos que no se entienden debido a la falta de robustez de las técnicas de visión por ordenador.

Un gesto que no se reconoce requiere repetición, pero un gesto que se malinterpreta ha de ser corregido. En este caso, ambos parámetros deberían ser evaluados individualmente.

Una métrica que se usa con frecuencia para evaluar la precisión es el número de gestos reconocidos correctamente en base al número de gestos realizados (Waldherr, 2000). Por ejemplo, en el trabajo de Manresa-Yee et al. (2005), se describe una evaluación de una interfaz basada en el movimiento de las manos para navegar en un mundo 3D. A partir de una secuencia de 40 gestos, 40 personas prueban la interfaz y se mide la precisión de cada gesto utilizando la ecuación Eq. 1.

$$p_{gesto} = \frac{n_{reconocidos}}{n_{total}}$$

(Eq.1)

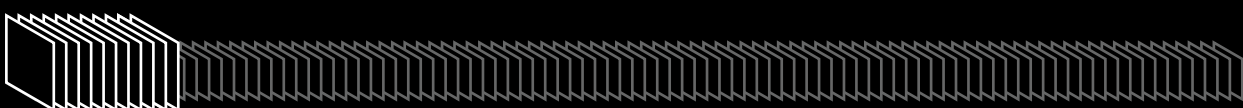
2.2 Eficiencia

En la validación de la eficiencia se utilizan diferentes métricas tales como el esfuerzo físico y mental del usuario, la duración del gesto, la facilidad con la que el gesto se aprende (learnability) o la facilidad con la que se recuerda (memorability).

2.2.1 Fatiga física

Cuando la interacción se realiza mediante movimientos corporales puede aparecer cansancio o fatiga, especialmente si el gesto requiere un esfuerzo físico. Las órdenes gestuales deben ser concisas, rápidas y, para minimizar el esfuerzo, se deben evitar los gestos que requieran mucha precisión por un largo período de tiempo (Baudel y Beaudouin-Lafon, 1993).

La fatiga es un atributo difícil de medir porque depende del usuario. Las métricas que se han ido utilizando para ello son normalmente cuestionarios que utilizan la escala de Likert o la escala Borg CR10 que es un método para valorar el esfuerzo percibido (Borg, 1982). Por ejemplo, en el trabajo



de De Silva et al. (2003) los autores realizan pruebas de usabilidad de un sistema controlado por el movimiento de la de nariz que reemplaza al ratón y, para evaluar el cansancio generado en el cuello, se utiliza la escala Borg CR10, donde 0 indicaba “sin esfuerzo” y 10 “mucho esfuerzo”. Otro ejemplo es el que se detalla en ISO 9241-9 (1998) que recomienda un cuestionario de esfuerzo que consta de doce preguntas con respuestas sobre la escala de Likert de 7 puntos y que permite evaluar los niveles de confort involucrados en el manejo de un sistema considerando la fatiga física en general y la fatiga en diferentes partes del cuerpo en particular, como por ejemplo sobre el brazo o el hombro.

Otra manera de evaluar la fatiga es mediante el control de las señales biológicas del usuario tales como el pulso o las señales electromiográficas (EMG). Para poder utilizar señales biológicas, se debe disponer del equipo necesario que permita realizar las mediciones. En el trabajo de Lukowicz et al. (2006) se presenta un sistema basado en sensores de fuerza resistivos colocados directamente sobre el cuerpo del usuario para medir la actividad muscular. Los autores del trabajo muestran una correlación entre la deformación del músculo (medido a través de los sensores) y la actividad muscular, especialmente la fatiga.

2.2.2 Duración

La duración de un gesto es el tiempo que el usuario necesita para realizar el gesto. Un gesto requiere una fase de preparación, una de ejecución y una de retracción (Kendon, 1988). La fase de ejecución requiere mantener el gesto durante un tiempo predeterminado para darle robustez al reconocimiento. La duración está estrechamente relacionada con las técnicas de visión por ordenador ya que es necesario configurar una duración mínima y una duración máxima para reconocer un gesto. Este factor afecta a la eficiencia del sistema y a la fatiga del usuario: cuanto más tiempo necesite el usuario para realizar el gesto, más probabilidad de que aumente la fatiga y disminuya el número de acciones de entrada al sistema que puedan ser realizados, por lo que puede disminuir el rendimiento.

La duración se evalúa midiendo el tiempo entre el inicio y el final de un gesto. El trabajo de Barclay et al. (2011) valida una interfaz gestual para controlar un centro multimedia y para evaluar la duración se calcula manualmente el tiempo entre el inicio y el final de un gesto. Del mismo modo, otros trabajos evalúan este factor calculando el tiempo necesario en realizar una tarea en lugar de medir la duración del gesto, especialmente cuan-

do se comparan interfaces diferentes. Por ejemplo, en el artículo de De Silva et al. (2003) se compara el tiempo necesario para realizar una tarea (la tarea de conexión multidireccional recomendada en ISO9241-9 (1998)) con el VBI que presentan y el ratón estándar.

2.2.3 Carga cognitiva

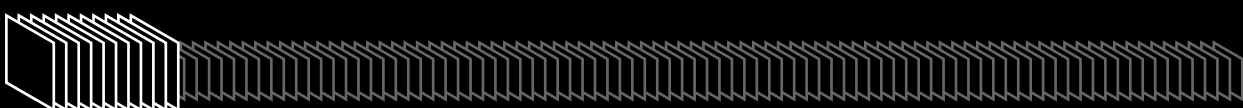
La carga cognitiva se refiere a la cantidad total de actividad mental impuesta sobre la memoria de trabajo. La evaluación de este atributo suele realizarse mediante evaluaciones subjetivas (cuestionarios como por ejemplo el NASA Task Load Index). En el trabajo de Bailly et al. (2012), los autores presentan un sistema incorporado en los zapatos que reconoce gestos discretos de la mano o gestos más amplios. Para medir la demanda mental, los participantes responden a las preguntas del cuestionario NASA-TLX (1988) sobre la interacción gestual.

Otro método para evaluar la carga cognitiva consiste en analizar la cantidad de información que una persona puede recordar mientras utiliza la interfaz. Esta última prueba se basa en la idea de que cuando se obliga a una persona a utilizar la memoria de trabajo u otros recursos cognitivos, la información se reemplaza o se pierde. El trabajo presentado en Sukeshini et al. (2011) describe una interfaz para la navegación visual de un modelo 3D del globo terrestre a través de gestos. La carga cognitiva se evalúa a través de un test post-tarea de memoria que sirve para determinar si los usuarios son capaces de recordar los símbolos que han visto, así como el orden y el lugar de aparición de los mismos, mientras navegaban.

2.2.4 Facilidad para aprender (Learnability) y Facilidad para recordar (Memorability)

La facilidad para aprender, o el tiempo de aprendizaje, es el tiempo y el esfuerzo requeridos para alcanzar un determinado nivel de rendimiento. La facilidad para recordar, o la retención en el tiempo, es la facilidad del sistema para ser utilizado correctamente por usuarios ocasionales (Nielsen et al., 2004). Cuanto más cercana al usuario sea la sintaxis de las operaciones, más fácil será recordar cómo utilizar la interfaz. Si el tiempo de aprendizaje es corto, la retención será menos importante (Shneiderman, 1998).

Para evaluar estos factores se utilizan comúnmente cuestionarios subjetivos. Por ejemplo, en el trabajo de Gonçalves et al. (2008) se utiliza un cuestionario con una escala Likert de 5 valores para evaluar la facilidad para



aprender. En el trabajo de Nielsen et al. (2004) la facilidad para recordar se mide utilizando una presentación donde van apareciendo cada 2 segundos el nombre de una función a realizar. Los participantes deben realizar el gesto asociado a la función de manera adecuada y correcta mientras ésta aparece. De no ser así, la presentación se reinicia. La métrica consiste en contar el número de veces que la presentación ha tenido que reiniciarse.

2.3 Satisfacción

Muchos de los factores clasificados bajo el atributo satisfacción se evalúan utilizando cuestionarios para identificar las percepciones y los sentimientos subjetivos de los usuarios acerca de la interfaz.

2.3.1 Naturalidad y Facilidad de intuición (Intuitiveness)

La naturalidad de los gestos se relaciona con la calidad de que éstos sean reales y que no requieran ningún elemento intermediario creado por personas. La facilidad de intuición es el uso instintivo de gestos basados en lo que uno siente que deben ser, incluso sin ningún tipo de razonamiento consciente. Ambos conceptos se suelen usar indistintamente, aunque la facilidad de intuición puede verse influenciada por experiencias previas (Sukeshini et al., 2011).

Los gestos naturales e intuitivos ayudan a que una interfaz sea fácil de descubrir por parte del usuario. Este es un aspecto recomendado (Saffer, 2009) que a su vez influencia a otros factores como la facilidad de aprender, la facilidad de recordar o la carga cognitiva.

Para evaluar este factor a menudo se utilizan cuestionarios subjetivos, aunque también se pueden encontrar algunas evaluaciones que valoran de manera cuantitativa la naturalidad, o algunos trabajos que se centran en descubrir los gestos más naturales.

En Gamberini et al. (2013) los autores evalúan cuantitativamente la naturalidad de determinados sistemas. La evaluación se basa en la identificación y el análisis de las interrupciones en las acciones del usuario durante su interacción con un sistema locomotor que el usuario controla con el movimiento de los pies con el objetivo de navegar en un entorno virtual mientras permanece sentado en una silla. Este trabajo se basa en los resultados de Winograd y Flores (1986) que consideran las interrupciones como una crisis en la interpretación de la situación en un determinado momento que hace que una persona suspenda sus acciones para encontrar una solución.

En el trabajo de Nielsen et al. (2004) se presentan dos métodos para identificar gestos que aseguren una correspondencia intuitiva y lógica con las funciones que representan: el método ascendente y el método descendente. El método ascendente consiste en descubrir los gestos a partir de las funciones, mientras que el método descendente parte de los gestos para identificar las funciones que les corresponden. Siguiendo el método ascendente, el trabajo de Höysniemi et al. (2004) analiza los movimientos que los niños prefieren y que resultan más intuitivos en diferentes contextos de juego (ej., nadar, correr, saltar) aplicando el método conocido como Mago de Oz.

2.3.2 Confort

El confort se define como un sentimiento placentero de estar relajado y libre de dolor.

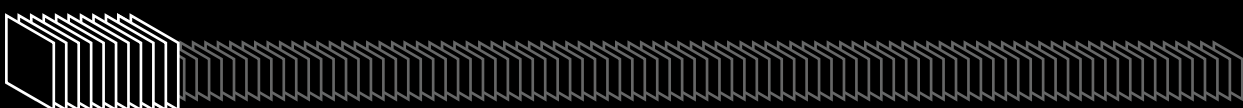
Para evaluar este factor a menudo se utilizan cuestionarios subjetivos (ej. con respuestas basadas en la escala Likert o el cuestionario basado en la escala Body Part Discomfort (BPD) (Corlett y Bishop, 1976), aunque también se pueden encontrar algunos trabajos orientados a la mejora de este factor mediante la identificación de zonas de confort. Por ejemplo, en Kölsch et al. (2003) se presenta un método para la evaluación subjetiva del confort postural en el que este factor se define como una postura que no requiere ningún tipo de movimiento de compensación de otras partes del cuerpo. En este trabajo sus autores analizan la postura del usuario para definir una zona de confort.

2.3.3 Facilidad de uso

El factor facilidad de uso significa que el usuario requiere poco esfuerzo general para utilizar el sistema.

Cuanto más fácil de usar sea la interfaz, más rápidamente se beneficiará el usuario. Para mejorar la facilidad de uso de una VBI se ha de tener en cuenta la experiencia previa del usuario y el diseño ha de ser familiar y consistente con sus expectativas (Lorés et al., 2001).

Este factor se puede valorar por medio de cuestionarios subjetivos. Por ejemplo el cuestionario SUS contiene una pregunta que hace referencia a la facilidad de uso del sistema (Wickeroth et al., 2009).



2.3.4 Experiencia del usuario y satisfacción

La experiencia del usuario incluye tanto aspectos pragmáticos (atributos eficiencia y efectividad) como aspectos hedonistas del sistema que se miden mediante indicadores subjetivos como la satisfacción de usuario y la calidad hedónica, que incluye factores como el divertimento o la estética. La calidad hedónica es el grado en que un sistema estimula al usuario por su carácter desafiante o nuevo o por la identificación de valores personales importantes para el usuario (Hassenzahl, 2004).

En base a esta definición, van Beurden et al. (2008) evalúan la experiencia de usuario mediante un cuestionario subjetivo con respuestas basadas en una escala semántica de 7 puntos (malo-bueno, fácil-difícil) para comparar los métodos de interacción basados en gestos con los métodos de interacción basados en dispositivos.

2.3.5 Aceptación social

En el trabajo de Montero et al. (2010) se revisan varias definiciones del factor de aceptación social propuestas en la literatura. Como resultado se propone una definición que se basa tanto en la manera en la que las personas se sienten al realizar la acción, como en la manera en que otras personas en el mismo entorno perciben las acciones del usuario:

- La aceptación social del usuario hace referencia a la impresión positiva o negativa de la tarea o de la tecnología desde la perspectiva del usuario.
- La aceptación social del espectador es una medida de las impresiones que provocan las acciones del usuario.

Desde estos dos puntos de vista, los gestos resultan socialmente aceptables si se consideran apropiados, tanto por el usuario, como por cualquier espectador, en el contexto en el que se llevan a cabo. También cabe la posibilidad de que el usuario considere la aceptación social de un gesto en base a como él mismo reaccionaría ante este mismo gesto si en lugar de usuario fuera espectador, creándose de este modo enlaces entre la aceptación social del usuario y la aceptación social del espectador.

Tal y como se subraya en el trabajo de Rico y Brewster (2010), puesto que el conjunto de gestos que pueden ser reconocidos de una manera fiable puede diferir bastante del conjunto que gestos que los usuarios desean adoptar, la aceptación social de un gesto debe ser evaluada antes de dedicar tiempo y esfuerzo en su implementación. Las personas validan la aceptación social cuando la motivación para hacer uso de la tecnología compite con las restric-

ciones del contexto social.

Existen muchos factores diferentes que podrían examinarse en relación a la aceptación social, tales como la cultura (UX Fellows, 2003), la discreción, la duración, el rendimiento del gesto, los rasgos de personalidad, la edad de los usuarios, el lugar y la audiencia. Con el objetivo de comprender como estos factores influyen en la aceptación social, ha habido diferentes iniciativas para evaluar como estos factores afectan a la voluntad del usuario a realizar determinados gestos (van Beurden et al., 2008; Rico y Brewster, 2010). Estas evaluaciones consisten en examinar un conjunto definido de gestos realizando entrevistas a diferentes grupos de participantes que deben responder unas preguntas después de ver un video de un gesto particular. Por ejemplo en el trabajo de Bragdon et al. (2011) se evalúa una interfaz para ser utilizada en reuniones considerando tacto y gestos y se valora la aceptabilidad social mediante cuestionarios subjetivos.

3 Conclusiones

Las VBI han suscitado gran interés, especialmente en contextos de ocio. El uso generalizado de este tipo de interfaces nos lleva a considerar su usabilidad como un atributo esencial que es necesario satisfacer para asegurar el éxito del gran número de aplicaciones que en la actualidad hacen uso de estas interfaces. A pesar de ello, debido a la novedad de estas interfaces, el diseño de su interacción es todavía inmaduro y aún no existen guías estándar disponibles. Es necesario proponer nuevas métricas para evaluar la usabilidad con el objetivo de mejorar el diseño y la implementación de las VBI.

En este trabajo hemos revisado las prácticas que actualmente se llevan a cabo para medir la usabilidad en las pruebas de evaluación de VBI. Con el objetivo de clasificar los diferentes factores que podrían ser considerados para evaluar la usabilidad de este tipo de interfaces, hemos revisado la literatura y presentado una taxonomía basada en los tres atributos de la usabilidad definidos en ISO 9241-9 (1998): efectividad, eficiencia y satisfacción. Cada uno de estos atributos se ha dividido en diferentes factores que han sido analizados para recopilar las métricas existentes que se utilizar para medirlos. En la tabla 1 se resumen los resultados principales de este trabajo.

Como trabajo futuro nos proponemos incluir la definición de nuevos factores y métricas que resulten apropiados para evaluar los aspectos de usabilidad de las interfaces basadas en visión.



Tabla 1. Resumen de los factores y métricas de usabilidad

Atributos	Factores	Definición	Métricas
Efectividad	Precisión y ratio de errores	Corrección en el reconocimiento de gestos	Número de gestos reconocidos en base al número de gestos realizados
Eficiencia	Fatiga física	Cansancio que aparece al interactuar con movimientos del cuerpo	Evaluación subjetiva (Escala Likert, escala Borg CR10) Señales biológicas del usuario
	Duración	Tiempo que el usuario necesita para realizar el gesto	Tiempo que pasa entre el inicio y el final del gesto
	Carga cognitiva	Cantidad total de actividad mental impuesta sobre la memoria de trabajo	Evaluación subjetiva (cuestionarios con respuestas basadas en la escala Likert, NASA Task Load Index) Cantidad de información que una persona puede recordar mientras utiliza la interfaz
	Facilidad para aprender y facilidad para recordar	La facilidad para aprender, o el tiempo para aprender, es el tiempo y el esfuerzo requeridos para alcanzar un determinado nivel de rendimiento al usar la interfaz La facilidad para recordar es la facilidad del sistema para ser utilizado correctamente por usuarios ocasionales	Evaluación subjetiva Número de veces que se reinicia una presentación porque el usuario se equivoca al efectuar el gesto asociado a la función que aparece en la presentación
Satisfacción	Naturalidad y facilidad de intuición	La naturalidad está relacionada con la calidad de ser real y no requerir ningún dispositivo adicional creado por personas La intuición es el uso instintivo de gestos basados en lo que uno siente que deben ser, incluso sin razonamiento consciente	Evaluación subjetiva Análisis de interrupciones Método ascendente: descubre gestos a partir de funciones Método descendente: parte de los gestos y encuentra las funciones relacionadas
	Confort	El confort se define como un sentimiento placentero de estar relajado y libre de dolor	Evaluación subjetiva (Escala Likert, Body Part Discomfort (BPD)) Definir la zona de confort para el uso de los gestos
	Facilidad de uso	Poco esfuerzo para utilizar el sistema	Evaluación subjetiva
	Experiencia de usuario y satisfacción	Aspectos pragmáticos y hedonistas del sistema	Evaluación subjetiva
	Aceptación social	Conveniencia del gesto, tanto desde el punto de vista del usuario como desde el punto de vista de los espectadores, en el contexto en el que se realiza	Evaluación subjetiva

Agradecimientos

Este trabajo ha sido parcialmente financiado por A1/037910/11 de MAEC-AE-CID, Ajudes grup competitiu UGIVIA 28/2011 del Govern de les Illes Balears, y TIN12-35427 del Gobierno de España. C. Manresa-Yee también agradece la financiación de la beca de movilidad CAS12/00199, Programa José Castillejo del Ministerio de Educación, Cultura y Deporte, mediante el Programa Nacional de Movilidad de Recursos Humanos del Plan Nacional de I-D+i 2008-2011, prorrogado por Acuerdo de Consejo de Ministros de 7 de octubre de 2011.

Referencias

Abran, A., Khelifi, A., Suryan, W. and Seffah, A. (2003). “Usability Meanings and Interpretations in ISO Standards”. *Software Quality Journal*, 11(4), 325-338

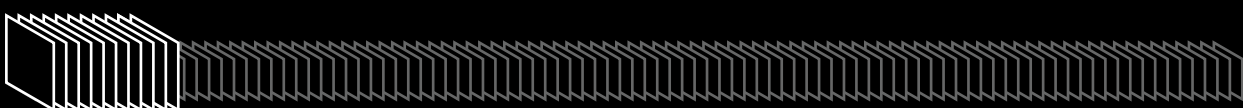
Bailly, G., Müller, J., Rohs, M., Wigdor, D. and Kratz, S. (2012). „ShoeSense: a new perspective on gestural interaction and wearable applications”. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, May 05-10, 2012, Austin, Texas, USA

Barclay, K., Wei, D., Lutteroth, C. and Sheehan, R. (2011). “A quantitative quality model for gesture based user interfaces”. *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, 31-39

Baudel T. and Beaudouin-Lafon, M. (1993). “Charade: remote control of objects using free-hand gestures”. *Commun. ACM* 36, 7, 28-35.

Borg, G. (1982). “Psychophysical Bases of Perceived Exertion”. *Medicine and Science in Sports and Exercise* (14), 377-381.

Bowman, D.A, McMahan, R.P. and Ragan, E.D. (2012). “Questioning naturalism in 3D user interfaces”. *Communications of the ACM* 55, 9 (September 2012), 78-88



Bragdon, A., DeLine, R., Hinckley, K. and Morris, R. (2011). “Code space: touch + air gesture hybrid interactions for supporting developer meetings”. Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS ‘11). ACM, New York, NY, USA, 212-221

Brooke, J. (1996). “SUS: a “quick and dirty” usability scale”. Eds. P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland. Usability Evaluation in Industry. London: Taylor and Francis.

Cabral, M.C., Morimoto, C.H. and Zuffo, M.K. (2005). “On the usability of gesture interfaces in virtual reality environments”. Proceedings of the 2005 Latin American conference on Human-computer interaction, 100-108

Corlett, E. N. and Bishop, R. P. (1976). “A Technique for Assessing Postural Discomfort”. Ergonomics 19(2), 175-182

De Silva, G. C. , Lyons, M. J. , Kawato, S. and Tetsutani, N. (2003). ”Human factors evaluation of a vision-based facial gesture interface”. Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop, 5, 52-59

Gamberini, L., Spagnoli, A., Prontu, L., Furlan, S., Martino, F., Solaz, B.R., Alcaniz, M., Lozano and J.A. (2013). “How natural is a natural interface? An evaluation procedure based on action breakdowns”. Personal and Ubiquitous Computing 17:69-79

Gonçalves, D., Jesus, R., Grangeiro, F., Romao, T. and Correia, N. (2008). “Tag around: a 3D gesture game for image annotation”. Proceedings International Conference on Advances in Computer Entertainment Technology, 259–262

Graetzel, C., Grange, S., Fong, T., and Baur, C. (2003). “A Non-Contact Mouse for Surgeon-Computer Interaction”. NCCR-COME Research Networking Workshop, Brauwald, Switzerland

Hassenzahl, M. (2004). “The interplay of Beauty, Goodness and Usability in Interactive products”. Human-Computer interaction. 19, 319-349

Hewett, B., Card, C., Gasen, M., et al. (1992). ACM SIGCHI curricula for human-computer interaction. <http://www.hcibib.org/>. Last visited: June 2013

Höysniemi, J., Hämäläinen, P., and Turkki, L. (2004). "Wizard of Oz prototyping of computer vision based action games for children". Proceedings of the 3rd International Conference on Interaction Design and Children (IDC'04), ACM Press, 27 -34

ISO. (1998). **ISO 9241-11:1998** Ergonomic requirements for office work with visual display terminals (vdts) – part 11: guidance on usability.

Kendon, A. (1988). "How gestures can become like words". Cross-Cultural Perspectives in Nonverbal Communication, 131-141

Kölsch, M., Beall, A. and Turk, M. 2003. "An Objective Measure for Postural Comfort". Proceedings. HFES 47th Annual Meeting

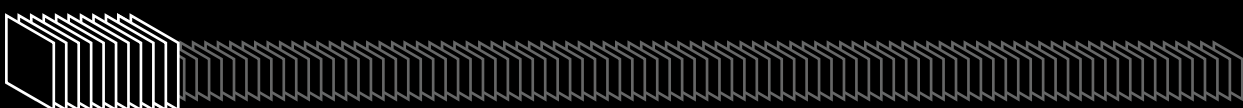
Lenman, S., Bretzner, L. and Thuresson, B. (2002). "Using Marking Menus to Develop Command Sets for Computer Vision Based Hand Gesture Interfaces". Proceedings of the Second Nordic Conference on Human-Computer Interaction, 239-242

Lorés, J., Granollers, T. and Lana, S. (2001). "Introducción a la interacción persona-ordenador". Diseño de sistemas interactivos centrados en el usuario. Editorial UOC

Lukowicz, P., Hanser, F., Szubski, C., Anser, F. and Scherberberger, W. (2006). "Detecting and interpreting muscle activity with wearable force sensors". Proceedings Pervasive 2006, LNCS 3968, 101-116

Manresa-Yee, C., Ponsa, P., Varona, J. and Perales, F.J. (2008). "User experience to improve the usability of a vision-based interface". Interacting with Computers 22 (6), 594-605

Manresa-Yee, C., Varona, J., Mas R. and Perales, F.J. (2005). "Hand Tracking and Gesture Recognition for Human-Com-



puter Interaction”. *Electronic Letters on Computer Vision and Image Analysis* 5(3), 96-104

Montero, C.S., Alexander, J., Marshall, M., and Subramanian, S. (2010). “Would you do that?—understanding social acceptance of gestural interfaces”. *Proceedings of the MobileHCI 2010*, 275-278

NASA TLX, NASA Ames Research Center, Moffet Field, California, 1988.

Nielsen, M., Störring, M., Moeslund, T.B. and Granum, E. (2004). “A procedure for developing intuitive and ergonomic gesture interfaces for HCI”. *Proceedings Gesture-Based Communication in Human-Computer Interaction. GW2003, LNAI 2915*, 409-420, Springer

Norman, D.A. (2010). “Natural user interfaces are not natural”. *Interactions* 17, 3 (May 2010), 6-10

Norman, D.A. and Nielsen, J. (2010). “Gestural interfaces: a step backward in usability”. *Interactions* 17(5) (September 2010), 46-49

Pavlovic, V.I., Sharma, R., and Huang, T-S. (1997). “Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review”. *IEEE transactions on pattern analysis and machine intelligence*, 19 (7)

Porta, M. (2002). “Vision-based user interfaces: methods and applications”. *International Journal of Human-Computer Studies*, 57(1), 27-73

Rico, J. and Brewster, S. (2010). “Usable gestures for mobile interfaces: evaluating social acceptability”. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, ACM, 887-896.

Saffer, D. (2009). “Designing Gestural Interfaces: Touchscreens and Interactive Devices”. O’Reilly Media, Sebastopol, CA

Shneiderman, B. (1998). “Designing the user interface: strategies for effective human-computer interaction”. Addison-Wesley. 3rd ed.

Sukeshini A. G., Joue, G. and Mittelberg, I. (2011). “Understanding naturalness and intuitiveness in gesture production: insights for touchless gestural interfaces”. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘11). ACM, 821-824

Tollmar, K., Demirdjian, D. and Darrell, T. (2004). “Navigating in virtual environments using a vision-based interface”. Proceedings of the third Nordic conference on Human-computer interaction (NordiCHI ‘04). ACM, New York, NY, USA, 113-120

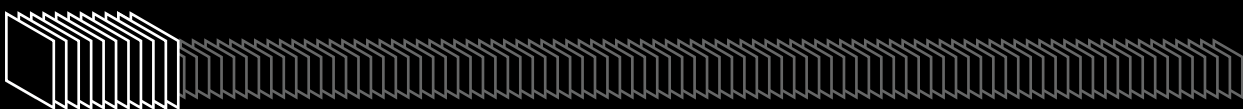
UX Fellows. (2003) Thumbs up to gesture-controlled Consumer Electronics? A cross-cultural study spanning 18 countries on spontaneous gesture behavior. In <http://www.uxfellows.com>. Last visited June 2013.

van Beurden, M.H.P.H., Ijsselsteijn, W.A. and de Kort, Y.A.W. (2012). “User Experience of Gesture Based Interfaces: A Comparison with Traditional Interaction Methods on Pragmatic and Hedonic Qualities”. Gesture and Sign Language in Human-Computer Interaction and Embodied Communication. Lecture Notes in Computer Science, Volume 7206, 36-47

Waldherr, S., Romero, R. and Thrun, S. (2000). “A gesture-based interface for human-robot interaction”. Autonomous Robots 9

Wickeroth, D., Benölken, P. and Lang, U. (2009). “Manipulating 3D Content using Gestures in Design Review Scenarios”. International Journal of Information Studies, 1(4), 242-250

Winograd, T. and Flores, F. (1986) Understanding computers and cognition. Ablex Publishing, Norwood



Sobre los autores

Cristina Manresa-Yee es Doctora en informática por la Universitat de les Illes Balears. Actualmente imparte docencia en la Universitat de les Illes Balears en diseño de interfaces de usuario e interacción persona-ordenador. Sus líneas de investigación se centran en interacción persona-ordenador, interfaces naturales, visión por computador e interfaces basadas en visión (VBI). Es miembro del Special Interest Group in Computer-Human Interaction (SIGCHI) de la ACM.

crisrina.manresa@uib.es

Rodríguez, Esperança Amengual es Doctora por la Universitat de les Illes Balears. Actualmente imparte docencia en ingeniería del software e ingeniería web en la Universitat de les Illes Balears. Sus líneas de investigación son la mejora de procesos software, ingeniería de la usabilidad y Brain Computer Interaction. Ha participado en el proyecto QuaSAR, un programa de mejora de procesos software en pequeñas y medianas empresas en la Islas Baleares. Es miembro del Grupo de Calidad de Software de ATI España y asesora ISO/IEC 15504.

eamengual@uib.es

Pere Ponsa Asensio es Doctor por la Universitat Politècnica de Catalunya (UPC) y Licenciado en Ciencias (Física) por la Universitat Autònoma de Barcelona (1994). Actualmente imparte docencia en la UPC en diseño centrado en el usuario, interacción persona-robot y teleoperación, interacción persona-sistema. Su investigación en los últimos años se ha llevado a cabo en la interacción de las personas con sistemas automatizados y la creación de guías de diseño de interfaces (gráficas, hápticas). Su investigación más reciente está orientada al análisis de sistemas complejos y en la interacción gestual. Es miembro de la asociación española IPO.

pedro.ponsa@upc.edu



© Leonardo Infante

La Usabilidad de las Interfaces Basadas en Visión

Las interfaces basadas en visión hacen uso de gestos para la comunicación del usuario con el sistema interactivo sin la necesidad de dispositivos que requieran contacto físico. El modelado de gestos suelen realizarse en laboratorios y es importante que se lleve a cabo la evaluación de su usabilidad, pero la gran diversidad de usos y aplicaciones de estas interfaces hace que resulte difícil decidir qué factores tener en cuenta cuando se mide la usabilidad. En este artículo se presenta una revisión de la literatura cuyo objetivo es clasificar y recopilar los factores y las métricas de usabilidad que se utilizan para validar las interfaces basadas en visión.