

RICH COMMUNICATION OVER INTERNET

Introduction

5 The present invention relates to an intuitive and user-friendly user interface for rich communication on a network which interacts in an efficient way with other applications and services. In particular, the present invention relates to rich and expressive real-time communication over the Internet supported by animated objects using the objects to increase the expressive and emotional bandwidth of instant messaging. The invention is suited for use on a broad range of Internet
10 terminal types, from mobile phones to PC's and TV's with set-top boxes.

Background of the invention

In recent years, there has been much diverse research, which explores the use of computing in ways that involve human emotion. This area is commonly referred to
15 as *affective computing*. This includes research on the use of emotions in human-computer interaction, artificial intelligence (AI) and agent architectures which are inspired by the mechanisms of emotion, the use of emotion in computer-mediated communication, the study of human emotion through computers and philosophical issues concerning, for example the extent to which it is meaningful to talk about
20 emotion in computational terms.

Emotional expression are often described as social and communicative by nature (Averill 90). Humans are, after all, fundamentally social beings. Infants rely completely on others to meet their needs at birth and throughout early childhood, and continue to rely on others to help meet their needs to varying degrees
25 throughout life. A primary function of emotions is to communicate state information to others, in order to enable them to assist in meeting the needs of the individual.

There are various dimensions associated with mediated interaction. It can be synchronous or asynchronous. The communication can be written, auditory or
30 visual. Mediated communication can develop its own form, syntax and context. One can see that writing for example has developed into a medium that can bring forth a whole range of emotions and feelings that are impossible to replicate using the spoken word in a face-to-face situation. In a similar way, telephonic interaction has its own style and form. This includes the tone of voice one uses and the way
35 that one replaces the visual with verbal gestures (Ling 1999, Ling 1996).

A part of the information richness in face-to-face interaction lies in its spur-of-the-moment nature. Conversation partners have a large set of para-communication

types available: various intended utterances, winks, nods, grounding and clearance signals. One of the things that makes synchronous face-to-face interaction particularly rich and also particularly precarious, is that the signs one 'gives off' are a large portion of the total message (Ling 1999).

- 5 Humans are experts at interpreting facial expressions and tones of voice, and making accurate inferences about others internal states from these clues. Controversy rages over anthropomorphism. The types of emotional needs, which the present invention aims at giving support for accommodating in computer-mediated communications, include the following:
- 10 - for attention – strong and constant in children, fading to varying degrees in adulthood
- to feel that one's current emotional state is understood by others (particularly strong during emotional response)
- to love and feel reciprocity of love
- 15 - to express affection, and feel reciprocated affection expressed
- for reciprocity of sharing personal disclosed information
- to feel connected to others
- to belong to a larger group
- for intimacy
- 20 - to feel that one's emotional responses are accepted by others
- to feel accepted by others
- to feel that emotional experiences and responses are 'normal'

Instant messaging (IM) is, like email and chat, a way for net users to keep in touch with each other. Unlike chat and email, IM allows users to see when their friends are online and to initiate instant, live communication.

25

The market for IM solutions is expected to show exceptional growth in the coming years driven by broadband telecommunication and cable offerings, always-on Internet connection from mobile phones as well as by changes in the business environment and in people's lifestyles.

- 30 IM type applications are expected to replace email as the main communication channel of the Internet over the next few years.

Present IM solutions are focused primarily on task and work related communication needs. The rapidly increasing accessibility of the Internet outside the work environment is creating a large and fast growing market for IM solutions better suited for private and social use.

35

A serious limitation of IM as a communication channel is the lack of support for para-communication types and expressions of emotion, affection, humour and irony.

5 The present invention overcomes the limitations of known IM communication by using an intuitive and user-friendly user interface for rich communication over the Internet. Instant messaging applications developed according to the present invention are also more effective in working together with other types of applications simultaneously active on the screen of the user terminal.

10 Applications developed according to the invention enable people to exchange both text, gestures and integrated text/gestures messages in real-time over the Internet. Two or more people can participate in the messaging sessions.

Applications developed according to the invention are well suited for deployment on a wide range of Internet terminal types – desktop PC's, TV's and mobile terminals.

15 Instant messaging applications such as ICQ, MSN messenger and Yahoo messenger let people communicate in real-time over the Internet. The communicated information is normally text-based. The text messages can be supplemented with 'emoticons' – small picture icons representing different expressions, gesture, moods or other non verbal messages.

20 The present invention increases the expressive and emotional bandwidth of instant messaging through the use of animated objects to present a wide range of emotional and affective expressions. Users are represented by avatars, which are animated objects controlled by a user.

25 In avatar chat sessions, users are normally represented in the form of animated character objects in a chat room or virtual world. The animated objects can normally be moved around in the chat space or virtual world. The animated objects cannot be moved outside of the frame representing the chat room. Nor can the animated objects interact with other active applications on the user screen.

30 US-5,880,731 describes the use of avatars with automatic gesturing and bounded interaction in an on-line chat session. This is a typical example on avatar chats with the graphical objects restricted to the frames of a specific program. Another example is shown in US-6,219,045 describing a scalable virtual world chat client-server system.

35 The advantages of the present invention compared to known avatar chat are primarily related to the animated objects being freely moveable on the whole user screen. It thus becomes possible for the animated objects to interact with other

objects and applications present on the user screen. The animated objects are less intrusive and distracting for the user when the user's primary focus is on another application, for instance a text processor.

On a PC, MS Windows, Linux or another operating systems will function as the user interface between a user and various other programs. These programs can interact with each other. A word processor program like MS Word will function as a user interface between a user and the spreadsheet program MS Excel if the user starts MS Excel as a linked object from within Word. The present invention will also represent a user interface between the user and other applications.

According to the invention, the users can place the animated object(s) representing themselves and other users on the user interface screen of their Internet terminal. Users can also place animated object(s) representing Internet based services on the user interface screen of their Internet terminal. The animated objects representing the different users and/or services can freely and independently be moved around and placed anywhere on the user interface screen of the users Internet terminals. Users can then communicate and share information with each other through interaction with the animated objects representing themselves and other users. Users can further communicate and interact with Internet based services through interaction with the animated objects representing themselves and the animated objects representing Internet based services. Groups comprising two or more users can share Internet based services through interaction with the animated objects representing themselves, the animated objects representing other users and the animated objects representing Internet based services. Users can communicate and share information through interaction between the animated objects and manifestations of other software applications on their terminals. The interaction of the users can be done by using a computer mouse, keyboard, remote control, pointing device or voice commands to make their representation (animated object) present information. The information is presented in the form of an animation sequence performed by the animated object, possibly in combination with animation sequences performed by the animated objects representing one or more of the user's communication partners. The animation sequences can be combined with text, audio, or other forms of information representations.

Detailed description

The present invention relates to an intuitive and user-friendly user interface for rich communication on a network and which interacts in an efficient way with other applications and services. In particular, the present invention relates to rich and

expressive real-time communication over the Internet supported by animated objects using the objects to increase the expressive and emotional bandwidth of instant messaging.

- 5 The present invention comprises thus a method for communicating synchronous information and gestures from a user on a terminal to a plurality of users on other terminals in a network, the method comprising the steps of:
- presenting users in the form of animated objects freely moveable on the terminals screens,
- 10 initiating, upon detecting objects which represent other users on the screen, in the proximity zone of the object representing the user, communication and interaction with said other terminals associated with respective other users, on the terminals, receiving signals from a user operated input device indicative of a specific action or expression to be represented as animation of the said object representing said user,
- 15 reconstructing and playing the received action or expression on the user terminals,
- transmitting to the terminals the received and interpreted signals from a user input device, describing user's initiated communication and animation, making thus this information available for other users.
- 20 In a preferred embodiment, the initiation is activated when objects placed freely on the screen, are moved closer to each other than 300 twips on the user screen. In another preferred embodiment, the initiation is activated when users interact with the objects representing themselves on the user screen. Further, in another preferred embodiment the initiation is activated when users
- 25 interact with an object representing another user on the user screen. In a preferred embodiment, the animation signals received are instructions to a skeletal animation system. In a preferred embodiment, the animations are represented as floating above the background and other applications on the screen by clipping the region on which
- 30 the animations are represented around the edge of the animated object. In a preferred embodiment, the animations are represented in the form of 3D renderings created from the animation signals by a processor on the user's terminal. In a preferred embodiment, the reconstruction includes receiving and interpreting
- 35 animation information sent from other users, checking if animation already exists on the user's terminal. In a preferred embodiment, the signals are transmitted in the form of XML encoded messages. In another preferred embodiment, the signals are transmitted in the form of SOAP

messages transmitted over HTTP.

In a preferred embodiment, the signals are transmitted over TCP or UDP protocols.

In a preferred embodiment, the input device is a computer mouse, keyboard, remote control, pointing device, VR peripherals, camera and/or voice commands

5 communicating the specific action or expression.

The invention comprises also a method for sharing information and applications between a plurality of users on terminals in a network, the method comprising the steps of:

10 presenting users in the form of animated objects freely moveable on the terminal screens,

initiating sharing of an application between a group of users by moving animated objects representing the users into the window area representing an application.

The invention further comprises a method for transmitting or making files
15 available to a user on a terminal in a network, the method comprising the steps of: presenting users in the form of animated objects freely moveable on the terminal screens,

moving the icon or other representation of the file to be shared into the proximity zone of the animated object representing the user.

20 The invention further comprises a method for initiating a synchronous communication session between a plurality of users on other terminals in a network, the method comprising the steps of:

presenting users in the form of animated objects freely moveable on the terminal screens,

25 initiating, upon detecting 2 or more objects which represent other users on the screen, in the proximity zone of the object representing the user, group communication and interaction with said other terminals associated with respective other users,

persisting the group to a storage structure on the network.

30 Seen from the client's point of view, the rich communication with gestures is achieved by presenting users, in a chat session, in the form of animated objects freely moveable on the user screen. The animated objects are selected and downloaded locally to the client from a server on the network. Communication and interaction with other users is initiated when objects on the user screen

35 representing other users are moved into the proximity zone of an object representing the user. The user can be represented in several instances at once allowing the user to participate in multiple proximity zones at once. By placing an object representing another user on the desktop, the other user is granted instant and continuous access to presence information from the user. Users who have their

representation on the desktop without being in the proximity zone of other characters will be able to broadcast status gestures to all users subscribing to information from the user by manipulating his screen representation. A user can at any time change his or her representation by manipulating the representation. It is possible to have different representations for different instances of the user.

Transmission of gestures can be initiated through manipulation of a text input box, a drop down menu, direct manipulation of the representations or direct access through shortcuts triggered by various physical interfaces. Gestures can be synchronized directly with text messages by adding the command into the text string. Gesture also can be accompanied by sound coordinated with the motion.

The representations can make gestures directed towards the screen. In a group situation with more than two participants the representations may form groups which interact synchronously towards another representation. After a gesture is sent, the representation will make a transition to an idle state which can reflect the last acted gesture.

The representations can alter size according to activity or user input. The terminal receives signals generated from a user operated input device indicative of a specific action or expression to be represented as animation of the object representing the user. Input devices may include a computer mouse,

keyboard, remote control, pointing device, VR (Virtual Reality) peripherals,

camera and/or voice commands. A command action which initiates a gesture can be typed in or accessed through a pull down menu accessed through the keyboard. The menu is controlled with mouse pointer, number or arrow keys. A gesture can also be suggested by the system as a result of interpreting the text input. The user can enter any number of animations into a text string. Some of the animations can also be directly altered in the text interface through a script. Some gestures can be influenced by the receiving representation through counter gestures. The counter gestures are made available in the interface in different situations, i.e.

representation starts an aggressive move and the receiving character responds by altering the initiated gesture.

The communication can also include sharing of applications (between two computers, both of which can view and interact in the session) and files, i.e. a user using a browser may share the browsing experience with several other users taking part in the communication session, and at the same time communicate expressions by inputting animation instructions to the client. An application sharing session can be initiated by dragging other user representations into the proximity zone of an application. One can also initiate a shared application by manipulating a representation of another user.

A received file can be received and presented visually by the relevant user representation.

In communication with text based IM applications, gestures sent to users are translated to hyperlinks. If the hyperlink is activated by the receiving system user, a web page is open with tools to send and receive gesture messages.

5

In a preferred embodiment of the invention the information describing the interaction is encoded in XML (eXtensible Markup Language) and routed between the users by a presence and notification server on the network. Alternative form of encoding the information as well as transmission of messages directly between users terminals can however be envisaged. The information describing the interaction contains animation instructions which are interpreted by software application on the users terminals. The type of terminal used, will decide the complexity and layout of the rendering of the information. The software application on a terminal with good graphic capabilities will render real-time 3D animation on the user terminal screen on the basis of skeletal animation instructions contained in the animation instructions. A low-end terminal with limited graphics capabilities, for instance a mobile phone, will show a sequence of pre-rendered images which are downloaded from the network on the basis of instructions contained in the information describing the interaction. On a terminal with text-only capabilities the interaction will be described in text. On a audio-only terminal the interaction will be described in audio. The signals in the form of XML encoded messages, SOAP (Simple Object Access Protocol) messages or other type of message encoding may be transmitted over for example TCP (Transmission Control Protocol) or UDP (User Datagram Protocol) protocols.

25

In one embodiment of the invention a presence and notification server connected to user terminals by means of a network that coordinates communication of information and gestures from one user on a terminal to a plurality of users on other terminals in the network by routing and sending information to users taking part in a communication session over the Internet. Information about participating users is stored in a data structure on a server in the network. The server keeps track of the types of terminals each user is using and adapts the information transmitted to each terminal accordingly. The data structure containing information about the users and their terminals could be part of the presence and notification server software or it could be part of a separate system communicating with the presence server, for instances an LDAP (Lightweight Directory Access Protocol) server. The above description is illustrative and not restrictive. In another embodiment of the invention the communication session is initiated over SIP (Session Initiation Protocol) while the near-real time communication and interaction between users is routed directly between users (Peer-to-peer).

40

With reference to the drawings, the following detailed description will explain how this is obtained. The drawings illustrate a preferred embodiment of the present invention, and it is obvious that a person skilled in the art may well derive other variations.

FIG. 1 shows a network diagram showing a plurality of users communicating with messages containing animation instructions and text being routed by an IM server. The clients have an IM application installed which is able to communicate with the IM server. The communication to and from the IM server is over TCP/IP. Users are offered a plurality of animated objects made available on an IM server. Users select one or more animated objects to represent themselves in computer mediated communication sessions. Messages describing the animations and information are sent from the user's terminal to the IM server. The IM server forwards the messages to terminals used by the communication partners. The messages are interpreted by the terminals and the information is presented both on the user terminal from where the message is sent and on the terminals of the active communication partners. The methods used for presenting the information are dependant on the capabilities of the Internet terminals used by the different communication partners. Information that is presented in the form of an animation sequence on one terminal, for instance a desktop PC or TV' set with set-top box, could, on for instance a mobile phone, be represented in text or audio form.

FIG. 2 shows a sequence of screens showing a user scenario in two steps where in scene 1 user A initiates an animation by sending a message describing the animation to an IM server. In scene 2 the message has been relayed from the IM server to user A's subscribers resulting in the animation being played on these user's screens.

FIG. 3 Is a flow chart that defines the logical steps implemented in the use scenario shown in fig.2. Scene 1 describes how messages are sent. Sending user initiates a specific action or expression of own avatar. The message containing the information about the information is then sent to the IM server. The server will then route the message to contacts currently subscribing to the information. Scene 2 how messages are received. Messages from the IM server arrives at users terminals. The messages is then decoded and information about animation is extracted. Each terminal will then check if the current animation is present locally on the terminal. If this is not the case, it will be downloaded from the IM server. The animation is then played on the users terminal screens.

FIG. 4 Illustrates some of the terminal types suitable for use in implementing the present invention.

FIG. 5 Is shows a schematic block diagram of a sample network configuration suitable for implementing the present invention with support for different types of terminals (examples of possible terminals). This is basically the same as described in Fig. 1, but with added servers serving other services like interactive TV, WEB, WAP etc.

The invention described herein is not restricted to the setup described, but may be implemented on any setup with users using any form of interactive service using a screen to present animations.

Claims

1. Method for communicating synchronous information and gestures from a user on a terminal to a plurality of users on other terminals in a network, the method comprising the steps of:
 - 5 a) presenting users in the form of animated objects freely moveable on the terminals screens,
 - b) initiating, upon detecting objects which represent other users on the screen, in the proximity zone of the object representing the user, communication and interaction with said other terminals associated with respective other users,
 - 10 c) on the terminals, receiving signals from a user operated input device indicative of a specific action or expression to be represented as animation of the said object representing said user,
 - d) reconstructing and playing the received action or expression on the user terminals,
 - 15 e) transmitting to the terminals the received and interpreted signals from a user input device, describing user's initiated communication and animation, making thus this information available for other users.
 - 20
2. Method according to claim 1, where in step b) the initiation is activated when objects placed freely on the screen, are moved closer to each other than 300 twips on the user screen.
- 25 3. Method according to claim 1, where in step b) the initiation is activated when users interact with the objects representing themselves on the user screen.
4. Method according to claim 1, where in step b) the initiation is activated when users interact with an object representing another user on the user screen.
- 30 5. Method according to claim 1, where in step c) the animation signals received are instructions to a skeletal animation system.

6. Method according to claim 1, where in step c) the animations are represented as floating above the background and other applications on the screen by clipping the region on which the animations are represented around the edge of the animated object.

5

7. Method according to claim 1, where in step c) the animations are represented in the form of 3D renderings created from the animation signals by a processor on the user's terminal.

10

8. Method according to claim 1, where in step d) the reconstruction includes receiving and interpreting animation information sent from other users, checking if animation already exists on the user's terminal.

15

9. Method according to claim 1, where in step e) the signals are transmitted in the form of XML encoded messages.

10. Method according to claim 1, where in step e) the signals are transmitted in the form of SOAP messages transmitted over HTTP.

20

11. Method according to claim 1, where in step e) the signals are transmitted over TCP or UDP protocols.

25

12. Method according to claim 1, where in step e) the input device is a computer mouse, keyboard, remote control, pointing device, VR peripherals, camera and/or voice commands communicating the specific action or expression.

30

13. A method for sharing information and applications between a plurality of users on terminals in a network, the method comprising the steps of:

- a. presenting users in the form of animated objects freely moveable on the terminal screens,
- b. initiating sharing of an application between a group of users by moving animated objects representing the users into the window area representing an application.

14. A method for transmitting or making files available to a user on a terminal in a network, the method comprising the steps of:

- a. presenting users in the form of animated objects freely moveable on the terminal screens,
- b. moving the icon or other representation of the file to be shared into the proximity zone of the animated object representing the user.

15. A method for initiating a synchronous communication session between a plurality of users on other terminals in a network, the method comprising the steps of:

- a. presenting users in the form of animated objects freely moveable on the terminal screens,
- b. initiating, upon detecting 2 or more objects which represent other users on the screen, in the proximity zone of the object representing the user, group communication and interaction with said other terminals associated with respective other users,
- c. persisting the group to a storage structure on the network.