

PHONE BASED DYNAMIC IMAGE ANNOTATIONBackground of the Invention5                   Field of the Invention

The present invention relates to the storage and retrieval of multimedia messages and more particularly to utilizing the store and forward capabilities of a voice mail system for multimedia messages containing  
10   dynamically annotated images.

Description of the Prior Art

The availability of hardware and software to generate, store and distribute digital media (text,  
15   voice, images, video, graphics, etc....) has enabled the development of multimedia messaging systems that give people the ability to communicate without having to be in the same place at the same time. In order for these systems to be effective, they must provide the means to  
20   convey as much information as possible in a simple and intuitive way. Within this context, the concept of image annotation which involves the ability to superimpose voice, text, graphics and mouse movements on a previously recorded image is appealing because it lets the sender  
25   act as if the receiver were present. For example, a person could verbally and graphically describe the way in which a document could be changed as if the other person was standing next to him/her. In general, dynamic image annotation extends the way in which people can use  
30   computers to communicate.

The use of annotation to enhance computer-based communication has been available in various forms. In its simplest way, text documents can be extended with special markers that resemble "post-it" notes where a  
35   person could add typewritten comments. More sophisticated systems enable the creation of annotations on images or documents that include sound and graphics.

However, in these it is assumed that the computers where the annotations are generated and played back possess the necessary hardware/software to perform audio recording and playback.

5        It is an object of the present invention to develop a unique annotation that can be used in environments where the computers used to compose and play back the annotation lack audio capabilities, but where a separate telephone device is available. The basic idea is to use  
10    the computer to record graphics and mouse movements (gestures) and a voice mail device to record the voice component of the message.

#### Summary of the Invention

15        The present invention extends the use of dynamic image annotation to PCs or workstations lacking full multimedia capabilities. The system enables the composition, storage and retrieval of multimedia messages containing dynamically annotated images by utilizing the  
20    store and forward capabilities of a voice mail system and a telephone device to carry the voice/sound portion of the message. The system includes protocols used to maintain the synchronization of the voice/sound components of the message during recording, storage and  
25    playback.

#### Brief Description of the Drawings

30        Figure 1 illustrates four potential multimedia scenarios.

Figure 2 illustrates the record protocol of one embodiment of the present invention.

Figure 3 illustrates the playback protocol of one embodiment of the present invention.

35        Figure 4 illustrates system components of a test of the present invention.

Detailed Description of the Invention

Figure 1 illustrates four scenarios in which the dynamic annotation of the present invention can be used within a voice mail environment. In scenario 1 a user with a multimedia PC 11 accesses the voice mail system 15, including dynamic image annotation, through a local area network 12. In scenario 2 a user with a non-multimedia PC 13 (i.e., a PC without speakers, microphone or telephone interface) accesses the services of the voice mail system 15 through the telephone 14 and the local area network 12. The telephone connection is achieved by means of a public branch exchange (PBX) 16. In scenario 3 a user of a portable non-multimedia PC 17 (i.e., a portable PC without speakers, microphone or telephone interface) accesses the services of the voice mail system 15 through the public telephone network 20 by means of a telephone 18 and modem 19. In scenario 4 a user with a multimedia PC 21 accesses the services of the voice mail system 15 remotely via the telephone network 20.

Of particular interest to the present invention are scenarios 2 and 3 which require the availability of a telephone device. More specifically, in scenario 2, the voice component or the annotated message is carried over the voice network while the graphics and gestures travel simultaneously over the local area network. In scenario 3, the voice component of the annotated message is again delivered to the voice mail system over the voice network. The rest of the annotation, including the image, graphic and mouse gestures, are also delivered to the voice mail system via a separate phone line. In an alternative implementation of scenario 3, the graphics and gesture components of the annotation can be recorded in the local PC and delivered to the voice mail system over the same telephone line that carried the voice.

A typical multimedia annotated message consists of one or more images, a voice/sound message and a timed sequence of gestures consisting of drawings, mouse

movements and viewing commands, all of which are synchronized in time. The objective of the present invention is to design a protocol that will enable the separation of the voice/sound component during  
5 transmission and storage and the reattachment of the voice/sound component during playback. The key problem in the development of this protocol is the identification of suitable synchronization mechanisms.

Figure 2 illustrates the recording portion of a  
10 protocol and the interconnections between the computer 24, the local area network 26, the voice mail system 28, the PBX 30 and the telephone device 32 of one embodiment of the present invention. The user begins a session by selecting the recording mode, then chooses the image(s)  
15 to be annotated. When the selection is completed the system will ask the voice mail system 28 to initiate a call to the appropriate telephone device 32 and then will enter a waiting state. It is assumed that there are mechanisms in place to determine the device to be called  
20 and to validate the user identification. The voice mail system 28 will initiate the call, wait for the user to answer and play a prompt indicating the reason for the call and asking the user to perform a specific action, e.g., push a button to confirm that he is ready to begin  
25 annotating the image(s). Upon receiving confirmation from the user, the voice mail system 28 will then instruct the annotation program on the PC 24 to go ahead with the annotation and will begin recording the voice message sent through the telephone device 32. Notice  
30 that there are real-time constraints to be satisfied in this step. More specifically, the system must be able to instantly notify the annotation program to go ahead. The program must also be able to return from the waiting state immediately. A series of synchronization marks  
35 ( $T_i$ ) are recorded during this step to keep track of the duration of the session. In particular,  $T_1$  refers to the time elapsed between the point where the annotation

begins and the first pause,  $T_2$  refers to the duration of the first pause and  $T_3$  to the duration of the second segment of the annotation. Notice that markers  $T_4$ ,  $T_5$ ,... may be needed if there are two or more pauses during the recording of the annotation. The user will indicate completion of the annotation by pressing a 'STOP' button and hanging up the telephone device 32. The session will be completed (from the user's point of view) when the user issues a 'SEND' command. The system will continue its communication with the voice mail system 28 in the background, sending the image(s), gestures and synchronization markers to the message store where they will be stored with the voice portion of the message.

Figure 3 illustrates the steps to be performed during playback. The interconnections between the computer 34, the local area network 36, the voice mail system 38, the PBX 40 and the telephone device 42 are as shown in Figure 2. In the first step, the user receives a list of the messages available and selects one to be played back. The next two steps in the process are similar to those described in the protocol above. In the third step, the user will receive the voice message over the telephone device 42 while the gesture portion of the annotation is being displayed on the screen of the computer 34. The synchronization markers collected during the recording session will be used to coordinate the two devices. More specifically, the gesture player will periodically compare the execution and real times. If necessary the gesture player may drop small gestures to speed up the process. The gesture player may also introduce small delays to re-synchronize the voice and gesture annotation.

As a test of the present invention, a simple concept system was created that simulates the functionality of a telephone and voice mail components. This system is illustrated in Figure 4. The system consists of three UNIX processors running on two separate workstations.

Processor 1 runs the image annotation software extended with remote procedure calls that simulate the communication that would take place with the voice mail server. Processor 2, running in the background, 5 simulates the functionality of the voice mail system; it receives and sends commands to the annotation program (processor 1) and telephone device (processor 3) as defined in the protocols. Finally, processor 3 is used to emulate a standard telephone device. It uses a 10 graphical user interface and the microphone and speakers of a second workstation to record and play the voice portion of the message. The above testing system shows that the desired functionality of the present invention is achievable.

15 As stated above, the present invention extends the use of dynamic image annotation to PCs or workstations lacking full multimedia capabilities. Additional benefits derived from the use of a telephone to transmit the voice/sound component of a multimedia message include 20 an improvement in the quality of the sound recorded and delivered and a reduction in traffic and bandwidth requirements of the local area network.

It is not intended that the present invention be limited to the hardware or software arrangement, or 25 operational procedures shown disclosed. This invention includes all of the alterations and variations thereto as encompassed within the scope of the claims as follows.



## CLAIMS:

1. A phone based dynamic image annotation system comprising:
  - 5 computer means comprising a computer screen;  
a local area network connected to said computer means;  
a voice mail system connected to said local area network;
  - 10 PBX means connected to said voice mail system;  
telephone means connected to said PBX means; and,  
protocol means for recording and playing back a message; wherein during said recording, a voice component of said message is decoupled so that said
  - 15 telephone means can carry said voice component of said message and wherein during said playing back, said voice component played over said telephone means is synchronized with a graphics and gesture component which is drawn overlapping an image of said message displayed
  - 20 on said computer screen.
2. A phone based dynamic image annotation system as claimed in claim 1 wherein said protocol means comprises:
  - 25 recording means; and,  
playback means.
3. A phone based dynamic image annotation system as claimed in claim 2 wherein said recording means comprises:
  - 30 loading means for loading said image on said computer screen;  
initiate means for asking said voice mail system to initiate a call to a device and to enter a waiting state; wherein said voice mail system initiates said call and
  - 35 upon answering of said call, plays a prompt indicating reason for said call and asks user to perform a specific action;





confirmation means for providing confirmation from said user;

instruction means for having said voice mail system instruct annotation program of said computer means to do  
5 annotation while said user begins recording a voice message through said telephone means; and,

communication means for having said voice mail system send images, gestures and synchronization markers to a message store to be stored with said voice message.  
10

4. A phone based dynamic image annotation system as claimed in claim 2 wherein said playback means comprises:

loading means for loading said image on said computer screen;

15 initiate means for asking said voice mail system to initiate a call to a device and to enter a waiting state; wherein said voice mail system initiates said call and upon answering of said call, plays a prompt indicating reason for said call and asks user to perform a specific  
20 action;

confirmation means for providing confirmation from said user;

transmission means for sending said voice messages over said telephone means while said images and gestures  
25 are displayed on screen of said computer means; and,

synchronization means for using said synchronization markers for coordinating said telephone means and said screen of said computer means.

30 5. A phone based dynamic image annotation system as claimed in claim 4 wherein said playback means further comprises:

retrieve messages means for said computer means to request a list of messages from said voice mail system;

35 message list means for said voice mail system to provide said list of messages;





select message means for said computer means to  
select one or more messages from said list of messages;  
and,

transmission and display means for said voice mail  
5 system to provide said one or more messages.

6. A phone based dynamic image annotation system  
comprising:

computer means comprising a computer screen;  
10 modem means connected to said computer means;  
telephone means connected to said modem means;  
a public telephone network connected to said modem  
means;

PBX means connected to said public telephone  
15 network;

a voice mail system connected to said PBX means;  
and,

protocol means for recording and playing back a  
message; wherein during said recording, a voice  
20 component of said message is decoupled so that said  
telephone means can carry said voice component of said  
message and wherein during said playing back, said voice  
component played over said telephone means is  
synchronized with a graphics and gesture component which  
25 is drawn overlapping an image of said message displayed  
on said computer screen.

7. A phone based dynamic image annotation system as  
claimed in claim 6 wherein said protocol means comprises:

30 recording means; and,  
playback means.

8. A phone based dynamic image annotation system as  
claimed in claim 7 wherein said recording means  
35 comprises:

loading means for loading said image on said  
computer screen;





initiate means for asking said voice mail system to initiate a call to a device and to enter a waiting state; wherein said voice mail system initiates said call and upon answering said call, plays a prompt indicating  
5 reason for said call and asks user to perform a specific action;

confirmation means for providing confirmation from said user;

instruction means for having said voice mail system  
10 instruct annotation program of said computer means to do annotation while said user begins recording a voice message through said telephone means; and,

communication means for having said voice mail system send images, gestures and synchronization markers  
15 to a message store to be stored with said voice message.

9. A phone based dynamic image annotation system as claimed in claim 7 wherein said playback means comprises:

loading means for loading said image on said  
20 computer screen;

initiate means for asking said voice mail system to initiate a call to a device and to enter a waiting state; wherein said voice mail system initiates said call and upon answering of said call, plays a prompt indicating  
25 reason for said call and asks user to perform a specific action;

confirmation means for providing confirmation from said user;

transmission means for sending said voice messages  
30 over said telephone means while said images and gestures are displayed on screen of said computer means; and,

synchronization means for using said synchronization markers for coordinating said telephone means and said screen of said computer means.

35





WO 97/11549

THIS SHEET HAS BEEN  
REPLACED

PCT/US96/13217

11

10. A phone based dynamic image annotation system as claimed in claim 9 wherein said playback means further comprises:

retrieve messages means for said computer means to  
5 request a list of messages from said voice mail system;

message list means for said voice mail system to provide said list of messages;

select message means for said computer means to select one or more messages from said list of messages;

10 and,

transmission and display means for said voice mail system to provide said one or more messages.

11. A method of transferring dynamically annotated  
15 images comprising the steps of:

utilizing computer means comprising a computer screen;

connecting said computer means to a local area network;

20 connecting a voice mail system to said local area network;

connecting PBX means to said voice mail system;

connecting telephone means to said PBX means; and,

utilizing protocol means for recording and playing  
25 back a message; wherein during said recording, a voice component of said message is decoupled so that said telephone means can carry said voice component of said message and wherein during said playing back, said voice component played over said telephone means is  
30 synchronized with a graphics and gesture component which is drawn overlapping an image of said message displayed on said computer screen.

12. A method of transferring dynamically annotated  
35 images as claimed in claim 11 wherein utilizing protocol means comprises the steps of:

This Sheet Has Been Replaced





recording said message; and,  
playing back said message.

13. A method of transferring dynamically annotated  
5 images as claimed in claim 12 wherein recording said  
message comprises the steps of:

loading said image on said computer screen;  
asking said voice mail system to initiate a call to  
a device and to enter a waiting state; wherein said voice  
10 mail system initiates said call and upon answering of  
said call, plays a prompt indicating reason for said call  
and asks user to perform a specific action;  
providing confirmation from said user;  
initiating said voice mail system to instruct  
15 annotation program of said computer means to do  
annotation while said user begins recording a voice  
message through said telephone means; and,  
instructing said voice mail system to send images,  
gestures and synchronization markers to a message store  
20 to be stored with said voice message.

14. A method of transferring dynamically annotated  
images as claimed in claim 12 wherein playing back said  
message comprises the steps of:  
25 loading said image on said computer screen;  
asking said voice mail system to initiate a call to  
a device and to enter a waiting state; wherein said voice  
mail system initiates said call, plays a prompt  
indicating reason for said call and asks a user to  
30 perform a specific action;  
providing confirmation from said user;  
sending said voice messages over said telephone  
means while said images and gestures are displayed on  
said screen of said computer means; and,  
35 synchronizing with said synchronization markers,  
said telephone means and said screen of said computer  
means.





WO 97/11549

THIS SHEET HAS BEEN  
REPLACED

PCT/US96/13217

13

15. A phone based dynamic image annotation system as claimed in claim 14 wherein playing back said messages further comprises the steps of:

initiating said computer means to request a list of  
5 messages from said voice mail system;

providing said list of messages from said voice mail system;

having said computer means select one or more messages from said list of messages; and,

10 transmitting said one or more messages from said voice mail system.

16. A method of transferring dynamically annotated images comprising the steps of:

15 utilizing computer means comprising a computer screen;

connecting modem means to said computer means;

connecting telephone means to said modem means;

20 connecting a public telephone network to said modem means;

connecting PBX means to said public telephone network;

connecting a voice mail system to said PBX means; and,

25 utilizing protocol means for recording and playing back a message;

wherein during said recording, a voice component of said message is decoupled so that said telephone means can carry said voice component of said message and

30 wherein during said playing back, said voice component played over said telephone means is synchronized with a graphics and gesture component which is drawn overlapping an image of said message displayed on said computer screen.

35

This Sheet Has Been Replaced





17. A method of transferring dynamically annotated images as claimed in claim 16 wherein utilizing said protocol means comprises the steps of:

5 recording said message; and,  
playing back said message.

18. A method of transferring dynamically annotated images as claimed in claim 17 wherein recording said message comprises the steps of:

10 loading said image on said computer screen;  
asking said voice mail system to initiate a call to a device and to enter a waiting state; wherein said voice mail system initiates said call and upon answering of said call, plays a prompt indicating reason for said call  
15 and asks user to perform a specific action;  
providing confirmation from said user;  
initiating said voice mail system to instruct annotation program of said computer means to do annotation while said user begins recording a voice  
20 message through said telephone means; and,  
instructing said voice mail system to send images, gestures and synchronization markers to a message store to be stored with said voice message.

25 19. A method of transferring dynamically annotated images as claimed in claim 18 wherein playing back said message comprises the steps of:

loading said image on said computer screen;  
asking said voice mail system to initiate a call to  
30 a device and to enter a waiting state; wherein said voice mail system initiates said call, plays a prompt indicating reason for said call and asks user to perform a specific action;  
providing confirmation from said user;  
35 sending said voice messages over said telephone means while said images and gestures are displayed on screen of said computer means; and,





WO 97/11549

PCT/US96/13217

15

synchronizing with said synchronization markers,  
said telephone means and said screen of said computer  
means.

- 5 20. A phone based dynamic image annotation system as  
claimed in claim 19 wherein playing back said messages  
further comprises the steps of:
- initiating said computer means to request a list of  
messages from said voice mail system;
- 10 providing said list of messages from said voice mail  
system;
- having said computer means select one or more  
messages from said list of messages; and,
- transmitting said one or more messages from said
- 15 voice mail system.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN  
REPLACED