



US 20040095376A1

(19) **United States**

(12) **Patent Application Publication**  
**Graham et al.**

(10) **Pub. No.: US 2004/0095376 A1**

(43) **Pub. Date: May 20, 2004**

(54) **TECHNIQUES FOR DISPLAYING  
INFORMATION STORED IN MULTIPLE  
MULTIMEDIA DOCUMENTS**

(75) Inventors: **Jamey Graham**, San Jose, CA (US);  
**Jonathan J. Hull**, San Carlos, CA  
(US); **Gregory J. Wolff**, Redwood City,  
CA (US)

Correspondence Address:

**TOWNSEND AND TOWNSEND AND CREW,  
LLP  
TWO EMBARCADERO CENTER  
EIGHTH FLOOR  
SAN FRANCISCO, CA 94111-3834 (US)**

(73) Assignee: **Ricoh Company, Ltd.**, Tokyo (JP)

(21) Appl. No.: **10/465,022**

(22) Filed: **Jun. 18, 2003**

**Related U.S. Application Data**

(63) Continuation-in-part of application No. 10/081,129,  
filed on Feb. 21, 2002.

Continuation-in-part of application No. 10/174,522,  
filed on Jun. 17, 2002.

(60) Provisional application No. 60/434,314, filed on Dec.  
17, 2002.

**Publication Classification**

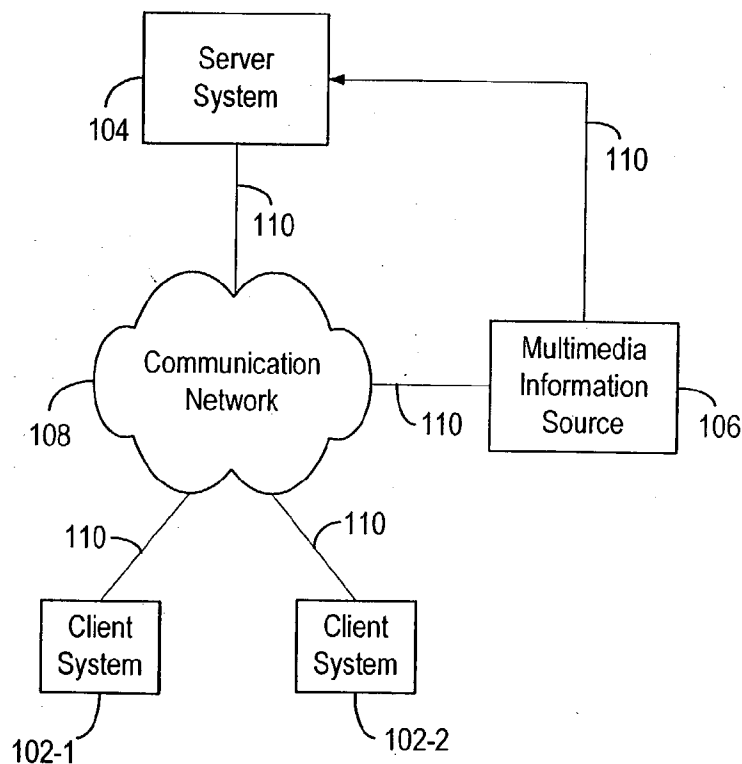
(51) **Int. Cl.<sup>7</sup> ..... G09G 5/00**

(52) **U.S. Cl. .... 345/716**

(57) **ABSTRACT**

Techniques for providing a graphical user interface (GUI) that displays a representation of stored information that may include information of one or more types. The displayed representation may include representations of information of the one or more types. The GUI enables a user to navigate and skim through the stored information and to analyze the contents of the stored information. The stored information may include information captured along the same timeline or along different timelines.

100



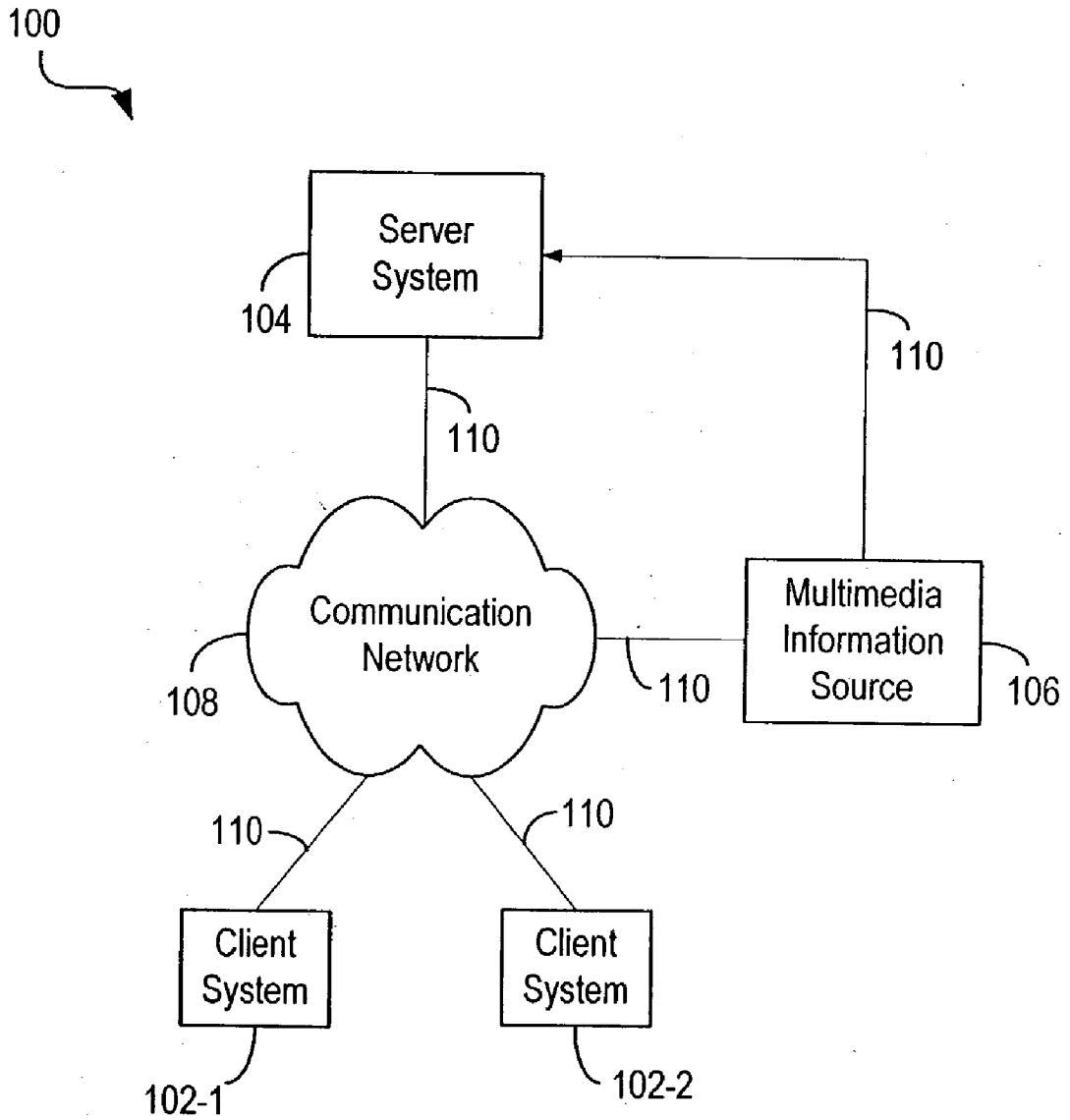


**Fig. 1**

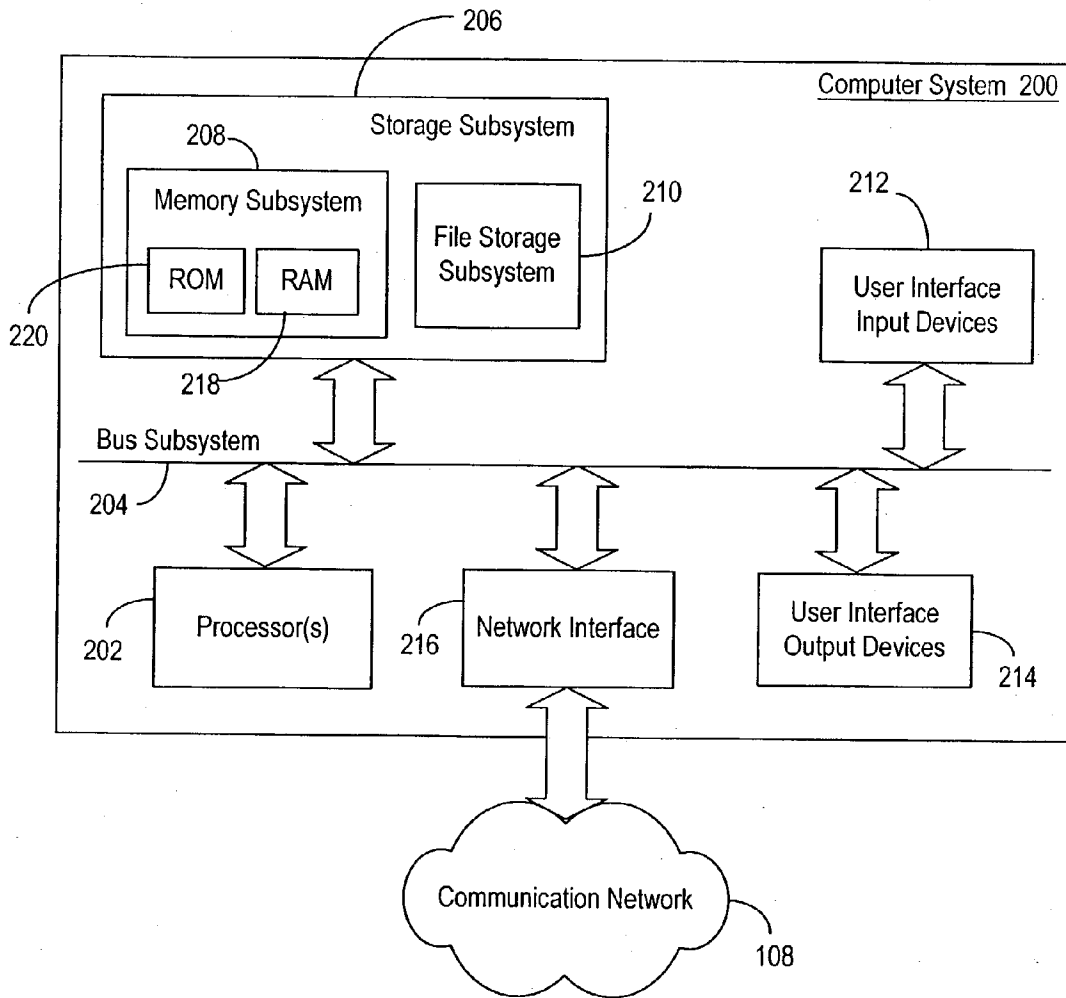


**Fig. 2**

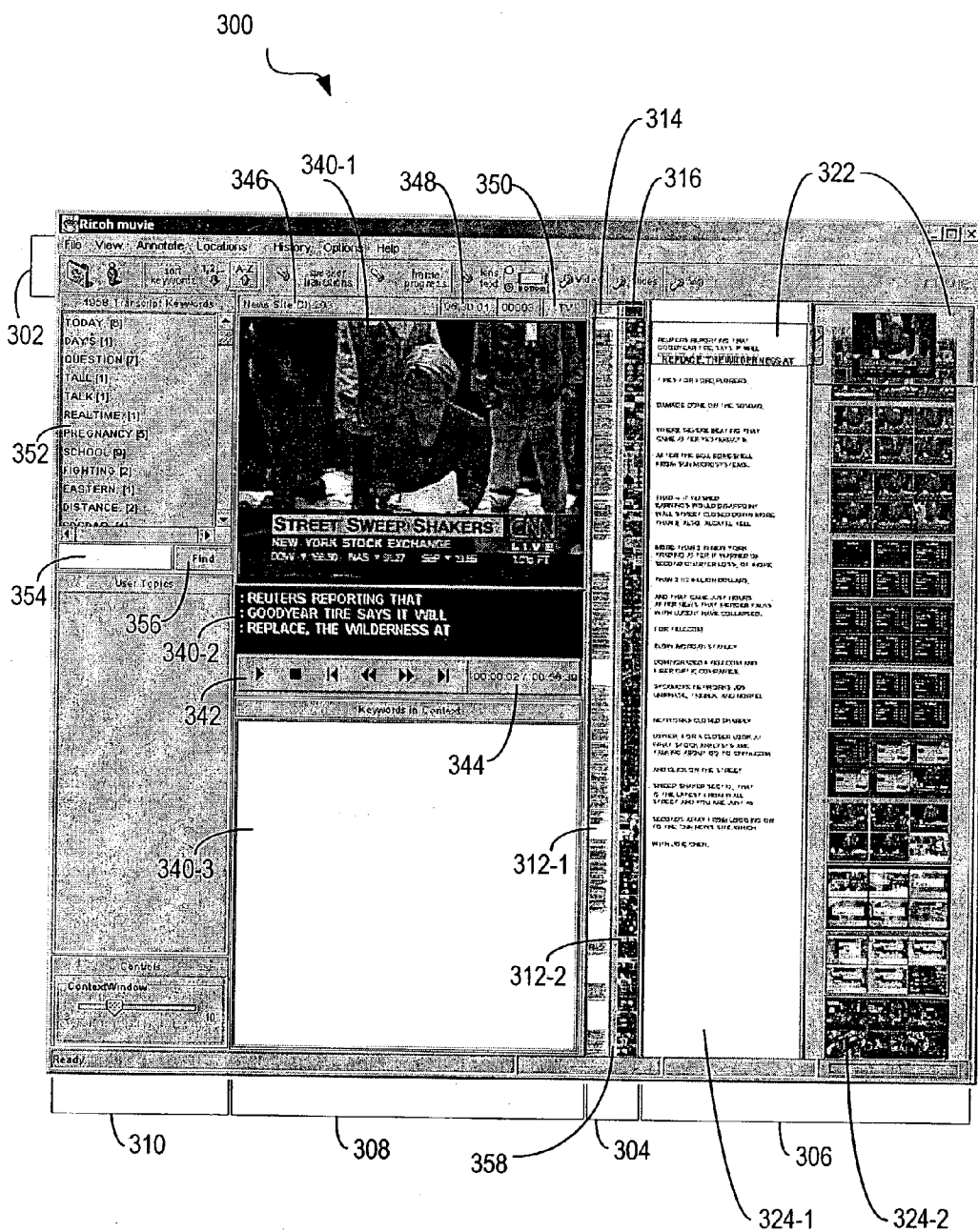


Fig. 3

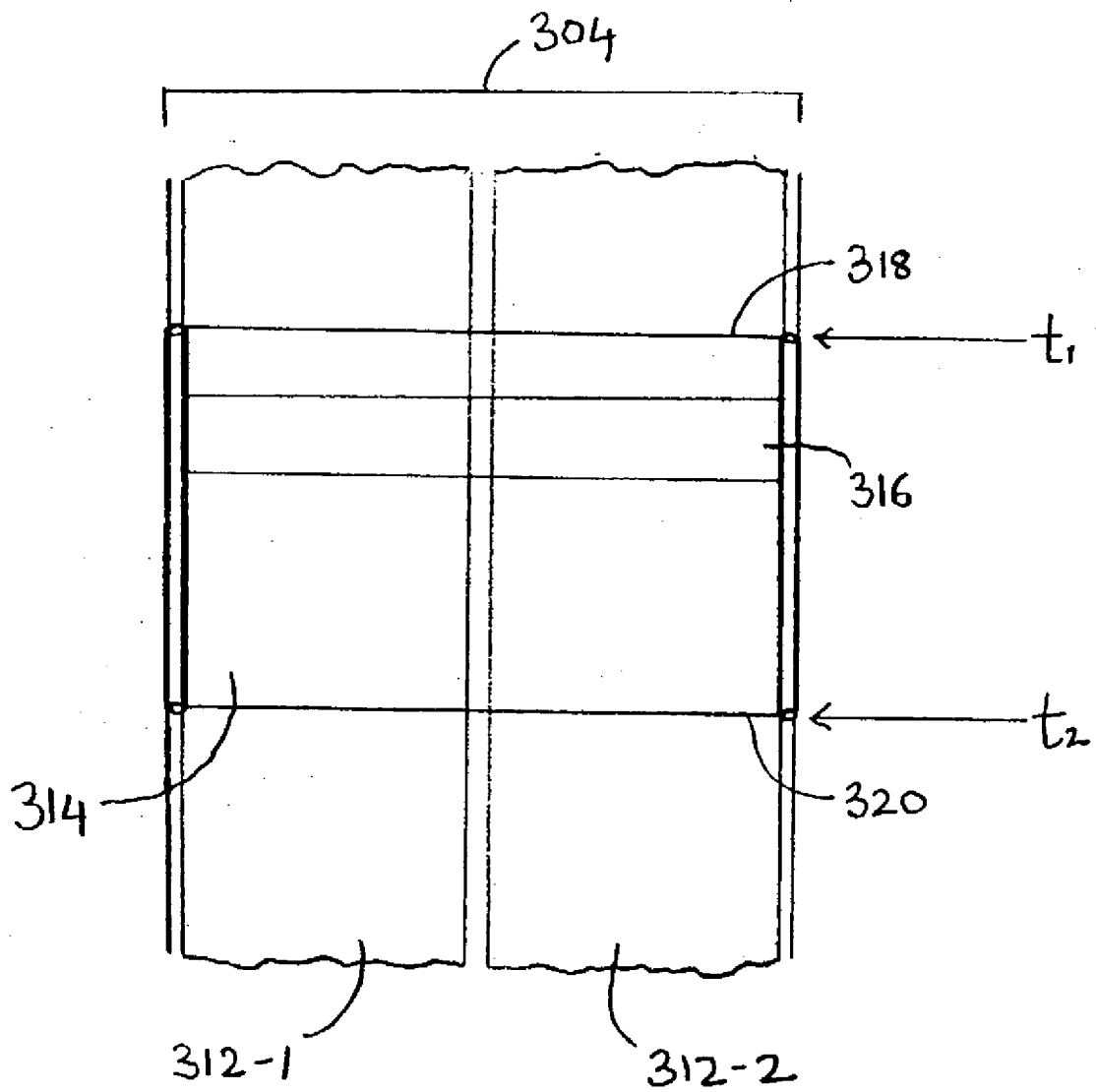


Fig. 4

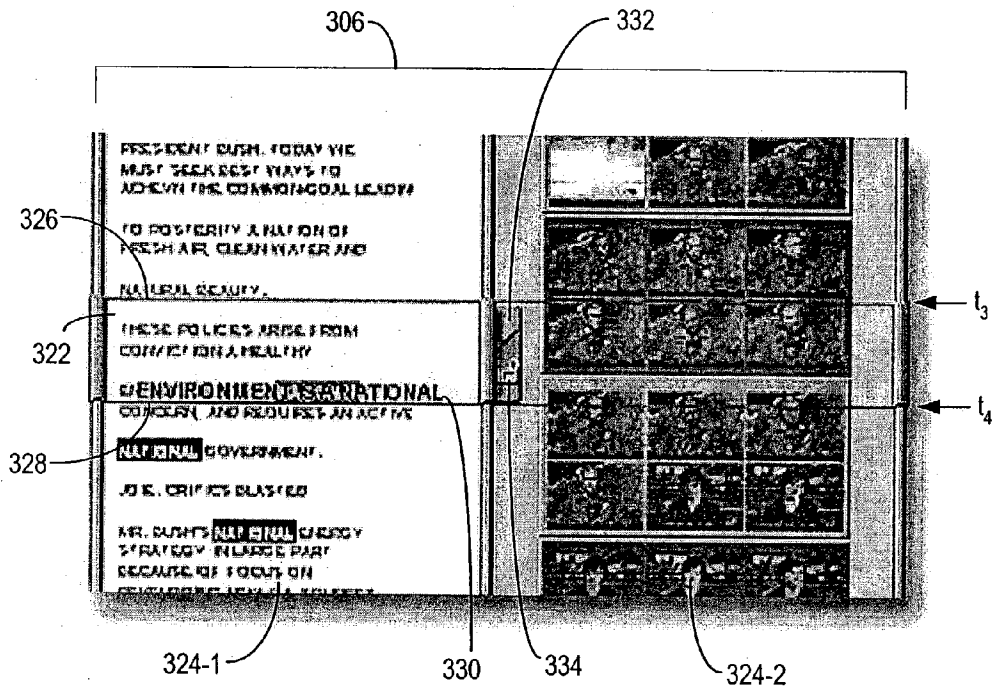


Fig. 5A

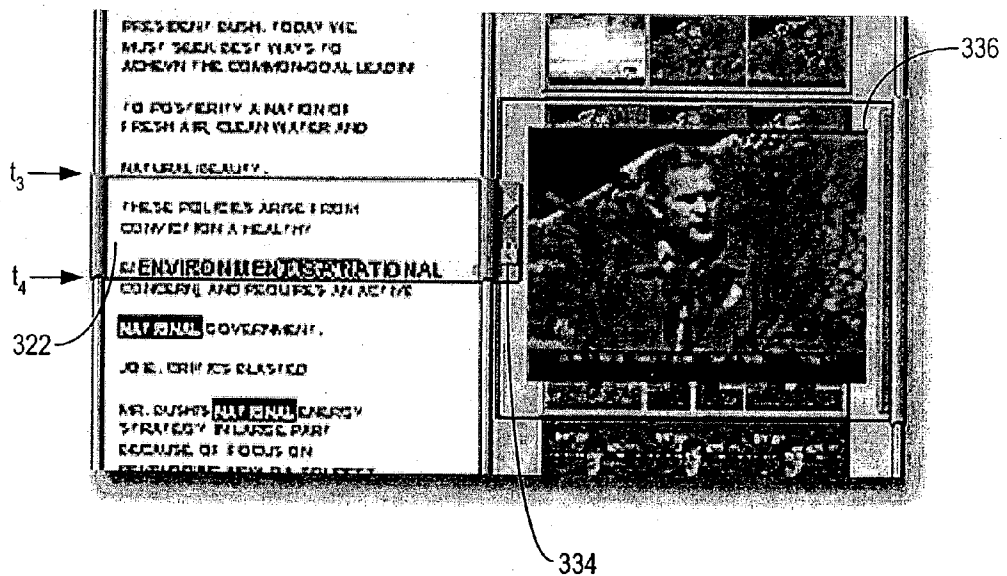


Fig. 5B

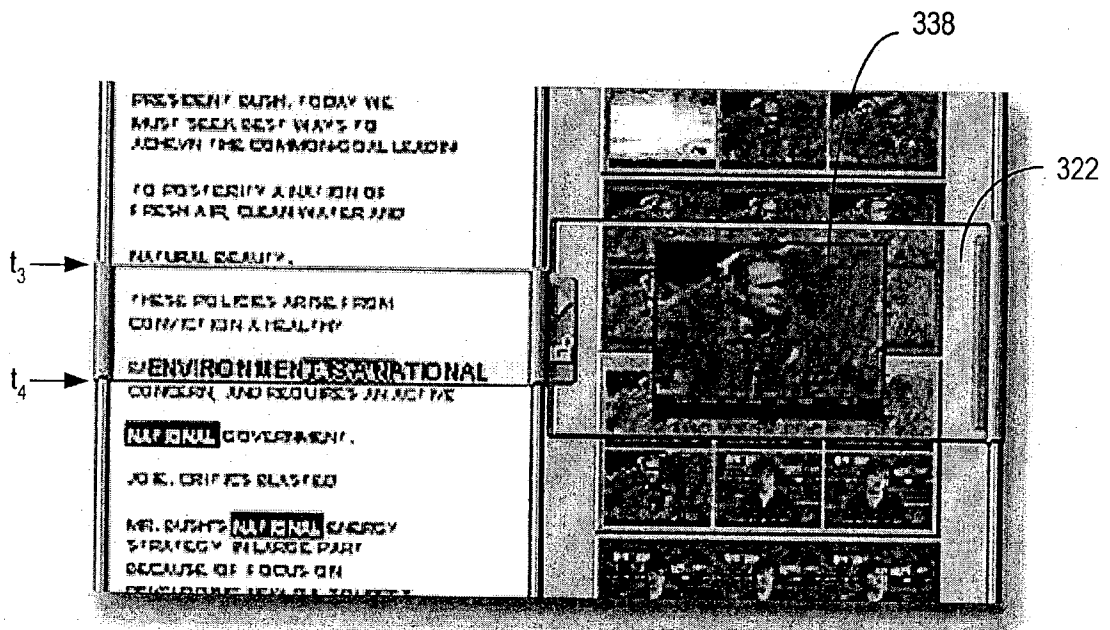


Fig. 5C

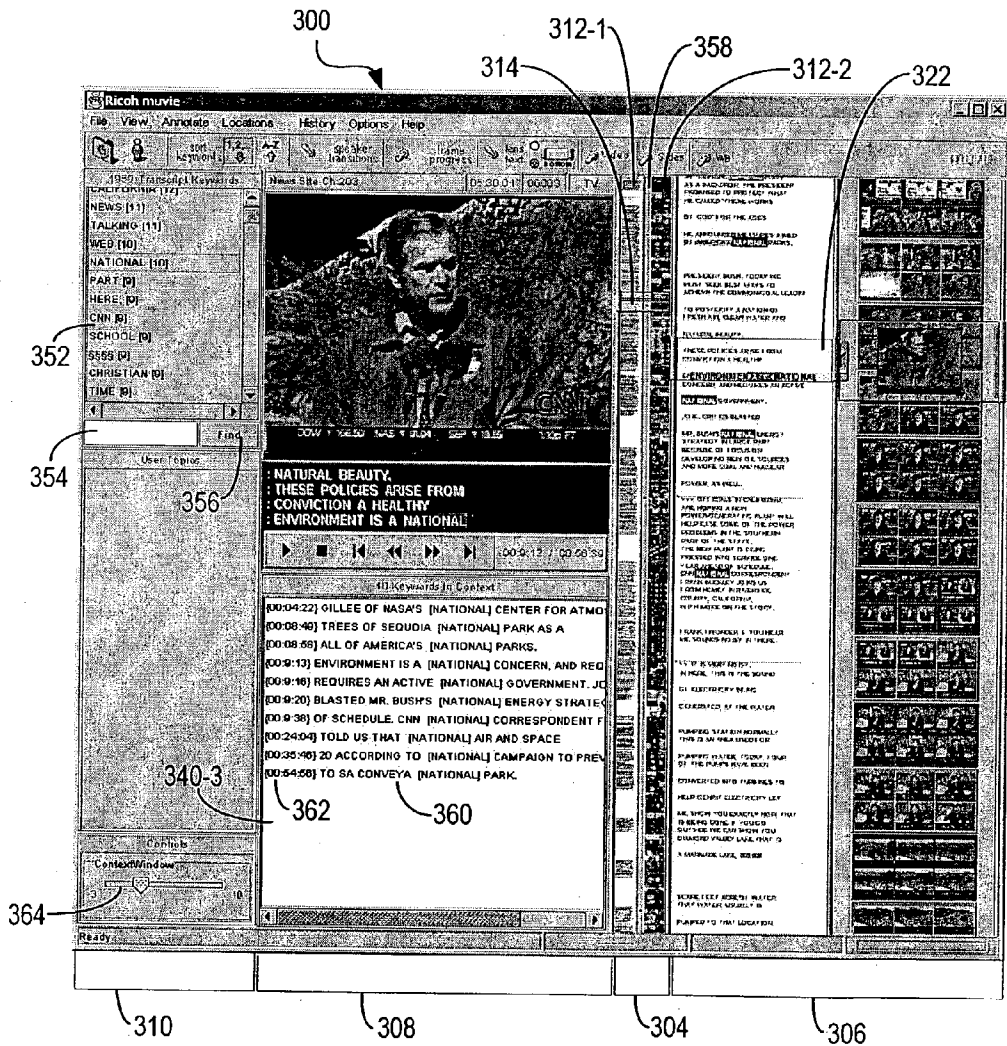


Fig. 6

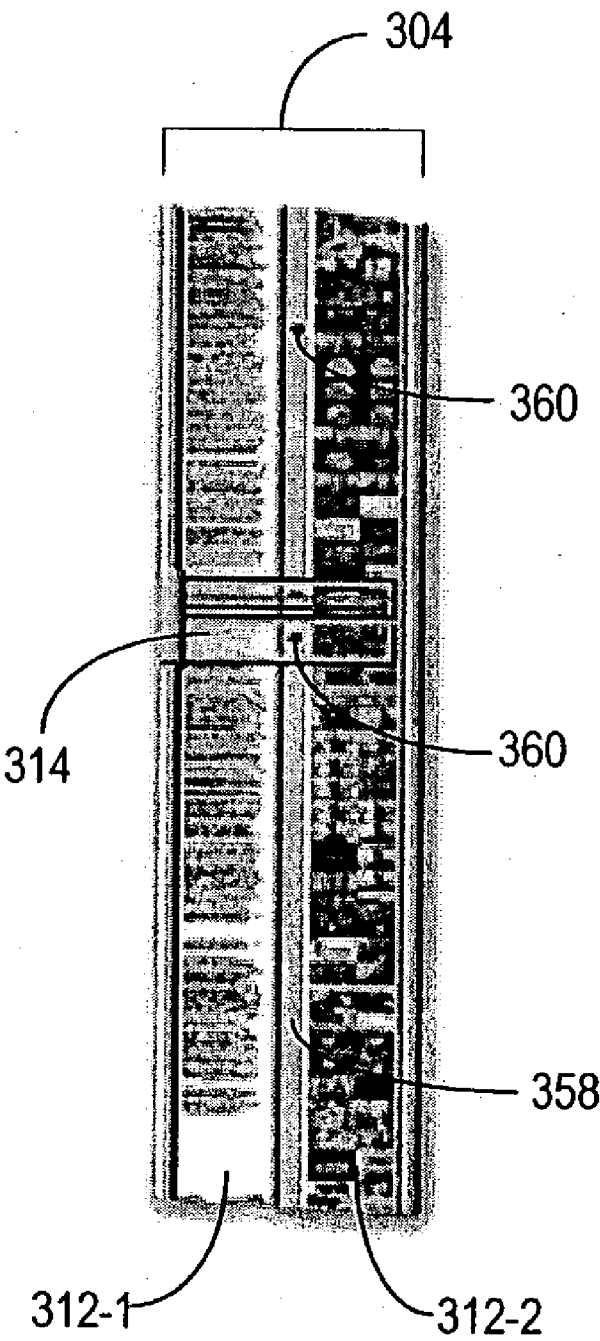


**Fig. 7**

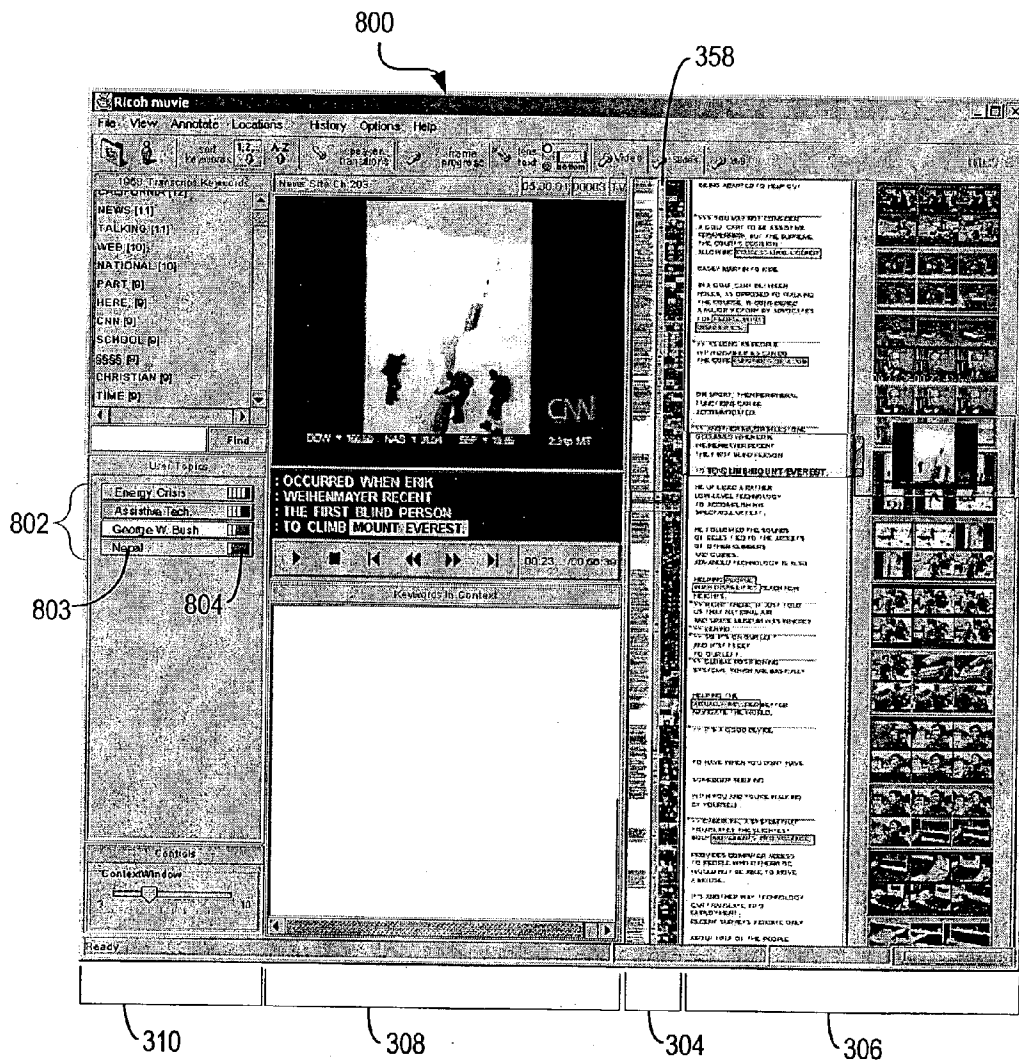


Fig. 8

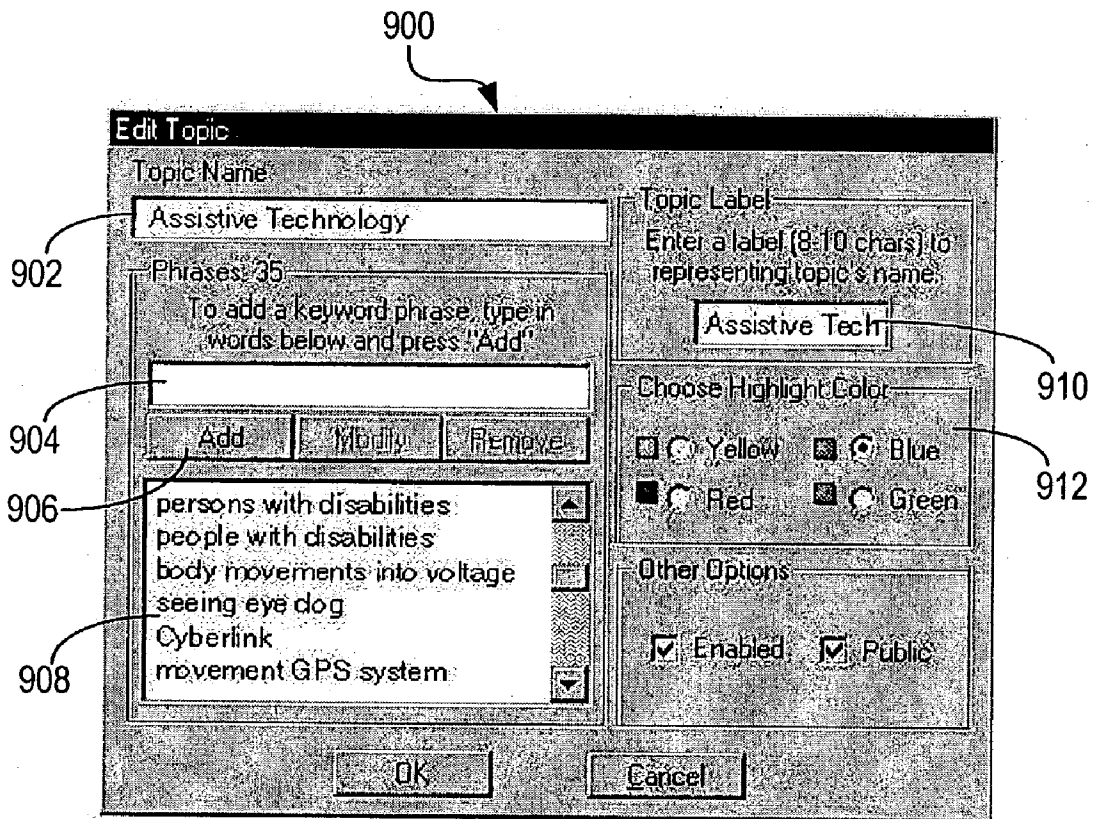


**Fig. 9**

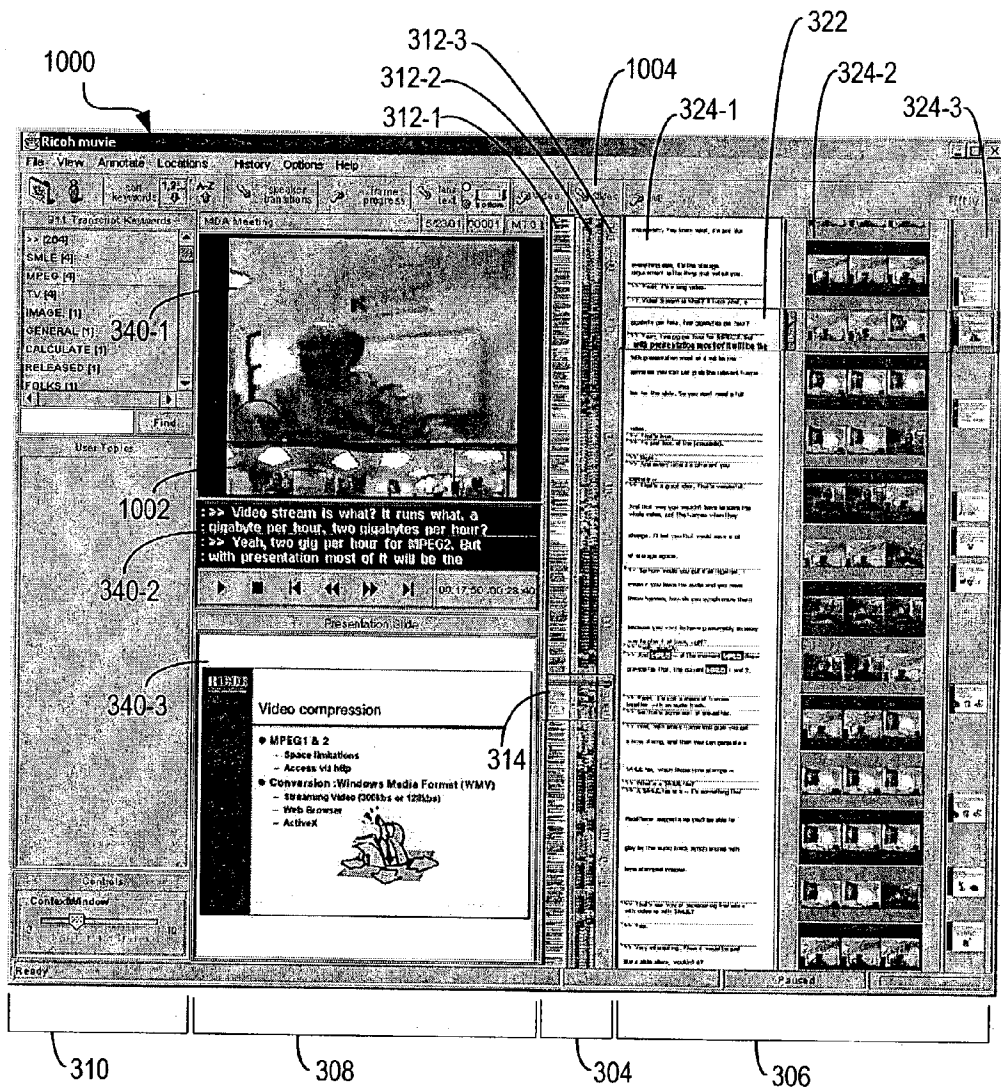


Fig. 10

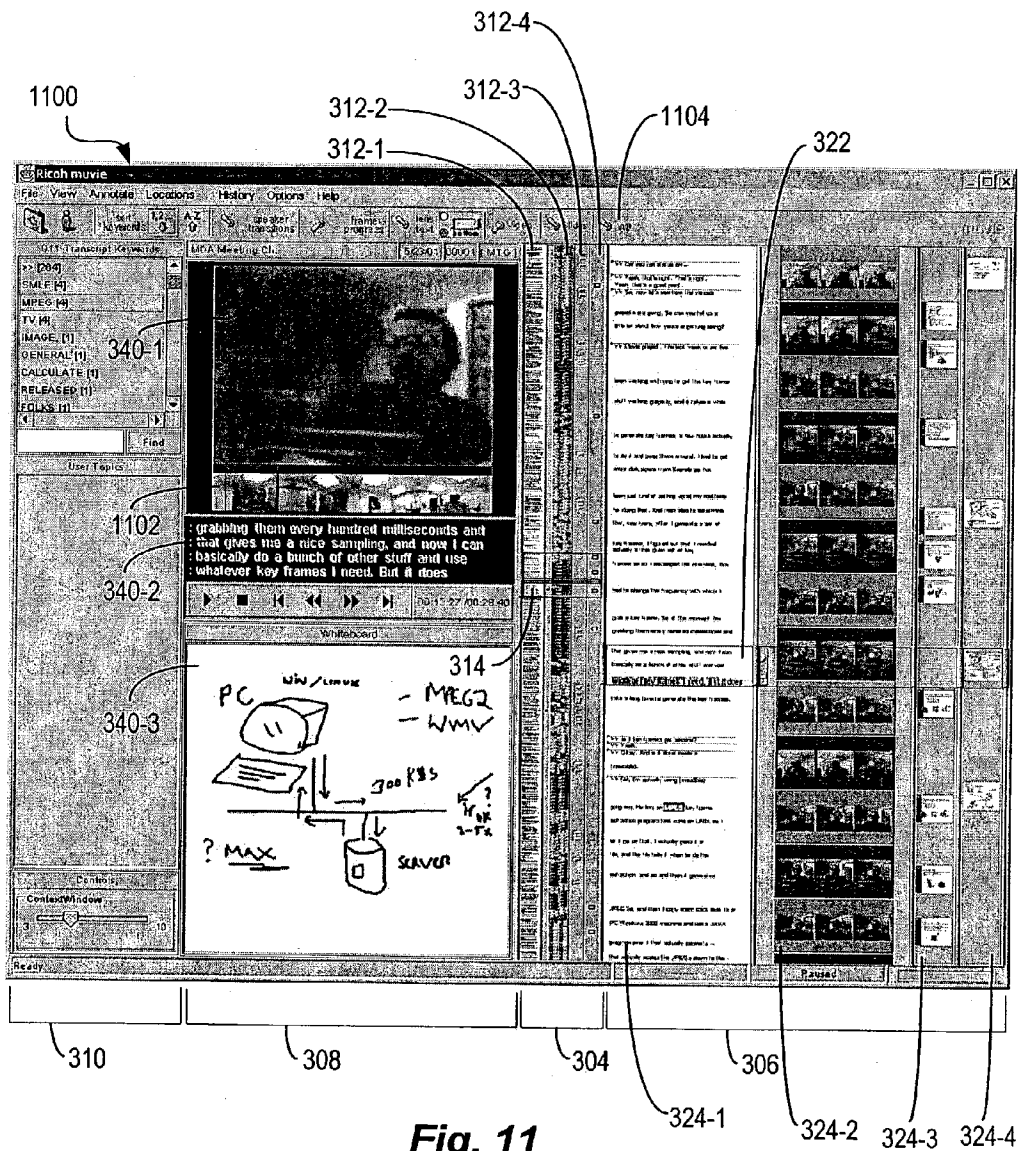


Fig. 11

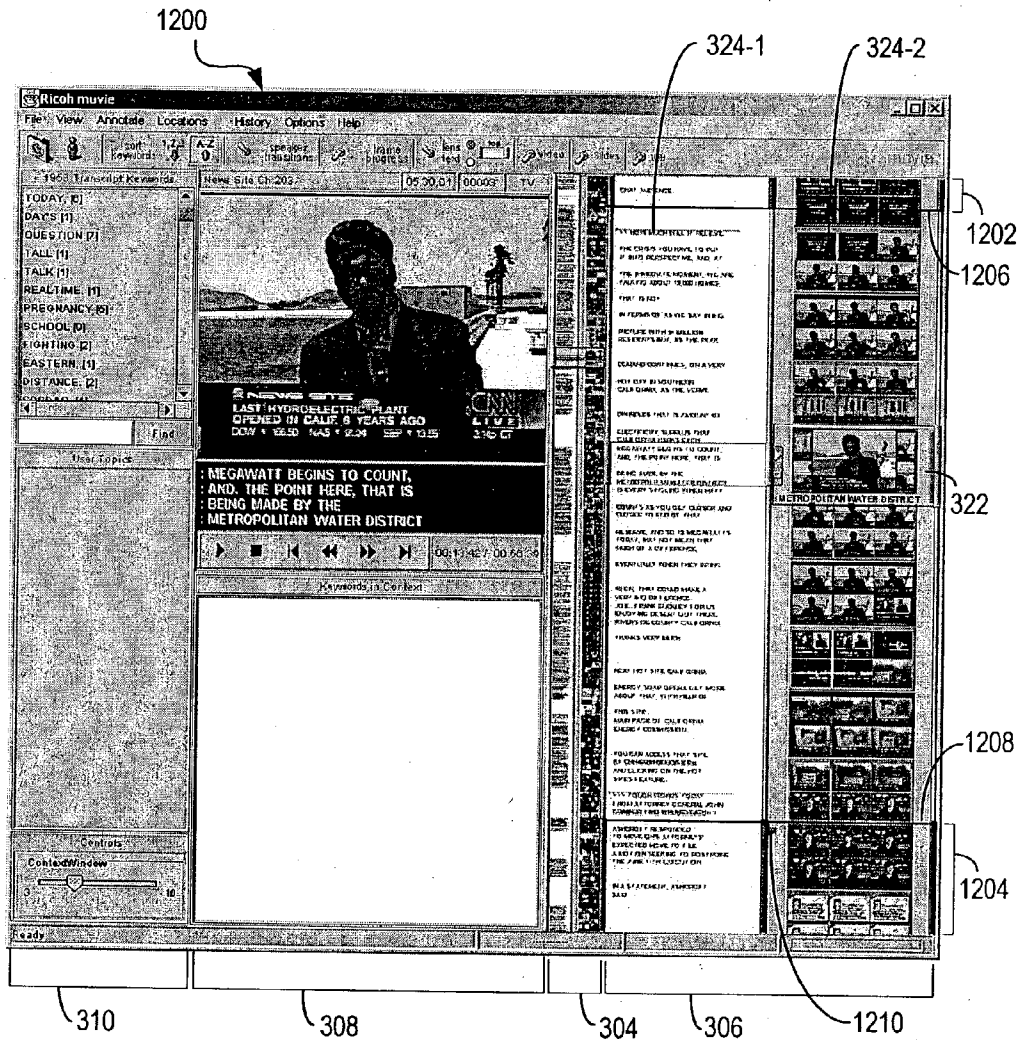


Fig. 12

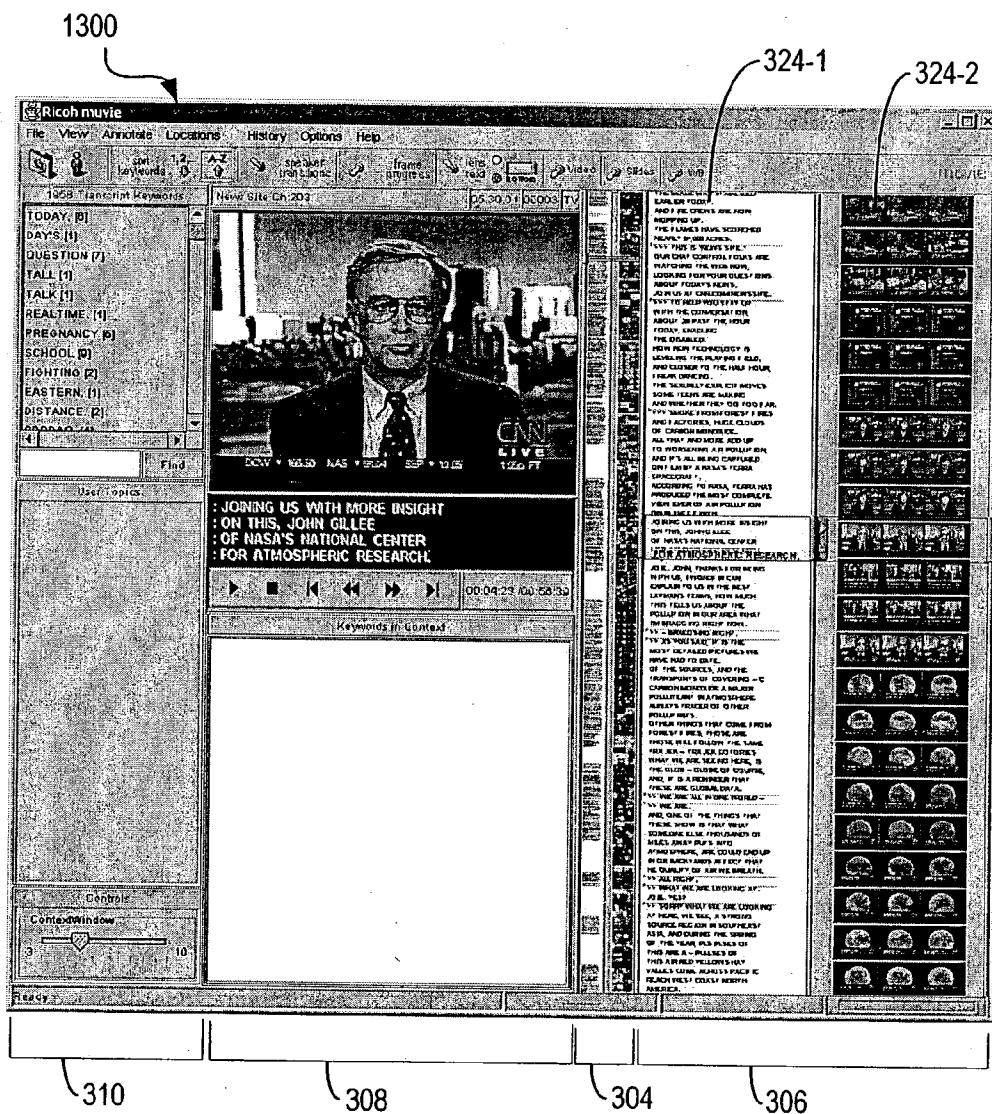


Fig. 13

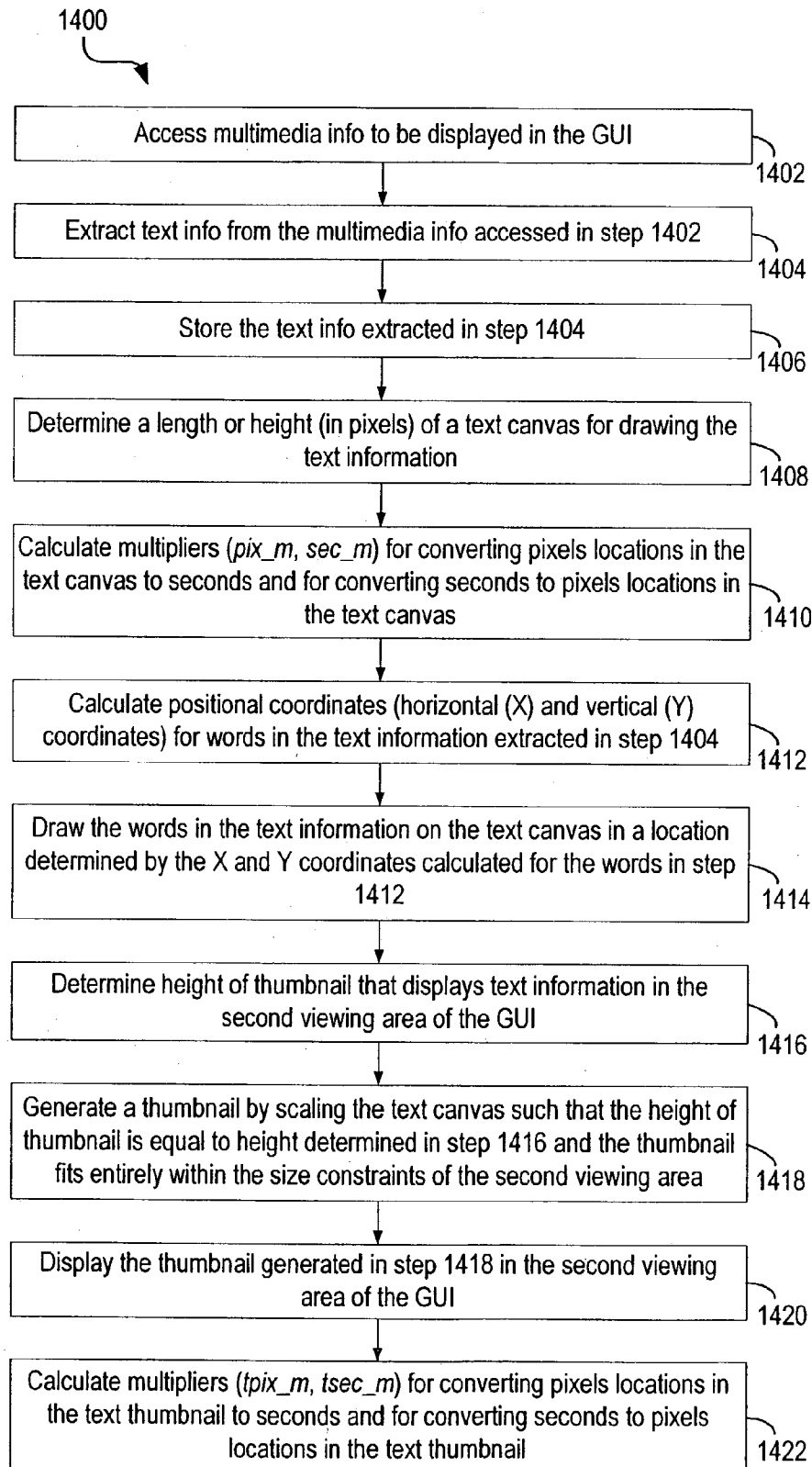


**Fig. 14**

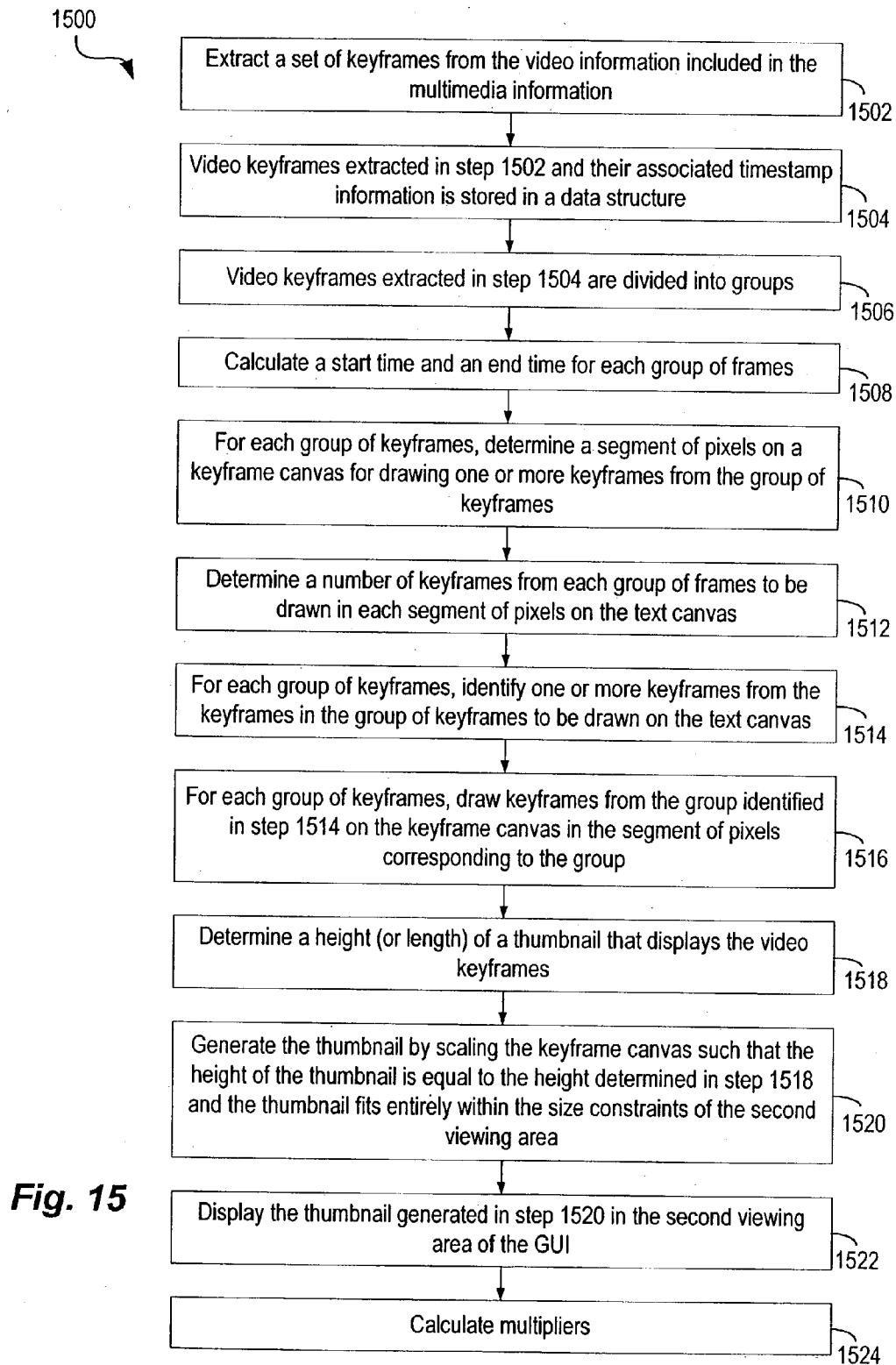


**Fig. 15**

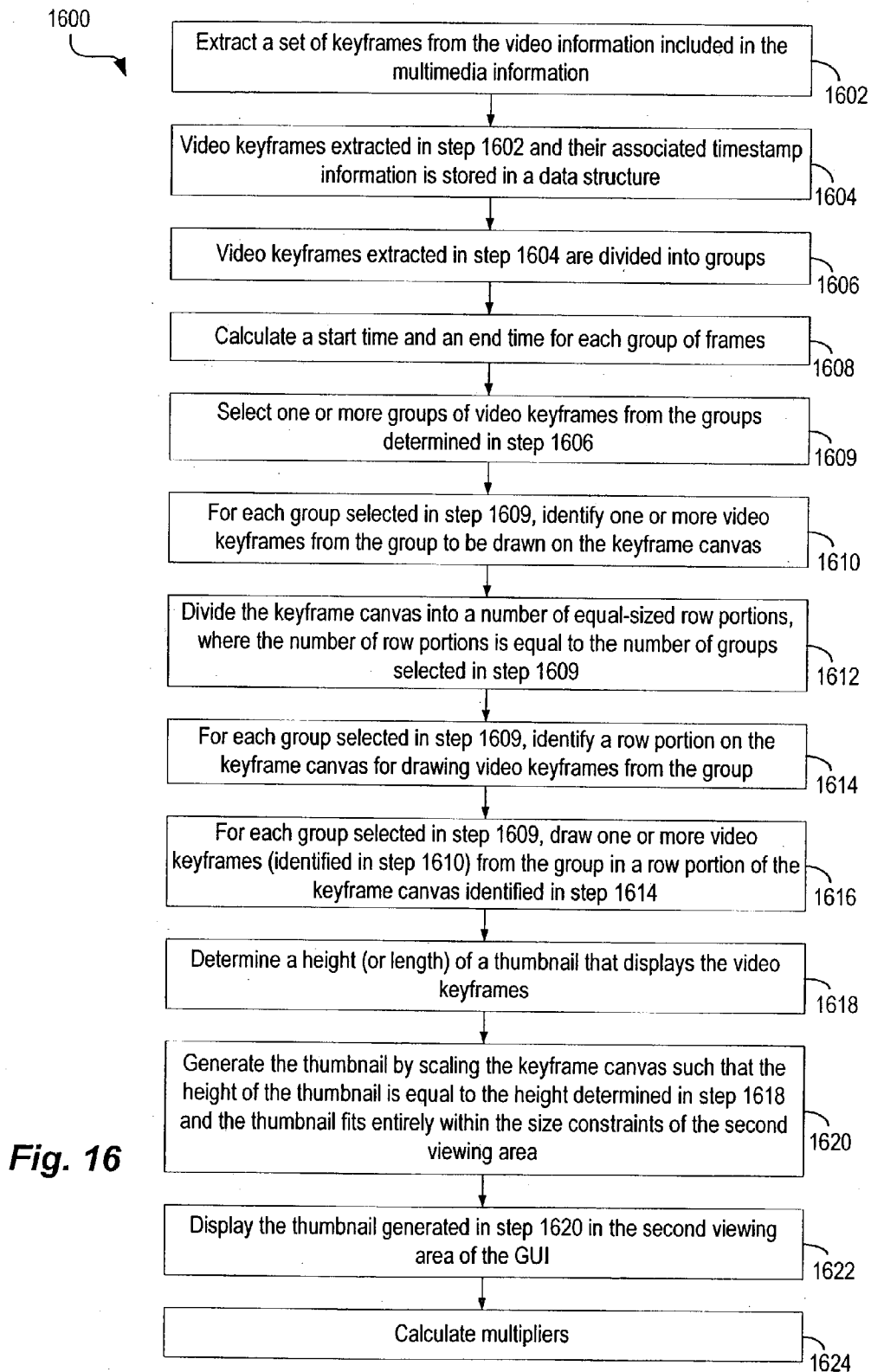


**Fig. 16**

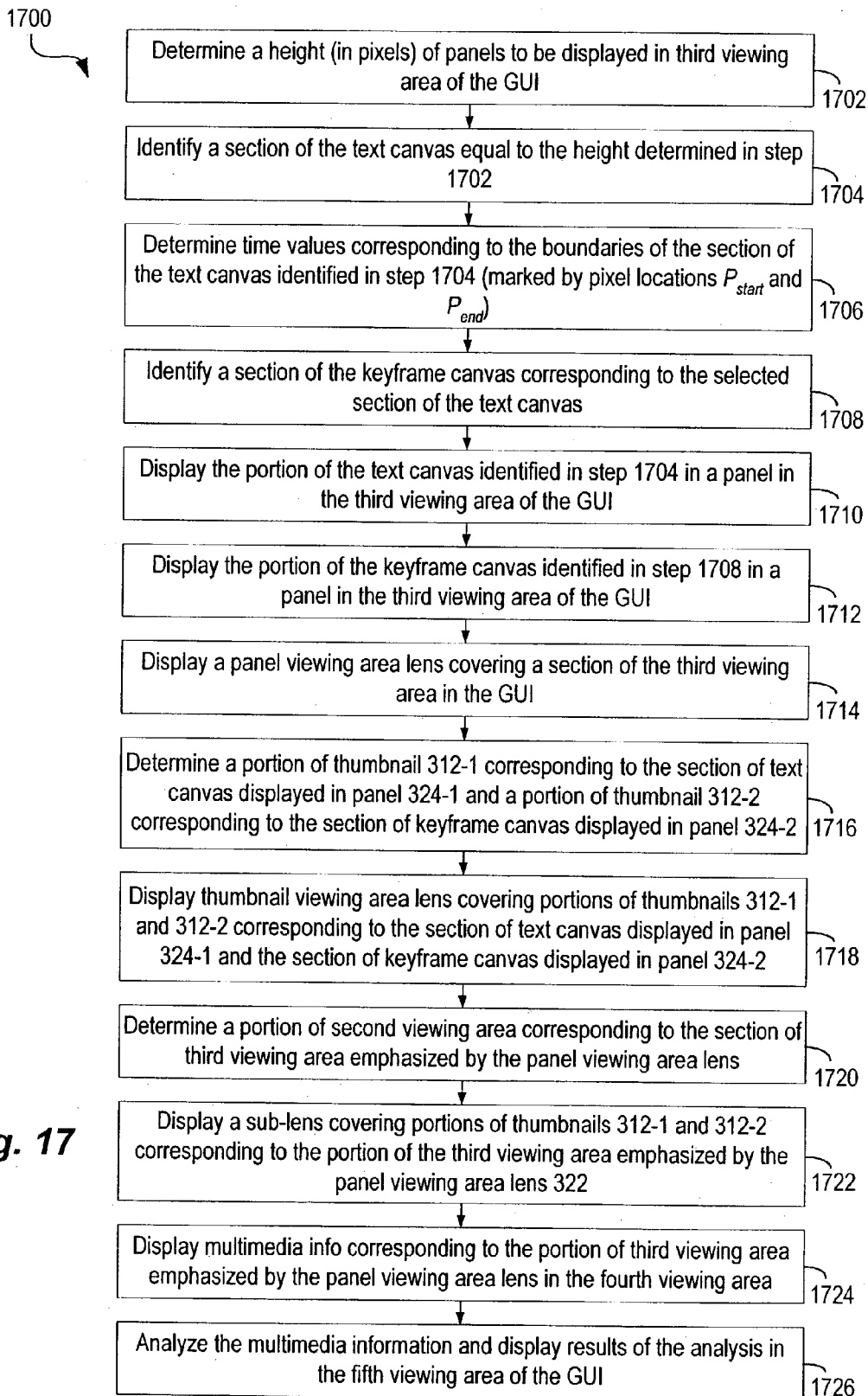


**Fig. 17**

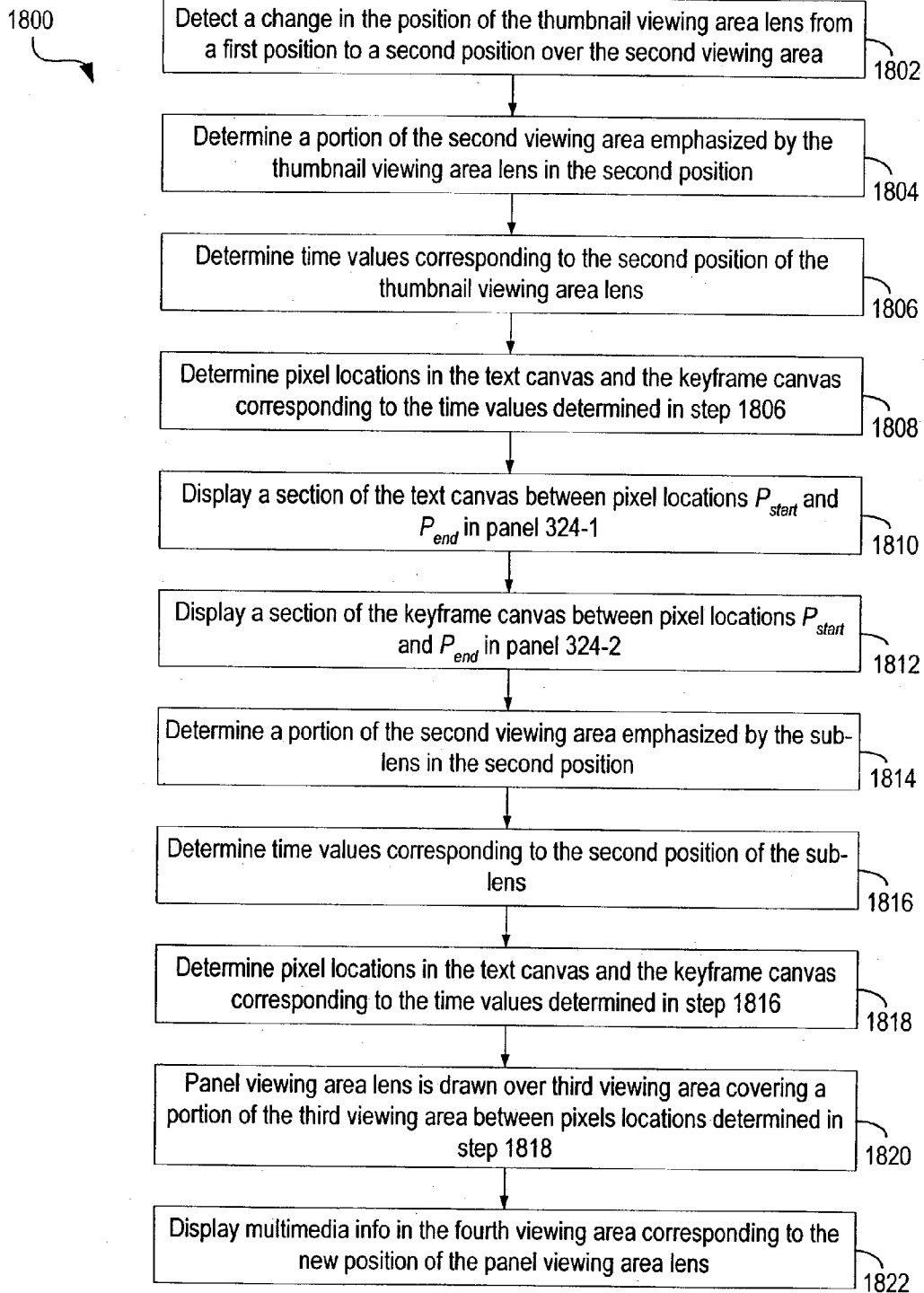


**Fig. 18**

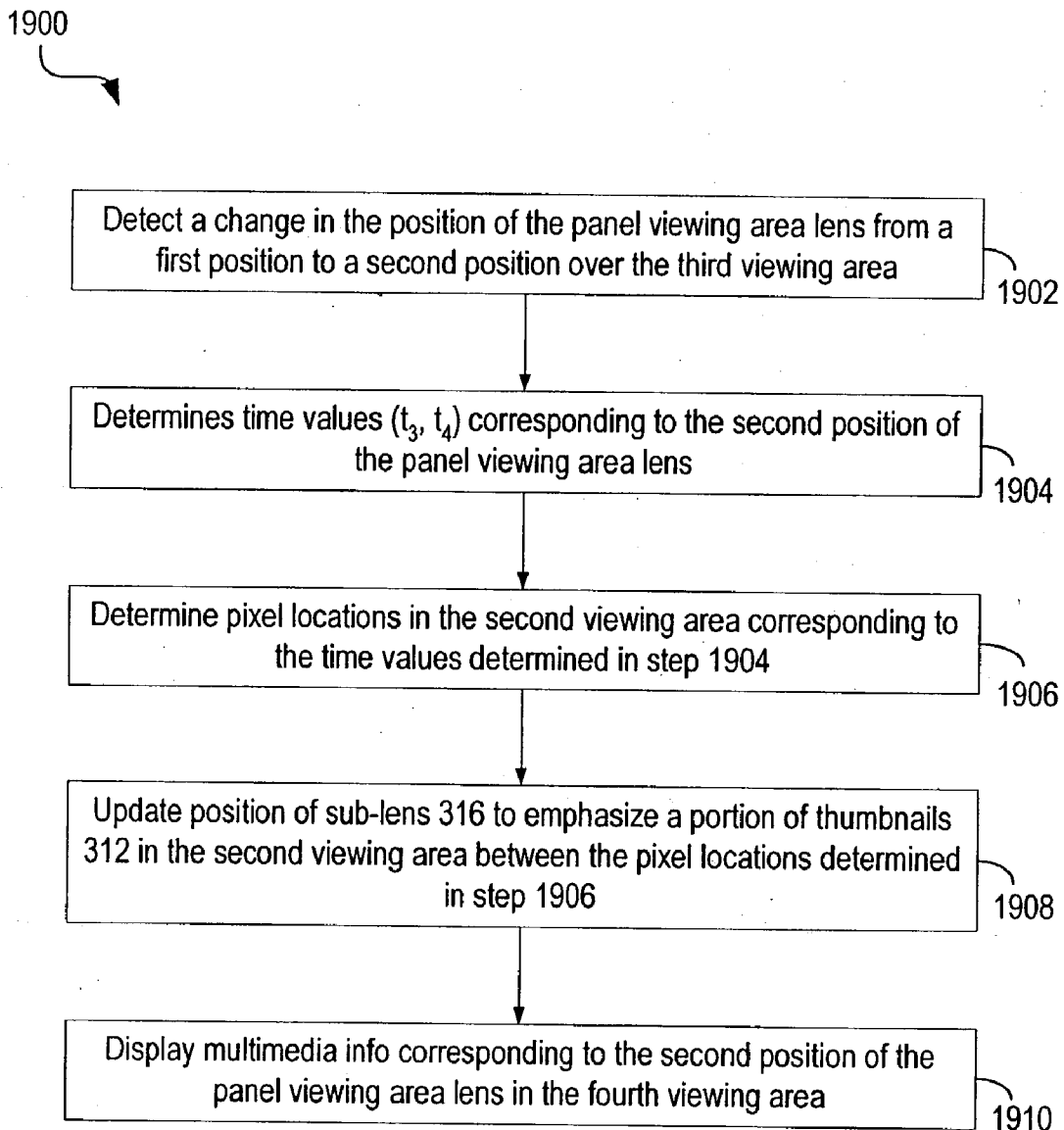


**Fig. 19**

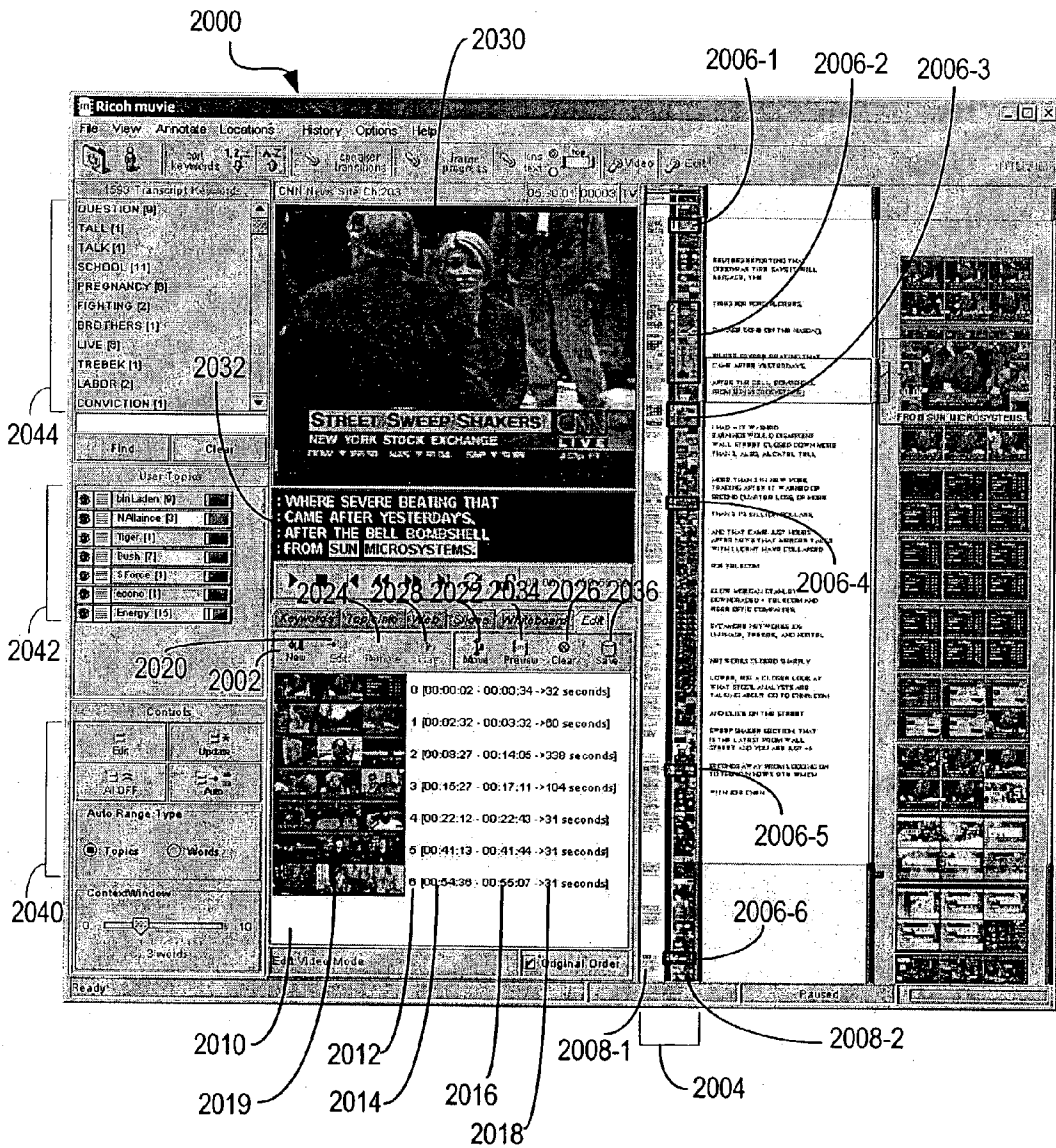


Fig. 20A

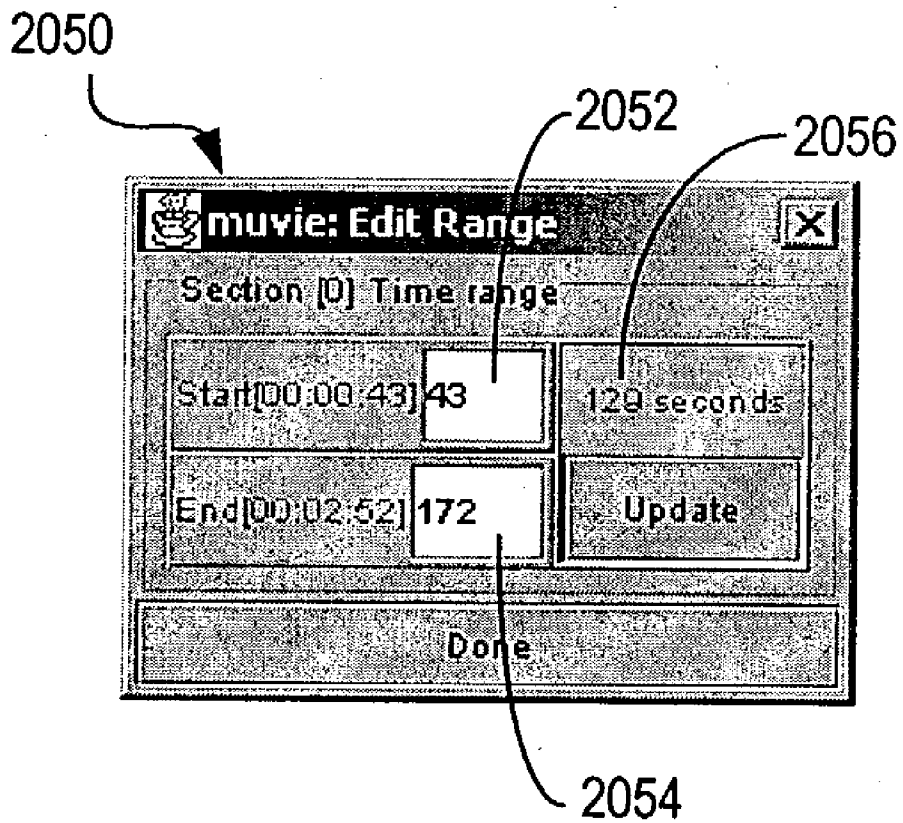


**Fig. 20B**

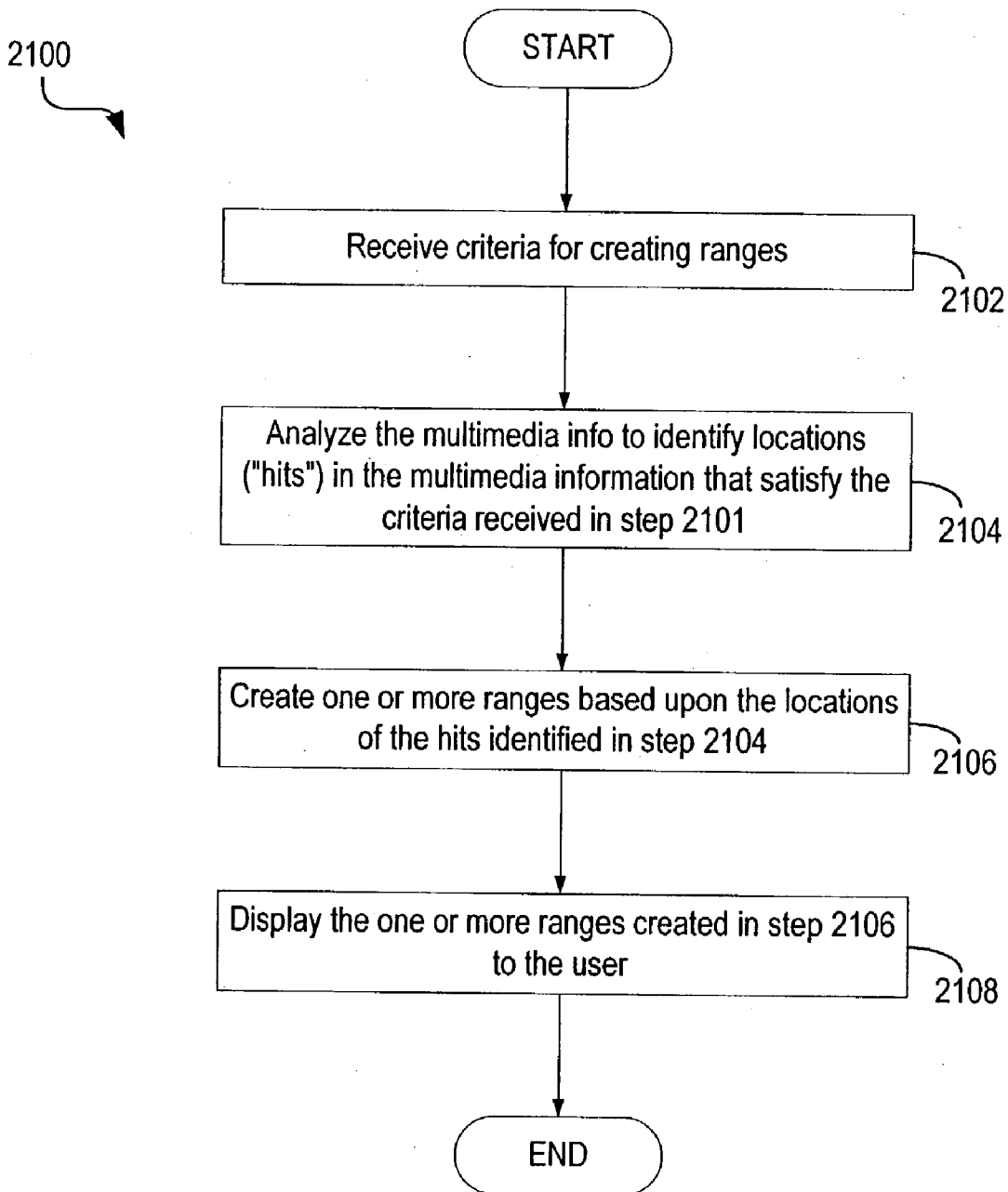


**Fig. 21**

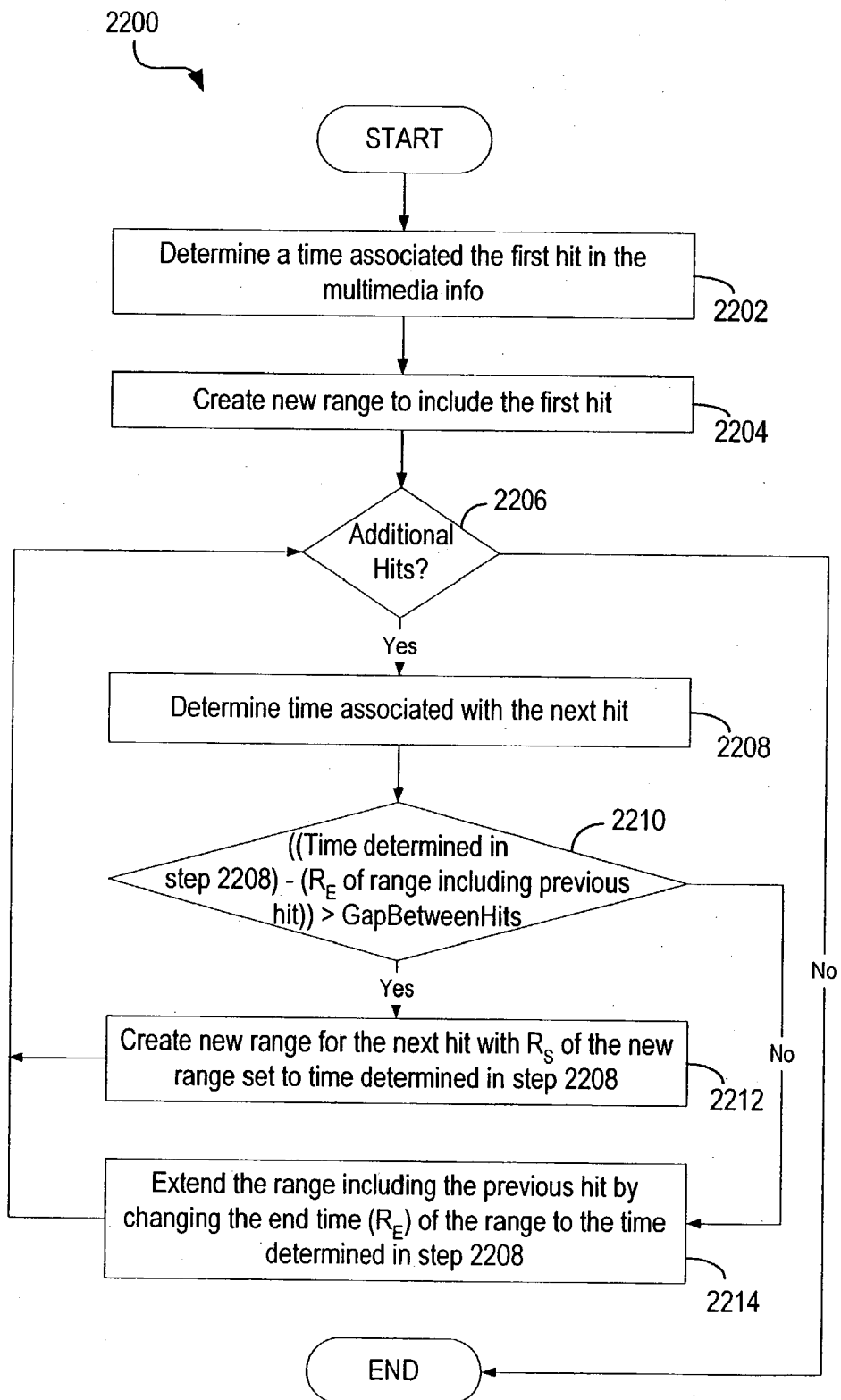
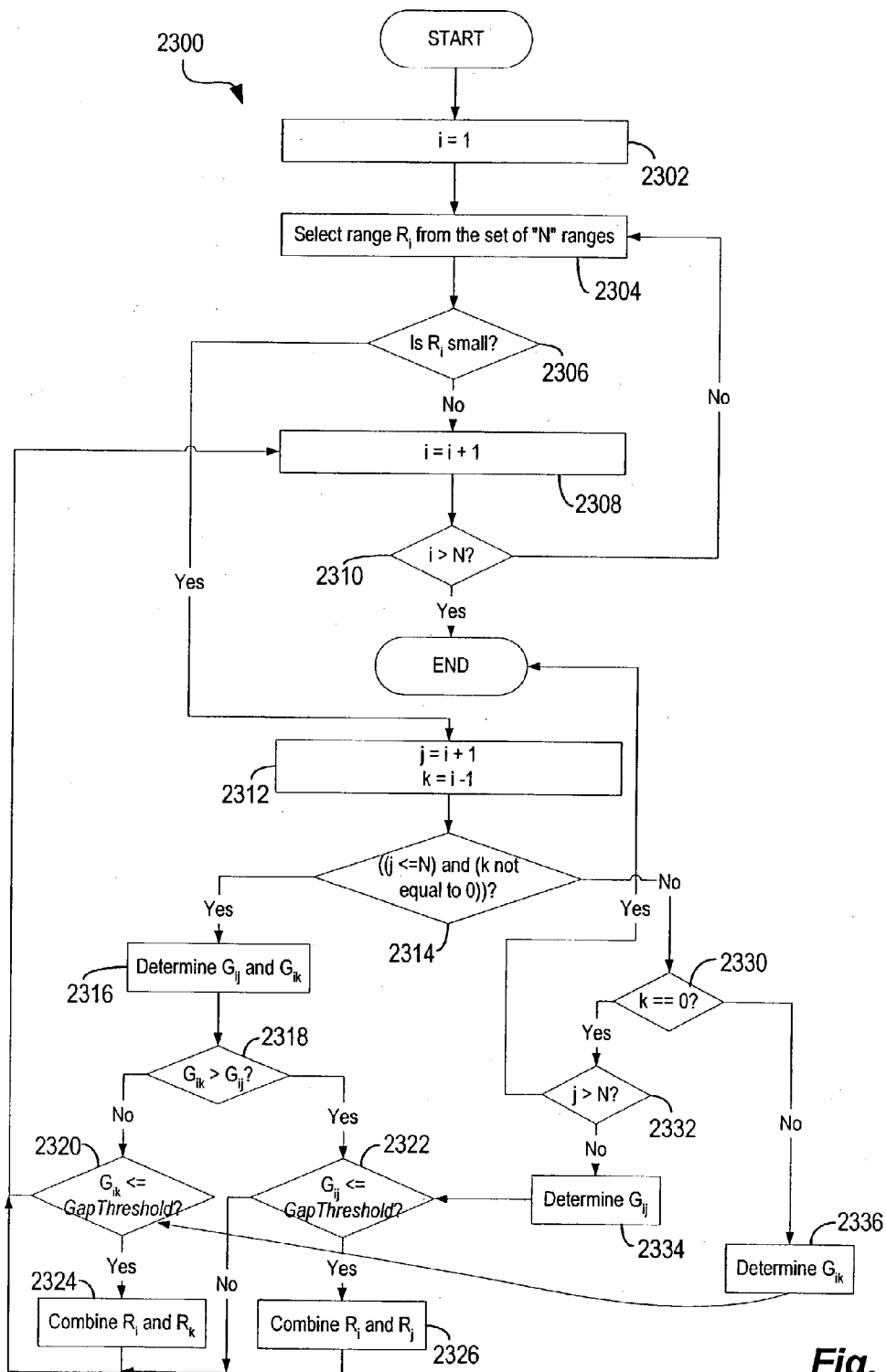


**Fig. 22**



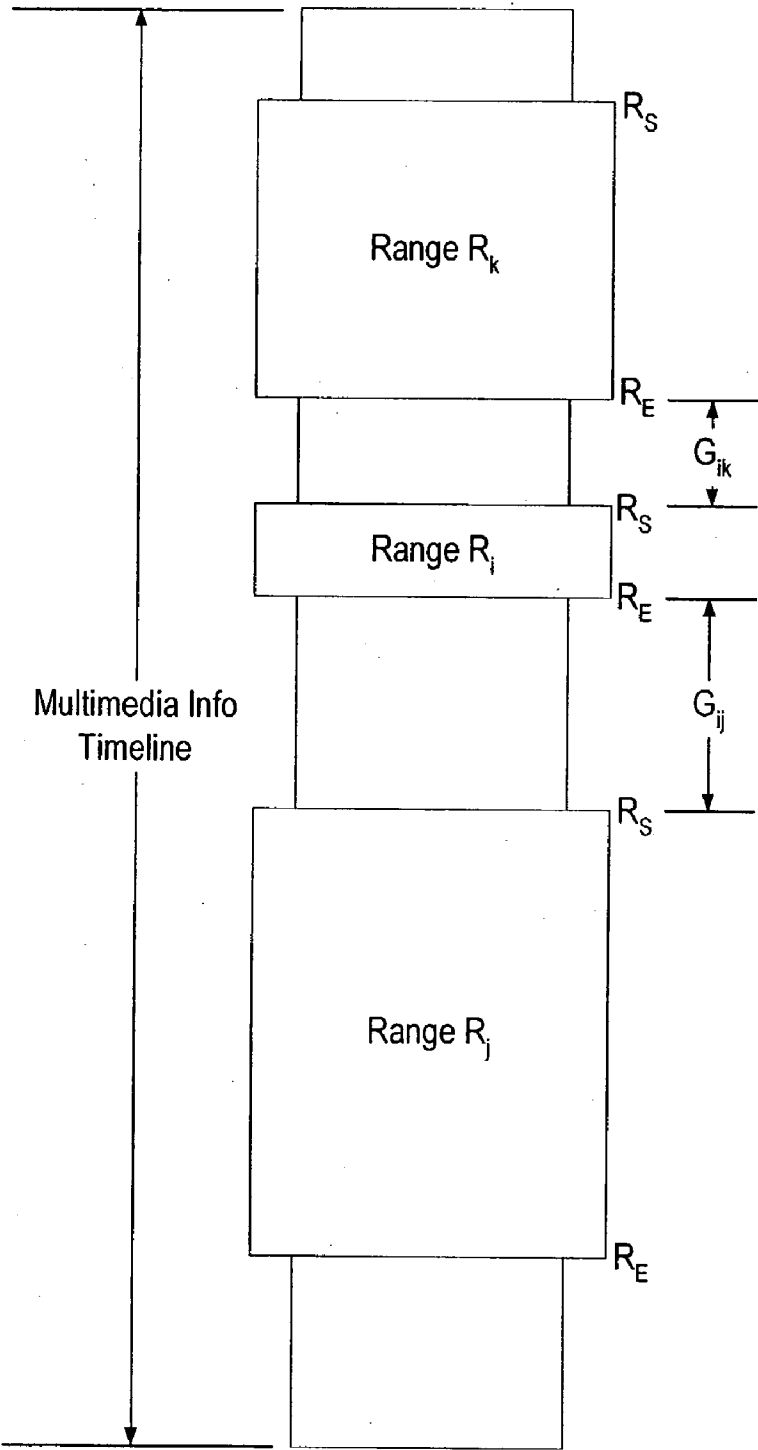


**Fig. 24**

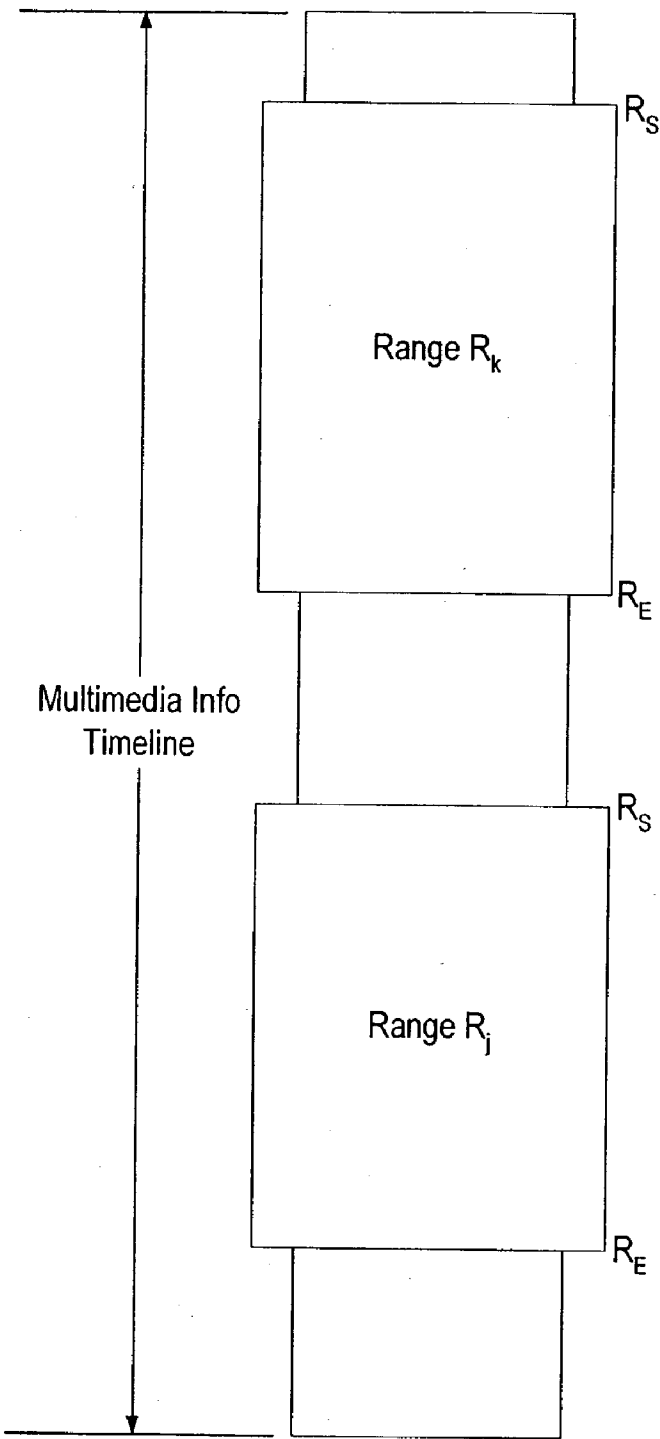


**Fig. 25A**

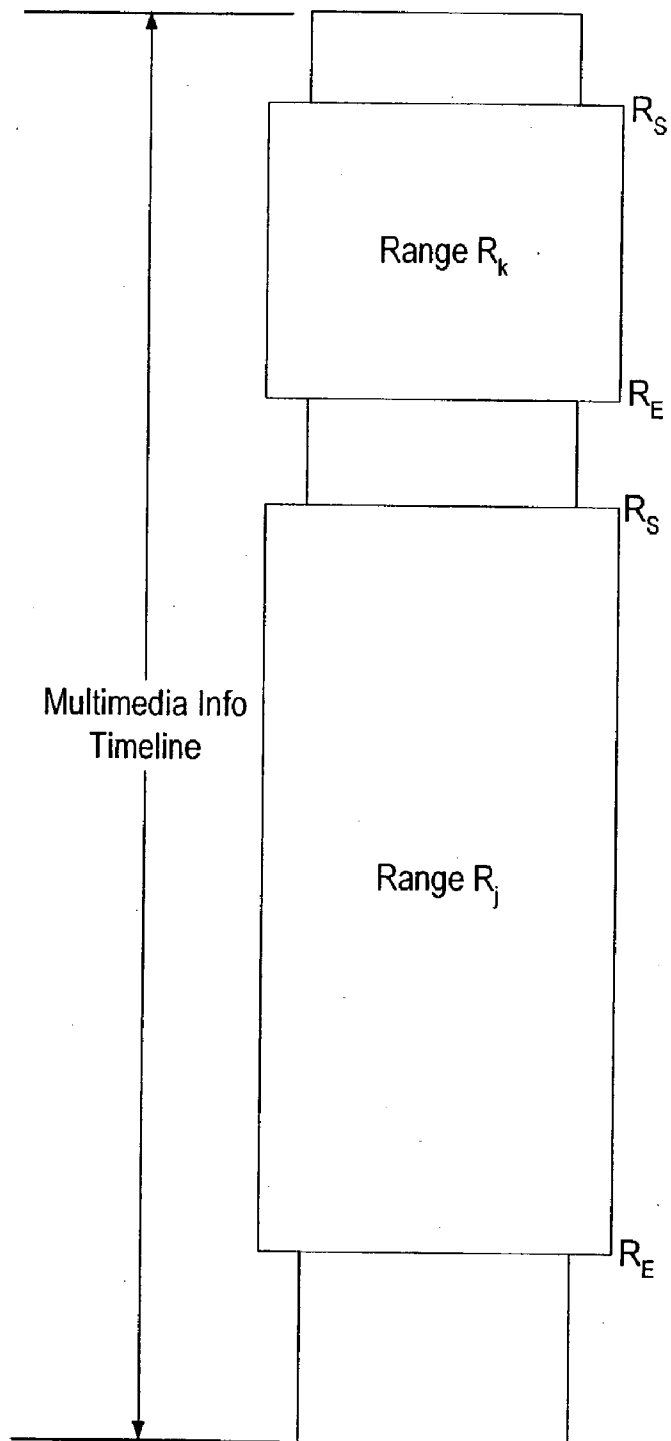


**Fig. 25B**

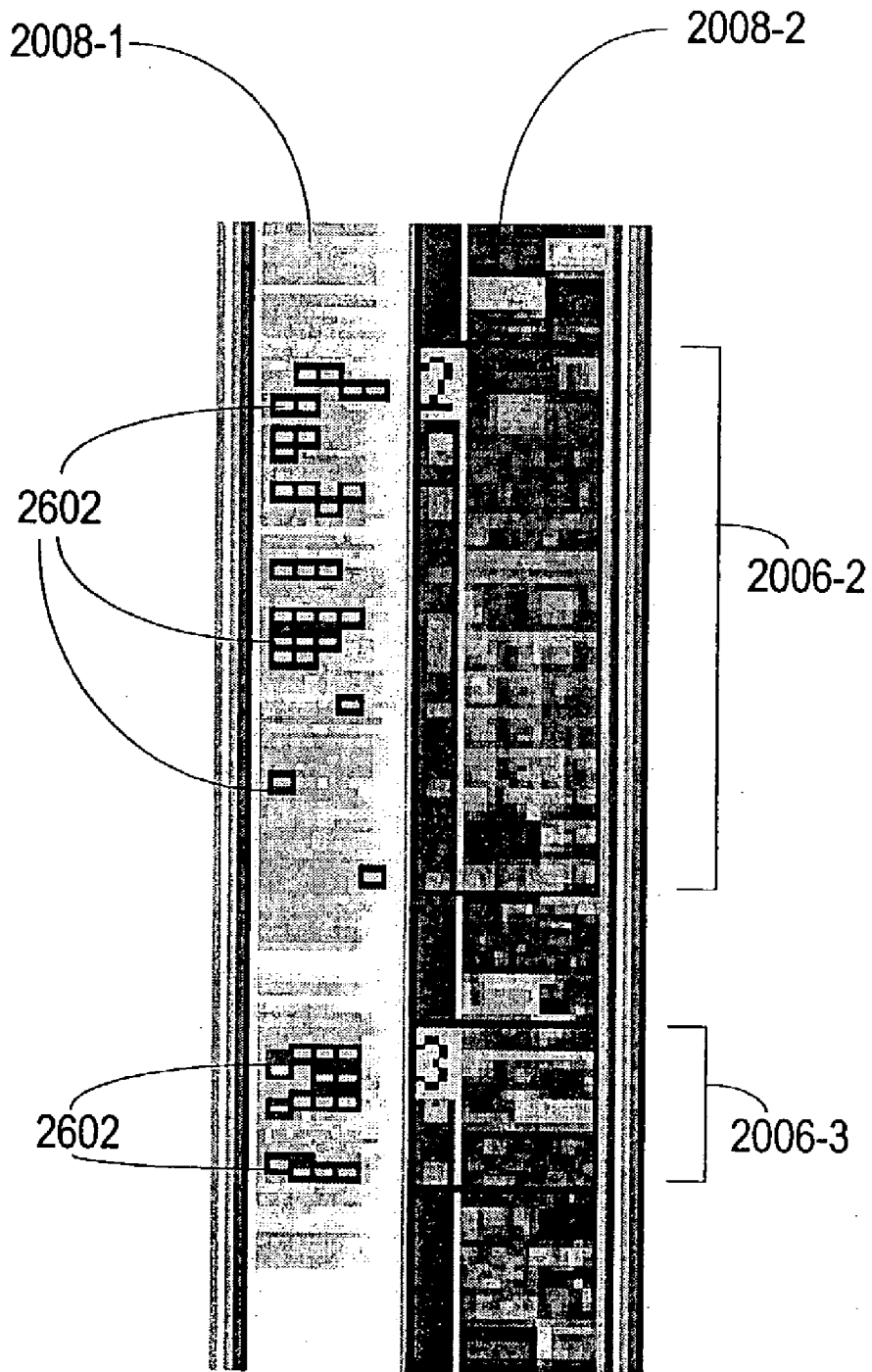


**Fig. 26**

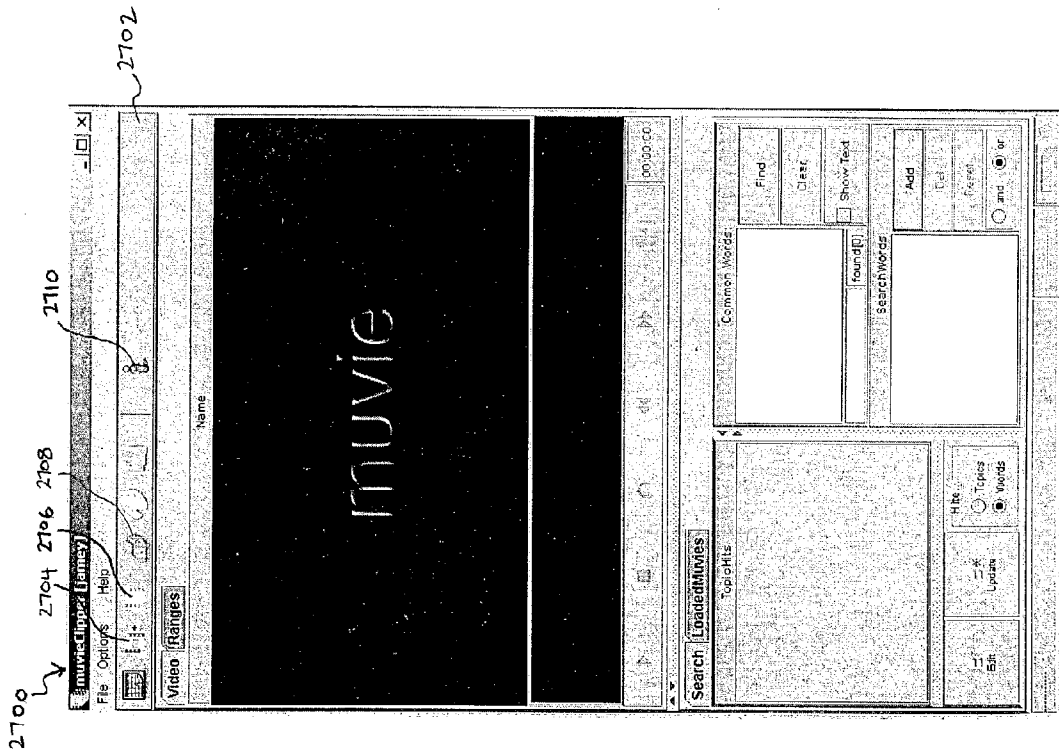


Fig. 27

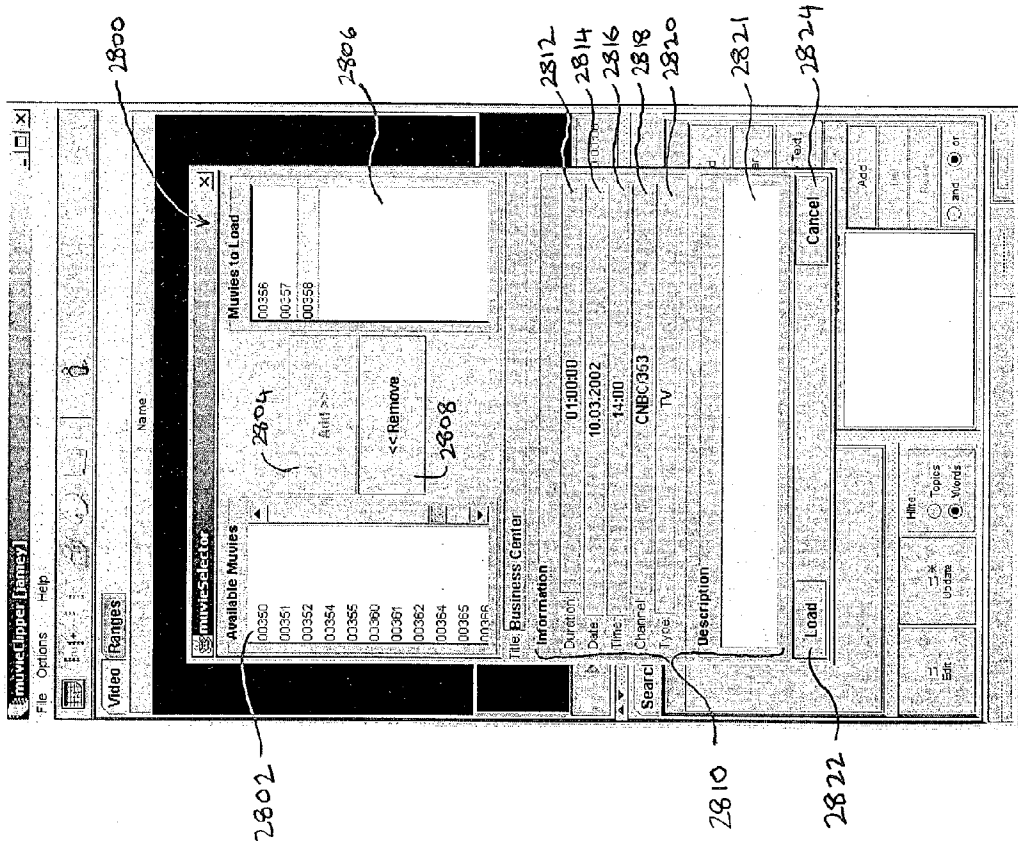
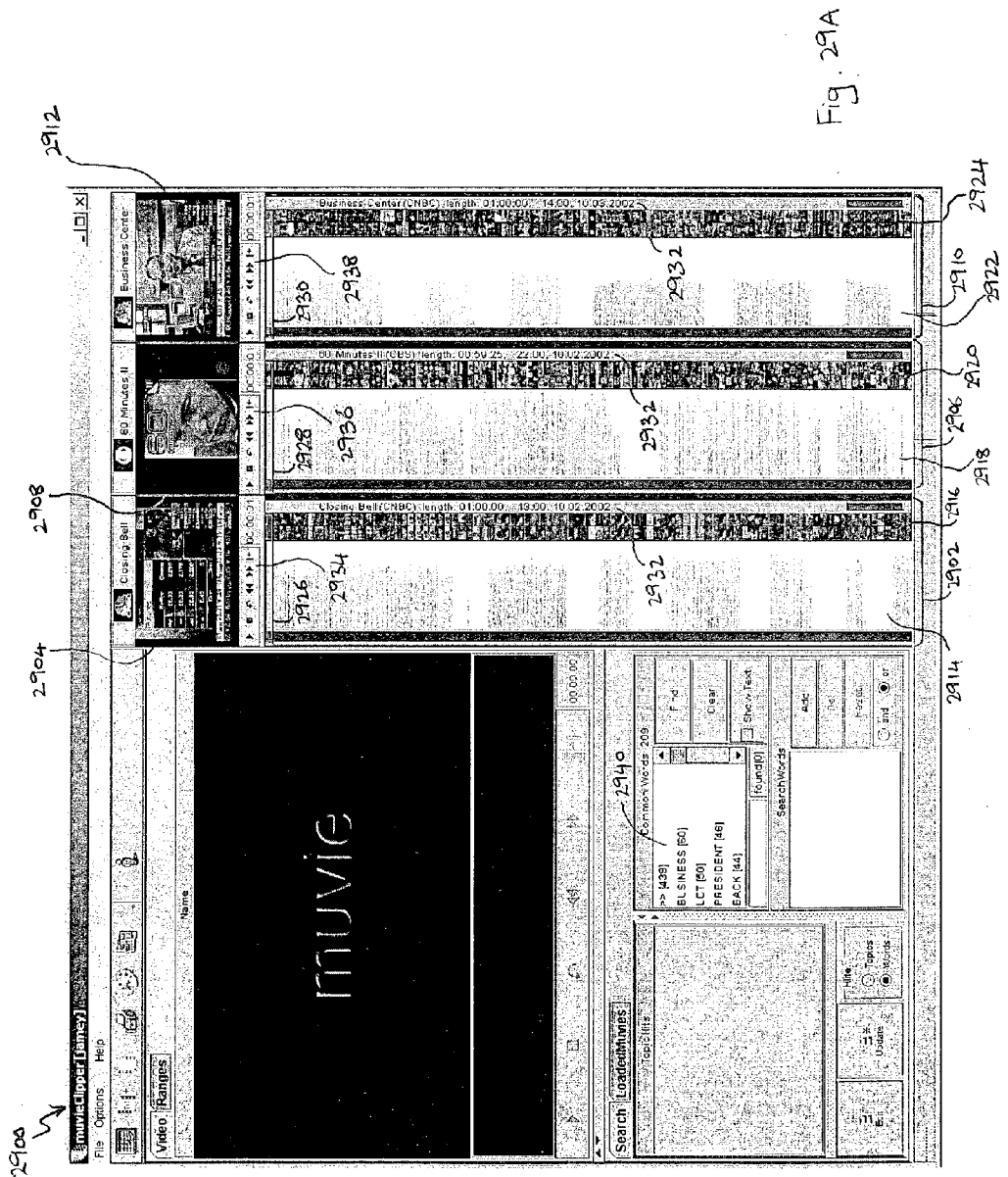
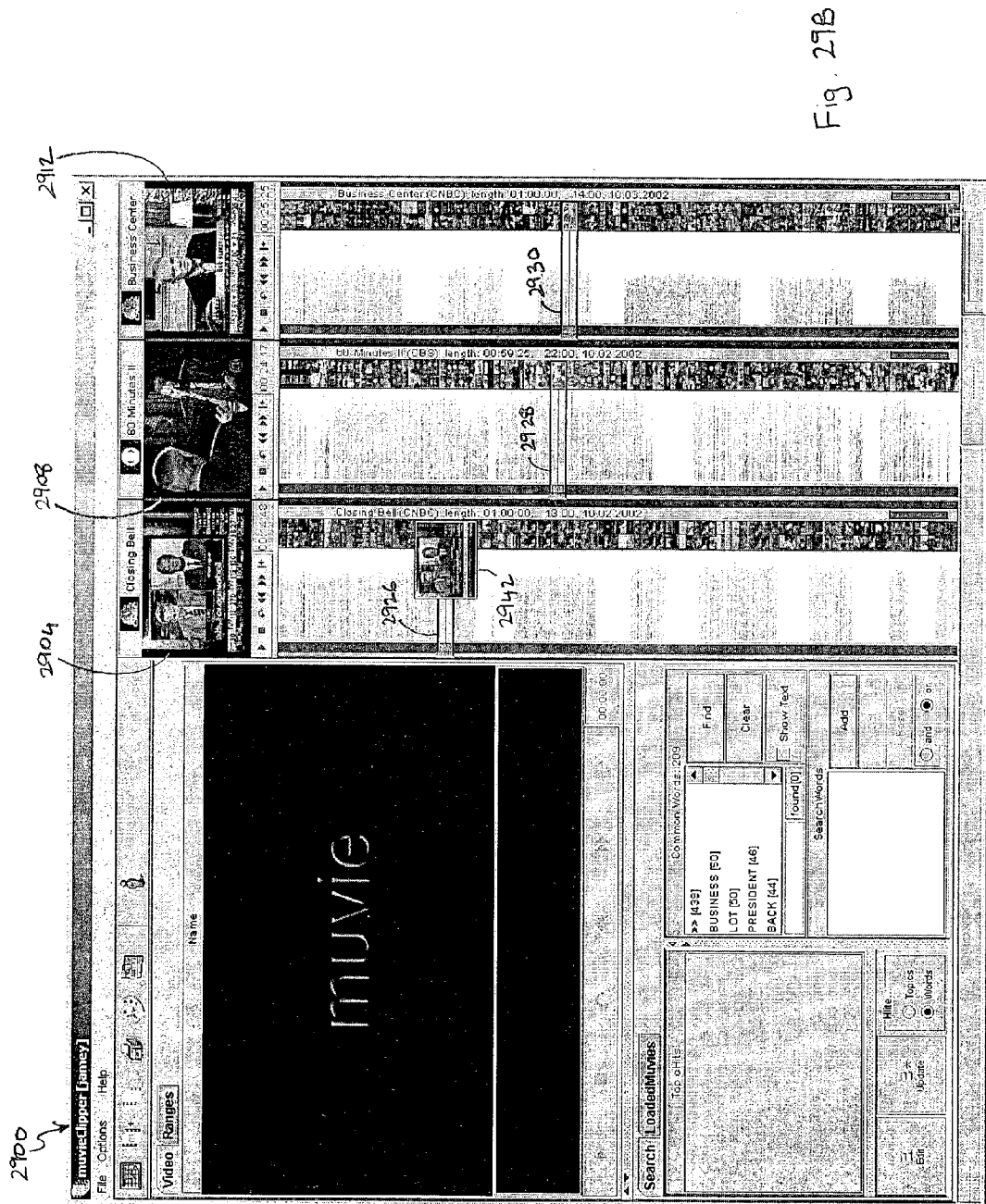
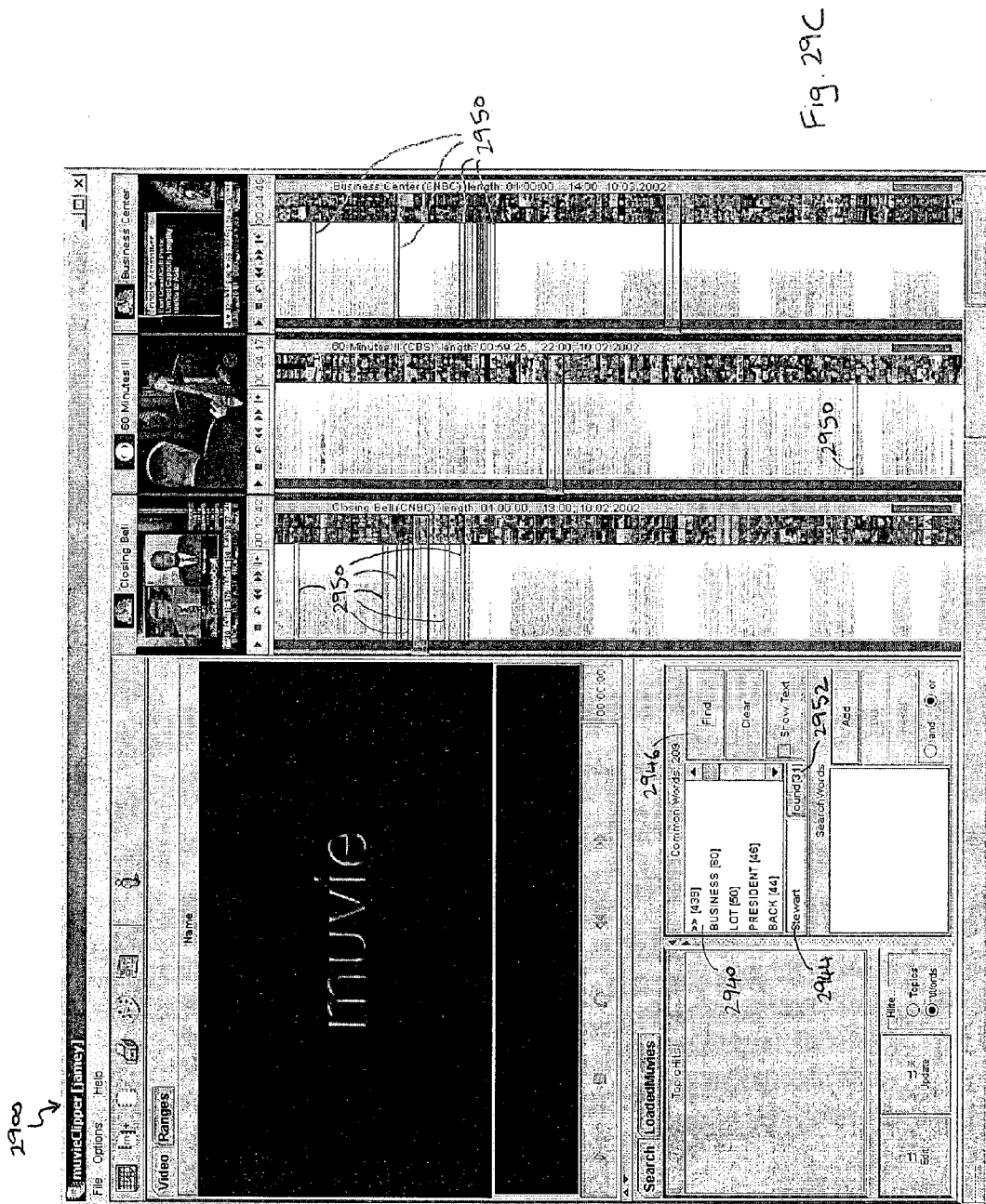
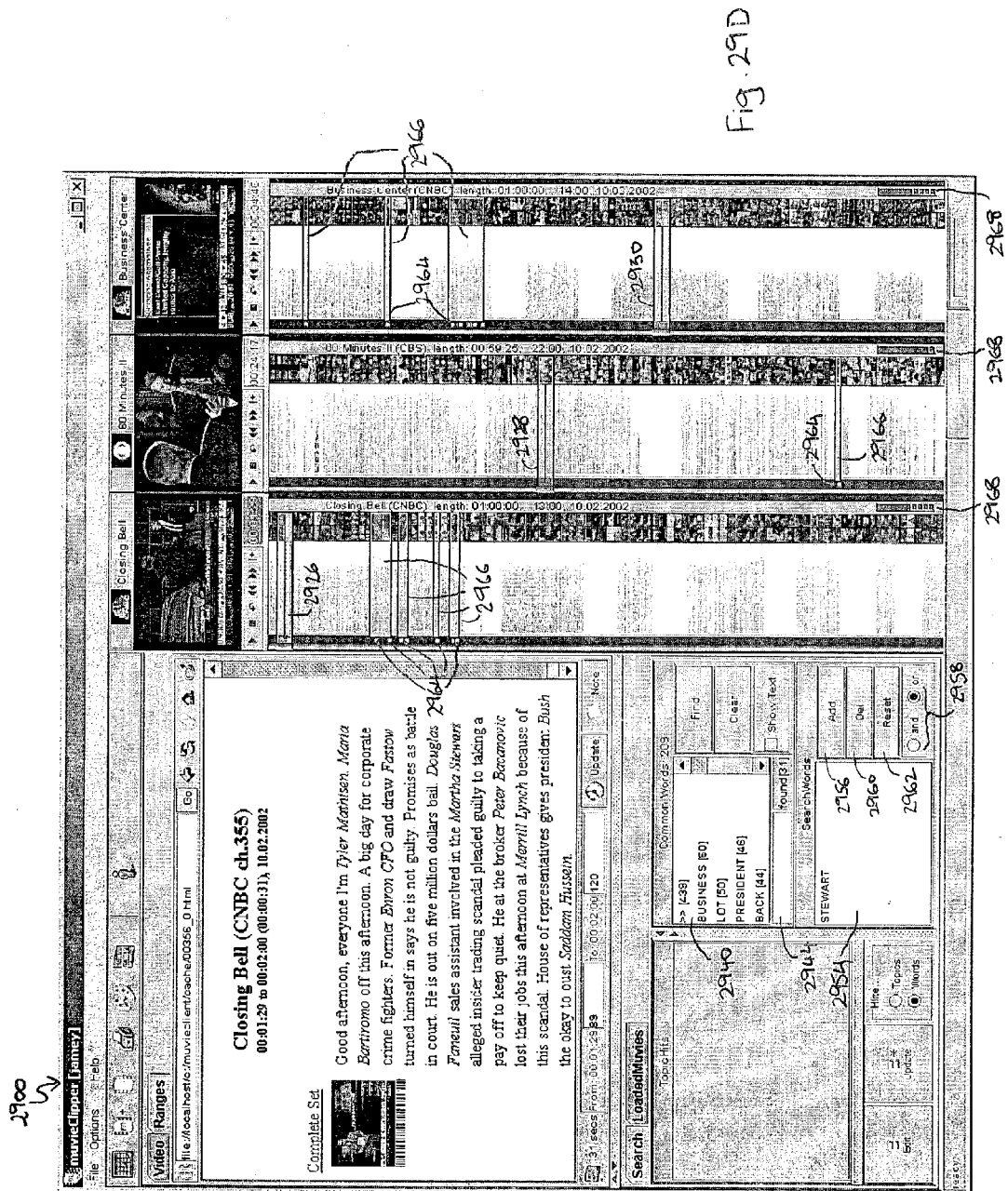


Fig. 28









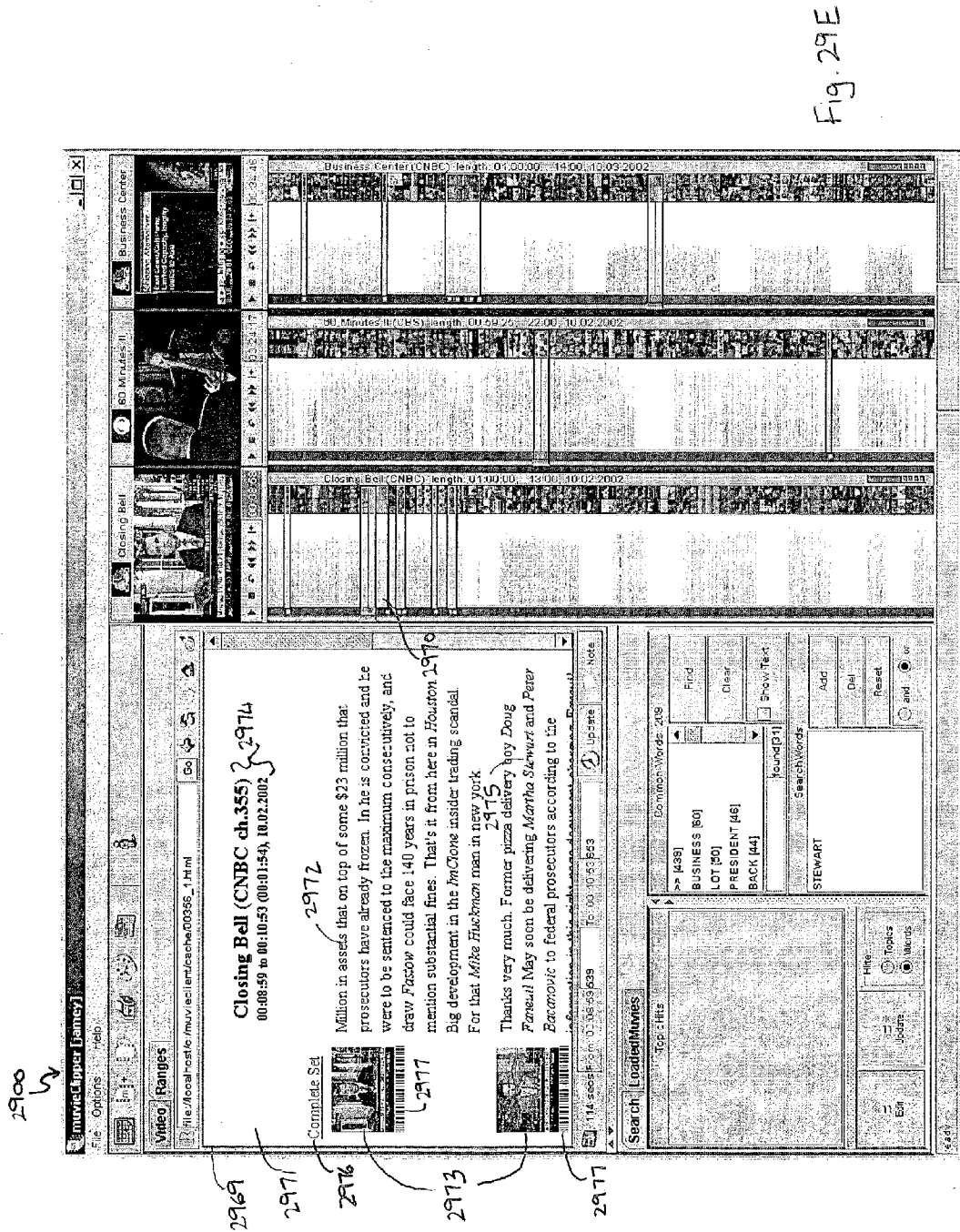


Fig. 29E

2900

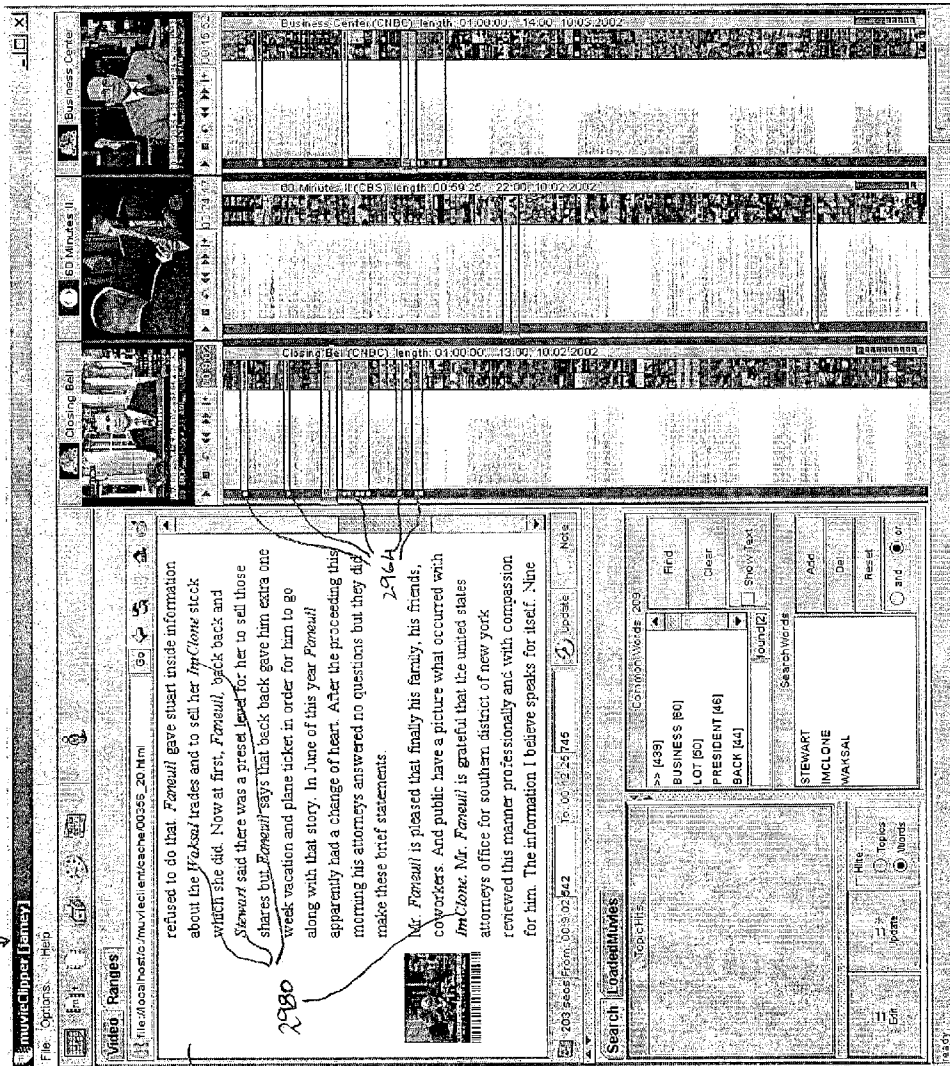


Fig. 29F

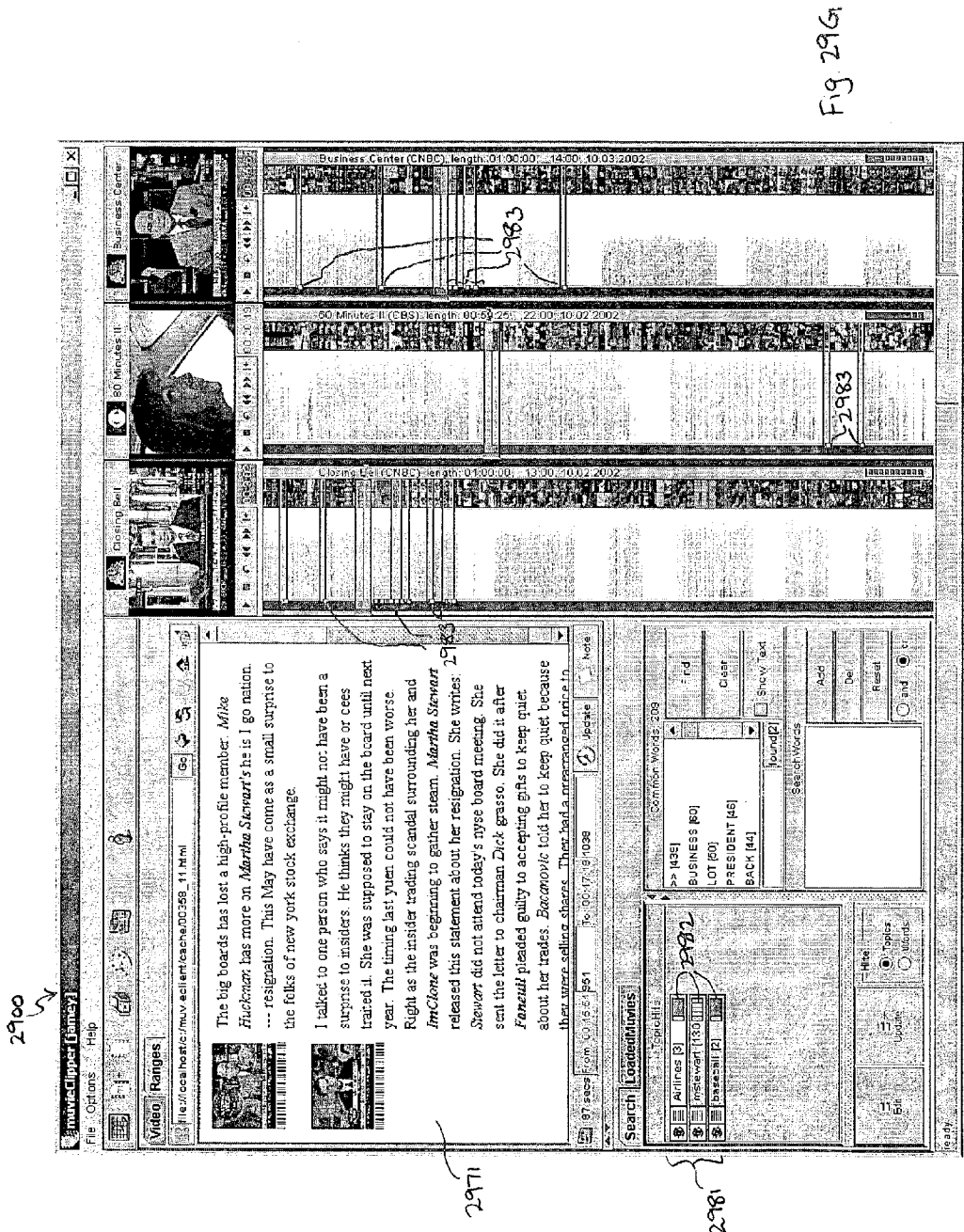
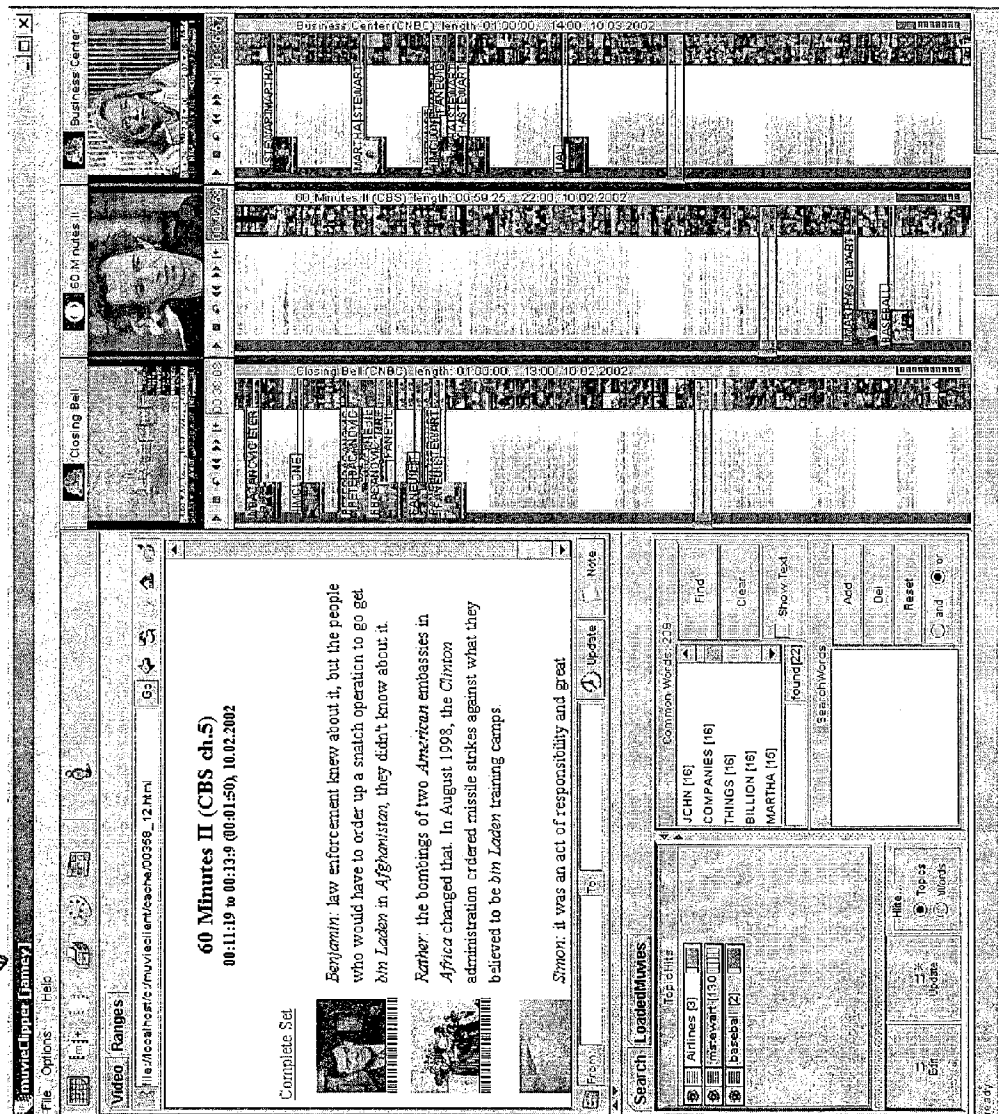
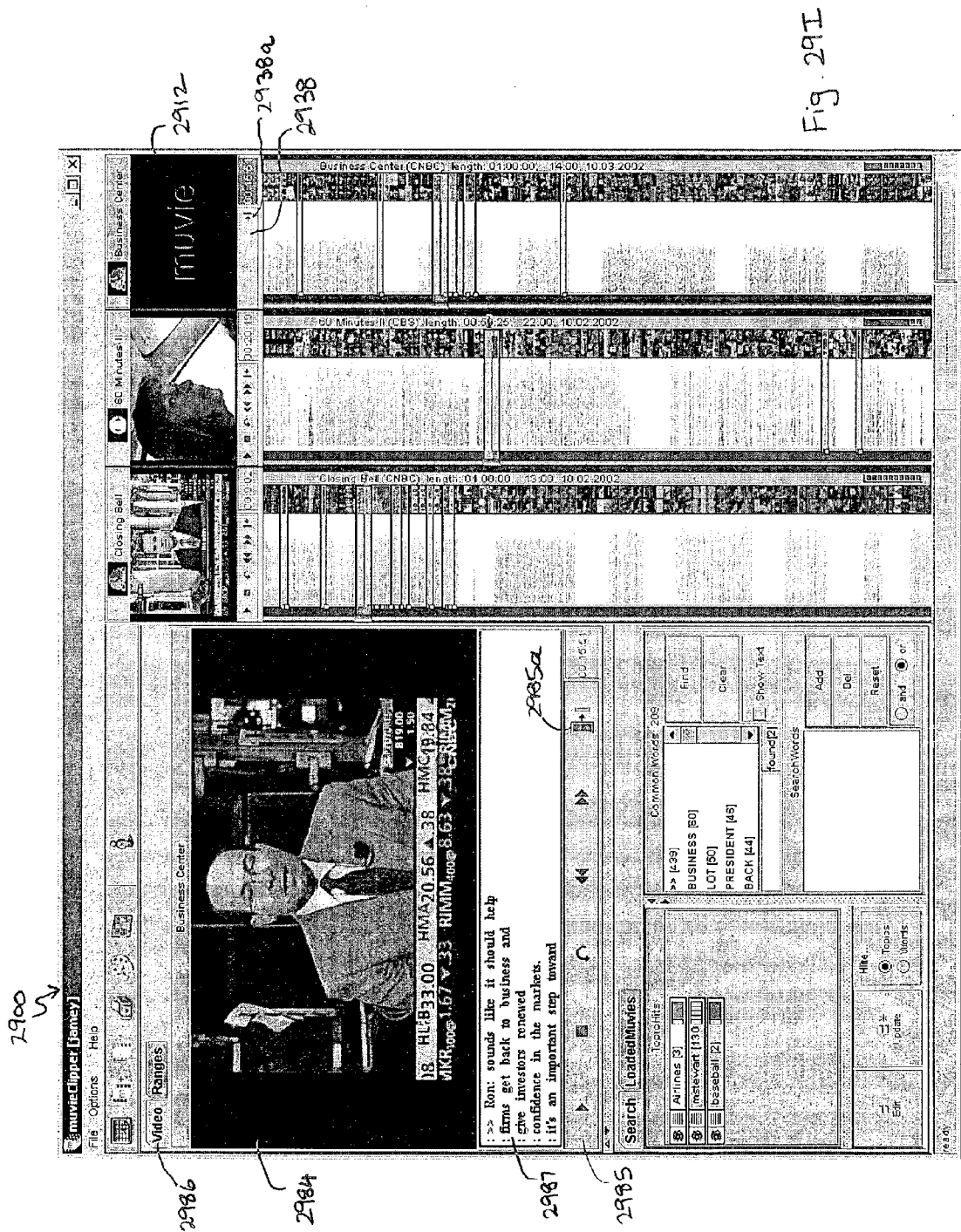


Fig 296





2900  
2988

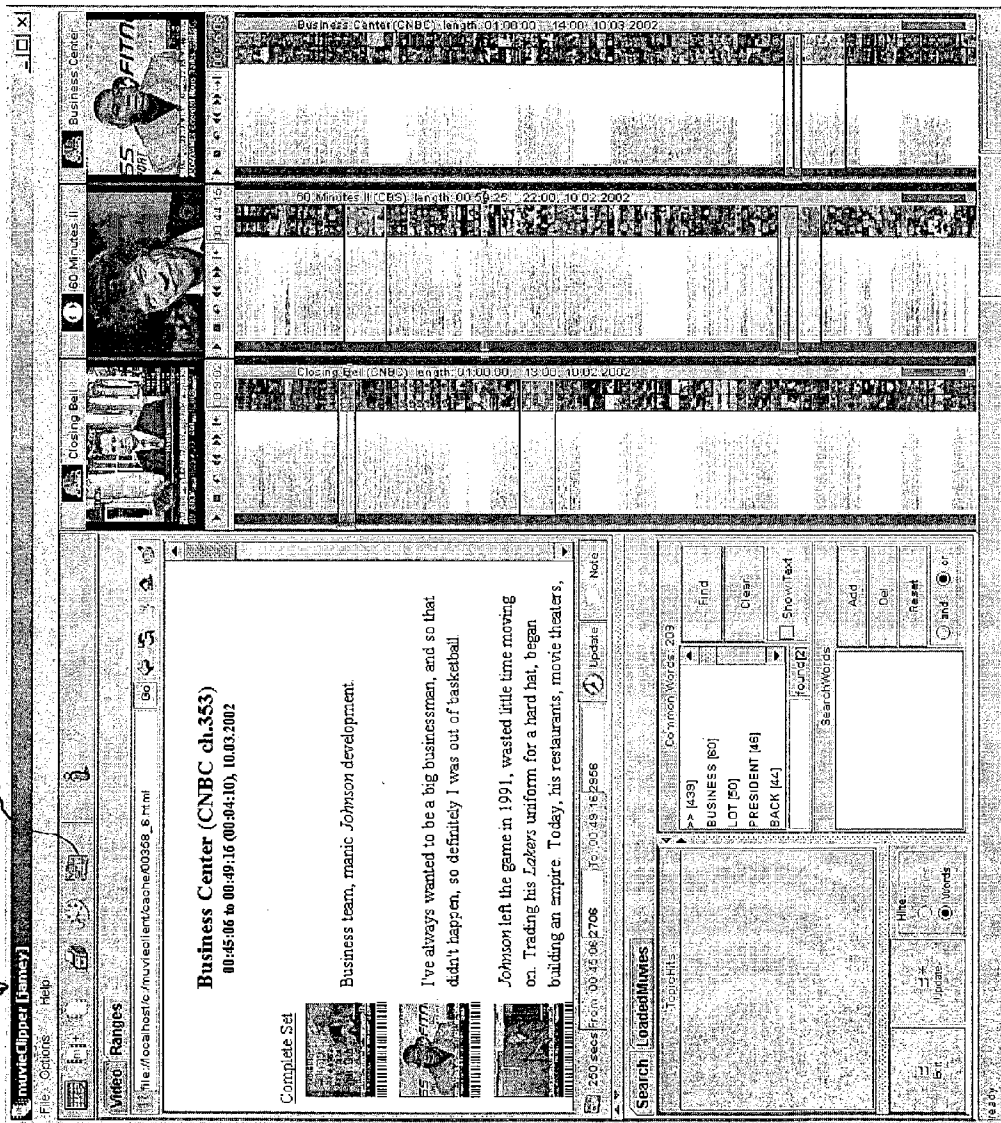
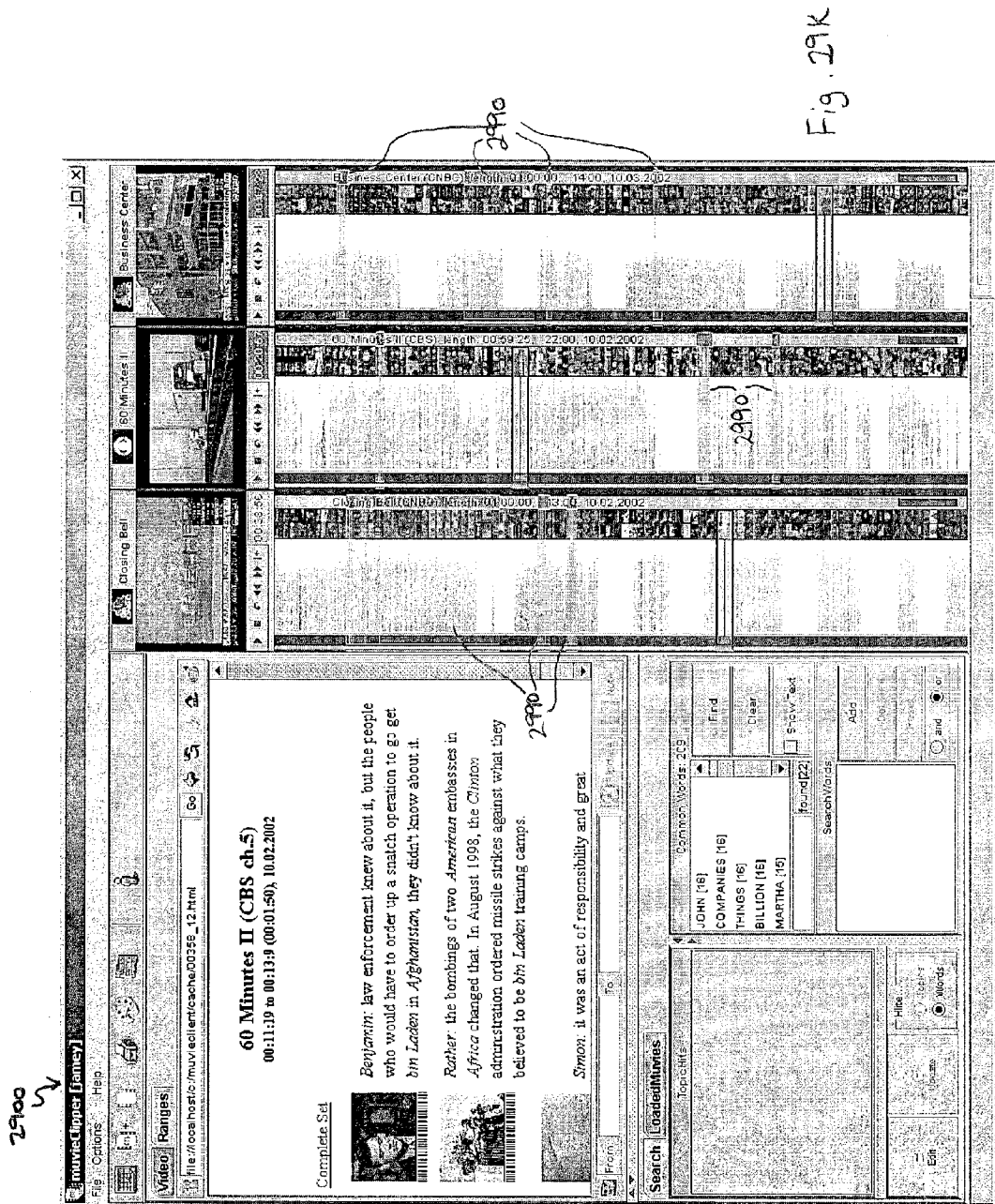
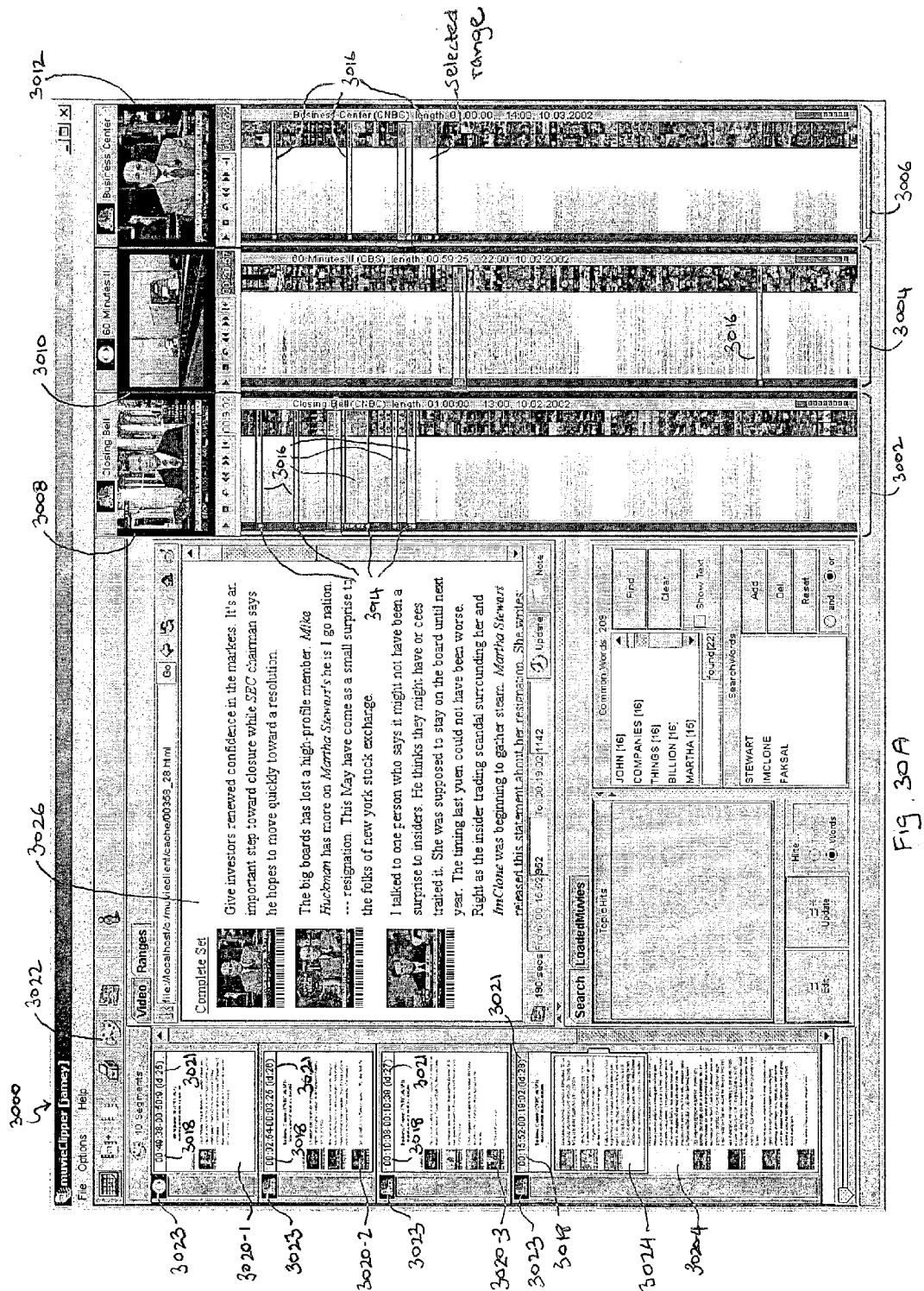
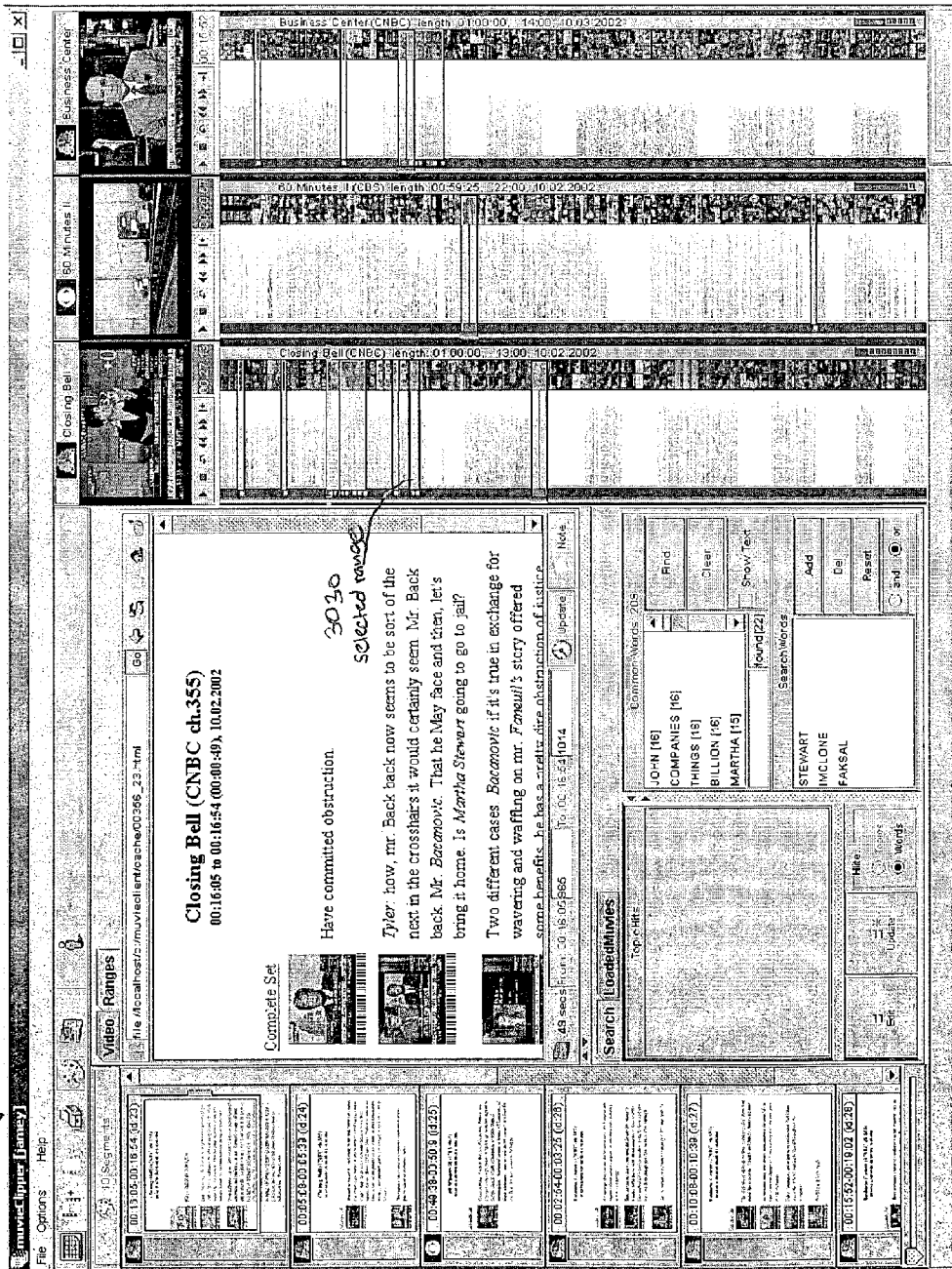


Fig. 29J







303

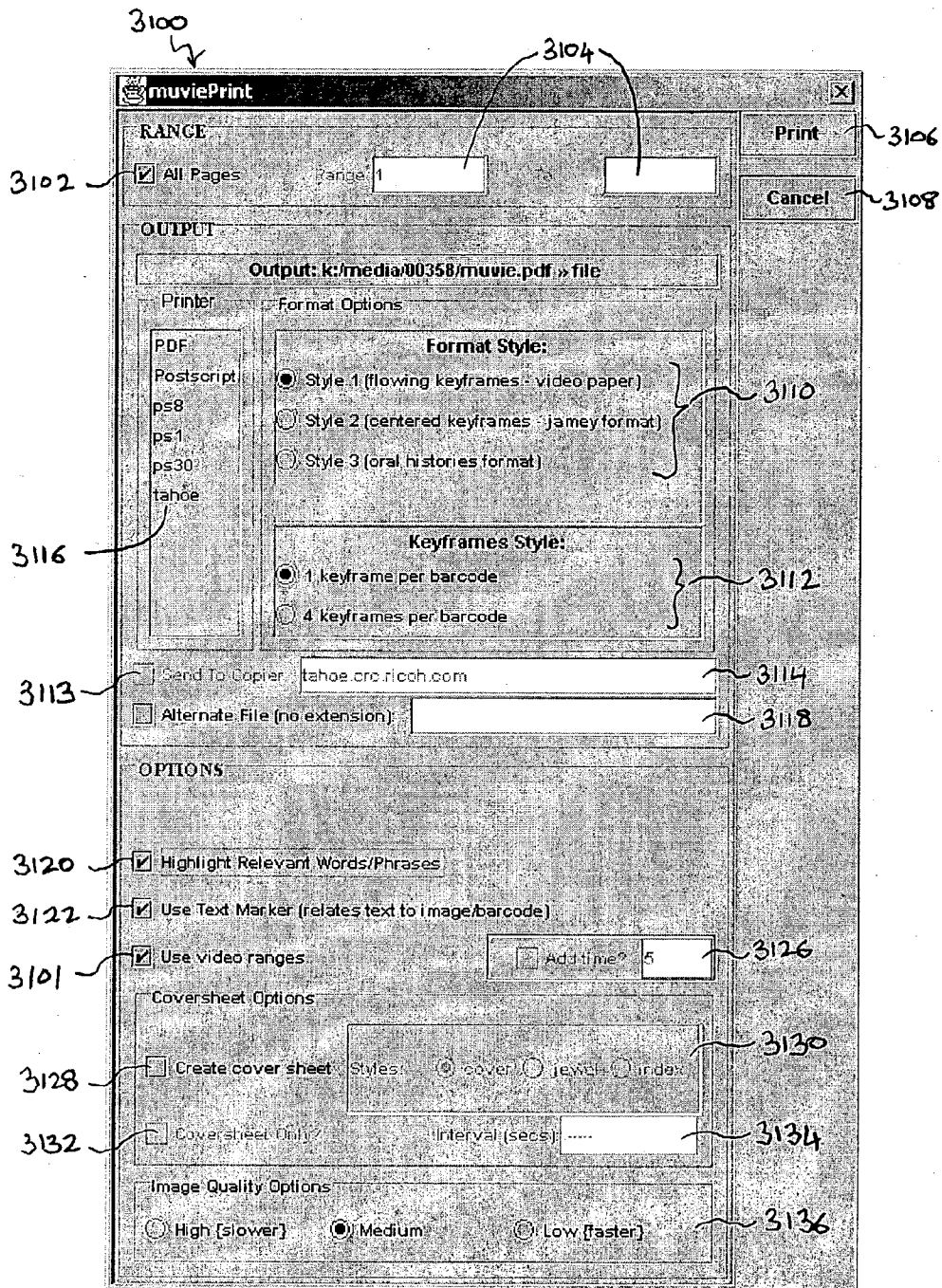


Fig. 31

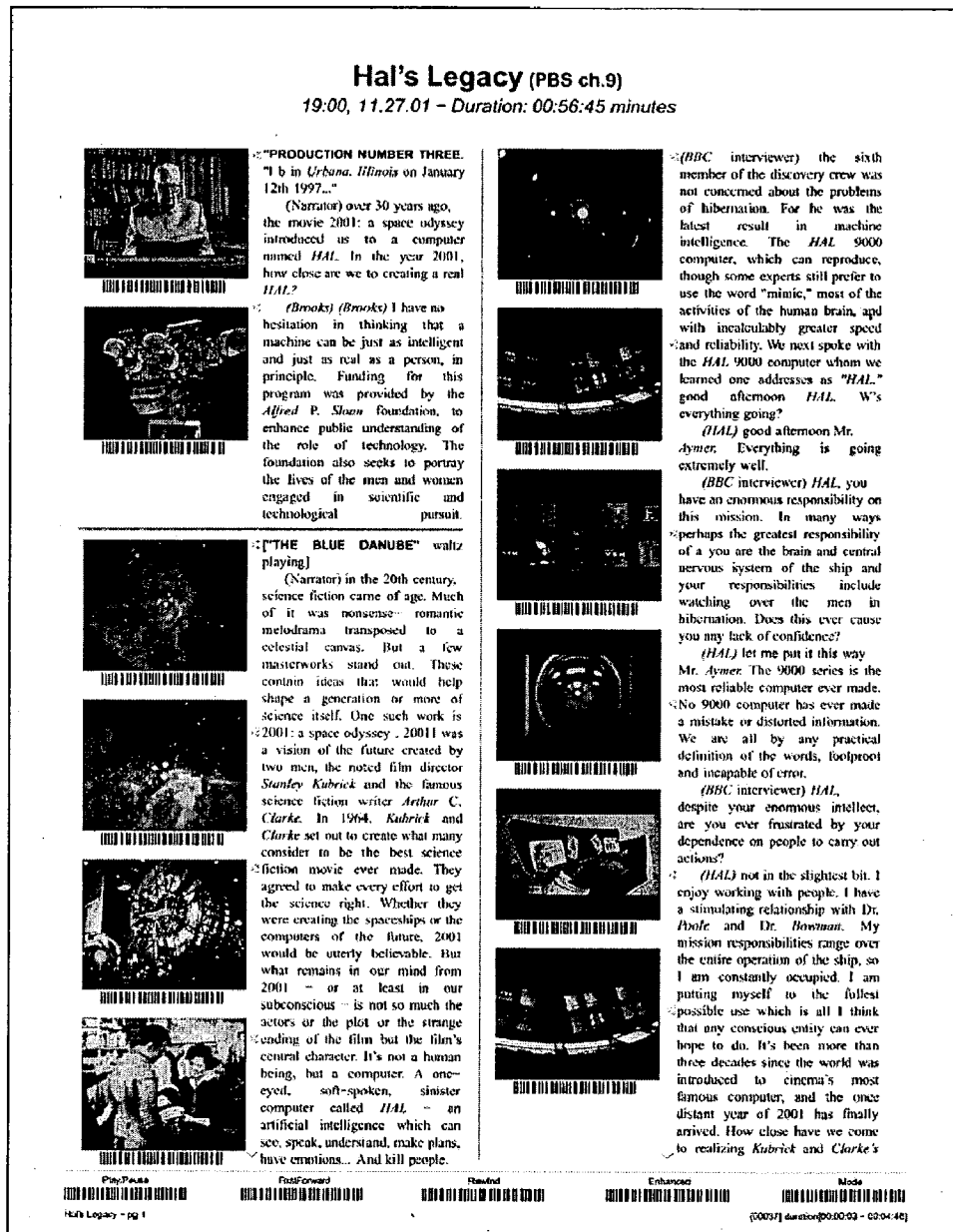


Fig. 32A

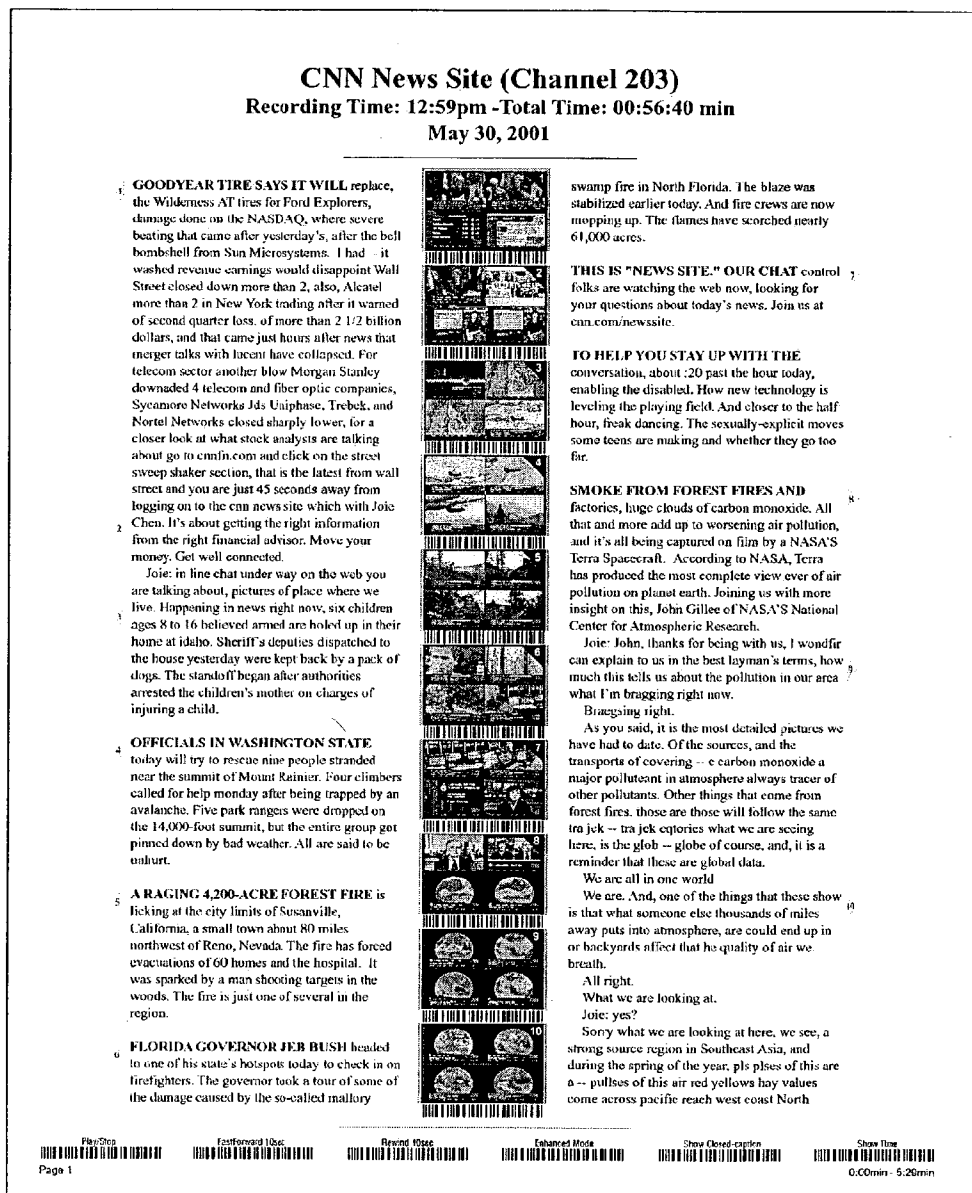


Fig. 32B

GRADUATE SCHOOL - 1970

**Saquin:** Let's talk about the exciting job I had at Bell Labs and why I would never give it up and come to Berkeley. It's actually interesting the way it started. I came there as a fresh Ph.D., basically had been told mostly what to do, and so came into this very structured environment and then was a supervisor. I was a department head. And I was kind of doing what they told me to do, tried to measure those interface states that were essentially smearing out the charge coming out of these charged-couple devices. And I didn't really want to back into the physics stuff I had been doing. I was trying to get into more electronics and and optical stuff and the like.



I vividly remember one day. I must have been there about six months already there was this older guy coming suddenly into my lab and kind of sat down on my desk, and I didn't know who he was. I guess I had seen him in the corridor but didn't quite know who he was and says, "So, Carlo, how are things going?" I said, "Oh, fine. Fine." And says "and then, 'Are you enjoying your life?'". "And I "said, "Oh, yeah. Definitely." He says, "And so what are you really doing here? And so I told him about the various little jobs. He says, "But, what's your real mission? What's your project?" I said, "Well," you know, "my supervisor hasn't told me yet what my project is." And he said, "Well, don't you have a Ph.D.?" And I said, "Yeah." "Well, we expect from our Ph.D.s that they can formulate their own project. After all, that is what Ph.D.s are all about. So, think about your own project." And so I said, "And who are you?" And he says, "Oh, sorry. I'm Gene "Gordon. I'm your lab director." And, I could have crawled under my desk. And he smiled and patted me on the back and walked out. And says, "Oh, me, gosh." that was not very impressive so I had here right to myself

Okay, Ph.D. So I really should decide myself what to do. Whaa. This is going to be difficult. And so it was just kind of a suddenly this insight that I should take responsibility for something. CCDs, at that point, had been proven in principle. I think we had working, at that time, the original eight-bit linear element. That was done before I came. And just as we came, we made a 32-bit linear device, a little bit longer and, still got the charge through that. And then we had an 8x8 array to prove you can actually move charge in two dimensions around. And so things have been progressing, and we said, well, we really should do something bigger, maybe 32x32. And there were some argument whether it was really worthwhile learning that much from a 32x32, or whether maybe we should go directly to 128x128, which would be like sort of a quarter of a TV image resolution. And it wasn't quite clear.

And everybody was kind of busy with various project of measuring the qualities of CCDs and doing various investigations.

1994 170; (Q55): 31 -- (Q20) > 43

interview: 16 July 2002

Fig. 32C



Fig. 33A



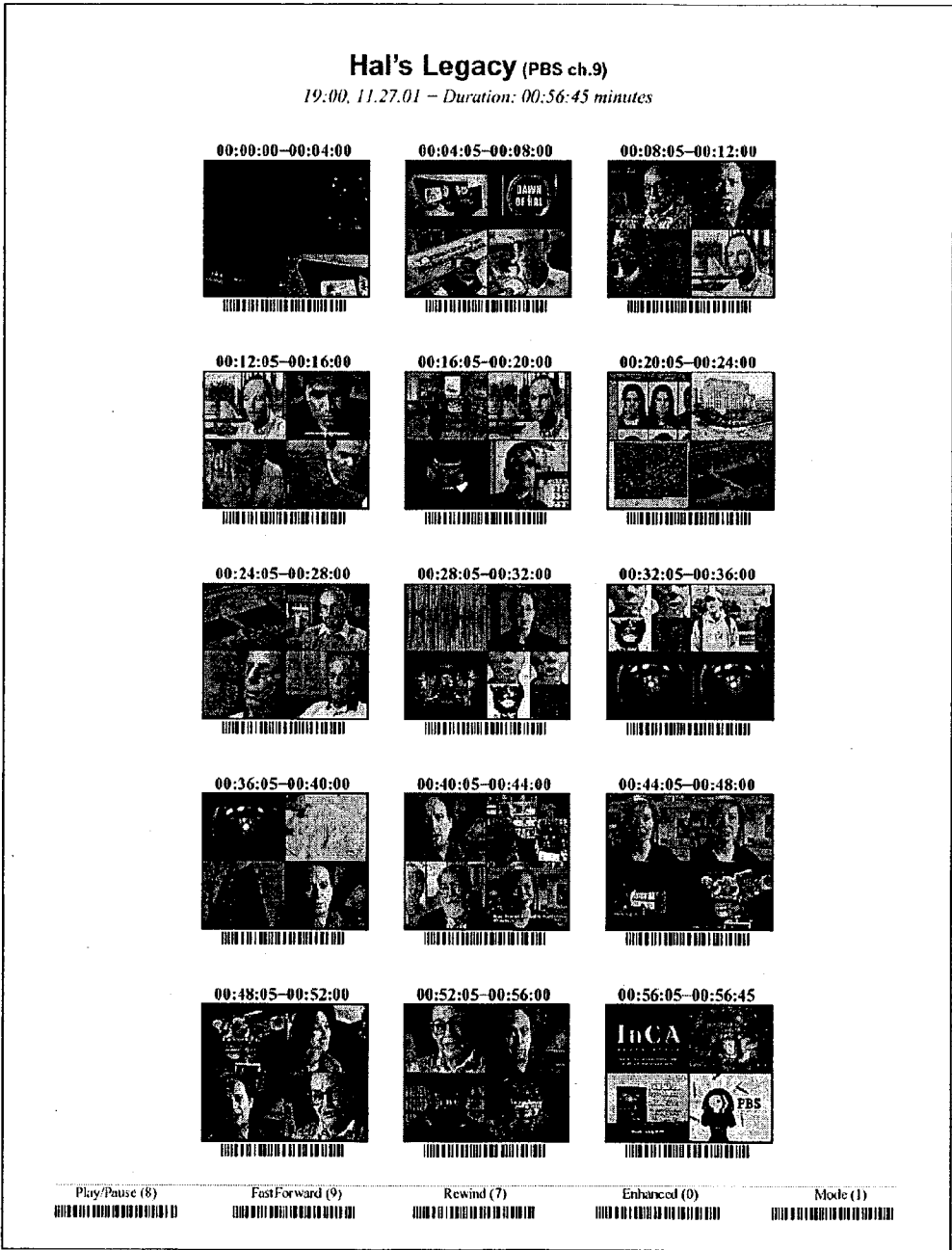


Fig. 34A



Fig. 34 B

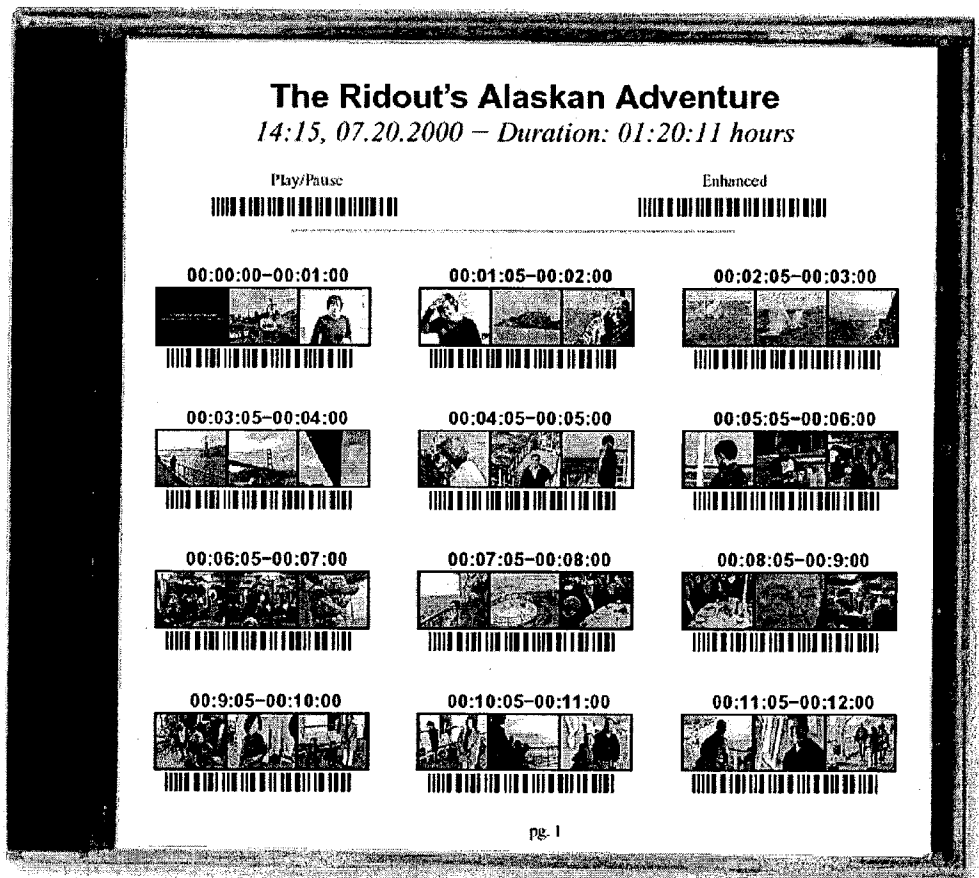


Fig. 34C

# Money and Markets (CNN/fn ch.358)

14:04, 09.10.2002 - Duration: 00:55:39 minutes

3500

3504  
(13) 00:01:21-00:02:52 (00:01:31)



3510

3508

3506

**FEDERAL FALSE STATEMENTS ACTS.** What they're saying to the justice department is we think that someone's been lying to us and it very well May have been *Martha Stewart*. The committee got a lot of inconsistent information from various parties about how that trade took place, how *Martha Stewart's* trade took place, and who called who and so on. So their recommendation to the justice department becomes part of the overall probe.

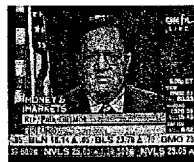
*Louise schiavone*, thank you for filling us in on the developments.



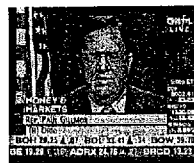
**JOINING US NOW** with his thoughts on today's development, congressman *Paul gilmore*, republican from *Ohio*, member of the house energy and commerce committee, sending that investigation over to the justice department as we just heard. Congressman, thank you for being with us.



Thank you, good to be here. Now, what you and the rest of the committee found out, do you think the serious enough to eventually merit legal charges against *Mrs. Stewart*?



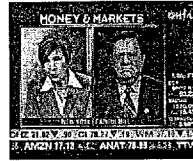
Think there's a very real possibility of that, but that is a decision for the department of justice. We did develop this information almost by accident in terms of the *ImClone* investigation. Certainly, nobody was targeting *Martha Stewart*. But based on the information that we've gathered and, really, the lack of a full response to some of the serious questions that were raised, the feeling was we ought to be referring this to the justice department, along with substantial amount of



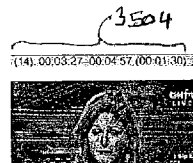
**WHAT WE'VE DONE IS**, as part of our jurisdiction, investigating both accounting practices and the way the *fda* operated, and insider trading, we developed this information, and felt it should appropriately go to justice. They'll make the determination from here.

What is the committee going to do now about the whole *ImClone* issue, as your colleague was talking about last night, the congresswoman was saying, look, there are important issues about a potentially important drug and the way it was -- its approval, or evaluation was handled by the *fda*?

Well, we'll continue to look at that. *Erbix*, I think it's a possibility it's still going to be approved. You have two issues here -- one is corporate governance, insider trading issue, and the other is the efficacy of the drug and how it's approved. I think any focus we have in the future in this respect is likely to be on *fda* and their process.



Sir, you mentioned in passing of course, your committee had not targeted *Stewart* per se, you found this information about possible insider trading as you were investigating *ImClone* and *Erbix* and the waksals. A lot of people out there in the regular public that I talk to are *Stewart* fans, they feel she has been picked on unfairly. And I wonder what they're going to say when they see this investigation of yours, you finally just sort of given up on and throw to the justice department, who doesn't say



**WITH STEWART'S ATTORNEYS** until the very last minute today in an effort to get *Stewart* to talk to the committee behind closed doors. Because of other ongoing probes her lawyers promised *Stewart* would take the fifth. In sending this case to justice, the committee essentially is getting the *Stewart* issue out of its life at a time when some were questioning

Play/Pause  
Money and Markets - pg 1

3512-1

FastForward

3512-2

Rewind

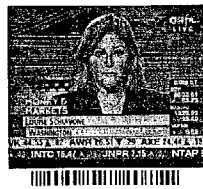
3512-3

Fig. 35A

Enhanced

3512-4

3512-5  
Mode  
(00:31:6) duration(00:01:18 - 00:00:21)



why Stewart had become a focal point and a potential liability. Republicans simply unloaded this case. Late today, Stewart's lawyers responded "any open questions relating to Mrs. Stewart's personal shares should be addressed with the SEC, which have the jurisdiction and expertise to resolve such matters." indeed, the committee today described the review of this as a side show. We, therefore, welcome the committee's decision. That statement, from Martha Stewart's lawyers. Bruce and Kathleen?

Louise, just because the committee refers to justice doesn't mean they have to actually take on the issue and make an investigation here. So what happened now?

Sort of like getting a very bad recommendation, bruce. It's being recommended to the justice department for a criminal inquiry under the federal false statements acts. What they're saying to the justice department is we think that someone's been lying to us and it very well May have been Martha Stewart. The committee got a lot of inconsistent information from various parties about how that trade took place, how Martha Stewart's trade took place, and

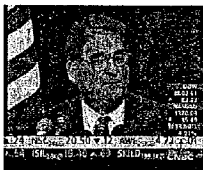
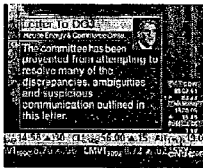


Fig. 35B

# Closing Bell (CNBC ch.355)

13:07, 09.10.2002 – Duration: 00:57:31 minutes

3504  
(16): 00:05:08-00:07:01 (00:01:53)



WEAK, THRIFTS AND GOLD. The dollar quite strong today. Let's go over to the NASDAQ. Leslie?

My question is, I worked at the NASDAQ.

Maria: she is not there.

Tyler: we'll pick it up later

There with Leslie Laroche. She was doing reporting on her own. The house committee investigating the Martha Stewart case looking for help from the justice department. Diana is live on capitol hill with this developing story.

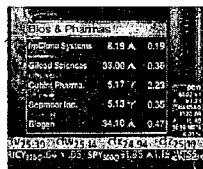
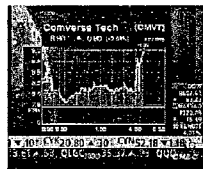
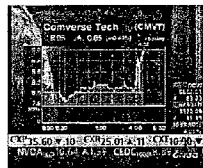
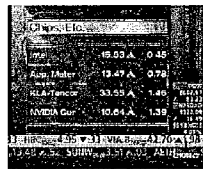
The chairman of the congressional committee investigating Martha Stewart says not only does he believe she traded her 4,000 shares of ImClone stock on insider information, but he also believes that during the committee investigation she may, through her attorneys, have lied to congress. They will not subpoena her because she has already said she would take the fifth amendment if they did. They are turning the entire thing over to the justice department with a strong recommendation that they investigate her for insider trading. Now, in this letter to attorney general John Ashcroft, the chairman says:

This raises a serious question as to whether Ms. Stewart's accounts were false, misleading and designed to conceal material facts. The justice department should conduct an investigation into the questions of whether or not the federal false statements act has been violated in this matter.

Now, the suspicious communications they are talking about revolve around phone calls on December 27th of last year between Ms. Stewart, her stock broker and the former CEO of ImClone. She sold the shares of her stock just one day before the company got some very bad news from the fda. She made \$230,000 on the deal.



3504  
(17): 00:08:35-00:07:06 (00:00:31)



THE FEDERAL FALSE STATEMENTS act has been violated in this matter.

Now, the suspicious communications they are talking about revolve around phone calls on December 27th of last year between Ms. Stewart, her stock broker and the former CEO of ImClone. She sold the shares of her stock just one day before the company got some very bad news from the fda. She made \$230,000 on the deal. Now, back to you.

Maria: thanks, diana. Coming up of a quick break, do

THROUGH TO YESTERDAY'S rally in the gold stocks that trade here at the NASDAQ of the all of those giving back if not much of the gains they packed on yesterday. Not a safe haven in gold today. In the biotech space, medical and pharmaceutical space, ImClone traded on the day as we got news -- I'm sorry. Converse tech. We had late breaking news, they came out after the bell, up side earnings surprise out of converse tech. They beat the street by 5 cents. That's posting a loss in the second quarter of 7 cents a share. And part because of shengni? Their surveillance products that they have. And that division showed strong gains there. Let's take a look at biotechs and some of the headlines there today. Gilead and cubist terminate add license agreement that was more important to cubist. ImClone 19 cents on the day. And I will finish out with a look at jetblue, a couple of headlines crossing, it surged late in the morning. This was on news that it is holding off on plans for secondary offering. So that really lifted fears that

Play/Pause  
Closing Bell - pg 1

FastForward

Rewind

Enhanced

Mode  
(00316) duration: 00:00:22 - 00:57:03

Fig. 35C



ORDER TO CONDUCT CAR bomb and other attacks on U.S. Facilities. These cells have been accumulating the *IntClone* matter and requesting the department of justice investigate whether *Martha Stewart* made false statements to the committee in violation of the federal false statements act. Congress has been

3504 (18) 00:32:42-00:33:56 (00:01:14) 3502



**MARIA: AS PROMISED**, let's see what stocks are moving tonight in extended hours trading. Thanks. We are just getting some comment from *Martha Stewart's* counsel this evening. Tonight the counsel for *Martha Stewart* is saying the following. They say that they have voluntarily provided more than 1500 pages of documents and have spent hours going ahead with additional requests for information. They say, however, it's unfortunate they say that some have prejudged this matter.

3504 (19) 00:57:06-00:57:31 (00:00:25) 3502



**MIDDAY DOUBTS AND** pushed stocks sharply higher toward the close. Home hrapbl security director said U.S. Officials around the world are taking measures against the risk of terrorist attack as the *Bush* administration raised the terror alert warning to code orange. *Sue*: did *Martha Stewart* make false statements to congress about the stock dealings in the chairman say they have asked the justice department to investigate that. Shares of biotech giant *Genentech* tumbled on the news that the company's experimental drug

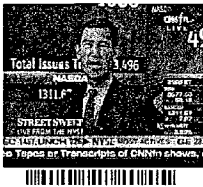
Play/Pause FastForward Rewind Enhanced Mode

Fig. 35D

# Street Sweep (CNN/fn ch.358)

12:00, 09.10.2002 - Duration: 01:00:00 hours

3504  
(22) 00:00:45-00:01:16 (00:00:31)



3502  
THE NASDAQ BOUNCES BACK from the red for the second time this session. Big caps leading the way higher.

Coming up, a terror warning for oil tankers, pushes crude prices higher. We'll tell you what the navy is saying and how the market is responding.



ON THE EVE OF the first



AND CONGRESSIONAL INVESTIGATORS speak out about their findings on Martha Stewart's ImClone sale. Find out what the kitchen is getting hotter. Our first guest is the portfolio manager at Sun America asset manager. It's Tuesday, September 10.

3504  
(24) 00:00:45-00:01:16 (00:00:31)



3502  
THIS PARTICULAR MEETING. A of things in the month of September that will be interesting for the market going forward.



NEXT ON "STREET sweep", the possibility of a U.S. Attack on Iraq has crude prices remaining firm. We'll talk with an oil analyst when we get back.

3502



LATER, DECISION DAY on Martha Stewart as a congressional panel announces the decision about a subpoena. And actually, not necessarily that, they took another tact. We'll have a live report from Washington coming up.

Play/Pause  
STREET SWEEP - pg 1

FastForward  
STREET SWEEP - pg 1

Rewind  
STREET SWEEP - pg 1

Enhanced  
STREET SWEEP - pg 1

Mode  
STREET SWEEP - pg 1

(00:315) duration[00:57:03 - 00:11:15]

Fig.35E

## TECHNIQUES FOR DISPLAYING INFORMATION STORED IN MULTIPLE MULTIMEDIA DOCUMENTS

### CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] The present application claims priority from and is a continuation-in-part (CIP) of the following applications, the entire contents of which are herein incorporated by reference for all purposes:

[0002] (1) U.S. Non-Provisional patent application Ser. No. 10/081,129, filed Feb. 21, 2002; and

[0003] (2) U.S. Non-Provisional application Ser. No. 10/174, 522, filed Jun. 17, 2002.

[0004] The present application also claims priority from and is a non-provisional application of U.S. Provisional Application No. 60/434,314 filed Dec. 17, 2002, the entire contents of which are herein incorporated by reference for all purposes.

### COPYRIGHT

[0005] A portion of the disclosure of this patent document contains material that is subject to copyright protection. The copyright owner has no objection to the xerographic reproduction by anyone of the patent document or the patent disclosure in exactly the form it appears in the U.S. Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

[0006] The present application also incorporates by reference for all purposes the entire contents of:

[0007] (1) U.S. Non-Provisional application Ser. No. 10/001,895 (Attorney Docket No.: 15358-006500US), entitled "PAPER-BASED INTERFACE FOR MULTIMEDIA INFORMATION" filed Nov. 19, 2001;

[0008] (2) U.S. Non-Provisional application Ser. No. 08/995,616 filed Dec. 22, 1997; and

[0009] (3) U.S. Non-Provisional Application No. \_\_\_\_\_ (Attorney Docket No. 15358-006640US) filed concurrently with this application.

### BACKGROUND OF THE INVENTION

[0010] The present invention relates to user interfaces for displaying information and more particularly to user interfaces for retrieving and displaying multimedia information that may be stored in one or more multimedia documents.

[0011] With rapid advances in computer technology, an increasing amount of information is being stored in the form of electronic (or digital) documents. These electronic documents include multimedia documents that store multimedia information. The term "multimedia information" is used to refer to information that comprises information of several different types in an integrated form. The different types of information included in multimedia information may include a combination of text information, graphics information, animation information, sound (audio) information, video information, slides information, whiteboard information, and other types of information. Multimedia information is also used to refer to information comprising one or more

objects wherein the objects include information of different types. For example, multimedia objects included in multimedia information may comprise text information, graphics information, animation information, sound (audio) information, video information, slides information, whiteboard information, and other types of information. Multimedia documents may be considered as compound objects that comprise video, audio, closed-caption text, keyframes, presentation slides, whiteboard capture information, as well as other multimedia type objects. Examples of multimedia documents include documents storing interactive web pages, television broadcasts, videos, presentations, or the like.

[0012] Several tools and applications are conventionally available that allow users to play back, store, index, edit, or manipulate multimedia information stored in multimedia documents. Examples of such tools and/or applications include proprietary or customized multimedia players (e.g., RealPlayer™ provided by RealNetworks, Microsoft Windows Media Player provided by Microsoft Corporation, QuickTime™ Player provided by Apple Corporation, Shockwave multimedia player, and others), video players, televisions, personal digital assistants (PDAs), or the like. Several tools are also available for editing multimedia information. For example, Virage, Inc. of San Mateo, Calif. (www.virage.com) provides various tools for viewing and manipulating video content and tools for creating video databases. Virage, Inc. also provides tools for face detection and on-screen text recognition from video information.

[0013] Given the vast number of electronic documents, readers of electronic documents are increasingly being called upon to assimilate vast quantities of information in a short period of time. To meet the demands placed upon them, readers find they must read electronic documents "horizontally" rather than "vertically," i.e., they must scan, skim, and browse sections of interest in one or more electronic documents rather than read and analyze a single document from start to end. While tools exist which enable users to "horizontally" read or skim electronic documents containing text/image information (e.g., the reading tool described in U.S. Non-Provisional patent application Ser. No. 08/995,616), conventional tools cannot be used to "horizontally" read or skim multimedia documents which may contain audio information, video information, and other types of information. None of the multimedia tools described above allow users to "horizontally" read or skim a multimedia document.

[0014] In light of the above, there is a need for techniques that allow users to skim or read a multimedia document "horizontally." Techniques that allow users to view, analyze, and navigate multimedia information stored in multimedia documents are desirable.

### BRIEF SUMMARY OF THE INVENTION

[0015] Embodiments of the present invention provide techniques for providing a graphical user interface (GUI) that displays a representation of stored information that may include information of one or more types. The displayed representation may include representations of information of the one or more types. The GUI enables a user to navigate and skim through the stored information and to analyze the contents of the stored information. The stored information may include information captured along the same timeline or along different timelines.

[0016] According to an embodiment of the present invention, a first representation of first stored information is displayed. The first stored information comprises information of a first type and information of a second type. The first representation comprises a representation of information of the first type included in the first stored information and a representation of the information of the second type included in the first stored information. One or more portions of the first representation are highlighted, the highlighted one or more portions of the first representation corresponding to portions of the first representation that include a first criterion.

[0017] According to another embodiment of the present invention, in addition to displaying a first representation of first stored information, a second representation of second stored information is displayed. The second stored information comprises information of a first type and information of a second type. The second representation comprises a representation of information of the first type included in the second stored information and a representation of information of the second type included in the second stored information. One or more portions of the second representation are highlighted, the highlighted one or more portions of the first representation corresponding to portions of the second representation that include the first criterion.

[0018] According to another embodiment of the present invention, techniques are provided for displaying multimedia information. A first thumbnail is displayed comprising a representation of information of a first type included in a first recorded information. A second thumbnail is displayed comprising a representation of information of a second type included in the first recorded information. A third thumbnail is displayed comprising a representation of information of a first type included in a second recorded information. A fourth thumbnail is displayed comprising a representation of information of a second type included in the second recorded information. According to one embodiment, one or more portions of the first thumbnail and the third thumbnail (or the second thumbnail and the fourth thumbnail) that comprise at least one word from the set of words (or a topic of interest) are highlighted.

[0019] According to yet another embodiment of the present invention, techniques are provided for displaying information included in a first recorded information and a second recorded information, the first recorded information comprising audio information and video information, the second recorded information comprising audio and video information. A first representation of information included in the first recorded information is displayed, the first representation comprising a first thumbnail and a second thumbnail, the first thumbnail comprising text information obtained from the audio information included in the first recorded information, the second thumbnail comprising one or more keyframes extracted from the video information included in the first recorded information. A second representation of information included in the second recorded information is displayed, the second representation comprising a third thumbnail and a fourth thumbnail, the third thumbnail comprising text information obtained from the audio information included in the second recorded information, the fourth thumbnail comprising one or more keyframes extracted from the video information included in the second recorded information. According to one embodi-

ment, one or more portions of the first representation and the second representation that include the user criterion are highlighted, wherein a highlighted portion of the first representation covers a section of the first thumbnail and the second thumbnail and a highlighted portion of the second representation covers a section of the third thumbnail and the fourth thumbnail.

[0020] According to another embodiment of the present invention, techniques are provided for displaying information. A representation of stored information is displayed. Information indicative of one or more portions of the stored information that have been output is received. One or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have been output are highlighted.

[0021] According to an embodiment of the present invention, techniques are provided for displaying information. A representation of stored information is displayed. Information indicative of one or more portions of the stored information that have been output is received. One or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have not been output are highlighted.

[0022] The foregoing, together with other features, embodiments, and advantages of the present invention, will become more apparent when referring to the following specification, claims, and accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0023] FIG. 1 is a simplified block diagram of a distributed network that may incorporate an embodiment of the present invention;

[0024] FIG. 2 is a simplified block diagram of a computer system according to an embodiment of the present invention;

[0025] FIG. 3 depicts a simplified user interface 300 generated according to an embodiment of the present invention for viewing multimedia information;

[0026] FIG. 4 is a zoomed-in simplified diagram of a thumbnail viewing area lens according to an embodiment of the present invention;

[0027] FIGS. 5A, 5B, and 5C are simplified diagrams of a panel viewing area lens according to an embodiment of the present invention;

[0028] FIG. 6 depicts a simplified user interface generated according to an embodiment of the present invention wherein user-selected words are annotated or highlighted;

[0029] FIG. 7 is a simplified zoomed-in view of a second viewing area of a GUI generated according to an embodiment of the present invention;

[0030] FIG. 8 depicts a simplified GUI in which multimedia information that is relevant to one or more topics of interest to a user is annotated or highlighted according to an embodiment of the present invention;

[0031] FIG. 9 depicts a simplified user interface for defining a topic of interest according to an embodiment of the present invention;

[0032] FIG. 10 depicts a simplified user interface that displays multimedia information stored by a meeting recording according to an embodiment of the present invention;

[0033] FIG. 11 depicts a simplified user interface that displays multimedia information stored by a multimedia document according to an embodiment of the present invention;

[0034] FIG. 12 depicts a simplified user interface that displays multimedia information stored by a multimedia document according to an embodiment of the present invention;

[0035] FIG. 13 depicts a simplified user interface that displays contents of a multimedia document according to an embodiment of the present invention;

[0036] FIG. 14 is a simplified high-level flowchart depicting a method of displaying a thumbnail depicting text information in the second viewing area of a GUI according to an embodiment of the present invention;

[0037] FIG. 15 is a simplified high-level flowchart depicting a method of displaying a thumbnail that depicts video keyframes extracted from the video information in the second viewing area of a GUI according to an embodiment of the present invention;

[0038] FIG. 16 is a simplified high-level flowchart depicting another method of displaying thumbnail 312-2 according to an embodiment of the present invention;

[0039] FIG. 17 is a simplified high-level flowchart depicting a method of displaying thumbnail viewing area lens 314, displaying information emphasized by thumbnail viewing area lens 314 in third viewing area 306, displaying panel viewing area lens 322, displaying information emphasized by panel viewing area lens 322 in fourth viewing area 308, and displaying information in fifth viewing area 310 according to an embodiment of the present invention;

[0040] FIG. 18 is a simplified high-level flowchart depicting a method of automatically updating the information displayed in third viewing area 306 in response to a change in the location of thumbnail viewing area lens 314 according to an embodiment of the present invention;

[0041] FIG. 19 is a simplified high-level flowchart depicting a method of automatically updating the information displayed in fourth viewing area 308 and the positions of thumbnail viewing area lens 314 and sub-lens 316 in response to a change in the location of panel viewing area lens 322 according to an embodiment of the present invention;

[0042] FIG. 20A depicts a simplified user interface that displays ranges according to an embodiment of the present invention;

[0043] FIG. 20B depicts a simplified dialog box for editing ranges according to an embodiment of the present invention;

[0044] FIG. 21 is a simplified high-level flowchart depicting a method of automatically creating ranges according to an embodiment of the present invention;

[0045] FIG. 22 is a simplified high-level flowchart depicting a method of automatically creating ranges based upon

locations of hits in the multimedia information according to an embodiment of the present invention;

[0046] FIG. 23 is a simplified high-level flowchart depicting a method of combining one or more ranges based upon the size of the ranges and the proximity of the ranges to neighboring ranges according to an embodiment of the present invention;

[0047] FIG. 24 depicts a simplified diagram showing the relationships between neighboring ranges according to an embodiment of the present invention;

[0048] FIG. 25A depicts a simplified diagram showing a range created by combining ranges  $R_i$  and  $R_k$  depicted in FIG. 24 according to an embodiment of the present invention;

[0049] FIG. 25B depicts a simplified diagram showing a range created by combining ranges  $R_i$  and  $R_j$  depicted in FIG. 24 according to an embodiment of the present invention;

[0050] FIG. 26 depicts a zoomed-in version of a GUI depicting ranges that have been automatically created according to an embodiment of the present invention;

[0051] FIG. 27 depicts a simplified startup user interface that displays information that may be stored in one or more multimedia documents according to an embodiment of the present invention;

[0052] FIG. 28 depicts a simplified window that is displayed when a user selects a load button according to an embodiment of the present invention;

[0053] FIGS. 29A, 29B, 29C, 29D, 29E, 29F, 29G, 29H, 29I, 29J, and 29K depict various user interfaces for displaying stored information according to embodiments of the present invention;

[0054] FIGS. 30A and 30B depict simplified user interfaces for displaying contents of one or more multimedia documents according to an embodiment of the present invention;

[0055] FIG. 31 depicts a simplified user interface that may be used to print contents of one or more multimedia documents or contents corresponding to ranges according to an embodiment of the present invention;

[0056] FIGS. 32A, 32B, and 32C depict pages printed according to styles selectable from the interface depicted in FIG. 31 according to an embodiment of the present invention;

[0057] FIGS. 33A and 33B depict pages printed using keyframe styles selectable from interface 31 depicted in FIG. 31 according to an embodiment of the present invention;

[0058] FIGS. 34A, 34B, and 34C depict examples of coversheets that may be printed according to an embodiment of the present invention; and

[0059] FIGS. 35A, 35B, 35C, 35D, and 35E depict a paper document printed for ranges according to an embodiment of the present invention.

#### DETAILED DESCRIPTION OF THE INVENTION

[0060] Embodiments of the present invention provide techniques for retrieving and displaying multimedia infor-

mation. According to an embodiment of the present invention, a graphical user interface (GUI) is provided that displays multimedia information that may be stored in a multimedia document. According to the teachings of the present invention, the GUI enables a user to navigate through multimedia information stored in a multimedia document. The GUI provides both a focused and a contextual view of the contents of the multimedia document. The GUI thus allows a user to “horizontally” read or skim multimedia documents.

**[0061]** As indicated above, the term “multimedia information” is intended to refer to information that comprises information of several different types. The different types of information included in multimedia information may include a combination of text information, graphics information, animation information, sound (audio) information, video information, slides information, whiteboard images information, and other types of information. For example, a video recording of a television broadcast may comprise video information and audio information. In certain instances the video recording may also comprise close-captioned (CC) text information which comprises material related to the video information, and in many cases, is an exact representation of the speech contained in the audio portions of the video recording. Multimedia information is also used to refer to information comprising one or more objects wherein the objects include information of different types. For example, multimedia objects included in multimedia information may comprise text information, graphics information, animation information, sound (audio) information, video information, slides information, whiteboard images information, and other types of information.

**[0062]** The term “multimedia document” as used in this application is intended to refer to any electronic storage unit (e.g., a file, a directory, etc.) that stores multimedia information. Various different formats may be used to store the multimedia information. These formats include various MPEG formats (e.g., MPEG 1, MPEG 2, MPEG 4, MPEG 7, etc.), MP3 format, SMIL format, HTML+TIME format, WMF (Windows Media Format), RM (Real Media) format, Quicktime format, Shockwave format, various streaming media formats, formats being developed by the engineering community, proprietary and customary formats, and others. Examples of multimedia documents include video recordings, MPEG files, news broadcast recordings, presentation recordings, recorded meetings, classroom lecture recordings, broadcast television programs, or the like.

**[0063]** FIG. 1 is a simplified block diagram of a distributed network 100 that may incorporate an embodiment of the present invention. As depicted in FIG. 1, distributed network 100 comprises a number of computer systems including one or more client systems 102, a server system 104, and a multimedia information source (MIS) 106 coupled to communication network 108 via a plurality of communication links 110. Distributed network 100 depicted in FIG. 1 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. For example, the present invention may also be embodied in a stand-alone system. In a stand-alone

environment, the functions performed by the various computer systems depicted in FIG. 1 may be performed by a single computer system.

**[0064]** Communication network 108 provides a mechanism allowing the various computer systems depicted in FIG. 1 to communicate and exchange information with each other. Communication network 108 may itself be comprised of many interconnected computer systems and communication links. While in one embodiment, communication network 108 is the Internet, in other embodiments, communication network 108 may be any suitable communication network including a local area network (LAN), a wide area network (WAN), a wireless network, an intranet, a private network, a public network, a switched network, or the like.

**[0065]** Communication links 10 used to connect the various systems depicted in FIG. 1 may be of various types including hardwire links, optical links, satellite or other wireless communications links, wave propagation links, or any other mechanisms for communication of information. Various communication protocols may be used to facilitate communication of information via the communication links. These communication protocols may include TCP/IP, HTTP protocols, extensible markup language (XML), wireless application protocol (WAP), protocols under development by industry standard organizations, vendor-specific protocols, customized protocols, and others.

**[0066]** Computer systems connected to communication network 108 may be classified as “clients” or “servers” depending on the role the computer systems play with respect to requesting information and/or services or providing information and/or services. Computer systems that are used by users to request information or to request a service are classified as “client” computers (or “clients”). Computer systems that store information and provide the information in response to a user request received from a client computer, or computer systems that perform processing to provide the user-requested services are called “server” computers (or “servers”). It should however be apparent that a particular computer system may function both as a client and as a server.

**[0067]** Accordingly, according to an embodiment of the present invention, server system 104 is configured to perform processing to facilitate generation of a GUI that displays multimedia information according to the teachings of the present invention. The GUI generated by server system 104 may be output to the user (e.g., a reader of the multimedia document) via an output device coupled to server system 104 or via client systems 102. The GUI generated by server 104 enables the user to retrieve and browse multimedia information that may be stored in a multimedia document. The GUI provides both a focused and a contextual view of the contents of a multimedia document and thus enables the multimedia document to be skimmed or read “horizontally.”

**[0068]** The processing performed by server system 104 to generate the GUI and to provide the various features according to the teachings of the present invention may be implemented by software modules executing on server system 104, by hardware modules coupled to server system 104, or combinations thereof. In alternative embodiments of the present invention, the processing may also be distributed between the various computer systems depicted in FIG. 1.

[0069] The multimedia information that is displayed in the GUI may be stored in a multimedia document that is accessible to server system 104. For example, the multimedia document may be stored in a storage subsystem of server system 104. The multimedia document may also be stored by other systems such as MIS 106 that are accessible to server 104. Alternatively, the multimedia document may be stored in a memory location accessible to server system 104.

[0070] In alternative embodiments, instead of accessing a multimedia document, server system 104 may receive a stream of multimedia information (e.g., a streaming media signal, a cable signal, etc.) from a multimedia information source such as MIS 106. According to an embodiment of the present invention, server system 104 stores the multimedia information signals in a multimedia document and then generates a GUI that displays the multimedia information. Examples of MIS 106 include a television broadcast receiver, a cable receiver, a digital video recorder (e.g., a TIVO box), or the like. For example, multimedia information source 106 may be embodied as a television that is configured to receive multimedia broadcast signals and to transmit the signals to server system 104. In alternative embodiments, server system 104 may be configured to intercept multimedia information signals received by MIS 106. Server system 104 may receive the multimedia information directly from MIS 106 or may alternatively receive the information via a communication network such as communication network 108.

[0071] As described above, MIS 106 depicted in FIG. 1 represents a source of multimedia information. According to an embodiment of the present invention, MIS 106 may store multimedia documents that are accessed by server system 104. For example, MIS 106 may be a storage device or a server that stores multimedia documents that may be accessed by server system 104. In alternative embodiments, MIS 106 may provide a multimedia information stream to server system 104. For example, MIS 106 may be a television receiver/antenna providing live television feed information to server system 104. MIS 106 may be a device such as a video recorder/player, a DVD player, a CD player, etc. providing recorded video and/or audio stream to server system 104. In alternative embodiments, MIS 106 may be a presentation or meeting recorder device that is capable of providing a stream of the captured presentation or meeting information to server system 104. MIS 106 may also be a receiver (e.g., a satellite dish or a cable receiver) that is configured to capture or receive (e.g., via a wireless link) multimedia information from an external source and then provide the captured multimedia information to server system 104 for further processing.

[0072] Users may use client systems 102 to view the GUI generated by server system 104. Users may also use client systems 102 to interact with the other systems depicted in FIG. 1. For example, a user may use user system 102 to select a particular multimedia document and request server system 104 to generate a GUI displaying multimedia information stored by the particular multimedia document. A user may also interact with the GUI generated by server system 104 using input devices coupled to client system 102. In alternative embodiments, client system 102 may also perform processing to facilitate generation of a GUI according to the teachings of the present invention. A client system 102 may be of different types including a personal computer, a

portable computer, a workstation, a computer terminal, a network computer, a mainframe, a kiosk, a personal digital assistant (PDA), a communication device such as a cell phone, or any other data processing system.

[0073] According to an embodiment of the present invention, a single computer system may function both as server system 104 and as client system 102. Various other configurations of the server system 104, client system 102, and MIS 106 are possible.

[0074] FIG. 2 is a simplified block diagram of a computer system 200 according to an embodiment of the present invention. Computer system 200 may be used as any of the computer systems depicted in FIG. 1. As shown in FIG. 2, computer system 200 includes at least one processor 202, which communicates with a number of peripheral devices via a bus subsystem 204. These peripheral devices may include a storage subsystem 206, comprising a memory subsystem 208 and a file storage subsystem 210, user interface input devices 212, user interface output devices 214, and a network interface subsystem 216. The input and output devices allow user interaction with computer system 200. A user may be a human user, a device, a process, another computer, or the like. Network interface subsystem 216 provides an interface to other computer systems and communication networks.

[0075] Bus subsystem 204 provides a mechanism for letting the various components and subsystems of computer system 200 communicate with each other as intended. The various subsystems and components of computer system 200 need not be at the same physical location but may be distributed at various locations within network 100. Although bus subsystem 204 is shown schematically as a single bus, alternative embodiments of the bus subsystem may utilize multiple busses.

[0076] User interface input devices 212 may include a keyboard, pointing devices, a mouse, trackball, touchpad, a graphics tablet, a scanner, a barcode scanner, a touchscreen incorporated into the display, audio input devices such as voice recognition systems, microphones, and other types of input devices. In general, use of the term "input device" is intended to include all possible types of devices and ways to input information using computer system 200.

[0077] User interface output devices 214 may include a display subsystem, a printer, a fax machine, or non-visual displays such as audio output devices. The display subsystem may be a cathode ray tube (CRT), a flat-panel device such as a liquid crystal display (LCD), a projection device, or the like. The display subsystem may also provide non-visual display such as via audio output devices. In general, use of the term "output device" is intended to include all possible types of devices and ways to output information from computer system 200. According to an embodiment of the present invention, the GUI generated according to the teachings of the present invention may be presented to the user via output devices 214.

[0078] Storage subsystem 206 may be configured to store the basic programming and data constructs that provide the functionality of the computer system and of the present invention. For example, according to an embodiment of the present invention, software modules implementing the functionality of the present invention may be stored in storage

subsystem **206** of server system **104**. These software modules may be executed by processor(s) **202** of server system **104**. In a distributed environment, the software modules may be stored on a plurality of computer systems and executed by processors of the plurality of computer systems. Storage subsystem **206** may also provide a repository for storing various databases that may be used by the present invention. Storage subsystem **206** may comprise memory subsystem **208** and file storage subsystem **210**.

[**0079**] Memory subsystem **208** may include a number of memories including a main random access memory (RAM) **218** for storage of instructions and data during program execution and a read only memory (ROM) **220** in which fixed instructions are stored. File storage subsystem **210** provides persistent (non-volatile) storage for program and data files, and may include a hard disk drive, a floppy disk drive along with associated removable media, a Compact Disk Read Only Memory (CD-ROM) drive, an optical drive, removable media cartridges, and other like storage media. One or more of the drives may be located at remote locations on other connected computers.

[**0080**] Computer system **200** can be of varying types including a personal computer, a portable computer, a workstation, a computer terminal, a network computer, a mainframe, a kiosk, a personal digital assistant (PDA), a communication device such as a cell phone, or any other data processing system. Server computers generally have more storage and processing capacity than client systems. Due to the ever-changing nature of computers and networks, the description of computer system **200** depicted in **FIG. 2** is intended only as a specific example for purposes of illustrating the preferred embodiment of the computer system. Many other configurations of a computer system are possible having more or fewer components than the computer system depicted in **FIG. 2**.

[**0081**] **FIG. 3** depicts a simplified user interface **300** generated according to an embodiment of the present invention for viewing multimedia information. It should be apparent that GUI **300** depicted in **FIG. 3** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[**0082**] GUI **300** displays multimedia information stored in a multimedia document. The multimedia information stored by the multimedia document and displayed by GUI **300** may comprise information of a plurality of different types. As depicted in **FIG. 3**, GUI **300** displays multimedia information corresponding to a television broadcast that includes video information, audio information, and possibly closed-caption (CC) text information. The television broadcast may be stored as a television broadcast recording in a memory location accessible to server system **104**. It should however be apparent that the present invention is not restricted to displaying television recordings. Multimedia information comprising other types of information may also be displayed according to the teachings of the present invention.

[**0083**] The television broadcast may be stored using a variety of different techniques. According to one technique, the television broadcast is recorded and stored using a satellite receiver connected to a PC-TV video card of server system **104**. Applications executing on server system **104**

then process the recorded television broadcast to facilitate generation of GUI **300**. For example, the video information contained in the television broadcast may be captured using an MPEG capture application that creates a separate metafile (e.g., in XML format) containing temporal information for the broadcast and closed-caption text, if provided. Information stored in the metafile may then be used to generate GUI **300** depicted in **FIG. 3**.

[**0084**] As depicted in **FIG. 3**, GUI **300** comprises several viewing areas including a first viewing area **302**, a second viewing area **304**, a third viewing area **306**, a fourth viewing area **308**, and a fifth viewing area **310**. It should be apparent that in alternative embodiments the present invention may comprise more or fewer viewing areas than those depicted in **FIG. 3**. Further, in alternative embodiments of the present invention one or more viewing areas may be combined into one viewing area, or a particular viewing area may be divided in multiple viewing areas. Accordingly, the viewing areas depicted in **FIG. 3** and described below are not meant to restrict the scope of the present invention as recited in the claims.

[**0085**] According to an embodiment of the present invention, first viewing area **302** displays one or more commands that may be selected by a user viewing GUI **300**. Various user interface features such as menu bars, drop-down menus, cascading menus, buttons, selection bars, buttons, etc. may be used to display the user-selectable commands. According to an embodiment of the present invention, the commands provided in first viewing area **302** include a command that enables the user to select a multimedia document whose multimedia information is to be displayed in the GUI. The commands may also include one or more commands that allow the user to configure and/or customize the manner in which multimedia information stored in the user-selected multimedia document is displayed in GUI **300**. Various other commands may also be provided in first viewing area **302**.

[**0086**] According to an embodiment of the present invention, second viewing area **304** displays a scaled representation of multimedia information stored by the multimedia document. The user may select the scaling factor used for displaying information in second viewing area **304**. According to a particular embodiment of the present invention, a representation of the entire (i.e., multimedia information between the start time and end time associated with the multimedia document) multimedia document is displayed in second viewing area **304**. In this embodiment, one end of second viewing area **304** represents the start time of the multimedia document and the opposite end of second viewing area **304** represents the end time of the multimedia document.

[**0087**] As shown in **FIG. 3**, according to an embodiment of the present invention, second viewing area **304** comprises one or more thumbnail images **312**. Each thumbnail image displays a representation of a particular type of information included in the multimedia information stored by the multimedia document. For example, two thumbnail images **312-1** and **312-2** are displayed in second viewing area **304** of GUI **300** depicted in **FIG. 3**. Thumbnail image **312-1** displays text information corresponding to information included in the multimedia information stored by the multimedia document being displayed by GUI **300**. The text displayed in thumbnail image **312-1** may represent a dis-

playable representation of CC text included in the multimedia information displayed by GUI 300. Alternatively, the text displayed in thumbnail image 312-1 may represent a displayable representation of a transcription of audio information included in the multimedia information stored by the multimedia document whose contents are displayed by GUI 300. Various audio-to-text transcription techniques may be used to generate a transcript for the audio information. The text displayed in a thumbnail image may also be a representation of other types of information included in the multimedia information. For example, the text information may be a representation of comments made when the multimedia information was recorded or viewed, annotations added to the multimedia information, etc.

[0088] Thumbnail image 312-2 displays a representation of video information included in the multimedia information displayed by GUI 300. In the embodiment depicted in FIG. 3, the video information is displayed using video keyframes extracted from the video information included in the multimedia information stored by the multimedia document. The video keyframes may be extracted from the video information in the multimedia document at various points in time using a specified sampling rate. A special layout style, which may be user-configurable, is used to display the extracted keyframes in thumbnail image 312-2 to enhance readability of the frames.

[0089] One or more thumbnail images may be displayed in second viewing area 304 based upon the different types of information included in the multimedia information being displayed. Each thumbnail image 312 displayed in second viewing area 304 displays a representation of information of a particular type included in the multimedia information stored by the multimedia document. According to an embodiment of the present invention, the number of thumbnails displayed in second viewing area 304 and the type of information displayed by each thumbnail is user-configurable.

[0090] According to an embodiment of the present invention, the various thumbnail images displayed in second viewing area 304 are temporally synchronized or aligned with each other along a timeline. This implies that the various types of information included in the multimedia information and occurring at approximately the same time are displayed next to each other. For example, thumbnail images 312-1 and 312-2 are aligned such that the text information (which may represent CC text information, a transcript of the audio information, or a text representation of some other type of information included in the multimedia information) displayed in thumbnail image 312-1 and video keyframes displayed in thumbnail 312-2 that occur in the multimedia information at a particular point in time are displayed close to each other (e.g., along the same horizontal axis). Accordingly, information that has a particular time stamp is displayed proximal to information that has approximately the same time stamp. This enables a user to determine the various types of information occurring approximately concurrently in the multimedia information being displayed by GUI 300 by simply scanning second viewing area 304 in the horizontal axis.

[0091] According to the teachings of the present invention, a viewing lens or window 314 (hereinafter referred to as "thumbnail viewing area lens 314") is displayed in second

viewing area 304. Thumbnail viewing area lens 314 covers or emphasizes a portion of second viewing area 304. According to the teachings of the present invention, multimedia information corresponding to the area of second viewing area 304 covered by thumbnail viewing area lens 314 is displayed in third viewing area 306.

[0092] In the embodiment depicted in FIG. 3, thumbnail viewing area lens 314 is positioned at the top of second viewing area 304 and emphasizes a top portion (or starting portion) of the multimedia document. The position of thumbnail viewing area lens 314 may be changed by a user by sliding or moving lens 314 along second viewing area 304. For example, in FIG. 3, thumbnail viewing area lens 314 may be moved vertically along second viewing area 304.

[0093] In response to a change in the position of thumbnail viewing area lens 314 from a first location in second viewing area 304 to a second location along second viewing area 304, the multimedia information displayed in third viewing area 306 is automatically updated such that the multimedia information displayed in third viewing area 306 continues to correspond to the area of second viewing area 304 emphasized by thumbnail viewing area lens 314. Accordingly, a user may use thumbnail viewing area lens 314 to navigate and scroll through the contents of the multimedia document displayed by GUI 300. Thumbnail viewing area lens 314 thus provides a context and indicates a location of the multimedia information displayed in third viewing area 306 within the entire multimedia document.

[0094] FIG. 4 is a zoomed-in simplified diagram of thumbnail viewing area lens 314 according to an embodiment of the present invention. As depicted in FIG. 4, thumbnail viewing area lens 314 is bounded by a first edge 318 and a second edge 320. Thumbnail viewing area lens 314 emphasizes an area of second viewing area 304 between edge 318 and edge 320. Based upon the position of thumbnail viewing area lens 314 over second viewing area 304, edge 318 corresponds to specific time " $t_1$ " in the multimedia document and edge 320 corresponds to a specific time " $t_2$ " in the multimedia document wherein  $t_2 > t_1$ . For example, when thumbnail viewing area lens 314 is positioned at the start of second viewing area 304 (as depicted in FIG. 3),  $t_1$  may correspond to the start time of the multimedia document being displayed, and when thumbnail viewing area lens 314 is positioned at the end of second viewing area 304,  $t_2$  may correspond to the end time of the multimedia document. Accordingly, thumbnail viewing area lens 314 emphasizes a portion of second viewing area 304 between times  $t_1$  and  $t_2$ . According to an embodiment of the present invention, multimedia information corresponding to the time segment between  $t_2$  and  $t_1$  (which is emphasized or covered by thumbnail viewing area lens 314) is displayed in third viewing area 306. Accordingly, when the position of thumbnail viewing area lens 314 is changed along second viewing area 304 in response to user input, the information displayed in third viewing area 306 is updated such that the multimedia information displayed in third viewing area 306 continues to correspond to the area of second viewing area 304 emphasized by thumbnail viewing area lens 314.

[0095] As shown in FIG. 4 and FIG. 3, thumbnail viewing area lens 314 comprises a sub-lens 316 which further emphasizes a sub-portion of the portion of second viewing

area **304** emphasized by thumbnail viewing area lens **314**. According to an embodiment the present invention, the portion of second viewing area **304** emphasized or covered by sub-lens **316** corresponds to the portion of third viewing area **306** emphasized by lens **322**. Sub-lens **316** can be moved along second viewing area **304** within edges **318** and **320** of thumbnail viewing area lens **314**. When sub-lens **316** is moved from a first location to a second location within the boundaries of thumbnail viewing area lens **314**, the position of lens **322** in third viewing area **306** is also automatically changed to correspond to the changed location of sub-lens **316**. Further, if the position of lens **322** is changed from a first location to a second location over third viewing area **306**, the position of sub-lens **316** is also automatically updated to correspond to the changed position of lens **322**. Further details related to lens **322** are described below.

[0096] As described above, multimedia information corresponding to the portion of second viewing area **304** emphasized by thumbnail viewing area lens **314** is displayed in third viewing area **306**. Accordingly, a representation of multimedia information occurring between time  $t_1$  and  $t_2$  (corresponding to a segment of time of the multimedia document emphasized by thumbnail viewing area lens **314**) is displayed in third viewing area **306**. Third viewing area **306** thus displays a zoomed-in representation of the multimedia information stored by the multimedia document corresponding to the portion of the multimedia document emphasized by thumbnail viewing area lens **314**.

[0097] As depicted in FIG. 3, third viewing area **306** comprises one or more panels **324**. Each panel displays a representation of information of a particular type included in the multimedia information occurring during the time segment emphasized by thumbnail viewing area lens **314**. For example, in GUI **300** depicted in FIG. 3, two panels **324-1** and **324-2** are displayed in third viewing area **306**. According to an embodiment of the present invention, each panel **324** in third viewing area **306** corresponds to a thumbnail image **312** displayed in second viewing area **304** and displays information corresponding to the section of the thumbnail image covered by thumbnail viewing area lens **314**.

[0098] Like thumbnail images **312**, panels **324** are also temporally aligned or synchronized with each other. Accordingly, the various types of information included in the multimedia information and occurring at approximately the same time are displayed next to each other in third viewing area **306**. For example, panels **324-1** and **324-2** depicted in FIG. 3 are aligned such that the text information (which may represent CC text information, a transcript of the audio information, or a text representation of some other type of information included in the multimedia information) displayed in panel **324-1** and video keyframes displayed in panel **324-2** that occur in the multimedia information at a approximately the same point in time are displayed close to each other (e.g., along the same horizontal axis). Accordingly, information that has a particular time stamp is displayed proximal to other types of information that has approximately the same time stamp. This enables a user to determine the various types of information occurring approximately concurrently in the multimedia information by simply scanning third viewing area **306** in the horizontal axis.

[0099] Panel **324-1** depicted in GUI **300** corresponds to thumbnail image **312-1** and displays text information cor-

responding to the area of thumbnail image **312-1** emphasized or covered by thumbnail viewing area lens **314**. The text information displayed by panel **324-1** may correspond to text extracted from CC information included in the multimedia information, or alternatively may represent a transcript of audio information included in the multimedia information, or a text representation of some other type of information included in the multimedia information. According to an embodiment of the present invention, the present invention takes advantage of the automatic story segmentation and other features that are often provided in close-captioned (CC) text from broadcast news. Most news agencies who provide CC text as part of their broadcast use a special syntax in the CC text (e.g., a ">>>" delimiter to indicate changes in story line or subject, a ">>" delimiter to indicate changes in speakers, etc.). Given the presence of this kind of information in the CC text information included in the multimedia information, the present invention incorporates these features in the text displayed in panel **324-1**. For example, a ">>>" delimiter may be displayed to indicate changes in story line or subject, a ">>" delimiter may be displayed to indicate changes in speakers, additional spacing may be displayed between text portions related to different story lines to clearly demarcate the different stories, etc. This enhances the readability of the text information displayed in panel **324-1**.

[0100] Panel **324-2** depicted in GUI **300** corresponds to thumbnail image **312-2** and displays a representation of video information corresponding to the area of thumbnail image **312-2** emphasized or covered by thumbnail viewing area lens **314**. Accordingly, panel **324-2** displays a representation of video information included in the multimedia information stored by the multimedia document and occurring between times  $t_1$  and  $t_2$  associated with thumbnail viewing area lens **314**. In the embodiment depicted in FIG. 3, video keyframes extracted from the video information included in the multimedia information are displayed in panel **324-2**. A special layout style (which is user-configurable) is used to display the extracted keyframes to enhance readability of the frames.

[0101] Various different techniques may be used to display video keyframes in panel **324-2**. According to an embodiment of the present invention, the time segment between time  $t_1$  and time  $t_2$  is divided into sub-segments of a predetermined time period. Each sub-segment is characterized by a start time and an end time associated with the sub-segment. According to an embodiment of the present invention, the start time of the first sub-segment corresponds to time  $t_1$  while the end time of the last sub-segment corresponds to time  $t_2$ . Server **104** then extracts a set of one or more video keyframes from the video information stored by the multimedia document for each sub-segment occurring between the start time and end time associated with the sub-segment. For example, according to an embodiment of the present invention, for each sub-segment, server **104** may extract a video keyframe at 1-second intervals between a start time and an end time associated with the sub-segment.

[0102] For each sub-segment, server **104** then selects one or more keyframes from the set of extracted video keyframes for the sub-segment to be displayed in panel **324-2**. The number of keyframes selected to be displayed in panel **324-2** for each sub-segment is user-configurable. Various different techniques may be used for selecting the video

keyframes to be displayed from the extracted set of video keyframes for each time sub-segment. For example, if the set of video keyframes extracted for a sub-segment comprises 24 keyframes and if six video keyframes are to be displayed for each sub-segment (as shown in **FIG. 3**), server **104** may select the first two video keyframes, the middle two video keyframes, and the last two video keyframes from the set of extracted video keyframes for the sub-segment.

[0103] In another embodiment, the video keyframes to be displayed for a sub-segment may be selected based upon the sequential positions of the keyframes in the set of keyframes extracted for sub-segment. For example, if the set of video keyframes extracted for a sub-segment comprises 24 keyframes and if six video keyframes are to be displayed for each sub-segment, then the 1st, 5th, 9th, 13th, 17th, and 21st keyframe may be selected. In this embodiment, a fixed number of keyframes are skipped.

[0104] In yet another embodiment, the video keyframes to be displayed for a sub-segment may be selected based upon time values associated with the keyframes in the set of keyframes extracted for sub-segment. For example, if the set of video keyframes extracted for a sub-segment comprises 24 keyframes extracted at a sampling rate of 1 second and if six video keyframes are to be displayed for each sub-segment, then the first frame may be selected and subsequently a keyframe occurring 4 seconds after the previously selected keyframe may be selected.

[0105] In an alternative embodiment of the present invention, server **104** may select keyframes from the set of keyframes based upon differences in the contents of the keyframes. For each sub-segment, server **104** may use special image processing techniques to determine differences in the contents of the keyframes extracted for the sub-segment. If six video keyframes are to be displayed for each sub-segment, server **104** may then select six keyframes from the set of extracted keyframes based upon the results of the image processing techniques. For example, the six most dissimilar keyframes may be selected for display in panel **324-2**. It should be apparent that various other techniques known to those skilled in the art may also be used to perform the selection of video keyframes.

[0106] The selected keyframes are then displayed in panel **324-2**. Various different formats may be used to display the selected keyframes in panel **324-2**. For example, as shown in **FIG. 3**, for each sub-segment, the selected keyframes are laid out left-to-right and top-to-bottom.

[0107] In an alternative embodiment of the present invention, the entire multimedia document is divided into sub-segments of a pre-determined time period. Each sub-segment is characterized by a start time and an end time associated with the sub-segment. According to an embodiment of the present invention, the start time of the first sub-segment corresponds to the start time of the multimedia document while the end time of the last sub-segment corresponds to the end time of the multimedia document. As described above, server **104** then extracts a set of one or more video keyframes from the video information stored by the multimedia document for each sub-segment based upon the start time and end time associated with the sub-segment. Server **104** then selects one or more keyframes for display for each sub-segment. Based upon the position of thumbnail viewing area lens **314**, keyframes that have been selected for

display and that occur between  $t_1$  and  $t_2$  associated with thumbnail viewing area lens **314** are then displayed in panel **324-2**.

[0108] It should be apparent that various other techniques may also be used for displaying video information in panel **324-2** in alternative embodiments of the present invention. According to an embodiment of the present invention, the user may configure the technique to be used for displaying video information in third viewing area **306**.

[0109] In GUI **300** depicted in **FIG. 3**, each sub-segment is 8 seconds long and video keyframes corresponding to a plurality of sub-segments are displayed in panel **324-2**. Six video keyframes are displayed from each sub-segment. For each sub-segment, the displayed keyframes are laid out in a left-to-right and top-to-bottom manner.

[0110] It should be apparent that, in alternative embodiments of the present invention, the number of panels displayed in third viewing area **306** may be more or less than the number of thumbnail images displayed in second viewing area **304**. According to an embodiment of the present invention, the number of panels displayed in third viewing area **306** is user-configurable.

[0111] According to the teachings of the present invention, a viewing lens or window **322** (hereinafter referred to as "panel viewing area lens **322**") is displayed covering or emphasizing a portion of overview region **306**. According to the teachings of the present invention, multimedia information corresponding to the area of third viewing area **306** emphasized by panel viewing area lens **322** is displayed in fourth viewing area **308**. A user may change the position of panel viewing area lens **322** by sliding or moving lens **322** along third viewing area **306**. In response to a change in the position of panel viewing area lens **322** from a first location in third viewing area **306** to a second location, the multimedia information displayed in fourth viewing area **308** is automatically updated such that the multimedia information displayed in fourth viewing area **308** continues to correspond to the area of third viewing area **306** emphasized by panel viewing area lens **322**. Accordingly, a user may use panel viewing area lens **322** to change the multimedia information displayed in fourth viewing area **308**.

[0112] As described above, a change in the location of panel viewing area lens **322** also causes a change in the location of sub-lens **316** such that the area of second viewing area **304** emphasized by sub-lens **316** continues to correspond to the area of third viewing area **306** emphasized by panel viewing area lens **322**. Likewise, as described above, a change in the location of sub-lens **316** also causes a change in the location of panel viewing area lens **322** over third viewing area **306** such that the area of third viewing area **306** emphasized by panel viewing area lens **322** continues to correspond to the changed location of sub-lens **316**.

[0113] **FIG. 5A** is a zoomed-in simplified diagram of panel viewing area lens **322** according to an embodiment of the present invention. As depicted in **FIG. 5A**, panel viewing area lens **322** is bounded by a first edge **326** and a second edge **328**. Panel viewing area lens **322** emphasizes an area of third viewing area **306** between edge **326** and edge **328**. Based upon the position of panel viewing area lens **322** over third viewing area **306**, edge **326** corresponds to specific time " $t_3$ " in the multimedia document and edge **328** corre-

sponds to a specific time “ $t_1$ ” in the multimedia document where  $t_1 > t_3$  and  $(t_1 \leq t_3 < t_4 \leq t_2)$ . For example, when panel viewing area lens 322 is positioned at the start of third viewing area 306,  $t_3$  may be equal to  $t_1$ , and when panel viewing area lens 322 is positioned at the end of third viewing area 306,  $t_4$  may be equal to  $t_2$ . Accordingly, panel viewing area lens 322 emphasizes a portion of third viewing area 306 between times  $t_3$  and  $t_4$ . According to an embodiment of the present invention, multimedia information corresponding to the time segment between  $t_3$  and  $t_4$  (which is emphasized or covered by panel viewing area lens 322) is displayed in fourth viewing area 308. When the position of panel viewing area lens 322 is changed along third viewing area 306 in response to user input, the information displayed in fourth viewing area 308 may be updated such that the multimedia information displayed in fourth viewing area 308 continues to correspond to the area of third viewing area 306 emphasized by panel viewing area lens 322. Third viewing area 306 thus provides a context and indicates the location of the multimedia information displayed in fourth viewing area 308 within the multimedia document.

[0114] According to an embodiment of the present invention, a particular line of text (or one or more words from the last line of text) emphasized by panel viewing area lens 322 may be displayed on a section of lens 322. For example, as depicted in FIGS. 5A and 3, the last line of text 330 “Environment is a national” that is emphasized by panel viewing area lens 322 in panel 324-1 is displayed in bolded style on panel viewing area lens 322.

[0115] According to an embodiment of the present invention, special features may be attached to panel viewing area lens 322 to facilitate browsing and navigation of the multimedia document. As shown in FIG. 5A, a “play/pause button” 332 and a “lock/unlock button” 334 are provided on panel viewing area lens 322 according to an embodiment of the present invention. Play/Pause button 332 allows the user to control playback of the video information from panel viewing area lens 322. Lock/Unlock button 334 allows the user to switch the location of the video playback from area 340-1 of fourth viewing area 308 to a reduced window on top of panel viewing area lens 322.

[0116] FIG. 5B is a simplified example of panel viewing area lens 322 with its lock/unlock button 334 activated or “locked” (i.e., the video playback is locked onto panel viewing area lens 322) according to an embodiment of the present invention. As depicted in FIG. 5B, in the locked mode, the video information is played back on a window 336 on lens 322. In the embodiment depicted in FIG. 5B, the portion of panel viewing area lens 322 over panel 342-2 is expanded in size beyond times  $t_3$  and  $t_4$  to accommodate window 336. According to an embodiment of the present invention, the video contents displayed in window 336 correspond to the contents displayed in area 340-1 of fourth viewing area 308.

[0117] According to an embodiment of the present invention, window 336 has transparent borders so that portions of the underlying third viewing area 306 (e.g., the keyframes displayed in panel 324-2) can be seen. This helps to maintain the user’s location focus while viewing third viewing area 306. The user may use play/pause button 332 to start and stop the video displayed in window 336. The user may change the location of panel viewing area lens 322 while the

video is being played back in window 336. A change in the location of panel viewing area lens 322 causes the video played back in window 336 to change corresponding to the new location of panel viewing area lens 322. The video played back in window 336 corresponds to the new time values  $t_3$  and  $t_4$  associated with panel viewing area lens 322.

[0118] FIG. 5C is a simplified example of panel viewing area lens 322 wherein a representative video keyframe is displayed on panel viewing area lens 322 according to an embodiment of the present invention. In this embodiment server 104 analyzes the video keyframes of panel 324-2 emphasized or covered by panel viewing area lens 322 and determines a particular keyframe 338 that is most representative of the keyframes emphasized by panel viewing area lens 322. The particular keyframe is then displayed on a section of panel viewing area lens 322 covering panel 324-2. In the embodiment depicted in FIG. 5C, the portion of panel viewing area lens 322 over panel 342-2 is expanded in size beyond times  $t_3$  and  $t_4$  to accommodate display of keyframe 338.

[0119] As described above, multimedia information corresponding to the section of third viewing area 306 covered by panel viewing area lens 322 (i.e., multimedia information occurring in the time segment between  $t_3$  and  $t_4$ ) is displayed in fourth viewing area 308. As depicted in FIG. 3, fourth viewing area 308 may comprise one or more sub viewing areas 340 (e.g., 340-1, 340-2, and 340-3). According to an embodiment of the present invention, one or more of sub-regions 340 may display a particular type of information included in the multimedia information corresponding to the section of third viewing area 306 emphasized by panel viewing area lens 322.

[0120] For example, as depicted in FIG. 3, video information corresponding to (or starting from) the video information emphasized by panel viewing area lens 322 in third viewing area 306 is displayed in sub viewing area 340-1. According to an embodiment of the present invention, video information starting at time  $t_3$  (time corresponding to the top edge of panel viewing area lens 322) may be played back in sub viewing area 340-1. In alternative embodiments, the video information played back in area 340-1 may start at time  $t_4$  or some other user-configurable time between  $t_3$  and  $t_4$ . The playback of the video in sub viewing area 340-1 may be controlled using control bar 342. Control bar 342 provides a plurality of controls including controls for playing, pausing, stopping, rewinding, and forwarding the video played in sub viewing area 340-1. The current time and length 344 of the video being played in area 340-1 is also displayed. Information identifying the name of the video 346, the date 348 the video was recorded, and the type of the video 350 is also displayed.

[0121] In alternative embodiments of the present invention, instead of playing back video information, a video keyframe from the video keyframes emphasized by panel viewing area lens 322 in panel 324-2 is displayed in sub viewing area 340-1. According to an embodiment of the present invention, the keyframe displayed in area 340-1 represents a keyframe that is most representative of the keyframes emphasized by panel viewing area lens 322.

[0122] According to an embodiment of the present invention, text information (e.g., CC text, transcript of audio information, text representation of some other type of infor-

mation included in the multimedia information, etc.) emphasized by panel viewing area lens **322** in third viewing area **306** is displayed in sub viewing area **340-2**. According to an embodiment of the present invention, sub viewing area **340-2** displays text information that is displayed in panel **324-1** and emphasized by panel viewing area lens **322**. As described below, various types of information may be displayed in sub viewing area **340-3**.

[0123] Additional information related to the multimedia information stored by the multimedia document may be displayed in fifth viewing area **310** of GUI **300**. For example, as depicted in **FIG. 3**, words occurring in the text information included in the multimedia information displayed by GUI **300** are displayed in area **352** of fifth viewing area **310**. The frequency of each word in the multimedia document is also displayed next to each word. For example, the word "question" occurs seven times in the multimedia information CC text. Various other types of information related to the multimedia information may also be displayed in fifth viewing area **310**.

[0124] According to an embodiment of the present invention, GUI **300** provides features that enable a user to search for one or more words that occur in the text information (e.g., CC text, transcript of audio information, a text representation of some other type of information included in the multimedia information) extracted from the multimedia information. For example, a user can enter one or more query words in input field **354** and upon selecting "Find" button **356**, server **104** analyzes the text information extracted from the multimedia information stored by the multimedia document to identify all occurrences of the one or more query words entered in field **354**. The occurrences of the one or more words in the multimedia document are then highlighted when displayed in second viewing area **304**, third viewing area **306**, and fourth viewing area **308**. For example, according to an embodiment of the present invention, all occurrences of the query words are highlighted in thumbnail image **312-1**, in panel **324-1**, and in sub viewing area **340-2**. In alternative embodiments of the present invention, occurrences of the one or more query words may also be highlighted in the other thumbnail images displayed in second viewing area **304**, panels displayed in third viewing area **306**, and sub viewing areas displayed in fourth viewing area **308**.

[0125] The user may also specify one or more words to be highlighted in the multimedia information displayed in GUI **300**. For example, a user may select one or more words to be highlighted from area **352**. All occurrences of the key-words selected by the user in area **352** are then highlighted in second viewing area **304**, third viewing area **306**, and fourth viewing area **308**. For example, as depicted in **FIG. 6**, the user has selected the word "National" in area **352**. In response to the user's selection, according to an embodiment of the present invention, all occurrences of the word "National" are highlighted in second viewing area **304**, third viewing area **306**, and third viewing area **306**.

[0126] According to an embodiment of the present invention, lines of text **360** that comprise the user-selected word(s) (or query words entered in field **354**) are displayed in sub viewing area **340-3** of fourth viewing area **308**. For each line of text, the time **362** when the line occurs (or the timestamp associated with the line of text) in the multimedia

document is also displayed. The timestamp associated with the line of text generally corresponds to the timestamp associated with the first word in the line.

[0127] For each line of text, one or more words surrounding the selected or query word(s) are displayed. According to an embodiment of the present invention, the number of words surrounding a selected word that is displayed in area **340-3** is user configurable. For example, in GUI **300** depicted in **FIG. 6**, a user can specify the number of surrounding words to be displayed in area **340-3** using control **364**. The number specified by the user indicates the number of words that occur before the select word and the number of words that occur after the selected word that are to be displayed. In the embodiment depicted in **FIG. 6**, control **364** is a slider bar that can be adjusted between a minimum value of "3" and a maximum value of "10". The user can specify the number of surrounding words to be displayed by adjusting slider bar **364**. For example, if the slider bar is set to "3", then three words that occur before a selected word and three words that occur after the selected word will be displayed in area **340-3**. The minimum and maximum values are user configurable.

[0128] Further, GUI **300** depicted in **FIG. 6** comprises an area **358** sandwiched between thumbnail images **312-1** and **312-2** that indicates locations of occurrences of the query words or other words specified by the user. For example, area **358** comprises markers indicating the locations of word "National" in thumbnail image **312-1**. The user can then use either thumbnail viewing area lens **314**, or panel viewing area lens **322** to scroll to a desired location within the multimedia document. **FIG. 7** depicts a simplified zoomed-in view of second viewing area **304** showing area **358** according to an embodiment of the present invention. As depicted in **FIG. 7**, area **358** (or channel **358**) comprises markers **360** indicating locations in thumbnail image **312-1** that comprise occurrences of the word "National". In alternative embodiments of the present invention, markers in channel **358** may also identify locations of the user-specified words or phrases in the other thumbnail images displayed in second viewing area **304**. In alternative embodiments, locations of occurrences of the query words or other words specified by the user may be displayed on thumbnail images **312** (as depicted in **FIG. 20A**).

[0129] As shown in **FIG. 6**, the position of thumbnail viewing area lens **314** has been changed with respect to **FIG. 3**. In response to the change in position of thumbnail viewing area lens **314**, the multimedia information displayed in third viewing area **306** has been changed to correspond to the section of second viewing area **304** emphasized by thumbnail viewing area lens **314**. The multimedia information displayed in fourth viewing area **308** has also been changed corresponding to the new location of panel viewing area lens **322**.

[0130] According to an embodiment of the present invention, multimedia information displayed in GUI **300** that is relevant to user-specified topics of interest is highlighted or annotated. The annotations or highlights provide visual indications of information that is relevant to or of interest to the user. GUI **300** thus provides a convenient tool that allows a user to readily locate portions of the multimedia document that are relevant to the user.

[0131] According to an embodiment of the present invention, information specifying topics that are of interest or are

relevant to the user may be stored in a user profile. One or more words or phrases may be associated with each topic of interest. Presence of the one or more words and phrases associated with a particular user-specified topic of interest indicates presence of information related to the particular topic. For example, a user may specify two topics of interest—"George W. Bush" and "Energy Crisis". Words or phrases associated with the topic "George Bush" may include "President Bush," "the President," "Mr. Bush," and other like words and phrases. Words or phrases associated with the topic "Energy Crisis" may include "industrial pollution," "natural pollution," "clean up the sources," "amount of pollution," "air pollution," "electricity," "power-generating plant," or the like. Probability values may be associated with each of the words or phrases indicating the likelihood of the topic of interest given the presence of the word or phrase. Various tools may be provided to allow the user to configure topics of interest, to specify keywords and phrases associated with the topics, and to specify probability values associated with the keywords or phrases.

[0132] It should be apparent that various other techniques known to those skilled in the art may also be used to model topics of interest to the user. These techniques may include the use of Bayesian networks, relevance graphs, or the like. Techniques for determining sections relevant to user-specified topics, techniques for defining topics of interest, techniques for associating keywords and/or key phrases and probability values are described in U.S. application Ser. No. 08/995,616, filed Dec. 22, 1997, the entire contents of which are herein incorporated by reference for all purposes.

[0133] According to an embodiment of the present invention, in order to identify locations in the multimedia document related to user-specified topics of interest, server 104 searches the multimedia document to identify locations within the multimedia document of words or phrases associated with the topics of interest. As described above, presence of words and phrases associated with a particular user-specified topic of interest in the multimedia document indicate presence of the particular topic relevant to the user. The words and phrases that occur in the multimedia document and that are associated with user specified topics of interest are annotated or highlighted when displayed by GUI 300.

[0134] FIG. 8 depicts an example of a simplified GUI 800 in which multimedia information that is relevant to one or more topics of interest to a user is highlighted (or annotated) when displayed in GUI 800 according to an embodiment of the present invention. GUI 800 depicted in FIG. 8 is merely illustrative of an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0135] In the embodiment depicted in FIG. 8, the user has specified four topics of interest 802. A label 803 identifies each topic. The topics specified in GUI 800 include "Energy Crisis," "Assistive Tech," "George W. Bush," and "Nepal." In accordance with the teachings of the present invention, keywords and key phrases relevant to the specified topics are highlighted in second viewing area 304, third viewing area 306, and fourth viewing area 308. Various different techniques may be used to highlight or annotate the keywords and/or key phrases related to the topics of interest. Accord-

ing to an embodiment of the present invention, different colors and styles (e.g., bolding, underlining, different font size, etc.) may be used to highlight words and phrases related to user-specified topics. For example, each topic may be assigned a particular color and content related to a particular topic might be highlighted using the particular color assigned to the particular topic. For example, as depicted in FIG. 8, a first color is used to highlight words and phrases related to the "Energy Crisis" topic of interest, a second color is used to highlight words and phrases related to the "Assistive Tech" topic of interest, a third color is used to highlight words and phrases related to the "George W. Bush" topic of interest, and a fourth color is used to highlight words and phrases related to the "Nepal" topic of interest.

[0136] According to an embodiment of the present invention, server 104 searches the text information (e.g., CC text, transcript of audio information, or a text representation of some other type of information included in the multimedia information) extracted from the multimedia information to locate words or phrases relevant to the user topics. If server 104 finds a word or phrase in the text information that is associated with a topic of interest, the word or phrase is annotated or highlighted when displayed in GUI 800. As described above, several different techniques may be used to annotate or highlight the word or phrase. For example, the word or phrase may be highlighted, bolded, underlined, demarcated using sidebars or balloons, font may be changed, etc.

[0137] Keyframes (representing video information of the multimedia document) that are displayed by the GUI and that are related to user specified topics of interest may also be highlighted. According to an embodiment of the present invention, server system 104 may use OCR techniques to extract text from the keyframes extracted from the video information included in the multimedia information. The text output of the OCR techniques may then be compared with words or phrases associated with one or more user-specified topics of interest. If there is a match, the keyframe containing the matched word or phrase (i.e., the keyframe from which the matching word or phrase was extracted by OCR techniques) may be annotated or highlighted when the keyframe is displayed in GUI 800 either in second viewing area 304, third viewing area 306, or fourth viewing area 308 of GUI 800. Several different techniques may be used to annotate or highlight the keyframe. For example, a special box may be drawn around a keyframe that is relevant to a particular topic of interest. The color of the box may correspond to the color associated with the particular topic of interest. The matching text in the keyframe may also be highlighted or underlined or displayed in reverse video. As described above, the annotated or highlighted keyframes displayed in second viewing area 304 (e.g., the keyframes displayed in thumbnail image 312-2 in FIG. 3) may be identified by markers displayed in channel area 358. In alternative embodiments, the keyframes may be annotated or highlighted in thumbnail image 312-2.

[0138] According to an embodiment of the present invention, as shown in FIG. 8, a relevance indicator 804 may also be displayed for each user topic. For a particular topic, the relevance indicator for the topic indicates the degree of relevance (or a relevancy score) of the multimedia document to the particular topic. For example, as shown in FIG. 8, the number of bars displayed in a relevance indicator associated

with a particular topic indicates the degree of relevance of the multimedia document to the particular topic. Accordingly, the multimedia document displayed in GUI 800 is most relevant to user topic "Energy Crisis" (as indicated by four bars) and least relevant to user topic "Nepal" (indicated by one bar). Various other techniques (e.g., relevance scores, bar graphs, different colors, etc.) may also be used to indicate the degree of relevance of each topic to the multimedia document.

[0139] According to an embodiment of the present invention, the relevancy score for a particular topic may be calculated based upon the frequency of occurrences of the words and phrases associated with the particular topic in the multimedia information. Probability values associated with the words or phrases associated with the particular topic may also be used to calculate the relevancy score for the particular topic. Various techniques known to those skilled in the art may also be used to determine relevancy scores for user specified topics of interest based upon the frequency of occurrences of words and phrases associated with a topic in the multimedia information and the probability values associated with the words or phrases. Various other techniques known to those skilled in the art may also be used to calculate the degree of relevancy of the multimedia document to the topics of interest.

[0140] As previously stated, a relevance indicator is used to display the degree or relevancy or relevancy score to the user. Based upon information displayed by the relevance indicator, a user can easily determine relevance of multimedia information stored by a multimedia document to topics that may be specified by the user.

[0141] FIG. 9 depicts a simplified user interface 900 for defining a topic of interest according to an embodiment of the present invention. User interface 900 may be invoked by selecting an appropriate command from first viewing area 302. GUI 900 depicted in FIG. 9 is merely illustrative of an embodiment of the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0142] A user may specify a topic of interest in field 902. A label identifying the topic of interest can be specified in field 910. The label specified in field 910 is displayed in the GUI generated according to the teachings of the present invention to identify the topic of interest. A list of keywords and/or phrases associated with the topic specified in field 902 is displayed in area 908. A user may add new keywords to the list, modify one or more keywords in the list, or remove one or more keywords from the list of keywords associated with the topic of interest. The user may specify new keywords or phrases to be associated with the topic of interest in field 904. Selection of "Add" button 906 adds the keywords or phrases specified in field 904 to the list of keywords previously associated with a topic. The user may specify a color to be used for annotating or highlighting information relevant to the topic of interest by selecting the color in area 912. For example, in the embodiment depicted in FIG. 9, locations in the multimedia document related to "Assistive Technology" will be annotated or highlighted in blue color.

[0143] According to the teachings of the present invention, various different types of information included in

multimedia information may be displayed by the GUI generated by server 104. FIG. 10 depicts a simplified user interface 1000 that displays multimedia information stored by a meeting recording according to an embodiment of the present invention. It should be apparent that GUI 1000 depicted in FIG. 10 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0144] The multimedia information stored by the meeting recording may comprise video information, audio information and possibly CC text information, slides information, and other type of information. The slides information may comprise information related to slides (e.g., a PowerPoint presentation slides) presented during the meeting. For example, slides information may comprise images of slides presented at the meeting. As shown in FIG. 10, second viewing area 304 comprises three thumbnail images 312-1, 312-2, and 312-3. Text information (e.g., CC text information, a transcript of audio information included in the meeting recording, or a text representation of some other type of information included in the meeting recording) extracted from the meeting recording multimedia information is displayed in thumbnail image 312-1. Video keyframes extracted from the video information included in the meeting recording multimedia information are displayed in thumbnail image 312-2. Slides extracted from the slides information included in the multimedia information are displayed in thumbnail image 312-3. The thumbnail images are temporally aligned with one another. The information displayed in thumbnail image 312-4 provides additional context for the video and text information in that, the user can view presentation slides that were presented at various times throughout the meeting recording.

[0145] Third viewing area 306 comprises three panels 324-1, 324-2, and 324-3. Panel 324-1 displays text information corresponding to the section of thumbnail image 312-1 emphasized or covered by thumbnail viewing area lens 314. Panel 324-2 displays video keyframes corresponding to the section of thumbnail image 312-2 emphasized or covered by thumbnail viewing area lens 314. Panel 324-3 displays one or more slides corresponding to the section of thumbnail image 312-3 emphasized or covered by thumbnail viewing area lens 314. The panels are temporally aligned with one another.

[0146] Fourth viewing area 308 comprises three sub-viewing areas 340-1, 340-2, and 340-3. Sub viewing area 340-1 displays video information corresponding to the section of panel 324-2 covered by panel viewing area lens 322. As described above, sub-viewing area 340-1 may display a keyframe corresponding to the emphasized portion of panel 324-2. Alternatively, video based upon the position of panel viewing area lens 322 may be played back in area 340-1. According to an embodiment of the present invention, time  $t_3$  associated with lens 322 is used as the start time for playing the video in area 340-1 of fourth viewing area 308. A panoramic shot 1002 of the meeting room (which may be recorded using a 360 degrees camera) is also displayed in area 340-1 of fourth viewing area 308. Text information emphasized by panel viewing area lens 322 in panel 324-1 is displayed in area 340-2 of fourth viewing area 308. One or more slides emphasized by panel viewing area lens 322

in panel 324-3 are displayed in area 340-3 of fourth viewing area 308. According to an embodiment of the present invention, the user may also select a particular slide from panel 324-3 by clicking on the slide. The selected slide is then displayed in area 340-3 of fourth viewing area 308.

[0147] According to an embodiment of the present invention, the user can specify the types of information included in the multimedia document that are to be displayed in the GUI. For example, the user can turn on or off slides related information (i.e., information displayed in thumbnail 312-3, panel 324-3, and area 340-3 of fourth viewing area 308) displayed in GUI 1000 by selecting or deselecting "Slides" button 1004. If a user deselects slides information, then thumbnail 312-3 and panel 324-3 are not displayed by GUI 1000. Thumbnail 312-3 and panel 324-3 are displayed by GUI 1000 if the user selects button 1004. Button 1004 thus acts as a switch for displaying or not displaying slides information. In a similar manner, the user can also control other types of information displayed by a GUI generated according to the teachings of the present invention. For example, features may be provided for turning on or off video information, text information, and other types of information that may be displayed by GUI 1000.

[0148] FIG. 11 depicts a simplified user interface 1100 that displays multimedia information stored by a multimedia document according to an embodiment of the present invention. It should be apparent that GUI 1100 depicted in FIG. 11 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0149] The multimedia document whose contents are displayed in GUI 1100 comprises video information, audio information or CC text information, slides information, and whiteboard information. The whiteboard information may comprise images of text and drawings drawn on a whiteboard. As shown in FIG. 11, second viewing area 304 comprises four thumbnail images 312-1, 312-2, 312-3, and 312-4. Text information (e.g., CC text information, or a transcript of audio information included in the meeting recording, or a text representation of some other type of information included in the multimedia information) extracted from the multimedia document is displayed in thumbnail image 312-1. Video keyframes extracted from the video information included in the multimedia document are displayed in thumbnail image 312-2. Slides extracted from the slides information included in the multimedia information are displayed in thumbnail image 312-3. Whiteboard images extracted from the whiteboard information included in the multimedia document are displayed in thumbnail image 312-4. The thumbnail images are temporally aligned with one another.

[0150] Third viewing area 306 comprises four panels 324-1, 324-2, 324-3, and 324-4. Panel 324-1 displays text information corresponding to the section of thumbnail image 312-1 emphasized or covered by thumbnail viewing area lens 314. Panel 324-2 displays video keyframes corresponding to the section of thumbnail image 312-2 emphasized or covered by thumbnail viewing area lens 314. Panel 324-3 displays one or more slides corresponding to the section of thumbnail image 312-3 emphasized or covered by thumbnail

viewing area lens 314. Panel 324-4 displays one or more whiteboard images corresponding to the section of thumbnail image 312-4 emphasized or covered by thumbnail viewing area lens 314. The panels are temporally aligned with one another.

[0151] Fourth viewing area 308 comprises three sub-viewing areas 340-1, 340-2, and 340-3. Area 340-1 displays video information corresponding to the section of panel 324-2 covered by panel viewing area lens 322. As described above, sub-viewing area 340-1 may display a keyframe or play back video corresponding to the emphasized portion of panel 324-2. According to an embodiment of the present invention, time  $t_3$  (as described above) associated with lens 322 is used as the start time for playing the video in area 340-1 of fourth viewing area 308. A panoramic shot 1102 of the location where the multimedia document was recorded (which may be recorded using a 360 degrees camera) is also displayed in area 340-1 of fourth viewing area 308. Text information emphasized by panel viewing area lens 322 in panel 324-1 is displayed in area 340-2 of fourth viewing area 308. Slides emphasized by panel viewing area lens 322 in panel 324-3 or whiteboard images emphasized by panel viewing area lens 322 in panel 324-4 may be displayed in area 340-3 of fourth viewing area 308. In the embodiment depicted in FIG. 11, a whiteboard image corresponding to the section of panel 324-4 covered by panel viewing area lens 322 is displayed in area 340-3. According to an embodiment of the present invention, the user may also select a particular slide from panel 324-3 or select a particular whiteboard image from panel 324-4 by clicking on the slide or whiteboard image. The selected slide or whiteboard image is then displayed in area 340-3 of fourth viewing area 308.

[0152] As described above, according to an embodiment of the present invention, the user can specify the types of information from the multimedia document that are to be displayed in the GUI. For example, the user can turn on or off a particular type of information displayed by the GUI. "WB" button 1104 allows the user to turn on or off whiteboard related information (i.e., information displayed in thumbnail image 312-4, panel 324-4, and area 340-3 of fourth viewing area 308) displayed in GUI 1000.

[0153] FIG. 12 depicts a simplified user interface 1200 that displays contents of a multimedia document according to an embodiment of the present invention. It should be apparent that GUI 1200 depicted in FIG. 12 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0154] As depicted in FIG. 12, preview areas 1202 and 1204 are provided at the top and bottom of third viewing area 306. In this embodiment, panel viewing area lens 322 can be moved along third viewing area 306 between edge 1206 of preview area 1202 and edge 1208 of preview area 1204. Preview areas 1202 and 1204 allow the user to preview the contents displayed in third viewing area 306 when the user scrolls the multimedia document using panel viewing area lens 322. For example, as the user is scrolling down the multimedia document using panel viewing area lens 322, the user can see upcoming contents in preview area 1204 and see the contents leaving third viewing area 306 in

preview area **1202**. If the user is scrolling up the multimedia document using panel viewing area lens **322**, the user can see upcoming contents in preview area **1202** and see the contents leaving third viewing area **306** in preview area **1204**. According to an embodiment of the present invention, the size (or length) of each preview region can be changed and customized by the user. For example, in GUI **1200** depicted in **FIG. 12**, a handle **1210** is provided that can be used by the user to change the size of preview region **1204**. According to an embodiment of the present invention, preview areas may also be provided in second viewing area **304**.

[0155] **FIG. 13** depicts a simplified user interface **1300** that displays contents of a multimedia document according to an embodiment of the present invention. It should be apparent that GUI **1300** depicted in **FIG. 13** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0156] As depicted in **FIG. 13**, text information is displayed in panel **324-1** of third viewing area **306** in compressed format, i.e., the white spaces between the text lines have been removed. This enhances the readability of the text information. The lines of text displayed in panel **324-1** are then used to determine the video frames to be displayed in panel **324-2**. According to an embodiment of the present invention, a timestamp is associated with each line of text displayed in panel **324-1**. The timestamp associated with a line of text represents the time when the text occurred in the multimedia document being displayed by GUI **1300**. In one embodiment, the timestamp associated with a line of text corresponds to the timestamp associated with the first word in the line of text. The lines of text displayed in panel **324-1** are then grouped into groups, with each group comprising a pre-determined number of lines.

[0157] Video keyframes are then extracted from the video information stored by the multimedia document for each group of lines depending on time stamps associated with lines in the group. According to an embodiment of the present invention, server **104** determines a start time and an end time associated with each group of lines. A start time for a group corresponds to a time associated with the first (or earliest) line in the group while an end time for a group corresponds to the time associated with the last line (or latest) line in the group. In order to determine keyframes to be displayed in panel **324-2** corresponding to a particular group of text lines, server **104** extracts a set of one or more video keyframes from the portion of the video information occurring between the start and end time associated with the particular group. One or more keyframes are then selected from the extracted set of video keyframes to be displayed in panel **324-2** for the particular group. The one or more selected keyframes are then displayed in panel **324-1** proximal to the group of lines displayed in panel **324-1** for which the keyframes have been extracted.

[0158] For example, in **FIG. 13**, the lines displayed in panel **324-1** are divided into groups wherein each group comprises 4 lines of text. For each group, the time stamp associated with the first line in the group corresponds to the start time for the group while the time stamp associated with the fourth line in the group corresponds to the end time for

the group of lines. Three video keyframes are displayed in panel **324-2** for each group of four lines of text displayed in panel **324-1** in the embodiment depicted in **FIG. 13**. According to an embodiment of the present invention, the three video keyframes corresponding to a particular group of lines correspond to the first, middle, and last keyframe from the set of keyframes extracted from the video information between the start and end times of the particular group. As described above, various other techniques may also be used to select the video keyframes that are displayed in panel **324-2**. For each group of lines displayed in panel **324-1**, the keyframes corresponding to the group of lines are displayed such that the keyframes are temporally aligned with the group of lines. In the embodiment depicted in **FIG. 13**, the height of keyframes for a group of lines is approximately equal to the vertical height of the group of lines.

[0159] The number of text lines to be included in a group is user configurable. Likewise, the number of video keyframes to be extracted for a particular group of lines is also user configurable. Further, the video keyframes to be displayed in panel **324-2** for each group of lines can also be configured by the user of the present invention.

[0160] The manner in which the extracted keyframes are displayed in panel **324-2** is also user configurable. Different techniques may be used to show the relationships between a particular group of lines and video keyframes displayed for the particular group of lines. For example, according to an embodiment of the present invention, a particular group of lines displayed in panel **324-1** and the corresponding video keyframes displayed in panel **324-2** may be color-coded or displayed using the same color to show the relationship. Various other techniques known to those skilled in the art may also be used to show the relationships.

[0161] GUI Generation Technique According to an Embodiment of the Present Invention

[0162] The following section describes techniques for generating a GUI (e.g., GUI **300** depicted in **FIG. 3**) according to an embodiment of the present invention. For purposes of simplicity, it is assumed that the multimedia information to be displayed in the GUI comprises video information, audio information, and CC text information. The task of generating GUI **300** can be broken down into the following tasks: (a) displaying thumbnail **312-1** displaying text information extracted from the multimedia information in second viewing area **304**; (b) displaying thumbnail **312-2** displaying video keyframes extracted from the video information included in the multimedia information; (c) displaying thumbnail viewing area lens **314** emphasizing a portion of second viewing area **304** and displaying information corresponding to the emphasized portion of second viewing area **304** in third viewing area **306**, and displaying panel viewing area lens **322** emphasizing a portion of third viewing area **306** and displaying information corresponding to the emphasized portion of third viewing area **306** in fourth viewing area **308**; and (d) displaying information in fifth viewing area **310**.

[0163] **FIG. 14** is a simplified high-level flowchart **1400** depicting a method of displaying thumbnail **312-1** in second viewing area **304** according to an embodiment of the present invention. The method depicted in **FIG. 14** may be performed by server **104**, by client **102**, or by server **104** and client **102** in combination. For example, the method may be

executed by software modules executing on server **104** or on client **102**, by hardware modules coupled to server **104** or to client **102**, or combinations thereof. In the embodiment described below, the method is performed by server **104**. The method depicted in **FIG. 14** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[**0164**] As depicted in **FIG. 14**, the method is initiated when server **104** accesses multimedia information to be displayed in the GUI (step **1402**). As previously stated, the multimedia information may be stored in a multimedia document accessible to server **104**. As part of step **1402**, server **104** may receive information (e.g., a filename of the multimedia document) identifying the multimedia document and the location (e.g., a directory path) of the multimedia document. A user of the present invention may provide the multimedia document identification information. Server **104** may then access the multimedia document based upon the provided information. Alternatively, server **104** may receive the multimedia information to be displayed in the GUI in the form of a streaming media signal, a cable signal, etc. from a multimedia information source. Server system **104** may then store the multimedia information signals in a multimedia document and then use the stored document to generate the GUI according to the teachings of the present invention.

[**0165**] Server **104** then extracts text information from the multimedia information accessed in step **1402** (step **1404**). If the multimedia information accessed in step **1402** comprises CC text information, then the text information corresponds to CC text information that is extracted from the multimedia information. If the multimedia information accessed in step **1402** does not comprise CC text information, then in step **1404**, the audio information included in the multimedia information accessed in step **1402** is transcribed to generate a text transcript for the audio information. The text transcript represents the text information extracted in step **1404**. The text information extracted in step **1404** may also be a text representation of some other type of information included in the multimedia information.

[**0166**] The text information determined in step **1404** comprises a collection of lines with each line comprising one or more words. Each word has a timestamp associated with it indicating the time of occurrence of the word in the multimedia information. The timestamp information for each word is included in the CC text information. Alternatively, if the text represents a transcription of audio information, the timestamp information for each word may be determined during the audio transcription process. Alternatively, if the text information represents a text representation of some other type of information included in the multimedia information, then the time stamp associated with the other type of information may be determined.

[**0167**] As part of step **1404**, each line is assigned a start time and an end time based upon words that are included in the line. The start time for a line corresponds to the timestamp associated with the first word occurring in the line, and the end time for a line corresponds to the timestamp associated with the last word occurring in the line.

[**0168**] The text information determined in step **1404**, including the timing information, is then stored in a memory

location accessible to server **104** (step **1406**). In one embodiment, a data structure (or memory structure) comprising a linked list of line objects is used to store the text information. Each line object comprises a linked list of words contained in the line. Timestamp information associated with the words and the lines is also stored in the data structure. The information stored in the data structure is then used to generate GUI **300**.

[**0169**] Server **104** then determines a length or height (in pixels) of a panel (hereinafter referred to as "the text canvas") for drawing the text information (step **1408**). In order to determine the length of the text canvas, the duration ("duration") of the multimedia information (or the duration of the multimedia document storing the multimedia document) in seconds is determined. A vertical pixels-per-second of time ("pps") value is also defined. The "pps" determines the distance between lines of text drawn in the text canvas. The value of pps thus depends on how close the user wants the lines of text to be to each other when displayed and upon the size of the font to be used for displaying the text. According to an embodiment of the present invention, a 5 pps value is specified with a 6 point font. The overall height (in pixels) of the text canvas ("textCanvasHeight") is determined as follows:

$$\text{textCanvasHeight} = \text{duration} * \text{pps}$$

[**0170**] For example, if the duration of the multimedia information is 1 hour (i.e., 3600 seconds) and for apps value of 5, the height of the text canvas (textCanvasHeight) is 18000 pixels (3600\*5).

[**0171**] Multipliers are then calculated for converting pixel locations in the text canvas to seconds and for converting seconds to pixels locations in the text canvas (step **1410**). A multiplier "pix\_m" is calculated for converting a given time value (in seconds) to a particular vertical pixel location in the text canvas. The pix\_m multiplier can be used to determine a pixel location in the text canvas corresponding to a particular time value. The value of pix\_m is determined as follows:

$$\text{pix\_m} = \text{textCanvasHeight} / \text{duration}$$

[**0172**] For example, if duration=3600 seconds and textCanvasHeight=18000 pixels, then  $\text{pix\_m} = 18000 / 3600 = 5$ .

[**0173**] A multiplier "sec\_m" is calculated for converting a particular pixel location in the text canvas to a corresponding time value. The sec\_m multiplier can be used to determine a time value for a particular pixel location in the text canvas. The value of sec\_m is determined as follows:

$$\text{sec\_m} = \text{duration} / \text{textCanvasHeight}$$

[**0174**] For example, if duration=3600 seconds and textCanvasHeight=18000 pixels, then  $\text{sec\_m} = 3600 / 18000 = 0.2$ .

[**0175**] The multipliers calculated in step **1410** may then be used to convert pixels to seconds and seconds to pixels. For example, the pixel location in the text canvas of an event occurring at time  $t = 1256$  seconds in the multimedia information is:  $1256 * \text{pix\_m} = 1256 * 5 = 6280$  pixels from the top of the text canvas. The number of seconds corresponding to a pixel location  $p = 231$  in the text canvas is:  $231 * \text{sec\_m} = 231 * 0.2 = 46.2$  seconds.

[**0176**] Based upon the height of the text canvas determined in step **1408** and the multipliers generated in step **1410**, positional coordinates (horizontal (X) and vertical (Y)

coordinates) are then calculated for words in the text information extracted in step 1404 (step 1412). As previously stated, information related to words and lines and their associated timestamps may be stored in a data structure accessible to server 104. The positional coordinate values calculated for each word might also be stored in the data structure.

[0177] The Y (or vertical) coordinate ( $W_y$ ) for a word is calculated by multiplying the timestamp ( $W_t$ ) (in seconds) associated with the word by multiplier  $\text{pix\_m}$  determined in step 1410. Accordingly:

$$W_y(\text{in pixels}) = W_t * \text{pix\_m}$$

[0178] For example, if a particular word has  $W_t=539$  seconds (i.e., the word occurs 539 seconds into the multimedia information), then  $W_y=539*5=2695$  vertical pixels from the top of the text canvas.

[0179] The X (or horizontal) coordinate ( $W_x$ ) for a word is calculated based upon the word's location in the line and the width of the previous words in the line. For example if a particular line (L) has four words, i.e., L:  $W_1 W_2 W_3 W_4$ , then

$$W_x \text{ of } W_1 = 0$$

$$W_x \text{ of } W_2 = (W_x \text{ of } W_1) + (\text{Width of } W_1) + (\text{Spacing between words})$$

$$W_x \text{ of } W_3 = (W_x \text{ of } W_2) + (\text{Width of } W_2) + (\text{Spacing between words})$$

$$W_x \text{ of } W_4 = (W_x \text{ of } W_3) + (\text{Width of } W_3) + (\text{Spacing between words})$$

[0180] The words in the text information are then drawn on the text canvas in a location determined by the X and Y coordinates calculated for the words in step 1412 (step 1414).

[0181] Server 104 then determines a height of thumbnail 312-1 that displays text information in second viewing area 304 of GUI 300 (step 1416). The height of thumbnail 312-1 (ThumbnailHeight) depends on the height of the GUI window used to displaying the multimedia information and the height of second viewing area 304 within the GUI window. The value of ThumbnailHeight is set such that thumbnail 312-1 fits in the GUI in the second viewing area 304.

[0182] Thumbnail 312-1 is then generated by scaling the text canvas such that the height of thumbnail 312-1 is equal to ThumbnailHeight and the thumbnail fits entirely within the size constraints of second viewing area 304 (step 1418). Thumbnail 312-1, which represents a scaled version of the text canvas, is then displayed in second viewing area 304 of GUI 300 (step 1420).

[0183] Multipliers are then calculated for converting pixel locations in thumbnail 312-1 to seconds and for converting seconds to pixel locations in thumbnail 312-1 (step 1422). A multiplier "tpix\_m" is calculated for converting a given time value (in seconds) to a particular pixel location in thumbnail 312-1. Multiplier  $\text{tpix\_m}$  can be used to determine a pixel location in the thumbnail corresponding to a particular time value. The value of  $\text{tpix\_M}$  is determined as follows:

$$\text{tpix\_M} = \text{ThumbnailHeight} / \text{duration}$$

[0184] For example, if duration=3600 seconds and ThumbnailHeight=900, then  $\text{tpix\_m}=900/3600=0.25$

[0185] A multiplier "tsec\_m" is calculated for converting a particular pixel location in thumbnail 312-1 to a corresponding time value. Multiplier  $\text{tsec\_m}$  can be used to determine a time value for a particular pixel location in thumbnail 312-1. The value of  $\text{tsec\_m}$  is determined as follows:

$$\text{tsec\_m} = \text{duration} / \text{ThumbnailHeight}$$

[0186] For example, if duration=3600 seconds and ThumbnailHeight=900, then  $\text{tsec\_m}=3600/900=4$ .

[0187] Multipliers  $\text{tpix\_m}$  and  $\text{tsec\_m}$  may then be used to convert pixels to seconds and seconds to pixels in thumbnail 312-1. For example, the pixel location in thumbnail 312-1 of a word occurring at time  $t=1256$  seconds in the multimedia information is:  $1256 * \text{tpix\_m} = 1256 * 0.25 = 314$  pixels from the top of thumbnail 312-1. The number of seconds represented by a pixel location  $p=231$  in thumbnail 312-1 is:  $231 * \text{tsec\_m} = 231 * 4 = 924$  seconds.

[0188] FIG. 15 is a simplified high-level flowchart 1500 depicting a method of displaying thumbnail 312-2, which depicts video keyframes extracted from the video information, in second viewing area 304 of GUI 300 according to an embodiment of the present invention. The method depicted in FIG. 15 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 15 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0189] For purposes of simplicity, it is assumed that thumbnail 312-1 displaying text information has already been displayed according to the flowchart depicted in FIG. 14. As depicted in FIG. 15, server 104 extracts a set of keyframes from the video information included in the multimedia information (step 1502). The video keyframes may be extracted from the video information by sampling the video information at a particular sampling rate. According to an embodiment of the present invention, keyframes are extracted from the video information at a sampling rate of 1 frame per second. Accordingly, if the duration of the multimedia information is 1 hour (3600 seconds), then 3600 video keyframes are extracted from the video information in step 1502. A timestamp is associated with each keyframe extracted in step 1502 indicating the time of occurrence of the keyframe in the multimedia information.

[0190] The video keyframes extracted in step 1502 and their associated timestamp information is stored in a data structure (or memory structure) accessible to server 104 (step 1504). The information stored in the data structure is then used for generating thumbnail 312-2.

[0191] The video keyframes extracted in step 1504 are then divided into groups (step 1506). A user-configurable time period ("groupTime") is used to divide the keyframes into groups. According to an embodiment of the present invention, groupTime is set to 8 seconds. In this embodiment, each group comprises video keyframes extracted

within an 8 second time period window. For example, if the duration of the multimedia information is 1 hour (3600 seconds) and 3600 video keyframes are extracted from the video information using a sampling rate of 1 frame per second, then if groupTime is set to 8 seconds, the 3600 keyframes will be divided into 450 groups, with each group comprising 8 video keyframes.

[0192] A start and an end time are calculated for each group of frames (step 1508). For a particular group of frames, the start time for the particular group is the timestamp associated with the first (i.e., the keyframe in the group with the earliest timestamp) video keyframe in the group, and the end time for the particular group is the timestamp associated with the last (i.e., the keyframe in the group with the latest timestamp) video keyframe in the group.

[0193] For each group of keyframes, server 104 determines a segment of pixels on a keyframe canvas for drawing one or more keyframes from the group of keyframes (step 1510). Similar to the text canvas, the keyframe canvas is a panel on which keyframes extracted from the video information are drawn. The height of the keyframe canvas ("keyframeCanvasHeight") is the same as the height of the text canvas ("textCanvasHeight") described above (i.e., keyframeCanvasHeight==textCanvasHeight). As a result, multipliers pix\_m and sec\_m (described above) may be used to convert a time value to a pixel location in the keyframe canvas and to convert a particular pixel location in the keyframe canvas to a time value.

[0194] The segment of pixels on the keyframe canvas for drawing keyframes from a particular group is calculated based upon the start time and end time associated with the particular group. The starting vertical (Y) pixel coordinate ("segmentStart") and the end vertical (Y) coordinate ("segmentEnd") of the segment of pixels in the keyframe canvas for a particular group of keyframes is calculated as follows:

$$\text{segmentStart} = (\text{Start time of group}) * \text{pix\_m}$$

$$\text{segmentEnd} = (\text{End time of group}) * \text{pix\_m}$$

[0195] Accordingly, the height of each segment ("segmentHeight") in pixels of the text canvas is:

$$\text{segmentHeight} = \text{segmentEnd} - \text{segmentStart}$$

[0196] The number of keyframes from each group of frames to be drawn in each segment of pixels on the text canvas is then determined (step 1512). The number of keyframes to be drawn on the keyframe canvas for a particular group depends on the height of the segment ("segmentHeight") corresponding to the particular group. If the value of segmentHeight is small only a small number of keyframes may be drawn in the segment such that the drawn keyframes are comprehensible to the user when displayed in the GUI. The value of segmentHeight depends on the value of pps. If pps is small, then segmentHeight will also be small. Accordingly, a larger value of pps may be selected if more keyframes are to be drawn per segment.

[0197] According to an embodiment of the present invention, if the segmentHeight is equal to 40 pixels and each group of keyframes comprises 8 keyframes, then 6 out of the 8 keyframes may be drawn in each segment on the text canvas. The number of keyframes to be drawn in a segment is generally the same for all groups of keyframes. For example, in the embodiment depicted in FIG. 3, six keyframes are drawn in each segment on the text canvas.

[0198] After determining the number of keyframes to be drawn in each segment of the text canvas, for each group of keyframes, server 104 identifies one or more keyframes from keyframes in the group of keyframes to be drawn on the keyframe canvas (step 1514). Various different techniques may be used for selecting the video keyframes to be displayed in a segment for a particular group of frames. According to one technique, if each group of video keyframes comprises 8 keyframes and if 6 video keyframes are to be displayed in each segment on the keyframe canvas, then server 104 may select the first two video keyframes, the middle two video keyframes, and the last two video keyframes from each group of video keyframes be drawn on the keyframe canvas. As described above, various other techniques may also be used to select one or more keyframes to display from the group of keyframes. For example, the keyframes may be selected based upon the sequential positions of the keyframes in the group of keyframes, based upon time values associated with the keyframes, or based upon other criteria.

[0199] According to another technique, server 104 may use special image processing techniques to determine similarity or dissimilarity between keyframes in each group of keyframes. If six video keyframes are to be displayed from each group, server 104 may then select six keyframes from each group of keyframes based upon the results of the image processing techniques. According to an embodiment of the present invention, the six most dissimilar keyframes in each group may be selected to be drawn on the keyframe canvas. It should be apparent that various other techniques known to those skilled in the art may also be used to perform the selection of video keyframes.

[0200] Keyframes from the groups of keyframes identified in step 1514 are then drawn on the keyframe canvas in their corresponding segments (step 1516). Various different formats may be used for drawing the selected keyframes in a particular segment. For example, as shown in FIG. 3, for each segment, the selected keyframes may be laid out left-to-right and top-to-bottom in rows of 3 frames. Various other formats known to those skilled in the art may also be used to draw the keyframes on the keyframe canvas. The size of each individual keyframe drawn on the keyframe canvas depends on the height (segmentHeight) of the segment in which the keyframe is drawn and the number of keyframes to be drawn in the segment. As previously stated, the height of a segment depends on the value of pps. Accordingly, the size of each individual keyframe drawn on the keyframe canvas also depends on the value of pps.

[0201] Server 104 then determines a height (or length) of thumbnail 312-2 that displays the video keyframes in GUI 300 (step 1518). According to the teachings of the present invention, the height of thumbnail 312-2 is set to be the same as the height of thumbnail 312-1 that displays text information (i.e., the height of thumbnail 312-2 is set to ThumbnailHeight).

[0202] Thumbnail 312-2 is then generated by scaling the keyframe canvas such that the height of thumbnail 312-2 is equal to ThumbnailHeight and thumbnail 312-2 fits entirely within the size constraints of second viewing area 304 (step 1520). Thumbnail 312-2, which represents a scaled version of the keyframe canvas, is then displayed in second viewing area 304 of GUI 300 (step 1522). Thumbnail 312-2 is

displayed in GUI 300 next to thumbnail image 312-1 and is temporally aligned or synchronized with thumbnail 312-1 (as shown in FIG. 3). Accordingly, the top of thumbnail 312-2 is aligned with the top of thumbnail 312-1.

[0203] Multipliers are calculated for thumbnail 312-2 for converting pixel locations in thumbnail 312-2 to seconds and for converting seconds to pixel locations in thumbnail 312-2 (step 1524). Since thumbnail 312-2 is the same length as thumbnail 312-1 and is aligned with thumbnail 312-1, multipliers “tpix\_m” and “tsec\_m” calculated for thumbnail 312-1 can also be used for thumbnail 312-2. These multipliers may then be used to convert pixels to seconds and seconds to pixels in thumbnail 312-2.

[0204] According to the method displayed in FIG. 15, the size of each individual video keyframe displayed in thumbnail 312-2 depends, in addition to other criteria, on the length of thumbnail 312-2 and on the length of the video information. Assuming that the length of thumbnail 312-2 is fixed, the height of each individual video keyframe displayed in thumbnail 312-2 is inversely proportional to the length of the video information. Accordingly, as the length of the video information increases, the size of each keyframe displayed in thumbnail 312-2 decreases. As a result, for longer multimedia documents, the size of each keyframe may become so small that the video keyframes displayed in thumbnail 312-2 are no longer recognizable by the user. To avoid this, various techniques may be used to display the video keyframes in thumbnail 312-2 in a manner that makes thumbnail 312-2 more readable and recognizable by the user.

[0205] FIG. 16 is a simplified high-level flowchart 1600 depicting another method of displaying thumbnail 312-2 according to an embodiment of the present invention. The method depicted in FIG. 16 maintains the comprehensibility and usability of the information displayed in thumbnail 312-2 by reducing the number of video keyframes drawn in the keyframe canvas and displayed in thumbnail 312-2. The method depicted in FIG. 16 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 16 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0206] As depicted in FIG. 16, steps 1602, 1604, 1606, and 1608 are the same as steps 1502, 1504, 1506, and 1508, depicted in FIG. 15 and explained above. After step 1608, one or more groups whose video keyframes are to be drawn in the keyframe canvas are then selected from the groups determined in step 1606 (step 1609). Various different techniques may be used to select the groups in step 1609. According to one technique, the groups determined in step 1606 are selected based upon a “SkipCount” value that is user-configurable. For example, if SkipCount is set to 4, then every fifth group (i.e., 4 groups are skipped) is selected in step 1609. The value of SkipCount may be adjusted based upon the length of the multimedia information. According to

an embodiment of the present invention, the value of SkipCount is directly proportional to the length of the multimedia information, i.e., SkipCount is set to a higher value for longer multimedia documents.

[0207] For each group selected in step 1609, server 104 identifies one or more keyframes from the group to be drawn on the keyframe canvas (step 1610). As described above, various techniques may be used to select keyframes to be drawn on the keyframe canvas.

[0208] The keyframe canvas is then divided into a number of equal-sized row portions, where the number of row portions is equal to the number of groups selected in step 1609 (step 1612). According to an embodiment of the present invention, the height of each row portion is approximately equal to the height of the keyframe canvas (“keyframeCanvasHeight”) divided by the number of groups selected in step 1609.

[0209] For each group selected in step 1609, a row portion of the keyframe canvas is then identified for drawing one or more video keyframes from the group (step 1614). According to an embodiment of the present invention, row portions are associated with groups in chronological order. For example, the first row is associated with a group with the earliest start time, the second row is associated with a group with the second earliest start time, and so on.

[0210] For each group selected in step 1609, one or more keyframes from the group (identified in step 1610) are then drawn on the keyframe canvas in the row portion determined for the group in step 1614 (step 1616). The sizes of the selected keyframes for each group are scaled to fit the row portion of the keyframe canvas. According to an embodiment of the present invention, the height of each row portion is more than the heights of the selected keyframes, and height of the selected keyframes is increased to fit the row portion. This increases the size of the selected keyframes and makes them more visible when drawn on the keyframe canvas. In this manner, keyframes from the groups selected in step 1609 are drawn on the keyframe canvas.

[0211] The keyframe canvas is then scaled to form thumbnail 312-2 that is displayed in second viewing area 304 according to steps 1618, 1620, and 1622. Since the height of the keyframes drawn on the keyframe canvas is increased according to an embodiment of the present invention, as described above, the keyframes are also more recognizable when displayed in thumbnail 312-2. Multipliers are then calculated according to step 1624. Steps 1618, 1620, 1622, and 1624 are similar to steps 1518, 1520, 1522, and 1524, depicted in FIG. 15 and explained above. As described above, by selecting a subset of the groups, the number of keyframes to be drawn on the keyframe canvas and displayed in thumbnail 312-2 is reduced. This in turn increases the height of each individual video keyframe displayed in thumbnail 312-2 thus making them more recognizable when displayed.

[0212] FIG. 17 is a simplified high-level flowchart 1700 depicting a method of displaying thumbnail viewing area lens 314, displaying information emphasized by thumbnail viewing area lens 314 in third viewing area 306, displaying panel viewing area lens 322, displaying information emphasized by panel viewing area lens 322 in fourth viewing area 308, and displaying information in fifth viewing area 310

according to an embodiment of the present invention. The method depicted in **FIG. 17** may be performed by server **104**, by client **102**, or by server **104** and client **102** in combination. For example, the method may be executed by software modules executing on server **104** or on client **102**, by hardware modules coupled to server **104** or to client **102**, or combinations thereof. In the embodiment described below, the method is performed by server **104**. The method depicted in **FIG. 17** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

**[0213]** As depicted in **FIG. 17**, server **104** first determines a height (in pixels) of each panel ("PanelHeight") to be displayed in third viewing area **306** of GUI **300** (step **1702**). The value of PanelHeight depends on the height (or length) of third viewing area **306**. Since the panels are to be aligned to each other, the height of each panel is set to PanelHeight. According to an embodiment of the present invention, PanelHeight is set to the same value as ThumbnailHeight. However, in alternative embodiments of the present invention, the value of PanelHeight may be different from the value of ThumbnailHeight.

**[0214]** A section of the text canvas (generated in the flowchart depicted in **FIG. 14**) equal to PanelHeight is then identified (step **1704**). The section of the text canvas identified in step **1704** is characterized by vertical pixel coordinate ( $P_{start}$ ) marking the starting pixel location of the section, and a vertical pixel coordinate ( $P_{end}$ ) marking the ending pixel location of the section.

**[0215]** Time values corresponding to the boundaries of the section of the text canvas identified in step **1704** (marked by pixel locations  $P_{start}$  and  $P_{end}$ ) are then determined (step **1706**). The multiplier  $sec\_m$  is used to calculate the corresponding time values. A time  $t_1$  (in seconds) corresponding to pixel location  $P_{start}$  is calculated as follows:

$$t_1 = P_{start} * sec\_m$$

**[0216]** A time  $t_2$  (in seconds) corresponding to pixel location  $P_{end}$  is calculated as follows:

$$t_2 = P_{end} * sec\_m$$

**[0217]** A section of the keyframe canvas corresponding to the selected section of the text canvas is then identified (step **1708**). Since the height of the keyframe canvas is the same as the height of the text canvas, the selected section of the keyframe canvas also lies between pixels locations  $P_{start}$  and  $P_{end}$  in the keyframe canvas corresponding to times  $t_1$  and  $t_2$ .

**[0218]** The portion of the text canvas identified in step **1704** is displayed in panel **324-1** in third viewing area **306** (step **1710**). The portion of the keyframe canvas identified in step **1708** is displayed in panel **324-2** in third viewing area **306** (step **1712**).

**[0219]** A panel viewing area lens **322** is displayed covering a section of third viewing area **306** (step **1714**). Panel viewing area lens **322** is displayed such that it emphasizes or covers a section of panel **324-1** and **324-2** displayed in third viewing area **306** between times  $t_3$  and  $t_4$  where ( $t_1 \leq t_3 < t_4 \leq t_2$ ). The top edge of panel viewing area lens **322** corresponds to time  $t_3$  and the bottom edge of panel viewing

area lens **322** corresponds to time  $t_4$ . The height of panel viewing area lens **322** (expressed in pixels) is equal to: (Vertical pixel location in the text canvas corresponding to  $t_4$ ) - (Vertical pixel location in the text canvas corresponding to  $t_3$ ). The width of panel viewing area lens **322** is approximately equal to the width of third viewing area **306** (as shown in **FIG. 3**).

**[0220]** A portion of thumbnail **312-1** corresponding to the section of text canvas displayed in panel **324-1** and a portion of thumbnail **312-2** corresponding to the section of keyframe canvas displayed in panel **324-2** are then determined (step **1716**). The portion of thumbnail **312-1** corresponding to the section of the text canvas displayed in panel **324-1** is characterized by vertical pixel coordinate ( $TN_{start}$ ) marking the starting pixel location of the thumbnail portion, and a vertical pixel coordinate ( $TN_{end}$ ) marking the ending pixel location of the thumbnail portion. The multiplier  $tpix\_m$  is used to determine pixel locations  $TN_{start}$  and  $TN_{end}$  as follows:

$$TN_{start} = t_1 * tpix\_m$$

$$TN_{end} = t_2 * tpix\_m$$

**[0221]** Since thumbnails **312-1** and **312-2** are of the same length and are temporally aligned to one another, the portion of thumbnail **312-2** corresponding to the sections of keyframe canvas displayed in panel **324-2** also lies between pixel locations  $TN_{start}$  and  $TN_{end}$  on thumbnail **312-2**.

**[0222]** Thumbnail viewing area lens **314** is then displayed covering portions of thumbnails **312-1** and **312-2** corresponding to the section of text canvas displayed in panel **324-1** and the section of keyframe canvas displayed in panel **324-2** (step **1718**). Thumbnail viewing area lens **314** is displayed covering portions of thumbnails **312-1** and **312-2** between pixels locations  $TN_{start}$  and  $TN_{end}$  of the thumbnails. The height of thumbnail viewing area lens **314** in pixels is equal to ( $TN_{end} - TN_{start}$ ). The width of thumbnail viewing area lens **314** is approximately equal to the width of second viewing area **304** (as shown in **FIG. 3**).

**[0223]** A portion of second viewing area **304** corresponding to the section of third viewing area **306** emphasized by panel viewing area lens **322** is then determined (step **1720**). In step **1720**, server **104** determines a portion of thumbnail **312-1** and a portion of thumbnail **312-2** corresponding to the time period between  $t_3$  and  $t_4$ . The portion of thumbnail **312-1** corresponding to the time window between  $t_3$  and  $t_4$  is characterized by vertical pixel coordinate ( $TNSub_{start}$ ) corresponding to time  $t_3$  and marking the starting vertical pixel of the thumbnail portion, and a vertical pixel coordinate ( $TNSub_{end}$ ) corresponding to time  $t_4$  and marking the ending vertical pixel location of the thumbnail portion. Multiplier  $tpix\_m$  is used to determine pixel locations  $TNSub_{start}$  and  $TNSub_{end}$  as follows:

$$TNSub_{start} = t_3 * tpix\_m$$

$$TNSub_{end} = t_4 * tpix\_m$$

**[0224]** Since thumbnails **312-1** and **312-2** are of the same length and are temporally aligned to one another, the portion of thumbnail **312-2** corresponding to the time period between  $t_3$  and  $t_4$  also lies between pixel locations  $TNSub_{start}$  and  $TNSub_{end}$  on thumbnail **312-2**.

**[0225]** Sub-lens **316** is then displayed covering portions of thumbnails **312-1** and **312-2** corresponding to the time

window between  $t_3$  and  $t_4$  (i.e., corresponding to the portion of third viewing area **306** emphasized by panel viewing area lens **322**) (step **1722**). Sub-lens **316** is displayed covering portions of thumbnails **312-1** and **312-2** between pixel locations  $TNSub_{start}$  and  $TNSub_{end}$ . The height of sub-lens **316** in pixels is equal to  $(TNSub_{end} - TNSub_{start})$ . The width of sub-lens **316** is approximately equal to the width of second viewing area **304** (as shown in **FIG. 3**).

[**0226**] Multimedia information corresponding to the portion of third viewing area **306** emphasized by panel viewing area lens **322** is displayed in fourth viewing area **308** (step **1724**). For example, video information starting at time  $t_3$  is played back in area **340-1** of fourth viewing area **308** in GUI **300**. In alternative embodiments, the starting time of the video playback may be set to any time between and including  $t_3$  and  $t_4$ . Text information corresponding to the time window between  $t_3$  and  $t_4$  is displayed in area **340-2** of fourth viewing area **308**.

[**0227**] The multimedia information may then be analyzed and the results of the analysis are displayed in fifth viewing area **310** (step **1726**). For example, the text information extracted from the multimedia information may be analyzed to identify words that occur in the text information and the frequency of individual words. The words and their frequency may be printed in fifth viewing area **310** (e.g., information printed in area **352** of fifth viewing area **310** as shown in **FIG. 3**). As previously described, information extracted from the multimedia information may be stored in data structures accessible to server **104**. For example, text information and video keyframes information extracted from the multimedia information may be stored in one or more data structures accessible to server **104**. Server **104** may use the information stored in these data structures to analyze the multimedia information.

[**0228**] Multimedia Information Navigation

[**0229**] As previously described, a user of the present invention may navigate and scroll through the multimedia information stored by a multimedia document and displayed in GUI **300** using thumbnail viewing area lens **314** and panel viewing area lens **322**. For example, the user can change the location of thumbnail viewing area lens **314** by moving thumbnail viewing area lens **314** along the length of second viewing area **304**. In response to a change in the position of thumbnail viewing area lens **314** from a first position to a second position over second viewing area **304**, the multimedia information displayed in third viewing area **306** is automatically updated such that the multimedia information displayed in third viewing area **306** continues to correspond to the area of second viewing area **304** emphasized by thumbnail viewing area lens **314** in the second location.

[**0230**] Likewise, the user can change the location of panel viewing area lens **322** by moving panel viewing area lens **322** along the length of third viewing area **306**. In response to a change in the location of panel viewing area lens **322**, the position of sub-lens **316** and also possibly thumbnail viewing area lens **314** are updated to continue to correspond to new location of panel viewing area lens **322**. The information displayed in fourth viewing area **308** is also updated to correspond to the new location of panel viewing area lens **322**.

[**0231**] **FIG. 18** is a simplified high-level flowchart **1800** depicting a method of automatically updating the informa-

tion displayed in third viewing area **306** in response to a change in the location of thumbnail viewing area lens **314** according to an embodiment of the present invention. The method depicted in **FIG. 18** may be performed by server **104**, by client **102**, or by server **104** and client **102** in combination. For example, the method may be executed by software modules executing on server **104** or on client **102**, by hardware modules coupled to server **104** or to client **102**, or combinations thereof. In the embodiment described below, the method is performed by server **104**. The method depicted in **FIG. 18** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[**0232**] As depicted in **FIG. 18**, the method is initiated when server **104** detects a change in the position of thumbnail viewing area lens **314** from a first position to a second position over second viewing area **304** (step **1802**). Server **104** then determines a portion of second viewing area **304** emphasized by thumbnail viewing area lens **314** in the second position (step **1804**). As part of step **1804**, server **104** determines pixel locations ( $TN_{start}$  and  $TN_{end}$ ) in thumbnail **312-1** corresponding to the edges of thumbnail viewing area lens **314** in the second position.  $TN_{start}$  marks the starting vertical pixel location in thumbnail **312-1**, and  $TN_{end}$  marks the ending vertical pixel location in thumbnail **312-1**. Since thumbnails **312-1** and **312-2** are of the same length and are temporally aligned to one another, the portion of thumbnail **312-2** corresponding to second position of thumbnail viewing area lens **314** also lies between pixel locations  $TN_{start}$  and  $TN_{end}$ .

[**0233**] Server **104** then determines time values corresponding to the second position of thumbnail viewing area lens **314** (step **1806**). A time value  $t_1$  is determined corresponding to pixel location  $TN_{start}$  and a time value  $t_2$  is determined corresponding to pixel location  $TN_{end}$ . The multiplier  $tsec\_m$  is used to determine the time values as follows:

$$t_1 = TN_{start} * tsec\_m$$

$$t_2 = TN_{end} * tsec\_m$$

[**0234**] Server **104** then determines pixel locations in the text canvas and the keyframe canvas corresponding to the time values determined in step **1806** (step **1808**). A pixel location  $P_{start}$  in the text canvas is calculated based upon time  $t_1$ , and a pixel location  $P_{end}$  in the text canvas is calculated based upon time  $t_2$ . The multiplier  $pix\_m$  is used to determine the locations as follows:

$$P_{start} = t_1 * pix\_m$$

$$P_{end} = t_2 * pix\_m$$

[**0235**] Since the text canvas and the keyframe canvas are of the same length, time values  $t_1$  and  $t_2$  correspond to pixel locations  $P_{start}$  and  $P_{end}$  in the keyframe canvas.

[**0236**] A section of the text canvas between pixel locations  $P_{start}$  and  $P_{end}$  is displayed in panel **324-1** (step **1810**). The section of the text canvas displayed in panel **324-1** corresponds to the portion of thumbnail **312-1** emphasized by thumbnail viewing area lens **314** in the second position.

[**0237**] A section of the keyframe canvas between pixel locations  $P_{start}$  and  $P_{end}$  is displayed in panel **324-2** (step

1812). The section of the keyframe canvas displayed in panel 324-2 corresponds to the portion of thumbnail 312-2 emphasized by thumbnail viewing area lens 314 in the second position.

[0238] When thumbnail viewing area lens 314 is moved from the first position to the second position, sub-lens 316 also moves along with thumbnail viewing area lens 314. Server 104 then determines a portion of second viewing area 304 emphasized by sub-lens 316 in the second position (step 1814). As part of step 1814, server 104 determines pixel locations (TNSub<sub>start</sub> and TNSub<sub>end</sub>) in thumbnail 312-1 corresponding to the edges of sub-lens 316 in the second position. TNSub<sub>start</sub> marks the starting vertical pixel location in thumbnail 312-1, and TNSub<sub>end</sub> marks the ending vertical pixel location of sub-lens 316 in thumbnail 312-1. Since thumbnails 312-1 and 312-2 are of the same length and are temporally aligned to one another, the portion of thumbnail 312-2 corresponding to second position of sub-lens 316 also lies between pixel locations TNSub<sub>start</sub> and TNSub<sub>end</sub>.

[0239] Server 104 then determines time values corresponding to the second position of sub-lens 316 (step 1816). A time value  $t_3$  is determined corresponding to pixel location TNSub<sub>start</sub> and a time value  $t_4$  is determined corresponding to pixel location TNSub<sub>end</sub>. The multiplier tsec\_m is used to determine the time values as follows:

$$t_3 = \text{TNSub}_{\text{start}} * \text{tsec\_m}$$

$$t_4 = \text{TNSub}_{\text{end}} * \text{tsec\_m}$$

[0240] Server 104 then determines pixel locations in the text canvas and the keyframe canvas corresponding to the time values determined in step 1816 (step 1818). A pixel location PSub<sub>start</sub> in the text canvas is calculated based upon time  $t_3$ , and a pixel location PSub<sub>end</sub> in the text canvas is calculated based upon time  $t_4$ . The multiplier pix\_m is used to determine the locations as follows:

$$\text{PSub}_{\text{start}} = t_3 * \text{tpix\_m}$$

$$\text{PSub}_{\text{end}} = t_4 * \text{tpix\_m}$$

[0241] Since the text canvas and the keyframe canvas are of the same length, time values  $t_1$  and  $t_2$  correspond to pixel locations PSub<sub>start</sub> and PSub<sub>end</sub> in the keyframe canvas.

[0242] Panel viewing area lens 322 is drawn over third viewing area 306 covering a portion of third viewing area 306 between pixels location PSub<sub>start</sub> and PSub<sub>end</sub> (step 1820). The multimedia information displayed in fourth viewing area 308 is then updated to correspond to the new position of panel viewing area lens 322 (step 1822).

[0243] FIG. 19 is a simplified high-level flowchart 1900 depicting a method of automatically updating the information displayed in fourth viewing area 308 and the positions of thumbnail viewing area lens 314 and sub-lens 316 in response to a change in the location of panel viewing area lens 322 according to an embodiment of the present invention. The method depicted in FIG. 19 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 19 is merely illustrative of an embodiment incorporating the present invention and does not limit the

scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. As depicted in FIG. 19, the method is initiated when server 1.04 detects a change in the position of panel viewing area lens 322 from a first position to a second position over third viewing area 306 (step 1902). Server 104 then determines time values corresponding to the second position of panel viewing area lens 322 (step 1904). In step 1904, server 104 determines the pixel locations of the top and bottom edges of panel viewing area lens 322 in the second position. Multiplier sec\_m is then used to convert the pixel locations to time values. A time value  $t_3$  is determined corresponding to top edge of panel viewing area lens 322 in the second position, and a time value  $t_4$  is determined corresponding to bottom edge of panel viewing area lens 322.

$$t_3 = (\text{Pixel location of top edge of panel viewing area lens 322}) * \text{sec\_m}$$

$$t_4 = (\text{Pixel location of bottom edge of panel viewing area lens 322}) * \text{sec\_m}$$

[0244] Server 104 then determines pixel locations in second viewing area 304 corresponding to the time values determined in step 1904 (step 1906). A pixel location TNSub<sub>start</sub> in a thumbnail (either 312-1 or 312-2 since they are aligned and of the same length) in second viewing area 304 is calculated based upon time  $t_3$ , and a pixel location TNSub<sub>end</sub> in the thumbnail is calculated based upon time  $t_4$ . The multiplier tpix\_m is used to determine the locations as follows:

$$\text{TNSub}_{\text{start}} = t_3 * \text{tpix\_m}$$

$$\text{TNSub}_{\text{end}} = t_4 * \text{tpix\_m}$$

[0245] Sub-lens 316 is then updated to emphasize a portion of thumbnails 312 in second viewing area 304 between pixel locations determined in step 1906 (step 1908). As part of step 1908, the position of thumbnail viewing area lens 314 may also be updated if pixels positions TNSub<sub>start</sub> or TNSub<sub>end</sub> lie beyond the boundaries of thumbnail viewing area lens 314 when panel viewing area lens 322 was in the first position. For example, if a user uses panel viewing area lens 322 to scroll third viewing area 306 beyond the PanelHeight, then the position of thumbnail viewing area lens 314 is updated accordingly. If the second position of panel viewing area lens 322 lies within PanelHeight, then only sub-lens 316 is moved to correspond to the second position of panel viewing area lens 322 and thumbnail viewing area lens 314 is not moved.

[0246] As described above, panel viewing area lens 322 may be used to scroll the information displayed in third viewing area 306. For example, a user may move panel viewing area lens 322 to the bottom of third viewing area 306 and cause the contents of third viewing area 306 to be automatically scrolled upwards. Likewise, the user may move panel viewing area lens 322 to the top of third viewing area 306 and cause the contents of third viewing area 306 to be automatically scrolled downwards. The positions of thumbnail viewing area lens 314 and sub-lens 316 are updated as scrolling occurs.

[0247] Multimedia information corresponding to the second position of panel viewing area lens 322 is then displayed in fourth viewing area 308 (step 1910). For example, video information corresponding to the second position of panel viewing area lens 322 is displayed in area 340-1 of fourth

viewing area **308** and text information corresponding to the second position of panel viewing area lens **322** is displayed in area **340-2** of third viewing area **306**.

[0248] According to an embodiment of the present invention, in step **1910**, server **104** selects a time “t” having a value equal to either  $t_3$  or  $t_4$  or some time value between  $t_3$  and  $t_4$ . Time “t” may be referred to as the “location time”. The location time may be user-configurable. According to an embodiment of the present invention, the location time is set to  $t_4$ . The location time is then used as the starting time for playing back video information in area **340-1** of fourth viewing area **308**.

[0249] According to an embodiment of the present invention, GUI **300** may operate in two modes: a “full update” mode and a “partial update” mode. The user of the GUI may select the operation mode of the GUI.

[0250] When GUI **300** is operating in “full update” mode, the positions of thumbnail viewing area lens **314** and panel viewing area lens **322** are automatically updated to reflect the position of the video played back in area **340-1** of fourth viewing area **308**. Accordingly, in “full update” mode, thumbnail viewing area lens **314** and panel viewing area lens **322** keep up or reflect the position of the video played in fourth viewing area **308**. The video may be played forwards or backwards using the controls depicted in area **342** of fourth viewing area **308**, and the positions of thumbnail viewing area lens **314** and panel viewing area lens **322** change accordingly. The multimedia information displayed in panels **324** in third viewing area **306** is also automatically updated (shifted upwards) to correspond to the position of thumbnail viewing area lens **314** and reflect the current position of the video.

[0251] When GUI **300** is operating in “partial update” mode, the positions of thumbnail viewing area lens **314** and panel viewing area lens **322** are not updated to reflect the position of the video played back in area **340-1** of fourth viewing area **308**. In this mode, the positions of thumbnail viewing area lens **314** and panel viewing area lens **322** remain static as the video is played in area **340-1** of fourth viewing area **308**. Since the position of thumbnail viewing area lens **314** does not change, the multimedia information displayed in third viewing area **306** is also not updated. In this mode, a “location pointer” may be displayed in second viewing area **304** and third viewing area **306** to reflect the current position of the video played back in area **340-1** of fourth viewing area **308**. The position of the location pointer is continuously updated to reflect the position of the video.

#### [0252] Ranges

[0253] According to an embodiment, the present invention provides techniques for selecting or specifying portions of the multimedia information displayed in the GUI. Each portion is referred to as a “range.” A range may be manually specified by a user of the present invention or may alternatively be automatically selected by the present invention based upon range criteria provided by the user of the invention.

[0254] A range refers to a portion of the multimedia information between a start time ( $R_S$ ) and an end time ( $R_E$ ). Accordingly, each range is characterized by an  $R_S$  and a  $R_E$  that define the time boundaries of the range. A range

comprises or identifies a portion of the multimedia information occurring between times  $R_S$  and  $R_E$  associated with the range.

[0255] FIG. **20A** depicts a simplified user interface **2000** that displays ranges according to an embodiment of the present invention. It should be apparent that GUI **2000** depicted in FIG. **20A** is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0256] As depicted in FIG. **20A**, GUI **2000** provides various features (buttons, tabs, etc.) that may be used by the user to either manually specify one or more ranges or to configure GUI **2000** to automatically generate ranges. In the embodiment depicted in FIG. **20A**, the user can manually specify a range by selecting “New” button **2002**. After selecting button **2002**, the user can specify a range by selecting a portion of a thumbnail displayed in second viewing area **2004**. One or more ranges may be specified by selecting various portions of the thumbnail. For example, in FIG. **20A**, six ranges **2006-1**, **2006-2**, **2006-3**, **2006-4**, **2006-5**, and **2006-6** have been displayed. One or more of these ranges may be manually specified by the user by selecting or marking portions of thumbnail **2008-2**.

[0257] In alternative embodiments, instead of selecting a portion of a thumbnail, a user can also specify a range by clicking on a location within a thumbnail. A range is then automatically generated by adding a pre-specified buffer time before and after the current clicked location. In this manner, a range can be specified by a single click. Multiple ranges may be specified using this technique.

[0258] In FIG. **20A**, each specified range is indicated by a bar displayed over thumbnail **2008-2**. An identifier or label may also be associated with each range to uniquely identify the range. In FIG. **20A**, each range is identified by a number associated with the range and displayed in the upper left corner of the range. The numbers act as labels for the ranges. Accordingly, information stored for a range may include the start time ( $R_S$ ) for the range, the end time ( $R_E$ ) for the range, and a label or identifier identifying the range. Information identifying a multimedia document storing information corresponding to a range may also be stored for a range.

[0259] Each range specified by selecting a portion of thumbnail **2008-2** is bounded by a top edge ( $R_{top}$ ) and a bottom edge ( $R_{bottom}$ ). The  $R_S$  and  $R_E$  times for a range may be determined from the pixel locations of  $R_{top}$  and  $R_{bottom}$  as follows:

$$R_S = R_{top} * tsec\_m$$

$$R_E = R_{bottom} * tsec\_m$$

[0260] It should be apparent that various other techniques may also be used for specifying a range. For example, in alternative embodiments of the present invention, a user may specify a range by providing the start time ( $R_S$ ) and end time ( $R_E$ ) for the range.

[0261] In GUI **2000** depicted FIG. **20A**, information related to the ranges displayed is GUI **2000** is displayed in area **2010**. The information displayed for each range in area **2010** includes a label or identifier **2012** identifying the range, a start time ( $R_S$ ) **2014** of the range, an end time ( $R_E$ )

**2016** of the range, a time span **2018** of the range, and a set of video keyframes **2019** extracted from the portion of the multimedia information associated with the range. The time span for a range is calculated by determining the difference between the end time  $R_E$  and the start time associated with the range (i.e., time span for a range =  $R_E - R_S$ ). In the embodiment depicted in **FIG. 20A**, the first, last, and middle keyframe extracted from the multimedia information corresponding to each range are displayed. Various other techniques may also be used for selecting keyframes to be displayed for a range. The information depicted in **FIG. 20A** is not meant to limit the scope of the present invention. Various other types of information for a range may also be displayed in alternative embodiments of the present invention.

[**0262**] According to the teachings of the present invention, various operations may be performed on the ranges displayed in GUI **2000**. A user can edit a range by changing the  $R_S$  and  $R_E$  times associated with the range. Editing a range may change the time span (i.e., the value of  $(R_E - R_S)$ ) of the range. In GUI **2000** depicted in **FIG. 20A**, the user can modify or edit a displayed range by selecting “Edit” button **2020**. After selecting “Edit” button **2020**, the user can edit a particular range by dragging the top edge and/or the bottom edge of the bar representing the range. A change in the position of top edge modifies the start time ( $R_S$ ) of the range, and a change in the position of the bottom edge modifies the end time ( $R_E$ ) of the range.

[**0263**] The user can also edit a range by selecting a range in area **2010** and then selecting “Edit” button **2020**. In this scenario, selecting “Edit” button **2020** causes a dialog box to be displayed to the user (e.g., dialog box **2050** depicted in **FIG. 20B**). The user can then change the  $R_S$  and  $R_E$  values associated with the selected range by entering the values in fields **2052** and **2054**, respectively. The time span of the selected range is displayed in area **2056** of the dialog box.

[**0264**] The user can also move the location of a displayed range by changing the position of the displayed range along thumbnail **2008-2**. Moving a range changes the  $R_S$  and  $R_E$  values associated with the range but maintains the time span of the range. In GUI **2000**, the user can move a range by first selecting “Move” button **2022** and then selecting and moving a range. As described above, the time span for a range may be edited by selecting “Edit” button and then dragging an edge of the bar representing the range.

[**0265**] The user can remove or delete a previously specified range. In GUI **2000** depicted in **FIG. 20A**, the user can delete a displayed range by selecting “Remove” button **2024** and then selecting the range that is to be deleted. Selection of “Clear” button **2026** deletes all the ranges that have been specified for the multimedia information displayed in GUI **2000**.

[**0266**] As indicated above, each range refers to a portion of the multimedia information occurring between times  $R_S$  and  $R_E$  associated with the range. The multimedia information corresponding to a range may be output to the user by selecting “Play” button **2028**. After selecting “Play” button **2028**, the user may select a particular range displayed in GUI **2000** whose multimedia information is to be output to the user. The portion of the multimedia information corresponding to the selected range is then output to the user. Various different techniques known to those skilled in the art

may be used to output the multimedia information to the user. According to an embodiment of the present invention, video information corresponding to multimedia information associated with a selected range is played back to the user in area **2030**. Text information corresponding to the selected range may be displayed in area **2032**. The positions of thumbnail viewing area lens **314** and panel viewing area lens **322**, and the information displayed in third viewing area **306** are automatically updated to correspond to the selected range whose information is output to the user in area **2030**.

[**0267**] The user can also select a range in area **2010** and then play information corresponding to the selected range by selecting “Play” button **2020**. Multimedia information corresponding to the selected range is then displayed in area **2030**.

[**0268**] The user may also instruct GUI **2000** to sequentially output information associated with all the ranges specified for the multimedia information displayed by GUI **2000** by selecting “Preview” button **2034**. Upon selecting “Preview” button **2034**, multimedia information corresponding to the displayed ranges is output to the user in sequential order. For example, if six ranges have been displayed as depicted in **FIG. 20A**, multimedia information corresponding to the range identified by label “1” may be output first, followed by multimedia information corresponding to the range identified by label “2”, followed by multimedia information corresponding to the range identified by label “3”, and so on until multimedia information corresponding to all six ranges has been output to the user. The order in which the ranges are output to the user may be user-configurable.

[**0269**] Multimedia information associated with a range may also be saved to memory. For example, in the embodiment depicted in **FIG. 20A**, the user may select “Save” button **2036** and then select one or more ranges that are to be saved. Multimedia information corresponding to the ranges selected by the user to be saved is then saved to memory (e.g., a hard disk, a storage unit, a floppy disk, etc.)

[**0270**] Various other operations may also be performed on a range. For example, according to an embodiment of the present invention, multimedia information corresponding to one or more ranges may be printed on a paper medium. Details describing techniques for printing multimedia information on a paper medium are discussed in U.S. application Ser. No. 10/001,895, (Attorney Docket No.: 15358-006500US) filed Nov. 19, 2001, the entire contents of which are herein incorporated by reference for all purposes.

[**0271**] Multimedia information associated with a range may also be communicated to a user-specified recipient. For example, a user may select a particular range and request communication of multimedia information corresponding to the range to a user-specified recipient. The multimedia information corresponding to the range is then communicated to the recipient. Various different communication techniques known to those skilled in the art may be used to communicate the range information to the recipient including faxing, electronic mail, wireless communication, and other communication techniques.

[**0272**] Multimedia information corresponding to a range may also be provided as input to another application program such as a search program, a browser, a graphics

application, a MIDI application, or the like. The user may select a particular range and then identify an application to which the information is to be provided. In response to the user's selection, multimedia information corresponding to the range is then provided as input to the application.

[0273] As previously stated, ranges may be specified manually by a user or may be selected automatically by the present invention. The automatic selection of ranges may be performed by software modules executing on server 104, hardware modules coupled to server 104, or combinations thereof. FIG. 21 is a simplified high-level flowchart 2100 depicting a method of automatically creating ranges according to an embodiment of the present invention. The method depicted in FIG. 21 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 21 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0274] As depicted in FIG. 21, the method is initiated when server 104 receives criteria for creating ranges (step 2102). The user of the present invention may specify the criteria via GUI 2000. For example, in GUI 2000 depicted in FIG. 20A, area 2040 displays various options that can be selected by the user to specify criteria for automatic creation of ranges. In GUI 2000 depicted in FIG. 20A, the user may select either "Topics" or "Words" as the range criteria. If the user selects "Topics", then information related to topics of interest to the user (displayed in area 2042) is identified as the range creation criteria. If the user selects "Words", then one or more words selected by the user in area 2044 of GUI 2000 are identified as criteria for automatically creating ranges. In alternative embodiments, the criteria for automatically creating ranges may be stored in a memory location accessible to server 104. For example, the criteria information may be stored in a file accessible to server 104. Various other types of criteria may also be specified according to the teachings of the present invention.

[0275] The multimedia information stored in the multimedia document is then analyzed to identify locations (referred to as "hits") in the multimedia information that satisfy the criteria received in step 2102 (step 2104). For example, if the user has specified that one or more words selected by the user in area 2044 are to be used as the range creation criteria, then the locations of the selected words are identified in the multimedia information. Likewise, if the user has specified topics of interest as the range creation criteria, then server 104 analyzes the multimedia information to identify locations in the multimedia information that are relevant to the topics of interest specified by the user. As described above, server 104 may analyze the multimedia information to identify locations of words or phrases associated with the topics of interest specified by the user. Information related to the topics of interest may be stored in a user profile file that is accessible to server 104. It should be apparent that various other techniques known to those skilled in the art may also

be used to identify locations in the multimedia information that satisfy the range criteria received in step 2102.

[0276] One or more ranges are then created based upon the locations of the hits identified in step 2104 (step 2106). Various different techniques may be used to form ranges based upon locations of the hits. According to one technique, one or more ranges are created based upon the times associated with the hits. Hits may be grouped into ranges based on the proximity of the hits to each other. One or more ranges created based upon the locations of the hits may be combined to form larger ranges.

[0277] The ranges created in step 2106 are then displayed to the user using GUI 2000 (step 2108). Various different techniques may be used to display the ranges to the user. In FIG. 20A, each range is indicated by a bar displayed over thumbnail 2008-2.

[0278] FIG. 22 is a simplified high-level flowchart 2200 depicting a method of automatically creating ranges based upon locations of hits in the multimedia information according to an embodiment of the present invention. The processing depicted in FIG. 22 may be performed in step 2106 depicted in FIG. 21. The method depicted in FIG. 22 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 22 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0279] As depicted in FIG. 22, the method is initiated by determining a time associated the first hit in the multimedia information (step 2202). The first hit in the multimedia information corresponds to a hit with the earliest time associated with it (i.e., a hit that occurs before other hits in the multimedia information). A new range is then created to include the first hit such that  $R_s$  for the new range is set to the time of occurrence of the first hit, and  $R_e$  for the new range is set to some time value after the time of occurrence of the first hit (step 2204). According to an embodiment of the present invention,  $R_e$  is set to the time of occurrence of the hit plus 5 seconds.

[0280] Server 104 then determines if there are any additional hits in the multimedia information (step 2206). Processing ends if there are no additional hits in the multimedia information. The ranges created for the multimedia information may then be displayed to the user according to step 2108 depicted in FIG. 21. If it is determined in step 2206 that additional hits exist in the multimedia information, then the time associated with the next hit is determined (step 2208).

[0281] Server 104 then determines if the time gap between the end time of the range including the previous hit and the time determined in step 2208 exceeds a threshold value (step 2210). Accordingly, in step 2210 server 104 determines if:  $(\text{Time determined in step 2208}) - (R_e \text{ of range including previous hit}) > \text{GapBetweenHits}$  wherein, GapBetweenHits represents the threshold time value. The threshold value is

user configurable. According to an embodiment of the present invention, GapBetweenHits is set to 60 seconds.

[0282] If it is determined in step 2210 that the time gap between the end time of the range including the previous hit and the time determined in step 2208 exceeds the threshold value, then a new range is created to include the next hit such that  $R_S$  for the new range is set to the time determined in step 2208, and  $R_E$  for the new range is set to some time value after the time determined in step 2208 (step 2212). According to an embodiment of the present invention,  $R_E$  is set to the time of occurrence of the hit plus 5 seconds. Processing then continues with step 2206.

[0283] If it is determined in step 2210 that the time gap between the end time of the range including the previous hit and the time determined in step 2208 does not exceed the threshold value, then the range including the previous hit is extended by changing the end time  $R_E$  of the range to the time determined in step 2208 (step 2214). Processing then continues with step 2206.

[0284] According to the method depicted in FIG. 22, a single range is created for hits in the multimedia information that occur within a threshold value ("GapBetweenHits") from the previous range. At the end of the method depicted in FIG. 22, one or more ranges are automatically created based upon the range criteria.

[0285] According to an embodiment of the present invention, after forming one or more ranges based upon the times associated with the hits (e.g., according to flowchart 2200 depicted in FIG. 22), one or more ranges created based upon the locations of the hits may be combined with other ranges to form larger ranges. According to an embodiment of the present invention, a small range is identified and combined with a neighboring range if the time gap between the small range and the neighboring range is within a user-configurable time period threshold. If there are two neighboring time ranges that are within the time period threshold, then the small range is combined with the neighboring range that is closest to the small range. The neighboring ranges do not need to be small ranges. Combination of smaller ranges to form larger ranges is based upon the premise that a larger range is more useful to the user than multiple small ranges.

[0286] FIG. 23 is a simplified high-level flowchart 2300 depicting a method of combining one or more ranges based upon the size of the ranges and the proximity of the ranges to neighboring ranges according to an embodiment of the present invention. The processing depicted in FIG. 23 may be performed in step 2106 depicted in FIG. 21 after processing according to flowchart 2200 depicted in FIG. 22 has been performed. The method depicted in FIG. 23 may be performed by server 104, by client 102, or by server 104 and client 102 in combination. For example, the method may be executed by software modules executing on server 104 or on client 102, by hardware modules coupled to server 104 or to client 102, or combinations thereof. In the embodiment described below, the method is performed by server 104. The method depicted in FIG. 23 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0287] In order to describe the processing performed in FIG. 23, it is assumed that "N" ranges ( $N \geq 1$ ) have been

created for the multimedia information displayed by the GUI. The ranges may have been created according to the processing depicted in flowchart 2200 in FIG. 22. Each range  $R_i$ , where  $(1 \leq i \leq N)$ , in the set of "N" ranges has a start time  $R_S$  and an end time  $R_E$  associated with it. For a range  $R_i$ , the neighbors of the range include range  $R_{(i-1)}$  and range  $R_{(i+1)}$ , where  $R_E$  of range  $R_{(i-1)}$  occurs before  $R_S$  of range  $R_i$  and  $R_E$  of range  $R_i$  occurs before  $R_S$  of range  $R_{(i+1)}$ . Range  $R_{(i-1)}$  is referred to as a range that occurs before range  $R_i$ . Range  $R_{(i+1)}$  is referred to as a range that occurs after range  $R_i$ .

[0288] As depicted in FIG. 23, the method is initiated by initializing a variable "i" to 1 (step 2303). A range  $R_i$  is then selected (step 2304). During the first pass through flowchart 2300, the first range (i.e., the range having the earliest  $R_S$  time) in the set of "N" ranges is selected. Subsequent ranges are selected in subsequent passes.

[0289] Server 104 then determines if range  $R_i$  selected in step 2304 qualifies as a small range. According to an embodiment of the present invention, a threshold value "SmallRangeSize" is defined and a range is considered a small range if the time span of the range is less than or equal to threshold value SmallRangeSize. Accordingly, in order to determine if range  $R_i$  qualifies as a small range, the time span of range  $R_i$  selected in step 2304 is compared to threshold time value "SmallRangeSize" (step 2306). The value of SmallRangeSize may be user-configurable. According to an embodiment of the present invention, SmallRangeSize is set to 8 seconds.

[0290] If it is determined in step 2306 that the range  $R_i$  selected in step 2304 does not qualify as a small range (i.e., the time span ( $R_E - R_S$ ) of range  $R_i$  is greater than the threshold value SmallRangeSize), then the range is not a candidate for combination with another range. The value of variable "i" is then incremented by one (step 2308) to facilitate selection of the next range in the set of "N" ranges. Accordingly, according to the teachings of the present invention depicted in FIG. 23, only ranges that qualify as small ranges are eligible for combination with other neighboring ranges.

[0291] After step 2308, server 104 determines if all the ranges in the set of "N" ranges have been processed. This is done by determining if the value of "i" is greater than the value of "N" (step 2310). If the value of "i" is greater than "N", it indicates that all the ranges in the set of ranges for the multimedia information have been processed and processing of flowchart 2300 ends. If it is determined in step 2310 that "i" is less than or equal to "N", then it indicates that the set of "N" ranges comprises at least one range that has not been processed according to flowchart 2300. Processing then continues with step 2304 wherein the next range  $R_i$  is selected.

[0292] If it is determined in step 2306 that range  $R_i$  selected in step 2304 qualifies as a small range (i.e., the time span ( $R_E - R_S$ ) of range  $R_i$  is less than or equal to the threshold value SmallRangeSize), the present invention then performs processing to identify a range that is a neighbor of range  $R_i$  (i.e., a range that occurs immediately before or after range  $R_i$  selected in step 2304) with which range  $R_i$  can be combined. In order to identify such a range, server 104 initializes variables to facilitate selection of ranges that are neighbors of range  $R_i$  selected in step 2304 (step 2312). A

variable “j” is set to the value (i+1) and a variable “k” is set to the value “(i-1)”. A variable “j” is used to refer to a range that is a neighbor of range  $R_i$  and occurs after range  $R_i$ , and a variable “k” is used to refer to a range that is a neighbor of range  $R_i$  and occurs before range  $R_i$ . FIG. 24 depicts a simplified diagram showing the relationship between ranges  $R_i$ ,  $R_j$ , and  $R_k$ . As shown in FIG. 24, range  $R_i$  occurs after range  $R_k$  (i.e.,  $R_S$  of  $R_i$  occurs after  $R_E$  of  $R_k$ ) and before range  $R_j$  (i.e.,  $R_E$  of  $R_i$  occurs before  $R_S$  of  $R_j$ ).

[0293] Server 104 then determines if the set of “N” ranges created for the multimedia information includes a range that is a neighbor of range  $R_i$  selected in step 2304 and occurs before range  $R_i$ , and a range that is a neighbor of range  $R_i$  and occurs after range  $R_i$ . This is done by determining the values of variables “j” and “k”. If the value of “j” is greater than “N”, it indicates that the range  $R_j$  selected in step 2304 is the last range in the set of “N” ranges created for the multimedia information implying that there is no range that occurs after range  $R_i$ . If the value of “k” is equal to zero, it indicates that the range  $R_i$  selected in step 2304 is the first range in the set of “N” ranges created for the multimedia information implying that there is no range that occurs before range  $R_i$ .

[0294] Accordingly, server 104 determines if range  $R_i$  has a neighboring range that occurs before  $R_i$  and a neighboring range that occurs after  $R_i$ . This is done by determining if the value of “j” is less than “N” and if the value of “k” is not equal to zero (step 2314). If the condition in step 2314 is satisfied, then it indicates that the set of “N” ranges comprises a range that is a neighbor of range  $R_i$  selected in step 2304 and occurs before range  $R_i$ , and a range that is a neighbor of range  $R_i$  and occurs after range  $R_i$ . In this case, processing continues with step 2316. If the condition in step 2314 is not satisfied, then it indicates that range  $R_i$  selected in step 2304 is either the first range in the set of “N” ranges implying that there is no range that occurs before range  $R_i$ , and/or that range  $R_i$  selected in step 2304 is the last range in the set of “N” ranges implying that there is no range that occurs after range  $R_i$ . In this case, processing continues with step 2330.

[0295] If the condition in step 2314 is determined to be true, server 104 then determines time gaps between ranges  $R_i$  and  $R_k$  and between ranges  $R_i$  and  $R_j$  (step 2316). The time gap (denoted by  $G_{ik}$ ) between ranges  $R_i$  and  $R_k$  is calculated by determining the time between  $R_S$  of range  $R_i$  and  $R_E$  of  $R_k$ , (see FIG. 24) i.e.,

$$G_{ik} = (R_S \text{ of } R_i) - (R_E \text{ of } R_k)$$

[0296] The time gap (denoted by  $G_{ij}$ ) between ranges  $R_i$  and  $R_j$  is calculated by determining the time between  $R_E$  of range  $R_i$  and  $R_S$  of  $R_j$ , (see FIG. 24) i.e.,

$$G_{ij} = (R_S \text{ of } R_j) - (R_E \text{ of } R_i)$$

[0297] According to the teachings of the present invention, a small range is combined with a neighboring range only if the gap between the small range and the neighboring range is less than or equal to a threshold gap value. The threshold gap value is user configurable. Accordingly, server 104 then determines the sizes of the time gaps to determine if range  $R_i$  can be combined with one of its neighboring ranges.

[0298] Server 104 then determines which time gap is larger by comparing the values of time gap  $G_{ik}$  and time gap

$G_{ij}$  (step 2318). If it is determined in step 2318 that  $G_{ik}$  is greater than  $G_{ij}$ , it indicates that range  $R_i$  selected in step 2304 is closer to range  $R_j$  than to range  $R_k$ , and processing continues with step 2322. Alternatively, if it is determined in step 2318 that  $G_{ik}$  is not greater than  $G_{ij}$ , it indicates that the time gap between range  $R_i$  selected in step 2304 and range  $R_k$  is equal to or less than the time gap between ranges  $R_i$  and  $R_j$ . In this case processing continues with step 2320.

[0299] If it is determined in step 2318 that  $G_{ik}$  is not greater than  $G_{ij}$ , server 104 then determines if the time gap ( $G_{ik}$ ) between range  $R_i$  and range  $R_k$  is less than or equal to a threshold gap value “GapThreshold” (step 2320). The value of GapThreshold is user configurable. According to an embodiment of the present invention, GapThreshold is set to 90 seconds. It should be apparent that various other values may also be used for Gap Threshold.

[0300] If it is determined in step 2320 that the time gap ( $G_{ik}$ ) between range  $R_i$  and range  $R_k$  is less than or equal to threshold gap value GapThreshold (i.e.,  $G_{ik} < \text{GapThreshold}$ ), then ranges  $R_i$  and  $R_k$  are combined to form a single range (step 2324). The process of combining ranges  $R_i$  and  $R_k$  involves changing the end time of range  $R_k$  to the end time of range  $R_i$  (i.e.,  $R_E$  of  $R_k$  is set to  $R_E$  of  $R_i$ ) and deleting range  $R_i$ . Processing then continues with step 2308 wherein the value of variable “i” is incremented by one.

[0301] If it is determined in step 2320 that time gap  $G_{ik}$  is greater than GapThreshold (i.e.,  $G_{ik} > \text{GapThreshold}$ ), it indicates that both ranges  $R_j$  and  $R_k$  are outside the threshold gap value and as a result range  $R_i$  cannot be combined with either range  $R_j$  or  $R_k$ . In this scenario, processing continues with step 2308 wherein the value of variable “i” is incremented by one.

[0302] Referring back to step 2318, if it is determined that  $G_{ik}$  is greater than  $G_{ij}$ , server 104 then determines if the time gap ( $G_{ij}$ ) between ranges  $R_i$  and  $R_j$  is less than or equal to the threshold gap value “GapThreshold” (step 2322). As indicated above, the value of GapThreshold is user configurable. According to an embodiment of the present invention, GapThreshold is set to 90 seconds. It should be apparent that various other values may also be used for GapThreshold.

[0303] If it is determined in step 2322 that the time gap ( $G_{ij}$ ) between ranges  $R_i$  and  $R_j$  is less than or equal to threshold gap value GapThreshold (i.e.,  $G_{ij} < \text{GapThreshold}$ ), then ranges  $R_i$  and  $R_j$  are combined to form a single range (step 2326). The process of combining ranges  $R_i$  and  $R_j$  involves changing the start time of range  $R_j$  to the start time of range  $R_i$  (i.e.,  $R_S$  of  $R_j$  is set to  $R_S$  of  $R_i$ ) and deleting range  $R_i$ . Processing then continues with step 2308 wherein the value of variable “i” is incremented by one.

[0304] If it is determined in step 2322 that time gap  $G_{ij}$  is greater than GapThreshold (i.e.,  $G_{ij} > \text{GapThreshold}$ ), it indicates that both ranges  $R_j$  and  $R_k$  are outside the threshold gap value and as a result range  $R_i$  cannot be combined with either range  $R_j$  or  $R_k$ . In this scenario, processing continues with step 2308 wherein the value of variable “i” is incremented by one.

[0305] If server 104 determines that the condition in step 2314 is not satisfied, server 104 then determines if the value of “k” is equal to zero (step 2330). If the value of “k” is equal to zero, it indicates that the range  $R_i$  selected in step 2304 is

the first range in the set of “N” ranges created for the multimedia information which implies that there is no range in the set of “N” ranges that occurs before range  $R_i$ . In this scenario, server **104** then determines if the value of variable “j” is greater than “N” (step **2332**). If the value of “j” is also greater than “N”, it indicates that the range  $R_i$  selected in step **2304** is not only the first range but also the last range in the set of “N” ranges created for the multimedia information which implies that there is no range in the set of ranges that comes after range  $R_i$ . If it is determined in step **2330** that “k” is equal to zero and that “j” > N in step **2332**, it indicates that the set of ranges for the multimedia information comprises only one range (i.e., N=1). Processing depicted in flowchart **2300** is then ended since no ranges can be combined.

[**0306**] If it is determined in step **2330** that “k” is equal to zero and that “j” is not greater than “N” in step **2332**, it indicates that the range  $R_i$  selected in step **2304** represents the first range in the set of “N” ranges created for the multimedia information, and that the set of ranges includes at least one range  $R_j$  that is a neighbor of range  $R_i$  and occurs after range  $R_i$ . In this case, the time gap  $G_{ij}$  between range  $R_i$  and range  $R_j$  is determined (step **2334**). As indicated above, time gap  $G_{ij}$  is calculated by determining the time between  $R_E$  of range  $R_i$  and  $R_S$  of  $R_j$ , i.e.,  $G_{ij} = (R_S \text{ of } R_j) - (R_E \text{ of } R_i)$ . Processing then continues with step **2322** as described above.

[**0307**] If it is determined in step **2330** that “k” is not equal to zero, it indicates that the range  $R_i$  selected in step **2304** represents the last range in the set of “N” ranges created for the multimedia information, and that the set of ranges includes at least one range  $R_k$  that is a neighbor of range  $R_i$  and occurs before range  $R_i$ . In this case, the time gap  $G_{ik}$  between range  $R_i$  and range  $R_k$  is determined (step **2336**). As indicated above, time gap  $G_{ik}$  is calculated by determining the time gap between  $R_S$  of range  $R_i$  and  $R_E$  of  $R_k$ , i.e.,

$$G_{ik} = (R_S \text{ of } R_i) - (R_E \text{ of } R_k)$$

[**0308**] Processing then continues with step **2320** as described above.

[**0309**] **FIG. 25A** depicts a simplified diagram showing a range created by combining ranges  $R_i$  and  $R_k$  depicted in **FIG. 24** according to an embodiment of the present invention. **FIG. 25B** depicts a simplified diagram showing a range created by combining ranges  $R_i$  and  $R_j$  depicted in **FIG. 24** according to an embodiment of the present invention.

[**0310**] As indicated above, the processing depicted in **FIG. 23** may be performed after one or more ranges have been created according to the times associated with the hits according to flowchart **2200** depicted in **FIG. 22**. According to an embodiment of the present invention, after the ranges have been combined according to flowchart **2300** depicted in **FIG. 23**, the ranges may then be displayed to the user in GUI **2000** according to step **2108** in **FIG. 21**.

[**0311**] According to an alternative embodiment of the present invention, after combining ranges according to flowchart **2300** depicted in **FIG. 23**, a buffer time is added to the start time and end time of each range. A user may configure the amount of time (BufferStart) to be added to the start time of each range and the amount of time (BufferEnd) to be added to the end time of each range. The buffer times are added to a range so that a range does not start immediately

on a first hit in the range and stop immediately at the last hit in the range. The buffer time provides a lead-in and a trailing-off for the information contained in the range and thus provides a better context for the range.

[**0312**] A buffer is provided at the start of a range by changing the  $R_S$  time of the range as follows:

$$R_S \text{ of range} = (R_S \text{ of range before adding buffer}) - \text{BufferStart}$$

[**0313**] A buffer is provided at the end of a range by changing the  $R_E$  time of the range as follows:

$$R_E \text{ of range} = (R_E \text{ of range before adding buffer}) + \text{BufferEnd}$$

[**0314**] **FIG. 26** depicts a zoomed-in version of GUI **2000** depicting ranges that have been automatically created according to an embodiment of the present invention. A plurality of hits **2602** satisfying criteria provided by the user are marked in thumbnail **2008-1** that displays text information. According to an embodiment of the present invention, the hits represent words and/or phrases related to user-specified topics of interest. As depicted in **FIG. 26**, two ranges **2006-2** and **2006-3** have been automatically created based upon locations of the hits. Range **2006-2** has been created by merging several small ranges according to the teachings of the present invention (e.g., according to flowchart **2300** depicted in **FIG. 23**).

[**0315**] Displaying Multimedia Information from Multiple Multimedia Documents

[**0316**] The embodiments of the present invention described above display representations of information that has been recorded (or captured) along a common timeline. The recorded information may include information of different types such as audio information, video information, closed-caption (CC) text information, slides information, whiteboard information, etc. The different types of information may have been captured by one or more capture devices.

[**0317**] As described above, a multimedia document may provide a repository for storing the recorded or captured information. The multimedia document may be a file that stores the recorded information comprising information of multiple types. The multimedia document may be a file that includes references to one or more other files that store the recorded information. The referenced files may store information of one or more types. The multimedia document may also be a location where the recorded information of one or more types is stored. For example, the multimedia document may be a directory that stores files comprising information that has been captured or recorded during a common timeline. According to an embodiment of the present invention, each file in the directory may store information of a particular type, i.e., each file may store a particular stream of information. Accordingly, for recorded information that comprises information of multiple types (e.g., a first type, a second type, etc.), the information of the various types may be stored in a single file, the information for each type may be stored in a separate file, and the like.

[**0318**] Since the different types of information have been captured along a common timeline, the representations of the information can be displayed in a manner such that the representations when displayed by the GUI are temporally aligned with each other. For example, interface **300** depicted

in **FIG. 3** displays multimedia information stored by a television broadcast recording multimedia document. The different types of information stored in the broadcast recording include video information, audio information, and possibly closed-caption (CC) text information. The video information, audio information, and CC text information are all captured along the same (or common) timeline possibly by different capture devices. For example, the audio information may have been captured using an audio information capture device (e.g., a microphone) and the video information may have been captured by a video information capture device (e.g., a video camera). The audio and video information might also have been captured by a single information capture device.

[0319] As described above, interface **300** displays text information that is a representation of the audio or CC text information included in the broadcast recording (or a text representation of some other type of information included in the multimedia information). Interface **300** also displays video keyframes extracted from the video information included in the broadcast recording. The displayed video keyframes are a representation of the video information stored in the multimedia document. Since the audio and video information are captured along the same timeline, the representations of the information can be displayed such that they are temporally aligned or synchronized with each other. For example, as described above, thumbnail images **312-1** and **312-2** are aligned such that the text information (which may represent a transcript of the audio information or the CC text information or a text representation of some other type of information included in the multimedia information) in thumbnail image **312-1** and video keyframes displayed in thumbnail **312-2** that occur at a particular point of time are displayed approximately close to each other along the same horizontal axis. This enables a user to determine various types of information in the television broadcast recording occurring approximately concurrently by simply scanning the thumbnail images in the horizontal axis. Likewise, panels **324-1** and **324-2** are temporally aligned or synchronized with each other such that representations of the various types of information occurring concurrently in the television broadcast recording are displayed approximately close to each other.

[0320] Embodiments of the present invention can also display recorded multimedia information that may be stored in multiple multimedia documents. The multimedia information in the multiple multimedia documents may have been captured along different timelines. For example, embodiments of the present invention can display representations of multimedia information from a television news broadcast captured or recorded during a first timeline (e.g., a morning newscast) and from another television news broadcast captured during a second timeline (e.g., an evening newscast) that is different from the first timeline. Accordingly, embodiments of the present invention can display multimedia information stored in one or more multimedia documents that may store multimedia information captured along different timelines. Each multimedia document may comprise information of different types such as audio information, video information, CC text information, whiteboard information, slides information, and the like.

[0321] The multiple multimedia documents whose information is displayed may also include documents that store

information captured along the same timeline. For example, the multiple multimedia documents may include a first television program recording from a first channel captured during a first timeline and a second television program recording from a second channel captured during the same timeline (i.e., the first timeline) as the first television program recording. Embodiments of the present invention can accordingly display representations of information from multiple multimedia documents that store information that may have been captured along the same or different timelines.

[0322] **FIG. 27** depicts a simplified startup user interface **2700** that can display information that may be stored in one or more multimedia documents according to an embodiment of the present invention. Interface **2700** is merely illustrative of an embodiment of the present invention and does not limit the scope of the present invention. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0323] As depicted in **FIG. 27**, interface **2700** comprises a toolbar **2702** including several user-selectable buttons. The buttons include a button **2704** for loading multimedia documents for display, a button **2706** for removing one or more previously loaded multimedia documents, a button **2708** for printing multimedia information from one or more loaded multimedia documents on a paper medium, a button **2710** for configuring user preferences, and other buttons that allow a user to perform actions, configure, customize, or control the manner in which information from one or more multimedia documents is displayed. Additional features of interface **2700** are described below in more detail.

[0324] In order to load one or more multimedia documents to be displayed, the user selects load button **2704**. **FIG. 28** depicts a simplified window **2800** that is displayed when the user selects load button **2704** according to an embodiment of the present invention. Window **2800** facilitates selection of one or more multimedia documents to be loaded and displayed according to the teachings of the present invention. As depicted in **FIG. 28**, information identifying one or more multimedia documents that are available to be loaded is displayed in box **2802** of window **2800**. Each multimedia document may be identified by an identifier (e.g., a filename, a location identifier such as a directory name). In the embodiment depicted in **FIG. 28**, each multimedia document is identified by a five digit code identifier. The user may select one or more multimedia documents to be loaded by highlighting the identifiers corresponding to the multimedia documents in box **2802** and then selecting "Add" button **2804**. The highlighted identifiers for the multimedia documents are then moved from box **2802** and displayed in box **2806** that displays multimedia documents selected for loading. A previously selected multimedia document can be deselected by highlighting the identifier for the multimedia document in box **2806** and then selecting "Remove" button **2808**.

[0325] Information related to the multimedia document corresponding to a highlighted identifier (highlighted either in box **2802** or **2806**) is displayed in information area **2810**. In the embodiment depicted in **FIG. 28**, the displayed information includes information **2812** indicating the duration of the multimedia document, information **2814** indicating the date on which the information in the multimedia

document was captured or recorded, information **2816** indicating the time of the recording, information **2818** identifying the television channel from which the information was recorded, and information **2820** indicating the type of recording. Other descriptive information that is available for the multimedia document (e.g., name of the TV program) might be displayed in description area **2821**.

[**0326**] The user can select "Load" button **2822** to load and display contents of multimedia documents identified by the identifiers displayed in box **2806**. As shown in **FIG. 28**, three multimedia documents have been selected and will be loaded upon selection of "Load" button **2822**. The selected multimedia documents may store multimedia information captured along the same or different timelines. Each selected multimedia document may comprise information of one or more types (e.g., audio information, video information, CC text information, whiteboard information, slides information, etc.). The types of information stored by one multimedia document may be different from the types of information stored by another selected multimedia document. The user can cancel the load operation by selecting "Cancel" button **2824**.

[**0327**] Other techniques may also be used for selecting and loading a multimedia document. For example, according to one technique, a user may scan a particular identifier (e.g., a barcode). The multimedia document (or portion of information stored by the multimedia document) corresponding to the scanned barcode may be selected and loaded.

[**0328**] **FIG. 29A** depicts a user interface **2900** after one or more multimedia documents have been loaded and displayed according to an embodiment of the present invention. Interface **2700** is merely illustrative of an embodiment of the present invention and does not limit the scope of the present invention. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[**0329**] As depicted in **FIG. 29A**, contents of three multimedia documents have been loaded and displayed. For each multimedia document, representations of information of different types stored by the multimedia document are displayed in a thumbar corresponding to the multimedia document. A video window is also displayed for each multimedia document. The thumbar for each multimedia document includes one or more thumbnail images displaying representations of the various types of information included in the multimedia document.

[**0330**] For example, in the **FIG. 29A**, a thumbar **2902** displays representations of information stored by a first multimedia document, a thumbar **2906** displays representations of information stored by a second multimedia document, and a thumbar **2910** displays representations of information stored by a third multimedia document. A video window **2904** is displayed for the first multimedia document, a video window **2908** is displayed for the second multimedia document, and a video window **2912** is displayed for the third multimedia document. In the embodiment depicted in **FIG. 29A**, the first, second, and third multimedia documents are recordings of television programs and each comprise audio information, video information, and possibly CC text information. This is however not intended to limit the scope of the present invention. A multimedia document displayed according to the teachings of the present invention may include different types of information.

[**0331**] Each thumbar displayed in **FIG. 29A** includes one or more thumbnail images. Each thumbnail image displays a representation of a type of information stored in the multimedia document. Since the three multimedia documents loaded in interface **2900** comprise audio, video, and possibly CC text information, thumbars **2902**, **2906**, and **2910** each include a thumbnail image displaying text information that is a representation of the audio information or CC text information (or a text representation of some other type of information included in the multimedia information) from the corresponding multimedia document and a thumbnail image displaying video keyframes representing the video information in the corresponding multimedia document. For example, thumbar **2902** includes a thumbnail image **2914** that displays text information representing the audio information (or CC text information) from the first multimedia document, and a thumbnail image **2916** displaying video keyframes extracted from the video information of the first multimedia document. Thumbar **2906** includes a thumbnail image **2918** that displays text information representing the audio information (or CC text information, or a text representation of some other type of information included in the multimedia information) from the second multimedia document, and a thumbnail image **2920** displaying video keyframes extracted from the video information of the second multimedia document. Thumbar **2910** includes a thumbnail image **2922** that displays text information representing the audio information (or CC text information, or a text representation of some other type of information included in the multimedia information) from the third multimedia document, and a thumbnail image **2924** displaying video keyframes extracted from the video information of the third multimedia document. Techniques for generating and displaying the thumbnail images have been previously described. Each thumbar is like the second viewing area depicted in **FIG. 3**.

[**0332**] The thumbnail images in a thumbar are aligned such that representations of information that occurs temporally concurrently in the multimedia document are displayed approximately close to each other along the same horizontal axis. Each thumbar represents information captured according to a common timeline. However, the timeline corresponding to one thumbar may be different from the timeline corresponding to another thumbar.

[**0333**] A lens ("thumbnail viewing area lens") is displayed for each thumbar covering or emphasizing a portion of the thumbar. As depicted in **FIG. 29A**, a thumbnail viewing area lens **2926** covers an area of thumbar **2902**, a thumbnail viewing area lens **2928** covers an area of thumbar **2906**, and a thumbnail viewing area lens **2930** covers an area of thumbar **2910**. The thumbnail viewing area lenses are initially positioned at the top of the thumbars (i.e., at the start of the multimedia documents) as depicted in **FIG. 29A**. As described above with respect to **FIG. 3**, each thumbnail viewing lens can be moved along the corresponding thumbar and can be used to navigate and scroll through the contents of the multimedia document displayed in the thumbar. Techniques for displaying a thumbnail viewing area lens and techniques for using the thumbnail viewing area lens to navigate and scroll through the contents of each multimedia document have been previously described. Each thumbnail viewing area lens may or may not comprise a sublens such as sublens **316** depicted in **FIG. 3**.

[0334] Descriptive information related to each multimedia document may also be displayed in the thumbbar corresponding to the multimedia document. The information may include information such as information indicating the duration of the multimedia document, information indicating the date on which the information in the multimedia document was captured or recorded, information indicating the time of the recording, information identifying the television channel or program from which the information was recorded, information indicating the type of recording, etc. As depicted in FIG. 29A, descriptive information 2932 for each multimedia document is displayed along a side of the corresponding thumbbar.

[0335] For each multimedia document, the video information may be played back in a video window corresponding to the multimedia document. The audio information accompanying the video information may also be out via an audio output device. For example, video information from the first multimedia document may be played back in video window 2904, video information from the second multimedia document may be played back in video window 2908, and video information from the third multimedia document may be played back in video window 2912. A control bar is provided with each video window for controlling playback of information in the associated video window. For example, the playback of video information in video window 2904 may be controlled using controls provided by control bar 2934, the playback of video information in video window 2908 may be controlled using controls provided by control bar 2936, and the playback of video information in video window 2912 may be controlled using controls provided by control bar 2938.

[0336] The contents of video information displayed in a video window for a multimedia document also depends on the position of the thumbnail viewing area lens over the thumbbar corresponding to the multimedia document. For example, contents of the video information displayed in video window 2904 depend upon the position of thumbnail viewing area lens 2926 over thumbbar 2902. As previously described, each thumbnail viewing area lens is characterized by a top edge corresponding to time  $t_1$  and a bottom edge corresponding to a time  $t_2$ . The playback of the video information in the video window is started at time  $t_1$  or  $t_2$  or some time in between times  $t_1$  and  $t_2$ . As the thumbnail viewing area lens is repositioned over a thumbbar, the video played back in the corresponding video window may change such that the video starts playing from time  $t_1$  or time  $t_2$  corresponding to the present position of thumbnail viewing area lens over the thumbbar, or some time in between  $t_1$  and  $t_2$ . It should be noted that the thumbnail viewing area lenses covering portions of the different thumbbars can be repositioned along the thumbbars independent of each other.

[0337] Each video window may also display information related to the multimedia document whose video information contents are displayed in the video window. The information may include for example, information identifying the television program for the recording, information identifying the time in the multimedia document corresponding to the currently played back content, etc.

[0338] According to an embodiment of the present invention, a list of words 2940 found in all the loaded multimedia documents (i.e., common words) is also displayed in an area

of interface 2900. The list of words 2940 includes words that are found in information of one or more types contained by the loaded multimedia documents. For example, the list of words displayed in FIG. 29A includes words that were found in the information from first multimedia document, the second multimedia document, and the third multimedia document. According to an embodiment of the present invention, text representations of information contained by the loaded multimedia documents are searched to find the common words. The text information may represent the CC text information, a transcription of the audio information, or a text representation of some other type of information stored in the multimedia documents. According to another embodiment of the present invention, the list of words may also include words determined from the video information contained by the multimedia documents. For example, the video keyframes extracted from the video information may also be searched to find the common words. The keyframes may be searched for the words. The number of occurrences of the words in the multimedia documents is also shown.

[0339] FIG. 29B depicts interface 2900 wherein the positions of the thumbnail viewing area lenses has been changed from their initial positions according to an embodiment of the present invention. As shown, the positions of thumbnail viewing area lenses 2926, 2928, and 2930 have been changed from their positions depicted in FIG. 29A. Since the positions of the thumbnail viewing area lenses affects the video information played back in the corresponding video windows, the contents in the video windows 2904, 2908, and 2912 have also changed. As a user moves a thumbnail viewing area lens over a thumbbar, a window, such as window 2942, is displayed on the lens. A video keyframe selected from the video keyframes extracted from the video information of the multimedia document between times  $t_1$  and  $t_2$  of the thumbnail viewing area lens is displayed in window 2942 as shown in FIG. 29B. The window 2942 disappears when the thumbnail viewing area lens is released by the user.

[0340] According to an embodiment of the present invention, the user can specify criteria and the contents of the multimedia documents that are loaded and displayed in the user interface may be searched to find locations within the multimedia documents that satisfy the user-specified criteria. Sections or locations of the multimedia document that satisfy the user specified criteria may be highlighted and displayed in interface 2900. According to an embodiment of the present invention, the user-specified criteria may include user-specified words or phrases, search queries comprising one or more terms, topics of interest, etc.

[0341] In interface 2900 depicted in FIG. 29C, a user may enter a word or phrase in input area 2944 and request that the contents of the multimedia documents be searched for the user-specified word or phrase by selecting "Find" button 2946. The word or phrase to be searched may also be selected from the common list of words 2940. In FIG. 29C, the user has specified word "Stewart".

[0342] The contents of the multimedia documents are then searched to identify locations and occurrences of the user-specified word or phrase. According to an embodiment of the present invention, text representations of information stored by the multimedia documents are searched to find locations of the user-specified word or phrase. Video keyframes may also be searched for the word or phrase. All

occurrences (“hits”) **2950** of the user-specified word or phrase in the various thumbars (i.e., in the thumbnail images in the thumbars) are highlighted as shown in **FIG. 29C**. Various different techniques may be used for highlighting the hits in the multimedia documents. For example, the individual hits may be highlighted. Ranges may also be determined based upon the hits (as describe above) and the ranges may be highlighted. Other techniques such as marks (see **FIG. 29D** described below) may also be used to mark the approximate locations of the hits. According to an embodiment of the present invention, as shown in **FIG. 29C**, colored rectangles may be drawn around lines in the thumbnail images to highlight lines that contain the search word or phrase. Video keyframes displayed in the thumbars that contain the search word or phrase may also be highlighted by drawing colored boxes around the video keyframes. Various other types of techniques may also be used. For example, if a multimedia document comprises slides information, then the slides displayed in the thumbars that contain the search word or phrase may be highlighted. The total number of occurrences **2952** of the word or phrase in the various multimedia documents is also displayed. For example, in **FIG. 29C**, the word “Stewart” occurs 31 times in the three multimedia documents.

**[0343]** The user may also form search queries comprising multiple terms (e.g., multiple words and/or phrases). As shown in **FIG. 29D**, the words or phrases that are included in a search query are displayed in area **2954**. The user can add a word or phrase to the search query by typing the word or phrase in input area **2944** (or by selecting a word from common list of words **2940**) and selecting “Add” button **2956**. The word or phrase is then added to the search query and displayed in area **2954**. In **FIG. 29D**, the word “Stewart” has been added to the search query. The user can delete or remove a word or phrase from the search query by selecting the word or phrase in area **2954** and selecting “Del” button **2960**. The user can reset or clear the search query using “Reset” button **2960**.

**[0344]** The user can also specify Boolean connectors for connecting the terms in a search query. For example, in the embodiment depicted in **FIG. 29D**, the words or phrases in a search query may be ANDed or ORed together based upon the selection of radio buttons **2958**. If the words are ORed together, then all locations of the words or phrases in the search query in the various multimedia documents are found and highlighted. If the words are ANDed together, then only those portions of the multimedia documents are highlighted as being relevant that contain all the words or phrases in the search query within a close proximity. The proximity measure may be user configurable. According to an embodiment of the present invention, the proximity measure corresponds to a number of words. For example, locations of search query words or phrases in the multimedia document are highlighted if they occur within a certain number o words of each other. Proximity can also be based upon time. In this embodiment, the locations of search query words or phrases in the multimedia document are highlighted if they occur within a specific length of time.

**[0345]** In **FIG. 29D**, the locations of the hits are shown by marks **2964** displayed in the thumbars. Each mark **2964** identifies a line in the text information printed in the thumbnails that contains the search query terms.

**[0346]** In the embodiment depicted in **FIG. 29D**, ranges have been formed based on the location of the hits. Techniques for forming ranges based on locations of hits has been previously described (see **FIGS. 20A, 20B, 21, 22, 23, 24, 25A, 25B, and 26**, and the associated description). The locations of the ranges are displayed using colored rectangles **2966**. Each rectangle identifies a range. The rectangular boxes representing the ranges thus identify portions of the multimedia documents that satisfy or are relevant to the user-specified criteria (e.g., words, phrases, topics of interest, etc.) that are used for searching the multimedia documents.

**[0347]** In embodiments of the present invention wherein contents of only one multimedia document are displayed (e.g., in **FIG. 3**), a range is identified by a start time ( $R_S$ ) and an end time ( $R_E$ ) that define the boundaries of the range, as previously described. In embodiments of the present invention wherein information from multiple multimedia documents is displayed, a range is defined by a start time ( $R_S$ ), an end time ( $R_E$ ), and an identifier identifying the multimedia document in which the range is present. Further, as described above, identifiers (e.g., a text code, numbers, etc.) may be used to identify each range. The range identifier for a range may be displayed in the rectangular box corresponding to the range or in some other location on the user interface.

**[0348]** Each thumbar in **FIG. 29D** also includes a relevance indicator **2968** that indicates the degree of relevance (or a relevancy score) of the multimedia document whose contents are displayed in the thumbar to the user-specified search criteria (e.g., user-specified word or phrase, search query, topics of interest, etc.). Techniques for determining the relevancy score or degree of relevance have been described above. According to an embodiment of the present invention, the degree of relevance for a thumbar is based upon the frequency of hits in the multimedia document whose contents are displayed in the thumbar. In the relevance indicators depicted in **FIG. 20D**, the degree of relevance of a multimedia document to the user-specified criteria is indicated by the number of bars displayed in the relevance indicators. Accordingly, the first and third multimedia documents whose contents are displayed in thumbars **2902** and **2910** are more relevant (indicated by four bars in their respective relevance indicators) to the current user-specified criteria (i.e., search query including the word “Stewart”) than the second multimedia document displayed in thumbar **2906** (only one bar in the relevance indicator). Various other techniques (e.g., relevance scores, bar graphs, different colors, etc.) may also be used to indicate the degree of relevance of the multimedia documents.

**[0349]** As previously described, various operations may be performed on ranges. The operations performed on a range may include printing a representation of the contents of the range on a paper document, saving the contents of the range, communicating the contents of the range, etc. Ranges can also be annotated or highlighted or grouped into sets. Ranges in a set of ranges can also be ranked or sorted according to some criteria that may be user-configurable. For example, ranges may be ranked based upon the relevance of each range to the user specified search criteria. According to an embodiment of the present invention, a range with higher number of hits may be ranked higher than

a range with a lower number of hits. Other techniques may also be used to rank and/or sort ranges.

[0350] The user may also select one or more ranges displayed by the user interface and perform operations on the selected ranges. According to an embodiment of the present invention, the user may select a range by clicking on a rectangle representing the range using an input device such as a mouse. In FIG. 29E, range 2970 has been selected by the user. The rectangle representing range 2970 may be highlighted (e.g., in a color different from the color of the rectangles representing the other ranges) to indicate that it has been selected. In the embodiment depicted in FIG. 29E, a web page 2971 is generated for the user-selected range and displayed in a window 2969 of user interface 2900. Web page 2971 comprises text information 2972 that represents the audio information or CC text information (or a text representation of some other type of information included in the multimedia information) for the selected range (i.e., text information that represents the audio information, CC text information, or other information occurring between times  $R_s$  and  $R_e$  of the selected range) and one or more video keyframes or images 2973 extracted from the video information corresponding to the selected range (i.e., video keyframes extract from video information occurring between times  $R_s$  and  $R_e$  of the selected range). According to an embodiment of the present invention, each image 2973 in web page 2971 is a hypertext link and when selected starts playback of video information from a time associated with the image.

[0351] As shown in FIG. 29E, a barcode 2977 may be printed for each image. The barcode may represent a time associated with the image and scanning the barcode using a barcode reader or scanner may cause playback of the video information from a time associated with the image and represented by the barcode. The playback may be displayed in the video window of a multimedia document corresponding to the selected range. Information 2974 identifying the multimedia document from which the range is selected is also displayed on web page 2971. Each barcode 2977 may also identify a start time and end time for a range. Scanning such a barcode using a barcode reader or scanner may cause playback of information corresponding to the range. Barcode 2977 may also represent a label or identifier that identifies a range. Upon scanning such a barcode, the range identifier represented by the scanned barcode may be used to determine the start and end times of the range and information corresponding to the range may then be played back.

[0352] If the user has identified user-specified criteria (e.g., a word or phrase, a topic of interest, a search query, etc.) for searching the multimedia documents, then occurrences of the user-specified criteria in web page 2971 are highlighted. For example, in FIG. 29E, the user has specified a search query containing the word "Stewart", and accordingly, all occurrences 2975 of the word "Stewart" in web page 2971 are highlighted (e.g., bolded).

[0353] A hypertext link 2976 labeled "Complete Set" is also included in web page 2971. Selection of "Complete Set" link 2971 causes generation and display of a web page that is based upon the contents of the various ranges depicted on the thumbars across all the multimedia documents that are displayed in interface 2900.

[0354] In alternative embodiments, other types of documents besides web pages may be generated and displayed

for the ranges. According to an embodiment of the present invention, a printable representation of the selected range or ranges may be generated and displayed. Further details related to generation and display of such a printable representation of multimedia information are described in U.S. application Ser. No. 10/001,895 filed Nov. 19, 2001, the entire contents of which are herein incorporated by reference for all purposes.

[0355] FIG. 29F depicts an interface 2900 in which the search query comprises multiple words, namely, "Stewart", "Imclone", and "Waksal", connected by the OR Boolean operator. All occurrences 2980 of the words in the search query are highlighted (e.g., bolded) in web page 2971. All occurrences or hits of the words in the thumbars are also marked using markers 2964. Ranges have been formed and displayed based upon the positions of the hits.

[0356] As previously described, the user-specified criteria for searching the multimedia documents may also include topics of interest. Accordingly, according to an embodiment of the present invention, the contents of the one or more multimedia documents may be searched to identify portions of the multimedia documents that are relevant to topics of interest that may be specified by a user.

[0357] FIG. 29G depicts a simplified user interface 2900 in which portions of the multimedia documents that are relevant to user specified topics of interest are highlighted. As shown in FIG. 29G, three topics of interest 2981 have been defined, namely "Airlines", "mstewart", and "baseball". Sections of the multimedia documents that are relevant to the topics of interest and that are displayed in the thumbars corresponding to the multimedia documents are highlighted using markers 2983. Ranges have been formed and displayed based upon the location of the hits. The ranges thus identify portions of the multimedia documents that are deemed relevant to the topics of interest. Portions of web page 2971 that are relevant to the topics of interest are also highlighted. Techniques for specifying topics of interest and techniques for determining portions of the multimedia documents that are relevant to one or more topics of interest have been described above and have also been described in U.S. application Ser. No. 10/001,895 filed Nov. 19, 2001, and U.S. Non-Provisional application Ser. No. 08/995,616 filed Dec. 22, 1997, the entire contents of which are herein incorporated by reference for all purposes. Relevance indicators 2982 are also displayed for each topic of interest indicating the relevance of the various multimedia documents to the topics of interest.

[0358] According to an embodiment of the present invention, a particular style or color may be associated with each topic of interest. For example, a first color may be associated with topic of interest "Airlines", a second color may be associated with topic of interest "mstewart", and a third topic of interest may be associated with the topic of interest "baseball". Portions of the multimedia document that are deemed to be relevant to a particular topic of interest may be highlighted using the style or color associated with the particular topic of interest. This enables the user to easily determine the portions of the multimedia documents that are relevant to a particular topic of interest.

[0359] FIG. 29H depicts another manner in which portions of the thumbars that are relevant to or satisfy or match user-specified criteria (e.g., words, phrases, search queries,

topics of interest, etc.) might be displayed according to an embodiment of the present invention. As shown in **FIG. 29H**, rectangular boxes are drawn on the thumbnail images in the thumbars to identify portions of the multimedia documents that are relevant to topics of interest **2981**. For each rectangular box covering a portion of a thumbar, the word or phrase or image that caused that portion of the thumbar to be chosen is displayed on the rectangular box. This enables the user to not only easily see which portions of the multimedia documents are relevant to the topics of interest (or any other user-specified criteria) but to also easily determine the word or phrase or image that resulted in the selection of the portions.

[**0360**] In **FIG. 29I**, the playback of video information for a particular multimedia document has been moved from the video window corresponding to the multimedia document to a larger video window. As shown in **FIG. 29I**, the video playback of the third multimedia document has been moved from video window **2912** to larger video window **2984**. A mode displaying the larger video window **2984** is activated by selecting "Video" tab **2986**. The switch of the display from video window **2912** to window **2984** may be performed by selecting a control **2938a** provided by control bar **2938**. A control bar **2985** comprising controls for controlling playback of the video information in window **2984** is also displayed below larger video window **2984**. Moving the playback of video information from the smaller video window **2912** (or **2904** or **2908**) to larger video window **2984** makes it easier for the user to view the video information playback. The video playback can be switched back to small window **2912** from window **2984** by selecting control **2938a** from control bar **2938** or by selecting control **2985a** from control bar **2985**.

[**0361**] Text information **2987** (e.g., CC text information, transcript of audio information, or a text representation of some other type of information included in the multimedia information) corresponding to the video playback is also displayed below larger video window **2984**. Text information **2987** scrolls along with the video playback. Each word in text information **2987** is searchable such that a user can click on a word to see how many times the selected word occurs in the contents of the multimedia documents and the locations where the word occurs. It should be noted that, as with the video playback in the smaller video window, the contents of the video played back in larger video window **2984** are affected by the position of thumbnail viewing area lens of the thumbar that displays a representation of the contents of the multimedia document whose video information is played back in larger video window **2984**.

[**0362**] As previously described, a user may also manually define ranges for one or more multimedia documents. Techniques for manually defining ranges for a multimedia document have been previously described. In the embodiment depicted in **FIG. 29J**, a button **2988** is provided that when selected initiates a mode of operation in which manual ranges can be defined by the user. According to an embodiment of the present invention, selection of button **2988** invokes a window like window **2050** depicted in **FIG. 20B**. In addition to the information depicted in **FIG. 20B**, the window also includes an entry field allowing the user to enter information identifying the multimedia document, from the multimedia documents loaded by the user interface, for which a range is to be defined. The user can also specify

the start and end times for the range. In an alternative embodiment, selection of button **2988** initiates a mode wherein the user can manually specify a range by clicking on a portion of one of the thumbars depicted in interface **2900** using an input device such as a mouse. Clicking a portion of a thumbar causes the display of a rectangular box representing the range. The user can manipulate the top and bottom edges of the rectangular box to configure the start time ( $R_S$ ) and end time ( $R_E$ ) for the range. The rectangular box itself can be moved along the thumbar.

[**0363**] According to an embodiment of the present invention, rectangular boxes representing ranges that are automatically generated (e.g., ranges generated based upon hits for user-specified criteria) and rectangular boxes representing manual ranges specified by a user may be displayed at the same time by interface **2900**. In order to differentiate between the manually generated and automatically generated ranges, different colors or styles may be used to display rectangular boxes that represent automatic ranges and boxes that represent manual ranges.

[**0364**] **FIG. 29K** depicts a user interface **2900** in which portions of the multimedia documents that the user has watched or played back are highlighted according to an embodiment of the present invention. In alternative embodiments, portions of the multimedia documents that the user has not watched or played back may be highlighted. As depicted in **FIG. 29K**, rectangular boxes **2990** are drawn on portions of the thumbars identifying portions of the multimedia documents displayed in user interface **2900** that have been watched or played by the user. The portions may have been played back or watched in smaller video windows **2904**, **2908**, or **2912**, or in larger video window **2984**, or using some output device. In this embodiment, information identifying portions of the stored multimedia information that have been output to a user (or alternatively, information identifying portions of the stored multimedia information that have not been output to a user) is stored. This feature of the present invention enables the user to easily see what sections of the multimedia documents the user has already viewed and which portions the user has yet to view. The boxes representing the viewed portions may be displayed in a particular color to differentiate them from other boxes displayed in interface **2900** such as boxes representing ranges.

[**0365**] **FIG. 30A** depicts another simplified user interface **3000** for displaying contents of one or more multimedia documents according to an embodiment of the present invention. As shown in **FIG. 30A**, contents of three multimedia documents are displayed. Three thumbars **3002**, **3004**, and **3006**, and three small video windows **3008**, **3010**, and **3012** are displayed. The contents of the multimedia documents have been searched for search query comprising terms "Stewart", "Imclone", and "Faksal". Portions of the thumbars that contain content relevant to the search query are identified by marks **3014**. Ranges have been formed based upon the hits, and rectangular boxes **3016** representing the ranges have been displayed. These features have already been described above.

[**0366**] In addition, user interface **3000** includes a number of web pages **3020-1**, **3020-2**, **3020-3**, etc. that are generated for the various ranges displayed on the thumbars by rectangular boxes. According to an embodiment of the present

invention, a web page is generated for each range displayed in the thumbars. The web pages (referred to as the “palette” view) shown in **FIG. 30A** are generated and displayed by selecting “Palette” button **3022**. Accordingly, the palette of web pages includes web pages generated for the various ranges. The palette of web pages may be displayed in the form of a scrollable list as shown in **FIG. 30A**. The user can add notes to the palette, annotate information to the web pages in the palette, and annotate parts of the multimedia documents with supplemental information. The ranges themselves may also be annotated. For example, annotation may be added to a range by adding a comment to the web page that displays information for the range.

[**0367**] In the embodiment depicted in **FIG. 30A**, each web page **3020** for a particular range comprises text information that represents the audio information, CC text information, or a text representation of some other type of information included for the particular range (i.e., text information that represents the transcribed audio information or CC text information occurring between times  $R_S$  and  $R_E$  of the particular range). The web page also includes one or more video keyframes or images extracted from the video information corresponding to the particular range. The images and the text information may be temporally synchronized or aligned. The images in the web page may be hypertext links and when selected start playback of video information from a time associated with the selected the image.

[**0368**] A barcode may be printed and associated with each image printed in a web page. The barcode may represent a time associated with the image and scanning the barcode using a barcode reader or scanner may cause playback of the video information from a time associated with the image and represented by the barcode.

[**0369**] For each range, information identifying the range and information identifying the multimedia document from which the range is selected may also displayed on the web page corresponding to the range. For example, in the embodiment depicted in **FIG. 30A**, the start and end times **3018** for each range are displayed on the web pages. Identifiers **3021** identifying the ranges are also displayed. Since the multimedia documents in **FIG. 30A** correspond to television video recordings, each web page corresponding to a range also displays an icon **3023** associated with the TV network that broadcast the information for the range.

[**0370**] Occurrences of user-specified criteria in each web page are highlighted. For example, in the embodiment depicted in **FIG. 30A**, the search query includes terms “Stewart”, “Imclone”, and “Faksal”, and occurrences of these terms in the web pages are highlighted. Various different techniques may be used for highlighting the terms such as bolding, use of colors, different styles, use of balloons, boxes, etc. As previously described, the search query terms may also be highlighted in the representations displayed in the thumbars.

[**0371**] A lens **3024** is displayed emphasizing or covering an area of a web page corresponding to a currently selected range. For example, in the embodiment depicted in **FIG. 30A**, a range has been selected in thumbar **3006**, and lens **3024** is displayed covering a portion of web page **3020-4** corresponding to the selected range. The portion of web page **3020-4** covered by lens **3024** is displayed in a larger window **3026**. The user can change the positions of lens

**3024** along the length of web page **3020-4**. The portion of the web page displayed in window **3026** is changed such that it continues to correspond to the portion of web page **3020-4** covered by lens **3024**. In this manner, the user may use lens **3024** to navigate the contents of the selected web page. The user may also use the scrollbar provided by window **3026** to scroll through the web page displayed in window **3026**. The position of lens **3024** over web page **3020-4** is changed such that it continues to correspond to the portion of web page displayed in window **3026**.

[**0372**] A user may chose another range by clicking on that range (i.e., by clicking on the rectangle representing the range) in the thumbars using an input device such as a mouse. In response, the position of lens **3024** is changed such that it is displayed over a web page in the palette view corresponding to the newly selected range. The portion of the web page in the palette view covered by lens **3024** is then displayed in window **3026**. For example, as depicted in **FIG. 30B**, a different range **3030** has been selected by the user. The user may select this range by clicking on a rectangular box corresponding to the range in thumbar **3002**. In response, lens **3024** is drawn covering a portion of web page **3020-1** corresponding to range **3030**. The portion of web page **3020-1** covered or emphasized by lens **3024** is displayed in window **3026**.

[**0373**] According to an embodiment of the present invention, a use may also select a range by selecting a web page corresponding to the range from the palette of web pages. The user may select a web page by clicking on that web page using an input device such as a mouse. In response, lens **3024** is displayed on the selected web page. The portion of the web page in the palette view covered by lens **3024** is displayed in window **3026**. The rectangular box representing the range corresponding to the newly selected web page is also highlighted to indicate selection of the range.

[**0374**] As described above, embodiments of the present invention can display representations of information stored by one or more multimedia documents that may have been recorded during the same or different timelines. The user can specify criteria such as words, phrases, search queries including multiple terms, topics of interest, etc., and portions of the multimedia documents that are relevant to or contain the user-specified criteria are highlighted using markers, boxes representing ranges, etc. Embodiments of the present invention can accordingly be used to compare the contents of multiple multimedia documents.

[**0375**] For example, the recordings of three different television programs such as “Closing Bell”, “60 Minute II”, and “Business Center” may be displayed as depicted in **FIG. 29A**, and searched for user-specified criteria (e.g., words, phrases, terms in the search query, topics of interest, etc.). Portions of the three programs that are relevant to or match the user-specified criteria are highlighted. Such an ability to search across multiple multimedia documents is not provided by conventional tools. Further, based upon the results of the searches displayed by the interface, the user can easily determine the relevance of the television programs to the user criteria. Embodiments of the present invention thus can be used to analyze the contents of the multimedia documents with respect to each other. The visualization of the search results is often useful for obtaining a feel for the contents of the multimedia documents.

**[0376]** As another example, if the user is interested in the Imclone/Martha Stewart scandal, the user can form a search query including the terms “Stewart”, “Imclone”, and “Wak-sal” (or other words related to the scandal) and portions of the representations of the multimedia documents that are displayed by the user interface and that contain the search query terms are highlighted using markers, colors, etc. Ranges may also be formed based upon the search hits and depicted on the interface using colored boxes to highlight the relevant sections. By viewing the portions of the multimedia documents highlighted in the interface, the user can easily determine how much information related to the scandal is contained in the multimedia documents and the locations in the multimedia documents of the relevant information. The user can also determine the distribution of the relevant information in the multimedia documents. The multimedia documents can also be compared to each other with regards to the search query. Embodiments of the present invention thus provide a valuable tool for a user who wants to analyze multiple multimedia documents.

**[0377]** The analysis and review of multiple multimedia documents is further facilitated by generating and displaying web pages corresponding to the ranges (that may be automatically generated or manually specified) displayed in the interface. The web pages generated for the ranges allow the user to extract, organize, and gather the relevant portions of the multiple multimedia documents.

**[0378]** Embodiments of the present invention also provide the user the ability to simultaneously watch a collection of multimedia documents. For example, the user can watch the contents of multiple video recordings or video clips. Various controls are provided for controlling the playback of the multimedia information. Portions of the multimedia documents played back by the user may be highlighted. The user can accordingly easily determine portions of the multimedia documents that the user has already viewed and portions that have not been viewed.

**[0379]** As previously described, several operations can be performed using ranges. These operations include, for example, printing a representation of the contents of a range on a paper document, saving the contents of a range, communicating the contents of a range, annotating a range, etc. Ranges may also be grouped (e.g., grouped into sets) and operations performed on the groups. For example, ranges in a set of ranges can also be ranked or sorted based upon some criteria that might be user-configurable. For example, ranges may be ranked based upon the relevance of each range to the user specified search criteria. According to an embodiment of the present invention, a range with higher number of hits may be ranked higher than a range with a lower number of hits. Other techniques may also be used to rank and/or sort ranges.

#### **[0380] Printing Multimedia Information**

**[0381]** As previously indicated, multimedia information from one or more multimedia documents displayed by the user interfaces described above may be printed on a paper medium to produce a multimedia paper document. Accordingly, a multimedia paper document may be generated for the one or more multimedia documents. The term “paper” or “paper medium” may refer to any tangible medium on which information can be printed, written, drawn, imprinted, embossed, etc.

**[0382]** According to an embodiment of the present invention, for each multimedia document, a printable representation is generated for the recorded information stored by the multimedia document. Since the recorded information may store information of different types such as audio information, video information, closed-caption (CC) text information, slides information, whiteboard information, etc., according to an embodiment of the present invention, the printable representation of the recorded information may comprise printable representations of one or more types of information. The printable representation for the recorded information, which may include printable representations for one or more types of information that make up the recorded information, can be printed on a paper medium to generate a multimedia paper document. Various different techniques may be used for generating a printable representation for the multimedia information. Examples of techniques for generating a printable representation and printing the printable representation on a paper medium to produce a multimedia paper document are described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001, the entire contents of which are herein incorporated by reference for all purposes.

**[0383]** The printable representation can then be printed on a paper medium. The term “printing” includes printing, writing, drawing, imprinting, embossing, and the like. According to an embodiment of the present invention, the printable representation is communicated to a paper document output device (such as a printer, copier, etc.) that is configured to print the printable version on a paper medium to generate a paper document. Various different techniques may be used for printing the printable representation on a paper medium. According to an embodiment of the present invention, the printing is performed according to the techniques described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001, the entire contents of which are herein incorporated by reference for all purposes.

**[0384]** In other embodiments of the present invention, instead of generating a multimedia paper document for the entire contents of the multimedia documents, a multimedia paper document may be generated only for the ranges displayed in the graphical user interface. In this embodiment, a printable representation is generated for multimedia information corresponding to the ranges, and the printable representation is then printed on a paper medium. Since multimedia information corresponding to a range may comprise information of one or more types, the printable representation of the multimedia information corresponding to the range may comprise printable representations of one or more types. Various different techniques may be used for generating a printable representation for the multimedia information corresponding to the ranges. For example, the described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001, maybe used.

**[0385]** **FIG. 31** depicts a simplified user interface **3100** that may be used to print contents of one or more multimedia documents or contents corresponding to ranges according to an embodiment of the present invention. Interface **3100** depicted in **FIG. 31** is merely illustrative of an embodiment of the present invention and does not limit the scope of the present invention. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

Graphical user interface **3100** may be invoked by selecting a command or button provided by the interfaces described above.

[**0386**] As depicted in **FIG. 31**, a user can specify that only information corresponding to the ranges is to be printed by selecting checkbox **3101**. If checkbox **3101** is not selected it implies that all the contents of the one or more multimedia documents that have been loaded are to be printed. The user can indicate that information corresponding to all the displayed ranges is to be printed by selecting checkbox **3102**. Alternatively, the user can specifically identify the ranges to be printed by entering the range identifiers in input boxes **3104**. For example, if the ranges are identified by numbers assigned to the ranges, then the user can enter the numbers corresponding to the ranges to be printed in boxes **3104**. If the range identifiers are serially numbers, a list of ranges may be specified.

[**0387**] Selection of "Print" button **3106** initiates printing of the contents of the ranges or contents of the loaded multimedia documents. User interface **3100** can be canceled by selecting "Cancel" button **3108**.

[**0388**] Several options are provided for controlling the manner in which information corresponding to the ranges or information from the loaded multimedia documents is printed. For example, a format style may be selected for printing the information on the paper medium. In the embodiment depicted in **FIG. 31**, the user can select from one of three different styles **3110** by selecting a checkbox corresponding to the style. **FIGS. 32A, 32B, and 32C** depict pages printed according to the three styles selectable from interface **31** depicted in **FIG. 31** according to an embodiment of the present invention. **FIG. 32A** depicts a page printed according to Style 1. **FIG. 32B** depicts a page printed according to Style 2. **FIG. 32C** depicts a page printed according to Style 3. Various other styles may also be provided in alternative embodiments of the present invention.

[**0389**] The user can also select different styles **3112** for printing keyframes extracted from video information. For example, in the embodiment depicted in **FIG. 31**, the user can select a style wherein one keyframe is printed for each barcode (or alternatively, one barcode will be printed for each printed keyframe) or multiple (e.g., 4) keyframes are be printed for each barcode. **FIGS. 33A and 33B** depict pages printed according to the two keyframes styles selectable from interface **31** depicted in **FIG. 31**. **FIG. 33A** depicts a page wherein one keyframe is printed per barcode. **FIG. 33B** depicts a page wherein four keyframes are printed per barcode. Other styles may also be provided in other embodiments of the present invention.

[**0390**] A list of printers **3116** (or any other paper document output device that can generate a print representation of multimedia information, e.g., a copier, a facsimile machine, etc.) is displayed. The user can select one or more printers from list **3116** to print the multimedia information on a paper medium. The user can select a specific copier for performing the printing (or copying) by selecting "Send to Copier" checkbox **3113** and identifying the copier to be used in box **3114**.

[**0391**] According to an embodiment of the present invention, the printable representation of multimedia information

corresponding to the selected ranges or to the multimedia documents can be stored in memory. For example, the printable representation can be stored as a PDF file. A name for the file can be specified in entry box **3118**.

[**0392**] According to an embodiment of the present invention depicted in **FIG. 31**, the user has the option of indicating whether sections of the printable representation that comprise words or phrases that satisfy or match user-specified criteria are to be highlighted when the printable representation is printed on the paper medium. The user may activate this option by selecting checkbox **3120**. When this option has been selected, words or phrases in the multimedia information corresponding to the multimedia documents or the selected ranges that are relevant to the topics of interest, or match words or phrases specified by the user or search query terms are highlighted when printed on paper. Various different techniques may be used for highlighting the word or phrases on paper.

[**0393**] A text marker that relates the barcodes to the printed text information may be printed by selecting checkbox **3122**.

[**0394**] As described above, the user can specify that only information corresponding to ranges is to be printed by selecting checkbox **3101**. If desired, for each range, the user can specify a buffer time period to be added to the start and end of the range in entry box **3126**. For example, if a buffer time period of 5 seconds is specified, for each range, information corresponding to 5 seconds before the range start and corresponding to 5 seconds after the range end is printed along with the information corresponding to the range.

[**0395**] Embodiments of the present invention can also print a cover sheet for the printed information (either for information corresponding to ranges of information corresponding to multimedia document contents). The user can specify that a coversheet should be printed in addition to printing the contents of the ranges or multimedia documents by selecting checkbox **3128**. The cover sheet may provide a synopsis or summary of the printed contents of the multimedia documents or ranges.

[**0396**] Various different techniques may be used for printing a coversheet. Different styles **3130** for the coversheet may be selected. **FIGS. 34A, 34B, and 34C** depict examples of coversheets that may be printed according to an embodiment of the present invention. Examples of techniques for generating and printing coversheets are described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001, the entire contents of which are herein incorporated by reference for all purposes. Examples of different coversheets are also described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001.

[**0397**] The coversheets may be used for several different purposes. As previously indicated, a coversheet provides a synopsis or summary of the printed contents of the multimedia documents or ranges. The coversheet may also provide a summary of information stored on a storage device. For example, for multimedia information stored on a CD, a coversheet may be generated based upon the contents of the CD that summarizes what contents of the CD. For example, as shown in **FIG. 34C**, a coversheet is generated and used as a cover for a jewel case that may store the CD. In the

embodiment depicted in **FIG. 34C**, the barcodes printed on the CD may be used to access or index into the multimedia information stored on the CD. Techniques for using the barcodes printed on the coversheet to access the multimedia information are described in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001. Various other uses of coversheets are also envisioned within the scope of the present invention.

**[0398]** A user may also elect to print only the coversheet and not the contents of the ranges or the multimedia documents by selecting checkbox **3132** in **FIG. 31**. This is useful for example when a cover sheet is to be generated for providing an index into information stored on a storage device.

**[0399]** The coversheets depicted in **FIGS. 34A, 34B, and 34C** each display a limited number of keyframes sampled (e.g., sampled uniformly every N seconds) from the multimedia information for which the coversheet is generated. The sampling interval may be specified by the user. For example, in the embodiment depicted in **FIG. 31**, the user can enter the sampling interval in entry box **3134**.

**[0400]** The user is also provided the ability to control the quality of the printed image. In the embodiment depicted in **FIG. 31**, the user can select one of three options **3136**.

**[0401]** **FIGS. 35A, 35B, 35C, 35D, and 35E** depict a paper document printed for ranges according to an embodiment of the present invention. The ranges may have been generated automatically or may have been manually specified by the user, as described above. The information corresponding to the ranges may be stored in one or more multimedia documents. The pages depicted in **FIGS. 325, 35B, 35C, 35D, and 35E** are merely illustrative of an embodiment of the present invention and do not limit the scope of the present invention. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

**[0402]** The document depicted in **FIGS. 35A, 35B, 35C, 35D, and 35E** is printed for ranges selected from three multimedia documents. The three multimedia documents are television program recordings, namely, "Money and Markets" program captured from the CNN/fn channel (Channel **358**), "Closing Bell" program captured from CNBC channel (Channel **355**) and "Street Sweep" program also captured from the CNN/fn channel (Channel **358**).

**[0403]** As depicted in **FIGS. 35A-E**, the contents for the ranges of the three recorded programs are printed sequentially. The contents of ranges from the "Money and Markets" program recording multimedia document are printed on the pages depicted in **FIGS. 35A and 35B**, the contents of ranges from the "Closing Bell" program recording multimedia document are printed on the pages depicted in **FIGS. 35C and 35D**, and the contents of ranges from the "Street Sweep" program recording multimedia document are printed on the page depicted in **FIG. 35E**.

**[0404]** Information **3500** identifying the multimedia documents from which the ranges are selected is printed as shown in **FIGS. 35A, 35C, and 35E**. In the embodiment depicted in **FIGS. 35A-E**, the information identifying each multimedia document includes the name of the television program, information identifying the channel from which the program was recorded, the duration of the recording, and the date and

time of the recording. Other types of information related to the multimedia documents may also be printed.

**[0405]** The start of each range is indicated by a bar **3502**. Accordingly, contents of two ranges have been printed from the "Money and Markets" multimedia document, contents of four ranges have been printed from the "Closing Bell" multimedia document, and contents of three ranges have been printed from the "Street Sweep" multimedia document. Information **3504** related to the range is also printed in each bar **3502**. In the embodiment depicted in **FIGS. 35A-E**, the information related to the ranges includes, an identifier for the range, a start time ( $R_s$ ) and end time ( $R_e$ ) for the range, and the span of the range. Other types of information related to each range may also be printed.

**[0406]** The information printed for each range includes text information **3506** and one or more images **3508**. The text information is a printable representation of the audio information (or CC text, or a text representation of some other type of information included in the multimedia information) corresponding to the range. Occurrences of words or phrases occurring in the printed text information that are relevant to topics of interest, or match user-specified words or phrases or search criteria are highlighted. For example, for the embodiment depicted in **FIGS. 35A-E**, the user has defined a search query containing terms "Stewart", "Imclone", and "Waksal". Accordingly, all occurrences of these search query terms are highlighted (using underlining) in the printed text sections for the various ranges. Various different techniques may also be used to highlight the words such as bolded text, different fonts or sizes, italicized text, etc.

**[0407]** Images **3508** printed for each range represent images that are extracted from the video information for the range. Several different techniques may be used for extracting video keyframes from the video information of the range and for identifying the keyframes to be printed. Examples of these techniques are described above and in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001. Various different styles may be used for printing the information. For example, the user may chose from styles **3110** and **3112** depicted in **FIG. 31**.

**[0408]** Barcodes **3510** are also printed for each range. In the embodiment depicted in **FIGS. 35A-E**, a barcode **3510** is printed for each image **3508** and is placed below the image. Various different styles may be used for printing the barcodes. For example, in the embodiment depicted in **FIG. 31**, two different styles **3112** are provided for printing barcodes, namely, a first style in which one barcode is printed per keyframe (as shown in **FIGS. 35A-E**) and a second style in which a barcode is printed for every four keyframes.

**[0409]** According to an embodiment of the present invention depicted in **FIGS. 35A-E**, each barcode printed below an image represents a time associated with the image. Barcodes **3510** provide a mechanism for the reader of the paper document to access multimedia information using the paper document. According to an embodiment of the present invention, scanning of a barcode using a device such as a scanner, barcode reader, etc. initiates playback of multimedia information from the multimedia document corresponding to the barcode from the time represented by the barcode. The playback may occur on any output device. For example, the information may be played back in a window of the previously described GUIs displayed on a computer screen.

[0410] Each barcode **3510** may also identify a start time and end time for a range. Scanning such a barcode using a barcode reader or scanner may cause playback of information corresponding to the range. Each barcode **3510** may also represent a label or identifier that identifies a range. In this embodiment, upon scanning such a barcode, the range identifier represented by the scanned barcode may be used to determine the start and end times of the range and information corresponding to the range may then be played back.

[0411] The document depicted in FIGS. **35A-E** thus provides a paper interface for accessing stored multimedia information. Further information related to using a paper interface for accessing multimedia information is discussed in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001. Other user-selectable identifiers such as watermarks, glyphs, text identifiers, etc. may also be used in place of barcodes in alternative embodiments of the present invention. The user-selectable identifiers might be printed in a manner that does not reduce or affect the overall readability of the paper document.

[0412] A set of barcodes **3512** are also printed at the bottom of each paper page of the paper document depicted in FIGS. **35A-E**. Barcodes **3512** allow a user to initiate and control playback of multimedia information using the paper document. According to an embodiment of the present invention, each barcode corresponds to a command for controlling playback of multimedia information. Five control barcodes **3512** are printed in the embodiment shown in FIGS. **35A-E**. Control barcode **3512-1** allows the user to playback or pause the playback. For example, a user may scan a barcode **3510** and then scan barcode **3512-1** to initiate playback of information from the time represented by scanned barcode **3510**. The user may rescan barcode **3512-1** to pause the playback. The playback can be fast forwarded by selecting barcode **3512-2**. The user may perform a rewind operation by selecting barcode **3512-3**. The playback can be performed in an enhanced mode by selecting barcode **3512-4**. Enhanced mode is an alternative GUI that provides additional viewing controls and information (e.g., a specialized timeline may be displayed, controls may be provided such that on screen buttons on a PDA may be used to navigate the information that is played back). Further details related to enhanced mode display are described in U.S. application Ser. No. 10/174, 522, filed Jun. 17, 2002, the entire contents of which are incorporated herein for all purposes. Specific modes of operation can be entered into by selecting barcode **3512-5**. Barcodes for various other operations may also be provided in alternative embodiments of the present invention. Information related to barcodes for controlling playback of information is discussed in U.S. patent application Ser. No. 10/001,895, filed Nov. 19, 2001.

[0413] Although specific embodiments of the invention have been described, various modifications, alterations, alternative constructions, and equivalents are also encompassed within the scope of the invention. The described invention is not restricted to operation within certain specific data processing environments, but is free to operate within a plurality of data processing environments. Additionally, although the present invention has been described using a particular series of transactions and steps, it should be apparent to those skilled in the art that the scope of the present invention is not limited to the described series of transactions and steps. For example, the processing for

generating a GUI according to the teachings of the present invention may be performed by server **104**, by client **102**, by another computer, or by the various computer systems in association.

[0414] Further, while the present invention has been described using a particular combination of hardware and software, it should be recognized that other combinations of hardware and software are also within the scope of the present invention. The present invention may be implemented only in hardware, or only in software, or using combinations thereof.

[0415] The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense. It will, however, be evident that additions, subtractions, deletions, and other modifications and changes may be made thereunto without departing from the broader spirit and scope of the invention as set forth in the claims.

What is claimed is:

1. A computer-implemented method of displaying information, the method comprising:

displaying a first representation of first stored information, the first stored information comprising information of a first type and information of a second type, the first representation comprising a representation of information of the first type included in the first stored information and a representation of the information of the second type included in the first stored information; and

highlighting one or more portions of the first representation, the highlighted one or more portions of the first representation corresponding to portions of the first representation that include a first criterion.

2. The method of claim 1 further comprising:

displaying a second representation of second stored information, the second stored information comprising information of a first type and information of a second type, the second representation comprising a representation of information of the first type included in the second stored information and a representation of information of the second type included in the second stored information; and

highlighting one or more portions of the second representation, the highlighted one or more portions of the first representation corresponding to portions of the second representation that include the first criterion.

3. The method of claim 2 wherein:

the first criterion comprises a word;

highlighting one or more portions of the first representation comprises highlighting one or more portions of the first representation that contain the word; and

highlighting one or more portions of the second representation comprises highlighting one or more portions of the second representation that contain the word.

4. The method of claim 2 wherein:

the first criterion comprises a topic of interest;

highlighting one or more portions of the first representation comprises highlighting one or more portions of the

first representation that are relevant to the topic of interest; and

highlighting one or more portions of the second representation comprises highlighting one or more portions of the second representation that are relevant to the topic of interest.

5. The method of claim 2 wherein:

the information of the first type included in the first stored information is audio information;

the information of the second type included in the first stored information is video information;

the information of the first type included in the second stored information is audio information; and

the information of the second type included in the second stored information is video information.

6. The method of claim 5 wherein:

the representation of information of the first type included in the first representation comprises text information obtained from transcribing the audio information included in the first stored information;

the representation of information of the second type included in the first representation comprises one or more video keyframes extracted from the video information included in the first stored information;

the representation of information of the first type included in the second representation comprises text information obtained from transcribing the audio information included in the second stored information; and

the representation of information of the second type included in the second representation comprises one or more video keyframes extracted from the video information included in the second stored information.

7. The method of claim 6 wherein:

highlighting one or more portions of the first representation comprises highlighting one or more lines from the text information included in the first representation that include the first criterion; and

highlighting one or more portions of the second representation comprises highlighting one or more lines from the text information included in the second representation that include the first criterion.

8. The method of claim 6 wherein:

highlighting one or more portions of the first representation comprises highlighting one or more video keyframes from the video keyframes included in the first representation that include the first criterion; and

highlighting one or more portions of the second representation comprises highlighting one or more video keyframes from the video keyframes included in the second representation that include the first criterion.

9. The method of claim 5 further comprising playing back the video information included in the first stored information and the video information included in the second stored information.

10. The method of claim 2 further comprising:

receiving a signal indicating selection of a first portion from the one or more highlighted portions of the first representation or the second representation; and

displaying a representation of the contents of the first selected portion.

11. The method of claim 10 wherein the representation of the contents of the first selected portion is a web page generated based upon the contents of the first selected portion.

12. The method of claim 2 further comprising:

displaying a representation for each portion in the one or more highlighted portions of the first representation and the second representation, wherein the representation for a portion is generated based upon the contents of the portion.

13. The method of claim 2 further comprising playing back the first type of information included in the first stored information and the first type of information stored in the second stored information.

14. A computer-implemented method of displaying multimedia information, the method comprising:

displaying a first thumbnail comprising a representation of information of a first type included in a first recorded information;

displaying a second thumbnail comprising a representation of information of a second type included in the first recorded information;

displaying a third thumbnail comprising a representation of information of a first type included in a second recorded information; and

displaying a fourth thumbnail comprising a representation of information of a second type included in the second recorded information.

15. The method of claim 14 further comprising:

receiving information identifying a set of one or more words; and

highlighting one or more portions of the first thumbnail and the third thumbnail that comprise at least one word from the set of words.

16. The method of claim 15 further comprising:

highlighting one or more portions of the second thumbnail and the fourth thumbnail that comprise at least one word from the set of words.

17. The method of claim 16 further comprising:

receiving information indicating selection of a first portion from the one or more highlighted portions of the second or the fourth thumbnail, the first portion corresponding to a portion of information of the second type included in the first recorded information or the second recorded information; and

outputting the information of the second type corresponding to the first portion.

18. The method of claim 14 further comprising:

receiving information identifying a set of one or more topics of interest; and

highlighting one or more portions of the first thumbnail and the third thumbnail that comprise content relevant to at least one topics of interest from the set of topics of interest.

**19.** The method of claim 18 further comprising:  
 highlighting one or more portions of the second thumbnail and the fourth thumbnail that comprise content relevant to at least one topics of interest from the set of topics of interest.

**20.** The method of claim 14 wherein:  
 the information of the first type included in the first recorded information is audio information;  
 the information of the second type included in the first recorded information is video information;  
 the information of the first type included in the second recorded information is audio information;  
 the information of the second type included in the second recorded information is video information;  
 the first thumbnail comprises text information obtained from the audio information included in the first recorded information;  
 the second thumbnail comprises one or more video keyframes extracted from the video information included in the first recorded information;  
 the third thumbnail comprises text information obtained from the audio information included in the second recorded information; and  
 the fourth thumbnail comprises one or more video keyframes extracted from the video information included in the second recorded information.

**21.** The method of claim 14 wherein:  
 the information of the first type included in the first recorded information is closed-caption text information;  
 the information of the second type included in the first recorded information is video information;  
 the information of the first type included in the second recorded information is closed-caption text information;  
 the information of the second type included in the second recorded information is video information;  
 the first thumbnail comprises text information obtained from the closed caption text information included in the first recorded information;  
 the second thumbnail comprises one or more video keyframes extracted from the video information included in the first recorded information;  
 the third thumbnail comprises text information obtained from the closed-caption text information included in the second recorded information; and  
 the fourth thumbnail comprises one or more video keyframes extracted from the video information included in the second recorded information.

**22.** A method of displaying information included in a first recorded information and a second recorded information, the first recorded information comprising audio information and video information, the second recorded information comprising

audio and video information, the method comprising:

displaying a first representation of information included in the first recorded information, the first representation comprising a first thumbnail and a second thumbnail, the first thumbnail comprising text information obtained from the audio information included in the first recorded information, the second thumbnail comprising one or more keyframes extracted from the video information included in the first recorded information; and

displaying a second representation of information included in the second recorded information, the second representation comprising a third thumbnail and a fourth thumbnail, the third thumbnail comprising text information obtained from the audio information included in the second recorded information, the fourth thumbnail comprising one or more keyframes extracted from the video information included in the second recorded information.

**23.** The method of claim 22 further comprising:

receiving information identifying a user criterion; and

highlighting one or more portions of the first representation and the second representation that include the user criterion, wherein a highlighted portion of the first representation covers a section of the first thumbnail and the second thumbnail and a highlighted portion of the second representation covers a section of the third thumbnail and the fourth thumbnail.

**24.** The method of claim 23 wherein highlighting the one or more portions of the first representation and the second representation comprises:

identifying locations in the first representation and the second representation that include the user criterion; and

determining the one or more portions of the first representation and the second representation to be highlighted based upon the identified locations.

**25.** The method of claim 23 wherein the user criterion is a set of one or more words or phrases.

**26.** The method of claim 23 wherein the user criterion is a set of one or more topics of interest.

**27.** The method of claim 23 further comprising:

displaying a document for each highlighted portion of the first representation and the second representation, wherein the document for a portion comprises text information from the section of the first or the third thumbnail covered by the portion and one or more keyframes from the section of the second or the fourth thumbnail covered by the portion.

**28.** The method of claim 23 further comprising:

receiving information indicating selection of a first highlighted portion from the one or more highlighted portions of the first representation;

displaying a document for first highlighted portion, wherein the document for the first highlighted portion comprises text information from the section of the first thumbnail covered by the first highlighted portion and one or more keyframes from the section of the second thumbnail covered by the first highlighted portion.

**29.** The method of claim 23 further comprising:  
 receiving information indicating selection of a first highlighted portion from the one or more highlighted portions of the first representation; and  
 outputting video information from the first recorded information corresponding to the first highlighted portion.

**30.** A computer-implemented method of displaying information, the method comprising:

displaying a representation of stored information;  
 receiving information indicative of one or more portions of the stored information that have been output; and  
 highlighting one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have been output.

**31.** The method of claim 30 wherein the stored information is video information.

**32.** The method of claim 30 wherein the stored information is audio information.

**33.** A computer-implemented method of displaying information, the method comprising:

displaying a representation of stored information;  
 receiving information indicative of one or more portions of the stored information that have been output; and  
 highlighting one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have not been output.

**34.** An apparatus for displaying information, the apparatus comprising:

a processor; and  
 a display;

wherein the processor is configured to:

display, on the display, a first representation of first stored information, the first stored information comprising information of a first type and information of a second type, the first representation comprising a representation of information of the first type included in the first stored information and a representation of the information of the second type included in the first stored information; and

highlight one or more portions of the first representation displayed on the display, the highlighted one or more portions of the first representation corresponding to portions of the first representation that include a first criterion.

**35.** The apparatus of claim 34 wherein the processor is further configured to:

display a second representation of second stored information on the display, the second stored information comprising information of a first type and information of a second type, the second representation comprising a representation of information of the first type included in the second stored information and a representation of information of the second type included in the second stored information; and

highlight one or more portions of the second representation, the highlighted one or more portions of the first representation corresponding to portions of the second representation that include the first criterion.

**36.** The apparatus of claim 35 wherein:

the first criterion comprises a word;

the processor is configured to highlight one or more portions of the first representation that contain the word; and

the processor is configured to highlight one or more portions of the second representation that contain the word.

**37.** The apparatus of claim 35 wherein:

the first criterion comprises a topic of interest;

the processor is configured to highlight one or more portions of the first representation that are relevant to the topic of interest; and

the processor is configured to highlight one or more portions of the second representation that are relevant to the topic of interest.

**38.** The apparatus of claim 35 wherein:

the information of the first type included in the first stored information is audio information;

the information of the second type included in the first stored information is video information;

the information of the first type included in the second stored information is audio information; and

the information of the second type included in the second stored information is video information.

**39.** The apparatus of claim 38 wherein:

the representation of information of the first type included in the first representation comprises text information obtained from transcribing the audio information included in the first stored information;

the representation of information of the second type included in the first representation comprises one or more video keyframes extracted from the video information included in the first stored information;

the representation of information of the first type included in the second representation comprises text information obtained from transcribing the audio information included in the second stored information; and

the representation of information of the second type included in the second representation comprises one or more video keyframes extracted from the video information included in the second stored information.

**40.** The apparatus of claim 39 wherein:

the processor is configured to highlight one or more lines from the text information included in the first representation that include the first criterion; and

the processor is configured to highlight one or more lines from the text information included in the second representation that include the first criterion.

**41.** The apparatus of claim 39 wherein:

the processor is configured to highlight one or more video keyframes from the video keyframes included in the first representation that include the first criterion; and

the processor is configured to highlight one or more video keyframes from the video keyframes included in the second representation that include the first criterion.

**42.** The apparatus of claim 38 wherein the processor is configured to cause playback of the video information included in the first stored information and the video information included in the second stored information.

**43.** The apparatus of claim 35 wherein the processor is configured to:

receive a signal indicating selection of a first portion from the one or more highlighted portions of the first representation or the second representation; and

display, on the display, a representation of the contents of the first selected portion.

**44.** The apparatus of claim 43 wherein the representation of the contents of the first selected portion is a web page generated based upon the contents of the first selected portion.

**45.** The apparatus of claim 35 wherein the processor is configured to display, on the display, a representation for each portion in the one or more highlighted portions of the first representation and the second representation, wherein the representation for a portion is generated based upon the contents of the portion.

**46.** The apparatus of claim 35 wherein the processor is configured to cause playback the first type of information included in the first stored information and the first type of information stored in the second stored information.

**47.** An apparatus for displaying multimedia information, the apparatus comprising:

a processor; and

a display; and

wherein the processor is configured to:

display, on the display, a first thumbnail comprising a representation of information of a first type included in a first recorded information;

display, on the display, a second thumbnail comprising a representation of information of a second type included in the first recorded information;

display, on the display, a third thumbnail comprising a representation of information of a first type included in a second recorded information; and

display, on the display, a fourth thumbnail comprising a representation of information of a second type included in the second recorded information.

**48.** The apparatus of claim 47 wherein the processor is configured to:

receive information identifying a set of one or more words; and

highlight one or more portions of the first thumbnail and the third thumbnail that comprise at least one word from the set of words.

**49.** The apparatus of claim 48 wherein the processor is configured to highlight one or more portions of the second thumbnail and the fourth thumbnail that comprise at least one word from the set of words.

**50.** The apparatus of claim 49 wherein the processor is configured to:

receive information indicating selection of a first portion from the one or more highlighted portions of the second

or the fourth thumbnail, the first portion corresponding to a portion of information of the second type included in the first recorded information or the second recorded information; and

cause the information of the second type corresponding to the first portion to be output.

**51.** The apparatus of claim 47 wherein the processor is configured to:

receive information identifying a set of one or more topics of interest; and

highlight one or more portions of the first thumbnail and the third thumbnail that comprise content relevant to at least one topics of interest from the set of topics of interest.

**52.** The apparatus of claim 51 wherein the processor is configured to highlight one or more portions of the second thumbnail and the fourth thumbnail that comprise content relevant to at least one topics of interest from the set of topics of interest.

**53.** The apparatus of claim 47 wherein:

the information of the first type included in the first recorded information is audio information;

the information of the second type included in the first recorded information is video information;

the information of the first type included in the second recorded information is audio information;

the information of the second type included in the second recorded information is video information;

the first thumbnail comprises text information obtained from the audio information included in the first recorded information;

the second thumbnail comprises one or more video keyframes extracted from the video information included in the first recorded information;

the third thumbnail comprises text information obtained from the audio information included in the second recorded information; and

the fourth thumbnail comprises one or more video keyframes extracted from the video information included in the second recorded information.

**54.** The apparatus of claim 47 wherein:

the information of the first type included in the first recorded information is closed-caption text information;

the information of the second type included in the first recorded information is video information;

the information of the first type included in the second recorded information is closed-caption text information;

the information of the second type included in the second recorded information is video information;

the first thumbnail comprises text information obtained from the closed-caption text information included in the first recorded information;

the second thumbnail comprises one or more video keyframes extracted from the video information included in the first recorded information;

the third thumbnail comprises text information obtained from the closed-caption text information included in the second recorded information; and

the fourth thumbnail comprises one or more video keyframes extracted from the video information included in the second recorded information.

**55.** An apparatus for displaying information included in a first recorded information and a second recorded information, the first recorded information comprising audio information and video information, the second recorded information comprising audio and video information, the apparatus comprising:

a display; and

a processor;

wherein the processor is configured to:

display, on the display, a first representation of information included in the first recorded information, the first representation comprising a first thumbnail and a second thumbnail, the first thumbnail comprising text information obtained from the audio information included in the first recorded information, the second thumbnail comprising one or more keyframes extracted from the video information included in the first recorded information; and

display, on the display, a second representation of information included in the second recorded information, the second representation comprising a third thumbnail and a fourth thumbnail, the third thumbnail comprising text information obtained from the audio information included in the second recorded information, the fourth thumbnail comprising one or more keyframes extracted from the video information included in the second recorded information.

**56.** The apparatus of claim 55 wherein the processor is configured to:

receive information identifying a user criterion; and

highlight one or more portions of the first representation and the second representation that include the user criterion, wherein a highlighted portion of the first representation covers a section of the first thumbnail and the second thumbnail and a highlighted portion of the second representation covers a section of the third thumbnail and the fourth thumbnail.

**57.** The apparatus of claim 56 wherein the processor is configured to:

identify locations in the first representation and the second representation that include the user criterion; and

determine the one or more portions of the first representation and the second representation to be highlighted based upon the identified locations.

**58.** The apparatus of claim 56 wherein the user criterion is a set of one or more words or phrases.

**59.** The apparatus of claim 56 wherein the user criterion is a set of one or more topics of interest.

**60.** The apparatus of claim 56 wherein the processor is configured to display, on the display, a document for each

highlighted portion of the first representation and the second representation, wherein the document for a portion comprises text information from the section of the first or the third thumbnail covered by the portion and one or more keyframes from the section of the second or the fourth thumbnail covered by the portion.

**61.** The apparatus of claim 56 the processor is configured to:

receive information indicating selection of a first highlighted portion from the one or more highlighted portions of the first representation; and

display, on the display, a document for first highlighted portion, wherein the document for the first highlighted portion comprises text information from the section of the first thumbnail covered by the first highlighted portion and one or more keyframes from the section of the second thumbnail covered by the first highlighted portion.

**62.** The apparatus of claim 56 wherein the processor is configured to:

receive information indicating selection of a first highlighted portion from the one or more highlighted portions of the first representation; and

cause video information from the first recorded information corresponding to the first highlighted portion to be output.

**63.** An apparatus for displaying information, the apparatus comprising:

a processor; and

a memory coupled to the processor, the memory configured to store a program for controlling the processor;

wherein the processor under control of the program is configured to:

cause a representation of stored information to be displayed;

receive information indicative of one or more portions of the stored information that have been output; and

cause one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have been output to be highlighted.

**64.** An apparatus for displaying information, the apparatus comprising:

a processor; and

a memory coupled to the processor, the memory configured to store a program for controlling the processor;

wherein the processor under control of the program is configured to:

cause a representation of stored information to be displayed;

receive information indicative of one or more portions of the stored information that have been output; and

cause one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have not been output to be highlighted.

**65.** A program stored on a computer-readable storage medium for displaying information, the program comprising:

code for displaying a first representation of first stored information, the first stored information comprising information of a first type and information of a second type, the first representation comprising a representation of information of the first type included in the first stored information and a representation of the information of the second type included in the first stored information; and

code for highlighting one or more portions of the first representation, the highlighted one or more portions of the first representation corresponding to portions of the first representation that include a first criterion.

**66.** The program of claim 65 further comprising:

code for displaying a second representation of second stored information, the second stored information comprising information of a first type and information of a second type, the second representation comprising a representation of information of the first type included in the second stored information and a representation of information of the second type included in the second stored information; and

code for highlighting one or more portions of the second representation, the highlighted one or more portions of the first representation corresponding to portions of the second representation that include the first criterion.

**67.** The program of claim 66 wherein:

the first criterion comprises a word;

the code for highlighting one or more portions of the first representation comprises code for highlighting one or more portions of the first representation that contain the word; and

the code for highlighting one or more portions of the second representation comprises code for highlighting one or more portions of the second representation that contain the word.

**68.** The program of claim 66 wherein:

the first criterion comprises a topic of interest;

the code for highlighting one or more portions of the first representation comprises code for highlighting one or more portions of the first representation that are relevant to the topic of interest; and

the code for highlighting one or more portions of the second representation comprises code for highlighting one or more portions of the second representation that are relevant to the topic of interest.

**69.** A program stored on a computer-readable storage medium for displaying multimedia information, the program comprising:

code for displaying a first thumbnail comprising a representation of information of a first type included in a first recorded information;

code for displaying a second thumbnail comprising a representation of information of a second type included in the first recorded information;

code for displaying a third thumbnail comprising a representation of information of a first type included in a second recorded information; and

code for displaying a fourth thumbnail comprising a representation of information of a second type included in the second recorded information.

**70.** The program of claim 69 further comprising:

code for receiving information identifying a set of one or more words; and

code for highlighting one or more portions of the first thumbnail and the third thumbnail that comprise at least one word from the set of words.

**71.** The program of claim 70 further comprising:

code for highlighting one or more portions of the second thumbnail and the fourth thumbnail that comprise at least one word from the set of words.

**72.** The program of claim 71 further comprising:

code for receiving information indicating selection of a first portion from the one or more highlighted portions of the second or the fourth thumbnail, the first portion corresponding to a portion of information of the second type included in the first recorded information or the second recorded information; and

code for outputting the information of the second type corresponding to the first portion.

**73.** A program stored on a computer-readable storage medium for displaying information included in a first recorded information and a second recorded information, the first recorded information comprising audio information and video information, the second recorded information comprising audio and video information, the program comprising:

code for displaying a first representation of information included in the first recorded information, the first representation comprising a first thumbnail and a second thumbnail, the first thumbnail comprising text information obtained from the audio information included in the first recorded information, the second thumbnail comprising one or more keyframes extracted from the video information included in the first recorded information; and

code for displaying a second representation of information included in the second recorded information, the second representation comprising a third thumbnail and a fourth thumbnail, the third thumbnail comprising text information obtained from the audio information included in the second recorded information, the fourth thumbnail comprising one or more keyframes extracted from the video information included in the second recorded information.

**74.** The program of claim 73 further comprising:

code for receiving information identifying a user criterion; and

code for highlighting one or more portions of the first representation and the second representation that include the user criterion, wherein a highlighted portion of the first representation covers a section of the first thumbnail and the second thumbnail and a highlighted portion of the second representation covers a section of the third thumbnail and the fourth thumbnail.

**75.** The program of claim 74 wherein the code for highlighting the one or more portions of the first representation and the second representation comprises:

code for identifying locations in the first representation and the second representation that include the user criterion; and

code for determining the one or more portions of the first representation and the second representation to be highlighted based upon the identified locations.

**76.** The program of claim 74 further comprising:

code for receiving information indicating selection of a first highlighted portion from the one or more highlighted portions of the first representation; and

code for outputting video information from the first recorded information corresponding to the first highlighted portion.

**77.** A program stored on a computer-readable storage medium for displaying information, the program comprising:

code for displaying a representation of stored information;

code for receiving information indicative of one or more portions of the stored information that have been output; and

code for highlighting one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have been output.

**78.** A program stored on a computer-readable storage medium for displaying information, the program comprising:

code for displaying a representation of stored information;

code for receiving information indicative of one or more portions of the stored information that have been output; and

code for highlighting one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have not been output.

**79.** An apparatus for displaying information, the apparatus comprising:

means for displaying a first representation of first stored information, the first stored information comprising information of a first type and information of a second type, the first representation comprising a representation of information of the first type included in the first stored information and a representation of the information of the second type included in the first stored information;

means for highlighting one or more portions of the first representation, the highlighted one or more portions of the first representation corresponding to portions of the first representation that include a first criterion.

means for displaying a second representation of second stored information, the second stored information comprising information of a first type and information of a second type, the second representation comprising a representation of information of the first type included

in the second stored information and a representation of information of the second type included in the second stored information; and

means for highlighting one or more portions of the second representation, the highlighted one or more portions of the first representation corresponding to portions of the second representation that include the first criterion.

**80.** An apparatus for displaying multimedia information, the apparatus comprising:

means for displaying a first thumbnail comprising a representation of information of a first type included in a first recorded information;

means for displaying a second thumbnail comprising a representation of information of a second type included in the first recorded information;

means for displaying a third thumbnail comprising a representation of information of a first type included in a second recorded information; and

means for displaying a fourth thumbnail comprising a representation of information of a second type included in the second recorded information.

**81.** An apparatus for displaying information included in a first recorded information and a second recorded information, the first recorded information comprising audio information and video information, the second recorded information comprising audio and video information, the apparatus comprising:

means for displaying a first representation of information included in the first recorded information, the first representation comprising a first thumbnail and a second thumbnail, the first thumbnail comprising text information obtained from the audio information included in the first recorded information, the second thumbnail comprising one or more keyframes extracted from the video information included in the first recorded information; and

means for displaying a second representation of information included in the second recorded information, the second representation comprising a third thumbnail and a fourth thumbnail, the third thumbnail comprising text information obtained from the audio information included in the second recorded information, the fourth thumbnail comprising one or more keyframes extracted from the video information included in the second recorded information.

**82.** An apparatus for displaying information, the apparatus comprising:

means for displaying a representation of stored information;

means for receiving information indicative of one or more portions of the stored information that have been output; and

means for highlighting one or more portions of the representation of the stored information corresponding to the one or more portions of the stored information that have been output.

\* \* \* \* \*