

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 October 2007 (25.10.2007)

PCT

(10) International Publication Number  
**WO 2007/120686 A2**

(51) International Patent Classification:  
*G06Q 10/00* (2006.01)

(21) International Application Number:  
PCT/US2007/008898

(22) International Filing Date: 10 April 2007 (10.04.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/790,769 10 April 2006 (10.04.2006) US  
60/790,770 10 April 2006 (10.04.2006) US

(71) Applicants (for all designated States except US): **QUI-VIDI** [FR/FR]; 74, Rue Du Rocher, F-75008 Paris (FR). **BUSINGER, Peter** [US/US]; 373 Park Avenue, Suite 205, Scotch Plains, NJ 07076 (US).

(71) Applicants and

(72) Inventors: **CHARDON, Gael** [FR/FR]; 27, Rue Clavel, F-75019 Paris (FR). **DUIZABO, Olivier** [FR/FR]; 74, Rue Du Rocher, F-75008 Paris (FR). **MAZARS, Gilles** [FR/FR]; 117, Rue Caulaincourt, F-75018 Paris (FR). **PRANDONI, Paolo** [IT/FR]; 11, Rue Des Recollets, F-75010 Paris (FR).

(74) Agent: **BUSINGER, Peter, A.**; 373 Park Avenue, Suite 205, Scotch Plains, NJ 07076 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

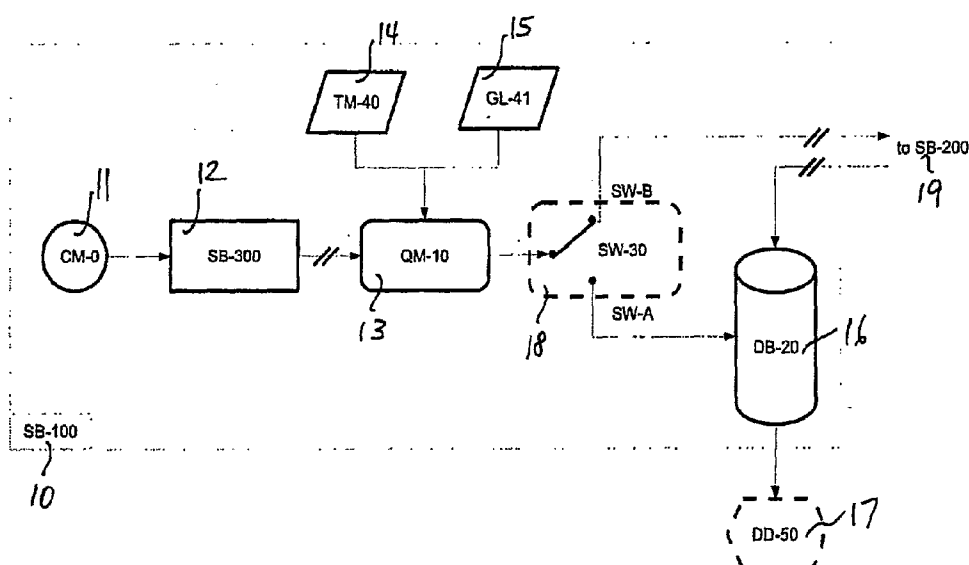
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: AUTOMATED MEDIA CONTENT ADAPTATION AND DELIVERY



(57) Abstract: In advertising, irrespective of the delivery medium, a concern with audience targeting can be addressed by context-based, usage-based and user-based techniques. In a fast-growing sector in the advertising industry, namely that of digital signage, user data for a display installations are obtained by methods designated as indirect, polling or rating. Beyond context- and usage-based targeting, for facilitating audience-based content adaptation or targeting, physical attributes of subjects in front of an installation can be ascertained in an automated fashion. The attributes can be combined with customary context and usage-pattern data in selecting messages suitable for display at the installation. The combined data can be supplied to a number of potential content providers with an invitation to bid in an auction of display time.



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## **AUTOMATED MEDIA CONTENT ADAPTATION AND DELIVERY**

### Technical Field

The invention concerns delivery of media content to an audience and,  
5 more particularly, of content comprising advertisements.

### Background of the Invention

In advertising, irrespective of the delivery medium, there is a primary concern with audience targeting. Subject to a constraint on deployment costs, the  
10 goal is to maximize the impact of a campaign by addressing each message to a subset of a potential audience, with the subset chosen for maximized relevancy of the message to members of the subset. The following are examples of choices for the audience subsets:

Context-based. The choice of the audience subset can be implicit in  
15 context targeted by a medium. For example, magazines directed to a female audience will mostly carry advertisement directed to women.

Usage-based. The audience subset may be determined by usage patterns. TV advertisement placed in early-afternoon time slots, for example, will be mostly directed to children.

20 User-based. For greater selectivity and efficacy, a campaign can be targeted to individuals or groups of individuals based on their known preferences and/or purchasing patterns. Examples are direct-mail and telephone marketing campaigns. While user-based subset selections are most powerful, often they are costly to deploy as they require detailed market-based analysis. But one medium  
25 where deployment costs can be kept low is the internet, as individual users can be characterized on the basis of their usage patterns, on a real-time basis even. A particularly fortuitous situation presents itself to the advertising services that most search engines offer: on entering queries, users implicitly declare an interest in a subject, so that advertising messages can be displayed on their screens according to a  
30 criterion of relevance to a query.

A fast-growing sector in the advertising industry is that of digital signage, including audiovisual out-of-home advertising installations utilizing digital display devices such as flat-screen monitors. The following represent examples of ways for obtaining user data for such display installations:

5                   Indirect Method. The visibility of an installation is estimated from exogenous parameters such as its location with respect to people flows or location attendance.

Polling Method. A random sample of a potential audience— in a shopping mall, for example— is explicitly asked about the visibility of an installation  
10                   by pollsters.

Ratings Method. A chosen sample of a potential audience— e.g. of households with a TV— is asked to keep a record of the visibility of an installation. This might include written diary entries, or might involve an automated method relying on a device which a user activates to identify him/herself.

15                   With the indirect method it is difficult to establish a confirmed baseline for assessing the validity of an estimation, and even more difficult to estimate OTS (Opportunity to See) or “impact” for a potential audience. Additionally, no reactions, opinions and other higher-level information about the effect of the installation on an audience can be inferred.

20                   Polling and rating methods, while potentially more powerful than the indirect method, still suffer from two major drawbacks. First, in the ratings method especially, where the population is chosen a priori, there will be statistical inaccuracy in the choice of the sample population. And second, these methods are obtrusive, requiring direct engagement with potential viewers. In particular, the measurement is  
25                   likely to be influenced adversely by the interaction with a pollster or the imposed adherence to a specified log-keeping protocol. Also, the methods are inaccurate because of an audience’s limited recall ability.

                  The need for accurate and punctual audience measurements is particularly strong in light of current proliferation of communications media. In view  
30                   of mounting encroachment by alternative media, more accurate home-usage TV

ratings are required. Rapidly growing use, of out-of-home display devices such as flat screens and dynamic billboards, calls for automatic audience measurement for validation purposes. Such needs are particularly pressing in fast-paced advertising where sufficiently fine resolution in time and localization are needed.

5

#### Summary of the Invention

Beyond context- and usage-based targeting, for facilitating audience-based content adaptation or targeting in digital signage and other display arrangements, physical attributes of subjects in front of an installation can be  
10 ascertained in an automated fashion. The attributes can include gender, age, clothing style and hairstyle, for example, which can be combined with customary context and usage-pattern data in selecting messages suitable for display at the installation. Customary data include geographical location and time. The combined data can be supplied to a number of potential content providers with an invitation for them to bid  
15 in an auction of display time.

#### Brief Description of the Drawing

Fig. 1 is a block schematic of a content adaptation system in accordance with an exemplary embodiment of the invention.

20 Fig. 2 is a block schematic of a functional module or subsystem of the content adaptation system of Fig. 1.

Fig. 3 is a block schematic of a functional module which can be coupled to the content adaptation system of Fig. 1, in accordance with a further exemplary embodiment of the invention.

25 Fig. 4 is a block schematic of a functional module or subsystem of the content adaptation system of Fig. 1, alternative to and more elaborate than the functional module of Fig. 2.

#### Detailed Description

Fig. 1 shows a content adaptation system 10 (also labeled as SB100 in the figure) including a video camera 11 (CM-0) which can be understood as placed in the vicinity of a signage installation or screen for capturing the scene in front of the screen. Signals from the video camera 11 are input to a subsystem 12 (SB-300) for processing into a series of high-level descriptors for any audience in front of the screen. Examples of descriptors are the number of people, the number of people actually watching the screen, audience gender breakdown, audience age breakdown, viewers' features, e.g. glasses, beard and hairstyle, and group dynamics such as family, crowd or single individuals.

From the subsystem 12 the generated descriptors are fed to a module 13 (QM-10) which is further provided with time input from clock 14 (TM-40) and with geographical identifiers from data support 15 (GL-41). Geographical identifiers can play a qualifier role similar to that of parameters employed in usage-based targeting methods, as there typically is a relationship between venue and typical audience base. Examples of geographical identifiers are designations of state, city and street, type of venue such as grocery store, apparel store or train station, and location within venue, e.g. produce aisle, ticket counter and the like.

Module 13 translates the audience data from subsystem 12, the time stamp, and the geographical index into a query for database 16 (DB-20), namely a repository of pre-tagged display material such as video files suitable for reproduction on display device 17 (DD-50). The tags associated with the display material are metadata which describe the suitability of the content for a given audience, time and location. The metadata are in standardized format and are exploitable via the standard database queries generated by module 13. Examples of queries – here translated into human language – are: “retrieve content suitable for a female audience at lunchtime in a grocery store”, “retrieve content suitable for a group of three young males on a Friday afternoon in a movie theater”, or “retrieve content suitable for a family group”.

Each item in database 16 is also marked by a unique owner's ID. The output data from module 13 can be differently routed by switch 18 (SW-30). If

switch 18 is in position SW-A, the system 10 can be said to function in local mode in which a query from module 13 is routed directly to the local database 16, for the results which best match the query to be sent to the display device 17. If the switch 18 is in position SW-B, the system 10 can be said to function in centralized mode in which a query from module 13 is sent to a subsystem 19 (SB-200) which modifies or augments the query before returning it for submission to the local database 16. Such modification most notably includes the addition of a specific owner's ID to the query parameters. The result which best matches the modified query and which exactly matches the added owner's ID parameter is sent to the display device 17.

Fig. 2 shows a subsystem 20 (also labeled as SB-300 in the figure) for analyzing the images captured by the camera 11 in order to extract audience data including high-level descriptors for the installation's audience. The camera signal is input to a module 21 (PD-10) for detecting the presence of people in front of the installation. Such detection triggers the module 22 (PT-11) for tracking the ensemble of detections and for establishing spatio-temporal relations for the detections. Further, a module 23 (PA-20) analyzes the detections to extract high-level descriptors such as individual height, body size, type of clothes, colors and the like, and a face detection module 24 (FD-21) extracts additional information from the detections such as gender, age, hair style and the like. A final module 25 (FE-22) extracts from the detections a directionality descriptor for the individual's gaze, which can be used to determine whether a person in the audience is actually looking at the installation. Module 25 can optionally connect to a local database 26 (DDB-51) which contains previously stored descriptors for known faces in order to perform face recognition and to output a named descriptor of the person. The aggregate of all the information extracted by the modules is the output of subsystem 20.

Fig. 3 shows a subsystem 30 (also labeled as SB-200 in the figure) which is active in centralized mode. The subsystem 30 includes a module 31 (CO-80) which is connected to a plurality of distinct, geographically distributed subsystems 32 (SB-100-1, SB-100-2, . . . , SB-100-N). Database queries constructed locally within each of the subsystems 32 are forwarded to module 31 via

communications links which can be standard TCP-IP links (Transmission Control Protocol – Internet Protocol). A database 33 (DB-81) mirrors the contents of the subsystems 32, so that queries issued from the subsystems 32 to the database 33 yield consistent results. The module 31 is connected to a plurality of bidding modules 34  
5 (UI-82-1, . . . , UI-82-M); again, the connections can be standard TCP-IP links. The bidding modules 82 are accessible to the owners of display material stored on database 33 and on the local databases 16 at each subsystem 10. Each module 34 is identified by the same owner ID as used in the database.

Upon receipt of a query from one of the subsystems 10, module 32  
10 retrieves the best-matched results from database 33 and notifies the bidding modules 34 which correspond to the owners' IDs associated with the results, thereby to trigger an auction of display time. On their examination of the results, the notified owners can respond by bidding for display time, and the highest bidder can be awarded control of the display for a period of time. Conveniently for awarding, module 31  
15 complements the query with the highest bidder's ID and returns the query to the local module 10. Only the display material marked with that bidder's ID will be forwarded to the display device(s) 17.

Fig. 4 shows a subsystem 40 (also labeled as SB-100 in the figure) alternative to the system of Fig. 1 and 2, in real time providing for enhanced audience  
20 data gathering. The subsystem 40 includes a camera 41 (CM-0) for capturing the scene in front of the installation for which audience measurement is sought. By wired or wireless connection, the camera 41 is connected to a module 42 (PD-10) for detecting the presence of people in front of the installation and for triggering the operation of subsequent systems modules. In operation, module 43 (PT-11) tracks  
25 individuals detected by module 42, while maintaining spatio-temporal coherence between detections. Module 44 (MA-12) analyzes the motion patterns of individuals to arrive at higher-level information such as the presence of groups of people, their average spacing, their mutual interactions and the like. The resulting data are time-stamped and recorded onto a local database 45 (DB-50).



Data of each person detected by module 43 is input also to one of a plurality of subsystems 46 (SB-101). The gateway to subsystem 46 is an analysis module 47 (PA-20) which, for each person in the audience, extracts biometric features such as height, body size, gait and the like. These data are stored in the database 45. The module 47 further specializes into a face detection module 48 (FD-21) which detects the face of the person and extracts a set of compact features for analysis by a module 49 (FE-22). There result a series of high-level descriptors for the person, e.g. the person's gender, age, hairstyle and the like. These data are stored on local database 45. Module 49 optionally can connect to a local database 50 (DB-51) which includes previously stored descriptors for known faces in order to facilitate face recognition and to output a named descriptor for the person. Module 49 also outputs a directionality descriptor for the person's gaze, which is used by decision module 51 (GD-23) to determine whether a person in the audience is actively looking at the installation. If so, a module 52 (AS-24) is triggered to start a log-taking activity for the attention span relative to the tracked audience member, optionally links the attention span to the content displayed by the installation at the time which is available from an external source 53 (CI-40). These data are stored on the local database 45 and updated for the entire duration of the attention span.

A complete system SB-200 of Fig. 2 can include a plurality of interconnected subsystems 40, with the camera 41 of each subsystem 40 located in the vicinity of a visual installation for which monitoring is sought. Interconnection between subsystems 40 can be effected by any suitable communication means, e.g. a TCP/IP link. The local data of each subsystem 40 is accessible by a main module 31 which gathers the audience measurement data from the individual subsystems 40 and stores it locally on a database 33. The data can be organized and arranged according to specific user requirement communicated to the module 31 via a control interface 34. Examples of queries include the global effective audience for the aggregate of installations, the global OTS for the aggregate installations, the effective audience for a single installation, the audience breakdown by location, the audience breakdown by

gender, the audience breakdown by age, the average attention span of the audience and the like.

Techniques as described above can be implemented in software, on general-purpose computing platforms, for example. For certain tasks, e.g. those of audience analysis modules, known routines can be used as described in the open literature. Database functionalities can be implemented by standard platforms such as one known as MySQL, for example. The bidding modules include a straightforward two-way data exchange interface which can be implemented as a web page with HTTP forms (Hypertext Transfer Protocol).

## CLAIMS

1. A method for automated authorization of delivery of media data for content display at least at one display station, comprising the steps of:

- 5           (a) obtaining information including audience information indicative of characteristics of audience present at the display station;
- (b) offering at least a portion of the information to a plurality of potential content providers, thereby to assist each potential content provider in estimating audience receptiveness to the respective potential content provider's display content;
- 10           (c) accepting auction bids for display time from the potential content providers; and
- (d) authorizing a high bidder's content display.

2. The method of claim 1, wherein audience characteristics include at least  
15 one of number of people, number of people watching a display device, audience gender breakdown, audience facial features, audience hairstyle, and audience socialization.

3. The method of claim 1, further comprising including location and time data  
20 in the information.

4. The method of claim 3, wherein the location data comprise at least one of state, city, street, type of venue and location within a venue.

25           5. The method of claim 1, wherein audience information comprises information derived from video of at least one of an actual audience and a potential audience.

30           6. The method of claim 5, wherein the video comprises video captured by a camera at the at least one display station.

7. The method of claim 5, wherein the audience information further comprises information derived from a database.

5           8. A service for automated authorization of delivery of media data for content display at least at one display station, comprising the steps of:

          (a) obtaining information including audience information indicative of characteristics of audience present at the display station;

          (b) offering at least a portion of the information to a plurality of potential  
10       content providers, thereby to assist each potential content provider in estimating audience receptiveness to the respective potential content provider's display content;

          (c) accepting auction bids for display time from the potential content providers; and

          (d) authorizing a high bidder's content display.

15

9. The service of claim 8, wherein audience characteristics include at least one of number of people, number of people watching a display device, audience gender breakdown, audience facial features, audience hairstyle, and audience socialization.

20

10. The service of claim 8, further comprising including location and time data in the information.

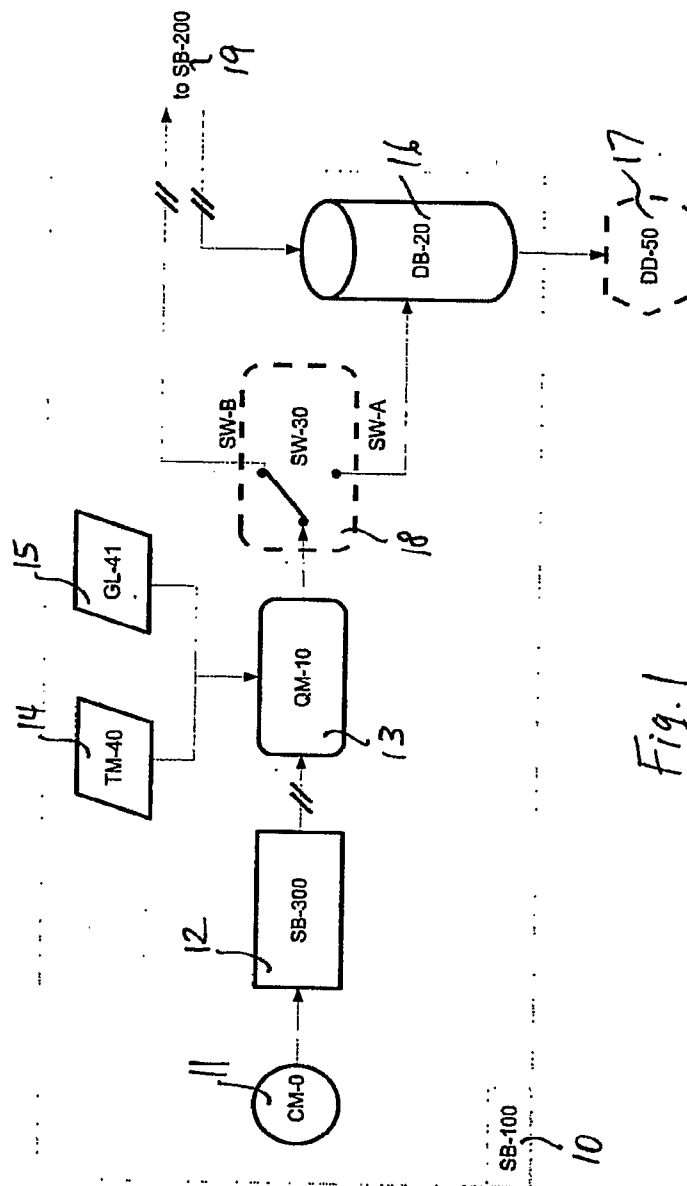
11. The service of claim 10, wherein the location data comprise at least one  
25       of state, city, street, type of venue and location within a venue.

30

12. The service of claim 8, wherein audience information comprises information derived from video of at least one of an actual audience and a potential audience.

13. The service of claim 12, wherein the video comprises video captured by a camera at the at least one display station.

14. The service of claim 12, wherein the audience information further  
5 comprises information derived from a database.



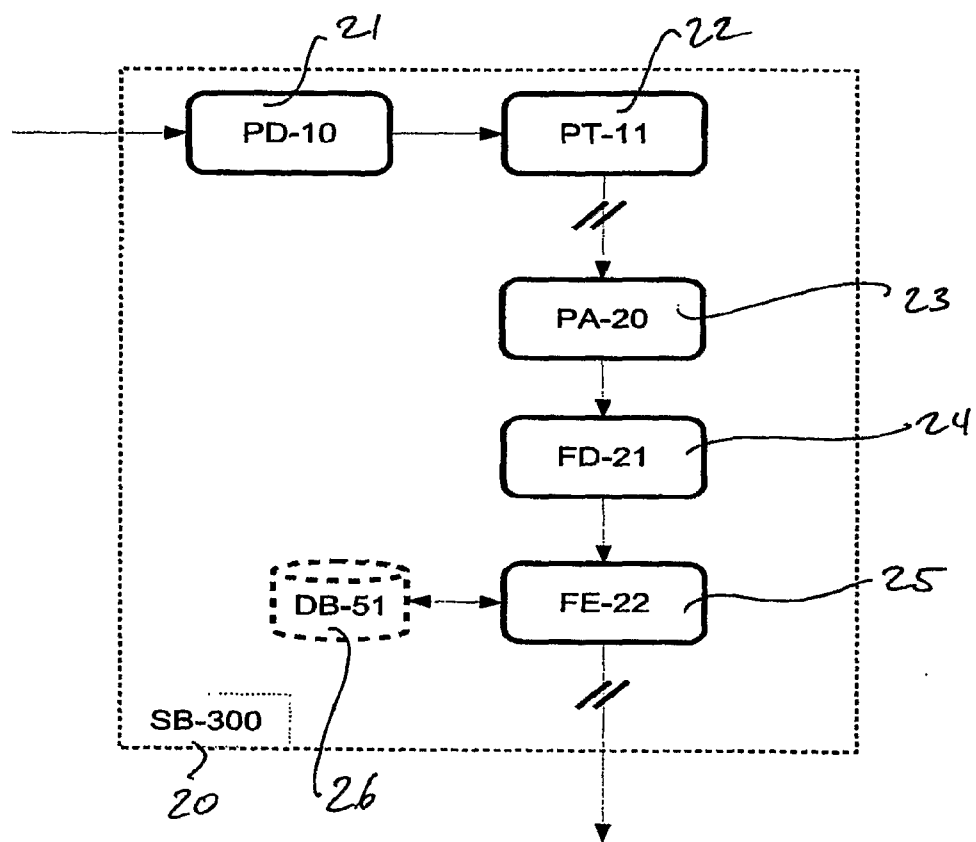
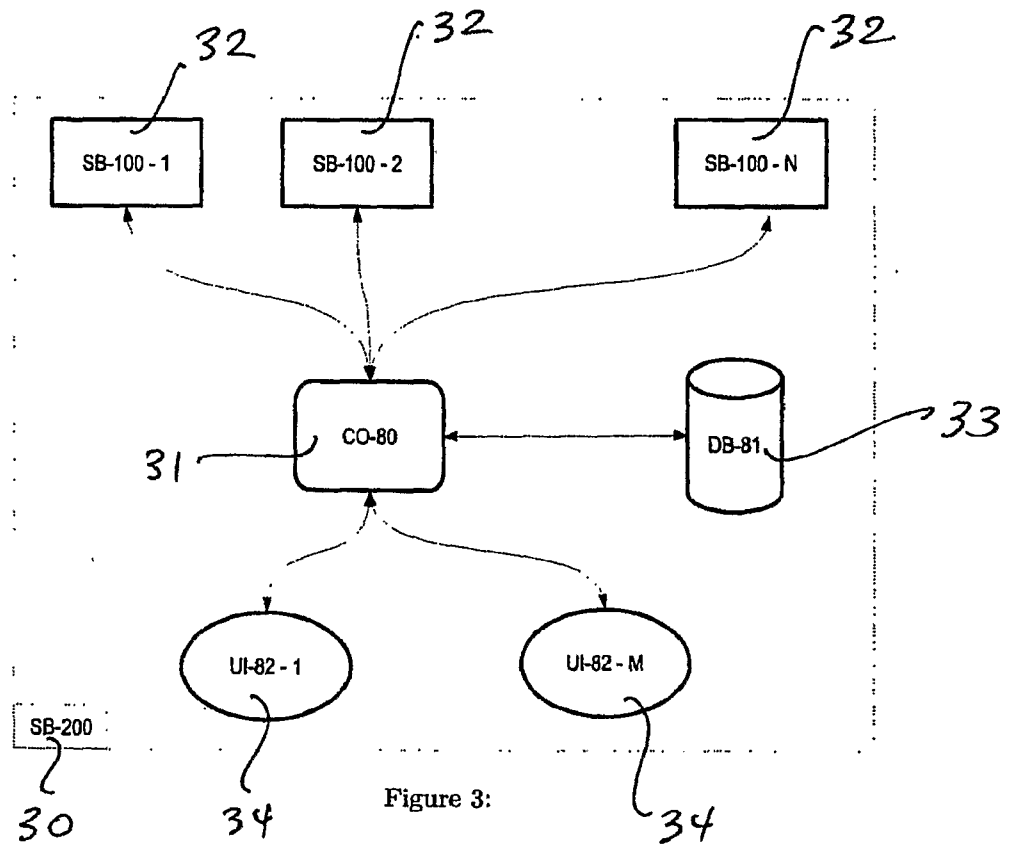


Figure 2:





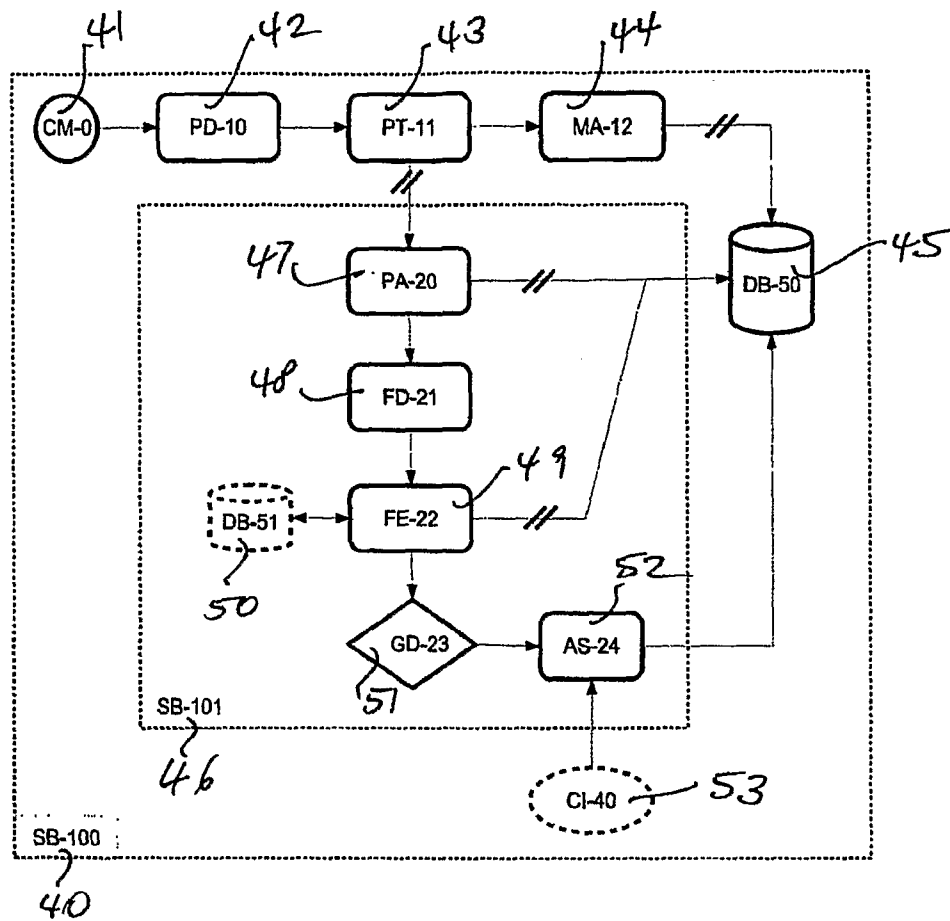


Fig. 4