(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 August 2006 (31.08.2006)

PCT

(10) International Publication Number
WO 2006/091306 A2

(51) International Patent Classification:
G06F 19/00 (2006.01)

(21) International Application Number:
PCT/US2006/002513

(22) International Filing Date: 19 January 2006 (19.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0503504.3 21 February 2005 (21.02.2005) GB

(71) Applicant *(for all designated States except US)*: **MOTOROLA, INC.** [US/US]; 1303 East Algonquin Road, Schaumburg, Illinois 60196 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TEH, Jonathan Soon Yew,** [GB/GB]; 12 Heathfield, Basingstoke Hampshire RG22 4PA (GB). **BRADY, Michael,** [GB/GB]; Information Engineering Building, Dept. Of Engineering Science, Univesity Of Oxford, Parks Road, Oxford Oxfordshire OX12BJ (GB). **DOLBEAN, Catherine Mary,** [GB/GB]; 27 The Polygon, Southampton Greater London SO15 2BP (GB).

(74) Agent: **NICHOLS, Daniel K.**; 1303 East Algonquin Road, Schaumburg, Illinois 60196 (US).

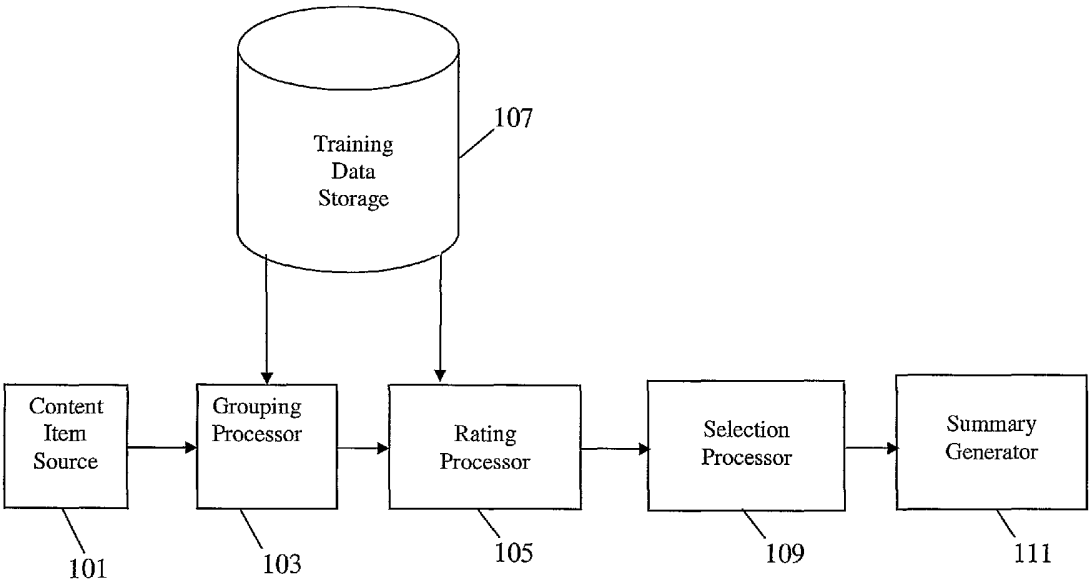
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC,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



(57) **Abstract:** An apparatus comprises a grouping processor (103) grouping sequential events of a content item into event groups. A training data storage (107) comprises training data in the form of information for a large number of training content items comprising event class sequences. The training data comprises training content summaries for each of the training content items. The training content summaries comprise training summary items which are linked to the event class sequences. A rating processor (105) determines a rating for each group in response to a probability characteristic of the training data. The rating processor (105) determines how often information for an event class sequence corresponding to the event group is included in the training content items relative to the frequency at which this event class sequence occurs in the training data. A selection processor (109) selects event groups.

WO 2006/091306 A2

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.

APPARATUS AND METHOD FOR GENERATING A CONTENT SUMMARY

Field of the invention

The invention relates to an apparatus and method for generating a content summary and in particular to automatic generation of content summaries for content items.

Background of the Invention

In recent years, the availability and provision of for example multimedia and entertainment content has increased substantially. For example, the number of available TV and radio channels has increased substantially and the popularity of the Internet has provided new content distribution means. Consequently, users are increasingly provided with a plethora of different types of multimedia content from different sources. In order to identify and select the desired content, the user must typically process large amounts of information which can be very cumbersome and impractical.

Accordingly, significant resources have been invested in research into techniques and algorithms that may provide an improved user experience and assist a user in identifying and selecting content. One such area of research is information filtering. Information filtering tackles the problem of information overload, which is a problem that is becoming more and more pressing as users are confronted with increasing volumes of e.g. multimedia data (including e.g. text, audio and video content

items), much of which is either unwanted or irrelevant. Accordingly, information filtering may provide functionality for selecting information which is of particular interest to the user. Furthermore, information filtering may include generation of new information that extracts, condenses or modifies available information in order to provide a more suitable information provision to the user.

Some of the problems that must be addressed by algorithms for information filtering include:

1) How to determine the importance of different pieces of information to the specific user;

2) How to allocate resources (e.g. a summary time limit) such that the user is provided with the most salient information;

3) How to provide a context of the pieces of information that are presented to the user; and

4) How to design the system in such a way as to allow domain-independent summarisation.

Although many algorithms for information filtering and in particular for content item summary generation have been proposed these tend to be suboptimal in one or more of these areas.

Many algorithms for generating content summaries are based on applying predefined rules, criterions or axioms to the content items. However, such algorithms tend to be

very inflexible and are typically not suited for different types of content items or applications. Furthermore, such methods tend to be complex and require accurate models, rules, criterions or axioms to be derived. However, the derivation of such models is very difficult and time consuming and tend to end up with models which are inaccurate for most content items. It furthermore results in suboptimal personalisation as the models or axioms may not be suited for accurate personalisation.

For instance, probabilistic models have been proposed for generation of summaries for video content. As a specific example, approaches for generation of summaries for video, often termed "video summarisation", for video sequences from football matches are presented in A. Ekin, A. M. Tekalp and R. Mehrotra "Automatic Soccer Video Analysis and Summarization", IEEE Transactions on Image Processing, vol 12, no 7, 2003 and J. Assfalg, M. Bertini, C. Colombo, A. del Bimbo and W. Nunziati, "Semantic Annotation of Soccer Highlights", Computer Vision and Image Understanding vol 92, no 2-3, 2003. In these examples, semantic labels are extracted from the video signal, and subsequently, all the semantic objects that are recognised in the video signal are included in the summary. Although these methods allow summary generation, the selection of what information to include is merely based on object or event recognition and is not optimised for the specific user.

As another example, Natural Language Processing techniques may be used for text summarisation. These techniques use term probability analysis to indicate the

significance of certain words, coupled with other importance indicators such as word position, lexical cues (such as "the best" or "in conclusion") and topic identification. However, such text-based techniques are inappropriate for many applications as more frequently occurring events are not necessarily more important (e.g. in the soccer domain, a goal rarely occurs, but is very important).

As yet another example, J. Conroy and D. P. O'Leary "Text summarization via Hidden Markov Models and pivoted QR matrix decomposition", University of Maryland. Department of Computer Science Technical Report, February 2001. CS-TR-4221 discloses the use of Hidden Markov Models (HMMs) for text summarisation in order to determine a probability that a given sentence, i, is in the summary, given that the previous sentence, i-1, is also in the summary. The probability of each sentence being included in the summary is calculated from a given set of features for each observed sentence. Sentence features include the position of the sentence in the document, the number of words in the sentence, the probability of those words occurring etc. A significant disadvantage of the described approach is that the number of HMM states (and hence algorithm complexity) required depends on the size of the summary. Thus, the complexity increases substantially with increasing summary lengths. Furthermore, a new model has to be developed each time the required summary length increases (i.e. for each additional sentence in the summary two extra states have to be added to the model). This results in a very inflexible system which is unsuitable for summaries of varying lengths.

Hence, an improved system for generating a content item summary would be advantageous and in particular a system allowing improved information selection, improved information presentation, improved information personalisation, improved applicability to different content, suitability for variable summary lengths, increased flexibility, facilitated implementation, improved performance and/or an improved user experience would be advantageous.

Summary of the Invention

Accordingly, the Invention seeks to preferably mitigate, alleviate or eliminate one or more of the above mentioned disadvantages singly or in any combination.

According to a first aspect of the invention there is provided an apparatus for generating a content summary for a content item having event information for a plurality of events belonging to event classes; the apparatus comprising: grouping means for grouping event classes of sequential events of the plurality of events into event groups; means for providing training data for a plurality of training content items, the training data comprising content item event class sequences and training content summaries comprising training summary items associated with the event class sequences; rating means for determining a rating for each of the event groups in response to a probability characteristic of the training data, the probability characteristic being indicative of a probability of inclusion of training

summary items related to event class sequences corresponding to the event groups in the training content summaries; selection means for selecting event groups for inclusion in the content summary in response to the rating of the event groups; and means for generating the content summary by including summary items associated with the selected event groups.

The invention may allow an improved generation of a content summary and/or may provide an enhanced user experience. The content summary may be efficiently generated from training data.

The summary generation may be achieved without requiring a specific model, rules, criterions or axioms to be defined. Rather, an automatic generation of a content item summary based on empirical data may for example be achieved.

The invention may e.g. provide an efficient system for generating variable length content summary. A content summary of different lengths may be generated by the same algorithm and may be based on the same training data.

The invention may e.g. provide an improved user experience. In particular, the invention may allow a more accurate content summary and/or may provide for an improved selection of content summary items for the content summary. The invention may additionally or alternatively allow a system which is suitable for a plurality of different content types and in particular to different content domains. The invention may in some

embodiments also provide improved context information for events for which summary items are included in the content summary.

The event information may in particular comprise event data for a variety of events of the content item. The event data may for example be structured according to one or more ontologies. For example the event information may comprise a number of instances belonging to of different event classes in an ontology. An event class sequence of the training content items may comprise event classes for one or more event instances.

The training data may include the training content items themselves or may comprise only the data, such as meta data, related to the content items and required for generating the content summary.

The training content summaries may for example have been generated manually or semi-automatically by selecting event class sequences of the training content items for which content summary items are included. For example, the training content summaries may have been generated manually for the training content items and the invention may allow an automatic generation of a content summary which resembles the characteristics of the manual generation. This may further be achieved without necessitating that criteria or principles of the manual generation are defined or even considered.

The training summary items may be the event classes of the event instances.

According to an optional feature of the invention, the probability characteristic comprises a probability indication for each of a plurality of event class sequences, the probability indication for a given event class sequence being indicative of a number of times training summary items for the given event class sequence is included in the training content summaries relative to a number of times the given event class sequence occurs in the training data.

This may e.g. allow an efficient system, improved content summaries and/or facilitated implementation. In particular, a low complexity algorithm may allow content summary generation which corresponds to the characteristics for the training data and in particular the content summary may reflect the underlying principles and criteria used when generating the training content summaries.

According to an optional feature of the invention, the rating means is arranged to divide at least a first event group into subsets of events and to determine a probability indication for the first event group in response to probability indications for each of the subsets. This may facilitate implementation and reduce the complexity of the apparatus. In particular, it may facilitate the rating of groups as training data probability information is required only for shorter event class sequences rather than for the whole event group.

According to an optional feature of the invention, the rating means is arranged to determine the rating for the

first event group by combining the probability indications for the subsets. This allows a low complexity, efficient and/or accurate rating.

According to an optional feature of the invention, the combining is by a multiplication of probability indications for the subsets. For example, the combining may be by a multiplication of probability indications for the events and event pairs in the event groups thus estimating the joint probability of occurrence of the event group, and the joint probability of inclusion of the event group. This may allow a particularly efficient and accurate rating.

According to an optional feature of the invention, each subset comprises a maximum of two events. This may provide for facilitated implementation and/or improved accuracy. In particular, it may allow the size of the subsets to be minimised, thereby reducing complexity, while allowing sequential information to be taken into account for the rating of event groups.

According to an optional feature of the invention, the probability characteristic comprises a conditional probability matrix comprising a probability indication for all event class pairs. This may e.g. allow a low complexity implementation.

The probability indication may for example be a conditional probability indication.

According to an optional feature of the invention, the rating means is arranged to determine ratings by using

Markov chain. This may e.g. allow a low complexity implementation while achieving accurate performance.

According to an optional feature of the invention, the grouping means is operable to group events in event groups in response to an event class sequence probability. This may provide an efficient and low complexity implementation. In particular, accurate grouping of associated sequential events suitable for combined decision on whether to include in the content summary may be enabled or facilitated. The grouping may in particular provide improved context information for the individual events.

The event class sequence probability may in particular be a joint probability of occurrence of the event classes of a group

According to an optional feature of the invention, the grouping means is arranged to include consecutive events in a first event group until the event class sequence probability for the first event group falls below a threshold. This allows an efficient and low complexity implementation with accurate performance.

According to an optional feature of the invention, the event sequence probability is determined in response to an event class sequence probability in the training data. This allows an efficient performance and in particular may allow the underlying principles and characteristics of the training data to be used for grouping of events. It may furthermore allow re-use of the training data

information for grouping of events and for rating of event groups.

The event sequence probability may for example be determined in response to the probability of the group's event classes occurring. In particular, the event sequence probability may be a joint probability which is determined using a Markov chain comprising the multiplication of the marginal probability of the first event in the group occurring, and the conditional probabilities of subsequent events in the group occurring, given that the previous event occurred.

According to an optional feature of the invention, the grouping means is operable to determine the event class sequence probability for a first event sequence in response to event sequence probabilities of subsets of event sequences of the first event sequence.

This may facilitate implementation and reduce the complexity of the apparatus. In particular, it may facilitate grouping of events as event class sequence probability information is required only for shorter event class sequences rather than for the whole event group.

For example, a marginal probability of an event may be estimated by the probability of occurrence of that event class in the training data content items. A conditional probability of an event occurring, given a first event has occurred is estimated by the ratio of the probability of occurrence of the event class pair to the probability of occurrence of the first event class. The joint

probability for a given event group may then be determined from the marginal probability and the conditional probabilities.

According to an optional feature of the invention, the grouping means is arranged to determine the event class sequence probability for the first event sequence by combining the event class sequence probabilities for the subsets of event sequences. This allows a low complexity, efficient and/or accurate rating. The combination may in particular be determined by multiplication of the event class sequence probabilities of the individual subsets. In particular, a marginal probability of the first event in the group may be multiplied by the conditional probability of subsequent events in the group, given the first event.

According to an optional feature of the invention, the subsets of event classes comprise event classes for a maximum of two events. This may provide for facilitated implementation and/or improved accuracy. In particular, it may allow the size of the subsets to be minimised, thereby reducing complexity, while allowing sequential information to be taken into account for the grouping of event groups

According to an optional feature of the invention, the apparatus further comprises an event conditional probability matrix comprising a conditional probability value for all event class pairs. The apparatus may also comprise a vector of marginal probability values for all possible event classes. This may e.g. allow a low complexity implementation.

According to an optional feature of the invention, the the grouping means is arranged to determine the event class sequence probability for a first event group by using a Markov chain. This may e.g. allow a low complexity implementation while achieving accurate performance.

According to an optional feature of the invention, the selection means is further arranged to select event groups in response to a size characteristic of the content summary.

For example, event groups may be selected until the combined size of the associated summary items exceed the size of the content summary. This may allow for a low complexity method of generating variable length content summaries wherein the most appropriate summary items for a given summary length are automatically selected.

According to an optional feature of the invention, the content item comprises event information structured in accordance with a first ontology and the training data comprise event information structured in accordance with a second ontology and the apparatus comprises means for associating instances of the first ontology with instances of the second ontology.

This may allow increased flexibility and may in particular allow that training data which is arranged according to a given ontology may be used for the generation of content summary for content items having event information structured in accordance with a

different ontology. The means for associating may for example be arranged to map, translate or link different ontology instances between the first and second ontology.

According to a second aspect of the invention, there is provided a method of generating a content summary for a content item having event information for a plurality of events belonging to event classes; the method comprising: grouping event classes of sequential events of the plurality of events into event groups; providing training data for a plurality of training content items, the training data comprising content item event class sequences and training content summaries comprising training summary items associated with the event class sequences; determining a rating for each of the event groups in response to a probability characteristic of the training data, the probability characteristic being indicative of a probability of inclusion of training summary items related to event class sequences corresponding to the event groups in the training content summaries; selecting event groups for inclusion in the content summary in response to the rating of the event groups; and generating the content summary by including summary items associated with the selected event groups.

These and other aspects, features and advantages of the invention will be apparent from and elucidated with reference to the embodiment(s) described hereinafter.

Brief Description of the Drawings

Embodiments of the invention will be described, by way of example only, with reference to the drawings, in which

FIG. 1 illustrates an apparatus for generating a content summary in accordance with some embodiments of the invention;

FIG. 2 illustrates a method of generating a content summary for a content item in accordance with some embodiments of the invention; and

FIG. 3 illustrates an example of an ontology for a soccer content item.

Detailed Description of Embodiments of the Invention

The following description focuses on embodiments of the invention applicable to generation of a summary for an audiovisual content item, such as a TV programme, a sports event, a movie etc. However, it will be appreciated that the invention is not limited to this application but may be applied to many other content items and content item types including for example text based and-or multimedia content items.

FIG. 1 illustrates an apparatus for generating a content summary in accordance with some embodiments of the invention.

In the example of FIG. 1, a content item is provided from a content item source 101. The content item source 101 may an internal or external content source. In the

specific example, the content item is an audiovisual sequence. The content of the audiovisual sequence comprises a number of events which are identified in event information. The event information may be provided integrally with the content item or may be provided separately. For example, the event information may be provided as embedded metadata in the audio visual sequence or may be provided as a separate file comprising event data.

The event information identifies a number of events in the content item. For example, the content item may be a full length sequence of a soccer match and the event information may identify a number of event instances such as goals, free kicks, penalties, throw ins, bookings etc. The event information may e.g. specify each event by a description, a start time and an end time. Each event has a corresponding event class. For example, an event instance may be a penalty to team A taken by player B. This event instance belongs to the event class of penalties. Similar, a first event instance may be a goal for Team A scored by Player B and a second event instance may be a goal for Team B scored by player C. Both of these event instances may belong to the event class of goals.

The content item source 101 is coupled to a grouping processor 103. The grouping processor 103 groups event classes for sequential events into event groups. Some of the event groups may only comprise event classes for a single event but at least some of the event groups comprise event classes for two or more events.

Specifically, the grouping processor 103 may apply a grouping criterion that results in event classes for associated events being combined into the same group. For example, if the content item is for a soccer match, a penalty event may be followed by a sending off event which may be followed by a goal event. In such cases, all three events may be grouped together thereby providing an event group comprising the whole event class sequence from a penalty incident to the resulting goal and including the player committing the penalty being sent off. Such grouping may provide improved context information for the individual events and may result in an improved content summary which more accurately provides the desired information to a user.

It will be appreciated that in some embodiments the grouping processor receives only the event information but not necessarily the content item itself.

The grouping processor 103 is coupled to a rating processor 105. Both the grouping processor 103 and the rating processor 105 are coupled to a training data storage 107.

The training data storage 107 comprises data related to a plurality of training content items which are used by the system to generate the content summary for the content item. The training data comprises data for training content items which, similarly to the content item to be summarised, comprises a number of events. For each training content item, the training data may specifically comprise event information that identifies event class

sequences in the content item. Each event class sequence may comprise event classes for one or more events.

The training content items may for example be full length sequences such as full length video sequences comprising entire soccer matches.

The training data furthermore comprises training content summaries and in particular each training content item has an associated training content summary. The training content summaries are made up of training summary items wherein each of the training summary items is associated with one or more of the event class sequences.

Typically, the training data comprises data for a large number of training content items for which summaries have been created manually by an operator defining events in the content item and subsequently selecting the most important event class sequences. Information for these event class sequences may then be included in the content summary by the operator including one or more summary items for the event class sequence in the content summary.

For example, the training data may comprise data for a large number of training content items corresponding to football matches. For each of these, events have manually been defined and a content summary has manually been created. For example, an operator may have identified all free-kicks, goals, penalties etc and may have selected some of these events and included a description in the content summary. For example, the operator may have selected all event class sequences which include a goal

event and may have included information related to these event class sequences in the training content summary. The operator may then have selected all free-kicks which result in a booking or a sending off and may have included information related to these event class sequences in the training content summary.

It should be noted that training data may be generated specifically for the purpose of creating training data. However, typically the manual creation of content summaries is performed for other purposes and in particular manual creation of content summaries may be widely used to provide summary information to users allowing them to identify and select between content items. Thus, manually generated content summaries may not only be used for information to a user but may also be used as learning/training data by the apparatus of FIG. 1.

It will be appreciated that in some embodiments, the training data storage 107 may comprise only the data needed for generating the content summary. However, in other embodiments the training data storage 107 may comprise additional information such as for example copies of some or all of the training content items themselves. Thus, in some embodiments the full length video sequences themselves may be stored in the training data storage 107 whereas only the associated meta data may be stored in the training data storage 107 in other embodiments.

The training data storage 107 may in some embodiments be used by the grouping processor 103 when grouping the

events. In addition, the training data is used by the rating processor 105 to rate the event groups generated by the grouping processor 103. In particular, the grouping processor 103 may determine a rating for each event group which is indicative of the desirability of including information related to the event group in the content summary. Thus, the rating may be indicative of the importance and/or the interest to a user for the given event group. For example, for a soccer game the rating of an event class sequence comprising a goal will typically be higher than one relating only to a throw in.

The rating processor 105 determines the rating for an event group by evaluating a probability characteristic of the training data. In particular, the rating processor evaluates a probability of the inclusion of training summary items which are related to the event class sequences of the event group in the training content summaries. Thus, if an event group comprises a given event class sequence, all sequences containing the same event classes occurring in the training data are identified. It is then evaluated how many of these were included in the corresponding content summaries.

For example, if an event group comprises the sequence penalty, sending off, goal, all instances of such an event class sequence in the training data are identified. The probability at which each of the event classes present in the event class sequence is referred to in the content summaries is then determined. The probability of inclusion of each event class may be considered as the conditional probability that information for the specific

event class is included in the summary given that it has occurred in the event class sequence.

n particular, the ratio of the probability of inclusion of an event class pair to the probability of inclusion of a first event class may be considered as the conditional probability that information for the second event class is included in the summary, given that the first event class has been included in the summary.. As described in more detail later, the joint probability of the whole group being included in the summary may then be calculated using a Markov chain to combine the marginal probability of the first event in the group with the conditional probabilities of the subsequent events in the group.

In the specific example, it is likely that many such sequences are included in the corresponding summaries and accordingly a high probability indication (or probability) is achieved.

However, if the event class sequence merely comprises a throw in event class, the rating processor 105 will typically find that it is relatively rare that information relating to throw in event classes are included in the summary. It will thus determine a much lower probability value.

The rating processor 105 determines the rating for each event group in response to the probability indication for the event group and may specifically determine an increasing rating for an increasing probability value. In

some embodiments, the probability value may be used directly as a rating.

The rating processor 105 is coupled to a selection processor 109 which selects event groups for inclusion in the content summary in response to the rating of the event groups. For example, the selection processor 109 may simply select all event groups which have a rating above a given threshold thereby ensuring that all events considered sufficiently important to a user are included in the content summary. This may provide for a variable length summary which is automatically adapted to characteristics of the content item. For example, the content summary for a goalless soccer match may be smaller than for a high scoring soccer match.

As another example, the selection processor 109 may select event groups in order of decreasing rating until the content summary reaches a desired size. This may allow relatively fixed size content summaries while ensuring that the most important events are included.

The selection processor 109 is coupled to a summary generator 111 which generates the content summary by including summary items associated with the selected event groups. For example, for each event group a predefined metadata text for each event may be included. For example, in the previously mentioned example, the summary generator 111 may generate the summary by including the text

"penalty"+"then"+"sending off"+ "then"+"goal"

in the content summary wherein the text "then" is automatically inserted between predetermined text strings defined for the individual events.

In other embodiments, the content summary may for example be an audiovisual sequence in itself and the summary generator 111 may for example generate an audiovisual sequence by merging clips of the content item where the clips are those corresponding to the individual event classes of the selected event groups.

The apparatus of FIG. 1 may thus allow an efficient and low complexity generation of a content summary for an unknown content item. The content summary is created without requiring any specific rules or criterion for selection of items for the summary to be defined. Rather, the apparatus automatically uses information of how summaries have been created for other content items to generate the summary. Thus, the explicit and implicit, conscious and unconscious rules, criteria and principles used when creating content summaries for the training content item will bias the content summary generation by the apparatus resulting in content summaries that may be identical or similar to the content summary that would have been achieved by a manual content summary generation. This is achieved without requiring that the criteria and principles used by the manual summary generation are explicitly derived or even considered.

Thus the apparatus may allow improved content summary generation. Furthermore, as the content summary generation does not rely on explicit rules but rather on the training content data, it may be suitable for many

content domains. For example, the training content items may comprise content items relating to many different content domains. When a new summary is generated, the apparatus automatically searches for event class sequences in the training data from the same application domain as specified by their metadata or event classes. If none can be found, other event class sequences in the training data may be used and mappings between the two application domain ontologies can be used. Also, the apparatus is particularly suitable for variable length summary generation as the approach does not require a new model to be built for different summary lengths. Rather a different summary length may be achieved merely by selecting a different number of event groups.

FIG. 2 illustrates a method of generating a content summary for a content item in accordance with some embodiments of the invention. In the following, a more detailed example of embodiments of the invention will be described with reference to the method of FIG. 2. The method is applicable to the apparatus of FIG. 1 and will be described with reference to this.

The method initiates in step 201 wherein event information for a content item comprising a full length sequence of a soccer match is received. The content item comprises event information wherein a number of events are represented by instances of an ontology.

In the example, semantic events are represented as instances of a domain ontology which may be encoded in a standard format (for example RDF (Resource Definition Framework) or OWL (web ontology language)). FIG. 3

illustrates an example of an ontology for the soccer domain. The ontology consists of classes (or event types) and properties or features providing metadata about those events. For example the Goal class has features by, duration, start_time, extra_time, from, taken and resulting_in.

In the example, the training data provided in the training data storage 107 comprises data for a number of full length content sequences, each consisting of a number of instances of an event ontology, along with summaries which consist of a subset of the events in each full length sequence. The semantic events may have been input directly as instances of classes in a standard ontology representation, for example using OWL, or may have been extracted from audio, video or text, using well-known techniques. The ontologies used for the training content item may be the same as the ontology used for the content item to be summarised or may be a different ontology. In the latter case, the apparatus may comprise means for linking different ontology classes of one ontology to corresponding ontology classes of the other ontology.

The training data may particularly comprise data for a large number of training data content items comprising full length soccer sequences but may also comprise data for content items relating to very different content items and domains.

Step 201 is followed by step 203 wherein the event instances are grouped into event groups by the grouping processor 103.

In the embodiment, the grouping processor 103 groups the event groups in response to event class sequence probability which specifically may be the joint probability of the event group. The joint probability of the event group is indicative of the probability of the events being associated with each other. The joint probability of the event group is determined in response to the training content items and in particular is determined by evaluating the probability of which consecutive event classes corresponding to the event classes in the event group are included in the training content summaries when they occur in the training content items. Sequential events may in particular be included until the joint probability falls below a given threshold.

In more detail, the events are grouped into causally related event groups using a Markov chain. This allows the system to provide context to the summary so that it does not consist solely of disjoint, unrelated events, but makes sense as a whole, and explains to the user, for example, what caused a player to be sent off the pitch, or how a goal came about.

This step uses the assumption that events commonly occurring in sequence in the training set are causally related. In order to reduce the complexity the event groups are divided into subsets comprising event pairs. In particular, in order to estimate the joint probability of an event group, a Markov chain is used, so that only the marginal probability of an event occurring and the conditional probability of a particular event occurring,

given that another event has occurred need be estimated, using frequencies of single events and event pairs. This means that only the number of event classes and pairs of event classes need to be counted, rather than the frequency of larger groups of event classes, which reduces complexity.

A conditional probability matrix, **X**, may be formed using the training data event class sequences and their corresponding summaries in the training data. The matrix **X** represents the probability of an event occurring in a training content item given that another event has just occurred. In the specific example, the matrix may be of size N^2, where N is the number of event classes (that is, each event E can take on any one of N symbols). E_t denotes the current event at time t, and E_{t-1} denotes the previous event at time t-1.

The elements of matrix **X**, $\mathbf{X_{ij}} = P(E_t=i \mid E_{t-1}=j)$, $\{i,j\} \in N$ are calculated as follows:

$$P(E_t = i \mid E_{t-1} = j) = \frac{P(E_t = i, E_{t-1} = j)}{P(E_{t-1} = j)}$$

where:

$$P(E_t = i, E_{t-1} = j) = \frac{frequency(event \quad pairs \quad (E_{t-1} = j, E_t = i))}{frequency(all \quad event \quad pairs \quad in \quad training \quad set)}$$

and:

$$P(E_{t-1} = j) = \frac{frequency \quad (event \quad E_{t-1} = j)}{frequency \quad (all \quad events \quad in \quad training \quad set)}$$

Thus, each element in the matrix represents the probability of an instance of the same class as the second event in the event pair occurring, given that an instance of the same class as the first event in the event pair occurred.

In order to group the test sequence events into event groups, it is assumed that the Markovian property holds within an event group. That is, all knowledge about the past is assumed to be contained in the previous state, namely:

$$P(E_t \mid E_{t-1}, E_{t-2} \ldots E_2, E_1) = P(E_t \mid E_{t-1})$$

A Markov chain is constructed beginning with the first event in the test sequence E_1, with subsequent events $E_2 \ldots E_t$, to calculate the joint probability of a number of events $E_t \ldots E_2, E_1$:

$$P(E_t, E_{t-1}, \ldots, E_1) = P(E_t \mid E_{t-1}) * P(E_{t-1} \mid E_{t-2}) \ldots P(E_2 \mid E_1) * P(E_1)$$

When the joint probability falls below a certain threshold (a suitable value may be 0.01) the events are considered to be a complete event group. In the example, there is no normalisation for long chains of events (e.g. by taking the t^{th} root) as it is frequently beneficial to

bias against very long sequences of events as these are less likely to be causally related.

It will be appreciated that any suitable method or criterion for grouping events may be used. For example, the grouping of sequential events may be performed manually or semi-manually.

Step 203 is followed by step 205 wherein the event groups are rated by the rating processor 105.

Specifically, the rating processor 105 evaluates the conditional probability of each event group being included in the summary, given that it has occurred in the training content item sequence. Typically, the more frequently information for a given event group is included in the summary, relative to how rare that grouping is, the higher the importance of the event class sequence and thus the higher the rating. Thus, the probability value may directly be used as a rating value.

In more detail, the rating processor 105 receives the event groups as grouped ontology instances from the grouping processor 103. In other embodiments, the grouping of events may for example be indicated by additional metadata (This additional input metadata could, for example, indicate the groupings chosen manually by a sports editor).

Using these event groups, a second conditional probability matrix, $\mathbf{Y}$, is calculated from the training set. Matrix element $\mathbf{Y}_{ij}$ represents the conditional probability of an event $E_t = i$ being included in the

summary, given that the previous event $E_{t-1}=j$ has been included and occurred immediately prior to E_t in the full length sequence. Specifically:

$$P(E_t^I \mid E_{t-1}^I) = \frac{P(E_t^I, E_t^I)}{P(E_1^I)}$$

$$= \frac{frequency(event\ pair\ (E_t, E_{t-1})\ in\ both\ summaries\ \&\ problem\ descriptions)}{frequency(event\ E_{t-1}\ in\ summaries)}$$

where E_t^I denotes the event at time t being included in a summary. Thus, whereas conditional probability matrix **X** indicated the probability of consecutive events being causally related, conditional probability matrix **Y** indicates the probability of such consecutive events being included in the training content summaries.

Then, each event group $G = \{E_t, E_{t-1}, \ldots E_2, E_1\}$ is assigned a rating, based on its probability of being included, given that it has occurred:

$$\Pr iority = P(G^I \mid G^o) = \frac{P(G^I, G^O)}{P(G^O)}$$

The denominator of this equation is calculated as a Markov chain using probabilities of events occurring in a training content item. The numerator is also computed as a Markov chain:

$$P(G^I, G^O) = P(E_t^{I,O} \mid E_{t-1}^{I,O}) * P(E_{t-1}^{I,O} \mid E_{t-2}^{I,O}) \ldots P(E_2^{I,O} \mid E_1^{I,O}) * P(E_1^{I,O})$$

where $E_t^{I,O}$ denotes an event at time t that is included in the summary and which occurred in the training content

item. Obviously, a single event that is included in the summary must have occurred in the training content item, however two events may be included as a pair in the summary, but not have occurred as a pair in the original full length sequence, due to editing points. In order to model these editing points, $P(G^I, G^O)$ is replaced by $P(G^I)$, which can be calculated by:

$$P(G^I) = P(E_t^I \mid E_{t-1}^I) * P(E_{t-1}^I \mid E_{t-2}^I) ... P(E_2^I \mid E_1^I) * P(E_1^I)$$

Therefore, the final rating for each event group G is calculated as:

$$Rating(G) = \frac{P(E_t^I \mid E_{t-1}^I) * P(E_{t-1}^I \mid E_{t-2}^I) ... P(E_2^I \mid E_1^I) * P(E_1^I)}{P(E_t^O \mid E_{t-1}^O) * P(E_{t-1}^O \mid E_{t-2}^O) ... P(E_2^O \mid E_1^O) * P(E_1^O)}$$

These calculations make the assumption that the first order Markovian property holds, not only between events in the content item, but also over events that are jointly included in the summary. Using this assumption, the problem of not having enough data to calculate probabilities of combinations of events that occur rarely or not at all (the so-called ''one-shot learning'' problem) may be alleviated substantially. For example, if the probabilities of all 5-event event groups are to be calculated a minimum of 20^5 = 3200 000 events would be required to get one occurrence of each possible combination of 5 events. However, as the conditional probability matrices are based only on frequencies of event pairs and single events, a much reduced training data volume may be acceptable.

Entries in the conditional probability matrix which are zero, because no examples of that particular event pair occur in the training data, may be given a small probability value in order to avoid that an event group is given a zero probability of occurrence which would result in an infinite rating.

Step 205 is followed by step 207 wherein event groups are selected in accordance with their rating. Specifically, the selection processor 109 may select a given number, N, of event groups and these may be selected as the N event groups which have the highest ratings.

As another example, the selection processor 109 may select event groups in response to a size characteristic of the content summary. For example, a fixed or maximum length summary may be required and the selection processor 109 may select event groups in order of decreasing rating until this length has been reached. As yet another example, the selection processor 109 may select all event groups having a rating above a given value resulting in summary of variable length depending on the importance/interest of the event groups.

Step 207 is followed by step 209 wherein a content summary is created by including items in the summary for each of the selected event groups. For example, a semantic description may be included for each of the selected event groups. This may for example be in the form of metadata (such as MPEG-7 metadata).

As another example, if the event descriptions contain time codes and duration information corresponding to the audio visual signal, the summary may be provided in the form of a compilation of the audio visual video clips related to each of the selected event groups. Thus, a summary in the form of a highlights video sequence may be generated.

Simulations and experiments have shown that the described approach may result in content summaries which are highly accurate and closely resemble those that would typically be obtained by a manual summary generation.

Furthermore, the method is applicable to information from any medium (not just text, or video alone). The importance of different pieces of information to the user is determined in a probabilistic manner, which avoids the need for manually specifying reasoning axioms. The method furthermore allows context to be incorporated in the summary, since groups of events that are causally related will all be included in the summary. Also, the approach avoids the requirement for separate models for different length summaries.

It will be appreciated that the above description for clarity has described embodiments of the invention with reference to different functional units and processors. However, it will be apparent that any suitable distribution of functionality between different functional units or processors may be used without detracting from the invention. For example, functionality illustrated to be performed by separate processors or controllers may be performed by the same

processor or controllers. Hence, references to specific functional units are only to be seen as references to suitable means for providing the described functionality rather than indicative of a strict logical or physical structure or organization.

The invention can be implemented in any suitable form including hardware, software, firmware or any combination of these. The invention may optionally be implemented at least partly as computer software running on one or more data processors and/or digital signal processors. The elements and components of an embodiment of the invention may be physically, functionally and logically implemented in any suitable way. Indeed the functionality may be implemented in a single unit, in a plurality of units or as part of other functional units. As such, the invention may be implemented in a single unit or may be physically and functionally distributed between different units and processors.

Although the present invention has been described in connection with some embodiments, it is not intended to be limited to the specific form set forth herein. Rather, the scope of the present invention is limited only by the accompanying claims. Additionally, although a feature may appear to be described in connection with particular embodiments, one skilled in the art would recognize that various features of the described embodiments may be combined in accordance with the invention. In the claims, the term comprising does not exclude the presence of other elements or steps.

Furthermore, although individually listed, a plurality of means, elements or method steps may be implemented by e.g. a single unit or processor. Additionally, although individual features may be included in different claims, these may possibly be advantageously combined, and the inclusion in different claims does not imply that a combination of features is not feasible and/or advantageous. Also the inclusion of a feature in one category of claims does not imply a limitation to this category but rather indicates that the feature is equally applicable to other claim categories as appropriate. Furthermore, the order of features in the claims do not imply any specific order in which the features must be worked and in particular the order of individual steps in a method claim does not imply that the steps must be performed in this order. Rather, the steps may be performed in any suitable order. In addition, singular references do not exclude a plurality. Thus references to "a", "an", "first", "second" etc do not preclude a plurality.

CLAIMS

1. An apparatus for generating a content summary for a content item having event information for a plurality of events belonging to event classes; the apparatus comprising:

grouping means for grouping event classes of sequential events of the plurality of events into event groups;

means for providing training data for a plurality of training content items, the training data comprising content item event class sequences and training content summaries comprising training summary items associated with the event class sequences;

rating means for determining a rating for each of the event groups in response to a probability characteristic of the training data, the probability characteristic being indicative of a probability of inclusion of training summary items related to event class sequences corresponding to the event groups in the training content summaries;

selection means for selecting event groups for inclusion in the content summary in response to the rating of the event groups; and

means for generating the content summary by including summary items associated with the selected event groups.

2. The apparatus claimed in claim 1 wherein the probability characteristic comprises a probability indication for each of a plurality of event class

sequences, the probability indication for a given event class sequence being indicative of a number of times training summary items for the given event class sequence is included in the training content summaries relative to a number of times the given event class sequence occurs in the training data.

3. The apparatus claimed in any previous claim wherein the rating means is arranged to divide at least a first event group into subsets of events and to determine a probability indication for the first event group in response to probability indications for each of the subsets.

4. The apparatus claimed in claim wherein 3 the rating means is arranged to determine the rating for the first event group by combining the probability indications for the subsets.

5. The apparatus claimed in claim 4 wherein the combining is by a multiplication of probability indications for the subsets.

6. The apparatus claimed in any of the claims 3 to 5 wherein each subset comprises a maximum of two events.

7. The apparatus claimed in claim 6 wherein the probability characteristic comprises a conditional probability matrix comprising a probability indication for all event class pairs.

8. The apparatus claimed in any previous claim wherein the rating means is arranged to determine ratings by using a Markov chain.

9. The apparatus claimed in any previous claim wherein the grouping means is operable to group events in event groups in response to an event class sequence probability.

10. The apparatus claimed in claim 9 wherein the grouping means is arranged to include consecutive events in a first event group until the event class sequence probability for the first event group falls below a threshold.

11. The apparatus claimed in claim 9 or 10 wherein the event sequence probability is determined in response to an event class sequence probability in the training data.

12 The apparatus claimed in any of the claims 9 to 11 wherein the grouping means is operable to determine the event class sequence probability for a first event sequence in response to event sequence probabilities of subsets of event sequences of the first event sequence.

13. The apparatus claimed in claim 12 wherein the grouping means is arranged to determine the event class sequence probability for the first event sequence by combining the event class sequence probabilities for the subsets of event sequences.

14. The apparatus claimed in claim 12 or 13 wherein the subsets of event classes comprise event classes for a maximum of two events.

15. The apparatus claimed in claim 14 wherein the apparatus further comprises an event conditional probability matrix comprising a conditional probability value for all event class pairs.

16. The apparatus claimed in any previous claim 9 to 15 wherein the grouping means is arranged to determine the event class sequence probability for a first event group by using a Markov chain.

17. The apparatus claimed in any previous claim wherein the selection means is further arranged to select event groups in response to a size characteristic of the content summary.

18. The apparatus claimed in any previous claim wherein the content item comprises event information structured in accordance with a first ontology and the training data comprise event information structured in accordance with a second ontology and wherein the apparatus comprises means for associating instances of the first ontology with instances of the second ontology.

19. A method of generating a content summary for a content item having event information for a plurality of events belonging to event classes; the method comprising:

grouping event classes of sequential events of the plurality of events into event groups;

providing training data for a plurality of training content items, the training data comprising content item event class sequences and training content summaries

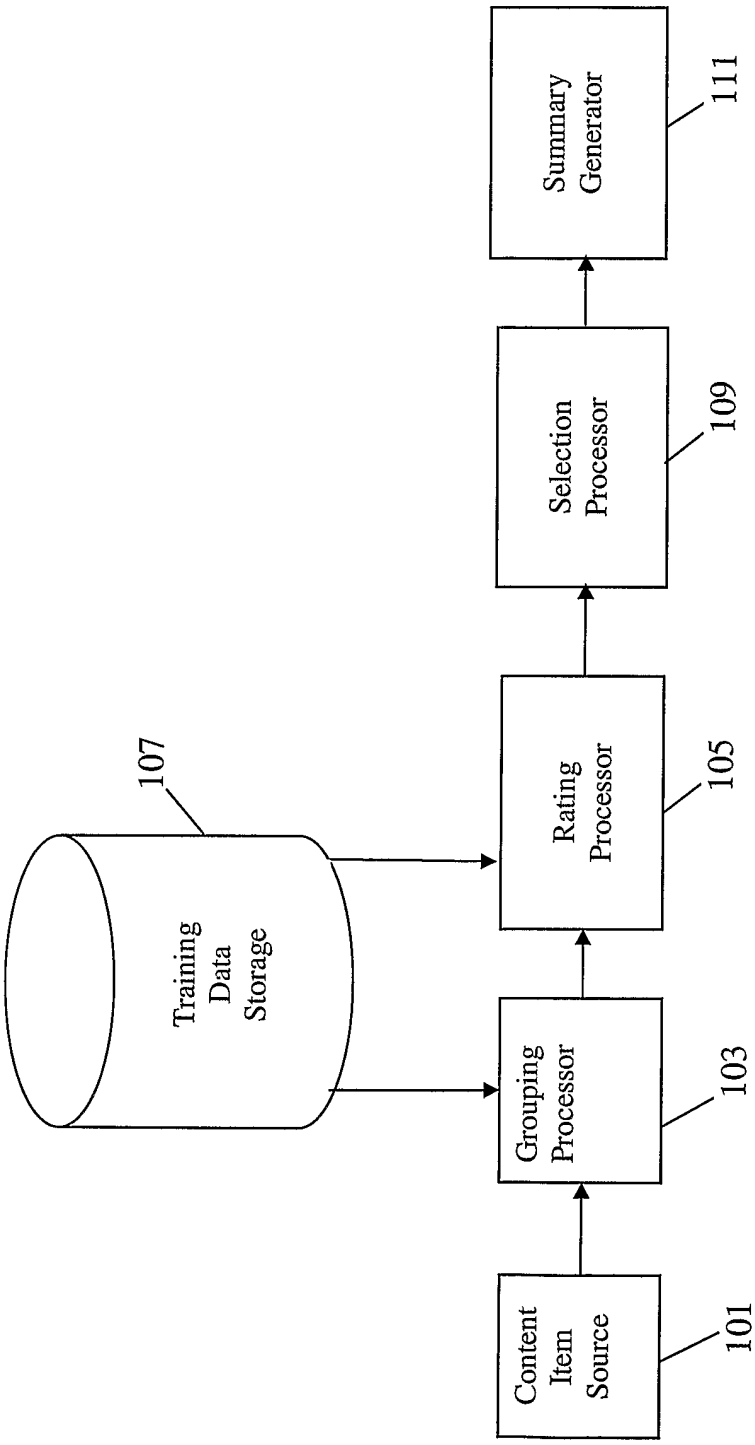
comprising training summary items associated with the event class sequences;

determining a rating for each of the event groups in response to a probability characteristic of the training data, the probability characteristic being indicative of a probability of inclusion of training summary items related to event class sequences corresponding to the event groups in the training content summaries;

selecting event groups for inclusion in the content summary in response to the rating of the event groups; and

generating the content summary by including summary items associated with the selected event groups.

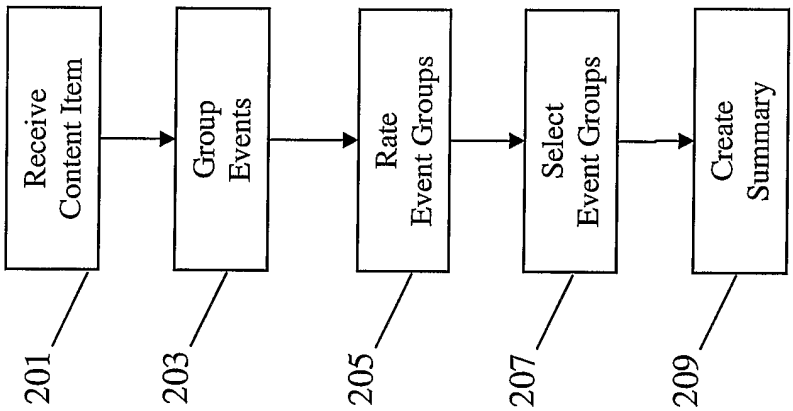
20. A computer program enabling the carrying out of a method according to claim 19.


Content Item Source
101
Grouping Processor
103
Rating Processor
105
Selection Processor
109
Summary Generator
111
Training Data Storage
107

100

FIG. 1

FIG.2


Receive Content Item
Group Events
Rate Event Groups
Select Event Groups
Create Summary
201
203
205
207
209

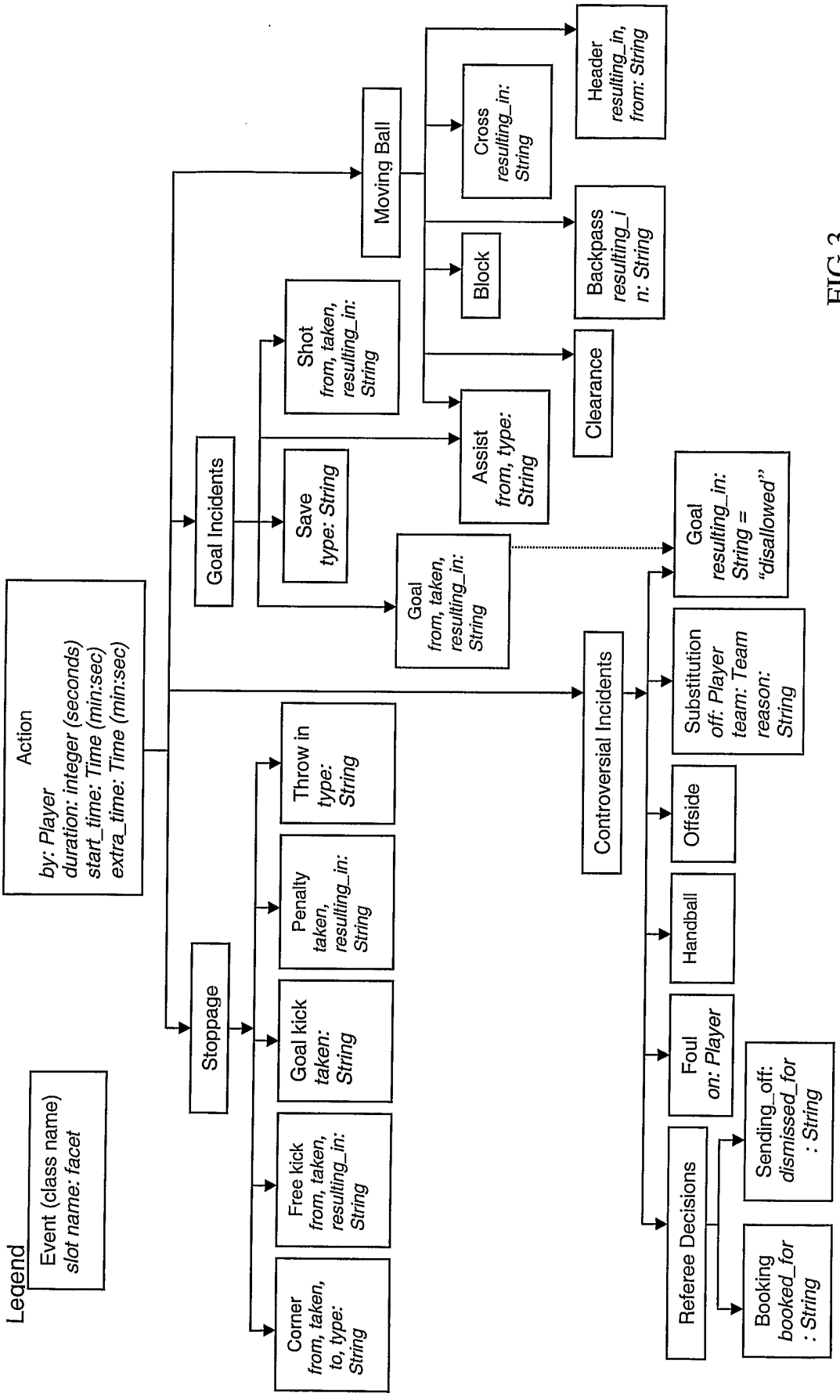

Legend
Event (class name)
slot name: facet
Action
by: Player
duration: integer (seconds)
start_time: Time (min:sec)
extra_time: Time (min:sec)
Stoppage
Corner
from, taken, to, type: String
Free kick
from, taken, resulting_in: String
Goal kick
taken: String
Penalty
taken, resulting_in: String
Throw in
type: String
Goal Incidents
Goal
from, taken, resulting_in: String
Save
type: String
Shot
from, taken, resulting_in: String
Moving Ball
Assist
from, type: String
Block
Clearance
Backpass
resulting_in: String
Cross
resulting_in: String
Header
resulting_in, from: String
Controversial Incidents
Goal
resulting_in: String = "disallowed"
Substitution
off: Player
team: Team
reason: String
Offside
Handball
Foul
on: Player
Referee Decisions
Sending_off:
dismissed_for
: String
Booking
booked_for
: String

FIG.3

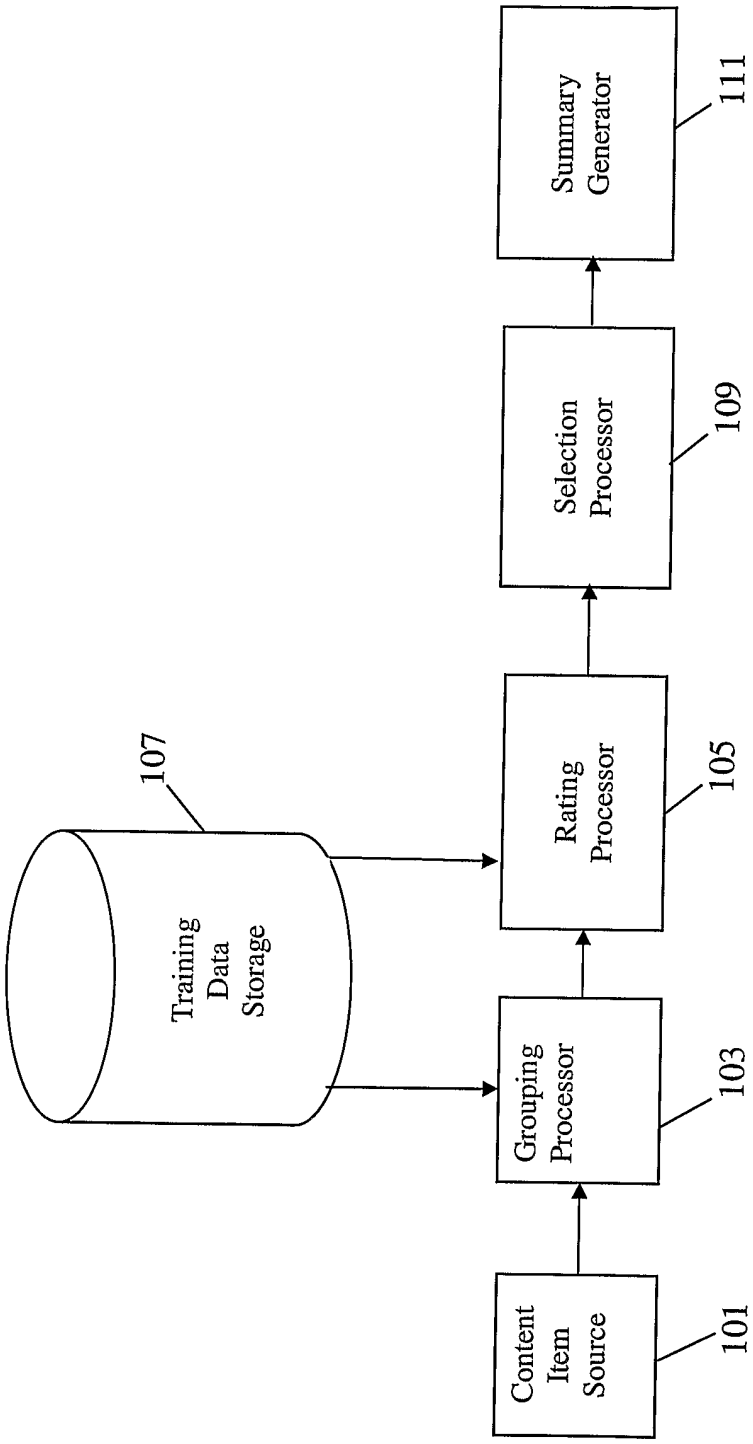
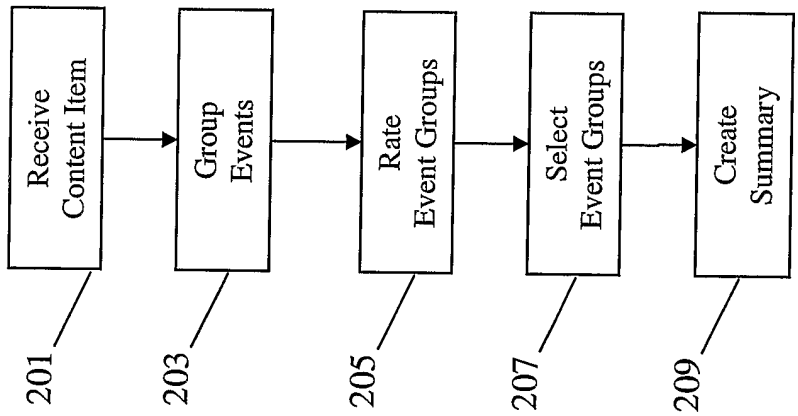

Content Item Source
101
Grouping Processor
103
Rating Processor
105
Selection Processor
109
Summary Generator
111
Training Data Storage
107
100

FIG. 1

FIG.2


Receive Content Item
Group Events
Rate Event Groups
Select Event Groups
Create Summary
201
203
205
207
209

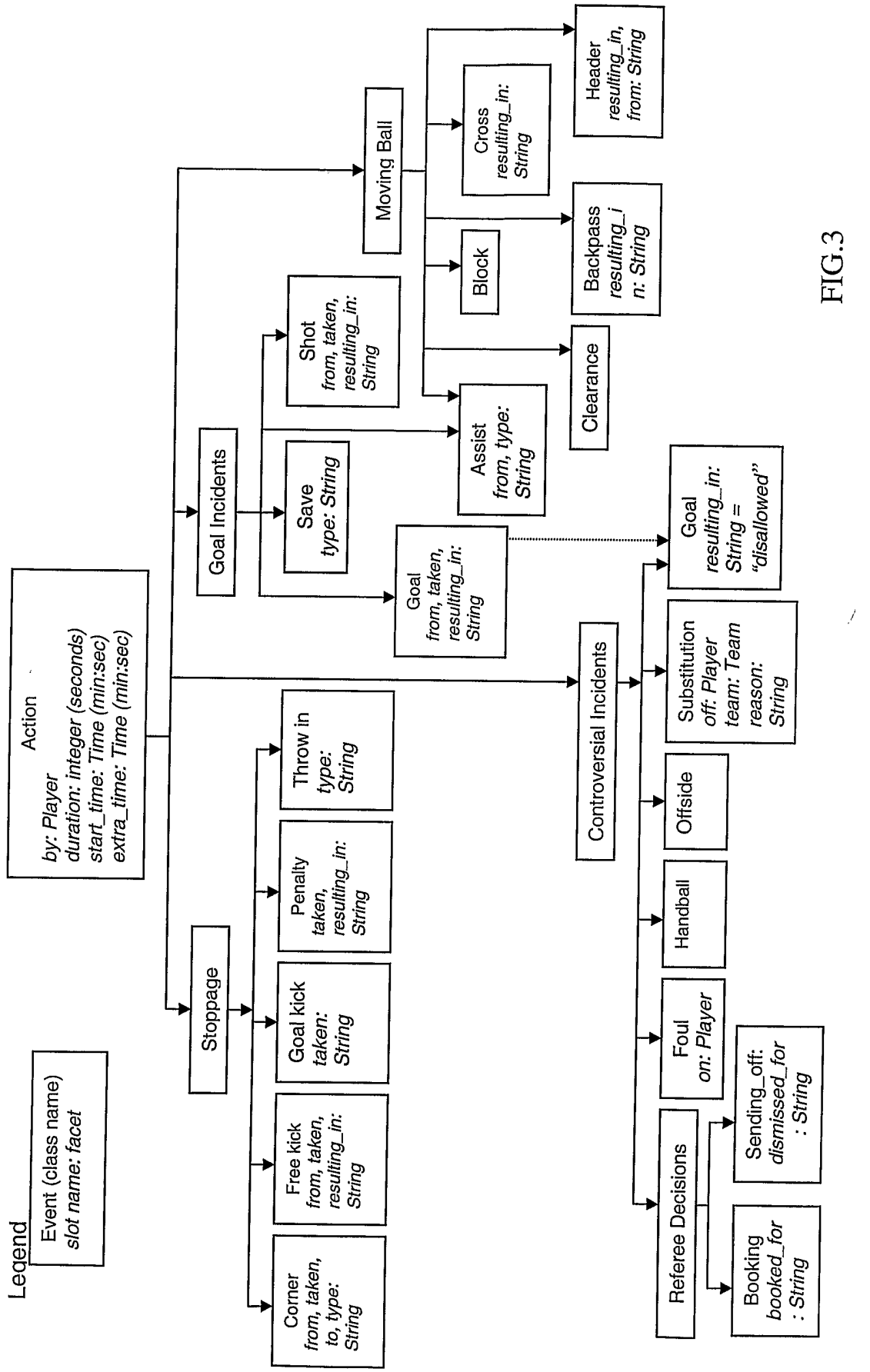

Legend
Event (class name)
slot name: facet
Action
by: Player
duration: integer (seconds)
start_time: Time (min:sec)
extra_time: Time (min:sec)
Stoppage
Corner
from, taken, to, type: String
Free kick
from, taken, resulting_in: String
Goal kick
taken: String
Penalty
taken, resulting_in: String
Throw in
type: String
Goal Incidents
Goal
from, taken, resulting_in: String
Save
type: String
Shot
from, taken, resulting_in: String
Moving Ball
Assist
from, type: String
Block
Cross
resulting_in: String
Backpass
resulting_in: String
Header
resulting_in, from: String
Clearance
Controversial Incidents
Goal
resulting_in: String = "disallowed"
Substitution
off: Player
team: Team
reason: String
Offside
Handball
Foul
on: Player
Referee Decisions
Booking
booked_for: String
Sending_off:
dismissed_for: String

FIG.3