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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 November 2006 (23.11.2006)

PCT

(10) International Publication Number
WO 2006/124077 A2

(51) International Patent Classification:
H04L 12/28 (2006.01)

(21) International Application Number:
PCT/US2006/000741

(22) International Filing Date: 6 January 2006 (06.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/131,608 18 May 2005 (18.05.2005) US

(71) Applicant *(for all designated States except US)*: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; Thomas J. Watson Research Center, P.O. Box 218, Yorktown, New York 10598 (US).

(71) Applicants and
(72) Inventors: **NAPHADE, Milind R.** [IN/US]; 911 Jefferson Blvd, Fishkill, New York 12524 (US). **YAN, Rong** [CN/US]; 5620 5th Ave, Apt 17A, Pittsburgh, Pennsylvania 15232 (US).

(74) Agents: **PETERSON, Charles, W., Jr.** et al.; LAW OFFICE OF CHARLES W PETERSON, JR, Suite 100, 11703 Bowman Green Dr., Reston, Virginia 20190 (US).

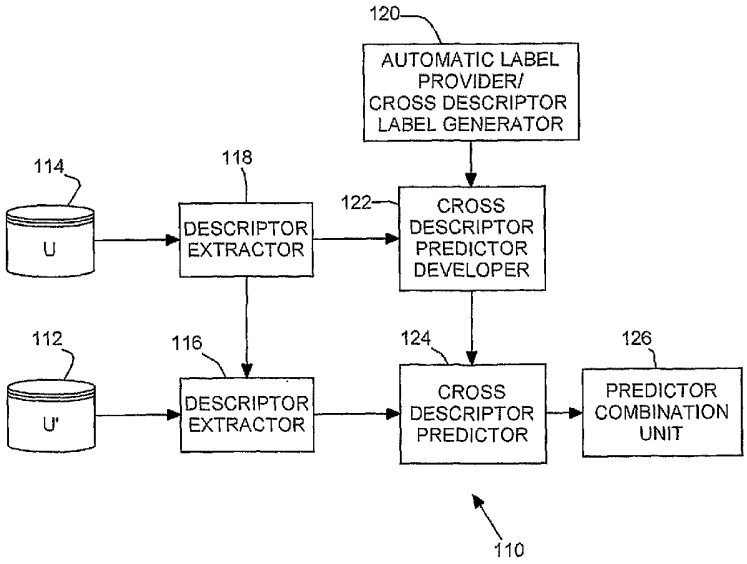
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, 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*

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.

(54) **Title:** CROSS DESCRIPTOR LEARNING SYSTEM, METHOD AND PROGRAM PRODUCT THEREFOR



(57) **Abstract:** A cross descriptor learning system, method and program product therefor. The system extracts descriptors from unlabeled exemplars. For each unlabeled exemplar, a cross predictor uses each descriptor to generate labels for other descriptor. An automatic label generator also generates labels for the same unlabeled exemplars or, optionally, for labeled exemplars. A label predictor results for each descriptor by combining labels from the cross predictor with labels from the automatic label generator.

WO 2006/124077 A2

CROSS DESCRIPTOR LEARNING SYSTEM, METHOD AND PROGRAM PRODUCT THEREFOR

GOVERNMENT RIGHTS

[0001] This invention was made with Government support under Contract No.: 2004*H839800*000 awarded by DST/ATP/Contracts. The Government has certain rights in this invention.

FIELD OF THE INVENTION

[0002] The present invention generally relates to information management and more particularly, to unstructured information management and automatic metadata enrichment.

Background Description

[0003] Increasingly, managing unstructured data content and information has required some type of semantic detection and indexing capability. Consequently, typical state of the art content management systems are increasingly relying on machine learning and classification techniques. These state of the art machine learning and classification techniques rely to varying degrees on human intervention to construct the detector (i.e., to teach the system how to classify) prior to use and, sometimes, during use. Also, machine learning and classification techniques may be classified as supervised, semi-supervised or unsupervised. Supervised machine learning and classification begins, for example, by iteratively classifying known examples or labeled exemplars. Semi-supervised machine learning and classification uses both labeled exemplars and unlabeled exemplars. Unsupervised machine learning and classification uses unlabeled exemplars. However, whether supervised or unsupervised, typical such techniques rely on human intervention or feedback to train the classifier to arrive at an acceptable result.

[0004] Whether supervised or unsupervised, learning and classification techniques may require considerable supervision as the semantic detector is being constructed, but that may not need a learning component during detection. Well known relevance feedback type techniques may be characterized as non-persistent lightweight binary classifiers that use incremental training to improve classification/retrieval performance. Relevance feedback classifier accuracy depends upon the number of exemplars provided, the level of feedback the classifier receives and the amount of time expended training. Statistical semantic modeling, for example, has significantly reduced the level of manual supervision needed over older relevance feedback techniques from lightweight classifiers to heavyweight classifiers. Unfortunately, these prior art training the classifier can be a time consuming and expensive proposition. So, these techniques consume large amounts of precious annotation time and require a considerable annotation effort during training to achieve acceptable annotation quality. As a result, it has become increasingly important to reduce human intervention in machine learning and classification, especially for state of the art media indexing and retrieval.

[0005] Consequently, to reduce human intervention time, disambiguation has been widely applied during annotation. Further, active learning with the system taking a pro-active role in selecting samples during annotation has maximized disambiguation and reduced the number of samples that need to be annotated by an order of magnitude. *See, e.g.*, M. Naphade et al., "Learning to Annotate Video Databases," *Proc. IS&T/SPIE Symp. on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases X*, San Jose, CA, Jan., 2002). An orthogonal approach for concepts with regional support, known as multiple instance learning, accepts annotations at coarser granularity. For example, a user can build a model for a regional concept (e.g., the sky) by selecting the region in an image that corresponds to the regional label. Once the regional concepts have been selected, the system learns from several possible positively and negatively annotated examples, how to represent the concept using regional features. *See, e.g.*, A.L. Ratan, O. Maron, W.E.L. Grimson, and T. Lozano Prez. *A framework for learning query concepts in image classification.* In CVPR, pp. 423--429, 1999.

[0006] Other useful tools include cross descriptor learning with multimodal semantic concept detection. *See, e.g.*, Naphade et al, "Probabilistic Multimedia Objects (Multijects): A Novel approach to Indexing and Retrieval in Multimedia Systems," *Proceedings of IEEE International Conference on Image Processing*, vol. 3, pp 536-540, Oct. 1998, Chicago, IL. For a semi-supervised example, where unlabeled exemplars are used in conjunction with labeled exemplars for classification, *see,* Naphade et al, "Classification using a Set of Labeled and Unlabeled Images," SPIE Photonics East, Internet Multimedia Management Systems, vol. 4210, pp 13-24, Boston, MA, Nov 2000. Also, unlabeled exemplars with multiple descriptors have been used with labeled exemplars in what is known as single view sufficiency. Single view sufficiency is useful when each descriptor is sufficient by itself for learning and to represent the metadata model. *See, e.g.*, Blum et al, "Combining labeled and unlabeled data with co-training," *Proceedings of Conference on Computational Learning Theory*, pp 92-100, 1998. Unfortunately, single view sufficiency requires making simplistic and unrealistic assumptions, i.e., that each descriptor in itself sufficiently represents the metadata and that all descriptors agree with each other in terms of the metadata characteristics. Descriptors for unstructured data (such as for reality based exemplars that support multiple descriptors, e.g., video, text, images and etc.) seldom satisfy single view sufficiency requirements. So, because of the constraints imposed by single view sufficiency, it has not been particularly useful on unstructured data and information. Consequently, none of these approaches all require some manual intervention in enriching metadata in the unlabeled exemplars, even for unlabeled exemplars that can be described using multiple descriptors.

[0007] Thus, there is a need for a system and method that is unconstrained by the restrictions of single view sufficiency and independent of the apparatus used for generating initial labels over the unlabeled exemplars and further, for a system and method for developing cross feature learning on unlabeled exemplars.

SUMMARY OF THE INVENTION

[0008] It is a purpose of the invention to facilitate unstructured data content and information management;

[0009] It is another purpose of the invention to relax single descriptor view sufficiency requirements in content management systems;

[0010] It is yet another purpose of the invention to automatically label unlabeled predictors without manually annotating any of the unstructured data;

[0011] It is yet another purpose of the invention to automatically characterize and label unstructured information;

[0012] It is yet another purpose of the invention to provide cross feature learning on unlabeled exemplars unconstrained by the restrictions of single view sufficiency and independent of the apparatus used for generating both the initial labels and the unlabeled exemplars.

[0013] The present invention relates to a cross descriptor learning system, method and program product therefor. The system extracts descriptors from unlabeled exemplars. For each unlabeled exemplar, a cross predictor uses each descriptor to generate labels for other descriptor. An automatic label generator also generates labels for the same unlabeled exemplars or, optionally, for labeled exemplars. A label predictor results for each descriptor by combining labels from the cross predictor with labels from the automatic label generator.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] The foregoing and other objects, aspects and advantages will be better understood from the following detailed description of a preferred embodiment of the invention with reference to the drawings, in which:

[0015] Figure 1 shows a flow diagram example of cross descriptor learning steps according to a preferred embodiment of the present invention;

[0016] Figure 2 shows an example of a preferred embodiment cross descriptor learning system for characterizing unlabeled exemplars according to the present invention;

[0017] Figure 3 shows an example of a descriptor extraction unit for extracting descriptors from unlabeled exemplars;

[0018] Figure 4 shows an example of a label generation unit with a bank of N label generators;

[0019] Figure 5 shows an example of developing persistent predictors from unlabeled exemplars;

[0020] Figure 6 shows an example of an automatic label generator;

[0021] Figure 7 shows an example wherein a set of labeled exemplars are generated from the unlabeled exemplars using a label generation service;

[0022] Figure 8 shows an example of cross descriptor construction data flow to automatically generate labels in a preferred embodiment cross descriptor prediction unit;

[0023] Figure 9 shows an example of a preferred cross descriptor predictor development unit that receives predictors generated by the cross descriptor predictor.

DESCRIPTION OF PREFERRED EMBODIMENTS

[0024] Turning now to the drawings, and more particularly Figure 1 shows a flow diagram example of cross descriptor learning steps 100 on a set (U) of unlabeled exemplars 102 according to a preferred embodiment of the present invention. Application of the cross descriptor learning steps 100 facilitates automatic metadata generation for unlabeled, unstructured data as well as refining the metadata based on exemplars in the unlabeled, unstructured data and representations of the exemplars by multiple descriptors. Further, cross descriptor learning according to the present invention has application to managing unstructured information encountered in numerous applications including search, federation across multiple search engines, automatic metadata enrichment of unstructured data of different modalities. Such unstructured data may include, for example only, text, audio, image, video, sensors, etc.

[0025] Cross descriptor learning begins in step 104 by creating representations for the unlabeled exemplars 102 by extracting as many as *N* descriptors (X_i) for each unlabeled exemplar, where N is at least 2 and *i* ranges from 1 to *N*. For each unlabeled exemplar each descriptor is used in step 106with other descriptors to automatically generate label predictors for some or all of the other *N-1* descriptors. Finally, in step 108 the predictors developed for the descriptors are combined to generate labels and corresponding confidences for all of the unlabeled exemplars 102. Thus, using labels for descriptors for other exemplars much more robustly refines the resulting predictors.

[0026] Accordingly, unlabeled exemplars are selected from unstructured information 102 and descriptors are determined 104 for the unlabeled exemplars. Each descriptor is used to automatically generate 106 labels for other ones of the unlabeled exemplars. So, by generating a label for each descriptor using some or all of the other descriptors, effectively, the unlabeled exemplars 102 serves as a source of *N* sets of up to *N-1* descriptors. Each label is automatically generated from one of the N sets and a predictor is generated from each label. Labels may be automatically

generated, for example, by a label generation unit using each descriptor for generating labels for some or all of the other descriptors. So for example with $N = 4$, an audio descriptor, image descriptor, and text descriptor may be combined for one unlabeled exemplar to automatically generate a video descriptor predictor; the same audio descriptor and text descriptor may be combined with a video descriptor for the unlabeled exemplar to automatically generate a image descriptor predictor; the audio descriptor and image descriptor may be combined with the video descriptor to automatically generate a text descriptor predictor; and the text descriptor and image descriptor may be combined with the video descriptor to automatically generate a audio descriptor predictor. Thus, the automatically generated label predictor for any descriptor is generated with confidences based on the processing of any other descriptor or combination of descriptors, excluding the descriptor itself. This leverages the information in metadata available in the descriptors that might otherwise insufficiently describe the exemplar. A predictor is developed for each descriptor from these automatically generated labels.

[0027] In another example, the unlabeled exemplars may be provided to a predictor or predictor service, e.g., Yahoo® Search and Google™ Search, to automatically generate the labels. Alternately, a separate set of labeled exemplars may be provided and used to automatically generate labels and predictors for the unlabeled exemplars. Regardless of how the labels are automatically generated, labels for each descriptor are used to develop predictors for other descriptors. Once predictors have been generated for each available descriptor, the predictors are combined for each descriptor 108 to generate a new set of predictors for the unlabeled exemplars.

[0028] Figure 2 shows an example of a preferred embodiment cross descriptor learning system 110 for generating and combining descriptors according to the present invention, e.g., according to the cross descriptor learning steps 100 of Figure 1. In this example distinct unlabeled exemplar sets 112, 114 are provided to the system 110, e.g., cached locally in local storage. The system 110 uses one set 112 to build predictors for the other set 114 and to automatically label resulting descriptors

from the second set 114. Descriptor extraction units 116, 118 extract identical sets of corresponding descriptors for each unlabeled exemplar set 112, 114. A label generation unit 120 automatically generates labels for all of the *N* extracted descriptors for each unlabeled exemplar. The label generation unit 120 uses each descriptor from descriptor extraction unit 118 to automatically generate labels for other ones of the *N* generated descriptors for each unlabeled exemplar. So, the labels generated for the descriptors for one exemplar are generated partially independent of the labels for other descriptors. A cross descriptor predictor development (or generation) unit 122 generates a predictor for each of the descriptors from the labels. A cross descriptor prediction unit 124 assigns a label to each of the descriptors based on the predictor from the cross descriptor predictor development unit 122. A predictor combination unit 126 combines the assigned labels for a more refined result.

[0029] Figure 3 shows an example of a descriptor extraction unit 130, e.g., units 116, 118 in Figure 2, for extracting descriptors from exemplars 132, e.g., from unlabeled exemplar sets 112, 114. The descriptor extraction unit 130 includes a bank of descriptor extractors 134-1, ..., 134-N operating in parallel to extract descriptors 136-1, ..., 136-N from each of the individual unlabeled exemplars 132. Preferably, the generated descriptors 136-1, ..., 136-N include image descriptors, audio descriptors, text descriptors and video characteristic descriptors, e.g., color, texture, shape, frequency and etc.

[0030] Figure 4 shows an example of a label generation unit 120 that automatically generates cross descriptor labels. In this example, the label generation unit 120 includes a bank of *N* label generators 140-1, ..., 140-N, one for each of the *N* descriptors. Each unlabeled exemplars 142 is passed to label generators 140-1, ..., 140-N. Each label generator 140-1, ..., 140-N automatically generates a corresponding label 144-1, ..., 144-N for each descriptor based on descriptor analysis of the *N-1* other descriptors, i.e., the *N-1* descriptors excluding the current descriptor that the particular label generator is labeling. It should be noted that, primarily, each label generator 140-1, ..., 140-N automatically generates labels without any human intervention, at least during generation.

[0031] So, for example, an audio descriptor and a video descriptor may be extracted from an unlabeled exemplar. The label generation unit 120 generates a label for the audio descriptor based on the analysis of the visual descriptor and vice-versa. Each label may include, for example, a character string and a numerical measure. The numerical measure may quantitatively represent the likelihood of the object represented by the character string being present in the visual exemplar. An "Airplane" label, for example, may be associated with detection confidences of 0.1 and 0.9 from the visual descriptor and the audio descriptor, respectively.

[0032] Figure 5 shows an example of developing persistent predictors 150 from unlabeled exemplars 114, substantially as described for Figure 2 with like elements labeled identically. In this example a single unlabeled exemplar set 112 is passed to a single descriptor extraction unit 118 that extracts a set of N descriptors for each unlabeled exemplar in the set 114. Again, the label generation unit 120 automatically generates labels for each of the *N* descriptors from the remaining *N-1* descriptors. The cross descriptor predictor development (or generation) unit 122 uses the *N* descriptors to generate a predictor for each. The resulting *N* predictors are stored 150 as persistent predictors that, subsequently, are used for prediction, e.g., over the current set 114 or other sets of unlabeled exemplars.

[0033] Figure 6 shows an example of an automatic label generator 140, e.g., one of 140-1, ..., 140-N in Figure 4. In this example, the label generator 140 includes a set of L manually labeled exemplars 1402. A descriptor extraction unit 1404, that is substantially identical to descriptor extractors 134-1, ..., 134-N in Figure 3, extracts the appropriate descriptors from all the labeled exemplars 1402. A label predictor development unit 1406 generates a label prediction unit 1408 based on the descriptors provided by the descriptor extraction unit 1404. The label prediction unit 1408 operates on descriptors received from another descriptor extraction unit 1410, e.g., substantially similar to one of descriptor extractors 134-1, ..., 134-N. So, the particular unlabeled exemplar from set 142 is passed to descriptor extraction unit 1410. Descriptor extraction unit 1410 extracts a descriptor and passes the extracted descriptor to label prediction unit 1408, which automatically generates a label 144 for

the descriptor. Preferably, the descriptor extraction unit 1410 extracts the exact same descriptor that was used for developing the label prediction unit 1406. Extracted descriptors are presented to the label prediction unit 1408, which generates labels 144 that may include a text string and a numeric measure of the likelihood of observing the label (represented by the text) in the unlabeled exemplar 142.

[0034] Figure 7 shows an example of an automatic label generator 140', in this example a service that accepts an unlabeled exemplar 142 and automatically generates labels 144. Suitable labeling services include any typical state of the art search engine, e.g., Google™ Search. When the system presents an unlabeled exemplar to a label provider service 140', the service 140' provides a list and ranking of labels for the unlabeled exemplar. Such labels can be supplied directly by the service provider to the cross descriptor predictor development unit 122 in Figure 2.

[0035] Figure 8 shows an example of data flow in automatically generating labels for $N = 2$ using a preferred embodiment cross descriptor prediction unit, e.g., 122 in Figure 2. Pairs of descriptors 136-1, 136-2 are extracted from unlabeled exemplars 142 by a pair of descriptor extractors, e.g., 134-1, ..., 134-N of Figure 3. Simultaneously, a pair of automatic label generators 140-1, 140-2 (e.g., 144-1, ..., 144-N in Figure 4) generate labels from a subset of the unlabeled exemplars 142. The extracted descriptors 136-1, 136-2 and corresponding automatically generated labels are each passed to a predictor developer 122-1, 122-2, e.g., in cross descriptor predictor developer 122 in Figure 2. Each predictor developer 122-1, 122-2 provides a cross-trained label predictor 150-1, 150-2 (e.g., in predictors 150 in Figure 5) for the particular descriptor 136-1, 136-2 with all predictors 150-1, 150-2 being created in parallel (e.g., in a batch submission) or sequentially.

[0036] Figure 9 shows a minimum descriptor size example (again, $N = 2$) of a preferred cross descriptor predictor development data flow on a pair of predictors 150-1, 150-2 in this example, generated by the cross descriptor predictor, e.g., 122 in Figure 5. Again, pairs of descriptors 136-1, 136-2 are extracted from unlabeled exemplars 142 in a pair of descriptor extractors, e.g., 134-1, ..., 134-N of Figure 3.

Also in this example, a pair of label generators 140-1, 140-2 automatically generate labels (e.g., 144-1, ..., 144-N in Figure 4) for all of the unlabeled exemplars 142. The extracted descriptors 136-1, 136-2 and predictors 150-1, 150-2 are each passed to a corresponding label predictor 124-1, 124-2, e.g., in the cross descriptor prediction unit 124 in Figure 2. The predictor combination unit 126 combines the predicted labels from label predictors 124-1, 124-2 with the automatically generated labels from label generators 140-1, 140-2 to refine the prediction result. It should be noted that, although shown for labeling exemplars with 2 descriptors in this example and the above examples, this is for a minimum example only and not intended as a limitation. The present examples are easily expanded to apply to exemplars with any number of descriptors.

[0037] Continuing this example, as the cross descriptor predictor development unit 124 is presented with unlabeled exemplars 142, descriptors (e.g., 134-1, ..., 134-N) are extracted from each of the unlabeled exemplars 142. The descriptors are passed to the label predictors 124-1, 124-2, which label the descriptors based on the respective predictor 150-1, 150-2. Simultaneously, the label generators 140-1, 140-2 automatically generate labels (e.g., 144-1, ..., 144-N) from remaining unlabeled exemplars 142 exclusive of the exemplar corresponding to the particular extracted descriptor 136-1, 136-2. The predictor combination unit 126 combines the predicted labels from label predictors 124-1, 124-2 with the automatically generated labels from label generators 140-1, 140-2 to refine the prediction result.

[0038] Advantageously, a preferred embodiment system provides cross descriptor learning to leverage unlabeled exemplars using multiple descriptors, where each individual descriptor may be insufficient in and of itself to describe a metadata model for any one of the unlabeled exemplars. Further, for unstructured information management, where only the unlabeled exemplars are available to a preferred embodiment system, the exemplars may still be labeled automatically. Also the present invention is not hampered by strict requirements for single descriptor views and, further, supports multiple descriptor views. So, a system constructed according to a preferred embodiment of the present invention automatically creates metadata

from unstructured information over multiple descriptors, even without prior manual annotation of the unstructured data.

[0039] Additionally, a preferred embodiment system provides a pseudo-federated search within a single system to automatically generate a set of labeled exemplars from the unlabeled exemplars. Optionally, labels may be generated using any well known search engine, such as Yahoo® Search and Google™ Search, to rank search results with text and hyperlinks. Thereafter, the system may build descriptors for the same unlabeled descriptors guided by the search results for a pseudo-federated search. So, an annotator for automatically labeling unlabeled exemplars according to the present invention may be built from the unlabeled exemplars, from a separate set of labeled exemplars or from the labels generated for the unlabeled exemplars. Thus, unlike prior techniques that required sets of labeled exemplars for any refinement over unlabeled exemplars, a preferred embodiment machine learning system can generate refined labels just from the unlabeled exemplars alone.

[0040] While the invention has been described in terms of preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims. It is intended that all such variations and modifications fall within the scope of the appended claims.

CLAIMS

What is claimed is:

1. A method of cross descriptor learning using unlabeled exemplars, said method comprising of the steps of:

a) extracting descriptors for each of a plurality of unlabeled exemplars, each extracted descriptor being a representation of a corresponding one of said plurality of unlabeled exemplars;

b) automatically generating labels from said each extracted descriptor, said each extracted descriptor being used to automatically generate labels for other ones of said descriptors;

c) developing a predictor for said each extracted descriptor from generated said labels; and

d) combining predictions across predictors, labels being generated from combined said predictions for each of said unlabeled exemplars.

2. A method as in claim 1, wherein the step (b) of automatically generating labels comprises providing said unlabeled exemplars to a label provider service.

3. A method as in claim 2, wherein said cross descriptor learning is performed as a search and wherein the label provider service provides one of a plurality of search engines.

4. A method as in claim 1, before the step (d) of combining further comprising:

d1) providing corresponding manually labeled exemplars, wherein said predictors being used to automatically generate said labels for said each descriptor are generated learning representations of any singularity or a plurality of said other ones of said descriptors with ones of said corresponding manually labeled exemplars.

5. A method as in claim 1, wherein predictions are automatically combined in the step (d) of combining predictions.

6. A method as in claim 1, wherein predictions are combined in step (d) responsive to statistical learning.

7. A method as in claim 6, wherein statistical learning comprises discriminant learning.

8. A method as in claim 7, wherein said statistical learning comprises a support vector machine learning.

9. A method as in claim 1, wherein the unlabeled exemplars are unstructured information comprising images, audio, text and video.

10. A method as in claim 1, wherein each of said labels include corresponding confidences quantitatively indicating the likelihood of an object represented by a label being present in corresponding unlabeled exemplar and classifying corresponding said unlabeled exemplars.

11. A method as in claim 1, wherein said labels and corresponding confidences rank said unlabeled exemplars.

12. A method as in claim 1, wherein ones of said confidences are binary indicators indicating the presence or absence of each label in a corresponding exemplar.

13. A method as in claim 1, wherein ones of said confidences are real values, each real value indicating the presence or absence of each label in a corresponding exemplar.

14. A cross descriptor learning system comprising:

storage storing a plurality of unlabeled exemplars;

a descriptor extraction unit extracting descriptors from said unlabeled exemplars;

an automatic descriptor labeler generating labels for each extracted descriptor from other extracted said descriptors;

a cross descriptor predictor predicting labels for said each extracted descriptor from said each extracted descriptor and labels generated for each extracted descriptor; and

a label combination unit combining labels for said each extracted descriptor from said cross descriptor predictor with labels from said automatic descriptor labeler, each one of the combined said labels being applied to a corresponding unlabeled exemplar.

15. A cross descriptor learning system as in claim 14, wherein said descriptor extraction unit comprises a plurality of descriptor extractors.

16. A cross descriptor learning system as in claim 15, wherein said plurality of descriptor extractors comprises an image descriptor extractor, an audio descriptor extractor, a text descriptor extractor and at least one video characteristic descriptor extractor.

17. A cross descriptor learning system as in claim 16, wherein at least one video characteristic descriptor extractor comprises, a color extractor, texture, a shape extractor, and a frequency extractor.

18. A cross descriptor learning system as in claim 15, wherein said automatic descriptor labeler is one of a plurality of automatic descriptor labelers each generating labels for descriptors extracted from a corresponding one of said plurality of descriptor extractors.

19. A cross descriptor learning system as in claim 18, wherein said cross descriptor predictor comprises a plurality of descriptor label predictors.

20. A cross descriptor learning system as in claim 14, wherein said automatic descriptor labeler comprises a descriptor labeling service.

21. A cross descriptor learning system as in claim 14, wherein said automatic descriptor labeler comprises:

storage storing a plurality of labeled exemplars;

an extractor extracting descriptors from said labeled exemplars; and

a label predictor developer developing a label predictor from extracted said descriptors from said plurality of labeled exemplars, said descriptors extracted from said unlabeled descriptors being automatically labeled by one said label predictor.

22. A computer-readable medium having stored thereon a plurality of instructions, the plurality of instructions including instructions that, when executed by a processor, cause the processor to:

a) extract descriptors for each of a plurality of unlabeled exemplars, each extracted descriptor being a representation of a corresponding one of said plurality of unlabeled exemplars;

b) automatically generate labels from said each extracted descriptor, said each extracted descriptor being used to automatically generate labels for other ones of said descriptors;

c) develop a predictor for said each extracted descriptor from generated said labels; and

d) combine predictions across predictors, labels being generated from combined said predictions for each of said unlabeled exemplars.

23. A computer-readable medium as in claim 22, wherein the step (b) of automatically generating labels causes the processor to provide said unlabeled exemplars to a label provider service.

24. A computer-readable medium as in claim 22, before the step (d) of combining further causing the processor to:

d1) extract descriptors from manually labeled exemplars;

d2) generate learning representations from descriptors extracted in step (d1);

d3) generate a label prediction device for each descriptor using said manually labeled exemplars and said descriptors extracted from said manually labeled exemplars; and

d4) generate labels for plurality of unlabeled exemplars.

25. A computer-readable medium as in claim 24, wherein the step (d4) of generating labels further causes the processor to:

i) extract descriptors from manually labeled exemplars; and

ii) present each extracted one of said descriptors to a corresponding developed said predictor, said corresponding developed predictor predicting a label for said each extracted one.

26. A computer-readable medium as in claim 22, wherein the step (d) of combining predictions causes the processor to automatically combine predictions responsive to statistical learning.

27. A computer-readable medium as in claim 26, causing the processor to use discriminant learning for said statistical learning.

28. A computer-readable medium as in claim 26, wherein the step (d) causes the processor to act as a support vector machine during said statistical learning.

29. A computer-readable medium as in claim 22, wherein said labels generated by the processor in step (d) include corresponding confidences quantitatively indicating the likelihood of an object represented by a label being present in corresponding unlabeled exemplar and, further, classifying corresponding said unlabeled exemplars.

30. A computer program product for cross descriptor learning, said computer program product comprising a computer usable medium having computer readable program code thereon, said computer readable program code comprising:

computer readable program code means for selecting each of a plurality of unlabeled exemplars;

computer readable program code means for extracting descriptors from each selected one of said unlabeled exemplars;

computer readable program code means for automatically generating labels for each extracted descriptor from other extracted said descriptors;

computer readable program code means for predicting labels for said each extracted descriptor from said each extracted descriptor and labels generated for each extracted descriptor; and

computer readable program code means for combining predicted said labels from said computer readable program code means for predicting labels with automatically generated said labels for said each extracted descriptor, each one of the combined said labels being applied to a corresponding unlabeled exemplar.

31. A computer program product as in claim 30, wherein said computer readable program code means for extracting descriptors comprises:

computer readable program code means for image descriptor extraction;

computer readable program code means for audio descriptor extraction;

computer readable program code means for text descriptor extraction; and

computer readable program code means for video characteristic descriptor extraction.

32. A computer program product as in claim 31, wherein said computer readable program code means for video characteristic descriptor extraction comprises:

computer readable program code means for color extraction;

computer readable program code means for texture extraction;

computer readable program code means for shape extraction; and

computer readable program code means for frequency extraction.

33. A computer program product as in claim 30, wherein said computer readable program code means for automatically generating labels comprises computer readable program code means for contacting a descriptor labeling service.

34. A computer program product as in claim 30, wherein said computer readable program code means for automatically generating labels comprises:

computer readable program code means for storing plurality of labeled exemplars;

computer readable program code means for extracting descriptors from said labeled exemplars; and

computer readable program code means for developing a label predictor from extracted said descriptors from said plurality of labeled exemplars, said descriptors extracted from said unlabeled descriptors being automatically labeled by one said label predictor.

35. A computer program product as in claim 30, further comprising:

computer readable program code means for generating corresponding confidences with each label, generated said corresponding confidences quantitatively indicating the likelihood of an object represented by said each label being present in corresponding unlabeled exemplar and, further, classifying corresponding said unlabeled exemplars.

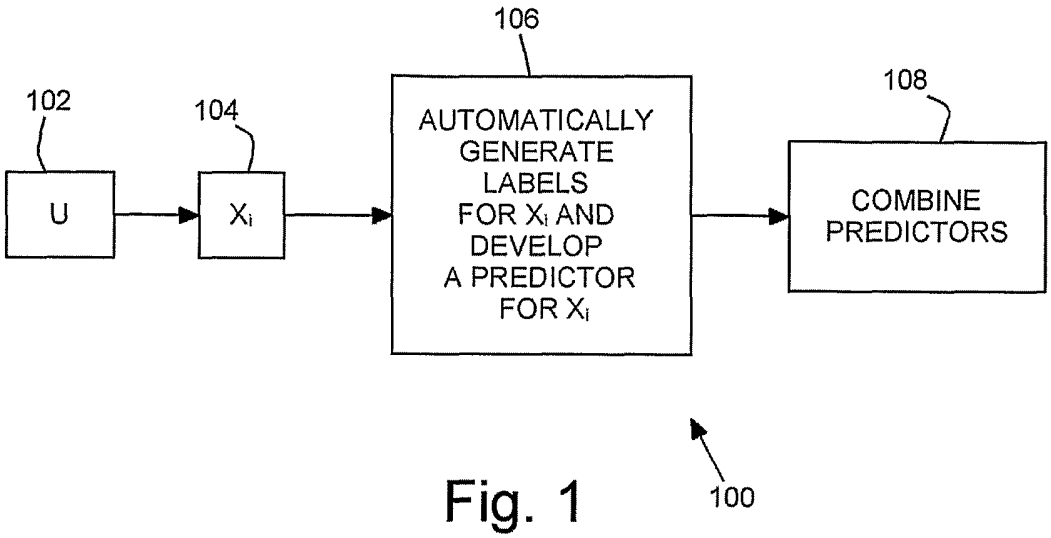

106
102
104
108
U
X_i
AUTOMATICALLY GENERATE LABELS FOR X_i AND DEVELOP A PREDICTOR FOR X_i
COMBINE PREDICTORS
100

Fig. 1

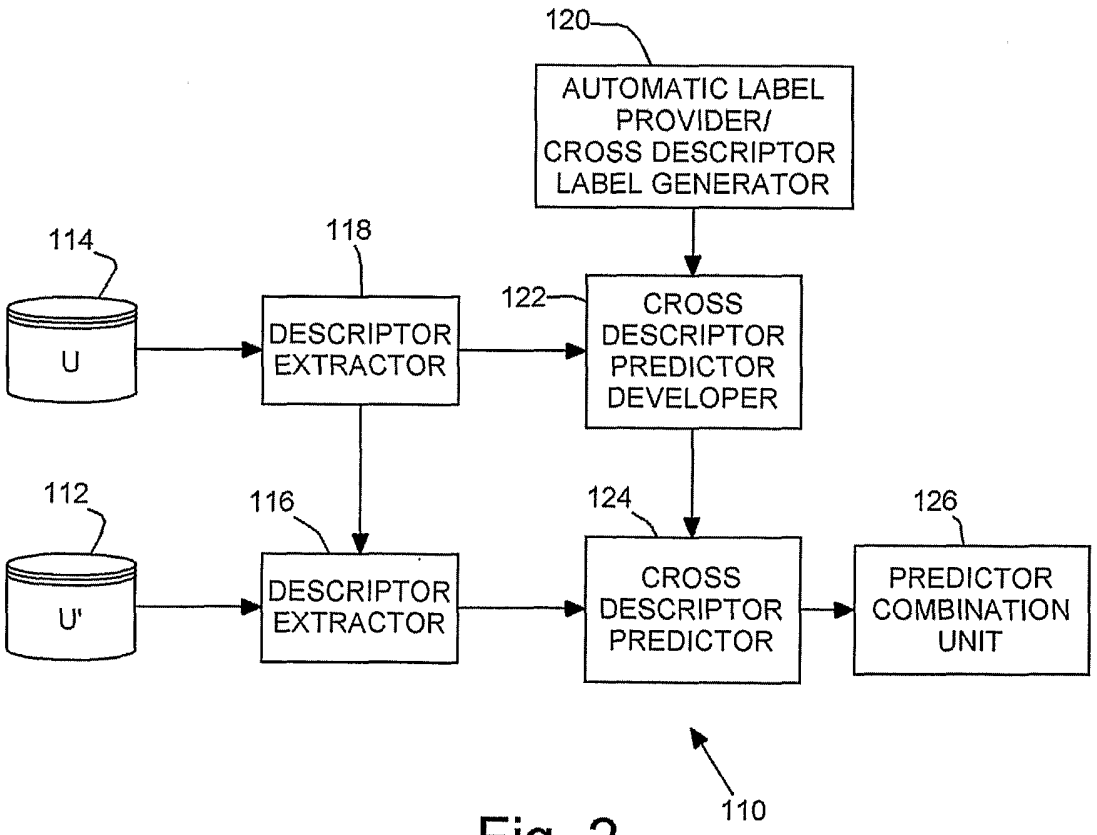

120
AUTOMATIC LABEL PROVIDER/ CROSS DESCRIPTOR LABEL GENERATOR
114
118
122
U
DESCRIPTOR EXTRACTOR
CROSS DESCRIPTOR PREDICTOR DEVELOPER
112
116
124
126
U'
DESCRIPTOR EXTRACTOR
CROSS DESCRIPTOR PREDICTOR
PREDICTOR COMBINATION UNIT
110

Fig. 2

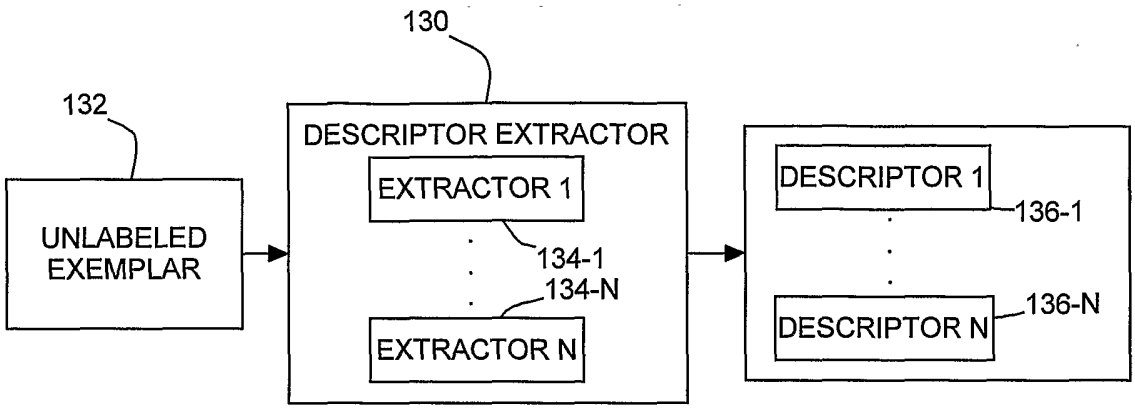

130
132
DESCRIPTOR EXTRACTOR
EXTRACTOR 1
134-1
134-N
EXTRACTOR N
UNLABELED EXEMPLAR
DESCRIPTOR 1
136-1
DESCRIPTOR N
136-N

Fig. 3

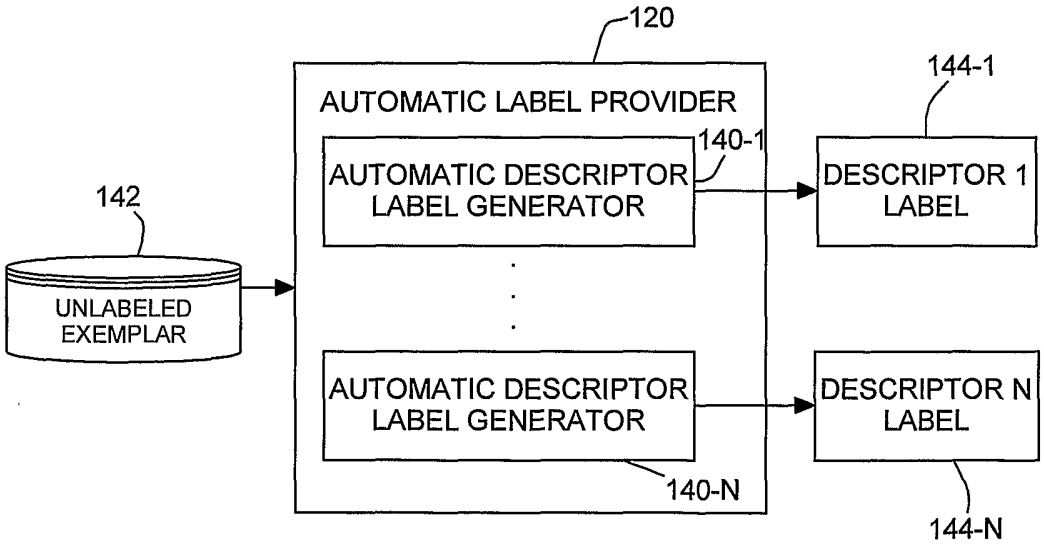

120
AUTOMATIC LABEL PROVIDER
140-1
AUTOMATIC DESCRIPTOR LABEL GENERATOR
142
UNLABELED EXEMPLAR
AUTOMATIC DESCRIPTOR LABEL GENERATOR
140-N
144-1
DESCRIPTOR 1 LABEL
DESCRIPTOR N LABEL
144-N

Fig. 4

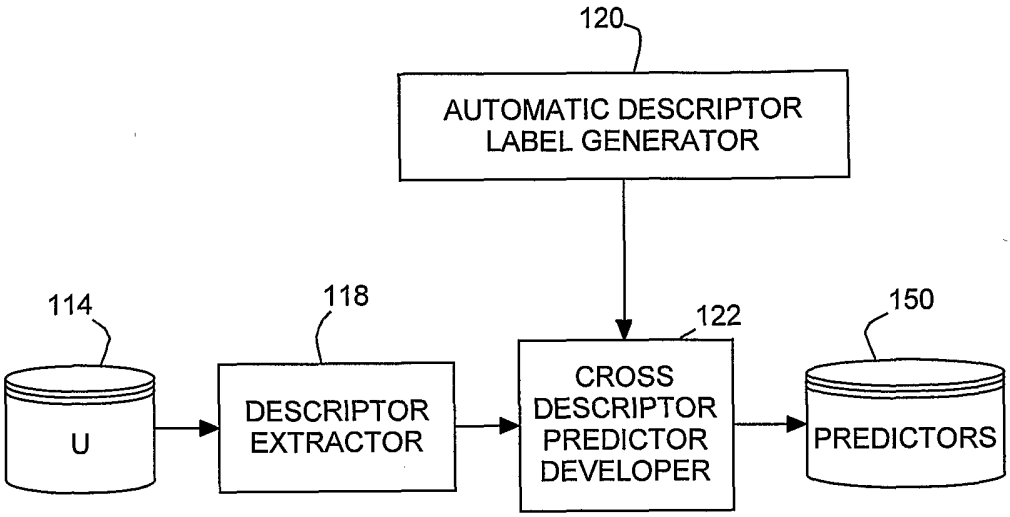

120
AUTOMATIC DESCRIPTOR LABEL GENERATOR
114
U
118
DESCRIPTOR EXTRACTOR
122
CROSS DESCRIPTOR PREDICTOR DEVELOPER
150
PREDICTORS

Fig. 5

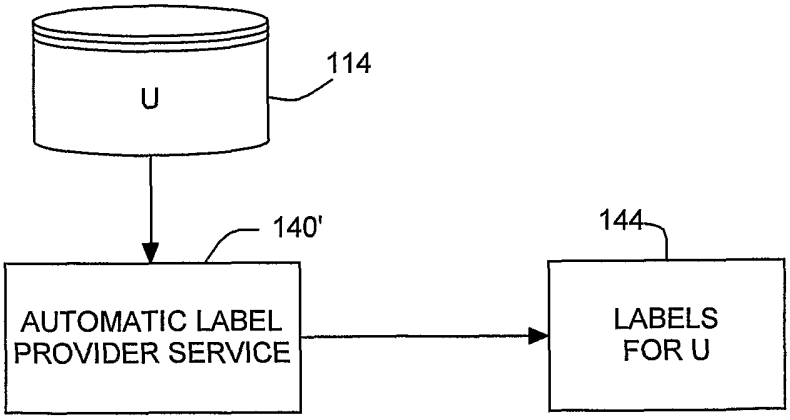

U
114
140'
AUTOMATIC LABEL PROVIDER SERVICE
144
LABELS FOR U

Fig. 7

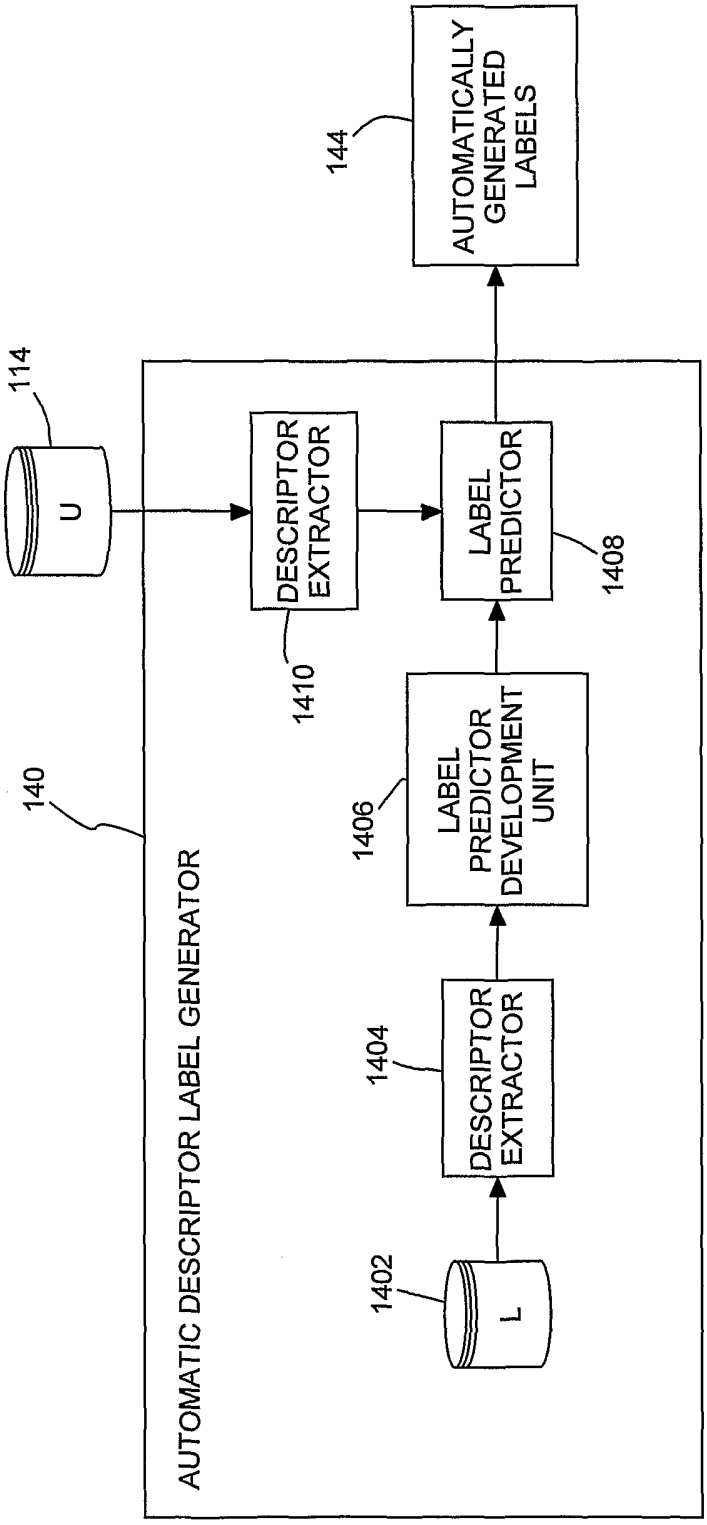

AUTOMATIC DESCRIPTOR LABEL GENERATOR
140
U
114
L
1402
DESCRIPTOR EXTRACTOR
1404
LABEL PREDICTOR DEVELOPMENT UNIT
1406
DESCRIPTOR EXTRACTOR
1410
LABEL PREDICTOR
1408
AUTOMATICALLY GENERATED LABELS
144

Fig. 6

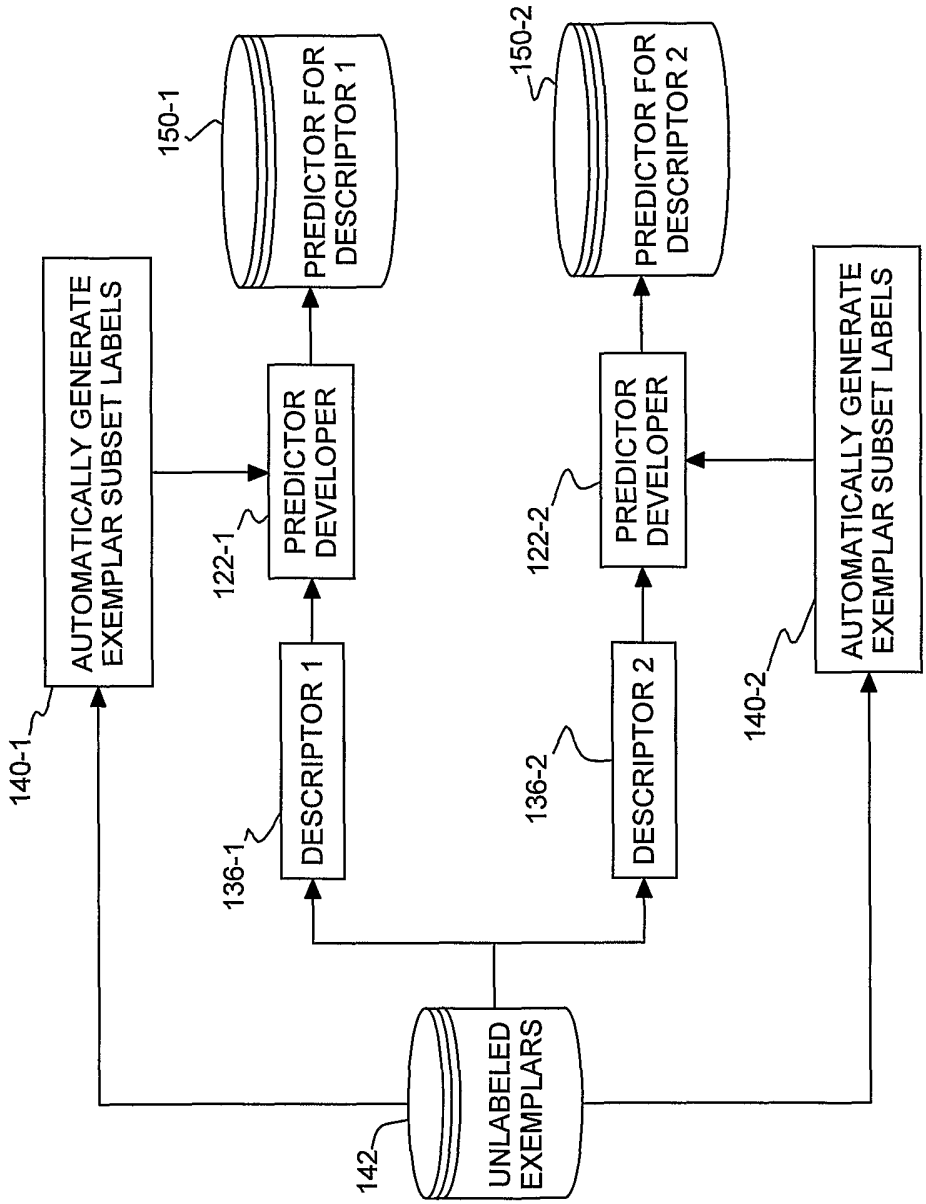

UNLABELED EXEMPLARS
142
AUTOMATICALLY GENERATE EXEMPLAR SUBSET LABELS
140-1
AUTOMATICALLY GENERATE EXEMPLAR SUBSET LABELS
140-2
DESCRIPTOR 1
136-1
DESCRIPTOR 2
136-2
PREDICTOR DEVELOPER
122-1
PREDICTOR DEVELOPER
122-2
PREDICTOR FOR DESCRIPTOR 1
150-1
PREDICTOR FOR DESCRIPTOR 2
150-2

Fig. 8

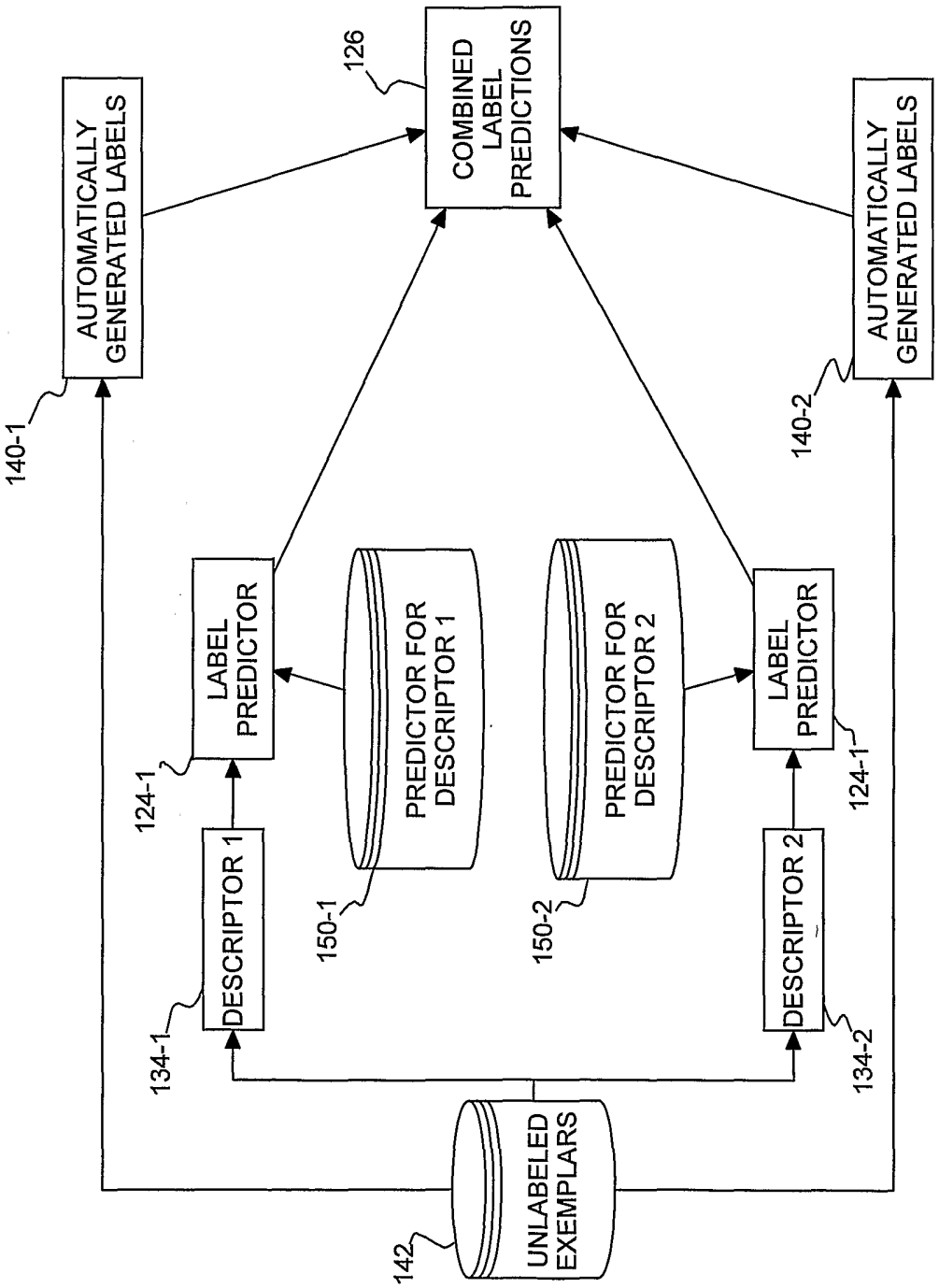

AUTOMATICALLY GENERATED LABELS
140-1
126
COMBINED LABEL PREDICTIONS
AUTOMATICALLY GENERATED LABELS
140-2
LABEL PREDICTOR
124-1
PREDICTOR FOR DESCRIPTOR 1
150-1
PREDICTOR FOR DESCRIPTOR 2
150-2
LABEL PREDICTOR
124-1
DESCRIPTOR 1
134-1
DESCRIPTOR 2
134-2
UNLABELED EXEMPLARS
142

Fig. 9