In one embodiment, a statistical pattern recognition engine determines content for a statistical pattern recognition task. The content and/or information related to the content is analyzed to determine a model to use in the statistical pattern recognition. For example, models may be classified in a plurality of domains based on different sets of data used to train models in the domain. The models may be classified based on knowledge sources used to generate the models, such as a news knowledge source, entertainment knowledge source, business knowledge source, etc. A model in the plurality of models is then determined based on the analysis where the determined model is classified in a domain. The model is then used by the statistical pattern recognition engine to perform the statistical pattern recognition task. For example, spoken words are transcribed into text using the determined model.
100

Domain #1

Model #1

Model #2

Model #3

Domain #2

Model #4

Model #5

Model #6

Fig. 2
Determine the content for statistical pattern recognition

Analyze the content and/or information related to the content to determine a domain

Determine a model associated with the domain
TECHNIQUES FOR MODEL OPTIMIZATION FOR STATISTICAL PATTERN RECOGNITION

CROSS REFERENCES TO RELATED APPLICATIONS


BACKGROUND

[0002] Particular embodiments generally relate to statistical pattern recognition and more specifically to model optimization for statistical pattern recognition.

[0003] Speech recognition involves transcribing spoken words into text automatically. Primarily, speech recognition has been applied to dictation, conversational systems, and surveillance. The task of transcribing spoken words into text for these applications is very different. For example, in the dictation application, the speech recognition system is dealing with a large vocabulary but the acoustic and vocabulary variability is limited as the dictation system is typically dealing with a single speaker. The tolerance for error is low as the speed of dictation should exceed that of the typing of the user. In the conversational system, such as those found in interactive voice response systems, the speech recognition is dealing with a small vocabulary with a small variability, such as asking users to answer very directed questions (e.g., yes or no questions). The acoustic variability, however, is high as the speech recognition system needs to act with a variety of speakers. The tolerance for error is low as an error may lead to an incorrect transaction. In the surveillance application, the speech recognition is dealing with a lower vocabulary with a high degree of acoustic and vocabulary variability. However, since the goal of the speech recognition system in surveillance is typically to reduce the amount of data that is manually processed, a high level of tolerance for error is allowed.

[0004] Because of the different characteristics for each application, different models are developed for speech recognition systems depending on the application it is used for. These models are typically geared toward the different characteristics of the applications. Because these characteristics do not generally change for the application being used, the models are somewhat static and are generally the same based on the application. For example, models for dictation systems are generally the same because of the characteristics of information processed for this application does not significantly change.

SUMMARY

[0005] Particular embodiments generally relate to model optimization for a statistical pattern recognition engine.

[0006] In one embodiment, a statistical pattern recognition engine determines content for a statistical pattern recognition task. The content and/or information related to the content is analyzed to determine a model to use in the statistical pattern recognition. For example, models may be classified in a plurality of domains based on different sets of data used to train models in the domain. The models may be classified based on knowledge sources used to generate the models, such as a news knowledge source, entertainment knowledge source, business knowledge source, etc. A model in the plurality of models is then determined based on the analysis where the determined model is classified in a domain. The model is then used by the statistical pattern recognition engine to perform the statistical pattern recognition task. For example, spoken words are transcribed into text using the determined model.

[0007] In one embodiment, a method for determining a model for a statistical pattern recognition engine is provided. The method comprises: determining content for analysis by the statistical pattern recognition engine; analyzing the content and/or information related to the content to determine a model in a plurality of models, wherein the determined model is classified in a domain in a plurality of domains, the domain including one or more models trained using information determined to include similar characteristics to the content; and providing the determined model to the statistical pattern recognition engine for statistical pattern recognition.

[0008] In another embodiment, a method for determining a model for a statistical pattern recognition engine is provided. The method comprises: receiving a plurality of files, each file including rich media content; for each of the files, performing the following: analyzing the content in the file and/or information related to the content to determine a domain that is determined to include similar characteristics to the content; determining a model in the determined domain based on the content in the file and/or information related to the content; and providing the determined model to the statistical pattern recognition engine for analysis.

[0009] In yet another embodiment, an apparatus configured to determine a model for a statistical pattern recognition engine is provided. The apparatus comprises: one or more processors; and logic encoded in one or more tangible media for execution by the one or more processors and when executed operable to: determine content for analysis by the statistical pattern recognition engine; analyze the content and/or information related to the content to determine a model in a plurality of models, wherein the determined model is classified in a domain in a plurality of domains, the domain including one or more models trained using information determined to include similar characteristics to the content; and provide the determined model to the statistical pattern recognition engine for statistical pattern recognition.

[0010] In another embodiment, an apparatus configured to determine a model for a statistical pattern recognition engine is provided. The apparatus comprises: one or more processors; and logic encoded in one or more tangible media for execution by the one or more processors and when executed operable to: receive a plurality of files, each file including rich media content; for each of the files, the logic when executed is further operable to: analyze the content in the file and/or information related to the content to determine a domain that is determined to include similar characteristics to the content; determine a model in the determined domain based on the content in the file and/or information related to the content; and provide the determined model to the statistical pattern recognition engine for analysis.
A further understanding of the nature and the advantages of the inventions disclosed herein may be realized by reference of the remaining portions of the specification and the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts an example of a statistical pattern recognition system according to one embodiment of the present invention.

FIG. 2 shows an example of models according to one embodiment of the present invention.

FIG. 3 depicts an example of a statistical pattern recognition system according to one embodiment of the present invention.

FIG. 4 shows an example of determining a model for an acoustic model according to one embodiment of the present invention.

FIG. 5 depicts a second example of determining a model for a language model according to one embodiment of the present invention.

FIG. 6 depicts a simplified flowchart of a method for performing statistical pattern recognition according to one embodiment of the present invention.

DESCRIPTION OF EXAMPLE EMBODIMENTS

System

FIG. 1 depicts an example of a statistical pattern recognition system 100 according to one embodiment of the present invention. As shown, statistical pattern recognition system 100 includes a statistical pattern recognition engine 102 and a model optimizer 104.

Statistical pattern recognition engine 102 is configured to analyze content using models. The engine is a statistical machine that uses statistical models to transcribe some pattern into a target. For example, statistical pattern engine may transcribe spoken words to text for speech recognition. Statistical pattern recognition engine 102 matches content against statistical models, such as acoustic, language, and/or semantic models to find the best match of words in the content’s audio stream. The statistical models may be any model that is generated from source data. The content may be any information, such as video, audio, conversations, surveillance information, or any other information that includes spoken words. Examples of content include webcasts, podcasts, commercials, TV shows, etc.

Acoustic models capture the way phonemes sound. The language models capture the way phonemes combine to form words and the way words combine to form phrases. Different types of models require different types of data. For example, acoustic models require data in the form of voice samples and associated transcripts. Language models require data in the form of text.

The performance and accuracy of statistical pattern recognition engine 102 may depend on how well the models used match the content being recognized. For example, acoustic models trained with noisy voice samples may work best with noisy content. Language models trained with speech text may not work very well with conversational speech. Rather, models trained with conversational speech may work better.

To provide models that are as close as possible to the content being recognized, model optimizer 104 is configured to determine a model to use based on the content received. Different models may be trained that have different characteristics. The different models may be partitioned into different domains. A domain may be based on different sets of data that are used to generate the model. For example, a domain may be defined as essentially a space that is similar by some metric. A news domain may have a news subject similarity, business domain may have a business subject similarity, entertainment domain may have an entertainment subject similarity, etc. Similarly, in the acoustic model case, it is acoustic/spectral similarity and in the language model case, it is topical similarity. Other domains may also be appreciated.

The models in the different domains are built using different knowledge sources. For example, a model in a business domain is created by using data from various business web sites. In one example, news reports from different news anchors on a web site, such as CNN.com, are used to train one or more models in a news domain. One of these models may be used when statistical pattern recognition of content from CNN is performed. Also, even if the statistical pattern recognition is not from CNN, the content may have similar characteristics to the data used to train the models from CNN because news content may have similar characteristics. For example, newscasts typically follow the same format, i.e., a newscaster is reading a story. A model trained with these characteristics may yield better statistical pattern recognition results than a model trained using people conversing in a conversational manner.

Accordingly, model optimizer 104 is configured to dynamically determine a model to use. The determination may be based on analysis of the content and/or information related to the content, such as metadata. Metadata may include a source of the content, a description of the content, data it was available, etc. Other examples of metadata may also be appreciated. In one example, if the content is news-based, then a model in a news domain may be chosen as the model to use for statistical pattern recognition engine 102. Although only one model is discussed as being selected, it will be understood that any number of models may be selected and used in a statistical pattern recognition analysis.

Real-time adaptation of the model may also be performed. For example, the model may be adapted based on information in the content or any available metadata. In one example, if it is determined that the content is about a specific news story, model optimizer 104 may modify the model to include data that is associated with the news story, such as a name of a person that is the subject of the story. This may be useful if the name of the person is not a popular name that may not have been used in the training of the model. Also, other characteristics in the model may be altered based on the content.

Once statistical pattern recognition engine 102 receives the model, it can perform the statistical pattern recognition with the content. For example, statistical pattern recognition engine 102 may transcribe spoken words in the
content to text using the determined model. The content may be received and statistical pattern recognition engine 102 may output the associated text for the spoken words with a time stamp for when the spoken words are spoken.

[0027] FIG. 2 shows an example of models according to one embodiment of the present invention. As shown, models 204 are partitioned into one or more domains. A model data determiner 202 is used to determine domain-specific information that is used to generate and update the models with training data.

[0028] As described above, a domain may be any information that is associated with a knowledge source. For example, specific domain models may be generated from knowledge sources in the entertainment, news, business, etc., domains. Each domain may be associated with any number of models 204. For example, a domain may be associated with a single model or there may be multiple models in a domain.

[0029] Model data determiner 202 is configured to determine information. Automatic gatherers of information may be used to determine information. For example, spiders may be used to search various web sites to determine information. This information may or may not be determined from rich media content. For example, a spider may determine words that are popular from text web sites or from video newscasts. Popular news stories may be determined and the keywords used in those stories are determined as words that may be more likely spoken. These words may be used to train the models in a domain.

[0030] Model data determiner 202 may then classify the information in various domains. For example, classification may be determined by measuring similarity to a domain. Model data determiner 202 may determine that information is news related and thus should be classified in the news domain. This may be determined using tags in the metadata, i.e., if the metadata says the information is news, or the metadata says the information is from cnn.com, then the information is classified in the news model. In other examples, the information is analyzed to determine the classification. For example, for acoustic models, the SNR is computed from the information to classify it.

[0031] In one embodiment, knowledge sources are determined for various domains. For example, the news domain may be associated with news knowledge sources and an entertainment domain may be associated with entertainment knowledge sources. This allows the specialization of models 204. For example, for a news domain, a CNN website may be used as a knowledge source. A spider may search the CNN website to determine domain-specific information. The CNN website may include various types of information, such as video newscasts, web page articles, etc. Characteristics of speakers may be determined. These characteristics may apply over content classified in the domain because different newscasts may include similar characteristics.

[0032] Accordingly, optimization of statistical pattern recognition may be provided because models have been optimized for specific subject matter. For example, if a model is trained with information from a news web site, it may be more relevant for news content. This is because characteristics for content in the news domain may be more like the training data culled from news knowledge sources. For example, the terms used, such as specific names for news stories, the speaking style, etc., may be better suited for news content.

[0033] Once the domain-specific information is determined, dynamic updater 206 is configured to update models in the domain. Not all of the models may be updated in the same way as each model may be different. In one example, if a new news anchor is being used in the CNN website, the acoustics for the speaker are determined and a new acoustic model 204 is then updated with the new acoustic information.

[0034] Updates to the models in a domain may be performed at various intervals. For example, the models 204 may be updated daily, hourly, in real-time, etc. The updating of the models provides more accurate models that optimize the statistical pattern recognition. For example, as a news story breaks and may become more popular, data for that news story may be determined and a model may be updated. Thus, when content for the news story is received by statistical pattern recognition engine 102, then the updated models 204 may be used to optimize the results of statistical pattern recognition of the content. Extensions to models 204 may also be used in lieu of updating the models.

[0035] Statistical pattern recognition of content will now be described. The statistical pattern recognition of content may be performed for any reasons. In one embodiment, statistical pattern recognition of content is performed to transcribe rich media content into text for the purpose of allowing searches of video content. These searches of video content are described in more detail in U.S. patent application Ser. No., entitled “TECHNIQUES FOR RENDERING ADVERTISEMENTS WITH RICH MEDIA”, filed concurrently, and incorporated by reference in its entirety for all purposes.

[0036] Statistical pattern recognition of different kinds of content requires different models to be used to optimize the statistical pattern recognition. This is different from applications, such as dictation, which typically deal with content that has similar characteristics. For example, it is expected that different users using a dictation application will always be dictating in a similar style has similar characteristics. However, statistical pattern recognition system 100 may be processing many different kinds of content that have different characteristics. Accordingly, the partitioning of models into domains classifies the models such that they are trained with similar characteristics for content being processed. Further, the dynamic determination of a domain/model for the content received provides better results when a large number of pieces of content are being processed.

[0037] FIG. 3 depicts a more detailed example of statistical pattern recognition system 100 according to one embodiment of the present invention. As shown, statistical pattern recognition engine 102 includes a domain determiner 302, a model determiner 304, and a model adapter 306.

[0038] Domain determiner 302 receives content and information related to the content. Domain determiner 302 analyzes the content and information related to the content to determine the appropriate domain for the content. For example, classification may be determined by measuring similarity to a domain. For example, the information related
to the content may indicate that the content is from a news source, such as Gannett. Domain determiner 302 may then determine that the content received is news and thus should be classified in the news domain. Other methods of determining a domain may also be appreciated. For example, analysis of the content may indicate that it is a news cast and it is classified in the news domain. Further, the content may be analyzed to determine the domain, e.g., the SNR is computed from the content to determine the domain.

Model determiner 304 is then configured to determine a model 204. In one embodiment, model determiner 304 determines a model 204 that is in the domain that was determined by domain determiner 302. Although the function of determining a domain and then determining a model the domain is described, it will be understood that separately determining of domain and then model may not be performed. Rather, a model may be determined or if a single model is associated with a single domain, then just a domain may be determined.

Model determiner 304 may determine the model in the domain using information about the content or information related to the content. For example, model determiner 304 may calculate characteristics from the content and compare them to characteristics for the models. The model that best matches the calculated characteristics may then be selected. For example, the model with the closest signal to noise ratio (SNR) to a SNR calculated from the content in the domain may be selected. Different methods for determining a model 204 will be described in more detail below with respect to examples for an acoustic model and language model.

Model adapter 306 is configured to adapt the determined model 204, as necessary. Model adapter 306 may adapt the determined model 204 based on the content and/or information related to the content. For example, if the content is about a specific person, the model may be adapted to include that person’s name. Also, if the content includes a specific newscaster, then the acoustic characteristics for the model may be adjusted based on the newscaster.

Model adapter 306 then outputs an optimized model for statistical pattern recognition engine 102. Text and time stamps for spoken words in the content are then determined.

**MODEL EXAMPLES**

As mentioned above, model determiner 304 determines a model to use. Examples of determining models for content will now be described.

**Acoustic Model Example**

FIG. 4 shows an example of determining a model for an acoustic model according to one embodiment of the present invention. As shown, model determiner 304 receives content and information related to the content. Model determiner 304 then determines a model 204. In one example, it is assumed a domain has been determined and models 204 may be chosen from the domain.

In one embodiment, models 204 may be trained with data that is partitioned according to signal-to-noise ratio (SNR) and pitch of voice samples as well as sex of the speaker in the samples. In other embodiments, the data may be partitioned according to the origin of the data, i.e. a certain web site. One or more knowledge sources may be used to train models 204.

Model determiner 304 determines the signal-to-noise ratio and pitch of the content. Model determiner 304 then determines a model 204 based on the SNR and pitch of the content. For example, the model closes to the computed SNR and pitch is determined. Also, multiple models that are closest to the computed SNR and pitch may be determined. If the SNR and/or pitch cannot be computed reliably and/or the variance for the SNR/pitch is too high, model determiner 304 may use a generic model.

Once model 204 is determined, model adapter 306 may adapt it. Model adapter 204 receives the content and information related to the content in addition to the determined model 204. It can then adapt the model 204 based on the content. For example, using methods such as maximum likelihood linear regression (MLLR), model adapter 306 adapts the model for the content. For example, the model is adapted based on characteristics for the speaker in the content.

**Language Model Example**

FIG. 5 depicts a second example of determining a model for a language model according to one embodiment of the present invention. As shown, models 204 in the different domains may be broken into more specific models, such as style models 502 and time-sensitive domain extension models 504.

Style models 502 are associated with speaking styles, such as reading style or conversational style. A reading style is when someone is reading content, such as in a newscast, and the conversational style is when people are conversing, such as holding a debate on a subject. The content may be analyzed to determine which specific speaking style is being used.

The time-sensitive domain extension models 504 are built at regular intervals using the latest information, such as text, culled from various knowledge sources. Although this is shown as a separate model, it will be understood that a model may be updated itself. Time-sensitive domain extensions are used to capture any information that may become relevant for a particular domain. Domain breaks, the names of the persons associated with the story may be determined as extensions to models in the domain.

During the statistical pattern recognition, model determiner 304 determines a model 204 by determining the most appropriate domain along with the speaking style model 502, and time sensitive domain extension model 504 to use.

Model adapter 306 then may adapt the models based on the content.

**Method for Performing Statistical Pattern Recognition**

FIG. 6 depicts a simplified flowchart 600 of a method for performing statistical pattern recognition according to one embodiment of the present invention. Step 602
determines the content for statistical pattern recognition. For example, content that may be later searched for using a search engine is used.

[0054] Step 604 analyzes the content and/or information related to the content to determine a domain. The domain may be one out of a plurality of domains where the domain may be based on different sets of data from knowledge sources that are related to the domain. Models of the domain may be generated using different characteristics from knowledge sources.

[0055] Step 606 determines a model associated with the domain. For example, the model is determined based on the content and/or information related to the content.

Normalization

[0056] In some cases, normalization of data culled from various knowledge sources may be performed. For example, the raw data from knowledge sources, such as web sites, may not be directly usable for training in models. Thus, the raw data is normalized. In one example, the normalization may be changing the sampling rate of audio. For example, audio on the web may be captured at a sampling rate of 48 kHz but the acoustic model requires the model to be in 16 kHz. The audio captured may then be performed using standard down-sampling techniques.

[0057] Other normalization may also be performed. For example, unprocessed raw HTML (hypertext transfer markup language) from a web site may not be usable for training language models. A training algorithm may only be able to take in sentences and prose within the HTML and not the symbols that are used in the language. In one example, the layout and the content of the HTML is intertwined within a file. A process may be used to determine the text/prose out of the HTML. For example, the process determines header patterns, footer patterns, etc. and uses trainable rules that can find beginning and ending of real content (as opposed to layout information). Normalization of text and language model training may also involve a set of rules that take into account document, word, sentence level constraints, in any order.

CONCLUSION

[0058] Accordingly, particular embodiments provide many advantages. For example, models may be determined that are optimized for content received. These models may be tailored toward the content because they are associated with a domain related to the content. Further, the models may be updated periodically. Thus, content for recent events may use models that have been recently updated. Additionally, the models determined may be adapted on the content received. Accordingly, this provides better statistical pattern recognition of the content.

[0059] Although the invention has been described with respect to specific embodiments thereof, these embodiments are merely illustrative, and not restrictive of the invention. For example, models for applications for any statistical pattern recognition may be optimized using embodiments of the present invention. Although examples for speech recognition are described, it will be understood that other statistical engines may be used.

[0060] Any suitable programming language can be used to implement the routines of embodiments of the present invention including C, C++, Java, assembly language, etc. Different programming techniques can be employed such as procedural or object oriented. The routines can execute on a single processing device or multiple processors. Although the steps, operations, or computations may be presented in a specific order, this order may be changed in different embodiments. In some embodiments, multiple steps shown as sequential in this specification can be performed at the same time. The sequence of operations described herein can be interrupted, suspended, or otherwise controlled by another process, such as an operating system, kernel, etc. The routines can operate in an operating system environment or as stand-alone routines occupying all, or a substantial part, of the system processing. Functions can be performed in hardware, software, or a combination of both. Unless otherwise stated, functions may also be performed manually, in whole or in part.

[0061] In the description herein, numerous specific details are provided, such as examples of components and/or methods, to provide a thorough understanding of embodiments of the present invention. One skilled in the relevant art will recognize, however, that an embodiment of the invention can be practiced without one or more of the specific details, or with other apparatus, systems, assemblies, methods, components, materials, parts, and/or the like. In other instances, well-known structures, materials, or operations are not specifically shown or described in detail to avoid obscuring aspects of embodiments of the present invention.

[0062] A “computer-readable medium” for purposes of embodiments of the present invention may be any medium that can contain, store, communicate, propagate, or transport the program for use by or in connection with the instruction execution system, apparatus, system or device. The computer readable medium can be, by way of example only but not by limitation, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, system, device, propagation medium, or computer memory.

[0063] Embodiments of the present invention can be implemented in the form of control logic in software or hardware or a combination of both. The control logic may be stored in an information storage medium, such as a computer-readable medium, as a plurality of instructions adapted to direct an information processing device to perform a set of steps disclosed in embodiments of the present invention. Based on the disclosure and teachings provided herein, a person of ordinary skill in the art will appreciate other ways and/or methods to implement the present invention.

[0064] A “processor” or “process” includes any human, hardware and/or software system, mechanism or component that processes data, signals or other information. A processor can include a system with a general-purpose central processing unit, multiple processing units, dedicated circuitry for achieving functionality, or other systems. Processing need not be limited to a geographic location, or have temporal limitations. For example, a processor can perform its functions in “real time,” “offline,” in a “batch mode,” etc. Portions of processing can be performed at different times and at different locations, by different (or the same) processing systems.

[0065] Reference throughout this specification to “one embodiment”, “an embodiment”, or “a specific embodiment” means that a particular feature, structure, or charac-
teristic described in connection with the embodiment is included in at least one embodiment of the present invention and not necessarily in all embodiments. Thus, respective appearances of the phrases “in one embodiment”, “in an embodiment”, or “in a specific embodiment” in various places throughout this specification are not necessarily referring to the same embodiment. Furthermore, the particular features, structures, or characteristics of any specific embodiment of the present invention may be combined in any suitable manner with one or more other embodiments. It is to be understood that other variations and modifications of the embodiments of the present invention described and illustrated herein are possible in light of the teachings herein and are to be considered as part of the spirit and scope of the present invention.

[0066] Embodiments of the invention may be implemented by using a programmed general purpose digital computer, by using application specific integrated circuits, programmable logic devices, field programmable gate arrays, optical, chemical, biological, quantum or nanoengineered systems, components and mechanisms may be used. In general, the functions of embodiments of the present invention can be achieved by any means as is known in the art. Distributed, or networked systems, components and circuits can be used. Communication, or transfer, of data may be wired, wireless, or by any other means.

[0067] It will also be appreciated that one or more of the elements depicted in the drawings/figures can also be implemented in a more separated or integrated manner, or even removed or rendered as inoperable in certain cases, as is useful in accordance with a particular application. It is also within the spirit and scope of the present invention to implement a program or code that can be stored in a machine-readable medium to permit a computer to perform any of the methods described above.

[0068] Additionally, any signal arrows in the drawings/figures should be considered only as exemplary, and not limiting, unless otherwise specifically noted. Furthermore, the term “or” as used herein is generally intended to mean “and/or” unless otherwise indicated. Combinations of components or steps will also be considered as being noted, where terminology is foreseen as rendering the ability to separate or combine is unclear.

[0069] As used in the description herein and throughout the claims that follow, “a”, “an”, and “the” includes plural references unless the context clearly dictates otherwise. Also, as used in the description herein and throughout the claims that follow, the meaning of “in” includes “in” and “on” unless the context clearly dictates otherwise.

[0070] The foregoing description of illustrated embodiments of the present invention, including what is described in the Abstract, is not intended to be exhaustive or to limit the invention to the precise forms disclosed herein. While specific embodiments of, and examples for, the invention are described herein for illustrative purposes only, various equivalent modifications are possible within the spirit and scope of the present invention, as those skilled in the relevant art will recognize and appreciate. As indicated, these modifications may be made to the present invention in light of the foregoing description of illustrated embodiments of the present invention and are to be included within the spirit and scope of the present invention.

[0071] Thus, while the present invention has been described herein with reference to particular embodiments thereof, a latitude of modification, various changes and substitutions are intended in the foregoing disclosures, and it will be appreciated that in some instances some features of embodiments of the invention will be employed without a corresponding use of other features without departing from the scope and spirit of the invention as set forth. Therefore, many modifications may be made to adapt a particular situation or material to the essential scope and spirit of the present invention. It is intended that the invention not be limited to the particular terms used in following claims and/or to the particular embodiment disclosed as the best mode contemplated for carrying out this invention, but that the invention will include any and all embodiments and equivalents falling within the scope of the appended claims.

1. A method for determining a model for a statistical pattern recognition engine, the method comprising:

determining content for analysis by the statistical pattern recognition engine;

analyzing the content and/or information related to the content to determine a model in a plurality of models, wherein the determined model is classified in a domain in a plurality of domains, the domain including one or more models trained using information determined to include similar characteristics to the content; and

providing the determined model to the statistical pattern recognition engine for statistical pattern recognition.

2. The method of claim 1, wherein the statistical pattern recognition engine is configured to perform a statistical pattern recognition analysis on the content using the determined model.

3. The method of claim 2, further comprising adapting the determined model in real-time based on the content and/or the information related to the content before providing the determined model to the statistical pattern recognition engine.

4. The method of claim 1, further comprising:

determining one or more additional models based on the analysis; and

providing the one or more additional models to the statistical pattern recognition engine for the statistical pattern recognition.

5. The method of claim 1, further comprising:

determining information related to a domain in the plurality of domains; and

modifying a model in the plurality of models associated with the domain based on the information.

6. The method of claim 5, further comprising normalizing the information before modifying the model.

7. The method of claim 1, wherein domains in the plurality of domains are associated with different knowledge sources.

8. The method of claim 1, wherein models in the plurality of domains are trained using different sets of data that include different characteristics.

9. The method of claim 1, wherein the determined model comprises a speaking style model and/or time extension model.
10. The method of claim 1, wherein the statistical pattern recognition comprises speech recognition.

11. A method for determining a model for a statistical pattern recognition engine, the method comprising:

- receiving a plurality of files, each file including rich media content;

- for each of the files, performing the following:
  - analyzing the content in the file and/or information related to the content to determine a domain that is determined to include similar characteristics to the content;
  - determining a model in the determined domain based on the content in the file and/or information related to the content; and
  - providing the determined model to the statistical pattern recognition engine for analysis.

12. The method of claim 11, wherein models in the plurality of domains are training used different sets of data that include different characteristics.

13. The method of claim 11, wherein the statistical pattern recognition comprises speech recognition.

14. An apparatus configured to determine a model for a statistical pattern recognition engine, the apparatus comprising:

- one or more processors; and

- logic encoded in one or more tangible media for execution by the one or more processors and when executed operable to:
  - determine content for analysis by the statistical pattern recognition engine;
  - analyze the content and/or information related to the content to determine a model in a plurality of models, wherein the determined model is classified in a domain in a plurality of domains, the domain including one or more models trained using information determined to include similar characteristics to the content; and
  - provide the determined model to the statistical pattern recognition engine for statistical pattern recognition.

15. The apparatus of claim 14, wherein the statistical pattern recognition engine is configured to perform a statistical pattern recognition analysis on the content using the determined model.

16. The apparatus of claim 15, wherein the logic when executed is further operable to adapt the determined model in real-time based on the content and/or the information related to the content before providing the determined model to the statistical pattern recognition engine.

17. The apparatus of claim 14, wherein the logic when executed is further operable to:

- determine one or more additional models based on the analysis; and

provide the one or more additional models to the statistical pattern recognition engine for the statistical pattern recognition.

18. The apparatus of claim 14, wherein the logic when executed is further operable to:

- determine information related to a domain in the plurality of domains; and

- modify a model in the plurality of models associated with the domain based on the information.

19. The apparatus of claim 18, wherein the logic when executed is further operable to normalize the information before modifying the model.

20. The apparatus of claim 14, wherein domains in the plurality of domains are associated with different knowledge sources.

21. The apparatus of claim 14, wherein models in the plurality of domains are training used different sets of data that include different characteristics.

22. The apparatus of claim 14, wherein the determined model comprises a speaking style model and/or time extension model.

23. The apparatus of claim 14, wherein the statistical pattern recognition comprises speech recognition.

24. An apparatus configured to determine a model for a statistical pattern recognition engine, the apparatus comprising:

- one or more processors; and

- logic encoded in one or more tangible media for execution by the one or more processors and when executed operable to:
  - receive a plurality of files, each file including rich media content;
  - for each of the files, the logic when executed is further operable to:
    - analyze the content in the file and/or information related to the content to determine a domain that is determined to include similar characteristics to the content;
    - determine a model in the determined domain based on the content in the file and/or information related to the content; and
    - provide the determined model to the statistical pattern recognition engine for analysis.

25. The apparatus of claim 24, wherein models in the plurality of domains are training used different sets of data that include different characteristics.

26. The apparatus of claim 24, wherein the statistical pattern recognition comprises speech recognition.