Abstract: Methods, devices, systems and tools are presented that allow the summarization of text, audio, and audiovisual presentations, such as movies, into less lengthy forms. High-content media files are shortened in a manner that preserves important details, by splitting the files into segments, rating the segments, and reassembling preferred segments into a final abridged piece. Summarization of media can be customized by user selection of criteria, and opens new possibilities for delivering entertainment, news, and information in the form of dense, information-rich content that can be viewed by means of broadcast or cable distribution: “on-demand” distribution, internet and cell phone digital video streaming, or can be downloaded onto iPod™ and other portable video playback devices.
Published: 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.
Preference Based Automatic Media Summarization

This application claims the benefit of U.S provisional application Nos. 60/745588, filed April 25, 2006, 60/890,214, filed February 16, 2007, and 80/892,311, filed March 1, 2007, all of which are specifically incorporated by reference in their entireties.

Field of the Invention

The invention relates to the audiovisual arts and to software and systems for manipulating data sets in the cinematic and literary fields.

Background

Audiovisual presentations are common and play a major role in modern life. Movies, demonstrations, television programs and other audiovisual information may be prepared and stored as files, which may be manipulated by computer or other electronic devices. Such files may be edited via a variety of technologies. See for example U.S Nos. 3,890,639, 5,367,341, 5,383,197, 6,052,508, and 6,965,723. The contents of each cited patent, and in particular the hardware, software, methods and systems for storing and editing such audio video files specifically are incorporated by reference in their entireties.

Audiovisual presentations prepared by these and other editing methods have proliferated greatly. An emerging problem is the length of many audiovisual presentations. Unfortunately, many movies, television programs and other series last an hour or more. Busy people often cannot watch such presentations. Some movies and television shows thus are not seen by some people because of their length. Accordingly, any method that can provide the basic information of a long audiovisual presentation in a shorter time would expand the audiovisual market and promote further business. At the same time, the advent of small screen video players, such as the video iPod™ and video cell phone, call for short format audiovisual presentations. The small-screen market for feature films and television programs certainly would improve with shorter running times in the same vein, long download times and large file sizes of full-length feature films are problematic to
media consumption. Despite these needs, attempts to systematize the shortening of audiovisual files have not progressed sufficiently.

**Summary of the Invention**

The problem of insufficient time for watching a long, detailed audiovisual presentation such as a movie, is alleviated by devices and methods that remove at least most of an audiovisual presentation, while preferentially maintaining important information. In an embodiment, a subject audiovisual presentation is provided having a video portion which may include, for example, segments identified as shots, images, frames, image sequences; as well as an audio portion which may include segments of spoken dialogue, and/or other non-dialogue audio, such as sound effects, musical scoring, and/or audience reactions. The identified components contained within said segments are rated by the frequency of their existence in the overall audiovisual presentation. The component rating optionally is further weighted by a database of user preferences and other factors, such as the perceived presence of stress or conflict. The segments are then rated based on the combined ratings of the components that they contain. The segments deemed to have less significant content are skipped, or abbreviated, creating a summarized output file.

An embodiment provides a system for the automatic summarization of an audiovisual presentation wherein the audiovisual presentation comprises multiple Dialogue Segments, Non-Dialogue Segments, Non-Dialogue Segments with Music, and Shot Segments. This system desirably comprises a storage medium for storing an electronic copy of the audiovisual presentation; a stored program that identifies the segments, evaluates and generate significance ratings for the audio and visual segments based on the presence of one or more Segment Components within the segments; stored and tagged memory locations for monitoring timecode locations and ratings of the segments as well as identification and significance ratings of the Segment Components they contain and, a stored program that filters out less-desirable segments based on a mathematical comparison of the Significance Values of the segments to a selectable Segment Significance Value Threshold,

These and other embodiments can address shortcomings of the prior art by one or more of. using input method step(s) to allow user preference customization; use of
simple but robust inductive machine teaming text pattern recognition method of
costent auto-summarization; definition of a multiplicity of parameters to fine-tune
segment boundaries and rating prioritization, including for example, "global formulate
experimentation " Definition of a multiplicity of parameters in particular allows for the
use of "global formulaic experimentation" to affect the interaction of the inputted
statistics in different ways thereby fine-tuning the output file.

Further embodiments provide methods to reduce or eliminate artifacts that can result
from auto-summarization, creating smoothly consumable output files; methods that
maintain the "entertainment value" of a full-length audiovisual presentation in a
summarized output file; and real-world solutions to the computational complexity
theory as the theory is applied to relationships between current microprocessor
technology and the end-user auto-summarization experience.

Other embodiments will be appreciated by a skilled artisan upon reading the
specification.

Brief Description of the Drawings

Figure 1 shows a method for shortening a text document based on noun usage.

Figure 2 is a news article that was shortened in an embodiment.

Figure 3 is the news article of Figure 2 that has been separated into Sentence
Segments.

Figure 4 is the news article of Figure 1 shortened by 48%(number of sentences).

Figure 5 is the news article of Figure 1 shortened by 75%(number of sentences).

Figure 6 is a matrix from which nouns and sentences were rated for automatic
summarization.

Figure 7 shows an example of a Segment Attribute Summary (Shot Segment).

Figure 8 shows an example of a Segment Attribute Summary (Dialogue Segment).
Figure 9 shows an example of Primary Significance Curve Overlays Plotted on 2-axis graph.

Figure 10 shows an example of a Secondary Significance Curve Overlays Plotted on 2-axis graph.

Figure 11 is a detail of Figure J, showing an example of a Frame Attribute Cross Section.

Figure 12 shows an example of a Frame Segment Summary.

Figure 13 shows an example of a typical motion picture and television audio track layout.

Figure 14 shows an example of using contour smoothing interpolation in the management of segment clusters or frame clusters.

Figure 15 shows secondary threshold zones as they relate to a Significance Curve Overlay.

**Detailed Description**

An embodiment provides a method for the automatic or manual summarization by at least 50% in time length, of an audiovisual presentation of at least 15 minutes in length, comprising the steps of: a) providing an audiovisual presentation, such as a feature-film, documentary video, documentary film, television program, or instructional video, stored on a linear optical or analogue medium, such as film or magnetic tape, or stored on a digital linear or non-linear medium, such as magnetic tape, hard disk drive, RAM (random access memory), or Flash RAM; b) identifying and dividing the audio portion of said audiovisual presentation into dialogue segments and non-dialogue segments; c) identifying and dividing the visual portion of said audiovisual presentation into shots or shot segments; d) taking into account at least 75% of the segments of b) and c), and rating them as to their significance to the audiovisual presentation; and e) deleting less significant dialogue segments, non-dialogue segments, and/or shot segments, and reassembling the more significant segments in their original order, thereby forming a summarized version at
least 50% shorter in time length, while preserving the story line, plot, message, and other important concepts of the unabridged audiovisual presentation.

An embodiment preserves at least most of the story of the unabridged audiovisual presentation in the summarized audiovisual presentation. A story line may be the plot of a book, play, or fsim. A story line also may be the plan or main story of a literary work. Segments that come together in a summarized audiovisual presentation generally have a relationship with a story line of the unabridged piece. In an embodiment, an unabridged audiovisual presentation is used to create a shorter output file that is referred to herein as a summarized audiovisual presentation. In an embodiment, the methods for this, as described herein produce an automatic choice of what material comprises the complimentary portion of the unabridged audiovisual presentation that can be removed without compromising the story line contained therein.

The derivative product of this method, or summarized audiovisual presentation is much different from a commercial, trailer, or highlight sequence. A commercial, trailer, or highlight sequence often is created by identifying highly desirable video shots and dialogue lines from an unabridged audio-visual presentation and adding those shots together to represent a mere sample of the original. In an embodiment described herein, portions of an original are selected for removal rather than for addition. Furthermore, the time order of segments from the unabridged piece generally is maintained and a storyline is followed. Still further, embodiments are not shortened for the sole intent of providing the viewer a tease, or glimpse of a story. In an embodiment, a summarized audiovisual presentation contains between 10% and 80% of the unabridged piece.

Although greedily summarized, in many instances, the summarized audiovisual presentation produced, conveys the story line and maintains the "entertainment value" of the full-length unabridged audiovisual presentation. This gives the viewer of the summarized audiovisual presentation the sensation of having watched the unabridged piece.

One or more shorter versions of an audiovisual presentation such as a movie are made by separating an unedited, longer version into segments, such as Dialogue
Segments, Non-dialogue Segments, and Shot Segments. The segments are characterized by significance. Higher significant segments are maintained preferentially, while lesser significant segments are not included in a reassembled version. The preservation of story sine meaning via this process was surprising, given how much (typically 70% or more) of the longer version was removed. This segmentation may be achieved by any of a number of techniques as will be appreciated by a skilled artisan.

Once separated, many, most, or preferably all of the segments are evaluated and rated according to significance. After rating, segments of higher significance desirably are grouped separately from segments of lesser significance. From this separation, characterization and optional regrouping, a much shorter version of the unabridged audiovisual presentation is made that maintains much of the original information content. The shorter summarized audiovisual presentation may be made, for example, by deleting less significant segments from the remainder, by grouping more significant segments into a new, separate file. In a particularly desirable embodiment, the input file remains uncut and the less significant segments are merely skipped during playback.

A skilled artisan readily will appreciate techniques for carrying out some of these activities, as represented, for example in U.S. 8,640,044, the contents of which, and particularly pertaining to storage and playback, is incorporated by reference in its entirety

Separation (the audiovisual presentation into segments)

A full-length, unabridged subject audiovisual presentation, is a self-contained video or film presentation between 5 minutes and 10 hours in length, and more typically between 30 minutes and five hours in length, consists of visual and synchronized narrative audio material, and is stored on a linear or non-linear media. Examples of a full-length audiovisual presentation include a feature-film, documentary video, documentary film, television program, or instructional video and combinations of such.
After initially storing an audiovisual presentation in a media such as a computer memory, or computer accessible memory, the audiovisual components preferably are separated into distinct segments. A segment is "distinct" in that the segment has a beginning point and an ending point. The beginning and end points may be determined manually or automatically.

A segment or segments of text desirably can be extracted from the spoken language or dialogue from a full-length audiovisual presentation by a variety of techniques, including but not limited to: a) speech recognition (see for example, U.S. 4,947,438, the contents of which, and particularly pertaining to speech recognition, are incorporated by reference in their entireties; and b) character recognition (see, for example, U.S. 4,989,258, the contents of which, and particularly pertaining to text character recognition, are incorporated by reference in their entireties. Computerized character recognition can gather all of or substantially all of on-screen text including subtitle information. In an embodiment, subtitle information is used that is imbedded on a DVD or other medium and can be used directly to create a text file of all subtitles from the full-length audiovisual presentation.

In an embodiment, segments are evaluated for signs of stress or conflict by the observation of volume and tone of speech, shape of mouth and eyebrows, and the comparison of Dialogue Segment Components to Verb and Adjective sets. Conflict, traditionally a key component of a story line, is weighted separately, typically more heavily, by a variable of user input.

A segment may comprise both audio information and corresponding visual information, or may consist of only audio, or only visual information. Preferably, an audio segment and a visual segment are linked by an associated tag and the audio segment alone initially is analyzed and categorized. An associated tag can for example be a number but can be more, and may specify the start and/or stop locations of the segment within the Sarger audiovisual presentation, and may specify the time duration of the segment. A tag may also refer to a stored table or matrix that contains information concerning the audio and/or visual segments associated with the tag.
Considered more closely, a segment often is a continuous set of data such as an audio sequence, visual sequence, and/or a combined visual and audio sequence. The actual beginning and/or end of a segment desirably may be determined by the information in the segment, in an embodiment, a segment's start and end points may be determined (set by a computer process, manually or a combination of both) by analyzing the presence of dialogue. A minimum quiet (non dialogue) space of time may be used to determine the segment border, in some compositions, a segment end may be determined by the presence of certain kinds of words, by the volume of a word, inflection used to express the word, by the co-presence of the word with video information, or by a change in video background scene. Combinations of such events also may be used. In an embodiment, Dialogue Segment division occurs when a new character or person begins speaking. Which character or person is speaking can be determined by analysis of the audio segment using voice recognition, or by analysis of the visual segment using face recognition and shape recognition (for mouth movement detection), or a combination thereof.

Non-dialogue information may be visual information, may be strategically placed pauses in dialogue, or other non-verbal content information.

A dialogue segment is a sentence spoken by a fictional or non-fictional character, the words making up a part in a drama or non-fiction story.

Dialogue information can be words uttered by a person, or other entity, such as a computer, talking animal or talking animated object. Narration that is not linked specifically to video movements may be included as dialogue. Dialogue may be aurai but also may be visual, such as a communication by a sign, a video screen, subtitle, or pantomime. In a desirable embodiment all such dialogue may be expressed, translated, or interpreted as words, which can be entered into a file for automatic or manual analysis.

In another embodiment that uses a term familiar to many in the Film and Television industry, segments of the subject audiovisual presentation are identified and analyzed as shots. A shot is defined in "The Complete Film Dictionary" (Konigsberg, 1987), as; "the continuous action on the [cinema] screen resulting in what appears to be a single run of the camera". The same source also defines a shot as: "the
building block of a film." In this embodiment at least most of the shots of the audiovisual presentation are analyzed for significant content.

Segment Characterization

In most instances dialogue segments were found more valuable than non-dialogue segments for maintaining the meaning of a larger unabridged audiovisual presentation when preparing a shorter version. Thus, in a most desirable embodiment, audio information initially is separated into dialogue and non-dialogue segments. This separation by itself is a valuable type of characterization.

Significance of a character or other noun may be determined by the number of times that character or noun is mentioned in the unabridged audiovisual presentation. The significance of a character may be further determined by the number of sentences and the amount of time that the character speaks as compared to the other characters.

Most desirably, words used in dialogue segments are reviewed and significance is determined by the presence of and/or frequency of specific words and/or word types. Nouns were found to be most useful in this context. For example, segments that contain specific nouns that refer to the same thing or person were scored based on the presence of the noun in the segment. This noun scoring used for several audiovisual pieces, too long to present here, was found to provide shortening while significantly maintaining meaning. The "presence" of the noun may be merely quantitative, such as the existence of or how many times the noun is uttered or otherwise displayed. The "presence" may be modified by other parameters, such as the length of the segment, or the existence of another noun or nouns that are associated with a strong story line.

In an embodiment, the audio portion of an audiovisual presentation is divided into three types of segments: Dialogue Segments, Non-dialogue Segments, and Non-dialogue Segments with Music. Non-dialogue Segments with Music sometimes contain material of little significance to the story line of the unabridged audio-visual presentation, but may contain material that arguably provides entertainment value.
An example may be a musical number or a car chase. Removal of car chases and/or musical numbers is a desirable step in one embodiment.

In an embodiment, assuming a shortening ratio of five to one, any random 20% portion of a non-dialogue segment with music can be used with the intention of adding entertainment value to the summarized audiovisual presentation. A crude method of shortening the Non-dialogue Segments with Music by this exemplified ratio of 1:5, is to use the first 10% and the last 10%, but to remove the middle 80% of the non-dialogue segment. Although this embodiment is exemplified with a 20% portion and a ratio of 1:5, any other range of percentages and ratios similarly may be used.

In the case of a musical number, at least the first verse and the final chorus of the number, or at least the final chorus, optionally may be included in the summarized audiovisual presentation, in whatever configuration does not exceed the overall percentage of desired summarization. Methods can be used to separate sections of songs into segments for rating and summarization (see for example, U.S. Application 20060210157, the contents of which, and particularly pertaining to the summarization of music videos, are incorporated by reference in their entirities.

In a most desirable embodiment a non-dialogue sequence can be summarized by utilizing the ratings of visual segment components. The visual presence of a component can be used to generate that segment’s Shot Significance Value. The Shot Significance Value of each shot is compared to a Shot Significance Value Threshold set by the user. Shots that have Shot Significance Values below the Shot Significance Value Threshold are skipped in the summarized output file.

In an embodiment, the shorter the shot segment, the higher the significance. Consider this: if 'a picture paints a thousand words', then in an embodiment, a cinematic shot could be said to do the same. If a two second shot and a ten second shot both 'paint' a thousand words, then the two second shot would 'paint' 500 words per second and the ten second shot would 'paint' 100 words per second, therefore, each second of the two second shot is more valuable than that of the ten second shot. This type of evaluation could represent a parameter of shot significance rating.
In an embodiment, longer Shot Segments are split into shorter segments of equal lengths. Each of the shorter Shot Segments may be individually analyzed for content significance. In a preferred embodiment, longer Shot Segments are broken into smaller segments, the boundaries determined by the entrance or exit of characters or objects, abrupt changes in speed of characters or objects, abrupt flashes of sight, or changes in scenery. If considered significant enough to keep, Song shots that contain little or no change in speed or presence of objects or characters will be cut at the beginning and end of the shot, equidistant from the middle of the shot. The remaining length of the shot will be a fraction of the original shot length coinciding with the overall ratio of the desired compression of the entire audiovisual presentation.

When dialogue is present, shots segments may be weighted as less significant or completely ignored as input, therefore increasing computer-processing efficiency.

**Noun Significance Rating Methods**

In an embodiment, a segment or segments of text can be extracted from the spoken language or dialogue from a full-length audiovisual presentation by a variety of techniques, including but not limited to; speech recognition, character recognition, and the retrieval of embedded subtitle information.

Any one or combination of these methods may be employed to create a transcript of all or substantially all of the dialogue in a full-length audiovisual presentation.

In many instances, terms used in text segments are compared with their use, and in particular, frequency of use, in larger regions of the starting material. For example, identified Word Segments may be compared with Word Segment usage in an adjacent or other larger region, such as an associated section of dialogue, including for example, multiple adjoining segments, or other regions that may be identified by function (e.g. all spoken by the same person, or spoken at the same location) or other characteristics. Noun frequency of usage is particularly favored and can be used profitably both for shrinkage of text documents as well as shrinkage of audiovisual documents. An example of noun frequency usage for strict text file abridgement is presented next, but the same and similar techniques may be used for
Abridging audiovisual files, by selecting desired audio segments, and including the
visual segments that are linked in time with the audio segments.

In an embodiment shown in Figure 1, text is summarized by identifying and rating the
nouns contained in a text document, such as a transcript, article, paragraph, book,
chapter, or essay. Said text may be generated from a speech or subtitles of an
audio or audiovisual work.

The first step in the embodiment of Figure 1 is comparing each Word Segment of a
text document to a user inputted set of nouns thought to be of significance to said
document. This set may be empty in some cases, where the user may be
unfamiliar with the contents of the segment of text being summarized. In the case of
the summarization of an audiovisual presentation, the names of story characters or
other nouns known to be of importance to the story can be inputted, manually or
automatically.

The second step is comparing each Word Segment of a text document to one or
more stored sets of nouns specific to a certain field of expertise (such as Medicine,
Law, Music, Post-production, Yoga, Buddhism, Business, etc). Which, if any, of
these sets of nouns are accessed is a function of user preference.

The third step is identifying each noun in a text document by comparing every Word
Segment in said document to a set of stored common nouns. Said set being part of
a locally stored database or remote, on-line database such as Word Net. Any words
identified as nouns by matching nouns in such stored sets will be added to a table
called the Found Noun Set. Once this set contains at least one entry, it will be the
first stored noun set used to compare to each word of the segment of text. The
found noun set will contain only one entry of each noun identified. Subsequent
matches to this set will result in a tally within a field of the set's table, in effect
counting the frequency of use of each entry. Each word of the segment of text being
summarized is compared to each of the above noun sets in the order described. If
no match is found then the word is skipped. If a match occurs, the search continues
immediately to the next word segment in the text document. The name of the noun
set from which the match occurred will be recorded in a field of the Found Noun Set
in the same line also containing the identified noun and its occurrence count.
Identification of the noun set from which the match occurred may weight the noun occurrence differently based on user preference.

A Named Entity Tagger may be employed to identify and tag entities and other proper nouns. For the purposes of simplifying analysis parameters, any noun that is capitalized will be considered a proper noun. Resulting rating errors will be within an acceptable range and will not noticeably affect the output file.

The fourth step is to assign a rating value to each noun by tabulating its frequency of use in the overall text document. The rating of each noun is considered the noun’s Noun Significance Value. Each pronoun in a document is identified by comparing every word in that document to a set of stored Pronoun Type Sets. Each pronoun is matched to its antecedent by observation of their proximity of placement in the sentence structure of the current sentence as well as the adjoining sentences preceding. The Noun Significance Value of each identified antecedent is increased based on the frequency of identified pronouns that describe it. Relative pronouns may be ignored.

In the case of the summarization of an audiovisual presentation, the character speaking the sentence will be identified by manual or automatic methods and this information will determine the subject of any first person pronouns.

The fifth step is identifying each Sentence Segment in the text document, in writing, a sentence usually begins with a capital letter and concludes with appropriate end punctuation. In speaking, a sentence exhibits characteristic patterns of stress, pitch, and pauses that may differ from adjacent sentences.

The sixth step is to assign a Sentence Segment Significance Value to each Sentence Segment based on formula relating the significance values of nouns contained in that sentence. In an embodiment, a mere summing of the Noun Significance Values contained in the Sentence Segment will provide the Sentence Segment Significance Value.

The seventh step is to compare each Sentence Segment Significance Value to a threshold set by user input. Sentences with ratings that fail below the threshold will
be added to a table of significant sentences. Sentences with ratings above the threshold will be added to a table of more significant sentences.

The eighth and final step is to reassemble the Sentence Segments from the table of more significant Sentence Segments or to merely skip the less significant Sentence Segments. The resulting output of this method represents a summarized version of the original segment of text.

Sentence Segment Ousters are groups of adjacent significant Sentence Segments or groups of adjacent non-significant Sentence Segments. In lengthy text documents, auto-summarization is likely to produce such clusters. Any single non-significant Sentence Segments surrounded by two significant Sentence Segment Clusters shall be included in the summary as if they were rated as significant. Likewise, any significant Sentence Segments surrounded by two non-significant Sentence Segment Clusters will be omitted from the summary as if they were rated as non-significant. The number of contained Sentence Segments needed to qualify said Sentence Segment Clusters as controllers in such cases, will be determined by a formula based on the total number of Sentence Segments in the over-all text document.

In the case of a short to medium length text document, such as a news article, the first sentence may be included in the summary whether or not it is rated as significant.

The method of noun analysis for shortening may be used to prioritize audio segments, but the same procedures may be used on text documents themselves. Such text shortening is another embodiment contemplated as part of the invention. A practical working example useful to illustrate a method for audiovisual piece shortening as well as regular text article shortening is shown in Figures 2 to 8. Figure 2 is a news article. Figure 3 is the same article separated into numbered sentences. Figure 4 is an example of the article summarized by 48%, via use of noun frequencies as described herein.

Figure 5 is an example of the article summarized by 75% by this method wherein the threshold of noun usage was set more stringently. Figure 6 is the matrix with which
the nouns and sentences were rated for automatic summarization. As will be appreciated from a comparative reading, the use of noun frequencies in this manner with text alone provides an advantageous shortening and is a desirable embodiment useful for the literary arts.

The noun rating method shown in Figures 1 - 6 was applied to several audiovisual recordings wherein the video segments were simply linked to the audio segments. Shortening was carried out via characterization of text in the audio segments, in the case of the summarization of an audiovisual presentation, each sentence is a dialogue segment, and desirably may be tagged with a segment marker, associating it with a table of segment attributes, including but not limited to the name of the character speaking the line, the line's beginning and ending timecodes, the components contained in that line, and the ratings of said components. Said components can be defined as Characters, Persons, Objects, Concepts, Events, Races, items, or Nouns. Timecode may be represented in hours, minutes, seconds, and/or frames. A value of 0 Hours, 0 Minutes, 0 Seconds, and 0 Frames (example, shown as 0:00:00:00) may be assigned to the first frame, or beginning, of an audiovisual presentation. Typically timecode progresses linearly to the end of an audiovisual presentation. A position or moment of the audiovisual presentation can be represented in this way to the accuracy of one Frame, by a given Timecode Value.

**Analysis Methods and Resources**

The determination of which character speaks a dialogue line may be carried out by computer recognition of audio and/or video information, or may be done manually. The software may employ voice pattern recognition to distinguish between different speakers. At times the software may not be able to accurately identify the speakers and may be confused by similar sounding speakers, in this case the software will make other logical deductions based on the known character and dialogue statistics to calculate the probability as to which character is speaking. Also, the software may employ a visual "face-recognition" algorithm that identifies on-screen characters by proportional measurements of and between their facial features. This will assist in the identification of which character is speaking. "Face-recognition" also can assist in calculating the "shot significance value" in non-dialogue segments. in an
embodiment, machine-learning methods, for example U.S. 5,819,247, and methods for boosting the performance of machine-Seaming classifiers, as in U.S. 7,024,033, can be employed in conjunction with image analysis systems, as referenced in U.S. 5,479,575, to identify objects and characters contained in various shot segments.

For the purpose of character identification, a database can be created, stored either locally or on-line, said database archiving relevant data about actors, celebrities, politicians, and other prominent or famous individuals, with pre-computed process-response statistical models, sample photos of various angles, voice imprint samples, resumes with character names. Data sets may also archive Film and Television credits, cross-referencing actors names with character names.

Another embodiment provides a system having a summarization program which for the purpose of character identification, can access a database, stored either locally or remotely, that archives relevant data about actors, celebrities, politicians, and other prominent or famous individuals, with pre-computed process-response statistical models, sample photos of various angles, voice imprint samples, and resumes with character names. Data sets may also archive Film and Television credits, cross-referencing actors names with character names.

Another embodiment provides a system that analyses media by a Disk Crawler program, in the background of computer functions or while the computer is otherwise idle. The functionality of the Disk Crawler program is prioritized by user input or stored user preference data. The purpose of the program is to perform the potentially lengthy and processor greedy analysis process in advance. This will enhance user experience by reducing wait time.

A most desirable embodiment provides a system, wherein a summarization program accesses a data base of pre-analyzed audiovisual presentation data, therefore bypassing the potentially lengthy and processor greedy analysis process. The actual customization and summary procedure that follows the analysis requires relatively simple calculations. This will not only save time, but will open the custom summarization market to smaller, cheaper, and less powerful processors, such as can be feasibly manufactured in cell phones and other portable devices. In this
embodiment, the analysis function and the summarizing function may be split into two independently operating systems.

Yet another embodiment provides a method wherein pre-analyzed audiovisual presentation data is accessed containing all sound waveform information at multiple points of an audiovisual presentation with an available variance of +/- .04%. Once an absolute sound waveform reference is matched at these points, a global adjustment ratio can compensate for unknown factors of speed variance that may potentially exist between the locally stored audiovisual presentation file and the remote pre-analyzed audiovisual presentation data.

In another embodiment, the audio portion of an audiovisual presentation created in front of a live audience is examined for purpose of detecting audience reactions. Defected audience reactions are separated into segments and are tagged and rated as to their type, length, and intensity, said ratings used in further rating segments and Segment Components of the audiovisual presentation.

In yet another embodiment, each Segment Component is compared to a set of User Preference Components created manually by the user, automatically discovered about the user, or a combination thereof. Components thus identified are weighted more in value than Components identified from other stored sets based on a factor of user input.

The audio of modern motion pictures and television programs is typically mixed in multi-track configurations as illustrated in Figure 13. This can be utilized in an advantageous way in the analysis of various audio components. Since different types of audio information typically reside on these tracks in a standardized fashion, the separate observation of each track of sound will reveal more simple sound waveforms than fully combined mono or stereo mixes, providing for analysis with reduced sound component interference. For example, dialogue is more easily recognizable when not obscured by music and background noise, in a multi-track audio mix of an audio-visual presentation, the center track typically contains most of the dialogue and natural sound effects. In a musical number or music video the center track typically contains the main vocals. The left and right tracks typically contain music, background atmospheres, and large sound effects. The surround
tracks usually contain similar information to the left and right tracks, but at lower levels. The LFE track contains only low frequency information which serves to add weight to music and sound effects. The presence of information on the LFE track may be used to aid in the detection of action and music scenes. A stereo track may be passed through a matrix decoding system which will split any center-panned information to a separate center track and any out of phase information to a separate surround track, essentially splitting the stereo pair of audio tracks into four audio tracks providing for easier analysis.

A well-known audio function called "Time Compression" is incorporated into most audio manipulation software, such as U.S. No. 5,883,804. Time Compression shortens the time of the audio file (or makes it faster) without raising the pitch. This is accomplished by breaking the sound file into tiny equal length segments and removing some of the segments in an alternating or otherwise repeating pattern. It can also be accomplished by speeding up the audio file, then lowering its pitch to the audio file's original reference pitch. Time compression can be performed on the output file of this invention to further shorten the summarized version.

**Time Ordered Assembly**

After segment characterization (which optionally may include segment modification, or fusion), the segments are combined to re-create a story that maintains at least some of the original unabridged audiovisual presentation's meaning. In an embodiment, the recombination occurs so that segments (or segment portions in the case of modified segments) are arranged in the summarized piece in their original time order.

Assembly into time order can be achieved by a number of techniques known to the skilled artisan. In an embodiment, undesired portions are removed from an unabridged audiovisual presentation. In an embodiment, segments from an audiovisual presentation are identified and stored in one or more files. Then, desired segments are taken from the one or more files and assembled into the final, shorter product. In an embodiment, segments of an unabridged audiovisual presentation are identified by determination of their starting and ending times (or, for example, starting time plus duration time) and making a table of this time information. A final
shortened audiovisual presentation can then be made by selecting segment times from the table and assembling desired segments in a desired order.

Typically, when making a shorter, summarized audiovisual presentation from an unabridged audiovisual presentation, the entire (unabridged) piece is first stored in electronic memory, such as a computer hard drive or other storage device. Desired segments, which optionally may be separated into distinct memory locations as separate (additions!) copies, are assembled in a time order. After or during assembly, the assemblage is exported into a digital storage medium such as a hard disk, flash-ram chip, DVD or streaming video internet files, as an integral, shorter piece.

In another embodiment, no separate copies of segments are stored outside of possible temporary file, and instead are assembled in a time order into a shorter piece by copying sections from the unabridged audiovisual presentation. In yet another embodiment, the entire unabridged audiovisual presentation (or at least most of it) is combined with information of segment location to make a shorter audiovisual presentation. In this latter embodiment a user can watch the unabridged audiovisual presentation, or portions of it, but also can watch an abridged form (or portions of it) by selecting the shorter version. Software that (for example) takes segment time information can be turned on or adjusted to provide the user either the long version, or a shorter version, including multiple shorter versions, as desired. A user, according to this embodiment can watch the unabridged version and interrupt such playback to watch instead a shorter version.

A customized summarization may be performed for individual users by the manipulation of various parameters and by reference of a personal preference database weighting the component significance values in various ways. This set of stored preferences may have been created manually by the user, automatically discovered about the user, or a combination of these methods.

Examples of user preferences may include negative or positive values for keywords or concepts, such as sex, violence, car crashes, swear words, body parts, kissing, face punches, shoes, hair, etc. In an embodiment, certain words when detected, such as swear words or names of former lovers, can be removed, beeped, or
reversed. Also, extreme negative values for keywords such as breasts, sex, and drug use, could be used to essentially clean up content for the consumption of children, old ladies, and religious people.

**Significance Curve Overlays**

A separate evaluation of audio and video segments and other types of segments could provide **Significance Curve Overlays**, which can be plotted on a two-axis graph representing **Significance Ratings** over time. (See Figs 1G&1) in an embodiment, said **Significance Curve Overlays** could be used to weight the primary method of segment rating. For example, during dialogue sections of an audiovisual presentation, the dialogue segments can be used as the primary method of segment rating while the **Significance Curve Overlays** for **Shots** and **Music** can be used merely as a weighting parameter. Examples of **Significance Curve Overlays** can include: **Dialogue Segments**, **Shot Segments**, Shot Length, **Music Segments** (volume, tempo, type), presence of transient sounds, volume (overall), audience reactions, emotions (various types), etc.

In yet another embodiment, all audio and visual segments are distinguished by their type (DiaSogue Segments, Music Segments, Audience Reaction Segments, Shot Segments, etc.). All data pertaining to Parameters and Components of said Segments will be tabled in a Segment Attribute Summary (Figs. 7&8). In an embodiment, the data of Segment Attribute Summaries are plotted as a **Significance Curve Overlay** on a two-axis graph representing **Significance Ratings** over time. (See Figs, 10&11) The **Significance Curve Overlays** optionally are used to weight the primary method of segment rating and to visualize the interaction between the various segment attributes in a time line.

The priority level of each segment type at a particular point in a time Sine will determine the weight of the significance of its value in relationship with other segment type values at said point. This priority level will be incorporated into a formula that generates a final value for that segment type for that moment.

In a preferred embodiment the entire audiovisual presentation can be divided into **Frame Segments** (usually 24 to 30 per second). A **Frame Segment Summary** (See
Fig. 12) contains the value data from each segment type, as it exists during that Frame Segment of the audiovisual presentation. In this embodiment, each Frame Segment would be rated individually based on said value data contained therein. Any Frame Segment or Frame Segment Cluster falling below a user inputted Frame Significance Value Threshold would be skipped.

Frame Segment Ousters are groups of adjacent Frame Segments. The boundaries of said dusters may be determined by their relationship to the primary and secondary segment types of that moment, in lengthy audiovisual presentations, auto-summarization is likely to produce such clusters. Any "small" duster of non-significant Frame Segments surrounded by two "large" significant Frame Segments may be included in the summarized audiovisual presentation as if they were rated as significant. Likewise, any "small" cluster of significant Frame Segments surrounded by two large \* non-significant Frame Segment Clusters may be omitted from the summarized audiovisual presentation as if they were rated as non-significant. The relative size of the aforementioned "small" and "large" clusters will be determined by a formula based on the total number of Frame Segments in the over-all full-length audiovisual presentation.

In an embodiment, Segment Clusters can be managed by the use of multiple Secondary Significance Thresholds above and below the Primary Significance Threshold, as illustrated in Figure 15. Segment values will fall into zones divided by the multipie Significance Thresholds. The zones may have negative and positive values based on their proximity to the Primary Significance Threshold. When Segment Clusters are omitted or included based on greater amounts of surrounding Segment Clusters, an attempt may be made to equalize the numbers of the segments artificially omitted and included based on opposing positive and negative zone values.

In a most desirable embodiment, Segment Clusters can be managed through a process of contour interpolation as illustrated in Figure 14. In a 2-axis plotting of final segment attributes, the graph will resemble a cross-section of mountains and valleys of land. The contours and slopes of the line connecting the plotted segment values may be traced by a program, which will resist a slope deviance greater than 90 degrees, particularly when a connected leg crosses the Primary Significance
Threshold from above or below. The resulting contour sine will effectively be smoothed, removing sharp peaks and crevices, reducing content relationship adjacency artifacts.

**Smoothing of Segment Boundaries**

Certain artifacts can occur in the output file of this automatic procedure, such as jump cuts, choppy sound edits, or concepts and visuals that lack sufficient introduction. Higher percentage of auto-summarization will produce higher numbers of artifacts. With time, users will surely become accustomed to consuming greedily summarized media and may forgive many of these artifacts. Even so, every attempt should be made to assure a more pleasurable consumption of summarized media.

Many of the issues of concept introduction and content relationship adjacency can be resolved by accurate segment boundary detection and the discovery of proper formulaic relationships between the various overlapping segment types.

General audio edits may be smoothed with simple cross-dissolves of various lengths. The smoothness of audio edits within rhythmic Music Segments can be improved by a system that analyzes music to detect musical beats, as in U.S. 7,132,595, the contents of which are incorporated by reference in their entireties.

In many cases, summarization requires that audio segments are skipped within the boundaries of a visual Shot Segments. This will most likely result in what is known in the film industry as a "jump-cut". In an embodiment, one proposed remedy is to perform a short cross-dissolve of the out-going to in-coming Shot Segment portions. In another preferred embodiment, alternating Shot Segment portions can be selected for a "digital zoom-in". The overall change in the shot framing will be perceived by the viewer of the summarized audiovisual presentation to be a different shot, therefore effectively reducing the jump-cut effect. In some cases, a combination of these techniques can be used.

**Other Considerations**

Embodiments described can be applied in various ways to summarize and/or condense an audiovisual presentation such as a feature film, documentary,
television program, or instructional video. However, embodiments also can perform text and language summarization alone (without video information), in various ways. Embodiments can be implemented to distribute condensed and summarized text and/or audiovisual presentations by internet download, internet streaming, cable on-demand viewing, video iPod download, video cell phone download or streaming, DVD or video tape rental or sales, cable broadcast, satellite broadcast, television broadcast, AM radio broadcast, and FM radio broadcast.

Embodiments can be distributed through a consumer operated vending kiosk, through a local hard wired network, or wireless network such as thai based on Bluetooth technology, via a digital newspaper, via on-line newspaper or magazine, or other web-based sources of text oriented media. Embodiments can be applied in a reading machine for the blind or visually impaired, to provide audio summaries of documents. In an embodiment, software may be made available for download, or a service may be provided for the analysis and summarization of media.

Embodiments have been described in relation to software implementation, but may also be implemented with specialized hardware.

**Additional Information for Practicing Embodiments**

Embodiments described herein can utilize a large variety of known methods of text and language summarization. The contents of the following U.S. patents are referenced for this purpose and are intended to be part of the specification for the use in practicing the invention, without reliance on the references for construing the meaning of claim terms.

The methods and devices described by others as found in U.S. Pat. No. 7,132,595, entitled "BEAT ANALYSIS OF MUSICAL SIGNALS", which discloses a system that analyzes music to detect musical beats and to rectify beats that are out of sync with the actual beat phase of the music; U.S. Pat. No. 5,257,186, entitled "DIGITAL COMPUTING APPARATUS FOR PREPARING DOCUMENT TEXT", which discloses a device for and method of summarizing a document by calculating a significance value of each sentence of the document data according to the nature of the connective relations between successive sentences; U.S. Pat. No. 5,384,703,
entitled "METHOD AND APPARATUS FOR SUMMARIZING DOCUMENTS ACCORDING TO THEME", which discloses a device for and method of summarizing a document that includes the step of selecting segments in the document based on frequently occurring complex expressions; U.S. Pat. No. 5,638,543, entitled "METHOD AND APPARATUS FOR AUTOMATIC DOCUMENT SUMMARIZATION", which discloses a device for and method of summarizing a document that includes the step of scoring sentences based on the number of stop words and strings of connected stop words, U.S. Pat. Nos. 5,924,108 and 6,349,316, each entitled "DOCUMENT SUMMARIZER FOR WORD PROCESSORS", which disclose a method of summarizing a document that includes the step of scoring sentences by summing the frequency counts of content words in a sentence divided by the total number of content words in the sentence; U.S. Pat. No. 5,978,820, entitled "TEXT SUMMARIZING METHOD AND SYSTEM", which discloses a device for and method of summarizing a document that includes the step of determining a number of attributes of the text such as the number of characters, the number of paragraphs, a specific character string, the frequency of occurrence of opinion sentences, imperative sentences, polite sentences, conversational sentences, and colloquial sentences; U.S. Pat. No. 6289,304, entitled "TEXT SUMMARIZATION USING PART-OF-SPEECH", which discloses a method of summarizing text by identifying "token breaks" and reassembling shortened sentences, U.S. Pat No. 8,338,034, entitled "METHOD, APPARATUS, AND COMPUTER PROGRAM PRODUCT FOR GENERATING A SUMMARY OF A DOCUMENT BASED ON COMMON EXPRESSIONS APPEARING IN THE DOCUMENT", which discloses a device for and method of summarizing a document that includes the step of looking up common expression information; U.S. Pat. No. 6,925,455, entitled "CREATING AUDIO-CENTRIC, IMAGE-CENTRIC, AND INTEGRATED AUDIO-VISUAL SUMMARIES", which discloses a method of creating a summary of a video program, the components of which are identified in accordance with a machine learning method which relies on previously-generated experience-based learning data to provide a probability that a given audio segment is suitable for inclusion in said summary; the same patent discloses an alternate method using a heuristics-based ranking of separated audio and video segments, resyncing selected segments utilizing a maximum-bipartite-matching approach, grouping image segments into a plurality of frame clusters based on a visual similarity and dynamic level, performing


Each cited reference is specifically incorporated by reference. Cited references show the level of the art and are intended to illustrate structures and methods that can be used as part of embodiments taught in this specification. The contents of the cited references are not intended for use in construing terms in the claims being.

Although embodiments have been described in relation to various implementations, together with modifications, variations, and extensions thereof, other implementations, modifications, variations, and extensions are within the scope of the invention. The claims therefore are not limited by the description contained herein or by the drawings.
Claims

1. A method for the summarization of an audiovisual presentation, comprising;
   a) providing an audiovisual presentation or a portion thereof, of at least 15 minutes in duration;
   b) characterizing audio portions of the audiovisual presentation as Dialogue Segments and Non-Dialogue Segments:
   c) optionally characterizing visual portions of the audiovisual presentation as Shot Segments;
   d) rating characterized segments by significance;
   e) using a significance value threshold to select segments that exceed the significance value threshold in their original time order, thereby forming a summarized version of the audiovisual presentation.

2. The method of claim 1, wherein at least 50% of the segments are rated in step c according to combined ratings of individual Segment Components contained in the segment.

3. The method of claim 2, wherein said Segment Components are defined as one or more of Characters, Persons, Objects, Concepts, Events, Places, Items, and Nouns.

4. The method of claim 2, wherein the rating of multiple individual Segment Components by significance is based on the frequency of the multiple Segment Components in the overall audiovisual presentation.

5. The method of claim 1 wherein each Dialogue Segment is a sentence associated with a corresponding visual segment of the audiovisual presentation.

6. A method for the automatic summarization of a text document, comprising:
   a) providing a storage medium to store an electronic copy of the text document;
b) using a Data Mining program to separate at least most of the text into Sentence Segments and Word Segments, and comparing each Word Segment in the document to a locally stored data base or remote on-line database, such as WorldNet,
c) assigning a Noun Significance Value to each noun in the text document based on the frequency of the noun's appearance;
d) assigning a Sentence Segment Significance Value to one or more sentences based on Noun Significance Values of one or more nouns contained in that sentence;
e) outputting a summarized document that lacks sentences segments that fall below a threshold of Sentence Segment Significance Value,

7. The method of claim 6, wherein a "Named Entity Tagger" recognizes and tags named entities for frequency rating and for pronoun antecedent matching.

8. The method of claim 6, wherein pronouns in the text document are identified by comparing words in the text document to a database of Pronoun Type Sets:, and wherein pronouns are matched to their antecedents from proximity of pronoun and antecedent in the structure of a given sentence.

9. The method of claim 8, wherein the Noun Significance Value of each identified antecedent is increased based on the frequency of identified pronouns that describe it.

10. The method of claim 8, wherein relative pronouns are ignored.

11. The method of claim 6, wherein Word Segments are compared to a database of user preference nouns and weighted more greatly in value for selecting segments than nouns identified from other stored sets.

12. The method of claim 11, wherein the user preference nouns are obtained from the user, automatically discovered about the user, or a combination thereof.
13. The method of claim 6, wherein *Word Segments* are looked up on one or more data bases of nouns from a field of expertise and wherein nouns identified as belonging to the reference data base are weighted more highly.

14. The method of claim 13, wherein the database is selected from the group consisting of a medical terminology word list, a legal word list, a music terminology word list, a film and television production terminology word list, a post-production terminology word list, a business terminology word list, and a religious terminology word list.

15. A system for the automatic summarization of an audiovisual presentation wherein the audiovisual presentation comprises multiple *Dialogue Segments*, *Non-Dialogue Segments*, *Non-Dialogue Segments with Music*, and *Shot Segments*; the system comprising:
   a) a storage medium for storing an electronic copy of the audiovisual presentation;
   b) a stored program that identifies the segments, evaluates and generate significance ratings for the audio and visual segments based on the presence of one or more *Segment Components* within the segments;
   c) stored and tagged memory locations for monitoring timecode locations and ratings of the segments as well as identification and significance ratings of the *Segment Components* they contain and,
   d) a stored program that filters out less-desirable segments based on a mathematical comparison of the *Significance Values* of the segments to a selectable *Segment Significance Value Threshold*.

16. The system of claim 15, wherein the *Segment Components* described in step c are defined as at least one of Characters, Persons, Objects, Concepts, Events, Places, Items, and Nouns.

17. The system of claim 15, wherein speech recognition is performed on *Dialogue Segments*, thereby creating time code referenced text transcripts of spoken dialogue in the audiovisual presentation.
18. The system of claim 15, wherein embedded text of closed caption subtitles is used to create at least part of the Dialogue Segments.

19. The system of claim 15, wherein embedded text is used for comparison to improve the accuracy of transcripts created by speech recognition.

20. The system of claim 15, further comprising an automatic method of text summarization, comprising the steps:
   i) providing a storage medium to store an electronic copy of the text document;
   ii) using a Data Mining program to separate at least most of the text into Sentence Segments and Word Segments, and comparing each Word Segment in the document to a locally stored data base or remote on-line data base, such as WordNet;
   iii) assigning a Noun Significance Value to the nouns in the text document based on the frequency of the noun's appearance,
   iv) assigning a Sentence Segment Significance Value to one or more sentences based on Noun Significance Values of nouns contained in that sentence;
   v) outputting a summarized document that Sacks sentence segments that fall below a threshold of Sentence Segment Significance Value.

21. The system of claim 15, further comprising the use of voice imprint recognition on Dialogue Segments to identify persons or characters speaking in Dialogue Segments for rating Character Significance.

22. The system of claim 15, further comprising the use of voice imprint recognition on Dialogue Segments to identify persons or characters speaking in Dialogue Segments for use in matching pronouns to antecedents.

23. The system of claim 15, further comprising the use of face recognition image analysis on Shot Segments in the identification of persons or characters speaking, for the purpose of Character Significance rating.
24. The system of claim 15, further comprising the use of face recognition image analysis on Shot Segments in the identification of persons or characters speaking, for the purpose of pronoun to antecedent matching,

25. The system of claim 15, further comprising shape and text recognition image analysis on Shot Segments to gather on-screen text, for the purpose of rating segments and Segment Components of the audiovisual presentation.

26. The system of claim 25, wherein the on-screen text is signage.

27. The method as described in any of claims 1 through 14, wherein segments are evaluated for signs of stress or conflict by the observation of at least one of volume of speech, tone of speech, shape of mouth and eyebrows.

28. The method of claim 27, further comprising a comparison of Dialogue Segment Components with Verb and Adjective sets for signs of negativity or stress.

29. The method as described in any of claims 1 through 14, wherein Conflict is weighted separately by a variable of user input.

30. The method as described in any of claims 1 through 14, wherein the audio portion of an audiovisual presentation that has been created in front of a live audience is further rated on the basis of audience reactions.

31. The method of claim 30, wherein detected audience reactions are separated into segments and are tagged and rated as to at least one of type, length, and intensity, and the generated ratings used in further rating segments and Segment Components of the audiovisual presentation.

32. The method as described in any of claims 1 through 14, wherein each Segment Component is compared to a set of User Preference Components created manually by the user, automatically discovered about the user, or a
combination thereof.

33. The method of claim 32, wherein identified Components are weighted more in value than Components identified from other stored sets based on a user input factor.

34. The method as described in any of claims 1 through 14, wherein all audio and visual segments are distinguished by type selected from the types; Dialogue Segments, Music Segments, Audience Reaction Segments, and Shot Segments.

35. The method as described in any of claims 1 through 14, wherein data pertaining to Parameters and Components of said Segments are tabled in a Segment Attribute Summary.

38. The method of claim 35, wherein the data of Segment Attribute Summaries are plotted as a Significance Curve Overlay on a two-axis graph representing Significance Ratings over time.

37. The method of claim 36, wherein the Significance Curve Overlays are used to weight the primary method of segment rating and to visualize the interaction between the various segment attributes in a time line.

38. The system as described in any of claims 15 through 26, comprising a summarization program for character identification, wherein the summarization program for the purpose of character identification can access a database, stored either locally or remotely, and which archives relevant data with pre-computed process-response statistical models.

39. The system as described in claim 38, comprising Data sets of archived Film and Tejβvision credits, and cross-referencing information that relates actor's names with character names.
40. The system as described in any of claims 15 through 20, that analyses media by a Disk Crawler program, as a background activity or while the computer is otherwise idle.

41. The system of claim 40, wherein the functionality of the Disk Crawler program is prioritized by user input or stored user preference data.

42. The System as described in any of claims 15 through 26, comprising a summarization program operating independently from the analyzing program that accesses a data base of pre-analyzed audiovisual presentation data, therefore bypassing potentially lengthy and processor greedy analysis process.

43. A portable hand held device comprising the system of claim 42.

44. The device of claim 43, wherein the device is a cell phone.

45. The method as described in any of claims 1 through 14, wherein pre-analyzed audiovisual presentation data is accessed containing all sound waveform information at multiple points of an audiovisual presentation with an available variance of +/- .04%.

46. The method of claim 45, wherein after an absolute sound waveform reference is matched at multiple points, a global adjustment ratio compensates for unknown factors of speed variance that may potentially exist between the locally stored audiovisual presentation file and the remote pre-analyzed audiovisual presentation data.

47. The method as described in any of claims 1 through 14, further comprising a procedure of visually smoothing audio segment deletions within visual segment boundaries, thereby reducing the "jump-cut" effect.

48. The method of claim 47, wherein alternating Shot Segment portions are selected for a 'digital zoom-in.'
A method of smoothing Segment Clusters through contour interpolation, comprising:
preparing a multiple axis plot of final segment attributes,
connecting plotted segment values to create lines; and
smoothing the lines to minimize content relationship adjacency artifacts.
Figure 1

"Noun Sets"

USER INPUTTED NOUNS
John Smith
Ring

Specialty
"Noun Sets"

LAW
MEDICINE
BUSINESS

COMMON NOUNS
Door
Hat
Milk

PRONOUNS
he, she, it, we, them, him, her, etc.

"FOUND NOUN SET"
Afghan convert 'arrives in Italy'

An Afghan man who escaped a possible death sentence for becoming a Christian has arrived in Italy where he has been granted asylum, says Italy's PM.

Afghan MPs had earlier demanded Abdul Rahman, 41, stay in the country.

"He is already in Italy. I think he arrived overnight," Silvio Berlusconi said on Wednesday, hours after his cabinet approved the asylum plea.

Mr Rahman was freed on Monday after being deemed mentally unfit to stand trial on a charge of apostasy.

Mr Rahman, who had been charged with rejecting Islam, had been held at a secret location since his release from Kabul's high security Pul-e-Charkhi prison.

Applications for political asylum in Italy normally take months to process, but Mr Berlusconi and several colleagues had said previously they favoured a quick decision in Mr Rahman's favour, says the BBC's David Willey in Rome.

'Pressure'

Suggestions he might be offered asylum have outraged politicians in Afghanistan.

I say that we are very glad to be able to welcome someone who has been so courageous
Silvio Berlusconi

The issue was discussed in the Afghan parliament on Wednesday, with almost all MPs in agreement that "his leaving Afghanistan must be prohibited", the AFP news agency reported.

Dr Assadullah Hymatyer, an MP from Logar province, told the BBC that parliament was planning to investigate the events that led to Mr Rahman's release.

"We will ask the judge to explain to us why he was released. In the beginning they said he was mentally fit. So why is he mentally unfit now?" he asked.

"If he is really mentally unfit, then that's a separate issue. But if not we will ask the judge why he allowed international pressure to influence him."

There had been an international outcry at the prospect of Mr Rahman being executed for his religious beliefs, but Afghan legislators said the decision to release him from trial for apostasy was "contrary to the laws in place in Afghanistan".

Ambiguities

Italy was among a number of countries which spoke out on Mr Rahman's behalf when news of his detention emerged.

Mr Berlusconi told Associated Press Television News: "I say that we are very glad to be able to welcome
someone who has been so courageous."

The case has highlighted ambiguities in Afghanistan's constitution over the interpretation of religious issues.

Conversion, or apostasy, is a crime under Afghanistan's Islamic law.

Mr Rahman, who converted 16 years ago while working as an aid worker for an international Christian group, was arrested after police discovered him with a Bible.

An ethnic Tajik originally from the Panjshir Valley, north of Kabul, Mr Rahman returned to Afghanistan a few years ago.

It is thought that he was denounced by relatives after returning to seek custody of his two daughters. His family alleged he forced them to read the Bible, something he has denied.

In an interview with the Italian newspaper La Repubblica, Mr Rahman said: "I have done nothing to repent, I respect Afghan law as I respect Islam. But I chose to become a Christian, for myself, for my soul. It is not an offence."

Story from BBC NEWS:
http://news.bbc.co.uk/go/pr/fr/-/2/hi/south_asia/4055748.stm

Published: 2006/03/29 16:55:16 GMT

© BBC NH/VI
FIGURE 3

Afghan convert "arrives in Italy"
Numbered

1. An Afghan man who escaped a possible death sentence for becoming a Christian has arrived in Italy where he has been granted asylum, says Italy's PM.

2. Afghan MPs had earlier demanded Abdul Rahman, 41, stay in the country.

3. "He is already in Italy. I think he arrived overnight," Silvio Berlusconi said on Wednesday, hours after his cabinet approved the asylum plea.

4. Mr Rahman was freed on Monday after being deemed mentally unfit to stand trial on a charge of apostasy.

5. Mr Rahman, who had been charged with rejecting Islam, had been held at a secret location since his release from Kabul's high security Pul-e-Charki prison.

6. Applications for political asylum in Italy normally take months to process, but Mr Berlusconi and several colleagues had said previously they favoured a quick decision in Mr Rahman's favour, says the BBC's David Willey in Rome.

7. Suggestions he might be offered asylum have outraged politicians in Afghanistan.

8. The issue was discussed in the Afghan parliament on Wednesday, with almost all MPs in agreement that "his leaving Afghanistan must be prohibited", the AFP news agency reported.

9. Dr Assadullah Hynatyar, an MP from Logar province, told the BBC that parliament was planning to investigate the events that led to Mr Rahman's release.

10. "We will ask the judge to explain to us why he was released. In the beginning they said he was mentally fit. So why is he mentally unfit now?" he asked.

11. "If he is really mentally unfit, then that's a separate issue.

12. But if not we will ask the judge why he allowed international pressure to influence him."

13. There had been an international outcry at the prospect of Mr Rahman being executed for his religious beliefs, but Afghan legislators said the decision to release him from trial for apostasy was "contrary to the laws in place in Afghanistan".

14. Italy was among a number of countries which spoke out on Mr Rahman's behalf when news of his detention emerged.
15. Mr Berlusconi told Associated Press Television News: "I say that we are very glad to be able to welcome someone who has been so courageous."

16. The case has highlighted ambiguities in Afghanistan's constitution over the interpretation of religious issues.

17. Conversion, or apostasy, is a crime under Afghanistan's Islamic law.

18. Mr Rahman, who converted 16 years ago while working as an aid worker for an international Christian group, was arrested after police discovered him with a Bible.

19. An ethnic Tajik originally from the Panjshir Valley, north of Kabul, Mr Rahman returned to Afghanistan a few years ago.

20. It is thought that he was denounced by relatives after returning to seek custody of his two daughters.

21. His family alleged he forced them to read the Bible, something he has denied.

22. In an interview with the Italian newspaper La Repubblica, Mr Rahman said: "I have done nothing to repent, I respect Afghan law as I respect Islam.

23. But I chose to become a Christian, for myself, for my soul.

24. It is not an offence."
Afghan convert “arrives in Italy”
Summarized by 48%

1. An Afghan man who escaped a possible death sentence for becoming a Christian has arrived in Italy where he has been granted asylum, says Italy’s PM.

3. “He is already in Italy. I think he arrived overnight,” Silvio Berlusconi said on Wednesday, hours after his cabinet approved the asylum plea.

6. Applications for political asylum in Italy normally take months to process, but Mr Berlusconi and several colleagues had said previously they favored a quick decision in Mr Rahman’s favor, says the BBC’s David Willey in Rome.

8. The issue was discussed in the Afghan parliament on Wednesday, with almost all MPs in agreement that “his leaving Afghanistan must be prohibited”, the AFP news agency reported.

9. Dr Assadullah Hymatyar, an MP from Logar province, told the BBC that parliament was planning to investigate the events that led to Mr Rahman’s release.

10. “We will ask the judge to explain to us why he was released. In the beginning they said he was mentally fit. So why is he mentally unfit now?” he asked.

13. There had been an international outcry at the prospect of Mr Rahman being executed for his religious beliefs, but Afghan legislators said the decision to release him from trial for apostasy was “contrary to the laws in place in Afghanistan”.

15. Mr Berlusconi told Associated Press Television News: “I say that we are very glad to be able to welcome someone who has been so courageous.”

18. Mr Rahman, who converted 16 years ago while working as an aid worker for an international Christian group, was arrested after police discovered him with a Bible.

20. It is thought that he was denounced by relatives after returning to seek custody of his two daughters.

21. His family alleged he forced them to read the Bible, something he has denied.

22. In an interview with the Italian newspaper La Repubblica, Mr Rahman said: “I have done nothing to repent, I respect Afghan law as I respect Islam.

23. But I chose to become a Christian, for myself, for my soul.
FIGURE 5

Afghan convert “arrives in Italy”
Summarized by 75%

1. An Afghan man who escaped a possible death sentence for becoming a Christian has arrived in Italy where he has been granted asylum, says Italy’s PM.

3. "He is already in Italy. I think he arrived overnight," Silvio Berlusconi said on Wednesday, hours after his cabinet approved the asylum plea.

13. There had been an international outcry at the prospect of Mr Rahman being executed for his religious beliefs, but Afghan legislators said the decision to release him from trial for apostasy was "contrary to the laws in place in Afghanistan".

18. Mr Rahman, who converted 16 years ago while working as an aid worker for an international Christian group, was arrested after police discovered him with a Bible.

21. His family alleged he forced them to read the Bible, something he has denied.

22. In an interview with the Italian newspaper La Repubblica, Mr Rahman said: "I have done nothing to repent, I respect Afghan law as I respect Islam."
Well, look what I found...
A dead man!
Wait, let me explain!
[phone ring]
Hey! [Gunshot]
FIGURE 8
Segment Attribute Summary
(Dialogue Segment)

Segment: 324 DX
Segment Significance Value: 112.36

<table>
<thead>
<tr>
<th>Character</th>
<th>Actor</th>
<th>Speaking</th>
<th>Spoken To</th>
<th>Character Significance Value</th>
<th>Conflict</th>
<th>Conflict Weighting</th>
<th>User Preference Weighting</th>
<th>Rating Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Jones</td>
<td>Jeff Simms II</td>
<td>1</td>
<td>1</td>
<td>75</td>
<td>1</td>
<td>15%</td>
<td>12%</td>
<td>31.75</td>
</tr>
<tr>
<td>Mary</td>
<td>Ashley Smith</td>
<td>1</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>15%</td>
<td>0%</td>
<td>23.85</td>
</tr>
<tr>
<td>Victor J</td>
<td>Jorge James</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0%</td>
<td>27%</td>
<td>20%</td>
<td>16.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noun</th>
<th>Mentioned</th>
<th>Noun Significance Value</th>
<th>Conflict</th>
<th>Conflict Weighting</th>
<th>User Preference Weighting</th>
<th>Rating Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>gun</td>
<td>1</td>
<td>27</td>
<td>1</td>
<td>15%</td>
<td>15%</td>
<td>35.1</td>
</tr>
</tbody>
</table>

Total Shot Significance: 107.01
Segment Length Weighting: 105%
Segment Significance Value: 112.36

Dialogue Line: Put down that damn gun, Mary! Are you idiots trying to kill me?
FIGURE 10

SECONDARY SIGNIFICANCE CURVE OVERLAY

VALUE
Figure 12

Frame Segment Summary

<table>
<thead>
<tr>
<th>Frame</th>
<th>129913</th>
</tr>
</thead>
<tbody>
<tr>
<td>timecode</td>
<td>1:12:14:28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PRIMARY SIGNIFICANCE</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURVE OVERLAYS</td>
<td></td>
</tr>
<tr>
<td>DIALOGUE SEG. CONTENT</td>
<td>67.01</td>
</tr>
<tr>
<td>SHOT SEG. CONTENT</td>
<td>55.73</td>
</tr>
<tr>
<td>AUDIENCE RECATION</td>
<td>46.22</td>
</tr>
<tr>
<td>Volume</td>
<td>75.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECONDARY SIGNIFICANCE</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURVE OVERLAYS</td>
<td></td>
</tr>
<tr>
<td>BRIGHTNESS</td>
<td>83.02</td>
</tr>
<tr>
<td>MUSIC VOLUME</td>
<td>62.95</td>
</tr>
<tr>
<td>OTHER NON-DIALOGUE VOL.</td>
<td>72.25</td>
</tr>
<tr>
<td>SHOT LENGTH</td>
<td>59.20</td>
</tr>
</tbody>
</table>

| Frame Sig. Value       |       |
| Total                 | 64.23 |
**FIGURE 13**

**TYPICAL MOTION PICTURE**
**AND TELEVISION**
**AUDIO TRACK LAYOUT**

<table>
<thead>
<tr>
<th>LEFT CHANNEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIGHT CHANNEL</td>
</tr>
<tr>
<td>CENTER CHANNEL</td>
</tr>
<tr>
<td>LFE CHANNEL</td>
</tr>
<tr>
<td>LEFT SURROUND CHANNEL</td>
</tr>
<tr>
<td>RIGHT SURROUND CHANNEL</td>
</tr>
</tbody>
</table>

5.1 AUDIO TRACK LAYOUT

---

**STEREO AUDIO TRACK LAYOUT**

![Diagram showing the conversion from stereo to 5.1 audio track layout]

- **LEFT CHANNEL**
- **RIGHT CHANNEL**

**DECODES TO:**

| LEFT CHANNEL |
| CENTER CHANNEL |
| RIGHT CHANNEL |
| SURROUND CHANNEL |

**FOUR CHANNEL LCRS AUDIO TRACK LAYOUT**