

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 January 2009 (22.01.2009)

PCT

(10) International Publication Number
WO 2009/011826 A2

(51) International Patent Classification:
H03G 3/30 (2006.01)

(21) International Application Number:
PCT/US2008/008591

(22) International Filing Date: 11 July 2008 (11.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/959,442 13 July 2007 (13.07.2007) US

(71) Applicant (for all designated States except US): **DOLBY LABORATORIES LICENSING CORPORATION** [US/US]; 100 Potrero Avenue, San Francisco, CA 94103-4813 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **SEEFELDT, Alan, Jeffrey** [US/US]; c/o Dolby Laboratories Licensing Corporation, 100 Potrero Avenue, San Francisco, CA 94103-4813 (US).

(74) Agents: **MENDENHALL, Larry et al.**; Gallagher & Lathrop, A Professional Corporation, 601 California Street, Suite 1111, San Francisco, Ca 94108-2805 (US).

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

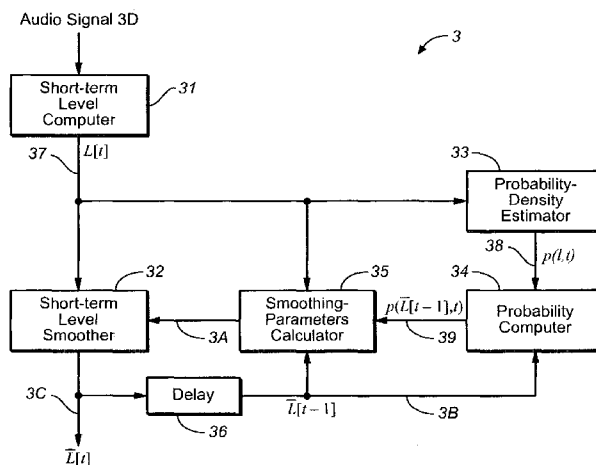
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: TIME-VARYING AUDIO-SIGNAL LEVEL USING A TIME-VARYING ESTIMATED PROBABILITY DENSITY OF THE LEVEL

FIG. 3



(57) Abstract: Methods, media and apparatus for smoothing a time-varying level of a signal. A method includes estimating a time-varying probability density of a short-term level of the signal and smoothing a level of the signal by using the probability density. The signal may be an audio signal. The short-term level and the smoothed level may be time series, each having current and previous time indices. Here, before the smoothing, computing a probability of the smoothed level at the previous time index may occur. Before the smoothing, calculating smoothing parameters using the probability density may occur. Calculating the smoothing parameters may include calculating the smoothing parameters using the smoothed level at the previous time index, the short-term level at the current time index and the probability of the smoothed level at the previous time index. Calculating the smoothing parameters may include calculating the smoothing parameters using breadth of the estimated probability density.

Description

Time-Varying Audio-Signal Level Using a Time-Varying Estimated Probability Density of the Level

Technical Field

5 The invention relates to audio signal processing. In particular, the invention relates to computing a time-varying measure of the level of an audio signal by smoothing the short-term level of the audio signal, where the parameters of the smoothing process are controlled, at least in part, by a time-varying estimate of the probability density of the short-term level.

10 ***References and Incorporation by Reference***

Certain techniques for measuring perceived (psychoacoustic) loudness useful in better understanding the present invention are described in published International patent application WO 2004/111994 A2, of Alan Jeffrey Seefeldt et al, published December 23, 2004, entitled “Method, Apparatus and Computer Program for Calculating and Adjusting the Perceived Loudness of an Audio Signal” and in “A New Objective Measure of Perceived Loudness” by Alan Seefeldt et al, Audio Engineering Society Convention Paper 6236, San Francisco, October 28, 2004. Said WO 2004/111994 A2 application and said paper are hereby incorporated by reference in their entirety.

Background Art

20 In audio signal processing, generating a time-varying measure of the audio signal level is often necessary. (Here the term “level” generically refers to a level measure such as peak level, rms level, loudness level, etc.) For example, a loudness meter may display a time-varying measure of an audio signal’s perceptual loudness, where this measure is smoothed significantly in order to indicate the average loudness over the past several

25 seconds. In another example, an Automatic Gain Control (AGC) process may compute a highly smoothed time-varying measure of an audio signal’s level and then use the resulting measure to generate a slowly varying gain which, when applied to the audio signal, automatically moves the average level of the audio closer to a desired target level.

- 2 -

In these two of many examples, the smoothed level measure may be computed by applying some form of smoothing filter to a short-term level measure. ("Short-term" means computed over a time interval significantly shorter than the interval over which the subsequent smoothing is effective.) For example, one might compute the rms level of the signal or the perceptual loudness level, as described in the WO 2004/111994 A2 application, over an interval of tens of milliseconds to generate the short-term level. The subsequent smoothing of this short-term level might then involve time constants on the order of several seconds. In the following discussion, this time-varying short-term level measure is represented as the signal $L[t]$, and the subsequently smoothed level measure is represented as $\bar{L}[t]$, where t represents the discrete time index.

Many different types of smoothing filters may be applied to $L[t]$ to generate $\bar{L}[t]$. One might use a finite impulse response (FIR) filter or a multi-pole infinite impulse response (IIR) filter. The particular filter employed is not critical. For illustrative purposes one may consider the commonly used fast-attack/slow-release single-pole IIR smoother. With such a filter, the smoothed level measure $\bar{L}[t]$ may be updated according to the equation:

$$\bar{L}[t] = \begin{cases} \alpha_{attack} \bar{L}[t-1] + (1 - \alpha_{attack}) L[t] & L[t] - \bar{L}[t-1] > 0 \\ \alpha_{release} \bar{L}[t-1] + (1 - \alpha_{release}) L[t] & L[t] - \bar{L}[t-1] \leq 0 \end{cases} \quad (1)$$

The smoothing coefficients α_{attack} and $\alpha_{release}$ may be chosen such that $\alpha_{attack} < \alpha_{release}$. This means that $\bar{L}[t]$ tracks $L[t]$ more quickly when $L[t]$ is increasing (attack) in comparison to when $L[t]$ is decreasing (release). For an AGC, one might, for example, choose α_{attack} corresponding to a time constant of one second and $\alpha_{release}$ corresponding to a time constant of four seconds. This way, $\bar{L}[t]$ varies quite slowly over time, and as a result, the corresponding gain which modifies the audio also varies slowly, thereby maintaining the short-term dynamics of the original audio. Problems may arise, however, when using such large time constants. Suppose that such an AGC is operating on the audio of a television set with the intent of maintaining a consistent average level across programming and across various channels. In such a situation, the content of the audio signal being processed by the AGC may abruptly change, when a channel is changed for

- 3 -

example, and the associated average level of the audio signal may therefore also abruptly change. With its large time constants, however, the AGC takes a considerable amount of time to converge to a new level and bring the modified level of the processed audio in line with the desired target level. During such adaptation time, a viewer of the television
 5 may perceive the level of the audio to be too loud or too soft. As a result, the viewer may quickly reach for the remote control to adjust the volume—only to find himself or herself fighting the AGC as it converges.

The prior art typically solves the problem just described using time constants that adapt based on the relative relationship of the short-term level $L[t]$ to the smoothed level
 10 $\bar{L}[t]$. For example, if the short-term level of the signal is significantly greater or less than the smoothed level as defined by some threshold boundaries around the smoothed level, then the smoothing operation switches to faster attack and/or release time constants, respectively, until the short-term level falls back within the threshold boundaries around the smoothed level. Subsequently, the system switches back to the original slower time
 15 constants. Equation 1 may be modified to implement this more sophisticated smoothing technique by including four cases rather than two:

$$\bar{L}[t] = \begin{cases} \alpha_{\text{attackFast}} \bar{L}[t-1] + (1 - \alpha_{\text{attackFast}}) L[t] & L[t] - \bar{L}[t] > \Delta L_{\text{fast}} \\ \alpha_{\text{attack}} \bar{L}[t-1] + (1 - \alpha_{\text{attack}}) L[t] & 0 < L[t] - \bar{L}[t] \leq \Delta L_{\text{fast}} \\ \alpha_{\text{release}} \bar{L}[t-1] + (1 - \alpha_{\text{release}}) L[t] & -\Delta L_{\text{fast}} \leq L[t] - \bar{L}[t] \leq 0 \\ \alpha_{\text{releaseFast}} \bar{L}[t-1] + (1 - \alpha_{\text{releaseFast}}) L[t] & L[t] - \bar{L}[t] < -\Delta L_{\text{fast}} \end{cases} \quad (2)$$

20

In Equation 2, $\alpha_{\text{attackFast}} < \alpha_{\text{attack}}$ and $\alpha_{\text{releaseFast}} < \alpha_{\text{release}}$ meaning that $\alpha_{\text{attackFast}}$ and $\alpha_{\text{releaseFast}}$ correspond to time constants faster than α_{attack} and α_{release} , respectively. If α_{attack} and α_{release} correspond to time constants of 1 and 4 seconds, respectively, then $\alpha_{\text{attackFast}}$ and $\alpha_{\text{releaseFast}}$ may be chosen, for example, corresponding to time constants of
 25 0.1 and 0.4 seconds, respectively (10 times faster). The fast time constant threshold ΔL_{fast} must be chosen judiciously so that switching to these faster time constants does not occur too often and result in unwanted instability of the smoothed level $\bar{L}[t]$. If, for

- 4 -

example, the level measures $L[t]$ and $\bar{L}[t]$ represent rms level in units of decibels, one might set ΔL_{fast} at 10 dB, an approximate doubling of perceived loudness.

Though an improvement over the smoothing in Equation 1, the smoothing of Equation 2 may still perform sub-optimally for many signals. In general, for any reasonable threshold ΔL_{fast} , signals may exist for which the original desired dynamics of the short-term level $L[t]$ fluctuate outside of the threshold boundaries around the average level $\bar{L}[t]$, thus causing the smoothing process to falsely switch into the fast attack or release mode.

To better understand cases for which the smoothing of Equation 2 performs as desired and for which it performs inadequately, one can imagine the distribution of the short-term level $L[t]$ over time. One may imagine this distribution as a time-varying probability density which predicts the probability of encountering any particular value of short-term level L within an interval of time around the current time index t . The duration of this interval should be commensurate with the slower set of time constants used in the smoothing filter of Equation 2.

Now consider the behavior of this probability density for the television-channel-change example described earlier. Assuming that the dynamic range of the short-term level for a given channel is somewhat limited, the probability density function of the short-term level $L[t]$ takes the form of a fairly narrow hump positioned around the smoothed level $\bar{L}[t]$. When the channel changes, and assuming that the average level of the new channel is significantly higher than the original, the probability density function will begin to change: the original hump decreases as a new hump grows positioned around the higher average level of the new channel.

FIG. 1 depicts such a probability density function at the beginning of the described transition. In this figure, the horizontal axis represents level and the vertical axis represents probability. The solid line represents the probability density of the short-term level at the beginning of the transition. One notes the decreasing hump on the left, representing the decreasing probability associated with the old channel selection, and the increasing hump on the right, representing the increasing probability associated with the new channel selection. At the beginning of the transition, the smoothed level $\bar{L}[t-1]$ still falls within the hump of the old channel selection, while the short-term level $L[t]$ falls within the hump of the new channel. In this diagram, the short-term level $L[t]$ is larger

- 5 -

than $\bar{L}[t-1]$ by an amount greater than ΔL_{fast} , and therefore, according to Equation 2, a fast time constant is used to update $\bar{L}[t]$ towards $L[t]$. This is the desired effect: the smoothed level $\bar{L}[t]$ adapts quickly to the higher level of the newly selected channel, traveling quickly across the gap dividing the two humps of the probability density.

5 FIG. 2 depicts the probability density of the short-term level for a very different audio signal. In this case, original dynamics of the signal are comparatively large, and therefore the hump of the probability density is spread quite wide. Such dynamics may be typical in a high-quality recording of jazz or classical music. Also in FIG. 2, the relationship between $\bar{L}[t-1]$ and $L[t]$ is exactly the same as in FIG. 1, but now both
10 values lie within the main hump of the probability density. Therefore, the switch to a fast time constant is not desired because this relation between $\bar{L}[t-1]$ and $L[t]$ is part of the typical dynamics of the signal. In this case, the smoothing described by Equation 2 is not appropriate.

Disclosure of the Invention

15 Herein are taught methods, media and apparatus for smoothing a time-varying level of a signal. A method includes estimating a time-varying probability density of a short-term level of the signal and smoothing a level of the signal by using the probability density. The signal may be an audio signal, while the smoothed level and the estimated probability density may vary over time at approximately the same rate.

20 The short-term level and the smoothed level may be time series, each having current and previous time indices. Here, before the smoothing, computing a probability of the smoothed level at the previous time index may occur. Before the smoothing, calculating smoothing parameters using the probability density may occur. Calculating the smoothing parameters may include calculating the smoothing parameters using the
25 smoothed level at the previous time index, the short-term level at the current time index and the probability of the smoothed level at the previous time index. Calculating the smoothing parameters may include calculating the smoothing parameters using breadth of the estimated probability density.

30 Calculating the smoothing parameters may include, when the smoothed level at the previous time index has no greater than a threshold probability, using first attack and release time constants, and otherwise, when the smoothed level at the previous time index

- 6 -

has greater than the threshold probability, using attack and release time constants, one of which is slower than its first-time-constants correspondent. Both of the used attack and release time constants may be slower than their corresponding first time constants.

The threshold probability may be variable. The variable threshold probability
5 may be a function of the estimated probability density.

Smoothing may include smoothing the short-term level by using the probability density.

Estimating the probability density may include fitting a parametric representation to past values of the short-term level. The fitting may include fitting a Gaussian mixture
10 model to the past values of the short-term level. Estimating the probability density may include computing a histogram of past values of the short-term level.

A computer-readable memory may contain a computer program for performing any one of the methods described herein.

A computer system may include a CPU, such a memory and a bus
15 communicatively coupling the CPU and the memory.

An audio-signal-level smoother may include a short-term-level computer for computing a short-term level of an audio signal, a level smoother for smoothing level of the audio signal using smoothing parameters, a probability-density estimator for
20 estimating a probability density of the audio signal, a probability computer for computing a probability of a delayed smoothed level, a smoothing-parameters calculator for calculating the smoothing parameters and a delay for delaying the smoothed level.

The probability-density estimator may include a probability-density estimator for estimating a probability density using a time series of the short-term level of the audio signal. The probability computer may include a probability computer for computing the
25 probability using the estimated probability density and the delayed smoothed level. The smoothing-parameters calculator may include a smoothing-parameters calculator for calculating smoothing parameters using the probability, the delayed smoothed level and the short-term level. The probability computer may include a probability computer for computing the probability using the estimated probability density and the delayed
30 smoothed level, and the smoothing-parameters calculator may include a smoothing-parameters calculator for calculating smoothing parameters using the probability, the delayed smoothed level and the short-term level.

The level smoother may include a level smoother for smoothing the short-term level of the audio signal.

- 7 -

The probability-density estimator may include a fitter for fitting a parametric representation to past values of the short-term level. The parametric-representation fitter may include a fitter for fitting a Gaussian mixture model to the past values of the short-term level. The probability-density estimator may include a histogram computer for
5 computing a histogram of past values of the short-term level.

Another method for smoothing a time-varying level of a signal may include receiving a signal having a short-term level and smoothing the time-varying level, where the short-term level and the smoothed time-varying level are time series, each having current and previous time indices, and the smoothing may include using slow time
10 constants where the difference between the smoothed short-term level at the previous time index and the short-term level at the current time index exceeds a threshold directing the use of fast time constants.

The various features of the present invention and its preferred embodiments may be better understood by referring to the following discussion and the accompanying
15 drawings in which like reference numerals refer to like elements.

Description of the Drawings

FIG. 1 depicts the probability density of the short-term level of an audio signal for which switching to a fast-smoothing time constant is appropriate.

FIG. 2 depicts the probability density of the short-term level of an audio signal for
20 which switching to a fast-smoothing time constant is inappropriate.

FIG. 3 depicts an overview block diagram of an embodiment of the invention.

FIG. 4 depicts the operation of an embodiment of the invention on an audio signal that switches between sources of significantly different levels.

FIG. 5 is a flowchart of the operation of the smoothed-level generator, according
25 to one embodiment of the invention.

Best Mode for Carrying Out the Invention

FIG. 3 is a block diagram of a smoothed-level generator 3, according to one embodiment of the invention.

The smoothed-level generator 3 includes a short-term-level computer 31, a short-term level smoother 32, a probability-density estimator 33, a probability computer 34, a
30 smoothing-parameters calculator 35 and a delay 36. The short-term-level computer 31

- 8 -

receives an audio signal 3D as input and produces a short-term level $L[t]$ 37 as its output. The probability-density estimator 33 receives the short-term level $L[t]$ 37 as input and produces an estimated probability density $p(\ell, t)$ 38 as its output. The probability computer 34 receives the estimated probability density $p(\ell, t)$ 38 and the delayed smoothed level measure $\bar{L}[t-1]$ 3C as its input and produces the probability $p(\bar{L}[t-1], t)$ 39 as its output. The smoothing-parameters calculator 35 receives the probability $p(\bar{L}[t-1], t)$ 39 as its input, as well as the delayed smoothed short-term level 3B and the short-term level $L[t]$ 37, and produces the smoothing parameters 3A as its output. The short-term level smoother 32 receives the smoothing parameters 3A and the short-term level $L[t]$ 37 as its input and produces the smoothed level measure $\bar{L}[t]$ 3C as its output. Finally, the delay 36 receives and delays the smoothed level measure $\bar{L}[t]$ 3C.

Of course, the smoothed-level generator 3 may be implemented in whole or in part as a general-purpose computer including programmed memory, as application-specific circuitry, as a programmable gate array, etc.

The operation of the smoothed-level generator 3 according to one embodiment is described next.

The smoothed-level generator 3 generates a smoothed level estimate $\bar{L}[t]$ by smoothing a short-term level estimate $L[t]$, controlled, at least in part, by information derived from a time-varying probability density estimate of the short-term level $L[t]$. The resulting smoothed level estimate $\bar{L}[t]$ is useful, for example, as a value displayed by a real-time loudness meter or for use in the control path in an audio dynamics processor such as an AGC.

FIG. 5 is a flowchart of the operation of the smoothed-level generator 3, according to one embodiment of the invention. First, the short-term level $L[t]$ of an audio signal is computed, step 505. This short-term level is smoothed, step 510, generating the smoothed level measure $\bar{L}[t]$.

Concurrently, from the short-term level $L[t]$, a time-varying estimate of the probability density of the short-term level is updated, step 515. This probability density estimate, $p(\ell, t)$, returns a probability between zero and one for an input level ℓ . The time index t indicates that this function changes over time as the probability density estimator 33 (FIG. 3) receives new values of the short-term level $L[t]$. Ideally, the time

- 9 -

interval over which the probability density estimator 33 considers values of $L[t]$ for creating the estimate $p(\ell, t)$ should be commensurate with the time constants from which the smoother 32 generates $\bar{L}[t]$. As a result, $\bar{L}[t]$ and $p(\ell, t)$ vary over time at approximately the same rate.

- 5 A probability of the smoothed level measure from the previous time index, $p(\bar{L}[t-1], t)$, is computed, step 520, by feeding the density estimate $p(\ell, t)$ with the level value $\bar{L}[t-1]$. This probability value $p(\bar{L}[t-1], t)$ represents information about the relationship between $\bar{L}[t-1]$ and many values of the short-term level around the time index t — not only the relationship between $\bar{L}[t-1]$ and the current short-term level
- 10 $L[t]$. For example, if $p(\bar{L}[t-1], t)$ is high, this indicates that $\bar{L}[t-1]$ is similar to many values of the short-term level in the recent past. Conversely, if $p(\bar{L}[t-1], t)$ is low, then $\bar{L}[t-1]$ is different from most values of the short-term level in the recent past.

- Finally, the probability value $p(\bar{L}[t-1], t)$, as well as the current short-term level $L[t]$ and the previous smoothed level measure $\bar{L}[t-1]$, feed into the smoothing-
- 15 parameter calculator 35 (FIG. 3) that uses these values to select the smoothing parameters for generating the current smoothed level measure $\bar{L}[t]$, step 525.

- Although FIG. 3 depicts $p(\bar{L}[t-1], t)$ controlling the smoothing parameters, other information derived from the density estimate $p(\ell, t)$ could be employed. For example, the smoothing-parameters calculator 35 may use a measure of the breadth of the density
- 20 estimate, an indicator of the dynamic range of the audio signal.

- Returning to the illustrated use of $p(\bar{L}[t-1], t)$, generally, when the density estimate $p(\ell, t)$ estimates the previous smoothed value $\bar{L}[t-1]$ to have low probability, then fast time constants should update $\bar{L}[t]$. If the probability of $\bar{L}[t-1]$ is relatively high, then normal slower time constants should be used. This way, if the majority of the
- 25 short-term level values have moved far away from the previous smoothed value $\bar{L}[t-1]$, then $\bar{L}[t]$ will move more rapidly towards such values. If $\bar{L}[t-1]$ is already near the majority of the short-term level values, then slower movement is appropriate.

 With reference to FIG. 2, an embodiment of the invention prevents the undesired switch to faster time constants because $\bar{L}[t-1]$ lies within an area of high probability

- 10 -

even though the short-term level $L[t]$ is relatively far away. For the channel-change example in FIG. 1, it is worthwhile to examine in more detail how this strategy still produces the desired fast adaptation. To assist, FIG. 4 depicts the transition between the two channels in five stages. A simple time-constant control strategy assumes that when

5 $p(\bar{L}[t-1], t)$ is greater than some threshold p_T , slow time constants are used to update $\bar{L}[t-1]$. When $p(\bar{L}[t-1], t)$ is less than this threshold p_T , fast time constants are used. (The time index has been removed from the notation in the figure to simplify the expressions.)

The first stage of FIG. 4, the initial steady state, depicts conditions just before

10 selecting the second, louder channel. The probability density of the short-term level is distributed closely around the average level of the first channel. The smoothed level $\bar{L}$ lies within the main hump of this density. Therefore the probability of $\bar{L}$, $p(\bar{L})$, is greater than the threshold p_T , and slow time constants are used to update $\bar{L}$.

The second stage, the beginning of transition, represents the state of the system

15 just after the audio has switched to the louder second channel. In this stage, the large hump of the probability density positioned around the average level of the first channel shrinks, and a new hump positioned around the average level of the louder second channel grows. The smoothed level $\bar{L}$ has begun to move right towards the average level of the second channel. Its probability, however, is still greater than p_T , and therefore the

20 slower time constants are still employed.

In the third stage, the middle of transition, the first hump of the probability density has continued to shrink, and the second hump has continued to grow. The smoothed level $\bar{L}$ has continued growing to the right towards the second channel, but now its probability falls below the threshold p_T . ($\bar{L}$ has moved into the low-probability gap between the

25 two humps.) At this point, fast time constants are used to update $\bar{L}$, causing it to rapidly move across the gap towards the hump of the second channel.

In the fourth stage, the end of transition, $\bar{L}$ has moved into the main hump of the second channel, and its probability is again greater than p_T . A switch back to the slower time constants now occurs.

30 Finally, in the fifth stage, the new steady state, the smoothed level $\bar{L}$ has settled at the average level of the second channel, and slow time constants continue to be used.

- 11 -

By using information derived from the time-varying probability density estimate $p(\ell, t)$, this invention embodiment adapts the time constants for smoothing an audio signal level measure in a way that responds quickly during true periods of transition yet remains stable for previously problematic signals.

5 An example of a specific embodiment of the invention is now described in detail. In this embodiment, the smoothing generating $\bar{L}[t]$ from $L[t]$ may be expressed as a one-pole filter with a time-varying coefficient $\alpha[t]$:

$$\bar{L}[t] = \alpha[t]\bar{L}[t-1] + (1 - \alpha[t])L[t]$$

10 (3)

This smoothing is adapted by varying the coefficient $\alpha[t]$ as a function of the probability density estimate $p(\ell, t)$. The derivation of this density estimate $p(\ell, t)$ from the short-term level $L[t]$ is first described. Then control of the coefficient $\alpha[t]$ is described.

15

Estimation of the Probability Density

The probability density estimate $p(\ell, t)$ captures the behavior of the short-term level $L[t]$ over a time interval extending into the recent past from the current time index t . To do this, the estimator 33 (FIG. 3) maintains a buffer (not shown) containing the past
20 T values of the short-term level $L[t]$. (The length T of this buffer may be chosen according to the desired application. For the AGC example discussed earlier, one may, for example, set the length of the buffer equal to four seconds.)

From this buffer, one might generate an estimate of the probability density in any number of ways. For example, one might fit a parametric representation such as a
25 Gaussian mixture model to the data in the buffer. Another technique computes a histogram of the buffered data. Each value in the buffer is assigned to a discrete bin. The density function $p(\ell, t)$ is computed by first finding the bin into which the value ℓ falls and then computing the probability as the number of points in the data buffer assigned to this bin divided by the total number of points in the data buffer. The bins of the
30 histogram may be defined to lie between some minimum level $L_{\min}$ and maximum level $L_{\max}$ with a spacing of ΔL_{bin} . For example, if the level is represented in units of dB, one

- 12 -

might set $L_{\min} = 1$ dB, $L_{\max} = 120$ dB, and $\Delta L_{bin} = 1$ dB, to yield a histogram with 120 bins.

The total number of bins B may be expressed generally as:

$$B = \text{round}\left(\frac{L_{\max} - L_{\min}}{\Delta L_{bin}}\right) + 1$$

5 (4)

At each time instant t , the number of samples in the length T data buffer falling into each bin of the histogram is counted. $H_t[b]$, $b = 1 \dots B$, is the histogram bins, and the bin b to which a level value ℓ is assigned may be computed according to the mapping function:

10

$$\text{LevelToBin}(\ell) = \begin{cases} B & \ell \geq L_{\max} \\ \text{round}\left(\frac{\ell - L_{\min}}{\Delta L_{bin}}\right) & L_{\min} < \ell < L_{\max} \\ 1 & \ell < L_{\min} \end{cases}$$

(5)

Algorithmically, the computation of the histogram $H_t[b]$ at every time t occurs according

15 to the following pseudo code:

for $b = 1$ to B ,

$H_t[b] = 0$

end

20

for $\tau = 0$ to $T-1$,

$b = \text{LevelToBin}(L[t - \tau])$

$H_t[b] = H_t[b] + 1$

end

25

The probability density estimate $p(\ell, t)$ is then given by the histogram at the bin to which ℓ is assigned divided by the total number of data points in the buffer:

- 13 -

$$p(\ell, t) = \frac{H_i[LevelToBin(\ell)]}{T}$$

(6)

The result is an estimated probability density $p(\ell, t)$ that is not a continuous function of the input level ℓ but rather a function that “stair-steps” at an interval of ΔL_{bin} . If the histogram spacing ΔL_{bin} is small enough, however, the probability density estimate is smooth enough for practical purposes.

Calculation of the Smoothing Filter Coefficient

Calculation of the adaptive filter coefficient $\alpha[t]$ is now described. Under normal conditions, when the probability of $\bar{L}[t-1]$ is relatively high, normal attack and release time constants (as in Equation 1) are used. When the probability of $\bar{L}[t-1]$ is low, faster attack and release time constants are used.

To smoothly interpolate between normal and fast time constants, a control signal $\beta[t]$ ranging between zero and one is generated based on the probability of $\bar{L}[t-1]$ computed from the density estimate $p(\ell, t)$:

$$\beta[t] = \begin{cases} 1 & p(\bar{L}[t-1], t) > p_r \\ \frac{p(\bar{L}[t-1], t)}{p_r} & p(\bar{L}[t-1], t) \leq p_r \end{cases}$$

(7)

20

If $p(\bar{L}[t-1], t)$ is greater than a threshold p_r , then the control signal $\beta[t]$ is equal to one. As $p(\bar{L}[t-1], t)$ falls below p_r , the control signal $\beta[t]$ falls to zero. Finally, the smoothing filter coefficient $\alpha[t]$ is computed by interpolating between normal and fast attack and release coefficients based on the control signal $\beta[t]$.

25

$$\alpha[t] = \begin{cases} \beta[t]\alpha_{attack} + (1 - \beta[t])\alpha_{attackFast} & L[t] - \bar{L}[t-1] > 0 \\ \beta[t]\alpha_{release} + (1 - \beta[t])\alpha_{releaseFast} & L[t] - \bar{L}[t-1] \leq 0 \end{cases}$$

(8)

- 14 -

When $\beta[t]$ is equal to zero ($\bar{L}[t-1]$ lies in an area of low probability), the fast attack and release coefficients are used. When $\beta[t]$ is equal to one ($\bar{L}[t-1]$ lies in an area of high probability), then the normal attack and release coefficients are used.

For the implementation of an AGC, the values of α_{attack} , $\alpha_{attackFast}$, $\alpha_{release}$, and
 5 $\alpha_{releaseFast}$, may be set corresponding to time constants of 1, 0.1, 4, and 0.4 seconds, respectively. The value of the probability threshold p_T may be set to a fixed value of 0.075.

Though this embodiment uses a fixed value of the threshold p_T it may also be advantageous in certain applications to make p_T a variable threshold which changes, for
 10 example, as a function of some measure of the density estimate $p(\ell, t)$. For example, p_T might vary with the maximum of $p(\ell, t)$ across level ℓ .

This specification incorporates by reference all publications and patent applications mentioned herein, to the same extent if the specification had specifically and individually incorporated by reference each such individual publication or patent
 15 application.

As this invention may be embodied in several forms without departing from the spirit of essential characteristics thereof, the present embodiment is therefore illustrative and not restrictive. Changes that fall within the metes and bounds of the claims, or the equivalence of such metes and bounds thereof are therefore intended to be embraced by
 20 the claims.

- 15 -

Claims

1. A method for smoothing a time-varying level of a signal, the method comprising:
 - estimating a time-varying probability density of a short-term level of the
 - 5 signal; and
 - smoothing a level of the signal by using the probability density.
2. The method of claim 1 wherein
10 the signal is an audio signal.
3. The method of claim 1 wherein
the smoothed level and the estimated probability density vary over time at
approximately the same rate.
- 15 4. The method of claim 1 wherein
the short-term level and the smoothed level are time series, each having
current and previous time indices and
before the smoothing, the following is performed:
 - 20 computing a probability of the smoothed level at the
previous time index.
5. The method of claim 1 wherein
before the smoothing, the following is performed:
 - 25 calculating smoothing parameters using the probability
density.
6. The method of claim 5 wherein
the calculating of smoothing parameters comprises
30 calculating the smoothing parameters using the smoothed
level at the previous time index, the short-term level at the current
time index and the probability of the smoothed level at the
previous time index.

- 16 -

7. The method of claim 5 wherein

the calculating of smoothing parameters comprises

calculating the smoothing parameters using breadth of the
estimated probability density.

5

8. The method of claim 6 wherein

the calculating of smoothing parameters comprises

when the smoothed level at the previous time index
has no greater than a threshold probability, using first attack
and release time constants; and

10

otherwise, when the smoothed level at the previous
time index has greater than the threshold probability, using
attack and release time constants, one of which is slower
than its first-time-constants correspondent.

15

9. The method of claim 8 wherein the using of attack and release constants
comprises

using attack and release time constants, both of which are slower than their
first-time-constants correspondents.

20

10. The method of claim 6 wherein

the threshold probability is variable.

11. The method of claim 10 wherein

25

the variable threshold probability is a function of the estimated probability
density.

12. The method of any one of claims 1 - 11 wherein the smoothing comprises
smoothing the short-term level by using the probability density.

30

13. The method of any one of claims 1 - 11 wherein estimating the probability
density comprises

fitting a parametric representation to past values of the short-term level.

- 17 -

14. The method of claim 13 wherein the fitting comprises
fitting a Gaussian mixture model to the past values of the short-term level.

15. The method of any one of claims 1 - 11 wherein estimating the probability
5 density comprises
computing a histogram of past values of the short-term level.

16. A computer-readable memory containing a computer program for performing
any one of the methods of claims 1 - 11 .

10

17. A computer system comprising:
a CPU;
the memory of claim 16 ; and
a bus communicatively coupling the CPU and the memory.

15

18. A audio-signal-level smoother comprising:
a short-term-level computer for computing a short-term level of an audio
signal;
a level smoother for smoothing level of the audio signal using smoothing
20 parameters;
a probability-density estimator for estimating a probability density of the
audio signal;
a probability computer for computing a probability of a delayed smoothed
level;
25 a smoothing-parameters calculator for calculating the smoothing
parameters; and
a delay for delaying the smoothed level.

19. The audio-signal-level smoother of claim 18 wherein the probability-density
30 estimator comprises
a probability-density estimator for estimating a probability density using a
time series of the short-term level of the audio signal.

- 18 -

20. The audio-signal-level smoother of claim 18 wherein the probability computer comprises

a probability computer for computing the probability using the estimated probability density and the delayed smoothed level.

5

21. The audio-signal-level smoother of claim 18 wherein the smoothing-parameters calculator comprises

a smoothing-parameters calculator for calculating smoothing parameters using the probability, the delayed smoothed level and the short-term level.

10

22. The audio-signal-level smoother of claim 19 wherein the probability computer comprises

a probability computer for computing the probability using the estimated probability density and the delayed smoothed level;

15

and

the smoothing-parameters calculator comprises

a smoothing-parameters calculator for calculating smoothing parameters using the probability, the delayed smoothed level and the short-term level.

20

23. The audio-signal-level smoother of claims 18 - 22 wherein the level smoother comprises

a level smoother for smoothing the short-term level of the audio signal.

25

24. The audio-signal level smoother of claims 18 - 22 wherein the probability-density estimator comprises

a fitter for fitting a parametric representation to past values of the short-term level.

30

25. The audio-signal level smoother of claim 24 wherein the parametric-representation fitter comprises

a fitter for fitting a Gaussian mixture model to the past values of the short-term level.

- 19 -

26. The audio-signal level smoother of claims 18 - 22 wherein the probability-density estimator comprises

a histogram computer for computing a histogram of past values of the short-term level.

5

27. A method for smoothing a time-varying level of a signal, the method comprising:

receiving a signal having a short-term level; and
smoothing the time-varying level,

10 wherein

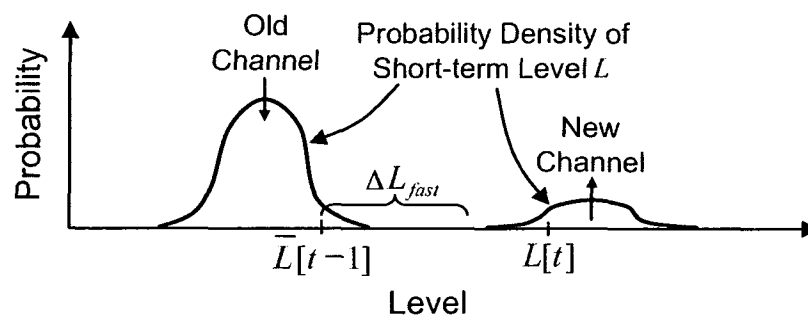
the short-term level and the smoothed time-varying level are time series, each having current and previous time indices and

the smoothing comprises

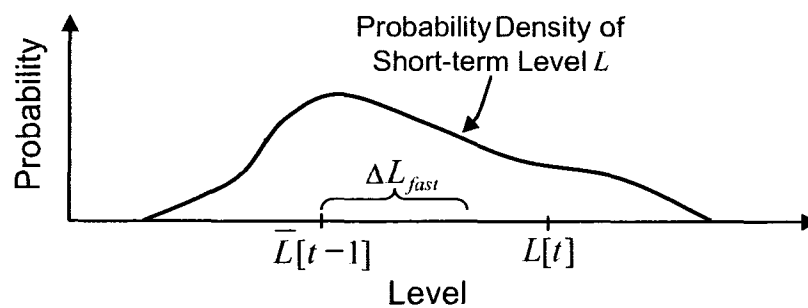
15 using slow time constants where the difference between the smoothed short-term level at the previous time index and the short-term level at the current time index exceeds a threshold directing the use of fast time constants.

1 / 4

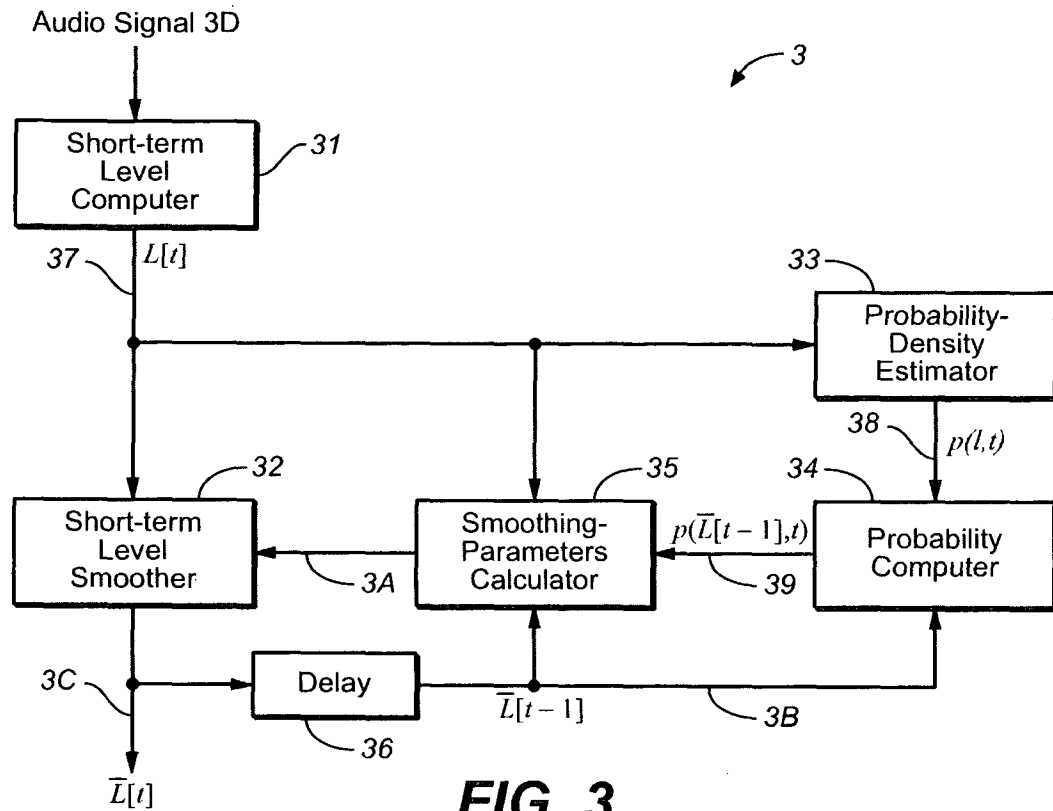
Fast Time Constant Appropriate

**FIG. 1**

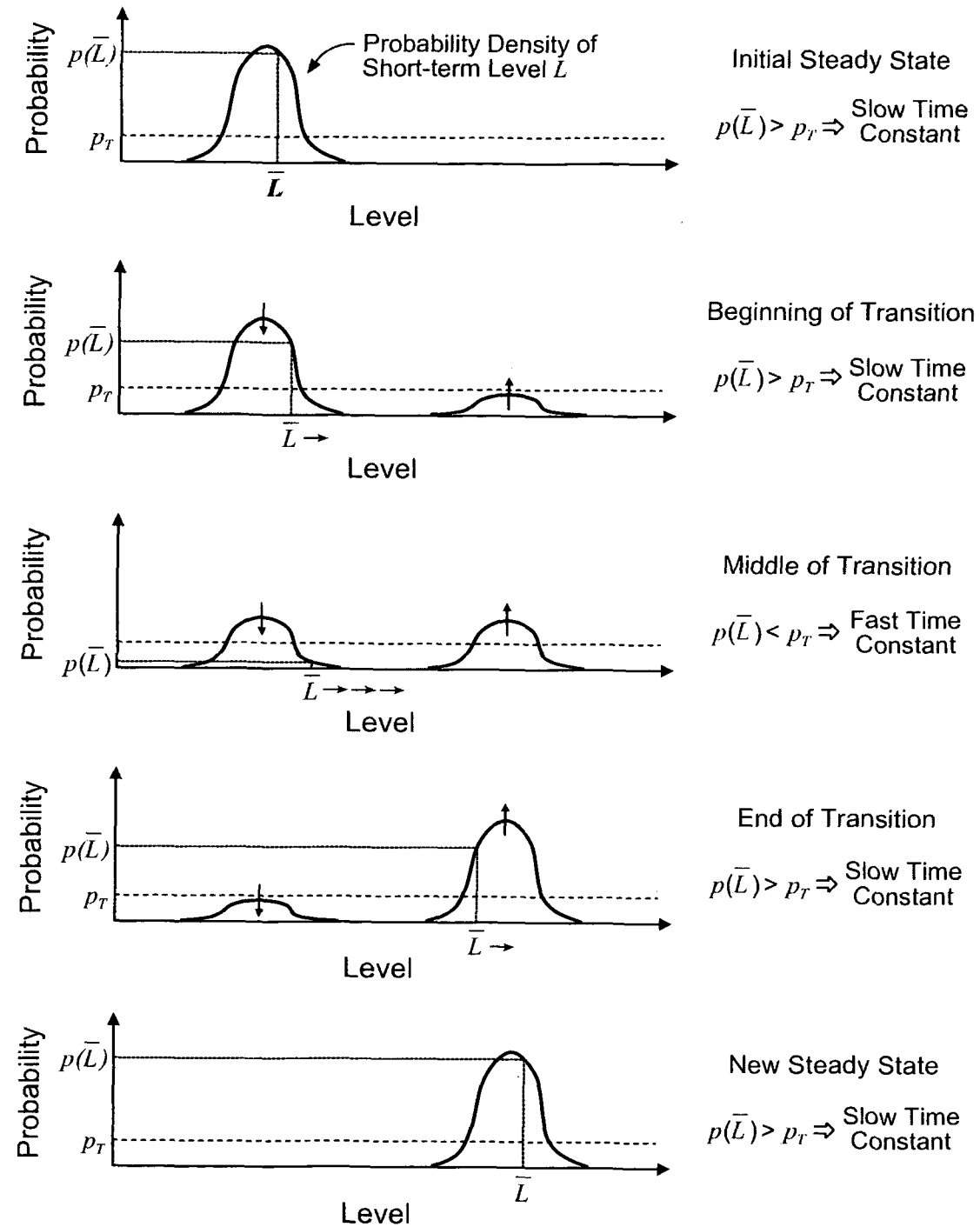
Fast Time Constant Inappropriate

**FIG. 2**

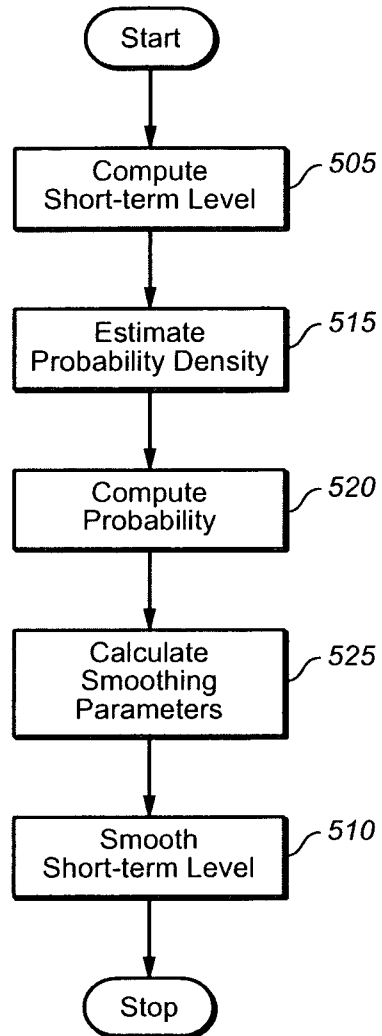
2 / 4

**FIG. 3**

3 / 4

**FIG. 4**

4 / 4

**FIG. 5**