



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI 0721733-1 A2**

(22) Data de Depósito: 14/06/2007
(43) Data da Publicação: 13/02/2013
(RPI 2197)



(51) *Int.Cl.:*
G06T 1/00

(54) **Título:** MÉTODO E APARELHO PARA DEFINIR UM LIMITE DE DETECÇÃO DADA UMA PROBABILIDADE FALSA DESEJADA

(73) **Titular(es):** Thomson Licensing

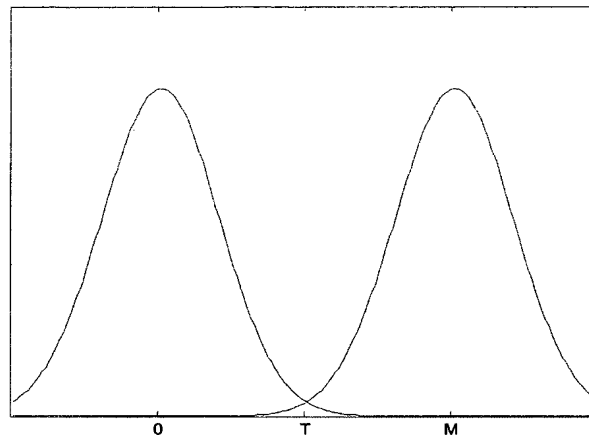
(72) **Inventor(es):** Jeffrey Adam Bloom, Jun Tian, Peter Georg Baum

(74) **Procurador(es):** NELLIE ANNE DAIEL-SHORES

(86) **Pedido Internacional:** PCT US2007014037 de 14/06/2007

(87) **Publicação Internacional:** WO 2008/153526de 18/12/2008

(57) **Resumo:** MÉTODO E APARELHO PARA DEFINIR UM LIMITE DE DETECÇÃO DADA UMA PROBABILIDADE FALSA DESEJADA. Um método e aparelho são descritos para a determinação de um limite, incluindo a seleção de uma probabilidade de falso resultado desejada, e a seleção de um valor limite a partir de uma tabela de limites correspondente à probabilidade de falso resultado desejada. São também descritos um método e aparelho para determinar um limite correspondente a uma probabilidade de falso resultado desejada, incluindo o cálculo de uma probabilidade de falso resultado, o cálculo de um valor de diferença entre a probabilidade de falso resultado calculada e a probabilidade de falso resultado desejada, a comparação do valor de diferença com um primeiro valor predeterminado e a comparação do valor de diferença com um segundo valor predeterminado, quando um resultado da primeira etapa de comparação é positivo.



“MÉTODO E APARELHO PARA DEFINIR UM LIMITE DE DETECÇÃO DADA UMA
PROBABILIDADE FALSA DESEJADA”

CAMPO DA INVENÇÃO

5 A presente invenção refere-se à definição de um limite de detecção e, em especial
à definição de um limite de detecção dada uma probabilidade de falso positivo desejada.

FUNDAMENTOS DA INVENÇÃO

10 É extremamente importante, em muitas aplicações de sinalização com marca d'á-
gua que a probabilidade de falso positivo de um detector de marca d'água fique abaixo de
um determinado valor. Um falso positivo ocorre quando o detector de marca d'água identifica
10 incorretamente um trabalho não sinalizado com marca d'água como sinalizado com marca
d'água. Um exemplo de um sistema no qual isso é importante é o uso de uma marca d'água
em um conteúdo de vídeo a fim de indicar a um gravador ou aparelho de gravação que o
material de vídeo não pode ser copiado. Quando um falso positivo ocorre durante uma gra-
vação, o dispositivo de gravação que monitora essa marca d'água concluirá incorretamente
15 que o conteúdo do vídeo não pode ser copiado e gravação será interrompida. Quando isso
ocorre durante a gravação de uma cerimônia de casamento, o fabricante da câmera enfren-
tará uma situação na qual os nubentes ficarão desolados. Quando isso acontece durante a
transmissão televisiva de um programa popular, os fabricantes do gravador enfrentarão mui-
tos clientes insatisfeitos. Assim, os fabricantes de gravador normalmente exigem que tal
20 marca d'água de controle de cópia tenha uma probabilidade de falso positivo próxima à pro-
babilidade da falha de um componente de hardware; na faixa dentre 1 erro em 10^9 detec-
ções para 1 erro em 10^{12} detecções.

Para melhor entender o problema da probabilidade de falso positivo, considere a i-
lustração da Figura 2a, que mostra duas distribuições. A distribuição à esquerda representa
25 os valores de detecção que podem ser esperados quando um detector de marca d'água é
aplicado a conteúdos que não contém marca d'água. A média dessa distribuição é zero e,
na maioria das vezes, espera-se que possa chegar a próximo de zero. Esta é uma curva de
probabilidade, para que a área sob a curva seja igual a 1,0. A distribuição à direita represen-
ta os valores de detecção que podem ser esperados quando o detector é aplicado a um con-
30 teúdo sinalizado com marca d'água. Neste caso, a média é M e o valor de detecção é ge-
ralmente perto de M. A função de um limite é distinguir entre amostras destas duas distribui-
ções. Considere o limite no ponto marcado T. Quando o valor de detecção excede T, pode-
se concluir que o conteúdo vem da distribuição à direita. Quando o valor de detecção é infe-
rior a T, pode-se concluir que o conteúdo vem da distribuição à esquerda. Como se pode
35 observar a partir da Figura 2a, existem algumas amostras da distribuição à direita que estão
abaixo do limite T. Essas são trabalhos marcados para os quais a detecção irá falhar, às
vezes chamados de falsos negativos. De maneira similar, existem algumas amostras da

distribuição à esquerda que excedem o limite. Estas são trabalhos não marcados que o detector irá rotular conforme marcados. Estes são os falsos positivos.

A probabilidade de falso negativo pode ser aperfeiçoada ao reduzir o limite. Esta ação, no entanto, simultaneamente, irá aumentar a probabilidade de uma indicação de falso positivo. No exemplo apresentado, as duas distribuições se sobrepõem. Isso significa que um limite que elimina os erros não poderá ser selecionado. Em muitas aplicações de sinalização com marca d'água, o limite de detecção T é selecionado com base em uma exigência de probabilidade de falso positivo específica à aplicação. Deve-se notar que a probabilidade de falso positivo é independente do algoritmo embutido na marca d'água. Ela é simplesmente a área sob a curva à esquerda, que é maior que o limite T.

A Figura 2b mostra uma vista mais próxima da curva à esquerda nas proximidades do limite. A área sob a curva maior que o limite T é mostrada sombreada. Esta área representa a probabilidade de falso positivo, a probabilidade de que uma marca d'água será detectada em um trabalho sem marcação.

Dada a exigência de uma probabilidade de falso positivo, seria vantajoso se ter um método e sistema para o estabelecimento de um limite mínimo que atenda à exigência da probabilidade de falso positivo.

Conforme usado no presente documento, "/" indica nomes alternativos para os componentes ou estruturas iguais ou semelhantes. Ou seja, uma "/" pode ser tomada no sentido de "ou", como usado no presente documento.

SUMÁRIO DA INVENÇÃO

A presente invenção resultou da tentativa de se resolver um problema no campo da sinalização com marca d'água e, em particular, no campo da sinalização digital com marca d'água de conteúdos de áudio e vídeo.

Esta probabilidade de falso positivo é bem entendida por detectores que usam uma medida de detecção de correlação normalizada. No entanto, em um esforço para aumentar a robustez da marca d'água, um número de pesquisadores propôs medidas de detecção que combinam múltiplas correlações normalizadas. Uma dessas medidas é a razão da maior correlação normalizada para a segunda maior quando um vetor extraído é comparado a múltiplos vetores de referência. Os presentes princípios se baseiam na análise da probabilidade de falso positivo da medida de detecção de razão de correlação sob uma suposição razoável sobre a distribuição de trabalhos não sinalizados com marca d'água.

Os pesquisadores investigaram este problema quando a medida de detecção é uma correlação normalizada. (M.L. Miller e J.A. Bloom, "Computing the Probability of False Watermark Detection", Proceedings of the Third International Workshop on Information Hiding, pp. 146-158, 1999). Acredita-se que ninguém tenha trabalhado sobre o problema quando a medida de detecção é a razão de duas correlações normalizadas.

Portanto, a presente invenção trata de um método e um sistema correspondente para configurar automaticamente o limite de detecção em um sistema de detecção de marca d'água quando a medida de detecção é a razão da maior correlação normalizada para a segunda maior correlação normalizada e uma probabilidade de falso positivo desejada é dada. Embora a presente invenção seja descrita em termos da determinação de um limite correspondente a uma probabilidade de falso positivo, esta não é assim limitada. O método e aparelho para a determinação de limites de acordo com a presente invenção podem ser usados para determinar um limite correspondente a qualquer probabilidade de falso resultado, por exemplo, uma probabilidade de falso negativo. Embora o método de determinação de limite de acordo com a presente invenção seja explicado em termos da detecção de marcas d'água, existem muitas outras áreas nas quais a determinação de limites é crítica, como na área da imagiologia médica, do registro de áudio ou imagem, do reconhecimento de áudio ou imagem, do reconhecimento de padrões, e da detecção, identificação e classificação de objetos. Na área do reconhecimento de padrões, incluindo a análise de assinatura de radar e sonar, a determinação de limites no cenário "identificar amigo ou inimigo" é um exemplo de utilização extremamente crítica da presente invenção. No cenário "identificar amigo ou inimigo", qualquer erro na identificação incorreta de amigo ou inimigo poderá ser mortal.

Um método e aparelho são descritos para a determinação de um limite, incluindo a seleção de uma probabilidade de falso resultado desejada e a seleção de um valor limite a partir de uma tabela de limites correspondente à probabilidade de um falso resultado desejada. São também descritos um método e aparelho para determinar um limite correspondente a uma probabilidade de falso resultado desejada, incluindo o cálculo de uma probabilidade de falso resultado, o cálculo de um valor de diferença entre a probabilidade de falso resultado calculada e a probabilidade de falso resultado desejada, a comparação do valor de diferença para um primeiro valor predeterminado valor e a comparação do valor de diferença com um segundo valor predeterminado quando o resultado da primeira etapa de comparação é positivo.

BREVE DESCRIÇÃO DOS DESENHOS

A presente invenção será melhor entendida a partir da descrição detalhada a seguir quando lida em conjunto com os desenhos em anexo. Os desenhos incluem as seguintes figuras brevemente descritas abaixo, nas quais números similares nas figuras representam elementos similares:

A Figura 1a ilustra a interpretação geométrica de uma fórmula analítica em que $m = 2$.

A Figura 1b ilustra a interpretação geométrica de uma fórmula analítica em que $m = 2$.

A Figura 1c mostra os resultados da comparação dos vetores sintéticos com as previsões utilizando a fórmula analítica.

A Figura 2a mostra duas distribuições - a probabilidade de valores de detecção para trabalhos marcados e não marcados.

5 A Figura 2b é uma vista mais próxima da distribuição de valores de detecção para trabalhos não marcados ilustrados na Figura 1 na área do limite, T.

A Figura 3a é um fluxograma de uma modalidade do método de determinação de limite de acordo com a presente invenção.

10 A Figura 3b é um diagrama em blocos do aparelho ou módulo de determinação de limite da presente invenção.

A Figura 4a é um fluxograma de uma modalidade do método de determinação de limite da presente invenção.

A Figura 4b é um diagrama em blocos do aparelho ou módulo de determinação de limite.

15 A Figura 5a é um fluxograma de uma modalidade alternativa do método de determinação de limite da presente invenção.

A Figura 5b é um diagrama em blocos de uma modalidade alternativa do aparelho de determinação de limite da presente invenção.

DESCRIÇÃO DETALHADA DAS MODALIDADES PREFERIDAS

20 Antes de o método e aparelho serem descritos, é necessário estabelecer uma base matemática. Uma questão crítica para muitas aplicações de sinalização com marca d'água é a probabilidade de que o detector de marca d'água identifique incorretamente um trabalho não marcado como sinalizado com marca d'água. Esta probabilidade de falso positivo é bem entendida por detectores que usam uma medida de detecção de correlação normalizada. No
25 entanto, em um esforço para aumentar a robustez da marca d'água, vários pesquisadores propuseram medidas de detecção que combinam múltiplas correlações normalizadas. Uma dessas medidas é a razão da maior correlação normalizada para a segunda maior quando um vetor extraído é comparado com múltiplos vetores de referência. Uma fórmula analítica
30 para a probabilidade de falso positivo é derivada e validada com dados empíricos. A correlação tem um teorema que relaciona o processamento do domínio espacial e o processamento do domínio de frequência através da Transformada de Fourier Discreta (DFT). Assim, as correlações podem ser realizadas nos domínios do tempo ou da frequência. A correlação é utilizada no sentido de determinar o grau de similaridade entre conjuntos de dados.

35 A sinalização com marca d'água digital é o processo de incorporação de uma mensagem invisível em um trabalho digital (como um áudio sem costura, uma imagem estática, ou uma sequência de imagens) com o objetivo de comunicar informações sobre esse trabalho. As aplicações incluem uma comunicação de direito autoral, a autenticação do conteúdo,

a dissuasão ou detecção de falsificação, um acompanhamento forense, ou um monitoramento de difusão.

Os erros são inevitáveis, mesmo nos mais bem desenhados sistemas de sinalização com marca d'água. Um erro de falso positivo ocorre quando um detector de marca d'água indica a presença de uma marca d'água em um trabalho não sinalizado com marca d'água. Tal erro pode levar à prevenção equivocada de uma operação legítima ou a denúncia de um cliente inocente. Assim, a falsa probabilidade tornou-se uma questão crítica para muitas aplicações de sinalização com marca d'água. Esta probabilidade depende do algoritmo de detecção de marca d'água, a maneira na qual o detector é usado, e a distribuição de trabalhos não sinalizados com marca d'água.

O problema de se analisar o comportamento da falsa detecção tem recebido pouca atenção na literatura relativa a marcas d'água. Um esquema de técnica anterior oferece um modelo para prever a probabilidade de falso positivo nos métodos de sinalização com marca d'água baseados na correlação e mostra que um espectro não branco de uma marca d'água faz com que o conteúdo da imagem interfira na detecção da marca d'água. Outro esquema da técnica anterior inclui a probabilidade de falsa detecção em sua estrutura no sentido de discutir os sistemas de sinalização com marca d'água. Ainda, um outro esquema da técnica anterior apresenta um método preciso para calcular a probabilidade de falso positivo ao utilizar um detector de correlação normalizada. Uma fórmula exata para a probabilidade de falso positivo é provida sob a suposição de que os vetores extraídos de trabalhos não sinalizados com marca d'água sejam extraídos a partir de uma distribuição radialmente simétrica. Um outro esquema da técnica anterior ainda estuda a probabilidade de falso positivo em exaustivas pesquisas geométricas. É mostrado que a imagem e a dependência de chave no detector de marca d'água resultam em uma diferente probabilidade de falso positivo para pesquisas geométricas.

Para muitas aplicações, um vetor extraído é comparado a um número de diferentes vetores de referência de marca d'água. Cada vetor de referência é associado a um símbolo de mensagem diferente. O símbolo de mensagem associado ao vetor de referência que tem a maior similaridade com o vetor extraído é relatado como o símbolo detectado. A certeza de detecção é o grau de similaridade. A medida de similaridade mais comum utilizado é a correlação normalizada. Ao usar a fórmula de um dos esquemas da técnica anterior, a exigência de probabilidade de falso positivo do aplicativo pode ser usada para definir um limite. Quando o valor de detecção excede o limite, o símbolo é relatado como presente, caso contrário, o detector não relata nenhum símbolo.

Em um esforço para melhorar a robustez de tais técnicas, alguns pesquisadores propuseram medidas de segurança que combinam o maior valor de correlação com o segundo maior valor de correlação. Uma dessas abordagens considera a diferença entre es-

ses dois valores como a segurança e outra considera a razão entre os dois valores como a segurança. Em ambos os casos, não existem análises de falsos positivos e, sem estas, não há como definir um limite e realizar uma comparação de robustez segura. Uma fórmula analítica é obtida aqui para a probabilidade de falso positivo para a medida da razão de correlação e os dados empíricos para apoiar esta análise são apresentados.

A probabilidade de uma detecção de marca d'água falsa é determinada pelo desenho do detector de marca d'água e a distribuição do conteúdo não sinalizado com marca d'água processado pelo mesmo. O algoritmo incorporado não é relevante para computar essa probabilidade, pois se presume que nenhuma marca d'água tenha sido incorporada.

Considere um detector de marca d'água de correlação normalizada específico, porém típico. O trabalho de entrada é primeiro processado para extrair um vetor de características n -dimensional, V . Exemplos de extração de marca d'água incluem várias combinações de transformadas de frequência, média de blocos, forma espectral, branqueamento e subamostragem. Estes processos são destinados a aumentar a robustez, aumentar a razão sinal para ruído, aumentar a eficiência de detecção, e/ou aplicar a distribuição assumida pela medida de detecção.

O vetor de características extraído V é então comparado com um conjunto de vetores de referência n -dimensionais $\{W_1, W_2, W_m\}$ $m \leq n$, para obter uma medição de detecção de razão de correlação D_v . Cada vetor de referência normalmente representa um dos símbolos de mensagem m e o símbolo associado ao vetor de referência mais similar ao vetor extraído é o símbolo relatado. Somente quando a medida de detecção excede um limite de detecção T , o detector relata uma detecção positiva. Caso contrário, nenhuma marca d'água é relatada como tendo sido detectada.

A fórmula exata para computar a medida de detecção é fundamental para determinar a probabilidade de falso positivo. A correlação normalizada é uma das técnicas mais comuns empregadas na medida de detecção. A correlação normalizada entre dois vetores n -dimensionais V e W_1 é

$$C_{V,W_1} = \frac{V \cdot W_1}{\sqrt{(V \cdot V)(W_1 \cdot W_1)}}$$

onde o símbolo \cdot é o produto interno. Duas outras medidas de detecção publicadas com base na correlação normalizada são a diferença entre o maior e a segundo maior correlação de magnitude e a razão entre a maior e a segunda maior correlação de magnitude.

Considere a medida de detecção de razão de correlação. O conjunto de $|C_{V,W_i}|$ é calculado para todos os vetores de referência m , W_i . Estes são, então, classificados do maior para o menor.

$$|C_{V,W_{i1}}| \geq |C_{V,W_{i2}}| \geq \dots \geq |C_{V,W_{im}}|$$

A medida de detecção de razão de correlação é então definido como:

$$D_V = \frac{|C_{V,W_{i1}}|}{C_{V,W_{i2}}},$$

e o símbolo de mensagem associado a W_{i1} é relatado, quando D_V ultrapassa o limite de detecção. Note que D_V é sempre maior ou igual a 1.

Em seguida, a medida de detecção de razão de correlação (1) é estudada e uma fórmula analítica é derivada para a probabilidade de falso positivo, ou seja, $P(D_V > T)$, a probabilidade de que $D_V > T$ para um trabalho não sinalizado com marca d'água, no qual T é o limite de detecção.

Considere o conjunto de vetores de referência (W_1, W_2, \dots, W_m) , onde $|W_i| = \sqrt{W_i \cdot W_i}$ e $W_i \cdot W_j = 0$ para $i \neq j$. Em outras palavras, o conjunto (W_1, W_2, \dots, W_m) é uma base ortonormal para um espaço de vetor m -dimensional. Quando $m < n$, este conjunto pode ser expandido para uma base ortonormal para o espaço vetorial n -dimensional, onde V e $\{W_i\}$ residem, ao acrescentar os vetores de comprimento unitário $n - m$ ao conjunto, onde cada um é ortogonal a todos os outros e aos de $\{W_i\}$. Uma base ortonormal de n dimensões é indicada como $\{W_1, W_2, \dots, W_m, e_{m+1}, \dots, e_n\}$.

Em álgebra linear, dois vetores em um espaço de produto interno é ortonormal se forem ortogonais e ambos de comprimento unitário. Os vetores ortogonais são perpendiculares uns aos outros. Ou seja, o seu produto interno é 0.

A projeção de dois vetores a partir de uma base ortonormal para outra representa uma rotação dos eixos e isso não muda o ângulo entre os vetores. Assim, a correlação normalizada entre dois vetores e a medida de detecção D_V são invariantes para transformações ortogonais. Considere a projeção de V sobre esta nova base $(W_1, W_2, \dots, W_m, e_{m+1}, e_{m+2}, \dots, e_n)$ e que os coeficientes neste espaço sejam indicados $\{v_1, v_2, \dots, v_n\}$. Neste caso, a correlação normalizada pode ser escrita

$$C_{V,W_i} = \frac{V \cdot W_i}{\sqrt{(V \cdot V)(W_i \cdot W_i)}} = \frac{v_i}{\sqrt{V \cdot V}}.$$

Uma vez que $\frac{1}{\sqrt{V \cdot V}}$ é um fator comum em todos os C_{V,W_i} para $i = 1, 2, \dots, m$, o mesmo pode ser abandonado e $|v_i|$ pode ser classificado diretamente,

$$|v_{i1}| \geq |v_{i2}| \geq \dots \geq |v_{im}|.$$

A medida de detecção de razão de correlação torna-se

$$D_V = \frac{|v_{i1}|}{v_{i2}} \quad (2)$$

e um evento de falso positivo $D_V > T$, é equivalente a

$$|v_{i1}| > T \cdot |v_j|, \text{ de todos } j \neq i1. \quad (3)$$

Se partirmos do pressuposto que cada $|v_i|$ tem a mesma probabilidade de ser o

maior (uma suposição razoável, dado que o caso no qual nenhuma marca d'água foi incorporada está sob investigação), neste caso a probabilidade de qualquer correlação normalizada ser grande o suficiente para causar um falso positivo é m vezes maior do que a probabilidade de V_1 causar um falso positivo. Isto é indicado como se segue:

$$5 \quad P_{fp} = P(D_V > T) = m \cdot P(D_V > T \mid il = 1). \quad (4)$$

A partir da equação (2), fica claro que um evento de falso positivo é uma invariante escalar. Ou seja, se para algum vetor de característica V , a sua medida de detecção D_V é maior que o limite T , neste caso para uma versão em escala $\alpha \cdot V = (\alpha \cdot V_1, \alpha \cdot V_2, \dots, \alpha \cdot v_n)$ onde α é uma constante diferente de zero, $D_{\alpha \cdot V}$ também é maior que T . Assim, V pode ser normalizado de tal modo que a sua norma 1 seja igual a $\sum_{i=1}^n |v_i| = 1$. Isto descreve um ponto no hiper-plano unitário $H_n = \{x_1, x_2, \dots, x_n \mid \sum_{i=1}^n |x_i| = 1\}$ (que na verdade é uma união de 2^n hiper-planos de dimensão $n - 1$). A distribuição de V no hiper-plano unitário H_n é importante para analisar o comportamento de detecção e, mais especificamente, a probabilidade de falso positivo.

15 Essa análise se aplica à utilização de características de detecção que produzem um vetor de característica normalizada V que, para trabalhos não sinalizados com marca d'água, pode ser aproximado por uma variável aleatória elaborada a partir de uma distribuição uniforme no hiper-plano unitário H_n . Como uma primeira hipótese, o vetor de característica normalizada V de um trabalho não sinalizado com marca d'água é uniformemente distribuído na hiper-plano unitário H_n . A hipótese 1, combinada com o fato de que um evento de falsos positivos é invariante escalar, sugere que a probabilidade de falso positivo $P(D_V > T)$ é igual à porção de vetores de característica normalizada no hiper-plano unitário H_n de tal modo que as suas medidas de detecção sejam maiores que o limite T ,

$$20 \quad P(D_V > T) = \frac{Area(V \in H_n, D_V > T)}{Area(H_n)}. \quad (5)$$

25 Um hiper-plano positivo COPIAR FÓRMULA e COPIAR FÓRMULA onde H_{n+} é um dos hiper-planos 2_n em H_n para o qual todas as coordenadas são não negativas, é definido. Neste caso, por simetria, a equação (5) é equivalente a

$$P(D_V > T) = \frac{Area(V \in H_n^+, D_V > T)}{Area(H_n^+)}. \quad (6)$$

30 A probabilidade de falso positivo na equação (6) é independente de n , a dimensão do espaço vetorial, onde V e W_1 residem. Uma prova disto segue, colocando um subscrito em P , V e H para indicar a dimensão, neste caso:

$$Area(V_{n+1} \in H_{n+1}^+, D_{V_{n+1}} > T) = Area\left(\sum_{i=1}^n v_i = 1 - v_{n+1}, v_{n+1} \geq 0, D_{V_n} > T\right),$$

onde $V_n = \{v_1, v_2, \dots, v_n\}$, e

$$Area(H_{n+1}^+) = Area\left(\sum_{i=1}^n v_i = 1 - v_{n+1}, v_{n+1} \geq 0\right).$$

Note que

$$Area\left(\sum_{i=1}^n (v_i = 1 - v_{n+1}, v_{n+1} \geq 0, D_V > T)\right) = \int_0^1 (1 - v_{n+1})^{n-1} \cdot Area(V_n \in H_n^+, D_{V_n} > T) dv_{n+1},$$

e

$$Area\sum_{i=1}^n (v_i = 1 - v_{n+1}, v_{n+1} \geq 0) = \int_0^1 (1 - v_{n+1})^{n-1} \cdot Area(H_n^+) dv_{n+1}.$$

5

Combinando todo o acima exposto, conclui-se que

$$\begin{aligned} P_{n+1}(D_{V_{n+1}} > T) &= \frac{Area(V_{n+1} \in H_{n+1}^+, D_{V_{n+1}} > T)}{Area(H_{n+1}^+)} \\ &= \frac{Area\left(\sum_{i=1}^n v_i = 1 - v_{n+1}, v_{n+1} \geq 0, D_{V_{n+1}} > T\right)}{Area\left(\sum_{i=1}^n v_i = 1 - v_{n+1}, v_{n+1} \geq 0\right)} \\ &= \frac{\int_0^1 (1 - v_{n+1})^{n-1} \cdot Area(V_n \in H_n^+, D_{V_n} > T) dv_{n+1}}{\int_0^1 (1 - v_{n+1})^{n-1} \cdot Area(H_n^+) dv_{n+1}} \\ &= \frac{Area(V_n \in H_n^+, D_{V_n} > T)}{Area(H_n^+)} \\ &= P_n(D_{V_n} > T). \end{aligned}$$

Assim, a probabilidade de falso positivo $P(D_V > T)$ é independente da dimensão n do vetor de característica. Em particular, $P(D_V > T)$ pode ser calculado ao definir $n = m$.

10

Uma fórmula analítica para $P(D_V > T)$ pode ser obtida através de uma interpretação geométrica da Equação 6, dada agora. Começando com $m = 2$, $H_2^+ = \{(x_1, x_2) | x_i > 0, \text{ e } x_1 + x_2 = 1\}$. Na Figura 1a, $O = (0, 0)$, $A = (1, 0)$, e $B = (0, 1)$. Todo o conteúdo não marcado produzirá uma detecção no segmento de linha AB . O segmento de laboratório AB é H_2^+ . O ponto $C = (c_1, c_2)$ onde $c_1 = T \cdot c_2$, representa o limite entre uma detecção positiva e uma negativa de detecção. Uma vez que todas as detecções positivas são detecções falsas, a probabilidade de um falso positivo é a razão do comprimento do segmento de linha AC para o todo, o comprimento do segmento de linha AB . Adicionar um ponto $D = (c_1, 0)$ sobre o segmento de linha AO . Esta é a projeção de C para x_1 . Então, uma vez que $|AO| = |AD| + |DO| = c_2 + c_1 = c_2 + Tc_2 =$

15

(T+1)c₂,

$$P(D_V > T \mid i|l = 1) = \frac{|AC|}{|AB|} = \frac{|AD|}{|AO|} = \frac{1}{T+1}.$$

Assim

$$P(D_V > T) = 2 \cdot P(D_V > T \mid i|l = 1) = \frac{2}{T+1}.$$

- 5 Quando $m = 3$, $H_3^+ = \{(x_1, x_2, x_3) \mid x_1 \geq 0, \text{ e } x_1 + x_2 + x_3 = 1\}$. Na Figura 1b, $O = (0, 0, 0)$, $A = (1, 0, 0)$, $B = (0, 1, 0)$ e $C = (0, 0, 1)$. O patch planar ABC é H_3^+ . O ponto $D = (d_1, d_2, 0)$ é um ponto no segmento de linha AB tal que $d_1 = T \cdot d_2$, e o ponto $E = (e_1, 0, e_3)$ é o ponto sobre o segmento de linha AC tal que $e_1 = T \cdot e_3$. O ponto F é a interseção de CD e o BE . Então:

$$\begin{aligned} P(D_V > T \mid i|l = 1) &= \frac{\text{Area}(ADFE)}{\text{Area}(ABC)} \\ &= \frac{\text{Area}(ADC) - \text{Area}(CEF)}{\text{Area}(ABC)} \\ &= \frac{\text{Area}(ADC)}{\text{Area}(ABC)} - \frac{\text{Area}(CEF)}{\text{Area}(BCE)} \cdot \frac{\text{Area}(BCE)}{\text{Area}(ABC)} \\ &= \frac{|AD|}{|AB|} - \frac{|EF|}{|EB|} \cdot \frac{|CE|}{|CA|} \\ &= \frac{1}{T+1} - \frac{1}{T+2} \cdot \frac{T}{T+1} \\ &= \frac{2}{(T+1)(T+2)}. \end{aligned}$$

10

$$\text{Então } P(D_V > T) = 3 \cdot P(D_V > T \mid i|l = 1) = \frac{6}{(T+1)(T+2)}.$$

Quando $m = 4$, $H_4^+ = \{(x_1, x_2, x_3, x_4) \mid x_1 \geq 0, \text{ e } x_1 + x_2 + x_3 + x_4 = 1\}$. Então:

$P(D_V > T \mid i|l = 1) =$

$$\begin{aligned} &\frac{\text{Area}(V \in H_4^+, v_1 > T \cdot v_2, v_1 > T \cdot v_3)}{\text{Area}(H_4^+)} - \frac{\text{Area}(V \in H_4^+, v_1 > T \cdot v_2, v_1 > T \cdot v_3, v_1 \leq T \cdot v_4)}{\text{Area}(H_4^+)} \\ &= \frac{2}{(T+1)(T+2)} - \frac{2 \cdot T}{(T+1)(T+2)(T+3)} \\ &= \frac{6}{(T+1)(T+2)(T+3)}. \end{aligned}$$

15

$$\text{Então } P(D_v > T) = \frac{24}{(T+1)(T+2)(T+3)}.$$

Continuando por dedução, pode-se obter uma fórmula analítica da probabilidade de falso positivo $P(D_v > T)$ para qualquer m , o que vem a ser:

$$P_{fp} = P(D_v > T) = \frac{m!}{\prod_{i=1}^{m-1} (T+i)}. \quad (7)$$

5 Para verificar se a fórmula da equação (7) está correta, os resultados preditos foram comparados com os resultados obtidos a partir de 100.000 vetores sintéticos extraídos de uma unidade de uma distribuição uniforme de hiper-plano unitária. A dimensão do vetor de característica é 1024 ($n = 1024$). Há 8 vetores ortogonais de marca d'água no detector ($m = 8$). A Figura 1c mostra os resultados da comparação dos vetores sintéticos contra as pre-
10 dições feitas pela equação (7). As predições da fórmula analítica corresponde aos resultados experimentais muito de perto abaixo a pelo menos $P_{fp} = 10^{-4}$

A Figura 3A é um fluxograma de uma modalidade do método de determinação de limite da presente invenção. Uma probabilidade de falso positivo desejada é aceita na etapa 305. Este requisito é específico ao ambiente de aplicação no qual a detecção de marca d'água ocorre. Um limite é determinado na etapa 310 e emitido para o detector de marca d'água na etapa 315. De acordo com a descrição acima, uma tabela de valores limites pode ser gerada tanto na inicialização ou em algum outro momento, antes ou durante a primeira de-
15 terminação de limite. A equação 7 acima é usada para gerar a tabela de valores limites (tabela de limites). Um gerador de tabela usa a equação 7, juntamente com um valor limite inicial ou mínimo, um limite máximo ou final e um tamanho de intervalo. A tabela de limites cria uma correspondência entre uma seleção de limites e as probabilidades de falso positivo que a mesma implica.

Para fins das modalidades da presente invenção descritas no presente documento, um valor limite inicial pode ser introduzido por um usuário ou selecionado. Quando selecio-
25 nado, um valor limite inicial possível é o limite inferior de uma probabilidade de falso negativo. Além disso, correspondentemente, para o valor limite final, quando selecionado, um valor limite final possível é um limite superior da dita probabilidade de falso positivo. Os incrementos ou intervalos limites utilizados nos cálculos são selecionados com base, entre outros fatores, no armazenamento disponível a fim de armazenar a tabela de limites ou na precisão
30 desejada dos cálculos.

A Figura 3b é um diagrama em blocos do aparelho ou módulo de determinação de limites da presente invenção. O aparelho ou módulo de determinação de limites aceita a probabilidade de falso positivo como entrada, determina o limite e transmite o limite para um detector de marca d'água.

35 A Figura 4a é um fluxograma de uma modalidade do método de determinação de

limite da presente invenção. Uma probabilidade de falso positivo desejada é aceita na etapa 405. O índice para a tabela de limites preparada na inicialização ou, de outra forma, no avanço é inicializado na etapa 410. A entrada de tabela é recuperada a partir da tabela com base no índice na etapa 415. Deve-se ter em mente que o índice pode não ser necessariamente inicializado para uma primeira entrada de tabela. Vários esquemas de pesquisa de tabelas são conhecidos no estado da técnica, incluindo um método binário, no qual o índice de tabela é inicializado aproximadamente no ponto intermediário da tabela. Isto é repetido de modo que toda a tabela não precise ser pesquisada. A entrada de tabela recuperada é comparada com a probabilidade de falso positivo desejada na etapa 420 e, se inferior à falsa probabilidade desejada, e não houver nenhuma outra entrada de limites superior ao limite recuperado e for menor que a probabilidade de falso positivo desejada, então o limite determinado é enviado para o gerador de marca d'água na etapa 425. Se puder existir entradas limites superiores à entrada de tabela de limites recuperada e também inferiores à probabilidade de falso positivo desejada, então o índice da tabela é ajustado na etapa 430.

A Figura 4b é um diagrama em blocos do aparelho ou módulo de determinação de limites 435, que inclui um módulo de pesquisa de tabela 440 e um comparador 445. O módulo de pesquisa de tabela recupera as entradas da tabela de limites a partir da tabela de limites previamente gerada usando um índice. O comparador 445 realiza a comparação e, se os critérios forem atendidos, emite o limite para o detector de marca d'água e, se os critérios não forem atendidos, em seguida, retorna para o módulo de pesquisa de tabela tendo ajustado o índice da tabela. O índice da tabela pode ser ajustado, quer pelo módulo de pesquisa de tabela ou pelo comparador ou por outro módulo (não mostrado) que executa essa função.

A Figura 5a é um fluxograma de uma modalidade alternativa do método de determinação de limite da presente invenção. Nesta modalidade, uma tabela de limites não é preparada com antecedência, mas o limite é determinado através de cálculos iterativos. Na etapa 505, a probabilidade de falso positivo desejada é recebida ou aceita como entrada. Um valor limite inicial é selecionado na etapa 510. Deve-se notar que o valor limite inicial pode também ser introduzido por um usuário. O limite corrente é inicializado para o valor limite inicial que foi selecionado ou introduzido na etapa 515. O valor limite corrente é usado para calcular a probabilidade de falso positivo na etapa 520. O valor da diferença entre a probabilidade de falso positivo desejada e a probabilidade de falso positivo calculada é calculado na etapa 525. O valor da diferença é comparado a 0 e a um valor predeterminado ϵ na etapa 530. Se o valor da diferença for maior que ou igual a 0 e menor que um valor predeterminado ϵ , então o sistema irá emitir o valor limite corrente, T, para um detector de marca d'água na etapa 540. Se o valor da diferença for negativo ou maior que o valor ϵ , então o limite corrente será modificado na etapa 535 de acordo com o sinal e a magnitude do valor

da diferença. O limite corrente modificado é utilizado para recalculer uma nova probabilidade de falso positivo usando o valor limite corrente. Esse processo continua até que o valor da diferença seja maior que ou igual a zero e abaixo do valor ϵ . Quando isso ocorre, o valor limite corrente, T , será enviado para um detector de marca d'água na etapa 540. O valor
5 limite corrente produz uma probabilidade de falso positivo, menor que a probabilidade de falso positivo necessária / desejada, de no máximo ϵ .

A Figura 5b é um diagrama em blocos de uma modalidade alternativa do aparelho de determinação de limite 545 da presente invenção, incluindo o módulo calculador de probabilidade de falso positivo 550 e o comparador 555. O calculador de probabilidade de falso
10 positivo 550 calcula um valor de probabilidade de falso positivo utilizando um valor limite (corrente) e também calcula a diferença entre um valor de falsa probabilidade desejada / necessária e o valor de probabilidade de falso positivo calculada. O comparador 555 realiza uma comparação entre o valor de diferença e 0 e também entre o valor de diferença e um valor predeterminado ϵ . Quando o valor da diferença é maior que ou igual a 0 e menor que
15 um valor predefinido ϵ , então o sistema enviará o valor limite corrente, T , para um detector de marca d'água. Quando o valor da diferença é negativo ou maior que o valor ϵ , então o limite corrente é modificado de acordo com o sinal e a magnitude do valor da diferença. O limite corrente modificado é usado para re-calculer uma nova probabilidade de falso positivo usando o valor limite corrente. Esse processo continua até que o valor da diferença seja
20 maior ou igual a zero ou abaixo do valor ϵ . Quando isso acontece, o valor limite corrente, T , será emitido para um detector de marca d'água. O valor limite corrente produz uma probabilidade de falso positivo menor que a probabilidade de falso positivo requerida / desejada, em no máximo ϵ .

É preciso entender que a presente invenção pode ser implementada em várias for-
25 mas de hardware (por exemplo, em um chip ASIC), software, firmware, processadores de uso especial, ou combinação dos mesmos, por exemplo, dentro de um servidor, um dispositivo intermediário (como um ponto de acesso sem fio ou um roteador sem fio) ou dispositivo móvel. De preferência, a presente invenção é implementada como uma combinação de hardware e software. Além disso, o software é de preferência implementado como um pro-
30 grama aplicativo tangivelmente incorporado em um dispositivo de armazenamento de programa. O programa aplicativo pode ser carregado para, e executado por, uma máquina compreendendo qualquer arquitetura adequada. De preferência, a máquina é implementada em uma plataforma de computador tendo um hardware, como, por exemplo, uma ou mais unidades de processamento central (CPU), uma memória de acesso aleatório (RAM), e in-
35 terfaces de entrada / saída (I/O). A plataforma de computador inclui ainda um sistema operacional ou código de micro-instruções. Os vários processos e funções descritos no presente documento podem fazer parte do código de micro-instrução ou parte do programa aplicativo

(ou uma combinação destes), executados através do sistema operacional. Além disso, vários outros dispositivos periféricos podem ser conectados à plataforma de computador, tal como um dispositivo de armazenamento de dados adicional e um dispositivo de impressão.

5 Deve-se ainda entender que, uma vez que alguns dos componentes de sistema
constituente e as etapas de método ilustradas nas figuras em anexo são de preferência im-
plementados em um software, as conexões em questão entre os componentes de sistema
(ou as etapas de processo) podem variar, dependendo da maneira na qual a presente in-
venção é programada. Dado os presentes ensinamentos, uma pessoa com habilidade sim-
10 ples na técnica relacionada será capaz de contemplar estas implementações ou configura-
ções, como também implementações ou configurações similares, da presente invenção.

REIVINDICAÇÕES

1. Método para determinar um limite, o dito método sendo **CARACTERIZADO** pelo fato de compreender as etapas de:

- selecionar uma probabilidade de falso resultado desejada; e

5 - selecionar um valor limite a partir de uma tabela de limites correspondente à dita probabilidade de falso resultado desejada.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de compreender ainda as etapas de:

- criar a dita tabela de limites;

10 - sendo que a dita etapa de criar compreende ainda:

- o cálculo de cada entrada de probabilidade de falso resultado na dita tabela de li-

mites usando $\frac{m!}{\prod_{i=1}^{m-1} (T = i)}$ e

- o armazenamento da dita probabilidade de falso resultado calculada na dita tabela de limites e do dito valor limite utilizado para calcular a dita entrada de probabilidade de falso resultado, quando T é o dito valor limite.

15

3. Método, de acordo com a reivindicação 2, **CARACTERIZADO** pelo fato de que a dita etapa de calcular é realizada utilizando um primeiro valor limite, um segundo valor limite, e um intervalo limite, incrementando iterativamente o dito primeiro valor limite por meio do dito intervalo limite e calculando cada dita entrada de probabilidade de falso resultado da dita tabela de limites até que o dito primeiro valor limite fique igual ao dito segundo valor limite.

20

4. Método, de acordo com a reivindicação 3, **CARACTERIZADO** pelo fato de que a dita probabilidade de falso resultado é uma probabilidade de falso positivo e ainda pelo fato de que o dito primeiro valor limite é um limite inferior de uma probabilidade de falso negativo, e o dito segundo valor limite segundo é um limite superior da dita probabilidade de falso positivo.

25

5. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a dita etapa de selecionar recupera o dito valor limite a partir da dita tabela de limites, sendo que o dito valor limite é um valor limite maior, tendo uma probabilidade de falso resultado menor que a dita probabilidade de falso resultado desejada.

30

6. Método para determinar um limite, o dito método sendo **CARACTERIZADO** pelo fato de compreender as etapas de:

- selecionar uma probabilidade de falso resultado desejada;

- calcular uma probabilidade de falso resultado;

35

- calcular um valor de diferença entre a dita probabilidade de falso resultado calculada e a dita probabilidade de falso resultado desejada;

- comparar o dito valor de diferença com um primeiro valor predeterminado; e
- comparar o dito valor de diferença com um segundo valor predeterminado quando um resultado da dita primeira etapa de comparar é positivo.

7. Método, de acordo com a reivindicação 6, **CHARACTERIZADO** pelo fato de que a dita probabilidade de falso resultado é uma probabilidade de falso positivo, e compreendendo ainda as etapas de:

- selecionar um primeiro valor limite; e
- modificar o dito primeiro valor limite quando os resultados de ambas as etapas de comparar não são positivos.

8. Aparelho para determinar um limite, **CHARACTERIZADO** pelo fato de compreender:

- um meio para selecionar uma probabilidade de falso resultado desejada; e
- um meios para selecionar um valor limite a partir de uma tabela de limites correspondente à dita probabilidade de falso resultado desejada.

9. Aparelho, de acordo com a reivindicação 8, **CHARACTERIZADO** pelo fato de compreender ainda:

- um meio para criar a dita tabela de limites;
- sendo que o dito meio para criar compreende ainda:
- um meio para calcular cada entrada de probabilidade de falso resultado na dita ta-

bela de limites usando $\frac{m!}{\prod_{i=1}^{m-1} (T = i)}$ e

- um meio para armazenar a dita probabilidade de falso resultado calculada na dita tabela de limites e o dito valor limite sendo usado para calcular a dita entrada de probabilidade de falso resultado, quando T é o dito valor limite.

10. Aparelho, de acordo com a reivindicação 9, **CHARACTERIZADO** pelo fato de que o dito meio para calcular é executado usando um primeiro valor limite, um segundo valor limite e um intervalo limite, iterativamente incrementando o dito primeiro valor limite por meio do dito intervalo limite e calculando cada dita probabilidade de falso resultado da dita tabela de limites até que o dito primeiro valor limite fique igual ao dito segundo valor limite.

11. Aparelho, de acordo com a reivindicação 10, **CHARACTERIZADO** pelo fato de que a dita probabilidade de falso resultado é uma probabilidade de falso positivo e ainda pelo fato de que o dito primeiro valor limite é um limite inferior de uma probabilidade de falso negativo e o dito segundo valor limite segundo é um limite superior da dita probabilidade de falso positivo.

12. Aparelho, de acordo com a reivindicação 8, **CHARACTERIZADO** pelo fato de que o dito meio para selecionar recupera o dito valor limite a partir da dita tabela de limites, sen-

do que o dito valor limite é um maior valor limite tendo uma probabilidade de falso resultado menor que a dita probabilidade de falso positivo desejada.

13. Aparelho para determinar um limite correspondente a uma probabilidade de falso resultado desejada, o aparelho sendo **CHARACTERIZADO** pelo fato de compreender:

- 5 - um meio para calcular uma probabilidade de falso resultado;
- um meio para calcular um valor de diferença entre a dita probabilidade de falso resultado calculada e a dita probabilidade de falso resultado desejada;
- um meio para comparar o dito valor de diferença com um primeiro valor predeterminado; e
- 10 - um meio para comparar o dito valor de diferença com um segundo valor predeterminado quando um resultado da dita primeira etapa de comparar é positivo.

14. Método, de acordo com a reivindicação 6, **CHARACTERIZADO** pelo fato de que a dita probabilidade de falso resultado é uma probabilidade de falso positivo, e compreende ainda:

- 15 - um meio para selecionar um primeiro valor limite; e
- um meio para modificar o dito primeiro valor limite quando os resultados das duas ditas etapas de comparar não são positivos.

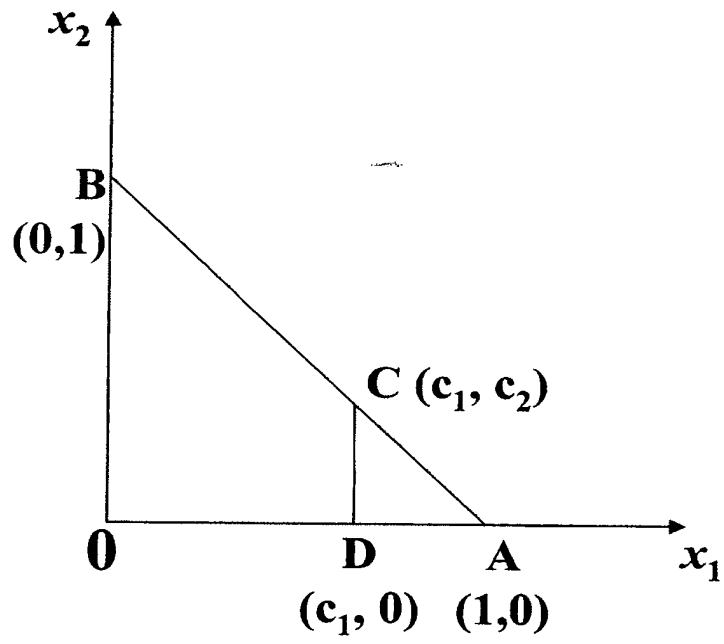


Fig. 1a

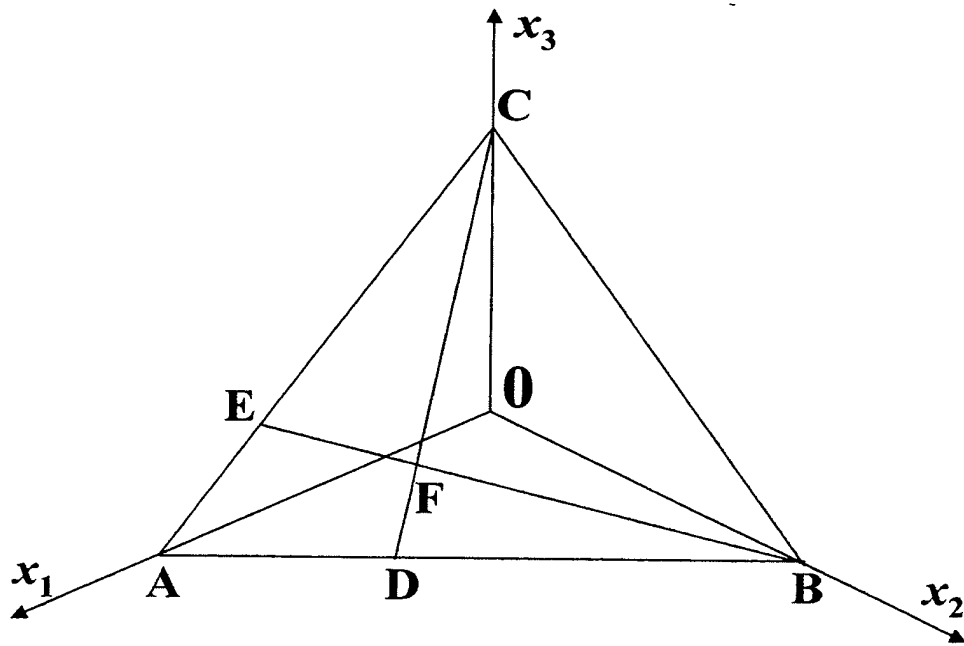


Fig. 1b

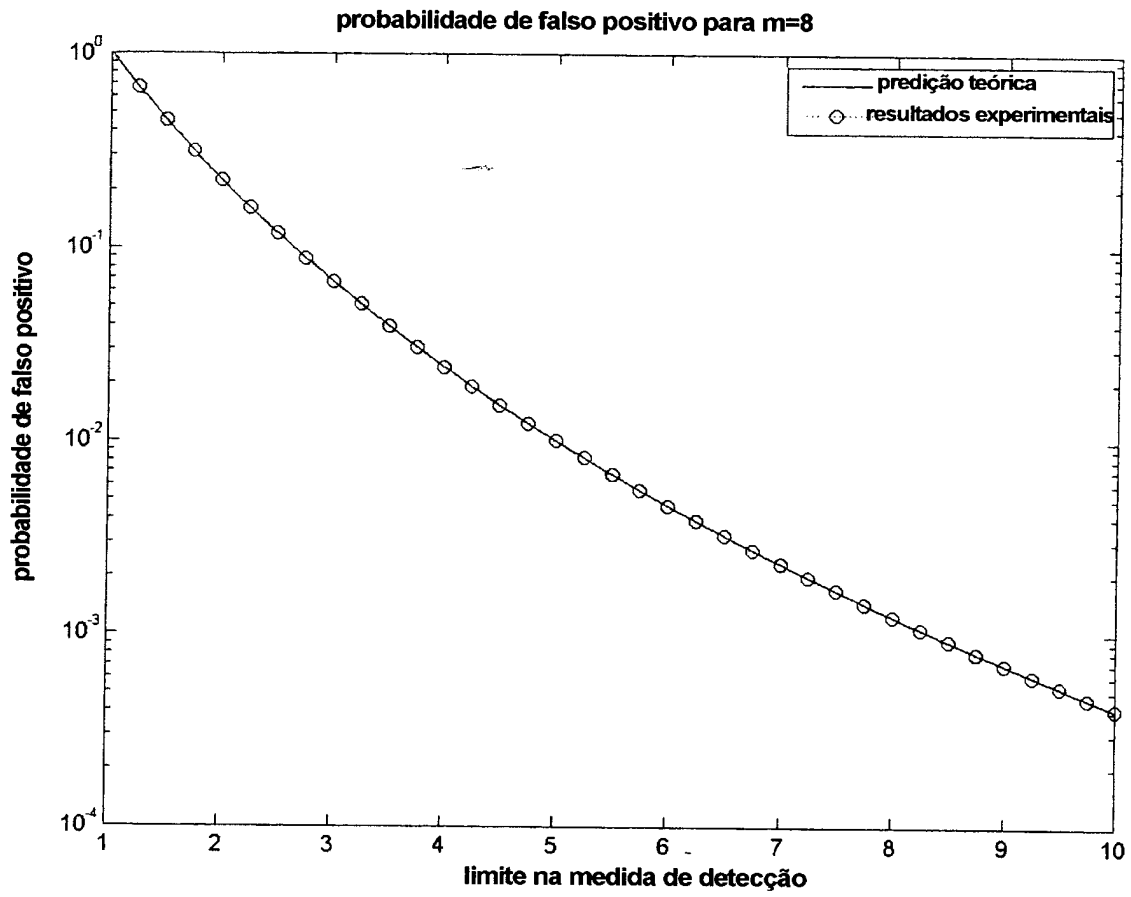


Fig. 1c

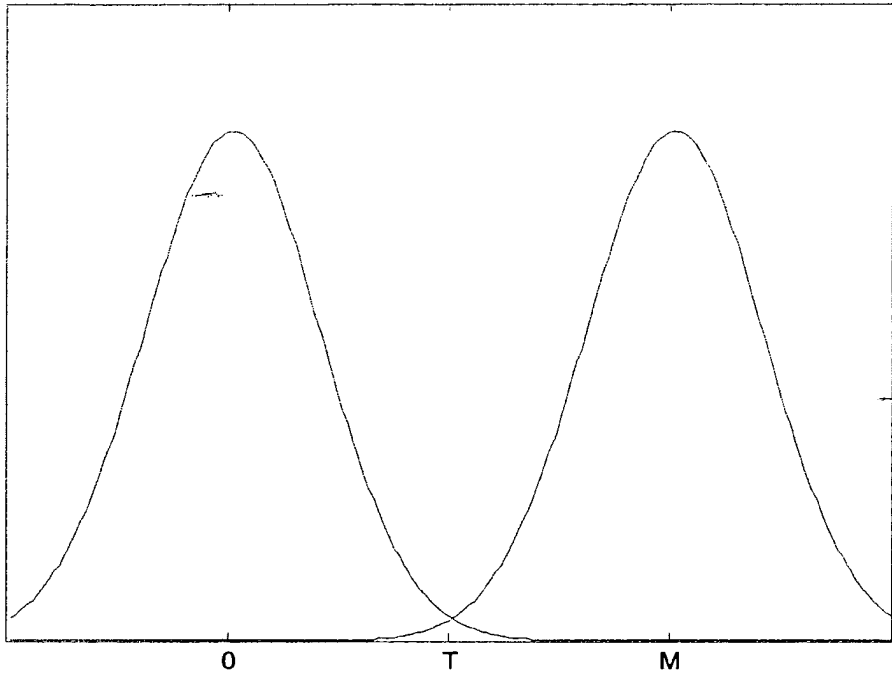


Fig. 2a

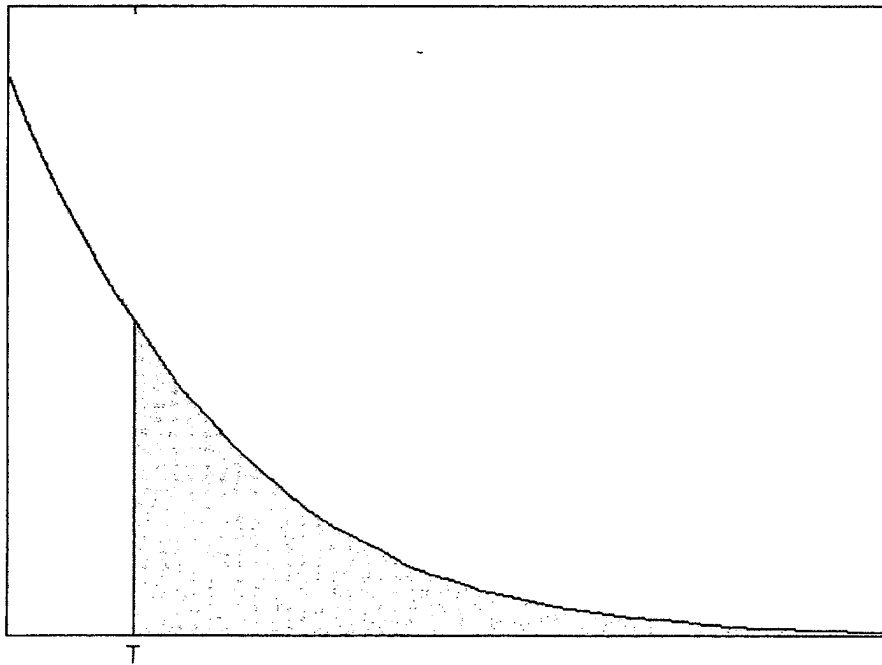


Fig. 2b

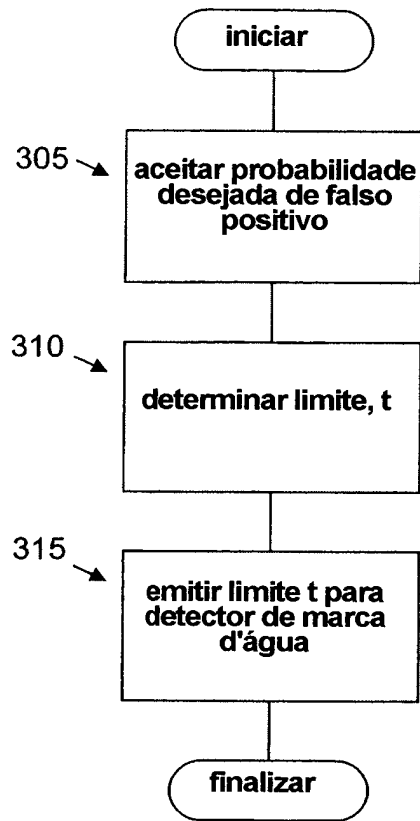


Fig. 3a

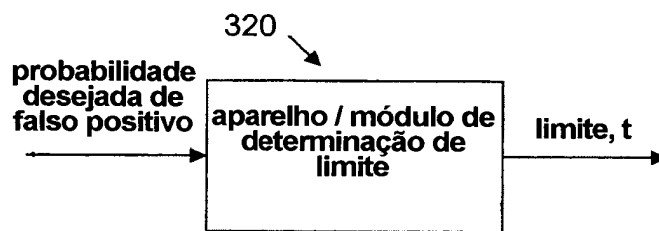


Fig. 3b

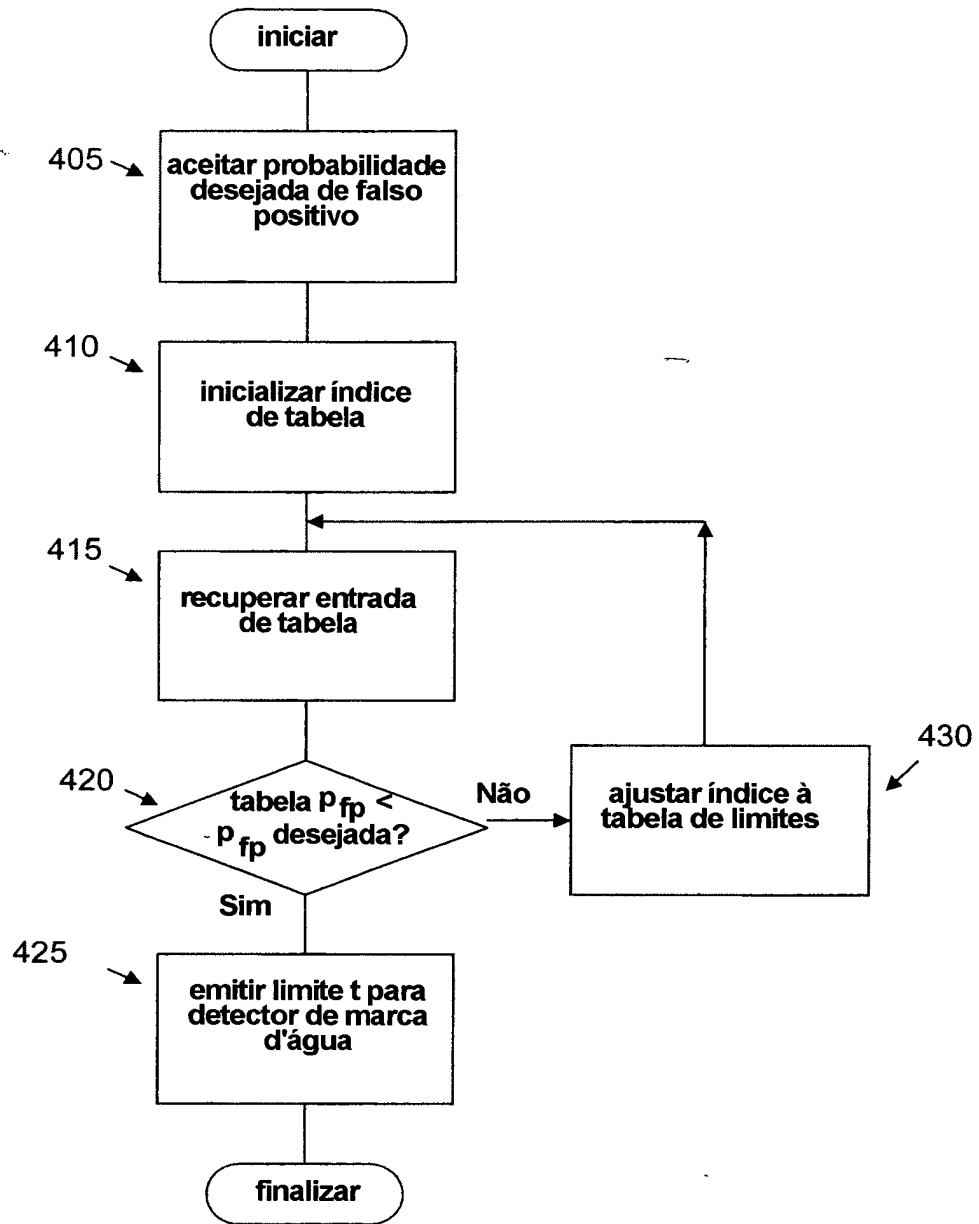


Fig. 4a

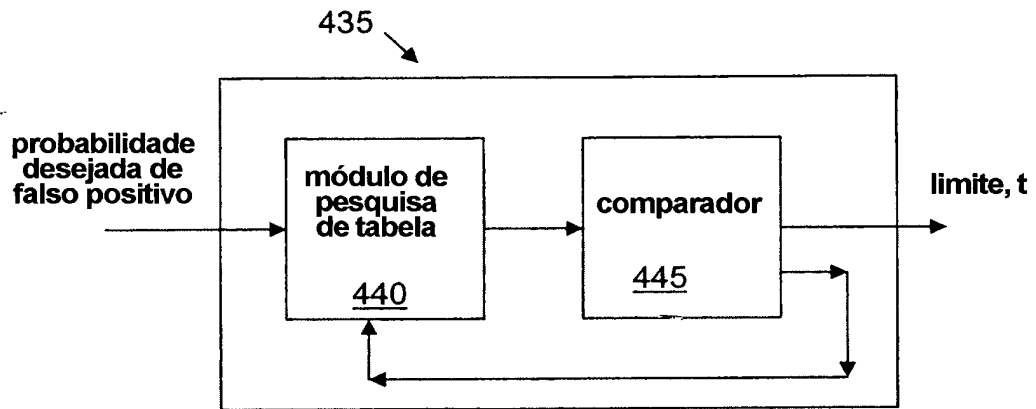


Fig. 4b

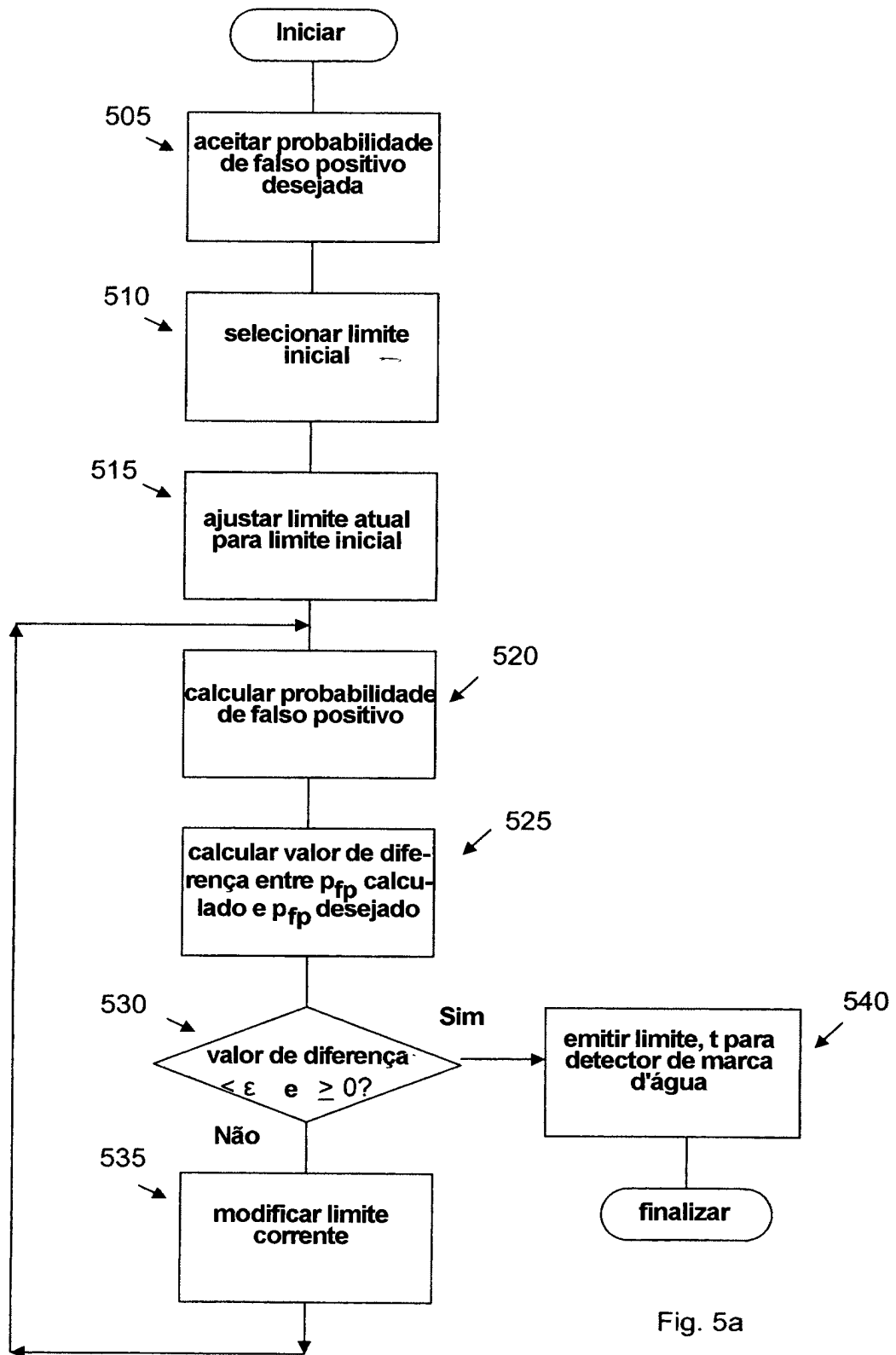


Fig. 5a

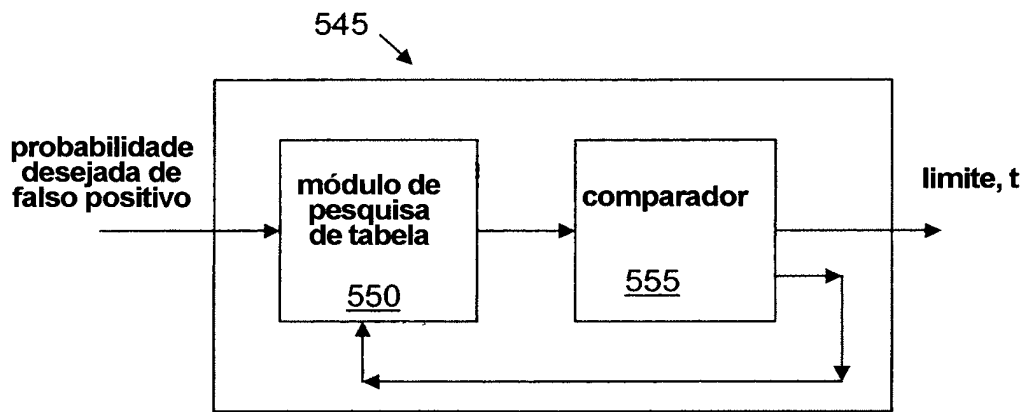


Fig. 5b

RESUMO

“MÉTODO E APARELHO PARA DEFINIR UM LIMITE DE DETECÇÃO DADA UMA PROBABILIDADE FALSA DESEJADA”

Um método e aparelho são descritos para a determinação de um limite, incluindo a
5 seleção de uma probabilidade de falso resultado desejada, e a seleção de um valor limite a partir de uma tabela de limites correspondente à probabilidade de falso resultado desejada. São também descritos um método e aparelho para determinar um limite correspondente a uma probabilidade de falso resultado desejada, incluindo o cálculo de uma probabilidade de falso resultado, o cálculo de um valor de diferença entre a probabilidade de falso resultado
10 calculada e a probabilidade de falso resultado desejada, a comparação do valor de diferença com um primeiro valor \bar{p} predeterminado e a comparação do valor de diferença com um segundo valor predeterminado, quando um resultado da primeira etapa de comparação é positivo.