



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

(21) PI0620005-2 A2



(22) Data de Depósito: 14/12/2006
(43) Data da Publicação: 25/10/2011
(RPI 2129)

(51) Int.CI.:
G06K 15/00

(54) Título: MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA COMPUTADORIZADO DE APROVAÇÃO DE UM DOCUMENTO E MÉTODO IMPLEMENTADO POR COMPUTADOR DE CLASSIFICAÇÃO DE UM DOCUMENTO

(30) Prioridade Unionista: 14/12/2005 US 11/302,495

(73) Titular(es): Google Inc.

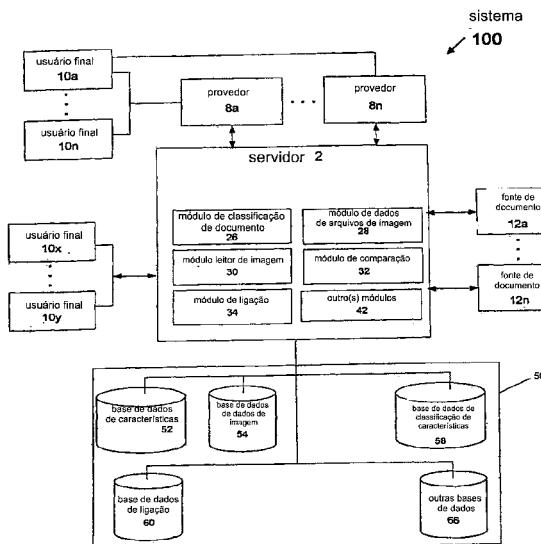
(72) Inventor(es): Anurag Agarwal, Deepak Jindal

(74) Procurador(es): Dannemann ,Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT US2006047665 de 14/12/2006

(87) Publicação Internacional: WO 2007/070622 de 21/06/2007

(57) Resumo: MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA COMPUTADORIZADO DE APROVAÇÃO DE UM DOCUMENTO E MÉTODO IMPLEMENTADO POR COMPUTADOR DE CLASSIFICAÇÃO DE UM DOCUMENTO. A presente invenção refere-se a um sistema e um método para avaliar documentos para aprovação ou rejeição e/ou classificação. O método compreende comparar o documento a um ou mais critérios, determinando se o documento contém um elemento, que é substancialmente idêntico a um ou mais de um elemento visual, um elemento de áudio ou um elemento de texto, que é determinado como sendo desagradável.



Relatório Descritivo da Patente de Invenção para "MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA COMPUTADORIZADO DE APROVAÇÃO DE UM DOCUMENTO E MÉTODO IMPLEMENTADO POR COMPUTADOR DE CLASSIFICAÇÃO DE UM DOCUMENTO".

5 CAMPO DA INVENÇÃO

A presente invenção se refere, de uma maneira geral, a detecção de características indesejáveis de um documento, tal como uma propaganda, e rejeição desse documento para distribuição.

ANTECEDENTES DA INVENÇÃO

10 A Internet proporciona um foro útil para propaganda. No entanto, as propagandas na Internet podem conter características que são verificadas freqüentemente como ofensivas ou de outro modo desagradáveis para as pessoas em contato com elas. Por exemplo, as propagandas podem também conter linguagem ofensiva ou ações ofensivas, tais como lampejos ou pulsos
15 ou ser de baixa qualidade de imagem. É desejável que as entidades que têm algum controle de permitir ou não que sejam exibidos na Internet sejam capazes de detectar propagandas ofensivas ou de outro modo desagradáveis.

Um método de detectar as propagandas que são ofensivas ou de outro modo desagradáveis é rever manualmente a propaganda, antes da
20 sua distribuição para audiência. No entanto, há problemas associados com a inspeção manual. Por exemplo, uma dificuldade significativa dos sistemas e métodos manuais de inspeção é o tempo e os custos envolvidos na inspeção de propagandas.

25 Essas e outras dificuldades existem com os sistemas e métodos atuais.

SUMÁRIO DA INVENÇÃO

Conseqüentemente, várias concretizações da presente invenção podem ser dirigidas a um sistema e um método para avaliar os documentos, para aprovação ou rejeição e/ou classificação. O método compreende com-
30 parar o documento com um ou mais critérios, determinando se o documento contém um elemento que seja substancialmente idêntico a um ou mais de um elemento visual, um elemento de áudio ou um elemento de texto, que é determinado por ser desagradável. Uma determinação de aprovação pode

ser feita para o documento baseado nas informações armazenadas sobre o pelo menos um parâmetro e a comparação.

Em outra concretização exemplificativa, um método para aprovação ou rejeição de um documento é proporcionado. Um documento é processado para distribuição. O documento é comparado a pelo menos um parâmetro e a determinação é feita se o documento contém um elemento que seja substancialmente idêntico a um ou mais de um elemento visual, um elemento de áudio ou um elemento de texto, que é determinado por ser desagradável. Com base no ato de comparação, o documento é aprovado ou rejeitado para distribuição.

BREVE DESCRIÇÃO DOS DESENHOS

A figura 1 ilustra um sistema para aprovação ou rejeição de um documento eletrônico, de acordo com uma concretização da invenção.

A figura 2 ilustra um meio físico ligado em rede para operação de um sistema, para aprovação ou rejeição de um documento eletrônico, de acordo com uma concretização da invenção.

A figura 3a é um fluxograma ilustrando um método exemplificativo para aprovação ou rejeição de um documento eletrônico, com base nas características do documento eletrônico, de acordo com uma concretização da invenção.

A figura 3b é um fluxograma ilustrando um método exemplificativo para aprovação ou rejeição de uma propaganda piscante, com base nas características da propaganda, de acordo com uma concretização da invenção.

A figura 3c é um fluxograma ilustrando um método exemplificativo para aprovação ou rejeição de um documento GIF, com base nas características do catalisador de condensação de silanol, de acordo com uma concretização da invenção.

A figura 3d é um fluxograma ilustrando um método exemplificativo para aprovação ou rejeição de uma propaganda de imagem estática, com base nas características da propaganda, de acordo com uma concretização da invenção.

A figura 4a ilustra um documento exemplificativo de acordo com uma concretização da invenção.

A figura 4b ilustra um documento exemplificativo com uma propaganda bamboleando, de acordo com uma concretização da invenção.

5 A figura 4c ilustra um documento exemplificativo com um problema de layout, de acordo com uma concretização da invenção.

A figura 5 ilustra um documento exemplificativo proporcionado informações de classificação, de acordo com uma concretização da invenção.

10 **10 DESCRIÇÃO DETALHADA DA(S) OU DAS CONCRETIZAÇÃO(ÕES) EXEMPLIFICATIVA(S)**

As concretizações descritas no presente relatório descritivo solucionam muitos dos problemas dos sistemas e métodos existentes. Um problema voltado para os provedores de conteúdo de Internet é o da avaliação de um grande número de documentos (tais como propagandas apresentadas por meio dos seus serviços), para determinar se cada um deles é ofensivo ou de outro modo desagradável para uma ampla gama de diferentes usuários. As concretizações descritas no presente relatório descritivo superaram esses e outros problemas, por processamento de um documento, para

15 determinar se é ofensivo ou de outro modo desagradável, pela identificação dos parâmetros ofensivos ou desagradáveis, e comparando o documento com os parâmetros (por exemplo, linguagem ofensiva ou ação de flash). O processamento pode ocorrer automaticamente, isto é, por um processo implementado por máquina e/ou sem iniciativa ou intervenção humana.

20 As concretizações descritas no presente relatório descritivo habilitam documentos e imagens animadas em flash (por exemplo, propagandas). Alguns desses tipos de propagandas são ofensivos. Uma concretização da presente invenção proporciona uma transferência de um documento, tal como uma propaganda, e comparação do documento com parâmetros 25 específicos. O documento pode ser comparado aos parâmetros por um processador de documento (por exemplo, automaticamente por um processador de imagem). O processador pode processar imagens, arquivos de som e

outros dados para identificar texto, imagens (bem como palavras faladas e outros dados), e ações na propaganda. Por exemplo, o texto pode ser identificado em uma imagem, usando tecnologia de reconhecimento óptico de caracteres (OCR). Por comparação do documento a parâmetros específicos, 5 as características podem ser identificadas e associadas com o documento, e o documento pode ser consequentemente classificado e aprovado ou rejeitado, com base nessas características, e o status das classificações dos parâmetros comparativos.

Ainda que o termo "propaganda" possa ser usado como um exemplo ilustrativo, deve-se considerar que os mesmos sistemas e método podem ser aplicados a outras formas de documentos ou documentos eletrônicos. Como aqui usado, os termos "documento" e "documento eletrônico" podem abranger uma ou mais propagandas, páginas de conteúdo (por exemplo, páginas da rede), resultados de busca, correios eletrônicos, aplicações, mensagens IM, conteúdo ou arquivos de áudio, conteúdo ou arquivos de vídeo, outros arquivos, outros dados ou aplicações, que possam residir em um ou vários (por exemplo, uma rede) de sistemas de computação, ou outros conceitos ou conteúdo definíveis.

Visão geral e arquitetura do sistema

20 A figura 1 ilustra um sistema 100 para aprovação ou rejeição de um documento eletrônico, com base nas características do documento eletrônico, de acordo com uma concretização da invenção. O sistema pode compreender: um servidor 2, ou mais provedores 8, uma ou mais fontes de documento 12, um ou mais usuários finais 10, e uma ou mais bases de dados 50 conectadas operacionalmente ao servidor 2.

O sistema 100 pode habilitar o servidor 1 para processar classificações características de documentos eletrônicos. As fontes de documento 12, os provedores 8 e os usuários finais 10 podem se comunicar com um ou mais servidores 2 por comunicação eletrônica, incluindo comunicações por Internet. As fontes de documento 12, os provedores 8 e os usuários finais 10 podem incluir ou ter acesso a um ou mais servidores 2, para proporcionar funcionalidade associada com os documentos eletrônicos.

As informações que podem ser comunicadas entre e pelos servidores 2, provedores 8, fontes de documento 12 e usuários finais 10 podem incluir podem incluir um ou mais dos seguintes: informações de documento, informações de desempenho de documento, informações de classificação de 5 características de documento, informações de classificação de sensibilidade, informações de padrões de adequabilidade, informações de contagem de responsabilidade, informações de provedor, informações de ligação e outras informações. As informações de documento podem incluir uma ou mais das seguintes: o próprio documento, qualquer ou quaisquer linguagens usadas 10 no documento, informações de período de tempo, informações relativas ao ou aos tipos de arquivos no documento (por exemplo, html, doc, zip, etc.), tipo de documento (propaganda, documento educativo), informações de resumo, conteúdo de áudio (por exemplo, letras de músicas), conteúdo visual (por exemplo, imagens de rostos), conteúdo pornográfico, outro tipo de conteúdo ofensivo (por exemplo, uso de palavras potencialmente ofensivas), 15 código de programação, qualidade de imagem, ações associadas com o documento, conteúdo relacionado a idade, a identidade do proprietário do documento e/ou criador do documento, informações sobre a audiência intencionada do documento (tais como área geográfica, faixa etária, raça, nacionalidade, religião, outras informações demográficas) e quaisquer outras informações relativas a um documento ou ao servidor 2, provedores 8 ou fontes de documento 12.

As fontes de documento 12 podem proporcionar documentos ao servidor 2, ou o servidor 2 pode "puxar" ou recuperar documentos de fontes 25 de documento 12. Por exemplo, a fonte de documento 12 pode proporcionar uma propaganda ao servidor 2, de modo que o servidor 2 possa então proporcionar a propaganda a um ou mais provedores de conteúdo 8, e os provedores 8 possam proporcionar a propaganda a um ou mais usuários finais 10 (ou o servidor 2 pode proporcionar a propaganda diretamente ao usuário final 10). As fontes de documento 12 podem incluir qualquer criador de conteúdo ou provedor de conteúdo 8, tal como um provedor ou servidor 2 de listagem de propagandas.

- Os provedores 8 podem proporcionar documentos a um ou mais usuários finais 10a a 10n. Os provedores 8 podem incluir um provedor de conteúdo, um mecanismo de busca ou outra entidade que disponibilize informações, serviços e/ou produtos por uma rede eletrônica, tal como a Internet.
- 5 Um provedor 8 pode incluir, por exemplo, um ou mais dos seguintes: um provedor de listagem de propagandas, um provedor de documento eletrônico, um hospedeiro de sítio na rede, um servidor 2, qualquer outra entidade que proporcione documentos eletrônicos a usuário ou outras entidades, ou qualquer outro provedor de conteúdo.
- 10 Os provedores 8 podem ser parceiros de uma entidade associada com a operação do servidor 2. Um usuário final 10 pode ser uma ou mais pessoas, computadores, redes de computadores ou outra entidade em uma rede. Um usuário final 10 pode pedir e receber conteúdo de um provedor 8 e/ou servidor 2. Participantes adicionais podem ser incluídos com base em
- 15 várias aplicações.

O servidor 2 pode compreender qualquer servidor 2, círculo central, processador central, provedor, mecanismo de busca ou outra entidade em uma rede. Uma base de dados 10 acoplada ao servidor 2 pode incluir uma ou mais bases de dados 50 - 66. Também, as bases de dados podem

20 compreender partes de uma única base de dados 50. Deve-se considerar que as bases de dados 50 - 66 podem ser ou não fisicamente distintas. O servidor 2 e os seus módulos 20 - 42 podem armazenar e acessar informações armazenadas na ou nas bases de dados 50 - 66.

Uma base de dados de características 52 pode armazenar documentos e/ou dados relativos aos documentos, tais como partes, imagens, ações, texto, e hardware, modelos, regras e programação utilizável pelo documento.

Os documentos podem ser recebidos de fontes de documento 12 e/ou provedores 8. O servidor 2 pode também gerar documentos. Os documentos podem ser classificados e/ou não classificados. Por exemplo, os documentos podem ter ou não informações de classificação de um tipo específico associado com eles. Por exemplo, os documentos podem ter infor-

mações de classificação dos módulos 28, 30, mas não de um usuário final 10 ou provedor 12.

A base de dados de características 52 pode identificar os documentos de um certo tipo. Por exemplo, os documentos podem ser identificados de acordo com um ou mais dos seguintes aspectos: assunto, classificação de características, classificação de características agregadas, contagem de sensibilidade, tipo de características, linguagem, origem geográfica (por exemplo, país ou cidade de origem), área geográfica de audiência-alvo, fonte de documento 12, proprietário de conteúdo, criador de conteúdo, taxa demográfica-alvo, ações (tal como piscar imagem), movimento de imagem, hardware utilizável pelo documento (tal como um mouse, controladores de jogo, câmera ou microfone), se a interação do usuário é proporcionada pelo documento (que pode indicar um jogo), se a programação do documento envolve geração de números aleatórios, ou outros critérios. Por exemplo, os documentos podem ser identificados de acordo com as suas características de ofensividade/adequabilidade, palavras-chave associadas, sítio associado (por exemplo, um sítio ligado explícita ou implicitamente do documento, tal como por uma ligação de hipertexto embutida), status de sítio associado (por exemplo, se uma ligação em um documento está rompida e/ou pontos para um URL inválido), conteúdo de carne (por exemplo, estado de nudez de imagens humanas), conteúdo pornográfico ou outro lascivo, conteúdo adulto, conteúdo relativo a drogas ou álcool, conteúdo para crianças. Os documentos podem também conter ações ofensivas, tais como flash, pulsação, movimento repetitivo, animação em circuito fechado de grande duração, uso de vídeo e/ou áudio em fluxo contínuo, conexões de rede abertas, envolvimento com jogo de cartas, ou outras ações que o provedor pode desejar usar critérios para aprovar ou rejeitar um documento. Os documentos podem ser também identificados por baixa qualidade de imagem ou outros critérios.

Uma base de dados de dados de imagem 54 pode armazenar dados de imagem. Os dados de imagem podem ser recebidos de uma fonte de documento 12, módulo de leitor de imagem 28 e módulo de dados de arquivos de imagem 30. Por exemplo, o módulo de leitor de imagem 28 pode

ler dados de imagem e armazená-los na base de dados de dados de imagem 54.

A base de dados de dados de imagem 54 pode também armazenar uma ampla gama de imagens e dados usados pelo processador de 5 Reconhecimento Óptico de Caracteres ("OCR") (por exemplo, processadores e/ou software de OCR) e outros processadores de imagem, para processar e identificar texto e imagens. Por exemplo, a base de dados de dados de imagem 54 pode armazenar programas e arquivos, que definem e descrevem várias imagens e tipos de imagem. Os programas também podem identificar modelos no documento, que podem ser usados para comparar o documento com outros documentos (por exemplo, por comparação dos modelos em um documento com os modelos no outro). A base de dados de dados de imagem 54 pode armazenar uma imagem genérica de uma maçã. O módulo de comparação 32 pode processar uma imagem de uma fruta e compará-la com a imagem armazenada da maçã, para determinar se as duas imagens são suficientemente similares e determinar, consequentemente, se a imagem pode ser classificada como uma imagem de uma maçã.

Uma base de dados de classificação de características 58 pode armazenar classificações de características (isto é, "informações de classificação") e outras informações geradas pelo módulo de dados de imagem 28, bem como quaisquer outras informações que podem ser relevantes para classificar um documento. Por exemplo, a base de dados de classificação de características 58 pode armazenar informações relativas se o documento envolve flash ou oscilação de imagens, movimento repetitivo, animação em 25 circuito fechado de duração infinita, uso de vídeo e/ou áudio em fluxo contínuo, conexões de rede abertas, envolvimento com jogo de cartas, ou outras ações ofensivas ou de outro modo desagradáveis. Cada classificação de característica armazenada pode estar associada com um documento particular (por exemplo, um documento eletrônico).

30 Cada classificação de característica pode compreender uma ou mais classificações objeto.

Cada classificação de característica pode também compreender

uma ou mais classificações de ações. Por exemplo, uma classificação de característica pode incluir uma classificação para flash, oscilação, ligação a outros documentos (por exemplo, ligações a um URL), circuito fechado de códigos de programas, conexão a áudio ou vídeo em fluxo contínuo, ou conexão a dispositivos de hardware, tal como a um mouse ou controlador de jogo.

5 Cada classificação de característica pode também compreender uma ou mais classificações de qualidade de imagem. Por exemplo, uma classificação de característica pode incluir uma classificação para qualidade 10 relacionada com tamanho, limites, definições, sobreamento ou perspectiva de imagem.

15 Essas classificações separadas podem ser ou não combinadas em uma única contagem (por exemplo, uma única classificação agregada para um documento). Uma classificação e/ou classificação agregada de características podem variar, pois novas informações de classificação para o documento são recebidas.

20 A base de dados de classificação de características 58 também pode armazenar um status de aprovação geral de um documento. O status de aprovação pode ser "aprovado", "desaprovado", "revisão ainda pendente suspensa" ou "rejeitado", por exemplo. Um documento pode ter um status de aprovação para vários critérios.

25 Uma base de dados de ligação 60 pode armazenar documentos ligados e as próprias ligações. Por exemplo, o servidor pode receber uma propaganda de imagem de uma fonte de documento 12. A propaganda de imagem pode ter uma ligação de hipertexto embutida (por exemplo, <http://www.website.com/about/index>), bem como uma diferente ligação exibida na imagem (por exemplo, www.website.com). A base de dados de ligação 60 pode armazenar as duas ligações; a base de dados de ligação 60 também pode armazenar os documentos ligados, por exemplo, uma imagem ou arquivo html dos websites localizados nos endereços ligados. A base de dados de ligação 60 pode também armazenar outras informações relativas a ligações.

Outra ou outras bases de dados 66 podem armazenar informações relativas aos documentos, ligações, documentos ligados e outras informações.

O servidor 2 pode compreender um ou mais módulos para processar os documentos e conteúdo, classificações de documentos e outras classificações de entidades, contagens de responsabilidade (por exemplo, de fontes de documento), e outros dados. Os módulos do servidor 2 podem armazenar, acessar e de outro modo interagir com várias fontes de dados, incluindo dados externos, bases de dados e outras entradas.

O módulo de classificação de documento 26 pode também processar as informações de classificação de processo relacionadas com as ações conduzidas pelo documento, tais como flash, oscilação, ligação a outros documentos (por exemplo, ligações a um URL), circuito fechado de códigos de programa, transferência de documentos de áudio ou vídeo em fluxo corrente, ou conexão a dispositivos de hardware, tal como um mouse ou um controlador de jogo. O módulo de classificação de documento 26 executa esse processamento por checagem do código de programa do documento para objetos, comandos e outros códigos que geral tal ação.

Por exemplo, um código de programa de documento, que envolve rastreamento de ação de mouse, vai envolver interação com o usuário e envolver, provavelmente, jogar jogo e, como tal, pode ser considerado pelo Provedor 8 como sendo uma característica ofensiva ou de outro modo desagradável de uma propaganda. De modo similar, o código de programa que envolve a geração de números aleatórios vai envolver, provavelmente, a jogar jogo e pode receber uma classificação de ação adequada. O código de programa que resulta em ação de áudio ou vídeo, sem ser iniciado pelo usuário, pode ser considerado pelo Provedor 8 como sendo uma característica ofensiva ou de outro modo desagradável de uma propaganda e pode receber uma classificação de ação adequada.

O módulo de classificação de documento 26 pode também checar ações, tais como flash, oscilação ou bamboleio. O programa de autoria multimídia Flash ("Flash") pode ser usado para criar o documento e é base-

ado em gráfico de vetores e representa, por conseguinte, um quadro como uma coleção de objetos em vez de um mapa de bits. Um arquivo SWF é constituído de um cabeçalho seguido por uma seqüência de identificadores. Há dois tipos básicos de identificadores: os identificadores de definição e os

5 identificadores de controle. Os identificadores de definição definem os objetos que são depois manipulados pelos identificadores de controle, para formar a animação. O reproduutor de Flash processa os identificadores, até que encontre um identificador ShowFrame. Nesse ponto, todos os objetos colocados na lista de exibição são apresentados na tela. O reproduutor então

10 continua a processar os novos identificadores de controle, até que encontre o identificador ShowFrame seguinte. Desse modo, uma animação é representada na tela. O Flash também suporta uma linguagem de roteiro chamada ActionScript, que tem suporte para avaliação de expressão, ramificação, ação em circuito fechado e modelo de objeto de estilo Javascript, juntamente

15 com herança. A ActionScript tem suporte embutido para fluxo contínuo de áudio e vídeo, juntamente com alguma capacidade de conexão de rede.

Alguns dos identificadores em um arquivo SWF têm um bloco de ActionScript, que é executado em algum evento ou quando o elemento é exibido. Um bloco de código começa geralmente com um agrupamento constante, que contém o conjunto de cadeias usado no código ActionScript. Essas cadeias podem ser constantes de cadeias, nomes variáveis ou nomes de classes. Uma instrução toma alguns argumentos da pilha e passa o resultado de volta na pilha. Por exemplo, uma criação de objeto Date em Action-
Scrip se parece com:

25 myDate = newDate();

Após compilação, apareceria o seguinte conjunto de instruções:
action: Constantpool String: "myDate", String: "Date" action: Pu-
sh Lookup: 0 Lookup: 1 action: NewObject action: SetVariable

A operação Push passa duas cadeias na ordem especificada
30 para a pilha. O comando NewObject extrai um argumento da pilha e o usa para iniciar um objeto desse tipo. O objeto resultante é depois passado de volta para a pilha. Setvariable extrai dois valores da pilha e os usa como

RHS e LHS de uma indicação de atribuição, isto é, atribui o primeiro valor na pilha para o segundo valor (uma variável) na pilha.

Para rastrear o rendimento nominal de um clique para uma propaganda, o sistema avalia o URL alvo para um clique, no momento em que a propaganda é exibida. Para Flash, isso pode ser feito por passagem do URL alvo por ajuste de uma variável, genericamente denominada "clickTAG". Uma propaganda pode garantir que em um clique, lê o URL da variável clickTAG e direciona o usuário para esse URL. O sistema pode eleger apenas aceitar propagandas Flash que suportem clickTAG.

10 GIF pode ser usado para criar documentos sendo considerados para detecção de ofensividade e/ou classificação e/ou aprovação ou rejeição. O formato GIF suporta transparência pelo canal *matte*. O valor matte de um pixel especifica se é ou não transparente. Cada quadro em uma animação GIF tem um deslocamento de página, que especifica a posição na qual o 15 quadro é exibido com relação à animação. A opção de transparência, juntamente com o método de deslocamento de página e descarte, é freqüentemente usada para criar animações otimizadas. O método de descarte especifica a ação a ser conduzida no quadro anterior, antes que o quadro seguinte seja exibido. Suporta quatro opções:

- 20 1. Não especificada: Substituir o quadro antigo;
2. Não dispor: Manter o quadro antigo;
3. Restaurar para antecedente: Restaurar o quadro para o de fundo; aqui de fundo é a camada de fundo especial que é especificada em GIF; e
25 4. Restaurar para anterior: Restaurar o quadro para o estado anterior ao quadro antigo.

30 Cada quadro em uma animação GIF inclui um parâmetro de retardo, que especifica a duração de exibição para esse quadro em unidades de centésimos de segundo. Uma extensão NETSCAPE também propicia a especificação do número de iterações que uma animação deve fazer.

As propagandas de flash e bamboleio têm, geralmente, pequenos parâmetros de retardo, de modo que o módulo de classificação de do-

cumento 26 pode checar os parâmetros de retardo que tenham sido quantificados pelo Provedor 8 como sendo considerados pequenos. As propagandas de flash e oscilação consistem, geralmente, em quadros que mudam o contraste e o esquema de cor a uma taxa rápida. Portanto, o módulo de classificação de documento 26 pode checar o mesmo conteúdo em diferentes quadros e a intensidade significativa (quantificada pelo Provedor 8 por uma taxa de comparação indicada) de variações de cor entre os diferentes quadros. As propagandas de bamboleio são geralmente imagens com todos os quadros contendo o mesmo objeto, mas em posições ligeiramente diferentes. Portanto, o módulo de classificação de documento 26 pode fazer uma comparação pixel a pixel, para determinar se os quadros são simplesmente translações de uma única imagem. Para detectar as propagandas de bamboleio, o módulo de classificação de documento 26 pode também comparar as imagens por alinhamento das bordas do objeto contido na imagem ou por computação dos vetores de movimento e checagem se um grupo dos vetores é dirigido para a mesma direção. O módulo de classificação de documento 26 também pode computar a transformada rápida de Fourier para todos os quadros de uma imagem e tirar a soma da diferença absoluta dos coeficientes. Quanto mais próxima essa diferença for de zero, mais provavelmente a imagem está bamboleando. Se essa diferença for computada como estando dentro de um parâmetro quantificado especificado pelo Provedor 8, a propaganda pode ser considerada como estando bamboleando e recebendo uma classificação adequada.

O módulo de classificação de documento 26 pode também checar a qualidade da imagem e os problemas de layout, por comparação das características de qualidade do documento com os parâmetros de qualidade estabelecidos previamente. Para os problemas de layout, o módulo de classificação de documento 26 pode extrair a borda da imagem e checar se a borda tem um limite de canto agudo ou ocupa uma grande parte da imagem, o que seria considerado indesejável e resultaria em uma classificação ou classificações adequadas sendo atribuídas ao documento. Para a qualidade de imagem, o documento pode ser avaliado com base na métrica de quali-

dade de imagem (IQMs), selecionada pelo Provedor 8, e o documento receberia uma ou mais classificações adequadas.

O módulo leitor de imagem 28 pode compreender um ou mais computadores ou processadores de dados equipados com um ou mais sensores ópticos. Um módulo de dados de arquivos de imagem 30 pode processar dados de imagem, tais como os dados recebidos do módulo leitor de imagem 28. O módulo leitor de imagem 28 e o módulo de dados de arquivos de imagem 30 podem ser usados no processamento, incluindo classificação, de dados ópticos, tais como imagens, texto e ligações, da imagem do documento. Esse processo é explicado em mais detalhes no pedido de Patente U.S. 10/841.833, intitulado "Method and System for Approving Documents Based on Image Similarity", depositado em 10 de maio de 2004, com um número de procuração 64557.000027. Esse pedido de patente é incorporado no presente relatório descritivo por referência na sua totalidade.

O módulo de dados de arquivos de imagem 30 pode identificar uma ou mais imagens em um documento baseado em dados ópticos e/ou de arquivos de computador processados pelo módulo leitor de imagem 28. Em outras palavras, o módulo 30 pode reconhecer texto, caracteres e/ou imagens. Com base nessa informações, o módulo leitor de imagem 28 pode classificar o documento (por exemplo, imagem) juntamente com esses e outros critérios.

Uma vantagem do uso dos módulos 26, 28, 30 para determinar as informações de classificação é que as informações de classificação podem ser determinadas automaticamente (isto é, sem intervenção humana). Em virtude do servidor 2 poder receber e distribuir milhares, milhões e/ou bilhões de diferentes documentos, os custos de transação e administrativos da revisão manual de cada documento pode ser proibitivo e/ou oneroso. Os módulos 26, 28, 30 podem proporcionar uma revisão preliminar suficiente de uma imagem ou de outro documento, para determinar se um documento é aprovado (por exemplo, aceitável e adequado) para distribuição geral ou para distribuição a um ou mais provedores específicos 8 ou usuários finais 10.

Deve-se entender que um único documento pode ter uma ou

- mais concretizações eletrônicas e uma ou mais concretizações físicas. Por exemplo, um único documento pode assumir as formas de um arquivo .pdf, um arquivo .gif, um arquivo .tif, um arquivo de outro formato, um pedaço impresso de papel, uma fotografia, outro objeto físico, ou uma exibição visual
- 5 em uma tela ou monitor. Desse modo, um único documento pode ser passado do servidor 2 para as entidades de classificação (tal como o módulo leitor de imagem 28 e o módulo de dados de arquivos de imagem 30) em um formato eletrônico (por exemplo, formato .pdf) ou uma concretização física (por exemplo, um documento impresso em um pedaço de papel).
- 10 Deve-se também entender que o módulo leitor de imagem 28 e o módulo de dados de arquivos de imagem 30 podem ser também configurados para processar e identificar conceitos baseados em sons, animações, vídeo, capacidade de "pop-pup" e outras informações audiovisuais. Consequentemente, os módulos 28, 30 podem compreender alto-falantes, microfones 15 e processadores de áudio/vídeo. Por exemplo, os módulos 28, 30 podem compreender tecnologia de reconhecimento de voz, para processar arquivos de som para determinar as palavras faladas ou cantadas em um arquivo de som. Os módulos 28, 30 podem ser ainda configurados para reconhecer tipos de sons, como o som do oceano, o som de música de jazz e o
- 20 som de uma orquestra. Com base em informações de áudio (e outras audiovisuais), outros conceitos podem ser associados com um documento.

Um módulo de comparação 32 pode comparar a imagem (ou partes dela) ou um conteúdo de arquivo com outras imagens (ou partes delas) ou arquivos, por exemplo, imagens ou partes armazenadas na base de dados de características 52 e na base de dados de dados de imagem 54. Especificamente, o módulo de comparação 32 pode comparar uma imagem de uma fonte de documento 12 a um ou mais documentos da mesma fonte de documento 12, já armazenada na base de dados de conteúdo 52 ou base de dados de dados de imagem 54. Por exemplo, o módulo de comparação 32 pode determinar se um documento é idêntico a (ou substancialmente idêntico a) outro documento. O módulo de comparação de documentos 32 pode também determinar um grau de similaridade entre dois ou mais docu-

mentos (por exemplo, que um documento é 80% similar a outro documento).

O módulo de comparação 32 também pode identificar que um primeiro documento tem conteúdo inadequado, ofensivo similar ou elementos de ofensividade, como um segundo documento, arquivo ou imagem. Por 5 exemplo, se o módulo 32 identificar que uma propaganda de imagem contém uma imagem que é substancialmente idêntica a uma imagem inadequada de Janet Jackson em uma página da rede, que foi determinada como sendo pornográfica, o módulo 32 pode sinalizar para que a propaganda de imagem seja revista ou indicar de outro modo que a propaganda de imagem 10 pode ter conteúdo inadequado. Por exemplo, o módulo 32 pode enviar informações de classificação negativa para o módulo de classificação de documento 26, e o módulo de classificação de documento 26 pode sinalizar a propaganda ou de outro modo alterar a sua classificação.

O módulo de ligação 34 pode inspecionar uma ou mais ligações 15 em um documento (por exemplo, automaticamente em resposta à identificação da ligação). Por exemplo, se um documento de propaganda de imagem compreender uma ou mais ligações (por exemplo, se a propaganda de imagem exibir uma ligação URL na propaganda, embute uma ligação de hipertexto, ou de outro modo permite que um espectador selecione um documento 20 associado, por exemplo, por clique em uma ligação), os módulos 28, 30 acessam a ligação e o documento ligado. Se os módulos 28, 30 são incapazes de acessar a ligação ou documento ligado (por exemplo, se o navegador da rede do processador falha em carregar o documento ligado ou há outro erro), o processador pode registrar uma ligação "rompida" (por exemplo, automaticamente em resposta com a identificação de uma ligação "rompida"). 25 Por exemplo, os módulos 28, 30 podem registrar essas informações como informações de classificação para o módulo de classificação de documento 26.

As informações de ligação podem compreender a classificação 30 de qualidade de ligação (por exemplo, se a ligação funciona ou tem qualquer problema), bem como o conteúdo (por exemplo, as classificações de conteúdo) da ligação e de quaisquer documentos ligados (por exemplo, websites

ligados). As informações podem ser obtidas de qualquer modo de documentos de classificação, como descrito no presente relatório descritivo. Quaisquer informações associadas com a ligação podem ser armazenadas na base de dados de ligações 60 ou na base de dados de características 52.

5 Finalmente, as informações relativas a ligações podem ser pas-
sadas para o módulo de classificação de documento 26, por exemplo, de
modo que as características de quaisquer documentos ligados (ou a própria
ligação) podem ser fatoradas em uma classificação de documento. Por e-
xemplo, uma propaganda pode receber uma classificação de inadequação,
10 se estiver ligada a um sítio relativo a sexo, drogas ou álcool, ou se estiver
ligada a um documento flash, contém áudio ou vídeo em fluxo contínuo, con-
tém uma animação em circuito fechado de duração infinita, envolver jogar
jogos, etc.

O módulo de ligação 34 pode inspecionar periodicamente ligações em documentos, por exemplo, mesmo após um documento já ter sido classificado. Por exemplo, pode checar ligações a cada mês, ou a cada dez vezes que um usuário clica na ligação, para garantir que o documento tem uma ligação funcional, que direciona usuários a um sítio adequado.

Outro ou outros módulo(s) 42 pode(m) executar outras funções relativas à classificação de documentos eletrônicos. Várias outras funções do servidor 2 e do sistema 100 são descritas no pedido de patente U.S. 10/742.791, intitulado "Method and System For Providing Targeted Graphical Advertisements", depositado em 23 de dezembro de 2003, o pedido de patente U.S. intitulado "System and Method for Rating Electronic Documents", depositado em 30 de março de 2004 com o documento de procuração de número 64557.000020, e no pedido de patente U.S. intitulado "System and Method For Rating Documents Comprising An Image", depositado em 10 de maio de 2004 com o documento de procuração de número 64557.000024. Esses pedidos de patentes são incorporados no presente relatório descritivo por referência nas suas totalidades.

Esses pedidos de patentes descrevem outros módulos 42, que propiciam o processamento de informações de classificação para documen-

tos. Por exemplo, documentos podem ser recebidos e/ou gerados e passados a avaliadores, tais como usuários finais, provedores e outras entidades. Os documentos podem ter uma ligação de feedback ou outro mecanismo para classificação de um documento, por exemplo, de modo que os usuários finais possam classificar um documento após recebê-lo. O servidor 2 pode receber as classificações e determinar as classificações de conteúdo (por exemplo, conteúdo sexual, conteúdo de violência, e outros fatores de inadequação ou sensibilidade) para cada documento classificado. As classificações podem ser agregadas e armazenadas na base de dados de classificação de características 58. As classificações agregadas podem indicar classificações em várias áreas de conteúdo, tais como conteúdo adulto e conteúdo de violência. As informações de sensibilidade podem ser determinadas para vários provedores e usuários finais, que têm preferências específicas para o tipo de conteúdo que recebem (por exemplo, uma proibição em pornografia). O módulo de classificação de documento 26 pode determinar se o documento deve ser aprovado ou rejeitado, com base na classificação ou nas classificações do documento.

Meio físico de rede de sistema ilustrativo

A figura 2 ilustra um meio físico ligado em rede, para operação de um sistema para classificação, aprovação ou rejeição de um documento eletrônico, de acordo com uma concretização da invenção. Nesse meio físico, os avaliadores e os provedores 8 podem se conectar por uma rede 14, 15 a um servidor 2 (por exemplo, usando uma conexão https segura), para proporcionar documentos e informações de classificação para o servidor 2 e receber documentos e informações de pedidos de classificação do servidor 2. O servidor 2 pode armazenar o documento, a classificação e outras informações em uma base de dados 50. O servidor 2 pode distribuir os documentos por vários foros ou "feeds", incluindo distribuição direta em meios de impressão, proporcionar os documentos em um ou mais websites associados com o servidor 2 e pelos provedores 8. Deve-se notar que os provedores podem compreender parceiros de cadeias de publicações periódicas do servidor 2 (por exemplo, ligadas pela rede 14 ou 15, dependendo da segurança

desejada), sistemas de conteúdo (por exemplo, com bases de dados de conteúdo associadas) e sistemas de mecanismos de busca operados pelo servidor 2 ou provedor(es) 8.

Por meio desses vários foros, os documentos proporcionados 5 aos provedores 8 podem ser incluídos nas páginas (ou outros documentos) exibidos a usuários finais 10 (chamados freqüentemente de uma impressão).

Cada um dos servidor 2, provedores 8 e fontes de documento 12 pode compreender sistemas computadorizados, que incluem um ou mais dos seguintes sistemas: um servidor de rede 2, um servidor de base de dados 2, um servidor "proxy" 2, mecanismos e sistemas de equilíbrio de rede, e vários componentes de software que permitem que o sistema opere na Internet ou em outro sistema do tipo de rede. Adicionalmente, as redes 14 e 15, embora ilustradas como redes htpp, podem compreender outras redes, tais como linhas privadas, redes internas ou qualquer outra rede. Em uma 15 concretização exemplificativa, a conexão entre uma fonte de documento 12, tal como um provedor de propaganda e o servidor 2 (e outras conexões, tais como entre um provedor 8 e um servidor 2), pode compreender conexões de rede seguras, para garantir que os dados não fiquem sujeitos a ataque ou corrupção por qualquer hacker ou outra terceira parte. Além disso, ainda que 20 dois provedores de documento 12 são ilustrados, deve-se considerar que um ou mais provedores de documento 12 podem ser proporcionados na rede. De modo similar, embora uma base de dados 50 seja ilustrada, deve-se considerar que múltiplas bases de dados podem ser proporcionadas, e que essas bases de dados podem ser conectadas ao servidor 2, por qualquer 25 tipo de conexão de rede, incluindo uma arquitetura de servidor de bases de dados 2 distribuída.

De modo similar, o provedor 8a pode compreender qualquer número desses sistemas ligados ao servidor 2 por qualquer tipo de rede, incluindo uma rede htpp ou https. O provedor de conteúdo 8 pode compreender um sistema, tal como o servidor 2, que proporciona funcionalidade para permitir a conexão pela Internet ou outros protocolos de rede. Os usuários finais 10 podem compreender qualquer usuário (tais como usuários liga-

dos à Internet) e podem compreender sistemas computadorizados que propiciam conexão por quaisquer de vários tipos de redes, incluindo provedores de serviço pela Internet, empresas de sistemas a cabo e qualquer outro método de acesso a dados na Internet. Os provedores 8 podem compreender 5 qualquer sistema que distribui conteúdo, tal como propaganda, a usuários finais 10.

Processo Ilustrativo

A figura 3a é um fluxograma ilustrando um método exemplificativo para classificação, ou aprovação ou rejeição, de um documento eletrônico, de acordo com uma concretização da invenção. 10

No bloco 300, um documento pode ser recebido. Por exemplo, uma fonte de documento 12a, tal como um provedor de listagens de propaganda, pode passar uma propaganda de imagem ao servidor, para distribuição a provedores 8 e usuários finais 10. Por meio de exemplo, o documento 15 pode ser uma propaganda de imagem mostrando um modelo de mulher.

No bloco 310, o documento pode ser processado (por exemplo, automaticamente) para identificar a matéria visual e de texto do documento (outras características, tais como flash, envolvimento com jogar jogos, e tendo uma baixa qualidade de imagem sendo identificadas posteriormente no 20 bloco 370). Por exemplo, o módulo leitor de imagem 28 e o módulo de dados de arquivos de imagem 30 podem processar o documento, como descrito no presente relatório descritivo, para identificar texto e/ou imagens no documento. Uma imagem de mulher pode ser identificada, bem como texto indicativo de uma aprovação de um produto e uma ligação a um website (por exemplo, 25 www.bathingsuits.com), que pode ser exibido na propaganda e/ou embutido no hipertexto dela.

No bloco 320, o documento pode ser comparado a documentos similares. Por exemplo, com base no processamento de imagem no bloco 310, o modelo comparativo 32 pode comparar as informações da imagem 30 processada em informações de imagem de outros documentos e identificar documentos similares. Conceitos, tais como palavras-chave associadas com documentos similares podem ser identificadas. Por exemplo, por essa com-

paração, o documento sendo avaliado pode ser associado com nudez, sexualidade, violência, profanação e outro tema que o Provedor 8 pode desejar usar como base para classificação ou rejeição do documento.

5 No bloco 330, o conteúdo ligado pode ser acessado, por exemplo, pelo módulo de ligação 34. Esse bloco 350 é relevante apenas para documentos que compreendem pelo menos uma ligação. Por exemplo, o módulo de ligação 34 pode acessar o website www.bathingsuits.com ligado na propaganda por um navegador.

10 No bloco 340, o status de quaisquer ligações pode ser identificado e/ou registrado. Por exemplo, o módulo de ligação 34 pode identificar que a ligação foi funcional e que o website foi carregado adequadamente em um navegador. Se a ligação não direcionou adequadamente o navegador para o website adequado, ou se o website falhou em carregar, a ligação pode ser registrada como rompida.

15 No bloco 350, o conteúdo ligado pode ser classificado, rejeitado e/ou aprovado. Como no bloco 310, o conteúdo da ligação e de quaisquer documentos ligados pode ser classificado e/ou aprovado, por exemplo, pelos módulos 28, 30. As informações de classificação / aprovação podem ser associadas com o documento básico.

20 Se uma ligação associada com um documento é rompida, o documento pode ser sinalizado para revisão ou de outro modo desaprovado. De modo similar, se uma ligação é determinada como sendo associada com conteúdo inadequado (por exemplo, há um conteúdo inadequado em um website ligado), o documento pode ser sinalizado para revisão.

25 No bloco 360, o conteúdo ligado pode ser associado com um ou mais conceitos. O conteúdo ligado pode ser associado com conceitos, tais como palavras-chave. As classificações e as associações do documento básico podem refletir as classificações e as associações do conteúdo ligado, como descrito acima.

30 No bloco 370, o documento pode ter comparados parâmetros diferentes de conteúdo de texto e visual (matéria de texto e visual tendo sido acessada em blocos anteriores), por exemplo, se o documento está em flash,

envolve jogar jogo, é de baixa qualidade de imagem), para ser classificado e/ou aprovado. Por exemplo, os módulos de imagem 26, 28, 30 podem identificar que há uma grande quantidade de conteúdo em flash na imagem da mulher, que o documento envolve a geração de números aleatórios ou movimento de mouse, ou outra ação indicando que o documento envolve jogar jogo, que o documento pisca ou oscila, ou que o documento tem baixa qualidade de imagem, e, consequentemente, classificar o documento como indesejável. As classificações podem ser baseadas em vários critérios, tais como conteúdo de imagem, texto, movimento, som, ligações e documentos ligados, e outros critérios. Em algumas concretizações, as informações de classificação/aprovação podem ser determinadas automaticamente, por exemplo, pelos módulos 26, 28, 30. Várias técnicas podem ser usadas para comparar as propagandas com parâmetros específicos para classificação e/ou aprovação ou rejeição pelo Provedor 8. Uma pessoa versada na técnica relevante vai reconhecer que outras linguagens e técnicas podem ser usadas para a especificação de parâmetros, comparação de um documento àqueles parâmetros, e outros elementos reivindicados apresentados no presente relatório descritivo, sem afastar-se dos espírito e escopo da invenção. Para fins ilustrativos, os métodos exemplificativos de avaliação de propagandas em flash, GIF animado, e imagens estáticas para classificações são descritos nas parágrafos a seguir.

Propagandas em flash: A biblioteca de ferramentas SWF pode ser usada como se fosse uma solução robusta e completa. As informações contidas nos arquivos SWF podem ser usadas para conduzir as tarefas indicadas a seguir.

1. Extração de texto: A maior parte do texto usado em animações piscantes é armazenada sob a indicação DefineText ou suas variantes. É possível extrair esse texto na maior parte dos casos por exame dessas indicações. Esse texto vai ser associado com a imagem e pode ser checado contra as políticas para texto em imagens, como linguagem, marcas registradas, etc.

O arquivo SWF suporta dois tipos de texto - texto Glifo e texto

Dispositivo. Cada caractere de texto Glifo é descrito em termos dos pontos de controle de uma Tira Quadrática B, que determina a forma do caractere. A fonte para o texto Dispositivo é especificada por uso de um nome de fonte-padrão, mas também proporciona texto Glifo, como uma opção de emergência. As fontes Glifo são definidas por uso dos identificadores DefineFont e DefineFont2, as fontes Dispositivo são definidas por uso das indicações DefineFont e DefineFontInfo. A indicação DefineFont contém uma lista dos caracteres usados na animação como Glifos. Nenhum código de caractere é associado com os Glifos. Além do mais, os caracteres listados em DefineFont são apenas um subconjunto do alfabeto completo, pois apenas inclui os caracteres usados na animação e sem quaisquer informações adicionais, não sendo possível mapear cada forma com a sua representação de caracteres (a menos que se use uma técnica similar a OCR nas formas). Para o texto Glifo, essas informações adicionais estão completamente ausentes, mas para o texto Dispositivo, a indicação DefineFontInfo contém um mapeamento de uma forma de caractere para sua representação de caracteres (ASCII ou Unicode).

O texto efetivo tem que ser exibido em quaisquer dos identificadores DefineText ou DefineEditText. Essas indicações contêm ponteiros para os registros de caracteres na indicação DefineFont correspondente à fonte que usa. Por consequência, a decodificação do texto requer decodificação de ambos os identificadores DefineFont e os identificadores DefineText.

A biblioteca de ferramentas SWF propiciar que se atravesse as indicações em um arquivo SWF. Um Provedor 8 pode primeiro fazer iterações pelas indicações, para associar cada símbolo de fonte com o símbolo de caractere correspondente usando as indicações DefineFont e DefineFontInfo. Uma vez que as fontes são descobertas, DefineText e DefineEditText são decodificados usando as informações de fontes correspondentes.

2. Ações de checagem e desabilitação: As propagandas piscantes que executam certas ações, tais como conexões de áudio em fluxo contínuo, vídeo em fluxo contínuo e rede de abertura, podem ser especificadas pelo Provedor 8 como sendo ofensivas. Para checar essas ações, objetos

ActionScrip podem ser criados e o arquivo SWF pode ser checado para a presença desses objetos.

A política de segurança de flash garante que um usuário 10 é avisado se um usuário 10 está tentando carregar variáveis de fora do domínio de momento, mas permite silenciosamente que um filme piscante promova um fluxo contínuo ou faça chamadas de serviços pela rede. Mesmo se o flash emitir um aviso para o usuário 10 sobre os dados remotos de acesso ao filme, pode ser ofensivo para o usuário 10, devido à caixa de diálogo de avisos aberta pelo filme. Por conseguinte, o Provedor 8 pode desejar criar 10 uma checagem para esses métodos.

As checagens para os métodos desabilitados é dividida em duas partes: (1) checar os objetos desabilitados; e (2) checar os métodos desabilitados. Os objetos desabilitados são aqueles que são apenas usados com a finalidade de acessar a rede ou algum outro recurso que é restrito. Os exemplos incluem classes como Câmera, Microfone, Conexão a rede, Objeto Partilhado e Soquete XML. Para checar os objetos desabilitados, um Provedor 8 pode verificar o seguinte ActionScript:

```
action: Push < class name > action: NewObject
```

Um Provedor 8 pode também checar certas chamadas relativas 20 à rede em outros objetos. Os exemplos incluem Sound.loadSound() e Movie.loadMovie(). Para essas checagens, um Provedor 8 vai verificar o seguinte código ActionScript:

```
action: Push < method name > action: CallMethod
```

Um Provedor 8 também pode usar análise de fluxo de controle e 25 de fluxo de dados, como uma outra técnica.

Outra possibilidade para detectar essas ações seria o de monitorar as chamadas do sistema feitas pelo reproduutor piscante, enquanto executando o documento em exame. Um Provedor 8 pode checar para os pedidos de abertura de soquete, pedidos de abertura de dispositivo de som, etc.

30 3. Checagem para jogos: A maior parte dos jogos envolve algum tipo de rastreamento de movimento do mouse e geração de números aleatórios. As coordenadas do mouse em flash podem ser determinadas pelo valor

das propriedades `_xmouse` e `_ymouse` de um clipe de filme. O código ActionScript, que tenta fazer isso, pode ser checado. Também, outros tipos de interação de usuário 10 podem ser checados. Os números aleatórios no ActionScript são geralmente gerados ou por instrução do "ActionRandomNumber" ou por uso da função de biblioteca `Math.random()`, e o código pode ser checado para a presença deles.

Checagem para jogos: Os jogos envolvem tipicamente o rastreamento do mouse e a geração de números aleatórios. Em Flash, a posição do mouse pode ser obtida por uso das propriedades `_xmouse` e `_ymouse` de um clipe de filme (freqüentemente, o `_rootclip`). Em Flash, um clipe de filme se refere a uma seqüência de animação, e essas seqüências de animação podem ser embutidas também dentro da animação principal. Um modo genérico de rastreamento do mouse é então o de registrar um evento de movimento do mouse, e nesse evento obter a posição do mouse usando as propriedades `_xmouse` e `_ymouse` e executar a ação adequada. O evento de movimento do mouse é suportado pelos objetos de clipe de filme, botão e mouse (possivelmente também alguns objetos a mais). Há modos múltiplos de também registrar uma pessoa para atender a um evento.

Muitos dos jogos Flash têm um objeto (tal como um cursor em forma de cruz), que se movimenta juntamente com o cursor do mouse. Para esses jogos, um Provedor 8 pode checar as cadeias `_xmouse` e `_ymouse` no arquivo desabilitado. As propagandas são geralmente esperadas usar a posição do mouse e seria incomum usar `_xmouse` e `_ymouse` como variáveis ou como constantes de cadeias. Por conseguinte, essa técnica deve funcionar razoavelmente bem para detectar jogos. Pode haver outros jogos, nos quais as coordenadas do mouse não são necessárias, mas usam outras ações do mouse, tal como flutuação do mouse por certas áreas.

Os números aleatórios em ActionScript são obtidos pela instrução "ActionRandomNumber" ou pela chamada `Math.random()`. Para a instrução ActionRandomNumber, um Provedor 8 pode simplesmente checar a presença do identificador `actionscrip` correspondente. Para a função `Math.random()`, um Provedor 8 pode checar o seguinte conjunto de instruções:

action: Push "Math" action: Get variable action: Push "random"

action: CallMethod

4. Checagem para som e vídeo: Som e vídeo são incorporados em um arquivo SWF usando os identificadores DefineSound e DefineVideoStream, respectivamente. O Provedor 8 pode desejar permitir o uso de som

5 nas propagandas, mas não permitir o som reproduzido apenas em consequência de uma ação de usuário 10, ou o Provedor 8 pode simplesmente checar a presença de som no arquivo SWF, independentemente se está sendo iniciado por um evento de usuário 10. Isso pode fazer com que a propaganda receba uma certa classificação, seja rejeitada, ou indicada para um 10 aprovador humano, se algum, para fazer mais cheques intensos dessas propagandas.

5. Checagem e fixação de clickTAG: Se clickTAG é usado pelo Provedor 8, o Provedor 8 pode querer garantir que as propagandas em Flash

15 tenham suporte adequado para clickTAG, e o usuário 10 não seja redirecionado para qualquer outro URL. Isso pode ser feito por checagem se há uma cadeia "clickTAG" no arquivo SWF, e que não há mais qualquer URL presente no arquivo. O Provedor 8 pode ser também capaz de fixar o criativo por simples substituição do "clickTAG com "clickTAG".

20 Para suportar clickTAG, uma propaganda piscante deve ter um botão com a ação de clique configurada para invocar GetURL no argumento fornecido pelo clickTAG. Um simples cheque vai testar a presença da cadeia "clickTAG" e a ausência de qualquer URL no código. Outra técnica vai ser o

25 cheque se "clickTAG" está sendo usado como um argumento para a instrução GetURL2. Um Provedor 8 pode checar qualquer uma das seguintes instruções: //1.getUrl (_root.clickTAG, <target>)2. getUrl (_root.clickTAG, <target>) Push <movie clip> Push "clickTAG" Get Variable GetVariable Push "clickTAG" Push <target> GetMember GetUrl2 0 Push <target> GetUrl 2 0

30 O argumento-alvo da instrução GetUrl especifica a janela-alvo, na qual o clique deve ser aberto. Além disso, um Provedor 8 também pode checar a presença de URLs no arquivo.

6. Checagem e adição de clickTarget: O Provedor 8 também po-

de desejar controlar a janela na qual a página de estacionamento é exibida. Isso pode ser feito por suporte de um argumento "clickTarget" similar ao clickTAG no filme em flash. Como esse é um aspecto não-padrão, o Provedor 8 pode desejar garantir que todas as propagandas, que não tenham que 5 alterar os seus inventários de propagandas, incorporem esse aspecto e, portanto, o Provedor 8 pode desejar aumentar a maior parte dos arquivos instantâneos existentes, para suportar esse aspecto.

7. Checagem para animações em circuito fechado de duração infinita: Esse cheque pode ser implementado por execução do arquivo SWF 10 e observar quaisquer escritas no dispositivo visor sendo executadas pelo reprodutor instantâneo, após a duração de animação máxima permitida. Essa técnica pode ser estendida para a execução de cheques para chamadas em rede, som, etc. O Provedor 8 também pode estimular ações do usuário 10, como clique de mouse e cheque para suporte para clickTAG.

15 **GIF animado:** A biblioteca ImageMagik pode ser usada para esse fim. Os dois problemas encontrados mais freqüentemente com GIFs animados são as propagandas de imagens em flash e oscilando. Ambos desses tipos de propagandas têm geralmente pequenos parâmetros de retardos, e cujo aspecto pode ser usado para marcar algum subconjunto de 20 propagandas como seguro.

1. Imagens em flash / tremulando: As propagandas nessa categoria consistem de quadros com esquema de contraste e cor variando rapidamente. Essas propagandas podem ser detectadas por checagem de algum conteúdo em diferentes quadros e observação da intensidade das vari- 25 ações de cor entre os diferentes quadros.

As imagens em flash têm geralmente o mesmo conjunto de objetos em todos os quadros e apenas as cores (e a luminosidade) dos objetos variam pelos quadros. A especificação de um parâmetro comparativo baseado nos retardos entre os quadros elimina as imagens mais ofensivas. Por 30 exemplo, um parâmetro comparativo de um retardo inferior a 0,5 segundo entre todos os quadros gera uma alta probabilidade da imagem de ser uma ofensiva.

Um Provedor 8 pode também identificar quadros similares na animação, que estão a uma distância de um quadro entre si. A identificação pode ser baseada na observação de que as animações que têm algum tipo de efeito flash têm quadros similares dispersos com quadros nos quais aparece um novo objeto na imagem, ou a cor de um objeto presente na imagem varia. Ambas dessas podem ser consideradas essencialmente como variações abruptas nas cores entre os quadros.

O Provedor 7 pode primeiro identificar quadros similares com um quadro entre eles, e depois comparar o quadro disperso com um dos quadros similares. A variação de contrates por pixel pode ser computada, como pode a freqüência na qual essa variação ocorreu. A fórmula de diferença de cor CIE 1976 pode ser usada para medir a diferença de cor percebida entre duas imagens.

O algoritmo pode ser melhorado por consideração de fatores, tais como a distribuição espacial de cores, as variações de cor ao longo das bordas, o mascaramento e outros idiossincrasias do Sistema Visual Humano (HVS).

2. Imagens bamboleando: As propagandas bamboleando são animações com cada quadro contendo o mesmo objeto, mas em posições ligeiramente diferentes, dando a impressão de um objeto bamboleando. Em muitas dessas propagandas, pode parecer que os quadros usam de fato a mesma imagem contra um fundo constante, mas em diferentes desvios de páginas. Essas propagandas podem ser detectadas fazendo-se uma comparação pixel a pixel. Se os quadros são simplesmente translações de uma única imagem, os coeficientes de transformada de Fourier rápida (FFT) dos quadros são os mesmos. Outra técnica é comparar as imagens por alinhamento das bordas do objeto contido na imagem.

Para animações bamboleando, que são criadas por variação, os deslocamentos de páginas da mesma imagem podem ser facilmente detectadas por comparação dos quadros efetivos usados na animação. Outras propagandas bamboleando consistem em quadros que são imagens diferentes dispostas no deslocamento de página (0,0). Em algumas dessas propa-

gandas, o objeto bamboleando é disposto contra um fundo transparente. Nesse casos, cada quadro pode ser recortado para remover esses pixels transparentes, e depois comparar as imagens recortadas.

- As técnicas baseadas nas transformadas de Fourier rápidas
- 5 (FTT) também executam bem as animações bamboleando. Os coeficientes de FFT para todos os quadros podem ser computados. Depois, a soma da diferença absoluta (SAD) dos coeficientes pode ser computada. Para as imagens bamboleando, essa diferença fica mais próxima de zero, pois os quadros envolvem basicamente translação. Para deslocar o efeito do nível
- 10 de iluminação na imagem, o provedor pode aplicar o limiar em SAD relativo à soma absoluta dos coeficientes de um dos quadros. Essa técnica pode ser tornada mais robusta por uso da transformada de Mellin - Fourier.

- Para detectar objetos em estrobo e móveis em geral, o Provedor 8 pode computar os vetores de movimento e checar se há um grupo deles
- 15 apontando na mesma direção. O Provedor 8 pode usar um algoritmo de vetor de movimento baseado em bloco. Para capturar os aspectos suficientemente bem em um bloco, grandes blocos de tamanho 12 ou mais podem ser usados. Para evitar os problemas de alinhamento errado dos blocos, podem ser usados blocos sobrepostos. Outra técnica é usar os métodos para rastreamento do fluxo óptico na biblioteca OpenCV. Outra técnica seria traçar os movimentos do objeto em flash pelas transformação, que são aplicadas nos objetos nos quadros subseqüentes. Se um objeto é transladado ou escalonado por um grau significativo entre os quadros, isso vai indicar a presença de estrobo.
- 20

- 25 **Imagens estáticas:** O Provedor 8 pode desejar checar as baixas qualidades de imagem e os problemas de layout nas imagens estáticas.

1. Problemas de layout: O Provedor 8 pode extrair a borda da imagem e checar se a borda tem um limite de canto agudo ou ocupa uma grande parte da imagem. As bordas da caixa de limite da imagem pode ser
- 30 primeiro identificadas por varredura ao longo de cada borda da imagem e descoberta da primeira linha paralela àquela borda que não é igual a ela. Essas linhas constituem a caixa de limite da imagem. Depois, um cheque

para se as bordas da caixa de limite correspondem às bordas na imagem original pode ser feito por computação do gradiente ao longo das bordas. Se a imagem tiver bordas comparáveis com a caixa de limite, ou se a caixa de limite é muito pequena, comparada com a imagem original, a imagem pode

5 ser classificada como tendo problemas de layout.

2. Qualidade de imagem: Uma classificação de qualidade sem referência de base pode ser usada. Nesse tipo de classificação, não há qualquer imagem "ideal", para comparar a qualidade da dada imagem. Sob essa abordagem exemplificativa, o Provedor 8 pode determinar alguma métrica de qualidade de imagem (IQMs), de modo a avaliar a dada imagens nessa IQMs. Também, as técnicas OCR podem ser empregadas para extrair texto de imagens. Além da extração de texto, pode ser também usado para determinar a qualidade de imagem. Para as propagandas que são supostas conter texto, se o OCR não pode decifrar qualquer texto significativo das imagens, o documento pode ser classificado e/ou rejeitado como tendo uma baixa qualidade de imagem.

Um Provedor 8 pode também comparar os parâmetros especificados de modo a detectar as imagens intencionadas para rastrear um usuário 10. Algumas propagandas podem conter caixas de texto, menu de opções, e botões que não são funcionais e estão apenas presentes como imagens. Um Provedor 8 pode usar técnicas de detecção de borda/canto, para detectar a presença desses objetos nas imagens.

Se um documento for indicado ou de outro modo não aprovado, com base nas suas características (por exemplo, se a propaganda contém 25 ações indesejáveis, tais como oscilação e bamboleio), o documento pode ser passado para um ou mais avaliadores para revisão humana. O documento pode ser aprovado posteriormente, após vários revisores humanos aprovarem o documento.

No bloco 380, o documento, as classificações, e o status de aprovação ou rejeição são passados, por exemplo, para o Provedor 8.

Aqueles versados na técnica vão considerar que os atos descritos podem ser conduzidos por hardware, software ou uma combinação de-

les, com ou sem intervenção humana, como pode ser incorporado em um ou mais sistemas de computação, tal como o sistema servidor 2 acoplados a entidades, tais como provedores, avaliadores, bases de dados e usuários finais 10. Além disso, deve-se considerar que nem todos os blocos devem 5 ser executados. Por exemplo, em uma concretização, o método pode começar no bloco 310 e terminar no bloco 370. Também, não é necessário que a ou as ação ou ações de cada bloco sejam conduzidas na ordem mostrada na figura 3. Qualquer ordem de desempenho pode ser considerada.

10 A figura 3b é um fluxograma ilustrando um método exemplificativo para classificação, ou aprovação ou rejeição, de uma propaganda piscante, de acordo com uma concretização da invenção.

No bloco 300, um documento de propaganda piscante pode ser recebido. Por meio de exemplo, a propaganda piscante é, nesse exemplo, um jogo, no qual o usuário 10 deve usar um mouse para jogar o jogo.

15 No bloco 310, o documento pode ser processado (por exemplo, automaticamente). As cadeias _xmouse e _ymouse são checadas no arquivo desmontado. As cadeias são encontradas e o documento é classificado como uma propaganda ofensiva ou rejeitado.

20 No bloco 320, o documento, a classificação e/ou o status de rejeição são passados, por exemplo, para o Provedor 8.

A figura 3c é um fluxograma ilustrando um método exemplificativo para aprovar ou rejeitar um documento GIF animado piscante, de acordo com uma concretização da invenção.

25 No bloco 300, um documento GIF animado pode ser recebido. Por meio de exemplo, o documento GIF animado é uma animação piscante.

No bloco 310, o documento pode ser processado (por exemplo, automaticamente). Os retardos entre os quadros é checado e verificado como sendo de 0,2 segundo entre todos os quadros. Esse retardo determinado é depois comparado contra um parâmetro comparativo de um retardo de 0,5 30 segundo. Em virtude do retardo determinado ser verificado como sendo inferior ao parâmetro comparativo, o documento é verificado como sendo uma propaganda piscante e é classificado como uma propaganda piscante ou

rejeitado.

No bloco 320, o documento, a classificação e/ou o status de rejeição são passados, por exemplo, para o Provedor 8.

A figura 3d é um fluxograma ilustrando um método exemplificativo para classificação, ou aprovação ou rejeição, de um documento de imagem estática, de acordo com uma concretização da invenção.

No bloco 300, uma imagem estática pode ser recebida. Por meio de exemplo, a imagem estática é uma imagem de uma lata de cola, com uma borda circundante.

No bloco 310, o documento pode ser processado (por exemplo, automaticamente). A borda é extraída da imagem e o percentual da imagem ocupado pela borda é determinado. Verificou-se que a borda ocupa 50% da imagem. O valor de 50% é comparado contra o parâmetro comparativo de 30%. Em virtude do valor determinado (50%) ser maior do que o parâmetro comparativo (30%), o documento é verificado como tendo um problema de layout, e o documento é classificado como propaganda ofensiva ou rejeitado.

No bloco 320, o documento, a classificação e/ou o status de rejeição são passados, por exemplo, para o Provedor 8.

Interface de usuário ilustrativa e resultados

A figura 4a ilustra um documento de propaganda de imagem exemplificativo 1, de acordo com uma concretização da invenção. A propaganda de imagem 1 pode compreender uma propaganda de título, outra propaganda que pode ser exibida em uma página da rede, ou outra propaganda gráfica que pode ser exibida por intermédio de um meio eletrônico. A propaganda de imagem 1 mostra imagens específicas 1A, texto de propaganda 1B e uma ou mais ligações 1C. As imagens específicas 1A compreendem um projetor de filme e imagens de um DVD, VHS e um recipiente de filme digital. O texto 1B compreende uma mensagem de propaganda, por exemplo, uma descrição de um produto ou serviço, uma sugestão a um cliente potencial, e/ou outro texto de propaganda. A ligação 1C pode compreender uma ligação a outro documento, tal como a página da rede de propaganda URL (ou parte dela). Por exemplo, a ligação 1C pode compreender

uma ligação de hipertexto embutida, e a ligação embutida pode ser associada com a ligação 1C exibida na imagem do documento 1. Em algumas concretizações, a seleção (por exemplo, o clique) no URL exibido ou outra ligação, enquanto visualizando o documento 1 (por exemplo, em um navegador da rede) pode direcionar o mecanismo do espectador para os documentos de visualização (por exemplo, um navegador da rede) para o conteúdo associado com a ligação (por exemplo, a página da rede da propaganda).

O documento pode ter vários tipos de conteúdo. Por exemplo, o documento pode ter palavras, imagens, sons e outras informações, bem como funções ou programas, que podem produzir dinamicamente palavras, imagens, sons e outras informações. Cada documento pode ter diferentes graus de vários tipos de conteúdo, tais como conteúdo sexual, conteúdo de violência, conteúdo relativo a drogas ou álcool, conteúdo financeiro, conteúdo relativo a adultos, conteúdo relativo a crianças e outro tipo de conteúdo.

O documento 1 pode exibir explicitamente a ligação 1C. Também, a ligação 1C pode ser embutida no documento (por exemplo, na programação do documento) ou em uma parte dele, de modo que a ligação 1C não fique visível. No presente relatório descritivo, a seleção do (por exemplo, clicar no) documento 1, uma imagem 1A, texto 1B, ou outra parte do documento pode direcionar o mecanismo de visualização do documento do usuário 10 para o ou os documentos. O próprio documento 1, as imagens 1A e o texto 1C podem compreender uma ou mais ligações 1C. Por exemplo, uma propaganda que apresenta uma pluralidade de produtos pode compreender uma ligação para cada produto, em que a seleção de (por exemplo, clicar em) uma imagem 1A, ícone 1A ou texto 1B, relativo a um produto específico, pode direcionar um navegador da rede (ou outro mecanismo de visualização de documento) a uma página em um sítio de comerciante associado com o produto específico (ou a outro documento). O documento 1 pode mudar rapidamente as cores, envolver fluxo contínuo de áudio ou vídeo não iniciado pelo usuário 10, ou outras ações que podem ser consideradas pelo Provedor 8 como sendo ofensivas ou de outro modo indesejáveis.

O documento pode ter vários tipos de conteúdo. Por exemplo, o

documento pode ter palavras, imagens, sons e outras informações, bem como funções ou programas, que podem produzir dinamicamente palavras, imagens, sons e outras informações. Cada documento pode ter diferentes graus de vários tipos de conteúdo, tais como conteúdo sexual, conteúdo de 5 violência, conteúdo relativo a drogas ou álcool, conteúdo financeiro, conteúdo relativo a adultos, conteúdo relativo a crianças e outro tipo de conteúdo.

A figura 4b ilustra vários quadros seqüenciais 10, 20, 30, 40, 50, 60, 70, 80, 90, 110, 120 de um documento de propaganda de vídeo exemplificativo 500 (que é uma propaganda em bamboleio), de acordo com uma 10 concretização da invenção. Os quadros 10, 20, 30, 40, 50, 60, 70, 80, 90, 110, 120 ilustram uma lata de cola 11, 21, 31, 41, 51, 61, 71, 81, 91, 111, 121, que é translada a diferentes posições dentro dos quadros 10, 20, 30, 40, 50, 60, 70, 80, 90, 110, 120, de modo que quando os quadros são exibidos seqüencialmente, a lata de cola 11, 21, 31, 41, 51, 61, 71, 81, 91, 111, 15 121 parece bambolear no documento de propaganda de vídeo 500. O quadro 10 ilustra uma lata de cola 11 em pé, o quadro 20 ilustra uma lata de cola ligeiramente inclinada para a direita, o quadro 30 ilustra uma lata de cola 31 novamente em pé, o quadro 40 ilustra uma lata de cola 41 inclinada ligeiramente para a esquerda, e a seqüência começa de novo no quadro 50. A 20 invenção vai detectar essa propaganda como uma ofensiva e, em uma concretização, sinalizar o documento de modo a informar para o Provedor 8 que a propaganda precisa ser examinada. Em outra concretização, a invenção vai rejeitar automaticamente a propaganda.

A figura 4c ilustra um documento de propaganda de imagem exemplificativo 1 (que é uma propaganda estática), de acordo com uma concretização da invenção. A propaganda de imagem 1 mostra uma imagem de uma lata de cola 2 e uma borda 6 em torno da imagem. A propaganda da imagem 1 tem um layout inferior; especificamente, o dimensionamento da propaganda resulta em uma grande área sendo ocupada pela borda 6 da 30 propaganda 1. Uma concretização da invenção vai detectar esse problema de layout e informar ao Provedor 8 que a propaganda precisa ser examinada. Em outra concretização, a invenção vai rejeitar automaticamente a pro-

paganda.

A figura 5 ilustra um documento eletrônico 1, de acordo com uma concretização da invenção. A figura 5 ilustra um documento proporcionando, em um monitor, ao Provedor 8, informações de classificação e aprovação / rejeição relativas a uma propaganda. A propaganda do documento 3 sendo classificada e aprovada ou rejeitada é proporcionada ao Provedor 8. As classificações proporcionadas incluem uma classificação de conteúdo 5, uma classificação de ação 7, uma classificação de ligação 9, uma classificação agregada 11 e o status de aprovação / rejeição 13.

Deve-se entender que o servidor, os processadores e os módulos descritos no presente relatório descritivo podem executar as suas funções (por exemplo, comparando um documento a parâmetros específico e/ou determinando as informações de classificação) automaticamente ou por meio de um sistema automático. Como usado no presente relatório descritivo, o termo "automaticamente" se refere a uma ação sendo conduzida por qualquer processo executável por máquina, por exemplo, um processo que não requer intervenção ou introdução humana.

As concretizações da presente invenção não são limitadas em escopo pelas concretizações específicas descritas no presente relatório descritivo. Por exemplo, embora muitas das concretizações descritas no presente relatório descritivo tenham sido descritas com referência a propagandas de imagem, os princípios aqui são igualmente aplicáveis a outros documentos, tais como websites. De fato, várias modificações das concretizações da presente invenção, além daquelas descritas no presente relatório descritivo, vão ser evidentes para aqueles versados na técnica da descrição precedente e dos desenhos em anexo. Desse modo, essas modificações são intencionadas para se encaixarem dentro do escopo das reivindicações em anexo apresentadas a seguir. Além disso, embora as concretizações da presente invenção tenham sido descritas no presente relatório descritivo, no contexto de uma implementação particular em um meio físico particular para um fim especial, aqueles versados na técnica vão reconhecer que a sua utilidade não é limitada a isso, e que as concretizações da presente invenção podem

ser implementadas benéficamente em qualquer número de meios físicos para qualquer número de finalidades. Conseqüentemente, as reivindicações apresentadas abaixo devem ser consideradas em vista das amplitude e es-
pírito totais das concretizações da presente invenção, como descrito no pre-
5 sente relatório descritivo.

REIVINDICAÇÕES

7. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que carrega um ou mais documentos de vídeo e um ou mais documentos de áudio, sem iniciação 5 pelo usuário.
8. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que obtém dados a partir de fontes diferentes do que o próprio documento.
- 10 9. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que abre conexões de rede.
- 15 10. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que executa repetidamente instruções em um número maior do que aquele previamente designado para comparação.
11. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que inclui geração de número aleatório simulado.
- 20 12. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a operação de determinação do código compreende determinar se o documento compreende código de computador que requer hardware que é substancialmente idêntico a um ou mais tipos de hardware 25 que foram pré-determinados para serem inaceitáveis.
13. Método, de acordo com a reivindicação 12, **caracterizado pelo fato de que** o um ou mais tipos de hardware compreende um mouse 30 de computador.
14. Método, de acordo com a reivindicação 12, **caracterizado pelo fato de que** o um ou mais tipos de hardware compreende um controla-

dor de jogo.

15. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende a etapa de comparar uma imagem do documento com um ou mais valores de qualidade de imagem que foram pré-determinados para serem aceitáveis.

16. Método, de acordo com a reivindicação 15, **caracterizado pelo fato de que** o um ou mais parâmetros de qualidade compreende um grau de definição de uma borda limite de um conteúdo visual do documento.

17. Método, de acordo com a reivindicação 15, **caracterizado pelo fato de que** o um ou mais parâmetros de qualidade compreende uma extensão que uma borda limite de um conteúdo visual do documento ocupa um conteúdo visual do documento.

18. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende a etapa de comparar um quadro de uma imagem da taxa de contraste de alteração do documento com uma taxa de contraste de alteração que foi pré-determinada para ser aceitável.

19. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende a etapa de comparar uma taxa de mudança de cor de um quadro de uma imagem com uma taxa de mudança de cor que foi pré-determinada para ser aceitável.

20. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende a etapa de comparar uma taxa na qual uma imagem de um documento translada com uma taxa de translação que foi pré-determinada para ser aceitável.

25 21. Sistema computadorizado para aprovação de um documento **caracterizado pelo fato de que** compreende:

um módulo de análise que opera em um ou mais processadores para analisar o conteúdo de um documento para identificar um ou mais porções visual, textual e de áudio; e

30 um módulo de determinação que opera no um ou mais processadores para determinar se o documento contém uma porção que é substancialmente idêntica a um ou mais de um segmento visual, um segmento de

áudio ou um segmento textual que foi pré-determinado para ser inaceitável.

22. Método implementado por computador de classificação de um documento **caracterizado pelo fato de que** compreende as etapas de:

5 armazenar o documento em memória de computador;

analisar o conteúdo do documento usando um ou mais processadores para identificar uma ou mais porções visuais, textuais ou de áudio;

determinar, usando um ou mais processadores, se o documento contém uma porção que é substancialmente idêntica a um ou mais de um segmento visual, um segmento de áudio ou um segmento textual que é pré-
10 determinado para ser inaceitável; e

15 classificar o documento com base no resultado da determinação.

23. Método, de acordo com a reivindicação 22, **caracterizado pelo fato de que** ainda compreende a etapa de exibir a classificação do documento.

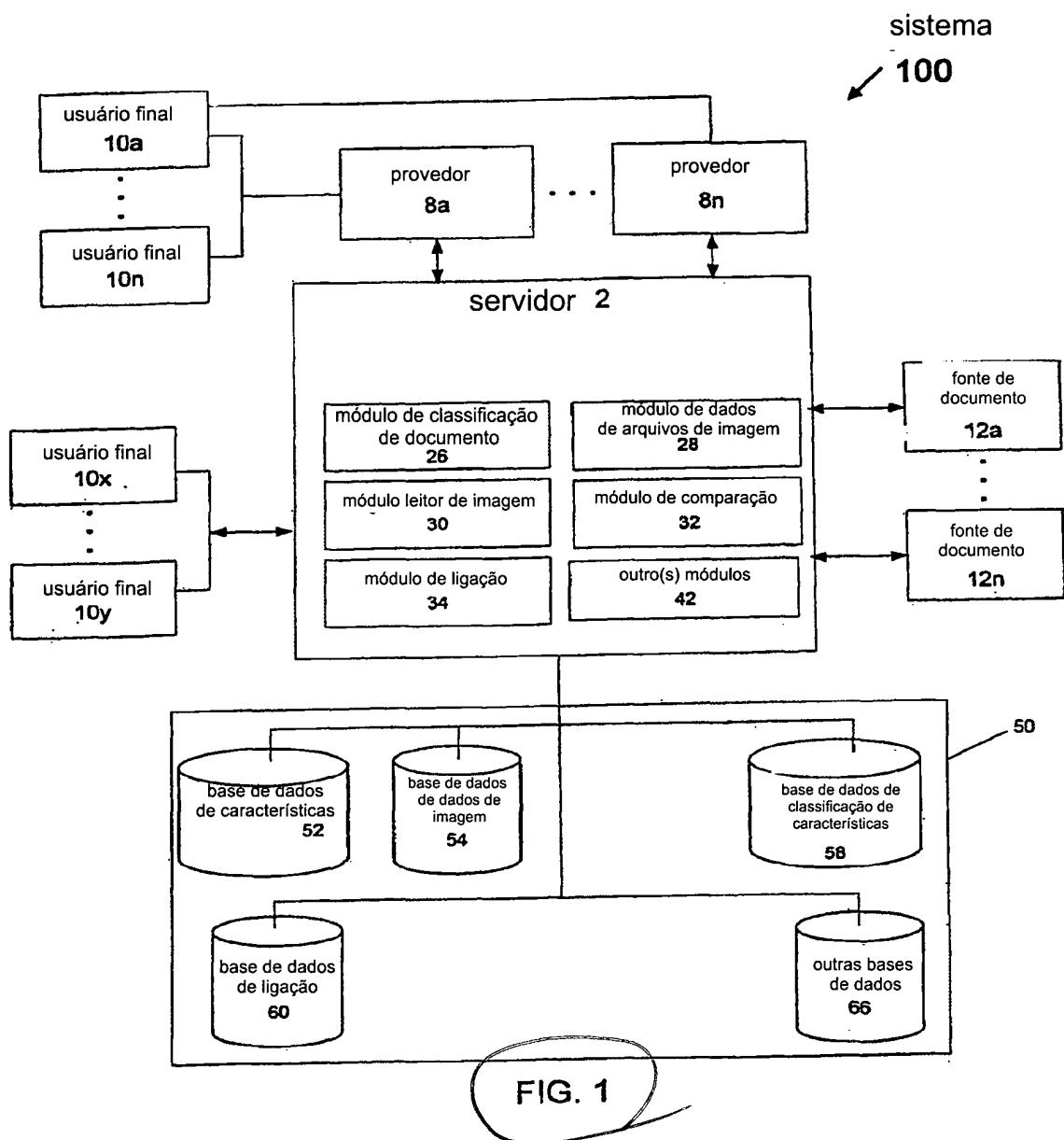
15 24. Método implementado por computador de classificação de um documento **caracterizado pelo fato de que** compreende as etapas de:

5 armazenar o documento em memória de computador;

analisar o conteúdo do documento usando um ou mais processadores para identificar uma ou mais porções visuais, textuais e de áudio;

20 determinar, usando um ou mais processadores, se o documento contém uma porção que é substancialmente idêntica a um ou mais de um segmento visual, um segmento de áudio ou um segmento textual que foi pré-determinado para ser inaceitável; e

25 indicar o documento para revisão com base na determinação.



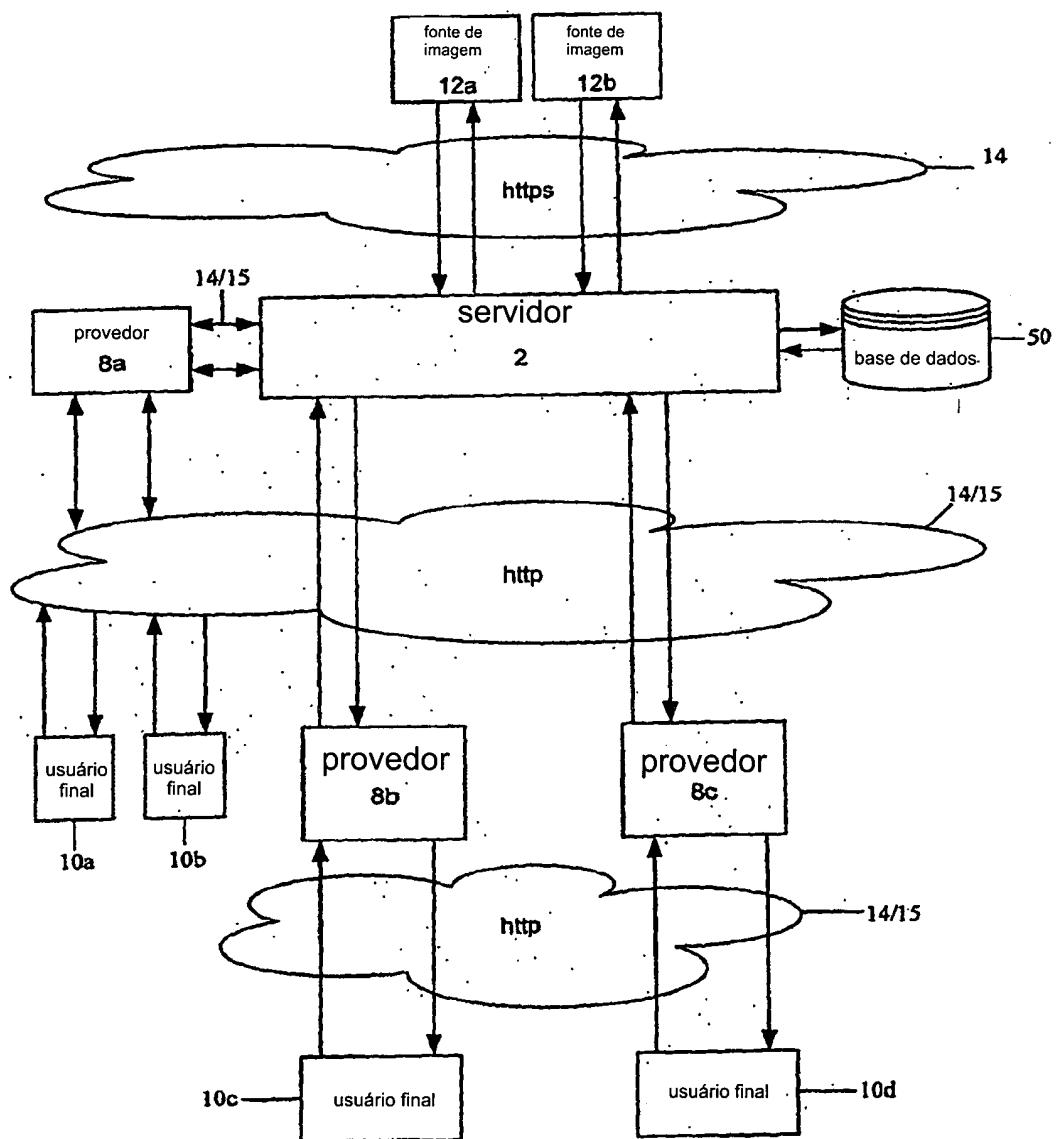


FIG. 2

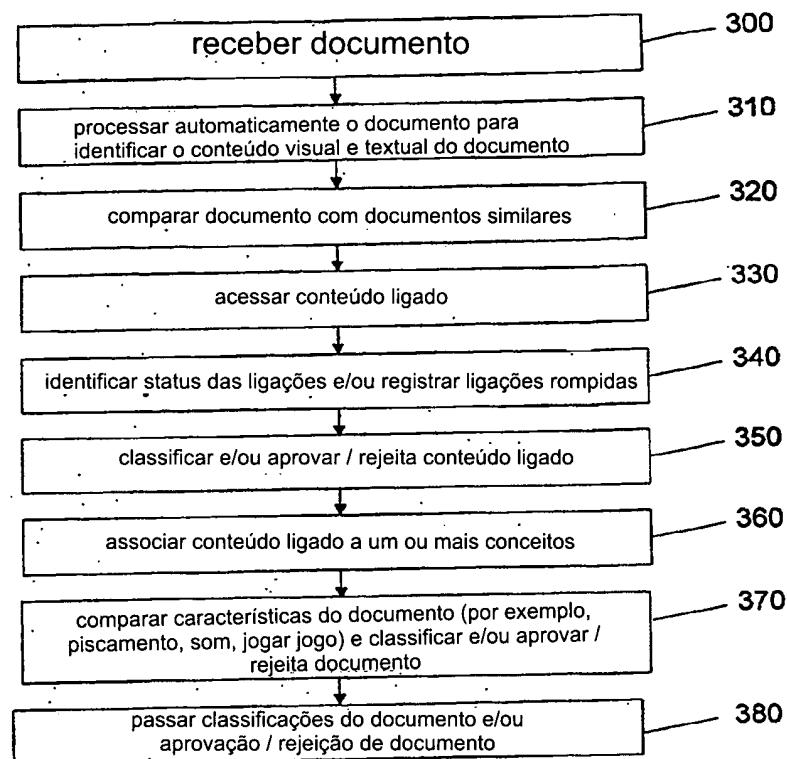


FIG. 3a

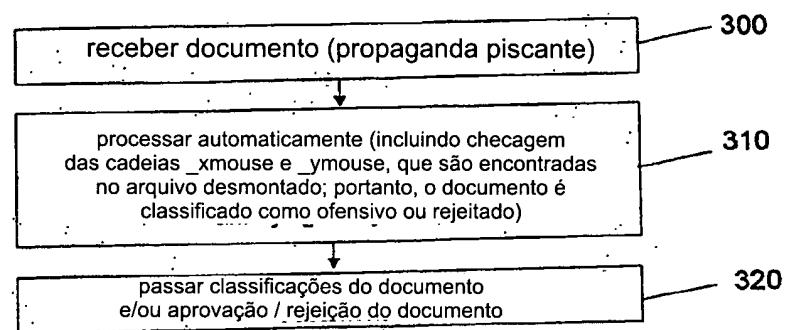


FIG. 3b

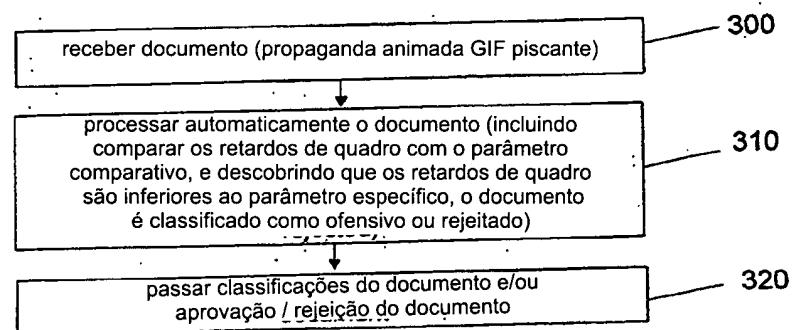


FIG. 3c

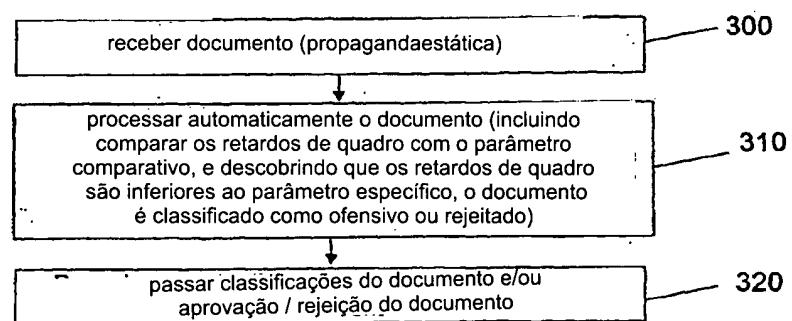


FIG. 3d

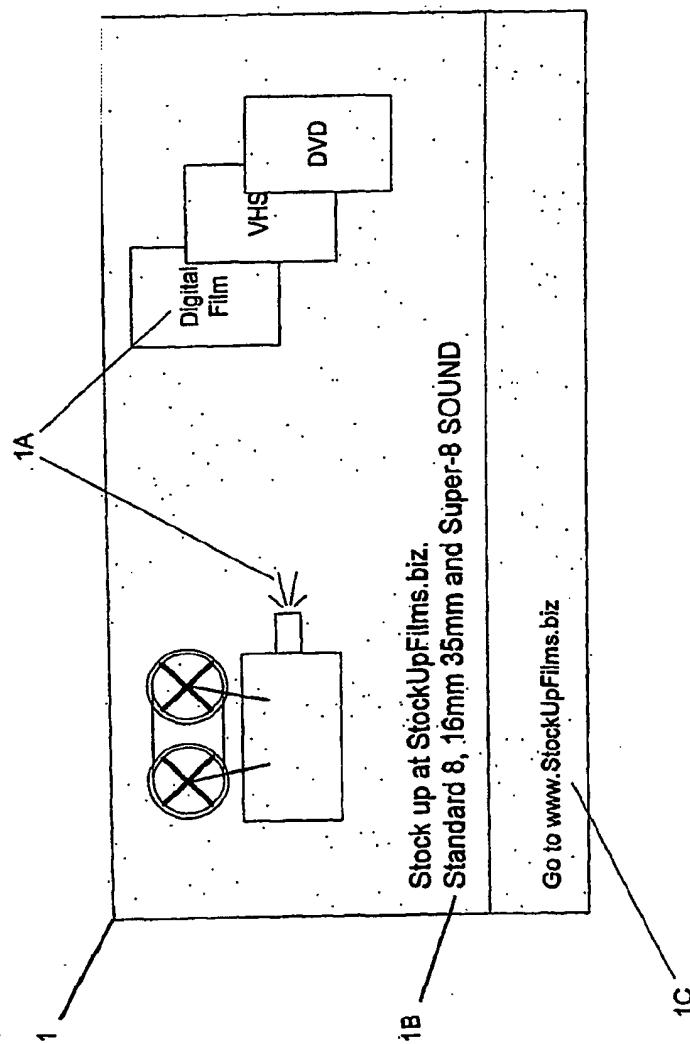


FIG. 4a

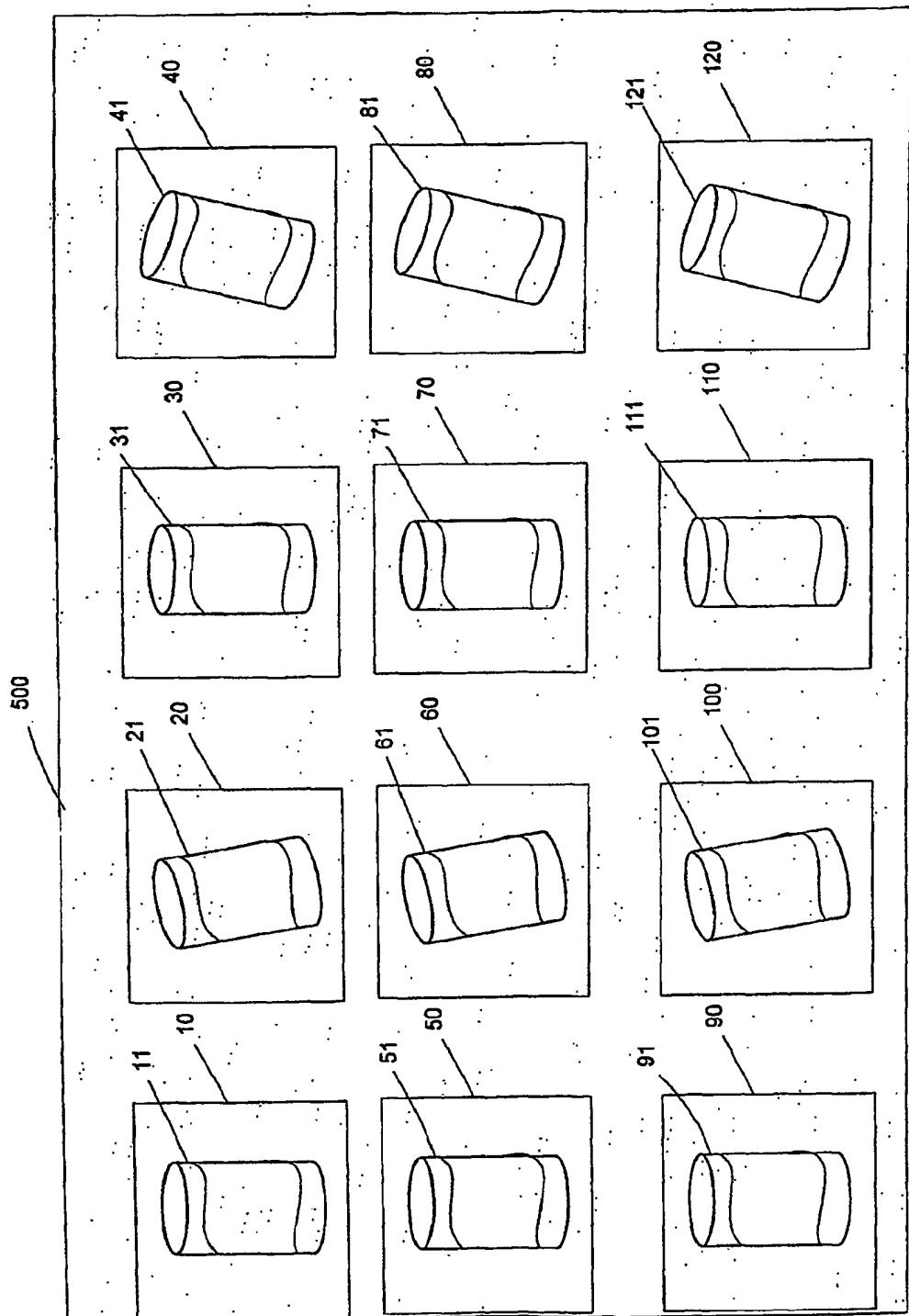


FIG. 4b

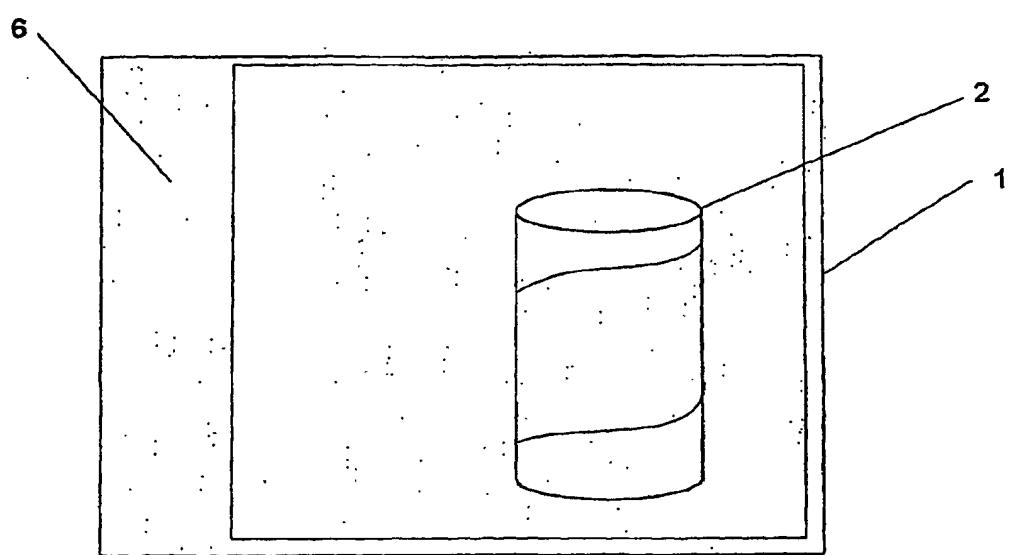


FIG. 4c

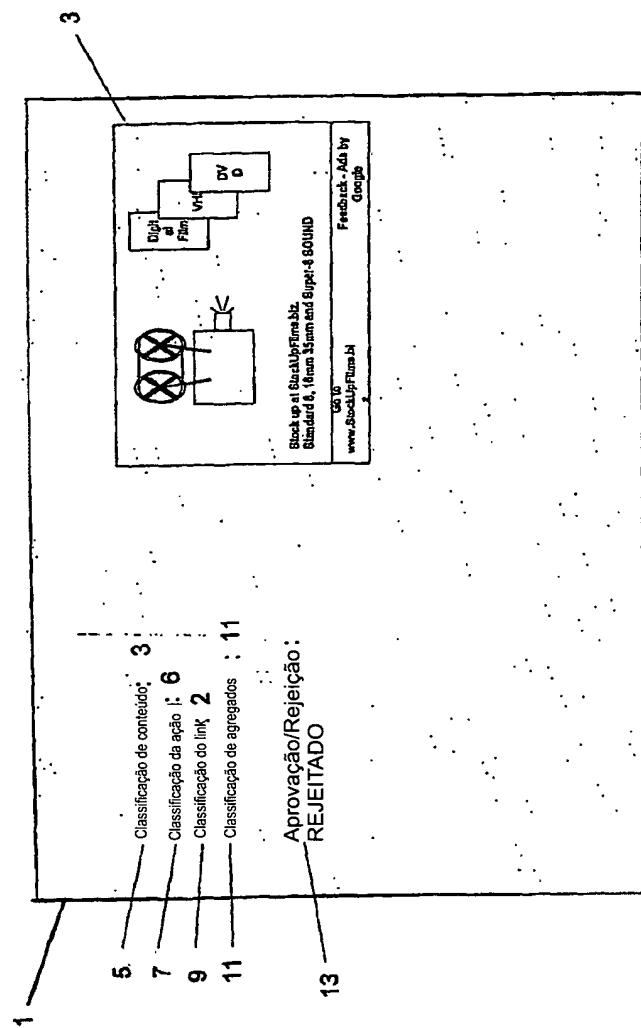


FIG. 5

RESUMO

Patente de Invenção: "**MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA COMPUTADORIZADO DE APROVAÇÃO DE UM DOCUMENTO E MÉTODO IMPLEMENTADO POR COMPUTADOR DE CLASSIFICAÇÃO DE UM DOCUMENTO**".

A presente invenção refere-se a um sistema e um método para avaliar documentos para aprovação ou rejeição e/ou classificação. O método compreende comparar o documento a um ou mais critérios, determinando se o documento contém um elemento, que é substancialmente idêntico a um ou mais de um elemento visual, um elemento de áudio ou um elemento de texto, que é determinado como sendo desagradável.