



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

(11) PI 0001662-4 B1

(22) Data do Depósito: 20/04/2000

(45) Data de Concessão: 17/11/2015

(RPI 2341)



(54) Título: SISTEMA DE TRANSMISSÃO INTERATIVA

(51) Int.Cl.: H04N 7/20; H04L 12/66

(30) Prioridade Unionista: 23/04/1999 US 60/130.676, 27/03/2000 US 09/536.276

(73) Titular(es): HUGHES ELECTRONICS CORPORATION

(72) Inventor(es): ASHA VELLAIKAL, YONGGUANG ZHANG, SON K. DAO

SISTEMA DE TRANSMISSÃO INTERATIVA

Campo técnico

A presente invenção refere-se, em geral, a transmissão interativa e mais particularmente a sistemas para produzir transmissões interativas.

Técnica antecedente

Tradicionalmente, a televisão tem sido uma forma de diversão passiva, com modos de interatividade muito limitadores. Atualmente, não há meio pelo qual um usuário possa interagir com os sinais de transmissão para acessar informações referentes ao teor de transmissão ou acionar outras funções sofisticadas. Isto é especialmente inadequado na atual era da informação, onde tem havido uma explosão de informações sobre tópicos incontáveis em várias formas. O computador pessoal se tornou recentemente o dispositivo preferido para acesso a informações, principalmente devido à crescente popularidade e crescimento da Internet. Desse modo, o usuário tem tipicamente de assinar um provedor de serviço de informações ou ter acesso direto à Internet. Em tal ambiente, se o usuário estiver interessado em coletar informações relacionadas ao teor de televisão, buscas independentes sobre serviços de informações disponíveis devem ser conduzidas.

Para evitar este procedimento incômodo e enriquecer o teor de televisão, é necessário um sistema mais simples para interação do usuário com sinais de transmissão e transmissão direta de informação. A fusão prevista da televisão e do Computador Pessoal (PC) como uma única fonte para diversão e informações facilitará qualquer implementação da invenção apresentada aqui. Por exemplo, PCs que são capacitados a receber sinais de transmissão são atualmente disponíveis. Além disso, o "Web PC" relativamente barato é outro exemplo de um dispositivo que aumenta a televisão comum. Além destes dispositivos, as redes de comunicação capazes de transportar dados de largura de banda elevada logo estarão disponíveis em

casas, permitindo sistemas de transmissão interativa. Estes avanços em infra-estrutura exigem novos métodos que facilitarão e utilizarão a união entre diversão e dispositivos de acesso de informações.

5 A televisão interativa inclui várias aplicações diferentes como vídeo sob encomenda, compras e outros. Utilizando a televisão interativa os usuários terão total controle para permitir que eles especifiquem ações diferentes como a escolha de um filme a ser mostrado ou item a ser
10 comprado. Várias arquiteturas de sistema diferentes foram propostas para tais sistemas. Contudo, estas exigem avanços tecnológicos tremendos nas áreas de comunicações e rede, servidores de vídeo, etc. e um grande investimento em infra-estrutura como colocação de novos cabos para casas
15 individuais.

Linguagens de marcação como HTML foram utilizadas para enriquecer documentos de texto padrão em formato de hipertexto incluindo etiquetas que precedem e seguem porções lógicas do documento. Quando o documento de hipertexto é
20 analisado, as etiquetas de marcação podem ser separadas dos dados em volta e estas declarações de marcação contêm as informações como hiperlinks, âncoras e anotações que ajudarão as funções de hipertexto (como acesso de base de dados), esquemas de representação e modalidades de interface. Também
25 foram propostos hipermídia nos quais elementos como texto, processos gráficos, fala, imagens e animação podem ser comunicados entre redes juntos. Hipermídia também pode ser definido de forma solta como dados de multimídia com hiperlinks.

30 Os sinais de transmissão de televisão comuns contêm apenas informações de vídeo e áudio produzindo um meio muito passivo de diversão com métodos limitados de interatividade. Nos últimos anos, uma quantidade limitada de interatividade foi fornecida adicionando-se informações aos sinais de
35 transmissão. As comunicações Wink® incluem dados relevantes adicionais no sinal de televisão como estatísticas

esportivas, ou informações e filmes que podem ser ativados pelo telespectador. Intericast®, Hypercast® e certos outros sistemas inserem dados de web no fluxo de vídeo e áudio. Contudo, estes sistemas não permitem que os telespectadores interajam entre eles próprios e, conseqüentemente, assistir a um filme em casa nunca é semelhante a uma experiência de teatro devido à falta de interação com a audiência. Adicionalmente, todos estes sistemas resultam em múltiplos usuários solicitando os mesmos dados do servidor de web simultaneamente criando uma tremenda carga no servidor e na rede. Portanto, há necessidade de um sistema e método pelos quais os meios de transmissão possam ser tornados ativos e o usuário possa então interagir com os sinais.

Sumário da Invenção

Um aspecto da invenção proposta é a adição de um novo fluxo de marcação de sinal de transmissão (BSMS) aos sinais de transmissão. Este fluxo adicional contém dados que converterão os fluxos de vídeo e áudio em fluxos de hipermídia. Como documentos de hipertexto, estas informações adicionais podem conter anotação, links e outras informações úteis ao usuário que pode ser capacitado por simples cliques de um controle remoto especializado para iniciar alguma ação como a recuperação de documentos associados em um canal de comunicação. Este esquema também intercala naturalmente com o padrão de compressão baseado em MPEG-4.

Em um aspecto adicional da invenção, um método de transmissão compreendendo as etapas de:

identificar e especificar objetos selecionáveis de um sinal de transmissão,

gerar um fluxo de marcação contendo o objeto; e

transmitir um sinal de transmissão interativo compreendendo o fluxo de marcação.

Uma finalidade da invenção é transformar os sinais de transmissão em uma fonte mais rica de informações e permitir que o usuário interaja com os sinais de transmissão, desse modo permitindo uma experiência de assistência interativa.

Uma vantagem da invenção é que não exigirá maiores alterações de hardware na infra-estrutura existente. Além disso, não requer armazenagem excessiva na extremidade de usuário como é comum em muitos sistemas de televisão interativos. O

5 acréscimo que é necessário na largura de banda de dados adicional pelo novo fluxo de dados é muito pequeno. Além disso, o esquema é excepcionalmente compatível com o padrão de compressão MPEG-4, embora ainda possível utilizar os padrões MPEG-1 e MPEG-2.

10 A presente invenção provê a um telespectador um alto grau de interatividade em termos de comunicação com outros telespectadores bem como obter informações adicionais de outras fontes, como a Internet.

 Estas e outras características e vantagens da presente

15 invenção serão evidentes para aqueles versados no estado da técnica a partir da seguinte descrição detalhada de modalidades preferidas, tomadas juntamente com os desenhos em anexo.

Breve Descrição dos desenhos

20 A figura 1 é um diagrama de blocos de nível elevado de um modelo de transmissão atual.

 A figura 2 é um diagrama ilustrando um sinal de transmissão típico.

 A figura 3 é um diagrama de blocos de nível elevado de um modelo de transmissão interativo de acordo com a presente

25 invenção.

 A figura 4 é um diagrama ilustrando o sinal de transmissão incluindo o fluxo de marcação de sinal de transmissão de acordo com a presente invenção.

30 A figura 5 é um diagrama dos canais de dados de multidifusão e hipermídia.

 A figura 6 é um diagrama da arquitetura de sistema implementada por um serviço de transmissão baseado em satélite.

35 A figura 7 é um diagrama dos componentes básicos do sistema incluindo o módulo de vídeo/repetição, o módulo de

interatividade e o módulo de comunicação.

A figura 8 é um diagrama do quadro de vídeo e o mapa de imagem correspondente.

Descrição das modalidades preferidas

5 Na seguinte descrição, os mesmos numerais de referência são utilizados para identificar componentes idênticos. A seguinte descrição é dirigida a um sistema baseado em satélite. O ensinamento da presente invenção pode ser aplicado a sistemas de transmissão convencionais ou ser
10 incorporado em sistemas de transmissão convencionais.

Com referência à figura 1, um sistema de transmissão por satélite conhecido 10 é ilustrado. O sistema 10 inclui um satélite 12 que é utilizado para gerar sinais de transmissão 14 para os usuários 16. Como ilustrado, os
15 sinais são transmitidos do satélite 12 para os usuários. Nenhuma interação é realizada pelos usuários.

Com referência agora à figura 2, o sinal de transmissão conhecido 14 contém apenas informações suficientes para transmitir uma imagem de vídeo e áudio sincronizado. O sinal
20 de transmissão 14 consiste em um fluxo de nível de sistema 18, um fluxo de vídeo 20 e um fluxo de áudio 22. Tipicamente, os dados de vídeo e áudio estarão em um formato comprimido, enquanto o fluxo do sistema conterá as informações de sincronização. Alguns sinais também conterão
25 informações de legenda fechada.

Com referência agora à figura 3, um sistema de transmissão por satélite interativo 24, de acordo com a invenção, é ilustrado. O sistema 24 tem um satélite 26 capaz tanto de enviar como receber sinais para e dos usuários 28.
30 O satélite 26 gera sinais de transmissão 30 modificados daqueles acima. Os usuários geram um sinal de retorno 32 para comunicar com o satélite 26 e interagem com o restante do sistema como será descrito abaixo.

Com referência agora à figura 4, o sinal de transmissão
35 30 da presente invenção inclui o fluxo do sistema 18, fluxo de áudio 20 e fluxo de vídeo 22 como mostrado acima na figura

2. Um fluxo adicional, um Fluxo de Marcação de Sinal de Transmissão 36 (BSMS), também é transmitido juntamente com os outros fluxos 18, 20, 22. BSMS 36 é uma linguagem de computador que pode ser entendida por um dispositivo de
5 repetição como será descrito adicionalmente abaixo. BSMS 36 carrega informações que permitirão que os sinais de transmissão suportem interatividade.

O BSMS 36 contém informações para servir várias finalidades descritas abaixo:

10 (1) Intensificação de informações de vídeo e áudio: A finalidade do BSMS 36 é carregar informações que podem ser capacitadas pelo usuário para melhorar a experiência de assistência. Por exemplo, BSMS 36 pode conter informações correlatas referentes a uma cena específica que pode ser
15 ativada se desejado pelo usuário; uma cena de filme pode conter os nomes dos atores que aparecem na cena naquele momento. Legenda fechada pode ser considerada parte do fluxo de marcação, intensificando o fluxo de áudio pela provisão de um transcrito na tela do vídeo do que está sendo falado.
20 Desse modo, a linguagem de marcação deve ter provisões para incluir estas informações adicionais, ativadas opcionalmente pelo usuário.

(2) Identificação e especificação de objetos selecionáveis: A linguagem de marcação tem a capacidade de
25 identificar e especificar objetos selecionáveis na tela do vídeo. Isto é importante, visto que podem haver ações definitivas que são associadas a estes objetos que são ativados quando o usuário seleciona aquele objeto. Alguns exemplos de objetos selecionáveis são palavras chave faladas
30 bem como partes específicas da tela de vídeo.

(3) Inclusão de funções relacionadas à escolha de objetos selecionáveis: além de ser capaz de especificar objetos selecionáveis dos fluxos de mídia, o fluxo de
35 marcação também deve conter funções relacionadas à escolha destes objetos. Esta é uma característica significativa do fluxo de marcação. Da perspectiva de hipermídia, uma das

funções importantes é hiperlinks, liga-se a informações associadas de várias fontes. Quando o usuário seleciona um objeto, a função correlata é invocada e, no caso de hiperlinks, as informações associadas são recuperadas e exibidas.

BSMS 36 mistura naturalmente com o esquema de compressão baseado em MPEG-4 padronizado. O padrão de compressão MPEG-4 é um sistema para comunicar objetos audiovisuais (AV). Ao contrário de padrões de compressão anteriores como MPEG-1 e MPEG-2 que são independentes de teor e baseados em bloco, o padrão MPEG-4 utiliza uma sintaxe de codificação baseada em objeto que é mais adequada para aplicações interativas. MPEG-4 inclui sintaxe de codificação baseado em teor, edição baseada em teor, acesso baseado em teor e interação de usuário. Na arquitetura MPEG-4, um ou mais objetos AV e suas relações espacial-temporal são transmitidos de uma fonte para um decodificador MPEG-4. Na fonte, os objetos AV são protegidos contra erro, multiplexados e transmitidos a jusante. No decodificador, os objetos AV são desmultiplexados, corrigidos em erro, descomprimidos, compostos e apresentados ao usuário final. É possível para o usuário final interagir com as informações e desse modo o decodificador MPEG-4 possui a funcionalidade para fornecer meios para interação de usuário. Além disso, a arquitetura MPEG-4 permite capacidade de download de ferramentas de decodificação e desse modo fornece aos desenvolvedores de teor a flexibilidade para criar estruturas de fluxo de bits personalizadas para adequar-se a necessidades de aplicação específicas. MPEG-4 coloca importância nos conceitos de um objeto de vídeo (VO), camadas de objeto de vídeo (VOL), e planos de objeto de vídeo (VOP). VO e VOP correspondem a entidades no fluxo de bits que o usuário pode acessar e manipular. O conceito VO é naturalmente adequado a BSMS 36. Cada objeto na tela pode ser facilmente associado a uma ação específica, como ligação a um URL. Após seleção daquele objeto (que é sustentado pela

natureza interativa de MPEG-4), a função correspondente pode ser ativada como descrito anteriormente. O Comitê de MPEG-4 definiu a Linguagem de Descrição de Sintaxe MPEG-4 (MSDL), incluindo uma especificação de classes de objetos que será
5 utilizada para aplicações específicas. A classe de objeto AV consiste em um método público denominado manipulação, que pode ser utilizado para manipular interação de usuário. Este método pode ser utilizado para analisar o fluxo e chegar na ação correspondente. Fornecendo um campo de atributo às
10 classes MPEG-4, qualquer ação pode ser especificada como um dos atributos ou propriedades do objeto.

Ao contrário de MPEG-4, MPEG-1 e MPEG-2 são esquemas de codificação baseados em bloco, que não incorporam o conceito de um objeto de vídeo. Desse modo, a inclusão de dados de
15 BSMS 36 com tais formatos é mais difícil. Os dois esquemas contêm a camada de sistema, que é tradicionalmente utilizada para manter informações de multiplexação e regulagem. O fluxo do sistema tem a provisão de incluir dados privados, que podem ser utilizados para carregar o BSMS. Contudo,
20 mapeamento entre elementos de imagem diferentes e funções devem ser manipulados em outro lugar visto que não há conceito de objeto embutido nestes esquemas de codificação.

Multidifusão é utilizada para se referir a informações de qualidade de membros de grupo, pelo que um grupo pode
25 consistir em todos os telespectadores que assistem um canal específico. Em multidifusão baseada em IP, cada grupo de multidifusão é especificado por um endereço de multidifusão. É possível ter-se um endereço de canal de multidifusão estático associado a cada canal. Por exemplo, o canal 4 pode
30 ter um endereço de multidifusão de 224.0.0.4. Este canal de multidifusão é um canal de dados que formará o meio pelo qual os usuários pode se comunicar mutuamente (tipo bate-papo), enviar mensagens e solicitar informações adicionais, como dados de web. Além deste canal de dados, os sinais de vídeo
35 e áudio serão aumentados com BSMB 36, que fornecerá a funcionalidade para especificar links de interesse e outras

informações de intensificação. Por exemplo, utilizando BSMS 36, é possível associar-se um link na Internet, que é de relevância ao que está sendo exibido. Um usuário pode ativar este link para receber dados que são então entregues via o canal de multidifusão. Outros usuários que são interessados nos mesmos dados também receberão os mesmos embora apenas uma única cópia tenha sido transmitida. Isto melhora consideravelmente o redimensionamento do sistema com relação à largura de banda disponível. Com o uso do BSMS 36, o canal de vídeo/áudio é transformado em um canal semi-hipermídia.

Com referência agora à figura 5, em uma variante da invenção, o sinal de transmissão 30 é compreendido de dois sinais. Um canal de hipermídia 38 e um canal de multidifusão lógico 40. O canal de hipermídia 38 é uma transmissão em um sentido da fonte de transmissão para os usuários. O BSMS 36 está contido no canal de hipermídia 38. O canal de multidifusão lógico 40 é um canal interativo em dois sentidos, compartilhados conjuntamente por todos os telespectadores de um canal específico. O canal de multidifusão lógico 40 acompanha cada canal de televisão para permitir aos usuários interagirem entre eles próprios com retardo de resposta limitada (através de caching), sem carregar excessivamente a rede e enquanto mantendo o sistema redimensionável. O canal de multidifusão lógica 40 conecta, desse modo, os usuários ou outros usuários e usuários à Internet.

Com referência agora à figura 6, a arquitetura de um sistema de transmissão baseado em satélite 46 é ilustrada. O sistema 46 inclui um centro de uplink 48 e um satélite 50 que multiplexam e transmitem os canais diferentes incluindo os canais de dados de multidifusão, áudio e vídeo. Um proxy de Centro de Operações de rede (NOC) 52 cuida de solicitações de usuário e programação, e recupera informações solicitadas da Internet 54 e outras fontes de dados associadas 56. Um sistema receptor 58 tendo uma antena de recepção 59 na extremidade do cliente através de um dispositivo de controle

61 coleta os dados de transmissão desmultiplexa os mesmos e orienta os mesmos para os dispositivos de saída apropriados ou software, além de responder à interação de usuário e enviar solicitações e outros dados através da conexão de rede. Como exemplo, o sistema receptor 58 pode incluir uma tela de exibição 60 incluindo vários sinais de tela incluindo uma porção de controle 62, uma porção de bate papo 64, uma porção de World Wide Web 66 e uma porção de vídeo 68.

Os três elementos básicos da arquitetura do sistema são descritos abaixo:

(1) Multiplexação no Centro de Uplink 48: Em transmissão de televisão por satélite tradicional, há um centro de uplink que gera os fluxos MPEG correspondendo a canais diferentes, multiplexa os mesmos e transmite os dados via um satélite de transmissão digital. Na arquitetura atual, o centro de uplink multiplexa o fluxo de marcação juntamente com os sinais de vídeo e áudio bem como aceita dados interativos de tempo real do proxy NOC 52 (para cada canal) e transmite todos estes nos canais de dados apropriados. O canal de hipermídia é armazenado localmente ou chega em tempo real de outra fonte como com um evento ao vivo. Embora haja um único canal de multidifusão lógica associado a cada canal de televisão, vários canais de dados podem ser multiplexados juntos.

(2) Proxy NOC 52: O proxy NOC 52 é a essência do sistema e fornece a função de coordenação principal entre solicitações de usuário e entradas e o canal de multidifusão. Os usuários comunicam diretamente ao proxy NOC 52, através de uma conexão de discagem ou via um provedor de serviço de Internet. (Embora a comunicação se baseie provavelmente em um protocolo IP, isto não é exigido pelo sistema).

Podem haver dois tipos de dados de usuário, uma entrada simples como no caso de uma aplicação de bate papo ou uma solicitação para alguns dados como no caso de dados da Internet. Nos dois casos os dados de usuário são encapsulados como um pacote IP e transmitidos para o proxy

NOC 52. É obrigatório para todos os dados de usuário conter as informações de endereço de multidifusão para as quais os dados apropriados devem ser enviados. Isto pode ser transparente para o usuário solicitando ao sistema receptor que inclua automaticamente as informações de canal no pacote. 5 Desse modo, o proxy recebe os pacotes e determina o curso de ação. No caso de uma aplicação onde o usuário está apenas entrando dados (como em bate papo), o proxy simplesmente envia o pacote para o grupo de multidifusão apropriado via o centro de uplink. Qualquer filtração ou moderação pode ser 10 realizada neste ponto. No caso de solicitações, a operação é mais complicada. Considere o cenário onde usuários solicitam informações de um link específico que foi incluído no BSMS 36. É altamente provável que múltiplos usuários solicitarão os mesmos dados. Para manter o sistema redimensionável 15 (tanto em complexidade como em largura de banda), apenas uma cópia dos dados deve ser baixada do servidor apropriado e multiplamente difundida via o canal apropriado. Desse modo, o proxy tem de seguir o rastro de quais dados foram solicitados ou foram previamente enviados. No caso de dados 20 que já foram solicitados, o proxy pode desconsiderar a solicitação uma vez que os dados já são transmitidos no canal de multidifusão comum como resultado de solicitação de outro usuário. No receptor há um cache que armazena todos os dados 25 vindos do canal de multidifusão específico sintonizado. Isto melhora em muito o desempenho, visto que o dispositivo de usuário enviará uma solicitação apenas se não ver o item solicitado no cache. Para suportar usuários que acabaram de entrar em um canal sem um acúmulo de cache é importante que o proxy programe retransmissões de alguns dados (por exemplo 30 dados que foram transmitidos no minuto anterior).

(3) O Sistema Receptor 58: O sistema receptor 58 executa a função de desmultiplexar e descomprimir o canal de dados apropriados dos fluxos de vídeo e áudio multiplexados. 35 O desmultiplexador extrai os canais de vídeo e áudio bem como o canal de multidifusão apropriado e envia os mesmos para os

dispositivos apropriados. No caso do sinal de hipermídia, um analisador BSMS separará os fluxos de sinais de marcação dos sinais de vídeo e de áudio. Os sinais de vídeo e áudio comprimidos são então enviados para um módulo que descomprime os dados e os apresenta no dispositivo de exibição apropriado. O dispositivo de exibição neste caso tem que ser mais sofisticado do que simplesmente um monitor NTSC uma vez que não tem dados de exibição de outras aplicações como uma sessão de bate papo, browser de rede e assim por diante. As solicitações de usuário e interação são realizadas via um controle remoto especializado bem como teclado sem fio. Tem que haver um dispositivo de controle de interatividade que toma a entrada de usuário e executa a ação apropriada, decifra a parte apropriada do sinal de marcação ou envia a entrada de usuário via um dispositivo de comunicação. No caso mencionado por último, é necessário que o sistema tenha uma pilha de comunicação e um modem além de um microprocessador de finalidade geral ou especializado. Além disso, o sistema tem um cache para armazenar todos os dados de entrada para o canal que o usuário está atualmente assistindo. Também, o módulo receptor contém um meio pelo qual os pacotes podem ser filtrados com base nos endereços de multidifusão de modo que apenas pacotes destinados àquele canal específico são extraídos. Informações detalhadas sobre os componentes do sistema são fornecidas abaixo.

Com referência agora à figura 7, o sistema receptor pode ter três módulos principais, o módulo de repetição de vídeo/áudio 74, o módulo de interatividade 76 e o módulo de comunicação 78.

O módulo de repetição de vídeo/áudio 74 consiste em dois dispositivos, um dispositivo de descompressão e repetição 80 que é acoplado a um dispositivo de exibição 82. O dispositivo de repetição 80 recebe o fluxo de transmissão e subseqüentemente desmultiplexa e descomprime o canal selecionado pelo usuário. Os fluxos de vídeo e áudio descomprimidos (após serem sincronizados utilizando as

informações de nível do sistema) são então transmitidos para o dispositivo de exibição de vídeo/áudio 82. Este dispositivo de exibição 82 pode ser um monitor de televisão normal com alto-falantes. Também é possível que o dispositivo de exibição seja um monitor PC capacitado por 5 sinal de transmissão. Observe que até o presente este módulo tem o mesmo receptor de topo ajustado de funcionalidade em sistemas de transmissão atuais. Além das funções acima, o dispositivo de repetição 80 também passa o fluxo de marcação 10 para o dispositivo de controle de interatividade que faz parte do módulo de interatividade como será discutido abaixo. O dispositivo de repetição 80 também é capaz de aceitar instruções do módulo interativo para modificar as propriedades dos dados de áudio e vídeo sendo transmitidos 15 para o dispositivo de exibição. Por exemplo, o usuário pode querer exibir informações extras sobre a cena que está contida no fluxo de marcação, sobrepostas no topo da exibição atual. Esta solicitação é comunicada pelo módulo de interatividade ao dispositivo de repetição juntamente com as 20 informações relevantes.

O módulo de interatividade 76 consiste em dois componentes, o dispositivo de controle de interatividade 84 e dispositivo de controle de usuário 86. O dispositivo de controle de usuário 86 é um controle remoto que provê ao 25 usuário capacidades como seleção de um objeto de vídeo ou áudio. Uma ação especificada pelo usuário como a seleção de um objeto de mídia é primeiramente capturado pelo dispositivo de controle de interatividade 84 e o fluxo de marcação (entrada no dispositivo de controle de interatividade pelo 30 dispositivo de repetição) e a seguir analisado para localizar a função correspondente. Se a função envolver comunicação com o mundo externo, a solicitação é então passada para o módulo de comunicação 78. Se a ação envolveu informações adicionais na tela de repetição, a solicitação é então 35 passada para o dispositivo de repetição juntamente com as informações correspondentes. Desse modo, o módulo de

interatividade 76 é responsável por capturar as ações do usuário e enviar comandos de controle para o módulo de comunicação ou o módulo de repetição para iniciar respostas para as funções especificadas pelo usuário.

5 O módulo de comunicação 78 permite acesso remoto a múltiplas fontes de informação. Desse modo, este módulo é responsável principalmente por comunicar-se com o mundo externo e está ligado a um canal de comunicação. O dispositivo de comunicação de rede 88 recebe as solicitações para informações remotas do dispositivo de controle de
10 interatividade 90 e a solicitação é enviada através do canal de comunicação. Este dispositivo também controla os dados que foram recuperados em resposta à solicitação e os envia apropriadamente a um dispositivo de exibição interativo. O
15 dispositivo de exibição interativo 90 pode ser idêntico ao dispositivo de repetição, como no caso de um PC capacitado para transmissão. Além de um monitor para exibição, este dispositivo também pode conter periféricos associados, como teclado ou mouse para maior interatividade.

20 Em hipermídia, as informações são ligadas com mais informações de várias fontes diferentes. As informações correlatas podem ser de tipos de dados diferentes, por exemplo vídeo de um servidor de vídeo informações do World Wide Web (WWW), áudio de um repositório de multimídia. Por
25 exemplo, um clip das notícias diárias pode ter links com informações de transmissão de vídeo anteriores referentes ao mesmo tópico. Estas transmissões podem ser comprimidas e armazenadas em um servidor de vídeo em um site específico. A nova transmissão pode conter links para as informações de
30 transmissão antigas que podem ser recuperados pelo usuário se desejado. Similarmente, um anúncio pode conter links para o site WWW da companhia pelo que o usuário pode coletar mais informações referentes ao produto anunciado. Neste esquema, o produtor de material de transmissão pré-definiu os links de
35 informações. Sacrificando um pouco da flexibilidade, o esquema oferece ao usuário acesso a informações correlatadas

com muito pouco esforço.

Mapas de imagem são utilizados comumente em documentos de hipertexto atualmente, permitindo que URLs diferentes sejam ligados com partes diferentes de uma imagem. Em um modo similar, é possível mapear a tela de vídeo de transmissão de tal modo que para um determinado quadro, parte diferente da tela terá ações diferentes associadas à mesma. Desse modo, o fluxo de marcação deve conter métodos para especificar porções diferentes da imagem e um método para anotar uma ação juntamente com a seleção daquele objeto.

Com referência agora à figura 8, um exemplo mostra uma tela de exibição 60 que tem uma transmissão de notícias que contém uma âncora 92, uma sub-tela 94 que ilustra cenas esportivas e um contato intermitente financeiro 96 juntamente com o mapa de imagem correspondente 98 indicando os links 100 identificados para as porções relevantes da tela.

Alem do vídeo, o áudio também pode conter certas partes que podem ter links associados ao mesmo. Por exemplo, considere a seguinte seção de áudio de uma transmissão. "Hoje em Nova York, o Presidente Clinton se dirigiu ao Conselho de Segurança das Nações Unidas. " Qualquer um dos termos sublinhados neste áudio pode conter um link com uma fonte de informações correlata. O áudio pode ser um meio difícil para hiperlink. Um método para reconhecimento de fala é utilizar a facilidade de legenda fechada permitindo que palavras especificadas como por sublinhar, possam ser selecionadas pelo usuário.

Outra possibilidade para se utilizar informações de áudio como um hiperlink é utilizar o sinal de áudio para criar um conjunto de palavras-chave que pode ser enviado para um motor de busca para obter links correlatos. Nestas modalidades, pode não haver necessidade de se ter endereços de fonte de informações correlatas enviadas juntamente com o sinal de transmissão.

Em uma variante com um servidor de fluxo de marcação, o fluxo de marcação associado ao fluxo de transmissão não é

transmitido com os fluxos de vídeo, áudio e sistema. Em vez disso, o fluxo de marcação é armazenado em um servidor juntamente com um índice que associa unicamente o mesmo a um segmento de vídeo/áudio específico. Por exemplo, o índice
5 pode ser o número de canal e um carimbo de hora indicando quando a cena foi mostrada. Quando o usuário submete uma ação como a seleção de um objeto, a ação é transmitida através do canal de comunicação para o servidor em alguma forma e o servidor de fluxo utiliza esta informação para
10 localizar o arquivo e função de marcação apropriados. Se a função envolve a recuperação de informações correlatas, o servidor de fluxo de marcação solicita então ao provedor de fonte para enviar os documentos necessários para aquele usuário específico. Tal implementação envolverá modificações
15 mínimas nos fluxos de transmissão atuais. A desvantagem deste esquema é a de redimensionamento. A base de dados de fluxos de marcação será muito grande dependendo do número de canais e um único servidor terá de processar todas as solicitações.

20 Como na variante anterior, outra implementação não necessita também de um fluxo de marcação associado, contudo esta variante é mais limitada em sua funcionalidade. O usuário pode encontrar informações correlatas formulando palavras-chave para uma cena que podem ser então enviadas a
25 um motor de busca. O motor de busca pode retornar informações sobre os vários repositórios de informações que conterão dados correlatos. A série de palavras-chave pode ser criada analisando as informações de legenda fechada ou analisando o fluxo de áudio. Nos dois casos, o dispositivo
30 de controle de interatividade deve conter processadores necessário que podem analisar o áudio e os fluxos de legenda fechada. A vantagem desta abordagem é que as informações de link correlatas não são pré-definidas. Contudo, esta abordagem requer um motor de busca, potência de processamento
35 adicional para analisar áudio ou fluxos de legenda fechada e mais ação do usuário para obter as informações desejadas.

Além disso, este esquema não provê um modo para associar links pré-definidos com sinais.

Um componente importante do sistema é a geração do teor de fluxo de marcação que corresponde às funções associadas à
5 seleção de objetos de vídeo e áudio diferentes. Isto requer um elemento humano no circuito para determinar as associações de função com cada objeto. Preferivelmente, a seleção de funções é realizada automaticamente. Os métodos que se seguem facilitarão a automação desta parte dos sistemas.

10 A porção de áudio do fluxo de meio pode ser entrada em um motor de reconhecimento de fala, que pode transcrever a mesma em texto. Os sistemas de reconhecimento de fala atuais podem obter transcrição de tempo real, de alta confiabilidade. Técnicas de processamento de linguagem podem
15 ser empregadas na saída para gerar palavras-chave que são representativas do teor de cena. As palavras-chave podem ser utilizadas agora em um motor de busca para encontrar os melhores links para aquela cena específica. Desse modo, a combinação de reconhecimento de fala, processamento de
20 linguagem e buscas de palavra-chave apresenta um método para gerar teor de fluxo de marcação.

Embora seja difícil analisar vídeo para gerar informações de semântica, o vídeo pode ser não obstante utilizado para auxiliar no contexto da busca. Por exemplo,
25 existem informações em vídeo, que não estão contidas ou mencionadas no segmento de áudio. Informações como cores dominantes, trabalho de câmara e informações de movimento podem ser facilmente extraídas analisando a porção de vídeo do fluxo. Estas informações podem ser utilizadas para
30 aumentar a busca. Por exemplo, se o segmento de áudio se refere a um carro, utilizando as informações de cor da porção de vídeo, o sistema pode deduzir qual é a cor do carro. Desse modo, as informações dos dados de vídeo também podem ser utilizadas para auxiliar no processo de busca na produção
35 de teor de fluxo de marcação.

Os exemplos discutidos aqui são exemplares e o âmbito

da invenção não deve ser limitado a estes exemplos. Deve ser entendido que a descrição acima se refere a certas modalidades preferidas da invenção e que inúmeras variações e modalidades alternadas ocorrerão para aqueles versados no estado da técnica que podem ser feitas sem se afastar do espírito e âmbito da invenção como exposto nas reivindicações que se seguem.

REIVINDICAÇÕES

1. Sistema de transmissão interativa (10), **caracterizado** pelo fato de compreender:

5 um proxy de centro de operação de rede (52) acoplado à uma fonte de dados;

um terminal (48) para receber um fluxo de dados do referido centro de operação de rede e transmitir um sinal de transmissão interativa, o sinal de transmissão interativa incluindo uma pluralidade de canais de hipermídia e um canal
10 de multidifusão associado com cada um dos canais de hipermídia; em que

cada canal de hipermídia inclui um fluxo de vídeo e um fluxo de marcação;

o canal de multidifusão é juntamente compartilhado
15 por uma pluralidade de usuários vendo o canal de hipermídia através de um endereço comum a todos da pluralidade de usuários e o canal de multidifusão transmite informação identificada pelo fluxo de marcação e requerida por qualquer um da pluralidade de usuários para todos da pluralidade de
20 usuários;

o sinal de transmissão interativa é recebido em um receptor tendo um cache para armazenar a informação transmitida através do canal de multidifusão para todos da pluralidade de usuários; e

25 o canal de multidifusão e o canal de hipermídia não são os mesmos canais.

2. Sistema, de acordo com a reivindicação 1, **caracterizado** pelo fato de que ainda compreende um receptor compreendendo:

30 um módulo de repetição de vídeo/áudio (74),
um módulo de interatividade (76), e
um módulo de comunicação (78).

3. Sistema, de acordo com a reivindicação 2, **caracterizado** pelo fato de que o referido módulo de repetição
35 de vídeo/áudio compreende um dispositivo de descompressão e repetição de fluxo de transmissão (80).

4. Sistema, de acordo com a reivindicação 3, **caracterizado** pelo fato de que o referido módulo de repetição de vídeo/áudio (74) compreende um dispositivo de exibição de vídeo áudio (82) acoplado ao referido dispositivo de descompressão e repetição de fluxo de transmissão.

5. Sistema, de acordo com a reivindicação 4, **caracterizado** pelo fato de que o referido dispositivo de exibição de vídeo/áudio (82) compreende uma tela (60) tendo uma pluralidade de sinais de tela (62), (64), (66), (68).

6. Sistema, de acordo com a reivindicação 5, **caracterizado** pelo fato de que os referido sinais de tela são selecionados do grupo que consiste em uma porção de controle (62), uma porção de bate papo (64), uma porção de world wide web (66) e uma porção de vídeo (68).

7. Sistema, de acordo com a reivindicação 2, **caracterizado** pelo fato de que o referido módulo de interatividade (76) compreende um dispositivo de controle de interatividade (84) acoplado ao referido módulo de repetição de vídeo/áudio (74).

8. Sistema, de acordo com a reivindicação 7, **caracterizado** pelo fato de que o referido módulo de interatividade (76) compreende um dispositivo de controle de usuário (86) acoplado ao dispositivo de controle de interatividade (76).

9. Sistema, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o referido proxy de centro de operações de rede (52) recebe ação de usuário a partir de um da pluralidade de usuários compreendendo uma seleção de um objeto no fluxo de vídeo e gera o fluxo de marcação do mesmo.

10. Sistema, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o fluxo de marcação é responsivo a uma solicitação do usuário transmitida através do canal de multidifusão e recebida pelo proxy de centro de operação de rede.

11. Sistema, de acordo com a reivindicação 1, **caracterizado** pelo fato de que uma solicitação do usuário é

transmitida através do canal de multidifusão e informação responsiva à solicitação do usuário é recebida através do canal de multidifusão.

12. Sistema, de acordo com a reivindicação 11,
5 **caracterizado** pelo fato de que a solicitação do usuário para a informação é transmitida somente se a informação não está no cache.

13. Sistema, de acordo com a reivindicação 1,
caracterizado pelo fato de que:

10 o fluxo de vídeo compreende uma pluralidade de objetos de vídeo incluindo objetos selecionáveis codificados por um sintaxe de codificação com base em objeto; e

o fluxo de marcação inclui informação identificando e especificando os objetos selecionáveis.

15 14. Sistema, de acordo com a reivindicação 13,
caracterizado pelo fato de que o fluxo de marcação ainda compreende funções relacionadas aos objetos selecionáveis.

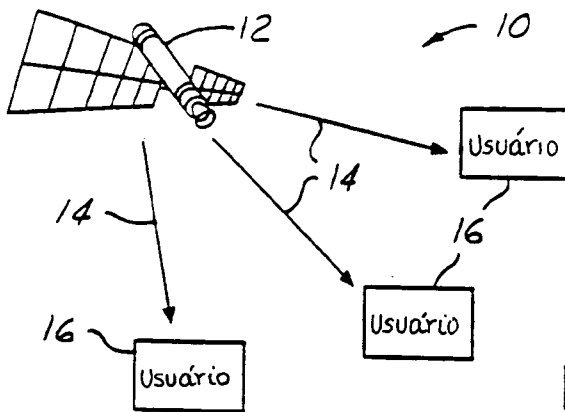


FIG. 1

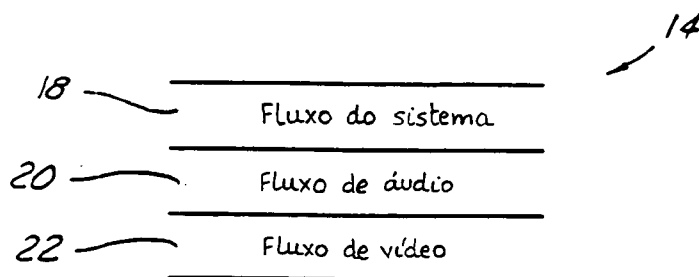


FIG. 2

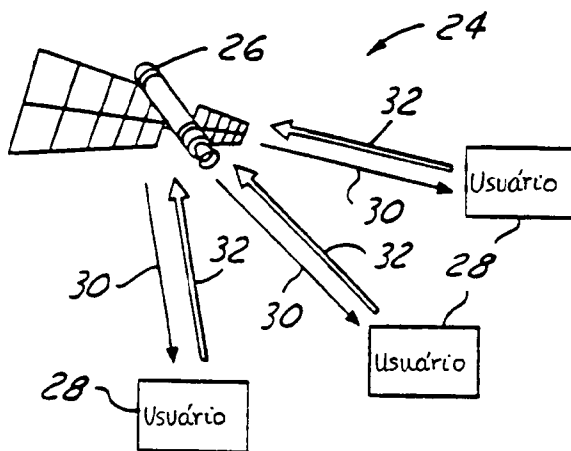


FIG. 3

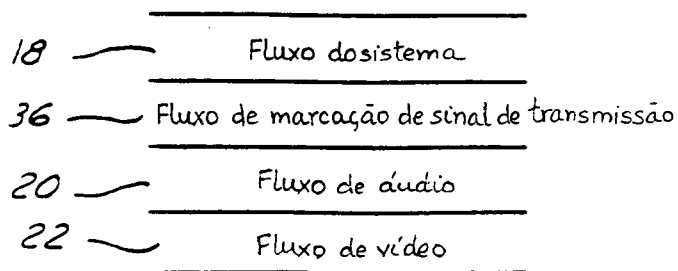


FIG. 4

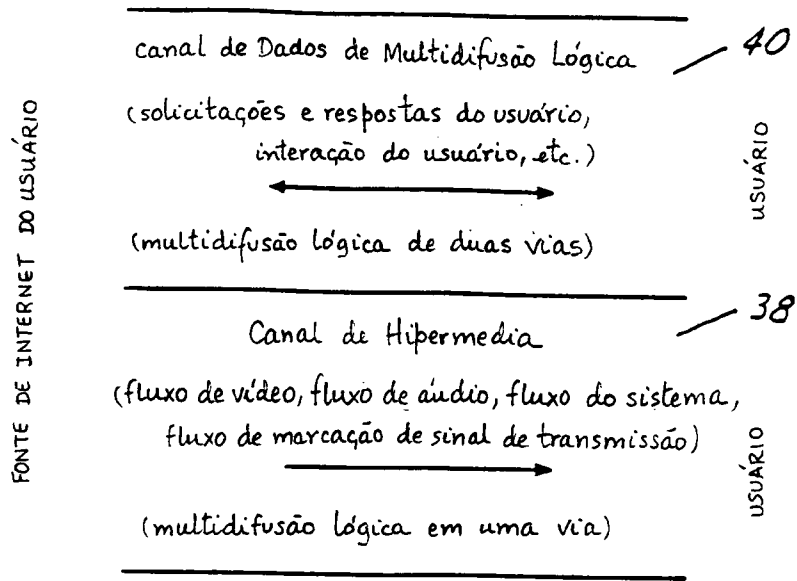


FIG. 5

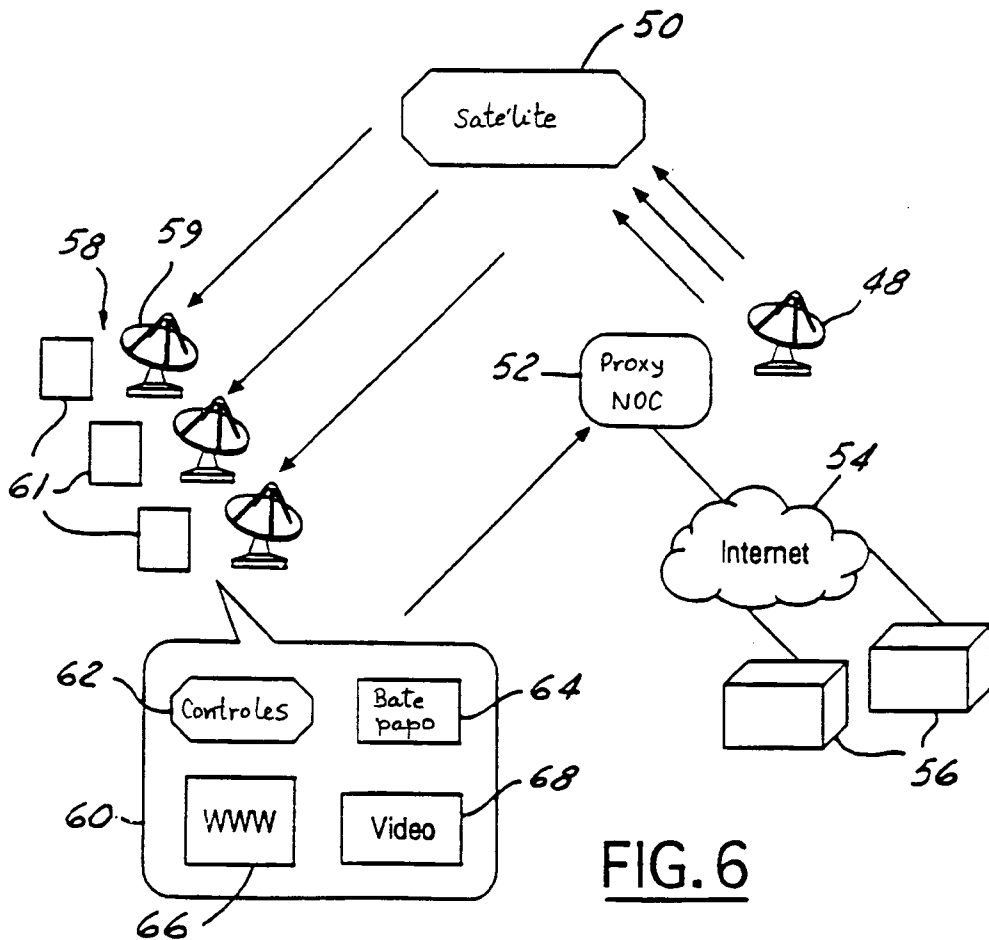


FIG. 6

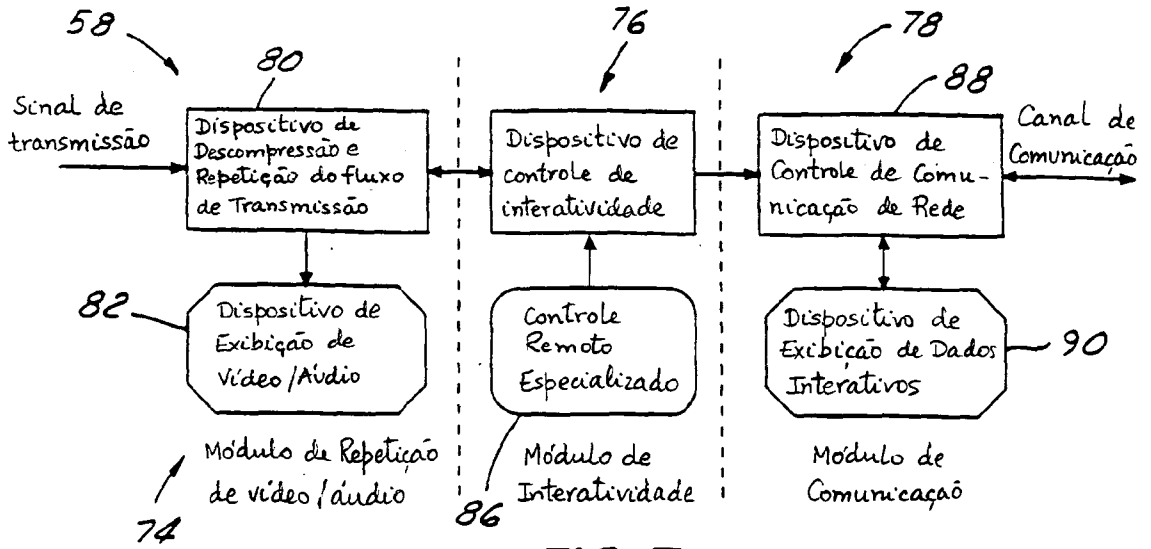


FIG. 7

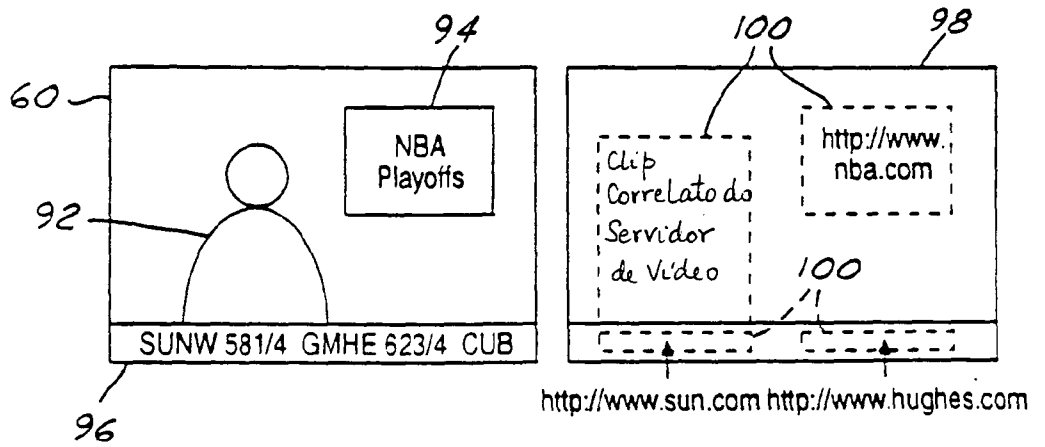


FIG. 8

SISTEMA DE TRANSMISSÃO INTERATIVA

Um sistema (10) para produzir e receber transmissão interativa tendo um sinal de transmissão com um canal de hipermídia (38) e um canal de dados de multidifusão lógica (40). O canal de dados de multidifusão lógica (40) tem um 5 fluxo de marcação de sinal de transmissão (BSMS) (36). Os usuários do sistema (28) têm um receptor (58) com um módulo de interatividade (76) e um módulo de comunicação (78). O BSMS permite que objetos selecionáveis pelo usuário sejam exibidos para os usuários do sistema. Os objetos de BSMS 10 podem ser gerados combinando reconhecimento de fala, processamento de linguagem (e opcionalmente vídeo) e buscas de palavras-chave. Selecionando objetos, várias informações referentes àqueles objetos podem ser exibidas para os 15 usuários. Uma porção do canal de multidifusão lógica pode ser empregada para fornecer interatividade entre os usuários. Uma variação elimina o BSMS empregando-se um servidor de fluxo de marcação para fornecer links de acesso ou gerando teor de link diretamente no receptor pelo mesmo processo.