

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

(21) **PI 0721501-0 A2**

(22) Data de Depósito: 29/06/2007
(43) Data da Publicação: 26/02/2013
(RPI 2199)



(51) *Int.Cl.:*
H04N 7/26

(54) **Título:** MÉTODO PARA CODIFICAR DADOS DE VÍDEO DE MANEIRA ESCALONÁVEL

(30) **Prioridade Unionista:** 18/04/2007 US 60/923.993

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

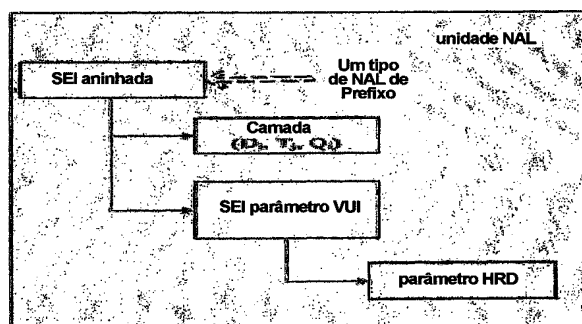
(72) **Inventor(es):** Juheng Yang, Lihua Zhu, Zhibo Chen

(74) **Procurador(es):** Nellie Anne Daniel-Shores

(86) **Pedido Internacional:** PCT CN2007002031 de 29/06/2007

(87) **Publicação Internacional:** WO 2008/128388de 30/10/2008

(57) **Resumo:** MÉTODO PARA CODIFICAR DADOS DE VIDEO DE MANEIRA ESCALONÁVEL. Trata-se de um método para codificar dados de vídeo de maneira escalonável de acordo com o padrão H.264/SVC, que compreende as etapas de: inserir uma mensagem de Informação de Aperfeiçoamento Suplementar aninhada escalonável para cada camada do fluxo de dados compreendendo pelo menos uma referência à camada e um vínculo a uma mensagem de Informação de Aperfeiçoamento Suplementar, seguindo a mensagem de Informação de Aperfeiçoamento Suplementar, inserir a mensagem de Informação de Aperfeiçoamento Suplementar para cada mensagem de Informação de Aperfeiçoamento Suplementar aninhada escalonável compreendendo a informação de usabilidade de vídeo para a camada.



"MÉTODO PARA CODIFICAR DADOS DE VÍDEO DE MANEIRA ESCALONÁVEL"

CAMPO DA INVENÇÃO

A invenção diz respeito a um método para codificar dados de vídeo de maneira escalonável.

5 ANTECEDENTES DA INVENÇÃO

A invenção se refere principalmente ao campo de codificação de vídeo quando os dados podem ser codificados de maneira escalonável.

10 A codificação dos dados de vídeo de acordo com várias camadas pode ser de grande auxílio quando os terminais aos quais os dados se destinam possuem capacidades diferentes, e, portanto, não conseguem decodificar todo o fluxo de dados, somente parte dele. Quando os dados de vídeo são codificados de acordo com várias camadas de maneira escalonável, o terminal receptor pode extrair, do fluxo de bits recebidos, os dados de acordo com seu perfil.

15 Atualmente, existem vários padrões de codificação de vídeo capazes de codificar dados de vídeo de acordo com diferentes camadas e/ou perfis. Dentre eles, podemos citar o H.264/SVC, também chamado de padrão ITU-T H.264.

No entanto, um problema existente é a sobrecarga que ele cria por transmitir mais dados do que o costuma ser necessário no lado receptor.

20 De fato, por exemplo, no H.264/SVC ou no MVC (SVC significa "codificação de vídeo escalonável" e MVC significa "codificação de vídeo com múltiplas visualizações), a transmissão de várias camadas solicita a transmissão de muitos cabeçalhos de modo a transmitir todos os parâmetros solicitados pelas diferentes camadas. Na versão atual do padrão, um cabeçalho compreende os parâmetros que correspondem a todas as camadas. Portanto, quando se precisa transmitir apenas a camada de base, todas as informações relacionadas às camadas de aperfeiçoamento têm de ser transmitidas. Por consequência, 25 gera-se uma grande sobrecarga na rede ao transmitir todos os parâmetros para todas as camadas, mesmo se todos os dados de camada não forem solicitados pelos diferentes dispositivos aos quais os dados de destinam.

A invenção propõe solucionar pelo menos uma dessas desvantagens.

30 SUMÁRIO DA INVENÇÃO

Para esse fim, a invenção propõe um método para codificar dados de vídeo de maneira escalonável de acordo com o padrão H.264/SVC. De acordo com a invenção, o método compreende as etapas de

35 - inserir uma mensagem de Informação de Aperfeiçoamento Suplementar aninhada escalonável para cada camada do fluxo de dados compreendendo pelo menos uma referência à camada e um vínculo a uma mensagem de Informação de Aperfeiçoamento Suplementar,

- seguindo a mensagem de Informação de Aperfeiçoamento Suplementar aninhada, inserir a referida mensagem de Informação de Aperfeiçoamento Suplementar para cada mensagem de Informação de Aperfeiçoamento Suplementar aninhada escalonável compreendendo a informação de usabilidade do vídeo para a referida camada.

5 De acordo com uma concretização preferida, a mensagem de Informação de Aperfeiçoamento Suplementar compreende uma referência ao Conjunto de Parâmetros de Sequência (SPS) ao qual a referida camada está vinculada.

De acordo com uma concretização preferida, a mensagem de Informação de Aperfeiçoamento Suplementar compreende a informação de usabilidade do vídeo, conforme definida no padrão H.264/SVC.

Em alguns métodos de codificação, todos os parâmetros para todas as camadas são transmitidos como um todo, não importam quantas camadas sejam transmitidas. Portanto, isso cria uma grande sobrecarga sobre a rede. Isso se deve principalmente ao fato de que alguns dos parâmetros dependem de certas camadas, enquanto outros são comuns a todas as camadas; portanto, um cabeçalho sendo definido para todos os parâmetros, todos parâmetros dependentes e independentes de camada são transmitidos juntos.

Graças à invenção, os parâmetros dependentes da camada são transmitidos apenas quando necessário, isto é, quando os dados codificados de acordo com essas camadas são transmitidos em vez de transmitir todo o cabeçalho compreendendo os parâmetros para todas as camadas.

BREVE DESCRIÇÃO DOS DESENHOS

Outras características e vantagens da invenção ficarão visíveis na descrição de uma concretização não-limitante da invenção, que será ilustrada com o auxílio dos desenhos em anexo:

25 - A Figura 1 representa a estrutura da unidade NAL usada para codificação de camadas escalonável de acordo com a técnica anterior,

- A Figura 2 representa uma concretização da estrutura conforme proposta na presente invenção,

30 - A Figura 3 representa uma visão geral do codificador de vídeo escalonável de acordo com uma concretização preferida da invenção,

- A Figura 4 representa uma visão geral do fluxo de dados de acordo com uma concretização preferida da invenção,

- A Figura 5 representa um exemplo de um fluxo de bits de acordo com uma concretização preferida da invenção.

35 DESCRIÇÃO DETALHADA DAS CONCRETIZAÇÕES PREFERIDAS

De acordo com a concretização preferida ora descrita, os dados de vídeo são codificados de acordo com o padrão H264/SVC. O SVC propõe a transmissão dos dados de

vídeo de acordo com vários níveis espaciais, temporais e de qualidade. Para um nível espacial, pode-se codificar de acordo com vários níveis temporais, e para cada nível temporal, de acordo com vários níveis de qualidade. Portanto, quando m níveis espaciais são definidos, n níveis temporais e O níveis de qualidade, os dados de vídeo podem ser codificados de acordo com $m \cdot n \cdot O$ níveis diferentes. De acordo com as capacidades do cliente, diferentes camadas são transmitidas até um certo nível correspondendo ao máximo das capacidades do cliente.

Como mostra a figura 1, que representa a técnica anterior à invenção, atualmente no SVC, a SPS é uma estrutura de sintaxe que contém elementos de sintaxe que se aplicam a zero ou mais sequências de vídeo inteiras codificadas, conforme determinado pelo conteúdo do elemento de sintaxe `seq_parameter_set_id` encontrados no conjunto de parâmetros de imagem referenciado pelo elemento de sintaxe `pic_parameter_set_id` encontrado em cada cabeçalho de fatia. No SVC, os valores de alguns elementos de sintaxe transmitidos na SPS são dependentes da camada. Esses elementos de sintaxe incluem, sem restrição, a informação de sincronização, parâmetros HRD (que significa "Decodificador de Referência Hipotético") e informações de restrição de fluxo de bits. Portanto, é necessário possibilitar a transmissão dos elementos de sintaxe supracitados para cada camada.

Um Conjunto de Parâmetros de Sequência (SPS) compreende todos os parâmetros necessários para todos os níveis correspondentes (espacial (D_i), temporal (T_i) e de qualidade (Q_i)) sempre que todas as camadas são transmitidas ou não.

A SPS compreende os parâmetros VUI (Informação de Usabilidade do Vídeo) para todas as camadas. Os parâmetros VUI representam uma quantidade de dados muito importante, uma vez que eles compreendem os parâmetros HRD para todas as camadas. Em aplicações práticas, uma vez que a taxa do canal é restringida, apenas certas camadas são transmitidas através da rede. Como a SPS representa um elemento de sintaxe básico no SVC, ela é transmitida como um todo. Portanto, não importa qual camada seja transmitida: os parâmetros HRD para todas as camadas são transmitidos.

Como mostra a figura 2, de modo a reduzir a sobrecarga do conjunto de Parâmetros de Sequência (SPS) para codificação de vídeo escalonável, a invenção propõe usar uma NAL de prefixo/sufixo `nesting_sei` e armazenar os parâmetros VUI em uma mensagem SEI.

O `scalable_nesting`, também chamado de SEI aninhada (e representado como NSEI nos desenhos), está agindo como o cabeçalho de uma unidade NAL do tipo prefixo/sufixo indicando as informações de camada. O `scalable_nesting` é vinculado, graças ao campo `vui_parameter_sei()`, à mensagem `vui_parameter_sei` compreendendo todas as propriedades da camada especificada pelo SEI aninhada.

A tabela 1 a seguir ilustra o `scalable_nesting` conforme definido pela NAL de prefixo-

xo/sufixo.

scalable_nesting (payloadSize) {	C	Descritor
all_pictures_in_au_flag	5	u(1)
if (all_pictures_in_au_flag == 0) {		
num pictures minus1	5	ue(v)
para(i = 0; i <= num_pictures_minus1; i++) {		
dependency_id[i]	5	u(3)
quality level[i]	5	u(2)
}		
temporal_level	5	u(3)
}		
while(!byte aligned())		
sei_nesting_zero_bit /* igual a 0 */	5	f(1)
vui parameter sei ()	5	
}		

Tabela 1

Uma mensagem SEI aninhada escalonável se refere a uma unidade de acesso. Quando presente, esta mensagem SEI aparece antes de qualquer unidade VCL NAL da unidade de acesso correspondente. A SEI aninhada escalonável está contida em uma unidade NAL. O escopo ao qual a mensagem SEI aninhada se aplica é indicado pelos elementos de sintaxe all_pictures_in_au_flag, e num_pictures_dependency_id[i] e quality_id[i], quando presentes.

- all_pictures_in_au_flag igual a 1 indica que a mensagem SEI aninhada se aplica a todas as imagens codificadas da unidade de acesso. all_pictures_in_au_flag igual a 0 indica que o escopo aplicável da mensagem SEI aninhada é sinalizado pelos elementos de sintaxe num_pictures, dependency_id[i] e quality[i].

- num_pictures_minus1 indica o número de imagens codificadas às quais a mensagem SEI aninhada se aplica.

- dependency_id[i] e quality_id[i] indicam, respectivamente, o dependency_id (nível espacial) e o quality_id da i-ésima imagem codificada à qual a mensagem SEI aninhada se aplica.

- sei_nesting_zero_bit é igual a 0.

A tabela a seguir ilustra a mensagem sei contendo os parâmetros específicos a cada camada.

vui_parameter_sei () {	C	Descritor
sequence_parameter_set_id	0	ue(v)

timing_info_present_flag	0	u(1)
se(timing_info_present_flag) {		
num_units_in_tick	0	u(32)
time_scale	0	u(32)
fixed_frame_rate_flag	0	u(1)
}		
nal_hrd_parameters_present_flag	0	u(1)
se(nal_hrd_parameters_present_flag)		
hrd_parameters()		
vcl_hrd_parameters_present_flag	0	u(1)
Se(vcl_hrd_parameters_present_flag)		
hrd_parameters()		
Se(nal_hrd_parameters_present_flag vcl_hrd_parameters_present_flag)		
low_delay_hrd_flag	0	u(1)
pic_struct_present_flag	0	u(1)
bitstream_restriction_flag	0	u(1)
Se(bitstream_restriction_flag) {		
motion_vectors_over_pic_boundaries_flag	0	u(1)
max_bytes_per_pic_denom	0	ue(v)
max_bits_per_mb_denom	0	ue(v)
log2_max_mv_length_horizontal	0	ue(v)
log2_max_mv_length_vertical	0	ue(v)
num_reorder_frames	0	ue(v)
max_dec_frame_buffering	0	ue(v)
}		
}		

Tabela 2

O sequence_parameter_set_id identifica o conjunto de parâmetros de sequência (SPS) para o qual mapeia a mensagem vui_parameter_sei atual e inclui as propriedades de parâmetro de sequência comuns para a camada atual.

5 Os outros parâmetros mencionados na tabela 2 são definidos no padrão H.264/SVC.

A tabela 3 a seguir ilustra a modificação a ser feita na definição existente do sei_payload conforme definido atualmente no padrão H.264/SVC. O vui_parameter_sei é definido como sendo do tipo 30. Em outras concretizações da invenção, ele pode ser qual-

quer outro campo ainda disponibilizado pelo padrão H.264/SVC.

sei_payload(payloadType, payloadSize) {	C	Descritor
se (payloadType=0)	C	Descritor
Buffering_period(payloadSize)		
.....		
se(payloadType==30)		
vui_parameter_sei(payloadSize)		
Senão		
Reserved_sei_message(payloadSize)		
.....		
}		

Tabela 3

A Figura 3 mostra uma concretização de um codificador de vídeo escalonável 1 de acordo com a invenção.

5 Um vídeo é recebido na entrada do codificador de vídeo escalonável 1.

O vídeo é codificado de acordo com diferentes níveis espaciais. Os níveis espaciais se referem principalmente a níveis diferentes de resolução do mesmo vídeo. Por exemplo, como a entrada de um codificador de vídeo escalonável, pode-se ter uma sequência GIF (352 por 288) ou uma sequência QCIF (176 por 144) que representa cada nível espacial.

10 Cada um dos níveis espaciais é enviado a um módulo de predição com compensação de movimento hierárquico. O nível espacial 1 é enviado ao módulo de predição com compensação de movimento hierárquico 2", o nível espacial 2 é enviado ao módulo de predição com compensação de movimento hierárquico 2' e o nível espacial n é enviado ao módulo de predição com compensação de movimento hierárquico 2.

15 Os níveis espaciais sendo codificados em 3 bits, usando o dependency_id, portanto, o número máximo de níveis espaciais é 8.

Uma vez realizada a compensação de movimento predita hierárquica, dois tipos de dados são gerados, sendo um deles o movimento, que descreve a disparidade entre as diferentes camadas, e o outro sendo a textura, que é o erro de estimação.

20 Para cada um dos níveis espaciais, os dados são codificados de acordo com uma camada de base e uma camada de aperfeiçoamento. Para o nível espacial 1, os dados são codificados através do codificador de camada de aperfeiçoamento 3" e do codificador de camada de base 4"; para o nível espacial 2, os dados são codificados através do codificador de camada de aperfeiçoamento 3' e do codificador de camada de base; para o nível espacial
25 1, os dados são codificados através do codificador de camada de aperfeiçoamento 3 e do codificador de camada de base 4.

Após a codificação, os cabeçalhos são preparados, e para cada uma das camadas

espaciais, são criadas mensagens SPS e PPS e várias mensagens NSEI-VUI_SEI.¹

Para o nível espacial 1, como representado na figura 3, a SPS e a PPS 5" são criadas, e um conjunto de NSEI-VUI_SEI¹₁, NSEI-VUI_SEI¹₂, ..., NSEI-VUI_SEI¹_m0, também é criado de acordo com esta concretização da invenção.

5 Para o nível espacial 2, como representado na figura 3, a SPS e a PPS 5" são criadas, e um conjunto de NSEI-VUI_SEI¹₁, NSEI-VUI_SEI²₂, ..., NSEI-VUI_SEI²_m0, também é criado de acordo com esta concretização da invenção.

10 Para o nível espacial n, como representado na figura 3, a SPS e a PPS 5 são criadas, e um conjunto de NSEI-VUI_SEIⁿ₁, NSEI-VUI_SEIⁿ₂, ..., NSEI-VUI_SEIⁿ_m0, também é criado de acordo com esta concretização da invenção.

Os fluxos de bits codificados pelos módulos de codificação da camada de base e pelos módulos de codificação da camada de aperfeiçoamento estão seguindo a pluralidade de cabeçalhos SPS, PPS e SUP_SPS no fluxo de bits global.

15 Na figura 3, 8" compreende SPS e PPS 5", NSEI-VUI_SEI¹₁, NSEI-VUI_SEI¹₂, ..., NSEI-VUI_SEI¹_m6" e o fluxo de bits 7", que constituem todos os dados codificados associados ao nível espacial 1.

Na figura 3, 8" compreende SPS e PPS 5', NSEI-VUI_SEI¹₁, NSEI-VUI_SEI²₂, ..., NSEI-VUI_SEI²_m6' e o fluxo de bits 7', que constituem todos os dados codificados associados ao nível espacial 2.

20 Na figura 3, 8" compreende SPS e PPS 5', NSEI-VUI_SEIⁿ₁, NSEI-VUI_SEIⁿ₂, ..., NSEI-VUI_SEIⁿ_m6 e o fluxo de bits 7, que constituem todos os dados codificados associados ao nível espacial n.

Os diferentes cabeçalhos NSEI-VUI_SEI estão em conformidade com os cabeçalhos descritos nas tabelas acima.

25 A Figura 4 representa um fluxo de bits, conforme codificado pelo codificador de vídeo escalonável da figura 3.

O fluxo de bits compreende um SPS para cada um dos níveis espaciais. quando m níveis espaciais são codificados, o fluxo de bits compreende SPS1, SPS2 e SPSm representados por 10, 10' e 10" na figura 4.

30 No fluxo de bits, cada SPS codificando a informação geral relativa ao nível espacial é seguida de um cabeçalho 10 do tipo NSEI-VUI_SEI, que por sua vez é seguido pelos dados de vídeo codificados correspondentes correspondendo, cada um correspondendo a um nível temporal e um nível de qualidade.

35 Portanto, quando um nível correspondendo a um nível de qualidade não é transmitido, o cabeçalho correspondente também não é transmitido, uma vez que há um cabeçalho NSEI-VUI_SEI correspondendo a cada nível.

Portanto, tomemos como exemplo a figura 5 para ilustrar o fluxo de dados a ser

transmitido.

A Figura 5 ilustra a transmissão dos seguintes níveis. As referências indicadas no fluxo de bits correspondem às referências usadas na figura 2.

As seguintes camadas são transmitidas:

- 5
 - camada espacial 1
 - nível temporal 1
 - o Nível de qualidade 1
 - nível temporal 2
 - o Nível de qualidade 1
- 10
 - camada espacial 2
 - nível temporal 1
 - o Nível de qualidade 1
 - camada espacial 3
 - nível temporal 1
 - o Nível de qualidade 1
 - nível temporal 2
 - o Nível de qualidade 1
 - nível temporal 3
 - o Nível de qualidade 1

20 Portanto, é possível notar que nem todos os parâmetros diferentes para todas as camadas são transmitidos, mas sim apenas os que correspondem às camadas solicitadas, uma vez que eles estão compreendidos nas mensagens NSEI-VUI_SEI e não mais nas mensagens SPS.

REIVINDICAÇÕES

1. Método para codificar dados de vídeo de maneira escalonável de acordo com o padrão H.264/SVC, **CARACTERIZADO** por compreender as etapas de:

- inserir uma mensagem de Informação de Aperfeiçoamento Suplementar aninhada escalonável para cada camada do fluxo de dados compreendendo pelo menos uma referência à camada e um vínculo a uma mensagem de Informação de Aperfeiçoamento Suplementar,

- seguindo a mensagem de Informação de Aperfeiçoamento Suplementar aninhada, inserir a referida mensagem de Informação de Aperfeiçoamento Suplementar para cada mensagem de Informação de Aperfeiçoamento Suplementar de aninhamento escalonável compreendendo a informação de usabilidade de vídeo para a referida camada.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a referida mensagem de Informação de Aperfeiçoamento Suplementar compreende uma referência ao Conjunto de Parâmetros de Sequência (SPS) ao qual a referida camada está vinculada.

3. Método, de acordo com a reivindicação 2, **CARACTERIZADO** pelo fato de que a mensagem de Informação de Aperfeiçoamento Suplementar compreende a informação de usabilidade do vídeo, conforme definida no padrão H.264/SVC.

Técnica Anterior

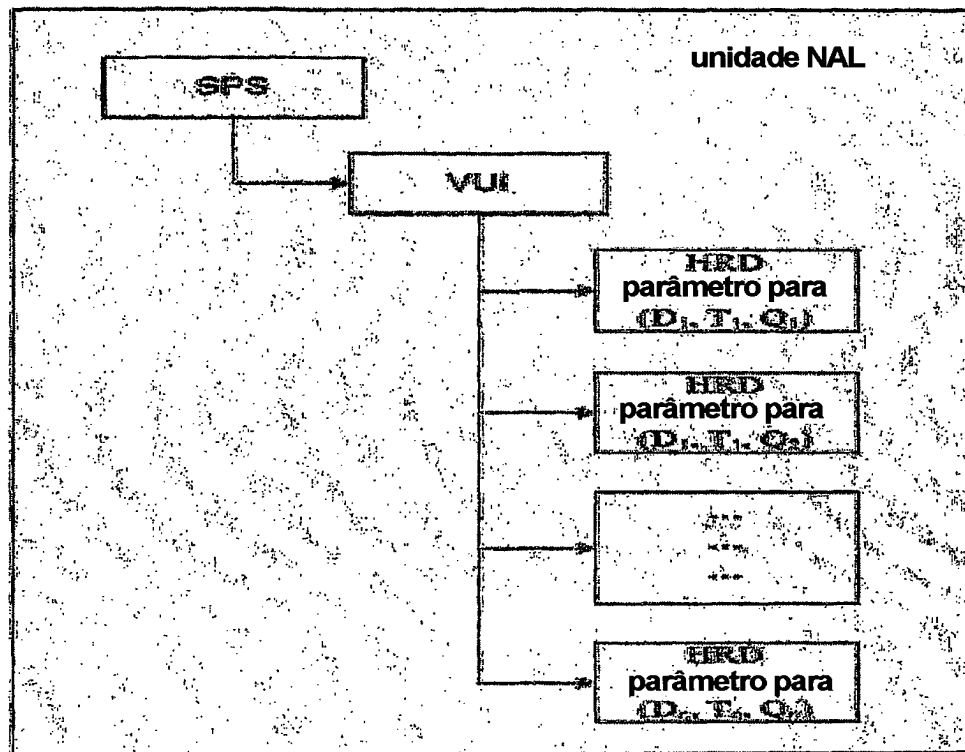


Fig 1

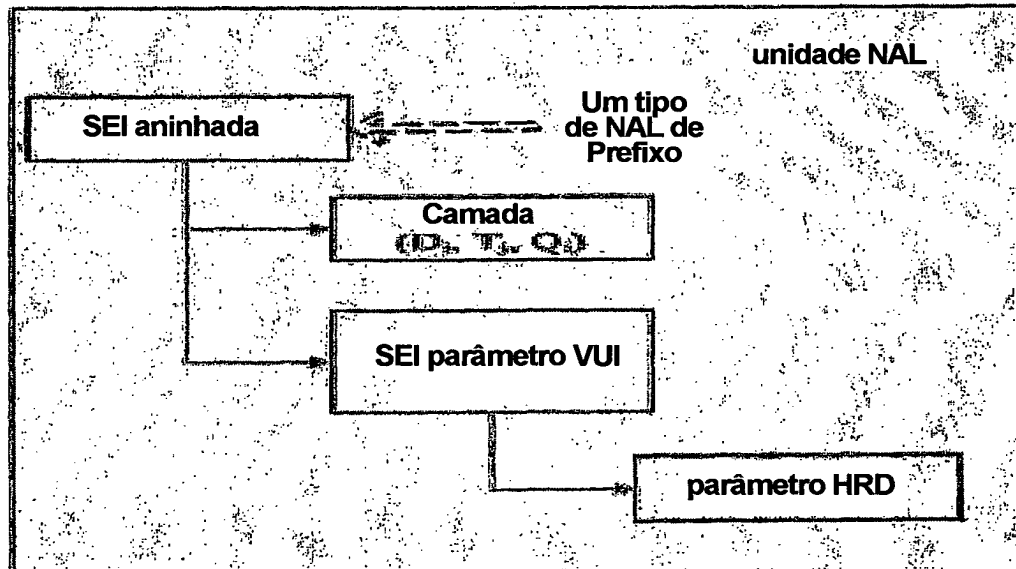


Fig 2

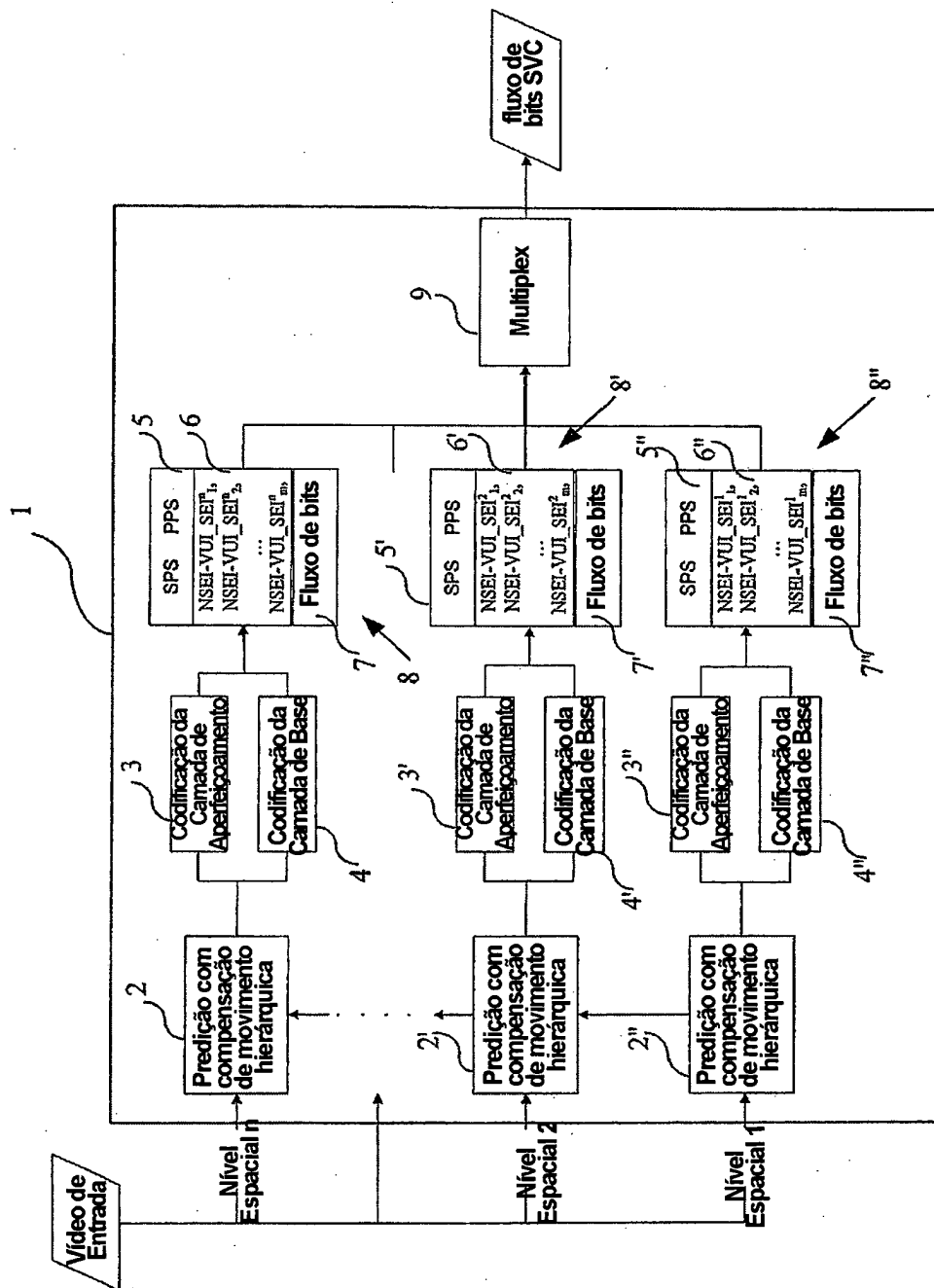


Fig 3

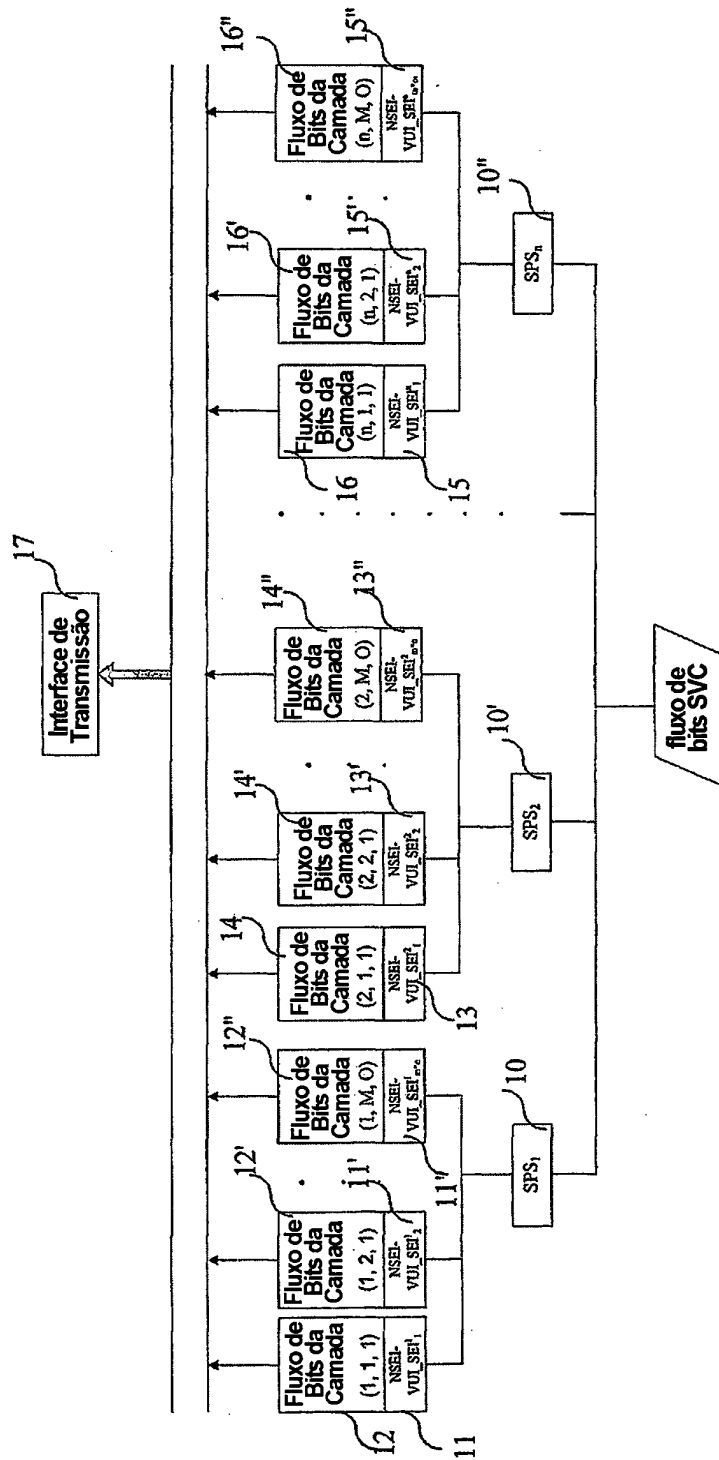


Fig 4

10	11	12	11'	12'	10'	13	14	10"	15	16	15'	16'	15"	16"
----	----	----	-----	-----	-----	----	----	-----	----	----	-----	-----	-----	-----

Vídeo Escalonável

**Fluxo de dados
codificado**

Fig 5

RESUMO

"MÉTODO PARA CODIFICAR DADOS DE VÍDEO DE MANEIRA ESCALONÁVEL"

Trata-se de um método para codificar dados de vídeo de maneira escalonável de acordo com o padrão H.264/SVC, que compreende as etapas de: inserir uma mensagem de

5 Informação de Aperfeiçoamento Suplementar aninhada escalonável para cada camada do fluxo de dados compreendendo pelo menos uma referência à camada e um vínculo a uma mensagem de Informação de Aperfeiçoamento Suplementar, seguindo a mensagem de In-

10 formação de Aperfeiçoamento Suplementar, inserir a mensagem de Informação de Aperfeiçoamento Suplementar para cada mensagem de Informação de Aperfeiçoamento Suple-

mentar aninhada escalonável compreendendo a informação de usabilidade de vídeo para a camada.