

PCT

Indus

SUPER

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

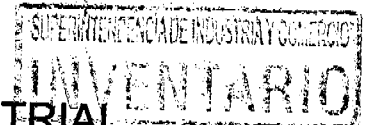


No. 11-001544-00000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR

Tra. 11 PATENTEIYII Lve: 379 FASENACIONALII

Act. 441 PRESENTACION Folios: 182



DELEGATURA PROPIEDAD INDUSTRIAL

SOLICITUD
REGISTRO DE PATENTE DE INVENCION
PCT

EXPEDIENTE No

11-1544

TITULO MÉTODO Y DISCRIMINADOR PARA
CLASIFICAR DISTINTOS SEGMENTOS DE UNA
SEÑAL

CLASIFICACION INTERNACIONAL _____

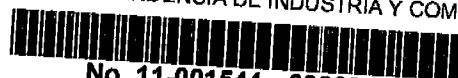
SOLICITANTE Fraunhofer-Gesellschaft zur
Förderung der angewandten Forschung e.V.

DOMICILIO Munich, Alemania

APODERADO JUANITA ACOSTA GOMEZ

BOGOTA D.C. _____

11-02-2011



①

SOLICITUD DE:

☒ Patente de Invención

☐ Patente de Modelo de Utilidad

☐ Capítulo I

☒ Capítulo II

②

SOLICITANTE

Nombre: **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**

Dirección: **Hansastraße 27c 80686 Munich, Alemania**

Nacionalidad o Domicilio: **Munich, Alemania**

Lugar de Constitución: **Alemania**

Teléfono: **3137800** Fax: **3122410**

E-mail: **general@cardenasycardenas.com**

IDENTIFICACION

☐ NIT ☐ NIT

☐ Otro ☒ Otro

Cuál
Número

③

REPRESENTANTE
O APODERADO

Nombre: **JUANITA ACOSTA GOMEZ**

Dirección: **Carrera 7 No 71-52 Torre B Piso 9**

Teléfono: **312 3600** Fax: **312 2410**

E-Mail: **general@cardenasvcardenas.com**

IDENTIFICACION

☐ NIT ☐ NIT

☒ CC ☐ Otro

Cual 52618636 BTA
Número 79686 C.S.J.

④

INVENTOR(ES)
(72)

Nombre: **VER ANEXO**

Dirección: **VER ANEXO**

Nacionalidad o Domicilio: **VER ANEXO**

⑤ PCT (86)

Solicitud Internacional No. **WO/2010/003521** Fecha **14.01.2010**

Publicación Internacional No. **PCT/EP2009/004339** Fecha **16.06.2009**

⑥

Título(54):

MÉTODO Y DISCRIMINADOR PARA CLASIFICAR DISTINTOS SEGMENTOS DE UNA SEÑAL

⑦

Clasificación Internacional (51):

⑧

Prioridad

(33)País de Origen

(32)Fecha

(31)N° de Solicitud

US

11.07.2008

61/079,875

SI ☒ NO ☐

⑨

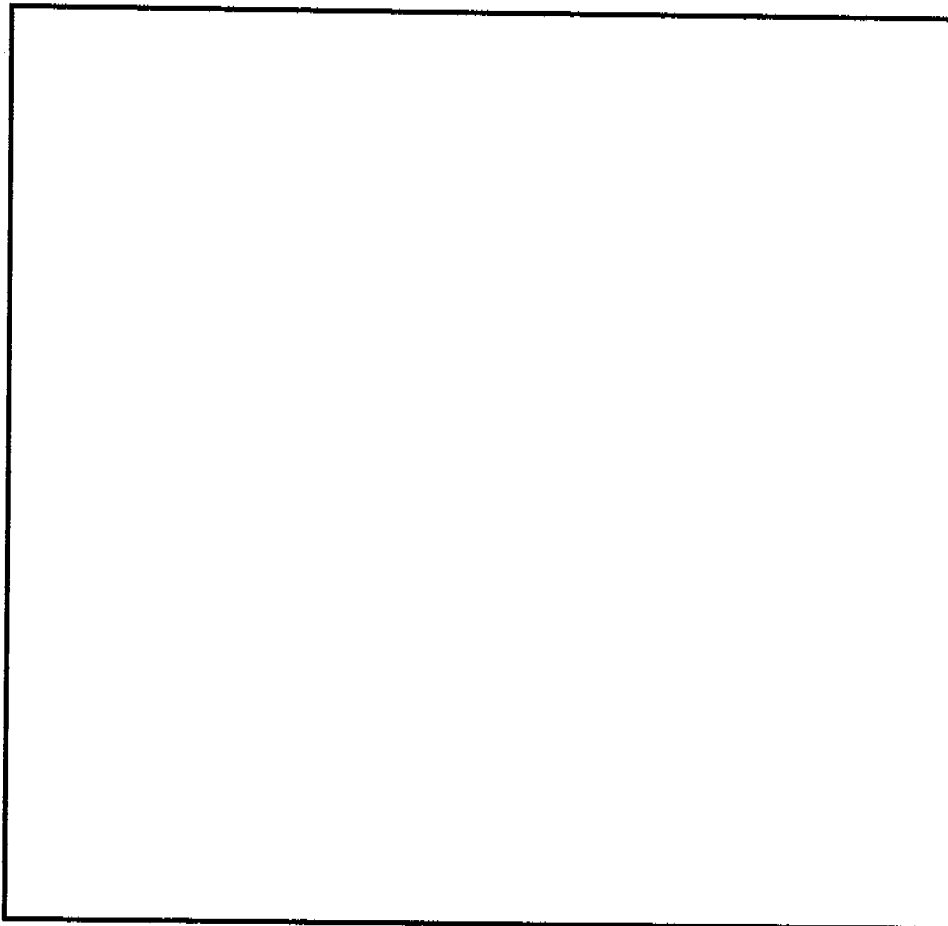
Comprobante de Pago No.: **11-667**

Fecha **Enero 4, 2011**

④ INVENTOR(ES) (72)	Nombre: 1. Guillaume Fuchs ✓ 2. Stefan Bayer ✓ 3. Jens Hirschfeld ✓ 4. Juergen Herre ✓ 5. Jérémie Lecomte ✓ 6. Frederick Nagel ✓ 7. Nikolaus Rettelbach ✓ 8. Stefan Wabnik ✓ 9. Yoshikazu Yokotani ✓
	Dirección: 1. Parkstrasse 12 90409 Nuernberg, Alemania 2. Dortmunder Strasse 14 90425 Nuernberg, Alemania 3. Steinweg 32 36266 Heringen, Alemania 4. Hallerstrasse 24 91054 Buckenhof, Alemania 5. Sulzbacher Strasse 39 90489 Nuernberg, Alemania 6. Wilhelmshavener Strasse 72 90425 Nuernberg, Alemania 7. Spessartstrasse 38 90427 Nuernberg, Alemania 8. Fichtenweg 5 98693 Ilmenau, Alemania 9. Kibune 336-101, Hamamatsu Shizuoka 434-0038, Japón
	Nacionalidad o Domicilio: 1. Frances 2. Alemán 3. Alemán 4. Alemán 5. Frances 6. Alemán 7. Alemán 8. Alemán 9. Japonés

10 ANEXOS

- ☒ Comprobante de pago de la tasa de presentación de la solicitud.
- ☒ Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- ☒ Poderes, si fuere el caso.
- ☐ Documento de cesión del inventor al solicitante o a su causante.
- ☒ Traducción de la solicitud internacional al castellano (Resumen, Resumen, descripción, reivindicaciones, dibujos)
- ☒ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA)
- ☐ Traducción del examen preliminar internacional efectuado por la IPEA.
- ☐ De ser el caso, copia del contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- ☐ De ser el caso, certificado de depósito del material biológico.
- ☒ Arte final 12x12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☒ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☒ Prioridad <http://www.wipo.int/pctdb/en/wo.jsp?WO=2010003521&IA=EP2009004339&DISPLAY=DOCS>
- ☒ Recibo de Reivindicaciones adicionales

11 FIGURA CARACTERISTICA**12**

Solicito la Concesión de la patente

NOMBRE: JUANITA ACOSTA GOMEZ

FIRMA:

C.C. 52618636
UsaquénTP. 79686
C.S.J.

RECIBO OFICIAL DE PAGO - 11 - 567
FECHA: 11/01/2011

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544- -00000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASC
Tra. 11 PATENTEIYI Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

FORMA DE PRESENTACION

DESCRIPCION	TIPO PAGO	CANTIDAD	MONEDA	VALOR PAGO	FECHA PAGO	VALOR TOTAL
CARRETERAS Y CARRETERAS S.A. INVERSIÓN	ALICUOTA DE IMPORTE	10000000	USD	10000000	11/01/2011	10000000

FORMA DE PRESENTACION

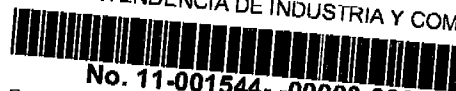
DESCRIPCION	TIPO PAGO	CANTIDAD	MONEDA	VALOR PAGO	FECHA PAGO	VALOR TOTAL
CARRETERAS Y CARRETERAS S.A. INVERSIÓN	ALICUOTA DE IMPORTE	10000000	USD	10000000	11/01/2011	10000000

GARDENAS & GARDENAS

RECIBO DE PAGO APLICADO AL EXPEDIENTE No. **ABRIL 2011**

RECIBO DE PAGO DE CASA N.º 11-001544-000000-0000
 FECHA: 2011-01-07 11:18:49

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544-000000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASC
 Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
 Act. 411 PRESENTACION Folios: 182

IMPORTE DE LA OPERACIÓN

FECHA	TIPO PAGO	NOMBRE	CÓDIGO	VALOR PAGO	VALOR RECIBO	VALOR TOTAL
-------	-----------	--------	--------	------------	--------------	-------------

2011-01-07	PAGO DE CASA	REPÚBLICA DE COLOMBIA	000000	11.000.000,00	11.000.000,00	11.000.000,00
------------	--------------	-----------------------	--------	---------------	---------------	---------------

IMPORTE DE LA OPERACIÓN

FECHA	TIPO PAGO	NOMBRE	CÓDIGO	VALOR PAGO	VALOR RECIBO	VALOR TOTAL
-------	-----------	--------	--------	------------	--------------	-------------

1. PAGO DE CASA

IMPORTE DE LA OPERACIÓN

11.000.000,00

11.000.000,00

IMPORTE DE LA OPERACIÓN

RECIBO DE PAGO

[Handwritten signature]

RECIBO DE PAGO DE CASA N.º 11-001544-000000-0000

IMPORTE DE LA OPERACIÓN

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544- -00000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

02042 7 6 8 5 1 2 4 7 1 1 2 3 3 3 3 3

IMPEDIENTE	TIPO ABUS	CANAL	CURSO	GR. PAIS	SECTOR ABUS	GR. PAIS
REPERTEJO DE CARRERAS P. I. DIMENSIONALES		REPORTE DE PROTECCION	REPERTEJO	REPERTEJO	REPERTEJO	REPERTEJO

NAME, RESIDENCE	PROPERTY	TOTAL PROPERTY
1. BROWN-01-01 WILMINGTON	2000 BROWNWOOD DRIVE WILMINGTON	20,000.00
	TOTAL	20,000.00

3042 2025/07/07 AM 0544

CARDENAS & CARDENAS

ABRICADOS

REPTES DE DATA APLICACAO AL EXPERIENTE 3-

[illegible][illegible]

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544- -00000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

[illegible][illegible]

RECEIVED BY POSTAL OFFICE OF THE UNITED STATES DEPARTMENT OF JUSTICE
JAN 10 1968

4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 10

DATE	DESCRIPTION	TOTAL COMMENTS
2010-01-01	Initial Review	1
2010-02-01	Second Review	2
2010-03-01	Third Review	3
2010-04-01	Fourth Review	4
2010-05-01	Fifth Review	5
2010-06-01	Sixth Review	6
2010-07-01	Seventh Review	7
2010-08-01	Eighth Review	8
2010-09-01	Ninth Review	9
2010-10-01	Tenth Review	10
2010-11-01	Eleventh Review	11
2010-12-01	Twelfth Review	12
2011-01-01	Thirteenth Review	13
2011-02-01	Fourteenth Review	14
2011-03-01	Fifteenth Review	15
2011-04-01	Sixteenth Review	16
2011-05-01	Seventeenth Review	17
2011-06-01	Eighteenth Review	18
2011-07-01	Nineteenth Review	19
2011-08-01	Twentieth Review	20
2011-09-01	Twenty-first Review	21
2011-10-01	Twenty-second Review	22
2011-11-01	Twenty-third Review	23
2011-12-01	Twenty-fourth Review	24
2012-01-01	Twenty-fifth Review	25
2012-02-01	Twenty-sixth Review	26
2012-03-01	Twenty-seventh Review	27
2012-04-01	Twenty-eighth Review	28
2012-05-01	Twenty-ninth Review	29
2012-06-01	Thirtieth Review	30
2012-07-01	Thirty-first Review	31
2012-08-01	Thirty-second Review	32
2012-09-01	Thirty-third Review	33
2012-10-01	Thirty-fourth Review	34
2012-11-01	Thirty-fifth Review	35
2012-12-01	Thirty-sixth Review	36
2013-01-01	Thirty-seventh Review	37
2013-02-01	Thirty-eighth Review	38
2013-03-01	Thirty-ninth Review	39
2013-04-01	Fortieth Review	40
2013-05-01	Forty-first Review	41
2013-06-01	Forty-second Review	42
2013-07-01	Forty-third Review	43
2013-08-01	Forty-fourth Review	44
2013-09-01	Forty-fifth Review	45
2013-10-01	Forty-sixth Review	46
2013-11-01	Forty-seventh Review	47
2013-12-01	Forty-eighth Review	48
2014-01-01	Forty-ninth Review	49
2014-02-01	Fiftieth Review	50
2014-03-01	Fifty-first Review	51
2014-04-01	Fifty-second Review	52
2014-05-01	Fifty-third Review	53
2014-06-01	Fifty-fourth Review	54
2014-07-01	Fifty-fifth Review	55
2014-08-01	Fifty-sixth Review	56
2014-09-01	Fifty-seventh Review	57
2014-10-01	Fifty-eighth Review	58
2014-11-01	Fifty-ninth Review	59
2014-12-01	Sixtieth Review	60
2015-01-01	Sixty-first Review	61
2015-02-01	Sixty-second Review	62
2015-03-01	Sixty-third Review	63
2015-04-01	Sixty-fourth Review	64
2015-05-01	Sixty-fifth Review	65
2015-06-01	Sixty-sixth Review	66
2015-07-01	Sixty-seventh Review	67
2015-08-01	Sixty-eighth Review	68
2015-09-01	Sixty-ninth Review	69
2015-10-01	Seventieth Review	70
2015-11-01	Seventy-first Review	71
2015-12-01	Seventy-second Review	72
2016-01-01	Seventy-third Review	73
2016-02-01	Seventy-fourth Review	74
2016-03-01	Seventy-fifth Review	75
2016-04-01	Seventy-sixth Review	76
2016-05-01	Seventy-seventh Review	77
2016-06-01	Seventy-eighth Review	78
2016-07-01	Seventy-ninth Review	79
2016-08-01	Eightieth Review	80
2016-09-01	Eighty-first Review	81
2016-10-01	Eighty-second Review	82
2016-11-01	Eighty-third Review	83
2016-12-01	Eighty-fourth Review	84
2017-01-01	Eighty-fifth Review	85
2017-02-01	Eighty-sixth Review	86
2017-03-01	Eighty-seventh Review	87
2017-04-01	Eighty-eighth Review	88
2017-05-01	Eighty-ninth Review	89
2017-06-01	Ninety Review	90
2017-07-01	Ninety-first Review	91
2017-08-01	Ninety-second Review	92
2017-09-01	Ninety-third Review	93
2017-10-01	Ninety-fourth Review	94
2017-11-01	Ninety-fifth Review	95
2017-12-01	Ninety-sixth Review	96
2018-01-01	Ninety-seventh Review	97
2018-02-01	Ninety-eighth Review	98
2018-03-01	Ninety-ninth Review	99

65-15349-10000-1407953

[illegible]

37. *Journal of the American Medical Association*, 1991; 265: 1033-1036.

REF: 10101517 001 0000

9220-180500

REF ID: A6730000

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176088-2

RECIBO OFICIAL DE CAJA : 11 - 677
FECHA : ENERO 4 DE 2011

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544- -00000-0000

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	VT. PAGO
CARDENAS Y CARDENAS P.T.	CONSIGNACION	BANCO DE BOGOTA	062754397	146326025	04/01/2011	\$,109,400.00

***** CONCEPTO *****

CANT.	CONCEPTO	TOTAL CONCEPTO
1	50009-01-01 SOLICITUDES	
	*** REIVINDICACION UNITARIA ADICIONAL	27,000.00
	TOTAL :	\$ 27,000.00

SON: VEINTISIETE MIL PESOS

RESPONSABLE :

CARDENAS & CARDENAS

ABOGADOS

RECIBO DE CAJA APLICADO AL EXPEDIENTE No.

[illegible]

Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIIII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

TOP SECRET

$$V_{\text{eff}} = V_0 + V_1 + V_2 + \dots$$

Page 12 of 12

$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

DATE OF BIRTH: 20/01/1950

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544- -00000-0000

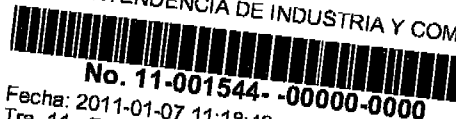
Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYI Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

CARDENAS & CARDENAS

ABOGADOS

REGION INDUSTRIAL DE PATI 11 - 669
CASA : PASO 1 DE PATI

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-001544-00000-0000

Fecha: 2011-01-07 11:18:49
Tra. 11 PATENTEIYII
Act. 411 PRESENTACION

Dep. 2020 DIR.NUEVASCR
Eve: 379 FASENACIONALII
Folios: 182

FORMA DE PRESENTACION

INDICADOR	TIPO DE	FECHA	CONCEPTO	VALOR	CLASE	VALOR
CARRENAS & CARDENAS S.R.L. PAGO DE PATI	FORMA DE PRESENTACION	2011-01-07	PRESENTE	11.000.000	11.000.000	11.000.000

FORMA DE PRESENTACION

INDICADOR	TIPO DE	FECHA	CONCEPTO	VALOR	CLASE	VALOR
CARRENAS & CARDENAS S.R.L. PAGO DE PATI	FORMA DE PRESENTACION	2011-01-07	PRESENTE	11.000.000	11.000.000	11.000.000

1. SOLICITUD DE SOLICITUD

2. PRESENTACION DE PATENTE

3. PAGO DE PATI

9000 VEINTISIETE MIL QUINIENTOS

SECRETARIA DE INDUSTRIA Y COMERCIO

(Handwritten signature)

CARRENAS & CARDENAS

REGION DE LA AMERICA LATINA Y EL CARIBE

SECRETARIA DE INDUSTRIA Y COMERCIO



Industria y Comercio

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DIVISION DE NUEVAS CREACIONES

TARJETA ARCHIVO PROPIETARIOS

PATENTE DE INVENCION ☐ PCT ☒ MODELO DE UTILIDAD ☐ DISEÑO INDUSTRIAL ☐

SUPERINTENDENCIA

(21) N° de solicitud:

(22) Fecha de solicitud:

(51) Clasificación Internacional:

(71) Solicitante(s) **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**

(72) Inventor(es) **Guillaume Fuchs, Stefan Bayer, Jens Hirschfeld...**

(74) Apoderado: **JUANITA ACOSTA GOMEZ**

(30) Prioridad	(31) No. Prioridad	(32) Fecha	(33) País
	61/079,875	11.07.2008	US

(85) Fecha inicio fase nacional: **11.01.2011**

(86) Datos relativos a la presentación de la solicitud PCT

Fecha de presentación de la solicitud: **16.06.2009**

No. de solicitud **PCT/EP2009/004339**

(54) Título **MÉTODO Y DISCRIMINADOR PARA CLASIFICAR DISTINTOS SEGMENTOS DE UNA SEÑAL**



Industria y Comercio

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DIVISION DE NUEVAS CREACIONES

TARJETA ARCHIVO TEMATICO

PATENTE DE INVENCION _____ **PCT** X **MODELO DE UTILIDAD** _____ **DISEÑO INDUSTRIAL** _____

SUPERINTENDENCIA

(21) N° de solicitud: _____ (22) Fecha de solicitud: _____
(51) Clasificación Internacional: _____
(71) Solicitante(s) **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**

(72) Inventor(es) **Guillaume Fuchs, Stefan Bayer, Jens Hirschfeld...**

(74) Apoderado: **JUANITA ACOSTA GOMEZ**

(30) Prioridad _____ (31) **No. Prioridad** **61/079,875** (32) **Fecha** **11.07.2008** (33) **País** **US**

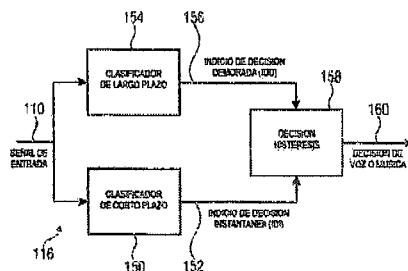
(85) Fecha límite inicio fase nacional: **11.01.2011**
(86) Datos relativos a la presentación de la solíc. PCT
Fecha de presentación de la solicitud: **16.06.2009**
No. de solicitud **PCT/EP2009/004339**

(87) Publicación internacional:
Fecha: **14.01.2010**
No. Publicación **WO/2010/003521**

(54) Título: **MÉTODO Y DISCRIMINADOR PARA CLASIFICAR DISTINTOS SEGMENTOS DE UNA SEÑAL**

(57) Resumen:

Para clasificar distintos segmentos de una señal que comprende segmentos de por lo menos un primer tipo y un segundo tipo, por ejemplo, segmentos de voz y de música, se clasifica la señal en un corto plazo (150) sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal y se entrega un resultado de clasificación de corto plazo (152). La señal se clasifica también en un largo plazo (154) sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal y se entrega un resultado de clasificación de largo plazo (156). Se combinan (158) el resultado de la clasificación de corto plazo (152) y el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si un segmento de la señal es del primer tipo o del segundo tipo.



(21) N° de solicitud: (22) Fecha de solicitud:				(51) Int Cl: (71) Solicitante(s) Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	
(30) Prioridad	(31) No. Prioridad 61/079,875	(32) Fecha 11.07.2008	(33) País US	(72) Inventor(es) Guillaume Fuchs, Stefan Bayer, Jens Hirschfeld...	
				(74) Apoderado: JUANITA ACOSTA GOMEZ	
(85) Fecha limite inicio fase nacional: 11.01.2011				(87) Publicación internacional	
(86) Datos relativos a la presentación de la solicitud PCT				Fecha: 14.01.2010	
Fecha de presentación de la solicitud: 16.06.2009				N° publicación: WO/2010/003521	
No. de solicitud PCT/EP2009/004339					
(54) Título: MÉTODO Y DISCRIMINADOR PARA CLASIFICAR DISTINTOS SEGMENTOS DE UNA SEÑAL					

Resumen:

Para clasificar distintos segmentos de una señal que comprende segmentos de por lo menos un primer tipo y un segundo tipo, por ejemplo, segmentos de voz y de música, se clasifica la señal en un corto plazo (150) sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal y se entrega un resultado de clasificación de corto plazo (152). La señal se clasifica también en un largo plazo (154) sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal y se entrega un resultado de clasificación de largo plazo (156). Se combinan (158) el resultado de la clasificación de corto plazo (152) y el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si un segmento de la señal es del primer tipo o del segundo tipo.

000854

- Colombia -

POWER OF ATTORNEY

I, Werner Liebler, confirm:

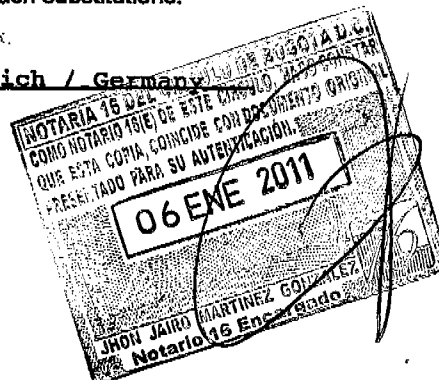
FIRST: That I act in my capacity of deputy chief of the Patent Department of **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**, legally constituted and in existence in accordance with the laws of Germany, and domiciled at Hansastrasse 27c, 80686 Muenchen, Germany.

SECOND: That in such capacity I grant hereby special, ample and sufficient Power of Attorney to DARIO CARDENAS-NAVAS and/or LUZ HELENA ADARVE-GOMEZ and/or EDUARDO CARDENAS-CABALLERO and/or BERNARDO P. CARDENAS-MARTINEZ and/or JUANITA ACOSTA GOMEZ and/or JULY GALINDO QUINTERO, registered lawyers of the firm CARDENAS & CARDENAS-LAWYERS, domiciled at Carrera 7 No. 71-52, Torre B, Piso 9 Bogotá, D.C., Colombia, to carry before the corresponding Colombian national Offices and authorities the proceedings to obtain trademarks, patents, utility models, industrial drawings and designs, deposits of commercial names and ensigns, as well as processing of renewals, assignments, change of name or domicile of the owner, waivers and licenses of the same for which purpose they are empowered to take before said authorities all necessary actions for the indicated purposes; to file applications, statements, claims, formulate descriptions, amendments, oppositions and appeals; pay all taxes and installments and make any other payments determined by law; to receive every document and values issuing the respective discharge; to fulfill any other requirements and take, in general, every measure that they shall deem opportune for the defense and protection of the Industrial Property of the company in the event of infringements, oppositions, proceedings for cancellation or annulment, imposition of fines, passing the antecedents to the courts; they are empowered to appear as plaintiff or defendant before the competent judges, being entitled to transact, submit to arbitration, waive, receive, resign, withdraw, appeal and all other legal powers which may be necessary, and it is hereby declared ever since now valid and good whatever such attorneys do for its benefit, granting them also power to substitute this proxy and if necessary to revoke such substitutions.

SIGNED AT Dezember 6th, 2010 ON Munich / Germany

BY Werner Liebler

Werner Liebler



File No. **H 2294** / 2010

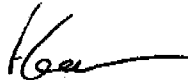
ag

I hereby certify that above is the true signature, subscribed in my presence of

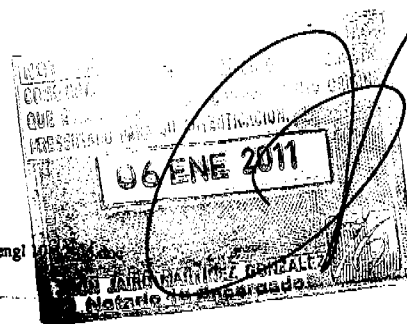
Mr. Dr. Werner Liebler,
born on May 6th, 1961,
business address: Hansastr. 27 c, 80686 Munich,
personally known to me.

Dr. Werner Liebler acting by virtue of
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
with its registered office in Munich,
business address: 80686 München, Hansastr. 27 c,
of power of attorney dated on July 13th, 2000, File No. 3076 K /2000, and
July 18th, 2000, File No. 3170 K /2000, by Notary Public, Dr. Helmut Keidel,
Munich, presented to me today in the original and which is attached as certi-
fied copy.

Munich, Dezember 6th, 2010



Thomas Haasen,
Notary Public, Munich



APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Land: Bundesrepublik Deutschland
- Diese öffentliche Urkunde
2. ist unterschrieben von Thomas Haasen
3. in seiner Eigenschaft als Notar
4. Sie ist versehen mit dem Siegel des Notars Thomas Haasen in München

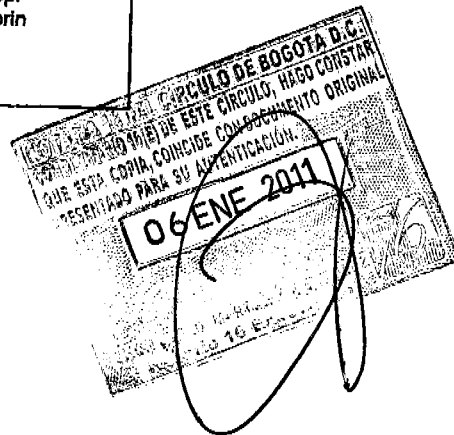
Bestätigt

5. in München
6. am 7. Dezember 2010
7. durch den Präsidenten des Landgerichts München I
8. unter Nr. 910 a - 9274/2010
9. Siegel

10. Unterschrift
Im Auftrag:

Christine Köppl

Christine Köppl
Amtsinspektorin



Certified Copy



Fraunhofer Gesellschaft

Fraunhofer-Gesellschaft Postfach 19 03 39 D-80603 München

Vorstand

Leonrodstraße 54
D-80636 München

Telefon +49 (0) 89/12 05-01
Telefax +49 (0) 89/12 05-700

Vollmacht

Hiermit bevollmächtigen wir

Herrn Dr. Werner Liebler
geb. 06.05.1961
geschäftsansässig: Patentstelle für
die Deutsche Forschung
Leonrodstraße 68
80636 München

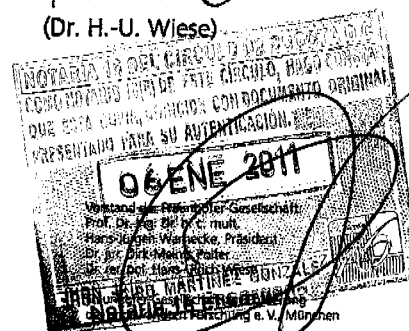
die

Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung e.V.,
Leonrodstraße 54, 80636 München

bis auf Widerruf in allen Angelegenheiten des Gesetzes über Arbeitnehmererfindungen
rechtsverbindlich zu vertreten. Darüberhinaus wird Herr Dr. Liebler bevollmächtigt, bei
Schutzrechtsverfahren im Ausland lokale Patentanwälte zu beauftragen.

(Dr. D.-M. Polter)

(Dr. H.-U. Wiese)



Bankverbindung: Deutsche Bank, München
Konto 7521 933 BLZ 700 700 10

URNr. 3076K

Ich beglaubige hiermit die Echtheit der vorstehenden, vor mir vollzogenen
Unterschrift von

Herrn Dr. Dirk-Meints Polter,
Vorstandsmitglied in München,

geschäftsansässig Leonrodstr. 54, 80636 München,
persönlich bekannt.

Hierzu bescheinige ich aufgrund Einsicht in das Vereinsregister des
Amtsgerichts München (VR-Nr. 4461) vom 12.07.2000, daß die Herren
Dr. Dirk-Meints Polter und Dr. Hans-Ulrich Wiese zur Vertretung des Vereins

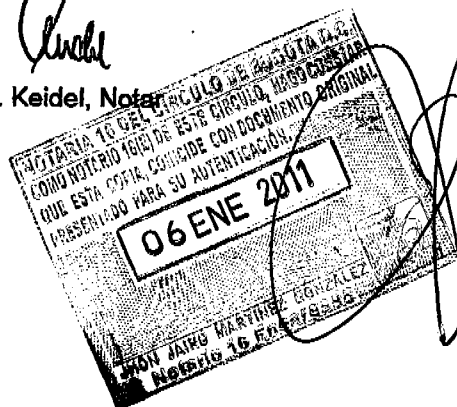
**Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung eingetragener Verein
mit dem Sitz in München**

als Vorstandsmitglieder gemeinsam berechtigt sind.

München, den 13.07.2000




Dr. Keidel, Notar



URNr. 3170 K

Ich beglaubige hiermit die Echtheit der vorstehenden, vor mir vollzogenen
Unterschrift von

Herrn Dr. Hans-Ulrich Wiese,
Vorstandsmitglied in München,

geschäftsansässig Leonrodstr. 54, 80636 München,
ausgewiesen durch amtlichen Lichtbildausweis.

Hierzu bescheinige ich aufgrund Einsicht in das Vereinsregister des
Amtsgerichts München –VR-Nr. 4461- vom 12.07.2000, dass die Herren
Dr. Dirk-Meints Polter und Dr. Hans-Ulrich Wiese zur Vertretung des Vereins

**Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung eingetragener Verein
mit dem Sitz in München**

als Vorstandsmitglieder gemeinsam berechtigt sind.

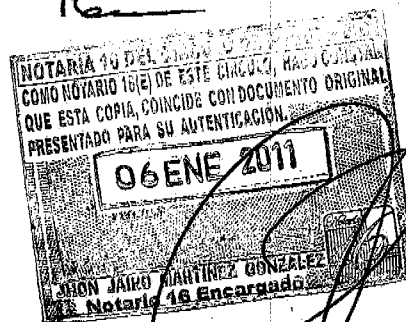
München, den 18.07.2000

Dr. Keidel, Notar

I hereby certify that this is a
true copy of the original document.

Munich, 06. Dez. 2010

Notary Public
in Munich



APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Land: Bundesrepublik Deutschland

Diese öffentliche Urkunde

2. ist unterschrieben von Thomas Haasen

3. In seiner Eigenschaft als Notar

4. Sie ist versehen mit dem Siegel des Notars Thomas Haasen in München

Bestätigt

5. in München

6. am 7. Dezember 2010

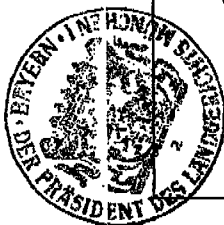
7. durch den Präsidenten des Landgerichts München I

8. unter Nr. 910 a - 9279/2010

9. Siegel

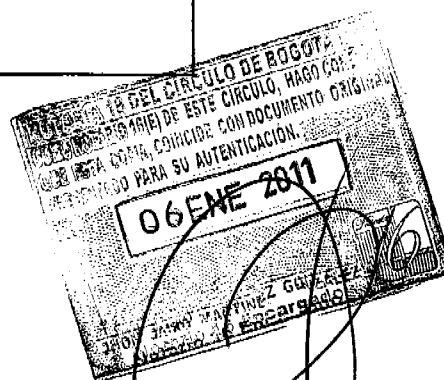
10. Unterschrift

Im Auftrag:



Christine Köppl

Christine Köppl
Amtsinspektorin



MARTHA SOFÍA SARMIENTO SOLANO - TRADUCTORA OFICIAL
Carrera 11D # 124 - 51 Apto. 510 Int. 2 - Teléfono 2139034 Bogotá - Colombia
E-mail: mssarmiento2@yahoo.com

Martha S. Sarmiento S.
Martha S. Sarmiento S.
Traductora Oficial Juramentada
Resolución No. 835 Minjusticia

TRADUCCIÓN OFICIAL DE LOS SIGUIENTES DOCUMENTOS: PODER DE ABOGADO otorgado por la Compañía Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., a la Firma Cárdenas y Cárdenas Abogados de Bogotá - Colombia, firmado por el señor Werner Liebler, Jefe Encargado del Departamento de Patentes de dicha Compañía, el 6 de diciembre de 2010. **CERTIFICACIÓN NOTARIAL**, Archivo No. H2294/2010, firmada por Thomas Haasen, Notario Público, en Munich, Alemania, con fecha del 6 de diciembre de 2010.

Estos documentos vienen acompañados por la **APOSTILLA DE LA CONVENCIÓN DE LA HAYA** No. 910 a - 9274/2010, en idioma alemán, con fecha del 7 de diciembre de 2010, firmada por Christine Köppl.

Traducción realizada por Martha Sofía Sarmiento Solano, Traductora e Intérprete Oficial Español - Inglés - Español, nombrada para el cargo mediante Resolución del Ministerio de Justicia, N° 835, del 3 de mayo de 1988.

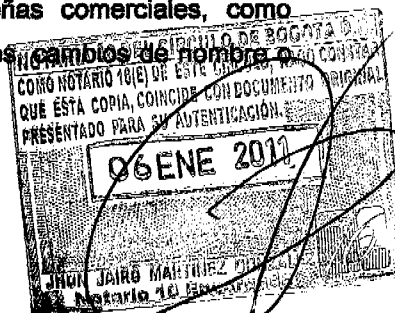
- COLOMBIA -

PODER DE ABOGADO

Yo, Werner Liebler, declaro lo siguiente:

PRIMERO: Que yo actúo en mi calidad de Jefe Encargado del Departamento de Patentes de Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., una Compañía legalmente constituida y en existencia de acuerdo con las leyes de Alemania y cuyo domicilio es Hansastrasse 27c, 80686 Muenchen Germany/Alemania.

SEGUNDO: Que haciendo uso de las facultades del cargo arriba mencionado, yo otorgo Poder de Abogado, especial, amplio y suficiente a **DARÍO CÁRDENAS NAVAS**, y/o a **LUZ HELENA ADARVE GÓMEZ**, y/o a **EDUARDO CÁRDENAS CABALLERO**, y/o a **BERNARDO P. CÁRDENAS MARTÍNEZ**, y/o a **JUANITA ACOSTA GÓMEZ**, y/o a **JULY GALINDO QUINTERO**, abogados registrados quienes trabajan para la firma **CÁRDENAS & CÁRDENAS - ABOGADOS**, con sede en la Carrera 7 No. 71-52, Torre B, Piso 9, en Bogotá, D.C., Colombia, para que gestionen en Colombia ante las Oficinas y Autoridades correspondientes la obtención de registros de marca comercial, patentes, modelos de utilidad, diseños y dibujos industriales, depósitos de nombres y enseñas comerciales, como también para el procesamiento de renovaciones, cesiones y cambios de nombre o



NO SE ASUME
RESPONSABILIDAD
DEL TERCERO

042637

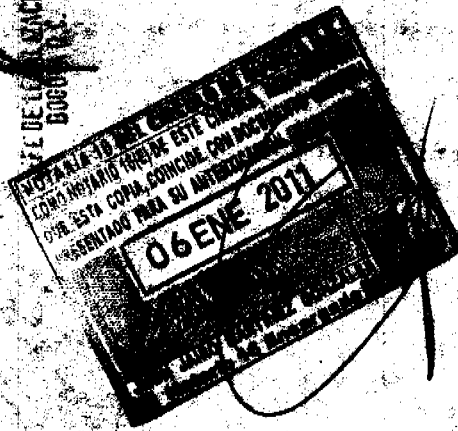
Ministerio de Gobierno

UNIDAD ADMINISTRATIVA ESPECIAL
BOGOTÁ D.C.

MINISTERIO DE
RELACIONES EXTERIORES

2000 ENE - E.A. 9 03

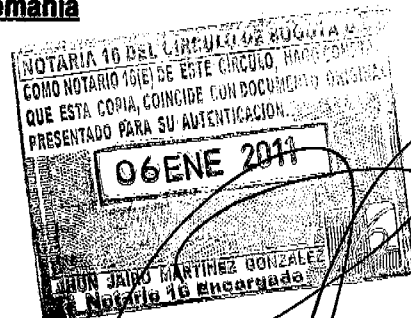
DEPARTAMENTO DE
BOGOTÁ D.C.



domicilio del titular, renuncias y licencias de lo arriba mencionado, propósitos para los cuales se les otorga Poder para que lleven a cabo ante dichas autoridades todas las acciones necesarias para los propósitos indicados, como radicar solicitudes, hacer declaraciones, reclamaciones, formular descripciones, realizar enmiendas o modificaciones, interponer recursos de oposición y apelaciones; pagar todos los impuestos y cuotas y hacer cualquier otro pago establecido por ley; recibir todos los documentos y dineros y emitir el descargo correspondiente; cumplir con cualquier otro requerimiento y tomar, en general, todas las medidas que consideren apropiadas y oportunas para defender y proteger la Propiedad Industrial de la Compañía en el evento de infracciones, oposiciones, procesos para cancelación o anulación, imposición de sanciones, presentar antecedentes ante las Cortes y Tribunales. Además, quedan facultados para comparecer como demandantes o demandados ante Jueces competentes, también se les da Poder para negociar, solicitar arbitraje, ceder, recibir, dimitir, rescindir o desistir, apelar, ratificar y todos los poderes legales adicionales que sean necesarios. También se declara que a partir de ahora todas las acciones que realicen dichos abogados respaldando este Poder de Representación y para beneficio de la Compañía son buenas y válidas. Adicionalmente se les otorga facultades para sustituir este Poder y si es necesario revocar dichas sustituciones.

FIRMADO EL: 6 de diciembre de 2010 EN: Munich, Alemania

Por: Werner Liebler
(Firma Autógrafa)



MARTHA SOFÍA SARMIENTO SOLANO - TRADUCTORA OFICIAL
Carrera 11D # 124 - 51 Apto. 510 Int. 2 - Teléfono 2139034 Bogotá - Colombia
E-mail: masarmiento2@yahoo.com

Archivo No. H2294/2010
ag

Yo, por la presente certifico que la anterior firma es verdadera, que fue suscrita en mi presencia por

El Sr. Dr. Werner Liebler, Nacido el 6 de mayo de 1961,
Dirección comercial: Hansastr. 27c, 80686 Munich,
a quien conozco personalmente.

El Dr. Werner Liebler actuando en representación de Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., con sus oficinas registradas en Munich,
Dirección comercial: 80686 München, Hansastr. 27c,
Hoy presentó ante mí en original los siguientes documentos: Poder de Abogado fechado el 13 de julio del 2000, Archivo No. 3076 K/2000, Poder de Abogado fechado el 18 de julio del 2000, Archivo No. 3170 K/2000, documentos expedidos y firmados por el Dr. Helmut Keidel, Notario Público de Munich, y como documento adjunto la Copia Certificada.

Munich, 6 de diciembre de 2010

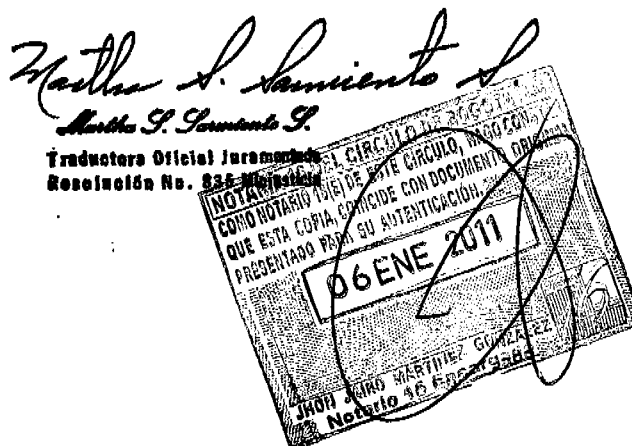
Firma Autógrafa
Thomas Haasen
Notario Público, Munich

Sello del Notario:
THOMAS HAASEN
NOTARIO EN MUNICH

H:\Grunert\Unterschriftbeglaubigungen\Dr. Liebler (Fraunhofer) engl 10-12-06.doc

CERTIFICO QUE ESTA ES UNA TRADUCCIÓN FIEL DEL INGLÉS AL ESPAÑOL DE LOS DOCUMENTOS QUE HE TENIDO A LA VISTA; QUE SOY TRADUCTORA E INTERPRETE OFICIAL ACREDITADA POR RESOLUCIÓN Nº 835 DEL 3 DE MAYO DE 1988 DEL MINISTERIO DE JUSTICIA Y QUE ESTOY REGISTRADA ANTE EL MINISTERIO DE RELACIONES EXTERIORES DE COLOMBIA.

Bogotá, D.C. 2 de enero de 2011



Copia autenticada

FhG – Fraunhofer Gesellschaft

Fraunhofer Gesellschaft Postfach 190339 D-80603 München

Junta Directiva Leonrodstrasse 54, D-80636 München

Tel.: +49 (0) 89 / 12 05 - 01 Telefax: +49 (0) 89 / 12 - 05 - 700

PODER

Por este medio, damos poder al

Señor Doctor Werner Liebler, nacido el 6 de mayo de 1961, con residencia comercial en
Patentstelle für die Deutsche Forschung, Leonrodstrasse 68, 80636 München
para representar, hasta nuevo aviso, con carácter obligatorio, a la firma

Fraunhofer-Gesellschaft zur Förderung der angewandten
Forschung e.V., Leonrodstrasse 54, 80636 München

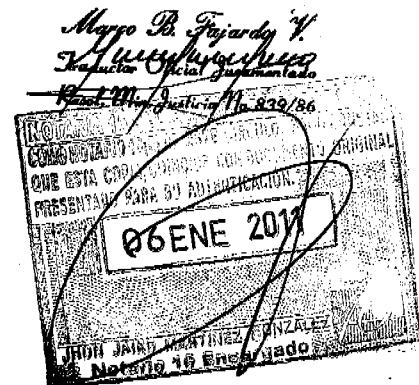
en todos los asuntos de ley sobre invenciones de los asalariados. Además, el señor Dr.
Liebler está autorizado a delegar abogados de patentes locales en el caso de procesos
legales de protección.

Dr. D.-M. Polter

(firma)

Dr. H.-U. Wiese

(firma)



URNo. 307 CK

Por este medio, autentico la veracidad de la firma anterior, trazada ante mí por el

Señor Dr. Dirk-Meints P o l t e r, miembro de la Junta directiva, en Munich,
con residencia comercial en Leonrodstrasse 68, 80636 München,
conocido en persona.

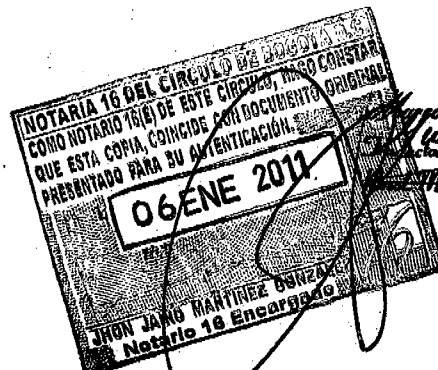
En ese sentido, certifico, con base en la inspección ocular en el registro de asociaciones
del Tribunal municipal de Munich (No. VR 4461), del 12 de julio de 2000, que los
señores Dr. Dirk-Meints Polter y Dr. Hans-Ulrich Wiese, como miembros de la Junta
directiva, están autorizados, en conjunto, para la representación de la asociación

**Fraunhofer-Gesellschaft zur Förderung der
Angewandten Forschung asociación registrada,
con domicilio en Munich.**

Munich, 13 de julio de 2000

Dr. Keidel, Notario

(firma y sello)



URNo. 3070 K

Por este medio, autentico la veracidad de la firma anterior, trazada ante mí por el

Señor Dr. Hans-Ulrich Wiese,

miembro de la Junta directiva, en Munich,

con residencia comercial en Leonrodstrasse 68, 80636 München,

identificado con documento oficial

conocido en persona.

En ese sentido, certifico, con base en la inspección ocular en el registro de asociaciones del Tribunal municipal de Munich (No. VR 4461), del 12 de julio de 2000, que los señores Dr. Dirk-Meints Polter y Dr. Hans-Ulrich Wiese, como miembros de la Junta directiva, están autorizados, en conjunto, para la representación de la asociación

**Fraunhofer-Gesellschaft zur Förderung der
Angewandten Forschung asociación registrada,
con domicilio en Munich.**

Munich, 18 de julio de 2000

Dr. Keidel, Notario

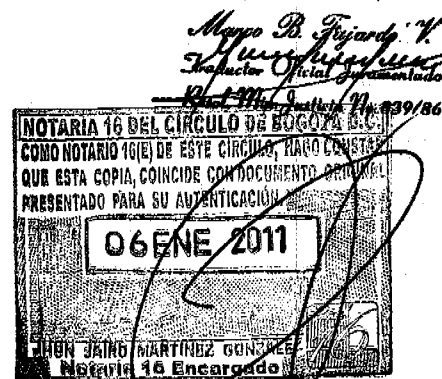
(firma y sello)

Por este medio certifico que el presente documento es una copia auténtica del original.

Munich, 6 de diciembre de 2010

El Notario, en Munich

(firma, sello y cordón)



APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Pais: **República Federal de Alemania**

el presente documento público

2. está firmado por Thomas Haasen

3. en su calidad de Notario

4. el documento está provisto del sello del Notario Thomas Haasen, en Munich

confirmado

5. en Munich

6. el 7 de diciembre de 2010

7. por el Presidente del Tribunal regional de Munich I

8. bajo el No. 910 a - 9279 / 2010

9. Sello:

10. Firma

en representación

Christine Köppl

Inspectora oficial

(firma, cordón y sello)

COPIA DEL DOCUMENTO ORIGINAL
EN LA APOSTILLA
042499
Mico Fossudo

NO SE ACOPIE
RESERVA EN LA
DEL TEXTO

2010 FEB - 11 A 8:16
JEFE DE LEGALIZACIONES
BOGOTA D.C.

RELACIONES EXTERIORES
MINISTERIO DE

06ENE 2011
JOHN JAIRO MORALES BONDALIZ
Notario Encargado
Margo B. Fajardo V.
Inspector Oficial Jurisdiccional
No. 839/86



Amtsgericht München -Registergericht-

VR 4461

Amtlicher aktueller Ausdruck aus dem Registerblatt

Datum der letzten Eintragung: 03.08.2010

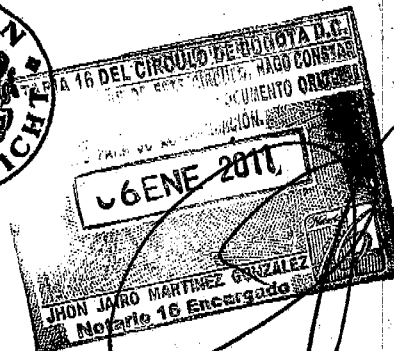
Datum des Abrufs: 19.10.2010

Ort und Tag der Ausstellung: München, den 19. Oktober 2010

Der Ausdruck bezeugt den Inhalt des Vereinsregisters.

Ersteller: Werdich, Justizangestellter,
Urkundsbeamter/Urkundsbeamtin der Geschäftsstelle

Dieser Ausdruck wird nicht unterschrieben und gilt als beglaubigte Abschrift.



Vereinsregister des Amtsgerichts München	Wiedergabe des aktuellen Registerinhalts Abruf vom 19.10.2010 08:33	Nummer des Vereins: VR 4461
-Ausdruck-	Seite 1 von 1	

Anzahl der bisherigen Eintragungen:

16

a) Name:

Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung eingetragener Verein

b) Sitz:

München

a) Allgemeine Vertretungsregelung:

Jeweils zwei Vorstandsmitglieder vertreten den Verein gemeinsam.

b) Vertretungsberechtigte und besondere Vertretungsbefugnis:

Vorstand: Prof. Dr. Buller, Ulrich, Golm, *01.06.1946

Vorstand: Prof. Dr.-Ing. Bullinger, Hans-Jörg, Stuttgart-Botnang, *13.04.1944

Vorstand: Prof. (Univ. Stellenbosch) Dr. Gossner, Alfred, Erharting, *15.10.1950

Vertretungsberechtigt in Gemeinschaft mit einem Vorstandsmitglied:

Besonderer Vertreter gemäß § 30 BGB für den Bereich Finanzen und Rechnungswesen: Meuer, Andreas, München, *12.06.1961

a) Satzung:

Eingetragener Verein

Satzung vom 26.03.1949

Zuletzt geändert durch Beschluss vom 20.05.2010

b) Sonstige Rechtsverhältnisse:

Der Verein hat im Wege der Abspaltung gemäß Spaltungs- und Übernahmevertrag vom 23.06.2009 sowie Beschluß seiner Mitgliederversammlung vom 24.06.2009 und Beschluß der Mitgliederversammlung des übertragenden Vereins vom 08.05.2009 Teile des Vermögens vom Forschungsinstitut für Pigmente und Lacke mit dem Sitz in Stuttgart (Amtsgericht Stuttgart VR 537) übernommen.

Die Forschungsgesellschaft für angewandte Naturwissenschaften e.V. mit dem Sitz in Bonn (Amtsgericht Bonn VR 2530) ist auf Grund des Verschmelzungsvertrages vom 23.06.2009 sowie des Beschlusses der Mitgliederversammlung vom 24.06.2009 und des Beschlusses der Mitgliederversammlung des übertragenden Vereins vom 04.06.2009 mit dem Verein verschmolzen.

Das Institut für Solare Energieversorgungstechnik Verein an der Universität Kassel mit dem Sitz in Kassel (Amtsgericht Kassel VR 2029) ist auf Grund des Angebots auf Abschluss eines Verschmelzungsvertrages vom 07.05.2009 und der Annahme des Angebots vom 23.06.2009 sowie des Beschlusses der Mitgliederversammlung vom 24.06.2009 und des Beschlusses der Mitgliederversammlung des übertragenden Vereins vom 07.05.2009 mit dem Verein verschmolzen.

5. a) Tag der letzten Eintragung:

03.08.2010



NOTARIA 16 DEL CIRCULO DE
COMO NOTARIO 16(E) DE ESTE CIRCULO
QUE ESTA COPIA, COINCIDE CON DOCUMENTO ORIGINAL
PRESENTADO PARA AUTENTICACION
10 ENE 2011
HON. JAIRO MARTINEZ GONZALEZ
Notario 16 Encarnación

Die Übereinstimmung dieses Ausdrucks mit dem Handelsregister
wird beglaubigt.

München, den 19. Oktober 2010

Amtsgericht München - Registergericht -

Jörg Werdich

Jörg Werdich, Justizangestellter
als Urkundsbeamter der Geschäftsstelle

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Land: Bundesrepublik Deutschland
Diese öffentliche Urkunde
2. ist unterschrieben von
Herrn Justizangestellter Jörg Werdich
3. in der Eigenschaft als
Urkundsbeamter der Geschäftsstelle
4. sie ist versehen mit dem Siegel des
Amtsgerichts München

Bestätigt

5. in München
6. am 19. Oktober 2010
7. durch den Richter am Amtsgericht
Dr. Ulrich Kühn als weiterem Vertreter des Präsidenten des Amtsgerichts München
8. unter Nr. GenA 9101 a Nr. 638/10
9. Stempel

10. Unterschrift

Dr. Ue

Dr. Ulrich Kühn



Tribunal municipal de Munich – Tribunal de registro

VR 4461

Extracto actual oficial de la hoja de registro

Fecha de la última anotación: 3 de agosto de 2010

Fecha de la petición de entrega: 19 de agosto de 2010

Lugar y fecha de expedición: Munich, 19 de octubre de 2010

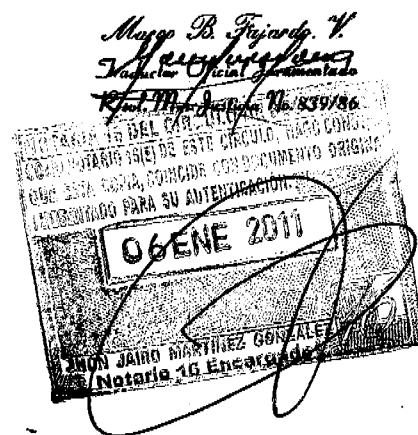
El extracto confirma el contenido del registro de la asociación

Proveedor: Werdich, funcionario judicial

Fedatario/funcionaria fedataria del despacho

El presente extracto no va firmado y rige como copia autenticada

(sello y cordón)



Registro de asociación del Tribunal municipal de Munich	Reproducción del contenido actual del Registro Petición del 19 de octubre de 2010 8.33 a.m.	Número de la asociación: VR 4461
- Extracto -	Página 1 de 1	

1. **Número de las anotaciones hechas hasta ahora:**

16

2. **a) Nombre:**

Fraunhofer-Gesellschaft zur Förderung de angewandten Forschung
eingetragener Verein

b) Domicilio

Munich

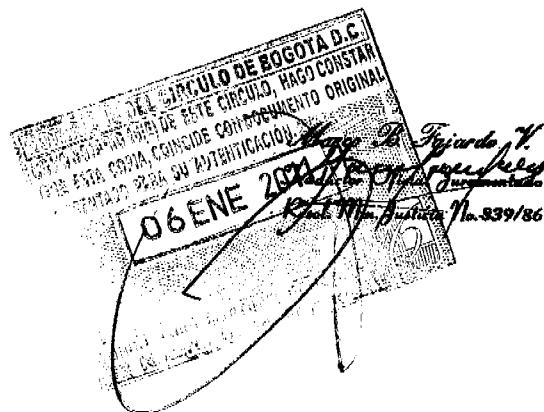
3. **a) Regulación general sobre representación:**

Cada vez dos miembros de la junta directiva representarán la asociación juntos

b) Autorizados a representar y permisos especiales de representación

Junta directiva: Prof. Dr. Buller, Ulrich, de Golm, nacido el 1 de junio de 1946

Junta directiva: Prof. Dr.-Ing. Bullinger, Hans-Jörg, de Stuttgart-Botnang, nacido
el 13 de abril de 1944



Junta directiva: Prof. (Universidad de Stellenbosch) Dr. Gossner, Alfred, de Erharting, nacido el 15 de octubre de 1950

Autorizado a representar junto con un miembro de la junta directiva:

Representante especial, según parágrafo 30 del código civil para el área de las finanzas y la contabilidad: Meuer, Andreas, de Munich, nacido el 12 de junio de 1961

4. a) Estatutos:

Sociedad registrada

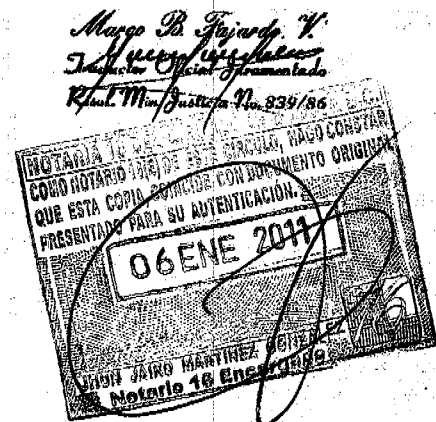
Estatutos del 26 de marzo de 1949

Última modificación según acuerdo del 20 de mayo de 2010

b) Otras relaciones jurídicas:

En el transcurso de la desmembración de la asociación cesionista, según contrato de desmembración y recepción, del 23 de junio de 2009, así como del acuerdo de su asamblea de socios, del 24 de junio de 2009, y del acuerdo de la asamblea de socios de la asociación cesionista, del 8 de mayo de 2009, la asociación tomó parte del patrimonio del Instituto para la investigación de pigmentos y lacas (Forschungsinstitut für Pigmente und Lacke), con domicilio en Stuttgart (Tribunal municipal de Stuttgart, referencia VR 537).

La Forschungsgesellschaft für angewandte Naturwissenschaften e.V. (Sociedad de investigación para ciencias naturales aplicadas, asociación registrada), con



domicilio en Bonn (Tribunal municipal de Bonn VR 2530) se fusionó con la asociación, con base en el contrato de fusión, del 23 de junio de 2009, así como del acuerdo de la asamblea de socios, del 24 de junio de 2009 y del acuerdo de la asamblea de socios de la asociación cesionista, del 4 de junio de 2009.

El Instituto para técnicas de aprovisionamiento de energía solar (Institut für Solare Energieversorgungstechnik) asociación de la Universidad de Kassel, con domicilio en Kassel (Tribunal municipal de Kassel VR 2029), se fusionó con la asociación, con base en la oferta de firma de un contrato de fusión, del 7 de mayo de 2009, y de la aceptación de tal oferta, del 23 de junio de 2009, así como del acuerdo de la asamblea de socios, del 24 de junio de 2009 y del acuerdo de la asamblea de socios de la sociedad cesionista, del 7 de mayo de 2009.

5. a) Fecha de la última anotación:

3 de agosto de 2010

Se autentica la coincidencia del extracto con el registro comercial.

Munich, 19 de octubre de 2010

Tribunal municipal de Munich – Tribunal de registro

Jörg Werdich, funcionario judicial

como funcionario fedatario del despacho

(firma, cordón y sello)



APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. País: **República Federal de Alemania**

el presente documento público

2. está firmado por el funcionario judicial Jörg Werdich

3. en su calidad de funcionario fedatario del despacho

4. el documento está provisto del sello del Tribunal municipal de Munich

confirmado

5. en Munich

6. 19 de octubre de 2010

7. por el Juez del tribunal municipal, Dr. Ulrich Kühn, como uno de los representantes
del Presidente del tribunal municipal de Munich

8. bajo el No. GenA 9101 a No. 636/10

9. Sello:

10. Firma

Dr. Ulrich Kühn

(firma y sello)

MINISTERIO DE
RELACIONES EXTERIORES
NO SE ASUME
RESPONSABILIDAD
DEL TEXTO
03827b
Copia firmada
DOCUMENTO DESEMPENA
ASFUNCIONES INDICADAS
BOGOTA D.C.
Jefe de LEGALIZACIONES

MINISTERIO DE RELACIONES EXTERIORES
RECIBO LEGALIZACION
No. 13843

NOTARIA 16 DEL CÍRCULO DE BOGOTÁ D.C.
COMO NOTARIO 16(B) DE ESTE CÍRCULO, HAGO CONSTAR
QUE ESTA COPIA, COINCIDE CON DOCUMENTO ORIGINAL
PRESENTADO PARA SU AUTENTICACION.
06 ENE 2011
JHON JAIR MARTINEZ GONZALEZ
Notario 16 Encargado

MÉTODO Y DISCRIMINADOR PARA CLASIFICAR DISTINTOS SEGMENTOS DE UNA SEÑAL

Fondo del Invento

5 El invento se refiere a un enfoque para la clasificación de distintos segmentos de una señal que comprende segmentos de por lo menos un primer tipo y de un segundo tipo. Las realizaciones del invento se refieren al campo de la codificación de audio y, en particular, a la discriminación de voz y música al codificar una señal de audio.

10 En el arte previo se conocen los esquemas de codificación en el dominio de la frecuencia tales como MP3 o AAC. Estos codificadores del dominio de la frecuencia se basan en una conversión del dominio del tiempo al dominio de la frecuencia, una fase de cuantificación, en la cual el error de cuantificación se controla usando la información de un módulo psicoacústico y una fase de codificación, en la
15 cual se codifica por entropía a los coeficientes espectrales cuantificados y a la correspondiente información lateral mediante unas tablas de códigos.

Por otro lado, existen unos codificadores que son muy apropiados para procesar la voz tales como el AMR-WB+ como se describe en el documento 3 GPP TS 26,290. Tales esquemas de codificación de voz llevan a cabo un filtrado de
20 predicción lineal (PL) de una señal del dominio del tiempo. Tal filtrado PL se deriva de un análisis de predicción lineal de la señal de entrada del dominio del tiempo. Luego se codifican los coeficientes de filtrado PL, que se dan como resultado, y se transmiten como información lateral. El proceso se conoce como codificación de predicción lineal (CPL). En la salida del filtro, la señal de predicción residual o la
25 señal de error de predicción, la cual se conoce también como la señal de excitación, se codifica mediante las fases de análisis por síntesis del codificador ACELP o, de manera alternativa, se codifica mediante un codificador transformador que utiliza una transformación de Fourier con una superposición. La decisión entre la codificación por ACELP y la codificación por excitación codificada por transformación, que se
220.300 llama también codificación XCT, se lleva a cabo usando un algoritmo de lazo cerrado o de lazo abierto.

Los esquemas de codificación de audio en el dominio de la frecuencia tales como el esquema de codificación tipo AAC de alta eficiencia, que combina un esquema de codificación tipo AAC y una técnica de réplica de ancho de banda espectral, se puede combinar también con una herramienta de codificación de
5 estéreo conjunto o de canales múltiples, la cual se conoce bajo el término "MPEG surround". Los esquemas de codificación en el dominio de la frecuencia tienen ventajas por el hecho de que con bajas velocidades de transferencia de bits muestran una alta calidad para señales de música. Sin embargo, las bajas velocidades de transferencia de bits son problemáticas para la calidad de señales de
10 voz.

Por otro lado, los codificadores de voz tales como el AMR-WB+ también tienen una fase de mejoramiento de alta frecuencia y una funcionalidad de estéreo. Los esquemas de codificación de voz muestran una alta calidad para señales de voz aún con bajas velocidades de transferencia de bits, pero muestran una baja calidad
15 para señales de música con bajas velocidades de transferencia de bits.

En vista de los esquemas de codificación disponibles que se describen más arriba, y de los cuales algunos son más apropiados para la codificación de voz y otros son más apropiados para la codificación de música, la segmentación y clasificación automáticas de una señal de audio a ser codificada son importantes
20 herramientas en muchas aplicaciones multimedia y se pueden usar con el fin de seleccionar un proceso apropiado para cada categoría diferente que ocurre en la señal de audio. El rendimiento total de la aplicación depende mucho de la confiabilidad de la clasificación de la señal de audio. De hecho, una clasificación equivocada puede generar selecciones y sintonizaciones no apropiadas de un los
25 procesos siguientes.

La Figura 6 muestra un diseño de un aparato de codificación convencional usado para codificar por separado música y voz, que depende de la discriminación de una señal de audio. El diseño del aparato de codificación comprende una rama de codificación de voz 100 que incluye un codificador de voz apropiado 102, por ejemplo
30 un codificador de voz AMR-WB+ tal como se describe en el "Codificador-decodificador (AMR-WB+) adaptivo, extendido de banda ancha de múltiples

velocidades" ("Extended Adaptive Multi-Rate – Wideband (AMR-WB+) codec"), 3GPP TS 26.290 V6.3.0, 2005-06, descripción técnica. El diseño del aparato de codificación comprende además una rama de codificación de música 104 que incluye un codificador de voz 106, por ejemplo un codificador de música tipo AAC tal como se describe, por ejemplo, en "La Codificación Genérica de Imágenes en Movimiento y de Audio Asociado: Codificación de Audio Avanzado" (Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding.) International Standard 13818-7, ISO/IEC JTC1/SC29/WG11 Grupo de Expertos en Imágenes en Movimiento (Moving Pictures Expert Group), 1997.

10 Las salidas de los codificadores 102 y 106 están conectados con la entrada de un multiplexor 108. Las entradas de los codificadores 102 y 106 se pueden conectar selectivamente con una línea de entrada 110 que trae una señal de audio de entrada. La señal de audio de entrada se aplica selectivamente al codificador de voz 102 o al codificador de música 106 mediante un conmutador 112
15 que se muestra esquemáticamente en la Figura 6 y que está controlado por un control de conmutación 114. El diseño del aparato de codificación comprende además a un discriminador de voz y música 116 que también recibe en una entrada la señal de audio de entrada y que emite por una salida una señal de control al control de conmutación 114. El control de conmutación 114 emite por otra salida
20 también una señal indicadora de modo sobre una línea 118 que se dirige a una segunda entrada del multiplexor 108 de modo que se puede enviar una señal indicadora de modo junto con una señal codificada. La señal indicadora de modo puede tener sólo un bit, que indica que un bloque de datos asociado con el indicador de modo es o una voz codificada o una música codificada, de modo que, por
25 ejemplo, no hace falta hacer una discriminación en un decodificador. En realidad, sobre la base del bit indicador de modo transmitido junto con los datos codificados al lado del decodificador se puede generar una señal conmutadora apropiada sobre la base del indicador de modo para dirigir los datos recibidos y codificados a un decodificador apropiado de voz o de música.

30 La Figura 6 muestra un diseño tradicional de un aparato de codificación que se usa para codificar digitalmente señales de voz y de música aplicadas a la

línea 110. En general, los codificadores de voz trabajan mejor con señales de voz y los codificadores de audio trabajan mejor con señales de música. Un esquema de codificación universal se puede diseñar usando un sistema de múltiples codificadores que cambian de un codificador a otro de acuerdo con la naturaleza de la señal de entrada. Aquí el problema no trivial es diseñar un clasificador de señal de entrada muy apropiado que accione el elemento conmutador. El clasificador es el discriminador de voz y música 116 que se muestra en la Figura 6. Usualmente la clasificación confiable de una señal de audio introduce un retardo importante, mientras, por otro lado, el retardo es un factor importante en las aplicaciones en tiempo real.

En general, se desea que el retardo algorítmico total introducido por el discriminador de voz y música es suficientemente corto para que permita que se usen los aparatos de codificación conmutados en una aplicación en tiempo real.

La Figura 7 muestra los retardos que se experimentaron en un diseño de aparato de codificación como se muestra en la Figura 6. Se supone que la señal aplicada sobre la línea de entrada 110 debe ser codificada sobre una base de cuadro de 1024 muestras con una tasa de muestreo de 16 kHz de modo que la discriminación de voz y música debe entregar una decisión en cada cuadro es decir, cada 64 milisegundos. La transición entre dos codificadores se realiza por ejemplo en una manera que se describe en el documento WO 2008/071353 A2 y el discriminador de voz y música no debe aumentar significativamente el retardo algorítmico de los decodificadores conmutados que en total es de unas 1600 muestras tomando en cuenta el retardo que se necesita para el discriminador de voz y música. Además se desea proveer la decisión de voz o música para el mismo cuadro en el cual se decide la conmutación del bloque de AAC. La situación se ilustra en la Figura 7, que muestra un bloque de AAC largo 120, que tiene una longitud de 2048 muestras, es decir el bloque largo 120 comprende dos cuadros de 1024 muestras, un bloque ACC corto 122 un cuadro de 1024 muestras y un súper cuadro AMR-WB+ 124 un cuadro de 1024 muestras.

En la Figura 7 se toman la decisión de conmutación de bloque de AAC y la decisión de voz o música sobre los cuadros 126 y 128, respectivamente, de 1024

muestras, que cubren el mismo periodo de tiempo. Las dos decisiones se toman en esta posición particular para permitir a la codificación que utilice al mismo tiempo ventanas de transición para pasar adecuadamente de un modo al otro. En consecuencia, un retardo mínimo de $512+64$ muestras está introducida por las dos
5 decisiones. Este retardo se debe sumar el retardo de 1024 muestras generadas por la superposición de 50% de la AAC MDCT, lo cual da como resultado un retardo mínimo de 1600 muestras. En una AAC convencional, sólo se hace la conmutación de bloque y el retardo es exactamente 1600 muestras. Se requiere este retardo para conmutar al mismo tiempo de un bloque largo a los bloques cortos cuando se
10 detectan componentes transitorios en el cuadro 126. Esta conmutación del largo de transformación es deseable para impedir un artefacto pre-eco. En cualquier caso (bloques largos o cortos) el cuadro decodificado 130 en la Figura 7 representa el primer cuadro entero que se puede restituir en el lado del decodificador.

En un aparato de codificación conmutado que utiliza una AAC como
15 codificador de música, la decisión de conmutación que viene de una fase de decisión debe evitar demasiado retardo adicional al retardo original de la AAC. El retardo adicional viene del cuadro anticipado 132 que es necesario para el análisis de señal en la fase de decisión. Con una tasa de muestreo de por ejemplo 16 kHz, el retardo de la AAC es de 100 ms mientras un discriminador convencional de voz y música
20 utiliza alrededor de 500 ms de anticipación, lo cual da como resultado una estructura de codificación conmutada con un retardo de 600 ms. Entonces, el retardo total correspondería a seis veces del retardo de la AAC original.

Los enfoques convencionales que se describen más arriba son desventajosos debido a que, para una clasificación confiable de una señal de audio,
25 se introduce un retardo no deseado de modo que existe la necesidad de un nuevo enfoque para la discriminación de una señal que incluye segmentos de distintos tipos, en lo cual un retardo algorítmica adicional introducida por el discriminador es suficientemente pequeña de modo que los aparatos de codificación conmutados se puedan usar también para una aplicación en tiempo real.

30 J. Wang, et. al. "Clasificación de voz y música en tiempo real con un árbol de decisión oblicuo jerárquico" ("Real-time speech/music classification with a

hierarchical oblique decision tree", ICASSP 2008, Congreso Internacional de Acústica de IEEE, Procesamiento de Voz y Señal (IEEE International Conference on Acoustics, Speech and Signal Processing), 2008, 31 de marzo de 2008 hasta el 4 de abril de 2008, describen un enfoque para la clasificación de voz y música usando rasgos distintivos de corto plazo y rasgos distintivos de largo plazo derivados de la misma cantidad de cuadros. Estos rasgos distintivos de corto plazo y rasgos distintivos de largo plazo se usan para clasificar la señal, pero sólo se explotan propiedades limitadas de los rasgos distintivos de corto plazo, por ejemplo no se explota la reactividad de la clasificación, aunque tienen un rol importante para la mayoría de las aplicaciones de codificación de audio.

Síntesis del Invento

Es un objetivo del invento proveer un enfoque mejorado para discriminar en una señal segmentos de distintos tipos, mientras se mantenga corta cualquier retardo introducida por la discriminación.

Este objetivo se obtiene mediante un método de acuerdo con la reivindicación 1 y por un discriminador de acuerdo con la reivindicación 14.

Una realización del invento provee un método para clasificar distintos segmentos de una señal, en el cual la señal comprende segmentos de por lo menos un primer tipo y un segundo tipo, y el método comprende:

clasificar en un corto plazo la señal sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal y entregar un resultado de clasificación de corto plazo;

clasificar en un largo plazo la señal sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídas de la señal y entregar un resultado de clasificación de largo plazo; y

combinar el resultado de la clasificación de corto plazo con el resultado de la clasificación de largo plazo para proveer una señal de salida que indica si un segmento de la señal es del primer tipo o del segundo tipo.

Otra realización del invento provee un discriminador que comprende:

un clasificador de corto plazo configurado para recibir una señal y proveer un resultado de clasificación de corto plazo de la señal sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal, en lo cual la señal comprende segmentos de por lo menos un primer tipo y un segundo tipo;

- 5 un clasificador de largo plazo configurado para recibir la señal y proveer un resultado de clasificación de largo plazo de la señal sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal; y

- 10 un circuito de decisión configurado para combinar el resultado de la clasificación de corto plazo con el resultado de la clasificación de largo plazo para proveer una señal de salida que indica si un segmento de la señal es del primer tipo o del segundo tipo.

- 15 Algunas realizaciones del invento proveen la señal de salida sobre la base de una comparación del resultado del análisis de corto plazo con el resultado del análisis de largo plazo.

- 20 Algunas realizaciones del invento se refieren a un enfoque para clasificar distintos segmentos de corto plazo no superpuestos de una señal de audio o bien como segmento de voz o como segmento de no voz o como otra categoría de segmento. Este enfoque se basa sobre la extracción de rasgos distintivos y el análisis de sus estadísticas sobre dos distintas longitudes de las ventanas de análisis. La primera ventana es larga y mira principalmente hacia el pasado. La primera ventana se utiliza para obtener un indicio de decisión confiable pero retardada para la clasificación de la señal. La segunda ventana es corta y considera principalmente el segmento que se procesa en ese instante o el segmento actual. La
25 segunda ventana se utiliza para obtener un indicio de decisión instantáneo. Los dos indicios de decisión se combinan de manera óptima preferiblemente utilizando una decisión de histéresis que recibe la información de memoria desde el indicio retardado y la información instantánea desde el indicio instantáneo.

- 30 Algunas realizaciones del invento utilizan rasgos distintivos de corto plazo tanto en el clasificador de corto plazo y en el clasificador de largo plazo de modo que los dos clasificadores explotan distintas estadísticas del mismo rasgo distintivo. El

clasificador de corto plazo sólo extrae la información instantánea porque sólo tiene acceso a un juego de rasgos distintivos. Por ejemplo, puede explotar el promedio de los rasgos distintivos. Por el otro lado, el clasificador de largo plazo tiene acceso a varios rasgos distintivos porque considera varios cuadros. Como una consecuencia, el clasificador de largo plazo puede explotar más rasgos distintivos de la señal explotando estadísticas sobre varios cuadros que el clasificador de corto plazo. Por ejemplo, el clasificador de largo plazo puede explotar la variación de los rasgos distintivos o la evolución en el tiempo. De ese modo, el clasificador de largo plazo puede explotar más información que el clasificador de corto plazo, pero introduce retardos o tiempos de espera. Sin embargo, los rasgos distintivos de largo plazo, a pesar de introducir retardos o esperas de tiempo hacen que los resultados de clasificación de largo plazo sean más robustos y confiables. En algunas realizaciones los clasificadores de corto plazo y de largo plazo pueden considerar los mismos rasgos distintivos de corto plazo, que se pueden calcular sólo una vez y pueden ser utilizadas por ambos clasificadores. De ese modo, en tal realización el clasificador de largo plazo puede recibir los rasgos distintivos de corto plazo directamente del clasificador de corto plazo.

El nuevo enfoque permite de esa manera obtener una clasificación más robusta, mientras se introduce un retardo de poca duración. A diferencia de los enfoques tradicionales, las realizaciones del invento limitan los retardos introducidos por la decisión de voz o música, mientras mantienen una decisión confiable. En una realización del invento, la anticipación se limita a 128 muestras que da como resultado un retardo total de 108 ms.

25

Breve Descripción de los Dibujos

Las realizaciones del invento se describirán más abajo haciendo referencia a los dibujos adjuntos, en los cuales:

- Figura 1 es un diagrama de bloque de un discriminador de voz y música de acuerdo con una realización del invento;
- 30 Figura 2 muestra la ventana de análisis utilizada por los clasificadores de largo plazo y de corto plazo del discriminador de la Figura 1;

- Figura 3 muestra una decisión de histéresis utilizada en el discriminador de la Figura 1;
- Figura 4 es un diagrama de bloque de un esquema de codificación ejemplar que comprende un discriminador de acuerdo con algunas realizaciones del invento;
- Figura 5 es un diagrama de bloque de un esquema de decodificación que corresponde al esquema de codificación de la Figura 4
- Figura 6 muestra un diseño de aparato de codificación convencional utilizado para codificar de manera separada voz y música dependiendo de una discriminación de una señal de audio; y
- Figura 7 muestra los retardos que se experimentan en el diseño de aparato de codificación que se muestra en la Figura 6.

Descripción de las Realizaciones del Invento

- La Figura 1 es un diagrama de bloque de un discriminador de voz y música 116 de acuerdo con una realización del invento. el discriminador de voz y música 116 comprende un clasificador de corto plazo 150 que recibe en una entrada del mismo una señal de entrada, por ejemplo una señal de audio que comprende segmentos de voz y de música. El clasificador de corto plazo 150 emite sobre una línea de salida 152 un resultado de clasificación de corto plazo, el indicio de decisión instantáneo. El discriminador 116 comprende además un clasificador de largo plazo 154 que también recibe la señal de entrada y emite sobre una línea de salida 156 el resultado de clasificación de largo plazo, el indicio de decisión retardada. Además se provee un circuito de histéresis 158 que combina las señales de salida del clasificador de corto plazo 150 y del clasificador de largo plazo 154 en una manera, que se describirá más abajo con más detalle para generar una decisión de voz o música que se emite a la línea de salida 160 y que se puede utilizar para controlar el futuro procesamiento de un segmento de una señal de entrada en una manera que se describió más arriba con respecto a la Figura 6, es decir la señal de decisión de voz o música 160 se puede utilizar para dirigir el segmento de señal de entrada, que ha sido clasificado, a un codificador de voz o un codificador de audio.

De ese modo, de acuerdo con las realizaciones del invento, se utilizan dos distintos clasificadores 150 y 154 en paralelo sobre la señal de entrada aplicada a los respectivos clasificadores vía la línea de entrada 110. Los dos clasificadores se llaman clasificador de largo plazo 154 y clasificador de corto plazo 150, en lo cual los dos clasificadores se distinguen analizando las estadísticas de los rasgos distintivos sobre las cuales funcionan a través de unas ventanas de análisis. Los dos clasificadores entregan las señales de salida 152 y 156, a saber los indicios de decisión instantáneos (IDI) y los indicios de decisión retardados (IDD). El clasificador de corto plazo 150 genera un IDI sobre la base de rasgos distintivos de corto plazo que tienen como meta capturar informaciones instantáneas con respecto a la naturaleza de la señal de entrada. Ellas están relacionados con los atributos de corto plazo de la señal que puede cambia rápidamente y en cualquier momento. En consecuencia, se espera que los rasgos distintivos de corto plazo sean reactivos y no introduzcan grandes retardos al proceso de discriminación en su totalidad. Por ejemplo, debido a que la voz se considera que es cuasiestacionario en duraciones de 5 a 20 ms, los rasgos distintivos de corto plazo se pueden calcular para cada cuadro de 16 ms en una señal que se toma con una tasa de muestreo de 16 kHz. El clasificador de largo plazo 154 genera los IDD sobre la base de rasgos distintivos que se dan como resultado de observaciones más largas de la señal (rasgos distintivos de largo plazo) y por lo tanto permiten lograr una clasificación más confiable.

La Figura 2 muestra las ventanas de análisis utilizadas por el clasificador de largo plazo 154 y el clasificador de corto plazo 150 que se muestran en la Figura 1. Suponiendo un cuadro de 1024 muestras con una tasa de muestreo de 16 kHz, la longitud de la ventana del clasificador de largo plazo 162 es de $4 \times 1024 + 128$ muestras, es decir, la ventana del clasificador de largo plazo 162 se extiende a lo largo de cuatro cuadros de la señal de audio y de unas 128 muestras adicionales que necesita el clasificador de largo plazo 154 para llevar a cabo su análisis. Este retardo adicional a la cual se refiere también como "anticipación" está indicada en la Figura 2 bajo el número de referencia 164. La Figura 2 muestra también la ventana del clasificador de corto plazo 166 que es de $1024 + 128$ muestras, es decir se extiende a

lo largo de un cuadro de la señal de audio y el retardo adicional que se necesita para analizar el segmento actual. El segmento actual está indicado con el número 128 como el segmento para el cual hace falta hacer la decisión de voz o música.

La ventana del clasificador de largo plazo indicado en la Figura 2 es
5 suficientemente larga para obtener el rasgo distintivo de modulación de energía de 4 Hz de la voz. La modulación de energía de 4 Hz es un rasgo distintivo importante y discriminatoria de la voz, la cual se explota tradicionalmente en los discriminadores robustos de voz y música usados, por ejemplo, por Scheirer E. y Slaney M., "Diseño y evaluación de un discriminador robusto de voz y música de múltiples rasgos
10 distintivos" ("Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator"), ICASSP'97, Munich, 1997. La modulación de energía de 4 Hz es un rasgo distintivo que sólo se puede extraer observando la señal sobre un segmento de tiempo largo. El retardo adicional que se introduce por el discriminador de voz y música es igual a la anticipación 164 de 128 muestras, la cual se necesita por cada
15 una de las clasificadores 150 y 154 para llevar a cabo un análisis de predicción lineal perceptiva tal como está descrito por H. Hermansky, "Análisis de predicción lineal perceptiva (plp) de la voz" ("Perceptive linear prediction (plp) analysis of speech"), Journal of the Acoustical Society of America, vol. 87, no. 4, pp. 1738 - 1752, 1990 y H. Hermansky, et al., "Análisis predictivo lineal de la voz basado en la percepción"
20 ("Perceptually based linear predictive analysis of speech"), ICASSP 5.509 - 512, 1985. De ese modo, cuando se utiliza el discriminador de la realización descrita más arriba en un diseño de codificación tal como se muestra en la Figura 7, el retardo total de los codificadores conmutados 102 y 106 será de $1600+128$ muestras, lo cual es igual a 108 milisegundos que se suficientemente bajo para aplicaciones en
25 tiempo real.

Ahora se hace referencia a la Figura 3 que describe la combinación de las señales de salida 152 y 156 de los clasificadores 150 y 154 del discriminador 116 para obtener una señal de decisión de voz o música 160. El indicio de decisión retardada IDD y el indicio de decisión instantáneo IDI de acuerdo con una realización
30 del invento se combinan utilizando una decisión de histéresis. Los procesos de histéresis se utilizan ampliamente para registrar decisiones de proceso para

estabilizarlas. La Figura 3 muestra una decisión de histéresis de dos estados como una función del IDD y del IDI para determinar si la señal de decisión de voz o música debe indicar que un segmento que se está procesando actualmente de la señal de entrada como un segmento de voz o un segmento de música. El ciclo de histéresis característico se puede ver en la Figura 3 y el IDI y el IDD están normalizados por los clasificadores 150 y 154 de tal manera que los valores están entre -1 y 1, en lo cual -1 significa que la probabilidad es totalmente del tipo música, y 1 significa que la probabilidad es totalmente del tipo voz.

La decisión se basa sobre el valor de una función $F(\text{IDI}, \text{IDD})$, unos ejemplos de la cual se describirán más abajo. En la Figura 3 $F_1(\text{IDD}, \text{IDI})$ indica un umbral, el cual $F(\text{IDI}, \text{IDD})$ debe cruzar para ir de un estado de música a un estado de voz. $F_2(\text{IDD}, \text{IDI})$ muestra un umbral, el cual $F(\text{IDI}, \text{IDD})$ debe cruzar para ir de un estado de voz a un estado de música. La decisión final $D(n)$ para un segmento actual o un cuadro actual que tiene el índice n , se puede calcular entonces sobre la base del siguiente pseudo código:

```
% Pseudo código de decisión de histéresis
If(D(n-1) == música)
    If(F(IDI, IDD) < F1(IDD, IDI))
        D(n) == música
    Else
        D(n) == voz
Else
    If(F(IDI, IDD) > F2(IDD, IDI))
        D(n) == voz
    Else
        D(n) == música

% Fin de pseudo código de decisión de histéresis
```

De acuerdo con algunas realizaciones del invento, la función $F(\text{IDI}, \text{IDD})$ y los umbrales, que se mencionan más arriba, se fijan de la siguiente manera:

```
F(IDI, IDD) = IDI
F1(IDI, IDD) = 0.4-0.4*IDD
F2(IDI, IDD) = -0.4-0.4*IDD
```

De manera alternativa, se pueden usar las definiciones siguientes:

```
F(IDI, IDD) = (2*IDI + IDD)/3
F1(IDI, IDD) = -0.75 * IDD
```

$$F2(IDI, IDD) = -0.75 * IDD$$

Cuando se utiliza la última decisión, el ciclo de histéresis se anula y la decisión se toma sólo sobre la base de un umbral adaptivo único.

El invento no se limita a la decisión de histéresis que se describe más arriba. En lo que sigue, se describirán las realizaciones adicionales para la combinación de los resultados de análisis para obtener la señal de salida.

Una determinación simple de umbrales se puede utilizar en lugar de la decisión de histéresis constituyendo el umbral de una manera que el mismo explota los rasgos distintivos tanto del IDD como del IDI. Se considera que el IDD es un indicio de discriminación más confiable porque vienen de una observación más larga de la señal. Sin embargo, el IDD se calcula parcialmente sobre la base de una observación del pasado de la señal. Un clasificador convencional, que sólo compara el valor de IDD con el umbral 0 y que clasifica un segmento como tipo voz cuando $IDD > 0$ o como tipo música en el caso contrario, tendrá una decisión retardada. En una realización del invento, podemos adaptar la determinación de umbrales explotando el IDI y hacer la decisión más reactiva. Por este propósito, se puede adaptar el umbral sobre la base del siguiente pseudo código:

```
% Pseudo código de la determinación adaptiva de umbrales
20 If(IDD > -0.5*IDI)
    D(n) == voz
Else
    D(n) == música
% Fin de la determinación adaptiva de umbrales
```

En otra realización se puede utilizar el IDD para hacer más confiable al IDI. El IDI se conoce como ser reactivo pero no tan confiable como el IDD. Además, la observación de la evolución del IDD entre el segmento pasado y el segmento actual puede dar otra indicación como el cuadro 166 en la Figura 2 influye sobre el IDD calculado para el segmento 162. La notación $IDD(n)$ se utiliza para el valor actual del IDD y $IDD(n-1)$ para el valor pasado. Utilizando ambos valores, $IDD(n)$ e $IDD(n-1)$, se puede hacer que el IDI sea más confiable utilizando un árbol de decisión tal como se describe en lo que sigue:

```
% Pseudo código del árbol de decisión
35 If(IDI > 0 && IDD(n) > 0)
    D(n) = voz
```

```
Else if (IDI < 0 && IDD(n) < 0)
    D(n) = música
Else if (IDI > 0 && IDD(n) - IDD(n-1) > 0)
    D(n) = voz
5 Else if (IDI < 0 && IDD(n) - IDD(n-1) < 0)
    D(n) = música
Else if (IDD > 0)
    D(n) = voz
Else
10 D(n) = música
% Fin del árbol de decisión
```

En el árbol de decisión, que se muestra arriba, la decisión se toma directamente si ambos indicios muestran la misma probabilidad. Si los dos indicios dan indicaciones contradictorias, miramos a la evolución del IDD. Si la diferencia
15 $IDD(n) - IDD(n-1)$ es positiva, podemos suponer que el segmento actual es del tipo voz. De otra manera, podemos suponer que el segmento actual es del tipo música. Si esta nueva indicación va en la misma dirección que el IDI, se toma la decisión final. Si ambos intentos fracasan en dar una decisión clara, se toma la decisión considerando sólo el indicio retardado IDD, porque la confiabilidad del IDI no se pudo
20 validar.

En lo que sigue se describirán los respectivos clasificadores 150 y 154 con más detalle de acuerdo con una realización del invento.

Tratando primero el clasificador de largo plazo 154, se nota que se extrae un respectivo conjunto de rasgos distintivos de igual manera desde cada sub-cuadro
25 de 256 muestras. El primer rasgo distintivo es el coeficiente cepstral de predicción lineal perceptiva (PLPCC) que está descrito por H. Hermansky, "Análisis de predicción lineal perceptiva (plp) de la voz" ("Perceptive linear prediction (plp) analysis of speech"), Journal of the Acoustical Society of America, vol. 87, no. 4, pp. 1738 - 1752, 1990 and H. Hermansky, et al., "Análisis predictivo lineal de la voz
30 basado en la percepción" ("Perceptually based linear predictive analysis of speech"), ICASSP 5.509 - 512, 1985. Los PLPCC son eficientes para la clasificación de personas que hablan usando la estimación de la percepción aditiva humana. Este rasgo distintivo se puede usar para discriminar voz y música y, de hecho, permite distinguir tanto los formantes de la voz como la modulación silábica de 4 Hz de la voz
35 observando la variación de los rasgos distintivos en el tiempo.

Sin embargo, para ser más robustos los PLPCC se combinan con otro rasgo distintivo que es capaz de capturar la información de frecuencia fundamental, que es otro rasgo distintivo importante de la voz y puede ser crítica para la codificación. De hecho, la codificación de voz se apoya en la suposición que una
5 señal de entrada es una señal pseudo-monoperiódico. Los esquemas de codificación de voz son eficientes para tales señales. Por otro lado, el rasgo distintivo de frecuencia fundamental de la voz perjudica mucho la eficiencia de codificación de aparatos de codificación de música. La fluctuación suave de retardo de frecuencia fundamental, dado por el vibrato natural de la voz, hace que la representación de
10 frecuencia en los aparatos de codificación de música no puede compactar eficientemente la energía que se requiere para obtener una alta eficiencia de codificación.

Se pueden determinar los siguientes rasgos distintivos de frecuencia fundamental:

15 Relación de energía de pulsos glotales:

Este rasgo distintivo de la relación de energía entre los pulsos glotales y la señal residual LPC. Los pulsos glotales se extraen desde la señal residual LPC usando un algoritmo de selección de picos. Normalmente, la señal residual LPC de un segmento sonoro muestra una gran estructura de tipo pulso que proviene de la
20 vibración glotal. Este rasgo distintivo es alto durante segmentos o sonoros.

Predicción de ganancia de largo plazo:

Normalmente se calcula la ganancia en los aparatos de codificación de voz (ve por ejemplo "Codificador-decodificador (AMR-WB+) adaptivo, extendido de banda ancha de múltiples velocidades" ("Extended Adaptive Multi-Rate – Wideband
25 (AMR-WB+) codec"), 3GPP TS 26.290 V6.3.0, 2005-06, Especificación técnica) durante la predicción de largo plazo. Este rasgo distintivo mide la periodicidad de la señal y se basa en la estimación de retardo de frecuencia fundamental.

Fluctuación de retardo de frecuencia fundamental:

Este rasgo distintivo determina las diferencias de la estimación de retardo
30 de frecuencia fundamental cuando se compara con el último sub-cuadro. Para la voz

sonora este rasgo distintivo debe ser bajo pero no en cero y debe evolucionar suavemente.

Una vez que el clasificador de largo plazo ha extraído el conjunto necesario de rasgos distintivos, se utiliza un clasificador estadístico sobre estos rasgos distintivos extraídos. El clasificador se ha entrenado primero extrayendo los rasgos distintivos de un conjunto de entrenamiento de voz y un conjunto de entrenamiento de música. Los rasgos distintivos extraídos se normalizan a un valor promedio de 0 y una variación de 1 sobre ambos conjuntos de entrenamiento. Para cada conjunto de entrenamiento, se recogen los rasgos distintivos extraídos y normalizados dentro de una ventana de clasificador de largo plazo y se modelan con un modelo de mezcla gaussiana (GMM) que usa cinco gaussianas. Al final de cada secuencia de entrenamiento, se obtienen y se guardan un conjunto de parámetros de normalización y dos conjuntos de parámetros GMM.

Para cada cuadro a ser clasificado, primero se extraen y se normalizan los rasgos distintivos con los parámetros de normalización. Se calculan la probabilidad máxima para voz (l_{ld_voz}) y la probabilidad máxima para música ($l_{ld_música}$) para los rasgos distintivos extraídos y normalizados utilizando el GMM de la categoría de voz y el GMM de la categoría de música, respectivamente. El indicio de decisión retardada IDD se calcula entonces de la siguiente manera:

$$IDD = (l_{ld_voz} - l_{ld_música}) / (abs(l_{ld_música}) + abs(l_{ld_voz}))$$

El IDD está limitado entre los valores -1 y 1 , y es positivo cuando la probabilidad máxima para voz es más alta que probabilidad máxima para música, $l_{ld_voz} > l_{ld_música}$.

El clasificador de corto plazo utiliza como un rasgo distintivo de corto plazo los PLPCC. Diferente al clasificador de largo plazo, este rasgo distintivo sólo se analiza en la ventana 128. Se explotan las estadísticas de este rasgo distintivo en este tiempo corto mediante un modelo de mezcla gaussiana (GMM) que usa cinco gaussianas. Dos modelos se entrenan, uno para música y el otro para voz. Vale la pena mencionar, que los dos modelos son diferentes de los modelos que se obtienen para el clasificador de largo plazo. Para cada cuadro a ser clasificado, primero se extraen los PLPCC y se calculan la probabilidad máxima para voz (l_{ld_voz}) y la

probabilidad máxima para música (lld_música) para el uso del GMM de la categoría de voz y del GMM de la categoría de música, respectivamente. Entonces se calcula el indicio de decisión instantáneo IDI de la siguiente manera:

$$\text{IDI} = (\text{lld_voz} - \text{lld_música}) / (\text{abs}(\text{lld_música}) + \text{abs}(\text{lld_voz}))$$

5 El IDI está limitado entre los valores -1 y 1.

De ese modo, el clasificador de corto plazo 150 genera los resultado de clasificación de corto plazo de la señal sobre la base del rasgo distintivo del "coeficiente cepstral de predicción lineal perceptiva" (PLPCC), y el clasificador de largo plazo 154 genera le resultado de clasificación de largo plazo de la señal sobre
10 la base del mismo rasgo distintivo "coeficiente cepstral de predicción lineal perceptiva" (PLPCC) y el rasgo distintivo o los rasgos distintivos adicionales mencionados más arriba, por ejemplo, el rasgo distintivo o los rasgos distintivos de las características de frecuencia fundamental. Más aún, el clasificador de largo plazo puede explotar distintas características del rasgo distintivo compartido, es decir los
15 PLPCC, si tienen acceso a una ventana de observación más larga. De ese modo, al combinar los resultados de clasificación de corto plazo y de largo plazo, se consideran suficientemente los rasgos distintivos de corto plazo para la clasificación, es decir, sus propiedades se explotan suficientemente.

Más abajo se describirá con más detalle otra realización para los
20 respectivos clasificadores 150 y 154.

Los rasgos distintivos de corto plazo analizados por el clasificador de corto plazo de acuerdo con esta realización corresponden principalmente a los coeficientes cepstrales de predicción lineal perceptiva (PLPCC) mencionados más arriba. Tanto los PLPCC como los MFCC (ver más arriba) se utilizan ampliamente en el
25 reconocimiento de voz y de la persona que habla. Se mantienen en uso los PLPCC porque comparten una gran parte de su funcionalidad de la predicción lineal (LP) que se utiliza en la mayoría de los aparatos de codificación de voz modernos y si ya están implementados también en un aparato de codificación de música conmutado. Los PLPCC pueden extraer le estructura de formato de la voz como hace también el
30 LP pero tomando en cuenta las consideraciones perceptivas. Los PLPCC son más independientes de la persona que habla y por lo tanto más importante con respecto a

la información lingüística. Se utiliza un grado de 16 en una señal de entrada tomado con una tasa de muestreo de 16 kHz.

Aparte de los PLPCC, se calcula una intensidad de sonorización como un rasgo distintivo de corto plazo. No se considera intensidad de sonorización como realmente discriminadora por si misma, pero es favorable en asociación con los PLPCC en la dimensión del rasgo distintivo. La intensidad de sonorización permite dibujar en la dimensión de rasgos distintivos por lo menos dos agrupamientos que corresponden respectivamente a las pronunciaciones sonoras y no sonoras de la voz. Se basa sobre un cálculo de calidad de sistema utilizando distintos parámetros, en particular, un contador de cruces por cero (zc), la distorsión espectral (tilt), la estabilidad de la frecuencia fundamental (ps) y la correlación normalizada de frecuencia fundamental (nc). Todos los cuatro parámetros están normalizados entre 0 y 1 en una manera que 0 corresponde a una señal típicamente no sonora y 1 corresponde a una señal típicamente sonora. En esta realización, la intensidad de sonorización está inspirada desde los criterios de clasificación de voz usados en el aparato de codificación VMR-WB descrito por Milan Jelinek y Redwan Salami, "Avances en la codificación de voz de banda ancha en el estándar vmr-wb" ("Wideband speech coding advances in vmr-wb standard"), IEEE Trans. on Audio, Procesamiento de voz e idioma (Speech and Language Processing), vol. 15, no. 4, pp. 1167 - 1179, Mayo de 2007. Se basa sobre un rastreador de frecuencia fundamental evolucionado sobre la base de una auto-correlación. Para el cuadro con el índice k, la intensidad de sonorización $u(k)$ tienen la siguiente fórmula:

$$v(k) = \frac{1}{5} (2 * nc(k) + 2 * ps(k) + tilt(k) + zc(k))$$

La capacidad discriminadora de los rasgos distintivos de corto plazo es evaluada por los modelos de mezcla gaussiana (GMMS) como un clasificador. Se aplican dos GMM, uno para la categoría de voz y el otro para la categoría de música. La cantidad de mezclas se hace a fin de evaluar el efecto sobre el rendimiento. Se calcula una decisión para cada segmento de cuatro cuadros sucesivos. El retardo total es entonces de 64 ms, lo cual es apropiado para una codificación de audio conmutada. Se puede observar, que el rendimiento mejora con la cantidad de mezclas. La brecha entre 1 GMM y 5 GMM es particularmente importante y se puede

explicar por el hecho de que la representación de formante de la voz es demasiado compleja para suficientemente definida por sólo una gaussiana.

	1 GMM	5 GMM	10 GMM	20 GMM
Voz	95.33	96.52	97.02	97.60
Música	92.17	91.97	91.61	91.77
Promedio	93.75	94.25	94.31	94.68

Tabla 1: Precisión de clasificación de rasgos distintivos de corto plazo en %

5

Tratando ahora el clasificador de largo plazo 154, se nota que muchos trabajos , por ejemplo, M. J. Carey, et. al. "Una comparación de los rasgos distintivos para la discriminación de voz y música" ("A comparison of features for speech and music discrimination"), Proc. IEEE Int. Conf. Acoustics, Procesamiento de voz y señal (Speech and Signal Processing), ICASSP, vol. 12, pp. 149 hasta 152, Marzo de 1999, consideran que las variaciones de los rasgos distintivos estadísticos sean más discriminatorias que los rasgos distintivos mismos. Como una regla general indicativa, se puede considerar que la música es más estacionaria y presenta generalmente menos variaciones. Al contrario, la voz se puede distinguir fácilmente por su modulación de energía de 4 Hz notable ya que la señal cambia periódicamente entre los segmentos sonoros y no sonoros. Más aún, la sucesión de distintos fonemas hace que los rasgos distintivos de voz estén menos constantes. En esta realización, se consideran dos rasgos distintivos de largo plazo, uno en base sobre el cálculo de una variación y el otro sobre la base de un conocimiento previo del contorno de frecuencia fundamental de la voz. Se adaptan los rasgos distintivos de largo plazo a la SMD (discriminación de voz y música) de poco retardo.

La variación móvil de los PLPCC consiste en el cálculo de la variación para cada conjunto de PLPCC a lo largo de una ventana de análisis superpuesta que cubre varios cuadros a fin de resaltar el último cuadro. Para limitar los tiempos de espera introducidos, la ventana de análisis es asimétrica y considera sólo el cuadro actual y la historia pasada. En una primera etapa, se calcula el promedio móvil $ma_m(k)$ de los PLPCC sobre por lo menos N cuadros como se describe en la siguiente fórmula:

$$ma_m(k) = \sum_{i=0}^{N-1} PLPC_m(k-i) \cdot w(i)$$

Donde $PLPP_m(k)$ es el coeficiente cepstral m sobre un total de M coeficientes provenientes del cuadro k . La variación móvil $mv_m(k)$ se define entonces como:

$$5 \quad mv_m(k) = \sum_{i=0}^{N-1} (PLPC_m(k-i) - ma_m(k))^2 \cdot w(i)$$

Donde w es una ventana del largo N , que en esta realización es una pendiente de rampa definida como:

$$w(i) = (N-i) / N \cdot (N+1) / 2$$

Finalmente se promedia la variación móvil sobre la dimensión cepstral:

$$10 \quad mv(k) = \frac{1}{M} \sum_{m=0}^M mv_m(k)$$

La frecuencia fundamental de la voz tiene unas propiedades notables y una parte de las mismas se puede observar sólo en ventanas de análisis largas. De hecho, la frecuencia fundamental fluctúa suavemente durante los segmentos sonoros pero es raras veces constante. Al contrario, la música presenta muy frecuentemente una frecuencia fundamental constante durante la duración completa de una nota y un cambio abrupto durante los componentes transitorios. Los rasgos distintivos de largo plazo abarcan esta característica observando el contorno de la frecuencia fundamental sobre un segmento de tiempo largo. Un parámetro de contorno de frecuencia fundamental $pc(k)$ se define como:

$$20 \quad pc(k) = \begin{cases} 0 & \text{si } |p(k) - p(k-1)| < 1 \\ 0,5 & \text{si } 1 \leq |p(k) - p(k-1)| < 2 \\ 1 & \text{si } 2 \leq |p(k) - p(k-1)| < 20 \\ 0,5 & \text{si } 20 \leq |p(k) - p(k-1)| < 25 \\ 0 & \text{si es de otra manera} \end{cases}$$

donde $p(k)$ es el retardo de la frecuencia fundamental calculado para el índice de cuadro k sobre la señal residual LP tomado con una tasa de muestreo de 16 KHz. A partir del parámetro de contorno de frecuencia fundamental se calcula una calidad de voz $sm(k)$ en una manera que espera que la voz demuestre un retardo de frecuencia fundamental de una suave fluctuación durante los segmentos sonoros y

una fuerte distorsión espectral hacia las frecuencias altas durante los segmentos no sonoros:

$$sm(k) = \begin{cases} nc(k) \cdot pc(k) & \text{si } v(k) \geq 0.5 \\ (1 - nc(k)) \cdot (1 - tilt(k)) & \text{si es de otra manera} \end{cases}$$

donde $nc(k)$, $tilt(k)$ y $v(k)$ están definidas como se indicó más arriba (ver el clasificador de corto plazo). Luego se pondera la calidad de voz con la ventana w definida más arriba y se integra sobre los últimos N cuadros:

$$ams(k) = \sum_{i=0}^N m(k-1)w(i)$$

El contorno de la frecuencia fundamental también es una indicación importante que una señal sea apropiada para una codificación de voz o de música. De hecho, los aparatos de codificación trabajan principalmente en el dominio del tiempo y suponen que la se es harmónica y cuasiestacionario den segmentos de tiempos cortos de aproximadamente 5 ms. De esta manera ellos pueden modelar eficientemente la fluctuación natural de la frecuencia fundamental de la voz. Al contrario, la misma fluctuación daña la eficiencia de codificadores de audio comunes que explotan las transformaciones lineales sobre ventanas de análisis largas. Luego se distribuye la energía principal de la señal sobre varios coeficientes transformados.

Tal como para los rasgos distintivos de corto plazo, también se evalúan los rasgos distintivos de largo plazo utilizando un clasificador estadístico, y de ese modo se obtienen un resultado de clasificación de largo plazo (IDD). Los dos rasgos distintivos se calculan utilizando $N = 25$ cuadros, por ejemplo, considerando 400 ms de historia pasada de la señal. Un análisis discriminante lineal (ADL) se aplica primero antes de utilizar 3 GMM en el espacio unidimensional reducido. La Tabla 2 muestra el rendimiento medido sobre los conjuntos de entrenamiento y de prueba para la clasificación de segmentos de cuatro cuadros sucesivos.

25

	Conjunto de entrenamiento	Conjunto de prueba
Voz	97,99	97,84
Música	95,93	95,44
Promedio	96,96	96,64

Tabla 2: Precisión de clasificación de rasgos distintivos de largo plazo en %

Los sistemas de clasificadores combinados de acuerdo con algunas realizaciones del invento combinan apropiadamente los rasgos distintivos de corto plazo y de largo plazo en una manera que aportan su propia contribución específica a la decisión final. Para este propósito se puede utilizar la fase de decisión final de histéresis, tal como se describe más arriba, donde el efecto de memoria está impulsado por el IDD o el indicio de discriminación de largo plazo (IDL P) mientras la entrada instantánea proviene del IDI o del indicio de discriminación de corto plazo (IDCP). Los dos indicios son las salidas de los clasificadores de largo plazo y de corto plazo como se muestra en la Figura 1. Se toma la decisión sobre la base del IDI pero se estabiliza la misma con el IDD que controla dinámicamente los umbrales que disparan un cambio de estado.

El clasificador de largo plazo 154 utiliza ambos, los rasgos distintivos de largo plazo y de corto plazo previamente definido por un ADSL seguido por 3 GMM. El IDD es igual a la relación logarítmica entre la probabilidad del clasificador de largo plazo para la categoría de voz y la para la categoría de música calculados sobre los últimos $4 \times K$ cuadros. El número de cuadros, que se toman en cuenta, puede variar con el parámetro K a fin de agregar más o menos efecto de memoria a la decisión final. Al contrario, el clasificador de corto plazo utiliza sólo los rasgos distintivos de corto plazo con 5 GMM que muestran un buen compromiso entre el rendimiento y la complejidad. El IDI es igual a la relación logarítmica entre la probabilidad del clasificador de largo plazo para la categoría de voz y la para la categoría de música calculados sólo sobre los últimos 4 cuadros.

A fin de evaluar el enfoque inventivo, en particular, para la codificación de audio conmutada, se evaluaron tres distintos tipos de rendimiento. Una primera medición de rendimiento es el rendimiento convencional de voz versus música (VvM). Se evalúa sobre un conjunto grande de productos de música y de voz. Una segunda medición de rendimiento se hace sobre un largo producto único que tienen segmentos de voz y de música que alternan cada 3 segundos. La precisión de discriminación se llama entonces el rendimiento de voz antes/después de música (VadM) y refleja principalmente la reactividad del sistema. Finalmente, la estabilidad de la decisión se evalúa llevando a cabo la clasificación sobre un conjunto grande de

productos de voz sobre música. La mezcla entre voz y música se hace en distintos niveles de un producto a otro. Entonces se obtiene el rendimiento de voz sobre música (VsM) calculando la relación de la cantidad de conmutaciones de categoría que ocurrió durante la cantidad total de cuadros.

5 Se utilizan los clasificadores de largo plazo y los clasificadores de corto plazo como referencias para evaluar los enfoques de clasificadores individuales convencionales. El clasificador de corto plazo muestra una buena reactividad, mientras tiene una estabilidad más baja y una capacidad de discriminación total más baja. Por otro lado el clasificador de largo plazo, especialmente cuando aumenta el
10 número de cuadros por $4 \times K$, puede alcanzar una mejor estabilidad y un mejor comportamiento de discriminación a costo de la reactividad para la decisión. En comparación con los enfoques convencionales recién arriba mencionados, los rendimientos de los clasificadores combinados de acuerdo con el invento tienen varias ventajas. Una ventaja es que mantienen un rendimiento de una buena voz
15 pura contra la discriminación de música mientras conserva la reactividad del sistema. Otra ventaja es la buena solución de compromiso entre la reactividad y la estabilidad.

 En lo que sigue, se hace referencia a las Figuras 4 y 5 que muestran esquemas de codificación y decodificación ejemplares, los cuales incluyen una fase
20 de discriminación o decisión que funciona de acuerdo con algunas realizaciones del invento.

 De acuerdo con el esquema de codificación ejemplar que se muestra en la Figura 4, se hace entrar una señal mono, una señal estéreo y una señal de múltiples canales en la fase de pre-procesamiento común 200.

25 La fase de pre-procesamiento común 200 puede tener una funcionalidad de estéreo unida, una funcionalidad "surround" y/o una funcionalidad de extensión de ancho de banda. En la salida de la fase 200 hay un canal mono, un canal estéreo o unos canales múltiples que forman la entrada para uno o más conmutadores 202. El conmutador 202 puede estar provisto para cada salida de la fase 200, cuando la
30 fase 200 tiene dos o más salidas, es decir, cuando la fase 200 emite una señal estéreo una señal de múltiples canales. A modo de ejemplo, el primer canal de una

señal estéreo puede ser un canal de voz y el segundo canal de la señal estéreo puede ser un canal de música. En este caso, la decisión en una fase de decisión 204 puede ser diferente entre los dos canales para el mismo instante de tiempo.

Se controla el conmutador 202 mediante la fase de decisión 204 la fase de
5 decisión comprende un discriminador de acuerdo con algunas realizaciones del invento y recibe, como una entrada, una señal de entrada hacia la fase 200 o una salida de señal desde la fase 200. De manera alternativa, la fase de decisión 204 puede recibir también una información lateral, que está incluida en la señal mono, la
10 señal estéreo y la señal de múltiples canales o está por lo menos asociada con tal señal donde existe la información la cual se generó, por ejemplo, cuando originalmente se produjo la señal mono, la señal estéreo y la señal de múltiples canales.

En una realización, la fase de decisión no controla la fase de pre-
procesamiento 200 y la flecha entre las fases 204 y 200 no existe. En otra
15 realización, el procesamiento en la fase 200 está controlado en cierto grado por la fase de decisión 204 a fin de fijar uno o más parámetros en la fase 200 sobre la base de la decisión. Sin embargo, esto no influye al algoritmo general en la fase 200 de modo que la funcionalidad principal en la fase 200 es activa con independencia de la decisión en la fase de decisión 204.

20 La fase de decisión 204 actúa el conmutador 202 a fin de alimentar la salida de la fase de pre-procesamiento común o sea en una porción de codificación de frecuencia 206 ilustrada en una rama superior de la Figura 4 o en una porción de codificación del dominio del LPC 208 ilustrada en una rama inferior de la Figura 4.

En una realización, el conmutador 202 conmuta entre las dos ramas de
25 codificación 206, 208. En otra realización, puede haber otras ramas de codificación tales como una tercera rama de codificación o aún una cuarta rama de codificación o aún más ramas de codificación. En una realización con tres ramas de codificación, la tercera rama de codificación puede ser similar a la segunda rama de codificación, pero incluye un codificador de excitación diferente del codificador de excitación 210
30 en la segunda rama de codificación 208. En tal realización la segunda rama de codificación comprende la fase LPC 212 y un codificador de excitación 210 basado

en un libro de códigos tal como en el ACELP, y la tercera rama de codificación comprende una fase LPC y un codificador de excitación que funciona con una representación espectral de la señal de salida de la fase LPC.

La rama de codificación del dominio de la frecuencia comprende un bloque
5 de conversión 214 que funciona para convertir la señal de salida de la fase de pre-
procesamiento común en un dominio espectral. El bloque de conversión puede incluir
un algoritmo MDCT un QMF, un algoritmo FFT, un análisis de tren de ondas o un
banco de filtros tales como un banco de filtros con una tasa de muestreo crítico que
tiene una cierta cantidad de canales de banco de filtros donde las señales de sub
10 banda en este banco de filtros pueden ser señales de valores reales o señales de
valores complejos. La salida del bloque de conversión 214 está codificada utilizando
un codificador de audio espectral 216 que puede incluir bloques de procesamiento
como se conocen desde el esquema de codificación tipo AAC.

La rama inferior de codificación 208 comprende un analizador de modelo
15 de fuente tal como el LPC 212 que emite dos tipos de señales. Una señal es una
señal de información LPC que se utiliza para controlar la característica de filtro de un
filtro de síntesis LPC. Esta información LPC se transmite hacia un decodificador. La
otra señal de salida de fase LPC 212 es una señal de excitación o una señal de
dominio del LPC que es una entrada en un codificador de excitación 210. El
20 codificador de excitación 210 puede ser cualquier codificador de modelo de filtro de
fuente tal como un codificador CELP, un codificador ACELP o cualquier otro
codificador que es capaz de procesar una señal de dominio del LPC.

Otra implementación de un codificador de excitación puede ser una
codificación de transformación de una señal de excitación. En tal realización la señal
25 de excitación no se codifica utilizando un mecanismo de libro de códigos ACELP,
sino se convierte la señal de excitación en una representación espectral y los valores
de la representación espectral tales como las señales de sub banda en el caso de un
banco de filtros o los coeficientes de frecuencia en el caso de una transformación tal
como una FFT se codifican para obtener una compresión de datos. Una
30 implementación de este tipo de codificador de excitaciones el modo de codificación
TCX que se conoce del AMR-WB+.

La decisión en la fase de decisión 204 puede ser adaptable a la señal de modo que la fase de decisión 204 lleva a cabo una discriminación de música y voz y controla el conmutador 202 de tal manera, que las señales de música forman la entrada para la rama superior 206 y las señales de voz forman la entrada para la rama inferior 208. En una realización, la fase de decisión 204 alimenta su información de decisión a una transmisión de bits de salida de modo que un decodificador puede utilizar esta información de decisión a fin de llevar a cabo las funciones correctas de decodificación.

Tal decodificador se ilustra en la Figura 5. Después de la transmisión se entra la señal emitida por el codificador de audio espectral 216 en un decodificador de audio espectral 218. La salida del decodificador de audio espectral 218 es la entrada para un convertidor de dominio del tiempo 220. La salida del codificador de excitación 210 de la Figura 4 es la entrada para un decodificador de excitación 222 que emite una señal de dominio del LPC. La señal de dominio del LPC es la entrada para una fase de síntesis de LPC 224 que recibe como otra entrada más la información LPC generada por la correspondiente fase de análisis LPC 212. La salida del convertidor de dominio del tiempo 220 y/o la salida de la fase de síntesis de LPC 224 son la entrada para un conmutador 226. El conmutador 226 está controlado por una señal de control de conmutador que, por ejemplo, fue generada por la fase de decisión 204, o el cual fue suministrado desde afuera tal como por un creador de la señal original mono, estéreo o de múltiples canales.

La salida del conmutador 226 es una señal mono completa que se ingresa subsiguientemente en una fase de post-procesamiento común 228 que puede llevar a cabo un procesamiento estéreo unido a un procesamiento de extensión de ancho de banda etc. De manera alternativa, la salida del conmutador también puede ser una señal estéreo o una señal de múltiples canales. Es una señal estéreo cuando el pre-procesamiento incluye una reducción de canales a dos canales. Inclusive, puede ser una señal de múltiples canales, cuando se lleva a cabo una reducción de canales a tres canales o ninguna reducción de canales, sino una replicación de banda espectral.

Dependiendo de la funcionalidad específica de la fase de post-procesamiento común, se emite una señal mono, una señal estéreo o una señal de múltiples canales, que tiene, cuando la fase de post-procesamiento común 228 lleva a cabo una operación de extensión de ancho de banda, un ancho de banda mayor
5 que la señal que entró en el bloque 228.

En una realización el conmutador 226 conmuta entre las dos ramas de decodificación 218, 220 y 222, 224. En otra realización, puede haber ramas de decodificación adicionales tales como una tercera rama de decodificación o aún una cuarta rama de decodificación o aún más ramas de decodificación. En una
10 realización con tres ramas de decodificación, la tercera rama de decodificación puede ser similar a la segunda rama de decodificación, pero incluye un decodificador de excitación que es diferente al decodificador de excitación 222 en la segunda rama de decodificación 222, 224. En una semejante realización, la segunda rama comprende la fase de LPC 224 y un decodificador de excitación basado en un libro
15 de códigos tal como en el ACELP, y la tercera rama comprende una fase de LPC y un decodificador de excitación que funciona sobre una representación espectral de la señal de salida de la fase de LPC 224.

En otra realización, la fase pre-procesamiento común comprende un bloque de estéreo surround/en conjunto que genera, como una salida, unos
20 parámetros de estéreo en conjunto y una señal mono de salida, que se genera mezclando hacia abajo la señal de entrada que es una señal que tiene dos o más canales. En general, la señal en la salida del bloque también puede ser una señal que tiene más canales, pero debido a la operación de mezclar hacia abajo, el número de canales en la salida del bloque será menor que el número de canales en
25 la entrada al bloque. En esta realización, la rama de codificación de frecuencia comprende una fase de conversión espectral y una fase conectada subsiguientemente de cuantificación y codificación. La fase de cuantificación y codificación puede incluir cualquiera de las funcionalidades que se conocen de los codificadores modernos del dominio de la frecuencia tales como los codificadores de
30 tipo AAC. Además se puede controlar la operación de cuantificación en la fase de cuantificación y codificación mediante un módulo psicoacústico que genera

información psicoacústica tal como un umbral de enmascaramiento psicoacústico sobre la frecuencia donde esta información se ingresa a esta fase. Preferiblemente, la conversión espectral se hace usando una función de MDCT que, aún más preferiblemente, es la función de MDCT deformada en el tiempo, donde la intensidad
5 o en general, la intensidad de deformación, puede ser controlada entre cero y una alta intensidad de deformación. En una intensidad de deformación cero, la función de MDCT es una función de MDCT sin complicaciones que se conoce en la materia. El codificador del dominio de LPC puede incluir un núcleo ACELP que calcula una ganancia de la frecuencia fundamental, un retraso de la frecuencia fundamental y/o
10 una información de libro de códigos tal como un índice de libro de códigos y una ganancia de códigos.

A pesar de que algunas figuras muestran diagramas de bloque de un aparato, se hace notar que estas figuras al mismo tiempo muestran un método, en el cual las funcionalidades de los bloques corresponden a las etapas de un método.

15 Algunas realizaciones del invento se describieron más arriba sobre la base de una señal de entrada de audio que comprende distintos segmentos o cuadros, se asocian los distintos segmentos o cuadros con información de voz o información de música. El invento no se limita a tales realizaciones, más bien el enfoque para clasificar distintos segmentos de una señal comprende segmentos de por lo menos
20 un primer tipo y un segundo tipo se puede aplicar también a señales de audio que comprenden tres o más distintos tipos de segmentos, cada uno de los cuales se desea codificar con distintos esquemas de codificación. Ejemplos para tales tipos de segmentos son:

– Segmentos estacionarios y no estacionarios pueden ser útiles para el
25 uso de distintos bancos de filtros, ventanas o adaptaciones de codificación. Por ejemplo, un componente transitorio se debe codificar con un banco de filtros de una fina resolución de tiempo, mientras una señal sinusoidal pura se debe codificar con un banco de filtros de una fina resolución de frecuencia.

– Sonoros / no sonoros: Los segmentos sonoros están bien tratados con
30 un codificador de voz como el CELP, pero para los segmentos no sonoros se gastan demasiados bits. La codificación paramétrica será más eficiente.

- Silencio / activo: Silencio se puede codificar con menos bits que un segmento activo.

Harmónico / no harmónico: Será favorable utilizar para la codificación de segmentos harmónicos una predicción lineal en el dominio de la frecuencia.

5 Además el invento no está limitado al campo de las técnicas de audio, más bien el enfoque descripto más arriba para clasificar una señal se puede aplicar a otros tipos de señales tales como señales de vídeo o señales de datos, en lo cual estas respectivas señales incluyen segmentos de distintos tipos que requieren un procesamiento diferente, como por ejemplo:

10 Se puede adaptar el presente invento para todas las aplicaciones en tiempo real que requieren una segmentación de una señal de tiempo. Por ejemplo, el reconocimiento de una cara desde una cámara de video de vigilancia puede ser basado sobre un clasificador que determina para cada píxel de un cuadro (aquí un cuadro corresponde a una imagen tomada en un momento e tiempo n) si pertenece a

15 la cara de una persona o no. La clasificación (es decir, la segmentación de la cara) se debe hacer para cada cuadro individual de la transmisión de video. Sin embargo, utilizando el presente invento, la segmentación del cuadro actual puede tomar en cuenta los sucesivos cuadros pasados para obtener una mejor precisión de la segmentación tomando la ventaja que las sucesivas imágenes están fuertemente

20 correladas. Entonces se pueden aplicar dos clasificadores. Uno considera sólo el cuadro actual y otro que considera un conjunto de cuadros incluyendo el cuadro actual y los cuadros pasados. Este último clasificador puede integrar un conjunto de cuadros y determinar la probabilidad de la región para la posición de la cara. La decisión de clasificador que se hace sólo sobre el cuadro actual se compara

25 entonces con las regiones de probabilidad. Entonces se valida la decisión o se modifica.

 Algunas realizaciones del invento utilizan un conmutador para conmutar entre ramas de modo que sólo una rama reciba una señal a ser procesada y que la otra rama no reciba la señal. Sin embargo, en una realización alternativa el

30 conmutador puede también estar dispuesto detrás de las fases de procesamiento o ramas, por ejemplo, el codificador de audio y el codificador de voz, de modo que

ambas ramas procesen la misma señal en paralelo. Se selecciona la salida de señal de una de estas ramas para la salida, por ejemplo, a ser escrito en una transmisión de bits de salida.

5 Mientras algunas realizaciones del invento se describieron sobre la base de señales digitales, en las cuales se determinaron los segmentos mediante una cantidad predeterminada de muestras obtenidas con una tasa de muestreo específica, el invento no está limitado a esas señales, más bien, se puede aplicar también a señales análogas en las cuales se determinaría el segmento mediante un rango de frecuencia específico o un periodo de tiempo específico de la señal
10 análoga. Además, algunas realizaciones del invento se describieron en combinación con codificadores que incluyen un discriminador. Se hace notar, que básicamente, el enfoque de acuerdo con las realizaciones del invento para clasificar señales se puede aplicar también a decodificadores que reciben una señal codificada, para lo cual se pueden clasificar distintos esquemas de codificación, permitiendo, de ese
15 modo, que se suministre la señal codificada a un decodificador apropiado.

Dependiendo de ciertos requerimientos de implementación de los métodos inventivos, los métodos inventivos se pueden implementar mediante hardware o software. Se puede llevar a cabo la implementación utilizando un medio de almacenamiento digital, en particular, un disco, un DVD o un CD, en el cual están
20 guardadas unas señales de control electrónicamente legibles, las cuales cooperan con sistemas de computación programables de modo que se llevan a cabo los métodos inventivos. Por lo tanto, el presente invento es, en general, un producto de programa de computación con un código de programa guardado sobre un portador legible por máquinas, este código de programa se usa para llevar a cabo los métodos
25 inventivos, cuando se ejecuta el producto de programa de computación en una computadora. En otras palabras, los métodos inventivos son, por lo tanto, un programa de computación que tiene un código de programa para llevar a cabo por lo menos uno de los métodos inventivos cuando se ejecuta el programa de computación en una computadora.

30 Las realizaciones que se describieron más arriba son puramente ilustrativas para los principios del presente invento. Se entiende que las

modificaciones y variaciones posibles de las disposiciones y de los detalles descriptos en la presente serán evidentes para los expertos en la materia. Por lo tanto, es la intención que el invento esté limitado sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados por la descripción y la explicación de las realizaciones en la presente.

En las realizaciones, la señal se describe, que comprenda una pluralidad de cuadros, en lo cual se evalúa el cuadro actual con respecto a una decisión de conmutación. Se hace notar que el cuadro actual de la señal que se está evaluando con respecto a una decisión de conmutación, puede ser un cuadro, sin embargo, el invento no está limitado a tales realizaciones. Más bien, un segmento de la señal también puede comprender una pluralidad de cuadros, es decir dos o más cuadros.

Además, en las realizaciones descriptas más arriba, tanto el clasificador de corto plazo como el clasificador de largo plazo utilizan el mismo rasgo distintivo o los mismos rasgos distintivos. Este enfoque se puede utilizar por distintas razones como le necesidad de calcular los rasgos distintivos de corto plazo sólo una vez y explotar el resultado por ambos clasificadores en distintas maneras que reduce la complejidad del sistema, por ejemplo, en la manera que se calcule el rasgo distintivo de corto plazo en uno de los clasificadores de corto plazo y de largo plazo y se suministra el resultado al otro clasificador. También la comparación entre los resultados de los clasificadores de corto plazo y de largo plazo pueden ser más importantes, ya que se puede deducir más fácilmente la contribución del cuadro actual en el resultado de clasificación de largo plazo comparándolo con el resultado de clasificación de corto plazo, debido a que los dos clasificadores comparten rasgos distintivos comunes

Sin embargo, el invento no se restringe a este enfoque y el clasificador de largo plazo no se restringe al uso de la misma característica o de las mismas rasgos distintivo que el clasificador de corto plazo, es decir tanto el clasificador de corto plazo como el clasificador de largo plazo pueden calcular su respectivo rasgo distintivo de corto plazo o sus respectivas rasgos distintivo de corto plazo que son distintos uno de otro.

Mientras las realizaciones descritas más arriba mencionan el uso de los PLPCC como rasgo distintivo de corto plazo, se hace notar que se pueden considerar otras características, por ejemplo la variabilidad de los PLPCC.

REIVINDICACIONES

1. Un método para clasificar distintos segmentos de una señal, en el cual la señal comprende segmentos de por lo menos un primer tipo y un segundo tipo, y
5 el método comprende:
 - clasificar en un corto plazo (150) la señal sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal y entregar un resultado de clasificación de corto plazo (152);
 - clasificar en un largo plazo (154) la señal sobre la base de por lo menos
10 un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídas de la señal y entregar un resultado de clasificación de largo plazo (156); y
 - combinar (158) el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si un segmento de la señal es del primer tipo o del segundo tipo.
- 15 2. El método de acuerdo con la reivindicación 1, en el cual la etapa de combinar comprende proveer la señal de salida sobre la base de una comparación del resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156)
3. El método de acuerdo con cualquiera de las reivindicaciones 1 o 2, en el
20 cual
 - se obtiene la por lo menos un rasgo distintivo de corto plazo analizando el segmento actual de la señal a ser clasificada;
 - se obtiene la por lo menos un rasgo distintivo de largo plazo analizando el segmento actual de la señal a ser clasificada y uno o más segmentos anteriores de la
25 señal; y
 - la señal de salida (160) indica si el segmento actual es del primer tipo o del segundo tipo.
4. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el cual

la por lo menos un rasgo distintivo de corto plazo se obtiene analizando una ventana de análisis (168) que tiene una primera longitud y usando un primer método de análisis; y

5 la por lo menos un rasgo distintivo de largo plazo se obtiene analizando una ventana de análisis (162) que tiene una segunda longitud y usando un segundo método de análisis, siendo la primera longitud más corta que la segunda longitud, y siendo los métodos de análisis primera y segunda diferentes.

10 5. El método de acuerdo con la reivindicación 4, en el cual la primera longitud se extiende a lo largo del segmento actual de la señal, la segunda longitud se extiende a lo largo del segmento actual de la señal y uno o mas segmentos anteriores de la señal, y las longitudes primera y segunda comprenden un periodo adicional (164) que cubre un periodo de análisis.

15 6. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el cual combinar (158) el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) comprende una decisión de histéresis sobre la base de un resultado combinado, en el cual el resultado combinado comprende el resultado de la clasificación de corto plazo (152) y el resultado de la clasificación de largo plazo (156), cada uno ponderado por un factor de ponderación predeterminado.

20 7. El método de acuerdo con cualquiera de las reivindicaciones 1 a 6, en el cual la señal es una señal digital y un segmento de la señal comprende un número predefinido de muestras obtenido a una tasa de muestreo específico.

8. El método de acuerdo con cualquiera de las reivindicaciones 1 a 7, en el cual

25 la señal comprende segmentos de voz y de música;

la por lo menos un rasgo distintivo de corto plazo comprende los parámetros PLPCC; y

la por lo menos un rasgo distintivo de largo plazo comprende información característica de frecuencia fundamental.

30 9. El método de acuerdo con cualquiera de las reivindicaciones 1 a 8, en el cual el rasgo distintivo de corto plazo usados para la clasificación de corto plazo y el

rasgo distintivo de corto plazo usados para la clasificación de largo plazo son las mismas o diferentes.

10. Un método para procesar una señal que comprende segmentos de por lo menos un primer tipo y un segundo tipo, en el cual el método comprende:

5 clasificar (116) un segmento de la señal con el método de acuerdo con cualquiera de las reivindicaciones 1 a 9;

 procesar (102; 206; 106; 208) el segmento de acuerdo con un primer proceso o un segundo proceso, dependiendo de la señal de salida (160) entregada por la etapa de clasificar (116); y

10 emitir el segmento procesado.

11. El método de acuerdo con la reivindicación 10, en el cual la señal comprende segmentos de voz y de música.

 se procesa el segmento con un codificador de voz (102) cuando la señal de salida (160) indica que el segmento es un segmento de voz; y

15 se procesa el segmento con un codificador de música (106) cuando la señal de salida (160) indica que el segmento es un segmento de música.

12. El método de acuerdo con la reivindicación 11, que comprende además:

 combinar (108) el segmento codificado e información de la señal de salida (160) que indica el tipo del segmento.

13. Un programa de computación que, cuando se ejecuta en una computadora, lleva a cabo el método de acuerdo con cualquiera de las reivindicaciones 1 a 12.

14. Un discriminador que comprende:

25 un clasificador de corto plazo (150) configurado para recibir una señal y proveer un resultado de clasificación de corto plazo (152) de la señal sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal, en lo cual la señal comprende segmentos de por lo menos un primer tipo y un segundo tipo;

 un clasificador de largo plazo (154) configurado para recibir una señal y
30 proveer un resultado de clasificación de largo plazo (156) de la señal sobre la base

de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal; y

un circuito de decisión (158) configurado para combinar el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si un segmento de la señal es del primer tipo o del segundo tipo.

15. El discriminador de acuerdo con la reivindicación 14, en el cual el circuito de decisión (158) está configurado para proveer la señal de salida sobre la base de una comparación del resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156).

16. El discriminador de acuerdo con cualquiera de las reivindicaciones 14 o 15, en el cual la señal comprende segmentos de voz y de música, y en el cual la señal de salida (160) indica si un segmento de una señal es un segmento de voz o un segmento de música.

17. Un aparato procesador de señal, que comprende:

una entrada (110) configurada para recibir una señal a ser procesada, en el cual la señal comprende segmentos de por lo menos un primer tipo y un segundo tipo;

un primer elemento procesador (102; 206) configurado para procesar segmentos de un primer tipo;

un segundo elemento procesador (104; 208) configurado para procesar segmentos de un segundo tipo;

un discriminador (116; 204) de acuerdo con cualquiera de las reivindicaciones 14 a 16 acoplado a la entrada;

y un dispositivo conmutador (112; 202) acoplado entre la entrada y los elementos procesadores primero y segundo y configurado para aplicar la señal proveniente de la entrada (110) a uno de los elementos procesadores primero y segundo dependiendo de la señal de salida (160) proveniente del discriminador (116).

18. Un codificador de audio, que comprende:

un aparato procesador de señal de acuerdo con la reivindicación 17,

en el cual la señal es una señal de audio que comprende segmentos de voz y de música, y

en el cual el primer elemento procesador comprende un codificador de voz (102) y el segundo elemento procesador comprende un codificador de música (106).

5

RESUMEN

Para clasificar distintos segmentos de una señal que comprende segmentos de por lo menos un primer tipo y un segundo tipo, por ejemplo, segmentos de voz y de música, se clasifica la señal en un corto plazo (150) sobre la

5 base de por lo menos un rasgo distintivo de corto plazo extraída de la señal y se entrega un resultado de clasificación de corto plazo (152). La señal se clasifica también en un largo plazo (154) sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal y se entrega un resultado de clasificación de largo plazo (156). Se combinan (158) el

10 resultado de la clasificación de corto plazo (152) y el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si un segmento de la señal es del primer tipo o del segundo tipo.

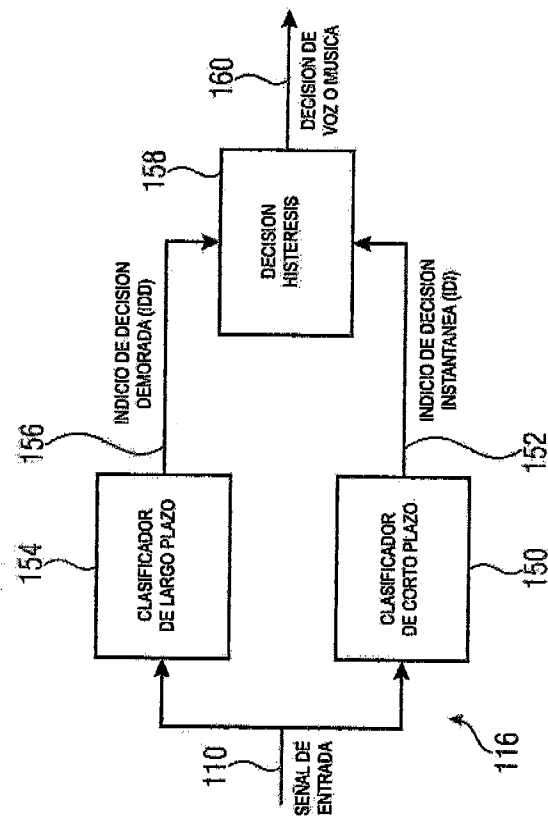


FIG. 1

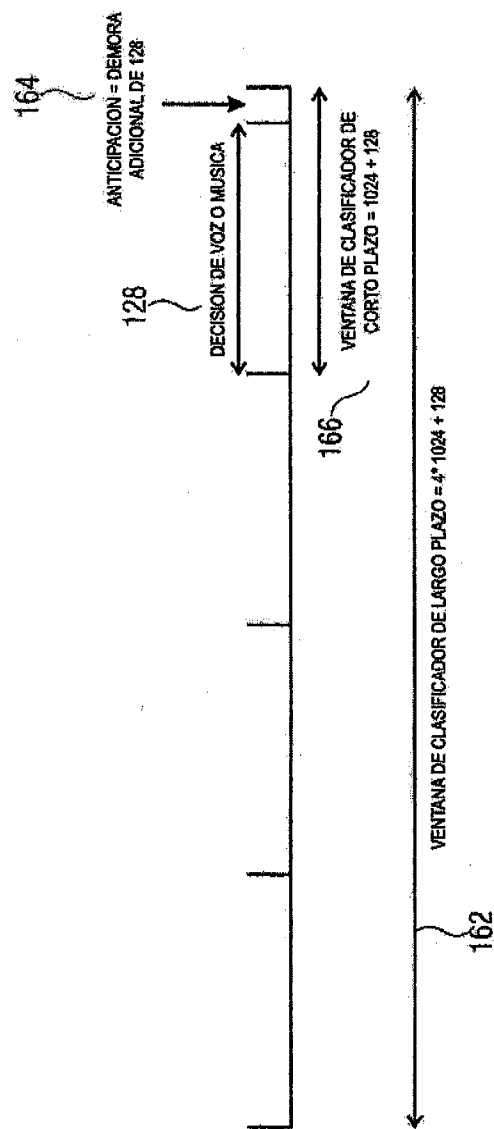


FIG. 2

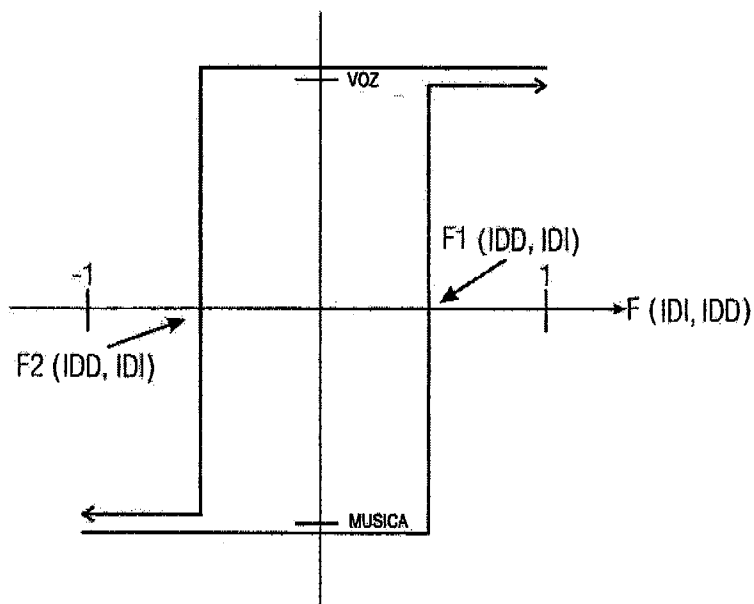


FIG. 3

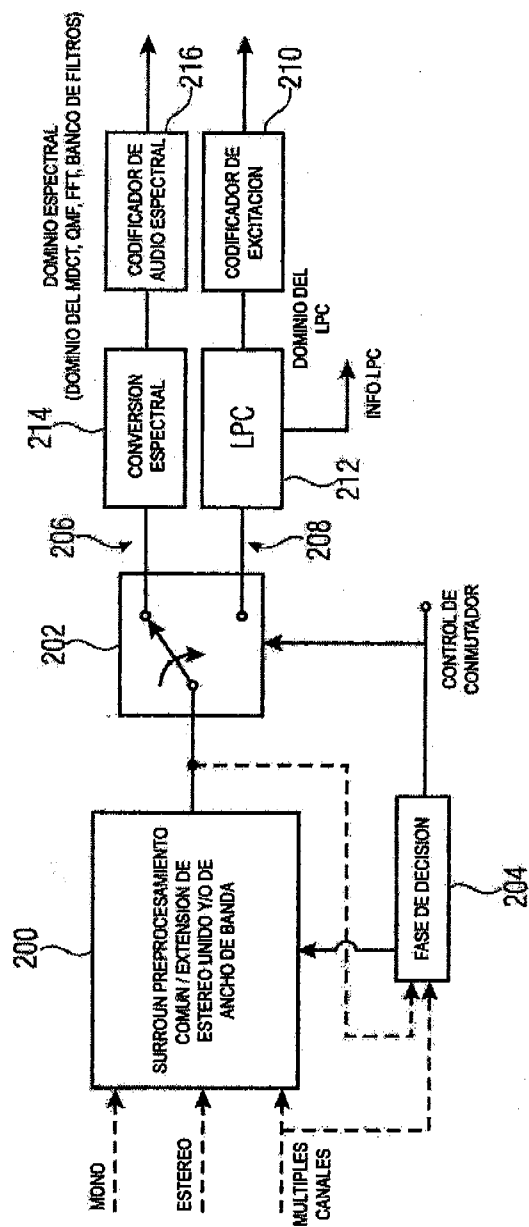


FIG. 4
(CODIFICADOR)

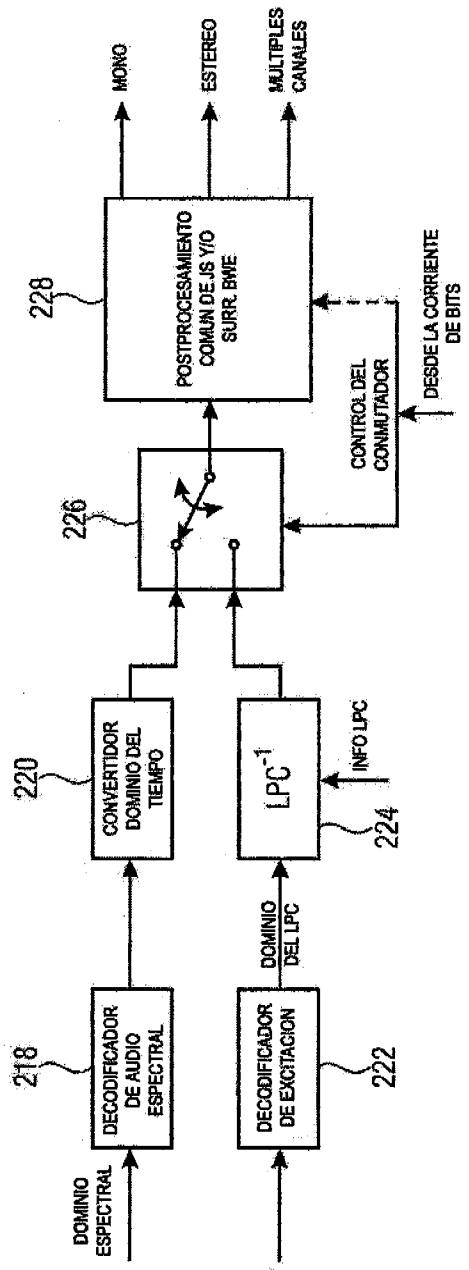


FIG. 5
(DECODIFICADOR)

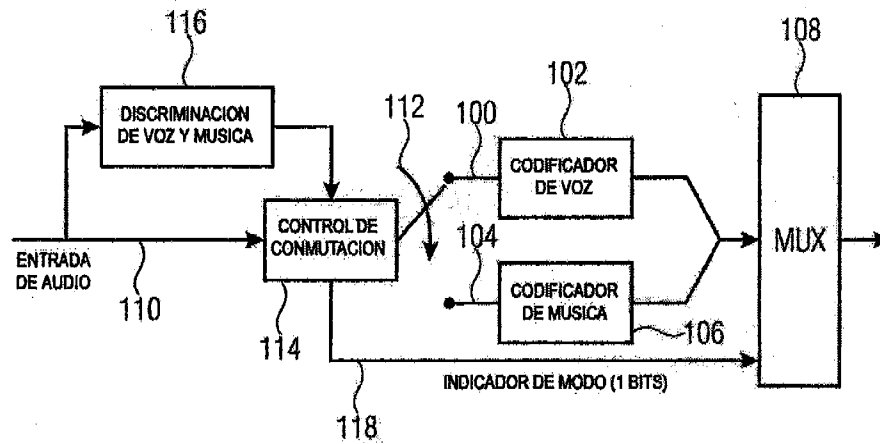


FIG. 6

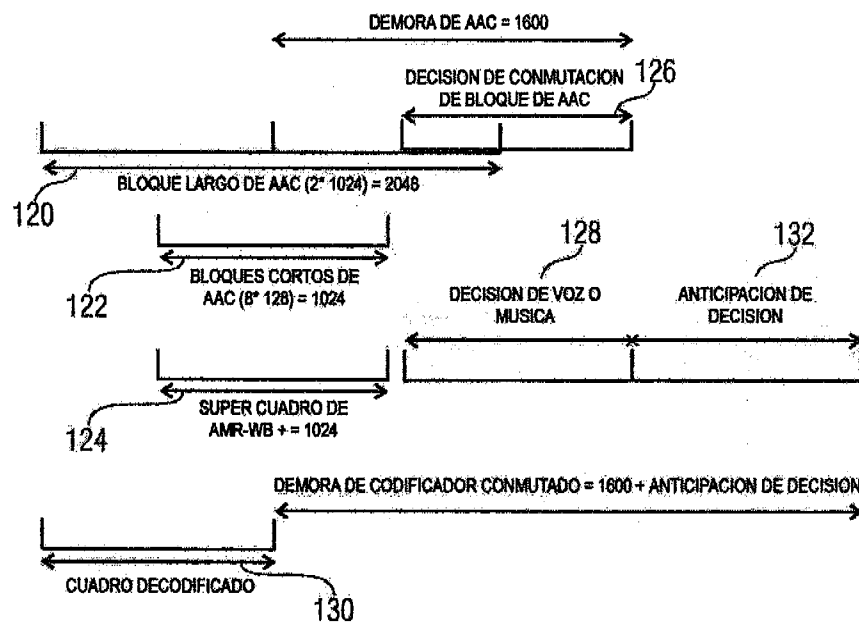


FIG. 7

Method and Discriminator for Classifying Different Segments of a Signal

5

Background of the Invention

The invention relates to an approach for classifying different segments of a signal comprising segments of at least a first type and a second type. Embodiments of the invention relate to the field of audio coding and, particularly, to the speech/music discrimination upon encoding an audio signal.

In the art, frequency domain coding schemes such as MP3 or AAC are known. These frequency-domain encoders are based on a time-domain/frequency-domain conversion, a subsequent quantization stage, in which the quantization error is controlled using information from a psychoacoustic module, and an encoding stage, in which the quantized spectral coefficients and corresponding side information are entropy-encoded using code tables.

On the other hand there are encoders that are very well suited to speech processing such as the AMR-WB+ as described in 3GPP TS 26.290. Such speech coding schemes perform a Linear Predictive filtering of a time-domain signal. Such a LP filtering is derived from a Linear Prediction analysis of the input time-domain signal. The resulting LP filter coefficients are then coded and transmitted as side information. The process is known as Linear Prediction Coding (LPC). At the output of the filter, the prediction residual signal or prediction error signal which is also known as the excitation signal is encoded using the analysis-by-synthesis stages of the ACELP encoder or, alternatively, is encoded using a transform encoder, which uses a Fourier transform with an overlap. The decision between the ACELP coding and the Transform Coded eXcitation coding which is also called TCX coding is done using a closed loop or an open loop algorithm.

Frequency-domain audio coding schemes such as the high efficiency-AAC encoding scheme, which combines an AAC coding scheme and a spectral bandwidth replication technique may also be combined to a joint stereo or a multi-channel coding tool which is known under the term "MPEG surround". Frequency-domain coding schemes are advantageous in that they show a high quality at low bit rates for music signals. Problematic, however, is the quality of speech signals at low bit rates.

On the other hand, speech encoders such as the AMR-WB+ also have a high frequency enhancement stage and a stereo functionality. Speech coding schemes show a high quality for speech signals even at low bit rates, but show a poor quality for music signals at low bit rates.

40

In view of the available coding schemes mentioned above, some of which are better suited for encoding speech and others being better suited for encoding music, the automatic segmentation and classification of an audio signal to be encoded is an important tool in many

multimedia applications and may be used in order to select an appropriate process for each different class occurring in an audio signal. The overall performance of the application is strongly dependent on the reliability of the classification of the audio signal. Indeed, a false classification generates mis-suited selections and tunings of the following processes.

5

Fig. 6 shows a conventional coder design used for separately encoding speech and music dependent on the discrimination of an audio signal. The coder design comprises a speech encoding branch 100 including an appropriate speech encoder 102, for example an AMR-WB+ speech encoder as it is described in "Extended Adaptive Multi-Rate – Wideband (AMR-WB+) codec", 3GPP TS 26.290 V6.3.0, 2005-06, Technical Specification. Further, the coder design comprises a music encoding branch 104 comprising a music encoder 106, for example an AAC music encoder as it is, for example, described in Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding. International Standard 13818-7, ISO/IEC JTC1/SC29/WG11 Moving Pictures Expert Group, 1997.

15

The outputs of the encoders 102 and 106 are connected to an input of a multiplexer 108. The inputs of the encoders 102 and 106 are selectively connectable to an input line 110 carrying an input audio signal. The input audio signal is applied selectively to the speech encoder 102 or the music encoder 106 by means of a switch 112 shown schematically in Fig. 6 and being controlled by a switching control 114. In addition, the coder design comprises a speech/music discriminator 116 also receiving at an input thereof the input audio signal and outputting a control signal to the switch control 114. The switch control 114 further outputs a mode indicator signal on a line 118 which is input into a second input of the multiplexer 108 so that a mode indicator signal can be sent together with an encoded signal. The mode indicator signal may have only one bit indicating that a datablock associated with the mode indicator bit is either speech encoded or music encoded so that, for example, at a decoder no discrimination needs to be made. Rather, on the basis of the mode indicator bit submitted together with the encoded data to the decoder side an appropriate switching signal can be generated on the basis of the mode indicator for routing the received and encoded data to an appropriate speech or music decoder.

30

Fig. 6 is a traditional coder design which is used to digitally encode speech and music signals applied to line 110. Generally, speech encoders do better on speech and audio encoders do better on music. A universal coding scheme can be designed by using a multi-coder system which switches from one coder to another according to the nature of the input signal. The non-trivial problem here is to design a well-suited input signal classifier which drives the switching element. The classifier is the speech/music discriminator 116 shown in Fig. 6. Usually, a reliable classification of an audio signal introduces a high delay, whereas, on the other hand, the delay is an important factor in real-time applications.

40

In general, it is desired that the overall algorithmic delay introduced by the speech/music discriminator is sufficiently low to be able to use the switched coders in a real-time application.

Fig. 7 illustrates the delays experienced in a coder design as shown in Fig. 6. It is assumed that the signal applied on input line 110 is to be coded on a frame basis of 1024 samples at a 16 kHz sampling rate so that the speech/music discrimination should deliver a decision ever frame, i.e. every 64 milliseconds. The transition between two encoders is for example effected in a manner as described in WO 2008/071353 A2 and the speech/music discriminator should not significantly increase the algorithmic delay of the switched decoders which is in total 1600 samples without considering the delay needed for the speech/music discriminator. It is further desired to provide the speech/music decision for the same frame where AAC block switching is decided. The situation is depicted in Fig. 7 illustrating an AAC long block 120 having a length of 2048 samples, i.e. the long block 120 comprises two frames of 1024 samples, an ACC short block 122 of one frame of 1024 samples, and an AMR-WB+ superframe 124 of one frame of 1024 samples.

In Fig. 7, the AAC block-switching decision and speech/music decision are taken on the frames 126 and 128 respectively of 1024 samples, which cover the same period of time. The two decisions are taken at this particular position for making the coding able to use at a time transition windows for going properly from one mode to the other one. In consequence, a minimum delay of $512+64$ samples is introduced by the two decisions. This delay has to be added to the delay of 1024 samples generated by the 50% overlap from the AAC MDCT which gives a minimal delay of 1600 samples. In a conventional AAC, only the block-switching is present and the delay is exactly 1600 samples. This delay is needed for switching at a time from a long block to short blocks when transients are detected in the frame 126. This switching of transformation length is desirable for avoiding pre-echo artifact. The decoded frame 130 in Fig. 7 represents the first whole frame which can be restituted at the decoder side in any case (long or short blocks).

In a switched coder using AAC as a music encoder, the switching decision coming from a decision stage should avoid adding too much additional delay to the original AAC delay. The additional delay comes from the lookahead frame 132 which is needed for the signal analysis in the decision stage. At a sampling rate of for example 16kHz, the AAC delay is 100 ms while a conventional speech/music discriminator uses around 500 ms of lookahead, which will result to a switched coding structure with a delay of 600 ms. The total delay will then be six times that of the original AAC delay.

Conventional approaches as described above are disadvantageous as for a reliable classification of an audio signal a high, undesired delay is introduced so that a need for a novel approach exists for discriminating a signal including segments of different types, wherein an additional algorithmic delay introduced by the discriminator is sufficiently low so that the switched coders may also be used for a real-time application.

J. Wang, et. al. "Real-time speech/music classification with a hierarchical oblique decision tree", ICASSP 2008, IEEE International Conference on Acoustics, Speech and Signal

Processing, 2008, March 31, 2008 to April 4, 2008 describes an approach for speech/music classification using short-term features and long term features derived from the same number of frames. These short-term features and long term features are used for classifying the signal, but only limited properties of the short-term features are exploited, for example the reactivity of the classification is not exploited, although it has an important role for most audio coding applications.

Summary of the Invention

It is an object of the invention to provide an improved approach for discriminating in a signal segments of different type while keeping low any delay introduced by the discrimination.

This object is achieved by a method of claim 1 and by a discriminator of claim 14.

One embodiment of the invention provides a method for classifying different segments of a signal, the signal comprising segments of at least a first type and a second type, the method comprising:

short-term classifying the signal on the basis of at least one short-term feature extracted from the signal and delivering a short-term classification result;

long-term classifying the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal and delivering a long-term classification result; and

combining the short-term classification result and the long-term classification result to provide an output signal indicating whether a segment of the signal is of the first type or of the second type.

Another embodiment of the invention provides a discriminator, comprising:

a short-term classifier configured to receive a signal and to provide a short-term classification result of the signal on the basis of at least one short-term feature extracted from the signal, the signal comprising segments of at least a first type and a second type;

a long-term classifier configured to receive the signal and to provide a long-term classification result of the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal;

a decision circuit configured to combine the short-term classification result and the long-term classification result to provide an output signal indicating whether a segment of the signal is of the first type or of the second type.

- 5 Embodiments of the invention provide the output signal on the basis of a comparison of the short-term analysis result to the long-term analysis result.

Embodiments of the invention concern an approach to classify different non-overlapped short time segments of an audio signal either as speech or as non-speech or further classes. The approach is based on the extraction of features and the analysis of their statistics over two different analysis window lengths. The first window is long and looks mainly to the past. The first window is used to get a reliable but delayed decision clue for the classification of the signal. The second window is short and considers mainly the segment processed at the present time or the current segment. The second window is used to get an instantaneous decision clue. The two decision clues are optimally combined, preferably by using a hysteresis decision which gets the memory information from the delayed clue and the instantaneous information from the instantaneous clue.

Embodiments of the invention use short-term features both in the short-term classifier and in the long-term classifier so that the two classifiers exploit different statistics of the same feature. The short-term classifier will extract only the instantaneous information because it has access only to one set of features. For example, it can exploit the mean of the features. On the other hand, the long-term classifier has access to several sets of features because it considers several frames. As a consequence, the long-term classifier can exploit more characteristics of the signal by exploiting statistics over more frames than the short-term classifier. For example, the long-term classifier can exploit the variance of the features or the evolution of features over the time. Thus, the long-term classifier may exploit more information than the short-term classifier, but it introduces delay or latency. However, the long-term features, despite introducing delay or latency, will make the long-term classification results more robust and reliable. In some embodiments the short-term and long-term classifiers may consider the same short-term features, which may be computed once and used by the both classifiers. Thus, in such an embodiment the long-term classifier may receive the short-term features directly from the short-term classifier.

The new approach thereby permits to get a classification which is robust while introducing a low delay. Other than conventional approaches, embodiments of the invention limit the delay introduced by the speech/music decision while keeping a reliable decision. In one embodiment of the invention, the lookahead is limited to 128 samples, which results of a total delay of only 108 ms.

40

Brief Description of the Drawings

Embodiments of the invention will be described below with reference to the accompanying drawings, in which:

- 5 Fig. 1 is a block diagram of a speech/music discriminator in accordance with an embodiment of the invention;
- Fig. 2 illustrates the analysis windows used by the long-term and the short-term classifiers of the discriminator of Fig. 1;
- 10 Fig. 3 illustrates the hysteresis decision used in the discriminator of Fig. 1;
- Fig. 4 is a block diagram of an exemplary encoding scheme comprising a discriminator in accordance with embodiments of the invention;
- 15 Fig. 5 is a block diagram of a decoding scheme corresponding to the encoding scheme of Fig. 4;
- Fig. 6 shows a conventional coder design used for separately encoding speech and music dependent on a discrimination of an audio signal; and
- 20 Fig. 7 illustrates the delays experienced in the coder design shown in Fig. 6.

Description of Embodiments of the Invention

25

Fig. 1 is a block diagram of a speech/music discriminator 116 in accordance with an embodiment of the invention. The speech/music discriminator 116 comprises a short-term classifier 150 receiving at an input thereof an input signal, for example an audio signal comprising speech and music segments. The short-term classifier 150 outputs on an output

30 line 152 a short-term classification result, the instantaneous decision clue. The discriminator 116 further comprises a long-term classifier 154 which also receives the input signal and outputs on an output line 156 the long-term classification result, the delayed decision clue. Further, an hysteresis decision circuit 158 is provided which combines the output signals from the short-term classifier 150 and the long-term classifier 154 in a manner as will be described

35 in further detail below to generate a speech/music decision signal which is output on line 160 and may be used for controlling the further processing of a segment of an input signal in a manner as is described above with regard to Fig. 6, i.e. the speech/music decision signal 160 may be used to route the input signal segment which has been classified to a speech encoder or to an audio encoder.

40

Thus, in accordance with embodiments of the invention two different classifiers 150 and 154 are used in parallel on the input signal applied to the respective classifiers via input line 110. The two classifiers are called long-term classifier 154 and short-term classifier 150, wherein

the two classifiers differ by analyzing the statistics of the features on which they operate over analysis windows. The two classifiers deliver the output signals 152 and 156, namely the instantaneous decision clue (IDC) and the delayed decision clue (DDC). The short-term classifier 150 generates the IDC on the basis of short-term features that have the aim to capture instant information about the nature of the input signal. They are related to short-term attributes of the signal which can rapidly and at any time change. In consequence the short-term features are expected to be reactive and not to introduce a long delay to the whole discriminating process. For example, since the speech is considered to be quasi-stationary on 5-20ms durations, the short-term features may be computed every frame of 16 ms on a signal sampled at 16 kHz. The long-term classifier 154 generates the DDC on the basis of features resulting from longer observations of the signal (long-term features) and therefore permits to achieve more reliable classification.

Fig. 2 illustrates the analysis windows used by the long-term classifier 154 and the short-term classifier 150 shown in Fig. 1. Assuming a frame of 1024 samples at a sampling rate of 16 kHz the length of the long-term classifier window 162 is $4 \times 1024 + 128$ samples, i.e., the long-term classifier window 162 spans four frames of the audio signal and additional 128 samples are needed by the long-term classifier 154 to make its analysis. This additional delay, which is also referred to as the "lookahead", is indicated in Fig. 2 at reference sign 164. Fig. 2 also shows the short-term classifier window 166 which is $1024 + 128$ samples, i.e. spans one frame of the audio signal and the additional delay needed for analyzing a current segment. The current segment is indicated at 128 as the segment for which the speech/music decision needs to be made.

The long-term classifier window indicated in Fig. 2 is sufficiently long to obtain the 4-Hz energy modulation characteristic of speech. The 4-Hz energy modulation is a relevant and discriminate characteristic of speech which is traditionally exploited in robust speech/music discriminators used as for example by Scheirer E. and Slaney M., "Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator", ICASSP'97, Munich, 1997. The 4-Hz energy modulation is a feature which can be only extracted by observing the signal on a long time segment. The additional delay which is introduced by the speech/music discriminator is equal to the lookahead 164 of 128 samples which is needed by each of the classifiers 150 and 154 to make the respective analysis like a perceptual linear prediction analysis as it is described by H. Hermansky, "Perceptive linear prediction (plp) analysis of speech," Journal of the Acoustical Society of America, vol. 87, no. 4, pp. 1738-1752, 1990 and H. Hermansky, et al., "Perceptually based linear predictive analysis of speech," ICASSP 5.509-512, 1985. Thus, when using the discriminator of the above embodiment in an encoder design as shown in Fig. 6, the overall delay of the switched coders 102 and 106 will be $1600 + 128$ samples which equals 108 milliseconds which is sufficiently low for real-time applications.

Reference is now made to Fig. 3 describing the combining of the output signals 152 and 156 of the classifiers 150 and 154 of the discriminator 116 for obtaining a speech/music decision

signal 160. The delayed decision clue DDC and the instantaneous decision clue IDC, in accordance with an embodiment of the invention, are combined by using a hysteresis decision. Hysteresis processes are widely used to post process decisions in order to stabilize them. Fig. 3 illustrates a two-state hysteresis decision as a function of the DDC and the IDC to determine whether the speech/music decision signal should indicate a currently processed segment of the input signal as being a speech segment or a music segment. The characteristic hysteresis cycle is seen in Fig. 3 and IDC and DDC are normalized by the classifiers 150 and 154 in such a way that the values are between -1 and 1, wherein -1 means that the likelihood is totally music-like, and 1 means that the likelihood is totally speech-like.

The decision is based on the value of a function $F(IDC, DDC)$ examples of which will be described below. In Fig. 3, $F1(DDC, IDC)$ indicates a threshold that $F(IDC, DDC)$ should cross to go from a music state to a speech state. $F2(DDC, IDC)$ illustrates a threshold that $F(IDC, DDC)$ should cross to go from the speech state to the music state. The final decision $D(n)$ for a current segment or current frame having the index n may then be calculated on the basis of the following pseudo code:

%Hysteresis Decision Pseudo Code

If($D(n-1) == \text{music}$)

 If($F(IDC, DDC) < F1(DDC, IDC)$)

$D(n) == \text{music}$

 Else

$D(n) == \text{speech}$

Else

 If($F(IDC, DDC) > F2(DDC, IDC)$)

$D(n) == \text{speech}$

 Else

$D(n) == \text{music}$

%End Hysteresis Decision Pseudo Code

In accordance with embodiments of the invention the function $F(IDC, DDC)$ and the above-mentioned thresholds are set as follows:

$F(IDC, DDC) = IDC$

$F1(IDC, DDC) = 0.4 - 0.4 * DDC$

$F2(IDC, DDC) = -0.4 - 0.4 * DDC$

Alternatively, the following definitions may be used:

$F(IDC, DDC) = (2 * IDC + DDC) / 3$

$F1(IDC, DDC) = -0.75 * DDC$

$F2(IDC, DDC) = -0.75 * DDC$

When using the last definition the hysteresis cycle vanishes and the decision is made only on the basis a unique adaptive threshold.

5 The invention is not limited to the hysteresis decision described above. In the following further embodiments for combining the analysis results for obtaining the output signal will be described.

A simple thresholding can be used instead of the hysteresis decision by making the threshold in a way that it exploits both the characteristics of DDC and IDC. DDC is considered to be a
10 more reliable discriminate clue because it comes from a longer observation of the signal. However, DDC is computed based partly on the past observation of the signal. A conventional classifier which only compares the value DDC to the threshold 0, and by classifying a segment as speech-like when $DDC > 0$ or as music-like otherwise, will have a delayed decision. In one
15 embodiment of the invention, we may adapt the thresholding by exploiting the IDC and make the decision more reactive. For this purpose, the threshold can adapted on the basis of the following pseudo-code:

% Pseudo code of adaptive thresholding

If($DDC > -0.5 * IDC$)

20 $D(n) = \text{speech}$

Else

$D(n) = \text{music}$

%End of adaptive thresholding

25 In another embodiment, the DDC may be used for making more reliable the IDC. The IDC is known to be reactive but not as reliable as DDC. Furthermore, looking to the evolution of the DDC between the past and current segment may give another indication how the frame 166 in Fig. 2 influences the DDC calculated on the segment 162. The notation $DDC(n)$ is used for the current value of the DDC and $DDC(n-1)$ for the past value. Using both values, $DDC(n)$ and
30 $DDC(n-1)$, IDC may be made more reliable by using a decision tree as it is described as follows:

% Pseudo code of decision tree

If($IDC > 0 \ \&\& \ DDC(n) > 0$)

35 $D(n) = \text{speech}$

Else if ($IDC < 0 \ \&\& \ DDC(n) < 0$)

$D(n) = \text{music}$

Else if ($IDC > 0 \ \&\& \ DDC(n) - DDC(n-1) > 0$)

$D(n) = \text{speech}$

40 Else if ($IDC < 0 \ \&\& \ DDC(n) - DDC(n-1) < 0$)

$D(n) = \text{music}$

Else if ($DDC > 0$)

$D(n) = \text{speech}$

Else

$D(n)=\text{music}$

%End of decision tree

- 5 In above decision tree, the decision is directly taken if the both clues show the same likelihood. If the two clues give contradictory indications, we look at the evolution of the DDC. If the difference $DDC(n)-DDC(n-1)$ is positive, we may suppose that the current segment is speech-like. Otherwise, we may suppose that the current segment is music-like. If this new indication goes to the same direction as the IDC, the final decision is then taken. If
10 the both attempts fail to give a clear decision, the decision is taken by considering only the delayed clue DDC since IDC reliability was not able to be validated.

In the following, the respective classifiers 150 and 154 in accordance with an embodiment of the invention will be described in further detail.

15

- Turning first of all to the long-term classifier 154 it is noted same is for extracting from every sub-frame of 256 samples a set of features. The first feature is the Perceptual Linear Prediction Cepstral Coefficient (PLPCC) as described by H. Hermansky, "Perceptive linear prediction (plp) analysis of speech," Journal of the Acoustical Society of America, vol. 87, no. 4, pp.
20 1738-1752, 1990 and H. Hermansky, et al., "Perceptually based linear predictive analysis of speech," ICASSP 5.509-512, 1985. PLPCCs are efficient for speaker classification by using human auditory perception estimation. This feature may be used to discriminate speech and music and, indeed permits to distinguish the characteristic formants of the speech as well as the syllabic 4-Hz modulation of the speech by looking to the feature variation over time.

25

- However, to be more robust, the PLPCCs are combined with another feature which is able to capture pitch information, which is another important characteristic of speech and may be critical in coding. Indeed, speech coding relies on the assumption that an input signal is a pseudo mono-periodic signal. The speech coding schemes are efficient for such a signal. On
30 the other hand, the pitch characteristic of speech harms a lot of the coding efficiency of music coders. The smooth pitch delay fluctuation given the natural vibrato of the speech makes the frequency representation in the music coders unable to compact greatly the energy which is required for obtaining a high coding efficiency.

- 35 The following pitch characteristic features may be determined:

Glottal Pulses Energy Ratio:

- This feature computes the ratio of energy between the glottal pulses and the LPC residual signal. The glottal pulses are extracted from the LPC residual signal by using a pick-peaking
40 algorithm. Usually, the LPC residual of a voiced segment shows a great pulse-like structure coming from the glottal vibration. The feature is high during voiced segments.

Long-term gain prediction:

It is the gain usually computed in speech coders (see e.g. "Extended Adaptive Multi-Rate – Wideband (AMR-WB+) codec", 3GPP TS 26.290 V6.3.0, 2005-06, Technical Specification) during the long-term prediction. This feature measures the periodicity of the signal and is based on pitch delay estimation.

5

Pitch delay fluctuation:

This feature determines the difference of the present pitch delay estimation when compared to the last sub-frame. For voiced speech this feature should be low but not zero and evolve smoothly.

10

Once the long-term classifier has extracted the required set of features a statistical classifier is used on these extracted features. The classifier is at first trained by extracting the features over a speech training set and a music training set. The extracted features are normalized to a mean value of 0 and a variance of 1 over both training sets. For each training set, the extracted and normalized features are gathered within a long-term classifier window and modeled by a

15

Gaussians Mixture Model (GMM) using five Gaussians. At the end of the training sequence a set of normalizing parameters and two sets of GMM parameters are obtained and saved.

For each frame to classify, the features are first extracted and normalized with the normalizing parameters. The maximum likelihood for speech (lld_speech) and the maximum likelihood for music (lld_music) are computed for the extracted and normalized features using the GMM of the speech class and the GMM of the music class, respectively. The delayed decision clue DDC is then calculated as follows:

20

$$25 \quad DDC = (lld_speech - lld_music) / (abs(lld_music) + abs(lld_speech))$$

DDC is bound between -1 and 1, and is positive when the maximum likelihood for speech is higher than the maximum likelihood for music, $lld_speech > lld_music$.

30

The short-term classifier uses as a short-term feature the PLPCCs. Other than in the long-term classifier, this feature is only analyzed on the window 128. The statistics on this feature are exploited on this short time by a Gaussians Mixture Model (GMM) using five Gaussians. Two models are trained, one for music, and another for speech. It is worth notifying, that the two models are different than the ones obtained for the long-term classifier. For each frame to

35

classify, the PLPCCs are first extracted and the maximum likelihood for speech (lld_speech) and the maximum likelihood for music (lld_music) are computed for using the GMM of the speech class and the GMM of the music class, respectively. The instantaneous decision clue IDC is then calculated as follows:

$$40 \quad IDC = (lld_speech - lld_music) / (abs(lld_music) + abs(lld_speech))$$

IDC is bound between -1 and 1.

Thus, the short-term classifier 150 generates the short-term classification result of the signal on the basis of the feature "Perceptual Linear Prediction Cepstral Coefficient (PLPCC)", and the long-term classifier 154 generates the long-term classification result of the signal on the basis of the same feature "Perceptual Linear Prediction Cepstral Coefficient (PLPCC)" and the above mentioned additional feature(s), e.g. pitch characteristic feature(s). Moreover, the long-term classifier can exploit different characteristics of the shared feature, i.e. PLPCCs, as it has access to a longer observation window. Thus, upon combining the short-term and long-term results the short-term features are sufficiently considered for the classification, i.e. its properties are sufficiently exploited.

Below a further embodiment for the respective classifiers 150 and 154 will be described in further detail.

The short-term features analyzed by the short-term classifier in accordance with this embodiment correspond mainly to the Perceptual Linear Prediction Cepstral Coefficients (PLPCCs) mentioned above. The PLPCCs are widely used in speech and speaker recognition as well as the MFCCs (see above). The PLPCCs are retained because they share a great part of the functionality of the Linear Prediction (LP) which is used in most of the modern speech coder and so already implemented in a switched audio coder. The PLPCCs can extract the formant structure of the speech as the LP does, but by taking into account perceptual considerations PLPCCs are more speaker independent and thus more relevant regarding the linguistic information. An order of 16 is used on the 16 kHz sampled input signal.

Apart from the PLPCCs, a voicing strength is computed as a short-term feature. The voicing strength is not considered to be really discriminating by itself, but is beneficial in association with the PLPCCs in the feature dimension. The voicing strength permits to draw in the features dimension at least two clusters corresponding respectively to the voiced and the unvoiced pronunciations of the speech. It is based on a merit calculation using different Parameters namely a Zero crossing Counter (zc), the spectral tilt (tilt), the pitch stability (ps), and the normalized correlation of the pitch (nc). All the four parameters are normalized between 0 and 1 in a way that 0 corresponds to a typical unvoiced signal and 1 corresponds to a typical voiced signal. In this embodiment the voicing strength is inspired from the speech classification criteria used in the VMR-WB speech coder described by Milan Jelinek and Redwan Salami, "Wideband speech coding advances in vmr-wb standard," IEEE Trans. on Audio, Speech and Language Processing, vol. 15, no. 4, pp. 1167-1179, May 2007. It is based on an evolved pitch tracker based on auto-correlation. For the frame index k the voicing strength $u(k)$ has the form below:

$$v(k) = \frac{1}{5}(2 * nc(k) + 2 * ps(k) + tilt(k) + zc(k))$$

The discriminating ability of the short-term features is evaluated by Gaussian Mixture Models (GMMS) as a classifier. Two GMMs, one for the speech class and the other for the music

class, are applied. The number of mixtures is made varying in order to evaluate the effect on the performance. Table 1 shows the accuracy rates for the different number of mixtures. A decision is computed for every segment of four successive frames. The overall delay is then equal to 64ms which is suitable for a switched audio coding. It can be observed that the performance increases with the number of mixtures. The gap between 1-GMMs and 5-GMMs is particularly important and can be explained by the fact that the formant representation of the speech is too complex to be sufficiently defined by only one Gaussian.

	1-GMMs	5-GMMs	10-GMMs	20-GMMs
Speech	95.33	96.52	97.02	97.60
Music	92.17	91.97	91.61	91.77
Average	93.75	94.25	94.31	94.68

Table 1: Short-term features classification accuracy in %

Turning now to the long-term classifier, it is noted that many works, e.g. M. J. Carey, et. al. "A comparison of features for speech and music discrimination," Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP, vol. 12, pp. 149 to 152, March 1999, consider variances of statistic features to be more discriminating than the features themselves. As a rough general rule, music can be considered more stationary and exhibits usually lower variance. On the contrary, speech can be easily distinguished by its remarkable 4-Hz energy modulation as the signal periodically changes between voiced and unvoiced segments. Moreover the succession of different phonemes makes the speech features less constant. In this embodiment, two long-term features are considered, one based on a variance computation and the other based on a priori knowledge of the pitch contour of the speech. The long-term features are adapted to the low delay SMD (speech/music discrimination).

The moving variance of the PLPCCs consists of computing the variance for each set of PLPCCs over an overlapping analysis window covering several frames in order to emphasize the last frame. To limit the introduced latency, the analysis window is asymmetric and considers only the current frame and the past history. In a first step, the moving average $ma_m(k)$ of the PLPCCs is computed over the last N frames as described as follows:

$$ma_m(k) = \sum_{i=0}^{N-1} PLPC_m(k-i) \cdot w(i)$$

where $PLP_m(k)$ is the mth cepstral coefficient over a total of M coefficients coming from the kth frame. The moving variance $mv_m(k)$ is then defined as:

$$mv_m(k) = \sum_{i=0}^{N-1} (PLPC_m(k-i) - ma_m(k))^2 \cdot w(i)$$

where w is a window of length N which is in this embodiment a ramp slope defined as follows:

$$w(i) = (N-i) / N \cdot (N+1) / 2$$

5

The moving variance is finally averaged over the cepstral dimension:

$$mv(k) = \frac{1}{M} \sum_{m=0}^M mv_m(k)$$

- 10 The pitch of the speech has remarkably properties and part of them can only be observed on long analysis windows. Indeed the pitch of speech is smoothly fluctuating during the voiced segments but is seldom constant. On the contrary, music exhibits very often constant pitch during the whole duration of a note and abrupt changes during transients. The long-term features encompass this characteristic by observing the pitch contour on a long time segment.

- 15 A pitch contour parameter $pc(k)$ is defined as:

$$pc(k) = \begin{cases} 0 & \text{if } |p(k) - p(k-1)| < 1 \\ 0,5 & \text{if } 1 \leq |p(k) - p(k-1)| < 2 \\ 1 & \text{if } 2 \leq |p(k) - p(k-1)| < 20 \\ 0,5 & \text{if } 20 \leq |p(k) - p(k-1)| < 25 \\ 0 & \text{otherwise} \end{cases}$$

- 20 where $p(k)$ is the pitch delay computed at the frame index k on the LP residual signal sampled at 16Hz. From the pitch contour parameter, a speech merit, $sm(k)$, is computed in a way that speech is expected to display a smoothly fluctuating pitch delay during voiced segments and a strong spectral tilt towards high frequencies during unvoiced segments:

$$sm(k) = \begin{cases} nc(k) \cdot pc(k) & \text{if } v(k) \geq 0.5 \\ (1 - nc(k)) \cdot (1 - tilt(k)) & \text{otherwise} \end{cases}$$

25

where $nc(k)$, $tilt(k)$, and $v(k)$ are defined as above (see the short term classifier). The speech merit is then weighted by the window w defined above and integrated over the last N frames:

$$ams(k) = \sum_{i=0}^N m(k-1)w(i)$$

30

- The pitch contour is also an important indication that a signal is suitable for a speech or an audio coding. Indeed speech coders work mainly in time domain and make the assumption that the signal is harmonic and quasi-stationary on short time segments of about 5ms. In this manner they may model efficiently the natural pitch fluctuation of the speech. On the contrary, the same fluctuation harms the efficiency of general audio encoders which exploit linear transformations on long analysis windows. The main energy of the signal is then spread over several transformed coefficients.
- 35

As for the short-term features, also the long-term features are evaluated using a statistical classifier thereby obtaining the long-term classification result (DDC). The two features are computed using $N = 25$ frames, e.g. considering 400 ms of past history of the signal. A Linear Discriminant Analysis (LDA) is first applied before using 3-GMMs in the reduced one-dimensional space. Table 2 shows the performance measured on the training and the testing sets when classifying segments of four successive frames.

	Training Set	Test Set
Speech	97.99	97.84
Music	95.93	95.44
Average	96.96	96.64

Table 2: Long-term features classification accuracy in %

The combined classifiers system according to embodiments of the invention combines appropriately the short-term and long-term features in way that they bring their own specific contribution to the final decision. For this purpose a hysteresis final decision stage as described above may be used, where the memory effect is driven by the DDC or long-term discriminating clue (LTDC) while the instant input comes from the IDC or short-term discriminating clue (STDC). The two clues are the outputs of the long-term and short-term classifiers as illustrated in Fig. 1. The decision is taken based on the IDC but is stabilized by the DDC which controls dynamically the thresholds triggering a change of state.

The long-term classifier uses both the long-term and short-term features previously defined with a LDA followed by 3-GMMs. The DDC is equal to the logarithmic ratio of the long-term classifier likelihood of the speech class and the music class computed over the last 4 X K frames. The number of frames taken into account may vary with the parameter K in order to add more or less memory effect in the final decision. On the contrary, the short-term classifier uses only the short-term features with 5-GMMs which show a good compromise between performance and complexity. The IDC is equal to the logarithmic ratio of the short-term classifier likelihood of the speech class and the music class computed only over the last 4 frames.

In order to evaluate the inventive approach, especially for a switched audio coding, three different kinds of performances were evaluated. A first performance measurement is the conventional speech against music (SvM) performance. It is evaluated over a large set of music and speech items. A second performance measurement is done on a large unique item having speech and music segments alternating every 3 seconds. The discriminating accuracy is then called speech after/before music (SabM) performance and reflects mainly the reactivity of the system. Finally, the stability of the decision is evaluated by performing the classification on a large set of speech over music items. The mixing between speech and music is done at different levels from one item to another. The speech over music (SoM) performance is then

obtained by computing the ratio of the number class switches that occurred over the total number of frames.

The long term classifier and the short-term classifier are used as references for evaluating conventional single classifier approaches. The short-term classifier shows a good reactivity while having lower stability and overall discriminating ability. On the other hand, the long-term classifier, especially by increasing the number of frames $4 \times K$, can reach better stability and discriminating behaviour by compromising the reactivity of the decision. When compared to the just mentioned conventional approach, the performances of the combined classifier system in accordance with the invention has several advantages. One advantage is that it maintains a good pure speech against music discrimination performance while preserving the reactivity of the system. A further advantage is the good trade-off between reactivity and stability.

In the following, reference is made to Figs. 4 and 5 illustrating exemplary encoding and decoding schemes which include a discriminator or decision stage operating in accordance with embodiments of the invention.

In accordance with the exemplary encoding scheme shown in Fig. 4 a mono signal, a stereo signal or a multi-channel signal is input into a common preprocessing stage 200.

The common preprocessing stage 200 may have a joint stereo functionality, a surround functionality, and/or a bandwidth extension functionality. At the output of stage 200 there is a mono channel, a stereo channel or multiple channels which is input into one or more switches 202. The switch 202 may be provided for each output of stage 200, when stage 200 has two or more outputs, i.e., when stage 200 outputs a stereo signal or a multi-channel signal. Exemplarily, the first channel of a stereo signal may be a speech channel and the second channel of the stereo signal may be a music channel. In this case, the decision in a decision stage 204 may be different between the two channels at the same time instant.

The switch 202 is controlled by the decision stage 204. The decision stage comprises a discriminator in accordance with embodiments of the invention and receives, as an input, a signal input into stage 200 or a signal output by stage 200. Alternatively, the decision stage 204 may also receive a side information which is included in the mono signal, the stereo signal or the multi-channel signal or is at least associated with such a signal, where information is existing, which was, for example, generated when originally producing the mono signal, the stereo signal or the multi-channel signal.

In one embodiment, the decision stage does not control the preprocessing stage 200, and the arrow between stage 204 and 200 does not exist. In a further embodiment, the processing in stage 200 is controlled to a certain degree by the decision stage 204 in order to set one or more parameters in stage 200 based on the decision. This will, however not influence the general

algorithm in stage 200 so that the main functionality in stage 200 is active irrespective of the decision in stage 204.

5 The decision stage 204 actuates the switch 202 in order to feed the output of the common preprocessing stage either in a frequency encoding portion 206 illustrated at an upper branch of Fig. 4 or an LPC-domain encoding portion 208 illustrated at a lower branch in Fig. 4.

10 In one embodiment, the switch 202 switches between the two coding branches 206, 208. In a further embodiment, there may be additional encoding branches such as a third encoding branch or even a fourth encoding branch or even more encoding branches. In an embodiment with three encoding branches, the third encoding branch may be similar to the second encoding branch, but includes an excitation encoder different from the excitation encoder 210 in the second branch 208. In such an embodiment, the second branch comprises the LPC stage 212 and a codebook based excitation encoder 210 such as in ACELP, and the third branch
15 comprises an LPC stage and an excitation encoder operating on a spectral representation of the LPC stage output signal.

20 The frequency domain encoding branch comprises a spectral conversion block 214 which is operative to convert the common preprocessing stage output signal into a spectral domain. The spectral conversion block may include an MDCT algorithm, a QMF, an FFT algorithm, Wavelet analysis or a filterbank such as a critically sampled filterbank having a certain number of filterbank channels, where the subband signals in this filterbank may be real valued signals or complex valued signals. The output of the spectral conversion block 214 is encoded using a spectral audio encoder 216, which may include processing blocks as known from the
25 AAC coding scheme.

30 The lower encoding branch 208 comprises a source model analyzer such as LPC 212, which outputs two kinds of signals. One signal is an LPC information signal which is used for controlling the filter characteristic of an LPC synthesis filter. This LPC information is transmitted to a decoder. The other LPC stage 212 output signal is an excitation signal or an LPC-domain signal, which is input into an excitation encoder 210. The excitation encoder 210 may come from any source-filter model encoder such as a CELP encoder, an ACELP encoder or any other encoder which processes a LPC domain signal.

35 Another excitation encoder implementation may be a transform coding of the excitation signal. In such an embodiment, the excitation signal is not encoded using an ACELP codebook mechanism, but the excitation signal is converted into a spectral representation and the spectral representation values such as subband signals in case of a filterbank or frequency coefficients in case of a transform such as an FFT are encoded to obtain a data compression.
40 An implementation of this kind of excitation encoder is the TCX coding mode known from AMR-WB+.

The decision in the decision stage 204 may be signal-adaptive so that the decision stage 204 performs a music/speech discrimination and controls the switch 202 in such a way that music signals are input into the upper branch 206, and speech signals are input into the lower branch 208. In one embodiment, the decision stage 204 feeds its decision information into an output
 5 bit stream, so that a decoder may use this decision information in order to perform the correct decoding operations.

Such a decoder is illustrated in Fig. 5. After transmission, the signal output by the spectral audio encoder 216 is input into a spectral audio decoder 218. The output of the spectral audio
 10 decoder 218 is input into a time-domain converter 220. The output of the excitation encoder 210 of Fig. 4 is input into an excitation decoder 222 which outputs an LPC-domain signal. The LPC-domain signal is input into an LPC synthesis stage 224, which receives, as a further input, the LPC information generated by the corresponding LPC analysis stage 212. The output of the time-domain converter 220 and/or the output of the LPC synthesis stage 224 are
 15 input into a switch 226. The switch 226 is controlled via a switch control signal which was, for example, generated by the decision stage 204, or which was externally provided such as by a creator of the original mono signal, stereo signal or multi-channel signal.

The output of the switch 226 is a complete mono signal which is subsequently input into a
 20 common post-processing stage 228, which may perform a joint stereo processing or a bandwidth extension processing etc. Alternatively, the output of the switch may also be a stereo signal or a multi-channel signal. It is a stereo signal, when the preprocessing includes a channel reduction to two channels. It may even be a multi-channel signal, when a channel reduction to three channels or no channel reduction at all but only a spectral band replication is
 25 performed.

Depending on the specific functionality of the common post-processing stage, a mono signal, a stereo signal or a multi-channel signal is output which has, when the common post-processing stage 228 performs a bandwidth extension operation, a larger bandwidth than the
 30 signal input into block 228.

In one embodiment, the switch 226 switches between the two decoding branches 218, 220 and 222, 224. In a further embodiment, there may be additional decoding branches such as a third decoding branch or even a fourth decoding branch or even more decoding branches. In an
 35 embodiment with three decoding branches, the third decoding branch may be similar to the second decoding branch, but includes an excitation decoder different from the excitation decoder 222 in the second branch 222, 224. In such an embodiment, the second branch comprises the LPC stage 224 and a codebook based excitation decoder such as in ACELP, and the third branch comprises an LPC stage and an excitation decoder operating on a spectral
 40 representation of the LPC stage 224 output signal.

In another embodiment, the common preprocessing stage comprises a surround/joint stereo block which generates, as an output, joint stereo parameters and a mono output signal, which

is generated by downmixing the input signal which is a signal having two or more channels. Generally, the signal at the output of block may also be a signal having more channels, but due to the downmixing operation, the number of channels at the output of block will be smaller than the number of channels input into block. In this embodiment, the frequency encoding

5 branch comprises a spectral conversion stage and a subsequently connected quantizing/coding stage. The quantizing/coding stage may include any of the functionalities as known from modern frequency-domain encoders such as the AAC encoder. Furthermore, the quantization operation in the quantizing/coding stage may be controlled via a psychoacoustic module which generates psychoacoustic information such as a psychoacoustic masking threshold over the

10 frequency, where this information is input into the stage. Preferably, the spectral conversion is done using an MDCT operation which, even more preferably, is the time-warped MDCT operation, where the strength or, generally, the warping strength may be controlled between zero and a high warping strength. In a zero warping strength, the MDCT operation is a straight-forward MDCT operation known in the art. The LPC-domain encoder may include an

15 ACELP core calculating a pitch gain, a pitch lag and/or codebook information such as a codebook index and a code gain.

Although some of the figures illustrate block diagrams of an apparatus, it is noted that these figures, at the same time, illustrate a method, wherein the block functionalities correspond to

20 the method steps.

Embodiments of the invention were described above on the basis of an audio input signal comprising different segments or frames, the different segments or frames being associated with speech information or music information. The invention is not limited to such

25 embodiments, rather, the approach for classifying different segments of a signal comprising segments of at least a first type and a second type can also be applied to audio signals comprising three or more different segment types, each of which is desired to be encoded by different encoding schemes. Examples for such segment types are:

- 30 - Stationary/non-stationary segments may be useful for using different filter-banks, windows or coding adaptation. For example a transient should be coded with a fine time resolution filter-bank while a pure sinusoid should be coded by a fine frequency resolution filter-bank.
- Voiced/unvoiced: voiced segments are well handled by speech coder like CELP but for
- 35 - unvoiced segments too much bits are wasted. The parametric coding will be more efficient.
- Silence/active: silence can be coded with fewer bits than active segments.
- Harmonic/non-harmonic: It will be beneficial to use for harmonic segments coding using a linear prediction in the frequency domain.

40 Also, the invention is not limited to the field of audio techniques, rather, the above-described approach for classifying a signal may be applied to other kinds of signals, like video signals or

data signals wherein these respective signals include segments of different types which require different processing, like for example:

5 The present invention may be adapted for all real time applications which need a segmentation of a time signal. For instance, a face detection from a surveillance video camera may be based on a classifier which determine for each pixel of a frame (here a frame corresponds to a picture taken at a time n) if it belongs to the face of a person or not. The classification (i.e., the face segmentation) should be done for each single frames of the video stream. However, using the present invention, the segmentation of the present frame can take into account the past
10 successive frames for getting a better segmentation accuracy taking the advantage that the successive pictures are strongly correlated. Two classifiers can be then applied. One considering only the present frame and another considering a set of frames including present and past frames The last classifier can integrate the set of frames and determine region of probability for the face position. The classifier decision done only on the present frame, will
15 then be compare to the probability regions. The decision may be then validated or modified.

Embodiments of the invention use the switch for switching between branches so that only one branch receives a signal to be processed and the other branch does not receive the signal. In an alternative embodiment, however, the switch may also be arranged after the processing stages
20 or branches, e.g. the audio encoder and the speech encoder, so that both branches process the same signal in parallel. The signal output by one of these branches is selected to be output, e.g. to be written into an output bitstream.

While embodiments of the invention were described on the basis of digital signals, the segments of which were determined by a predefined number of samples obtained at specific
25 sampling rate, the invention is not limited to such signals, rather, it is also applicable to analog signals in which the segment would then be determined by a specific frequency range or time period of the analog signal. In addition, embodiments of the invention were described in combination with encoders including the discriminator. It is noted that, basically, the approach
30 in accordance with embodiments of the invention for classifying signals may also be applied to decoders receiving an encoded signal for which different encoding schemes can be classified thereby allowing the encoded signal to be provided to an appropriate decoder.

Depending on certain implementation requirements of the inventive methods, the inventive
35 methods may be implemented in hardware or in software. The implementation may be performed using a digital storage medium, in particular, a disc, a DVD or a CD having electronically-readable control signals stored thereon, which co-operate with programmable computer systems such that the inventive methods are performed. Generally, the present invention is therefore a computer program product with a program code stored on a machine-
40 readable carrier, the program code being operated for performing the inventive methods when the computer program product runs on a computer. In other words, the inventive methods are, therefore, a computer program having a program code for performing at least one of the inventive methods when the computer program runs on a computer.

The above described embodiments are merely illustrative for the principles of the present invention. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to
5 be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.

In the above embodiments the signal is described as comprising a plurality of frames, wherein a current frame is evaluated for a switching decision. It is noted that the current segment of the
10 signal which is evaluated for a switching decision may be one frame, however, the invention is not limited to such embodiments. Rather, a segment of the signal may also comprise a plurality, i.e. two or more, frames.

Further, in the above described embodiments both the short-term classifier and the long-term classifier used the same short-term feature(s). This approach may be used for different reasons, like the need to compute the short-term features only once and to exploit same by the two classifiers in different ways which will reduce the complexity of the system, as e.g. the short-term feature may be calculated by one of the short-term or long-term classifiers and provided to the other classifier. Also, the comparison between short-term and long-term
15 20 classifier results may be more relevant as the contribution of the present frame in the long-term classification result is more easily deduced by comparing it with the short-term classification result since the two classifiers share common features.

The invention is, however, not restricted to such an approach and the long-term classifier is
25 not restricted to use the same short-term feature(s) as the short-term classifier, i.e. both the short-term classifier and the long-term classifier may calculate their respective short-term feature(s) which are different from each other.

While embodiments described above mentioned the use of PLPCCs as short-term feature, it is
30 noted that other features may be considered, e.g. the variability of the PLPCCs.

Claims

1. A method for classifying different segments of a signal, the signal comprising segments of at least a first type and a second type, the method comprising:

5

short-term classifying (150) the signal on the basis of at least one short-term feature extracted from the signal and delivering a short-term classification result (152);

10

long-term (154) the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal and delivering a long-term classification result (156); and

15

combining (158) the short-term classification result (152) and the long-term analysis classification (156) to provide an output signal (160) indicating whether a segment of the signal is of the first type or of the second type.

2. The method of claim 1, wherein the step of combining comprises providing the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).

20

3. The method of claim 1 or 2, wherein

the at least one short-term feature is obtained by analyzing a current segment of the signal which is to be classified;

25

the at least one long-term feature is obtained by analyzing the current segment of the signal and one or more preceding segments of the signal; and

30

the output signal (160) indicates whether the current segment is of the first type or of the second type.

4. The method of one of claims 1 to 3, wherein

35

the at least one short-term feature is obtained by analyzing an analysis window (168) of a first length and a first analysis method; and

40

the at least one long-term feature is obtained by analyzing an analysis window (162) of a second length and second analysis method, the first length being shorter than the second length, and the first and second analysis methods being different.

5. The method of claim 4, wherein the first length spans a current segment of the signal, the second length spans the current segment of the signal and one or more preceding

segments of the signal, and the first and second lengths comprise an additional period (164) covering an analysis period.

- 5 6. The method of one of claims 1 to 5, wherein combining (158) the short-term classification result (152) and the long-term classification result (156) comprises a hysteresis decision on the basis of a combined result, wherein the combined result comprises the short-term classification result (152) and the long-term classification result (156), each weighted by a predefined weighting factor.
- 10 7. The method of one of claims 1 to 6, wherein the signal is a digital signal and a segment of the signal comprises as predefined number of samples obtained at a specific sampling rate.
- 15 8. The method of one of claims 1 to 7, wherein
the signal comprises speech and music segments;
the at least one short-term feature comprises the PLPCCs parameters; and
20 the at least one long-term feature comprises pitch characteristic information.
- 25 9. The method of one of claims 1 to 8, wherein the short-term feature used for short-term classification and the short-term feature used for long-term classification are the same or different.
- 30 10. A method for processing a signal comprising segments of at least a first type and a second type, the method comprising:
classifying (116) a segment of the signal in accordance with the method of one of claims 1 to 9;
dependent on the output signal (160) provided by the classifying step (116), processing (102, 206; 106, 208) the segment in accordance with a first process or a second process; and
35 outputting the processed segment.
- 40 11. The method of claim 10, wherein
the signal comprises speech and music segments;
the segment is processed by a speech encoder (102) when the output signal (160) indicates that the segment is a speech segment; and

the segment is processed by a music encoder (106) when the output signal (160) indicates that the segment is a music segment.

5 12. The method of claim 11, further comprising:

combining (108) the encoded segment and information from the output signal (160) indicating the type of the segment.

10 13. A computer program for performing, when running on a computer, the method of one of claims 1 to 12.

14. A discriminator, comprising:

15 a short-term classifier (150) configured to receive a signal and to provide a short-term classification result (152) of the signal on the basis of at least one short-term feature extracted from the signal, the signal comprising segments of at least a first type and a second type;

20 a long-term classifier (154) configured to receive a signal and to provide a long-term classification result (156) of the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal; and

25 a decision circuit (158) configured to combine the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether a segment of the signal is of the first type or is of the second type.

30 15. The discriminator of claim 14, wherein the decision circuit (158) configured to provide the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).

35 16. The discriminator of claim 14 or 15, wherein the signal comprises speech and music segments, and wherein the output signal (160) indicates whether a segment of a signal is speech segment or a music segment.

17. A signal processing apparatus, comprising:

a input (110) configured to receive a signal to be processed, wherein the signal comprises segments of at least a first type and of a second type;

40 a first processing stage (102; 206), configured to process segments of the first type;

a second processing stage (104; 208) configured to process segments of the second type;

a discriminator (116; 204) of one of claims 14 to 16 coupled to the input; and

a switching device (112; 202) coupled between the input and the first and second processing stages and configured to apply the signal from the input (110) to one of the first and second processing stages dependent on the output signal (160) from the discriminator (116).

18. An audio encoder, comprising:

a signal processing apparatus of claim 17,

wherein the signal is an audio signal comprising speech and music segments, and

wherein the first processing stage comprises a speech encoder (102) and the second processing stage comprises a music encoder (106).

Method and Discriminator for Classifying
Different Segments of a Signal

5

Abstract

For classifying different segments of a signal which comprises segments of at least a first type and second type, e.g. audio and speech segments, the signal is short-term classified (150) on the basis of the at least one short-term feature extracted from the signal and a short-term
10 classification result (152) is delivered. The signal is also long-term classified (154) on the basis of the at least one short-term feature and at least one long-term feature extracted from the signal and a long-term classification result (156) is delivered. The short-term classification result (152) and the long-term classification result (156) are combined (158) to provide an
15 output signal (160) indicating whether a segment of the signal is of the first type or of the second type.

Fig. 1

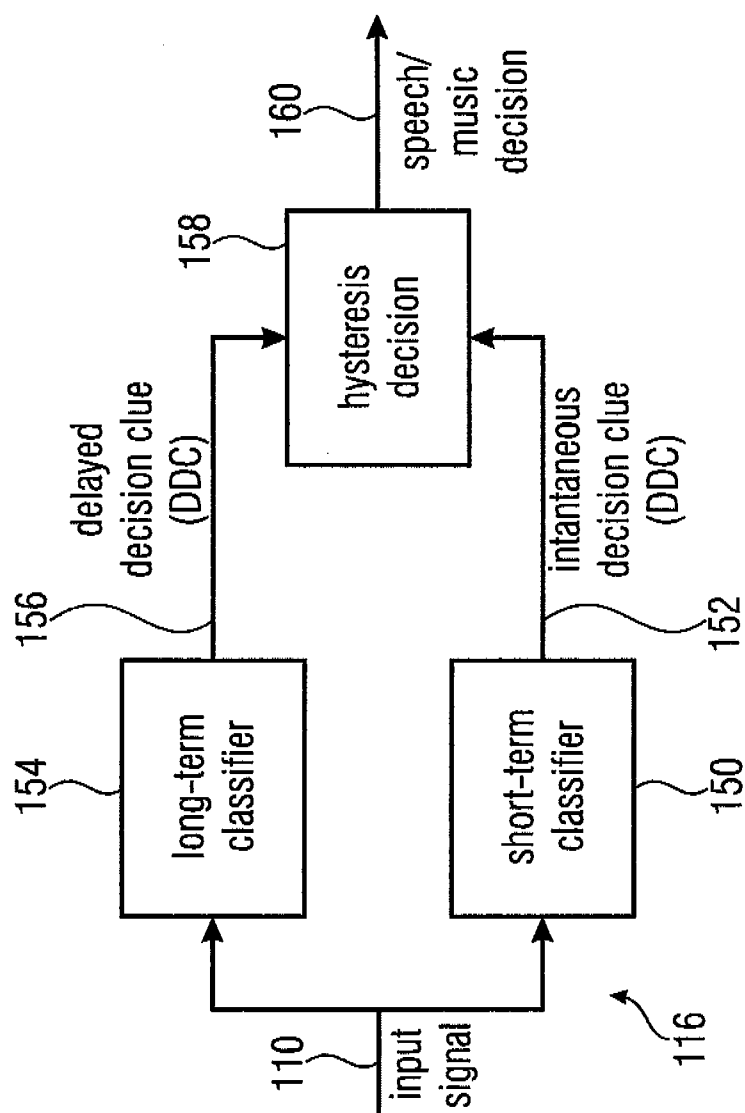


FIGURE 1

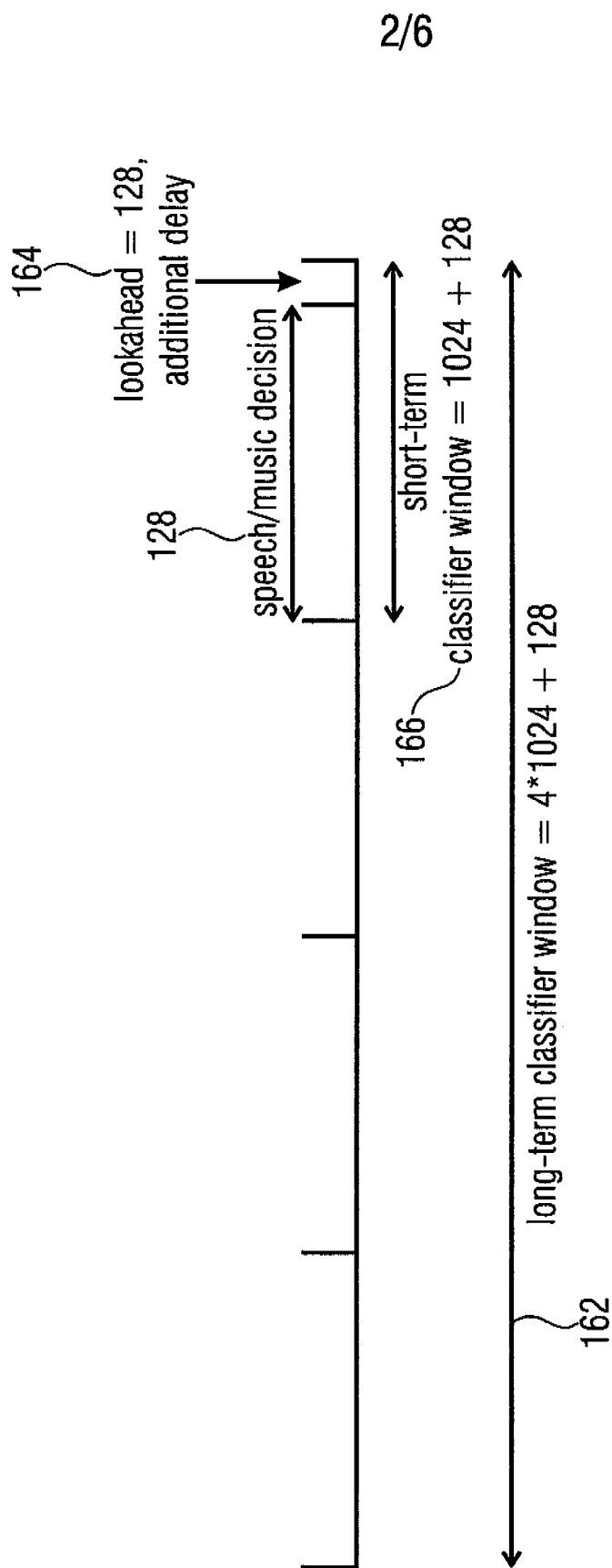


FIGURE 2

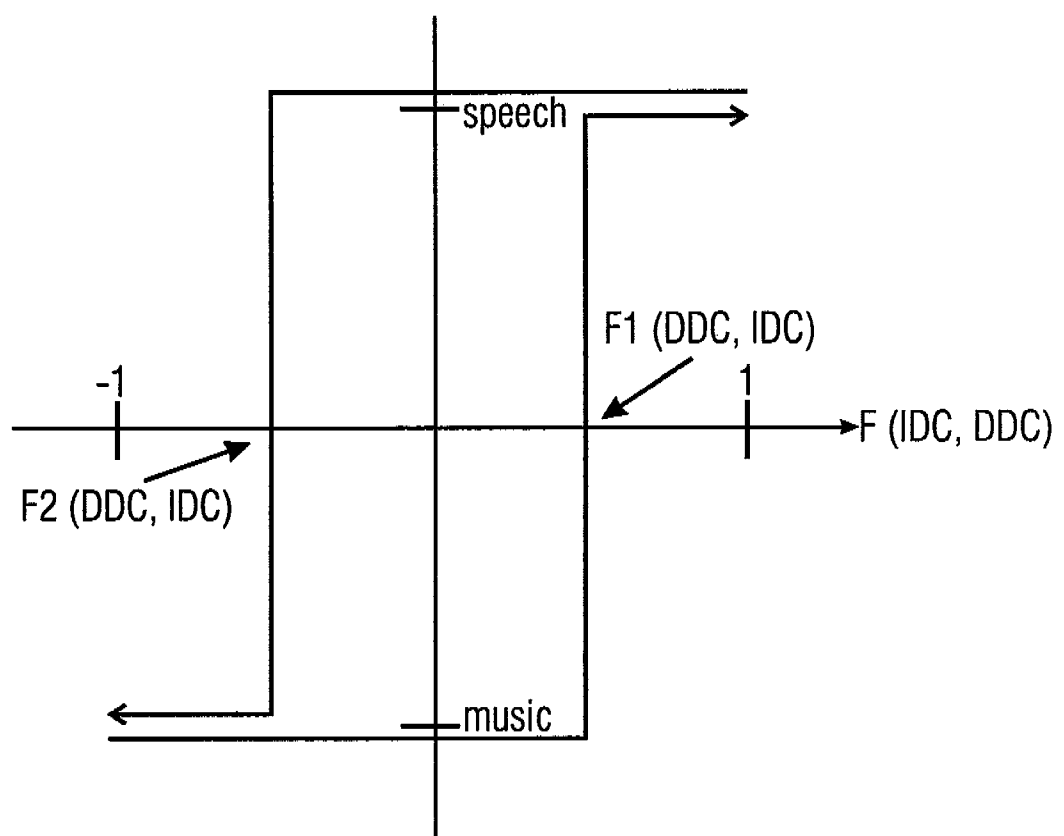


FIGURE 3

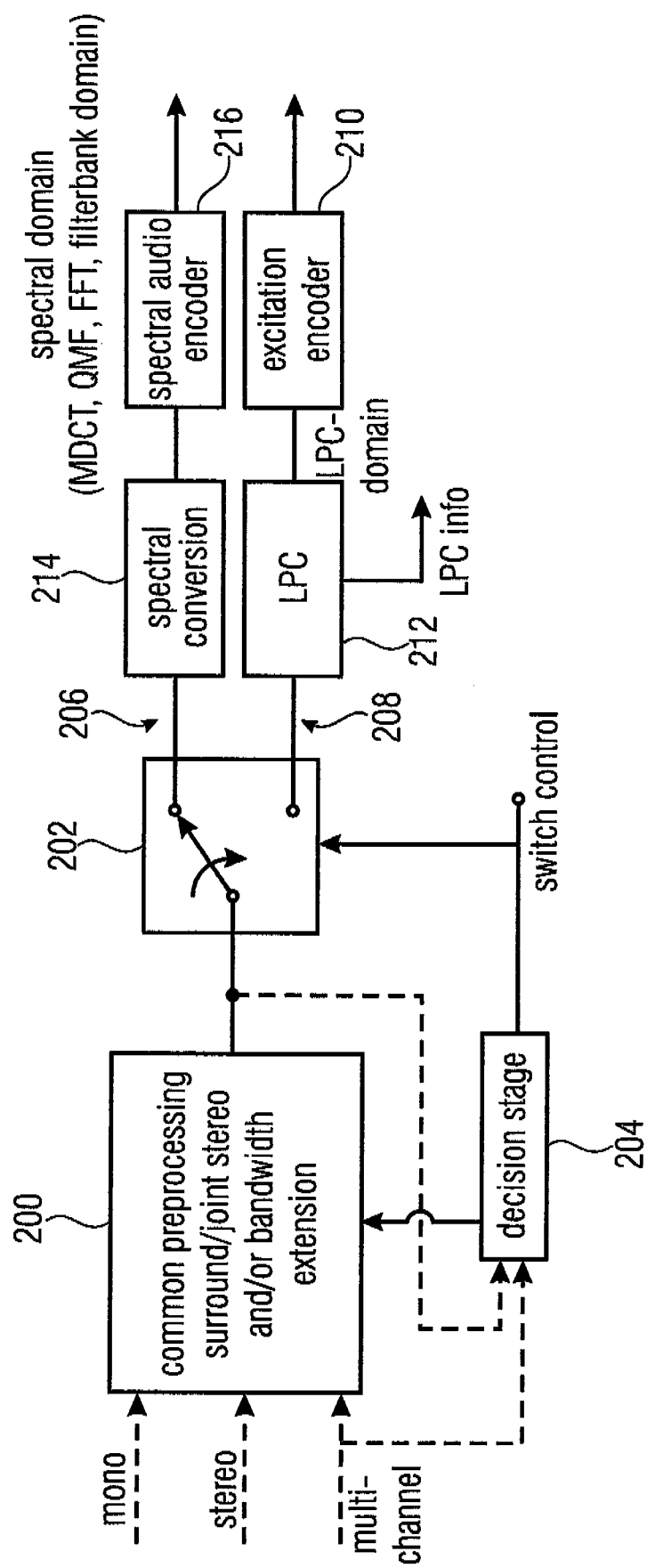


FIGURE 4
(ENCODER)

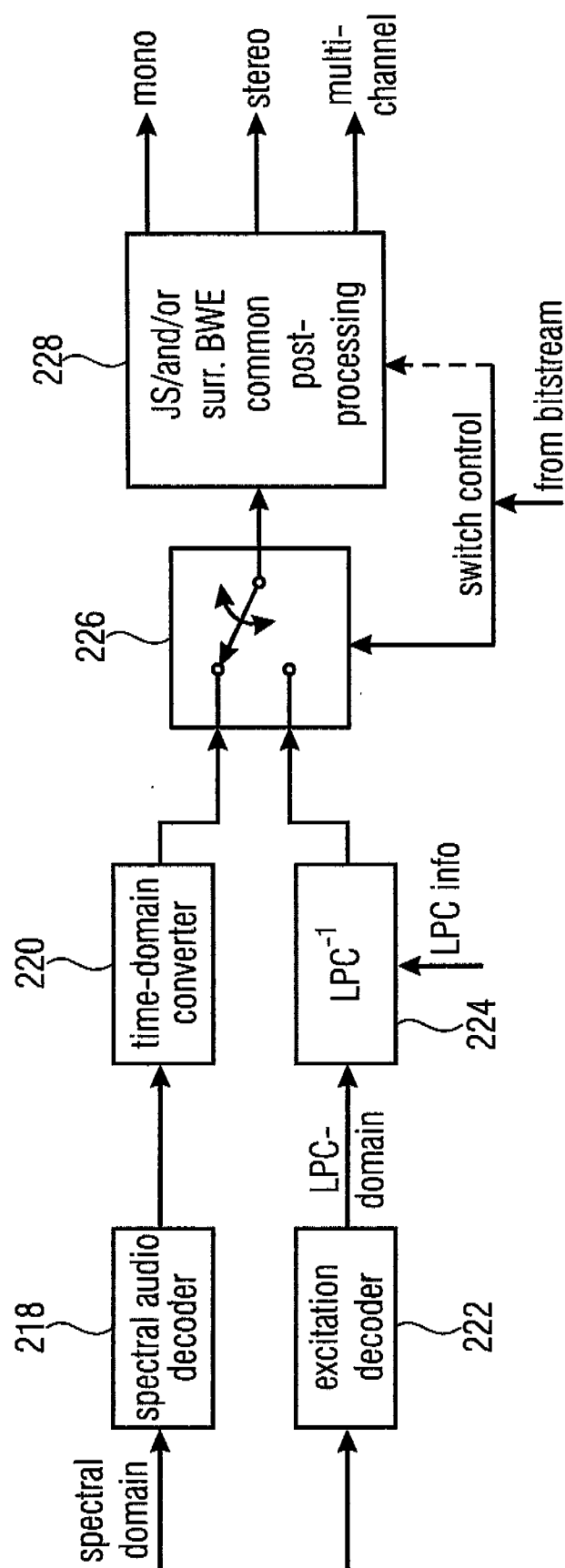


FIGURE 5
(DECODER)

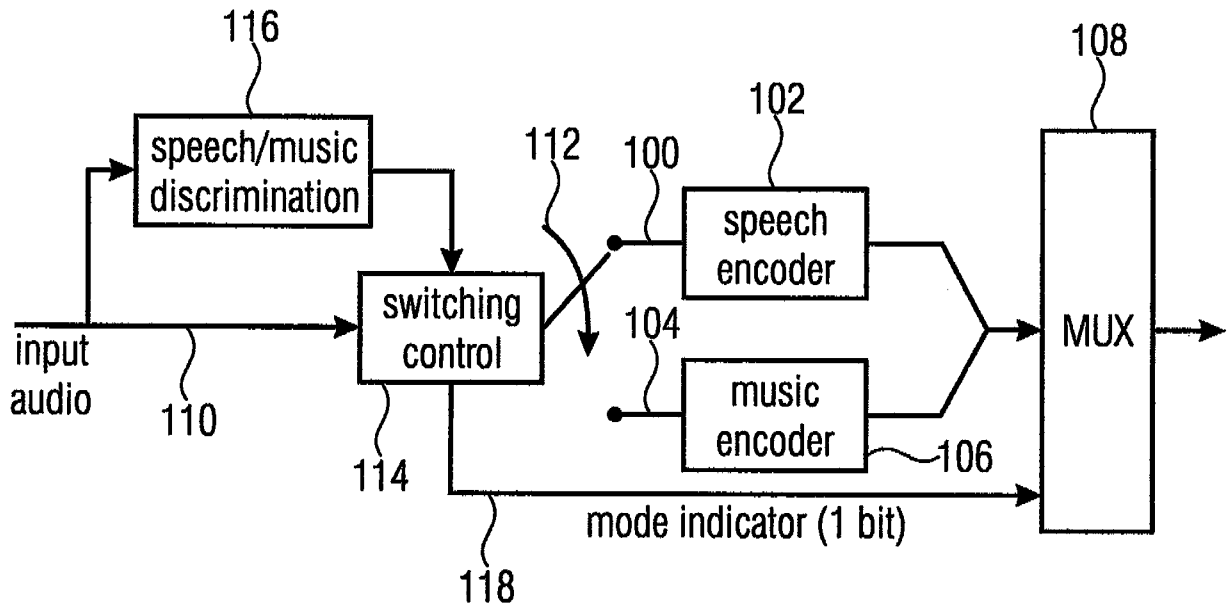


FIGURE 6

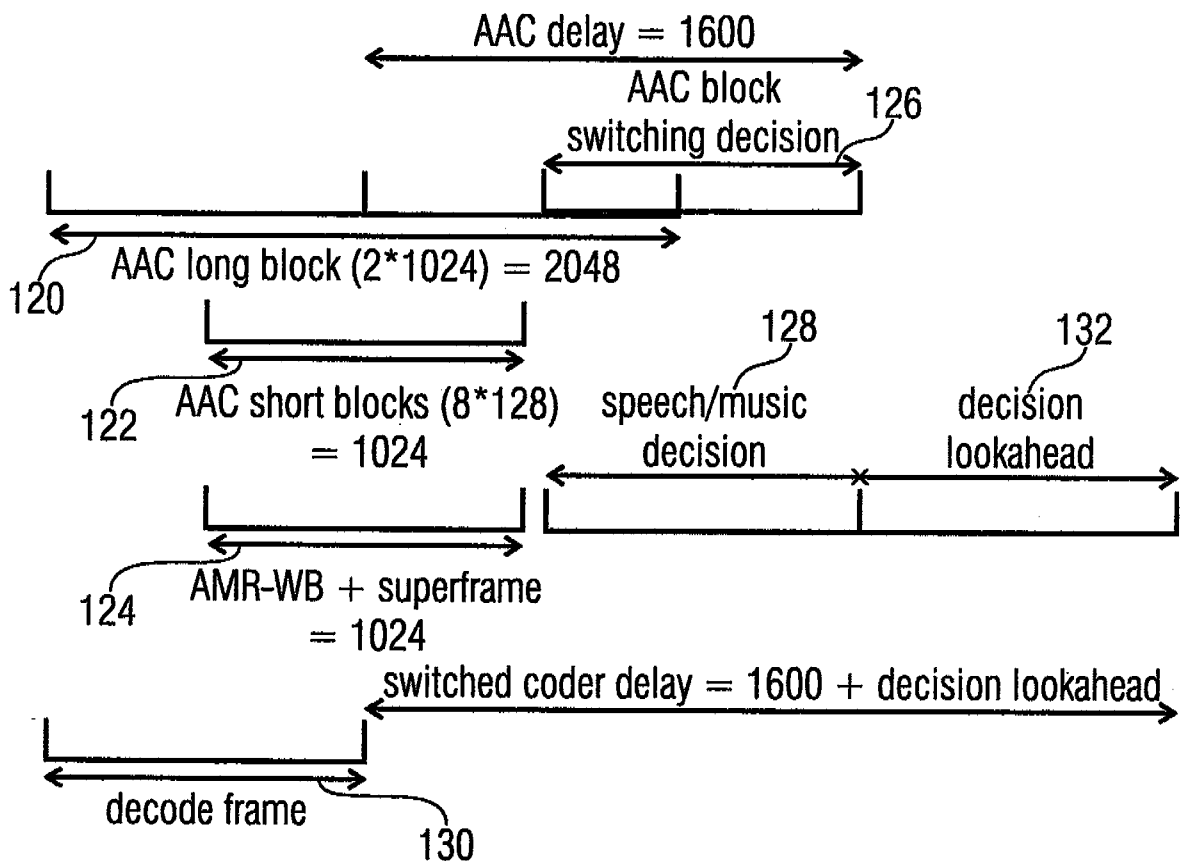


FIGURE 7

FIGURE FOR ABSTRACT

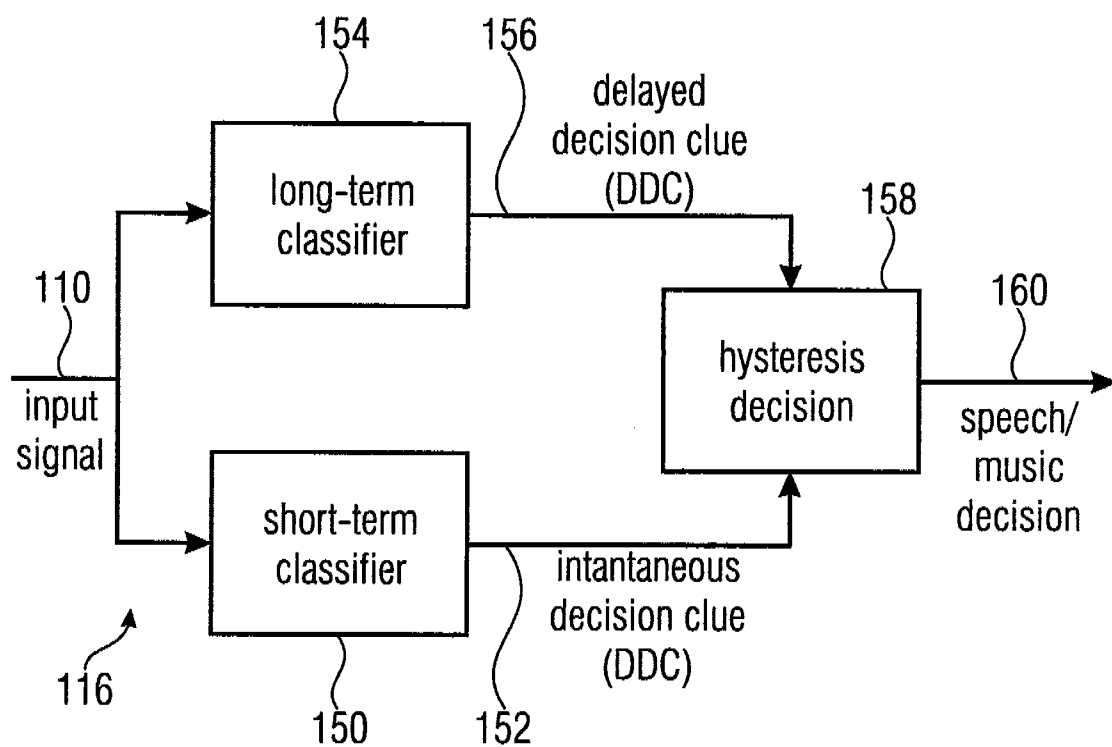


FIGURE 1

REIVINDICACIONES

1. Un método para clasificar distintos segmentos de una señal de audio, en el cual la señal de audio comprende segmentos de voz y música, y el método
5 comprende: clasificar en un corto plazo (150) la señal de audio sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal de audio para determinar si un segmento actual de la señal de audio es un segmento de voz o un segmento de música, y entregar un resultado de clasificación de corto plazo (152) que indica que el segmento actual de la señal de audio es un segmento de voz o un
10 segmento de música; clasificar en un largo plazo (154) la señal de audio sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídas de la señal de audio para determina si un segmento actual de la señal de audio es un segmento de voz o un segmento de música, y entregar un resultado de clasificación de largo plazo (156) que indica que
15 el segmento actual de la señal de audio es un segmento de voz o un segmento de música; y combinar (158) el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si el segmento actual de la señal de audio es un segmento de voz o un segmento de música.

20

2. El método de acuerdo con la reivindicación 1, en el cual la etapa de combinar comprende proveer la señal de salida sobre la base de una comparación del resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156)

25

3. El método de acuerdo con cualquiera de las reivindicaciones 1 o 2, en el cual se obtiene por lo menos un rasgo distintivo de corto plazo analizando el segmento actual de la señal de audio a ser clasificada; y se obtiene por lo menos un rasgo distintivo de largo plazo analizando el segmento actual de la señal de audio a
30 ser clasificada y uno o más segmentos anteriores de la señal de audio.

4. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el cual el por lo menos un rasgo distintivo de corto plazo se obtiene analizando una ventana de análisis (168) que tiene una primera longitud y usando un primer método de análisis; y la por lo menos un rasgo distintivo de largo plazo se obtiene analizando
5 una ventana de análisis (162) que tiene una segunda longitud y usando un segundo método de análisis, siendo la primera longitud más corta que la segunda longitud, y siendo los métodos de análisis primera y segunda diferentes.

5. El método de acuerdo con la reivindicación 4, en el cual la primera
10 longitud se extiende a lo largo del segmento actual de la señal de audio, la segunda longitud se extiende a lo largo del segmento actual de la señal de audio y uno o mas segmentos anteriores de la señal de audio, y las longitudes primera y segunda comprenden un periodo adicional (164) que cubre un periodo de análisis.

15 6. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el cual combinar (158) el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) comprende una decisión de histéresis sobre la base de un resultado combinado, en el cual el resultado combinado comprende el resultado de la clasificación de corto plazo (152) y el
20 resultado de la clasificación de largo plazo (156), cada uno ponderado por un factor de ponderación predeterminado.

7. El método de acuerdo con cualquiera de las reivindicaciones 1 a 6, en el cual la señal de audio es una señal digital y un segmento de la señal de audio
25 comprende un número predefinido de muestras obtenido a una tasa de muestreo específico.

8. El método de acuerdo con cualquiera de las reivindicaciones 1 a 7, en el cual el por lo menos un rasgo distintivo de corto plazo comprende parámetros
30 PLPCC; y el por lo menos un rasgo distintivo de largo plazo comprende información característica de frecuencia fundamental.

9. El método de acuerdo con cualquiera de las reivindicaciones 1 a 8, en el cual el rasgo distintivo de corto plazo usados para la clasificación de corto plazo y el rasgo distintivo de corto plazo usados para la clasificación de largo plazo son las mismas o diferentes.

5

10. Un método para procesar una señal de audio que comprende segmentos de voz y música, en el cual el método comprende: clasificar (116) un segmento actual de la señal de audio con el método de acuerdo con cualquiera de las reivindicaciones 1 a 9; procesar (102; 206; 106; 208) el segmento actual de acuerdo con un primer proceso o un segundo proceso, dependiendo de la señal de salida (160) entregada por la etapa de clasificar (116); y emitir el segmento procesado.

11. El método de acuerdo con la reivindicación 10, en el cual se procesa el segmento con un codificador de voz (102) cuando la señal de salida (160) indica que el segmento es un segmento de voz; y se procesa el segmento con un codificador de música (106) cuando la señal de salida (160) indica que el segmento es un segmento de música.

12. El método de acuerdo con la reivindicación 11, que comprende además: combinar (108) el segmento codificado e información de la señal de salida (160) que indica el tipo del segmento.

13. Un programa de computación que, cuando se ejecuta en una computadora, lleva a cabo el método de acuerdo con cualquiera de las reivindicaciones 1 a 12.

14. Un discriminador que comprende: un clasificador de corto plazo (150) configurado para recibir una señal de audio y para determinar si un segmento actual de la señal de audio es un segmento de voz o un segmento de música, y proveer un resultado de clasificación de corto plazo (152) de la señal de audio sobre la base de por lo menos un rasgo distintivo de corto plazo extraída de la señal de audio, el

5 resultado de clasificación de corto plazo (152) que indica que el segmento actual de la señal de audio es un segmento de voz o un segmento de música, en el cual la señal de audio comprende segmentos de voz y música; un clasificador de largo plazo (154) configurado para recibir una señal de audio y para determinar si un segmento
10 actual de la señal de audio es un segmento de voz o un segmento de música, y proveer un resultado de clasificación de largo plazo (156) de la señal de audio sobre la base de por lo menos un rasgo distintivo de corto plazo y por lo menos un rasgo distintivo de largo plazo extraídos de la señal de audio, el resultado de la clasificación de largo plazo (156) que indica que el segmento actual de la señal de audio es un
15 segmento de voz o un segmento de música; y un circuito de decisión (158) configurado para combinar el resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156) para proveer una señal de salida (160) que indica si el segmento actual de la señal de audio es un segmento de voz o un segmento de música.

15

15. El discriminador de acuerdo con la reivindicación 14, en el cual el circuito de decisión (158) está configurado para proveer la señal de salida sobre la base de una comparación del resultado de la clasificación de corto plazo (152) con el resultado de la clasificación de largo plazo (156).

20

25 16. Un aparato procesador de una señal de audio, que comprende: una entrada (110) configurada para recibir una señal de audio a ser procesada, en el cual la señal de audio comprende segmentos de voz y música; un primer elemento procesador (102; 206) configurado para procesar segmentos de voz; un segundo
30 elemento procesador (104; 208) configurado para procesar segmentos de música; un discriminador (116; 204) de acuerdo con la reivindicación 14 o 15 acoplado a la entrada; y un dispositivo conmutador (112; 202) acoplado entre la entrada y los elementos procesadores primero y segundo y configurado para aplicar la señal de audio proveniente de la entrada (110) a uno de los elementos procesadores primero y segundo dependiendo de la señal de salida (160) proveniente del discriminador (116).

17. Un codificador de audio, que comprende: un aparato procesador de una señal de audio de acuerdo con la reivindicación 16, en el cual el primer elemento procesador comprende un codificador de voz (102) y el segundo elemento procesador comprende un codificador de música (106).

Claims

- as attached to IPER – clean copy -

1. A method for classifying different segments of an audio signal, the audio signal comprising speech and music segments, the method comprising:
 - short-term classifying (150) the audio signal on the basis of at least one short-term feature extracted from the audio signal to determine whether a current segment of the audio signal is a speech segment or a music segment, and delivering a short-term classification result (152) indicating that the current segment of the audio signal is a speech segment or a music segment;
 - long-term classifying (154) the audio signal on the basis of at least one short-term feature and at least one long-term feature extracted from the audio signal to determine whether a current segment of the audio signal is a speech segment or a music segment, and delivering a long-term classification result (156) indicating that the current segment of the audio signal is a speech segment or a music segment; and
 - combining (158) the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether the current segment of the audio signal is a speech segment or a music segment.
2. The method of claim 1, wherein the step of combining comprises providing the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).
3. The method of claim 1 or 2, wherein
 - the at least one short-term feature is obtained by analyzing a current segment of the audio signal which is to be classified; and
 - the at least one long-term feature is obtained by analyzing the current segment of the audio signal and one or more preceding segments of the audio signal.
4. The method of one of claims 1 to 3, wherein
 - the at least one short-term feature is obtained by analyzing an analysis window (168) of a first length and a first analysis method; and
 - the at least one long-term feature is obtained by analyzing an analysis window (162) of a second length and second analysis method, the first length being shorter than the second length, and the first and second analysis methods being different.

5. The method of claim 4, wherein the first length spans a current segment of the audio signal, the second length spans the current segment of the audio signal and one or more preceding segments of the audio signal, and the first and second lengths comprise an additional period (164) covering an analysis period.
6. The method of one of claims 1 to 5, wherein combining (158) the short-term classification result (152) and the long-term classification result (156) comprises a hysteresis decision on the basis of a combined result, wherein the combined result comprises the short-term classification result (152) and the long-term classification result (156), each weighted by a predefined weighting factor.
7. The method of one of claims 1 to 6, wherein the audio signal is a digital signal and a segment of the audio signal comprises as predefined number of samples obtained at a specific sampling rate.
8. The method of one of claims 1 to 7, wherein
 - the at least one short-term feature comprises PLPCCs parameters; and
 - the at least one long-term feature comprises pitch characteristic information.
9. The method of one of claims 1 to 8, wherein the short-term feature used for short-term classification and the short-term feature used for long-term classification are the same or different.
10. A method for processing an audio signal comprising speech and music segments, the method comprising:
 - classifying (116) a current segment of the audio signal in accordance with the method of one of claims 1 to 9;
 - dependent on the output signal (160) provided by the classifying step (116), processing (102, 206; 106, 208) the current segment in accordance with a first process or a second process; and
 - outputting the processed segment.
11. The method of claim 10, wherein
 - the segment is processed by a speech encoder (102) when the output signal (160) indicates that the segment is a speech segment; and

the segment is processed by a music encoder (106) when the output signal (160) indicates that the segment is a music segment.

12. The method of claim 11, further comprising:

combining (108) the encoded segment and information from the output signal (160) indicating the type of the segment.

13. A computer program for performing, when running on a computer, the method of one of claims 1 to 12.

14. A discriminator, comprising:

a short-term classifier (150) configured to receive an audio signal and to determine whether a current segment of the audio signal is a speech segment or a music segment, and to provide a short-term classification result (152) of the audio signal on the basis of at least one short-term feature extracted from the audio signal, the short-term classification result (152) indicating that the current segment of the audio signal is a speech segment or a music segment, the audio signal comprising speech and music segments;

a long-term classifier (154) configured to receive a audio signal and to determine whether a current segment of the audio signal is a speech segment or a music segment, and to provide a long-term classification result (156) of the audio signal on the basis of at least one short-term feature and at least one long-term feature extracted from the audio signal, the long-term classification result (156) indicating that the current segment of the audio signal is a speech segment or a music segment; and

a decision circuit (158) configured to combine the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether the current segment of the audio signal is a speech segment or a music segment.

15. The discriminator of claim 14, wherein the decision circuit (158) configured to provide the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).

16. An audio signal processing apparatus, comprising:

a input (110) configured to receive a audio signal to be processed, wherein the audio signal comprises speech and music segments;

a first processing stage (102; 206), configured to process speech segments;

a second processing stage (104; 208) configured to process music segments;

a discriminator (116; 204) of claim 14 or 15 coupled to the input; and

5

a switching device (112; 202) coupled between the input and the first and second processing stages and configured to apply the audio signal from the input (110) to one of the first and second processing stages dependent on the output signal (160) from the discriminator (116).

10

17. An audio encoder, comprising:

an audio signal processing apparatus of claim 16,

15

wherein the first processing stage comprises a speech encoder (102) and the second processing stage comprises a music encoder (106).

Claims

1. A method for classifying different segments of an audio signal, the audio signal comprising speech and music segments, the method comprising:
 - 5 short-term classifying (150) the audio signal on the basis of at least one short-term feature extracted from the audio signal to determine whether a current segment of the audio signal is a speech segment or a music segment, and delivering a short-term classification result (152) indicating that the current segment of the audio signal is a speech segment or a music segment;
 - 10
 - 15 long-term classifying (154) the audio signal on the basis of at least one short-term feature and at least one long-term feature extracted from the audio signal to determine whether a current segment of the audio signal is a speech segment or a music segment, and delivering a long-term classification result (156) indicating that the current segment of the audio signal is a speech segment or a music segment; and
 - 20
 - 25 combining (158) the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether the current segment of the audio signal is a speech segment or a music segment.
 - 30
2. The method of claim 1, wherein the step of combining comprises providing the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).
- 35
3. The method of claim 1 or 2, wherein

the at least one short-term feature is obtained by analyzing a current segment of the audio signal which is to be classified; and

5 the at least one long-term feature is obtained by analyzing the current segment of the audio signal and one or more preceding segments of the audio signal.

4. The method of one of claims 1 to 3, wherein

10

the at least one short-term feature is obtained by analyzing an analysis window (168) of a first length and a first analysis method; and

15

the at least one long-term feature is obtained by analyzing an analysis window (162) of a second length and second analysis method, the first length being shorter than the second length, and the first and second analysis methods being different.

20

5. The method of claim 4, wherein the first length spans a current segment of the audio signal, the second length spans the current segment of the audio signal and one or more preceding segments of the audio signal, and the first and second lengths comprise an additional period (164) covering an analysis period.

25

6. The method of one of claims 1 to 5, wherein combining (158) the short-term classification result (152) and the long-term classification result (156) comprises a hysteresis decision on the basis of a combined result, wherein the combined result comprises the short-term classification result (152) and the long-term classification result (156), each weighted by a predefined weighting factor.

30

35

7. The method of one of claims 1 to 6, wherein the audio signal is a digital signal and a segment of the audio

signal comprises as predefined number of samples obtained at a specific sampling rate.

- 5
8. The method of one of claims 1 to 7, wherein
- the at least one short-term feature comprises PLPCCs parameters; and
- the at least one long-term feature comprises pitch characteristic information.
- 10
9. The method of one of claims 1 to 8, wherein the short-term feature used for short-term classification and the short-term feature used for long-term classification are the same or different.
- 15
10. A method for processing an audio signal comprising speech and music segments, the method comprising:
- 20
- classifying (116) a current segment of the audio signal in accordance with the method of one of claims 1 to 9;
- dependent on the output signal (160) provided by the
- 25
- classifying step (116), processing (102, 206; 106, 208) the current segment in accordance with a first process or a second process; and
- outputting the processed segment.
- 30
11. The method of claim 10, wherein
- the segment is processed by a speech encoder (102) when the output signal (160) indicates that the
- 35
- segment is a speech segment; and

the segment is processed by a music encoder (106) when the output signal (160) indicates that the segment is a music segment.

- 5 12. The method of claim 11, further comprising:

combining (108) the encoded segment and information from the output signal (160) indicating the type of the segment.

10

13. A computer program for performing, when running on a computer, the method of one of claims 1 to 12:

14. A discriminator, comprising:

15

a short-term classifier (150) configured to receive an audio signal and to determine whether a current segment of the audio signal is a speech segment or a music segment, and to provide a short-term classification result (152) of the audio signal on the basis of at least one short-term feature extracted from the audio signal, the short-term classification result (152) indicating that the current segment of the audio signal is a speech segment or a music segment, the audio signal comprising speech and music segments;

20

25

a long-term classifier (154) configured to receive a audio signal and to determine whether a current segment of the audio signal is a speech segment or a music segment, and to provide a long-term classification result (156) of the audio signal on the basis of at least one short-term feature and at least one long-term feature extracted from the audio signal, the long-term classification result (156) indicating that the current segment of the audio signal is a speech segment or a music segment; and

30

35

5 a decision circuit (158) configured to combine the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether the current segment of the audio signal is a speech segment or a music segment.

10 15. The discriminator of claim 14, wherein the decision circuit (158) configured to provide the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).

15 16. An audio signal processing apparatus, comprising:

a input (110) configured to receive a audio signal to be processed, wherein the audio signal comprises speech and music segments;

20 a first processing stage (102; 206), configured to process speech segments;

a second processing stage (104; 208) configured to process music segments;

25 a discriminator (116; 204) of claim 14 or 15 coupled to the input; and

30 a switching device (112; 202) coupled between the input and the first and second processing stages and configured to apply the audio signal from the input (110) to one of the first and second processing stages dependent on the output signal (160) from the discriminator (116).

35 17. An audio encoder, comprising:

an audio signal processing apparatus of claim 16,

wherein the first processing stage comprises a speech encoder (102) and the second processing stage comprises a music encoder (106).

document D1 (the article to Tancerel et al.) or document D2 (US 2003/0101050 A1). For the following reasons, this conclusion is respectfully traversed.

The present application is concerned with a novel and inventive approach for classifying different segments of a signal, for example, music and speech in an input signal to provide for a switching signal to a switching element that supplies the input signal segments to the appropriate coders/decoders. This basic approach of discriminating between speech and music is well known in the art. However, as is outlined in the introductory portion of the present application, especially for real-time applications, conventional approaches are disadvantageous as the delay introduced by the discriminator for switching is undesirable (see pages 4 and 5 of the present application).

Therefore, it is the object underlying the present invention to provide for an improved approach for discriminating segments of a different type in a signal, while keeping any delay introduced by the discrimination low. In accordance with the present invention, this is achieved by providing a new approach for classifying different segments, which leaves the conventional approaches of only doing a "single" discrimination considering both short-term and long-term features of the signal. Rather as is clearly defined in the claim language, two different classifications are done, namely a first classification that operates on the basis of at least one short-term feature, the so-called short-term classification and a second classification that also operates on this short-term feature and, in addition, on a long-term feature, which is called the long-term classification. It is important to note that the claim language and also the specification clearly specify that two different classification processes are executed, which result in two different classification results, which are then combined to obtain an indication of the type of the segment of the signal currently analyzed.

Thus, in accordance with the present invention, short-term features are used both in a short-term classifier and in a long-term classifier so that the two classifiers exploit different statistics of the same feature, for example, the short-term classifier will extract only the instantaneous information, because it has access to only one set of features, while, on the other hand, the long-term classifier has access to several sets of features and may exploit more characteristics of the signal. The long-term classifier may receive the short-term feature directly from the short-term classifier.

By means of the inventive approach, the delay introduced is limited, while keeping a reliable decision, since the short-term classification result that may be used for switching is made more robust by providing, in addition, the long-term classification result so that only a minor delay is introduced at any time and switching occurs substantially while the current frame is processed, i.e. unnecessary delays are avoided by the inventive approach.

Turning now to the disclosure of the prior art documents cited in the International Search Report, after reviewing same, the applicant feels that same do not go beyond the conventional discriminator approaches already described in the introductory portion and considered disadvantageous as same also introduce undesired delays.

Actually, neither document D1 (the article to Tancerel et al.) nor document D2 (US 2003/0101050 A1) describe an approach allowing for a reduction of the delay by exploiting information about the quality of the signal (speech or music) obtained by a short-term feature analysis. In short, document D1 uses a VAD (voice activity detector, see Fig. 2) and document D2 uses a feature extractor 220 that may also extract short-term features (see. Fig. 2), however the VAD and feature extractor, respectively, are used only for determining an appropriate switching instant. In accordance with the teachings of documents D1 and D2 the VAD and feature extractor are not exploited for obtaining information about the quality of the signal (speech or music), which is in contradiction to the teachings of the present invention. As a consequence, neither of these documents can achieve a low delay classification. Note e.g. that Fig. 5b of document D2 explicitly shows that the short-term analysis does not help to reduce the overall delay.

Document D1 describes a combined speech and audio coding by discrimination and the conventional encoder structure is given in Fig. 1 showing the discriminator acting on a switch for selecting one of two encoders for an input signal dependent on whether the input signal is a speech signal or a music signal. The speech/music discrimination is described in further detail in chapter 2 on pages 154 and 155 and Fig. 2 thereof shows a high-level description of the speech/music discriminator. As can be seen, this discriminator has three blocks, namely the parameter examination block, the long-term statistics and normalization block and the speech/music identification block. As can be seen in Fig. 2, these three blocks are arranged in series and output the signal S/NS, i.e. the classifying signal indicating a speech/non-speech decision. Thus, what document D1

actually teaches is that the input signal is received at the discriminator and first of all parameters for doing the long-term statistics are extracted in the parameter examination block and the respective long-term statistics are done over a length of 500 ms (see section 2.2 of document D1) in order to generate a final feature vector on the basis of which the classification is done in the speech/music classification block. Thus, despite the fact that in the parameter estimation block, also the short-term energy of the input signal every 5 ms on frames of lengths of 15 ms is computed, the basic approach of the discriminator of document D1 is such that a single classification step is executed on the basis of the parameters determined, which are then applied to a long-term statistic module that operates over 500 ms that, in accordance with the teachings of document D1, is a good compromise between global discrimination efficiency and local switching accuracy. Thus, at best, document D1 describes a discriminator comprising a single classifier, namely a long-term classifier introducing a delay of 500 ms. Nothing in document D1 teaches or suggests that an additional "branch" should be provided in parallel to the long-term classifier that operates only on the basis of short-term features to provide a short-term classification result in a manner as is taught by the present invention. As to this aspect, it is noted that while document D1 shows in Fig. 2 thereof a discriminator having a parallel branch including the VAD decision block (voice activity detector block), this block does not qualify as a short-term classifier as defined by the present invention. Actually, this block is only used for constraining the switching decision in low-energy frames according to the VAD decision, so that the discrimination is effective only in these parts of the signal, as is outlined in section 2.4 of document D1. However, this does not suggest using a short-term classifier, the output of which is used for switching and is made more robust by combining it with the classification results of a long-term classifier. Actually, the VAD block outputs an activity flag that enables the output of the discrimination signal from the long-term discriminator dependent on whether the current part of the signal is a low energy frame or not.

Thus, when starting from document D1 as being one possible closest prior art, it is to be emphasized that nothing in this document suggests providing an additional classifier operating on short-term features only in a manner as discussed above and as defined in the claim language.

Basically, the same is true for document D2 that describes a real-time speech and music classifier. When looking at the disclosure of document D2, especially Fig. 2 thereof, the

classifier comprises a look-ahead buffer 210 that receives audio samples and forwards the samples to the feature extractor 220 that extracts long-term features and short-term features, which are both introduced into a classification module 230, which outputs classified frames. As is outlined in paragraph [0011], the feature extractor extracts a set of long-term and short-term features from each frame in the buffer, which are then provided to the classification module that first detects a potential switching time according to the short-term feature of the current coding frame and then classifies each frame according to the long-term features and determines whether to switch the operation mode of the coder for the classified frame at the potential switching time according to the predefined switch criterion. Thus, as can be seen from Fig. 2, by means of the classifier 111, a delay by buffering the input samples is introduced, which is undesired in accordance with the teachings of the present invention. Document D2 does not suggest providing a further classifier operating only on the short-term features without any delay so that, basically, this document, like document D1, does not go beyond the conventional approach of providing a delayed classification result that is determined on the basis of long-term features and short-term features derived from an audio sample. Document D2, like document D1, however, does not suggest using two parallel classifications, namely a short-term classification operating only on the basis of a short-term feature and a long-term classification operating both on the long-term and the short-term features, which are then combined in a manner as described above.

Thus, document D2, like document D1, does not suggest the inventive approach of using two classifiers for providing two classification results, which are then combined for determining an output signal indicating the type of a currently processed segment of a signal.

Therefore, contrary to the conclusion given in the Written Opinion attached to the International Search Report, the applicant feels that the prior art documents cited neither anticipate nor render obvious the subject matter of the present application so that this subject matter is clearly patentable over the available art.

3. In view of the above comments, it is kindly requested to issue a favorable International Preliminary Examination Report.

Should the Examiner still have concerns with regard to the allowability of the subject matter of the present application, it is requested to issue a Written Opinion or discuss this matter during a telephone conversation.

Respectfully submitted,



Tankred Zimmermann

Encl.:
New claims 1 to 18

Claims

1. A method for classifying different segments of a
signal, the signal comprising segments of at least a
5 first type and a second type, the method comprising:

short-term classifying (150) the signal on the basis
of at least one short-term feature extracted from the
signal and delivering a short-term classification
10 result (152);

long-term classifying (154) the signal on the basis of
at least one short-term feature and at least one long-
term feature extracted from the signal and delivering
15 a long-term classification result (156); and

combining (158) the short-term classification result
(152) and the long-term analysis classification (156)
to provide an output signal (160) indicating whether a
20 segment of the signal is of the first type or of the
second type.
2. The method of claim 1, wherein the step of combining
comprises providing the output signal on the basis of
25 a comparison of the short-term classification result
(152) to the long-term classification result (156).
3. The method of claim 1 or 2, wherein
30 the at least one short-term feature is obtained by
analyzing a current segment of the signal which is to
be classified;

the at least one long-term feature is obtained by
35 analyzing the current segment of the signal and one or
more preceding segments of the signal; and

the output signal (160) indicates whether the current segment is of the first type or of the second type.

4. The method of one of claims 1 to 3, wherein
5 the at least one short-term feature is obtained by analyzing an analysis window (168) of a first length and a first analysis method; and
10 the at least one long-term feature is obtained by analyzing an analysis window (162) of a second length and second analysis method, the first length being shorter than the second length, and the first and second analysis methods being different.
15
5. The method of claim 4, wherein the first length spans a current segment of the signal, the second length spans the current segment of the signal and one or more preceding segments of the signal, and the first and second lengths comprise an additional period (164) covering an analysis period.
20
6. The method of one of claims 1 to 5, wherein combining (158) the short-term classification result (152) and the long-term classification result (156) comprises a hysteresis decision on the basis of a combined result, wherein the combined result comprises the short-term classification result (152) and the long-term classification result (156), each weighted by a predefined weighting factor.
25
30
7. The method of one of claims 1 to 6, wherein the signal is a digital signal and a segment of the signal comprises as predefined number of samples obtained at a specific sampling rate.
35
8. The method of one of claims 1 to 7, wherein

the signal comprises speech and music segments;

the at least one short-term feature comprises the PLPCCs parameters; and

5

the at least one long-term feature comprises pitch characteristic information.

9. The method of one of claims 1 to 8, wherein the short-term feature used for short-term classification and the short-term feature used for long-term classification are the same or different.

10

10. A method for processing a signal comprising segments of at least a first type and a second type, the method comprising:

15

classifying (116) a segment of the signal in accordance with the method of one of claims 1 to 9;

20

dependent on the output signal (160) provided by the classifying step (116), processing (102, 206; 106, 208) the segment in accordance with a first process or a second process; and

25

outputting the processed segment.

11. The method of claim 10, wherein

30

the signal comprises speech and music segments;

the segment is processed by a speech encoder (102) when the output signal (160) indicates that the segment is a speech segment; and

35

the segment is processed by a music encoder (106) when the output signal (160) indicates that the segment is a music segment.

12. The method of claim 11, further comprising:

5 combining (108) the encoded segment and information from the output signal (160) indicating the type of the segment.

13. A computer program for performing, when running on a computer, the method of one of claims 1 to 12.

10

14. A discriminator, comprising:

15 a short-term classifier (150) configured to receive a signal and to provide a short-term classification result (152) of the signal on the basis of at least one short-term feature extracted from the signal, the signal comprising segments of at least a first type and a second type;

20 a long-term classifier (154) configured to receive a signal and to provide a long-term classification result (156) of the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal; and

25

30 a decision circuit (158) configured to combine the short-term classification result (152) and the long-term classification result (156) to provide an output signal (160) indicating whether a segment of the signal is of the first type or is of the second type.

15. The discriminator of claim 14, wherein the decision circuit (158) configured to provide the output signal on the basis of a comparison of the short-term classification result (152) to the long-term classification result (156).

35

16. The discriminator of claim 14 or 15, wherein the signal comprises speech and music segments, and wherein the output signal (160) indicates whether a segment of a signal is speech segment or a music segment.
- 5
17. A signal processing apparatus, comprising:
- 10 a input (110) configured to receive a signal to be processed, wherein the signal comprises segments of at least a first type and of a second type;
- a first processing stage (102; 206), configured to process segments of the first type;
- 15 a second processing stage (104; 208) configured to process segments of the second type;
- a discriminator (116; 204) of one of claims 14 to 16 coupled to the input; and
- 20 a switching device (112; 202) coupled between the input and the first and second processing stages and configured to apply the signal from the input (110) to one of the first and second processing stages dependent on the output signal (160) from the discriminator (116).
- 25
18. An audio encoder, comprising:
- 30 a signal processing apparatus of claim 17,
- wherein the signal is an audio signal comprising speech and music segments, and
- 35 wherein the first processing stage comprises a speech encoder (102) and the second processing stage comprises a music encoder (106).

PATENT COOPERATION TREATY

PCT/EP2009/004339

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT

25. SEP. 2009

To:

ZIMMERMANN, Tankred
SCHOPPE, ZIMMERMANN, STÖCKELER &
ZINKLER
Postfach 246
82043 Pullach bei München
ALLEMAGNE

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year) 15 September 2009 (15.09.2009)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference FH090608PCT	
International application No. PCT/EP2009/004339	International filing date (day/month/year) 16 June 2009 (16.06.2009)
International publication date (day/month/year) Not yet published	Priority date (day/month/year) 11 July 2008 (11.07.2008)
Applicant FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. et al.	

- By means of this Form, which replaces any previously issued notification concerning submission or transmittal of priority documents, the applicant is hereby notified of the date of receipt by the International Bureau of the priority document(s) relating to all earlier application(s) whose priority is claimed. Unless otherwise indicated by the letters "NR", in the right-hand column or by an asterisk appearing next to a date of receipt, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
- (If applicable)* The letters "NR" appearing in the right-hand column denote a priority document which, on the date of mailing of this Form, had not yet been received by the International Bureau under Rule 17.1(a) or (b). Where, under Rule 17.1(a), the priority document must be submitted by the applicant to the receiving Office or the International Bureau, but the applicant fails to submit the priority document within the applicable time limit under that Rule, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
- (If applicable)* An asterisk (*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b) (the priority document was received after the time limit prescribed in Rule 17.1(a) or the request to prepare and transmit the priority document was submitted to the receiving Office after the applicable time limit under Rule 17.1(b)). Even though the priority document was not furnished in compliance with Rule 17.1(a) or (b), the International Bureau will nevertheless transmit a copy of the document to the designated Offices, for their consideration. In case such a copy is not accepted by the designated Office as the priority document, Rule 17.1(c) provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

Priority date	Priority application No.	Country or regional Office or PCT receiving Office	Date of receipt of priority document
11 July 2008 (11.07.2008)	61/079,875	US	24 August 2009 (24.08.2009)

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Authorized officer

Yolaine Cussac

Facsimile No. +41 22 338 82 70

e-mail PT05.PCT@WIPO.INT

Telephone No. +41 22 338 74 05

Form PCT/IB/304 (October 2005)

1/D6Z3NFAD0

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

PCT

To: SCHOPPE, ZIMMERMANN, STÖCKELER & ZINKLER Attn. Zimmermann, Tankred Postfach 246 D-82043 Pullach bei München ALLEMAGNE	EINGETRAGEN 31 JULI 2009
---	-----------------------------

NOTIFICATION OF TRANSMITTAL OF
THE INTERNATIONAL SEARCH REPORT AND
THE WRITTEN OPINION OF THE INTERNATIONAL
SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference FH090608PCT	FOR FURTHER ACTION See paragraphs 1 and 4 below
International application No. PCT/EP2009/004339	International filing date (day/month/year) 16/06/2009
Applicant FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ...	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

Filing of amendments and statement under Article 19:

The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

When? The time limit for filing such amendments is normally two months from the date of transmittal of the International Search Report.

Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes
1211 Geneva 20, Switzerland, Facsimile No.: (41-22) 338.82.70

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

- ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.
☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. Reminders

Shortly after the expiration of 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90bis.1 and 90bis.3, respectively, before the completion of the technical preparations for international publication.

The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

Within 19 months from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later); otherwise, the applicant must, within 20 months from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

In respect of other designated Offices, the time limit of 30 months (or later) will apply even if no demand is filed within 19 months.

See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the International Searching Authority  European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Rosa Poquet
--	--

NOTES TO FORM PCT/ISA/220

These Notes are intended to give the basic instructions concerning the filing of amendments under article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the *PCT Applicant's Guide*, a publication of WIPO.

In these Notes, "Article", "Rule", and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions, respectively.

INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report and the written opinion of the International Searching Authority, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only (see *PCT Applicant's Guide*, Volume I/A, Annexes B1 and B2).

The attention of the applicant is drawn to the fact that amendments to the claims under Article 19 are not allowed where the International Searching Authority has declared, under Article 17(2), that no international search report would be established (see *PCT Applicant's Guide*, Volume I/A, paragraph 296).

What parts of the international application may be amended?

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

When?

Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

Where not to file the amendments?

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

How?

Either by cancelling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Section 205(b)).

The amendments must be made in the language in which the international application is to be published.

What documents must/may accompany the amendments?

Letter (Section 205(b)):

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.

NOTES TO FORM PCT/ISA/220 (continued)

The letter must indicate the differences between the claims as filed and the claims as amended. It must, in particular, indicate, in connection with each claim appearing in the international application (it being understood that identical indications concerning several claims may be grouped), whether

- (i) the claim is unchanged;
- (ii) the claim is cancelled;
- (iii) the claim is new;
- (iv) the claim replaces one or more claims as filed;
- (v) the claim is the result of the division of a claim as filed.

The following examples illustrate the manner in which amendments must be explained in the accompanying letter:

1. [Where originally there were 48 claims and after amendment of some claims there are 51]:
"Claims 1 to 29, 31, 32, 34, 35, 37 to 48 replaced by amended claims bearing the same numbers; claims 30, 33 and 36 unchanged; new claims 49 to 51 added."
2. [Where originally there were 15 claims and after amendment of all claims there are 11]:
"Claims 1 to 15 replaced by amended claims 1 to 11."
3. [Where originally there were 14 claims and the amendments consist in cancelling some claims and in adding new claims]:
"Claims 1 to 6 and 14 unchanged; claims 7 to 13 cancelled; new claims 15, 16 and 17 added." or
"Claims 7 to 13 cancelled; new claims 15, 16 and 17 added; all other claims unchanged."
4. [Where various kinds of amendments are made]:
"Claims 1-10 unchanged; claims 11 to 13, 18 and 19 cancelled; claims 14, 15 and 16 replaced by amended claim 14; claim 17 subdivided into amended claims 15, 16 and 17; new claims 20 and 21 added."

"Statement under article 19(1)" (Rule 46.4)

The amendments may be accompanied by a statement explaining the amendments and indicating any impact that such amendments might have on the description and the drawings (which cannot be amended under Article 19(1)).

The statement will be published with the international application and the amended claims.

It must be in the language in which the international application is to be published.

It must be brief, not exceeding 500 words if in English or if translated into English.

It should not be confused with and does not replace the letter indicating the differences between the claims as filed and as amended. It must be filed on a separate sheet and must be identified as such by a heading, preferably by using the words "Statement under Article 19(1)."

It may not contain any disparaging comments on the international search report or the relevance of citations contained in that report. Reference to citations, relevant to a given claim, contained in the international search report may be made only in connection with an amendment of that claim.

Consequence if a demand for international preliminary examination has already been filed

If, at the time of filing any amendments and any accompanying statement, under Article 19, a demand for international preliminary examination has already been submitted, the applicant must preferably, at the time of filing the amendments (and any statement) with the International Bureau, also file with the International Preliminary Examining Authority a copy of such amendments (and of any statement) and, where required, a translation of such amendments for the procedure before that Authority (see Rules 55.3(a) and 62.2, first sentence). For further information, see the Notes to the demand form (PCT/IPEA/401).

If a demand for international preliminary examination is made, the written opinion of the International Searching Authority will, except in certain cases where the International Preliminary Examining Authority did not act as International Searching Authority and where it has notified the International Bureau under Rule 66.1bis(b), be considered to be a written opinion of the International Preliminary Examining Authority. If a demand is made, the applicant may submit to the International Preliminary Examining Authority a reply to the written opinion together, where appropriate, with amendments before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later (Rule 43bis.1(c)).

Consequence with regard to translation of the international application for entry into the national phase

The applicant's attention is drawn to the fact that, upon entry into the national phase, a translation of the claims as amended under Article 19 may have to be furnished to the designated/elected Offices, instead of, or in addition to, the translation of the claims as filed.

For further details on the requirements of each designated/elected Office, see the *PCT Applicant's Guide*, Volume II.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference FH090608PCT	FOR FURTHER ACTION see Form PCT/ISA/220 as well as, where applicable, item 5 below.	
International application No. PCT/EP2009/004339	International filing date (day/month/year) 16/06/2009	(Earliest) Priority Date (day/month/year) 11/07/2008
Applicant FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ...		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 4 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the language, the international search was carried out on the basis of:

- ☒ the international application in the language in which it was filed
☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))

b. ☐ This international search report has been established taking into account the rectification of an obvious mistake authorized by or notified to this Authority under Rule 91 (Rule 43.6bis(a)).

c. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, see Box No. I.

2. ☐ Certain claims were found unsearchable (See Box No. II)

3. ☐ Unity of invention is lacking (see Box No. III)

4. With regard to the title,

- ☒ the text is approved as submitted by the applicant
☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract,

- ☒ the text is approved as submitted by the applicant
☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the drawings,

- a. the figure of the drawings to be published with the abstract is Figure No. 1
☒ as suggested by the applicant
☐ as selected by this Authority, because the applicant failed to suggest a figure
☐ as selected by this Authority, because this figure better characterizes the invention
b. ☐ none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/004339

A. CLASSIFICATION OF SUBJECT MATTER

INV. G10L11/00 G10L19/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TANCEREL L ET AL: "Combined speech and audio coding by discrimination" SPEECH CODING, 2000. PROCEEDINGS. 2000 IEEE WORKSHOP ON SEPTEMBER 17-20, 2000, PISCATAWAY, NJ, USA, IEEE, 17 September 2000 (2000-09-17), pages 154-156, XP010520073 ISBN: 978-0-7803-6416-5 abstract section 1, left-hand column, lines 4-12 section 2 figures 1, 2 ----- -/-	1-18

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"G" document member of the same patent family

Date of the actual completion of the international search

22 July 2009

Date of mailing of the international search report

31/07/2009

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Tilp, Jan

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/004339

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/101050 A1 (KHALIL HOSAM ADEL [US] ET AL) 29 May 2003 (2003-05-29) paragraphs [0001] - [0003] paragraphs [0006] - [0009] paragraphs [0031] - [0038] paragraphs [0059] - [0062] paragraph [0071] figures 1-4b	1-18

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2009/004339

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2003101050	A1	29-05-2003	NONE

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:

see form PCT/ISA/220

PCT

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY
(PCT Rule 43bis.1)**

Date of mailing
(day/month/year) see form PCT/ISA/210 (second sheet)

Applicant's or agent's file reference
see form PCT/ISA/220

FOR FURTHER ACTION
See paragraph 2 below

International application No.
PCT/EP2009/004339

International filing date (day/month/year)
16.06.2009

Priority date (day/month/year)
11.07.2008

International Patent Classification (IPC) or both national classification and IPC
INV. G10L11/00 G10L19/14

Applicant
FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ...

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. FURTHER ACTION

If a demand for international preliminary examination is made, this opinion will usually be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA:



European Patent Office

D-80298 Munich
Tel. +49 89 2399 - 0
Fax: +49 89 2399 - 4465

Date of completion of
this opinion

see form
PCT/ISA/210

Authorized Officer

Tilp, Jan

Telephone No. +49 89 2399-4753



**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/EP2009/004339

Box No. I Basis of the opinion

1. With regard to the **language**, this opinion has been established on the basis of:
 - ☒ the international application in the language in which it was filed
 - ☐ a translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1 (b)).
2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))
3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application and necessary to the claimed invention, this opinion has been established on the basis of:
 - a. type of material:
 - ☐ a sequence listing
 - ☐ table(s) related to the sequence listing
 - b. format of material:
 - ☐ on paper
 - ☐ in electronic form
 - c. time of filing/furnishing:
 - ☐ contained in the international application as filed.
 - ☐ filed together with the international application in electronic form.
 - ☐ furnished subsequently to this Authority for the purposes of search.
4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
5. Additional comments:

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/EP2009/004339

Box No. V Reasoned statement under Rule 43b/s.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes: Claims	<u>5, 6, 8</u>
	No: Claims	<u>1-4, 7, 9-18</u>
Inventive step (IS)	Yes: Claims	
	No: Claims	<u>5, 6, 8</u>
Industrial applicability (IA)	Yes: Claims	<u>1-18</u>
	No: Claims	

2. Citations and explanations

see separate sheet

Re Item V

Reasoned statement with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

- 1 Reference is made to the following document:

D1: XP010520073

2 **INDEPENDENT CLAIMS**

The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of the independent claims 1, 10, 13, 14, 17 and 18 is not new in the sense of Article 33(2) PCT.

- 2.1 The document D1 discloses (the references in parentheses applying to this document) - applying the wording of claim 1:

"A method for classifying different segments of a signal, the signal comprising segments of at least a first type and a second type (abstract, lines 2-5), the method comprising:

short-term classifying (150) the signal on the basis of at least one short-term feature extracted from the signal and delivering a short-term classification result (152) (figure 2 and section 2, page 154, line 11 - page 155, line 1);

long-term [classifying] (154) the signal on the basis of at least one short-term feature and at least one long-term feature extracted from the signal and delivering a long-term classification result (156) (figure 2, section 2.1, lines 4-7, 9-15); and

combining (158) the short-term classification result (152) and the long-term analysis classification (156) to provide an output signal (160) indicating whether a segment of the signal is of the first type or of the second type (figure 2)."

Consequently, claim 1 fails to fulfill the requirements of the PCT with respect to novelty (Article 33(2) PCT).

- 2.2 The document D1 discloses (the references in parentheses applying to this document) - applying the wording of claim 10:

"A method for processing a signal comprising segments of at least a first type and a second type (abstract), the method comprising:

classifying (116) a segment of the signal in accordance with the method of one of claims 1 to 9 (see references cited above against claim 1);

dependent on the output signal (160) provided by the classifying step (116), processing (102, 206; 106, 208) the segment in accordance with a first process or a second process; and outputting the processed segment (figure 1)."

Consequently, claim 10 fails to fulfill the requirements of the PCT with respect to novelty (Article 33(2) PCT).

- 2.3 The subject-matter of claims 13 and 14 corresponds in terms of computer program features and device features to that of claim 1. The objections raised above against said latter claim therefore also apply, *mutatis mutandis*, to claims 13 and 14, which thus fail to fulfill the requirements of the PCT with respect to novelty (Article 33(2) PCT).
- 2.4 The subject-matter of claims 17 and 18 is already disclosed by D1 (see e.g. figure 1, abstract, lines 2-5, and section 1, left-hand column, lines 6-12). Consequently, claims 17 and 18 also fail to fulfill the requirements of the PCT with respect to novelty (Article 33(2) PCT).

3 DEPENDENT CLAIMS

The dependent claims do not contain any features which, in combination with the features of any claim to which they refer, meet the requirements of the PCT in respect of novelty (claims 2-4, 7, 9, 11, 12, 15 and 16) or inventive step (claims 5, 6, 8), see document D1 and the corresponding passages cited in the search report.

Possible steps after receipt of the international search report (ISR) and written opinion of the International Searching Authority (WO-ISA)

General information

For all international applications filed on or after 01/01/2004 the competent ISA will establish an ISR. It is accompanied by the WO-ISA. Unlike the former written opinion of the IPEA (Rule 66.2 PCT), the WO-ISA is not meant to be responded to, but to be taken into consideration for further procedural steps. This document explains about the possibilities.

Amending claims under Art. 19 PCT

Within 2 months after the date of mailing of the ISR and the WO-ISA the applicant may file amended claims under Art. 19 PCT directly with the International Bureau of WIPO. The PCT reform of 2004 did not change this procedure. For further information please see Rule 46 PCT as well as form PCT/ISA/220 and the corresponding Notes to form PCT/ISA/220.

Filing a demand for international preliminary examination

In principle, the WO-ISA will be considered as the written opinion of the IPEA. This should, in many cases, make it unnecessary to file a demand for international preliminary examination. If the applicant nevertheless wishes to file a demand this must be done before expiry of 3 months after the date of mailing of the ISR/ WO-ISA or 22 months after priority date, whichever expires later (Rule 54bis PCT). Amendments under Art. 34 PCT can be filed with the IPEA as before, normally at the same time as filing the demand (Rule 66.1 (b) PCT).

If a demand for international preliminary examination is filed and no comments/amendments have been received the WO-ISA will be transformed by the IPEA into an IPRP (International Preliminary Report on Patentability) which would merely reflect the content of the WO-ISA. The demand can still be withdrawn (Art. 37 PCT).

Filing informal comments

After receipt of the ISR/WO-ISA the applicant may file informal comments on the WO-ISA directly with the International Bureau of WIPO. These will be communicated to the designated Offices together with the IPRP (International Preliminary Report on Patentability) at 30 months from the priority date. Please also refer to the next box.

End of the international phase

At the end of the international phase the International Bureau of WIPO will transform the WO-ISA or, if a demand was filed, the written opinion of the IPEA into the IPRP, which will then be transmitted together with possible informal comments to the designated Offices. The IPRP replaces the former IPER (international preliminary examination report).

Relevant PCT Rules and more information

Rule 43 PCT, Rule 43bis PCT, Rule 44 PCT, Rule 44bis PCT, PCT Newsletter 12/2003, OJ 11/2003, OJ 12/2003

Bitte beachten Sie, dass angeführte Nichtpatentliteratur (wie z. B. wissenschaftliche oder technische Dokumente) je nach geltendem Recht dem Urheberrechtsschutz und/oder anderen Schutzarten für schriftliche Werke unterliegen könnte. Die Vervielfältigung urheberrechtlich geschützter Texte, ihre Verwendung in anderen elektronischen oder gedruckten Publikationen und ihre Weitergabe an Dritte ist ohne ausdrückliche Zustimmung des Rechtsinhabers nicht gestattet.

Veuillez noter que les ouvrages de la littérature non-brevets qui sont cités, par exemple les documents scientifiques ou techniques, etc., peuvent être protégés par des droits d'auteur et/ou toute autre protection des écrits prévue par les législations applicables. Les textes ainsi protégés ne peuvent être reproduits ni utilisés dans d'autres publications électroniques ou imprimées, ni rediffusés sans l'autorisation expresse du titulaire du droit d'auteur.

Please be aware that cited works of non-patent literature such as scientific or technical documents or the like may be subject to copyright protection and/or any other protection of written works as appropriate based on applicable laws. Copyrighted texts may not be copied or used in other electronic or printed publications or re-distributed without the express permission of the copyright holder.

XS CPRTENFRDE

FH090608PCT

1/6

PCT REQUEST

Original (for SUBMISSION)

0	For receiving Office use only	
0-1	International Application No.	PCT/EP 2009 / 0 0 4 3 3 9
0-2	International Filing Date	1 6 JUN 2009 (1 6 . 0 6 . 0 9)
0-3	Name of receiving Office and "PCT International Application"	RO/EP
0-4	Form PCT/RO/101 PCT Request Prepared Using	PCT-SAFE [EASY mode] Version 3.51.038.214 MT/FOP 20090301/0.20.5.15
0-5	Petition The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty	
0-6	Receiving Office (specified by the applicant)	European Patent Office (EPO) (RO/EP)
0-7	Applicant's or agent's file reference	FH090608PCT
I	Title of invention	METHOD AND DISCRIMINATOR FOR CLASSIFYING DIFFERENT SEGMENTS OF A SIGNAL
II	Applicant	
II-1	This person is	applicant only
II-2	Applicant for	all designated States except US
II-4	Name	FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V.
II-5	Address	Hansastraße 27c 80686 München Germany
II-6	State of nationality	DE
II-7	State of residence	DE
III-1	Applicant and/or inventor	
III-1-1	This person is	applicant and inventor
III-1-2	Applicant for	US only
III-1-4	Name (LAST, First)	FUCHS, Guillaume
III-1-5	Address	Parkstraße 12 90409 Nürnberg Germany
III-1-6	State of nationality	FR
III-1-7	State of residence	DE

FH090608PCT

2/6

PCT REQUEST

Original (for SUBMISSION)

III-2	Applicant and/or inventor	
III-2-1	This person is	applicant and inventor
III-2-2	Applicant for	US only
III-2-4	Name (LAST, First)	BAYER, Stefan
III-2-5	Address	Dortmunder Straße 14 90425 Nürnberg Germany
III-2-6	State of nationality	DE
III-2-7	State of residence	DE
III-3	Applicant and/or inventor	
III-3-1	This person is	applicant and inventor
III-3-2	Applicant for	US only
III-3-4	Name (LAST, First)	NAGEL, Frederik
III-3-5	Address	Wilhelmshavener Straße 72 90425 Nürnberg Germany
III-3-6	State of nationality	DE
III-3-7	State of residence	DE
III-4	Applicant and/or inventor	
III-4-1	This person is	applicant and inventor
III-4-2	Applicant for	US only
III-4-4	Name (LAST, First)	HERRE, Jürgen
III-4-5	Address	Hallerstraße 24 91054 Buckenhof Germany
III-4-6	State of nationality	DE
III-4-7	State of residence	DE
III-5	Applicant and/or inventor	
III-5-1	This person is	applicant and inventor
III-5-2	Applicant for	US only
III-5-4	Name (LAST, First)	RETTELACH, Nikolaus
III-5-5	Address	Spessartstraße 38 90427 Nürnberg Germany
III-5-6	State of nationality	DE
III-5-7	State of residence	DE

FH090608PCT

3/6

PCT REQUEST

Original (for SUBMISSION)

III-6	Applicant and/or inventor	applicant and inventor US only WABNIK, Stefan Fichtenweg 5 98693 Ilmenau Germany DE DE
III-6-1	This person is	
III-6-2	Applicant for	
III-6-4	Name (LAST, First)	
III-6-5	Address	
III-6-6	State of nationality	
III-6-7	State of residence	
III-7	Applicant and/or inventor	applicant and inventor US only YOKOTANI, Yoshikazu Kibune 336-101, Hamamatsu, Shizuoka, 434-0038 Japan JP JP
III-7-1	This person is	
III-7-2	Applicant for	
III-7-4	Name (LAST, First)	
III-7-5	Address	
III-7-6	State of nationality	
III-7-7	State of residence	
III-8	Applicant and/or inventor	applicant and inventor US only HIRSCHFELD, Jens Steinweg 32 36266 Heringen Germany DE DE
III-8-1	This person is	
III-8-2	Applicant for	
III-8-4	Name (LAST, First)	
III-8-5	Address	
III-8-6	State of nationality	
III-8-7	State of residence	
III-9	Applicant and/or inventor	applicant and inventor US only LECOMTE, Jérémie Sulzbacher Straße 39 90489 Nürnberg Germany FR DE
III-9-1	This person is	
III-9-2	Applicant for	
III-9-4	Name (LAST, First)	
III-9-5	Address	
III-9-6	State of nationality	
III-9-7	State of residence	

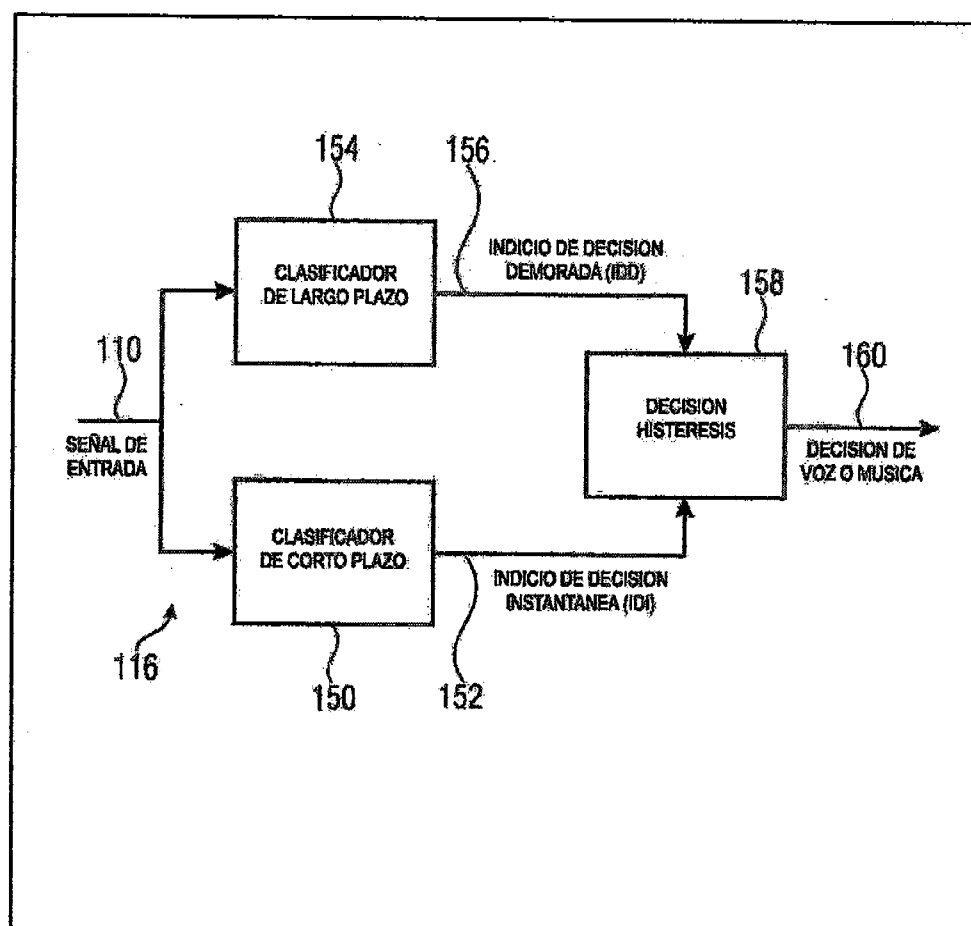
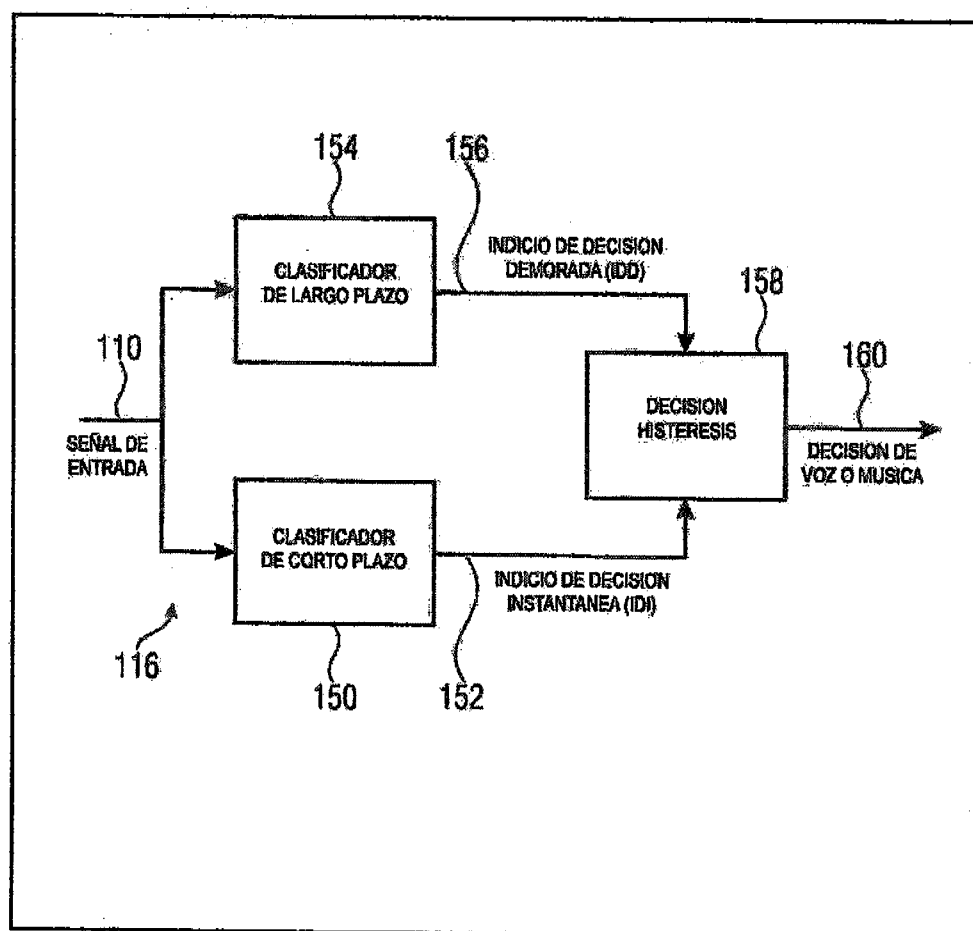
FH090608PCT

4/6

PCT REQUEST

Original (for SUBMISSION)

IV-1	Agent or common representative; or address for correspondence The person identified below is hereby/ has been appointed to act on behalf of the applicant(s) before the competent International Authorities as:	agent
IV-1-1	Name (LAST, First)	ZIMMERMANN, Tankred
IV-1-2	Address	SCHOPPE, ZIMMERMANN, STÖCKELER & ZINKLER Postfach 246 82043 Pullach bei München Germany
IV-1-3	Telephone No.	089/7904450
IV-1-4	Facsimile No.	089/7902215
IV-1-5	e-mail	mail@schoppe-zimmermann.com
IV-2	Additional agent(s)	additional agent(s) with same address as first named agent
IV-2-1	Name(s)	SCHOPPE, Fritz; STÖCKELER, Ferdinand; ZINKLER, Franz; SCHENK, Markus; HERSINA, Günter; BURGER, Markus; GÖTTLING, Stefan
V	DESIGNATIONS	
V-1	The filing of this request constitutes under Rule 4.9(a), the designation of all Contracting States bound by the PCT on the international filing date, for the grant of every kind of protection available and, where applicable, for the grant of both regional and national patents.	
VI-1	Priority claim of earlier national application	
VI-1-1	Filing date	11 July 2008 (11.07.2008)
VI-1-2	Number	61/079,875
VI-1-3	Country	US
VI-2	Incorporation by reference : where an element of the international application referred to in Article 11(1)(iii)(d) or (e) or a part of the description, claims or drawings referred to in Rule 20.5(a) is not otherwise contained in this international application but is completely contained in an earlier application whose priority is claimed on the date on which one or more elements referred to in Article 11(1)(iii) were first received by the receiving Office, that element or part is, subject to confirmation under Rule 20.6, incorporated by reference in this international application for the purposes of Rule 20.6.	



183

Industria y Comercio SUPERINTENDENCIA	SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO  No. 11-001544- -00000-0000
---	--

PATENTE DE INVENCION ☐ MODE ☐ Fecha: 2011-01-07 11:18:49 Dep. 2020 DIR.NUEVASCR)
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 182

☐	Indicación que se solicita una patente.
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Descripción de la invención
☐	Dibujos de ser estos pertinentes
☐	Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)
Completa ☐ Incompleta ☐	

PATENTE DE INVENCION PCT ☐ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

☐	Indicación que se solicita una PCT
☐	Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
☐	Dibujos de ser estos pertinentes
☐	Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)
Completa ☐ Incompleta ☐	

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

☐	Indicación que se solicita Diseño industrial
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
☐	Comprobante de pago de las tasas establecidas
Completa ☐ Incompleta ☐	

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

☐	Indicación que se solicita un esquema de trazado
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Representación gráfica de un esquema de trazado
☐	Comprobante de pago de las tasas establecidas
Completa ☐ Incompleta ☐	



Bogotá D.C.,
2020



Doctor(a)
ACOSTA GOMEZ JUANITA
Apoderado(a) y/o Representante de
FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG E.V.

REFERENCIA	Radicación No.	11 001544
	Solicitud internacional	PCT/EP2009/004339
	Fecha inicio fase nacional	11/02/2011
	Trámite	11
	Evento	379
	Actuación	430
	Oficio No.	03931

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante, conforme a lo establecido por el artículo 26 literal k) de la Decisión 486 de la Comisión de la Comunidad Andina. Cuando se trate de documentos otorgados en el extranjero, estos deberán legalizarse de conformidad con lo dispuesto por los artículos 259 y 260 del Código de Procedimiento Civil. Cuando se trate de las declaraciones a que hace referencia la Regla 4.17 ii) del Tratado de Cooperación en Materia de Patentes PCT., el alcance de las tales declaraciones se determinará según indiquen claramente una cesión de derechos del inventor al solicitante o su causante, de conformidad con lo dispuesto en el numeral 5.2.15 del Capítulo Quinto, Título X de la Circular Única de Superintendencia de Industria y Comercio.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Unica No.10 de 2001.


JOSE LUIS SALAZAR LOPEZ
Director de Nuevas Creaciones (c)

El anterior requerimiento fue notificado por fijación en lista de fecha 04 MAR 2011
adpa11