

(19)
(12)

(KR)
(B1)

(51) 。 Int. Cl. ⁷
G06F 17/00

(45)
(11)
(24)

2003 04 21
10 - 0380947
2003 04 07

(21) 10 - 2000 - 0002364
(22) 2000 01 19

(65) 2000 - 0076488
(43) 2000 12 26

(30) 09/234,663 1999 01 19 (US)

(73) 10504

(72) 95070 가 13591
94303 3770
95120 7190

(74)

:

(54)

가 (歌詞 : speech on music) 가 가

가 .

1

1 ,

2 ,

3 ,

4 가 ,

5 가 ,

6 5 ,

7 가 ,

8 7 ,

9 가 ,

10 9 ,

11 , , ,

12 " "

,

13 , ,

,

14 ,

15 ,

16 .

10 : 12 :

14 : 16 :

17 : 18 :
20 : 22 :
24 : 25 :
26 : 28 :
29 :

(gisting), (skimming),

가
가

가

가
가

(source text)가

가

가

가

가

가

가

(

,"land"

(rec

all)

"lamb"

가 (漸減)

),

(precision)"

(recall) 가

가

가

가

가

가

가 ,

(fast forward)

(,)

8063184

10049189

702351

5,199,077

가

8063184

10049189

8087292A

EP702351A

(phrases)

5,655,058

EP780777A

(audio event)

Erling IEEE Multimedia 1996 "Content - Based Classification, Search, and Retrieval of Audio" (, "Muselefish" ())

"Content - Based Classification, Search, and Retrieval of Au

ally) (Muselefish)가 (heuristic

(15 sec)

Pfeiffer ACM Multimedia 96(1996)

"Automatic Audio Content Analysis"

(, "MoCA()")

가 , (MoCA)

(domain specific),

(violenc

e) (video event)

가
가 90%

(3

)
,가

가
()

(ZCR),
(F)

(SE)

(

)

()

(dominant frequency)

N%

, N

1

N%

가

가

가

/

1 (10)

(12) (12)

International Business Machines Corporation(IBM) (12) IBM

AS400

(12) (Unix) OS/2 NT AIX 3. 2. 5. 가

128 MB IBM RS/6000 250

(12) (12) 가 1

(14) (12) RAM

6) 가 (12) 가 (17) 1 (1

DASD

가 C++

1 (10) (18) / (20) (22) /

(24) 가 (22) (24) 가,

(12) (25) (26)

(28) (26) (28)

(14) (26)

(25) (26) (12)

(25) (wide area network)

(14) (29) (29)

5,293,584

"Large Vocabulary Continuous Speech Recognition"

15 (300)

(302) 가

(304) 16 (306) (3

06) (18)(1) (304)

()

15 (308) , (310)(16)
 , (302) 가
 가 " " 15 (312) 가
 가 (306) (324) 가
 (316), (314) (clip)
 (play) (320) (318)
 , FFT (14) (30) (FFT)
 (31) 2 가
 100 (14) (30)
 (31) , 가,
 (32) (speech on music) 가
 (33) (32) (34) (가
)가 (33) 30 가 (com
 merical) (33) / 가
 " " , 4
 (35) (29)(1) (3 (alpha -
 6) (29) (36)

(35) " (newscasts)"
 가 /

(36) 가 가 " " 가 " 가
 가 .가 가

(37) , DO 가 N (, N
 , 3) 가 가 90% 가 (29)
 (37) N
 가 (37)

DO (38) (39) (39)
 가 ,가 가 가
 (lattice)가 (40) (Okapi)
 가
 (40)

가 (out - of - vocabulary word)
 , Okapi 가
 () (corpus) (thesaurus)

,가 () 가
 가 (inverse document frequency) 가

$$l_1 = 0.5 \quad l_2 = 1 + l_1 (\quad),$$

$$l_d = \quad d \quad , l \quad ,$$

$$q_k = \dots k \dots ,$$

$$C_d(q_k) = \dots "d" \dots "k" \dots ,$$

$$EC_d(q_k) = E_{dk} \dots "d" \dots "k" \dots ,$$

$$C_q(q_k) = \dots "q" \dots k \dots ,$$

$$E_q(q_k) = E_{qk} \dots "q" \dots k \dots ,$$

$$n(q_k) = \dots q_k \dots ,$$

$$n'(q_k) = \dots q_k \dots ,$$

$$Q' = \dots , N \dots ,$$

$$p_i(q_k) = \dots k \dots i \dots \text{가} \dots ,$$

$$e_i(q_k) = k \dots i \dots \text{가}$$

$$k = \text{idf}(q_k):$$

$$\text{idf}(q_k) = \log \{ (N - n'(q_k) + 1) / (n'(q_k) + 1) \}$$

$$"q" = S(d, q) \dots "d" \dots :$$

$$S(d, q) = \{ E_{dk} * E_{qk} * \text{idf}(q_k) \} / \{ \dots + \dots (i_d / i') + E_{dk} \} \quad k=1 \dots Q'$$

$$E_{dk} = \dots "d" \dots \{ p_i(q_k) * e_i(q_k) \} \quad i=1 \dots Q'$$

$$E_{qk} = \dots "q" \dots \{ p_i(q_k) * e_i(q_k) \} \quad i=1 \dots Q'$$

가 , 가 , e_i(q_k) , "e" , 가
 "+word" (prefix symbol)
 e_i(q_k) 0 1 가 가 ,
 0 1 , e_i(q_k)

$$3 \dots (44)$$

$$k \dots (46) \dots (48) \dots DO \text{가}$$

$$(RS_i) \dots (46) \dots (ZCR), \dots (E)$$

RSi (squared) . RSi i , , i , , , 4 가 .
 1 (R1) 0 - 1000 Hz , 2 (R2) 1000 - 8000 Hz , 3 (R3)
 8000 - 16,000 Hz , 4 (R4) 16,000 Hz .

가 (harmonicity)(
)가 (調性 : tonality) 가 .

(48) , . (quot
 ient) . (NRi) .

$$NR_i = (RS_i - \text{mean}(RS_i)) / SRS_i$$

4 가 . 4 - 10 , , (threshold) (heuristics)
 가 / 가 (50)
 ple) 1 , DO 가 (52) , 가 90% (sam 가
 . 44 KHz 100 ms , 20

가 " (silence)" 90% 1 (R1) , (54) , (54)
 (52) , 가 (56) 가 , (52)
 , 가 (60) .

5 가 (62) ,
 DO 가 (64) ,
 3 (R3) 가 15% 가 .
 , Speedch Freq 가 (66) 가 .

(66) 가 (64) , (ZCR) 6
 가 (70) (68) 가 (70) 가
 가 (72) 가 (68) ,
 (74) 가 , (64) , 가
 6 .

가
 가

6 , 3가 (73) 14 ()
) " " 가 () , 40%
 (74) 가 , Speech Freq (74) , 40%가
 5 (64) 가 . , 2
 , ZCR (76) 가 , (76) , 20%
 5 (68) 가 . 6 (74) 가
 , (73) , (78) ,
 (80) 가 , 6 ,

7 , 가 가 (82)
 , DO 가 (84) , 가 15%
 가 , 3 (R3) (86) 가 .
 (86) 가 (84) , (ZCR) 5
 (88) 가 . , ZCR 가 (90)
 가 . (90) 가 (88) ,
 3 (NR3)(3 (48)) 100,000
 (92) 가 . , EN 가 (94) 가
 가 . (94) 가 (92) ,
 , (96) 가 ,
 (98) 가 (84) . , 가 , 8

3가 80%
 (100) , (102)
 가 . , 가 (100) 1 , 2
 , ZCR (104) 가 . 2 가 , 95%
 , 3 (106) 가 .

(106) , EN 80% 가 .
 가 " " (102) . 3가 가 , 가
 (108) .
 9 , 가 (SOM) 가
 . DO 가 (112) ,
 3 (R3) 가 15%
 가 . , SOM 가 (114) 가 .

(114) 가 (112) , (ZCR)
 5 10 (116) 가 , SOM ZCR
 (118) 가 . (118) 가 (116) ,
 3 (NR3)가 90,000
 (120) 가 . , SOM EN (122) 가 . (12
 2) 가 (12) , 가 (126) 가
 (124) 가 , , (112) 가 10 .

), " " , 2가 , (70%
) 가 . (128) , SOM ZCR 가 (1
 30) 가 가 , 가 (128) 1 , (1
 2 1 (subtest) (132) 가 , (1
 32) , SOM 50% .
 1 가 SOM ZCR 15%
 N (134) 2 가 , SOM E
 10% (136) . 2
 3가 가 , (130)
 (132, 134, 136) 가 ,
 (138) 가 . ,

2 , , (prestored)
 , 가 . 11 ,
 , 가 .
 (140) , (142) , ,
 (144) , / " " (144)

, (144)
 . 가 , / ,
 (scoll) , , /

, 12 13 ,
 " " 가 가 .
 12 (146) , 가 (pitch)
 1 (branch) ,

5,293,584 (148) (1)
50) ()

(150) (146) 2 ()
152) 13 "가
" " " 12

(154) 3 () (156)

13 , 3 /
(160) 1
0

(162) 가
8

- R1 - 100Hz 3,000Hz, R2 - 3,000Hz 4,000Hz,
- R3 - 4,000Hz 5,000Hz, R4 - 5,000Hz 6,000Hz,
- R5 - 6,000Hz 6,500Hz, R6 - 6,500Hz 7,000Hz,
- R7 - 7,000Hz 7,500Hz, R8 - 7,500Hz 8,000Hz.

RnFreq = n(n = 1 8) 가 RnFreq (8
가) 가

(164) 가 (histogram) 가 (166)
1%

(168) 가 1% 가
100 가 (172)
(170) (168) 가 (172)

(174) 가 1%
가 100 가 (172)

(172) . (176) . , (174)

13 가 , 3 가 , (176) (178) ,
 . (180) 가 , (180) (178) , (172) .

14 6 " " DO 가 . (202)
 , (200) (f_R) 가 (f_R) 가

2)가 , (f_R) , (f₁) , f₂ = (i/i+1) * f₁ (, i
 , (f_R) f_R = f₁ / i . (F_R) (f₂) . f₁ f₂가 (202) .

04) (202) 가 , (206) 가 (2 (202)
 가 , (208) .

" 가 , (204) (210) . (210) , "
 (212) . , 2 가 . " , "
 6 (214) " "

(57)

1.

(temporal sequence)

(recall) , 가

2.

가 가 가

가

3.

2

4.

3

N x%

5.

2 ,

6.

2 ,

7.

6 ,

8.

6 ,

(gisting)

9.

2 ,

가

가 ,

10.

2 ,

11.

10 ,

(dominant frequency)

가 N%(N) ,

가 N%(N) ,

12.

11 ,

가 ,

13.

14.

13 ,

15.

13 ,

가

가,

,

,

,

,

16.

15 ,

가

,

,

,

가 ,

,

17.

16 ,

,

가

N%(N)

,

,

가

N%(N)

,

.

18.

17

가

19.

가

가

가

20.

19

가

21.

19

가 .

22.

21 ,

가 ,

가 .

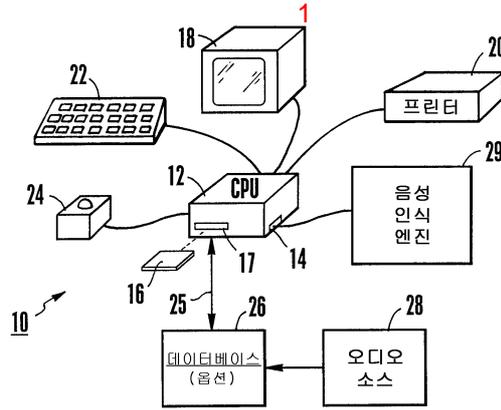
23.

22 ,

가 N%(N) ,

가 N%(N) ,

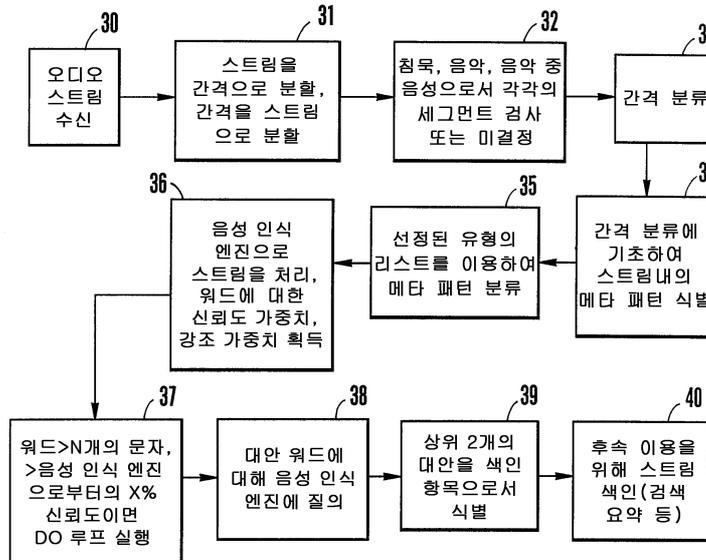
가 .



본원의 시스템

2

색인 논리



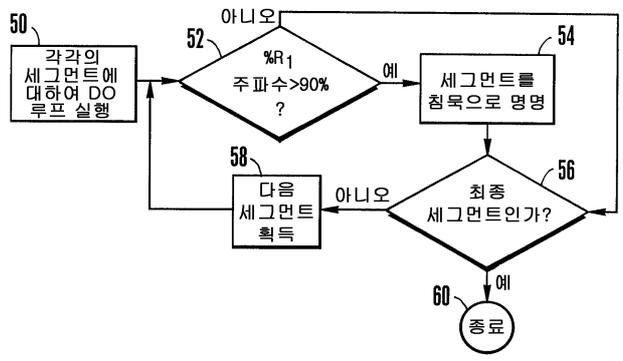
3

오디오 특징 검사



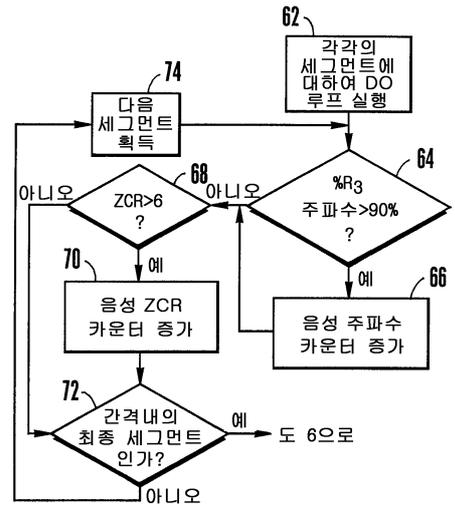
4

침묵 검사

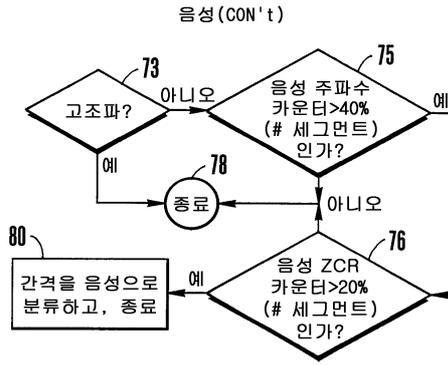


5

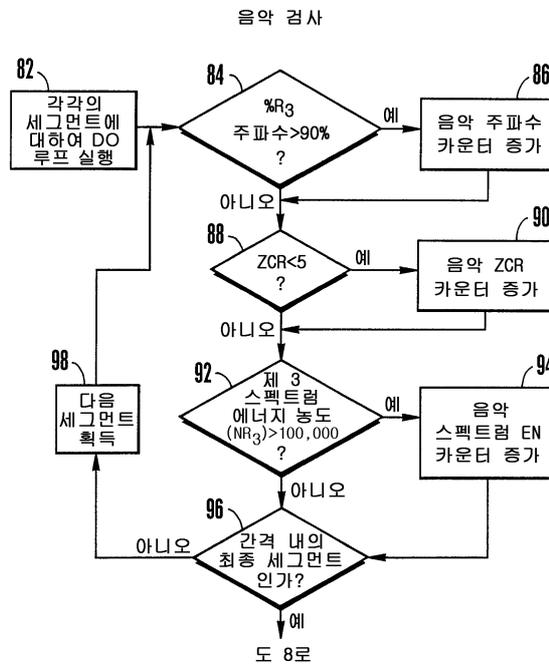
음성 검사



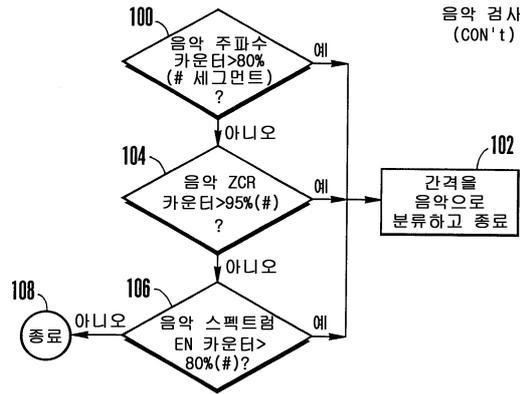
6



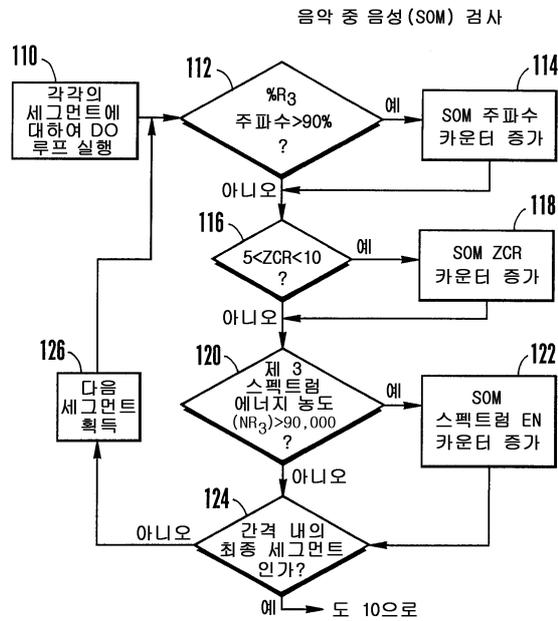
7



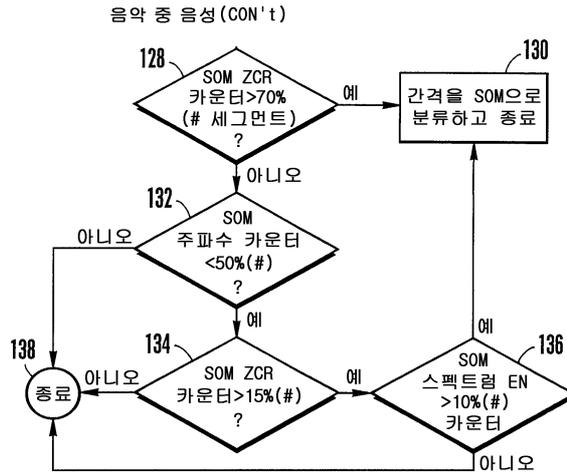
8



9

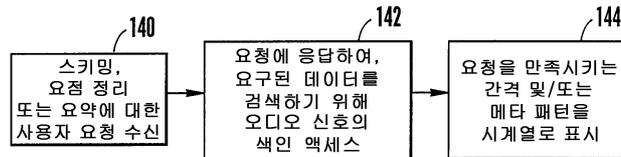


10

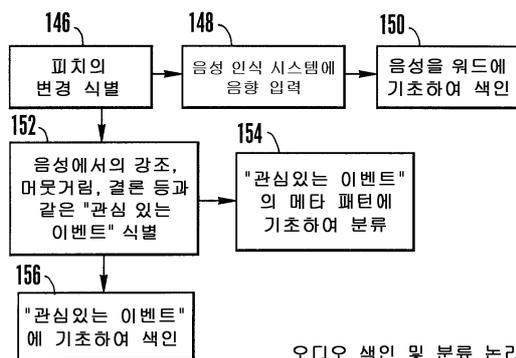


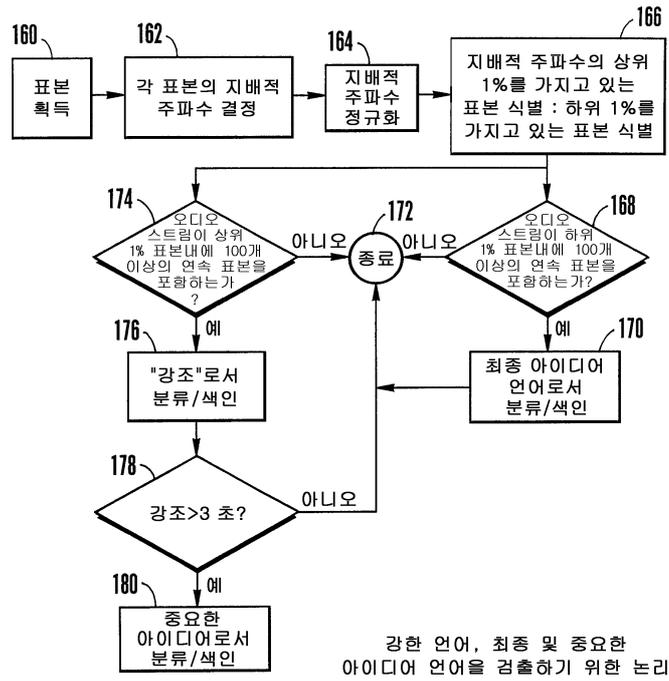
11

스키밍, 요약 정리 또는 요약 논리



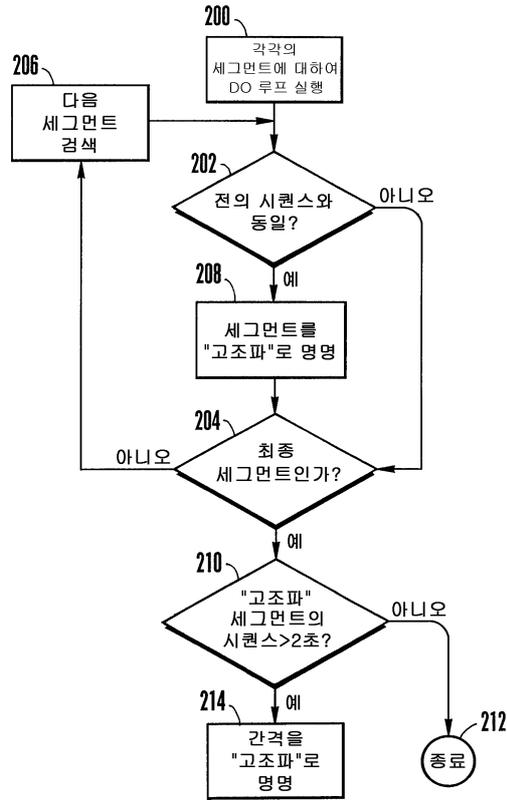
12





14

고조파 검출 논리



15

요약 및 스키밍 논리



16

사용자 인터페이스

