

303451

(由本局填寫)

承辦人代碼：
大類：
IPC分類：

A6
B6

本案已向：

日本國(地區) 申請專利，申請日期：1995-5-12 案號：7-114752，有 無主張優先權

有關微生物已寄存於：

，寄存日期：

，寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

經濟部中央標準局員工消費合作社印製

五、發明說明(1)

發明之背景

a) 發明之領域

本發明係關於一種利用只具有比語音輸入信號(input speech signal)少的資訊量來傳送或存儲語音資訊的系統及方法。本發明尤其是關於一種從語音輸入信號中抽出(extract)顯示其特徵之參數，再傳送或存儲(transmit or store)所抽出之參數，再根據所傳送或存儲之參數來合成(synthesize)本來之語音信號的系統及方法。本發明更具體而言是關於一種為了要聽覺性抑制(supress)被合成之語音信號(語音合作信號：Synthesized speech signal)中之量化噪音(quantizing noise)的語音加工過濾器，及為了要改善語音了解性等之所希望品質的語音強調過濾器。

b) 習知技術之描述

圖 31 顯示語音分析/合成系統之一例構成。該圖之系統係由分析單元(analyzing unit)100及合成單元(synthesizing unit)200所構成，而分析單元100係由分析器(analyzer)101及編碼器(coder)102所構成，而合成單元200係由解碼器(decoder)201及合成器(synthesizer)202所構成。單元100及200在某用途上，是透過通信通道(communication channel)來連接；此情況一般兩者皆是被遠程配置。單元100及200在其他之用途上，是透過記憶媒體(storage media)來傳送接收資訊；此情況有時兩者被構成單一裝置，而有時則被構成各別裝置。分析器101

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(2)

係由使用者所提供的語音輸入信號中，抽出顯示該語音輸入信號之特徵的頻譜資訊。被抽出之頻譜資訊可利用編碼器102來編碼，並透過通信通道或記憶媒體供給合成單元200，再利用解碼器201解碼。合成器202係根據被解碼之頻譜資訊來合成語音信號。具有如此構成的系統優點，是被傳送或存儲之信號的資訊量會少。此乃是被傳送或存儲的信號比被編碼化之頻譜資訊，亦即比語音輸入信號會導致其資訊量少之信號的因素。

圖32顯示合成單元200之變形例：該變形例具有後處理過濾器(post filter)203，係對依合成器202而所得之語音信號(以下稱為語音合成信號)，藉由根據被解碼之參數群施以規定之加工處理，來生成被加工之語音信號(以下稱為語音加工合成信號)者。該後處理過濾器203在某用途上，係為了要聽覺性抑制(suppress)語音合成信號中之量化噪音(quantizing noise)而被使用，在其他用途上，係為了改善語音了解性等之主觀品質而被使用。以下之說明中，係將此種後處理過濾器稱作為語音加工過濾器或語音強調過濾器。具備如此之過濾器203的合成單元200，特別適用在語音編碼解碼系統(voice coding/decoding system)或語音對話系統(voice recognition and response system)上。

在可作為過濾器203來用的過濾器雖有各式各樣，但是其中亦強調語音素特徵者，係具有抑制量化噪音及改善主觀品質等之效果特大的優點。在揭示如此之過濾器的習知

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(3)

技術文獻中，例如有如以下之文獻。在文獻1及文獻2中所
示的過濾器，皆在從分析單元100中接受線性預測碼（
linear prediction codes:LPC）作為上述之頻譜資訊的合成
單元200上，使用語音加工過濾器203。在文獻3中所示
的過濾器，則是在從分析單元100中接受自相關係數（
autocorrelation constants）作為上述之頻譜資訊的合成
單元200上，使用語音加工過濾器203。接著在文獻4中所
示的過濾器，是在從分析單元100中接受Mel-倒頻譜（Mel-
cepstrum）作為上述之頻譜資訊的合成單元200上，使用語
音加工過濾器203。

日本專利特開昭64-13200號公報（以下稱為文獻1）；

日本專利特表平5-500573號公報（以下稱為文獻2）；

日本專利特開平2-82710號公報（以下稱為文獻3）；及

「考慮誤傳送之自適應Mel-倒頻譜語音編碼系統」，日
本音響學會，平成6年（1994年）度春季研究發表會演講論
文集，分冊I，第257～258頁，（1944-03）（以下稱為文獻
4）。

圖33顯示由文獻1所揭示之過濾器的概略構成：該過濾
器203係在由合成器202所供給的語音合成信號之外，還輸
入從解碼器201所解碼的LPC。在此所謂的LPC，係指由線
性預測編碼所得的 α 參數（ α parameters）：所謂線性預測
編碼，係指將人類之發聲機構模化後之例如8次～12次之
過濾器的過濾係數（filter coefficients），亦即係根據
語音輸入信號波形之抽樣值且依照線性預測法（linear

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

五、發明說明(4)

prediction method), 來決定 α 參數的方法, 可利用圖 31 所示之分析器 101 來實行。

圖 33 所示之過濾器 203, 具有過濾器 204, 藉由過濾語音合成信號來生成語音半加工合成信號者; 及過濾器 205, 藉由過濾語音半加工合成信號來生成語音加工合成信號者; 過濾器 204 及 205, 皆以 α 參數作為過濾係數來使用。但是在過濾器 204 中所使用的 α 參數並非從解碼器 201 所供給的 α 參數 α_i (但是 $i=1, 2, \dots, p$; p 為預測次數), 而是藉由校正係數 ν 來校正 α_i 的 $\alpha_{1i} = \alpha_i / \nu^{-i}$ 。同樣, 在過濾器 205 中所使用的 α 參數, 係為藉由校正係數 n 來校正 α_i 的 $\alpha_{2i} = \alpha_i / n^{-i}$ 。藉由校正係數 ν 及 n 來校正 α_i 的處理, 可各別利用 LPC 校正器 206 及 207 來實行。

現在, 各別藉由過濾器 204 及 205 來實現將語音合成信號轉換成語音加工合成信號用的傳遞函數 (transmission function) $H(z)$ 之分母及分子; 亦即, 可將過濾器 204 作為 LPC 合成過濾器 (LPC filter) 204, 將過濾器 205 作為 LPC 反相過濾器 (inverse-LPC filter) 205。可更進一步將以 α_i 為過濾係數所用的過濾做成以下的式子。

$$A(z) = \sum_{i=0}^p (\alpha_i z^{-i}) \dots \dots \dots (1)$$

但是 z 為 z 轉換運算子 (z transformation operator)。在過濾器 204 及 205 中所使用的過濾係數由於如上述各別為 $\alpha_{1i} = \alpha_i / \nu^{-i}$ 及 $\alpha_{2i} = \alpha_i / n^{-i}$, 所以過濾器 204

(請先閱讀背面之注意事項再填寫本頁)

表

訂

五、發明說明(5)

及 205 之傳遞函數可各別表示成 $1/A(z/\nu)$ 及 $A(z/n)$ 。因而，將語音合成信號轉換成語音加工合成信號用的傳遞函數 $H(z)$ ，可以下式表現。

$$H(z) = A(z/n) / A(z/\nu) \dots \dots \dots (2)$$

圖 34 顯示文獻 2 所揭示的過濾器之概略構成：在該過濾器 203 中，在 LPC 校正器 206 上所生成的 $\alpha 1_i$ ，係藉由 LPC/ACC 轉換器 208 從 LPC 領域 (domain) 轉換至自相關領域，並藉由 ACC 校正器 209 在自相關領域內進行帶寬擴充 (bandwidth expansion)，再藉由 ACC/LPC 轉換器 210 且依照磊賓遜 (Rebinson) 之歸納法從自相關領域轉換至 LPC 領域。過濾器 205 係輸入由此所得的 $\alpha 2_i$ 。另外，該圖中雖廢除圖 33 所示的 LPC 校正器 207，但是在文獻 2 中亦教示其具備 LPC 校正器 207 且藉由 LPC/ACC 轉換器 208、ACC 校正器 209 及 ACC/LPC 轉換器 210 再次校正其輸出之 $\alpha 2_i$ 的構成。

圖 35 顯示文獻 3 所揭示之過濾器的概略構成：該過濾器 203 係具有在文獻 1 之構成中追加 ACC/LPC 轉換器 211 及 212 的構成。ACC/LPC 轉換器 211 係輸入自相關係數作為頻譜資訊，再將已輸入之自相關係數從自相關領域轉換成 LPC 領域。ACC/LPC 轉換器 212，係在輸入 ACC/LPC 轉換器 211 之自相關係數中輸入下位 m 次 (low order) 之部分 (m 次，但是 $m < p$)，再將已輸入之自相關係數從自相關領域轉換成 LPC 領域。LPC 校正器 206 及 207 係以和文獻 1 同樣的方法，校正各別在 ACC/LPC 轉換器 211 及 212 中所得的 α 參數。另外，在該構成輸入之自相關係數，亦可為以解碼器 201 解碼者 (亦

(請先閱讀背面之注意事項再填寫本頁)

衣

訂

85年11月18日 修正
補充

A7
B7

五、發明說明 (6)

即以分析器 101 算出而以編碼器 102 編碼的自相關係數)，或亦可根據以解碼器 201 解碼之他種參數來算出解碼器 201 或合成器 202 者。

圖 36~圖 38 顯示文獻 1~文獻 3 所揭示的語音加工(或強調)過濾器之對數功率頻譜特性(log-power vs. frequency spectrum characteristics)。該等之圖中，A~D 依序表示合成器 202 之特性 $=1/A(z)$ 、過濾器 204 之特性 $=1/A(z/\nu)$ 、過濾器 205 之逆特性(inverse characteristics) $=1/A(z/\eta)$ 、傳遞函數 $H(z)=A(z/\eta)/A(z/\nu)$ 。從式子(2)中就可明白，又從圖 36~38 中亦可明白，過濾器 204 係作為強調語音合成信號之頻譜的語音素，同時抑制該頻譜之谷的過濾器功能，而過濾器 205 係作為消除因過濾器 204 而所導入之頻譜斜率的過濾器功能。過濾器 204 之強調及抑制的程度，係 ν 愈大就愈強，而 ν 愈小則愈弱。另外，文獻 1 中，是假設 η 及 ν 能滿足 $0 \leq \eta \leq \nu < 1$ 。再者，圖 36、37 及 38 係各別顯示 $\nu = 0.8$ 、 $\eta = 0.5$ 之例， $\nu = 0.8$ 、使用 1200Hz 落後窗(lag window)之帶寬擴充處理之例，及 $p=10$ 、 $m=4$ 、 $\nu = 0.95$ 、 $\eta = 0.95$ 之例。

又，從圖 36 和圖 37 之比較中可明白，或從圖 36 和圖 38 之比較中可明白，若依據文獻 2 或文獻 3 所揭示的語音加工(或強調)過濾器，來比較文獻 1 所揭示者的話，則可增強過濾器 205 消除頻譜斜率的效果。亦即，文獻 1 所揭示之技術中利用過濾器 205 不能充分消除依過濾器 204 所授與的頻譜斜率。又，該頻譜斜率由於係和時間同時變化所以未能在

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

修正
補充
75年11月18日

五、發明說明(7)

固定之高通頻譜強調處理下消除，亮度就會和時間同時變化。相對於此，若依據文獻2及文獻3所揭示的技術，則可增強頻譜之山谷構造的強調且可使頻譜斜率更加平坦。此關係著藉可依過濾器203來防止明瞭度(亮度)及自然性的劣化。

但是，文獻2及3所揭示的技術對文獻1所揭示的技術而言一方面雖是改良技術，但是另一方面具有不良點。例如在文獻2所揭示的技術中，雖亦依據分析單元101之構成及標準方式，但是所獲得之語音加工合成信號經常有伴隨獨特失真之問題。此是因在自相關領域上進行非常強之頻譜平滑化(smoothing)處理而結果導致在語音素很強的近旁頻譜發生很大失真的因素。此結果，即使比起文獻1所揭示的技術語音加工合成信號的品質亦會發生不良的情況。又，在文獻3所揭示的技術中，起因於自相關領域中之過濾次數的減低，所以屢次語音數位置有時會大移動，有時多個語音數會成一個。如此不穩定的頻譜變化，會給語音加工合成信號帶來失真。例如若比較圖38所示之特性B和特性C，則可了解會出現最低頻率之語音素的移動現象；及正中央之二個語音素會成一個的現象。再者，如此原因所造成之大的語音素移動，由於有時會和時間同時發生或不發生，所以語音加工合成之音色就會不自然地搖晃變化。

又，在文獻1~文獻3所揭示之技術中，其共同的問題點是具有設計自由度(特性操作·調整之自由度)低的問題。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(8)

例如文獻1所揭示的技術，只有在頻譜斜率或其時間變動的問題變成不太顯著的範圍內使 ν 和 n 變化，才不會使過濾器203之特性改變太大。又，文獻2所揭示的技術，若加大過濾器204之語音素強調效果並將 ν 或落後窗頻率之可變範圍設定很大，則上述之失真，亦即起因於自相關領域之頻譜平滑化處理的失真會變大。因而，由於不得不限定 ν 或落後窗頻率之可變範圍，所以不會將過濾器203之特性改變太大。再者，文獻3所揭示的技術，由於係將過濾次數之有限整數值作為控制變數(control variable)，所以自身會使特性之自由度變低。

圖39顯示文獻4所揭示之語音加工(或強調)過濾器203的構成：在該圖之過濾器203和前述之各習知技術相差甚大之點，係因從解碼器201輸入Mel-倒頻譜作為頻譜資訊，和藉由將校正已輸入之Mel-倒頻譜而所得的校正Mel-倒頻譜作為過濾次數而用的過濾來將語音合成信號轉換成語音加工合成信號。亦即，語音合成信號係可藉由將利用Mel-倒頻譜校正器214所生成的校正Mel-倒頻譜作為過濾係數而用的過濾器213來過濾。更具體而言，Mel-倒頻譜校正器214，係藉由將已輸入之Mel-倒頻譜的一次成分置換成0而其他成分放大 β 部，來生成校正Mel-倒頻譜。過濾器213係將該校正Mel-倒頻譜作為過濾次數來用並過濾語音合成，再將所得之信號以語音加工合成輸出。另外，過濾器213由於係為將校正Mel-倒頻譜作為過濾係數而用的過濾器，所以稱為Mel-對數頻譜近似(MLSA)過濾器。

A7
修正
補充
1975年11月18日

五、發明說明(9)

在此所謂的 Mel-倒頻譜，係指依據直交轉換(orthogonal transformation)語音輸入信號之對數頻譜(log spectrum)而利用分析器 101 所算出的參數。一般是不能將文獻 1~文獻 3 之技術原狀適用在將語音資訊轉換成 Mel-倒頻譜再傳送或存儲的系統上；亦即，若將 Mel-倒頻譜等之倒頻譜系統的參數轉換至 LPC 領域的話，則由於頻譜形狀會失真(distort)大，所以語音合成信號之再分析的 LPC 之算出就變成必要。加上，即使是如此所算出的 LPC，由於在和因分析原語音所得的 LPC 之間亦有失真，所以無法獲得那麼良好的語音加工特性。相對於此，在使用文獻 4 之方法的情況時，就可防止該失真。

逆言之，在文獻 4 所揭示的技術中，有連接性(connectability)不佳，亦即在使用倒頻譜系統以外的參數群來合成語音信號的系統中有使用困難的問題。作為如此之系統，例如有使用 LPC、LSP(line spectrum pairs)、PARCOR(partial autocorrelation coefficients)等之參數群的系統。LPC、LSP 及 PARCOR 由於多採用於語音之編碼解碼上，所以此問題點得重要。假設在作為頻譜資訊而輸入 Mel-倒頻譜的合成單元 200 上，使用以 α 參數為過濾係數而用的語音加工過濾器的話，如前述般，頻譜形狀會伴隨從 LPC 領域轉換至 Mel-倒頻譜領域而失真。當然藉由再分析語音合成信號而再次算出 Mel-倒頻譜可某程度地消除此失真。然而，如此一來即使算出 Mel-倒頻譜，但若比起以分析器 101 所得的 Mel-倒頻譜的話亦會含有很多的

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(10)

失真。亦即無法獲得那樣良好的語音加工特性。

發明之概述

本發明之第一目的，係在於實現在被容許之頻譜斜率的範圍內可獲得良好的語音素強調效果之語音加工(或強調。以下省略)過濾器。本發明之第二目的，係在於實現不會使可知覺之準位的失真在語音素構造上發生，並可獲得良好的語音素強調效果之語音加工過濾器。本發明之第三目的，係在於實現以此習知少的構成手段就可實現和習知同等之語音素強調效果的語音加工過濾器。本發明之第四目的，係在於實現可選擇進行亮度之控制，處理量之削減，了解性之改善等的語音加工過濾器。本發明之第五目的，係在於實現設計自由度高的語音加工過濾器。本發明之第六目的，係在於實現適於將LSP、PARCOR、對數截面積比(log area ratio:LAR)等從分析單元側作為頻譜資訊輸入的合成單元之語音加工過濾器。本發明之第七目的，係在於實現在將LSP、PARCOR、LAR等作為頻譜資訊輸入時，不同進行頻譜之再分析或參數轉換就可獲得良好之連接性的語音加工過濾器。本發明之第八目的，係在於使用可達成第一至第七目的之語音加工過濾器來實現語音合成系統。若依據本發明之第一態樣，則藉由利用過濾係數所限定之傳遞函數來過濾語音合成信號，就可生成語音加工合成信號。該過濾係數係以多次元向量來表現，可根據屬於規定領域且關於語音輸入信號的頻譜資訊來生成，以便可按照上述頻譜資訊且比上述語音合成信號還能強調上述語

(請先閱讀背面之注意事項再填寫本頁)

訂

錄

五、發明說明(11)

音加工合成信號的語音素特徵。再者，作為頻譜資訊，可使用LSP資訊、PARCOR資訊及LAR資訊中之任一種。生成過濾係數用的運算，在LSP資訊、PARCOR資訊及LAR資訊之特質上，關於各別之次元之運算會成為附屬於關於其他次元之運算性質的運算。因而，使用本態樣中所生成之過濾係數的過濾器，會比使用根據LPC資訊所生成的過濾係數的過濾器還變得更穩定的過濾器。加上，在傳送或存儲LPS資訊、PARCOR資訊或LAR資訊之系統上適用本態樣之情況，由於沒必要進行頻譜之再分析或參數轉換，所以可獲得良好的連接性。

本發明中之過濾，亦可在LPC領域、LSP領域、PARCOR領域及LAR領域中之任一領域上進行；亦即，在本發明中之過濾係數，亦可屬於LPC領域、LSP領域、PARCOR領域及LAR領域中之任一領域。若依據本發明之第二態樣，則首先藉由在其所屬領域內校正頻譜資訊來生成校正頻譜資訊，其次藉由將該校正頻譜資訊從該所屬領域轉換成LPC領域來生成過濾係數，再使用所獲得之過濾係數在LPC領域過濾。在當進行該校正時由於可導入各種之校正係數，所以若依據本態樣，則按照使用者所要求之過濾特性(加工語音合成信號之特性)可比習知還能自由調整過濾係數之生成動作。

若依據本發明之第三態樣，則為了使語音加工合成信號之語音素的峯值變小，可校正頻譜資訊。因而，在被容許之頻譜斜率之範圍內可獲得良好的語音素強調效果，並可

(請先閱讀背面之注意事項再填寫本頁)

家

訂

線

五、發明說明(12)

在語音素構造上不發生知覺準位之失真而獲得良好的語音素強調效果等。

作為校正之方法，第一，可揭示按照校正係數而比例分配關於語音輸入信號之頻譜資訊；及屬於和該頻譜資訊同一領域之參考資訊的方法。本方法可在頻譜資訊為LSP資訊時使用。若依據本方法，則可依參考資訊之設定方法逐次進行：使語音加工合成信號之頻譜平坦化的校正；將被固定之頻譜斜率給與語音加工合成信號的校正；將反映平均噪音頻譜之頻譜斜率附與語音加工合成信號的校正（亦即將噪音頻譜以外之語音頻譜做若干強調的校正）；將反映頻譜資訊成為過去之經歷的頻譜斜率附與語音加工合成信號的校正（亦即強調語音頻譜之變動部分的校正等）等。據此，可進行亮度之控制、資訊處理量之削減、了解性之改善等。又，若依據本方法，則藉由本發明之過濾器可一併實現其他附隨之過濾處理（例如固定的高通強調處理）。

作為校正之方法，第二，有關於語音輸入信號之構成頻譜資訊的多個之每一次元，在該頻譜資訊上乘以校正係數或其冪次方的方法。本方法可在頻譜資訊為PARCOR資訊及LAR資訊中之任一種時使用。即使依據本方法亦可獲得同樣的效果。另外，在頻譜資訊為PARCOR資訊時，使用頻譜資訊乘以校正係數之冪次方的方法，且使該冪次方附屬於頻譜資訊之次元。

作為校正之方法，第三，有在關於語音輸入信號之表現頻譜資訊的多個次元中，擴充相鄰接之次元間之距離的方

（請先閱讀背面之注意事項再填寫本頁）

表

訂

五、發明說明(13)

法。更具體而言，在相鄰接之次元間的距離在參考距離之下時，就將該距離擴充至該參考距離以上，之後為了使頻譜資訊全體之範圍可成為和擴充前同樣的範圍，而使該距離與全部之次元相關並做均等壓縮的方法。本方法可在頻譜資訊為LSP資訊時使用。即使依據本方法，在頻譜斜率平坦化之點上，亦可獲得同樣的效果。另外，第一及第三之校正方法，可組合在一起。在那時亦可選擇使用第一校正方法及第三校正方法，亦可兩者同時使用。

作為第一至第三之校正方法的實施形態，第一，要有換算表，用來將關於語音輸入信號之頻譜資訊對應校正頻譜資訊並存儲，再對應頻譜資訊所提供者而生成校正頻譜資訊；第二，要有神經網路，用來依學習而掌握將頻譜資訊轉換成校正頻譜資訊的能力，以便在提供關於語音輸入信號之頻譜資訊時可生成校正頻譜資訊。該等之換算表及神經網路，係以可區分獲得關於語音輸入信號之頻譜資訊的所屬領域並設在不相互重覆之多個範疇之每一個上，或者每一範疇藉由係數轉換等一面轉換其動作而一面使用者為較佳。如此一來，可實現依範疇分割而成之自適應型控制，同時可減低在範疇境界中的失真。另外，亦可每一範疇使用第一至第三之校正方法以外的校正方法。

若依據在LSP領域、PARCOR領域及LAR領域中之任一領域上進行過濾的第四態樣，則關於語音輸入信號之頻譜資訊可在其所屬領域內校正，且據此所得之校正頻譜資訊可當作過濾係數來使用。若依據本態樣，則由於不需要關於校

(請先閱讀背面之注意事項再填寫本頁)

表

訂

線

五、發明說明(14)

正頻譜資訊之領域轉換，所以可利用比習知少的構成要素來實現和習知同等之語音素強調效果等。

若依據本發明之第五態樣，則可在語音加工合成信號之語音素比語音合成信號之語音素還強調下進行過濾，若依據第六態樣，則可抑制在第五態樣中語音加工合成信號所附與的頻譜斜率。

若依據本發明之第七態樣，則可以多次元向量表現，並可基於屬於規定領域且關於語音輸入信號之頻譜資訊來生成語音合成信號，之後再基於頻譜資訊實行有關上述之各態樣的處理。若依據本發明之第八態樣，則可以多次元向量表現，並可基於屬於規定領域且關於語音輸入信號之第一頻譜資訊來生成語音合成信號，而第一頻譜資訊可轉換成屬於和其所屬領域相異之領域的第二頻譜資訊，之後再基於第二頻譜資訊實行有關上述之各態樣的處理。若依據本發明之第九態樣，則可以多次元頻譜來表現，並可基於屬於規定領域且關於語音輸入信號之第一頻譜資訊來生成語音合成信號，可藉由分析語音合成信號來生成第二頻譜資訊，之後再基於第二頻譜資訊實行有關上述之各態樣的處理。若依據本發明之第十態樣，則先進行第七至第九態樣的處理，就可實行分析語音輸入信號之頻譜資訊或第一頻譜資訊的生成，和頻譜資訊或第一頻譜資訊之存儲或傳送。

圖式之簡單說明

圖1及圖2皆顯示在本發明之較佳實施形態中有關利用

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明 (15)

LSP之實施形態之語音加工過濾器的構成方塊圖。

圖 3顯示語音分析/合成系統之一例構成的方塊圖。

圖 5為依分配而生成校正 LSP之方法的說明圖。

圖 4、圖 6、圖 7、圖 8、圖 10、圖 12、圖 13、圖 14、圖 15、圖 16、圖 17、圖 18及圖 19係各別顯示 LSP校正方法之一例的方塊圖。

圖 9及圖 11皆顯示在本發明之較佳實施形態中利用 LSP之實施形態之對數功率頻譜特性的圖表；其中圖 9顯示在圖 1之構成中使用依分配而生成校正 LSP之方法時的特性，而圖 11顯示在圖 2之構成中使用依相鄰次元間距離擴充而生成校正 LSP之方法時的特性。

圖 20及圖 21係各別顯示在本發明之較佳實施形態中有關在 LSP領域實行過濾之實施形態的語音加工過濾器的構成方塊圖。

圖 22顯示在本發明之較佳實施形態中有關利用 PARCOR之實施形態之語音加工過濾器的構成方塊圖。

圖 23顯示在本發明之較佳實施形態中利用 PARCOR之實施形態之對數功率頻譜特性的圖表。

圖 24及圖 25係各別顯示在本發明之較佳實施形態中有關在 PARCOR領域實行過濾之實施形態的語音加工過濾器的構成方塊圖。

圖 26顯示在本發明之較佳實施形態中有關利用 LAR之實施形態之語音加工過濾器的構成方塊圖。

圖 27顯示在本發明之較佳實施形態中利用 LAR之實施形

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(16)

態之對數功率頻譜特性的圖表。

圖 28 及 圖 29 係各別顯示在本發明之較佳實施形態中有關在 LAR 領域實行過濾之實施形態的語音加工過濾器的構成方塊圖。

圖 30 顯示在本發明之較佳實施形態中有關利用多個參數之實施形態之語音加工過濾器的構成方塊圖。

圖 31 顯示語音分析/合成系統之一例構成的方塊圖。

圖 32 顯示語音加工過濾器之使用方法的方塊圖。

圖 33、圖 34 及 圖 35 係各別顯示在文獻 1、文獻 2 及文獻 3 中所揭示之語音加工過濾器的構成方塊圖。

圖 36、圖 37 及 圖 38 係各別顯示在文獻 1、文獻 2 及文獻 3 中所揭示之語音加工過濾器之對數功率頻譜特性的圖表。

圖 39 顯示在文獻 4 中所揭示之語音加工過濾器的構成方塊圖。

較佳實施例之詳細說明

以下，參照圖面說明本發明之實施形態。另外，在和圖 31～圖 39 所示之習知技術相同或對應的構成構件上附相同符號而省略說明。又，在各實施形態共同的構成構件上附相同的符號而省略說明。

a) 利用 LSP 之實施形態

圖 1 及 圖 2 顯示在有關本發明之過濾器 203 的較佳實施形態中，輸入 LSP 作為頻譜資訊的二個實施形態。圖 1 所示之實施形態，係除了過濾器 204 及 205 之外，還具備 LSP 校正器 216 及 217 暨 LSP/LPC 轉換器 218 及 219。又，圖 2 所示之實

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

85年11月18日 修正
補充

五、發明說明(17)

施形態，係除了過濾器204及205之外，還具備LSP校正器216及LSP/LPC轉換器218。

該等之實施形態，係可以圖32所示之構成的合成單元200或圖3所示之構成的合成單元200來使用。亦即，在使用輸出LSP之解碼器201作為頻譜資訊時，如圖32所示，可將解碼器201之輸出直接供給過濾器203，對此在使用輸出LSP以外之資訊的解碼器201作為頻譜資訊時，如圖3所示，在利用轉換器215將解碼器201之輸出轉換成LSP領域(LSP domain)之後，就有必要供給過濾器203。另外，亦可能將轉換器215組入解碼器201或合成器202中。

LSP校正器216及217，係從解碼器201或轉換器202A，輸入多次元頻譜LSP ω_i ，藉由遵循規定之方法校正該 ω_i ，來生成已校正之LSP ω_{h1i} 及 ω_{h2i} 。LSP/LPC轉換器218及219，各別藉由將 ω_{h1i} 及 ω_{h2i} 從LSP領域轉換成LPC領域，生成已校正之 α 參數 α_{1i} 及 α_{2i} 。過濾器204及205，各別將 α_{1i} 及 α_{2i} 作為過濾係數來使用並依序過濾語音合成信號。其結果，可從過濾器205輸出語音加工合成信號。在此，當將過濾器204及205之傳遞函數各別表示成 $1/A_1(z)$ 及 $A_2(z)$ 時，圖1之過濾器203的傳遞函數 $H(z)$ ，係成為下式

$$H(z) = A_2(z) / A_1(z) \dots \dots \dots (3)$$

所表示的函數，而圖2之過濾器203的傳遞函數 $H(z)$ ，係成為下式

$$H(z) = 1 / A_1(z) \dots \dots \dots (4)$$

(請先閱讀背面之注意事項再填寫本頁)

裝
訂
線

經濟部中央標準局員工消費合作社印製

85年11月8日 修正
補充^{A7}

五、發明說明(18)

所表示的函數。

如此在本發明之利用LSP之實施形態中，藉由校正當作頻譜資訊輸入的LSP ω_i ，並將已校正之LSP ω_{h1i} （及 ω_{h2i} ）從LSP領域轉換成LPC領域，生成已校正之 α 參數之過濾係數 α_{1i} 及 α_{2i} 。具有如此構成之利用LSP之實施形態的第一優點，是過濾器203之特性比較穩定。例如，在文獻1～文獻3所揭示的技術中，各個 i 由於必須獨立進行 α_{1i} 及 α_{2i} 的生成運算，所以過濾器203的特性容易造成不穩定。相對於此，若依據本發明之利用LSP之實施形態，則關於LSP一般會成立

$$0 < \omega_1 < \omega_2 < \dots < \omega_p < \pi \dots \quad (5)$$

之順序關係接著由於 α_{1i} 及 α_{2i} 之生成運算並非各個 i 獨立，所以不容易使過濾器203之特性穩定化。

利用LSP之實施形態的第二優點，是很容易適用在當作頻譜資訊傳送或存儲LSP的系統上。尤其是近年所開發之語音編碼解碼系統大多將LSP當作頻譜資訊使用。本發明之利用LSP之實施形態，很容易適用在此種語音編碼解碼系統上。亦即，由於不需要頻譜之再分析或參數轉換，故和如文獻4所示輸入Mel-倒頻譜並據此決定過濾係數的習知技術相異，可對該種系統獲得良好的連接性。

又，從上述說明中可明白，在本發明之利用LSP之實施形態中之過濾器203的傳遞函數 $H(z)$ ，係藉由如何進行為了獲得過濾係數 α_{1i} 及 α_{2i} 之LSP校正運算及LSP/LPC轉換運算來控制。作為LSP校正運算之較佳方法，可各別揭

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

五、發明說明(19)

示第一，分配校正，第二相鄰次元間距離擴充。

其中分配校正，係將滿足 $0 \leq \nu \leq n < 1$ 之校正係數 ν 、 n 作為分配比使用，以分配 ω_i 的方法。圖 1 之構成中實施該方法時，LSP 校正器 216 及 217，例如如圖 4 所示，係構成具有分配運算器 220 及斜率設定部 221 的機能構造。分配運算器 220，係按照下面之分配式

$$\omega_{h1i} = \omega_i \times (1 - \nu) + \omega_{fi} \times \nu \text{ 或}$$

$$\omega_{h2i} = \omega_i \times (1 - n) + \omega_{fi} \times n \dots\dots\dots (6)$$

但是 $i = 1, 2, \dots\dots\dots p$

生成 ω_{h1i} 或 ω_{h2i} 。斜率設定部 221，係根據預測次數 p 在分配運算器 220 上設定 ω_{fi} 。另外，以 LSP 校正器 216 及 217 使用的 ω_{fi} 亦可為相異值。又，亦可將分配之 ω_i 的校正適用在圖 2 之構成上。

分配校正之第一優點，係可獲得良好的語音素強調效果。亦即，由於藉由分配校正所生成的 ω_{h1i} 及 ω_{h2i} ，若從 LSP 領域轉換成 LPC 領域時語音素會衰微，所以可獲得良好的語音素強調效果。在此所謂“語音素會衰微”，係指“語音素之峰值會變小”，亦即“頻譜之山谷構造在留下某種程度之下頻譜特性會平坦化”。

分配校正之第二優點，係可以頻率帶通變更語音合成信號之加工程度等，並可進行對應使用者要求之自由度高的特性設計。尤其是藉由加上 ν 及 n 並按照使用者之要求設計 ω_{fi} ，就可實現各種特性的過濾器 203。該自由度之高度，係關係到可容易在被容許的頻譜斜率之範圍內獲得在

(請先閱讀背面之注意事項再填寫本頁)

表

訂

經濟部中央標準局員工消費合作社印製

五、發明說明(20)

習知以上之良好的語音素強調效果之效果。

在 ωf_i 之設定方法上有幾個方法。第一被揭示者，是將表示平坦頻譜之 LSP 設定在 ωf_i 的方法。按照該方法所實現的斜率設定部 221，係按照下式

$$\omega f_i = \pi \times i / (p+1) \dots \dots \dots (7)$$

設定 ωf_i ，以便 ωf_i 之相鄰次元間距離 ($= \omega f_i - \omega f_{i-1}$) 可成為 $\pi / (p+1)$ 之一定值。圖 5 係以 ωhli 之生成為例概念性顯示按照式 (7) 設定 ωf_i 時的分配校正動作。但是在此係假設 $p=10$ 。在該方法中，具有斜率設定部 221 之機能簡單之優點。

第二被揭示者，係將表示固定斜率頻譜之 LSP 設定在 ωf_i 的方法。按照該方法所實現的斜率設定部 221，係如圖 6 所示輸入 ω_i ，並按照式 (7) 之右邊附加 ω_i 之一次項的下式

$$\omega f_i = \pi \times i / (p+1) + \delta (\omega_i) \dots \dots \dots (7a)$$

設定 ωf_i ，以便可按照 ω_i 線性增加或減少 ωf_i 之相鄰次元間距離。此情況分配校正動作會變成如何之動作，從上述之說明及圖 5 之揭示，身為從業者就可容易著手處理。該方法中，第一，在過濾器 203 之特性上由於可以附與大致固定的斜率，所以具有可依 ωf_i 之調整來控制亮度的優點。第二、由於可將普通之語音素強調處理中前後所進行的固定性高通強調處理之特性包含在該過濾器 203 之傳遞函數 $H(z)$ 中，所以具有可削減處理量的優點。

第二被揭示者，係將表示平均噪音頻譜的 LSP 利用分配

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(21)

處理等校正的LSP，設定在 ωf_i 的方法。按照該方法所實現的斜率設定部221，如圖7所示，可按照下式

$$\omega f_i = \omega_i' \times (1 - \nu') + \omega_i' \times \nu' \text{ 或}$$

$$\omega f_i = \omega_i' \times (1 - \eta') + \omega_i' \times \eta' \dots\dots (7b)$$

但是， $i=1, 2, \dots, p$

並藉由根據分配比 ν' 或 η' 校正表示平均噪音頻譜的LSP ω_i' 設定 ωf_i 。該方法之優點，由於可若干強調噪音頻譜以外語音頻譜，所以了解性會變佳。另外， ω_i' 係藉由利用平均運算器223來平均化，而可獲得利用圖7所示之判定器222來判定噪音區間的 ω_i 。又，施行於 ω_i' 上之校正處理，以不將那麼極端的頻譜變動提供給語音加工合成信號之下設定為佳。例如，若事先使 ωf_i 衰微，則可使極端的頻譜變動發生在語音加工合成信號上。

第四被揭示者，係將利用分配處理等校正在動作開始後至現在或者在過去規定期間內之 ω_i 平均值的LSP，設定在 ωf_i 上的方法。按照該方法所實現的斜率設定部221，如圖8所示係利用平均運算器223求出過去LSP ω_i 之平均值 ω_i' ，再根據該 ω_i' 及分配比 ν' 或 η' ，並按照式(7b)設定 ωf_i 。該方法之優點，由於可強調語音之頻譜變動量，所以了解性會佳。另外，即使在實施該方法之際，亦以為了不將那麼極端的頻譜變動提供給語音加工合成信號上而校正 ω_i' 等的考量為較佳。

圖9顯示按照式(6)及(7)校正 ω_i 時之圖1過濾器203的對數功率頻譜特性。圖中A~D依序為合成器202之特性 $=1/$

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(22)

$A(z)$ 、過濾器 204 之特性 $=1/A_1(z)$ 、過濾器 205 之逆特性 $=1/A_2(z)$ 、過濾器 203 之傳遞函數 $H(z)=A_2(z)/A_1(z)$ 。另外， $\nu=0.5$ ， $\eta=0.8$ 。該圖之特性 D 比圖 36 之特性 D，係相當於在留下某種程度之下使頻譜之山谷構造平坦化的特性。如此在圖 9 中可比圖 36 還獲得良好的語音素強調效果。又，該圖之特性 D 比圖 37 之特性 D，關於頻譜之山谷構造的失真會少。更且，在該圖之特性 D 上，不會出現在圖 38 之特性 B 及 C 所觀測到的最低頻率語音素之移動及中央二語音素之單一化的二種現象。另外，即使使用在 LSP 領域具有使語音素衰微效果的其他處理來取代分配處理，亦可產生同樣的優點。

又，發明者係將從關於按照式 (6) 及 (7) 所示之方法校正 ω_i 之實施形態的過濾器 203 中所獲得的加工合成音，和從關於前述各習知技術之過濾器 203 中所獲得的加工合成音做聽覺比較。其結果，確認本實施形態之語音加工過濾器比較能抑制亮度劣化，且亦不會發生獨特之失真者或音色之不穩。

LSP 校正運算之第二較佳方法的相鄰次元間距離擴充，如圖 10 所示可藉由為了在 S_i 空間 (space) 上之相鄰次元間距離 S_i-S_{i-1} 變得比在 ω_i 空間上之相鄰次元間距離 $\omega_i-\omega_{i-1}$ (參照圖 5) 還寬而將 ω_i 從 ω_i 空間轉至 S_i 空間的擴充器 224；及根據 ω_i 及 S_i 求出 ω_{h1i} 的均等壓縮器 225 來實施。 S_i 亦和 ω_i 同樣希望留意到多次元向量之點。在圖 2 之構成中實施該方法時，均等壓縮部 225 係按照下式

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(23)

$$\omega_{h1i} = (\omega_i + S_i) / S_{p+1} \times \pi \dots\dots\dots (8)$$

求出 ω_{h1i} ，擴充器 224 係按照下式

$$S_i = S_{i-1} + \max(\omega_i - \omega_{i-1}, th) \dots\dots (9)$$

但是， $i = 1, 2, \dots, p+1$

$$\omega_0 = 0, \omega_{p+1} = \pi, S_0 = 0$$

th: 臨限值

求出 S_i 。

如此，相鄰次元間距離擴充係為在藉由式(9)之最右邊第2項所定義下，根據 $\omega_i - \omega_{i-1}$ 和 th 之比較結果，在第 $i-1$ 之次元和第 i 之次元間至少確保距離 th 的處理。藉由該處理在 S_i 空間上關於 $i+1$ 以上之次元的 LSP 可一次移位至只相當 $th - (\omega_i - \omega_{i-1})$ 量上的位置。又，式(8)之右邊所包含的因素 π / S_{p+1} ，係為對應在 ω_i 空間上之 LSP 範圍 $0 \sim \pi$ 和在 S_i 空間上之 LSP 範圍 $0 \sim S_{p+1}$ 的比，均等壓縮相鄰次元間距離用的因素。另外，本發明並非只藉由該定義式做限定解釋，若是擴充相鄰次元間距離小的部份則亦可採用其他之定義式。又，亦可將相鄰次元間距離擴充之 ω_i 的無轉換適用在圖1之構成上。此情況可更增加過濾器 203 之特性自由度。

圖11顯示將該方法適用在圖2之過濾器 203 上時之對數功率頻譜特性。圖中 A~C 依序為合成器 202 之特性 $= 1/A(z)$ 、 $th = 0.3$ 時過濾器 204 之特性 $= 1/A1(z; th = 0.3)$ 、 $th = 0.4$ 時過濾器 204 之特性 $= 1/A1(z; th = 0.4)$ 。從該圖可明白，若依據本方法則只要以過濾器 204 (亦即過濾器 205 及非對應

(請先閱讀背面之注意事項再填寫本頁)

裝
訂

經濟部中央標準局員工消費合作社印製

五、發明說明(24)

此之構成)就可獲得比圖36及圖37還沒有特別遜色的特性；亦即以比習知少的過濾次數亦可實現良好的語音加工過濾特性，又，以更少的構成要素就可實現和習知同等的語音素強調效果。又，發明者係將在本實施形態所獲得的加工合成音和利用各習知技術所獲得的加工合成音做聽覺比較。其結果，確認在使用本實施形態之語音加工過濾器時可獲得不會比習知還遜色的音質。

分配校正及相鄰次元間距離擴充之二種校正方法，由於不是相互排斥者，所以可兩者併用。例如，LSP校正器216及217中亦可一方實行分配校正，而另一方實行相鄰次元間距離擴充。或者，如圖12所示，亦可作為藉由轉換機構228及229選擇使用分配校正 ω_1 之分配校正器226及擴充LSP之相鄰次元間距離擴充器227的構成。分配校正器226亦可為前述圖4，圖6、圖7及圖8中之任一構成。或者，如圖13所示亦可為級聯連接分配校正器226及相鄰次元間距離擴充器227的構成。若依據以該等單一的LSP校正器併用分配校正器226及相鄰次元間距離擴充器227的構成，則可更增加過濾器203之特性自由度。另外，亦可交換圖13中之分配校正器226及相鄰次元間距離擴充器227的順序。當然亦可組合分配校正及相鄰次元間距離擴充的雙方或者其他之處理。

更且，亦可利用LSP校正器216及217實行 ω_1 自適應的處理(ω_1 adaptive process)。作為藉由分配校正 ω_1 的處理進行 ω_1 自適應的方法而言，例如有為了以免相互

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

89年11月18日 修正
補充

五、發明說明(25)

重覆在 ω_i 空間上而分割成多個部分空間(以下稱為範疇)，且各個範疇皆準備(轉換) ν 、 n 的方法。該情況，亦可如以LSP校正器216-1(或217-1)對應第一範疇，LSP校正器216-2(或217-2)對應第二範疇，...LSP校正器216-N(或217-2)對應第N範疇在各個範疇設置LSP校正器(參照圖14)，或者亦可準備單一LSP校正器216(或217)利用校正係數轉換器230按照範疇或 i 來轉換 ν 、 n (參照圖15)。 ω_i 自適應處理之優點，係在於有關增強語音素強調的話會發生失真的範疇，可進行只對該範疇減弱語音素強調的柔軟性處理；據此就可均等改善過濾器203的特性。另外，由於 ω_i 是多次元向量，所以這裏所謂的範疇是指多次元向量空間。

在LSP校正器216及217中之 ω_i 校正處理，如圖16所示係以利用換算表231來實現為較佳。亦即，事先準備對應 ω_i 和 ω_{h1i} 或 ω_{h2i} 的換算表231，在提供 ω_i 時LSP校正器216或217會輸出對應該 ω_i 的 ω_{h1i} 或 ω_{h2i} 。利用換算表231之優點，係可縮短處理時間，而此優點在使用比較複雜式作為 ω_i 校正處理的原理式時會變得比較顯著。

在LSP校正器216及217中之 ω_i 校正處理，如圖17所示，亦可利用預先學習(learn)藉由式(6)等所提供 ω_i 校正特性完成後的神經網路232來實現。利用神經網路232的第一優點係可縮短處理時間，而此優點在使用比較複雜式作為 ω_i 校正處理的原理式時會變得比較顯著。利用神經網路

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(26)

232的第二優點，係比利用換算表231時，由於沒有必要記憶換算表231所以可節省儲存容量。

利用換算表231的方法或利用神經網路232的方法，各別可以各個範疇準備 ν 、 n 的方法(前述)：亦即，如圖18所示，可將圖16所示之構成或圖17所示之構成和圖14所示之構成組合，或者如圖19所示，可將圖16所示之構成或圖17所示之構成和圖15所示之構成組合。若依據如此所組合的方法，則在範疇境界中之失真會變少。所謂“範疇境界中之失真”，是指在某個範疇和其他範疇之境界近旁僅變動 ω_i 的結果，由於 ν 、 n 會急速變化(亦即校正偶而變強偶而變弱)，所以是在語音加工合成信號或語音半加工合成信號中所出現的失真。尤其是在 ω_i 空間之範疇分割較粗(rough)時，該失真就會更易顯著。若依據圖18及圖19所示的構成，則即使該分割為若干粗，藉由換算表231之校正或神經網路232之學習，亦可比較容易補償該粗度。

本發明之利用LSP之實施形態，並非只限定進行LPC合成過濾及LPC反相過濾的構成，亦可使用LPC以外之參數作為過濾係數。例如，如圖20及圖21所示，亦可使用以原狀將 ω_{h1i} (及 ω_{h2i})作過濾係數使用的LSP合成過濾器233(及LSP反相過濾器234)來實現本發明，該構成之優點係可廢止LSP/LPC轉換器218及219。

b) 利用PARCOR之實施形態

圖22顯示輸入PARCOR作為頻譜資訊的實施形態。該實施形態除了LPC合成過濾器204及LPC反相過濾器205之外，選

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(27)

具有 PARCOR 校正器 235 及 236 暨 PARCOR/LPC 轉換器 237 及 238。
 PARCOR 校正器 235，係從解碼器 201 或轉換器 215 輸入 PARCOR ϕ_i 作為頻譜資訊，藉由校正該 ϕ_i 來生成校正 PARCOR ϕ_{h1i} 。PARCOR 校正器 236 亦同樣而生成校正 PARCOR ϕ_{h2i} 。PARCOR/LPC 轉換器 237 係藉由將 ϕ_{h1i} 從 PARCOR 領域轉換成 LPC 領域來生成 LPC 合成過濾器 204 之過濾係數 α_{1i} 。PARCOR/LPC 轉換器 238，亦藉由將 ϕ_{h2i} 從 PARCOR 領域轉換成 LPC 領域來生成 LPC 反相過濾器 205 之過濾係數 α_{2i} 。

PARCOR 校正器 235 及 236，例如係使用滿足 $0 \leq n \leq \nu \leq 1$ 的校正係數 ν 及 n ，並按照下式

$$\begin{aligned}\phi_{h1i} &= \phi_i \times \nu^{(i \times i)} \\ \phi_{h2i} &= \phi_i \times n^{(i \times i)} \dots \dots \dots (10)\end{aligned}$$

但是， $i=1, 2, \dots, P$

來生成 ϕ_{h1i} 及 ϕ_{h2i} 。藉由此種校正就可在 PARCOR 領域上使語音素衰微。

因而，若依據本實施形態，則可獲得和前述利用 LPC 之實施形態同樣的特性改善效果（改善語音素強調效果，或改善該強調程度之調整能力等），又，可按照使用者要求來自由操作，設定過濾器 203 的特性。當然本發明並非只藉由式 (10) 來限定，亦可採用在 PARCOR 領域上使語音素衰微的其他處理。更且，關於 PARCOR 之方面由於成立下式

$$-1 < \phi_i < 1 \dots \dots \dots (11)$$

，所以和利用 LSP 之實施形態同樣，即使在利用 PARCOR 之

(請先閱讀背面之注意事項再填寫本頁)

訂

85年1月修正補充

五、發明說明(28)

實施形態中過濾器203亦會比習知還穩定。加上，在適用傳送或存儲PARCOR作為頻譜資訊的系統時，由於不需要頻譜之再分析或參數轉換，所以可獲得良好的連接性。

圖23顯示圖22之過濾器203的對數功率頻譜特性。該圖中A~D係依序為合成器202之特性 $=1/A(z)$ 、過濾器204之特性 $=1/A_1(z)$ 、過濾器205之逆特性 $=1/A_2(z)$ 、過濾器203之特性 $=A_2(z)/A_1(z)$ 。但是， $\nu=0.98$ ， $\eta=0.9$ 。從圖23和圖36之比較中可明，若依據本實施形態，則比文獻1所示之構成，可稍強顯現頻譜之山谷構造。又，發明者藉加工合成音之聽覺比較，確認在使用本實施形態之過濾器203時不會發生獨特之失真音或音色之不穩定而可獲得良好的語音素強調效果。

以和利用LSP之實施形態同樣的觀點來看，可構成該利用PARCOR之實施形態的細部，對一個從業者而言，從本案之揭示中就自可明白。又，如圖24所示省略LPC反相過濾及關於此之構成要素，或如圖25所示設置PARCOR合成過濾器239及PARCOR反相過濾器240並形成以校正PARCOR ϕ_{h1} 及 ϕ_{h2} 作為其過濾係數使用的構成，對一個從業者而言，若根據本案之揭示則亦很容易。

c) 利用LAR之實施形態

圖26顯示輸入LAR作為頻譜資訊的實施形態。該實施形態，除了LPC合成過濾器204及205以外，還具有LAR校正器241及242暨LAR/LPC轉換器243及244。LAR校正器241，係從解碼器201或轉換器215輸入LAR ϕ_i 作為頻譜資訊，再

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

89年11月18日 修正
補充

五、發明說明(29)

藉由校正該 ϕ_i 生成校正 LAR ϕ_{h1i} 。LAR校正器 242，亦同樣而生成校正 LAR ϕ_{h2i} 。LAR/LPC轉換器 243係藉由將 ϕ_{h1i} 從 LAR領域轉換成 LPC領域來生成 LPC合成過濾器 204 的過濾係數 α_{1i} 。LAR/LPC轉換器 244，亦藉由將 ϕ_{h2i} 從 LAR領域轉成 LPC領域來生成 LPC反相過濾器 205的過濾係數 α_{2i} 。

LAR校正器 241及 242，例如，係使用滿足 $0 \leq n \leq \nu < 1$ 的校正係數 ν 及 n ，並按照下式

$$\begin{aligned} \phi_{h1i} &= \phi_i \times \nu \\ \phi_{h2i} &= \phi_i \times n \end{aligned} \dots\dots\dots (12)$$

但是， $i=1, 2, \dots\dots\dots, P$

，來生成 ϕ_{h1i} 及 ϕ_{h2i} 。藉由此種校正，可在 LAR領域上使語音素衰微。

因而，若依據本實施形態，則可獲得和前述之利用 LPC 之實施形態及利用 PARCOR之實施形態同樣的特性改善效果(改善語音素強調效果，或改善該強調程度之調整能力等)

，又，可按照使用者要求來自由操作，設定過濾器 203 的特性。當然本發明並非只受限定於式(12)，若為在 LAR領域上具有使語音素衰微之效果的處理，則亦可採用其他處理。更且，在利用 LAR之實施形態中由於可經常保證過濾器 203 的穩定性，所以過濾器 203 會比習知還穩定。加上，在適用傳送或存儲 LAR作為頻譜資訊的系統時，由於不需要頻譜之再分析或參數轉換，所以可獲得良好的連接性。

圖 27 顯示圖 26 之過濾器 203 的對數功率頻譜特性。該圖

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

經濟部中央標準局員工消費合作社印製

85年11月18日 修正
補充

A7
B7

五、發明說明 (30)

中 A~D 係依序為合成器 202 之特性 $=1/A(z)$ 、過濾器 204 之特性 $=1/A_1(z)$ 、過濾器 205 之逆特性 $=1/A_2(z)$ 、過濾器 203 之特性 $=A_2(z)/A_1(z)$ 。但是， $\nu = 0.9$ ， $\eta = 0.7$ 。從圖 27 和圖 36 之比較中可明白，若依據本實施形態，則比文獻 1 所示之構成，在某種程度留下頻譜之山谷構造之下可使頻譜平坦化，因而可獲得良好的語音素強調效果。又，即使比圖 37，圖 27 有關頻譜之山谷構造的失真亦會少。更且，從圖 38 之特性 B 和特性 C 之比較中可明白的中央 2 語音素合為一的現象，不會在圖 27 中出現。發明者藉由加工合成音之聽覺比較，確認在使用本實施形態之過濾器 203 時亦不發生獨特之失真音或音色之不穩定而可獲得良好的語音素強調效果。

在和利用 LSP 之實施形態或利用 PARCOR 之實施形態同樣的觀點來看，可構成該利用 LAR 之實施形態的細部，對一個從業者而言，從本案之揭示中自可明白。又，如圖 28 所示省略 LPC 反相過濾器及關於此之構成要表，或如圖 29 所示設置 LAR 合成過濾器 246 及 LAR 反相過濾器 247 並形成以校正 LAR ψh_{1i} 及 ψh_{2i} 作為其過濾係數使用的構成，對一個從業者而言，若根據本案之揭示，則亦可很容易。

d) 補充

選擇性組合上述之利用 LSP 之實施形態、利用 PARCOR 之實施形態及利用 LAR 之實施形態，從業者根據本案之揭示就可容易完成。又，將本發明之實施形態和習知之 LPC 利用裝置組合，從業者根據本案之揭示亦可容易完成。該等

(請先閱讀背面之注意事項再剪本頁)

裝

訂

線

五、發明說明(31)

之各種組合，有助於實現在各實施形態單獨中所無法實現之自由度高的過濾器203。例如，如圖30所示在各以和文獻1同樣的方法，決定過濾器204之過濾係數 α_{1i} ，或者以和利用PARCOR之實施形態同樣的方法，決定過濾器205之過濾係數 α_{2i} 的構成中，頻譜斜率會比圖36之特性D還更少，並可獲得語音素近旁之失真比圖37之特性D還更少的特性之過濾器203。

又，亦可採用在過濾器203之前或後者和過濾器203並排插入另一過濾器，再進行間距強調處理、高通強調處理、語音素強調處理等的構成。

(請先閱讀背面之注意事項再填寫本頁)

不

訂

四、中文發明摘要(發明之名稱： 語音加工或強調用之過濾器暨使用)

該過濾器之各種裝置、系統及方法

本發明之目的係在於提供一種語音加工或強調用之過濾器暨使用該過濾器之各種裝置、系統及方法。藉由過濾語音合成信號來生成語音加工合成信號。根據以多次元向量所表現的頻譜資訊來決定過濾係數，以便語音加工合成信號之語音素特徵可比語音合成信號強調且可按照頻譜資訊來強調。可使用LSP資料、PARCOR資訊及LAR資訊中之任一種資訊來作為頻譜資訊。可提高為了改善語音合成信號中之量化噪音的聽覺性抑制及該語音合成信號之了解性等而所採用的語音加工過濾器之特性自由度。在所容許之頻譜斜率之範圍內，不會產生知覺準位的失真，並可獲得良好的語音素強調效果。

英文發明摘要(發明之名稱：)

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

六、申請專利範圍

1. 一種過濾器，其為具備：

過濾機構，用來藉由以過濾係數所限定的傳遞函數來過濾語音合成信號，藉以生成語音加工合成信號；及

過濾係數生成機構，係以多次元向量表現，根據屬於規定領域且關於語音輸入信號的頻譜資訊，來生成上述過濾係數，以便可按照上述頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號的語音素特徵；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一種。

2. 如申請專利範圍第1項之過濾器，其中上述過濾係數屬於LPC領域。

3. 如申請專利範圍第2項之過濾器，其中上述過濾係數生成機構包含有：

校正機構，係藉由在上述規定領域內校正上述頻譜資訊來生成校正頻譜資訊；及

將上述校正頻譜資訊從上述規定領域轉換成LPC領域藉以生成過濾係數的機構。

4. 如申請專利範圍第3項之過濾器，其中上述校正機構，包含有校正上述頻譜資訊的平坦化機構，以便使上述語音加工合成信號之語音素峰值變小。

5. 如申請專利範圍第4項之過濾器，其中，

上述頻譜資訊為LSP資訊，

而上述平坦化機構包含有分配機構，係藉由按照校正係數來分配屬於和上述頻譜資訊相同領域的參考資訊，和該

(請先閱讀背面之注意事項再填寫本頁)

訂

六、申請專利範圍

頻譜資訊，藉以生成上述校正頻譜資訊者。

6. 如申請專利範圍第5項之過濾器，其中上述分配機構，係用來分配上述參考資訊和上述頻譜資訊，以便使上述語音加工合成信號之頻譜得以平坦化。

7. 如申請專利範圍第5項之過濾器，其中上述分配機構，係用來分配上述參考資訊和上述頻譜資訊，以便被固定之頻譜斜率得以附與上述語音加工合成信號上。

8. 如申請專利範圍第5項之過濾器，其中上述分配機構，係用來分配上述參考資訊和上述頻譜資訊，以便反映平均噪音頻譜的頻譜斜率得以附與上述語音加工合成信號上。

9. 如申請專利範圍第5項之過濾器，其中上述分配機構，係用來分配上述參考資訊和上述頻譜資訊，以便上述頻譜資訊反映過去所經歷的過程之頻譜斜率得以附與上述語音加工合成信號上。

10. 如申請專利範圍第4項之過濾器，其中，
上述頻譜資訊為PARCOR資訊及LAR資訊中之任一種，
而上述平坦化機構，包含有藉由構成上述頻譜資訊之多個之每一次元，在該頻譜資訊上乘以校正係數或其冪次方，來生成上述校正頻譜資訊的機構。

11. 如申請專利範圍第10項之過濾器，其中，
上述冪次方屬於上述次元。

12. 如申請專利範圍第3項之過濾器，其中，
上述頻譜資訊為LSP資訊，

而上述校正機構，包含有距離擴充機構，係藉由在表現

(請先閱讀背面之注意事項再填寫本頁)

第

訂

號

六、申請專利範圍

上述頻譜資訊的多個次元中，擴充相鄰接之次元間距離，而生成上述校正頻譜資訊。

13. 如申請專利範圍第12項之過濾器，其中，

上述距離擴充機構，包含有：

擴充機構，係在上述相鄰接之次元間的距進在參考距離以下時，將該距離擴充至該參考距離以上者；及

壓縮機構，係在利用上述擴充機構擴充上述相鄰接之次元間的距離之後，關於所有之次元均等壓縮上述距離，以便上述頻譜資訊全體之範圍變成和擴充前為同樣的範圍。

14. 如申請專利範圍第3項之過濾器，其中，

上述頻譜資訊為LSP資訊，

而上述校正機構，包含有：

分配機構，係按照校正係數來分配屬於和上述頻譜資訊相同領域之參考資訊，和該頻譜資訊者；

距離擴充機構，係在表現上述頻譜資訊之多個次元中，擴充相鄰接之次元間的距離者；及

轉換機構，係藉由選擇性使用上述分配機構及上述擴充機構中之任一機構來生成上述校正頻譜資訊者。

15. 如申請專利範圍第3項之過濾器，其中，

上述頻譜資訊為LSP資訊，

而上述校正機構，包含有：

分配機構，係按照校正係數來分配屬於和上述頻譜資訊相同領域之參考資訊，和該頻譜資訊者；

距離擴充機構，係在表現上述頻譜資訊之多個次元中，

(請先閱讀背面之注意事項再填寫本頁)

訂

六、申請專利範圍

擴充相鄰接之次元間的距離者；及

級聯連接機構，係併用上述分配機構及上述擴充機構來生成上述校正頻譜資訊者。

16. 如申請專利範圍第3項之過濾器，其中上述校正機構，包含有換算表，係對應上述校正頻譜資訊來存儲上述頻譜資訊者，而該換算表係按照上述頻譜資訊所提供者而生成該生成之校正頻譜資訊。

17. 如申請專利範圍第3項之過濾器，其中上述校正機構，包含有神經網路，係藉由學習而掌握將上述頻譜資訊轉換成上述校正頻譜資訊的能力者，而該神經網路，係按照上述頻譜資訊所提供者而生成該生成的校正頻譜資訊。

18. 如申請專利範圍第3項之過濾器，其中，
上述校正機構包含有多個範疇特定校正機構，係因區分上述規定領域而所得且以不相互重覆之多個之每一範疇所設者；

各範疇特定校正機構，包含有：

在上述範疇內校正上述頻譜資訊藉以生成校正頻譜資訊的機構；及

將上述校正頻譜資訊從上述規定領域轉換成LPC領域藉以生成過濾係數的機構。

19. 如申請專利範圍第18項之過濾器，其中上述校正機構，包含有多個換算表，係以上述多個之每一範疇所設且對應上述校正頻譜資訊而存儲上述頻譜資訊者，而各換算表，係按照屬於對應之範疇的上述頻譜資訊所提供者而生

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

六、申請專利範圍

成該生成的校正頻譜資訊。

20.如申請專利範圍第18項之過濾器，其中上述校正機構，包含有多個神經網路，係以上述多個之每一範疇所設且藉由學習來掌握將上述頻譜資訊轉換成上述校正頻譜資訊的能力，而各神經網路係按照屬於對應之範疇的上述頻譜資訊所提供者來生成該生成的校正頻譜資訊。

21.如申請專利範圍第3項之過濾器，其中上述校正機構包含有：

藉由隨著校正係數在上述規定領域內校正上述頻譜資訊來生成校正頻譜資訊的機構；

藉由將上述校正頻譜資訊從上述規定領域轉換成LPC域來生成過濾係數的機構；及

按照上述頻譜資訊是屬於因區分上述規定領域而所得且在不相互重覆之多個範疇中之哪一個範疇，而藉以調整上述校正係數的機構。

22.如申請專利範圍第21項之過濾器，其中上述校正機構，包含有換算表，係用來對應上述校正頻譜資訊存儲上述頻譜資訊，而該換算表係按照上述頻譜資訊所提供者來生成該生成的校正頻譜資訊。

23.如申請專利範圍第21項之過濾器，其中上述校正機構，包含有神經網路，係藉由學習來掌握將上述頻譜資訊轉換成上述校正頻譜資訊的能力，而該神經網路，係按照上述頻譜資訊所提供者來生成該生成的校正頻譜資訊。

24.如申請專利範圍第1項之過濾器，其中上述過濾係數

(請先閱讀背面之注意事項再填寫本頁)

表

訂

承

六、申請專利範圍

是屬於LSP領域、PARCOR領域及LAR領域中之任一領域。

25.如申請專利範圍第24項之過濾器，其中上述過濾係數生成機構，包含有：

藉由在上述規定領域內校正上述頻譜資訊來生成校正頻譜資訊的校正機構；及

將上述校正頻譜資訊作為過濾係數而供給過濾機構的機構。

26.如申請專利範圍第1項之過濾器，其中過濾機構包含有合成過濾器，係語音加工合成信號之語音素為了比語音合成信號之語音素還能強調而實現上述傳遞函數之分母者。

27.如申請專利範圍第26項之過濾器，其中過濾機構，更包含有反相過濾器，係利用上述合成過濾器來抑制在上述語音加工合成信號上所附與的頻譜斜率者。

28.一種語音合成裝置，其為包含有：

以多次元向量表現，再根據屬於規定領域且關於語音輸入信號之頻譜資訊來生成語音合成信號的機構；

藉由以過濾係數所限定的傳遞函數過濾語音合成信號藉以生成語音加工合成信號的機構；及

為了按照上述頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號語音素特徵，而根據上述頻譜資訊生成上述過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

29.一種語音合成裝置，其為包含有：

(請先閱讀背面之注意事項再填寫本頁)

訂

六、申請專利範圍

以多次元向量表現，再根據屬於規定領域且關於語音輸入信號的第一頻譜資訊來生成語音合成信號的機構；

將上述第一頻譜資訊轉換成屬於和上述規定領域相異的領域之第二頻譜資訊的機構；

藉由以過濾係數所限定的傳遞函數過濾語音合成信號藉以生成語音加工合成信號的機構；及

為了按照上述第二頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號的語音素特徵，而根據上述第二頻譜資訊生成上述過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

30. 一種語音合成裝置，其為包含有：

以多次元向量表現，再根據屬於規定領域且關於語音輸入信號的第一頻譜資訊來生成語音合成信號的機構；

藉由分析上述語音合成信號生成第二頻譜資訊的機構；

藉由以過濾係數所限定的傳遞函數過濾語音合成信號藉以形成語音加工合成信號的機構；及

為了按照上述第二頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號之語音素特徵，而根據上述第二頻譜資訊生成過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

31. 一種語音存儲傳送系統，其為具備有：

藉由分析語音輸入信號，生成以多次元向量表現，並屬

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

六、申請專利範圍

於規定領域且關於上述語音輸入信號之頻譜資訊的機構；

存儲或傳送上述頻譜資訊的機構；

根據被存儲或傳送的上述頻譜資訊生成語音合成信號的機構；

藉由以過濾係數所限定的傳遞函數過濾上述語音合成信號藉以生成語音加工合成信號的機構；及

為了按照上述頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號之語音素特徵，而根據上述頻譜資訊生成上述過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

32.一種語音存儲傳送系統，其為具備有：

藉由分析語音輸入信號，生成以多次元向量表現，並屬於規定領域且關於上述語音輸入信號之第一頻譜資訊的機構；

存儲或傳送上述第一頻譜資訊的機構；

根據被存儲或傳送的上述第一頻譜資訊而生成語音合成信號的機構；

將上述第一頻譜資訊轉換成屬於和上述規定領域相異之領域之第二頻譜資訊的機構；

藉由以過濾係數所限定的傳遞函數過濾上述語音合成信號藉以生成語音加工合成信號的機構；及

為了按照上述第二頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號之語音素特徵，而根據上述第

(請先閱讀背面之注意事項再填寫本頁)

表

訂

六、申請專利範圍

二頻譜資訊生成上述過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

33.一種語音存儲傳送系統，其為具備有：

藉由分析語音輸入信號，生成以多次元向量表現，並屬於規定領域且關於上述語音輸入信號之第一頻譜資訊的機構；

存儲或傳送上述第一頻譜資訊的機構；

根據被存儲或傳送的上述第一頻譜資訊而生成語音合成信號的機構；

藉由分析上述語音合成信號而生成第二頻譜資訊的機構；

藉由以過濾係數所限定的傳遞函數過濾上述語音合成信號藉以生成語音加工合成信號的機構；及

為了按照上述第二頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號之語音素特徵，而根據上述第二頻譜資訊生成上述過濾係數的機構；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

34.一種語音加工方法，其為具有：

第一步驟，係藉由以過濾係數所限定的傳遞函數過濾語音合成信號藉以生成語音加工合成信號者；及

第二步驟，係以多次元向量表現，並為了按照上述頻譜資訊且比上述語音合成信號還能強調上述語音加工合成信號之語音素特徵，而根據屬於規定領域且關於語音合成信

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

六、申請專利範圍

號的頻譜資訊生成上述過濾係數者，其可先實行第一步驟；其中，

上述頻譜資訊，係為LSP資訊、PARCOR資訊及LAR資訊中之任一資訊。

(請先閱讀背面之注意事項再填寫本頁)

訂

公告本

303451

303451

申請日期	85.2.29
案 號	85102394
類 別	G10L 7/02, H04B 1/06, C16

(以上各欄由本局填註)

A4
C4
85年11月6日 修正補充

發明專利說明書

一、發明 名稱	中 文	語音加工或強調用之過濾器暨使用該過濾器之各種裝置、系統及方法
	英 文	
二、發明 創作人	姓 名	田 崎 裕 久
	國 籍	日 本
	住、居所	日本國東京都千代田區丸の内2丁目2番3號 三菱電機株式會社内
三、申請人	姓 名 (名稱)	三菱電機股份有限公司 (三菱電機株式會社)
	國 籍	日 本
	住、居所 (事務所)	日本國東京都千代田區丸の内2丁目2番3號
	代 表 人 姓 名	北 岡 隆

經濟部中央標準局員工消費合作社印製

裝 訂 線

85.11.18 修正
補充

85.11.18
修正本

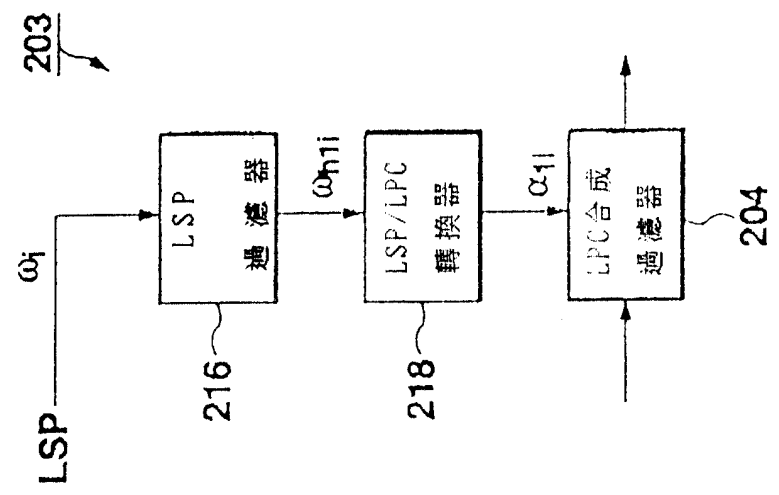


圖 2

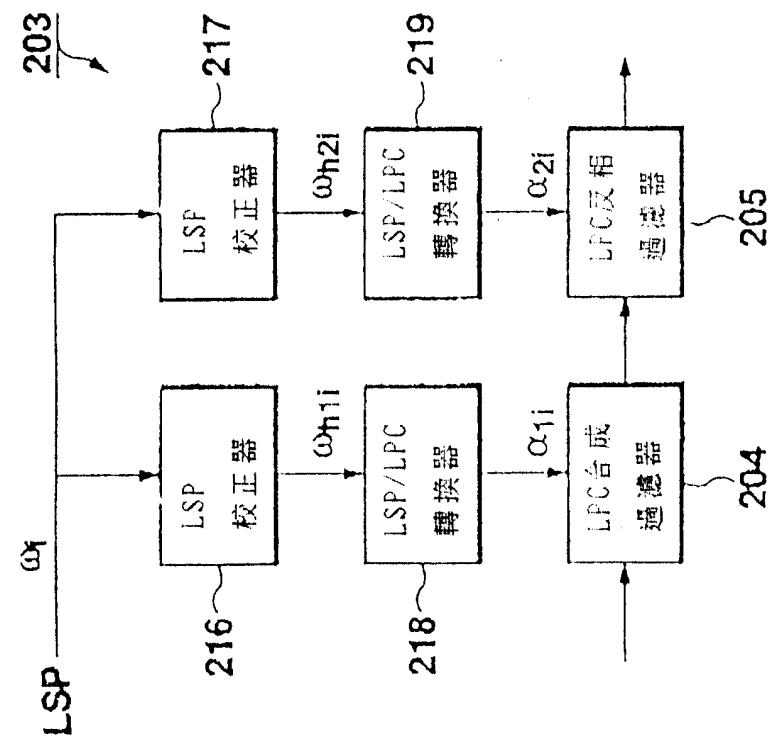
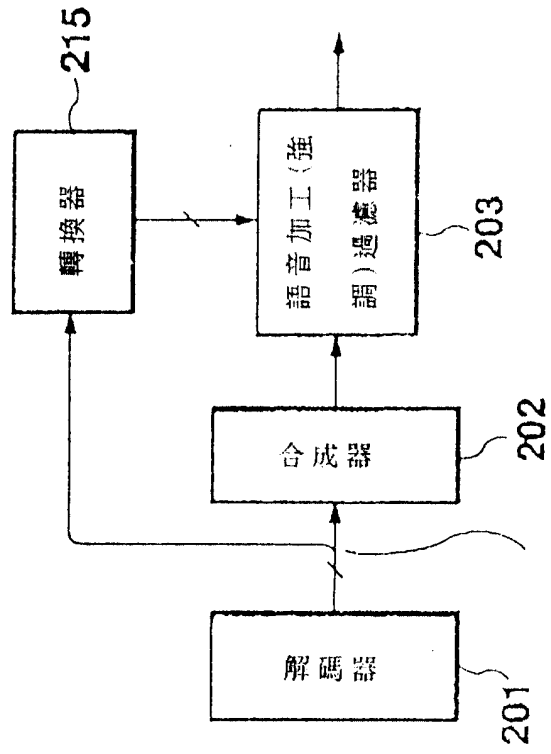


圖 1



200

圖 3

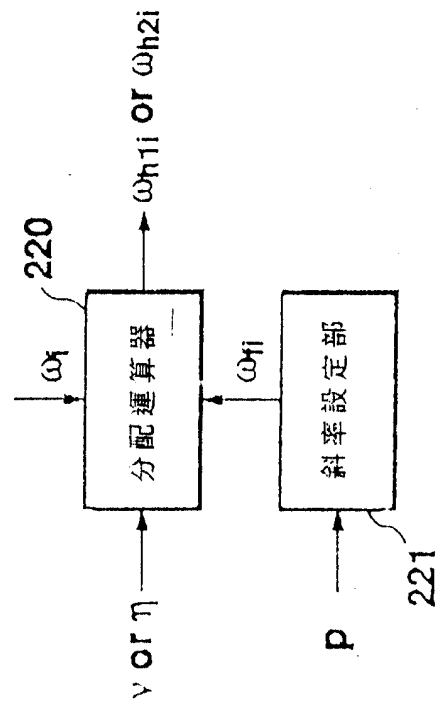


圖 4

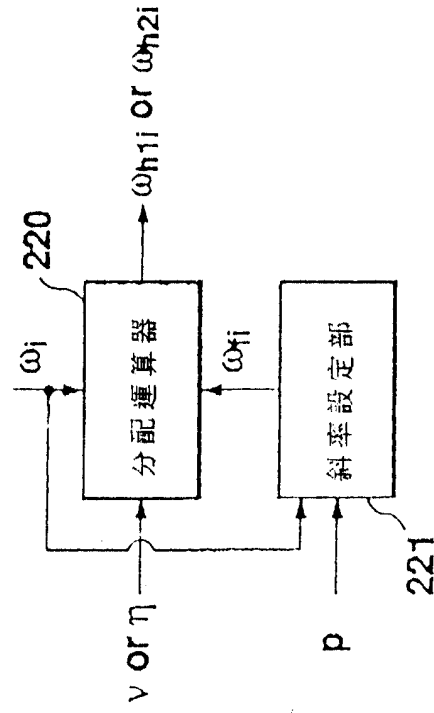


圖 6

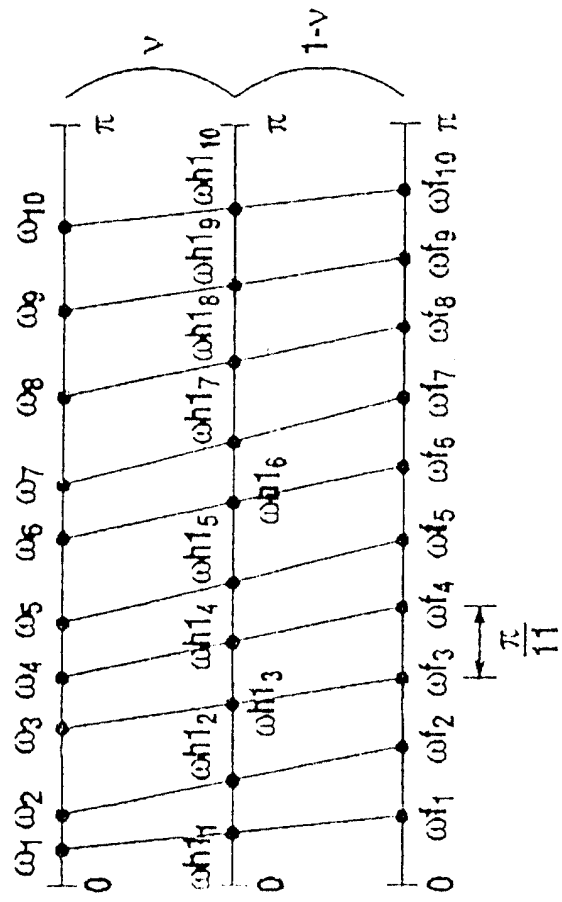


圖 5

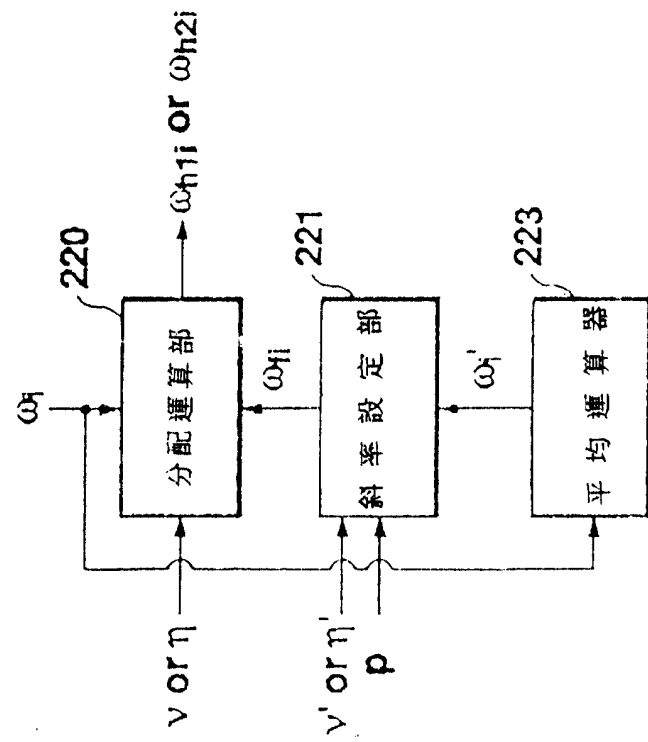


圖 8

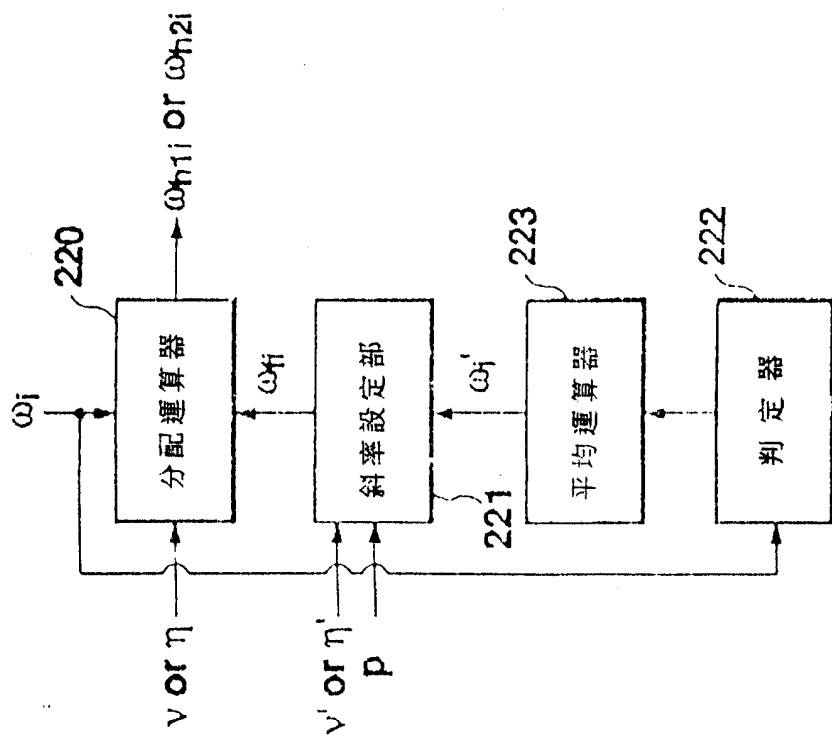
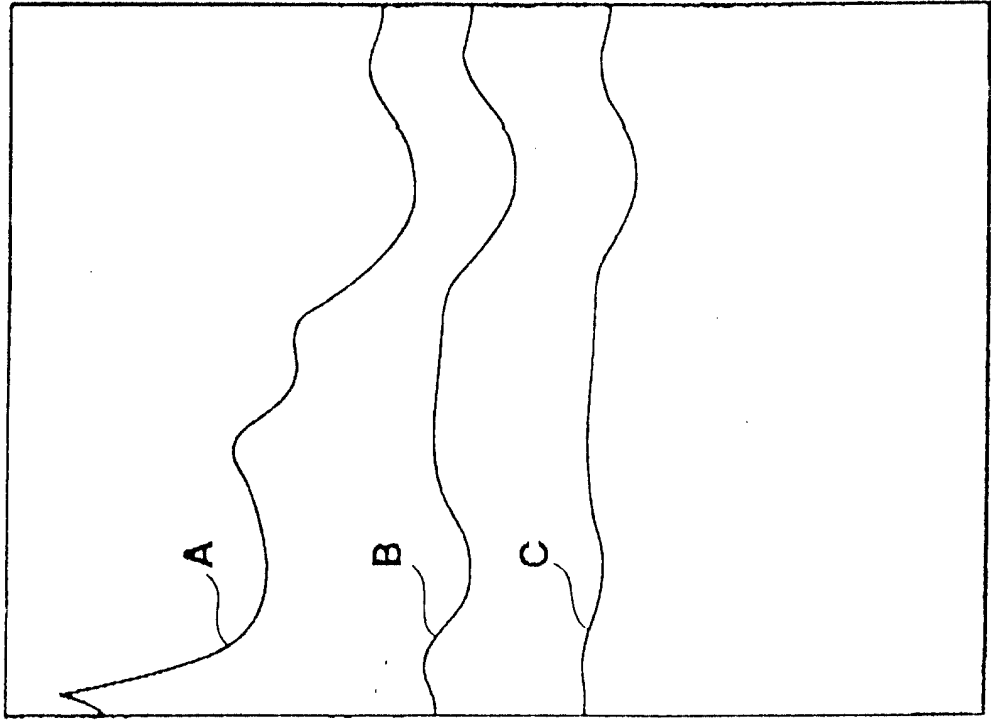
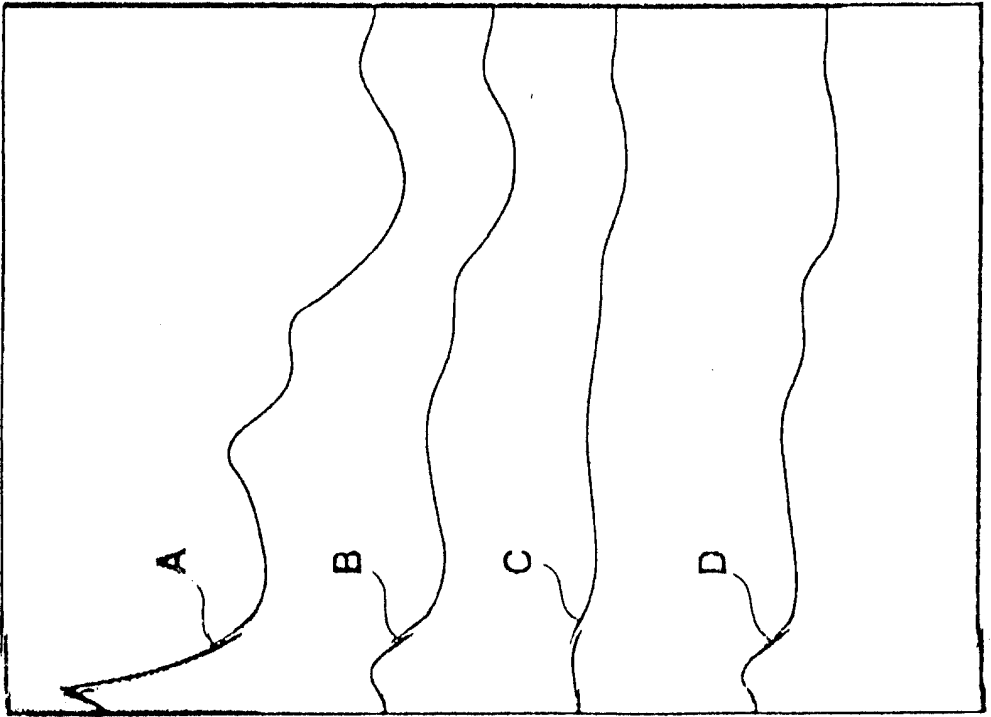


圖 7



對數功率



對數功率

頻率

圖 11

頻率

圖 9

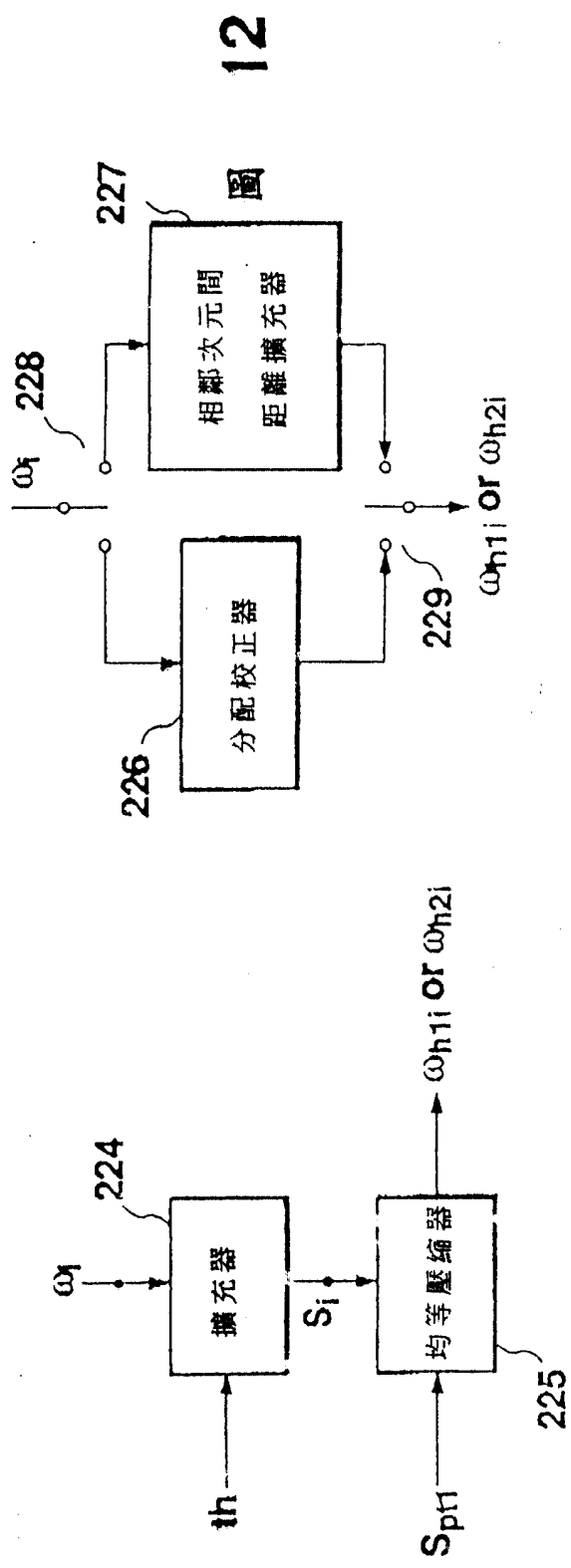


圖 10

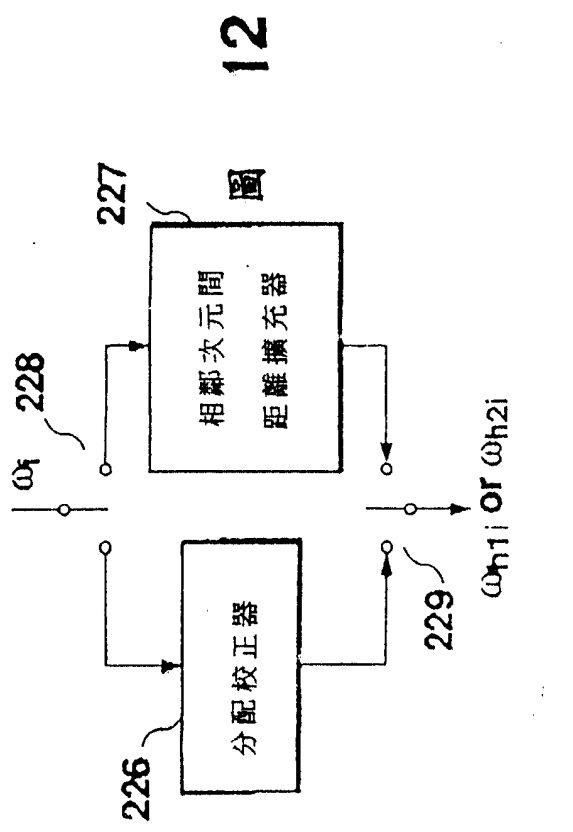


圖 12

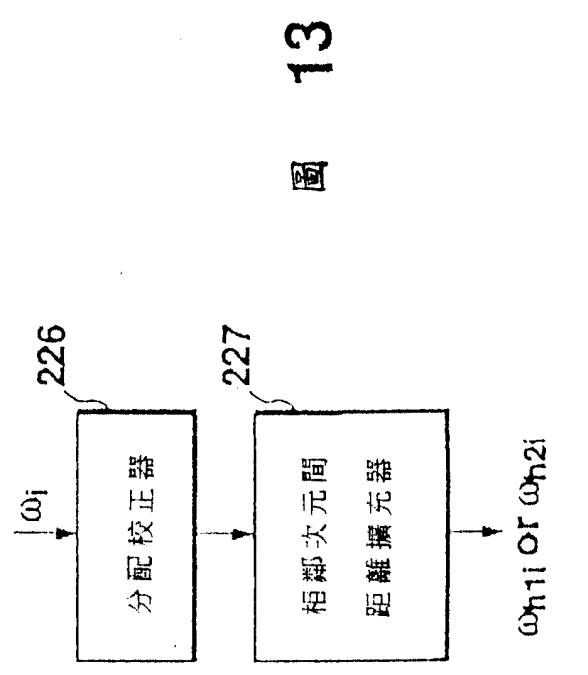


圖 13

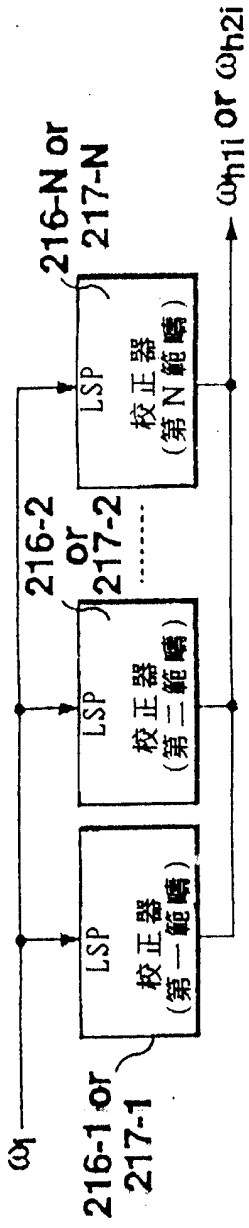


圖 14

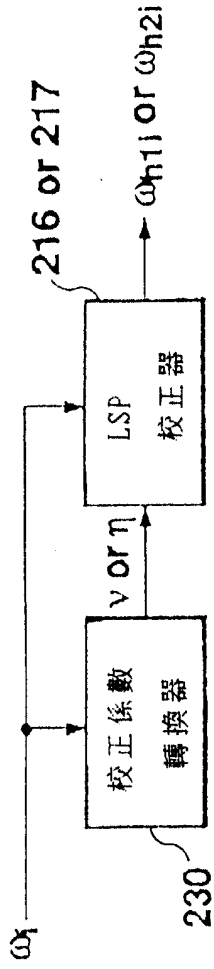


圖 15

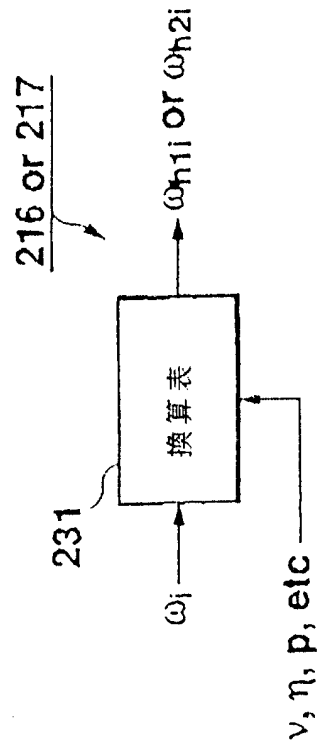


圖 16

216 or 217

232

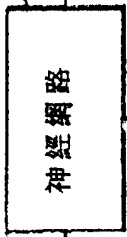


圖 17

$v, \eta, p, \text{ etc}$

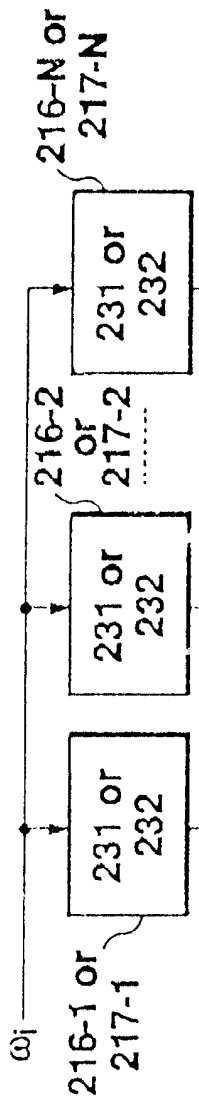


圖 18

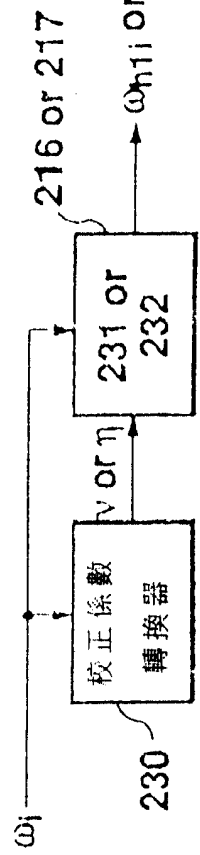


圖 19

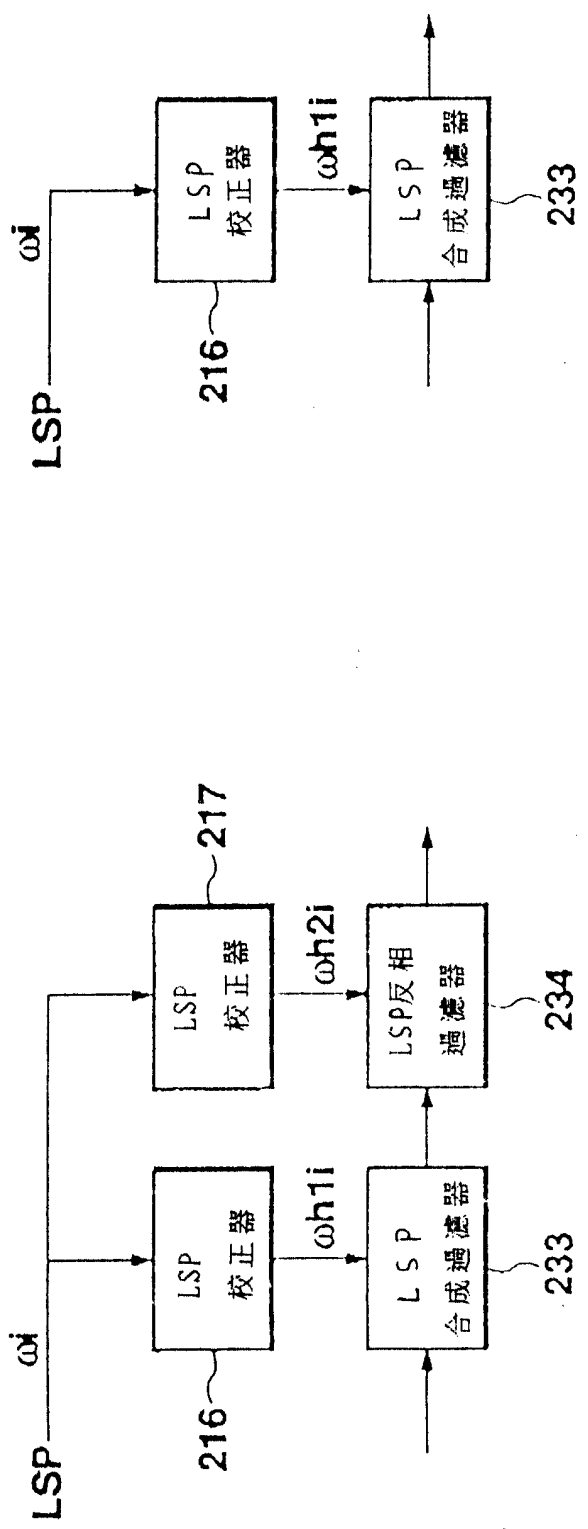


圖 20

圖 21

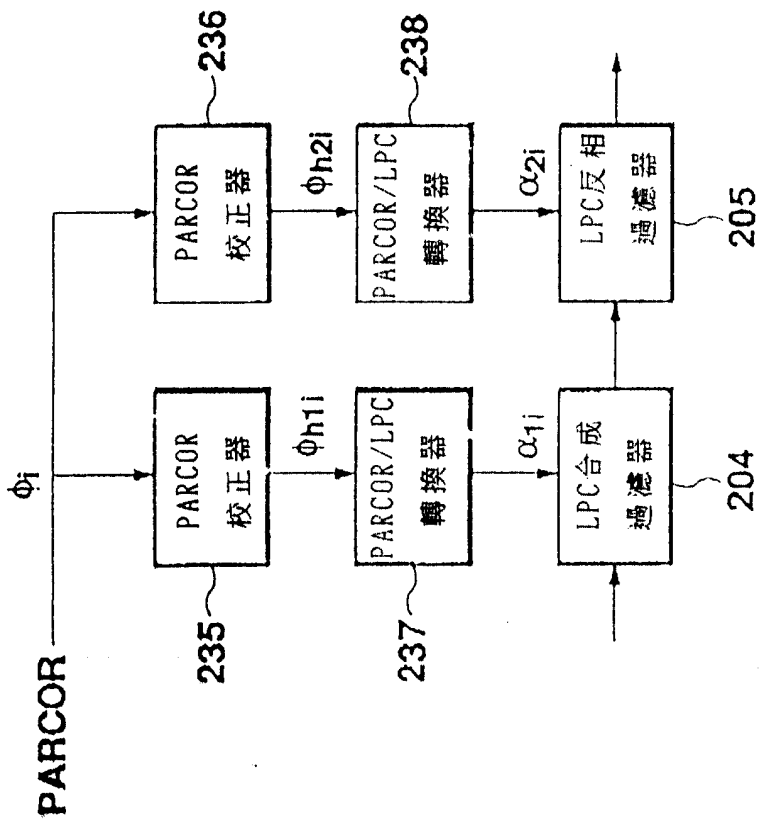


圖 22

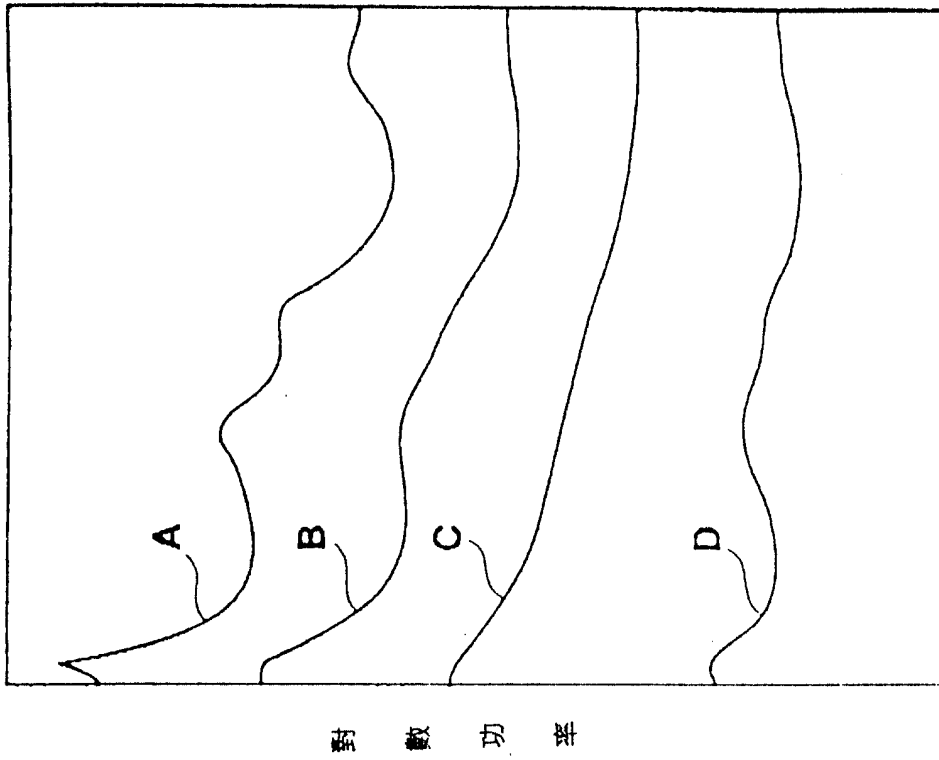


圖 23

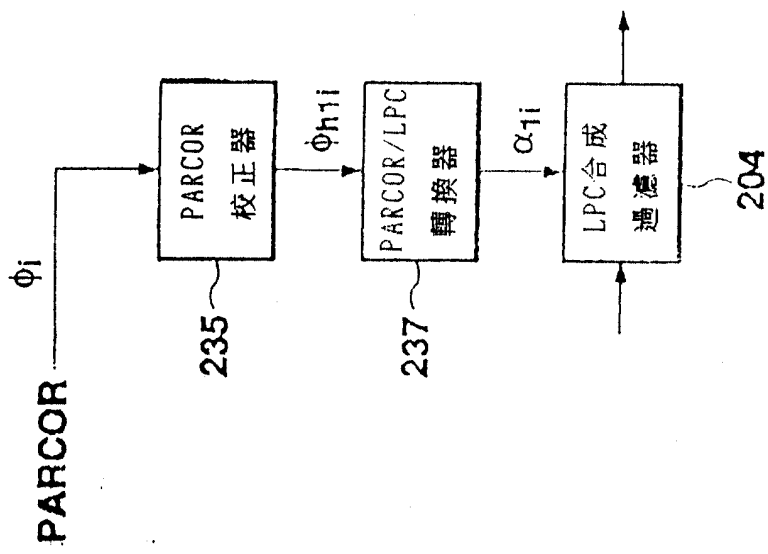


圖 24

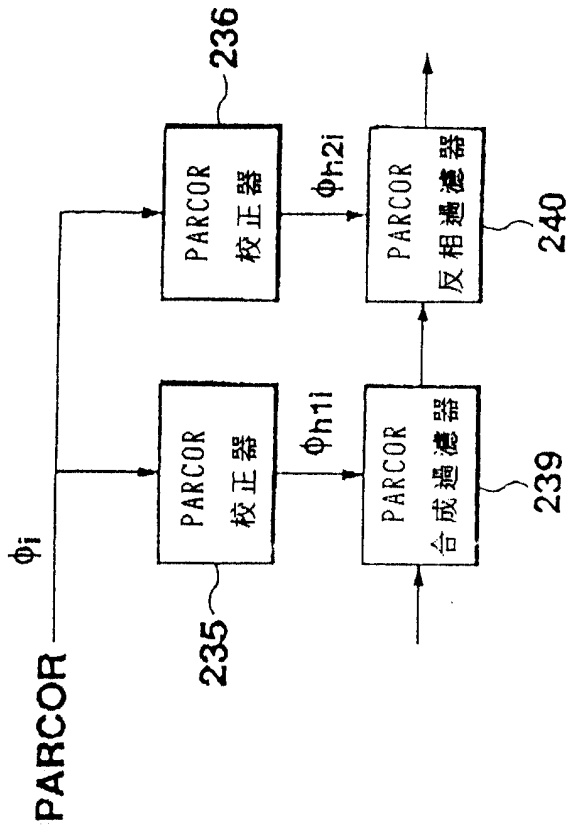


圖 25

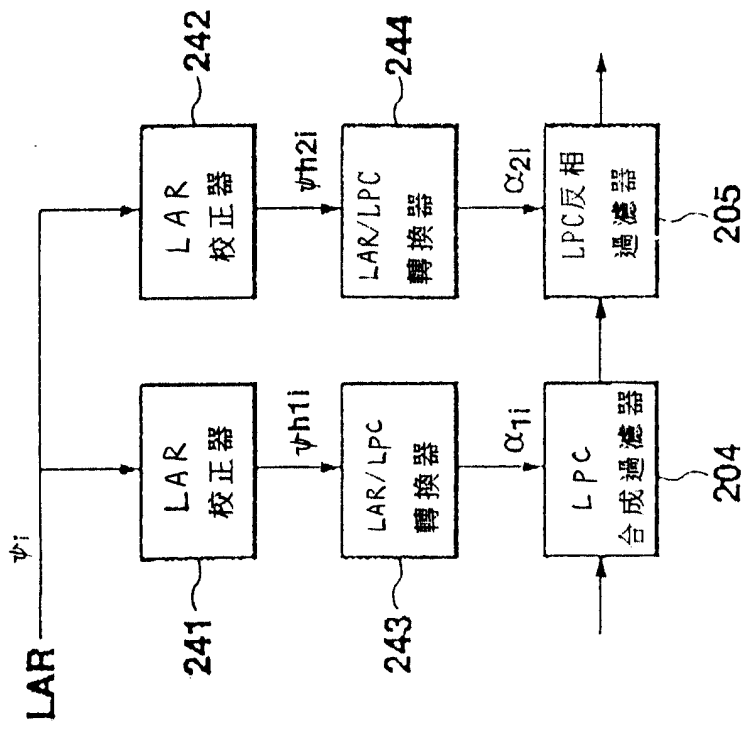
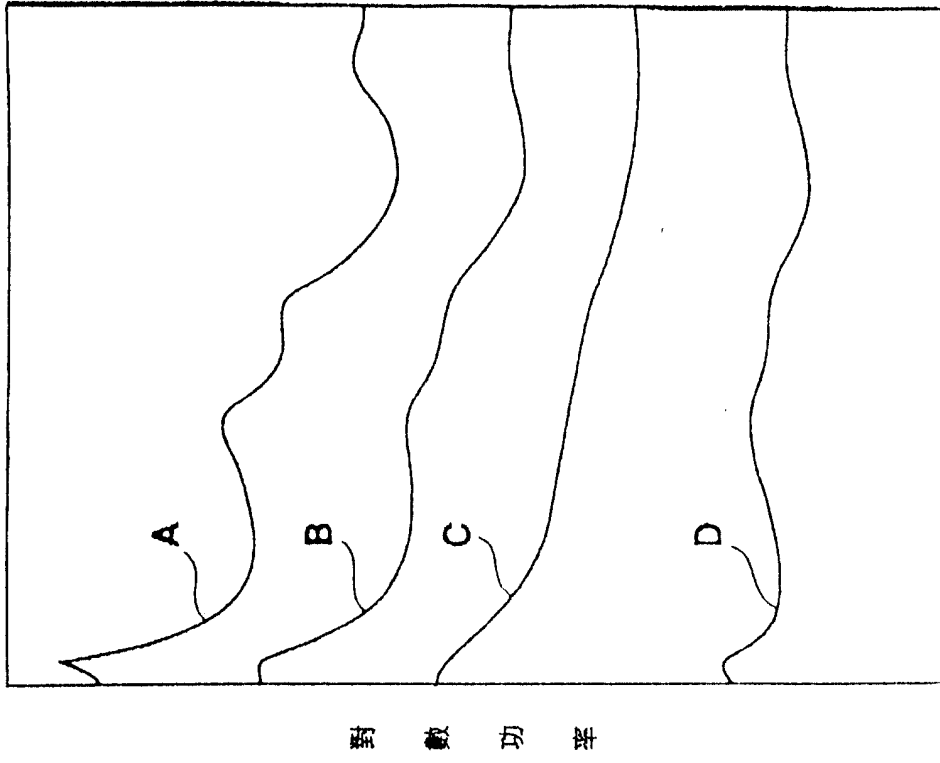


圖 26



頻率

圖 27

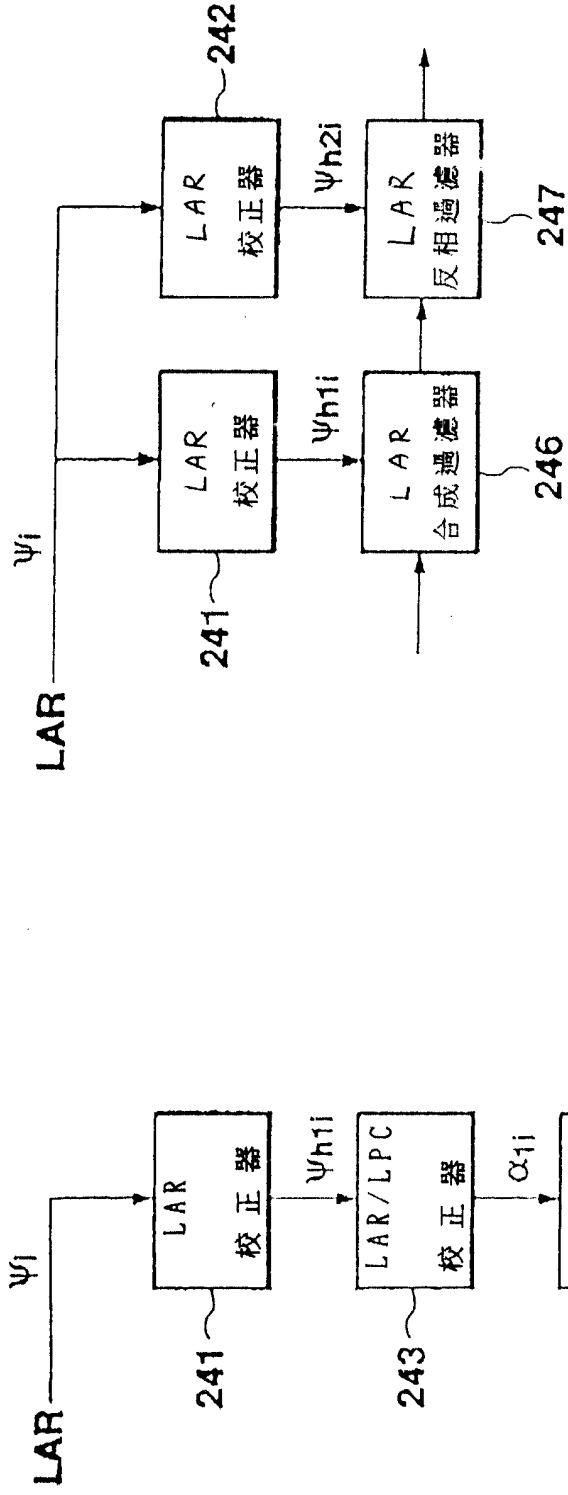


圖 28

圖 29

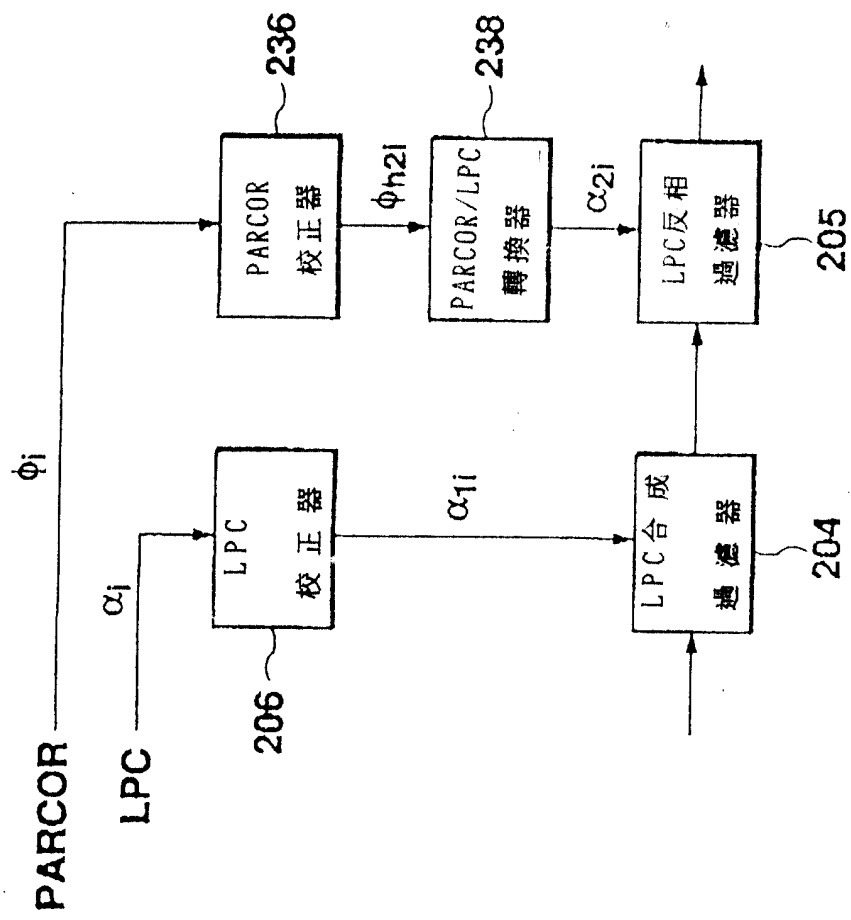
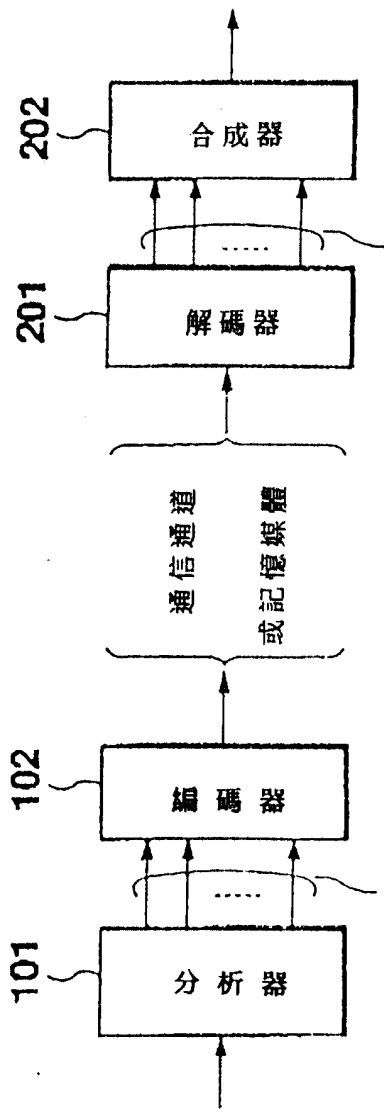


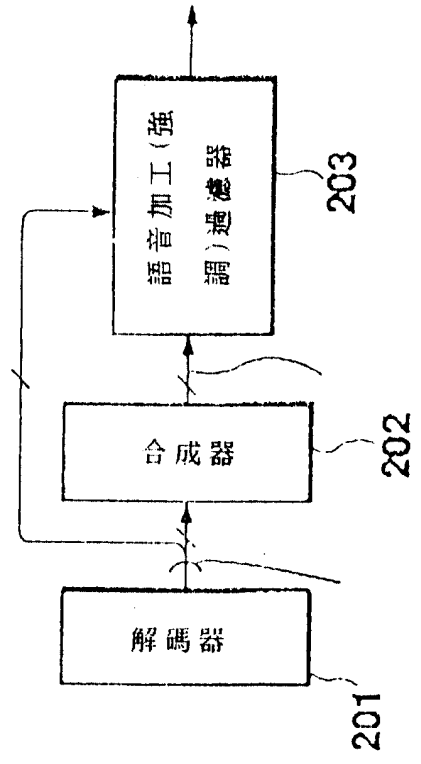
圖 30



100

圖 31

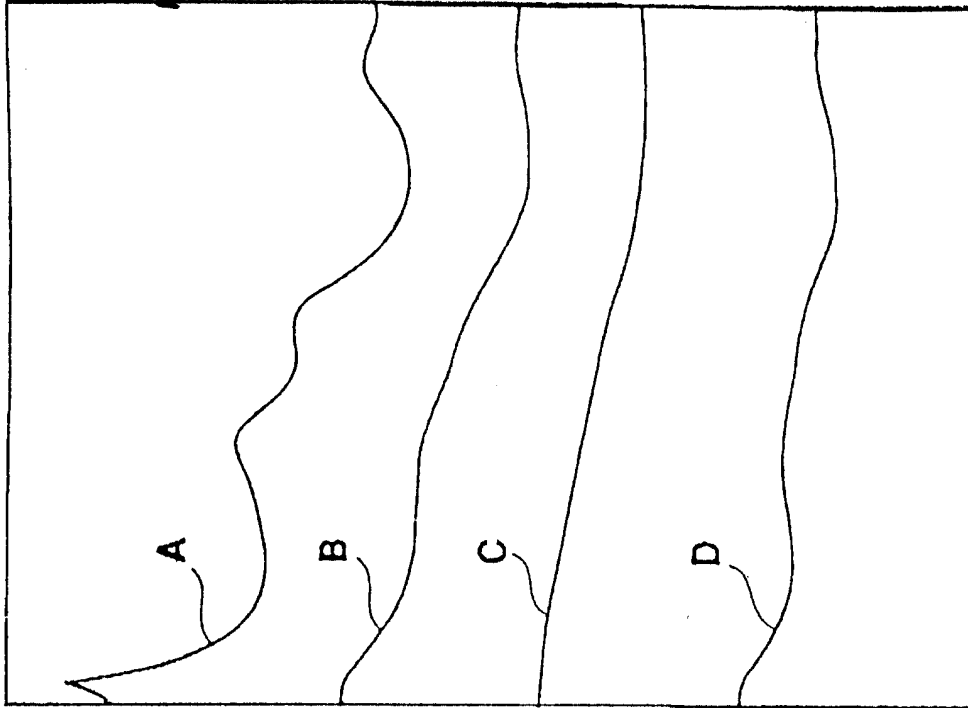
200



200

圖 32

303451



對數功率

頻率

圖 36

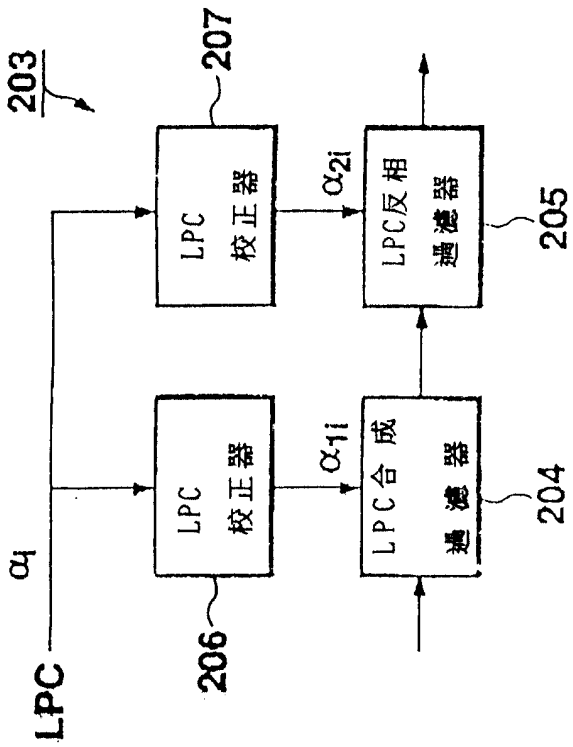


圖 33

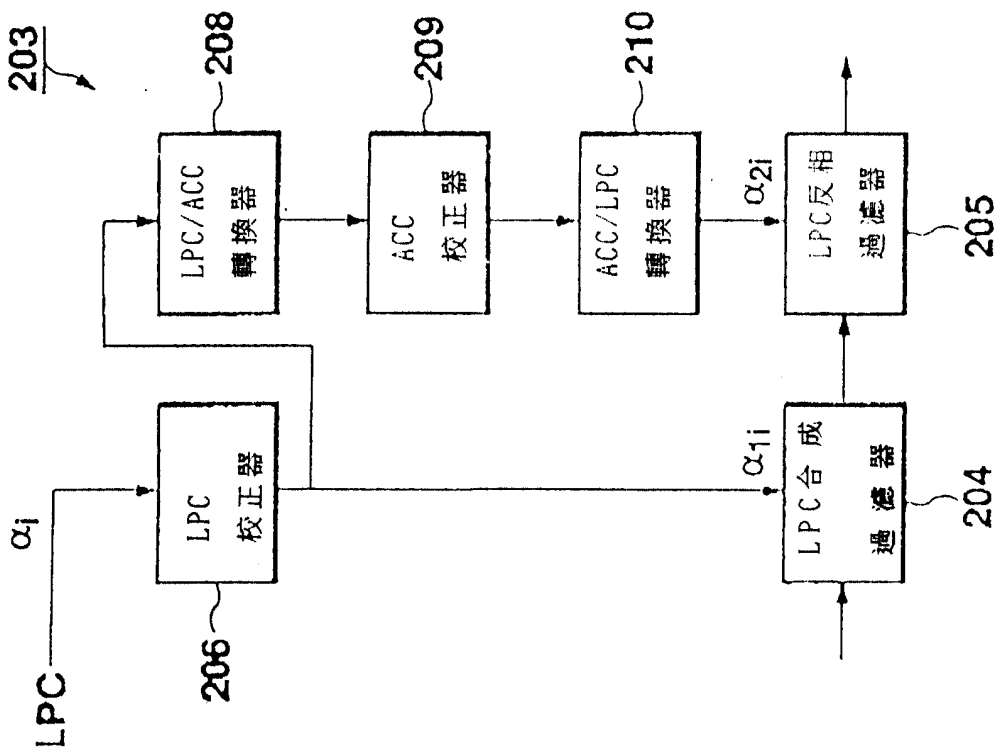
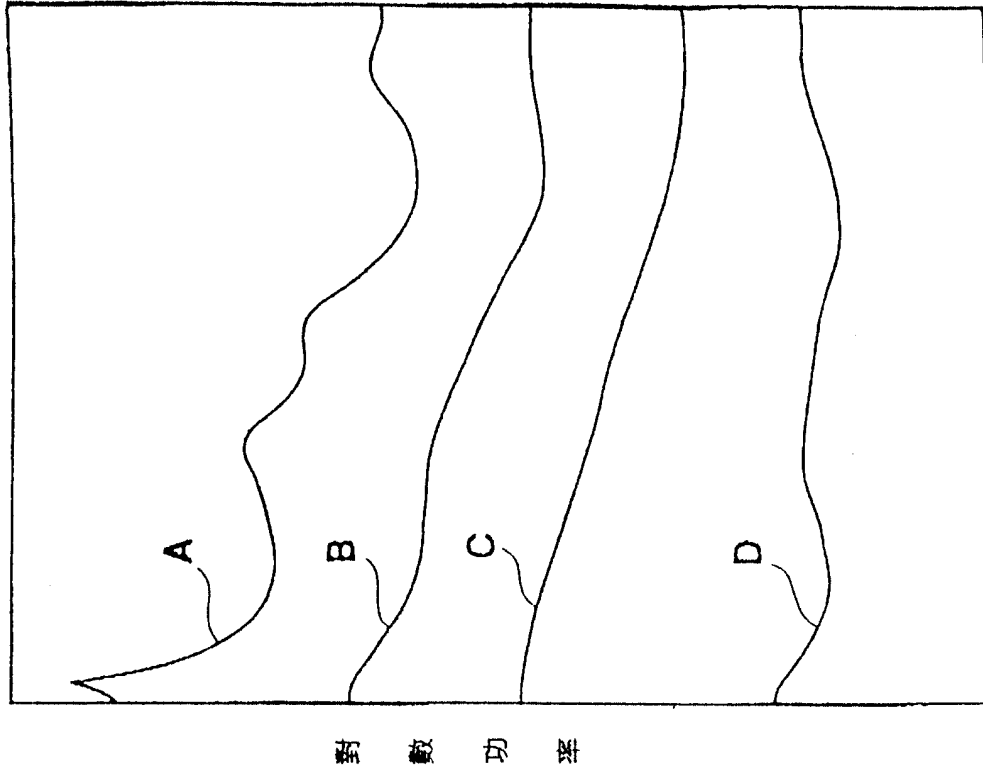
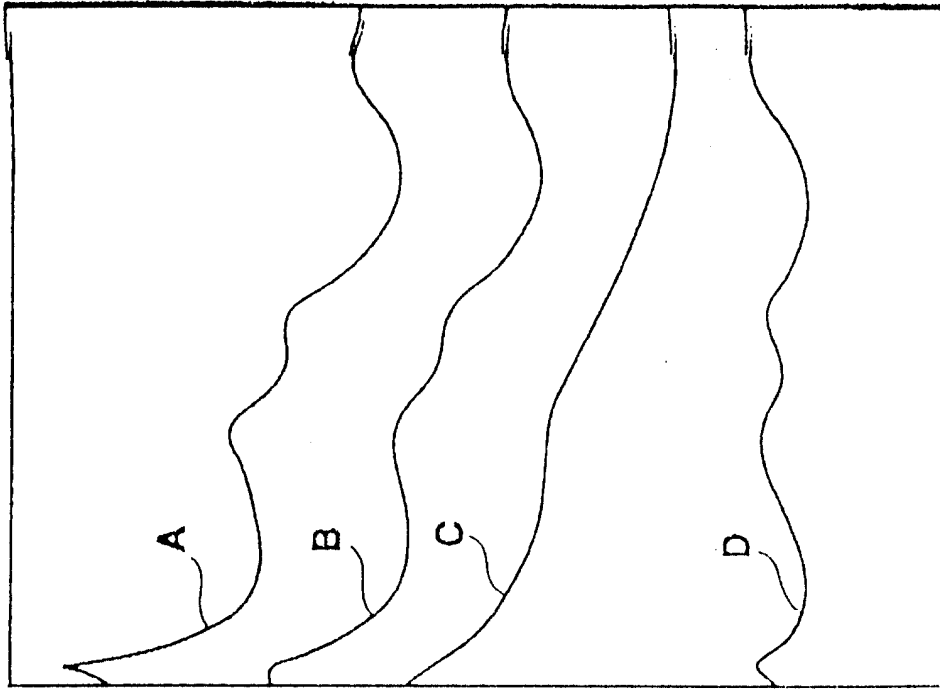


圖 34



頻率

圖 37



頻率

圖 38

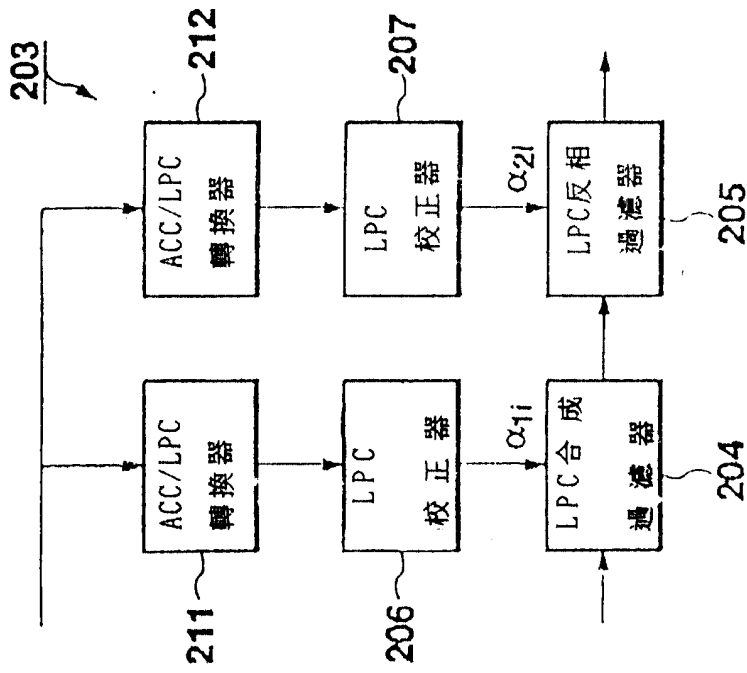


圖 35

303451

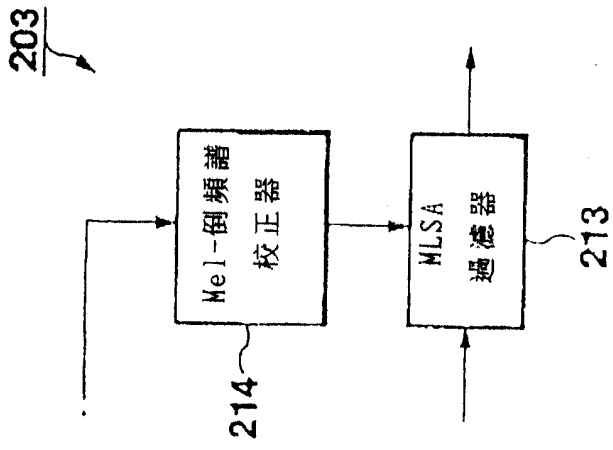


圖 39