



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I854496 B

(45) 公告日：中華民國 113 (2024) 年 09 月 01 日

(21) 申請案號：112106274

(22) 申請日：中華民國 112 (2023) 年 02 月 21 日

(51) Int. Cl. : G10L19/018 (2013.01)

H04N21/233 (2011.01)

(71) 申請人：瑞昱半導體股份有限公司 (中華民國) REALTEK SEMICONDUCTOR CORP. (TW)
新竹科學園區創新二路二號

(72) 發明人：趙盈盈 CHAO, YING-YING (TW)

(74) 代理人：吳豐任；戴俊彥；高銘良

(56) 參考文獻：

TW I337501B

TW 201232323A

TW 201532035A

US 2018/0018748A1

審查人員：施孝欣

申請專利範圍項數：10 項 圖式數：6 共 27 頁

(54) 名稱

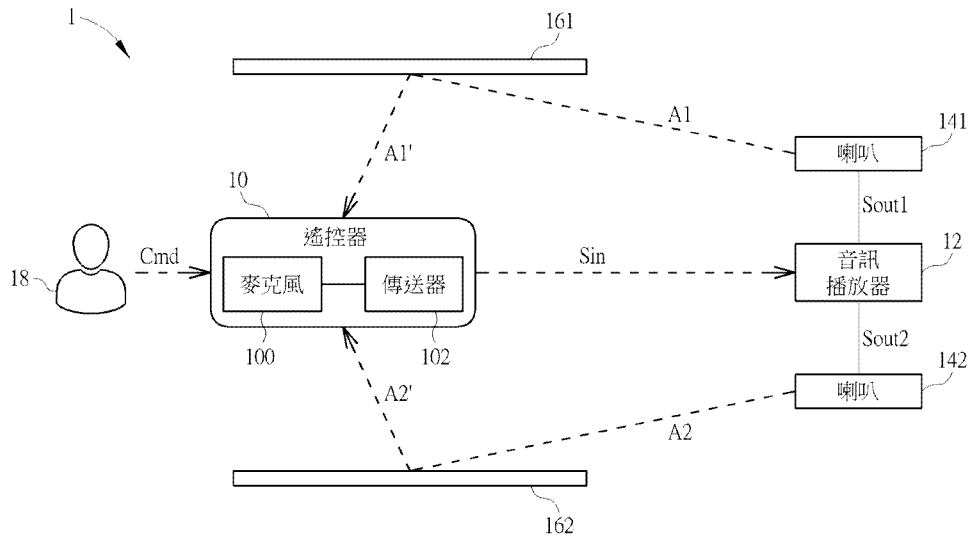
以浮水印為基礎的音訊處理方法及音訊播放器

(57) 摘要

一種以浮水印為基礎的音訊處理方法，適用於包含接收器、類比數位轉換器、處理器、數位類比轉換器及傳送器的音訊播放器。方法包含接收器接收音訊輸入訊號，類比數位轉換器將音訊輸入訊號轉換為數位輸入訊號，處理器偵測數位輸入訊號的能量，若能量超出臨界值，處理器偵測數位輸入訊號中是否包含浮水印，若否，處理器將浮水印嵌入預設數位語音訊號以產生一數位輸出訊號，再經由數位類比轉換器將該數位輸出訊號轉換為音訊輸出訊號，最後傳送器輸出音訊輸出訊號以進行播放，以讓使用者得知音訊播放器已正確辨識出指令。

A watermark-based audio processing method for use in the audio player including a receiver, an analog-to-digital converter, a processor, a digital-to-analog converter and a transmitter includes the receiver receiving an audio input signal, the analog-to-digital converter converting the audio input signal into a digital input signal including a keyword, the processor detecting an energy of the digital input signal, if the energy exceeds a threshold, the processor detecting whether a watermark is included in the digital input signal, if not, the processor embedding the watermark into a predefined digital audio signal to generate a digital output signal, then the digital-to-analog converter converting the digital output signal into an audio output signal, and finally, the transmitter outputting the audio output signal for playback, for a user to know that the audio player has accurately identified a command.

指定代表圖：



第1圖

符號簡單說明：

1:音訊系統

10:遙控器

100:麥克風

102:傳送器

12:音訊播放器

141,142:喇叭

161,162:障礙物

18:使用者

A1,A2,A1',A2':重放
聲波

Cmd:語音指令

Sin:音訊輸入訊號

Sout1,Sout2:音訊輸出
訊號



I854496

【發明摘要】

公告本

【中文發明名稱】以浮水印為基礎的音訊處理方法及音訊播放器

【英文發明名稱】WATERMARK-BASED AUDIO PROCESSING METHOD AND AUDIO PLAYER

【中文】

一種以浮水印為基礎的音訊處理方法，適用於包含接收器、類比數位轉換器、處理器、數位類比轉換器及傳送器的音訊播放器。方法包含接收器接收音訊輸入訊號，類比數位轉換器將音訊輸入訊號轉換為數位輸入訊號，處理器偵測數位輸入訊號的能量，若能量超出臨界值，處理器偵測數位輸入訊號中是否包含浮水印，若否，處理器將浮水印嵌入預設數位語音訊號以產生一數位輸出訊號，再經由數位類比轉換器將該數位輸出訊號轉換為音訊輸出訊號，最後傳送器輸出音訊輸出訊號以進行播放，以讓使用者得知音訊播放器已正確辨識出指令。

【英文】

A watermark-based audio processing method for use in the audio player including a receiver, an analog-to-digital converter, a processor, a digital-to-analog converter and a transmitter includes the receiver receiving an audio input signal, the analog-to-digital converter converting the audio input signal into a digital input signal including a keyword, the processor detecting an energy of the digital input signal, if the energy exceeds a threshold, the processor detecting whether a watermark is included in the digital input signal, if not, the processor embedding the watermark into a predefined digital audio signal to generate a digital output signal, then the digital-to-analog converter converting the digital output signal into an audio output signal, and finally, the transmitter outputting the audio output signal for playback, for a user to know that the audio player has accurately identified a command.

【指定代表圖】第(1)圖。

【代表圖之符號簡單說明】

1: 音訊系統

10: 遙控器

100: 麥克風

102: 傳送器

12: 音訊播放器

141, 142: 喇叭

161, 162: 障礙物

18: 使用者

A1, A2, A1', A2': 重放聲波

Cmd: 語音指令

Sin: 音訊輸入訊號

Sout1, Sout2: 音訊輸出訊號

【特徵化學式】

無

【發明說明書】

【中文發明名稱】以浮水印為基礎的音訊處理方法及音訊播放器

【英文發明名稱】WATERMARK-BASED AUDIO PROCESSING METHOD AND AUDIO PLAYER

【技術領域】

【0001】 本發明關於音訊訊號處理，特別是一種以浮水印(watermark)為基礎的音訊處理方法及音訊播放器。

【先前技術】

【0002】 機上盒(set-top box, STB)為一種網路終端裝置，可在傳統的電視實現線上點播、電視直播、瀏覽網頁、收看影片或遊戲等功能。機上盒能透過語音指令執行各種操作。然而，目前市面上的機上盒在收到語音指令後僅會於電視螢幕顯示語音指令以供使用者確認，使用方式不直覺，造成使用者不便。

【發明內容】

【0003】 本發明實施例提供一種以浮水印為基礎的音訊處理方法，適用於包含接收器、類比數位轉換器、處理器、數位類比轉換器及傳送器的音訊播放器。方法包含：接收器接收音訊輸入訊號；類比數位轉換器將音訊輸入訊號由類比轉換為數位形式以產生數位輸入訊號；處理器偵測數位輸入訊號的能量；若能量超出臨界值，處理器便偵測數位輸入訊號中是否包含浮水印；若數位輸入訊號中未包含浮水印，則處理器將浮水印嵌入預設數位語音訊號以產生數位輸出訊號；數位類比轉換器將數位輸出訊號由數位轉換為類比形式以產生音訊輸出訊號；及傳送器輸出音訊輸出訊號以進行播放。

【0004】 本發明實施例另提供一種以浮水印為基礎的音訊播放器，包含接收

器、類比數位轉換器、處理器、數位類比轉換器及傳送器。接收器，用以接收音訊輸入訊號；類比數位轉換器，耦接於接收器，用以將音訊輸入訊號由類比轉換為數位形式以產生數位輸入訊號；處理器，耦接於類比數位轉換器，用以偵測數位輸入訊號的能量，及若能量超出臨界值，便偵測數位輸入訊號中是否包含浮水印，若數位輸入訊號中未包含浮水印，則理器將浮水印嵌入預設數位語音訊號，以產生數位輸出訊號；數位類比轉換器，耦接於處理器，用以將數位輸出訊號由數位轉換為類比形式以產生音訊輸出訊號；及傳送器，耦接於數位類比轉換器，用以輸出音訊輸出訊號以進行播放。

【圖式簡單說明】

【0005】

第1圖係為本發明實施例中之一種以浮水印為基礎的音訊系統的方塊圖。

第2圖係為第1圖中之音訊播放器的方塊圖。

第3圖係為第1圖中之音訊播放器之音訊處理方法的流程圖。

第4圖係為第1圖中之音訊播放器之音訊輸入訊號的示意圖。

第5圖係為第3圖中之浮水印嵌入步驟的流程圖。

第6圖係為第3圖中之浮水印偵測步驟的流程圖。

【實施方式】

【0006】 第1圖係為本發明實施例中之一種以浮水印為基礎的音訊系統1的方塊圖。音訊系統1可接收使用者18的語音指令並依據語音指令執行相應功能。此外，音訊系統1可將至少一個浮水印嵌入符合裝置中服務語言設定的預設數位語音訊號以產生數位輸出訊號，並播放嵌入浮水印的音訊輸出訊號供使用者18確認，藉以提高系統可用性(accessibility)及使用性(user-friendliness)。若使用者確認

語音指令不正確，使用者可說出關鍵字"取消指令"或其他等效關鍵字，讓音訊系統1取消上一個語音指令的動作。每個浮水印可包含一組浮水印位元。由於播放出的音訊輸出訊號包含至少一個浮水印，因此音訊系統1可判定此語音音訊並非來自使用者18，且不會重複執行語音指令。

【0007】 音訊系統1可包含遙控器10、音訊播放器12、喇叭141及142。遙控器10可透過無線連接耦接於音訊播放器12，音訊播放器12可透過有線或無線連接耦接於喇叭141及142。無線連接可為藍芽連接、無線網路(Wi-Fi)連接、或其他無線連接。有線連接可為通用序列匯流排(Universal Serial Bus, USB)或其他有線連接。音訊播放器12可例如整合於機上盒(set-top box, STB)或智慧音箱控制器。若音訊播放器12整合於機上盒，機上盒也能透過音訊播放器12接收來自使用者18的語音指令以控制電視執行各種操作。在一些實施例中，喇叭141及142可和音訊播放器12分開設置。在另一些實施例中，喇叭141及142可整合入音訊播放器12。在一些實施例中，音訊系統1可設置於室內環境，包含障礙物161及162。障礙物161及162可為牆壁、天花板、桌椅或其他障礙物。當聲波碰撞到障礙物161及/或162時會減弱聲波。

【0008】 遙控器10可包含彼此耦接的麥克風100及傳送器102。麥克風100可接收使用者18發出的語音指令Cmd。語音指令Cmd可包含關鍵字，例如播放音樂、調整音量、播報天氣、設定鬧鐘等。傳送器102可依據語音指令Cmd產生並傳送音訊輸入訊號Sin至音訊播放器12，音訊輸入訊號Sin可包含關鍵字。音訊輸入訊號Sin可為脈衝編碼調變(pulse-code modulation, PCM)訊號，且可為時域訊號。音訊播放器12可執行關鍵字的相應功能以及將P個浮水印的P組浮水印位元嵌入預設數位語音訊號以產生音訊輸出訊號Sout1及Sout2，P為正整數，例如P等於

15。每組浮水印位元可相同，且可依據預設字串的美國資訊交換標準(American standard code for information interchange, ASCII)碼產生。例如，預設字串可為”Realtek”，且每組浮水印位元可依據”Realtek”的ASCII碼產生(例如“0010 0101 0001 1100 0100 0101 1011”)，總共包含28位元。預設數位語音訊號可相應於關鍵字，且可為符合音訊系統1中服務語言設定的內建數位語音訊號。例如，服務語言設定可為中文，預設數位語音訊號可為相應於關鍵字”播放音樂”的內建數位語音訊號。音訊輸出訊號Sout1及Sout2可為相異的立體聲道訊號或相同的單聲道訊號。另外，音訊輸出訊號Sout1及Sout2皆可為脈衝編碼調變訊號，且可為時域訊號。喇叭141及142可分別依據音訊輸出訊號Sout1及Sout2產生重放聲波A1及A2，且重放聲波A1及A2可分別經過空氣傳導及由障礙物161及162反彈以產生衰減後的重放聲波A1’及A2’。麥克風100亦可接收衰減後的重放聲波A1’及A2’，且傳送器102亦可依據衰減後的重放聲波A1’及/或A2’產生並傳送音訊輸入訊號Sin至音訊播放器12。若音訊播放器12偵測到音訊輸入訊號Sin包含至少一組浮水印位元(例如“0010 0101 0001 1100 0100 0101 1011”)則不會再次產生音訊輸出訊號Sout1及Sout2，也不會再次執行關鍵字的相應功能。

【0009】 第2圖係為第1圖中之音訊播放器12的方塊圖。音訊播放器12可包含依序耦接的接收器120、類比數位轉換器(analog-to-digital converter, ADC)122、處理器124、數位類比轉換器(digital-to-analog converter, DAC)126及傳送器128。接收器120可接收音訊輸入訊號Sin，且類比數位轉換器122可依據預設取樣率將音訊輸入訊號Sin由類比轉換為數位形式，經由處理器124分段後產生數位輸入訊號Din。例如，若預設取樣率為每秒48kHz，則類比數位轉換器121可對音訊輸入訊號Sin進行取樣以於每秒產生48k個數位資料以產生數位化的音訊輸入訊號。處理器124可依據預設窗大小(window size)對數位化的音訊輸入訊號進行分段以產

生數位輸入訊號，並偵測數位輸入訊號的能量，偵測數位輸入訊號中的浮水印或將浮水印嵌入預設數位語音訊號，及依據關鍵字執行相應功能。在一些實施例中，預設窗大小可為預設值，例如3秒。在另一些實施例中，處理器124可依據語音指令Cmd的長度設置預設窗大小。在一些實施例中，處理器124可產生嵌入浮水印的數位輸出訊號Dout，且數位類比轉換器126可將數位輸出訊號Dout由數位轉換為類比形式以產生音訊輸出訊號，且傳送器128可依據音訊輸出訊號分別產生音訊輸出訊號Sout1及Sout2，及分別傳送音訊輸出訊號Sout1及Sout2至喇叭141及142以進行播放。使用者18可經由聆聽音訊輸出訊號Sout1及Sout2得知音訊播放器12已正確辨識出語音指令Cmd。

【0010】 第3圖係為音訊播放器12之音訊處理方法300的流程圖。音訊處理方法300包含步驟S300至S314，其中步驟S300至S303用以判斷數位輸入訊號Din是否包含關鍵字；若包含關鍵字，進行步驟S304及S306以判斷數位輸入訊號Din的能量是否超出臨界值Eth；若超出臨界值Eth，進行步驟S307以判斷數位輸入訊號Din是否包含浮水印；若未包含浮水印，表示數位輸入訊號Din為使用者18的輸入語音，繼續步驟S308至S314以播放包含浮水印的音訊輸出訊號Sout1及Sout2。任何合理的步驟改變、順序或調整都落在本公開內容的範圍內。步驟S300至S314解釋如下：

【0011】 步驟S300: 接收器120接收音訊輸入訊號Sin；

【0012】 步驟S302: 類比數位轉換器122將音訊輸入訊號Sin由類比轉換為數位形式以產生數位輸入訊號Din；

【0013】 步驟S303: 處理器124判斷數位輸入訊號Din是否包含關鍵字？若是，繼續步驟S304；若否，結束方法300。

- 【0014】 步驟S304: 處理器124偵測數位輸入訊號Din的能量E；
- 【0015】 步驟S306: 處理器124判斷能量E是否超出臨界值Eth? 若是，繼續步驟S307；若否，結束方法300。
- 【0016】 步驟S307: 處理器124判斷數位輸入訊號Din是否包含浮水印? 若否，繼續步驟S308；若是，結束方法300。
- 【0017】 步驟S308: 處理器124執行語音指令Cmd；
- 【0018】 步驟S310: 處理器124將浮水印嵌入預設數位語音訊號以產生數位輸出訊號Dout；
- 【0019】 步驟S312: 數位類比轉換器將數位輸出訊號Dout由數位轉換為類比形式以產生音訊輸出訊號Sout1/Sout2；
- 【0020】 步驟S314: 傳送器將音訊輸出訊號Sout1/Sout2傳送至喇叭141/142以進行播放；結束方法300。
- 【0021】 在步驟S300，接收器120從遙控器10接收音訊輸入訊號Sin。在步驟S302，類比數位轉換器122以預設取樣率(例如48kHz)對音訊輸入訊號Sin進行取樣以產生數位輸入訊號Din1。在步驟S306，若能量E超出臨界值Eth，可視為數位輸入訊號Din由使用者18以正常音量產生或由較大音量的重放聲波A1'及/或A2'產生；若能量E未超出臨界值Eth，可視為數位輸入訊號由重放聲波A1'及/或A2'產生或由使用者18以較小音量產生。若能量E未超出臨界值Eth，不予處理；若能量E超出臨界值Eth，則由步驟307偵測數位輸入訊號Din是否包含浮水印，來判斷其為使用者18的輸入語音(不包含浮水印)或者為重放聲波(包含浮水印)。在步驟S307及S310，浮水印可例如為"Realtek"的ASCII碼(例如"0010 0101 0001 1100 0100 0101 1011")。以下搭配第1、2、4圖解釋音訊處理方法300。第4圖為數位輸入訊號Din的示意圖，其中橫軸表示時間t，縱軸表示振幅M。

【0022】 針對時間 t_1 及 t_2 之間的第一段數位輸入訊號Din1，處理器124判斷第一段數位輸入訊號Din1包含關鍵字(例如”增加音量”)(步驟S303)，因此偵測第一段數位輸入訊號Din1的能量E(步驟S304)。時間 t_1 及 t_2 之間的時段可為3秒，因此第一段數位輸入訊號Din1包含 $144k(=48k*3)$ 個取樣。第一段音訊輸入訊號Din1的面積越大則能量E越大表示音量越大。

【0023】 接著，處理器124比較第一段數位輸入訊號Din1的能量E及臨界值Eth(步驟S306)。由於第一段數位輸入訊號Din1的能量E超出臨界值Eth，因此處理器124繼續判斷第一段數位輸入訊號Din1是否包含浮水印(步驟S307)。判斷數位輸入訊號Din是否包含浮水印的方法可如第6圖所示，在後續段落會詳細說明。由於處理器124判斷第一段數位輸入訊號Din1未包含浮水印，因此繼續步驟S308。

【0024】 在步驟S308，處理器124依據關鍵字”增加音量”增加播放音量，藉以執行語音指令Cmd。在步驟S310，由於第一段數位輸入訊號Din1未包含浮水印，因此處理器124將浮水印嵌入預設語音數位訊號(例如”增加音量”的內建數位語音訊號)。在一些實施例中，嵌入浮水印的方式可為將P組(例如15組)浮水印位元(例如“0010 0101 0001 1100 0100 0101 1011”)嵌入預設語音數位訊號以產生第一組數位輸出訊號Dout。嵌入浮水印的方式可如第5圖所示，在後續段落會詳細說明。

【0025】 在步驟S312，數位類比轉換器126將第一組數位輸出訊號Dout轉換為音訊輸出訊號Sout及Sout2。在步驟S314，傳送器128將音訊輸出訊號Sout1及Sout2

傳送至喇叭141及142以進行播放，並結束方法300。

【0026】 針對時間 t_3 及 t_4 之間的第二段數位輸入訊號Din2，處理器124判斷第二段數位輸入訊號Din2包含關鍵字(例如”增加音量”)(步驟S303)，因此偵測第二段數位輸入訊號Din2的能量E(步驟S304)。時間 t_3 及 t_4 之間的時段可為3秒，因此第二段數位輸入訊號Din2包含 $144k(=48k*3)$ 個取樣。

【0027】 接著，處理器124比較第二段數位輸入訊號Din2的能量E及臨界值Eth(步驟S306)。由於第二段數位輸入訊號Din2的能量E超出臨界值Eth，因此處理器124繼續判斷第二段數位輸入訊號Din2是否包含浮水印(步驟S307)。由於處理器124判斷第二段數位輸入訊號Din2包含浮水印，因此第二段數位輸入訊號Din2可能由較大音量的重放聲波A1’及/或A2’產生，接著結束方法300。

【0028】 針對時間 t_5 及 t_6 之間的第三段數位輸入訊號Din3，處理器124判斷第三段數位輸入訊號Din3包含關鍵字(例如”增加音量”)(步驟S303)，因此偵測第三段數位輸入訊號Din3的能量E(步驟S304)。時間 t_5 及 t_6 之間的時段可為3秒，因此第三段數位輸入訊號Din3包含 $144k(=48k*3)$ 個取樣。

【0029】 接著，處理器124比較第三段數位輸入訊號Din3的能量E及臨界值Eth(步驟S306)。由於第三段數位輸入訊號Din3的能量E未超出臨界值Eth，因此結束方法300。

【0030】 第5圖係為第3圖中之浮水印嵌入步驟S310的流程圖。步驟S310包含步驟S500至S510，用以將P組浮水印位元嵌入一段符合裝置中服務語言設定的預

設數位語音訊號，以產生一組數位輸出訊號Dout。任何合理的步驟改變、順序或調整都落在本公開內容的範圍內。步驟S500至S510解釋如下：

【0031】 步驟S500: 將預設數位語音訊號切為P段子音訊訊號，及將每段子音訊訊號分為I組時域取樣；

【0032】 步驟S502: 對I組時域取樣進行時域至頻域轉換以產生I組初始轉換係數，每組初始轉換係數包含1個直流係數及N個交流係數；

【0033】 步驟S504: 依據嵌入係數 α 及該組浮水印位元中之第i位元 $w(i)$ 更新第i組初始轉換係數的第m係數 $Coef(i,m)$ 以產生第i組更新轉換係數；

【0034】 步驟S506: 判斷i是否小於I? 若是，繼續步驟S508；若否，繼續步驟S510；

【0035】 步驟S508: 將i更新為i+1；

【0036】 步驟S510: 將更新後的P*I組更新轉換係數進行頻域至時域轉換以產生一組數位輸出訊號Dout；繼續步驟S312。

【0037】 在步驟S500，處理器124先將要嵌入浮水印的預設數位語音訊號(例如”增加音量”的內建數位語音訊號)切為P段(例如P=15)子音訊訊號，每段子音訊訊號嵌入一組浮水印位元(例如“0010 0101 0001 1100 0100 0101 1011”)，接著將每段子音訊訊號分為I組(例如I=28)時域取樣，每組時域取樣嵌入一個浮水印位元。依據前面段落內容，每段數位輸入訊號皆可包含144k個取樣，因此每段子音訊訊號可包含9600(=144000/15)個取樣，且每組時域取樣可包含342(=9600/28)個取樣。

【0038】 在步驟S502，由第一段子音訊訊號開始，處理器124依據離散餘弦轉

換(discrete cosine transform, DCT)或其他種類的離散傅立葉轉換將28組時域取樣進行時域至頻域轉換以產生28組初始轉換係數，每組初始轉換係數嵌入一個浮水印位元。每組初始轉換係數包含1個直流係數 $\text{Coef}(i,0)$ 及 N 個(例如31個)交流係數 $\text{Coef}(i,n)$ ，其中直流係數 $\text{Coef}(i,0)$ 表示第 i 組初始轉換係數的直流係數，交流係數 $\text{Coef}(i,n)$ 表示第 i 組初始轉換係數的第 n 交流係數， i 及 n 為正整數， $i \leq I$ ， $n < N$ 。在一些實施例中，處理器124可將第 i 組初始轉換係數的係數 $\text{Coef}(i,0)$ 至 $\text{Coef}(i,N)$ 儲存於本地記憶體。

【0039】 在步驟S504，處理器124從 $i=1$ 開始，更新第 i 組初始轉換係數的第 m 交流係數 $\text{Coef}(i,m)$ 以產生第 i 組更新轉換係數的第 m 更新交流係數 $\text{Coef}'(i,m)$ ， $i \leq I$ ， $m < N$ ，如公式Eq(1)所示：

$$\text{【0040】} \quad \text{Coef}'(i,m) = \text{Coef}(i,m) * (1 + \alpha * w(i)) \quad \text{Eq(1)}$$

【0041】 例如，嵌入係數 α 可為1，若 $i=1$ ，該組浮水印位元(“0010 0101 0001 1100 0100 0101 1011”)中之第1位元為”0”，因此處理器124計算係數 $(1 + \alpha * w(i))$ 為 $1(=1+1*0)$ ，及產生等於交流係數 $\text{Coef}(i,m)$ 的更新交流係數 $\text{Coef}'(i,m)$ 。處理器124可依據公式Eq(1)使用係數 $(1 + \alpha * w(i))$ 產生一或多個更新交流係數。在一些實施例中， m 為1，交流係數亦可稱為基頻係數，處理器124可依據公式Eq(1)使用係數 $(1 + \alpha * w(i))$ 產生第 i 組更新轉換係數的更新交流係數 $\text{Coef}'(i,1)$ 。在另一些實施例中， m 為1至3，處理器124可依據公式Eq(1)使用係數 $(1 + \alpha * w(i))$ 產生第 i 組更新轉換係數的更新交流係數 $\text{Coef}'(i,1)$ 至 $\text{Coef}'(i,3)$ 。大致而言，若嵌入係數 α 的值增加，則浮水印位元 $w(i)$ 對係數 $\text{Coef}(i,m)$ 的改變可能會隨之增加，增加浮水印位元的可偵測性，但使用者越容易察覺。此外，更新越低頻的交流係數所造成的訊號失真越小，使用者越不易察覺。

【0042】 接著在步驟S506，由於 $i=1$ ，則處理器124判定 i 小於 $I(1<28)$ ，因此在步驟S510， i 被設為 $2(=1+1)$ 。接著處理器124重複步驟S504至S510的迴圈直到 i 到達28為止，藉以更新一段子音訊訊號的 I 組初始轉換係數。此外，處理器124會對 P 段子音訊訊號重複步驟S502至S510以更新 P 段子音訊訊號的 $P*I$ 組初始轉換係數。在步驟S512，處理器124依據反離散餘弦轉換(inverse discrete cosine transform, IDCT)或其他種類的反離散傅立葉轉換將更新後的 $P*I$ 組初始轉換係數由頻域轉換至時域以產生一組數位輸出訊號 D_{out} 。在一些實施例中，處理器124可於本地記憶體保留係數 $Coef(0,0)$ 至 $Coef(I,N)$ 以供後續使用。

【0043】 第6圖係為第3圖中的步驟S307之浮水印偵測的詳細流程圖。步驟S307包含步驟S600至S612，用以偵測一段數位輸入訊號 D_{in} 中是否包含該組浮水印位元。任何合理的步驟改變、順序或調整都落在本公開內容的範圍內。步驟S600至S612解釋如下：

【0044】 步驟S600：將數位輸入訊號 D_{in} 切為 P 段子音訊訊號，及將每段子音訊訊號分為 I 組時域取樣；

【0045】 步驟S602：對 I 組時域取樣進行時域至頻域轉換以產生 I 組後續轉換係數，每組後續轉換係數包含1個直流係數及 N 個交流係數；

【0046】 步驟S604：依據第 i 組初始轉換係數的直流係數 $Coef(i,0)$ 及第 i 組後續轉換係數的直流係數 $Coef'(i,0)$ 估計衰減係數 δ ；

【0047】 步驟S606：依據衰減係數 δ 、嵌入係數 α 、第 i 組初始轉換係數的第 m 交流係數 $Coef(i,m)$ 、及第 i 組後續轉換係數的第 m 交流係數 $Coef'(i,m)$ 產生一組估計浮水印位元的一第 i 估計位元 $w'(i)$ ；

【0048】 步驟S608：判斷 i 是否小於 I ？若是，繼續步驟S610；若否，繼續步驟

S612 ;

【0049】 步驟S612: 將*i*更新為*i+1* ;

【0050】 步驟S614: 依據P組估計浮水印位元判斷數位輸入訊號Din是否包含浮水印。

【0051】 在步驟S600，處理器124先將要進行浮水印偵測的一段數位輸入訊號Din切為P段(例如P=15)子音訊訊號，接著將每段子音訊訊號分為I組(例如I=28)時域取樣。依據前面段落內容，每段子音訊訊號可包含9600(=144000/15)個取樣，且每組時域取樣可包含342(=9600/28)個取樣。在步驟S602，由第一段子音訊訊號開始，處理器124依據離散餘弦轉換(discrete cosine transform, DCT)或其他種類的離散傅立葉轉換將28組時域取樣進行時域至頻域轉換以產生28組後續轉換係數，每組後續轉換係數1個直流係數Coef(i,0)及N個(例如31個)交流係數Coef(i,n)，其中直流係數Coef(i,0)表示第i組後續轉換係數的直流係數，交流係數Coef(i,n)表示第i組後續轉換係數的第n交流係數，i及n為正整數， $i \leq I$ ， $n < N$ 。在一些實施例中，處理器124可將第i組後續轉換係數的係數Coef'(i,0)至Coef'(i,N)儲存於本地記憶體。

【0052】 在步驟S604，處理器124從*i=1*開始計算第i組後續轉換係數的衰減係數 δ ，衰減係數 δ 相對於聲波傳播的衰減程度。處理器124從本地記憶體獲取第i組初始轉換係數的直流係數Coef(i,0)及第i組後續轉換係數的直流係數Coef'(i,0)，並依據以下公式Eq(2)使用直流係數Coef(i,0)及直流係數Coef'(i,0)計算衰減係數 δ 。

【0053】 $\delta = \text{Coef}'(i,0) / \text{Coef}(i,0)$ Eq(2)

【0054】 由於直流成分在聲波傳播時失真最小，因此依據直流係數Coef'(i,0)

第 12 頁，共 15 頁(發明說明書)

所產生的衰減係數 δ 較準確。

【0055】 在步驟S606，處理器124從本地記憶體獲取第 i 組初始轉換係數的第 m 交流係數 $\text{Coef}(i,m)$ 及第 i 組後續轉換係數的第 m 交流係數 $\text{Coef}'(i,m)$ ，及依據公式Eq(3)使用嵌入係數 α 、衰減係數 δ 、交流係數 $\text{Coef}(i,m)$ 及交流係數 $\text{Coef}'(i,m)$ 產生一組估計浮水印位元中之第 i 估計位元 $w'(i)$ 。

$$\text{【0056】 } w'(i) = (\text{Coef}'(i,m) * \alpha) / ((\delta * \text{Coef}(i,m)) - 1) \quad \text{Eq(3)}$$

【0057】 處理器124可依據公式Eq(3)使用一或多個交流係數產生第 i 估計位元 $w'(i)$ 。在一些實施例中，處理器124可依據公式Eq(3)使用1個交流係數中產生第 i 估計位元 $w'(i)$ 。舉例而言， m 為1，處理器124可依據公式Eq(3)使用交流係數 $\text{Coef}(i,1)$ 及交流係數 $\text{Coef}'(i,1)$ 產生第 i 估計位元 $w'(i)$ 。

【0058】 在另一些實施例中，處理器124可依據公式Eq(3)使用多個交流係數中產生第 i 估計位元 $w'(i)$ 。舉例而言， m 為1至3，處理器124可依據公式Eq(3)使用交流係數 $\text{Coef}(i,1)$ 至 $\text{Coef}(i,3)$ 及交流係數 $\text{Coef}'(i,1)$ 至 $\text{Coef}'(i,3)$ 分別產生3個初始第 i 位元，及依據3個初始第 i 位元產生第 i 估計位元 $w'(i)$ ，例如取3個初始第 i 位元的眾數、平均數或其他合適的統計數值作為第 i 估計位元 $w'(i)$ 。

【0059】 接著在步驟S608，由於 $i=1$ ，則處理器124判定 i 小於 $I(1<28)$ ，因此在步驟S610， i 被設為 $2(=1+1)$ ，接著處理器124重複步驟S604至S610的迴圈直到 i 到達28為止，藉以產生1組估計浮水印位元的28個估計位元。此外，處理器124會對剩餘的第2段至第 P 段子音訊訊號重複步驟S602至S610以產生剩餘的14組估計浮水印位元。

【0060】 在步驟S612，處理器124將15組估計浮水印位元和1組真正的浮水印位元相比以判斷該段數位輸入訊號Din是否包含浮水印，該組真正的浮水印位元可儲存於本地記憶體。處理器124可分別比較1組估計浮水印位元的該I個估計位元及該組浮水印位元的該I個位元以產生相似度，及對於15組估計浮水印位元產生15個相似度，相似度可以百分比表示。若至少一個相似度超出預設比例，則處理器124可判定該段數位輸入訊號Din包含浮水印。若15組估計浮水印位元的15個相似度皆未超出預設比例，則判定該段數位輸入訊號Din不包含浮水印。例如，預設比例可為70%，若該組真正的浮水印位元為“0010 0101 0001 1100 0100 0101 1011”，第一組估計浮水印位元為”0**101** 0**110** 0**011** 1100 0000 0001 0011”，粗體數字代表錯誤位元，因此第一組估計浮水印位元包含9個錯誤位元，相似度為68%，未超出預設比例(68%<70%)。若第二組估計浮水印位元為” 0011 0110 0011 1100 0000 0001 0011”，則第二組估計浮水印位元包含7個錯誤位元，相似度為75%，超出預設比例(75%>70%)，因此處理器124判定該段數位輸入訊號Din包含浮水印。

【0061】 實施例提供以浮水印為基礎的音訊處理方法及音訊播放器，將一組浮水印位元嵌入符合裝置中服務語言設定的數位語音音訊，以供使用者確認，藉以提高系統可用性及使用性。

以上所述僅為本發明之較佳實施例，凡依本發明申請專利範圍所做之均等變化與修飾，皆應屬本發明之涵蓋範圍。

【符號說明】

【0062】

1: 音訊系統

10: 遙控器
100: 麥克風
102: 傳送器
12: 音訊播放器
141, 142: 喇叭
161, 162: 障礙物
18: 使用者
120: 接收器
122: 類比數位轉換器
124: 處理器
126: 數位類比轉換器
128: 傳送器
300: 音訊處理方法
S300至S314, S500至S510, S600至S612: 步驟
A1, A2, A1', A2': 重放聲波
Cmd: 語音指令
Din, Din1, Din2, Din3: 數位輸入訊號
Dout: 數位輸出訊號
M: 振福
Sin: 音訊輸入訊號
Sout1, Sout2: 音訊輸出訊號
t, t1, t2, t3, t4, t5, t6: 時間

【發明申請專利範圍】

【請求項1】 一種以浮水印為基礎的音訊處理方法，適用於一音訊播放器，該音訊播放器包含一接收器、一類比數位轉換器、一處理器、一數位類比轉換器及一傳送器，該方法包含：

該接收器接收一音訊輸入訊號；

該類比數位轉換器將該音訊輸入訊號由類比轉換為數位形式以產生一數位輸入訊號；

該處理器偵測該數位輸入訊號的一能量；

若該能量超出一臨界值，該處理器便偵測該數位輸入訊號中是否包含一浮水印；

若該數位輸入訊號中未包含該浮水印，則該處理器將該浮水印嵌入一預設數位語音訊號以產生一數位輸出訊號；

該數位類比轉換器將該數位輸出訊號由數位轉換為類比形式以產生一音訊輸出訊號；及

該傳送器輸出該音訊輸出訊號以進行播放，

其中，該預設數位語音訊號之來源不同於該數位輸入訊號。

【請求項2】 如請求項1所述之方法，其中

該浮水印包含一組浮水印位元，該組浮水印位元包含 I 個位元， I 為一正整數；及

該處理器將該預設數位語音訊號嵌入該浮水印包含：

將該預設數位訊號切為 P 段子音訊訊號， P 為一正整數，及將每段子音訊訊號分為 I 組時域取樣；

對該 I 組時域取樣進行一時域至頻域轉換以產生 I 組初始轉換係數，每組

初始轉換係數包含一直流係數及 N 個交流係數， N 為一正整數；
依據一嵌入係數及該組浮水印位元中之一第 i 位元更新一第 i 組初始轉換係數的一第 m 交流係數以產生一第 i 組更新轉換係數， i 為一正整數且 i 小於或等於 I ， m 為一正整數且 m 小於或等於 N ；及
依據 $P \times I$ 組更新轉換係數進行一頻域至時域轉換以產生該數位輸出訊號。

【請求項3】 如請求項2所述之方法，其中 $m=1$ ，且該第 m 初始交流係數為一基頻係數。

【請求項4】 如請求項1所述之方法，其中該數位輸入訊號包含一關鍵字，且該預設數位語音訊號相應於該關鍵字。

【請求項5】 如請求項4所述之方法，另包含：
若該能量未超出該臨界值，該處理器不執行該關鍵字的一相應功能。

【請求項6】 如請求項1所述之方法，該類比數位轉換器將該音訊輸入訊號由類比轉換為數位形式以產生該數位輸入訊號包含：
該類比數位轉換器將該音訊輸入訊號由類比轉換為數位形式以產生一數位化的音訊輸入訊號；及
對該數位化的音訊輸入訊號進行分段以產生該數位輸入訊號。

【請求項7】 如請求項1所述之方法，其中該處理器偵測該數位輸入訊號中是否包含該浮水印包含：

將該數位輸入訊號切為 P 段子音訊訊號，及將每段子音訊訊號分為 I 組時域取樣；

將該 I 組時域取樣進行一時域至頻域轉換以產生 I 組後續轉換係數，每組後續轉換係數具有一直流係數及 N 個交流係數， N 為一正整數；

依據一第 i 組後續轉換係數的該直流係數及該第 i 組初始轉換係數的該直流係數估計一衰減係數；

依據該衰減係數、該嵌入係數、該第 i 組後續轉換係數的一第 m 交流係數、及該第 i 組初始轉換係數的該第 m 交流係數產生一組估計浮水印位元的一第 i 估計位元；及

依據 P 組估計浮水印位元判斷該數位輸入訊號是否包含該浮水印。

【請求項8】 如請求項7所述之方法，其中依據該 P 組估計浮水印位元判斷該數位輸入訊號是否包含該組浮水印位元包含：

分別比較該組估計浮水印位元的該 I 個估計位元及該組浮水印位元的該 I 個位元以產生一相似度；及

若該相似度超出一預設比例，則判定該數位輸入訊號包含該浮水印。

【請求項9】 如請求項7所述之方法，其中依據該 P 組估計浮水印位元判斷該數位輸入訊號是否包含該組浮水印位元包含：

分別比較該組估計浮水印位元的該 I 個估計位元及該組浮水印位元的該 I 個位元以產生一相似度；及

若該 P 組估計浮水印位元的 P 個相似度皆未超出一預設比例，則判定該數位輸入訊號不包含該浮水印。

【請求項10】 一種以浮水印為基礎的音訊播放器，包含：

一接收器，用以接收一音訊輸入訊號；

一類比數位轉換器，耦接於該接收器，用以將該音訊輸入訊號由類比轉換為數位形式以產生一數位輸入訊號；

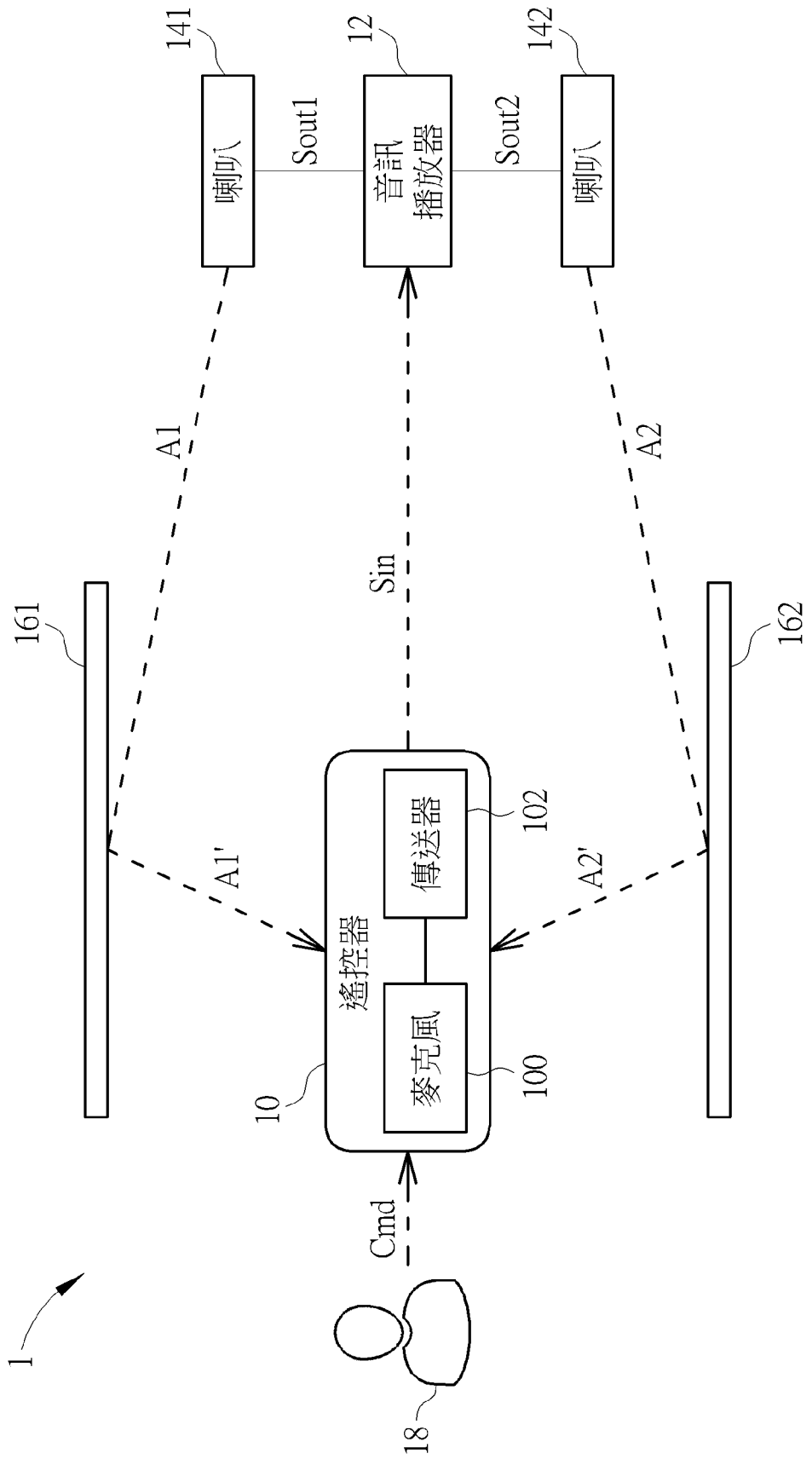
一處理器，耦接於該類比數位轉換器，用以偵測該數位輸入訊號的一能量，及若該能量超出一臨界值，便偵測該數位輸入訊號中是否包含一浮水印，若該數位輸入訊號中未包含該浮水印，則該處理器將該浮水印嵌入一預設數位語音訊號，以產生一數位輸出訊號；

一數位類比轉換器，耦接於該處理器，用以將該數位輸出訊號由數位轉換為類比形式以產生一音訊輸出訊號；及

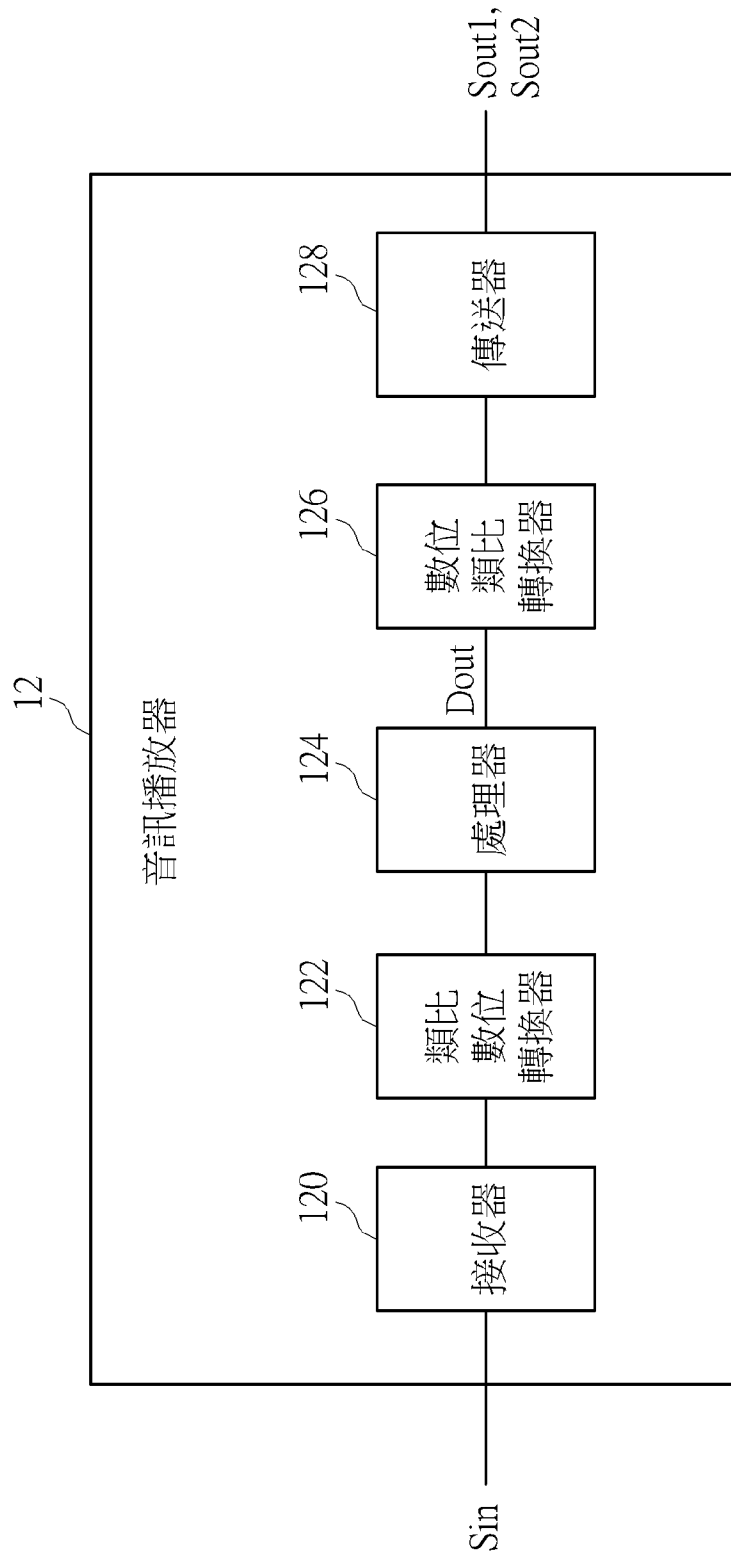
一傳送器，耦接於該數位類比轉換器，用以輸出該音訊輸出訊號以進行播放

其中，該預設數位語音訊號之來源不同於該數位輸入訊號。

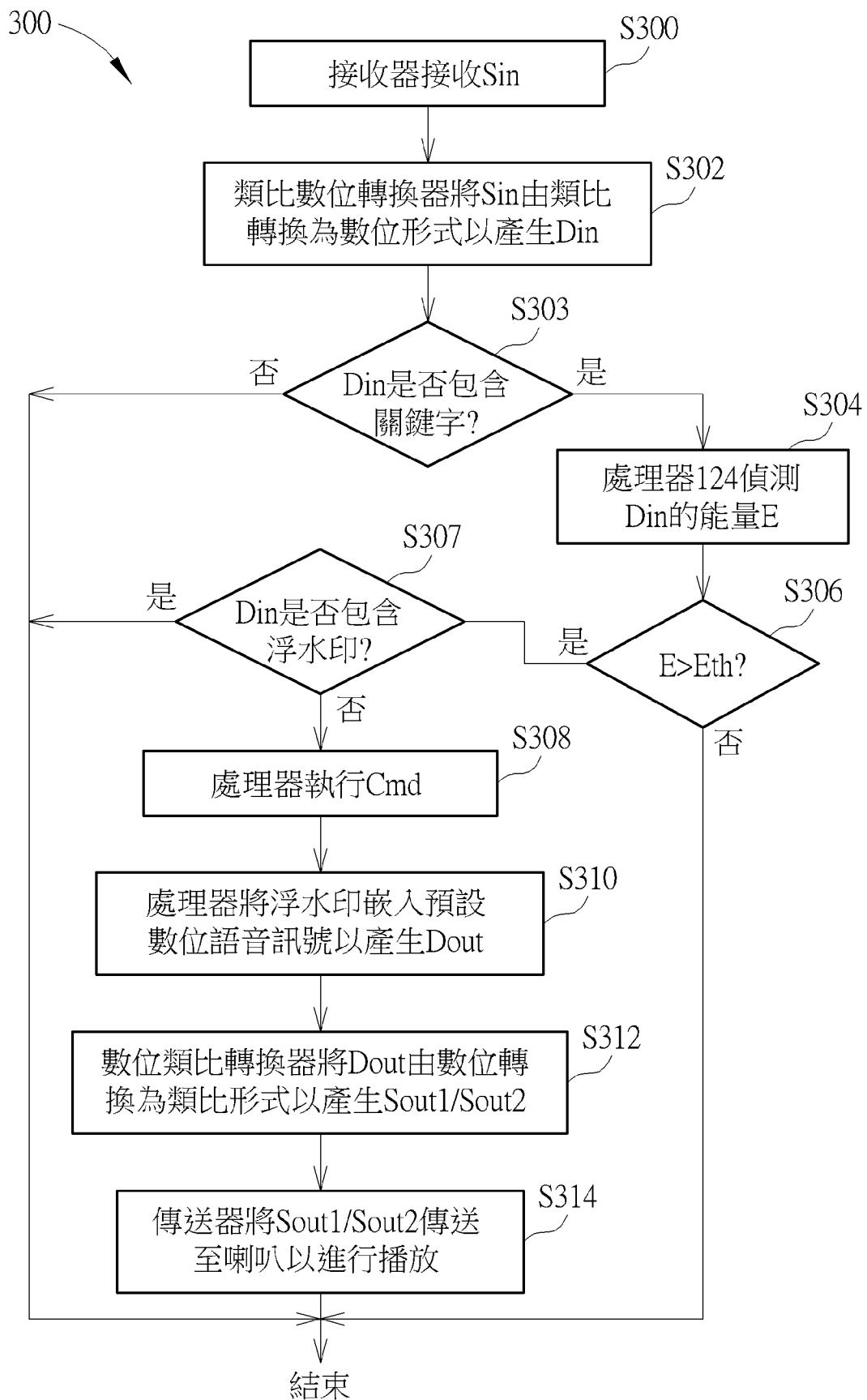
【發明圖式】



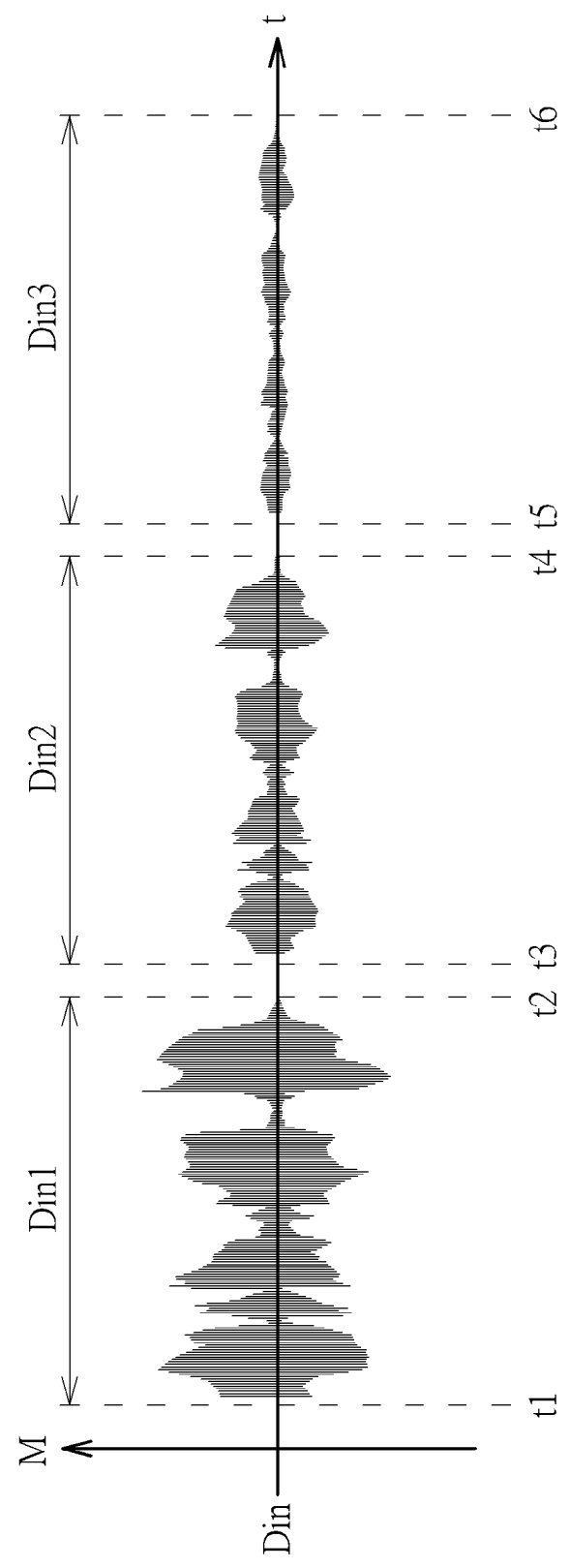
第1圖



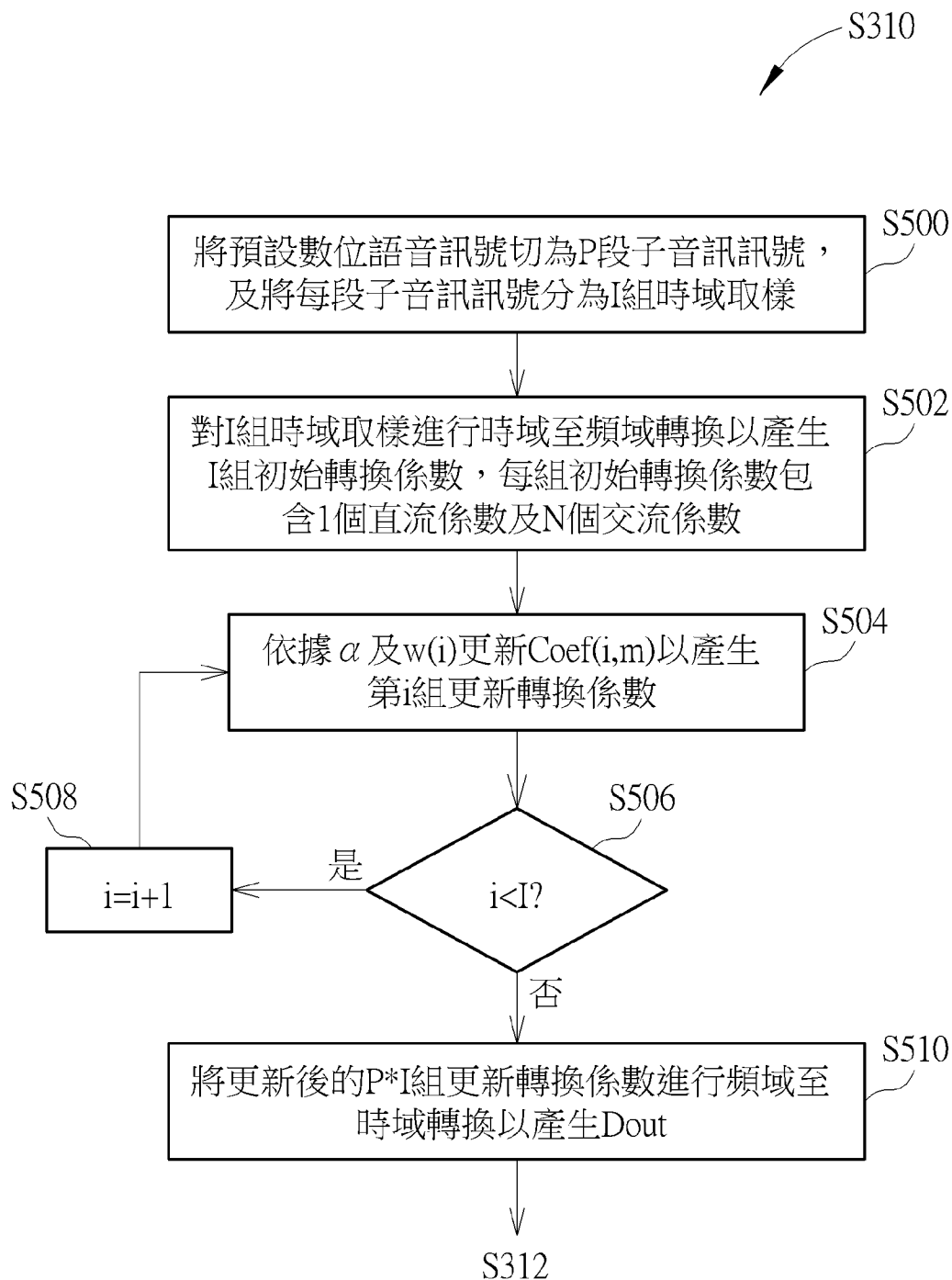
第2圖



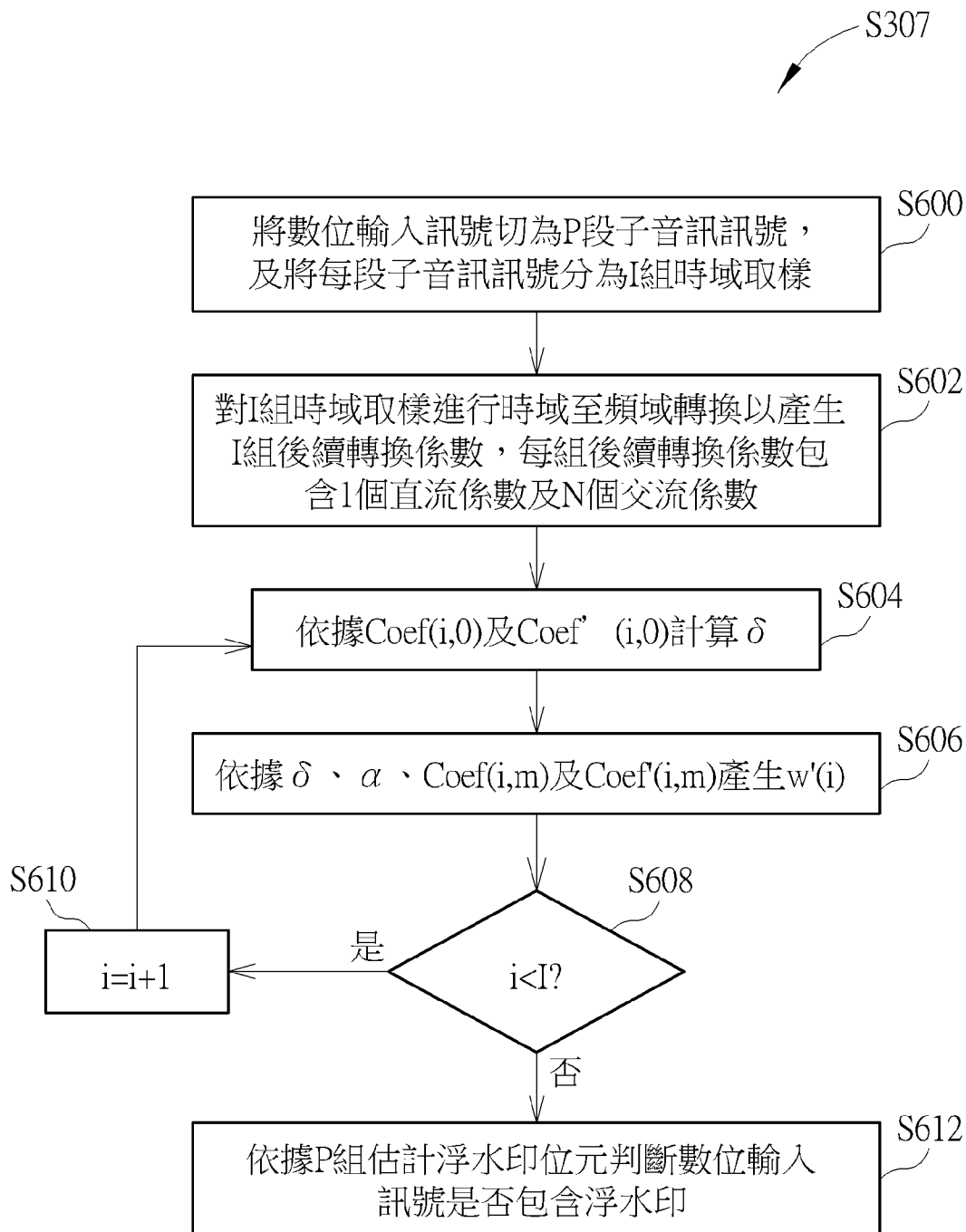
第3圖



第4圖



第5圖



第6圖