



(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 I751463 B

(45)公告日：中華民國 111 (2022) 年 01 月 01 日

(21)申請案號：108146768

(22)申請日：中華民國 108 (2019) 年 12 月 19 日

(51)Int. Cl.：

G10L21/0316(2013.01)**G10L21/038 (2013.01)**

(30)優先權：2018/12/21

歐洲專利局

18215691.9

2019/04/01

歐洲專利局

19166643.7

(71)申請人：弗勞恩霍夫爾協會(德國) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG
DER ANGEWANDTEN FORSCHUNG E.V. (DE)

德國

(72)發明人：迪斯曲 薩斯洽 DISCH, SASCHA (DE)；斯特姆 邁克爾 STURM, MICHAEL (DE)

(74)代理人：劉哲郎

(56)參考文獻：

TW 200830278A

CN 1965352A

US 2008/0027711A1

審查人員：陳守德

申請專利範圍項數：20 項 圖式數：12 共 51 頁

(54)名稱

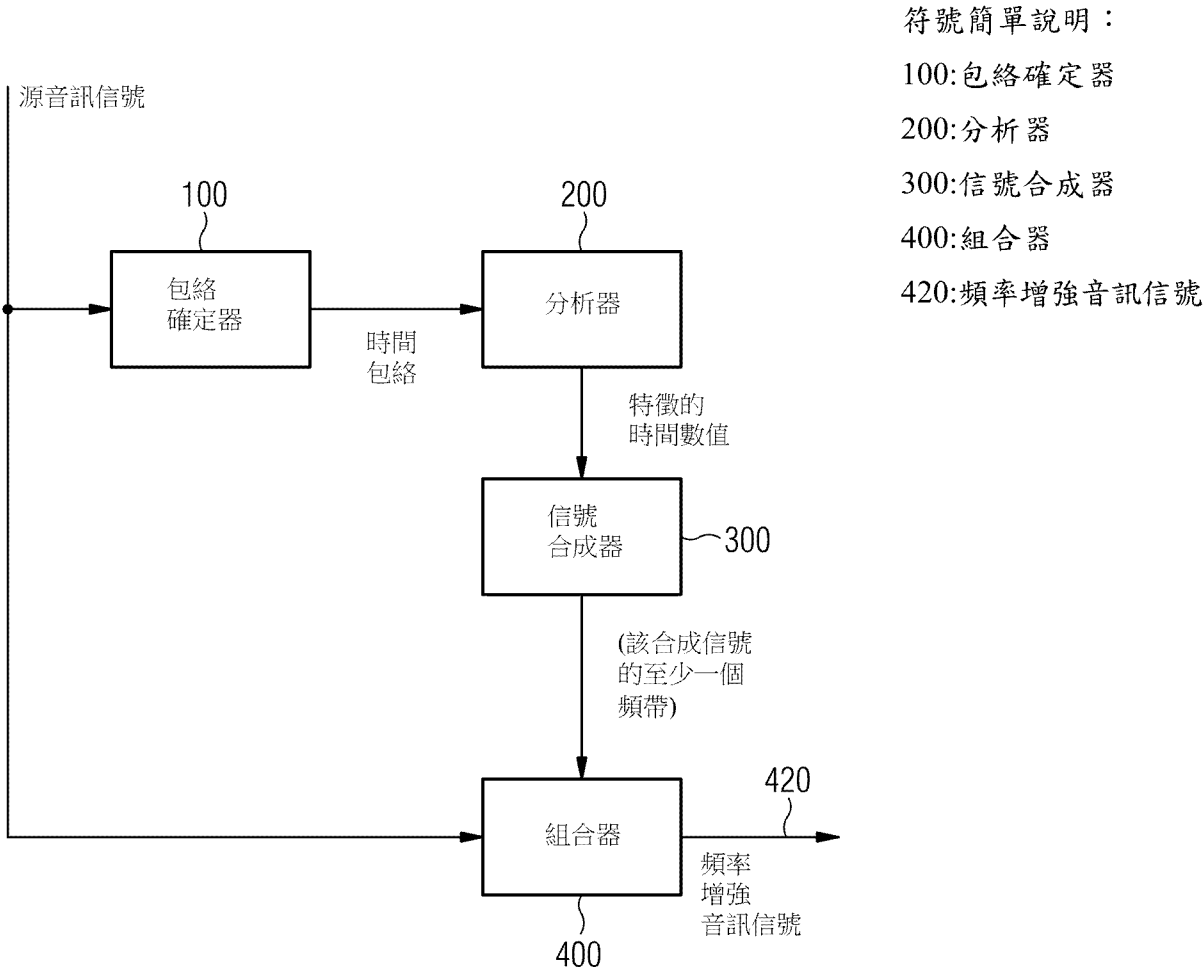
使用脈衝處理用於產生一頻率增強音訊信號的音訊處理器及方法

(57)摘要

一種音訊處理器，用於從一源音訊信號產生一頻率增強音訊信號，其包含：一包絡確定器，用於確定該源音訊信號的至少一部分的一時間包絡；一分析器，用於分析該時間包絡，以確定該時間包絡的多個特定特徵的多個數值；一信號合成器，用於產生一合成信號，該產生包含放置有關於該確定後的時間數值的多個脈衝，其中，在該多個脈衝放置的位置，使用從有關於該多個時間數值的該時間包絡的多個振幅所導出的多個權重值來加權該多個脈衝；以及一組合器，用於組合不包含在該源音訊信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

An audio processor for generating a frequency enhanced audio signal from a source audio signal, comprises: an envelope determiner for determining a temporal envelope of at least a portion of the source audio signal; an analyzer for analyzing the temporal envelope to determine temporal values of certain features of the temporal envelope; a signal synthesizer for generating a synthesis signal, the generating comprising placing pulses in relation to the determined temporal values, wherein the pulses are weighted using weights derived from amplitudes of the temporal envelope related to the temporal values, where the pulses are placed; and a combiner for combining at least a band of the synthesis signal that is not included in the source audio signal and the source audio signal to obtain the frequency enhanced audio signal.

指定代表圖：



【第 1 圖】



公告本

【發明摘要】

I751463

【中文發明名稱】 使用脈衝處理用於產生一頻率增強音訊信號的音訊處理器及方法

【英文發明名稱】 Audio Processor and Method for Generating a Frequency Enhanced Audio Signal Using Pulse Processing

【中文】

一種音訊處理器，用於從一源音訊信號產生一頻率增強音訊信號，其包含：一包絡確定器，用於確定該源音訊信號的至少一部分的一時間包絡；一分析器，用於分析該時間包絡，以確定該時間包絡的多個特定特徵的多個數值；一信號合成器，用於產生一合成信號，該產生包含放置有關於該確定後的時間數值的多個脈衝，其中，在該多個脈衝放置的位置，使用從有關於該多個時間數值的該時間包絡的多個振幅所導出的多個權重值來加權該多個脈衝；以及一組合器，用於組合不包含在該源音訊信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

【英文】

An audio processor for generating a frequency enhanced audio signal from a source audio signal, comprises: an envelope determiner for determining a temporal envelope of at least a portion of the source audio signal; an analyzer for analyzing the temporal envelope to determine temporal values of certain features of the temporal envelope; a

signal synthesizer for generating a synthesis signal, the generating comprising placing pulses in relation to the determined temporal values, wherein the pulses are weighted using weights derived from amplitudes of the temporal envelope related to the temporal values, where the pulses are placed; and a combiner for combining at least a band of the synthesis signal that is not included in the source audio signal and the source audio signal to obtain the frequency enhanced audio signal.

【指定代表圖】 第1圖

【代表圖之符號簡單說明】

- 100 包絡確定器
- 200 分析器
- 300 信號合成器
- 400 組合器
- 420 頻率增強音訊信號

【發明說明書】

【中文發明名稱】 使用脈衝處理用於產生一頻率增強音訊信號的音訊處理器及方法

【英文發明名稱】 Audio Processor and Method for Generating a Frequency Enhanced Audio Signal Using Pulse Processing

【技術領域】

【0001】 本發明涉及音訊信號處理，特別是從一源音訊信號產生一頻率增強音訊信號的概念。

【先前技術】

【0002】 音訊信號的儲存或傳送通常受到嚴格的位元率限制。過去，當只有非常低的位元率可用時，編碼器被迫大幅度地降低傳送的音訊帶寬。如今，現代音訊編解碼器能夠通過使用頻寬延伸(bandwidth extension, BWE)方法對寬頻帶信號進行編碼(參考文獻[1-2])。

【0003】 這些演算法依賴於高頻內容(high-frequency, HF)的一參數化表示，其是通過轉置到該HF頻譜區域(“修補”)從該解碼信號的該波形編碼低頻部分(low-frequency, LF)所產生的。為此，首先產生一“原始”補丁，接著後處理所驅動的參數應用在該“原始”補丁上。

【0004】 通常，所述後處理應用於調整重要的感知屬性，其在高頻產生的過程中尚未通過轉置來考慮，因此需要對該產生的“原始”補丁進行調整。

【0005】然而，如果例如複製到某個目標區域的一補丁中的頻譜細部結構與原始內容的頻譜細部結構有很大差異，可能會產生不期望的偽像，並且降低該解碼音訊信號的感知品質。通常，在這些情況下，所應用的後處理無法完全校正該“原始”補丁的錯誤屬性。

【0006】本發明的目的在於通過間隙填充或估計高頻信號“原始”補丁內容的一新穎的信號自適應產生來改善感知品質，其感知地適用於LF信號。

【0007】通過已經獲得的一感知地適用的“原始”信號，否則必需的後驗校正措施被最小化。此外，一感知地適用的原始信號可以允許在LF和HF之間選擇比傳統方法更低的一交叉頻率(參考文獻[3])。

【0008】在BWE方案中，在一給定的所謂交叉頻率之上的該HF頻譜區域的重建通常基於頻譜修補。通常，該HF區域由多個堆疊的補丁組成，並且這些補丁中的每一個源自於低於該給交叉頻率的該LF頻譜的帶通(band-pass, BP)區域。

【0009】最新技術的系統通過將一組相鄰的子頻帶係數從一源頻譜複製到該目標頻譜區域，有效地在一濾波器組或一時間頻率轉換表示中執行該修補。

【0010】在下一個步驟中，調整音調、噪度及該頻譜包絡，使得其與該原始HF信號的該感知屬性及該包絡非常相似，該原始HF信號在該編碼器中被測量到並且作為BWE輔助資訊在該位元流中傳送。

【0011】頻譜頻帶複製(Spectral Band Replication, SBR)是現代音訊編解碼器，如高效先進音訊編碼(High Efficiency Advanced Audio Coding, HE-AAC)，中所採用的一種眾所周知的BWE，並且使用以上概述的技術(參考文獻[1])。

【0012】 智慧間隙填充(Intelligent Gap Filling, IGF)表示了現代編解碼器，如MPEG-H 3D音訊或3gpp EVS編解碼器(參考文獻[2])，中的一種半參數化編碼技術。由於低位元率限制，IGF可以應用於填充由該編碼器中的量化處理引入的頻譜空洞。

【0013】 通常，如果受限的位元預算不允許透明編碼，則頻譜空洞會首先出現在該信號的該高頻(HF)區域，並且越來越多地影響整個頻譜範圍，以實現最低比特率。

【0014】 在解碼器側，通過IGF使用以一半參數化方式從低頻(LF)內容中產生的合成HF內容及由附加參數化輔助資訊，如頻譜包絡調整及一頻譜“白化水平”，所控制的後處理來替換這樣的頻譜空洞。

【0015】 然而，在所述後處理之後，仍然可能存在剩餘不匹配，其可能導致偽像的感知。這樣的一種不匹配通常可以包括：

- 諧波不匹配：由於錯置的諧波分量而導致的拍打偽像
- 相位不匹配：類似脈衝的激勵信號的分散導致語音或黃銅信號中的嗡嗡聲明顯消失
- 音調不匹配：誇大或太少音調

【0016】 因此，頻率及相位校正方法已經被提出，以通過附加的後處理來校正這些類型的不匹配(參考文獻[3])。在本發明中，我們已經建議避免在“原始”信號中引入這些偽像，而不是像在現有技術方法中那樣在後處理步驟中修復它們。

【0017】 BWE的其他實施方式通常基於用於估計該HF信號的時域技術，通常通過在時域LF波形上應用一非線性函數，如整流、平方或一幕函數。如此

一來，通過使LF失真，產生了子音與不協和泛音的混和物，其可用於作為“原始”信號，以恢復HF內容。

【0018】這裡，尤其是諧波不匹配是一個問題，因為對於和弦內容，這些技術會產生所需諧波泛音的密集混合物，不可避免地會與不期望的非諧波分量混合。

【0019】儘管後處理可能會容易增加雜訊，但一旦將它們引入“原始”估計的HF中，則完全無法消除不想要的非諧波音調分量。

【發明內容】

【0020】本發明的一目的是提供一種改善的概念，用於從一源音訊信號產生一頻率增強音訊信號。

【0021】這個目的可以通過請求項1的一種音訊處理器、請求項19用於處理一音訊的一種方法、或請求項20的一種電腦程式來實現。

【0022】本發明基於發現通過間隙填充或估計高頻(HF)信號“原始”補丁內容的一新穎的信號自適應產生來獲得音訊頻寬延伸或間隙填充或通常是頻率增強的一改善感知品質。通過獲得一感知地適用的“原始”信號，否則必需後驗校正措施可以被最小化甚至消除。

【0023】指示為波形包絡同步脈衝激勵(Waveform Envelope Synchronized Pulse Excitation, WESPE)的本發明實施例是基於在一時域中一類脈衝列信號，其中該實際脈衝放置被同步到一時域包絡。後者是從該低頻(LF)信號所導出，該低頻信號可以在一核心編碼器的一輸出上獲得或從一源音訊信號的任何其他來源獲得。

【0024】 根據本發明的一方面的一音訊處理器，其配置成用於從該源音訊信號產生該頻率增強音訊信號，並且包含一包絡確定器，用於確定該源音訊信號的至少一部分的一時間包絡。一分析器配置成用於分析該時間包絡，以確定該時間包絡的多個特定特徵的多個數值。這些數值可以是時間數值、或能量、或有關於該多個特徵的其它數值。一信號合成器被設置用於產生一合成信號，其中該合成信號的該產生包含放置有關於該確定後的時間數值的多個脈衝，其中，在該多個脈衝放置的位置，使用從該時間包絡的多個振幅所導出的多個權重值來加權該多個脈衝，其中該多個振幅有關於該多個時間數值。具有一組合器，用於組合不包含在該源音訊信號的該合成信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

【0025】 本發明的優點在於，與從該源音訊信號有點“盲目”地產生較高頻率相比，例如通過使用一非線性處理等，本發明提供了一種易於控制的程序，其通過確定該源信號的該時間包絡及通過在該時間包絡的多個特定特徵，例如該時間包絡的局部最大值或該時間包絡的局部最小值，上放置多個脈衝，或通過在該時間包絡的兩個局部最小值之間放置多個脈衝，或對應於該時間包絡的多個特定特徵的任何其它關係。一脈衝具有一頻率內容，其通常在所考慮的整個頻率範圍內是平坦的。因此，即使使用理論上不理想但接近理想的脈衝，在感興趣的頻率範圍內，這種非理想脈衝，即與理想狄拉克形狀不一致的脈衝，的頻率內容仍然相對平坦，例如，在IGF(智慧間隙填充)的情況下，在0和20 kHz之間、或在音頻延伸的情況下，在例如8 kHz至16 kHz或20 kHz之間的頻率範圍內，其中該源信號是受頻寬限制的。

【0026】 因此，由這種脈衝組成的該合成信號提供了一密集且易於控制的高頻內容。通過在每一個時間包絡放置幾個脈衝，該時間包絡例如從該源音訊信號的一幀中提取，來獲得該頻譜域的整形，因為相對於該特定特徵所放置的不同脈衝的該頻率內容在該頻域中彼此重疊，以便於至少匹配該主要特徵或通常是該音訊信號的該時間包絡的該多個特定特徵。由於脈衝所代表的該頻譜值的相位相互鎖定，並且較佳地，通過該信號合成器放置多個正脈衝或多個負脈衝，在不同的脈衝中由一個別脈衝所代表的該頻譜值的相位彼此鎖定。因此，獲得了具有非常有用的頻域特性的一控制的合成信號。通常，該合成信號是一寬帶信號，其在整個現有音訊頻率範圍內延伸，即，也延伸到該LF範圍內。為了實際產生該最終信號，該最終信號最終地組合了用於頻率增強的該源音訊信號、該合成信號的至少一個頻帶，例如該高頻頻帶、或通過一帶通所確定的一信號被提取並且加入到該源音訊信號。

【0027】 本發明概念具有在時域中完全執行的潛力，即，無需任何特定轉換。該時域可以是典型時域，也可以是線性預測編碼(linear prediction coding, LPC)濾波後的時域，即，一時域信號，其已經頻譜地白化並且最終需要使用一LPC合成濾波器進行處理，以重新引入該原始頻譜形狀，以便於對音訊信號渲染有用。因此，包絡確定、分析、信號合成、合成信號頻帶的提取以及最終組合都可以在該時域中進行，這樣就可以避免任何通常引起延遲的時間頻譜轉換或頻譜時間轉換。然而，本發明的上下文在幾個程序上也是靈活的，例如包絡確定、信號合成及組合也可以在該頻譜域中部分地或全部地執行。因此，本發明的實施方式，即，本發明所需的特定程序是在時域還是在頻譜域中實施的，總是可以完全適應於一特定應用所需的一典型解碼器設計的相應框架。在一

LPC語音編碼器的上下文中，本發明的上下文甚至是靈活的，例如，執行一LPC激勵信號(例如，TCX信號)的一頻率增強。該合成信號與該源音訊信號的組合是在該LPC時域中執行的，並且從該LPC時域到該正常時域的該最終轉換是通過一LPC合成濾波器來執行的，其中，具體地，在該LPC合成濾波器階段內執行該合成信號的一典型較佳的包絡調整，用於該合成信號的該至少一個頻帶所代表的對應頻譜部分。因此，通常需要的後處理運算在一單一個濾波階段中與包絡調整組合。這樣的後處理操作可能涉及LPC合成濾波、語音解碼器已知的去加重濾波、例如低音後處理的其它後處理操作、或基於在TCX解碼器或其他解碼器中的長期預測(Long Term Prediction, LTP)的其它聲音增強後處理程序。

【0028】 隨後參考圖式討論本發明的優選實施例，其中：

【圖式簡單說明】

【0029】

第1圖是根據本發明一音訊處理器的一實施例的一框圖。

第2圖是第1圖的該包絡確定器的一較佳實施例的一更詳細描述。

第3a圖是用於計算一子頻帶或全頻帶的該時間包絡的一實施例。

第3b圖是一時間包絡的產生的一替代實施方式。

第3c圖示出了用於使用一希伯特轉換來確定該分析信號的一實施例的一流程圖。

第4圖示出了第1圖的該分析器的一較佳實施方式。

第5圖示出了第1圖的該信號合成器的一較佳實施方式。

第6圖示出了在一核心解碼器上下文中作為一元件或方法的該音訊處理器的一較佳實施例。

第7圖示出了一較佳實施方式，其在該LPC域中執行該合成信號與該源音訊信號的該組合。

第8圖示出了本發明的另一個實施例，其中在該頻譜域中執行該高通濾波或帶通濾波、該包絡調整及該源音訊信號與該合成信號的該組合。

第9a圖示出了對於一聲音項目“德國男性語音”的頻率增強處理中的數個信號。

第9b圖示出了用於該聲音項目“德國男性語音”的一頻譜圖。

第10a圖示出了對於一聲音項目“音調管”的頻率增強處理中的數個信號。

第10b圖示出了用於該聲音項目“音調管”的一頻譜圖。

第11a圖示出了對於一聲音項目“麥當娜時尚”的頻率增強處理中的數個信號。

第11b圖示出了用於該聲音項目“麥當娜時尚”的一頻譜圖。

第12圖示出了另一個實施例，其在某種程度上類似於第6圖的實施例。

【實施方式】

【0030】 第1圖示出了一種音訊處理器，用於從一源音訊信號在一組合器400的輸出產生一頻率增強音訊信號420，該源音訊信號一方面輸入到一包絡確定器100中，另一方面輸入到該組合器400中。

【0031】 該包絡確定器100配置成用於確定該源音訊信號的至少一部分的一時間包絡。該包絡確定器可以使用該全頻帶源音訊信號，也可以僅使用該源音訊信號的一頻帶或一部分，該頻帶或該部分具有一特定較低的邊界頻率，例如100、200或500赫茲(Hz)。該時間包絡從該包絡確定器100轉發到一分析器

200，用於分析該時間包絡以確定該時間包絡的多個特定特徵的多個數值。該多個特定特徵例如可以是該時間包絡的局部最大值、該時間包絡的局部最小值、該時間包絡的零交叉點或兩個局部最小值之間的點或兩個局部最大值之間的點，其中，例如這些特徵之間的點是與相鄰特徵具有相同時間距離的數值。因此，這些特定特徵也可以是兩個局部最小值或兩個局部最大值之間的中間點。然而，在較佳實施例中，優選地使用例如一曲線微積分(**curve calculus**)處理來確定該時間包絡的局部最大值。該時間包絡的該多個特定特徵的時間數值被轉發到一信號合成器300用於一產生合成信號。該合成信號的該產生包含放置有關於該確定後的時間數值的多個脈衝，其中在放置之前或放置之後，在該多個脈衝放置的位置，使用從該時間包絡的多個振幅所導出的多個權重值來加權該多個脈衝，該多個振幅有關於從該分析器接收的該多個時間數值或有關於該多個時間數值。

【0032】該合成信號的至少一個頻帶、或該合成信號的該全高頻帶、或該合成信號的數個單獨且不同的頻帶、甚至該整個合成信號被轉發到該組合器400，用於組合不包含在該源音訊信號的該合成信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

【0033】在一較佳實施例中，該包絡確定器被配置成如第2圖所示。在這個實施例中，該源音訊信號或該源音訊信號的至少一部分被分解成複數個子頻帶信號，如框105所示。如框110所示，一個或多個或甚至所有子頻帶被選定或使用，用於確定每一個(選定的)子頻帶的多個單獨時間包絡，如框120所示。如第2圖中的框125所示，該多個時間包絡被標準化或被濾波，並且，如框130所示，該多個單獨時間包絡相互組合，以在該包絡確定器的該輸出中獲得

該最終時間包絡。該最終時間包絡可以是通過如第2圖所示的該程序所確定的一組合後包絡。取決於該實施方式，可以提供一附加濾波階段115，以利於對該單獨選定的子頻帶進行標準化或濾波。如果使用所有子頻帶，則對所有這樣的子頻帶進行標準化或濾波，如在第2圖中指示該附加濾波階段115的框所示。當在框120中來自該多個時間包絡的該多個子頻帶已經被標準化或相應地被濾波時，框125所指示的該標準化程序可以被旁通，並且不執行該確定後的時間包絡的標準化或濾波的這個程序非常有用。自然地，也可以同時執行該附加濾波階段115的程序、框125的程序，或者甚至也可以僅執行確定用於每一個(選定的)子頻帶120的該時間包絡的程序。

【0034】在進一步的實施方式中，在框105中的該分解不能被執行，但是可以由具有一低交叉頻率的一高通濾波來替代，例如交叉頻率為20、50、100或例如低於500 Hz的一頻率，並且從這個高通濾波的結果只有一單一時間包絡被確定。自然地，該高通濾波也是可以避免的，並且從該源音訊信號僅有一單一時間包絡(通常是該源音訊信號的幀)被導出，其中，最好在典型地重疊的幀中處理該源音訊信號，但是甚至也可以使用非重疊的幀。在框110中指示的一選定在一特定場景中被實現，例如，當確定一特定子頻帶信號無法滿足有關子頻帶信號特徵的特定標準時，或出於任何原因而被排除在該最終時間包絡的該確定之外時。

【0035】第5圖示出了該信號合成器300的一較佳實施方式。該信號合成器300從該分析器200接收該多個特徵的時間數值及附加地關於該包絡的其他資訊作為一輸入。在框310中，第5圖所示的該信號合成器300從有關於該時間數值的該時間包絡中導出縮放因子。因此，框310一方面接收包絡資訊，例如包

絡振幅，另一方面接收時間數值。例如，使用一壓縮函數來執行該縮放因子的推導，該壓縮函數例如一平方根函數、一幂小於1的一幂函數或一對數函數。

【0036】該信號合成器300包含在該多個時間值放置305多個脈衝的該程序，其中，較佳地，僅放置負脈衝或僅放置正脈衝，以便於具有與一脈衝相關的該相關頻譜值的同步相位。然而，在其它實施例中，通常當該基帶信號的該音調不夠高時，多個脈衝的一隨機放置，例如，取決於從典型可用的間隙填充或頻寬延伸輔助資訊所導出的其它標準。負脈衝或正脈衝的該放置可以通過該原始波形的該極性來控制。該多個脈衝的極性可以選擇，使得其等於具有最高波峰因子的該原始波形的該極性。換句話說，這意味著正峰值是由正脈衝所模擬，反之亦然。

【0037】在步驟315中，使用框310的結果來縮放由框305所獲得的該脈衝，並且對該脈衝執行一可選的後處理320。該脈衝信號可以使用，並且如框325所示，該脈衝信號是經過高通濾波或帶通濾波的，以便於獲得該脈衝信號的一頻率頻帶，即，至少獲得轉發到該組合器的該合成信號的該頻帶。然而，一可選的頻譜包絡調整330應用於由如框325所示的該濾波階段所輸出的該信號，其中這樣的頻譜包絡調整是通過一特定包絡函數、或從輔助資訊所導出的包絡參數的一特定選擇、或者作為替代方案，例如在盲頻寬延伸應用中從該源音訊信號所導出。

【0038】第6圖示出了用於產生一頻率增強音訊信號的一音訊處理的一種音訊處理器或方法的一較佳實施例。稱為波形包絡同步脈衝激勵(Waveform Envelope Synchronized Pulse Excitation, WESPE)的發明方法是基於一類脈衝列信號的產生，其中該實際脈衝放置同步到一專用時域包絡。這個所謂的公共包絡

是從該LF信號所導出的，該LF信號是通過一組帶通信號在核心解碼器20的該輸出所獲得的，其各自的包絡組合成一個公共包絡。

【0039】 第6圖示出了該WESPE處理進入具有頻寬延伸(bandwidth extension, BWE)特徵的一音訊解碼器的一典型整合，其也是這個新技術的一較佳實施例。這個實施方式在持續時間例如為20毫秒(ms)的時間幀進行操作，可選地，幀之間具有時間重疊，例如50%。

【0040】 新提出的WESPE BWE的優點

- 減少粗糙度和拍打偽像
- 信號的諧波連續
- 保留脈衝
- 保留脈衝
- 也可以處理音樂
- BWE交叉已經可以開始於2 kHz或更低
- 關於音調、音調對準、諧波、相位的自調節BWE

【0041】 該WESPE處理包括以下步驟：

【0042】 1. 時間包絡估計(如包絡確定器100所示)：從該核心解碼器20獲得的該LF信號被分離(105)成帶通信號的一集合。接下來，為每一個帶通信號確定該時間包絡120。可選地，可以對各個包絡進行標準化(normalization)或濾波。然後，將所有時間包絡組合到一個公共包絡中130。較佳地，組合操作是一平均化處理(averaging process)。

【0043】 2. 同步脈衝放置：通過曲線微積分(curve calculus)的應用來分析205從步驟1中得到的該公共包絡，較佳地，是針對其局部最大值的位置分析。

該獲得的最大值候選可選地對它們210的時間距離進行後選定(post-selected)或穩定化。一狄拉克脈衝放置305在該估計“原始”信號中，用於在每一個最大值位置的產生高頻信號(HF)。可選地，這個處理可以由輔助資訊支持。

【0044】 3. 從包絡導出各自的脈衝大小縮放：通過從該公共包絡所導出的時間權重值來加權315在先前的步驟2中組裝的該脈衝列。

【0045】 4. 後處理、HF提取或間隙填充選擇：可選地，對在步驟3中產生的該“原始”信號進行後處理320，例如，通過雜訊相加，並且對其進行濾波(如框325所示)，用於作為BWE中的HF或作為一間隙填充目標圖塊信號。

【0046】 5. 能量調整：調整330來自步驟4的濾波信號的頻譜能量分佈，用於作為BWE中的HF或作為一間隙填充目標圖塊信號。在此，使用來自該期望能量分佈上的該位元流的輔助資訊40。

【0047】 6. HF或間隙填充信號與LF的混合：最後，根據常規的BWE或間隙填充原理，將來自步驟5的該調整信號與該核心編碼器輸出30混合(如該組合器400所示)，即通過一HP濾波並且補充該LF，或在間隙填充頻譜區域中填充頻譜空洞。

【0048】 在下文中，將進一步給出示例信號及其對處理結果的影響，進一步說明WESPE處理中包括的每個步驟的功能。

【0049】 正確的時間公共包絡估計是WESPE的一關鍵部分。該公共包絡允許估計平均值以及每個時間幀的代表性感知屬性。

【0050】 如果該LF信號的音調為 f_0 且具有很強的 Δf_0 間隔的泛音線頻譜，因此音調非常高，則在各個帶通信號的每一個中都會出現幾條線，如果它們的通帶寬度可以容納它們，則通過所有帶通頻帶內的跳動產生強的相干包絡調

製。時間包絡的平均將保留在帶通包絡中發現的這種相干包絡調製結構，並將在相距 $\Delta T_0 = 1/(\Delta f_0)$ 的近似等距位置處產生強峰值。後來，通過應用曲線微積分，強脈衝將被放置在這些峰值位置，從而形成一脈衝列，其具有由位置 $n * \Delta f_0$ ， $n = 1 \dots N$ 處的離散等距線組成的頻譜。

【0051】 如果一個強音調信號根本沒有泛音，或者該帶通濾波器的該頻寬不能在每個單獨的頻帶中容納一個以上的泛音，則該調製結構不會出現在所有帶通信號中，因此不會控制該平均公共包絡。該結果脈衝放置將基於大部分不規則間隔的最大值，並且會因此吵雜。

【0052】 對於在該公共包絡信號中表現出一隨機局部最大值放置的吵雜 LF 信號，情況也是如此：這會導致偽隨機脈衝放置。

【0053】 瞬態事件被保留下來，因為在這樣的情況下，所有帶通信號都共享一個時間對齊的公共最大值，因此這也會出現在該公共包絡中。

【0054】 帶通應該確定尺寸，這樣它們可以跨越一個感知的頻帶，並且可以容納至少 2 個泛音，用於需要解決的最高頻率。為了獲得更好的平均，帶通可能會有一些過渡頻帶重疊。這樣一來，該估計信號的該音調本質上適應於該 LF 信號。帶通可能會排除非常低的頻率，例如低於 20 Hz。

【0055】 同步時間脈衝放置和縮放是 WESPE 的另一個關鍵貢獻。該同步脈衝放置繼承了在該公共包絡的該時間調製中的該代表性感知屬性，並且壓印成一可感知的原始全頻帶信號。

【0056】 請注意，通過在關鍵頻帶包絡中的調製的評估，已知人們對高頻內容的感知會起作用。如之前已經詳細描述的，同步到該公共 LF 包絡的時間

脈衝放置會強制LF和HF之間在感知上相關的時間和頻譜結構的相似性和對齊性。

【0057】 如果信號音調強烈且音調清晰，像音調管，WESPE可以通過附加的可選的穩定化來確保脈衝放置完全等距，從而導致該“原始”信號的一非常高音調的HF泛音頻譜。

【0058】 以該公共包絡對該脈衝進行加權可以確保在強脈衝中保留主要的調製，而不太重要的調製會導致弱脈衝，這進一步有助於該“原始”信號對LF信號的固有適應的WESPE特性。

【0059】 在一吵雜信號的情況下，如果脈衝放置及加權變得越來越隨機，這將導致一逐漸吵雜的“原始”信號，這是非常需要的特性。

【0060】 其餘的處理步驟、HF提取、能量調節和混合是進一步的步驟，這些都是將新穎的WESPE處理整合到一編解碼器中以適應BWE或間隙填充的全部功能所必需的。

【0061】 第3a圖示出了用於確定該時間包絡的一較佳實施方式。如框135所示，使用一希伯特轉換來確定該分析信號。框135的輸出，即希伯特轉換信號，用於計算如框140所示的該包絡 $ENV(t)$ 。為此，該包絡的計算是將特定時刻的該原始源音訊信號的時間數值進行平方，將該特定時刻的該對應希伯特轉換值進行平方，並且相加該些平方後的值，並且為了每一個各自的時刻的相加結果計算平方根。通過這樣的程序，可以以與該原始源音訊信號 $a(t)$ 相同的取樣解析度來確定該時間包絡。自然地，當進入框135及140的該輸入是由框105所獲得的，或由框110選擇的，或由第2圖的該附加濾波階段115標準化後及濾波後的一子頻帶信號時，執行相同的程序。

【0062】 用於計算一時間包絡的另一個程序在第3b圖的框145及150示出。為此，對該源音訊信號或來自該源音頻的子頻帶的波形進行整流(145)，並且對該整流後波形進行低通濾波(150)，並且該低通濾波的結果是該源音訊信號的嘎包絡或是一各自子頻帶信號的包絡，其與來自其他子頻帶的其他此類包絡組合，較佳地通過第2圖的框130所示的平均。

【0063】 C. Jarne在2017年3月20日發表的“Simple empirical algorithm to obtain signal envelope in the three steps”示出了用於計算時間包絡的其它程序，例如通過具有有限支持的一滑動窗口計算該波形的一瞬時均方根(RMS)值。其它程序包含計算該波形的一分段線性近似，其中通過在移動通過該數據的一窗口中找到並且連接該波形的多個峰值所創建的該振幅包絡。進一步的程序依賴於確定該源音訊信號或該子頻帶信號中的永久峰值以及通過內插法導出該包絡。

【0064】 用於計算該時間包絡的其他程序包含解釋代表該包絡的輔助資訊、或根據從時間雜訊塑形(Temporal Noise Shaping, TNS)所知的一時域幀所導出的一組頻譜值，在該頻譜域中進行一預測。

【0065】 第3c圖示出了使用第3a圖中指示為135的該希伯特轉換來確定該分析信號的一較佳實施方式。例如，He Lei等人發表在INTER_SPEECH 2016-1447，第530-534頁的“A Praat-Based Algorithm to Extract the Amplitude Envelope and Temporal Fine Structure Using the Hilbert Transform”說明了用於計算這類希伯特轉換的程序。在步驟136中，從該信號 $a(t)$ ，例如該源音訊信號或一子頻帶信號，計算一複合頻譜。在步驟137中，選擇該複合頻譜的一正部或刪除負

部。在步驟138中，將該複合頻譜的該正部乘以"-j"，並且在步驟139中，該相乘結果轉換成該時域，並且通過採用該虛部來獲得該分析信號 $\hat{a}(t)$ 。

【0066】自然地，還有許多確定時間包絡的程序，需要注意的是，時間包絡並不一定要實際上“包絡”該時域信號，但是，當然可以是該時域信號的某些最大值或最小值大於或小於這個時間點上的該相應的包絡值。

【0067】第4圖示出了確定該時間包絡的該多個特定特徵的該時間數值的程序的一較佳實施例。為此，將一平均時間包絡引入到框205中，以確定該多個特徵的多個初始時間數值。這些初始時間數值可以是例如該時間包絡中實際找到的最大的時間數值。如框120所示，通過一最佳化函數、或通過一輔助資訊、或通過選擇或操縱該原始特徵，從該原始時間數值中或從該“初始”時間數值中導出放置該實際脈衝的該多個特徵的該最終時間數值。較佳地，框210，根據一處理規則或一最佳化函數來操縱該初始數值。特別是，實現該最佳化函數或處理規則，使得多個時間數值放置在具有一光柵間距T的一光柵中。特別是，在該時間包絡中的該光柵間距T和/或該光柵的該位置使得該多個時間數值及該多個初始時間數值之間的一偏差值具有一預定特性，其中，在實施例中，該偏差值是平方差的一總和，和/或該預定特性是該最小值。因此，在確定該多個初始時間數值之後，放置等距時間數值的一光柵，其與該多個初始時間數值的非恆定光柵盡可能地匹配，但是現在顯示出一清晰且理想的音調行為。與非上取樣域相比，可以在具有一更精細時間粒度的該上取樣域中確定該光柵，或者也可以使用分數延遲，用於以子樣本精度進行脈衝放置。

【0068】第7圖示出了在LPC處理的情況下，本發明的另一實施例。例如，如第1圖或第6圖所示，第7圖的該音訊處理器包含該包絡確定器100、該分

析器200(均未在第7圖中示出)和該信號合成器300。然而，與第6圖相反，該核心解碼器輸出數據，即該LF輸出30不是一時域音訊信號，而是在該LPC時域中的一音訊信號。通常可以在一轉換編碼激勵(Transform Coded eXcitation, TCX)編碼器中找到這種數據作為一內部信號表示。

【0069】 由第7圖中的該音訊解碼器20產生的該TCX數據被轉發到第7圖中所示作為該LPC域加法器405的該混合器。該信號合成器產生TCX頻率增強數據。因此，在這個實施例中，由該信號合成器產生的該合成信號是從作為一TCX數據信號的該源音訊信號中所導出。因此，在框405的該輸出，可獲得一頻率增強音訊信號，然而其仍在該LPC時域中。一隨後連接的LPC合成濾波器410執行該LPC時域信號到該時域的該轉換。

【0070】 該LPC合成濾波器配置用以如果需要附加地執行一去加重，並且附加地，這個時域濾波器配置成用以也對於該合成信號頻帶執行該頻譜包絡調整。因此，第7圖中的該LPC合成濾波器410不僅執行由該音頻解碼器20輸出的該TCX數據頻率範圍的合成濾波，而且還不包含在由該音頻解碼器20輸出的該TCX數據的該頻譜頻帶中的該數據執行頻譜包絡調整。通常，這個數據也可以從該編碼音訊信號10獲得，通過該音訊解碼器20，該音頻解碼器20提取用於該核心頻率範圍的該LPC數據40a，並且附加地提取該頻譜包絡調整，用於該高頻帶、或用於IGF(智慧間隙填充)第7圖40b所指示一個或多個頻帶。因此，第1圖中的該組合器或混和器通過該LPC域加法器405與第7圖隨後連接的LPC合成濾波器410來實現，使得指示為420的該LPC合成濾波器410的該輸出是該頻率增強的時域音訊信號。與第6圖的該程序相反，其中在通過該組合器400執行該混合

操作之前執行該頻譜包絡調整330，第7圖在混合或組合兩個信號之後執行高頻帶的包絡調整或填充頻帶。

【0071】第8圖示出了第6圖中示出的該程序的一進一步實施方式。基本上，第6圖的實施方式是在該時域中執行的，使得標記為320、325、330、400的項目完全在時域中執行。可替代地，第8圖的實施方式依賴用於該低頻帶的一頻譜轉換105，然而，這是可選的措施，但是在第8圖中用於該低頻帶的該頻譜轉換105有利於使用第6圖中該帶通濾波器組105的實施方式。附加地，第8圖的實施方式包括一頻譜轉換器345，用於該轉換該脈衝處理器340的輸出，該脈衝處理器340通常包括第6圖的該脈衝放置305和該脈衝縮放315。第8圖中的該脈衝處理器340附加地可以包含一穩定器框210作為一可選特徵，以及該極值搜索框205作為可選特徵。

【0072】然而，如框325所示的高通濾波、包絡調整330及該低頻帶與該高頻帶的組合的程序是通過一合成濾波器組來完成，即在該頻譜域中完成的，並且第8圖中的該合成濾波器組(實現在該組合器400中)的該輸出為該時域頻率增強音訊信號420。然而，如果該組合器400被實現為用於組合不同頻帶的一簡單組合器，則該組合器400的該輸出也可以是一全頻譜域信號，該信號通常由隨後以任何所需方式進一步處理的頻譜值的框組成。

【0073】以下，給出了三個示例的特徵信號，這些特徵信號已使用WESPE BWE進行了頻寬延伸。取樣率為32 kHz，使用具有513行單面頻譜的DFT(第8圖中的框105)提取8個重疊的帶通信號。為了實現4 kHz高通濾波(第8圖中的框325)、該頻譜包絡調整(第8圖中的框330)及LF與HF的該混合(第8圖中的該組合器400)，具有50%重疊的一相似的DFT/IDFT(第8圖中的框345)被組織為

16個均勻縮放因子頻帶。該頻譜圖中顯示的該結果信號是從DC到4 kHz的非編碼PCM，並且由WESPE在從4 kHz到16 kHz的範圍內產生。

【0074】 第9a圖顯示了該波形的一簡短摘錄(一1024個樣本框)、該公共包絡及由WESPE產生的同步及縮放脈衝放置。具有輕微分散的大脈衝大約等距離地放置在一寬週期結構中。

【0075】 第9b圖描繪了整個測試項目的該頻譜圖。混濁語音的該垂直脈衝結構在LF和HF之間保持一致對齊，而擦音則表現出類似HF結構的一雜訊。

【0076】 因此，第9a圖顯示了WESPE如何模擬語音脈衝，顯示一波形、公共包絡及一脈衝產生，其中該項目為“德國男性語音”。第9b圖顯示了WESPE如何模擬語音脈衝並且顯示一頻譜圖。該項目是“德國男性語音”。

【0077】 第10a圖顯示了該波形的一簡短摘錄(一1024個樣本框)、該公共包絡及由WESPE產生的同步及縮放脈衝放置。不同的尖銳脈衝等距地排列在狹窄的周期性結構中。第10b圖描繪了整個測試項目的該頻譜圖。音調管的水平線結構在LF和HF之間保持對齊，但是HF也有些吵雜，並且可以從額外的穩定中受益。

【0078】 第10a圖顯示了WESPE如何模擬諧波，並且顯示一波形、一公共包絡及一脈衝產生。該項目是“音調管”。第10b圖顯示了WESPE如何模擬諧波，並顯示一頻譜圖。該項目是“音調管”。

【0079】 第11a圖顯示了測試項目“麥當娜時尚”的波形的一簡短摘錄(一1024個樣本框)、該公共包絡及由WESPE產生的同步及縮放脈衝放置。脈衝的放置和縮放幾乎具有隨機結構。第11b圖描繪了整個測試項目的該頻譜圖。流

行音樂的該垂直瞬態結構在LF和HF之間保持一致地對齊，而該HF音調通常較低。

【0080】第11a圖顯示了WESPE如何模擬一吵雜混合物，並且顯示一波形、一公共包絡及一脈衝產生。該項目是“時尚”。第11b圖顯示了WESPE如何模擬一吵雜混合物，並顯示一頻譜圖。該項目是“時尚”。

【0081】第9a、10a、11a圖的第一張圖示出了一低頻帶源信號的1024個樣本的一框的一波形。附加地，該分析濾波器對提取樣本的框的影響如下所示，在框的開始處，即在樣本0，該波形等於0，並且在框的結束處，即在樣本1023，該波形也等於0。這樣的波形可以獲得，例如在第1圖的該包絡確定器100的該輸入或第6圖的30。第9a、9b、9c圖中的垂直軸始終表示一時域振幅，而在這些圖中水平軸始終表示時域變量，尤其是，樣本數量通常從0延伸到1023，持續一個框。

【0082】第9a、10a、11a圖的第二張圖示出了該平均低頻帶包絡，尤其是，僅該低頻帶包絡的該正部。自然地，該低頻帶包絡通常是對稱的，並且也延伸到該負範圍內。然而，僅需要該低頻帶包絡的該正部。從第9a、10a、11a圖可見，在這個實施例中，僅在排除該框的前幾個樣本和該框的最後兩個樣本時，計算該包絡，然而，這根本不是問題，因為優選地以重疊的方式計算這些框。因此，第9a、10a、11a圖的第二張圖典型地示出，例如第1圖的該包絡確定器100的該輸出或第2圖的框130的該輸出。

【0083】第9a、10a、11a圖的第三張圖示出了該脈衝縮放後，即在處理之後，的該合成信號，其中該脈衝放置在該包絡的該特徵的該時間包絡的該時間數值處，並且已經通過該包絡的該對應的振幅加權。第9a、10a、11a圖示出

了該放置的脈衝僅從樣本256延伸到樣本768。因此，由該加權後脈衝組成的該信號僅在512個樣本上延伸，並且在這些樣本之前以及這些樣本之後不具有任何部分，即，覆蓋幀的中間部分。這反映了前一幀具有重疊並且後一幀也具有重疊的情況。為了以後續框產生一脈衝信號，在下一個四分之一和最後一個四分之一丟失的情況下，也會處理來自下一個框的脈衝信號，並且因此，來自下一個框的該脈衝信號將緊接在第9a、10a、11a圖中的該當前塊的所示該脈衝信號之後。由於不需要該脈衝信號的任何重疊/相加操作，因此該程序非常有效率。然而，如果需要的話，也可以執行相對於該脈衝信號的從一幀到下一幀的任何重疊/相加過程或任何交叉衰落程序。

【0084】 第9b、10b、11b圖示出了多個頻譜圖。水平軸表示時間，但不是如第9a、10a、11a圖所示的相對於樣本的時間，而是相對於DFT框數的時間。垂直軸示出了從對應圖的底部的低頻到對應圖的頂部的高頻的頻譜。水平範圍從0延伸至16 kHz，使得下方四分之一代表該原始信號，上方四分之三代表該合成信號。因此，第9b、10b、11b圖示出了該頻率增強音訊信號，而這些圖的僅下方四分之一示出了該源音訊信號。

【0085】 這些圖指示出該低頻帶結構在高頻帶中得到了很好的反映。關於第10圖這是特別可見的，第10b圖示出了一音調管，其中在第10b圖中從左到右一個接一個地播放該音調管的三個不同音調。第10b圖左側的第一部分是該音調管的最低音調、中間部分是該音調管的中間音調，而第10b圖的右側部分是該音調管的最高音調。該音調管的特徵在於一非常高的音調頻譜，並且似乎本發明在很好地複製較高的12 kHz的諧波結構中特別有用。

【0086】對於第三個測試項目，顯而易見的是，通過本發明程序，用於這種流行音樂項目的該低頻帶結構很好地轉換到該高頻率範圍。

【0087】第12圖示出了另一實施例，其在某種程度上類似於第6圖的實施例。因此，第6圖中相似的符號表示了第12圖中相似的項目。除了第6圖中的特徵外，第12圖的實施例附加地包含一LF/HF分解器160、一隨機雜訊或偽隨機雜訊產生器170，例如雜訊表等，以及一能量調整器180。

【0088】該LF/HF分解器160將該時間包絡分解成一LF包絡及一HF包絡。較佳地，通過低通濾波來確定該LF包絡，並且通過從該LF包絡中減去該HF包絡來確定該HF包絡。

【0089】該隨機雜訊或偽隨機雜訊產生器170產生一雜訊信號，並且該能量調整器180將該雜訊能量調整成該HF包絡的能量，其同樣在框180中被估計。通過該加法器335將具有調整成該HF包絡(沒有來自該LF包絡的任何貢獻)的該能量的一能量的這個雜訊加入到該加權後脈衝列，如框315輸出。然而，例如，該處理框或步驟315、335的順序也可以改變。

【0090】另一方面，如框160所確定的，關於項目205至315的程序僅應用於該LF包絡。

【0091】依賴將該全頻帶包絡分解至少兩個部分的一較佳實施例包含以下依下列順序或其它可行順序的框或步驟。

【0092】時間包絡估計(如該包絡確定器100所示)：整流；通過使用，例如一函數 $x^{0.75}$ 來壓縮；將該包絡分離160成一LF包絡及一HF包絡。該LF包絡是通過低通濾波所獲得的，其中一交叉頻率例如為2-6 kHz。在一個實施例中，該HF包絡是該原始包絡與一較佳地延遲調整的LF包絡之間的該差值。

【0093】 同步脈衝放置300。通過例如一曲線微積分來分析從上述步驟中所導出的該LF包絡，並且在LF包絡最大值位置上完成一脈衝放置。

【0094】 從包絡導出各自的脈衝大小縮放315：通過從該LF包絡導出的該時間權重值來加權上述步驟中組裝的該脈衝列。

【0095】 估計該HF包絡的該能量，並且將相同能量的隨機雜訊加入335到該加權後脈衝列。

【0096】 後處理、HF提取或間隙填充選擇：對在上述步驟中在框335的輸出所產生的該“原始”信號進行可選地後處理320，例如通過雜訊相加，並且對其進行如框325所示的濾波，用於作為BWE中的HF或作為一間隙填充目標圖塊信號。

【0097】 能量調整330：調整上述步驟中概述的來自能量估計的濾波信號的頻譜能量分佈，用於作為BWE中的HF或作為一間隙填充目標圖塊信號。在此，較佳地使用來自該期望能量分佈上的該位元流的輔助資訊。

【0098】 HF或間隙填充信號與LF的混合(由該組合器400所執行)：最後，根據常規的BWE或間隙填充原理，將來自步驟5的該調整信號與該核心編碼器輸出混合，即通過一HP濾波並且補充該LF，或在間隙填充頻譜區域中填充頻譜空洞。

【0099】 這裡要提到的是，如前所述的所有替代方案或方面以及由所附權利要求書中的獨立請求項限定的所有方面可以單獨使用，即，除了預期的替代方案、目的或獨立請求項外，沒有任何其他替代方案或目的。然而，在其他實施例中，兩個或多個替代方案或方面或獨立請求項可以彼此組合，並且在其他實施例中，所有方面或替代方案和所有獨立請求項可以彼此組合。

【0100】 根據本發明的編碼的音訊信號可以被儲存在數位儲存介質或非暫時性儲存介質上，或者可以在諸如無線傳輸介質的傳輸介質或諸如網路的有線傳輸介質上傳輸。

【0101】 儘管在設備的上下文中描述了一些方面，但是很明顯，這些方面也代表了對應方法的描述，其中框或設備對應於方法步驟或方法步驟的特徵。類似地，在方法步驟的上下文中描述的方面也表示對相應裝置的相應框或項目或特徵的描述。

【0102】 取決於某些實施要求，本發明的實施例可以以硬體或軟體來實現。實施例可以使用數位儲存介質來執行，例如儲存有電子可讀控制信號的軟碟、DV、藍光、CD、ROM、PROM、EPROM、EEPROM或快閃記憶體，它們與可編程電腦系統配合(或能夠配合)，從而執行相應的方法。因此，數位儲存介質可以是電腦可讀的。

【0103】 根據本發明的一些實施例包括具有電子可讀控制信號的數據載體，該電子可讀控制信號能夠與可程式化電腦系統合作，從而執行本文描述的方法之一。

【0104】 通常，本發明的實施例可以被實現為具有程式代碼的電腦程式產品，當電腦程式產品在電腦上運作時，該程式代碼可操作用於執行方法之一。程式代碼可以例如被儲存在機器可讀載體上。

【0105】 其他實施例包括儲存在機器可讀載體上的或非暫時性儲存介質上的，用於執行本文描述的方法之一的電腦程式。

【0106】換句話說，因此，本發明方法的實施例是一種電腦程式，該電腦程式具有當電腦程式在電腦上運行時用於執行本文描述的方法之一的程式代碼。

【0107】因此，本發明方法的另一實施例是一種數據載體(或數位儲存介質、或電腦可讀介質)，其包括記錄在其上的用於執行本文所述方法之一的電腦程式。

【0108】因此，本發明方法的另一實施例是表示用於執行本文所述方法之一的電腦程式的數據流或信號序列。數據流或信號序列可以例如被配置為經由數據通信連接，例如經由網路來傳輸。

【0109】另一實施例包含處理裝置，例如電腦或可程式化邏輯元件，其被配置為或適於執行本文描述的方法之一。

【0110】另一實施例包含一種電腦，該電腦上安裝了用於執行本文描述的方法之一的電腦程式。

【0111】在一些實施例中，可程式化邏輯元件(例如現場可程式化閘陣列)可以用於執行本文描述的方法的一些或全部功能。在一些實施例中，現場可程式化閘陣列可以與微處理器協作以便執行本文描述的方法之一。通常，該方法首選地由任何硬體設備執行。

【0112】上面描述的實施例僅用於說明本發明的原理。應當理解，本文描述的佈置和細節的修改和變化對於本領域的其他技術人員將是顯而易見的。因此，本發明的意圖僅由即將來臨的專利權利要求的範圍限制，而不受通過本文的實施方式的描述和解釋而給出的具體細節的限制。

【0113】 參考文獻

[1]Dietz, M., Liljeryd, L., Kjörling, K., and Kunz, O., "Spectral Band Replication, a Novel Approach in Audio Coding," in Audio Engineering Society Convention 112, 2002.

[2]Disch, S., Niedermeier, A., Helmrich, C. R., Neukam, C., Schmidt, K., Geiger, R., Lecomte, J., Ghido, F., Nagel, F., and Edler, B., "Intelligent Gap Filling in Per-ceptual Transform Coding of Audio," in Audio Engineering Society Convention 141, 2016.

[3]Laitinen M-V., Disch S., Oates C., Pulkki V. "Phase derivative correction of bandwidth extended signals for perceptual audio codecs." In 140th Audio Engineering Society International Convention 2016, AES 2016. Audio Engineering Society. 2016.

[4]Atti, Venkatraman, Venkatesh Krishnan, Duminda A. Dewasurendra, Venkata Chebiyyam, Shaminda Subasingha, Daniel J. Sinder, Vivek Rajendran, Imre Varga, Jon Gibbs, Lei Miao, Volodya Grancharov and Harald Pobloth. "Super-wideband bandwidth extension for speech in the 3GPP EVS codec." 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2015): 5927-5931.

【符號說明】

【0114】

- 10 編碼音訊信號
- 20 解碼器
- 30 輸出

40	輔助資訊
40a	LPC數據
40b	包絡資訊
100	包絡確定器
105	框
110	框
115	附加濾波階段
120	框
125	框
130	框
135	框
136	步驟
137	步驟
138	步驟
139	步驟
140	框
145	框
150	框
160	LF/HF分解器
170	隨機雜訊或偽隨機雜訊產生器
180	能量調整器
200	分析器

- 205 框
- 210 框
- 300 信號合成器
- 305 放置
- 310 框
- 315 步驟
- 320 可選的後處理
- 325 框
- 330 可選的頻譜包絡調整
- 335 加法器
- 340 脈衝處理器
- 345 頻譜轉換器
- 400 組合器
- 405 LPC域加法器
- 410 LPC合成濾波器
- 420 頻率增強音訊信號

【發明申請專利範圍】

【請求項1】 一種音訊處理器，用於從一源音訊信號產生一頻率增強音訊信號，其包含：

一包絡確定器，用於確定該源音訊信號的至少一部分的一時間包絡；

一分析器，用於分析該時間包絡，以確定該時間包絡的多個特定特徵的多個時間數值；

一信號合成器，用於產生一合成信號，該產生包含放置有關於該確定後的時間數值的多個脈衝，其中，在該多個脈衝放置的位置，使用從有關於該多個時間數值的該時間包絡的多個振幅所導出的多個權重值來加權該多個脈衝；以及

一組合器，用於組合不包含在該源音訊信號的該合成信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

【請求項2】 如請求項1所述之音訊處理器，其中該分析器配置成用以確定局部最大值或局部最小值的該多個時間數值作為該多個特定特徵。

【請求項3】 如請求項1所述之音訊處理器，其中該包絡確定器配置成用以將該源音訊信號分解成複數個子頻帶信號；

計算該複數個子頻帶信號的一選定的子頻帶信號的一選定的時間包絡，該選定的時間包絡為該時間包絡；或

從該複數個子頻帶信號的至少兩個子頻帶信號來計算至少兩個時間包絡，並且組合該至少兩個子頻帶信號，以獲得一組合後時間包絡作為該時間包絡。

【請求項4】 如請求項3所述之音訊處理器，其中該包絡確定器配置成用以在組合之前對該選定的子頻帶信號或時間包絡進行標準化或濾波；或

其中該組合包含一平均運算；或

其中該包絡確定器配置成用以從該複數個子頻帶信號的全部子頻帶信號來計算多個時間包絡；或

其中該包絡確定器配置成用以確定該源音訊信號的一單個寬帶時間包絡為該時間包絡。

【請求項5】 如請求項1所述之音訊處理器，其中該包絡確定器配置成用以確定該時間包絡通過：

使用一包絡跟隨器，其配置成用以對一波形進行整流並且對該整流後波形進行低通濾波；或

使用計算出的一數位波形的絕對值或絕對值的幕，並且隨後對一結果進行低通濾波；或

使用通過具有一定義後窗口寬度的一滑動窗口的該波形的一瞬時均方根值的計算；或

對該波形使用一分段線性近似，其中通過找到並且連接在一滑動窗口移動中的該波形的多個峰值來確定該時間包絡，該滑動窗口移動通過該分段線性近似的一結果；或

使用一希伯特轉換，其用於產生用於該波形的一分析信號，並且使用平方運算、加法運算及平方根運算從該源音訊信號及該分析信號計算該包絡時間。

【請求項6】 如請求項1所述之音訊處理器，其中該分析器配置成用以計算該多個特定特徵的多個初始時間數值；以及

使用一最佳化函數、或使用與該源音訊信號相關的一輔助資訊、或根據一處理規則選擇或操縱該附加數值，從該多個初始時間數值中導出該多個時間數值。

【請求項7】 如請求項6所述之音訊處理器，其中該處理規則或該最佳化函數被實現，使得該多個時間數值放置在具有一光柵間距的一光柵中，其中在該時間包絡中的該光柵間距及該光柵的一位置使得該多個時間數值及該多個初始時間數值之間的一偏差值具有一預定特性。

【請求項8】 如請求項7所述之音訊處理器，其中該偏差值是平方差的一總和，並且其中該預定特性是一最小特性。

【請求項9】 如請求項1所述之音訊處理器，其中該信號合成器配置成用以放置僅多個正脈衝或僅多個負脈衝，以獲得一脈衝列；及
隨後加權該脈衝列中的該多個脈衝；或
使用與在一脈衝列中的該多個脈衝的該多個時間數值相關的該多個對應權重值來加權僅該多個負脈衝或僅該多個正脈衝；及
放置該多個加權後脈衝在各自的時間數值上，以獲得該脈衝列。

【請求項10】 如請求項1所述之音訊處理器，其中該信號合成器配置成用以使用一壓縮函數從該多個振幅所導出該多個權重值，該壓縮函數是選自於以下函數組：

一幕小於1的一幕函數、一對數函數、一平方根函數及一非線性函數，用於減少較高的數值並且增加較低的數值。

【請求項11】 如請求項1所述之音訊處理器，其中該信號合成器配置成用以執行一後處理功能，該後處理功能包含以下功能組中的至少一個：雜訊相加、一缺失諧波的相加、反向濾波及包絡調整。

【請求項12】 如請求項1所述之音訊處理器，其中該包絡確定器配置成用以將該時間包絡分解成一低頻部分及一高頻部分，

其中該分析器配置成用以使用該時間包絡的該低頻部分進行分析。

【請求項13】 如請求項12所述之音訊處理器，其中該信號合成器配置成用以產生能量調整後的雜訊，並且將該能量調整後的雜訊加入到包含加權後脈衝或非加權後脈衝的一信號，以獲得一合成信號。

【請求項14】 如請求項1所述之音訊處理器，其中該信號合成器配置成用以對一信號進行高通濾波或進行帶通濾波，該信號包含多個該放置且加權後脈衝，以至少獲得不包含在該源音訊信號的該合成信號的該至少一個頻帶並且以該合成信號的該頻帶來執行一頻譜包絡調整，或其中使用從與該源音訊信號相關的輔助資訊所導出的包絡調整值、或使用從該源音訊信號所導出的包絡調整值、或根據一預定包絡調整函數來執行該頻譜包絡調整。

【請求項15】 如請求項1所述之音訊處理器，其中該源音訊信號是一時域音訊信號；

其中該合成信號的該至少一個頻帶是一時域音訊信號；及

其中該組合器配置成用以使用該合成信號的該至少一個頻帶與該源音訊信號的對應的樣本的一逐一樣本相加來執行一時域組合。

【請求項16】 如請求項1所述之音訊處理器，其中該源音訊信號是一LPC域中的一激勵信號；

其中該合成信號的該至少一個頻帶是該LPC域中的一激勵信號；

其中該組合器配置成用以通過在該LPC域中的一逐一樣本相加來組合該源音訊信號與該至少一個頻帶；

其中該組合器配置成用以使用一LPC合成濾波器對該逐一樣本相加的一結果進行濾波，以獲得該頻率增強音訊信號；以及

其中該LPC合成濾波器由與作為輔助資訊的該源音訊信號相關的LPC數據所控制，並且其中該LPC合成濾波器額外地由包絡資訊所控制，用於該合成信號的該至少一個頻帶。

【請求項17】 如請求項1所述之音訊處理器，其中該分析器、該信號合成器及該組合器在一時域或一LPC時域中運作。

【請求項18】 如請求項1所述之音訊處理器，其中該包絡確定器配置成用以應用一頻譜轉換來提取用於一幀序列的複數個帶通信號；

其中該信號合成器配置成用以應用一頻譜轉換，以提取該合成信號的該至少一個頻帶，並且對該至少一個頻帶執行一包絡調整，以及

其中該組合器配置成用以在該頻譜域中組合並且應用一轉換至一時域，以獲得該頻率增強音頻信號。

【請求項19】 一種音訊處理方法，用於從一源音訊信號產生一頻率增強音訊信號，其包含：

確定該源音訊信號的至少一部分的一時間包絡；

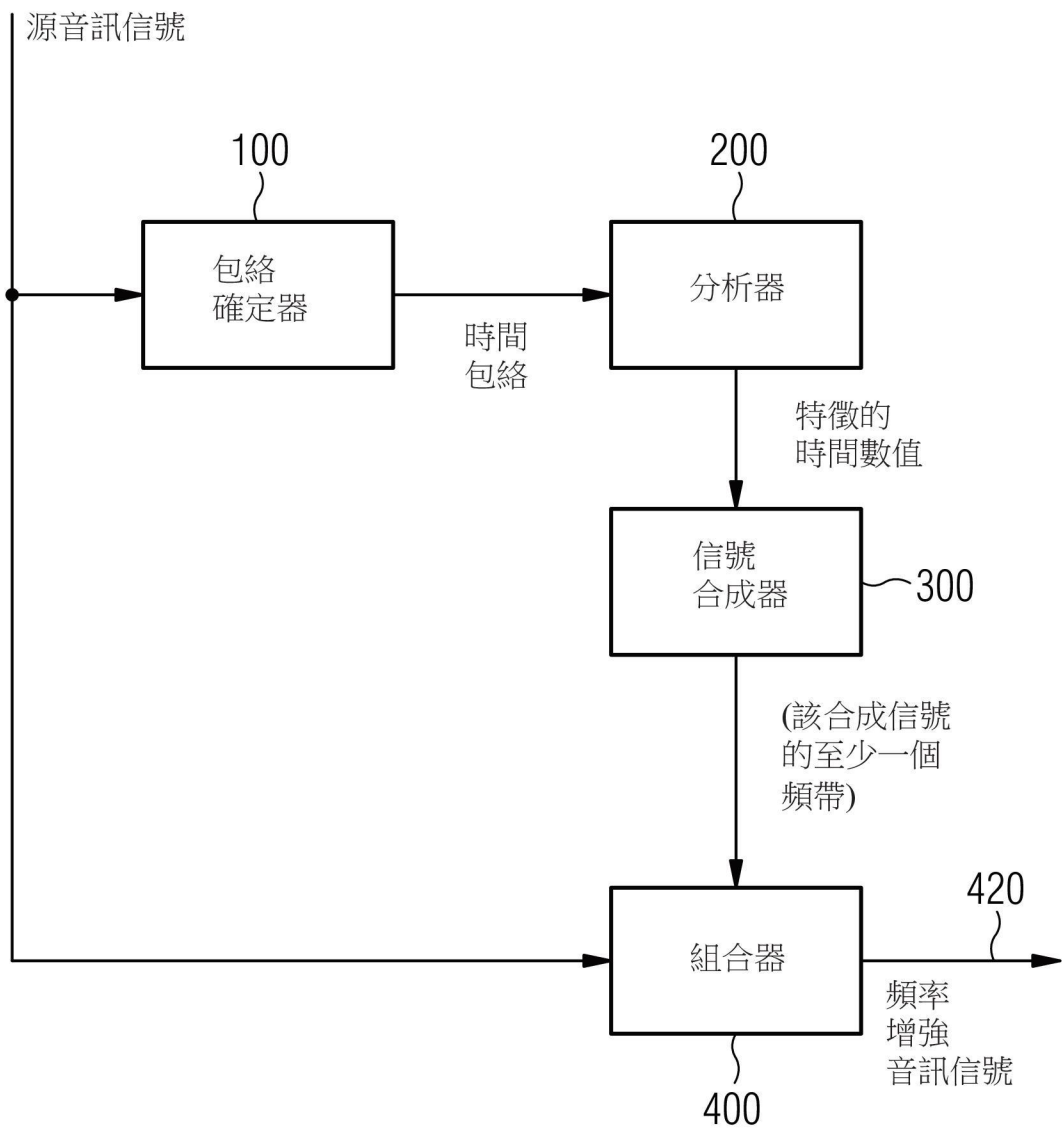
分析該時間包絡，以確定該時間包絡的多個特定特徵的多個時間數值；

在一合成信號中放置有關於該確定後的時間數值的多個脈衝，其中，在該合成信號中，在該多個脈衝放置的位置，使用從有關於該多個時間數值的該時間包絡的多個振幅所導出的多個權重值來加權該多個放置後脈衝；以及

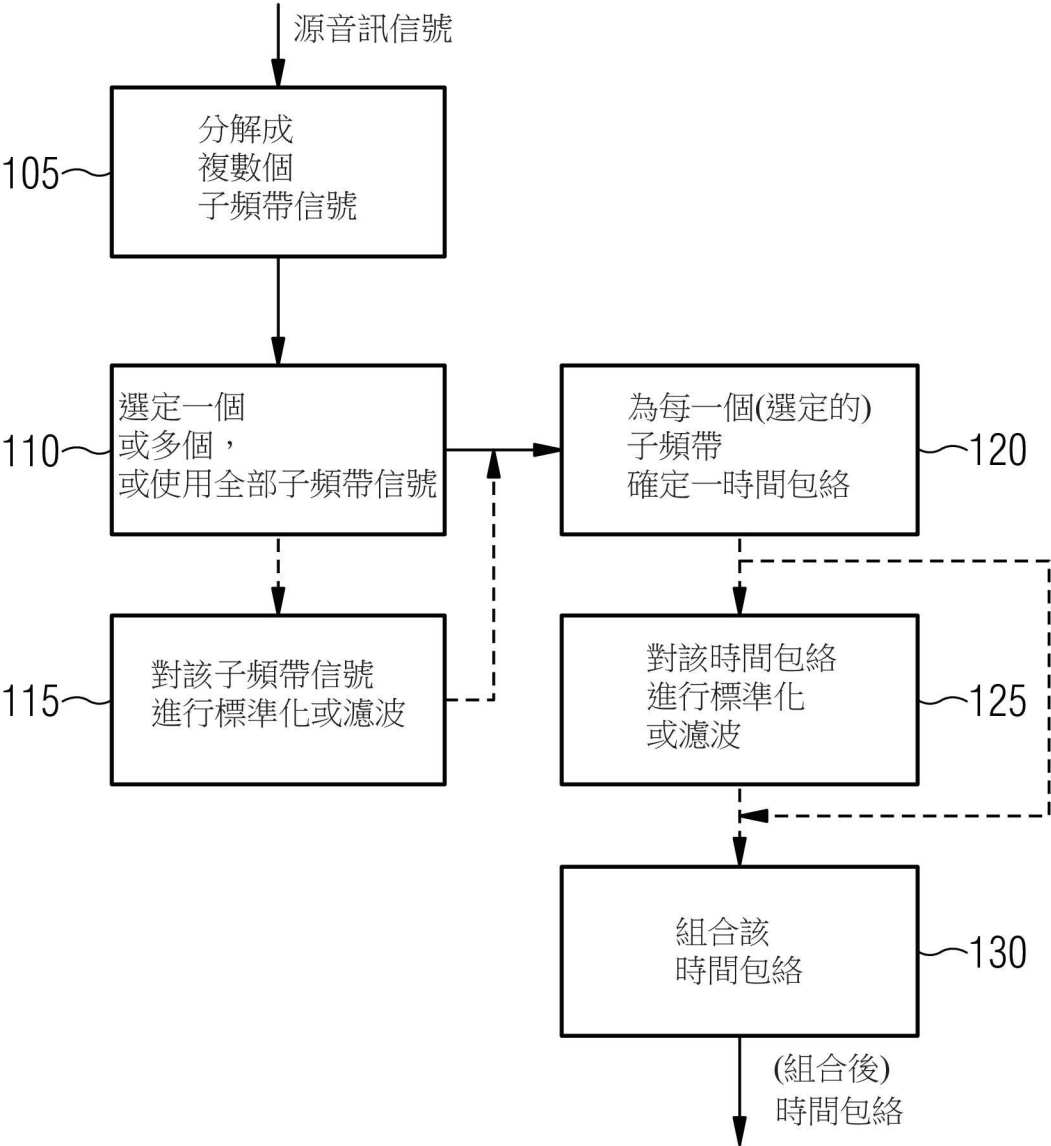
組合不包含在該源音訊信號的該合成信號的至少一個頻帶與該源音訊信號，以獲得該頻率增強音訊信號。

【請求項20】 一種電腦程式，當在一電腦或一處理器上運行時，該電腦程式用於執行如請求項19所述之方法。

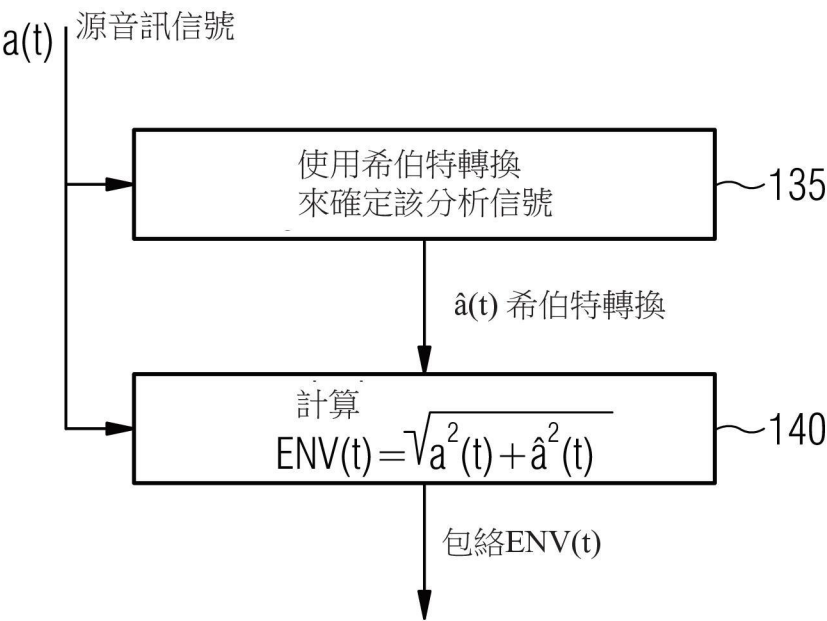
【發明圖式】



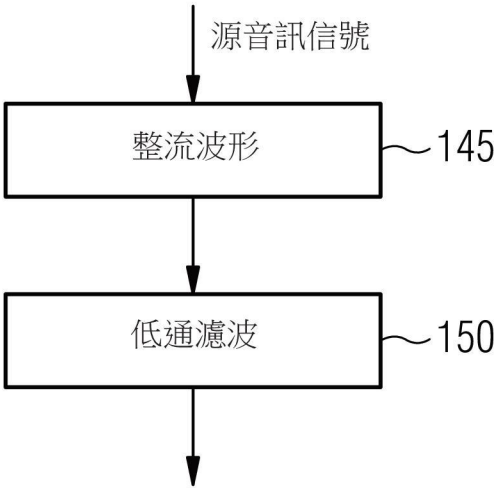
【第 1 圖】



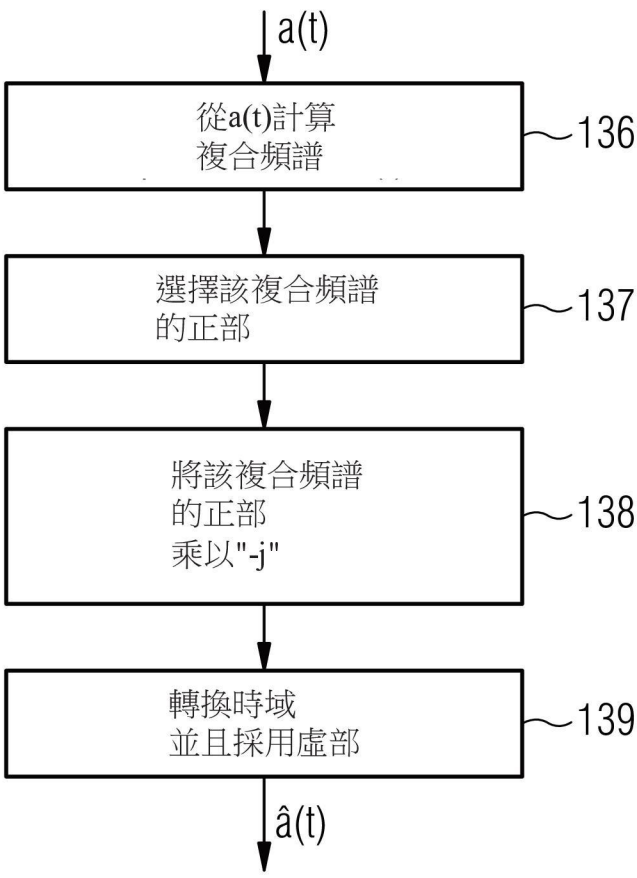
【第 2 圖】



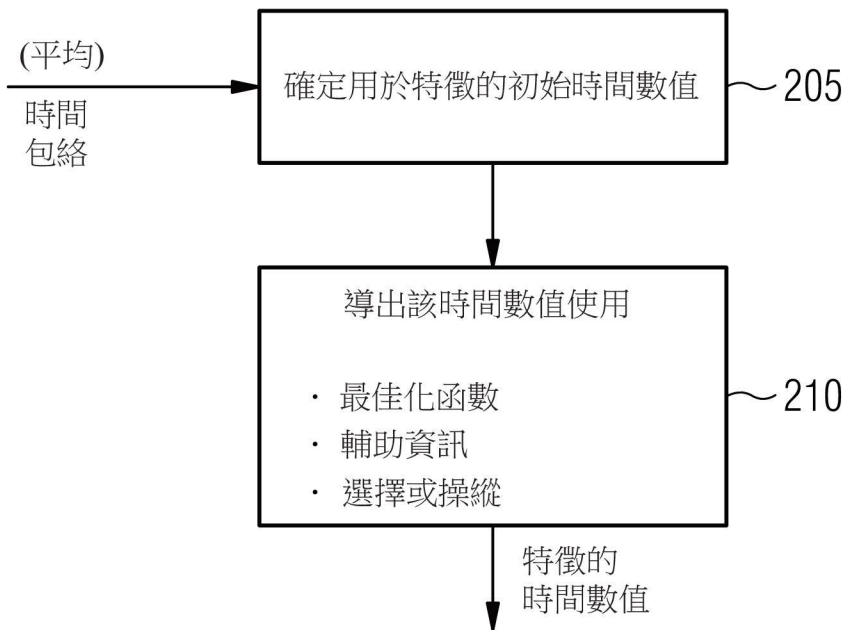
【第 3a 圖】



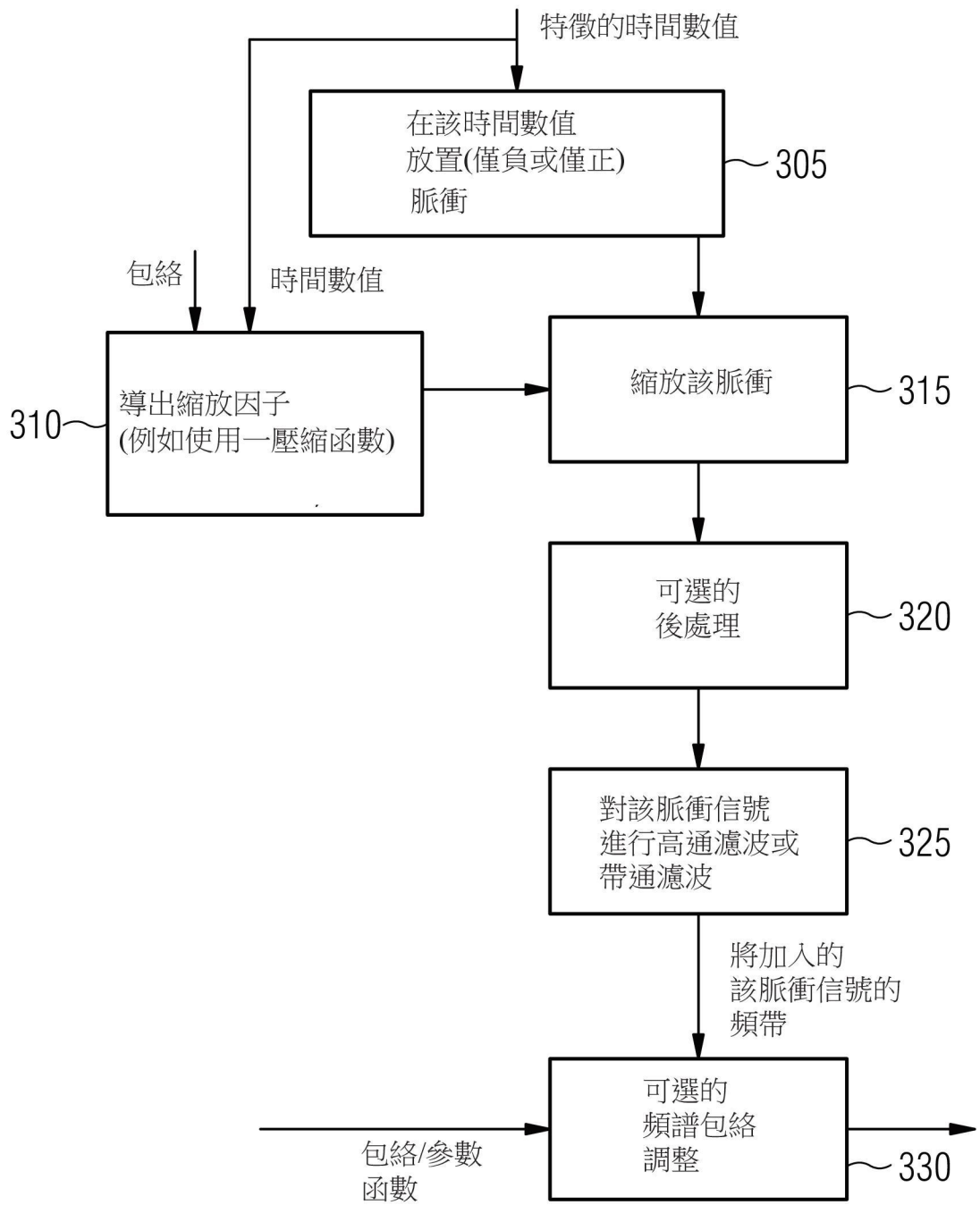
【第 3b 圖】



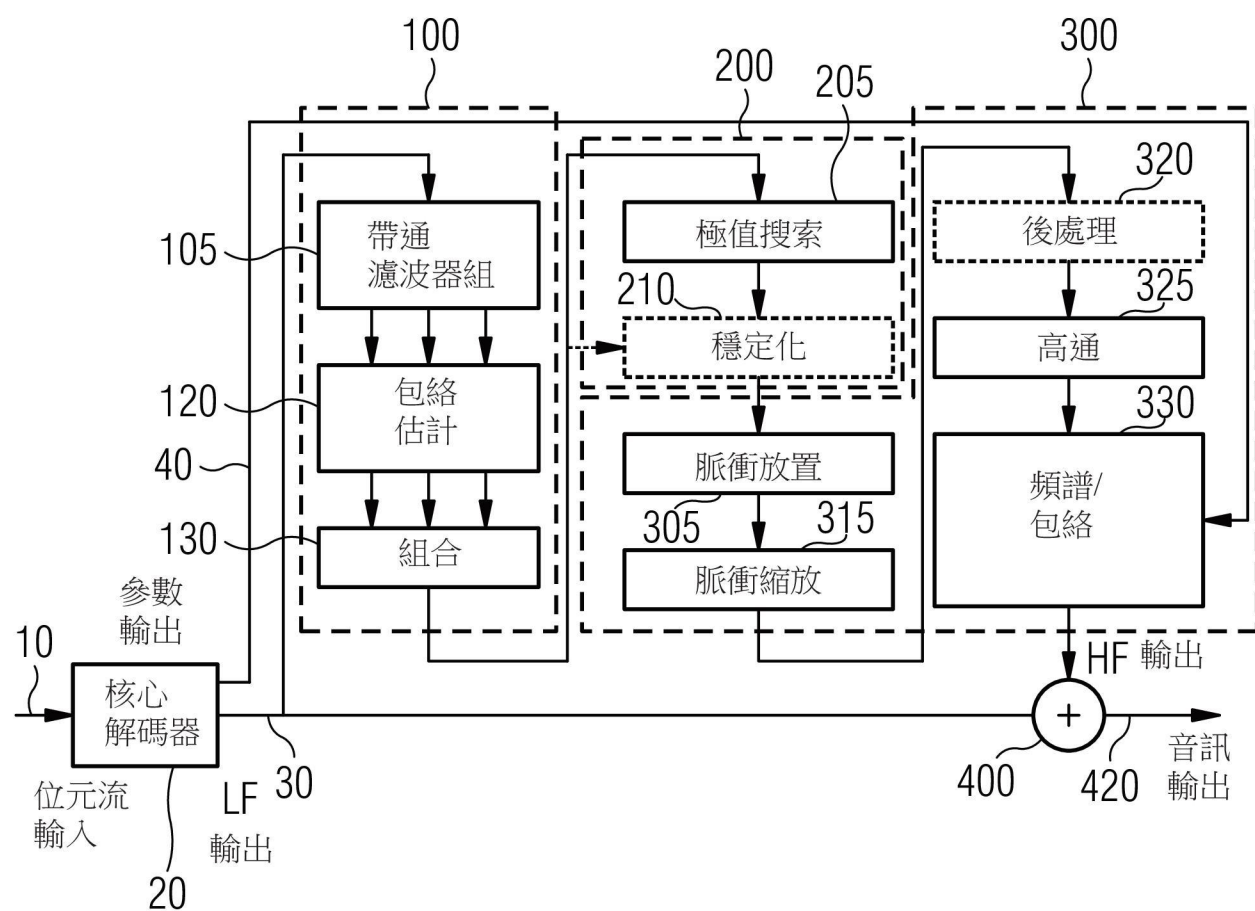
【第 3c 圖】



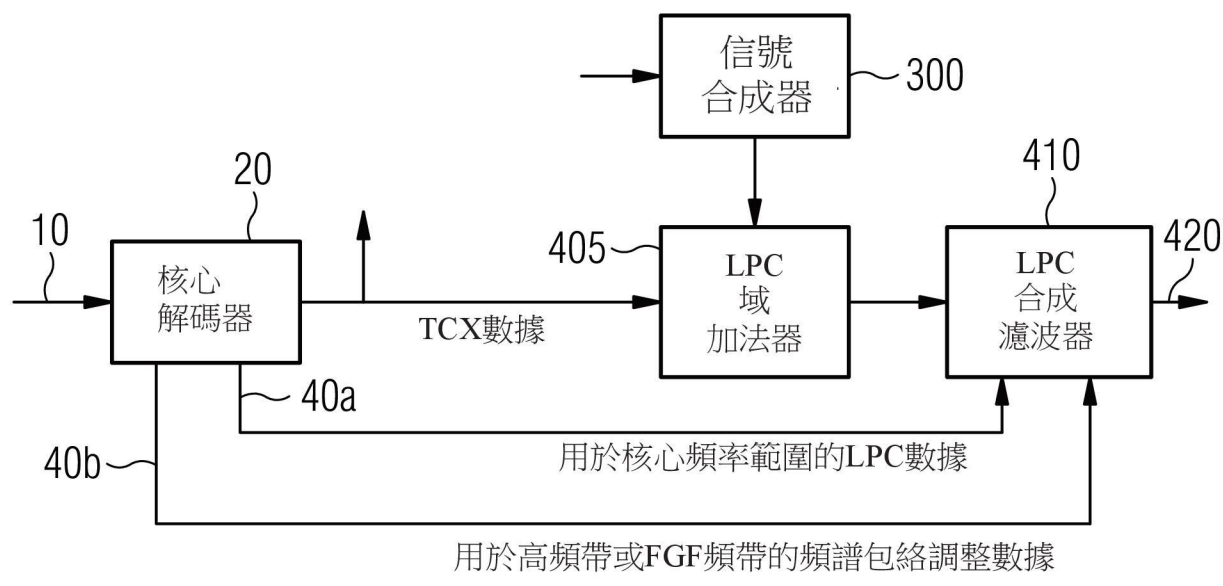
【第 4 圖】



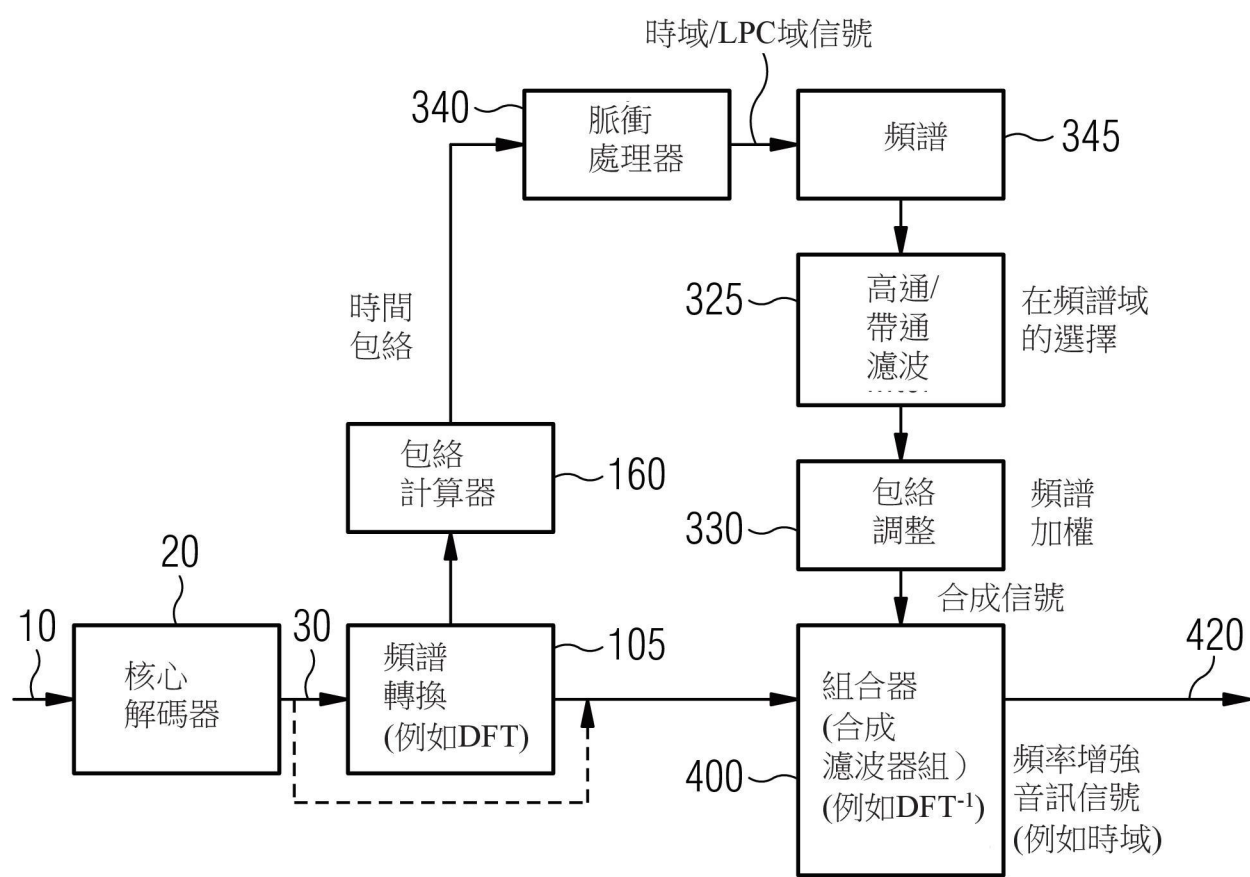
【第 5 圖】



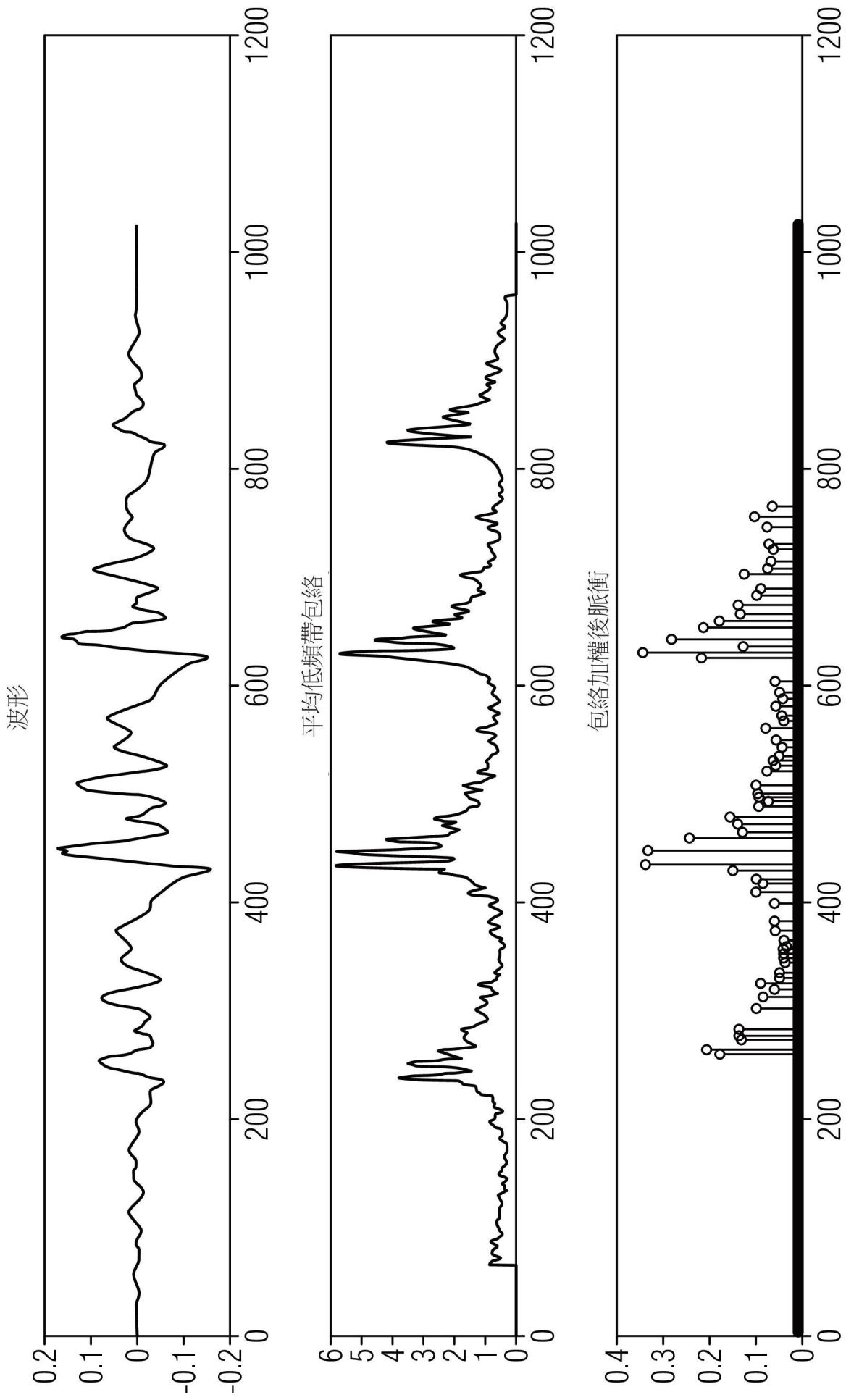
【第 6 圖】



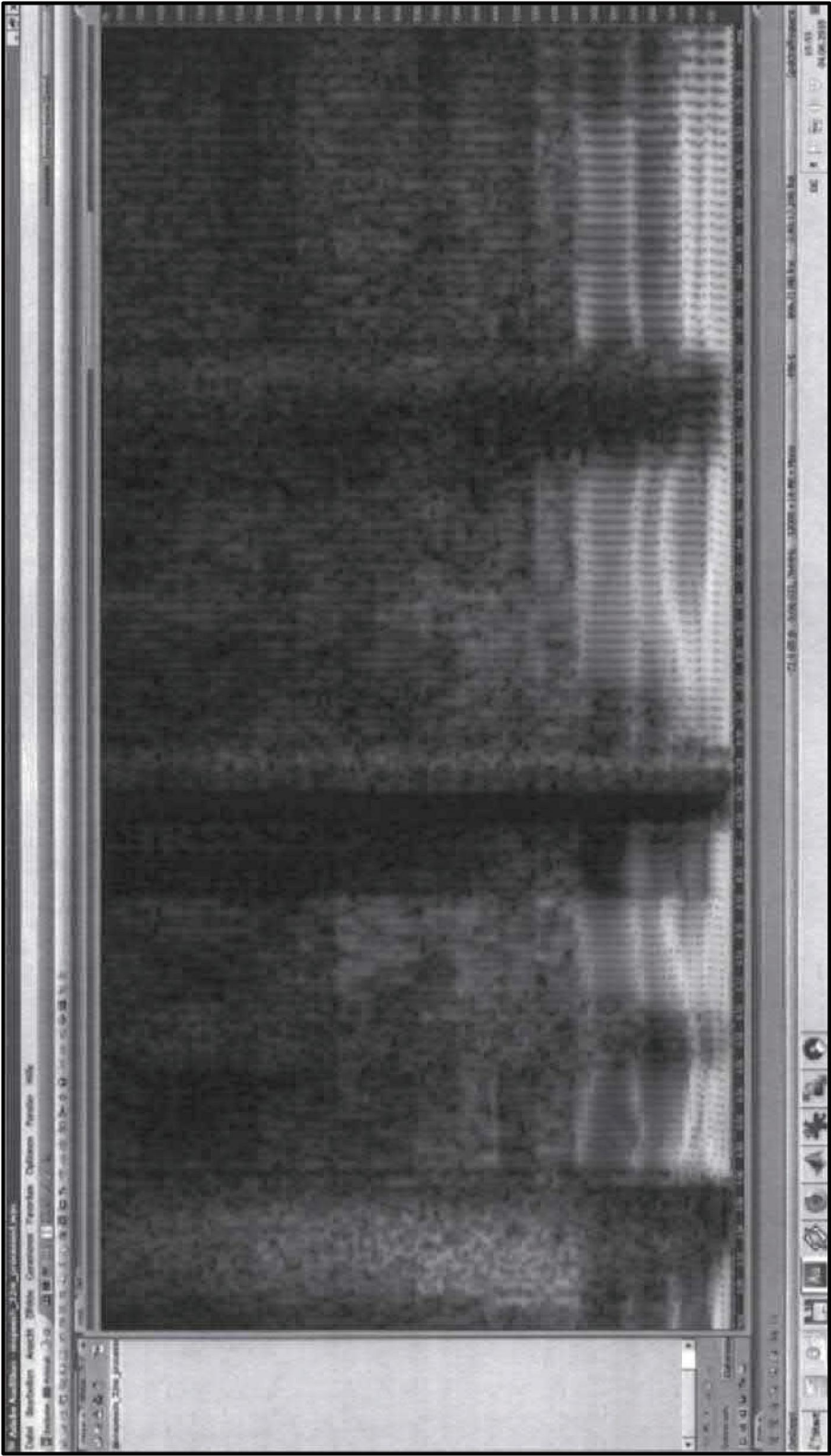
【第 7 圖】



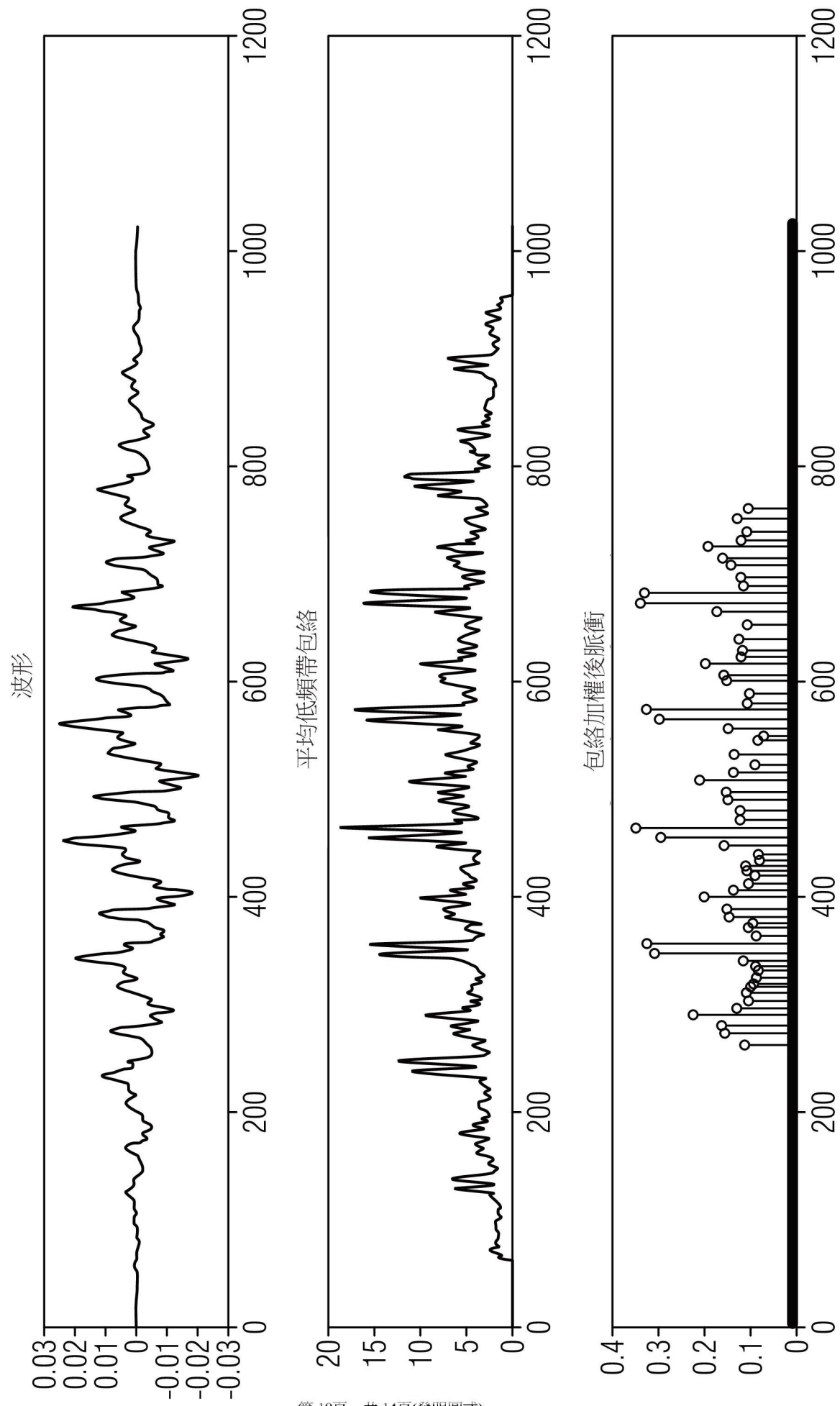
【第 8 圖】



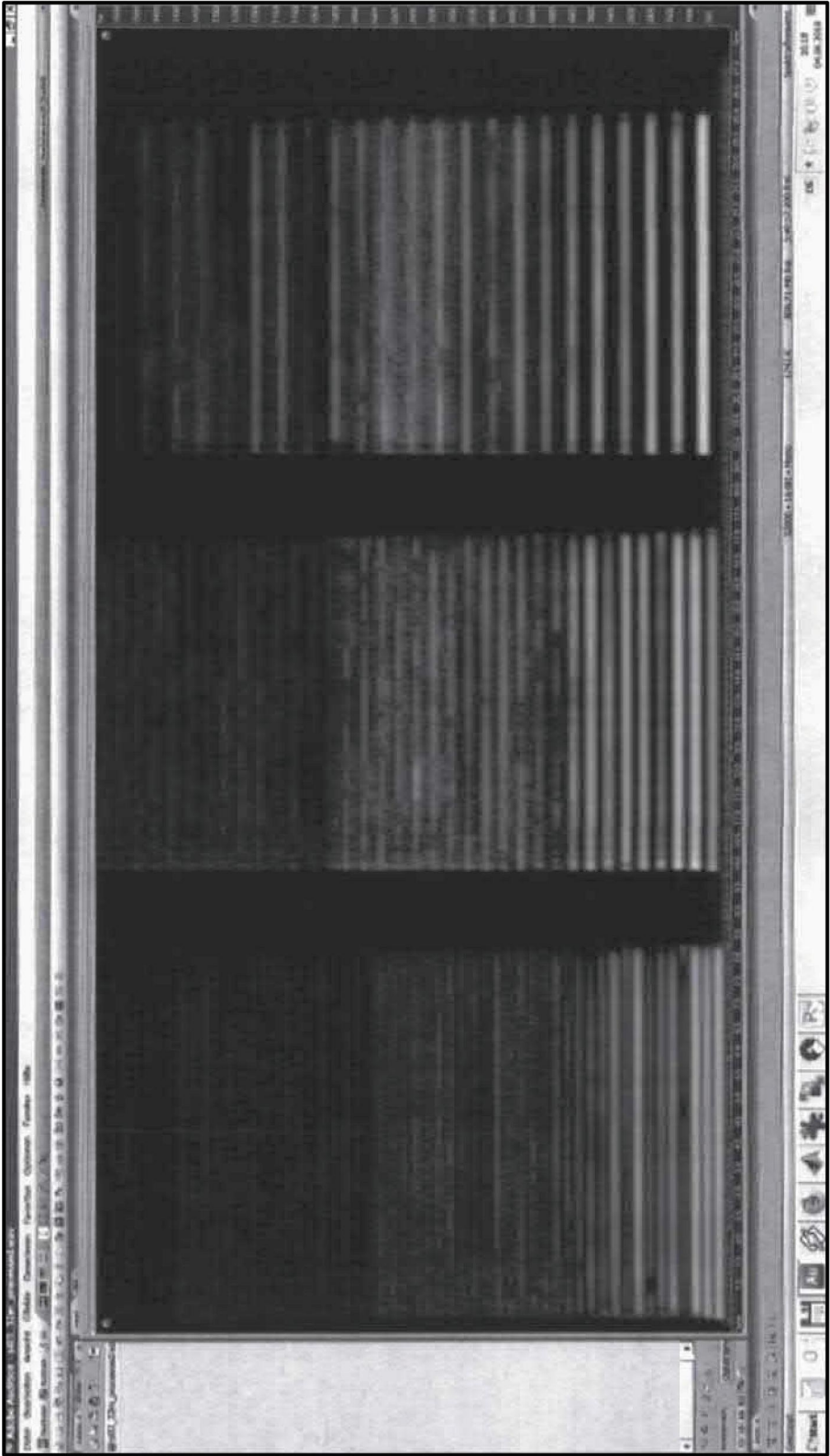
【第 9a 圖】



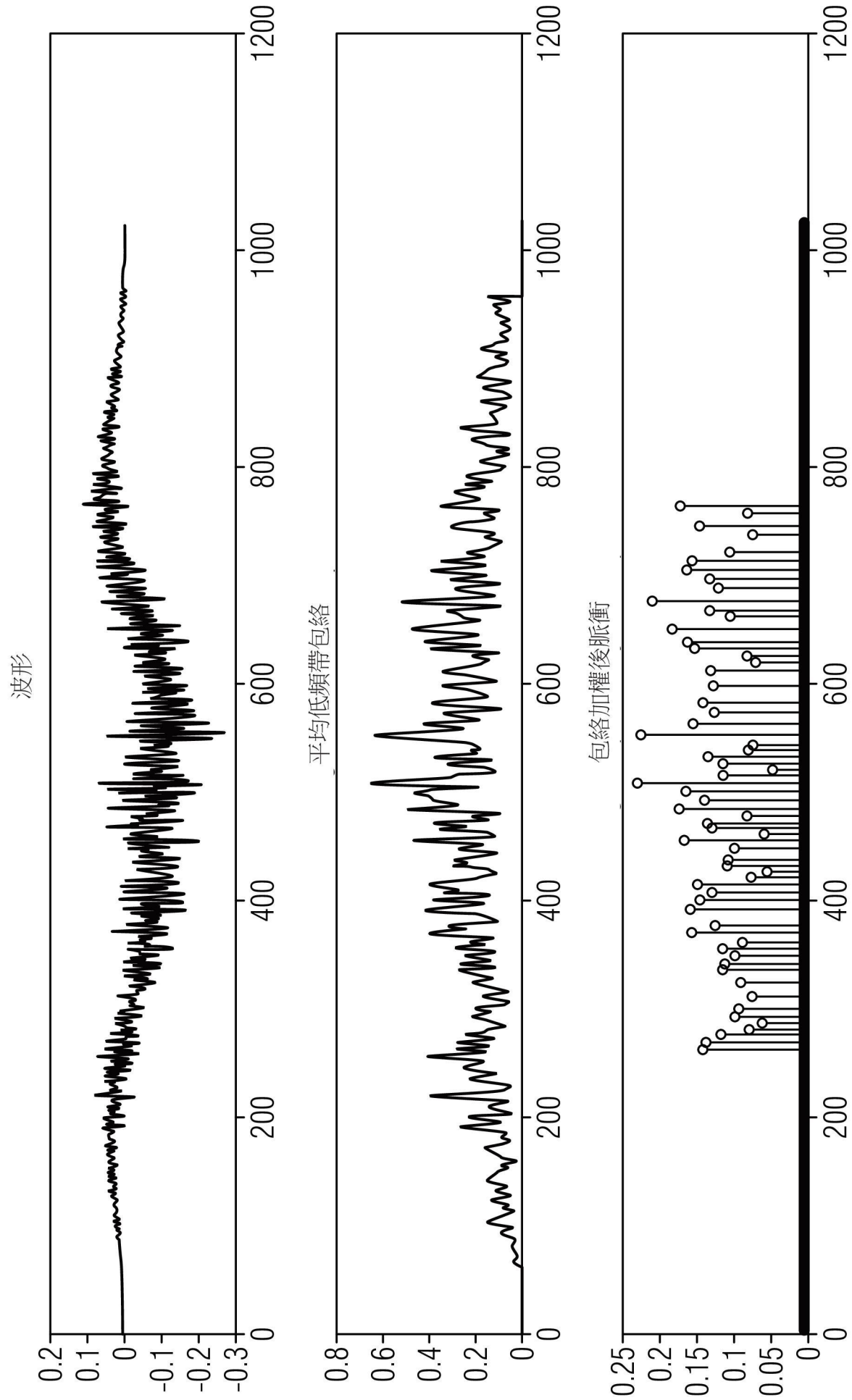
【第 9b 圖】



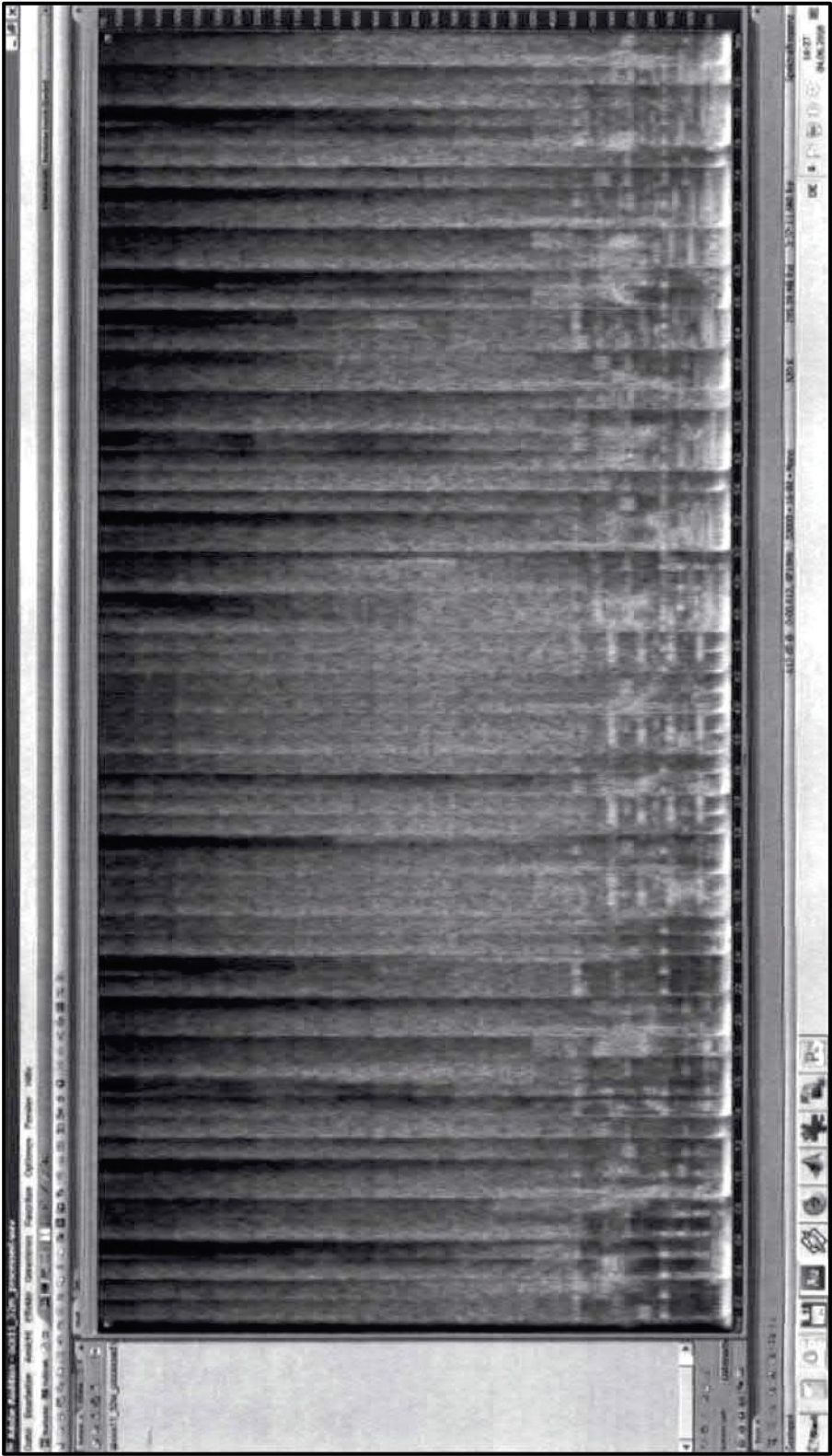
【第 10a 圖】



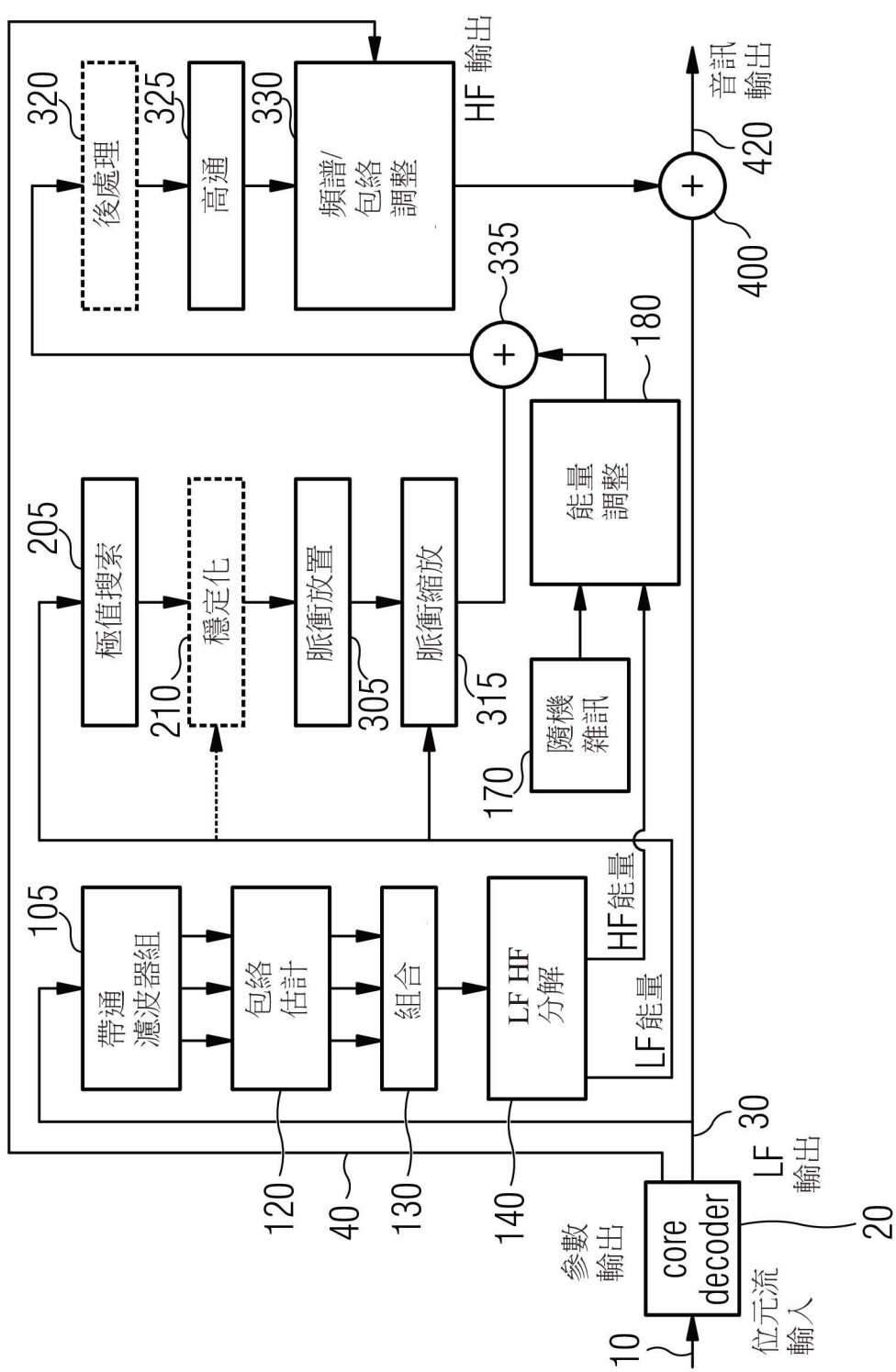
【第 10b 圖】



【第 11a 圖】



【第 11b 圖】



【第12圖】