I702824

(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 I702824 B**

(45)公告日： 中華民國 109 (2020) 年 08 月 21 日

(21)申請案號：107135727　(22)申請日： 中華民國 107 (2018) 年 10 月 11 日

(51)Int. Cl.：　*H04N19/105 (2014.01)*　*H04N19/176 (2014.01)*

(30)優先權：2017/10/13　歐洲專利局　17196402.6

(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：帕法佛　強納森 PFAFF, JONATHAN (DE)；海勒　菲利浦 HELLE, PHILIPP (DE)；瑪尼里　多米尼克 MANIRY, DOMINIQUE (DE)；韋根　湯瑪士 WIEGAND, THOMAS (DE)；沙梅克　沃伊切赫 SAMEK, WOJCIECH (DE)；卡爾騰施塔德勒　史帝芬 KALTENSTADLER, STEPHAN (DE)；施瓦茨　海科 SCHWARZ, HEIKO (DE)；馬爾皮　德特勒弗 MARPE, DETLEV (DE)；錫克曼　米斯嘉 SIEKMANN, MISCHA (DE)；偉肯　馬丁 WINKEN, MARTIN (DE)

(74)代理人：邱珍元

(56)參考文獻：

EP　2063644A2　　WO　2016/199330A1

審查人員：賴韻曲

申請專利範圍項數：28 項　　圖式數：10　　共 68 頁

(54)名稱

用於逐塊圖片編碼的幀內預測模式概念

(57)摘要

根據一第一觀點，經由使一逐塊圖片編解碼器支持幀內預測模式的一集合來實現一改進的壓縮效率，據其，一圖片的一當前區塊的該幀內預測信號經由將當前區塊的相鄰樣本的一集合應用到一神經網路上來決定。本申請的一第二觀點是，附加地或替代地，基於神經網路的幀內預測模式的該開銷，經由應用相鄰樣本的一集合在其上的一專用神經網路的該使用，來決定幀內預測模式的集合的每一個的一秩或一機率值，可以使模式選擇更有效率，秩或機率值用於從多個幀內預測模式中選擇一個幀內預測模式，該多個幀內預測模式包括或者與幀內預測模式的該集合一致。

In accordance with a first aspect, an improved compression efficiency is achieved by letting a block-wise picture codec support a set of intra-prediction modes according to which the intra-prediction signal for a current block of a picture is determined by applying a set of neighboring samples of the current block onto a neural network. A second aspect of the present application is that, additionally or alternatively to the spending of neural network-based intra-prediction modes, the mode selection may be rendered more effective by the usage of a neural network dedicated to determine a rank or a probability value for each of the set of intra-prediction modes by applying a set of neighboring samples thereonto with the rank or probability value being used for the selection of one intra-prediction mode out of the plurality of intra-prediction modes including or coinciding with the set of intra-prediction modes.

指定代表圖：

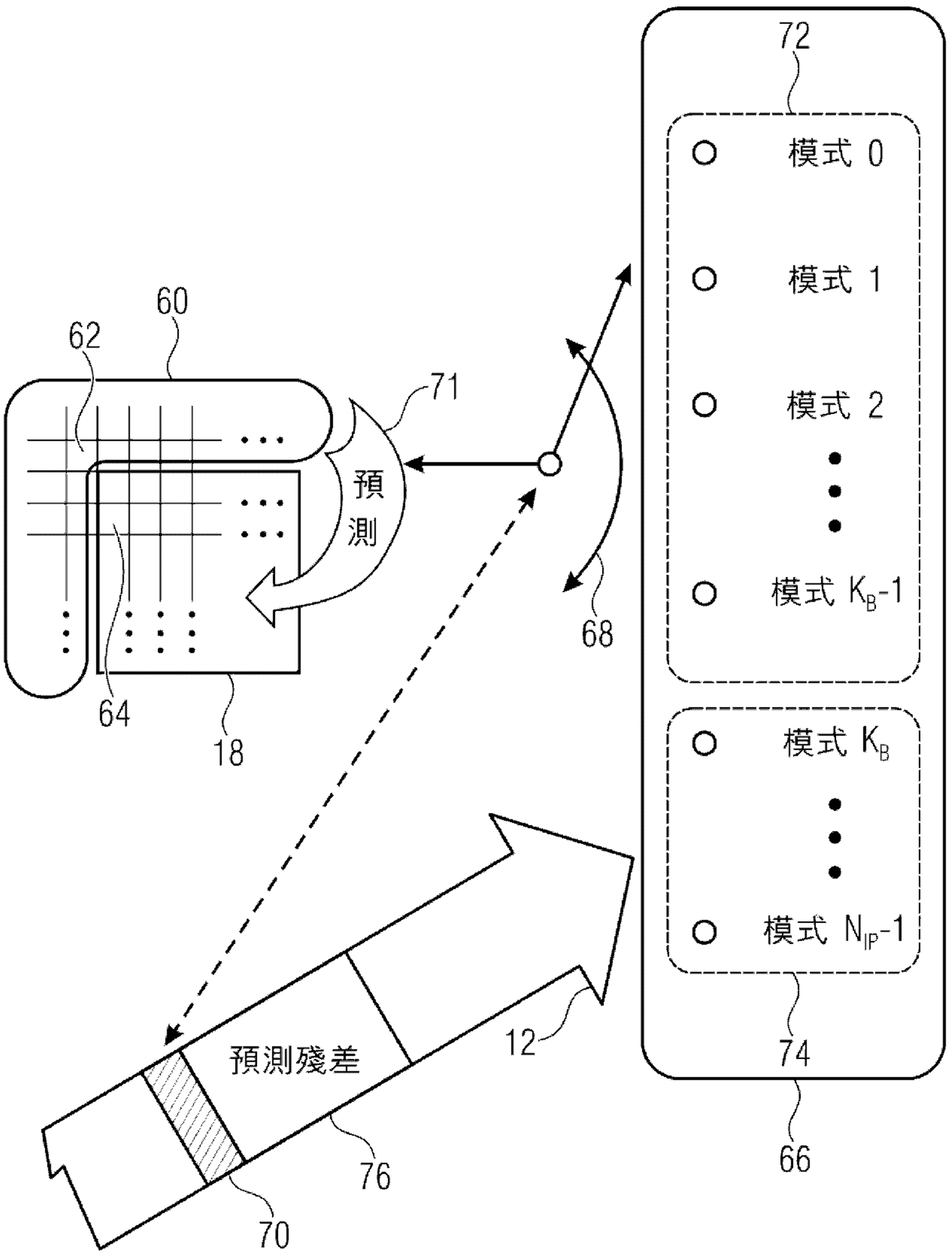

圖 5

符號簡單說明：

12・・・資料流

18・・・區塊

60・・・集合

62・・・相鄰樣本

64・・・樣本

66・・・多個預測模式

68・・・選擇

70・・・輔助信息信號通知

71・・・預測

72・・・集合

74・・・集合

74・・・模式

76・・・部分

專利案號：107135727

公告本

I702824

【發明摘要】

【中文發明名稱】用於逐塊圖片編碼的幀內預測模式概念

【英文發明名稱】INTRA-PREDICTION MODE CONCEPT FOR BLOCK-WISE PICTURE CODING

【中文】

根據一第一觀點，經由使一逐塊圖片編解碼器支持幀內預測模式的一集合來實現一改進的壓縮效率，據其，一圖片的一當前區塊的該幀內預測信號經由將當前區塊的相鄰樣本的一集合應用到一神經網路上來決定。本申請的一第二觀點是，附加地或替代地，基於神經網路的幀內預測模式的該開銷，經由應用相鄰樣本的一集合在其上的一專用神經網路的該使用，來決定幀內預測模式的集合的每一個的一秩或一機率值，可以使模式選擇更有效率，秩或機率值用於從多個幀內預測模式中選擇一個幀內預測模式，該多個幀內預測模式包括或者與幀內預測模式的該集合一致。

【英文】

In accordance with a first aspect, an improved compression efficiency is achieved by letting a block-wise picture codec support a set of intra-prediction modes according to which the intra-prediction signal for a current block of a picture is determined by applying a set of neighboring samples of the current block onto a neural network. A second aspect of the present application is that, additionally or alternatively to the spending of neural network-based intra-prediction modes, the mode selection may be

rendered more effective by the usage of a neural network dedicated to determine a rank or a probability value for each of the set of intra-prediction modes by applying a set of neighboring samples thereonto with the rank or probability value being used for the selection of one intra-prediction mode out of the plurality of intra-prediction modes including or coinciding with the set of intra-prediction modes.

【指定代表圖】圖5

【代表圖之符號簡單說明】

12：資料流

18：區塊

60：集合

62：相鄰樣本

64：樣本

66：多個預測模式

68：選擇

70：輔助信息信號通知

71：預測

72：集合

74：集合

74：模式

76：部分

【特徵化學式】無

【發明說明書】

【中文發明名稱】用於逐塊圖片編碼的幀內預測模式概念

【英文發明名稱】INTRA-PREDICTION MODE CONCEPT FOR BLOCK-WISE PICTURE CODING

【技術領域】

【0001】 本申請涉及用於逐塊圖片編碼的一改進幀內預測模式概念，例如可用於一視頻編解碼器，諸如 HEVC 或 HEVC 的任何後續解碼器。

【先前技術】

【0002】 幀內預測模式廣泛被用於圖像和視頻編碼。在視頻編碼中，幀內預測模式與諸如幀間預測模式(諸如運動補償預測模式)的其他預測模式競爭。在幀內預測模式中，基於相鄰樣本來預測一當前區塊，亦即，該相鄰樣本就編碼器側而言為已經編碼的樣本，並且就解碼器側而言為已經被解碼。坦率地說，相鄰樣本值被推斷到該當前區塊中，以便形成該當前區塊的一預測信號，針對該當前區塊，其中該預測殘差在該資料流中被發送。該預測信號越好，則該預測殘差越低，因此，編碼該預測殘差需要一更少的位元數目。

【0003】 為了效率，應該考慮若干觀點，以便在逐塊圖像編碼環境中形成用於幀內預測的一有效幀工作。例如，該編解碼器支持的幀內預測模式的數量越大，該輔助信息速率消耗越大，以便將該選擇以信號通

知該解碼器。另一方面，所支持的幀內預測模式的該集合應該能夠提供一良好的預測信號，亦即可產生一低預測殘差的一預測信號。

【發明內容】

【0004】 本申請試圖提供一幀內預測模式概念，如果使用該改進的幀內預測模式概念，則可允許一逐塊圖片編解碼器更有效地壓縮。

【0005】 該目的經由本申請的獨立請求項的主題實現。

【0006】 本揭露的較佳實施例係隨後參照附圖描述。

【0007】 根據本申請的一第一觀點，經由使一逐塊圖片編解碼器支持幀內預測模式的一集合，一改進的壓縮效率可被實現，據其，一圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一集合應用到一神經網路上來決定。該集合可以形成完全支持的多個幀內預測模式或僅形成其適當的子集合。在該所支持的多個幀內預測模式中，一個被選擇用於該當前區塊，並且使用該一個幀內預測模式(即，所選擇的一個)來預測該當前區塊。可以向該資料流提供該當前區塊的一索引，其指示該所選擇的幀內預測模式。提供具有多於一個基於神經網路的幀內預測模式的一逐塊圖片編解碼器，在所述幀內預測模式之間對一當前區塊執行一選擇，其能夠設計這些基於神經網路的幀內預測模式，以便為具有一增加的似然性的一特定區塊找到低預測誤差的一預測信號。另外，可以以如下一方式執行該神經網路幀內預測模式設計：如果用於幀內預測模式選擇的輔助信息係被耗費，則可以保持該輔助信息開銷低。這經由該幀內預測模式設計可以用幀內預測圖像區塊之間的相互不同的選擇頻率的方式，而自由地生成該基於神經網路的幀內預測模式的事實而變得可行。換句話說，花費多個基於神經網路的幀內預測模式

使得能夠以這樣的方式設計這些模式，使得它們的選擇的統計頻率適合於該選擇信號通知，從而與在啟發式設計的幀內預測模式集合中進行選擇相比，導致一熵編碼意義下低的輔助信息開銷。例如，指向所選擇的幀內預測模式的一索引可以使用可變長度代碼或使用熵編碼來編碼，並且基於神經網路的幀內預測模式可以以這樣的一方式來設計，使得它們的選擇頻率或頻率分佈是適應於該可變長度碼，或者它們的頻率分佈適合於熵編碼的基礎樣本統計，使得該幀內預測模式中的該模式選擇的該平均信號通知開銷減小或最小化。

【0008】 本申請的一第二觀點是，附加地或替代地，基於神經網路的幀內預測模式的該開銷，經由應用相鄰樣本的一集合在其上的一專用神經網路的該使用，來決定幀內預測模式的該集合的每一個的一秩或一機率值，可以使該模式選擇更有效率，該秩或機率值用於從該多個幀內預測模式中選擇一個幀內預測模式，該多個幀內預測模式包括幀內預測模式的該集合或者與幀內預測模式的該集合一致。用於在該幀內預測模式中進行選擇的該輔助信息開銷可能可以完全沒有留存，或者可以經由使用該神經網路使其更有效率。

【0009】 就該上述神經網路的該設計而言，本申請提供了許多用於適當地決定其參數的實施例。

【0010】 本揭露主題的各種目的、特徵、面向和優點將從以下對優選實施例的詳細描述以及附圖中變得更加明顯，附圖中相同的附圖標記表示相同的部件。

【0011】 所示實施例在附圖中以例子，而非限制的方式顯示，其中相同的參考標號表示相似的元件。

【圖式簡單說明】

【0012】

圖1係顯示說明用於將一圖片編碼成一資料流的一編碼器的一示意性方塊圖，其作為可以實現本申請的實施例的一個一般範例。

圖2係顯示根據圖1的一編碼器的一更具體範例的一方塊圖。

圖3係顯示說明適合於圖1的該編碼器並且用作一解碼器的一範例的一解碼器的一示意性方塊圖，其中本申請的實施例可以被實現。

圖4係顯示圖3的一解碼器的的一更具體範例的一方塊圖，其適合於圖2的該編碼器。

圖5係顯示說明根據本申請的實施例的編碼器和解碼器中關於使用幀內預測處理一區塊的該操作模式的一示意圖。

圖6係顯示說明根據本申請的一實施例的一解碼器的一示意性方塊圖，該解碼器包括若干基於神經網路的幀內預測模式。

圖7a係顯示說明根據支持基於神經網路的幀內預測模式的一實施例的一編碼器和一解碼器的操作模式以及這些模式的基於神經網路的排序的一示意圖，其中在該資料流內傳輸一索引到基於神經網路的幀內預測模式的一有序列表，以及指示要使用的幀內預測模式是否是基於神經網路的幀內預測模式的該集合的成員的一斑點。不用說，可以使用可變長度編碼對該索引進行編碼，以便利用由決定90所決定的不同頻率。

圖7b係顯示與圖7a不同的一示意圖，其中未使用該斑點信號（fleck signalization）。

圖7c係顯示與圖7b不同的一示意圖，其中不使用一神經網路來控制該模式排序。

圖7d係顯示與7a不同的一示意圖，其中模式信號化中的該神經網路輔助用於控制在熵編/解碼中的該機率分佈估計而不是模式排序。

圖8係顯示根據一實施例的用於設計基於神經網路的幀內預測模式的一集合之一裝置。

圖9a係顯示說明根據一個實施例的編碼器和解碼器的該操作模式的一示意圖，根據該實施例，一神經網路用於對所支持的幀內預測模式進行排序，而不管是否基於神經網路。

圖9b係顯示與圖9a不同的一示意圖，其中該神經網路基礎用於控制用於該索引的熵解碼/編碼到所支持的幀內預測模式該集合的該機率分佈估計。

圖10係顯示根據一實施例的用於設計一神經網路的一裝置，該神經網路用於輔助和選擇用於基於區塊的圖像編碼的幀內預測模式的一集合。

【實施方式】

【0013】 在下文中，描述了各種實施例，其有助於在使用幀內預測時實現更有效的壓縮。一些實施例經由花費基於神經網路的幀內預測模式的一集合來實現該壓縮效率的提高。後者可以被添加到例如啟發式設計的其他幀內預測模式，或者可以被專門提供。其他實施例使用一神經網路以在多個幀內預測模式中執行一選擇。甚至其他實施例也利用了剛才討論的兩個特徵。

【0014】 為了容易理解本申請的以下實施例，該描述從可能的編碼器和解碼器的一陳述開始，其中可適合於構建本申請的隨後概述的實施例。圖 1 係顯示用於將一圖片 10 逐塊編碼成一資料流 12 的一裝置。該

裝置使用附圖標記 14 表示，並且可以是一靜止圖片編碼器或一視頻編碼器。換句話說，當該編碼器 14 被配置為將包括圖片 10 的一視頻 16 編碼到資料流 12 中時，圖片 10 可以是該視頻 16 中的一當前圖片，或者編碼器 14 可以將圖片 10 專門編碼到資料流 12 中。

【0015】 如上所述，編碼器 14 以逐塊方式或以區塊為基礎執行編碼。為此，編碼器 14 將圖像 10 細分為區塊，其為編碼器 14 將圖像 10 編碼成資料流 12 的單元。圖像 10 細分成區塊 18 的可能範例在下面更詳細地闡述。通常，該細分可以最終成為具有恆定大小的區塊 18，其係以行和列排列的區塊陣列表示，或不同區塊大小的區塊 18，例如經由使用一分層多樹細分來啟動一多樹細分，其係從圖片 10 的該整個圖片區域或從圖片 10 的一預分區域進行細分而到一樹區塊陣列，其中這些範例不應被視為排除將圖片 10 細分為區塊 18 的其他可能方式。

【0016】 此外，編碼器 14 是一預測編碼器，其被配置為將圖片 10 預測編碼至資料流 12。對於一某個區塊 18，這意味著編碼器 14 決定區塊 18 的一預測信號，並將該預測殘差(亦即該預測信號偏離區塊 18 內的該實際圖像內容的該預測誤差)編碼到資料流 12 中。

【0017】 編碼器 14 可以支持不同的預測模式，以便導出一特定區塊 18 的該預測信號。在以下實施例中重要的該預測模式是幀內預測模式，根據該幀內預測模式，區塊 18 的該內部係從圖像 10 中在空間上相鄰的已編碼樣本而被預測。圖片 10 到資料流 12 的該編碼以及相應的解碼過程可以基於在區塊 18 之間定義的一特定編碼順序 20。例如，該編碼順序 20 可以以一光柵掃描順序(a raster scan order)遍歷區塊 18，

例如從頂部到底部逐列掃描，並且例如從左到右遍歷每一列。在基於分層多樹的細分的情況下，可以在每一個層級別內應用光柵掃描順序，其中可以應用一深度優先遍歷順序，亦即，在一特定層級別的區塊內的葉子筆記根據編碼順序 20 可以在具有相同父區塊的級別之相同層級別的區塊之前。取決於該編碼順序 20，一區塊 18 的相鄰的已編碼樣本可以通常位於區塊 18 的一側或多側。在這裡給出的例子的情況下，例如，一區塊 18 的相鄰的已編碼樣本位於區塊 18 的該頂部和該左側。

【0018】 幀內預測模式可能不是編碼器 14 所支持的唯一模式。在編碼器 14 是一視頻編碼器的情況下，例如，編碼器 14 還可以支持幀內預測模式，根據該幀內預測模式，從一先前編碼的視頻圖像 16 暫時地預測一區塊 18。這一種幀內預測模式可以是一運動補償預測模式，根據該運動補償預測模式，對於這樣的一區塊 18 用信號通知一運動向量，指示該部分的一相對空間偏移，由該部分可從中導出區塊 18 的該預測信號作為一副本。附加地或替代地，其他非幀內預測模式也可以是可用的，例如在編碼器 14 是多視圖編碼器的情況下的視圖間預測模式(inter-view prediction modes)，或者非預測模式，根據該非預測模式，區塊 18 的該內部按原樣編碼，即沒有任何預測。

【0019】 在開始將本申請的描述聚焦於幀內預測模式之前，針對一可能的基於區塊的編碼器的一更具體的實施例，即針對編碼器 14 的一可能實現係相關於圖 2 所描述的，然後呈現一解碼器的兩個對應的實施例，分別用於符合圖 1 和 2 的。

【0020】 圖 2 係顯示圖 1 的編碼器 14 的一可能實現，即該編碼器

被配置為使用變換編碼來編碼該預定殘差的實現，儘管這幾乎是一範例，並且本申請不限於那種預測殘差編碼。根據圖 2，編碼器 14 包括一減法器 22，其被配置為從該入站信號(即圖片 10、或者在區塊的基礎上係為當前區塊 18)中減去該相應的預測信號 24，以便獲得該預測殘差信號 26，然後由一預測殘差編碼器 28 編碼到一資料流 12 中。該預測殘差編碼器 28 由一有損編碼級 28a 和一無損編碼級 28b 組成。該有損編碼級 28a 接收該預測殘差信號 26 並且包括一量化器 30，該量化器 30 量化該預測殘差信號 26 的樣本。如上所述，本範例使用該預測殘差信號 26 的變換編碼，因此，該有損編碼級 28a 包括連接在減法器 22 和量化器 30 之間的一變換級 32，以便將一這種頻譜分解的預測殘差 26 變換為代表該殘差信號 26 的該變換係數，量化器 30 的一量化發生在該變換係數上。該變換可以是 DCT、DST、FFT、Hadamard 變換等。然後，經過變換和量化的該預測殘差信號 34 通過該無損編碼級 28b 進行無損編碼，該無損編碼級 28b 是一個熵編碼器，對量化預測殘差信號 34 進行熵編碼，而為資料流 12。編碼器 14 還包括連接到量化器 30 的該輸出的該預測殘差信號重建級 36，以便從該變換和量化的預測殘差信號 34 重建該預測殘差信號，其也以解碼器可用的方式進行，亦即，考慮量化器 30 該編碼損失。至此，該預測殘差重建級 36 包括一去量化器 38，其執行該量化器 30 的該逆量化，接著是一個逆變換器 40，其執行相對於該變換器 32 執行的該變換的該逆變換，例如，諸如與任何上述特定變換例子相反的逆頻譜分解。編碼器 14 包括一加法器 42，其將逆變換器 40 輸出的該重建的預測殘差信號和該預測信號 24 相加，以

便輸出一重建信號，即重建例子。該輸出被饋送到編碼器 14 的一預測器 44，然後編碼器 14 基於此預測器決定該預測信號 24。該預測器 44 支持上面已經參考圖 1 討論的所有預測模式。圖 2 還顯示在編碼器 14 是一視頻編碼器的情況下，編碼器 14 還可以包括一環路濾波器 46，其具有濾波器以完全重建圖像，在經過濾波之後，相對於幀間預測區塊，形成預測器 44 的參考圖像。

【0021】 如上所述，編碼器 14 係基於區塊操作。對於該隨後的描述，感興趣的區塊基礎是將一個細分圖像 10 成區塊，對於其中，分別從預測器 44 或編碼器 14 支持的一集合或多個幀內預測模式中選擇該幀內預測模式，並且單獨地執行該所選擇的幀內預測模式。然而，圖片 10 被細分的其他種類的區塊也可以存在。例如，圖片 10 是幀間編碼還是幀內編碼的該上述決定可以以一粒度或以和區塊 18 不同的區塊為單位進行。例如，該幀間/幀內模式的決定可以在圖片 10 被細分的編碼區塊的一級別執行，並且將每個編碼區塊細分為預測區塊。針對已經確定使用幀內預測的編碼區塊的預測區塊，預測區塊的每一個被細分為一幀內預測模式決策。為此，對於這些預測區塊中的每一個，其將被決定應該將哪個支持的幀內預測模式用於該相應的預測區塊。這些預測區塊將形成這裡感興趣的區塊 18。預測器 44 將不同地處理與幀間預測相關聯的編碼區塊內的預測區塊。經由決定一運動向量並從該運動向量指向的該參考圖像中的一位置複製該區塊的該預測信號，可以從參考圖像中對它們進行幀間預測。另一個區塊細分係涉及細分為變換區塊，在變換器 32 和逆變換器 40 係以變換區塊為單元而執行變換。例如，變換區塊可

以是進一步細分編碼區塊的該結果。自然地，這裡列出的範例不應被視為限制性的，並且也存在其他範例。僅為了完整性，應注意，細分為編碼區塊可以例如使用多樹細分，並且也可以經由使用多樹細分進一步細分編碼區塊來獲得預測區塊和/或變換區塊。

【0022】 在圖 3 中描繪了適合於圖 1 的編碼器 14 的逐塊解碼的一解碼器或裝置。該解碼器 54 與編碼器 14 相反，亦即它以一區塊方式從資料流 12 解碼圖像 10，並為此支持多個幀內預測模式。以上關於圖 1 討論的所有其他可能性也對解碼器 54 有效。為此，解碼器 54 可以是一靜止圖片解碼器或一視頻解碼器，並且解碼器 54 也支持所有該預測模式和預測可能性。編碼器 14 和解碼器 54 之間的區別主要在於編碼器 14 根據某些優化建議而選擇或抉擇編碼決策，例如，以便最小化可能取決於編碼率和/或編碼失真的某些成本函數。這些編碼選項或編碼參數之一可以涉及在可用或支持的幀內預測模式中，要用於一當前區塊 18 的該幀內預測模式的一選擇。然後，編碼器 14 可以針對資料流 12 內的當區前塊 18 用信號通知該所選擇的幀內預測模式，其中解碼器 54 使用資料流 12 中的該信號通知對區塊 18 重新進行該選擇。同樣地，將圖像 10 細分為區塊 18 可以在編碼器 14 內進行優化，並且可以在資料流 12 內傳送相應的細分信息，其中解碼器 54 基於該細分信息將圖像 10 的該細分恢復為區塊 18。綜上所述，解碼器 54 可以是在一基於區塊的基礎上操作的一預測解碼器，並且除了幀內預測模式之外，解碼器 54 可以支持其他預測模式，例如在例如解碼器 54 是一視頻解碼器的情況下的幀間預測模式。在解碼中，解碼器 54 還可以使用關於圖 1 討論的

該編碼順序 20，並且當在編碼器 14 和解碼器 54 兩者都遵循該編碼順序 20 時，該相同的相鄰樣本在編碼器 14 和解碼器 54 兩者處都可用於一當前區塊 18。因此，為了避免不必要的重複，編碼器 14 的該操作模式的該描述也應當適用於解碼器 54，只要圖片 10 細分為區塊係被關注，例如，到目前為止預測係被關注，並且到目前為止預測殘差的編碼係被關注。不同之處在於編碼器 14 經由優化選擇一些編碼選項或編碼參數以及在資料流 12 內信號通知或插入資料流 12 中，然後由解碼器 54 從該資料流 12 導出該編碼參數以重新進行預測、細分等等。

【0023】 圖 4 係顯示圖 3 的該解碼器 54 的一可能實現，亦即一個適合於如圖 2 所示的圖 1 的編碼器 14 的該實現。由於圖 4 的編碼器 54 的許多元件與圖 2 的相應編碼器中出現的元件相同，所以在圖 4 中使用具有撇號的該相同附圖標記以指示這些元件。特別地，加法器 42’、可選的環路濾波器 46’和預測器 44’以與它們在圖 2 的編碼器中該相同的方式連接到一預測環路中。該重建(亦即去量化的和重新變換的預測殘差信號應用於添加的 42’)係由熵解碼器 56 的一序列導出，該熵解碼器 56 反轉熵編碼器 28b 的該熵編碼，接著是由去量化器 38’和逆變換器 40’組成的該殘差信號重建級 36’，正如編碼側的情況一樣。該解碼器的輸出是圖像 10 的該重建。圖像 10 的該重建可以直接在加法器 42’的該輸出端獲得，或者另外地在環路濾波器 46’的該輸出端獲得。一些後置濾波器可以被安排在該解碼器的輸出處，以便對圖像 10 的重建進行一些後濾波，俾改善圖像品質，但是該選項在圖 4 中沒有被示出。

【0024】 再次，關於圖 4，上面提到的與圖 2 相關的描述對於圖 4 也是有效的，除了僅該編碼器執行該優化任務和關於編碼選項的該相關決定。然而，關於區塊細分、預測、去量化和重新變換的所有描述對於圖 4 的該解碼器 54 也是有效的。

【0025】 在繼續描述本申請的可能實施例之前，關於上述例子，應作出一些註釋。儘管上面沒有明確提到，但是顯然區塊 18 可以具有任何形狀。例如，它可以是矩形或二次形狀。此外，儘管對編碼器 14 和解碼器 54 的該操作模式的該上面描述經常提到一「當前區塊」18，但是很清楚，編碼器 14 和解碼器 54 相應地對於要為其選擇一幀內預測模式的每一個區塊起作用。如上所述，也可以存在其他區塊，但是該以下描述聚焦於圖像 10 被細分的那些區塊 18，針對其，一幀內預測模式被選擇。

【0026】 為了總結要為其選擇一幀內預測模式的一特定區塊 18 的該情況，請參考圖 5。圖 5 係顯示一當前區塊 18，即當前要被編碼或被解碼的區塊。圖 5 係顯示相鄰樣本 62 的一集合 60，即具有空間上與區塊 18 相鄰的樣本 62。區塊 18 內的該樣本 64 將被預測。因此，要導出的該預測信號是針對區塊 18 內的每個樣本 64 的一預測。如上所討論，多個 66 預測模式可用於每個區塊 18，並且如果區塊 18 將被幀內預測，則該多個 66 模式僅包括幀間預測模式。在編碼器和解碼器側執行一選擇 68，以便基於該相鄰樣本集合 60 決定多個 66 中的一個幀內預測模式，以被用於預測 71 區塊 18 的該預測信號。下面進一步描述的該實施例係與該可用的幀內預測模式 66 並且與關於選擇 68 建議的該操作模

式不同，例如，是否相對於區塊 18 關於選擇 68 在該資料流 12 中設置了輔助信息。然而，這些實施例的該描述從提供數學細節的一具體描述開始。根據該初始實施例，對要被幀內預測的一特定區塊 18 的該選擇係與對應的輔助信息信號通知 70 相關聯，並且該資料流和該多個 66 幀內預測模式包括基於神經網路的幀內預測模式的一集合 72 以及啟發式設計的進一步幀內預測模式的一集合 74。模式 74 的該幀內預測模式之一可以是例如一 DC 預測模式，根據該 DC 預測模式，基於該相鄰樣本集合 60 可決定一些平均值，並且將該平均值分配給區塊 18 內的所有樣本 64。附加地或替代地，集合 74 可以包括幀間預測模式，其可以被稱為角度幀間預測模式(angular inter-prediction modes)，根據該幀間預測模式，該相鄰樣本集合 60 的樣本值沿著一特定幀內預測方向被複製到區塊 18 中，該幀內預測方向在這種角度幀內預測模式之間係不同。圖 5 係顯示除了關於多個 66 幀內預測模式中的該選擇 68 的該可選地存在的輔助信息 70 之外，該資料流 12 還包括預測殘差編碼的一個部分 76，編碼可以如上所述，可選地包括在變換域中具有量化的變換編碼。

【0027】 具體地，為了容易理解本申請的一特定實施例的該以下描述，圖 6 係顯示編碼器和解碼器處的一幀內預測區塊的一般操作模式。圖 6 係顯示區塊 18 以及基於執行該幀內預測的該相鄰樣本集合 60。應當注意，該集合 60 可以在基數方面而在該多個 66 幀內預測模式的該幀內預測模式之間變化，亦即，根據相應幀內預測模式的實際使用的集合 60 的該樣本數量，以用於決定區塊 18 的預測信號。然而，這是為了

容易理解，未在圖 6 中顯示。圖 6 係顯示編碼器和解碼器針對集合 72 的每個基於神經網路的幀內預測模式具有 80_0 到 80_{KB-1} 的神經網路。集合 60 被應用於該相應的神經網路，以便在集合 72 中導出該相應的幀內預測模式。除此之外，圖 6 相當代表性地顯示一個區塊 82，其基於輸入提供亦即相鄰樣本的該集合 60，以及集合 74 的一個或多個幀內預測模式的一個或多個預測信號，例如，該 DC 模式預測信號和/或角度幀內預測模式預測信號。以下描述將顯示如何有利地決定具該神經網路 80i 的該參數，其中 $i=0、\ldots、K_B-1$。在下文中闡述的該特定實施例還向編碼器和解碼器提供另一神經網路 84，其基於相鄰樣本的一集合 86，俾為集合 72 內的每個基於神經網路的幀內預測模式專門提供一機率值，相鄰樣本的該集合 86 可以與集合 60 重合或不重合。該神經網路 84 幫助使用於該模式選擇的該輔助信息 70 更有效時，因此該機率值被提供。例如，在下面描述的實施例中，假設一可變長度代碼用於指向該幀內預測模式之一，並且至少就集合 72 而言，由該神經網路 84 提供的該機率值致能可使用輔助信息 70 內的該可變長度代碼，作為排序的幀內預測模式的一有序列表的一索引，其係根據神經網路 84 輸出的該機率值，而針對集合 72 內的基於神經網路的幀內預測模式，從而優化或降低該輔助信息 70 的該碼率。對此，如圖 6 所示，該模式選擇 68 有效地執行，其係取決於由該另外的神經網路 84 提供的該機率值以及資料流 12 內的該輔助信息 70。

【0028】 1.訓練執行幀內預測的神經網路的該參數的演算法

【0029】 令$B \subset \mathbb{Z}^2$為一視頻幀的一區塊，即區塊 18。假設B具有M個像素。對於一固定顏色分量，讓im成為在B上一視頻信號的該內容。我們將im視為$\mathbb{R}^M$的一個元素。假設存在B的一個鄰域$B_{rec} \subset \mathbb{Z}^2$，其具有$L$個像素並且其上已經重建的圖像$rec \in \mathbb{R}^L$是可用的，即樣本集合 60 和 86，儘管它們可以替代地不同。經由一幀內預測函數，我們指的是函數F: $\mathbb{R}^L \to \mathbb{R}^M$。我們將$F(rec)$視為$im$的一預測因子。

【0030】 接下來描述的是經由一資料驅動的優化方法，設計可以在一典型的混合視頻編碼標準(即集合 72)中出現的若干區塊B的幀內預測函數的一演算法。為了實現這一目標，我們考慮了以下主要設計特徵：

1.在我們進行的該優化演算法中，我們希望使用該成本函數的一個良好近似，特別是涉及可以預期用於發信號通知該預測殘差的該位元數目。

2.我們希望聯合訓練數個幀內預測，以便能夠處理不同的信號特徵。

3.當訓練幀內預測時，必須考慮發信號通知使用哪種幀內模式所需的該位元數目。

4.我們希望保留已定義的幀內預測的一集合，例如該 HEVC 幀內預測，並訓練我們的預測，作為補充預測。

5.一典型的混合視頻編碼標準通常支持幾個區塊形狀，該給定區塊 B 可以被分成幾個區塊形狀。

【0031】 在接下來的四個部分中，有可能描述一個人如何處理這些要求。更準確地說，在 1.1 節中，我們將描述如何處理第一個項目。在

第 1.2 節中，描述瞭如何處理第 2 至第 3 項目。在 1.4 節中，描述瞭如何考慮項目 4。最後，在第 1.5 節中，描述瞭如何處理最後一項目。

【0032】 **1.1 訓練近似於一視頻編解碼器的該速率函數的一損失函數的演算法**

【0033】 用於決定在一視頻編解碼器中使用的未知參數的一資料驅動方法通常被設置為一優化演算法，該演算法試圖最小化在一給定訓練樣例集合上的一預定義損失函數。通常，對於在實踐中工作的數值優化演算法，該後一種損失函數應滿足一些平滑度要求。

【0034】 另一方面，像 HEVC 這樣的視頻編碼器在經由最小化該速率-失真成本(Rate-Distortion costs) $D+\lambda\cdot R$來做出決定時表現最佳。此處，D是該解碼視頻信號的該重建誤差，R是該速率，即編碼該視頻信號所需的該位元數目。此外，$\lambda\in\mathbb{R}$是一個拉格朗日參數(Lagrangian Parameter)，其取決於該所選擇的量化參數。

【0035】 該真正的函數$D+\lambda\cdot R$通常非常複雜，並且不能經由一閉合表達式給出，而可以提供一資料驅動的優化演算法。因此，我們經由一分段平滑函數逼近該整個函數$D+\lambda\cdot R$或至少近似該速率函數R。

【0036】 更確切地說，如前所述，讓B為一視頻幀 10 的一給定區塊1/，並且設im是一個固定顏色分量在B上的該相應視頻信號。假設B具有M個像素。然後，對於一預測候選者$pred\in\mathbb{R}^M$，我們考慮該預測殘差$res:=(im-pred)\in\mathbb{R}^M$。對於一給定的量化參數(Quantization Parameter)和一給定的變換，令$R(res)$是一真實視頻編碼器用信號通知

res的該量化變換所需的該速率。此外，令 $D(res)$為res的逆量化和逆變換產生的該重建誤差。然後我們想要決定函數$\mathrm{H}, \tilde{\mathrm{R}}: \mathbb{R}^M \to \mathbb{R}$，它們是分段平滑的，使得$H(res)$用作為$D(res) + \lambda \cdot R(res)$的一良好近似並且使得$\tilde{R}(res)$用作$R(res)$的一良好近似值。

【0037】 我們固定了一些$N \in \mathbb{N}$並固定了預定義的“架構”，即分段平滑函數

$$\mathcal{H}: \mathbb{R}^M \times \mathbb{R}^N \to \mathbb{R}, \quad \mathcal{R}: \mathbb{R}^M \times \mathbb{R}^N \to \mathbb{R}$$

然後尋找$\Phi_1, \Phi_2 \in \mathbb{R}^{\mathrm{N}}$，這樣我們就將我們的函數H和$\tilde{\mathrm{R}}$建模為

$$H(res) = \mathcal{H}(res, \Phi_1), \quad \tilde{R}(res) = \mathcal{R}(res, \Phi_2)。$$

【0038】 為了決定該權重Φ_1和Φ_2，在使用該給定混合視頻編碼標準的一典型編碼器上，我們收集了預測殘差res_j、$j \in \mathcal{J}$和該相應的速率-失真值$(D + \lambda R)(res_j)$的訓練樣例的一巨大集合，分別僅為某些有限大的索引集合$\mathcal{J}$的速率值$R(res_j)$。然後我們嘗試找到Φ_1和Φ_2，它們可使下面表達式最小化或至少使表達式變小

$$\sum_{j \in \mathcal{J}} ((D + \lambda R)(res_j) - \mathcal{H}(res_j, \Phi_1)(res_j))^2, \sum_{j \in \mathcal{J}} (R(res_j) - \mathcal{R}(res, \Phi_2)(res_j))^2。$$

【0039】 對於該任務，我們通常使用一(隨機)梯度下降方法。

【0040】 **1.2 一固定區塊形狀的預測訓練**

【0041】 在本節中，我們描述了我們設置的演算法，用於為一給定的區塊$B18$、集合 72 的幀內預測和已經重建的樣本的區域B_{rec} 60 設計K_B幀內預測。

【0042】 我們假設我們給出了預測的一預定義“架構”。我們的意思是，對於某些固定的$T \in \mathbb{N}$，我們給出了一個函數

$$F^B: \mathbb{R}^L \times \mathbb{R}^T \to \mathbb{R}^M \quad (1)$$

並且我們想要決定“權重” $\Theta_{B,1}, \ldots, \Theta_{B,K_B} \in \mathbb{R}^T$，使得我們的幀內預測給出為

$$F^B_{\Theta_{B,k}}: \mathbb{R}^L \to \mathbb{R}^M$$

其中我們把$rec \in \mathbb{R}^L$放在

$$F^B_{\Theta_{B,k}}(rec) := F^B(rec, \Theta_{B,k})。 \quad (2)$$

【0043】 以下部分提供了這方面的詳細信息。公式(2)中的函數定義了圖 6 中該神經網路 80_0-80_{KB} -1。

【0044】 接下來，我們經由使用一第二參數相關函數來模擬我們嘗試設計的該幀內模式的該信號化成本

$$G^B: \mathbb{R}^L \times \mathbb{R}^T \to \mathbb{R}^{K_B}。 \quad (3)$$

【0045】 再次，對於$\Psi_B \in \mathbb{R}^L$，我們定義

$$G^B_{\Psi_B}: \mathbb{R}^L \to \mathbb{R}^{K_B}$$

經由

$$G^B_{\Psi}(rec) := G^B(rec, \Psi_B)。 \quad (4)$$

【0046】 再次，在 1.3 節中給出了一個範例，其中公式(4)的函數表示圖 6 的神經網路 84。

【0047】 我們假設我們有一個函數

$$\mathcal{M}^B: \mathbb{R}^{K_B} \to \mathbb{R}^{K_B}。$$

【0048】 例如，該函數定義了用於輔助信息 70 的一個 VLC 代碼長度分佈。即，由輔助信息 70 與更多集合 72 的電腦輔助設計(cad)考量相關聯的代碼長度。

【0049】 然後我們定義

$$\mathcal{M}^{B}_{\Psi_B}:\mathbb{R}^{L}\rightarrow\mathbb{R}^{K_B}$$

經由

$$\mathcal{M}^{B}_{\Psi_B}(rec):=\mathcal{M}^{B}(G^{B}_{\Psi_B}(rec))。$$

【0050】 目前，該第 k 個元件

$$\left(\mathcal{M}^{B}_{\Psi_B}(rec)\right)_{k}\in\mathbb{R}$$

$\mathcal{M}^{B}_{\Psi_B}(rec)$應當模擬我們訓練的第 k 個幀內模式所需的該位元數目。

【0051】 如果$\tilde{R}$是 2.1 節中定義的該函數，則對於給定的$\Theta^{B}_{1},\ldots,\Theta^{B}_{K_B},\Psi^{B}\in\mathbb{R}^{T}$、重建圖像$rec\in\mathbb{R}^{L}$和原始圖像$im\in\mathbb{R}^{M}$，我們讓

$$k^{B}_{opt}(im,rec,\Theta^{B}_{1},\ldots,\Theta^{B}_{K_B},\Psi^{B})\in\{1,\ldots,K_B\}$$

表示該最小$k\in\{1,\ldots,K_B\}$，其具有該屬性

$$\tilde{R}(im-F^{B}_{\Theta_{B,k}}(rec))+\left(\mathcal{M}^{B}_{\Psi_B}(rec)\right)_{k}\leq\tilde{R}(im-F^{B}_{\Theta_{B,l}}(rec))+\left(\mathcal{M}^{B}_{\Psi_B}(rec)\right)_{l}$$

對於所有$l\in\{1,\ldots,K_B\}$。

【0052】 由於$\mathcal{M}$模擬一幀內模式的該信號通知的位元的該真實數目，因此其梯度為零或未定義。因此，$\mathcal{M}$都不足以經由基於一梯度下降的演算法來優化該權重Ψ_B。因此，我們還經由使用該 softmax 函數將該函數$G^{B}_{\Psi_B}$變換為一機率分佈來調用一幀內模式的該交叉熵。我們回顧該後一個函數的該定義。對於$x\in\mathbb{R}^{T}$，令x_i表示 x的第 i 個分量。然後將該 softmax 函數$\sigma:\mathbb{R}^{K_B}\rightarrow(0,1)^{K_B}$定義為

$$(\sigma(x))_k := \frac{e^{x_k}}{\sum_{l=1}^{K_B} e^{x_l}} \text{。}$$

【0053】 針對梯度更新，我們將嘗試最小化該殘差率以及該模式k_{opt}的該交叉熵相對於該後者機率分佈的該總和。因此，我們將該區塊B的損失函數Loss^B定義為

$$Loss^B(im, rec, \Theta^B_1, \ldots, \Theta^B_{k_B}, \Psi^B)$$
$$:= \tilde{R}(im - F^B_{\Theta^B_{k^B_{opt}}}(rec)) - \log_2\left(\left(\sigma(G^B_{\Psi_B}(rec))\right)_{k^B_{opt}}\right), \qquad (5)$$

其中

$$k^B_{opt} = k^B_{opt}(im, rec, \Theta^B_1, \ldots, \Theta^B_{K_B}, \Psi^B) \text{。}$$

【0054】 給定公式(5)中的該損失函數，我們經由一資料驅動的優化以決定該權重$\Theta^B_1, \ldots, \Theta^B_{K_B}$。因此，如果對於一有限的大索引集合$\mathcal{I}_B$，我們給出一組訓練樣例

$$\amalg_{i\in\mathcal{I}_B}\ (rec_i, im_i)$$

在B上的圖像im_i和在B_{rec}上的相應重建圖像rec_i，我們應用一優化演算法，例如，基於該(隨機)梯度下降法，以找到權重$\Theta^B_1, \ldots, \Theta^B_{k_B}, \Psi^B \in \mathbb{R}^T$，俾最小化下列該表達式

$$\sum_{i\in\mathcal{I}_B} Loss^B(im_i, rec_i, \Theta^B_1, \ldots, \Theta^B_{k_B}, \Psi^B) \text{。} \qquad (6)$$

【0055】 **1.3 函數F^B_Θ和G^B_Θ的規範**

【0056】 在本節中，我們更精確地定義該函數F^B_Θ和G^B_Θ的形式。再次，回想神經網路 80 和 84 的一些定義。這些函數中的每一個由一系列

的函數組合所組成，這些函數組合是：1)一仿射變換Aff；或 2)一非線性活化函數Act。

【0057】 經由一仿射變換$Aff: \mathbb{R}^m \to \mathbb{R}^n$，我們指的是以下面形式的一映射

$$Aff(x) = L(x) + b \ ,$$

其中$L: \mathbb{R}^m \to \mathbb{R}^n$是一個線性變換，即滿足

$$L(\lambda \cdot x_1 + x_2) = \lambda \cdot L(x_1) + L(x_2)$$

對於所有$\lambda \in \mathbb{R}$，$x_1, x_2 \in \mathbb{R}^m$，並且其中$b \in \mathbb{R}^n$。每個線性映射$L: \mathbb{R}^m \to \mathbb{R}^n$完全由$\mathbb{R}^{n \times m}$中的一矩陣決定，即唯一地對應於一向量$\Theta_L \in \mathbb{R}^{m \cdot n}$。因此，每個仿射函數$Aff: \mathbb{R}^m \to \mathbb{R}^n$完全由$m \cdot n + n$個權重決定，即由一向量$\Theta \in \mathbb{R}^{m \cdot n + n}$決定。對於每個$\Theta \in \mathbb{R}^{m \cdot n + n}$，我們將為該上述方式中對應於$\Theta$的唯一仿射變換寫為$Aff_\Theta$。

【0058】 經由一非線性活化函數$Act: \mathbb{R}^n \to \mathbb{R}^n$，我們指的是以下列形式表達的一函數

$$(Act(x))_i = \rho(x_i) \ 。$$

【0059】 這裡，$((Act(x))_i$表示$Act(x)$的該第 i 個分量，x_i表示 x的該第 i 個分量。最後，$\rho: \mathbb{R} \to \mathbb{R}$可以是這種形式

$$\rho(z) = \begin{cases} z, \text{if } z \geq 0 \\ 0, \text{if } z < 0 \end{cases} \ , \qquad (7)$$

或形式

$$\rho(z) = (z) = \begin{cases} z, \text{if } z > 0 \\ e^z - 1, \text{otherwise} \end{cases} \ , \qquad (8)$$

儘管這些範例不應解釋為將本申請的實施方案限制於這些明確的範例。也可以使用其他公式，例如$\rho(z) = \log(1 + e^z)$或任何其他非線性函數。例如，$\rho: \mathbb{R} \to \mathbb{R}$可以另外地是一分段平滑函數。

【0060】 我們的函數F_Θ^B現在看起來如下。我們假設對於一固定的$k \in \mathbb{N}$，我們給出$m_1, \ldots, m_k \in \mathbb{N}$和$n_1, \ldots, n_k \in \mathbb{N}$，其中$m_1 = L$、$n_k = M$，這樣使得

$T = (m_1 \cdot n_1 + n_1) + (m_2 \cdot n_2 + n_2) + \cdots + (m_k \cdot n_k + n_k)$。

【0061】 這裡，$T \in \mathbb{N}$，$L \in \mathbb{N}$和$M \in \mathbb{N}$如公式(1)中所示。然後，對於$\Theta_1 \in \mathbb{R}^{m_1 \cdot n_1 + n_1}, \ldots, \Theta_k \in \mathbb{R}^{m_k \cdot n_k + n_k}$，其中$\Theta = (\Theta_1, \ldots, \Theta_k) \in \mathbb{R}^T$，我們定義

$F_\Theta^B := Aff_{\Theta_k} \circ \rho \circ Aff_{\Theta_{k-1}} \circ \rho \circ Aff_{\Theta_{k-2}} \circ \ldots \circ \rho \circ Aff_{\Theta_1}$。

【0062】 因此，F_Θ^B將描述使用參數Θ參數化的一神經網路 80_i。它將是線性函數Aff_{Θ_j}和非線性函數ρ的一序列，其在本範例中，其交替地被應用於該序列中，其中該參數Θ包括Aff_{Θ_j}中的該線性函數權重。在線性函數Aff_{Θ_j}和非線性函數ρ的該序列中，線性函數Aff_{Θ_j}後跟非線性函數ρ的該對將表示一神經元階層，例如，諸如第 j 階層，在神經網路的前饋方向上，該神經元階層 j 之前的前驅節點的該數量由Aff_{Θ_j}的維度 m 所決定，Aff_{Θ_j}的該行(column)數目和該神經元階層 j 本身的神經元的該數目由Aff_{Θ_j}的維度 n 決定，其係該列(row)數目。Aff_{Θ_j}的每一列(row)包含該權重，該權重控制分別活化該 m 個前驅神經元中的每一個的一信號強度有多強，其被轉發到該神經元階層 j 的該相應神經元，其係對應於該相應的列(row)。 ρ 控制神經元階層 j 的每個神經元，其前驅神經元活化的線性組合的該非線性映射到其自身的活化。在上面的例子中，有 k 個這樣的神經元階層。每一個神經元階層的該數量可以變化。

神經元階層 k 的該數量可以在各種神經網路 80_j 之間變化，即對於不同的 j。注意，該非線性函數可能根據神經元階層或甚至每個神經元或某些其他單位而有變化。

【0063】 類似地，我們的函數G_Ψ^B如下所示。我們假設對於一固定的$k' \in \mathbb{N}$，我們給出$m_{1'}, \ldots, m_{k'} \in \mathbb{N}$和$n_{1'}, \ldots, n_{k'} \in \mathbb{N}$，其中$m_{1'} = L$、$n_{k'} = K_B$，這樣使得

$$T = (m'_1 \cdot n'_1 + n'_1) + (m'_2 \cdot n'_2 + n'_2) + \cdots + (m'_{k'} \cdot n'_{k'} + n'_{k'}) \text{ 。}$$

【0064】 這裡，$T \in \mathbb{N}$、$L \in \mathbb{N}$和$K_B \in \mathbb{N}$如公式(3)中所示。然後，對$\Psi_1 \in \mathbb{R}^{m_{1'} \cdot n_{1'} + n_{1'}}, \ldots, \Psi_{k'} \in \mathbb{R}^{m_{k'} \cdot n_{k'} + n_{k'}}$，其中$\Psi = (\Psi_1, \ldots, \Psi_{k'}) \in \mathbb{R}^T$，我們定義

$$G_\Psi^B := Aff_{\Psi_{k'}} \circ \rho \circ Aff_{\Psi_{k'-1}} \circ \rho \circ Aff_{\Psi_{k'-2}} \circ \ldots \circ \rho \circ Aff_{\Psi_1} \text{ 。}$$

【0065】 因此，G_Ψ^B將描述使用參數Ψ參數化的一神經網路 84。它將是線性函數Aff_{Ψ_j}和非線性函數ρ的一序列，正如上面關於該預測信號計算的該神經元階層所描述的那樣。神經網路 84 的神經元階層k'的該數量可以與神經網路 80i 的神經元階層k的該數量中的一個或多個不同。

【0066】 **1.4 將現有預測考慮在內的訓練**

【0067】 我們擴展了該上一節的演算法，以便我們可以訓練預測，以補充已經存在的幀內預測。

【0068】 亦即，令

$$\mathcal{F}_{fix}^B := \{F_{1,fix}^B, \ldots, F_{\tilde{K}_B,fix}^B\}$$

其係為已經可用的固定幀內預測函數的一集合。例如，$\mathcal{F}_{fix}^B$可以包括

HEVC 的 DC 或平面預測以及根據 HEVC 定義的角度預測；所有這些預測還可以包括該重建樣本的一初步平滑。此外，我們假設我們有一個函數

$$\mathcal{L}_{fix}: \mathbb{R}^M \times \mathbb{R}^L \times \{1, \dots, \widetilde{K}_B\} \to \mathbb{R}$$

使得$\mathcal{L}_{\text{fix}}(\text{im}, rec, \text{k})$模擬在給定該原始圖像$im$的情況下，應用於$rec$的該第k個幀內預測函數$F^B_{k,fix}$的該損失。

【0069】 然後我們將該損失函數從公式(5)擴展到該損失函數

$$Loss^B_{\mathcal{F}^B_{fix}}(im, rec, \Theta^B_1, \dots, \Theta^B_{k_B}, \Psi^B)$$

$$:= \min(Loss^B(im, rec, \Theta^B_1, \dots, \Theta^B_{k_B}, \Psi^B), \min_{k \in \{1,\dots,\widetilde{K}_B\}} \mathcal{L}_{fix}(im, rec, k))。(9)$$

【0070】 保留上一節該末尾的該符號，我們決定權重$\Theta^B_1, \dots, \Theta^B_{k_B}, \Psi^B \in \mathbb{R}^T$經由最小化

$$\sum_{\text{i} \in \mathcal{I}_B} Loss^B_{\mathcal{F}^B_{fix}}(im_i, rec_i, \Theta^B_1, \dots, \Theta^B_{k_B}, \Psi^B) \quad (10)$$

在訓練樣例的一個大的集合上。

【0071】 為該目的，我們通常首先經由優化公式(6)以找到權重，然後初始化這些權重以找到可優化公式(10)的該權重。

【0072】 **1.5 聯合訓練幾種區塊形狀的預測**

【0073】 在本節中，在我們的預測的該訓練中，我們描述了如何考慮在一典型的視頻編碼標準中，通常可以以各種方式將區塊分割成更小的子區塊並對該較小的子區塊執行一幀內預測。

【0074】 即，假設對於某些$S \in \mathbb{N}$，我們給出可容許區塊$B_i \subset \mathbb{Z}^2$的一集合

$$\mathcal{BL} = \{B_1, \dots, B_S\}$$

以及區域的一集合

$$\mathcal{BL}^{rec} = \{B_1^{rec}, \dots, B_S^{rec}\}$$

使得每個$B_i^{rec} \subset \mathbb{Z}^2$是$B_i$的一鄰域。通常，$B_i^{rec}$是$B_i$的左側和上方的兩個矩形的一個聯集。

【0075】 我們假設存在一區塊$B_{max} \in \mathcal{BL}$，針對每個$i \in \{1, \dots, S\}$使得$B_i \subseteq B_{max}$。設$\mathcal{P}(\mathcal{BL})$為$\mathcal{BL}$的該冪集(power set)。然後對於$B \in \mathcal{BL}$，我們假設一集合

$$\mathcal{BL}(B) \subseteq \mathcal{P}(\mathcal{BL})$$

被給出，使得對於每個$Y = \{\mathrm{B}_{\mathrm{i}_1}, \dots, \mathrm{B}_{\mathrm{i}_k}\} \in \mathcal{BL}(\mathrm{B})$，該區塊$B$可以寫成一個不相交的聯集

$B = \coprod_{B_i \in Y} B_i$ 。

【0076】 於一給定的顏色分量，令im為B_{max}上的一圖像，其經由限制，針對每個$B_i \in \mathcal{BL}$我們視為B_i上的一圖像$im|_{B_i}$。此外，假設在$\mathrm{U}_{i=1}^{S}\ \mathrm{B}_i^{rec}$上存在一重建圖像$rec$，其經由限制，針對每個$\mathrm{B}_i^{rec} \in \mathcal{BL}^{rec}$我們視為$\mathrm{B}_i^{rec}$上的一圖像$rec|_{\mathrm{B}_i^{rec}}$。

【0077】 保留 1.2 節的該符號，對於每個$B \in \mathcal{BL}$，我們尋找$\Theta_1^B, \dots, \Theta_{K_B}^B \in \mathbb{R}^T$作為$\mathrm{K}_B$幀內預測函數$F_{\Theta_k}^B$的該權重集合，並且我們尋求$\Psi^B \in \mathbb{R}^T$作為該模式預測函數$G^B$的權重。我們共同決定針對所有$\mathrm{B} \in \mathcal{BL}$的這些權重如下。對於$\mathrm{B} \in \mathcal{BL}$和給定的權重集合$\{\Theta_k^{B'}\}$，$B' \in \mathcal{BL}$，$B' \subseteq B$，$1 \le k \le K_{B'}$，我們把

$$\Theta^B := \coprod_{B' \in \mathcal{BL}: B' \subseteq B} \coprod_{k=1}^{K_{B'}} \Theta_k^{B'}, \quad \Psi^B := \coprod_{B' \in \mathcal{BL}: B' \subseteq B} \Psi^{B'} \text{ 。}$$

【0078】 此外，對於$B' \subset B$，我們將$\Theta^B|_{B'} \subset \Theta^B$定義為

$$\Theta^B|_{B'} := \coprod_{B'' \in \mathcal{BL}: B'' \subseteq B'} \coprod_{k=1}^{K_{B''}} \Theta_k^{B''}, \quad \Psi^B|_{B'} := \coprod_{B'' \in \mathcal{BL}: B'' \subseteq B'} \Psi^{B''} \text{ 。}$$

【0079】 如在 1.4 節中，我們假設對於每個$\mathrm{B} \in \mathcal{BL}$，幀內預測函數的一可能空集合$\mathcal{F}_{\mathrm{fix}}^{\mathrm{B}}$是可用的。我們令

$$\mathcal{F}_{fix} := \coprod_{B \in \mathcal{B}} \mathcal{F}_{fix}^B \text{ 。}$$

【0080】 然後我們如下定義一損失函數$Loss_{\mathcal{F}_{fix}}^{B,total}$。經由集合的該含括，我們在該集合$\mathcal{BL}$上有一個排序≤。令

$$\mathcal{BL}^{\min} \subseteq \mathcal{BL}$$

為$\mathcal{BL}$中所有最小元素的該集合。對於$\mathrm{B} \in \mathcal{BL}^{\min}$，我們把

$$Loss_{\mathcal{F}_{fix}}^{B,total}(im, rec, \Theta^B, \Psi^B) := Loss_{\mathcal{F}_{fix}^B}^B(im, rec, \Theta^B, \Psi^B) \text{ ，}$$

其中，該後者的函數如公式(9)所示。

【0081】 接下來，讓$B \in \mathcal{BL}$並假設已經為$B' \subseteq B$的所有 $B' \in \mathcal{BL}$定義了 $Loss^{B,total}$ 。

【0082】 然後，我們定義

$$Loss_{\mathcal{F}_{fix}}^{B,total}(im, rec, \Theta^B, \Psi^B) := \min(Loss_{\mathcal{F}_{fix}^B}^B(im, rec, \Theta_1^B, \ldots, \Theta_{K_B}^B, \Psi_{K_B}),$$

$$\min_{Y \in \mathcal{BL}(B)} \sum_{B' \in Y} Loss_{\mathcal{F}_{fix}}^{B',total}(im, rec, \Theta^B|_{B'}, \Psi^B|_{B'})) \text{ 。}$$

【0083】 最後，給出訓練樣例的一固定集合

$$\coprod_{\mathrm{i} \in \mathcal{I}_B} (rec_{\mathrm{i}}, im_{\mathrm{i}})$$

其係為在$\mathrm{B}_{\max}$上圖像im_{i}的固定集合，我們經由最小化或至少使該表達式變小來決定$\Theta^{B_{\max}}$、$\Psi^{B_{\max}}$

$$\sum_{i \in \mathcal{I}} Loss_{\mathcal{F}_{fix}}^{B_{\max},total}(im_i, rec_i, \Theta^{B_{\max}}, \Psi^{B_{\max}}) \text{ 。}$$

【0084】 我們通常經由首先單獨地對每個$\mathrm{B} \in \mathcal{BL}$最小化公式(9)來

初始化權重Θ_k^B、Ψ^B。

【0085】 **2.將經過訓練的神經網路集成到一視頻編解碼器中**

【0086】 我們考慮一混合視頻編碼標準，其中對於一給定的顏色分量，在一給定區塊$B \subset \mathbb{Z}^2$上的一視頻信號的該內容將由一解碼器生成。設M是B的像素數目。此外，令$B_{rec} \subset \mathbb{Z}^2$是$B$的一固定鄰域，使得該解碼器在其處理上具有在$B_{rec}$上的一重建圖像$rec$。設$L$是$B_{rec}$的該像素數目。然後我們將$rec$視為$\mathbb{R}^L$的一個元素。我們假設該編解碼器經由該當前區塊$B$ 10 上的預測編碼來操作。然後，我們聲稱對於一解碼器可以執行的以下步驟的版權，以便在B上生成一預測信號$pred$，我們將其視為$\mathbb{R}^M$的一個元素：

1. 該解碼器具有在其處理下的固定數量K_B、$T \in \mathbb{N}$、函數

$F^B: \mathbb{R}^L \times \mathbb{R}^T \to \mathbb{R}^M$, 亦即 $80_1 \ldots 80_{(CB-1)}$及 $G_B: \mathbb{R}^L \times \mathbb{R}^T \to \mathbb{R}^{K_B}$ ，亦即 84 ，

以及權重$\Theta_1, \ldots, \Theta_{K_B} \in \mathbb{R}^T$和一權重$\Psi \in \mathbb{R}^T$，其中後者權重是經由前一章節中描述的一訓練演算法預先決定的。

2. 該解碼器從該位元流重建作為輔助信息 70 的一部分的一旗標，並指示該以下選項中的恰好之一是否為真：

(i)該預測之一$F^B_{\Theta_{B,1}}, \ldots, F^B_{\Theta_{B,K_B}}$將被使用，亦即集合 72 之外的一模式；

(ii)沒有使用該預測 $F^B_{\Theta_{B,1}}, \ldots, F^B_{\Theta_{B,K_B}}$ ，亦即例如 74 之外的一個

這裡，該函數$F^B_{\Theta_{B,K_B}}$如公式(2)中所示。

3. 如果步驟 2 中的選項二為真，則該解碼器繼續進行該給定區塊 10，如在該基礎混合視頻編碼標準中那樣。

4. 如果步驟 2 中的選項一為真，則該解碼器將根據公式(4)定義的該函數G_{Ψ}^{B}，亦即 84，應用於該重建的圖像rec。令$X:=(x_1,\ldots,x_{K_B})\in\mathbb{R}^{K_B}$被定義為$\mathrm{X}:=\mathrm{G}_{\Psi}^{\mathrm{B}}(rec)$。然後以一種方式改變該標準，使得該解碼器經由該以下兩個選項中的一個，以定義一數字$m\in\{1,\ldots,K_B\}$：

(i)該解碼器在該集合$\{1,\ldots,\mathrm{K_B}\}$上定義一機率分$P_{G_{\Psi}^{B}(rec)}$如下所示

$$P_{G_{\Psi}^{B}(rec)}(i):=\frac{e^{x_i}}{\sum_{j=1}^{K_B}e^{x_j}}$$ 。

並且使用該後者的機率分佈$P_{G_{Psi}^{B}(rec)}$來解析一索引$k\in\{1,\ldots,K_B\}$，其也是來自該資料流 12 的該基礎標準中使用的該熵編碼引擎的輔助信息 70 的一部分，並定義$m:=k$。

(ii)該解碼器定義一排列

$$\sigma:\{1,\ldots,K_B\}\rightarrow\{1,\ldots,K_B\}$$

由將$\sigma(1):=k_{\max}^{1}$歸納，其中$k_{\max}^{1}\in\{1,\ldots,K_B\}$是最小數目，對於所有$\mathrm{k}\in\{1,\ldots K_B\}$，$x_{k_{max}^{1}}\geq\mathrm{x}_k$，並且經由把$\sigma(\mathrm{l}+1):=k_{max}^{l+1}$，其中$k_{max}^{l+1}\in\{1,\ldots,\mathrm{K}_B\}\backslash\{\sigma(1),\ldots,\sigma(l)\}$是最小數目，針對所有$k\in\{1,\ldots,K_B\}\backslash\{\sigma(1),\ldots,\sigma(l)\}$，使得具有$x_{k_{max}^{l+1}}\geq x_k$。

然後，該解碼器從該位元流 12 重建一個唯一索引$i\in\{1,\ldots,K_B\}$，其也是資料流 12 的一部分並且設置$m:=\sigma(i)$。

在解析該後者索引 i 的該代碼設計中，如果$\sigma(i_1)\leq\sigma(i_2)$，要求信號通知一索引$i_1\in\{1,\ldots,K_B\}$所需的位元數目小於或等於發信號通知一索引

的位元數$i_2 \in \{1, \ldots, K_B\}$，並且如果該熵編碼引擎使用的所有涉及的基礎機率被設置為相等的機率。

5. 如果步驟 2 中的選項一為真並且如果該解碼器根據該前一步驟 4 決定了該索引m，則該解碼器生成 71 該預測信號$pred \in \mathbb{R}^M$作為 $pred = F_{\Theta_m}^B(rec)$，亦即使用該選定的神經網路 80_m。然後，該解碼器如在該基礎混合視頻編碼標準中那樣使用$pred$作為預測信號進行。

【0087】 將設計基於一資料驅動學習方法的幀內預測函數集成到一現有的混合視頻編解碼器中。該描述有兩個主要部分。在該第一部分中，我們描述了用於幀內預測函數的一離線訓練的一具體演算法。在該第二部分中，我們描述了一視頻解碼器如何使用該後者的預測函數以便為一給定區塊生成該預測信號。

【0088】 因此，以上在章節 1.1 至 2 中描述的內容尤其是用於對來自一資料流 12 的一圖片 10 進行逐塊解碼的一裝置。該裝置 54 支持多個幀內預測模式，其至少包括幀內預測模式的一集合 72，根據該幀內預測模式，經由應用該當前區塊 18 的相鄰樣本的一第一集合 60 到一神經網路 80_i 上，來決定該圖片 10 的該當前區塊 18 的該幀內預測信號。該裝置 54 被配置為從多個 66 幀內預測模式中選擇 68 用於該當前區塊 18 的一個幀內預測模式，並且使用該一個幀內預測模式預測 71 該當前區塊 18，亦即使用被選中該相應的神經網路 80_m。儘管在第 2 章節中呈現的該解碼器，其具有在多個 66 幀內預測模式內的幀內預測模式 74，該多個 66 幀內預測模式支持另外的在集合 72 中的基於神經網路的幀內預測模式，但這僅僅是一個範例並且不需要是這種情況。此外，

章節 1 和 2 中的該以上描述可以變化的，其中該解碼器 54 不使用、並且不包括該另外的神經網路 84。關於該上述優化，這意味著在 1.2 節中呈現的用於找出k_{opt}^{B}的在該內部質量的該第二加法器不是必須為一個函數 M^B的串聯，其應用於任何機率值神經網路函數 G^B。相反，該優化演算法以一種方式決定該神經網路 80_i的合適參數，使得該選擇頻率將適當地遵循 M^B的一碼率指示。例如，該解碼器 54 可以使用一可變長度碼從資料流 12 解碼區塊 18 的一索引，其碼長以 M^B指示，並且該解碼器 54 將基於該索引執行選擇 68。該索引將是該輔助信息 70 的一部分。

【0089】 以上在第 2 章節中提出的該描述的另一替代方案是，該解碼器 54 可替代地導出基於神經網路的幀內預測模式的該集合 72 中的一排序，其取決於與該當前區塊 18 的一鄰域相關的該資料流的一第一部分，以便獲得幀內預測模式的一有序列表，並從幀內預測模式的該有序列表中最終選擇該幀內預測模式以使用，其取決於不同於該第一部分的該資料流的一第二部分。該「第一部分」可以例如涉及與當前區塊 18 相鄰的一個或多個區塊相關的一編碼參數或預測參數。然後，該“第二部分”可以是一索引，例如，指向該基於神經網路的幀內預測模式集合 72 或者是該基於神經網路的幀內預測模式集合 72 的一索引。當與上面概述的章節 2 一致時，該解碼器 54 包括該另外的神經網路 84，其對於幀內預測模式的該集合 72 的每一個幀內預測模式，經由將相鄰樣本的集合 86 應用於其上、並對這些機率值進行排序來決定一機率值，以便決定集合 72 的每一個幀內預測模式的一秩，從而獲得幀內預測模式的一有序列表。然後將該資料流 12 中作為輔助信息 70 的一部分的一索

引用作該有序列表的一索引。這裡，可以使用可變長度代碼對該索引進行編碼，其中 M^B 表示該代碼長度。並且如上面在章節 2 中所解釋的，在項目 4i 中，根據一另一替代實施例，解碼器 54 可以使用由該另外的神經網路 84 決定的該剛剛提到的機率值用於集合 72 的每一個基於神經網路的幀內預測模式，以便有效地執行索引的熵編碼到集合 72 中。特別地，該索引的該符號字母表是該輔助信息 70 的一部分並且用作對集合 72 的一索引，其將包括針對集合 72 內的每一個模式的一符號或值，以及由神經網路 84 提供的該機率值，在根據以上描述的神經網路 84 設計的情況下，其將提供導致有效熵編碼的機率值，其中這些機率值緊密地表示該實際符號的統計特性。對於該熵編碼，例如可以使用算術編碼，或機率區間劃分熵(PIPE)編碼。

【0090】 有利地，對於集合 72 的任何幀內預測模式，不需要附加信息。每個神經網路 80_i 針對編碼器和解碼器，一旦有利地根據例如在章節 1 和 2 中的上述描述被參數化，其在沒有資料流中的任何附加指導的情況下導出該當前區塊 18 的預測信號。如上所述，除了集合 72 中的該基於神經網路的模式之外，其他幀內預測模式的存在是可選的。它們已在上面用集合 74 表示。在這方面，應當注意，選擇集合 60 的一種可能的方式，亦即形成用於幀內預測 71 的該輸入的相鄰樣本的該集合，其可以如此被選擇使得針對集合 74 的該幀內預測模式，該集合 60 是相同的，亦即啟發式的，針對基於神經網路的幀內預測模式的集合 60，在集合 60 中所包括的相鄰樣本的數量方面更大，並且影響幀內預測 71。換句話說，與集合 74 的其他模式相比，對於基於神經網路的幀內預測

模式 72，集合 60 的該基數可以更大。例如，集合 74 的任何幀內預測模式的集合 60 可以僅包括沿著一直延伸到區塊 18 的側面的一條一維直線的相鄰樣本，例如左邊一側和上邊一側。基於神經網路的幀內預測模式的集合 60 可以覆蓋在區塊 18 的剛剛提到的側面旁邊延伸的一個 L 形部分，但是比用於集合 74 的幀內預測模式的集合 60 寬一個樣本寬度。以此種方式，基於神經網路的幀內預測模式可產生具有一相應較低預測殘差的一更好幀內預測。

【0091】 如上面在章節 2 中所描述的，在該資料流 12 內傳送到一個幀內預測區塊 18 的該輔助信息 70 可以包括一斑點，該斑點通常指示針對區塊 18 所選幀內預測模式是否為集合 72 的成員或者為集合 74 的成員。然而，該斑點對於輔助信息 70 僅僅是可選的，指示例如包括集合 72 和 74 的一整個多個 66 幀內預測模式的一索引。

【0092】 在下文中，關於圖 7a 至圖 7d 簡要討論了剛才討論的替代方案。該等圖式同時定義解碼器和編碼器，亦即就其在一幀內預測區塊 18 中的功能而言。關於一幀內編碼區塊 18 的編碼器操作模式和解碼器操作模式之間的差異在於，一方面是該編碼器執行所有或至少一些可用的幀內預測模式 66 的事實，以便在 90 處決定最佳的一個，例如一些成本函數最小化的方式，並且該編碼器形成資料流 12，亦即編碼資料到資料流 12，而該解碼器分別經由解碼和讀取從中導出資料。圖 7a 係顯示該上述替代方案的該操作模式，根據該操作模式，用於區塊 18 的該輔助信息 70 內的一旗標 70a 指示在步驟 90 中該編碼器決定幀內預測模式為區塊 18 的該最佳模式，其是否在集合 72 內，亦即是基於神經網

路的幀內預測模式，或在集合 74 內，亦即是基於非神經網路的幀內預測模式之一。該編碼器相應地將旗標 70a 插入資料流 12 中，而該解碼器從中取回旗標 70a。圖 7a 假設所決定的幀內預測模式 92 在集合 72 內。然後，該分離的神經網路 84 決定用於集合 72 的每一個基於神經網路的幀內預測模式的一機率值，並且使用這些機率值集合 72，或者更確切地說，其中基於神經網路的幀內預測模式根據它們的機率值被排序，例如按其機率值的降序排列，從而得到幀內預測模式的一有序列表 94。作為輔助信息 70 的一部分的一索引 70b 然後由該編碼器編碼到資料流 12 中，並由該解碼器從其中解碼。因此，該解碼器能夠決定集合 72 和集合 74 的哪個。在要使用的該幀內預測模式位於集合 72 中的情況下，要用於區塊 18 的該幀內預測模式位於並且執行集合 72 的排序 96。在該所決定的幀內預測模式位於集合 74 中的情況下，也可能在資料流 12 中發送一索引。因此，該解碼器能夠經由相應地控制該選擇 68，使用所決定的幀內預測模式來生成用於區塊 18 的預測信號。

【0093】 圖 7b 係顯示一替代方案，根據該替代方案，該旗標 70a 不存在於資料流 12 中。相反，該有序列表 94 將不僅包括集合 72 的該幀內預測模式，而且還包括集合 74 的幀內預測模式。輔助信息 70 內的該索引將是該更大有序列表的一索引，並指示該所決定的幀內預測模式，亦即該一模式被決定為優化 90。在神經網路 84 僅在 72 內提供基於神經網路的幀內預測模式的機率值的情況下，集合 72 的幀內預測模式之間相對於集合 74 的幀內預測模式的該排序可以經由其他方式決定，例如，不可避免地在順序列表 94 中，將集合 72 的該基於神經網路

的幀內預測模式安排在集合 74 的該模式之前，或者相對於彼此安排它們交替排列。也就是說，該解碼器能夠從資料流 12 導出該索引，使用該索引 70 作為索引到該有序列表 94 中，並使用神經網路 84 輸出的該機率值從該多個幀內預測模式 66 導出該有序列表 94。圖 7c 係顯示一個另一種變型。圖 7c 係顯示不使用旗標 70a 的一情況，但是可以另外地使用該旗標。圖 7c 所針對的問題涉及編碼器和解碼器都不使用神經網路 84 的該可能性。相反，該排序 96 是經由其他方式導出的，例如相對於一個或多個相鄰區塊 18 在資料流 12 內傳送的編碼參數，亦即與這樣的一個或多個相鄰區塊有關的一資料流 12 的部分 98。

【0094】 圖 7d 係顯示圖 7a 的一個另一變體，即根據其使用熵編碼對該索引 70b 進行編碼並使用熵解碼從資料流 12 解碼的該變體，通常使用附圖標記 100 表示。用於該熵編碼 100 的該樣本統計或該機率分佈由神經網路 84 輸出的該機率值控制，如上所述，這使得索引 70b 的該熵編碼非常有效。

【0095】 對於所有實施例 7a 至 7d，其確實可以不存在集合 74 模式。因此，該相應的模塊 82 可能丟失，並且旗標 70a 無論如何都是不必要的。

【0096】 此外，儘管未在任何圖式中顯示，但顯然即使沒有任何顯式信令 70，在該編碼器和解碼器處的該模式選擇 68 也可彼此同步，亦即不花費任何輔助信息。相反，該選擇可以從其他方法導出，例如經由不可避免地取該有序列表 94 中的該第一個，或者經由基於與一個或多個相鄰區塊有關的編碼參數將該索引導出到該有序列表 94 中。圖 8 係

顯示用於設計要用於基於區塊的圖像編碼的集合 72 的幀內預測模式的該集合的一裝置。該裝置 108 包括一可參數化網路 109，其繼承或包括神經網路 80_0 至 80_{KB-1} 的可參數化版本以及神經網路 84。這裡，在圖 8 中，繪示為單獨的單元，即神經網路 840 用於提供基於神經網路的幀內預測模式 0 的該機率值，至神經網路 84_{KB-1} 用於提供與該基於神經網路的內部預測模式 K_{B-1} 相關聯的該機率值。用於參數化神經網路 84 的該參數 111 和用於參數化神經網路 80_0 至 80_{KB-1} 的該參數 113 被一更新器 110 輸入或應用到這些神經網路的相應參數輸入。裝置 108 可以訪問一儲存器或多個圖像測試區塊 114 以及相應的相鄰樣本集合 116。成對的這些區塊 114 及其相關的相鄰樣本集合 116 被裝置 108 順序地使用。具體地，一當前圖片測試區塊 114 被應用於可參數化神經網路 109，使得神經網路 80 為集合 72 的每一個基於神經網路的幀內預測模式提供一預測信號 118，並且每一個神經網路 80 為這些模式的每一個提供一機率值。至此，這些神經網路使用它們的當前參數 111 和 113。

【0097】 在以上描述中，rec 已被用於表示該圖片測試區塊 114，並且$F_{\Theta_B}^{B}(rec)$是模式 B 的該預測殘差 118，並且機率值$G_{\Psi_B}^{B}(rec)$是該機率值 120。對於每個模式 0、... K_{b-1}，裝置 108 包括一成本估計器 122，其基於針對該相應模式所獲得的該預測信號 118，來計算該相應模式的一成本估計。在該上面的例子中，成本估計器 122 計算了在 1.2 節中該不等式的該左側和右側所示的該成本估算。也就是說，這裡，成本估計器 122 還針對每個模式使用該對應的機率值 120。然而，這不需要是如上所述的情況。然而，成本估計在任何情況下都是兩個加法項的總和，

其中一個加法項是在該上述不等式中的該項目$\tilde{R}$所表示的該預測殘差的該編碼成本的一估計，以及另一個加法項估計指示該模式的該成本編碼。為了計算與該預測殘差相關的該編碼成本的該估計，該成本估計器 122 還獲得該當前圖片測試區塊 114 的該原始內容。該神經網路 80 和 84 在其輸入處具有該相應的相鄰樣本集合 116。由成本估計器 122 輸出的該成本估計 124 由一最小成本選擇器 126 接收，該最小成本選擇器 126 決定該模式最小化或具有與之相關的最小成本估計。在上面的數學符號中，這是k_{opt}^{B}。該更新器接收該最佳模式並使用一編碼成本函數，該編碼成本函數具有一第一加法項以及一第二加法項，該第一加法項係取決於針對最低編碼估計的該幀內預測模式所獲得的該預測信號 118 形成殘差率估計，該第二加法項係形成一模式以信號通知輔助信息速率估計，其取決於該預測信號和針對由選擇器 126 指示的最低編碼成本估計的該幀內預測模式所獲得的該機率值。如上所述，這可以使用一梯度距離來完成。因此，該編碼成本函數是可微分的，並且在上面的數學表示中，該函數的一個例子在公式(5)中給出。此處，與該模式信號通知輔助信息速率估計有關的該第二加法項則計算用於最低編碼成本估計的該幀內預測模式的該交叉熵。

【0098】 因此，該更新器 110 尋求更新參數 111 和 113 以便減少該編碼成本函數，然後這些更新的參數 111 和 113 被可參數化神經網路 109 所使用，以便處理該多個 112 中的下一個圖像測試區塊。如上面關於章節 1.5 所討論的，可以存在一種機制，主要控制那些圖像測試區塊 114 及其相關的相鄰樣本集合 116，其被應用於該幀內預測的該遞歸

更新過程，在速率失真意義上，較佳地在沒有任何區塊子劃分的情況下完成，從而避免在圖像測試區塊的基礎上對該參數 111 和 113 進行過多優化，無論如何，以子區塊為單位的一編碼則更具成本效益。

【0099】 到目前為止，上述討論的實施例主要涉及編碼器和解碼器在其支持的幀內預測模式 66 內具有基於神經網路的幀內預測模式的一集合的情況。根據參考圖 9a 和 9b 討論的該實施例，這不一定是必須的情形。圖 9a 試圖概述根據一個實施例的一編碼器和一解碼器的該操作模式，其中聚焦於與上面關於圖 7a 提出的描述的不同之處的一種方式，以提供其描述。所支持的多個 66 幀內預測模式可以包括或不包括基於神經網路的幀內預測模式，並且可以包括或不包括基於非神經網路的幀內預測模式。分別由編碼器和解碼器所包含的圖 9a 中的該模塊 170，為了替每個該支持的模式 66 提供該相應的預測信號，因此不一定是神經網路。如上所述，這種幀內預測模式可以是基於神經網路的，或者它們可以是啟發式激勵、並且基於一 DC 幀內預測模式或一角度幀內預測模式或任何其他模式來計算該預測信號。因此，這些模塊 170 可以表示為預測信號計算機。然而，根據圖 9a 的該實施例的編碼器和解碼器包括一神經網路 84。該神經網路 84 基於該相鄰樣本集合 86 計算該所支持的幀內預測模式 66 的機率值，使得多個 66 幀內預測模式可以變為該有序列表 94。用於區塊 18 的資料流 12 內的該索引 70 指向該有序列表 94。因此，該神經網路 84 有助於降低用於幀內預測模式信號通知的該輔助信息速率。

【0100】 圖 9b 係顯示圖 9a 的一替代方案，其中代替該排序，該索

引 70 的熵解碼/編碼 100 用於控制其機率或簡單統計，亦即控制在編碼器/解碼器中熵解碼/編碼的該熵機率分佈，其係根據針對多個 66 中的每一個模式為該神經網路 84 決定的該機率值。

【0101】 圖 10 係顯示用於設計或參數化神經網路 84 的一裝置。因此，它是用於設計一神經網路的一裝置 108，神經網路用於輔助在幀內預測模式的一集合 66 中進行選擇。這裡，對於集合 66 中的每個模式，存在相應的神經網路區塊一起形成神經網路 84，並且裝置 108 的該可參數化神經網路 109 僅可參照這些區塊進行參數化。對於每種模式，還存在該預測信號計算機 170，然而，根據圖 10，該預測信號計算機 170 不需要可參數化。因此，基於由相應的該預測信號計算機 170 計算的該預測信號 118，圖 10 的裝置 108 來計算每種模式的成本估計，並且可選地，基於由針對此模式的該相應神經網路區塊所決定的相應機率值，來計算每種模式的成本估計。在所該得到的成本估計 124 的基礎上，該最小成本選擇器 126 選擇該最小成本估計的該模式，並且該更新器 110 更新該神經 84 網路的該參數 111。

【0102】 以下係關於圖 7a 至圖 7d 和圖 9a 和圖 9b 該描述的註釋。圖 9a 和 9b 的該實施例的一共同特徵是也用於圖 7a 至圖 7d 的一些實施例，其事實是該神經網路值的該機率值，以便改善或減少用於在該優化過程 90 中在該編碼器側該決定的模式向該解碼器發信號通知的該輔助信息 70 的開銷。如上面關於圖 7a 至圖 7d 的該實施例所示，然而，應該清楚的是，圖 9a 和圖 9b 的該實施例可以改變到一個程度，亦即相對於該模式選擇根本沒有在資料流 12 中花費輔助信息 70 的程度。相

反，神經網路 84 為每個模式輸出的該機率值可用於不可避免地同步在編碼器和解碼器之間的該模式選擇。在那種情況下，關於該模式選擇，在該編碼器側將不存在優化決定 90。而是，將在編碼器和解碼器側以該相同的方式決定在集合 66 中的被使用的該模式。關於圖 7a 至圖 7d 的相應實施例，一類似的陳述也是如此，當改變時，以便在資料流 12 中不使用任何輔助信息 70。回到圖 9a 和圖 9b 的該實施例，然而，有趣的是，在該解碼器側的該選擇過程 68 係依賴於該神經網路輸出的該機率值，其中該排序或該機率分佈估計對該機率值的依賴性改變了對該輔助信息的該解釋，就該編碼器而言，對該機率值的該依賴性不僅影響將該輔助信息 70 編碼到資料流 12 中，例如使用對該有序列表的一索引的相應可變長度編碼、或者根據該神經網路的機率值得一機率分佈估計而使用熵編碼/解碼，而且影響該優化步驟 90：這裡，用於發送輔助信息 70 的該碼率可以被考慮，並且因此可以影響該決定 90。

【0103】 儘管已經在一裝置的該上下文中描述了一些觀點，但是顯然這些觀點也表示該對應方法的一描述，其中一區塊或設備對應於一方法步驟或一方法步驟的一特徵。類似地，在一方法步驟的該上下文中描述的觀點還表示一對應裝置的一對應區塊或項目或特徵的一描述。一些或所有方法步驟可以由(或使用)一硬體設備執行，例如一微處理器、一可編程計算機或一電子電路。在一些實施例中，一個或多個該最重要的方法步驟可以由這樣的一裝置執行。

【0104】 本揭露的該編碼資料流可以存儲在一數位存儲媒體上，或者可以在諸如一無線傳輸媒體的一傳輸媒體或諸如因網際網路的一有線傳輸媒體上傳輸。

【0105】 根據某些實現要求，本揭露的實施例可以用硬體或軟體實現。T 可以使用數位存儲媒體來執行該實現，例如一軟盤、一 DVD、一藍光、一 CD、一 ROM、一 PROM、一 EPROM、一 EEPROM 或一 FLASH 記憶體，其具有存儲的電子可讀控制信號在其上，與一可編程計算機系統協作(或能夠協作)，從而執行相應的方法。因此，數位存儲媒體可以是計算機可讀的。

【0106】 根據本揭露的一些實施例，其包括具有電子可讀控制信號的一資料載體，其能夠與一可編程計算機系統協作，從而執行本文所述的方法之一。

【0107】 通常，本揭露的實施例可以實現為具有一程式代碼的一計算機程式產品，當該計算機程式產品在一計算機上運行時，該程式指令可操作用於執行這些方法之一。程式指令可以例如儲存在機器可讀媒體上。

【0108】 其他實施例包括用於執行存儲在機器可讀媒體上的本文所述方法之一的計算機程式。

【0109】 換句話說，本揭露方法的一實施例因此是具有一程式代碼的一計算機程式，當該計算機程式在一計算機上運行時，該程式代碼用於執行本文描述的方法之一。

【0110】 因此，本揭露方法的一另一實施例是一資料載體(或數位

存儲媒體、或計算機可讀取媒體)，其包括在其上記錄的用於執行本文描述的方法之一的計算機程式。該資料載體，數位存儲媒體或記錄媒體通常是有形的和/或非過渡的。

【0111】 因此，本揭露方法的一另一實施例是表示用於執行本文所述方法之一的計算機程式的一資料流或依信號序列。該資料流或該信號序列可以例如被配置為經由一資料通信連接，例如經由網際網路傳輸。

【0112】 一另一實施例包括一處理裝置，例如計算機或可編程邏輯設備，其被配置為或適於執行本文所述的方法之一。

【0113】 一另一實施例包括一計算機，其上安裝有用於執行本文所述方法之一的該計算機程式。

【0114】 根據本揭露的一另一實施例包括一種裝置或系統，其被配置為將用於執行本文所述方法之一的計算機程式傳送(例如，電子地或光學地)給一接收器。該接收器可以是例如一計算機、一移動設備、一記憶體設備等。該裝置或系統可以例如包括用於將計算機程式傳送到接收器的檔案伺服器。

【0115】 在一些實施例中，一可程式化邏輯器件(例如，一現場可編輯閘陣列)可用於執行本文描述的該方法的一些或全部功能。在一些實施例中，一現場可編輯閘陣列可以與一微處理器協作，以便執行本文描述的方法之一。通常，該方法較佳地由任何硬體設備執行。

【0116】 這裡描述的該裝置可以使用一硬體設備，或使用一計算機，或使用一硬體設備和一計算機的一組合來實現。

【0117】 這裡描述的該裝置或這裡描述的該裝置的任何組件可以

至少部分地以硬體和/或軟體實現。

【0118】 這裡描述的該方法可以使用一硬體設備，或使用一計算機，或使用一硬體設備和一計算機的一組合來執行。

【0119】 本文描述的該方法或本文描述的該裝置的任何組件可以至少部分地由硬體和/或軟體執行。

【0120】 上述範例僅代表了本揭露該原理的一說明。應理解，本領域其他技術人員將理解本文所述的佈置和細節的任何修改和變化。上述實施例僅係為了方便說明而舉例而已，本揭露所主張之權利範圍自應以申請專利範圍該為準，而非僅限於上述實施例。

【0121】 在借助附圖對本揭露的實施例進行詳細說明之前，應當注意的是，在不同的附圖中，相同的、功能上相同的和相等的元件、物件和/或結構被提供有相同的附圖標記，使得不同實施例中的這些元件的描述是可互換和/或相互適用的。

【符號說明】

【0122】

10：圖片

12：資料流

14：裝置/編碼器

16：視頻

18：區塊

20：編碼順序

22：減法器

24：預測信號

26：預測殘差信號

28：預測殘差編碼器

28a：有損編碼級

28b：無損編碼級/熵編碼器

30：量化器

32：變換級

34：量化預測殘差信號

36：預測殘差信號重建級

38、38’：去量化器

40、40’：逆變換器

42、42’：加法器

44、44’：預測器

46、46’：環路濾波器

54：解碼器/裝置

56：熵解碼器

36’：殘差信號重建級

62：相鄰樣本

60：集合

64：樣本

68：選擇

66：多個預測模式

71：預測

70：輔助信息信號通知/輔助信息

72：集合

74：集合/模式

76：部分

80_0~80_{KB-1}：神經網路

82：區塊/模塊

84：神經網路

86：集合

90：處/步驟/優化/優化決定/決定

70a：旗標

92：幀內預測模式

94：有序列表

70b：索引

96：排序

98：部分

100：變體/熵解碼/編碼/熵編碼

70：顯式信令

108：裝置

109：可參數化網路/可參數化神經網路

111、113：參數

110：更新器

114：圖像測試區塊

116：相鄰樣本集合

118：預測信號/預測殘差/預測信號

80：神經網路

120：機率值

122：成本估計器

126：最小成本選擇器

112：該多個

170：模塊/預測信號計算機

124：成本估計

【發明申請專利範圍】

【第1項】 一種用於對來自一資料流的一圖片進行逐塊解碼的裝置，該裝置支持多個幀內預測模式，其包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

根據該資料流的一第一部分決定幀內預測模式的該集合的一排名，該第一部分與該當前區塊的一鄰域相關，以便獲得幀內預測模式的一有序列表，以及

根據除該第一部分之外的該資料流的一第二部分，從幀內預測模式的該有序列表中選擇該一個幀內預測模式。

【第2項】 如申請專利範圍第1項所述之裝置，其被配置為：

使用一可變長度代碼從該資料流解碼一索引，以及

使用該索引執行該選擇。

【第3項】 一種用於對來自一資料流的一圖片進行逐塊解碼的裝置，該裝置支持多個幀內預測模式，其包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一秩（rank），以便獲得幀內預測模式的一有序列表，

從該資料流解碼一索引到該有序列表中，以及

使用該索引和該有序列表執行該選擇。

【第4項】 如申請專利範圍第3項所述之裝置，其被配置為使用一可變長度代碼從該資料流將該索引解碼到該有序列表。

【第5項】 一種用於對來自一資料流的一圖片進行逐塊解碼的裝置，該裝置支持多個幀內預測模式，其包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一機率值，

使用該機率值從該資料流將一索引熵解碼到幀內預測模式的該集合，以及

使用該索引執行該選擇。

【第6項】 如申請專利範圍第1項所述之裝置，其中，相鄰樣本的該第一集合與相鄰樣本的該第二集合重合。

【第7項】 如申請專利範圍第1項所述之裝置，其中，幀內預測模式該的集合中的每一個，專門地從該當前區塊的相鄰樣本的該第一集合唯一地決定該當前區塊的該幀內預測信號。

【第8項】 如申請專利範圍第1項所述之裝置，其中，該多個幀內預測模式還包括幀內預測模式的一另一集合，其包括一DC模式中的一個或多個，根據該DC模式，該當前區塊被一DC值填充；或者一定向模式，根據該定向模式，經由沿著一預定方向將從該當前區塊的相鄰樣本的一第三集合所導出的該當前區塊旁邊的一空間樣本值分佈投影到該當前區塊中來決定該幀內預測信號。

【第9項】 如申請專利範圍第1項所述之裝置，其中，該多個幀內預測模式還包括幀內預測模式的一另一集合，據其，該當前區塊的幀內預測信號係由使用一線性函數從該當前區塊的相鄰樣本的一第三集合所決定，其中相鄰樣本的該第一集合的一基數高於相鄰樣本的該第三集合的一基數。

【第10項】 如申請專利範圍第9項所述之裝置，其被配置為從該多個幀內預測模式中選擇該一個幀內預測模式時，響應於該資料流中的一個旗標，以便從幀內預測模式的該集合或幀內預測模式的該另一集合選擇該一個幀內預測模式。

【第11項】 一種用於將一圖片逐塊編碼成一資料流的裝置，該裝置支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，該圖片的一當前區塊的該幀內預測信號

經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊一個幀內預測模式，並且

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

取決於該資料流的一第一部分以決定幀內預測模式的該集合的一排序，其與該當前區塊的一鄰域相關，以便獲得預測模式的一有序列表，以及

從幀內預測模式的該有序列表中選擇該一個幀內預測模式，以及

形成除該第一部分之外的該資料流的一第二部分，以指示幀內預測模式的該有序列表中該選擇。

【第12項】 如申請專利範圍第11項所述之裝置，其被配置為：

使用一可變長度代碼將一索引編碼到該資料流中，該索引表示該選擇。

【第13項】 一種用於將一圖片逐塊編碼成一資料流的裝置，該裝置支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，該圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊一個幀內預測模式，並且

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一秩，以便獲得幀內預測模式的一有序列表，以及

將一索引編碼到該資料流，其指到該有序列表中，

其中，該索引從該有序列表中選擇該一個幀內預測模式。

【第14項】 如申請專利範圍第13項所述之裝置，其被配置為使用一可變長度代碼從該資料流將該索引解碼到該有序列表。

【第15項】 一種用於將一圖片逐塊編碼成一資料流的裝置，該裝置支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，該圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該裝置被配置為：

從該多個幀內預測模式中選擇該當前區塊一個幀內預測模式，並且

使用該一個幀內預測模式預測該當前區塊；其中，該裝置被配置為：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一機率值，

使用該機率值將一索引熵編碼到該資料流中，其指到幀內預測模式的該集合中，該索引指向該一個幀內預測模式。

【第16項】 如申請專利範圍第11項所述之裝置，其中，相鄰樣本的該第一集合與相鄰樣本的該第二集合重合。

【第17項】 如申請專利範圍第11項所述之裝置，其中，幀內預測模式的該集合中的每一個，專屬地從該當前區塊的相鄰樣本的該第一集合中唯一地決定該當前區塊的該幀內預測信號。

【第18項】 如申請專利範圍第11項所述之裝置，其中，該多個幀內預測模式還包括幀內預測模式的一另一集合，其包括一DC模式中的一個或多個，根據該DC模式，該當前區塊被一DC值填充，或者一定向模式，根據該定向模式，經由沿著一預定方向將從當前區塊的相鄰樣本的一第三集合所導出的該當前區塊旁邊的一空間樣本值分佈投影到該當前區塊中來決定該幀內預測信號。

【第19項】 如申請專利範圍第11項所述之裝置，其中，該多個幀內預測模式還包括幀內預測模式的一另一集合，據其，該當前區塊的幀內預測信號係由使用一線性函數從該當前區塊的相鄰樣本的一第三集合所決定，其中相鄰樣本的該第一集合的一基數高於相鄰樣本的該第三集合的一基數。

【第20項】 如申請專利範圍第17項所述之裝置，其被配置為向該資料流提供一個旗標，以指示該一個幀內預測模式是幀內預測模式的該集合還是幀內預測模式的該另一集的成員。

【第21項】 一種用於從一資料流逐塊解碼一圖片的方法，其支持多個幀內預測模式，其包括幀內預測模式的至少一集合，據其，該圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

根據該資料流的一第一部分決定幀內預測模式的該集合的一排名，該第一部分與該當前區塊的一鄰域相關，以便獲得幀內預測模式的一有序列表，以及

根據除該第一部分之外的該資料流的一第二部分，從幀內預測模式的該有序列表中選擇該一個幀內預測模式。

【第22項】 一種用於從一資料流逐塊解碼一圖片的方法，其支持多個幀內預測模式，其包括幀內預測模式的至少一集合，據其，該圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一秩（rank），以便獲得幀內預測模式的一有序列表，

從該資料流解碼一索引到該有序列表中，以及

使用該索引和該有序列表執行該選擇。

【第23項】 一種用於從一資料流逐塊解碼一圖片的方法，其支持多個幀內預測模式，其包括幀內預測模式的至少一集合，據其，該圖片的一當前區塊的該幀內預測信號經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一機率值，

使用該機率值從該資料流將一索引熵解碼到幀內預測模式的該集合，以及

使用該索引執行該選擇。

【第24項】 一種用於將一圖片逐塊編碼成一資料流的方法，其支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

取決於該資料流的一第一部分以決定幀內預測模式的該集合的一排序，其與該當前區塊的一鄰域相關，以便獲得預測模式的一有序列表，以及

從幀內預測模式的該有序列表中選擇該一個幀內預測模式，以及

形成除該第一部分之外的該資料流的一第二部分，以指示幀內預測模式的該有序列表中該選擇。

【第25項】 一種用於將一圖片逐塊編碼成一資料流的方法，其支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一秩，以便獲得幀內預測模式的一有序列表，以及

將一索引編碼到該資料流，其指到該有序列表中，

其中，該索引從該有序列表中選擇該一個幀內預測模式。

【第26項】 一種用於將一圖片逐塊編碼成一資料流的方法，其支持多個幀內預測模式，該多個幀內預測模式包括幀內預測模式的至少一集合，根據該幀內預測模式，經由將該當前區塊的相鄰樣本的一第一集合應用到一神經網路上來決定用於該圖片的一當前區塊的該預測信號，該方法包括：

從該多個幀內預測模式中選擇該當前區塊的一個幀內預測模式，以及

使用該一個幀內預測模式預測該當前區塊；其中，該方法包括：

對於幀內預測模式的該集合的每一個幀內預測模式，經由將相鄰樣本的一第二集合應用到一另一神經網路上來決定一機率值，

使用該機率值將一索引熵編碼到該資料流中，其指到幀內預測模式的該集合中，該索引指向該一個幀內預測模式。

【第27項】 一種計算機程式，具有一程式代碼，用於在計算機上運行時執行根據申請專利範圍第21至26項中任一項所述的方法。

【第28項】 一種資料流產品，其包含使用根據申請專利範圍第24或26項的方法所生成的資料流。

【發明圖式】

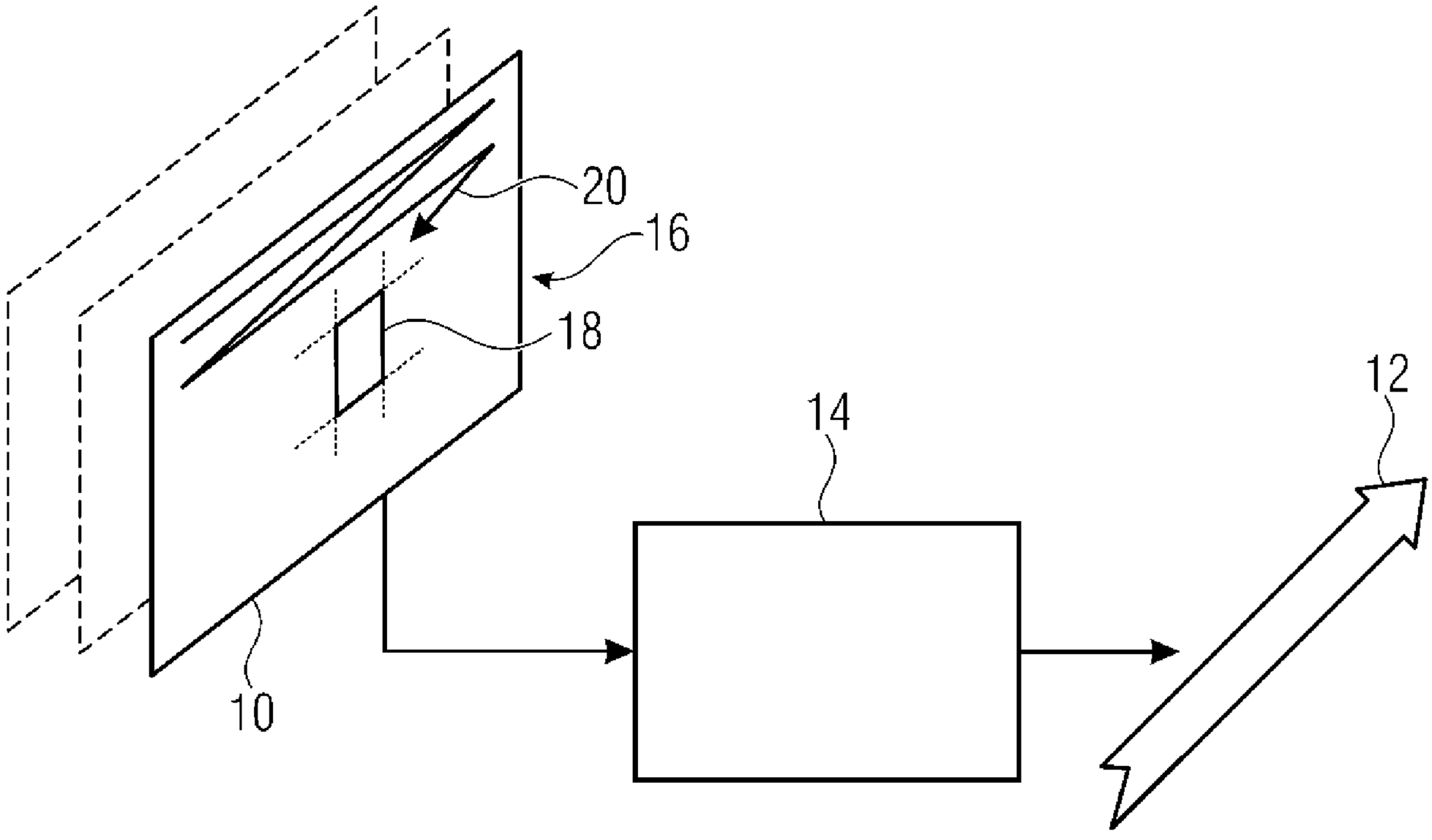
20
16
18
14
12
10

圖 1

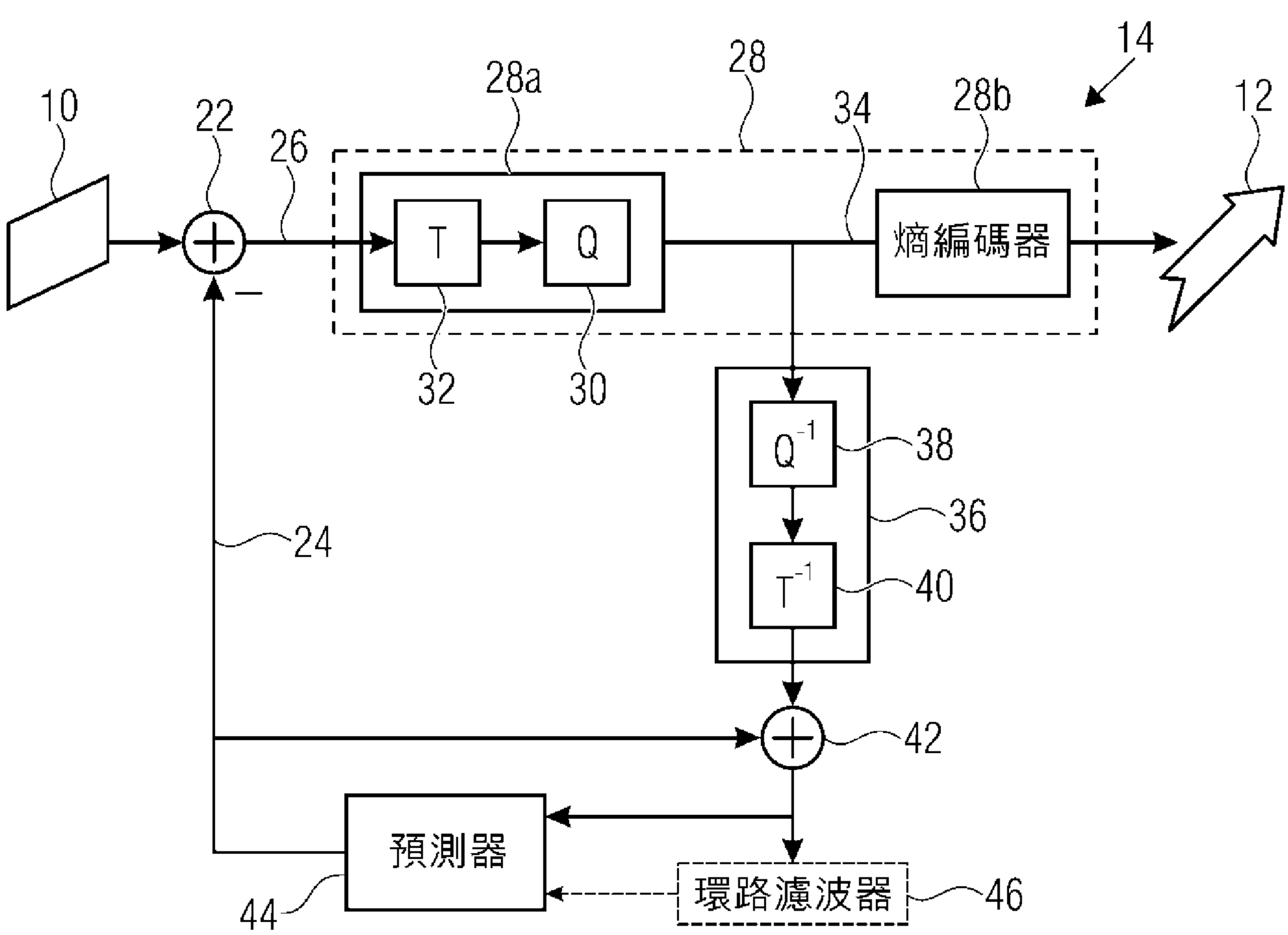
14
10
22
26
28a
28
34
28b
12
T
Q
熵編碼器
32
30
Q^{-1}
38
36
T^{-1}
40
24
42
預測器
44
環路濾波器
46

圖 2

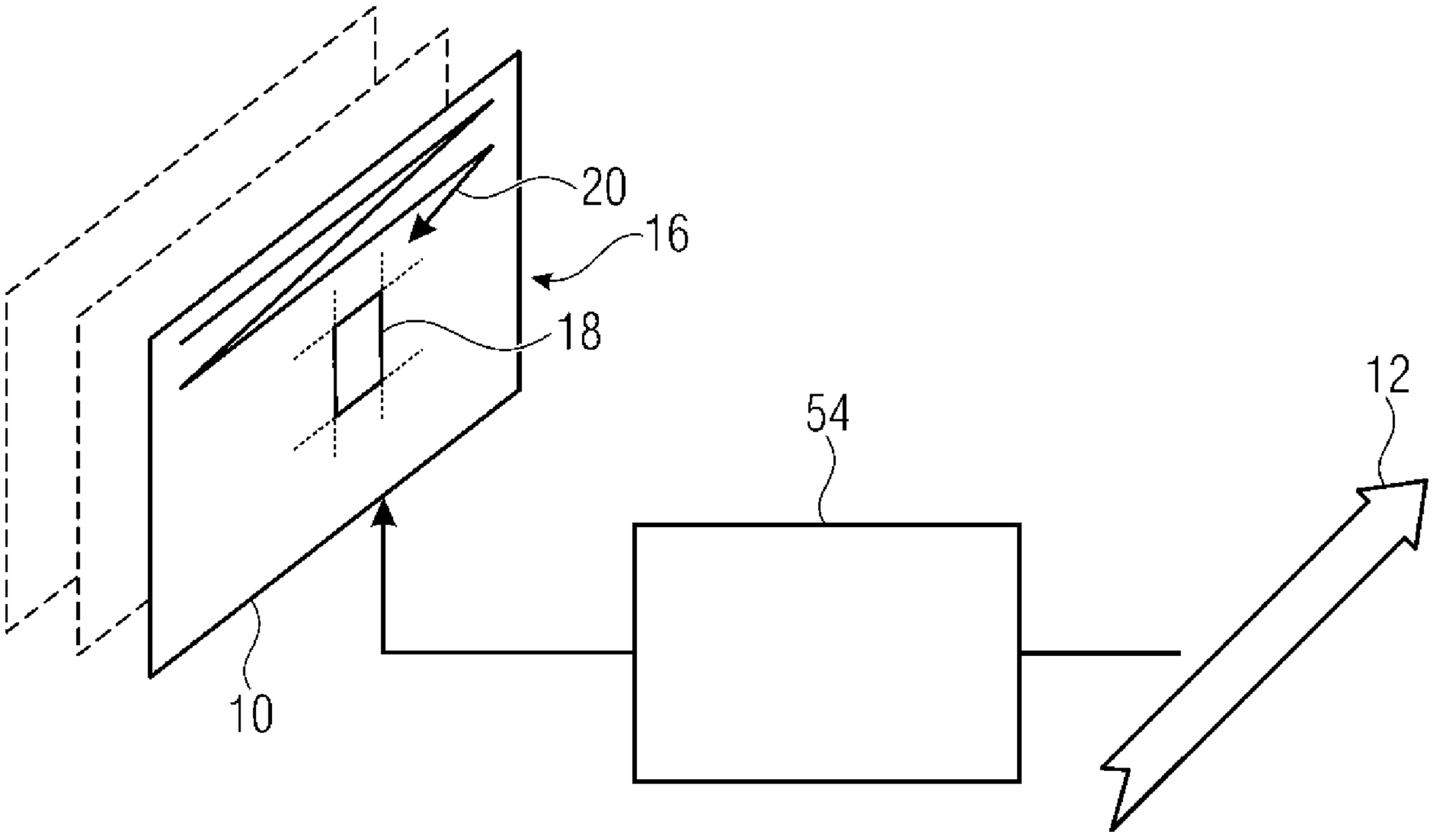
20
16
18
54
12
10

圖 3

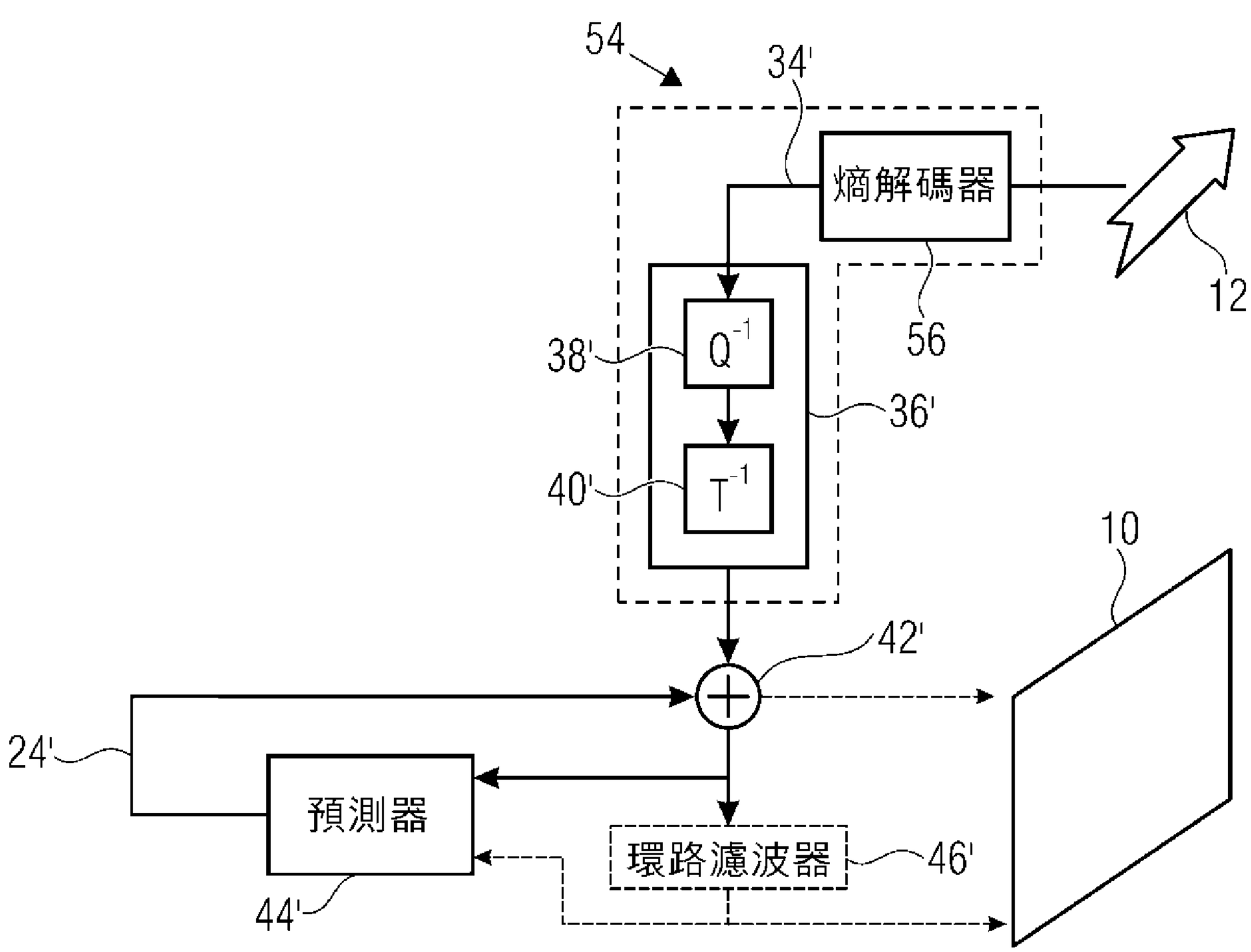
54
34'
熵解碼器
12
56
38'
Q^{-1}
36'
40'
T^{-1}
10
42'
24'
預測器
環路濾波器
46'
44'

圖 4

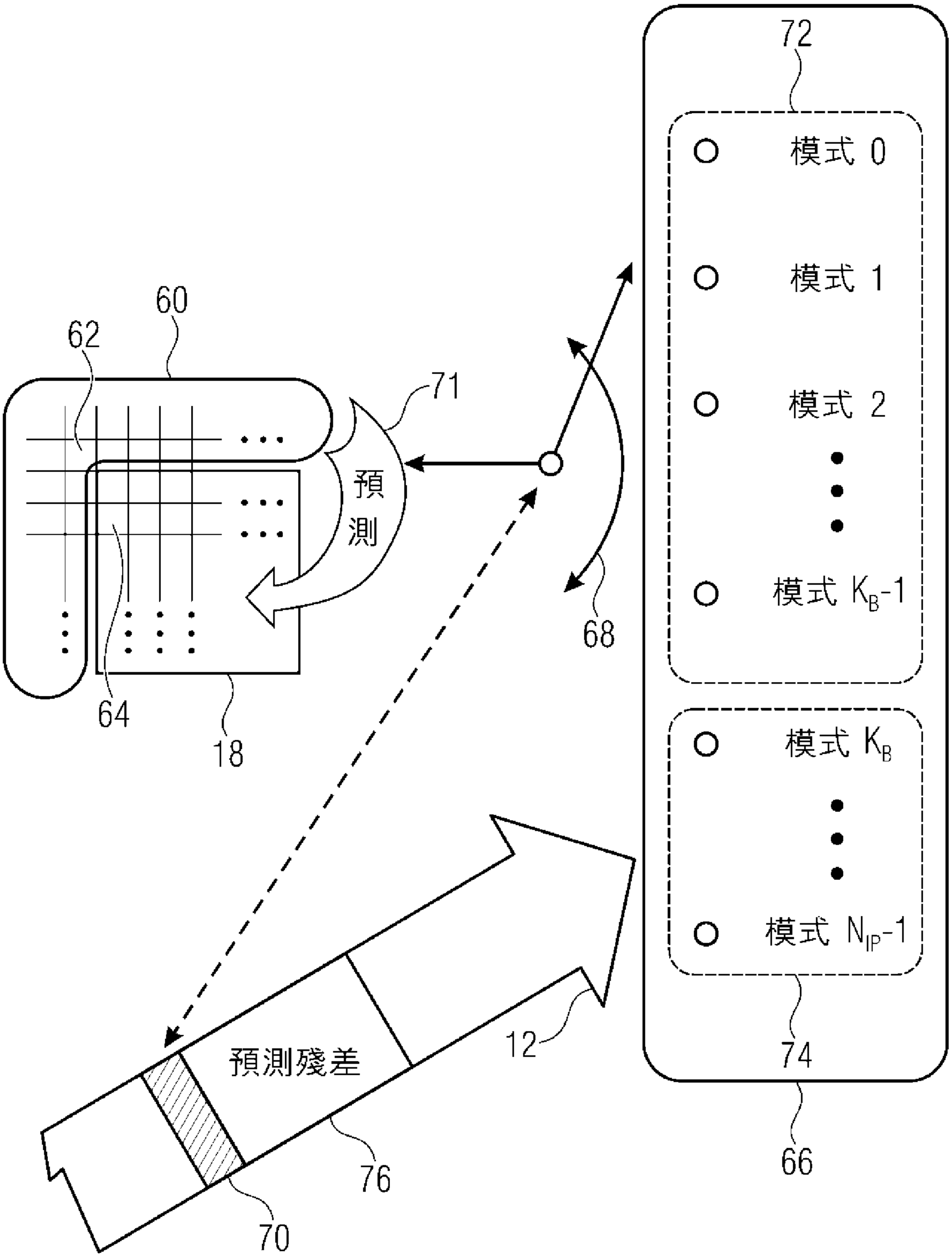
72
模式 0
模式 1
模式 2
模式 K_B-1
模式 K_B
模式 N_{IP}-1
74
66
60
62
71
預測
68
64
18
預測殘差
12
76
70

圖 5

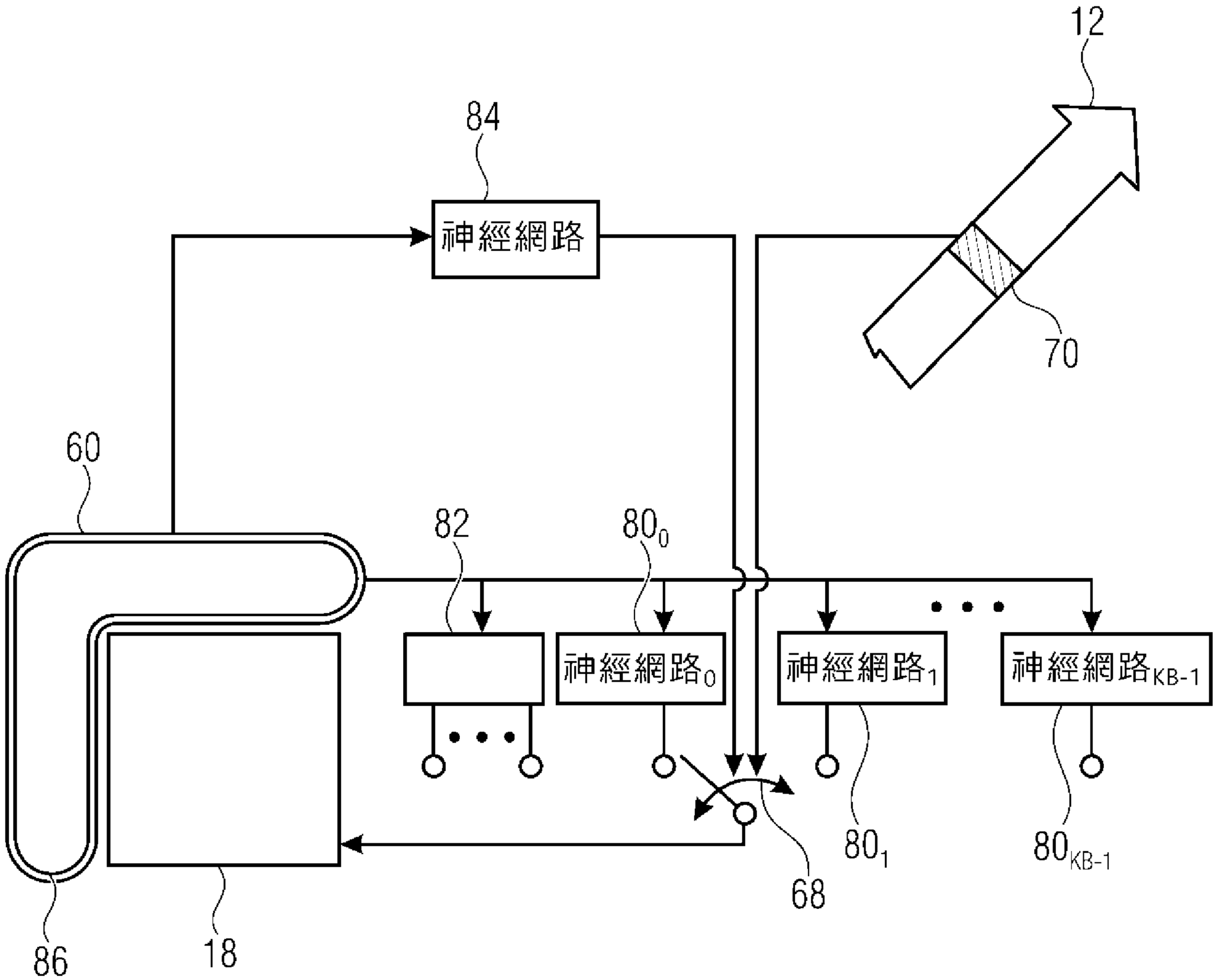

圖 6

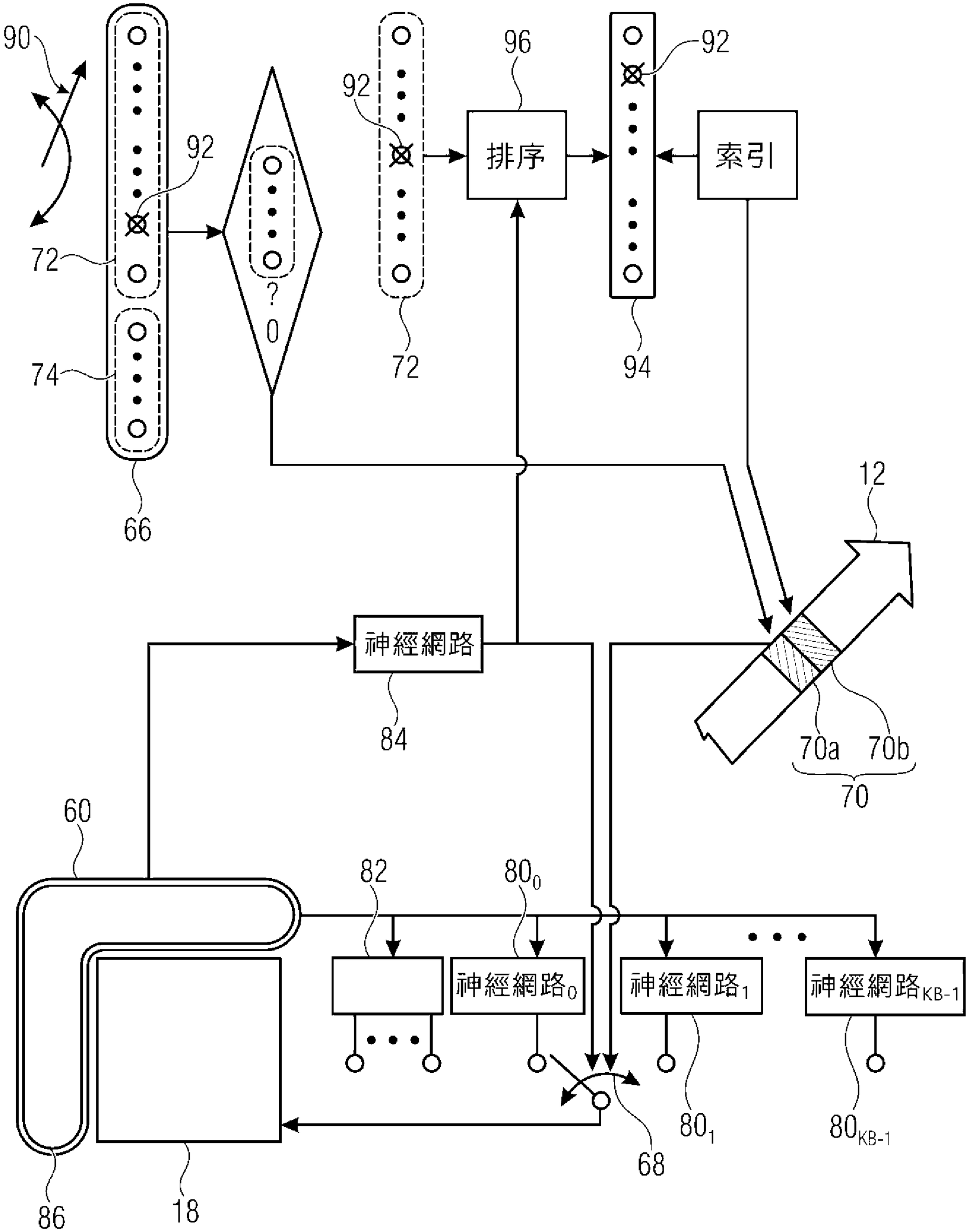
90
92
72
74
66
?
0
92
72
96
排序
92
94
索引
12
神經網路
84
70a
70b
70
60
82
80_0
神經網路$_0$
神經網路$_1$
神經網路$_{KB-1}$
80_1
80_{KB-1}
68
86
18

圖 7a

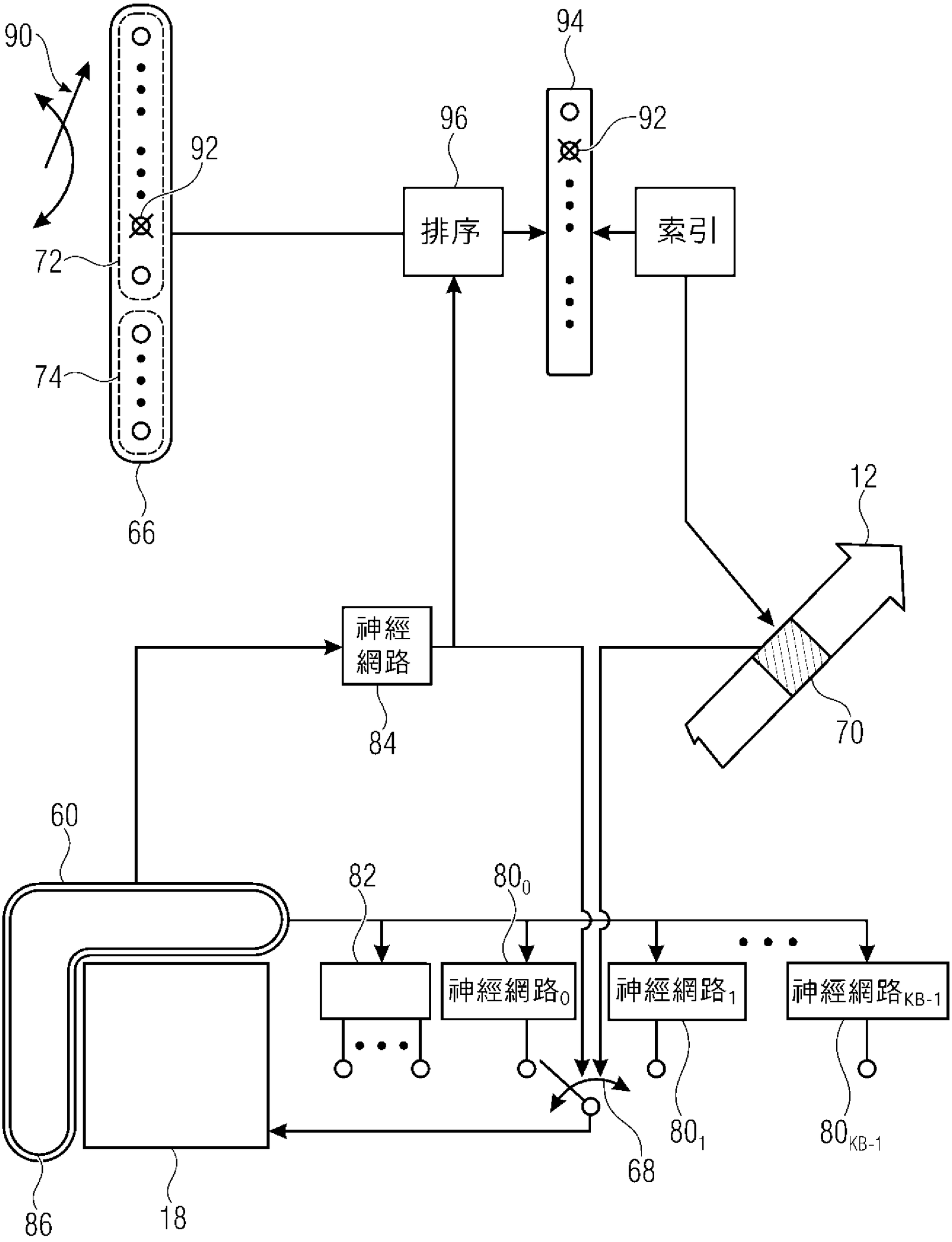
90
92
72
74
66
96
94
92
排序
索引
神經
網路
84
12
70
60
82
80₀
神經網路₀
神經網路₁
神經網路KB-1
80₁
80KB-1
68
86
18

圖 7b

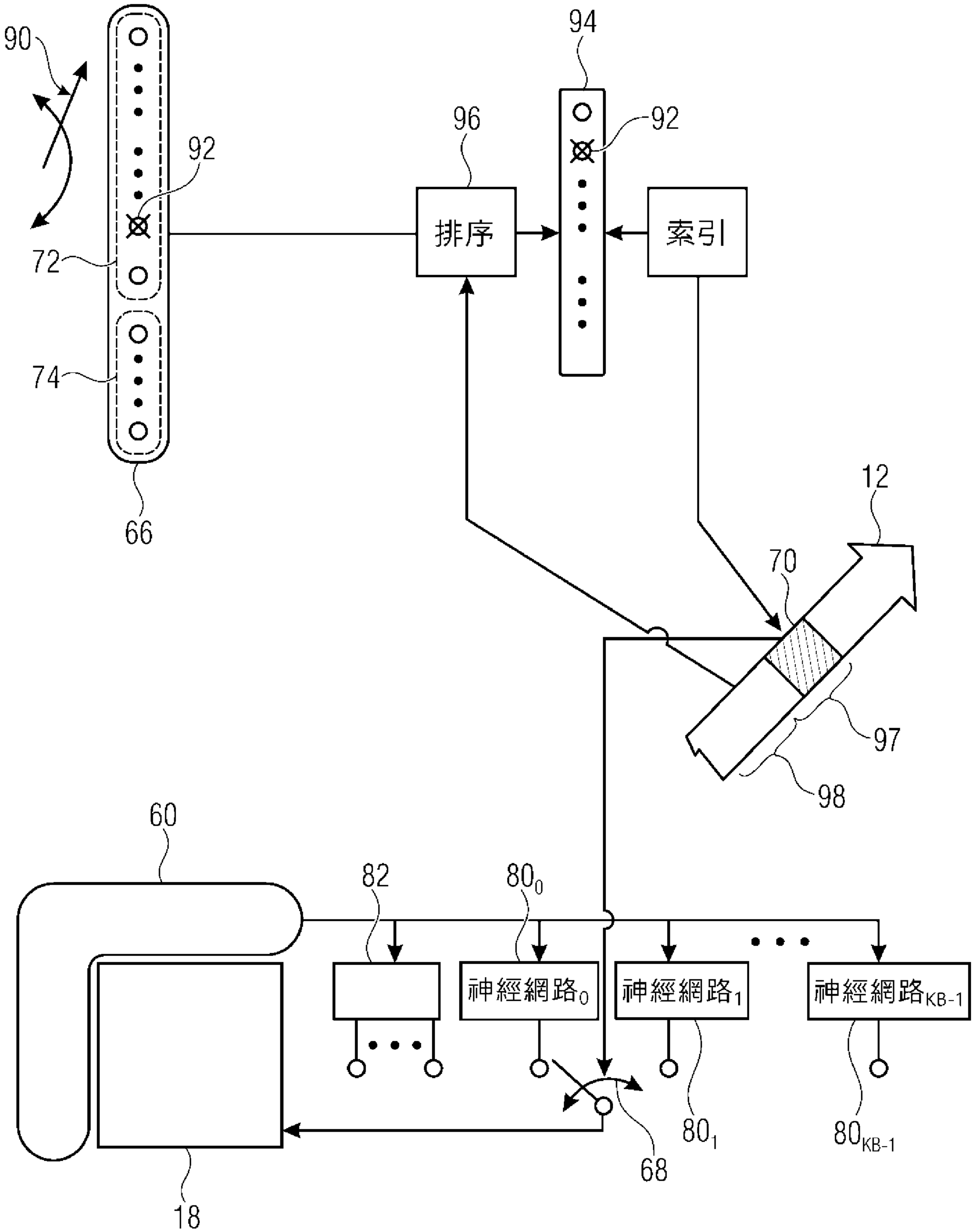
90
92
72
74
66
94
96
92
排序
索引
12
70
97
98
60
82
80$_0$
神經網路$_0$
神經網路$_1$
神經網路$_{KB-1}$
80$_1$
80$_{KB-1}$
68
18

圖 7c

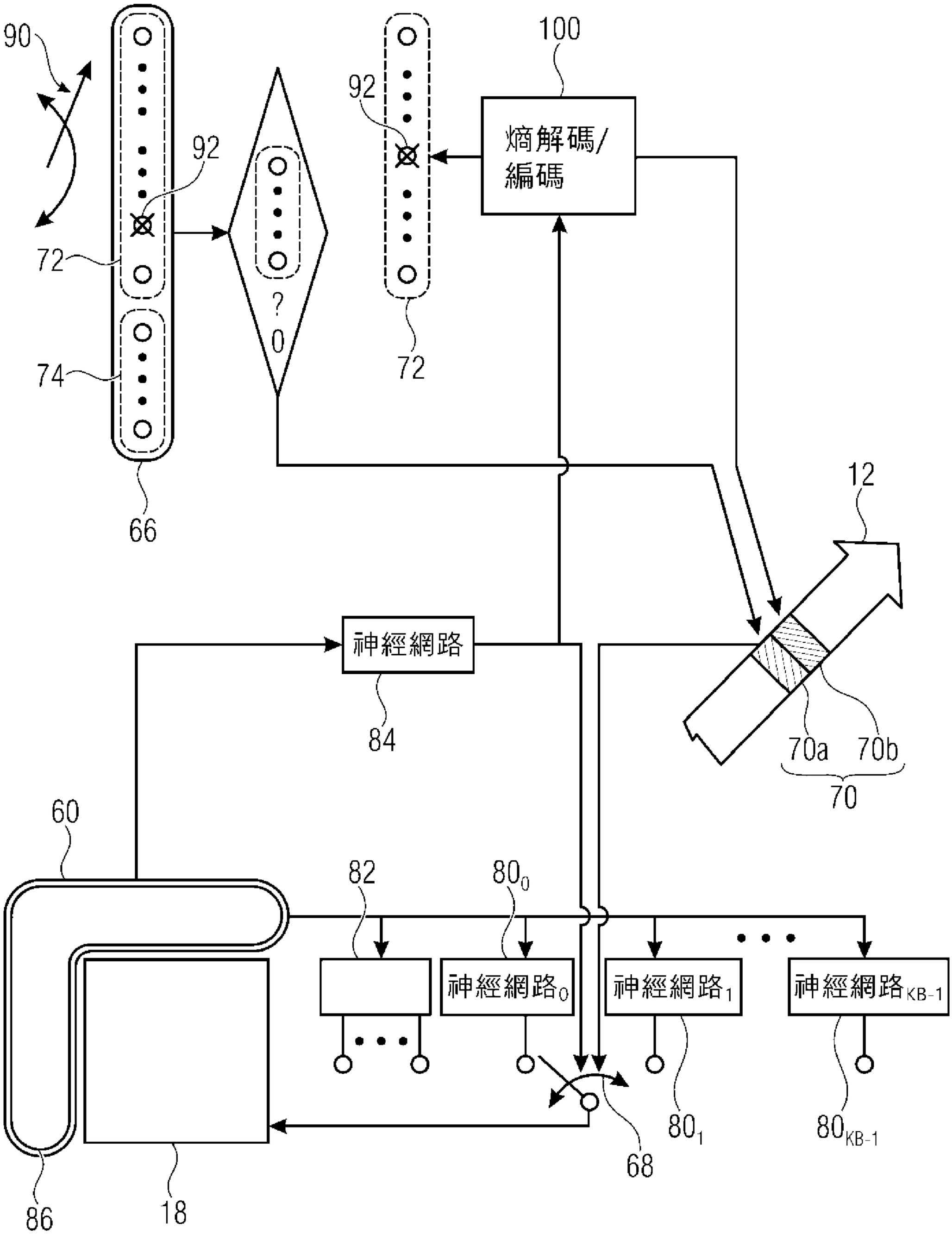
90
92
72
74
66
?
0
92
72
100
熵解碼/
編碼
12
70a
70b
70
神經網路
84
60
82
80_0
神經網路$_0$
神經網路$_1$
神經網路$_{KB-1}$
80_1
80_{KB-1}
68
86
18

圖 7d

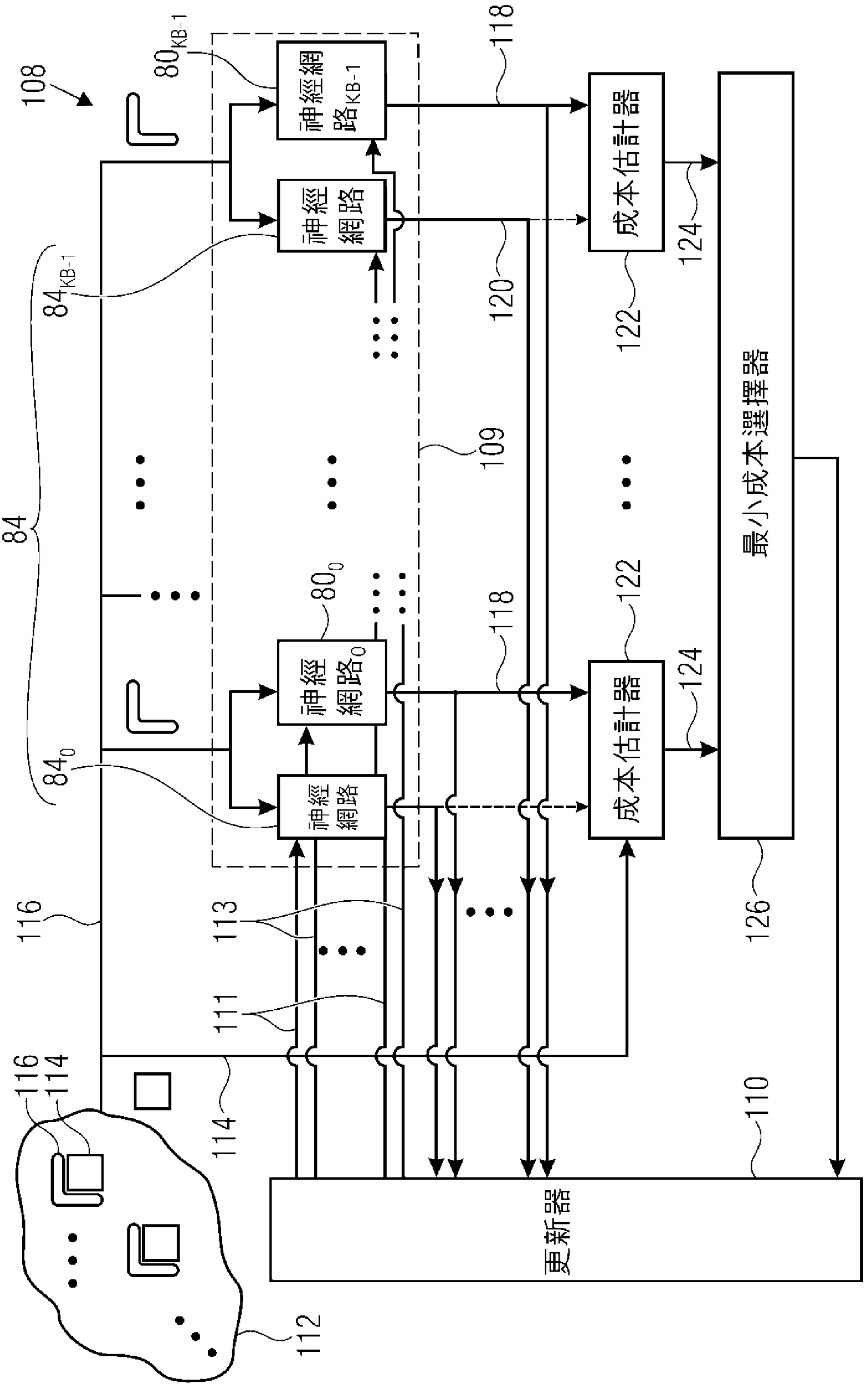
108
80_{KB-1}
神經網路$_{KB-1}$
118
84_{KB-1}
神經網路
120
成本估計器
122
124
84
80_0
109
118
122
最小成本選擇器
84_0
神經網路$_0$
神經網路
124
116
113
111
126
116
114
114
110
更新器
112

圖 8

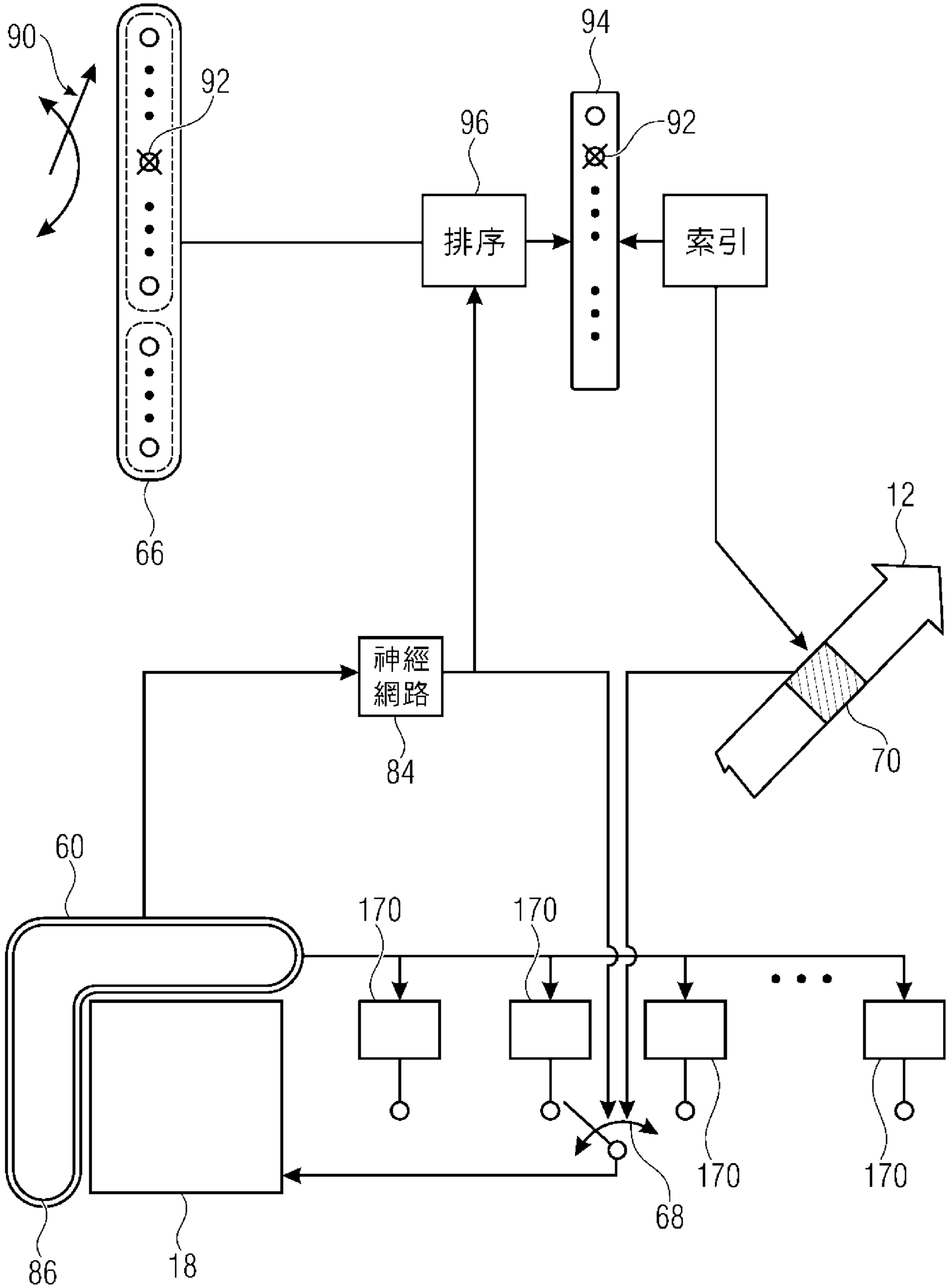

圖 9a

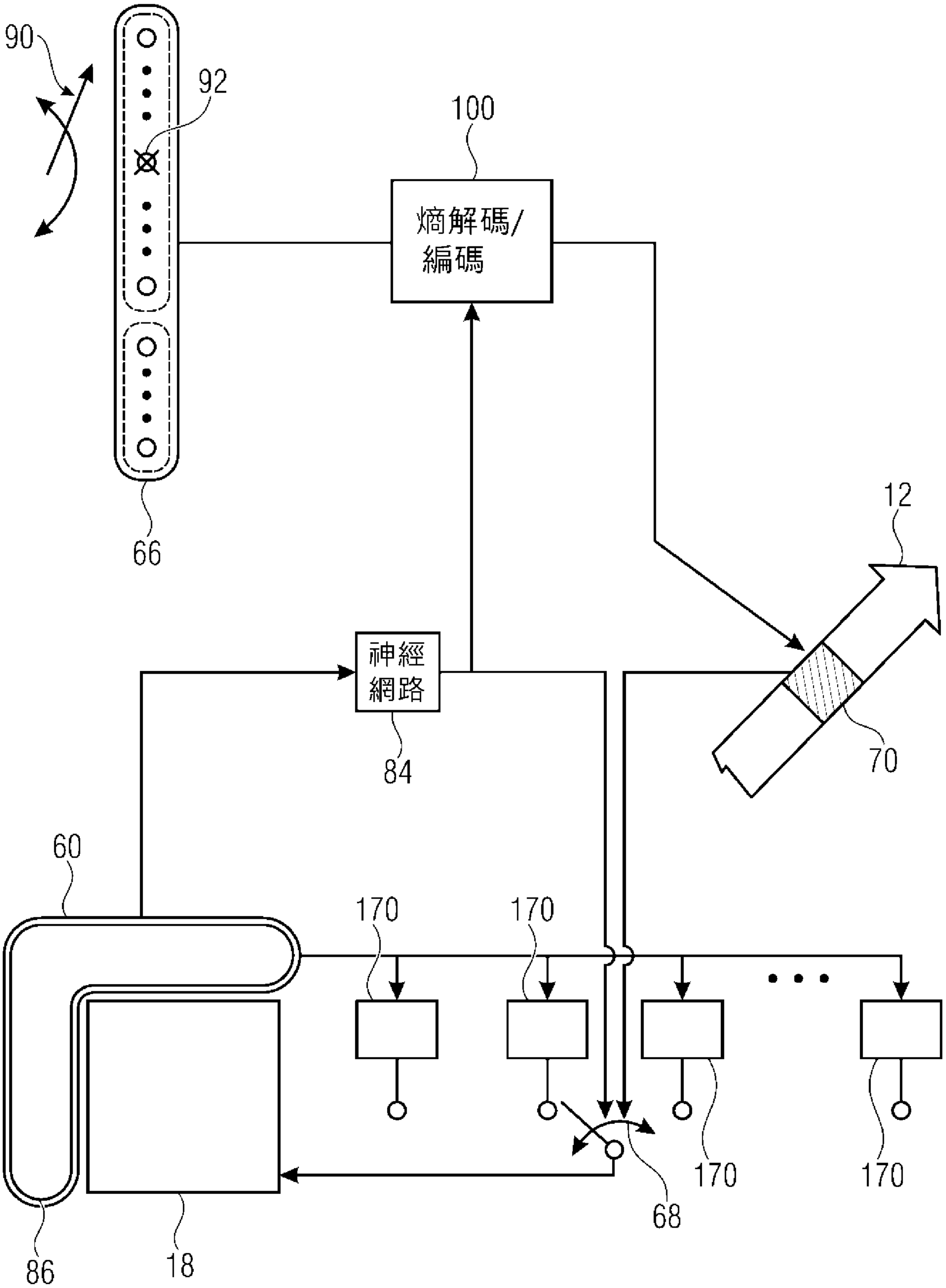

圖 9b

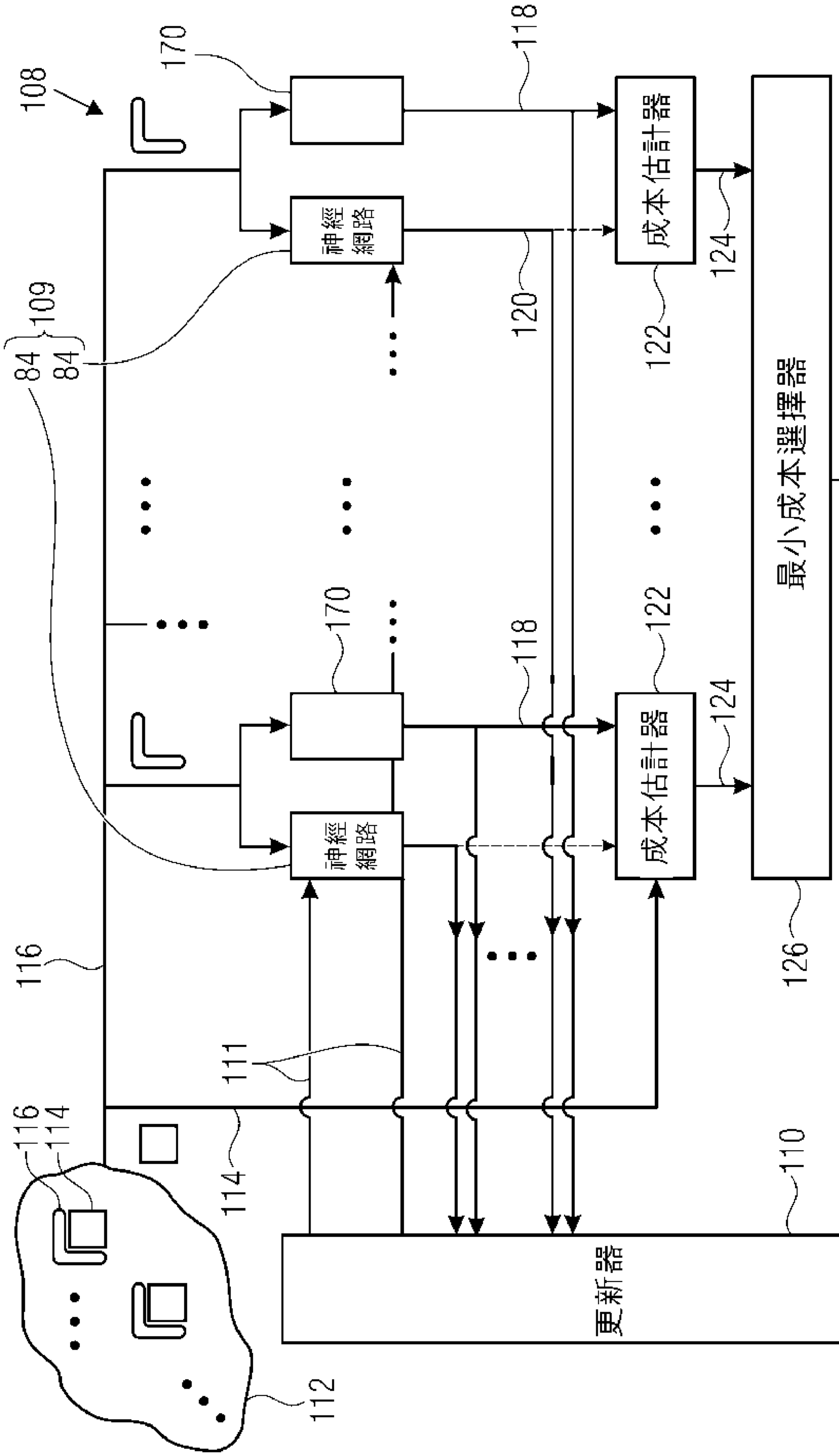
108
170
118
109
84
84
神經網路
成本估計器
120
122
124
最小成本選擇器
110
116
111
114
126
更新器
112

圖 1