



(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 I883294 B

(45)公告日：中華民國 114 (2025) 年 05 月 11 日

(21)申請案號：110146050

(22)申請日：中華民國 110 (2021) 年 12 月 09 日

(51)Int. Cl. : H04N19/124 (2014.01)

H04N19/172 (2014.01)

H04N19/186 (2014.01)

H04N19/42 (2014.01)

H04N19/91 (2014.01)

(30)優先權：2020/12/10 美國 63/124,016

2020/12/30 美國 63/131,802

2021/12/08 美國 17/643,383

(71)申請人：美商高通公司(美國) QUALCOMM INCORPORATED (US)

美國

(72)發明人：艾吉爾梅茲 希勒米艾內斯 EGILMEZ, HILMI ENES (TR)；辛格 安基泰許庫瑪

SINGH, ANKITESH KUMAR (IN)；克班 莫哈美德塞伊德 COBAN, MUHAMMED

ZEYD (US)；卡克基維克茲 瑪塔 KARCZEWICZ, MARTA (US)

(74)代理人：李世章

(56)參考文獻：

CN 108184129A CN 111861877A

US 2019/0220746A1 US 2020/0053388A1

網路文獻 Y. Ye et al., "[DNNVC] AhG on deep neural networks based video coding," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 20th Meeting, by teleconference, 7 - 16 Oct. 2020. Document: JVET-T0121 (version 1 - date 2020-10-13 11:01:46) https://jvet-experts.org/doc_end_user/current_document.php?id=10518

審查人員：賴韻曲

申請專利範圍項數：60 項 圖式數：9 共 115 頁

(54)名稱

用於基於神經網路的視訊譯碼的前端架構

(57)摘要

本文描述了用於使用神經網路系統來處理視訊資料的技術。例如，一種程序可以包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值。該程序可以包括：由編碼器子網路的第二迴旋層產生與訊框的至少一個色度通道相關聯的輸出值。該程序可以包括：由第三迴旋層基於與訊框的亮度通道相關聯的輸出值和與訊框的至少一個色度通道相關聯的輸出值來產生訊框的組合表示。該程序可以包括：基於訊框的組合表示來產生經編碼的視訊資料。

Techniques are described herein for processing video data using a neural network system. For instance, a process can include generating, by a first convolutional layer of an encoder sub-network of the neural network system, output values associated with a luminance channel of a frame. The process can include generating, by a second convolutional layer of the encoder sub-network, output values associated with at least one chrominance channel of the frame. The process can include generating, by a third convolutional

layer based on the output values associated with the luminance channel of the frame and the output values associated with the at least one chrominance channel of the frame, a combined representation of the frame. The process can further include generating encoded video data based on the combined representation of the frame.

指定代表圖：

符號簡單說明：

632:N 通道色度輸出

634:N 通道亮度輸出

638:1x1 迴旋層

639:輸出值

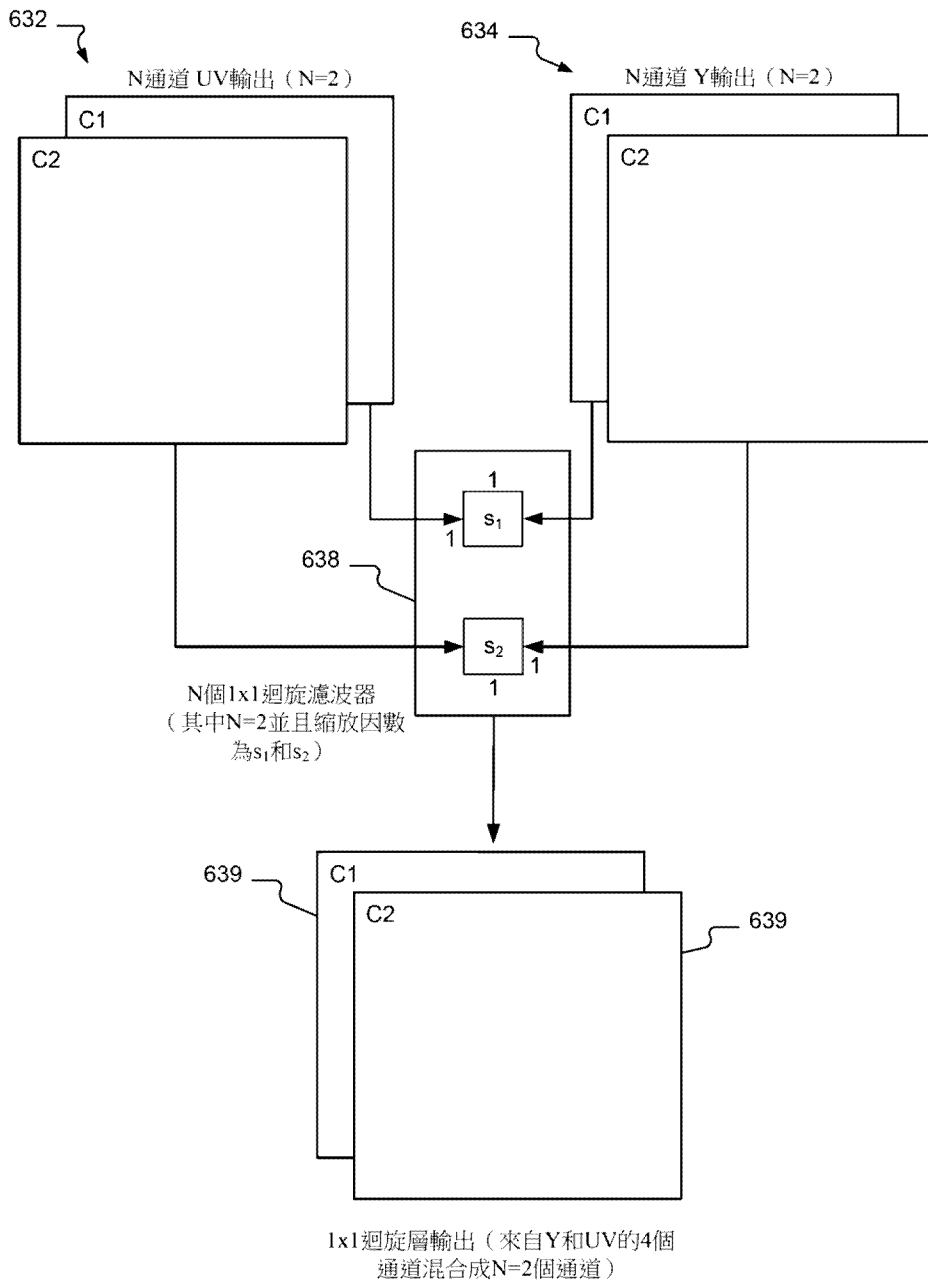


圖6B



I883294

【發明摘要】

【中文發明名稱】用於基於神經網路的視訊譯碼的前端架構

【英文發明名稱】A FRONT-END ARCHITECTURE FOR NEURAL NETWORK
BASED VIDEO CODING

【中文】

本文描述了用於使用神經網路系統來處理視訊資料的技術。例如，一種程序可以包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值。該程序可以包括：由編碼器子網路的第二迴旋層產生與訊框的至少一個色度通道相關聯的輸出值。該程序可以包括：由第三迴旋層基於與訊框的亮度通道相關聯的輸出值和與訊框的至少一個色度通道相關聯的輸出值來產生訊框的組合表示。該程序可以包括：基於訊框的組合表示來產生經編碼的視訊資料。

【英文】

Techniques are described herein for processing video data using a neural network system. For instance, a process can include generating, by a first convolutional layer of an encoder sub-network of the neural network system, output values associated with a luminance channel of a frame. The process can include generating, by a second convolutional layer of the encoder sub-network, output values associated with at least one chrominance channel of the frame. The process can include generating, by a third convolutional layer based on the output values associated with the luminance channel

of the frame and the output values associated with the at least one chrominance channel of the frame, a combined representation of the frame. The process can further include generating encoded video data based on the combined representation of the frame.

【指定代表圖】第（ 6B ）圖。

【代表圖之符號簡單說明】

6 3 2 : N 通道色度輸出

6 3 4 : N 通道亮度輸出

6 3 8 : 1 x 1 迴旋層

6 3 9 : 輸出值

【特徵化學式】

無

【發明說明書】

【中文發明名稱】用於基於神經網路的視訊譯碼的前端架構

【英文發明名稱】A FRONT-END ARCHITECTURE FOR NEURAL NETWORK
BASED VIDEO CODING

【技術領域】

【0001】 概括而言，本案內容係關於影像和視訊譯碼，包括對影像及/或視訊的編碼（或壓縮）和解碼（解壓縮）。例如，本案內容的各態樣係關於用於使用基於端到端機器學習（例如，神經網路）的影像和視訊譯碼系統來處理亮度-色度（YUV）輸入格式（例如，4:2:0 YUV輸入格式、4:4:4 YUV輸入格式、4:2:2 YUV輸入格式等）及/或其他輸入格式的技術。

【先前技術】

【0002】 許多設備和系統允許視訊資料被處理和輸出以供消費。數位視訊資料包括大量資料，以滿足消費者和視訊提供者的需求。例如，視訊資料的消費者期望高品質視訊，包括高保真度、高解析度、高畫面播放速率等。結果，滿足這些需求所需要的大量視訊資料為處理和儲存視訊資料的通訊網路和設備帶來了負擔。

【0003】 各種視訊譯碼技術可以用於對視訊資料進行壓縮。視訊譯碼的一個目標是將視訊資料壓縮為使用較低位元速率的形式，同時避免或最小化對視訊品質的降級。隨著不斷發展的視訊服務變得可用，需要具有更好的譯碼效率的編碼技術。

【發明內容】

【0004】 描述了用於使用一或多個機器學習系統來對影像及/或視訊內容進行譯碼（例如，編碼及/或解碼）的系統和技術。例如，提供了一種基於端到端機器學習（例如，神經網路）的影像和視訊譯碼（E2E-NNVC）系統，其可以處理YUV（數位域YCbCr）輸入格式（以及在一些情況下，其他輸入格式），在一些情況下，具體為4:2:0 YUV輸入格式。E2E-NNVC系統可以處理包含多個訊框的獨立訊框（亦被稱為影像或圖片）及/或視訊資料。YUV格式包括亮度通道（Y）和一對色度通道（U和V）。U和V通道可以相對於Y通道進行二次取樣，而不會對視覺品質產生顯著或明顯的影響。在YUV格式中，通道之間的相關性減小，這可能與其他色彩格式（例如，紅綠藍（RGB）格式）不同。本文描述的系統和技術的各態樣提供了前端架構（例如，新的子網路），以適應針對RGB輸入格式設計的E2E-NNVC（以及在一些情況下，針對其他輸入格式設計的E2E-NNVC）中的YUV 4:2:0輸入格式（以及在一些情況下，其他輸入格式）。前端架構適用於許多E2E-NNVC架構。

【0005】 在一個說明性實例中，提供了一種處理視訊資料的方法。該方法包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；由該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；由第三迴旋層基於與該訊框的該亮

度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生經編碼的視訊資料。

【0006】 在另一實例中，提供了一種用於處理視訊資料的裝置，其包括：記憶體；及處理器（例如，在電路中實現的），其耦合到該記憶體。在一些實例中，一個以上的處理器可以耦合到該記憶體，並且可以用於執行該等操作中的一或多個操作。該處理器被配置為：使用神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；使用該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；使用第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生經編碼的視訊資料。

【0007】 在另一實例中，提供了一種用於對視訊資料進行編碼的非暫時性電腦可讀取媒體，其具有儲存在其上的指令，該等指令在被一或多個處理器執行時使得該一或多個處理器進行以下操作：使用神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；使用該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；使用第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色

度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生經編碼的視訊資料。

【0008】 在另一實例中，提供了一種用於處理視訊資料的裝置。該裝置包括：用於經由神經網路系統的編碼器子網路的第一迴旋層來產生與訊框的亮度通道相關聯的輸出值的單元；用於經由該編碼器子網路的第二迴旋層來產生與該訊框的至少一個色度通道相關聯的輸出值的單元；用於經由使用第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示的單元；及用於基於該訊框的該組合表示來產生經編碼的視訊資料的單元。

【0009】 在一些態樣中，該第三迴旋層包括 1×1 迴旋層。該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0010】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：使用該編碼器子網路的第一非線性層來處理與該訊框的該亮度通道相關聯的該等輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值。在此類態樣中，該組合表示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0011】 在一些態樣中，該訊框的該組合表示是由該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0012】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：對該經編碼的視訊資料進行量化。

【0013】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：對該經編碼的視訊資料進行熵譯碼。

【0014】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：將該經編碼的視訊資料儲存在記憶體中。

【0015】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：在傳輸媒體上向至少一個設備發送該經編碼的視訊資料。

【0016】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：獲得經編碼的訊框；由該神經網路系統的解碼器子網路的第一迴旋層產生與經編碼的訊框的亮度通道相關聯的經重構的輸出值；及由該解碼器子網路的第二迴旋層產生與經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。

【0017】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：使用該解碼器子網路的第三迴旋層來將經編碼的訊框的該亮度通道與經編碼的訊框的該至少一個色度通道分離。

【0018】 在一些態樣中，該解碼器子網路的該第三迴旋層包括 1×1 迴旋層。該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0019】 在一些態樣中，該訊框包括視訊訊框。在一些態樣中，該至少一個色度通道包括色度藍色通道和色度紅色通道。在一些態樣中，該訊框具有亮度色度 (YUV) 格式。

【0020】 在一個說明性實例中，提供了一種處理視訊資料的方法。該方法包括：獲得經編碼的訊框；由該解碼器子網路的第一迴旋層將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離；由神經網路系統的解碼器子網路的第二迴旋層產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；由該解碼器子網路的第三迴旋層產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及產生輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【0021】 在另一實例中，提供了一種用於處理視訊資料的裝置，其包括：記憶體；及處理器（例如，在電路中實現的），其耦合到該記憶體。在一些實例中，一個以上的處理器可以耦合到該記憶體，並且可以用於執行該等操作中的一或多個操作。該處理器被配置為：獲得經編碼的訊框；使用該解碼器子網路的第一迴旋層來將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離；使用神經網路系統的解碼器子網路的第二迴旋層產生與該經

編碼的訊框的該亮度通道相關聯的經重構的輸出值；使用該解碼器子網路的第三迴旋層產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及產生輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【0022】 在另一實例中，提供了一種用於對視訊資料進行編碼的非暫時性電腦可讀取媒體，其具有儲存在其上的指令，該等指令在被一或多個處理器執行時使得該一或多個處理器進行以下操作：獲得經編碼的訊框；使用該解碼器子網路的第一迴旋層來將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離；使用神經網路系統的解碼器子網路的第二迴旋層產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；使用該解碼器子網路的第三迴旋層產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及產生輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【0023】 在另一實例中，提供了一種用於處理視訊資料的裝置。該裝置包括：用於獲得經編碼的訊框的單元；用於經由該解碼器子網路的第一迴旋層來將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離的單元；用於經由神經網路系統的解碼器子網路的第二迴旋層產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值的單元；用於經由該解碼器子網路的第三迴旋層產生

與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值的單元；及用於產生輸出訊框的單元，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值

【0024】 在一些態樣中，該解碼器子網路的該第一迴旋層包括 1×1 迴旋層。該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0025】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：使用該解碼器子網路的第一非線性層來處理與該經編碼的訊框的亮度通道相關聯的值，其中與該亮度通道相關聯的該等經重構的輸出值是基於該第一非線性層的輸出來產生的；及使用該解碼器子網路的第二非線性層來處理與該經編碼的訊框的該至少一個色度通道相關聯的值，其中與該至少一個色度通道相關聯的該等經重構的輸出值是基於該第二非線性層的輸出來產生的。

【0026】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：對該經編碼的訊框的取樣進行去量化。

【0027】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：對該經編碼的訊框的取樣進行熵解碼。

【0028】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：將該輸出訊框儲存在記憶體中。

【0029】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：顯示該輸出訊框。

【0030】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：由該神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；由該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；由該編碼器子網路的第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生該經編碼的訊框。

【0031】 在一些態樣中，該編碼器子網路的該第三迴旋層包括 1×1 迴旋層。該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0032】 在一些態樣中，上述用於處理視訊資料的方法、裝置和電腦可讀取媒體亦包括：使用該編碼器子網路的第一非線性層來處理與該訊框的亮度通道相關聯的輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；其中該組合表示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0033】 在一些態樣中，該訊框的該組合表示是由該編碼器子網路的該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0034】 在一些態樣中，該經編碼的訊框包括經編碼的視訊訊框。

【0035】 在一些態樣中，該至少一個色度通道包括色度藍色通道和色度紅色通道。

【0036】 在一些態樣中，該經編碼的訊框具有亮度 - 色度 (YUV) 格式。

【0037】 在一些態樣中，該裝置可以是以下各項或可以是以下各項的一部分：行動設備（例如，行動電話或所謂的「智慧型電話」、平板電腦或其他類型的行動設備）、網路連接的可穿戴設備、擴展現實設備（例如，虛擬實境 (VR) 設備、增強現實 (AR) 設備或混合現實 (MR) 設備）、個人電腦、膝上型電腦、伺服器電腦（例如，視訊伺服器或其他伺服器設備）、電視、車輛（或車輛的計算設備或系統）、照相機（例如，數位照相機、網際網路協定 (IP) 照相機等）、多照相機系統、機器人設備或系統、航空設備或系統、或其他設備。在一些態樣中，該裝置亦包括用於擷取一或多個影像或視訊訊框（或圖片）的至少一個照相機。例如，該裝置可以包括用於擷取包括視訊訊框的一或多個影像及 / 或一或多個視訊的一個照相機（例如，RGB 照相機）或多個照相機。在一些態樣中，該裝置包括用於顯示一或多個影像、視訊、通知或其他可顯示資料的顯示

器。在一些態樣中，該裝置包括發射器，其被配置為在傳輸媒體上向至少一個設備發送一或多個視訊訊框及/或語法資料。在一些態樣中，上述裝置可以包括一或多個感測器。在一些態樣中，處理器包括神經處理單元（NPU）、中央處理單元（CPU）、圖形處理單元（GPU）或其他處理設備或部件。

【0038】 該發明內容既不意欲標識所要求保護的主題的關鍵或必要特徵，亦不意欲單獨用於決定所要求保護的主題的範疇。經由參照本專利的整個說明書的適當部分、任何或所有附圖以及每個請求項，應當理解該主題。

【0039】 在參考以下說明書、請求項和附圖之後，前述內容以及其他特徵和實施例將變得更加顯而易見。

【圖式簡單說明】

【0040】 下文參照以下附圖來詳細地描述本案的說明性實施例：

【0041】 圖 1 示出片上系統（SOC）的實例實現方式；

【0042】 圖 2 A 示出全連接神經網路的實例；

【0043】 圖 2 B 示出局部連接神經網路的實例；

【0044】 圖 2 C 示出迴旋神經網路的實例；

【0045】 圖 2 D 示出被設計為從影像中辨識視覺特徵的深度迴旋網路（DCN）的詳細實例；

【0046】 圖 3 是示出深度迴旋網路（DCN）的方塊圖；

【0047】 圖 4 是示出根據一些實例的系統的實例的示意圖，該系統包括可操作以使用基於神經網路的系統來執行影像及 / 或視訊譯碼（編碼和解碼）的設備；

【0048】 圖 5 是示出根據一些實例的用於具有紅 - 綠 - 藍（RGB）格式的輸入的基於端到端神經網路的影像和視訊譯碼系統的實例的示意圖；

【0049】 圖 6 A 是示出根據一些實例的前端神經網路架構的實例的示意圖，該前端神經網路架構可以是基於端到端神經網路的影像和視訊解碼系統的一部分；

【0050】 圖 6 B 是示出根據一些實例的 1 x 1 迴旋層的實例操作的示意圖；

【0051】 圖 6 C 是示出根據一些實例的前端神經網路架構的另一實例的示意圖，該前端神經網路架構可以是基於端到端神經網路的影像和視訊譯碼系統的一部分；

【0052】 圖 6 D 是示出根據一些實例的前端神經網路架構的另一實例的示意圖，該前端神經網路架構可以是基於端到端神經網路的影像和視訊譯碼系統的一部分；

【0053】 圖 6 E 是示出根據一些實例的前端神經網路架構的另一實例的示意圖，該前端神經網路架構可以是基於端到端神經網路的影像和視訊譯碼系統的一部分；

【0054】 圖 6 F 是示出根據一些實例的前端神經網路架構的另一實例的示意圖，該前端神經網路架構可以是基於端到端神經網路的影像和視訊譯碼系統的一部分；

【0055】 圖 7 是示出根據一些實例的用於處理視訊資料的程序的實例的流程圖；

【0056】 圖 8 是示出根據一些實例的用於處理視訊資料的程序的另一實例的流程圖；及

【0057】 圖 9 示出可以實現本文描述的各種技術的實例計算設備的實例計算設備架構。

【實施方式】

【0058】 下文提供了本案內容的某些態樣和實施例。如對於本發明所屬領域中具有通常知識者將顯而易見的，這些態樣和實施例中的一些項可以獨立地應用，並且其中的一些項可以相結合地應用。在以下描述中，出於解釋的目的，闡述了具體細節以便提供對本案的實施例的透徹理解。然而，將顯而易見的是，可以在沒有這些具體細節的情況下實施各個實施例。附圖和描述不意欲是限制性的。

【0059】 隨後的描述僅提供了實例實施例，並且不意欲限制本案內容的範疇、適用性或配置。確切而言，對實例實施例的隨後描述將向本發明所屬領域中具有通常知識者提供用於實現實例實施例的可行描述。應當理解的是，在不背離如在所附的申請專利範圍中闡述的本案的精神和範疇的情況下，可以對元素的功能和佈置進行各種改變。

【0060】 數位視訊資料可以包括大量資料，尤其是隨著對高品質視訊資料的需求持續增長。例如，視訊資料的消費者通常期望越來越高品質的、具有高保真度、高解析度、高畫面播放速率等的視訊。然而，滿足此類需求所需要的

大量視訊資料可能為通訊網路以及處理和儲存視訊資料的設備帶來了顯著負擔。

【0061】 各種技術可以用於對視訊資料進行譯碼。可以根據特定的視訊譯碼標準來執行視訊譯碼。實例視訊譯碼標準包括高效率視訊譯碼（H E V C）、改進的視訊譯碼（A V C）、運動影像專家組（M P E G）譯碼、以及多功能視訊譯碼（V V C）等。視訊譯碼通常使用預測方法（諸如訊框間預測或訊框內預測），預測方法利用在視訊影像或序列中存在的冗餘。視訊解碼技術的一個共同目標是將視訊資料壓縮為使用較低位元速率的形式，同時避免或最小化視訊品質的降級。隨著對視訊服務的需求增長以及新的視訊服務變得可用，需要具有更好的譯碼效率、效能和速率控制的譯碼技術。

【0062】 本文描述了用於使用一或多個機器學習（M L）系統來執行影像及/或視訊譯碼的系統、裝置、程序（亦被稱為方法）和電腦可讀取媒體（統稱為「系統和技術」）。通常，M L是人工智慧（A I）的一子集。M L系統可以包括演算法和統計模型，電腦系統可以使用這些演算法和統計模型，以在無需使用顯式指令的情況下經由依賴於模式和推理而執行各種任務。M L系統的一個實例是神經網路（亦被稱為人工神經網路），其可以包括一組相互連接的人工神經元（例如，神經元模型）。神經網路可以用於各種應用及/或設備，諸如影像及/或視訊譯碼、影像分析及/或電

腦視覺應用、網際網路協定 (IP) 照相機、物聯網路 (IoT) 設備、自主車輛、服務機器人等。

【0063】 神經網路中的各個節點可以經由採用輸入資料並且對資料執行簡單運算來類比生物神經元。對輸入資料執行的簡單運算的結果被選擇性地傳遞給其他神經元。權重值與網路之每一者向量和節點相關聯，並且這些值約束輸入資料如何與輸出資料相關。例如，可以將每個節點的輸入資料乘以相應的權重值，並且可以對乘積求和。可以經由可選的偏置來調整乘積的總和，並且可以將啟動函數應用於結果，從而產生節點的輸出訊號或「輸出啟動」（有時被稱為啟動圖或特徵圖）。權重值最初可以由經由網路的訓練資料的反覆運算流來決定（例如，在訓練階段期間建立權重值，在訓練階段中，網路學習如何經由其典型的輸入資料特性來辨識特定類別）。

【0064】 存在不同類型的神經網路，諸如迴旋神經網路 (CNN)、遞迴神經網路 (RNN)、產生性對抗網路 (GAN)、多層感知器 (MLP) 神經網路等。例如，迴旋神經網路 (CNN) 是一種類型的前饋人工神經網路。迴旋神經網路可以包括人工神經元的集合，每個人工神經元具有感受野（例如，輸入空間的空間局部化區域）並且共同平鋪輸入空間。RNN的工作原理是保存層的輸出，並且將該輸出回饋回輸入，以幫助預測層的結果。GAN是一種形式的產生性神經網路，其可以學習輸入資料中的模式，使得神經網路模型可以產生新的合成輸出，這些合成輸出可

能合理地來自原始資料集。GAN可以包括一起操作的兩個神經網路，包括產生合成輸出的產生性神經網路和評估輸出真實性的判別性神經網路。在MLP神經網路中，可以將資料饋送給輸入層，並且一或多個隱藏層為資料提供抽象層次。隨後，可以基於經抽象的資料在輸出層上進行預測。

【0065】 在分層神經網路架構（當存在多個隱藏層時，其被稱為深度神經網路）中，第一層人工神經元的輸出成為第二層人工神經元的輸入，第二層人工神經元的輸出成為第三層人工神經元的輸入，依此類推。例如，可以對CNN進行訓練以辨識特徵的層次結構。CNN架構中的計算可以分佈在可以被配置在一或多個計算鏈中的一群處理節點上。這些多層架構可以一次訓練一個層，並且可以使用反向傳播進行微調。

【0066】 在一些態樣中，本文描述的系統和技術包括被設計用於處理具有亮度-色度（YUV）輸入格式的輸入資料的基於端到端ML（例如，使用神經網路架構）的影像和視訊譯碼（E2E-NNVC）系統。YUV格式包括亮度通道（Y）和一對色度通道（U和V）。U通道可以被稱為色度（或色度）-藍色通道，並且V通道可以被稱為色度（或色度）-紅色通道。在一些情況下，亮度（Y）通道或分量亦可以被稱為亮度通道或分量。在一些情況下，色度（U和V）通道或分量亦可以被稱為色度通道或分量。YUV輸入格式可以包括YUV 4:2:0、YUV 4:4:4、YUV 4:2:2等。在一些情況下，本文描述的系統和技術可以被設計為處理其他

輸入格式，諸如具有 Y - 色度藍色 (C b) - 色度紅色 (C r) (Y C b C r) 格式、紅 - 綠 - 藍 (R G B) 格式及 / 或其他格式的資料。本文描述的 E 2 E - N N V C 系統可以對包括多個訊框的獨立訊框 (亦被稱為影像或圖片) 及 / 或視訊資料進行編碼及 / 或解碼。

【0067】 在許多情況下，E 2 E - N N V C 系統被設計為負責學習用於熵譯碼 (解碼器子網路) 的經量化的隱變數 (l a t e n t) 的概率模型的自動編碼器子網路 (編碼器子網路) 和第二子網路 (在一些情況下，亦被稱為超先驗網路) 的組合。在一些情況下，可能存在解碼器的其他子網路。此類 E 2 E - N N V C 系統架構可以被視為變換加量化模組 (或編碼器子網路) 和熵建模子網路模組的組合。

【0068】 大多數 E 2 E - N N V C 系統架構被設計為以非二次取樣輸入格式操作，諸如 R G B 、 Y U V 4 : 4 : 4 或其他非二次取樣輸入格式。然而，視訊譯碼標準 (諸如 H E V C 和 V V C) 被設計為在其各自的主簡檔中支援 Y U V 4 : 2 : 0 色彩格式。為了支援 4 : 2 : 0 Y U V 格式，必須修改被設計為以非二次取樣輸入格式操作的 E 2 E - N N V C 架構。

【0069】 本文描述的系統和技術提供了用於處理一或多個適用於現有 E 2 E - N N V C 架構的特定色彩格式 (例如，Y U V 4 : 2 : 0 色彩格式) 的前端架構 (例如，子網路) 。系統和技術考慮了 Y 和 U V 通道的不同特性以及解析度的差異。例如，訊框或訊框的一部分的 Y 和 U V 通道可以輸入到神經網路系統的編碼器子網路的兩個單獨的神經網路層。在一些

實例中，兩個神經網路層包括迴旋層。在一些態樣中，兩個單獨的神經網路層的輸出由編碼器子網路的一對非線性層或運算元處理。這對非線性層或運算元可以包括廣義除數正規化（GDN）層或運算元、參數校正線性單元（PReLU）層或運算元及/或其他非線性層或運算元。使用編碼器子網路的額外的神經網路層來對兩個獨立神經網路層的輸出（或非線性層或運算元的輸出）進行組合。

【0070】 在一些實例中，額外的神經網路層是 1×1 迴旋層。 1×1 迴旋層執行Y分量和UV分量的每圖元或每值交叉通道混合（例如，經由產生線性組合），從而產生改進譯碼效能的交叉分量（例如，交叉亮度和色度分量）預測。例如，Y分量和UV分量的交叉通道混合使Y分量與U和V分量去相關，這導致改進的譯碼效能（例如，提高譯碼效率）。在一些情況下， 1×1 迴旋層可以包括N個 1×1 迴旋濾波器（其中N等於與輸入到 1×1 迴旋層的通道數量相對應的整數值）。每個 1×1 迴旋濾波器具有各自的縮放因數，該縮放因數應用於Y分量的對應的第n通道和UV分量的對應的第n通道。

【0071】 額外的神經網路層（例如， 1×1 迴旋層）的輸出可以由編碼器子網路的一或多個非線性層及/或一或多個額外的神經網路層（例如，迴旋層）處理。量化引擎可以對由編碼器子網路的最終神經網路層輸出的特徵執行量化，以產生經量化的輸出。熵編碼引擎可以對來自量化引擎的經量化的輸出進行熵編碼以產生位元串流。神經網路

系統可以輸出位元串流以用於儲存、用於傳輸到另一設備、伺服器設備或系統等。

【0072】 神經網路系統的解碼器子網路或另一神經網路系統（另一設備）的解碼器子網路可以解碼位元串流。例如，解碼器子網路的熵解碼引擎可以對位元串流進行熵解碼，並且將經熵解碼的資料輸出到去量化引擎。去量化引擎可以對資料進行去量化。經去量化的資料可以由解碼器子網路的一或多個神經網路層（例如，迴旋層）及/或一或多個逆非線性層處理。例如，在被一或多個迴旋層和一或多個逆非線性層處理之後， 1×1 迴旋層可以處理資料。 1×1 迴旋層可以將資料劃分為Y通道特徵和組合UV通道特徵。Y通道特徵和組合UV通道特徵可以由兩個最終神經網路層（例如，兩個迴旋層）處理，以及在一些情況下由兩個最終逆非線性層處理。例如，第一最終神經網路層可以處理Y通道特徵並且輸出經重構的訊框的每個圖元或取樣的經重構的Y通道（例如，亮度取樣或圖元）。第二最終神經網路層可以處理組合UV通道特徵，並且可以輸出經重構的訊框的每個圖元或取樣的經重構的U通道（例如，色度藍色取樣或圖元）和經重構的訊框的每個圖元或取樣的經重構的V通道（例如，色度紅色取樣或圖元）。

【0073】 將參照附圖描述關於系統和技術的額外的細節。

【0074】 圖1示出片上系統（SOC）100的實例實現方式，SOC 100可以包括被配置為執行本文描述的一或多個功能的中央處理單元（CPU）102或多核心CPU。參數或變數

(例如，神經訊號和突觸權重)、與計算設備相關聯的系統參數(例如，具有權重的神經網路)、延遲、頻段資訊、任務資訊以及其他資訊可以被儲存在與神經處理單元(NPU) 108相關聯的記憶體塊中，被儲存在與CPU 102相關聯的記憶體塊中，被儲存在與圖形處理單元(GPU) 104相關聯的記憶體塊中，被儲存在與數位訊號處理器(DSP) 106相關聯的記憶體塊中，被儲存在記憶體塊118中，及/或可以分佈在多個塊中。在CPU 102處執行的指令可以是從與CPU 102相關聯的程式記憶體載入的，或者可以是從記憶體塊118載入的。

【0075】 SOC 100亦可以包括針對特定功能定製的額外處理塊，諸如GPU 104、DSP 106、連接塊110(其可以包括第五代(5G)連接、第四代長期進化(4G LTE)連接、Wi-Fi連接、USB連接、藍芽連接等)、以及多媒體處理器112(例如，其可以偵測和辨識手勢)。在一種實現方式中，在CPU 102、DSP 106及/或GPU 104中實現NPU。SOC 100亦可以包括感測器處理器114、影像訊號處理器(ISP) 116及/或導航模組120(其可以包括全球定位系統)。

【0076】 SOC 100可以是基於ARM指令集的。在本案內容的一態樣中，被載入到CPU 102中的指令可以包括用於在與輸入值和濾波器權重的乘法乘積相對應的查閱資料表(LUT)中搜尋被儲存的乘法結果的代碼。被載入到CPU 102中的指令亦可以包括用於當偵測到乘法乘積的查閱資

料表命中時在乘法乘積的乘法運算期間禁用乘法器的代碼。另外，被載入到CPU 102中的指令可以包括用於當偵測到乘法乘積的查閱資料表缺失時儲存輸入值和濾波器權重的計算出的乘法乘積的代碼。

【0077】 SOC 100及/或其部件可以被配置為使用根據本文所論述的本案內容的各態樣的機器學習技術來執行視訊壓縮及/或解壓縮（亦被稱為視訊編碼及/或解碼，其被統稱為視訊譯碼）。經由使用深度學習架構來執行視訊壓縮及/或解壓縮，本案內容的各態樣可以提高設備上的視訊壓縮及/或解壓縮的效率。例如，使用所描述的視訊譯碼技術的設備可以使用基於機器學習的技術來對視訊更高效地壓縮，可以將經壓縮的視訊發送給另一設備，並且另一設備可以使用本文描述的基於機器學習的技術來對經壓縮的視訊更高效地解壓縮。

【0078】 如前述，神經網路是機器學習系統的實例，並且可以包括輸入層、一或多個隱藏層和輸出層。從輸入層的輸入節點提供資料，由一或多個隱藏層的隱藏節點執行處理，並且經由輸出層的輸出節點產生輸出。深度學習網路典型地包括多個隱藏層。神經網路的每個層可以包括可以包括人工神經元（或節點）的特徵圖或啟動圖。特徵圖可以包括濾波器、核等。節點可以包括用於指示這些層中的一或多個層的節點的重要性的一或多個權重。在一些情況下，深度學習網路可以具有一系列的多個隱藏層，其中初

期的層用於決定輸入的簡單且低級的特性，以及後期的層建立更複雜且抽象的特性的層次結構。

【0079】 深度學習架構可以學習功能的層次結構。例如，若以視覺資料來提供，則第一層可以學習辨識輸入串流中的相對簡單的特徵，諸如邊緣。在另一實例中，若以聽覺資料來提供，則第一層可以學習辨識特定頻率中的頻譜功率。採用第一層的輸出作為輸入的第二層可以學習辨識特徵的組合，諸如用於視覺資料的簡單形狀或用於聽覺資料的聲音的組合。例如，較高層可以學習表示視覺資料中的複雜形狀或者聽覺資料中的詞語。再高層可以學習辨識常見的視覺物件或口語短語。

【0080】 當被應用於具有自然層次結構的問題時，深度學習架構可能表現得尤其好。例如，對機動運載工具的分類可以得益於首先學習辨識輪子、擋風玻璃和其他特徵。這些特徵可以在較高層處以不同的方式進行組合，以辨識汽車、卡車和飛機。

【0081】 神經網路可以被設計為具有多種連接模式。在前饋網路中，資訊是從較低層傳遞到較高層的，其中給定層之每一者神經元與較高層中的神經元進行通訊。如前述，可以在前饋網路的連續層中建立層次表示。神經網路亦可以具有循環或回饋（亦被稱為自頂向下）連接。在循環連接中，來自給定層中的神經元的輸出可以被傳送給同一層中的另一個神經元。循環架構可以有助於辨識跨一個以上的按順序被遞送給神經網路的輸入資料區塊的模式。從給

定層的神經元到較低層的神經元的連接被稱為回饋（或自頂向下）連接。當辨識高級概念可以輔助辨別輸入的特定低級特徵時，具有許多回饋連接的網路可以是有說明的。

【0082】 神經網路的各層之間的連接可以是全連接的或者局部連接的。圖 2 A 示出全連接神經網路 202 的實例。在全連接神經網路 202 中，第一層中的神經元可以將其輸出傳送給第二層之每一者神經元，使得第二層之每一者神經元將接收來自第一層之每一者神經元的輸入。圖 2 B 示出局部連接神經網路 204 的實例。在局部連接神經網路 204 中，第一層中的神經元可以連接到第二層中的有限數量的神經元。更一般而言，局部連接神經網路 204 的局部連接層可以被配置為使得一層之每一者神經元將具有相同或相似的連接模式，但是具有可能帶有不同的值（例如，210、212、214 和 216）的連接強度。局部連接的連線性模式可以在較高層中造成空間上不同的感受野，這是因為給定區域中的較高層神經元可以接收經由訓練而調諧到針對網路的總輸入的受限部分的屬性的輸入。

【0083】 局部連接神經網路的一個實例是迴旋神經網路。圖 2 C 示出迴旋神經網路 206 的實例。迴旋神經網路 206 可以被配置為使得與針對第二層之每一者神經元的輸入相關聯的連接強度被共享（例如，208）。迴旋神經網路可以非常適於其中輸入的空間位置有意義的問題。根據本案內容的各態樣，迴旋神經網路 206 可以用於執行視訊壓縮及 / 或解壓縮的一或多個態樣。

【0084】 一種類型的迴旋神經網路是深度迴旋網路 (DCN)。圖2D示出DCN 200的詳細實例，DCN 200被設計為從影像226中辨識視覺特徵，影像226是從諸如車載相機之類的影像擷取裝置230輸入的。可以對當前實例的DCN 200進行訓練以辨識交通標誌和在交通標誌上提供的數位。當然，DCN 200可以被訓練用於其他任務，例如辨識車道標線或辨識交通燈。

【0085】 可以利用有監督學習來訓練DCN 200。在訓練期間，可以向DCN 200呈現影像（諸如限速標誌的影像226），並且隨後可以計算前向傳遞以產生輸出222。DCN 200可以包括特徵提取部分和分類部分。在接收到影像226時，迴旋層232可以將迴旋核（未圖示）應用於影像226以產生特徵圖的第一集合218。作為一實例，用於迴旋層232的迴旋核可以是產生 28×28 特徵圖的 5×5 核。在本實例中，因為在特徵圖的第一集合218中產生四個不同的特徵圖，所以在迴旋層232處對影像226應用四個不同的迴旋核。迴旋核亦可以被稱為濾波器或迴旋濾波器。

【0086】 特徵圖的第一集合218可以由最大池化層（未圖示）二次取樣以產生特徵圖的第二集合220。最大池化層減小特徵圖的第一集合218的尺寸。亦即，特徵圖的第二集合220的尺寸（諸如 14×14 ）小於特徵圖的第一集合218的尺寸（諸如 28×28 ）。減小的尺寸向後續層提供類似的資訊，同時減少記憶體消耗。特徵圖的第二集合220可以

經由一或多個後續迴旋層（未圖示）被進一步迴旋以產生一或多個後續特徵圖集合（未圖示）。

【0087】 在圖 2 D 的實例中，對特徵圖的第二集合 2 2 0 進行迴旋以產生第一特徵向量 2 2 4。此外，對第一特徵向量 2 2 4 進一步迴旋以產生第二特徵向量 2 2 8。第二特徵向量 2 2 8 的每個特徵可以包括對應於影像 2 2 6 的可能特徵（諸如，「符號」、「60」和「100」）的數字。softmax 函數（未圖示）可以將第二特徵向量 2 2 8 中的數位轉換為概率。照此，DCN 200 的輸出 2 2 2 是影像 2 2 6 包括一或多個特徵的概率。

【0088】 在本實例中，輸出 2 2 2 中的針對「符號」和「60」的概率高於輸出 2 2 2 的其他項（諸如，「30」、「40」、「50」、「70」、「80」、「90」和「100」）的概率。在訓練之前，由 DCN 200 產生的輸出 2 2 2 可能是不正確的。因此，可以計算輸出 2 2 2 與目標輸出之間的誤差。目標輸出是影像 2 2 6 的地面真值（例如，「符號」和「60」）。隨後，可以調整 DCN 200 的權重，使得 DCN 200 的輸出 2 2 2 與目標輸出更緊密地對準。

【0089】 為了調整權重，學習演算法可以計算針對權重的梯度向量。梯度可以指示關於在權重被調整的情況下誤差將增加或減少的量。在頂層，梯度可以直接地對應於將倒數第二層中的啟動神經元和輸出層中的神經元進行連接的權重的值。在較低層中，梯度可以取決於權重的值和所計算出的較高層的誤差梯度。隨後可以調整權重以減小誤

差。這種調整權重的方式可以被稱為「反向傳播」，因為其涉及經由神經網路的「後向傳遞」。

【0090】 在實踐中，權重的誤差梯度可以經由少量實例來計算，使得所計算出的梯度接近真實誤差梯度。這種近似方法可以被稱為隨機梯度下降法。可以重複隨機梯度下降，直到整個系統的可實現誤差率已經停止下降或直到誤差率已經達到目標水平。在學習之後，可以向DCN呈現新影像，並且經由網路的前向傳遞可以產生可以被認為是DCN的推斷或預測的輸出222。

【0091】 深度信念網路（DBN）是包括多層隱藏節點的概率模型。DBN可以用於提取訓練資料集的分層表示。DBN可以經由將受限玻爾茲曼機（RBM）的各層進行疊加來獲得。RBM是一種類型的人工神經網路，其可以學習在輸入集合上的概率分佈。由於RBM可以在沒有關於每個輸入應當被分類到的類別的資訊的情況下學習概率分佈，因此RBM通常在無監督學習中使用。使用混合的無監督和有監督範式，DBN的底部RBM可以以無監督的方式來訓練並且可以充當特徵提取器，以及頂部RBM可以以有監督的方式來訓練（基於來自前一層的輸入和目標類別的聯合分佈）並且可以充當分類器。

【0092】 深度迴旋網路（DCN）是迴旋網路的網路，其被配置有額外的池化和非線性（例如，正規化）層。DCN在許多工作上都已經實現了最先進的效能。DCN可以使用有監督學習（其中輸入目標和輸出目標兩者對於許多範例都

是已知的)來訓練，並且經由使用梯度下降方法來修改網路的權重。

【0093】 DCN可以是前饋網路。此外，如前述，從DCN的第一層中的神經元到下一較高層中的一組神經元的連接是在第一層中跨越神經元共享的。DCN的前饋和共享連接可以用於快速處理。例如，DCN的計算負擔可以比包括循環或回饋連接的類似尺寸的神經網路的計算負擔小得多。

【0094】 對迴旋網路的每個層的處理可以被認為空間不變範本或基投影。若首先將輸入分解為多個通道(諸如彩色影像的紅色、綠色和藍色通道)，則在該輸入上訓練的迴旋網路可以被認為是三維的，其中兩個空間維度沿著沿著影像的軸線，以及第三維度擷取色彩資訊。迴旋連接的輸出可以被認為是形成後續層中的特徵圖，其中特徵圖(例如，220)的每個元素接收來自前一層中的一系列神經元(例如，特徵圖218)和來自多個通道之每一者通道的輸入。可以利用非線性(諸如整流 $\max(0, x)$)來進一步處理特徵圖中的值。來自鄰近神經元的值可以被進一步池化(這對應於下取樣)，並且可以提供額外的局部不變性和降維。

【0095】 圖3是示出深度迴旋網路350的實例的方塊圖。深度迴旋網路350可以包括基於連線性和權重共享的多個不同類型的層。如圖3所示，深度迴旋網路350包括迴旋塊354A、迴旋塊354B。迴旋塊354A、迴旋塊354B中的各者可以被配置為具有迴旋層(CONV)356、正規化層(LNorm)358和最大池層(MAX POOL)360。

【0096】 迴旋層 356 可以包括一或多個迴旋濾波器，其可以被應用於輸入資料 352 以產生特徵圖。儘管僅圖示兩個迴旋塊 354 A、迴旋塊 354 B，但是本案內容並不限於此，而相反，根據設計偏好，任何數量的迴旋塊（例如，塊 354 A、塊 354 B）可以被包括在深度迴旋網路 350 中。正規化層 358 可以對迴旋濾波器的輸出進行正規化。例如，正規化層 358 可以提供白化或橫向抑制。最大池化層 360 可以提供在空間內的下取樣聚合以用於局部不變性和降維。

【0097】 例如，深度迴旋網路的並行濾波器組可以被載入在 SOC 100 的 CPU 102 或 GPU 104 上，以實現高效能和低功耗。在替代實施例中，並行濾波器組可以被載入在 SOC 100 的 DSP 106 或 ISP 116 上。此外，深度迴旋網路 350 可以存取可以在 SOC 100 上存在的其他處理塊，諸如分別專用於感測器和導航的感測器處理器 114 和導航模組 120。

【0098】 深度迴旋網路 350 亦可以包括一或多個全連接層，諸如層 362 A（被標記為「FC1」）和層 362 B（被標記為「FC2」）。深度迴旋網路 350 可以進一步包括邏輯回歸（LR）層 364。在深度迴旋網路 350 的每個層 356、358、360、362 A、362 B、364 之間是要被更新的權重（未圖示）。這些層（例如，356、358、360、362 A、362 B、364）之每一者層的輸出可以充當深度迴旋網路 350 的這些層（例如，356、358、360、362 A、362 B、364）中的隨後層的輸入，以從在迴旋塊 354 A 的開始提供

的輸入資料 352 (例如, 影像、音訊、視訊、感測器資料、及/或其他輸入資料)學習分層特徵表示。深度迴旋網路 350 的輸出是針對輸入資料 352 的分類得分 366。分類得分 366 可以是概率集合, 其中每個概率是包括輸入資料來自特徵集合的特徵的概率。

【0099】 如上文所提到的, 數位視訊資料可以包括大量資料, 這可能為通訊網路以及處理和儲存視訊資料的設備帶來顯著負擔。例如, 記錄未經壓縮的視訊內容通常導致隨著所記錄的視訊內容的解析度增加而大大增加的大的檔尺寸。在一個說明性實例中, 以 1080p/24 (例如, 在寬度上具有 1920 個圖元和在高度上具有 1080 個圖元的解析度, 其中每秒擷取 24 個訊框) 記錄的未經壓縮的每通道 16 位元視訊可能佔用每訊框 12.4 百萬位元組或每秒 297.6 百萬位元組。以每秒 24 個訊框的 4K 解析度而記錄的未經壓縮的每通道 16 位元視訊可能佔用每訊框 49.8 百萬位元組或每秒 1195.2 百萬位元組。

【0100】 網路頻寬是大視訊檔可能變得有問題的另一約束。例如, 視訊內容通常在無線網路上 (例如, 經由 LTE、改進的 LTE、新無線電 (NR)、WiFiTM、藍芽TM 或其他無線網路) 遞送, 並且可能構成消費者網際網路流量的很大一部分。儘管在無線網路中的可用頻寬量態樣取得了進步, 但是可能仍然期望減少用於在這些網路中遞送視訊內容的頻寬量。

【0101】 由於未經壓縮的視訊內容可能導致大檔，大檔可能涉及用於實體儲存的相當大的記憶體和用於傳輸的相當大的頻寬，因此可以利用視訊譯碼技術來壓縮以及隨後解壓縮此類視訊內容。

【0102】 為了減小視訊內容的大小——並且因此減小儲存視訊內容所涉及的儲存量以及傳送視訊內容所涉及的頻寬量，可以根據特定的視訊譯碼標準（諸如，H E V C、A V C、M P E G、V V C 等等）來執行視訊譯碼技術。視訊譯碼通常使用預測方法（諸如訊框間預測、訊框內預測等），預測方法利用在視訊影像或序列中存在的冗餘。視訊譯碼技術的一個共同目標是將視訊資料壓縮為使用較低位元速率的形式，同時避免或最小化視訊品質的降級。隨著對視訊服務的需求增長以及新的視訊服務變得可用，需要具有更好的譯碼效率、效能和速率控制的譯碼技術。

【0103】 一般而言，編碼設備根據視訊譯碼標準來對視訊資料進行編碼，以產生經編碼的視訊位元串流。在一些實例中，經編碼的視訊位元串流（或「視訊位元串流」或「位元串流」）是一系列的一或多個經編碼的視訊序列。編碼設備可以經由將每個圖片分割為多個切片來產生圖片的編碼表示。一切片是獨立於其他切片的，使得對該切片中的資訊進行編碼，而不依賴於來自在同一圖片內的其他切片的資料。切片包括一或多個切片段，其包括獨立的切片段以及（若存在的話）依賴於先前切片段的一或多個從屬切片段。在 H E V C 中，隨後將切片分割為亮度取樣和色度取

樣的譯碼樹塊 (CTB)。亮度取樣的CTB和色度取樣的一或多個CTB連同用於取樣的語法一起被稱為譯碼樹單元 (CTU)。CTU亦可以被稱為「樹塊」或「最大譯碼單元」 (LCU)。CTU是用於HEVC編碼的基本處理單元。CTU可以被拆分為尺寸不同的多個譯碼單元 (CU)。CU包含被稱為譯碼塊 (CB) 的亮度和色度取樣陣列。

【0104】 亮度和色度CB可以被進一步拆分為預測塊 (PB)。PB是亮度分量或色度分量的取樣塊，其使用相同的運動參數來進行訊框間預測或塊內複製 (IBC) 預測 (當可用或被啟用以供使用時)。亮度PB和一或多個色度PB連同相關聯的語法一起形成預測單元 (PU)。對於訊框間預測，運動參數集 (例如，一或多個運動向量、參考索引等) 是在用於每個PU的位元串流中用訊號通知的，並且用於亮度PB和一或多個色度PB的訊框間預測。運動參數亦可以被稱為運動資訊。CB亦可以被分割為一或多個變換塊 (TB)。TB表示色彩分量的取樣的正方形塊，其中殘差變換 (例如，在一些情況下，相同的二維變換) 被應用於其以對預測殘留訊號進行譯碼。變換單元 (TU) 表示亮度和色度取樣的TB以及對應的語法元素。下文更加詳細地描述變換譯碼。

【0105】 根據HEVC標準，可以使用TU來執行變換。可以基於在給定CU內的PU的尺寸來設定TU的尺寸。TU可以具有與PU相同的尺寸或者小於PU。在一些實例中，可以使用被稱為殘差四叉樹 (RQT) 的四叉樹結構來將與CU

相對應的殘差取樣細分為更小的單元。RQT的蔓葉線節點可以對應於TU。與TU相關聯的圖元差值可以被變換以產生變換係數。隨後可以由編碼設備對變換係數進行量化。

【0106】 一旦視訊資料的圖片被分割為CU，編碼設備就使用預測模式來預測每個PU。隨後從原始視訊資料中減去預測單元或預測塊以獲得殘差（下文描述）。對於每個CU，可以使用語法資料在位元串流內用訊號通知預測模式。預測模式可以包括訊框內預測（或圖片內預測）或訊框間預測（或圖片間預測）。訊框內預測利用在圖片內在空間上相鄰的取樣之間的相關性。例如，使用訊框內預測，每個PU是從在相同圖片中的相鄰影像資料進行預測的，使用例如DC預測以找到用於PU的平均值，使用平面預測以使平面表面適配PU，使用方向預測以從相鄰資料進行推斷，或者使用任何其他適當的預測類型。訊框間預測使用在圖片之間的時間相關性，以便推導出用於影像取樣塊的運動補償預測。例如，使用訊框間預測，每個PU是使用運動補償預測來從在一或多個參考圖片（按照輸出次序在當前圖片之前或之後）中的影像資料進行預測的。例如，可以在CU級別處作出是使用圖片間預測還是使用圖片內預測來對圖片區域進行解碼的決策。

【0107】 在使用訊框內預測及/或訊框間預測執行預測之後，編碼設備可以執行變換和量化。例如，在預測之後，編碼設備可以計算與PU相對應的殘差值。殘差值可以包括在正被編碼的當前區塊（PU）與用於預測當前塊的預測塊

(例如，當前塊的預測版本)之間的圖元差值。例如，在產生預測塊(例如，實行訊框間預測或訊框內預測)之後，編碼設備可以經由從當前塊中減去由預測單元產生的預測塊來產生殘差塊。殘差塊包括圖元差值集合，其對在當前塊的圖元值與預測塊的圖元值之間的差進行量化。在一些實例中，可以用二維塊格式(例如，圖元值的二維矩陣或陣列)來表示殘差塊。在此類實例中，殘差塊是對圖元值的二維表示。

【0108】 使用塊變換來對在執行預測之後可能剩餘的任何殘差資料進行變換，塊變換可以是基於離散餘弦變換、離散正弦變換、整數變換、小波變換、其他適當的變換函數、或其任何組合的。在一些情況下，可以將一或多個塊變換(例如，尺寸 32×32 、 16×16 、 8×8 、 4×4 或其他適當的尺寸)應用於每個CU中的殘差資料。在一些實施例中，可以將TU用於由編碼設備實現的變換和量化程序。具有一或多個PU的給定CU亦可以包括一或多個TU。如下文進一步詳細描述的，殘差值可以使用塊變換而被變換為變換係數，並且隨後可以使用TU進行量化和掃描，以產生用於熵譯碼的序列化變換係數。

【0109】 編碼設備可以執行對變換係數的量化。量化經由對變換係數進行量化以減少用於表示係數的資料量來提供進一步的壓縮。例如，量化可以減小與係數中的一些或所有係數相關聯的位元深度。在一個實例中，具有 n 位元值的係數可以在量化期間向下捨入為 m 位元值，其中 n 大於 m 。

【0110】 一旦執行了量化，則經編碼的視訊位元串流包括經量化的變換係數、預測資訊（例如，預測模式、運動向量、塊向量等）、分割資訊以及任何其他適當的資料（諸如其他語法資料）。經編碼的視訊位元串流的不同元素隨後可以由編碼設備進行熵編碼。在一些實例中，編碼設備可以利用預定義的掃描次序來掃描經量化的變換係數，以產生可以被熵編碼的序列化向量。在一些實例中，編碼設備可以執行自我調整掃描。在掃描經量化的變換係數以形成向量（例如，一維向量）之後，編碼設備可以對該向量進行熵編碼。例如，編碼設備可以使用上下文自我調整可變長度譯碼、上下文自我調整二進位算術譯碼、基於語法的上下文自我調整二進位算術譯碼、概率區間分割熵譯碼或另一種適當的熵編碼技術。

【0111】 編碼設備可以儲存經編碼的視訊位元串流及/或可以在通訊鏈路上向接收設備（其可以包括解碼設備）發送經編碼的視訊位元串流資料。解碼設備可以經由熵解碼（例如，使用熵解碼器）並且提取構成經編碼的視訊資料的一或多個經編碼的視訊序列的元素，從而對經編碼的視訊位元串流資料進行解碼。解碼設備隨後可以重新縮放經編碼的視訊位元串流資料並且對其執行逆變換。殘差資料隨後被傳遞給解碼設備的預測階段。解碼設備隨後使用訊框內預測、訊框間預測、IBC及/或其他類型的預測來預測區塊（例如，PU）。在一些實例中，將預測與逆變換的輸出（殘差資料）相加。解碼設備可以將經解碼的視訊輸出

到視訊目標設備，視訊目標設備可以包括用於將經解碼的視訊資料顯示給內容的消費者的顯示器或其他輸出設備。

【0112】 經由各種視訊譯碼標準定義的視訊譯碼系統和技術（例如，上述 H E V C 視訊譯碼技術）可能能夠保留原始視訊內容中的大部分資訊，並且是可以基於訊號處理和資訊理論概念而先驗定義的。然而，在一些情況下，基於機器學習（ M L ）的影像及 / 或視訊系統可以提供優於非基於 M L 的影像和視訊譯碼系統（諸如基於端到端神經網路的影像和視訊譯碼（ E 2 E - N N V C ）系統）的優點。如前述，許多 E 2 E - N N V C 系統被設計為自動編碼器子網路（編碼器子網路）和負責學習在用於熵譯碼的經量化的隱變數上的概率模型的第二子網路的組合。這種架構可以被視為變換加量化模組（編碼器子網路）和熵建模子網路模組的組合。

【0113】 圖 4 圖示了系統 4 0 0 ，其包括被配置為使用 E 2 E - N N V C 系統來執行視訊編碼和解碼的設備 4 0 2 。設備 4 0 2 耦合到照相機 4 0 7 和儲存媒體 4 1 4 （例如，資料存放裝置）。在一些實現方式中，照相機 4 0 7 被配置為將影像資料 4 0 8 （例如，視訊資料串流）提供給處理器 4 0 4 ，以供 E 2 E - N N V C 系統進行編碼。在一些實現方式中，設備 4 0 2 可以耦合到及 / 或可以包括多個照相機（例如，雙照相機系統、三個照相機或其他數量的照相機）。在一些情況下，設備 4 0 2 可以耦合到麥克風及 / 或其他輸入設備（例如，鍵盤、滑鼠、諸如觸控式螢幕及 / 或觸控板之類的觸摸輸入設備、及 / 或其他輸入設備）。在一些實例中，照相機 4 0 7 、

儲存媒體 414、麥克風及/或其他輸入設備可以是設備 402 的一部分。

【0114】 設備 402 亦經由傳輸媒體 418（例如，一或多個無線網路、一或多個有線網路、或其組合）耦合到第二設備 490。例如，傳輸媒體 418 可以包括由無線網路、有線網路、或者有線網路和無線網路的組合提供的通道。傳輸媒體 418 可以形成基於封包的網路的一部分，諸如區域網路、廣域網或諸如網際網路之類的全球網路。傳輸媒體 418 可以包括路由器、交換機、基地台或可以用於促進從源設備到接收設備的通訊的任何其他設備。無線網路可以包括任何無線介面或無線介面的組合，並且可以包括任何合適的無線網路（例如，網際網路或其他廣域網、基於封包的網路、Wi-FiTM、射頻（RF）、UWB、Wi-Fi 直連、蜂巢、長期進化（LTE）、WiMaxTM 等）。有線網路可以包括任何有線介面（例如，光纖、乙太網路、電力線乙太網路、基於同軸電纜的乙太網路、數位訊號線（DSL）等）。有線網路及/或無線網路可以使用各種設備來實現，諸如基地台、路由器、存取點、橋接器、閘道、交換機等。經編碼的視訊位元串流資料可以根據諸如無線通訊協定之類的通訊標準進行調制，並且被發送給接收設備。

【0115】 設備 402 包括一或多個處理器（本文被稱為「處理器」）404，其耦合到記憶體 406、第一介面（「I/F 1」）412 和第二介面（「I/F 2」）416。處理器 404 被配置為從照相機 407、從記憶體 406 及/或從儲存媒體 414 接收影

像資料 408。處理器 404 經由第一介面 412（例如，經由記憶體匯流排）耦合到儲存媒體 414，並且經由第二介面 416（例如，網路周邊設備、無線收發機和天線、一或多個其他網路周邊設備、或其組合）耦合到傳輸媒體 418。

【0116】 處理器 404 包括 E2E-NNVC 系統 410。E2E-NNVC 系統 410 包括編碼器部分 462 和解碼器部分 466。在一些實現方式中，E2E-NNVC 系統 410 可以包括一或多個自動編碼器。編碼器部分 462 被配置為接收輸入資料 470 並處理輸入資料 470，以至少部分地基於輸入資料 470 來產生輸出資料 474。

【0117】 在一些實現方式中，E2E-NNVC 系統 410 的編碼器部分 462 被配置為對輸入資料 470 執行失真壓縮以產生輸出資料 474，使得輸出資料 474 具有與輸入資料 470 相比較少的位元。可以對編碼器部分 462 進行訓練以對輸入資料 470（例如，影像或視訊訊框）進行壓縮，而不使用基於任何先前表示（例如，一或多個先前重構的訊框）的運動補償。例如，編碼器部分 462 可以僅使用來自視訊訊框的視訊資料來壓縮該視訊訊框，而不使用先前重構的訊框的任何資料。由編碼器部分 462 處理的視訊訊框在本文中可以被稱為經訊框內預測的訊框（I 訊框）。在一些實例中，可以使用傳統視訊譯碼技術（例如，根據 HEVC、VVC、MPEG-4 或其他視訊譯碼標準）來產生 I 訊框。在此類實例中，處理器 404 可以包括被配置為執行基於塊的訊框內預測（諸如上文關於 HEVC 標準該的訊框內預測）的視訊譯

碼設備（例如，編碼設備）或者與其耦合。在此類實例中，處理器 404 可以不包括 E2E-NNVC 系統 410。

【0118】 在一些實現方式中，可以對 E2E-NNVC 系統 410 的編碼器部分 462 進行訓練以使用基於先前表示（例如，一或多個先前重構的訊框）的運動補償來壓縮輸入資料 470（例如，視訊訊框）。例如，編碼器部分 462 可以使用來自視訊訊框的視訊資料並且使用先前重構的訊框的資料來壓縮視訊訊框。由編碼器部分 462 處理的視訊訊框在本文中可以被稱為經訊框內預測的訊框（P 訊框）。運動補償可以用於經由描述來自先前重構的訊框的圖元如何移動到當前訊框中的新位置上以及殘差資訊來決定當前訊框的資料。

【0119】 如圖所述，E2E-NNVC 系統 410 的編碼器部分 462 包括神經網路 463 和量化器 464。神經網路 463 可以包括產生中間資料 472 的一或多個迴旋神經網路（CNN）、一或多個全連接神經網路、一或多個閘控循環單元（GRU）、一或多個長短期記憶（LSTM）網路、一或多個迴旋 RNN、一或多個迴旋 GRU、一或多個迴旋 LSTM、一或多個 GAN、其任何組合、及 / 或其他類型的神經網路架構。中間資料 472 被輸入到量化器 464。在圖 6A - 圖 6E 中圖示可以被包括在編碼器部分 462 中的部件的實例。

【0120】 量化器 464 被配置為執行對中間資料 472 的量化以及在一些情況下執行對其的熵譯碼，以產生輸出資料 474。輸出資料 474 可以包括經量化的（並且在一些情況

下，經熵編碼的)資料。由量化器464執行的量化操作可以導致從中間資料472產生量化碼(或表示由E2E-NNVC系統410產生的量化碼的資料)。量化碼(或表示量化碼的資料)亦可以被稱為隱碼(latent code)或隱變數(表示為 z)。被應用於隱變數的熵模型在本文中可以被稱為「先驗」。在一些實例中，可以使用當根據現有的視訊譯碼標準來編碼及/或解碼視訊資料時執行的現有的量化及/或熵譯碼操作，來執行量化和熵譯碼操作。在一些實例中，量化及/或熵譯碼操作可以由E2E-NNVC系統410來完成。在一個說明性實例中，可以使用有監督訓練來訓練E2E-NNVC系統410，其中在訓練期間，殘差資料被用作輸入，並且量化碼和熵碼被用作已知輸出(標籤)。

【0121】 E2E-NNVC系統410的解碼器部分466被配置為(例如，直接從量化器464及/或從儲存媒體414)接收輸出資料474。解碼器部分466可以處理輸出資料474，以至少部分地基於輸出資料474來產生對輸入資料470的表示476。在一些實例中，E2E-NNVC系統410的解碼器部分466包括神經網路468，其可以包括一或多個CNN、一或多個全連接神經網路、一或多個GRU、一或多個長短期記憶(LSTM)網路、一或多個迴旋RNN、一或多個迴旋GRU、一或多個迴旋LSTM、一或多個GAN、其任何組合、及/或其他類型的神經網路架構。在圖6A-圖6E中示出可以被包括在解碼器部分466中的部件的實例。

【0122】 處理器 404 被配置為將輸出資料 474 發送給傳輸媒體 418 或儲存媒體 414 中的至少一者。例如，可以將輸出資料 474 儲存在儲存媒體 414 處，以供解碼器部分 466 稍後檢索和解碼（或解壓縮）以產生對輸入資料 470 的表示 476 作為經重構的資料。經重構的資料可以用於各種目的，諸如以用於重播已被編碼/壓縮以產生輸出資料 474 的視訊資料。在一些實現方式中，可以在與解碼器部分 466 匹配的另一解碼器設備（例如，在設備 402 中、在第二設備 490 中或在另一設備中）處解碼輸出資料 474，以產生對輸入資料 470 的表示 476 作為經重構的資料。例如，第二設備 490 可以包括與解碼器部分 466 匹配（或基本上匹配）的解碼器，並且可以經由傳輸媒體 418 將輸出資料 474 發送給第二設備 490。第二設備 490 可以處理輸出資料 474，以產生對輸入資料 470 的表示 476 作為經重構的資料。

【0123】 系統 400 的部件可以包括電路或其他電子硬體及/或可以使用其來實現，電子電路或其他電子硬體可以包括一或多個可程式設計電子電路（例如，微處理器、圖形處理單元（GPU）、數位訊號處理器（DSP）、中央處理單元（CPU）、及/或其他適當的電子電路），及/或可以包括電腦軟體、韌體或其任何組合，及/或使用其來實現，以執行本文描述的各种操作。

【0124】 儘管系統 400 被示為包括某些部件，但是本發明所屬領域中具有通常知識者將領會的是，系統 400 可以包括比在圖 4 中所示的部件更多或更少的部件。例如，系統

400亦可以包括輸入設備和輸出設備（未圖示），或者可以是包括輸入設備和輸出設備的計算設備的一部分。在一些實現方式中，系統400亦可以包括以下各項或者可以是包括以下各項的計算設備的一部分：一或多個記憶體設備（例如，一或多個隨機存取記憶體（RAM）部件、唯讀記憶體（ROM）部件、快取緩衝記憶體部件、緩衝器部件、資料庫部件、及/或其他記憶體設備）、與一或多個記憶體設備相通訊及/或電連接到其的一或多個處理設備（例如，一或多個CPU、GPU及/或其他處理設備）、用於執行無線通訊的一或多個無線介面（例如，包括用於每個無線介面的一或多個收發機和基頻處理器）、用於在一或多個硬接線連接上執行通訊的一或多個有線介面（例如，諸如通用序列匯流排（USB）輸入之類的序列介面、照明連接器、及/或其他有線介面）、及/或在圖4中未圖示的其他部件。

【0125】 在一些實現方式中，系統400可以由計算設備本端實現及/或被包括在計算設備中。例如，計算設備可以包括行動設備、個人電腦、平板電腦、虛擬實境（VR）設備（例如，頭戴式顯示器（HMD）或其他VR設備）、增強現實（AR）設備（例如，HMD、AR眼鏡或其他AR設備）、可穿戴設備、伺服器（例如，在軟體作為服務（SaaS）系統中或其他基於伺服器的系統）、電視機、及/或具有執行本文描述的技術的資源能力的任何其他計算設備。

【0126】 在一個實例中，E2E-NNVC系統410可以被併入可攜式電子設備，可攜式電子設備包括：記憶體406，其

耦合到處理器 404 並且被配置為儲存可由處理器 404 執行的指令；及無線收發機，其耦合到天線和處理器 404 並且可操作以將輸出資料 474 發送給遠端設備。

【0127】 E2E-NNVC 系統被通常設計為處理 RGB 輸入。在以下項中描述了針對 RGB 輸入的影像和視訊譯碼方案的實例：J. Balle, D. Minnen, S. Singh, S. J. Hwang, N. Johnston, 「Variational image compression with a scale hyperprior」, ICLR, 2018 (被稱為「J. Balle 論文」), 以及 D. Minnen, J. Balle, G. Toderici, 「Joint Autoregressive and Hierarchical Priors for Learned Image Compression」, CVPR 2018 (被稱為「D. Minnen 論文」), 據此經由引用的方式將其全部內容併入並且用於所有目的。

【0128】 圖 5 是示出 J. Balle 論文中描述的 E2E-NNVC 系統實例的示意圖。圖 5 的 E2E-NNVC 系統中的 g_a 和 g_s 子網路分別對應於編碼器子網路 (例如, 編碼器部分 462) 和解碼器子網路 (例如, 解碼器部分 466)。圖 5 的 g_a 和 g_s 子網路被設計用於三通道 RGB 輸入, 其中所有三個 R、G 和 B 輸入通道皆經過相同的神經網路層 (迴旋層和廣義除法正規化 (GDN) 層) 並且由其處理。神經網路層可以包括執行迴旋運算的迴旋層和實現局部除法正規化的 GDN 及 / 或逆 GDN (IGDN) 非線性層。局部除法正規化是一種類型的變換, 其已被證明特別適合於影像的密度建模和壓

縮。E 2 E - N N V C 系統（諸如圖 5 所示）以具有類似統計特性的輸入通道為目標，諸如 R G B 資料（其中不同 R、G 和 B 通道的統計特性相似）。

【0129】 儘管 E 2 E - N N V C 系統典型地被設計為處理 R G B 輸入，但是大多數影像和視訊譯碼系統使用 Y U V 輸入格式（例如，在許多情況下，Y U V 4 2 0 輸入格式）。Y U V 格式的資料的色度（U 和 V）通道可以相對於亮度（Y）通道進行二次取樣。二次取樣對視覺品質產生最小的影響（例如，對視覺品質沒有顯著或明顯的影響）。二次取樣格式包括 Y U V 4 2 0 格式、Y U V 4 2 2 格式及 / 或其他 Y U V 格式。在 Y U V 格式中，通道之間的相關性降低，這可能與其他色彩格式（例如，R G B 格式）不同。此外，亮度（Y）和色度（U 和 V）通道的統計是不同的。例如，與亮度通道相比，U 和 V 通道具有較小的方差，而在例如 R G B 格式中，不同 R、G 和 B 通道的統計特性更相似。視訊轉碼器 - 解碼器（或轉碼器）是根據資料的輸入特性來設計的（例如，轉碼器可以根據資料的輸入格式來對資料進行編碼及 / 或解碼）。例如，若訊框的色度通道被二次取樣（例如，色度通道的解析度是亮度通道的一半），則當轉碼器預測用於運動補償的訊框的塊時，亮度塊的寬度和高度皆將是色度塊的兩倍。在另一實例中，轉碼器可以決定將針對色度和亮度來對多少圖元進行編碼或解碼等等。

【0130】 若將 R G B 輸入資料（如前述，大多數 E 2 E - N N V C 系統被設計為處理）替換為 Y U V 4 : 4 : 4 輸入資料（其中所

有通道具有相同尺寸)，則由於亮度（Y）和色度（U和V）通道的不同的統計特性，E2E-NNVC系統處理輸入資料的效能降低。如前述，色度（U和V）通道以一些YUV格式進行二次取樣，諸如在YUV420的情況下。例如，對於具有YUV4:2:0格式的內容，U和V通道解析度為Y通道解析度的一半（由於寬度和高度減半，U和V通道的大小為Y通道的四分之一）。此類二次取樣可能導致輸入資料與E2E-NNVC系統的輸入不相容。輸入資料是E2E-NNVC系統嘗試進行編碼及/或解碼的資訊（例如，包括三個通道（包括亮度（Y）和色度（U和V）通道）的YUV訊框）。許多基於神經網路的系統假設輸入資料的所有通道尺寸相同，並且因此將所有輸入通道饋送到同一網路。在此類情況下，可以（例如，使用矩陣加法）添加某些操作的輸出，在這種情況下，通道的尺寸必須相同。

【0131】 在一些實例中，為了解決此類問題，可以將Y通道二次取樣為四個半解析度Y通道。四個半解析度Y通道可以與兩個色度通道組合，從而形成六個輸入通道。六個輸入通道可以輸入或饋送到針對RGB輸入設計的E2E-NNVC系統。此類方法可以解決關於亮度（Y）和色度（U和V）通道的解析度差異的問題。然而，亮度（Y）通道與色度（U和V）通道之間的固有差異仍然存在，從而導致差的譯碼（例如，編碼及/或解碼）效能。

【0132】 如前述，本文描述了用於使用一或多個基於ML的系統來執行影像及/或視訊譯碼的系統和技術。本文描述

的系統和技術提供被設計用於處理具有亮度-色度（YUV）輸入格式（例如，YUV420、YUV444、YUV422等）的輸入資料的前端架構（例如，新的子網路，諸如在基於端到端神經網路的影像和視訊譯碼（E2E-NNVC）系統中）。在一些實例中，前端架構被配置為在針對RGB輸入格式設計的E2E-NNVC中適應YUV 4:2:0輸入格式。如前述，前端架構適用於許多E2E-NNVC架構（例如，包括在J. Balle論文和D. Minnen論文中描述的架構）。本文描述的系統和技術考慮亮度（Y）通道和色度（U和V）通道的不同特性以及亮度（Y）通道和色度（U和V）通道的解析度差異。E2E-NNVC系統可以對包括多個訊框的獨立訊框（或影像）及/或視訊資料進行編碼及/或解碼。

【0133】 在一些實例中，本文描述的系統和技術最初可以將Y通道和UV通道輸入或饋送到兩個分離的層中。隨後，E2E-NNVC系統可以在某一數量的層之後（例如，在第一對迴旋和非線性層或其他層之後，如下文描述的圖6A-圖6E所示）對與Y通道和UV通道相關聯的資料進行組合。由於U和V色度分量相對於亮度（Y）通道進行二次取樣，因此可以跳過第一迴旋層中的二次取樣，並且特定大小（例如，具有為 $(N/2+1) \times (N/2+1)$ 的大小）的迴旋（例如，CNN）核心可以用於二次取樣的輸入色度（U和V）通道。則與用於色度（U和V）通道的核心相比，具有不同大小的CNN核心（例如， $N \times N$ CNN核心）可以用於亮度（Y）通道。可以使用迴旋層（例如， 1×1 迴旋層）來對前

端架構的兩個分支（分別攜帶亮度和色度通道或分量資訊）進行組合，該迴旋層跨通道對值進行組合。對 1×1 迴旋層的使用可以提供如本文描述的各種益處，包括提高解碼效率。

【0134】 圖 6 A - 圖 6 F 示出神經網路系統的前端架構的實例。在一些實例中，圖 6 A - 圖 6 F 的前端架構可以是被設計用於處理（編碼及 / 或解碼）具有 Y U V 4:2:0 格式的資料的 E 2 E - N N V C 系統的一部分。例如，前端架構可以被配置用於處理 Y U V 4:2:0 格式的輸入資料。圖 6 A、圖 6 C、圖 6 D 和圖 6 E 的前端架構在 1×1 迴旋層之後應用了兩個不同的非線性運算元。例如，在圖 6 A 的架構中使用廣義除數正規化（G D N）運算元，而在圖 6 C - 圖 6 E 的架構中應用參數校正線性單元（P R e L U）非線性運算元。在一些實例中，如圖 6 A 和圖 6 C - 圖 6 F 所示的類似神經網路架構可以用於對其他類型的 Y U V 內容（例如，具有 Y U V 4:4:4 格式、Y U V 4:2:2 格式等的內容）及 / 或具有其他輸入格式的內容進行編碼及 / 或解碼。

【0135】 例如，圖 6 A 是示出可以被配置為直接利用 4:2:0 輸入（Y、U 和 V）資料工作的前端神經網路系統或架構的實例的示意圖。如圖 6 A 所示，在神經網路系統的編碼器子網路處，使用 1×1 迴旋層 6 0 6 來對分支亮度和色度通道（亮度（Y）通道 6 0 2 和色度（U 和 V）通道 6 0 4）進行組合，並且隨後應用 G D N 非線性運算元 6 0 8。在神經網路系統的解碼器子網路上執行類似的操作，但順序相反。例如，如

圖 6 A 所示，應用逆 GDN (IGDN) 運算元 609，使用 1×1 迴旋層 613 來分離 Y 和 U、V 通道，並且使用相應的 IGDN 615、IGDN 616 和迴旋層 617、迴旋層 618 來處理分離的 Y 通道和 U、V 通道。

【0136】 例如，圖 6 A 的神經網路系統的編碼器子網路中的前兩個神經網路層包括第一迴旋層 611 (表示為 $N_{cov} | 3 \times 3 | \downarrow 1$)、第二迴旋層 610 (表示為 $N_{conv} | 5 \times 5 | \downarrow 2$)、第一 GDN 層 614 和第二 GDN 層 612。圖 6 A 的前端神經網路架構的解碼器子網路中的最後兩個神經網路層包括第一逆 GDN (IGDN) 層 616、第二逆 GDN (IGDN) 615、用於產生訊框的經重構的色度 (U 和 V) 分量的第一迴旋層 618 (表示為 $2_{conv} | 3 \times 3 | \uparrow 1$)、以及用於產生訊框的經重構的亮度 (Y) 分量的第二迴旋層 617 (表示為 $1_{conv} | 5 \times 5 | \uparrow 2$)。「 N_{conv} 」標記法是指給定迴旋層的輸出通道數量 (N) (對應於輸出特徵數量) (其中 N 值定義了輸出通道數量)。 3×3 和 5×5 標記法指示相應的迴旋核心的大小 (例如， 3×3 核心和 5×5 核心)。「 $\downarrow 1$ 」和「 $\downarrow 2$ 」標記法是指跨距值，其中 $\downarrow 1$ 是指為 1 的跨距 (用於下取樣，如「 \downarrow 」所指示)，以及 $\downarrow 2$ 是指為 1 的跨距 (用於下取樣)。「 $\uparrow 1$ 」和「 $\uparrow 2$ 」標記法是指跨距值，其中 $\uparrow 1$ 是指為 1 的跨距 (用於上取樣，如「 \uparrow 」所指示)，以及 $\uparrow 2$ 是指為 1 的跨距 (用於上取樣)。

【0137】 例如，迴旋層 610 經由以跨距值 2 在水平和垂直維度上應用 5×5 迴旋濾波器來將輸入亮度通道 602 下取樣 4

倍。迴旋層 610 的結果輸出是 N 個特徵值陣列（對應於 N 個通道）。迴旋層 611 經由以跨距值 1 在水平和垂直維度應用 3×3 迴旋濾波器來處理輸入色度（ U 和 V ）通道 604。迴旋層 611 的結果輸出是 N 個特徵值陣列（對應於 N 個通道）。由迴旋層 610 輸出的特徵值陣列具有與由迴旋層 611 輸出的特徵值陣列相同的尺寸。隨後， GDN 層 612 可以處理由迴旋層 610 輸出的特徵值，並且 GDN 層 614 可以處理由迴旋層 611 輸出的特徵值。

【0138】 隨後， 1×1 迴旋層 606 可以處理由 GDN 層 612、 GDN 層 614 輸出的特徵值。 1×1 迴旋層 606 可以產生與亮度通道 602 和色度通道 604 相關聯的特徵的線性組合。線性組合操作作為 Y 分量和 UV 分量的每值交叉通道混合來操作，從而產生增強譯碼效能的交叉分量（例如，交叉亮度和色度分量）預測。 1×1 迴旋層 606 的每個 1×1 迴旋濾波器可以包括應用於亮度通道 602 的對應的第 N 通道和色度通道 604 的對應的第 N 通道的相應的縮放因數。

【0139】 圖 6B 是示出 1×1 迴旋層 638 的實例操作的示意圖。如前述， N 表示輸出通道數量。如圖 6B 所示，提供 $2N$ 個通道作為 1×1 迴旋層 638 的輸入，包括 N 通道色度（組合 U 和 V ）輸出 632 和 N 通道亮度（ Y ）輸出 634。在圖 6B 的實例中， N 的值等於 2，其指示具有 N 通道色度輸出 632 的值的兩個通道和具有 N 通道亮度輸出 634 的值的兩個通道。參照圖 6A， N 通道色度輸出 632 可以是來自 GDN 層 614 的輸出，並且 N 通道亮度輸出 634 可以是來自 GDN 層

612的輸出。然而，在其他實例中，N通道色度輸出632和N通道亮度輸出634可以從其他非線性層（例如，分別從圖6D的pReLU層652和pReLU層654、分別從圖6E的pReLU層662和pReLU層664）輸出或直接從迴旋層輸出（例如，分別從圖6F的迴旋層670和迴旋層671輸出）。

【0140】 1x1迴旋層638處理2N個通道並且執行2N個通道的特徵線性組合，並且隨後輸出N通道特徵或係數集合。1x1迴旋層638包括兩個1x1迴旋濾波器（基於 $N=2$ ）。第一1x1迴旋濾波器示為具有 s_1 值，並且第二1x1迴旋濾波器示為具有 s_2 值。 s_1 值表示第一縮放因數，並且 s_2 值表示第二縮放因數。在一個說明性實例中， s_1 值等於3，並且 s_2 值等於4。1x1迴旋層638的1x1迴旋濾波器中的各者都具有跨距值1，這指示縮放因數 s_1 和縮放因數 s_2 將應用於UV輸出632和Y輸出634之每一者值。

【0141】 例如，將第一1x1迴旋濾波器的縮放因數 s_1 應用於UV輸出632的第一通道（C1）之每一者值和Y輸出634的第一通道（C1）之每一者值。一旦經由第一1x1迴旋濾波器的縮放因數 s_1 對UV輸出632的第一通道（C1）的每個值和Y輸出634的第一通道（C1）的每個值進行縮放，就將經縮放的值組合成輸出值639的第一通道（C1）。將第二1x1迴旋濾波器的縮放因數 s_2 應用於UV輸出632的第二通道（C2）之每一者值和Y輸出634的第二通道（C2）之每一者值。在經由第二1x1迴旋濾波器的縮放因數 s_2 對UV輸出632的第二通道（C2）的每個值和Y輸出634的第

二通道 (C 2) 的每個值進行縮放之後，將經縮放的值組合成輸出值 6 3 9 的第二通道 (C 2) 。因此，將四個 Y 通道和 U V 通道 (兩個 Y 通道和兩個組合的 U V 通道) 混合或組合成兩個輸出通道 C 1 和輸出通道 C 2 。

【0142】 返回到圖 6 A ， 1 x 1 迴旋層 6 0 6 的輸出由編碼器子網路的額外 G D N 層和額外迴旋層處理。量化引擎 6 2 0 可以對編碼器子網路的最終神經網路層 6 1 9 輸出的特徵執行量化，以產生經量化的輸出。熵編碼引擎 6 2 1 可以對來自量化引擎 6 2 0 的經量化的輸出進行熵編碼，以產生位元串流。如圖 6 A 所示，熵編碼引擎 6 2 1 可以使用超先驗網路產生的先驗來執行熵編碼。神經網路系統可以輸出位元串流以用於儲存、用於傳輸到另一設備、伺服器設備或系統，及 / 或以其他方式輸出位元串流。

【0143】 神經網路系統的解碼器子網路或另一神經網路系統 (另一設備) 的解碼器子網路可以解碼位元串流。例如，如圖 6 A 所示，解碼器子網路的熵解碼引擎 6 2 2 可以對位元串流進行熵解碼，並且將經熵解碼的資料輸出到去量化引擎 6 2 3 。熵解碼引擎 6 2 2 可以使用超先驗網路產生的先驗來執行熵解碼，如圖 6 A 所示。去量化引擎 6 2 3 可以對資料進行去量化。經去量化的資料可以由解碼器子網路的多個迴旋層和多個逆 G D N (I G D N) 處理。

【0144】 在被 I G D N 層 6 0 9 處理之後， 1 x 1 迴旋層 6 1 3 可以處理資料。 1 x 1 迴旋層 6 1 3 可以包括 2 N 個迴旋濾波器，其可以將資料劃分為 Y 通道特徵和組合 U V 通道特徵。例如，

可以使用 1×1 迴旋層 613 的 $2N$ 個 1×1 迴旋（導致縮放）來處理由 IGDN 層 609 輸出的 N 個通道之每一者通道。對於與應用於 N 個輸入通道的輸出通道（總共 $2N$ 個輸出通道）相對應的每個縮放因數 n_i ，解碼器子網路可以在 N 個輸入通道上執行求和，從而產生 $2N$ 個輸出。在一個說明性實例中，對於縮放因數 n_1 ，解碼器子網路可以將縮放因數 n_1 應用於 N 個輸入通道，並且可以對結果求和，這導致一個輸出通道。解碼器子網路可以針對 $2N$ 個不同的縮放因數（例如，縮放因數 n_1 、縮放因數 n_2 至縮放因數 n_{2N} ）執行該操作。

【0145】 由 1×1 迴旋層 613 輸出的 Y 通道特徵可以由 IGDN 615 處理。由 1×1 迴旋層 613 輸出的組合 UV 通道特徵可以由 IGDN 616 處理。迴旋層 617 可以處理 Y 通道特徵並且輸出經重構的訊框的每圖元或取樣的經重構的 Y 通道（例如，亮度取樣或圖元），示為經重構的 Y 分量 624。迴旋層 618 可以處理組合 UV 通道特徵，並且可以輸出經重構的訊框的每圖元或取樣的經重構的 U 通道（例如，色度藍色取樣或圖元）和經重構的訊框的每圖元或取樣的經重構的 V 通道（例如，色度紅色取樣或圖元），示為經重構的 U 和 V 分量 625。

【0146】 圖 6C 是示出可以被配置為直接利用 4:2:0 輸入（Y、U 和 V）輸入資料進行操作的前端神經網路系統或架構的另一實例的示意圖。如圖 6C 所示，在神經網路系統的編碼器子網路處，使用 1×1 迴旋層 648 來對分支亮度和色度通道（亮度通道 642 和色度通道 644）進行組合（類似於上

文關於圖 6 A 的 1×1 迴旋層 606 描述的)，並且隨後應用 pReLU 非線性運算元 649。在其他實例中，可以應用除 pReLU 非線性運算元之外的非線性運算元。類似的操作由圖 6 C 的神經網路系統的解碼器子網路執行（類似於上文關於圖 6 A 描述的），但順序相反（例如，應用 pReLU 運算元，使用 1×1 迴旋層來分離 Y 和 U、V 通道，並且使用相應的逆 IGDN 和迴旋層來處理分離的 Y 和 U、V 通道）。

【0147】 與圖 5 中描述的 E2E-NNVC 系統（基於神經網路的轉碼器）相比，經由在 g_a （編碼器側）和對應的 g_s （解碼器側）中的前兩個網路層中單獨處理 Y 通道和 UV 通道來修改圖 6 A 和圖 6 C 的前端架構的輸入處理。用於處理 Y 分量的第一迴旋層（例如，圖 6 A 中的迴旋層 610 和圖 6 C 中的迴旋層 646）（表示為 $\text{Nconv} | 5 \times 5 | \downarrow 2$ ）可以與圖 5 中的第一迴旋層 510 相同或相似。類似地，圖 6 A 和圖 6 C 的解碼器子網路的用於產生經重構的亮度(Y)分量的第二迴旋層（表示為 $\text{conv} | 5 \times 5 | \uparrow 2$ ）可以與圖 5 的系統中的解碼器子網路 g_s 的最後一個迴旋層相同或相似。與圖 5 的系統不同，U 和 V 色度通道由圖 6 A 和圖 6 C 的架構使用單獨的迴旋層（例如，單獨的 CNN，諸如圖 6 A 中的迴旋層 611 或圖 6 C 中的迴旋層 647）（表示為 $\text{Nconv} | 3 \times 3 | \downarrow 1$ ）來處理，其中核心的大小為圖 6 A 中的 $\text{Nconv} | 5 \times 5 | \downarrow 2$ 迴旋層 610 或圖 6 C 中的 $\text{Nconv} | 5 \times 5 | \downarrow 2$ 迴旋層 646 的 Y 核心的一半（並且無下取樣，對應於等於 1 的跨距），之後跟有特定 GDN

層（一個 G D N 層用於亮度 Y，以及一個 G D N 層用於色度 U 和 V）。

【0148】 在圖 6 A 和圖 6 C 的迴旋層（第一對 C N N N c o n v | 5 x 5 | ↓ 2 和 N c o n v | 3 x 3 | ↓ 1 層）和 G D N 層之後，亮度（Y）通道和色度（U 和 V）通道（例如，輸入通道的變換或濾波版本）的表示或特徵具有相同的尺寸，並且隨後使用圖 6 A 的 1 x 1 迴旋層 6 0 6 或圖 6 C 的 1 x 1 迴旋層 6 4 8 進行組合。例如，在 Y U V 4 : 2 : 0 格式中，亮度（Y）通道在每個維度上的大小是色度（U 和 V）通道的兩倍。當經由 2 對色度（U 和 V）通道進行二次取樣時，基於處理這些通道產生的輸出變為與亮度通道的 c o n v 2 d 輸出相同的維度（因為亮度通道未被二次取樣）。通道的單獨正規化解決了亮度和色度通道的方差的差異。如前述，隨後可以在使用額外三個迴旋層之前應用非線性運算元（例如，使用 G D N 6 0 8 或 p R e L U 6 4 9），直到到達量化步驟為止。

【0149】 在圖 6 A 和圖 6 C 中的架構的解碼器子網路中，使用單獨的 I G D N 和迴旋層來分開地產生經重構的亮度（Y）分量和經重構的色度（U 和 V）分量。例如，圖 6 A 的用於產生經重構的色度（U 和 V）分量 6 2 5 的迴旋層 6 1 8（解碼器子網路的 2 c o n v | 3 x 3 | ↑ 1 層）的核心大小約為在用於產生經重構的亮度（Y）分量 6 2 4 的迴旋層 6 1 7（解碼器子網路的 1 c o n v | 5 x 5 | ↑ 2 層）中使用的核心大小的一半（並且無上取樣，對應於等於 1 的跨距）。

【0150】 圖 6 D 是示出可以被配置為直接利用 4:2:0 輸入 (Y、U 和 V) 輸入資料進行操作的前端神經網路架構的另一實例的示意圖。如圖 6 D 所示，在編碼器側，使用 1 x 1 迴旋層來對分支亮度和色度通道進行組合，並且隨後應用 pReLU 非線性運算元。與圖 6 A 和圖 6 C 所示的架構相比，亮度和色度分支中的 GDN 層被替換為 PReLU 運算元。

【0151】 圖 6 E 是示出可以被配置為直接利用 4:2:0 輸入 (Y、U 和 V) 輸入資料進行操作的前端神經網路架構的另一實例的示意圖。如圖 6 E 所示，在編碼器側，使用 1 x 1 迴旋層來對分支亮度和色度通道進行組合，並且隨後應用 pReLU 非線性運算元。與圖 6 A、圖 6 C 和圖 6 D 所示的架構相比，圖 6 E 中的架構的所有 GDN 層皆被替換為 PReLU 運算元。

【0152】 圖 6 F 是示出可以被配置為直接利用 4:2:0 輸入 (Y、U 和 V) 輸入資料進行操作的前端神經網路架構的另一實例的示意圖。如圖 6 F 所示，在編碼器側，使用 1 x 1 迴旋層來對分支亮度和色度通道進行組合。與圖 6 A - 圖 6 E 所示的架構相比，所有 GDN 層皆被完全移除，並且在迴旋層之間不使用非線性啟動操作。

【0153】 圖 6 C - 圖 6 F 中示出的神經網路架構設計可以用於減少 GDN 層 (例如，如圖 6 C 的架構所示) 或完全移除 GDN 層 (例如，如圖 6 E 和圖 6 F 的架構所示)。

【0154】 在一些實例中，本文描述的系統和技術可以用於在基於神經網路的譯碼系統的輸入處使用迴旋（例如，CNN）和正規化級組合的其他編碼器-解碼器子網路。

【0155】 圖7是示出使用本文描述的一或多個機器學習技術來處理視訊的程序700的實例的流程圖。在方塊702處，程序700包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值。

【0156】 在方塊704處，程序700包括：由編碼器子網路的第二迴旋層產生與訊框的至少一個色度通道相關聯的輸出值。在方塊706處，程序700包括：由第三迴旋層基於與訊框的亮度通道相關聯的輸出值和與訊框的至少一個色度通道相關聯的輸出值來產生訊框的組合表示。在一些情況下，第三迴旋層包括1x1迴旋層（例如，圖6A-圖6F的編碼器子網路的1x1迴旋層），其包括一或多個1x1迴旋濾波器。在方塊708處，程序700包括：基於訊框的組合表示來產生經編碼的視訊資料。

【0157】 在一些實例中，程序700包括：使用編碼器子網路的第一非線性層來處理與訊框的亮度通道相關聯的輸出值。程序700可以包括：使用編碼器子網路的第二非線性層來處理與訊框的至少一個色度通道相關聯的輸出值。在此類實例中，組合表示是基於第一非線性層的輸出和第二非線性層的輸出來產生的。在一些情況下，訊框的組合表示是由第三迴旋層（例如，圖6A-圖6F的編碼器子網路的

1 x 1 迴旋層) 使用第一非線性層的輸出和第二非線性層的輸出作為輸入來產生的。

【0158】 在一些實例中，程序 700 包括：(例如，使用量化引擎 620) 對經編碼的視訊資料進行量化。在一些實例中，程序 700 包括：對經編碼的視訊資料進行熵譯碼(例如，使用熵編碼引擎 621)。在一些實例中，程序 700 包括：將經編碼的視訊資料儲存在記憶體中。在一些實例中，程序 700 包括：在傳輸媒體上向至少一個設備發送經編碼的視訊資料。

【0159】 在一些實例中，程序 700 包括：獲得經編碼的訊框。程序 700 可以包括：由神經網路系統的解碼器子網路的第一迴旋層產生與經編碼的訊框的亮度通道相關聯的經重構的輸出值。程序 700 可以進一步包括：由解碼器子網路的第二迴旋層產生與經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。在一些實例中，程序 700 包括：使用解碼器子網路的第三迴旋層來將經編碼的訊框的亮度通道與經編碼的訊框的至少一個色度通道分離。在一些情況下，解碼器子網路的第三迴旋層包括 1 x 1 迴旋層(例如，圖 6 A - 圖 6 F 的解碼器子網路的 1 x 1 迴旋層)，其包括一或多個 1 x 1 迴旋濾波器。

【0160】 在一些實例中，訊框包括視訊訊框。在一些實例中，至少一個色度通道包括色度藍色通道和色度紅色通道。在一些實例中，訊框具有亮度-色度(YUV)格式。

【0161】 圖 8 是示出使用本文描述的一或多個機器學習技術來處理視訊的程序 800 的實例的流程圖。在方塊 802 處，程序 800 包括：獲得經編碼的訊框。在方塊 804 處，程序 800 包括：由解碼器子網路的第一迴旋層將經編碼的訊框的亮度通道與經編碼的訊框的至少一個色度通道分離。在一些情況下，解碼器子網路的第一迴旋層包括 1×1 迴旋層（例如，圖 6 A - 圖 6 F 的解碼器子網路的 1×1 迴旋層），其包括一或多個 1×1 迴旋濾波器。在方塊 806 處，程序 800 包括：由神經網路系統的解碼器子網路的第二迴旋層產生與經編碼的訊框的亮度通道相關聯的經重構的輸出值。在方塊 808 處，程序 800 包括：由解碼器子網路的第三迴旋層產生與經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。在方塊 810 處，程序 800 包括：產生輸出訊框，該輸出訊框包括與亮度通道相關聯的經重構的輸出值和與至少一個色度通道相關聯的經重構的輸出值。

【0162】 在一些實例中，程序 800 包括：使用解碼器子網路的第一非線性層來處理與經編碼的訊框的亮度通道相關聯的值。與亮度通道相關聯的經重構的輸出值是基於第一非線性層的輸出來產生的。程序 700 可以包括：使用解碼器子網路的第二非線性層來處理與經編碼的訊框的至少一個色度通道相關聯的值。與至少一個色度通道相關聯的經重構的輸出值是基於第二非線性層的輸出來產生的。

【0163】 在一些實例中，程序 800 包括：（例如，經由去量化引擎 623）對經編碼的訊框的取樣進行去量化。在一

些實例中，程序800包括：（例如，經由熵解碼引擎622）對經編碼的訊框的取樣進行熵解碼。在一些實例中，程序800包括：將輸出訊框儲存在記憶體中。在一些實例中，程序800包括：顯示輸出訊框。

【0164】 在一些實例中，程序800包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值。程序700可以包括：由編碼器子網路的第二迴旋層產生與訊框的至少一個色度通道相關聯的輸出值。程序700進一步可以包括：由編碼器子網路的第三迴旋層基於與訊框的亮度通道相關聯的輸出值和與訊框的至少一個色度通道相關聯的輸出值來產生訊框的組合表示。程序700可以包括：基於訊框的組合表示來產生經編碼的訊框。在一些情況下，編碼器子網路的第三迴旋層包括 1×1 迴旋層（例如，圖6A - 圖6F的編碼器子網路的 1×1 迴旋層），其包括一或多個 1×1 迴旋濾波器。

【0165】 在一些實例中，程序800包括：使用編碼器子網路的第一非線性層來處理與訊框的亮度通道相關聯的輸出值，以及使用編碼器子網路的第二非線性層來處理與訊框的至少一個色度通道相關聯的輸出值。在此類實例中，組合表示是基於第一非線性層的輸出和第二非線性層的輸出來產生的。在一些實例中，訊框的組合表示是由編碼器子網路的第三迴旋層使用第一非線性層的輸出和第二非線性層的輸出作為輸入來產生的。

【0166】 在一些實例中，經編碼的訊框包括經編碼的視訊訊框。在一些實例中，至少一個色度通道包括色度藍色通道和色度紅色通道。在一些實例中，經編碼的訊框具有亮度-色度（YUV）格式。

【0167】 在一些實例中，本文描述的程序（例如，本文描述的程序700、程序800及/或本文描述的其他程序）可以由計算設備或裝置（諸如具有在圖9中所示的計算設備架構900的計算設備）來執行。在一個實例中，程序700及/或程序800可以由計算設備來執行，其中計算設備架構900實現圖6A-圖6F種所示的神經網路架構之一。在一些實例中，計算設備可以包括行動設備（例如，行動電話、平板計算設備）、可穿戴設備、擴展現實設備（例如，虛擬實境（VR）設備、增強現實（AR）設備或混合現實（MR）設備）、個人電腦、膝上型電腦、視訊伺服器、電視機、車輛（或車輛的計算設備）、機器人設備及/或具有用於執行本文描述的程序（包括程序700及/或程序800）的資源能力的任何其他計算設備。

【0168】 在一些情況下，計算設備或裝置可以包括各種部件，諸如一或多個輸入設備、一或多個輸出設備、一或多個處理器、一或多個微處理器、一或多個微型電腦、一或多個發射器、接收器或組合發射器-接收器（例如，被稱為收發機）、一或多個照相機、一或多個感測器及/或被配置為執行本文描述的程序的其他部件。在一些實例中，計算設備可以包括顯示器、被配置為傳送及/或接收資

料的網路介面、其任何組合及/或其他部件。網路介面可以被配置為傳送及/或接收基於網際網路協定（IP）的資料或其他類型的資料。

【0169】 計算設備的部件可以在電路中實現。例如，部件可以包括及/或可以使用電子電路或其他電子硬體來實現，該等電子電路或其他電子硬體可以包括一或多個可程式設計電子電路（例如，微處理器、圖形處理單元（GPU）、數位訊號處理器（DSP）、中央處理單元（CPU）、神經處理單元（NPU）、及/或其他適當的電子電路），及/或可以包括電腦軟體、韌體或其任何組合及/或使用其來實現，以執行本文描述的重種操作。

【0170】 程序700和程序800被示為邏輯流程圖，邏輯流程圖的動作表示可以用硬體、電腦指令或其組合來實現的一系列操作。在電腦指令的背景重，該動作表示被儲存在一或多個電腦可讀取儲存媒體上的電腦可執行指令，該電腦可執行指令在由一或多個處理器執行時，執行所記載的操作。通常，電腦可執行指令包括執行特定功能或實現特定資料類型的常式、程式、物件、部件、資料結構等。描述操作的順序並不意欲被解釋為限制，並且可以以任意順序及/或並行地組合任何數量的所描述的操作以實現程序。

【0171】 補充地，本文描述的程序（包括程序700、程序800及/或本文描述的其他程序）可以在被配置有可執行指令的一或多個電腦系統的控制下執行，並且可以被實現為在一或多個處理器上共同執行的代碼（例如，可執行指令、

一或多個電腦程式、或一或多個應用)，經由硬體來實現，或其組合。如上文所提到，代碼可以例如以包括可由一或多個處理器執行的複數個指令的電腦程式的形式儲存在電腦可讀或機器可讀儲存媒體上。電腦可讀取儲存媒體或機器可讀儲存媒體可以是非暫時性的。

【0172】 圖9示出可以實現本文描述的各種技術的實例計算設備的實例計算設備架構900。在一些實例中，計算設備可以包括行動設備、可穿戴設備、擴展現實設備（例如，虛擬實境（VR）設備、增強現實（AR）設備或混合現實（MR）設備）、個人電腦、膝上型電腦、視訊伺服器、車輛（或車輛的計算設備）或其他設備。例如，計算設備架構900可以實現圖6的系統。計算設備架構900的部件被示為使用連接905（例如，匯流排）彼此電通訊。實例計算設備架構900包括處理單元（CPU或處理器）910和計算設備連接905，計算設備連接905將包括計算設備記憶體915（諸如，唯讀記憶體（ROM）920和隨機存取記憶體（RAM）925）的各種計算設備部件耦合到處理器910。

【0173】 計算設備架構900可以包括高速記憶體的快取記憶體，快取記憶體直接與處理器910連接、靠近處理器910或作為處理器910的一部分整合。計算設備架構900可以將資料從記憶體915及/或存放裝置930複製到快取記憶體912，以供處理器910快速存取。以這種方式，快取記憶體可以提供效能提升，其避免處理器910在等待資料時延遲。這些模組和其他模組可以控制或被配置為控制處理器

910 以執行各種行動。其他計算設備記憶體 915 亦可以可供使用。記憶體 915 可以包括具有不同效能特性的多種不同類型的記憶體。處理器 910 可以包括任何通用處理器以及硬體或軟體服務（諸如被儲存在存放裝置 930 中的服務 1 932、服務 2 934 和服務 3 936），硬體或軟體服務被配置為控制處理器 910 以及其中軟體指令被併入處理器設計中的專用處理器。處理器 910 可以是自包含系統，其包含多個核心或處理器、匯流排、記憶體控制器、快取記憶體等。多核心處理器可以是對稱的或非對稱的。

【0174】 為了實現使用者與計算設備架構 900 的互動，輸入設備 945 可以表示任何數量的輸入機制，諸如用於語音的麥克風、用於手勢或圖形輸入的觸控式螢幕、鍵盤、滑鼠、運動輸入、語音等。輸出設備 935 亦可以是本發明所屬領域中具有通常知識者已知的多種輸出機制中的一或多個，諸如顯示器、投影儀、電視機、揚聲器設備等。在一些情況下，多模態計算設備可以使得使用者能夠提供多種類型的輸入以與計算設備架構 900 進行通訊。通訊介面 940 通常可以控制和管理使用者輸入和計算設備輸出。對在任何特定硬體設定上的操作沒有限制，並且因此，此處的基本特徵可以容易地被改進的硬體或軟體配置（隨著它們被開發）所取代。

【0175】 存放裝置 930 是非揮發性記憶體，並且可以是硬碟或可以儲存可由電腦存取的資料的其他類型的電腦可讀取媒體，諸如磁帶、快閃記憶卡、固態記憶體設備、數位

多功能磁碟、盒式磁帶、隨機存取記憶體（RAM）925，唯讀記憶體（ROM）920、以及其的混合物。存放裝置930可以包括用於控制處理器910的服務932、服務934、服務936。預期其他硬體或軟體模組。存放裝置930可以連接到計算設備連接905。在一個態樣中，執行特定功能的硬體模組可以包括被儲存在電腦可讀取媒體中的軟體部件，該等軟體部件與執行該功能所必要的硬體部件（諸如處理器910、連接905、輸出設備935等）相連接。

【0176】 本案內容的各態樣適用於包括或耦合到一或多個主動深度感測系統的任何適當的電子設備（諸如安全系統、智慧型電話、平板電腦、膝上型電腦、車輛、無人機或其他設備）。儘管下文關於具有或耦合到一個光投影器的設備進行了描述，但是本案內容的各態樣適用於具有任何數量的光投影器的設備，並且因此不限於特定設備。

【0177】 術語「設備」不限於一個或特定數量的實體物件（諸如一個智慧型電話、一個控制器、一個處理系統等）。如本文所使用的，設備可以是具有可以實現本案內容的至少一些部分的一或多個部件的任何電子設備。儘管以下描述和實例使用術語「設備」來描述本案內容的各個態樣，但是術語「設備」不限於特定配置、類型或數量的物件。補充地，術語「系統」不限於多個部件或特定實施例。例如，系統可以在一或多個印刷電路板或其他基板上實現，並且可以具有可移動或靜態的部件。儘管以下描述和實例

使用術語「系統」來描述本案內容的各個態樣，但是術語「系統」不限於特定配置、類型、或數量的物件。

【0178】 在以上描述中提供了具體細節以提供對本文提供的實施例和實例的透徹理解。然而，本發明所屬領域中具有通常知識者將理解的是，可以在沒有這些具體細節的情況下實施這些實施例。為了解釋清楚，在一些情況下，本文的技術可以被呈現為包括包含如下的功能方塊的單獨的功能方塊，這些功能方塊包括設備、設備部件、以軟體體現的方法中的步驟或常式、或者硬體和軟體的組合。除了在各圖中所示及/或本文描述的部件之外，亦可以使用額外的部件。例如，電路、系統、網路、程序和其他部件可以以方塊圖形式被示為部件，以便不會在不必要的細節上模糊這些實施例。在其他情況下，公知的電路、程序、演算法、結構和技術可能被示為不具有不必要的細節，以便避免模糊這些實施例。

【0179】 上文可能將各個實施例描述為程序或方法，該程序或方法被圖示為流程圖、流程示意圖、資料串流圖、結構圖或方塊圖。儘管流程圖可以將操作描述為順序的程序，但是這些操作中的許多操作可以並行或同時執行。另外，可以重新排列操作的次序。程序在其操作完成時被終止，但是可能具有未被包括在圖中的額外步驟。程序（`process`）可以對應於方法、函數、程序（`procedure`）、子常式、副程式等。當程序對應於函數時，其終止可以對應於該函數返回到調用函數或主函數。

【0180】 根據上述實例的程序和方法可以使用電腦可執行指令來實現，電腦可執行指令被儲存在電腦可讀取媒體中或者以其他方式可從電腦可讀取媒體得到。此類指令可以包括例如指令或資料，指令或資料使得通用電腦、專用電腦或處理設備執行或者以其他方式將其配置為執行特定功能或特定的一組功能。可以經由網路存取所使用的電腦資源的部分。電腦可執行指令可以是例如二進位檔案、諸如組合語言之類的中間格式指令、韌體、原始程式碼等。

【0181】 術語「電腦可讀取媒體」包括但不限於可攜式或非可攜式存放裝置、光學存放裝置以及能夠儲存、包含或攜帶指令及/或資料的各種其他媒體。電腦可讀取媒體可以包括可以在其中儲存資料的非暫時性媒體，並且不包括無線地或在有線連接上傳播的載波及/或暫時性電子訊號。非暫時性媒體的實例可以包括但不限於磁碟或磁帶、光學儲存媒體（諸如快閃記憶體）、記憶體或記憶體設備、磁碟或光碟、快閃記憶體、提供有非揮發性記憶體的USB設備、網路存放裝置、壓縮光碟（CD）或數位多功能光碟（DVD）、其任何適當的組合等等。電腦可讀取媒體可以具有儲存在其上的代碼及/或機器可執行指令，其可以表示程序、函數、副程式、程式、常式、子常式、模組、套裝軟體、軟體部件、或指令、資料結構或程式語句的任何組合。程式碼片段可以經由傳遞及/或接收資訊、資料、引數、參數或記憶體內容來耦合到另一程式碼片段或硬體電路。資訊、引數、參數、資料等可以經由包括記憶體共享、訊

息傳遞、符記傳遞、網路傳輸等的任何適當的單元來傳遞、轉發或發送。

【0182】 在一些實施例中，電腦可讀存放裝置、媒體和記憶體可以包括包含位元串流等的電纜或無線訊號。然而，當提及時，非暫時性電腦可讀取儲存媒體明確地不包括諸如能量、載波訊號、電磁波和訊號本身之類的媒體。

【0183】 實現根據這些揭示內容的程序和方法的設備可以包括硬體、軟體、韌體、中介軟體、微代碼、硬體描述語言或其任何組合，並且可以採用多種形狀因數中的任何一種。當用軟體、韌體、中介軟體或微代碼來實現時，用於執行必要任務的程式碼或程式碼片段（例如，電腦程式產品）可以被儲存在電腦可讀或機器可讀取媒體中。處理器可以執行必要任務。形狀因數的典型實例包括膝上型電腦、智慧型電話、行動電話、平板設備或其他小型形狀因數的個人電腦、個人數位助理、機架式設備、獨立設備等。本文描述的功能亦可以體現在周邊設備或外掛程式卡中。經由額外的舉例，這種功能亦可以在單個設備中執行的不同晶片或不同程序之間的電路板上實現。

【0184】 指令、用於傳送此類指令的媒體、用於執行它們的計算資源以及用於支援此類計算資源的其他結構是用於提供在本案內容中描述的功能的實例單元。

【0185】 在前面的描述中，參考本案的特定實施例描述了本案的各態樣，但是本發明所屬領域中具有通常知識者將認識到的是，本案不限於此。因此，儘管本文已經詳細描

述了本案的說明性實施例，但是應理解的是，可以以其他方式不同地體現和採用本發明構思，並且所附的請求項意欲被解釋為包括此類變型，除了由現有技術限制的變型。可以單獨地或共同地使用上述應用的各個特徵和態樣。此外，在不背離本說明書的更寬泛的精神和範疇的情況下，實施例可以在除了本文描述的環境和應用之外的任何數量的環境和應用中使用。因此，說明書和附圖被認為是說明性的而不是限制性的。出於說明的目的，以特定次序描述了方法。應當領會的是，在替代實施例中，可以以與所描述的次序不同的次序來執行所述方法。

【0186】 本發明所屬領域中具有通常知識者將領會的是，在不背離本說明書的範疇的情況下，本文中使用的小於（「 $<$ 」）和大於（「 $>$ 」）符號或術語可以分別用小於或等於（「 \leq 」）以及大於或等於（「 \geq 」）符號來替換。

【0187】 在將部件描述為「被配置為」執行某些操作的情況下，這種配置可以例如經由以下方式來實現：將電子電路或其他硬體設計為執行該操作，將可程式設計電路（例如，微處理器或其他適當的電路）程式設計為執行該操作，或其任何組合。

【0188】 短語「耦合到」代表直接或間接地實體連接到另一部件的任何部件、及/或直接地或者間接地與另一部件通訊的任何部件（例如，經由有線或無線連接及/或其他適當的通訊介面而連接到另一部件）。

【0189】 記載集合中的「至少一個」及/或集合中的「一或多個」的請求項語言或其他語言指示該集合中的一個成員或者該集合中的多個成員滿足該請求項。例如，記載「A和B中的至少一個」或「A或B中的至少一個」的請求項語言意指A、B、或者A和B。在另一實例中，記載「A、B和C中的至少一個」或「A、B或C中的至少一個」的請求項語言意指A、B、C、或A和B、或A和C、或B和C、或A和B和C。語言集合中的「至少一個」及/或集合中的「一或多個」不將該集合限制為在該集合中列出的條目。例如，記載「A和B中的至少一個」或「A或B中的至少一個」的請求項語言可以意指A、B、或A和B，並且可以另外地包括在A和B的集合中未列出的條目。

【0190】 結合本文揭示的實施例描述的各種說明性的邏輯方塊、模組、電路和演算法步驟可以被實現為電子硬體、電腦軟體、韌體或其組合。為了清楚地說明硬體和軟體的這種可互換性，上面已經對各種說明性的部件、方塊、模組、電路和步驟圍繞其功能進行了整體描述。至於這種功能被實現為硬體還是軟體取決於特定的應用和被施加在整個系統上的設計約束。具有普通知識者可以針對每種特定應用以不同的方式來實現所描述的功能，但是這種實現決策不應當被解釋為導致背離本案的範疇。

【0191】 本文描述的技術亦可以用電子硬體、電腦軟體、韌體或其任何組合來實現。此類技術可以在各種設備中的任何一種中實現，諸如通用電腦、無線通訊設備手機或具

有多種用途（包括在無線通訊設備手機和其他設備中的應用）的積體電路設備。被描述為模組或部件的任何特徵皆可以在整合邏輯裝置中一起實現，或者分別作為個別但是可交互動操作的邏輯裝置來實現。若用軟體來實現，則該等技術可以至少部分地由電腦可讀取資料儲存媒體來實現，電腦可讀取資料儲存媒體包括程式碼，程式碼包括在被執行時執行上述方法中的一或多個方法的指令。電腦可讀取資料儲存媒體可以形成電腦程式產品的一部分，電腦程式產品可以包括包裝材料。電腦可讀取媒體可以包括記憶體或資料儲存媒體，諸如隨機存取記憶體（R A M）（諸如同步動態隨機存取記憶體（S D R A M））、唯讀記憶體（R O M）、非揮發性隨機存取記憶體（N V R A M）、電子可抹除可程式設計唯讀記憶體（E E P R O M）、快閃記憶體、磁或光資料儲存媒體等。補充地或替代地，該等技術可以至少部分地由以指令或資料結構的形式攜帶或傳送程式碼並且可以由電腦存取、讀取及/或執行的電腦可讀通訊媒體（諸如傳播的訊號或波）來實現。

【0192】 程式碼可以由處理器執行，處理器可以包括一或多個處理器，諸如一或多個數位訊號處理器（D S P）、通用微處理器、特殊應用積體電路（A S I C）、現場可程式設計邏輯陣列（F P G A）或其他等效的整合或個別邏輯電路。此類處理器可以被配置為執行在本案內容中描述的任何技術。通用處理器可以是微處理器，但是在替代方式中，處理器可以是任何一般的處理器、控制器、微控制器或狀態

機。處理器亦可以被實現為計算設備的組合，例如，DSP和微處理器的組合、複數個微處理器、一或多個微處理器與DSP核心的結合、或任何其他此類配置。因此，如本文所使用的術語「處理器」可以代表任何前述結構、前述結構的任何組合、或適於實現本文描述的技術的任何其他結構或裝置。

【0193】 本案內容的說明性實例包括：

【0194】 態樣1：一種處理視訊資料的方法，該方法包括：由神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；由該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；由第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生經編碼的視訊資料。

【0195】 態樣2：根據態樣1之方法，其中該第三迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0196】 態樣3：根據態樣1或2中任一項所述的方法，亦包括：使用該編碼器子網路的第一非線性層來處理與該訊框的該亮度通道相關聯的該等輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；其中該組合表示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0197】 態樣 4：根據態樣 3 之方法，其中該訊框的該組合表示是由該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0198】 態樣 5：根據態樣 1 至 4 中任一項所述的方法，亦包括：對該經編碼的視訊資料進行量化。

【0199】 態樣 6：根據態樣 1 至 5 中任一項所述的方法，亦包括：對該經譯碼的視訊資料進行熵解碼。

【0200】 態樣 7：根據態樣 1 至 6 中任一項所述的方法，亦包括：將該經編碼的視訊資料儲存在記憶體中。

【0201】 態樣 8：根據態樣 1 至 7 中任一項所述的方法，亦包括：在傳輸媒體上向至少一個設備發送該經編碼的視訊資料。

【0202】 態樣 9：根據態樣 1 至 8 中任一項所述的方法，亦包括：獲得經編碼的訊框；由該神經網路系統的解碼器子網路的第一迴旋層產生與經編碼的訊框的亮度通道相關聯的經重構的輸出值；及由該解碼器子網路的第二迴旋層產生與經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。

【0203】 態樣 10：根據態樣 9 之方法，亦包括：使用該解碼器子網路的第三迴旋層來將經編碼的訊框的該亮度通道與經編碼的訊框的該至少一個色度通道分離。

【0204】 態樣 11：根據態樣 10 之方法，其中該解碼器子網路的該第三迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0205】 態樣 1 2：根據態樣 1 至 1 1 中任一項所述的方法，其中該訊框包括視訊訊框。

【0206】 態樣 1 3：根據態樣 1 至 1 2 中任一項所述的方法，其中該至少一個色度通道包括色度藍色通道和色度紅色通道。

【0207】 態樣 1 4：根據態樣 1 至 1 3 中任一項所述的方法，其中該訊框具有亮度色度（YUV）格式。

【0208】 態樣 1 5：一種用於處理視訊資料的裝置。該裝置包括：記憶體；及處理器，其耦合到該記憶體並且被配置為：使用神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；使用該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；使用第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生經編碼的視訊資料。

【0209】 態樣 1 6：根據態樣 1 5 之裝置，其中該第三迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0210】 態樣 1 7：根據態樣 1 5 或 1 6 中任一項所述的裝置，其中該處理器被配置為：使用該編碼器子網路的第一非線性層來處理與該訊框的該亮度通道相關聯的該等輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；其中該組合表

示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0211】 態樣 18：根據態樣 17 之裝置，其中該訊框的該組合表示是由該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0212】 態樣 19：根據態樣 15 至 18 中任一項所述的裝置，其中該處理器被配置為：對該經編碼的視訊資料進行量化。

【0213】 態樣 20：根據態樣 15 至 19 中任一項所述的裝置，其中該處理器被配置為：對該經編碼的視訊資料進行熵解碼。

【0214】 態樣 21：根據態樣 15 至 20 中任一項所述的裝置，其中該處理器被配置為：將該經編碼的視訊資料儲存在記憶體中。

【0215】 態樣 22：根據態樣 15 至 21 中任一項所述的裝置，其中該處理器被配置為：在傳輸媒體上向至少一個設備發送該經編碼的視訊資料。

【0216】 態樣 23：根據態樣 15 至 22 中任一項所述的裝置，其中該處理器被配置為：獲得經編碼的訊框；使用該神經網路系統的解碼器子網路的第一迴旋層來產生與經編碼的訊框的亮度通道相關聯的經重構的輸出值；及使用該解碼器子網路的第二迴旋層來產生與經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。

【0217】 態樣 2 4：根據態樣 2 3 之裝置，其中該處理器被配置為：使用該解碼器子網路的第三迴旋層來將經編碼的訊框的該亮度通道與經編碼的訊框的該至少一個色度通道分離。

【0218】 態樣 2 5：根據態樣 2 4 之裝置，其中該解碼器子網路的該第三迴旋層包括 1 x 1 迴旋層，該 1 x 1 迴旋層包括一或多個 1 x 1 迴旋濾波器。

【0219】 態樣 2 6：根據態樣 1 5 至 2 5 中任一項所述的裝置，其中該訊框包括視訊訊框。

【0220】 態樣 2 7：根據態樣 1 5 至 2 6 中任一項所述的裝置，其中該至少一個色度通道包括色度藍色通道和色度紅色通道。

【0221】 態樣 2 8：根據態樣 1 5 至 2 7 中任一項所述的裝置，其中該訊框具有亮度色度（Y U V）格式。

【0222】 態樣 2 9：根據態樣 1 5 至 2 8 中任一項所述的裝置，其中該處理器包括神經處理單元（N P U）。

【0223】 態樣 3 0：根據態樣 1 5 至 2 9 中任一項所述的裝置，其中該裝置包括行動設備。

【0224】 態樣 3 1：根據態樣 1 5 至 3 0 中任一項所述的裝置，其中該裝置包括擴展現實設備。

【0225】 態樣 3 2：根據態樣 1 5 至 3 1 中任一項所述的裝置，亦包括顯示器。

【0226】 態樣 3 3：根據態樣 1 5 至 2 9 中任一項所述的裝置，其中該裝置包括電視機。

【0227】 態樣 3 4：根據態樣 1 5 至 3 3 中任一項所述的裝置，其中該裝置包括：照相機，其被配置為擷取一或多個視訊訊框。

【0228】 態樣 3 5：一種儲存指令的電腦可讀取儲存媒體，該等指令在被執行時使得一或多個處理器執行根據態樣 1 至 1 4 之操作中的任一項操作。

【0229】 態樣 3 6：一種裝置，包括用於執行根據態樣 1 至 1 4 之操作中的任一項操作的單元。

【0230】 態樣 3 7：一種處理視訊資料的方法，該方法包括：獲得經編碼的訊框；由該解碼器子網路的第一迴旋層將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離；由神經網路系統的解碼器子網路的第二迴旋層產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；由該解碼器子網路的第三迴旋層產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及產生輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【0231】 態樣 3 8：根據態樣 3 7 之方法，其中該解碼器子網路的該第一迴旋層包括 1 x 1 迴旋層，該 1 x 1 迴旋層包括一或多個 1 x 1 迴旋濾波器。

【0232】 態樣 3 9：根據態樣 3 7 或 3 8 中任一項所述的方法，亦包括：使用該解碼器子網路的第一非線性層來處理與該經編碼的訊框的亮度通道相關聯的值，其中與該亮度

通道相關聯的該等經重構的輸出值是基於該第一非線性層的輸出來產生的；及使用該解碼器子網路的第二非線性層來處理與該經編碼的訊框的該至少一個色度通道相關聯的值，其中與該至少一個色度通道相關聯的該等經重構的輸出值是基於該第二非線性層的輸出來產生的。

【0233】 態樣 40：根據態樣 37 至 39 中任一項所述的方法，亦包括：對該經編碼的訊框的取樣進行去量化。

【0234】 態樣 41：根據態樣 37 至 40 中任一項所述的方法，亦包括：對該經編碼的訊框的取樣進行熵解碼。

【0235】 態樣 42：根據態樣 37 至 41 中任一項所述的方法，亦包括：將該輸出訊框儲存在記憶體中。

【0236】 態樣 43：根據態樣 37 至 42 中任一項所述的方法，亦包括：顯示該輸出訊框。

【0237】 態樣 44：根據態樣 37 至 43 中任一項所述的方法，亦包括：由該神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；由該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；由該編碼器子網路的第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生該經編碼的訊框。

【0238】 態樣 45：根據態樣 44 之方法，其中該編碼器子網路的該第三迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0239】 態樣 46：根據態樣 44 或 45 中任一項所述的方法，亦包括：使用該編碼器子網路的第一非線性層來處理與該訊框的亮度通道相關聯的輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；其中該組合表示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0240】 態樣 47：根據態樣 46 之方法，其中該訊框的該組合表示是由該編碼器子網路的該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0241】 態樣 48：根據態樣 37 至 47 中任一項所述的方法，其中該經編碼的訊框包括經編碼的視訊訊框。

【0242】 態樣 49：根據態樣 37 至 48 中任一項所述的方法，其中該至少一個色度通道包括色度藍色通道和色度紅色通道。

【0243】 態樣 50：根據態樣 37 至 49 中任一項所述的方法，其中該經編碼的訊框具有亮度-色度（YUV）格式。

【0244】 態樣 49：一種用於處理視訊資料的裝置。該裝置包括：記憶體；及處理器，其耦合到該記憶體並且被配置為：獲得經編碼的訊框；使用該解碼器子網路的第一迴旋層來將該經編碼的訊框的亮度通道與該經編碼的訊框的至少一個色度通道分離；使用神經網路系統的解碼器子網路的第二迴旋層來產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；使用該解碼器子網路的第三迴旋層

來產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及產生輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【0245】 態樣 50：根據態樣 49 之裝置，其中該解碼器子網路的該第一迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0246】 態樣 51：根據態樣 49 或 50 中任一項所述的裝置，其中該處理器被配置為：使用該解碼器子網路的第一非線性層來處理與該經編碼的訊框的亮度通道相關聯的值，其中與該亮度通道相關聯的該等經重構的輸出值是基於該第一非線性層的輸出來產生的；及使用該解碼器子網路的第二非線性層來處理與該經編碼的訊框的該至少一個色度通道相關聯的值，其中與該至少一個色度通道相關聯的該等經重構的輸出值是基於該第二非線性層的輸出來產生的。

【0247】 態樣 52：根據態樣 49 至 51 中任一項所述的裝置，其中該處理器被配置為：對該經編碼的訊框的取樣進行去量化。

【0248】 態樣 53：根據態樣 49 至 52 中任一項所述的裝置，其中該處理器被配置為：對該經編碼的訊框的取樣進行熵解碼。

【0249】 態樣 5 4：根據態樣 4 9 至 5 3 中任一項所述的裝置，其中該處理器被配置為：將該輸出訊框儲存在記憶體中。

【0250】 態樣 5 5：根據態樣 4 9 至 5 4 中任一項該的裝置，其中該處理器被配置為：顯示該輸出訊框。

【0251】 態樣 5 6：根據態樣 4 9 至 5 5 中任一項所述的裝置，其中該處理器被配置為：由該神經網路系統的編碼器子網路的第一迴旋層產生與訊框的亮度通道相關聯的輸出值；由該編碼器子網路的第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；由該編碼器子網路的第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的組合表示；及基於該訊框的該組合表示來產生該經編碼的訊框。

【0252】 態樣 5 7：根據態樣 5 6 之裝置，其中該編碼器子網路的該第三迴旋層包括 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【0253】 態樣 5 8：根據態樣 4 4 或 5 7 中任一項所述的裝置，其中該處理器被配置為：使用該編碼器子網路的第一非線性層來處理與該訊框的亮度通道相關聯的輸出值；及使用該編碼器子網路的第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；其中該組合表示是基於該第一非線性層的輸出和該第二非線性層的輸出來產生的。

【0254】 態樣 59：根據態樣 58 之裝置，其中該訊框的該組合表示是由該編碼器子網路的該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的輸出作為輸入來產生的。

【0255】 態樣 60：根據態樣 49 至 59 中任一項所述的裝置，其中該經編碼的訊框包括經編碼的視訊訊框。

【0256】 態樣 61：根據態樣 49 至 60 中任一項所述的裝置，其中該至少一個色度通道包括色度藍色通道和色度紅色通道。

【0257】 態樣 62：根據態樣 49 至 61 中任一項所述的裝置，其中該經編碼的訊框具有亮度-色度（YUV）格式。

【0258】 態樣 63：根據態樣 49 至 62 中任一項所述的裝置，其中該處理器包括神經處理單元（NPU）。

【0259】 態樣 64：根據態樣 49 至 63 中任一項所述的裝置，其中該裝置包括行動設備。

【0260】 態樣 65：根據態樣 49 至 64 中任一項所述的裝置，其中該裝置包括擴展現實設備。

【0261】 態樣 66：根據態樣 49 至 65 中任一項所述的裝置，亦包括顯示器。

【0262】 態樣 67：根據態樣 49 至 63 中任一項所述的裝置，其中該裝置包括電視機。

【0263】 態樣 68：根據態樣 49 至 67 中任一項所述的裝置，其中該裝置包括：照相機，其被配置為擷取一或多個視訊訊框。

【0264】 態樣 69：一種儲存指令的電腦可讀取儲存媒體，該等指令在被執行時使得一或多個處理器執行根據態樣 37 至 48 之操作中的任一項操作。

【0265】 態樣 70：一種裝置，包括用於執行根據態樣 37 至 48 之操作中的任一項操作的單元。

【符號說明】

【0266】

100: SOC

102: CPU

104: GPU

106: DSP

108: 神經處理單元 (NPU)

110: 連接塊

112: 多媒體處理器

114: 感測器處理器

116: 影像訊號處理器 (ISP)

118: 記憶體塊

120: 導航模組

202: 全連接神經網路

204: 全連接神經網路

206: 迴旋神經網路

210: 值

212: 值

214: 值

2 1 6 : 值
2 1 8 : 第一集合
2 2 0 : 第二集合
2 2 2 : 輸出
2 2 4 : 第一特徵向量
2 2 6 : 影像
2 2 8 : 第二特徵向量
2 3 0 : 影像擷取裝置
2 3 2 : 迴旋層
3 5 0 : 深度迴旋網路
3 5 2 : 輸入資料
3 5 4 A : 迴旋塊
3 5 4 B : 迴旋塊
3 5 6 : 層
3 5 8 : 層
3 6 0 : 層
3 6 2 A : 層
3 6 2 B : 層
3 6 4 : 層
3 6 6 : 分類得分
4 0 0 : 系統
4 0 2 : 設備
4 0 4 : 處理器
4 0 6 : 記憶體

- 407: 照相機
- 408: 影像資料
- 410: E2E - NNVC 系統
- 412: 第一介面 (「 I / F 1 」)
- 414: 儲存媒體
- 416: 第二介面 (「 I / F 2 」)
- 418: 傳輸媒體
- 462: 編碼器部分
- 463: 神經網路
- 464: 量化器
- 466: 解碼器部分
- 468: 神經網路
- 470: 輸入資料
- 472: 中間資料
- 474: 輸出資料
- 476: 表示
- 490: 第二設備
- 510: 第一迴旋層
- 602: 亮度 (Y) 通道
- 604: 色度 (U 和 V) 通道
- 606: 1 x 1 迴旋層
- 608: GDN 非線性運算元
- 609: 逆 GDN (IGDN) 運算元
- 610: 第二迴旋層

- 6 1 1 : 迴旋層
- 6 1 2 : 第二 G D N 層
- 6 1 3 : 1 x 1 迴旋層
- 6 1 4 : 第一 G D N 層
- 6 1 5 : I G D N
- 6 1 6 : I G D N
- 6 1 7 : 迴旋層
- 6 1 8 : 迴旋層
- 6 1 9 : 最終神經網路層
- 6 2 0 : 量化引擎
- 6 2 1 : 熵編碼引擎
- 6 2 2 : 熵解碼引擎
- 6 2 3 : 去量化引擎
- 6 2 4 : Y 分量
- 6 2 5 : U 和 V 分量
- 6 3 2 : N 通道色度輸出
- 6 3 4 : N 通道亮度輸出
- 6 3 8 : 1 x 1 迴旋層
- 6 3 9 : 輸出值
- 6 4 2 : 亮度通道
- 6 4 4 : 色度通道
- 6 4 6 : 迴旋層
- 6 4 7 : 迴旋層
- 6 4 8 : 1 x 1 迴旋層

649: p R e L U 非線性運算元

652: p R e L U 層

654: p R e L U 層

662: p R e L U 層

664: p R e L U 層

670: 迴旋層

671: 迴旋層

700: 程序

702: 方塊

704: 方塊

706: 方塊

708: 方塊

800: 程序

802: 方塊

804: 方塊

806: 方塊

808: 方塊

810: 方塊

900: 計算設備架構

905: 連接

910: 處理器

912: 快取記憶體

915: 計算設備記憶體

920: 唯讀記憶體 (R O M)

9 2 5 : 隨 機 存 取 記 憶 體 (R A M)

9 3 0 : 存 放 裝 置

9 3 2 : 服 務

9 3 4 : 服 務

9 3 5 : 輸 出 設 備

9 3 6 : 服 務

9 4 0 : 通 訊 介 面

9 4 5 : 輸 入 設 備

C 1 : 第 一 通 道

C 2 : 第 二 通 道

g_a : 子 網 路

g_s : 子 網 路

G D N : 正 規 化

I G D N : 逆 正 規 化

p R e L U : 參 數 校 正 線 性 單 元

S_1 : 值

S_2 : 值

U : 色 度

V : 色 度

Y : 亮 度

【生物材料寄存】

國 內 寄 存 資 訊 (請 依 寄 存 機 構 、 日 期 、 號 碼 順 序 註 記)

無

國 外 寄 存 資 訊 (請 依 寄 存 國 家 、 機 構 、 日 期 、 號 碼 順 序 註 記)

無

【發明申請專利範圍】

【請求項 1】 一種處理視訊資料的方法，該方法包括以下步驟：

由一神經網路系統的一編碼器子網路的一第一迴旋層產生與一訊框的一亮度通道相關聯的輸出值；

由該編碼器子網路的一第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；

由一第三迴旋層基於與該訊框的該亮度通道相關聯的該等輸出值和與該訊框的該至少一個色度通道相關聯的該等輸出值來產生該訊框的一組合表示；及

基於該訊框的該組合表示來產生經編碼的視訊資料。

【請求項 2】 根據請求項 1 之方法，其中該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 3】 根據請求項 1 之方法，亦包括以下步驟：

使用該編碼器子網路的一第一非線性層來處理與該訊框的該亮度通道相關聯的該等輸出值；及

使用該編碼器子網路的一第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；

其中該組合表示是基於該第一非線性層的一輸出和該第二非線性層的一輸出來產生的。

【請求項 4】 根據請求項 3 之方法，其中該訊框的該組合表示是由該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的一輸出作為輸入來產生的。

【請求項 5】 根據請求項 1 之方法，亦包括以下步驟：
對該經編碼的視訊資料進行量化。

【請求項 6】 根據請求項 1 之方法，亦包括以下步驟：
對該經編碼的視訊資料進行熵譯碼。

【請求項 7】 根據請求項 1 之方法，亦包括以下步驟：
將該經編碼的視訊資料儲存在記憶體中。

【請求項 8】 根據請求項 1 之方法，亦包括以下步驟：
在一傳輸媒體上向至少一個設備發送該經編碼的視訊資料。

【請求項 9】 根據請求項 1 之方法，亦包括以下步驟：
獲得一經編碼的訊框；
由該神經網路系統的一解碼器子網路的一第一迴旋層產生與該經編碼的訊框的一亮度通道相關聯的經重構的輸出值；及

由該解碼器子網路的一第二迴旋層產生與該經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。

【請求項 10】 根據請求項 9 之方法，亦包括以下步驟：
使用該解碼器子網路的一第三迴旋層來將該經編碼的訊框的該亮度通道與該經編碼的訊框的該至少一個色度通道分離。

【請求項 11】 根據請求項 10 之方法，其中該解碼器子網路的該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 12】 根據請求項 1 之方法，其中該訊框包括一視

訊訊框。

【請求項 13】根據請求項 1 之方法，其中該至少一個色度通道包括一色度藍色通道和一色度紅色通道。

【請求項 14】根據請求項 1 之方法，其中該訊框具有一亮度色度（YUV）格式。

【請求項 15】一種用於處理視訊資料的裝置，包括：

一記憶體；及

一處理器，其耦合到該記憶體並且被配置為進行以下操作：

使用一神經網路系統的一編碼器子網路的一第一迴旋層產生與訊框的一亮度通道相關聯的輸出值；

使用該編碼器子網路的一第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；

使用一第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的一組合表示；及

基於該訊框的該組合表示來產生經編碼的視訊資料。

【請求項 16】根據請求項 15 之裝置，其中該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 17】根據請求項 15 之裝置，其中該處理器被配置為進行以下操作：

使用該編碼器子網路的一第一非線性層來處理與該訊

框的該亮度通道相關聯的該等輸出值；及

使用該編碼器子網路的一第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；

其中該組合表示是基於該第一非線性層的一輸出和該第二非線性層的一輸出來產生的。

【請求項 18】根據請求項 17 之裝置，其中該訊框的該組合表示是由該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的一輸出作為輸入來產生的。

【請求項 19】根據請求項 15 之裝置，其中該處理器被配置為：

對該經編碼的視訊資料進行量化。

【請求項 20】根據請求項 15 之裝置，其中該處理器被配置為：

對該經編碼的視訊資料進行熵編碼。

【請求項 21】根據請求項 15 中任一項該的裝置，其中該處理器被配置為：

將該經編碼的視訊資料儲存在記憶體中。

【請求項 22】根據請求項 15 之裝置，其中該處理器被配置為：

在一傳輸媒體上向至少一個設備發送該經編碼的視訊資料。

【請求項 23】根據請求項 15 之裝置，其中該處理器被配置為進行以下操作：

獲得一經編碼的訊框；

使用該神經網路系統的一解碼器子網路的一第一迴旋層來產生與該經編碼的訊框的一亮度通道相關聯的經重構的輸出值；及

使用該解碼器子網路的一第二迴旋層來產生與該經編碼的訊框的至少一個色度通道相關聯的經重構的輸出值。

【請求項 24】根據請求項 23 之裝置，其中該處理器被配置為：

使用該解碼器子網路的一第三迴旋層來將該經編碼的訊框的該亮度通道與該經編碼的訊框的該至少一個色度通道分離。

【請求項 25】根據請求項 24 之裝置，其中該解碼器子網路的該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 26】根據請求項 15 之裝置，其中該訊框包括一視訊訊框。

【請求項 27】根據請求項 15 之裝置，其中該至少一個色度通道包括一色度藍色通道和一色度紅色通道。

【請求項 28】根據請求項 15 之裝置，其中該訊框具有一亮度-色度 (YUV) 格式。

【請求項 29】根據請求項 15 之裝置，其中該處理器包括一神經處理單元 (NPU)。

【請求項 30】根據請求項 15 之裝置，其中該裝置包括一行動設備。

【請求項 3 1】根據請求項 1 5 之裝置，亦包括：一顯示器和被配置為擷取一或多個訊框的一照相機中的至少一者。

【請求項 3 2】一種處理視訊資料的方法，該方法包括以下步驟：

獲得一經編碼的訊框；

由一解碼器子網路的一第一迴旋層將該經編碼的訊框的一亮度通道與該經編碼的訊框的至少一個色度通道分離；

由一神經網路系統的該解碼器子網路的一第二迴旋層產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；

由該解碼器子網路的一第三迴旋層產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；

及

產生一輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【請求項 3 3】根據請求項 3 2 之方法，其中該解碼器子網路的該第一迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 3 4】根據請求項 3 2 之方法，亦包括以下步驟：

使用該解碼器子網路的一第一非線性層來處理與該經編碼的訊框的該亮度通道相關聯的值，其中與該亮度通

道相關聯的該等經重構的輸出值是基於該第一非線性層的一輸出來產生的；及

使用該解碼器子網路的一第二非線性層來處理與該經編碼的訊框的該至少一個色度通道相關聯的值，其中與該至少一個色度通道相關聯的該等經重構的輸出值是基於該第二非線性層的一輸出來產生的。

【請求項 35】根據請求項 32 之方法，亦包括以下步驟：

對該經編碼的訊框的取樣進行去量化。

【請求項 36】根據請求項 32 之方法，亦包括以下步驟：

對該經編碼的訊框的取樣進行熵解碼。

【請求項 37】根據請求項 32 之方法，亦包括以下步驟：

將該輸出訊框儲存在記憶體中。

【請求項 38】根據請求項 32 之方法，亦包括以下步驟：

顯示該輸出訊框。

【請求項 39】根據請求項 32 之方法，亦包括以下步驟：

由該神經網路系統的一編碼器子網路的一第一迴旋層產生與訊框的一亮度通道相關聯的輸出值；

由該編碼器子網路的一第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；

由該編碼器子網路的一第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的一組合表示；及

基於該訊框的該組合表示來產生該經編碼的訊框。

【請求項 40】根據請求項 39 之方法，其中該編碼器子網

路的該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 41】根據請求項 39 之方法，亦包括以下步驟：

使用該編碼器子網路的一第一非線性層來處理與該訊框的該亮度通道相關聯的輸出值；及

使用該編碼器子網路的一第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；

其中該組合表示是基於該第一非線性層的一輸出和該第二非線性層的一輸出來產生的。

【請求項 42】根據請求項 41 之方法，其中該訊框的該組合表示是由該編碼器子網路的該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的一輸出作為輸入來產生的。

【請求項 43】根據請求項 32 之方法，其中該經編碼的訊框包括一經編碼的視訊訊框。

【請求項 44】根據請求項 32 之方法，其中該至少一個色度通道包括一色度藍色通道和一色度紅色通道。

【請求項 45】根據請求項 32 之方法，其中該經編碼的訊框具有一亮度-色度 (YUV) 格式。

【請求項 46】一種用於處理視訊資料的裝置，包括：

一記憶體；及

一處理器，其耦合到該記憶體並且被配置為進行以下操作：

獲得一經編碼的訊框；

第 8 頁(發明申請專利範圍)

使用一解碼器子網路的一第一迴旋層來將該經編碼的訊框的一亮度通道與該經編碼的訊框的至少一個色度通道分離；

使用一神經網路系統的該解碼器子網路的一第二迴旋層來產生與該經編碼的訊框的該亮度通道相關聯的經重構的輸出值；

使用該解碼器子網路的一第三迴旋層來產生與該經編碼的訊框的該至少一個色度通道相關聯的經重構的輸出值；及

產生一輸出訊框，該輸出訊框包括與該亮度通道相關聯的經重構的輸出值和與該至少一個色度通道相關聯的經重構的輸出值。

【請求項 47】 根據請求項 46 之裝置，其中該解碼器子網路的該第一迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 48】 根據請求項 46 之裝置，其中該處理器被配置為進行以下操作：

使用該解碼器子網路的一第一非線性層來處理與該經編碼的訊框的該亮度通道相關聯的值，其中與該亮度通道相關聯的該等經重構的輸出值是基於該第一非線性層的一輸出來產生的；及

使用該解碼器子網路的一第二非線性層來處理與該經編碼的訊框的該至少一個色度通道相關聯的值，其中與該至少一個色度通道相關聯的該等經重構的輸出值是基

於該第二非線性層的一輸出來產生的。

【請求項 49】根據請求項 46 之裝置，其中該處理器被配置為：

對該經編碼的訊框的取樣進行去量化。

【請求項 50】根據請求項 46 之裝置，其中該處理器被配置為：

對該經編碼的訊框的取樣進行熵解碼。

【請求項 51】根據請求項 46 之裝置，其中該處理器被配置為：

將該輸出訊框儲存在記憶體中。

【請求項 52】根據請求項 46 之裝置，其中該處理器被配置為：

顯示該輸出訊框。

【請求項 53】根據請求項 46 之裝置，其中該處理器被配置為進行以下操作：

由該神經網路系統的一編碼器子網路的一第一迴旋層產生與訊框的一亮度通道相關聯的輸出值；

由該編碼器子網路的一第二迴旋層產生與該訊框的至少一個色度通道相關聯的輸出值；

由該編碼器子網路的一第三迴旋層基於與該訊框的該亮度通道相關聯的輸出值和與該訊框的該至少一個色度通道相關聯的輸出值來產生該訊框的一組合表示；及

基於該訊框的該組合表示來產生該經編碼的訊框。

【請求項 54】根據請求項 53 之裝置，其中該編碼器子網

路的該第三迴旋層包括一 1×1 迴旋層，該 1×1 迴旋層包括一或多個 1×1 迴旋濾波器。

【請求項 55】根據請求項 53 之裝置，其中該處理器被配置為進行以下操作：

使用該編碼器子網路的一第一非線性層來處理與該訊框的該亮度通道相關聯的輸出值；及

使用該編碼器子網路的一第二非線性層來處理與該訊框的該至少一個色度通道相關聯的輸出值；

其中該組合表示是基於該第一非線性層的輸出和該第二非線性層的一輸出來產生的。

【請求項 56】根據請求項 55 之裝置，其中該訊框的該組合表示是由該編碼器子網路的該第三迴旋層使用該第一非線性層的輸出和該第二非線性層的一輸出作為輸入來產生的。

【請求項 57】根據請求項 46 之裝置，其中該經編碼的訊框包括一經編碼的視訊訊框。

【請求項 58】根據請求項 57 之裝置，其中該至少一個色度通道包括一色度藍色通道和一色度紅色通道。

【請求項 59】根據請求項 46 之裝置，其中該經編碼的訊框具有一亮度 - 色度 (YUV) 格式。

【請求項 60】根據請求項 46 之裝置，亦包括：一顯示器和被配置為擷取一或多個視訊訊框的一照相機中的至少一者。

【發明圖式】

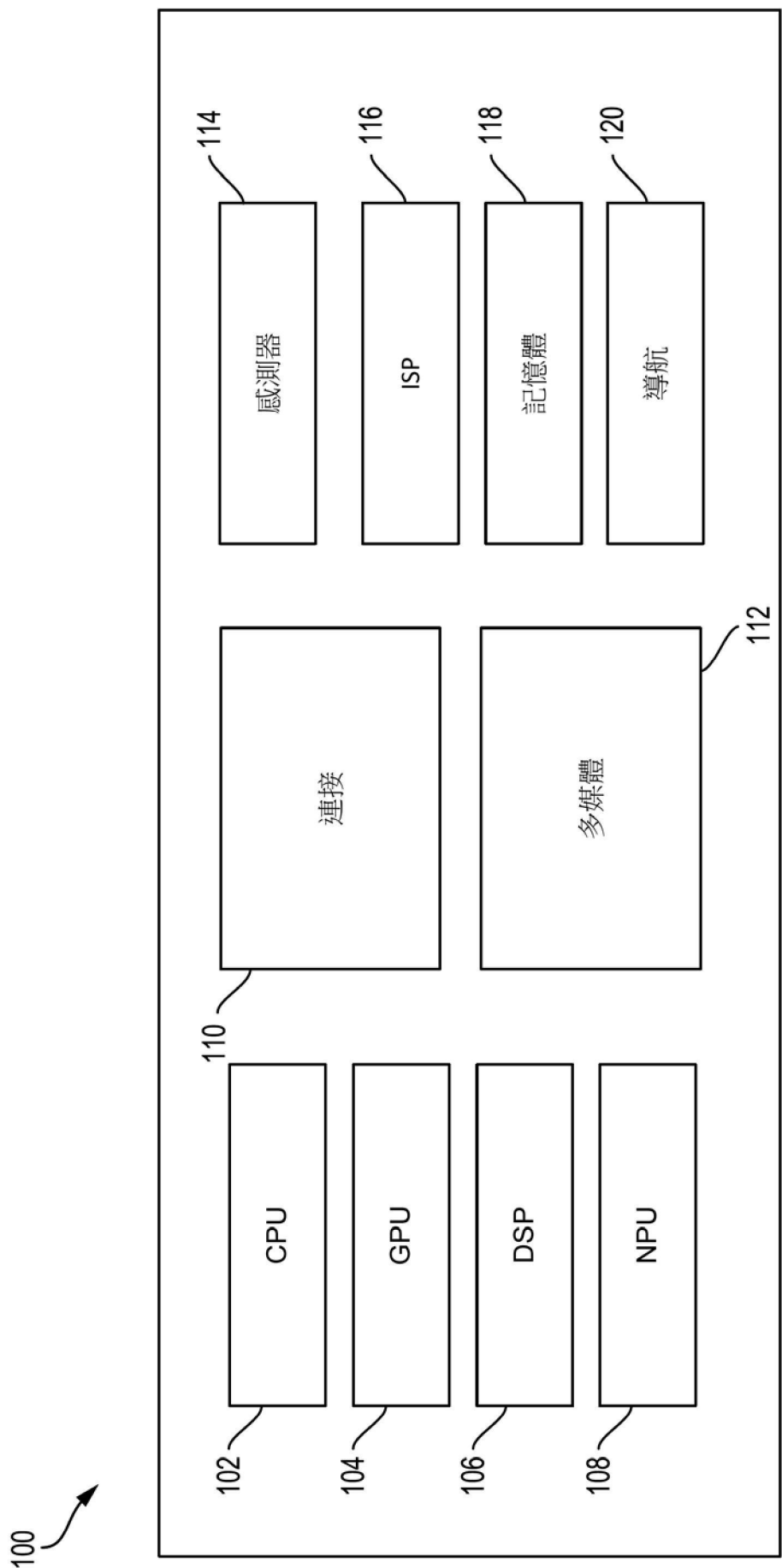


圖1

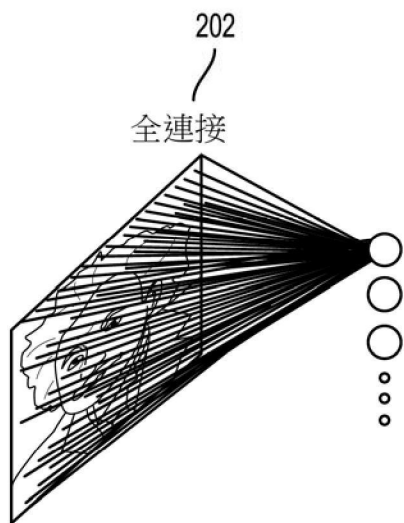


圖2A

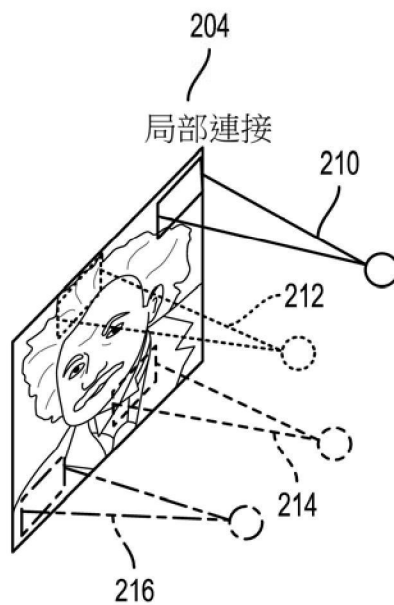


圖2B

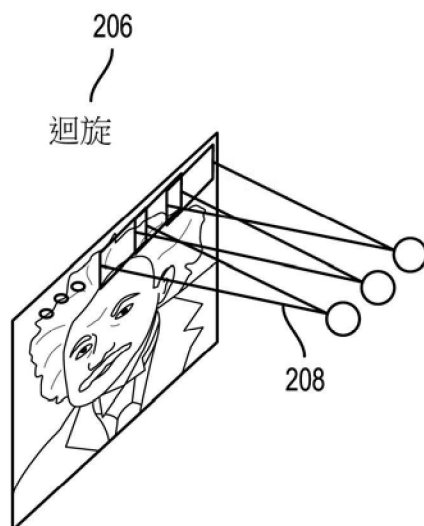


圖2C

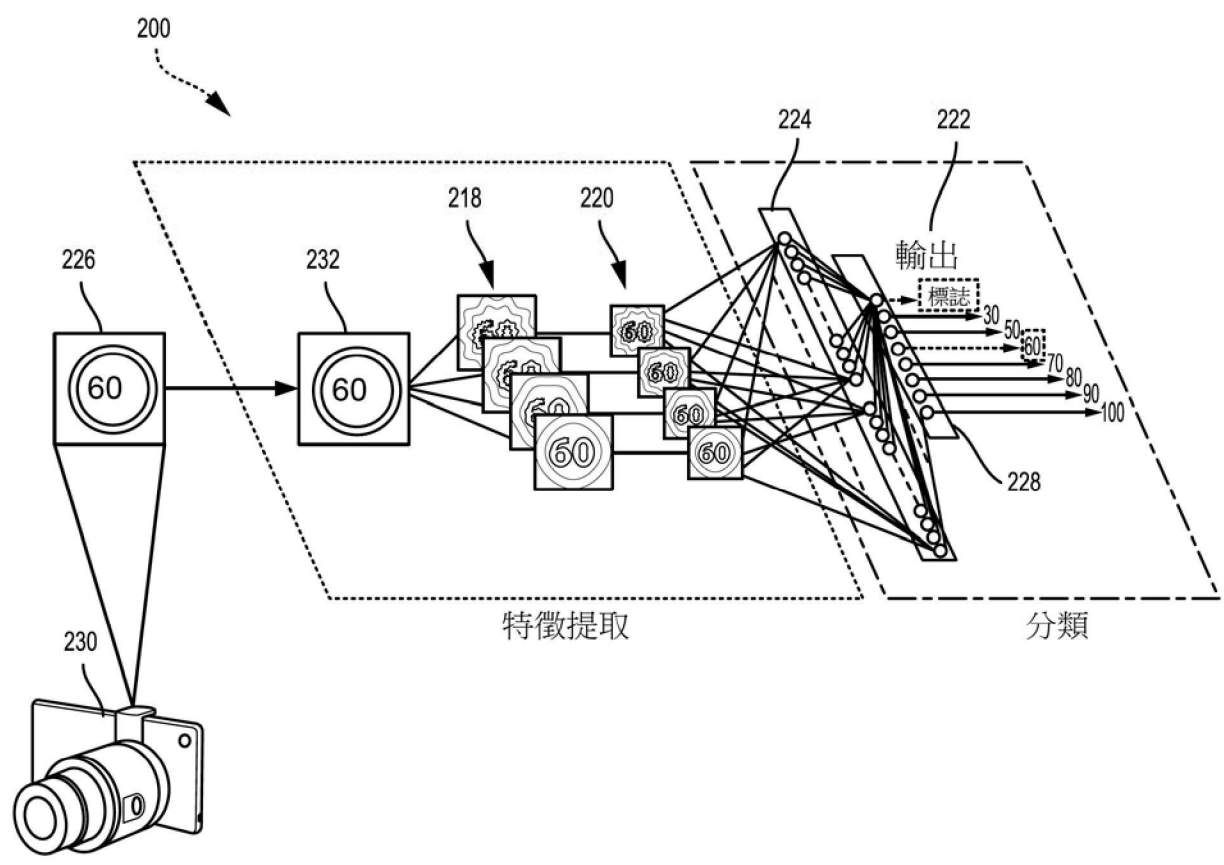


圖2D

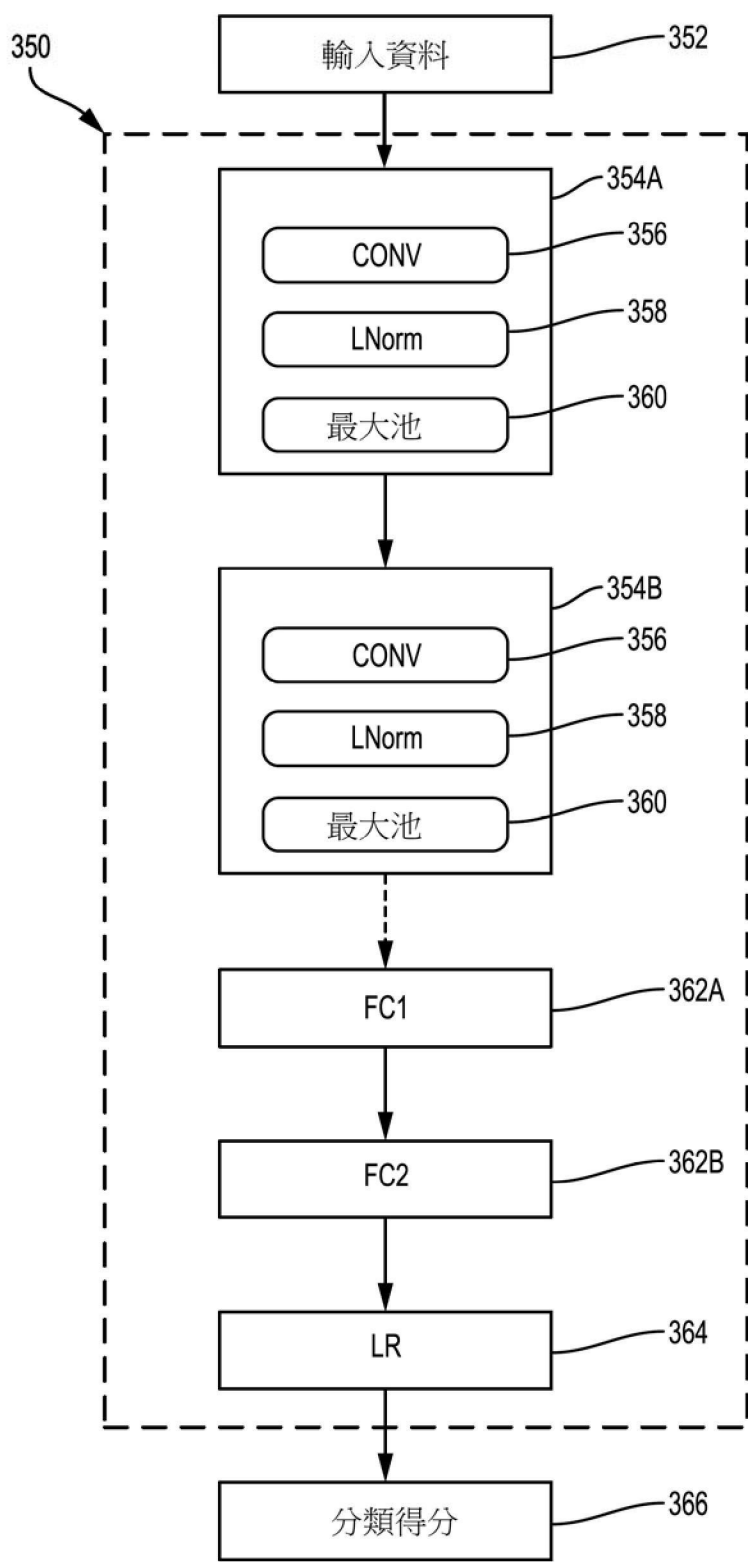


圖3

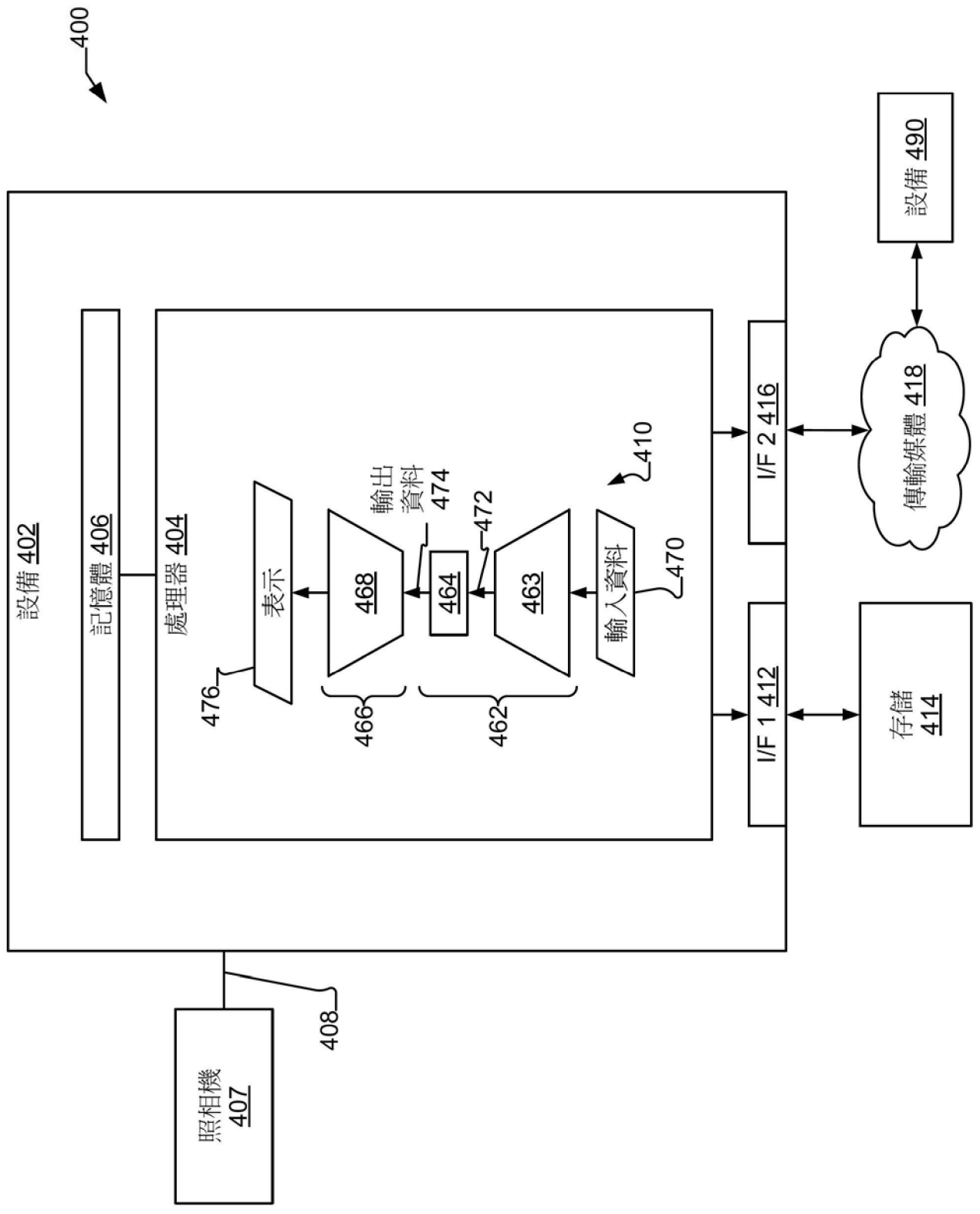


圖4

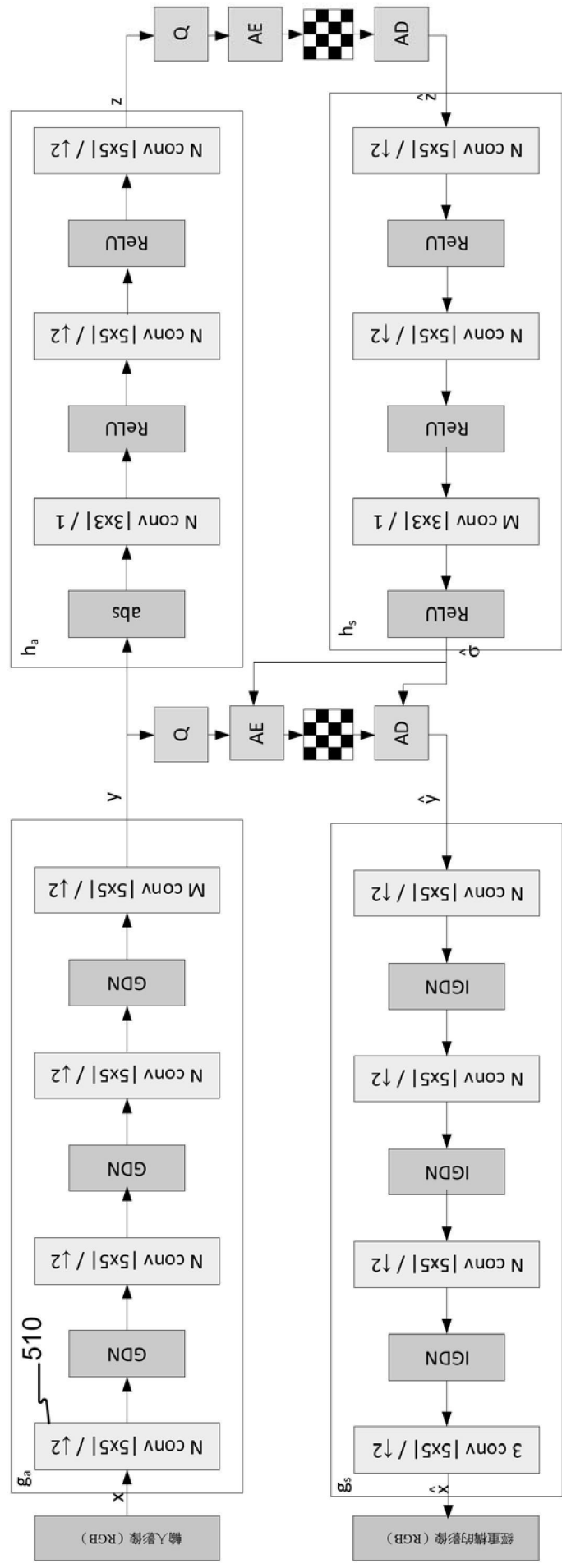


圖5

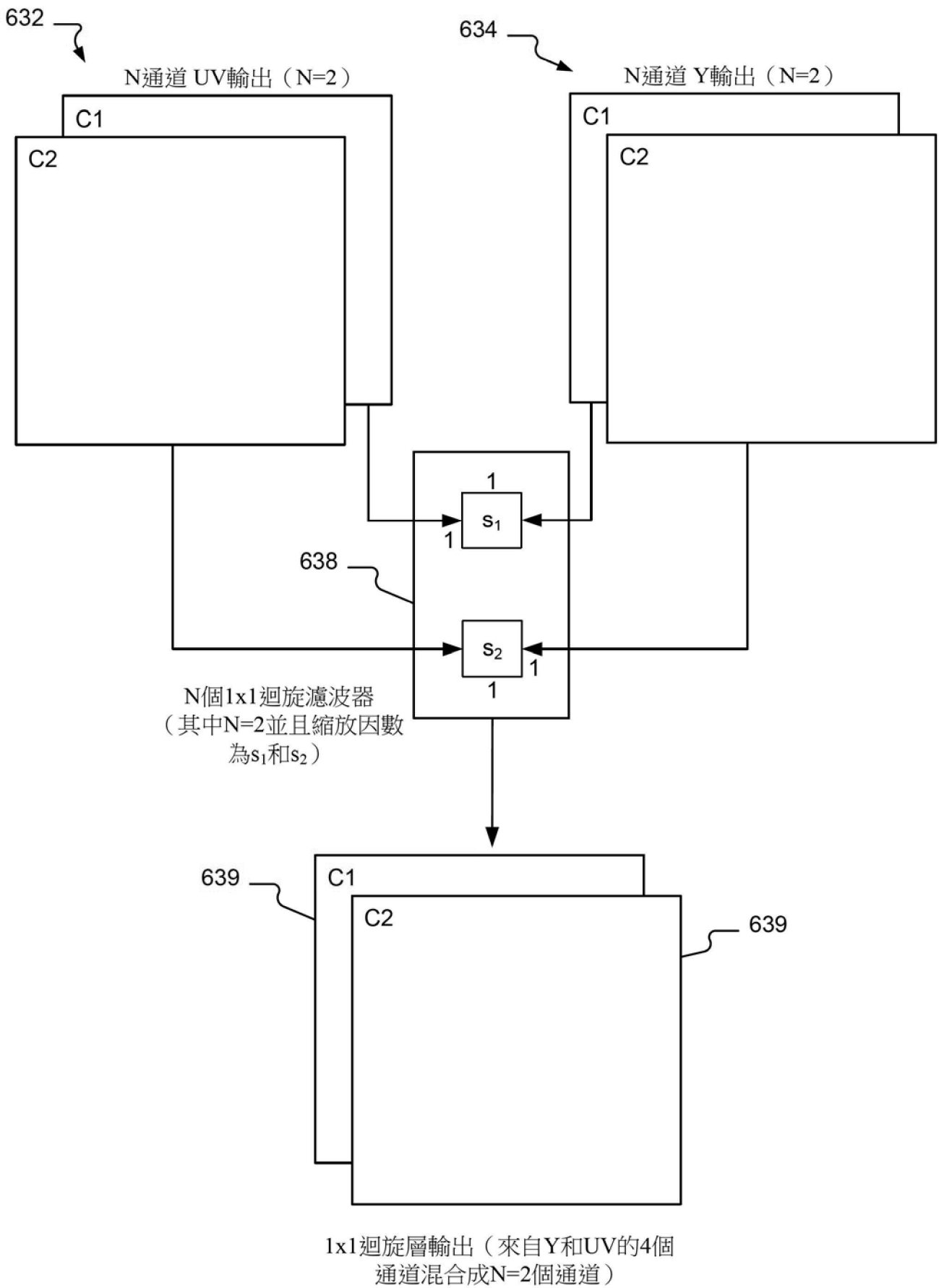


圖6B

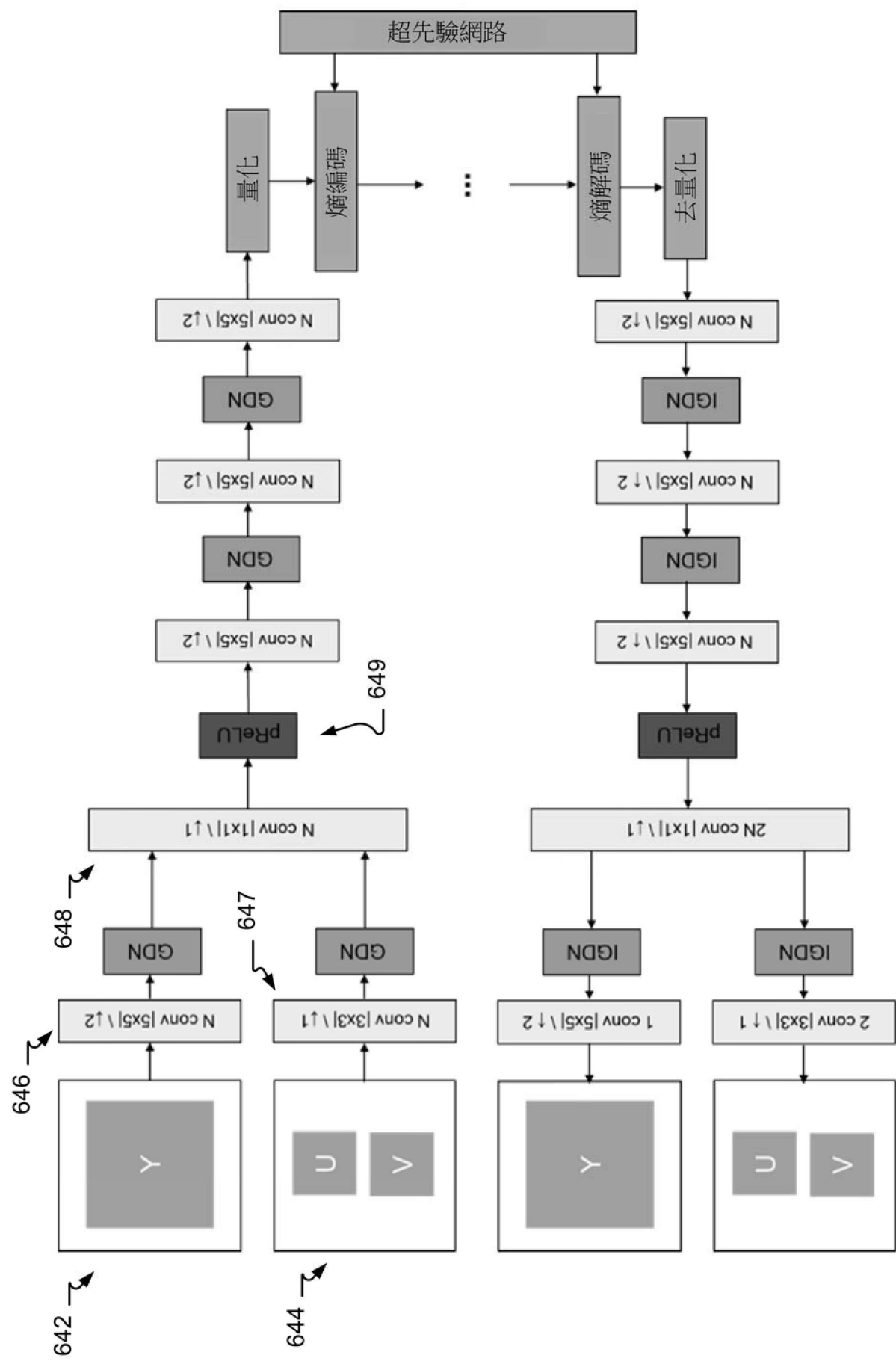


圖6C

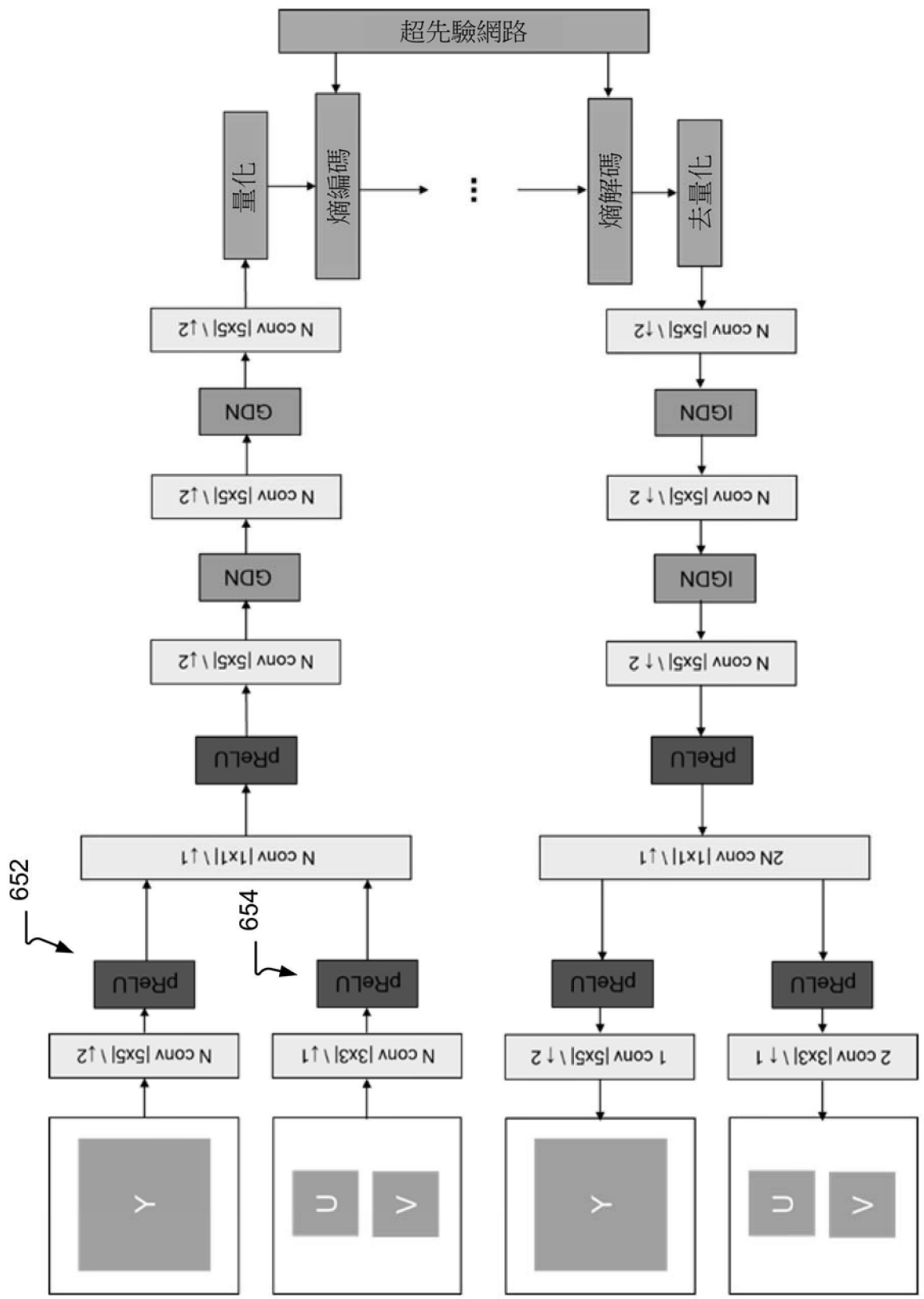


圖6D

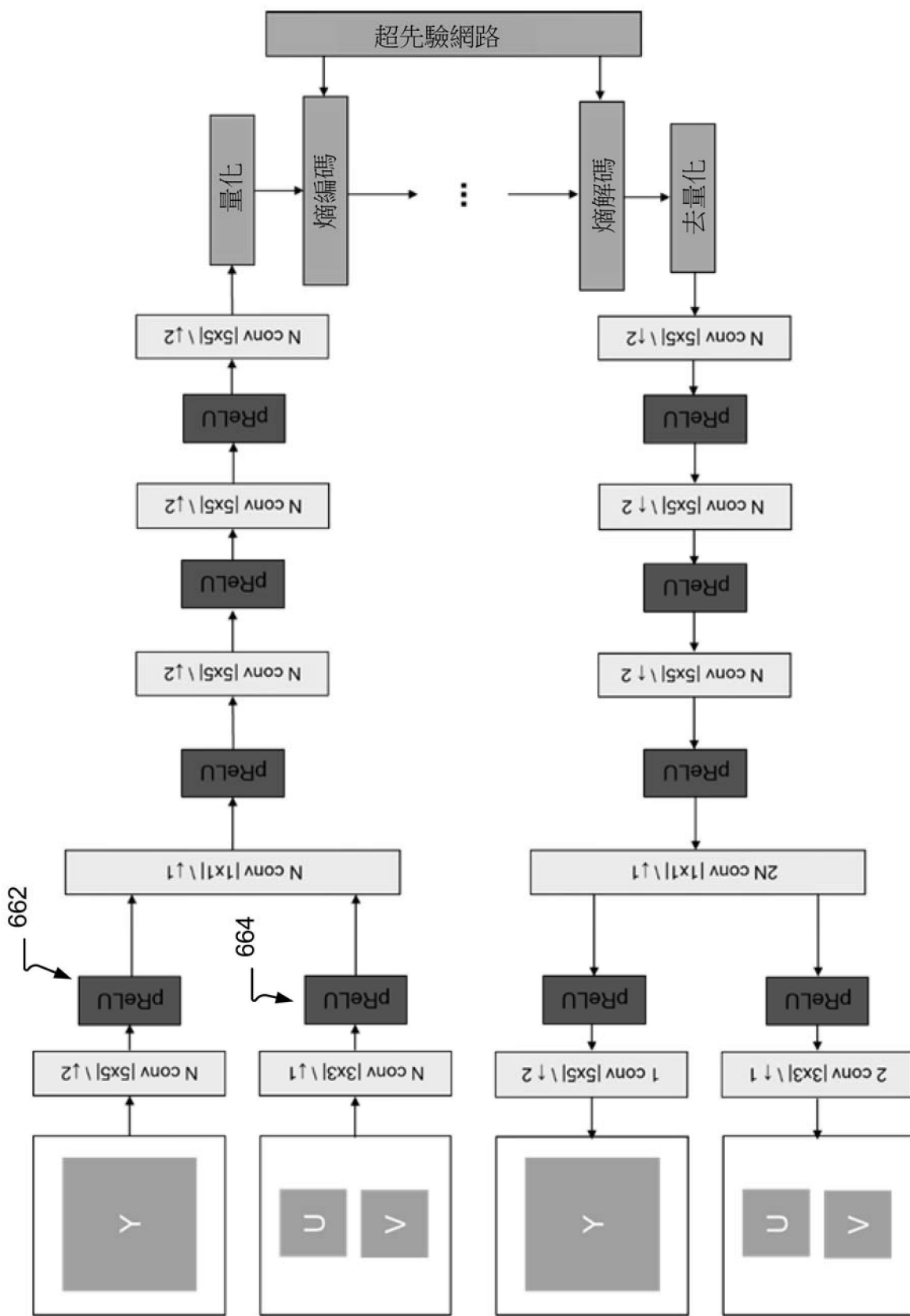


圖6E

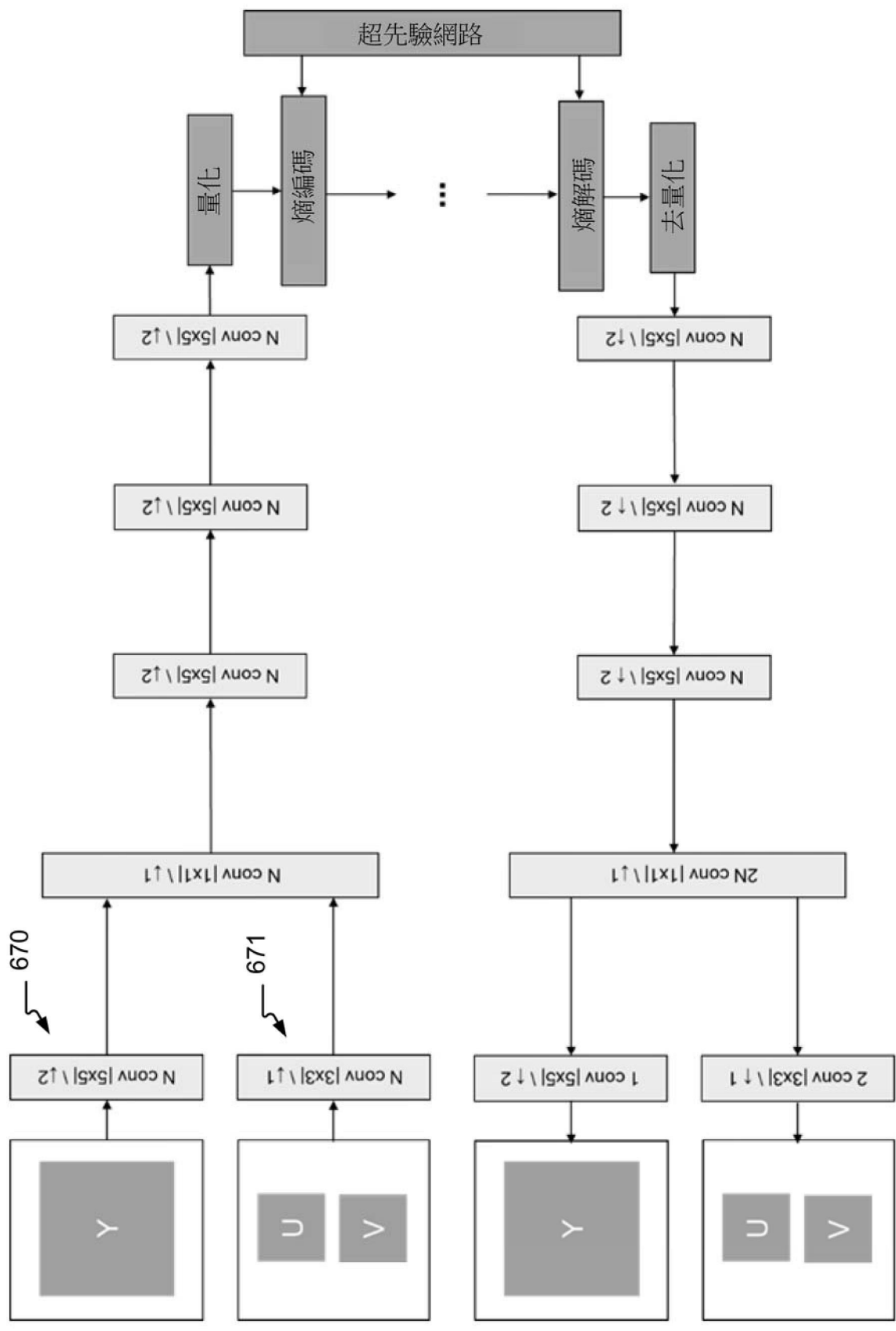


圖6F

700

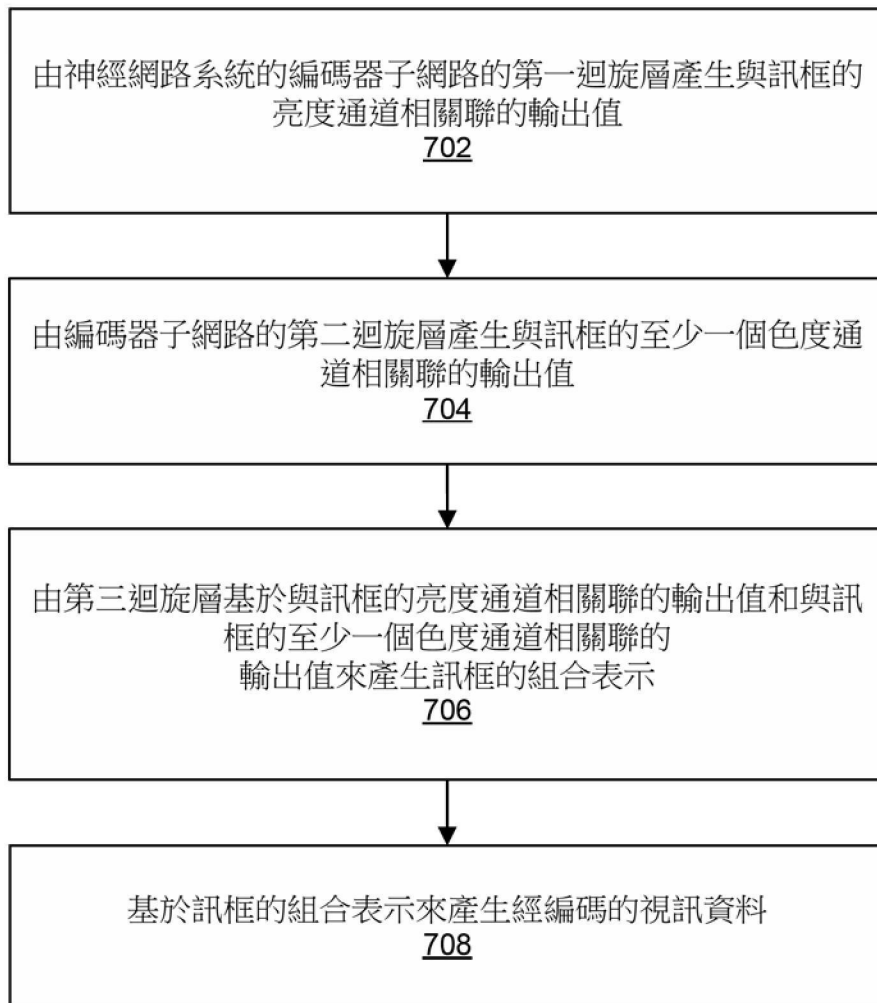


圖7

800

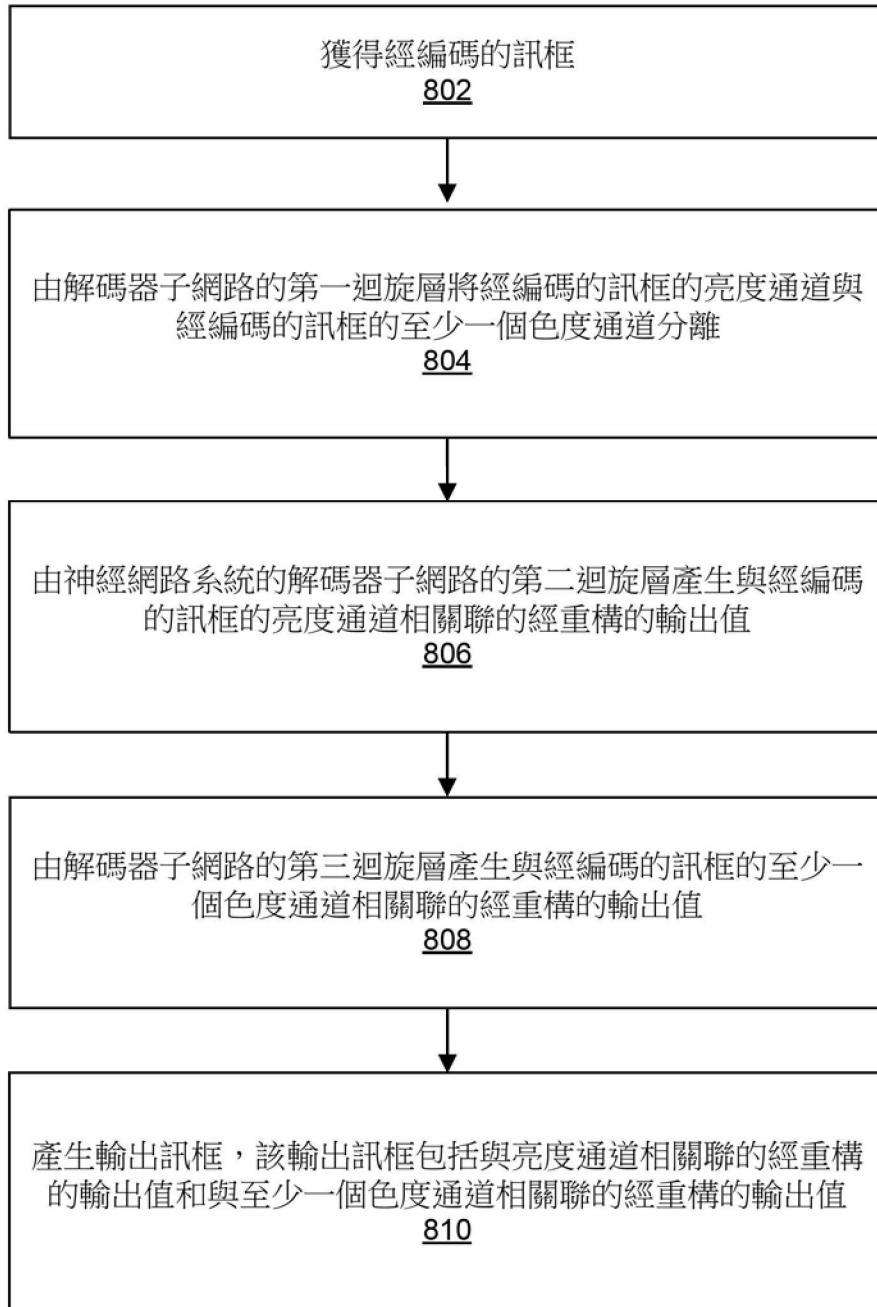


圖8

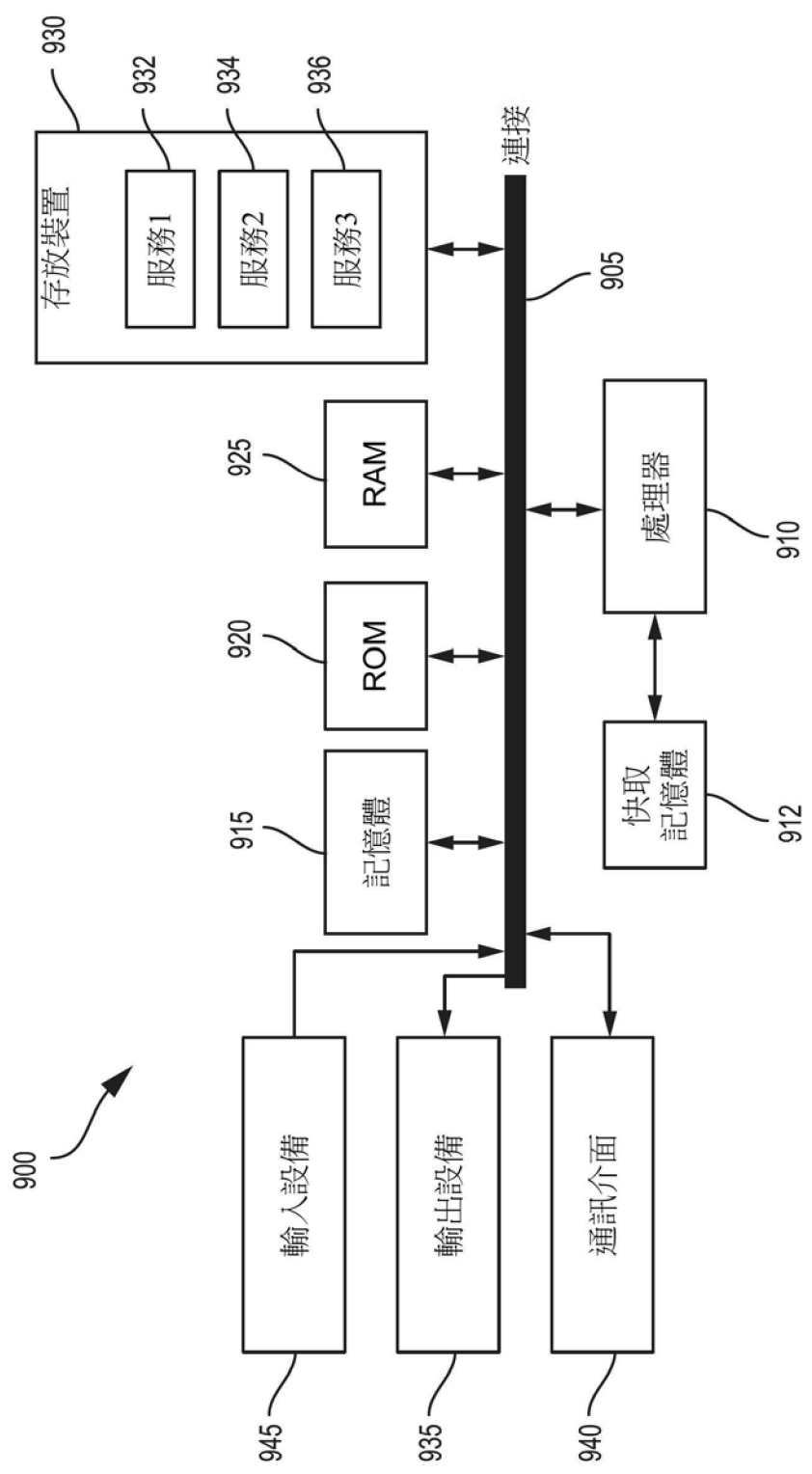


圖9