



(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 I661363 B

(45)公告日：中華民國 108 (2019) 年 06 月 01 日

(21)申請案號：106132484

(22)申請日：中華民國 106 (2017) 年 09 月 21 日

(51)Int. Cl. : G06K9/78 (2006.01)

G10L25/63 (2013.01)

G06F17/30 (2006.01)

G06N3/08 (2006.01)

(30)優先權：2017/08/05 中國大陸

201710663432.4

(71)申請人：鴻海精密工業股份有限公司 (中華民國) HON HAI PRECISION INDUSTRY CO., LTD. (TW)

新北市土城區自由街2號

(72)發明人：張學琴 ZHANG, XUE-QIN (CN)

(56)參考文獻：

TW I365416B

TW I447660B

TW I456515B

CN 104484656A

審查人員：李榮祥

申請專利範圍項數：項 圖式數： 共頁

(54)名稱

智慧型機器人及人機交互方法

SMART ROBOT AND HUMAN-COMPUTER INTERACTION METHOD

(57)摘要

本發明涉及一種智慧型機器人及人機交互方法。該方法應用在該智慧型機器人中。該方法包括步驟：從獲取的語音資訊及圖像資訊中識別一目標物件；確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯；從該目標物件對應的相關事件資訊中提取關鍵資訊；根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及該目標物件的情緒類別確定出一回饋模型；及執行該回饋模型。

The present invention relates to a smart robot and a human-computer interaction method. The method is applied in the smart robot. The method includes: identifying a target object from a voice information and an image information; determining a basic information of the target object and a relation event information of the target object, and binding the basic information with the relation event information; extracting a key information from the relation event information of the target object; determining an emotional category of the target object according to the extracted key information by a neural network analysis algorithm, retrieving a preset public basic knowledge base according to the key information, and determining a feedback model according to the retrieval result, the emotional category of the target object by a depth learning algorithm; and performing the feedback.

指定代表圖：

符號簡單說明：

S501~S506 . . . 步驟

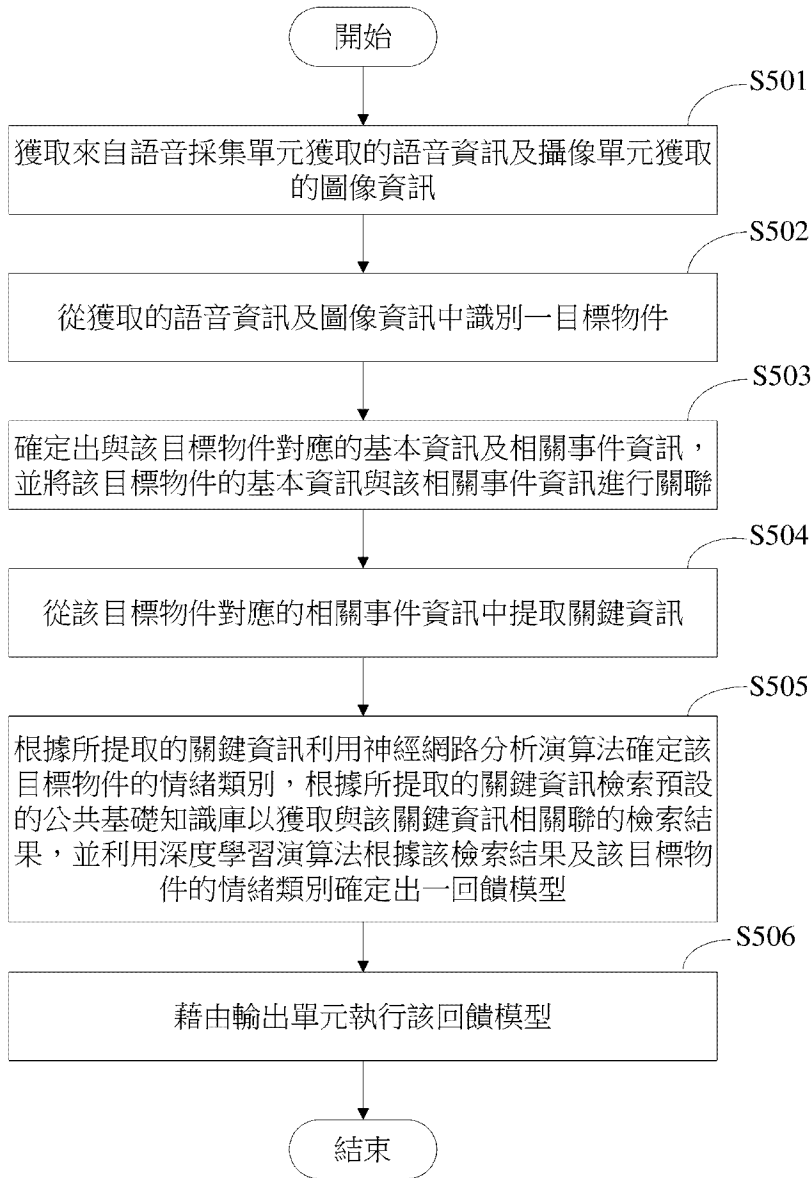


圖5

## 【發明說明書】

【中文發明名稱】 智慧型機器人及人機交互方法

【英文發明名稱】 SMART ROBOT AND HUMAN-COMPUTER INTERACTION  
METHOD

### 【技術領域】

【0001】本發明涉及機器人領域，尤其涉及一種智慧型機器人及人機交互方法。

### 【先前技術】

【0002】現有技術中，機器人與人之間的交互主要涉及簡單的人機對話，或在人的指令下完成特定的工作內容。然而，能夠使機器人從社會的倫理道德、法律法規、天文地理及人際關係（如，家庭關係、同事關係、朋友關係）等方面深入地理解用戶的情感和情緒的問題並與用戶進行交互的方案卻很少。

### 【發明內容】

【0003】鑒於以上內容，有必要提供一種智慧型機器人及人機交互方法以深入地理解用戶的情感和情緒的問題並與用戶進行交互。

【0004】一種智慧型機器人，包括攝像單元、語音採集單元、輸出單元及處理單元，該處理單元用於：

【0005】獲取來自該語音採集單元獲取的語音資訊及該攝像單元獲取的圖像資訊；

【0006】從獲取的語音資訊及圖像資訊中識別一目標物件；

【0007】確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯；

【0008】從該目標物件對應的相關事件資訊中提取關鍵資訊；

【0009】根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果及該目標物件的情緒類別確定出一回饋模型，其中，該回饋模型是指用於控制該智慧型機器人與該目標物件進行交互的指令集；及

【0010】藉由該輸出單元執行該回饋模型。

【0011】優選地，該處理單元從該語音資訊中識別一聲紋特徵及從該圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別對應的目標物件。

【0012】優選地。該處理單元控制該語音採集單元採集該目標物件的語音資訊作為該目標物件的相關事件資訊。

【0013】優選地，該處理單元識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，並將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。

【0014】優選地，該處理單元控制該攝像單元獲取該目標物件的圖片資訊作為該目標物件的相關事件資訊。

【0015】優選地，該處理單元獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

【0016】優選地，該處理單元還用於設定該智慧型機器人的情感方向，該處理單元從該目標物件的相關事件資訊中提取關鍵資訊，根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果、該目標物件的情緒類別及設定的該智慧型機器人的情感方向確定出對應的回饋模型。

【0017】優選地，該輸出單元包括語音輸出單元及表情輸出單元，該執行模組藉由控制該語音輸出單元輸出語音資訊，及藉由該表情輸出單元輸出表情動作的方式執行該回饋模型。

【0018】一種人機交互方法，應用在一智慧型機器人中，該方法包括步驟：

【0019】獲取一語音採集單元獲取的語音資訊；

【0020】獲取一攝像單元獲取的圖像資訊；

【0021】從獲取的語音資訊及圖像資訊中識別一目標物件；

【0022】確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯；

【0023】從該目標物件對應的相關事件資訊中提取關鍵資訊；

【0024】根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果及該目標物件的情緒類別確定出一回饋模型，其中，該回饋模型是指用於控制該智慧型機器人與該目標物件進行交互的指令集；及

【0025】藉由一輸出單元執行該回饋模型。

【0026】優選地，該方法在步驟“從獲取的語音資訊及圖像資訊中識別一目標物件”還包括：

【0027】從該語音資訊中識別一聲紋特徵及從該圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別對應的目標物件。

【0028】優選地，該方法在步驟“確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯”中包括：

【0029】控制該語音採集單元採集該目標物件的語音資訊作為該目標物件的相關事件資訊。

【0030】優選地，該方法還包括步驟：

【0031】識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，並將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。

【0032】優選地，該方法在步驟“確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯”中包括：

【0033】控制該攝像單元獲取該目標物件的圖片資訊作為該目標物件的相關事件資訊。

【0034】優選地，該方法還包括步驟：

【0035】獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

【0036】優選地，該方法還包括步驟：

【0037】藉由控制一語音輸出單元輸出語音資訊，及藉由一表情輸出單元輸出表情動作的方式執行該回饋模型。

【0038】本案中的智慧型機器人及人機交互方法能夠深入地理解用戶的情感和情緒的問題並與用戶進行交互，提高了用戶的體驗感。

#### 【圖式簡單說明】

##### 【0039】

圖 1 為本發明一實施方式中人機交互系統的應用環境圖。

圖 2 為本發明一實施方式中智慧型機器人的功能模組圖。

圖 3 為本發明一實施方式中人機交互系統的功能模組圖。

圖 4 為本發明一實施方式中基本資訊表的示意圖。

圖 5 為本發明一實施方式中人機交互方法的流程圖。

**【實施方式】**

**【0040】**請參考圖 1，所示為本發明一實施方式中人機交互系統 1 的應用環境圖。該人機交互系統 1 應用在一智慧型機器人 2 中。該智慧型機器人 2 與一伺服器 3 通訊連接。該人機交互系統 1 用於控制該智慧型機器人 2 與用戶進行交互。

**【0041】**請參考圖 2，所示為本發明一實施方式中智慧型機器人 2 的功能模組圖。該智慧型機器人 2 包括攝像單元 22、語音採集單元 23、壓力感測器 24、輸出單元 25、味道感測器 26、通訊單元 27、處理單元 28 及存儲單元 29。該攝像單元 22 攝取智慧型機器人 2 周圍環境的圖像並將攝取的圖像傳送給處理單元 28。例如，該攝像單元 22 可以攝取智慧型機器人 2 周圍的人、動物或靜止物體的畫面，及將獲取的人、動物或靜止物體的畫面傳送給該處理單元 28。本實施方式中，該攝像單元 22 可以為一攝像頭、3D 光場相機等。該語音採集單元 23 用於接收智慧型機器人 2 周圍的語音資訊並將接收的語音資訊傳送給處理單元 28。在本實施方式中，該語音採集單元 23 可以為麥克風、麥克風陣列等。該壓力感測器 24 用於檢測用戶對該智慧型機器人 2 的按壓力資訊及將檢測出的按壓力資訊傳送給處理單元 28。

**【0042】**該輸出單元 25 包括語音輸出單元 251、表情輸出單元 252、運動驅動單元 253 及顯示單元 254。該語音輸出單元 251 用於在該處理單元 28 的控制下輸出語音資訊。在本實施方式中，該語音輸出單元 251 可以為揚聲器。表情輸出單元 252 用於在該處理單元 28 的控制下輸出表情動作。在一實施方式中，該表情輸出單元 252 包括設於機器人頭部可開合的眼簾和嘴巴及設於眼簾內可轉動的眼球。該運動驅動單元 253 用於在該處理單元 28 的控制下控制該智慧型機器人 2 的控制下驅動該智慧型機器人 2 移動。在一實施方式中，該運動驅動單元 253 包括兩軸或四軸驅動輪。該顯示單元 254 用於顯示表情圖像，如高興、苦惱、憂鬱表情等。該味道感測器 26 用於檢測氣味資訊。該通訊單元 27 用於供該

智慧型機器人 2 與一伺服器 3 (如圖 1 所示) 通訊連接。在一實施方式中，該通訊單元 27 可以為 WIFI 通訊模組、Zigbee 通訊模組及 Blue Tooth 通訊模組。

【0043】該存儲單元 29 用於存儲該智慧型機器人 2 的程式碼及資料資料。例如，該存儲單元 29 可以存儲預設人臉圖像、預設語音及人機交互系統 1。本實施方式中，該存儲單元 29 可以為該智慧型機器人 2 的內部存儲單元，例如該智慧型機器人 2 的硬碟或記憶體。在另一實施方式中，該存儲單元 29 也可以為該智慧型機器人 2 的外部存放裝置，例如該智慧型機器人 2 上配備的插接式硬碟，智慧存儲卡 (Smart Media Card, SMC)，安全數位 (Secure Digital, SD) 卡，快閃記憶體卡 (Flash Card) 等。

【0044】本實施方式中，該處理單元 28 可以為一中央處理器 (Central Processing Unit, CPU)，微處理器或其他資料處理晶片，該處理單元 28 用於執行軟體程式碼或運算資料。

【0045】請參考圖 3，所示為本發明一實施方式中人機交互系統 1 的功能模組圖。本實施方式中，該人機交互系統 1 包括一個或多個模組，所述一個或者多個模組被存儲於該存儲單元 29 中，並被該處理單元 28 所執行。人機交互系統 1 包括感知模組 101、識別模組 102、分析模組 103、執行模組 104 及設置模組 105。在其他實施方式中，該人機交互系統 1 為內嵌在該智慧型機器人 2 中的程式段或代碼。

【0046】該感知模組 101 獲取來自該語音採集單元 23 獲取的語音資訊及該攝像單元 22 獲取的圖像資訊。

【0047】該識別模組 102 用於從獲取的語音資訊及圖像資訊中識別一目標物件。在一實施方式中，該識別模組 102 對獲取的語音資訊進行語音信號預處理，例如進行去噪處理，使得語音辨識時更加準確。在一實施方式中，該識別模組 102 從語音資訊中識別一聲紋特徵及從圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別該目標物件。該目標物件包括人及動物。例如，該存儲單元 29 中存儲一第一對應關係表 (圖中未示)，該第一對應關係表中定



義了聲紋特徵、臉部特徵及目標物件的對應關係，該識別模組 102 根據識別出的聲紋特徵、臉部特徵及該第一對應關係表確定該目標物件。在另一實施方式中，該識別模組 102 也可以僅從獲取的語音資訊中識別該目標物件。例如，該存儲單元 29 中存儲一第二對應關係表（圖中未示），該第二對應關係表中定義了聲紋特徵及目標物件的對應關係。該識別模組 102 根據識別出的聲紋特徵及該第二對應關係表確定該目標物件。在其他實施方式中，該識別模組 102 也可以僅從獲取的圖像資訊中識別該目標物件。例如，該存儲單元 29 中存儲一第三對應關係表（圖中未示）。該第三對應關係表中定義該臉部特徵與該目標物件的對應關係，該識別模組 102 根據識別出的臉部特徵及該第三對應關係表確定該目標物件。在一實施方式中，該識別模組 102 將獲取的圖像與存儲在存儲單元 29 中的圖片集進行比對確定出獲取的圖像中的臉部特徵，該識別模組 102 根據識別出的臉部特徵及該第三對應關係表確定該目標物件。

**【0048】**在一實施方式中，該目標物件還包括靜止物體，例如桌子，椅子，房子等物體。該識別模組 102 還用於將藉由從該攝像單元 22 中獲取的圖片與一存儲在存儲單元 29 中的圖片集進行比對，並根據比對結果確定出該圖片中包含的靜止物體。其中，該圖片集中的每一幅圖片對應包含一靜止物體。例如，該識別模組 102 將獲取的圖片與存儲的圖片集進行比對後確定該圖片中的物體與存儲的圖片集中一幅圖片所包含的房子一致時，該識別模組 102 識別出該獲取的圖片中的目標物件為房子。

**【0049】**該識別模組 102 還用於確定出與該目標物件相關的資訊，例如，基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯。在一實施方式中，所述相關的資訊包括，但不限於，該目標物件的基本資訊、及與該目標物件相關的事件資訊。在一實施方式中，當目標物件為人時，該目標物件的基本資訊包括，但不限於，用戶的姓名、年齡、身高、體重、體型（例如，大體型、中體型、小體型）。當目標物件為物體時，該目標物件的基本資訊包括，但不限於，目標物件的名稱、位置、功能屬性。例如，該存

儲單元 29 存儲一基本資訊表 T1 (參考圖 4)，該基本資訊表 T1 定義了目標物件與基本資訊的對應關係。該識別模組 102 根據該目標物件及該基本資訊表 T1 確定出與該目標物件對應的基本資訊。

**【0050】** 在一實施方式中，與該目標物件相關的事件資訊可以是指發生在某個時間或某個地點的與該目標物件相關的事件。在一實施方式中，當識別出目標物件時，該識別模組 102 控制該語音採集單元 23 採集該目標物件的語音資訊作為該目標物件的相關事件資訊。在另一實施方式中，當識別出目標物件時，該識別模組 102 控制該攝像單元 22 獲取該目標物件的圖片資訊作為相關事件資訊。在其他實施方式中，當識別出目標物件時，該識別模組 102 同時將藉由語音採集單元 23 採集的語音資訊及藉由攝像單元 22 攝取的圖片資訊作為該目標物件的相關事件資訊。

**【0051】** 該分析模組 103 用於從該目標物件對應的相關事件資訊中提取關鍵資訊。例如，當該目標物件的相關事件資訊為語音資訊時，該分析模組 103 識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，及將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。本實施方式中，該文本資料中的關鍵資訊包括關鍵字、關鍵字、或關鍵語句。當該當目標物件的相關事件資訊為圖片資訊時，該分析模組 103 獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

**【0052】** 該分析模組 103 根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果及該目標物件的情緒類別確定出一回饋模型。本實施方式中，該回饋模型是指用於控制該智慧型機器人 2 與該目標物件進行交互的指令集。本實施方式

中，該公共基礎知識庫可以包括，但不限於人文倫理知識庫、法律法規知識庫、道德情操知識庫、宗教知識庫、天文地理知識庫。在一實施方式中，該公共基礎知識庫存儲在該智慧型機器人 2 的存儲單元 29 中。該智慧型機器人 2 可以直接訪問該存儲單元 29 中的公共基礎知識庫。在其他實施方式中，該公共基礎知識庫存儲在伺服器 3 中。該智慧型機器人 2 藉由通訊單元 27 訪問該伺服器 3 中的公共基礎知識庫。

**【0053】**本實施方式中，該深度學習演算法包括，但不限於，“神經詞袋模型”、“遞迴神經網路”、“迴圈神經網路”、“卷積神經網路”。本實施方式中，該目標物件的情緒類別包括高興、悲傷、憤怒、平和、暴躁等情緒。例如，當用戶微笑著對智慧型機器人 2 說“這些花真漂亮啊！”時，該感知模組 101 藉由該語音採集單元 23 獲取使用者的語音資訊，及藉由該攝像單元 22 獲取包含使用者的圖像資訊。該識別模組 102 根據使用者的語音資訊識別出使用者的聲紋特徵及根據使用者的圖像資訊識別出使用者的臉部特徵。該識別模組 102 根據該識別出的聲紋特徵及臉部特徵識別出目標物件為使用者。該識別模組 102 確定出使用者的基本資訊，及將使用者發出的語音資訊“這些花真漂亮啊！”及使用者微笑的圖像資訊作為使用者的相關事件資訊，並將使用者的基本資訊與使用者的相關事件資訊進行關聯。該分析模組 103 從使用者發出的語音資訊“這些花真漂亮啊！”提取出關鍵資訊為“花”、“漂亮”，及從使用者微笑的圖像資訊中提取出關鍵資訊為“微笑表情”。該分析模組 103 根據所提取的該些關鍵資訊“花、漂亮、微笑表情”利用神經網路分析演算法確定該目標物件的情緒類別為高興。該分析模組 103 根據所提取的上述關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及高興的情緒類別確定出對應的回饋模型。

**【0054】**在一實施方式中，該設置模組 105 用於設定該智慧型機器人 2 的情感方向。在一實施方式中，該設定的情感方向包括勵志型、安慰型、諷刺型、幽默型等。該分析模組 103 從該目標物件的相關事件資訊中提取關鍵資訊，根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根

據所提取的關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果、該目標物件的情緒類別及該設定的智慧型機器人 2 的情感方向確定出對應的回饋模型。本實施方式中，該情緒類別包括，但不限於，高興、憤怒、悲傷、愉快等情緒。

**【0055】** 在一實施方式中，該感知模組 101 還用於藉由該壓力感測器 24 感測用戶輸入的按壓力，及藉由該味道感測器 26 檢測該智慧型機器人 2 周圍環境的氣味資訊。該分析模組 103 還用於根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果、該目標物件的情緒類別、該感測的按壓力及該感測的氣味資訊確定出對應的回饋模型。

**【0056】** 該執行模組 104 用於藉由該輸出單元 25 執行該回饋模型。在一實施方式中，該執行模組 104 藉由該輸出單元 25 輸出語音資訊及輸出表情動作的方式來執行該回饋模型以實現該智慧型機器人 2 與目標物件之間的交互。具體的，該執行模組 104 藉由語音輸出單元 251 輸出語音資訊、藉由表情輸出單元 252 輸出表情動作的方式執行該回饋模型。例如，當用戶微笑著對智慧型機器人 2 說“這些花真漂亮啊！”時，該識別模組 102 識別出該目標物件為使用者。該分析模組 103 從使用者發出的語音資訊“這些花真漂亮啊！”提取出關鍵資訊為“花”、“漂亮”，及從使用者微笑的圖像資訊中提取出關鍵資訊為“微笑表情”。該分析模組 103 根據所提取的該些關鍵資訊“花、漂亮、微笑表情”利用神經網路分析演算法確定該目標物件的情緒類別為高興。該分析模組 103 根據所提取的關鍵資訊“花、漂亮、微笑表情”檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及高興的情緒類別確定出對應的回饋模型。該回饋模型為控制該智慧型機器人 2 輸出語音資訊“這些花真的很漂亮，我也很喜歡！”及輸出笑容的表情動作的指令。該執行模組 104 藉由語音輸出單元 251 輸出“這些花真的很漂亮，我也很喜歡！”的語音資訊及藉由表情輸出單元 252 控制設於智慧型機器

人 2 頭部中的眼簾和嘴巴開合及設於眼簾內的眼球轉動輸出笑容表情動作，從而實現該智慧型機器人 2 與用戶進行交互。

【0057】在其他實施方式中，該執行模組 104 還藉由運動驅動單元 253 控制該智慧型機器人 2 以預設移動的方式及控制該顯示單元 254 顯示一預設表情圖像的方式來執行該回饋模型以實現該智慧型機器人 2 與目標物件之間的交互。

【0058】請參考圖 5，所示為本發明一實施方式中人機交互方法的流程圖。該方法應用在智慧型機器人 2 中。根據不同需求，該流程圖中步驟的順序可以改變，某些步驟可以省略或合併。該方法包括如下步驟。

【0059】S501：獲取來自語音採集單元 23 獲取的語音資訊及攝像單元 22 獲取的圖像資訊。

【0060】S502：從獲取的語音資訊及圖像資訊中識別一目標物件。

【0061】在一實施方式中，該智慧型機器人 2 對獲取的語音資訊進行語音信號預處理，例如進行去噪處理，使得語音辨識時更加準確。在一實施方式中，該智慧型機器人 2 從語音資訊中識別一聲紋特徵及從圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別該目標物件。該目標物件包括人及動物。在一實施方式中，該目標物件還包括靜止物體，例如桌子，椅子，房子等物體。該智慧型機器人 2 還用於將藉由從該攝像單元 22 中獲取的圖片與一存儲在存儲單元 29 中的圖片集進行比對，並根據比對結果確定出該圖片中包含的靜止物體。

【0062】S503：確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯。

【0063】在一實施方式中，所述相關資訊包括，但不限於，該目標物件的基本資訊、及與該目標物件相關的事件資訊。在一實施方式中，當目標物件為人時，該目標物件的基本資訊包括，但不限於，用戶的姓名、年齡、身高、體重、體型（例如，大體型、中體型、小體型）。當目標物件為物體時，該目標物件的基本資訊包括，但不限於，目標物件的名稱、位置、功能屬性。例如，

該存儲單元 29 存儲一基本資訊表 T1（參考圖 4），該基本資訊表 T1 定義了目標物件與基本資訊的對應關係。該智慧型機器人 2 根據該目標物件及該基本資訊表 T1 確定出與該目標物件對應的基本資訊。

**【0064】** 在一實施方式中，與該目標物件相關的事件資訊可以是指發生在某個時間或某個地點的與該目標物件相關的事件。該智慧型機器人 2 可以控制該語音採集單元 23 採集該目標物件的語音資訊作為該目標物件的相關事件資訊。在另一實施方式中，該智慧型機器人 2 可以控制該攝像單元 22 獲取該目標物件的圖片資訊作為相關事件資訊。在其他實施方式中，該智慧型機器人 2 同時將藉由語音採集單元 23 採集的語音資訊及藉由攝像單元 22 攝取的圖片資訊作為該目標物件的相關事件資訊。

**【0065】** S504：從該目標物件對應的相關事件資訊中提取關鍵資訊。

**【0066】** 例如，當該目標物件的相關事件資訊為語音資訊時，該智慧型機器人 2 識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，及將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。當該當目標物件的相關事件資訊為圖片資訊時，該智慧型機器人 2 獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

**【0067】** S505：根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果及該目標物件的情緒類別確定出一回饋模型。

**【0068】** 本實施方式中，該回饋模型是指用於控制該智慧型機器人 2 與該目標物件進行交互的指令集。本實施方式中，該公共基礎知識庫可以包括，但不限於人文倫理知識庫、法律法規知識庫、道德情操知識庫、宗教至少庫、天

文地理知識庫。在一實施方式中，該公共基礎知識庫存儲在該智慧型機器人 2 的存儲單元 29 中。該智慧型機器人 2 可以直接訪問該存儲單元 29 中的公共基礎知識庫。在其他實施方式中，該公共基礎知識庫存儲在伺服器 3 中。該智慧型機器人 2 藉由通訊單元 27 訪問該伺服器 3 中的公共基礎知識庫。

**【0069】**本實施方式中，該深度學習演算法包括，但不限於，“神經詞袋模型”、“遞迴神經網路”、“迴圈神經網路”、“卷積神經網路”。本實施方式中，該目標物件的情緒類別包括高興、悲傷、憤怒、平和、暴躁等情緒。例如，當用戶微笑著對智慧型機器人 2 說“這些花真漂亮啊！”時，該對智慧型機器人 2 藉由該語音採集單元 23 獲取使用者的語音資訊，及藉由該攝像單元 22 獲取包含使用者的圖像資訊。該對智慧型機器人 2 根據使用者的語音資訊識別出使用者的聲紋特徵及根據使用者的圖像資訊識別出使用者的臉部特徵。該對智慧型機器人 2 根據該識別出的聲紋特徵及臉部特徵識別出目標物件為使用者。該對智慧型機器人 2 確定出使用者的基本資訊，及將使用者發出的語音資訊“這些花真漂亮啊！”及使用者微笑的圖像資訊作為使用者的相關事件資訊，並將使用者的基本資訊與使用者的相關事件資訊進行關聯。該對智慧型機器人 2 從使用者發出的語音資訊“這些花真漂亮啊！”提取出關鍵資訊為“花”、“漂亮”，及從使用者微笑的圖像資訊中提取出關鍵資訊為“微笑表情”。該對智慧型機器人 2 根據所提取的該些關鍵資訊“花、漂亮、微笑表情”利用神經網路分析演算法確定該目標物件的情緒類別為高興。該對智慧型機器人 2 根據所提取的上述關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及高興的情緒類別確定出對應的回饋模型。

**【0070】**S506：藉由輸出單元 25 執行該回饋模型。

**【0071】**在一實施方式中，該智慧型機器人 2 藉由該輸出單元 25 輸出語音資訊及輸出表情動作的方式來執行該回饋模型以實現該智慧型機器人 2 與目標物件之間的交互。具體的，該智慧型機器人 2 藉由語音輸出單元 251 輸出語音資訊、藉由表情輸出單元 252 輸出表情動作的方式執行該回饋模型。例如，當用

戶微笑著對智慧型機器人 2 說“這些花真漂亮啊！”時，該智慧型機器人 2 識別出該目標物件為使用者。該智慧型機器人 2 從使用者發出的語音資訊“這些花真漂亮啊！”提取出關鍵資訊為“花”、“漂亮”，及從使用者微笑的圖像資訊中提取出關鍵資訊為“微笑表情”。該智慧型機器人 2 根據所提取的該些關鍵資訊“花、漂亮、微笑表情”利用神經網路分析演算法確定該目標物件的情緒類別為高興。該智慧型機器人 2 根據所提取的關鍵資訊“花、漂亮、微笑表情”檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及高興的情緒類別確定出對應的回饋模型。該回饋模型為控制該智慧型機器人 2 輸出語音資訊“這些花真的很漂亮，我也很喜歡！”及輸出笑容的表情動作的指令。該智慧型機器人 2 藉由語音輸出單元 251 輸出“這些花真的很漂亮，我也很喜歡！”的語音資訊及藉由表情輸出單元 252 控制設於智慧型機器人 2 頭部中的眼簾和嘴巴開合及設於眼簾內的眼球轉動輸出笑容表情動作，從而實現該智慧型機器人 2 與用戶進行交互。

【0072】在其他實施方式中，該智慧型機器人 2 還藉由運動驅動單元 253 控制該智慧型機器人 2 以預設移動的方式及控制該顯示單元 254 顯示一預設表情圖像的方式來執行該回饋模型以實現該智慧型機器人 2 與目標物件之間的交互。

【0073】在一實施方式中，該方法還包括步驟：

【0074】設定該智慧型機器人的情感方向；

【0075】從該目標物件的相關事件資訊中提取關鍵資訊；

【0076】根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別；

【0077】根據所提取的關鍵資訊檢索預設的公共基礎知識庫；及

【0078】並利用深度學習演算法根據檢索結果、該目標物件的情緒類別及該設定的智慧型機器人 2 的情感方向確定出對應的回饋模型。在一實施方式中，該設定的情感方向包括勵志型、安慰型、諷刺型、幽默型等。

【0079】在一實施方式中，該方法還包括步驟：



【0080】藉由壓力感測器 24 感測用戶輸入的按壓力；

【0081】藉由該味道感測器 26 檢測智慧型機器人 2 周圍環境的氣味資訊；

及

【0082】根據所提取的關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果、該目標物件的情緒類別、該感測的按壓力及該感測的氣味資訊確定出對應的回饋模型。

【0083】綜上所述，本發明符合發明專利要件，爰依法提出專利申請。惟，以上所述者僅為本發明之較佳實施方式，舉凡熟悉本案技藝之人士，於爰依本發明精神所作之等效修飾或變化，皆應涵蓋於以下之申請專利範圍內。

#### 【符號說明】

##### 【0084】

人機交互系統	1
智慧型機器人	2
伺服器	3
攝像單元	22
語音採集單元	23
壓力感測器	24
輸出單元	25
味道感測器	26
通訊單元	27
處理單元	28
存儲單元	29
語音輸出單元	251
表情輸出單元	252

運動驅動單元	253
顯示單元	254
感知模組	101
識別模組	102
分析模組	103
執行模組	104
設置模組	105
基本資訊表	T1
步驟	S501~S506

**【生物材料寄存】**

**【0085】** 無。



I661363

## 【發明摘要】

【中文發明名稱】 智慧型機器人及人機交互方法

【英文發明名稱】 SMART ROBOT AND HUMAN-COMPUTER INTERACTION

METHOD

## 【中文】

本發明涉及一種智慧型機器人及人機交互方法。該方法應用在該智慧型機器人中。該方法包括步驟：從獲取的語音資訊及圖像資訊中識別一目標物件；確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯；從該目標物件對應的相關事件資訊中提取關鍵資訊；根據所提取的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫，並利用深度學習演算法根據檢索結果及該目標物件的情緒類別確定出一回饋模型；及執行該回饋模型。

## 【英文】

The present invention relates to a smart robot and a human-computer interaction method. The method is applied in the smart robot. The method includes: identifying a target object from a voice information and an image information; determining a basic information of the target object and a relation event information of the target object, and binding the basic information with the relation event information; extracting a key information from the relation event information of the target object; determining an emotional category of the target object according to the extracted key information by a neural network analysis algorithm, retrieving a preset public basic knowledge base according to the key information, and determining a feedback model according to the

retrieval result , the emotional category of the target object by a depth learning algorithm; and performing the feedback.

【指定代表圖】:第(5)圖。

【代表圖之符號簡單說明】

步驟	S501~S506
----	-----------

【特徵化學式】

無。

## 【發明申請專利範圍】

【第 1 項】一種智慧型機器人，包括攝像單元、語音採集單元、輸出單元及處理單元，其改良在於，該處理單元用於：

獲取來自該語音採集單元獲取的語音資訊及該攝像單元獲取的圖像資訊；

從獲取的語音資訊及圖像資訊中識別一目標物件；

確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯，其中，將該目標物件的語音資訊及圖片資訊作為該目標物件的相關事件資訊；

從該目標物件對應的語音資訊及圖片資訊中提取關鍵資訊；

根據從所述語音資訊及所述圖片資訊中提取出的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果、該目標物件的情緒類別及該智慧型機器人預先設定的情感方向確定出一回饋模型，其中，該回饋模型是指用於控制該智慧型機器人與該目標物件進行交互的指令集；及

藉由該輸出單元執行該回饋模型。

【第 2 項】如申請專利範圍第 1 項所述的智慧型機器人，其中，該處理單元從該語音資訊中識別一聲紋特徵及從該圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別對應的目標物件。

【第 3 項】如申請專利範圍第 1 項所述的智慧型機器人，其中，該處理單元識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，並將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。

【第 4 項】如申請專利範圍第 1 項所述的智慧型機器人，其中，該處理單元獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作

資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

【第 5 項】如申請專利範圍第 1 項所述的智慧型機器人，其中，該處理單元還用於設定該智慧型機器人的情感方向為勵志型、安慰型、諷刺型或幽默型。

【第 6 項】如申請專利範圍第 1 項所述的智慧型機器人，其中，該輸出單元包括語音輸出單元及表情輸出單元，該執行模組藉由控制該語音輸出單元輸出語音資訊，及藉由該表情輸出單元輸出表情動作的方式執行該回饋模型。

【第 7 項】一種人機交互方法，應用在一智慧型機器人中，其改良在於，該方法包括步驟：

獲取一語音採集單元獲取的語音資訊；

獲取一攝像單元獲取的圖像資訊；

從獲取的語音資訊及圖像資訊中識別一目標物件；

確定出與該目標物件對應的基本資訊及相關事件資訊，並將該目標物件的基本資訊與該相關事件資訊進行關聯，其中，將該目標物件的語音資訊及圖片資訊作為該目標物件的相關事件資訊；

從該目標物件對應的語音資訊及圖片資訊中提取關鍵資訊；

根據從所述語音資訊及所述圖片資訊中提取出的關鍵資訊利用神經網路分析演算法確定該目標物件的情緒類別，根據所提取的關鍵資訊檢索預設的公共基礎知識庫以獲取與該關鍵資訊相關聯的檢索結果，並利用深度學習演算法根據該檢索結果、該目標物件的情緒類別確定出一回饋模型及該智慧型機器人預先設定的情感方向，其中，該回饋模型是指用於控制該智慧型機器人與該目標物件進行交互的指令集；及

藉由一輸出單元執行該回饋模型。

【第 8 項】如申請專利範圍第 7 項所述的人機交互方法，其中，該方法在步驟“從獲取的語音資訊及圖像資訊中識別一目標物件”還包括：

從該語音資訊中識別一聲紋特徵及從該圖像資訊中識別臉部特徵，並根據該聲紋特徵及該臉部特徵識別對應的目標物件。

【第 9 項】如申請專利範圍第 7 項所述的人機交互方法，其中，該方法還包括步驟：

識別所獲取的語音資訊，將所識別的語音資訊轉化為文本資料，提取出該文本資料中的關鍵資訊，並將該文本資料中的關鍵資訊作為該相關事件的關鍵資訊。

【第 10 項】如申請專利範圍第 7 項所述的人機交互方法，其中，該方法還包括步驟：

獲取該圖片資訊中包含的面部表情資訊及肢體動作特徵資訊，對所獲取的面部表情資訊進行面部表情特徵提取後確定出面部表情特徵參數，對所獲取的肢體動作資訊進行肢體特徵提取後確定出肢體特徵參數，並將該面部表情特徵參數及肢體特徵參數作為該相關事件資訊的關鍵資訊。

【第 11 項】如申請專利範圍第 7 項所述的人機交互方法，其中，該方法還包括步驟：

藉由控制一語音輸出單元輸出語音資訊，及藉由一表情輸出單元輸出表情動作的方式執行該回饋模型。

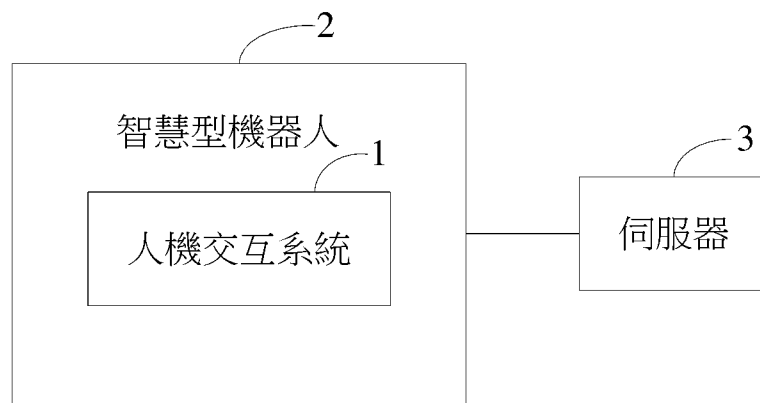


圖1



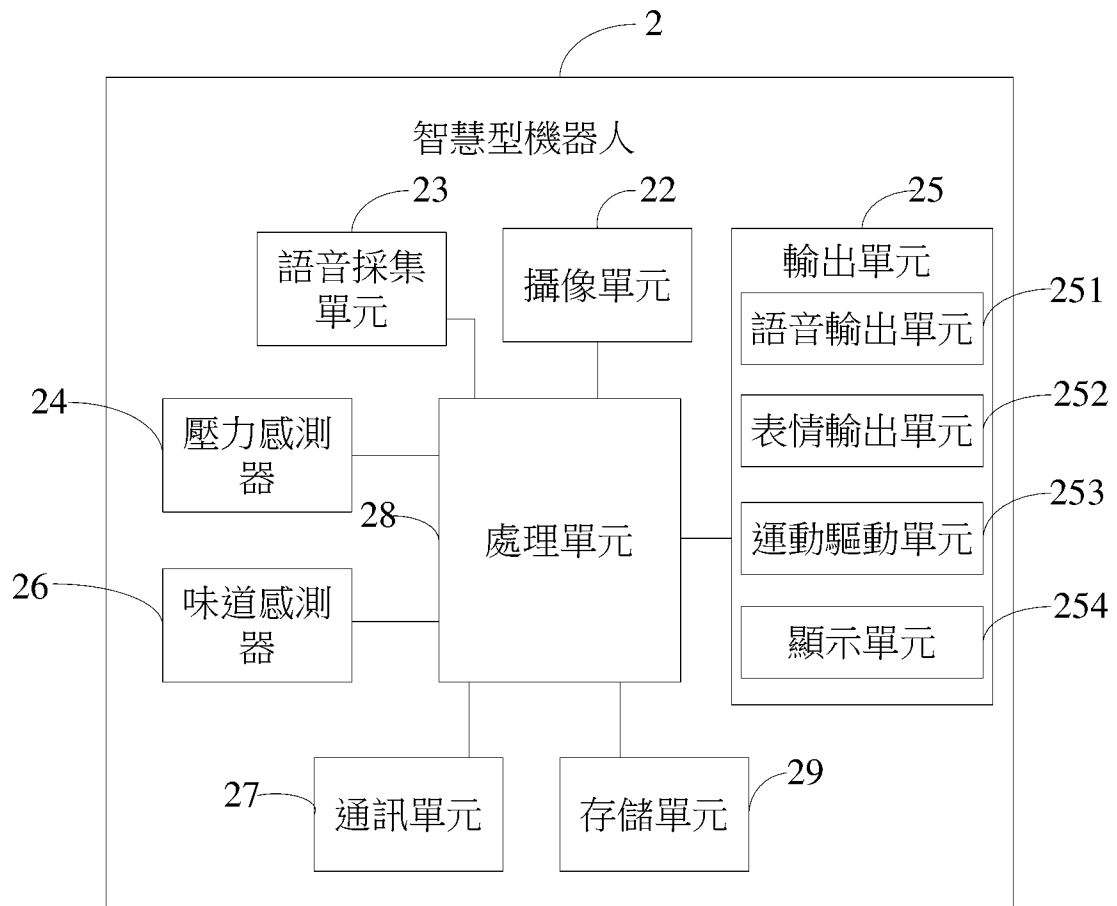


圖2

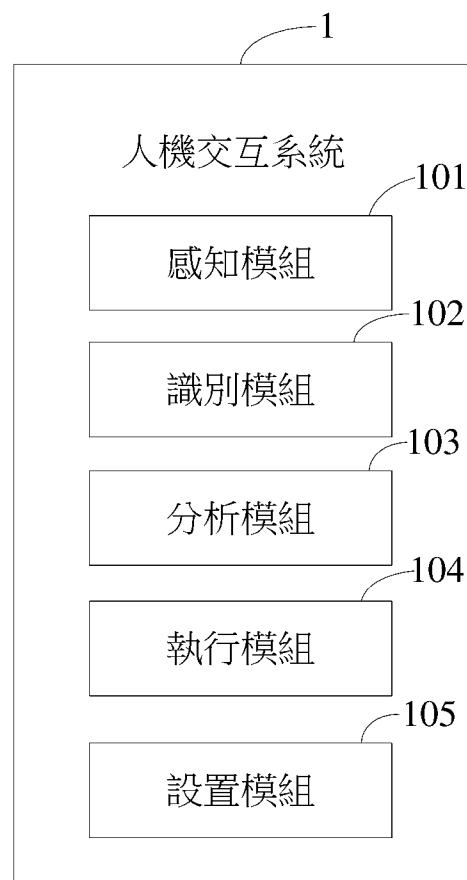


圖3

T1

目標物件	姓名	年齡	身高	體重	體型
使用者A	李某	27	176	65	中型
使用者B	張某	25	180	75	大型
...	...	...	...	...	...

圖4

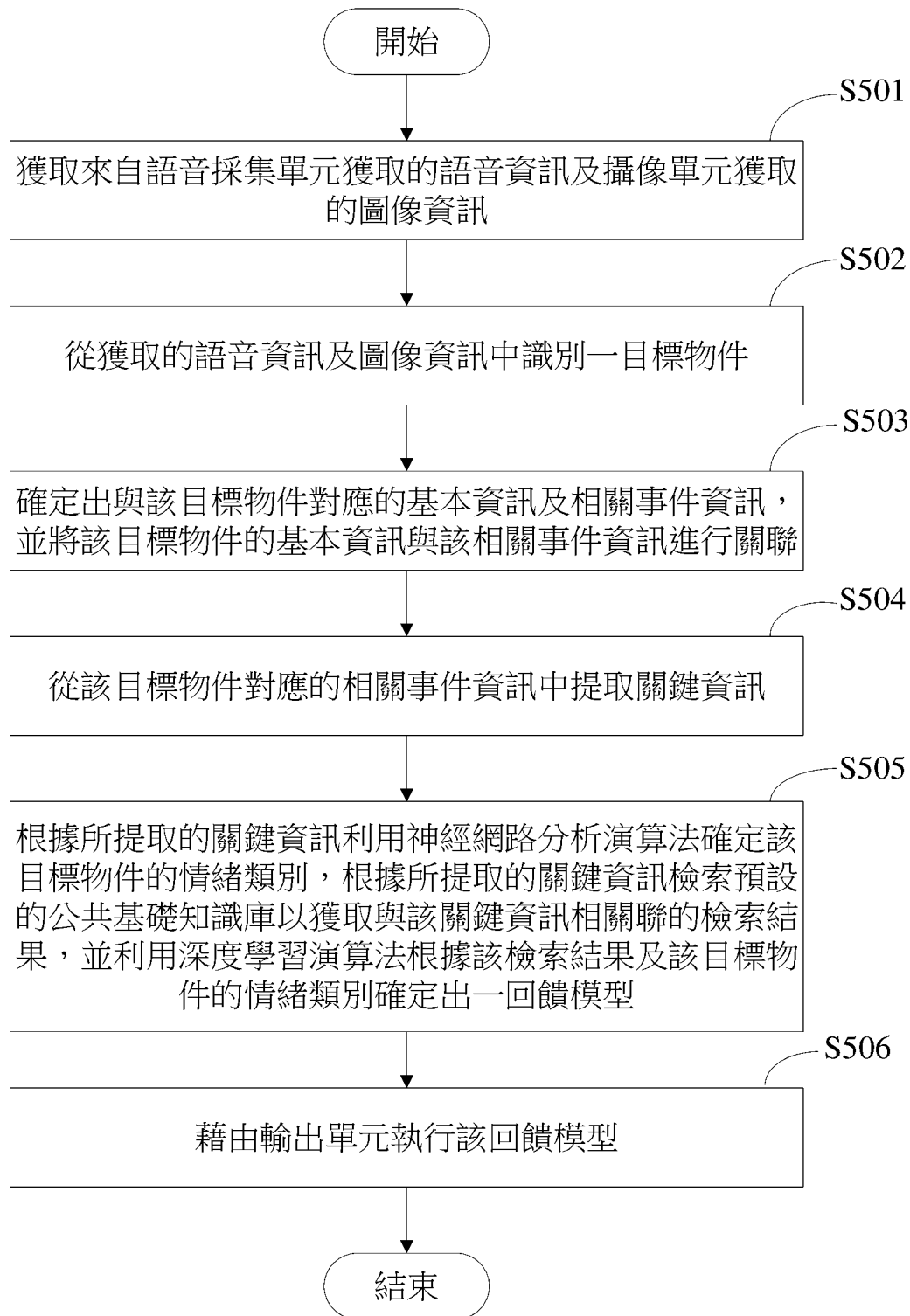


圖5