

92117207

發明專利說明書 200401943

※ 申請案號：92117027

※ 申請日期：92年6月23日

※IPC 分類：G03B 35/08

壹、發明名稱：(中文/英文)

全向性攝影機及麥克風陣列之整合性設計

INTEGRATED DESIGN FOR OMNI-DIRECTIONAL CAMERA AND
MICROPHONE ARRAY

貳、申請人：(共 1 人)

姓名或名稱：(中文/英文)

美商·微軟公司

Microsoft Corporation

代表人：(中文/英文)

丹可萊斯

Dan Crouse

住居所或營業所地址：(中文/英文)

美國華盛頓州列德蒙微軟路 1 號

One Microsoft Way, Building 8, Redmond, WA 98052-6399, USA

國籍：(中文/英文)

美國/USA

參、發明人：(共 4 人)

姓名：(中文/英文)

1. 羅斯卡特勒/Ross Cutler

2. 艾文澁須夫/Ivan Tashev

3. 安歐普古柏塔/Anoop Gupta

4. 楊如意/Yong Rui

住居所地址：(中文/英文)

- 1.美國華盛頓州杜維爾市東北第 277 廣場 16031 號
16031 277th PL. NE, Duvall, WA 98019, U.S.A.
- 2.美國華盛頓州柯克藍市東北第 135 街 13407 號
13407 NE 135th Street, Kirkland, WA 98034-5527, U.S.A.
- 3.美國華盛頓州烏頓維爾市東北第 129 街 19908 號
19908 NE 129th Street, Woodinville, WA 98072, U.S.A.
- 4.美國華盛頓州撒瑪米西市東北第 239 廣場 2508 號
2508 239th Pl. NE, Sammamish, WA 98074, U.S.A.

國籍：(中文/英文)

- 1.美國/USA
- 2.保加利亞/Bulgaria
- 3.美國/USA
- 4.中華人民共和國/PROC

肆、聲明事項：

◎本案申請前已向下列國家(地區)申請專利 主張國際優先權：

【格式請依：受理國家(地區)；申請日；申請案號數 順序註記】

美國；2002 年 6 月 27 日；10/184,499

住居所地址：(中文/英文)

- 1.美國華盛頓州杜維爾市東北第 277 廣場 16031 號
16031 277th PL. NE, Duvall, WA 98019, U.S.A.
- 2.美國華盛頓州柯克藍市東北第 135 街 13407 號
13407 NE 135th Street, Kirkland, WA 98034-5527, U.S.A.
- 3.美國華盛頓州烏頓維爾市東北第 129 街 19908 號
19908 NE 129th Street, Woodinville, WA 98072, U.S.A.
- 4.美國華盛頓州撒瑪米西市東北第 239 廣場 2508 號
2508 239th Pl. NE, Sammamish, WA 98074, U.S.A.

國籍：(中文/英文)

- 1.美國/USA
- 2.保加利亞/Bulgaria
- 3.美國/USA
- 4.中華人民共和國/PROC

肆、聲明事項：

◎本案申請前已向下列國家(地區)申請專利 主張國際優先權：

【格式請依：受理國家(地區)；申請日；申請案號數 順序註記】

美國；2002 年 6 月 27 日；10/184,499

玖、發明說明：

【發明所屬之技術領域】

本發明關於一種整合性全向攝影機與麥克風陣列。更特別是，本發明係指一用於電話會議與會議記錄之整合性全向攝影機與麥克風陣列。

【先前技術】

視訊會議系統以往在商業上僅具有有限度的成功。此歸因於許多因素。特別是，在此等系統中通常有各種技術上之缺失。不良的攝影機視線與不足之影像解析度使參加與會者難以看到發言人。此情形會因不準確之發言人偵測(特別是在遙控/傾斜/送大攝影機)造成攝影機未指向發言人而更形惡化。此外，不良之視頻壓縮技藝常導致不良視頻影像品質且使影像顯示“波動”。

用於電話會議之捕捉裝置與系統傾向於集中於少數對視訊會議與會議觀看較具價值之主要資料來源。這些包括視頻資料、聲頻資料及電子文件或顯示在電腦監視器上之陳述。就現行用以共享文件與陳述之各種特定軟體解決方案，捕捉聲頻與視頻資料之改良方式是特別令人關注的。

現有三種不同捕捉視頻資料之方法：遙控/傾斜/送大(PTZ)攝影機、鏡面架構全向性攝影機與攝影機陣列。雖然PTZ攝影機是目前最流行之選擇，但具有二主要限制。首先，只可捕捉有限之視野。如果其鏡頭拉得太近，就會失去會議室之內容；如果拉得太遠，人們的表情則變得不清

楚。其次，因為控制馬達需要時間移動攝影機，攝影機對會議之反應(如在發言者間切換)嫌太慢。事實上，PTZ攝影機無法移動太大或太快，否則觀眾可能會十分困擾。

因為這些缺點及近年來在鏡頭/稜鏡架構全向視覺感測器技術的提升，研究人員已開始再思考捕捉與分析視訊之方式。例如，貝喜(BeHere)公司提供在娛樂、新聞與運動網路廣播之360度網際網路視頻技術。遠端使用者可用此介面控制與其他觀眾獨立之個人化360度攝影機角度，以獲得一種“臨場”經驗。雖然此方式克服PTZ攝影機所面臨之視野有限與攝影機反應緩慢的二種困難，以今日現行技術與市場需求來架構這些型式之裝置將頗為昂貴。此外，這些鏡面/稜鏡架構全向攝影機深受低解析度(即使具有1MP之感測器)與離焦問題之苦，因而導致次級之視頻品質。

在另一方式中，多個較不昂貴攝影機或視頻感測器將可組合以形成一全向性攝影機陣列。例如，一習知系統利用四個國際電視系統委員會(NTSC)以建構一會議室之全景。然而，此設計具有些缺點。首先，NTSC攝影機提供相當低品質之視頻訊號。此外，該四攝影機需求四視頻捕捉電路板以便在分析、傳輸或記錄前先數位化。對四個視頻捕捉板之需求增加了此系統之成本與複雜度，且使其更難製造與維護。

除了對視訊捕捉問題外，在一會議室中捕捉高品質聲頻也是挑戰。該聲頻捕捉系統需要移除各種雜訊與迴音。

其亦必須調整不同位準之輸入訊號增益。通常有三種方式可滿足這些需求。最簡單之方式是使用密閉式麥克風(如經由頭戴式耳機)，但此對使用者/發言者是累贅而有所侵犯。第二種方式是置放一麥克風於會議桌上。此可防止多重聲音路徑且是目前最常用以記錄會議聲音內容之方式。這些系統使數(通常三)個超心型曲線麥克風以提供全向性特徵。第三種方式係提供一桌上型電話會議系統。在此方式中，一單向性麥克風係裝設在一指向發言者之 PTZ 攝影機之上方。該攝影機/麥克風組合係由一使用分離群組之麥克風以施行聲源定位之電腦所控制。然而此方式會需求二組分離之麥克風。

【發明內容】

本發明關於一種可克服上述在視訊會議與會議記錄系統中之限制的系統與方法。明確言之，本系統與方法利用一整合式全向攝影機與麥克風陣列以達成此工作。

在大多數通識中，本發明由一薄至對於人類發聲頻率範圍(50 至 4000 赫)之聲音係可透過且連接一攝影機陣列至一麥克風陣列的圓桿所組成。結果，可消除聲音之繞射與陰影。

該整合性全向攝影機與麥克風陣列使用設計供解決上述視訊會議中各問題的一 360 度攝影機。該 360 度攝影機可位於會議桌之中央，其比起典型視訊會議系統(其中攝影機是在室內的一端)，可提供與會者絕佳之攝影機取景點。

該攝影機係升高於桌上以提供與會者接近正面的一取景點。此外該整合性全向攝影機與麥克風陣列提供足夠解析度，使一遠端觀眾可看到與會者之面部表情(如，一有效具體實施例具有 3000 x 480 之解析度)。該攝影機可以是任何全向型式，即利用一攝影機陣列或一具有雙曲線鏡頭之單一視頻感測器。

該麥克風陣列係在一平面組態。該等麥克風最好是裝設在一麥克風陣列基座上，以致可儘可能地靠近桌面以消除來自桌子反射之聲音。如上述，該攝影機是以一薄圓桿連接至該麥克風陣列，其對於人類聲聲頻率範圍(即，約 50 至 4000 赫)之聲音係可透過。此提供從說話者至所有在該陣列中麥克風的一直接路徑，使其聲源的定位(決定發言者之位置)與形成聲束(藉由過濾非由發言者方向來之聲音而改進發言者之聲音品質)效果最佳。該整合性麥克風陣列是用以施行即時聲源定位，且該攝影機陣列是協同電腦視覺架構之人體偵測與追蹤，以準確地偵測發言者在影像中之位置。該聲頻與視頻偵測可用於自動化攝影機管理，以及大幅度地改進視頻壓縮(即藉由在面部區域使用比背景更多位元)。

整合性全向攝影機與麥克風陣列之輸出最好是連接至個人電腦，在其中可進行譬如影像聯結與壓縮、聲源定位、聲束形成及攝影機管理等之應用。

整合性攝影機與麥克風陣列的一有效具體實施例使用一 1394 匯流排以傳輸視頻至一 PC，及類比電纜以傳輸聲

頻至一個人電腦(PC)。在此具體實施例中使用五個可提供優異視頻品質之 IEEE 1394 攝影機且只需一單一 1394 卡。另一替代具體實施例使用一單一印刷電路板(PCB)用於所有攝影機與麥克風，使得所有聲頻與視頻可透過單一 1394 電纜傳輸。該 1394 電纜也提供電源，因此在攝影機與 PC 間只需要一單一電纜。

所使用之麥克風可以是全向或單向性(雖然最好是全向)，因為其等會對所有關注之角度有一致的反應。所需最少之麥克風數量為 3，然而本發明一較佳具體實施例使用 8 個以增加聲源定位準確性、較佳之聲束形成與整個聲頻系統之強健性。該等麥克風最好等向地置放於圍繞該圓形、平面麥克風基座周邊的一圓上，雖然其他組態也係可行。使用愈多麥克風會有愈佳之全向聲頻涵蓋與信號雜訊比。然而較多數量的麥克風其成本與複雜性則另需考量。此外，具有愈多麥克風時，處理該等聲頻信號變得愈複雜。為減低桌子雜訊，麥克風可裝設在一橡皮殼體內，而隔音器可置於麥克風下方。

該攝影機可使用一鏡頭遮蓋，向上為正常操作模式，而向下為隱私模式。另一選擇是，當在隱私模式時，攝影機感測器之快門可關閉或是該攝影機可在電子系統上分離以關閉該攝影機。當遇到隱私模式時，該等麥克風最好也關閉。在記錄時，攝影機上之燈會打開讓使用者知道攝影機在操作中。當攝影機在隱私模式時燈會關閉。

各種整合性全向攝影機與麥克風陣列設計之替代具體

實施例均可行。此部份是由於該系統之模組化本質。例如，在一使用全向攝影機之一具體實施例中，利用多個視頻感測器以達到 360 度涵蓋。或者，在本發明另一具體實施例中會使用一全向攝影機，其利用一視頻感測器與一可捕捉來自 360 度光線的雙曲線鏡頭以達到全景涵蓋。再者，該等攝影機之設置可由其等自行使用(升高至該聲音可透過之圓桿上)以提供與會者之正面影像。或者其可與上述麥克風陣列整合。或者，其他之攝影機設計也可用以結合該圓桿。連接該攝影機與麥克風陣列之桿也不需要是圓柱狀，只要其足夠薄使得在(50 至 4000 赫)範圍不會使聲音產生繞射。

同理如前述，在一具體實施例中，該麥克風陣列是由等距離置放在一圍繞一圓形之圓周上且儘可能接近桌面以達成到室內任何發言者的一清晰路徑，使離開桌面反射之聲音最小的各麥克風所組成。然而，其他可使用該聲音可透過之桿整合一全向攝影機設置的麥克風組態也是可行。此外，可使用前述全向麥克風陣列而無須任何攝影機以達成最佳 360 度聲音的涵蓋。此涵蓋對於聲源定位與聲束形成特別有用，因為可最小化或消除多重路徑之問題。

利用本發明攝影機與麥克風陣列之一具體實施例，使用電腦以最佳化該影像資料與聲頻信號。攝影機之數位影像輸出與麥克風陣列之聲頻輸出(經由一類比對數位轉換器)會經由一路徑進入電腦。該電腦施行各種功能以加強且使用該影像與聲頻輸入。例如，一全景影像過濾器會將在

全向攝影機內各種感測器所取得之影像聯結。此外，該影像資料可經壓縮使其更相容供透過一網路(譬如網際網路)廣播，或儲存在一電腦可讀媒體中，最好是經由一可分割待傳輸及/或記錄之視頻與聲頻輸出的分割器。視需要該影像資料也可輸入一人員偵測器/追蹤器以改進攝影機管理。例如，含有該發言者影像/視頻之部份可經識別(且連結該聲頻信號)，使得當其等發言時在視訊會議中顯示之攝影機影像可指向發言者。此外，發言者定位可藉由提供面部比背景較高之解析度以改進視訊壓縮。

該聲頻輸入也可用於各種目的。例如，該聲頻可用於聲源定位，使得任何時間均可最佳化在發言者方向之聲音。此外，在電腦中可使用一聲束形成模組以改進該聲頻之聲束形狀，藉以進一步改進對一特定方向聲音的過濾。一雜訊減低與自動增益控制模組也可藉由減低該雜訊與調整該增益以更佳地捕捉來自一發言者之聲頻信號(而非室內之背景雜訊)，而改進該信號對雜訊比。各個此等影像與聲頻處理模組可單獨，或組合或無須使用。

該視頻或聲頻信號(不論是否經加強)可被廣播至另一視訊會議地點或網際網路。其等也可被儲存至一電腦可讀媒體供後續觀看。

上述整合性全向攝影機與麥克風陣列主要應用為視訊會議與會議記錄。藉由整合該麥克風陣列與全向攝影機，在視頻與聲頻間所需之校正可大幅簡化(一準確製造之攝影機與麥克風將無須校正)，且可達到以一單一裝置從一會

議室收集聲頻與視頻資訊。

【實施方式】

1.0 代表性操作環境

在以下本發明較佳具體實施例的說明中，會參考形成本發明一部份之附圖，且其中所顯示係以可實現本發明的特定具體實施例加以示範。請瞭解可以使用其他具體實施例及可達成結構上之改變而不脫離本發明之範疇。

第 1 圖示範可實施本發明之適當計算系統環境 100 之範例。該計算系統環境 100 僅是一適當計算環境之範例，且不希望使人聯想到任何對本發明使用或功能的限制。同時不希望計算環境 100 被解讀為對於代表性操作環境 100 中所示範組件之一或其組合具有從屬性或必需性。

本發明可用各種其他一般目的或特殊目的之計算系統環境或組態操作。適於使用本發明之習知計算系統、環境及/或組態的範例包括(但不限於)個人電腦、伺服器電腦、手持或膝上型裝置、多重微處理器系統、以微處理器為架構之系統、視訊轉接器、可程式消費者電子裝置、網路 PC、迷你型電腦、主機型電腦、包括上述任何系統或裝置之分散式計算環境與其類似物。

本發明可在由一電腦所執行之電腦可執行指令(譬如程式模組)中之一般上下文中加以描述。大體上，程式模組包括執行特定工作或實施特定抽取資料型式之常式、程

式、物件、組件、資料結構等等。本發明也可在工作係藉由經一通信網路連結之遠端處理裝置施行的分散式計算環境中實施。在一分散式計算環境中，程式模組可位於包括記憶體儲存裝置之區域與遠端電腦儲存媒體中。

請參考第 1 圖，用以實施本發明的一代表性系統包括如電腦 110 型式之一般電腦。電腦 110 之組件可包括(但不限於)一處理單元 120、一系統記憶體 130 及一將包括系統記憶體之各種系統組件耦合至處理單元 120 之系統匯流排 121。系統匯流排 121 可為任何數種型式匯流排結構，包括一記憶體匯流排或記憶體控制器、一周邊匯流排與一使用任何各種匯流排結構之區域匯流排。藉由實例說明而非限制，此架構包括工業標準架構 (ISA) 匯流排、微通道架構 (MCA) 匯流排、增強工業標準架構 (EISA) 匯流排、視頻電子標準協會 (VESA) 區域匯流排、與周邊組件內部連線 (PCI) 匯流排 (亦稱為 Mezzanine 匯流排)。

電腦 110 通常包括各種電腦可讀媒體。電腦可讀媒體可以是由電腦 110 存取之任何可用媒體，且包括揮發性與非揮發性、可移與不可移媒體二者。以實例說明(而非限制)，電腦可讀媒體可至少包含電腦儲存媒體與通信媒體。電腦儲存媒體包括揮發性與非揮發性、可移與不可移媒體，能以任何用於儲存譬如電腦可讀指令、資料結構、程式模組或其他資料之資訊的方法或技術實施。電腦儲存媒體包括(但不限於)RAM、ROM、EEPROM、快閃記憶體或其他記憶體技術，CD-ROM、數位多功能碟片 (DVD) 或其他

光碟儲存、磁卡匣、磁帶、磁碟儲存或其他磁性儲存裝置，或任何其他可用以儲存需求資訊且可由電腦 110 存取之媒體。通信媒體通常包括電腦可讀指令、資料結構、程式模組或在調變資料信號(譬如載波或其他傳輸機構)內之其他資料，且包括任何資訊傳遞媒體。“調變資料信號”一詞意指一信號具有能經設定或改變之一或多數特徵，以便將資訊編碼在該信號中。舉例而言(而非限制)，通信媒體包括有線媒體(譬如線網路或直接佈線連接)與無線媒體(譬如音響、射頻、紅外線及其他無線媒體)。以上之任何組合也包括在電腦可讀媒體之範疇。

系統記憶體 130 包括揮發性及/或非揮發性記憶體型式之電腦儲存媒體，譬如唯讀記憶體(ROM)131 與隨機存取記憶體(RAM)132。含有可協助電腦 110 內元件之間傳送資訊(譬如在開機期間)之基本常式之一基本輸入/輸出系統(BIOS)通常係儲存於 ROM 131 內。RAM 132 通常含有可由處理單元 120 立即存取及/或目前正在其上操作之資料及/或程式模組。舉例而言(而非限制)，第 1 圖示範操作系統 134、應用程式 135、其他程式模組 136 與程式資料 137。

電腦 110 也可包括其他可移/不可移、揮發性/非揮發性媒體。舉例而言(而非限制)，第 1 圖示範一可讀自或寫至一不可移、非揮發性磁性媒體的一硬碟驅動器 141、一用於讀自與寫至一可移、非揮發性磁碟 152 之磁碟驅動器 151、與一用於讀自與寫至一可移、非揮發式光碟 156(譬如一 CD ROM 或其他光學媒體)之光碟驅動器 155。其他可

用於該代表性操作環境中之可移/不可移、揮發性/非揮發性電腦儲存媒體包括(但不限於)磁帶卡匣、快閃記憶卡、數位多功能碟片、數位錄影帶、固態隨機存取記憶體、固態唯讀記憶體與其他類似物。該硬碟驅動器 141 通常是經由一如介面 140 之不可移記憶體介面連接至系統匯流排 121，而磁碟驅動器 151 與光碟驅動器 155 通常是藉由一如介面 150 之可移式記憶體介面連接至系統匯流排 121。

以上討論與示範於第 1 圖中之驅動器及其相關電腦儲存媒體提供電腦 110 之電腦可讀指令、資料結構、程式模組或其他資料的儲存。(例如)在第 1 圖中所示範之硬碟驅動器 141 是作為儲存操作系統 144、應用程式 145、其他程式模組 146 與程式資料 147。請注意這些組件可與操作系統 134、應用程式 135、其他程式模組 136 與程式資料 137 相同或不同。在此特定同數字之操作系統 144、應用程式 145、其他程式模組 146 與程式資料 147，是用以示範(至少)其等是不同版本。一使用者可將命令與資訊經由輸入裝置(譬如鍵盤 162 與一般稱為滑鼠、軌跡球或觸摸板之指向裝置 161)輸入電腦 110。其他輸入裝置(未顯示)可包括一麥克風、搖桿、遊戲墊、萬用碟、掃描器或其類似物。這些與其他輸入裝置通常經由一耦合至系統匯流排 121 之使用者輸入介面 160 連接至處理單元 120，但也可藉由其他介面與匯流排，譬如一平行連接埠、遊戲連接埠、統一串列匯流排(USB)連接。一螢幕 191 或其他型式之顯示裝置也可經由一介面(譬如一視頻介面 190)連接至系統匯流排

121，而後與視頻記憶體 186 通信。除螢幕 191 外，電腦通常也包括其他周邊輸出裝置，譬如喇叭 197 與列表機 196 可經由一輸出周邊介面 195 連接。本發明中特別重要的是能夠捕捉一序列影像 164 之攝影機 163(譬如一數位/電子靜態或視頻攝影機，或薄膜/圖片掃描器)，可當作個人電腦 110 之一輸入裝置。再者，雖然只描述一攝影機，但可包括多個攝影機作為個人電腦 110 之輸入裝置。來自一或多個攝影機之影像 164 會經由一適當之攝影機介面 165 輸入個人電腦 110。此介面 165 係連接至系統匯流排 121，藉以容許經一路徑將影像傳至且儲存於隨機存取記憶體 132 內，或連接至電腦 110 之一或其他資料儲存裝置。然而，應注意影像資料可從任何上述電腦可讀媒體輸入電腦 110，而無須使用攝影機 163。

電腦 110 可在一使用至一或多數遠端電腦(譬如一遠端電腦 180)之邏輯連接的網路化或分散式環境中運作。遠端電腦 180 可以是一個人電腦、一伺服器、一路由器、一網路個人電腦、一同級裝置或其他共用網路節點，且通常包括許多或所有上述與電腦 110 有關之元件(雖然第 1 圖中只示範一記憶體儲存裝置 181)。第 1 圖中描述之邏輯連接包括一區域網路(LAN)171 與一廣域網路(WAN)173，但也可包括其他網路/匯流排。此網路化環境係位於家中、辦公室、泛企業電腦網路、內部網路與網際網路中之共用處所。

當使用在區域網路化環境中時，電腦 110 係經由一網路介面或適配器 170 連接至區域網路 171。當在廣域網路

環境中使用時，電腦 110 通常包括一數據機 172 或其他用於透過廣域網路 173(譬如網際網路)建立通信的構件。數據機 172(可為內建或外接)係經由使用者輸入介面 160 或其他適當機構連接至系統匯流排 121。在一網路化環境中，所顯示連接至個人電腦 110 之程式模組或其部份，可儲存於遠端記憶體儲存裝置中。藉由舉例(而非限制)，第 1 圖顯示遠端應用程式 185 係駐留在記憶體裝置 181 上。應瞭解所顯示之網路連接是代表性，且其他在電腦間建立一通信鏈路之其他構件均可使用。

現已討論過代表性操作環境，本說明書剩餘部份將致力於說明包含本發明之程式模組。

3.0 整合性全向攝影機與麥克風陣列

在此段落中將討論經由一聲音可透過圓桿連接之整合性全向攝影機與麥克風陣列。

3.1 概述

本發明陳述如何最佳地整合一全向性攝影機與一麥克風陣列。設計之目標為：

1. 該麥克風陣列之設計應提供從一發言人至所有在陣列內之麥克風的一清晰路徑。如果一清晰路徑不可行，則所使用之聲源定位與聲束形成演算法會變得特別困難且獲得品質降低之結果。
2. 該麥克風陣列之設計應使該麥克風儘可能接近桌

面或其他表面，以防止聲音從該表面反射，而使任何聲音定位與聲束形成結果品質降低。

3. 該攝影機陣列應升高以提供與會者的一接近正面影像。該影機陣列應小至不明顯。
4. 該麥克風陣列之幾何性質(二麥克風之位置與距離)應容許該聲頻處理演算法達到在工作頻寬中之良好聲束形狀(直接到該發言人)，藉以提供高品質聲音。

3.2 整合設計

該整合性攝影機與麥克風陣列利用一圓桿連接麥克風基座與攝影機陣列。該桿在聲音上對於人類演講之頻率範圍(50至4000赫)而言應視為可透過的。

如第2圖所示，該整合性攝影機與麥克風陣列202是希望置放在會議室桌204之中央。

該設計提供從任何特定發言人或聲源至所有麥克風的一清晰路徑，且置放該麥克風靠近桌面以避免聲音從該桌反射之多重路徑問題。此外，該設計會將攝影機從桌面升高，因此提供所有與會者的一正面或接近正面之影像。

該整合性全向攝影機與麥克風陣列確保一良好聲束形狀，可藉由過濾而只留一方向之聲音而改進發言者之聲音品質。再者，該攝影機與麥克風之整合性本質係優點，因為其消除了重覆校正之需要。由於攝影機與麥克風被整合成一單一裝置，只需要一次初始化校正。同時，由於該攝影機與麥克風可以是一精簡、固定設計，其比二分離之攝

影機與麥克風(在會議桌上需要分開之電纜與額外空間)會減少干擾。

3.3 系統組件

整合性全向攝影機與麥克風陣列的一具體實施例示於第 3A 與 3B 圖。此設計中之組件包括一全向攝影機 302、一圓桿 304、一麥克風基座 306、麥克風 308、一麥克風前放大器 310 與一類比對數位(A/D)轉換器(未顯示)。

3.3.1 全向攝影機

目前有各種全向攝影機技術。這些包括一攝影機型式，其中多個視頻感測器是以背對背方式緊密地封裝在一起。另一全向攝影機型式包含一具有一可捕捉 360 度光線之雙曲線鏡頭的單一視頻感測器。本發明之整合性攝影機與麥克風陣列設計可使用任何此類全向攝影機。最好該攝影機頭部 302 足夠小，以致當設置於一會議室桌或其他表面時不會干擾。

如果是使用一多個感測器攝影機之組態，則可包含複數個攝影機或視頻感測器。較佳之數目為 8 個。這些感測器最好是以背對背方式置放，使得各感測器之投影中心是等角距離分散。例如使用 8 個感測器，則各感測器將是與相鄰感測器隨 45 度。然而，如果需要以不同距離捕捉影像時，也可能包含不同鏡頭與不同攝影機佈置。例如，此情況會如同在一矩形或橢圓形會議桌。具有較長、較窄視野

之鏡頭可用於較長距離，而較寬、較短視野之鏡頭可用以捕捉在短距離之影像。在此情形下該攝影機感測器可能不是等向地繞該攝影機頭部置放。具有較寬視野之攝影機感測器可置放於比具有較窄視野之攝影機感測器要更遠。或者，可使用具有一可變視野(可旋轉縮小或放大以調整一特定位置)之攝影機。

本發明一有效具體實施例(如第 4 圖所示)使用一精簡型多個感測器設計，其中多個微型感測器 402 是以背對背方式繞一圓形攝影機 404 頭部之圓周設置。攝影機 404 之底部最好具有吸音材料以避免發生聲音從桌子反射至攝影機至麥克風。在此具體實施例中使用 8 個攝影機感測器。此精簡型設計(其中該攝影機感測器是以背對背方式緊密地封裝在一起)改進了影像聯結。在此具體實施例中，攝影機頭 404 是約 50 毫米寬，而連接該攝影機頭至基座之圓桿 406 是約 250 公分高。以此高度置放攝影機 404 可使其獲得所有與會者之正面影像。或者，攝影機頭 404 可低於此，而攝影機感測器 402 可稍微朝上。此也使得攝影機較少受干擾。

3.3.2 圓柱

請參考第 3A 與 3B 圖中之具體實施例，該全向攝影機 302 係附接至一較佳是中空之圓桿 304，其係附接至一麥克風基座 306。該圓柱直徑(D1)應足夠薄，以使圓柱 304 在聲音上對希望使用之相關頻率是可透過的。最好該攝影機

電纜是通過圓柱 304，以致不會造成任何額外的聲音障礙。圓柱高度(H1)應足夠高以提供每一個繞著會議桌坐之人士的一正面影像，但不應太高以致造成對攝影機之干擾。該圓柱高度視需要應可調整。例如在一視訊會議應用中，該圓柱高度可調整以考慮各種桌子及參與者高度。

請再參考第 3A 與 3B 圖中所示之有效具體實施例，圓柱 304 具有 2 公分或更少之直徑(D1)，其可容許約 50 至 4000 赫之聲頻(相對於人聲)通過該圓桿，基本上不干涉從任何發言之與會者至基座內之所有麥克風。在此具體實施例中，該圓桿高度(H1)係 14 公分。

3.3.3 麥克風基座

通常，麥克風基座是支撐該麥克風、麥克風前放大器及 A/D 轉換器。其連接該圓桿且提供一攝影機電纜之連接插座。該麥克風基座具有較低輪廓，以使桌面與麥克風間之距離最小。此基座提供從各麥克風至與會者間一直接路徑。

在第 3A 與 3B 圖中所示之整合性攝影機與麥克風陣列之有效具體實施例中，麥克風基座 306 之直徑 D3 是比攝影機頭 302 之直徑 D2 寬。此提供該整合性全向攝影機與麥克風陣列穩定性，避免其易於被碰倒。基座之高度(H3)是相當小。此尺寸最好是足夠低以保持該嵌入麥克風 308 足夠靠近桌面以避免多重路徑問題，但需夠高以容許麥克風 308 嵌入基座內。對於此有效具體實施例，D3 是 16 公

分、H3 是 1.5 公分、H2 是 6 公分而 D2 是 10 公分。

3.3.4 麥克風

所使用之麥克風可為全向或單向，但全向麥克風是較佳，因為其等對所有關注之聲音角度均可獲得一致的響應。需求麥克風之最少數量為 3，雖然本發明之具體實施例使用 8 個以增加聲源定位準確性、較佳聲束形成與整體聲頻系統之強健性。

為減低桌面雜訊，麥克風可裝設在一橡皮殼體內，且隔音器可置放於該麥克風下用於相同之目的。

再請參考第 3A 與 3B 圖所示之有效具體實施例，麥克風 308 是繞該平面麥克風基座 306 上一圓之圓周等向置放。在此具體實施例中包含 8 個麥克風 308。一般而言，使用愈多麥克風，全向聲頻涵蓋性與信號雜訊比愈佳。然而，須考量較多麥克風之成本與複雜性。此外，愈多麥克風則處理該聲頻信號變得愈複雜。在第 3A 與 3B 圖所示之有效具體實施例中，從圓柱 306 中心至各麥克風 308 中心之距離(D5)是 7 公分。

3.3.5 麥克風前放大器，A/D 轉換器

該麥克風前放大器 310 與類比對數位(A/D)轉換器(未顯示)最好是整合於麥克風基座 306 內，如第 3B 圖所示。在此具體實施例中，前放大器 310 之寬度(D4)是 5.901 公分。該麥克風前放大器放大來自於麥克風之信號，以標準

化該信號幅度供後續 A/D 轉換器使用。該 A/D 轉換器轉換來自於攝影機之類比信號成為數位。

在此具體實施例中，來自麥克風之信號的信號取樣是彼此同步於 1 毫秒內，以有助於聲源定位與聲束形成。

3.4 隱私模式

該攝影機可使用一鏡頭遮蓋，其正常操作模式時開啟，而在隱私模式時關閉。另一選擇是，當在隱私模式時，攝影機感測器之快門可關閉或是該攝影機可與電子系統分離以關閉該攝影機。當遇到隱私模式時，該等麥克風最好也關閉。在記錄時，攝影機上的一燈會打開讓使用者知道攝影機在操作中。當攝影機在隱私模式時燈會關閉。

4.0 模組化本質之替代具體實施例

各種整合性全向攝影機與麥克風陣列設計之替代具體實施例均係可行。此部份是由於該系統之模組化本質。

例如，可包含各種攝影機之具體實施例。在一具體實施例中，使用之全向攝影機包含多個視頻感測器以達到攝影機之 360 度涵蓋。或者在本發明另一具體實施例中，會使用一全向攝影機，其利用一視頻感測器與一可捕捉來自 360 度光線的雙曲線鏡頭以達到涵蓋全景。再者，該等攝影機設置可由其等自行使用(升高至該聲音可透過之圓桿上)以提供與會者之正面影像。或者該等攝影機可與一麥克風陣列整合。另一選擇是其他全向攝影機設置也可用以結

合該圓桿及/或麥克風陣列。

同理，可包含各種麥克風組態。在一具體實施例中，該麥克風陣列是由等距離置放在一圍繞一圓形之圓周上且儘可能接近桌面，以達成到室內任何發言者的一清晰路徑的各麥克風所組成。然而，其他使用該聲音可透過圓桿整合一全向攝影機設置之麥克風組態也屬可行。或者，可使用前文討論之全向麥克風陣列而無須任何攝影機以達成最佳 360 度的聲音涵蓋。此涵蓋對於聲源定位與聲束形成特別有用，因為可最小化或消除多重路徑之問題。

在整合性攝影機與麥克風陣列之一具體實施例中，影像聯結與壓縮是在一 PC 上施行。一替代具體實施例是在具有一場可程式化閘陣列 (FPGA) 或其他閘陣列之攝影機中施行影像聯結與壓縮。該設計使用一 USB 介面以建立攝影機與 PC 之介面，且提供該 PC 更多 CPU 循環以進行其他工作，譬如影像壓縮及記錄/廣播該會議。

5.0 代表性有效具體實施例

包含本發明之攝影機 502 與麥克風陣列 504 之一有效具體實施例如第 5 圖所示。攝影機 502 之數位影像輸出與麥克風陣列 504 之聲音輸出會經由一類比對數位轉換器 506 之路徑進入電腦 508。電腦 508 施行各種功能以加強且使用該影像與聲音輸入。例如，一全景過濾器模組 510 會將在全向攝影機 502 內各種感測器所取得之影像聯結。此外，該影像資料可經一壓縮模組 512 壓縮以使其更相容供

透過一網路(譬如網際網路)廣播 514, 或儲存在一電腦可讀媒體 516 中(最好是經由一分割器 520)。視需要該影像資料也可輸入一個人偵測器/追蹤器模組 522 以改進攝影機管理 524。例如, 含有該發言者影像/視頻之部份可經識別(且連結該聲頻信號), 使得由直接指向該發言者之攝影機/感測器捕捉之影像會被廣播或儲存於碟片。

該聲頻輸入也可用於各種目的。例如, 該聲頻可被輸入至一聲源定位模組 526, 使得來自發言者之聲音可被隔離。此外, 一聲束形成模組 528 可用於電腦 508 中以改進該聲頻之聲束形狀。一雜訊減低與自動增益控制模組 530 也可藉由減低該雜訊與調整該增益, 以更佳地捕捉來自一發言者之聲頻信號(而非室內之背景雜訊), 而改進該信號對雜訊比。

如上述, 該視頻或聲頻信號(不論是否經加強)可被廣播至另一視訊會議地點或網際網路。其等也可被儲存至一電腦可讀媒體供後續觀看。

上文所呈現之本發明說明僅用於示範與說明之目的。本發明不希望被視為毫無遺漏或使本發明受限於所揭示之精確形式。例如, 上述整合性攝影機與麥克風陣列之具體實施例可應用於一監視系統。根據上述教示將可以有許多修改與變化。因此希望本發明範疇之限制不在於此詳細說明, 而是依據隨附之申請專利範圍。

【圖式簡單說明】

本發明之特定特色、觀點與優勢可參考以下說明、隨
附申請專利範圍與圖式而更佳瞭解，其中：

第 1 圖係描述構成實施本發明之代表性系統的一般計
算裝置之圖式；

第 2 圖係描述該整合性攝影機與麥克風陣列在一會議
桌上之較佳位置的圖式；

第 3A 圖係該整合性攝影機與麥克風陣列之一具體實
施例的側視圖；

第 3B 圖係第 3A 圖中所示該整合性攝影機與麥克風陣
列之具體實施例的上視圖；

第 4 圖係該整合性攝影機與麥克風陣列之另一具體實
施例的立體圖；

第 5 圖係本發明使用電腦以加強所捕捉視頻與聲頻資
料的一有效具體實施例。

【元件代表符號簡單說明】

100 電腦系統	110 電腦
120 處理單元	121 系統匯流排
130 系統記憶體	131 唯讀記憶體
132 隨機存取記憶體	134 操作系統
135 應用程式	136 其他程式模組
137 程式資料	140 介面
141 硬碟驅動器	144 操作系統
145 應用程式	146 其他程式模組

- | | | | |
|-----|---------|-----|---------------|
| 147 | 程式資料 | 150 | 介面 |
| 151 | 磁碟驅動器 | 152 | 磁碟 |
| 155 | 光碟驅動器 | 156 | 光碟 |
| 160 | 使用者輸入介面 | 161 | 滑鼠 |
| 162 | 鍵盤 | 163 | 攝影機 |
| 164 | 影像 | 165 | 攝影機介面 |
| 170 | 網路介面 | 171 | 區域網路 |
| 172 | 數據機 | 173 | 廣域網路 |
| 180 | 遠端電腦 | 185 | 遠端應用程式 |
| 190 | 視頻介面 | 191 | 螢幕 |
| 192 | 攝影機 | 194 | 攝影機介面 |
| 195 | 輸出周邊介面 | 196 | 列表機 |
| 197 | 喇叭 | 202 | 整合性攝影機與麥克風 |
| 204 | 會議桌 | 302 | 全向攝影機 |
| 304 | 圓桿 | 306 | 麥克風基座 |
| 308 | 麥克風 | 310 | 麥克風前放大器 |
| 402 | 感測器 | 404 | 攝影機 |
| 502 | 全向攝影機 | 504 | 麥克風陣列 |
| 506 | A/D 轉換器 | 508 | 電腦 |
| 510 | 全景過濾器模組 | 512 | 壓縮模組 |
| 514 | 網路 | 516 | 電腦可讀媒體 |
| 520 | 分割器 | 522 | 人員偵測器/追蹤器模組 |
| 524 | 管理 | 526 | 聲源定位模組 |
| 528 | 聲束形成模組 | 530 | 雜訊減低與自動增益控制模組 |

伍、中文發明摘要：

本發明提供一種具有整合性麥克風陣列之全向攝影機(一 360 度攝影機)。此一攝影機之主要應用是視訊會議與會議記錄，且該裝置經設計置於一會議室桌上。該麥克風陣列係在一平面組態，且該等麥克風是儘可能地靠近桌面以消除來自桌子之反射聲音。該攝影機是以一薄圓桿連接至該麥克風陣列基座，其對於頻率範圍(50 至 4000 赫)之聲音係可透過。此提供一條從說話者至所有在該陣列中麥克風之直接路徑，因而可用於聲源的定位(決定說話者之位置)與聲束形成(藉由過濾只留由一特殊方向來之聲音而改進說話者之聲音品質)。該攝影機陣列係從桌子升高以提供與會者一接近正面之影像。

陸、英文發明摘要：

An omni-directional camera (a 360 degree camera) is proposed with an integrated microphone array. The primary application for such a camera is videoconferencing and meeting recording, and the device is designed to be placed on a meeting room table. The microphone array is in a planar configuration, and the microphones are located as close to the desktop as possible to eliminate sound reflections from the table. The camera is connected to the microphone array base with a thin cylindrical rod, which is acoustically invisible to the microphone array for the frequency range [50-4000] Hz. This provides a direct path from the person talking to all of the microphones in the array, and can therefore be used for sound source localization (determining the location of the talker) and beam-forming (improving the sound quality of the talker by filtering only sound from a particular direction). The camera array is elevated from the table to provide a near frontal viewpoint of the meeting participants.

拾、申請專利範圍：

1. 一種用於捕捉聲頻與視頻資料之系統，至少包含：
 - 一或多數捕捉視頻影像資料之攝影機；
 - 一捕捉聲頻信號之麥克風陣列；及
 - 一將該一或多數攝影機連接至該麥克風陣列之圓柱，該圓柱對人聲之頻率範圍內係聲音可透過。
2. 如申請專利範圍第 1 項所述之系統，其中該一或多數攝影機捕捉超過 360 度之影像。
3. 如申請專利範圍第 1 項所述之系統，更包含一麥克風陣列基座，且其中該等麥克風是裝設於該麥克風基座上。
4. 如申請專利範圍第 3 項所述之系統，其中該麥克風陣列基座之高度是足夠小以定位該麥克風，以使該陣列從任何表面拾取之聲音反射最小。
5. 如申請專利範圍第 1 項所述之系統，其中該圓柱對聲音可透過之頻率範圍是 50 至 4000 赫。
6. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是輸入一電腦，該電腦使用該等聲頻信號用於聲源定位，該聲源定位決定產生該等聲源信號之聲音的方向。

7. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是輸入一電腦，該電腦使用該等聲頻信號用於聲束形成，該聲束形成藉由過濾來自於所有其他方向之聲音改進來自一方向之聲音的品質。
8. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是輸入一電腦，該電腦使用人類偵測與追蹤演算法以準確地偵測已捕捉在該視頻影像資料中之一人，且將該人之該視頻影像資料與由該人發言所產生之該等聲頻信號產生關聯。
9. 如申請專利範圍第 1 項所述之系統，其中該視頻影像信號是輸入一電腦，該電腦使用該視頻影像資料將影像連結在一起。
10. 如申請專利範圍第 1 項所述之系統，其中在該麥克風陣列中之麥克風是全向性。
11. 如申請專利範圍第 1 項所述之系統，其中在該麥克風陣列中之麥克風是單向性。
12. 如申請專利範圍第 1 項所述之系統，其中該麥克風陣列

至少包含三麥克風。

13. 如申請專利範圍第 1 項所述之系統，其中該麥克風陣列至少包含八麥克風。

14. 如申請專利範圍第 1 項所述之系統，其中該麥克風陣列至少包含等向置放於繞一平面麥克風基座之該圓周的一圓上之複數個麥克風。

15. 如申請專利範圍第 14 項所述之系統，其中至少一麥克風是裝設於一橡皮殼體中以保護其免於無關的聲音反射。

16. 如申請專利範圍第 14 項所述之系統，其中該隔音器是置於至少一麥克風下以保護其免於無關的聲音反射。

17. 如申請專利範圍第 1 項所述之系統，其中該攝影機更包含一鏡頭蓋，當在一正常操作模式是朝上而在隱私模式則朝下。

18. 如申請專利範圍第 17 項所述之系統，其中當該攝影機是在該隱私模式作用時會關閉該等麥克風。

19. 如申請專利範圍第 1 項所述之系統，其中當該攝影機作用時在攝影機上的一燈會亮起。
20. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是透過一網路傳遞。
21. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是儲存至一電腦可讀媒體。
22. 如申請專利範圍第 1 項所述之系統，其中該等視頻信號是透過一網路傳遞。
23. 如申請專利範圍第 1 項所述之系統，其中該等視頻信號是儲存至一電腦可讀媒體。
24. 如申請專利範圍第 1 項所述之系統，其中該等視頻信號是使用一 1394 匯流排傳輸至一電腦。
25. 如申請專利範圍第 1 項所述之系統，其中該等聲頻信號是使用一類比電纜傳輸至一電腦。
26. 如申請專利範圍第 1 項所述之系統，其中該等攝影機是 IEEE 1394 攝影機。

27. 一種用於電話會議與會議記錄之捕捉視頻與聲頻的方法，至少包含下列步驟：

以一全向攝影機捕捉在會議中的一人之影像；

以一麥克風陣列捕捉發生在該會議之聲音的聲頻信號，該麥克風陣列被容置於一以聲音可透過之桿連接至該全向攝影機的基座中。

28. 如申請專利範圍第 27 項所述之方法，其中該麥克風陣列是置放於一桌上且其中在該會議中之該等人是圍繞該桌坐著。

29. 如申請專利範圍第 28 項所述之方法，其中該麥克風陣列是儘可能靠該桌，使得來自於該桌被該陣列拾取之反射聲音最小。

30. 如申請專利範圍第 27 項所述之方法，其中該麥克風陣列至少包含三或更多之麥克風。

31. 如申請專利範圍第 30 項所述之方法，其中該麥克風陣列至少包含平均分佈於與該麥克風基座之該圓周相鄰的一圓內之麥克風。

32. 如申請專利範圍第 27 項所述之方法，其中該全向攝影機至少包含多數視頻感測器。
33. 如申請專利範圍第 32 項所述之方法，其中該等視頻感測器是平均分佈於圍繞一圓之該圓周。
34. 如申請專利範圍第 27 項所述之方法，其中該全向攝影機至少包含一具有從 360 度捕捉光線的一雙曲線鏡頭之單一視頻感測器。
35. 如申請專利範圍第 27 項所述之方法，其中該聲音可透過之桿是中空的。
36. 如申請專利範圍第 27 項所述之方法，其中該聲音可透過之桿是附接於一麥克風基座，且其中攝影機電纜通過該聲音可透過之桿，因此該等電纜不會造成聲音障礙。
37. 如申請專利範圍第 27 項所述之方法，其中該聲音可透過之桿的外徑是 2 公分或更少。
38. 如申請專利範圍第 27 項所述之方法，其中該聲音可透過之桿具有一直徑以容許從 50 至 4000 赫之聲頻(相對於人聲)通過該圓桿，實質上不干擾從一人朝向該麥克

風陣列內所有麥克風之發言。

39. 如申請專利範圍第 27 項所述之方法，其中該桿之高度是可調的。

40. 一種整合性全向攝影機與麥克風陣列，至少包含：

- 一全向攝影機；
- 一升高該攝影機以提供最佳攝影機涵蓋之聲音可透過桿；及
- 一麥克風陣列。

41. 如申請專利範圍第 40 項所述之整合性全向攝影機與麥克風陣列，其中該全向攝影機包含多數視頻感測器以達到 360 攝影機涵蓋。

42. 如申請專利範圍第 40 項所述之整合性全向攝影機與麥克風陣列，其中該全向攝影機包含一視頻感測器與一從 360 度捕捉光線的一雙曲線鏡頭以達到涵蓋全景。

43. 如申請專利範圍第 40 項所述之整合性全向攝影機與麥克風陣列，其中該聲音可透過之桿升高該攝影機以提供圍繞一桌坐之人員的一正面影像。

44. 如申請專利範圍第 40 項所述之整合性全向攝影機與麥克風陣列，其中該麥克風陣列至少包含複數個由等距離圍繞一圓形之該圓周置放的麥克風。

45. 如申請專利範圍第 40 項所述之整合性全向攝影機與麥克風陣列，其中該麥克風陣列是座落於一表面上，儘可能地靠近該表面以使從該表面之聲音反射最小。

46. 如申請專利範圍第 41 項所述之整合性全向攝影機與麥克風陣列，其中具有較長、較窄視野之視頻感測器是用以捕捉在較長距離處之影像，而具有較寬、較短視野之視頻感測器可用以捕捉在短距離處之影像。

47. 一種全向麥克風陣列，至少包含：

複數個等距置放在一圍繞一平面麥克風基座之該圓周的一圓上之麥克風；且其中

該麥克風基座是足夠低以定位該麥克風陣列，使得由該陣列從該麥克風基座座落之任何表面所拾取之聲音反射最小。

48. 如申請專利範圍第 47 項所述之麥克風陣列，其中由該等複數個麥克風捕捉之聲頻信號會輸入一電腦，該電腦使用該等聲頻信號用於聲源定位，該聲源定位決定產生

該等聲源信號之聲音的方向。

49. 如申請專利範圍第 47 項所述之麥克風陣列，其中由該等複數個麥克風捕捉之聲頻信號會輸入一電腦，該電腦使用該等聲頻信號用於聲束形成，該聲束形成藉由過濾來自所有其他方向之聲音以改進來自一方向之聲音品質。

50. 如申請專利範圍第 47 項所述之麥克風陣列，其中在該麥克風陣列中之該等麥克風是全向性。

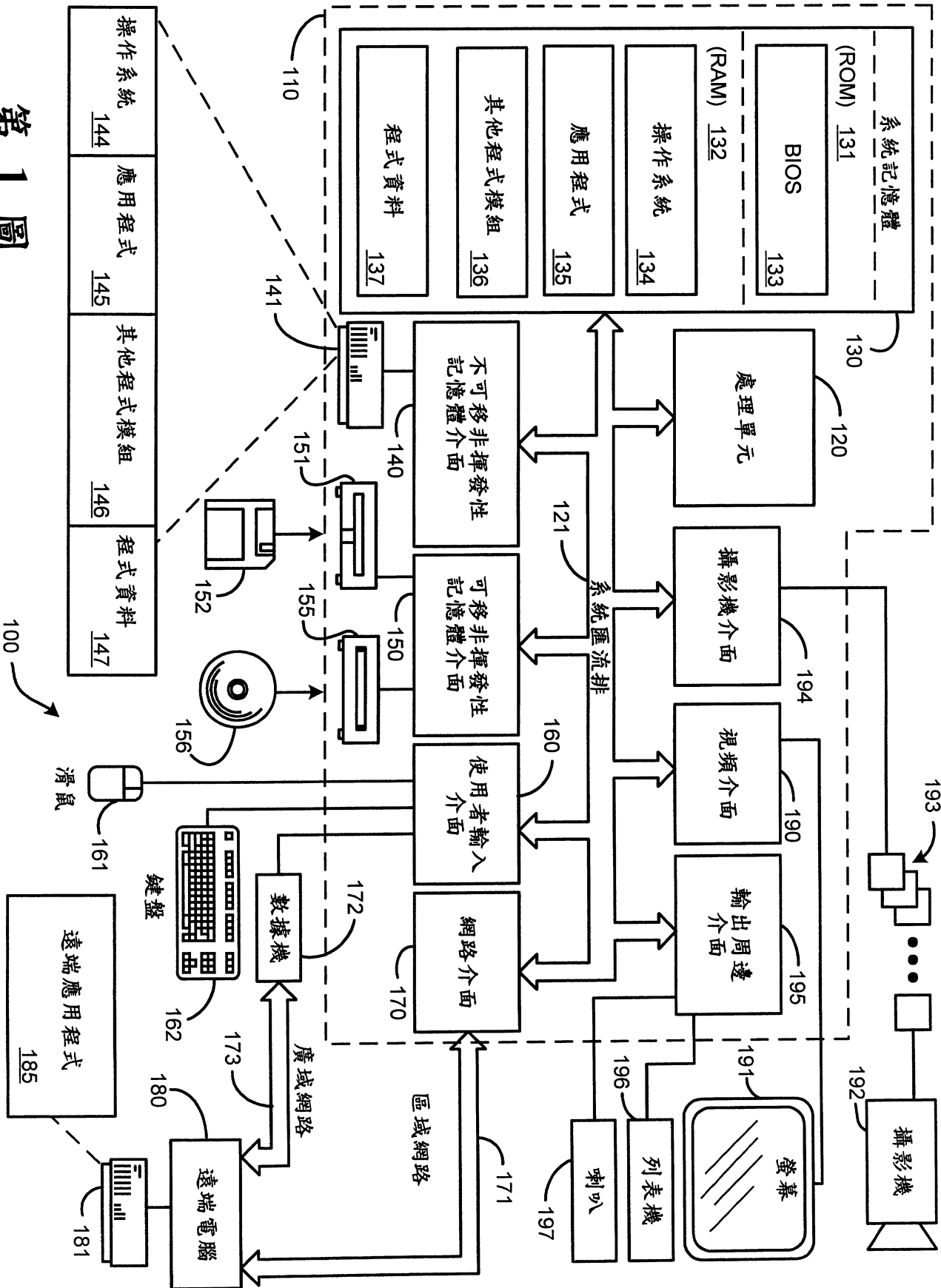
51. 如申請專利範圍第 47 項所述之麥克風陣列，其中在該麥克風陣列中之該等麥克風是單向性。

52. 如申請專利範圍第 47 項所述之麥克風陣列，其中該麥克風陣列至少包含八麥克風。

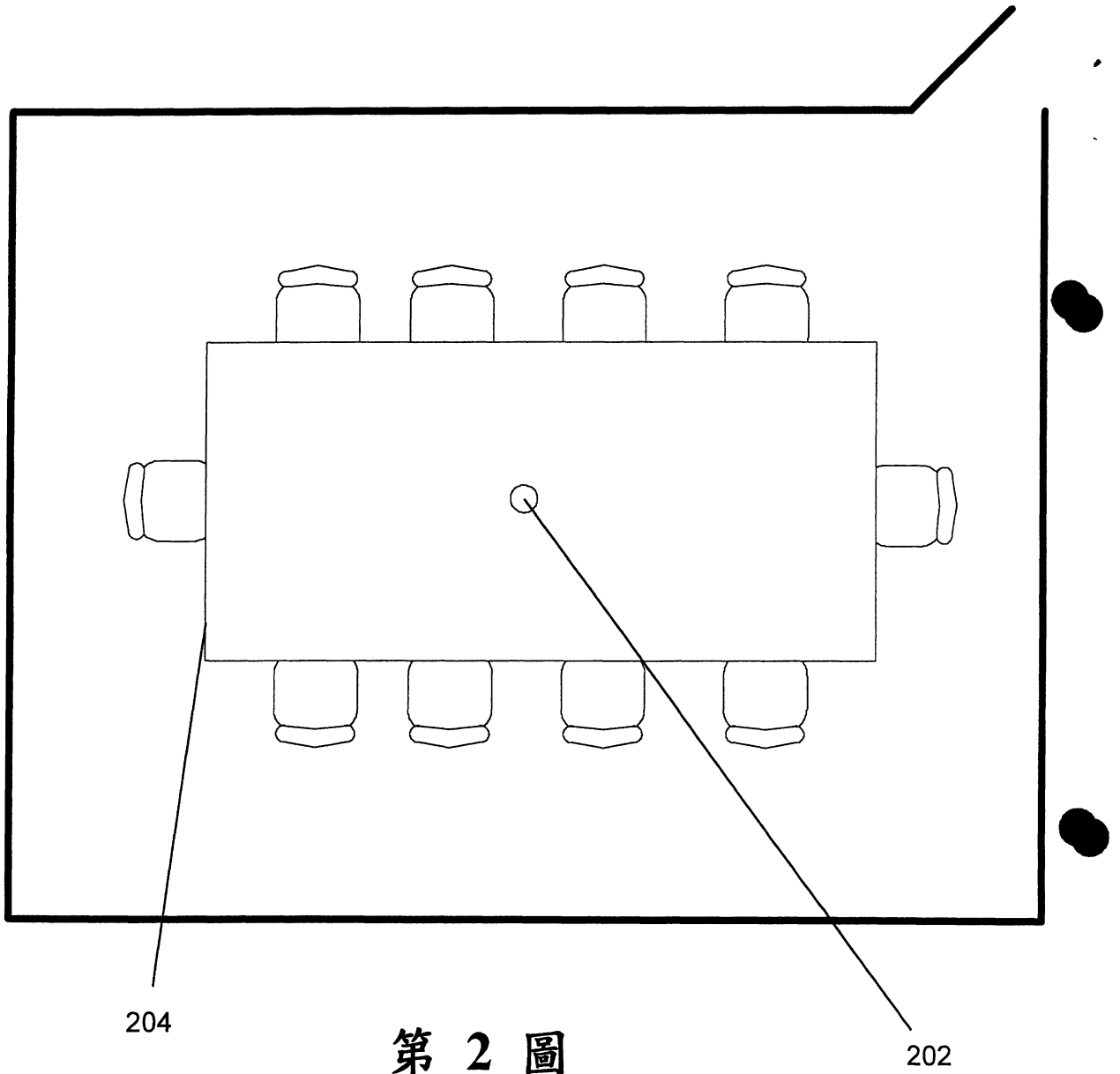
53. 如申請專利範圍第 47 項所述之麥克風陣列，其中至少一麥克風是裝設於一橡皮殼體中以保護其免於無關的聲音反射。

54. 如申請專利範圍第 47 項所述之麥克風陣列，其中隔音器是置於至少一麥克風下以保護其免於無關的聲音反

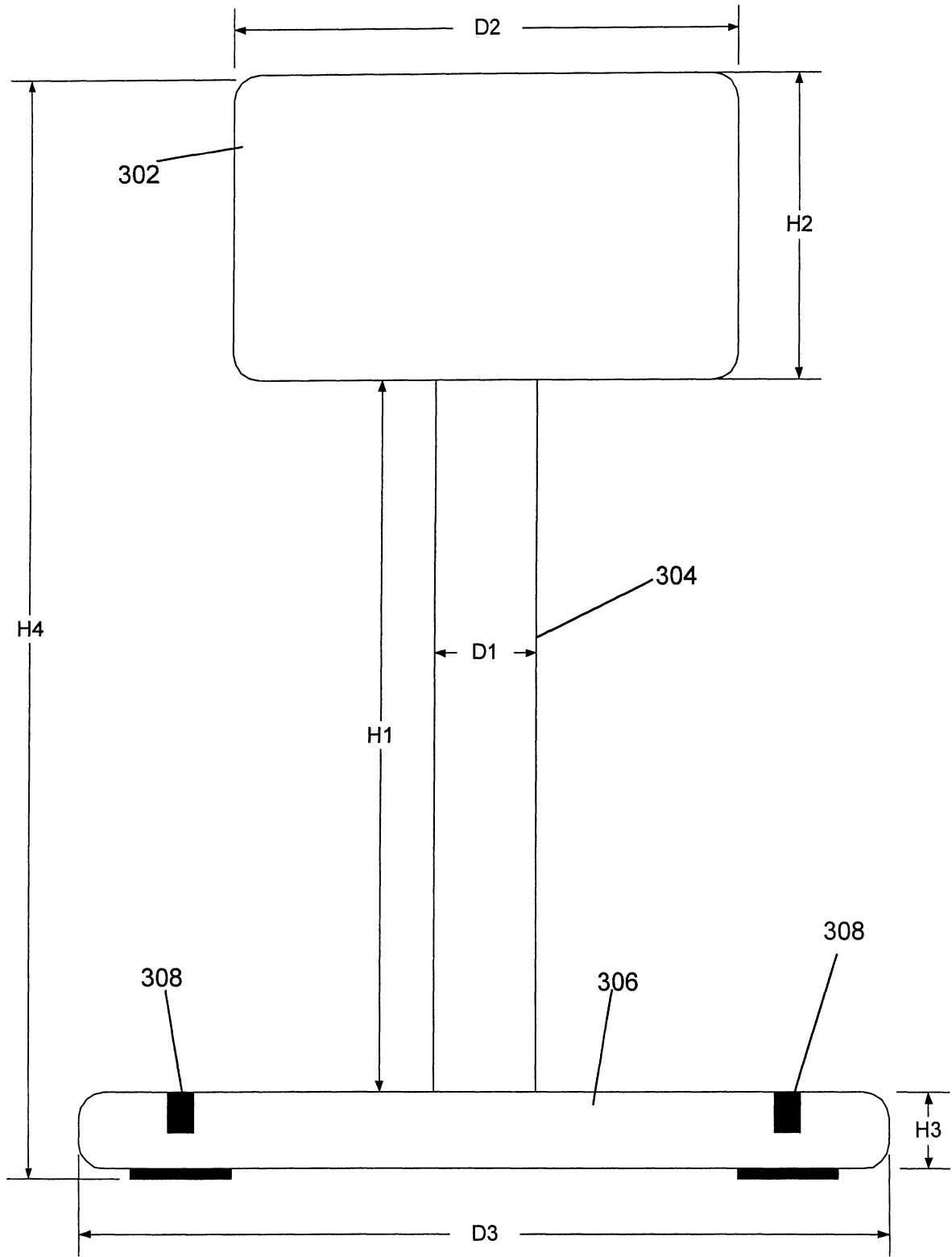
射。



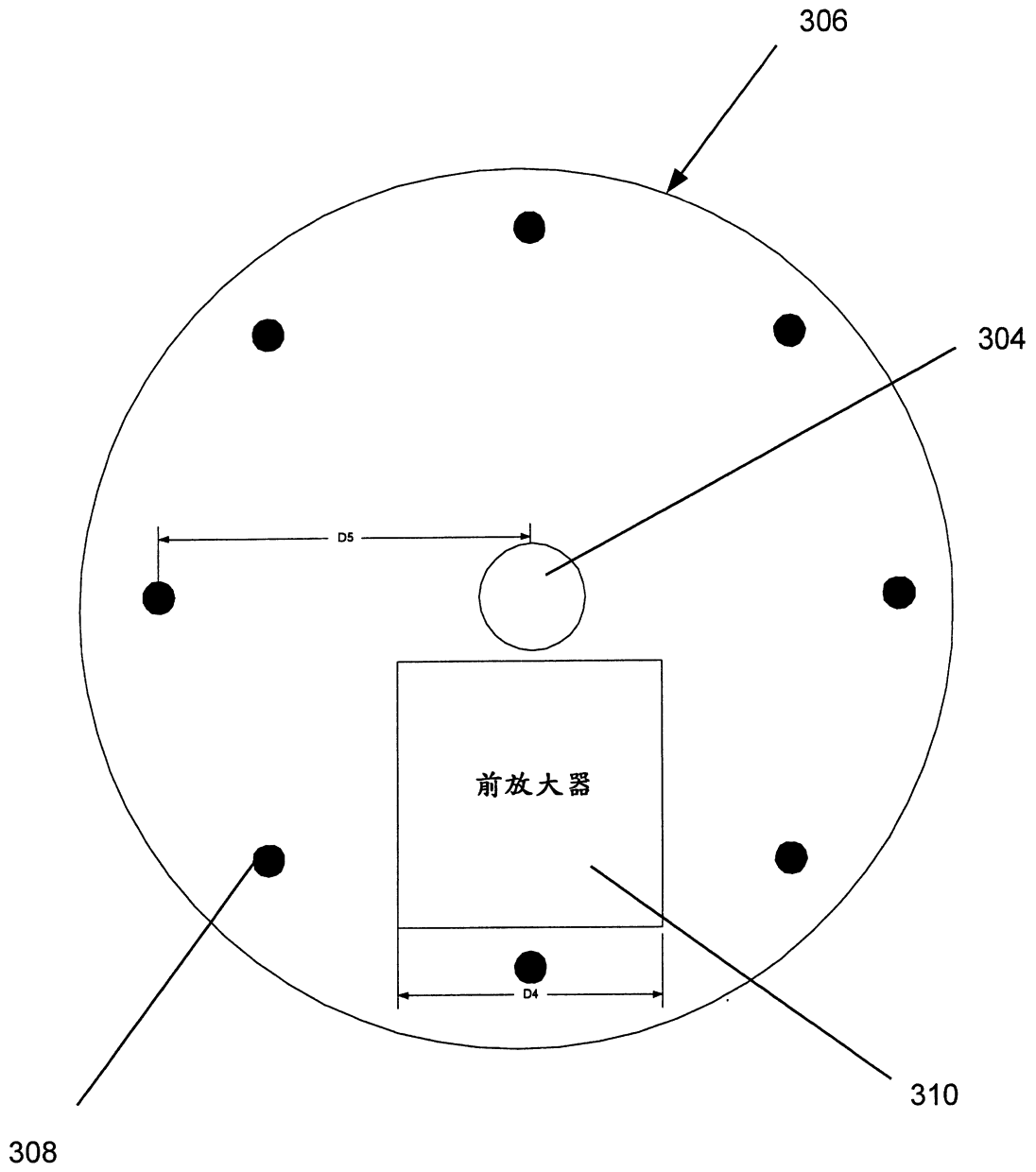
第 1 圖



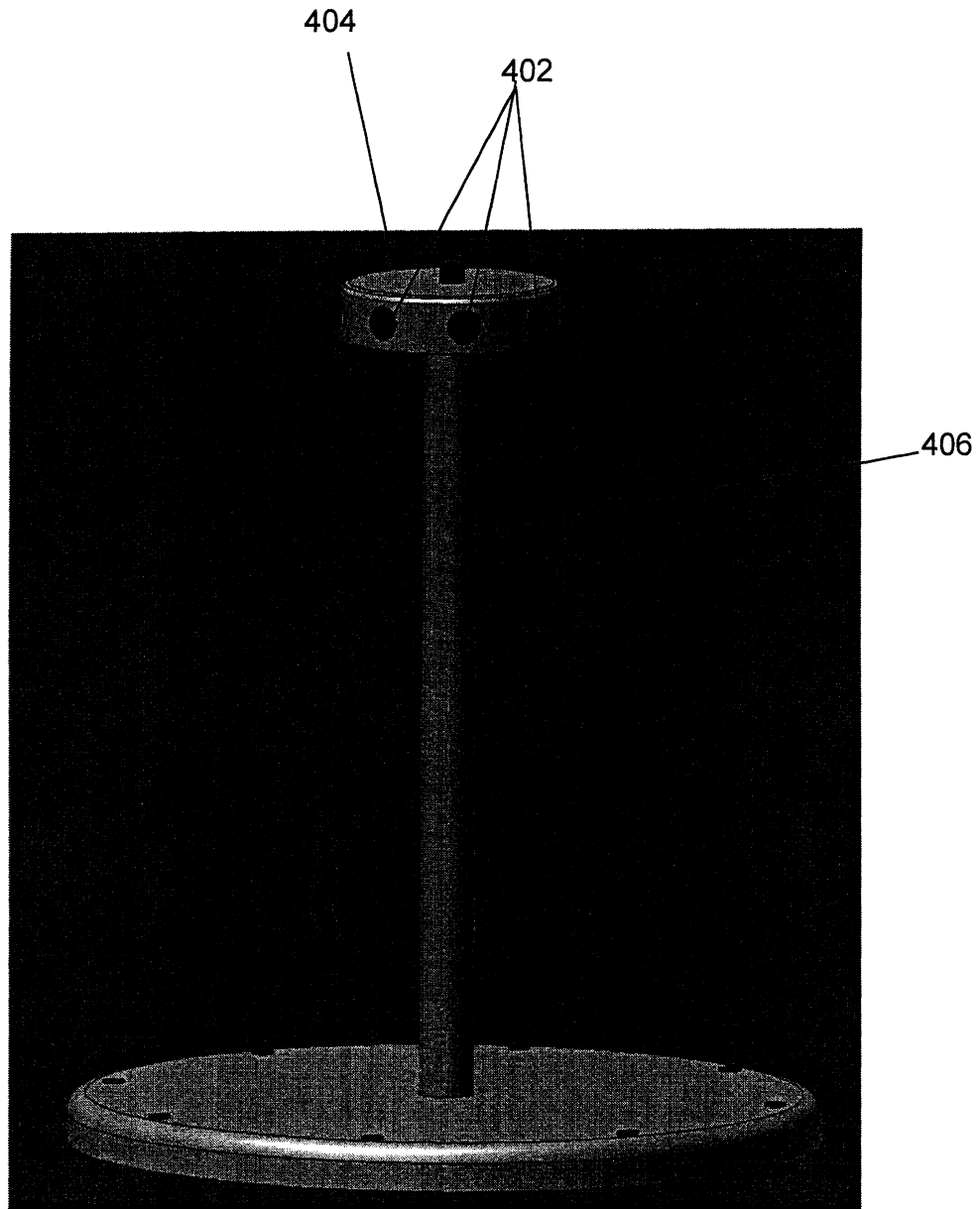
第 2 圖



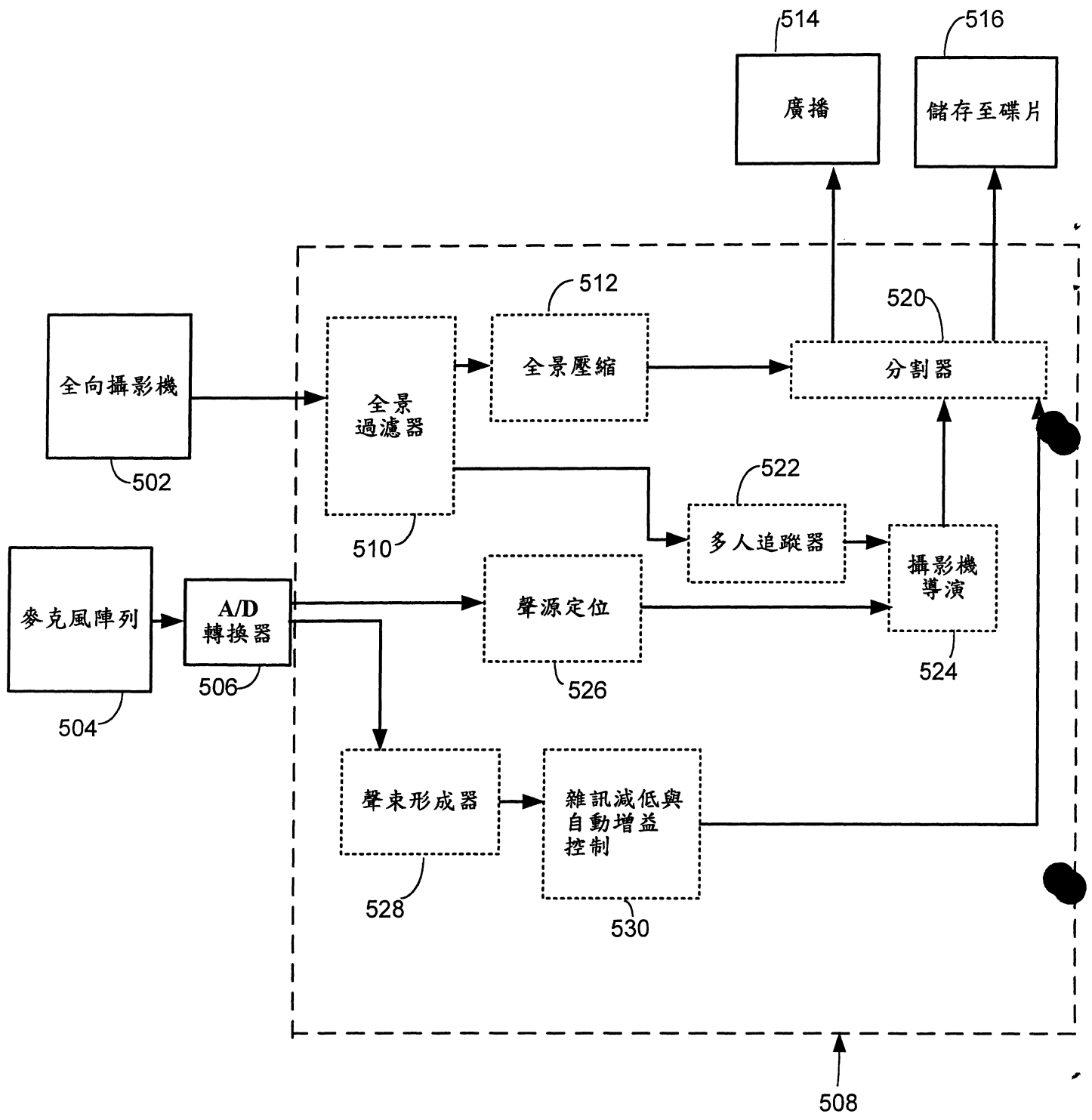
第 3A 圖



第 3B 圖



第 4 圖



第 5 圖

柒、指定代表圖：

(一)、本案指定代表圖為：第 3A 圖。

(二)、本代表圖之元件代表符號簡單說明：

302 全向攝影機

304 圓桿

306 麥克風基座

308 麥克風

捌、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無